Contents lists available at ScienceDirect



# **Building and Environment**



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

# An envelope airtightness predictive model for residential buildings in Spain



Irene Poza-Casado<sup>a</sup>, Pilar Rodríguez-del-Tío<sup>b</sup>, Miguel Fernández-Temprano<sup>b</sup>, Miguel-Ángel Padilla-Marcos<sup>a</sup>, Alberto Meiss<sup>a,\*</sup>

<sup>a</sup> GIR Arquitectura & Energía, Dpto. Construcciones Arquitectónicas, Ingeniería del Terreno y Mecánica de los Medios Continuos y Teoría de Estructuras, Universidad de Valladolid. E.T.S. de Arquitectura, Avenida Salamanca, 18, 47014, Valladolid, Spain

<sup>b</sup> Dpto. Estadística e Investigación Operativa, E. de Ingenierías Industriales - Sede Doctor Mergelina, Paseo Prado de la Magdalena 3-5, 47011, Valladolid, Spain

#### ARTICLE INFO

Keywords: Predictive model Airtightness Blowerdoor Dwellings Database Statistical analysis

#### ABSTRACT

The need for airtightness control is a reality given its impact on buildings' energy use and Indoor Air Quality (IAQ). For the past few years, this fact has resulted in energy performance regulations establishment that involves airtightness requirements in many countries in Europe and North America. In this sense, efforts should not only be focused on new buildings, but also existing ones. Considering that around 90% of the built stock in the EU is expected to still be standing in 2050 and that almost 75% of the buildings are energy inefficient, attention must be paid to retrofitting actions.

Airtightness predictive models have become useful in the decision-making process and to estimate input values in energy performance simulation tools. So far, several predictive models have been developed in different countries. However, specific construction systems and practices lead to a lack of consensus regarding the impact of different factors on airtightness performance. Therefore, the applicability of existing models is limited to their specific contexts.

This paper presents a predictive model for envelope airtightness, which was developed from a database that contains a fully characterised representative sample of the residential building stock in Spain. A General Linear Model (GLM) was considered to assess significant variables related to the age of the building, typology, building state, construction system, and dimensions. As a result, a predictive model is presented and validated. Overall, even if some limitations were identified, the relevance of the model proposed is warranted from the statistical point of view. The airtightness predictive model presented offers a procedure for airtightness estimation of residential buildings in Spain.

# 1. Introduction

The energy crisis has led to a need to reduce energy consumption and, thus, greenhouse gas emissions. In this sense, the European Union (EU) is committed to establishing a sustainable, competitive and decarbonised energy system by 2050 as part of the Green Deal [1]. In this context, buildings' energy performance is crucial, and the EU has already established a legislative framework to address this issue through the Energy Performance of Buildings Directive (EPBD) [2]. The EPBD determined that all new buildings must be nearly Zero-Energy Buildings (nZEB) and established long-term renovation strategies to improve the existing building stock. In this sense, the Renovation Wave aims to renovate 35 million inefficient buildings by 2030 [3].

However, the low replacement rate of existing buildings into efficient ones under current energy standards (0.4-1.2%) highlights the need for strategies that contribute to the renovation of national building stocks [4]. This is key to shifting to a low-carbon building stock considering that 85–95% of buildings in the EU are expected to still be standing in 2050 [3].

In this regard, the reduction of heat transmission through the building envelope has been a priority. On the contrary, air infiltration is often overlooked in spite of the great energy impact on the overall energy performance of buildings, which can account for 10%–30% of the heating demand depending on climate [5–13].

Understanding airtightness as the main building feature that impacts air infiltration [14] is decisive and, therefore, airtightness assessment turns key to evaluating energy performance and prioritising efforts towards renovation strategies. This involves not only airtightness quantification, but also the identification of the main leakage paths and factors that condition global performance.

https://doi.org/10.1016/j.buildenv.2022.109435

Received 11 May 2022; Received in revised form 8 July 2022; Accepted 19 July 2022 Available online 2 August 2022 0360-1323/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article

<sup>\*</sup> Corresponding author. E.T.S. de Arquitectura, Avenida Salamanca, 18, 47014, Valladolid, Spain. *E-mail address: alberto.meiss@uva.es* (A. Meiss).

<sup>0360-1323/© 2022</sup> The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1.1. State of the art

# Building and Environment 223 (2022) 109435

This issue has been broadly approached in the literature from the analysis of airtightness databases in the USA [15], Canada [16], France [17], the UK [18,19], and Belgium [20]. Among others, the year of

construction, building geometry-related factors, materials used, type of construction and climate seem to be influencing factors for airtightness. Overall, common trends can be identified although some contradictions among different analyses were found due to the fact that construction practices vary considerably among regions, and differing approaches

# Table 1

Summary of previous predictive models developed from airtightness databases.

Country	Database origin	Number of cases	Type of building	Metric	Methodology	Model variables	R <sup>2</sup>	Ref.
USA (Mcwilliams & Jung, 2006)	Different sources from research programs, weatherization assistance programs (WAP) and energy rating programs	100,000	Single-family houses. The database does not contain equally distributed data that is representative or geographically uniform	NL [-]	Multiple linear regression	Climate zone, floor area, height, basement type, energy program, occupant income, testing age, floor leakage	0.3	[23]
USA (Chan et al., 2013)	Same source as Mcwilliams & Jung, 2006	134,000	Same source as Mcwilliams & Jung, 2006	NL [-]	Multiple linear regression	Year built, climate zone, WAP, energy efficiency rate, floor area, height, foundation type, duct location	0.68	[15]
Croatia (Krstić et al., 2014)	Research study	58	Representative dwellings in terms of year of construction in the area of Osijek- Baranja County	n <sub>50</sub> [h <sup>-1</sup> ]	Neural networks	Input parameters: wall thickness, ceiling/floor thickness, wall material, ceiling/floor material, thermal insulation thickness, quality of installed joinery, maintenance quality, opening frame material, type of opening glazing, share of transparent envelope, share of exposed envelope	-	[27]
Netherlands (Bramiana et al., 2016)	Blower door organizations' databases	320	Single-family houses across the country at different stages of the construction process	w <sub>10</sub> [dm <sup>3</sup> / (s·m <sup>2</sup> )]	ANOVA and multiple linear regression	Year of construction, total leakage	0.421	[25]
UK (Pan, 2010)	Building companies	287	Post-2006 new-built dwellings, reasonable geographical coverage, different typologies	q <sub>50</sub> [m <sup>3</sup> / (hm <sup>2</sup> )]	One-2-3-way ANOVA, multiple linear regression	Dwelling type, management context, building method + interaction between building method, dwelling type	0.49	[26]
Canada (Khemet and Richman, 2018)	Residential energy efficiency program	Subset of 330,000 of a 900,000 cases database	Low-rise single-family detached dwellings throughout Canada built between the late 1700s to 2016	n <sub>50</sub> [h <sup>-1</sup> ] and NL [–]	Univariate analyses and multiple linear regression. Development of two models	3-variable model: building volume, year, height. 8-variable model: building volume, year, height, heating degree days, ceiling, foundation, window, wall	3-variable model: 0.32 8-variable model: 0.46	[16]
Canada (Khemet and Richman, 2021)	National airtightness database	2297	Light framed, detached homes in a cold climate zone in Canada	n <sub>50</sub> [h <sup>-1</sup> ] and NL [-]	Multiple linear regression. Development of seven models from subsets	Volume, Rim Joist Ratio, Fenestration Perimeter to Height Ratio, Ceiling to Wall Ratio, Above Grade Height, Perimeter Ratio, Shell Area, Exposed Area to Condition Floor Area, Ceiling Area, Builder Identification	0.262–0.865	[24]
Spain-France (Montoya et al., 2010)	CETE de Lyon database	251	Dwellings in France from 1983 onwards	C' [m <sup>3</sup> / (s Pa <sup>2/</sup> <sup>3</sup> ]	Multiple linear regression	Structure type, floor area, age of the building, number of stories	0.37	[28]
Spain (Fernández- Agüera et al., 2016)	Research study	45	Seven multifamily buildings (2011–2013) in southern Spain. Low- income, open gallery- type buildings	n <sub>50</sub> [h <sup>-1</sup> ]	ANOVA and multiple linear regression with variable selection using factorial analysis	Mean floor area, volume, mean window area, mean window perimeter	0.98	[29]
Spain (Fernández- Agüera et al., 2019)	Research study	Model 1: 98 Model 2: 53	Low-income multifamily buildings in Southern Spain	${}^{n_{50}}$ [h <sup>-1</sup> ]	PCA and clustering. Development of two models for buildings built before and after 1979. Separate multiple linear regression in each cluster	Window area, window perimeter, window type, bathroom window, blinds, separate kitchen, winter severity, exposure, façade type, net floor area, general condition	Model 1: 0.887 Model 2: 0.626	[30]

regarding metrics, parameters definition and statistical approach [21, 22].

Deep analysis of statistical relationships between the variables of airtightness datasets led to the development of airtightness predictive models over the past few years in the USA [15,23], Canada [16,24], the Netherlands [25], the UK [26], or Croatia [27], and Spain [28–31] (Table 1). Although airtightness estimation cannot be seen in any way as a replacement for on-site measurement [32], the tightening of building regulations has strongly increased the interest in tools to analytically assess airtightness. Predictive models are useful to consider the infiltration phenomenon in energy performance (EP) simulation tools in order to control costs and time in the decision-making process before building construction and retrofitting actions.

The analysis of previous approaches shows that the purposes and origin of the airtightness databases that were used for the development of predictive models determined their size and type of buildings included. Government energy programs resulted in huge datasets in the USA and Canada [15,16,23], although those studies could be lacking certain building types and uniform distribution. In addition, those databases from different sources did not contain standardised and complete information. On the other hand, the limited number of cases and restricted contexts in the case of Croatian and Spanish models [27–30] hindered their applicability to specific populations.

Most of the models studied contained airtightness information of existing dwellings, which turns critical when evaluating the impact of potential improvements and retrofitting actions before assuming any investment [31]. But it is important to highlight that understanding the performance of new buildings is also useful to evaluate current practices and assess design at an early stage, such as in the models developed in the UK and Canada [24,26]. The latter introduces several builder-specific, geometric-based and temporally independent models.

The available data concerning airtightness results of the databases and different national requirements explain the diverse metrics used by the models. This adds difficulty when comparing the outcomes. Turning to the methodology, aside from some exceptions, ANOVA combined with multivariate linear regression seem to be the widespread approach to building predictive models.

In line with previous analyses, common variables are present in most of the models, which account for the location or climate conditions [15, 16,23,30], the age of the building [15,16,23,25,28], dimension-related characteristics [15,16,23,24,27–30], building systems [15,16,23,26, 27,30], and, to a lesser extent, type of construction [26,28], participation in an energy program [15,23], its conservation state [27,30], or workmanship [24,26].

As previous authors mentioned, interactions among factors and the influence of supervision and workmanship add difficulties to the analysis of factors that impact airtightness [21]. This could also be explained by the lack of data or representativeness and is reflected in their fitting quality ( $R^2$ ), which is not too high in most cases, except for very uniform and specific datasets, such as the case of Spanish ones [29,30].

In Spain, specifically, previous approaches regarding airtightness estimation based on specific regions and typologies were developed [28–31]. Montoya et al. [28] conceived a model to predict air leakage of single-family dwellings in Catalonia. Arguing that construction procedures are similar in Catalonia and France, the model was based on data from 251 dwellings in France. Another example is an algorithm that was developed by Ibáñez-Puy and Alonso [31] to predict the airtightness of new and retrofitted dwellings. However, the algorithm is not fully detailed in the paper or published elsewhere to the authors' knowledge, making it very difficult to be applied in practice or compared to others. For this reason, the model was not included in Table 1. The authors revealed that the most important factor was workmanship, which added difficulties to its integration into a model. Similar conclusions were reached by Fernández-Agüera et al. [29,30] from the study of low-income multifamily buildings in Southern Spain.

All the models previously assessed in Spain were obtained from

samples that consider specific regions and typologies, and, thus, their applicability is restricted to those populations. On an international scale, other mentioned models were developed from built stocks with construction typologies, characteristics and configurations that vary among regions. This seems to be a drawback that makes airtightness predictive models not easily reproducible nor exportable to different contexts.

Therefore, there is a lack of knowledge in this field and, in this sense, a predictive model from representative experimental data is proposed to estimate the level of airtightness of existing dwellings in Spain, which is applicable at a national level. This model will shed light on remaining open questions on a scientific level, such as understanding the factors that most impact airtightness in the national-built stock. The model presented could easily be applied or adapted to other Southern European countries where construction systems have similarities and there has also been a lack of specific limitations concerning envelope airtightness [29,33,34], and its methodology, which considers interactions among variables in a natural way, could be reproduced for other datasets.

The outline of the rest of this paper is as follows. In the next section, the airtightness database is described and the regulatory framework in Spain is considered and analysed showing that the estimations obtained from this framework are not consistent with real data. Then, in subsequent sections, the methodology used for the predictive model considered here is described and applied to the data. Moreover, the predictive model generated is exposed and validated in full detail. The final sections discuss the results and present the final conclusions obtained.

# 2. Context in Spain

#### 2.1. Airtightness database

The proposed model was developed from results gathered in INFILES national airtightness database [35]. The database consisted of a representative sample of the existing residential building stock in Spain. For this purpose, a sampling method was carried out by means of a non-probabilistic quota sampling scheme considering climate zone, age and typology as control variables [36].

Airtightness tests were conducted in 400 dwellings in Spain whose owners allowed voluntarily the performance of the test. The final distribution of the dwellings followed strictly the criteria set by the sampling design [37,38]. Although the sample may seem limited compared to other previous approaches detailed in Section 1.1, the cases are uniformly distributed and are representative of the national built stock.

Four simplified climate zones were established according to their representativeness: Mediterranean climate (209 cases, 52.1%), Continental zone (129 cases, 32.2%), Oceanic zone (47 cases, 11.7%) and the Canary Islands (16 cases, 4%).

Regarding the age of the building, airtightness tests were performed in dwellings built between 1880 and 2015. The periods of major construction activity in Spain during the decades 1960–1979 (148 cases, 36.9% of the sample), and the period 1980–2006, including the years of the real state bubble, just before its bursting in 2007 (158 cases, 39.4% of the sample) were the most represented ones. According to the planned sampling scheme, 325 cases were apartments (81%), and 76 cases were single-family houses (19%).

The selected cases of the sample were characterised, not only in terms of airtightness but also through a whole process of feature data gathering. The need to capture and manage a great volume of data led to the development of a specific tool called *infilAPP* [35,36], which constituted the core of the database. A wide characterisation of different parameters, more than 140, were stored in the database, including identification information, configuration, construction of the envelope, and building systems.

Global descriptive results, their distribution and the impact of different parameters on airtightness can also be found in the literature [37,39,40]. However, so far, no estimation model has been proposed from the database.

## 2.2. Regulatory framework

The traditional dependence of ventilation on air renewal supplemented by window airing and air infiltration has hindered Mediterranean countries with temperate climates, and, specifically, Spain, awareness of the impact of the lack of airtightness. Currently, mandatory controlled ventilation systems that guarantee adequate IAQ in residential buildings make air infiltration no longer necessary as an air renewal source. Therefore, the need to comply with EPBD has led to a change in the scenario, and, in order to meet energy requirements, the improvement of airtightness of the building envelope becomes a priority to achieve nZEB.

Until 2019, airtightness was only considered establishing maximum permeability values for doors and windows of the building envelope depending on the winter severity of the climate zone [41–43].

In December 2019, a new update of DB HE1 [44] came into force introducing whole envelope airtightness. This requirement is only applicable for new and retrofitted dwellings for private use with a floor area greater than 120  $m^2$ . Limit airtightness values are established considering the air change rate at 50 *Pa* ( $n_{50}$ ) as reference metric, which vary between 3 and 6  $h^{-1}$  depending on the compacity of the dwelling. Compliance with limits can be proved either by performing pressurization tests or by means of reference values.

Pressurization tests can be performed according to Method 2 described on UNE-EN ISO 9972 [45], considering no further guidelines. It must be noted, though, that there is currently no quality control system nor tester scheme, in contrast with other countries like France [46], the UK [47], Germany [48] or Belgium [20]. Therefore, the consistency of results could be compromised.

Compliance with airtightness limits can be also proved analytically from reference values. The air change rate at 50 Pa ( $n_{50}$ ) can be estimated through reference values provided by regulations, as shown in Equation (1):

$$n_{50} = 0.629 \cdot \frac{C_0 \cdot A_0 + C_h \cdot A_h}{V}$$

$$\tag{1}$$

where:  $n_{50}$  is the air change rate at 50  $Pa[h^{-1}]$ ; *V* is the internal volume of a building or part of a building  $[m^3]$ ;  $C_0$  is the airflow coefficient of the opaque part of the thermal envelope at a reference pressure of 100 Pa $[m^3/hm^2]$ . Reference values are assigned depending on the type of building. For new or existing buildings with improved airtightness,  $C_0$  is 16  $m^3/hm^2$ , whereas for existing buildings a value of 29  $m^3/hm^2$  is assumed;  $A_0$  is the sum of areas of the opaque thermal building envelope with heat exchange with the outdoor air. Therefore, internal partitions and the envelope area in contact with other adjacent spaces or buildings are excluded  $[m^2]$ ;  $C_h$  is the permeability of doors and windows on the thermal building envelope at a reference pressure of 100  $Pa [m^3/hm^2]$ , according to laboratory testing results provided by the manufacturer;  $A_h$ is the sum of the area of the doors and windows of the thermal building envelope  $[m^2]$ .

In practice, this analytical approach is generally preferred by designers and construction companies and, thus, tests are only performed to comply with voluntary energy programmes (Passivhaus, BREEAM, LEED, etc.), or by some contractors who wish to ensure the quality of construction.

#### 2.3. Regulatory analytical model assessment

Since airtightness testing is still not a widespread approach to prove compliance with envelope airtightness limitations, it is crucial to validate the current airtightness estimation model proposed by regulations. The study of the estimation method proposed by Spanish regulations as an alternative to pressurization test performance revealed some inconsistencies regarding envelope area definition and reference values for airflow coefficients. In addition, results obtained from the model were compared to values obtained from pressurization tests performed in a representative sample of existing dwellings, as described in Section 2.1. It must be noted, though, that the sample of dwellings tested was built before current regulations came into force. Therefore, cases were not subject to whole envelope airtightness limitation.

The model suitability for existing dwellings was evaluated considering an initial sample of 400 dwellings. After an outlier detection procedure, 8 dwellings were dropped, so that a final sample of 392 observations was analysed. Fig. 1 shows the results of a correlation analysis between the  $n_{50}$  values obtained in the pressurization tests and the values obtained using the CTE model. In Fig. 1 it can be seen that there is a lack of linear association between the values of the CTE model and the test values as the correlation is as low as 0.15. However, it can also be observed that both variables are highly asymmetric so it might be argued that the lack of linear association is due to this feature.

For this reason, a second correlation analysis was performed with the logarithms of these two  $n_{50}$  values. Fig. 2 shows the results of this second analysis. It can be seen that the transformed variables are now symmetric and approximately normal but that the correlation is still as low as 0.20, which confirms the low linear association between the two  $n_{50}$  values. Notice that, taking into account the relationship between the correlation coefficient and the coefficient of determination  $R^2$  of the simple regression model, this means that only 100 x  $0.20^2 = 4\%$  of the variability of the measured  $n_{50}$  values is explained by the CTE model. Therefore, it can be concluded that CTE model does not fit well the real values obtained from test results.

In this sense, the analysis performed proved that the CTE model is not suitable for the estimation of the airtightness of existing buildings. Although on-site testing is the only reliable way to assess airtightness for energy certification of buildings, an accurate estimation can be useful for energy performance calculation tools, and also in order to prioritise efforts and determine strategies for building design and renovation of the existing building stock. In this sense, the airtightness characterisation of the envelope turns essential to reaching the decarbonisation objectives set.

The model to be presented in this study can be used on a national scale since it considers a much bigger representative sample (400 dwellings) than previous approaches, it is not restricted to a single city or climate zone, and includes a wide range of building typologies as will be apparent when the different variables measured are described (see Table 2 below). This model could constitute an alternative for airtightness estimation of residential buildings in Spain in a more accurate way.

## 3. Methods

## 3.1. Airtightness testing

The airtightness of the cases assessed was measured by means of fan pressurization tests, commonly known as blower-door tests, according to the International Standard ISO 9972 [49]. Further and specific guidelines were gathered in a specific protocol that was developed, whose definition was verified and validated, and broadly detailed by Feijó et al. [36]. In multi-family buildings, individual-unit tests were performed considering no guarded-zone tests.

Results were obtained from the commonly referred to as the powerlaw (Equation (2)), which describes the relationship between the airflow rate and pressure difference [50]:

$$Q_{pr} = c_L \cdot \left(\Delta p_r\right)^n \tag{2}$$

where:  $Q_{pr}$  is the airflow rate of the opening at a reference pressure difference  $[m^3 / h]$ ;  $C_L$  is the air leakage coefficient  $[m^3 / (h Pa^n)]$ ;  $\Delta p_r$  is the reference pressure difference [Pa]; *n* is the airflow exponent [-].

Results at a reference pressure of 50 *Pa* were then normalised by the volume of the premises under study (*V*), obtaining the air change rate at the reference pressure ( $n_{50}$ ) as in Equation (3).



Fig. 1. Correlation analysis between the n<sub>50</sub> values obtained from pressurization tests and those computed using the CTE model.



Fig. 2. Correlation analysis between the logarithms of the n<sub>50</sub> values obtained in the pressurization tests and those computed using the CTE model.

$$n_{50} = \frac{Q_{50}}{V}$$
(3)

where:  $n_{50}$  is the air change rate at 50  $Pa[h^{-1}]$ ;  $Q_{50}$  is the air leakage rate at 50  $Pa[m^3/h]$ ; V is the internal volume  $[m^3]$ .

It expresses the exchange of the equivalent volume of indoor air per hour  $[h^{-1}]$ . It is of particular interest when referring to ventilation airflows. In Spain, the airtightness obtained from the calculation method from reference values is expressed through the metric  $n_{50}$ , and results obtained are introduced in the official EP calculation tool LIDER/CAL-ENER (HULC) in order to verify the requirements established by regulations. Furthermore, although there is no agreement among countries, this metric has also been used in previous approaches as shown in Table 1, which allows comparability.

#### 3.2. Statistical model development

In order to establish a predictive model for airtightness, we considered a General Linear Model (GLM), which is a statistical methodology that allows assessing which variables (both categorical and quantitative) have a significant influence on the response variable. In the model, we have not only considered the main effects of the explanatory variables but also possible first-order interactions among them. The formulation of this model can be written as in Equation (4):

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i + \sum_{i < j} \tau_{ij} (X_i X_j) + \varepsilon$$
(4)

where: Y is the response variable to be predicted,  $X_i$  with i = 1, ..., p are the explanatory variables,  $\beta_i$  are the main effects of the explanatory variables on the response,  $\tau_{ij}$  are the first-order interactions among variables  $X_i$  and  $X_i$ , and  $\epsilon$  are the random independent homoscedastic

normal perturbations. For the qualitative explanatory variables, the usual decomposition in dummy indicator variables has been considered.

After fitting the full model, outlier detection and elimination were performed and then a stepwise procedure was considered. This procedure starts with the model containing all variables and then an iterative procedure is performed. In this procedure, in each step the less significant variable (the one with highest p-value) is dropped, provided its p-value is higher than 0.05, to simplify the model and drop those explanatory variables that are not significant enough in the model. The model is then refitted without that variable and the next step performed. In this process model parsimony was considered in each step, i.e. no main effect was dropped if the corresponding variable had any significant interaction present in the model. The process stops when all variables in the model are significant, either through their main effects or through an interaction with other variables. In this way, it is ensured that all variables in the final model are significant. Residual analyses were also performed at each step to check the GLM assumptions of linearity, homoscedasticity, independence and normality.

In this study, due to the asymmetry exposed by  $n_{50}$  in the initial study appearing in Fig. 1, we considered log ( $n_{50}$ ) as response variable Y in model (4).

As for possible explanatory variables  $X_i$  to be included in the predictive model, variables related to location, age of the building, building typology, state, building systems, and dimensions were considered. First-order interactions among these variables were also considered. Table 2 contains a list of the variables initially considered, detailing which ones had a significant impact on the response variable. These significant variables are fully described below.

The relationship among variables and significance of the assessed variables on airtightness results were addressed through statistical analysis. The following variables were significant and, therefore, considered in the model proposed:

#### Table 2

Variables considered classified according to their type and their significance in the GLM model.

Type of variable	Variables in the final model	Variables dismissed in the final model
Location	Climate zone (CTE)	City Winter severity climate Summer severity climate Simplified climate zone
Age of the building	Period of construction	Year of construction Decades of construction Applied regulations
Type of building	Typology	Position within the building Height Number of floors Property developer Number of rooms/bathrooms Layout of the floorplan
Building state	Retrofitting state	Improvement of thermal bridges Identified cracks Closed balconies Integrated balconies Kitchen refurbishment Bathroom refurbishment Improvement of the envelope
Building system	False ceiling Window permeability Window material Shutter position	Envelope layer composition Outer cladding Insulation of the envelope Air chamber Windows opening system Double window Shutter type Partitioning system Heating system Cooling system Ventilation system Adventitious openings Ductwork Kitchen hood exhaust
Dimensions	Share of windows Share of opaque envelope	Floor area Volume Envelope area Compacity Ceiling height Share of wet rooms Windows joint length Window area Share of joint length

- Climate zone: climate was considered according to DB HE1 [44] regarding winter (zones A to E and  $\alpha$  and summer severity (1–4). Climate severity combines degree-days and solar radiation in each location. From the international perspective, these zones would have the following equivalence in the Köppen-Geiger climate classification [51]: A3 = Csa, B4 = BSk-Csa, C1 = Csb-Cfb, C2 = Csa, C3 = BSk, D2 = Csb,  $\alpha$ 3 = BSh.
- Period of construction: the age of the building is related to Energy Performance Regulations (EPR) over time. This fact was assessed by considering cases built before and after the first national regulations that established measures related to energy performance were implemented in 1980 [42].
- Typology: dwellings were classified as single-family or multi-family buildings given the impact that different construction systems and envelope features may entail. This variable is key in Spain, where multi-family housing prevails.
- Retrofitting estate: dwellings tested could be in their original state, or the envelope could have been retrofitted by their owners to a variable extent (windows replacement, external/internal insulation layer, etc.).
- False ceiling: the presence of this element can lead to the concealment of construction imperfections and, thus, leakages. A simplified characterisation was addressed considering dwellings with no false

ceiling (FC0), dwellings with false ceiling only in corridor, kitchen and bathroom (FC1), and dwellings with false ceiling in all the rooms (FC2).

- Window permeability: the air permeability of windows was assessed according to UNE-EN 12207 [52] and classified as Class 0 (not tested windows), Class 1 (up to 50  $m^3/h m^2$ ), Class 2 (up to 27  $m^3/h m^2$ ), Class 3 (up to 9  $m^3/h m^2$ ), or Class 4 (up to 3  $m^3/h m^2$ ). It must be noted, though, that this information was not always available and could be just estimated from visual inspection.
- Window material: the impact of window frame material was considered (aluminium, PVC, wood, steel). The most representative material was considered when more than one type of window was found.
- Shutter position: shutters are widely used in Spain, and they have an important impact on the envelope airtightness since they constitute a discontinuity of the envelope. Rolling shutters were classified regarding their position: non-integrated shutters, external shutters, internal shutters, and no shutters, according to Fig. 3. The most common solution is external shutters integrated into the inner layer of the envelope, whereas non-integrated shutters make reference to cases that originally had no shutter, and it is added constituting no additional leakages.
- Share of windows: it is the sum of the area of doors and windows related to the total envelope area. This parameter is closely related to  $A_h$  in the model proposed by Spanish regulations. This is a quantitative variable  $[m^2]$ .
- Share of opaque envelope: it is the sum of areas of the opaque thermal building envelope with heat exchange with the outdoor air related to the total envelope area of the dwelling. This parameter is closely related to  $A_0$  in the model proposed by Spanish regulations. This is a quantitative variable  $[m^2]$ .

#### 4. Predictive model results

All analyses in this section: descriptive study, model estimation, variable selection and model validation, were performed with IBM SPSS software [53].

## 4.1. Descriptive study

The outlier detection procedure mentioned in the previous section resulted in the elimination of 8 observations that had anomalous log  $(n_{50})$  values possibly due to measurement errors. Therefore, in the final model 392 observations are considered. Table 3 contains a descriptive study of the explanatory variables in the final model while Table 4 gives a more detailed descriptive study of the initial response variable  $n_{50}$  and the final transformed response variable log  $(n_{50})$  and Fig. 4 shows histograms of these two variables.

### 4.2. Predictive model

The final predictive model selected according to the procedure described in the statistical model development section resulted in a model containing 10 main effects and 2 interactions. The ANOVA table corresponding to this model is shown in Table 5. This table shows the variability of the response variable explained by each of the explanatory variables and interactions included in the model and whether this explained variability is statistically significant or not. It can be observed that all main effects are significant at the usual 0.05 level, except the main effect of the variable "Share of opaque envelope". However, the interaction of this variable with both "Typology" and "Period of construction" is significant at that 0.05 level. These interactions mean that the effect of "Share of opaque envelope" on log ( $n_{50}$ ) is significantly different for the different levels of "Typology" (that are single-family and multifamily dwellings) and for those dwellings built before and



Fig. 3. Shutter position classification.

## Table 3

Descriptive study for the explanatory variables n the final model.

	Value	Ν	%
Retrofitting state	Original	271	69.13%
	Retrofitted	121	30.87%
Climate zone	A3	33	8.42%
	B4	85	21.68%
	C1	47	11.99%
	C2	85	21.68%
	C3	112	28.57%
	D2	16	4.08%
	α3	14	3.57%
Period of construction	Before 1980	219	55.87%
	Since 1980	173	44.13%
Window permeability	Class 0 or 1	46	11.73%
	Class 2	196	50.00%
	Class 3	117	29.85%
	Class 4	33	8.42%
Window material	Steel	5	1.28%
	Aluminium	263	67.09%
	Wood	54	13.78%
	PVC	70	17.86%
Shutter position	P.01	19	4.85%
	P.02	290	73.98%
	P.03	21	5.36%
	P.04	62	15.82%
False ceiling	FC0	85	21.68%
	FC1	245	62.50%
	FC2	62	15.82%
Typology	Multifamily	317	80.87%
	Single-family	75	19.13%
Share of windows	Mean		5.18
	Std. Dev.		2.04
Share of opaque envelope	Mean		25.07
	Std. Dev.		17.52

from 1980 onwards.

As a final caution, notice that we are not claiming that the variables that have been dropped in the selection procedure do not have any influence on airtightness. Their effect, as usual in multivariate statistical studies, may already be collected in the model by the variables that are already present in it. Table 6 contains the  $\beta_i$  and  $\tau_{ij}$  coefficients of the equation of the final GLM model appearing in Equation (4) that can be used for predicting airtightness, together with the significance level of each coefficient. As usual in many studies, the convention used here is that p-values between 0.10 and 0.05 showed weak significance, p-values between 0.05 and 0.01 showed strong significance, and p-values less than 0.01 show very strong evidence of significance. We can see that, for example, the log ( $n_{50}$ ) value of a building in its original state is estimated to increase by 0.15 over a retrofitted building in the case that the other variables are the same.

Fig. 5 contains the residual analysis for this final GLM model. The graph shows that the main hypotheses of the model (linearity and homoscedasticity) can be assumed since no curvature or other shape is observed in the graph. Moreover, a single observation studentized residual appears outside the [-3,3] interval, which is completely compatible with the absence of significant outliers in the model.

#### 5. Discussion

The GLM model allows the identification and analysis of factors with a significant impact on the level of airtightness. The influence of features such as the state of the dwelling or window permeability was evaluated through the model. While trends regarding these factors did not differ from what may be expected, the model additionally quantifies their effect. For example, from the model, it is estimated that the value of log  $n_{50}$  decreases by 0.150 if the dwelling is retrofitted and no other changes are made in the construction. Moreover, the analysis of other variables such as the climate zone and building systems (window material, shutter position or false ceiling) revealed the effect of different configurations on the level of airtightness of the building envelope. For example, it is estimated from the model that using wood as window material instead of PVC in a building increases  $\log n_{50}$  by 0.321 provided there are no other changes in the building.

Special attention should be paid to interactions. That is the case of the period of construction and typology, which present interaction with the share of opaque envelope. When analysing the main effect of the period of construction, the model showed better airtightness performance for dwellings in the group "before 1980" than for more recent buildings. A similar effect was also pointed out by McWilliams et al. [23] in the model they developed. It should be noted, though, that in this case the main effects should not be interpreted when interactions are present as they may lead to wrong conclusions. In this case, the interactions

#### Table 4

Descriptive study for the variable  $n_{50}$  and the final transformed response variable log ( $n_{50}$ ).

	Ν	Minimum	Maximum	Mean	Standard deviation	Lower quartile	Median	Upper quartile
n <sub>50</sub>	392	1.1930	39.4217	7.2238	4.2981	4.3371	6.2763	9.1672
Log (n <sub>50</sub> )	392	0.18	3.67	1.8291	0.5463	1.4672	1.8368	2.2156



Fig. 4. Histograms for the variable  $n_{50}$  and the final transformed response variable log  $(n_{50})$ .

#### Table 5

ANOVA table for the final GLM model showing the variability of the response explained by each of the variables and interaction included in the model and its statistical significance.

Source	Type III Sum of Squares	Degrees of freedom	Mean Square	F-value	p- value
Corrected model	44.899 <sup>a</sup>	24	1.871	9.565	.000
Intercept	23.171	1	23.171	118.470	.000
Retrofitting state	1.377	1	1.377	7.039	.008
Climate zone	10.363	6	1.727	8.831	.000
Period of construction	2.219	1	2.219	11.343	.001
Window permeability	4.940	3	1.647	8.420	.000
Window material	2.216	3	.739	3.778	.011
Shutter position	2.058	3	.686	3.507	.016
False ceiling	2.854	2	1.427	7.297	.001
Typology	1.518	1	1.518	7.762	.006
Share of windows	1.862	1	1.862	9.522	.002
Share of opaque envelope	0.340	1	.340	1.736	.188
Period of construction * Share of opaque envelope	2.468	1	2.468	12.620	.000
Typology * Share of opaque envelope	0.834	1	.834	4.262	.040
Error	71.779	367	.196		
Corrected Total	116.677	391			

<sup>a</sup>  $R^2 = .385$  (Adjusted  $R^2 = .345$ ).

indicate that the effect of the share of opaque envelope is significantly higher for buildings in the group "before 1980" than in the group "after 1980", and also for single-family dwellings with respect to multi-family ones.

Overall, the variables that build the model are in line with previously developed models addressed in Section 1.1, related to climate conditions, the age of the building, dimension-related characteristics, building systems type of building, and conservation state. It considers, in addition, variables that make reference to the singularities of the Spanish national built stock such as the effect of the position of rolling shutters, or the role of the share of the envelope to outdoors in the case of multifamily buildings.

The  $R^2$  value of this model is 0.385 so the model explains 38.5% of the variability of the response. While this value may seem not too high, it is in line with previous approaches as detailed in Table 1. It should be noticed that, as opposed to other studies with much higher  $R^2$  values [29], a sample representing faithfully the population of dwellings in very different zones of Spain is considered. Thus, the value is even higher than what could be expected. Notice also that, from the statistical point of view, the relevance of the model is warranted by the significance of the coefficients and by the validity of its residual analysis as explained before (Fig. 4).

In contrast, the analytical model proposed by regulations in Spain

Table 6
Equation of the final GLM predictive model for airtightness.

ParameterCoefficientParameterCoefficientIntercept $0.543^{**}$ Shutter position. P01 $0.142$ Retrofitting state. $0.150^{***}$ Shutter position. P02 $0.089$ Original $-0.255^{**}$ Retrofitting state. $0^a$ Shutter position. P03 $-0.255^{**}$ Retrofitted $-0.256^{***}$ Shutter position. P04 $0^a$ Climate zone. A3 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. C3 $0.563^{***}$ Typology. Single-family $0^a$ Climate zone. C3 $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of paque envelope $0.010^{***}$ Period of $0^a$ Period of construction. Before 1980 $0^a$ Period of $0^a$ Period of construction. Before 1980 * Share of opaque envelope $0^a$
Intercept $0.543^{**}$ Shutter position. P01 $0.142$ Retrofitting state. $0.150^{***}$ Shutter position. P02 $0.089$ Original         -         - $-0.255^{**}$ Retrofitting state. $0^{"0}$ Shutter position. P03 $-0.255^{**}$ Retrofitted         -         - $-0.255^{**}$ Climate zone. A3 $0.456^{***}$ Shutter position. P04 $0^{"0}$ Climate zone. B4 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^{"0}$ Climate zone. C3 $0.553^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. D2 $0.027$ Typology. Single-family $0^{"0}$ Climate zone. 03 $0^{"0}$ Share of opaque envelope $0.001$ construction.         Before 1980         -         - $980^{*}$ Share of opaque $0.010^{***}$ Period of $0^{"0}$ Period of construction. Before $0.010^{***}$
Retrofitting state. Original $0.150^{***}$ Shutter position. P02 $0.089$ Retrofitting state. Retrofitted $0^a$ Shutter position. P03 $-0.255^{**}$ Retrofitted $0^a$ Shutter position. P04 $0^a$ Climate zone. A3 $0.456^{***}$ Shutter position. P04 $0^a$ Climate zone. B4 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. C3 $0.553^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. D2 $0.027$ Typology. Single-family $0^a$ Climate zone. D3 $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction. $Before 1980$ Period of construction. Before $1980^{*}$ Share of opaque $0.010^{***}$ Period of $0^a$ Period of construction. Before $1980^{*}$ Share of opaque $0.010^{***}$ Window $0.591^{***}$ Period of construction. After $0^a$
OriginalIRetrofitting state. $0^a$ Shutter position. P03 $-0.255^{**}$ Retrofitted $-0.255^{**}$ Climate zone. A3 $0.456^{***}$ Shutter position. P04 $0^a$ Climate zone. B4 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. D2 $0.027$ Typology. Single-family $0^a$ Climate zone. D3 $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction. $Before 1980$ $Period of construction. Before0.010^{***}Period of0^aPeriod of construction. Before0.010^{***}1980envelopePeriod of construction. After0^a$
Retrofitting state. Retrofitted $0^a$ Shutter position. P03 $-0.255^{**}$ $-0.255^{**}$ Retrofitted-0.255^{**}Shutter position. P04 $0^a$ Climate zone. A3 $0.456^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. D2 $0.027$ Typology. Multifamily $0.444^{***}$ Climate zone. C3 $0.636^{***}$ Typology. Single-family $0^a$ Climate zone. $\alpha3$ $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction. Before 1980-Period of construction. Before 1980 * Share of opaque $0.010^{***}$ Pasoenvelope-980 * Share of opaque $0^a$ Window $0.591^{***}$ Period of construction. After $0^a$
Retrofitted         1           Climate zone. A3         0.456***         Shutter position. P04         0 <sup>a</sup> Climate zone. B4         0.651***         False ceiling. FC0         -0.260***           Climate zone. C1         0.311**         False ceiling. FC1         -0.252***           Climate zone. C2         0.646***         False ceiling. FC2         0 <sup>a</sup> Climate zone. C3         0.563***         Typology. Single-family         0 <sup>a</sup> Climate zone. D2         0.027         Typology. Single-family         0 <sup>a</sup> Climate zone. a3         0 <sup>a</sup> Share of windows         0.035***           Period of         -0.291***         Share of opaque envelope         0.001           construction.         Before 1980         Period of construction. Before         0.010***           Period of         0 <sup>a</sup> Period of construction. Before         0.010***           1980         envelope         Window         0.591***         Period of construction. After         0 <sup>a</sup>
Climate zone. A3 $0.456^{***}$ Shutter position. P04 $0^{a}$ Climate zone. B4 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^{a}$ Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. C3 $0.563^{***}$ Typology. Single-family $0^{a}$ Climate zone. C3 $0.027$ Typology. Single-family $0^{a}$ Climate zone. $\alpha$ $0^{a}$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction.         Before 1980         Period of construction. Before $0.010^{***}$ Period of $0^{a}$ Period of construction. Before $0.010^{***}$ 1980         share of opaque $0.010^{***}$ $1980^{*}$ Period of construction. After
Climate zone. B4 $0.651^{***}$ False ceiling. FC0 $-0.260^{***}$ Climate zone. C1 $0.311^{**}$ False ceiling. FC1 $-0.252^{***}$ Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0^a$ Climate zone. C3 $0.563^{***}$ Typology. Single-family $0^a$ Climate zone. C3 $0.027$ Typology. Single-family $0^a$ Climate zone. $\alpha$ $0^a$ Share of windows $0.035^{**x}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction.         Before 1980         Period of construction. Before $0.010^{***}$ Period of $0^a$ Period of construction. Before $0.010^{***}$ 1980         share of opaque $0.010^{***}$ $1980^*$ Share of opaque $0^a$
$\begin{array}{ccccc} \mbox{Climate zone. C1} & 0.311^{**} & \mbox{False ceiling. FC1} & -0.252^{***} \\ \mbox{Climate zone. C2} & 0.646^{***} & \mbox{False ceiling. FC2} & 0^a \\ \mbox{Climate zone. C3} & 0.563^{***} & \mbox{Typology. Multifamily} & 0.444^{***} \\ \mbox{Climate zone. D2} & 0.027 & \mbox{Typology. Single-family} & 0^a \\ \mbox{Climate zone. D2} & 0.027 & \mbox{Typology. Single-family} & 0^a \\ \mbox{Climate zone. C3} & 0^a & \mbox{Share of windows} & 0.035^{***} \\ \mbox{Period of} & -0.291^{***} & \mbox{Share of opaque envelope} & 0.001 \\ \mbox{construction.} & & & \\ \mbox{Before 1980} & & & \\ \mbox{Period of} & 0^a & \mbox{Period of construction. Before} & 0.010^{***} \\ \mbox{construction. Since} & & 1980 * \mbox{Share of opaque} \\ \mbox{1980} & & & \\ \mbox{envelope} & & \\ \mbox{Window} & 0.591^{***} & \mbox{Period of construction. After} & 0^a \end{array}$
Climate zone. C2 $0.646^{***}$ False ceiling. FC2 $0^a$ Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. D2 $0.027$ Typology. Single-family $0^a$ Climate zone. D2 $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction. $Before 1980$ Period of construction. Before $0.010^{***}$ Period of $0^a$ Period of construction. Before $0.010^{***}$ 1980 $1980^*$ Share of opaque $0.010^{***}$ Poindow $0.591^{***}$ Period of construction. After $0^a$
Climate zone. C3 $0.563^{***}$ Typology. Multifamily $0.444^{***}$ Climate zone. D2 $0.027$ Typology. Single-family $0^a$ Climate zone. $\alpha 3$ $0^a$ Share of windows $0.035^{***}$ Period of $-0.291^{***}$ Share of opaque envelope $0.001$ construction. $-0.291^{***}$ Share of construction. Before $0.001^{***}$ Period of $0^a$ Period of construction. Before $0.010^{***}$ Period of $0^a$ envelope $0.010^{***}$ 1980envelope $0^a$ envelopeWindow $0.591^{***}$ Period of construction. After $0^a$
$\begin{array}{ccccc} \mbox{Climate zone. D2} & 0.027 & Typology. Single-family & 0^a \\ \mbox{Climate zone. } \alpha 3 & 0^a & Share of windows & 0.035^{***} \\ \mbox{Period of} & -0.291^{***} & Share of opaque envelope & 0.001 \\ \mbox{construction.} & & & & & \\ \mbox{Before 1980} & & & & & \\ \mbox{Period of} & 0^a & Period of construction. Before & 0.010^{***} \\ \mbox{construction. Since} & 1980 * Share of opaque \\ \mbox{1980} & & & & \\ \mbox{envelope} & & & \\ \mbox{Window} & 0.591^{***} & Period of construction. After } & 0^a \end{array}$
Climate zone. α3     0 <sup>a</sup> Share of windows     0.035***       Period of     -0.291***     Share of opaque envelope     0.001       construction.     -     Share of opaque envelope     0.001       Before 1980     -     -     -       Period of     0 <sup>a</sup> Period of construction. Before     0.010***       construction. Since     1980 * Share of opaque     -     0.010***       1980     envelope     -     -       Window     0.591***     Period of construction. After     0 <sup>a</sup>
Period of construction.     -0.291***     Share of opaque envelope     0.001       construction.     Period of construction. Before     0.010***       Period of onstruction. Since     1980 * Share of opaque     0.010***       1980     envelope     envelope       Window     0.591***     Period of construction. After     0 <sup>a</sup>
construction.         Before 1980         Period of       0°         Period of construction. Before       0.010***         construction. Since       1980 * Share of opaque         1980       envelope         Window       0.591***    Period of construction. After
Before 1980     Period of construction. Before     0.010***       Position for the structure of the st
Period of     0 <sup>a</sup> Period of construction. Before     0.010***       construction. Since     1980 * Share of opaque     1980       1980     envelope     envelope       Window     0.591***     Period of construction. After     0 <sup>a</sup>
construction. Since     1980 * Share of opaque       1980     envelope       Window     0.591***       Period of construction. After     0 <sup>a</sup>
1980envelopeWindow0.591***Period of construction. After0a
Window 0.591*** Period of construction. After 0 <sup>a</sup>
permeability. Class 1980 * Share of opaque
0 or 1 envelope
Window 0.311*** Typology. Multifamily * Share -0.008**
permeability. Class of opaque envelope
Window 0.227** Typology. Single-tamily * 0"
permeability. Class Share of opaque envelope
3
permeability. Class
4 Window motorial 0.147
Steel
Mindow motorial 0.005
Aluminium
Window material 0 321***
Wood
Window material 0 <sup>a</sup>
PVC

\* stands for p-value  $\leq$  0.1, \*\* for p-value  $\leq$  0.05 and \*\*\* for p-value  $\leq$  0.01. <sup>a</sup> This parameter is set to 0 as it corresponds to the reference class of the

variable.

could only explain 4% of the variability, as explained before. It should be highlighted, though, that parallelism among variables of both models was identified:

- The effect of the variable "Retrofitting state" can be reflected in the variable *C*<sub>0</sub> of CTE model (airflow coefficient of the opaque part of the thermal envelope for new or existing buildings with/without improved airtightness).
- "Window permeability" makes reference to the same concept as  $C_h$  (permeability of doors and windows).



Fig. 5. Residual analysis for the final GLM model proposed.

- The effect of "Share of opaque envelope" can be identified in  $A_0$  (sum of areas of the opaque thermal building envelope with heat exchange with the outdoor air).

"Share of windows" can be associated to  $A_h$  (sum of the area of the doors and windows.

The improvement of the proposed model can be explained because it considers a wider range of variables including their main effects and some interactions, resulting in a more accurate model for the sample chosen. Therefore, we may conclude that, while the variables considered in current regulations are interesting, this study suggests that, if a closed formula for airtightness is to be established, other variables should also be taken into account to improve the accuracy of the predicted values.

However, the model involves some limitations. First, the sample size is limited because cases were selected proportionally to the existing built stock in Spain. This fact might lead to the underrepresentation of some dwelling types, adding uncertainty to the model. For instance, even though the main climate zones were represented, not all the local variants could be addressed.

Another limitation is the fact that some of the characterization variables addressed are not easy to evaluate from visual inspection in a standardised way. This is the case of window permeability, which, when no information was provided, was estimated by the technician. Also, the retrofitting state of the dwelling can involve different criteria and improvements can be carried out to a different extent leading to a diverse degree of impact on the airtightness of the envelope. Characterization data introduction into a database by different agents always introduces unknown bias. However, in order to avoid inconsistencies, a standardised protocol was developed, and the technicians were accordingly trained. Furthermore, all cases were reviewed and verified upon completion to limit bias [36].

Lastly, the wide scatter in the airtightness results obtained could be in part due to the workmanship effect. As other authors have previously claimed [21,29,31], workmanship and supervision seem to be parameters with a great impact on airtightness performance. Although previously addressed for specific contexts [24,26], challenges related to the thoroughness of workmanship and supervision still remain and are difficult to consider in predictive tools, thus hindering more consistent models [21,26,27,29,31,32].

# 6. Conclusions

The urge for more energy-efficient building stocks has raised awareness of the airtightness performance of the building envelope. The reduction of envelope permeability has become a priority in many countries, and this is reflected in EP standards.

In Spain, whole building airtightness requirements were recently introduced for the first time. This approach, although not too stringent, can be seen as a way to raise awareness and positive progress towards energy-efficient buildings. However, compliance can be justified through analytical estimation based on a calculation model, which was proved to be inaccurate.

In this context, a GLM to predict the envelope airtightness was presented based on real test results obtained from a representative sample of existing dwellings in Spain. The model considers several variables related to location, age of the building, building typology, state, building systems, and dimensions.

The methodology used to develop the model, although based on widespread strategies, offers added value regarding the origin of representative data, full characterization of the cases, standardised procedures, and the assessment of both quantitative and qualitative interactions.

Still, predictive models like the one proposed encounter some limitations such as representativeness, lack of data and predictive quality partially due to the great impact of workmanship. In spite of reasonable uncertainty, the model is robust, and it provides valuable knowledge regarding the airtightness of dwellings and the factors that most impact its performance. Therefore, it is intended as a useful tool during the design stage, to evaluate the impact of certain retrofitting actions, or as a method to provide realistic hypothetical input values in energy performance simulation tools.

Nevertheless, inherent limitations hinder the possibility of substituting on-site testing with the predictive model. Only on-site testing can provide accurate and reliable airtightness performance data, so models should not be seen as tools to use in certification processes. In this line, adequate design and careful workmanship and supervision seem to be some of the challenges to overcome in the near future in order to accomplish airtight buildings.

The development of the proposed model opens a window for remaining questions and other issues to be addressed by future research. First, the methodology proposed could involve an opportunity for other contexts, especially in countries with mandatory test performance to prove compliance with regulations, where huge and fast-growing airtightness datasets are available. In Spain, the growing interest in the energy performance of buildings seems to point towards more airtightness data of new and retrofitted dwellings in the near future, so that knowledge of current building practices may be approached. In addition, the applicability of the model in other contexts, especially in building stocks around the Mediterranean area, could be assessed. Lastly, the combination of different airtightness datasets draws the possibility of building a wider model, which could involve greater representativeness from the international perspective.

### CRediT authorship contribution statement

Irene Poza-Casado: Writing - review & editing, Writing - original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Pilar Rodríguezdel-Tío: Writing - review & editing, Writing - original draft, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Miguel Fernández-Temprano: Writing - review & editing, Writing - original draft, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Miguel-Ángel Padilla-Marcos: Software, Resources, Conceptualization. Alberto Meiss: Writing - original draft, Writing - original draft, Visualization, Validation, Supervision, Meth-Formal odology, Investigation, analysis, Data curation, Conceptualization.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

# Data availability

Data will be made available on request.

### Acknowledgements

The airtightness data assessed in this work was part of the study developed in the frame of INFILES Project "Energy impact of the airtightness level in residential buildings in Spain: Analysis and characterization of the infiltrations" funded by the Economy and Competitiveness Ministry of the Spanish Government (reference BIA2015-64321-R).

#### References

- [1] European Commission, COMMUNICATION from the COMMISSION to the EUROPEAN PARLIAMENT, the COUNCIL, the EUROPEAN ECONOMIC and SOCIAL COMMITTEE and the COMMITTEE of the REGIONS Stepping up Europe's 2030 Climate Ambition Investing in a Climate-Neutral Future for the Benefit of Our, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:5202 ODC0562.
- [2] European Parliament, Directive 2018/844 Amending Directive 2010/31/EU on the Energy Performance of Buildings and Directive 2012/27/EU on Energy Efficiency, 2018.
- [3] European Commission, Renovation Wave: the European Green Deal Factsheet, 2020, https://doi.org/10.2833/535670.
- [4] European Commission Directorate-General for Energy, Comprehensive Study of Building Energy Renovation Activities and the Uptake of Nearly Zero-Energy Buildings in the EU, 2019. Final report, https://data.europa.eu/doi/10.2 833/14675.
- [5] X. Zheng, E. Cooper, M. Gillott, C.J. Wood, A practical review of alternatives to the steady pressurisation method for determining building airtightness, Renew. Sustain. Energy Rev. 132 (2020), 110049, https://doi.org/10.1016/j. rser.2020.110049.
- [6] J. Jokisalo, J. Kurnitski, M. Korpi, T. Kalamees, J. Vinha, Building leakage, infiltration, and energy performance analyses for Finnish detached houses, Build. Environ. 44 (2009) 377–387, https://doi.org/10.1016/j.buildenv.2008.03.014.

- [7] R. Simson, T. Rebane, M. Kiil, M. Thalfeldt, J. Kurnitski, The impact of infiltration on heating systems dimensioning in Estonian climate, in: 12th Nord. Symp. Build. Phys, NSB 2020, 2020, https://doi.org/10.1051/e3sconf/202017205004.
- [8] S. Domínguez-Amarillo, J. Fernández-Agüera, M.Á. Campano, I. Acosta, Effect of airtightness on thermal loads in legacy low-income housing, Energies 12 (2019) 1–14, https://doi.org/10.3390/en12091677.
- [9] J. Huang, J. Hanford, F. Yang, Residential Heating and Cooling Loads Component Analysis, 1999. Berkeley, USA.
- [10] B. Jones, P. Das, Z. Chalabi, M. Davies, I. Hamilton, R. Lowe, A. Mavrogianni, D. Robinson, J. Taylor, Assessing uncertainty in housing stock infiltration rates and associated heat loss: English and UK case studies, Build. Environ. 92 (2015) 644–656, https://doi.org/10.1016/j.buildenv.2015.05.033.
- [11] T. Kalamees, Air tightness and air leakages of new lightweight single-family detached houses in Estonia, Build. Environ. 42 (2007) 2369–2377, https://doi.org/ 10.1016/j.buildenv.2006.06.001.
- [12] A. Meiss, J. Feijó-Muñoz, The energy impact of infiltration: a study on buildings located in north central Spain, Energy Effic 8 (2015) 51–64, https://doi.org/ 10.1007/s12053-014-9270-x.
- [13] I. Poza-Casado, A. Meiss, M.Á. Padilla-Marcos, J. Feijó-Muñoz, Airtightness and energy impact of air infiltration in residential buildings in Spain, Int. J. Vent. (2020), https://doi.org/10.1080/14733315.2020.1777029.
- [14] M.H. Sherman, R. Chan, Building Airtightness: Research and Practice, 2004 https://doi.org/LBNL-53356.
- [15] W.R. Chan, J. Joh, M.H. Sherman, Analysis of air leakage measurements of US houses, Energy Build. 66 (2013) 616–625, https://doi.org/10.1016/j. enbuild.2013.07.047.
- [16] B. Khemet, R. Richman, A univariate and multiple linear regression analysis on a national fan (de)Pressurization testing database to predict airtightness in houses, Build. Environ. 146 (2018) 88–97, https://doi.org/10.1016/j. buildenv.2018.09.030.
- [17] A.B. Mélois, B. Moujalled, G. Guyot, V. Leprince, Improving building envelope knowledge from analysis of 219,000 certified on-site air leakage measurements in France, Build. Environ. 159 (2019), 106145, https://doi.org/10.1016/j. buildenv.2019.05.023.
- [18] B. Cope, Statistics, analysis and conclusions from 250,000 blower door tests, including ventilation types, in: 38th AIVC Conf. - 6th TightVent Conf. - 4th Vent. Conf, Ventilating Heal. Low-Energy Build., 2017, pp. 160–167. https://www.aivc. org/resource/statistics-analysis-and-conclusions-250000-blower-door-tests-incl uding-ventilation-types.
- [19] J. Love, J. Wingfield, A.Z.P. Smith, P. Biddulph, T. Oreszczyn, R. Lowe, C.A. Elwell, 'Hitting the target and missing the point': analysis of air permeability data for new UK dwellings and what it reveals about the testing procedure, Energy Build. 155 (2017) 88–97, https://doi.org/10.1016/j.enbuild.2017.09.013.
  [20] M. De Strycker, L. Van Gelder, V. Leprince, Quality framework for airtightness
- [20] M. De Strycker, L. Van Gelder, V. Leprince, Quality framework for airtightness testing in the Flemish Region of Belgium - feedback after three years of experience, in: 39th AIVC Conf. Smart Vent. Build, 2018. Jua-les-Pins, France.
- [21] M. Prignon, G. Van Moeseke, Factors influencing airtightness and airtightness predictive models: a literature review, Energy Build. 146 (2017) 87–97, https:// doi.org/10.1016/j.enbuild.2017.04.062.
- [22] I. Poza-Casado, V.E.M. Cardoso, R.M.S.F. Almeida, A. Meiss, N.M.M. Ramos, M.Á. Padilla-Marcos, Residential buildings airtightness frameworks: a review on the main databases and setups in Europe and North America, Build, Environ. Times 183 (2020), 107221, https://doi.org/10.1016/j.buildenv.2020.107221.
- [23] J. Mcwilliams, M. Jung, Development of a Mathematical Air-Leakage Model from Measured Data, 2006 https://doi.org/LBNL-59041.
- [24] B. Khemet, R. Richman, An empirical approach to improving preconstruction airtightness estimates in light framed, detached homes in Canada, J. Build. Eng. 33 (2021), 101433, https://doi.org/10.1016/j.jobe.2020.101433.
- [25] C.N. Bramiana, A.G. Entrop, J.I.M. Halman, Relationships between building characteristics and airtightness of Dutch dwellings, Energy Proc. 96 (2016) 580–591, https://doi.org/10.1016/j.egypro.2016.09.103.
- [26] W. Pan, Relationships between air-tightness and its influencing factors of post-2006 new-build dwellings in the UK, Build, Environ. Times 45 (2010) 2387–2399, https://doi.org/10.1016/j.buildenv.2010.04.011.
- [27] H. Krstić, Ž. Koški, I.I. Otković, M. Španić, Application of neural networks in predicting airtightness of residential units, Energy Build. 84 (2014) 160–168, https://doi.org/10.1016/j.enbuild.2014.08.007.
- [28] M.I. Montoya, E. Pastor, F.R. Carrié, G. Guyot, E. Planas, Air leakage in Catalan dwellings: developing an airtightness model and leakage airflow predictions, Build, Environ. Times 45 (2010) 1458–1469, https://doi.org/10.1016/j. buildenv.2009.12.009.
- [29] J. Fernández-Agüera, S. Domínguez-Amarillo, J.J. Sendra, R. Suárez, An approach to modelling envelope airtightness in multi-family social housing in Mediterranean Europe based on the situation in Spain, Energy Build. 128 (2016) 236–253, https://doi.org/10.1016/j.enbuild.2016.06.074.
- [30] J. Fernández-Agüera, S. Domínguez-Amarillo, J.J. Sendra, R. Suarez, Predictive models for airtightness in social housing in a Mediterranean region, Sustain. Cities Soc. 51 (2019), 101695, https://doi.org/10.1016/j.scs.2019.101695.
- [31] M. Ibanez-Puy, J. Alonso, Airtightness in Spanish residential buildings. Case study, in: 2019 IEEE Int. Conf. Eng. Technol. Innov, IEEE, Valbonne Sophia-Antipolis, France, France, 2019, pp. 1–6, https://doi.org/10.1109/ICE.2019.8792809.
- [32] T.-O. Relander, S.B. Holøs, J.V. Thue, Airtightness estimation—a state of the art review and an en route upper limit evaluation principle to increase the chances that wood-frame houses with a vapour- and wind-barrier comply with the airtightness requirements, Energy Build. 54 (2012) 444–452, https://doi.org/ 10.1016/j.enbuild.2012.07.012.

- [33] V.E.M. Cardoso, M. Lurdes Simões, N.M.M. Ramos, R.M.S.F. Almeida, M. Almeida, J.N.D. Fernandes, A labelling strategy to define airtightness performance ranges of naturally ventilated dwellings: an application in southern Europe, Energy Build. (2022), 112266, https://doi.org/10.1016/j.enbuild.2022.112266.
- [34] S. Ferrari, V. Zanotto, Building Energy Performance Assessment in Southern Europe, Springer International Publishing, Cham, 2016, https://doi.org/10.1007/ 978-3-319-24136-4.
- [35] J. Feijó-Muñoz, A. Meiss, I. Poza-Casado, M.Á. Padilla-Marcos, M. Rabanillo-Herrero, A. Royuela-del-Val, M.J. Dios-Viéitez, V. Echarri-Iribarren, C. Pardal, V. J. Del Campo-Díaz, R.A. González-Lezcano, R. Assiego de Larriva, M. Montesdeoca Calderín, J. Fernández Agüera, Permeabilidad al aire de los edificios residenciales en España. Estudio y caracterización de sus infiltraciones, Ediciones Asimétricas, 2019.
- [36] J. Feijó-Muñoz, I. Poza-Casado, R.A. González-Lezcano, C. Pardal, V. Echarri-Iribarren, R. Assiego De Larriva, J. Fernández-Agüera, M.J. Dios-Viéitez, V.J. Del Campo-Díaz, M. Montesdeoca Calderín, M.Á. Padilla-Marcos, A. Meiss, Methodology for the study of the envelope airtightness of residential buildings in Spain: a case study, Energies 11 (2018) 704, https://doi.org/10.3390/ en11040704.
- [37] I. Poza-Casado, A. Meiss, M.Á. Padilla-Marcos, J. Feijó-Muñoz, Preliminary analysis results of Spanish residential air leakage database, in: 39th AIVC - 7th TightVent 5th Vent. Conf. "Smart Vent. Build, 2018. Antibes Juan-les-Pins, https ://www.aivc.org/resources/collection-publications/aivc-conference-proceedings -presentations.
- [38] I. Poza-Casado, Airtightness Performance of the Building Envelope of Dwellings in Spain. Characterisation and Energy Impact of Air Infiltration, Universidad de Valladolid, 2021. https://uvadoc.uva.es/handle/10324/52115.
- [39] J. Feijó-Muñoz, C. Pardal, V. Echarri-Iribarren, J. Fernández-Agüera, R. Assiego de Larriva, M. Montesdeoca Calderín, I. Poza-Casado, M.Á. Padilla-Marcos, A. Meiss, Energy impact of the air infiltration in residential buildings in the Mediterranean area of Spain and the Canary islands, Energy Build. 188–189 (2019) 226–238, https://doi.org/10.1016/j.enbuild.2019.02.023.
- [40] J. Feijó-Muñoz, R.A. González-Lezcano, I. Poza-Casado, M.Á. Padilla-Marcos, A. Meiss, Airtightness of residential buildings in the Continental area of Spain, Build. Environ. 148 (2019) 299–308, https://doi.org/10.1016/j. buildenv.2018.11.010.
- [41] Presidencia del Gobierno, Decreto 1490/1975, de 12 de junio, por el que se establecen medidas a adoptar en las edificaciones con objeto de reducir el consumo

de energía, 1975, pp. 15001–15003. https://www.boe.es/buscar/doc.php? id=BOE-A-1975-14782.

- [42] Ministerio de Obras Públicas y Urbanismo, Gobierno de España, Norma Básica de Edificación NBE-CT-79, Condiciones térmicas en los edificios, Spain, 1979. https:// boe.es/boe/dias/1979/10/22/pdfs/A24524-24550.pdf.
- [43] Ministerio de Vivienda, Gobierno de España, Código técnico de la Edificación (CTE), 2006. https://www.codigotecnico.org/pdf/Documentos/Parte1/DBAnteri ores/CTE\_PARTE1\_marzo2006.pdf.
- [44] Ministerio de Fomento, Gobierno de España, Código técnico de la Edificación (CTE). Documento básico HE 1: Limitación de la demanda energética, 2019. htt ps://www.codigotecnico.org/images/stories/pdf/ahorroEnergia/DBHE.pdf.
- [45] Aenor, UNE-EN ISO 9972: Prestaciones térmicas de los edificios. Determinación de la permeabilidad al aire de los edificios. Método de presurización con ventilador, 2019 (ISO 9972:2015).
- [46] S. Charrier, A. Bailly, F.R. Carrié, Building Airtightness in France: Regulatory Context, Control Procedures, Results, 2015, in: http://qualicheck-platform.eu/2 016/03/fact-sheet-07-building-airtightness-in-france-regulatory-context-controlprocedures-results/.
- [47] Department for Communities and Local Government, Welsh Government, Minimum Technical Competence for Air Tightness Testing in Buildings for England and Wales, 2018. https://www.gov.uk/guidance/competent-person-sch eme-current-schemes-and-how-schemes-are-authorised#current-schemes.
- [48] FliB, Prüfungsordnung des Fachverbandes Luftdichtheid im Bauwesen e. V. (FliB) zur Erlangung der Qualifikation "Zertifizierter Prüfer der Gebäude-Luftdichtheit im Sinne der Energieeinsparverordnung,", 2016.
- [49] ISO, ISO 9972, Thermal Performance of Buildings. Determination of Air Permeability of Buildings. Fan Pressurization Method (ISO 9972:2015), 2015. https://www.iso.org/standard/55718.html.
- [50] M.H. Sherman, A power-law formulation of laminar flow in short pipes, J. Fluids Eng. Trans. ASME. 114 (1992) 601–605, https://doi.org/10.1115/1.2910073.
- [51] Agencia Estatal de Meteorología (AEMET), Atlas climático ibérico (Iberian climate atlas), Ministerio de Medio Ambiente y Medio Rural y Marino de España, 2011. http://www.aemet.es/documentos/es/conocermas/publicaciones/Atlas-climatolo gico/Atlas.pdf.
- [52] Aenor, UNE-EN 12207:2017, Ventanas Y Puertas, Permeabilidad al aire. Clasificación, 2017.
- [53] IBM Corporation, IBM® SPSS® Statistics, 2019.