

Article

Open Data-Driven Reconstruction of Power Distribution Grid: A Land Use-Based Approach

Mohannad Babli ^{1,2,*} , Tobias Gebhard ³  and Eva Brucherseifer ^{1,2} 

¹ Department of Computer Science, Darmstadt University of Applied Sciences, Schöfferstraße 8, 64295 Darmstadt, Germany

² European University of Technology, European Union

³ Institute for the Protection of Terrestrial Infrastructures, German Aerospace Center (DLR), Mornewegstraße 30, 64293 Darmstadt, Germany

* Correspondence: mohannad.babli@h-da.de

Abstract: Disruptive events and the rapid evolution of urban energy systems highlight the need for robust methods to reconstruct critical infrastructure networks. Comprehensive, up-to-date power grid representations are essential for both researchers developing methods for analysing and optimising power systems and first responders requiring approximate data for urgent decisions. However, traditional grid reconstruction approaches often rely on incomplete data, expert knowledge, or closed datasets, limiting their utility during emergencies. This study proposes a novel automated method for reconstructing medium-voltage (MV) power grids. The novelty of the proposed method lies in combining OpenStreetMap energy and land-use data in a unified and automated framework, thereby reducing the need for expert input. The proposed method employs a systematic aggregation of data, an estimation of energy demand, and the application of algorithmic techniques to generate synthetic MV grid models that functionally represent real networks, capturing key topological features. The resulting outputs include visual representations to support decision-makers in simulating "what-if" scenarios and ensuring rapid operational awareness. In a step toward eliminating reliance on proprietary data, our approach broadens access to critical infrastructure insights across diverse urban contexts, contributing to critical infrastructure resilience and potentially supporting both energy system research and crisis management. A case study demonstrates that a medium-sized city's MV grid can be reconstructed in minutes without expert knowledge or geographically constrained datasets, underscoring the method's deployment potential and practical value for emergency scenarios.

Keywords: automated distribution grid reconstruction; critical infrastructure; OpenStreetMap; land use; crisis management



Academic Editor: Jen-Hao Teng

Received: 26 February 2025

Revised: 27 March 2025

Accepted: 28 March 2025

Published: 31 March 2025

Citation: Babli, M.; Gebhard, T.; Brucherseifer, E. Open Data-Driven Reconstruction of Power Distribution Grid: A Land Use-Based Approach. *Electronics* **2025**, *14*, 1414. <https://doi.org/10.3390/electronics14071414>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Critical Infrastructures (CIs) are indispensable for daily life yet remain vulnerable to crises such as natural disasters or targeted attacks, which can disrupt vital services like energy and transportation [1–3]. Because these events often occur without warning, the lack of reliable, up-to-date infrastructure data hinders rapid crisis management, particularly in municipalities with lower levels of digitisation that lack comprehensive digital representations of their CIs.

Although open geographic data from crowdsource and community-driven projects like OpenStreetMap (OSM) can inform risk management [4], their varying completeness and consistency pose significant challenges to effective decision making. Consequently,

emergency services, healthcare facilities, and utilities can face operational delays that jeopardise public safety and incur substantial economic losses during crises [5,6].

It is imperative for a wide range of stakeholders to bridge the data gaps in the aforementioned model to reconstruct an approximate representation of existing power grids. Such stakeholders include those who need to assess disaster impacts on grid topology, make infrastructure decisions when up-to-date grid models are unavailable, or simulate cascading failure propagation and rebalancing of power systems [7]. However, grid-mapping approaches often rely on expert knowledge, proprietary data, or manual intervention, which limits their scalability in dynamic emergency scenarios [8–10]. In such situations, even an approximate grid representation can be indispensable for rapid decision making.

Recent advancements in data integration and Digital Twin (DT) technologies [11,12] offer potential solutions to enhance CIs resilience. A DT is bidirectionally coupled through a twinning mechanism to ensure the virtual replica is kept up-to-date. Utilising the replica, related tools can implement smart functionalities, simulation, and control actions on the real object, aligning with the aforementioned paradigm. This study contributes to crisis management and infrastructure resilience in the preparing and reacting phases. Building on our earlier work [11], we propose an automated toolbox for reconstructing the Medium Voltage (MV) power grid, based solely on OSM.

The ubiquity of the open-source geospatial data provided by OSM renders it a preeminent resource for researchers in the field of distribution grids [10,13–15]. Its extensive global coverage and standardised format make it well suited for distribution grid research. The database, which is collaboratively maintained, includes geo-located information on roads and power network elements, often derived from aerial imagery, on-site surveys, or other openly licensed data. Although completeness varies by region, two main advantages make OSM particularly valuable for grid reconstruction. Firstly, it is actively maintained and continuously expanding, ensuring the longevity of tools that rely on its evolving dataset. Secondly, its worldwide coverage supports broader applicability, enabling methods to scale across diverse urban contexts. The present study introduces an approach centred on open data, leveraging existing power infrastructure and land-use data from OSM to estimate and augment missing grid elements. By eschewing proprietary or expert-restricted sources, this method provides a scalable solution for reconstructing MV power grids across multiple urban environments.

The primary contributions of this work are as follows:

- Firstly, a fully open-data automated methodology for estimating urban energy demand and reconstructing MV power distribution grids is presented. This method utilises solely publicly available OSM data.
- Secondly, the elimination of reliance on proprietary or geographically specific datasets is proposed, thus ensuring generalisability, broad applicability and scalability across various urban contexts globally, significantly expanding the potential use cases, especially in crisis scenarios.
- The innovative integration of power and land-use data from OSM has been utilised to estimate and identify missing secondary substations systematically, resulting in more accurate reconstruction results. This approach differs from previous methodologies, which used either OSM power or land-use data alone without combining them, required external expert input, or relied on incomplete OSM infrastructure mapping.

Specifically, the following steps are taken: (i) Retrieve and preprocess publicly available geospatial information. (ii) Estimate power demand and identify gaps in secondary substations via a land-use-based approach. (iii) Position missing secondary substations. (iv) Employ a constraint-programming solver to produce realistic MV networks. (v) Generate layered visual outputs for rapid interpretation and “what-if” simulations. (vi) The

output consists of geo-referenced MV grids that closely follow street infrastructure, resulting in highly detailed topology rather than simple straight lines. To the best of our knowledge, no existing approach offers this level of automation in reconstructing grid models by solely utilising data from OSM without expert input or geographically constrained datasets. This research distinctly advances previous methodologies by systematically automating previously manual tasks, addressing notable gaps in existing approaches related to data completeness, scalability, and practical applicability, particularly in emergency response contexts.

We evaluate our approach through a case study of Darmstadt, a medium-sized German city. This evaluation demonstrates the processing of open geospatial data to reconstruct representative MV grid models efficiently, identify potential discrepancies, and address data gaps in the OSM transformer inventory, especially in commercial and industrial zones with lower mapping completeness (data coverage and availability). Performance metrics, visual outputs, and statistical assessments compare land use power demand estimation to the OSM power transformer data to identify potential data gaps.

The structure of the paper is as follows. The subsequent section reviews related work on grid reconstruction. Section 3 presents preliminaries and the method overview, followed by data acquisition and preprocessing in Sections 4 and 5, respectively. Section 6 expounds upon secondary substation estimation based on land use. The generation of the MV synthetic energy grid is delineated in Section 7. The visualisation approach is outlined in Section 8, and the evaluation case study, discussion, implications and limitations are presented in Section 9. Finally, Section 10 concludes and discusses future directions.

2. Related Work

The automated modelling of power grids has gained increasing attention in recent years, driven by the need for scalable and efficient solutions to address growing urban energy demands [16]. Traditionally, grid reconstruction relied on static assumptions and proprietary datasets. In response, recent research has shifted towards open-data and algorithmic approaches.

A plethora of open-source tools and initiatives [8,10,17–21] extract power grid information from OSM, thereby demonstrating the potential for automated modelling to streamline energy distribution network reconstruction and support adaptive urban planning [22]. However, many works have centred on High Voltage (HV) networks [23], while MV and Low Voltage (LV) grids remain challenging due to data scarcity, sensitivity, and proprietary restrictions. Furthermore, the incomplete mapping of underground cables and smaller-scale components frequently restricts the applicability of these open-data-based methods [24], emphasising the intricacy and variability of contemporary urban energy systems.

Kisse et al. [25] introduced a Geographic Information System (GIS)-based framework that combined OSM data with Pandapower to model power and gas distribution grids integrated with heat pump systems. Although their synthetic grid models provided valuable insights, such as investment costs and CO₂ reductions, the approach still required manual editing of transformer locations where source data were insufficient, limiting its scalability and full automation. In a similar vein, Dierich et al. [9] and Fekete [4] integrated qualitative stakeholder insights with GIS analyses to assess critical infrastructure interdependencies and cascading effects during crises. While these mixed-methods approaches enhanced situational awareness, they were dependent on localised data, expert input, and proprietary GIS tools, which can impede generalisability.

Focusing specifically on MV grids, Gebhard et al. [10] leveraged OSM substation and street network data, Capacitated Vehicle Routing Problem (CVRP)-based optimisation, Voronoi diagrams, and Delaunay triangulation (a method that naturally connects points

that are spatially close) to generate cost-optimal MV topologies. Their method incorporated optional manual adjustments, for instance, by including proprietary or uncertain load data as a list of custom locations. Although land use data were discussed, they were not integrated into the automation process. Their case study demonstrated that even incomplete OSM data for secondary substations can yield realistic grid topologies. However, the quality of the result depends on the quality of the regional OSM data.

Tomaselli et al. [8] further advanced the field by generating ensembles of LV grid topologies to capture uncertainty through probabilistic methods. However, their approach relied on expert-provided substation inputs [25]. Similarly, [21] assumed secondary substations as given. In another study by Tomaselli et al. [26], external references were employed to extract or synthesise transformer coordinates, thereby reducing manual intervention and the reliance on external datasets, though not entirely eliminating it.

Meanwhile, Baecker et al. [13] presented a methodology that combines OSM-derived building data with statistical information from sources like CORINE Land Cover [27] and TABULA [28], relying on German census-based assumptions [29] for household counts and external studies for peak loads [15,30]. Despite its innovative nature and ability to generate highly detailed results, this approach is contingent on region-specific data for residential and non-residential peak loads, as well as expert-derived building typologies. This limitation raises concerns regarding its generalisability to other regions.

Beyond these approaches, land use data have proven valuable for infrastructure modelling [31–34], providing essential spatial insights by classifying geographic areas according to their social functions. However, most MV grid reconstruction efforts employ land-use data to define load areas or generate virtual MV-LV stations, often without integrating OSM-derived power infrastructure data [35–37]. For instance, the DINGO tool [35] employs OSM-based land-use classifications to delineate load zones, followed by clustering or Voronoi-based assignment of supply points to HV-MV stations. While some methods use the Traveling Salesman Problem (TSP) [37] or the CVRP [35,36] to optimise grid layouts, these approaches generally model direct point-to-point connections and have focused on rural settings.

Whilst earlier studies such as [36,37] have also proposed methods for generating synthetic power grids, these differ significantly from our approach in terms of application scope, input data, and methodological philosophy as they do not consider the street network, focus on rural settings, and rely on certain land-use data. The focus of [36] is on idealised long-term planning scenarios primarily in suburban and rural settings, utilising demographic, census, and Corine Land Cover datasets in conjunction with probabilistic modelling techniques. In a similar manner, the study in [37] employs deterministic heuristic methods in rural areas characterised by agriculture and forestry, relying primarily on clustering and shortest-path optimisation without explicitly incorporating existing transformer locations from OSM. In contrast, our methodology uniquely focuses on rapid grid reconstruction for emergency or crisis scenarios in urban environments. We exclusively utilise publicly available OSM data, explicitly leveraging actual MV-LV transformer positions, systematically augmenting only incomplete mapping areas and reconstructing the MV grid through deterministic constraint optimisation.

In summary, despite the significant progress that has been made in the field of automated grid reconstruction, three main challenges persist:

- Focus on HV networks: The reconstruction of MV and LV remains under-explored and less automated.
- Reliance on manual inputs, proprietary data, expert knowledge, or geographically constrained datasets: Many methods still require expert intervention, closed-source datasets, or regional assumptions that limit scalability and reproducibility.

- Underutilisation of land use data for urban demand and substation estimation: While the use of OSM data in research is increasing, its land use layer is rarely used to estimate missing grid elements directly in dense urban environments.

The present work addresses these gaps by presenting a synthetic grid reconstruction methodology that integrates both OSM land use data and existing power infrastructure, automating spatial energy demand estimation and MV grid topology generation. The objective of the proposed approach is to deliver a scalable solution for urban environments, with the aim of reducing reliance on proprietary sources and manual refinements. The potential application of this approach extends to crisis management and rapid decision-making scenarios.

3. Method

The section provides a contextual framework for the modelling approach that will subsequently be employed. The preliminaries of the problem are introduced in Section 3.1, and the method overview is presented in Section 3.2.

3.1. Preliminaries

Figure 1 provides a schematic representation of an electrical distribution grid, illustrating its integration within the broader context of the power system. For this paper, the MV grid, including its nodes and lines, as well as the primary and secondary substations, is relevant.

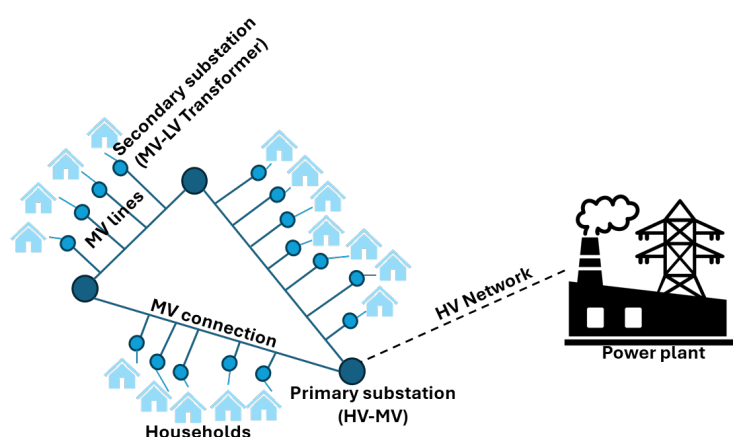


Figure 1. Illustration of a distribution grid.

Electrical grids consist of buses, substations and lines at varying voltage levels. Transformers in substations facilitate voltage conversion, thereby facilitating the transmission of electrical energy. While electrical power transmission over large distances typically occurs at the extra-high voltage level (above 200 kV), distribution grids can be categorised into HV, MV, and LV levels. For instance, in Germany, the MV level is commonly defined as below 50 kV, while LV is defined as below 1 kV. Residential and commercial customers are connected at the LV layer, while industrial ones are often connected to the MV grid. Traditionally, power generation was exclusively connected to the transmission grid. However, due to the ongoing energy transition, distribution grids are undergoing a progressive transformation with an increasing integration of local generation sources such as wind and solar power.

At urban scales, two substation types are particularly relevant to this work. The first is known as a *primary substation*, which contains transformers that facilitate power conversion between the HV and MV levels. The second is known as a *secondary substation* (a transformer), which converts MV to LV. Due to their network character, distribution

grids can be modelled as undirected, weighted, geometric graphs: $\mathcal{G} = (V, E)$. Each node $v \in V$ corresponds to a bus or substation, connected via power lines (mostly cables in urban areas), represented as edges $e \in E$. The edge weights can be interpreted as the cost of the line. The network is modelled as undirected because power lines allow for a bidirectional power flow.

3.2. Method Overview

The proposed approach, illustrated in Figure 2, provides an automated method for reconstructing a synthetic MV power grid by leveraging land-use and power infrastructure data from OSM. The final output of this process is a graph-based representation of the MV network, as outlined below:

- Nodes are used to represent primary and secondary substations. Each node is associated with a specific geographic location and a set of relevant attributes, including, but not limited to, voltage and type.
- Edges represent MV lines aligned with actual street paths, thereby ensuring spatial realism.

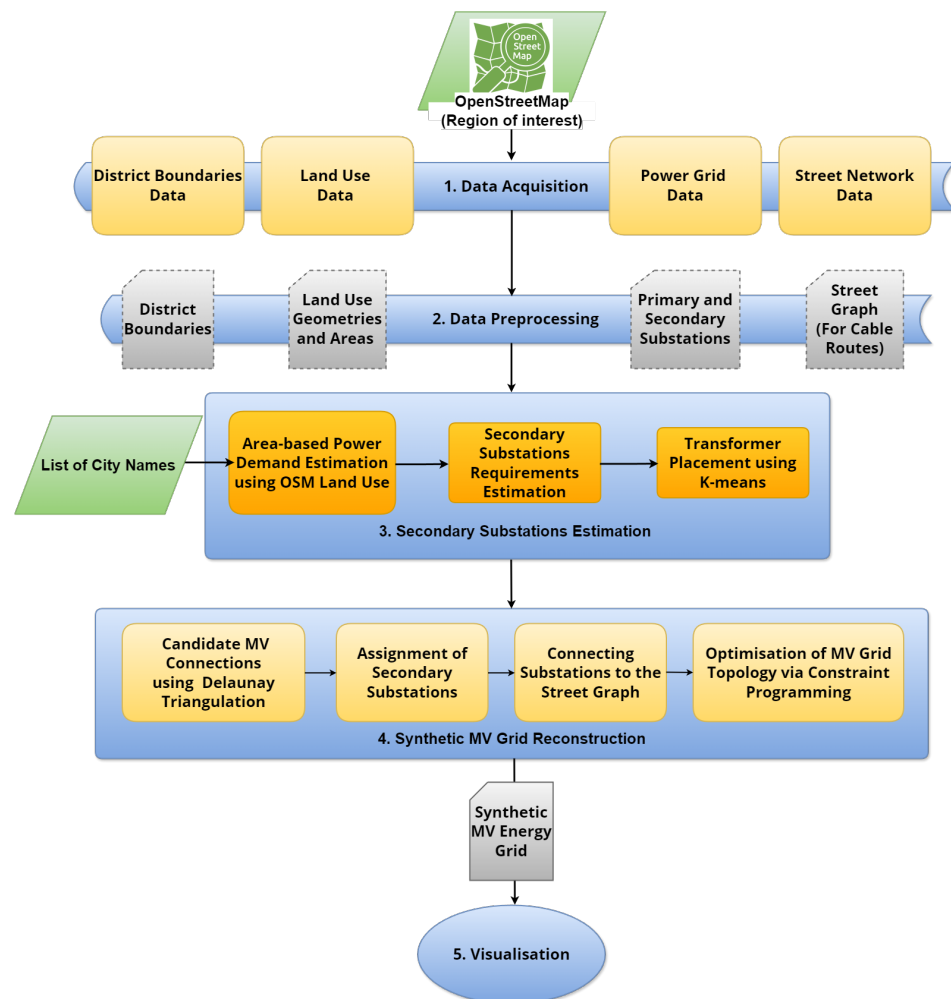


Figure 2. Flowchart describing the workflow of the proposed approach.

By framing the grid in this manner, the approach allows stakeholders (e.g., crisis response teams) to identify approximate cable locations and assess which specific roads might carry MV lines. For instance, in the event of a damaged street or compromised primary substation, the method can be employed to gauge the potential for disruption to the power supply rapidly. The methodology minimises total cable length subject to

capacity constraints (via an optimisation problem detailed in Section 7.4) to yield a plausible, cost-effective MV layout.

The approach in Figure 2 is structured around four key steps, each of which generates outputs that are essential for subsequent analyses:

1. **Data acquisition:** Administrative district boundaries, land-use features, power grid data, and street network information are retrieved from OSM.
Result: A collection of raw geospatial data layers—district polygons, land-use polygons, existing substations, and street network.
2. **Data preprocessing:** The geospatial data are cleaned, filtered, and standardised into a consistent coordinate system.
Result: A validated and harmonised dataset, complete with corrected geometries, classified substations, and uniform spatial referencing.
3. **Secondary substation estimation:** Land-use areas are aggregated, and a specialised workflow is applied to detect coverage gaps in the existing MV grid and estimate where additional transformers may be required.
Result: An updated substation inventory that bridges gaps in the MV grid by incorporating newly placed secondary substations.
4. **Synthetic MV grid reconstruction:** Network cost is minimised subject to capacity constraints, ensuring that cables align with realistic street paths for a plausible MV network topology.
Result: A fully constructed synthetic MV network—modelled as a weighted graph where edges follow the street geometry—enabling subsequent analysis or scenario-based evaluations (e.g., identifying the impact of street closures or a primary substation compromise on the power supply).

These steps support demand estimation, coverage validation, and optimisation of grid layout. Algorithm 1 provides a high-level overview of the reconstruction process, while the subsequent sections elaborate on each step in detail. Therefore, the algorithm's product is a spatially explicit MV grid with edges reflecting actual streets, which can serve as a practical resource for crisis managers and municipal planners.

Algorithm 1 Steps for MV Grid Reconstruction using Land Use Estimation

- 1: **Input:** Region of interest (area name)
 - 2: **Output:** Optimised placement of secondary stations and synthetic power grid topology
 - 3: **Step 1: Data Acquisition (Fetch Data from OSM)**
 - 4: Fetch district boundary data, land use data, power grid data, street network data
 - 5: **Step 2: Data preprocessing**
 - 6: Preprocess district geometries, land use polygons, and classify substations
 - 7: **Step 3: Secondary Substation Estimation (Based on Land Use Data)**
 - 8: 3.1 Assign transformer's capacity
 - 9: 3.2 Calculate Land Use Areas
 - 10: 3.3 Derive state demand per square meter estimations using land use data
 - 11: 3.4 Calculate power demand per area using land use data
 - 12: 3.5 Estimate substations according to land use and supply–demand discrepancy
 - 13: 3.6 Placement of synthetic transformers
 - 14: **Step 4: Synthetic MV Grid Reconstruction**
 - 15: 4.1 Generates candidate MV connections between primary substations
 - 16: 4.2 Assignment of secondary substations to candidate MV connections
 - 17: 4.3 Connecting substations to the street graph
 - 18: 4.4 Generation of street network edges and calculating the actual street distances
 - 19: 4.5 Optimise MV grid via Constraint Programming (CP) formulation
 - 20: **Step 5: Visualisation**
 - 21: Generate layered visuals to facilitate further analysis.
-

The presented methodology was carefully selected and developed based on several critical criteria to overcome the limitations identified in previous works. Specifically, the following list provides a summary of the most important aspects of our methodology:

- Our method prioritises complete automation and scalability using solely OSM data.
- The employment of an OSM land-use transformer estimation strategy eliminates the necessity for manual transformer input, proprietary knowledge, or geographically constrained datasets, thereby ensuring the reproducibility and broad applicability of the method.
- The employment of K-means clustering for synthetic transformer placement provides systematic, algorithmically derived spatial locations without the need for subjective manual adjustments.
- Constraint-programming optimisation that ensures realistic, operationally viable MV grid configurations are aligned with actual urban street networks, improving upon heuristic or manual topology definitions in the literature.

These methodological choices uniquely position our approach as an innovative, open-data automated solution for rapid and scalable reconstruction of MV power grids.

4. Data Acquisition from OSM

Our approach relies on OSM as the only source for geospatial data to obtain four main data layers: district boundaries, land use data, power grid infrastructure, and street networks. The OSMnx library automates data retrieval. This section corresponds to *Step 1* in Algorithm 1.

District Boundary Data: Districts are defined as municipal administrative subdivisions serving governance and planning functions. In this study, district boundaries are extracted from OSM to define the study area, thereby enabling spatially disaggregated analyses of infrastructure, land use and demand. In OSM, these subdivisions are tagged with `boundary = administrative` and a corresponding `admin_level` (https://wiki.openstreetmap.org/wiki/Key:admin_level, accessed on 25 February 2025). In this work, we focus on and retrieve the relevant polygons for `admin_level = 10`, as it typically represents the smallest formally recognised administrative units.

Land Use Data: Within the extracted district boundaries, we retrieve land use features: `landuse = residential`, `landuse = commercial`, and `landuse = industrial`. These categories form the basis for estimating urban energy demand as we associate each land use type with different demand patterns.

Power Grid Data: Power grid infrastructure, including substations and transformers, is queried using OSM tags `power = substation` and `power = transformer`. Due to their good coverage in OSM and their manageable number for one city, primary substations can be assumed as given. Relevant attributes such as name, voltage, and frequency are included. Geometries are processed with the Shapely library, extracting coordinates for point features and centroids for polygonal features.

Street Network Data: Due to the limited availability of MV power distribution lines in OSM, they are assumed to follow street layouts, as supported by prior studies [8,10,24]. Therefore, the street network from OSM is used as the base graph for modelling possible cable routes, assuming a radial grid topology. The street network data are retrieved via OSMnx based on predefined modes (e.g., `walk` for pedestrian paths and `drive` for vehicle-accessible roads). This network is used later to compute realistic distances between substations and to model cable routes.

5. Data Preprocessing

This section corresponds to *Step 2* in Algorithm 1. It refines and standardises the raw OSM-extracted data layers—district boundaries, land use polygons, power grid elements, and street networks—so they can reliably support the subsequent grid reconstruction. In particular, it tackles the data inconsistencies and missing attributes often observed in OSM, a community-driven platform. Such crowd-sourced data can contain partially overlapping polygons, inconsistent geometric definitions, and attribute discrepancies (e.g., missing voltages or incorrect substation tags).

District Boundaries: District polygons, extracted initially from OSM, may overlap or extend beyond the main city area due to heterogeneous mapping inputs or different `admin_level` definitions. Filtering is applied to retain districts fully contained within the city boundary using `geometry.within()`. Overlapping and subsumed regions are systematically removed using `geometry.covers()`. Then, we resolve invalid geometries. Geometries are checked for errors such as self-intersections and gaps. As a topological cleaning technique, a zero-width buffer operation (`buffer(0)`) in Shapely is applied to reconstruct invalid polygons, enforcing topological consistency. This step aims to prevent computational errors in area calculations and spatial analysis. Next, each district is transferred into a uniform coordinate system suited for distance and area measurements. This ensures that the polygons do not double-count or omit any city areas and are comparable and measurable on the same spatial scale. In this study, EPSG:32632 is used, as it is well suited for regions such as Germany, where UTM Zone 32 provides accurate spatial measurements.

Land Use Geometries and Areas: Although we initially extracted residential, commercial and industrial land use features in the acquisition step, these polygons are stored in OSM's global latitude–longitude format. For more precise local-area calculations, geometries are transformed to EPSG:4326, the coordinate reference system used by OSM for geospatial data representation. Relevant land use features are extracted using OSMnx's `features_from_polygon()` function and clipped to district boundaries using the GeoPandas' `clip()` function. As a result, each district ends up with its own clearly defined land use category data, which can then be analysed independently for energy demand or other purposes.

Primary and Secondary Substations: Only a few features have assigned voltage and frequency values in OSM for the power grid. Therefore, retrieved power grid features must be preprocessed.

- Features with a frequency value that differs from the main grid frequency (50 Hz in Europe) are excluded, for example, to avoid elements from the railway power system.
- Substations without names are assigned placeholder identifiers based on OSM IDs.
- Due to inconsistencies in OSM tagging (e.g., `interchanging power = substation` and `power = transformer`), a voltage-based strategy ensures consistent classification.
- Features with voltage above 50 kV are categorised as primary substations, while those below 50 kV or with missing voltage values are classified as secondary substations. A default voltage of 20 kV is assigned to elements with undefined voltage.
- Each substation is then spatially attributed to its corresponding district and land use category, enabling a quantitative assessment, i.e., determination of substation counts per land use within each district.

Street Graph: For computational efficiency, the static street network graph is cached using the `pickle` library, which can be reloaded while validating it against a predefined freshness period. This allows quick reloading for further processing or distance computations, avoiding redundant online queries and reducing runtime in repeated simulations.

This two-step acquisition–preprocessing arrangement helps ensure that OSM-extracted features are both comprehensive (broadly acquired) and accurate (cleaned, consistently measured, and clipped) before moving to the next step in the grid reconstruction process.

6. Secondary Substation Estimation Based on Land Use

This section corresponds to *Step 3* in Algorithm 1 and details an open-data automated method based solely on OSM to estimate and position the missing secondary substations.

6.1. Transformer Capacities

As noted by [13], LV transformer capacities are scarcely available in OSM. Due to the unreliability of OSM metadata, their study employed building density as a proxy to select standard transformer sizes conforming to standard utility practices. Since our approach deliberately does not incorporate buildings to avoid reliance on external, geographically constrained datasets for building-area-dependent peak load estimations and instead relies solely on OSM land use data, we adapt their methodology to determine transformer capacities. Specifically, we select 160 kVA for residential areas (low to medium demand), 250 kVA for commercial zones (medium to high demand) and 400 kVA for industrial areas (high demand).

6.2. Land Use Area Calculation

Each land use category is aggregated within each district to obtain cohesive polygons (e.g., all residential polygons unified). We then compute these unified polygons' total area (in square meters). This step effectively yields each district's final land use areas, forming the spatial foundation for subsequent demand estimation.

6.3. Demand per Square Meter Using OSM Land Use

This subsection describes an open-data automated method to derive demand per square meter for various land use types using OSM data.

We gather transformer data and land use areas from multiple cities to estimate energy demand for each land use category. This ensures that derived demand values reflect state-level conditions. In this study, 25 counties and independent cities in Hesse, Germany, encompassing 299 districts, are employed for this purpose. The input data can be readily extended to include additional cities, enabling a broader application of the proposed methodology. Moreover, the approach exhibits high scalability, as incorporating further cities requires minimal effort—simply adding their names to the input list. Our process follows these steps:

1. **Gather Data from Other Cities:** We collect district-level OSM data (transformers, land use areas) for a set of cities.
2. **Compute Capacity per Land Use:** In each district, we multiply the number of OSM transformers (attributed to that land use) by their standard capacities (see Section 6.1), yielding a total transformer capacity.
3. **Obtain Demand per Square meter:** We divide the total capacity by the district's land use area (in square meters), obtaining a “demand density” (kVA/m²) for that land use type.
4. **Average Across Districts and Cities:** Each city's data are aggregated to produce a city-level average for each land use type. We then compute the state-level average by merging these city-level values, effectively deriving a baseline demand per square meter for each land use category.

This open-data automated approach bypasses reliance on geographically constrained or proprietary datasets, instead leveraging freely available OSM. It is parallelised for efficiency, taking under two minutes (93.34 s in our setup) to process hundreds of districts.

6.4. Power Demand per Area and Number of Secondary Substation Estimation

We obtain the total load for each land use by multiplying the baseline demand per square meter (Section 6.3) by the land use area. Next, we divide this load by the transformer's effective capacity (adjusted for future growth and conservative loading) and round up to the nearest whole number, revealing the required count of secondary substations. Comparing this estimate with the actual OSM-based substation inventory highlights any deficit, i.e., the missing transformers needed to meet the estimated demand.

Because OSM data can be incomplete, our method may estimate fewer transformers than OSM indicates (as shown in the discussion in Section 9 for residential areas with multi-story buildings). In such cases, we do not remove existing substations; rather, we focus on bridging deficits. To mitigate possible underestimation, we apply a 10% growth factor (aligned with forecasts projecting load increases [38–40]) and an 80% loading threshold to account for thermal stress [41]. These measures ensure we do not drastically under-represent real-world demand conditions and possibly reflect common planning assumptions that may have been considered in the grid's original design.

6.5. Transformer Placement

This subsection outlines how missing transformers are added to areas where our land use-based estimation indicates deficits. We employ a grid-based approach to represent demand as discrete “load zones” (Section 6.5.1), then apply K-means clustering to identify suitable transformer locations (Section 6.5.2).

6.5.1. Load Zone Generation

A spatially explicit representation of energy demand is required to guide the placement of new transformers. A “load zone” is a small, discrete portion of a land use polygon, essentially a point representing some fraction of the total demand in that area. To create these zones, we overlay a regular grid on each district's land use polygon wherever the estimated number of needed transformers exceeds what OSM currently shows. Each grid cell within the polygon is assigned a portion of the total demand, producing a set of load zones with associated load values. By distributing demand uniformly among these generated grid points, we avoid lumping all consumption into a single location and using external, geographically constrained peak load estimations or proprietary datasets. Although this approach does not capture accurate variation, it provides a transparent and scalable way to represent spatial demand patterns using only OSM. The result is a set of load zones with associated load values.

6.5.2. K-Means Clustering for Synthetic Transformer Placement

We combine these load zones with the coordinates of existing transformers in the same district and land use. For each district and land use category with missing transformers, we create one cluster per missing transformer using K-Means clustering. Each cluster centre (the centroid of a group of zones and existing transformer points) becomes a new synthetic transformer location. This ensures new transformers are placed in areas with loads, are respectful of existing transformer positions, and are likely to reduce the distance between load pockets and transformer sites (helping lower technical losses). An optional post-processing step can be applied to “snap” each new transformer to the nearest valid street node to ensure realistic siting and viable locations.

Although we do not visualise these intermediate steps here, Section 9 of the evaluation presents a map illustrating how newly placed transformers align with the identified load zones and existing infrastructure. Transformer placement is implemented utilising Shapely and scikit-learn libraries for spatial operations and K-Means clustering, respectively. The outcome of this step is a list of new transformer locations to cover previously unmet demand, which is then added to the secondary substation inventory for the synthetic MV grid reconstruction.

7. Synthetic MV Grid Reconstruction

This section outlines the method for generating a synthetic MV grid and corresponds to Step 4 in Algorithm 1. The process involves generating candidate MV connections and the geometric assignment of secondary substations to candidate MV connections between primary substations (HV-MV) in Section 7.1, connecting substations to the street graph in Section 7.2, generating street distances and street network edges between substations in Section 7.3, and finally, the optimisation of MV grid topology in Section 7.4.

7.1. Candidate MV Connections and Assignment of Secondary Substations

We assume that each MV line is treated as a direct link between two distinct primary substations. This configuration permits each point to be served by two separate sides and from different primary substations, and all secondary substations are connected to the grid by exactly one line, not consisting of junctions. This keeps the network model focused on primary substations and their attached secondary substations.

- **Generating Candidate MV Connections:** To identify potential MV routes, we adopt the proximity-based approach [10]. First, we gather the geographic coordinates of all primary substations. Next, we apply Delaunay triangulation, which connects points so no substation lies inside the circumcircle of any triangle, resulting in line segments between spatially close substations. We collect each segment to form a set of candidate MV connections. A visualisation of the Delaunay triangulation is shown in Figure 4a.
- **Assigning Secondary Substations:** Each secondary substation is assigned to the nearest candidate line based on Euclidean distance. This ensures every secondary node attaches to exactly one MV connection, providing the basis for subsequent integration, where realistic street network edges, distances and operational constraints will be incorporated to refine the synthetic grid topology.

7.2. Connecting Substations to the Street Graph

The electrical grid can be modelled as a graph, with nodes representing supply points (primary and secondary substations) and edges corresponding to cable lines following the street network. This process connects each substation to its nearest node within the street graph, aligning the grid with actual urban infrastructure.

7.3. Generating Street Distances and Street Network Edges Between Substations

With candidate connections established, we next integrate these with real street network data to ensure the grid's alignment with urban infrastructure. To model realistic MV grid topologies, we compute the shortest path distances between the connected substations (Section 7.2) based on the street network rather than geodesic distances using the NetworkX library. This generates the required distance matrix for optimisation while ensuring cost-effective grid reconstruction.

7.4. Optimisation of MV Grid Topology via Constraint Programming

Reconstructing an efficient MV grid topology requires minimising total cable length while ensuring that transformer capacities and physical cable limits are not exceeded. We address this routing challenge by formulating it as a constraint-programming problem [42,43] that explicitly encodes route continuity, capacity, and load distribution constraints.

7.4.1. Problem Formulation Using Constraint Programming

We model the grid as a sequence of nodes where primary substations serve as terminal nodes (node 0 for the start and node $(n - 1)$ for the end), and secondary substations (each with a capacity requirement ψ_l in kVA) occupy intermediate positions. The cost of connecting any two nodes i and j is defined by the street-network distance d_{ij} (computed in Section 7.2), and the total load along the route must not exceed a cable capacity B (e.g., 10,000 kVA). Formally, we define:

Variables: Binary decision variables x_{ij} for all $i \neq j$, where $x_{ij} = 1$ if the route travels directly from node i to node j , and auxiliary variables u_i (for $i = 1, \dots, n - 1$) that track the accumulated load at node i .

Constraints: (i) *Route continuity:* Each node (except the last) must have exactly one outgoing edge, and each node (except the first) exactly one incoming edge:

$$\sum_{j \neq i} x_{ij} = 1 \quad \text{and} \quad \sum_{i \neq j} x_{ij} = 1.$$

(ii) *Capacity and subtour elimination:* For all $i, j \in \{1, \dots, n - 1\}$ with $i \neq j$, if the route travels from node i to node j (i.e., $x_{ij} = 1$), then the cumulative load at node j must be at least the cumulative load at node i plus the demand at node j (i.e., ψ_j), without exceeding the cable capacity:

$$u_i + \psi_j \leq u_j + B(1 - x_{ij}),$$

Finally, the (iii) *overall capacity* constraint limits the total load of secondary substations in the route:

$$\sum_{l \in \text{secondary}} \psi_l \leq B$$

The objective function: We aim to minimise the total cable length:

$$\min \sum_{i \neq j} d_{ij} x_{ij},$$

where d_{ij} are the street-network distances between nodes i and j . These definitions yield a *constraint-programming* model that enforces route continuity, capacity restrictions, and load distribution while minimising total cable length.

7.4.2. Solver Utilisation

We assemble a distance matrix from precomputed street-network paths and a demand vector $\{\psi_l\}$ for each secondary substation. The solver then uses the variables $\{x_{ij}\}$ and $\{u_i\}$ to find a sequence of nodes (primary-to-primary) that satisfies all constraints while minimising total cable length. The final solution includes the ordered route, total distance, and auxiliary outputs (e.g., the street segments).

7.4.3. Computational Time and Data Requirements

We selected CP due to its capability to handle spatially explicit optimisation problems, enforcing detailed operational constraints (e.g., route continuity, sub-tour elimination, and

cable capacity) simultaneously. CP provides feasibility checks and flexible integration of street-network distances, transformer capacities, and explicit grid topology constraints.

The computational time in our scheme is primarily determined by the complexity of each candidate instance. For each candidate, the number of nodes (i.e., the primary substations plus assigned secondary substations) directly affects the size of the distance matrix and the number of decision variables and constraints in the CP model.

A maximum time of (90 to 300) seconds per candidate instance is set, which ensures that the solver terminates even if an optimal solution is not found. In practice, many instances converge faster than this limit. The CP solver is designed to handle combinatorial problems efficiently within these time constraints.

To further manage computational time, we run candidate instances in parallel using a `ThreadPoolExecutor`, provided by the `concurrent.futures` python package to launch parallel tasks. This parallelisation helps to scale the process when dealing with multiple candidates simultaneously, effectively reducing the overall processing time.

Unlike machine learning or deep learning methods that require large datasets for training to achieve acceptable accuracy, our approach is based on constraint programming. There is no training phase; rather, the “accuracy” of our solution is about feasibility and optimality, given the constraints. The CP model relies on input data such as candidate connections, street distances, and substation information. Acceptable solution quality is achieved if the street distance data accurately reflect the real-world network, the candidate connection data correctly capture the possible configurations, and the assignment mapping (secondary substations to primary pairs) is appropriately defined.

8. Visualisation

An interactive map presents the grid model as distinct, toggleable layers for clarity and analysis using Folium. Administrative boundaries are outlined in black dashed lines with an orange fill, while land use categories are colour-coded for differentiation. Primary and secondary substations are marked with ID, voltage, and classification, with additional secondary substations appearing as black markers. Each substation is linked to the nearest street node, and demand clusters are visualised as filled polygons. MV grid routes are displayed using actual street paths, allowing interactive exploration of network expansion scenarios. The colour-coded, toggleable layers ensure clear differentiation and can support decision-makers analysing the grid.

9. Evaluation and Discussion

9.1. Case Study for a German City

We apply our methodology to Darmstadt, a medium-sized city with a population of over 168,457 (<https://www.darmstadt.de/standort/statistik/statistik-aktuell>, accessed on 25 February 2025).

Figure 3 illustrates the study area in Darmstadt, showing its administrative division into nine distinct districts. It includes detailed land-use categories derived from OSM data: residential zones indicated in cyan, commercial zones in yellow, and industrial zones in magenta. These land-use classifications form the basis for subsequent demand estimation, allowing accurate identification of potential areas requiring additional substations due to gaps in existing infrastructure mapping.

Figure 4a presents an intermediate visualisation of the power grid’s primary substations and the triangles generated via the Delaunay triangulation. The figure depicts candidate MV connections as blue polylines linking HV-MV substations, which are shown as larger blue circle markers.

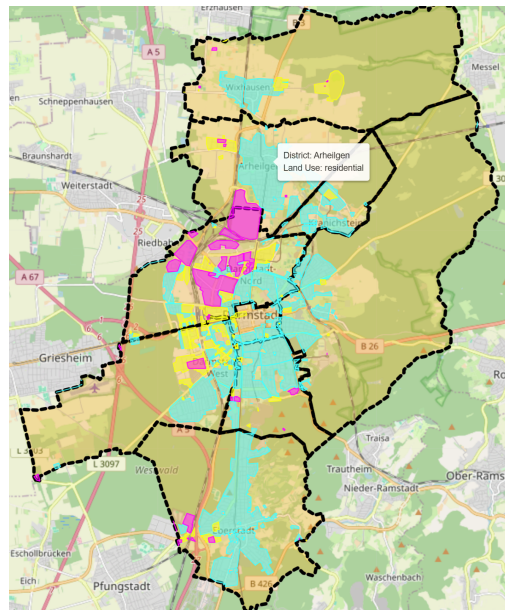
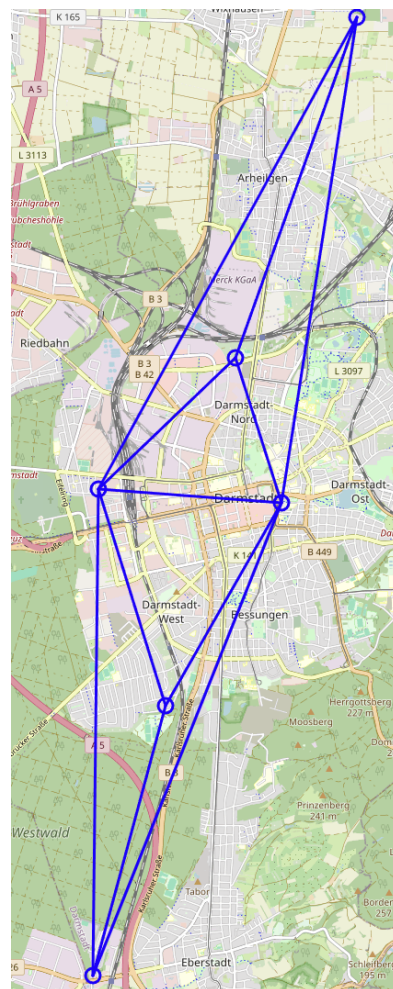
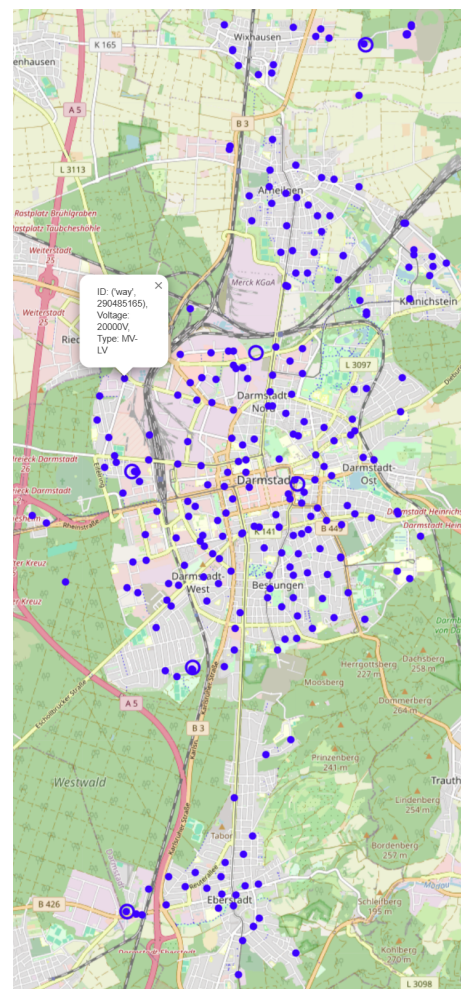


Figure 3. Darmstadt: district administrative boundaries and land use data: residential zones (cyan), commercial zones (yellow), and industrial zones (magenta).



(a)



(b)

Figure 4. (a) Primary substations with the Delaunay triangulation for MV connections. (b) Darmstadt OSM power data: primary (large circles) and secondary (small circles) substations.

Figure 4b visualises power infrastructure elements extracted directly from OSM, highlighting six primary substations (larger blue-outlined circles) and 258 secondary substations (smaller filled blue circles). Each marker is labelled explicitly with the substation ID, voltage level, and classification (HV-MV for primary and MV-LV for secondary substations). This detailed visualisation enables immediate differentiation between substations, clearly reflecting their roles and spatial distribution within the grid structure. Using this infrastructure data, we subsequently construct a complete street-network graph and precompute the shortest paths between all substation pairs, caching the results for computational efficiency.

Figure 5 demonstrates the augmentation procedure for addressing transformer coverage gaps, specifically for the Darmstadt-Nord district. Existing MV-LV transformers extracted from OSM are marked in blue, whereas newly added synthetic transformer locations, determined via K-Means clustering based on estimated demand from land-use data, are marked in black. The visual overlay explicitly illustrates how the automated placement integrates with existing transformer distributions, ensuring comprehensive spatial coverage.

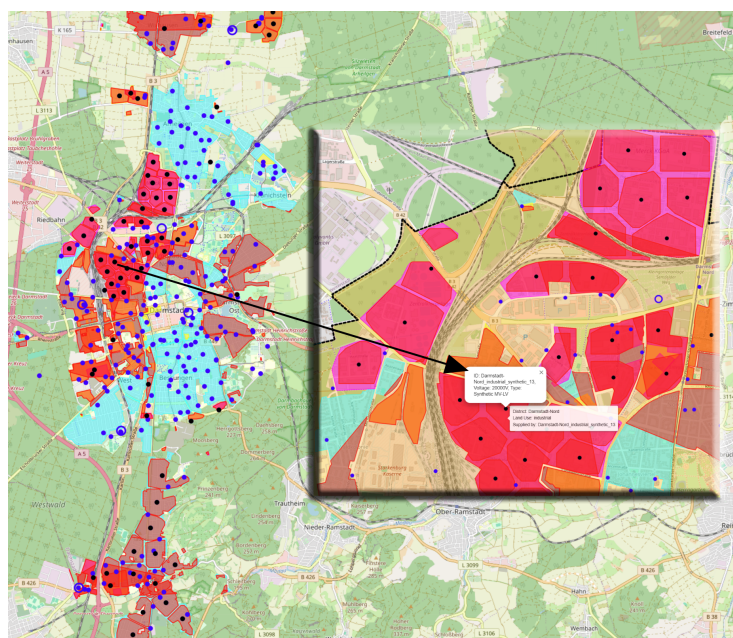


Figure 5. K-Means clustering: the new transformers' placement (black circles) in Darmstadt-Nord, overlaid with existing transformers (blue).

Figure 6 shows the resulting synthetic MV grid topology generated by our constraint-programming solver for the entire Darmstadt area. The optimised grid, consisting of realistic cable routes aligned along street paths, has a total computed cable length of 163,836 m. This figure highlights the reconstructed MV grid's practical usability and spatial realism, supporting rapid interpretation, operational scenario analysis, and decision making for crisis management and infrastructure resilience purposes.

Although our primary goal is to evaluate the model's reliability and to demonstrate the behaviour for city-level scenario analyses rather than optimise speed, we note that the method runs efficiently using a standard laptop with an Intel i7-12700H CPU (Intel Corporation, Santa Clara, CA, USA). For instance, processing district boundaries and land use features required a few seconds, while estimating loads took roughly 3 s. Fetching and processing substations took 0.12 s. The initial computation of street-network distances, including footpaths (664 s), is a one-time cost; after caching, subsequent runs complete in a fraction of this time (14.18 s). Substation connections to the street graph are established in 8.35 s. Applying the transformer placement algorithm took about 37 s, and the solvers had

a max search time limit of 300 s per solution where satisfactory solutions are obtained in 0.32 s, while an optimal solution is obtained in approximately 96 s for a medium-sized city, as shown in the case study.

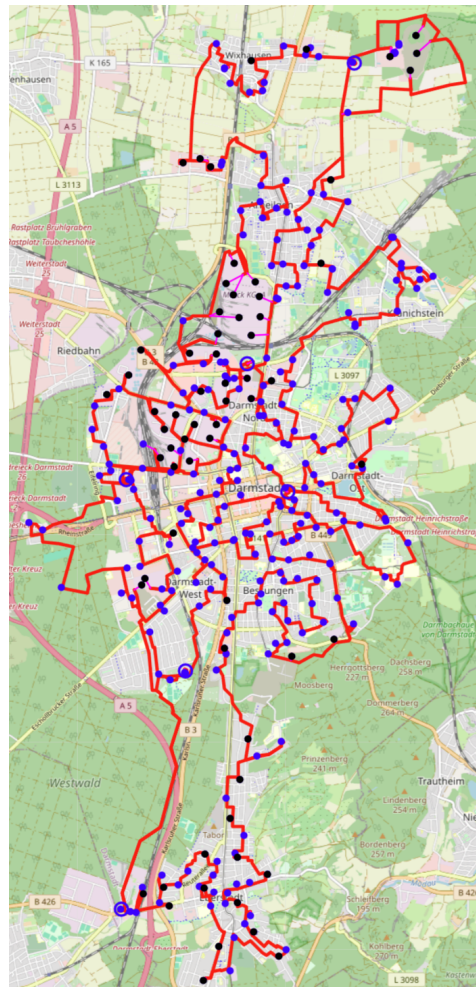


Figure 6. The Resulting MV synthetic grid for Darmstadt has a total cable length of 163,836 m. Cables are highlighted in red, existing OSM substations are shown in blue, and newly added substations are indicated in black.

9.2. OSM Data Quality Assessment

Table 1 summarises the results of a comparison for Darmstadt between the estimated required number of secondary substations according to land use and the actual number retrieved from OSM data to estimate the missing number needed to cover the estimated demand.

We then summarise these discrepancies across residential, commercial, and industrial land uses in Table 2, computing mean difference, mean absolute deviation (MAD), root mean square error (RMSE), and coverage. We report the mean difference primarily to indicate whether our approach tends to over- or underestimate transformer counts overall. However, large positive and negative errors may offset each other, making the mean difference appear smaller. To address this, we also provide the MAD and the RMSE, both of which are less sensitive to sign and reflect the true scale of discrepancies.

Table 1. Comparison of estimated transformers from our land-use approach and OSM data.

District	Metric	OSM #	Estimated #	Net
Eberstadt	Residential	24	38	−14
Eberstadt	Commercial	2	3	−1
Eberstadt	Industrial	2	3	−1
Eberstadt	Total	28	44	−16
Darmstadt-West	Residential	19	19	0
Darmstadt-West	Commercial	16	18	−2
Darmstadt-West	Industrial	1	3	−2
Darmstadt-West	Total	36	40	−4
Darmstadt-Mitte	Residential	17	6	+11
Darmstadt-Mitte	Commercial	4	2	+2
Darmstadt-Mitte	Industrial	0	0	0
Darmstadt-Mitte	Total	21	8	+13
Darmstadt-Ost	Residential	18	19	−1
Darmstadt-Ost	Commercial	3	1	+2
Darmstadt-Ost	Industrial	2	1	+1
Darmstadt-Ost	Total	23	21	+2
Arheilgen	Residential	28	27	+1
Arheilgen	Commercial	2	5	−3
Arheilgen	Industrial	0	5	−5
Arheilgen	Total	30	37	−7
Bessungen	Residential	31	24	+7
Bessungen	Commercial	0	2	−2
Bessungen	Industrial	0	2	−2
Bessungen	Total	31	28	+3
Darmstadt-Nord	Residential	25	18	+7
Darmstadt-Nord	Commercial	15	17	−2
Darmstadt-Nord	Industrial	16	43	−27
Darmstadt-Nord	Total	56	78	−22
Wixhausen	Residential	9	10	−1
Wixhausen	Commercial	4	9	−5
Wixhausen	Industrial	8	1	+7
Wixhausen	Total	21	20	+1
Kranichstein	Residential	11	9	+2
Kranichstein	Commercial	1	1	0
Kranichstein	Industrial	0	0	0
Kranichstein	Total	12	10	+2

Table 2. Summary statistics of transformer comparison.

Land Use	Mean Difference	MAD	RMSE	Coverage (%)
Residential	+1.33	4.89	6.85	107.1
Commercial	−1.22	2.11	2.47	81.0
Industrial	−3.22	5.00	9.51	50.0

The reported coverage percentages reflect the completeness of OSM data relative to our land-use-based estimations. Specifically, coverage greater than 100% indicates that the existing OSM dataset contains more substations than our approach estimates necessary, likely due to local conditions (e.g., multi-story residential buildings). Conversely, lower coverage percentages, such as 81% for commercial and 50% for industrial zones, indicate substantial data gaps in OSM compared to our estimates, highlighting areas where our

synthetic reconstruction contributes the most significant practical value. Overall, residential areas show slight overrepresentation in OSM (107% coverage), possibly reflecting higher mapping activity, or our approach does not capture multi-story buildings. Commercial areas are underrepresented by about 19%, while industrial zones (50% coverage) reveal the most significant gaps in OSM's inventory. These discrepancies underscore the need for robust estimation methods (like ours) that can supplement incomplete OSM data, especially when assessing power infrastructure in crisis scenarios. The reported coverage percentages reflect the completeness of OSM data relative to our land-use-based estimations.

9.3. Implications and Limitations

In practice, overrepresentation of residential OSM transformers may result from community-driven mapping. At the same time, underrepresentation in commercial and industrial districts may stem from partially mapped private substations subject to industry-internal handling or restricted-access facilities. Although our land-use-based approach mitigates these gaps, it also relies on simplifying assumptions—such as uniform load distribution across each land-use area.

The assumption of uniformly distributed loads preserves model generality and circumvents reliance on proprietary or region-specific consumption data; however, it may not accurately capture local variations within a single land use. Similarly, our methodology depends solely on OSM land-use data, foregoing occupant-level coincidence factors, detailed building footprints, and explicit kW-to-kVA conversions. This strategy enhances scalability and ensures applicability in regions lacking high-resolution datasets, yet it can underestimate demand when OSM data are incomplete. To counter this, we incorporate conservative margins via a growth factor and a loading threshold.

Another aspect is that, since our method relies solely on OSM data (both existing energy infrastructure and land use), it can leverage comparable cities—within the same country or even abroad—to derive global-state demand estimates. This is particularly advantageous where OSM data are absent for secondary energy infrastructure (e.g., transformer counts). However, it also underscores the need for research on data-absent contexts that avoid reliance on proprietary or highly localised inputs.

Where apparent excesses arise (e.g., a district with more OSM transformers than estimated), local factors such as multi-story buildings or incomplete neighbour data may be at play. Consequently, we focus on bridging deficits rather than redistributing surplus units. Future investigations could explore inter-district supply dependencies or integrate occupant-level/building-height data where freely available.

Our open-data automated approach offers robust, scalable solutions but may undermine local precision. Higher-accuracy methods can incorporate occupant profiles and floor areas to refine load estimations, though they typically rely on geographically constrained or proprietary data. This trade-off underscores how purely OSM-based models minimise external dependencies and facilitate large-scale implementations, albeit with certain simplifying assumptions. For even greater fidelity, future work might refine these margins or incorporate new open datasets (e.g., nighttime satellite imagery, population density) to capture demand patterns more accurately. To further demonstrate the practical effectiveness and applicability in real-world emergency scenarios, future work would include explicit scenario analyses, such as simulations of infrastructure disruptions (e.g., a primary substation failure or major road blockages), to highlight the synthetic grid's utility in rapid operational decision making. Moreover, obtaining feedback from domain experts (e.g., emergency responders, urban planners, distribution system operators) regarding the synthetic grids' usability and accuracy would strengthen confidence in their value. Such qualitative validation would complement our quantitative evaluations and help demon-

strate the broader operational effectiveness and decision support capabilities enabled by our open-data automated approach.

10. Conclusions

This paper presents an open-data automated approach for reconstructing MV power grids exclusively leveraging publicly available OSM infrastructure and land-use data. Unlike traditional methods relying on proprietary datasets and manual input, our scalable and deterministic approach facilitates rapid grid reconstruction, particularly for crisis management in urban contexts.

Through a detailed case study in Darmstadt, our method demonstrated its effectiveness by reconstructing a realistic synthetic MV grid topology aligned with street paths, totalling approximately 163,836 m of cable. An analysis comparing our transformer estimates to existing OSM infrastructure revealed varying degrees of data completeness; residential zones exhibited slight surplus coverage (107%), commercial zones moderate coverage (81%), and industrial zones showed significant mapping gaps (50%). This highlights areas where our approach adds substantial value. Computationally, our model efficiently performs demand estimation within seconds and synthesises representative grids in minutes, confirming its practical applicability for rapid operational decision making during emergencies.

While our methodology incorporates simplifying assumptions—such as uniform load distribution—these choices ensure scalability and independence from proprietary datasets. To enhance accuracy further, future research may refine these assumptions, explore inter-district dependencies, and integrate additional open datasets like nighttime satellite imagery and population density maps, capturing more nuanced demand variations. Integrating real-time smart meter and sensor data with open-source grid models is a promising future direction to boost model accuracy and responsiveness. Such data-enabled digital twins could perform on-the-fly load calibration, anomaly detection, and adaptive network reconfiguration in response to live conditions, greatly enhancing the value of synthetic grid models for both researchers and practitioners. Nonetheless, achieving this broadly will require overcoming data accessibility challenges, so developing methods that gracefully incorporate live data when available—while still functioning with open data alone—is key to scalable adoption.

The present approach to routing and cost efficiency can be extended in multiple directions. Future work could incorporate local power plants and end users (e.g., residential houses) to encompass the medium- and low-voltage grid comprehensively. In addition, exploring different optimisation techniques, cost parameters, and varying algorithmic settings such as load balancing would refine the routing methodology further and adapt the network design for diverse real-world scenarios.

Author Contributions: Conceptualisation, M.B. and E.B.; methodology, M.B. and E.B.; software, M.B.; validation, M.B.; formal analysis, M.B.; investigation, M.B. and E.B.; resources, M.B.; data curation, M.B.; writing—original draft preparation, M.B. and E.B.; review, E.B., M.B. and T.G.; writing—editing, M.B., T.G. and E.B.; writing—revision, M.B., E.B. and T.G.; visualisation, M.B.; supervision, E.B.; project administration, E.B.; funding acquisition, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been conducted in the context of the “diresCity” project, “Applied Methods for Digital and Resilient Cities”, funded by the German Aerospace Center (DLR). We acknowledge support by the Deutsche Forschungsgemeinschaft (DFG—German Research Foundation) and the Open Access Publishing Fund of Hochschule Darmstadt—University of Applied Sciences.

Data Availability Statement: The data utilised in this study are sourced from OpenStreetMap, and detailed descriptions of its usage are provided in the article.

Acknowledgments: The authors acknowledge using Grammarly (Grammarly Inc., San Francisco, CA, USA; accessed on 25 February 2025), DeepL Write (DeepL SE, Cologne, Germany; accessed on 25 February 2025), and ChatGPT (OpenAI, San Francisco, CA, USA; GPT-4, accessed on 25 February 2025) to enhance grammar and sentence coherence. All content, interpretations, and conclusions are the sole responsibility of the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CI	Critical Infrastructure
OSM	OpenStreetMap
DT	Digital Twin
HV	High Voltage
MV	Medium Voltage
LV	Low Voltage
GIS	Geographic Information System
CVRP	Capacitated Vehicle Routing Problem
CP	Constraint Programming

References

1. United Nations Office for Disaster Risk Reduction. The Disaster Risk Reduction (DRR) Glossary. 2022. Available online: <https://www.undrr.org/drr-glossary> (accessed on 25 February 2025).
2. Federal Ministry of the Interior (Germany). National Strategy for Critical Infrastructure Protection (CIP Strategy). 2009. Available online: https://www.bmi.bund.de/SharedDocs/downloads/EN/publikationen/2009/kritis_englisch.pdf?__blob=publicationFile&v=1 (accessed on 25 February 2025).
3. Cybersecurity and Infrastructure Security Agency (CISA) and U.S. Department of Homeland Security. A Guide to Critical Infrastructure Security and Resilience. 2019. Available online: <https://cncpic.mai.gov.ro/sites/default/files/2020-03/Guide-Critical-Infrastructure-Security-Resilience-110819-508v2.pdf> (accessed on 25 February 2025).
4. Fekete, A. Critical infrastructure cascading effects. Disaster resilience assessment for floods affecting city of Cologne and Rhein-Erft-Kreis. *J. Flood Risk Manag.* **2020**, *13*, e312600. [CrossRef]
5. Cámara Valencia. Valencia Chamber Report on Damages in the 87 Industry Municipalities Affected by DANA. 2024. Available online: <https://www.camaravalencia.com/wp-content/uploads/2025/02/Informe-danos-ocasionados-por-la-DANA-en-la-industria-de-la-zona-afectada.pdf> (accessed on 25 February 2025).
6. Ouyang, M. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliab. Eng. Syst. Saf.* **2014**, *121*, 43–60. [CrossRef]
7. Hoff, R.; Sparks, R.; Chester, M.; Mustafa, A.; Johnson, N.; Birchfield, A.; McPhearson, T.; Li, R.; Ahmad, N.; Searles, I. Cascading Failure Propagation and Perfect Storms in Interdependent Infrastructures. *ASCE OPEN Multidiscip. J. Civ. Eng.* **2025**, *3*, 04025001. [CrossRef]
8. Tomaselli, D.; Stursberg, P.; Metzger, M.; Steinke, F. Representing topology uncertainty for distribution grid expansion planning. In Proceedings of the 27th International Conference on Electricity Distribution (CIRED 2023), CIRED, Rome, Italy, 12–15 June 2023. [CrossRef]
9. Dierich, A.; Tzavella, K.; Setiadi, N.J.; Fekete, A.; Neisser, F.M. Enhanced Crisis-Preparation of Critical Infrastructures through a Participatory Qualitative-Quantitative Interdependency Analysis Approach. In Proceedings of the ISCRAM, Valencia, Spain, 19–22 May 2019; Franco, Z.; González, J.J.; Canós, J.H., Eds.; ISCRAM Association: Brussels, Belgium, 2019.
10. Gebhard, T.; Tundis, A.; Steinke, F. Automated Generation of Urban Medium-voltage Grids using OpenStreetMap Data. In Proceedings of the 2024 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Dubrovnik, Croatia, 14–17 October 2024. [CrossRef]
11. Brucherseifer, E.; Winter, H.; Mentges, A.; Mühlhäuser, M.; Hellmann, M. Digital Twin conceptual framework for improving critical infrastructure resilience. *at-Automatisierungstechnik* **2021**, *69*, 1062–1080. [CrossRef]

12. Gebhard, T.; Sattler, B.J.; Gunkel, J.; Marquard, M.; Tundis, A. Improving the resilience of socio-technical urban critical infrastructures with digital twins: Challenges, concepts, and modeling. *Sustain. Anal. Model.* **2025**, *5*, 100036. [\[CrossRef\]](#)
13. Baecker, B.R.; Candas, S.; Tepe, D.; Mohapatra, A. Generation of low-voltage synthetic grid data for energy system modeling with the pylovo tool. *Sustain. Energy Grids Netw.* **2025**, *41*, 101617. [\[CrossRef\]](#)
14. Verheggen, L.; Ferdinand, R.; Moser, A. Planning of low voltage networks considering distributed generation and geographical constraints. In Proceedings of the 2016 IEEE International Energy Conference (ENERGYCON), Leuven, Belgium, 4–8 April 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6. [\[CrossRef\]](#)
15. Schlömer, G.; Blaufuß, C.; Hofmann, L. Modelling of Low-Voltage Grids with the Help of Open Data. In Proceedings of the 2016 NEIS Conference: Conference on Sustainable Energy Supply and Energy Storage Systems, Hamburg, Germany, 15–16 September; Springer: Berlin/Heidelberg, Germany, 2017; pp. 21–25. [\[CrossRef\]](#)
16. Safari, A.; Daneshvar, M.; Anvari-Moghaddam, A. Energy Intelligence: A Systematic Review of Artificial Intelligence for Energy Management. *Appl. Sci.* **2024**, *14*, 11112. [\[CrossRef\]](#)
17. Banze, T.; Kneiske, T.M. Open data for energy networks: Introducing DAVE—A data fusion tool for automated network generation. *Sci. Rep.* **2024**, *14*, 1938. [\[CrossRef\]](#)
18. Caetano, H.O.; Desuó, L.; de SS Fogliatto, M.; Ribeiro, V.P.; Balestieri, J.A.; Maciel, C.D. A Bayesian Hierarchical Model to create synthetic Power Distribution Systems. *Electr. Power Syst. Res.* **2024**, *235*, 110706. [\[CrossRef\]](#)
19. Gaugl, R.; Wogrin, S.; Bachhiesl, U.; Frauenlob, L. GridTool: An open-source tool to convert electricity grid data. *SoftwareX* **2023**, *21*, 101314. [\[CrossRef\]](#)
20. Medjroubi, W.; Müller, U.P.; Scharf, M.; Matke, C.; Kleinhans, D. Open data in power grid modelling: New approaches towards transparent grid models. *Energy Rep.* **2017**, *3*, 14–21. [\[CrossRef\]](#)
21. Çakmak, H.K.; Janecke, L.; Weber, M.; Hagenmeyer, V. An optimization-based approach for automated generation of residential low-voltage grid models using open data and open source software. In Proceedings of the 2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Novi Sad, Serbia, 10–12 October 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6. [\[CrossRef\]](#)
22. Heitkoetter, W.; Medjroubi, W.; Vogt, T.; Agert, C. Comparison of open source power grid models—combining a mathematical, visual and electrical analysis in an open source tool. *Energies* **2019**, *12*, 4728. [\[CrossRef\]](#)
23. Xiong, B.; Fioriti, D.; Neumann, F.; Riepin, I.; Brown, T. Modelling the high-voltage grid using open data for Europe and beyond. *Sci. Data* **2025**, *12*, 277. [\[CrossRef\]](#)
24. Domingo, C.M.; San Roman, T.G.; Sanchez-Miralles, A.; Gonzalez, J.P.P.; Martinez, A.C. A reference network model for large-scale distribution planning with automatic street map generation. *IEEE Trans. Power Syst.* **2010**, *26*, 190–197. [\[CrossRef\]](#)
25. Kisse, J.M.; Braun, M.; Letzgus, S.; Kneiske, T.M. A GIS-Based planning approach for urban power and natural gas distribution grids with different heat pump scenarios. *Energies* **2020**, *13*, 4052. [\[CrossRef\]](#)
26. Tomaselli, D.; Stursberg, P.; Metzger, M.; Steinke, F. Learning probability distributions over georeferenced distribution grid models. *Electr. Power Syst. Res.* **2024**, *235*, 110636. [\[CrossRef\]](#)
27. European Environment Agency. *CORINE Land Cover—European Union’s Copernicus Land Monitoring Service Information*; Report; European Environment Agency (EEA): Copenhagen, Denmark, 2018. [\[CrossRef\]](#)
28. Loga, T.; Stein, B.; Diefenbach, N.; Born, R. *Deutsche Wohngebäudetypologie. Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden*; zweite erweiterte Auflage; Technical Report; Institut Wohnen und Umwelt GmbH: Darmstadt, Germany, 2015. Available online: https://www.iwu.de/fileadmin/publikationen/gebaeudebestand/episcopes/2015_IWU_LogaEtAl_Deutsche-Wohngeb%C3%A4udetypologie.pdf (accessed on 25 February 2025).
29. Statistische Ämter des Bundes und der Länder. *Zensus 2011: Vielfältiges Deutschland*. Zensus. 2016. Available online: https://www.zensus2011.de/SharedDocs/Downloads/DE/Publikationen/Aufsaeetze_Archiv/2016_12_NRW_Zensus_Vielfalt.pdf?__blob=publicationFile&v=2 (accessed on 25 February 2025).
30. Wille-Haussmann, B.; Fischer, D.; Köpfer, B.; Bercher, S.; Engelmann, P.; Ohr, F. Synthetische Lastprofile für eine effiziente Versorgungsplanung für Nicht-Wohngebäude. Fraunhofer ISE. 2020. Available online: <https://www.tib.eu/en/search/id/TIBKAT:1737777061/> (accessed on 25 February 2025).
31. Alhamwi, A.; Medjroubi, W.; Vogt, T.; Agert, C. OpenStreetMap data in modelling the urban energy infrastructure: A first assessment and analysis. *Energy Procedia* **2017**, *142*, 1968–1976. [\[CrossRef\]](#)
32. Hülk, L.; Wienholt, L.; Cußmann, I.; Müller, U.P.; Matke, C.; Kötter, E. Allocation of annual electricity consumption and power generation capacities across multiple voltage levels in a high spatial resolution. *Int. J. Sustain. Energy Plan. Manag.* **2017**, *13*, 79. [\[CrossRef\]](#)
33. Alhamwi, A.; Medjroubi, W.; Vogt, T.; Agert, C. Development of a GIS-based platform for the allocation and optimisation of distributed storage in urban energy systems. *Appl. Energy* **2019**, *251*, 113360. [\[CrossRef\]](#)
34. Chang, S.; Wang, Z.; Mao, D.; Guan, K.; Jia, M.; Chen, C. Mapping the essential urban land use in changchun by applying random forest and multi-source geospatial data. *Remote Sens.* **2020**, *12*, 2488. [\[CrossRef\]](#)

35. Amme, J.; Pleßmann, G.; Bühler, J.; Hülk, L.; Kötter, E.; Schwaegerl, P. The eGo grid model: An open-source and open-data based synthetic medium-voltage grid model for distribution power supply systems. *J. Phys. Conf. Ser.* **2018**, *977*, 012007. [CrossRef]
36. Tran, J.; Pfeifer, P.; Wirtz, C.; Wursthorn, D.; Vennegeerts, H.; Moser, A. Modelling of synthetic power distribution systems in consideration of the local electricity supply task. In Proceedings of the Cired, AIM, Madrid, Spain, 3–6 June 2019. [CrossRef]
37. Kays, J.; Seack, A.; Smirek, T.; Westkamp, F.; Rehtanz, C. The generation of distribution grid models on the basis of public available data. *IEEE Trans. Power Syst.* **2016**, *32*, 2346–2353. [CrossRef]
38. Federal Ministry for Economic Affairs and Energy. Monitoring the adequacy of resources in European electricity markets. German Federal Ministry for Economic Affairs and Climate Action (BMWK). 2021. Available online: <https://www.bmwk.de/Redaktion/DE/Publikationen/Studien/angemessenheit-der-ressourcen-an-den-europaeischen-strommaerkten.html> (accessed on 25 February 2025).
39. Statista Research Department. Electricity Consumption in Germany from 2000 to 2022. Statista. 2024. Available online: <https://www.statista.com/statistics/383650/consumption-of-electricity-in-germany/> (accessed on 25 February 2025).
40. Wilson, J.D.; Zimmerman, Z.; Gramlich, R. Strategic Industries Surging: Driving Us Power Demand. Grid Strategies, LLC. 2024. Available online: <https://gridstrategiesllc.com/wp-content/uploads/National-Load-Growth-Report-2024.pdf> (accessed on 25 February 2025).
41. Diahovchenko, I.; Petrichenko, R.; Petrichenko, L.; Mahnitko, A.; Korzh, P.; Kolcun, M.; Čonka, Z. Mitigation of transformers' loss of life in power distribution networks with high penetration of electric vehicles. *Results Eng.* **2022**, *15*, 100592. [CrossRef]
42. Dechter, R. *Constraint Processing*; Morgan Kaufmann: San Francisco, CA, USA, 2003.
43. Sapena, O.; Onaindia, E.; Garrido, A.; Arangu, M. A distributed CSP approach for collaborative planning systems. *Eng. Appl. Artif. Intell.* **2008**, *21*, 698–709. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.