The application of survival analysis for service life prediction of building materials: a proof of concept

BUYLE Matthias[1, a *], BRAET Johan[2, 3, b] and AUDENAERT Amaryllis[1, 2, c]

[1] Energy and Materials in Infrastructure and Buildings (EMIB), Applied Engineering, University of Antwerp, Groenenborgerlaan 171 – B-2020 Antwerp, Belgium

[2] Department Engineering Management, Applied Economics, University of Antwerp, Prinsstraat 13, B-2000 Antwerp, Belgium

[3] Department of Product Development, Design sciences, University of Antwerp, Prinsstraat 13, 2000 Antwerp

[a]matthias.buyle@uantwerpen.be, [b]johan.braet@uantwerpen.be, [c]amaryllis.audenaert@uantwerpen.be

*corresponding author

Keywords: Service life prediction; Survival analysis; Life cycle assessment; buildings

Abstract.

Improved energy efficiency in new buildings due to stricter regulations results in the growing importance of the environmental impact of materials. Not only the initial construction has a significant contribution to the total impact, but replacements during the use phase as well. To date limited data are available on the service life of building components and the corresponding number of replacements. The data that are available are typically deterministic values. This study is a proof of concept on the application of survival analysis for service life prediction of building materials to compute stochastic distributions in order to reduce uncertainty in building related environmental studies.

A survey was conducted asking building residents for the replacement frequency of components such as wall and floor finishes, kitchen and bathroom furniture, etc. This resulted in 99 responses. A survival analysis was applied to calculate non-parametric Kaplan-Meier curves. These curves are fitted to obtain theoretical parameterized survival and hazard functions. The hazard functions feed a Monte Carlo simulation in order to obtain the number of replacements over a certain period of time. 1000 runs per component result in a stochastic distribution of the expected number of replacements, presented by their mean and a 90% confidence interval.

The analysis was performed per component and more subdivided according to the additional parameter 'room type' in the case of surface finishes (wall, floor and ceiling). This way more detailed information was obtained compared to the existing literature. The simulations indicate that in most cases, there are discrepancies between literature data and the computed results in this study.

Introduction

The growing environmental awareness of the last decades resulted in identifying the construction sector as one of the major targets for improvement. The building sector is responsible for nearly 40% of the global energy consumption, 30% of raw material use, 25% of solid waste production, 25% of water use, 12% of land use, and 33% of the related global greenhouse gas (GHG) emissions [1], [2]. This awareness resulted in Europe in energy regulations such as Energy Performance of Buildings Directive (EPBD) 2002/91/EC and the revised EPBD 2010/31/EU issued by the European Union [3], [4]. However, actions to improve energy efficiency of buildings require additional material, so the entire life cycle has to be taken into account to avoid just shifting the problem. Blengini and DiCarlo investigated a low-energy dwelling in Italy and found that although the operational energy was 10 times lower compared to the reference standard house, the total environmental impact was only

reduced by a factor two [5]. Additionally, when the level of insulation and energy efficiency increases, the share of material related impacts increases, both in relative and absolute terms. Buyle et al. analysed the influence of building type, level of insulation and different technical services in order to improve the environmental profile of Flemish dwellings [6]. Taking into account the current energy regulations, multiple non-hierarchic actions for improvement were relevant. A combination of a compact building design with one of the two following possible ways to reach a similar environmental optimum: firstly by following the current regulations for insulation complemented with the most efficient technical services, and secondly by an extensive reduction of energy losses - entailing a reduced (but not negligible) influence of the efficiency of technical services on the results. Himpe et al. came to similar conclusions for the Belgian situation by performing a life cycle energy assessment [7]. Another recent Belgian study by Stephan et al. pointed out that a passive house was preferable to a standard one, even if the embodied energy exceeds the operational energy in the passive scenario [8], [9]. All of these studies point out that, in particular for energy efficient buildings, the environmental impact related to materials is growing in importance.

Often materials are only taken into account at the initial construction of a building, however environmental impacts related to maintenance and the replacements of building components can contribute up to 50% of the initial construction [6]. This is in particular valid for components with a relative short service life and easily to replace e.g., paint. In current environmental studies, typically deterministic values from literature are applied. For example Vissering et al. developed a method to estimate the life expectancy of building components, resulting in one reference value per component [10]. A more detailed guideline developed by Building Cost Information Service (BCIS) estimates the life expectancies of common building components based on the experience of building experts [11]. For each component, a mean, average, median minimum and median maximum value is presented. Engineering approaches exist as well, starting from a reference value which is adjusted depending on case specific parameters. The factor method is an example where factors such as design level, outdoor environment and in-use conditions are taken into account to calculate an estimated service life based on a general reference service life [12], [13]. Van Nunen tried to improve the factor method by adding more criteria and by replacing deterministic values by distributions. However, only a conceptual framework was presented, but no data [14]. Limited research is available applying a bottom-up approach, focussing on data collection by surveying residents instead of building experts.

In this context, the goal of this research is to develop a method to compute stochastic distributions of the service life of building components based on survey data. Accordingly the number of replacements can be predicted in order to gain insight in the magnitude of the overall uncertainty in building related environmental studies. As a starting point, there was focused on components with a short reference service life which could easily be replaced. The main reason for this selection was the assumption that aesthetical criteria could play an important role as well, making the technical service life less relevant for those materials. To process the results of the survey, the methodology of survival analysis was applied to calculate non-parametric Kaplan-Meier curves. These curves were fitted to obtain theoretical parameterized survival and hazard functions. The hazard functions feed a Monte Carlo simulation in order to obtain the number of replacements over a certain period of time. The analysis was performed per component and more subdivided per room type in the case of surface finishes. The results were represented by their mean and 90% confidence interval and compared with the benchmark values of BCIS. This research is a proof of concept, so developing an exhaustive dataset for all building components is beyond its scope.

Methods

The developed model consists of three distinct steps: data collection (1), processing the data to compute theoretical parameterized survival and hazard curves (2) and finally running a Monte Carlo analysis to estimate the number of replacements over the total life span of a building (3). In line with other Belgian research, the total life span was assumed to be 60 years [15], [16]. A sensitivity analysis

was performed by extending the buildings life span to 120 year, but new insights were created so these results are not presented in the paper. The included building components are windows, floor finishes, wall finishes, ceiling finishes and fixed furniture. These components were subdivided per type of material in the case of windows and per room type for the rest as additional criteria. The included building components and their characteristics are listed in Table 1.

Table 1. Included building components and their characteristics

component	Material properties	Additional properties
windows	PVC frame wooden frame aluminium frame	-
floor finishes	Wood (laminate, parquet, …) soft and resilient (carpet, vinyl, PVC, …) ceramic tiles	living room bedroom(s) bathroom kitchen
ceiling finishes - paint	paint decorative covering (wood, plastic, …)	living room bedroom(s) bathroom kitchen
Wall finishes - paint	paint wallpaper ceramic tiles	living room bedroom(s) bathroom kitchen
fixed furniture	bathroom kitchen	-

Data collection. The purpose of this step was to collect data on the actual service life of the selected building components. Not only the service life was part of the survey, but more detailed information was collected as well: which were the actual materials before and after an intervention, in which room the intervention took place, which interventions were combined, etc. After a pilot survey was decided not only to ask for data on replacements that had actually taken place, but the respondents had to make an estimation on the time left to the next replacement as well. An online survey was selected as an easiest and relevant tool to reach individual residents. Potential respondents were selected in the region of Antwerp, Belgium and contacted by a (physical) letter, e-mail or social media. This resulted in about 110 responses, out of which 99 were relevant. The characteristics of the respondents are presented in Table 2. When the characteristics of the sample are compared with the Flemish population, it is clear the sample is not representative for multiple criteria such as age and sex. However, since the goal of the study was to verify the feasibility of a model, there was opted to continue with the presented dataset and improve the input data in future research.

Table 2. Summary of the characteristics of the respondents, compared with Flemish population

characteristics respondents		samples	share in sample	Flemish population [17], [18]
Sex				
	Male	39	39,4%	49,3%
	Female	60	60,6%	50,7%
Age				
	< 20	1	1,0%	22,1%
	20-25	38	38,4%	5,8%
	26-30	17	17,2%	6,3%
	31-35	11	11,1%	6,1%
	36-40	7	7,1%	7,0%
	41-45	6	6,1%	7,8%
	46-50	3	3,0%	7,8%
	51-55	5	5,1%	7,1%
	56-60	6	6,1%	6,4%
	61-65	2	2,0%	5,8%
	66-70	1	1,0%	4,6%
	71-75	1	1,0%	4,6%
	76-80	1	1,0%	4,0%
	> 80	0	0,0%	4,7%
level of education				
	secondary education	19	19,2%	54,7%
	higher education - non-University	38	38,4%	12,3%
	higher education - University	41	41,4%	12,8%
	others	1	1,0%	20,2%
family structure				
	single	20	20,2%	24,6%
	single parent	4	4,0%	7,0%
	couple with (resident) children	32	32,3%	34,4%
	couple without (resident) children	38	38,4%	32,7%
	reconstituted family	2	2,0%	-
	others	3	3,0%	1,4%
Ownership status				
	owner	68	68,7%	74,4%
	renter	30	30,3%	24,1%
	others	1	1,0%	1,5%

Theoretical parameterized survival and hazard curves. The results of the survey were processed by applying a survival analysis. This is a branch of statistics which is often applied in a biomedical context with as main interest the time until a certain event occurs, typically death [19]. Despite the medical origin of the method, it has been applied in an engineering context as well. Haldi et al. developed a model to simulate user behaviour related to window openings by office occupants [20], [21]. The event of interest in this case was the opening and closure of a window. Long term observations were available to validate the model. Other models were tested as well in this study, but survival analysis, combined with Markov chains to model the arrival and departure of the occupants, turned out to be the most reliable approach.

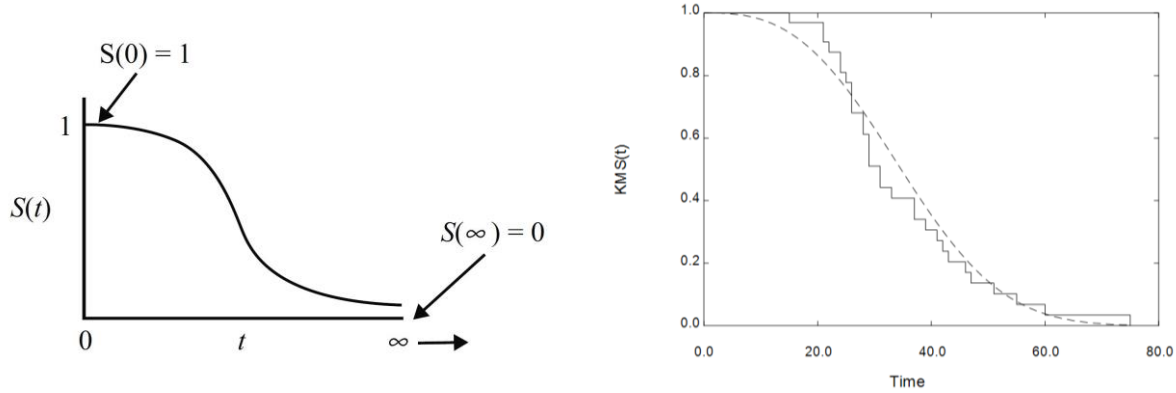
One of the strengths of survival analysis is the possibility to include observations containing only partial information such as survival for at least a certain period of time. Such non-comprehensive observations are called censored data and even if the event of interest was not observed during the study period, such observations still contain relevant information. In this research only right censored observations occur. This means that the status at the start of the observation period is known (the moment of a previous replacement), but the next replacement is not observed when the study period has ended. The only thing which is known is that at that time, the replacement did not take place yet. Theoretically, a survival analysis is characterized by two functions, the survivor function S(t) and the hazard function h(t) (see Eq. 1, Eq. 2 and Fig. 1a). The survivor function S(t) gives the probability that a subject survives longer than some specified time t. The hazard function h(t) gives the instantaneous potential per unit time for the event to occur, given calculated that the subject has survived up to time t [19]. If one of the two functions is known, the other can be as well. In reality, the theoretical functions are not known and a non-parametric Kaplan–Meier (KM) estimator is mostly

applied to analyse the data (Eq. 3, Fig. 1b). KM curves are typically step functions. However, in order to carry out further calculations, a continuous parametric function was needed similar to the theoretical functions. Therefore the Kaplan-Meier curves were fitted using a maximum likelihood regression. The distribution used in this study is the Weibull distribution. Previous research identified the this as the most relevant distribution, so no other distributions are taken into account [20].

$$S(t) = \exp[-\int_0^t h(u)du] \qquad \text{(Eq. 1)}$$

$$h(t) = [\frac{dS(t)/dt}{S(t)}] \qquad \text{(Eq. 2)}$$

$$\hat{S}(t_{(f-1)}) = \prod_{i=1}^{f-1} \hat{P}r(T > t_{(i)} | T \geq t_{(i)}) \qquad \text{(Eq. 3)}$$



a) **b)**

Fig. 1. a) theoretical survival curve [19] **b) Kaplan-Meier Curve of window replacements with PVC frame (step function) with fitted Weibull distribution (dashed line)**

The survey described in the previous step resulted in two lists. The first one included only historical data, so without the estimations. Most of the data is right censored, since not many components have been replaced twice and the exact service life is therefore unknown. The second one included the total estimated service life, so in this case all data is uncensored. The disadvantage of the historical data is that for most of the components included in the survey, not enough replacements were observed to come to relevant results. Therefore all simulations have been based on the estimated data.

Model and Monte Carlo simulations. To estimate the service life of a building component and the number of replacements over a certain period of time, a model has been developed in Matlab to create virtual timelines. The time interval is one year and each year a (memoryless) random number is generated to evaluate if a component is replaced or not. The criteria whether a replacement takes place or not at time t + 1 after the survival up to time t, is based on the conditional failure rate of the hazard function at time t. If a replacement took place, the timeline is reset. If not, a new random number is generated to evaluate t + 2. This procedure is repeated until the end of the observed period is reached and the number of replacements can be counted. To process the results statistically, a Monte Carlo simulation has been performed with 1000 independent runs of this timeline. The average and a 90% confidence interval have been calculated. The boundaries of the confidence interval are rounded up to the next integer, since it is assumed replacements cannot be subdivided. At the end, the results were compared with the reference values of BCIS.

Results and discussion

After processing the surveys, 7 aggregated and 34 detailed datasets were composed. The first category focusses on building level, without taking into account the additional properties. The latter is subdivided according to the additional properties, namely the specific room of the replacements. The results are presented in Table 3. The model calculates in the first place the number of

replacements. The same exercise could be done for the service life as well, but in the case of life cycle studies the replacements are more important so only this figures are presented. As can be seen, there is a wide range in sample size influencing the accuracy of the curve fitting. In particular the components with a short reference service life have a relative high sample size.

Table 3. Estimated service life and number of replacements of the selected building components

component	Properties component	sample size	replacements BCIS	Calculated replacements		
			average	average	lower limit	upper limit
windows						
	PVC frame	60	1,88	1,32	1	2
	wooden frame	20	1,82	2,23	1	4
	Aluminium frame	25	1,54	2,33	1	4
Floor finishes - wood						
	All rooms	135		2,78	1	5
	living room	54		2,37	1	4
	bedroom(s)	64	1,58	3,02	1	5
	bathroom	6		4,50	3	6
	kitchen	11		2,18	2	3
Floor finishes - soft and resilient						
	All rooms	37		2,05	0	4
	living room	3		-	-	-
	bedroom(s)	17	3,75	1,49	0	3
	bathroom	6		2,03	1	4
	kitchen	11		4,83	2	8
Floor finishes - tiles						
	All rooms	188		1,52	1	3
	living room	31		0,92	0	1
	bedroom(s)	10	1,20	-	-	-
	bathroom	76		1,83	1	3
	kitchen	71		1,50	0	3
Ceiling finishes - paint						
	All rooms	362		3,82	2	6
	living room	96		4,28	2	7
	bedroom(s)	92	8,57	3,83	2	6
	bathroom	82		3,54	1	6
	kitchen	92		3,68	2	6
Ceiling finishes - decorative covering						
	All rooms	45		1,41	0	3
	living room	13		1,18	0	3
	bedroom(s)	9	1,13	1,46	0	3
	bathroom	14		1,58	1	3
	kitchen	9		1,67	0	3
Wall finishes - paint						
	All rooms	244		5,31	3	8
	living room	81		5,69	3	9
	bedroom(s)	79	8,57	5,83	4	8
	bathroom	19		7,18	6	9
	kitchen	65		4,12	2	7
Wall finishes - wallpaper						
	All rooms	106		6,03	4	9
	living room	46		6,79	4	10
	bedroom(s)	38	7,50	5,36	3	8
	bathroom	0		-	-	-
	kitchen	22		5,49	3	8
Wall finishes - tiles						
	All rooms	86		1,90	1	3
	living room	0		-	-	-
	bedroom(s)	1	2,22	-	-	-
	bathroom	69		1,96	1	3
	kitchen	16		1,54	1	3
fixed furniture						
	bathroom	95	1,71	2,36	1	4
	kitchen	93	3,33	1,82	1	3

The results are very heterogeneous making it hard to draw general conclusions. In addition, the variance is high, resulting in a very wide diameter of the 90% confidence interval. Only for painted ceilings and kitchen furniture there is a significant difference between the calculated replacements and the BCIS results (paint: 8,57 replacements vs. 3,82 at building level; kitchen: 3,33 vs. 1,82). Nonetheless some relevant observations can be made. Basically all wall finishes are replaced less according to the calculations compared to BCIS, while for the other components the results are more related to the room type. The indoor finishes are subdivided per room type, which means the technical properties of the used materials are identical. Apparently other motivations affect the results to a big extent, in some cases up to a factor 2 (wall finishes – paint: bedroom vs. the rest) or 3 (floor finishes - soft and resilient: bedroom vs. kitchen). Possible motivations are economic or aesthetic reasons. Also the combination of interventions might shorten or extend the real service life of components.

The most important result of this research is the observation that there is a wide range in the number of replacements and that they are not uniform for different parameters of a building, such as room type for example. So one single deterministic value per building component is unsuitable for a detailed prediction of the number of replacements. Of course the presented results should be interpreted with care. A major limitation of this research is the relative small sample size. Before the model can be used to provide robust results, the survey data should be extended and improved. In the first place, the sample should be representative for the (in this case Flemish) population. A second limitation regarding the data quality is the limited amount of observed replacements. Since the presented results are mainly based on the predictions of the respondents, these are too much a reflection of the surveyors' opinions instead of on objective observations. In future research, more data should be collected to address these two limitations.

On the other hand, the simulated results emphasize the importance of tuning the results more specific to their field of application. Besides improving data quality, there are some other important research opportunities as well. In this study, only room type has been analysed as an additional parameter. However, other criteria could be included as well. User profiles could help to gain insight which groups tend to postpone or accelerate interventions. Criteria such as age, ownership status, and level of education or family status could result in different replacement rates. Also building related criteria such as building type and size could be analysed more in detail. Such results could assist to develop more customized guidelines for reducing environmental impacts. As example, industrial, flexible and demountable (IFD) buildings typically have a higher initial impact, but the benefits are due to the lower impacts of future replacements. The user profiles could identify for which groups this structural concept could be environmentally beneficial and which groups fit better with traditional non-flexible concepts.

A second research opportunity lies in capturing the cross impact of multiple interventions, which could have a big influence as well. At this point, the building components are treated as independent entities, but the model is able to include interference between components. It is possible to compute virtual time lines at building level as well. If a component is replaced, it can be assumed this will increase the probability other components will be replaced at the same time i.e., painting walls after installing a new kitchen. Or interventions can be postponed after the technical service life of a component is reached when a combined intervention is scheduled. The main issue in this case is the data collection, since it is hard to quantify to what extent decisions are mutually influencing each other. Which is the cause and which the effect? An attempt was undertaken to include such effects, but the survey turned out to be not detailed enough to capture these effects properly.

Conclusion

Over the last years, the energy efficiency of buildings has greatly improved. In environmental studies taken into account the entire life cycle of the building, this typically results in an increasing importance of material related impacts. These impacts take place both at the initial construction and

during the use phase when maintenance and replacements take place. The key factor for estimating the number of replacements and corresponding impacts is the service life prediction of the building components. In current research often single deterministic values are applied without taken into account the corresponding uncertainties. In this study a model is proposed to compute a distribution for the replacement of components, based on survey data. The data are analysed by applying a survival analysis and the results are fitted to obtain parameterized hazard curves. This curves feed a Monte Carlo model to calculate a set of 1000 timelines. For each run, the number of replacements were counted and for the total result, the mean and 90% confidence interval were presented.

Despite the limitations of the survey data, the results point out multiple clear discrepancies between the calculated and the BCIS reference values. So even without taken more detailed characteristics into account, the model indicates the need for more detailed information regarding the estimated service life. When the results are analysed more in detail by subdividing the sample based on the room type, the results are in many cases even more heterogeneous. This research is a proof of concept rather than an exhaustive list of service life distributions and the concept could be extended to the lifespan of the entire building as well. More research is needed to come to robust results with more complex models (e.g. autoregressive moving average), however the model itself is able to increase insight in the magnitude of the uncertainties related to service life prediction of building components.

References

[1] United Nations Environmental Programme and Sustainable Buildings and Climate Initiative, "Buildings and climate change: a summary for decision-makers," Paris, 2009.

[2] United Nations Environmental Programme and Sustainable Buildings and Climate Initiative, "Common carbon metric for measuring energy use and Reporting Greenhouse Gas Emissions from building operations," Paris, 2010.

[3] European Parliament, "Directive 2002/91/EC of the European Parliament and of the Council on the energy performance of buildings," *Off. J. Eur. Union*, vol. 4.1.2003, no. L 1, pp. 65–71, 2002.

[4] European Parliament, "Directive 2010/31/EU of the European Parliament and of the Council on the energy performance of buildings (recast)," *Off. J. Eur. Union*, vol. 18.6.2010, no. L 153, pp. 13–35, 2010.

[5] G. A. Blengini and T. Di Carlo, "The changing role of life cycle phases, subsystems and materials in the LCA of low energy buildings," *Energy Build.*, vol. 42, pp. 869–880, 2009.

[6] M. Buyle, A. Audenaert, J. Braet, and W. Debacker, "Towards a More Sustainable Building Stock: Optimizing a Flemish Dwelling Using a Life Cycle Approach," *Buildings*, vol. 5, no. 2, pp. 424–448, 2015.

[7] E. Himpe, L. Trappers, W. Debacker, M. Delghust, J. Laverge, A. Janssens, J. Moens, and M. Van Holm, "Life cycle energy analysis of a zero-energy house," *Build. Res. Inf.*, vol. 41, no. June, pp. 435–449, 2013.

[8] A. Stephan, R. H. Crawford, and K. de Myttenaere, "A comprehensive assessment of the life cycle energy demand of passive houses," *Appl. Energy*, vol. 112, pp. 23–34, 2013.

[9] A. Stephan, R. H. Crawford, and K. De Myttenaere, "Towards a comprehensive life cycle energy analysis framework for residential buildings," *Energy Build.*, vol. 55, pp. 592–600, 2012.

[10] C. L. Vissering, R. Janssen, M. A. A. M. Liebregts, A. Staub, and H. van Nunen, *Levensduur van bouwproducten - methode voor referentiewaarden [Life expectancy of building components - method for reference values]*. Rotterdam: SBR, 2011.

[11] BCIS, *Life Expectancy of Building Components. Surveyors' experiences of buildings in use. A practical guide*, 2nd ed. London: BCIS, 2006.

[12] P. J. Hovde, "The Factor Method For Service Life Prediction From Theoretical Evaluation To Practical Implementation," in *The 9th International Conference on Durability of Building Materials and Components (9DBMC), 17-21 march 2002,Brisbane, Australia*, 2002, pp. 1–

10.
[13] International Organization for Standardization, *ISO 15686-8:2008 Buildings and constructed assets -- Service-life planning -- Part 8: Reference service life and service-life estimation*. 2008, p. 36.
[14] H. van Nunen, *Assessment of the Sustainability of Flexible building – The Improved Factor Method: service lift prediction of buildings in the Netherlands, applied to life cycle assessment*. BOXTEL: Uitgeverij Æneas, 2010.
[15] W. Debacker, K. Allacker, F. De Troyer, A. Janssen, L. Delem, K. Peeters, L. De Nocker, C. Spirinckx, and J. Van Dessel, "Environmental Profile of Building elements," Mechelen, Belgium, 2013.
[16] K. Allacker, "Environmental and economic optimisation of the floor on grade in residential buildings," *Int. J. Life Cycle Assess.*, vol. 17, no. 6, pp. 813–827, Mar. 2012.
[17] Belgian Federal Government, "http://statbel.fgov.be," 2016. [Online]. Available: http://statbel.fgov.be. [Accessed: 09-Sep-2016].
[18] K. Heylen, M. Le Roy, S. Vanden Broucke, B. Vandekerckhove, and S. Winters, "Wonen in Vlaanderen: De resultaten van de woonsurvey 2005 en de uitwendige Woonschouwing 2005 [Living in Flanders: The results of the housing survey 2005 and the exterior housing inspection 2005]," Brussels, Belgium, 2007.
[19] D. G. Kleinbaum and M. Klein, *Survival Analysis. A Self-Learning Text*, Third., vol. 53, no. 9. New York, USA: Springer, 2013.
[20] F. Haldi and D. Robinson, "Interactions with window openings by office occupants," *Build. Environ.*, vol. 44, no. 12, pp. 2378–2395, 2009.
[21] M. Schweiker, F. Haldi, M. Shukuya, and D. Robinson, "Verification of stochastic models of window opening behaviour for residential buildings," *J. Build. Perform. Simul.*, vol. 5, pp. 55–74, 2012.