*Article*

# Data Extraction Algorithm for Energy Performance Certificates (EPC) to Estimate the Maximum Economic Damage of Buildings for Economic Impact Assessment of Floods in Flanders, Belgium

**Samuel Van Ackere *** [ID]**, Jochem Beullens, Alain De Wulf and Philippe De Maeyer**

Department of Geography, Ghent University, 9000 Ghent, Belgium; jochem.beullens@ugent.be (J.B.);
Alain.DeWulf@ugent.be (A.D.W.); Philippe.DeMaeyer@ugent.be (P.D.M.)
**\*** Correspondence: samuel.vanackere@ugent.be; Tel.: +32-(0)92-644-700

check for updates

**Abstract:** Floods cause major disruptions to energy supply and transportation facilities and lead to significant impacts on the society, economy, and environment. As a result, there is a compelling need for resilience and adaptation against extreme flood events under a changing climate. An accurate focal priority analysis of how societies can adapt to these changing events can provide insight into practical solutions. Besides the social, ecological, and cultural impact assessments of floods, an accurate economic impact analysis is required to define priority zones and priority measures. Unfortunately, studies show that economic impact assessments can be highly inaccurate because of the margin of error in economic value estimation of residential and industrial buildings, as they account for a large part of the total economic damage value. Therefore, tools that can accurately estimate the maximum economic damage value (or replacement value) of residential and industrial buildings are imperative. This paper outlines a methodology to estimate the maximum economic value of buildings by using a data extraction algorithm for Energy Performance Certificates (EPC), through which the replacement value can be calculated for all of the buildings in Flanders, and in addition, across Europe.

**Keywords:** flood impact assessment tool; FLIAT; data extraction; EPC; Energy Performance Certificates; disruption potential; urban floods; flood damage assessment; Flanders; Ghent; adaptation plan; priority plan; open source program languages

## 1. Introduction

Worldwide, a number of trends suggest that the severity and frequency of pluvial, fluvial, groundwater, and coastal floods are likely to increase in the future [1]. The disruption because of these flood events can have a tremendous economic impact on a society, both directly and indirectly. Therefore, it is important that flood damage assessments are used for supporting the priority adaptation guidelines, flood insurance, measures, and policy recommendations.

Flanders, the northern, low-lying region in Belgium, is vulnerable to flooding as a result of the rivers that pass through the inland. Additionally, Flanders is vulnerable to floods from the North Sea, by which overtopping can cause coastal floods in the coastal areas [2]. Because of the increasing severity and frequency of these floods since the 21st century [3], the Flemish government has changed its flood management strategy from a flood control approach to a risk-based approach. Instead of focusing on engineered artificial protection against a certain water level, the new focus lies in assuring protection "against the consequences of a flood, while considering its probability" [2], thus avoiding serious economic damage.

Nevertheless, in Belgium, it is still permissible to live and build in flood-prone areas. To prevent residents from moving to and building in these areas, a notification obligation was enacted in 2013 [4]. Subsequently, the notary, the broker, and the owner are obliged to inform the buyer or renter of buildings and building plots of the flood risk. The Flemish government determines and registers the flood risk and updates this information regularly. Although people are informed of this flood risk of a property prior to the sale or lease, they continue to build and live in flood-prone areas. These findings can be linked to the population's overall lack of knowledge concerning the possible impact of floods.

Yet the Flanders flood management strategies not only focus on reducing flood vulnerability, but also focus on how to build and live with water and the natural flood phenomenon [5], taking a cue from the Netherlands' Delta Programme of 2008 entitled 'Working Together with Nature—A living land builds for its future' [6]. Hereby, water robust building or resilient building covers all types of adapted building techniques to minimize flood damages on the building scale, including both wet- and dry-proofing [7]. Unfortunately, in reality, the actual compliance with adaptive cities, flood-safe building, and living is substandard, and there is still considerable work to be done with regard to resilience against extreme events under a changing climate.

Besides the social, ecological, and cultural impact assessments of floods, an accurate economic impact analysis is needed in order to define priority zones and priority measures. The so-called Standard Method is the most commonly used method to assess the real economic impact due to floods [8]. This method assesses the economic damage for every building, infrastructure, etc., individually, based on one or more of the flood characteristics (water depth, horizontal and vertical flood velocity, duration, etc.), the maximum damage per object, and a damage function. The damage function is the ratio between the flood characteristics and the fraction of the economic damage that occurs to the building, infrastructures, etc.

This Standard Method is used in a variety of flood models (e.g., LATIS for Belgium [2], FLEMO for Germany [9], the Multi-Coloured Manual for the UK [10], HAZUS for the USA [11], and SSM-2017 for the Netherlands [12]). Cammerer et al. (2013) showed that these flood damage models differ because each region, country, and flood type calls for a tailored flood damage model for each specific group, within the range of these many variables [13].

The Standard Method considers four elements, namely, the estimated maximum damage $s_i$ (total value) per unit in category $i$, the flood characteristics $d_j$ for a given return period (such as water depth, horizontal and vertical flood velocity, duration, etc.), the number of flooded units $n_{i,j}$ in each category $i$ in cell $j$, and the damage factor $f_{ij}$ for category i in cell $j$, dependent on flood characteristics $d_j$ (where $0 \leq f_{ij}(d_j) \leq 1$). The general formula that is used to predict flood damage is given by Egorova et al., (2008) [14], as follows:

$$S_{w,t} = \sum_{i=1}^{m} s_i \sum_{j=1}^{n} f_{ij}(d_j) n_{ij} \tag{1}$$

The economic damage is usually calculated for all categories, including buildings, infrastructure, cropland, meadow, etc. This is because each category has its own damage function and maximum economic damage. The most common flood models (Hazus [11], LATIS [2], SSM-2017 [12], etc.) use a relative damage function, where the fraction of the maximum damage is a function of the flood characteristics' intensity.

This method is visualized in Figure 1, in which each color of the cells represents a land use category (buildings, infrastructures, etc.). For every category, it is possible to link a maximum damage value to these categories.
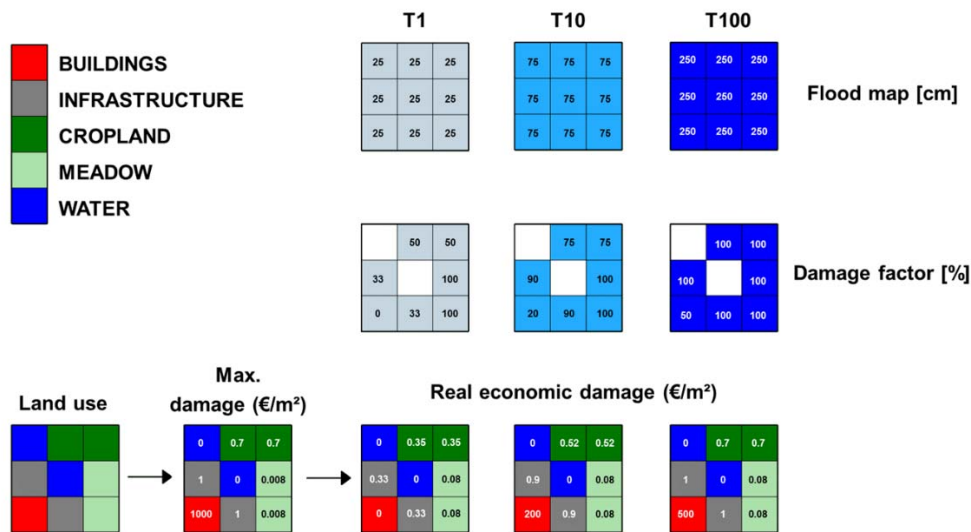
**Figure 1.** The Standard Method visualized for the LATIS flood impact analysis in Flanders, Belgium (adapted from [15]).

Every grid cell has a specific water depth from which a damage factor can be derived, using a damage function per specific land use category. Multiplying the maximum economic damage values per cell with the corresponding damage factor computes the real economic damage (see Figure 2). The consideration of uncertainty in damage analysis may be crucial for reliable results. Good decision-making can only succeed when all of the relevant uncertainties are taken into account [16,17]. Therefore, the uncertainties of the depending factors of the Standard Method need to be taken into account.
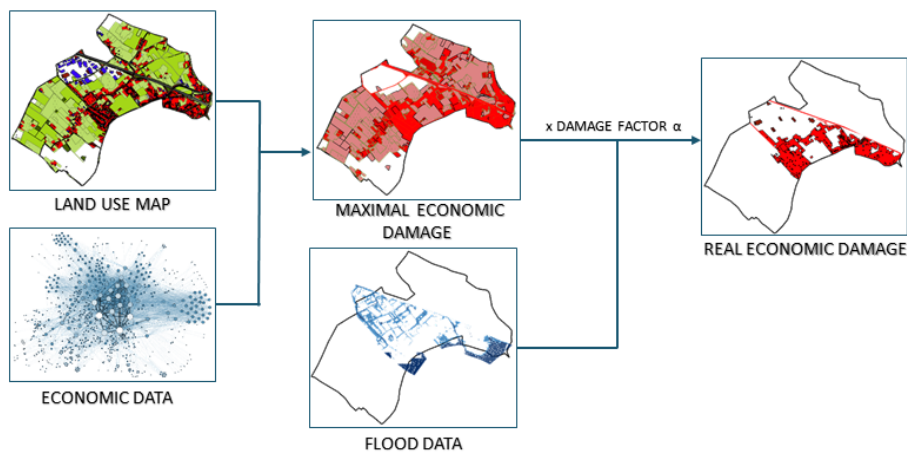


**Figure 2.** Real economic impact calculation (sample of the Flood Impact Assessment Tool (FLIAT) methodology [18]).

The economic damage assessment of buildings requires an estimation of the maximum economic value of buildings (see Figure 2). Hereby, "the maximum damage corresponding to each land use or object [ ... ] is defined as a replacement value" [19]. Thus, in order to determine the expected damage to buildings for a given flood, the replacement value of the buildings ought to be used, not the original value of purchase [20]. The replacement value or replacement cost is the value that an entity would have to pay in order to replace an asset at the current time, according to its worth.

Every flood model uses a different approach to estimate the maximum economic value of buildings. For instance, for the HAZUS-MH model, Shultz S. (2017) reports that the use of default

HGBS data, used in the HAZUS-MH model, should not be used for cost-benefit analyses or related feasibility analyses of natural hazard mitigation efforts. This is particularly due to the high bias and highly exaggerated structural damage reduction estimates [21]. Therefore, HAZUS-MH developers recommend that users import a more detailed replacement value data from external sources, using the Building Import Tool. These findings indicate that the majority recognizes the limitations in accuracy and the possibilities of optimization.

Furthermore, Kleist L. et al., (2006) has reported that the FLEMO model first calculates the total reconstruction cost of residential buildings per community, using the following equation [9]:

$$RCc = \sum_{i,j} LA_{c,i,j} \; x \; GFS_i \; x \; NHK_{i,j} \; x \; (1 + \; Anc_i) \tag{2}$$

with

*RC*: reconstruction cos*t* [€]
*c*: community index *c e* {1, . . . , 13,415}
*i*: building type *i e* {1, . . . , 38}
*j*: data quality *j e* {1, . . . , 4}
*LA*: living area [m$^2$]
*GFS*: type-specific conversion factor between living area and gross floor space
*NHK*: type-specific construction costs per m$^2$ gross floor space
*Anc*: incidental construction cost.

After calculating the total construction cost per community, the average replacement cost per community is calculated by dividing the total cost by the total living area per community. Although the real replacement value for every building can widely differ compared to the calculated average replacement value per community, this building stock replacement value dataset, computed in 2006, is still used in recent economic impact assessment studies [9].

In 2017, SSM-2017 replaced HIS-SSM [22] for the Netherlands [12]. The most notable differences include the update of the population and business data and the implementation of new insights, particularly regarding (indirect) trading loss and lower maximum damage amounts for homes. SSM-2017 uses a fixed replacement value of 1000 €/m$^2$ for buildings (low-rise, mid-build, high-rise, and single-family home), in order to calculate the economic impact due to floods [12].

Vanneuville, W. et al., (2002) has reported that for the LATIS model in Flanders, the maximum damage is determined on the basis of the average market value of all of the homes within a certain area [23], because there is a lack of replacement data for the buildings. Unfortunately, using the market value as a measure of structure value has some major shortcomings, as it incorporates the location in this value [19]. A building that is located near a lake would have a higher market value than an identical building that is located in a slum. In contrast with the replacement value, the market value is much more dependent upon subjective characteristics, such as visual characteristics (e.g., a building's interior and exterior appearance); proximity to schools, stores and other amenities; and surrounding noise disturbance; etc.

Clearly, every flood model uses a different way to estimate the maximum economic value of buildings. As a result, there are two types of shortcomings. Firstly, the market value (used in the LATIS model) and replacement value are distinctly different concepts, which are estimated using different criteria. In contrast with the replacement value, the market value is the estimated price at which a property would be sold on the open market, between a willing buyer and a willing seller, under all conditions for a fair sale. Therefore, it is important that the replacement value is used to calculate the economic damage that is as a result of floods.

Secondly, individual replacement costs will substantially differ from average replacement costs (used in the HAZUS-MH [11] and FLEMO [9] model) or a default replacement cost (used in the SSM [12] model), as each building differs from the 'average house' of a region. Therefore, to the extent

that it is possible, it is of utmost importance that the individual replacement cost of each building be considered.

When the right damage function is taken into account, the calculation of the economic damage due to floods is more accurate by using the exact replacement value instead of the space-averaged market value. Yet, it is important to take into account on the basis of which data these damage functions are developed. Attention must be paid when damage functions are developed in such a way that they reflect observed damages using house market values as an input, as they may not be used to calculate the economic damage due to floods in combination with calculated replacement values.

## 2. Methods

Unfortunately, in Flanders and across Europe, there is no existing database with replacement values of each individual building. Therefore, the data input of the technical and constructional information of buildings is crucial to calculating the replacement value. Fortunately, this data can be found in the Energy Performance Certificates (EPC), which are required for housing units in the European Union, and inform potential buyers or tenants about the energy efficiency of the building in which they are interested [24].

Since the introduction of the Energy Performance of Buildings Directive in 2002 (Directive 2002/91/EC) and the recast that was reinforced in 2010 (Directive 2010/31/EU) [25], all of the EU member states (MS) must comply with this directive by enforcing the necessary laws and regulations in order to improve the energy performance of buildings within the community. A certified energy inspector performs the building energy audit and creates an EPC of the building unit (e.g., apartment, residential building, etc.) after collecting all of the relevant data.

These EPCs include the following:

- administrative information (e.g., address, street name, zip code, municipality, cadastral department, cadastral section, and cadastral numbers);
- general information (e.g., building type, construction year, type of construction, nature of construction, form of construction, volume, etc.);
- the building envelope and building geometry (e.g., useful floor area, heated floor area);
- the composition of walls, floors, roof, doors, and windows, etc.;
- type and dimension of rooms;
- energy performance information (e.g., annual energy consumption per year);
- recommendations and expected energy savings;
- estimated $CO_2$ emissions;
- the building service (heating and domestic hot water) system;
- length of pipes for heated water (m);
- and airtightness of the building shell.

All of these parameters provided a rich insight into the technical and constructional information of buildings and could be used to calculate the replacement value. Because this document was an official document that was filled in by a certified energy inspector, the quality of the EPC data, and thus the input data for the calculation of the replacement value, was insured [26].

### 2.1. Data Extraction Algorithm

A Python script was developed that extracted data out of Energy Performance Certificates (EPC) and saved this data in a relational database that was linked to the geometric data of the buildings (see Figure 3). To extract the wanted data out of the EPC XML file, the ElementTree XML API, which was developed by Fredrik Lundh, was used [27]. Citing Frederik Lundh, "The ElementTree library provides an Element type, which is a simple but flexible container object, designed to store hierarchical data structures such as simplified XML infosets" [28].

The first step of the algorithm (see Figure 4) was to extract the administrative data, which contained location information of the building, from the EPC (e.g., address, street name, zip code, municipality, cadastral department, cadastral section, and cadastral numbers).
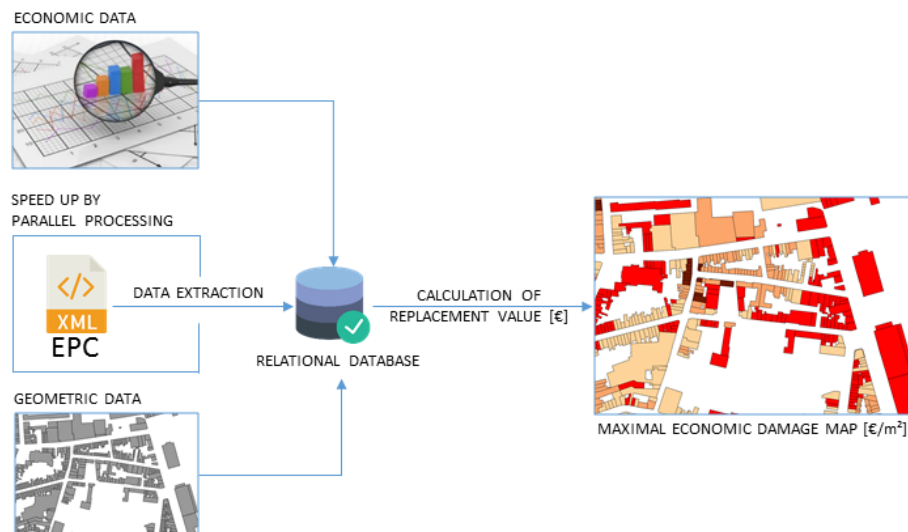


**Figure 3.** Data extraction methodology of the developed Python algorithm.

This location information made it possible to link the structural information, the function of the rooms and their dimensions, etc., to the vector data of all of the buildings in Flanders. By using the Central Reference Address Database of Flanders [29], it became possible to convert the address into the exact geographical coordinates of the building. In addition to this database, the Google Maps Geocoding API [30] or other geocoding API's could be used to convert addresses (like 'Krijgslaan 281, 9000 Ghent, Belgium') into geographic coordinates (like latitude 51.024135 and longitude 3.710557).

```
e = xml.etree.ElementTree.parse('epcfile.xml').getroot()
address =
e.find('./m__modelerProxy/m__building/m__common/m__energeticProxy/plot/
project/projectInformation/address')
streetname =  address.find('address').text
zip code = address.find('postalCode').text
municipality = address.find('town').text
CadastralData= address.find('cadastralData')
CadastralDepartment= CadastralData.find('cadastralDepartment').text
CadastralSection= CadastralData.find('cadastralSection').text
CadastralNumbers= CadastralData.find('cadastralNumbers').text
```

**Figure 4.** Code segment of the Energy Performance Certificates (EPC) extraction script.

After geocoding the address into geographical coordinates, the following extracted data could be linked to the data of the Flanders Large-scale Reference 3D Building Archive [31]. This archive contained the geometric data, the area, and the perimeter length of the footprint of the building. Additionally, the minimum and average Digital Terrain Model height, the maximum and 99 percentile height, and the maximum and 99 percentile absolute height of the building could be found in this 3D building archive, which was computed on the basis of LiDAR data [32].

Secondly, general information was extracted out of EPC, such as:

- construction year,
- volume,
- gross floor surface,
- total construction surface (including floor, wall, and roof surface),
- destiny (single-family house, apartment, loft, studio, holiday home in recreational domain, or collective housing),
- nature of construction (open buildings, semi-open building, or closed buildings),
- type of construction (heavy, moderately heavy, light, and half heavy),
- the orientation of frontage,
- a period of validity of the EPC, and
- annual primary energy consumption.

The type of construction indicated the inertia of the building and related to the quantity of the massive construction parts (see Table 1).

**Table 1.** The type of construction.

| Type of Construction | |
|---|---|
| Heavy | >90% of the construction is massive [1], of all of the horizontal, vertical, and inclined construction parts |
| Half heavy | >90% of the construction is massive [1], of all of the vertical and inclined construction parts |
| Moderately heavy | 50–90% of the construction is massive [1] |
| Light | <50% of the construction is massive [1] |

[1] Structural parts are considered massive if their mass is at least 100 kg/m$^2$.

Because the exact location and the footprint area was known, accompanied by an idea of the weight of the construction, it became possible to estimate the type of foundation and its replacement cost, taking into account the bearing capacity of the soil at that exact location [31].

This detailed information of the building inventory (building type, building destiny, and type of construction) in Flanders made it possible to link more applied damage functions to these well-defined categories, and to reduce the uncertainty of the impact assessment.

Subsequently, the type and function of rooms, quantity, and dimensions were extracted from the EPC. The types of rooms are divided into the following categories (see Table 2):

**Table 2.** Fixed categories of rooms in the Energy Performance Certificates (EPC).

| Type of Room | Area of Room Known |
|---|---|
| Living room (or analogue room) | Yes |
| Sleep, study, playroom (or analogue room) | Yes |
| Bathroom, laundry, drying room | Yes |
| Kitchen | Yes |
| Open kitchen | No |
| Washroom | No |
| Non-residential space | Yes |
| Corridor, staircase, hall (or analogue room) | No |
| Storage room | No |
| Basement or attic | No |
| Garage | No |
| Special room | No |
| Other type | No |
| Non-real spaces | No |

In this way, the number and size of the different types of rooms could be a benchmark to better estimate the number of inhabitants in each building, which could be helpful for determining socio-human impact assessments. Additionally, the number and size of the different types of rooms were used to estimate the value of the household goods, which gave better insight, versus only using the ground surface of the building as a dependent factor. Likewise, the presence or absence of a cellar was extractable with this data extraction possibility. Unfortunately, the exact location of these rooms in the building were not definable from this data. Although the connection of some rooms could be derived from data on the flow-in and flow-out openings, it was not possible to determine if a room was located on the ground floor or not.

Because the energy inspector had to give each room a name (e.g., basement), it was possible to find out the exact function of the room after analyzing these names (see example Figure 5), when this was not clear from the general category (e.g., basement or attic). Since the EPC file was an official document, the given name by energy inspectors was generally well-chosen and covered a perfect description of the room. Hereby, the function of the room could be deducted from the name of the room, when the selected category of the room was taken into account. Yet, attention should be given to typographical errors by filtering out frequently occurring errors.

```
x =
e.findall('./m__modelerProxy/m__building/m__common/m__energeticProxy/pl
ot/buildings/Building/protectedVolumes/ProtectedVolume/envelope/paroiRo
les/ParoiRole/paroi/volumes/Volume/envelopes/FunctionalPart/energeticSe
ctor/ventilationZone/systems/VentilationSystem/flowLocations/FlowLocati
on/wetSpaceLinks/*/space')
def findSpaces(root, l=[]):
    for child in list(root):
        findSpaces(child, l)

    name = root.find('./name')
    if name is not None:
        l.append(name.text)
    return l
spaces = [s for y in x for  s in findSpaces(y)]

def isBedroom(roomname):
    return 'sleep' in roomname.lower()
number_of_bedrooms = len(list(filter(isBedroom, set(spaces))))
```

**Figure 5.** Code segment of Python extraction script to define the existing rooms (e.g., number of bedrooms).

For non-residential rooms, there were multiple categories and subcategories from which to choose, as follows:

- restaurant and bars (with subcategories restaurants, cafeteria, bars, and kitchen),
- hotels, motels, holiday centers (with subcategories meeting rooms, lobby, bedrooms, etc.),
- office buildings (with subcategories office, meeting room, etc.),
- public spaces (with subcategories library, departure hall, etc.),
- public gathering places (with subcategories churches and other religious buildings, courthouses, etc.),
- retail trade (with subcategories mall, laundry service, hairdresser's, etc.),
- sports and leisure (with subcategories disco, changing rooms, sports club, etc.)
- workspaces (with subcategories photo studio, pharmacy, computer room, etc.),
- educational institutions (with subcategories classrooms and multipurpose hall),
- healthcare (with subcategories infirmary, examination room, operation room, etc.),

- correctional institution (with subcategories cells, day room, guard posts, waiting room, etc.),
- space not for human occupation (with subcategories Washroom stairwell, shower room, loading, and unloading space for industrial buildings, etc.).

This information was of enormous value for estimating the replacement value of the content of the buildings where a company was located, since, for each non-residential room, the specific area was defined. Via the Flemish company database [33], it could be determined whether or not the company or the specific function could still be linked to the building. In this way, the changes as a result of relocation, abandonment, and others, could be noted and taken into account.

After this, structural information of the building was extracted, such as the composition; use of material; and dimensions of the walls, floors, and the roof, which made it possible to calculate the replacement value of the integral structure of the building. Because the energy inspector could choose from a wide range of predefined construction materials out the EPC library, it became feasible to calculate the total replacement cost per construction element when a replacement value was assigned for every predefined construction material in the EPC library.

The composition of the structure of the building was also a criterion in order to estimate the flood resilience of the building—some structures were built to resist a flood event over a long period, while some were not [34]. Some buildings were built, for example, with materials which would allow fast recovery, namely, flood resistant doors, concrete block partitions, lime plaster, or magnesium oxide board as finishing skirting, which was sealed with an internal cavity membrane, etc. Unfortunately, because the EPC embedded only the rough composition of the walls, floors, roofs, etc., it was sometimes difficult to determine if the composition was flood resilient (e.g., lack of knowledge of the presence of construction films or the presence of rubbers between door frame).

In the final step, the orientation, dimensions, and type of doors and windows were extracted from the EPC, whereby it was possible to complete the calculation of the replacement value of the buildings.

## *2.2. 3D Module*

To date, it was possible for the energy inspector to model the building in 3D with the '3D Module' in the EPC (see Figure 6), which was added to the EPC software. This 3D module made it possible to embed geometric and topological properties to the EPC. Unfortunately, the inventory of geometric and topological data in the EPC was not yet obligated, but could already be used in the existing software.

When an EPC contained this geometric and topological data, it became possible to extract the exact location of doors, windows, garage doors, etc. Additionally, the exact position and dimensions of the basement could be extracted out of the EPC, whereby the location and dimensions of the cellar holes and basement stairs were known. Likewise, the exact slope of the roofs could be extracted out of this 3D model. Unfortunately, the location of rooms was not embedded in this 3D inventory, which made it still impossible to determine if a room was located on the ground floor or not. Yet the addition of the 3D module to the EPC was of enormous value when it was filled in to calculate the total replacement cost of the building. Unfortunately, because the inventory of the geometric and topological data in the EPC was not yet obligated, energy inspectors rarely used this module.
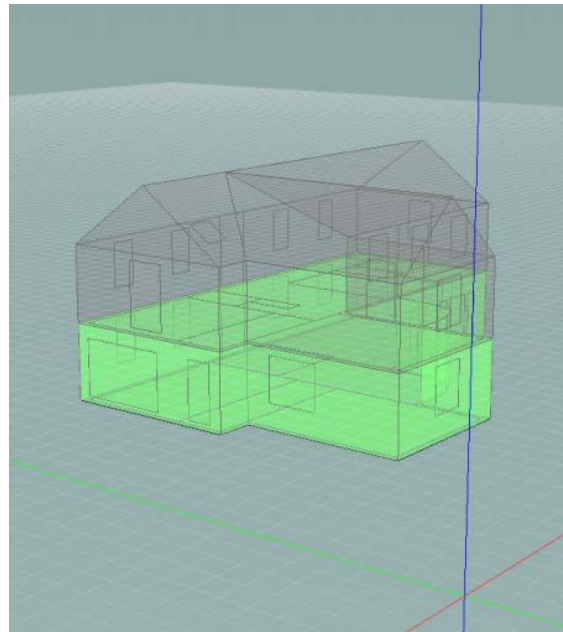
**Figure 6.** Geometric modeling with the 3D module in the EPC software of Belgium Government [35].

### 2.3. Relational Database Flood Impact Assessment Tool (FLIAT)

To handle the high variety of multi-variate data, there was a need for a powerful relational database (see Figure 7) that could be linked to a flood impact assessment tool. For this reason, in 2017, a new impact assessment tool, called FLIAT (www.fliat.be), was developed by the author so as to calculate the impact of floods in Flanders [18]. This tool used a vector approach methodology as well as a raster approach methodology, and was programmed in Python using the advanced PostgreSQL database with a PostGIS extension. The architecture of the FLIAT object-relational database management system was essential to save the extracted data out of the EPC, through which the total replacement value could be calculated by taking the higher described multiple input parameters into account.

In contrast with other flood impact assessment tools (Hazus, LATIS, etc.), the user could use FLIAT as a cloud computing tool, and downloading all of the necessary source data to use the flood impact assessment tool was not required.



**Figure 7.** The Flood Impact Assessment Tool (FLIAT) comes with a relational database that can embed a variety of detailed data sets.

*2.4. Speed up Extraction by Parallel Processing*

By using a database to store general to technical and constructional information of each unique building unit, it became possible to use parallel processing to speed up the extraction process (see Figure 3). As a result, the execution of processes were carried out simultaneously, using multiple cores of a desktop, laptop, or server. This task parallelism focused on distributing extraction tasks across different processors, in which each processor carried out the same extraction script but on a different EPC file. To run this task parallelism smoothly, the multiprocessing package was used, which was part of a pathos that was developed by Mike McKerns [36].

## 3. Case Study: Building Example

In order to demonstrate the capabilities of the data extraction algorithm, the data of a real EPC document is presented as an example. The samples of interesting and relevant data are published underneath all of the data that is extracted with the algorithm. Hereby, no effective replacement cost is calculated, since the extensive database with construction prices is not set up and an estimate of the operating costs has yet to be done.

In Table 3, the extracted administrative information of the building is shown. After geocoding the address into geographical coordinates, the following extracted data is linked to the data of the Flanders Large-scale Reference 3D building archive (see Table 4).

**Table 3.** Extracted administrative data out of the EPC.

| Administrative Data | |
| --- | --- |
| Region | Flanders |
| Address | Bartstraat 432 [1] |
| Zip code | 9230 |
| Municipality | Wetteren |
| Cadastral data | 4-B-6c2 [2] |

[1] Other data is shown because of privacy reasons. [2] Combination of cadastral department, cadastral section, and cadastral number.

**Table 4.** Data of the Flanders Large-scale Reference 3D building archive.

| H dtm min (m) | H dtm gem (m) | H dsm max (m) | H dsm p99 (m) | Hn max (m) | Hn p99 (m) | Length (m) | Area (m$^2$) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 9.76 | 9.80 | 20.46 | 19.32 | 10.70 | 9.52 | 44.95 | 109.07 |

Table 5 indicates that the building is a semi-detached, new construction, with a total volume of 764 m$^3$. The half-heavy type of construction indicates that it is a light construction, such as a timber frame construction.

Table 6 shows all of the extracted types of rooms and their accompanying dimensions, if available. From this information, it can be deduced that there is no basement present in the building, which is very important information when looking at the impact of floods.

**Table 5.** Extracted general data out of EPC document.

| General Data | |
| --- | --- |
| Building type | Single-family house |
| Construction year | 2015 |
| Form of construction | Semi-detached building |
| Nature of construction | New construction (or equivalent) |
| Clarification | Complete rebuilding |
| Type of construction | Half-heavy |
| Volume | 764 m$^3$ |
| Gross floor surface | 220.5 m$^2$ |
| Total loss surface | 473.385 m$^2$ |

**Table 6.** Extracted types of functions and accompanying dimensions of rooms out of EPC document.

| Name | Type of Room | Area (m²) [1] |
|------|------|------|
| Entrance | Corridor, staircase, hall (or analogue room) | NDA |
| Washroom 1 | Washroom | NDA |
| Kitchen | Open kitchen | NDA |
| Laundry Room | Bathroom, laundry, drying room | 6.8 |
| Living space + desk | Living room (or analogue room) | 46 |
| Bathroom | Bathroom, laundry, drying room | 12 |
| Bedroom 1 | Sleep, study, playroom (or analogue room) | 19.7 |
| Bedroom 2 | Sleep, study, playroom (or analogue room) | 11 |
| Shower | Bathroom, laundry, drying room | 2.7 |
| Attic | Basement or attic | NDA |
| Washroom 2 | Washroom | NDA |
| Cool storage | Storage room | NDA |
| Bicycle storage | Storage room | NDA |
| Night hall | Corridor, staircase, hall (or analogue room) | NDA |

[1] NDA: No data available.

Table 7 indicates that the entrance door has a height of 2.25 m and a width of 1.7 m. Its material is aluminium and it consists of a transparent element in the middle, with a surface of 3 m². The heat transfer coefficient (U value) of 1 W/m²K and a sun accession factor (g value) of 0.6 of the transparent element indicates that the glazing is a super-insulating double glazing, which is fixed in the door. This data makes it possible to estimate the replacement value of the door after an extensive database with prices set up (see Section 4). Additionally, it is possible to locate the entrance door by its orientation, and it is indicated that there is a door threshold present. Unfortunately, the height of thresholds are not embedded in the EPC, which would be very interesting information in order to calculate the exact water depth that can flow into the house.

**Table 7.** Extracted data of one door out of the EPC.

| Door | |
|------|------|
| Name | Entrance door |
| Type | Aluminum door |
| Surface | 2.25 m × 1.7 m |
| Thickness frame | 60 mm |
| Slope | 90° |
| Orientation | 15° |
| Boundary | Outside environment |
| Door profile type | Fixed |
| Ventilation grille | No |
| Element partly in glass | Yes |
| Transparent surface | 3 m² |
| Single Glass | No |
| Glazing with coating | No |
| Window profile type | Fixed |
| U value glazing | 1 W/m²K |
| g value | 0.6 |
| Threshold | 1.7 m |

Analogously, the replacement value of windows can be estimated. Table 8 indicates that the window is a super-insulating double glazing (U value of 1 W/m²K) with an aluminium framing and a surface of 1.2 m². Unfortunately, the EPC rapporteur is not obligated to fill in the height and width separately, but can also choose to fill in the surface of doors and windows immediately, which makes it more difficult to estimate the replacement value of this window.

**Table 8.** Extracted data of one window out of the EPC.

| Window | |
|---|---|
| Name | Window 1 |
| Type | Aluminum Frame |
| Surface | 1.2 m$^2$ |
| Slope | 90° |
| Orientation | −165° |
| Boundary | Outside environment |
| Burglary risk | No |
| Blinds | Yes |
| Glazing with coating | No |
| Window profile type | Fixed |
| U value glazing | 1 W/m$^2$K |
| g value glazing | 0.6 |

Both Tables 9 and 10 indicate that the building has a partly pitched roof and a partly flat roof (with a slight slope), with a surface of 88.60 m$^2$ and 49.35 m$^2$, respectively. Although the exact slope of the two roofs is not extractable out of the EPC, it is still possible to estimate the replacement value of the roofs, if the composition of the roofs is known (see Table 11).

Tables 11 and 12 show a large part of the composition of the pitched roof and the floor, respectively, without mentioning the finishing construction elements, such as roof tiles, battens, wind and rain protective films or plates, etc. Because of the lack of information of these finishing construction elements, an estimate will have to be made in order to include these materials in the replacement cost. After estimating the square meter prices of these construction materials, the total replacement cost can be estimated by multiplying by respectively the known floor surface and the roof surface.

**Table 9.** Extracted data of pitched roof out of the EPC.

| Pitched Roof | |
|---|---|
| Name | Pitched roof |
| Total surface | 88.60 m$^2$ |
| Slope | Between 0° and 60° |

**Table 10.** Extracted data of flat roof out of the EPC document.

| Flat Roof | |
|---|---|
| Name | Flat roof with slight slope |
| Total surface | 49.35 m$^2$ |
| Slope | Between 0° and 60° |

**Table 11.** Extracted data of composition of the pitched roof out of the EPC.

| # | Type of Layer | Fraction | Type of Material | Thickness (m) |
|---|---|---|---|---|
| 1 | Layer consists of one homogeneous material | 100% | Insulation/Pavatex/Isolair | 0.02 |
| 2 | Composed | 10% 90% | Timber of hardwood, deciduous and coniferous wood Cellulose (in situ) insulation | 0.3 0.3 |
| 3 | Layer consists of one homogeneous material | 100% | Non-ventilated air layer (air layer) | 0.03 |
| 4 | Layer consists of one homogeneous material | 100% | Plasterboard between two layers of cardboard (non-homogenous building materials) | 0.014 |

**Table 12.** Extracted data of composition of floor out of the EPC.

| # | Type of Layer | Type of Material | Thickness (m) |
|---|---|---|---|
| 1 | Layer consists of one homogeneous material | Heavy normal reinforced concrete (Stony components without joints) | 0.2 |
| 2 | Layer consists of one homogeneous material | Isotry (group)/isotry | 0.12 |
| 3 | Layer consists of one homogeneous material | Light concrete in tiling, solid slabs or coating (stony building parts without joints) | 0.07 |
| 4 | Layer consists of one homogeneous material | Grès tiles | 0.01 |

## 4. Discussion

### 4.1. Database with Construction Material Prices

In order to calculate the total replacement cost based on the information that is extracted from EPC files, an extensive database with construction material prices must be developed. This database has to embed, for every construction material, a substantiated square meter price. Hereby, the difficult task is to find a generally accepted, average substantiated square meter price for every construction material, taking into account all of the outliers for each material. For composed construction elements, the fraction needs to be taken into account of the construction element, to calculate the replacement cost correctly. Thus, in order to finally calculate the total replacement cost, this database needs to be developed in the future.

### 4.2. Additional Estimated Costs

Although studies have shown that building components constitute between 35% and 60% of the total construction cost [37,38], which makes it an important component with an enormous effect on the total replacement cost of the construction, the preparation (architectural design, stability study, earthworks, etc.) and the labor cost, etc. need to also be estimated. This has not been analyzed in this study and, therefore, future research around the elaboration of this estimate is needed.

Likewise, finishing (construction) materials that are not included in the EPC file need to be considered in order to estimate the total reconstruction cost. Because the EPC file does not include, for example, which type of roof tiles are used in the building, a fixed surcharge needs to be agreed upon for every construction element. This fixed surcharge has not been studied in this paper and, therefore, requires further analysis and research in the future.

Additionally, estimating the replacement cost of the household effects on the basis of the extracted data out of EPC (e.g., types of rooms and their surface) needs to be studied in the future.

### 4.3. Residential Building Stock with an Accompanying EPC

In 2014, 32.1% of the entire residential building stock of the Netherlands and 32.5% of the entire residential building stock of the UK had an accompanying EPC [39], and the number of new certificates that are added every year is over 1 million for the UK [26]. In 2015, Poland, Norway, Belgium, and Denmark had 18%, 22%, 24%, and 32%, respectively, of the entire residential building stock with an accompanying EPC [39]. To date, in Flanders, there are approximately 3.52 million unique addresses that are listed in the Central Reference Address File (CRAB) [29]. For every unique address, an Energy Performance Certificate will eventually be drawn up.

In July 2017, 1.1 million certificates were submitted in Flanders, which accounts for 31.5% of all of the building patrimonial in Flanders. Unfortunately, because not all of the buildings have an accompanying EPC file, it is necessary to use the sample data to perform a multivariate regression analysis to estimate the real replacement value of the remaining buildings, without an EPC. Thus, future research on the calculation of the replacement value of the building stock of Flanders, or other case

study areas in Europe, are still needed in order to develop a replacement value equation using multivariate regression analysis.

## 5. Conclusions

Flood damage assessments are critical to supporting priority adaptation guidelines, flood insurance, measures, and policy recommendations. Therefore, it becomes necessary to calculate the economic impact due to floods. To calculate the economic impact due to floods of buildings, infrastructure, etc., the so-called Standard Method considers four elements, namely: (1) the estimated replacement value; (2) the flood characteristics (such as water depth, horizontal and vertical flood velocity, duration, etc.); (3) the number of flooded elements; and (4) the damage factor.

Studies show that economic impact assessments can be highly inaccurate because of the margin of error in the economic value estimation of residential and industrial buildings, as they account for a large part of the total economic damage value. Therefore, tools that can accurately estimate the maximum economic damage value (or replacement value) of residential and industrial buildings are imperative.

Unfortunately, in Flanders and across Europe, there is no existing database with the replacement values of each individual building. Therefore, the data input of the technical and constructional information of buildings is crucial for calculating the replacement value. Fortunately, this data can be found in the Energy Performance Certificates, which are required for housing units in the European Union, and inform potential buyers or tenants about the energy efficiency of the building in which they are interested.

This paper presents a powerful algorithm that extracts the data out of Energy Performance Certificates (EPCs). These EPCs must be drawn up in Flanders from 2009 for every building that is built, rented, or sold. Because this EPC includes a tremendous amount of detailed data of the general structure, the amount, type, and dimensions of rooms, etc., it becomes possible to calculate the replacement cost or maximum damage value of each building. This algorithm is an innovative and necessary tool to create valid focal priority adaptation plans and build flood-resilient communities

**Author Contributions:** S.V.A. developed the EPC extraction script and wrote the paper; J.B., A.D.W., and P.D.M. upgraded the text where needed.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

## References

1. Hammond, M.J.; Chen, A.S.; Djordjević, S.; Butler, D.; Mark, O. Urban flood impact assessment: A state-of-the-art review. *Urban Water J.* **2015**, *12*, 14–29. [CrossRef]
2. Kellens, W.; Vanneuville, W.; Verfaillie, E.; Meire, E.; Deckers, P.; De Maeyer, P. Flood Risk Management in Flanders: Past Developments and Future Challenges. *Water Resour. Manag.* **2013**, *27*, 3585–3606. [CrossRef]
3. Chen, Y.; Zhou, H.; Zhang, H.; Du, G.; Zhou, J. Urban flood risk warning under rapid urbanization. *Environ. Res.* **2015**, *139*, 3–10. [CrossRef] [PubMed]
4. Ministerie van de Vlaamse Overheid. *Decreet Betreffende Het Integraal Waterbeleid*; Ministerie van de Vlaamse Overheid: Brussels, Belgium, 2003. (In Dutch)
5. Coördinatiecommissie Integraal Waterbeleid. *Overstromingsveilig Bouwen en Wonen*; Coördinatiecommissie Integraal Waterbeleid: Aalst, Belgium, 2011. (In Dutch)
6. Deltacommissie. *Working Together with Water: A Living Land Builds for Its Future*; Deltacommissie: Den Haag, The Netherlands, 2008.

7.  Tempels, B. Flood Resilience: A Co-Evolutionary Approach. Ph.D. Thesis, Faculty of Engineering and Architecture, Ghent University, Ghent, Belgium, November 2016.

8.  De Bruijin, K.M. Resilience and Flood Risk Management: A Systems Approach Applied to Lowland Rivers. Ph.D. Thesis, Civil Engineering and Geosciences, Delft, The Netherlands, January 2005.

9.  Kreibich, H.; Botto, A.; Merz, B.; Schröter, K. Probabilistic, Multivariable Flood Loss Modeling on the Mesoscale with BT-FLEMO. *Risk Anal.* **2017**, *37*, 774–787. [CrossRef] [PubMed]

10. Ossa-Moreno, J.; Smith, K.M.; Mijic, A. Economic analysis of wider benefits to facilitate SuDS uptake in London, UK. *Sustain. Cities Soc.* **2017**, *28*, 411–419. [CrossRef]

11. Scawthorn, C.; Flores, P.; Blais, N.; Seligson, H.; Tate, E.; Chang, S.; Mifflin, E.; Thomas, W.; Murphy, J.; Jones, C.; Lawrence, M. HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment. *Nat. Hazards Rev.* **2006**, *7*, 72–81. [CrossRef]

12. Slager, K.; Wagenaar, D. Standaardmethode Schade aan LNC-Waarden als Gevolg van Overstromingen. 2017. Available online: https://library.wur.nl/WebQuery/wurpubs/320700 (accessed on 10 July 2018). (In Dutch)

13. Cammerer, H.; Thieken, A.H.; Lammel, J. Adaptability and transferability of flood loss functions in residential areas. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 3063–3081. [CrossRef]

14. Egorova, R.; van Noortwijk, J.M.; Holterman, S.R. Uncertainty in flood damage estimation. *Int. J. River Basin Manag.* **2008**, *6*, 139–148. [CrossRef]

15. Deckers, P.; Broidio, S.; Verwaest, T.; De Maeyer, P.; Mostaert, F. *LATIS: Van Overstromingskaarten Naar Schadekaarten en Risicokaarten*; Vereniging leraars aardrijkskunde: Antwerp, Belgium, 2013. (In Dutch)

16. USACE. *Guidelines for Risk and Uncertainty Analysis in Water Resources Planning*; Institute for Water Resources: Fort Belvoir, VA, USA, 1992; Volume 1, 173p.

17. Aven, T.; Pörn, K. Expressing and interpreting the results of quantitative risk analyses. Review and discussion. *Reliab. Eng. Syst. Saf.* **1998**, *61*, 3–10. [CrossRef]

18. Van Ackere, S. Flood Impact Assessment Tool (FLIAT). Available online: www.fliat.be (accessed on 4 July 2018).

19. Vanderkimpen, P.; Peeters, P.; Deckers, P. The impact of individual buildings on urban flood risk analysis. In *SimHydro 2010: Hydraulic Modeling and Uncertainty*; Société Hydrotechnique de France: Nice, France, 2–4 June 2010.

20. Kellens, W.; Deckers, P.; Saleh, H.; Vanneuville, W.; De Maeyer, P.; Allaert, G.; Sutter, R. De A GIS tool for flood risk analysis in Flanders (Belgium). *WIT Trans. Inf. Commun. Technol.* **2008**, *39*, 21–27. [CrossRef]

21. Shultz, S. Accuracy of HAZUS General Building Stock Data. *Nat. Hazards Rev.* **2017**, *18*, 4017012. [CrossRef]

22. De Moel, H.; Aerts, J.C.J.H. Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates. *Nat. Hazards* **2011**, *58*, 407–425. [CrossRef]

23. Vanneuville, W.; Maeghe, K.; De Maeyer, P.; Bogaert, F.; Mostaert, P. *Methodologie en Case Study Denderbekken Risicobenadering bij Waterbeheersingplannen*; EurOBIS: Antwerp, Belgium, 2002.

24. Florio, P.; Teissier, O. Estimation of the energy performance certificate of a housing stock characterised via qualitative variables through a typology-based approach model: A fuel poverty evaluation tool. *Energy Build.* **2015**, *89*, 39–48. [CrossRef]

25. Hårsman, B.; Daghbashyan, Z.; Chaudhary, P. On the quality and impact of residential energy performance certificates. *Energy Build.* **2016**, *133*, 711–723. [CrossRef]

26. Arcipowska, A.; Anagnostopoulos, F.; Mariottini, F.; Kunkel, S. *Energy Performance Certificates across the EU. A Mapping of National Approaches*; BPIE: Bruxelles, Belgium, 2014.

27. Lundh, F. The ElementTree XML API. Available online: https://docs.python.org/2/library/xml.etree.elementtree.html (accessed on 24 March 2018).

28. Lundh, F. ElementTree—Python Wiki. Available online: https://wiki.python.org/moin/ElementTree (accessed on 29 March 2018).

29. Flanders Information Agency. Central Reference Address Database. Available online: https://www.agiv.be/international/en/products/crab-en (accessed on 2 May 2018).

30. Google Developers. Google Maps Geocoding API. Available online: https://developers.google.com/maps/documentation/geocoding/intro (accessed on 2 May 2018).

31. Flanders Information Agency. 3D GRB. Available online: https://download.agiv.be/Producten/Detail?id=971&title=3D (accessed on 4 May 2018).

32. Chen, K.; Xue, F.; Lu, W. Development of 3D Building Models Using Multi-Source Data: A Study of High-Density Urban Area in Hong Kong. In *LC3 2017: Volume I, Proceedings of the Joint Conference on Computing in Construction (JC3), Heraklion, Greece, 4–7 July 2017*; Heriot-Watt University School of the Built Environment: Edinburgh, UK, 2017; pp. 611–618. [CrossRef]

33. Federal Government—Department of Economy. Flemish Company Database. Available online: https://economie.fgov.be/nl/themas/ondernemingen/kruispuntbank-van (accessed on 5 May 2018).

34. American Society of Civil Engineers. Flood Resistant Design and Construction. Available online: https://ascelibrary.org/doi/pdf/10.1061/9780784413791 (accessed on 10 July 2018).

35. Belgian Government. Handleiding EPB-Software 3G, 3D-Module, Brussels, Belgium. Available online: http://document.environnement.brussels/opac_css/elecfile/UserManualPEB3D-ModelerNL (accessed on 5 May 2018). (In Dutch)

36. McKerns, M. Multiprocessing Python Package. Available online: https://github.com/uqfoundation/pathos (accessed on 24 March 2018).

37. Bourne, L.S. The Geography of Housing. *Real Estate Econ.* **1986**, *14*, 173–174. [CrossRef]

38. Windapo, A.O.; Cattell, K. The South African Construction Industry: Perceptions of Key Challenges Facing Its Performance, Development and Growth. *J. Constr. Dev. Ctries.* **2013**, *18*, 65–79.

39. Enerdata Consulting. Share of New Dwellings with EPCs. Available online: http://www.zebra-monitoring.enerdata.eu/overall-building-activities/share-of-new-dwellings-with-epcs.html (accessed on 24 March 2018).