Energy Reports 3 (2017) 14–21

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models

Wided Medjroubi[a,*], Ulf Philipp Müller[b], Malte Scharf[b,c], Carsten Matke[a], David Kleinhans[a]

[a] NEXT ENERGY, EWE Research Centre for Energy Technology at The University of Oldenburg, Carl-von-Ossietzky-Str. 15, 26129 Oldenburg, Germany

[b] Flensburg University of Applied Sciences, Centre for Sustainable Energy Systems, Germany

[c] Wuppertal Institute for Climate, Environment and Energy, Research Group Future Energy and Mobility Structures, Germany

ARTICLE INFO

Article history:
Received 30 June 2016
Received in revised form
14 October 2016
Accepted 7 December 2016

Keywords:
Transmission network modelling
Open data
OpenStreetMap
Transmission grid models
Transmission grid data
Open source software

ABSTRACT

In order to analyse the mid- and long-term impacts of energy related policies, different modelling approaches can be derived. However, the results of even the best energy system model will highly depend on the underlying input data. First, in this contribution the importance and availability issues of grid data in the context of energy system modelling are highlighted. Second, this paper focuses on power grid modelling based on open and publicly available data from OpenStreetMap using open source software tools. Two recent approaches developed to build electrical transmission network models using openly available data sources are presented and discussed. The proposed methods provide transparent assumptions, simplifications and documentation of grid modelling. This results in the ability of scientists and other stakeholders to validate, discuss or reproduce the results of energy system models. Thus the new open approaches offer a unique opportunity to increase transparency, comparability and reproducibility of results in energy system modelling.

1. Introduction

The constant increase of the renewable energy source shares in the European energy mix implies a transformation of the energy system structure and operation. This transformation is directly induced by the shift from central to de-central energy production and from unidirectional to bidirectional energy flows. In Germany, the goals of the energy transition are a share of 40%–50% of renewable energy production until 2025 and a share of 50%–60% in 2035.[1] For Europe, the objectives is to have a 86% share of renewable in electricity production.[2] Those ambitious goals require a strong and modern electricity grid implying the modernisation and expansion of the current grid (Nikoletatos and Tselepis, 2015; EnerNex Corporation, 2011; Sims et al., 2011; American Physical Society, 2010). The EU scenarios assume investments of more than 50 billion EUR in the transmission grid infrastructure implying fundamental changes in the electricity grid.

The progress of the electrical grid expansion has direct political, economical, socio-cultural and environmental impacts. On the medium and long term, it may also influence investment decisions in current and future generation facilities. Due to these factors, electrical grid extension additionally requires public acceptance (Roland Berger Strategy Consultants, 2015). As a basis for political and strategic decisions, different studies on the impacts of renewable energy sources on the future design of energy systems were commissioned (RESTORE2050, 2014; Bode, 2010; EmployRES Project Consortium, 2009; European Commission, 2007). In addition, the funding of research projects addressing this particular topic increased significantly (Bundesministerium, 2015; RWTH Aachen University, 2014; BMU, 2010).

The process of expanding transmission systems (including but not limited to: grid operation, stability, control, congestion management and integration studies) heavily relies on modelling. Moreover, public acceptance of such expansion depends on the transparency of the underlying strategies and decisions (Ciupuliga and Cuppen, 2013). Those factors added to the emergence of wholesale electricity markets imply an increasing importance of energy system modelling tools in general and electricity market

* Corresponding author.
E-mail address: wided.medjroubi@next-energy.de (W. Medjroubi).

1 Federal Ministry for Economic Affairs and Energy.

2 EU Commission: Energy Roadmap 2050.

http://dx.doi.org/10.1016/j.egyr.2016.12.001

modelling tools in particular. These tools involve a representation, or a model, of the electrical grid. However, all grid models rely on the quality and availability of the electrical grid data they represent.

In this contribution, we present and discuss recent efforts towards developing open grid models and using open data in electrical grid modelling. In Section 2, a general introduction to grid models, their relevance and use in load flow calculations and energy system models are presented. An overview on the data requirements for grid models and grid data availability are also introduced. The current status of open grid models and data are discussed in Section 3. Section 4 introduces OpenStreetMap and the power data available in its database. Two approaches used to derive grid models from OpenStreetMap data are also presented: a deterministic and heuristic approach including their key features. The comparison of the two approaches, as well as their advantages and limitations are discussed. Finally, in Section 5 conclusions are presented.

2. Grid models and their input data

2.1. Grid modelling

Energy system models analyse energy systems or a number of sub-systems (e.g. power system). These models became increasingly relevant for decisions in energy acquisition and trading. According to Hoffman and Wood (1976), "Energy system models are formulated using theoretical and analytical methods from several disciplines including engineering, economics, operations research and management science". Multiple solution methods involving "mathematical programming (especially linear programming), econometrics and related methods of statistical analysis and network analysis" are used to solve the models equations (Hoffman and Wood, 1976). Energy system modelling involves modelling the transmission grid in an appropriate way using grid models.

In general, grid models are developed and used for different purposes and can have many focuses. In this paper, the authors differentiate between two applications where grid models are used, namely *grid simulation models* and *electricity system models*. The focus of grid simulation models lies on the technical and physical behaviour of electric grids. Electricity system models are tools used to manage, plan and extend electricity systems, manage electricity demand as well as electricity trading. They model the physical behaviours of the underlying grids as well as market mechanisms.

Grid models are in general classified in four different types (Pfluger, 2014; Ventosa et al., 2005; Villasana et al., 1985). The simplest grid model is *the single-node model*, also known as *the copper plate approach* (Tande et al., 2008; Schaffner and Mihalic, 2005). This simple approach assumes an unconstrained electrical grid, and is mostly used in economical models. The second type of grid models is *the transshipment model* (Nygard et al., 2011), where a various number of *nodes* or *regions* are defined. Between the defined regions an exchange of power is possible and is constrained by the net transfer capacity. This model neglects the existence of the physical power flow principles. A more sophisticated model is the (linear) *DC model* (Schavemaker and Sluis, 2008), where a network of several nodes and interconnections (or power lines) is defined. Using Kirchhoff's law, active power flows can be determined, which depend on the resistance and the maximal capacity of the power lines. A more realistic way to model AC grids is to consider an *AC power flow model* (Schavemaker and Sluis, 2008). In this approach the active and reactive power flows are modelled. Therefore, the reactance of the power lines is highly relevant and both capacitive and inductive behaviour of the power lines are considered. However, the increase of detail and accuracy is usually accompanied with an increased complexity resulting in longer simulation times. An overview of the characteristics and level of details of the grid data needed for the above cited models is presented in the following section.

2.2. Data requirements of grid simulation models

The input data needed by grid models (called grid data) and their requirements highly differ depending on the type of the grid model used. In *single-node models*, active power data of generation and demand are needed. Information about grid topology and its electric parameters are neglected. *Transshipment models* require an abstract grid topology using nodes, power lines and regions. The only electrical parameters considered are the net transfer capacities of power lines connecting the different regions considered. Active power of demand and generation are considered spatially differentiating between the defined nodes or regions. Losses can be addressed to each line as an assumed percentage (Wiese, 2015). In *DC models* in addition to topological information of the grid, electric parameters of power lines have to be defined. These parameters, which are essential to perform DC load flow analysis are the resistance (R [Ω]) and thermal limit current (I_{th} [A]) of each power line.

The *AC power flow model* requires by far most input data in the scope of the relevant grid model types defined in this paper. Global frequencies (f [Hz]) and nominal voltage levels (V_n [kV]) are set. Each node needs to be defined either as a *PQ* or as a *PV* bus which states whether reactive power (Q [var]) or voltage (V [kV]) are known (Rendel, 2015; Milano, 2010) at this bus. Demand and generation data have to be additionally specified by their reactive power or voltage behaviour. The power lines, apart from their active resistance, are defined by their reactive behaviour, hence information of their reactance (X[Ω]) is necessary. Specifying this behaviour either the capacitance (C [nF]) or inductance (L [H]) have to be defined as well (Agricola et al., 2012). The thermal limit currents are also crucial for determining utilisation rates of power lines. In order to simulate multiple voltage levels at once, transformer data are also needed. Apart from the normal line characteristic (R, X, C), additional parameters define the transformer's behaviour allowing to modify the magnitude and the phase angle of the voltages at the nodes being connected. Most importantly the primary and secondary voltage rating, iron losses, tap ratio and phase shifts have to be additionally provided for transformers (Milano, 2010).

2.3. Grid data availability

After highlighting the requirements for grid data in the previous section we discuss here its availability for scientific and technical studies. Energy systems modelling accuracy and results depend of the input data available. In the case of grid modelling, this data includes in general the structure and electrical properties of the grid, called *grid data* commonly generated using a *grid model*. Other data are necessary in energy systems modelling including electricity generation capacities, electrical loads estimation and electricity price data. Note that, the details and scale of the required data depends on the modelling temporal and spatial scales. Generally, neither grid models nor grid data are publicly available. Moreover, the details and derivation of grid data used in modelling are not accessible. However, data is an important resource in energy system modelling as models are only of use if they are provided with valid input data. How can the quality of energy system modelling be evaluated if there are no or few information about input data, its quality, and the assumptions and simplification considered in their derivation (Egerer et al., 2014; Wiese et al., 2014)? How can the results of simulations be verified, validated or compared to other models and assumptions? These issues extend to the question of reproducibility of the results which is at best limited. On the other hand, data collection, maintenance and update are tedious processes which have to be performed by different institutions. This process is inefficient as there is no possibility to reuse data from other sources although data has been collected in many publicly financed projects. Institutions must dedicate human and technical resources for data collections tasks which constitute an overhead to be considered when conducting research. This is more of a burden when the goal of the research project is not data collection per se, but a pre-processing step for simulations. Due to the previous mentioned reasons, public and scientific acceptance of energy models results and the ensuing political and economic decisions are therefore limited. This can lead to socio-economic problems on a national and international level. A prominent example is the resistance to the extension of the power network in Germany (Neukirch, 2014; Cotton and Devine-Wright, 2013) and to the erection of wind turbines (Jobert et al., 2007). These issues motivated the present work and the use of publicly available data in deriving power grid models and input data for power flow simulations.

A central reason for grid data unavailability and lacking of transparency is data accessibility. Two aspects are often cited justifying non-disclosure of grid data. The security aspect, as sensitive data, especially power plants location can be used to harm these facilities and disrupt their operations (McDaniel and McLaughlin, 2009; Tang and McMillin, 2008; Salmeron et al., 2004). The economic aspect, as grid data (which can be used in energy system models) can be used to calculate revenues of power operators (Egerer et al., 2014). Currently grid operators seem to be rather protective with the actual grid data and often do not agree to publish it. This lack of data extends also to load data, renewable energy sources locations and power plants locations and specifications. In some projects or bilateral contracts, research institutes may obtain limited access to particular grid datasets.

Grid data availability in Europe slightly improved with the release of partial datasets by ENTSO-E,[3] the German Network Agency[4](BNetzA), and the British National Grid[5] data. The first two datasets are however released under publication and distribution restrictions and are not geo-referenced, which makes them virtually unusable as they cannot be related to generation and load centres. In 2011, the amendment of the *National Energy Legislation* in the UK[6] made third parties access to electricity (and gas and oil) planning data accessible. This resulted in an impressive disclosure of different datasets relating to the transmission network. The power grid released includes geo-locations of overhead lines, towers, transmission cables, underground cables and electrical substations at different voltage levels.

Some transmission operators in Europe partially released datasets of their respective networks. As examples we cite the following operators datasets. The Austrian operator (APG) released a static network dataset as a pdf document. It includes the power lines (represented by their connecting substations), their length, voltage level and their electrical properties. Although the name of the substations is indicated, their geo-coordinates are not provided. The same is valid for the static network dataset released by TenneT Germany, Amprion Transnet BW, 50 Hz and RTE, although the data format is more convenient (dataset provided as csv file). Other operators are still lagging behind and offer only pdf maps of their grid. A very important information which seems to be always lacking are the geo-coordinates of the grid elements (substations, transformers and lines). Despite the partial releases, the problem remains that grid models and data are typically not transparent, implying a discriminating distribution of information.

3. Open grid models and open grid data

Many publications mention the data problematic in energy modelling in general. In her work on cross-border congestion modelling and management in electricity markets, Zhou (2003) mentions the difficulty in obtaining grid data. The author goes through the painstaking work of digitising and geo-referencing the ENTSO-E (formerly UCTE) transmission network map (status 2001). The power plant data and load information were obtained from different public databases. The electrical properties of power lines were deduced using standard value taken from the literature. Using a network topology as close as possible to the *real* transmission grid was important as cross-border congestion management was investigated. The grid data were made available under power factory format.[7] The author conclude that grid data were insufficient and does not represent the actual status of the power grid. An updated model introducing several changes was published more recently (Hutcheon and Bialek, 2013).

In Wiese et al. (2014) the lack of transparency of energy models used and its impact on the (public) acceptance of the results are discussed. As well as the importance of open source modelling and a review of energy models are presented. The degree of openness of available models is discussed, in terms of source code and solvers. Wiese introduces the open source model *renpass* (Renewable Energy pathways Simulation System),[8] developed at Flensburg University. Renpass simulates the electricity supply and use of the grid infrastructure with high time and spatial resolution. The code uses *MySQL* database and is written in *R*. *EnergyMap*[9] data are used to determine installed capacities and renewable energy sources locations. The grid data are derived from the ENTSO-E grid map. However, no information is provided about how the data is derived. Other datasets needed for the simulations (power plants data, costs, etc.) are taken from publicly available databases. *renpass* is freely available under the *GNU GPL3* license.

In Egerer et al. (2014), Egerer et al. address both the grid modelling and data issues. They highlight electricity markets modelling transparency as a major contributor to understand electricity markets. Transparency is set as *a requirement* for public acceptance. The recent development of grid data in the industry, Transmission System Operators (TSOs) and at academic level is also mentioned. The paper's objective is "to provide insights into the public availability of data sources used for electricity market and transmission network modelling of the German and European power system". Information about where to obtain load data, cross-border flows, electricity generation and price data is provided. Some data are only available as figures and tables, and grid data are not available. A short description about grid data (for Germany) being obtained using grid operators maps and OpenStreetMap is presented. The details of how grid data are obtained are not available. The authors aim to encourage other actors to contribute by publishing grid data. A real-world application is presented using the dataset discussed using the ELMOD model (Leuthold et al., 2008). An estimate of the hourly electricity prices for Germany is presented as a case study. For

3 European Network of Transmission System Operators for Electricity, ENTSO-E.
4 German Federal Network Agency.
5 National Grid: Data Explorer.
6 Department of Energy and Climate Change.
7 Bialek Continental European Transmission Network Dataset.
8 renpass: Renewable Energy pathways Simulation.
9 energyMap.info.

Europe, the data are used to compare generation mix and trade flows between European countries.

A trend towards open source grid models and open grid data is developing in the electricity system modelling community.[10] The awareness of the importance of public acceptance as well as the (scientific) reproducibility of the model's results have dramatically increased. Many recent scientific modelling projects like renpass, GENESYS (Alvarez et al., 2013), SciGRID (Medjroubi et al., 2014), open_eGo (NEXT ENERGY, 2016), openMod.sh (Europa University, 2015) and Open Power System Data (Open Power System Data, 2015) and OpenGridMap[11] focus on open source software and/or open grid data. Moreover, several grid simulation software packages are available in the open source domain such as MATPOWER (Zimmerman et al., 2011) which is written in Matlab and PYPOWER (Lincoln, 2015) and PyPSA (FIAS, 2015) both written in Python.

4. OpenStreetMap based power grid models

Recent efforts to build and use the open and freely OpenStreetMap data to model grid data are presented in this Section. Two approaches are highlighted namely *SciGRID* and *osmTGmod*.

4.1. OpenStreetMap

OpenStreetMap (OSM) is "a free and editable map of the world" (Bennet, 2010). OSM is a Voluntary Geographic Information (VGI) project aiming to build a free and accessible geographic database of the world. Created in 2004 by Steve Cost, the OSM database is continuously growing in data details, diversity and number of users.[12] OSM data includes geo-referenced features which can be mapped, e.g. roads, buildings, power plants, wind turbines, administrative boundaries, waterways, forests, etc. Data are collected on a voluntary basis by data "mappers". There exist many ways to collect geo-references data (Ramm et al., 2011; Bennet, 2010). After data are gathered, it is linked to the OSM database by using editing tools (like JOSM or Potlatch2). The database is available under the share alike Open Database License 1.0 (ODbL)[13] allowing data download, modification and sharing. This is a very important advantage for models using OSM data as it permits sharing and distributing the input data as well as the resulting data.

OSM data are hierarchically structured and have three types: nodes, ways and relations. The nodes are points in space which are defined by their geographical coordinates. Ways are an ordered list of nodes, which define none-closed features (like a transmission line) or closed features (like buildings, power plants, electrical substations). A relation is the most complex data type in OSM and is an ordered list of nodes, ways and even other relations. Relations are used to represent a spatial or logical association relating their different components, e.g. a bus route that contains multiple bus stops and road parts. All OSM data types are associated with *tags*, dictionary-like entries having a *key* and a *value* text attributes. Tags describe a specific detail of the data, e.g. a node representing a power carrying tower have the `key=value` combination `power=tower`. The data format is XML-based and follows an XML schema definition which lists OSM data types (Ramm et al., 2011). An excerpt of an XML-data representation of a substation is shown on Fig. 1. Several open source tools can be used to process OSM data.


```
"tags": {
  "access": "no",
  "addr:city": "Dollern",
  "addr:country": "DE",
  "addr:housename": "Umspannwerk Dollern",
  "addr:postcode": "21739",
  "addr:street": "Hagener Weg",
  "barrier": "fence",
  "frequency": "50",
  "name": "Dollern",
  "name:en": "Dollern Substation",
  "power": "substation",
  "source": "landsat",
  "substation": "transmission",
  "voltage": "380000;220000;110000"
}
```


Fig. 1. An excerpt of the XML definition of a way defining an electrical substation in OpenStreetMap (OSM-ID = 24502420). Credits: ©OpenStreetMap contributors.

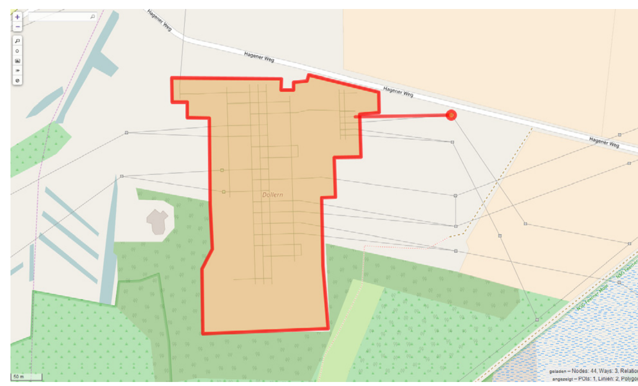

Fig. 2. An example of power elements in OSM highlighted in red: substation (closed-way OSM-ID = 24502420) connected to a power line (way OSM-ID = 318302070) and a power pole (node OSM-ID = 377980387). Credits: ©OpenStreetMap contributors, overpass-turbo.eu. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.2. Power (Grid) data in OpenStreetMap

Power data in OSM is represented by the OSM types mentioned earlier (*nodes*, *ways* and *relations*). Nodes can represent line-carrying towers and electrical poles. Transmission lines and underground cables are represented by open ways (see Fig. 2), while substations, generators and power plants are closed ways (see Fig. 2). Power relations represent electrical circuits and are constituted of one or several towers, substations and transmission lines (see Fig. 3). OSM power relations have the `key=value` combination `route=power`. On the 12.07.2015 there are 3111 power relations in Germany. Available power data in OSM include, beside the geo-location of the different components, additional information about e.g. the voltage level, frequency, the transmission line wire types and numbers, and transmission operator.

There are several *quality assurance tools* available to insure a better quality of OSM data by detecting bugs automatically.[14] However, power data extracted from OSM can be incomplete, missing, outdated or even erroneous. This needs to be taken into account when using the derived data in grid models. Nevertheless, the coverage and the quality of OSM data are in continuous progress (Zielstra and Zipf, 2010; Haklay, 2010; Ather, 2009).

10 Open Modelling Initiative.

11 OpenGridMap.com.

12 OpenStreetMap Wiki: Statistics.

13 Open Database License 1.0.

14 OpenStreetMap Wiki: available quality assurance Tools.

Fig. 3. Example of a power relation in OSM: relation with OSM-ID = 1188848 is composed of two substations (represented by square-like closed ways in red) and seven transmission lines (represented by the linear features in red colour). Credits: ©OpenStreetMap contributors, overpass-turbo.eu. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Concerning the extra high voltage levels (voltages equal or greater than 220 kV) the power data coverage in Germany is around 95% (the comparison is based on the total length of the transmission network[15]). For the high, medium and low voltage levels (voltages below 220 kV) the data quality and coverage decreases. This is because of the higher extent and complexity of the power grid at the lower and medium voltages.[16] Moreover, the access to the lower voltage power elements (especially underground cables and transformers boxes). This should be considered when using OSM data for modelling the medium and low voltage power grid.

4.3. Building power grid model using OpenStreetMap

In this section, two different approaches for generating grid models based on OSM data are presented. The two approaches were developed by the authors of this paper. *SciGRID*, developed at NEXT ENERGY - EWE Research Centre for Energy Technology (in the Energy System Analysis Division), consists of an automatised process to filter, extract and abstract OSM data to a power grid model in a transparent and traceable way. Its focus lies on *attribute-based abstraction* using exclusively *power relations*. *osmTGmod*, developed at the Wuppertal Institute for Climate, Environment and Energy (at the research group future energy and mobility structures) in cooperation with Flensburg University of Applied Sciences at the Centre for Sustainable Energy Systems (ZNES), is based on a similar process to filter and extract OSM data, but focuses on a *geometry-based abstraction* in an attempt to address the limited power relations coverage. The two approaches will be compared in Section 4.4.

4.3.1. SciGRID approach

The *SciGRID* project was initiated in October 2014 (Medjroubi et al., 2014) to address lack of transmission grid data for scientific purposes. In its first release, the SciGRID abstraction code, the documentation and a dataset of the power transmission network (for Germany) were made available. SciGRID uses OSM power data and filters for extra high voltage levels making use of the *power* tag. It exclusively makes use of *power relations* which include information on the electrical circuits of the transmission network. The procedure used in *SciGRID* is completely automatised and supports OSM data updates. This approach is chosen as it allows

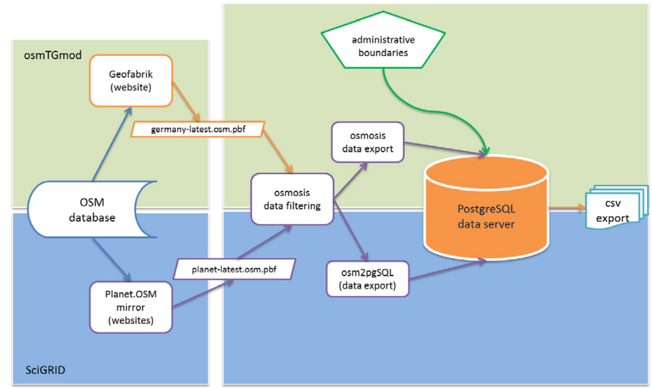

Fig. 4. SciGRID and osmTGmod work flow charts displaying the different steps and open source software involved in both models.

for reproducibility and transparency in building a transmission network using OSM power data. The different stages involved in the *SciGRID* code are briefly introduced in the following (see also Fig. 4).

In *SciGRID* the raw OSM data is downloaded and filtered with respect to the *power* tag using the open source command-line Java tool *Osmosis*.[17] The filtered data are the extracted from the raw data and particularly *power relations* are used. The power-filtered data are exported to relational databases (or tables) using the open source command-line program *osm2pgsql*[18] which converts OSM data to postGIS-enabled *PostgreSQL* (PostgreSQL, 1996) databases. Once the data are stored in the *PostgreSQL* database *SQL* queries are used to abstract the filtered substations to their geometrical centres. They constitute the vertices of the transmission network. The transmission lines between two vertices are abstracted to direct connections with individual lengths calculated from the detailed layout in OSM, also performed using *SQL* queries. The abstracted transmission lines constitute the links (or edges) of the transmission network. Additionally, the voltage level, number of cables and wires of the transmission lines are adopted from OSM data. The electrical properties of transmission lines (resistance, reactance, capacitance and maximum thermal limit current) are then modelled using the following equations:

$$R\ [\Omega/\text{km}] = C_r \div \left(\frac{wires}{wires_{typical}}\right) \div \left(\frac{cables}{3}\right) \quad (1)$$

$$X\ [\Omega/\text{km}] = C_x \div \left(\frac{wires}{wires_{typical}}\right) \div \left(\frac{cables}{3}\right) \quad (2)$$

$$C\ [\text{nF/km}] = C_c \cdot \left(\frac{wires}{wires_{typical}}\right) \cdot \left(\frac{cables}{3}\right) \quad (3)$$

$$I_{th}[\text{A}] = C_I \cdot \left(\frac{wires}{wires_{typical}}\right) \cdot \left(\frac{cables}{3}\right) \quad (4)$$

where, $wires_{typical}$ is the number of wires in a transmission cable which is typically 2 for transmission lines of 220 kV and 4 for transmission lines of 380 kV. The coefficients: C_r, C_x, C_c and C_I are typical values listed in Deutsche Energy-Agentur (2012) and Oeding and Rüdiger (2004).

An example of the transmission network obtained for Germany is shown in Fig. 5. The resulting vertices and links data are available to download as csv files.

15 German Federal Ministry of Economic Affairs and Energy.
16 German Federal Association of Energy and Water Management (BDEW).
17 OpenStreetMap Wiki: Osmosis tool.
18 OpenStreetMap Wiki: Osm2pgSQL tool.

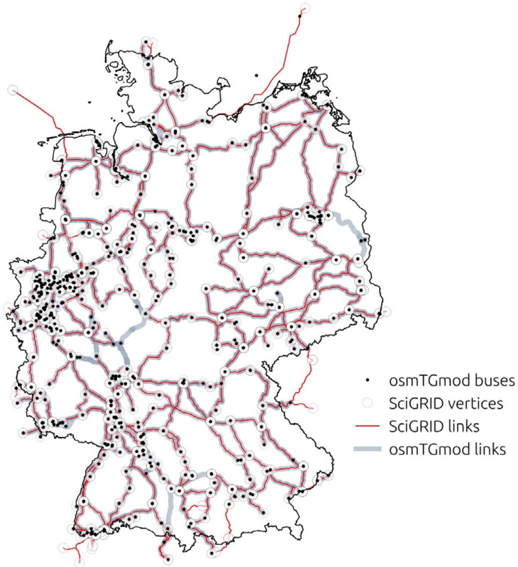

Fig. 5. Topological transmission network for Germany obtained by SciGRID and osmTGmod. Data status: 08 May 2016.

4.3.2. osmTGmod approach

Similarly to *SciGRID*, *osmTGmod* is a model of the German transmission grid based on OSM data. Due to the lack of grid data for scientific purposes, *osmTGmod* was developed at the *Wuppertal Institut* for internal use. However, since there is scientific interest in methods for generating grid models based on open data, *osmTGmod* was made available (in a preliminary version) under the Apache License, Version 2.0.[19]

osmTGmod uses OSM data for Germany obtained from the *Geofabrik* web-server[20] (see Fig. 4). *Osmosis* is then used to filter the raw OSM data and import it into a *PostgreSQL* database (contrary to *SciGRID* which uses *osm2pgsql* for this step). Within this database, the power grid is abstracted from the raw data. The abstraction process is implemented in *SQL* and PostgreSQL's procedural language *pl/pgSQL*, whereas *osmTGmod*'s user interface is implemented in *Python*. The resulting grid is stored in the database and can be visualised using *QGis* or exported as csv tables (see Fig. 4). Furthermore, *osmTGmod* allows the user to enter grid development measures, that can be included in the abstraction process.

As not all transmission lines are covered by power relations, *osmTGmod* pursues the philosophy of evaluating ways *and* relations in a complementary way: As a first step, *osmTGmod* evaluates as much information as possible from the available power relations and abstracts an electrical grid from them. This approach is rather similar to the *SciGRID* model. Since relations are not completely mapped, this grid model is incomplete. In order to use as much data as available for the grid, as a second step, missing information is derived from ways. All the transmission lines that are not covered by relations are added to the grid. Since no wiring-information can be derived from ways, these *remaining* transmission lines are *wired* using heuristics. The described process allows the generation of a mostly circuit-specific model, without, however, neglecting information about transmission lines that are not covered by relations. This procedure may be even more valuable when modelling grids in areas where the coverage of relations within OSM is as complete as in the case of Germany.

osmTGmod abstracts the grid based on the geometry of the power relations (and the remaining transmission lines). The topology of the grid is derived from the geometries of these objects.[21] The geometry of substation objects is used to connect transmission lines and thus, abstract network nodes. *osmTGmod* makes use of heuristics in certain cases. E.g. it is assumed that adjacent transmission lines must have the same number of conductors so that they can inherit information from their *neighbours* or that transmission lines, which end without connection very close to a substation are connected to that substation. During the abstraction, underlying assumptions, heuristics and potential data errors are documented in a log-file. At the end of the abstraction process, the model is tested, whether it represents one connected or several unconnected grids. All grids which are not connected to the main grid are deleted. This simplification allows *osmTGmod* to be computable by AC load flow simulation software such as *MATPOWER* or *PYPOWER*.

4.4. Comparison of SciGRID and osmTGmod

A preliminary and qualitative comparison between *SciGRID* and *osmTGmod* is attempted within this chapter. A comparison of the two approaches with a reference grid model is however not possible due to the absence of such model. However, a quantitative assessment and comparison is planned for both approaches in the future.

Applying the *SciGRID* and *osmTGmod* approaches on the OSM dataset for 08 May 2016 resulted in the following number of vertices (buses) and links (branches): *SciGRID*: 515 vertices 833 and links, *osmTGmod*: 461 vertices and 1256 links (see Fig. 5). The number of vertices obtained by both approaches is comparable. The higher number of links in *osmTGmod* is due to three reasons: first, links in *osmTGmod* are separated according to their voltage level which means that "duplicates" can exists if a link carries more than one voltage, while in *SciGRID* the links are not separated according to their voltage level. Second, *osmTGmod* include transformer lines which are not mapped in OSM and not included in SciGRID but which are automatically added to connect vertices of different voltage levels in *osmTGmod*. Third and most importantly, *osmTGmod* captures more links due to using all available OSM power data and not only power relations as in SciGRID. Note also that, *SciGRID* includes cross-border vertices and links.

The main difference between the *SciGRID* and *osmTGmod* approaches lies in the primary data used for the abstraction process. *SciGRID* is based on power relations only and *osmTGmod* uses power relations and non-relations data by applying heuristics. While *osmTGmod* adds missing non-relations data by expanding information from neighbouring elements, no data are added in *SciGRID*. This implies that, in principle *osmTGmod* in its present structure can be extended to derive lower voltage grids, and *SciGRID* is more suitable for deriving high voltage grids. Moreover, *osmTGmod* can be used to derive grid data in regions where the availability of power relations is limited.

Using power relations results in a fully automated abstraction process without additional assumptions as transmission circuits are explicitly available. Another advantage in *SciGRID* geometrical errors do not induce errors in the model. However, using only relation data limits the completeness of the model to the regions where such relations are present and cannot resolve missing data.

Applying heuristic approaches can result in different grid topologies depending on the assumptions used in these approaches. *osmTGmod* assumes that the electrical network topology,

19 osmTGmod github repository.

20 Geofabrik download website

21 In OSM, relations have no particular geometry. However, their geometry can be derived from ways which are represented as *members* of a relation.

which is not covered by relations, can be derived from the geographical topology. Although few topological and geometrical inaccuracies can occur, they constitute further sources of errors. The use of information about lines and substations does not provide information about circuits. However, both *SciGRID* and *osmTGmod* suffer from shared issues as they are both using OSM data as input.

Note that a heuristic approach to include OSM power data not covered by *power relations* and, hence, not included in *SciGRID* is currently developed in the context of the *SciGRID project*. This open source approach, called *GridKit* is pre-release with documentation and available for download.[22] *GridKit* uses spatial and topological analysis to transform power mapped objects from OSM into a network model but to the contrary of *osmTGmod* does not make use of power relations.

4.5. Issues with OSM data

OSM relies in VGI collected data which implies issues with the data quality, completeness and accuracy. Several publications dealt with these issues for different applications, as an example for streets mapping (Haklay, 2010), buildings footprints (Fan et al., 2014) and land use (Arsanjani et al., 2015). The risks associated with these issues have to be taken into account when using OSM data including the power data. Moreover, assumptions and simplifications are needed to complete or infer missing or erroneous OSM data. These assumptions need to be made available to the users of applications and models built with or using OSM data. Some attempts were made to contribute to the quality assessment of OSM data (Mooney et al., 2010; Barron et al., 2014; Mondzech and Sester, 2011). We discuss here some issues encountered while deriving grid models for the special case of using OSM power data.

Power tags

An important issue is that mandatory tags for power elements in OSM are insufficient for their detailed use. For example, there are no mandatory tags for mapping power substations, only 4 recommended tags.[23] The same is true about power lines and power plants. Transmission lines are also mapped with different voltage levels, which should match the cables tags. An example is a transmission line with *voltage* tag *110000;220000*. The tag *cables* should also include two values separated by a semi-colon to indicate the number of cables per voltage level, for example cables = 6;3 for a line having two circuits of the first voltage and one circuit of the second voltage. However, this recommendation is not followed by mappers in all cases resulting in the difficulty of defining the number of circuits present.

Circuit routing

Transmission lines are not mapped in OSM with the *electrical* branching details referring to an explicit definition and allocation of electrical circuits (see Fig. 6) (OSM Wiki, 2014). The *electrical* details of circuits are important when running power flow simulations. Apart from being relevant for the simulation at the normal operating point it is especially of interest when performing simulations on the compliance of the $n - 1$ criterion.

The way power data are mapped and tagged is very important and must be taken into account when deriving a grid model using OSM data. This constitute a challenge for power grid modellers and users of the model data. However, one can encourage power data mapping by informing OSM communities and especially

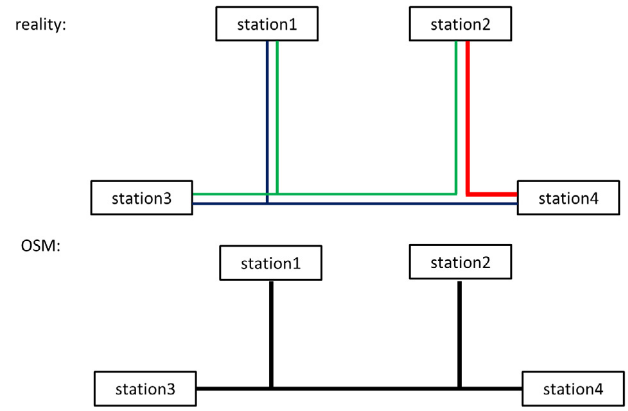

Fig. 6. Central circuit routing issue when using OSM data.
Source: Figure adapted from OSM Wiki (2014).

the mappers about this issue and explaining the data usage. Such attempts have been already conducted.[24] Only solutions accepted by the OSM community have a chance to be effective and sustainable. Therefore, power data improvement has to be done by mappers and should be mirrored back to the OSM database.

Wrong/missing data

One example of missing data are the missing details of substations which are not represented in OSM. The different elements presented in a power substation which can be transformers, switches and circuit breakers are important when more modelling details are needed. Due to the layout of power substations, it is not possible to extract such details from OSM database.

5. Conclusion

The current energy debate in Europe include the importance of developing optimal architectures for future electricity markets (Bundesministerium, 2016; ACER, 2012). An important factor to evaluate the different possible scenarios is energy system modelling. In this contribution we discussed the importance of transmission grid modelling and data issues in energy system modelling. We presented the use of open and publicly available data as an alternative data source to enhance energy system modelling transparency. Two approaches were introduced, which used OSM data in two different ways. They were compared and their strengths and weaknesses discussed. Although some power data is missing, the quality of the data obtained in SciGRID and osmTGmod is sufficient in modelling a realistic transmission system. However, data validation with other available datasets is a necessary step towards setting open grid models as standard models.

Although the OSM database is not a power database, OSM power data quality still is the best alternative to date in deriving open grid models and grid data. The objective of this contribution is to enhance transparency of transmission network modelling and data issues and to open the debate for more actors in the energy sector.

Acknowledgements

The authors thank Arjuna Nebel for constructive comments and valuable additions to the paper, in particular with respect to

22 GridKit github repository.
23 Substation tag in OpenStreetMap.
24 SciGRID webpage articles: Power Relations in OpenStreetMap, and OSM Data Coverage.

the openTGmod-model. W. Medjroubi, C. Matke and D. Kleinhans acknowledge funding by the German Federal Ministry of Education and Research (BMBF) (funding initiative *Zukunftsfähige Stromnetze*, project SciGRID, funding code 03SF0471).

References

Agency for the Cooperation of Energy Regulators (ACER), 2012. Remit (regulation on wholesale energy markets integrity and transparency), available at http://www.acer.europa.eu/en/remit.

Agricola, A.Cl., Höflich, B., Richard, P., Völker, J., Rehtanz, C., Greve, M., Gwisdorf, B., Kays, J., Noll, T., Schwippe, J., Seack, A., Teuwsen, J., Brunekreeft, G., Meyer, R., Liebert, V., 2012. Ausbau- und Innovationsbedar der Stromverteilnetze in Deutschland bis 2030 (kurz: dena-Verteilnetzstudie): Endbericht.

Alvarez, R., Thien, T., Moos, M., Chen, H., Cai, Z., Leuthold, M., Sauer, D., Moser, A., 2013. Optimal allocation and capacity of renewable energies, storage systems and transmission grid in the future European power system. In: ETG-Kongress. VDE Verlag GmbH.

American Physical Society, , 2010. Integrating Renewable Electricity on the Grid, Tech. rep.. Committee on Energy and Environment, online available at: https://www.aps.org/policy/reports (last access 30.09.2016).

Arsanjani, J., Mooney, P., Zipf, A., Schauss, A., 2015. Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets. In: OpenStreetMap in GIScience. Springer, pp. 37–58.

Ather, A., 2009. A Quality Analysis of OpenStreetMap Data (ME Thesis), University College London, London, UK.

Barron, C., Neis, P., Zipf, A., 2014. A comprehensive framework for intrinsic OpenStreetMap quality analysis. Trans. GIS 18 (6), 877–895.

Bennet, J., 2010. OpenStreetMap: Be Your Own Cartographer. Pakt Publishing.

Bode, S., 2010. Integration der erneuerbaren Energien in das Stromversorgungssystem. Wirtschaftsdienst 10, 643–660.

Bundesministerium für Wirtschaft und Energie, Bundesministerium für Bildung und Forschung, Zukunftsfähige Stromnetze: Forschungsinitiative der Bundesregierung, available at: http://forschung-stromnetze.info/ (2015).

Bundesministerium für Wirtschaft und Energie, BMWi, 2016. Electricity Market 2.0, An Electricity Market for Germany's Energy Transition, available at: http://www.bmwi.de/.

Ciupuliga, A., Cuppen, E., 2013. The role of dialogue in fostering acceptance of transmission lines: the case of a France–Spain interconnection project. Energy Policy 60, 224–233.

Cotton, M., Devine-Wright, P., 2013. Putting pylons into place: a UK case study of public perspectives on the impacts of high voltage overhead transmission lines. J. Environ. Plann. Manage. 56 (8), 1225–1245.

Deutsche Energy-Agentur (dena), , 2012. Ausbau- und Innovationsbedarf der Stromverteilnetze in Deutschland bis 2030, Tech. rep.. dena.

Egerer, J., Gerbaulet, C., Ihlenburg, R., Kunz, F., Reinhard, B., von Hirschhausen, C., Weber, A., Weibezahn, J., 2014. Electricity sector data for policy-relevant modeling: Data documentation and applications to the German and European electricity markets, Data Documentation 72, DIW Berlin, German Institute for Economic Research.

EmployRES Project Consortium, , 2009. EmployRES: The Impact of Renewable Energy Policy on Economic Growth and Employment in the European Union, Tech. rep.. European Commission, available at: http://www.isi.fraunhofer.de.

EnerNex Corporation, , 2011. Eastern Wind Integration and Transmission Study, Tech. rep.. EnerNex Corporation, available at: http://www.nrel.gov/.

Europa University, 2015. Flensburg, openMod.sh: Open Source Energy Model of Schleswig-Holstein.

European Commission, 2007. An Energy Policy for Europe, available at: http://eur-lex.europa.eu/.

Fan, H., Zipf, A., Qing, F., Neis, P., 2014. Quality assessment for building footprints data on OpenStreetMap. Int. J. Geogr. Inf. Sci. 28 (4), 700–719.

Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU), , 2010. Cost and Benefit Effects of Renewable Energy Expansion in the Power and Heat Sectors, Tech. rep.. BMU.

F.I. for Advanced Studies, 2015. PyPSA: Python for Power System Analysis, available at: http://http://pypsa.org/.

Haklay, M., 2010. How good is OpenStreetMap information? A comparative study of OpenStreetMap and ordinance survey datasets for London and the rest of England. Environ. Plann. B 37, 682–703.

Hoffman, K., Wood, D., 1976. Energy system modeling and forecasting. Annu. Rev. Energy 1 (1), 423–453.

Hutcheon, N., Bialek, J., 2013. Updated and validated power flow model of the main continental European transmission network. In: PowerTech (POWERTECH), 2013 IEEE Grenoble, pp. 1–5.

Jobert, A., Laborgne, P., Mimler, S., 2007. Local acceptance of wind energy: Factors of success identified in French and German case studies. Energy Policy 35 (5), 2751–2760.

Leuthold, F., Weigt, H., Von Hirschhausen, C., 2008. ELMOD-a Model of the European Electricity Market, Tech. rep.. Dresden University of Technology.

Lincoln, R., 2015. Pypower 5.0.1.

McDaniel, P., McLaughlin, S., 2009. Security and privacy challenges in the smart grid. IEEE Secur. Priv. 7 (3), 75–77.

Medjroubi, W., Matke, C., Kleinhans, D., 2014. SciGIRD: an Open Source Reference Model of European Transmission Networks for Scientific Analysis, online available at: http://scigrid.de/ (last access 30.09.16).

Milano, F., 2010. Power System Modelling and Scripting. Springer-Verlag, London.

Mondzech, J., Sester, M., 2011. Quality analysis of OpenStreetMap data based on application needs. Cartographica 46 (2), 115–125.

Mooney, P., Corcoran, P., Winstanley, A., 2010. Towards quality metrics for OpenStreetMap. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, pp. 514–517.

Neukirch, M., 2014. Konflikte um den Ausbau der Stromnetze: Status und Entwicklung heterogener Protestkonstellationen, Tech. rep.. Stuttgarter Beiträge zur Organisations-und Innovationsforschung, SOI Discussion Paper.

NEXT ENERGY, 2016. EWE Research Centre for Energy Technology, Research Project open_eGo, available at: https://www.next-energy.de/.

Nikoletatos, J., Tselepis, S., 2015. Renewable Energy Integration in Power Grids: Technology Brief, Tech. rep.. International Renewable Energy Association.

Nygard, K., Ghosn, S., Chowdhury, M., Loegering, D., McCulloch, R., Ranganathan, P., 2011. Optimization models for energy reallocation in a smart grid. In: Computer Communications Workshops, INFOCOM WKSHPS, 2011 IEEE Conference, pp. 186–190.

Oeding, D., Rüdiger, O.B., 2004. Elektrische Kraftwerke und Netze, Vol. 6. Springer.

Open Power System Data (OPSD), 2015. A Free and Open Data Platform for Power System Modelling, available at: http://open-power-system-data.org/.

OSM Wiki, 2014. Proposed features/Power routing proposal, available at: http://wiki.openstreetmap.org/wiki/.

Pfluger, B., 2014. Assessment of Least-Cost Pathways for Decarbonising Europe's Power Supply: a Model-Based Long-Term Scenario Analysis Accounting for the Characteristics of Renewable Energies. KIT Scientific Publishing.

PostgreSQL Global Development Group, 1996. PostgreSQL, available at http://www.postgresql.org.

Ramm, F., Topf, J., Chilton, S., 2011. OpenStreetMap: Using and Enhancing the Free Map of the World. UIT Cambridge, Cambridge.

Rendel, T., 2015. Erweiterung und Plausibilisierung eines Modells für die integrierte Simulation des europäischen Verbundnetzes und Strommarktes (Ph. d. thesis), Gottfried Wilhelm Leibniz Universität Hannover.

RESTORE2050, 2014. Regenerative stromversorgung und speicher bedarf in 2050, Tech. rep. RWTH Aachen.

Roland Berger Strategy Consultants, 2015. Study Regarding Grid Infrastructure Development: European Strategy for Raising Public Acceptance, Tech. rep.. European Commission, available at: https://ec.europa.eu/.

RWTH Aachen University, , 2014. Speicher- und Netzausbaubedarf bei Betrachtung des gesamten Europäischen Raums mit hohem Anteil Erneuerbarer Energien, Tech. rep.. ISEA, IAEW.

Salmeron, J., Wood, K., Baldick, R., 2004. Analysis of electric grid security under terrorist threat. IEEE Trans. Power Syst. 2 (19).

Schaffner, C., Mihalic, R., 2005. Economic evaluation of controllable devices in the Slovenian electric power system–a case study. In: Proc. of 2005 Power Systems Computation Conference, PSCC'05, Liege, Belgium.

Schavemaker, P., Sluis, L.V.D., 2008. Electrical Power System Essentials. John Wiley & Sons.

Sims, R., Mercado, P., Krewitt, W., Bhuyan, G., Flynn, D., Holttinen, H., Jannuzzi, G., Khennas, S., Liu, Y., OMalley, M., Nilsson, L.J., Ogden, J., Ogimoto, K., Outhred, H., Ulleberg, ., van Hull, F., 2011. Integration of Renewable Energy into Present and Future Energy Systems, Tech. rep.. IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation.

Tande, J., Korps, M., Warland, L., Uhlen, K., Hulle, F.V., 2008. Impact of tradewind offshore wind power capacity scenarios on power flows in the European HV network. In: Seventh International Workshop on Large-Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms, Madrid, Spain, May.

Tang, H., McMillin, B., 2008. Security of information flow in the electric power grid. In: Critical Infrastructure Protection. Springer, pp. 43–56.

Ventosa, M., Baíllo, A., Ramos, A., Rivier, M., 2005. Electricity market modeling trends. Energy Policy 33 (7), 897–913.

Villasana, R., Garver, L., Salon, S., 1985. Transmission network planning using linear programming. IEEE Trans. Power Appar. Syst. 2, 349–356.

Wiese, F., 2015. renpass: Renewable Energy Pathways Simulation System, Open Source as an Approach to Meet Challenges in Energy Modeling (Ph.D. thesis), University of Flensburg.

Wiese, F., Bökenkamp, G., Wingenbach, C., Hohmeyer, O., 2014. An open source energy system simulation model as an instrument for public participation in the development of strategies for a sustainable future. Wiley Interdiscip. Rev.: Energy Environ. 3 (5), 490–504.

Zhou, Q., 2003. Cross-border Congestion Management in the Electricity Market (Ph.D. thesis), Durham University.

Zielstra, D., Zipf, A., 2010. Quantitative studies on the data quality of OpenStreetMap in Germany. In: Proceedings of the Sixth International Conference on Geographic Information Science, GIScience, Zurich, Switzerland, pp. 20–26.

Zimmerman, R.D., Murillo-Sanchez, C.E., Thomas, R.J., 2011. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE Trans. Power Syst. 26 (1), 12–19.