

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ENGENHARIA GEOGRÁFICA, GEOFÍSICA E ENERGIA



## **Energy Performance Certification: National and Regional Characterization of Residential Building Stock**

André João da Silva Penas

**Mestrado Integrado em Engenharia da Energia e Ambiente**

Dissertação orientada por:

Professora Doutora Marta J. N. Oliveira Panão

2019

## Acknowledgements

*I want to express my deepest gratitude to Professor Marta Panão and to Doc. Student Raquel Figueiredo for supervising, who were always supportive and available to aid me solve the emerging problems, for their patience aswell. To ADENE for providing the data.*

*I'm thankful towards my family and Rita Raimundo for always holding my back.*

## Abstract

The global energy demand is partially due to the energy demand in the residential sector, which represents 30%. Part of this energy demand in residential buildings is related to heating and cooling systems, so the user can achieve conditions for thermal comfort. The quantification and characterization of the energy consumption in a building stock can be used as information to predict consumption behavior, to map energy poverty, usage scenarios and as impact study for energy policies.

The Residential Building Stock in Portugal has not been deeply characterised yet, so a computational model, which provides a rapid and easy study of the database available to the date could potentially be helpful to identify construction and energy related patterns. To address the need for characterization and analysis, a computational tool was built using MATLAB. This tool arrangement and data treatment requirement will be further explained so it can be modified if needed. The data is part of the Portuguese Energy Performance Certification, which first implementation dates from 2006, proceeded by an update, in 2013.

The questions one can analyze about the Residential Building Stock are extensive, therefore the most pertinent questions will be answered during this work, as to understand the evolution of the Residential Building Stock in time and space, using this analyzing tool. Every question studied had its national sample characterised to frame it. The answers to the questions showed that energy performance related parameters, Energy Certification Coefficient improved over the periods, in particular the periods when a regulation was introduced or updated. This proves that energy performance regulations had a positive effect on the Residential Building Stock towards the concept of nearly-zero energy buildings (NZEB) on private households in 2020.

**Keywords:** residential buildings, building stock characterization, energy performance certification, EPC database, Portuguese case study

## Resumo alargado

O consumo energético do setor residencial representa cerca de 30% do consumo global de energia. Parte deste consumo de energia é relativo a sistemas de aquecimento e arrefecimento, para proporcionar condições de conforto térmico ao utilizador, necessário ao seu bem-estar. A quantificação e caracterização do consumo energético em edifícios pode ser utilizada como material para estudos em cenários de utilização, previsão de comportamento energético, mapeamento de pobreza energética e estudo de impacto de políticas energéticas. Uma das soluções já testada para a caracterização do consumo energético foi a elaboração de um modelo estocástico, que gera dados sintéticos com base nas características de desempenho energético de uma amostra reduzida. Estes dados além de necessitarem de uma atualização, carecem de tratamento focado na vertente regional (NUTS III) e nacional, bem como no contexto e no ano de construção da habitação.

O objetivo primário relativo ao tema proposto passa pelo tratamento e análise dos dados em bruto provenientes da Certificação Energética de Edifícios feita pela ADENE - Agência para a Energia no âmbito do Sistema de Certificação Energética, cujo conteúdo são dados relativos à certificação energética de cada elemento da amostra de edifícios residenciais, como parâmetros físicos e indicadores de desempenho energético. Outro objetivo é a criação de um modelo computacional que permita caracterizar e analisar diversos parâmetros do parque residencial, possibilitando também uma análise combinada de vários desses parâmetros. O programa a desenvolver pretende atribuir um certo nível de rapidez na obtenção de resultados através da automatização dos processos de análise. Através desta caracterização será possível comparar as variáveis em função da região ou do ano de construção, assim como a avaliação do impacto de medidas de política energética, o progresso realizado na área de certificação energética ou a disparidade regional da certificação energética.

A metodologia implementada baseia-se essencialmente no modelo de análise e caracterização estatística. Este modelo está a ser desenvolvido no software MATLAB. A parte em foco nesta primeira abordagem ao trabalho foi o código desenvolvido, sendo que toda a análise será aprofundada numa segunda parte do trabalho.

Os dados relativos ao Sistema de Certificação Energética fornecidos pela ADENE – Agência para a Energia estão organizados em ficheiros de informação diferenciada, com a informação de cada entrada correspondente a um certificado energético, divididos em ‘Geral’ e ‘Envolvente’. Contêm uma série de informações sobre o certificado de cada uma das habitações, relacionados com a localização do mesmo, características estruturais, zona climática, necessidades de energia e índices de desempenho.

O tratamento destes dados pelo modelo computacional desenvolvido veio a mostrar ao longo da sua elaboração uma panóplia de resultados, que foram utilizados para responder a perguntas colocadas. O parâmetro sobre o qual recaía cada pergunta foi inicialmente caracterizado num âmbito nacional, de

modo a aferir sem distinção de contexto (novo, existente ou grande intervenção), ano de construção ou região NUTS III e que constituem os três tipos de filtros disponíveis. Após o enquadramento foi possível observar, através da análise por variável, a disparidade entre ambientes urbanos e rurais no qual é destacada a área útil do pavimento, sendo que esta é como seria de esperar maior em regiões tipicamente rurais. A surpresa está na área de pavimento útil que para o ‘Algarve’ e a ‘Grande Lisboa’ uma distribuição de probabilidade equiparável, apesar de uma muito maior densidade de alojamentos em ‘Grande Lisboa’. Foi possível observar através do ano de construção e contexto o impacto nos indicadores de desempenho energético que a alteração da Legislação foi tendo ao longo dos anos, como é o caso da transmitância térmica das janelas e do coeficiente global de desempenho energético, mostrando a forte contribuição para um melhor desempenho energético do edificado português.

**Palavras-chave:** edifícios habitacionais, caracterização do parque residencial, certificação de desempenho energético, base de SCE, caso de estudo Português

# Contents

|                                                                                               |      |
|-----------------------------------------------------------------------------------------------|------|
| <b>Abstract</b> .....                                                                         | iii  |
| <b>Resumo alargado</b> .....                                                                  | iv   |
| <b>List of Figures</b> .....                                                                  | viii |
| <b>List of Tables</b> .....                                                                   | xi   |
| <b>List of Nomenclature</b> .....                                                             | xiii |
| 1. Introduction.....                                                                          | 1    |
| 1.1 Framework .....                                                                           | 1    |
| 1.2 Objectives .....                                                                          | 2    |
| 1.3 Structure.....                                                                            | 3    |
| 2. State of the Art.....                                                                      | 4    |
| 3. Methods .....                                                                              | 6    |
| 3.1 <i>Input</i> .....                                                                        | 6    |
| 3.2 Processes .....                                                                           | 10   |
| 3.2.1 General.....                                                                            | 10   |
| 3.2.2 Envelope.....                                                                           | 15   |
| 3.2.3 Ventilation.....                                                                        | 19   |
| 3.2.4 Main Function .....                                                                     | 20   |
| 3.3 Output display .....                                                                      | 21   |
| 3.4 Challenges.....                                                                           | 25   |
| 3.5 Fast Guide .....                                                                          | 25   |
| 4 - Method Applicability .....                                                                | 27   |
| 4.1 How Air Changes Per Hour (RPH) differs in the Portuguese RBS? .....                       | 27   |
| 4.2 How does the building energy performance ( $RNt$ ) evolve in time?.....                   | 32   |
| 4.3 How does the building energy performance ( $RNt$ ) will evolve in the recent future?..... | 38   |
| 4.4 What is the building energy performance ( $RNt$ ) in different contexts? .....            | 40   |
| 4.5 How does Net Floor Area changes regionally, or in context?.....                           | 42   |
| 4.6 How the window-to-floor area ratio changes over the contexts? .....                       | 49   |
| 4.7 How does the window U-value change over the years and contexts? .....                     | 52   |
| 4.8 Walls U-value by Region and Climatic Zone .....                                           | 56   |
| 5 Conclusions.....                                                                            | 62   |
| References .....                                                                              | 65   |
| Appendix A – Rules and Recommendations .....                                                  | 66   |



## List of Figures

|                                                                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1 - General Analysis Flowchart. ....                                                                                                                                                                                  | 12 |
| Figure 2 - Energy Certification Coefficient ( $RNt$ ) Probability Distribution Function and Histogram for a National Sample. ....                                                                                            | 13 |
| Figure 3 - General Analysis with Variables Flowchart. ....                                                                                                                                                                   | 14 |
| Figure 4 - Envelopeanalysis.m Flowchart. ....                                                                                                                                                                                | 18 |
| Figure 5 - Envelope Analysis with Variables Flowchart. ....                                                                                                                                                                  | 19 |
| Figure 6 - Flowchart of File Relationship for Non-filtered Variable Results. ....                                                                                                                                            | 20 |
| Figure 7 - Flowchart of File Relationship for Filtered Variable Results. ....                                                                                                                                                | 21 |
| Figure 8 - Energy Certification Coefficient ( $RNt$ ) probability distribution for the National sample. ...                                                                                                                  | 22 |
| Figure 9 - Boxplot of Energy Certification Coefficient ( $RNt$ ) for National Sample with Quartile Values (Q1, Median and Q3) and. Non-outlier Values (L1 and L2) displayed. ....                                            | 23 |
| Figure 10 - Existing vs New Energy Certification Coefficient ( $RNt$ ) probability distribution. ....                                                                                                                        | 23 |
| Figure 11 - Existing vs New Boxplot comparison for Energy Certification Coefficient ( $RNt$ ). ....                                                                                                                          | 24 |
| Figure 12 - Energy Certification Coefficient ( $RNt$ ) for ‘Beira Interior Sul’ for the construction period ‘1946-1969’. ....                                                                                                | 24 |
| Figure 13 - MATLAB main.m Script First Input. ....                                                                                                                                                                           | 26 |
| Figure 14 - MATLAB main.m Script Second Input. ....                                                                                                                                                                          | 26 |
| Figure 15 – MATLAB Envelopereg.m Script Input. ....                                                                                                                                                                          | 26 |
| Figure 16 - Air Changes Per Hour (RPH) at National Level Probability Distribution Function and Histogram (Left). Air Changes Per Hour (RPH) at National Level Probability Distribution Function (Right). ....                | 27 |
| Figure 17 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Grande Porto’(Left). Air Changes Per Hour (RPH) Probability Distribution Function in 'Grande Porto' (Right). ....                 | 28 |
| Figure 18 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Grande Lisboa’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function in 'Grande Porto'(Right). ....                | 28 |
| Figure 19 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Beira Interior Norte’(Left). Air Changes Per Hour (RPH) Probability Distribution Function in 'Beira Interior Norte' (Right). .... | 29 |
| Figure 20 - Comparison of Parameter RPH in Regions ‘Grande Lisboa’ and ‘Grande Porto’ with ‘Beira Interior Norte’. ....                                                                                                      | 30 |
| Figure 21 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of ‘Existing’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'Existing’ (Right). ....                        | 30 |
| Figure 22 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of ‘New’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'New’(Right). ....                                   | 31 |
| Figure 23 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of ‘Major Renovation’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'New’(Right). ....                      | 31 |
| Figure 24 - Comparison of RPH Between All Contexts. ....                                                                                                                                                                     | 32 |
| Figure 25 - $RNt$ for Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram. ....                                                                                                      | 33 |
| Figure 26 - $RNt$ for Retrofitted Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram. ....                                                                                          | 34 |

|                                                                                                                                                                                                                                   |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 27 - RNt for 'Existing' Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram. ....                                                                                                  | 34 |
| Figure 28 - RNt for 'Existing' Buildings Constructed in '1961-1990' Probability Distribution Function and Histogram. ....                                                                                                         | 35 |
| Figure 29 - RNt for 'Existing' Buildings Constructed in '1991-2005' Probability Distribution Function and Histogram. ....                                                                                                         | 35 |
| Figure 30 - RNt for Buildings Constructed in '2006-2012' Probability Distribution Function and Histogram (Left). RNt for Buildings Constructed in '2013-2018' Probability Distribution Function and Histogram (Right). ....       | 36 |
| Figure 31 - RNt Comparison for Buildings Constructed in Periods '2006-12' and '2013-2018'.....                                                                                                                                    | 36 |
| Figure 32 - RNt 'Before 1919-1960', '1961-1990', '1991-2005', '2006-2012' and '2013-2018' Comparison of Probability Distribution Functions. ....                                                                                  | 37 |
| Figure 33 - RNt for Buildings with Pre-certificate Probability Distribution Function and Histogram (Left). RNt Pre-certificate Sample Probability Distribution Function (Right). ....                                             | 38 |
| Figure 34 - RNt Pre-certificate Sample and 'New from 2013-2018' Probability Distribution Functions Comparison. ....                                                                                                               | 39 |
| Figure 35 - RNt for Context 'Existing' Probability Distribution Function and Histogram (Left). RNt for Context 'Existing' Probability Distribution Function (Right) .....                                                         | 40 |
| Figure 36 - RNt for Context 'New' Probability Distribution Function and Histogram (Left). RNt for Context 'New' Probability Distribution Function (Right).....                                                                    | 40 |
| Figure 37 - RNt for Context 'Major Renovation' Probability Distribution Function and Histogram (Left). RNt for Context 'Major Renovation' Probability Distribution Function (Right).....                                          | 41 |
| Figure 38 - RNt 'Existing', 'New' and 'Major Renovation' Contexts Comparison of Probability Distribution Functions.....                                                                                                           | 41 |
| Figure 39 - Net Floor Area ( $m^2$ ) in Portugal Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Portugal Probability Distribution Function (Right). ....                                     | 42 |
| Figure 40 - Net Floor Area ( $m^2$ ) in 'Grande Lisboa' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Grande Lisboa' Probability Distribution Function (Right). ....                       | 43 |
| Figure 41 - Net Floor Area ( $m^2$ ) in 'Grande Porto' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Grande Porto' Probability Distribution Function (Right). ....                         | 43 |
| Figure 42 - Net Floor Area ( $m^2$ ) in 'Algarve' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Algarve' Probability Distribution Function (Right). ....                                   | 44 |
| Figure 43 - Net Floor Area ( $m^2$ ) Comparison Between Regions Aforementioned. ....                                                                                                                                              | 45 |
| Figure 44 - Net Floor Area ( $m^2$ ) in Context 'Existing' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Context 'Existing' Probability Distribution Function (Right). ....                 | 46 |
| Figure 45 - Net Floor Area ( $m^2$ ) in Context 'New' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Context 'New' Probability Distribution Function (Right). ....                           | 46 |
| Figure 46 - Net Floor Area ( $m^2$ ) in Context 'Major Renovation' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Context 'Major Renovation' Probability Distribution Function (Right). .... | 47 |
| Figure 47 - Net Floor Area ( $m^2$ ) of Contexts 'Major Renovation', 'New' and 'Existing' Probability Distribution Functions.....                                                                                                 | 47 |
| Figure 48 - Window-to-Floor Area Ratio in Portugal Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Portugal Probability Distribution Function (Right). ....                                          | 49 |
| Figure 49 - Window-to-Floor Area Ratio in Context 'Existing' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'Existing' Probability Distribution (Right). ....                               | 50 |

|                                                                                                                                                                                                                                       |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 50 - Window-to-Floor Area Ratio in Context 'New' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'New' Probability Distribution (Right) .....                                             | 50 |
| Figure 51 – Window-to-Floor Area Ratio in Context 'Major Renovation' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'Major Renovation' Probability Distribution (Right) .....                   | 50 |
| Figure 52 - Window to Floor Area Ratio for Contexts 'Existing', 'New' and 'Major Renovation' Probability Distribution Functions Comparison. ....                                                                                      | 51 |
| Figure 53 - Window U-value ( $WK.m2$ ) in Portugal Probability Distribution and Histogram (Left). Window U-value ( $WK.m2$ ) in Portugal Probability Distribution Function (Right).....                                               | 52 |
| Figure 54 - Window U-value ( $WK.m2$ ) in Context 'Existing' Probability Distribution Function and Histogram (Left). Window U-value ( $WK.m2$ ) in Context 'Existing' Probability Distribution Function (Right). ....                 | 53 |
| Figure 55 - Window U-value ( $WK.m2$ ) in Context 'New' Probability Distribution Function and Histogram (Left). Window U-value ( $WK.m2$ ) in Context 'New' Probability Distribution Function (Right). ....                           | 53 |
| Figure 56 - Window U-value ( $WK.m2$ ) in Context 'Major Renovation' Probability Distribution Function and Histogram (Left). Window U-value ( $WK.m2$ ) in Context 'Major Renovation' Probability Distribution Function (Right). .... | 54 |
| Figure 57 - Window U-value ( $WK.m2$ ) for Contexts 'Existing', 'New' and 'Major Renovation' Probability Distribution Functions.....                                                                                                  | 54 |
| Figure 58 - Window U-value ( $WK.m2$ ) Probability Distribution Functions of Periods in Legend. ...                                                                                                                                   | 55 |
| Figure 59 - Walls U-value ( $WK.m2$ ) in Portugal Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in Portugal Probability Distribution Function (Right). ....                                       | 56 |
| Figure 60 - Walls U-value ( $WK.m2$ ) in 'Alentejo Litoral' Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in 'Alentejo Litoral' Probability Distribution Function (Right). ....                   | 57 |
| Figure 61 - Walls U-value ( $WK.m2$ ) in 'Alto Trás-os-Montes' Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in 'Alto Trás-os-Montes' Probability Distribution Function (Right). ....             | 57 |
| Figure 62 – Comparison Between Probability Distribution Function of Walls U-value ( $WK.m2$ ) 'Alto Trás-os-Montes' and 'Alentejo Litoral' (Left). Boxplot Diagram for Walls U-value ( $WK.m2$ ) for the same Regions (Right).....    | 58 |
| Figure 63 - Walls U-value ( $WK.m2$ ) in 'Grande Lisboa' Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in 'Grande Lisboa' Probability Distribution Function (Right). ....                         | 58 |
| Figure 64 - Walls U-value ( $WK.m2$ ) in 'Algarve' Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in 'Algarve' Probability Distribution Function (Right). ....                                     | 59 |
| Figure 65 - Walls U-value ( $WK.m2$ ) in 'Grande Porto' Probability Distribution Function and Histogram (Left). Walls U-value ( $WK.m2$ ) in 'Grande Porto' Probability Distribution Function (Right). ....                           | 59 |
| Figure 66 - Walls U-value ( $WK.m2$ ) in 'Grande Porto' Probability Distribution Function. ....                                                                                                                                       | 60 |
| Figure 67 - Walls U-value ( $WK.m2$ ) in 'Grande Lisboa', 'Grande Porto' and 'Algarve' Probability Distribution Function Comparison. ....                                                                                             | 60 |
| Figure 68 - Walls U-value ( $WK.m2$ ) in 'Grande Lisboa' and 'Algarve' Probability Boxplot Comparison. ....                                                                                                                           | 61 |

## List of Tables

|                                                                                                                                  |    |
|----------------------------------------------------------------------------------------------------------------------------------|----|
| Table 1 - Example of a .csv file (generalcalculatedvalues.csv) with its parameter names and limits. ..                           | 10 |
| Table 2 - General EPC data.....                                                                                                  | 11 |
| Table 3 - General .csv Files and MATLAB Script for Domains. ....                                                                 | 11 |
| Table 4 - Envelope EPC data.....                                                                                                 | 15 |
| Table 5 - Envelope .csv Files and MATLAB Script for Domains.....                                                                 | 16 |
| Table 6 - Energy Classification According to <i>RNt</i> Parameter, adapted from [13].....                                        | 33 |
| Table 7 - Distribution Function and Distribution Parameters and Medians for Each Period in Figure 32.<br>.....                   | 37 |
| Table 8 - Distribution Function and Distribution Parameters and Medians for Each Context in Figure<br>38. ....                   | 41 |
| Table 9 - Distribution Function and Distribution Parameters for Each Region Net Floor Area in Figure<br>43. ....                 | 45 |
| Table 10 - Median and Quartiles of Regions in Figure 43 by Net Floor Area in <i>m</i> <sup>2</sup> . ....                        | 45 |
| Table 11 - Distribution Function and Distribution Parameters for Each Context Net Floor Area in Figure<br>47. ....               | 47 |
| Table 12 - Median and Quartiles for all Contexts by Net Floor Area.....                                                          | 48 |
| Table 13 - Distribution Function, Distribution Parameters and Median for Each Context Window-to-<br>Floor Area in Figure 52..... | 51 |
| Table 14 - Distribution Function, Distribution Parameters and Median for Each Context Window U-<br>value in Figure 57. ....      | 54 |
| Table 15 - Distribution Function and Distribution Parameters for Each Period in Figure 58. ....                                  | 55 |
| Table 16 - Distribution Function and Distribution Parameters for Each Region Studied in Sub-Chapter<br>4.8. ....                 | 61 |
| Table 17 - Reference Values for Heating Season (Part I), adapted from [14]. ....                                                 | 66 |
| Table 18 - Reference Values for Heating Season (Part II), adapted from [14]. ....                                                | 66 |
| Table 19 - Criteria for Climatic Zones in Heating Season, adapted from [14]. ....                                                | 66 |



## List of Acronyms

EPBD - Energy Performance Building Directive

EPC - Energy Performance Certification

GD – Graus-dia (Degrees-day)

NUTS III – Nomenclatura das Unidades Territoriais para Fins Estatísticos (Nomenclature of Territorial Units for Statistical Purposes)

NZEB – Nearly-zero Energy Buildings

RBS - Residential Building Stock

RCCTE – Regulamento das Características de Comportamento Térmico dos Edifícios (Regulation of Thermal Behaviour Characteristics of Buildings)

RECS - Regulamento de Desempenho Energético dos Edifícios de Comércio e Serviços (Services and Commerce Building Energy Performance Regulation)

REH - Regulamento de Desempenho Energético dos Edifícios de Habitação (Habitational Building Energy Performance Regulation)

RPH - Renovations per Hour

SCE - Sistema de Certificação Energética (Energy Certification System)

## List of Nomenclature

*RNt* - Energy Certification Coefficient

*U<sub>wall</sub>* - Thermal transmission of the walls

*U<sub>wdn</sub>* - Thermal transmission of the windows

*U<sub>roof</sub>* – Thermal transmission of the roof

*g<sub>t</sub>* - Solar radiation transmittance of the shading and glass systems

*g<sub>t vi</sub>* - Solar radiation transmittance of the glass system



# 1. Introduction

## 1.1 Framework

The Portuguese residential building stock (RBS) represents a complex sample given its heterogeneity caused by regional variations of the housing materials, technics used and different energy needs. Such building stock requires to be characterised regionally and chronologically, to get a most accurate snapshot of how the RBS evolved in time and the result of policies taken to regulate its energy performance. Regionally, it is important to characterise how energy performance and insulation properties of dwellings differ in space, like urban/rural contrasts and patterns for different climate regions.

The Characteristics Regulation of Building Thermal Behaviour (RCCTE, acronym in Portuguese) , approved by Decree-Law n°40/90 from 1990, was the first legal document to regulate project requirements for new buildings and major renovations in Portugal, to guarantee thermal comfort in buildings and air quality. This first RCCTE version aimed to limit the potential consumption, however, at that time using a climatization equipment was uncommon and its introduction resulted in a growth in energy consumption. A RCCTE updated version had to be created to consider a more proper usage of climatization equipment and thus more strict energy requirements. This update is Decree-Law n°80/2006, from 2006, which also introduces the necessity of air renovation in buildings [2].

The most recent legislation is Decree-Law n°118/2013 of 20<sup>th</sup> August, approved an update and uniformization to EPC, for residential buildings the Habitational Building Energy Performance Regulation (REH) and for service and commercial buildings the Services and Commerce Building Energy Performance Regulation (RECS) which are a complement to the same regulation. This actualization brought new requirements of thermal quality and separates REH and RECS, this gives a specialized technical approach to both REH and RECS. Besides the separation by type of building, in both REH and RECS there are specific requirements according to the context of the building, divided into “New”, “Existing” and “Major Renovation”.

The REH defines the standards for construction and maintenance of residential buildings that consider an optimized management of the consumption in order to guarantee the thermal comfort to the user, also it includes the calculation norms for the energy needs of cooling and heating, primary energy consumption and respective energy classification [3].

The building energy performance evaluation lays, for residential buildings, in its thermal behavior and in its systems efficiency. For service and commerce buildings it is added to the mentioned above the conduction and maintenance of the technical systems. In addition to the energy efficiency, it considers the quantity of the renewable energy contribution, the using of passive systems and performance optimization in buildings. Further the concept of a nearly zero energy building (NZEB), which will be the standard by the year 2020 for private owned buildings, or 2018 for new buildings of public entities and for large interventions in existing buildings [3].

## 1.2 Objectives

The goal of this work is the treatment and analysis of the raw data coming from the Energy Performance Certification (EPC) managed by ADENE, which content are data related to the EPC of each element of the building stock, such as physical properties and energy performance indicators. Most importantly, the creation and development of a computational model with MATLAB tool allow users simple access to the RBS characteristics, whose purpose is to characterise and analyze various parameters of the RBS, allowing also a joint analysis in both time and region. The developed model is pretended to give quick results by automatization in the analysis and characterization processes. Through this characterization it will be possible to compare the variables according to the region, year of construction or context as well as to measure the impact of energy policies, energy certification performance progress over the years or its disparity between regions.

It is not intended to display all possible results during this work, but rather show possible answers to questions proposed. As the range of questions is unlimited, the objective is to set a defined range of questions to exemplify the purpose and usefulness of the computational model developed.

## 1.3 Structure

The dissertation is divided into six chapters, including Chapter 1 - Introduction, which gives a context for the proposed work. Chapter 2 - State of the Art refers to the development of similar studies, that provide the most recent updates to the technology concerning EPC and RBS characterization, analysis and impacts. In Chapter 3 - Methods, the data input is explained qualitatively and quantitatively; then, the building of the computational model, its processes and its capabilities are also described with examples in the output subsection. A Fast Guide gives a simple tutorial on how to run the computational model and which inputs are needed for each output. In Chapter 4 - Method applicability, a few questions are to be answered, due to the broadness of possible outcomes. Finally, in Chapter 5 - Conclusions, where conclusions are taken concerning the EPC database characteristics and RBS analysis, answers from the previous chapter are summarized, and the computational model capabilities to analyze the RBS are described.

## 2. State of the Art

The EPC is part of a strategy imposed by the Energy Performance of Buildings Directive (European Union's legislative instrument) and is an important tool to improve the energy efficiency in buildings. The main goals of the EPC are to give information about the dwelling efficiency, consumption and energy costs to the owners, to target energy efficiency upgrades so to reduce energy and costs. Finally, the EPC is an instrument to be used in real estate transactions, to let costumers aware of the dwelling energy performance [4]. European Union member states have their own EPC data collection and EPC calculation approaches, making the comparison of two different EPC systems difficult, still it is important to learn from other EPCs to understand about the EPC one is currently working on [5].

The impact of the energy EPC policy was studied, after some years of implementation, in some countries. Murphy [6] studied the impact of EPC of two distinct populations focused on the Dutch RBS, the first sample received the EPC certificate during property transaction and the second population was not in that situation. The main goal was to evaluate the influence of the EPC in these two previously described situations. It was found that the influence of the EPC at the property acquisition mattered for 10% of the buyers surveyed, a weak influence. It is also stated that after the property acquisition 22% were influenced by the EPC to apply energy efficiency measures. One result suggests that the most cited reason by house owners for not requesting an EPC was apathy, with 51% of owners not considering an EPC necessary, also financial barriers might have an impact. Finally, 27% of EPC house owners and 50% of house owners without EPC did not take any energy efficiency measure due to consider their habitations to be already efficient.

In Hardy and Glew [7], it is aimed to quantify within the UK RBS (with over 15 million dwellings covered) errors in EPCs and it was estimated that between 36% and 62% of the EPCs contain errors, caused mostly by EPC accessors disagreeing in some parameters as in terms of wall and floor types. This problem can be improved using machine learning in the future. We can expect with this information that errors in the EPC could be common, and so, we can associate some mistakes in the Portuguese EPC that we might face to the same cause.

Usually there is a gap between theoretical energy performance and actual measured energy use in buildings. Wilde [8] investigates the dimension of this gap and how to possibly shorten it. Its causes differ from dwelling to dwelling or a combination of different causes, Wilde's literature review suggests that the main cause for the existence of this gap is in the design itself, due to miscommunication between client and design team, though the design itself sometime can be energy inefficient [8].

To overcome some of those problems Fonseca and Panão [9] tested a physics-based stochastic model using Monte Carlo for residential energy building consumption characterization in Portugal, which generates synthetic data based on energy performance features of a short sample. The EPC classes distribution generated by the model was satisfactory with differences among classes below 3.1%. Relevant simulations were performed, for example one of the scenarios simulated improvements in 30% of windows, decreasing windows U-value could improve the classification of at least one level in 15.5% of the dwellings in the RBS [9]. Also, in the same work, the RBS in Portugal is roughly characterised by histograms in terms of dwellings per area in region, by year of construction, net floor area, conservation, consumption and other structural characteristics. Besides the need to update these data, a data treatment focused in the regional (NUTS III), as well as national, year of construction and context of the dwelling is lacking.

In a study of characterization of thermal performance and nominal heating gap of the RBS the Energy Performance of Buildings Directive (EPBD) in Portugal mainland, the same area studied with this work, but with data from 2012, Magalhães and Leal [10] explore the EPC database to characterise the gap between the nominal energy needs and the actual energy use. This study proved the effectiveness of the EPBD with improvements in energy performance after the legislation of 2006. The database used is the same used in this work, the SCE of ADENE from the year 2012, just before the introduction of REH, where most dwellings was built in the periods of the 1980 and 1990 decades. Meanwhile, the parameters  $N_{ic}$ ,  $N_{tc}$  and  $N_{vc}$ , which represent useful energy needs for heating, annual total primary energy needs and useful energy needs for cooling, respectively, were studied. The study showed a steady decrease in the first two parameters since '1951-1960', the last parameter had a decline since that period, but recently at the beginning of the 21<sup>st</sup> century increased, possibly due to an increase of window area, less use of shading and architectural changes.  $N_{tc}$  decreased with improvements in building construction materials and techniques after 1950, significant improvements were also seen in 1980 and 1990, also after in 2006. This paper shows that EPBD databases can be helpful to get insights on energy performance in the RBS, to identify the most critical types of dwellings and to track the evolution of the building stock or the impact of regulations.

## 3. Methods

The model proposed to analyse the RBS in Portugal has multiple MATLAB files, *.csv* files where the most useful data is stored and other *.csv* as auxiliary files. Most of them will have the purpose and construction explained below, as the rest are similar.

The model's final files are composed of sixteen MATLAB scripts including one main file, two data preparation files and one auxiliary file, additionally, seventeen *.csv* files with all data that have been used. File and scripts for every domain (National, Region, Construction Year, Context or Mixed) for general certificate information (*see Table 3*) (*General*) are differentiated from those that focus on dwelling envelope characteristics (*see Table 5*) (*Envelope*) and how these files and scripts are related with each others is presented further in Figures 6 and 7, for the first case and for the second, respectively.

The present chapter is separated by *Input*, *Processes* and *Output Display*. *Input* is where data transformation process is explained and which data has been so far analysed, present in the files. The next, *Processes* is where MATLAB processes are explained, how the model treats the data and how it organizes it so its results can be displayed. Finally, *Output* consists on displaying results as pretended.

### 3.1 *Input*

The data more suitable to characterise the residential building stock is the database of Energy Performance Certification (EPC) collected by ADENE – *Agência para a Energia* - from 2008 to 2018. The files in which the information was stored were divided into three origin groups with different information. The EPC database contains several differentiated files. These files were originally in *.xlsx* format and can be separated into three groups. The first group consists in one file only with general information of every dwelling, called *General*. Differently from *General*, where one entry corresponds to just only one EPC number, there are eight files with information about Envelope characteristics, in which there are several entries for the same EPC number. Each of these files is named from *Envelope1* to *Envelope8*, the multiple entries for one EPC number contain information of one feature as walls, windows or roofs. Finally, the last file *Ventilation* has data regarding dwelling ventilation, here like in *General*, each entry is for one EPC number. The sample has different sizes among these files, and every group was considered to constitute a representative sample of the Portuguese RBS.

None of the files were ready to be analysed and characterised in their untreated state, as there were headers, some document's description and invalid filling data had to be removed or treated. Also the files memory was enormous, therefore a data treatment was necessary to proceed. A more simple file was created, *data* file, to store the most important and necessary data of *General*. In the original format the actual data did not start in the file's first line, to be treated in MATLAB the data should be starting in the first line of the file. This was done by transferring the EPC number in the original *General* to *data*, starting in its first line. Fulcral data was transferred to *data*, as Region, Construction Year, and Context information. These data have a major role in the analysis and characterization, thus, they are called **variables**; it is in function of these variables that the physical and energy properties will be analysed, characterised and compared. Those physical and energy properties are named **parameters**. Invalid, non-numerical or empty values were filled with zero values, otherwise MATLAB file importation could not be executed, popping up an error message.

The variables are:

- **Region:** Regions are important to show regional RBS differences: different regions have distinct climates, physical properties and building materials which will have an influence in dwelling's energy efficiency. The data available shows information from 28 Regions (NUTS III) from Portugal mainland.
- **Context:** The RBS can be divided into three types of dwelling contexts. *Existing* refers to dwellings that already existed when they had the EPC done. *New* is for dwellings built from scratch since 2006, year when EPC was mandatory for new residential buildings. At last, *Major Renovation* concerns dwellings that existed before and were renovated.
- **Construction Year:** Refers to the year the dwelling was built. The construction year was divided into nine categories before EPC was implemented from '*Before 1919*' to '*2001-2005*', in between these periods are '*1919-1945*', '*1946-1960*', '*1961-1970*', '*1971-1980*', '*1981-1990*', '*1991-1995*', '*1996-2000*', '*2001-2005*' and '*after 2006*' each construction year was considered individually. Multiple periods can be studied simultaneously if continuous in time. When a building is not yet built, a classification of '*Pre-certificate*' is given to it. A pre-certificate is a document with the same characteristics as the EPC, where it is certified that the project fullfills the regulations, for dwellings that have not been yet built or retrofitted. This Pre-certificate gives a picture of how will be the RBS in the near future.

The three variables were not just copied and pasted to *data*. All non-numerical values were converted into numerical values. After creating *data* file it was necessary to make the files lighter and simpler as possible, to achieve this for every variable was developed a code of numbers. All non-numerical values were converted into numerical values to easen the reading process of the variables, e.g. an association number was given to the construction periods ‘Before 1919’ to ‘2001-2005’ corresponding to its index number. The auxiliar files then contain the code numbers for each variable are *contextcode.csv* for Context, *periodcode.csv* for Construction Year and *regcode.csv* for Regions.

To import the numerical value file *data* into MATLAB by the load function it was converted to .csv. Similarly the auxiliar files, with numerical and character values were converted to .csv to be imported by the *importdata* function.

The parameters are the object of the study since they describe certain characteristics of a dwelling. Some relevant parameters presented in *General* were included in *data*. These parameters are *Nic Value*, *Nvc value*, *Ntc value*, *Ntc limit*, *Energy Certification Coefficient (Rnt)* and *Net Floor Area*. Totalizing in the end ten distinct columns [11].

- ***Nic value*** : Useful energy needs for heating  $\left[ \frac{kWh}{\frac{m^2}{year}} \right]$ .
- ***Nvc value*** : Useful energy needs for cooling  $\left[ \frac{kWh}{\frac{m^2}{year}} \right]$ .
- ***Ntc value*** : Annual total primary energy needs  $\left[ \frac{kWh}{\frac{m^2}{year}} \right]$ .
- ***Ntc limit*** : Is the reference value for the annual limit of primary energy needs for a certain dwelling  $\left[ \frac{kWh}{\frac{m^2}{year}} \right]$ .
- ***Energy certification coefficient (Rnt)*** : This parameter results from the ratio between *Ntc value* and *Ntc limit*, allowing the measurement of the actual energy consumption of the dwelling to its reference energy consumption (dimentionless). In the REH it is required that buildings of context ‘New’ must be lower or equal to 1.00 of *Energy Certification Coefficient (Rnt)* and dwellings facing a *Major Renovation* must be lower or equal to *Rnt* of 1.50.
- ***Net floor area*** : Refers to the useful floor area of a dwelling [ $m^2$ ].

The same problem was faced when extracting the values from each of the eight *Envelope* files. For each *Envelope* original file another one was created, named from *Envelopetreated1* to *Envelopetreated8*, the files are heavier due to cointaing more than one piece of information for every EPC number. The *Envelopetreated* files created have eight columns, from the left to the right they are: *EPC number*, *envelope area*, *glazing g-value (gt vi)*, *shading g-value (gT)*, *window area (Awdn)*, *window thermal*

*transmittance (Uwdn), walls and roofs thermal transmittance (Uwalls and Uroofs, respectively)*. These files of *Envelope* are first treated with *testvalue.m* to create *calculatedvalues.csv*, a file with the same, but treated, parameters as all the other *Envelope* files plus *Window to Floor Area Ratio* parameter.

- **EPC number** : Certificate number to allocate the following values to a dwelling.
- **Envelope area** : The area of the surrounding envelope [ $m^2$ ].
- **Glazing g-value (gt vi)** : Solar radiation transmittance of the glass, an important parameter to calculate solar gains.
- **Shading g-value (gT)** : Solar radiation transmittance of the shading and glass systems.
- **Window area** : The area of the window [ $m^2$ ].
- **Window thermal transmittance (Uwdn)** : Thermal transmission of the windows, the lower this value is more heat resistant they are [ $\frac{W}{K}$ ].
- **Walls thermal transmittance (Uwalls)** : Thermal transmission of the walls [ $\frac{W}{K}$ ].
- **Roofs thermal transmittance (Uroofs)** : Thermal transmission of the roofs [ $\frac{W}{K}$ ].
- **Window to floor area ratio** : Given by the ratio between *Window area* and *Net floor area* of the same dwelling.

Finally, regarding the data coming from ventilation data file *Ventilation.csv*, similarly to *data* includes four variables: Region, Context, Construction Year and the *Air Changes Per Hour* (RPH) – the latter is the only parameter that exclusively belongs to *Ventilation.csv*.

- **Air changes per hour (RPH)** : This parameter refers to the number of air changes in a certain room or house, its value is correspondent to the volume of air renovated compared to the volume of the same space [RPH].

To keep the record of every parameter, numerical associations were also made and condensed into three other auxiliar files. They are also used to label further in the computational model and they contain the higher and lower limit values of every parameter, to eliminate all absurd or possible faulty values. As only one parameter is being studied in *Ventilation* (the RPH), it does not require separate files for each variable, and its limits and labels are established directly in MATLAB ventilation files. One example of this kind of files is presented in *Table 1*, where *generalcalculatedvalues.csv* can be found.

Table 1 - Example of a .csv file (*generalcalculatedvalues.csv*) with its parameter names and limits.

|             | Certificate Number | Window U-value (U <sub>wdn</sub> ) | Window to Floor Area Ratio | Glazing g-value (g <sub>T vi</sub> ) | Shading g-value (g <sub>T</sub> ) | Window area (m <sup>2</sup> ) | Envelope area (m <sup>2</sup> ) | Walls U-value | Roofs U-value |
|-------------|--------------------|------------------------------------|----------------------------|--------------------------------------|-----------------------------------|-------------------------------|---------------------------------|---------------|---------------|
| Lower limit | 0                  | 0.01                               | 0.01                       | 0.01                                 | 0.01                              | 0.01                          | 0.01                            | 0.01          | 0.01          |
| Upper limit | 0                  | 7                                  | 1                          | 1                                    | 1                                 | 80                            | 500                             | 5             | 5             |

In the first row of *Table 1*, the parameter names for labelling are shown. In the second and third rows, the lower limit (0.01 to eliminate absurd small values near zero) and the upper limits of each parameter are defined, respectively.

## 3.2 Processes

Processes section explains the working processes of the model from the nature of data used for *General*, *Envelope* and *Ventilation* to the assembly of the functions of each script, the different methods for extraction, filtering and analyzing to each specific set of data. Processes are divided into *General*, *Envelope* and *Ventilation* with several bullet points to describe each script.

### 3.2.1 General

The largest sample is *General* with 712,182 EPCs. It is distributed according to the variables Construction Year and Context in *Table 2*. Despite of the sample size difference between *General* and *Envelope* (*Table 4*), one can consider both samples to be similarly distributed in terms of Context and Year of Construction apart from the lack of EPCs for ‘2013-2018’ period (in which *Envelope* has around 37 thousand EPCs less than *General*). There were a few invalid EPCs (much less than 1% of the total sample) attributed with Context ‘New’ and Construction Year from earlier periods, which could not be possible and had to be ignored. The period with more EPCs is ‘1980-1990’, this result is supported by Magalhães and Leal [10].

Table 2 - General EPC data.

| <b>General</b>  |                |               |               |                |             |
|-----------------|----------------|---------------|---------------|----------------|-------------|
| Period          | Existing       | New           | Retrofitted   | Total          | %           |
| Before 1919     | 9,768          | 0             | 2,195         | 11,963         | 2%          |
| 1919 - 1945     | 25,865         | 0             | 3,172         | 29,037         | 4%          |
| 1946 - 1960     | 41,171         | 0             | 5,314         | 46,485         | 7%          |
| 1961 - 1970     | 38,032         | 0             | 1,362         | 39,394         | 6%          |
| 1971 - 1980     | 62,024         | 0             | 1,069         | 63,093         | 9%          |
| 1981 - 1990     | 126,912        | 0             | 1,094         | 128,006        | <b>18%</b>  |
| 1991 - 1995     | 59,348         | 0             | 271           | 59,619         | 8%          |
| 1996 - 2000     | 70,518         | 0             | 276           | 70,794         | 10%         |
| 2001 - 2005     | 94,736         | 0             | 172           | 94,908         | 13%         |
| 2006 - 2012     | 55,765         | 1,946         | 1,799         | 59,510         | 8%          |
| 2013 - 2018     | 27,716         | 10,019        | 1,754         | 39,489         | 6%          |
| Pre-certificate | 523            | 68,859        | 97            | 69,479         | 10%         |
| <b>Total</b>    | <b>612,378</b> | <b>80,824</b> | <b>18,575</b> | <b>711,777</b> | <b>100%</b> |
| <b>%</b>        | <b>86%</b>     | <b>11%</b>    | <b>3%</b>     | <b>100%</b>    | <b>-</b>    |

Below, it will be explained how the files above are treated in order to get the desired results. The first MATLAB script created was *Generalanalysis.m*, this file allows the analysis and characterization of the parameters in *General* at a national level, without filtering a specific variable value.

Table 3 shows how *General* files and scripts are related to each other and to which domain belong. To analyse a certain domain of *General* a few MATLAB scripts and .csv files are required, these files are presented for each domain.

Table 3 - General .csv Files and MATLAB Script for Domains.

| <b>General</b>    |                   |                       |
|-------------------|-------------------|-----------------------|
| <b>Domain:</b>    | <b>.csv File:</b> | <b>Matlab Script:</b> |
| National          | ---               | Generalanalysis.m     |
| Region            | regcode.csv       | Generalreg.m          |
| Context           | contextcode.csv   | Generalcontext.m      |
| Construction year | periodcode.csv    | Generalyear.m         |
| Mixed             | All above         | Generalmix.m          |

The files *regcode.csv*, *contextcode.csv* and *periodcode.csv* serve mainly for labelling purposes. The *generaldata.csv* has the parameter and variable values. Finally, the *generallegend.csv* includes the parameter names, also for labelling, and each parameter lower and upper limits. Each MATLAB script will initialize the analysis for each domain.

- **General Analysis**

General Analysis presents the data without any variable filter, just as it is for the whole sample, thus a sample for the entire Portuguese RBS. The following process is described in this paragraph and represented by a flowchart in *Figure 1*. When initializing *Generalanalysis.m*, the required files *generaldata.csv* and *generallegend.csv* are imported, creating data matrixes: *data* and *parametersread*, respectively. After that the parameters in *data* are displayed with their corresponding numerical value (also corresponding to the column in the file) in the command window, using *fprintf*. Here, the user is able to introduce a number for the chosen parameter, due to the use of *input* function.

The chosen numerical value (*k1*) corresponds to a column of the imported matrix *data* that will go through a logical test to check if its values are between the upper and lower limits previously defined ('*parametersread.data(1,k1)*' and '*parametersread.data(2,k1)*'). Consequently, a logical array (*B*) with all valid values is created. The vector *y1* is the vector of the parameter values in the positions of *data* where the logical array has value 1 (true). After discarding the values outside the defined range (i.e. for which the logical test gives the value zero), the vector *y1* includes only the valid values and is ready to be further analysed.

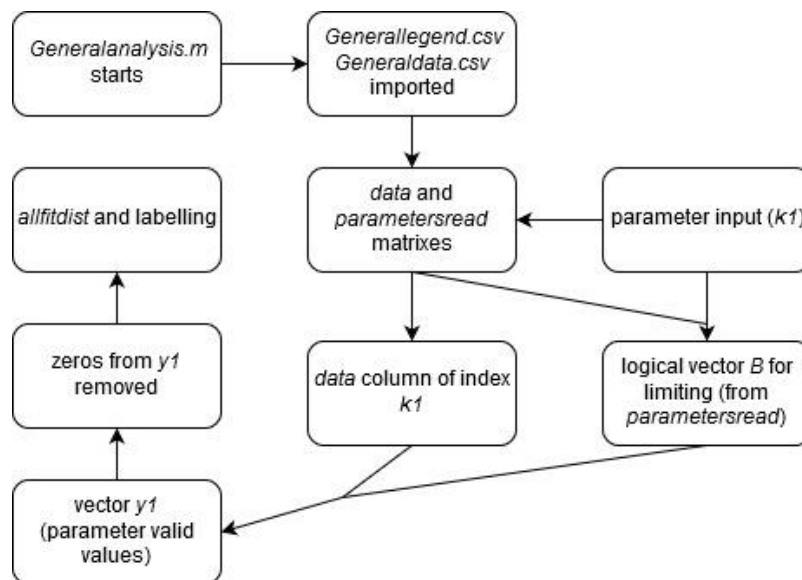


Figure 1 - General Analysis Flowchart.

Then, the *y1* vector is called by the function *allfitdist*. *Allfitdist* is a statistical data processing function that finds the best fitting distribution curve, calculates statistical values such as probability distribution parameters and plots it together with a histogram. The best fitting distribution curve is set by a criteria,

this criteria is the distribution with the lowest  $NLogL$  (negative loglikelihood for estimating the best fitting probability distribution). The probability distribution curve and histogram are plotted in the same figure.

Figure 2 shows an example of the resulting output of the model for the parameter  $RNt$  considering the national RBS, where the best fitting for the data was the probability distribution function ‘rician’.

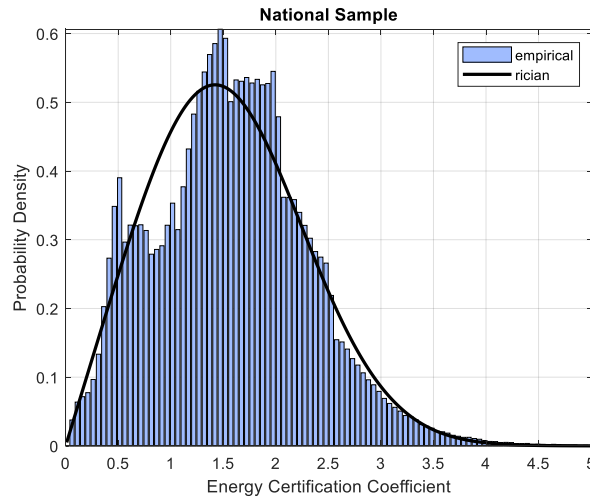


Figure 2 - Energy Certification Coefficient ( $RNt$ ) Probability Distribution Function and Histogram for a National Sample.

- **General Analysis with Variable Filters**

The next step is the introduction of Region, Year of Construction and Context variable constraints resulting in the creation of *Generalreg.m*, *Generalyear.m* and *Generalcontext.m*, and later in *Generalmix.m*. Here, one variable at a time is used to filter the data. Similarly to what was done in *General.m*, *generallegend.csv* and *generaldata.csv* are imported but now, for each variable, a different additional file is also imported: *regcode.csv* for *Generalreg.m*, *periodcode.csv* for *Generalyear.m* and *contextcode.csv* for *Generalcontext.m*.

The procedure used in each of these new MATLAB files is the same, but it includes this additional file and some additional logical code to filter out the pretended output. Additionally, to the parameter input now the variable must be also an input. The available options for both the parameter and variable are printed in the command window, where the user must introduce the chosen options. To simplify, the input for the variable value will be named *Option1*. After the user selects the parameter and variable, and after creating a vector with the parameter value limits, the *find* function finds the entries with *Option1* in the chosen variable column of *generaldata.csv* (e.g. for Context, it finds the rows with the value 1, which corresponds to existing buildings). Then, a logical array is generated, this logical array is used to return the parameter values in matrix *data* with the same index position where there are true values. Zero values in the logical test are again discarded. The procedure from now on is the same as of *General.m*. Here as done in *Generalyear.m* and *Generalreg.m*, two values for variable must be chosen with the goal of comparing both.

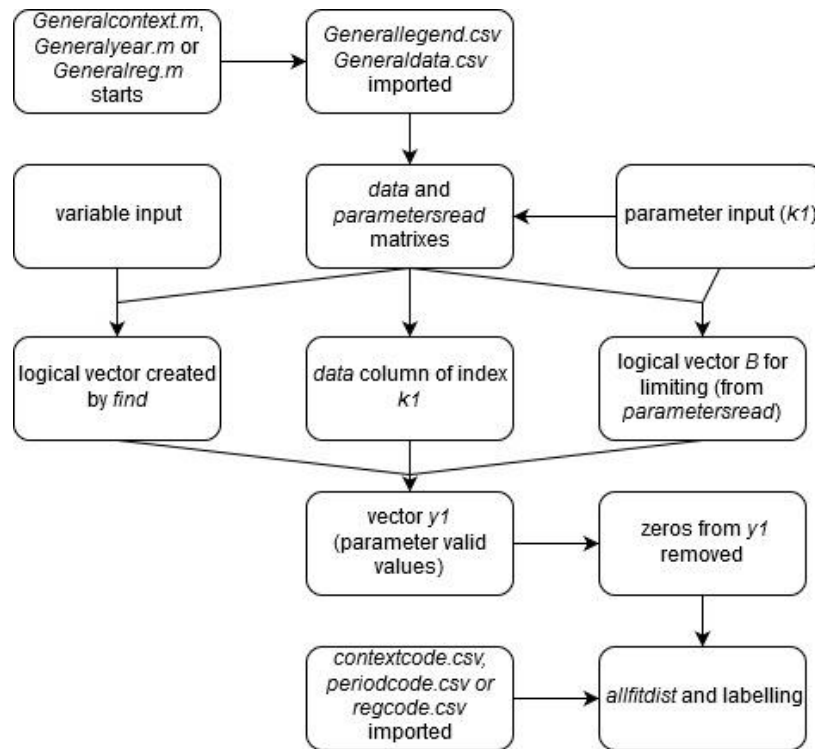


Figure 3 - General Analysis with Variables Flowchart.

- **General Mixed Analysis**

At last, there is still one MATLAB file for the analysis of *General* data. Here, the mixed analysis enables the analysis of the data using a combined filter for two of the three variables. It asks for one parameter to analyse, then asks the user to choose a value for the three variables. Once these options are chosen three different crossed analyses are made: Context is crossed with Year of Construction, Region with Context and Region with Year of Construction. Henceforth the impact of one variable can be evaluated in another, or vice versa, to deeply understand how the RBS is constituted. This analysis is done by using *ismember* function. Each variable option is introduced to create logical arrays, then each of these logical arrays is multiplied with one another and with the vector with values within the parameter limits, finally the product of these logical arrays returns the parameter values, from the values in matrix *data*, in the index position which it is true. The continuation of this analysis is already known, except this time the results are simply the figure generated with *allfitdist* and the probability distribution curve. Its flowchart is identical as in *Figure 3*, except for the use of *ismember* instead of *find* to create a logical vector, together with data column index of *k1* and arrays B1, B2 and B3, will result in the output vectors of *y1*, *y2* and *y3*.

### 3.2.2 Envelope

Envelope data files contain 492,313 EPCs, the lowest figure of EPCs. Its distribution is shown in *Table 4*. As stated before, *Envelope* has around 37 thousand EPCs less than *General* for 2013-2018 period and had a residual number of invalid EPCs ( $\ll 1\%$ ) for ‘New’ buildings dating from periods earlier than 2006.

Table 4 - Envelope EPC data.

| <b>Envelope</b> |                |               |              |                |             |
|-----------------|----------------|---------------|--------------|----------------|-------------|
| Period          | Existing       | New           | Retrofitted  | Total          | %           |
| Before 1919     | 6,922          | 0             | 1,214        | 8,136          | 2%          |
| 1919 - 1945     | 18,109         | 0             | 1,634        | 19,743         | 4%          |
| 1946 - 1960     | 29,050         | 0             | 2,871        | 31,921         | 7%          |
| 1961 - 1970     | 26,748         | 0             | 764          | 27,512         | 6%          |
| 1971 - 1980     | 43,673         | 0             | 578          | 44,251         | 10%         |
| 1981 - 1990     | 89,567         | 0             | 656          | 90,223         | <b>20%</b>  |
| 1991 - 1995     | 41,569         | 0             | 161          | 41,730         | 9%          |
| 1996 - 2000     | 49,815         | 0             | 181          | 49,996         | 11%         |
| 2001 - 2005     | 67,009         | 0             | 114          | 67,123         | 15%         |
| 2006 - 2012     | 21,470         | 1,804         | 987          | 24,261         | 5%          |
| 2013 - 2018     | 2,216          | 12            | 5            | 2,233          | 1%          |
| Pre-certificate | 351            | 38,785        | 59           | 39,195         | 9%          |
| <b>Total</b>    | <b>396,499</b> | <b>40,601</b> | <b>9,224</b> | <b>446,324</b> | <b>100%</b> |
| <b>%</b>        | <b>89%</b>     | <b>9%</b>     | <b>2%</b>    | <b>100%</b>    | <b>-</b>    |

- **Envelope Preparation**

The most complex procedure in preparing the data to be analysed is the transformation of *Envelope* files into a shorter file with only essential information. This file is *calculatedvalues.csv* and was prepared using *testvalue.m*. The latter function will produce the parameters *envelope area* and *window area* that will help to calculate *window U-value (U<sub>wdn</sub>)*, *Window-to-Floor Ratio*, *glazing g-value (gt<sub>vi</sub>)*, *shading g-value (gT)*, *walls and roofs U-value*. This *testvalue.m* script does not ask for input, by running every parameter will be treated without any distinction.

Table 5 shows how *Envelope* files and scripts are related to each other and each domain. To analyse a certain domain of *Envelope* a few MATLAB scripts and .csv files are required, these files are presented for each domain.

Table 5 - Envelope .csv Files and MATLAB Script for Domains.

| Envelope          |                 |                    |
|-------------------|-----------------|--------------------|
| Domain:           | .csv File:      | Matlab Script:     |
| National          | ---             | Envelopeanalysis.m |
| Region            | regcode.csv     |                    |
| Context           | contextcode.csv | Envelopecontext.m  |
| Construction year | periodcode.csv  | Envelopeyear.m     |
| Mixed             | All above       | Envelopepemis.m    |

The .csv files regcode.csv, contextcode.csv and periodcode.csv serve mainly for labelling purposes, *generaldata.csv* has the variable values and *calculatedvalues.csv* has the parameter values, *calculatedvalueslegend.csv* has the parameter names and their limits. Each MATLAB file will initialize the analysis for each domain.

The first parameter determined was *window U-value (Uwdn)*, naturally calculating *window area* first. For this purpose, a matrix with three columns was created with information from *Envelopetreated* files with: certificate number and *window area* of every element and multiplication of each *window area* element with each element of *Uwdn*, named *Envelopematrix*. Applying the function *unique* to the first column of *Envelopematrix*, the certificate numbers will become unique and organized in increasing order. Concerning the other two columns, the function *acummarray* sums the values in second and third columns by the same certificate number in first column. Then, we will get a matrix with a unique certificate number and the total *window area* of each certificate and the average *Uwdn* value for each certificate. Finally, to get the *Uwdn Average* parameter one must divide the third and the second column. At the end of this step both *Uwdn Average* and *Window Area* parameters are calculated and transferred to *calculatedvalues.csv*. The equation is given for dwelling *k* (each certificate number) by:

$$Uwdn_{dwelling\ k} = \frac{\sum(Uwdn_{element} \times Area_{element})}{(Window\ Area)_{dwelling\ k}} \quad [1]$$

The next parameter prepared is *Window to Floor Area Ratio*, as total *window area* is already calculated, the *net floor area* parameter is taken from *generaldata.csv* and then the first parameter is divided by the second. Creating, thus, *Window to Floor Area Ratio* column values. This parameter follows the equation below for dwelling *k*:

$$Windowtofloor_{dwelling\ k} = \frac{(Window\ Area)_{dwelling\ k}}{(Floor\ Area)_{dwelling\ k}} \quad [2]$$

Parameters like  $gt\ vi$  and  $gT$ , have the same calculation method. There is a matrix with 3 columns starting with certificate number, *window area* and the products between *window area* with  $gT$  or *window area* with  $gt\ vi$ , for each element. Then, as previously made for *Uwdn*, using *unique* and *accumarray* a matrix is generated with columns for the unique certificate numbers, the total *window area* for each certificate number and the total  $gT$  or  $gt\ vi$ . To get an average  $gT$  and  $gt\ vi$  its total values for each dwelling are divided by its total *window area* values. The following equations, *Equation 3* and *4*, give the average  $gT$  and  $gt\ vi$  for dwelling  $k$  windows.

$$gt\ vi_{dwelling\ k} = \frac{\sum(gt\ vi_{element} \times Window\ Area_{element})}{(Window\ Area)_{dwelling\ k}} \quad [3]$$

$$gT_{dwelling\ k} = \frac{\sum(gT_{element} \times Window\ Area_{element})}{(Window\ Area)_{dwelling\ k}} \quad [4]$$

The *envelope area* is calculated exactly as *window area*. Afterwards the total *envelope area* is essential to calculate both *walls and roofs U-value*. Again, we create a large matrix with the certificate numbers, *envelope area* and *U-value* whether the element is a wall or a covering, then the product between these two, and a column with 0,1 and 2 to allow identifying if an element is a wall (for 1), a covering (for 2) or none of them (for 0). We use function *ismember* twice to create two logical arrays to separate walls and roofs, these arrays are multiplied with the third column of the matrix, the one with the product of *envelope area* and U-values. Thus, the result will be two vectors, one with the product of *envelope area* and walls and roofs U-values. Then, finally, these files go through the same process of *unique* and *accumarray* functions and are ready to be transferred to *calculatedvalues.csv*. The following equations demonstrate the method for average wall and roof thermal transmittance calculation, in dwelling  $k$ .

$$U_{wall\ dwelling\ k} = \frac{\sum(U_{wall\ element} \times Area_{element})}{(Roof\ Area)_{dwelling\ k}} \quad [5]$$

$$U_{roof\ dwelling\ k} = \frac{\sum(U_{roof\ element} \times Area_{element})}{(Roof\ Area)_{dwelling\ k}} \quad [6]$$

- **Envelope Analysis**

Once all parameters are calculated, the MATLAB functions *Envelopeanalysis.m*, *Envelopereg.m*, *Envelopeyear.m*, *Envelopecontext.m* and *Envelopemix.m* are explained. *Envelopeanalysis.m* works the same as *General.m* as schematized in *Figure 4*.

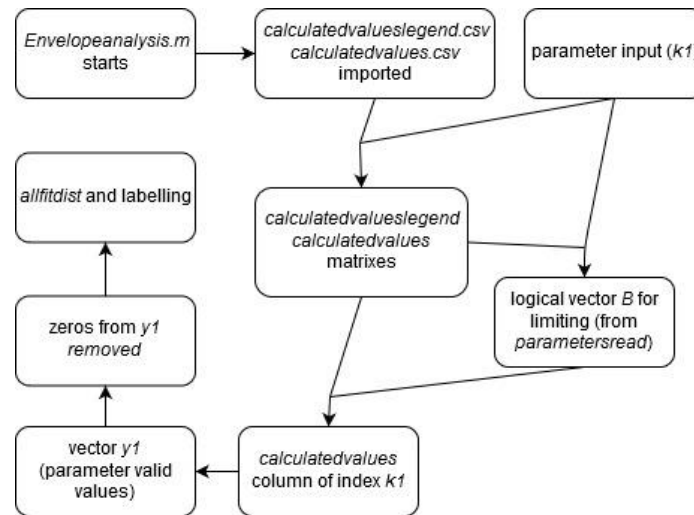


Figure 4 - *Envelopeanalysis.m* Flowchart.

In continuation to *Envelope.m*, Envelope MATLAB scripts that require a variable input must import *generaldata.csv*. There will be a logical test with *find* to check which values belong to the chosen input variable from *generaldata.csv*, to filter these values creating logical array *C1*, an array that will be multiplied with another logical array, the upper and lower limit filtering array, creating vector *y1*. Apart from this particularity it works like *Generalreg.m*, *Generalyear.m*, *Generalcontext.m*, *Generalmix.m*, respectively. The flowchart in *Figure 5* schematizes this process. *Envelopemix.m* scheme would resemble just like *Figure 5* with the difference of generating and calculating at the same time three different logical arrays (*C1*, *C2*, *C3*) and three output vectors (*y1*, *y2*, *y3*).

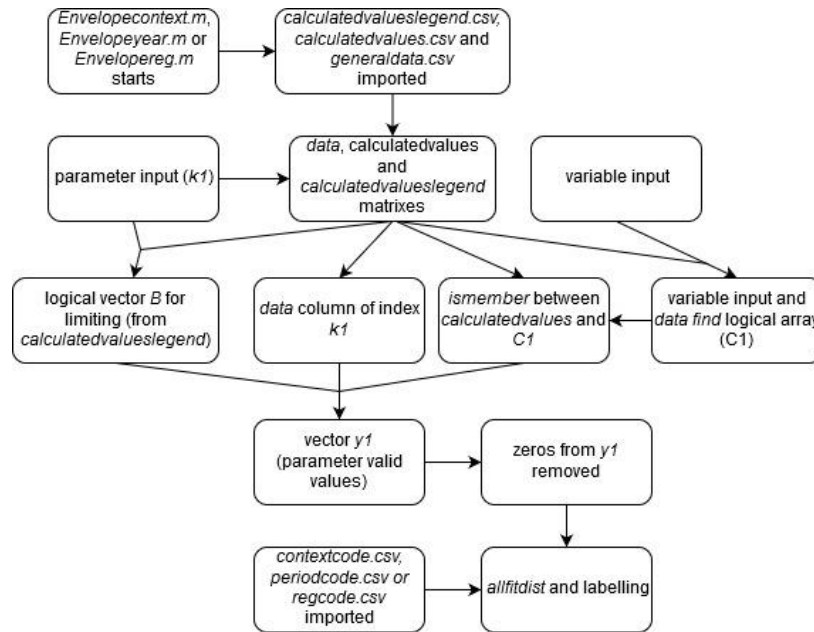


Figure 5 - Envelope Analysis with Variables Flowchart.

### 3.2.3 Ventilation

The last set of data is related to Ventilation, which differently from the other, has only two files and one extra to prepare, *Ventilation.m* and *Ventilationconstraints.m*, the preparation file is *Ventilationprepare*. It has 686,840 EPCs. The only parameter studied here is Air changes per hour (RPH), this is a parameter related to ventilation. Ventilation can be intentional or unintentional, intentional air flow depends on fans, grids or windows, while unintentional ventilation is due to infiltrations, which is uncontrollable airflow and depends on air tightness of building envelope.

*Ventilationprepare.m* finds which certificate numbers are in Ventilation data original files and gives back the variables of the same certificates, using *find* of which will be presented in *Ventilation.csv* next to the parameter *RPH*. The file *Ventilation.m* does not need any additional file due to using only one parameter, the limits and labels are written directly in MATLAB, its only output is the National analysis for *RPH*. The code structure is identical to *General.m* and *Envelope.m*.

What differences *Ventilationconstraints.m* from the other variable analysis files is that all variables can be studied with solely this file, instead of having three differentiated files. This is due to the information lightness of having just one parameter. After running this file, the user will be asked to select a variable. After making this decision, the file will run and filter the values using again *find* for the input made. Since the variable values are already in *Ventilation.csv*, there is no need to cross it with *generaldata.csv* as before. Then, the process repeats using *allfitdist* to elaborate the distribution functions and histograms.

### 3.2.4 Main Function

The aim of this main function denominated *main.m* is to guide the user to initialize the appropriate file. This way the user will get in touch to the files without knowing them, as it asks which part of the data is to analyse and in which domain (national, mixed or by each variable), then it runs the file for the selected input. It is explained in the chapter 3.5 - *Fast Guide*.

The following figures represent how files and scripts interact with each other. *Figure 6* contains a flowchart of the interaction of non-filtered variable (national analysis) results like *General.m*, *Envelope.m* and *Ventilation.m*. We can observe which .csv data files are imported to each script, and their relationship with each other. *Figure 7* shows the same but for filtered variable results, where variables play a role in data analysis, like *Generalyear.m*, *Envelopereg.m* or *Envelopemix.m*.

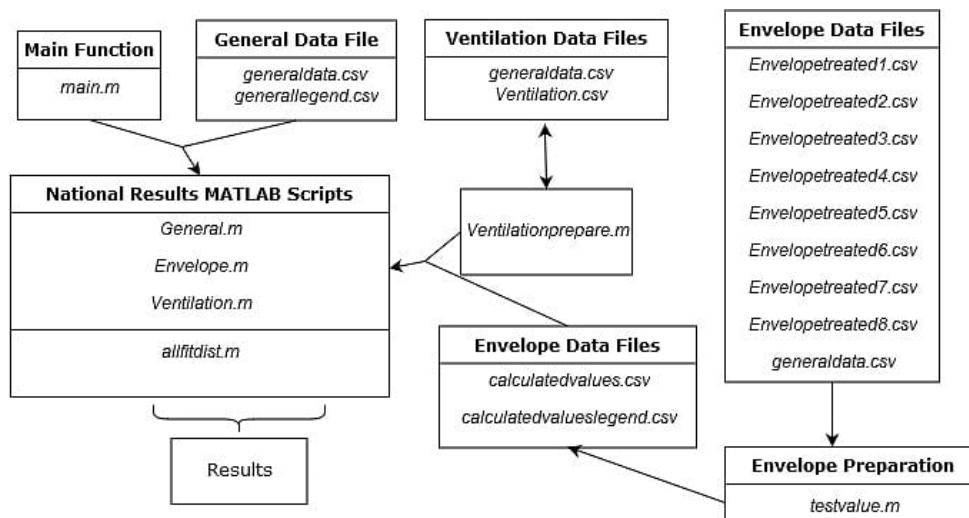


Figure 6 - Flowchart of File Relationship for Non-filtered Variable Results.

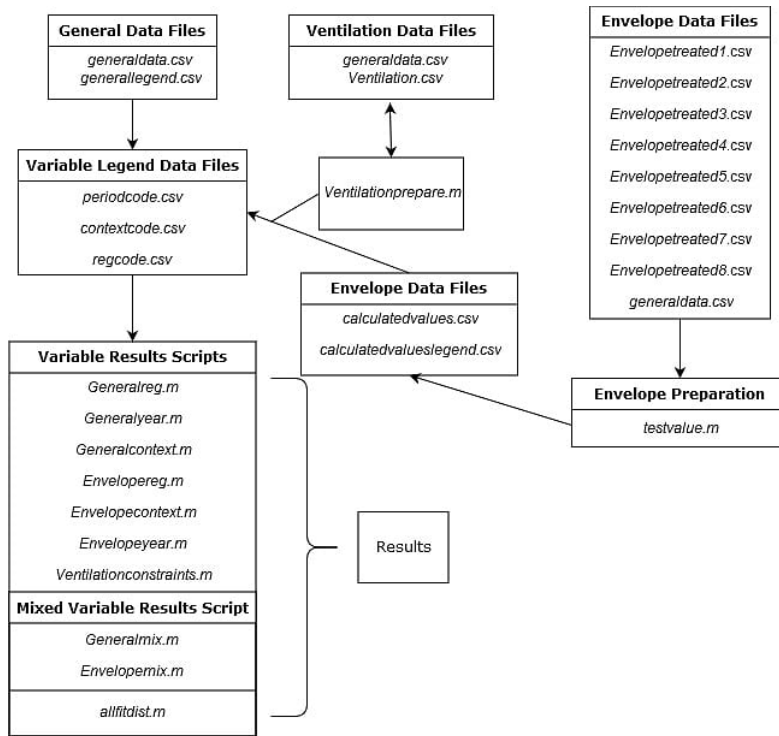


Figure 7 - Flowchart of File Relationship for Filtered Variable Results.

### 3.3 Output display

Output display demonstrates briefly the model results and capabilities as the model is finished by now. The goal is to prepare the user to what the outputs will look like and what to expect from the model.

After having the probability distribution parameters calculated, the objective is to isolate the probability distribution curve from the histogram. The variable *dist1* gets the probability distribution name attributed to it, stored inside it in *DistributionName*. Then *makedist* function generates a probability distribution object using the probability distribution parameters specified by the probability distribution in *dist1*, as they are inside *dist1* in the ParameterNames section, for example rician are *non-centrality* and *scale* (*s* and *sigma*).

The x axis consists on an array starting in the lower limit value and ending in the upper limit value of the parameter, adding 0.01 to each value from the initial to the final. On the other hand, the y axis array is formed using *pdf* function, which returns a probability distribution function with the introduction of probability distribution parameters, name and x axis array.

A figure is started and the arrays  $x1$  and  $y1$  are put in a `plot` function to display its results, with the corresponding labelling and title. Later a legend is attached to the figure containing the probability distribution function name and each one of the probability distribution parameters.

The result can be seen in the following figure, where the probability distribution curve is plotted without the histogram, differently from *Figure 2*, its legend is presented with function name and parameters as we can observe, found in *dist1*.

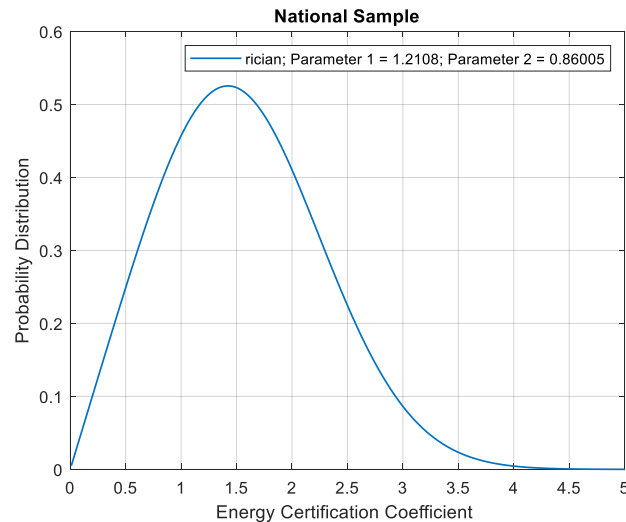


Figure 8 - Energy Certification Coefficient ( $Rnt$ ) probability distribution for the National sample.

One more statistical evaluation of the sample is to be done, this last statistical analysis is made through `boxplot` function. A boxplot is a way to display the distribution in a sort of box limited according to median, first quartile, third quartile, minimum non-outlier value and maximum non-outlier value. The `boxplot` function has as inputs the array  $y$  previously calculated, and the automatic label for  $k1$  input. Also, a textbox is created where important values are displayed to give a more accurate perception to the user, these values are calculated separately with `prctile` function and are median (*median*), first quartile ( $Q1$ ), third quartile ( $Q3$ ), minimum non-outlier value ( $L1$ ) and maximum non-outlier value ( $L2$ ).

One result is presented in *Figure 9*. The blue box is where the values between  $Q1$  and  $Q3$  are constrained, the red line represents the median and the two lines, above and below, are the maximum and minimum non-outlier values, respectively.

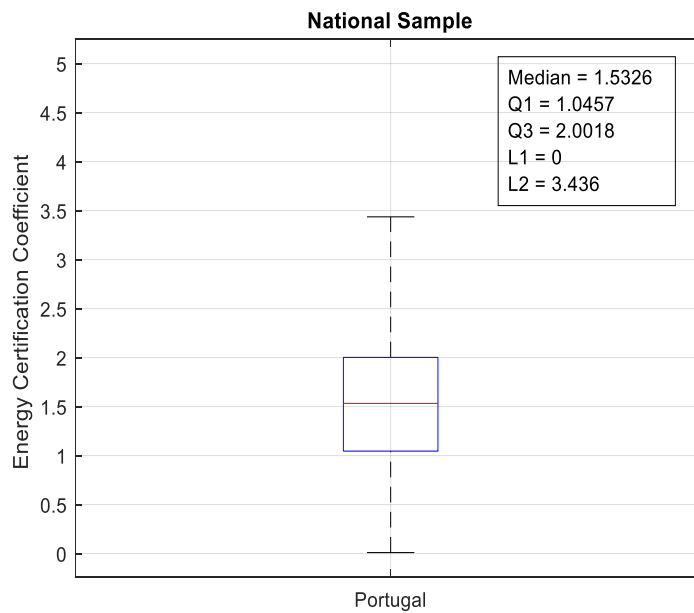


Figure 9 - Boxplot of Energy Certification Coefficient ( $RNt$ ) for National Sample with Quartile Values ( $Q1$ , Median and  $Q3$ ) and Non-outlier Values ( $L1$  and  $L2$ ) displayed.

Results from the “General Analysis with Variables” – are displayed in the next two figures. In *Figure 10* we have a comparison between contexts ‘New’ and ‘Existing’ for  $RNt$ . It is possible to observe - with the help of boxplots in *Figure 11* – a significant difference between the values of  $RNt$  in ‘New’ and in ‘Existing’, having almost all new dwellings a  $RNt$  below 1, while the RBS elements that existed before have a wider range of values, with an average of 1.66.

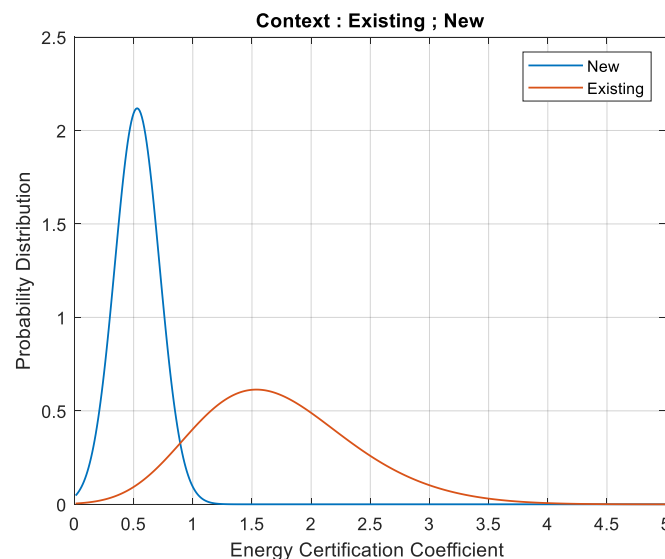


Figure 10 - Existing vs New Energy Certification Coefficient ( $RNt$ ) probability distribution.

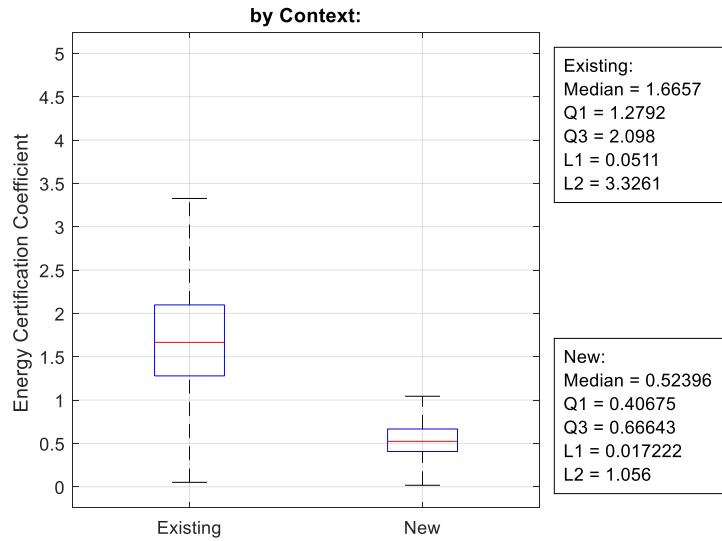


Figure 11 - Existing vs New Boxplot comparison for Energy Certification Coefficient ( $Rnt$ ).

Figure 12 shows one example of a possible result of the mixed analysis of a Region and a Construction Year, where the  $Rnt$  of the dwellings in Region “Beira Interior Sul” built in the period ‘1945-1960’ is displayed.

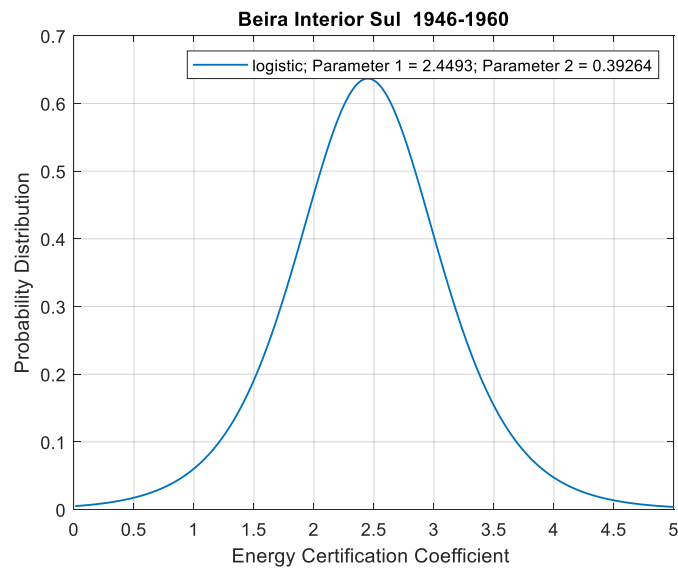


Figure 12 - Energy Certification Coefficient ( $Rnt$ ) for ‘Beira Interior Sul’ for the construction period ‘1946-1969’.

### 3.4 Challenges

This MATLAB work suffered deep changes since the time it was started, as every new MATLAB function discovered through the model construction phase could disruptively restructure what had already been made at the time. The processes explained until here are totally different from what they were in an earlier phase, the earlier model draft was based on multiple iteration cycles. This type of approach was time consuming, as some MATLAB files had a running time from thirty minutes to one hour against the one minute to five minutes span currently. These time-consuming iteration cycles through the model construction phase were gradually replaced by MATLAB functions that could do the same or similar task in seconds. In the earliest draft the parameter studied was *Net Floor Area* due to the greater sensibility we have compared with other parameters, to allow identifying and correcting mistakes, or to improve MATLAB code performance. Thus, contributing to create better working and functional MATLAB model. Nevertheless, other challenges faced include which file format was the best to import data files, the preference went easily to .csv because it is simple to make and accept numerical and character information. Finding the best automatic representation of labels in figures took some time of research, as most of examples were too simpler than what was pretended and differed from type of figure.

### 3.5 Fast Guide

To initialize the computational model, the function *main.m* must be initialized. After this step the user will be asked in the command window which domain would like to analyse between *General*, *Envelope* or *Ventilation* and when choosing the desired option another question will pop up (see *Figure 13*) in the command window asking which of the three variables is to analyse or if its national sample output or mixed between variables (see *Figure 14*). Finally, the file, which is satisfied by these two inputs, will run the desired script and ask for further input (see *Figure 15*). Then the desired output is generated.

In case of making a mistake when choosing the input in command window, rather than waiting until the process ends it is recommended, to just input one option that is not listed, the program will find an error and stop. If the file is already running the option to interrupt is to click in 'Pause' then in 'Quit Debugging'.

This *main.m* allows the user to find the right MATLAB file to generate the desired output out of the many files available, using *main.m* allows the usage of the computational model accessible to everybody without having to know each individual script.

```
Command Window
1 - Data from Envelope
2 - Data from General
3 - Data from Ventilation
fx Data: |
```

Figure 13 - MATLAB main.m Script First Input.

```
Command Window
datafrom =
    1
What to analyse:
1 - National Sample
2 - Context
3 - Region
4 - Construction year
5 - Mixed
fx Option: |
```

Figure 14 - MATLAB main.m Script Second Input.

```
Command Window
19 - Ave
20 - Alentejo Central
21 - Minho-Lima
22 - Baixo Alentejo
23 - Alto Alentejo
24 - Entre Douro e Vouga
25 - Pinhal Interior Norte
26 - Douro
27 - Beira Interior Norte
28 - Pinhal Interior Sul
fx Region 1: |
```

Figure 15 - MATLAB Envelopereg.m Script Input.

## 4 - Method Applicability

In *Chapter 4* it is pretended to demonstrate how can the computational model answer to questions concerning the RBS, showing examples of solutions to possible questions. The possible questions are obviously vast so it will be impossible to answer to all. Here questions will range the work because there are unlimited possibilities, for the goal is to provide a tool, the computational model, to provide the answer for questions in future RBS related projects.

### 4.1 How Air Changes Per Hour (RPH) differs in the Portuguese RBS?

Air changes per hour (RPH) is given by the ratio between the volume air removed or added to a space by the volume of that same space. Parameter RPH changes only slightly when analysed in all RBS for a multitude of possibilities in each variable. But patterns have been identified regionally between more urban regions and more rural regions, as expected. In *Figure 16* we analyse the entire RBS sample with probability distribution and histogram for the air changes per hour.

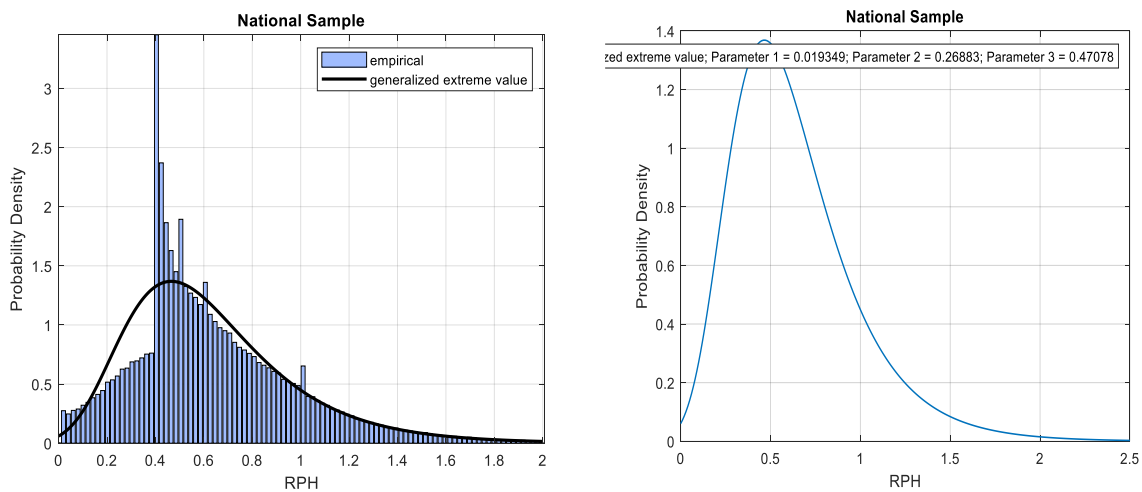


Figure 16 - Air Changes Per Hour (RPH) at National Level Probability Distribution Function and Histogram (Left). Air Changes Per Hour (RPH) at National Level Probability Distribution Function (Right).

The fitting probability distribution function is a *generalized extreme value* with distribution parameters *shape*, *scale* and *location* of 0.0193, 0.269 and 0.471, respectively.

Representing urban areas, the regions ‘Grande Lisboa’ and ‘Grande Porto’ were selected due to hosting the most populous cities in Portugal and the most habitational dense of all NUTS III, ‘Grande Lisboa’ with 495.6 dwellings/km<sup>2</sup> and with 408.9 dwellings/km<sup>2</sup> ‘Grande Porto’ in 2015 [12]. Figure 17 and Figure 18 show the RPH distribution function for ‘Grande Porto’ and ‘Grande Lisboa’, respectively.

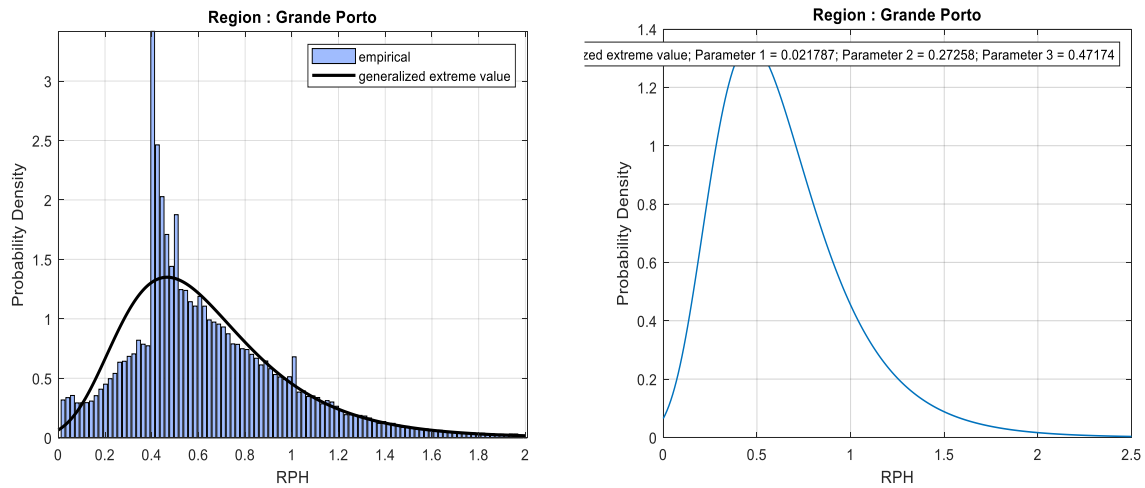


Figure 17 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Grande Porto’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function in ‘Grande Porto’ (Right).

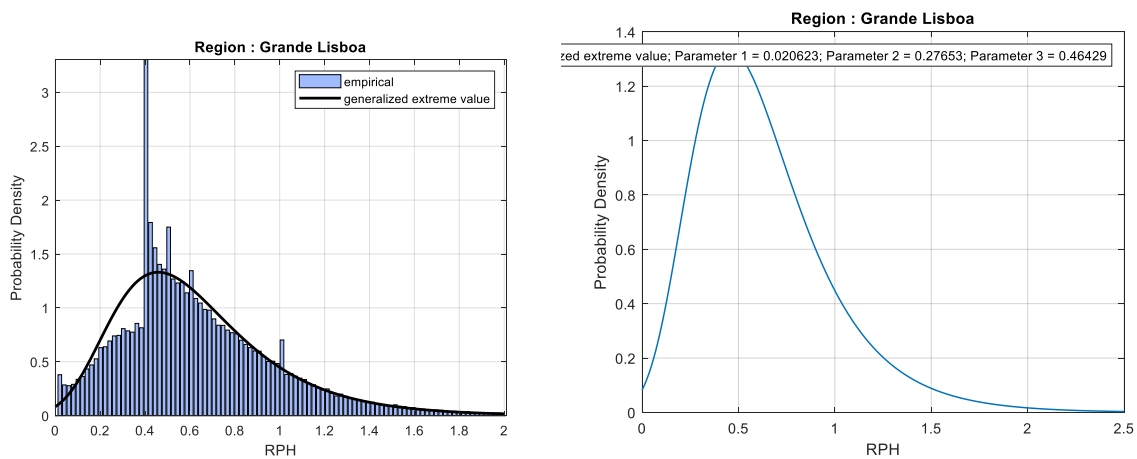
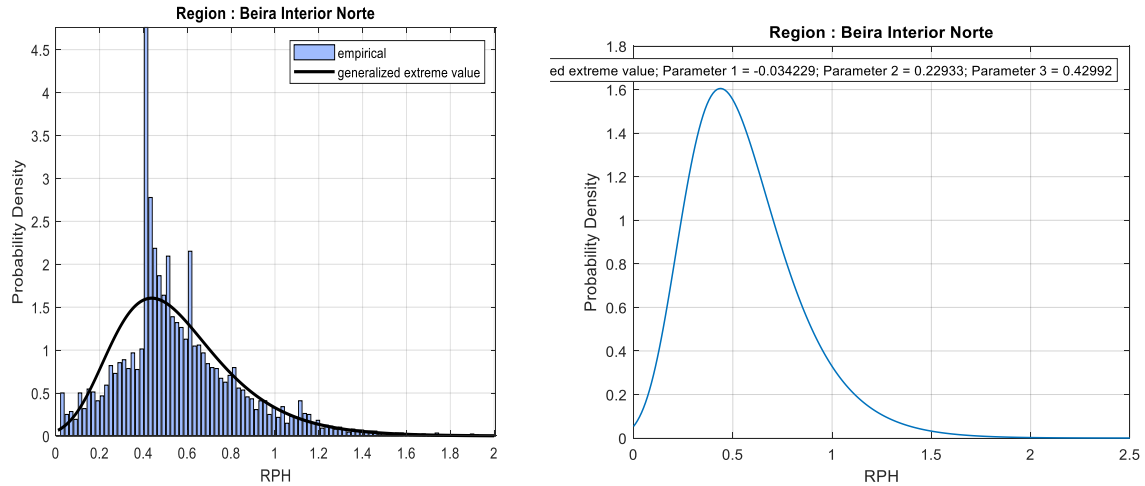


Figure 18 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Grande Lisboa’ (Left). Air Changes Per Hour (RPH) Probability Distribution Function in ‘Grande Porto’ (Right).

As we can observe both curves are a generalized extreme value with similar probability distribution curve parameters. The previous figures characterise the RPH parameter of typical urban areas in the RBS. Now we take the example of one rural region and compare with these urban regions. The rural region taken as an example is ‘*Beira Interior Norte*’, a region with low habitational building density, of  $28.8 \frac{\text{dwellings}}{\text{km}^2}$  (this value includes also ‘*Serra da Estrela*’ another region characterised by low residential building density [12]), this example is presented in *Figure 19*. These two regions are compared together with ‘*Beira Interior Norte*’ in *Figure 20*.



*Figure 19 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram in ‘Beira Interior Norte’ (Left), Air Changes Per Hour (RPH) Probability Distribution Function in ‘Beira Interior Norte’ (Right).*

Probability distribution functions are found to be *generalized extreme value* for regions whether more urban or rural. When comparing directly ‘*Grande Lisboa*’ and ‘*Grande Porto*’ with ‘*Beira Interior Norte*’ we come across this minor difference. As shown in *Figure 20*, the distribution of RPH in ‘*Grande Lisboa*’ is rather sparser with the typical peak around 0.5 RPH lower than in ‘*Beira Interior Norte*’. It means that in ‘*Grande Lisboa*’ the RPH is a little higher than in ‘*Beira Interior Norte*’. This might be because buildings in urban areas are taller than in rural areas, and there the wind is more intense, resulting in higher RPH. This pattern has been consistently observed while using the computational model.

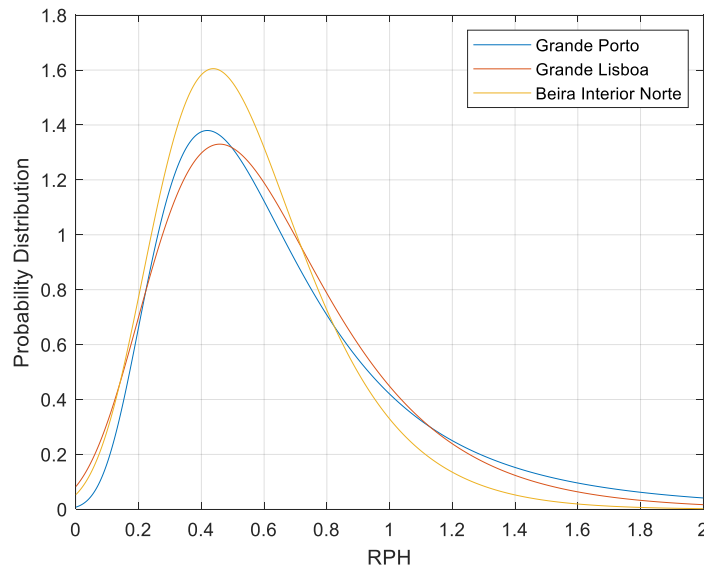


Figure 20 - Comparison of Parameter RPH in Regions 'Grande Lisboa' and 'Grande Porto' with 'Beira Interior Norte'.

The regional differences in RPH show a pattern, now for the same parameter it is pretended to analyze in terms of context to observe the variations between distinct contexts, the three contexts are expected to be differentiated in a more abruptly way. 'Existing' and 'New' are characterised in Figures 21 and 22.

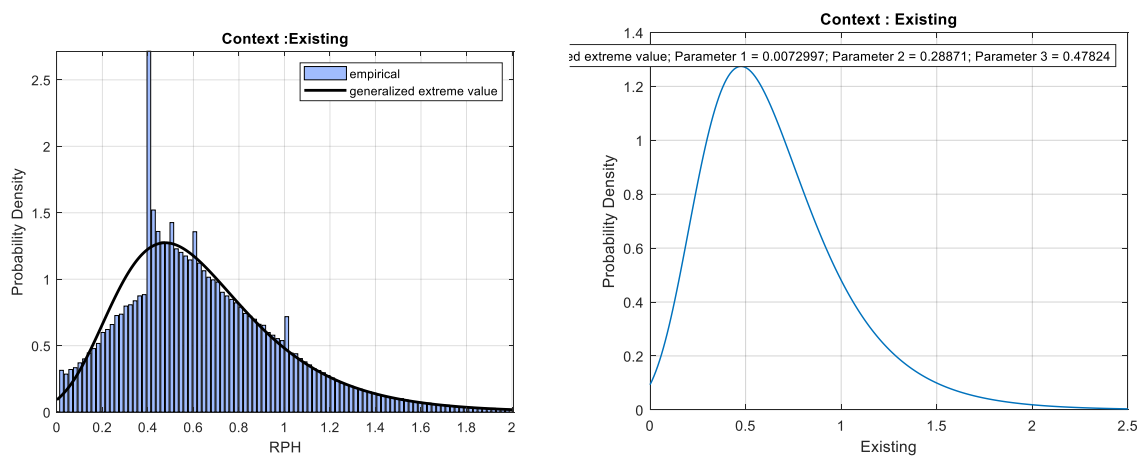


Figure 21 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of 'Existing' (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'Existing' (Right).

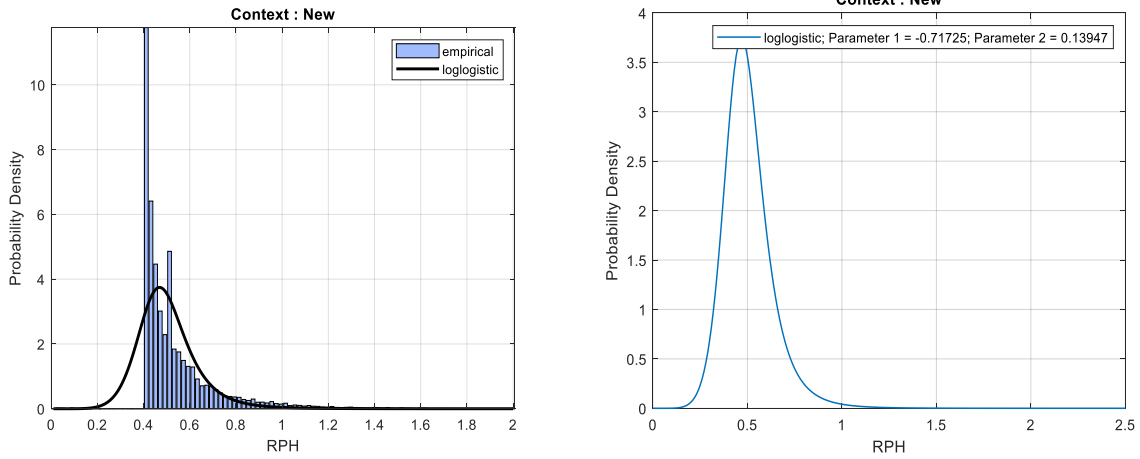


Figure 22 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of 'New' (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'New'(Right).

RPH of buildings that went through a 'Major Renovation' is presented in Figure 23. All probability distribution functions are in Figure 24.

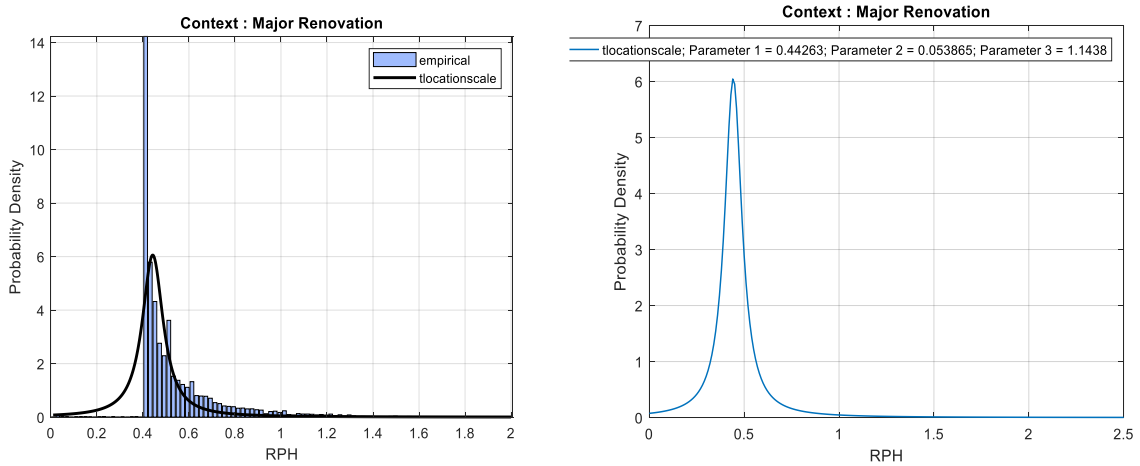


Figure 23 - Air Changes Per Hour (RPH) Probability Distribution Function and Histogram of 'Major Renovation' (Left). Air Changes Per Hour (RPH) Probability Distribution Function of 'New'(Right).

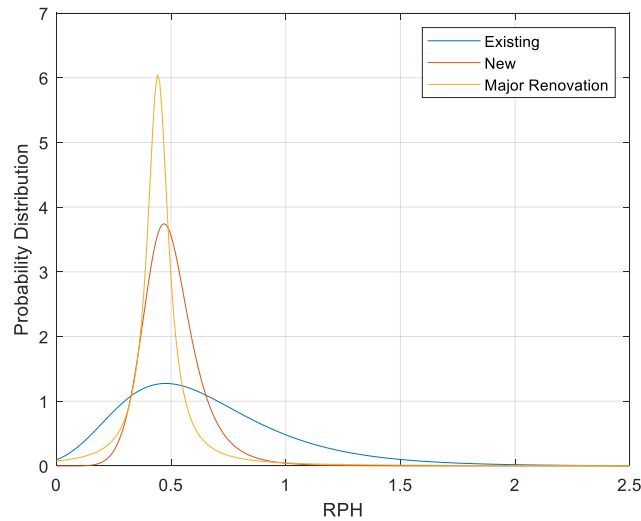


Figure 24 - Comparison of RPH Between All Contexts.

When comparing both probability distribution functions, ‘Existing’ buildings have a higher diversity in results due to excess of ventilation usually found in older construction and lack of ventilation, peaked around 0.5 with a wider distribution function from 0 to 1.5. ‘New’ constructions have an RPH value between 0 and 1 with a peak around 0.4. Due to an increased importance added to REH about air quality, recent dwellings are found to be more air tightened and allow the necessary airflow. ‘Existing’ curve is a *generalized extreme value*, while ‘New’ is a *loglogistic*. For the context of ‘Major Renovation’, the best fitting is a probability distribution function *tlocationsscale*. Although ‘Major Renovation’ and ‘New’ might appear to have the same shape when compared we notice that they are quite different with different distribution probability functions *tlocationsscale* and *loglogistic*. ‘Major Renovation’ has a greater incidence around 0.4 RPH and the RPH values a bit lower than ‘New’. This improvement is due to improvements in building construction and air quality regulations, air infiltration points are insulated to a point where air flow is enough to fulfil the requirements of air quality and thermal losses are minimized.

## 4.2 How does the building energy performance ( $Rnt$ ) evolve in time?

The  $Rnt$  determines the energy performance of a dwelling and is the ratio between  $Ntc$  value, the annual total primary energy needs, and  $Ntc$  limit, a reference limit value of the annual total primary energy needs [13]. The equation below calculates  $Rnt$  for dwelling  $k$ .

$$Rnt_k = \frac{Ntc_{value_k}}{Ntc_{limit_k}} \quad [7]$$

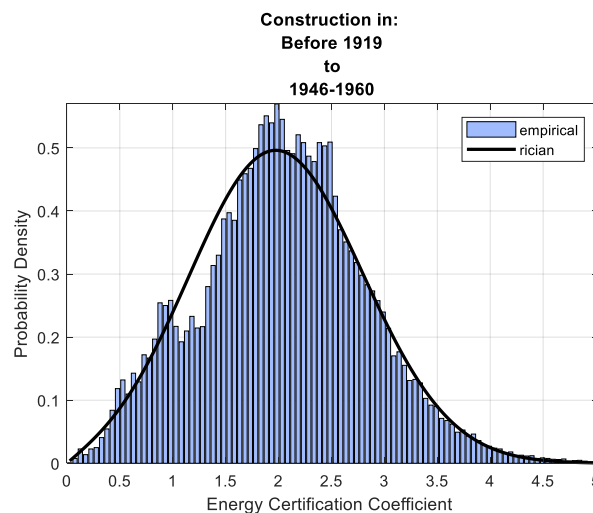
According to *Table 6*, adapted from [13], the lower this figure the best the performance is in a scale from F to A+, where F accounts for a  $RNt$  value of 3.00 and A+ below 0.25 [13].

*Table 6 - Energy Classification According to  $RNt$  Parameter, adapted from [13].*

| Classification | $RNt$                  |
|----------------|------------------------|
| A+             | $RNt \leq 0.25$        |
| A              | $0.25 < RNt \leq 0.50$ |
| B              | $0.50 < RNt \leq 0.75$ |
| B-             | $0.75 < RNt \leq 1.00$ |
| C              | $1.00 < RNt \leq 1.50$ |
| D              | $1.50 < RNt \leq 2.00$ |
| E              | $2.00 < RNt \leq 2.50$ |
| F              | $2.50 < RNt \leq 3.00$ |
| G              | $3.00 > RNt$           |

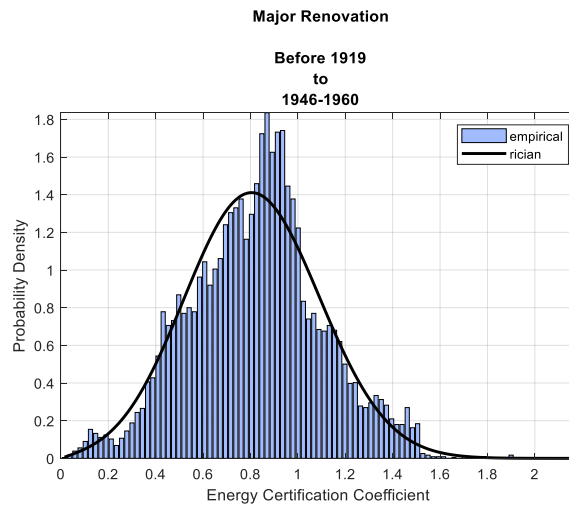
First, we have a snapshot of the entire RBS sample of how parameter  $RNt$  is distributed. In *Figure 2* we have the probability distribution and histogram, *Figure 8* the probability distribution function alone, which is a *rician*. Finally, in *Figure 9* a boxplot shows statistical values. This *rician* probability distribution curve has, according boxplot diagram in *Figure 9*, a median of 1.53 (rank D), and the third quartile (Q3) of 2.00 (rank E). It shows that a lot of dwellings in Portugal have a poor energy performance, when analyzing all RBS sample, with dwellings built in a huge variety of construction year periods. But when analyzing each period individually and comparing with different construction year periods, we will observe how energy policy legislation has impacted the RBS over time.

First, the oldest buildings are presented for the period of construction ‘*Before 1919-1960*’, *Figure 25*.



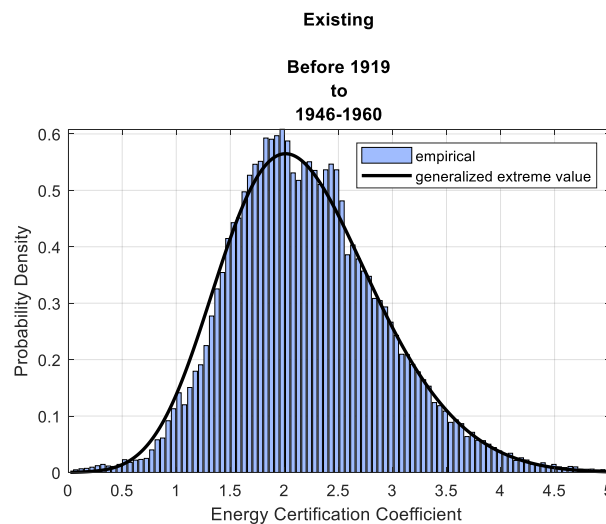
*Figure 25 -  $RNt$  for Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram.*

The histogram allows us to identify a local maximum in *Figure 25* between *RNt* of 0.5 and 1.0. Observing the ‘Major Renovation’ distribution function (*Figure 26*), it is possible to conclude that such local maximum may be originated by retrofitted buildings constructed in the studied period (‘Before 1919-1960’) that have an energy certification coefficient lower than 1.5. Using information from *Table 2* (see subsection 3.2.1), it is estimated that buildings gone under ‘Major Renovation’ represent 12.2% of the sample for this period, while in RBS total sample they represent just 3%.



*Figure 26 - RNt for Retrofitted Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram.*

If we exclude ‘Major Renovation’ from dwellings built during this period (only ‘Existing’), we will have a distribution as in *Figure 27*.



*Figure 27 - RNt for 'Existing' Buildings Constructed in 'Before 1919-1960' Probability Distribution Function and Histogram.*

Without ‘Major Renovation’ the probability distribution function differed from a *rician* to a *generalized extreme value*.

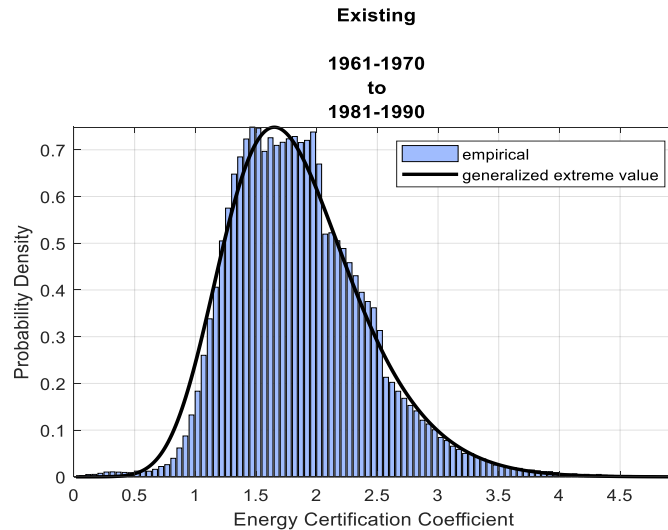


Figure 28 -  $RN_t$  for 'Existing' Buildings Constructed in '1961-1990' Probability Distribution Function and Histogram.

'Existing' dwellings from '1961-1990' have a distribution curve of *generalized extreme value*, in Figure 28. There is a reason why these periods were chosen, 'Before 1919 – 1960' shows the RBS before a period of time, '1961-1990' when construction was booming due to population growth, many Portuguese returned from Africa after the Colonial War and rural exodus, dwellings in this period were often mass produced and the cheapest possible to construct, thus these dwellings are often characterised by low quality.

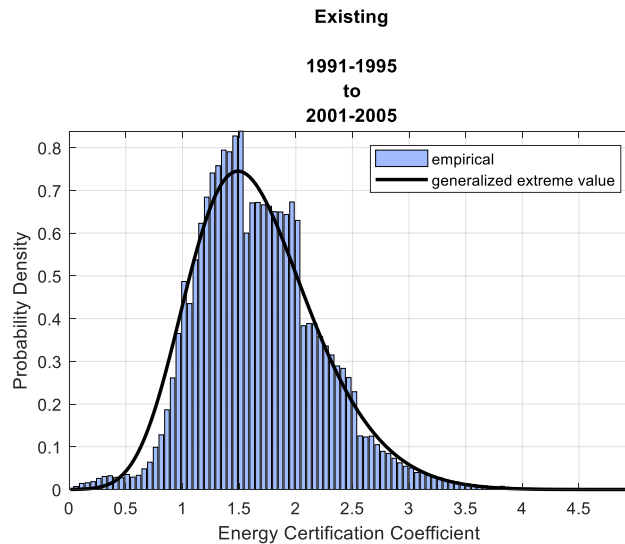


Figure 29 -  $RN_t$  for 'Existing' Buildings Constructed in '1991-2005' Probability Distribution Function and Histogram.

Also, for the period of '1991-2005' in Figure 29, with only 'Existing' dwellings the distribution function is a *generalized extreme value* with a median of 1.60. In '1991-2005' RCCTE was already approved (in 1990), as the first regulation it is expected a significant improvement in energy performance.

The next two periods ‘2006-2012’ and ‘2013-2018’ will have the contexts ‘*New*’ and ‘*Major Renovation*’ considered because buildings retrofitted and new buildings were built in the same period, as the goal is to characterise the buildings from each period. These periods are both represented in *Figure 30 (Left and Right)*. Their probability distribution functions are in *Figure 31*. ‘2006-2012’ marks the update of the RCCTE of 1990, in 2006. Finally, the REH was introduced in 2013, then, the last period studied is ‘2013-2018’.

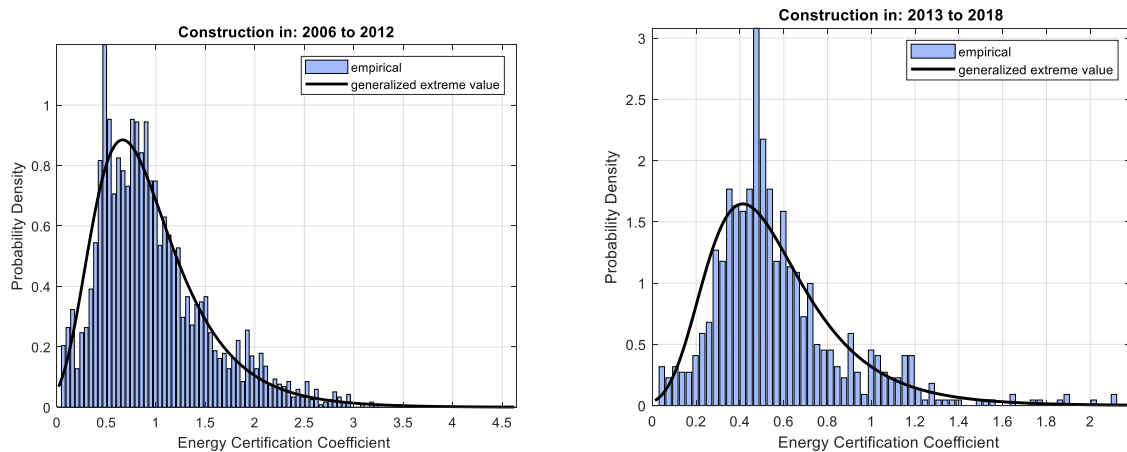


Figure 30 - *RN<sub>t</sub>* for Buildings Constructed in '2006-2012' Probability Distribution Function and Histogram (Left). *RN<sub>t</sub>* for Buildings Constructed in '2013-2018' Probability Distribution Function and Histogram (Right).

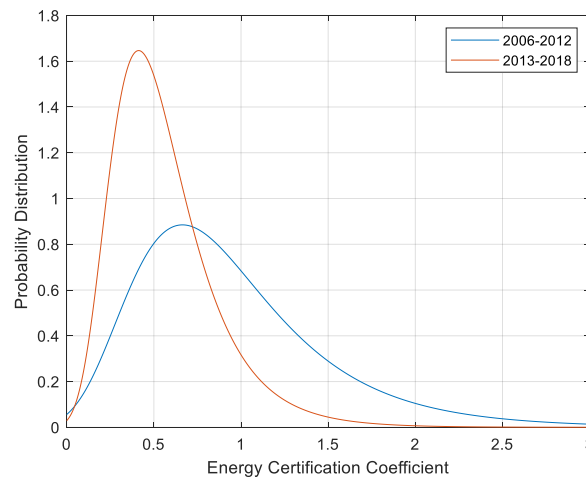
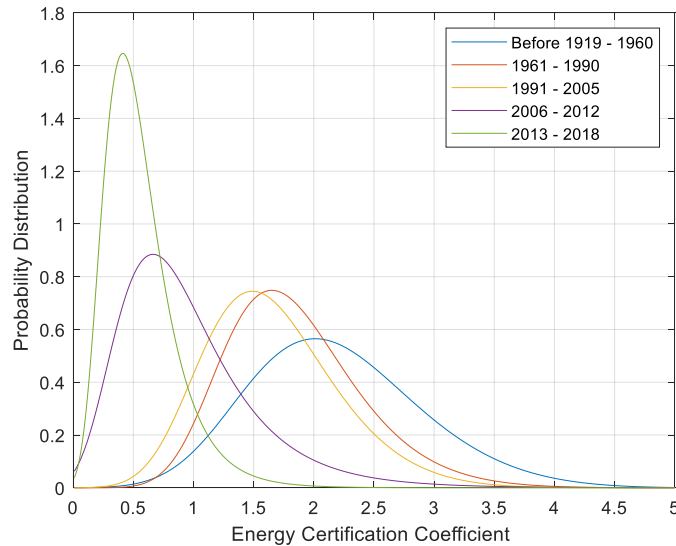


Figure 31 - *RN<sub>t</sub>* Comparison for Buildings Constructed in Periods '2006-12' and '2013-2018'.

Both ‘2006-2012’ and ‘2013-2018’ have also a *generalized extreme value*, with medians of 0.83 and 0.49, respectively.

Studying these periods, we hope to recognize the improvements, especially after the introduction of new regulations. Using each period probability distribution parameters, the distribution functions could be all reproduced in the same figure. With *Figure 32* we can observe how the parameter  $RNt$  has been evolving over the years.



*Figure 32 -  $RNt$  'Before 1919-1960', '1961-1990', '1991-2005', 2006-2012' and '2013-2018' Comparison of Probability Distribution Functions.*

As expected the RBS energy performance is improving over the years. We start with a flatter curve in 'Before 1919-1960' with a median of 2.13 to a curve in '2013-2018' with a huge incidence peak around 0.4, its median is 0.49. The probability distribution function parameters of the functions present in Figure 32, are summarised in *Table 7*, alongside with the respective medians.

*Table 7 - Distribution Function and Distribution Parameters and Medians for Each Period in Figure 32.*

| <b>Period:</b>            | <b>Distribution Function</b>     | <b>Dist. Parameter Names</b>  | <b>Dist. Parameter Values</b> | <b>Median</b> |
|---------------------------|----------------------------------|-------------------------------|-------------------------------|---------------|
| Existing Before 1919-1990 | <i>generalized extreme value</i> | <i>shape; scale; location</i> | -0.1578;0.6598;1.899;         | 2.13          |
| Existing 1961-1990        |                                  |                               | -0.0971;0.4940;1.6009;        | 1.78          |
| Existing 1991-2005        |                                  |                               | -0.1249;0.4979;1.4244;        | 1.60          |
| 2006-2012                 |                                  |                               | 0.0500;0.4163;0.6844;         | 0.83          |
| 2013-2018                 |                                  |                               | 0.03858;0.2236;0.4212;        | 0.49          |

### 4.3 How does the building energy performance ( $RNt$ ) will evolve in the recent future?

Pre-certificates are given to projects of dwelling that are going to be built. So, when we analyse the pre-certificates, we get a draft of what the RBS is evolving towards in the near future. Pre-certificates in *generaldata.csv* have no construction year (it is 0) because these dwellings have not been built yet, so to analyse them we start *Generalmix.m* and choose value 0 for the variable construction year and 2 in the variable context, then we will have a sample of buildings in context ‘New’ with any construction year defined, this sample corresponds to pre-certificates. The result pre-certificate analysis with parameter  $RNt$  can be observed in the *Figure 33*.

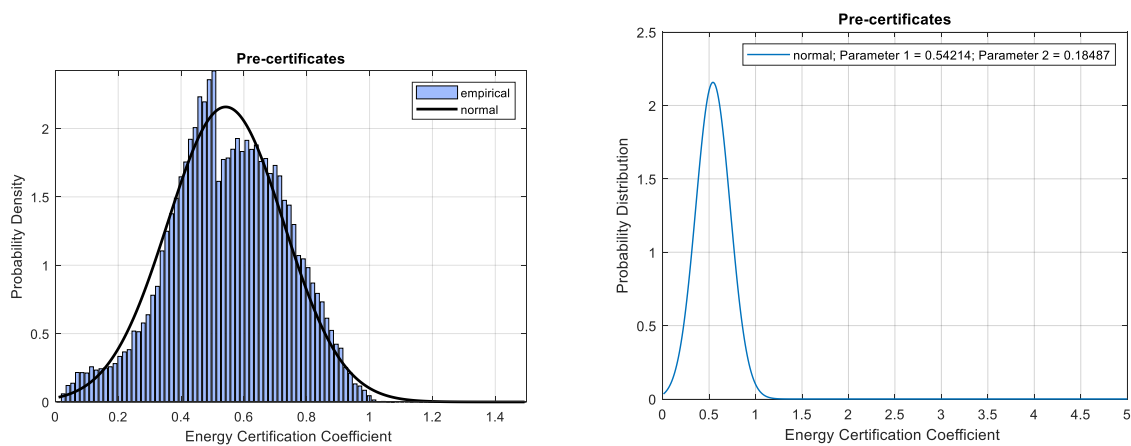


Figure 33 -  $RNt$  for Buildings with Pre-certificate Probability Distribution Function and Histogram (Left).  $RNt$  Pre-certificate Sample Probability Distribution Function (Right).

The probability distribution of  $RNt$  in pre-certificates is a *normal* distribution, the distribution parameters are *location* 0.542 and *scale* 0.185, the median is 0.541. For ‘New’ dwellings from the period ‘2013-2018’, the distribution is a *rician*, the parameters *non centrality* and *shape* are 0.398 and 0.206, respectively. We can compare both distribution functions in *Figure 34*.

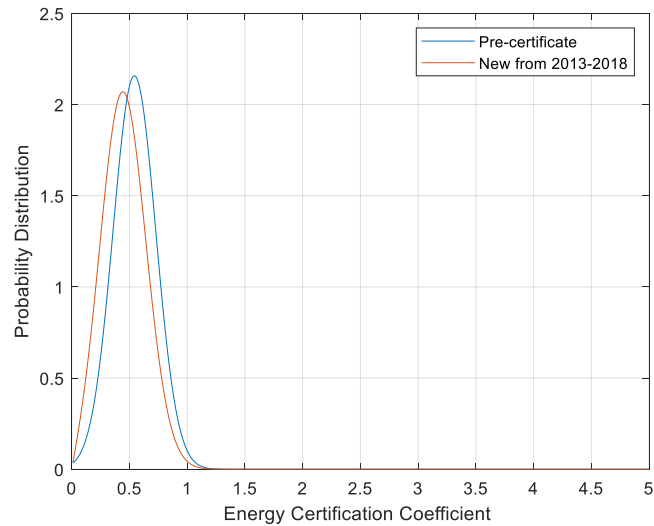


Figure 34 -  $RN_t$  Pre-certificate Sample and 'New from 2013-2018' Probability Distribution Functions Comparison.

According to the data output, dwellings to be built will have a slightly higher  $RN_t$  than those that were built in '2013-2018', this means that the energy performance is not increasing. Yet, this  $RN_t$  value compared to the rest of the periods studied still represents a huge improvement in the RBS energy performance. This might be due to when experts are preparing the pre-certificate, they do not know exactly what or which equipments will be used, therefore a default value is used, that will may not correspond to reality.

### 4.4 What is the building energy performance ( $R_{Nt}$ ) in different contexts?

We have observed that  $R_{Nt}$  differs from period to period, so it is also expected to differ from context to context. Context ‘Existing’ contains dwellings that existed before the introduction of norms that regulate energy efficiency in buildings, therefore its energy performance should be lower than dwellings that had energy performance regulations. This affirmation is supported by the next figures, *Figure 35, 36 and 37*.

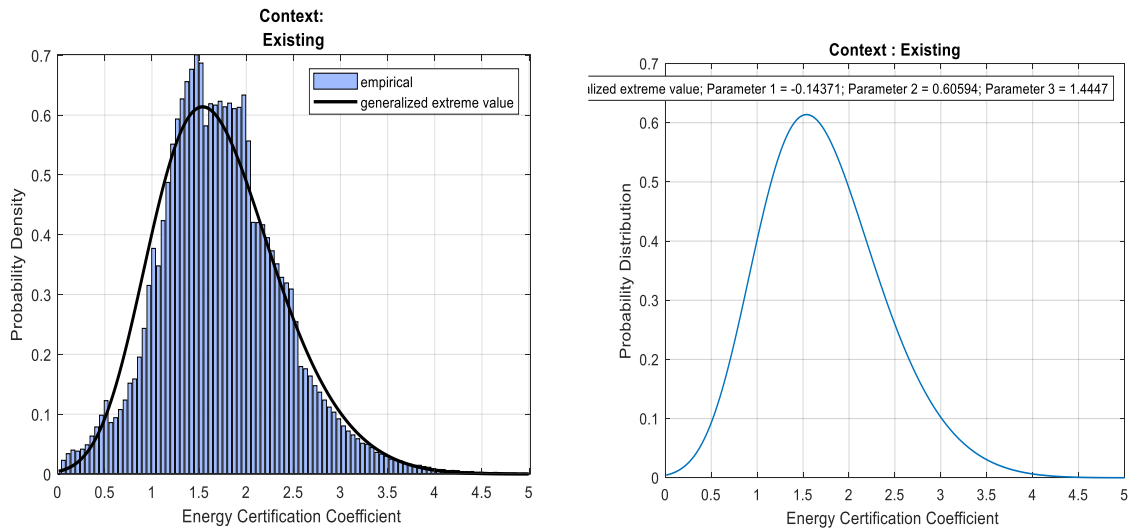


Figure 35 -  $R_{Nt}$  for Context ‘Existing’ Probability Distribution Function and Histogram (Left).  $R_{Nt}$  for Context ‘Existing’ Probability Distribution Function (Right)

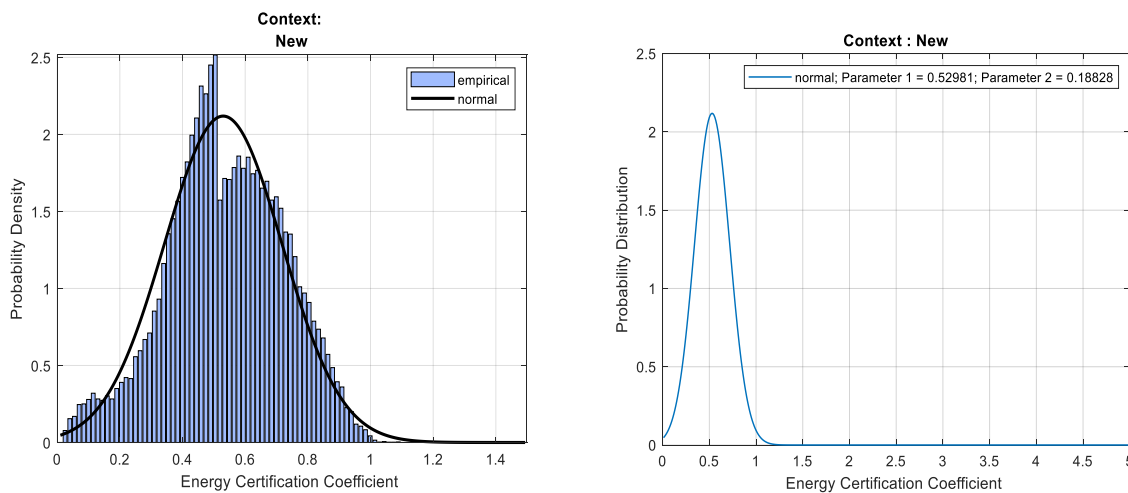


Figure 36 -  $R_{Nt}$  for Context ‘New’ Probability Distribution Function and Histogram (Left).  $R_{Nt}$  for Context ‘New’ Probability Distribution Function (Right)

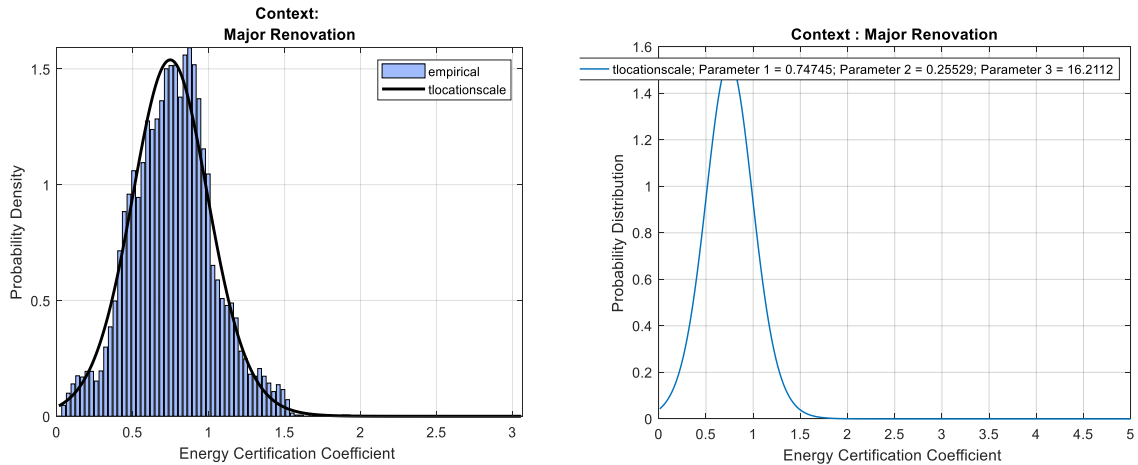


Figure 37 - Rnt for Context 'Major Renovation' Probability Distribution Function and Histogram (Left). Rnt for Context 'Major Renovation' Probability Distribution Function (Right).

The results of this context analysis are for 'Existing' a *generalized extreme value* with a median of 1.67. 'New' has for distribution function a *normal* distribution with median of 0.524. Finally, for 'Major Renovation' is represented a *tlocationscale* with 0.750 of median. Using the distribution parameters of these probability distribution functions, which can be found in Table 8, it was possible to replicate the probability distribution functions in the same figure, in Figure 38. In the table below we also find the medians.

Table 8 - Distribution Function and Distribution Parameters and Medians for Each Context in Figure 38.

| Context:         | Distribution Function            | Dist. Parameter Names                       | Dist. Parameter Values  | Median |
|------------------|----------------------------------|---------------------------------------------|-------------------------|--------|
| Existing         | <i>generalized extreme value</i> | <i>shape; scale; location;</i>              | -0.14371;0.60594;1.4447 | 1.67   |
| New              | <i>normal</i>                    | <i>location; scale;</i>                     | 0.52981;0.18828;        | 0.52   |
| Major Renovation | <i>tlocationscale</i>            | <i>location; scale; degrees of freedom;</i> | 0.74745;0.25529;16.2112 | 0.75   |

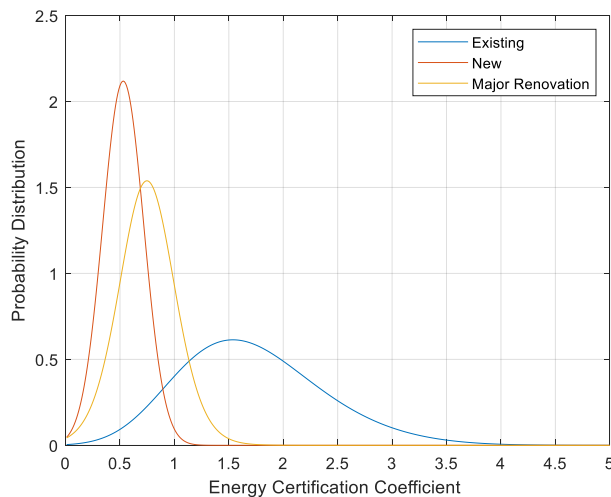


Figure 38 - Rnt 'Existing', 'New' and 'Major Renovation' Contexts Comparison of Probability Distribution Functions.

'Existing' has the higher *RNT* while 'New' has the lowest due to the introduction of regulation to new buildings, 'Major Renovation' is much closer to 'New' than it is to 'Existing'.

#### 4.5 How does Net Floor Area changes regionally, or in context?

Net floor area refers to the sum of the areas, measured by the inner perimeter of all useful surfaces that are under REH. Net floor area across Portuguese RBS follows the distributions observed in *Figure 39*.

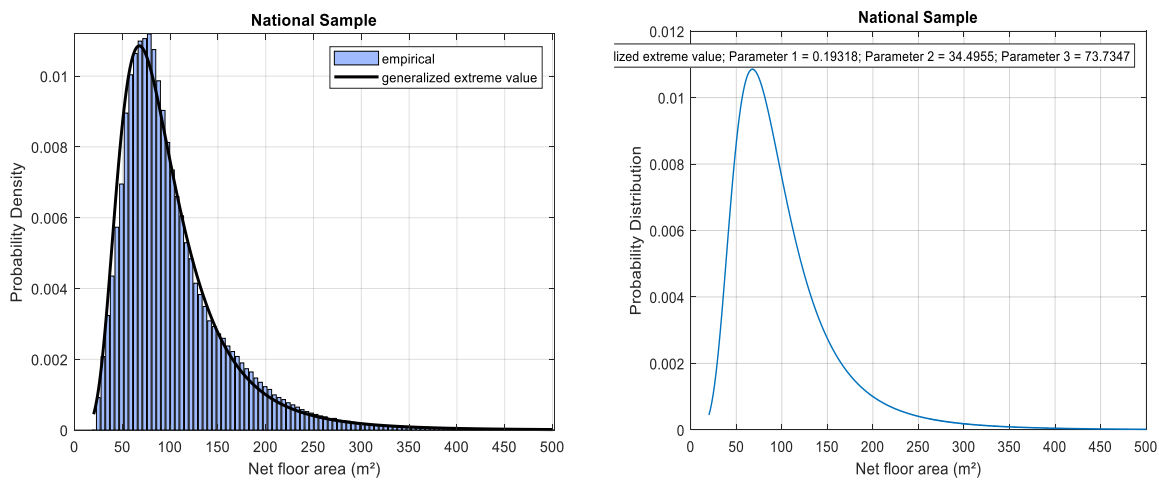


Figure 39 - Net Floor Area ( $m^2$ ) in Portugal Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Portugal Probability Distribution Function (Right).

The sample was limited between 20 and 500  $m^2$  to clear absurd values, some close to 0 and some enormous. The resulting probability distribution function is a *generalized extreme value* with a median of 86.0  $m^2$  ( $Q1 = 63.0 m^2$ ;  $Q3 = 122.6 m^2$ ;  $L1 = 0 m^2$ ;  $L2 = 212.0 m^2$ ).

Using this *Net Floor Area* to compare within regions, it is obvious that dwellings in urban areas have smaller areas than rural ones, something to be shown further. *Figures 40* and *41* show the Net floor area of the regions 'Grande Lisboa' and 'Grande Porto', which are the most urbanized regions with 495.6  $\frac{dwellings}{km^2}$  for the first and 408.9  $\frac{dwellings}{km^2}$  for the second. Both probability distribution functions are *generalized extreme value* and look similar with a difference as 'Grande Porto' dwellings have a bit more net floor area, with 'Grande Lisboa' having a higher incidence in smaller net floor areas.

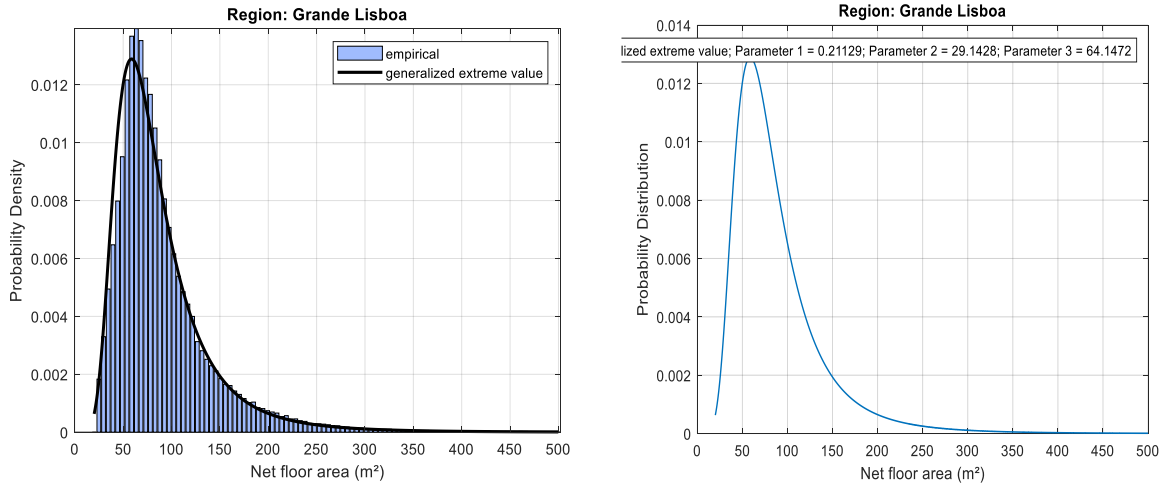


Figure 40 - Net Floor Area ( $m^2$ ) in 'Grande Lisboa' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Grande Lisboa' Probability Distribution Function (Right).

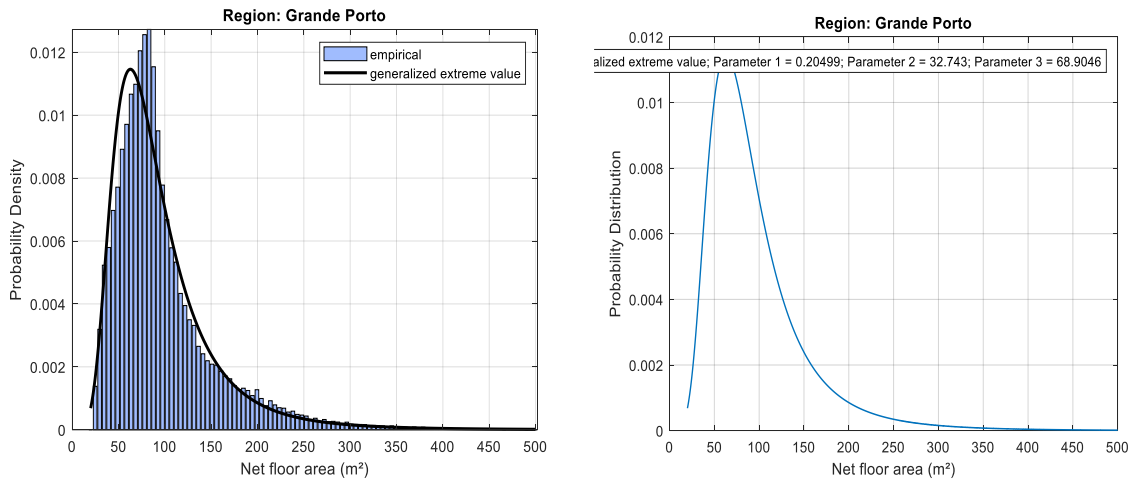


Figure 41 - Net Floor Area ( $m^2$ ) in 'Grande Porto' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Grande Porto' Probability Distribution Function (Right).

'Grande Lisboa' has a median net floor area of  $74.75 m^2$  where in 'Grande Porto' is  $81.16 m^2$ . These values do not differ that much, like was expected. However, one result that is not obvious regards the net floor area in the region 'Algarve'.

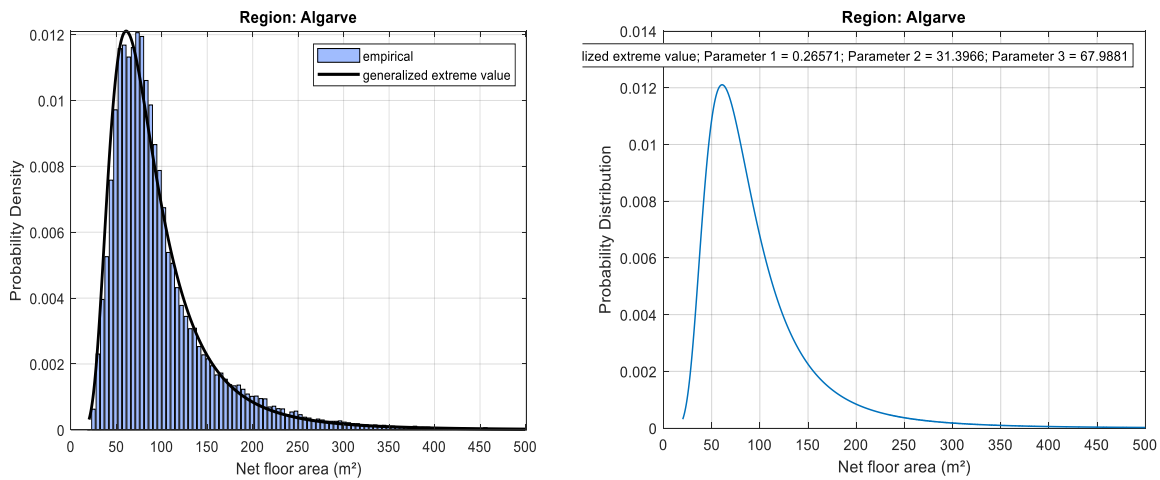


Figure 42 - Net Floor Area ( $m^2$ ) in 'Algarve' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in 'Algarve' Probability Distribution Function (Right).

Its function is also a *generalized extreme value*, 'Algarve' probability distribution function describes a shape almost equal to 'Grande Lisboa'. This outcome was not expected, it is known that 'Algarve' has some level of urbanization, with  $76.5 \frac{\text{dwellings}}{\text{km}^2}$  compared to  $495.6 \frac{\text{dwellings}}{\text{km}^2}$  in 'Grande Lisboa', especially due to tourism in Summer, but here we can observe, as in Figure 43, that its net floor area is smaller than 'Grande Porto'. The median in 'Algarve' is  $79.8 m^2$  compared to  $74.75 m^2$  in 'Grande Lisboa' and  $81.16 m^2$  in 'Grande Porto'.

As mentioned before, in rural areas net floor area tends to be higher than in urban areas like 'Grande Lisboa', this is supported through two comparisons with examples of more rural regions. In Figure 43 these regions are 'Baixo Alentejo' and 'Alto Trás-os-Montes'. These rural regions are compared with the other regions from Figures 40,41 and 42.

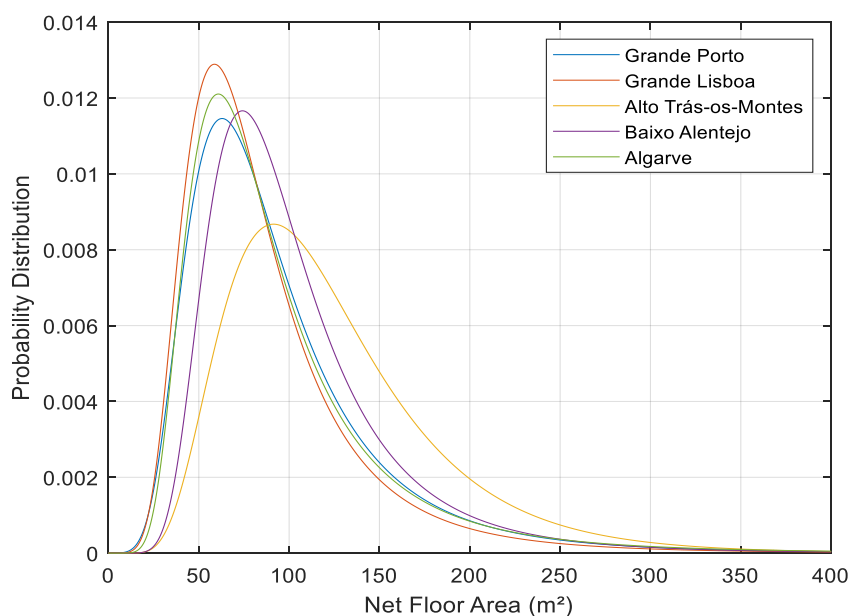


Figure 43 - Net Floor Area ( $m^2$ ) Comparison Between Regions Aforementioned.

As expected, both rural regions had more net floor area than ‘Grande Lisboa’. ‘Baixo Alentejo’ with a median of  $89.92 m^2$  and ‘Alto Trás-os-Montes’ with a more expressive result of  $112.52 m^2$ . The distribution functions and their respective parameters are presented in Table 9. The medians and quartiles are summed in a table, Table 10.

Table 9 - Distribution Function and Distribution Parameters for Each Region Net Floor Area in Figure 43.

| Period:             | Distribution Function            | Dist. Parameter Names          | Dist. Parameter Values |
|---------------------|----------------------------------|--------------------------------|------------------------|
| Alto Trás-os-Montes | <i>lognormal</i>                 | <i>loglocation; logscale</i>   | 4.721; 0.454;          |
| Grande Porto        | <i>generalized extreme value</i> | <i>shape; scale; location;</i> | 0.205; 32.743; 68.905; |
| Grande Lisboa       |                                  |                                | 0.211; 29.143; 64.147; |
| Baixo Alentejo      |                                  |                                | 0.177; 32.018; 79.343; |
| Algarve             |                                  |                                | 0.265; 31.397; 67.988; |

Table 10 - Median and Quartiles of Regions in Figure 43 by Net Floor Area in  $m^2$ .

| Context:            | Median ( $m^2$ ) | Q1 and Q3 ( $m^2$ ) |
|---------------------|------------------|---------------------|
| Alto Trás-os-Montes | 112.52           | 83.24 and 152.23    |
| Grande Porto        | 81.16            | 59.85 and 112.46    |
| Grande Lisboa       | 74.75            | 55.84 and 104.43    |
| Baixo Alentejo      | 89.92            | 69.40 and 124.25    |
| Algarve             | 79.80            | 58.28 and 113.20    |

Regarding context this time for the same parameter, *Net floor area*, it changes significantly in the three possible contexts. For buildings of all contexts we have the distributions in *Figure 44*, *Figure 45* and *46*.

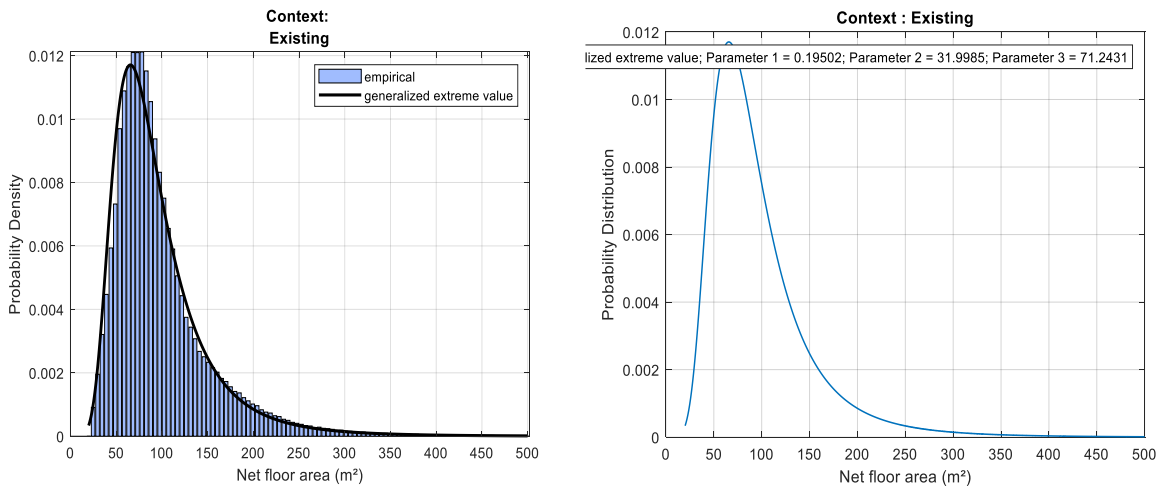


Figure 44 - Net Floor Area ( $m^2$ ) in Context 'Existing' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Context 'Existing' Probability Distribution Function (Right).

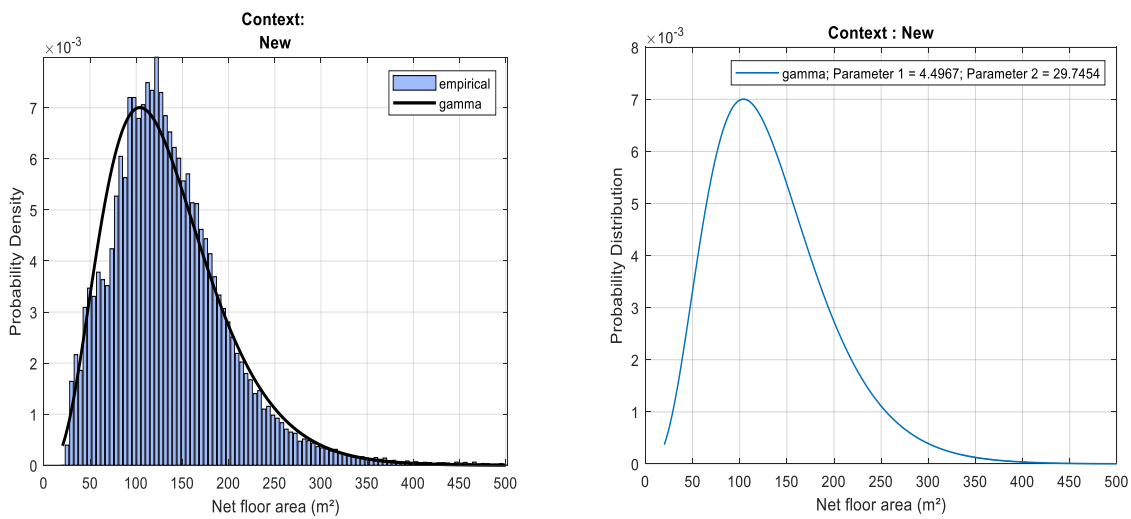


Figure 45 - Net Floor Area ( $m^2$ ) in Context 'New' Probability Distribution Function and Histogram (Left). Net Floor Area ( $m^2$ ) in Context 'New' Probability Distribution Function (Right).

If these functions are put in the same figure, in *Figure 47*, it is evident that ‘New’ dwellings have larger net floor areas than ‘Existing’, this huge difference can be confirmed in the table resultant from the *boxplot* diagram, *Table 11*. ‘Existing’ probability distribution function is a *generalized extreme value* contrasting with a *gamma* for ‘New’. In *Figure 47* we notice that ‘New’ contains a wider range of values some even further  $350 m^2$  with a median of  $124.7 m^2$  while in ‘Existing’ values are more concentrated and have a median of  $82.75 m^2$ . This analysis is pursued and ‘Major Renovation’ is analysed, starting with histogram and distribution functions in *Figure 46*. Its distribution function is ‘*inverse gaussian*’

and its peak is around 50 m<sup>2</sup>. We now compare Net floor area probability distribution function in 'Major Renovation' with the other contexts, in Figure 47.

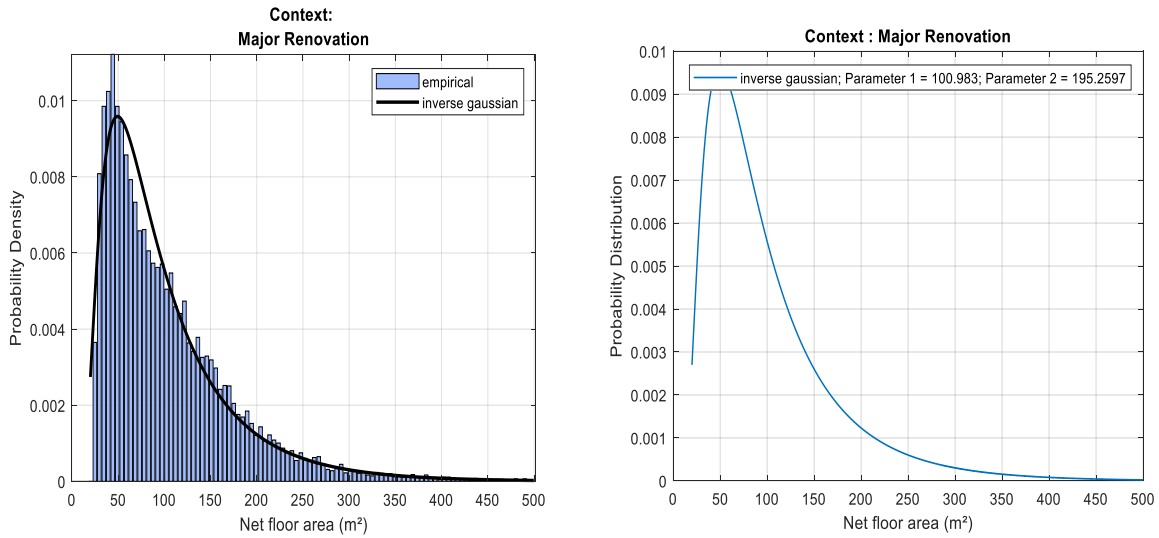


Figure 46 - Net Floor Area (m<sup>2</sup>) in Context 'Major Renovation' Probability Distribution Function and Histogram (Left). Net Floor Area (m<sup>2</sup>) in Context 'Major Renovation' Probability Distribution Function (Right).

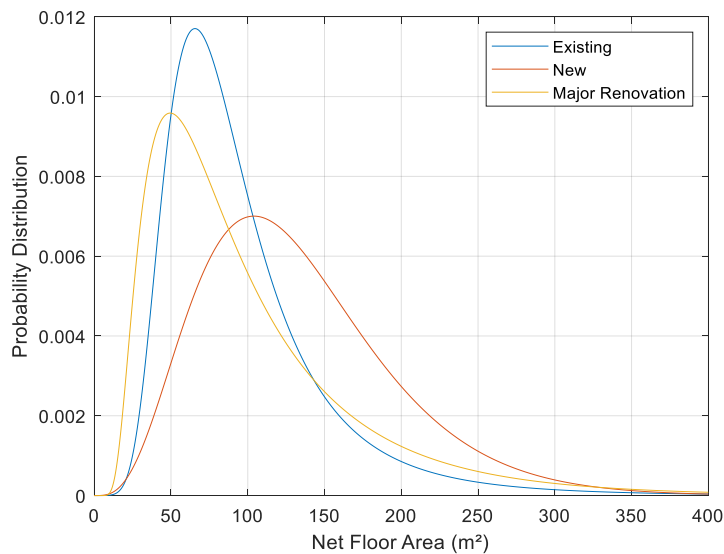


Figure 47 - Net Floor Area (m<sup>2</sup>) of Contexts 'Major Renovation', 'New' and 'Existing' Probability Distribution Functions.

The following table summarizes the information of the distribution functions.

Table 11 - Distribution Function and Distribution Parameters for Each Context Net Floor Area in Figure 47.

| Context:         | Distribution Function            | Dist. Parameter Names          | Dist. Parameter Values |
|------------------|----------------------------------|--------------------------------|------------------------|
| Existing         | <i>generalized extreme value</i> | <i>shape; scale; location;</i> | 0.195; 31.999; 71.243; |
| New              | <i>gamma</i>                     | <i>shape; scale;</i>           | 4.497; 29.745;         |
| Major Renovation | <i>inverse gaussian</i>          | <i>scale; shape;</i>           | 100.983; 195.260;      |

Besides having its maximum before 'New', the probability distribution function of 'Major Renovation' has a maximum before 'Existing', but also at higher net floor areas there is a greater occurrence in 'Major Renovation' than in 'Existing'. This could mean that smaller apartments are getting retrofitted or that bigger dwellings are getting divided into smaller ones due to real state pressure. Medians and quartiles (Q1 for first quartile and Q3 for third quartile) of each context, extracted from *boxplot* diagrams in *Table 12*.

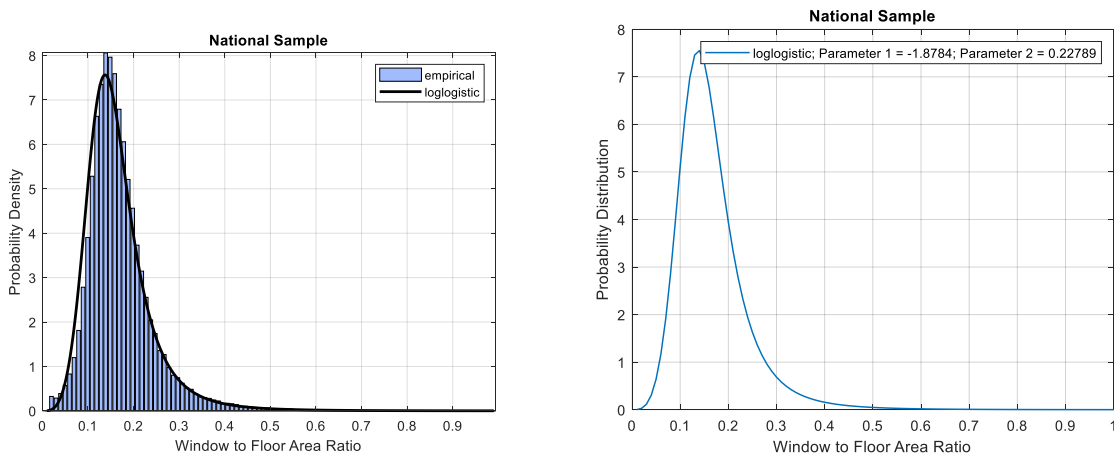
*Table 12 - Median and Quartiles for all Contexts by Net Floor Area.*

| <b>Context:</b>  | <b>Median (<math>m^2</math>)</b> | <b>Q1 and Q3 (<math>m^2</math>)</b> |
|------------------|----------------------------------|-------------------------------------|
| Existing         | 82.75                            | 61.8 and 114.60                     |
| New              | 124.7                            | 90.4 and 166.77                     |
| Major Renovation | 81.47                            | 48.34 and 132.82                    |

The first quartile, median and third quartiles are all higher in 'New' than in 'Major Renovation', one is  $124.7 m^2$  the other is  $81.47 m^2$ . shows that the first quartile and median are inferior in 'Major Renovation' than 'Existing', with the median of 'Existing' already mentioned of  $82.75 m^2$ . Meanwhile, the third quarter of 'Major Renovation' is significantly higher at  $132.82 m^2$  compared to  $114.6 m^2$  of 'Existing'.

## 4.6 How the window-to-floor area ratio changes over the contexts?

In this section, the parameter studied is the window-to-floor area ratio in the different contexts, a parameter which is defined by the division of window area with net floor area. Higher window-to-floor area ratios mean a higher natural lighting, solar heat gains and natural ventilation in the dwelling - solar heat gains are heavily impacted by window orientation, glazing and shading. Higher window-to-floor area ratio also means a bigger heat loss due to transmission, as glass has a higher heat transfer coefficient (lower thermal resistance) than walls, thus, an additional challenge to keep indoor thermal comfort. This parameter is usually between 0 and 1, though values greater than 1 are not impossible, yet rather uncommon (for example, a dome or the crystal palace in Madrid). The window-to-floor area ratio probability distribution and histogram of the Portuguese RBS are disclosed in *Figure 48 (Left)*. The probability distribution function alone for the same sample with its distribution characteristics is in *Figure 48 (Right)*.



*Figure 48 – Window-to-Floor Area Ratio in Portugal Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Portugal Probability Distribution Function (Right).*

The functions displayed above is a *loglogistic* with those parameters in legend of *Figure 48 (Right)*. The median is 0.153, the first quartile 0.121 and third quartile 0.194.

Now that the parameter is characterised and framed, it will be characterised according to each one of the contexts. The context '*Existing*' contains most of the RBS and has a wider range of periods of construction represents the group where this parameter is lowest. As we can observe in the following figures, *Figure 49* is where it is represented in distribution and histogram.

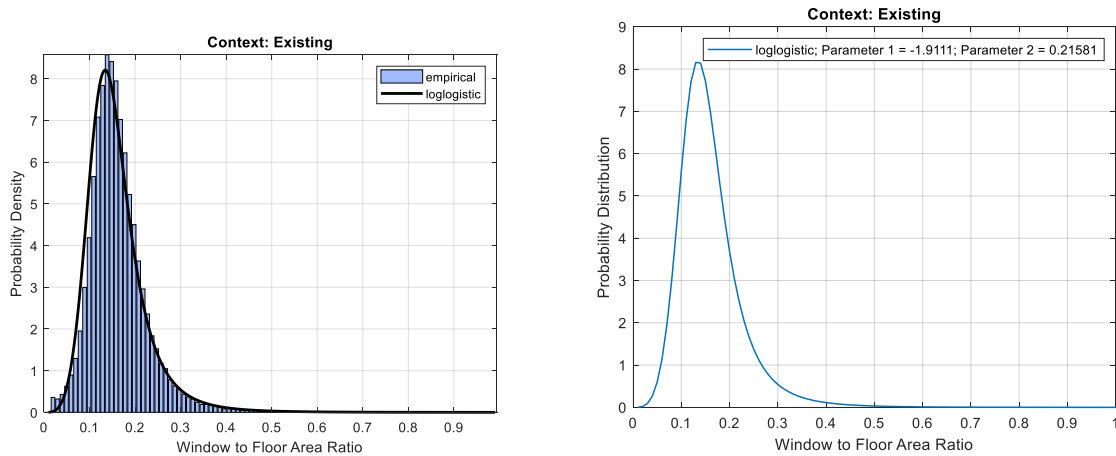


Figure 49 - Window-to-Floor Area Ratio in Context 'Existing' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'Existing' Probability Distribution (Right).

Similarly, the same functions for 'New' and 'Major Renovation' are in Figure 50 and 51.

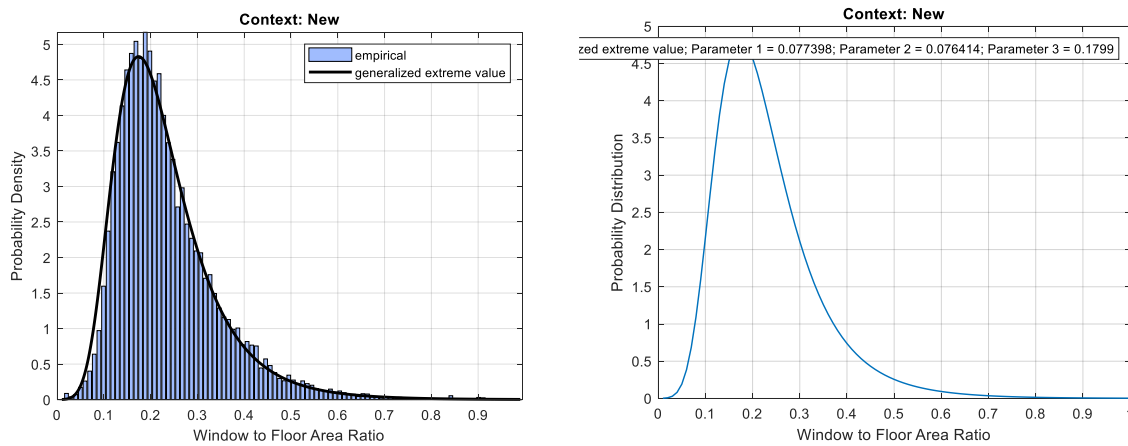


Figure 50 - Window-to-Floor Area Ratio in Context 'New' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'New' Probability Distribution (Right)

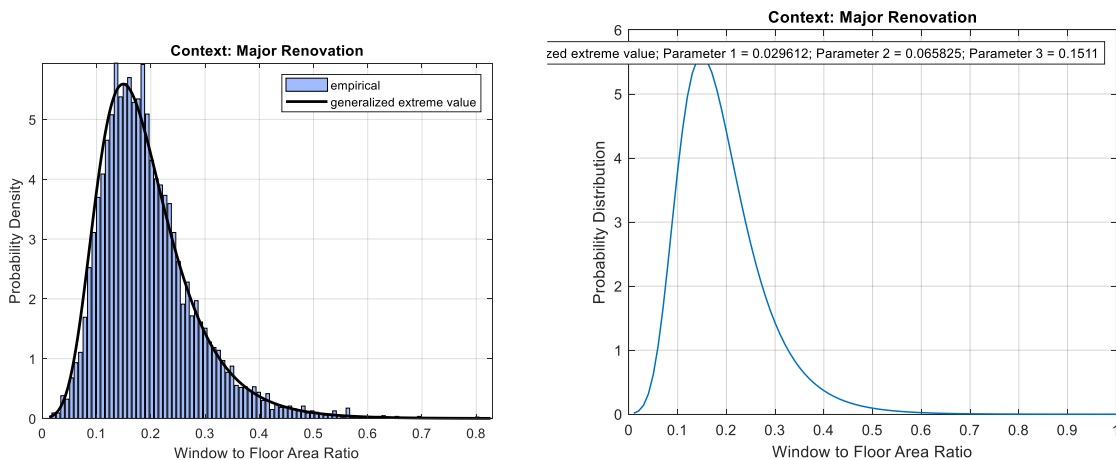
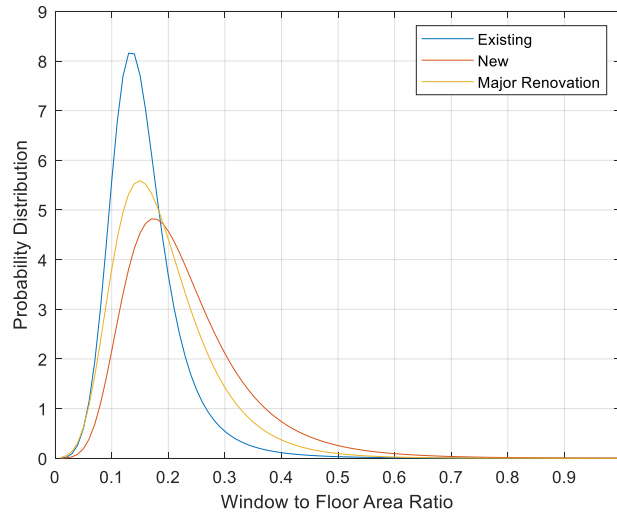


Figure 51 - Window-to-Floor Area Ratio in Context 'Major Renovation' Probability Distribution and Histogram (Left). Window-to-Floor Area Ratio in Context 'Major Renovation' Probability Distribution (Right)

Here we can observe that both ‘*Major Renovation*’ and ‘*New*’ have larger window to floor area ratio than ‘*Existing*’. For ‘*Existing*’ the distribution function is a *loglogistic* while for the other contexts it is a *generalized extreme value*. The distribution parameters in *Table 13* generated the distributions in *Figure 52*.



*Figure 52 - Window to Floor Area Ratio for Contexts 'Existing', 'New' and 'Major Renovation' Probability Distribution Functions Comparison.*

When the three contexts are compared, we see that, as was said, ‘*Existing*’ has the lower values with a median of 0.149, followed by ‘*Major Renovation*’ with a median of 0.176 and close is ‘*New*’ with 0.206. Which tells us that the Portuguese RBS is moving towards larger window-to-floor area ratios. In *Table 13* are the distribution function names and parameters, with the medians, for each context.

*Table 13 - Distribution Function, Distribution Parameters and Median for Each Context Window-to-Floor Area in Figure 52.*

| <b>Context:</b>  | <b>Distribution Function</b>     | <b>Dist. Parameter Names</b>   | <b>Dist. Parameter Values</b> | <b>Median</b> |
|------------------|----------------------------------|--------------------------------|-------------------------------|---------------|
| Existing         | <i>loglogistic</i>               | <i>loglocation; logscale;</i>  | 1.91;0.216;                   | 0.149         |
| New              | <i>generalized extreme value</i> | <i>shape; scale; location;</i> | 0.0774;0.0764; 0.180;         | 0.206         |
| Major Renovation |                                  |                                | 0.0296;0.0658;0.151;          | 0.176         |

## 4.7 How does the window U-value change over the years and contexts?

The window U-value ( $U_{wnd}$ ) is the window heat transfer coefficient  $\frac{W}{K.m^2}$ . The lower  $U_{wnd}$  is, the greater the thermal insulation is. With a lower  $U_{wnd}$  we expect to save energy due to thermal insulation, especially for dwelling with great window to floor area ratios. This parameter's RBS distribution is evidenced in *Figure 53*.

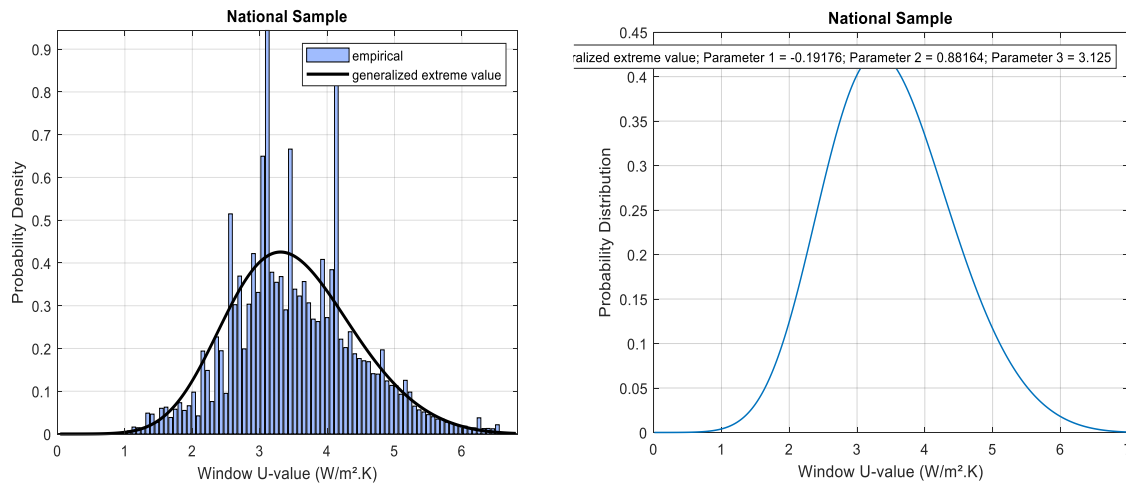


Figure 53 - Window U-value ( $\frac{W}{K.m^2}$ ) in Portugal Probability Distribution and Histogram (Left). Window U-value ( $\frac{W}{K.m^2}$ ) in Portugal Probability Distribution Function (Right).

$U_{wnd}$  at a national level is characterised by a *generalized extreme value* with a first quartile of  $2.9 \frac{W}{K.m^2}$  median as  $3.4 \frac{W}{K.m^2}$  and third quartile of  $4.1 \frac{W}{K.m^2}$ . Now with this snapshot of  $U_{wnd}$  contexts are characterised (*Figure 54,55 and 56*).

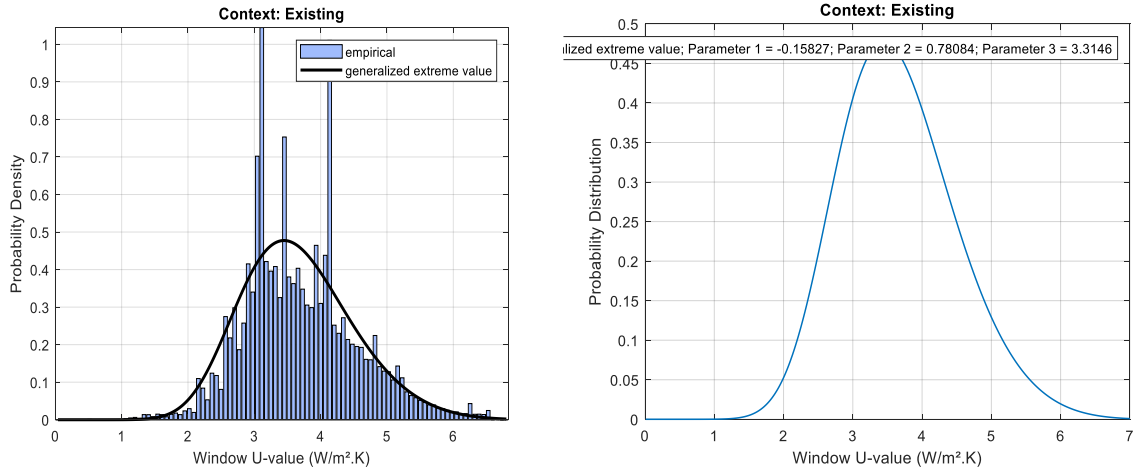


Figure 54 - Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'Existing' Probability Distribution Function and Histogram (Left). Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'Existing' Probability Distribution Function (Right).

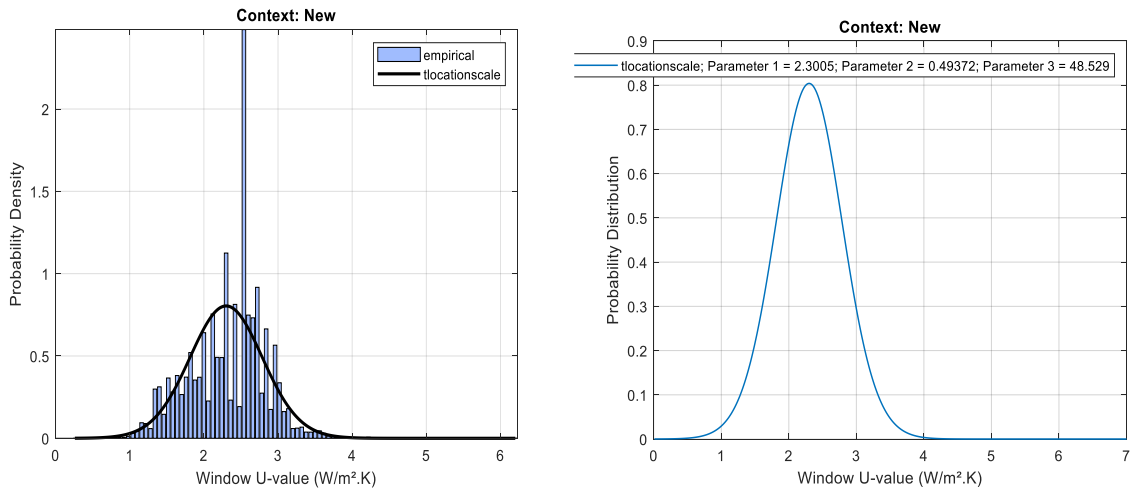


Figure 55 - Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'New' Probability Distribution Function and Histogram (Left). Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'New' Probability Distribution Function (Right).

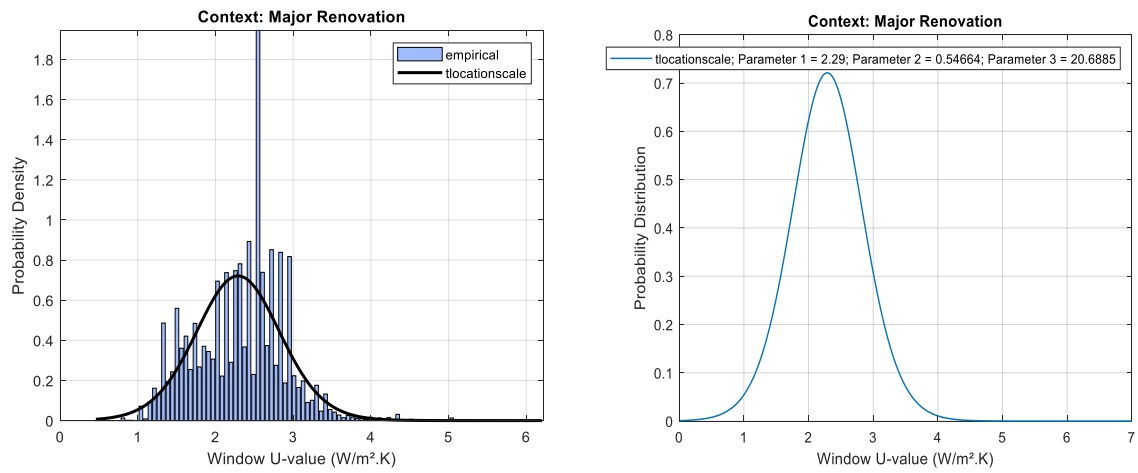


Figure 56 - Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'Major Renovation' Probability Distribution Function and Histogram (Left).  
 Window U-value ( $\frac{W}{K.m^2}$ ) in Context 'Major Renovation' Probability Distribution Function (Right).

Again, the parameter differs between contexts, in 'Existing' is a *generalized extreme value*, while for 'New' and 'Major Renovation' is a *tlocationscale*. In Table 14 are the probability distribution function parameters which will generate the probability distribution functions in Figure 57.

Table 14 - Distribution Function, Distribution Parameters and Median for Each Context Window U-value in Figure 57.

| Context:         | Distribution Function            | Dist. Parameter Names                       | Dist. Parameter Values |
|------------------|----------------------------------|---------------------------------------------|------------------------|
| Existing         | <i>generalized extreme value</i> | <i>shape; scale; location;</i>              | -0.158; 0.781; 3.314;  |
| New              | <i>tlocationscale</i>            | <i>location; scale; degrees of freedom;</i> | 2.30; 0.494; 48.529;   |
| Major Renovation |                                  |                                             | 2.29; 0.547; 20.689;   |

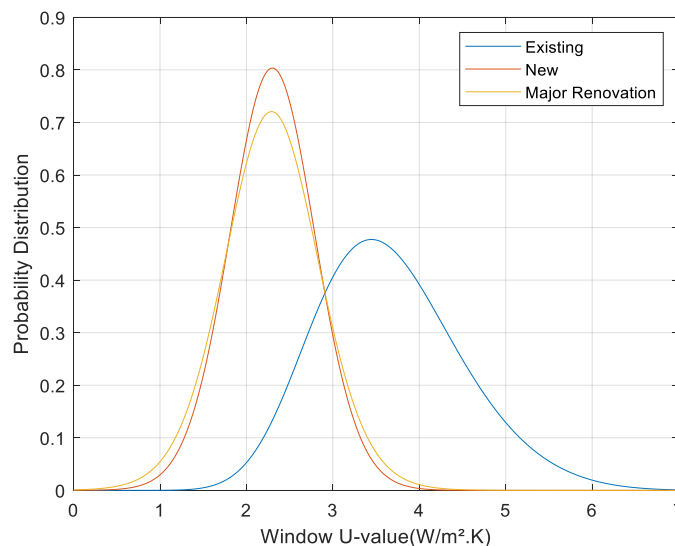


Figure 57 - Window U-value ( $\frac{W}{K.m^2}$ ) for Contexts 'Existing', 'New' and 'Major Renovation' Probability Distribution Functions.

Looking at *Figure 57*, while the ‘Existing’ RBS is worse insulated with a median of  $3.55 \frac{W}{K.m^2}$ , the contexts ‘Major Renovation’ and ‘New’ are identical: the median of ‘New’ is  $2.40 \frac{W}{K.m^2}$  and ‘Major Renovation’ is  $2.37 \frac{W}{K.m^2}$ . We conclude that, except for ‘Existing’, the windows in the RBS are becoming more insulated – with the increased importance due to having greater window to floor area ratios -, thus contributing to a better energy performance. To support this affirmation, we will now study the evolution of the probability distribution function of this parameter over the periods when regulations were imposed.

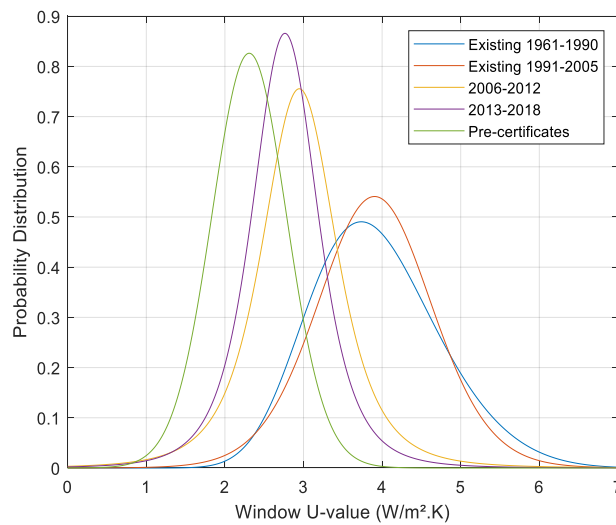


Figure 58 - Window U-value ( $\frac{W}{K.m^2}$ ) Probability Distribution Functions of Periods in Legend.

The distribution function names, and their parameter values for the periods in *Figure 58* are in *Table 15*.

Table 15 - Distribution Function and Distribution Parameters for Each Period in Figure 58.

| Period:            | Distribution Function            | Dist. Parameter Names                       | Dist. Parameter Values |
|--------------------|----------------------------------|---------------------------------------------|------------------------|
| Existing 1961-1990 | <i>generalized extreme value</i> | <i>shape; scale; location</i>               | -0.165; 0.761; 3.600;  |
| Existing 1991-2005 | <i>tlocationscale</i>            | <i>location; scale; degrees of freedom;</i> | 3.904; 0.728; 18.923;  |
| 2006-2012          |                                  |                                             | 2.946; 0.492; 3.491;   |
| 2013-2018          |                                  |                                             | 2.764; 0.433; 3.988;   |
| Pre-certificates   |                                  |                                             | 2.309; 0.480; 42.378;  |

*Figure 58* shows that, although  $U_{wdn}$  increased from ‘1961-1990’ to ‘1991-2005’, which demonstrates a downgrade in thermal insulation, possibly explained by the higher number of major renovations in the first period. After the latter period from ‘2006-2012’ decreased significantly and has been decreasing since then. ‘Pre-certificates’ show that buildings to be constructed will have better thermal window insulation than in ‘2013-2018’, this represents a considerable step in improving energy performance in buildings. By the same order of periods medians are:  $3.87 \frac{W}{K.m^2}$ ,  $3.90 \frac{W}{K.m^2}$ ,  $2.94 \frac{W}{K.m^2}$ ,  $2.76 \frac{W}{K.m^2}$ ,  $2.31 \frac{W}{K.m^2}$ .

## 4.8 Walls U-value by Region and Climatic Zone

The parameter  $U$ -value of walls is the heat transfer coefficient of a wall  $\frac{W}{K.m^2}$ , determines how walls insulate the dwelling from the outside and vice-versa. Like  $U_{wdn}$ , the lower values for  $U$ -value represent a better insulated surface, contributing to a better energy performance. Its national sample for probability distribution function and histogram are in *Figure 59*.

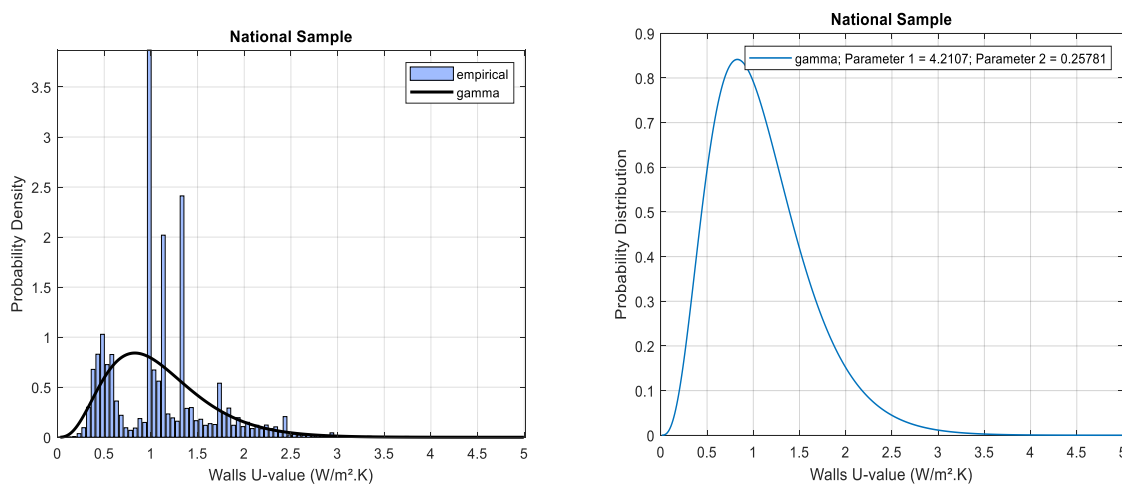


Figure 59 - Walls  $U$ -value ( $\frac{W}{K.m^2}$ ) in Portugal Probability Distribution Function and Histogram (Left). Walls  $U$ -value ( $\frac{W}{K.m^2}$ ) in Portugal Probability Distribution Function (Right).

The RBS Walls  $U$ -value is characterised by a *gamma* distribution with *shape* and *scale* of 4.21 and 0.258. The median is  $1.02 [\frac{W}{K.m^2}]$ ;  $Q1 = 0.64 [\frac{W}{K.m^2}]$ ;  $Q3 = 1.30 [\frac{W}{K.m^2}]$ ;  $L1 = 0 [\frac{W}{K.m^2}]$ ;  $L2 = 2.29 [\frac{W}{K.m^2}]$ .

Climatic zones in Portugal are divided by NUTSIII region, defined by three distinct climatic zones for summer ( $V1$ ,  $V2$ ,  $V3$ ) and three for winter ( $I1$ ,  $I2$ ,  $I3$ ), according to heating degrees-days (HDD) by an order of law, where a table with HDD value for each NUTSIII region (*Table 17 and 18*), then in *Table 19* HDD values are related to climatic zones, in this case for the heating season.

First, the climate zone  $I1$ , where less heating energy is required, is analysed using the example of 'Alentejo Litoral' region (1089 HDD)

First, the walls  $U$ -value of the climate zone  $I1$ , where less heating energy is required, is analysed using the example of 'Alentejo Litoral' region (1089 GD) in *Figure 60*.

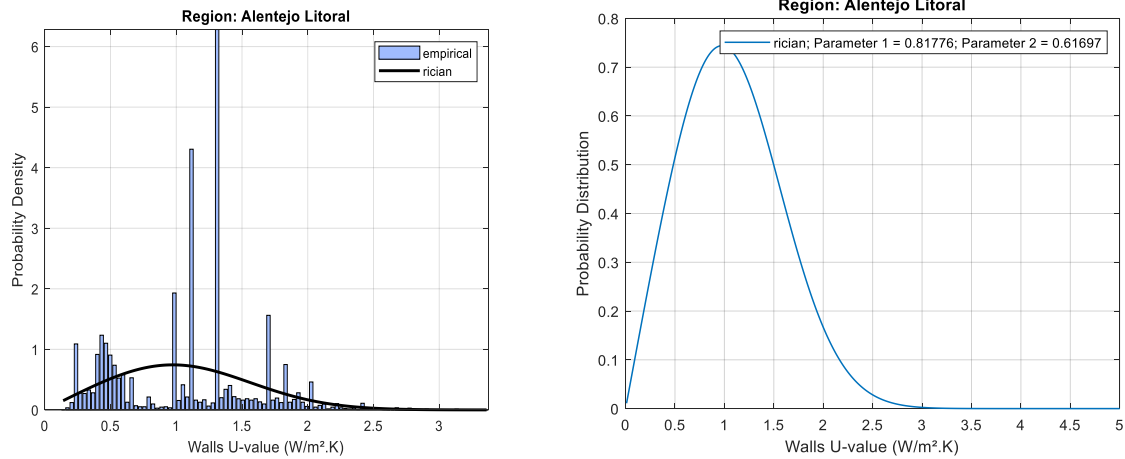


Figure 60 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Alentejo Litoral' Probability Distribution Function and Histogram (Left). Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Alentejo Litoral' Probability Distribution Function (Right).

The same is done for a region belonging to climatic zone *I3*, in this case 'Alto Trás-os-Montes' with 2015 HDD, according to Table 17 and 19. These two regions were chosen because they both represent rural regions with contrasting climatic zones. The figure for histogram and distribution functions is Figure 61.

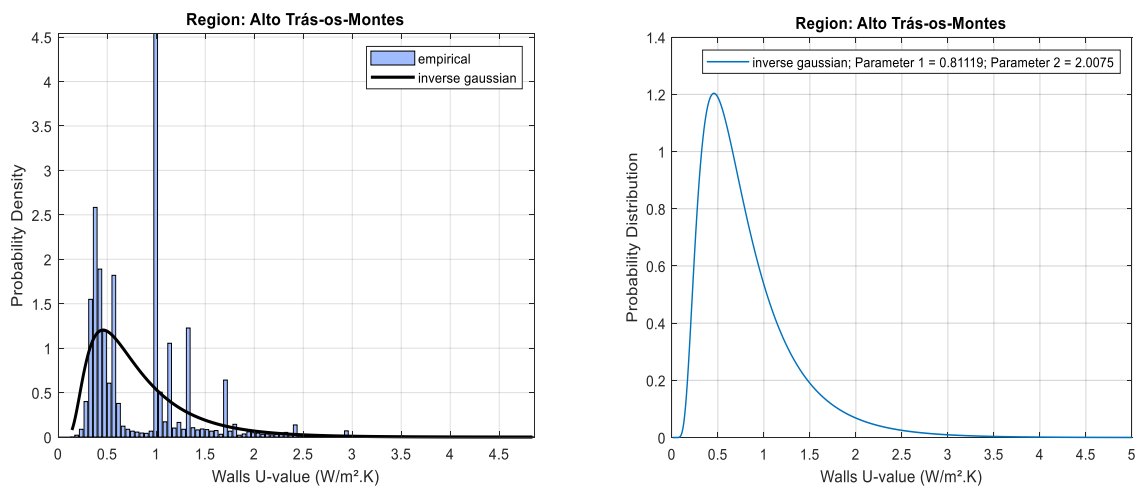


Figure 61 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Alto Trás-os-Montes' Probability Distribution Function and Histogram (Left). Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Alto Trás-os-Montes' Probability Distribution Function (Right).

In Figure 60 we see a *rician* distribution while in Figure 61 it is an *inverse gaussian*. The statistical parameters such as quartiles are found in the boxplot in Figure 62, along with their probability distribution functions comparison.

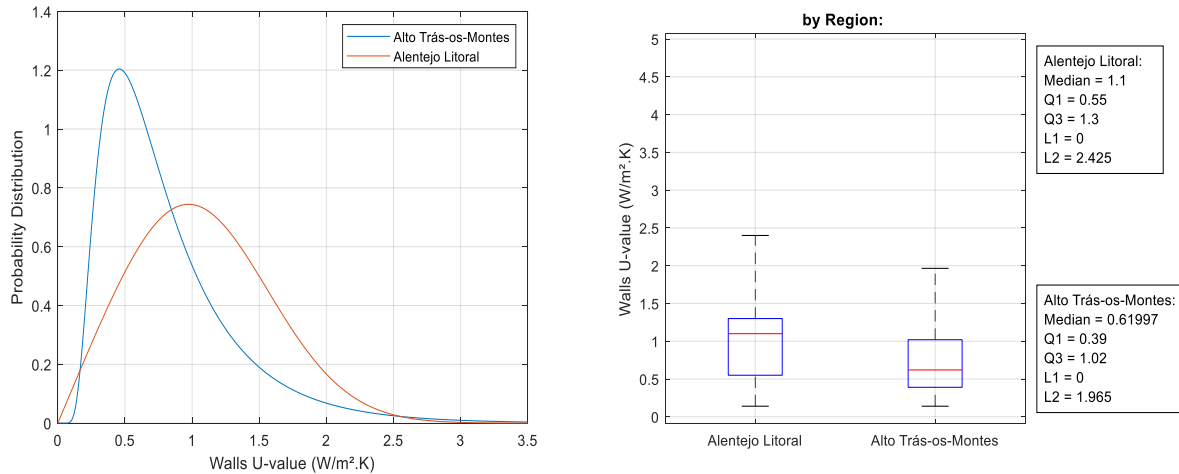


Figure 62 – Comparison Between Probability Distribution Function of Walls U-value ( $\frac{W}{K.m^2}$ ) ‘Alto Trás-os-Montes’ and ‘Alentejo Litoral’ (Left). Boxplot Diagram for Walls U-value ( $\frac{W}{K.m^2}$ ) for the same Regions (Right).

Analysing Figure 62 the region of climatic zone II ‘Alentejo Litoral’ has higher walls U-value (median of  $1.10 \frac{W}{K.m^2}$ ) than in ‘Alto Trás-os-Montes’ of climatic zone I3 (median of  $0.620 \frac{W}{K.m^2}$ ). Therefore, dwellings in the latter region with higher HDD have better wall insulation than ‘Alentejo Litoral’ where winters are milder (lower HDD). The next goal of this sub-chapter was to compare urbanized regions with different climates, the closest possible was comparing ‘Grande Porto’ with ‘Algarve’, however the level of urbanization between these two is incomparable, although ‘Algarve’ is extremely dense along the coast these data refers to the entire region, and because HDD between these regions do not differ much as expected. ‘Grande Porto’ has 1250 HDD and ‘Algarve’ 987 HDD, according to Tables 17 and 18, while ‘Grande Lisboa’ in between with 1071 HDD, as the three regions do not differ that much in HDD (all belong to climatic region I1), the comparison is made between the three, first each one them is represented in Figure 63, 64 and 65.

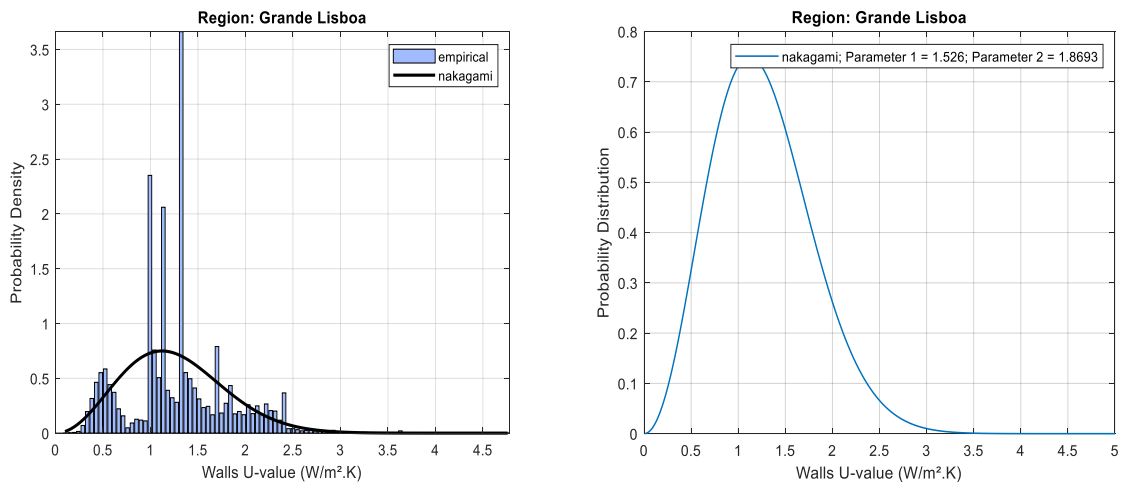


Figure 63 - Walls U-value ( $\frac{W}{K.m^2}$ ) in ‘Grande Lisboa’ Probability Distribution Function and Histogram (Left). Walls U-value ( $\frac{W}{K.m^2}$ ) in ‘Grande Lisboa’ Probability Distribution Function (Right).

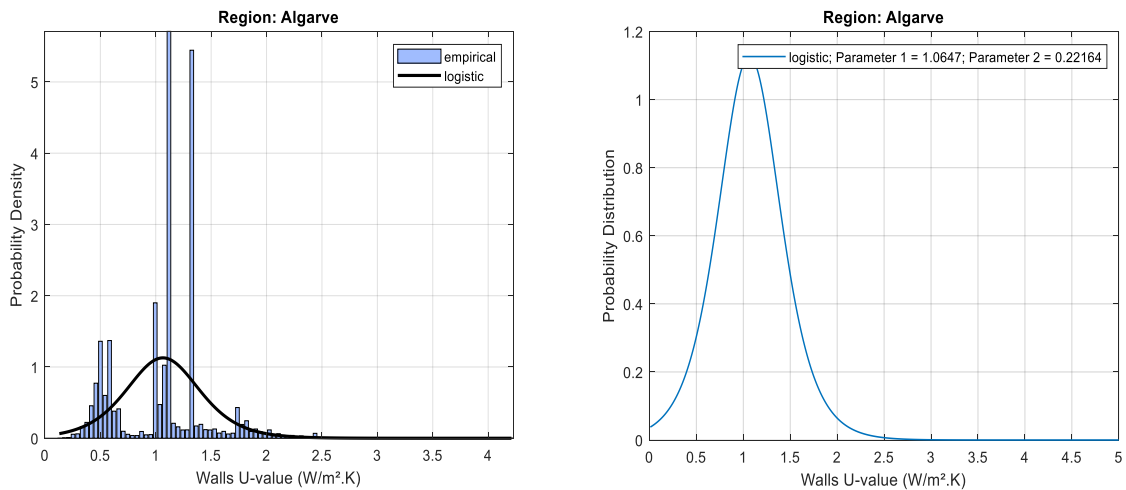


Figure 64 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Algarve' Probability Distribution Function and Histogram (Left). Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Algarve' Probability Distribution Function (Right).

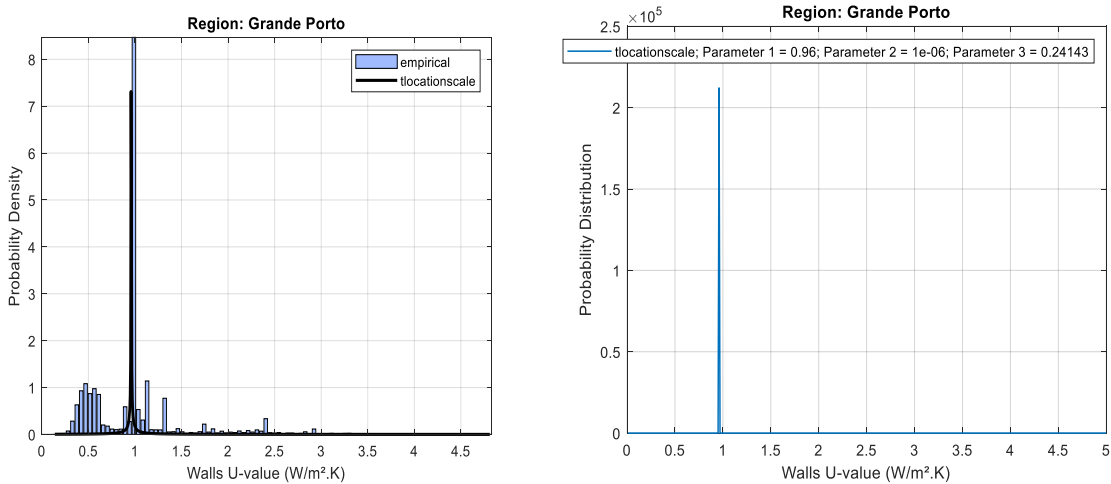


Figure 65 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Grande Porto' Probability Distribution Function and Histogram (Left). Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Grande Porto' Probability Distribution Function (Right).

Once more, an example of the importance of showing both histogram and probability distribution function is evidenced in *Figure 65*. We observe that in *Figure 65 (Left)* there are values of  $U_{walls}$ , around 0.5 that will not be noticed in the probability distribution function on the right, because of the large peak we see around 1, this way the characterization is flawed because the probability distribution function does not show the exact distribution of values, which is something that could be improved. So, the comparison of these three regions was tried, but unsuccessfully, because of the enormous probability distribution in 'Grande Porto' peak overlaps the other regions probability distributions. Therefore, as the probability distribution comparison between these three regions would look equal as *Figure 65 (Right)*, meanwhile the region's walls U-value has a median of 0.96, which is close to the national median for the same parameter. To overcome this problem the second-best fitting curve, with the second closest to zero  $NLogL$ , was plotted separately and resulted in *Figure 66*.

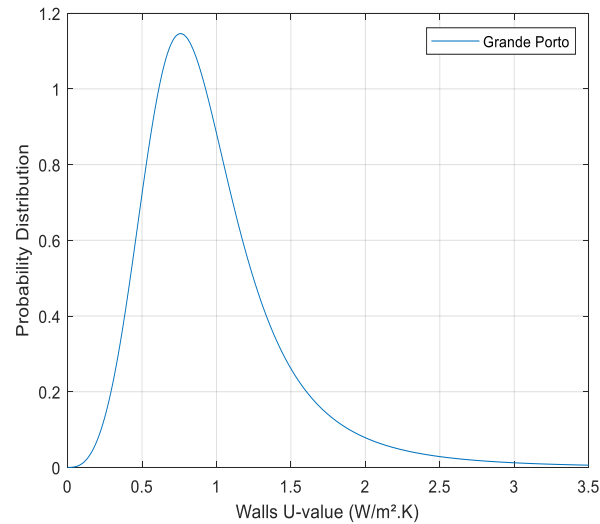


Figure 66 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Grande Porto' Probability Distribution Function.

This new probability distribution function is a loglogistic, with its distribution function parameters in *Table 16*. We are now able to compare all regions in one figure, the comparison made was then between the three regions, which is in *Figure 67*, with the boxplot figure comparison, *Figure 68*.

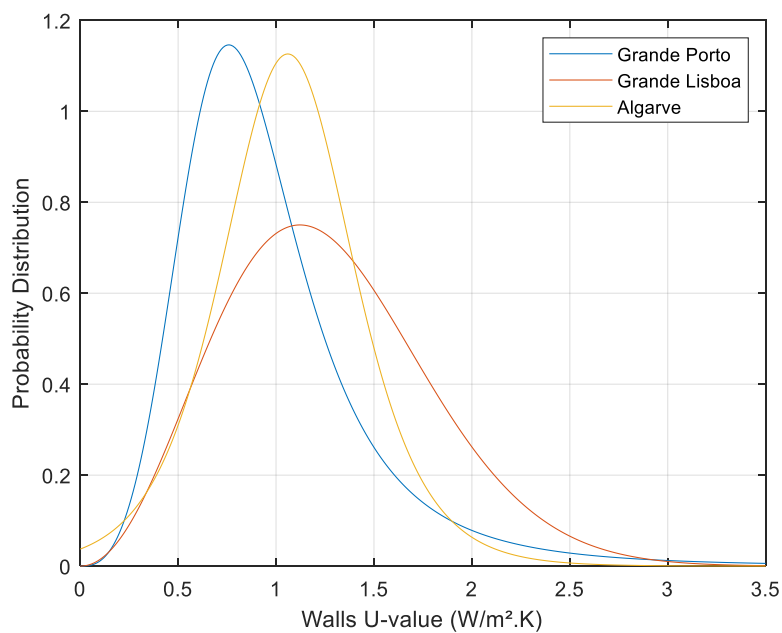


Figure 67 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Grande Lisboa', 'Grande Porto' and 'Algarve' Probability Distribution Function Comparison.

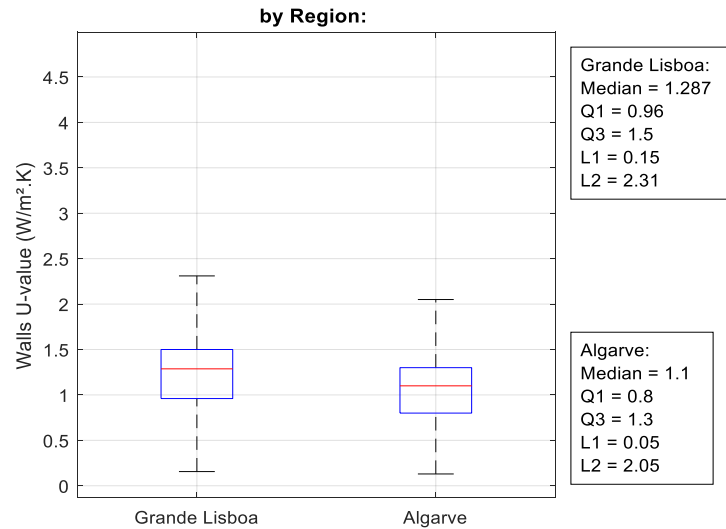


Figure 68 - Walls U-value ( $\frac{W}{K.m^2}$ ) in 'Grande Lisboa' and 'Algarve' Probability Boxplot Comparison.

We observe in Figure 67 that, although lower HDD occur in 'Algarve', this region has better thermal insulation in walls than 'Grande Lisboa' where HDD is higher, in Figure 68, medians are  $1.29 \frac{W}{K.m^2}$  in 'Grande Lisboa' and  $1.10 \frac{W}{K.m^2}$  in 'Algarve'. By this, we conclude that thermal insulation of walls in 'Grande Lisboa' is poor, if we compare it to the national sample in Figure 59 and if we compare it to another region with similar HDD values such as 'Algarve' or 'Alentejo Litoral'. When it comes to 'Grande Porto' the quality is higher than the other two, despite the few more HDD its walls have much better thermal insulation, closer to the national sample in Figure 59. Below we find a table with the probability distribution functions characteristics of the regions studied in this sub-chapter.

Table 16 - Distribution Function and Distribution Parameters for Each Region Studied in Sub-Chapter 4.8.

| Region              | Distribution Function | Dist. Parameter Names  | Dist. Parameter Values |
|---------------------|-----------------------|------------------------|------------------------|
| Alto Trás-os-Montes | inverse gaussian      | scale; shape;          | 0.811; 2.01;           |
| Grande Porto        | loglogistic           | location; scale;       | -0.130;0.267;          |
| Grande Lisboa       | nakagami              | shape; scale;          | 1.53; 1.87             |
| Alentejo Litoral    | rician                | non-centrality; scale; | 0.818; 0.617;          |
| Algarve             | logistic              | location; scale;       | 1.06; 0.222;           |

## 5 Conclusions

The computational model gives easy access to the information and data analysis of the energy performance certificates in the Portuguese residential building stock. To demonstrate the applicability of such tool, the development of the computational model was used to enable the answer for some questions about the characteristics of the residential building stock in Portugal. We can conclude that the RBS is not homogenous and that it has changed over time, recently at a faster pace than before. This change is happening towards a future where our dwellings will have higher energy performances in some decades, since the majority of dwellings in Portugal are categorized as '*Existing*', the category where most energy inefficient buildings are found and will tend to stay inefficient if not retrofitted nor energy efficiency measures are taken. However, '*New*' and '*Major Renovation*' were highly energy efficient due to REH implementations and the way paved by previous regulation before REH.

More specifically, parameters are not homogeneous along the residential building stock, renovations per hour sees a disparity between typical urban and rural regions, higher infiltrations were observed in urban regions due to the presence of higher buildings where the wind is more intense, while in rural areas lower buildings will have lower infiltrations. Differences were observed also in context, '*Existing*' saw a huge range of values from 0 to 1.5 renovations per hour, peaking around 0.5. REH regulation of air quality is verified when we analyse '*New*' and '*Major Renovation*', where in contrast to '*Existing*', renovations per hour value was concentrated around 0.4 to most dwellings. This is probably due to the improvement in building construction and air quality regulations, air infiltrations are minimized due to improvements in insulation in windows and doors, to balance thermal losses due to air flow and air quality.

The energy performance indicator measures the quotient of the primary energy needs and the energy needs reference value for the same dwelling, therefore it is the most relevant indicator to measure energy performance in buildings. This parameter was higher (the higher is worse) in '*1961-1990*' slowly reduced in the following period of '*1991-2005*', but it really experienced a sudden drop in the periods when the regulations were introduced, in '*2006-2012*' the SCE was introduced and in '*2013-2018*' this parameter reduced a little with the REH regulation. Apart from the periods in the past of which energy certification coefficient was studied, the future of the RBS was analysed, by studying the pre-certificates of buildings that have a project but are not yet built, in this sample it is observed that this energy performance parameter is set to increase slightly when compared to the previous period of '*2013-2018*', still it suits the requirement of a energy certification coefficient lower than 1 set in REH. Obviously, energy certification coefficient changed in the different contexts, in context '*Existing*' many dwellings were never under any regulation so its values were distributed in a wider range mostly between 0 and 4,

meanwhile ‘*New*’ and ‘*Major Renovation*’ were under REH, so they were required to have at least 1 in RNt value for newly constructed dwellings and 1.5 for retrofitted dwellings. The distribution of RNt was exactly 0 to 1 for ‘*New*’ and from 0 to 1.5 to ‘*Major Renovation*’.

The next parameter is the net floor area, this parameter is interesting to study regionally where we expect to see differences between rural and urban regions. The most populated and most building dense regions were analysed, first ‘*Grande Lisboa*’ and ‘*Grande Porto*’ with a median net floor area of  $74.75 \text{ m}^2$  and  $81.16 \text{ m}^2$  that were found to be smaller than rural regions like ‘*Alto Trás-os-Montes*’ with a median of  $112.52 \text{ m}^2$ , explained by the cheaper floor area when buying a house in a rural region and the by the greater land area available to build. Region ‘*Algarve*’ was analysed, and it seems that net floor area in this region and in the most urbanized region ‘*Grande Lisboa*’ is practically the same in distribution shape and in median. The net floor area also differs depending on the context. It was found that among contexts ‘*New*’ is the one with the largest net floor area by far with a median of  $124.70 \text{ m}^2$  a figure incredibly high if we consider that ‘*Existing*’ has a median of  $82.75 \text{ m}^2$ , right behind is ‘*Major Renovation*’ with a median of  $81.47 \text{ m}^2$ .

Window-to-floor area is lower for ‘*Existing*’ with median of 0.149, followed by ‘*Major Renovation*’ and ‘*New*’, medians of 0.176 and 0.206, respectively. As ‘*New*’ dwellings have more window area than ‘*Existing*’ this could lead to concerns about insulation. But with this increase in window area per floor area we can also observe that ‘*New*’ dwellings have lower window heat transfer coefficient (window U-value) than ‘*Existing*’, therefore more insulated, the medians for both are  $2.40 \frac{W}{K.m^2}$  and  $3.55 \frac{W}{K.m^2}$ , respectively. ‘*Major Renovation*’ is close to ‘*New*’ with  $2.37 \frac{W}{K.m^2}$ . We have seen how window U-value changes through the contexts, it is expected that this parameter decreases in with time, so in period ‘1961-1990’ median was  $3.87 \frac{W}{K.m^2}$ , surprisingly it then got a little worse in ‘1991-2005’ with  $3.90 \frac{W}{K.m^2}$ . With the introduction of the regulation in 2006 this number saw a huge decrease to  $2.94 \frac{W}{K.m^2}$  in ‘2006-2012’, afterwards this parameter improves with  $2.76 \frac{W}{K.m^2}$  in period ‘2013-2018’, pre-certificates reach  $2.31 \frac{W}{K.m^2}$ .

U-value of the walls was analysed, first for the national sample with a median of  $1.02 \frac{W}{K.m^2}$ . For region and climatic zone, we usually find better values of wall thermal insulation for regions where HDD is higher, like in region ‘*Alto Trás-os-Montes*’. In more urbanized regions, ‘*Grande Porto*’ and ‘*Algarve*’, were found to be really close the national sample, and ‘*Grande Lisboa*’ when compared to regions where HDD are similar show a poorer walls U-value, while ‘*Grande Porto*’ has the walls better insulated than the other two. Analysing this parameter, it was found a gap in this analysis, as the probability distribution function omits important values if there is an accentuated peak, then precious information ends lost. A

future improvement would pass through this, as it is necessary to find a more reliable method for data analysis and characterization.

As proposed, this work pretended to create a model that could characterise and analyse the database of the energy performance certification. This goal is achieved when it becomes possible to display probability distribution functions, histograms and other statistical information of a certain parameter for the variables asked. Many questions can come across and answers will also appear, while working on data analysis, these answers can directly be linked to the effect of policies, other answers are the product of the climate of different regions, or to urban and rural disparities. The most relevant parameters were studied, however, many other parameters are in the energy performance certification database and this work could be pursued in this way, to include more parameters in the computational models and relate them to other parameters.

## References

- [1] “Electricity Information 2018 Overview.” [Online]. Available: <https://www.iea.org/statistics/electricity/>. [Accessed: 05-Feb-2019].
- [2] “Decreto-Lei n°80/2006,” pp. 2468–2513, 2006. (in Portuguese)
- [3] “Decreto-Lei 118/2013, 2013-08-20 - DRE.” [Online]. Available: <https://dre.pt/pesquisa/-/search/499237/details/maximized>. [Accessed: 06-Feb-2019].
- [4] S. Kunkel, *ENERGY PERFORMANCE CERTIFICATES ACROSS THE EU*. .
- [5] E. K. Brand, B. Von Manteuffel, and A. Hermelink, “Energy Performance Certificate Database in Denmark,” 2018.
- [6] L. Murphy, “The influence of the Energy Performance Certificate : The Dutch case,” vol. 67, pp. 664–672, 2014.
- [7] A. Hardy and D. Glew, “An analysis of errors in the Energy Performance certificate database,” *Energy Policy*, vol. 129, no. March, pp. 1168–1178, 2019.
- [8] P. De Wilde, “Automation in Construction The gap between predicted and measured energy performance of buildings : A framework for investigation,” *Autom. Constr.*, vol. 41, pp. 40–49, 2014.
- [9] J. N. B. Fonseca and M. J. N. Oliveira Panão, “Monte Carlo housing stock model to predict the energy performance indicators,” *Energy Build.*, vol. 152, pp. 503–515, 2017.
- [10] Sara Magalhães and Vítor Leal, Characterization of thermal performance and nominal heating gap of the residential building stock using the EPBD-derived databases: The case of Portugal mainland, *Energy and Building*, Volume 70, February 2014, Pages 167-179.
- [11] A. C. L. Fernando Moreira da Silva, Helena Maria Bártolo, Paulo Bártolo, Rita Almendra, Filipa Roseta, Henrique Amorim Almeida, *Challenges for Technology Innovation: An Agenda for the Future: Proceedings*.
- [12] Instituto Nacional de Estatística INE, “Estatísticos Regionais – Informação estatística à escala regional e municipal.” 2015. (in Portuguese)
- [13] A. Energia, “Guia Prático do Certificado Energético da Habitação,” 2015. (in Portuguese)
- [14] Order No 15793-F/2013, “Despacho (extrato) n.º 15793-F/2013, de 3 de dezembro,” *Diário da República N.º 234, 2.ª série*, vol. 35088, no. 26, pp. 26–31, Lisboa, Portugal (in Portuguese).

## Appendix A – Rules and Recommendations

Table 17 - Reference Values for Heating Season (Part I), adapted from [14].

|                       | z        | M            |             | GD        |            | $\theta_{ext,i}$ |            | $G_{s,al}$                    |
|-----------------------|----------|--------------|-------------|-----------|------------|------------------|------------|-------------------------------|
|                       | REF<br>m | REF<br>meses | a<br>mês/km | REF<br>°C | a<br>°C/km | REF<br>°C        | a<br>°C/km | kWh/m <sup>2</sup><br>por mês |
| Minho-Lima            | 268      | 7,2          | 1           | 1629      | 1500       | 8,2              | -5         | 130                           |
| Alto Trás-os-Montes   | 680      | 7,3          | 0           | 2015      | 1400       | 5,5              | -4         | 125                           |
| Cávado                | 171      | 6,8          | 1           | 1491      | 1300       | 9,0              | -6         | 125                           |
| Ave                   | 426      | 7,2          | 0           | 1653      | 1500       | 7,8              | -6         | 125                           |
| Grande Porto          | 94       | 6,2          | 2           | 1250      | 1600       | 9,9              | -7         | 130                           |
| Tâmega                | 320      | 6,7          | 0           | 1570      | 1600       | 7,8              | -5         | 135                           |
| Douro                 | 579      | 6,9          | 0           | 1764      | 1400       | 6,3              | -4         | 135                           |
| Entre Douro e Vouga   | 298      | 6,9          | 1           | 1544      | 1400       | 8,4              | -5         | 135                           |
| Baixo Vouga           | 50       | 6,3          | 2           | 1337      | 1100       | 9,5              | -5         | 140                           |
| Baixo Mondego         | 67       | 6,3          | 0           | 1304      | 1000       | 9,7              | -5         | 140                           |
| Beira Interior Norte  | 717      | 7,5          | 0           | 1924      | 1000       | 6,3              | -3         | 135                           |
| Beira Interior Sul    | 328      | 5,4          | 1           | 1274      | 1800       | 9,1              | -6         | 140                           |
| Cova da Beira         | 507      | 7,1          | 0           | 1687      | 1400       | 7,5              | -5         | 140                           |
| Serra da Estrela      | 553      | 7,5          | 0           | 1851      | 1600       | 7,0              | -5         | 135                           |
| Dão - Lafões          | 497      | 7,3          | 0           | 1702      | 1900       | 7,5              | -6         | 135                           |
| Pinhal Interior Norte | 361      | 6,8          | 0           | 1555      | 1600       | 8,3              | -5         | 140                           |
| Pinhal Interior Sul   | 361      | 6,7          | 1           | 1511      | 1500       | 8,4              | -4         | 145                           |
| Pinhal Litoral        | 126      | 6,6          | 0           | 1323      | 1900       | 9,6              | -5         | 140                           |
| Oeste                 | 99       | 5,6          | 0           | 1165      | 2200       | 10,3             | -8         | 145                           |
| Médio Tejo            | 168      | 5,9          | 0           | 1330      | 1300       | 9,5              | -4         | 145                           |
| Lezíria do Tejo       | 73       | 5,2          | 3           | 1135      | 2700       | 10,2             | -7         | 145                           |
| Grande Lisboa         | 109      | 5,3          | 3           | 1071      | 1700       | 10,8             | -4         | 150                           |
| Península de Setúbal  | 47       | 4,7          | 0           | 1045      | 1500       | 10,7             | -4         | 145                           |
| Alto Alentejo         | 246      | 5,3          | 2           | 1221      | 1200       | 9,6              | -3         | 145                           |
| Alentejo Central      | 221      | 5,3          | 2           | 1150      | 1100       | 10,0             | -4         | 150                           |
| Alentejo Litoral      | 88       | 5,3          | 2           | 1089      | 1100       | 10,8             | -2         | 150                           |
| Baixo Alentejo        | 178      | 5,0          | 0           | 1068      | 1000       | 10,7             | -2         | 155                           |

Table 18 - Reference Values for Heating Season (Part II), adapted from [14].

|              | z        | M            |             | GD        |            | $\theta_{ext,i}$ |            | $G_{s,al}$                    |
|--------------|----------|--------------|-------------|-----------|------------|------------------|------------|-------------------------------|
|              | REF<br>m | REF<br>meses | a<br>mês/km | REF<br>°C | a<br>°C/km | REF<br>°C        | a<br>°C/km | kWh/m <sup>2</sup><br>por mês |
| Algarve      | 145      | 4,8          | 0           | 987       | 1800       | 11,3             | -6         | 155                           |
| R.A. Açores  | 10       | 2,9          | 1           | 604       | 1500       | 14,4             | -7         | 110                           |
| R.A. Madeira | 380      | 3,2          | 1           | 618       | 1500       | 14,8             | -7         | 105                           |

Table 19 - Criteria for Climatic Zones in Heating Season, adapted from [14].

| Crítério | GD ≤ 1300 | 1300 < GD ≤ 1800 | GD > 1800 |
|----------|-----------|------------------|-----------|
| Zona     | I1        | I2               | I3        |

