Introduction

The field of construction management oversees the whole process of design, construction, and maintenance of buildings and infrastructure. It has been well established that quality failures and the occurrence of errors in the design and construction has become an important feature of the procurement process in construction and systematically leads to schedule delays, time overruns, unexpected costs, rework, litigation costs, and other intangible costs of poor quality (Love and Li, 2010; Love et al., 2010). Quality control framework strategies aim to reduce the likelihood for unexpected costs in the design and building process. For example, in 1990 a compulsory Danish Building Defects Fund was introduced in Denmark, which reduced the number of construction defects by 30 up to 40% (de Freitas, 2013). Errors can be differentiated into design errors, construction errors, or e.g. by material and product defaults, and originate from individual or managerial errors, or stem from wider external influences (Atkinson, 2010). Significant research efforts have been conducted to study how errors arise, which causes can be found for reworks, and how quality control can help to reduce the frequency. However, Love et al. (2016) question the theoretical underpinning and highlight conceptual flaws in rework causation. Furthermore, they conclude that – regardless of decades of research on rework causation - rework remains a prevailing and chronic problem. Hammarlund and Josephson (1991) list defective workmanship, defects in products, and insufficient work separation as the top three causes for failure costs, which highlights the relevance to study the technical quality of building products and components to reduce excess failure costs. Quality control framework approaches should address both the causal factors for active failures, and as well, assess the effectiveness of design reviews to catch defects and errors in time (Minato, 2010). Evidently, the associated appraisal costs of the quality control frameworks should be compensated by a larger reduction in failure costs. In order to reduce the costs associated with rework, premature deterioration, building defects, and litigation, it is important to study the occurrence of these issues in a systematic qualitative and quantitative way. The Construction Industry Development Agency in Australia has estimated the direct cost of rework in the construction sector to be 10% or higher (CIDA, 1995). The Building Research Establishment found that in the UK about half of the errors in buildings could be traced back to the design process, whereas 40% originated from the construction process (BRE, 1981).

Damage databases

Databases of building defects have proven to be useful for drawing conclusions on underlying causes of building defects and for identifying potential improvement actions to reduce the occurrence of building defects. The analysis of databases of building defects allows to perform risk analyses considering the relative frequency of damage cases, as well as the (financial) consequences of typical problems. Furthermore, this allows to:

- Pinpoint relevant research areas to improve existing products
- Highlight the need for additional training, technical guidelines, building codes and quality control measures in specific areas
- Quantify typical costs associated with design and construction errors
- Assess the impact and effectiveness of enhanced quality control measures
- Provide additional insights to assess the expected service life of materials and components

Although no two buildings are realized in the same way, a systematic approach and the collection of empirical data, can connect the different cases (de Freitas 2013). In different countries initiatives have been taken to create and analyse databases on building defects (Carretero-Ayuso et al. 2017; Duncan and Ward 2017). Forcada et al. (2014) analysed 3647 construction defects in Spain, and found that the stability of the structure and inappropriate installation of roofs and facades were the most common defects, and mainly caused by poor workmanship. In contrast, in Norway it was found that two thirds of all defects were found in the building envelope and mainly relate to water ingress (de Freitas, 2013). In France, the AQC (l'Agence de la Qualité de la Construction) launched Sycodés (SYstème de Collecte des DESordres) and created statistical reports on building defects in the building industry using data from insurance companies. Roof related building defects make up the largest group of defects, followed by interior flooring. However, when the repair cost is considered, internal flooring is by far the most important type of defect (AQC, 2016). Throughout different countries, costs related to technical construction errors range between 5 and 10% of the overall building cost (de Freitas, 2013).

Furthermore, databases on service life prediction and reference service life can be considered either as counterpart or complementary to the database on defects. The most elaborate reference constitutes the work by Daniotti et al. (2008). This reference service life database has been developed to collect a series of grids in which data is stored and indexed, comprising the duration in years, the failure mode, and the selection of the several levels of factors to be used in factor methods.

Factorial methods using regression analysis, and more recently statistical black-box approaches based on neural networks, have been introduced in the field of service life prediction by a categorical analysis of large sets of similar buildings by assessing the condition of the building materials and components (e.g. Chew et al, 2010; Shohet and

Paciuk, 2010; Gaspar and de Brito, 2007). Typically, external factors such as climate, orientation, surroundings, height, and distance to the coast are studied in these models. Based on a multi-factor approach and regression analysis these studies aim to quantify the remaining service life and required maintenance actions. It should be noted that some studies omit those cases where defects can be attributed directly to obvious design or construction errors, to ensure that the analysis is (only) valid for a correct installation. Next to that, these studies typically focus on less recent buildings in order to highlight the impact of external parameters on material degradation. For example Pereira et al. (2020) studied 52 buildings between 8 and 66 years old, and used multiple regression analysis to predict the urgency of repair. A study by Bordalo et al. (2010) on ceramic tiling systems in Lisbon concluded that 42% of the variance in degradation could be explained by environmental agents, whereas 58% was attributed to design and installation errors. Research by Hammarlund and Josephson (1991) in Sweden indicates that quality failures after completion of a project amounts to 4% of the construction cost, of which 51% of the costs are design related, 26% due to installation errors, and 10% to material failure. Furthermore, these studies typically limit the external parameters to climatological conditions and do not differentiate in quality of materials, workmanship, or other external parameters that may affect the occurrence of construction problems. Perhaps regional differences in the occurrence of material degradation and defects may be caused by differences in construction quality, complexity of construction and building envelope interfaces in urban areas, or even the variability in construction cost and average wages in different areas. These external parameters might challenge a reliable statistical analysis of building defects that only considers materials, construction, and climate, rendering them prone to unquantified bias-effects.

Also in Belgium, an immense amount of information is available concerning building defects. Details about the causes, the damage caused by the defects, the associated repair cost, the liability, etc. can be found in reports of insurance companies, lawyer offices and consultants, since these are appointed to deal with damage claims.

Statistical analysis of these data may reveal unidentified patterns and provide insights into the dominant defects in terms of building component, type of damage, cause, time of occurrence, geographical location, building exposure, cost, and the like. For property developers, architects and construction managers, the results of the analysis may provide a series of lessons on the critical aspects in order to minimize building defects. The insights derived from the analysis can be used by building professionals, leading to innovative ideas regarding materials, design concepts, construction methods, and quality control systems. As a consequence, the quality of the buildings can be enhanced. Higher quality implies better durability, longer service life, lower life cycle cost, lower waste production, and higher sustainability.

Methods

General methodology

The analysis is based on the information that is available in a structured way in the database of the insurance company (see section Database of building defects). Based upon the postal code of the affected building, the information in the database was enriched with statistical information of the municipality in which the building is located. The enriched information is either demographical, geographical, climatological, or related to the typical building characteristics in the municipality.

For each of these external factors on municipality level, this study investigates to which extent they impact the relative occurrence of the different building damage categories. Distinction was made between the damage categories moisture problems, stability problems, neighbour damage problems, HVAC problems, acoustics problems and energy efficiency problems.

Database of building defects

The primary data source for this study is the claims database from an insurance company which is a specialized in supporting and protecting architects and engineers that have a high level of professional liability in the conceptual and control phase of the building process. Note that in Belgium there is a compulsory professional indemnity insurance (professional civil liability insurance) for architects, there is no widespread building defects cover, and there is a decennial liability for structural solidity defects and major pathologies. Please refer to (Elios, 2010) for an overview of national liability and insurance systems in the EU member states. It must be considered that unconscious selection bias is present, as the insurance company mainly focuses on the insurance of the professional liability of architects and contractors. Hence, failures in which an insurance company is not involved (e.g. contractor makes a mistake and repairs it) are not covered here. By consequence, the analysis does not provide insights into the complete overall fault cost or construction failures. Nonetheless, this database is a relevant sample of the real ‘population’ of construction failures, as it can be assumed that the more complex cases are covered, and as well the important failures in terms of associated costs. The insurance company is the market leader in insurance of building professionals with a share of about 60 to 65% of the architects in Belgium. The database contains claims entered from 1991 till 2019. Figure 1 shows for each municipality the number of claims in the database, divided by the number of building permits granted in the period 1996-2018. The figure shows that this ratio is not uniform across the Belgian territory. This does not necessarily imply that regional differences (such as topography, climate, traditions and regulation) impact the probability of occurrence of building defects. As the market share of the insurance company is not uniform across Belgium, the variance in the number of claims can also be the result of the variance in market share.

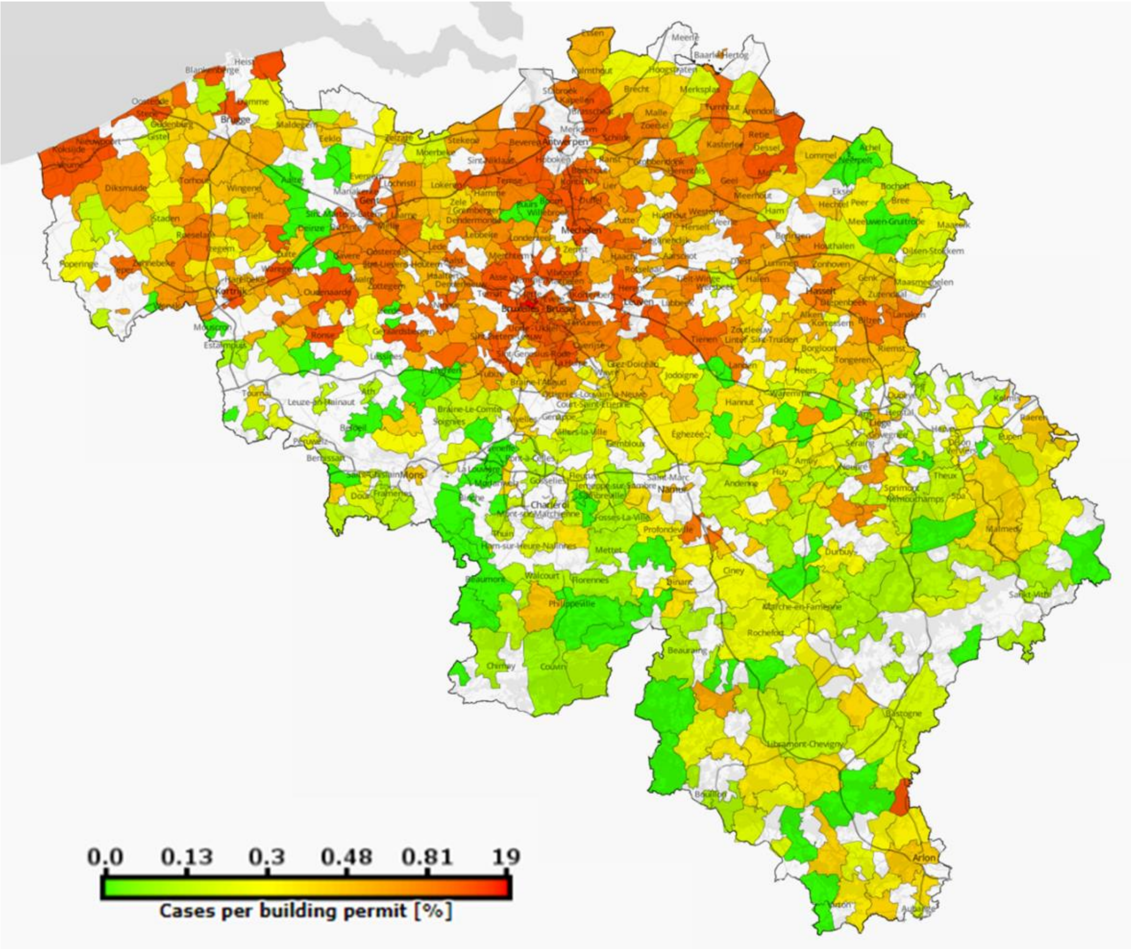

Figure 1. Geographical distribution of the number of cases per building permit.

For the analysis in this study, the type of damage and postal code is used. Underlying details regarding the claim are available in the database as well, but as this information is not documented in a structured way, it cannot be used as a source for statistical analysis. The database also contains information on the costs related to each aspect of the claim. Considering the confidential nature of this information, the impact on financials is not in scope of this study.

The most important field for this study is the type of damage. There is however some ambiguity in the utilization of this field as it is a single value that mostly is used to document the nature of the damage, but in other claims it refers to the location of the damage. As this study is about building defects, only the cases where the field was actually used to document the type of building damage were taken into account. The type of damage was mapped into one of the following damage categories:

- Moisture problems;
- Stability problems;
- Neighbour damage problems;
- HVAC problems;
- Acoustics problems;
- Energy efficiency problems.

As it is the aim to enrich the cases with information from external data sources based on postal code, the study was limited to cases for which a valid postal code was documented.

Figure 2 provides the overview of the exclusion criteria and the number of excluded cases.

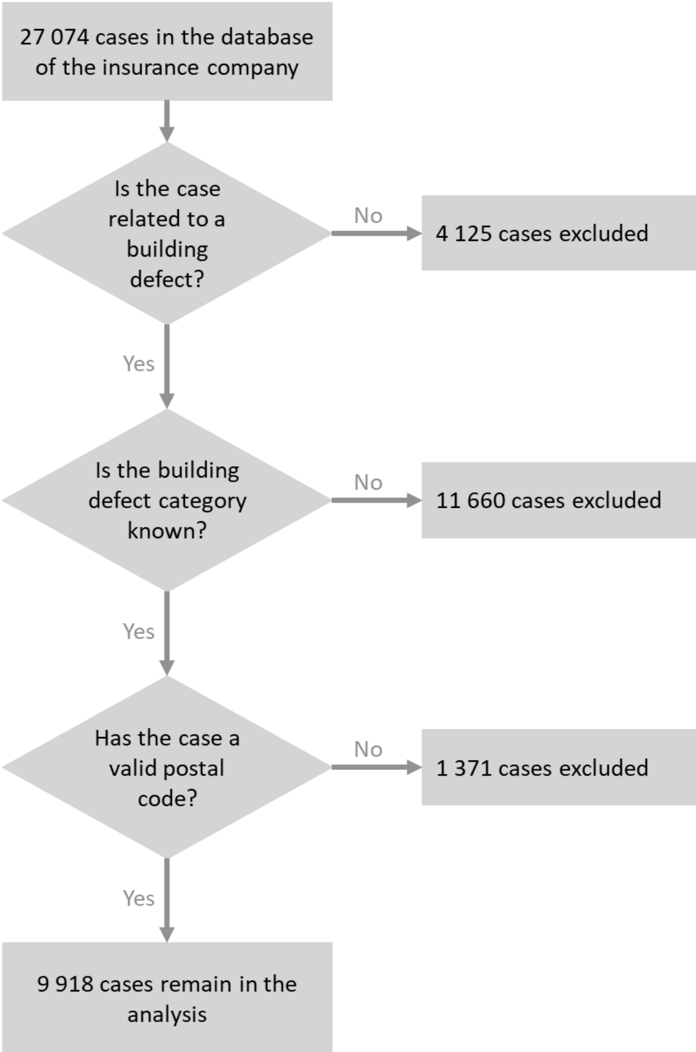

Figure 2. Flowchart of exclusion criteria and the number of excluded cases.

Enrichment of each claim with information from external data sources

As illustrated in Figure 3 for the damage category moisture problems, there are regional differences in the relative frequency at which the different damage categories occur. For each municipality the share of moisture problems was calculated based on the total number of cases.

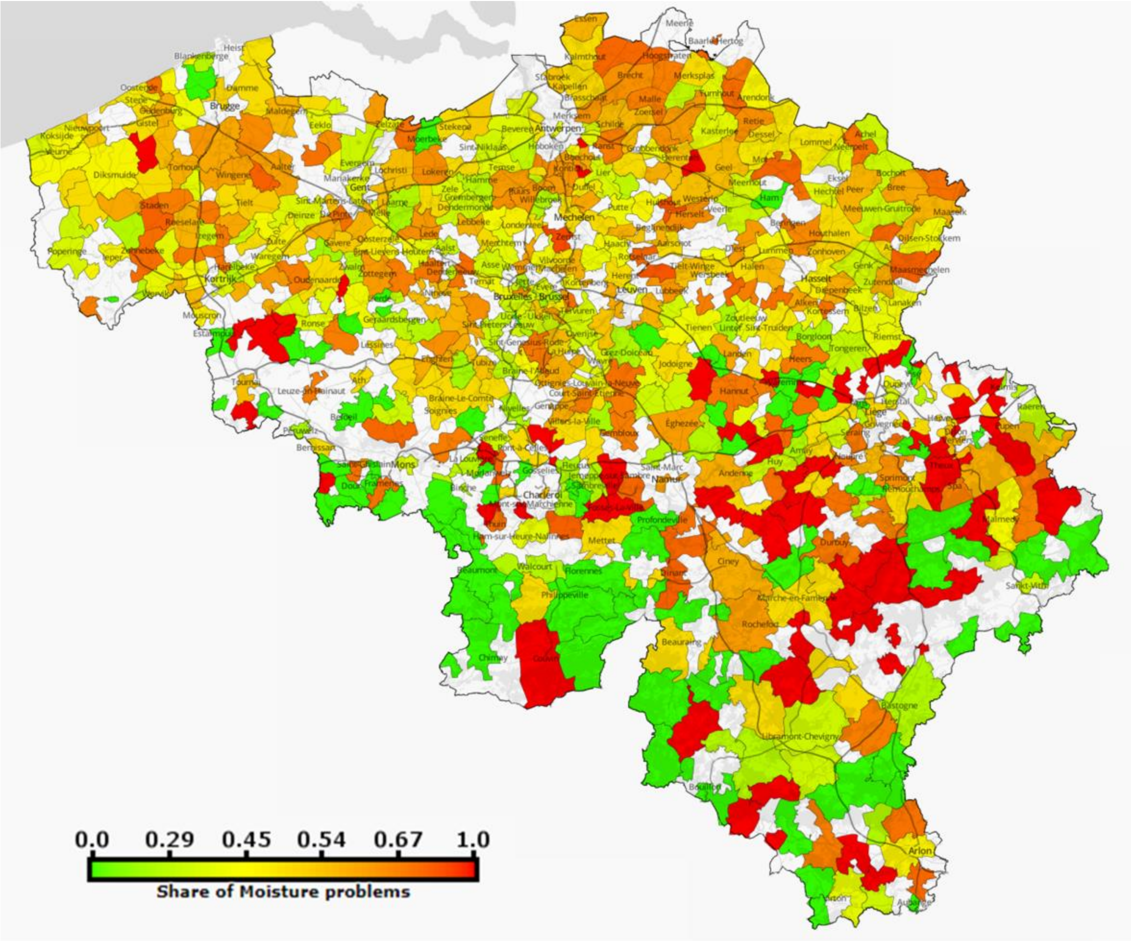

Figure 3. Geographical distribution of the share of moisture problems.

This finding shows that it is worthwhile to look for potential underlying factors that could explain these differences. In factor methods and neural network approaches typically materials, workmanship and boundary conditions are considered. In this study we assess the potential impact of demographical, geographical and climatological parameters. The demographical information might provide additional insights related to population density and urban fabric, and as well, typical income of residents, types of buildings, and size of the building plot. Correlations with these parameters may require additional analysis to see whether existing approaches may be susceptible to bias effects in this regard. As these parameters are not readily available in the database of the insurance company, other information sources are used to enrich the data in the database.

For this study, the climate atlas of the RMI website (RMI 2020b) was used to add meteorological information. The Belgian climate atlas shows the geographical distribution of the normal of various meteorological parameters concerning air temperature, precipitation, solar radiation, thunderstorms and snow. Evidently, Belgium is a small country with only one climate zone, but nevertheless there are still differences in important parameters that may affect the occurrence of certain damage types. Belgium has a temperate maritime climate with a mean temperature of 10.4°C, an average yearly rainfall of 825mm, less than 50 days of frost and less then 10 days of snow. Monthly mean minimum temperatures stay above freezing point, and in summer monthly max temperatures stay below 25°C.

The values for each municipality are the average of the values calculated for each grid point located above the municipal territory. This was derived by RMI from different interpolation schemes depending on the available weather stations per variable. The climate statistics can be interpreted as averages for the entire territory of each municipality (RMI 2020c). The climatological parameters for which the impact is analysed are the following:

- Temperature (annual average);
- Number of frost days (annual);
- Amount of precipitation (annual);
- Wind speed (annual average);
- Solar radiation (daily value of the annual average);
- Number of days with thunderstorms (annual);
- Number of days with snow per year (annual).

Statbel is the statistics agency of the Belgian government. It collects, produces and distributes reliable and relevant figures about the Belgian economy, society and territory (Statbel 2020). The collection of data is based on administrative data sources and surveys, the production takes place in a scientific, high-quality manner and the statistics are distributed in a timely and customer-friendly manner. The parameters obtained from Statbel are the following:

- Income of the inhabitants (average);
- Building value (median);
- Population density;
- Lot size (average);
- Number of common walls (average);
- Single-residence / Multi-residence / Non-residential buildings (share in all buildings).

The seismic zoning map from the Royal Observatory of Belgium was used to define the earthquake sensitivity for each municipality in Belgium. The seismic zoning map for Belgium was published in the Belgian national annex (NBN EN 1998-1 ANB) to the European building code Eurocode 8 (EN 1998-1). This map classifies Belgian communes into five seismic zones (0 to 4), corresponding to different values of the reference peak ground acceleration (PGA) to be taken into account in the design of structures for earthquake resistance (ROB 2020).In this study, the zone number is used as the parameter that identifies the earthquake sensitivity.

Also the following geographical information was gathered for each municipality:

- Geographical height (average);

- Distance to coast (km from centre of municipality).

Statistical analysis

As no information is available on how many building activities did not lead to a claim in the database, it is not possible to calculate the probability of occurrence of a building defect and neither to calculate the probability of occurrence of a specific damage category. However, it is possible to calculate the relative share of each damage category in the overall number of claims in the database. These relative shares are the dependent variables in the inferential statistics analysis. The independent variables in the statistical analysis are the parameters from the external data sources mentioned before.

In a first step, scatter plots were generated to study the impact of each parameter on the output values, and the significance level of linear and logarithmic correlations was evaluated. Based on the results, the input value was maintained, or converted to the $\log_{10}$ value of the input. It should be noted that this is not based on a physical understanding of the impact of the input variables on the output, but mere an approach to identify to what extent these correlations become apparent in the data-analysis, and avoid a lack of constant variance in the residuals. The full database of 9918 cases was used for this. As the different independent variables (the external factors) interfere with each other, a multiple-variable linear regression is used to assess the impact of the external factors on the relative occurrence of each damage category. The multiple-variable regression allows to pinpoint which parameters have a dominant impact on the output variables, but does not necessarily entail a reliable model with predictive value. Given that some input variables where not available (e.g. median building value for some locations), the total dataset of 9918 was reduced to 8942 datapoints for which all input variable were available to conduct the multi-variate analysis.

For the external factors where there appeared to be a logarithmic instead of linear relation, the $\log_{10}$ logarithm of this external factor was used as independent variable. This appeared to be the case for Population density, Distance to coast and Geographical height.

A p-value lower than 0,05 is used as criterion for statistical significance.

Eliminating redundant external factors

An attention point is that a multiple-variable linear regression only works correctly if the input parameters have no strong interdependence. If for instance two input parameters have a 100% correlation between each other, a multiple-variable linear regression cannot distinguish the individual impact of either of these factors and cannot be calculated.

In order to cope with this issue of multicollinearity, a solution is to remove the input parameters which depend too strongly on the other input parameters (and hence are also redundant in the analysis). To identify the most likely redundant input parameters, a correlation table (see Table 1) was made between all input and output parameters. Input parameters with a mutual correlation are flagged as suspects for being too dependent on the other input parameters.

Table 1. Correlation factors between all input and output parameters.
The correlations higher than 0,7 are highlighted.

	Income of inhabitants [€/year]	Building value [€]	Population density [#/km²] (10log)	Lot size [m²]	Distance to coast [km] (10log)	Earthquake sensitivity	Geographical height [m] (10log)	Number of common walls [#]	Single-residence building	Multi-residence building	Non-residential building	Temperature [°C]	Frost days [#/year]	Precipitation [mm/year]	Wind speed [m/s]	Solar radiation [kWh/m²/year]	Thunderstorm days [#/year]	Snow days [#/year]	Moisture problems	Neighbor damage problems	Stability problems	HVAC problems	Acoustics problems	Energy efficiency problems
Income of inhabitants [€/year]	1,00																							
Building value [€]	0,48	1,00																						
Population density [#/km²] (10log)	-0,37	0,30	1,00																					
Lot size [m²]	0,27	-0,17	-0,77	1,00																				
Distance to coast [km] (10log)	-0,19	-0,15	0,06	0,34	1,00																			
Earthquake sensitivity	-0,17	-0,27	-0,22	0,20	0,16	1,00																		
Geographical height [m] (10log)	-0,02	-0,03	-0,17	0,41	0,65	0,34	1,00																	
Number of common walls [#]	-0,47	0,12	0,81	-0,79	-0,15	-0,33	-0,42	1,00																
Single-residence building	0,55	-0,14	-0,60	0,58	0,16	0,18	0,42	-0,72	1,00															
Multi-residence building	-0,34	0,38	0,73	-0,76	-0,28	-0,16	-0,34	0,69	-0,85	1,00														
Non-residential building	-0,54	-0,31	0,06	0,04	0,12	-0,11	-0,30	0,32	-0,62	0,12	1,00													
Temperature [°C]	0,11	0,15	0,34	-0,20	-0,12	-0,35	-0,59	0,37	-0,18	0,14	0,14	1,00												
Frost days [#/year]	-0,06	-0,24	-0,42	0,37	0,31	0,43	0,60	-0,51	0,30	-0,30	-0,12	-0,85	1,00											
Precipitation [mm/year]	-0,10	0,00	-0,19	0,17	0,19	0,10	0,38	-0,20	0,02	-0,07	0,06	-0,68	0,67	1,00										
Wind speed [m/s]	0,09	0,10	-0,08	-0,19	-0,74	-0,11	-0,37	0,10	-0,04	0,10	-0,09	0,00	-0,26	-0,20	1,00									
Solar radiation [kWh/m²/year]	0,18	-0,04	-0,29	-0,11	-0,82	-0,11	-0,41	-0,04	0,07	-0,01	-0,12	-0,02	-0,22	-0,25	0,80	1,00								
Thunderstorm days [#/year]	-0,06	0,13	0,15	0,08	0,57	0,28	0,46	-0,07	0,02	0,05	-0,10	-0,25	0,45	0,30	-0,53	-0,66	1,00							
Snow days [#/year]	-0,11	-0,05	-0,24	0,22	0,34	0,42	0,68	-0,37	0,16	-0,08	-0,18	-0,90	0,87	0,64	-0,21	-0,22	0,54	1,00						
Moisture problems	0,06	0,01	-0,08	0,09	0,00	0,03	0,02	-0,10	0,07	-0,07	-0,02	-0,03	0,04	0,02	0,01	0,01	0,01	0,03	1,00					
Neighbor damage problems	-0,07	0,00	0,12	-0,16	-0,07	-0,07	-0,09	0,16	-0,12	0,12	0,04	0,08	-0,11	-0,05	0,03	0,02	-0,06	-0,10	-0,54	1,00				
Stability problems	0,00	-0,02	-0,04	0,07	0,06	0,05	0,07	-0,06	0,06	-0,05	-0,03	-0,05	0,06	0,03	-0,05	-0,03	0,05	0,06	-0,51	-0,30	1,00			
HVAC problems	-0,02	0,01	0,03	-0,01	0,02	0,00	0,00	0,02	-0,02	0,02	0,01	0,01	0,00	0,01	-0,01	-0,03	0,01	0,00	-0,16	-0,10	-0,09	1,00		
Acoustics problems	0,00	0,02	0,03	-0,04	-0,01	-0,02	-0,02	0,03	-0,03	0,03	0,00	0,01	-0,02	-0,02	-0,01	0,00	0,00	-0,01	-0,11	-0,06	-0,06	-0,02	1,00	
Energy efficiency problems	-0,01	0,00	-0,02	0,04	0,02	0,02	0,03	-0,03	-0,01	-0,01	0,03	-0,05	0,05	0,04	-0,02	-0,02	0,03	0,06	-0,08	-0,05	-0,05	-0,01	-0,01	1,00

Legend: Input parameter Output parameter

Starting with the highest correlations, the dependency of the suspected input parameters on the other input parameters was investigated. This was done by checking the R^2 value of the regression between the potentially redundant input parameter and the other input parameters. For our study, we have used 0,8 as cut-off value for R^2. After removing the redundant independent variables, the following external factors remained in the analysis:

- Income of inhabitants;
- Building value;
- Earthquake sensitivity;
- Geographical height (logarithmic);
- Number of common walls;
- Non-residential building;
- Temperature;
- Precipitation;
- Wind speed;
- Solar radiation;
- Thunderstorm days.

Use of averages per municipality

Note that the values for these external factors are available as an average on municipality level. For the geographical and climatological factors, this average is a very good approximation of the actual value for the concerned building given the negligible variation on that scale. For the parameters “Income of inhabitants”, “Building value”, “Number of common walls” and “Non-residential building” this is however not necessarily the case. There can be important individual differences between the different buildings in the same municipality. One will for instance in each municipality have buildings with zero and two common walls, regardless of the average number of common walls in the municipality.

However, this does not significantly affect the validity of the conclusions of the statistical analysis. When the linear regression shows that in municipalities with a higher average number of common walls there is a statistically lower occurrence of moisture problems, this lower occurrence of moisture problems can only be explained by a lower probability of occurrence of moisture problems for the individual buildings in that municipality. Obviously, this implies that all impacts found on relative occurrences have to be interpreted as probabilities. This study will hence only make statements on the probability that a damage will be of a certain category.

Results and discussion

Relative occurrence of each damage category

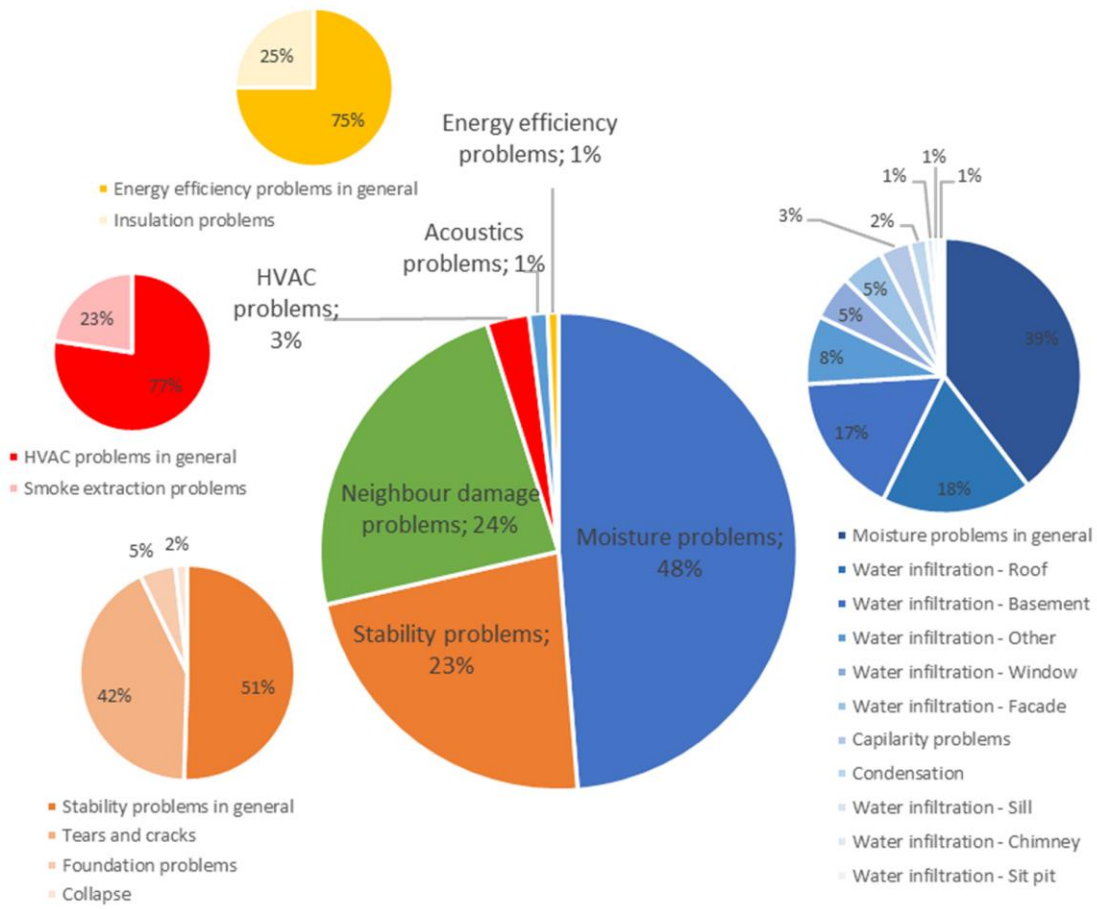

Figure 4. Distribution of the damage categories and major subcategories.

The shares of the six damage categories are visualized in Figure 4. Moisture problems have the biggest share of the relevant cases. This category accounts for 48% of all cases. Next in line are the categories neighbour damage problems and stability problems, both having a share of about 23%. HVAC problems, acoustics problems and energy efficiency problems are smaller categories, together having a share of less than 5%.

It should be noted that the temporal evolution of total number of claims is affected by the company structure, the variable market share over time, the work volume in the construction industry (business cycle), evolution of construction standards and guidelines, prevention measures, and the like. The relative share of the different damage categories over time show some fluctuation, but is relatively stable: e.g. the share of moisture problems ranges from 30% to 57% in the period 1991-2019. The largest variability in claims relates to neighbour damage, varying between 2 and 27%.

Finally, 47% of all claims were filed before the provisional acceptance, 13% between provisional and final acceptance, and 40% within 10 years after the final acceptance. Evidently, the fact that there are no claims beyond that point is due to the insurance framework that covers a 10 year period. For moisture problems 68% of the claims are filed after final acceptance, compared to 46% for stability problems. It appears that moisture issues come to front in a later stage, perhaps because some problems require an accumulative effect (buffering capacity of building materials, slow initiation of biological degradation, cyclic nature of some degradation phenomena), or perhaps because some moisture problems only occur during extreme (wind driven) rain loads that occur only every few years.

Normalization of variables

In order to obtain comparable values for the impact of each parameter in the multiple-variable linear regression, all data were normalised first. This normalisation is done for each dependent and independent variable separately to obtain an average of 0 and a standard deviation of 1 (see Table 2). The table also comprises the average value and standard deviation for the different variables. Given that Belgium is a small country, climatological, geographical and demographical variability may be limited, rendering it more complex to find significant correlations in the analysis.

Table 2. Input variables: units, average, standard deviation (STD), description, and source.

Parameter	**Unit**	**Average**	**STD**	**Description**	**Source**
Income of inhabitants	€/year	19586	2664	Average total net taxable income per inhabitant	Statbel
Building value	€	266215	7813 0	Median of the house sales prices in the municipality	Statbel
Earthquake sensitivity	-	1,33	0,79	Eurocode 8 seismic zone number (0-5)	NBN EN 1998-1 ANB
Geographical height	m (10log)	1,46	0,5	Average height above sea level of municipality	Topographic map Belgium
Number of common walls	#	0,83	0,37	0: detached / 1: semi-detached / 2: terraced	Statbel
Multi-residence building	-	0,84	0,07	0: no / 1: yes	Statbel
Non-residential building	-	0,16	0,07	0: no / 1: yes	Statbel
Temperature	°C	10,43	0,31	Averag annual temperature	RMI
Precipitation	mm/year	849,33	50,6	Average annual amount of precipitation	RMI
Wind speed	m/s	3,95	0,57	Average annual wind speed	RMI
Solar radiation	kWh/m²/ year	1026,87	22,2 9	Daily value of the annual average global solar radiation	RMI
Thunderstorm days	#/year	15,17	1,66	Based on Belgian Lightning Location System	RMI

Multiple-variable linear regression for moisture problems

Table 3 shows the result of the Multiple-variable analysis for moisture problems.

Table 3. Multiple-variable analysis for moisture problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,10826
R Square	0,01172
Adjusted R Square	0,01050
Standard Error	0,99468
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	104,7838	9,5258	9,6280	0,0000
Residual	8931	8836,2162	0,9894		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	0,03400	0,01951	1,74263	0,08143	-0,00425	0,07225
Building value [€]	0,00890	0,01483	0,60036	0,54828	-0,02016	0,03797
Earthquake sensitivity	0,00630	0,01315	0,47889	0,63203	-0,01948	0,03208
Geographical height [m] (10log)	-0,01004	0,01807	-0,55538	0,57865	-0,04547	0,02539
Number of common walls [#]	-0,09778	0,01749	-5,59012	0,00000	-0,13206	-0,06349
Non-residential building	0,03443	0,01441	2,38871	0,01693	0,00618	0,06268
Temperature [°C]	0,00002	0,02153	0,00104	0,99917	-0,04218	0,04222
Precipitation [mm/year]	0,00497	0,01689	0,29407	0,76871	-0,02814	0,03807
Wind speed [m/s]	0,05736	0,01854	3,09394	0,00198	0,02102	0,09371
Solar radiation [kWh/m²/year]	-0,03306	0,02338	-1,41358	0,15752	-0,07889	0,01278
Thunderstorm days [#/year]	0,01833	0,01604	1,14241	0,25331	-0,01312	0,04977

Note that the input and output parameters are all based on the normalized values. This means for instance that an increase of the number of common walls per building by 1 standard deviation (i.e. 0,369 extra common walls, see Table 2) leads to a decrease of 0,09778 standard deviations of the probability that a building defect is a moisture problem. Based on Table 2 these 0,09778 standard deviations correspond to a decrease of 4,88 percentage points of the probability that a building defect is a moisture problem.

In Table 3 the input parameters that have a p-value below 0,05 are highlighted. For these input parameters there is more than 95% chance that they do indeed impact the probability of moisture problems in the same sense (positive or negative) as the sign of the coefficient. In the next paragraphs, the underlying mechanisms that can explain this impact are discussed.

Number of common walls

Having an additional common wall decreases the probability that a building defect is a moisture problem by 13,23 percentage points. Number of common walls per building is the parameter with the highest absolute value of the coefficient. This means that this parameter, out of all the other parameters in the table, will have the highest influence on whether or not a moisture problem will occur. This impact can perhaps be explained by the fact that less common walls implies extra surface through which water can enter the building. On the other hand, the presence of common walls typically refers to a more urban context and more complex building details, and a higher likelihood for neighbour damage.

Wind speed

An increase in annual average wind speed by 1m/s, increases the probability that a building defect is a moisture problem by 5,01 percentage points. This corresponds to a difference of 10,51 percentage points difference between the least and the most windy municipalities of Belgium. Wind speed is the parameter with the second highest absolute value of the coefficient. This can be explained by the fact that wind speed has a linear correlation with wind driven rain load on walls, and a quadratic relationship with the driving rain wind pressure. The former quantifies the absolute rain load on a façade, whereas the latter yields the pressure difference acting upon the façade leading to an increased risk in water infiltration. The wind pressure during rain events is a relevant parameter as it is the main driving force for water ingress for a number of façade systems (Van Den Bossche et al., 2013). Instant water ingress in roof and wall assemblies is often reported, but an excess of rain water deposition in moisture buffering materials may also lead to slower moisture related defects such as frost damage, wood decay, and mould growth (Vandemeulebroucke et al., 2021). Furthermore, high wind speeds may cause elements to detach from the structure and in that way expose the building component to uncontrolled rain water infiltration.

Non-residential buildings

For non-residential buildings the probability that a building defect is a moisture problem is 24,61 percentage points higher than for residential buildings. This is a very high difference that can be explained by multiple factors. First of all, a lot of Non-residential buildings have a roof coverage out of bigger elements than the roof coverage generally found on residential buildings. These bigger elements are typically made out of metal (zinc, aluminium, copper). These metal elements might be a reason behind the increased share of moisture problems. A second factor is the fact that non-residential buildings are often detached and therefore less likely to cause neighbour damage, entailing a lower sensitivity to the number of common walls. A last factor is that non-residential buildings are in general more thoroughly studied from a structural stability point of view, which will perhaps reduce the likelihood of stability problems, which in turn increases the relative share of moisture problems.

Precipitation

One might expect that precipitation is an important driver for moisture problems. This does not come to front in the statistical analysis here. This could in theory be the result of the use of adapted building design or techniques in areas with more precipitation. It seems however more likely that the real explanation is that the variation in precipitation in Belgium is so limited (the standard deviation is only 5,9% of the average) that the impact is negligible.

Multiple-variable linear regression for neighbour damage problems

Similar to the analysis of moisture problems, the Multiple-variable analysis for neighbour damage problems was done (Table A.1 in Appendix A). It was found that the parameter Number of common walls has the highest impact on whether or not a neighbour damage problem will occur. Also for the parameters Non-residential buildings and Thunderstorm days a significant correlation is found.

Number of common walls per building

Having an extra common wall increases the probability that a building defect is a neighbour damage problem by 16,15 percentage points. Evidently, when there are more walls in common with your neighbour and when one lives closer to them, there is a higher likelihood that damage is inflicted on their building when working on your own building.

Non-residential buildings

For Non-residential buildings the probability that a building defect is a neighbour damage problem is 26,44 percentage lower than for residential buildings. As discussed in the previous section, number of common walls is not reported for non-residential buildings, and the reduced neighbour interaction comes to front here.

Thunderstorm days

Having an extra thunderstorm day per year, reduces the probability that a building defect is a neighbour damage problem by 1,04 percentage points. This corresponds to a difference of 9,63 percentage points between the municipalities with the highest and the lowest number of Thunderstorm days. Apparently, more Thunderstorm days lead to less neighbour damage problems. A potential explanation might be found in an increased probability of other damage categories, but there is insufficient statistical evidence for this, as it was not a significant parameter for moisture problems. As a conclusion, there is either an unknown underlying cause or the found correlation is a coincidence.

Multiple-variable linear regression for stability problems

For stability problems Wind speed and Solar radiation have the highest impact on whether or not a stability problem will occur (see Table A.2 in Appendix A). Also for the parameters Geographical height and Number of common walls a significant correlation is found.

Geographical height

Having the $\log_{10}$ Geographical height increase by 1 (which is the same as increasing the Geographical height by a factor 10), increases the probability that a building defect is a stability problem by 3,42 percentage points. Buildings with a high Geographical height in Belgium are situated on the slopes of hills in the southern part of the country. It makes sense that buildings situated on slopes are more sensitive to stability problems.

Number of common walls

An increase of 1 standard deviation (i.e. 0,369 extra walls) leads to a decrease of 1,49 percentage points of the probability that a building defect is related to stability problems. Easier said: having an extra common wall decreases the probability that a building defect is a stability problem by 4,04 percentage points. It could be that common walls contribute to the stability of the building. However, this negative correlation could just be the side-effect of the positive correlation between Number of common walls and neighbour damage problems.

Wind speed and Solar radiation

It is difficult to find underlying technical reasons why Wind speed and Solar radiation would impact stability. Due to the fact that both factors have an almost equal but opposite effect and that both parameters are rather highly correlated (correlation factor 0,80), their effect will mostly be neutralized. It is likely that the impact detected by the linear regression calculation is an artefact caused by the correlation between those two factors (i.e. the effect of collinearity). Therefore it was decided to disregard these two parameters.

Multiple-variable linear regression for HVAC problems

The multiple-variable linear regression analysis shows that only Solar Radiation has significant influence on the relative likelihood of HVAC problems (see Table A.3 in Appendix A). An increase of 1 standard deviation (i.e. 22,29) leads to a decrease of 0,95 percentage points of the probability that a building defect is related to HVAC problems. Easier said: Having the Solar radiation increase by 1 kWh/m²/year, decreases the probability that a building defect is an HVAC problem by 0,043 percentage points. Apparently, the additional sunshine that reduces the load (and hence the potential problems) on the heating system, has a higher impact than the fact that additional sunshine increases the load (and hence the potential problems) on the cooling system. This might be explained by the fact that only a limited number of buildings have a cooling system in Belgium. On the other hand, again the collinearity reported above may be affecting the results.

Multiple-variable linear regression for acoustics problems

Wind speed has the highest influence on whether or not acoustics problems will occur (see Table A.4 in Appendix A). Also for the parameter Number of common walls a significant correlation is found.

Impact of Wind speed

The reason why Wind speed results in a lower probability that a building defect is an acoustics problem is probably a side-effect of the fact that Wind speed results in a higher probability that a building defect is a moisture problem. Because of the higher number of moisture problems, the relative occurrence of acoustics problems is lower.

Impact of Number of common walls

An increase of 1 standard deviation (i.e. 0,369 extra walls) leads to an increase of 0,46 percentage points of the probability that a building defect is related to acoustics problems. Easier said: having an extra common wall increases the probability that a building defect is an acoustics problem by 1,24 percentage points.

These results make sense since for buildings with common walls, the acoustic insulation in respect to the other building is very important to minimize noise traveling from one building to the next. However, the impact is smaller than what might be anticipated.

Multiple-variable linear regression for energy efficiency problems

The Temperature has the highest impact on whether or not an energy efficiency problem will occur (see Table A.5 in Appendix A). Also for the parameters Non-residential buildings, Number of common walls and Building value a significant correlation is found.

Impact of Temperature

Having the temperature increased by 1°C, reduces the probability that a building defect is an energy efficiency problem by 2,36 percentage points. A logical explanation can be found in the fact that in colder regions, either the insulation or either the heating system needs to be more performant. It is hence not illogical to have more complaints related to energy efficiency.

Impact of Number of common walls

Having an extra common wall decreases the probability that a building defect is an energy efficiency problem by 0,97 percentage points. This might be explained by the fact that an additional common wall implies less heat losses thanks to a smaller surface through which the heat escapes the building, and heat gains from neighbouring buildings that might blur existing problems. On top of that, buildings with common walls are very common in cities. Studies have shown that the temperature in cities is usually higher than in the surrounding rural areas. On average, this difference rises to a few degrees, there can be peaks up to 7 to 8 °C and more (Caluwaerts *et al*. 2020). However, the decrease might also be caused by the relative increase of the share of neighbour problems for semi-detached and terraced houses.

Impact of Non-residential buildings

For Non-residential buildings the probability that a building defect is an energy efficiency problem is 6,84 percentage points higher than for residential buildings. This can be explained by the fact that Non-residential buildings are often larger and need more (complex) and tailor-made technical installations. Also note that non-residential buildings are often detached and that in that case they do not benefit of the energy efficiency benefits of common walls.

Impact of Building value

Having the Building value increase by € 100 000, increases the probability that a building defect is an energy efficiency problem by 0,0039 percentage points. This might be explained by the fact that more expensive buildings are often larger and therefore have more complex technical installations.

Risk of bias

The fact that the database of an insurance company is used inherently leads to different kinds of bias.

The cases reported to the insurance company only represent a portion of the building defects. The following cases of building defects will in general not be found in the database of the insurance company:

- Building defects for which one of the involved parties (e.g. the contractor) himself resolves the defect without involving an insurance company;
- Building defects for which only the insurance company of the contractor is involved, and not the one of the architect or engineer;
- Building defects detected after more than 10 years (the period of the liability in Belgium);
- Building defects caused by non-professionals (e.g. owner, tenants).

It is important to note that the market share of the insurance company is not the same all over Belgium. As a consequence, some geographical areas are more represented in the analysis than others.

An important drawback of using the claims database of an insurance company as the starting point of the study, is that almost the whole analysis is based on problem cases only, without data on how often things go well under given circumstances. In the analysis on the occurrence of the different damage categories, it is therefore clearly mentioned that these occurrence frequencies are relative to the overall number of cases and hence are not an indication of the absolute number. As a result of this approach, impact of an external factor on the number of cases of one damage category automatically also impacts the relative occurrence of the other parameters (as the total number of cases will change).

Discussion

The results of the statistical analysis on the data from the Belgian insurance company is summarised in Table 4. It shows the output of the linear regression expressed as the slope of the linear relation. Higher impacts (steeper slopes) have been given a darker colour in this table. The values are shown only for statistically significant correlations. Note that all values relate to buildings in Belgium only.

In order to be able to compare the different input parameters, the table uses the normalized values for the independent variables. The dependent variables (the probabilities of relative occurrence of each problem category) are expressed in percentage points (not normalized).

Table 4. Overview of the slope value results from linear regression of normalized input parameters.

	Moisture problems	Neighbor damage problems	Stability problems	HVAC problems	Acoustics problems	Energy efficiency problems
Income of inhabitants [€/year]						
Building value [€]						0,30%
Earthquake sensitivity						
Geographical height [m] (10log)			1,70%			
Number of common walls [#]	-4,88%	5,96%	-1,49%		0,46%	-0,36%
Non-residential building	1,72%	-1,85%				0,48%
Temperature [°C]						-0,73%
Precipitation [mm/year]						
Wind speed [m/s]	2,87%		-2,36%		-0,50%	
Solar radiation [kWh/m²/year]			2,31%	-0,95%		
Thunderstorm days [#/year]		-1,72%				

The external factor with the highest impact is “Number of common walls”. Not surprisingly, this is the input parameter that is the most closely linked to the architectural and technical aspects of the building. The most significant climatological factor appears to be the wind speed.

It is important to realize that all impact quantification reported above refers to probabilities. Even though the mentioned impacts are all statistically relevant with a p-value under 0,05 and that this impact is sometimes more than 30 percentage points, the R^2 values of the linear analysis are always less than 0,03. This means that less than 3% of the variance in the type of building defect can be explained by the variance of the input parameters. Hence, factors outside the statistical model are the main driver for building defects. These factors are primarily human factors such as execution, which introduces a large variability, but evidently also material properties and configuration of building components. This highlights that the multi-linear regression model does not provide any predictive power. Also reduced-order models for each type of damage do not entail any predictive power due to the low R^2-values. These conclusions are in line with typical literature in construction management, highlighting the dominant impact of design errors and execution errors in general. The contrast with research on building maintenance and factorial studies lies in the selection of cases. Damage cases from an insurance company only include defects originating within the first 10 years after completion, whereas degradation studies typically focus on older buildings, excluding evident design and execution errors, and for which material degradation affected by climatic parameters does in fact dominate the performance and damage risk (see e.g. Pereira et al, 2020).

Conclusions

The database of an insurance company with 27074 claims between 1991 and 2019 was studied for buildings in Belgium. Damage claims were filtered and categorized uniformly, with a focus on building defects. It was found that moisture problems account for 48% of all claims, followed by neighbour damage (24%), stability problems (23%), HVAC problems (3%), acoustics (1%) and energy efficiency problems (1%). Within the category of moisture problems, roofs and basements come to front as most susceptible building components. Furthermore, 47% of all claims were filed before the provisional acceptance, 13% between provisional and final acceptance, and 40% within 10 years after the final acceptance (decennial insurance framework). For moisture problems 68% of the claims were filed after final acceptance, compared to 46% for stability problems: moisture problems need more time to appear.

A linear regression analysis was done on the database, combined with meteorological, geographical, and demographical parameters. For meteorological parameters little relevant correlations were found: only wind speed seems to increase the relative share of moisture problems. However, it should be noted that the climate variability in Belgium is very limited. An increased geographical height entails a higher risk for stability problems, as this means for Belgium in practice that more constructions are built on slopes. The demographical data again showed little correlation with the occurrence of damage claims. Only the “number of common walls”, which is a proxy for population density, evidently leads to a higher number of neighbour damage claims.

The methodology presented in this paper allows to study which parameters affect the occurrence of building defects and insurance claims. The results allow to target specific fields for quality control measures. Furthermore, a wider adoption of this approach would allow an interesting comparison of construction markets, enable a study on the effectiveness of varying quality control frameworks, and provide a better understanding of how different construction management traditions affect the occurrence of building defects in the field of construction. However, the results indicate that meteorological, geographical and demographic parameters have no significant impact on the relative occurrence of different defects in insurance claims in Belgium. Most likely design errors and execution errors will dominate defects given that only a time frame of 10 years after completion is considered in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

AQC, 2016. Sycodés Pathology. www.qualiteconstruction.com

Atkinson A, 2010. Human error in the management of building projects. Construction management and economics, Vol 16 (3): 339-349 doi.org/10.1080/014461998372367

Bordalo R., de Brito J., Gaspar P.L., Silva A., 2010. Service life prediction modelling of adhesive ceramic tiling systems. Building Research & Information, Vol 39 (1): 66-78. doi.org/10.1080/09613218.2010.532197

BRE (1981) Quality Control on Building Sites, Building Research Establishment, Current Paper Publication n. 7/81, HMSO, London.

Caluwaerts S, Hamdi R, Top S, Lauwaet D, Berckmans J, Degrauwe D, et al., 2020. The urban climate of Ghent, Belgium : a case study combining a high-accuracy monitoring network with numerical simulations. *URBAN CLIMATE*; 31.

Carretero-Ayuso, M.J., A. Moreno-Cansado, and J. de Brito, 2017. Failure and damage determination of building roofs. *Revista De La Construccion*, 16(1): p. 13.

Chew M.Y.L., Nayanthara De Silva, Tan S.S., 2010. A neural network approach to assessing building façade maintainability in the tropics. Construction management and economics, Vol 22 (6): 581-594. doi.org/10.1080/01446190310001631019

CIDA (1995) Measuring Up or Muddling Through: Best Practice in the Australian Non-Residential Construction Industry, Construction Industry Development Agency and Masters Builders Australia, Sydney, Australia, pp. 59–63

Daniotti B., Lupica Spagnolo S., Hans J., Chorier J. (2008), "Service Life Estimation using Reference Service Life Databases and Enhanced Factor Method", Proceeding of the 11th International Conference on Durability of Building Materials and Components, Vol. IV, pp. 1459-1466, 11-14 May 2008, Istanbul, Turkey.

de Freitas V.P., 2013. A state-of-the-art report on building pathology. *Internation Council for Research and Innovation in Building and Construction*, p.5-195.

Duncan, A. and L. Ward, 2017. A building pathology system in New Zealand-what is possible? *BRANZ Study Report SR366: Judgeford,New Zealand*, p. 36.

Elios, 2010. Liability and insurance regimes in the construction sector: national schemes and guidelines to stimulate innovation and sustainability. Publication in the Official Journal of the European Union of 02/12/2008.

Forcada N., Macarulla M., Gangolells M., Casals M., 2014. Assessment of construction defects in residential buildings in Spain. Building Research & Information Vol. 42 (5): 629-640. doi.org/10.1080/09613218.2014.922266

Gaspar P.L., de Brito J., 2007. Service life estimation of cement-rendered façades. Building Research & Information, Vol 36 (1): 44-55. doi.org/10.1080/09613210701434164

Hammarlund, Y. and Josephson, P.E. (1991) Sources of quality failures in building, in Proceedings of the 1991 European Symposium on Management, Quality and Economics in Housing and other Building Sectors, pp. 671–9.

Love P.E.D., Li H., 2010. Quantifying the causes and costs of rework in construction. Construction management and economics, Vol 18 (4): 479-490. doi.org/10.1080/014461900050024897

Love P.E.D., Mandal P., Li H., 2010. Determining the causal structure of rework influences in construction. Construction management and economics, Vol 17 (4): 505-517. doi.org/10.1080/014461999371420

Love P.E.D., Edwards D.J., Smith J., 2016. Rework causation: emergent theoretical insights and implications for research. Journal of Construction Engineering and Management. Vol. 142 (6). Doi.org/10.1061/(ASCE)CO.1943-7862.0001114

Minato T, 2010. Representing causal mechanism of defective designs: a system approach considering human errors. Construction management and economics, Vol 21 (3): 297-305. doi.org/10.1080/0144619032000056135

Pereira C, Silva A, de Brito J, Silvestre JD, Urgency of repair of building elements: Prediction and influencing factors in façade renders, Construction and Building Materials, Volume 249, 2020. doi.org/10.1016/j.conbuildmat.2020.118743

RMI – Royal Meteorological Institute, 2020 (a). Wat doet het KMI? [online]. Available from: https://www.meteo.be/nl/over-het-kmi/presentatie, [Accessed 19 March 2020]

RMI – Royal Meteorological Institute, 2020 (b). Klimaatatlas [online]. Available from: https://www.meteo.be/nl/klimaat/klimaatatlas, [Accessed 19 March 2020]

RMI – Royal Meteorological Institute, 2020 (c). Klimaat in uw gemeente [online]. Available from: https://www.meteo.be/nl/klimaat/klimaatatlas/klimaat-in-uw-gemeente, [Accessed 19 March 2020]

ROB – Royal Observatory of Belgium, 2020. Eurocode 8 [online].Available from: http://seismologie.oma.be/nl/onderzoek/seismologie/seismische-gevarenkans-en-risico/eurocode-8, [Accessed 8 April 2020]

Shohet I.M., Paciuk M., 2010. Service life prediction of exterior cladding components under standard conditions. Construction management and economics, Vol 22 (10): 1081-1090. doi.org/10.1080/0144619042000213274

Statbel – Belgian Statistical Office, 2020. Over Statbel [online].Available from: https://statbel.fgov.be/nl/over-statbel, [Accessed 19 March 2020]

Vandemeulebroucke I., Caluwaerts S., Van Den Bossche N., 2021. Factorial Study on the Impact of Climate Change on Freeze-Thaw Damage, Mould Growth and

Wood Decay in Solid Masonry Walls in Brussels. Buildings Vol 11 (134). doi.org/10.3390/buildings11030134

Van Den Bossche N., Lacasse M.A., Janssens A., 2013. A uniform methodology to establish test parameters for watertightness testing. Part I: a critical review. Building and Environment 63: 145-156. doi.org/10.1016/j.buildenv.2012.12.003

W086, C., Building pathology_A state-of-the-art report. 1993, *International Council for Research and Innovation in Building and Construction*. p. 78.

Figure 1. Geographical distribution of the number of cases per building permit.

Figure 2. Flowchart of exclusion criteria and the number of excluded cases.

Figure 3. Geographical distribution of the share of moisture problems.

Figure 4. Distribution of the damage categories and major subcategories.

Table 1. Correlation factors between all input and output parameters.

Table 2. Averages and standard deviations of dependent and independent variables.

Table 3. Multiple-variable analysis for moisture problems.

Table 4. Overview of the slope value results from linear regression of normalized input parameters.

Table A.1. Multiple-variable analysis for neighbour damage problems.

Table A.2. Multiple-variable analysis for stability problems.

Table A.3. Multiple-variable analysis for HVAC problems.

Table A.4. Multiple-variable analysis for acoustics problems.

Table A.5. Multiple-variable analysis for energy efficiency problems.

Appendix A

Table A.1. Multiple-variable analysis for neighbour damage problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,16846
R Square	0,02838
Adjusted R Square	0,02718
Standard Error	0,98626
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	253,7302	23,0664	23,7135	0,0000
Residual	8931	8687,2698	0,9727		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	-0,02349	0,01935	-1,21389	0,22482	-0,06141	0,01444
Building value [€]	-0,02419	0,01470	-1,64526	0,09995	-0,05301	0,00463
Earthquake sensitivity	-0,01926	0,01304	-1,47675	0,13978	-0,04483	0,00631
Geographical height [m] (10log)	-0,01481	0,01792	-0,82617	0,40873	-0,04994	0,02032
Number of common walls [#]	0,13914	0,01734	8,02266	0,00000	0,10514	0,17313
Non-residential building	-0,04314	0,01429	-3,01855	0,00255	-0,07115	-0,01512
Temperature [°C]	0,02516	0,02135	1,17845	0,23865	-0,01669	0,06700
Precipitation [mm/year]	0,00891	0,01675	0,53223	0,59458	-0,02391	0,04174
Wind speed [m/s]	-0,00922	0,01838	-0,50168	0,61590	-0,04526	0,02681
Solar radiation [kWh/m²/year]	0,00254	0,02319	0,10962	0,91271	-0,04291	0,04799
Thunderstorm days [#/year]	-0,04022	0,01591	-2,52884	0,01146	-0,07140	-0,00904

Table A.2. Multiple-variable analysis for stability problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,08979
R Square	0,00806
Adjusted R Square	0,00684
Standard Error	0,99652
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	72,0853	6,5532	6,5991	0,0000
Residual	8931	8868,9147	0,9930		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	-0,01403	0,01955	-0,71749	0,47309	-0,05235	0,02429
Building value [€]	-0,00367	0,01485	-0,24697	0,80494	-0,03279	0,02545
Earthquake sensitivity	0,00912	0,01318	0,69184	0,48906	-0,01671	0,03495
Geographical height [m] (10log)	0,04063	0,01811	2,24376	0,02487	0,00513	0,07613
Number of common walls [#]	-0,03576	0,01752	-2,04052	0,04133	-0,07011	-0,00141
Non-residential building	-0,00496	0,01444	-0,34323	0,73144	-0,03326	0,02335
Temperature [°C]	0,00000	0,02157	0,00015	0,99988	-0,04228	0,04228
Precipitation [mm/year]	-0,00443	0,01692	-0,26172	0,79354	-0,03760	0,02874
Wind speed [m/s]	-0,05657	0,01857	-3,04539	0,00233	-0,09298	-0,02016
Solar radiation [kWh/m²/year]	0,05534	0,02343	2,36219	0,01819	0,00942	0,10126
Thunderstorm days [#/year]	0,02906	0,01607	1,80832	0,07059	-0,00244	0,06056

Table A.3. Multiple-variable analysis for HVAC problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,04257
R Square	0,00181
Adjusted R Square	0,00058
Standard Error	0,99965
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	16,2043	1,4731	1,4741	0,1335
Residual	8931	8924,7957	0,9993		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	-0,01696	0,01961	-0,86489	0,38712	-0,05540	0,02148
Building value [€]	0,01437	0,01490	0,96466	0,33474	-0,01484	0,04358
Earthquake sensitivity	0,00769	0,01322	0,58143	0,56096	-0,01823	0,03360
Geographical height [m] (10log)	-0,00977	0,01817	-0,53808	0,59053	-0,04538	0,02583
Number of common walls [#]	0,00416	0,01758	0,23678	0,81283	-0,03030	0,03862
Non-residential building	-0,00305	0,01449	-0,21084	0,83302	-0,03145	0,02534
Temperature [°C]	-0,00035	0,02164	-0,01603	0,98721	-0,04276	0,04207
Precipitation [mm/year]	0,00318	0,01697	0,18730	0,85143	-0,03009	0,03645
Wind speed [m/s]	0,03437	0,01863	1,84483	0,06510	-0,00215	0,07090
Solar radiation [kWh/m²/year]	-0,05803	0,02350	-2,46916	0,01356	-0,10409	-0,01196
Thunderstorm days [#/year]	-0,00768	0,01612	-0,47633	0,63385	-0,03928	0,02392

Table A.4. Multiple-variable analysis for acoustics problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,05057
R Square	0,00256
Adjusted R Square	0,00133
Standard Error	0,99928
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	22,8653	2,0787	2,0817	0,0184
Residual	8931	8918,1347	0,9986		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	0,01113	0,01960	0,56792	0,57011	-0,02729	0,04956
Building value [€]	0,01865	0,01490	1,25201	0,21060	-0,01055	0,04785
Earthquake sensitivity	-0,00248	0,01321	-0,18796	0,85092	-0,02839	0,02342
Geographical height [m] (10log)	-0,01898	0,01816	-1,04526	0,29593	-0,05458	0,01661
Number of common walls [#]	0,04146	0,01757	2,35938	0,01833	0,00701	0,07590
Non-residential building	0,00034	0,01448	0,02352	0,98123	-0,02804	0,02872
Temperature [°C]	-0,03741	0,02163	-1,72967	0,08372	-0,07981	0,00499
Precipitation [mm/year]	-0,02916	0,01697	-1,71867	0,08571	-0,06242	0,00410
Wind speed [m/s]	-0,04516	0,01863	-2,42431	0,01536	-0,08167	-0,00864
Solar radiation [kWh/m²/year]	0,01878	0,02349	0,79925	0,42416	-0,02727	0,06483
Thunderstorm days [#/year]	-0,00242	0,01612	-0,15034	0,88050	-0,03401	0,02917

Table A.5. Multiple-variable analysis for energy efficiency problems.

The highlighted parameters have a p-value smaller than 0,05.

df = Degrees of freedom; SS = Sum of squares; MS = Mean squares; F = F-ratio.

Regression Statistics	
Multiple R	0,08060
R Square	0,00650
Adjusted R Square	0,00527
Standard Error	0,99730
Observations	8942

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	58,0863	5,2806	5,3092	0,0000
Residual	8931	8882,9137	0,9946		
Total	8942	8941			

Normalized parameters	*Coëfficiënts*	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0,00000					
Income of inhabitants [€/year]	-0,00922	0,01956	-0,47139	0,63737	-0,04757	0,02913
Building value [€]	0,03483	0,01487	2,34270	0,01917	0,00569	0,06397
Earthquake sensitivity	0,00088	0,01319	0,06705	0,94654	-0,02497	0,02674
Geographical height [m] (10log)	-0,02712	0,01812	-1,49674	0,13450	-0,06265	0,00840
Number of common walls [#]	-0,04090	0,01754	-2,33201	0,01972	-0,07527	-0,00652
Non-residential building	0,05466	0,01445	3,78219	0,00016	0,02633	0,08298
Temperature [°C]	-0,08326	0,02159	-3,85744	0,00012	-0,12558	-0,04095
Precipitation [mm/year]	-0,02597	0,01693	-1,53363	0,12516	-0,05916	0,00722
Wind speed [m/s]	-0,00830	0,01859	-0,44651	0,65524	-0,04474	0,02814
Solar radiation [kWh/m²/year]	-0,01867	0,02345	-0,79643	0,42580	-0,06463	0,02729
Thunderstorm days [#/year]	0,00682	0,01608	0,42418	0,67144	-0,02471	0,03835