Technische Universität München

Fakultät für Informatik

Master's Thesis in Informatik

Inference of High Voltage Power Grids based on Crowdsourced Data

Johannes Leimhofer

Technische Universität München

Fakultät für Informatik

Master's Thesis in Informatik

Inference of High Voltage Power Grids based on Crowdsourced Data

Inferenz von Hochspannungs-Stromnetzen basierend auf Crowdsourced Daten

Author: Johannes Leimhofer

Supervisor: Prof. Dr. Hans-Arno Jacobsen

Advisor: Dipl.-Ing. (Univ.) Jose Adan Rivera Acevedo

Submission: 26.08.2016

I confirm that this master's thesis is my own work and I have documented all sources and material used.

München, 26.08.2016

(Johannes Leimhofer)

Abstract

OpenGridMap is an open-source project that crowdsources realistic power grid data to be used for research purposes. In this thesis, we propose an approach for the automatic generation of power gird simulation models from crowdsourced data. The proposed approach orders the crowdsourced data into power circuit relations, which are then used to produce a CIM description file and subsequently a power grid simulation model. We provide experiments which demonstrate the effectiveness of our approach on OpenGridMap data. Given the large amount of crowdsourced data available, our approach has the potential to generate power grid simulation models of larger size, more variety and more accuracy than the currently available state-of-the-art test power grids.

Inhaltsangabe

OpenGridMap ist ein Open-Source Projekt, dass realistische Stromnetzdaten mittels Crowdsourcing sammelt und diese für wissenschaftliche Zwecke zur Verfügung stellt. Diese Arbeit stellt einen Ansatz zur automatischen Generierung von Stromnetzmodellen, basierend auf crowdsourced Daten, vor. Der vorgestellte Ansatz versucht die ungeordneten Daten in Relationen zu bringen, welche Stromkreis-Sektionen von Hochspannungsnetzen beschreiben. Diese Stromkreis-Relationen werden dann verwendet, um CIM-Dateien und infolgedessen Stromnetz-Simulationsmodelle zu generieren. Um die Effektivität unseres Ansatzes zu unterstreichen, führen wir Experimente für verschiedene Regionen und Spannungsebenen durch. Verglichen mit den derzeit verfügbaren Stromnetz-Testmodellen, verfügt unser, auf großen Crowsourcing-Datenmengen basierender, Ansatz über das Potential größere, vielseitigere und genauere Stromnetz-Simulationsmodellen zu generieren.

Acknowledgment

“*It always seems impossible until it’s done.*” (Nelson Mandela)

In this sense, I want to thank all people that supported me throughout my studies. Besides my girlfriend Sonja, I want to explicitly thank my parents for their patience and mental support during my studies. Moreover, I want to thank my supervisor Jose for his ideas and commitment during the production of this thesis. Also, many thanks to the colleagues of the OpenGridMap project for the uncomplicated and productive collaboration.

Contents

List of Figures

List of Tables

Abbreviations

CDPSM	Common Distribution Power System Model.
CIM	Common Information Model.
CPSM	Common Power System Model.
ENTSO-E	European Network of Transmission System Operators for Electricity.
EPRI	Electric Power Research Institute.
EVs	Electric Vehicles.
GPS	Geographic Positioning System.
IEC	International Electrotechnical Commission.
kV	kilovolts.
kWh	kilowatt-hours.
MW	megawatts.
OGM	OpenGridMap.
OSM	OpenStreetMap.
OWL	Web Ontology Language.
p2p	point-to-point.
RDF	Resource Description Framework.
TC	Technical Committee.
TSO	Transmission System Operator.
UML	Unified Modeling Language.
W3C	World Wide Web Consortium.
WGS	World Geodetic System.

XML	Extensible Markup Language.

Chapter 1

Introduction

1.1 Motivation

The complexity of the German high voltage power grid, as seen at Fig. 1.1, increased enormously the past years. Aside from the variety of renewable energy sources that emerged in the course of the Energiewende [10], many new energy consumers appeared due to the growing establishment of Electric Vehicles (EVs) [11]. Although renewable energy sources and EVs are usually attached to the low voltage distribution network [3], there establishment also challenges the enhancement of the high voltage transmission grid in order to balance supply and demand.

In general, the simulation based on high voltage power grid models enables the identification of vulnerabilities and resilience issues in the transmission system. To obtain such a model, the network's topology and its characteristics, the properties of its nodes and links, have to be known. For security reasons, the topology of transmission grids—the locations of generators, transformers, substations, and lines—is mostly not publicly available. As a consequence, other data sources have to be consulted that help to reveal the topology of transmission grids. Note that in the following the transmission grid is used synonymously for the high voltage power grid.

Figure 1.1: German High Voltage Power Grid as of 2014/01/01 [1].

1.2 Problem Statement

The lack of publicly available information on power grid topologies is a common problem in related research fields [12, 13]. To address this problem the research project OpenGridMap (OGM)[1] [14] has been started, which functions as a public data source for power grid topology data. The core of OGM is an open community for crowdsourcing power grid data with the target to create an open platform for inferring realistic power grids based on actual data. In future, the OGM platform should be a playground for researches and practitioners providing tools for power grid simulation studies. Fig. 1.2 shows the OGM platform for the area of Munich.

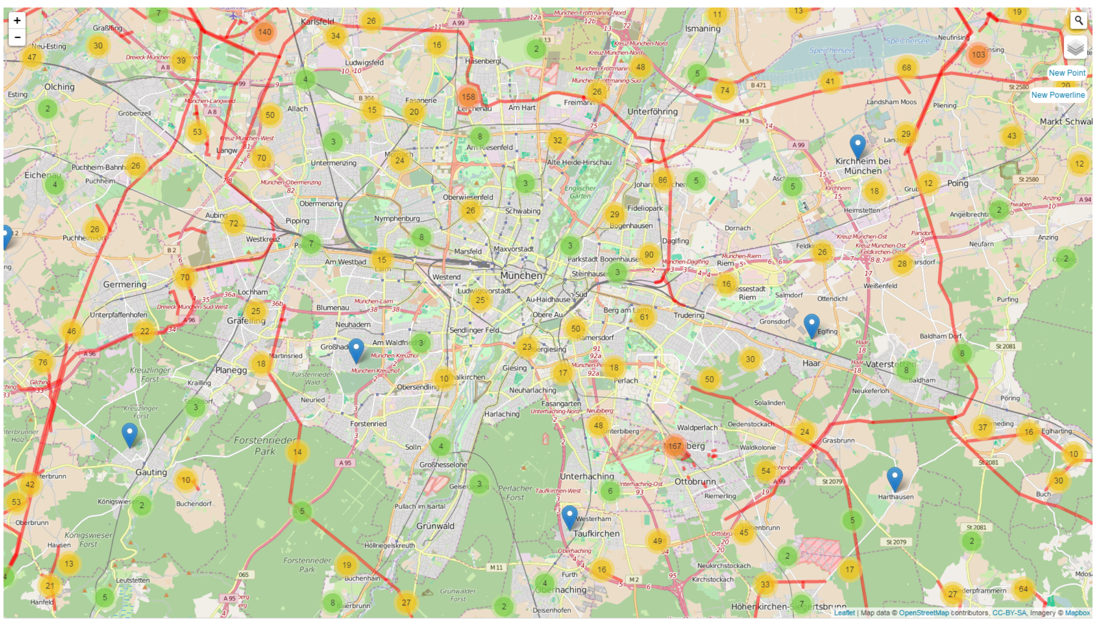


Figure 1.2: OpenGridMap for the area of Munich.

The most valuable data source for OGM is the crowdsourced OpenStreetMap (OSM) project, which is built by a community of mappers that contribute geographical data of various infrastructural networks, including the power grid. The Extensible Markup Language (XML)-organized OSM data is structured by *Relation*-, *Way*-, and *Node*-objects. *Relations* express the relationship between objects, *Ways* describe open or closed paths, and *Nodes*—the smallest structural entities—contain geographical positions. All these objects can carry further descriptive elements, such as attribute- and tag-elements. Power grid relevant objects contain power tags such as <tag k=’power’

[1] http://opengridmap.com/

v='generator' />. For example, an arbitrary power pole is represented by an OSM *Node* object, like <node id='1988705117' lat='48.4687278' lon='9.2840263'><tag k='power' v='pole'/></node>.

OSM represents power circuits in form of relations according to the OSM power routing proposal [15]. Listing 1.1 exemplary depicts the OSM relation *5581803* in its XML format. As the code reveals, a common OSM power circuit relation is composed of two transmission end points, e.g. substations (*role="substation"*) or power plants (*role="plant"*), and several power lines (*role="line"*). The single power line elements represent subparts of the real power line connecting the end points of the transmission section. Fig. 1.3 gives a geographical representation of the relation, showing a power plant and a substation (denoted in red color) connected by a power line (denoted in black color).

Listing 1.1: OSM relation *5581803* in XML format.

```xml
<osm version="0.6" ... >
        <relation id="5581803" ... >
                <member type="way" ref="11341038" role="plant"/>
                <member type="way" ref="234156601" role="line"/>
                ...
                <member type="way" ref="234156595" role="line"/>
                <member type="way" ref="23025610" role="substation"/>
                <tag k="cables" v="3"/>
                <tag k="frequency" v="50"/>
                <tag k="operator" v="50Hertz"/>
                <tag k="route" v="power"/>
                <tag k="type" v="route"/>
                <tag k="voltage" v="380000"/>
        </relation>
</osm>
```

Germany's transmission system is covered by more than 90 per cent with such relations [16], whereas other countries' transmission systems are only little or not at all covered with power circuit relations. The automatic identification of such power circuit relations would be a major progress in the transmission grid inference, since it enables to extract power circuits from geographical information, which in turn allows to construct simulation models. Most of the countries only lack on the power circuit relations, but not on the information (OSM nodes and ways) such relations are built upon. Therefore

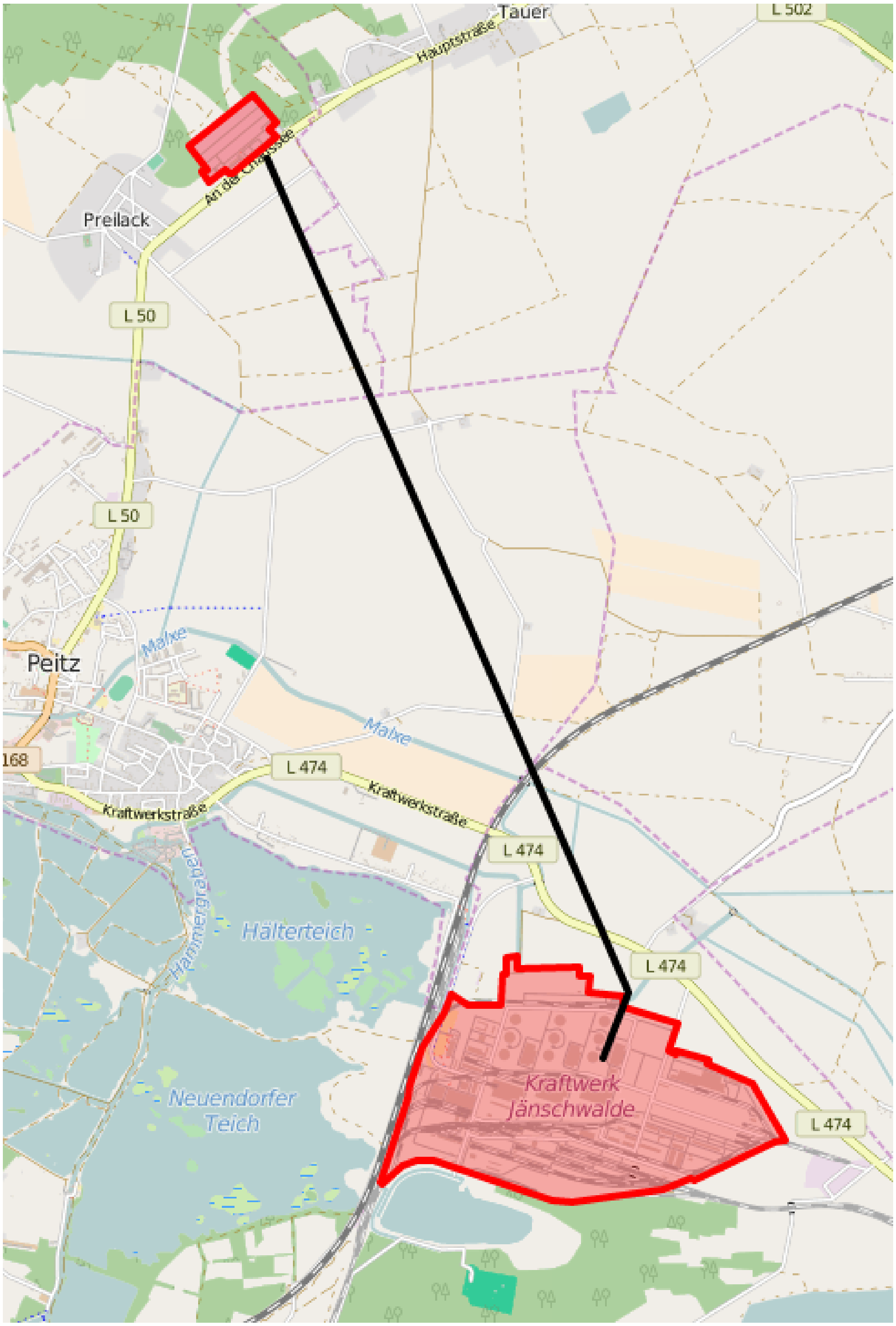


Figure 1.3: OSM relation (id=5581803) representing a power circuit of the transmission system.

one aspect of the thesis is to provide an engine that is able to construct power circuit relations based on the available OSM data. Since OSM data is not reliable, the engine needs to handle incomplete data by e.g. supporting the power circuit inference with decisions based on power grid design rules. The emerging engine can then be used for the inference of OSM power circuit relations of any country or region and the engine's performance is to be validated by comparing already existing German power circuit relations with the generated ones. The evaluation of resulting power system topologies is performed by a comparison with corresponding, already existing maps. Besides other sources, transmission system maps are provided by the European Network of Transmission System Operators for Electricity (ENTSO-E) [17]. Moreover, emerging power grid topologies are to be validated against structural properties of transmission systems, as published by the German Transmission System Operator (TSO) community [18].
Note that we refer to the generated relations as point-to-point (p2p) circuits throughout the thesis, to be able to distinguish them from already existing OSM power circuit relations.

Another focus of the thesis lies on the transformation of the emerged p2p circuits into MATLAB models that allow power flow simulations. To do so, the relation elements have to be represented by parametrized sub-models according to their function in the transmission grid. Hence, the sub-models may vary, whether the element maps a generator, a substation, a power line, or some other power circuit component. Moreover, each sub-model is subject to parametrization, meaning that generators with varying nominal powers are represented with the same sub-model, but parametrized with their corresponding nominal power. The determination of these parameters is not trivial, since OSM data often lacks in required information or appears to be incomplete or inconsistent. Another challenge in the context of transmission system modeling is the estimation of loads attached to substations, which are required for power flow simulations. The majority of existing approaches are designed for load estimations in the distribution grid. Due to the high level of detail this approaches bring along, they are not appropriate for the estimation of transmission grid loads that have to cover large geographical regions. Therefore, other, higher-level load models have to be developed.

1.3 Approach

In general, the thesis' work flow can be divided in three sequential phases, as illustrated at Fig. 1.4. The thesis bases on **OSM power circuit relations**, which have been described at Sec 1.2. As already mentioned, only the German transmission grid is almost completely covered by these relations. To enable the inference of the transmission grid for any country or region, the thesis aims to develop an inference engine, which extracts p2p circuits based on OSM data for the corresponding geographical area. Thereafter, **Common Information Model (CIM) models** are built upon these relations. The CIM provides a framework for the modeling of power grid elements, which is widely used for information exchange in the energy industry [19]. The CIM models are the basis for the generation of parametrized **MATLAB models**, which finally allow to conduct power flow simulations with MATLAB's simulation engine Simulink[2].

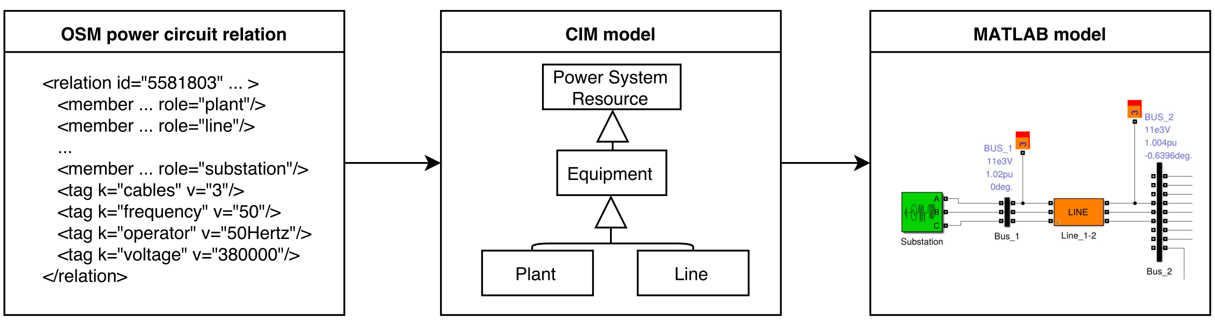


Figure 1.4: Thesis' work flow.

The thesis' work flow, as shortly described above, requires the following steps to take:

1. **Inference of power circuits:** The inference of transmission systems is performed by constructing p2p circuits based on OSM power ways as described at the OSM power routing proposal. OSM already contains an almost complete and reliable representation of the transmission system in form of power circuit relations for Germany. The already existing relations are used to validate the p2p circuits that have been constructed by the inference engine.

2. **Creation of CIM models:** The p2p circuits that emerged from the inference are mapped to CIM models. The CIM models are aligned with a predefined domain model representing the single power circuit elements (entities) and its characteristics (attributes). The domain model considers exactly those characteristics that are

[2] `http://de.mathworks.com/products/simulink/`

relevant for proper power system simulation. As already mentioned in Sec. 1.2, OSM data often lacks in completeness and consistency. To circumvent this problem, we integrate other, more reliable data sources and establish estimation methodologies for electrical characteristics. As mentioned in Sec. 1.2, a challenge in the context of transmission grid modeling is the estimation of loads attached to substations. The approach we follow is to geographically partition the modeled region by means of Voronoi partitions based on substations. The partitioning results in one supply zone for each substation and each zone's load can be estimated based on its population [12]. To complete the estimation, further data sources are consulted building a relation among population and power consumption for considered regions. However, all proposed measures contribute to get fully parametrized, ready-for-simulation, power system models.

3. **Creation of MATLAB models**: The created CIM models are transformed to MATLAB models by accordingly combining model templates of the MATLAB library Simscape Power Systems[3]. These model templates support parametrization and are available for a variety of power grid components. Consequently, the derived power system model can be simulated by means of the simulation engine Simulink, which in turn enables to identify vulnerabilities and resilience issues of the underlying system.

4. **Evaluation of power system models:** The performance of our approach is evaluated by means of several examples that try to reconstruct familiar transmission grids. The topologies are compared with existing transmission system maps, whereas the simulation ability of the derived transmission grid models is evaluated by performing load flow simulations.

5. **OGM integration:** For the integration of this work into the OGM project, MATLAB models for selected transmission systems[4] are free to download at the OGM platform.

1.4 Contributions

By addressing the problems mentioned at Sec. 1.2, following contributions are made:

- The approach proposed at Sec. 1.3 allows to estimate power network topologies

[3] `http://de.mathworks.com/products/simpower/`
[4] `https://github.com/OpenGridMap/transnet/tree/master/models`

based on open, crowdsourced geographical data. In the bigger picture, this approach enables to reveal network topologies that are not intended to be publicly available. Hence, simulations based on the emerged topologies enable research on the vulnerability and resilience of any network.

- The provided inference engine reduces the workload for manual OSM contributors, since it builds relation-like p2p circuits automatically. This inferred p2p circuits can be fed back to OSM to serve the community.
  Moreover, the OGM platform, which is fed with the results of this work, is going to be a valuable playground for researchers and practitioners in corresponding research fields.
- The MATLAB models of transmission grids, which will be free to download at the OGM website, allow to simulate the power flow over arbitrary regions, which is key for identifying energy shortages or surpluses. As a consequence, network operators can coordinate themselves by using this models to overcome power grid issues. Also energy shortcomings that affect single countries can be identified and corresponding measures can be initiated.
- The CIM models, which are created as an intermediate level of abstraction, represent a standardized model for information exchange in the energy industry. We therefore provide a common basis model, which is interpretable by various other related software programs.

1.5 Organization

The thesis is organized as follows: Chap. 2 introduces background information on the design of power grids with a focus on transmission grids. This chapter also gives basic information on the creation of CIMs and MATLAB models. Chap. 3 lists related works on the inference of power grid models and the construction of CIMs. Consequently, Chap. 4 highlights the realization of the single steps of the proposed approach, which are mentioned at Chap. 1.3. The resulting generated transmission grid models are then evaluated in Chap. 5. Finally, the thesis is summarized at Chap. 6 by concluding the results and discussing problems that emerged during the thesis.

Chapter 2

Background

The following chapter discusses various topics that contribute to the general understanding of the thesis. This work is part of the research project OpenGridMap (OGM), which is described at Sec. 2.1. Background information on the design of power grids and relevant transmission grid characteristics are presented in Sec. 2.2. How OSM represents power grids is discussed within Sec. 2.3 in detail. The way CIM models are built is presented in 2.4, whereas Sec. 2.5 covers the power model construction in Simulink.

2.1 OpenGridMap

The aim of the OpenGridMap project is to provide tools for researchers and practitioners to produce realistic input data for simulation studies. In [14], the project has been introduced and the crowdsourcing approach to collect large amounts of power grid data (over 3 million devices collected) has been described.

Since its initial introduction there have been major developments on the project. One of the major goals of the project was the support of data collection and verification. To improve the support for data collection, the OpenGridMap mobile app[1] has been introduced, as shown at Fig. 2.1 [2].

To improve the collection and verification of data also the OpenGridMap web application has been updated, which offers a pan and zoom searchable map seen in Fig. 1.2. The new web application also supports adding, removal, editing and viewing of power devices

[1] https://play.google.com/store/apps/details?id=tanuj.opengridmap&hl=en

Figure 2.1: OpenGridMap's Android app [2].

and their data. Moreover, the web application provides a management platform for the verification of data quality. Every power device data submission must first be verified by an expert-in-the-loop to guarantee data quality. The new web application provides a support for the verification process. An example of a verified submission to OpenGridMap from the Android app can be seen in Fig. 2.2 [2].

Using data available from OpenStreetMap [20] and data collected with the mobile app, the OpenGridMap database has information for over 3 million power devices form all over the world. Thanks to several contributors, the number of collected devices keeps increasing each day [2].

In this work, we extend the original work and present our approach for structuring crowdsourced data to produce power grid simulation models. The long-term vision is to provide such simulation models for free via the OGM website, as illustrated at Fig. 2.3. The platform should provide a mechanism to select a region by specifying a selection box and should enable the user to download the inferred power grid model in various formats.

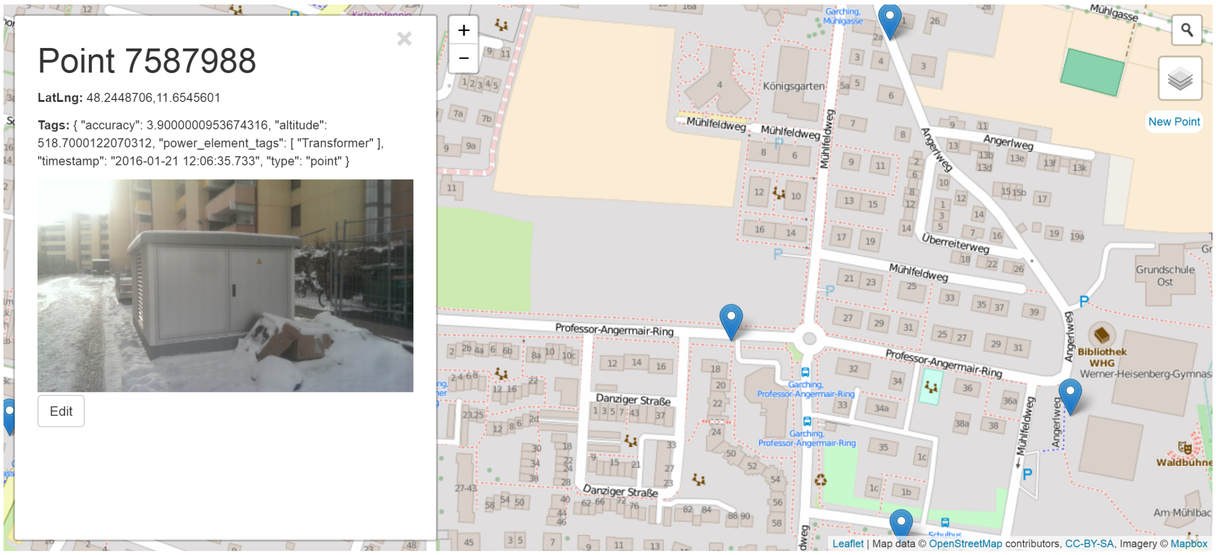


Figure 2.2: OpenGridMap's crowdsourced transformer [2].

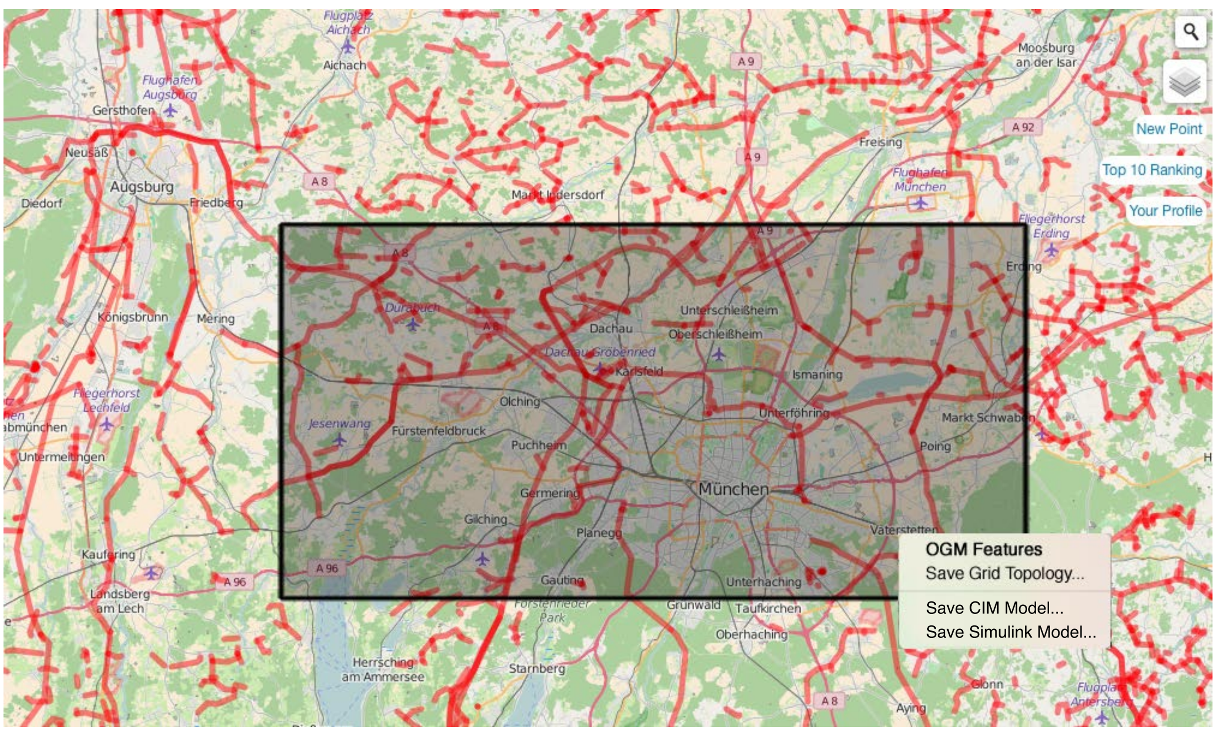


Figure 2.3: Vision: Free-to-download power grid simulation models for any region at OGM website.

2.2 Power Grid Design

Some background information on the design and characteristics of transmission grids is necessary to understand how our inference approach works. The material of this section is aligned with [3].

Power delivery grids can be divided into two general layers: the transmission and the distribution system. The former transports power at high voltages over long distances, whereas the latter carries power at lower voltages over short distances. Literature does not come up with a particular voltage limit, which divides transmission and distribution grids, but [3] declares transmission systems to carry power with a voltage ranging from 60 kilovolts (kV) to 500kV, whereas more local distribution systems work with voltages in the low tens of kV. Distribution systems supply domestic, commercial and industrial customers, while transmission systems supply the distribution system. Transformers serve as boundary between those two systems.

The basic structure and components of the transmission and distribution network are illustrated at Fig. 2.4. The thick vertical lines represent buses, which are the key connection points, especially installed in substations and generators. For reasons of simplicity, the diagram uses one-line connections instead of actual three-phase lines. The power flows from left to right at voltage levels that are typical for the USA.
Two plants generate power at 21kV on the left side of the diagram. Transformers connect the generators to the transmission system and step up the voltage to 230kV. The circuit breakers on both sides of the transformer are indicated with squares. Their main function is to separate the generators from the transmission system on demand.
The figure shows high-voltage transmission on two voltage levels: 230kV and 60kV. Both transmission systems are connected with step-down transformers at a transmission substation in the center of the diagram. At the right side, the transformer of a distribution substation steps down the voltage from 60kV to 12kV, indicating the start of the distribution system. So-called distribution feeders branch out from the substation to supply local regions. Finally, another distribution transformer steps down the voltage at each of the three phases to 120V for the end customer.

The topology of transmission and distribution systems is another important characteristic. While transmission systems tend to have a more interconnected network configuration, distribution systems' lines branch out sequentially in a so-called radial configuration. In a network configuration any two points are usually connected by more than one path, such

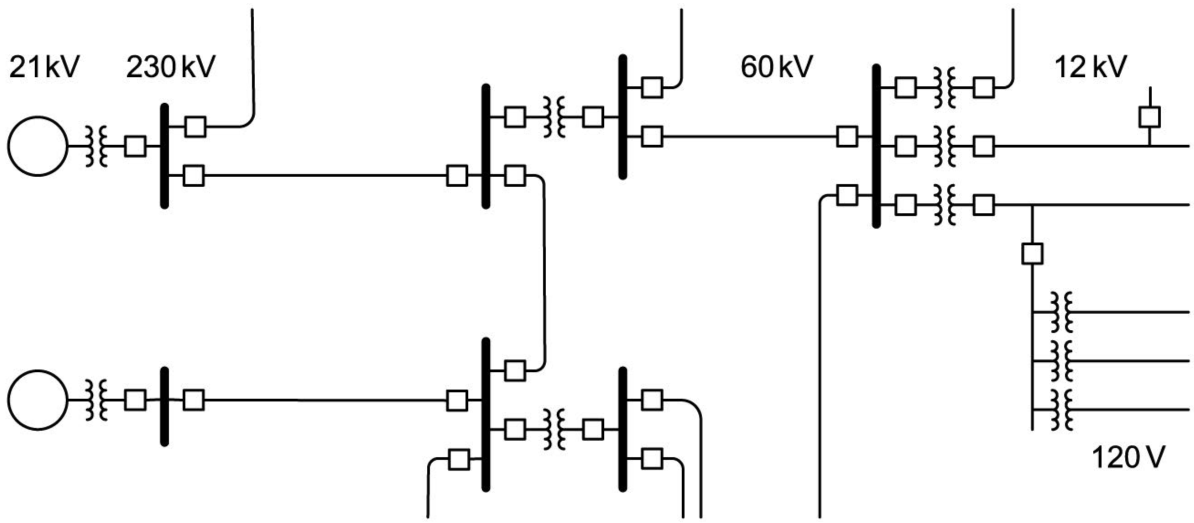


Figure 2.4: Basic power system structure in one-line diagram [3].

that loops arise. This design characteristic is important to counteract fluctuating loads, since it enables to fed power from different locations into the transmission system and additionally allows the power to flow in both directions. The network configuration also provides redundancy, which is a crucial advantage in the design of transmission systems. Multiple lines allow the power to flow at various paths, such that a damaged line does not affect the operation as long as the remaining lines have enough capacity to carry the load.

Throughout the power system, substations exist in diverse sizes. Their main purpose is to build an interface between various levels of the power system and to enable switching and reconfiguration of connections among transmission and distribution lines. Different high-voltage circuits meet at a largest-scale transmission substation. A distribution substation is attached to the transmission system on one side and supplies a few or dozens of distribution circuits on the other side depending on the area.
The major hardware within a substation is the transformer, which connects the high- and low-voltage parts of the grid. Additionally, substations have circuit breakers and switches. While breakers open automatically in the case of a fault, switches are installed to enable a planned and controlled separation of system parts.

In general, the transmission system is operated with three-phase transmission, meaning that each transmission line has three conductors. A transmission line is considered to be balanced, which means it is assumed that every phase serves the same load. The effect

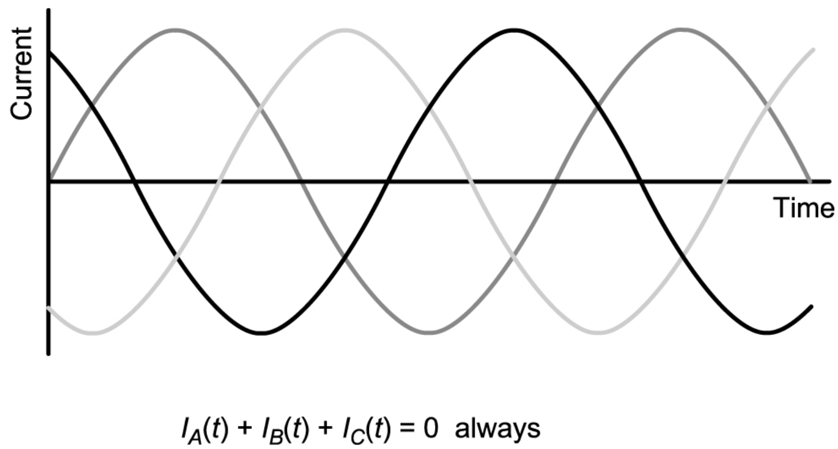


Figure 2.5: Phases' currents over time in a balanced three-phase a.c. system[3].

of this characteristic is that the phase currents always add up to zero, as illustrated at Fig. 2.5. This condition is only met if the phases are separated in timing by 120° (one-third of a cycle) and their respective amplitudes are equal. Therefore, a balanced system does not need a forth, combined conductor for residual currents. In practice, it is difficult ot attach the exact same load to each phase. Nevertheless, system operators try to distribute the load equally, such that only small residual currents arise, which can flow through the soil over the grounding of power poles.

In this work, we infer only the substation-to-substation and generator-to-substation connections of the transmission system. To do so, only the locations and voltage levels of transformers and generators are required. Therefore, we restrict the inference of substation equipments to transformers only and waive further hardware like buses, circuit breakers and switches for reasons of simplicity.
Since we do not consider distribution systems, we can assume to have a balanced system, which is especially important for the generation of the Simulink model, as described in Sec. 4.3.

Finally, a closer look is taken on the electrical characteristics of the power equipments, used in the simulation models, to allow a realistic hardware parametrization.
The conductors of overhead transmission and distribution lines are typically made of lightweight and relatively cheap aluminum. In contrast, underground lines consist of more expensive, better-conducting, isolated copper for reasons of heat dissipation. Therefore underground and overhead power lines have different electrical properties. In general, the impedance of a power line is a combination of resistance, reactance and capacitance, as the equivalent circuit diagram of a transmission line segment at Fig. 2.6 illustrates: The

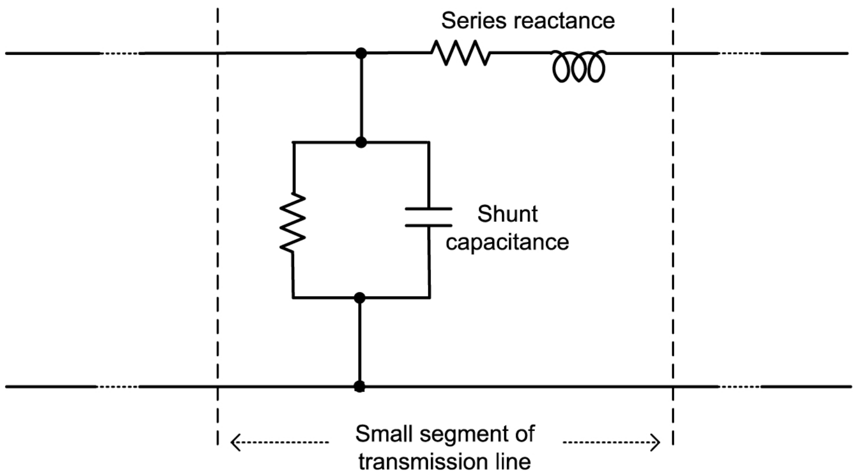


Figure 2.6: Transmission-line modeling [3].

inductance is in series, whereas the capacitance is in parallel to the resistance. Hence, a transmission line's electrical parameters can be characterized in terms of an equivalent resistance, inductance, and capacitance on a per-meter basis.
While the electrical characteristics of common power line types can be assumed, the parameters of transformers and generators are highly individual and therefore not considered in this thesis.

2.3 OSM Power Infrastructure Model

OpenStreetMap (OSM) is a crowdsourcing project, which targets to collect free-to-use geographical data to enable the construction of maps and tools that rely on geographical data [20].

OSM can be classified as a Geographic Information System (GIS), which is designed to collect, manage, and present geographical data. The reference coordinate system that is used for representing geographical information is the World Geodetic System (WGS) 84, which is a standard earth projection for use in cartography and navigation, including the Geographic Positioning System (GPS). The WGS84 projection is based on the WGS-84-ellipsoid, which approximates the surface of the earth. This ellipsoid is defined within a spatial Cartesian coordinate system and determines locations with geographical length and width angles (longitudes and latitudes). The latitude specifies the north-south position of a point, which ranges from 0° at the equator to 90° at the north or south pole. In contrast, the longitude defines the east-west position of a point, which ranges

from 0° at the prime meridian to +180° eastward and −180° westward [6].
To represent a geographical location on a two-dimensional map, a map projection is used that converts geodetic coordinates (e.g. longitudes and latitudes) to two-dimensional coordinates on a map. Popular web services like Google Maps use the Spherical Mercator (code EPSG:3857) projection, just as OSM, for the presentation of geographical information on web platforms. The projection describes coordinates (x/y) in meters and is therefore useful for geographical distance approximations [6]. Another advantage of a protected system is that geometric calculations—for example, whether a point is geometrically within a polygon—are considerably simpler compared to calculations based on the three-dimensional surface of the earth [7]. Therefore, this work uses the Spherical Mercator projection to performantly inference power grid topologies and to plot resulting grid topology models on two-dimensional planes.

The modeling of power grids in OSM has been initiated with the project *Power networks*, which has the following aims [21]: "This is a project to map and document all power facilities on the planet. This could be power lines, towers and substations. But also gas or oil lines."
Various sub-projects for different countries have been started to adapt the general power infrastructure modeling guidelines to the more specific characteristics of particular countries. The countries that explicitly established sub-projects are Austria, Czech Republic, Denmark, France, Germany, Italy, Moldova, Netherlands, Philippines, South Africa, Switzerland, United Kingdom, and the United States according to the *Power networks* project website [21]. Note that the lack of a separate sub-project does not mean that there is no power grid data available for the respective country.
Fig. 2.7 depicts the geographical distribution of power-tagged OSM objects, showing that OSM power data exists all over the world. Especially, North America, Europe and wide parts of Asia are densely covered with OSM data.

As in Sec. 1.2 already briefly mentioned, OSM builds upon three basic structure objects in its data model [22]:

- **Nodes** map single points on the earth's surface and are therefore the basic geographic object of OSM. A node's location is defined by geographical coordinates—latitudes and longitudes—in its attribute list. A node can either be used to define standalone features, or to define the shape or path of an OSM way. Lst. 2.1 illustrates the node with unique identifier (OSM-ID) *1057565829*.

Listing 2.1: Simplified representation of OSM node *1057565829* in XML format.

```xml
<node id="1057565829" lat="48.2609453" lon="13.0025783"/>
```

Figure 2.7: Geographical distribution of the *power* key [4].

- **Ways** are spanned by an ordered list of OSM nodes and can be either open or closed. The node list of a closed way begins and ends with the same node. This characteristic makes it useful to model objects with areas such as buildings, substations, or parks. In contrast, open ways are used to model unclosed routes, such as roads, railway tracks, and power lines. Lst. 2.2 depicts the OSM way *91086220*, which represents a substation of the German transmission system: It is spanned by a list of nodes, whereby the node *1057565829* appears at the beginning and at the end of the list; hence, it is a closed way.

Listing 2.2: Simplified representation of OSM way *91086220* in XML format.

```xml
<way id="91086220">
        <nd ref="1057565829"/>
        <nd ref="1057565847"/>
        ...
        <nd ref="1833918810"/>
        <nd ref="1057565829"/>
        <tag k="name" v="Umspannwerk Simbach"/>
        <tag k="operator" v="Tennet, E.on Netz"/>
        <tag k="power" v="substation"/>
        <tag k="substation" v="transmission"/>
        <tag k="voltage" v="220000"/>
</way>
```

- **Relations** are used to define geographical or logical relationships and consist of an ordered list of members. Members can be nodes, ways or other relations. Optionally, members can have a role attribute that describes its role within the relation. The usage of relations range from bus routes, over power transmission routes to country borders. The structure of the relation *3732378* is shown at Lst. 2.3, indicating a power transmission route. The route is built upon way members: The first and last way members are substations (*role="substation"*), linked with a set of members that represent interconnected power lines (*role="line"*); hence the relation describes a substation-to-substation power circuit, similar to Fig. 1.3.

Listing 2.3: Simplified representation of OSM relation *3732378* in XML format.

```
<relation id="3732378">
<member type="way" ref="77694901" role="substation"/>
<member type="way" ref="77694907" role="line"/>
...
<member type="way" ref="91086222" role="line"/>
<member type="way" ref="91086220" role="substation"/>
<tag k="cables" v="3"/>
<tag k="from" v="Altheim"/>
<tag k="ref" v="234"/>
<tag k="route" v="power"/>
<tag k="to" v="Simbach"/>
<tag k="type" v="route"/>
<tag k="voltage" v="220000"/>
<tag k="wires" v="single"/>
</relation>
```

The listings above make extensive use of key-value pairs, which are referred to as tags. Tags describe specific properties of OSM objects (nodes, ways, and relations). Both key and value fields are free text fields. However, in practice the OSM community usually defines conventions for proper tagging. One such convention represents the *power routing proposal* [15] for the tagging of power circuits. These power circuits should be modeled as OSM relations with the following most important tags:

- Tags *type=route* and *route=power* to indicate power circuit relations
- Tag *voltage=v* to indicate the voltage levels *v* in a semicolon-separated list, e.g. *v=380000;220000*
- Tag *operator=o* to indicate the operator *o*

- Tag *cables=c* to indicate the number c of electrically separable cables; one three-phase a.c. circuit has three cables
- Tag *wires=n* to indicate the number n of conductors of a cable
- Tag *frequency=f* to indicate the frequency f of the power circuit

One relation that is aligned with the *power routing proposal* is illustrated at Lst. 2.3. Usually, such relations model power circuits from station (generator or substation) to station. Therefore, these relations would be of great help for the inference of power grid topologies, since they not only offer geographical, but also logical relationships among the relation-included OSM objects. Nevertheless, other ways to find such relationships have to be found, since this OSM power circuit relations are only well established in Germany and lack in power grid coverage for the rest of the world.

The thesis' transmission grid topology inference approach, which is described in detail at Sec. 4.1, waives the power circuit relations and tries to construct relation-similar station-to-station connections based on power-relevant OSM ways. As Lst. 2.2 reveals, such ways are tagged with *power=** tags. To get a feeling for the amount of power-relevant data in OSM, Tab. 2.1 lists the most common power tags of OSM objects. It shows the value, the total frequency, the share with respect to all power-tagged OSM objects, the OSM object type, and the official description for each tag. As the table illustrates, power *towers* and *poles* occur the most with a total share of more than 90 per cent of all OSM power objects. The remaining rows of the table represent OSM way objects that are used for the inference of station-to-station connections. Stations are modeled as closed ways and can in terms of this thesis either be a *generator* or a *substation* object. The *line*, *minor_line*, and *cable* objects serve as connections between stations and are normally, but not necessarily, spanned by *towers* and *poles*.

The thesis' inference approach is highly dependent on the comprehensive tagging of power lines. Lines that do not have a voltage tag are not considered in the inference. To assess the amount of unusable data, Tab. 2.2 lists the coverage of tags that represent electrical properties: 54 per cent of all power lines are *voltage*-tagged. The rest, 46 per cent of power lines, is not considered by our inference approach, since it relies on voltage-tagged power lines. In fact, this problem is a major limitation for a comprehensive power grid inference and demonstrates the importance of consistent tagging.

Table 2.1: Most common *power*-tagged OSM objects as of 20 July, 2016 [4].

Value	Frequency	Share (%)	Type	Description
tower	8336265	65.59	node	For towers or pylons carrying high voltage electricity cables. Normally constructed from steel latticework.
pole	3134321	24.66	node	Power line pole or small mast
line	390143	3.07	way	High-voltage power lines supported by towers or pylons.
generator	291042	2.29	way	A "device" converting kind of power to another.
substation	178427	1.40	way	Power substation.
minor_line	169474	1.33	way	Minor power cables, supported by poles and not towers/pylons.
cable	8967	0.07	way	Cabling of low / high voltage lines

Table 2.2: Coverage of electrical property tags at OSM power lines and stations [7].

Power Line Tag	Coverage (%)	Power Station Tag	Coverage (%)
frequency	31	frequency	9
voltage	54	voltage	15
cables	55	operator	32

2.4 CIM Modeling

In the course of energy market deregulations, the need for a common data exchange medium between power system operators has dramatically increased to guarantee a reliable operation of interconnected power networks among different utilities. For this reason, the Common Information Model (CIM) has been introduced, which represents a standard model for the information exchange in the energy industry. According to the *CIM Survey 2010* [23] held by the Gartner Group, the CIM is widely applied in the electric utility domain of Europe and the USA. One of its main purpose is to provide a vocabulary by means of a large domain ontology for intra-utility communication [24].

The CIM has been initially designed by the Electric Power Research Institute (EPRI) in the mid 90s. While it was originally developed for a specific problem, it became more comprehensive over time and grew to an international standard, which was published by the International Electrotechnical Commission (IEC). The CIM is still under development by the Technical Committee (TC) 57 and its working groups [24]. However, the most important CIM-related standards are the *IEC 61970* [25] and the

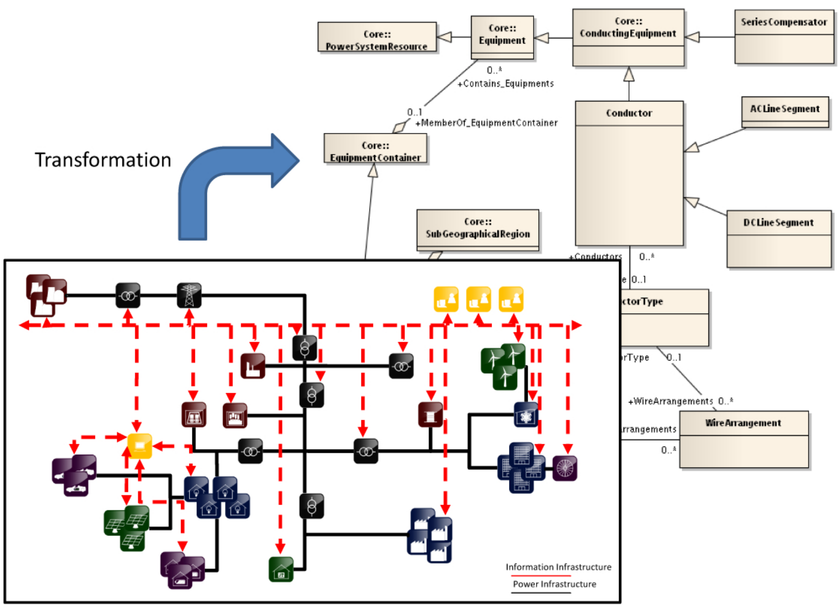


Figure 2.8: Transformation of real world power objects into the CIM data model [5].

IEC 61968 [26]:

- The **IEC 61970** "Energy management system application program interface (EMS-API)" [25] contains the basic data model with the objects required for power network modeling. Furthermore, the following CIM serialization strategies are specified to make this model usable: Resource Description Framework (RDF), Extensible Markup Language (XML), and Web Ontology Language (OWL).
- The **IEC 61968** "Application integration at electric utilities - System interfaces for distribution management" [26] defines virtual objects for business use cases like billing, markets, or network extension plans.

The core of CIM is its data model, which allows to convert real power domain objects into a Unified Modeling Language (UML) data model, as illustrated at Fig. 2.8. The objects are derived from different UML packages that focus on different domains. The single packages are organized in three data model subparts: *IEC 61970-301–"CIM Base"*, *IEC 61968-11–"Distribution Information Exchange Model"*, and *IEC 62325-301–"Data Model for Market Extension"*. The *CIM Base* part contains the integral components to model power equipments, while the *Data Model for Market Exchange* part provides building blocks for the modeling of locations [5]. The components provided by both parts are sufficient for the scope of this work.

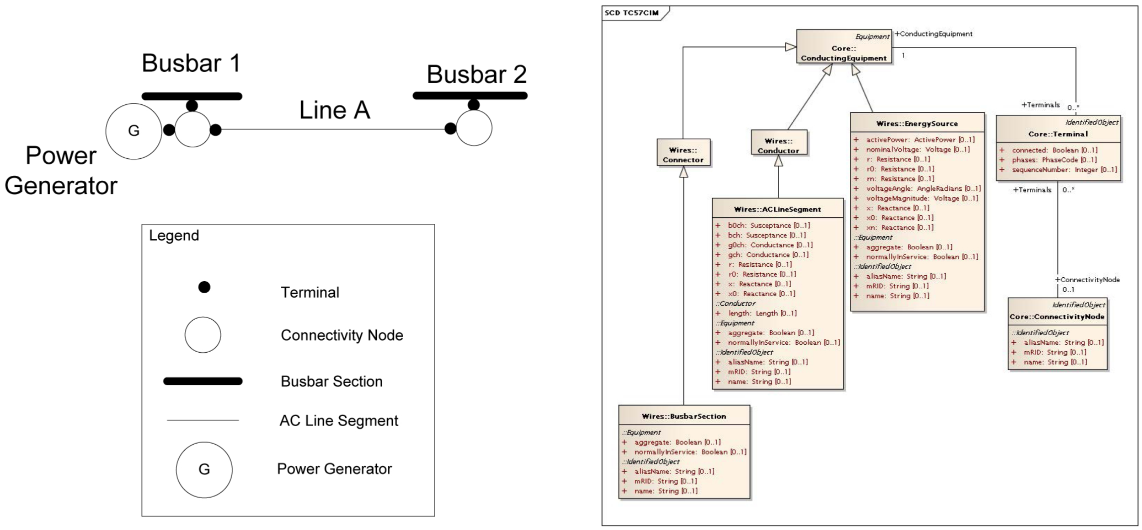
As explained above, CIM relies on a comprehensive data model. The variety of objects allows to model any power-related real-world relationship. However, the extensive amount of objects complicates the application of CIM models. Hence, CIM profiles have been introduced for a more specific, context-dependent level of abstraction. CIM profiles represent subsets of the CIM data model, which include all relevant classes and associations for the considered scenario. They further allow to define class attributes as either being optional or mandatory and can restrict associations with cardinalities. These profiles are a common means to address specific scenarios with the CIM data model. While CIM profiles can be created individually by enterprises or users, official profiles exist that have a greater audience and therefore a higher influence. According to [5], the most important official profiles are:

- **CPSM:** The Common Power System Model (CPSM) is applied for the exchange of transmission system models in the USA.
- **CDPSM:** The Common Distribution Power System Model (CDPSM) is applied for the exchange of distribution system models in Europe.
- **ENTSO-E:** The ENTSO-E profile is applied for the exchange of transmission system models in Europe.

Since this work focuses on the modeling of transmission systems in Europe, we align our CIM models with the ENTSO-E CIM Profile 1 [9].

Depending on the use case, several standardized serialization strategies exist that map the CIM to an applicable format. The most relevant serialization strategy for this thesis is the XML/RDF standard. It builds upon the Extended Markup Language (XML) and extends XML with semantic information through the use of the Resource Description Framework (RDF), which focuses on the definition of resources. RDF is a meta data interchange standard for the web, which is specified by several World Wide Web Consortium (W3C) recommendations. In the context of the CIM, RDF is used to serialize grid topology information [24].

For a better understanding, the creation of a CIM model in XML/RDF format is discussed based on Fig. 2.9a. The example shows a single line diagram of a little power grid with selected CIM objects: The objects, that represent real, physical components are the *Power Generator*, the *AC Line Segment*, and the *Busbar Section*. The other, virtual CIM objects—the *Terminal* and *ConnectivityNode* objects—are used to connect the physical equipments, because the CIM serialization does not allow to connect physical objects natively.
Starting at the left-hand side, a *Power Generator* and a *Busbar Section* are located at



(a) Single line diagram [5].

(b) Used CIM classes [5].

Figure 2.9: CIM elements of a simple power grid section.

the same location and are connected by a *ConnectivityNode* through their *Terminals*. An *ACLineSegment* connects the generator with another *Busbar Section* on the right-hand side, which represents a substation for example. Note that the connection of the second *Busbar Section* to the *ACLineSegment* requires another *ConnectivityNode* that connects the physical components at its *Terminals*.

To serialize the power grid section of Fig. 2.9a in the RDF format, the data model of Fig. 2.9b is used to map the CIM objects to RDF objects: Except of the power generator, which is mapped to an *EnergySource* object, all other objects are mapped to classes with similar or equal names. With help of associations, the relations between objects can be expressed and therefore allow to model power grid topologies.

Listing 2.4: RDF-serialized CIM *Terminal* [5].

```
<cim:Terminal rdf:ID="X">
        <cim:Terminal.ConductingEquipment rdf:resource="#XA"/>
        <cim:Terminal.ConnectivityNode rdf:resource="#XB"/>
</cim:Terminal>
```

A RDF-serialized CIM *Terminal* is illustrated at Lst. 2.4. Each CIM object is uniquely identified by *rdf:ID*. This identifiers are used to cross-reference other objects. For example, this terminal builds a connection between the *ConductingEquipment* with *rdf:ID="XA"* (e.g. a bus bar) and the *ConnectivityNode* with *rdf:ID="XB"*. The corresponding

ConnectivityNode serialization is depicted at 2.5: No subtags are specified—connections to it are indirectly specified by *ConnectivityNode*-references at *Terminals*.

Listing 2.5: RDF-serialized CIM *ConnectivityNode* [5].

```
<cim:ConnectivityNode rdf:ID="XB"></cim:ConnectivityNode>
```

While the virtual CIM objects have a rather slim structure, physical CIM objects often contain lots of subtags that represent its electrical characteristics. Lst. 2.6 illustrates a power line in form of an *ACLineSegment*, which defines the line's length, resistance, reactance, shunt susceptance, and its name. Another example for a serialized physical CIM object is given at Lst. 2.7, which indicates a serialized power generator with its name and active power.

Listing 2.6: RDF-serialized CIM *ACLineSegment* [5].

```
<cim:ACLineSegment rdf:ID="A A1">
        <cim:Conductor.length>2500</cim:Conductor.length>
        <cim:Conductor.r>0.3125</cim:Conductor.r>
        <cim:Conductor.x>0.28</cim:Conductor.x>
        <cim:Conductor.bch>235.45</cim:Conductor.bch>
        <cim:IdentifiedObject.name>Line A</cim:IdentifiedObject.name>
</cim:ACLineSegment>
```

Listing 2.7: RDF-serialized CIM *EnergySource* [5].

```
<cim:EnergySource rdf:ID="A G">
        <cim:IdentifiedObject.name>Power Generator</cim:IdentifiedObject.name>
        <cim:activePower>400</cim:activePower>
</cim:EnergySource>
```

The exact approach we pursue to generate CIM models based on OSM point-to-point connections is explained in detail at Sec. 4.2.

2.5 Simulink Power System Simulation

> Simulink® is a block diagram environment for multidomain simulation and Model-Based Design. It supports simulation, automatic code generation, and

> continuous test and verification of embedded systems.
>
> Simulink provides a graphical editor, customizable block libraries, and solvers for modeling and simulating dynamic systems. It is integrated with MATLAB®, enabling you to incorporate MATLAB algorithms into models and export simulation results to MATLAB for further analysis. [27]

The modeling of power systems in Simulink is achieved with the Simscape Power Systems™ library, which provides components and analysis tools for the modeling and simulation of electrical power systems [28]. The library offers to model complete electrical networks, which in turn enables to perform stability and vulnerability analysis.

For transmission system components, which are the scope of this work, the Simscape Power Systems library offers configurable block templates [28]:

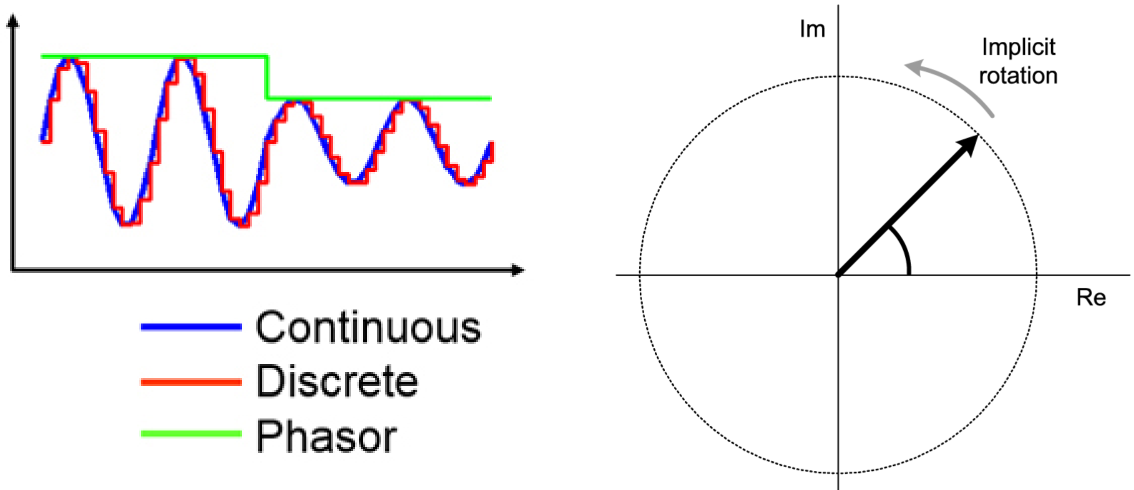
- **Transmission Line** models for single and multiphase overhead transmission lines and underground cables.
- **Transformer** models with different connection types (delta-delta, wye-delta, zigzag-delta-wye, ...) for single and multiphase connection.
- **Load** models in the desired level of granularity.
- **Generator** models in form of synchronous and asynchronous machines.

To build a power system model, the basic component model blocks have to be put on the Simulink plane and parametrized according to their electrical properties. Thereafter, the blocks have to be connected accordingly and a so-called *Powergui* block has to be added to enable the simulation of the power system. The *Powergui* environment offers three different simulation methods for the analysis of power grid models: "

- **Continuous** solution method using Simulink variable-step solvers.
- **Discretization** for solution at fixed time steps.
- **Phasor** solution method using Simulink variable-step solvers.

" [29]

As Fig. 2.10a illustrates, the continuous method performs highly accurate simulations, whereas the accuracy of a discrete simulation is highly dependent on the selection of the time step size. With phasor simulation, a set of algebraic equations at a fixed frequency is applied, instead of differential equations. Since the engine solves much simpler equations, the phasor simulation is much faster and is therefore recommended for large networks. In



(a) Continuous, discrete, or phasor simulation [28].

(b) A phasor [24].

Figure 2.10: Simulation methods with focus on phasor.

contrast to continuous and discrete simulation, phasor simulation only analyzes changes of magnitude and phase for all voltages and currents in phasors [28]. The phasor notation is a simplified way to characterize a sine wave by using its magnitude and phase angle in relation to a reference. A rotating vector or phasor, as depicted at Fig. 2.10b, is a scaled line, whose length represents the magnitude and its direction indicates the phase angle. Phasors are especially useful, if not only one sine cycle is to be analyzed, but rather to assess the effect of many cycles. Therefore, phasors are widely used in the steady state analysis of a.c. power systems [3]. Note that we also use phasors notation for our power system simulations.

Among various power system analysis tools, the Simscape Power Systems library enables to do load flow calculations based on the underlying power system model. Note that load flow and power flow analysis can be used synonymously, whereby power flow is the more general term [3].
Power flow analysis aims to describe the operating state of an entire power system of generators, transmission lines, and loads. It determines unknown quantities based on known quantities, such as the amount of power generated and consumed at certain locations. The quantities of interest are mostly voltages, which have in case of a.c. power systems magnitude and phase angle. With the knowledge of voltages, the currents flowing over the transmission lines can be easily determined. Due to the size and complexity of a real transmission system, there is no closed-form solution for the power flow problem. Only numerical methods can solve the problem by iterative approximation [3].

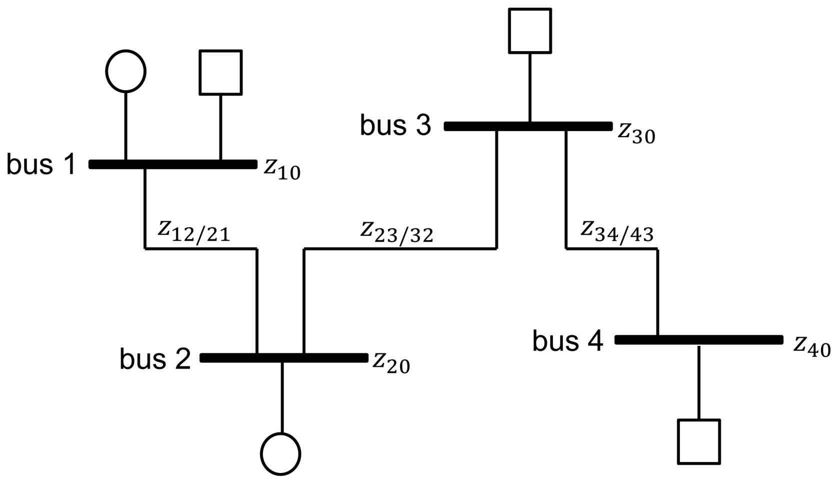


Figure 2.11: One-line diagram for a power system, aligned with [3].

The power flow analysis starts with the identification of electrically distinct points, which are called nodes or buses. The electrical distinction emerges from different electrical potentials at different points of the underlying system. Each bus is acting either as generator that injects power or as load that consumes power at the system. Buses are connected by three-phase transmission lines, which are mostly modeled as single lines, because it is assumed that the same power injection or consumption is happening at each phase. The impedance of a power line is assumed to be constant, since its mostly dependent on its physical characteristics [3].

An example for a one-line diagram is considered at Fig. 2.11. The thick horizontal lines represent buses, circles indicate generators, and squares are loads. The transmission lines' and buses' impedances are indicated by the constants z_x. At this point, known and unknown bus variables have to be determined and brought together in an equation system that has as much distinct equations as it has unknown variables. To completely describe the voltage at any bus in an a.c. system, voltage magnitude and angle have to be specified. Analogously, to describe the currents in each branch, two numbers have to be obtained: the current magnitude and its angle. The same applies for the amount of power transferred at any node of an a.c. circuit: Real and reactive power have to be considered. Hence, an a.c. circuit is completely determined if two pieces of information are known for each bus. In practice, the power injected or consumed from a bus and some bus voltages are known, whereas currents are not known at all [3].
Depending on the variables given, different bus types are distinguished [3]:

- At the **Load Bus**, it is assumed that the load is implicitly given by the consumer's power demand. Therefore real (P) and reactive (Q) power are specified and voltage angle θ and magnitude V are unknown.

- At the **Generator Bus**, it is assumed that the real power P and the generator bus voltage V are known, whereas the voltage angle θ and the reactive power Q are unknown.
- The **Slack Bus** is usually also a generator bus, but varies in power generation to balance the system. Therefore, the real (P) and reactive (Q) power is not given, but instead it is assumed that the voltage angle θ and magnitude V are known. In general, one slack bus is required to enable power flow analysis in a power system.

Assuming that two input variables are given for each bus and all constant characteristics of the system—e.g. transmission line impedances and the system frequency—are known, the system's operating state can be determined. To do so, power flow equations are specified that formulate the electrical relationships between the buses. The set of these equations builds a power flow problem [3].

As already mentioned above, the solution of the power flow problem requires iterative approximation. Several algorithms exist for solving the power flow problem, whereby the most popular ones are the Gauss-Seidel and the Newton-Raphson algorithms. The Simscape Power Systems library uses the Newton-Raphson algorithm to solve the power flow problem [28]. For further information regarding power flow equations and solution methods, the interested reader is referred to [3]. Note that the applicability of our power system simulation models is tested by performing load flow calculations.

Chapter 3

Related Work

The work's approach is based on the combination of two research lines. On the one hand, it builds upon existing works in the inference of power grid topologies [12, 16, 30]. On the other hand it makes use of previous efforts to generate power grid simulation models from CIM files [31, 32, 33]. The novelty of this approach is the combination of both research lines' results to obtain a comprehensive approach that generates simulation-ready power grid models from crowdsourced geographical data. In the following, relevant contributions the approach builds upon are described.

3.1 Power Grid Topology Inference

In [30] a simulation model of the German transmission system was created manually. The identification of grid elemenents, such as generators, substations and power lines was achieved by labeling the 220kV and 380kV grid elements with placemarks on Google Earth. The satellite pictures were not only used to extract the transmission grid topology, but to determine the voltage level of power towers. The shadow of a tower reveals the number of traversers and subsequently the number of parallel conductor lines it carries. This information is sufficient to determine the voltage level of a tower and consequently the voltage level of the power line it is part of. Pursuing this approach, a model of the complete German transmission system was built with Google Earth placemarks. Its equipments have been exported to Microsoft Excel equipment lists to complete the equipment parameters. Based on the equipment lists, a PSS©NETOMAC model has been created using MATLAB. Eventually, the PSS©NETOMAC model has been tested by conducting load flow simulations. The results of the load flow simulations were then fed to a power flow plot, which helped to identify highly loaded tracks.

Similarly to this work, [12] automatically extracted the 380kV transmission grid's topology of the German state Baden-Württemberg using OSM power relations. Based on the emerged model, a dynamic load flow simulation has been performed using the self-developed simulation tool *ePowSim*. The work also included a quality analysis of OSM data evaluating its availability, completeness, correctness, uniquity, and consistency. The analysis resulted in OSM being a valuable source for the topology of power grids. Nonetheless, OSM data is not complete and therefore not sufficient to obtain comprehensive power grid models according to [12]. For this reason, the transmission grid inference integrated other, mostly governmental data sources to compensate for missing OSM data. Furthermore, [12] divided the considered geographical region in Voronoi partitions to enable a dynamic load estimation based on the population-density at each partition. For this task cities with a population of more than 10.000 people have been mapped to a 2D-grid and the population density has been interpolated with a radial gaussian distribution. Dynamic load estimates emerged from combining standard load profiles for households, businesses, and industry with the population density approximations for the considered partition.

SciGRID[16] is an open-source research project for the development of automated power grid model generation methods. SciGRID also relies on OSM power relations to infer the topology of transmission grids. Due to this dependency, SciGRID encourages the establishment of further OSM relations to steadily improve the transmission network coverage. Therefore, the SciGRID website provides a map showing OSM substations all over Europe. To help contributors, the map highlights substations that are not covered by an OSM power relation yet. With the release of GridKit, the project also offers a toolkit for power grid inference that is independent of OSM relations. Instead, it uses any other power-tagged OSM object to construct power grids. The results of these projects are graphs representing the topology of the German transmission network. Substations are mapped as the graph's vertices, whereas power lines are mapped as the graph's edges. The edges are approximated as direct point-to-point connections and do not represent the real course of power lines.

3.2 CIM to Power Grid Simulation Model Mapping

The work in [31] provides solutions for the transformation of CIM models to proprietary data formats that are supported by widely used legacy simulation tools. In particular, the work provided a Java tool that was capable of transforming a CIM model to a PSS©E formatted data model. One of the addressed problems was the non-trivial transition of a node-breaker (CIM model) to a bus-brunch (PSS©E model) topology. Furthermore the work proposed ways to convert literal-stored CIM attribute values to percentage or ratio values by means of mathematical formulas. The conversion was required to enable a problem-free import of CIM files into considered proprietary tools.

Similarly, [32] proposed a CIM model extension to provide all data requirements for power flow simulations of distribution networks. The proposed CIM model modifications were applied while demonstrating the conversion of the IEEE radial test feeder to a CIM model, which in turn fulfilled all requirements for a power flow simulation model. One of the major problems addressed by [32] was the inconvenient CIM representation of unbalanced networks, which are typical for distribution systems. Thus [32] introduced new CIM classes to account for the cumbersome modeling requirements.

A co-simulation framework for smart grid analysis that uses CIM models as intermediate communication media was proposed by [33]. The intermediate CIM model not only covered structural information on the modeled power system, but included data derived from actor-driven SCADA systems. The resulting CIM model was then transformed into the glm format, which is the standard input format for the smart grid analysis and simulation tool GridLab-D. The mapping from the node-breaker (CIM) to the bus-branch (glm) data model format was performed by a self-written Java application. Using the GridLab-D solver, power flow calculations were performed to evaluate the behavior of new equipments.

Chapter 4

Approach

As Sec. 1.3 already introduced, the thesis' work flow passes three major phases. The basic work flow is aligned with Figure 1.4: Sec. 4.1 describes the inference of p2p circuits, which are similar to OSM power circuit relations. Consequently, Sec. 4.2 discusses the CIM model generation process based on p2p circuits. Finally, Sec. 4.3 deals with the generation of Simulink models based on CIM models.

4.1 Power Grid Topology Inference

OSM power circuit relations are a proposed way of modeling the point-to-point connections of power system components. Since only the German transmission system is comprehensively modeled with these relations in OSM, we provide the inference engine Transnet[1] that tries to construct such point-to-point connections automatically. Transnet is a Python script that relies on OSM data in a local database. Before we discuss Transnet in Sec. 4.1.2 in detail, we provide some information on the preparation of OSM data and its import into a local database in Sec. 4.1.1.

4.1.1 Data Preparation

The data preparation work flow is depicted at Fig. 4.1: At first, OSM data for the considered region is downloaded. Thereafter, relevant power infrastructure data is filtered.

[1] `https://github.com/OpenGridMap/transnet`

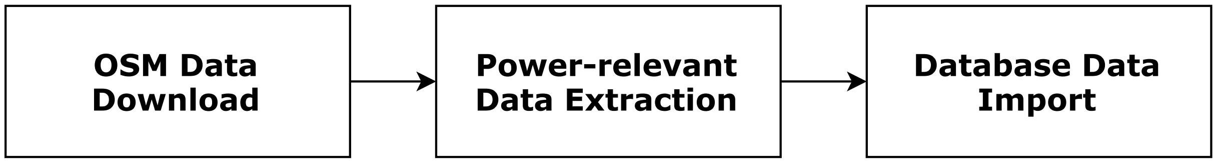


Figure 4.1: Data preparation work flow.

Finally, the filtered data is imported into the database, which serves as the base for Transnet.

1. **OSM Data Download:** OSM data dumps are available at the Geofabrik[2] download server. The server hosts daily-updated OSM data extracts for all continents and its countries. For each data dump, a so-called "poly"-file is available describing the extent of the respective region in a set of coordinates. These sets build geographical (multi-)polygons, which in turn represent the boundaries of continents and countries. The data preparation process requires to download both the data dump file and the respective poly-file.

2. **Power-relevant Data Extraction:** The downloaded data dumps contain a lot of data that is not of interest for the inference of power grids. As already mentioned at Sec. 2.3, power-relevant OSM objects (nodes, ways, and relations) contain OSM tags with a "power" key. To exploit this characteristic, the command line tool osmosis[3] is used, which is capable of filtering OSM data by tags. Hence, calling osmosis with the respective power-tag filters results in extracts that only contain power-relevant OSM objects. Additionally, osmosis accepts poly-files as input to extract only the OSM objects that are geographically within the respective region.

3. **Database Data Import:** OSM uses a PostgreSQL[4] database to persist its data. PostgreSQL is an open-source relational database, which encourages to work with geographical data by means of the extension PostGIS[5]. PostGIS supports geospatial objects and allows to formulate location queries in SQL. The PostgreSQL with PostGIS installation is from now referred to as PostGIS database. The powerful support of spacial data and various useful tools make the PostGIS database our first choice for the local database installation. One useful tool, in this context, is the osm2pgsql[6] tool, which enables to import OSM data dumps into PostGIS databases. The import can be configured with help of so-called "style"-files. A style-file defines

[2] `http://download.geofabrik.de`
[3] `http://wiki.openstreetmap.org/wiki/Osmosis`
[4] `https://www.postgresql.org`
[5] `http://postgis.net`
[6] `http://wiki.openstreetmap.org/wiki/Osm2pgsql`

Table 4.1: Relevant tables produced by osm2pgsql data import.

Table Name	Content
Tables of raw OSM objects	
planet_osm_nodes	Raw OSM node objects
planet_osm_ways	Raw OSM way objects
planet_osm_rels	Raw OSM relation objects
Tables of prepared OSM objects according to style file	
planet_osm_polygon	Closed OSM ways (areas); e.g. substations
planet_osm_line	Open OSM ways; e.g. power lines
planet_osm_point	OSM nodes

how OSM objects end up in the database tables and what columns have to be created from OSM tags. Our style-file[7] configures osm2pgsql to create additional columns for the tags *voltage, name, ref, generator:output:electricity,* and *plant:output:electricity*. While the style file only affects table columns, the database schema produced by osm2pgsql is always the same. The relevant part of the database schema is illustrated at Tab. 4.1. The tool produces three tables that contain the raw OSM objects: *planet_osm_nodes*, *planet_osm_ways*, and *planet_osm_rels*. In contrast, the tables *planet_osm_polygon*, *planet_osm_line*, and *planet_osm_point* contain style-file aligned data. Additionally, the latter tables provide PostGIS geometry columns in Spherical Mercator projection (see Sec. 2.3) for handy spacial SQL queries.

In summary, the data preparation consists of downloading, filtering, and importing power-relevant OSM data into a PostGIS database. At this point, the database is prepared to serve Transnet with OSM data.

4.1.2 Transnet

Sec. 4.1.1 explains how databases with power-relevant OSM data are set up. These databases supply the Transnet inference engine with electrical infrastructure data.
In particular, Transnet relies on OSM way objects that represent substations, generators and power lines. Each OSM object can be identified by its unique OSM-ID. Moreover, OSM objects are labeled with key-value pairs (tags) that provide meta-information. Power-relevant OSM objects have power-tags that inform us about the power equipment type. As Tab. 4.2 illustrates, different literals are valid to indicate a particular equipment type. Substations are identified by having either *'power'='substation'*, *'power'='sub_station'*, or *'power'='station'* tags. The tags *'power'='generator'* or

[7] `https://github.com/OpenGridMap/transnet/blob/master/util/power.style`

Table 4.2: OSM power equipment tags.

Equipment type	Identifying power tag ('power'=?)
substation	'substation', 'sub_station', 'station'
generator	'generator', 'plant'
power line	'line', 'cable', 'minor_line'

'power'='plant' label OSM ways that represent generators. OSM ways having the tags *'power'='line'*, *'power'='cable'*, or *'power'='minor_line'* refer to power lines. Further important tags for the power grid inference are the tags *'voltage'='*'*, *'name'='*'*, and *'generator:output'='*'*, which define the power equipments' operating voltage, name, and generator power output respectively.

Based on these OSM ways, Transnet tries to derive parts of the transmission system in form of point-to-point (p2p) connections. These points can either be OSM substations or generators; we refer to both equipments as stations. As already explained at Sec. 2.3, stations are represented as closed OSM ways, indicating an area. In contrast, power lines are modeled as open OSM ways. The basic idea for the gathering of this p2p connections is to look for power lines that geographically intersect the areas of two different stations within a specific region. The two stations and the connecting power line indicate a section of the power grid, which we refer to as p2p circuit. By merging all these found p2p circuits, we derive the power grid for the considered region. Note that a power line that connects two stations does not necessarily consist of only one, but can consist of several OSM way objects. Fig. 4.2 depicts such an incidence, by showing the OSM way objects that connect the stations in different colors: The first and the last OSM way objects are denoted in black, whereas the OSM way object in the middle is denoted in red color. All these parts of the connection are labeled as power lines and together indicate a 380kV transmission line.

Depending on the input parameters shown at Tab. 4.3, Transnet's power grid inference builds upon different sets of OSM objects: The script can be called either for a substation (input parameter *-s*) or a region (input parameters *-p* and *-b*). If the parameter *-s* is specified, the script is called for a substation and requires the so-called *SSID*, which is the substation's unique OSM-ID. In this case, Transnet considers all OSM stations and power lines with at most 300 kilometers distance from the considered substation. To execute the script for a particular region, either the path to a poly-file with input parameter *-p*, or a bounding polygon in Well-known text (WKT)[8] format can be specified. In this mode, all stations and power lines within the respective region are considered for the inference.

[8] `http://www.opengeospatial.org/standards/wkt-crs`

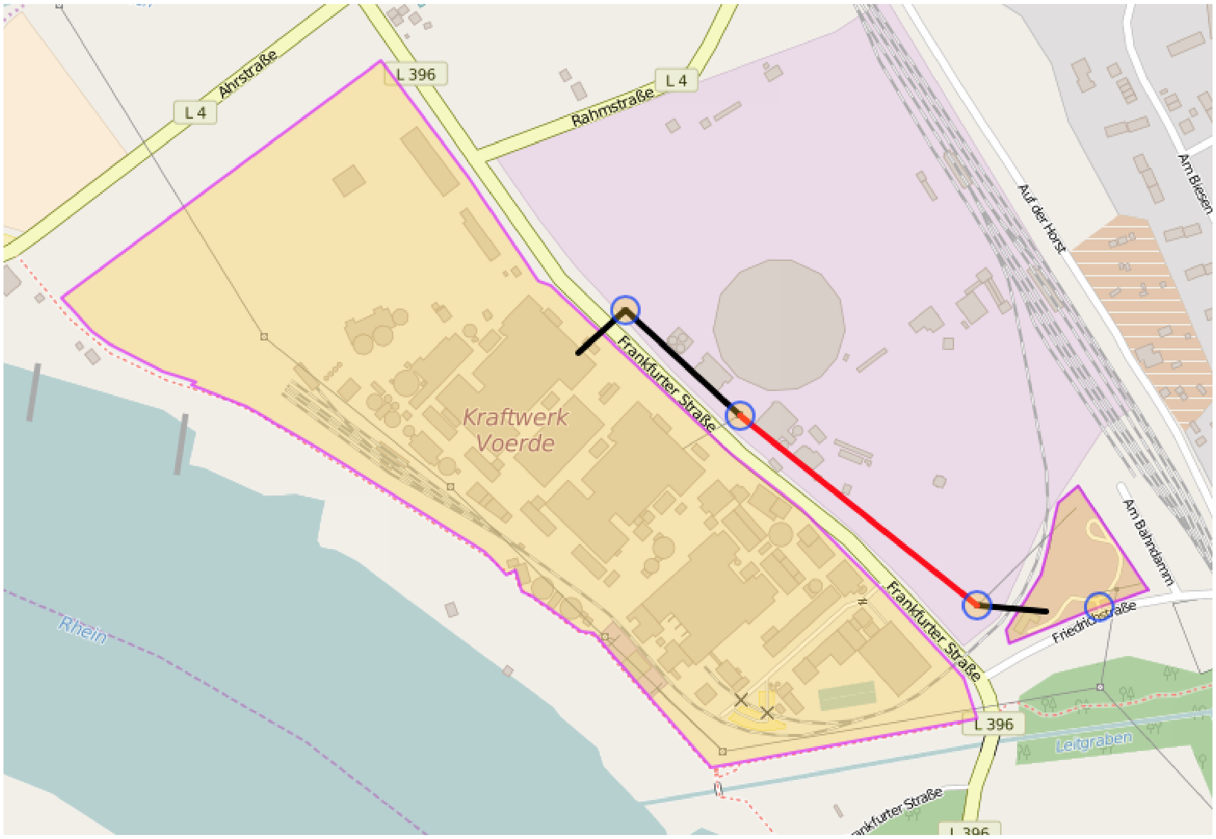


Figure 4.2: 380kV transmission line consisting of several OSM way objects.

Finally, the parameter -V specifies the voltage levels of the power grid for which the inference should be done. The single voltage levels have to be specified in Volts, separated by a pipe symbol (|). For example, the parameter value "380000|220000|110000" lets Transnet infer the transmission grid for the high voltage levels 380kV, 220kV, and 110kV.

Given the set of considered stations, the set of power lines, and the considered voltage levels as input, the Transnet inference, described at Alg. 1, works as follows: A list C is created, which serves as container for found p2p circuits. For each voltage level, power lines with the corresponding operating voltage are queried and put into set L. Furthermore, stations that are geographically intersected with the power lines of set L are queried and put into set S. For each substation s in S, find power lines in L that geographically intersect station s. For each line l in L, create an initial p2p power circuit list c consisting of the station s and the line l such that c = *[s,l]*. Now follow the line l recursively by calling the recursive function *follow_power_line()* until another station *s'* is found. The result of the function call is the list of found p2p circuits *FC* that emerged from power line *l*. A p2p circuit is a list of OSM objects in the form of $[s, l, ...l', s']$. The OSM objects s and *s'* indicate the transmission section end points in

Table 4.3: Important Transnet input parameters.

Name	Parameter	Description
VOLTAGE_LEVELS	-V	Infer the power grid for the specified voltage levels.
SSID	-s	Infer the power grid for a particular substation by entering its SSID. The substation ID (SSID) corresponds to the substation's OSM-ID.
POLY	-p	Infer the power grid for a particular region, defined by the specified poly file.
BOUNDING_POLYGON	-b	Infer the power grid for a particular region, defined by the specified polygon in WKT format.

form of stations, whereas *l, ..., l'* represent OSM ways that are parts of the complete power line. Thereafter, the power circuits of FC are copied to the container of all found p2p circuits C. Before the algorithm continues with the inference for the next voltage level, duplicate p2p circuits are removed from the list C. Duplicates are encountered by comparing the p2p circuit endpoints. After the deletion of duplicates, a connection between station s and station *s'* exists only once in C for all *s,s'* of S. After the inference has been performed for all voltage levels, all found p2p circuits reside in container C. Note that if Transnet is called for a single substation (input parameter -*s*), the inference is only done for the corresponding substation.

The introduced algorithm is prone to be stuck in power line loops. To counteract this problem, Transnet remembers the OSM nodes of already passed power lines during the inference. Once it encounters a node that has been already covered, it stops the inference branch for the line that happened to run into a loop. The same principle applies to pseudo-circuits that start and end with the same station. When the inference for a pseudo-circuit arrives at a station, which is the same as the circuit's start station, the inference is stopped and the corresponding circuit path is dropped.
Another weakness of the algorithm is that it infers each p2p circuit twice, since the inference is done for both, the start and end station of a p2p circuit. For this reason, Transnet remembers the power lines from which valid p2p circuits already could be inferred for each station. Assume a p2p circuit in the form $[s, l, ...l', s']$ has already been found during the inference. Stations s and *s'* indicate the start and end station and *l*, *l'* represent the lines that intersect station s and *s'* respectively. Furthermore, assume that Transnet remembers the intersecting power lines from which valid p2p circuits could be

Algorithm 1: Power grid topology inference

```
Data: OSM stations, OSM power lines, and voltage levels
Result: List of found p2p circuits
C = []; // container list for found p2p circuits
foreach voltage level v do
    L = set of power lines with corresponding voltage level v;
    S = set of stations that are geographically intersected by any line of L;
    foreach station s in S do
        SL = set of power lines in L that geographically intersect station s;
        foreach line l in SL do
            c = [s,l]; // init power circuit with s and l
            FC = follow_power_line(S, L, s, l, c);
            C.update(FC); // update C with found circuits in FC
    Remove duplicate p2p connections from C;
function follow_power_line(stations S, lines L, start_station s, from_line l, circuit c)
    if line l intersects another station s' of S; s' != s then
        // found a p2p circuit
        c.append(s');
        return [c];
    else
        C = []; // container list for found p2p circuits
        CL = set of power lines in L that are connected to line l;
        foreach cl in CL do
            c.append(cl);
            FC = follow_power_line(S, L, s, cl, c);
            C.update(FC); // update C with found circuits in FC
        return C;
```

already inferred for each station. Therefore, Transnet knows that a p2p circuit has been already found for power line *l* at station *s* and for power line *l'* at station *s'*. When the inference is done for station *s'*, Transnet remembers that a p2p circuit has been already extracted for power line *l'* and therefore skips the inference for this particular line. This mechanism prevents Transnet from inferring each p2p circuit twice.

Transnet is implemented as a Python script and makes heavily use of the Python package Shapely[9], which supports the manipulation and analysis of geometric objects in the Cartesian plane. For performance reasons, the package converts geographical elements to geometric objects, since geometric calculations are much simpler and therefore faster than geographical calculations. Among various other functions, Shapely provides the

[9] `https://pypi.python.org/pypi/Shapely`

possibility to check whether two geometries intersect. This function is used by Transnet to determine whether the end points of a power line intersect the area of any station. Moreover, this functionality helps to identify whether two power lines are connected, which is exploited by Transnet to follow power line paths.

Transnet focuses on the extraction of transmission system p2p circuits. Hence, it does not consider further power components like busbars or switches within substations. The inner-substation components are not relevant for the inference of transmission system topologies and are therefore not considered in this work.

4.2 CIM Generation

Background information on CIM and the modeling process is given at Sec. 2.4. For the purpose of this work, the CIM model serves as an intermediary product, from which further models can be built upon. The CIM model is a well-known way to model power systems and can be used in various commercial software tools to allow power system simulations. To be applicable to commercial software, the CIM model has to meet certain syntactical and semantical requirements. These requirements can be met with CIM profiles, which can be seen as guidelines for proper CIM construction. Throughout the thesis, the ENTSO-E CIM Profile 1 (see [9]) is used, which has been released to allow a standardized modeling of European transmission grids. Among the CIM objects to use, the profile also determines whether object attributes are mandatory or optional. To account for the mandatory parameters, Sec. 4.2.2 describes the references and methods we use for parameter estimations. In Sec. 4.2.1 we list the used CIM objects and demonstrate how these objects are combined to build a CIM model that represents transmission systems.

4.2.1 Model Construction

The CIM model generation is also implemented with a Python script using the Python package PyCIM[10]. PyCIM is a Python implementation of the CIM and brings along the following features: "

- Support for IEC 61970 15v13 and IEC 61968 11v05,
- Legacy support for IEC 61970 14v15 and IEC 61968 10v31,

[10] `https://pypi.python.org/pypi/PyCIM`

Table 4.4: Used CIM classes and OSM-to-CIM object mapping.

Purpose	**CIM Object**	**Description**
Substation	Substation	Container for substation equipment
	PowerTransformer	Transformer in a Substation
	TransformerWinding	Winding of a PowerTransformer
Generator	GeneratingUnit	Container for generator equipment
	SynchronousMachine	Synchr. generator in a GeneratingUnit
Power Line	ACLineSegment	Power line
Load	EnergyConsumer	Load
	LoadResponseCharacteristic	Load's characteristic response
Connection	Terminal	ConductingEquipment connection point
	ConnectivityNode	Node, connecting two or more Terminals

- Profiles of the CIM, including: - Common Power Systems Model (CPSM) (CIM v14) - Common Distribution Power System Model (CDPSM) (CIM v14 and v15) - European Network of Transmission System Operators for Electricity (ENTSO-E) (CIM v14),
- Class and attribute documentation integrated as Python doc-strings,
- Transparent bi-directional reference handling using Python properties,
- CIM RDF/XML parsing and serialisation according to IEC 61970-552.

" [34]

As already mentioned, our CIM models are aligned with the ENTSO-E CIM Profile 1, for which PyCIM provides a separate sub-package containing Python classes for each CIM object of this profile. Additionally, PyCIM is capable of persisting CIM models in the RDF/XML format, which is a standard model format for data interchange on the Web. For this reason, the end product of the CIM model generation process is a CIM file in RDF/XML format.

To transform the p2p circuits that we extracted in Sec. 4.1 to CIM models, the members of the circuits have to be mapped to corresponding CIM objects. Tab. 4.4 shows the set of used CIM classes that are required for a simple representation of the modeled transmission grid sections. The transformation Python script simply maps all p2p circuit members (OSM substations, generators, and power lines) to the corresponding CIM class equivalents. A OSM substation is therefore modeled as a CIM *Substation* instance containing a *PowerTransformer* instance, which has as much *TransformerWindings* instances as its voltage levels. Each *TransformerWinding* instance provides a *Terminal* instance that is to be connected to another equipment. An OSM generator is modeled as a CIM *GeneratingUnit* instance containing a *SynchronousMachine* instance, which also provides a *Terminal* instance for connections. The power lines that connect two transmission end points are modeled as *ACLineSegment* instances, which provide *Terminal* instances on both sides. To construct a connection amongst terminals, *ConnectivityNode* instances are used, which can connect two or more terminals. For example, the connection between *TransformerWinding* instances of two different substations A and B is represented with the following CIM objects chain: *TransformerWinding(A) - Terminal - ConnectivityNode - Terminal - ACLineSegment - Terminal - ConnectivityNode - Terminal - TransformerWinding(B)*. To construct valid power circuits, loads have to be attached. Loads are modeled with *EnergyConsumer* instances that rely on *LoadResponseCharacteristic* instances. *LoadResponseCharacteristic* instances characterize the dynamic behavior of loads. For example, the behavior of a load demand can follow a mathematical function or can just be static with a constant number of Watts. For reasons of simplicity, we configure loads to have a static behavior with a constant power demand.

The algorithm used for the construction of CIM files is shown at Alg. 2: A container O is initialized, which collects the created CIM objects. Each time a CIM object is created, it is automatically appended to O. As data input serve the extracted p2p circuits. Each p2p circuit has a list of members that represent the actual power circuit and the circuit's voltage level. The members list contains Python objects in the form *[s1,l,s2]*. The members *s1* and *s2* indicate the circuit end points, which can either be substation or generator objects. The member l represents the connecting power line object. For each p2p circuit, a container N is initialized, which gathers the *ConnectivityNodes* of the circuit end points to enable a connection through *ACLineSegments* later on. For both of the stations *s1* and *s2*, check whether the station represents a substation or a generator. If the station is a substation, verify if it is already covered in the CIM object container O. If O does not contain the particular substation, create a new CIM *Substation* object o. Thereafter, create a new CIM *PowerTransformer* instance that refers to its parent substation o. Now, check if the newly-created or already existing *PowerTransformer* already has a *TransformerWinding* with the voltage that corresponds to the p2p circuit's voltage. If this is not the case, create a new *TransformerWinding* instance for the respective voltage and let it refer to its parent transformer. Continue with adding a *ConnectivityNode* to the *TransformerWinding*'s terminal and add it to the *ConnectivityNode* container N.

If the station is a generator, verify if it is already covered in the CIM object container O. If O does not contain the particular generator, create a new CIM *GeneratingUnit* object

Algorithm 2: CIM construction algorithm.

Data: List of p2p circuits C; each circuit contains a OSM object member list *[s1,l,s2]* and the circuit voltage

Result: CIM file

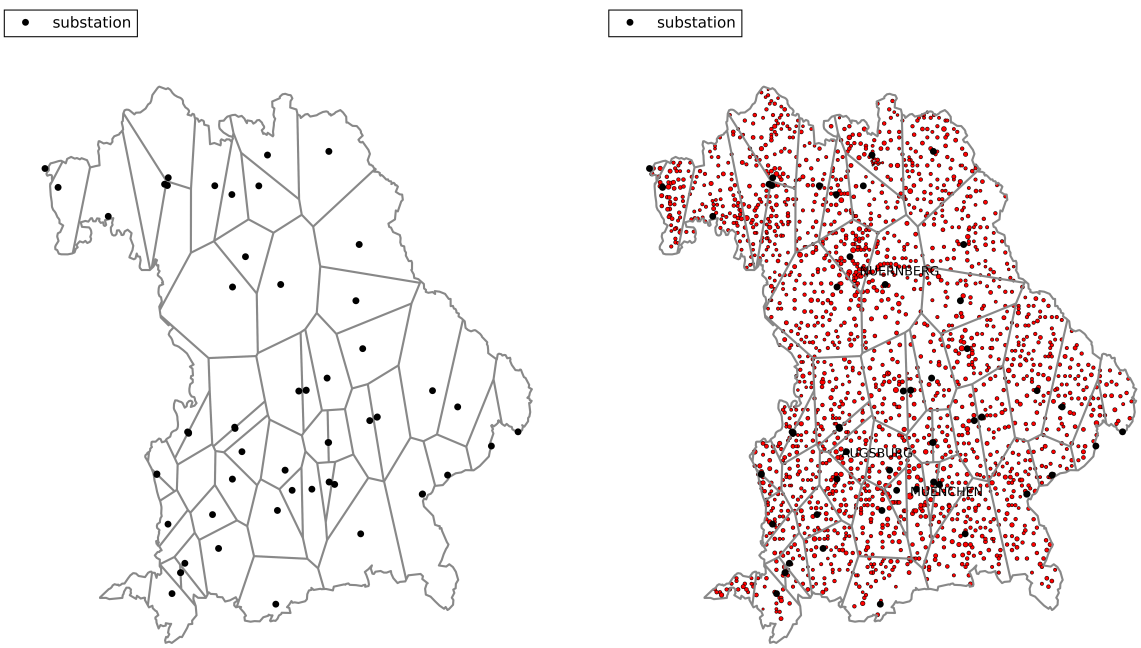
```
O = []; // container list for CIM objects; + indicates addition to list
foreach circuit c in C do
    N = []; // container for ConnectivityNodes
    foreach s in [s1, s2] do
        if s is a substation then
            o = O.get(s);
            if o is None then
                o = new CIM Substation()+;
                new CIM PowerTransformer(o)+;
            p = o.getPowerTransformer();
            tw = p.getTransformerWinding(c.voltage);
            if tw is None then
                tw = new CIM TransformerWinding(p, c.voltage)+;
            N.append(add_connectivity_node(O, tw));
        else
            // s is a generator
            o = O.get(s);
            if o is None then
                o = new CIM GeneratingUnit()+;
                new CIM SynchronousMachine(o)+;
            g = o.getSynchronousMachine();
            N.append(add_connectivity_node(O, g));
    create new CIM ACLineSegment(N)+;
foreach CIM object o in O do
    if o is a PowerTransformer then
        if o.numWindings() < 2 then
            v = voltage of TransformerWinding 1;
            new CIM TransformerWinding(o, v)+;
        tw = p.getTransformerWindingWithLowestVoltage();
        n = add_connectivity_node(O, tw);
        e = new CIM EnergyConsumer(new CIM LoadResponseCharacteristic())+;
        new CIM Terminal(e, n)+;
write O to CIM file;

function add_connectivity_node(CIM object container O, conducting equipment e)
    n = new CIM ConnectivityNode()+;
    new CIM Terminal(e, n)+;
    return n;
```

o. Thereafter, create a new *SynchronousMachine* instance, which refers to its parent o. Now add a *ConnectivityNode* to the *SynchronousMachine*'s terminal and add it to the *ConnectivityNode* container N.
Once both stations, *s1* and *s2* have been modeled and its respective *ConnectivityNodes* are included in N, create a new CIM *ACLineSegment* object, which uses the *ConnectivityNodes* of N as connection points.
Once all p2p circuits have been processed, start with attaching loads to every substation. Therefore, check for each *PowerTransformer* in the container O whether it has less than two transformer windings. If so, create a new CIM *TransformerWinding* object with the voltage of the already existing transformer winding. Note that this behavior ensures that transformers are modeled with at least two transformer windings. Now get the transformer winding with the lowest voltage and add a *ConnectivityNode* to it. Continue with the creation of a CIM *EnergyConsumer* object that relies on a CIM *LoadResponseCharacteristic* object. Finish the load attachment by creating a *Terminal* instance for the *EnergyConsumer*, which refers to the previously created *ConnectivityNode*.
Finally, publish the CIM model by writing all CIM objects that are contained in O to a CIM file in XML/RDF format.

4.2.2 Parameterization

So far, the basic CIM objects and the combination of those to form a CIM model have been described. All these used CIM classes are defined in the CIM standard IEC 61970 [25]. To add scientific value to the CIM, the single CIM objects have to be parametrized by means of CIM object attributes. Tab. A.2 lists the used CIM objects and their mandatory attributes according to the ENTSO-E CIM Profile 1 [9]. These attributes often describe electrical characteristics, which are mostly not given by OSM tags. The only electrical properties we can derive from OSM are the power line voltages and the generator outputs. The substation transformer winding voltages are implicitly determined by the voltage levels of power lines directly attached to it. All other electrical properties have to be estimated or looked up in appropriate reference literature. Transmission System Operators (TSOs) might have estimates for the loads attached to their substations. However, these estimates are mostly not publicly available and relying on region-dependent data sources would diminish the general applicability of the thesis' approach.



(a) Without villages and cities. **(b)** With villages and cities.

Figure 4.3: Voronoi partitions based on substations in the region of Bavaria.

To estimate loads that are attached to substations, we pursue a similar approach to [12]. Given a region and a set of substations within this region, we search for an adequate partitioning of this region to have one supply partition for each substation. Within each partition, we count the population and multiply the total population with the average per-head power consumption of an citizen of the respective country. This methodology results in a load estimation for the supply region of each substation.

Assume a 2-D plane indicating a specific region and a set of points (substation locations) mapped on the plane. One way to find distinct partitions for given points on a 2-D plane is to construct a Voronoi Diagram. A Voronoi Diagram is a disassembly of space into regions, which are determined by a predefined set of points (so-called "centers") within the respective region. Each region is specified by exactly one center and contains all points of the space, which are closer to the center of the region than to any other center according to the Euclidean Distance. Such regions are also referred to as Voronoi Regions. The set of all points, which have more than one closest center, build the borders of the regions and consequently the Voronoi Diagram [35].
Fig. 4.3a illustrates an applied Voronoi partitioning on the region of Bavaria. The

substations, marked in black, represent the Voronoi Region centers and the gray lines indicate the emerging Voronoi Diagram. Normally, the area of the Voronoi Regions would be infinite at the exterior of the considered space. To counteract this issue, the affected regions are cut at the exterior, which results in an exact overlay with Voronoi Regions over the considered space. In the case of Fig. 4.3a, the area of Bavaria is exactly partitioned by Voronoi Regions. Each substation is assigned a region, for which it is the authoritative power supply.

The next step towards load estimates for substations is to approximate the power consumption within the substation regions. In fact, one of the main indicators for the power demand of a region is its population density [36]. Although various factors influence the power demand, we restrict our load estimations to be only dependent on the population of the considered region. Therefore, data sources have to be acquired that provide population figures for geographical regions. One such data source is the free-of-charge OpenGeoDB[11], which provides administrative data (postal codes, areas, inhabitants) for villages and cities in Central Europe [37]. The OpenGeoDB platform provides its data in form of MySQL[12] data dumps, which can be easily imported to local MySQL databases. In addition to the administrative data, OpenGeoDB also provides geographical coordinates of village and city locations.
Given the OpenGeoDB data, we can visually map these locations (marked in red) onto the area of Bavaria, as illustrated at Fig. 4.3b. Based on the geographical location, each city or village can be assigned to a substation partition. This mapping enables the determination of the total population within a partition by simply counting all village and city inhabitants within the respective area.
To get the power demand for any substation region, we multiply the region's population by the average power consumption of a German. According to [38], a German's average energy consumption has been 7.381 kilowatt-hours (kWh) in the year 2015. To retrieve the power consumption in Watts, Eq. 4.1 has to be applied, which converts the Energy E, specified in kWh, to Power P, specified in watts (W), for the time period t, specified in hours h. This formula applied to the German average per-head energy consumption for the year 2015 results in an average per-head power consumption of approximately 0.843 Watts ($1000 * 7.381 kWh/(365 * 40) h \simeq 0.843 W$).

$$P = \frac{1000 * E}{t} \tag{4.1}$$

[11]http://opengeodb.giswiki.org/wiki/OpenGeoDB
[12]https://www.mysql.com

With the knowledge of the population within a substation partition and the per-head power consumption of a citizen, the load estimation is calculated according to Eq. 4.2: The load L_s, attached to substation s, is estimated by multiplying the substation region's population P_s with the average per-head power consumption c.

$$L_s = P_s * c \tag{4.2}$$

Although this load estimation approach is quite practical, it brings along several problems that are discussed at Chap. 6. However, the value L_s is used to parametrize constant substation loads. In terms of CIM, the value of L_s is mapped to the *pConstantPower* attribute of the corresponding CIM *LoadResponseCharacteristic* object. Also, note that the attribute *exponentModel* is always set to *'false'* in order to indicate a constant power load.

The electrical properties of transmission lines have been theoretically discussed at Sec. 2.2. Considering the mandatory CIM attributes for an *ACLineSegment* object at Tab. A.2, the electrical characterization of a power line requires to specify at least the positive sequence shunt susceptance *bch*, the positive sequence series resistance r, and the positive sequence series reactance x. Tab. A.1 lists the operating parameters of a standard 110kV power line. For reasons of simplicity, we use these characteristics for the parameterization of all power lines. Note that the values are specified in per-kilometer units. Therefore, we also make use of the optional *ACLineSegment* attribute *length* to specify the length of power lines, which in turn enables to calculate the total resistance, reactance, and susceptance.

The electrical parametrization of transformers, or especially transformer windings is difficult, since OSM provides no information on the equipment types installed in substations. For this reason, most of the required attributes are filled with default values of Simscape Power Systems block components, as Tab. A.2 shows. One of the parameters that can be determined is the rated voltage attribute *ratedU*, which is specified with the voltage of the connected power line. Also, the attribute *windingType* can be determined depending on the transformer winding voltage. The value range is limited to *'primary'*, *'secondary'*, and *'ternary'*, since the Simscape Power Systems library offers transformer blocks with at most three windings. The winding with the highest voltage level indicates the *primary* winding type. The *'secondary'* winding type is assigned to the winding with the second highest voltage. In case of a three-winding transformer, the winding with the lowest voltage level is labeled with the *'ternary'* winding type.

According to Tab. A.2, *GeneratingUnits* are characterized with three attributes: *maxOperatingP*, *minOperatingP*, and *nominalP*. The nominal power, represented by attribute *nominalP*, is sporadically available through OSM tags. If the nominal power is specified in OSM, the value is mapped to the attribute *nominalP*. Otherwise, *nominalP* is specified with a very high number of Watts to represent a generator that is attached to a slack bus. As explained at Sec. 2.5, at least one slack bus is required to enable power flow analysis. In fact, most of the OSM generators lack in the nominal voltage specification. Therefore, the power grid simulation models might have various slack buses. The parameters *maxOperatingP* and *minOperatingP* are power limits dispatchers can specify for a generator. Since we have no information about these values, we assume them to be equal to the nominal power. In terms of this work, the *SynchronousMachine* within a *GeneratingUnit* is the actual generator. Considering Tab. A.2, the mandatory attributes to specify for a *SynchronousMachine* are the *operatingMode*, *qPercent*, *r*, *ratedS*, and *Type*. While the parameters *operatingMode*, *qPercent*, *r* and *Type*, can be specified statically with *'generator'*, *'100'*, the Simulink default value, and *'generator'*, the parameter *ratedS* is dynamically specified with its nominal power according to OSM tags.

The last entry of Tab. A.2 describes the CIM *BaseVoltage* object. This object is an optional attribute of the class *ConductingEquipment*, which is a superclass for all conduction equipment classes, including the *TransformerWinding* class, the *SynchronousMachine* class, the *EnergyConsumer* class, and the *ACLineSegment* class. We exploit this fact to assign each conducting CIM object a voltage level by specifying the attribute *BaseVoltage*. The CIM *BaseVoltage* class itself requires two mandatory attributes: *isDC* and *nominalVoltage*. While *'false'* is constantly assigned to *isDC*, the *nominalVoltage* attribute carries the actual voltage level information in Watts.

Other, non-electrical meta-information derived by OSM tags include substation and generator names. Moreover, the geographical features of the OSM objects are exploited to measure the length of power lines and to reproduce the geographical arrangement at the succeeding generation of Simulink models. While the length of power lines can be simply mapped to the optional *length* parameter of CIM *ACLineSegment* objects, the geographical coordinates of the circuit components require the usage of two CIM classes of standard IEC 61968 [26]: The CIM *PositionPoint* class can represent geographical coordinates by means of its attributes *xPosition* and *yPosition*. This *PostionPoints* are accepted by CIM *Location* objects to define a location. CIM *Location* objects can be

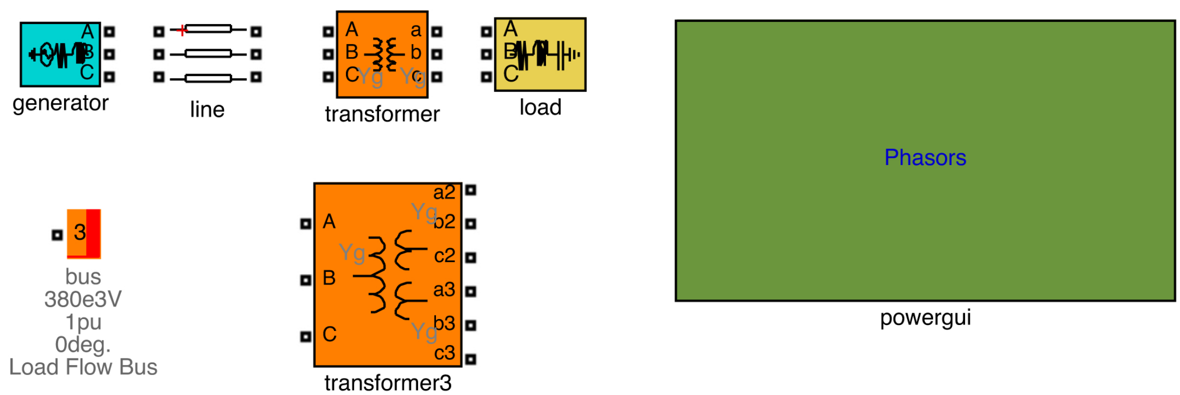


Figure 4.4: Preconfigured Simscape Power Systems block templates.

attached as attribute to any equipment by means of the inherited attribute *Location* of the superclass *PowerSystemResource*. For each power system component (substation, generator, or power line), a *Location* object is specified using the geographic centroid coordinates of the component.

4.3 Simulink Model Generation

Simulink and its ability to build power system simulation models with the Simscape Power Systems library is described at Sec. 2.5. In this section, the power system model generation process based on CIMs is examined.

The power system modeling is performed by a MATLAB script, which programmatically constructs Simulink models based on CIM files. The results are ready-for-simulation Simulink models consisting of various interconnected Simscape Power Systems blocks.

Only a few blocks are required to represent high voltage power systems. These blocks have been preconfigured and stored in a block template file. Fig. 4.4 depicts the blocks of the template file: The blue *generator* block refers to a *Three-Phase Source* of the Simscape Power Systems library. Power *lines* are represented by a *Distributed Parameters Line* block. In the scope of this work, transformers can either have two or three windings. For this reason, two different orange-colored blocks are specified indicating two-winding and three-winding transformers. The former refers to a *Three-Phase Transformer (Two Windings)* and the latter to a *Three-Phase Transformer (Three Windings)* block respectively. The yellow-colored *load* block refers to the *Three-Phase Series RLC Load* block. To measure voltage magnitudes and phase angles, the *bus* block is required, which is a *Load Flow Bus* block. Finally, the *powergui* block (marked in green) is necessary to enable power system simulation tools such as the load flow analysis.

With help of this block templates, Simulink models are built based on CIM files. The algorithm for the automatic generation of the Simulink models is illustrated at Alg. 3: At first, obtain the list of CIM objects O by parsing the given CIM (XML) file. Note that the xml_read[13] MATLAB script has been exploited for the parsing task. Then create a new Simulink model m and open it. To have a reference for the positioning of Simulink blocks, remember the geographical centroid c, which is available by means of an extra CIM *PositionPoint* object. Now start to place blocks on the model based on the CIM objects.

For each CIM *PowerTransformer* in O, copy either the three-winding or two-winding transformer block template depending on the number of windings to the model m. Position the emerging transformer block relative to the model centroid c by considering the corresponding CIM *Location* object.

For each *TransformerWinding* in O, find the already created parent transformer block and specify either its *Winding1*, *Winding2*, or *Winding3* parameter depending on whether the transformer winding is of *primary*, *secondary*, or *ternary* type respectively. The *WindingX* parameters require a list of values in the form *[Vrms,R,L]*, where *Vrms* stands for the phase-to-phase voltage, R is the resistance, and L indicates the inductance of the transformer windings. These values are provided by the corresponding CIM *TransformerWinding* attributes *ratedU* (voltage), r (resistance), and x (reactance, including inductance).

For each SynchronousMachine s in O, copy the generator block template to the model m. Parametrize the emerging block according to s by specifying the block parameters *Voltage*, *BaseVoltage*, and *Pref* with the CIM attributes *BaseVoltage*, *BaseVoltage*, and *ratedS* respectively. Thereafter, position the block relative to the model centroid c by considering the corresponding CIM *Location* object.

For each *EnergyConsumer* e in O, copy the load block template to the model m. Parametrize the emerging block according to e by specifying the block parameters *NominalVoltage* and *ActivePower* with the CIM attributes *BaseVoltage* and *pConstantPower* of the *LoadResponseCharacteristic* object that refers to e. Thereafter, position the block relative to the model centroid c by considering the corresponding CIM *Location* object.

For each *ACLineSegment* l in O, copy the line block template to the model m.

[13] `https://www.mathworks.com/matlabcentral/fileexchange/12907-xml-io-tools/content/xml_read.m`

Algorithm 3: Simulink power system model generation algorithm.

Data: CIM file

Result: Simulink power system model

```
O = CIM objects; // derived by means of CIM (XML) parser
m = new Simulink model;
open(m);
c = model centroid; // retrieved from PositionPoint in O
foreach PowerTransformer t in O do
    if t has three windings then
        b = add three-winding transformer block to model m;
    else
        b = add two-winding transformer block to model m;
    set position of b relative to model centroid c;
foreach TransformerWinding w in O do
    t = find transfomer block of w;
    parametrize t according to attributes of w;
foreach SynchronousMachine s in O do
    b = add generator block to model m;
    parametrize b according to attributes of s;
    set position of b relative to model centroid c;
foreach EnergyConsumer e in O do
    b = add load block to model m;
    parametrize b according to attributes of e;
foreach ACLineSegment l in O do
    b = add line block to model m;
    parametrize b according to attributes of l;
    set position of b relative to model centroid c;
foreach ConnectivityNode n in O do
    E = find set of equipments that are connected trough n;
    connect equipments of E with each other;
    b = add a bus block to model m;
    connect the first equipment of E with b;
foreach BaseVoltage v in O do
    create a legend entry for v;
b = add powergui block to model m;
set position of b to be on the left top of m;
save model m;
export m to compatible version for older MATLAB versions;
```

Parametrize the emerging block according to l by specifying the block parameters *Resistance*, *Capacitance*, *Inductance*, and *Length* with the CIM attributes r, bch, x, and *length* respectively. Thereafter, position the block relative to the model centroid c by considering the corresponding CIM *Location* object.
For each *ConnectivityNode* n in O, find the set of equipments E that are connected through n. Connect the equipments by adding a line for each phase from the first equipment to any other equipment of E. Additionally, manipulate the orientation of the involved model blocks such that straight connections between block terminals can be established. Moreover, copy the bus block template to the model m and connect it with the first equipment of E.
Finally, create a legend on the left top of the model by creating an entry for each *BaseVoltage* v in O. Furthermore, copy the powergui block template and place it on the left top of the model m. Save the model m and export it to a version that is compatible with older MATLAB distributions.
Note that each Simulink block has to have a unique name, under which it can be referenced. To ensure name uniqueness we define the format *EquipmentType_OSMID_OSMName* for block names. The *EquipmentType* can either be T for transformer, G for generator or L for load. The *OSMID* is the OSM way's unique OSM-ID and the *OSMName* is derived from OSM *name* tags. Power lines do not have a unique identifier, since they are composed of several OSM way objects having different OSM-IDs. Instead, power lines can be uniquely identified by combining the OSM-IDs of the equipments they connect and the voltage level of the power circuit they belong. The OSM-IDs and the circuit voltage are implicitly given by the names of connectivity nodes, which span the power line. Hence, power line names are specified in the form *CircuitName_CN1_CN2*, where *CN1* and *CN2* stand for the spanning connectivity node names and the *CircuitName* is derived from the *name* tag of the first OSM power line part.

An important requirement for the generated Simulink models is that the block positions on the Simulink plane have to be similar to the corresponding geographical locations of the power system equipments. As already mentioned, the CIM files contain *PositionPoint* objects that specify geographical coordinates in form of latitude and longitude angles for each equipment. To map these locations to the 2-D Simulink plane, we convert latitudes and longitudes to Cartesian x- and y-coordinates considering the Azimuthal Equidistant projection. In contrast to other 2-D projections, the Azimuthal Equidistant projection maps the x- and y-coordinates relative to a specified center, from which distances and directions are correctly shown [6]. Fig 4.5 shows such a projection with the north pole as center. The dynamic center of this projection is especially useful for the mapping of our

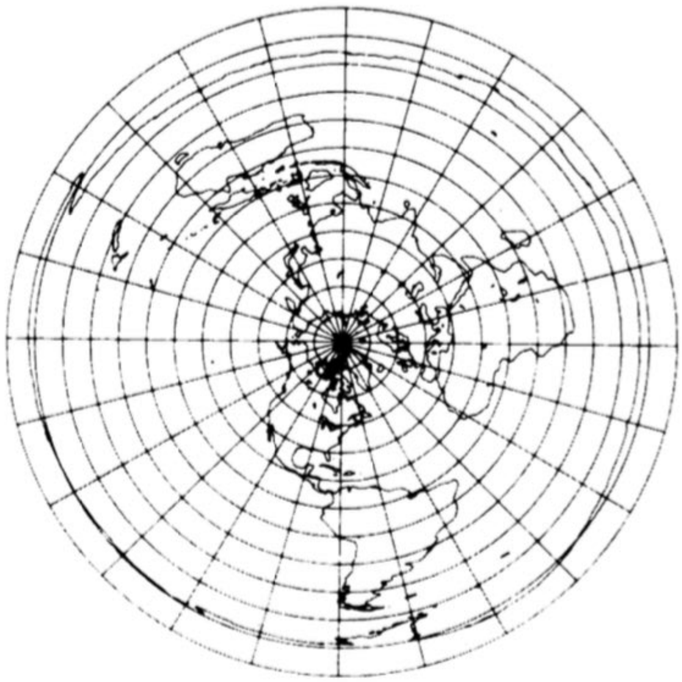


Figure 4.5: Polar aspect Azimuthal Equidistant projection [6].

models, since we mostly only represent power systems of countries or single continents, but not of the whole world. Therefore, our models do not get larger as the considered region's boundary and the model's blocks are positioned with respect to the dynamic center. In this work, the dynamic center is defined as the geographical center of all substation locations. Note that the conversion from latitudes and longitudes to x- and y-coordinates is performed by the Spherical2AzimuthalEquidistant[14] MATLAB script.

[14] `https://www.mathworks.com/matlabcentral/fileexchange/28848-spherical-to-azimuthal-equidistant/content/Spherical2AzimuthalEquidistant/Spherical2AzimuthalEquidistant.m`

Chapter 5

Evaluation

In Chap. 4, we introduced an approach to automatically infer power system simulation models based on crowdsourced OSM data. Within this chapter, the approach's ability to generate such models and the quality of the derived power system models are assessed. To illustrate the capability of the approach, we examine high voltage power grid inference examples with increasing level of difficulty.

Starting with the inference of a simple generator-to-substation connection at Sec. 5.1, we evaluate the ability to infer all connections emerging from a particular substation at Sec. 5.2. Thereafter, we continue with the inference of larger power systems, including the Bavarian and German transmission grids at Sec. 5.3 and Sec. 5.4 respectively. These sections also include approaches to validate the inferred transmission system topologies. Finally, we take a look on transmission grids of non-German regions in Sec. 5.5 and end this chapter with the consideration of the European transmission grid in Sec. 5.6. Note that the introduced models are based upon OSM data as of July 1$^{\text{st}}$, 2016. Besides OSM data for Europe[1], the following section additionally makes use of OSM data for Japan[2].

5.1 Generator-to-Substation Connection

A valid 380kV transmission line section is modeled by the OSM relation (OSM-ID=)*1527969*, which is geographically depicted at Fig. 5.1. The single OSM object members of the relation, marked in black, are listed at Lst. 5.1. The relation indicates a

[1] `http://download.geofabrik.de/europe-latest.osm.pbf`
[2] `http://download.geofabrik.de/asia/japan-latest.osm.pbf`

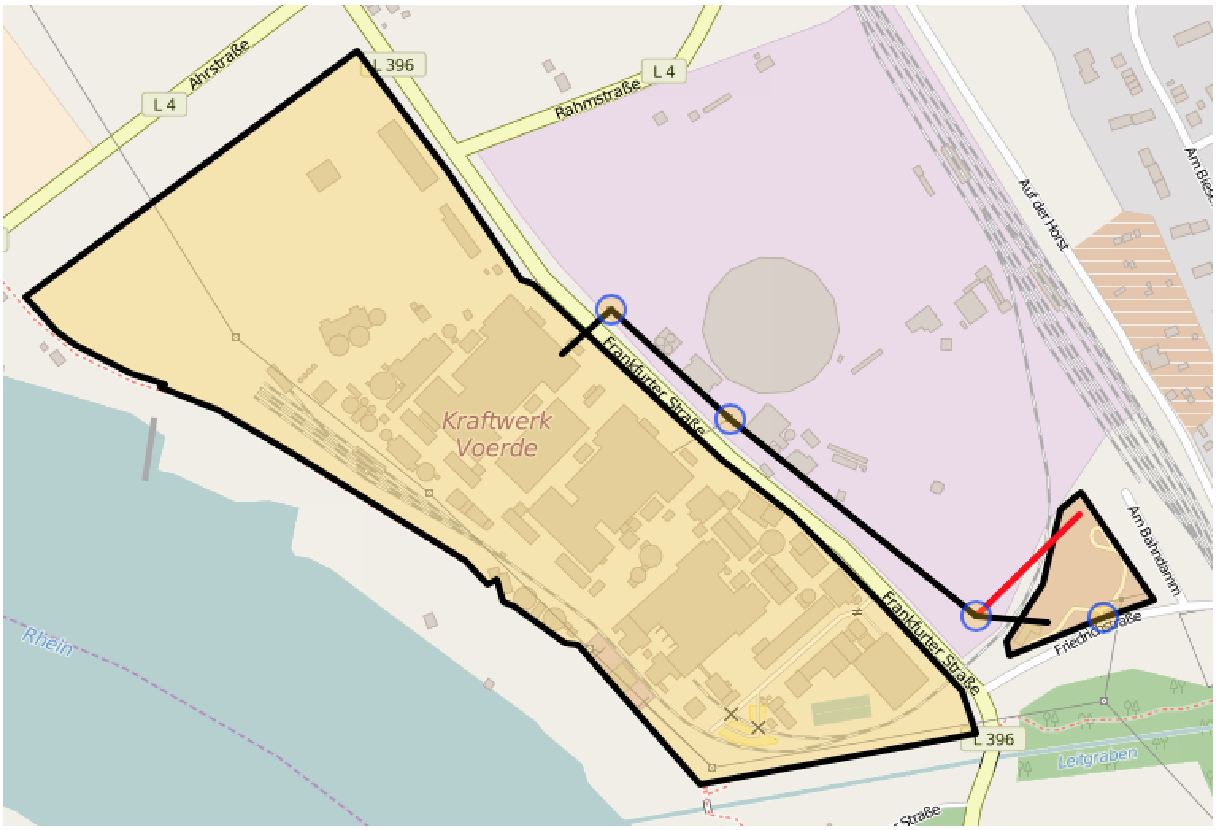


Figure 5.1: 380kV transmission section modeled by OSM relation *1527969*.

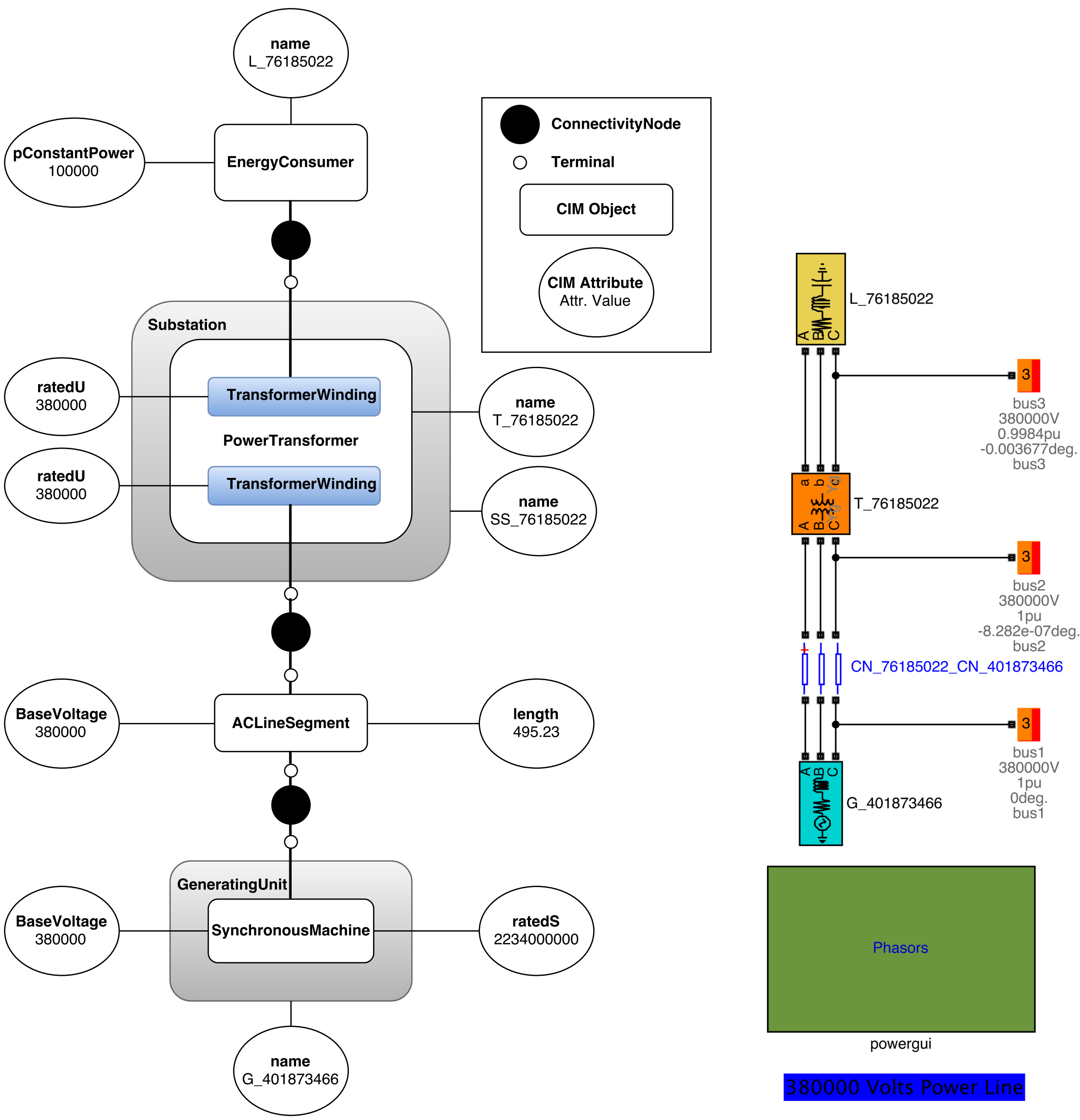
connection between generator *401873466* and substation *76185022*. An additional power line, OSM way *107175380*, is denoted in red color to show that there is more than one path connecting the two stations.

Listing 5.1: OSM relation *1527969* in XML format.

```
<relation id=”1527969” ... >
<member type=”way” ref=”401873466” role=”plant”/>
<member type=”way” ref=”107271574” role=”line”/>
<member type=”way” ref=”374424756” role=”line”/>
<member type=”way” ref=”374424755” role=”line”/>
<member type=”way” ref=”76185022” role=”sub_station”/>
...
<tag k=”route” v=”power”/>
<tag k=”type” v=”route”/>
<tag k=”voltage” v=”380000”/>
</relation>
```

Listing 5.2: Transnet logging excerpt, indicating a found p2p connection.

Circuit 2
Station – ID: 76185022 Type: substation Name: Station Zensenbusch Voltage: 380000
Line – ID: 107175380 Type: line Name: 380kV Voerde A Ref: 4167 Voltage: 380000
Line – ID: 374424756 Type: line Name: 380kV Voerde A+B Ref: 4167 Voltage: 380000
Line – ID: 107271574 Type: line Name: 380kV Voerde B Ref: 4167 Voltage: 380000
Station – ID: 401873466 Type: plant Name: Kraftwerk Voerde



(a) CIM. (b) Simulink model.

Figure 5.2: Models for the considered 380kV transmission circuit.

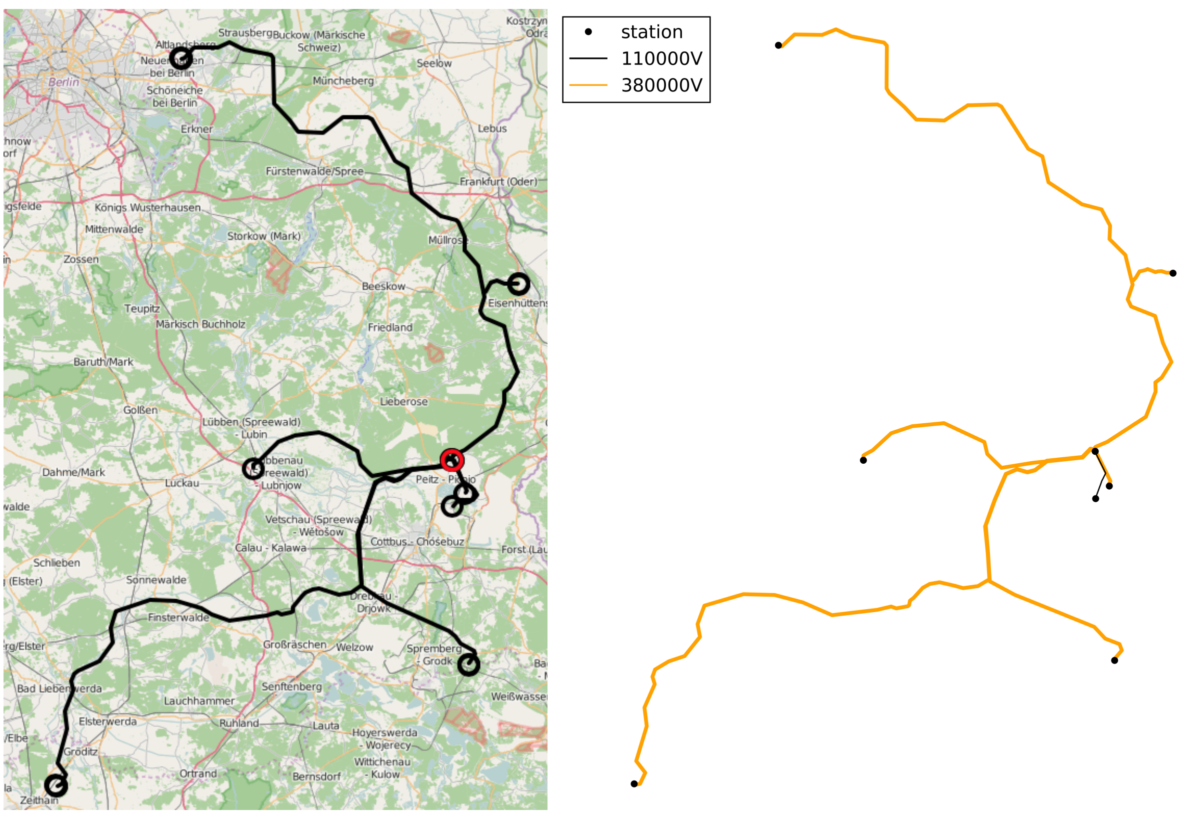
To reconstruct the p2p connection modeled by the OSM relation, we launch Transnet with input parameter *-s 76185022* to start a inference for substation *76185022*. Lst. 5.2 presents an excerpt of the Transnet logging during the respective execution, showing that the inference engine found a similar p2p circuit. The p2p circuit contains the OSM ways *401873466*, *374424756*, *374424755*, and *76185022*, just as the considered OSM relation. But substation *76185022* is attached through OSM way *107175380*, instead of OSM way *374424755* at the relation. Nonetheless, the 380kV connection between station *401873466* and *76185022* has been found—no matter of which OSM ways it consists. As 4.1.2 already explained, Transnet has a mechanism to prevent the duplicate inference of the same p2p circuit. Due to this mechanism, the OSM relation could not be exactly reconstructed, since another p2p circuit connecting generator *401873466* and substation *76185022* has already been found in advance.

Based on the found p2p circuit a CIM is constructed, which is schematically depicted at Fig. 5.2a. Note that the figure does not show all CIM objects and attributes for reasons of simplicity—the complete CIM for this p2p circuit is attached at B.1 in RDF/XML format. However, the CIM represents basic p2p circuit equipments with the *GeneratingUnit*, *Substation*, and *ACLineSegment* CIM objects. To attach a load to the substation, a second *TransformerWinding* object has been added to the *PowerTransformer*. This winding is connected to the artificially created *EnergyConsumer* object that represents the load *L_76185022*. Since the power system inference has been done for a particular substation and not for a region, the load estimation approach of Sec. 4.2.2 could not be applied. Therefore, a constant power demand of *100000* = 100kW is assumed. Further properties like the *BaseVoltage* and *ratedU* attributes indicate that the modeled power circuit operates at *380000* = 380kV. The *ratedS* attribute, which belongs to the *SynchronousMachine* object stands for the plant's nominal power of *2234000000* = 2234 megawatts (MW). Moreover, the length of the transmission line connecting both stations is specified by the *length* attribute of the *ACLineSegment* with *459.23* meters. Finally, the remaining CIM attributes indicate unique equipment names.

Based on the CIM, a Simulink simulation model, as presented at Fig. 5.2b, is generated. The model consists of the component blocks that build the considered power circuit: The generator *G_401873466*, the power line *CN_76185022_CN_401873466*, the transformer *T_76185022*, and the load *L_76185022* block. By means of the *powergui* block, a load flow simulation has already been performed. The results are displayed by the three buses *bus1*, *bus2*, and *bus3*. All three buses are configured to use 380kV as base voltage. At *bus1*, which is directly attached to the generator, we encounter neither a voltage drop, nor a phase angle shift. There is also no remarkable voltage drop at *bus2*, due to the negligible line losses caused by the short transmission line. However, we encounter a little phase angle shift at *bus2*. Finally, at *bus3* we derive a recognizable voltage drop caused by the transformer losses: The measured voltage is *0.9984pu* of 380kV, which are $0.9984 * 380000 = 379392$ volts.

5.2 Substation Star Network

In this example we consider all transmission connections emerging from substation *23025610*. A geographical representation of the considered transmission network is shown at Fig. 5.3a. The substation *23025610* is marked in red color, whereas all emerging transmission connections are black-colored. We refer to this example as substation star network, since we derive a star-like topology.



(a) Map. **(b)** Inferred topology.

Figure 5.3: Substation *23025610* star network.

To infer the transmission grid around substation *23025610*, we launch Transnet with input parameter *-s 23025610* to start an inference run for this particular station. The resulting

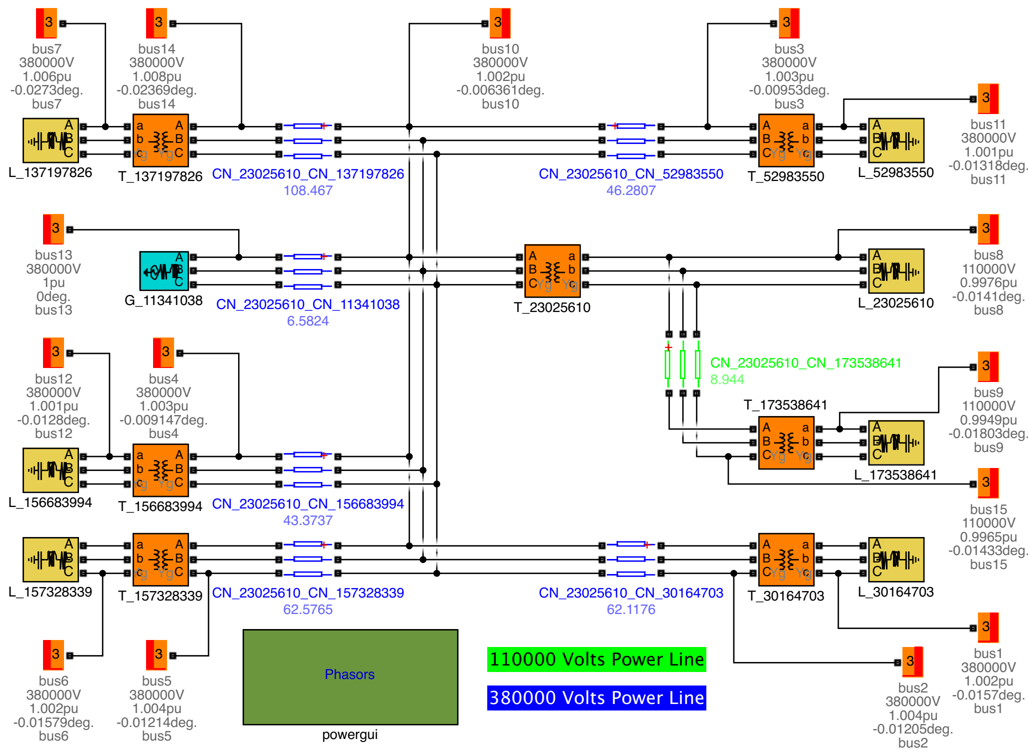


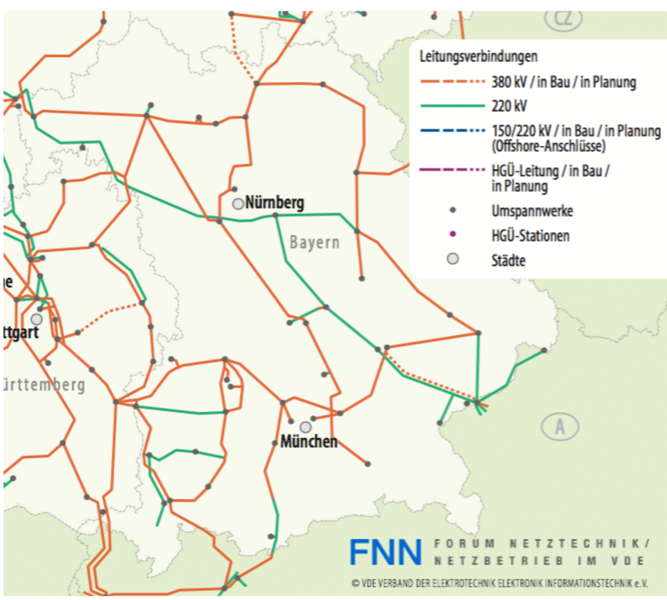
Figure 5.4: Simulink model of the substation *23025610* star network.

topology plot is depicted at Fig. 5.3b, which has a very similar structure compared to the geographical representation at Fig. 5.3a. The legend of the topology plot reveals that black-filled circles represent stations (substations or generators) and the power lines are colored depending on their voltage level. In this case, 380kV power lines are orange-colored, whereas 110kV power lines are marked with black color.

Based on the CIM file, which has been generated by the Transnet execution, a Simulink simulation model has been constructed, which is similar to Fig. 5.4. Due to space issues, the figure presents a compact version of the star network around substation *23025610* without equipment block positions that correspond to their geographical locations. Power lines are again colored depending on the voltage level: 380kV lines are colored blue, whereas 110kV lines are colored light-green. Their length is displayed right below their name. For example, *CN_23025610_CN_137197826* is a 380kV transmission line with a length of *108.467* kilometers. The transformer block *T_23025610* in the center of the model is representing the transformer of substation *23025610*. It steps down the voltage from 380kV to 110kV and therefore indicates the border of the 380kV and 110kV transmission subsystems. Moreover, Fig. 5.4 illustrates the condition after a load flow simulation. Therefore, the single buses already carry measurements of the actual voltages and phase angles after load flow calculations.

5.3 Bavarian Transmission Grid

While the previous examples ran Transnet for a single substation, this example illustrates the inference for a particular region. To do so, Transnet can be launched with the input parameter *-p*, which additionally requires to specify the path to a poly file. Since we aim to infer the transmission system of Bavaria, the corresponding poly file[3] describing the extent of Bavaria's border, is used.



(a) Map [1].

(b) Inferred topology.

Figure 5.5: Bavarian 380kV and 220kV transmission system.

Fig. 5.5 compares the inferred transmission system topology (Fig. 5.5b) with an existing map of the Bavarian high voltage grid as of 2014 (Fig. 5.5a). Although the reference map is just an approximation of the topology, it suffices to identify transmission sections that have not been found by Transnet. In general, the structures of the grids are very similar. A

[3] http://download.geofabrik.de/europe/germany/bayern.poly

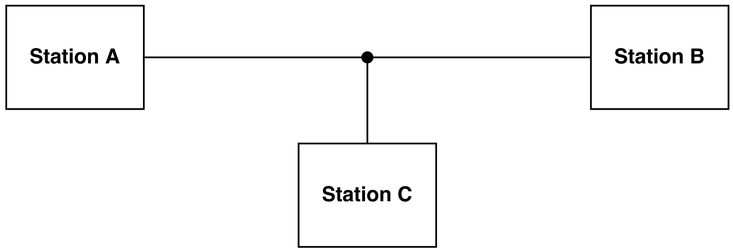


Figure 5.6: Three stations connected by a T-junction.

few transmission sections at the border are not covered, due to the fact that the inference only considers substations and power lines within the specified region. Apparently, several transmission sections in the south of Munich are labeled with the wrong voltage level. The reason for this issue are wrong voltage tags of the underlying OSM ways.

The visual comparison with existing transmission system maps gives a rough estimate for the quality of the inferred topology. However, to better quantify the quality of our work, we establish a metric that evaluates the coverage of existing OSM relations. As already mentioned, the German and consequently the Bavarian transmission system are comprehensively modeled with OSM relations. These relations also model connections from one station to another, which allows us to compare the already existing OSM relations with our inferred p2p circuits. Therefore, the emerging metric measures how many OSM relations could have been covered by the inferred p2p circuits.
Since one OSM relation may model more than one p2p circuit, the validation approach requires a case discrimination: If an OSM relation models exactly one p2p circuit, meaning that it has exactly two stations in its members list, it can be directly validated against the inferred p2p circuits by comparing the end point stations. If an OSM relation represents more than one p2p circuit, a direct validation against the inferred p2p circuits is not possible. In that case, we split affected OSM relations in single p2p circuits to enable a direct comparison.
Another problem regarding this validation approach arises from Transnet's mechanism to prevent the inference of duplicate p2p circuits (see Sec. 4.1.2). Assume three stations A, B, and C are connected with a T-junction, as illustrated at Fig. 5.6. Furthermore, imagine Transnet starts with the power system inference for station A and extracts the p2p circuits A-B and A-C. Then Transnet continues with the inference for station B and encounters that B's power line is already occupied with a p2p circuit. For this reason, the inference for the power line at station B is skipped and the inference continues at station C. Again, Transnet recognizes that station C's power line is also already occupied by a p2p circuit and therefore skips the corresponding inference. Although this transmission section is completely described by the p2p circuits A-B and A-C, the

implicit p2p circuit B-C is missing, which is a problem for the p2p circuit validation. To circumvent this problem, we enhance the validation mechanism to also test the OSM relation end points against the transitive closures of the inferred p2p circuits' stations. Note that the transitive closure of a station s is the set of stations that are directly or indirectly connected to s.

The complete algorithm for the validation of inferred p2p circuits is described at Alg. 4: Given all OSM relations R and the inferred p2p circuits C for a given region, the algorithm determines how many OSM relations are covered by inferred p2p circuits in per cent. To do so, a counter *ncr* is initialized with zero to keep record of the amount of covered relations. For each relation r in R, build a list S that contains the stations within r. Now, for each circuit c in C, check whether the relation r's voltage equals the circuit c's voltage. If the voltages match, build the transitive closure *th1* for the first station *s1* of the circuit. Now check for each station pair p in S, if both the first (*p1*) and the second station (*p2*) of p are contained in the transitive closure *th1*. If so, increment the counter *nhp*, which keeps track of the number of covered station pairs. After looping over all station pairs, check whether the number of covered station pairs *nhp* equals the number of total station pairs (*length(S) - 1*) in relation r. If this is the case, the relation r is fully represented by inferred p2p circuits and the counter *ncr* is incremented. Once, all relations have been processed, determine the coverage *cov* of all OSM relations by dividing the number of covered relations *ncr* by the total number of relations (length(R)). To get the coverage in per cent, multiply the result with hundred.

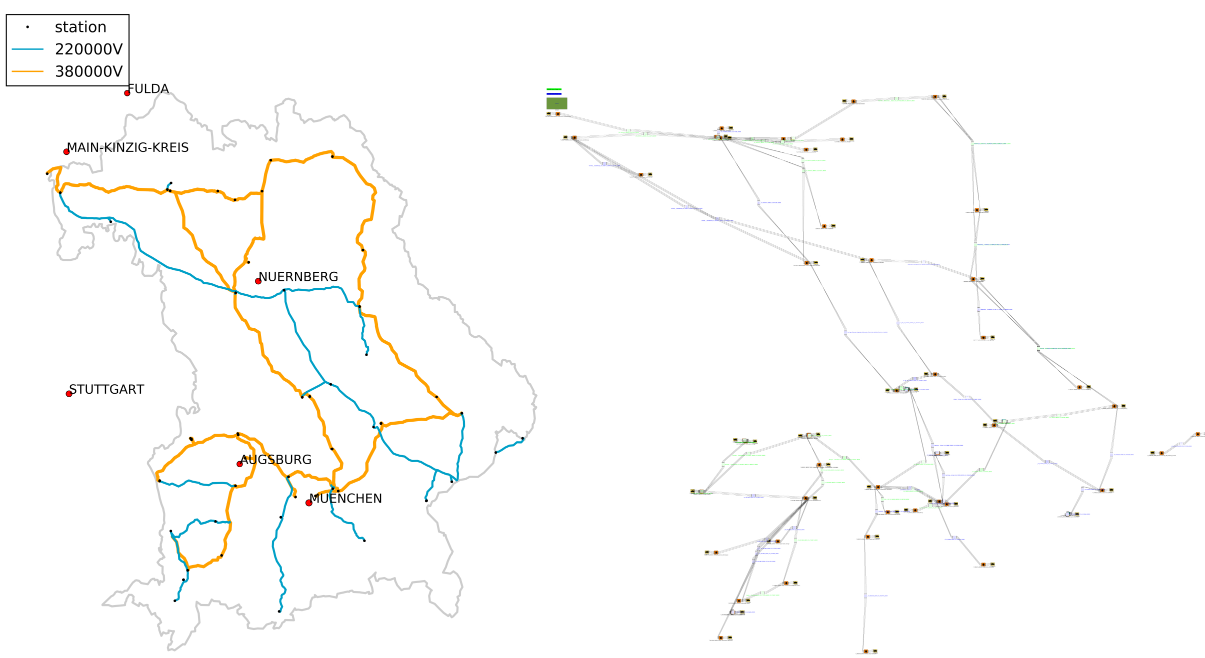
The application of the introduced algorithm to the inferred Bavarian 220kV and 380kV transmission system results in a relation coverage of 95 per cent. This means that Transnet is capable of reconstructing 95 per cent of the already existing, manually created OSM relations, which model the Bavarian 220kV and 380kV transmission grid topology.

Besides the Bavarian transmission grid topology, the Transnet run also delivers the corresponding CIM. Based on the CIM, a respective Simulink model has been generated, which is depicted at 5.7b. The size of the emerging simulation model is too large to perform a detailed analysis. Nevertheless, we compare the topology of the Simulink model with the topology plot created by Transnet (Fig. 5.7a) and notice that both structures are highly similar. Moreover, this comparison illustrates that the introduced projection of the power equipment locations on the 2-D Simulink plane accomplishes power system simulation models that geographically correspond to their real geospatial topology.

Algorithm 4: Validation of inferred p2p circuits against OSM relations.

Data: OSM relations R and inferred p2p circuits C for a given region
Result: Percentage of covered OSM relations by p2p circuits

```
ncr = 0; // number of covered relations
foreach relation r in R do
    S = list of stations in r;
    foreach circuit c in C do
        if c.voltage == r.voltage then
            s1 = c.members[0]; // first circuit member is a station
            th1 = transitive closure of s1;
            nhp = 0; // number of hit p2p circuits
            foreach station pair p in S do
                p1 = p[0]; // station 1 of station pair
                p2 = p[1]; // station 2 of station pair
                if p1 in th1 and p2 in th1 then
                    nhp += 1;
            if nhp == length(S) - 1 then
                ncr += 1;
                break;
cov = ncr/length(R) * 100;
```

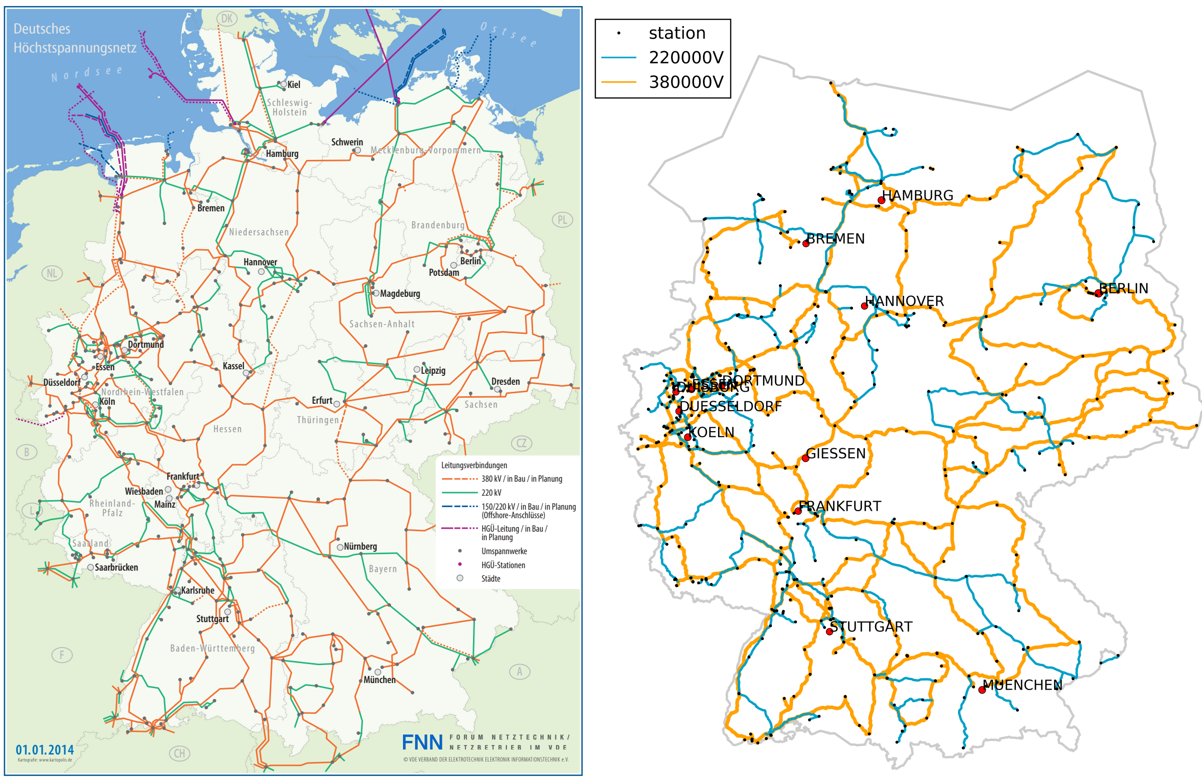


(a) Inferred topology. **(b)** Simulink model.

Figure 5.7: Bavarian 380kV and 220kV transmission system.

5.4 German Transmission Grid

To show that our approach also works for larger power systems, we evaluate the derivation of the German 220kV and 380kV transmission system in this section. The transmission system topology inference has been performed by running Transnet with input parameter *-p* and the corresponding poly file[4] modeling the shape of Germany's border.



(a) Map [1]. **(b)** Inferred topology.

Figure 5.8: German 380kV and 220kV transmission system.

For the first validation step, we perform a visual comparison of the emerged topology plot (Fig. 5.8b) with an existing transmission system map (Fig. 5.8b). Again, we notice that both transmission system topologies are very similar. Also, we encounter missing transmission system sections towards the German border at the inferred topology due to the fact that the inference engine only considered stations and power lines within Germany.
In the second validation phase, we exploit structural characteristics of the German transmission grid, which are published at [18] by the German TSOs. The published

[4] `http://download.geofabrik.de/europe/germany.poly`

Table 5.1: Evaluation Summary for Germany and Bavaria

Region	Covered OSM relations (%)	Covered power circuit length (%)
Bavaria	95	-
Germany	89	73

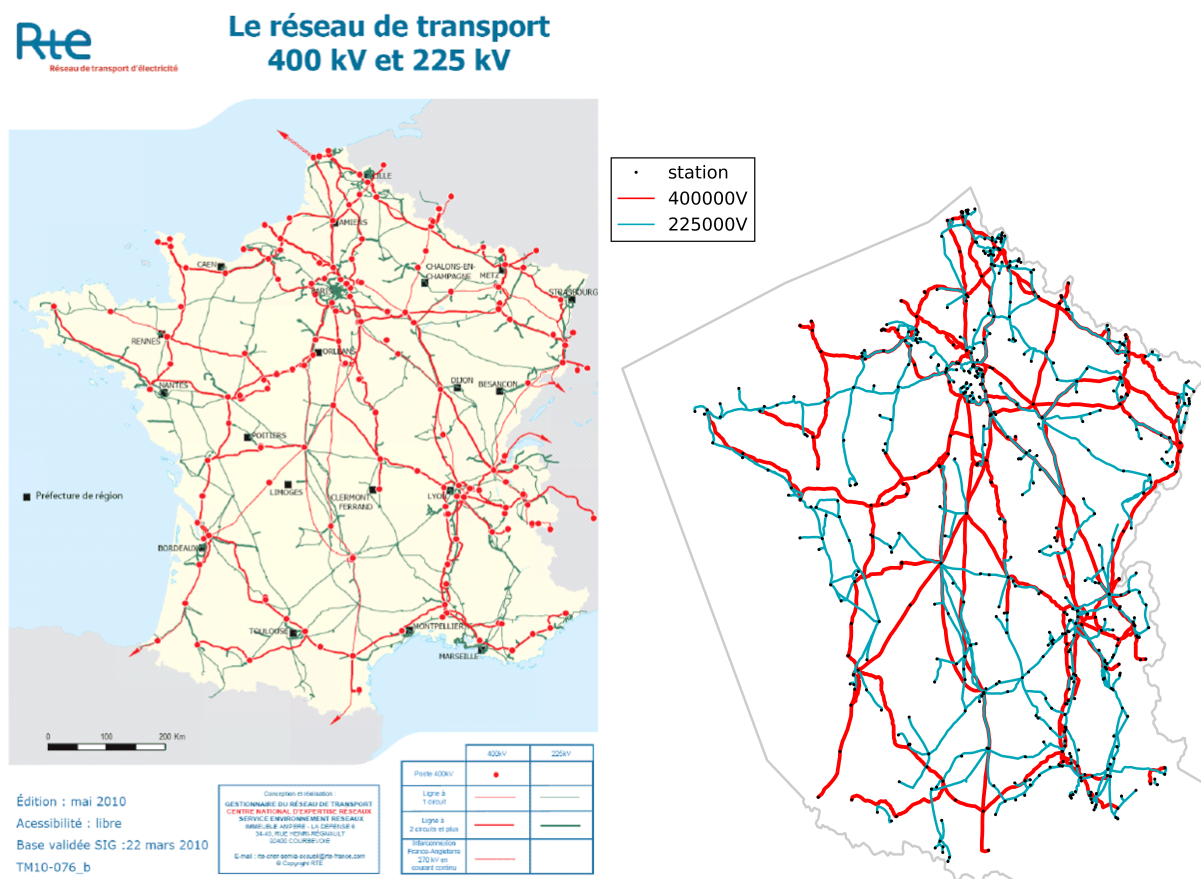
data allow us to compare the total power circuit length of the German high voltage grid with the total length of the inferred p2p circuits. According to Germany's TSOs, the total length of installed power circuits is 34813 kilometers [18]. In contrast, the total length of the inferred p2p circuits is 25510 kilometers, which would indicate a coverage of approximately 73 per cent. The reason for the rather low coverage can be explained by the different measurement methodologies. The TSOs consider the length of each power circuit, meaning that a p2p circuit with more than one circuit installed is counted multiple times. In contrast, we count the length of each inferred p2p circuit only once. Hence, this metric seems not explanatory to assess the extent of the inferred grid topology.
For that reason, we proceed with the third validation step, which is the application of Alg. 4 to the inferred transmission system topology. The outcome of this validation step is that the inferred p2p circuits cover approximately 89 per cent of the existing OSM relations that model the German 220kV and 380kV transmission system. We believe this figure to be the most representative metric regarding the quality of the inferred transmission grid, since various related projects, as discussed at Sec. 3.1, consider OSM relations a reasonable data source for the German transmission system. The validation results for the inferred transmission topology of Bavaria and Germany are summarized at Tab. 5.1. Note that structural characteristics for Bavaria's transmission system could not be found. Hence, a comparison of the total power circuit length according to TSOs with the total length of inferred p2p circuits was not possible.

Further automatically generated products of the inference engine include a CIM and a Simulink model for the German 220kV and 380kV transmission grid. These models are not further analyzed due to their extensive size.

5.5 Transmission Grids of Other Countries

To proof that the thesis' approach is not only applicable to Germany, but generally to any other region, we examine transmission grid topologies of various other countries in this section.

While the transmission systems of Central Europe mostly operate at voltage levels of 220kV and 380kV, the Western Europe's transmission systems work mainly at 225kV and 380kV. An integral part of the transmission system in Western Europe is represented by the French transmission system, which is shown at Fig. 5.9a. To infer the transmission grid for France, Transnet has to be launched with the input parameter *-p* and the poly file[5] that corresponds to the French border. The result of the inference is depicted at Fig. 5.9b. When comparing the inferred topology with the transmission system map at Fig. 5.9a, we recognize a high similarity of both grids. Again, several border-crossing transmission sections are missing at the inferred topology due to the fact that only stations and power lines within France are considered. In fact, the inferred topology almost completely covers the reference map's transmission system topology, which in turn indicates that power-relevant OSM data is well established for France.



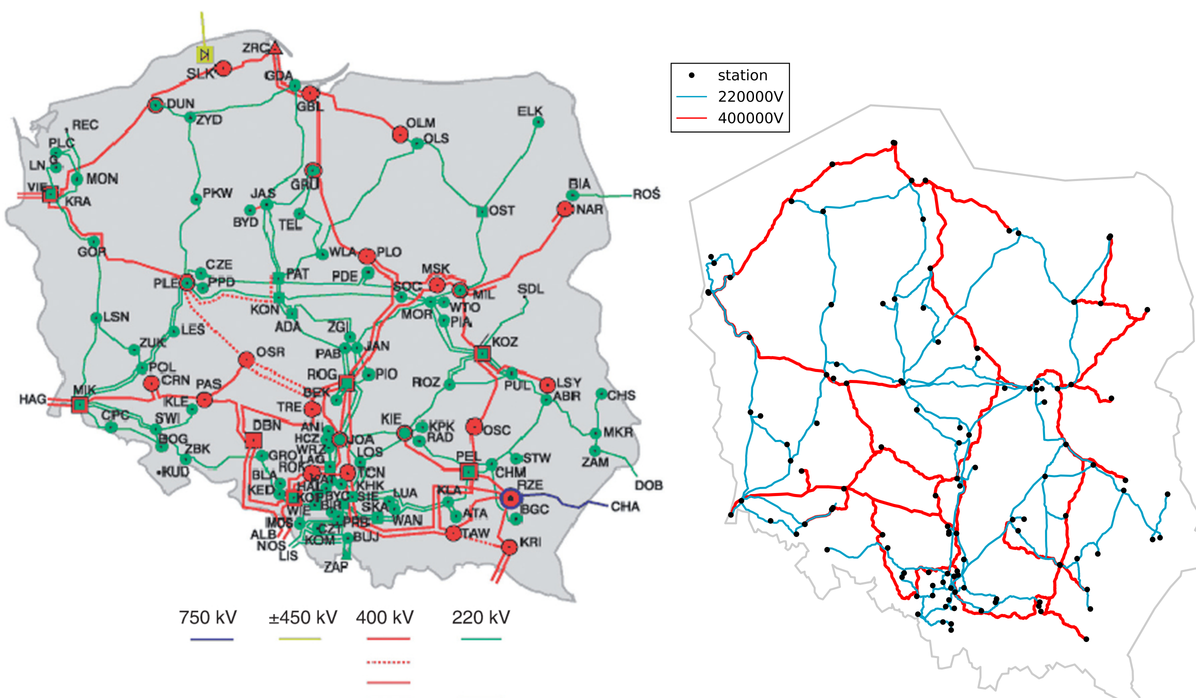
(a) Map [39]. **(b)** Inferred topology.

Figure 5.9: French 400kV and 225kV transmission system.

The transmission systems of Eastern Europe mostly operate at 220kV and 400kV. One

[5] `http://download.geofabrik.de/europe/france.poly`

representative example for the Eastern Europe transmission grid is illustrated at Fig. 5.10a, which depicts the Polish transmission system map for the voltages 750kV, 450kV, 400kV, and 220kV. The majority of the transmission lines transmit energy on the voltage levels 400kV and 220kV. To let Transnet infer the transmission topology of Poland, the inference engine has to be called with parameter *-p* and the poly file[6] that represents the shape of the Polish border. The result of the corresponding Transnet execution is shown at Fig. 5.10b. Once more, we discover a highly similar structure of the inferred topology and the transmission system map for the 220kV and 400kV power grid of Poland. Again, the inferred topology almost completely covers the reference map's transmission system topology, which in turn indicates that power-relevant OSM data is well established for Poland.



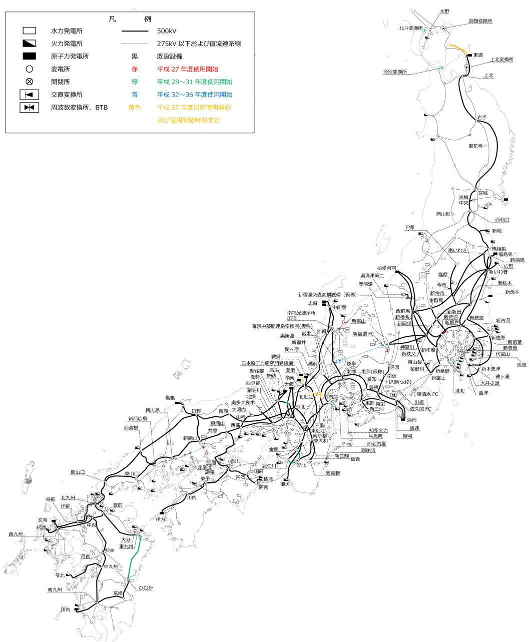
(a) Map [40]. **(b)** Inferred topology.

Figure 5.10: Polish 400kV and 220kV transmission system.

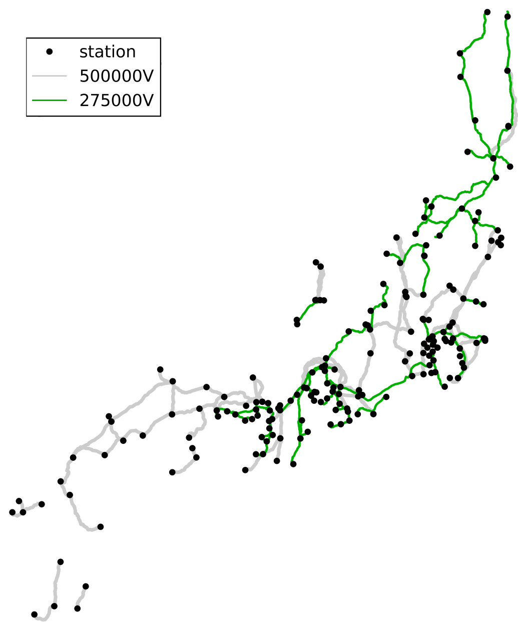
To assess the applicability of the thesis' approach for regions other than Europe, we present the results of the Japanese transmission system topology inference. Fig. 5.11b demonstrates the extracted 500kV and 275kV transmission system topology for Japan. Besides, Fig. 5.11a shows the Japanese transmission system map as of 2015. The comparison of both figures indicates a high similarity among the structure of the inferred

[6] `http://download.geofabrik.de/europe/poland.poly`

topology and the map's transmission infrastructure. However, especially at the north and south of Japan, several transmission connections could not be inferred due to invalid OSM tagging. Nevertheless, the comparison underlines the general applicability of the power system inference approach for any region and that power-relevant OSM data is also well established for Japan.



(a) Map [41].



(b) Inferred topology.

Figure 5.11: Japanese 500kV and 275kV transmission system.

5.6 European Transmission Grid

In the previous section, we assessed the inference approach's ability to extract transmission grids for particular countries and only a few voltage levels. In this section, we demonstrate the approach's applicability to even larger power systems with various voltage levels on the example of the European transmission grid.

Fig. 5.12 gives an overview of the European transmission grid topology. Apparently, the transmission system operates at different voltage levels depending on the region in

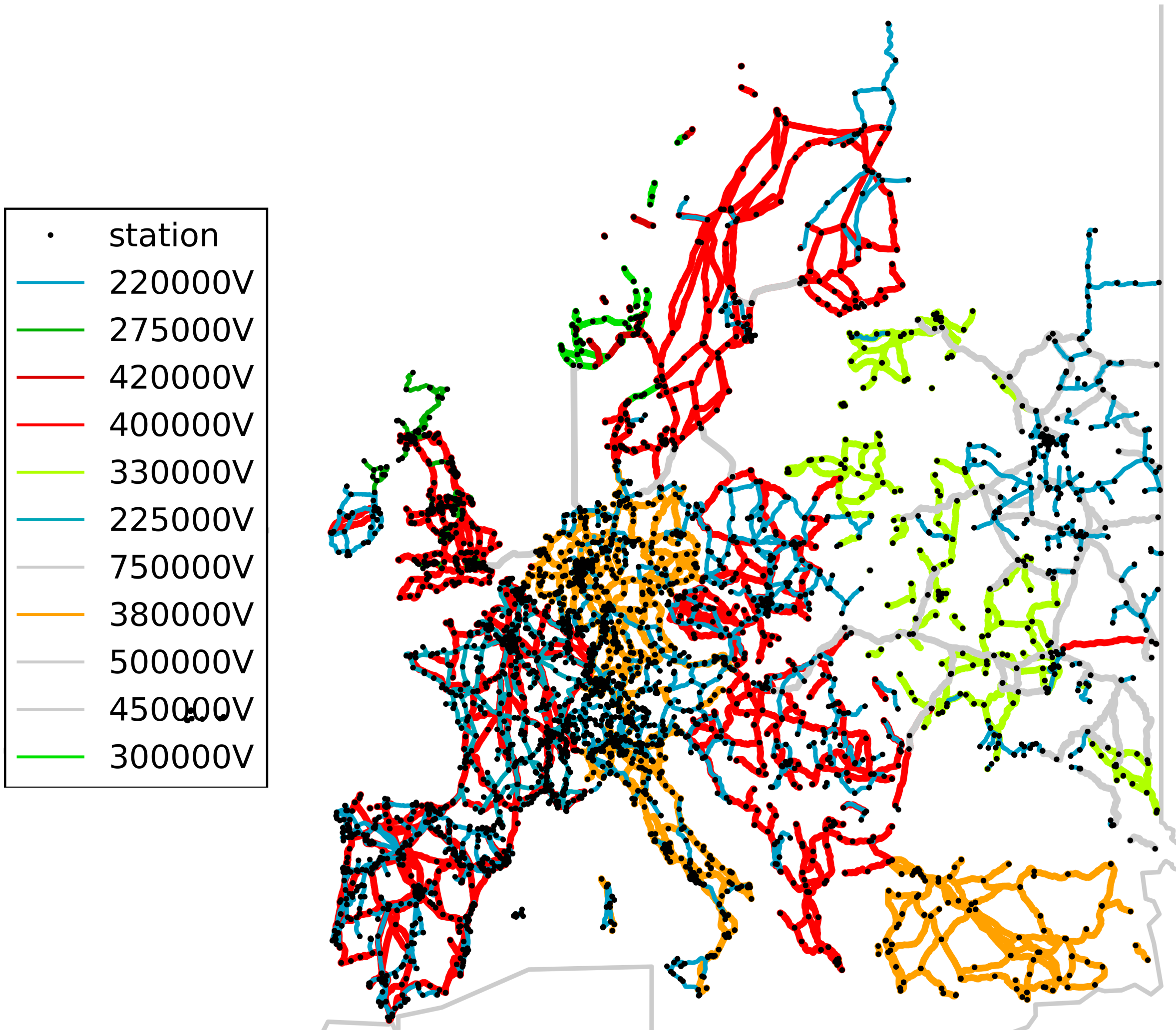


Figure 5.12: Inferred European transmission system topology.

Europe. While the transmission system in Western Europe works with the voltage levels 225kV and 400kV, Central Europe's transmission system operates at 220kV and 380kV. Towards Eastern Europe, we encounter the voltage levels 220kV and 400kV, whereas the easternmost parts of Europe mostly transmit energy at 330kV and 750kV.

Another interesting aspect implicitly illustrated by Fig. 5.12 is the density and availability of power-relevant OSM data throughout Europe. While the figure depicts a dense transmission system infrastructure in Western and Central Europe, the transmission system density decreases towards Eastern Europe. The reason behind this fact is the varying availability and quality of power-relevant OSM data in Europe. The amount of valid-tagged OSM data is high enough in Western and Central Europe to reconstruct its transmission system topology reasonably. In contrast, the easternmost regions of Europe lack in sufficient power-relevant data to build reasonable transmission system topology models.

Chapter 6

Conclusions

In this thesis we proposed an approach to automatically generate power grid simulation models based on crowdsourced data. The generation process starts with the construction of point-to-point circuits that represent typical transmission system sections like substation-to-substation or substation-to-generator connections. The inferred circuits are then transformed to a CIM format, which is the standard information exchange format between transmission system operators and a common input format for power grid analysis tools. Finally our approach provides a ready-for-simulation power grid model by mapping the obtained CIM into a Simulink model. We demonstrated the effectiveness of our approach for several use cases in Chap. 5 and showed that a comprehensive simulation model can be generated from crowdsourced data. As an important component of the open-source OpenGridMap project, the method presented in this thesis is under constant development and improvement[1].
Moreover, we outlined related projects and literature to our work in Chap. 3 and provided an extensive knowledge-base in Chap. 2 to better understand the core topics of the thesis.

Although our approach is applicable for any region, it does not necessarily mean that our inference yields good results for every spot on the earth. The completeness of our models depends on the availability of power-related OSM data, which varies from region to region. That not being enough, the power-related OSM data has to follow the OSM mapping rules for power systems to be eligible for our inference. However, during our research we encountered several problems that impaired the quality of our work due to unreliable OSM data.
First of all, our inference approach relies on the availability of OSM ways and the validity

[1] `https://github.com/OpenGridMap/transnet`

of power-relevant tags. According to Tab. 2.2 only 54 per cent of OSM ways, which are labeled as power lines, have a voltage tag. The remaining power lines do not have a voltage tag and can not be considered at the inference of power systems. Most of the lines without a voltage tag in OSM belong to the low-level distribution system and therefore have no noticeable impact on the inference of transmission systems. Nevertheless, this problem is a huge limitation, once distribution systems are to be inferred.

Another limitation arises from our approach expecting consistent voltage tags along a power line path. When parts of a power line path change voltages due to inconsistent tagging, our inference engine can not obtain the corresponding transmission section.

Moreover, our inference approach relies on the geospatial intersection of OSM power lines and stations to extract p2p connections. Due to OSM modeling mistakes, power lines spanning a potential p2p connection might not even geospatially touch stations. In such cases, the inference of potential p2p connections fails due to gaps between power lines and stations.

One more problem arises from substations that are geographically located within generators. Starting from a substation, our inference approach follows an emerging power line path until another station - a substation or a generator - is intersected. In case of a substation being geographically located within a generator, the inference stops when encountering the generator, although the power line path would further lead to the substation. Hence, potential connections to substations located within generators can not be inferred, because such connections already end at the substation-wrapping generator.

While the problems listed above affect the quality of the power system structure inference, the following limitations have an impact on the parametrization of the generated CIMs and power system simulation models.

In general, OSM data is a valuable data source for the inference of power system topologies. Nevertheless, OSM data is not sufficient to fully specify the electrical characteristics of the inferred power systems. The information on electrical properties through OSM tags is very limited. Besides the voltage tags of OSM power lines and substations, power output tags of OSM generators are sporadically available to specify their nominal power output. To build a reasonable power system simulation model, a minimum set of electrical characteristics has to be known. For the scope of this thesis, the minimum set's extent is defined by Tab. A.2, which lists the used power system equipments and their mandatory electrical characteristics in terms of CIM objects and attributes. At the moment, OSM provides only a little contribution to specify the required properties.

In the context of electrical parameter estimations, we also introduced a methodology for the estimation of loads attached to transmission substations at Sec. 4.2.2. Although we believe that the approach builds a good basis for proper load estimations, several limitations are addressed in the following. To approximate the population within a certain region, we exploit a data source that only provides administrative data for Central Europe. This circumstance limits the applicability of the thesis' inference approach to Central Europe, unless other global data sources are acquired. Moreover, the load estimations consider the average power consumption of a German citizen in 2015, which is not generally representative for people all over the world at any time in future. Additionally, the estimations do not involve important load influences like the level of industrialization, seasonality, or the time of the day.

As part of our future work, we plan to use the proposed method to generate larger and more complex simulation models and also improve their quality by considering other relevant data sources. Also, we plan to make our inference engine more robust against invalid tagged OSM data. In regions, where power-related OSM data is not well established, we try to develop a means to automatically create OSM objects based on satellite pictures. In the long term, we aim to provide complete power system models for any region in the world, including distribution networks.

Appendices

Appendix A

Parametrization

A.1 Transmission Line Characteristics

Table A.1: Standard operating parameters of 110kV overhead transmission lines [8].

Parameter	**Overhead line Al/St 265/35**	**Description**
R_{20} mΩ/km	109.5	Series resistance for 20° Celsius
R_{40} mΩ/km	118.3	Series resistance for 40° Celsius
X Ω/km	0.381	Series reactance
G nS/km	40	Susceptance
C nF/km	9.4	Capacity

A.2 CIM Object Attributes

Table A.2: Used mandatory CIM object attributes according to ENTSO-E CIM Profile 1 [9].

CIM object	**Mandatory parameter**	**Parameter description**	**Value**
Substation	SubGeographicalRegion	The geographical region this equipment belongs to.	'EU'
TransformerWinding	b	Magnetizing branch susceptance.	Simscape default
	connectionType	The type of connection of the winding.	Wye
	g	Magnetizing branch conductance	Simscape default
	r	Positive sequence series resistance of the winding.	Simscape default
	ratedS	The normal apparent power rating for the winding.	Simscape default
	ratedU	The rated voltage (phase-to-phase) of the winding.	See Sec. 4.2.2
	windingType	The type of winding.	See Sec. 4.2.2
	x	Positive sequence series reactance of the winding.	Simscape default
GeneratingUnit	maxOperatingP	This is the maximum operating active power limit.	See Sec. 4.2.2
	minOperatingP	This is the minimum operating active power limit.	See Sec. 4.2.2
	nominalP	The nominal power of the generating unit.	See Sec. 4.2.2
SynchronousMachine	operatingMode	Current mode of operation.	'generator'
	qPercent	Percent of the coordinated reactive control from this machine.	'100'
	r	Positive sequence resistance of the synchronous machine.	Simscape default
	ratedS	Nameplate apparent power rating for the unit.	See Sec. 4.2.2
	Type	Modes that this synchronous machine can operate in.	'generator'
LoadResponseCharacteristic	exponentModel	If false, a constant power demand is assumed.	'false'
	pConstantPower	Portion of active power load modeled as constant power.	See Sec. 4.2.2
ACLineSegment	bch	Positive sequence shunt (charging) susceptance.	See Sec. 4.2.2
	r	Positive sequence series resistance of the entire line section.	See Sec. 4.2.2
	x	Positive sequence series reactance of the entire line section.	See Sec. 4.2.2
BaseVoltage	isDC	False indicates alternating current.	'false'
	nominalVoltage	The PowerSystemResource's base voltage.	See Sec. 4.2.2

Appendix B

Results

B.1 CIMs

Listing B.1: CIM example of a transmission p2p circuit.


```xml
<rdf:RDF xmlns:cim="http://iec.ch/TC57/2010/CIM-schema-cim15#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
	<cim:SubGeographicalRegion rdf:ID="651c183a-5d3c-11e6-ac94-f45c899d1079">
	<cim:SubGeographicalRegion.Region rdf:resource="#"/>
	</cim:SubGeographicalRegion>
	<cim:PositionPoint rdf:ID="651e3fb5-5d3c-11e6-b9af-f45c899d1079">
	<cim:PositionPoint.xPosition>7.4581543712</cim:PositionPoint.xPosition>
	<cim:PositionPoint.zPosition>1</cim:PositionPoint.zPosition>
	<cim:PositionPoint.yPosition>51.4823604737</cim:PositionPoint.yPosition>
	<cim:PositionPoint.Location rdf:resource="#651e420c-5d3c-11e6-9a45-f45c899d1079"/>
	</cim:PositionPoint>
	<cim:Location rdf:ID="651e420c-5d3c-11e6-9a45-f45c899d1079"/>
	<cim:PositionPoint rdf:ID="651e46e6-5d3c-11e6-a7c8-f45c899d1079">
	<cim:PositionPoint.xPosition>6.68934281886</cim:PositionPoint.xPosition>
	<cim:PositionPoint.yPosition>51.5764229288</cim:PositionPoint.yPosition>
	<cim:PositionPoint.Location rdf:resource="#651e4811-5d3c-11e6-b459-f45c899d1079"/>
	</cim:PositionPoint>
	<cim:Location rdf:ID="651e4811-5d3c-11e6-b459-f45c899d1079"/>
	<cim:Substation rdf:ID="651e4cfd-5d3c-11e6-9475-f45c899d1079">
	<cim:IdentifiedObject.name>SS_76185022</cim:IdentifiedObject.name>
	<cim:PowerSystemResource.Location rdf:resource="#651e4811-5d3c-11e6-b459-f45c899d1079"/>
	<cim:Substation.Region rdf:resource="#651c183a-5d3c-11e6-ac94-f45c899d1079"/>
	</cim:Substation>
	<cim:PowerTransformer rdf:ID="651e4d80-5d3c-11e6-ae47-f45c899d1079">
	<cim:IdentifiedObject.name>T_76185022_380000_Station_Zensenbusch</cim:IdentifiedObject.name>
	<cim:Equipment.EquipmentContainer rdf:resource="#651e4cfd-5d3c-11e6-9475-f45c899d1079"/>
	</cim:PowerTransformer>
	<cim:BaseVoltage rdf:ID="651e4ec0-5d3c-11e6-9a88-f45c899d1079">
	<cim:BaseVoltage.nominalVoltage>380000</cim:BaseVoltage.nominalVoltage>
	</cim:BaseVoltage>
	<cim:TransformerWinding rdf:ID="651e50e6-5d3c-11e6-a32b-f45c899d1079">
	<cim:IdentifiedObject.name>TW_76185022_380000</cim:IdentifiedObject.name>
	<cim:TransformerWinding.x>1.0</cim:TransformerWinding.x>
	<cim:TransformerWinding.ratedS>5000000</cim:TransformerWinding.ratedS>
	<cim:TransformerWinding.r>1.0</cim:TransformerWinding.r>
	<cim:TransformerWinding.ratedU>380000</cim:TransformerWinding.ratedU>
	<cim:TransformerWinding.connectionType rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#WindingConnection.Yn"/>
	<cim:TransformerWinding.windingType rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#WindingType.primary"/>
	<cim:ConductingEquipment.BaseVoltage rdf:resource="#651e4ec0-5d3c-11e6-9a88-f45c899d1079"/>
	<cim:TransformerWinding.PowerTransformer rdf:resource="#651e4d80-5d3c-11e6-ae47-f45c899d1079"/>
	</cim:TransformerWinding>
	<cim:ConnectivityNode rdf:ID="651e51b8-5d3c-11e6-a47b-f45c899d1079">
	<cim:IdentifiedObject.name>CN_76185022_380000</cim:IdentifiedObject.name>
	</cim:ConnectivityNode>
	<cim:Terminal rdf:ID="651e535c-5d3c-11e6-a877-f45c899d1079">
	<cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
	<cim:Terminal.ConductingEquipment rdf:resource="#651e50e6-5d3c-11e6-a32b-f45c899d1079"/>
	<cim:Terminal.ConnectivityNode rdf:resource="#651e51b8-5d3c-11e6-a47b-f45c899d1079"/>
	</cim:Terminal>
	<cim:PositionPoint rdf:ID="651e5685-5d3c-11e6-abf2-f45c899d1079">
	<cim:PositionPoint.xPosition>6.68146363662</cim:PositionPoint.xPosition>
	<cim:PositionPoint.yPosition>51.5776527735</cim:PositionPoint.yPosition>
```

```
<cim:PositionPoint.Location rdf:resource="#651e579e-5d3c-11e6-93ec-f45c899d1079"/>
</cim:PositionPoint>
<cim:Location rdf:ID="651e579e-5d3c-11e6-93ec-f45c899d1079"/>
<cim:GeneratingUnit rdf:ID="651e5c3a-5d3c-11e6-9c8a-f45c899d1079">
<cim:IdentifiedObject.name>G_401873466</cim:IdentifiedObject.name>
<cim:GeneratingUnit.maxOperatingP>2234000000.0</cim:GeneratingUnit.maxOperatingP>
<cim:GeneratingUnit.nominalP>2234000000.0</cim:GeneratingUnit.nominalP>
<cim:GeneratingUnit.genControlMode rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#GeneratorControlMode.setpoint"/>
<cim:GeneratingUnit.genOperatingMode rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#GeneratorOperatingMode.fixed"/>
<cim:GeneratingUnit.genControlSource rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#GeneratorControlSource.unavailable"/>
<cim:PowerSystemResource.Location rdf:resource="#651e579e-5d3c-11e6-93ec-f45c899d1079"/>
</cim:GeneratingUnit>
<cim:SynchronousMachine rdf:ID="651e5cc5-5d3c-11e6-9748-f45c899d1079">
<cim:IdentifiedObject.name>G_401873466_Kraftwerk_Voerde</cim:IdentifiedObject.name>
<cim:SynchronousMachine.r>0.01</cim:SynchronousMachine.r>
<cim:SynchronousMachine.x>0.01</cim:SynchronousMachine.x>
<cim:SynchronousMachine.ratedS>2234000000.0</cim:SynchronousMachine.ratedS>
<cim:SynchronousMachine.operatingMode rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#SynchronousMachineOperatingMode.generator"/>
<cim:SynchronousMachine.type rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#SynchronousMachineType.generator"/>
<cim:ConductingEquipment.BaseVoltage rdf:resource="#651e4ec0-5d3c-11e6-9a88-f45c899d1079"/>
<cim:SynchronousMachine.GeneratingUnit rdf:resource="#651e5c3a-5d3c-11e6-9c8a-f45c899d1079"/>
</cim:SynchronousMachine>
<cim:ConnectivityNode rdf:ID="651e5d8c-5d3c-11e6-b30e-f45c899d1079">
<cim:IdentifiedObject.name>CN_401873466_380000</cim:IdentifiedObject.name>
</cim:ConnectivityNode>
<cim:Terminal rdf:ID="651e5e7d-5d3c-11e6-a1f4-f45c899d1079">
<cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
<cim:Terminal.ConductingEquipment rdf:resource="#651e5cc5-5d3c-11e6-9748-f45c899d1079"/>
<cim:Terminal.ConnectivityNode rdf:resource="#651e5d8c-5d3c-11e6-b30e-f45c899d1079"/>
</cim:Terminal>
<cim:PositionPoint rdf:ID="651ea5a8-5d3c-11e6-9d1a-f45c899d1079">
<cim:PositionPoint.xPosition>6.68661911749</cim:PositionPoint.xPosition>
<cim:PositionPoint.yPosition>51.5769402838</cim:PositionPoint.yPosition>
<cim:PositionPoint.Location rdf:resource="#651ea757-5d3c-11e6-992b-f45c899d1079"/>
</cim:PositionPoint>
<cim:Location rdf:ID="651ea757-5d3c-11e6-992b-f45c899d1079"/>
<cim:ACLineSegment rdf:ID="651eaa63-5d3c-11e6-84d4-f45c899d1079">
<cim:IdentifiedObject.name>380kV_Voerde_A_CN_76185022_380000_CN_401873466_380000</cim:IdentifiedObject.name>
<cim:Conductor.length>495.228882832</cim:Conductor.length>
<cim:ACLineSegment.r>0.3257</cim:ACLineSegment.r>
<cim:ACLineSegment.x0>0.88025</cim:ACLineSegment.x0>
<cim:ACLineSegment.x>0.3153</cim:ACLineSegment.x>
<cim:ACLineSegment.r0>0.5336</cim:ACLineSegment.r0>
<cim:ConductingEquipment.phases rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#PhaseCode.A"/>
<cim:PowerSystemResource.Location rdf:resource="#651ea757-5d3c-11e6-992b-f45c899d1079"/>
<cim:ConductingEquipment.BaseVoltage rdf:resource="#651e4ec0-5d3c-11e6-9a88-f45c899d1079"/>
</cim:ACLineSegment>
<cim:Terminal rdf:ID="651eabae-5d3c-11e6-b7d2-f45c899d1079">
<cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
<cim:Terminal.ConductingEquipment rdf:resource="#651eaa63-5d3c-11e6-84d4-f45c899d1079"/>
<cim:Terminal.ConnectivityNode rdf:resource="#651e51b8-5d3c-11e6-a47b-f45c899d1079"/>
</cim:Terminal>
<cim:Terminal rdf:ID="651eac9c-5d3c-11e6-a007-f45c899d1079">
<cim:Terminal.sequenceNumber>2</cim:Terminal.sequenceNumber>
<cim:Terminal.ConductingEquipment rdf:resource="#651eaa63-5d3c-11e6-84d4-f45c899d1079"/>
<cim:Terminal.ConnectivityNode rdf:resource="#651e5d8c-5d3c-11e6-b30e-f45c899d1079"/>
</cim:Terminal>
<cim:Terminal rdf:ID="651eebc7-5d3c-11e6-acda-f45c899d1079">
<cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
<cim:Terminal.ConductingEquipment rdf:resource="#651eea9c-5d3c-11e6-9be6-f45c899d1079"/>
<cim:Terminal.ConnectivityNode rdf:resource="#651e51b8-5d3c-11e6-a47b-f45c899d1079"/>
</cim:Terminal>
<cim:TransformerWinding rdf:ID="651ef582-5d3c-11e6-8569-f45c899d1079">
<cim:IdentifiedObject.name>TW_76185022_380000</cim:IdentifiedObject.name>
<cim:TransformerWinding.x>1.0</cim:TransformerWinding.x>
<cim:TransformerWinding.ratedS>5000000</cim:TransformerWinding.ratedS>
<cim:TransformerWinding.r>1.0</cim:TransformerWinding.r>
<cim:TransformerWinding.ratedU>380000</cim:TransformerWinding.ratedU>
<cim:TransformerWinding.connectionType rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#WindingConnection.Yn"/>
<cim:TransformerWinding.windingType rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#WindingType.secondary"/>
<cim:ConductingEquipment.BaseVoltage rdf:resource="#651e4ec0-5d3c-11e6-9a88-f45c899d1079"/>
<cim:TransformerWinding.PowerTransformer rdf:resource="#651e4d80-5d3c-11e6-ae47-f45c899d1079"/>
</cim:TransformerWinding>
<cim:ConnectivityNode rdf:ID="651ef685-5d3c-11e6-b503-f45c899d1079">
<cim:IdentifiedObject.name>CN_76185022_380000</cim:IdentifiedObject.name>
</cim:ConnectivityNode>
<cim:Terminal rdf:ID="651ef7b3-5d3c-11e6-815c-f45c899d1079">
<cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
<cim:Terminal.ConductingEquipment rdf:resource="#651ef582-5d3c-11e6-8569-f45c899d1079"/>
<cim:Terminal.ConnectivityNode rdf:resource="#651ef685-5d3c-11e6-b503-f45c899d1079"/>
</cim:Terminal>
<cim:LoadResponseCharacteristic rdf:ID="651ef8ca-5d3c-11e6-bb77-f45c899d1079">
<cim:LoadResponseCharacteristic.pConstantPower>100000</cim:LoadResponseCharacteristic.pConstantPower>
</cim:LoadResponseCharacteristic>
<cim:EnergyConsumer rdf:ID="651efb21-5d3c-11e6-be4d-f45c899d1079">
<cim:IdentifiedObject.name>L_76185022</cim:IdentifiedObject.name>
<cim:ConductingEquipment.phases rdf:resource="http://iec.ch/TC57/2010/CIM-schema-cim15#PhaseCode.A"/>
<cim:ConductingEquipment.BaseVoltage rdf:resource="#651e4ec0-5d3c-11e6-9a88-f45c899d1079"/>
```

```
    <cim:EnergyConsumer.LoadResponse rdf:resource="#651ef8ca-5d3c-11e6-bb77-f45c899d1079"/>
    </cim:EnergyConsumer>
    <cim:Terminal rdf:ID="651efc2e-5d3c-11e6-92b4-f45c899d1079">
    <cim:Terminal.sequenceNumber>1</cim:Terminal.sequenceNumber>
    <cim:Terminal.ConductingEquipment rdf:resource="#651efb21-5d3c-11e6-be4d-f45c899d1079"/>
    <cim:Terminal.ConnectivityNode rdf:resource="#651ef685-5d3c-11e6-b503-f45c899d1079"/>
    </cim:Terminal>
</rdf:RDF>
```

Bibliography

[1] VDE Verband der Elektrotechnik Elektronik Informationstechnik e.V., "Deutsches Höchstspannungsnetz Übersichtsplan," https://www.vde.com/de/fnn/dokumente/documents/uebersichtsplan-2014.pdf, 2015, accessed: 2015-11-12.

[2] J. Rivera, J. Leimhofer, and H.-A. Jacobsen, "OpenGridMap: Towards Automatic Power Grid Simulation Model Generation from Crowdsourced Data," 2016, to appear.

[3] A. Von Meier, *Electric power systems: a conceptual introduction.* John Wiley & Sons, 2006.

[4] OpenStreetMap, "power - for marking and tagging facilities for the generation and distribution of electrical power," http://taginfo.openstreetmap.org/keys/power, 2016, accessed: 2016-07-12.

[5] M. Specht and S. Rohjans, *ICT and Energy Supply: IEC 61970/61968 Common Information Model.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 99–114. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-34916-4_6

[6] J. P. Snyder, *Map projections–A working manual.* US Government Printing Office, 1987, vol. 1395.

[7] B. Wiegmanns, "Improving the topology of an electric network model based on Open Data," Master's Thesis, Energy and Sustainability Research Institute, University of Groningen, The Netherlands, 2015.

[8] L. Hofman and B. R. Oswald, "Vergleich Erdkabel – Freileitung im 110-kV-Hochspannungsbereich," http://www.energie.brandenburg.de/media/bb1.a.2865.de/Gutachten_Vergleich_Erdkabel_Freileitung_110kV_Hochspannungsbereich_technische_Aspekte.pdf, 2010, accessed: 2016-08-04.

[9] European Network of Transmission System Operators for Electricity (ENTSO-E), "Common Information Model (CIM) - Model Exchange Profile 1," https://www.entsoe.eu/Documents/CIM_documents/Grid_Model_CIM/140610_ENTSO-E_CIM_Profile_v1_UpdateIOP2013.pdf, 2014, accessed: 2016-07-04.

[10] T. Sattich, "Mutually beneficial? Germany's energy transition and the internal

electricity market," in *Proceedings of the 10th International Conference on the European Energy Market (EEM)*. IEEE, 2013, pp. 1–6.

[11] Y.-J. Liu, T.-P. Chang, H.-W. Chen, T.-K. Chang, and P.-H. Lan, "Power quality measurements of low-voltage distribution system with smart electric vehicle charging infrastructures," in *Proceedings of the 16th International Conference on Harmonics and Quality of Power (ICHQP)*. IEEE, 2014, pp. 631–635.

[12] H. K. Cakmak, H. Maass, F. Bach, U. Künhapfel, and V. Hagenmeyer, "Ein Ansatz zur automatisierten Erstellung umfangreiecher und komplexer Simluationsmodelle für elektrische Übertragungsnetze aus OpenStreetMap-Daten," *Manuskript für at - Smartgrids*, vol. 03, 2015.

[13] S. Soltan and G. Zussman, "Generation of Synthetic Spatially Embedded Power Grid Networks," *ArXiv e-prints*, Aug. 2015.

[14] J. Rivera, C. Goebel, D. Sardari, and H.-A. Jacobsen, *OpenGridMap: An Open Platform for Inferring Power Grids with Crowdsourced Data*. Cham: Springer International Publishing, 2015, pp. 179–191. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-25876-8_15

[15] OpenStreetMap, "Power routing proposal," http://wiki.openstreetmap.org/wiki/Proposed_features/Power_routing_proposal, 2016, accessed: 2016-07-13.

[16] W. Medjroubi, C. Matke, and D. Kleinhans, "SciGRID - An Open Source Reference Model for the European Transmission Network (v0.2)," Nov. 2015. [Online]. Available: http://www.scigrid.de

[17] European Network of Transmission System Operators for Electricity (ENTSO-E), "ENTSO-E Transmission System Map," https://www.entsoe.eu/map/Pages/default.aspx, 2016, accessed: 2016-08-17.

[18] 50hertz, Ampiron, TransnetBW, and TenneT, "Strukturdaten," https://www.netztransparenz.de/de/strukturdaten.htm, 2016, accessed: 2016-08-10.

[19] T. Berry, "Standards for energy management system application program interfaces," in *Proceedings of the International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*. IEEE, 2000, pp. 156–161.

[20] M. M. Haklay and P. Weber, "OpenStreetMap: User-Generated Street Maps," *IEEE Pervasive Computing*, vol. 7, no. 4, pp. 12–18, Oct. 2008.

[21] OpenStreetMap, "WikiProject Power Networks," http://wiki.openstreetmap.org/wiki/WikiProject_Power_networks, 2016, accessed: 2016-07-13.

[22] ——, "OpenStreetMap Wiki," http://wiki.openstreetmap.org/wiki/Main_Page, 2016, accessed: 2016-07-12.

[23] EPRI, "Development of the Common Information Model for Distribution and A Survey of Adoption," http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001020103, 2010, accessed: 2016-07-26.

[24] M. Uslar, M. Specht, S. Rohjans, J. Trefke, and J. M. González, *The Common Information Model CIM: IEC 61968/61970 and 62325-A practical introduction to the CIM.* Springer Science & Business Media, 2012.

[25] I. TC57, "IEC 61970: Energy management system application program interface (EMS-API)," https://webstore.iec.ch/publication/6210, IEC, Tech. Rep., December 2013.

[26] ——, "IEC 61968: Application integration at electric utilities - System interfaces for distribution management," https://webstore.iec.ch/publication/6197, IEC, Tech. Rep., October 2012.

[27] MathWorks, "Simulation and Model-Based-Design," http://de.mathworks.com/products/simulink/, 2016, accessed: 2016-07-27.

[28] ——, "Simscape Power Systems," http://de.mathworks.com/products/simpower/, 2016, accessed: 2016-07-27.

[29] ——, "Choosing an Integration Method," http://www.mathworks.com/help/physmod/sps/powersys/ug/choosing-an-integration-method.html, 2016, accessed: 2016-07-27.

[30] S. Altschaffl, R. Witzmann, and T. Ahndorf, "Generating a PSS™ NETOMAC model of the German Transmission Grid from Google Earth and visualizing load flow results," in *Proceedings of the 2014 IEEE International Energy Conference (ENERGYCON).* IEEE, 2014, pp. 603–609.

[31] A. W. McMorran, G. W. Ault, I. M. Elders, C. E. Foote, G. M. Burt, and J. R. McDonald, "Translating CIM XML power system data to a proprietary format for system simulation," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 229–235, 2004.

[32] X. Wang, N. N. Schulz, and S. Neumann, "CIM extensions to electrical distribution and CIM XML for the IEEE radial test feeders," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1021–1028, 2003.

[33] D. Bytschkow, M. Zellner, and M. Duchon, "Combining SCADA, CIM, GridLab-D and AKKA for smart grid co-simulation," in *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society.* IEEE, 2015, pp. 1–5.

[34] Python Software Foundation, "Pycim 15.13.4," https://pypi.python.org/pypi/PyCIM, 2016, accessed: 2016-08-01.

[35] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams.* John Wiley & Sons, 2009, vol. 501.

[36] Z. Mohamed and P. Bodger, "Forecasting electricity consumption in New Zealand using economic and demographic variables," *Energy*, vol. 30, no. 10, pp. 1833–1843, 2005.

[37] OpenGeoDb, "OpenGeoDB — OpenGeoDb, Die freie Geoinformatik-Wissensdatenbank," http://opengeodb.giswiki.org/index.php?title=OpenGeoDB&oldid=13813, 2015, accessed: 2016-07-11.

[38] Statista, "BDEW. (n.d.). Pro-Kopf-Stromverbrauch in Deutschland in den Jahren 1995 bis 2015 (in Kilowattstunden)," http://de.statista.com/statistik/daten/studie/240696/umfrage/pro-kopf-stromverbrauch-in-deutschland/, 2015, accessed: 2016-07-11.

[39] Sénat, "Le réseau de transport en France," http://www.senat.fr/rap/r09-506/r09-5068.html, 2010, accessed: 2016-07-12.

[40] POLISH INFORMATION AND FOREIGN INVESTMENT AGENCY, "Energy Sector in Poland," http://www.paiz.gov.pl/files/?id_plik=19610, 2012, accessed: 2016-08-11.

[41] OCCTO, "平成27年度供給計画の取りまとめ," https://www.occto.or.jp/pressrelease/2015/files/150703_kyoukyukeikaku_kai.pdf, 2015, accessed: 2016-08-11.

[42] E. Allen, D. Kosterev, and P. Pourbeik, "Validation of power system models," in *2010 IEEE Power and Energy Society General Meeting.* IEEE, 2010, pp. 1–7.

[43] A. W. McMorran, "An introduction to iec 61970-301 & 61968-11: The common information model," *University of Strathclyde*, vol. 93, p. 124, 2007.

[44] N. Ammann, U. Pugfelder, and D. Fischer, "Energienetze in Bayern," https://www.bihk.de/bihk/Anhaenge/bihkrepository/energienetze-studie-langfassung.pdf, 2013, accessed: 2016-08-09.