Groningen Earthquakes Structural Upgrading

Data Documentation Exposure Database Version 5

ARUP

Date January 2018

Editors Jan van Elk & Dirk Doornhof

General Introduction

To be able to assess the risk for the buildings and community in Groningen resulting from induced earthquakes knowledge of full occupied build stock in the region of the Groningen field is required. For this purpose, an exposure database was built. An earlier version of this database (V2) (Ref. 1) was used for the Hazard and Risk Assessment of November 2015 (Ref. 2) and Hazard and Risk Assessment for Winningsplan 2016 (Ref. 3). Early 2017, an update of the databased was issued (Ref. 4 and 5)

This document provides information on the data used to create the Exposure Database (EDB) V5 delivered in September 2017. For each dataset, a description is provided, along with the contents, processing requirements and the limitations. The datasets used for the EDB are categorised as follows:

- Source data Datasets which have been received and maintained by external sources such as government departments.
- Project data Datasets which have been produced within the project such as inspection datasets and desktop studies. This includes project information produced by Arup and external consultants.
- Processed data Datasets which Arup has created utilising source datasets, assumptions and analysis to provide information that is not available from external sources.

This database was used in the hazard and risk assessment (Ref. 6) of November 2017.

As an Appendix, analysis the exposure database and additional data visualisations have been included.

References

1. Groningen Earthquakes - Structural Upgrading, Exposure Database V2, ARUP, Ikumi Nakanishi, Thomas Paul, Dirk Jan Oostwoud Wijdenes, September 2015
2. Hazard and Risk Assessment for Induced Seismicity in Groningen Interim Update November 2015, Nederlandse Aardolie Maatschappij BV (Jan van Elk and Dirk Doornhof, eds), 1st November 2015.
3. Winningsplan Groningen 2016, Nederlandse Aardolie Maatschappij BV, 1st April 2016.
4. Exposure Database: V3 Post-analysis report, 229746_031.0_REP1011_Rev0.03_ISSUE, Arup (several staff members), 25th January 2017.
5. Exposure Database: V3 Post-analysis report, 229746_031.0_REP1011_Rev0.03_ISSUE, Arup (several staff members), 25th January 2017.
6. Assessment of Hazard, Building Damage and Risk for Induced Seismicity Groningen, November 2017, NAM, Jan van Elk and Dirk Doornhof, November 2017.

NAM

<table>
<tr><td rowspan="2">Title</td><td colspan="2" rowspan="2">Groningen Earthquakes Structural Upgrading
Data Documentation Exposure Database Version 5</td><td>Date</td><td>January 2018</td></tr>
<tr><td>Initiator</td><td>NAM</td></tr>
<tr><td>Autor(s)</td><td>Several</td><td>Editors</td><td colspan="2">Jan van Elk and Dirk Doornhof</td></tr>
<tr><td>Organisation</td><td>ARUP</td><td>Organisation</td><td colspan="2">NAM</td></tr>
<tr><td>Place in the Study and Data Acquisition Plan</td><td colspan="4"><u>Study Theme:</u> Description Groningen Building Stock
<u>Comment:</u>
To be able to assess the risk for the buildings and community in Groningen resulting from induced earthquakes knowledge of full occupied build stock in the region of the Groningen field is required. For this purpose, an exposure database was built. An earlier version of this database (V2) (Ref. 1) was used for the Hazard and Risk Assessment of November 2015 (Ref. 2) and Hazard and Risk Assessment for Winningsplan 2016 (Ref. 3). Early 2017, an update of the databased was issued (Ref. 4 and 5)
This document provides information on the data used to create the Exposure Database (EDB) V5 delivered in September 2017. For each dataset, a description is provided, along with the contents, processing requirements and the limitations. The datasets used for the EDB are categorised as follows:
Source data Datasets which have been received and maintained by external sources such as government departments.
Project data Datasets which have been produced within the project such as inspection datasets and desktop studies. This includes project information produced by Arup and external consultants.
Processed data Datasets which Arup has created utilising source datasets, assumptions and analysis to provide information that is not available from external sources.
This database was used in the hazard and risk assessment (Ref. 6) of November 2017.
As an Appendix, analysis the exposure database and additional data visualisations have been included.</td></tr>
<tr><td>Directliy linked research</td><td colspan="4">(1) Seismic Response of Buildings
(2) Risk Assessment</td></tr>
<tr><td>Used data</td><td colspan="4">Various databases (BAG, Dataland, AHN, etc.), drawing data and building inspections (RVS, EVS)</td></tr>
<tr><td>Associated organisation</td><td colspan="4">NAM</td></tr>
<tr><td>Assurance</td><td colspan="4">Inspections</td></tr>
</table>

Client: Nederlandse Aardolie Maatschappij

Arup Project Title: Groningen Earthquakes Structural Upgrading

Data Documentation EDB V5

229746_052.0_REP2014

Rev.0.09 ISSUE_DEF | 18 January 2018

Job number 255090

Arup bv
Postal address:
PO Box 57145
1040 BA Amsterdam
Visitor address:
Naritaweg 118
1043 CA Amsterdam
The Netherlands

ARUP

Contents

1 Introduction

This document provides information on the data used to create the Exposure Database (EDB) V5 delivered in September 2017 [1]. For each dataset, a description is provided, along with the contents, processing requirements and the limitations. The datasets used for the EDB are categorised as follows:

- **Source data**
  Datasets which have been received and maintained by external sources such as government departments.
- **Project data**
  Datasets which have been produced within the project such as inspection datasets and desktop studies. This includes project information produced by Arup and external consultants.
- **Processed data**
  Datasets which Arup has created utilising source datasets, assumptions and analysis to provide information that is not available from external sources.

2 Data overview

In this section, the geo-referenced datasets that fall within the EDB V5 study area, also known as the Risk Assessment Area, are briefly described to give a general overview on the contents and how these are used in the EDB V5 [1].

2.1 Source Data

2.1.1 BAG

The '*Basisregistratie Adressen en Gebouwen*' (BAG) or Base Registration Addresses and Buildings datasets, produced by Kadaster [2], provide information on registered addresses and buildings, along with related information such as building year and status. The BAG datasets are used as the base building information and to provide the unique identifier for buildings which is used in analysis and to join building datasets.

The BAG datasets also provide information on the building coordinates, building year, footprint area and the number of addresses per building for the EDB V5 extract.

The BAG datasets used for the EDB V5 were updated in April 2017 and filtered to contain only existing buildings within the study area.

2.1.2 Dataland

Dataland provides real estate information based on addresses and it is used to provide insight into building use. The Dataland dataset used for EDB V5 was released in April 2015 [3]. Therefore, the version in use is related to the BAG address information of March 2015.

2.1.3 Parcel data

The parcel information or '*percelen en grenzen*' [4] provides the spatial geometry of parcel (lot) information and was used for the development of both the V2 and V3 releases of the database. This information is used to help identify buildings which are sited on the same parcel, especially buildings without an address.

2.1.4 Vs30

The Vs30 dataset [5], produced by Deltares, contains information on a time averaged value of the Vs (shear wave velocity) over the top 30m of the soil. This information is used to provide information on the ground conditions of buildings and used in the foundation analysis. Further information on the development of the map can be found in the Deltares report [6] that accompanies the data.

This set of Vs30 information supersedes TNO's Vs30 information which was used in the V0 and V1 versions of the exposure databases but remains unchanged from V2.

2.1.5 AHN

The '*Actueel Hoogtebestand Nederland*' [7] is a detailed height grid for the Netherlands, obtained by using laser altimetry: the latest version of the AHN which covers the scope area, the AHN2, is implemented in the EDB V5. There are filtered as well as unfiltered grids available in different resolutions. The filtered grid shows just the terrain, the unfiltered grid includes all objects such as buildings, vegetation, cars, etc. Arup uses for both the filtered as well as the unfiltered data at 0.5m grid.

The data for the study area have been released in 2009 (i.e. AHN data for the study area is not available post 2009). The data remains the same from previous versions of the EDB but the analysis which uses the dataset has been updated for V5, this concerns the gutter height and roof analysis processed datasets.

2.1.6 Rijksmonumenten

The Rijksdienst voor Cultureel Erfgoed (RCE) provides the '*Rijksmonumenten-kaartlagen*' [8]; information on the locations and descriptions of officially registered national monuments. This data was used as an input for the Building Use analysis to identify churches (with monument status) within the study area. This dataset was updated for V5 in December 2016.

2.1.7 Nationale Atlas Volksgezondheid

The Rijksinstituut voor Volksgezondheid en Milieu provides information about public health and environment through their online platform '*Nationale Atlas Volksgezondheid*' [9]. The locations of the hospitals have been collected by Arup from this platform. This dataset was updated for V5 in December 2016.

2.1.8 Basisregister Instellingen

The Dienst Uitvoering Onderwijs (DUO) manages the *'Basisregister Instellingen'* (BRIN) [10], which contains information on all schools in the Netherlands. This dataset is used to identify primary schools, high schools, special educational schools, vocational schools, colleges and university buildings. The School Registry is updated every month and can be extracted per province. The School Registry that is used for this EDB V5 was last updated on December 2016.

2.2 Project Datasets

2.2.1 RVS inspections

The *Rapid Visual Screening* (RVS) is a preliminary building assessment process [11], designed with the principles of existing international guidelines (FEMA) [12], aiming to collect building information from the public realm (without entering the property boundaries). The RVS focuses on providing a safety assessment of inhabitants, identify external High Risk Building Elements (HRBE) and provides input for future assessment activities as initial step of a tiered approach.

RVS inspections carried out by Arup (up to November 2015) were included in EDB V5.

2.2.2 EVS inspections

The *Extended Visual Screening* (EVS) is a structural assessment based on the visual inspection of the building's interior and exterior [13]. The EVS focuses on identifying potential falling hazards and significant structural damage and deformations. The inspection also includes the collection and recording of structural information and construction details where visible.

For EDB V5, EVS inspections carried out by Arup (up to November 2015) were included.

2.2.3 Desktop Visual Inspections, JBG

The *Desktop Visual Inspections* undertaken by JBG [14], is a desk study, designed to collect building information using Google Streetview [15] and additional building pictures produced by the company Horus. These visual inspections were delivered on July 2017 and focused on habitable buildings within the 0.2g PGA contour [16] except for terraced buildings which were covered by the technical drawing data collection (Chapter 2.2.4).

2.2.4 Technical Drawing Data Collection, Arup

The *Technical Drawing Data Collections* are undertaken by Arup to provide information about construction and internal features of selected buildings [17]. Drawings of selected buildings were obtained from the relevant municipalities and construction information was then collected from them in a digital format. The technical drawing data used in EDB V5 covered terraced buildings within the 0.2g PGA contour [17].

2.2.5 Arup Expert Inspections

The *Arup Expert Inspections* dataset is a collection of three inspections performed by Arup through the collation of project information and desktop studies, which have been formatted to allow for use in the database and the building classification process.

2.2.6 Population, NAM

The *population* dataset was provided by NAM and was received in September 2017 [18]. It includes calculated population within and nearby buildings for day and night periods. The dataset is included in the EDB V5 extract.

2.3 Processed Datasets

Processed datasets have been created to serve as input into another analysis or as a direct field within the EDB extract. These additional datasets are described herein.

2.3.1 Building Use

Building Use information provides the primary and secondary use of a building and a flag to identify if a building contains a residential use. This is used as a direct data field in the exposure database extract as well as input into the structural system classification process as inference modifier [1].

For EDB V5, the Building Use analysis has been updated to include the updated and additional source data alongside an improved methodology which uses the additional source datasets and validates specific uses relating to the building typologies. This includes using the national monument register (2.1.6) to identify churches and the school registry (2.1.8) to identify schools. Additionally, the parcel information (2.1.3) is used to appoint buildings, without an addresses, such as sheds, or an unclear Building Use, to a building with particular use sharing the same parcel.

2.3.2 Adjacency and unit count

In the *adjacency* analysis the spatial relation between separate buildings is determined; i.e. how a building relates to neighbouring buildings, if there are any. The result of this analysis is a set of parameters which are input for the building typology classification (such as number of buildings in a block, number of neighbouring buildings, etc.). Furthermore, a flag (single separate flag) is assigned to blocks that contain one building with a different building year but are otherwise homogeneous.

The results of the adjacency analysis are part of the EDB extract as well as input into the structural system classification process.

2.3.3 Exposed footprint length

The *exposed footprint length* captures the length of the building's footprint which are exterior facing (i.e. not including walls between buildings). The exposed footprint length is a data field in the EDB extract.

2.3.4 Gutter height

The *gutter height* is the height of the building's walls (i.e. where the gutter would be located) excluding sloped roof planes. In the case where the building has several different gutter heights, the average is used weighted by the length of the wall. The gutter height is data field in the EDB V5 extract and is an input into the Structural System classification process.

2.3.5 Roof steepness

The *roof steepness* aims to capture whether the roof of a building is mainly flat, non-steep or steep. To do so, the angle of the roof slope is recorded per volume and assigned a percentage of roof area within an angle domain. The domains are based on a 10° angle step starting from 0° and reaching 90° angles. The roof steepness is an input into the Structural System classification process.

2.3.6 Roof count

The *roof count* aims to capture whether the building has a simple (i.e. one roof ridge) or complex roof structure (multiple discontinuous roof ridges). The presence of one or more roofs was assigned based on the identification of one or more roof parts that are considered as flat and/or the identification of one or more roof ridges. The roof count is an input into the Structural System classification process.

2.3.7 Maximum enclosed rectangle

The *maximum enclosed rectangle* (MER) information captures the dimensions of the largest possible rectangles to fit within a building footprint. These rectangles and their dimensions provide information about the building footprint and can be used as an indication for the likelihood of different Structural Systems. The MER dimensions are therefore used as an input into the Structural System classification process.

3 Source Data

3.1 BAG

The '*Basisregistratie Adressen en Gebouwen'* (Base Registration Addresses and Buildings) [2] or BAG datasets provides information on registered addresses and buildings (defined as 'pand' in BAG). The BAG datasets are the key source datasets used to relate other datasets to a building level alongside providing a basis for various analyses.

The BAG data is delivered in the form of several key-related datasets which relate to how the BAG records and structures the buildings and addresses register [19]. These datasets and their descriptions can be found below.

***Openbare ruimte, nummeraanduiding and woonplaats* [Address]**
The combination of *openbare ruimte, woonplaats* and *nummeraanduiding* is the geographic location as made up by street name, house number, letter, postal code and town. Combined this information forms the address. Addresses can only be assigned to an *addresseerbaar object* (i.e. addressable object) of either a *standplaats, verblijfsobject* or *ligplaats* (described below). Each address has a unique ID (from the *nummeraanduiding* set) and is geometrically described as a point feature.

***Verblijfsobjecten* (VBO) [Use]**
The *verblijfsobject* or VBO is the smallest unit of function (use) within one or more building objects. The VBO can be accessed through its own and lockable entrance from a public road, yard or shared space, is subject to property law and functionally independent from other VBOs. Each VBO has a unique ID. It is an addressable object and each VBO has a one to one relationship with an address ID. Note that no VBO can exist without a relationship to a building.

***Standplaats and ligplaats* [Standing and Berth place]**
Standplaats and *ligplaats* are the smallest unit of function (use) related to a 'standing place' (i.e. registered caravan point) or 'berth' (i.e. registered boat house berth). In both cases, the objects at a standing place and berth are not permanently connected to the ground (i.e. movable) and are considered not relevant to the scope of the project.

***Panden* (Buildings)**
The *pand* is the smallest unit that is directly and permanently connected to the ground, able to be entered, lockable and structurally independent. Each building has a unique id. The *pand* is equivalent to a 'building' for the purposes of the EDB and for ease, *panden* or *pand* will be referred to as buildings or building respectively. As noted above, buildings are related to VBOs and accordingly, addresses. Note that while a VBO cannot exist without a relationship to a building, a building may have no relationship to a VBO.

The diagram below (Figure 1) shows the relationship between the datasets.

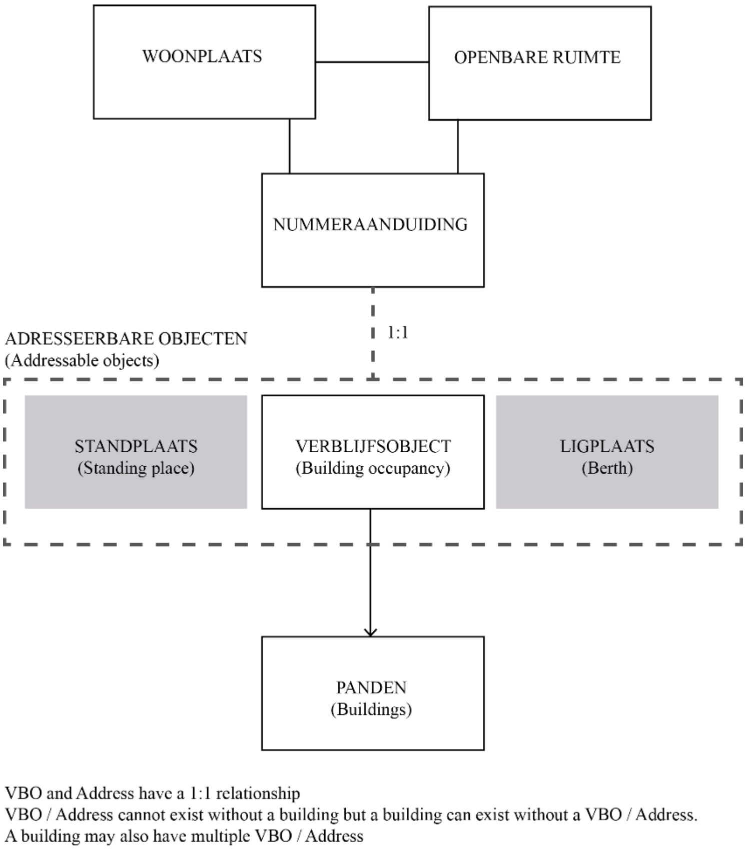


Figure 1 Schematic overview of relationships within the BAG.

The relationship can also be explained through examples. One building may contain a number of VBOs and thus a number of addresses, as in a case of an apartment block. A building may also contain no VBO and thus no address, as in the case of a shed behind a house or farm.

For the EDB V5, the BAG datasets from April 2017 were used.

3.1.1 Requirements

The BAG information requires some processing to only contain the relevant buildings required for the study. This includes filtering historic or inactive records, non-building objects (i.e. standing place and berth) and clipping to the required study area. The BAG information was also further processed to contain clear relationships between the three pieces of relevant BAG information: Address, VBO and Buildings.

The output schema of the three datasets are as follows:

Table 1: BAG address schema.

Address		
Column Name	**Type**	**Description**
nummer_id	Text	Unique identification code assigned per address by BAG
Straatnaam	Text	Street name
Huisnummer	Integer	House number
Huisletter	Text	House letter
toevoeging	Text	Extra house numbers or letters
woonplaats	Text	City
postcode	Text	Postcode
type	Text	Type of *addresseerbaar object* (i.e. addressable object)
vbo_id	Text	Unique VBO identification code related to the address
pand_id	Text	Unique *pand* identification code related to the address

Table 2: BAG VBO schema

VBO		
Column Name	**Type**	**Description**
vbo_id	Text	Unique identification code assigned per VBO by BAG
pand_id	Text	Unique *pand* identification code related to the VBO
gebr_doel1	Text	Main use
gebr_doel2	Text	Secondary use
oppervlak	Double	Usable floor area

Table 3: BAG Building schema

Building		
Column Name	**Type**	**Description**
pand_id	Text (50)	Unique identification code assigned per building by BAG
bouwjaar	Text (30)	Building construction year
shape_area	Double	Footprint area of the building outline

3.1.2 Post-processing and assumptions

To process the BAG dataset prior to use for EDB, the following main steps were taken:

1. Filter out the inactive / historic records from all datasets

The following definition queries are used to filter out the inactive and historic records (including demolished buildings) for each of the datasets:

Address	datum_eind = 0 AND "inactief" = 'N' AND NOT "status" IN ('Naamgeving ingetrokken')
VBO	datum_eind =0 AND "inactief" = 'N' AND NOT "status" IN ('Niet gerealiseerd verblijfsobject', 'Verblijfsobject ingetrokken')
Buildings	datum_eind = 0 AND "inactief" = 'N' AND NOT "status" IN ('Bouwvergunning verleend', 'Niet gerealiseerd pand', 'Pand gesloopt')

These filters are used as suggested from the BAG dictionary [19] to result in only existing addresses, VBO and buildings. Note that the following statuses are included as the building is assumed to still exist:

- *Sloopvergunning verleend* – Demolition permit is granted
- *Verblijfsobject / pand buiten gebruik* – Building is out of use.

2. Filter out the objects outside of the project scope (i.e. standing place and berth)

The address dataset is the only dataset out of the three which contains information about standing place and berth. To filter these out, the following definition query is used:

Address	NOT type IN (‘Ligplaats’, ‘Standplaats’)

This left only addresses for VBO objects.

3. Clip the three datasets to the study area scope.

The three datasets are clipped to the EDB V5 risk assessment boundary.

4. Confirm the relationships of the three datasets (addresses, VBO and buildings).

The filtered address dataset is joined with the filtered VBO dataset to confirm a one to one relationship. The relationship between the VBO and the address dataset is established through the main address.

Establishing a relationship between the VBO and address dataset enables a relationship between the address and building dataset to be created. This has been done by relating the address identification numbers to a building identification number (through VBO). The address dataset is summarised by unique building identification numbers to identify the number of addresses per building. This is then joined to the building dataset and the sum of the number of addresses is checked to confirm that all addresses could be joined to the building dataset.

3.1.3 Results

The study area of the exposure database includes 245,758 active addresses and 257,174 buildings as per the new BAG. Of these buildings, 164,032 buildings have addresses (with several buildings such as apartments blocks containing multiple addresses) leaving a total of 93,142 buildings with no address. These buildings are usually sheds, barns or other secondary buildings related to a building with an address.

3.1.4 Limitations

The BAG dataset is a continuously evolving dataset that is updated as buildings are demolished and built. While it is used as the reference for the existence of buildings and their building outline (as it is the best data available that is governed by the Dutch land registry, Kadaster), it may not have full coverage. This is evident when looking at the results of the updates between the two BAG versions and the addition of several buildings which are not new builds but were missing from the previous set.

Some additional limitations have been observed:

- Small drawing errors are identified in the BAG data: e.g. minor gaps between BAG buildings, where in reality these are adjacent structures, have been observed. Similarly, small overlaps of BAG buildings do exist in the data, where there should not be any overlap.
- Some polygons of historic buildings (demolished or radically renovated) are not removed from the dataset. Additionally, there are some polygons observed that represent underground parking garages while the polygons of the historic buildings are not visible, which makes it difficult to measure the overall influence of the redundant polygons.
- It is observed, when a building might be renovated or extended, that the building year gets updated inconsistently throughout the catalogue.

3.2 Dataland

Dataland provides real estate information based on addresses. Dataland is the precursor of the various governmental *basisregistraties* set up in 2009, including BAG. Dataland therefore still includes similar data and classifications as is in the BAG dataset. Additionally, missing data is also complemented by BAG data. The relation between Dataland and the governmental *basisregistraties* can be seen in Figure 2.

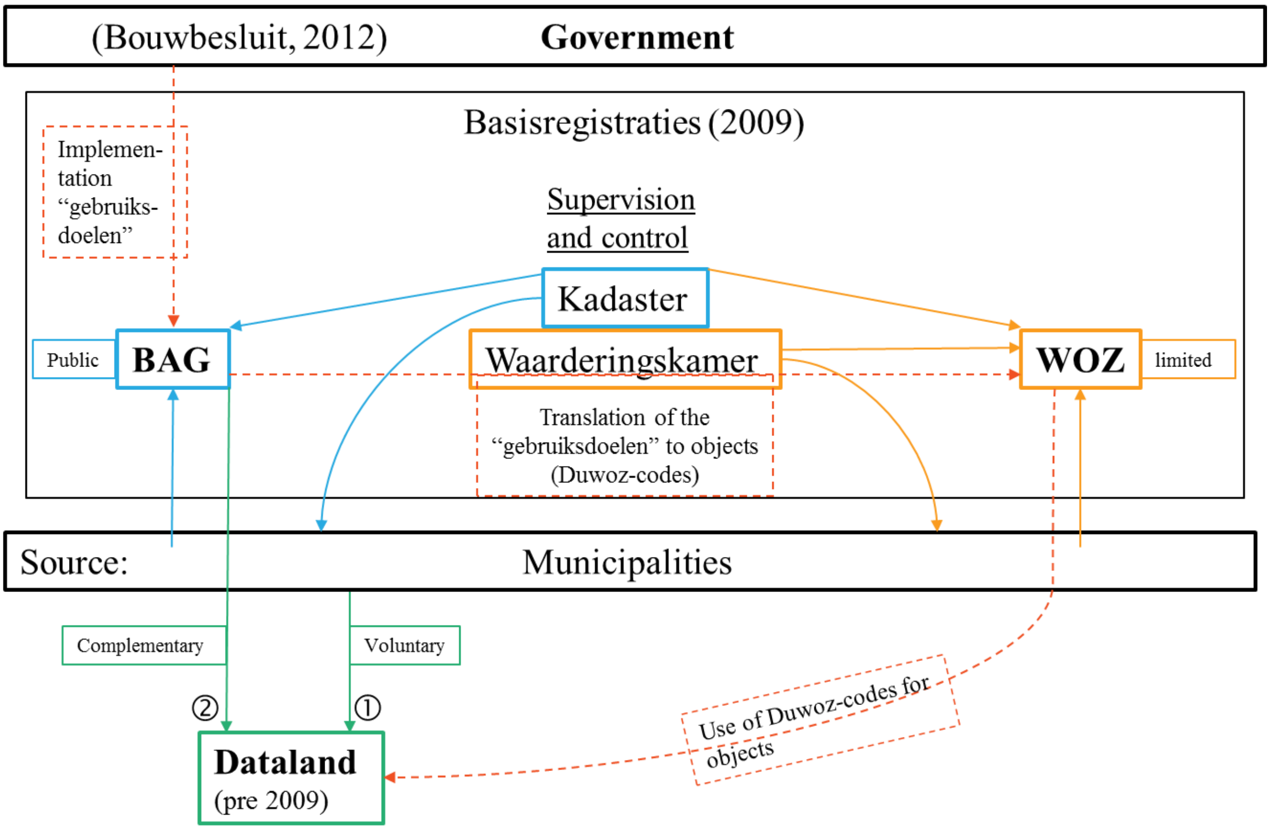


Figure 2: Relation between BAG and Dataland.

3.2.1 Requirements

Table 4 lists the complete available dataset of the Dataland schema. Highlighted fields have been removed as they are either duplicates from another data source (i.e. BAG) or contain limited and unreliable information as described by the provider.

Table 4: Dataland schema.

Column Name	Type	Description
gemeentecode	Numeric	Code name municipality
gemeentenaam	Text	Name municipality
wijkcode	Numeric	Code of neighbourhood
wijkcode omschrijving	Text	Code according StUF-TAX
postcode	Text	Postal ZIP

woonplaatsnaam	Text	Name of village/town in municipalities
straatcode	Numeric	Streetcode according StUF-TAX
straatnaam	Text	Name of street (TNT)
huisnummer	Numeric	Number of house
huisnummertoev	Text	Addition to number of house
huisnummeraand	Text	Empty column
huisletter	Text	Additional letter to number of house
locatieomschrijving	Text	Empty column
identificatie nummeraanduiding	Text	BAG identification number
Type openbareruimte	Text	Type of public space
Indicatie hoofd/nevenadres	Text	H = Hoofdadres, N= Nevenadres
Indicatie geconstateerd adres	Text	Object included only on basis of actual notification sight.
Datum begin geldigheid adres	Text	Date starting address
Datum einde geldigheid adres	Text	Date address ceases to exist
Aanduiding adres in onderzoek	Text	Address registration is under investigation
x-coördinaat	Numeric	x- coordinate RDNEW
y-coördinaat	Numeric	y- coordinate RDNEW
bouwjaar	Numeric	Building year (GFO or BAG)
bouwjaarklasse	Text	Building year period (GFO)
inhoud	Numeric	Volume of building
inhoud bruto/netto	Text	See Data Dictionary
geregistreerd woonoppervlak	Numeric	Area
geregistreerd woonoppervlak bruto/netto	Text	See Data Dictionary
gebruiksoppervlak	Text	See Data Dictionary
oppervlaktewijziging	Numeric	Changes is area
oppervlaktewijzigingsdatum	Text	Date changes in area
geregistreerd niet-woonoppervlak	Numeric	Area not for living
geregistreerd niet-woonoppervlak bruto/netto	Text	See Data Dictionary
geregistreerd oppervlak	Numeric	See Data Dictionary
geregistreerd oppervlak bruto/netto	Text	See Data Dictionary
Aantal woonlagen	Text	Number of storeys
Indicatie geconstateerd adresseerbaar object	Text	Indication object is registered because it is ‘viewed’
Datum begin geldigheid adresseerbaar object	Text	Begin date registration
Datum einde geldigheid adresseerbaar object	Text	End date validity registration

Aanduiding adresseerbaar object in onderzoek	Text	Object registration is under investigation
objectstatus	Text	Status of life cycle of object
objectstatus omschrijving	Text	Status in use
complexinformatie	Text	Describes if object has relationship with another object
bouwkundige bestemming actueel	Numeric	Code for type of building (used for adjacency)
bouwkundige bestemming omschrijving	Text	Type of building
gebruiksklasse	Numeric	Simplified code for use building (Dataland)
gebruiksklasse omschrijving	Text	Building Use
Gebruiksdoel/ligplaats/standplaats	Text	Code for mobile home, mooring
monumentaanduiding	Text	Monument indication, Gov,Prov, Municipality
monumentaanduiding omschrijving	Text	Description monument
soort woonobject	Numeric	Code for use building calculation WOZ
soort woonobject omschrijving	Text	Alternative description for WOZ
gebruik/eigendom status	Text	Code for ownership calculation WOZ
gebruik/eigendom omschrijving	Text	Building owner description
waardepeildatum	Numeric	Date WOZ
waardeklasse	Numeric	Class WOZ
waardeklasse ondergrens	Text	Lower bandwidth WOZ
waardeklasse bovengrens	Text	Upper bandwidth WOZ
waardeontwikkeling	Text	WOZ trend
ozb vrijstelling	Numeric	WOZ waiver
ozb vrijstelling omschrijving	Text	Waiver description
pand	Text	BAG identification
verblijfsobject/ligplaats/standplaats	Text	BAG identification

3.2.2 Post-processing

The inclusion of Dataland datasets into the EDB has been processed following the procedure described below:

1. Adding field names

The fieldnames are added to the dataset using ArcGIS ModelBuilder. The dataset in csv format received are opened in ArcGIS as a table. In the ModelBuilder the table-to-table tool are used to store the table in a FGDB (file geodatabase). In the table-to-table tool the fieldnames are copied from the csv-file into the field map window. This process lead to the fieldnames being stored as aliases.

2. Join to BAG and delete duplicates

The received Dataland relates to the BAG version of March 2015. This version of BAG is used to establish a relationship between the datasets and clean any duplicates.

To join the Dataland dataset to the BAG addresses, the field 'identification_nummeraanduiding' can be joined to the field 'nummer_id' from the BAG dataset. The field 'identification_nummeraanduiding' appears to have duplicate values and null values. According to Dataland, the duplicates were wrongly provided by the municipalities. This occurs when there is more than one WOZ registration at the same address.

A visual check on the duplicate values of identification_nummeraanduiding shows that they are identical across the entire row. Duplicates are therefore deleted using the Model Builder tool: delete_identical.

The unknowns (nulls) are also the result of coupling the WOZ database to the dataset. Because WOZ is not used in the GIS analyses the records are filtered out by using following expression:

identificatie_nummeraanduiding	<> '\N'

Where there is an identification_nummeraanduiding (address id) but no pand_id (i.e. building_id), these are filtered out as they were found to be representative of boat houses, mobile homes and caravans.

3.2.3 Results

As visible in Figure 3, the total extent of the dataset includes the municipalities: Aa en Hunze, Appingedam, Assen, Bedum, Bellingwedde, De Marne, Delfzijl, Eemsmond, Groningen, Grootegast, Haren, Hoogezand-Sappemeer, Leek, Loppersum, Menterwolde, Noordenveld, Oldambt, Pekela, Slochteren, Stadskanaal, Ten Boer, Tynaarlo, Veendam, Winsum, Zuidhorn.

Figure 3: Geographic extent of Dataland information used.

Table 5 shows the comparison of the resulting Dataland dataset to the EDB V5 version of BAG (April 2017) within the scope area.

Table 5: Comparison of address and building counts of BAG (2017) and Dataland (2015) within the EDB V5 Risk Assessment Boundary.

BAG		Dataland	
Buildings	**Addresses**	**Buildings**	**Addresses**
257,174	249,096	156,327	235,568

3.2.4 Limitations

It is worth noting that not all buildings in the EDB study area are included in the Dataland dataset, and many records in the dataset have unknown or missing values for the building description, from which the use and objects are usually derived. In addition, Dataland only provides information on buildings with an address so that buildings without an address do not have corresponding Dataland information.

3.3 Parcel

The parcel data was received from Kadaster on the 25th of August 2015. No processing was done to the raw dataset and this section only describes the extent and content of the dataset received.

3.3.1 Data

The data delivered is a polygon shape file containing 261,244 features with the following relevant schema:

Table 6: Parcel schema

Column Name	Type	Description
AKR_UID	Double	The unique identification number per parcel as given by Kadaster

3.3.2 Extent

The extent of the parcel data provided by Kadaster (through NAM) is shown in Figure 4.

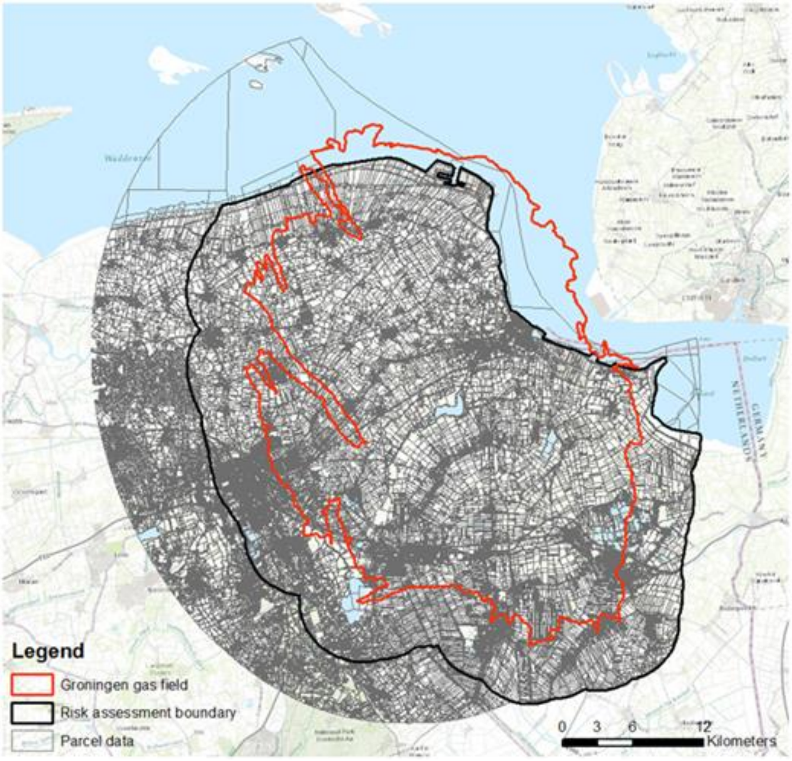


Figure 4: Geographic extent of Parcel.

3.3.3 Limitations

The dataset is mainly used for its parcel polygons. While the extents of the parcel information cover the scope area, there are gaps in the coverage.

3.4 Vs30

The *Vs30* information was received in a GIS format from Deltares[1] [5]. No processing has been done to the *Vs30* data and this section only describes the extent and content of the dataset.

3.4.1 Data

The data delivered was a polygon shape file with the following schema:

Table 7: Vs30 schema.

Column Name	Type	Description
MEAN_Vs	Double	The mean Vs30 value based on the GSG model [5].
STD_Vs	Double	The standard deviation based on the GSG model [5].
GEOL_AREA	Integer	Defined geological areas.

3.4.2 Extent

The extent of the *Vs30* information as provided by Deltares corresponds to a 5km buffer around the Groningen field as shown in Figure 5.

[1] Deltares is an independent institute for applied research in the field of subsurface including seismology: www.deltares.nl

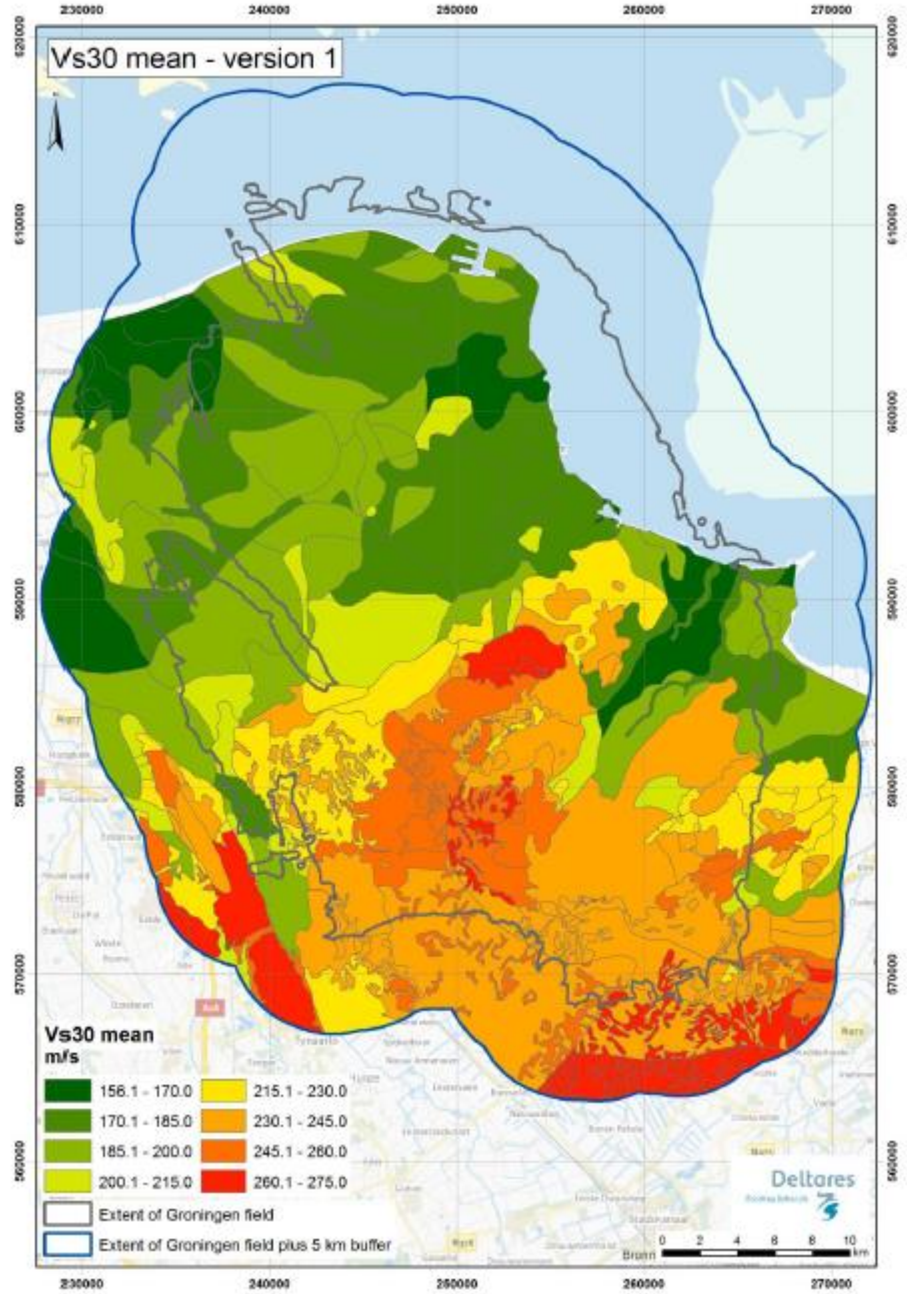


Figure 5: Geographic extent of Vs30 [5].

3.5 AHN

The latest version of the *Actueel Hoogtebestand Nederland* or AHN [7] is a detailed height model for the Netherlands obtained through laser altimetry. There are different models available in different resolutions. The terrain model just describes the height of the terrain excluding any object on the terrain. The complete model includes all objects such as buildings, vegetation, cars, etc. Arup uses both the terrain model and the complete model with a of resolution of 0.5m x 0.5m.

The actuality of the data for the study area is 2009 (i.e. AHN data for the study area is not available post 2009).

3.5.1 Requirements

The height data on buildings (from the complete height model) is required for determining the gutter height, number of storeys and information on the roof. The height data from vegetation however is not required for the analysis and causes noise in the height data when looking at buildings. For instance in the case of overhanging trees (see Figure 6).

The terrain model is also used because it provides information of the height of the terrain directly surrounding the buildings. To calculate the relative height of buildings (the top of the building minus the ground height at the base level of the building) the ‘no data’ cells in the terrain model need to be extrapolated.

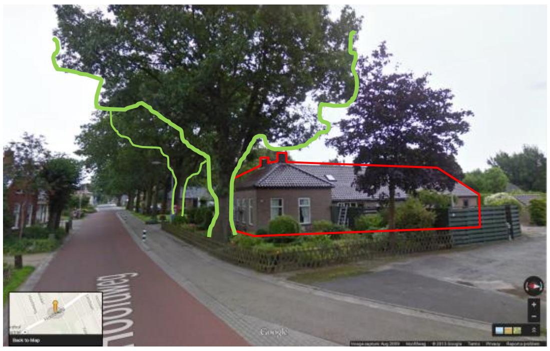



Figure 6 Example of overhanging trees creating noise in the height data.

3.5.2 Post-processing

The terrain model and the complete model filtering can be separated into vegetation filtering and the creation of the terrain grid.

3.5.2.1 Vegetation filtering

To filter the vegetation from the unfiltered height data the method described by Hewett [20] was applied.

1. Preparing the data

To reduce computation time, the unfiltered AHN grid of raster data is first clipped to the area covered by building outlines.

Since the height of the ground surrounding the buildings is also required for further analysis, a buffer is applied before the height raster is clipped to the BAG building outline polygons.

2. Data enhancement and simplification

After the raster has been prepared for filtering, the height data of the raster is divided by 2.5m, which is the assumed approximate floor-to-floor height. This provides a raster that is the approximate number of storeys. By converting the number of storeys raster to integer values, the cells can be grouped by their number of storeys.

3. Extract to polygon and clean-up

Next, the raster is converted to polygons. These polygons are then selected by a nominated minimum building area value of $3m^2$. This makes sure the small polygons caused by vegetation are removed.

3.5.2.2 Terrain grid

To create a terrain grid the following steps are taken:

1. Extrapolating empty cells

Extrapolating ‘no data’ cells from the filtered AHN grid is done by applying a raster cell calculation. For any cell without a value the mean is taken within a three-cell radius. This step is repeated until all the ‘no data’ cells have been assigned a value.

2. Calculating the base height of a building

The mean is calculated for all (extrapolated) cells which intersect with a building outline. This value is then assigned to all these cells.

3.5.3 Results

The final is a height raster with most of the noise removed that is caused by vegetation, as visible in Figure 7.

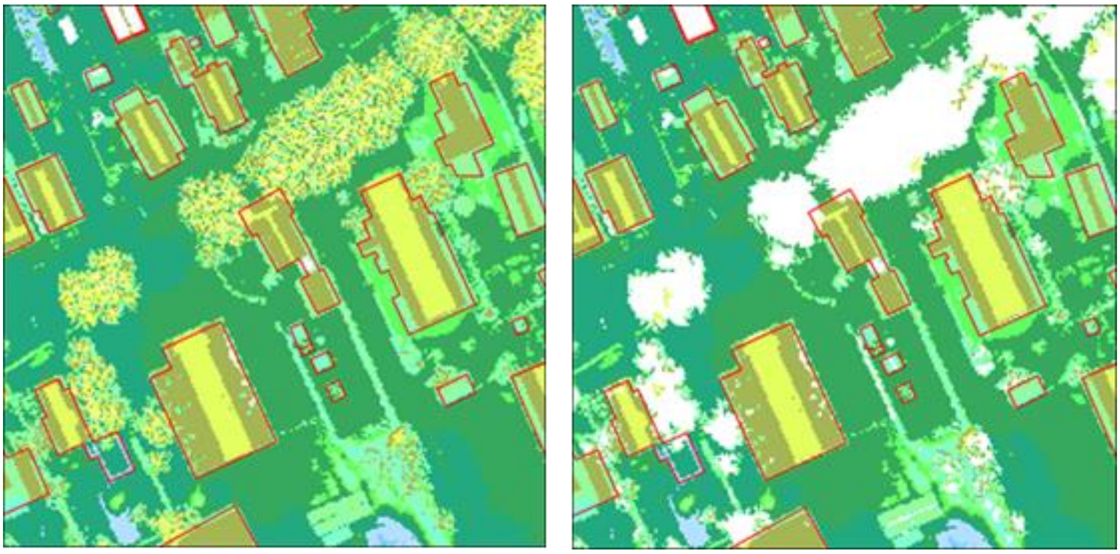


Figure 7 Building outlines overlaying height data before (left) and after (right) filtering. Note: in this example the raster data is not clipped to the building outline.

3.5.4 Limitations

The height data from AHN2, does not have data for buildings built after 2009. In addition, some height data can be missing for buildings, for example because of overhanging trees. This results in 3,829 buildings not having AHN2 height data.

3.6 Rijksmonumenten

3.6.1 Data

The '*Rijksmonumenten-kaartlagen'* dataset provided by the Rijksdienst voor het Cultureel Erfgoed (RCE) has been formatted into Table 8 below.

Table 8: RCE Schema

Rijksmonumenten		
Column Name	**Type**	**Description**
RIJKSMONNR	Number	Number Rijksmonument
NAAM	Text	Name of monument
TYPEMONUM	Text	Building or archaeological site of garden
CBSCATEGOR	Text	Description used by CBS
CBSCODE	Number	Code used by CBS
OORSPRFUNC	Text	Original function
SUBCATNR	Number	Code to specify description
SUBCATOMS	Text	More specific description
HFDCATCODE	Number	Main category code
HFDCATOMS	Text	Main category description
TYPECHOBJ	Text	Type description
BEGBOUWJR	Date	Year of start construction
EINDBOUWJR	Date	Year of end construction
GRS_DATUM	Number	-
INSCHRDAT	Number	Date of registration as monument
GEMEENTE	Text	Municipality
GEMEENTENR	Number	Number of municipality
PROVINCIE	Text	Province
PROVCODE	Text	Code for Province
PLAATS	Text	Place name
SITUERING	Text	Info about location to main object
STRAAT	Text	Street name
HUISNUMMER	Number	House number
TOEVOEGING	Text	Additional house number
POSTCODE	Text	Postal ZIP
BAG_PLAATS	Text	Place name according BAG
X_COORD	Number	RD_X_Coordinate
Y_COORD	Number	RD_Y_Coordinate
COORDHERK	Text	Source coordinates
KICH_URL	Text	URL with official document description monument
STATUS	Text	Protection status monument
EXTRACTDAT	Text	Date of last update

3.6.2 Post-processing

The RCE dataset has no BAG ID (whether it be address, VBO or building) since monuments are not only buildings. Where the monument listed is a building, the coordinates listed are usually situated in the location of the building.

The following steps were carried out to assign the RCE dataset to a building ID:

1. The '*Rijksmonumenten-kaartlagen'* are downloaded and clipped to the EDB V5 study area.
2. Filter on 'church-like' buildings as the building types of interest for the EDB.

   Monuments which were buildings and 'church-like' were identified using the expression:

'Church-like' buildings	'"TYPEMONUM" = 'Geb' AND ("CBSCATEGOR" = 'Kerk-onderdl./object' OR "CBSCATEGOR" = 'Kerkelijke gebouwen'

3. Spatially join RCE information to buildings from BAG

   It appeared that some buildings were part of several monument-categories. For instance, a church can have a monument status because of its building, or/and because of a bell-tower or organ or other part of it. Therefore, a one-to-many type of spatial join was used. However, several monuments could not be joined to the BAG, due to the absence of any coordinates located inside the identified building polygon or into an erroneous one:

   - The location of the monument was a house nearby that serves as an administrative address.
   - The monument consisted of several elements, for example the church and a tower or the church and the rectory, but only one of these buildings indicated the location of the monument.

4. **Google Streetview validation of the church-like monuments**

   The monuments identified were validated through Google Streetview [15] and where possible, the location of the point representing the monument was corrected by moving the point to the church or by duplicating the point for each building related to the monument. After these edits, the RCE dataset was then spatially joined to the buildings from BAG again to be assigned a building ID.

5. **The 'church-like' buildings which were rectories were filtered out using the expression below.**

'Church-like' buildings	"SUBCATOMS" IS NOT 'Kerkelijke dienstwoning'

3.6.3 Results

There are 166 monumental church structures in the scope area.

3.6.4 Limitations

The RCE provide only a list with national monuments. There are also local monuments, registered by municipalities although there is no central institution that issues and maintain this information. The following limitations have been observed:

- The data from the RCE is of good quality but by joining it to the BAG, it appears that a few monuments could not be joined. A thorough scan of these cases showed that it concerned mostly monuments that are not physical buildings. Other reasons might be that the structure has no BAG ID (i.e. it is a ruin) or does not exist anymore.

- The point location of some monuments is manually moved so that it falls inside a building-polygon. If this data is updated one should be aware of the corrections: e.g. the church of Loppersum, which is located on Kerkpad 8 in the RCE file but moved to the actual location in the post-processed dataset.

3.7 Nationale Atlas Volksgezondheid

The Rijksinstituut voor Volksgezondheid en Milieu (RIVM) provides information about public health and environment through their online platform '*Nationale Atlas Volksgezondheid'* (1).

3.7.1 Data

The dataset was provided in a CSV format and contains the schema presented in Table 9:

Table 9: RIVM Schema.

Column Name	Type	Description
Ziekenhuisnummer	Text	Hospital unique ID from RIVM
Organisatienummer	Text	Organization unique ID from RIVM
Naam organisatie	Text	Name Organisation
Naam ziekenhuis	Text	Name Hospital
Soort ziekenhuis	Text	Type Hospital
Adres	Text	Address
Postcode	Text	Postal Zipcode
Postcode	Text	Concatenated Postal Zipcode
Plaats	Text	Municipality

3.7.2 Post-processing

The RIVM dataset has no BAG ID (whether it be address, VBO or building) but was assigned a building id using the following steps:

1. The data was cleaned and formatted.
2. The BAG ID was joined to the dataset based on the following query:

RIVM .BAG BUILDING ID	select BAG BUILDING ID from BAG, RIVM where BAG.HUISNUMMER = RIVM.HUISNUMMER AND BAG. HUISLETTER = RIVM. HUISLETTER AND BAG. TOEVOEGING = RIVM. TOEVOEGING AND BAG. POST CODE = RIVM. POST CODE

3. Lastly, the table was clipped to the EDB study area.

3.7.3 Results

The RIVM dataset contains 8 hospitals within the EDB study area.

3.7.4 Limitations

Hospitals are often described with a single address in the RIVM dataset although the hospital function may be hosted in several adjacent structures. Therefore, when matching the address to a BAG ID, not all buildings belonging to the hospital district may be identified as having a hospital function.

3.8 Basisregister Instellingen

The Dienst Uitvoering Onderwijs (DUO) manages the *'Basisregister Instellingen'* [10], which contains information on all schools in the Netherlands. This dataset is used to identify primary schools, high schools, special educational schools, vocational schools, colleges and university buildings. The School Registry is updated every month and can be extracted per province.

3.8.1 Data

The data is provided in an Excel spreadsheets format. The schema of the output is presented in the following Table 10:

Table 10: DUO Schema.

Column Name	Type	Description
PROVINCIE	Text	Province
BRIN_NUMMER	Text	Unique code per school
VESTIGINGSNUMMER	Text	Unique code per school location
NAME	Text	Name of the school
STRAATNAAM	Text	Street name
HUISNUMMER	Number	House number
TOEVOEGING	Text	Additional house number
HUISLETTER	Text	Additional house letter
POSTCODE	Text	Postal ZIP
GEMEENTENAAM	Text	Municipality
DENOMINATIE	Number	Denomination

3.8.2 Post-processing

As the DUO dataset had no BAG ID (whether it be address, VBO or building), it had to be assigned a BAG ID. It had to be further post-processed manually as some colleges and university only had one main address to represent multiple buildings. The following steps were performed:

1. The data are cleaned and the format is aligned to BAG in ArcGIS.
2. Schools located in municipalities outside the research scope are filtered out.
3. The BAG ID is joined to the DUO dataset using the following query:

DUO.BAG BUILDING ID	select BAG BUILDING ID from BAG, DUO where BAG.HUISNUMMER = DUO.HUISNUMMER AND BAG. HUISLETTER = DUO. HUISLETTER AND BAG. TOEVOEGING = DUO. TOEVOEGING AND BAG. POST CODE = DUO. POST CODE

4. The schools with an BAG ID are joined back to the total list of schools. Schools with no BAG ID are identified and manually checked.
5. Schools with no BAG ID in the municipalities on the edge of the research scope are checked based on post code and filtered out.

6. Schools in municipalities within the research scope are checked in the BAG viewer and edited to align it to the right BAG address.
7. Again, the BAG ID is joined to the remaining dataset based on the following query:

DUO.BAG BUILDING ID	select BAG BUILDING ID from BAG, DUO where BAG.HUISNUMMER = DUO.HUISNUMMER AND BAG. HUISLETTER = DUO. HUISLETTER AND BAG. TOEVOEGING = DUO. TOEVOEGING AND BAG. POST CODE = DUO. POST CODE

8. Finally, the two output datasets are appended together and clipped to the research scope area.

3.8.3 Results

The DUO dataset contains the 395 educational facilities within the EDB scope area.

3.8.4 Limitations

When joining the BAG ID's some gaps in the datasets have been identified. In particular, part of the buildings do not have a BAG ID assigned due to mismatching address details. Additional limitations have also been found during the process:

- The DUO dataset includes colleges and a university. If these buildings meet the definition of a school can be questioned. These buildings could be structurally different and the behaviour of the population of the buildings could also be different if compared to primary or high schools.
- Larger educational institutions such as universities are often spread across several buildings but described with a single address. When matching the address to a BAG ID, not all educational buildings may be identified.

4 Project Data

4.1 RVS Inspection Data

The Rapid Visual Screening (RVS) is a preliminary building assessment process designed to collect building information from the public realm (without entering the property boundaries), with a clear focus on high risk building elements. The information collected is being used to assess the seismic risk of buildings, and building elements, and to prioritise them further for more detailed assessments and/or mitigation measures.

The specific objectives of the screening are:

- Perform a safety assessment, to evaluate the safety of inhabitants and to ensure safe conditions for the inspectors to carry out the RVS screening;
- Identify external High Risk Building Elements (HRBE's), such as chimneys and parapets, which could pose a life safety risk during a seismic event;
- Gathering of additional information on site that will allow to evaluate the building performance during a seismic event.

In addition to the above, building information is collected based on specific requests of stakeholders (e.g. possibility to fit solar panels).
The outcome of the RVS assessment is used to prioritise the follow up actions, consisting of either measures on individual HRBE's or an Extended Visual Screening (EVS).

The method of acquiring RVS inspection data is described in more detail in the dedicated literature [11] [13].

4.1.1 Data

Table 11 gives a description of the data provided to the inspectors and collected during the RVS process. Note that it is not the full dataset, but an extract of it which were made available after it was imported into the project database.

Table 11: RVS Schema.

Field ID	Field Name	Source	Field Type	Field Description	Comments
001	Address	GIS	Text	House number-Additional House Number Street/Public Space_City Name	
002	Unique Reference	GIS	Text	Address ID_Premise ID	
003	Status	RVS	List	Status of the record in the inspection process.	
004	Street/Public Space	GIS	Text	Street name.	Source: BAG (03/2015)
005	House Number	GIS	Integer	House number.	Source: BAG (03/2015)
006	Additional House Number	GIS	Text	Extra house numbers or letters.	Source: BAG (03/2015)
007	Postcode	GIS	Text	Postal code.	Source: BAG (03/2015)
008	House Number (manual)	RVS	Integer	Corrected house number, as assessed during inspection.	
009	City Name	GIS	Text	Name of village/town in municipalities.	Source: BAG (03/2015)
010	Additional House Number (manual)	RVS	Text	Corrected extra house numbers or letters, as assessed during inspection.	
011	Postcode (manual)	RVS	Text	Corrected postal code, as assessed during inspection.	
012	City Name (manual)	RVS	List	Corrected name of village/town in municipalities, assessed during inspection.	
013	Street/Public Space (manual)	RVS	Text	Corrected street name, as assessed during inspection.	
014	Recommendation Engineer	RVS	List	Recommendation given by the engineer after review of the inspection report.	
015	Facade Height (YL)	RVS	Double	Height (m) of the left façade.	Pre 06/2014 the average height was recorded, later the maximum height.
016	Facade Length (YL)	RVS	Double	Maximum length (m) of the left façade.	
017	Total Area of Facade(s) (YL)	RVS	Double	Total area of the left façade, including openings.	
018	Certainty (YL)	RVS	Double	Indicates the certainty with which the inspector could provide the data regarding the left façade.	
019	Facade Height (YR)	RVS	Double	Height (m) of the left façade.	Pre 06/2014 the average height was recorded, later the maximum height.

020	Facade Length (YR)	RVS	Double	Maximum length (m) of the right façade.	
021	Total Area of Facade(s) (YR)	RVS	Double	Total area of the right façade, including openings.	
022	Certainty (YR)	RVS	Double	Indicates the certainty with which the inspector could provide the data regarding the right façade.	
023	Facade Height (XB)	RVS	Double	Height (m) of the street facing (front) façade.	Pre 06/2014 the average height was recorded, later the maximum height.
024	Facade Length (XB)	RVS	Double	Maximum length (m) of the rear façade.	
025	Total Area of Facade(s) (XB)	RVS	Double	Total area of the rear façade, including openings.	
026	Certainty (XB)	RVS	Double	Indicates the certainty with which the inspector could provide the data regarding the rear façade.	
027	Facade Height (XF)	RVS	Double	Height (m) of the rear façade.	Pre 06/2014 the average height was recorded, later the maximum height.
028	Facade Length (XF)	RVS	Double	Maximum length (m) of the street facing (front) façade.	
029	Total Area of Facade(s) (XF)	RVS	Double	Total area of the street facing (front) façade, including openings.	
030	Certainty (XF)	RVS	Double	Indicates the certainty with which the inspector could provide the data regarding the street facing (front) façade.	
031	X Front Inspection possible	RVS	List	Indicates the possibility of inspecting the street facing (front) façade.	
032	Openings (YL)	RVS	Integer	Opening percentage of the left façade, considered for the most unfavourable shear-plan at ground floor.	
033	Openings (YR)	RVS	Integer	Opening percentage of the right façade, considered for the most unfavourable shear-plan at ground floor.	
034	Openings (XB)	RVS	Integer	Opening percentage of the rear façade, considered for the most unfavourable shear-plan at ground floor.	
035	Openings (XF)	RVS	Integer	Opening percentage of the street facing (front) façade, considered for the most unfavourable shear-plan at ground floor.	
036	X Front Reason if not possible	RVS	List	Reason why the inspection of the street facing (front) façade was not possible.	

037	Y Right Inspection possible	RVS	List	Indicates the possibility of inspecting the right façade.	
038	Y Right Reason if not possible	RVS	List	Reason why the inspection of the right façade was not possible.	
039	X Back Inspection possible	RVS	List	Indicates the possibility of inspecting the rear façade.	
040	X Back Reason if not possible	RVS	List	Reason why the inspection of the rear façade was not possible.	
041	Y Left Inspection possible	RVS	List	Indicates the possibility of inspecting the left façade.	
042	Y Left Reason if not possible	RVS	List	Reason why the inspection of the left façade was not possible.	
043	Walls out of plane	RVS	List	Indicates the presence of a HRBE 1, assessed during inspection.	
044	Recommendation Wall Out of Plane	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
045	Column crack(s) or slenderness issues	RVS	List	Indicates the presence of a HRBE 2, assessed during inspection.	
046	Recommendation column crack(s) or slenderness issues	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
047	Wall cracks	RVS	List	Indicates the presence of a HRBE 3, assessed during inspection.	
048	Recommendation wall cracks	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
049	Deflected lintels	RVS	List	Indicates the presence of a HRBE 4, assessed during inspection.	
050	Recommendation Deflected Lintels	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
051	Wall ties damage	RVS	List	Indicates the presence of a HRBE 12, assessed during inspection.	
052	Recommendation wall ties	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
053	Balcony(s) present	RVS	List	Indicates the presence of a HRBE 6, assessed during inspection.	
054	Parapet(s) present	RVS	List	Indicates the presence of a HRBE 6, assessed during inspection.	
055	Cantilevered elements present	RVS	List	Indicates the presence of a HRBE 6, assessed during inspection.	
056	Canopy(s) present	RVS	List	Indicates the presence of a HRBE 6, assessed during inspection.	
057	Recommendation Balcony-Parapets-Canopies	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	

058	Slender chimney(s) present	RVS	List	Indicates the presence of a HRBE 7, assessed during inspection.	
059	Recommendation Slender Chimneys	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
060	Damaged chimney(s) present	RVS	List	Indicates the presence of a HRBE 8, assessed during inspection.	
061	Recommendation Damaged Chimney(s)	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
062	Unsafe roof cladding	RVS	List	Indicates the presence of a HRBE 9, assessed during inspection.	
063	Recommendation unsafe roof cladding	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
064	Mortar damage	RVS	List	Indicates the presence of a HRBE 10, assessed during inspection.	
065	Masonry dormer(s) present	RVS	List	Indicates the presence of a HRBE 11, assessed during inspection.	
066	Recommendation Dormers	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
067	Lack of ties in cavity walls	RVS	List	Likelihood of the presence of adequate wall ties within cavity walls.	Assumption based on building year; a lack of ties is assumed prior 1991.
068	Recommendation Mortar Damage	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
069	Recommendation lack of ties in cavity walls	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
070	Other damages	RVS	List	Indicates the presence of a HRBE 13, assessed during inspection.	
071	Other Recommendations	RVS	List	Given the extent of the encountered HRBE a recommendation for further action is given.	
072	Inspection Possible	RVS	List	Indicates if it was possible to carry out the inspection.	
073	Abandonment	RVS	List	Indicates if an object is out of use.	
074	Reason inspection not performed	RVS	List	Reason why the inspection was not possible.	
075	Address ID	GIS	Text	Unique identification code assigned per address by BAG.	Source: BAG (03/2015)
076	Premises ID	GIS	Text	Unique identification code assigned per building by BAG.	Source: BAG (03/2015)
077	Latitude (Y)	GIS	Double	Y coordinate of address point.	Y coordinate in WGS84

078	Longitude (X)	GIS	Double	X coordinate of address point.	X coordinate in WGS84
079	PGA	GIS	Text	PGA value.	Source: Shell P&T PGA (09/2013)
080	Address Use 1	GIS	Text	Main use of the building.	Source: (batch 1-7) Dataland, (batch 8ff.) Source: BAG (03/2015)
081	Address Use 2	GIS	Text	Secondary use of the building.	Source: (batch 1-7) Dataland, (batch 8ff.) Source: BAG (03/2015)
082	BAG Address Use	GIS	Text	Main use of the building.	Source: BAG (03/2015)
083	Status of Premises	GIS	Text	Status of lifecycle of building (i.e. from planned to demolished).	Source: BAG (03/2015)
084	Building Year	GIS	Integer	Building construction year.	Source: BAG (03/2015)
085	Importance Class	GIS	Text	Classification according to Eurocode 8, depending on the consequences of collapse for human life and the importance of the building for public safety.	
086	Occupancy Class	GIS	Text	Assumed population classification.	Source: Bridgis (March 2013)
087	Address Use 1 (manual)	RVS	List	Corrected main use, as assessed during inspection.	
088	Address Use 2 (manual)	RVS	Text	Corrected secondary use, as assessed during inspection. (obsolete)	
089	Address Use 2 (list)	RVS	List	Corrected secondary use, as assessed during inspection.	
090	BAG Address Use (manual)	RVS	List	Corrected main use, as assessed during inspection.	
091	Status of Premises (manual)	RVS	List	Corrected status of lifecycle of building, as assessed during inspection.	
092	Building Year (manual)	RVS	Integer	Corrected building construction year, as assessed during inspection.	Corrections were rounded at 5 years.
093	Importance Class (manual)	RVS	List	Corrected Eurocode 8 classification, as assessed during inspection.	
094	Occupancy Class (manual)	RVS	List	Corrected population classification, as assessed during inspection.	
095	Main Wall Material	GIS	Text	Main construction material of the outer walls.	Where building year < 1960 = wood and ≥ 1960 = concrete.
096	Main Wall Material (manual)	RVS	List	Corrected main construction material of the outer walls, as assessed during inspection.	

097	Ground Floor Material	GIS	Text	Construction material of the ground floor.	Where building year < 1960 = wood and ≥ 1960 = concrete.
098	Ground Floor Material (manual)	RVS	List	Corrected construction material of the ground floor, as assessed during inspection.	
099	Higher Floor Material	GIS	Text	Construction material of upper floors.	Where building year < 1960 = wood and ≥ 1960 = concrete.
100	Higher Floor Material (manual)	RVS	List	Corrected construction material of the upper floors, as assessed during inspection.	
101	Area Building Footprint	GIS	Double	Area (m2) of the building outline polygon.	Source: BAG (03/2015)
102	Building Height	GIS	Text	Height (m) of the building.	Source: Algemeen Hoogtebestand Nederland (2009)
103	Number of Storeys	GIS	Double	Number of building layers.	Calculated as total building height / 3.31.
104	Area Building Footprint (manual)	RVS	Double	Corrected area (m2) of the 'main' building outline polygon, as assessed during inspection.	
105	Building Height (manual)	RVS	Text	Corrected height (m) of the building, as assessed during inspection.	Sometimes small height differences compared to the value provided by GIS are registered. There are two possible reasons for this: 1) terraced houses and semi-detached houses are made identical 2) subtraction of chimney heights >1m
106	Number of Storeys (manual)	RVS	Double	Corrected number of building layers, as assessed during inspection.	A building consisting of one storey means a building in which only the ground floor provides habitable space.
107	Horizontal Irregularity	RVS	List	Horizontal or plan irregular structures are those in which seismic response is not only translational but also torsional, and is a result of stiffness and/or mass eccentricity in the structure.	
108	Vertical Irregularity	RVS	List	Changes in structural system along the height, changes in story height, setbacks, changes in materials and unanticipated participation of non-structural components.	
109	Storey Height	RVS	Double	Average height (m) of a storey.	

110	Basement present	RVS	List	Indicates the presence of a basement, as assessed during inspection.	
111	Emergency confirmed?	RVS	List	Second opinion by the engineer regarding the emergency situation, after review of the inspection material.	
112	Safety Label	RVS	List	Indicates to what extent the encountered situation poses a possible safety risk for inhabitants and/or the inspector.	
113	Emergency Situation	RVS	List	Indicates if the encountered situation poses an immediate safety risk for the inhabitants and/or the inspector.	
114	Emergency intervention	RVS	List	Planned time frame for emergency interventions.	
115	S-Curve Value In Plane	RVS	Double	Based on information collected during the inspection, this automatically calculated value indicates the in plane seismic fragility of the building.	
116	S-Curve Value Out of Plane	RVS	Double	Based on information collected during the inspection, this automatically calculated value indicates the out of plane seismic fragility of the building.	
117	S-Curve version	RVS	Text	Version of the S-score calculation tool to be used.	Dependent on the available input parameters.
118	Presence Of Adequate Wall Ties	RVS	List	Indicates the presence of a HRBE 5, assessed during inspection.	
119	Joint Structure	RVS	List	Structure with at least one shared wall between building parts.	
120	Foundation Type	RVS	List	Type of foundation.	
121	Structure Certainty	RVS	List	Indicates the certainty with which the inspector could provide the data regarding the structural characteristics.	
122	Roof Type	RVS	List	Roof construction type.	
123	Cavity Walls	RVS	List	Indicates if the façade consists of cavity walls.	
124	Presence of Wall Floor/Roof Ties	RVS	List	Visual observation from the outside of wall/roof ties.	As opposed to HRBE 5, this value is used for the calculation of the S-score.
125	Wall Thickness	RVS	Double	Total wall thickness.	
126	Thickness of Inner Leaf	RVS	Double	Thickness (mm) of the inner leaf of cavity walls.	
127	Thickness of Outer Leaf	RVS	Double	Thickness (mm) of the outer leaf of cavity walls.	

128	Maintenance Level	RVS	List	Impression of the general status of maintenance of the building structure.	
129	Deterioration of masonry	RVS	List	Indicates the extent to which the masonry is considered to be deteriorated.	
130	Deterioration of mortar over joints	RVS	List	Indicates the extent to which the mortar is considered to be deteriorated.	
131	Deterioration of concrete elements	RVS	List	Indicates the extent to which concrete elements are considered to be deteriorated.	
132	Deterioration of metal connections	RVS	List	Indicates the extent to which metal connections are considered to be deteriorated.	
133	Deterioration of wooden elements	RVS	List	Indicates the extent to which wooden elements are considered to be deteriorated.	
134	Solar Cells Present	RVS	List	Indicates the presence of solar cells.	
135	Roof suitable for Solar Cells	RVS	List	Indicates the possibility of installing solar cells on the roof.	
136	Consequence class	GIS	Text	The possible consequences of failure in terms of risk to life, injury, potential economic losses.	Replaces the previously used ‘Importance Class’.
137	Consequence class (manual)	RVS	List	Corrected consequence class as assessed during inspection.	
138	Slender chimney(s) - Height [m]	RVS	Double	Height (m) of the most unfavourable slender chimney.	
139	Slender chimney(s) - Short side [m]	RVS	Double	Short side in cross section (m) of the most unfavourable slender chimney.	
140	Slender chimney(s) - Long side [m]	RVS	Double	Long side in cross section (m) of the most unfavourable slender chimney.	
141	HRBE 7 - Slenderness ratio	RVS	Double	Calculated as chimney height/short side.	
142	Sloped chimney flue - Probability	RVS	List	Indicates the probability of a sloped chimney flue inside the building, based on observations regarding roof shape, location of chimneys and the facades.	
143	Sloped chimney flue - Additional recommendations	RVS	List	Given the probability of a sloped chimney flue inside the building, a recommendation for further action is given.	

4.1.2 Results

The number of buildings with RVS inspections that was included in EDB V5 was 15,859.

4.1.3 Data limitations

As per definition, the RVS is designed to collect building information from the public realm. This restriction reflects on the type of data that can be collected as well as on the confidence level. In addition, the aim of RVS activities give priority to safety rather than to an accurate evaluation of each attribute into a building that often is simply assumed, introducing a not reproducible nor standard variation due to the expertise of the inspectors and the status of the premises.

The inspection itself has so far been assisted by inspection tools. These tools have undergone a series of changes since their first release. These changes and the confidence of the collected inspection data have been summarized in Table 12. Together they give an indication about the quality of the collected information.

Table 12: Changes and the confidence of the collected RVS inspection data.

Field ID	Field Name	Source	Confidence	Dependencies	Date introduced	Changes made?	Type of change	Change description
001	Address	GIS			released 12/2013	No		
002	Unique Reference	GIS			released 12/2013	No		
003	Status	RVS	n/a		released 12/2013	Yes	Extra / modified choice list items	
004	Street/Public Space	GIS			released 12/2013	No		
005	House Number	GIS			released 12/2013	No		
006	Additional House Number	GIS			released 12/2013	No		
007	Postcode	GIS			released 12/2013	No		
008	House Number (manual)	RVS	Observed		released 12/2013	No		
009	City Name	GIS			released 12/2013	No		
010	Additional House Number (manual)	RVS	Observed		released 12/2013	No		
011	Postcode (manual)	RVS	n/a		released 12/2013	No		
012	City Name (manual)	RVS	Observed		released 12/2013	No		
013	Street/Public Space (manual)	RVS	Observed		released 12/2013	No		

014	Recommendation Engineer	RVS	n/a		released 12/2013	No		
015	Facade Height (YL)	RVS	manual	018	released 19/03/2014	Yes	Inspection instructions	Changed to maximum wall height; in the period before the change (approx. starting June 2014) this field isn't used, since not considered in the S-score calculation v8.
016	Facade Length (YL)	RVS	manual	018	released 11/03/2014	No		
017	Total Area of Facade(s) (YL)	RVS	manual	018	released 12/2013	No		
018	Certainty (YL)	RVS	manual		released 12/2013	No		
019	Facade Height (YR)	RVS	manual	022	released 19/03/2014	Yes	Inspection instructions	Changed to maximum wall height; in the period before the change (approx. starting June 2014) this field isn't used, since not considered in the S-score calculation v8.
020	Facade Length (YR)	RVS	manual	022	released 11/03/2014	No		
021	Total Area of Facade(s) (YR)	RVS	manual	022	released 12/2013	No		
022	Certainty (YR)	RVS	manual		released 12/2013	No		
023	Facade Height (XB)	RVS	manual	026	released 19/03/2014	Yes	Inspection instructions	Changed to maximum wall height; in the period before the change (approx. starting June 2014) this field isn't used, since not considered in the S-score calculation v8.
024	Facade Length (XB)	RVS	manual	026	released 11/03/2014	No		
025	Total Area of Facade(s) (XB)	RVS	manual	026	released 12/2013	No		
026	Certainty (XB)	RVS	manual		released 12/2013	No		
027	Facade Height (XF)	RVS	manual	030	released 19/03/2014	Yes	Inspection instructions	Changed to maximum wall height; in the period before the change (approx. starting

								June 2014) this field isn't used, since not considered in the S-score calculation v8.
028	Facade Length (XF)	RVS	manual	030	released 11/03/2014	No		
029	Total Area of Facade(s) (XF)	RVS	manual	030	released 12/2013	No		
030	Certainty (XF)	RVS	manual		released 12/2013	No		
031	X Front Inspection possible	RVS	Observed		released 12/2013	No		
032	Openings (YL)	RVS	manual	018	released 12/2013	Yes	Field type	Choice list to number field.
033	Openings (YR)	RVS	manual	022	released 12/2013	Yes	Field type	Choice list to number field.
034	Openings (XB)	RVS	manual	026	released 12/2013	Yes	Field type	Choice list to number field.
035	Openings (XF)	RVS	manual	030	released 12/2013	Yes	Field type	Choice list to number field.
036	X Front Reason if not possible	RVS	Observed		released 12/2013	No		
037	Y Right Inspection possible	RVS	Observed		released 12/2013	No		
038	Y Right Reason if not possible	RVS	Observed		released 12/2013	No		
039	X Back Inspection possible	RVS	Observed		released 12/2013	No		
040	X Back Reason if not possible	RVS	Observed		released 12/2013	No		
041	Y Left Inspection possible	RVS	Observed		released 12/2013	No		
042	Y Left Reason if not possible	RVS	Observed		released 12/2013	No		
043	Walls out of plane	RVS	Observed		released 12/2013	No		
044	Recommendation Wall Out of Plane	RVS	Observed		released 12/2013	No		
045	Column crack(s) or slenderness issues	RVS	Observed		released 12/2013	No		
046	Recommendation column crack(s) or slenderness issues	RVS	Observed		released 12/2013	No		

047	Wall cracks	RVS	Observed		released 12/2013	No		
048	Recommendation wall cracks	RVS	Observed		released 12/2013	No		
049	Deflected lintels	RVS	Observed		released 12/2013	No		
050	Recommendation Deflected Lintels	RVS	Observed		released 12/2013	No		
051	Wall ties damage	RVS	Observed		released 11/03/2014	No		
052	Recommendation wall ties	RVS	Observed		released 12/2013	No		
053	Balcony(s) present	RVS	Observed		released 12/2013	No		
054	Parapet(s) present	RVS	Observed		released 12/2013	No		
055	Cantilevered elements present	RVS	Observed		released 21/05/2014	No		
056	Canopy(s) present	RVS	Observed		released 12/2013	No		
057	Recommendation Balcony-Parapets-Canopies	RVS	Observed		released 12/2013	No		
058	Slender chimney(s) present	RVS	Observed		released 12/2013	Yes	Inspection instructions	Procedure of reporting not slender, not damaged chimney's is changed; first all not slender chimneys were registered as HRBE 8 (reported as follows: Recommendation= 'No action', Presence of HRBE = 'No', but photographs and/or description provided). After the change, these chimneys were reported as HRBE 7
059	Recommendation Slender Chimneys	RVS	Observed		released 12/2013	No		
060	Damaged chimney(s) present	RVS	Observed		released 12/2013	Yes		Procedure of reporting not slender, not damaged chimney's is changed; first all not slender chimneys were registered as HRBE 8 (reported as follows:

								Recommendation= 'No action', Presence of HRBE = 'No', but photographs and/or description provided). After the change, these chimneys were reported as HRBE 7
061	Recommendation Damaged Chimney(s)	RVS	Observed		released 12/2013	No		
062	Unsafe roof cladding	RVS	Observed		released 12/2013	No		
063	Recommendation unsafe roof cladding	RVS	Observed		released 12/2013	No		
064	Mortar damage	RVS	Observed		released 11/03/2014	No		
065	Masonry dormer(s) present	RVS	Observed		released 12/2013	No		
066	Recommendation Dormers	RVS	Observed		released 12/2013	No		
067	Lack of ties in cavity walls	RVS	Assumed		released 12/2013	No		
068	Recommendation Mortar Damage	RVS	Observed		released 12/2013	No		
069	Recommendation lack of ties in cavity walls	RVS	Observed		released 12/2013	No		
070	Other damages	RVS	Observed		released 12/2013	No		
071	Other Recommendations	RVS	Observed		released 12/2013	No		
072	Inspection Possible	RVS	Observed		released 12/2013	No		
073	Abandonment	RVS	Observed		released 12/2013	No		
074	Reason inspection not performed	RVS	Observed		released 12/2013	No		
075	Address ID	GIS			released 12/2013	No		
076	Premises ID	GIS			released 12/2013	No		
077	Latitude (Y)	GIS			released 12/2013	No		
078	Longitude (X)	GIS			released 12/2013	No		

079	PGA	GIS			released 12/2013	Yes	Definition	From December 2014 till June 29th 2015 (batch 1-6) KNMI values were provided; for earlier records (Loppersum) and batches after June 29th 2015 (batch 7 ff.) Shell values were given.
080	Address Use 1	GIS			released 12/2013	No		
081	Address Use 2	GIS			released 12/2013	No		
082	BAG Address Use	GIS			released 12/2013	No		
083	Status of Premises	GIS			released 12/2013	No		
084	Building Year	GIS			released 12/2013	No		
085	Importance Class	GIS			released 12/2013	No		
086	Occupancy Class	GIS			released 12/2013	No		
087	Address Use 1 (manual)	RVS	Observed		released 11/03/2014	Yes	Extra/ modified choice list items	
088	Address Use 2 (manual)	RVS	Observed		released 11/03/2014	Yes	Other	Made obsolete.
089	Address Use 2 (list)	RVS	Observed		released 19/03/2014	Yes	Other	Made obsolete.
090	BAG Address Use (manual)	RVS	Observed		released 12/2013	No		
091	Status of Premises (manual)	RVS	Observed		released 12/2013	No		
092	Building Year (manual)	RVS	Assumed		released 12/2013	No		
093	Importance Class (manual)	RVS	Observed		released 12/2013	No		
094	Occupancy Class (manual)	RVS	Assumed		released 12/2013	No		
095	Main Wall Material	GIS			released 12/2013	Yes	Other	Value no longer provided.
096	Main Wall Material (manual)	RVS	Assumed		released 12/2013	No		
097	Ground Floor Material	GIS			released 12/2013	Yes	Other	Value no longer provided.

098	Ground Floor Material (manual)	RVS	Assumed		released 12/2013	No		
099	Higher Floor Material	GIS			released 12/2013	Yes	Other	Value no longer provided.
100	Higher Floor Material (manual)	RVS	Assumed		released 12/2013	No		
101	Area Building Footprint	GIS			released 12/2013	No		
102	Building Height	GIS			released 12/2013	No		
103	Number of Storeys	GIS			released 12/2013	No		
104	Area Building Footprint (manual)	RVS	Observed		released 12/2013	Yes	Inspection instructions	Initially the footprint of the whole premise was accounted for; after the change only the footprint of the 'main' building is assessed (implications on S-score).
105	Building Height (manual)	RVS	Observed		released 12/2013	Yes	Inspection instructions	In some cases the building height provided by GIS might include chimneys; the change only needs to be registered if the difference is >1m.
106	Number of Storeys (manual)	RVS	Observed		released 12/2013	Yes	Field type	Double to integer > attic storeys that were previously counted as half storeys were are now added as full storeys.
107	Horizontal Irregularity	RVS	Observed		released 12/2013	No		
108	Vertical Irregularity	RVS	Observed		released 12/2013	No		
109	Storey Height	RVS	Observed		released 11/03/2014	No		
110	Basement present	RVS	Assumed		released 11/03/2014	No		
111	Emergency confirmed?	RVS	n/a		released 29/07/2014	No		
112	Safety Label	RVS	Observed		released 04/06/2014	No		
113	Emergency Situation	RVS	Observed		released 11/03/2014	No		

114	Emergency intervention	RVS	Assumed		released 04/06/2014	No		
115	S-Curve Value In Plane	RVS	Assumed		released 12/2013	No		
116	S-Curve Value Out of Plane	RVS	Assumed		released 11/03/2014	No		
117	S-Curve version	RVS	n/a		released 21/05/2014	No		
118	Presence Of Adequate Wall Ties	RVS	Assumed		released 12/2013	No		
119	Joint Structure	RVS	Assumed		released 12/2013	Yes	Inspection instructions	All terraced, semi-detached and linked buildings are reported as joint structures. In 2013 it is only set to linked when a solid/joined party wall was expected, since this cannot be observed, a different approach is used
120	Foundation Type	RVS	Assumed		released 12/2013	No		
121	Structure Certainty	RVS	manual		released 12/2013	No		
122	Roof Type	RVS	Assumed		released 12/2013	No		
123	Cavity Walls	RVS	Assumed		released 12/2013	No		
124	Presence of Wall Floor/Roof Ties	RVS	Assumed		released 12/2013	No		
125	Wall Thickness	RVS	Assumed		released 11/03/2014	No		
126	Thickness of Inner Leaf	RVS	Assumed		released 12/2013	No		
127	Thickness of Outer Leaf	RVS	Assumed		released 12/2013	No		
128	Maintenance Level	RVS	Observed		released 12/2013	No		
129	Deterioration of masonry	RVS	Observed		released 12/2013	No		
130	Deterioration of mortar over joints	RVS	Observed		released 12/2013	No		
131	Deterioration of concrete elements	RVS	Observed		released 11/03/2014	No		
132	Deterioration of metal connections	RVS	Observed		released 12/2013	No		

133	Deterioration of wooden elements	RVS	Observed		released 12/2013	No		
134	Solar Cells Present	RVS	Observed		released 12/2013	No		
135	Roof suitable for Solar Cells	RVS	Observed		released 11/03/2014	Yes	Other	Bug fix: if 134 was set to "Yes", it was not possible to fill this field.
136	Consequence class	GIS			released 16/03/2015	No		
137	Consequence class (manual)	RVS	Observed		released 16/03/2015	No		
138	Slender chimney(s) - Height [m]	RVS	Observed		released 16/03/2015	No		
139	Slender chimney(s) - Short side [m]	RVS	Observed		released 16/03/2015	No		
140	Slender chimney(s) - Long side [m]	RVS	Observed		released 16/03/2015	No		
141	HRBE 7 - Slenderness ratio	RVS	Observed		released 16/03/2015	No		
142	Sloped chimney flue - Probability	RVS	Assumed		released 16/03/2015	No		
143	Sloped chimney flue - Additional recommendations	RVS	Assumed		released 16/03/2015	No		

4.2 EVS Inspection Data

The EVS is a structural assessment based on the visual inspection of the building, internal and external, recording information required to determine structural upgrading measures for the building. The EVS focuses on identifying potential falling hazards and significant structural damage and deformations. The inspection also includes the collection and recording of structural information and construction details, where visible. However, the inspection does not include invasive investigation, though the need for this type of investigations may be identified as a follow-up action.
The specific objectives of the EVS are to:

- Collect (initial) building information in preparation for potential future structural upgrading design works;
- Confirm the condition of HRBEs (High Risk Building Elements) identified during the RVS (Rapid Visual Screening) and identify and describe any additional HRBEs which could not be identified during the RVS;
- Validate data collected during the RVS and collect further data;
- Identifying the existing structure, to carry out a preliminary seismic evaluation according to Tier 1 of ASCE 41-13.

Only the primary residential building on a given address is fully assessed during the EVS, while a limited assessment is executed for other buildings on the property. A detailed description of the scope of work can be found in [13].

4.2.1 Data

Prior to conducting the on-site screening, a desk study is carried out by the inspection and engineering team. The objective of the Desk Study is to gather all available technical information from different sources (municipality archives, building owners, etc.), including the RVS report if available for the address. In particular, the parameters that define the risk associated with HRBE's are being reviewed, so the relevant details can be screened during the visit.
Details about the Desk Study process and the minimum required data to be collected can be found in [13].
The information obtained during the Desk Study is validated and supplemented by a visual screening on-site, but limited to nonintrusive investigations. A preliminary seismic evaluation will be performed and a screening for HRBE's executed, comprising the inspection of HRBE's identified during the RVS, supplemented where additional HRBE's are identified.
The final deliverable of this process in an EVS report. The EVS report summarizes the information obtained from the visual screening, and provides a complete but brief detailed description of the building and its structure including photographs and drawings. It will at least contain the following information:

- General information (for instance: address data building age, etc.)
- Building description
- Structural description
- Screening validation of HRBE
- Final HRBE recommendations

- Drawings of the building
- Safety Assessment Form
- RVS HRBE recommendations
- ASCE 41-13 Tier 1 checklist

Although providing sufficient information for further structural analyses, the unstructured data contained in the EVS report cannot be directly imported into database and can therefore not be analysed. Hence an attempt has been made during a repair action in December 2015 to extract the most relevant data from the to date available EVS reports, with focus on the information required for the assignment of building typologies [21]. The below table gives a description of the extracted data.

Table 13: EVS Schema.

Field ID	Field Name	Field Type	Field Description	Comments
001	Building name (if applicable)	Text	Denomination of public or historic buildings.	
002	Street	Text	Street name.	
003	Street number	Text	House number.	
004	Post Code	Text	Postal code.	
005	Town	Text	Name of village/town in municipalities.	
012	BAG object-ID	Text	Unique identification code assigned per building by BAG.	
013	Building year	Text	Building construction year.	
014	Address Use	Menu	Main use of the building.	
015	Mixed Use?	Menu	Building with two or more use functions.	
016	Adjacency	Menu	Spatial relation between separate buildings.	
017	Apartment?	Menu	Building with two or more addresses.	
018	Aggregation	Menu	Connection between immediately adjacent but separate buildings.	
019	Presence of secondary buildings	Text	Number of secondary buildings (Sheds, garages etc.).	
020	Presence of extension	Text	Extension built later than the main building.	
022	Shape in plan	Menu	Building geometry in section.	
024	Presence of basement	Menu	Presence of a building layer which is fully or partially below ground.	
025	Foundation type	Menu	Distinction between shallow and deep foundation.	
026	Foundation system	Menu	Specification of the main foundation system.	
027	Number of storeys above ground, excluding attic	Text	Number of building layers.	A building consisting of one storey means a building in which only

				the ground floor provides habitable space.
028	Number of habitable attic storeys	Text	Number of habitable building layers which are fully or partially under the roof.	
029	Gutter height - Above ground excluding roof	Text	Average height from ground level to gutter of the main building.	
030	Building height - Above ground to top of roof (excluding chimneys etc)	Text	Average height from ground level to ridge of the roof of the main building.	
031	Ground storey - Inter-storey height	Text	Height of the first storey.	
032	Roof form	Menu	Shape of the main roof.	
034	Roof type	Menu	Predominant construction material of the main roof.	
035	Roof system	Menu	Specification of the structural system of the main roof.	
036	Presence of gable walls	Text	Number of gables on the building.	
037	Presence of dormer	Text	Number of dormers on the building.	
039	Vertical load-bearing system	Menu	Main vertical support system.	
040	Vertical load-bearing material	Menu	Main structural material of vertical support system.	
041	Stability system	Menu	Main lateral stability system.	
042	Stability material	Menu	Main structural material of lateral stability system.	
043	Presence of internal load bearing walls	Text	Number of internal load bearing walls.	
044	Presence of internal non load-bearing walls	Text	Number of internal non load bearing walls.	
045	Floor system - Ground floor	Menu	Main structural system of the ground floor.	
047	Floor system - Upper floors	Menu	Main structural system of the higher floors.	

4.2.2 Results

The number of buildings with EVS inspections that was included in EDB V5 was 392.

4.2.3 Data limitations

During the repair action from December 2015, the above fields were supplemented with a confidence value per field. This confidence value relates to the certainty with which the data could be retrieved from the report and ranges from ‘assumed’ to ‘verified’. The data collection has not been conceived to be used separately from the report and/or for other users than the inspectors.

EVS inspections are recorded manually (i.e. with no data structure) and stored as a report. To allow for the EVS information to be used in data processing, the reports were interpreted by engineers and recorded in a structured data format. While check processes were implemented during the translation of the reports to structured data, several misinterpretations and human errors (including inconsistent field values) in the data were included.

4.3 Drawing Data (TBDB)

Drawing data was collected and stored in the 'technical building database' (TBDB) for terraced and semi-detached buildings within the 0.2g PGA contours (KNMI 2015) as identified from preliminary studies.

The collection method included retrieving drawings (architectural and structural where possible) and extracting relevant building attributes into a structured format to be hosted in a database. The primary purpose of the drawing data collection is to provide detailed information which may assist with the inspection process and structural assessment.

As the TBDB data collection process is ongoing, the data set used in the EDB V5 is the status of the data collection as per June 2017.

4.3.1 Data

Table 14 show the fields that were made available after the extract of the TBDB was imported into the project database. Where values have been identified from a predefined list, the field has been marked with a comment of 'list'.

Table 14: TBDB Schema.

Field ID	Field Name	Field Type	Field Description	Comments
1000	IntendedUse	Text	Intended Use / address use of the building	List
1025	MultipleAddressBuilding	Boolean	Flag to whether it is a multiple address building	
1040	Architect	Text	Architect noted in the drawings	
1045	StructuralEngineer	Text	Structural engineer noted in the drawings	
1050	ConstructionCompany	Text	Construction company noted in the drawings	
1060	ArchitectType	Text	Architectural type as noted in the drawings	
1061	TypeMain	Text	Architectural type class as defined by the drawing data collection team	
1062	TypeSub1	Text	Architectural type subclass as defined by the drawing data collection team	
3000	Adjacency	Text	Adjacency	List
3010	HorzIrregularities	Text	Horizontal Irregularities / Shape in Plan	List
3015	VerticalIrregularitiess	Text	Vertical Irregularities	List
3200	BuildingUnitHeight	Numeric	Building / Unit Height	
3205	StoreyHeightGroundFloor	Numeric	Storey Height - Ground Floor	
3206	StoreyHeightFirstFloor	Numeric	Storey Height - First Floor	

3207	StoreyHeightAttic	Numeric	Storey Height - Attic	
3210	GutterHeight	Numeric	Gutter Height	
3216	LoadBearingSpaceWidth	Numeric	Building / Unit Width	
3217	NumberofLoadBearingSpaces	Integer	Load-Bearing Space (centre to centre, mm)	
3225	RoofInclination	Integer	Inclination / Slope of Roof	
3250	NumberofStoreysNonAttic	Integer	Number of Storeys (above ground, excluding attic)	
3255	NumberofStoreysAttic	Integer	Number of Attic storeys	
3265	NumberofBasementLevels	Integer	Number of basement levels	
3405	PresenceofExtensions	Boolean	Presence of Extensions	List
3410	PresenceofBasement	Boolean	Presence of basement	List
3411	PresenceofSouterain	Boolean	Presence of souterrain	List
3412	PresenceofSoftStorey	Boolean	Presence of soft storey	List
3415	PresenceofDormer	Boolean	Presence of dormer	List
3420	PresenceofGableWall	Boolean	Presence of gable wall	List
3421	PresenceofURMChimney	Boolean	Presence of URM Chimney	List
3425	RoofShape	Text	Roof Shape	List
3430	RoofType	Text	Roof Type	List
3435	RoofSystem	Text	Roof System	List
3440	FloorTypeGroundFloor	Text	Floor type - Ground Floor	List
3445	FloorSystemGroundFloor	Text	Floor system - Ground Floor	List
3451	FloorTypeFirstFloor	Text	Floor type - First Floor	List
3452	FloorSystemFirstFloor	Text	Floor system - First Floor	List
3453	FloorTypeSecondFloor	Text	Floor type - Second Floor	List
3454	FloorSystemSecondFloor	Text	Floor system - Second Floor	List
3456	FloorTypeAtticFloor	Text	Floor type - Attic Floor	List
3457	FloorSystemAtticFloor	Text	Floor system - Attic Floor	List
3458	FloorTypeSecondAtticFloor	Text	Floor type - Second Attic Floor	List
3459	FloorSystemSecondAtticFloor	Text	Floor system -Second Attic Floor	List
3465	PresenceofWallTiesFloorRoof	Boolean	Wall ties floor/roof	
3470	FoundationType	Text	Foundation Type	List
3475	FoundationSystem	Text	Foundation System	List
3480	VerticalSupportSystemType	Text	Vertical support system / Gravity Load support system	List
3485	VerticalSupportSystemSystem	Text	Vertical support system / Gravity load support system - Material	List
3490	LateralSupportSystemTypeFrontBack	Text	Lateral support system – Front Back	List

3495	LateralSupportSystemSystemFrontBack	Text	Lateral support system Material – Front Back	List
3500	LateralSupportSystemTypeLeftRight	Text	Lateral support system – Left Right	List
3505	LateralSupportSystemSystemLeftRight	Text	Lateral support system Material – Left Right	List
3510	PresenceofInternalLoadBeauringWalls	Boolean	Presence of internal structural walls	
3515	PresenceofInternalNonLoadBearingWalls	Boolean	Presence of internal non-structural (partition) walls	
3530	WallTypeExteriorWall	Text	Exterior Wall Type	List
3531	WallSystemInnerLeafExteriorWall	Text	Exterior Wall System Inner Leaf	List
3532	WallSystemOuterLeafExteriorWall	Text	Exterior Wall System Outer Leaf	List
3533	WallTypePartyWall	Text	Party Wall Type	List
3534	WallSystemPartyWall	Text	Party Wall System	List
3535	WallTypeInternalWall	Text	Internal Wall Type	List
3536	WallSystemInternalWall	Text	Internal Wall System	List
3805	GroundLevelOpeningPercFront	Numeric	Ground level opening percentage Front	Calculated based of corresponding redrawn elevation
3810	GroundLevelOpeningPercBack	Numeric	Ground level opening percentage Back	Calculated based of corresponding redrawn elevation
3815	GroundLevelOpeningPercLeft	Numeric	Ground level opening percentage Left	Calculated based of corresponding redrawn elevation
3820	GroundLevelOpeningPercRight	Numeric	Ground level opening percentage Right	Calculated based of corresponding redrawn elevation

4.3.2 Results

The TBDB data used in the EDB V5 contains information for 9470 buildings which are mainly semi-detached and terraced buildings within the 0.2g PGA contour (KNMI 2015).

4.3.3 Data limitations

The TBDB data requires architectural and structural drawings to be interpreted by engineers and recorded in a structured data format. While check processes were implemented during the data collection process, misinterpretations and human errors (including inconsistent field values) in the data were observed. It should also be noted that the TBDB parameters were not designed to match a format suitable for the EDB processes, therefore assumptions while converting the original data collected were sometimes required. The validity of these assumptions has been assessed at a regional scale, but could lead to misleading interpretation of the specific building.

4.4 Visual Inspections (JBG)

Desktop visual inspections were undertaken by sub-contractors JBG on buildings within the 0.2g PGA contour (KNMI, 2015) using Google Streetview or Horus photos (collected street imagery for the project by Horus). Buildings which were identified to be terraced or semi-detached buildings as per the Exposure Database V3 were not included as these were expected to be covered by the drawing data collection.

The primary purpose of these visual inspections was twofold:

- Validate the types of buildings which had been assigned to Exposure Database V3 Intermediate classes, and
- Collect and assess additional exterior building characteristics which may have an influence on the building structural system to help refine the classification process.

To assist with the data collection process, certain fields were prefilled which could be verified or corrected by the visual inspectors. Other parameters were provided with pre-set choice lists. The data was collected using a SharePoint tool developed by JBG and the buildings were processed in batches. The data was checked by both JBG reviewers and Arup and specific buildings which either contained incorrect information or where advice was required were further verified and processed.

A high-level description of the Visual Inspections process, definitions used by inspectors, tools and data can be found in the manual [22].

4.4.1 Data

Table 15 provides an overview of the fields collected including whether prefilled information was provided. For each of the relevant fields, a corresponding confidence field was added to understand if the characteristic was observed or assumed.

Table 15: Visual Inspections Schema.

Description (Field)	Type of Field	Prefill	High level Definition	Source
Visual validation possible	List		Indicates whether it was possible to perform the visual the inspection.	Google Streetview or Horus photograph
Original building type	List		Type of building at construction.	Google Streetview or Horus photograph
Original building subtype	List		Sub-type of building at construction.	Google Streetview or Horus photograph
Current Building Use (population)	List		Current use of the building.	Google Streetview or Horus photograph
Additional use in the same building (population)	List		Secondary use of the building.	Google Streetview or Horus photograph
Construction year	Number	Y	Building construction year.	Google Streetview or Horus photograph
Multiple address building	List	Y	Presence of more than one address in the building.	Google Streetview or Horus photograph
Adjacency	List		High level spatial relationship of the building to other buildings.	All

Touching other buildings	Number	Y	Number of buildings touching the inspected building.	All
Touching same building	Number	Y	Number of buildings touching the inspected building which appear the same as the inspected building.	Google Streetview or Horus photograph
Shape in Plan	List		Shape of building footprint.	BAGviewer
Presence of soft storey	List		Presence of a soft storey	Google Streetview or Horus photograph
Presence of souterrain	List		Presence of a split basement level	Google Streetview or Horus photograph
Number of storeys (above ground, excluding attic)	Number		Number of building layers above ground, excluding attic.	Google Streetview or Horus photograph
Number of attic storeys	Number		Number of building layers within the roof.	Google Streetview or Horus photograph
Number of chimneys	Number		Number of chimneys.	Google Streetview or Horus photograph
Number of gable wall	Number		Number of gable walls, defined as triangular wall area delimited by inclined roof planes.	Google Streetview or Horus photograph
Number of parapets (trapgevels)	Number		Number of parapets.	Google Streetview or Horus photograph
Dominant roof shape	List		Dominant roof shape of the building.	Google Streetview or Horus photograph
Dominant roof shape in plan	List		Dominant roof shape of the building in plan.	All
Secondary roof shape	List		Secondary roof shape of the building.	Google Streetview or Horus photograph
Secondary roof shape in plan	List		Secondary roof shape of the building in plan.	All
Exterior material	List		Main façade material.	Google Streetview or Horus photograph
Presence of secondary buildings	Number		Presence of additional buildings on the plot.	All
Related BAG ID of secondary buildings	Number		Building ID of related, non-touching buildings. Multiple values/field possible.	BAGviewer

4.4.2 Results

The number of buildings with visual inspections that was included in EDB V5 was 12,174.

4.4.3 Data limitations

Some difficulties with the gathered parameters are identified during the validation processes. The following issues were observed:

- For some parameters it was noticed that there were various interpretations among the inspectors. The parameters were this issue was mainly the case are the combination of Dominant roof type and Secondary roof type, Shape in plan, Additional use, Original building type and Number of attic storeys. Especially, the roof types regularly showed issues.
- Some parameters, for example 'Number of touching other buildings', have been implemented differently than the parameter definition states. Which can cause issues, as the prefilled parameter might wrongly be revised.
- Inconsistencies between depending parameters within one visual inspection. For example, there were several adjacency parameters which could contradict each other within a building record.
- The confidence flags have not always been filled in.
- Sometimes fields were left blank were a value is expected. It is not always clear what the reason or meaning is of a blank field.
- In the final outputs some unexpected values were found in several parameter fields, for example text in numerical fields or unlisted choices.
- Similarly, it was noticed that the new records for secondary buildings sometimes refer to multiple secondary buildings.

4.5 Arup Expert

The Arup Expert dataset is a collection of three tables created through the collation of project information and desktop studies which have been formatted to allow for use in the database and the building classification process. The three tables correspond to the following studies:

- Index buildings [23]

  The index buildings are selected buildings which have been structurally analysed. The buildings' corresponding structural layout and structural systems (GEM string format [1], [25]) are captured.

- Soft storey prioritisation [22]

  The mixed-use buildings (RECA) within the 0.2g PGA contour (KNMI 2012) were part of a desktop study to prioritise for soft storey vulnerability. The outcomes of this study along with their corresponding structural layout (also collected as part of the desktop study) are captured.

- Drawing data [24]

  The drawing information from the technical building database (TBDB) was translated into structural systems (GEM string format) [1], [25]. The validated results were captured.

4.5.1 Data

The following tables describes the contents and the schema of the three datasets.

Table 16: Index building table format.

Column name	Type	Description
building_name	Text	Name of the building
pand_id	Text (50)	The building identification number provided by BAG
address	Text	Street name, house number and postcode as provided by BAG
gem_1_material_dx	Text	Material of the lateral load-resisting system in X direction
gem_2_ls_dx	Text	Lateral load-resisting system in X direction
gem_3_material_dy	Text	Material of the lateral load-resisting system in Y direction
gem_4_ls_dy	Text	Lateral load-resisting system in Y direction
gem_11_exterior_walls	Text	Presence of cavity external walls
gem_13_floor	Text	Material of the building floor system (Ground floor not considered)
geom_class	Text	Geometry class [1]
sly_class	Text	Structural layout [1]
source	Text	Source of information

Table 17: Soft storey prioritisation table format.

Column name	Type	Description
pand_id	Text (50)	The building identification number provided by the Basisregistratie Adressen en Gebouwen (BAG)
soft_storey_vulnerability_cluster	Text	Outcome of soft-storey priorisation [22]
classification_path	Integer	Soft-Storey priorisation - Flowchart classification path
soft_storey_flag	Text	Soft-Storey high risk flag
source	Text	Source of information
ref	Text	Reference document
geometrical_class	Text	Visually assigned geometrical class [1]
structural_layout	Text	Visually assigned building layout [1]

Table 18: Drawing data GEM strings table format.

Column name	Type	Description
pand id	Text (50)	The building identification number provided by BAG
gem_1_material_dx	Text	Material of the lateral load-resisting system in X direction
gem_2_ls_dx	Text	Lateral load-resisting system in X direction
gem_3_material_dy	Text	Material of the lateral load-resisting system in Y direction
gem_4_ls_dy	Text	Lateral load-resisting system in Y direction
gem_11_exterior_walls	Text	Presence of cavity external walls
gem_13_floor	Text	Material of the building floor system (Ground floor not considered)
gem_string	Text	Complete GEM string
source	Text	Source of information

4.5.2 Results

The number of buildings with covered by the Arup Expert Dataset in the EDB V5 were:

- Index buildings – 13 buildings
- Soft storey prioritisation – 342 buildings
- Drawing data – 9470 buildings

4.5.3 Data limitations

The data collection of the soft storey prioritisation study was undertaken using Google Streetview and inherits the inaccuracies of Google Streetview. The content collected is also engineering judgement based.

The TBDB GEM strings are affected by the assumptions made on the translation process. The resulting GEM strings will also inherit any inaccuracies of the TBDB data as input [1].

4.6 Population Data

The population dataset contains information on estimated population per building. This information was provided by NAM and was updated for the EDB V5 deliverable.

4.6.1 Data

Table 19 describes the contents and the schema of the population dataset.

Table 19: Population data table format.

Column Name	Type	Description
pand_id	Text	Unique identification code, assigned per building by BAG
sum_pop_in_day	Double	# People inside the building, during day time
sum_pop_run_day	Double	# People running outside at the event of an earthquake that are estimated to be in the at-risk zone from debris falling outside a building, during day time
sum_pop_pas_day	Double	# People passing-by or stay present in the at-risk zone from debris falling outside a building, during day time
sum_pop_in_night	Double	# People inside the building, during night time
sum_pop_run_night	Double	# People running outside at the event of an earthquake that are estimated to be in the at-risk zone from debris falling outside a building, during night time
sum_pop_pas_night	Double	# People passing-by or stay present in the at-risk zone from debris falling outside a building, during night time

4.6.2 Results

Population data was provided for 163,996 buildings. This is majority of the buildings within the EDB V5 scope area which contains addresses (164,032).

4.6.3 Data limitations

The reliability of the population data relies on several assumptions. These assumptions have been made using the best available data but may contain inaccuracies due to input data or misalignment of input data versions. As the population data is used for a regional assessment, it was deemed suitable.

The following observations have been identified in the data:

- The field ‘pop_run’ (day and night) are calculated based on a factor applied to ‘pop_in’ with the result not subtracted from the ‘pop_in’. Care should be taken not to count the population twice when using the two fields together.
- The fields ‘pop_in’ and ‘pop_run’ are not related to the field ‘pop_pas’ as they come from separate datasets and assumptions.

- The distribution for 'pop_in_day' appear to be inconsistent within building functions with observed issues around functions such as schools and hospital which seem low.

5 Processed Data

5.1 Building Use and Pre-class

Building Use information provides the primary and secondary Use of a building and a flag to identify whether a building contains a Residential Use. These are used as direct data fields in the exposure database extract as well as input into the structural system classification process as ‘pre-class’, which is an aggregated Building Use classification.

For V5, the Building Use analysis has been updated to include the updated and additional source data alongside an improved methodology which utilises the additional source dataset and validates specific uses relating to the building typologies. This includes using open source data to identify churches and the school registry to find schools. Additionally, the parcel information is used to appoint buildings, with no addresses, such as sheds, or an unclear Building Use, to a building with particular Use sharing the same parcel.

5.1.1 Data

Table 20 describes the contents and the schema of the resulting dataset.

Table 20: Building Use and pre-class table format.

Column Name	Type	Description
pand_id	Text	Unique identification code assigned per building by BAG
AKR_UID	Text	Unique parcel id of the parcel the building sits on
use_1	Text	The first building use of a building
use_2	Text	The second Building Use of a building
has_residential_use	Text	Indication of whether the building has a residential use
special_use	Text	Flag of whether the building is a school, church or hospital
flag_parcel_source	Short	Flag of whether parcel data was used to assign Building Use
parcel_use	Text	Building Uses aggregated to assign a use to the parcel
no_postcode	Short	Flag when a building has an address without postal code, which indicates that the building is a garage
preclass	Text	Aggregated uses to a pre-class

5.1.2 Methodology

To have a reproducible and traceable data creation method, ArcGIS’s ModelBuilder and FME Workbench was used. In both ModelBuilder and FME Workbench the data can be loaded, processed and exported in an automated way.

5.1.2.1 Input data

For this analysis, the BAG building information have been used as a reference.

- Basisregistratie Adressen en Gebouwen (Kadaster, April 2017);
- Address use data (Dataland, June 2016);

- Rijksmonumenten data (RCE, April 2017);
- Nationale Atlas Volksgezondheid versie 4.18 (RIVM, December 2016);
- School Register (DUO, April 2017);
- Not Monumental churches dataset (Reliwiki and Google Maps, November 2015; updated January 2017);
- Parcel geometry (Kadaster, August 2015).

5.1.2.2 Operations

The operations undertaken can be summarised as five steps which are further explained in the following text.

1. Identifying a first and second use per VBO from BAG information.
2. Identifying a first and second uses per building through BAG and Dataland information
3. Identifying special uses and garage boxes
4. Identifying uses through parcel data
5. Aggregating uses to a pre-class.

Identifying a first and second use per VBO from BAG information

When the VBO had multiple *gebruiksdoelen*, two *gebruiksdoelen* were selected based on a ranking. This ranking is like the ranking used later in the analysis.

Table 21 Ranking of the *gebruiksdoelen*

Rank	Use
1	*woonfunctie*
2	*winkelfunctie*
3	*kantoorfunctie*
4	*bijeenkomstfunctie*
5	*gezondheidszorgfunctie*
6	*onderwijsfunctie*
7	*logiesfunctie*
8	*sportfunctie*
9	*celfunctie*
10	*industriefunctie*
11	*overige gebruiksfunctie*

Identifying first and second uses and 'has residential' per building (BAG and Dataland)

BAG's Address Use information provides 11 categories (as ranked above) to describe the possible Address Uses. The BAG's Address Use information has full coverage over the study area.

Dataland's Address Use information is more specific using approximately 470 categories. These categories describe the buildings uses more precisely. The dataset also provides some insight to multiple uses per address.

Both datasets are used to assign the main and secondary uses per building but require some interpretations to assign the uses as per the required field values. BAG's raw field values and translations are as per below:

Table 22: Translation and interpretation of the BAG *gebruiksdoelen*.

BAG Use (Dutch)	BAG Use (English)	Interpretation
Logiesfunctie	Accommodation Function	CommercialBusiness
Celfunctie	Cell Function	PublicGovernmental
Onderwijsfunctie	Educational Function	Education
Gezondheidszorgfunctie	Healthcare Function	HealthCare
Industriefunctie	Industrial Function	Industrial
Bijeenkomstfunctie	Meeting Function	(refer to Dataland)
Kantoorfunctie	Office Function	CommercialBusiness
Overigegebruiksfunctie	Other Function	(refer to Dataland)
Woonfunctie	Residential Function	Residential
Winkelfunctie	Shop Function	CommercialBusiness
Sportfunctie	Sport Function	Recreation

As the Dataland dataset's taxonomy is not corresponding to the one used in the EDB, a mapping table is used to help classify the 470 categories to the required field values. The mapping table can be provided upon request.

BAG's Uses are used as the default Address Use except for several cases in which Dataland's Uses are referred to. This is due to BAG's full coverage and the limited amount of required interpretations.

While BAG was the default dataset, Dataland uses were referred to in the following cases:

- Where BAG's uses refer to '*Bijeenkomstfunctie'* (meeting function) and '*Overige gebruiksfunctie'* (other use function). '*Bijeenkomstfunctie'* and '*Overige gebruiksfunctie'* were too broad as a use to be classified into any of the required field values. Dataland was referred to provide a better understanding on the use.
- Where Dataland identifies Agriculture or Religious buildings as BAG's uses does not identify Agricultural or Religious buildings
- Where a single building with a single address is identified by Dataland as having two uses where in the case BAG's uses only found one use per address.

The steps undertaken as part of this stage of the analysis are as follows:

- The BAG and Dataland use information was joined to an address ID and BAG ID. Building Use information was categorised into the required field values where possible.
- The buildings with addresses were divided by buildings with single addresses and multiple.

- The buildings with single addresses were assigned either BAG or Dataland uses.
- The buildings with multiple addresses were checked to see if only a single use was evident (i.e. only one use was identified on a building level). If so, this was assigned as the first use of the building.
- For each building with multiple addresses and multiple uses, the uses were aggregated with a corresponding sum of the floor area ('oppervlak') per address/use.
- The floor area of each use per building was given a ranking. The first and second ranking then indicated the first and second use per building. For example, a building with two addresses and uses of commercial and residential that each had an area of 100sqm and 50sqm would be given a rank of 1 and 2 respectively. This would result in a first use of commercial and a second use of residential.
- Where a building has multiple uses in which each has an equal area, an extra step was taken to prioritise uses. The following ranking was used:

Table 23: Ranking of the Building Uses where multiple uses are found.

Rank	Use
1	Residential
2	CommercialBusiness
3	Educational
4	Health Care
5	Recreation
6	PublicGovernmental
7	Religious
8	Industrial
9	Agricultural
10	NotVerblijfsobject
11	Unknown

- The results were then appended together to create a dataset with the total buildings of the study area.
- Where a building has an Address Use that is flagged as being residential, either through BAG by BAG as 'Woonfunctie' (Residential) or according to the Dataland classification sheet as having residential ('has_res') then the building is deemed to have a residential use.

<u>Identifying special uses</u>

Churches, hospitals and schools are perceived as buildings with special uses. However, while BAG and Dataland uses have assigned two uses where possible, it is noted that:

- not all buildings with a Use of Health Care are 'hospitals'
- not all buildings with a Use of Religious are 'churches'
- not all buildings with a Use of Educational are 'schools'.

To select the specific buildings of hospitals, churches and school, alternative data sources are used where possible. This includes the *Rijksmonumenten* dataset to select church buildings; the *Basisregister Instellingen* to identify schools; and the *Nationale Atlas Volksgezondheid* for the identification of hospitals within the region. The identification of a special use (hospital, school or church) was linked to the Building Use dataset through BAG ID matching.

Additionally, a flag has been created for buildings with an address without postal code. It is observed that these buildings are in general garage boxes or sheds, as they fall outside the zip code zones.

Identifying unknown or missing uses with Parcel Data

The BAG and Dataland datasets can only assign uses to buildings that have an address. Approximately 37% of the building stock within the study area does not have an address. The Parcel Data is used to help identify a potential use for these buildings by identifying a relationship of the buildings without an address to a building with an address and use on the same parcel.

The following steps are undertaken:

- The Building Use dataset extended with a unique parcel id through a spatial join.
- The parcels that only contain buildings which had no Use assigned (i.e. only contained buildings without addresses) are filtered out. For these buildings identifying a building use through the above mentioned procedure is not possible.
- The parcel with address and first use information is then analysed. First parcels which contain only one use are analysed, excluding the first Uses of '*NotVerblijfsobjec*t' and 'Unknown'. These parcels are given a 'parcel use' of the single Use (i.e. when two buildings are located on a parcel, one having an address and a first use of residential and the other not having an address, the 'parcel use' would be assigned as 'Residential'). Where a parcel contains multiple uses (i.e. has two buildings, one with a use of commercial and another of residential), these are assigned a 'parcel use' of 'multiple'.
- This processed parcel dataset is then joined back to the building dataset so it can be identified whether a building is on a parcel with an identified parcel Use. The following assumptions are then made to buildings which either have no address (and thus no use previously defined) or have a Use of '*OtherVerblijfsobject*' or 'Unknown':
    - o If the building with no address has a footprint ⩽ 20sqm, then the building is assigned a use of 'Shed'
    - o If the building had no address and the parcel Use has been identified as 'Residential', then the building is being assigned a Use of 'Shed'.
- Finally, an extra field is added to flag those buildings that are assigned a potential use through the parcel data.
- The resulting table is then exported as the Building Use dataset.

Aggregating uses to pre-classes

The pre-class analysis is based on the Building Use of the buildings in the research scope. It is a preparation for the adjacency classification which is only applicable to the buildings which are assigned a residential typology.

The pre-classes and the rules that are used to assign the pre-class can be seen in Table 24.

Table 24: Queries governing the pre-class classification.

Pre-Class	Database query
SHED	main_building_use is NULL OR main_building_use = 'NotVerblijfsobject' OR main_building_use = 'Shed' OR no_postcode = '1'
UNK	main_building_use = 'Unknown' OR main_building_use = 'OtherVerblijfsobject'
HOS	special_use = 'Hospital'

SCH	special_use = 'School'
CHU	special_use = 'Church'
RES	main_building_use = 'Residential'
REC	has_residential_use = 'Res' AND (main_building_use = 'CommercialBusiness' OR '' secondary_building_use = 'CommercialBusiness' OR main_building_use = 'Health Care' OR '' main_building_use = 'Educational' OR main_building_use = 'PublicGovernmental' OR secondary_building_use = 'Health Care')
AGR	main_building_use = 'Agricultural'
IND	main_building_use = 'Industrial'
COM	main_building_use = CommericalBusiness OR main_building_use = Healthcare OR main_building_use = PublicGovernmental OR main_building_use = Religious OR main_building_use = Education OR main_building_use = 'Recreation'

5.1.3 Result

The pre-class analysis has been performed for all the building within the EDB V5 scope, which is summarised in Table 25.

Table 25 Summary of the results of the pre-class analysis following the Building Use analysis.

Pre-class	Building count
SHED	129253
IND	3232
REC	4521
COM	5472
AGR	2287
SCH	396
CHU	297
UNK	328
RES	167681
HOS	8

5.1.4 Validation

The results have been checked mainly through desktop studies using the BAG viewer and Google Streetview [15] and by comparing against previous EDB version's Building Use results.

5.1.5 Data limitations

The results of analysis on Building Use and pre-class dataset inherits any limitations from the BAG and Dataland datasets as listed in Section 2.1 and 2.2 respectively. It should be noted that the Dataland dataset does not have full coverage of the study area and is not an audited dataset.

Not all 'shed / garages' have been identified through the stages of analysis. This is predominately because there are several garages which have an address but their use has not been defined by either BAG or Dataland. These buildings with addresses cannot be picked up through the current parcel parameters as they are within their own parcel (i.e. have no other use within the parcel and are larger than 20sqm). The diagram below shows the residential buildings in pink with the grey buildings as being identified as sheds though the parcel analysis. The black buildings currently have unknown uses.

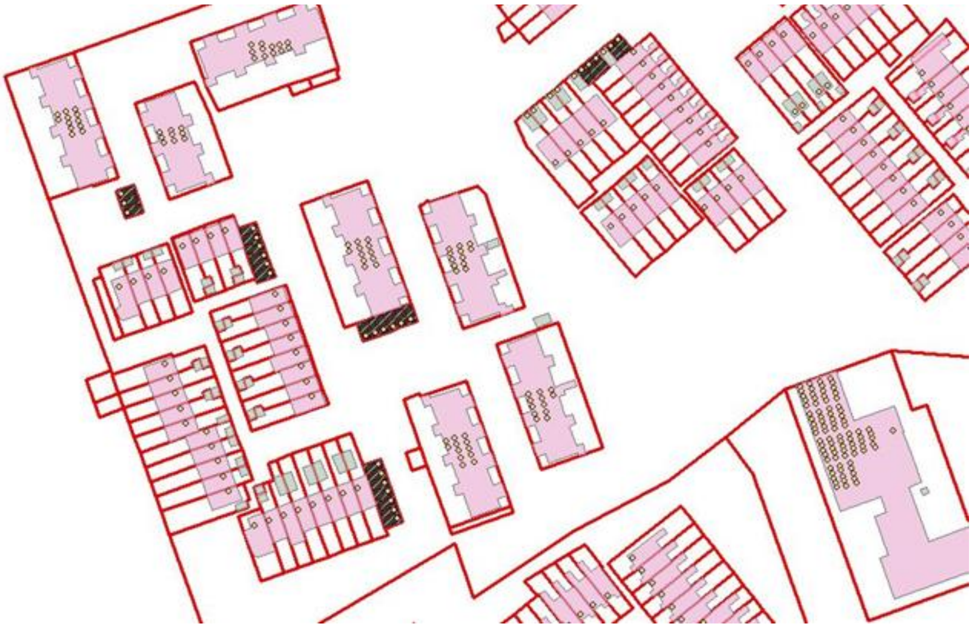


Figure 8: Results of the Building Use analysis indicating that not all sheds or garages have been identified (in black). Buildings which have been assigned as a 'shed' are in grey.

5.2 Adjacency

In the adjacency analysis the spatial relation between separate buildings is determined; i.e. how a building relates to neighbouring buildings, if there are any. The result of this analysis is a set of parameters which are input for the building typology classification (such as number of buildings in a block, number of neighbouring buildings, etc.). A single separate flag is assigned to blocks that contain one building with a different building year but are otherwise homogeneous.

The results of the adjacency analysis are part of the EDB extract as well as input into the structural system classification process.

5.2.1 Content

Table 26 describes the contents and the schema of the resulting dataset.

Table 26: Adjacency table format.

Column Name	Type	Description
pand_id	Text	Unique identification code assigned per building by BAG
bldg_in_block	Short	Number of buildings in a block (cluster of all touching buildings).
adjacency_class	Text	Adjacency classification
ts	Double	Result of combining parameters building year, continuous roof ridge and footprint area indicating that the building is touching similar
to	Double	Result of combining parameters building year, continuous roof ridge and footprint area indicating that the building is touching other
bldg_in_blockpart	Short	Number of buildings in a block part (cluster of touching identical buildings).
nbr_blockpart	Short	Number of neighbours in a block part
nbr_block	Short	Number of neighbours in a block
nbr_without_vbo	Short	Number of neighbours without vbo
bldg_within_05m	Short	Number of buildings within 0.5 meter of building (excluding neighbours)
bldg_within_05m_without_vbo	Short	Number of buildings without vbo within 0.5 meter of building (excluding neighbours)
block_flag	Short	Flag when part of a block
blockpart_flag	Short	Flag when part of a block part
mixed block	Short	Flag for buildings in a mixed block
end buildings	Short	Flag for buildings at the end of a block part

5.2.2 Methodology

To have a reproducible and traceable data creation method, ArcGIS's ModelBuilder and FME Workbench was used. In both ModelBuilder and FME Workbench the data can be loaded, processed and exported in an automated way.

5.2.2.1 Input data

For this analysis, the following input data is used:

- Basisregistratie Adressen en Gebouwen (April 2017);
- AHN height data (October 2016);
- Dataland Address Use (June 2016).

5.2.2.2 Operations

To understand whether a building is touching a similar building, the buildings are tested on building year, continuous roof ridges and footprint area.

The required characteristics are determined for each individual building through a spatial analysis of all the buildings within the scope area. The analysis is divided into three main parts:

1. Creating blocks of touching building and block parts of similar buildings
2. Classify the buildings into an adjacency classification
3. Assessing intermediate parameters

Creating blocks and block parts

1. Snap buildings which are within 10cm of each other.

   Previous analysis has found that the BAG building outline datasets contains small drawing errors such as small gaps between terraced buildings. By snapping the buildings together within a tight threshold, these gaps can be removed.

2. Identifying blocks

   Buildings touching each other are 'dissolved' to create a polygon (block) per set of adjacent buildings and are assigned a block id. Each of the dissolved buildings which make up the block are also assigned the relevant block ID.

3. Identifying block parts

   Block parts are identified in a similar fashion to the blocks but with the addition of selected parameters as a requirement that needs to be met. This is performed with the following parameters:

   - Building year
   - Presence of a continuous roof ridge
   - Similar footprint

   This results to three sets of block parts being created for touching buildings with the same building year, continuous roof ridge and footprint area.

   The footprint area and roof ridge parameters required a classification to be made to allow for the dissolve grouping.

4. Assessing the number of buildings per block (unit count) and block part.

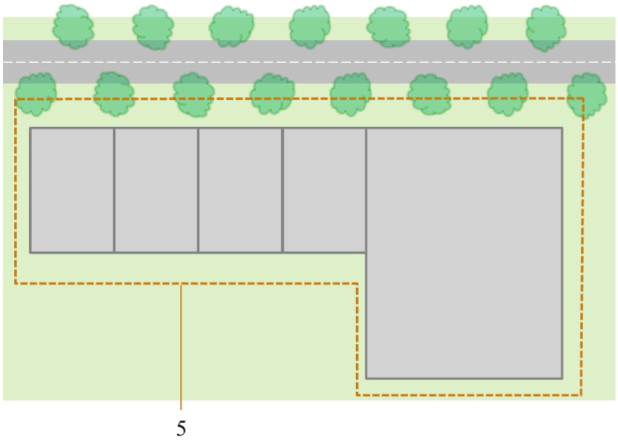
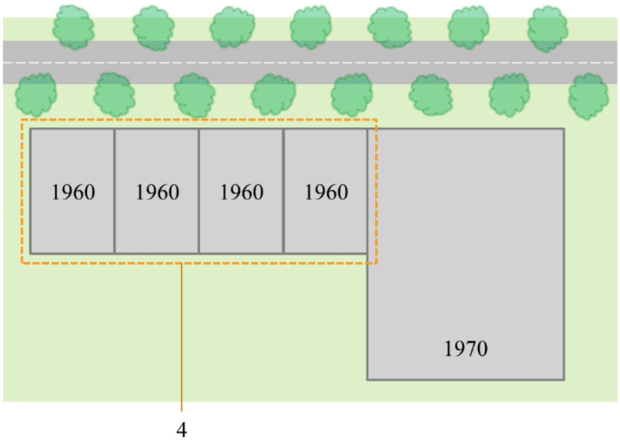
The number of buildings was counted against the block ID and block part IDs. This represents the number of building per block (i.e. unit count) and block parts.

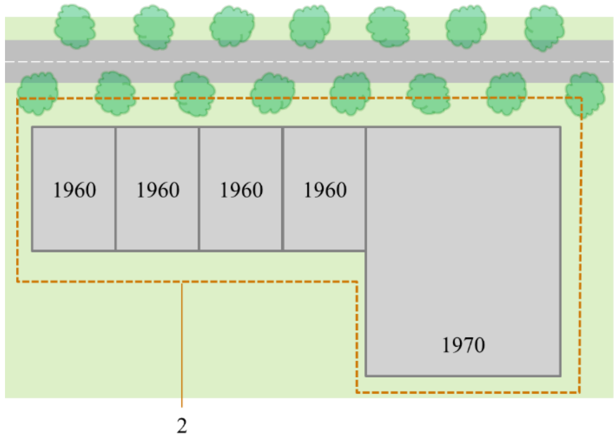
5. Assessing the number of similar buildings per block

   The number of buildings with the same building year, roof ridge and footprint area class is counted against the block id. This represents the number of buildings with similar parameters per block, regardless of whether they are touching or not.

6. Assigning intermediate parameters

   Intermediate parameters were calculated to be used for the adjacency classification. These used the following definitions:

Intermediate parameter	Description	Example
bldg_in_block	Number of buildings in block	
bldg_in_blockpartY	Number of buildings in block part with similar building year	
bldg_in_blockpartR	Number of buildings in block part with similar roof ridge height	
bldg_in_blockpartA	Number of buildings in block part with similar footprint area classification	
count_bouwjaar	Count of building year in block	
count_roof	Count of building year in block	

count_area	Count of building year in block	

Classify the buildings into an adjacency classification

1. Identifying freestanding / detached buildings

   Building which are freestanding are identified where the number of buildings in a block is only one.

2. Identifying similar buildings

   Similar buildings were identified when the three parameters used to identify block-parts (i.e. building year, roof ridge and footprint size)were present. Of these similar buildings, semi-detached buildings could also be identified when the number of buildings within a block part was two.

3. Identifying different buildings touching

   Dis-similar buildings were identified when none of the above parameters were not present.

4. Weighting parameters to assign an adjacency classification

   The remaining buildings which were not classified after the above steps required further investigation. These included buildings where a changed building year and footprint area was a result of a renovation but the building was still like its neighbours. In these cases, the three parameters were weighted as percentage of likelihood to assign a final adjacency classification based on its quality as an indicator of touching similar buildings.

5. Assign final adjacency classification

   Using the three parameters and its percentage of likelihoods as independent pieces of evidence of touching similar buildings, the overall likelihood is calculated by combining the probabilities using the following formula:

$$P(s) = \frac{p1 * p2 * p3 \dots pn}{\left(p1 * \mathrm{p2} \dots pn + (1 - p1)(1 - p2) \dots (1 - pn)\right)}$$

$$P(o) = \frac{(1 - p1)(1 - \mathrm{p2}) \dots (1 - pn)}{\left((1 - p1)(1 - p2) \dots (1 - pn) + p1 * \mathrm{p2} \dots pn\right)}$$

Where P(s) and P(o) are the probability of 'touching similar' and 'touching other' respectively.
Most of adjacent buildings can then be classified to 'touching similar' or 'touching other'. Where the probabilities of either one lies between 0.4 and 0.6, the classification is then set to 'touching', only.

Assessing adjacency parameters

This final section processes the adjacency classification and added additional spatial parameters.

1. Dataland building description

   Dataland contains extensive building description including the description of terraced and semi-detached houses. These two types of building descriptions are used to help revise buildings which have been classed as 'touching' from the prior process. These buildings are re-classified to 'touching similar'.

2. Final similarity check

   A final check is on the overall result to ensure that the assigned classes make sense i.e. that 'touching similar' buildings do not have neighbouring buildings which are classed as 'touching other' or 'detached'. The source of the inconsistency is mainly from Dataland information.

3. Creating final block parts

   The final set of block parts are created using touching similar buildings. Each of these block parts are assigned a block part ID.

4. The final adjacency parameters are calculated including:

   a. Assessing the number of neighbouring buildings within the block, block part and buildings without an occupancy (*verblijfsobject* as per BAG).

   b. Assessing the number of buildings with and without occupancy within a 0.2m range.

   c. Identifying end buildings of a block and mixed blocks (i.e. blocks which contain both similar and dis-similar buildings).

5.2.3 Results

The adjacency analysis has been performed for all buildings covering the EDB V5 assessment area.

Table 27 Summary of the results of the adjacency class following the adjacency analysis.

Adjacency class	Building count
Touching Similar	123418
Touching Other	16829
Touching	4564
Freestanding	112363

5.2.4 Validation

Both the intermediate and final results have been checked mainly through desktop studies using the BAG viewer and Google Streetview and compared with the previous EDB version of the analysis.

5.2.5 Data limitations

The results of this analysis are expected to inherit the limitations of the input data sets. The impact on the results includes gaps in analysis (i.e. AHN2 data which was used as the basis of roof ridges does not contain information for post 2009 buildings) or errors in the geometry (i.e. building outlines from BAG may not have been drawn correctly). The full description and limitations of the BAG and AHN datasets can be found in Section 2.1 and 2.5.

5.3 Exposed Footprint

The exposed footprint length captures the length of the building's footprint which are exterior facing (i.e. not including walls between buildings). The exposed footprint length is a data field in the EDB extract.

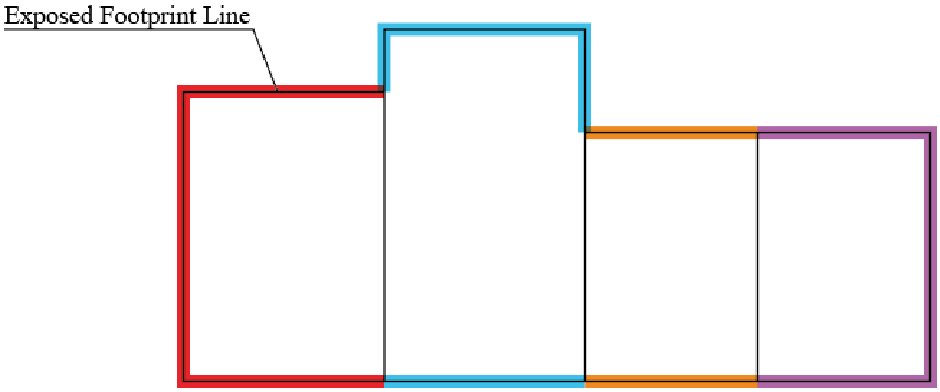


Figure 9. Example of the exposed footprint length/s of different buildings.

5.3.1 Data

Table 28 describes the contents and the schema of the resulting dataset.

Table 28 Exposed footprint table format.

Column Name	Type	Description
pand_id	Text	Unique BAG building ID
exposed_footprint_length	Double	The length of the exposed, outer wall per building in meters.

5.3.2 Methodology

To have a reproducible and traceable data creation method, ArcGIS's ModelBuilder was used. Through ModelBuilder, the data can be loaded, processed and exported in an automated way.

5.3.2.1 Input data

For this analysis, the BAG building outlines was used.

- Basisregistratie Adressen en Gebouwen (April, 2017)

5.3.2.2 Operations

To calculate the exposed wall length, the following operations are performed:

1. Buildings are dissolved to urban blocks by removing shared or touching walls. This creates an outline around groups of adjacent buildings.

2. The outlines of the urban block are split at every overlapping vertices from the individual building outlines. This turns the polygons of the blocks in lines. The image below depicts this process.

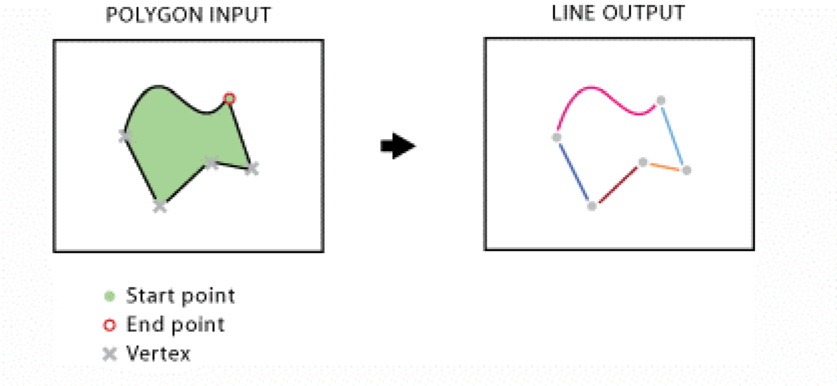


Figure 10: Split Line at Vertices tool in ArcGIS.

3. Subsequently, the resulting lines get assigned the corresponding BAG building ID by joining the buildings and lines spatially, when a line segment overlaps.

4. The lines with the same BAG building ID are dissolved, so that one polyline feature is created per building.

5. Finally, the length of each line per building is calculated.

5.3.3 Results

This analysis has been performed for all 257,174 buildings.

In the figure below an example of the expected output can be seen. Every colour represents another exposed wall.

Figure 11: Results for exposed wall analysis

ID	LENGTH	UNIT
1	24.2	m
2	12.3	m
3	12.3	m
4	12.4	m
5	12.3	m
6	12.3	m
7	24.1	m
8	12.3	m

5.3.4 Validation

A visual check on a number of buildings has been performed and a sample of buildings has been measured through GIS against the results.

5.3.5 Data limitation

The creation of the exposed wall length is based on a spatial analysis of the buildings dataset from BAG. Although this is a reliable dataset, it does have few limitations which are listed below:

- Minor gaps between the building outlines, where in reality these are adjacent/touching buildings, have been observed. This would result in a larger exposed wall length.
- The BAG dataset is a continuously evolving dataset that is updated as buildings are demolished and built. While it is currently the best data available with regards to buildings and outlines, it may still be subject to errors and missing data.
- The building outline does not always distinguish multiple building structures. This means that the lengths extracted for a building are not necessarily of the same building structure.

Full description of the BAG dataset and its limitations can be found in Section 2.1.

5.4 Gutter height

Gutter height is defined as the height between the beginning and the end of an exterior wall of a building where the wall intersects with the roof structure. The gutter height of a building is not always represented by a unique value as each one of its walls might have a different gutter height, which can also vary throughout the length of the wall itself.

The gutter height algorithm developed is returning an average height (AvHeight) per wall along with its length (L). This information is then processed to return an average weighted gutter height for each unit under study.

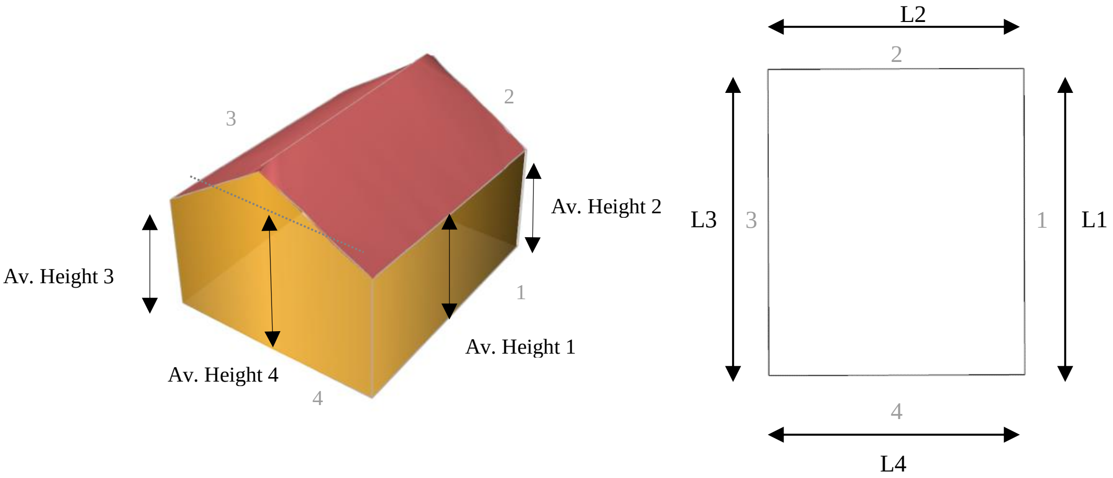


Figure 12: Definition of the average gutter height and corresponding wall length

The gutter height is one of the main geometric features used in the building classification process and a field in the extract [1].

5.4.1 Data

Table 29 describes the contents and the schema of the resulting dataset.

Table 29 Gutter height table format

Column Name	Type	Description
pand_id	Text	Unique BAG building ID
gutter_height	String	The length and average height per outer wall per building in meters

The gutter height is returned in the following format:

(L1|AvHeight1;L2|AvHeight2;L3|AvHeight3;…;Ln|AvHeightn)

Where “L” and “AvHeight” are length and average height respectively.

5.4.2 Methodology

5.4.2.1 Software

McNeel Rhinoceros3D parametric modelling environment Grasshopper3D was used for a reproducible and traceable data creation method. This allows the data to be loaded, processed and exported in an automated way, while taking advantage of Rhinoceros 3d modelling environment and Grasshopper3D's visual programming capabilities.

5.4.2.2 Input data

For this analysis, the following input data is used:

- Basisregistratie Adressen en Gebouwen (April 2017);
- AHN height data (October 2016).

5.4.2.3 Operations

The gutter height calculation algorithm (shown in Figure 13) is based on the identification of points close to each segment of the footprint outline and the subsequent calculation of the average height of the points per segment. The final gutter height per building is the average of the segment average heights("AvHeight$_i$") weighted by the respective lengths of the segment ("L$_i$"), i.e. is calculated through the following formula:

$$gutter\_height = \frac{(L1 * AvHeight1) + (L2 * AvHeight2) + \cdots + (Ln * AvHeightn)}{L1 + L2 + \cdots + Ln}$$

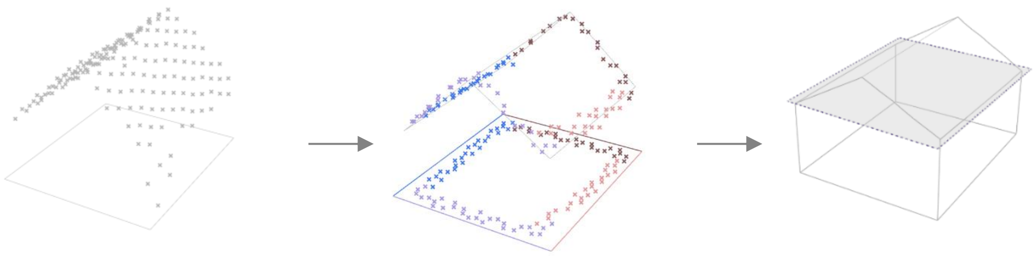


Figure 13: Operations undertaken to calculate the gutter height.

5.4.3 Results

The analysis has been performed on 239 039 buildings. For the rest of the buildings there was either insufficient amount of point cloud data, poor quality point cloud data, or there was no point cloud data available due to the building being built in a later date than when the point cloud scan was generated (2009).

5.4.4 Validation

To validate the accuracy of the results of this parameter, a comparison between the calculated and inspection data was performed, as documented in a dedicated note [35]. The study showed a sufficient agreement between inspection and algorithmic data – especially given the fact that the definition was not the same for the processes.

5.4.5 Data limitations

This algorithm uses as input the BAG and AHN datasets therefore related limitations are inherited from these datasets as described in sections 3.1.4 and 3.5.4 and mainly relate to the absence of point cloud data for buildings built after 2009 and to occasional observed inaccuracies on the footprint outline data of BAG.

The point cloud data is based on a 0.5m grid. The resolution of this grid can result in inaccuracies in the produced geometries. The impact of the resolution to the gutter height accuracy is expected to be limited.

Additionally, the AHN data does not account for roof overhangs leading to slight misalignment of the roof profile with the actual wall. This is expected to be a case for a very limited amount of buildings.

It should also be noted that, after filtering points related to overhanging trees, the geometry on that area is an approximation based on extending the plane of the roof geometry. Therefore, if below the trees the roof is discontinuous (e.g. on the presence of dormer) this would not be captured by the algorithm. The impact of this is expected to be only significant in buildings of small size.

5.5 Roof Steepness

The roof steepness captures whether the roof of a building is predominantly flat, inclined but not steep or inclined and steep, through the measuring of roof inclinations. The ratio of points within each angle domain is recorded. The angle domains are based on 10° inclination steps starting from 0° and reaching 90° degrees. The predominant inclination domain is subsequently recorded as the main steepness class for each processed building.

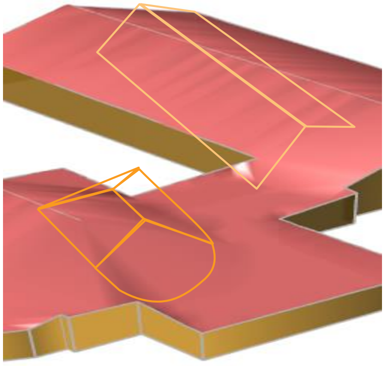


Figure 14: Example building with multiple roof compartments of varying steepness.

The roof steepness is a geometric feature contributing in the Structural Layout classification process [1].

5.5.1 Data

Table 30 describes the contents and the schema of the resulting dataset.

Table 30: Roof steepness table format.

Column Name	Type	Description
pand_id	Text	Unique BAG building ID
steep1_0_10	Double	Percentage of roof area belonging to this domain
steep2_10_20	Double	Percentage of roof area belonging to this domain
steep3_20_30	Double	Percentage of roof area belonging to this domain
steep4_30_40	Double	Percentage of roof area belonging to this domain
steep5_40_50	Double	Percentage of roof area belonging to this domain
steep6_50_60	Double	Percentage of roof area belonging to this domain
steep7_60_70	Double	Percentage of roof area belonging to this domain
steep8_70_80	Double	Percentage of roof area belonging to this domain
steep9_80_90	Double	Percentage of roof area belonging to this domain

5.5.2 Methodology

5.5.2.1 Software

As described in 5.4.2.1, McNeel Rhinoceros3D parametric modelling environment Grasshopper3D was used for a reproducible and traceable data creation method.

5.5.2.2 Input data

For this analysis, the following input data are used:

- Basisregistratie Adressen en Gebouwen (April 2017);
- AHN height data (October 2016).

5.5.2.3 Operations

The algorithmic calculation of roof steepness is based on the identification of slope per point. This is based on the generation of a Delaunay mesh from the pointcloud data and the subsequent calculation of the mesh normal vector at each point. The amount of points at each angle domain are then counted with the most predominant angle domain defining the steepness class for the building.

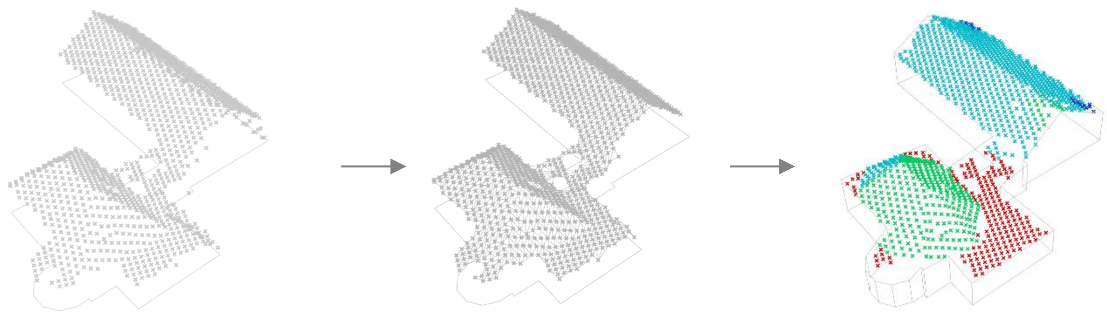


Figure 15 Operation steps undertaken to calculate the roof steepness.

5.5.3 Results

As described in 5.4.3 not all buildings within the region had available pointcloud data. This resulted in buildings whose steepness could not be assessed[2].

5.5.4 Validation

For the dataset used in the EDB V5, validation checks were performed through comparison of the calculated values against manual angle calculations using technical drawings. The visual classification between flat and steep roofs in the inspected buildings was in agreement with the algorithmic results.

[2] . In EDB V5, the steepness was only used as a distinction between similarly sized Structural Layouts which then eventually received the same inference rules. If next versions assign different inference rules to WBW (“Warehouse”) Structural Layouts and to WBB (“Barns”) an alternative source for roof steepness should be identified - e.g. through drawing data.

5.5.5 Data limitations

In addition to the inherent limitations expressed in Section 5.4.5, buildings with complex roofs of varying steepness cannot be fully classified into flat or steep. Further study is needed to identify ways to separate complex roof structures into separate roof segments and subsequently classify each segment by the predominant inclination.

5.6 Roof Count

The roof count captures whether the building has a simple (i.e. one roof) or complex (multiple) roof structures. The count is calculated as the sum of the identified roof ridges and the horizontal roof parts per roof. An algorithm capable of measuring adjacent types of roof has been applied with the capability of identifying and exclude ridges that might result from small openings, overhangs, point cloud inconsistencies or flat parts that are result of small extensions of the units or connection points of two or more separate volumes, as shown in Figure 16

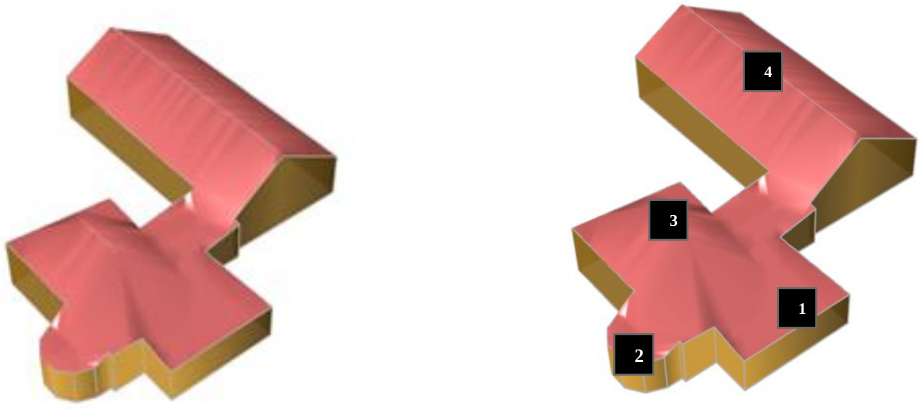


Figure 16: Example of unit with 4 identified separate roofs. two flat roofs (1,2), two roof ridges (3,4) and a connection point that is not accounted as additional roof.

The roof count is used as input into the EDB V5 Structural Layout classification process, with certain Structural Layouts featuring the description “Complex” based on a count of roofs larger than 1. In EDB V5 this doesn’t alter the inference set used but it is implemented as a distinction that could be further explored in later versions.

5.6.1 Data

Table 31: Roof count table format describes the contents and the schema of the resulting dataset.

Table 31: Roof count table format

Column Name	Type	Description
pand_id	Text	Unique BAG building ID
roofcount	Integer	Number of roofs identified

5.6.2 Methodology

5.6.2.1 Software

As described in 5.4.2.1, McNeel Rhinoceros3D parametric modelling environment Grasshopper3D was used for a reproducible and traceable data creation method.

5.6.2.2 Input data

For this analysis, the following input data is used:

- Basisregistratie Adressen en Gebouwen (April 2017);
- AHN height data (October 2016).

5.6.2.3 Operations

The roof count calculation algorithm (shown in Figure 17) is based on the generation of a Delaunay mesh from the height map point cloud of each building. Grouping the mesh faces by their normal vector values can help identify faces that are approximately horizontal within a given tolerance. The count of point groups above a certain size threshold gives the roof count.

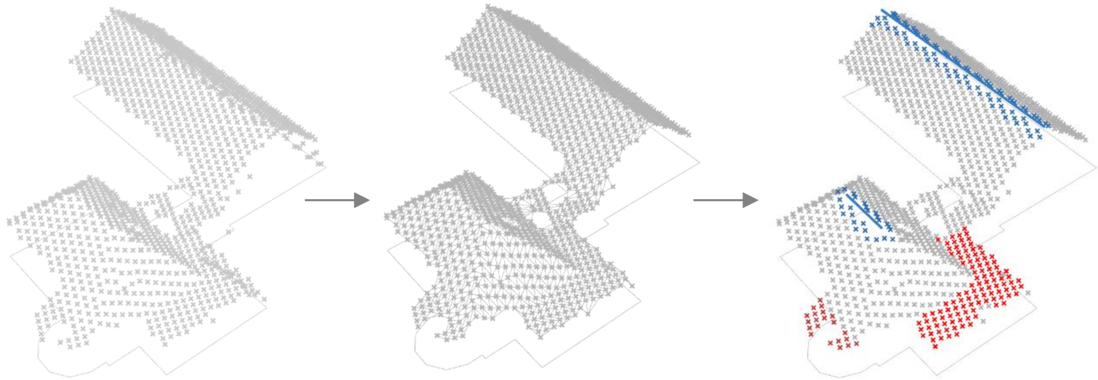


Figure 17 Operation steps undertaken to calculate the roof count.

5.6.3 Results

As described in 5.4.3 not all buildings within the region had available pointcloud data. This results in buildings whose roof count could not be assessed[3].

5.6.4 Validation

For the dataset used in the EDB V5, validation checks were performed through comparison of the calculated values against visually assessed roof counts. The visual classification between “simple” or “complex” roofs (i.e. roofs with multiple roof ridges or flat parts) in the inspected buildings was in agreement with the algorithmic results.

5.6.5 Data limitations

In addition to the inherent limitations expressed in Section 5.4.5, given the complexity of certain roofs, the presence of a chimney or other discontinuity can disrupt the roof ridge separation.

[3] . In EDB V5, given that roof count is mostly used as a distinction between similarly sized Structural Layouts with identical inference rules. If next versions assign different inference rules e.g. to buildings classified as UHO (“House”) Structural Layouts and different inference rules to buildings classified as UHC (“House Complex”), an alternative source for roof steepness should be identified –e.g. through drawing data.

5.7 Maximum Enclosed Rectangle

The “Maximum Enclosed Rectangle” (MER) values capture the dimensions of the largest possible rectangles to fit within a building footprint. These rectangles and their dimensions are expected to provide information about the likely structures within a building outline. Figure 18 shows a few samples of the identification of the MER (color-coded in red) in footprints of different shape (rectangular, L-shaped and T-shaped).

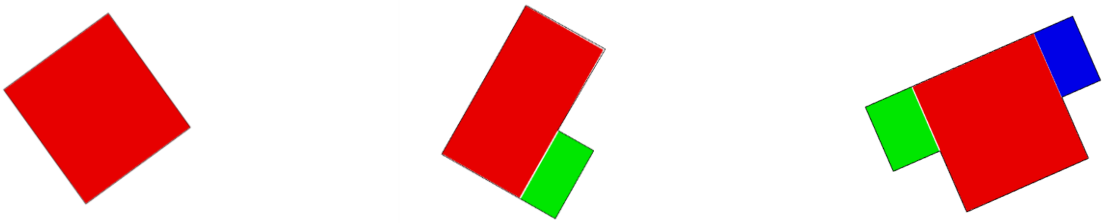


Figure 18: Examples of building footprints and their corresponding MERs. The red rectangle shows the largest fitting rectangle while the green rectangle notes the second largest (i.e. the largest rectangle fitting in the remaining area of the footprint outline).

5.7.1 Data

Table 32describes the contents and the schema of the resulting dataset.

Table 32: MER table format

Column Name	Type	Description
pand_id	Text	Unique BAG building ID
mer_1_x	Double	Length of the largest rectangle within the building footprint
mer_1_y	Double	Width of the largest rectangle within the building footprint
mer_1_area	Double	Area of the largest rectangle within the building footprint
mer_2_x	Double	Length of the second largest rectangle within the building footprint
mer_2_y	Double	Width of the second largest rectangle within the building footprint
mer_2_area	Double	Area of the second largest rectangle within the building footprint

5.7.2 Methodology

5.7.2.1 Software

As described in 5.4.2.1, McNeel Rhinoceros3D parametric modelling environment Grasshopper3D was used for a reproducible and traceable data creation method.

5.7.2.2 Input data

For this analysis, the following input data is used:

- Basisregistratie Adressen en Gebouwen (April 2017)

5.7.2.3 Operations

The algorithm performs a sequence of operations on the footprint outline polygon to calculate the dimensions of the MER. First, the footprint outline is divided into segments and then rectangles are drawn connecting all division points. The largest rectangle that is fully included within the polygon outline is identified as the maximum enclosed rectangle (see Figure 19.

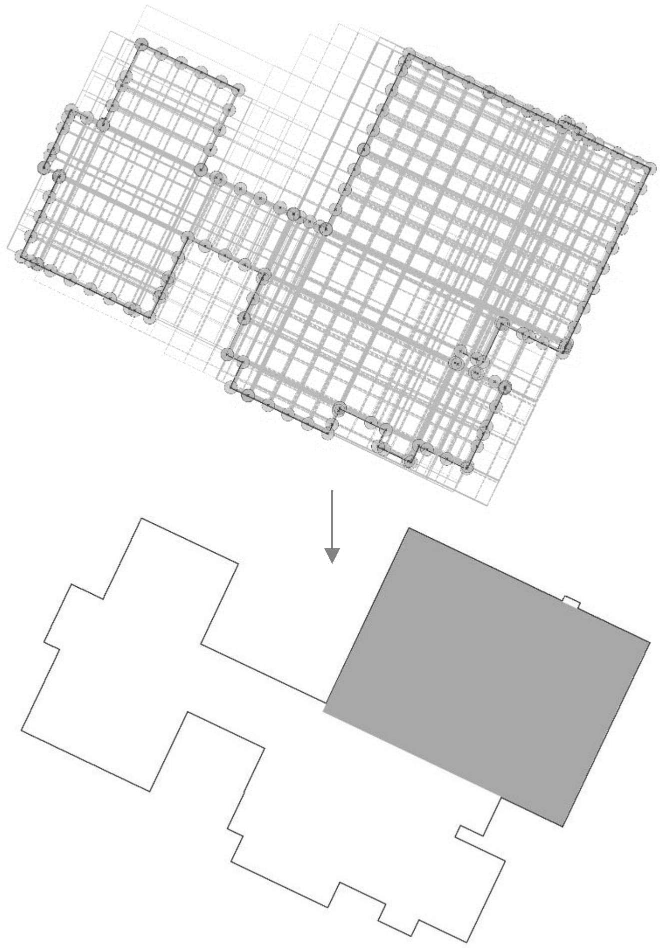


Figure 19 Operation steps undertaken to calculate the dimensions of the Maximum Enclosed Rectangle.

These operations can be repeated for the area that remains after subtracting the MER from the total footprint to identify the second largest enclosed rectangle fitting in the polygon.

5.7.3 Results

The areas and dimensions of the Maximum Enclosed Rectangles have been captured for all buildings in the EDB V5 scope area. Example footprints and the identified MER’s are shown in Figure 20.

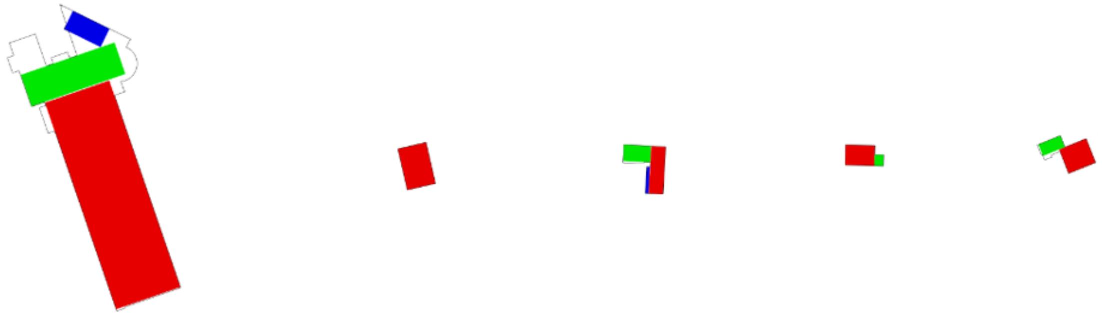


Figure 20 Examples of Max Enclosed Rectangle identified in building footprint outlines of different sizes (in red). The second and third largest fitting rectangles are also shown in green and blue rectangles respectively.

For a limited amount of cases the algorithm needed to be revised to ignore erroneous polygon segments of the building outlines or to modify the tolerances for large or very small footprints. An example of both cases is shown in Figure 21 .

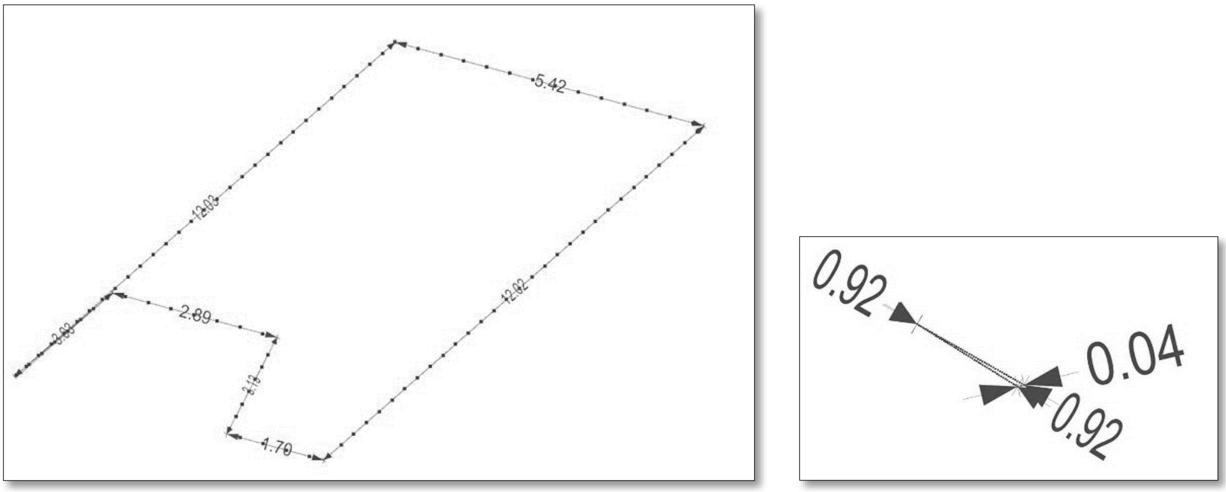


Figure 21 Examples of BAG footprint outlines that required recalibrated tolerances (the dimensions shown are in metres).

5.7.4 Validation

Random samples of footprint outlines have been checked visually and were found to be adequately precise in calculating the dimensions of the maximum enclosed rectangle given an input polygon.

5.7.5 Data limitations

The creation of the MER dataset is based on the building outlines from BAG. Some inaccuracies on the outlines of buildings in this dataset have been observed. It should be noted that it has been observed that the building outlines provided by BAG also include parts of the building corresponding to overhangs.
Full description of the BAG dataset and its limitations can be found in Section 2.1

6 References

[1] **Arup**. *Exposure Database V5_Tech Note*. September 2017. Arup report number: 229746_052.0_AUX2008

[2] **Kadaster**. *Landelijke Voorziening Basisregistraties Addressen en Gebouwen*. 2017.

[3] **Dataland**. DL_ALBGF_Arup_20150409.csv, 2015.

[4] **Kadaster.** *Percelen en grenzen*.

[5] **Deltares**. *Geological Schematisation of the shallow subsurface of Groningen*. 2015. 1209862-005-GEO-0004.

[6] **Kruiver, Dedem, Romijn, de Lange, Korff**. *An integrated shear-wave velocity model for the Groningen gas field, The Netherlands*., 2017 Bull Earthquake Eng .

[7] **AHN**. *Actueel Hoogtebestand Nederland*. 2009 Amersfoort

[8] **RCE** Rijksdienst voor Cultureel Erfgoed, *Rijksmonumenten-kaartlagen*. 2016.

[9] **RIVM**. *De Nationale Atlas Volksgezondheid*. 2016.

[10] **DUO**, Dienst Uitvoering Onderwijs. *Basisregister Instellingen*. 2016.

[11] **Arup**. *Engineering File on RVS process*. 2015. Arup report number: 229746_042.0_NOT106.

[12] **FEMA**. *Rapid Visual Screening of Buildings for Potential Seismic Hazards*. 2015. https://www.fema.gov/media-library/assets/documents/15212.

[13] **Arup**. *Engineering Inspection Protocol EVS for CC1b buildings V0.0*. 2014. Arup report number: 229746_033.0_REP104.

[14] **Arup**. *Visual Inspections Manual*. July 2017. Arup report number: 229746_052.0_NOT2016.

[15] **Google**. *Google Maps*. 2017. www.maps.google.com.

[16] **KNMI**. *Seismische hazardkaart Groningen 2015*. 2015.

[17] **Arup**. *Definities Building Datalijst*. March 2017, Arup report number: 229746_047.0_REP2035.

[18] **NAM**. *Population Data*. 2017. Arup report number: 229746_INC_052.0_NOT2020.

[19] **Kadaster**. *Processenhandboek BAG basisregistraties adressen en gebouwen*. 2013.

[20] **Hewett, M**. *Automating Feature Extraction with the ArcGIS Spatial Analyst Extension*. 2005, ESRI International User Conference, p. 22.

[21] **NAM**. *Exposure Model V1: Updated Typologies and Inference Rules*. 2015. Arup report number: 229746_REF1734.

[22] **Arup**. *Visual Inspections Manual*. 2017. Arup report number: 229746_052.0_NOT2016.

[23] **Arup**. *Structural Layout and System of Index Buildings*. September 2017. Arup report number: 229746_031.0_NOT2033.

[24] **Arup**. *EDB V5 Inspection Mapping Documentation*. November 2017. Arup report number: 229746_052.0_MEM2004.

[25] **GEM Foundation**. About GEM. Global Earthquake Model. https://www.globalquakemodel.org/gem/.

Client: Nederlandse Aardolie Maatschappij

Arup Project Title: Groningen Earthquakes Structural Upgrading

Exposure Database V5 Post-Analysis Report

229746_052.0_REP2018

ISSUE_DEF Rev.0.02 | 18 January 2018

Job number 255090

Arup bv
Postal address:
PO Box 57145
1040 BA Amsterdam
Visitor address:
Naritaweg 118
1043 CA Amsterdam
The Netherlands
www.arup.com

ARUP

1 Contents

2 Introduction

The exposure database (EDB) is one of Arup's key deliverables for the Groningen Earthquakes Structural Upgrading (GESU)'s Hazard and Risk Analysis and is used as input into the Hazard and Risk Modelling completed by NAM's consultants. The EDB is a project database and contains aggregate information on number/type of buildings and on exposed population specific to the Hazard and Risk Modelling, and it is delivered in the form of an extract as described in the companion note [1]. It consists mainly of the building typology classifications and building attributes, within a 5 km buffer around the Slochteren gas field outline, as agreed with NAM.

The EDB is updated periodically in line with key dates provided by NAM. The development of each EDB is in collaboration with Arup's GESU Engineering Studies - Hazard and Risk team (WS 31), Arup's GESU GIS team (WP 52.2) and independent consultants Pinho and Crowley, with feedback from the client (NAM). This is the fifth update of the exposure database and supersedes V0 (July 2014) [2], V1 (March 2015) [3], V2 (September 2015) [4] and V3 (March 2016) [5]. This report documents the updates from V3 and the process of producing the exposure database V5 [1], dated September 2017 for Risk Model V5.

This report aims to provide further insight into the EDB classification process and results through complementary data analysis and visualisations. More specifically it will focus on the following areas:

- Classification Process overview;
- Building data-mining operations;
- Insights on the EDB V5 Extract;
- Inspection data inclusion.
- Evaluation and validation of data in EDB V5.

3 Classification Process Overview

The current approach to develop fragility models for assessing the seismic risk in the Groningen region requires the building stock to be classified into typologies which are expected to have distinct structural performance and failure consequences [6]. Contrary to previous versions of the EDB, those building typologies are no longer a direct and only consequence of their building function and adjacency but they have been more closely related to their structural material and lateral resisting system based on the Global Earthquake Model (GEM) taxonomy system [7]. Through the use of this uniform taxonomy system across the full process the use of available inspection data is improved. This data is now directly mapped to the GEM taxonomy system and in the classification process (as described in a dedicated technical note [8]) this data can be used to update the inferred probabilities with building specific information.

The newly developed classification of the building stock can be described in two main processes:

- Assignment of Structural Layout;
- Assignment of Structural System.

The Structural Layouts of the first process, and related characterization of various buildings attributes, are used as input for the second process. In the 2nd process the most appropriate Structural Systems are determined, which are eventually captured in the EDB extract. Both of these processes are multi-staged and include multiple workflows to allow for the use of project knowledge and inspection datasets. The overall classification diagram which includes the various workflows is shown in Figure 1.

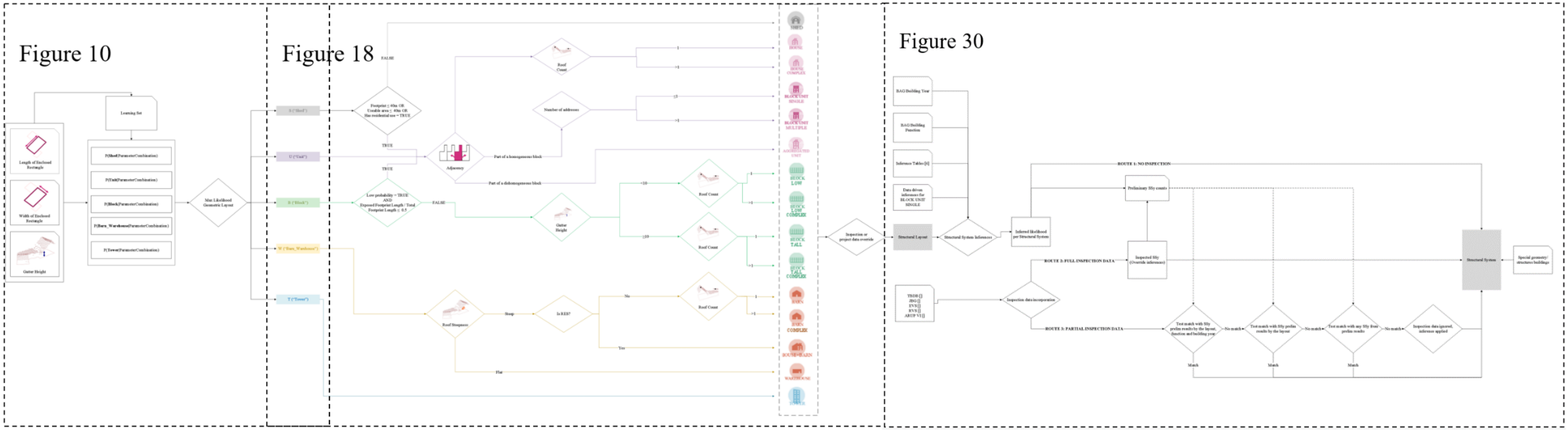


Figure 1: Classification process: from data-mining to Structural Systems. The figures noted on the flow-chart are provided later in this report and show the denoted flowchart areas in higher resolution.

The extensive workflow can be further organised in distinct phases, as exemplified in Figure 2:

1. Building data-mining and geometrical characterization (Figure 10)
2. Classification into Structural Layouts (Figure 18)
3. Structural System Inference (Figure 30)
4. Incorporation of available Inspection Data (Figure 30)
5. Final Structural System Assignment (Figure 30)

This process sequence will form the outline of this report, whereby each phase will be further explained in the following sections of this report.

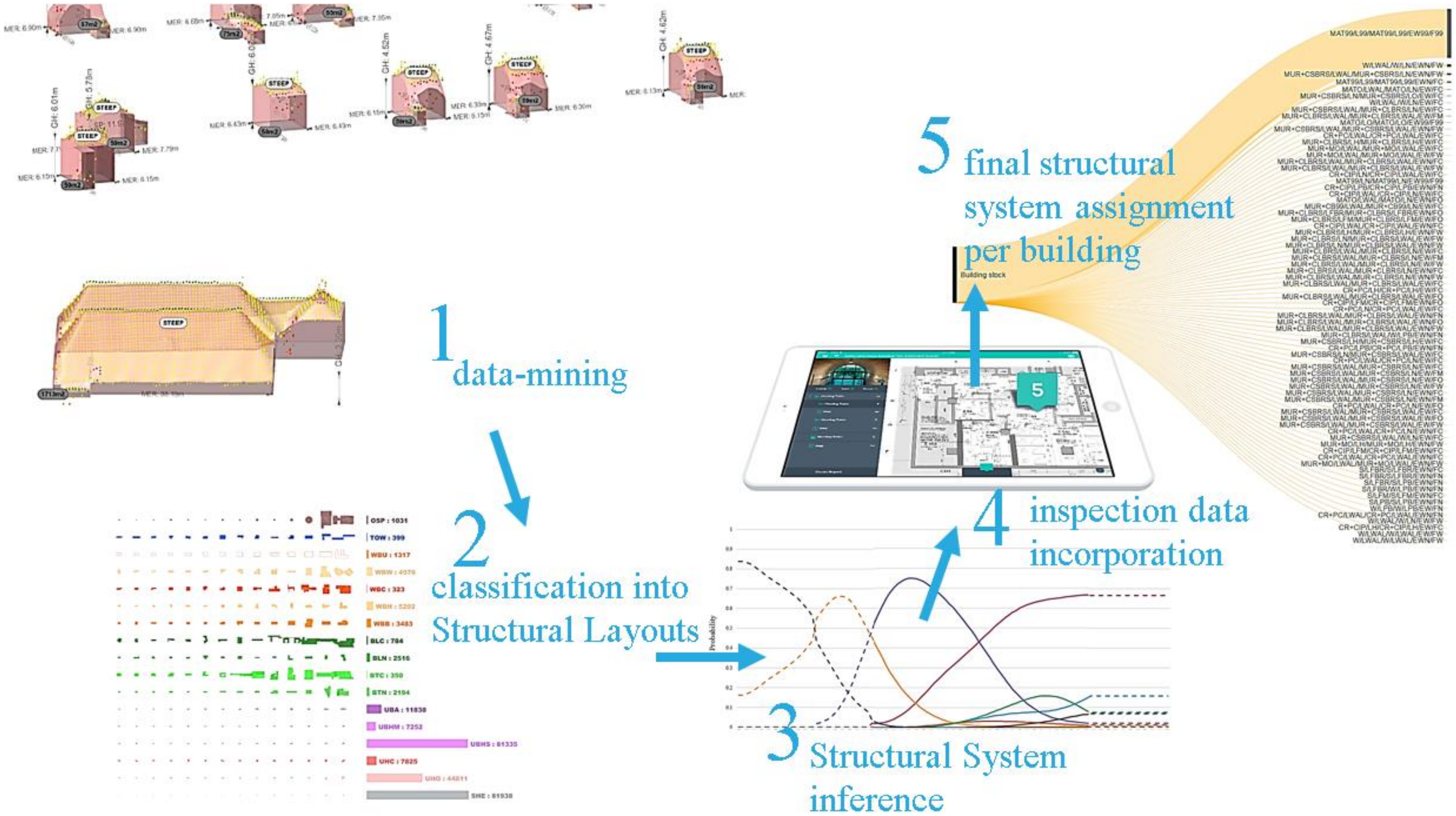


Figure 2 Diagram of EDB V5 calculation phases.

4 Building Data-Mining and Geometrical Characterization

In this first phase, the following main geometric parameters are generated through algorithmic processing of two main datasets, the AHN height map [8] and the BAG building footprint outlines [10]:

1. Width of Maximum Enclosed Rectangle (MER_width) within the footprint outline of a building
2. Length of Maximum Enclosed Rectangle (MER_length) within the footprint outline of a building
3. Gutter Height
4. Roof Steepness
5. Roof Ridge Count

The data-mining algorithms (described in the *EDB V5 Data Documentation* note [8]) allow the clustering of buildings in geometrically uniform classes based on their geometrical parameters instead of their functional use (which was primarily used in EDB V3). The correlation between the various architectural forms of a building and their structural design has been deemed as a more traceable, accurate and measurable starting point for the classification process in the EDB V5 [1]. Applications of the data-mining algorithms, on 3 buildings of different sizes, are shown in Figure 3, while the extraction of building data in a larger set of buildings (urban scale) is shown in Figure 4. This set of buildings will be subsequently used in the document to show the evolution over the building characterisation in the EDB V5 calculation process (Figure 4, Figure 28, Figure 32 and Figure 40).

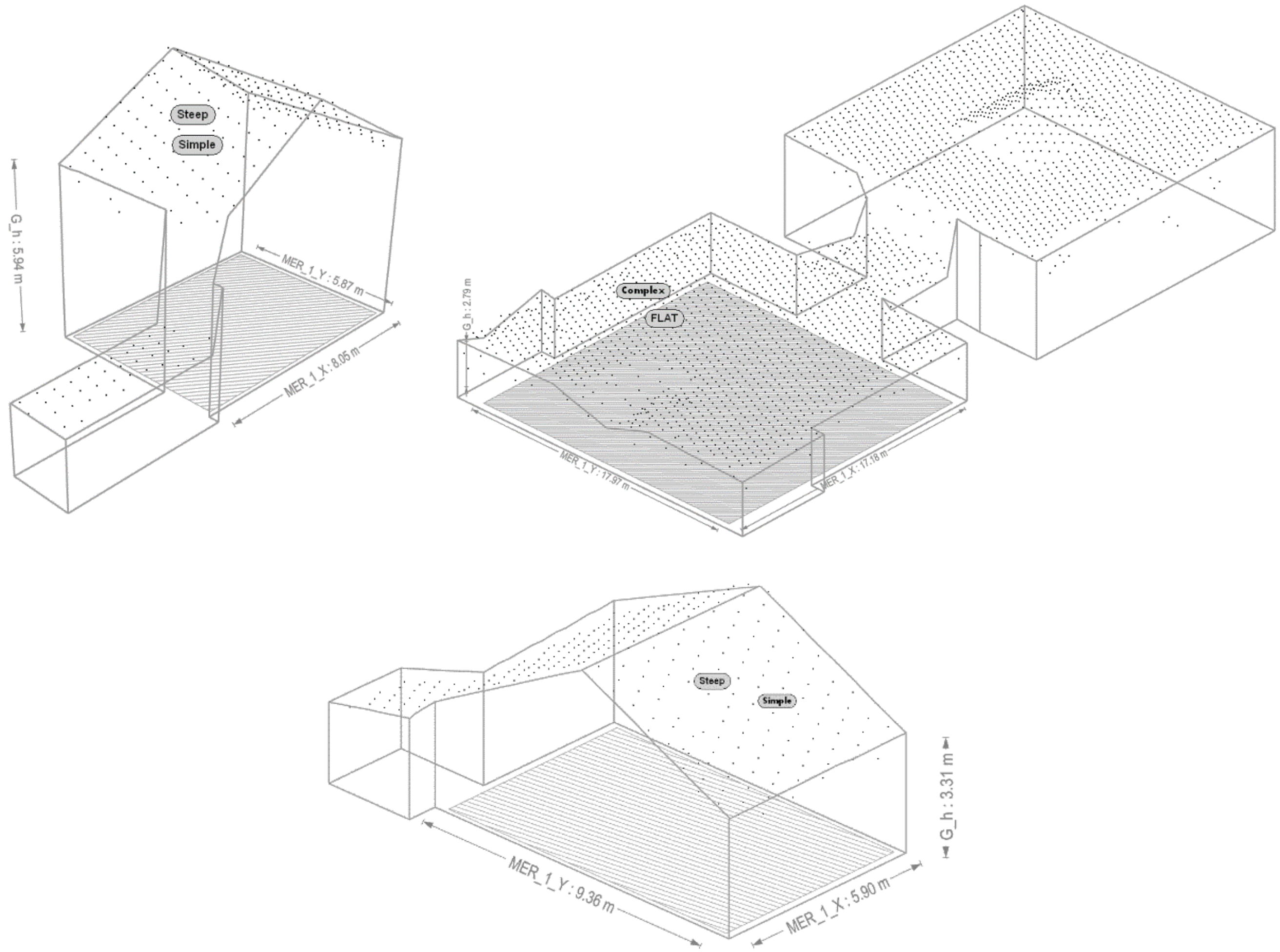


Figure 3 Example application of Arup's data-mining algorithms on three test buildings.

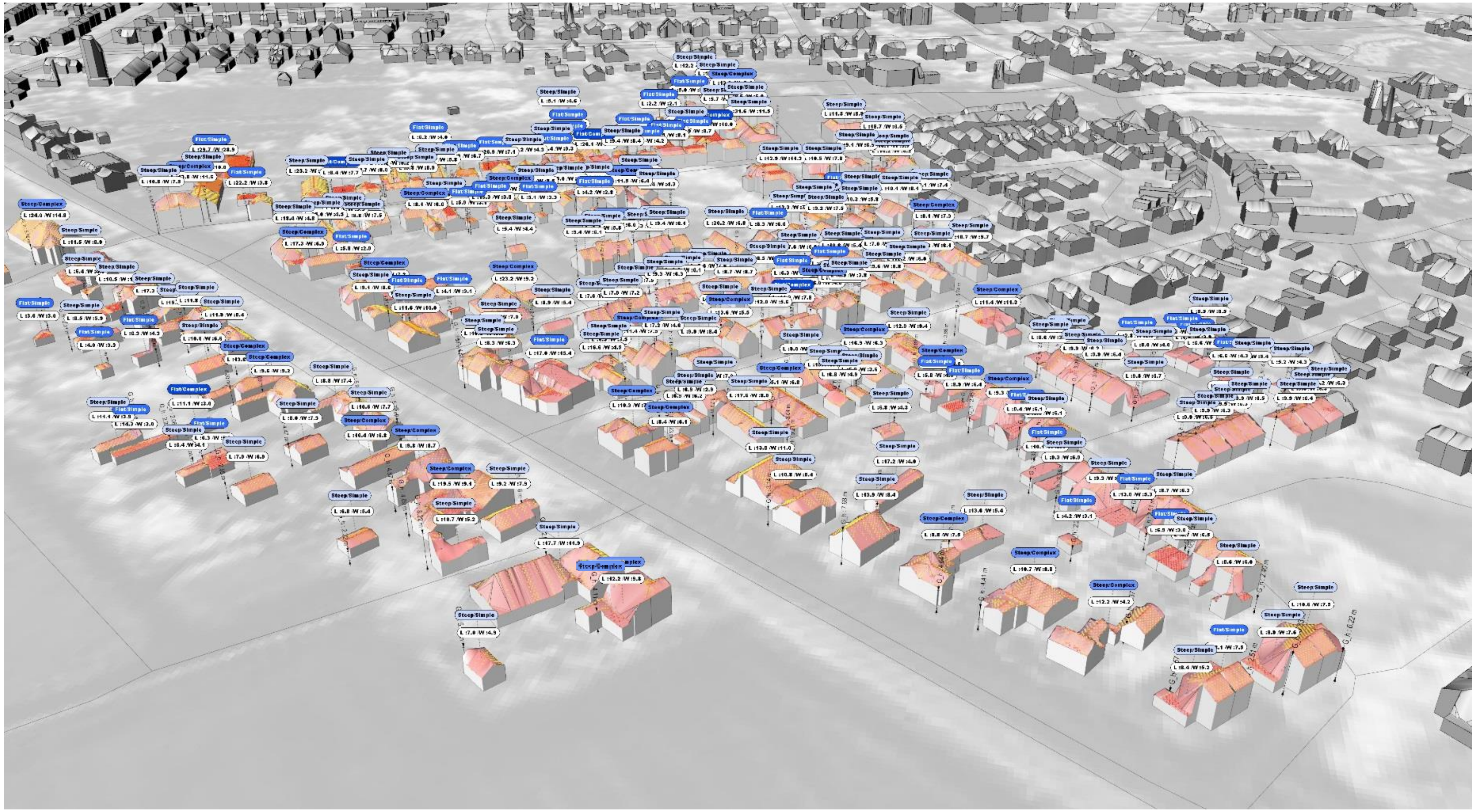


Figure 4 Visualisation of data-mining concerning width (W), length (L), roof steepness (high-steep, low-steep and flat roofs) and roof complexity.

4.1 Data Overview

The data extracted demonstrates the expected large variety in geometries (sizes, combinations of attributes) in the building stock analysed. Figure 6 shows the calculated Maximum Enclosed Rectangle length (mer_length) and width (mer_width) of the footprints on the x and y axis and the calculated gutter height on the z axis, for all buildings of the study area – with the exception of 510 buildings having dimensions beyond the domains of the graph. The length and width of the max enclosed rectangle algorithm have been observed to be consistently calculated and its' use could be potentially expanded to consider the possible architectonical combinations likely given the 2nd and 3rd largest fitting rectangles, as shown in Figure 5.

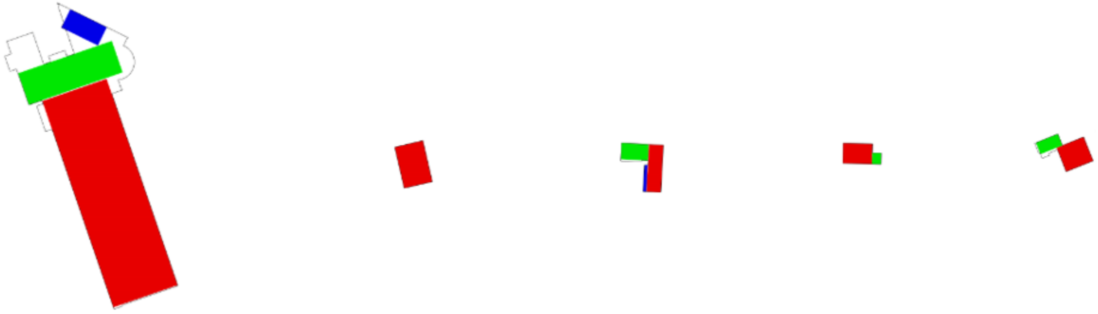


Figure 5 Example demonstration of the identified Max Enclosed Rectangle in building footprint outlines of different sizes (in red). The second and third largest fitting rectangles are also shown in green and blue, when surpassing a size threshold, respectively.

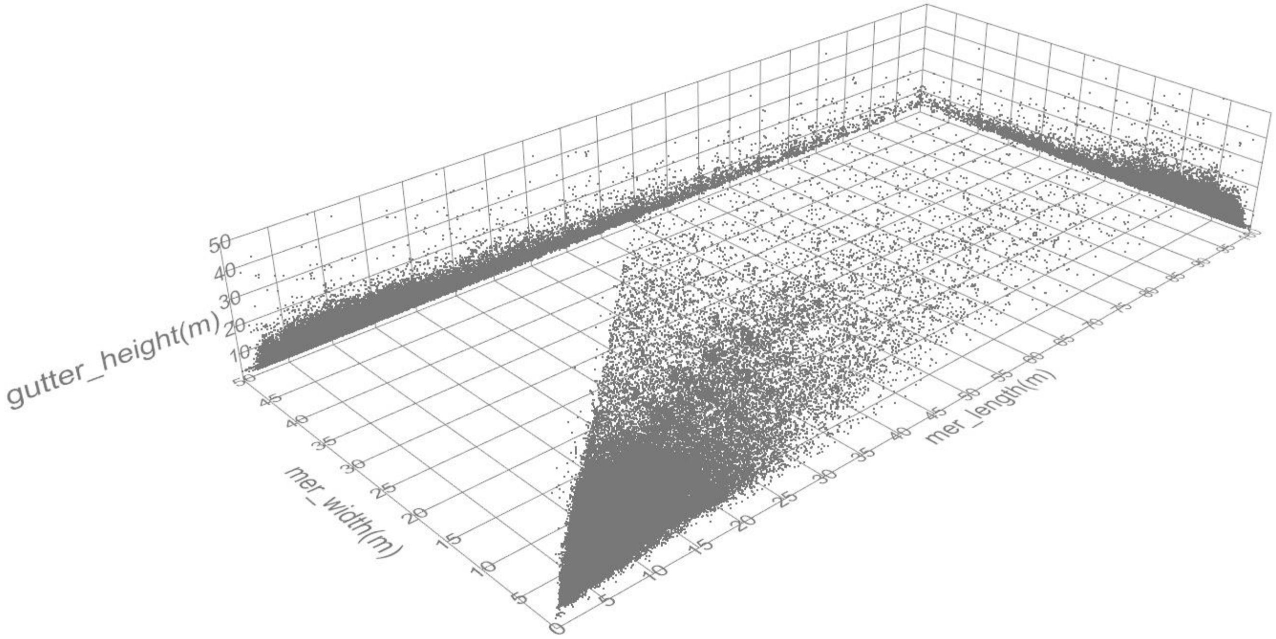


Figure 6 Representation of the 256 664 buildings in the area (510 buildings had at least one of the 3 dimensions larger than maximum value of the axis and have been omitted for visualisation purposes).

As expected, the vast majority of the building stock is characterized by an overall volume included within the 25x25m area and below 20m in height (almost 95% of the total). Figure 7[1] clearly shows that more than 60% of the buildings in the study area has both width and length

[1] It should be noted that mer_length is always considered to be the larger of the two dimensions of the enclosed rectangle, which is why the upper left part of the graph is not populated.

smaller than 10m. The calculated gutter height for all buildings[2], is shown in Figure 8: only 5 134 were calculated to have a gutter height of more than 10m (less than 2% of the buildings in the area).

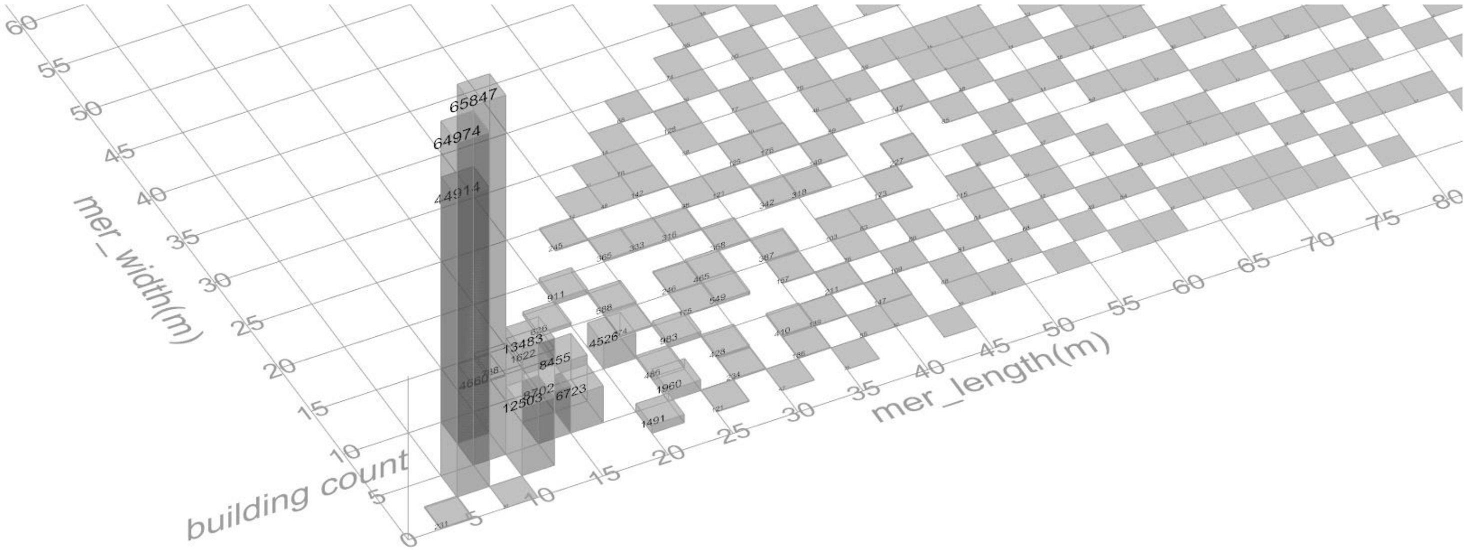


Figure 7 Building count per mer_width and mer_length dimension domain (zoomed where more than 95% of buildings are included).

[2] It is reminded here that for buildings where height map point clouds were not available (i.e. all buildings built after 2009), an alternative methodology was used based on the usable area divided by the footprint area, as described in section 5.1 below.

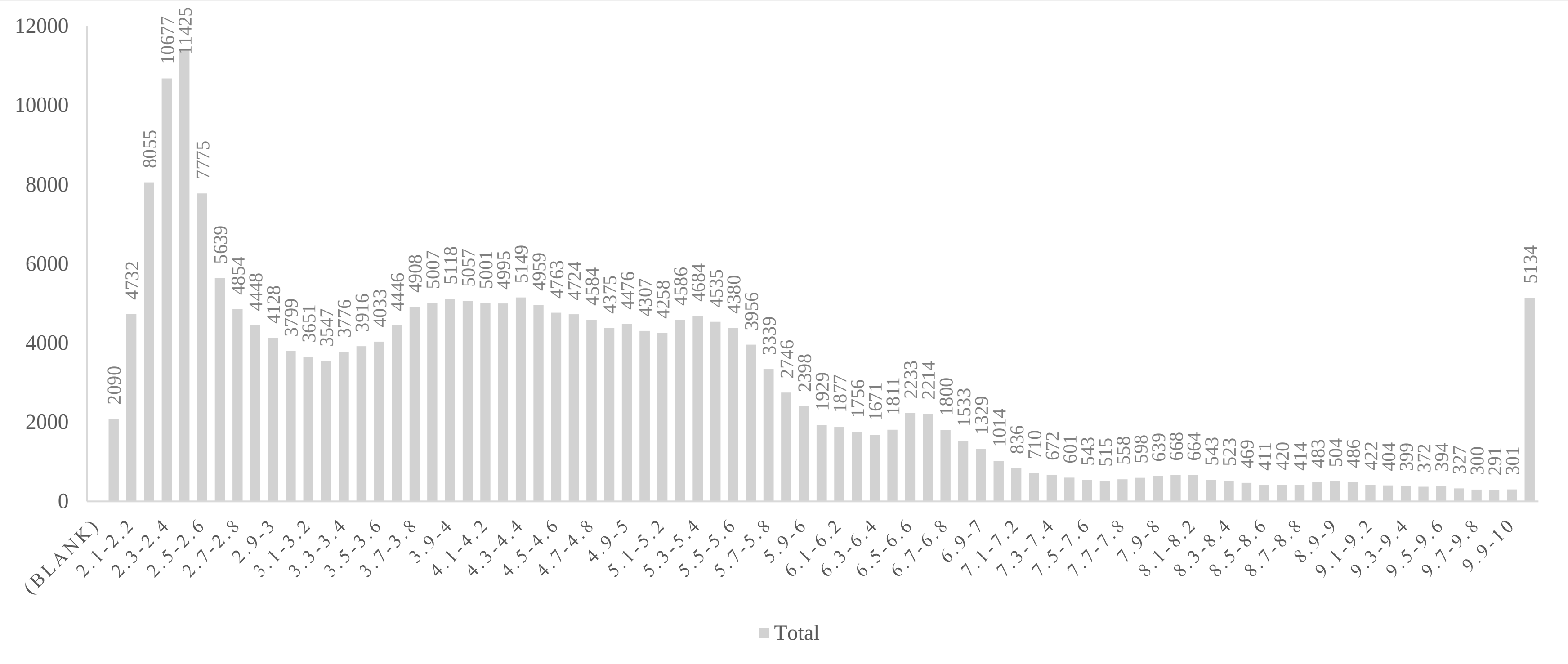


Figure 8 Calculated gutter height of all buildings in the EDB V5

In Figure 9 the additional parameters calculated in this phase are presented: the amount of roof ridges identified (“roof counts”), the calculated roof steepness class, the presence of residential use and the calculated building adjacency. A comparison of the data-mining algorithm for the gutter height against inspection data has been documented in a dedicated exercise [12], where the trend found in the algorithmically calculated data was found to be adequately aligned with the trend based on measurements on corresponding technical building drawings.

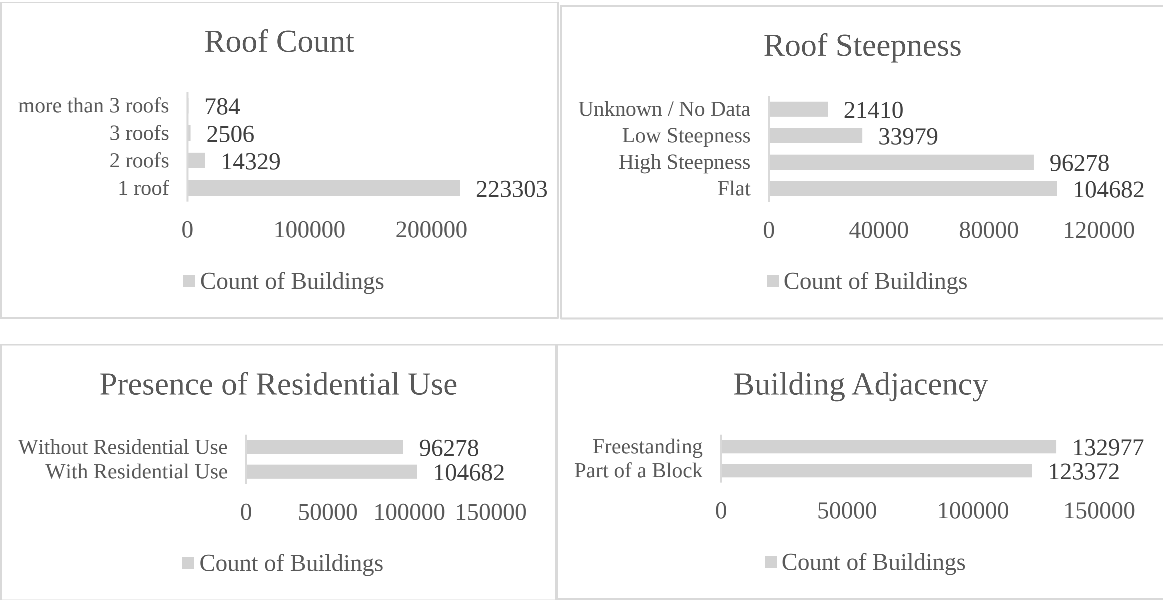


Figure 9 Roof ridge count, roof steepness, presence of residential use and building adjacency results for all buildings in the EDB V5[3].

[3] As shown in the graph, there are 73 564 with “0” roof ridges identified, which is mostly the case for flat roof buildings, although the possibility of the algorithm missing a ridge due to pointcloud data noise cannot be excluded.

5 Structural Layout Classification

The first of the two classification processes consists of the classification into defined Structural Layouts. These are classes based on the expected arrangement of vertical and lateral resisting systems in response to specific geometrical and architectural needs (e.g. maximum floor span length, number of storeys, building height, etc.). This classification contains two main categorisation stages followed by the inclusion of building specific inspection data, as used in EDB V5 [1][4]:

1. Geometry classification:
   Geometric parameters are used to classify the building stock into five Geometric Layout classes

2. Structural Layout class:
   The five Geometric Layout classes are further sub-classified into 17 Structural Layouts (as reported in Table 1) using additional building attributes. Note that Structural Layout class *Special* is only assigned through dedicated project or inspection data.

3. Inspection and project data inclusion:
   Inspection and project data are used to override the Structural Layout resulting from previous classification stages.

[4] To keep track of data gaps, outliers and inclusion of additional data, a series of flags have been implemented throughout the process which are used to create the 'SL_FLAG' field (a flag corresponding to classification confidence) within the final extract.

Table 1: Structural Layouts in EDB V5

Structural Layout class in the EDB V5 extract	Structural Layout short description	Geometric grouping intention
SHE	Shed	Buildings with small dimensions.
UHO	House	Medium sized buildings with small span which are freestanding.
UHC	House Complex	Medium sized buildings which are freestanding formed by multiple small span structures.
UBHS	Block Unit Single	Medium sized buildings which are touching other similar buildings, forming a uniform[5] or homogeneous block and has a single address.
UBHM	Block Unit Multiple	Medium sized buildings which are touching other similar buildings, forming a uniform or homogeneous block and has multiple addresses
UBA	Aggregated Unit	Medium sized buildings which are touching different buildings, forming an inhomogeneous block.
BTN	Block Tall	Buildings expected to contain horizontal and vertical repetitions of unit structures without large spans. ‘Tall’ is designated where a gutter height is larger than 10m.
BTC	Block Tall Complex	Buildings formed by multiple structures expected to contain horizontal and vertical repetitions of unit structures without large spans. ‘Tall’ is designated where a gutter height is larger than 10m.
BLN	Block Low	Buildings expected to contain horizontal and vertical repetitions of unit structures without large spans. ‘Low’ is designated where a gutter height is smaller than 10m.

[5] In this case uniformity-homogeneity signifies a block of building units that are part of the same design and were built as part of the same construction project.

BLC	Block Low Complex	Buildings formed by multiple structures expected to contain horizontal and vertical repetitions of unit structures without large spans. ‘Low’ is designated where a gutter height is smaller than 10m.
WBB	Barn	Single storey buildings with a pitched roof with a large span.
WBH	Barn with House	Buildings which are formed by WBB and UHO structures, combined.
WBC	Barn Complex	Buildings which are formed by multiple WBB structures.
WBW	Warehouse	Buildings with flat roof, large span structure.
WBU	Barn / Warehouse	Buildings with a large span structure and pitched roof.
TOW	Tower	Multi-storey/tall buildings.
OSP	Special Geometries	Buildings with indications of atypical structural characteristics.

5.1 Geometric Layout Classification

All buildings within the EDB V5 are initially classified in the following five Geometric Layouts [1]:

- S (*Shed* Geometric Layout)
- U (*Unit* Geometric Layout)
- B (*Block* Geometric Layout)
- W (*Barn/Warehouse* Geometric Layout)
- T (*Tower* Geometric Layout)

To assign each specific building Geometric Layouts, the classification process follows the algorithmic steps described below.

1. Geometric parameters are calculated using BAG building footprint and AHN height map data. These parameters include:
    a. Length of maximum enclosed rectangle
    b. Width of maximum enclosed rectangle
    c. Gutter height.

    Where height data are not available or a gutter height is deemed unreliable (i.e. when lower than 2.0m), the useable floor area provided by BAG is used to estimate the height of the building. The details of the analysis used to create the geometric parameters can be found in the Data Documentation note [11].

2. Reference buildings have been identified for each Geometric Layout to create a Learning Set (i.e. a set of buildings used to train the algorithmic classifier).

3. The buildings within the Learning Set and their respective geometric parameters have then been used to create a lognormal probability density function per each of the 3 geometric parameters and for each Geometric Layout. The resulting probability density functions are shown in Figure 13.

Each building has been classified to the most likely Geometric Layout based on its geometric parameters and the corresponding likelihood of each Geometric Layout. This is done using the probability density distributions of step *3* and Bayes' theorem, without taking into account the correlation between the geometric classification parameters of [1][6] (Annex A4).

[6] To help with some identified special cases of Geometric Layout misclassification, two additional checks were implemented after the 4 steps described in this document. Additional information can be found in the dedicated note [1].

The process to classify the building stock on these five Geometric Layouts is shown diagrammatically in Figure 10, and on the scatterplot visualisations of Figure 11.

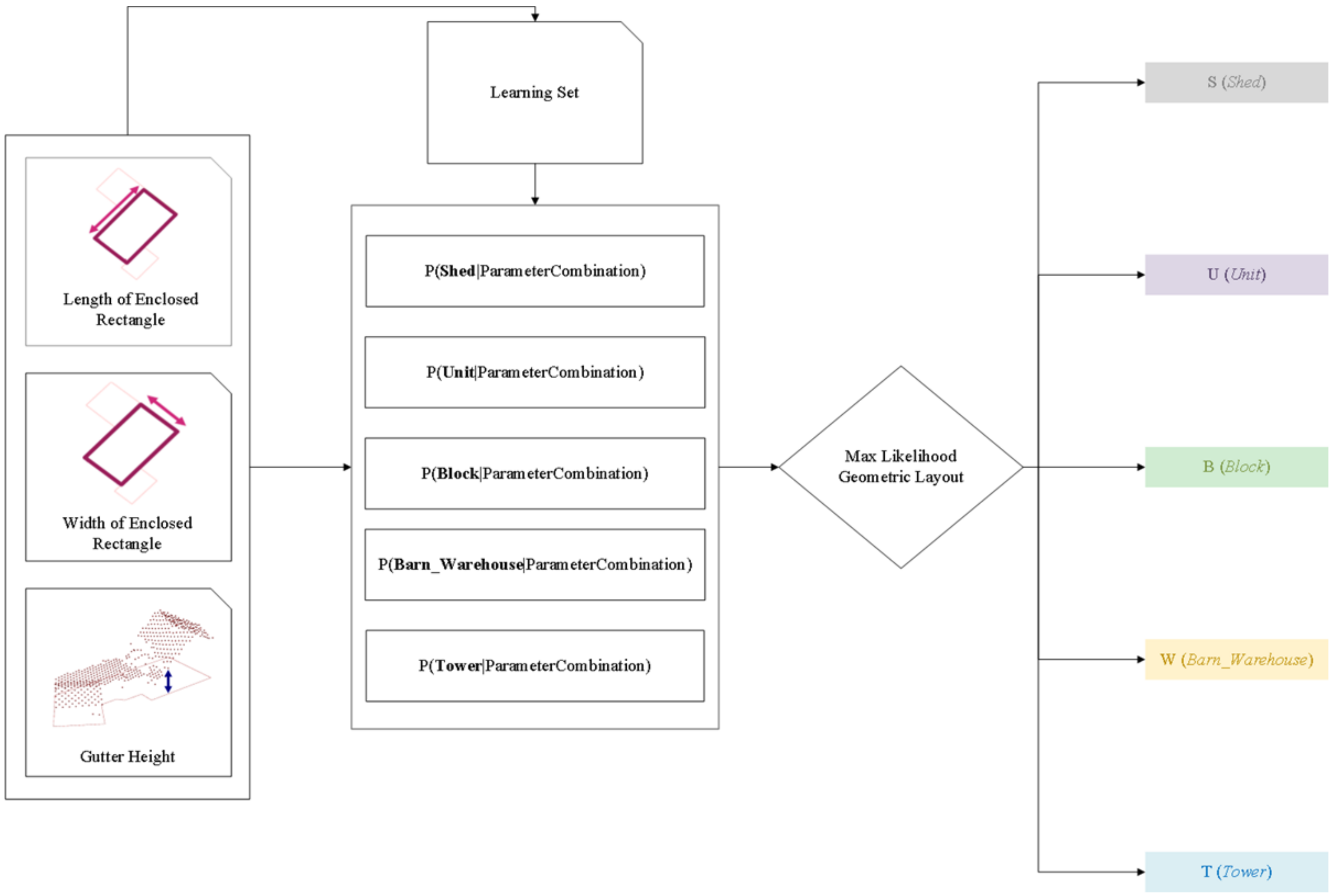


Figure 10: Stage 1 – Assigning Geometric Layouts for the Structural Layout classification

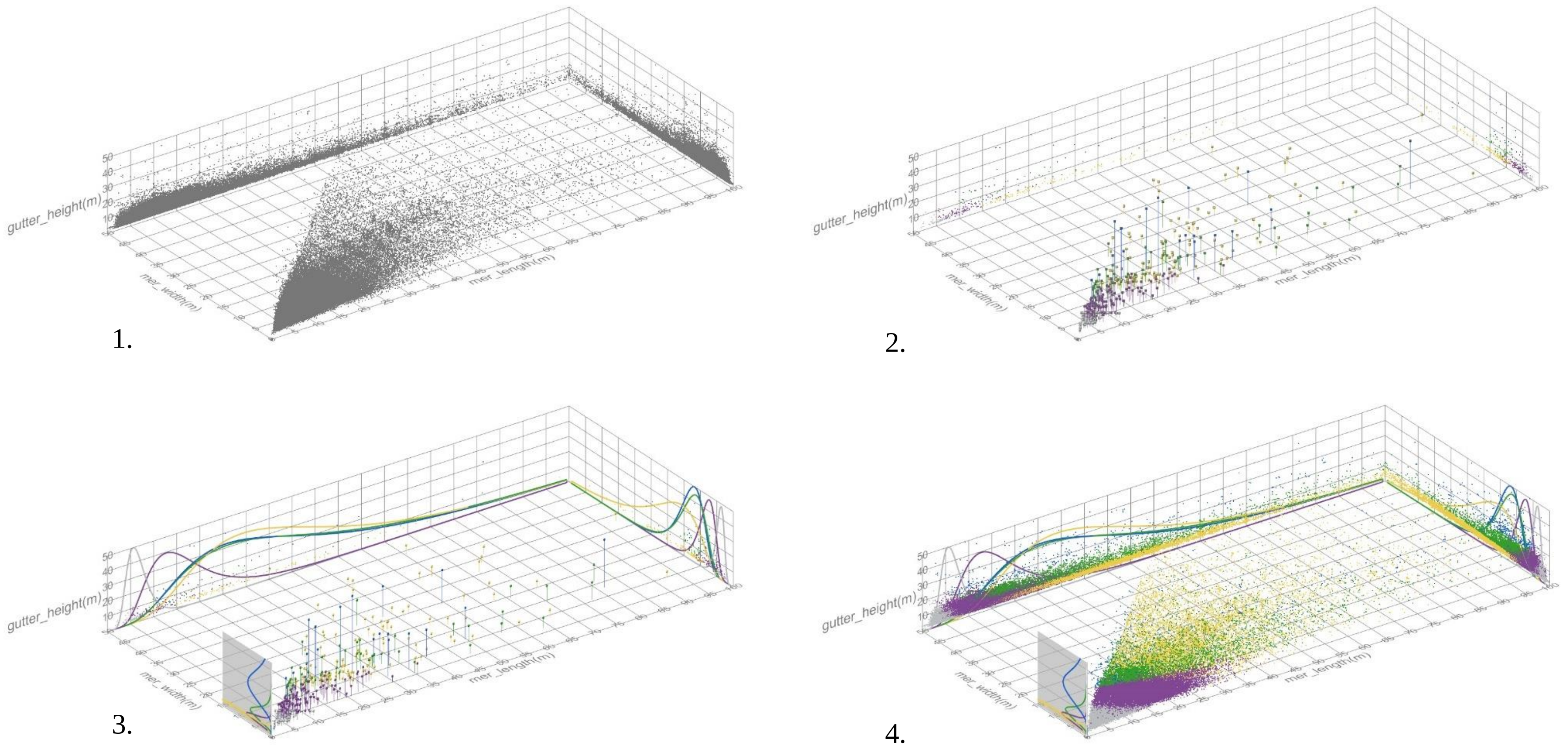


Figure 11 Algorithmic steps for the classification of buildings within the EDB V5 scope in the one of the five EDB V5 Geometric Layouts.

5.1.1 Learning set and classifier functions

376 reference buildings formed the Learning Set for Geometric Layout classification used in EDB V5 [1] (steps *2* and *3* of the process illustrated in Figure 11, and as presented in the previous section). The Learning Set has been populated with the aim of offering a fair coverage of the Geometric Layouts with regards to the geometric variability and recurrence of the geometric characteristics. This resulted in a weighted sampling, with more samples on buildings with more recurrent dimensions (e.g. more samples for buildings below 25x25m MER dimensions).

The average dimensions of Geometric Layouts resulting from the Learning Set created are shown in Figure 12, while their the log-normal probability distributions are presented in Figure 13. These distributions are plotted against their respective learning set in Figure 14 (step *3* of the Geometric Layout classification process). Based on these probability density functions the rest of the building stock is subsequently classified (step *4* of Figure 11).

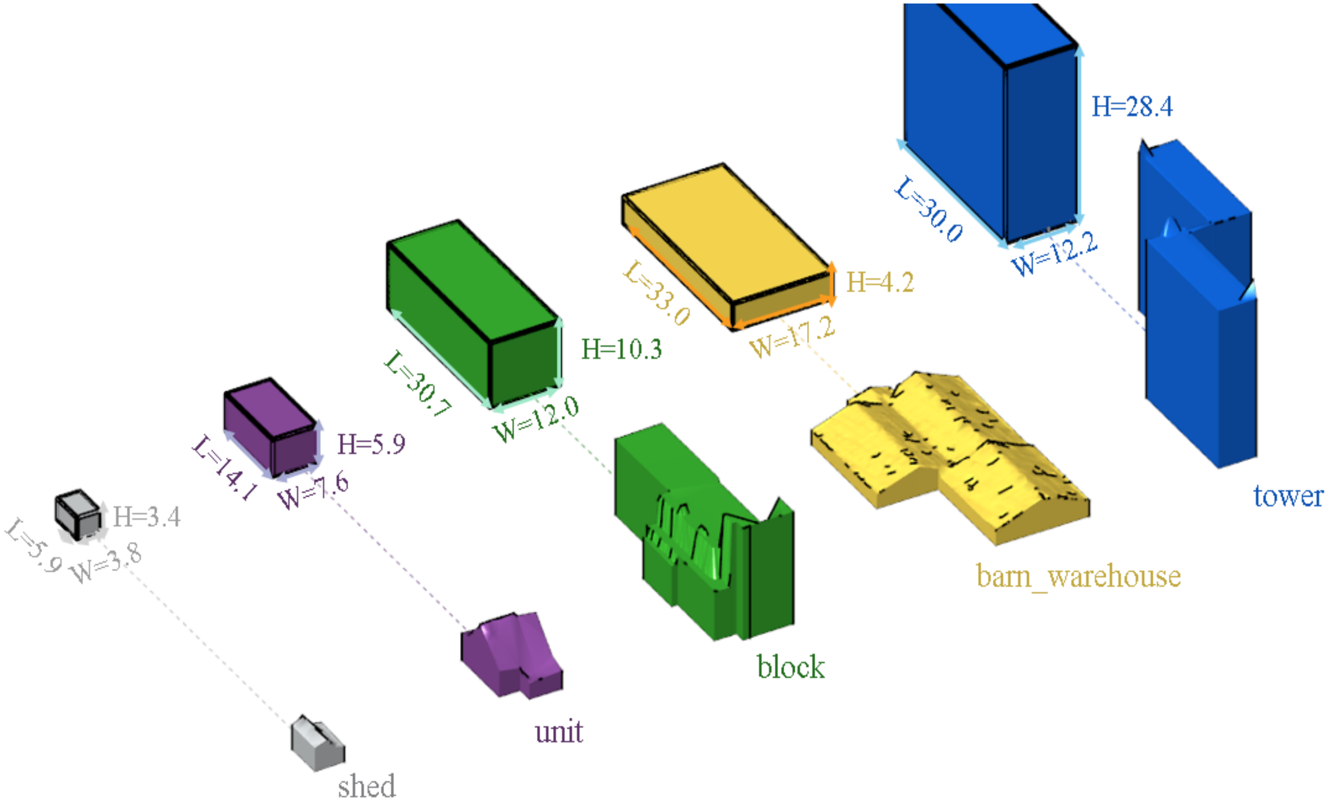


Figure 12 Average Learning Set dimensions (length-L, width-W and gutter height-H) for each EDB V5 Geometric Layout

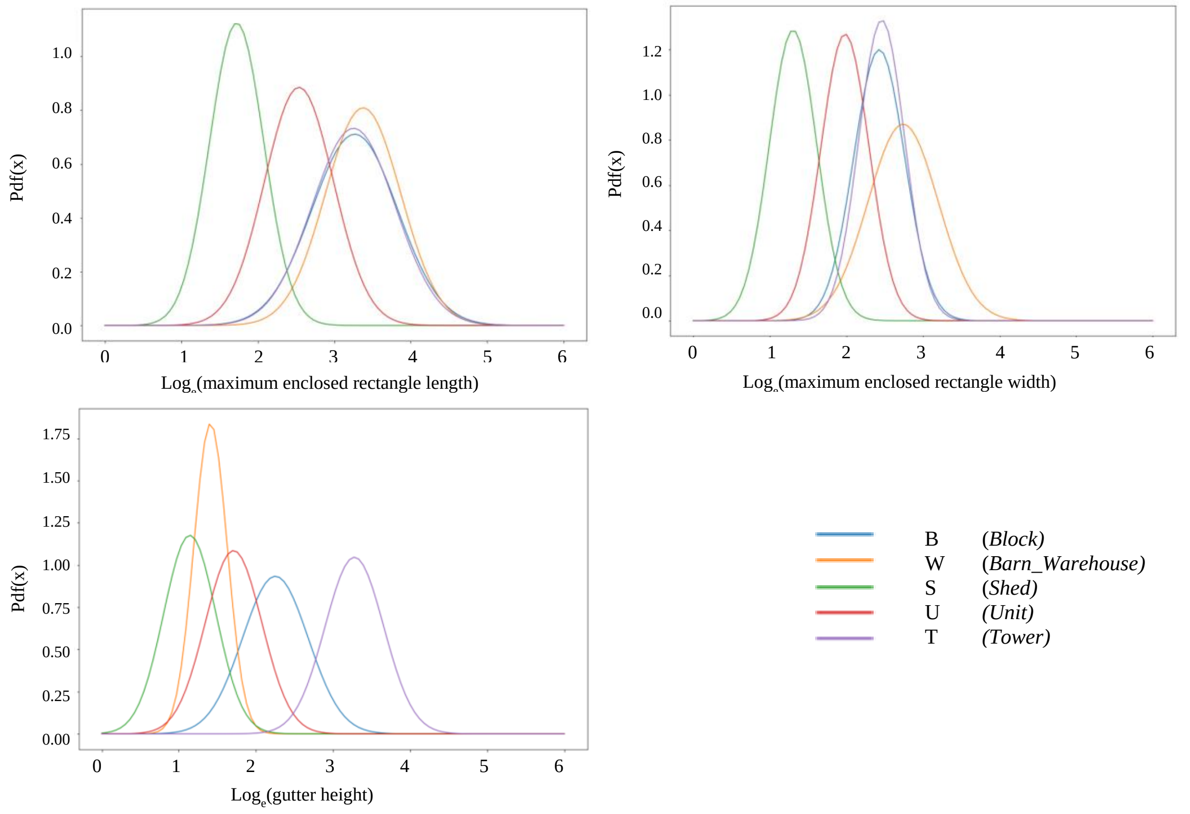


Figure 13: Probability density function per geometrical layout based on the length – top left, width - top right and gutter height - bottom left of the maximum enclosed rectangle.

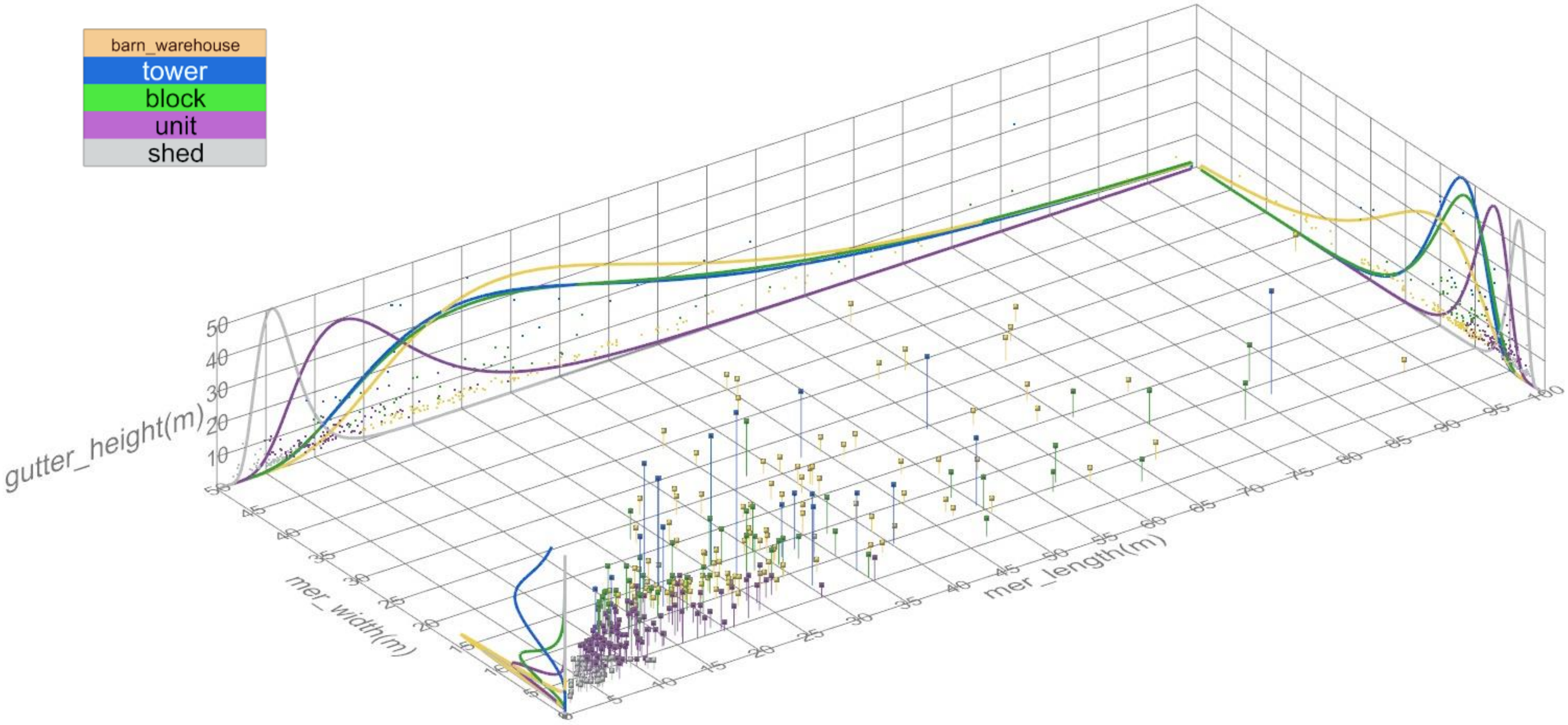


Figure 14 Learning set scatter plot and resultant probability density functions per parameter.

The final part of the process has been exemplified in Figure 15, where, given the developed probability density functions, the Geometrical Layout is assigned according to the MER width and length dimensions for four different gutter height values. The radius of the circles in these visualisations corresponds to the calculated confidence of the classification, related to the likelihood of the first most likely Geometric Layout divided by the likelihood of the second. As expected, at the boundaries between Geometric Layouts the confidence reduces. This has been captured in a flag corresponding to classification confidence ("sl_flag") in EDB V5 [1].

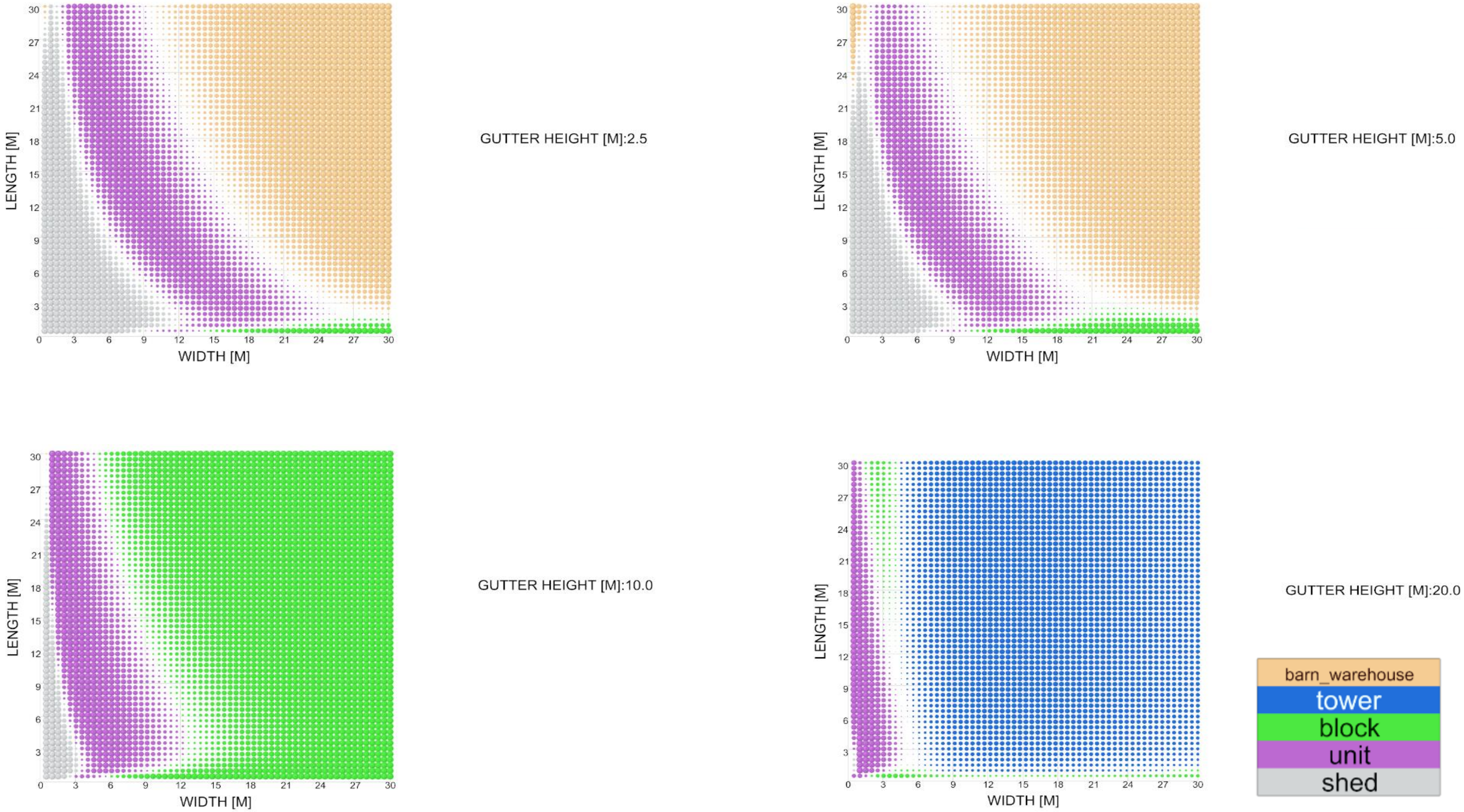


Figure 15 Classification maps for four different gutter heights.

5.1.2 Geometric layout results and validation

The resulting classification in the five Geometric Layouts is shown in Figure 16 for the total building stock. As also shown in Figure 15, the area where Units are more likely is relatively small, however it corresponds to the dimensions where the density of the buildings in the Groningen region is highest. This results in Units being the predominant Geometrical Layout. On the other side of the spectrum, Towers are the predominant class when gutter heights is over 20m but, as presented in Figure 8. However, the amount of buildings with such a height is relatively limited, low-rise buildings being the most typical and recurrent types in the area.

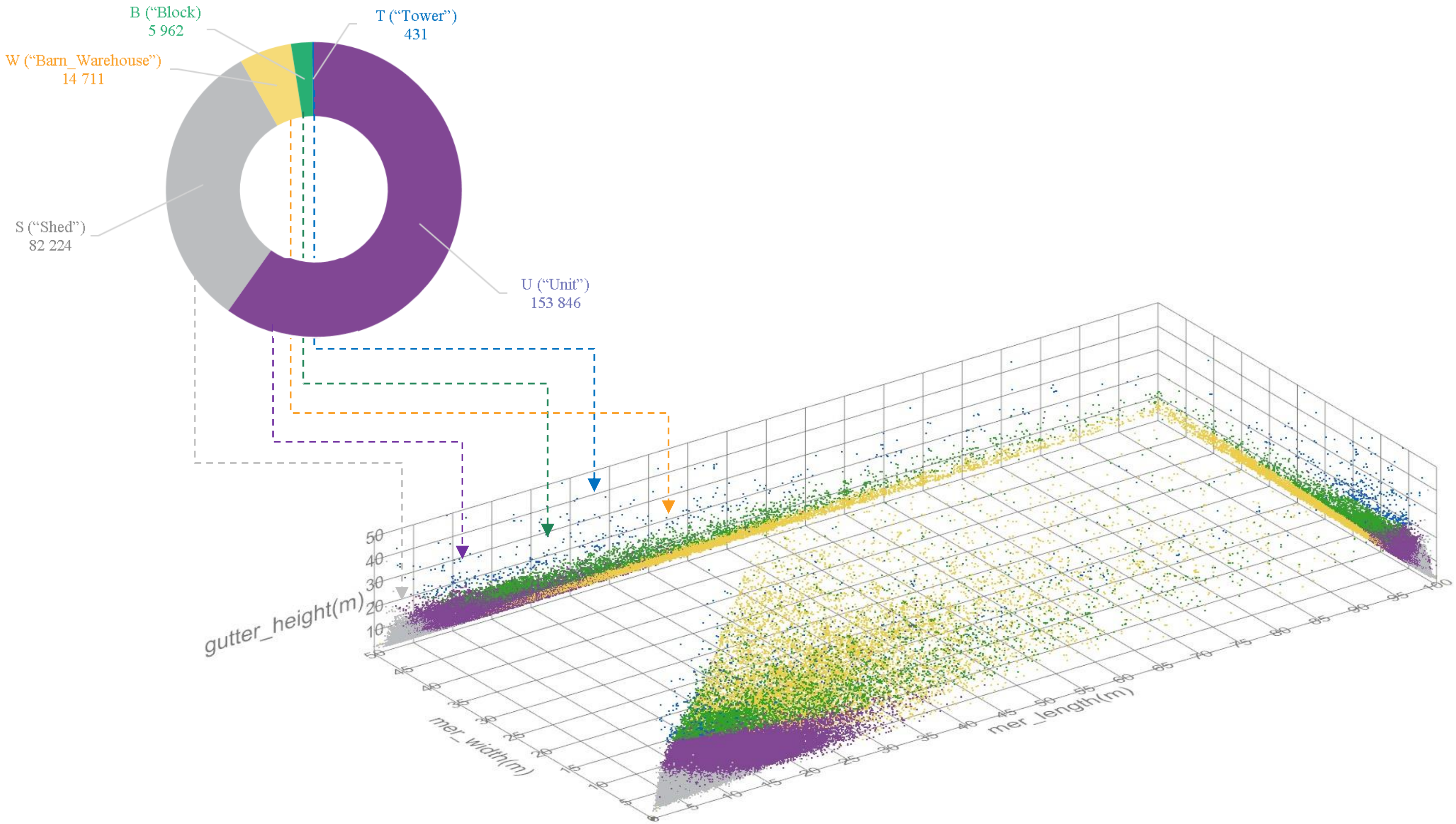


Figure 16 EDB V5 Geometric Layout classification

In Figure 16, the relative density in different areas of the scatter plot cannot be precisely appreciated. For this reason histograms per geometric parameter have been added in Annex B. The Geometric Layout classes can be analysed not only with regards to their geometric parameters but also with regards to other parameters like the number of addresses (found in the same Annex) and the building year (shown in the histogram of Figure 17).

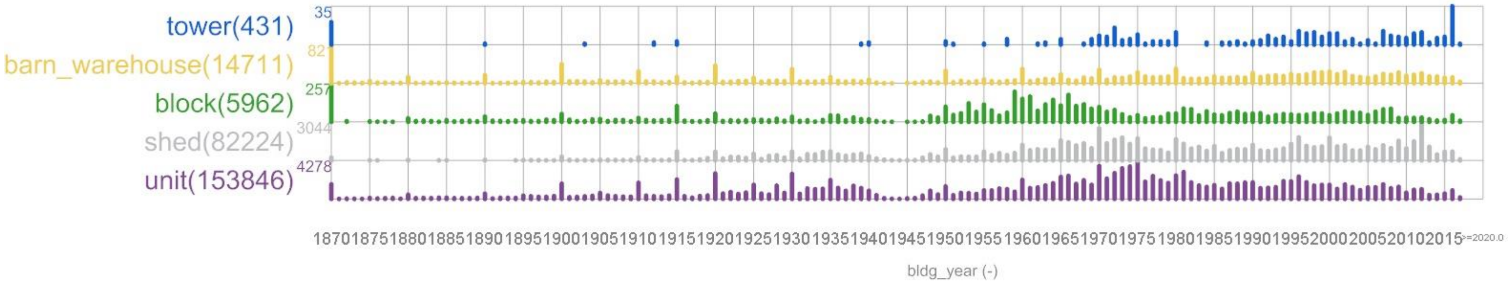


Figure 17 Histogram shown the count of buildings per year for each Geometric Layout. It should be noted that the y-axis of each Geometric Layout is re-normalised for demonstration purposes.

The 376 buildings in the learning set have been tested against their most likely class according to the Geometric Layout classification algorithm described, as applied in the EDB V5. This results in 84% of the buildings within the learning set being correctly classified, while only 16% of the set do not fully replicate the expected Layout. All buildings within the latter group belong to areas of either low confidence (14 out of the 60 non-matching buildings) or are outliers of their class but in regions of high confidence for other Layouts. These results have been deemed satisfactory for the regional approach used in the Risk Model.

5.2 Structural Layout

The additional classification parameters, such as building function, footprint area, roof steepness and count, adjacency, number of addresses and gutter height help to further classify buildings belonging to each Geometric Layout into the EDB V5 Structural Layouts, as provided in Table 1. The full process for the definition of these Structural Layout is summarised in the process flow in Figure 18. In this process, important distinctions in the Layouts are made, especially to buildings belonging to the Unit Geometric Layout, where buildings are classified into UHO (House), UHC (House Complex), UBH (Block Unit) or UBA (Aggregate Unit) Structural Layouts. Other distinctions, e.g. WBC (Barn Complex) instead of WBB (Barn), play a less significant role in the final classification towards the assignment of Structural Systems, but support further characterization of the building stock and could guide the selection of reference structures. The resulting Structural Layout classification is visualised at an urban scale in Figure 19. This is done by using an example subset of the building stock, but the same process has been executed at the same detailed level for the complete study area.

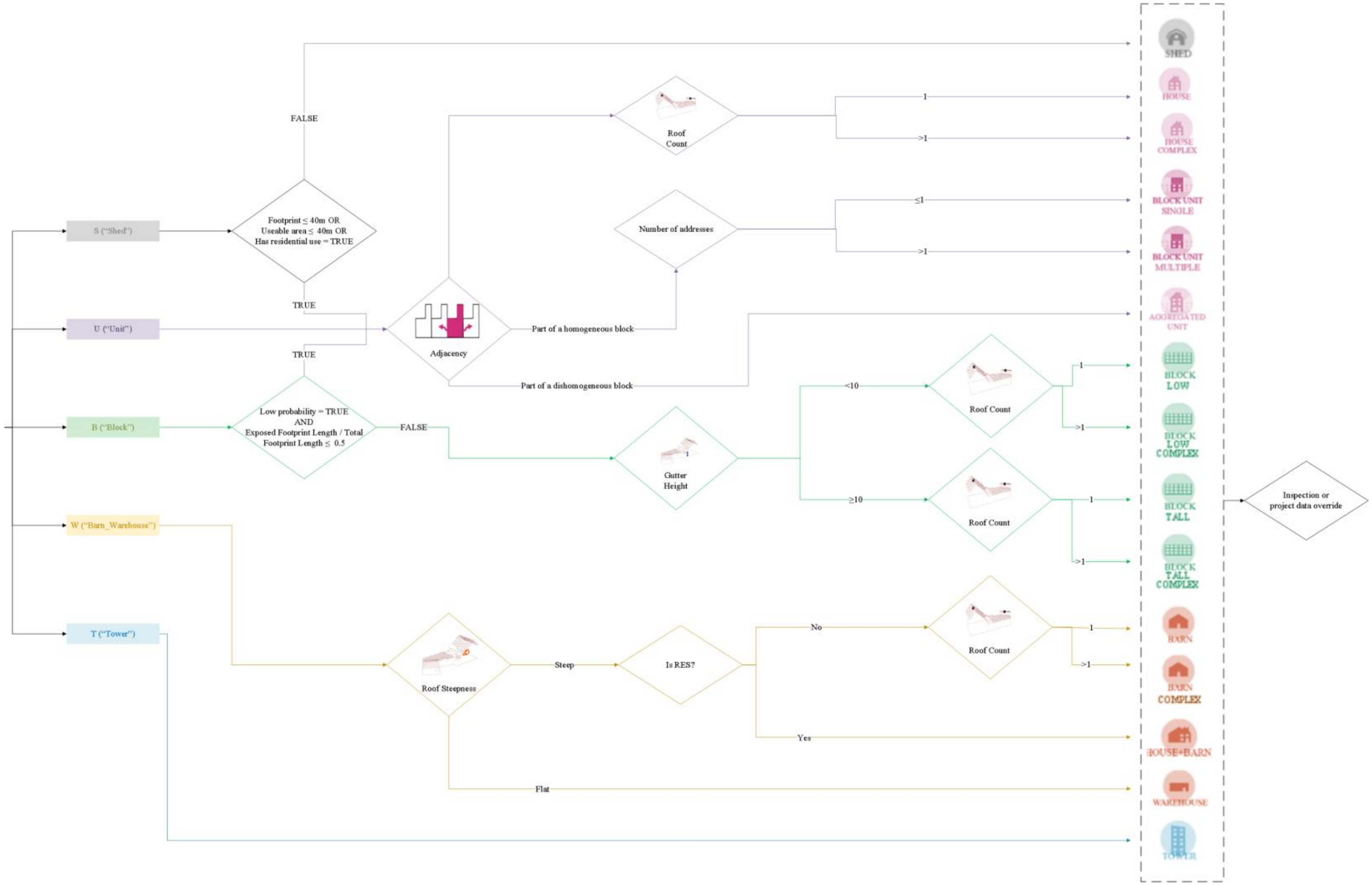


Figure 18: Structural Layout classification flowchart

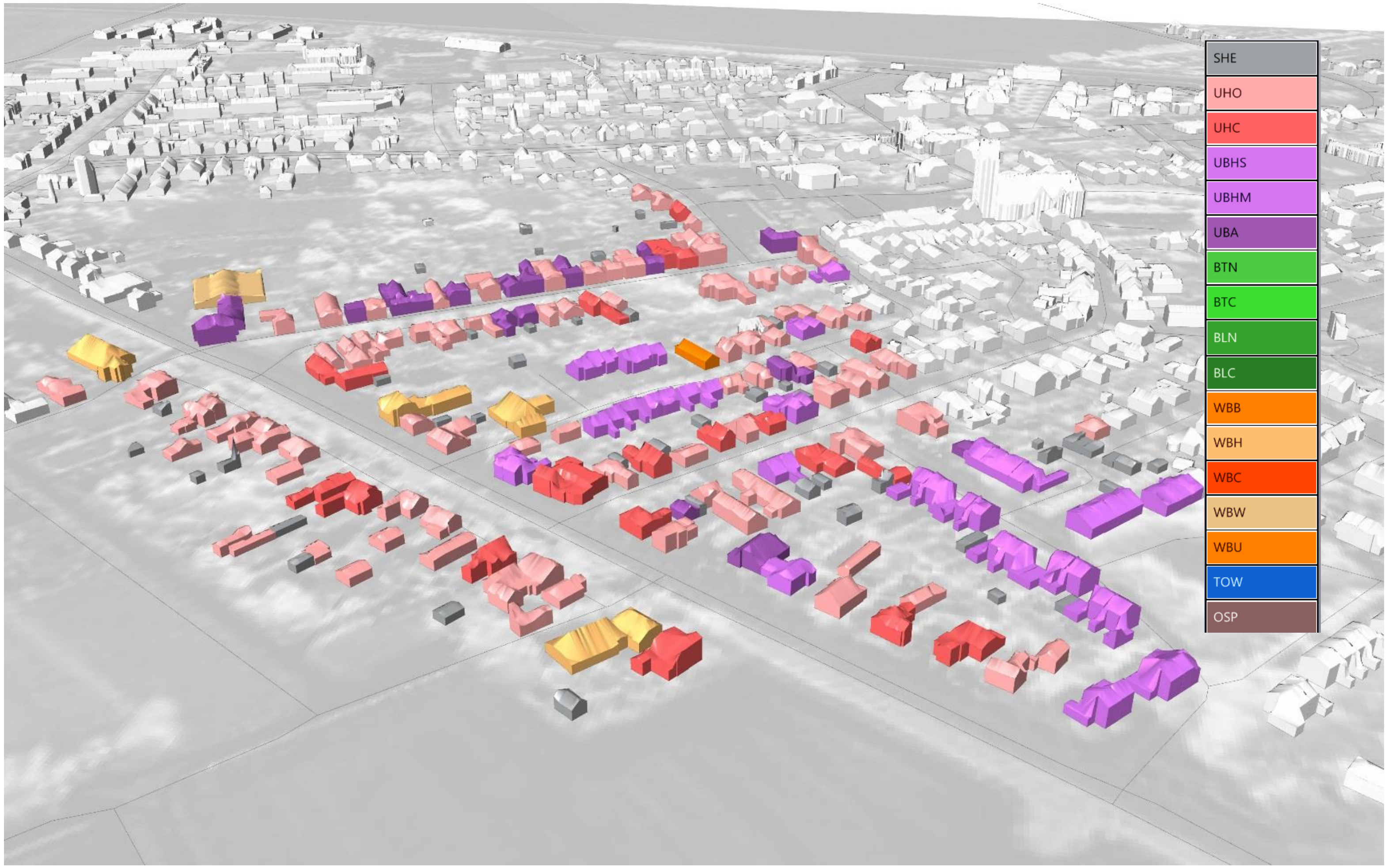


Figure 19 Structural Layouts of the example buildings within the region (same buildings shown previously in Figure 4.)

5.2.1 Structural layout results and validation

The final counts per Structural Layout class in EDB V5 are shown in Figure 20, together with one example building per class. As shown in the figure, buildings in the SHE *(Shed)* Structural Layout, with relatively small dimensions, represent a big portion of the building stock (approximately 32% of the buildings in the area, although unlikely to have population exposed in them), followed by another 32% of buildings being in the UBHS *(Block Unit Single)* Structural Layout, and where most of the refinement for the current version of the EDB has been focused.

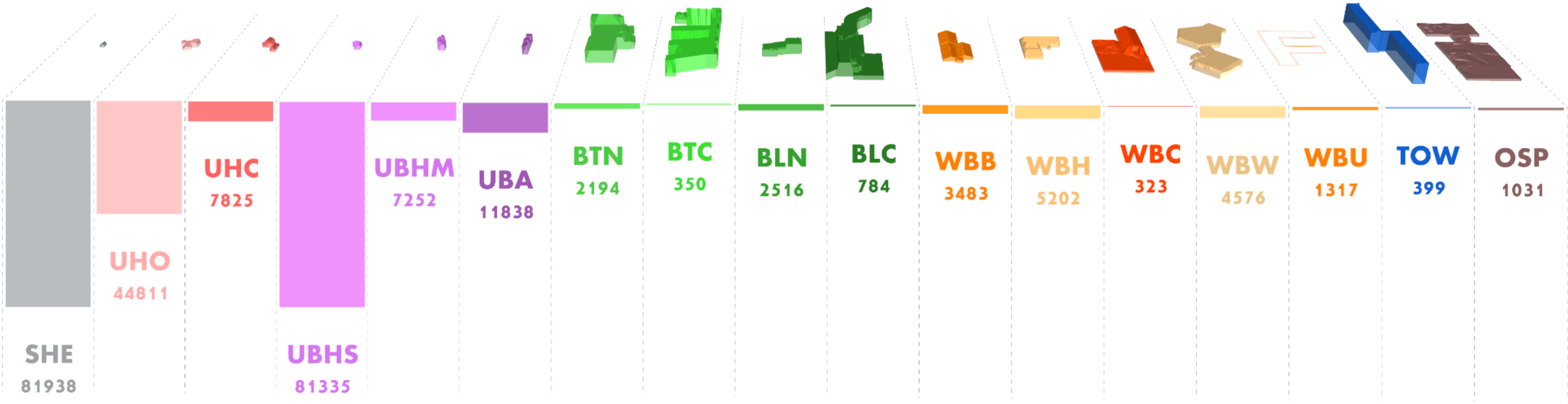


Figure 20 Count of buildings per Structural Layout visualisation together with an example building per class, for size reference.

Figure 21 shows the amount of buildings from each of the five EDB V5 Geometric Layouts[7] that are classified to each of the EDB V5 Structural Layouts, in order to demonstrate the direct link between these two classes.

The resulting classification in the 17 Structural Layout is shown in Figure 22 for the total building stock[8], against the most relevant geometric attributes that drive the Geometric Layout classification process.

The most recurring building years per Structural Layout are shown in the histogram of Figure 23. This diagram confirms many trends that would be intuitively expected for example:

- Buildings in the TOW *(Tower)* Structural Layout are predominantly built after the 1970.
- UBH *(Block Units)*, WBW *(Warehouse)* and all Block Structural Layouts buildings are predominantly built after 1945
- On the other hand, UHO *(House)*, UBA *(Aggregate Unit)* and WBH *(Barn with House)* buildings have a big portion of them built before 1940.
- All Structural Layouts have a dip in their recurrence between 1940 and 1945[9].
- The WBU *(Barn_Warehouse)* Structural Layout is mainly assigned after 2009 because after that period there is no AHN height map data available for the region, therefore it could not be distinguished if a building is best fitting to the WBB *(Barn)* or the WBW *(Warehouse)* Structural Layout[10].

---

[7] Additional example of building geometries (generated using AHN height map and footprint outline BAG data) per Structural Layout are shown in the Annex A
[8] Excluding the 557 buildings that have dimensions beyond the domains of this graph
[9] Expected due to the World War II and its effects in the region.
[10] As described in Table 1 the main distinction for these two types has been the roof steepness, as calculated from the AHN height map.

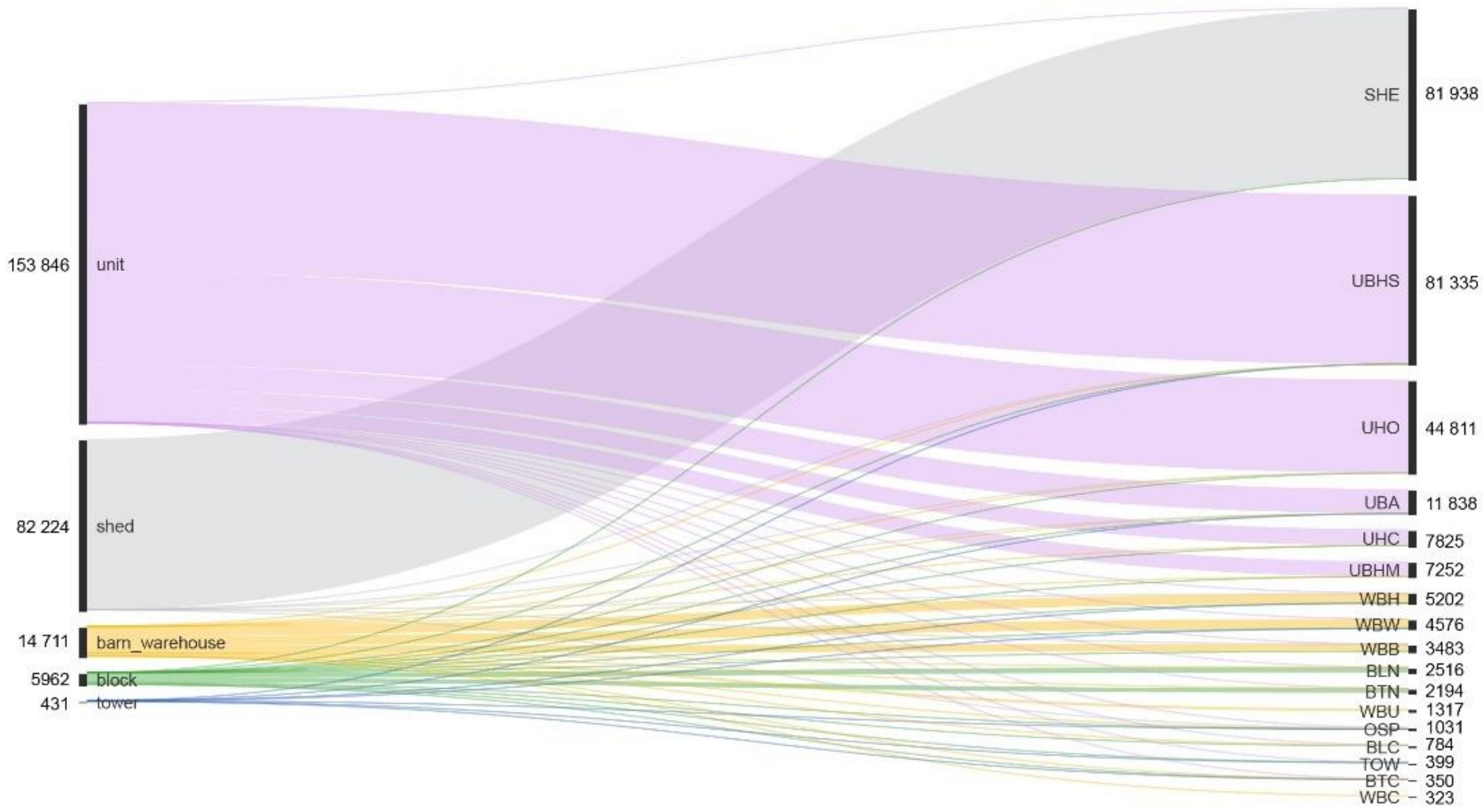


Figure 21 Sankey diagram demonstrating the “flow” of buildings from each Geometric Layout to each of the EDB V5 Structural Layouts [1]

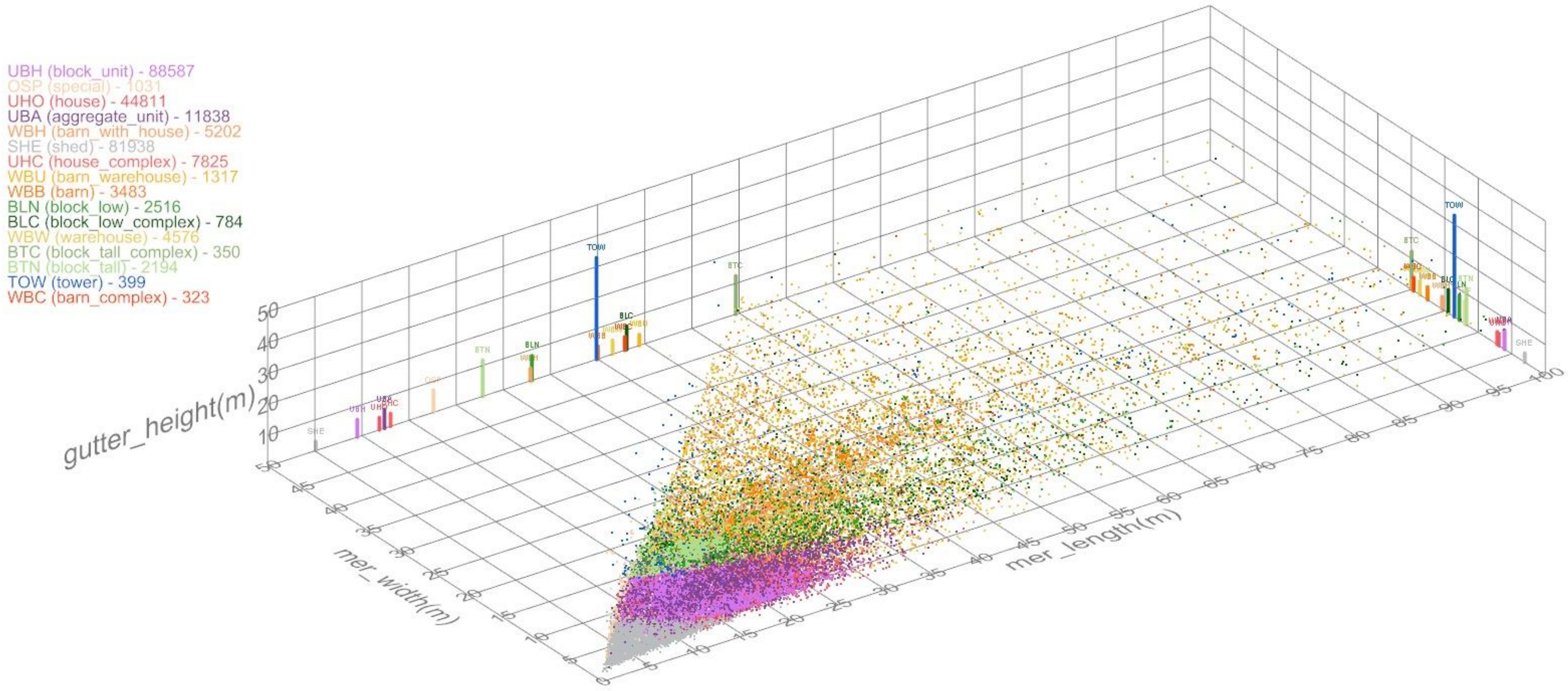


Figure 22 Structural Layout of each building (represented by a data point), together with the average dimensions of buildings within each Structural Layout.

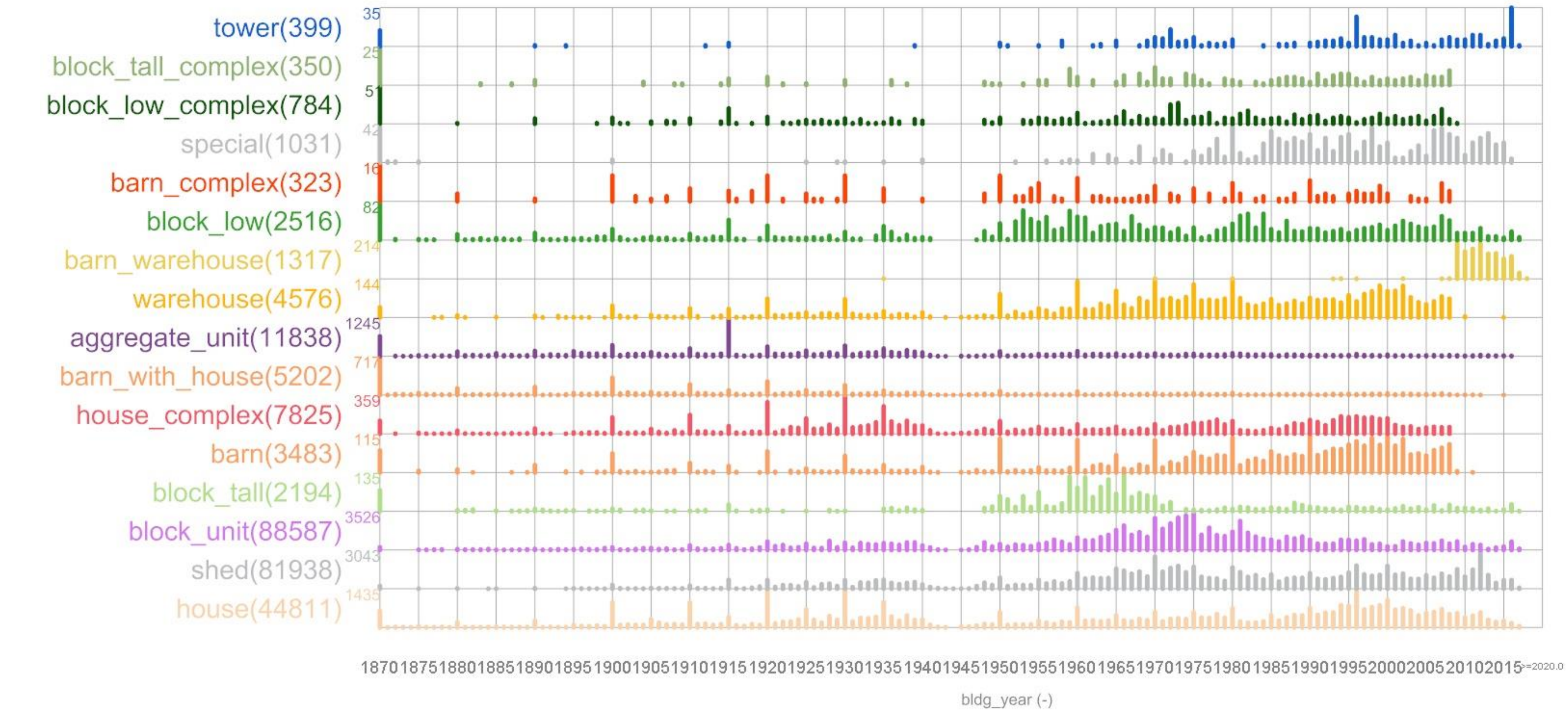


Figure 23 Histogram presenting the count of buildings per building year within each Structural Layout. The y-axis is renormalized in each Structural Layout for visualisation purposes

In the EDB V5, 16 977 buildings have either been visually checked or used a source dataset that includes Structural Layout indications (e.g. drawings from TBDB [13] and visual inspections

[14]). From this group 86% of inspected buildings matched the Geometric Layout classification, while 79% of these buildings matched the Structural Layout assigned trough the automated classification algorithms herein discussed and summarised graphically in Figure 24. While the misclassification from UHO *(House)* to UHC *(House Complex)* would have no impact on the EDB V5 process, the low matching ratio shown for UBA *(Aggregate Units)* would need further study to truly capture the subtle difference from other forms of building's adjacency. The OSP *(Special)* Structural Layout could only be assigned through visual inspection and therefore have all been considering as non-matching compared to the automated classification. Given the accuracy of the available data [11] (where also erroneous classification from visual inspections has been observed), these matching percentages are considered satisfactory for the overall classification process.

As shown in Figure 25, different sources have also been found to have different matching accuracy. Buildings inspected through the TBDB [13] data collection have been always found in agreement with a UBH *(Block Unit)* classification, while the Dataland farmhouse classification ([11]) is in agreement with a WBH *(Barn with House)* classification in 73% of the buildings. Finally, Structural Layouts inspected by JBG [14] agree with EDB V5 classifications in 62% of the buildings. Further studies are therefore encouraged, in order to investigate in even more detail the types of mismatches and to assess the consistent assignment of Structural Layout definitions throughout the various inspection activities.

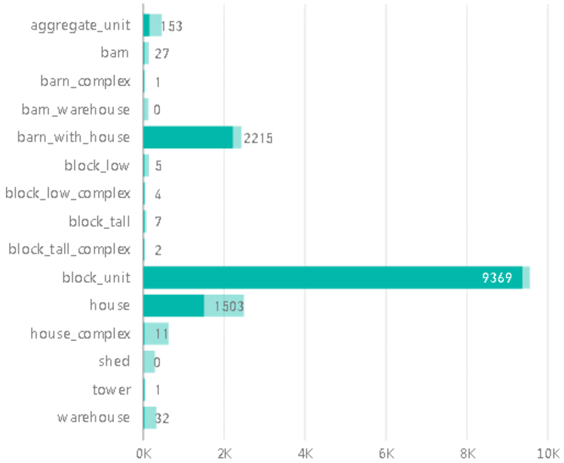


Figure 24 Number of correct assignments per Structural Layout. The length of the deep green in comparison to the light green demonstrates the matching ratio for the algorithm.

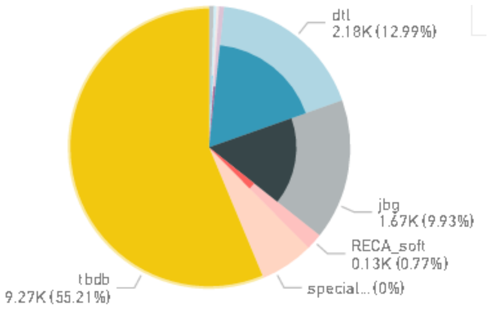


Figure 25 Number of matching assignments depending on the inspection source they are compared to (% indicates size of dataset compared to total amount of inspections).

More information on the inspection sources can be found in [11]. For inspection sources with low match ratio it could be studied further if the mismatches are due to misclassifications by the algorithm or by the inspector.

6 Structural System

The last part of the EDB V5 classification process combines the Structural Layout class per each individual building with its respective year of construction (as shown in Figure 23) and any available inspection data, so to estimate its most likely Structural Systems [1]. This conclusive process is shown in Figure 26, which also outlines the sub-processes described in this section.

The Structural System description is based on the GEM taxonomy system [5] and, as per the development of the classification process described in [1], the EDB V5 Structural System strings are reporting six GEM elements, selected in accordance with the EDB V3 Hazard & Risk typology descriptions [5].

Structural Systems can be assigned to a building in the following ways:

1. Inference based Structural System assignment:
    a. Structural Layout based inferences (Section 6.1): Structural Systems are assigned through judgement-based inferences based on the Structural Layout and building year with a function modifier. The original inferences [5] have been modified to suit the new classification system [14].
    b. Data-driven inferences for UBHS buildings: For buildings assigned a UBHS Structural Layout, data-driven inferences are applied. The data-driven inferences are based on available internal structural data from project information.

2. Inspection data Structural System assignment (Section 6.2)
    a. Full inspection data: Where inspection data is able to fully assign a Structural System string, it overwrites the inferred one.
    b. Partial inspection data: Where inspection data is only able to assign a partial Structural System string, the classification process gap-fills the missing information based on known full Structural System strings within the building stock.
3. Special geometry Structural System assignment (Section 6.2):
    a. Where structures have been identified as having a special geometry, a specific Structural System is assigned. This includes windmills, silos, masts, glass houses, water towers and parking garages.

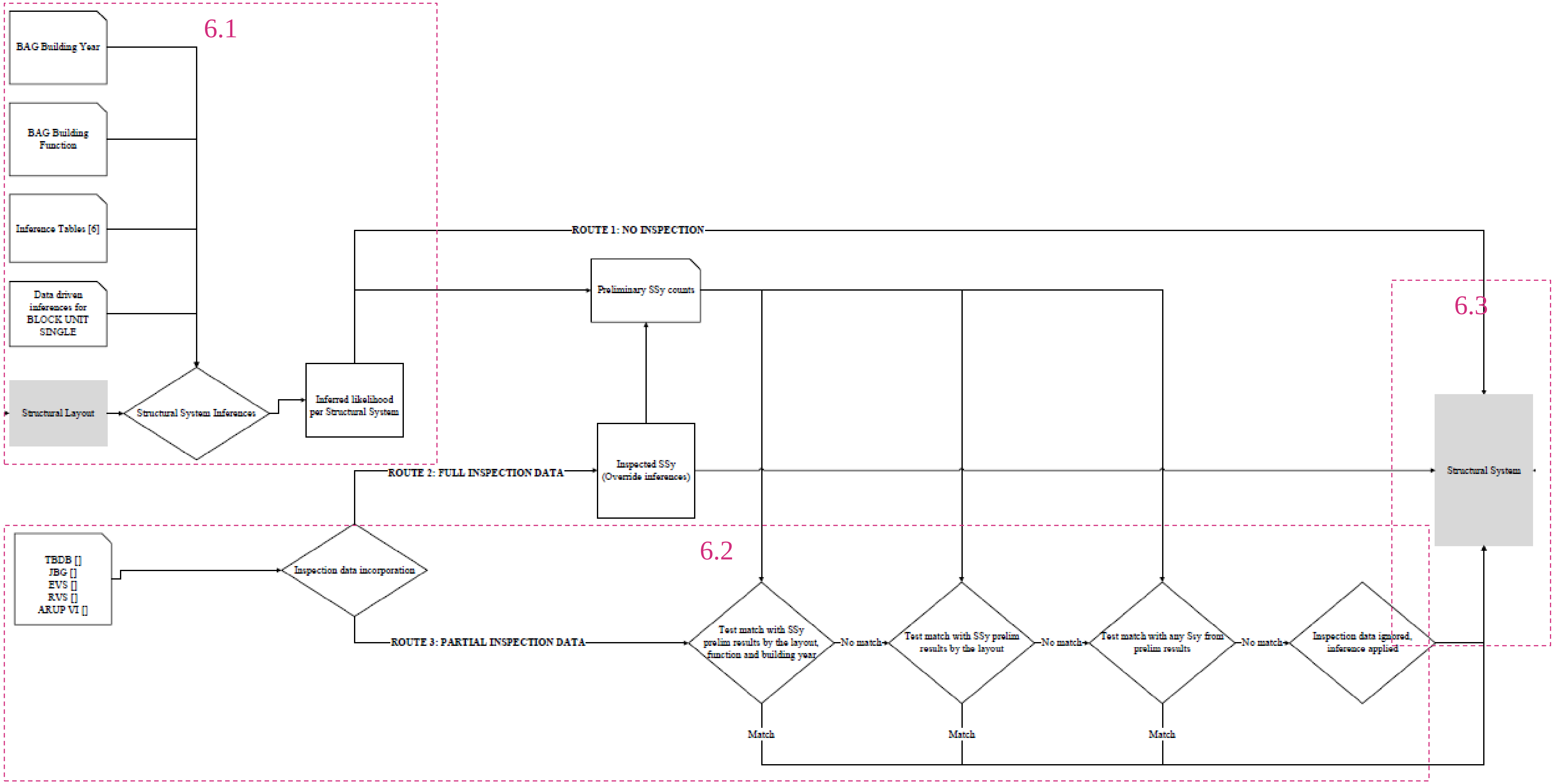


Figure 26 Structural System classification processses

To help with the readability of the next sections, the GEM taxonomy labels used in EDB V5 [1], are added below (Table 2).

Table 2 GEM labels used in EDB V5

Dx / Dy: Material of lateral load resisting system Position 1 and 3			
element_type	**element_subtype**	**element_code**	**element_id**
			00
Unknown material		MAT99	01
Other Material		MATO	02
Concrete		CR	10
Concrete	Precast concrete	CR+PC	11
Concrete	Cast in place concrete	CR+CIP	12
Masonry		MUR	20
Masonry	Clay Bricks	MUR+CLBRS	21
Masonry	***Calcium Silicate***	***MUR+CSBRS***	22
Masonry	Masonry unit / other	MUR+MO	23
Masonry	Concrete blocks / unknown type	MUR+CB99	24
Masonry	Stone, unknown technology	MUR+ST99	25
Wood		W	30
Steel		S	40
Dx / Dy: Type of lateral load resisting system Position 2 and 4			
			00
Unknown lateral load-resisting system		L99	01
Other		LO	02

No lateral load-resisting system		LN	03
Dual frame-wall system		LDUAL	10
Post and beam		LPB	20
Moment frame		LFM	30
Braced frame		LFBR	40
Hybrid lateral load-resisting system		LH	50
Walls		LWAL	60
Exterior wall material **Position 5**			
			00
Unknown material of exterior wall		EW99	01
Presence of outer leaf	Other	EWO	02
No outer leaf cavity walls		EWN	03
Presence of outer leaf		EW	10
Material of floor system **Position 6**			
			00
Unknown floor material		F99	01
Other		FO	02
No elevated floor material (single storey)		FN	03
Concrete		FC	10
Masonry		FM	20
Timber		FW	30
Steel		FME	40

6.1 Inferred Structural Systems

A set of age-based inference rules is applied on each building specific Structural Layout (and building function) for the total building stock [15], using the process shown in Figure 27. The visualisation at an urban scale is summarised in the example set of buildings in Figure 28 while the expected counts per Structural System in the region are reported in Figure 29.

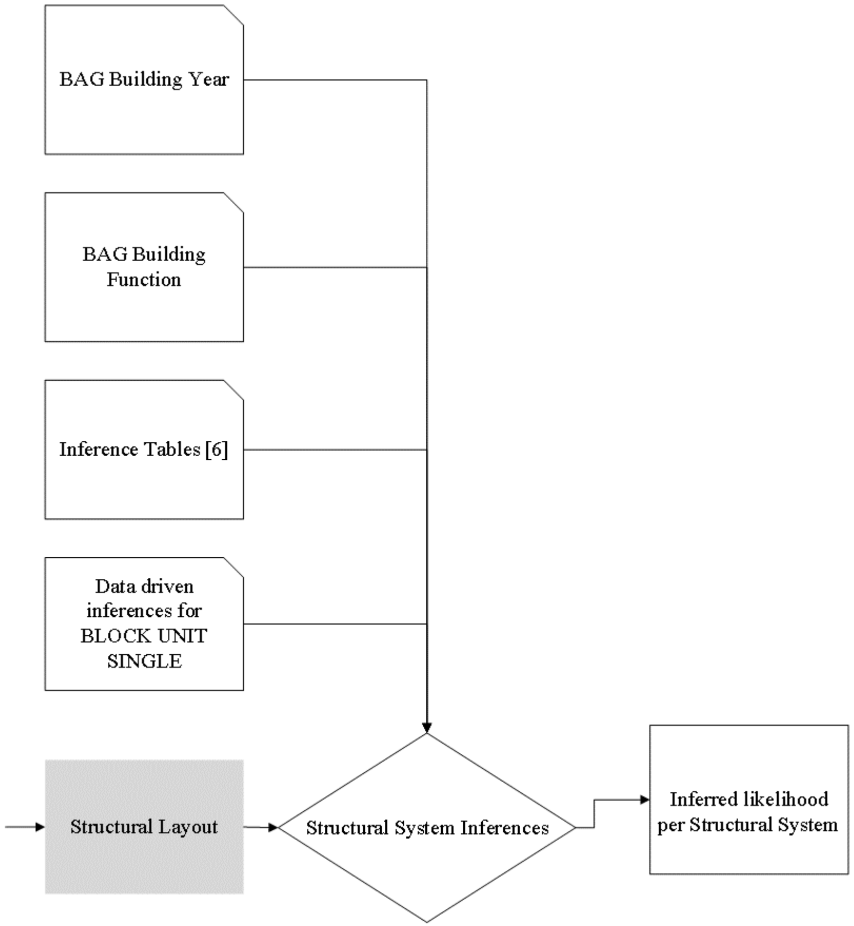


Figure 27 Inference–based Structural Layout assignment per building.

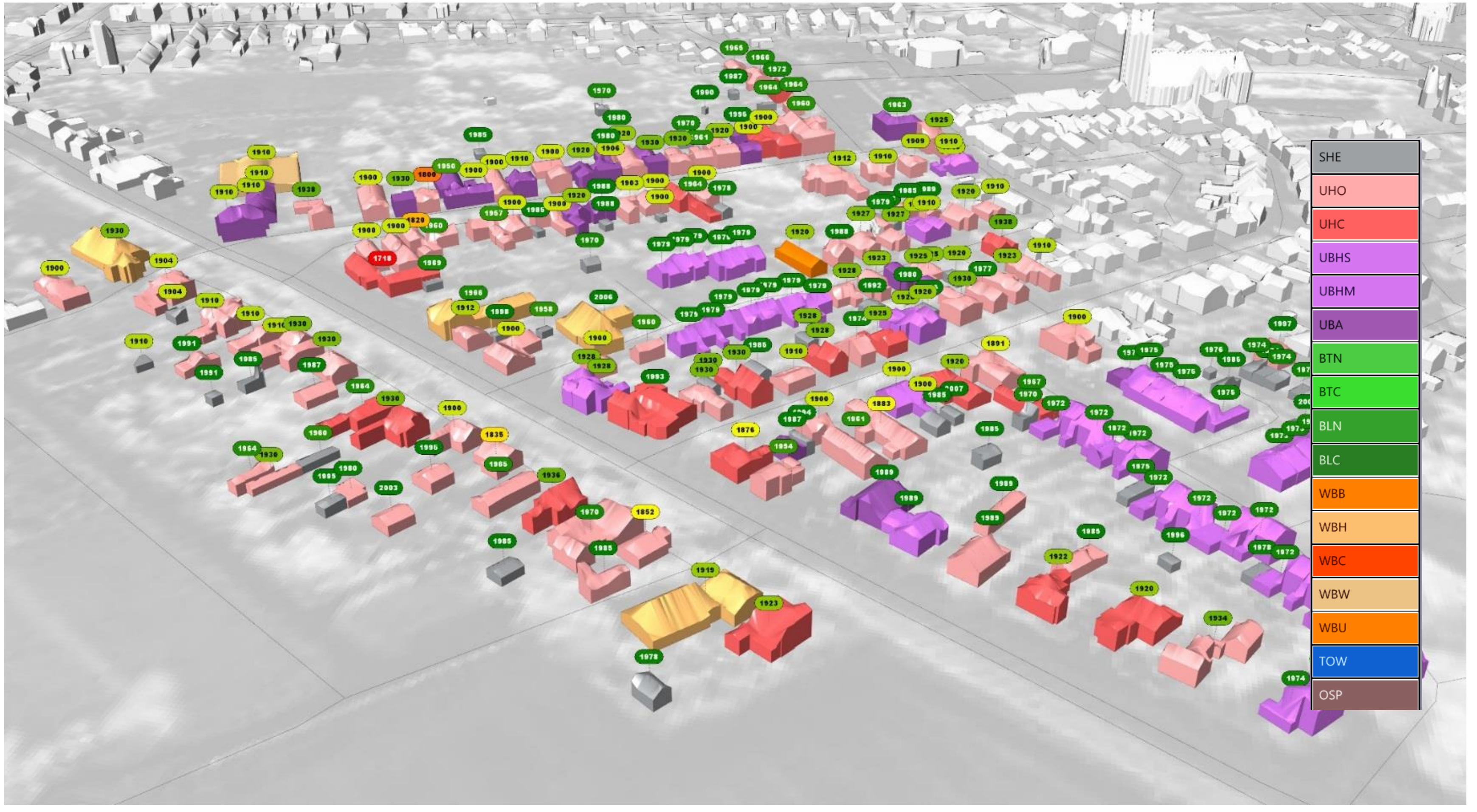


Figure 28 Structural Layout and building year data overlaid on a subset buildings (same as the ones shown in Figure 4)

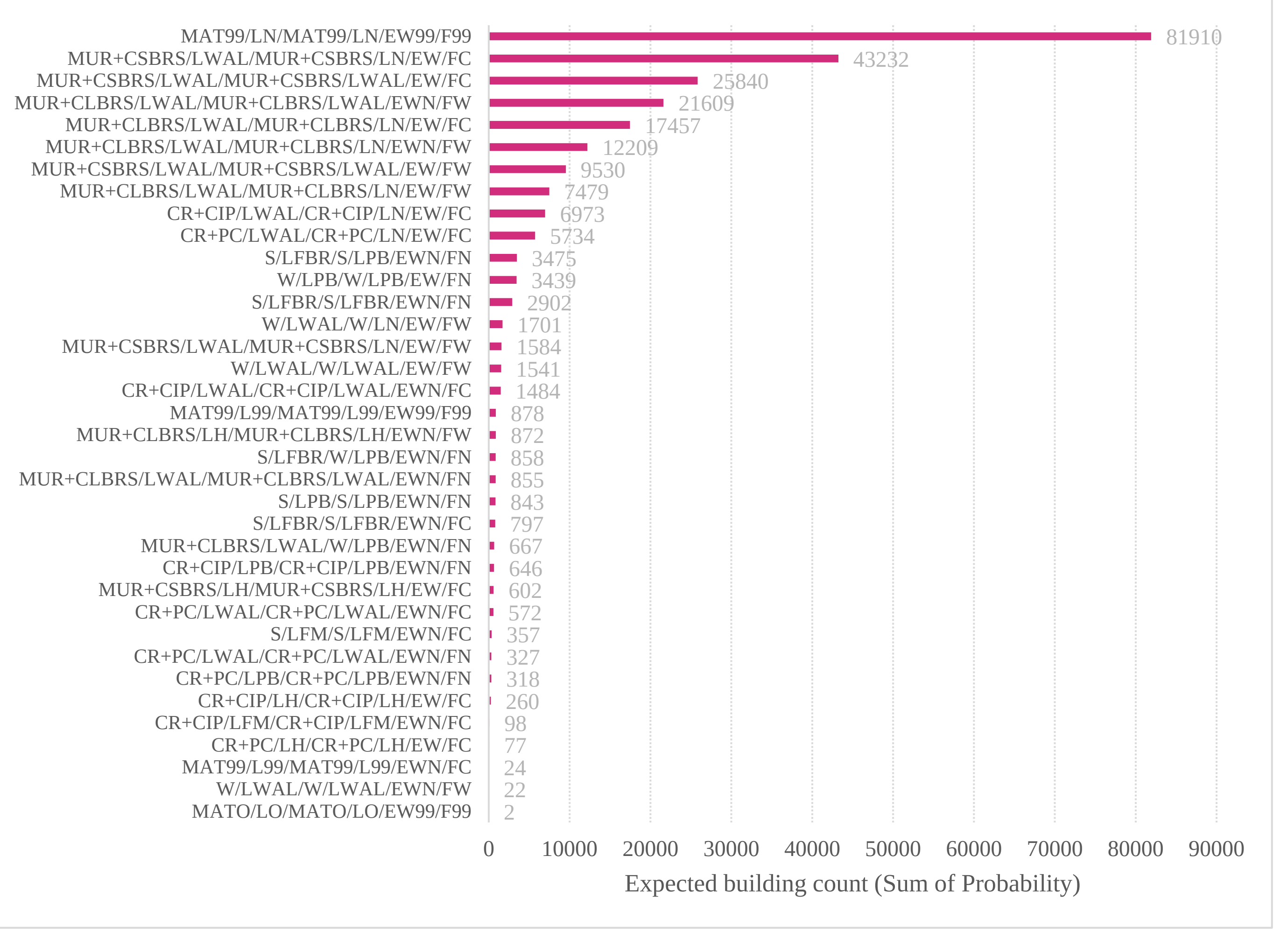


Figure 29 Expected building count for the EDB V5 Structural Systems before the incorporation of inspection data.

6.2 Inspection Data Incorporation

The inferred Structural Systems have been checked against all available inspection data. When required structural characteristics are available through inspections, the inferred Structural Systems are replaced by the inspected ones, as schematized in Figure 30.

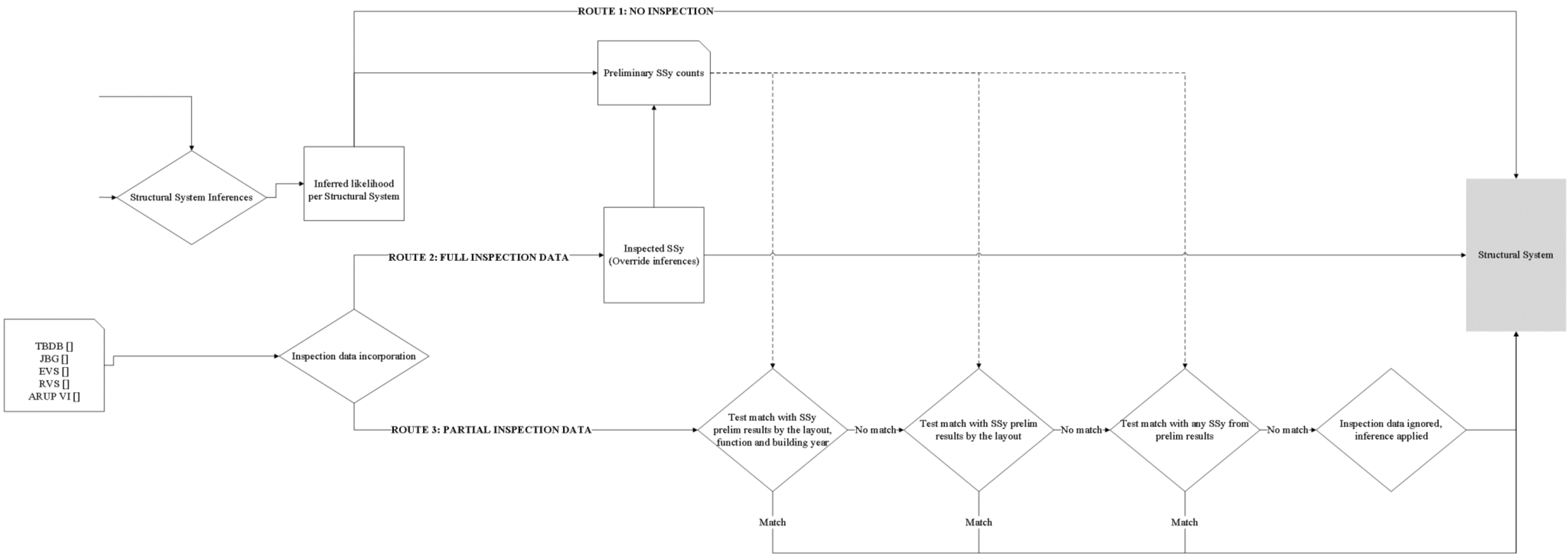
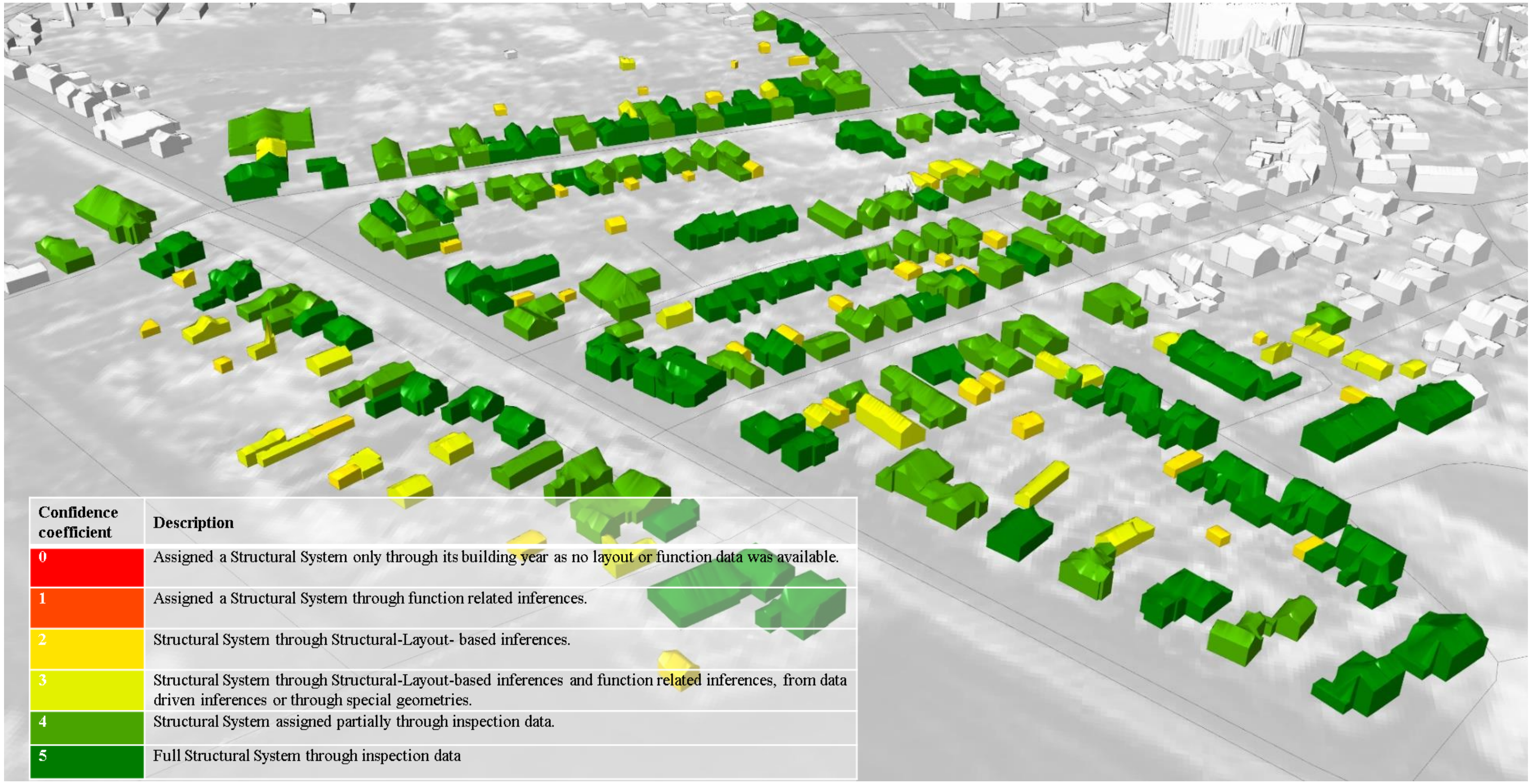


Figure 30 Inclusion of Inspection data flowchart.

A total of 26 847 buildings, more than 10% of the building stock of the study area (and more than 15% of the buildings expected to be populated), benefit from additional structural information gathered through inspections. As shown in Figure 31, inspection data is predominantly available within the 0.2g contour of the KNMI Hazard map of 2015 [16], while the coverage outside the 0.2g contour is currently very limited. In areas like the one shown in Figure 32, most buildings are color-coded green for having inspection data incorporated, in most cases even with all GEM Structural System fields completed.

Figure 31 Buildings with inspections within the 0.2g contour of the 2015 KNMI Hazard map [16]



Confidence coefficient	Description
0	Assigned a Structural System only through its building year as no layout or function data was available.
1	Assigned a Structural System through function related inferences.
2	Structural System through Structural-Layout- based inferences.
3	Structural System through Structural-Layout-based inferences and function related inferences, from data driven inferences or through special geometries.
4	Structural System assigned partially through inspection data.
5	Full Structural System through inspection data

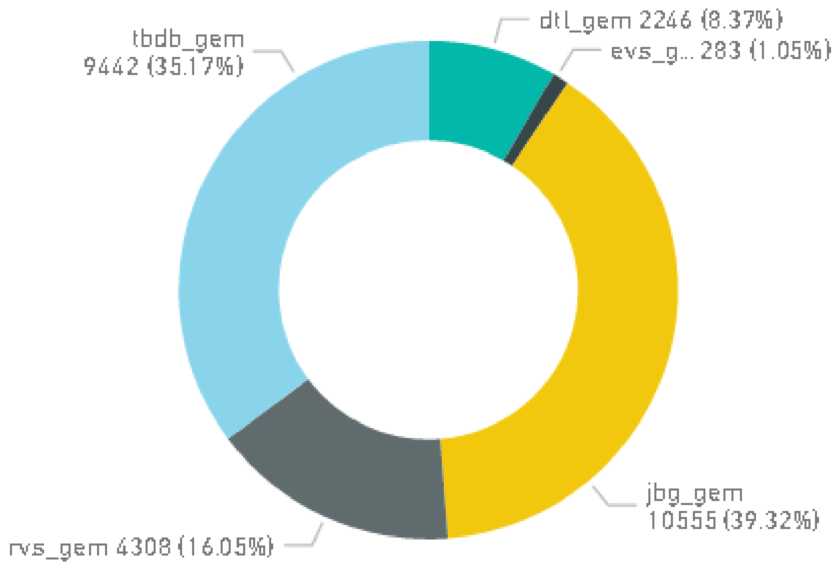
Figure 32 Set of buildings in the region color-coded based on the confidence coefficient (same buildings as the ones shown in Figure 4)

As shown in Figure 33, inspections by the engineering firm JBG [14] are the most predominant source of inspection data (approximately 40% of the inspected buildings in EDB V5 [1]). It should be noted that part of the inspection data does not cover all six data fields that comprise a Structural System in EDB V5 or the structural characteristics required for the Structural Layout classification. The main reason being that the majority of the inspection data has been gathered through inspections taking place either from the street or through desk studies based on streetview images. The count of full GEM strings and partial GEM strings retrieved from each inspection source is noted in Figure 33, while the full GEM strings and partial GEM strings per inspection source are shown in Figure 35 and Figure 36 respectively.

All inspection labels found in more than 50 buildings are also ranked in Figure 34. The empty fields of partial inspections are left empty in this Figure, leading to a corresponding GEM label with consecutive slashes (“//”). The EDB V5 algorithms then auto-fills the empty fields using conditional probabilities of the previously inferred Structural Systems (described in Section 6.1).

The combination of different inspection sources, in an example area of the EDB with comparatively larger inspection coverage, is shown in Figure 37.

Finally, Figure 38 shows that the large majority of buildings within the 0.2g contour -if SHE *(Shed)* buildings are excluded- have already been inspected through at least one of the inspection data sources. The Structural Layout of the remaining uninspected buildings is also shown in the same graph.



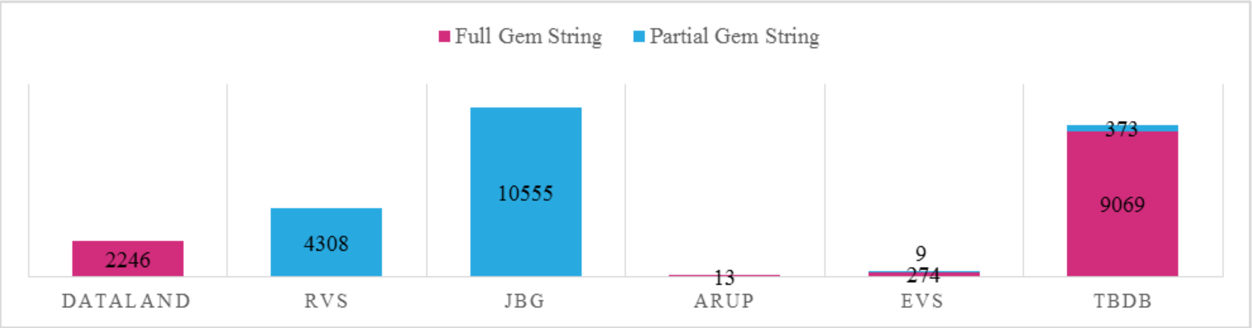


Figure 33 Count of buildings per inspection source incorporated in EDB V5 with breakdowun of full and partial GEM strings per inspection source.

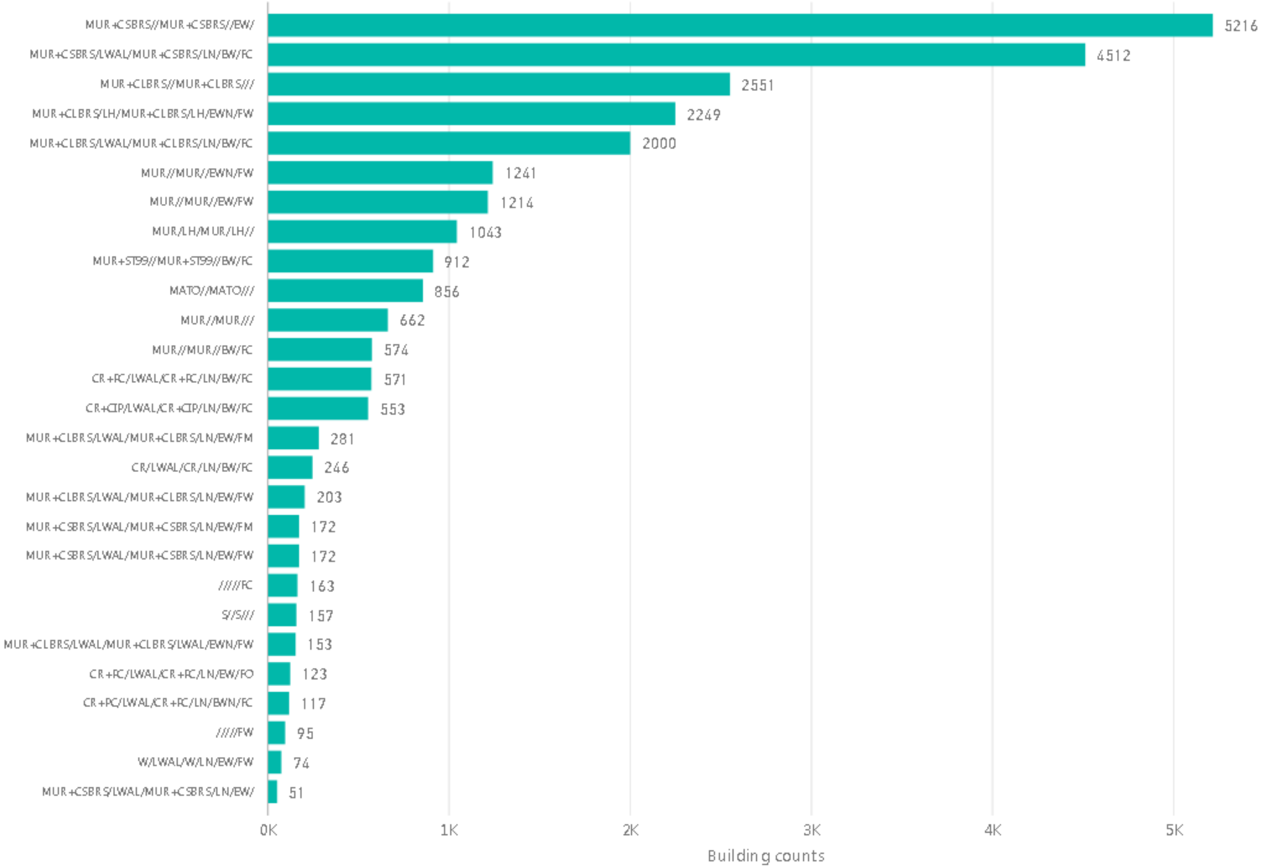


Figure 34 Most recurring GEM label strings retrieved from inspection data[11].

[11] For example the most recurring inspection label has been MUR+CSBRS//MUR+CSBRS//EW//, where structural system and floor material are left empty (due to the inspection taking place using street view images)

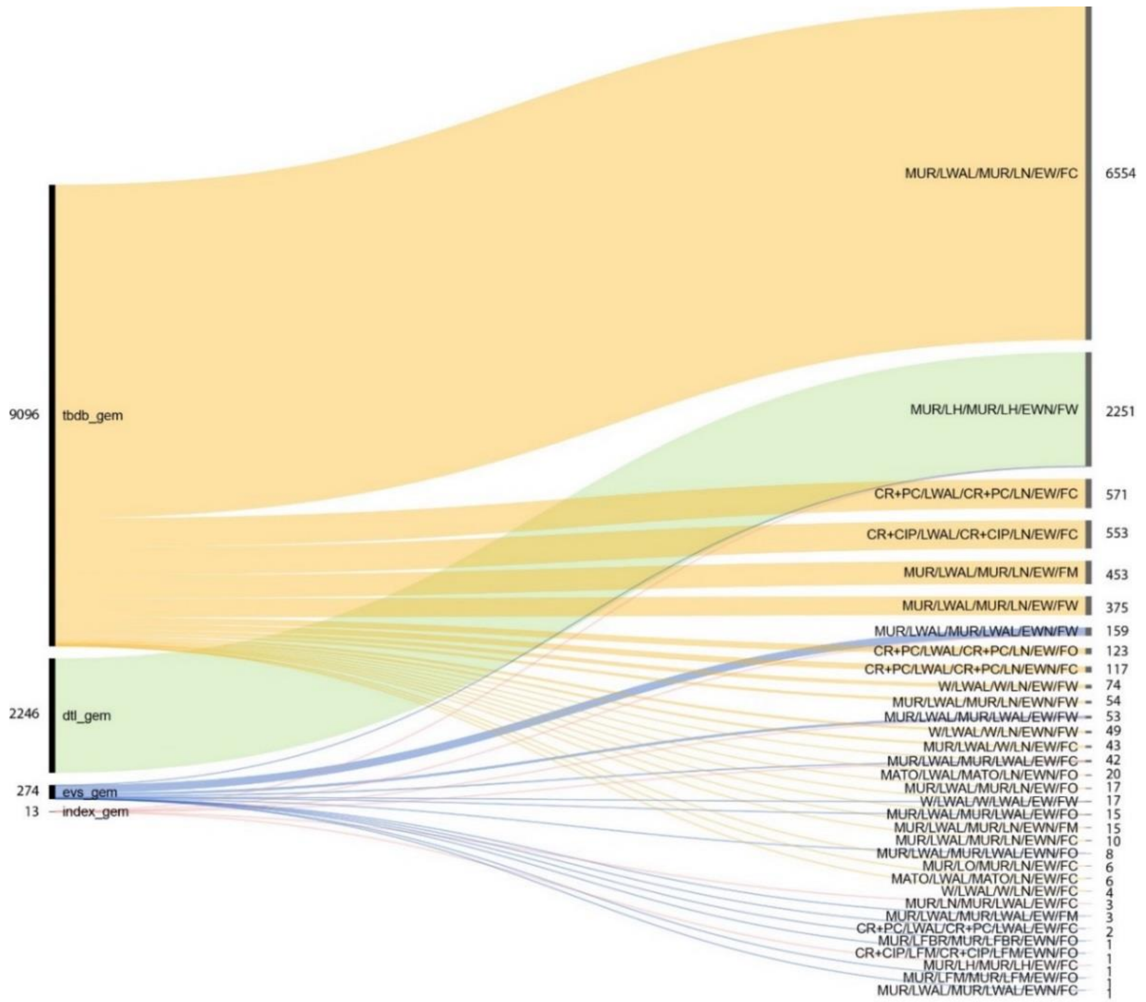


Figure 35 Full GEM string building count per inspection source.

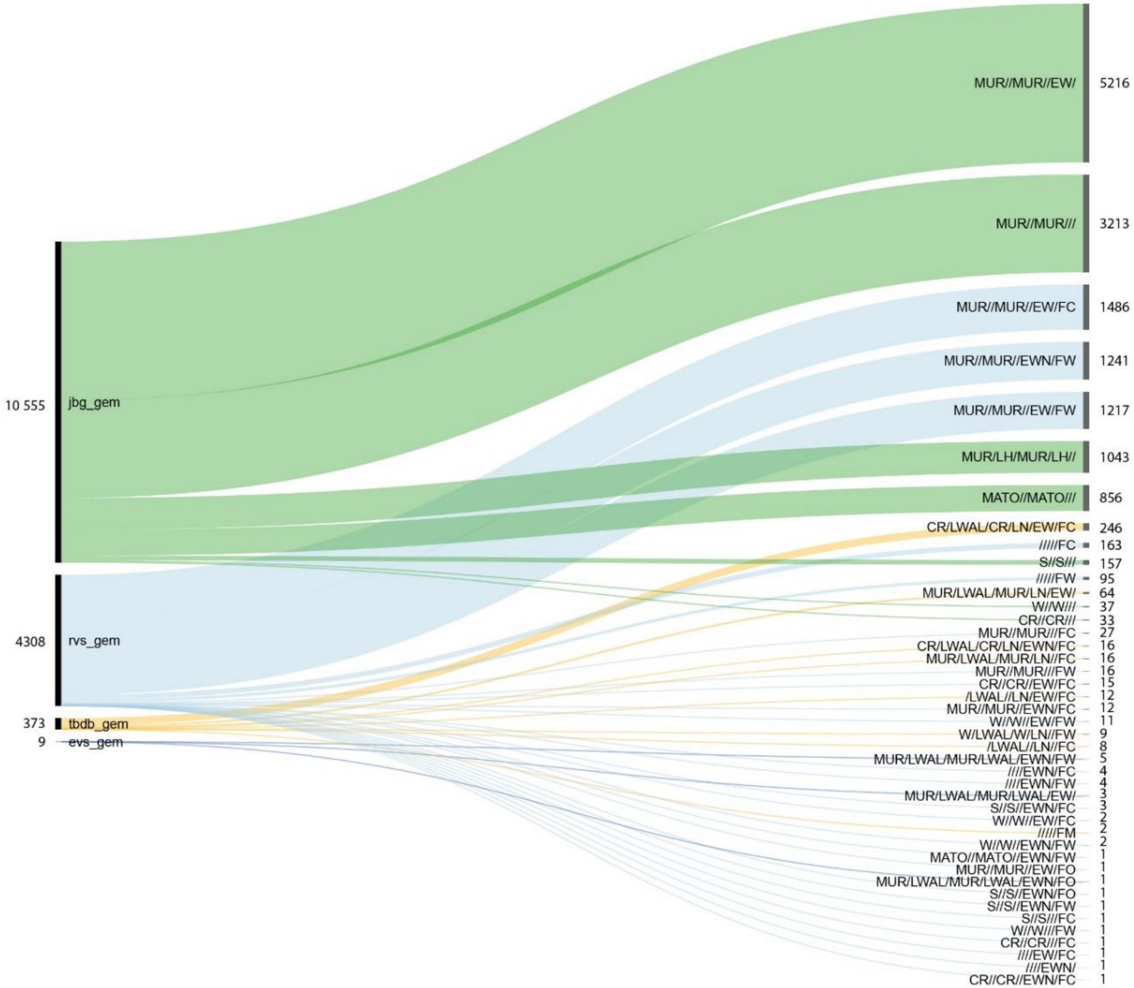
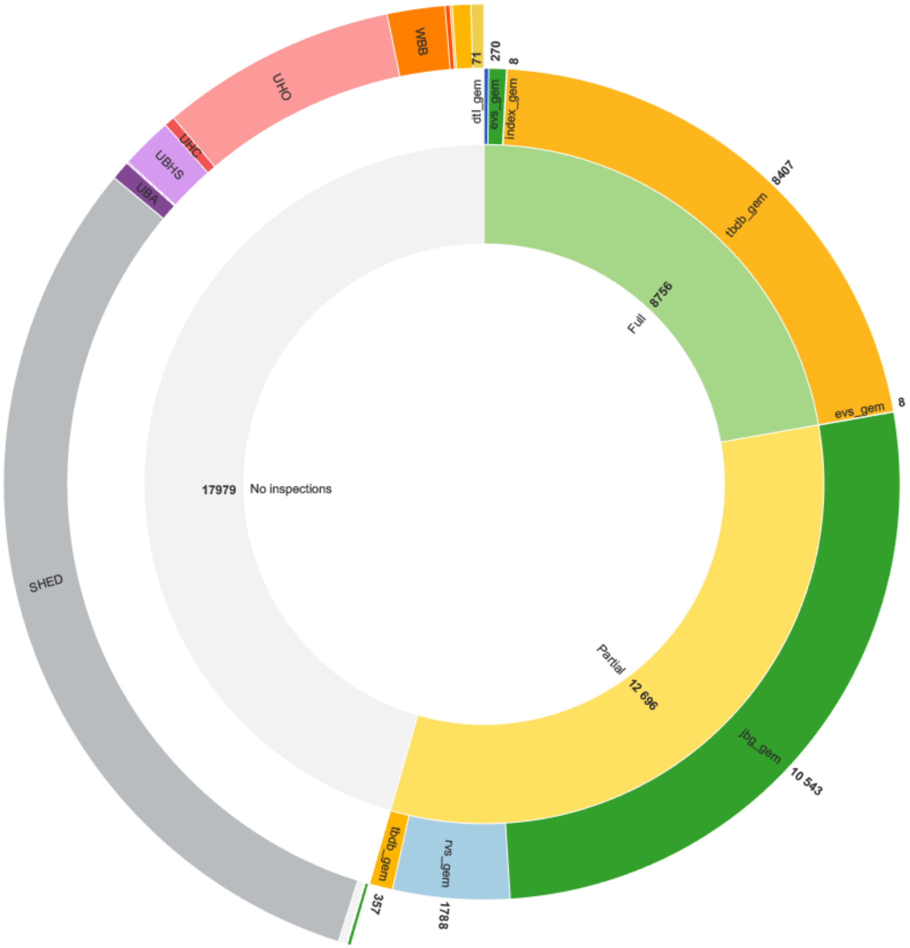


Figure 36 Partial GEM string building count per inspection source.

Figure 37 Combination of different inspection sources.



Data source	Count	%
Drawing data (tbdb)	8764	22%
EVS inspection	278	1%
Desk study, Arup (index buildings)	8	0%
Desk study, JBG	10555	27%
RVS inspection	1790	5%
Farmhouses, Dataland address description	71	0%
No inspections	17965	46%
	39431 buildings	**100%**

Figure 38 Coverage of inspection data in the buildings within the 0.2g contour of the 2015 KNMI Hazard Map [16]

6.3 Resulting Structural Systems Classification

Excluding the Structural Systems resulting from the SHE Layout (MAT99 / LN / MAT99 / LN / EW99 / F99), the final count of buildings assigned to each Structural System in EDB V5 [1], summarised in Figure 39, highlight *Calcium Silicate Masonry (MUR+CSBRS) Wall System (LWAL)* in one direction and *No Lateral Load Resisting System (LN)* on the other direction with an *Outer Leaf (EW)* and a *Concrete Floor (FC)* as the most predominant Structural System. More than 40 000 buildings are expected to have this combination of structural characteristics in the scope area.

As visible when comparing the resulting combinations in Figure 39 with the inferred Structural Systems prior to the use of inspection data in Figure 29, the count of combinations of lateral resisting systems, structural materials and floor types increases from 36 to more than double that number through the incorporation of inspection data. On the other hand the counts of the most predominant types are minimally affected, as shown by the comparison of the blue and the red bars in this graph, corresponding to the expected counts before and after the inspection data incorporation respectively. The visualisation, at urban scale, of the final assignment of the process presented is shown in Figure 40.

In Figure 39, the count of buildings per Structural Material cannot be easily calculated, as multiple characteristic variations are present. For this reason the Structural Systems per material on the DX direction have been aggregated in Figure 41. The bar charts have been color-coded per age bracket to demonstrate also the influence of construction year to the expected Structural Material.

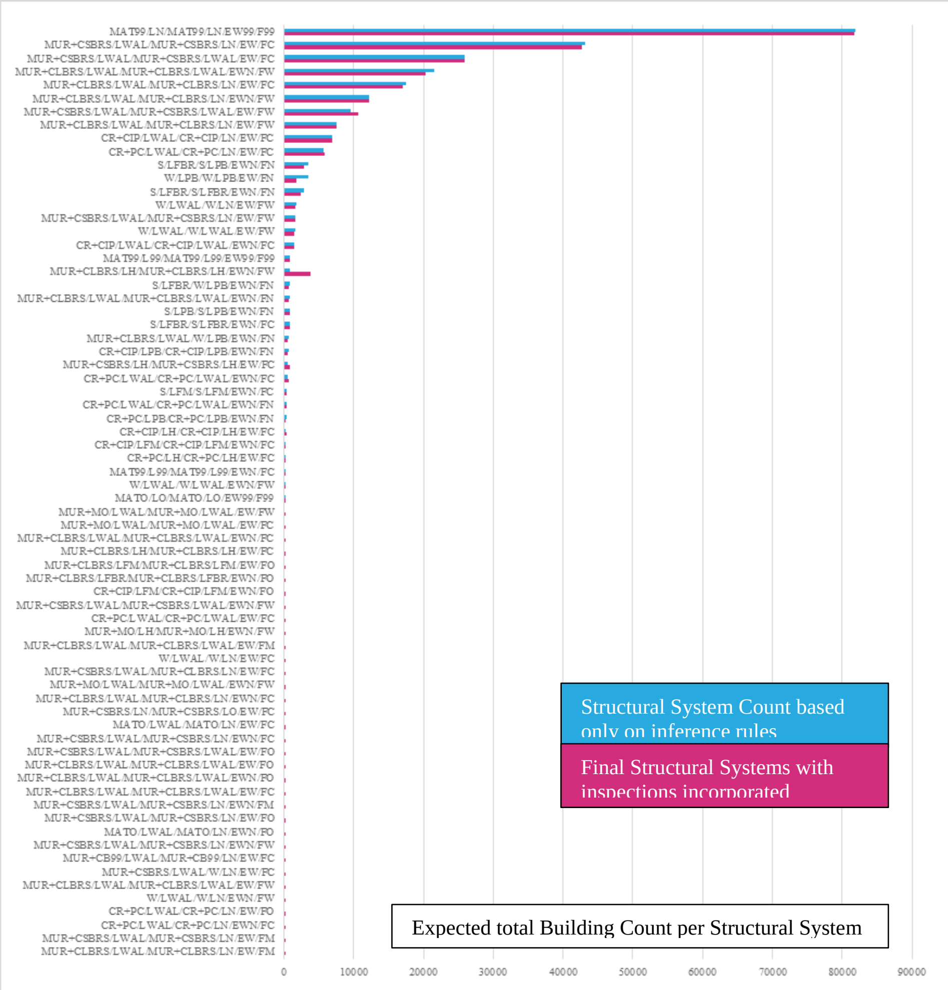


Figure 39 Conclusive count of buildings per Structural System in EDB V5 [1] – red lines, compared to pre-inspection building count – blue lines.

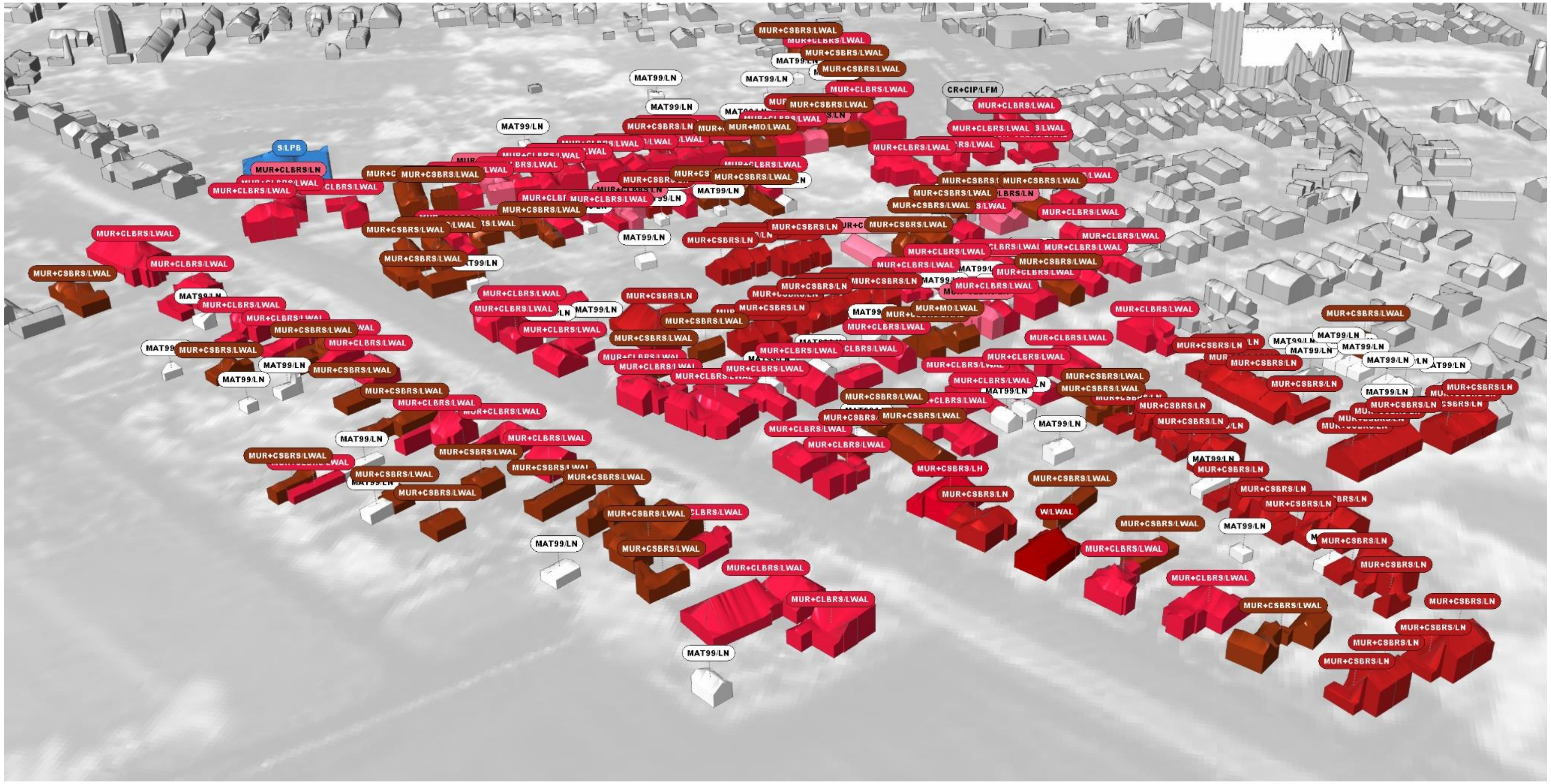
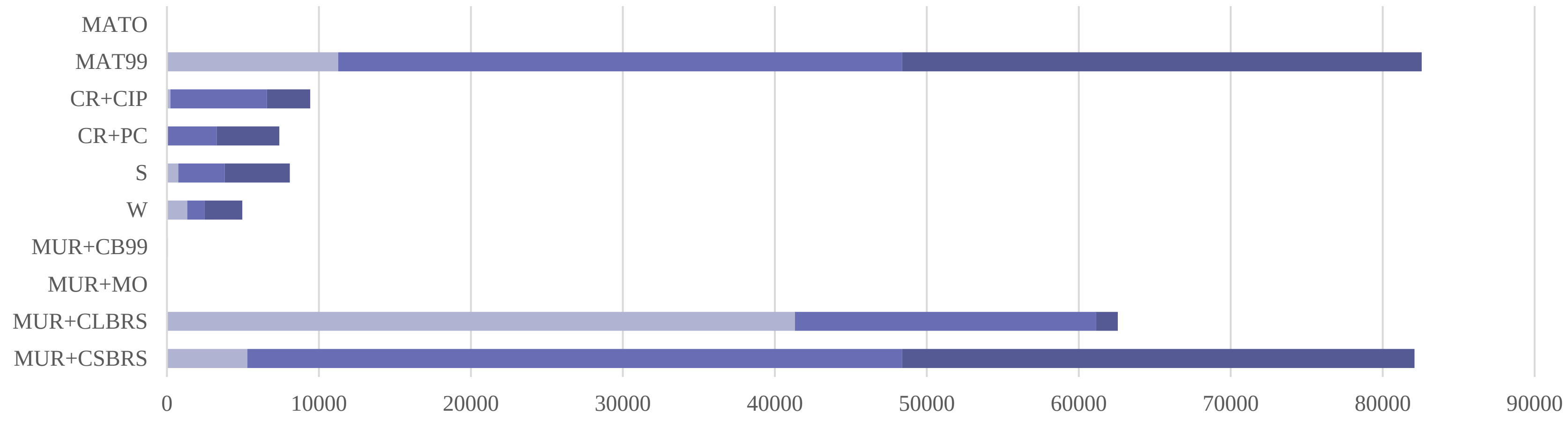


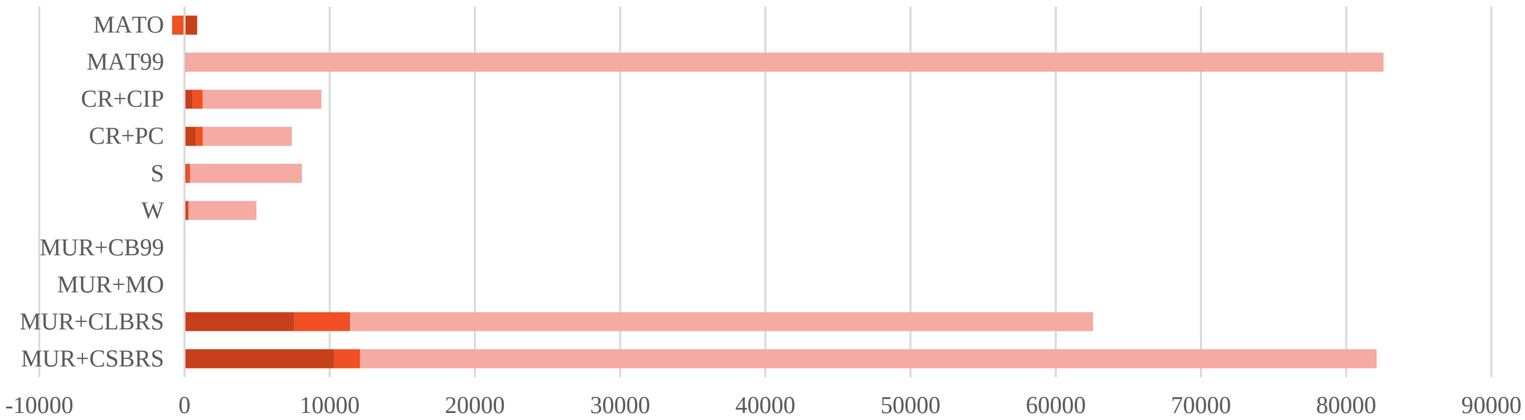
Figure 40 Final Structural System per building for a subset of buildings (same as the ones shown in Figure 4) – only material and structural system in the DY direction of the most likely Structural System are shown for visualisation purposes.



	MUR+CSBRS	MUR+CLBRS	MUR+MO	MUR+CB99	W	S	CR+PC	CR+CIP	MAT99	MATO
building_year<1945	5292	41325	5	0	1340	749	57	234	11270	3
1945<=building_year<=1985	43088	19802	2	39	1139	3041	3224	6339	37101	20
building_year>1985	33717	1446	1	0	2478	4295	4111	2853	34199	6

Figure 41 Expected total count of buildings per material organised in three building year domains

In Figure 42 the total count is also color-coded differently if the structural material has been inspected or inferred. As shown there, the count of buildings without additional information collected through inspections is still large, although the trend in coverage changes according to the investigation area considered (within or outside the zone presented in Figure 31).



	MUR+CSBRS	MUR+CLBRS	MUR+MO	MUR+CB99	W	S	CR+PC	CR+CIP	MAT99	MATO
Inspected (Full Inspections)	10300	7568	8	39	206	163	813	554	0	883
Inspected (Partial Inspections)	1797	3835	0	0	92	221	446	689	90	-856
Inferred	70000	51170	0	0	4659	7701	6133	8183	82479	2

Figure 42 Expected total count of buildings per material organised and ratios visualisation of information source.

6.4 Structural System Validation

Several automated validation checks have been devised throughout the process flow to ensure the correct performance of the algorithm. In 94% of the inspected buildings the combination of Structural Material and the most likely Structural System final string has been accurately assessed.

In order to validate that the inferences provide results consistent to the received inspections it has been checked in how many cases the inspected Structural System -for complete inspection sources- matched the most likely Structural System before inspections are incorporated. This results in a correct match for 57% of the buildings. However, using the same validation set, this percentage increase to 90% when the inspected Structural System is checked against any of the 10 most likely Structural Systems (again before inspections are incorporated).

The matching test above corresponds to a correct prediction of all six GEM labels of a Structural System. If the match is limited to only the Structural Material and the Structural System in the main direction the initial match rate increases to 65%. If the prediction of only the Structural System in the main direction is tested then the initial success rate is 80%. The matching percentages across different Structural Layouts for the three levels described are shown in Table 3. The Structural Layouts with relatively low percentages are also the ones with little available inspections [11]

Table 3 Counts of buildings with full inspection data available per Structural Layout and corresponding match to the most likely Structural System.

	# full inspected buildings	% full inspected buildings matching top1 data-driven classification grouped	% full inspected buildings matching top1 data-driven classification grouped on gem part 3&4	% full inspected buildings matching top1 data-driven classification grouped on gem part 4
WBH	2260[12]	0%	0%	0%
BTN	2	0%	50%	50%
UHO	167	65%	95%	99%
WBB	1	0%	0%	0%
UHC	41	71%	100%	100%
UBA	31	6%	16%	16%
BLN	2	0%	0%	50%
UBHM	57	93%	95%	98%
UBHS	9304	70%	80%	99%
BLC	2	50%	100%	100%
WBW	3	0%	0%	0%

[12] The inference for WBH *(Barn with House)* doesn't yet assign the hybrid Structural System only, leading to the 0% match. This is mitigated by using Dataland data to assign the final Structural System.

A similar validation check has been also repeated on buildings of the UBH *(Block Unit)* Structural Layout in order to test the improvement between the newly developed data-driven inferences of EDB V5 [14] in comparison to judgement-based inferences derived with the default Layout process. The data-driven inferences applied have not been designed to completely match the inspection data, in order to avoid generation of artificial spread of attributes in the building stock (as a consequence of the variability presented in Figure 35). The data-driven inferred Structural Systems are resulting in a correct assignment for 86% of the buildings, when Inspection data from the TBDB [13] is available as reference. With the original expert-judgement inferences, used in previous version of the EDB, the assigned Structural Systems were correctly assigned in only 63% of the cases, when compared to the inspected Structural System.

Finally it has been verified that the data-driven inference methodology used for the UBHS *(Block Unit Simple)* Structural Layout captured adequately the frequency of the most predominant Structural Systems, through a comparison of the total count of buildings per Structural System based on inspections and based on data-driven inferences. This comparison returns very similar results, as presented in Figure 43.

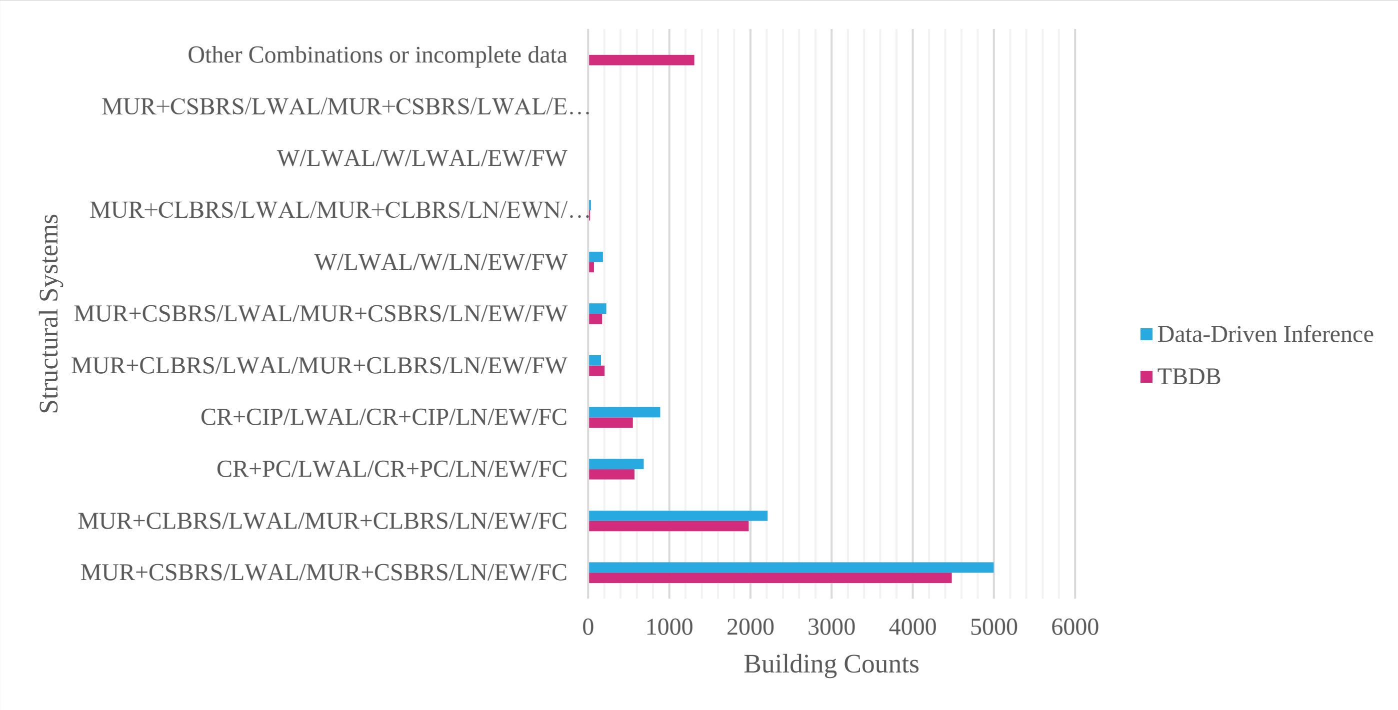


Figure 43 Comparison of total building Counts per Structural System through TBDB inspection and through applying the data-driven inferences at the same subset of buildings.

7 Conclusions and Recommendations

This report presented the data process sequence that took place in the EDB V5 calculation algorithm. It also visualised intermediate results from each analysis phase to better demonstrate the value of each step in the total calculation scheme.

It describes going from the initial data mining and source data retrieval to the final assignment of Structural Systems per building. The way in which the EDB V5 calculation engine combines inspection data, for buildings where it is available and the use of probabilistically inferred Structural Systems, for buildings without complete inspection data.

The inspection data output from the various inspection sources used in EDB V5 has been analysed, which can inform the next steps in inspection data gathering activities.

Additionally, initial validation activities for the EDB V5 calculation engine have been provided per calculation stage, proving an adequate effectiveness of the devised methodologies in providing quality data input for NAM's Hazard and Risk model for the Groningen earthquake region.

Despite the fact that this model is calibrated to perform on a regional scale, it has been shown how several activities contributing to the EDB development largely improve the building-to-building knowledge in the region, especially as far as geometric characterisation and high-level structural layout classification is concerned.

The further integration of the Exposure Database developments with NAM's Hazard & Risk calculation engine and results can facilitate the definition of most impactful further developments.

Based on current findings, an important recommendation is the improvement of the inspection data incorporation process, which can be predominantly facilitated by a definition alignment between inspection parties. A consistent and real-time data gathering process executed during the inspection process would directly improve the accuracy of the Exposure Database, not only by correct assignment of individual buildings, but also facilitating the development of data driven inferences for other typologies.

8 References

[1] **Arup.** *Exposure Database V5_Tech Note*. September 2017. Arup report number: 229746_052.0_AUX2008

[2] **Arup.** *Seismic Risk Study - Earthquake Scenario-Based Risk Assessment.* Arup report REP/229746/SR001

[3] **Arup.** *White Paper Exposure database v1*. May 2015. Arup report number: 229746_031.2_REP1001

[4] **Arup.** *Exposure Database V2*. September 2015. Arup report number: 229746_031.0_REP1003

[5] **Arup**. *Exposure Database V3*. March 2016. Arup report number: 229746_031.0_REP1010

[6] **Crowley H. and Pinho R**. *Report on the v5 Fragility and Consequence Models for the Groningen Gas Field.* October 2017.

[7] **GEM Foundation.** *About GEM. Global Quake Model*. [Online] 2017. https://www.globalquakemodel.org

[8] **Arup**. *EDB V5 Inspection Mapping Documentation*. November 2017. Arup report number: 229746_052.0_MEM2004

[9] **AHN.** *Actueel Hoogtebestand Nederland.* AHN, 2009.

[10] **Kadaster.** *Landelijke Voorziening Basisregistraties Addressen en Gebouwen.* s.l.: Kadaster, 2017.

[11] **Arup**. *EDB V5 Data Documentation*. December 2017. Arup report number: 229746_052.0_REP2014

[12] **Arup.** *Comparison of EDB V5 height data to inspection-based storey count*. September 2017. Arup report number: 229746_031.0_NOT2037

[13] **Arup.** *Definities Building Datalijst.* March 2017. Arup report number: 229746_047.0_REP2035

[14] **Arup.** *Visual Inspection Manual.* July 2017. Arup report number: 229746_052.0_NOT2016

[15] **Arup**. *EDB V5 Structural System Inference Documentation.* October 2017. Arup report number: 229746_031.0_NOT2039

[16] **Dost, B., Spetzler. B.,** *Probabilistic Seismic Hazard Analysis for Induced Earthquakes in Groningen – Update 2015;* October 2015.

Annex A

Count of Structural Systems and Example Buildings per Structural Layout

In this section example buildings and the most recurrent Structural Systems per Structural Layout will be provided. To begin with, the example buildings are shown all together in Figure 44 where the total count per Structural Layout in EDB V5 is also reminded. In this image the difference in scale between Structural Layouts becomes obvious. Due to the fact that it is not easy to discern the exact geometries, this section will study the groups in more detail, showing the geometries also from a closer point per Structural Layout group.

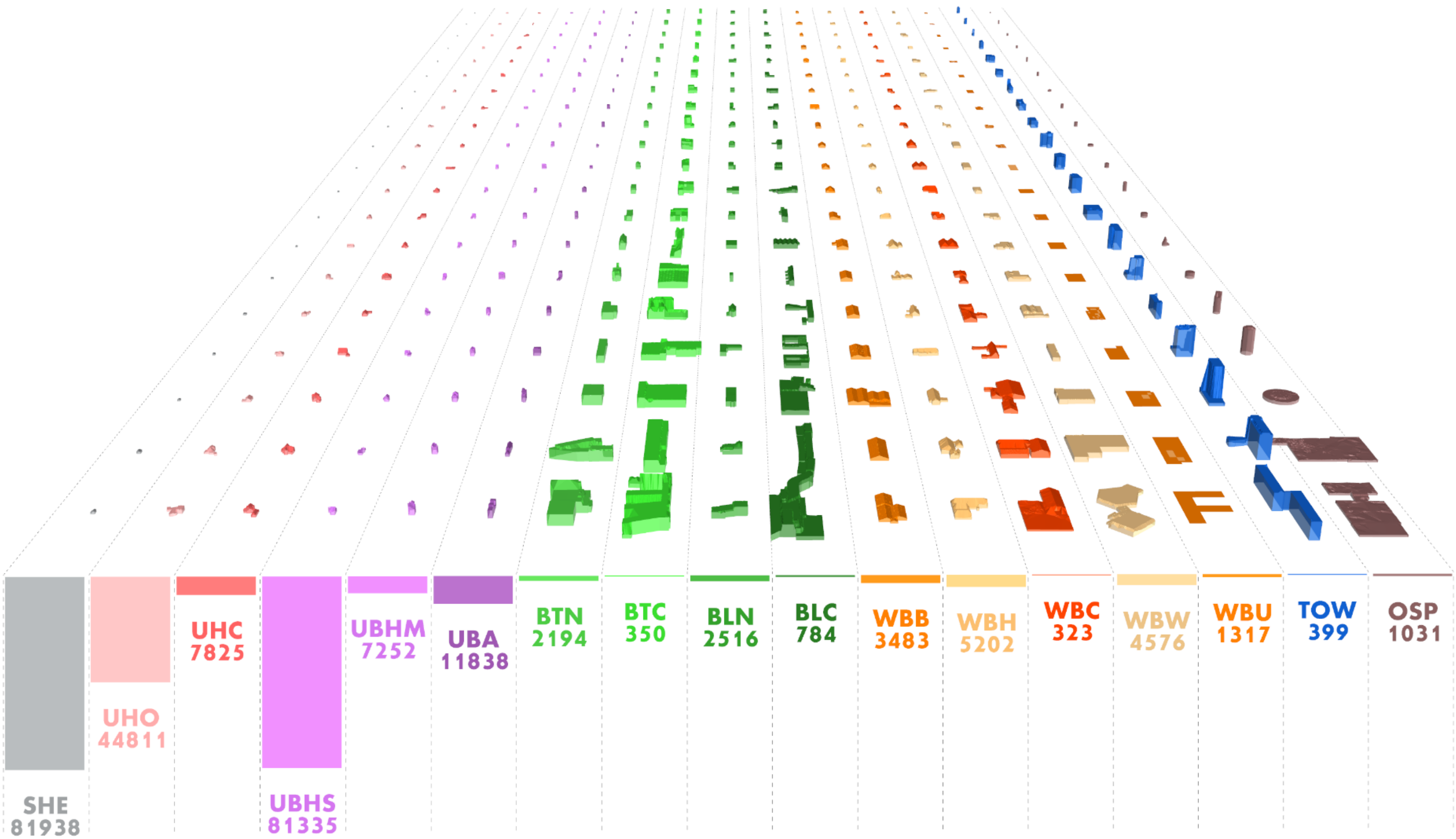


Figure 44 Expected building count and example building geometries classified in each of the EDB V5 Structural Layouts

Figure 45 shows example point cloud-generated geometries for the relatively smaller in size Structural Layouts: SHE (“Shed”), UHO (“House”), UHC (“House Complex”), UBHS (“Block Unit Single”), UBHM (“Block Unit Multiple”) and UBA (“Aggregated Unit”). Subsequently the most predominant Structural Systems for the “House” Structural Layouts (UHO and UHC) are provided in Figure 46.

The same combination of characteristics is also the most likely among “Block Unit” Structural Layouts (UBHM, UBHS and UBA), with the only difference being that in this Layout *no lateral load-resisting system (LN)* is inferred to be present in the DY direction.

Building of the SHE (“Shed”) Structural Layout are almost entirely represented by the MAT99/LN/MAT99/LN/EW99/F99 on which they are inferred given that they are currently considered of relatively lower importance due to their very small size and related population present.

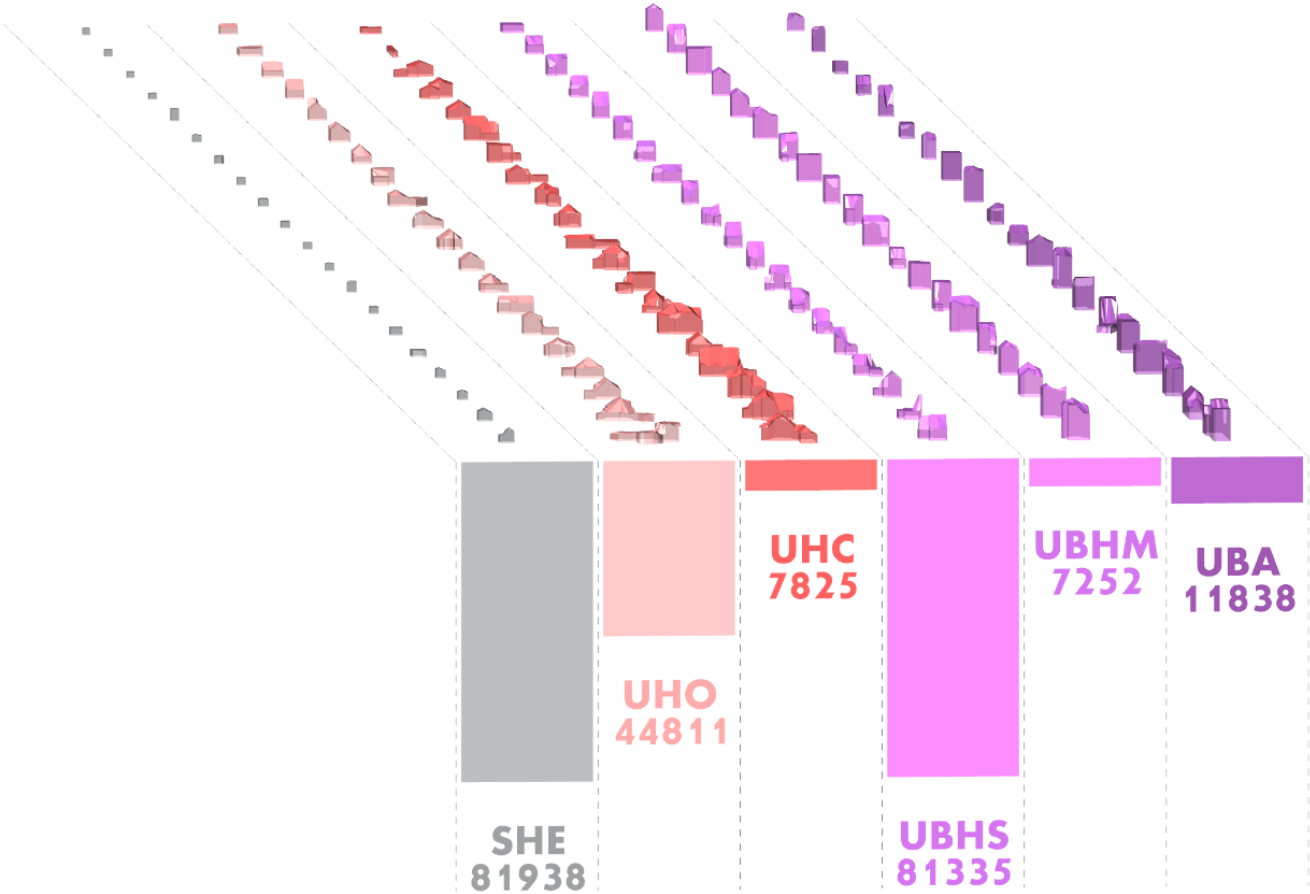


Figure 45 Example buildings and building count for the SHE, UHO, UHC, UBHS, UBHM and UBA Structural Layouts

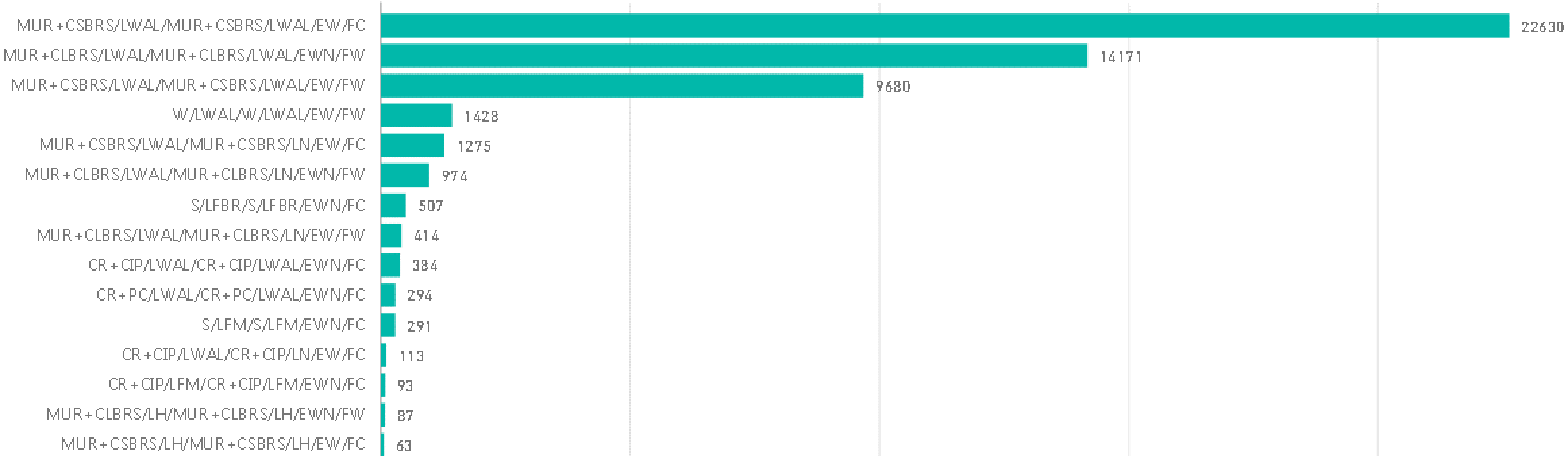


Figure 46 Expected building count for Structural Systems for UHO and UHC Structural Layout. Only Structural Systems expected to be present in more than 50 UHO and UHC buildings are shown in this graph.

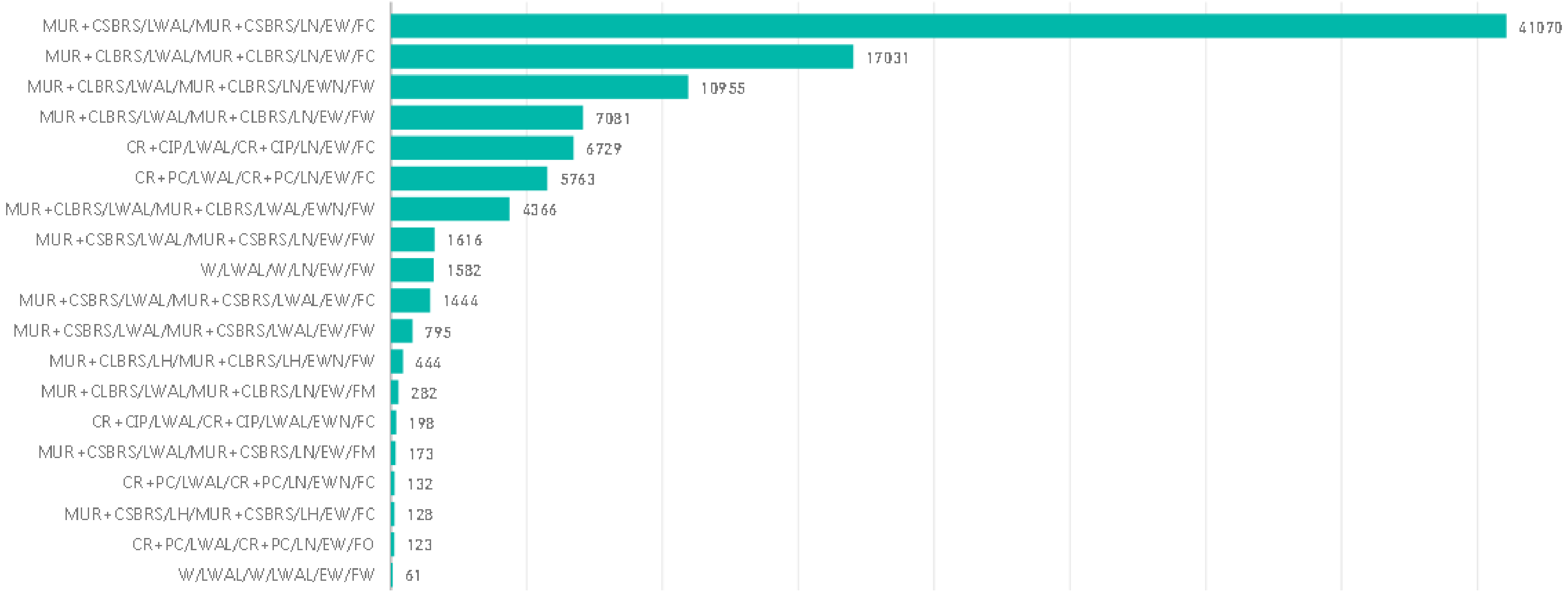


Figure 47 Expected building count for Structural Systems for UBHM, UBHS and UBA Structural Layout. Only Structural Systems expected to be present in more than 50 UBHM, UBHS and UBA buildings are shown in this graph.

Moving towards Structural Layouts of, on average, larger dimensions, the "Block" Structural Layouts are studied. Figure 48 shows example point-cloud generated geometries[13].

[13] It should be noted that buildings in the following similar figures Figure 48, Figure 50, Figure 52, are scaled down by half in comparison to Figure 45. Figure 44 can be used to appreciate the actual difference scale if the same scale is kept for all Structural Layouts.

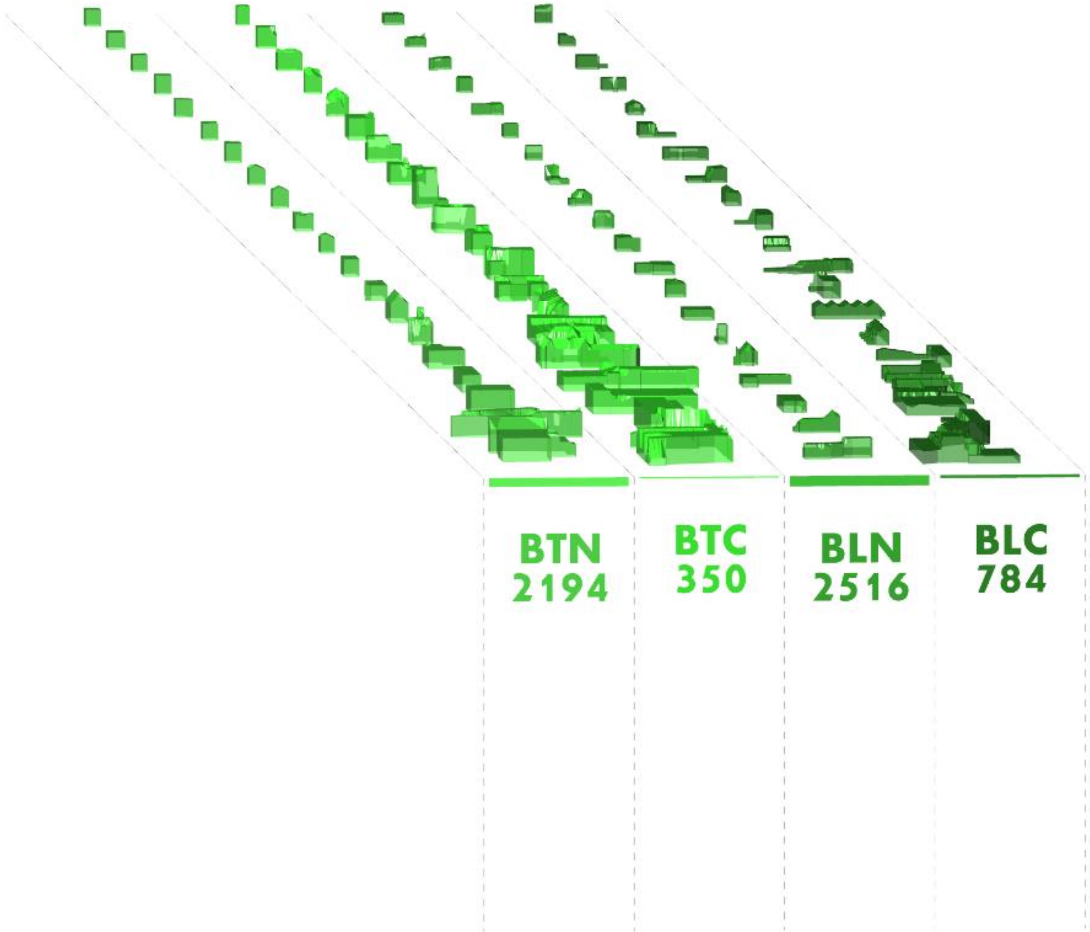


Figure 48 Example buildings and building count for the BTN, BTC, BLN and BLC Structural Layouts

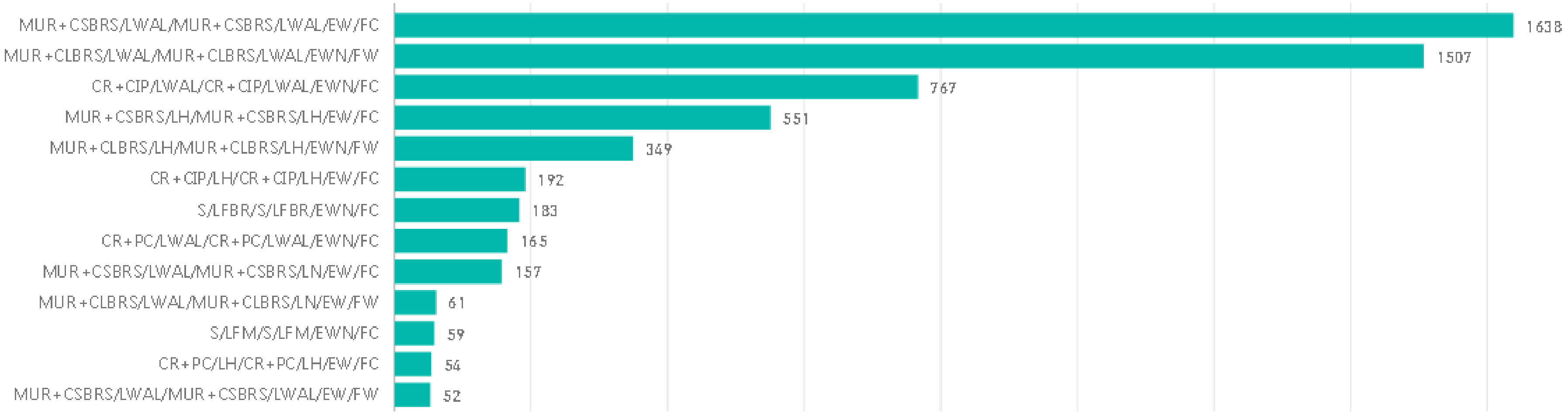


Figure 49 Expected building count for Structural Systems for BLC, BLN, BTC and BTN Structural Layout. Only Structural Systems expected to be present in more than 50 BLC, BLN, BTC and BTN buildings are shown in this graph.

On buildings in the "Barn" and "Warehouse" Structural Layouts, *frame systems (LFBR, LPB)* made with *Steel (S)* or *Timber (W)* structural elements become much more predominant. The Structural System which is expected to be most often present is a *Steel(S)* structure with a *braced frame system (LFBR)* in one direction and a *post-and-beam (LPB)* system in the other direction. More than 2 800 buildings with these characteristics are expected to be in the region. More than 40% of these buildings are expected be featuring a large flat roof – which is a characteristic of the WBW ("Warehouse") Structural Layout.

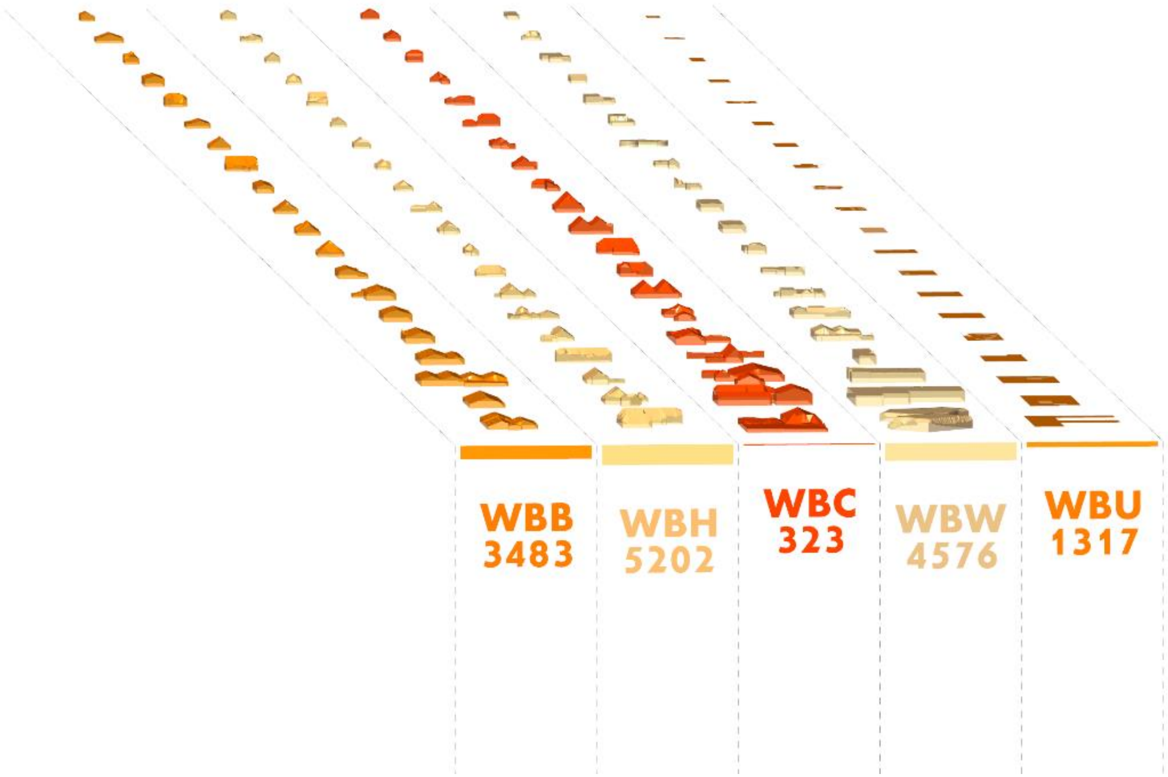


Figure 50 Example buildings and building count for the WBB, WBH, WBC, WBW and WBU Structural Layouts.

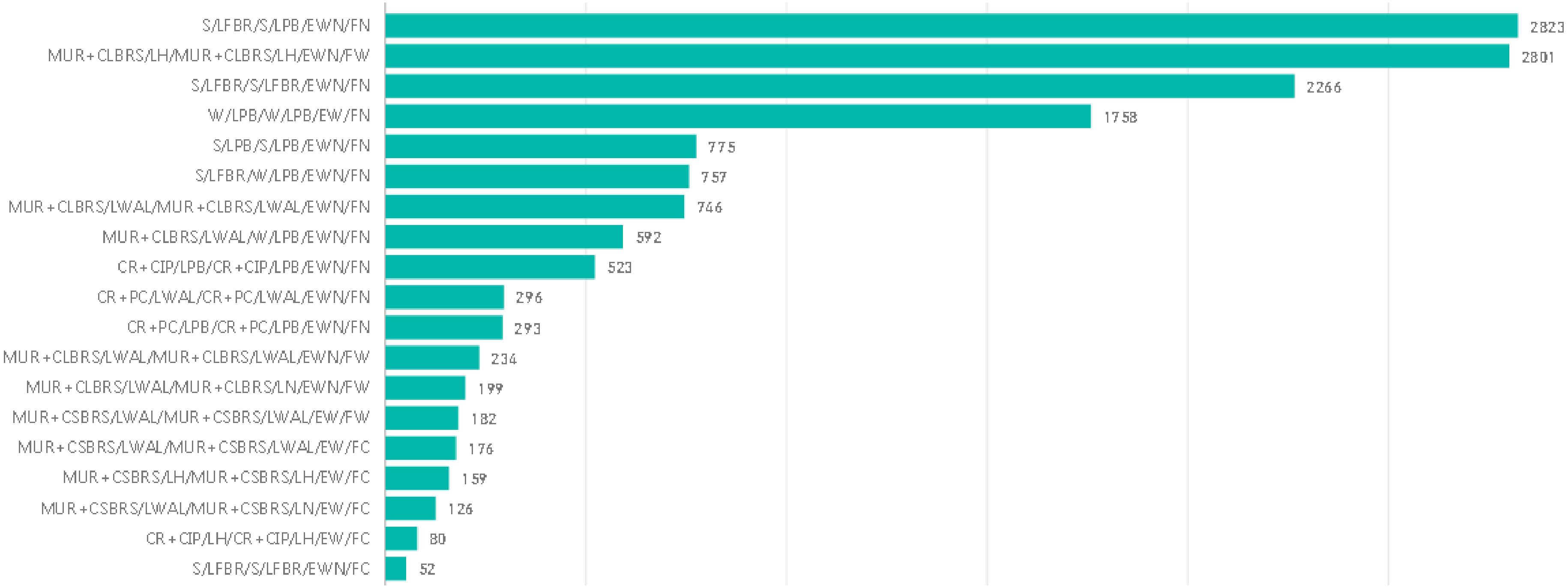


Figure 51 Expected building count for Structural Systems for WBB, WBC, WBH, WBU and WBW Structural Layout. Only Structural Systems expected to be present in more than 50 WBB, WBC, WBH, WBU and WBW buildings are shown in this graph.

Finally Figure 52 shows buildings of the TOW ("Tower") Structural Layout having a relatively larger height as well as buildings of the OSP ("Special") Structural Layout, featuring buildings visually identified as non-standard cases, e.g. for buildings with circular footprint. As shown in Figure 53 Buildings in the TOW Structural Layout are inferred to be Concrete (CR) or Steel(S) structures.

For OSP buildings an *unknown structural system (L99)* was assigned unless inspection data was present (see Figure 54).

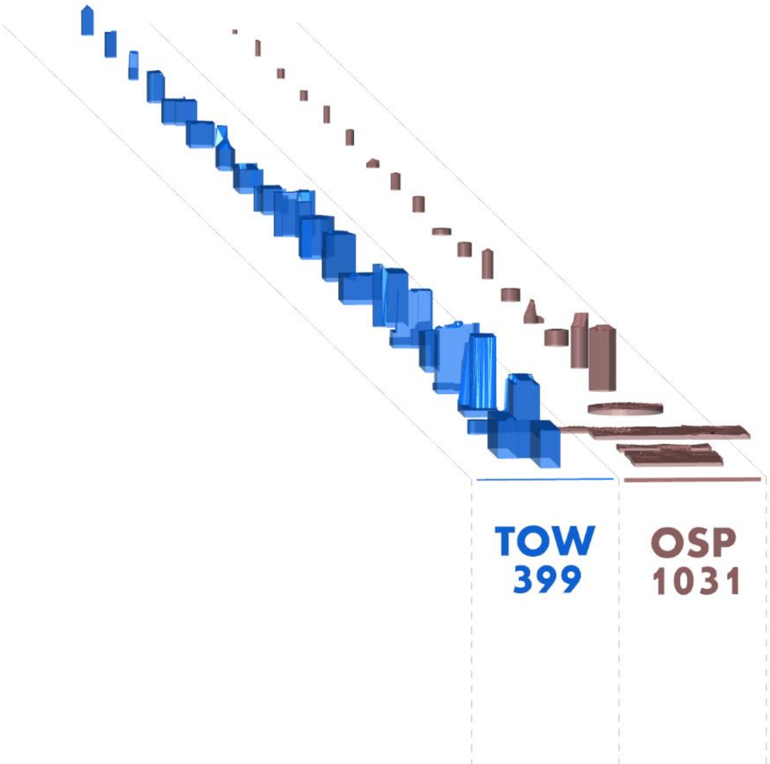


Figure 52 Example buildings and building count for the TOW and OSP Structural Layouts.

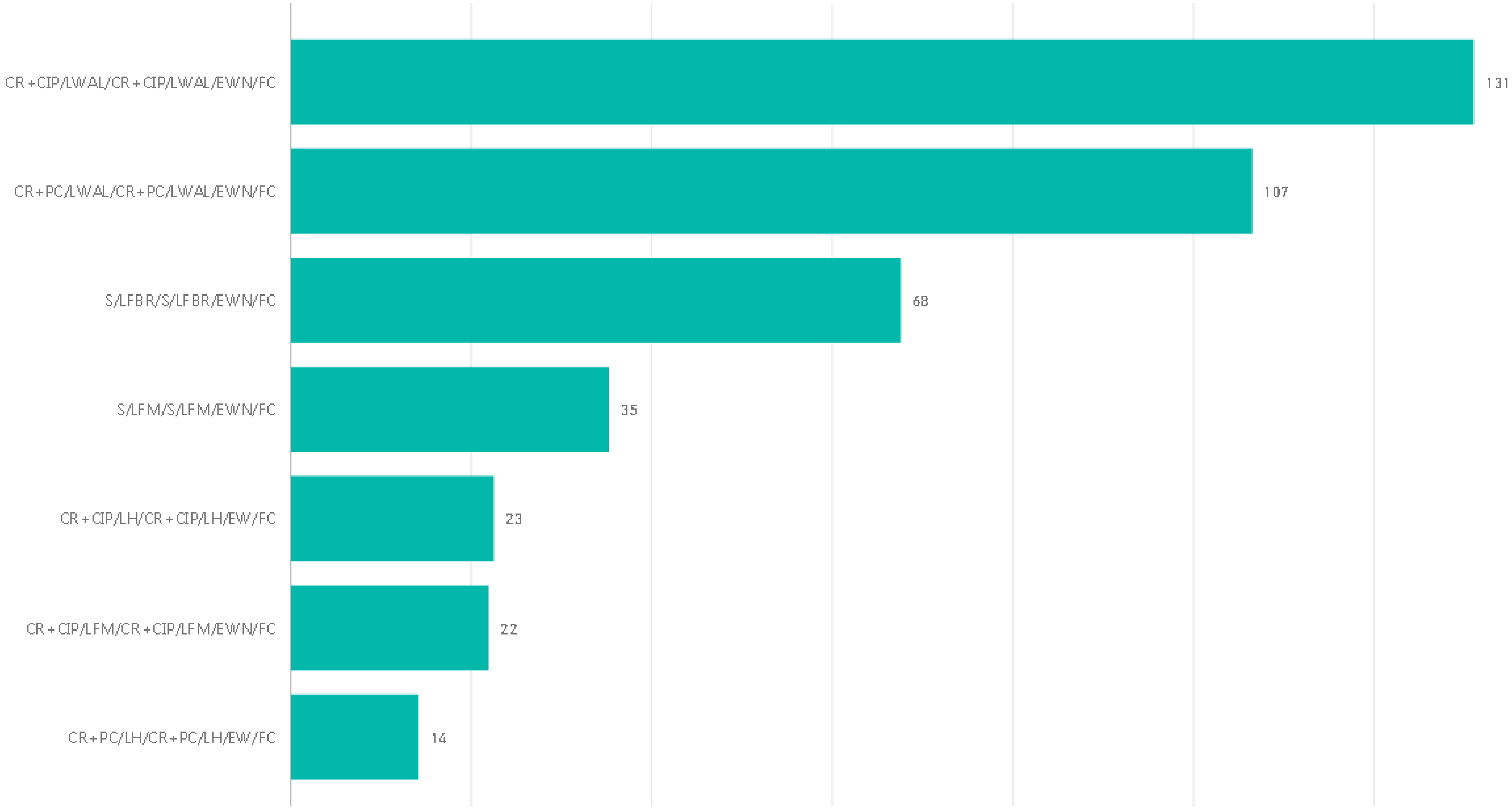


Figure 53 Expected building count for Structural Systems for the TOW Structural Layout.

Figure 54 Expected building count for Structural Systems for OSP Structural Layout. Only Structural Systems expected to be present in more than 50 OSP buildings are shown in this graph.

Based on the above study it is demonstrated how the classification in different Structural Layout classes resulted in the assignment of appropriate Structural Systems, through the use of corresponding age-based inference rules

Annex B

Additional Data Visualisations

In this section additional data analysis has been performed on buildings within the Geometric Layouts and Structural Layouts of EDB V5. The relative geometric uniformity of buildings within the EDB V5 classes can therefore be appreciated.

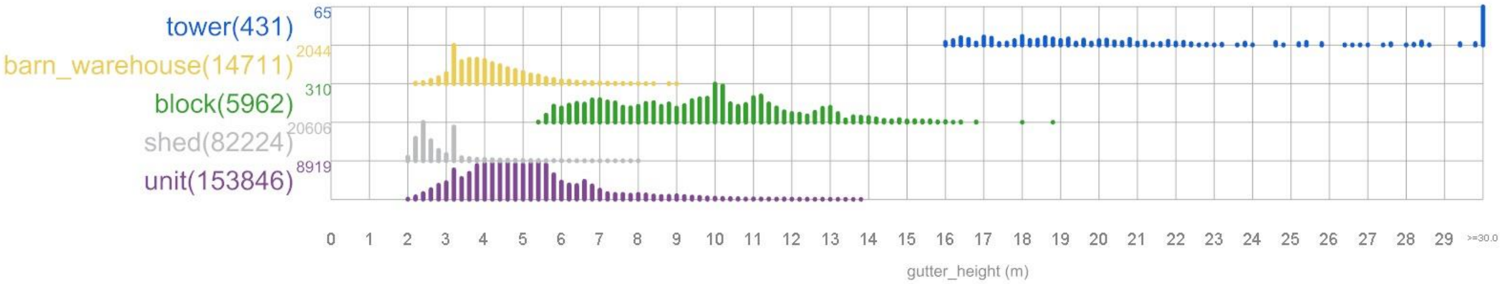


Figure 55 Histogram of building counts per gutter height domain for each Geometric Layout. The y-axis is renormalized in each Geometric Layout for visualisation purposes.

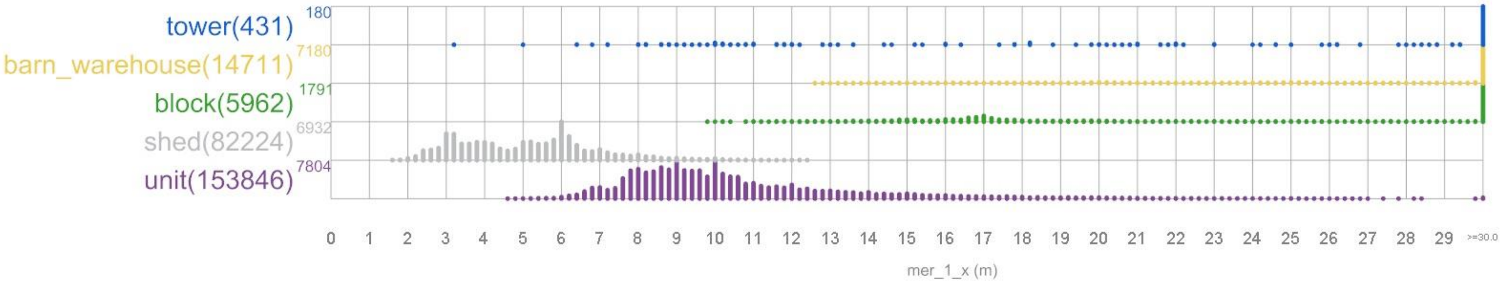


Figure 56 Histogram of building counts per MER length domain for each Geometric Layout. The y-axis is renormalized in each Geometric Layout for visualisation purposes.

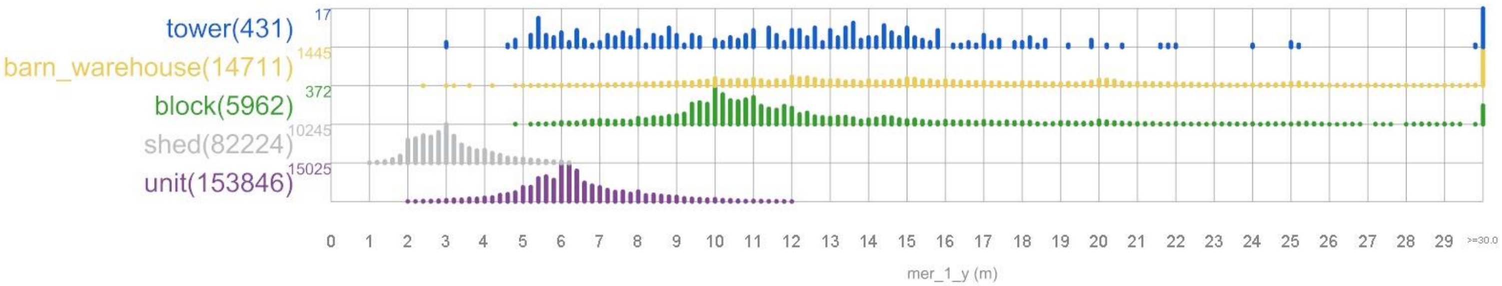


Figure 57 Histogram of building counts per MER width domain for each Geometric Layout. The y-axis is renormalized in each Geometric Layout for visualisation purposes.

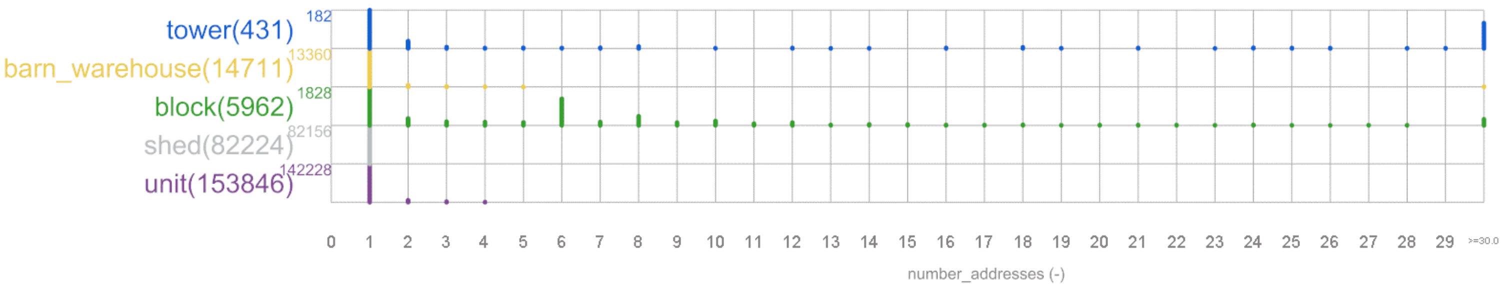


Figure 58 Histogram of building counts per number of addresses for each Geometric Layout. The y-axis is renormalized in each Geometric Layout for visualisation purposes.

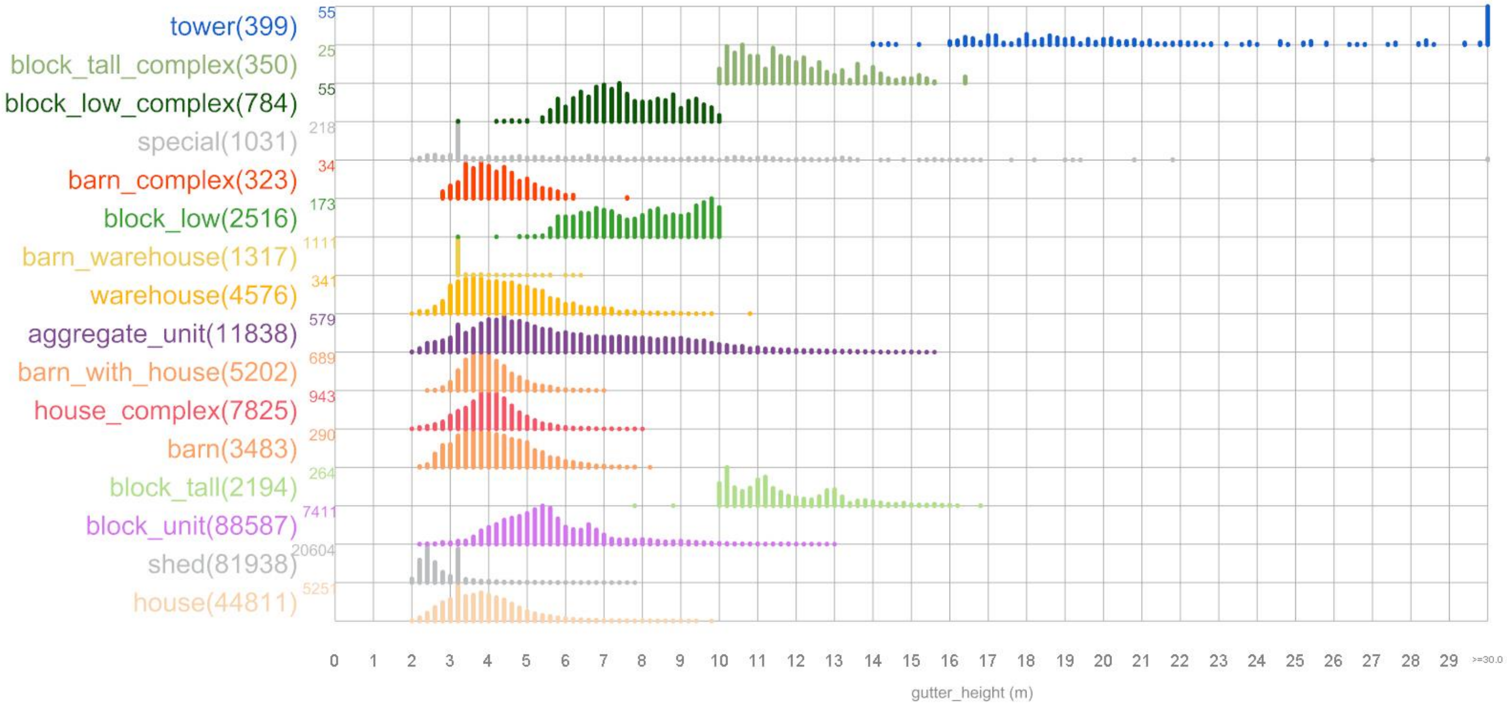


Figure 59 Histogram of building counts per building height domain for each Structural Layout. The y-axis is renormalized in each Structural Layout for visualisation purposes.

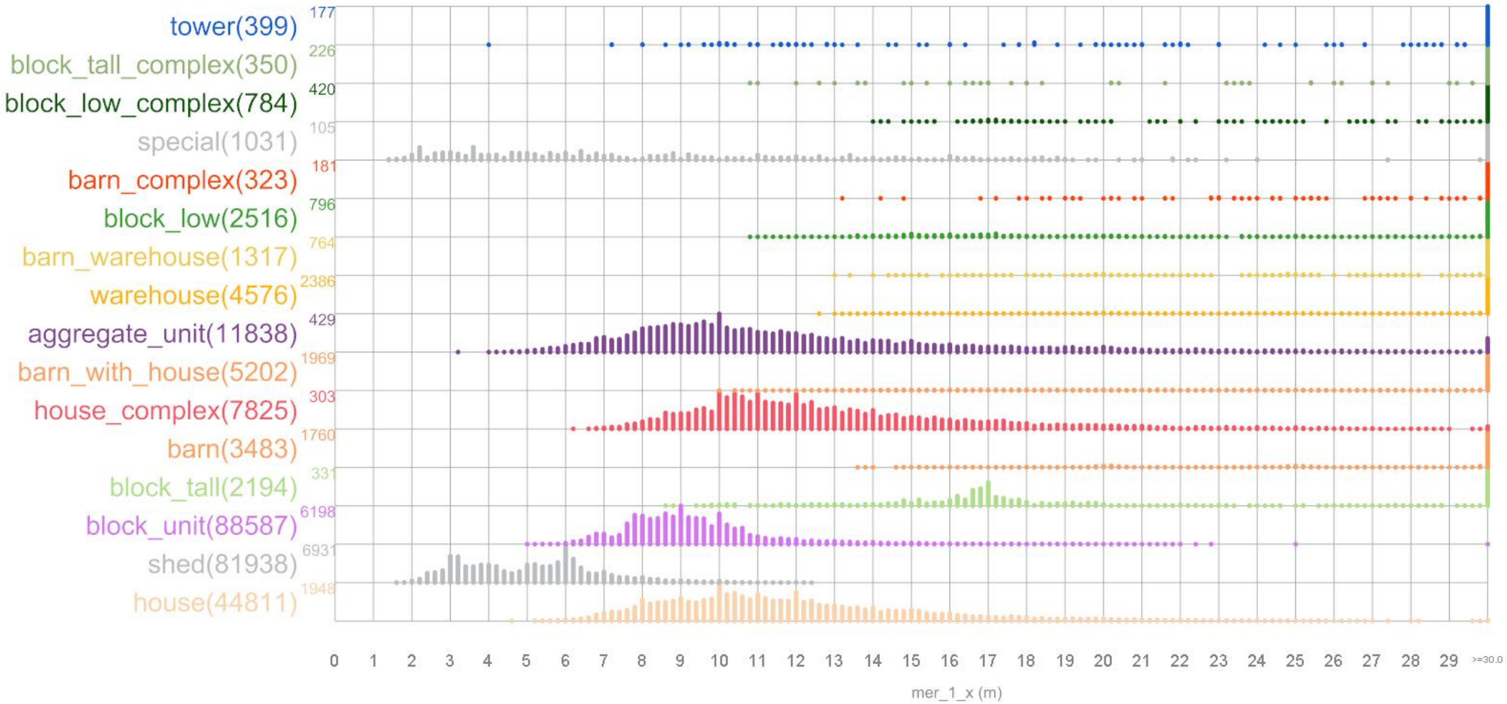


Figure 60 Histogram of building counts per MER length domain for each Structural Layout. The y-axis is renormalized in each Structural Layout for visualisation purposes.

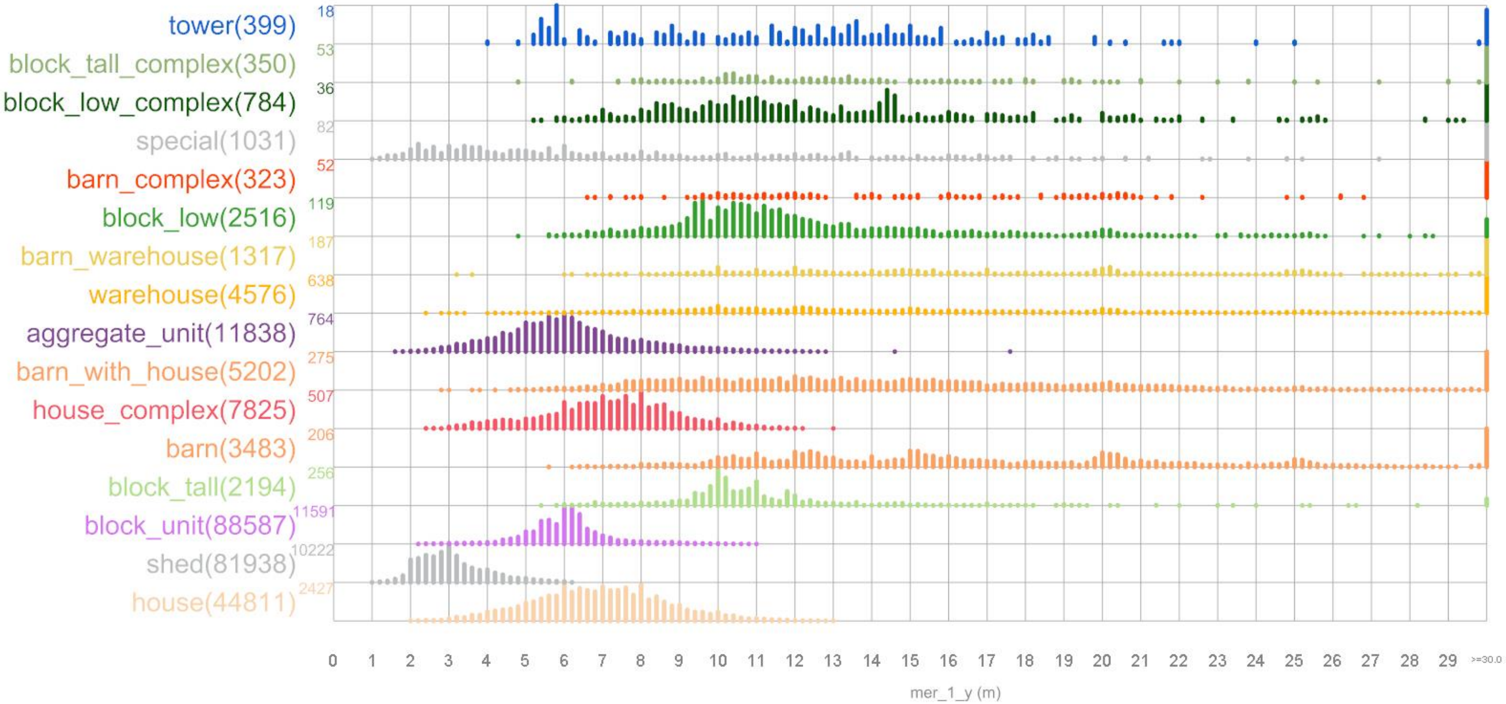


Figure 61 Histogram of building counts per MER width domain for each Structural Layout. The y-axis is renormalized in each Structural Layout for visualisation purposes.

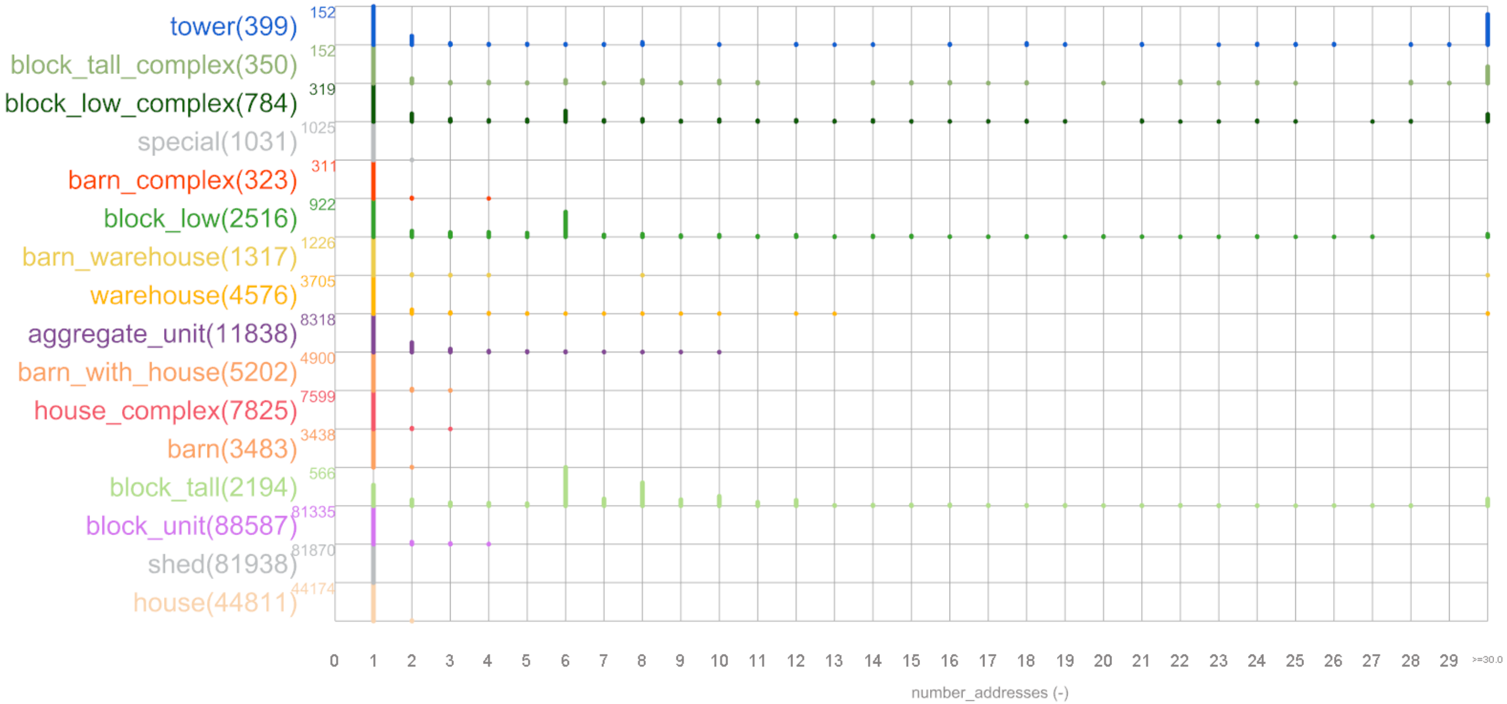


Figure 62 Histogram of building counts per number of addresses for each Structural Layout. The y-axis is renormalized in each Structural Layout for visualisation purposes.

In these histograms each parameter is visualised separately. The following scatter plot visualisations show the three geometric parameters at the same time for each Structural Layout of EDB V5. The arithmetic mean per scatter plot is also visualised in these Figures.

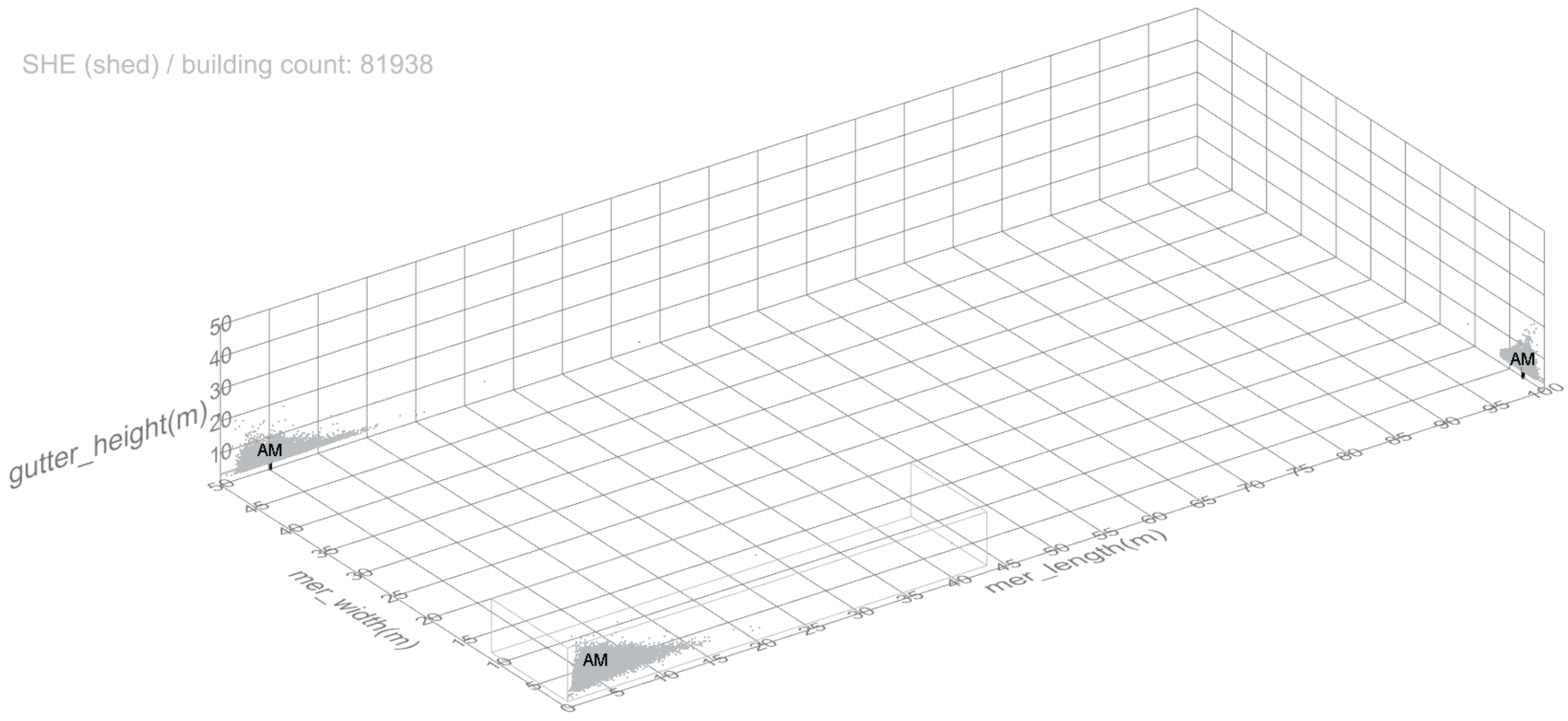


Figure 63 Scatter plot based on the MER width, MER length and gutter height values of buildings within the SHE Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

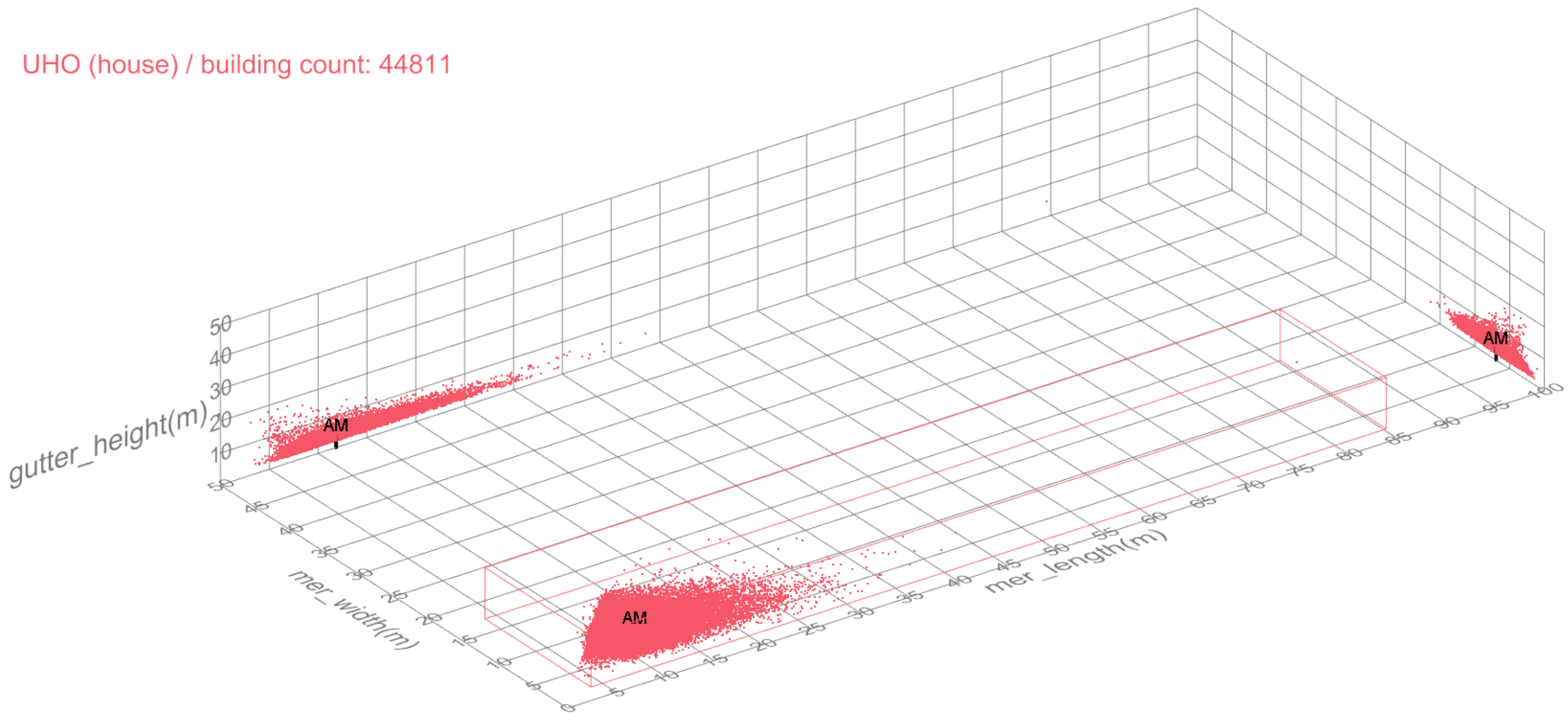


Figure 64 Scatter plot based on the MER width, MER length and gutter height values of buildings within the UHO Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

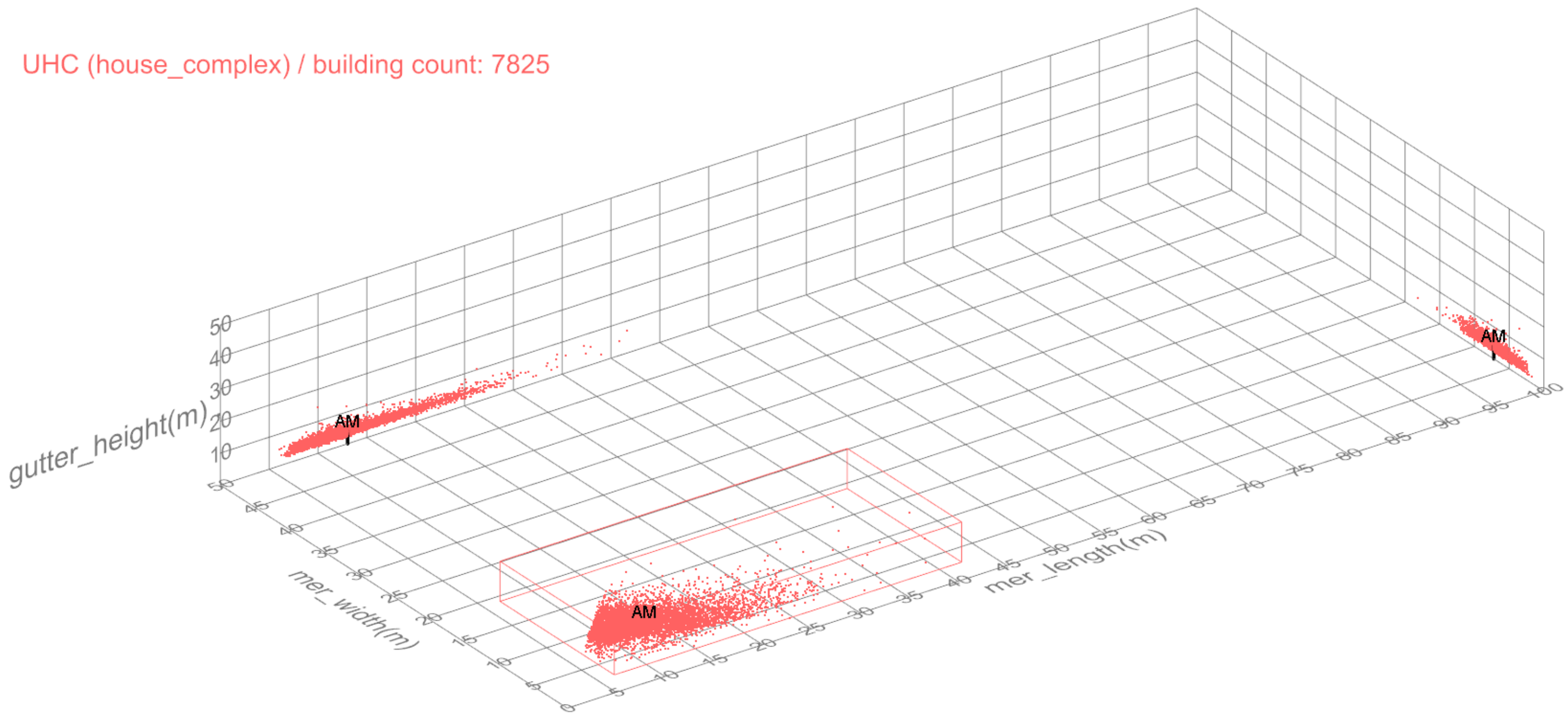


Figure 65 Scatter plot based on the MER width, MER length and gutter height values of buildings within the UHC Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

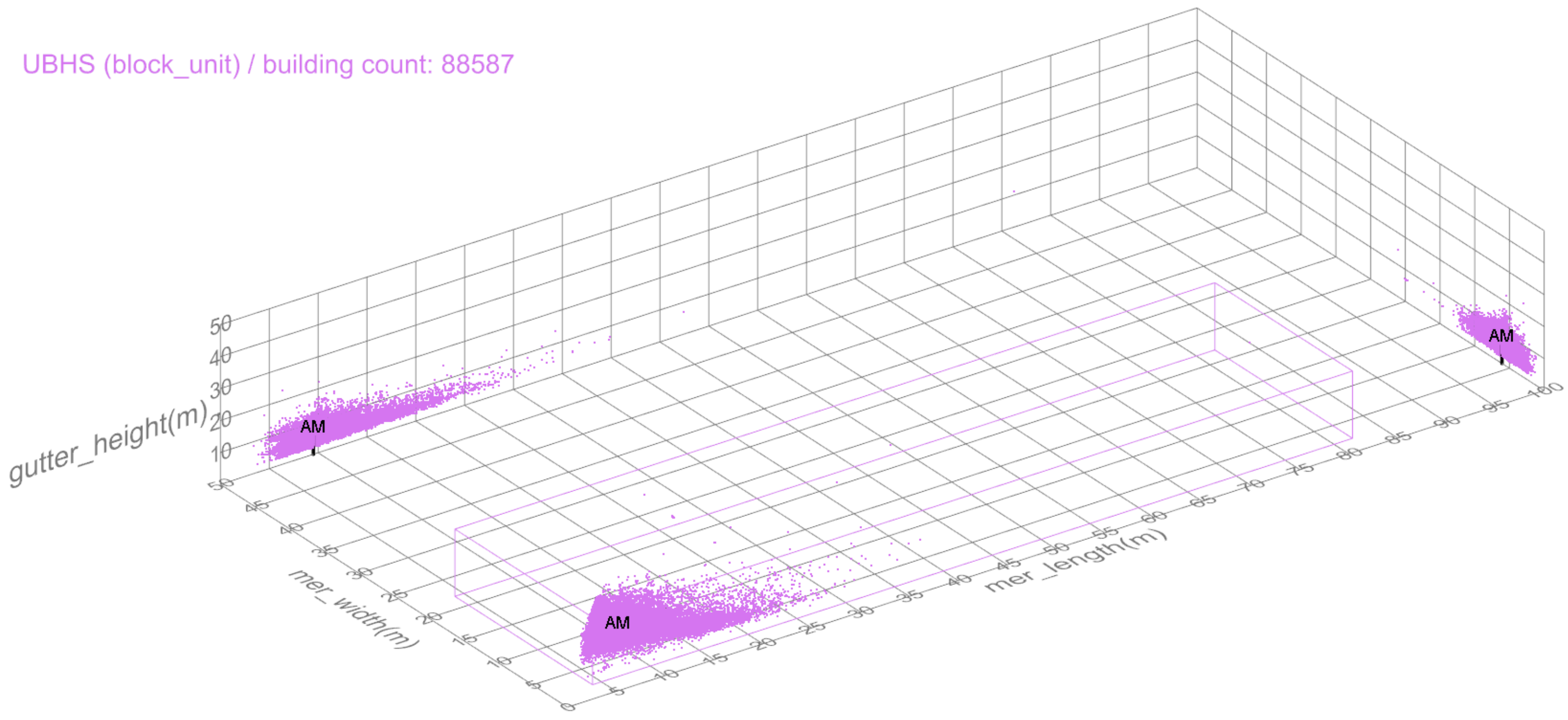


Figure 66 Scatter plot based on the MER width, MER length and gutter height values of buildings within the UBH Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

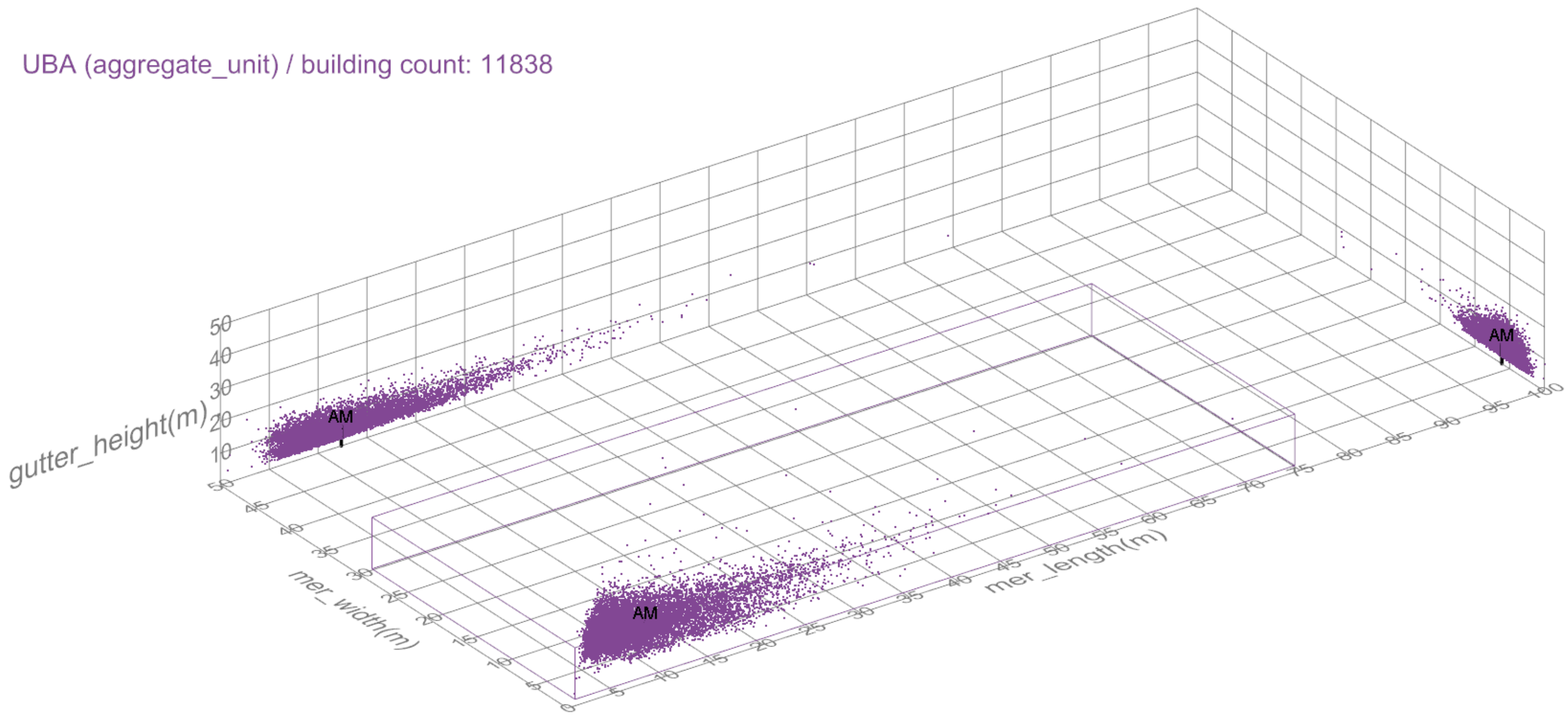


Figure 67 Scatter plot based on the MER width, MER length and gutter height values of buildings within the UBA Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

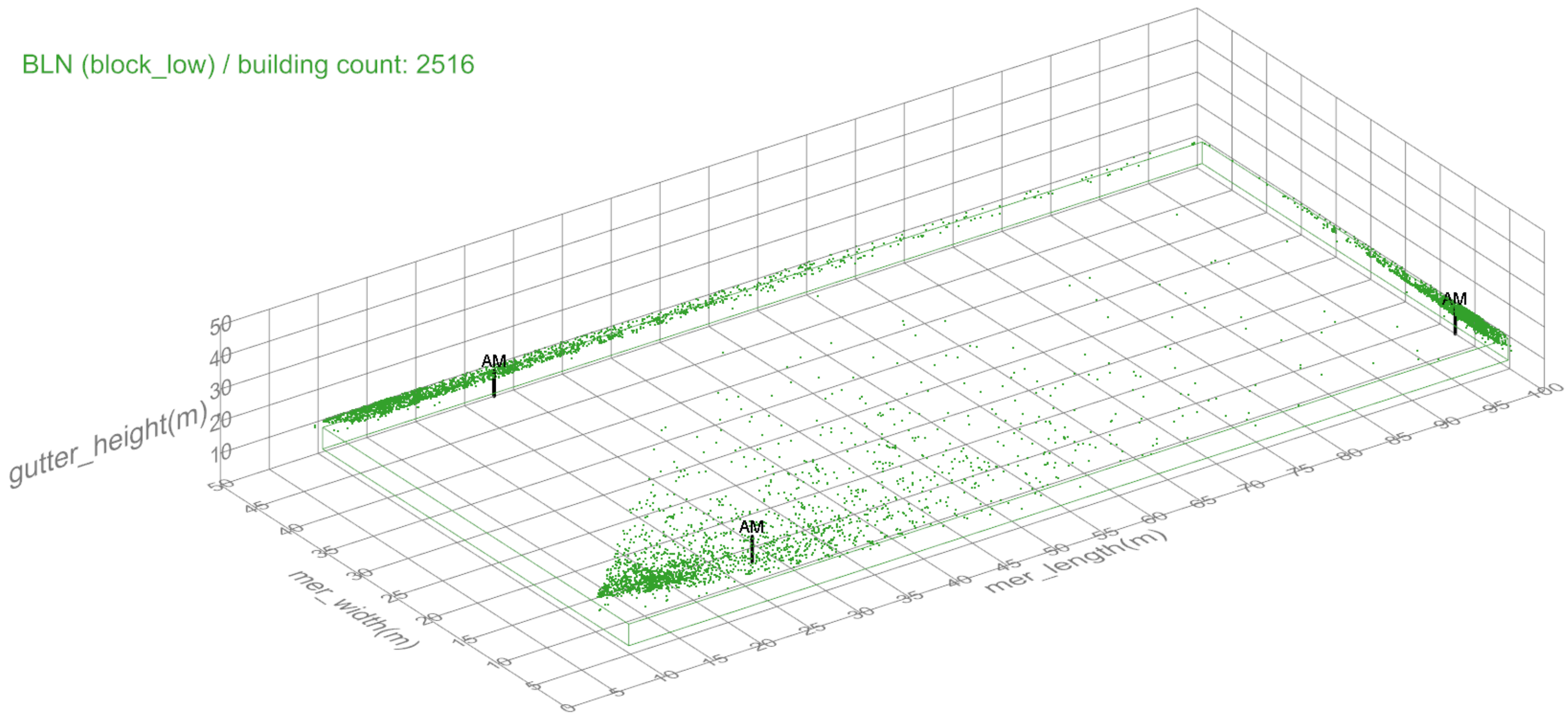


Figure 68 Scatter plot based on the MER width, MER length and gutter height values of buildings within the BLN Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

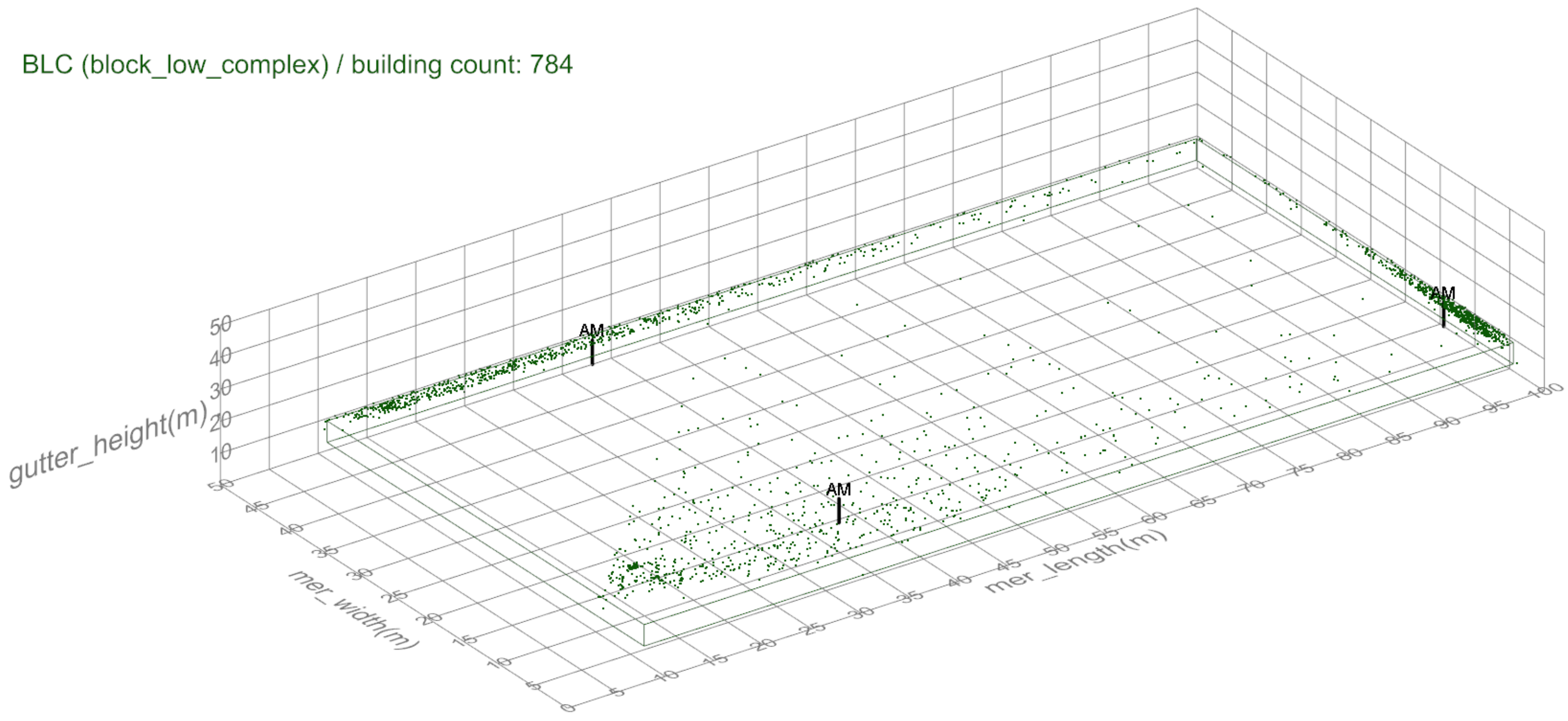


Figure 69 Scatter plot based on the MER width, MER length and gutter height values of buildings within the BLC Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

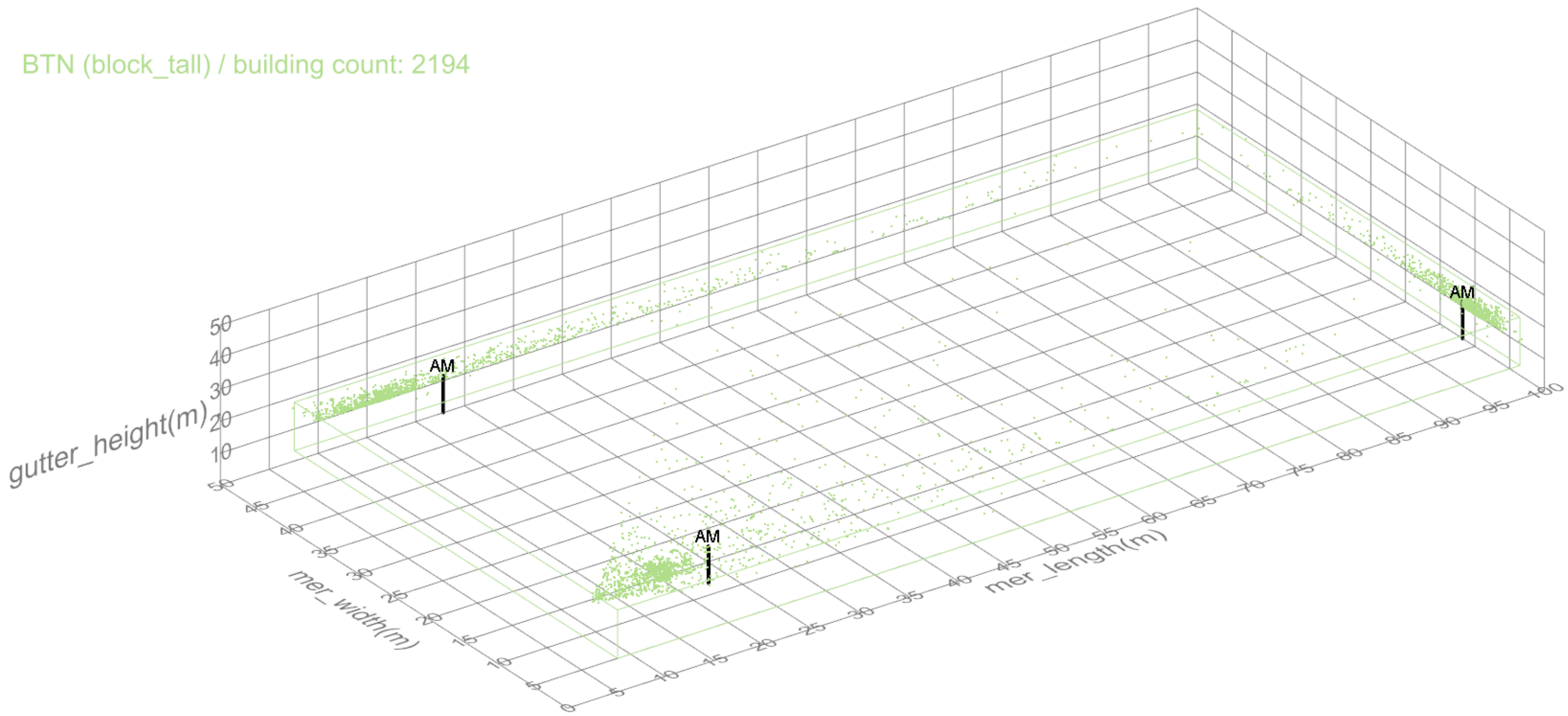


Figure 70 Scatter plot based on the MER width, MER length and gutter height values of buildings within the BTN Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

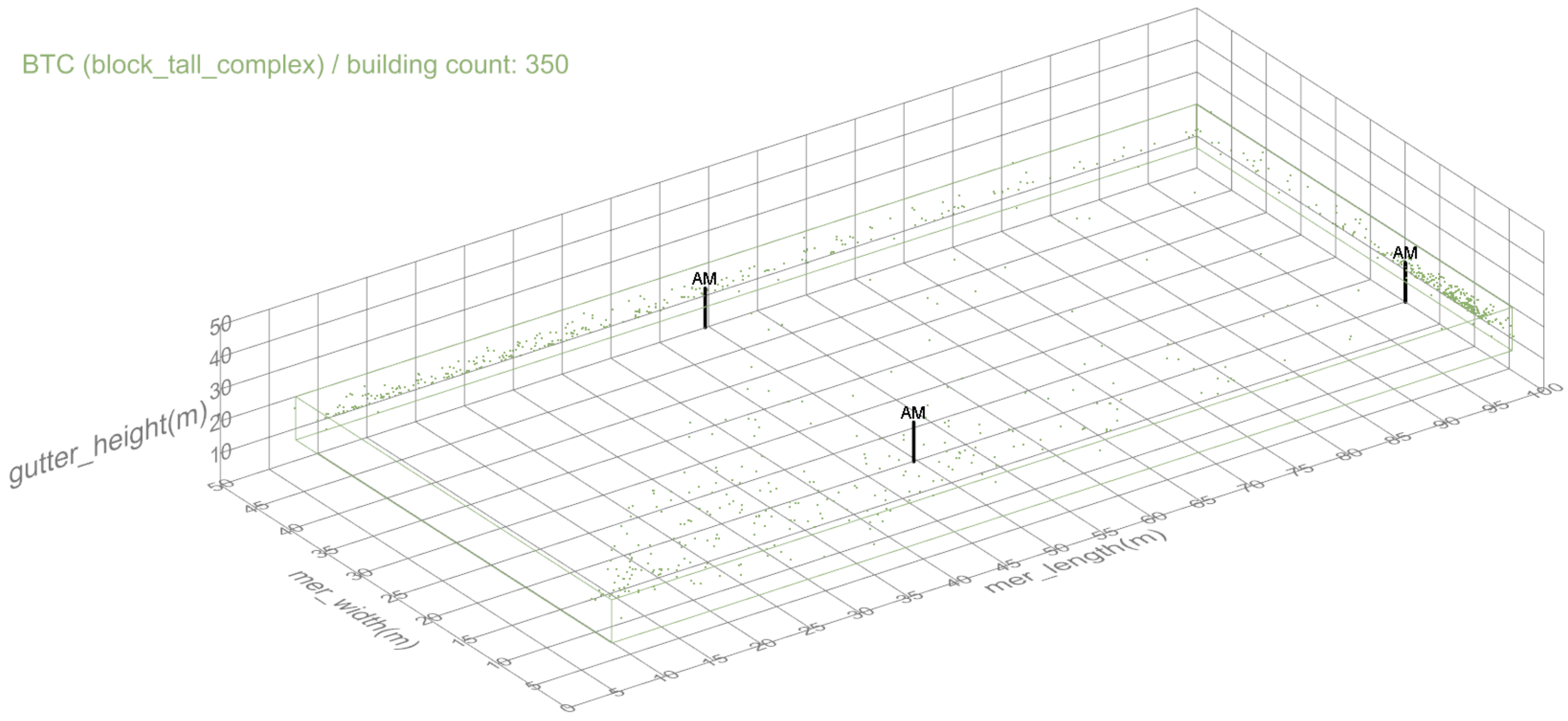


Figure 71 Scatter plot based on the MER width, MER length and gutter height values of buildings within the BTC Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

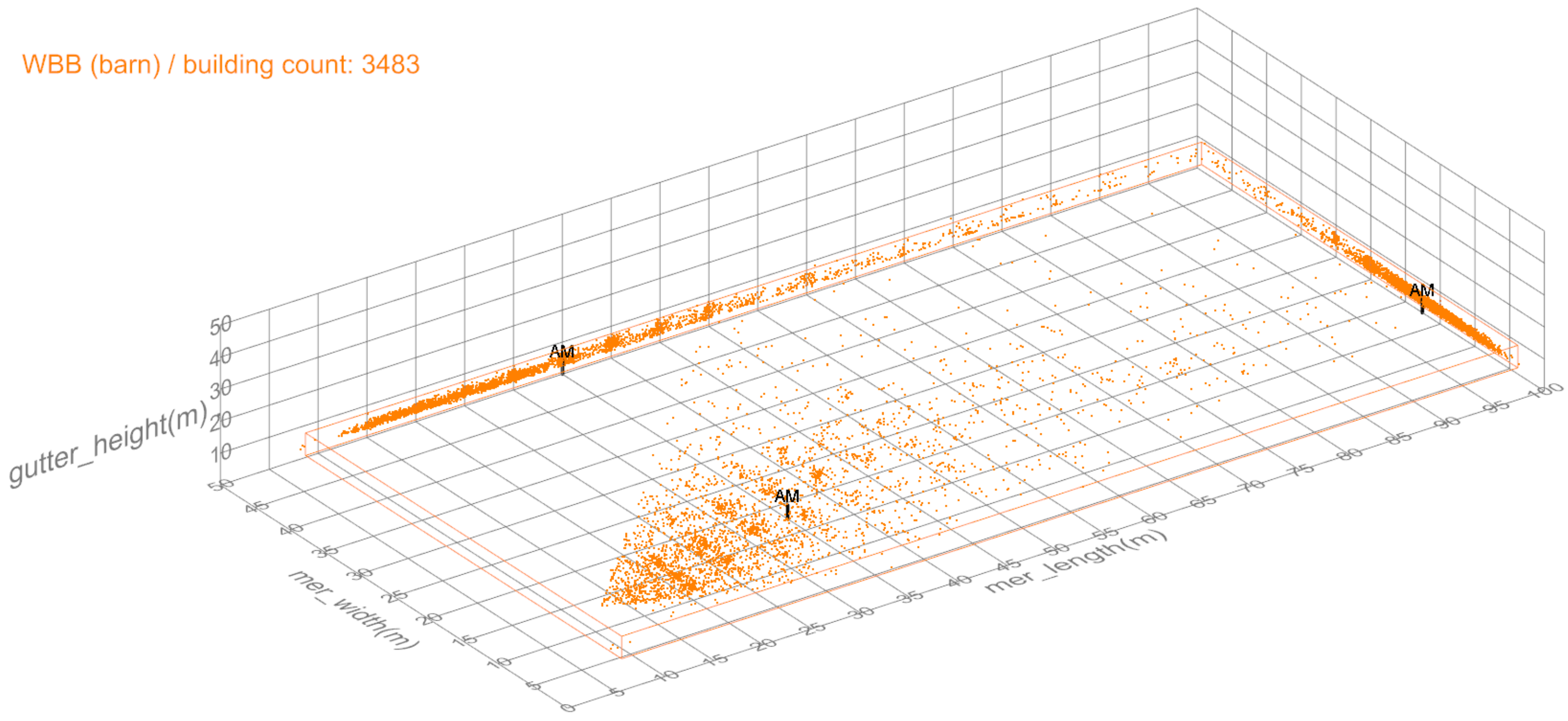


Figure 72 Scatter plot based on the MER width, MER length and gutter height values of buildings within the WBB Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

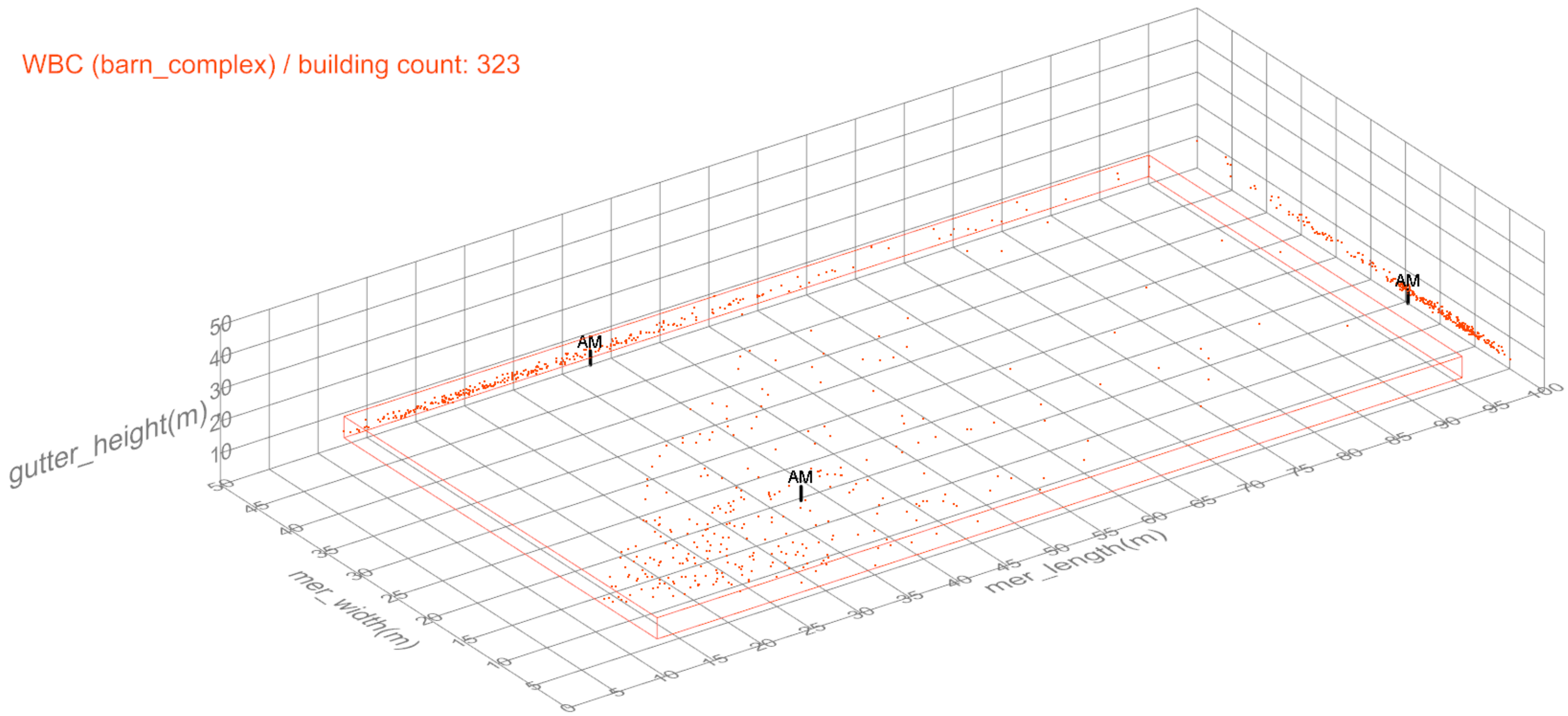


Figure 73 Scatter plot based on the MER width, MER length and gutter height values of buildings within the WBC Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

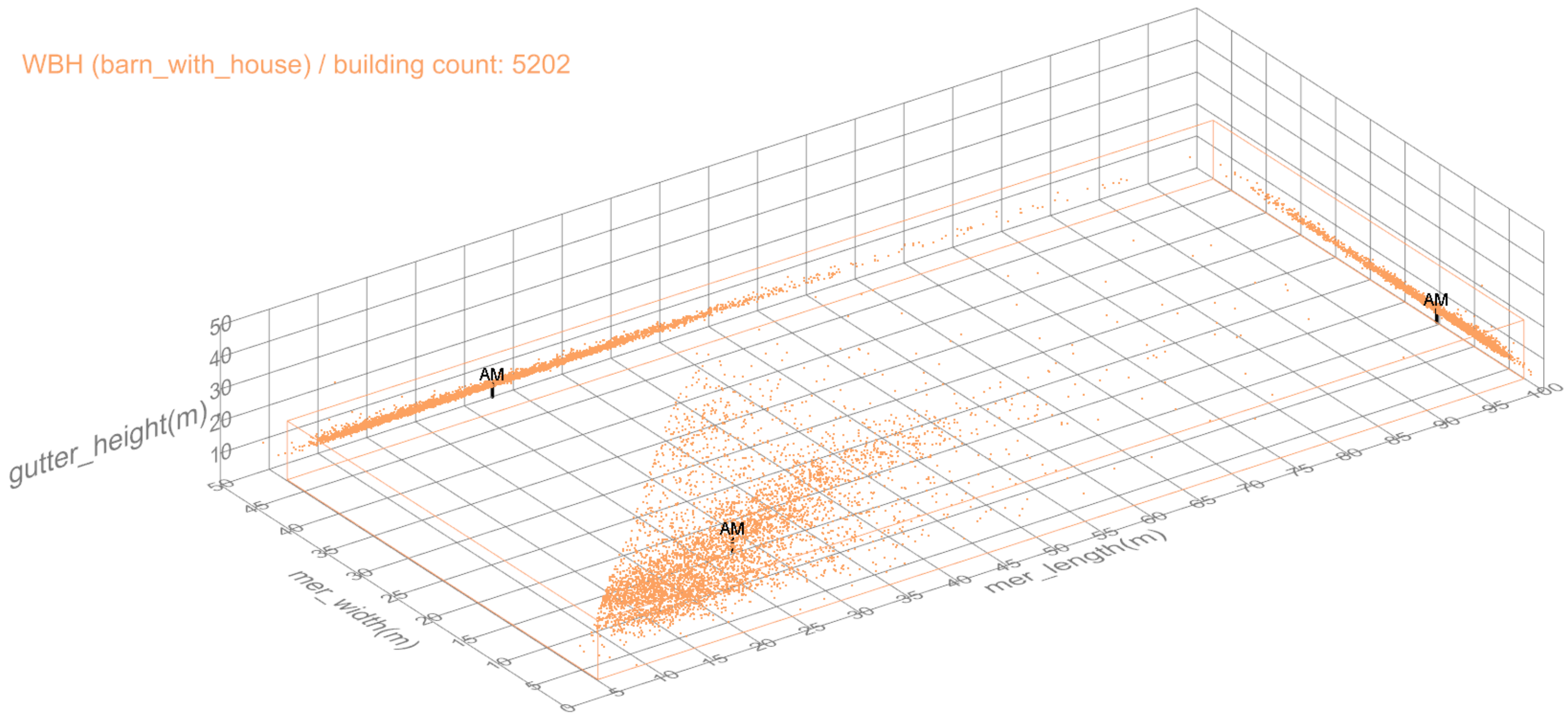


Figure 74 Scatter plot based on the MER width, MER length and gutter height values of buildings within the WBH Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

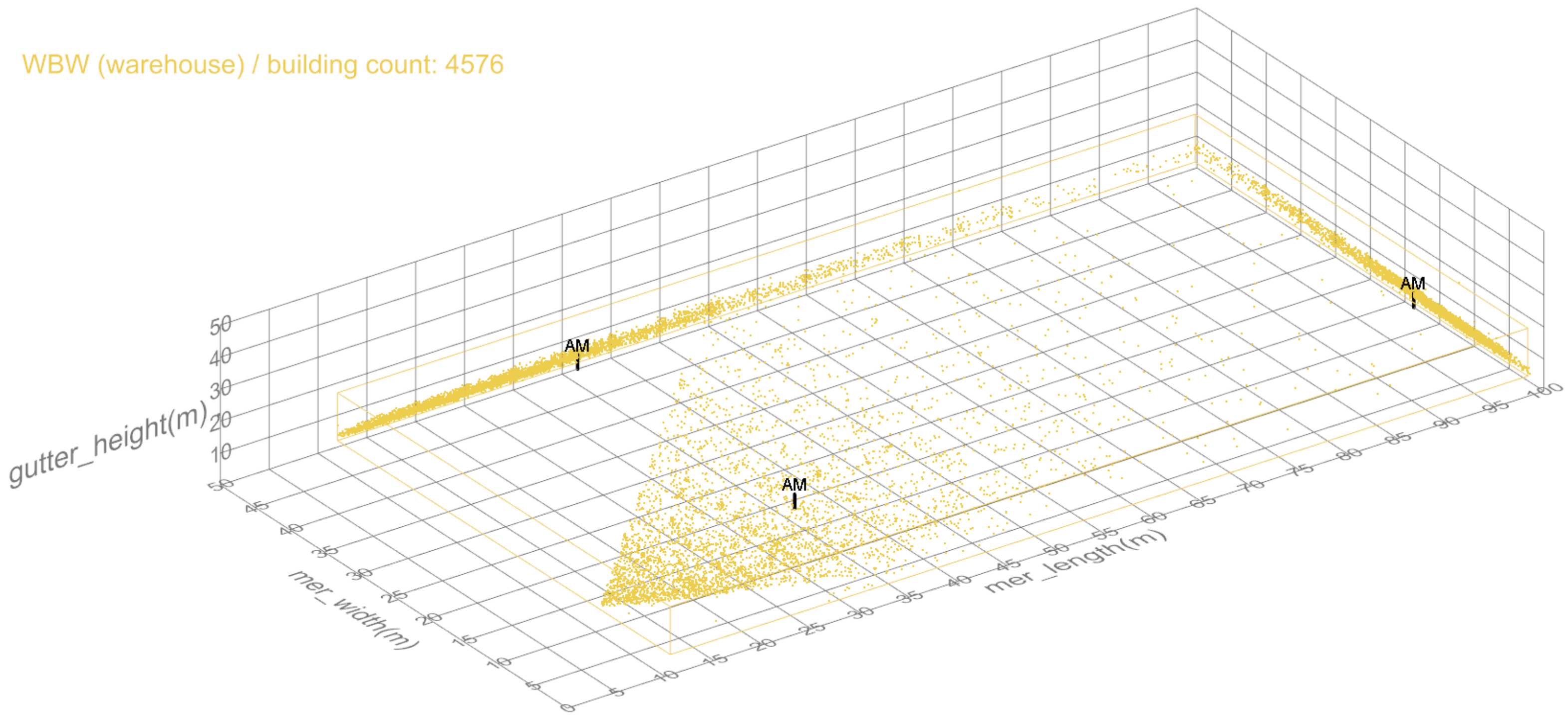


Figure 75 Scatter plot based on the MER width, MER length and gutter height values of buildings within the WBW Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

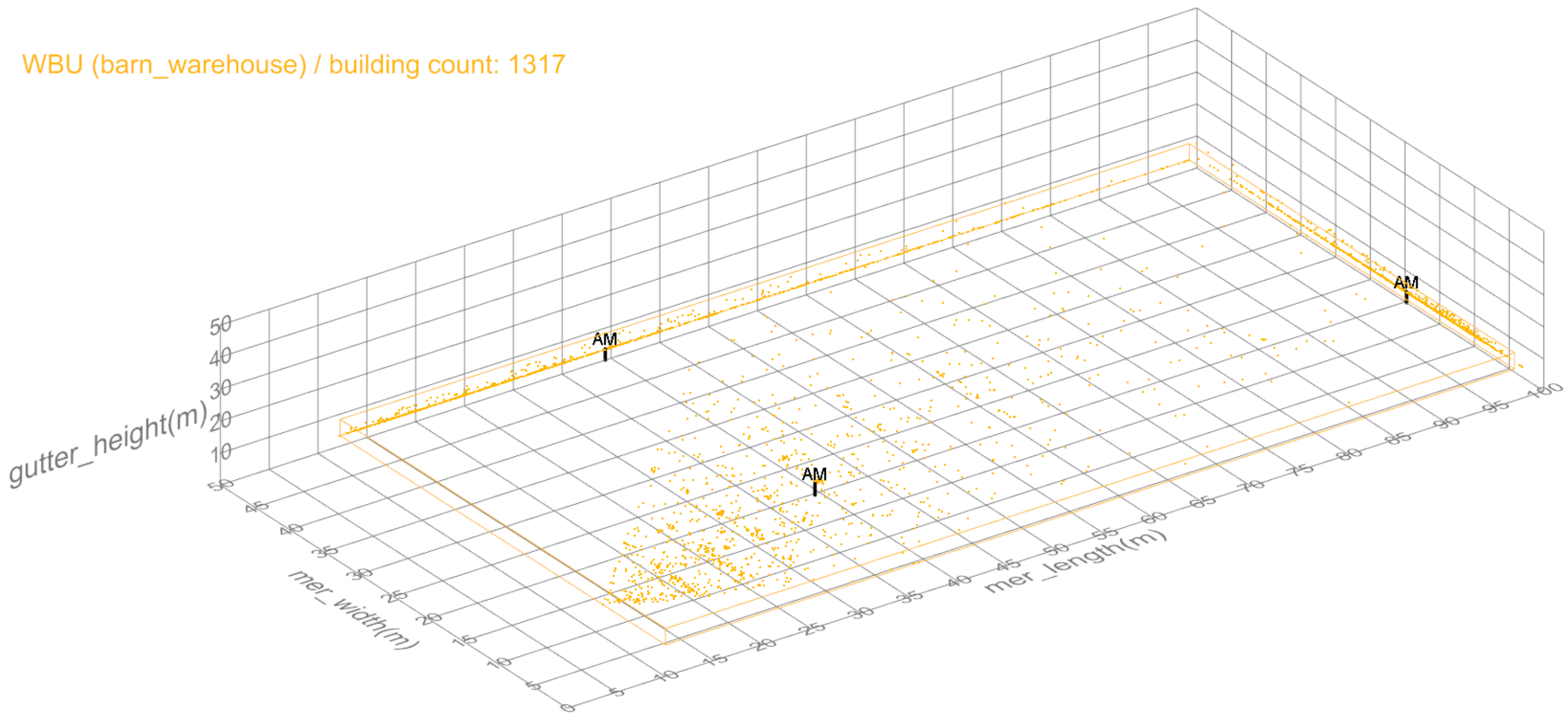


Figure 76 Scatter plot based on the MER width, MER length and gutter height values of buildings within the WBU Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

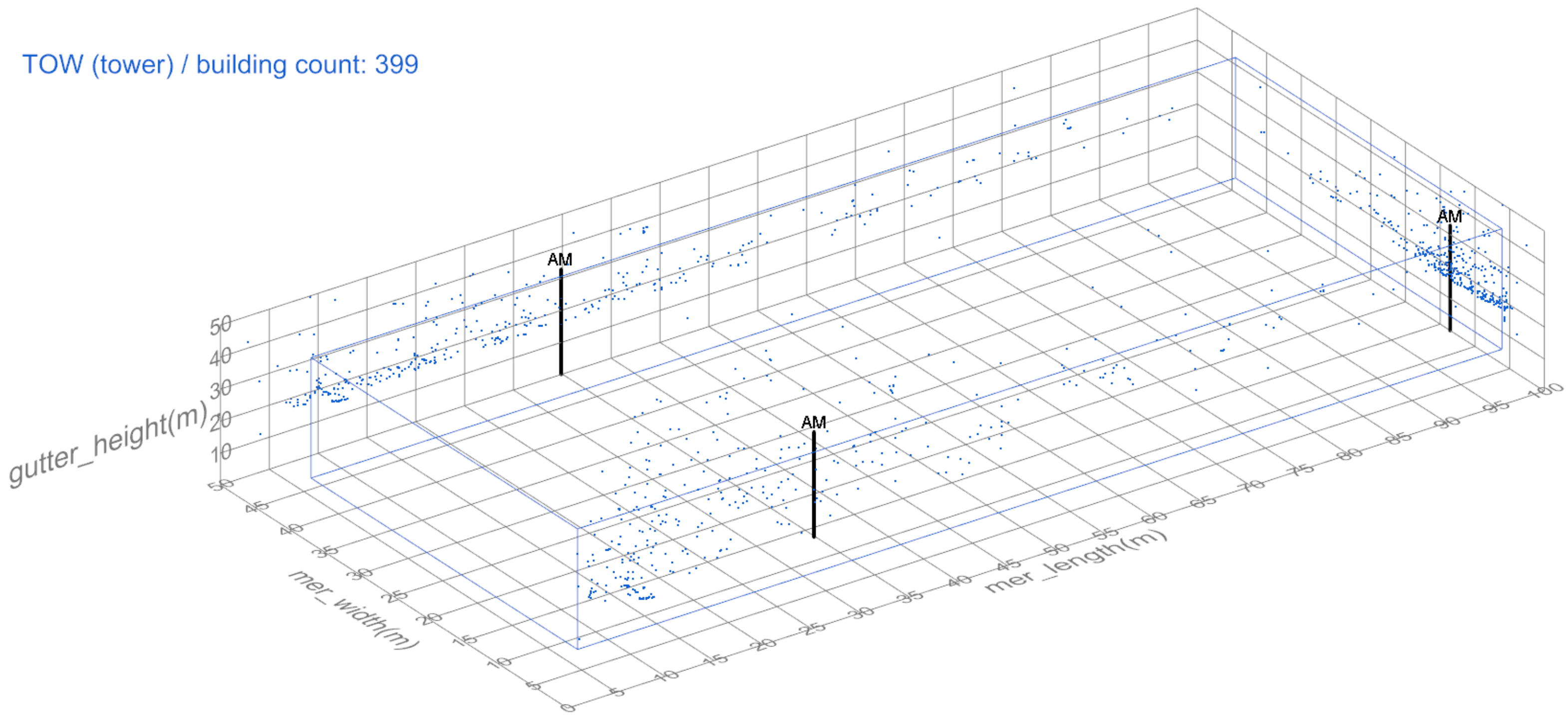


Figure 77 Scatter plot based on the MER width, MER length and gutter height values of buildings within the TOW Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.

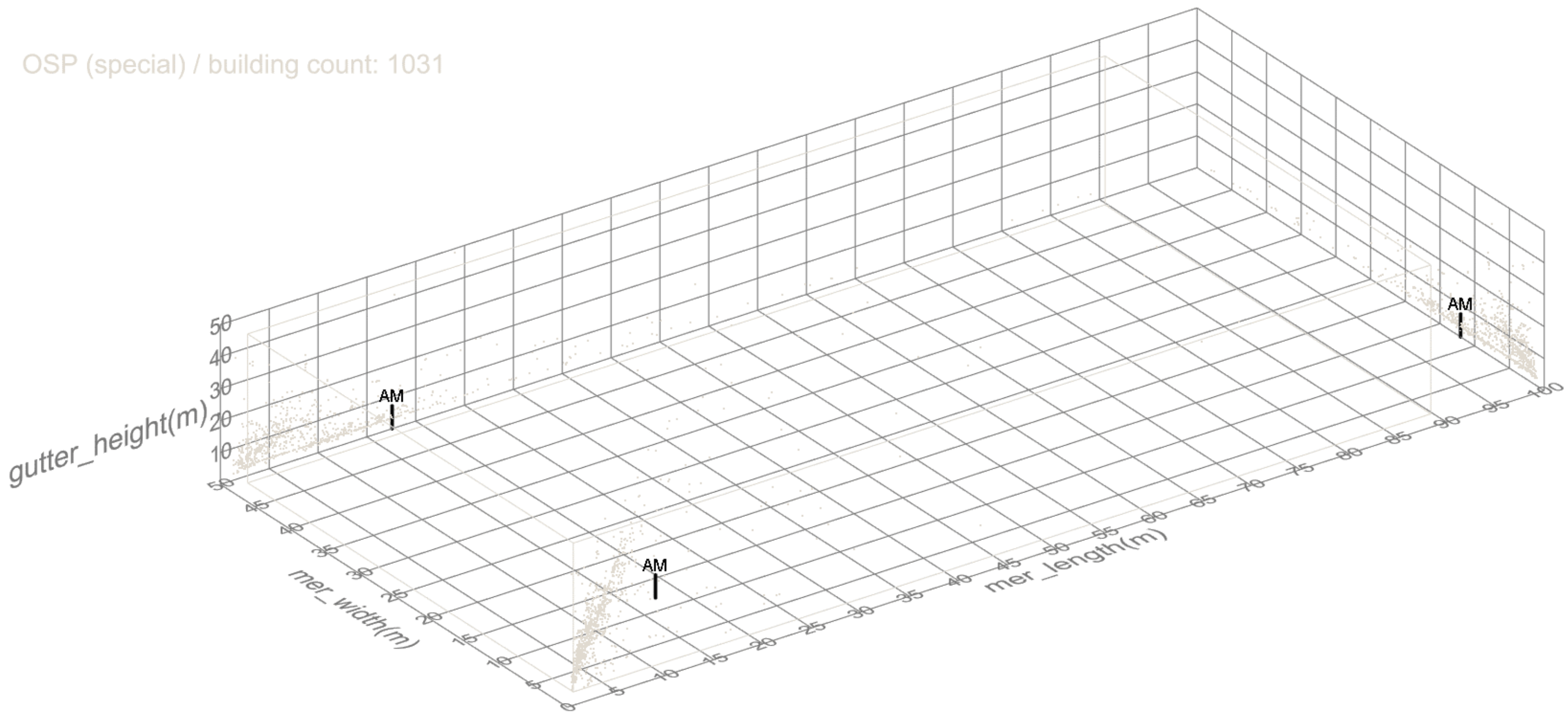


Figure 78 Scatter plot based on the MER width, MER length and gutter height values of buildings within the OSP Structural Layout. *AM* signified the coordinates of a point whose coordinates are the arithmetic mean values for each of these three geometric parameters.