Deriving and Benchmarking Building Renovation Passport Data Requirements Against CityGML and the Energy ADE

Daan H. Schlosser[1], Hiba Doi[1], Jantien E. Stoter[1], Giorgio Agugiaro[1]

[1] 3D Geoinformation group, Department of Urbanism, Faculty of Architecture and the Built Environment,
Delft University of Technology, Delft, The Netherlands
daan.schlosser@gmail.com, (h.d.doi, j.e.stoter, g.agugiaro)@tudelft.nl

Keywords: Building Renovation Passport, Digital Building Logbook, CityGML, Energy ADE, Data interoperability, Semantic mapping, Building stock, Open data.

Abstract

Reaching the European Union's 2050 climate targets depends on renovating a building stock responsible for a large share of the EU's energy use and greenhouse-gas emissions. The recast Energy Performance of Buildings Directive (EU/2024/1275) proposes two instruments: the Digital Building Logbook (DBL), a repository for building data across the whole life cycle, and the Building Renovation Passport (BRP), a stepwise renovation roadmap that draws on those data. Both require a data model to make building data machine-readable and interoperable, yet few European initiatives provide one: most publish semi-structured spreadsheets, software templates, or closed schemas, so their coverage cannot be assessed and each new implementation starts from scratch.
This paper tests whether an open, formally specified geospatial data model can cover the data needs of a BRP, based on CityGML 2.0 paired with the Energy Application Domain Extension 3.0. From a comparative analysis of 14 heterogeneous European initiatives, we synthesise a so-called Minimum Set of Required Data (MSRD) of 275 attributes, benchmark it field by field against the candidate schema, and test the mapping at two scales on Dutch building-stock data. CityGML 2.0 with Energy ADE 3.0 already covers roughly 95% of the MSRD. The few missing fields are specific and isolated, most concentrated in the renovation-advice module. These findings indicate that CityGML and the Energy ADE already meet most BRP data needs, with the remaining gaps addressable through targeted improvements in further developments of the Energy ADE.

1. Introduction

The building sector accounts for roughly 40% of energy use and 36% of energy-related greenhouse-gas emissions in the European Union. Reaching the 2050 climate targets therefore depends on renovating its largely old and inefficient stock (European Commission, 2020). An Energy Performance Certificate (EPC) records how a building performs at a single point in time, but it does not tell the owner how and when to improve it. Planning a renovation needs a forward-looking, building-specific plan that an EPC does not provide. The revised European Directive on Energy Performance of Buildings (EU/2024/1275) answers this need with two complementary instruments. It also mandates that all EU Member States establish Building Renovation Passport (BRP) frameworks by May 2026 (European Parliament, 2024). A *Digital Building Logbook* (DBL) can be seen as a central repository for all relevant building data across its life cycle. A *Building Renovation Passport* is a building-specific, long-term renovation roadmap, consisting of a sequence of energy-efficiency renovations, that draws on and is stored in or accessible via a DBL.

Both instruments need a data model (i.e. a formal definition of what building data are relevant, stored in a reusable, exchangeable form) to be machine-readable and interoperable. In practice, though, few existing initiatives provide one. The term data model itself appears mainly in recent academic proposals (Gómez-Gil et al., 2024, Malinovec Puček et al., 2023), while most operational logbooks and passports express their data as semi-structured spreadsheets, tabular templates, or fields embedded in software rather than as an open, formally specified schema. Despite the legislative momentum, fragmentation persists: there is still no consensus on what data a BRP should contain, nor on how these data should be structured for cross-border and cross-domain interoperability (Karami et al., 2025), and, more fundamentally, no common model exists (Gómez-Gil et al., 2022). The practical consequence is that their coverage and completeness cannot be systematically assessed, and each new initiative defines its own set of attributes from scratch, with no shared, reusable model to build on.

An open, formally specified (geospatial) data model for urban energy modelling already exists which is adjacent to this use case: the OGC City Geography Markup Language (CityGML) paired with its Energy Application Domain Extension (Energy ADE) (Agugiaro and Padsala, 2025). However, whether this data model is sufficient as a BRP data model has not yet been tested. This paper answers this question through three contributions. First, it synthesises a so-called *Minimum Set of Required Data* (MSRD) from 14 existing European initiatives. Second, it benchmarks the MSRD against the CityGML 2.0 + Energy ADE 3.0 (beta 8) data model, mapping the set field by field to establish how much of it the schema can carry and where the gaps remain. Third, it tests the data model at two scales on real Dutch building-stock data: an audit-depth test on a single owner-occupier dwelling, and a city-scale breadth test on entire municipalities populated from open sources. The MSRD and the two-scale test pipeline are released openly (Schlosser, 2026a).

2. Related Work and Background

2.1 EPCs, DBLs, and BRPs

The literature is split between two framings of how these two instruments, the Building Renovation Passport (BRP) and the

Digital Building Logbook (DBL), relate. Earlier research projects positioned the BRP as the overarching instrument containing a logbook and a roadmap, while more recent EU policy instruments position the DBL as the overarching repository that stores renovation passports among other data (Hartenberger et al., 2021, Buchholz and Lützkendorf, 2023). Following the recast EPBD and the most recent policy positioning, this work treats the BRP as a building-specific renovation roadmap that draws on the broader DBL repository; the BRP therefore needs only a subset of the data a DBL would hold.

Besides defining the instrument, the 2024 EPBD recast[1] also specifies what a renovation passport must contain:

- the building's current energy performance and applicable renovation requirements;
- a staged renovation roadmap presented graphically with an explanation of step order;
- per-step estimates of energy savings, emission reductions, and bill savings;
- the energy-performance class reached;
- the share of renewable energy after renovation;
- wider information on circularity, whole-life-cycle emissions, health, comfort, indoor environmental quality, and climate-adaptation capacity;
- pointers to financial support and advisory services.

These mandated items, together with the building data needed to derive them, define what a BRP draws on. This paper takes this definition as the working framework.

2.2 The assessed initiatives

We assessed 14 initiatives, deliberately heterogeneous in scope and origin (Table 1). They fall into four groups. Three are official government instruments at national or regional level (the German iSFP, the Flemish Woningpas, the French CLÉA). Six are data-model proposals from EU-funded research projects (iBRoad, ALDREN, X-tendo, EUB SuperHub, BUILDCHAIN, OneClickRENO). Two are academic or framework studies (Gómez-Gil et al., the Ecorys EU DBL framework). The remaining three are Dutch operational tools that consume building data without proposing a data model of their own (Hoomdossier, the Vitec Vabi EPA software, and BuurtWarmteWijzer).

Including these data-consumer tools alongside the BRP proposals is deliberate. It lets the synthesis compare initiatives by how they treat data, rather than review BRPs alone.

Earlier comparative work is narrower. Gómez-Gil et al. (2022) set four DBL proposals side by side, and Karami et al. (2025) semantically compared logbook and EPC models. Both efforts cover fewer cases than the 14 initiatives assessed here.

[1] Art. 12 and Annex VIII.

ID	Initiative	Type / scope
I1	EUB SuperHub DBL	Data-model proposal
I2	ALDREN	Data-model proposal
I3	Gómez-Gil et al.	Academic proposal
I4	iBRoad	Data model + roadmap
I5	Woningpas	Implementation (Flanders)
I6	X-tendo	Data-model proposal
I7	EU DBL (Ecorys)	Framework study
I8	BUILDCHAIN	Data-model proposal
I9	OneClickRENO	Data model + tool
I10	Vabi EPA	EPC software (NL)
I11	BuurtWarmteWijzer	Decision-support tool (NL)
I12	CLÉA	Implementation (France)
I13	iSFP	Implementation (Germany)
I14	Hoomdossier	Implementation (NL)

Table 1. The 14 assessed Digital Building Logbook and Building Renovation Passport initiatives.

2.3 CityGML and the Energy ADE as candidate data model

CityGML is an open OGC standard for the representation, storage, and exchange of semantic 3D city models, defined as a UML conceptual model on the ISO 191xx series. Its defining feature is the coupling between geometry and semantics: for example, `WallSurface`, `RoofSurface`, and `GroundSurface` are semantically specified elements (thematic surfaces) that compose the hull of a building, each also possessing a geometry. Three mechanisms make it particularly relevant to a BRP data model: Application Domain Extensions (ADEs) allow to define and add new classes and properties without breaking core conformance; `ExternalReferences` allow additional, optional identifiers from external registers to be attached to any `CityObject`; and the possibility, inherited from GML, to associate units of measure to specific values whenever necessary, as is the case, for example, of a U-value or an annual energy demand. Most of what a passport records is attribute data rather than geometry, but a (geo)spatial model is nevertheless useful, as it offers the optional possibility to also include geometries. Every attribute attaches to a building object of a defined class, inheriting a stable identity, a clear spatial hierarchy from dwelling to neighbourhood, and a consistent system of units. The optional integration of (geo)spatial data reflects current trends in the digitalisation of the built environment, in which large-scale semantic 3D city models and BIM models (mainly for selected new buildings) steadily enrich the available building data.

The Energy ADE enriches CityGML with the data needed for urban building energy modelling, and has an established ten-year history of use for that purpose. The latest public draft, Energy ADE 3.0 beta 8 (as of May 2026), is the implementation target of this work. Although CityGML 3.0 is the latest release, the Energy ADE is still based on CityGML 2.0. In practice, CityGML 2.0 remains the most used version worldwide: the GML exchange encoding of CityGML 3.0 was only adopted in 2023, mainstream tooling has begun supporting it only recently, and virtually all large open city models published to date are still provided in CityGML 2.0 (Kutzner et al., 2020). A description of the current development status of the Energy ADE 3.0 is beyond the scope of this paper; information can be found in (Agugiaro and Padsala, 2025), as well as in the resources available on the project's GitHub repository. The current Energy ADE 3.0 contains classes such as `Intervention` (to account for Inspection, Maintenance and Refurbishment operations), `EnergyPerformanceCertificate`, together with modules to model thermal zones, devices (e.g. solar panels), and layered constructions and their physical properties, such multi-layered walls and roofs.

The Energy ADE has been used for urban energy modelling in several national and international projects (Agugiaro and Padsala, 2025); semantic 3D city models have likewise been populated from open data at national scale, for example in the Netherlands (León Sánchez, 2025). This prior work focuses, however, on energy simulation, not the renovation passport. What has not been tested is whether the schema covers, field by field, the data that a BRP must record, and whether that coverage holds once the model is populated from open building-stock data.

3. Methodology

The research is organised as a conceptual analysis followed by two implementation tests. The conceptual analysis distils the MSRD from the 14 surveyed initiatives and maps it onto the candidate data model. The first test applies the resulting mapping in depth to one owner-occupier reference building; the second applies the mapping in breadth on a city scale using an ingestion of the open Dutch registers.

3.1 Deriving the MSRD

For each initiative, we extracted the stated data requirements, adapting the analysis each time to the specific initiative. These requirements, typically already organised into the initiative's own thematic groups, were decomposed into individual attributes. Many initiatives lack a formally specified schema; these structures, therefore, had to be reconstructed from a wide range of sources, such as Excel templates, user interfaces, project deliverables, interviews with stakeholders and even YouTube videos, for lack of better-documented sources.

The attributes from all 14 initiatives were then laid side by side and semantically aligned. Fields naming the same concept under different terms, data types, or levels of detail were merged into a single entry. The aligned attributes were grouped into thematic modules, with each module emerging where the same kind of attribute recurs across initiatives, such as a dimensions module or an envelope-and-construction module. The 13 resulting modules were in turn grouped into three conceptual layers to capture the logical dependency chain of a BRP: a *repository layer* of fundamental building information, an *assessment layer* of performance indicators, and a *functional layer* of renovation-related outputs. This layering reflects the fact that a building cannot be assessed without the underlying repository data, and a renovation roadmap cannot be advised without both the data and the assessment. The outcome is a unified comparison grid in which each attribute is scored for inclusion across the 14 initiatives. The inclusion count and frequency then feed the relevance filter (Figure 1 shows this grid for the identification and administrative-details module; the grids for all 13 modules are given in (Schlosser, 2026a)).

Starting from the comprehensive, organised list of all attributes gathered from the 14 initiatives, the MSRD was determined by assessing each field against an empirically defined relevance filter. This filter draws on three specific inputs: (i) the frequency with which an attribute recurs across initiatives, treated as a relative signal rather than a hard threshold; (ii) the regulatory mandate set out in EPBD Annex VIII; and (iii) practitioner findings from a parallel study that interviewed 17 respondents across 9

Feature	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	Total	%
Building address details	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	**14**	**100%**
Building identification (Cadastral/INSPIRE)	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			**11**	**79%**
Ownership details and occupancy details	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	**13**	**93%**
Geospatial Coordinates (latitude, longitude, altitude)	✓	✓	✓	✓	✓	✓	✓								**7**	**50%**
DBL author/ assessor qualifications and information	✓	✓		✓		✓		✓	✓	✓			✓	✓	**9**	**64%**
Details on the building with building-unit relation			✓	✓	✓		✓		✓	✓	✓				**7**	**50%**
Timestamps on collection, provision or modification**			✓	✓		✓		✓	✓	✓			✓		**7**	**50%**
Building owner address details			✓							✓		✓	✓	✓	**5**	**36%**

Figure 1. The cross-initiative comparison grid for the identification and administrative-details module: each attribute (row) is scored for inclusion across the 14 initiatives (I1–I14), and the inclusion count and frequency (right-hand columns) feed the relevance filter.

stakeholder categories on which Dutch building-data sources they consult and trust (Schlosser, 2026b). These inputs are consolidated through a reasoned judgement of BRP applicability rather than a single automatic rule. For each field, the relevance designation reflects where the three inputs converge or diverge.

3.2 Mapping to the candidate data model

We systematically mapped the MSRD onto the CityGML 2.0 + Energy ADE 3.0 model, covering all classes, codelists, enumerations, and their relations. For each field, we recorded whether it could be represented and, if so, by which class, attribute, or association. Each field was assigned one of two coverage outcomes: *covered*, when the existing schema already accommodates it, or a *gap*, when it does not and a targeted addition (a new class, attribute, association, or codelist) would be needed. (Figure 2 shows the result for one module; the full per-module mapping and a classification of each gap by type are given in (Schlosser, 2026a)). While the found gaps constitute the findings of this research, the targeted improvements proposed to close them are framed as input to the ongoing development of the Energy ADE.

3.3 Two-scale implementation test

Two implementation tests applied the candidate target model to Dutch data. The *per-building* (audit-depth) test used one owner-occupier dwelling. We used the ISSO 82.1 intake protocol that Dutch energy-performance advisors must follow as a checklist for what to collect. This ensured that no thematic area was overlooked. The data came from building plans and detail drawings, the existing energy label, metered consumption and production, and conversations with the resident. The intake deliberately stayed within realistic limits to mirror what an energy-performance advisor actually collects in practice. The dwelling was visited once, the roof was not climbed, and no specialised equipment was used. The data were encoded in a CityGML 2.0 + Energy ADE 3.0 GML document covering the dwelling's zones, schedules, devices, layered constructions, and a material library.

The *city-scale* (breadth) test populated the same schema from open data across whole municipalities. The evaluation rests

Feature	%	MSRD	Module	Class	Attribute	Type
Building address details	**100%**	Yes	Core*	Address	Multiple	Multiple
Building identification (Cadastral/INSPIRE)	**79%**	Yes	Core	_CityObject	identifier	Code [0..1]
Ownership details and occupancy details	**93%**	Yes	Building Building	_AbstractBuilding _AbstractBuilding	bdgOwnerName bdgOwnershipType	CharacterString [0..1] OwnershipTypeValue [0..1]
Geospatial Coordinates (latitude, longitude, altitude)	**50%**	Yes	Core	_CityObject	referencePoint	GM_Point [0..1]
DBL author/ assessor qualifications and information	**64%**	Yes	DataTypes	Metadata	Multiple	Multiple
Details on the building with building-unit relation	**50%**	Yes	Building	BuildingUnit	N/A (done by composition)	
Timestamps on collection, provision or modification**	**50%**	Yes	Core* Core*	_CityObject _CityObject	creationDate terminationDate	Date [0..1] Date [0..1]
Building owner address details	**36%**	No	✗			

* CityGML class; unmarked Module entries belong to the Energy ADE.

Figure 2. Field-by-field mapping of the identification and administrative-details module onto CityGML 2.0 + Energy ADE 3.0 (beta 8): for each attribute, its frequency across initiatives, its inclusion in the MSRD, and the target class, attribute, and data type.

on a small area in Emmer-Compascuum (about 41.5 ha, ≈950 buildings), selected because it is one of the few areas where all the thematic layers used in this work are jointly available: it is specifically targeted by the open solar-panel dataset, and it is covered by the AHN6 elevation data required for the tree reconstructions. The area integrates multiple thematic layers (buildings, energy labels, topography, neighbourhood statistics, solar panels, and trees), assembled from open Dutch national datasets: the addresses-and-buildings register (BAG), the national 3D building models (3DBAG; LoD0–LoD2 geometries), the energy-label register (EP-Online), the large-scale topographic base map (BGT), neighbourhood statistics at six-digit postcode level (CBS Postcode-6), open solar-panel footprints (Luimstra, 2025), and individual trees (CFTree) (Alting, 2025). The solar and tree layers appear only in this test area.

4. Results

4.1 The Minimum Set of Required Data

Of the 352 distinct fields compared across the 14 initiatives, 275 are included in the MSRD. The rest are excluded as out of scope for a BRP, too granular to act on, too subjective to record as structured data, or already derivable from other required fields (for instance, a count of identical openings, which the surface and opaque areas already convey for an energy assessment). The 275 fields are organised into 13 thematic modules and the three conceptual layers.

4.2 Coverage benchmarking

Benchmarked against the MSRD, the CityGML 2.0 + Energy ADE 3.0 (beta 8) schema provides a mapping target for 262 of the 275 MSRD fields, roughly 95% (Table 2, Figure 3). Coverage is complete for 8 of the 13 modules, including building identification, general building characteristics, documents, dimensions, and ancillary data. The remaining modules are well accommodated: technical systems (97%), legal and financial data (96%), operation and use (92%), and envelope and construction (93%). The one substantial shortfall occurs where the

Module	Comp.	MSRD	Cov.	Impr.
1. Identification & admin.	8	7	7	0
2. General building char.	24	19	19	0
3. Envelope, constr. & mat.	60	43	40	3
4. Technical systems	30	29	28	1
5. Attachable documents	4	4	4	0
6. Dimensions	19	19	19	0
7. Operation & use	13	12	11	1
8. Legal & finance	34	25	24	1
9. Ancillary data & context	25	11	11	0
10. Energy performance	17	17	17	0
11. Indoor env. quality	46	26	26	0
12. Smart readiness	45	37	37	0
13. Renovation advice	27	26	19	7
Total	**352**	**275**	**262**	**13**

Table 2. The MSRD across the 13 modules. *Comp.*: distinct fields compared across the 14 initiatives; *MSRD*: fields taken in by the relevance filter; *Cov.*: MSRD fields with a mapping target in the schema; *Impr.*: fields requiring a proposed improvement.

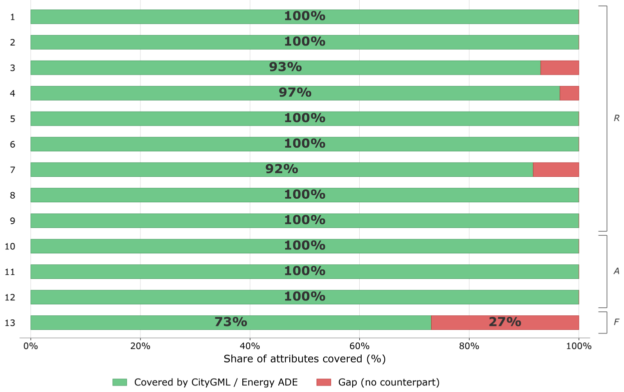


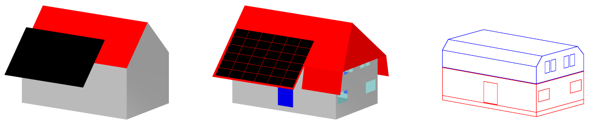
Figure 3. Coverage of the relevance-filtered MSRD by CityGML 2.0 + Energy ADE 3.0 (beta 8), per module, grouped into the three conceptual layers: Repository (R, modules 1–9), Assessment (A, modules 10–12), and Functional (F, module 13).

BRP concept is most novel: the renovation-advice module that carries the staged roadmap has a coverage of 73%.

The 13 uncovered fields are concentrated in a few thematic modules rather than spread across the schema: seven are in the renovation-advice module, three in envelope and construction, and one each in technical systems, operation and use, and legal and finance data. Twelve are addressed as targeted improvement proposals (Section 5.1). The one remaining field—public subsidies and funding programmes—is deliberately left as an open question. The obstacle here is not a simple missing attribute, but the immaturity of the Energy ADE's overall cost model. Currently, financial data is only handled via the Resources module, which is designed for physical flows and carries only basic expense and revenue values. Properly representing complex financial instruments like subsidies requires a broader structural redesign of how the standard handles economics, making it an open design question for future development rather than a quick targeted fix.

4.3 Per-building implementation at audit depth

The per-building pipeline produced a single CityGML 2.0 + Energy ADE 3.0 document, `NL-single-family-house.gml`,



(a) LoD2 (b) LoD3 (c) Thermal zones

Figure 4. The owner-occupier reference building visualized in the KITModelViewer. It carries all four exterior levels of detail (LoD0 to LoD3) in one document; (a) and (b) show it at LoD2 and LoD3, and (c) shows the thermal model as a wireframe, with the two heated storeys carried as two `nrg3:ZonePart` instances in distinct colours, red (lower) and blue (upper).

which validates against the Energy ADE 3.0 (beta 8) XSDs. The document carries one `bldg:Building` with one `nrg3:BuildingUnit`. The four exterior levels of detail (LoD0 footprint, LoD1 solid, LoD2 thematic surfaces, LoD3 surfaces with 13 openings) are all present on that one building (Figure 4). The thermal model is one `nrg3:Zone` divided into two `nrg3:ZonePart` instances, one per heated storey, each carrying its heating and cooling setpoints (Figure 4c). The envelope is modelled surface by surface, as 42 thematic boundary surfaces. Each surface carries its own physical properties and is built from 12 layered constructions that reference a shared 22-entry material library. The same document also holds the technical systems (a ground-source `nrg3:HeatPump` with its distribution, storage, and EV charging), the rooftop `nrg3:PhotovoltaicCollector`, the occupants, and the energy resources. All of these populate the schema together without conflict.

Three capabilities of the data model, deliberately exercised by this test, stood out. First, the schema allows a single document to hold the building at every level of detail and at full audit depth at once. Second, the same file stores measured and modelled data side by side: a metered house-consumption series, a metered photovoltaic-production series, and an independent NTA 8800 simulated production series for the same array, all as `nrg3:MonthlyTimeSeries`. The measured generation can then be read directly against the modelled estimate (Figure 5). The two series end roughly 20% apart. The point of this demonstration is the side-by-side comparison itself. Third, every city object can carry `validFrom` and `validTo` dates that record the time extent of its validity. This lets the model hold a building's history, not just its current state. When a component is replaced, for example a gas boiler by a heat pump, the boiler is kept with an end date rather than deleted, and the heat pump is added with a start date. Past and present components then coexist in the same model, which is exactly the kind of renovation history a passport accumulates over a building's life.

Building the file raised two issues. The first resolves within the schema. The photovoltaic array sits on the roof, but the roof appears twice in the file: as the LoD2 surface and as the LoD3 surface. These are separate objects with different identifiers. The Energy ADE's `installedOn` relation links the array to the surface it sits on, yet a single relation can target only one of these objects. It therefore cannot mean "on the roof" across both levels of detail at once. One viable solution is to specify the relation as one `installedOn` for each roof surface the array covers in each representation.

The second is a finding about the schema. The same quantity can be stored more than once: the dwelling's usable floor area is 122.0 m^2 in the BAG register but 104.2 m^2 as measured

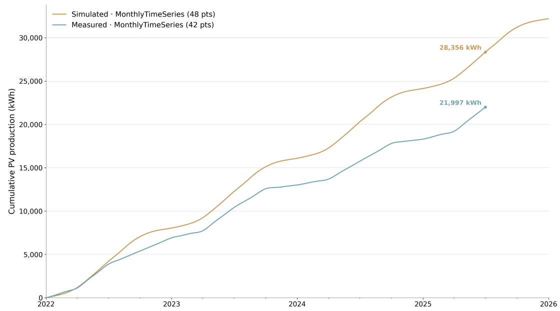


Figure 5. Cumulative metered photovoltaic production against the independent NTA 8800-simulated production for the owner-occupier array, both carried as `nrg3:MonthlyTimeSeries` in the same document.

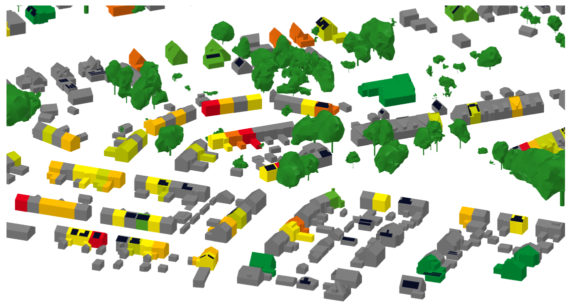


Figure 6. The integrated Emmer-Compascuum small-area scene: 3DBAG geometry, EP-Online energy labels (building colour), rooftop solar, and CFTree vegetation, assembled on a single CityGML model (KITModelViewer).

from the LoD3 geometry. Because `nrg3:QualifiedArea` can repeat on the same object, both values coexist, each tagged with its own `nrg3:source`. What the schema does not yet allow is richer, structured metadata per value, such as the measurement method, responsible party, or date. A small improvement addressing this is proposed in Section 5.1.

4.4 City-scale implementation

To demonstrate that the data model can scale to any desired extent, the city-scale pipeline was populated automatically from national open data across five test areas. These included the small area in Emmer-Compascuum integrating multiple thematic layers (Section 3.3; Figure 6) and four full-extent municipality runs (Emmer-Compascuum, Delft, Groningen, and Zwolle; Figure 7 shows the Delft run). All five validated successfully against the Energy ADE 3.0 (beta 8) XSD file. The pipeline confirms that the mapping holds regardless of the size of the implementation.

Each source mapped onto the schema with most fields landing in native CityGML or Energy ADE slots. From BAG, each `pand` became a `bldg:Building` and each `verblijfsobject` a `nrg3:BuildingUnit`. The 3DBAG tiles supplied the LoD0–LoD2 geometries. Storey count was mapped to the native `bldg:storeysAboveGround` attribute, while roof height and volume were carried as Energy ADE qualified attributes: `bldg:Building` offers only a single, unqualified `bldg:measuredHeight` and no volume attribute, whereas the qualified attributes can repeat and carry an explicit height reference and source. Per-surface thickness and heat capacity were left empty, because the open registers provide no information about layered constructions. EP-Online labels became

Figure 7. The municipality of Delft: building footprints coloured by the average EP-Online energy label of the building units they contain (green best, red worst, grey no label).

`nrg3:EnergyPerformanceCertificate` instances. Here, the support for units of measure proved its value: EP-Online mixes per-area NTA 8800 figures and legacy absolute totals in one column, but because every `nrg3:Energy` resource carries its `uom`, downstream calculations are prevented from erroneously aggregating fundamentally different metrics. CBS Postcode-6 statistics became `nrg3:UrbanFunctionArea` aggregates, rooftop panels `nrg3:GenericSolarCollector` instances, and trees `veg:SolitaryVegetationObject`. Fields with no native home but a documented future use were mapped to CityGML generic attributes rather than dropped. These include the EP-Online building subtype and renewable-energy share, the CBS dwelling counts, and secondary tree metrics. This avoids data losses where a specific Dutch field lacks a native energy-focused equivalent.

4.5 Cross-cutting analysis

Bringing the Dutch sources together in one model does more than confirm fit; it makes the data quality of the stock visible. The integrated model places the values recorded by different registers on the same building, each tagged with its source, so a cross-register comparison becomes a direct query on one model.

Year-of-construction disagreement. Across all four areas, BAG and EP-Online agree exactly on the year of construction only about 60% of the time, and agree to within a few years about 86% of the time (Figure 8). That two authoritative national registers differ on a basic, supposedly objective fact for a large share of the stock is itself a result. The integration makes the disagreement systematically visible and addressable. Part of this disagreement can be explained by the different definitions used by the two registers. The BAG records the year in which a building was originally completed; this year is never updated, even after a major renovation or extension (Kadaster, n.d.a). The year stored with an energy label is recorded by the energy assessor following the ISSO 82.1 recording protocol. This protocol prescribes a fixed order of sources: the year stated on the building-permit application; if this is not available, the year in which the permit was granted; and only if both are unknown, the completion year registered in the BAG, from which the assessor may deviate only when that year is demonstrably incorrect (ISSO, 2021, Kadaster, n.d.b). Because a permit precedes completion, the recorded year can differ from the BAG year by a few years even when both entries are correct. In addition, some buildings consist of parts built in different periods, and some addresses are linked to several BAG buildings; in these cases a single representative year is recorded. The two registers therefore do not record the same information, which explains part of the observed mismatches.

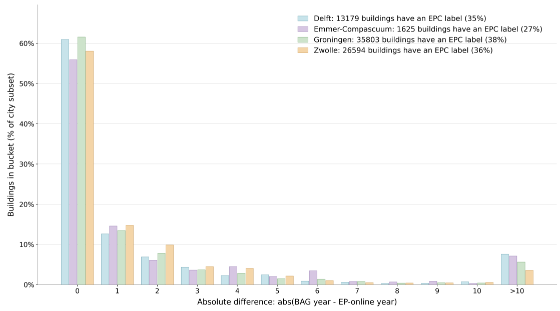


Figure 8. Agreement between the BAG and EP-Online year of construction across the four areas: share of buildings whose recorded years agree exactly, within one year, within three, and beyond.

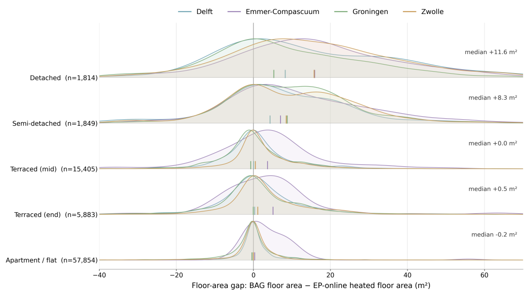


Figure 9. Signed floor-area gap (BAG usable area minus EP-Online heated area, m^2) per building typology, with the four areas overlaid; the right-hand value is the median gap across the four areas combined.

Placeholder and rounding artefacts. Part of the disagreement is structural. One year, 1930, is over-represented in EP-Online, recorded for 7% to 12% of labelled buildings in each area. A genuine building "boom" would show in BAG as well, but BAG dates those same buildings across several centuries, so 1930 is most likely a placeholder rather than a real construction year. A milder version is rounding. About a quarter of EP-Online construction years end in zero, far more than among the BAG and 3DBAG years for the same buildings. Because both registers describe the same stock, this excess of round years in EP-Online points to rounding rather than to a genuine clustering of construction dates; even where EP-Online gives a year it is often a rounded estimate.

Floor-area gap varies by building type. BAG and EP-Online also disagree on floor area, and the integration shows the gap follows a clear pattern rather than random noise (Figure 9). BAG records a usable floor area per building unit, while EP-Online records the heated floor area its label is computed on. The integration reveals a positive median gap of roughly 12 m^2 for detached dwellings, alongside negative gaps for other properties. The exact causes remain hypotheses rather than verified facts; even if these specific explanations are ultimately adjusted, the integrated model proves its value by making this systematic discrepancy visible and measurable.

What the integration enables. The same integration also enables two things that no single source provides on its own. First, the dwelling-to-certificate link is built into the model as an explicit relation, rather than a join across separate registers that each analysis must redo; about two-thirds of the building units carry such a link. Second, five independent data themes (geometry, energy performance, rooftop solar, vegetation, and neighbourhood aggregates) come together on one CityGML model, attached to the same buildings through shared identifiers and spatial relations rather than in separate disjoint files. The same content was also loaded into a 3DCityDB database, via the 3DCityDB Importer/Exporter and an FME Workbench from the Energy ADE 3.0 repository. The populated model can therefore be held as relational tables as well as a GML file, two equivalent lossless enconding of the same conceptual model.

5. Discussion

This paper investigated whether CityGML 2.0 + Energy ADE 3.0 (beta 8) can serve as the data model for a Building Renovation Passport, along two lines. Structurally, the conceptual mapping (Section 4.2) leaves only 13 of the 275 MSRD fields without a target, with 8 of the 13 modules covered in full, including the whole assessment layer. Empirically, the implementation tests show that the schema held the data at both scales, with the per-building file and the five city-scale runs all validating. The answer is therefore yes: each of the 13 remaining gaps is a specific omission that a targeted addition can close, not a limitation of the model's structure.

5.1 Proposed improvements

The major gap is the renovation-advice module. The schema does address renovation to some extent: the `nrg3:Intervention` feature records a single renovation measure with an action codelist, type, and start/end dates, and a single measure maps cleanly. What the schema does not provide is the roadmap *as* a roadmap: the ordering of measures into a sequence, the per-step descriptors of renovation depth, priority, and motivation, the scheduling attributes (estimated duration, best opportunity), and the expected per-step change in energy demand and cost. Two suggestions may close the current gap. The first, simpler approach stays inside the Energy ADE, treating the roadmap as an ordered set of `nrg3:Intervention` features with the missing per-step attributes and a container class that groups the steps into one named passport. The second approach, more complex but more powerful, could adopt the Scenario ADE (Auer et al., 2026) (also in development). The Scenario ADE adds scenario and simulation modules, and models a roadmap as a base version, a scenario referencing it, and a per-step transition. We recommend building the first, queryable version of the roadmap on the existing `nrg3:Intervention` class, and reusing the Scenario ADE rather than defining a competing class. Two design questions remain open: how fine-grained a step should be (per-component or per-building), and how a step should link to the building element it modifies.

The remaining proposals are smaller. The three envelope gaps (two damage-observation fields, an opening-level airtightness class) suggest a small set of codelist-based condition attributes on the construction and surface classes and an airtightness-class attribute on the opening. The single technical-systems gap is a codelist-based control-and-automation attribute on the relevant system classes. The single operation gap is a free-text auditor-recommendation attribute, which may instead belong on the renovation step it concerns if the roadmap is modelled through interventions. The audit-depth metadata finding suggests adding a `gml:id` to `nrg3:Metadata` so a source record for each value can be referenced from the value it describes, a small change with a wide effect. Finally, the output audit found a codelist gap: `HeightTypeValue` enumerates only absolute-elevation reference points and has no member for a height measured above ground level, so a `maxHeightAboveGround` member is proposed. More broadly, in the case of the Netherlands, we suggest a systematic review of the codelists against the values a Dutch implementation actually uses.

5.2 Implications and limitations

The proposed improvements mark where the regulatory definition of a BRP goes beyond what the existing open data model covers. The renovation-roadmap gap is exactly the part the EPBD recast emphasises, and the proposals therefore serve as recommendations to the policy and standardisation communities alike. On the Dutch side, the city-scale pipeline already exercises the open data that the national *Landelijke Voorziening Gebouwgegevens* (LVG) is shaping. The data-quality patterns reported here indicate where a national facility joining these sources must reconcile them.

A first limitation concerns the 14 reviewed initiatives. Despite being heterogeneous in scale and scope, they are not exhaustive of European developments. The selection was driven by what was openly available, so the MSRD is the field set this comparison yields. Reconstructing undocumented models from spreadsheets and interfaces also carries its own uncertainty. The coverage figure is exact only against this paper's own MSRD, which was reconstructed from the initiatives and narrowed by a discretionary relevance filter. While the EPBD recast mandates the high-level contents of a BRP, no external, authoritative reference for the specific field-level data requirements exists to measure it against. The exact composition of the MSRD therefore relies on the authors' reasoned judgement to translate those broad mandates, alongside the surveyed initiatives and practitioner needs, into a granular data model. The implementation tests used Dutch building-stock data only, so the EU-broad data-model claim is stated as one credible case rather than a settled result. The mapping is held against the beta 8 release and may need revision if a later beta or the 3.0 final changes the schema.

6. Conclusion

This paper tested whether an open, formally specified geospatial data model can serve as the data model for a Building Renovation Passport. It derived a Minimum Set of Required Data of 275 fields, organised into 13 modules and three layers, from 14 European DBL/BRP initiatives and benchmarked that set against CityGML 2.0 + Energy ADE 3.0 (beta 8), both conceptually and on real Dutch building-stock data at two scales. The schema covers roughly 95% of the MSRD, with 8 of 13 modules fully covered, and the two-scale implementation confirms that the mapping holds under real data: the schema produced XSD-valid output ranging from a single audited dwelling to large-scale extents. As a result of building upon CityGML, the data model can easily scale up to hold massive amounts of building data for entire municipalities or even regions, with practically no structural limit when deployed within a spatial database. The 13 uncovered fields do not indicate a structural shortcoming of the schema: they are concentrated in a few modules, most of them in the renovation-advice module that carries the staged roadmap, the BRP's defining feature, and each can be addressed with a targeted improvement. Bringing the Dutch open-data sources together in one data model also made the data-quality problems of that stock visible and tractable, exposing different interpretations of the construction year, placeholder and rounding artefacts, and different methods to measure the floor area, which the source datasets, read apart, do not reveal.

CityGML and Energy ADE 3.0 as an open, formally specified data model already covers most of what a BRP needs, and the remaining gaps can be closed by the minor improvements proposed here. It is therefore a near-complete starting point for a BRP data model. Once Energy ADE 3.0 is finalised, we plan to port the model to CityGML 3.0. This will enable further testing against other datasets and prepare the approach for the gradual migration of existing datasets to the latest version of the standard.

ACKNOWLEDGEMENTS

This publication is part of the project Accelerating Building RENOvation and Decarbonization Through DATa Integration (RenoDat) with file number KICH1.ED09.23.003 of the research programme KIC - MISSION Digitalisation of Energy Renovations in the Built Environment 2023 which is (partly) financed by the Dutch Research Council (NWO) under the grant ID https://doi.org/10.61686/XQBEG97133.

References

Agugiaro, G., Padsala, R., 2025. A proposal to update and enhance the CityGML Energy Application Domain Extension. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W6-2025, 1–8. https://doi.org/10.5194/isprs-annals-X-4-W6-2025-1-2025.

Alting, N., 2025. From Point Clouds to Porous Crowns: A Scalable Approach for CFD-Ready Urban Tree Reconstruction. MSc thesis, Geomatics, Delft University of Technology. https://resolver.tudelft.nl/uuid:adc9c299-b8f7-4854-b004-dcf2393d06fb. Code: CFTree, https://github.com/NoahAlting/CFTree.

Auer, F., León-Sánchez, C., Agugiaro, G., 2026. A glimpse of the CityGML Scenario Application Domain Extension. *Proceedings of the 5th 4TU/14UAS Research Day on Digitalization of the Built Environment*, Enschede, The Netherlands.

Buchholz, M., Lützkendorf, T., 2023. European building passports: developments, challenges and future roles. *Buildings & Cities*, 4(1), 902–919. https://doi.org/10.5334/bc.355.

European Commission, 2020. A Renovation Wave for Europe - greening our buildings, creating jobs, improving lives. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2020) 662 final, European Commission, Brussels. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0662.

European Parliament, 2024. Directive 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast) (Text with EEA relevance). http://data.europa.eu/eli/dir/2024/1275/oj.

Gómez-Gil, M., Karami, S., de Almeida, J.-P., Cardoso, A., Espinosa-Fernández, A., López-Mesa, B., 2024. Envisaging a European Digital Building Renovation Logbook: Proposal of a Data Model. *Applied Sciences*, 14(19), 8903. https://doi.org/10.3390/app14198903.

Gómez-Gil, M., López Mesa, B., Espinosa Fernández, A., 2022. European Digital Building Logbook: definition of its functionalities to maximize its potential. *EESAP13 International Conference 2022, 5-6 October Donostia-San Sebastián: Akten liburua - Libro de actas - Proceedings book*, 63–71. https://dialnet.unirioja.es/servlet/articulo?codigo=8979984.

Hartenberger, U., Ostermeyer, Y., Lützkendorf, T., 2021. The Building Passport: A Tool for Capturing and Managing Whole Life Data and Information in Construction and Real Estate. Technical report, Global Alliance for Buildings and Construction. https://globalabc.org/resources/publications/building-passport-tool-capturing-and-managing-whole-life-data-and.

ISSO, 2021. *ISSO-publicatie 82.1: Energieprestatie woningen en woongebouwen. Methode 2020, opnameprotocol*. 6 edn, ISSO, Rotterdam.

Kadaster, n.d.a. Praktijkhandleiding BAG: Oorspronkelijk bouwjaar. https://imbag.github.io/praktijkhandleiding/attributen/oorspronkelijk-bouwjaar.

Kadaster, n.d.b. Praktijkhandleiding BAG: Welke rol speelt het bouwjaar bij het toekennen van een energielabel? https://imbag.github.io/praktijkhandleiding/artikelen/welke-rol-speelt-het-bouwjaar-bij-het-toekennen-van-een-energielabel.

Karami, S., de Almeida, J.-P., Ellul, C., Cardoso, A., 2025. Semantic mapping-based comparison of digital building LogBook and energy performance certificate models - envisaging data reuse. *Developments in the Built Environment*, 22, 100674. https://doi.org/10.1016/j.dibe.2025.100674.

Kutzner, T., Chaturvedi, K., Kolbe, T. H., 2020. CityGML 3.0: New Functions Open Up New Applications. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(1), 43–61.

León Sánchez, C. A., 2025. Enhancing urban energy applications through semantic 3D city models and open data. PhD thesis, Delft University of Technology.

Luimstra, G., 2025. Annotated High-Resolution Aerial Imagery of the Dutch Landscape for Solar Panel Detection and Segmentation. https://zenodo.org/records/14860030. Presented at ESANN 2025, the 33rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.

Malinovec Puček, M., Khoja, A., Bazzan, E., Gyuris, P., 2023. A Data Structure for Digital Building Logbooks: Achieving Energy Efficiency, Sustainability, and Smartness in Buildings across the EU. *Buildings*, 13(4), 1082. https://doi.org/10.3390/buildings13041082.

Schlosser, D. H., 2026a. Investigation on the data model requirements for Building Renovation Passports. MSc thesis, Geomatics, Delft University of Technology. https://resolver.tudelft.nl/uuid:fe05529d-fe1c-49e4-8d98-1a0acb2bfa93.

Schlosser, D. H., 2026b. Towards a Dutch Building Renovation Passport: Data Needs and Expectations in Practice. MSc thesis, Management in the Built Environment, Delft University of Technology. https://resolver.tudelft.nl/uuid:0374afe3-c411-408e-b5ce-486831e5033d.