The German Electricity System in 2030 - Data on Consumption, Generation, and the Grid

Frederik vom Scheidt
Karlsruhe Institute of Technology
frederik.scheidt@kit.edu

Carmen Müller
Karlsruhe Institute of Technology
carmen.mueller@student.kit.edu

Philipp Staudt
Karlsruhe Institute of Technology
philipp.staudt@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology
christof.weinhardt@kit.edu

Abstract

This documentation describes a data set for the German electricity system in 2030 and the methodology behind certain figures. The data set includes 1.) topological data on the transmission grid, the generation infrastructure and the demand, 2.) a time series of electricity demand in 2030 in an hourly resolution, 3.) a time series of renewable generation in 2030 in an hourly resolution, and 4.) dispatchable generation capacity and costs in 2030. All data is spatially resolved, at the level of 485 nodes.

1. Introduction

Long-term planning of electricity systems requires adequate data. This data set includes estimated data for the 2030 electricity consumption, generation, and transmission capacity in Germany. The procedure of data acquisition and preprocessing is documented in detail in the following chapters.

2. Electricity grid

We build our estimation of the electricity grid in 2030 in two steps. First, we acquire data on the existing electricity grid from SciGRID [1]. The SciGRID data set is based on the ”power” relations data available in OpenStreetMap and was built by DLR Institute for Networked Energy Systems in a project funded by the federal government of Germany. Second, we consult the confirmed version of the Electricity Network Development Plan of the Federal Network Agency [2] which covers planned measures for expanding the grid until 2030. These measures include both the capacity increase of existing lines, as well as the installation of new lines and nodes. It should be noted that the existing grid data from [1] does not completely match with the ”2018 starting grid” of the Network Development Plan. For this reason, multiple lines from the Network Development Plan ”starting grid” are manually added to the 2018 grid. To control for any errors, additional grid data was obtained from the Federal Network Agency. Unfortunately, this data set lacks crucial data, e.g. for some nodes no location information is available, neither in the form of coordinates nor in the form of postal codes or place names. This prohibits rigorous double checking of the data. The existing data represents the to the best of our knowledge the intended grid topology in 2030. The final model of the 2030 grid includes 485 nodes and 663 lines.

Next, line capacity is calculated according to [3], i.e. by assigning fixed capacity values in Megawatt (MW) to lines based on their voltage. For several new measures announced in the confirmation of the Network Development Plan, the exact value of expected new voltage or capacity is opaque and needs to be estimated based on the project description. For these cases data validity is limited.

Last, any missing coordinates of nodes are determined using the Google Maps Geocoding API [4].

The list of nodes and the associated coordinates is stored in the file ”nodeCoordinates.csv”. The connections between nodes through lines is stored as an incidence matrix in the file ”incidence.csv”. The grid capacity values of all lines in MW are stored in the file "gridCapa.csv".

3. Electricity consumption

In the next step, the hourly electricity consumption in 2030 for each of the 485 network nodes is estimated. For this purpose, an hourly consumption time series is derived and spatially decomposed.

3.1. Hourly demand on national level

The European Network of Transmission System Operators for Electricity (ENTSO-E) offers national forecasts for hourly demand in 2030 [5]. Three different scenarios exist. For this data set, we chose the scenario "EUCO30", which assumes that the European Council's climate and energy targets for 2030 will be achieved. "EUCO30" includes three different variations of hourly 2030 time series, with each variation representing the climatic conditions of a different year: 1982, 1984 and 2007. We average these three forecasts as shown in Equation 1.

$$Consumption_{h,2030} = \frac{1}{3} * (Consumption_{h,2030,1972} + Consumption_{h,2030,1974} + Consumption_{h,2030,2007}) \quad (1)$$

The total electricity consumption in the ENTSO-E's EUCO30 scenario (577 TWh) differs from the Network Development Plan scenario (543,9 TWh) [2]. To achieve consistency with the grid and generation data for 2030, we therefore refactor the hourly generation as show in Equation 2.

$$Consumption_{h,2030,refactored} = Consumption_{h,2030} * \frac{543,9}{577} \quad (2)$$

The final, refactored demand time series in MWh is stored in the file "demandHourly.csv".

3.2. Spatial distribution of demand

Next, the hourly demand values are spatially distributed across Germany. For this, the load factors of the more than 11,000 municipalities in Germany are determined. This is done according to the top-down approach from [6], who argue that the gross domestic product (GDP) of a county can be used as a proxy for its electricity consumption.

We acquire a data set of all politically independent municipalities with associated postal code, county and the number of inhabitants from the Federal Statistical Office [7]. In addition, we acquire 2030 forecasts of the GDP at the district level from [8]. To break down district forecasts to the municipal level (m), we assume that a municipality's GDP share corresponds to its population share in its respective district (d) (Equation 3).

$$GDP_m = GDP_d * \frac{Inhabitants_m}{Inhabitants_d} \forall m \in d \quad (3)$$

Now the load factor and thus the share of each municipality can be calculated:

$$LoadFactor_m = \frac{GDP_m}{\sum_{m \in d} GDP} \quad (4)$$

Next, the load factors of the municipalities are assigned to the grid node that is geographically closest to them. The distances between the coordinates of the network nodes and the coordinates of the municipalities are calculated with the Haversine formula. If more than one municipality is assigned to a node, the load factors of the assigned municipalities are summed up. Equation 5 displays this. NN_node refers to the set of municipalities for which a given node is the nearest node.

$$LoadFactor_{node,2030} = \sum_{m \in NN_{node}} LoadFactor_{m,2030} \quad (5)$$

The load factor of each node is stored in the file "loadShare.csv".

4. Electricity generation

On the generation side, our methodology distinguishes between non-dispatchable renewable generation (i.e., solar PV, onshore wind, and offshore wind) and dispatchable generation.

4.1. Electricity generation from renewable energy sources

Similar to demand data, renewable electricity generation data with temporal and spatial granularity is needed for system analysis.

4.1.1. Hourly electricity generation from renewable energy sources

Assuming that the temporal distribution of electricity generation from solar PV and wind power plants will not change substantially, we use historical time series for the 2030 estimates. Data on hourly renewable power generation in 2016, 2017, and 2018 are available from [9]. The hourly values of the three years are averaged. Since 2016 was a leap year, 29.02.2016 was removed from the data before averaging.

$$Generation_{h,avg161718} = \frac{1}{3} * (Generation_{h,2016} + Generation_{h,2017} + Generation_{h,2018}) \tag{6}$$

However, the amount of total electricity generation will change in 2030, compared to 2016, 2017, and 2018. Therefore, we refactor the hourly generation values. This refactorization is done separately for solar PV and wind.

On average, the total power generated by PV systems in 2016, 2017 and 2018 was 35.34 TWh. The network development plan predicts 950 full load hours with a projected capacity of 91.3 GW [10]. This results in a power generation of 86.7 TWh in 2030, according to equation 7.

Total electricity generation from wind turbines in 2016, 2017, and 2018 averaged 108.6 TWh. For onshore wind power generation, 2133 full-load hours are predicted for 2030 at a projected capacity of 81.5 GW[10]. For offshore wind power generation, 4328 full-load hours are expected with a projected capacity of 17 GW [10]. This results in a total electricity generation from wind power of 247.4 TWh in 2030, according to equation 7.

$$Generation_{h,2030,refactored} = Generation_{h,avg161718} * \frac{TotalGeneration_{2030}}{TotalGeneration_{avg161718}} \tag{7}$$

The resulting time series in MWh are stored in the files ”solarHourly.csv” and ”windHourly.csv”, respectively.

4.1.2. Spatial distribution of electricity generation from renewable energy resources

In the next step, the shares of solar and wind generation at each node are determined. For this, we use data from [11], which documents the installed capacity of solar and wind generation capacity, together with a postal code. We calculate the generation node shares separately for 2016, 2017, and 2018, since installed capacity (slightly) changes over the years. For 2018, the shares are calculated as follows. Those plants whose commissioning took place after 2018 are filtered out from the data set. For the remaining active plants we derive the coordinates from the given postal codes. Based on these coordinates we then assign each plant to its closest node using the Haversine formula. For 2016 and 2017, the process is the same.

By adding up the total installed capacity of the photovoltaic systems, the share of installed capacity for each node can be determined for the years 2016, 2017 and 2018:

$$GenerationSharePV_{node,year} = \frac{InstalledCapacityPV_{node,year}}{TotalGenerationPV_{year}} \tag{8}$$

Similarly, the generation factors of the nodes in the three years can be calculated for wind production:

$$GenerationShareWind_{node,year} = \frac{InstalledCapacityWind_{node,year}}{TotalGenerationWind_{year}} \tag{9}$$

Subsequently, average values for the three years 2016, 2017 and 2018 are calculated for the generation shares of PV electricity generation and wind electricity generation, respectively. These averages are used to forecast the spatial

Table 1: Fuel prices and carbon factors of various power plant types

Fuel type	Fuel price [EUR/MWh]	Carbon factor [tCO_2/MWh]
Waste	0	0.329
Lignite	4	0.4
Natural Gas	27.88	0.201
Nuclear	3	0
Oil	36.37	0.28
Hard coal	11.67	0.337
Cogeneration gas	0	0.201
Pumped hydro	35	0
Reservoir hydro	0	0

distribution of PV power generation and wind power generation for the year 2030. This means that we assume that the spatial distribution will not change in 2030, compared to today. Arguably, where a lot of capacity is already installed today, additional capacity is more likely to be added for three different reasons: Firstly, legal, political and social acceptance hurdles have already been overcome in these places. Secondly, existing plants can be replaced (”repowering”) with more powerful ones, once they reach the end of their lifetime. And thirdly, the ambient conditions are favourable for installing the respective renewable technology.

The resulting shares of solar PV and wind are stored in the files ”solarShareNodes.csv” and ”windShareNodes.csv”, respectively.

4.2. Dispatchable electricity generation

4.2.1. Construction of power plant set

A list of the 828 conventional power plants installed in Germany in 2020 is available from [12]. Some of the listed power plants are not located in Germany or will only be commissioned after 2030, these are not considered further. Besides, we exclude power plants with the status *Stillgelegt (Decommissioned)*, *Vorläufig Stillgelegt ohne StA (Temporarily Decommissioned without notice)*, *Endgültig Stillgelegt mit StA (Permanently Decommissioned with notice)*, and *Endgültig Stillgelegt ohne StA (Permanently Decommissioned without notice)*.

For seven natural gas power plants that are in the planning stage, no nominal power is stated. For these power plants, the net nominal power of scenario B2030 is taken from the approval of the scenario framework for the Network Development Plan [13]. For the plants with unknown commissioning date, we assume that they will be commissioned before 2030. In addition, all nuclear power plants are excluded, since Germany is set to phase out nuclear power until 2022. Combined Heat and Power (CHP) plants with the energy source ”Other” are assumed to use ”natural gas”. Other plants whose energy source are ”Other” or ”Other Storage”, are excluded since no clear efficiency can be determined for these power plants.

After preprocessing, 718 of the initial 828 power plants are left in the data set. These plants have a combined total capacity of 70,175 MW.

4.2.2. Calculation of marginal costs of power plants

In addition to the generation capacity and location of the conventional power plants, their marginal costs are determined. These consist of the costs for the energy source, the emission costs and, in the case of hard coal, the transport costs (Equation 10).

$$MarginalCosts = c_{Transport} + c_{Fuel} + c_{Emissions} \tag{10}$$

The transport costs for hard coal are differentiated by state level (NUTS 1), as reported in [14].

The fuel costs are derived from the costs of the input fuel and the efficiency factor [15] of the respective power plant. The estimated input fuel prices for 2030 are displayed in 1. Since fuel prices are subject to changes these estimates should be updated regularly.

$$c_{Fuel} = \frac{p_{InputFuel}}{efficiency} \tag{11}$$

Table 2: Reference years and efficiency factors of various power plant types

Plant type	Reference year	Reference value	Reference value hard coal
Natural gas	2014	0.3883	0.4791
Waste	2014	0.33	0.4791
Pumped hydro	2004	0.75	0.4492
Lignite	2014	0.4336	0.4791
Mineral oil products	2011	0.3971	0.4701

The emission costs depend on the fuel specific CO_2 factors, the plant efficiency and the CO_2 price. According to the climate protection program 2030 of the German Federal Government, the CO_2 prices in 2030 should amount to 60 Euro per ton of CO_2 [16]. Estimated carbon factors are displayed in 1.

$$c_{Emissions} = Factor_{CO_2} * efficiency * PriceCO_2 \tag{12}$$

For power plants that are active today, the efficiency can be derived from [15]. For power plants that are yet to be built or commissioned, we estimate efficiency as follows.

First, the efficiency of hydro power plants in 2030 is assumed to be 90 percent based on [17].

Second, for hard coal power plants without stated efficiency values, we extrapolate the efficiency according to the approach by [18]. The respective regression line is described in Equation 13. Here t_{inb} denotes the year of commissioning of the power plant.

$$\eta_{SK(t_{inb})} = 0.2982 * (t_{inb} - 1950) + 28.81 \tag{13}$$

Third, for the other types of power plants, we assume that efficiency factors will develop similar to hard coal (Equation 14). t_{ref} is a reference year for which the efficiency is known. Table 2 shows the used reference years and values. For hard coal, the reference values are determined with the regression approach outlined above.

$$\eta_{k(t_{inb})} = \eta_{k,ref} * \frac{\eta_{SK(t_{inb})}}{\eta_{SK(t_{ref})}} \tag{14}$$

4.2.3. Allocation of power plants to grid nodes

We assign coordinates to all plants, based on their postal code. Then, the plants are assigned to the closest network nodes via their coordinates. For this purpose, the distance between all coordinates of the power plants and all coordinates of the nodes is determined using the Haversine formula.

CHP plants play a special role: These cannot be assigned to a postal code, but are distributed all over Germany. Their production correlates with the heat demand in Germany [19]. For this reason, we assume that the generation capacity of CHP plants is distributed proportionally to the gross domestic product as described in chapter 3.2. Thus, each node is allocated a respective share of CHP generation capacity.

To reduce computational complexity of subsequent calculations, we classify all plants according to their marginal costs, in steps of 10 €/MWh. This results in 23 cost classes, ranging from 0 €/MWh to 220 €/MWh.

To account for potential demand response measures, curtailment measures and demand side flexibility from, e.g. storage, and to ensure feasibility of the model, we add an additional hypothetical generation capacity of 1,000 MW at each grid node, with marginally higher marginal costs than all other conventional power plants (221 €/MWh).

The resulting generation capacity in MW per node and cost class is stored in the file "convGenData.csv".

5. Conclusions, limitations and future work

In this work, we describe the estimation of data to enable the modelling of the German electricity system in 2030. The data set is created in the course of an analysis of the effects of hydrogen production on redispatch costs [20] and can be used for various modelling projects.

While sources and assumptions are chosen carefully, the data set exhibits several limitations. We want to reiterate the most crucial ones below.

Currently, the data set of SciGRID is not perfectly compatible with the data set provided by the Federal Network Agency. We expect that the accuracy of the grid representation in 2030 can thus be improved through collaboration with the Federal Network Agency or the four Transmission System Operators in Germany.

For the forecast of the load factors of the municipalities in Chapter 3.2 the top-down approach developed by [6] is used, in which the gross domestic product is used as an indicator for the electricity consumption of the municipalities. This is an imperfect proxy that could be improved, e.g. by incorporating information on different industry sectors, and their respective energy intensity in each municipality. For this, tools like the Python Toolkit *disaggregator* [21] can be used.

The renewable power generation estimates are based on the assumption that the spatial distribution of generation will not change until 2030. This assumption can be validated and compared with bottom-up analyses such as [22].

6. Acknowledgements

This work was supported by the German Research Foundation (DFG) as part of the Research Training Group GRK 2153: Energy Status Data - Informatics Methods for its Collection, Analysis and Exploitation.

References

[1] C. Matke, W. Medjroubi, and D. Kleinhans, "SciGRID - An Open Source Reference Model for the European Transmission Network (v0.2)." Retrieved from `http://www.scigrid.de`, 2016.

[2] Bundesnetzagentur, "Bestätigung des Netzentwicklungsplans Strom für das Zieljahr 2030." Retrieved from `https://www.netzentwicklungsplan.de/sites/default/files/paragraphs-files/NEP2019-2030_Bestaetigung.pdf`, 2019.

[3] Kießling F., P. Nefzger, and U. Kaintzyk, *Freileitungen: Planung, Berechnung, Ausführung*. Springer-Verlag, 2011.

[4] Google Maps, "Geocoding API." `https://developers.google.com/maps/documentation/geocoding/overview`, 2020.

[5] European Network of Transmission System Operators for Electricity, "Transparency Platform." Retrieved from `https://tyndp.entsoe.eu/maps-data`, 2018.

[6] M. Robinius and D. Stolten, *Lastmodellierung und -visualisierung mittels Geoinformationssystemen*. Conference paper, 2014.

[7] Statistisches Bundesamt, "Gemeindeverzeichnis." Retrieved from `https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/_inhalt.html;jsessionid=B67512182102023E2B2FA4DC8207E3CF.internet8731`, 2018.

[8] Ifo Institut für Wirtschaftsforschung, "regionalisierten strukturdatenprognose 2030," 2012.

[9] Bundesnetzagentur, "Marktdaten." Retrieved from `https://www.smard.de/home/downloadcenter/download_marktdaten/726#!?downloadAttributes=\%7B\%22selectedCategory\%22:1,\%22selectedSubCategory\%22:8,\%22selectedRegion\%22:\%22DE\%22,\%22from\%22:1546297200000,\%22to\%22:1577833199999,\%22selectedFileType\%22:\%22CSV\%22\%7D`, 2018.

[10] Deutsche Übertragungsnetzbetreiber, "Netzentwicklungsplan Strom 2030, Zweiter Entwurf: Version 2019," 2019.

[11] Deutsche Übertragungsnetzbetreiber, "EEG-Anlagenstammdaten." Retrieved from `https://www.netztransparenz.de/EEG/Anlagenstammdaten`, 2018.

[12] Bundesnetzagentur, "Bnetza-power-plant-list." Retrieved from `https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Versorgungssicherheit/Erzeugungskapazitaeten/Kraftwerksliste/kraftwerksliste-node.html`, 2019.

[13] Deutsche Übertragungsnetzbetreiber, "Genehmigung des Szenariorahmens 2019-2030." Retrieved from `https://www.netzentwicklungsplan.de/sites/default/files/paragraphs-files/Szenariorahmen_2019-2030_Genehmigung.pdf`, 2018.

[14] J. Egerer, C. Gerbaulet, R. Ihlenburg, F. Kunz, B. Reinhard, C. von Hirschhausen, A. Weber, and J. Weibezahn, "Electricity sector data for policy-relevant modeling: Data documentation and applications to the german and european electricity markets," DIW Data Documentation 72, Berlin, 2014.

[15] Open Power System Data, "Data package conventional power plants." Retrieved from `https://doi.org/10.25832/conventional_power_plants/2018-12-20`, 2018.

[16] Bundesregierung, "Klimaschutzprogramm 2030 der Bundesregierung zur Umsetzung des Klimaschutzplans 2050." Retrieved from `https://www.bundesregierung.de/resource/blob/975226/1679914/e01d6bd855f09bf05cf7498e06d0a3ff/2019-10-09-klima-massnahmen-data.pdf?download=1`, 2019.

[17] O. Ellaban, H. Abu-Rub, and F. Blaabjerg, “Renewable energy resources: Current status, future prospects and their enabling technology,” *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 748–764, 2014.

[18] J. Schröter, *Auswirkungen des eurpäischen Emissionshandelsystems auf den Kraftwerkseinsatz in Deutschland*. Diplomarbeiten Agentur, 2004.

[19] Bundesministerium für Wirtschaft und Energie, “Evaluierung der Kraft-Wärme-Kopplung: Analysen zur Entwicklung der Kraft-WärmeKopplung in einem Energiesystem mit hohem Anteil erneuerbarer Energien,” 2019.

[20] F. vom Scheidt, J. Qu, P. Staudt, D. Mallapragada, and C. Weinhardt, “The effects of electricity tariffs on cost-minimal hydrogen supply chains and their impact on electricity prices and redispatch costs,” in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021 (forthcoming).

[21] F. P. Gotzens, P. A. Verwiebe, and M. Held, “Demandregio disaggregator.” Retrieved from `https://github.com/DemandRegioTeam/disaggregator/`, 2020.

[22] M. Robinius, *Strom- und Gasmarktdesign zur Versorgung des deutschen Straßenverkehrs mit Wasserstoff*. Dissertation, RWTH Aachen, 2015.