Modelling the high-voltage grid using open data for Europe and beyond

Bobby Xiong*, Davide Fioriti, Fabian Neumann, Iegor Riepin, Tom Brown

*presenting author: xiong@tu-berlin.de

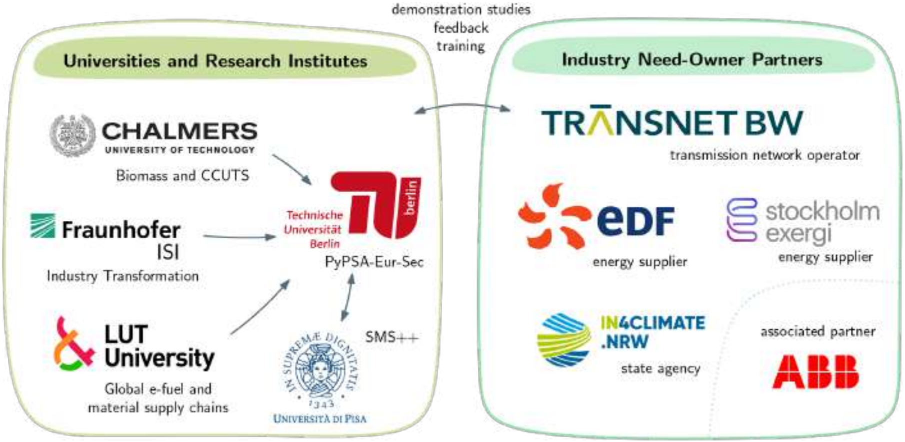
Scan QR code or visit:
https://www.nature.com/articles/s41597-025-04550-7

MITEI Research (Hybrid/Virtual) Seminar, February 17, 2026

Technische Universität Berlin

Supported by: Federal Ministry for Economic Affairs and Energy on the basis of a decision by the German Bundestag

RESILIENT — Project partners

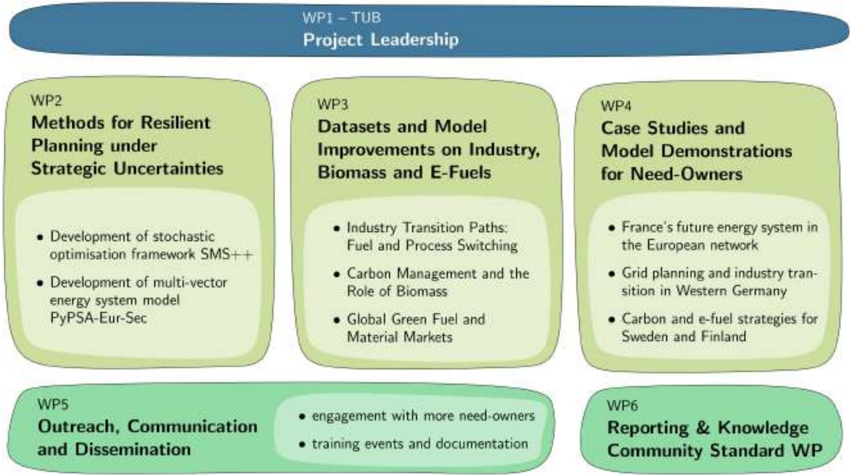


Funded via **CETPartnership 2022** Call — **BMWK** for all German partners.

Source: `https://resilient-project.github.io/`

Technische Universität Berlin

RESILIENT — Work packages



Source: https://resilient-project.github.io/

Motivation

- Conclusion drawn from models only as good as underlying data and assumptions
- With few exceptions, existing datasets are either (i) not open or provide no licensing [4], (ii) not georeferenced [8], or (iii) are not up-to-date or too complex to update [3, 7, 9]
- Data provided on the ENTSO-E online map [4] is highly stylised in terms of line routes and topological connectivity. In some cases (e.g., Madrid, Spain), it contains clearly visible errors.


scientific data

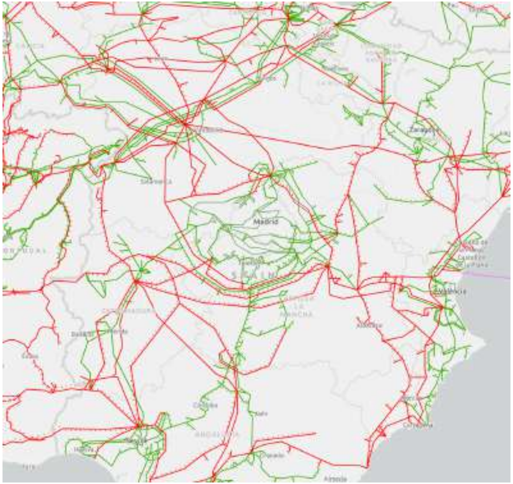
OPEN

DATA DESCRIPTOR

Modelling the high-voltage grid using open data for Europe and beyond

Source: Screenshot taken from https://www.nature.com/articles/s41597-025-04550-7 [12]

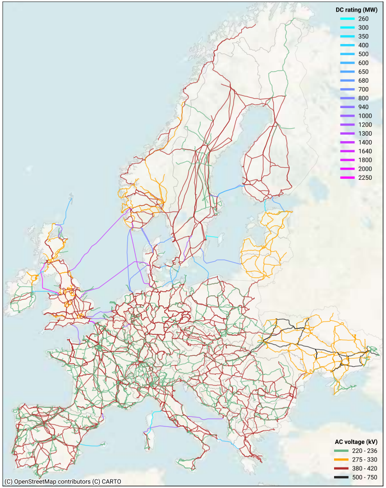
Motivation — Madrid example



Source: Screenshot taken from `https://www.entsoe.eu/data/map/` (left) and own illustration (right), interactive demo: `https://bxio.ng/assets/osm`.

Proposed solution

- Dataset [13] contains a topologically connected representation of the European high-voltage grid (220 kV to 750 kV) constructed using OpenStreetMap data, see OpenInfraMap for visualisation [5]
- Transparent, open-source and reproducible workflow completely implemented in Python/PyPSA-Eur
- By accessing the Overpass turbo API [10], the dataset can be updated anytime to include new AC and DC lines, substations, and transformers
- AC lines mapped using pandapower's standard line type library [11]. Nominal capacity is set to 70 % of the technical capacity to account for n-1 security approximation [2]
- Includes all 39 European HVDC connections with their nominal rating that are commissioned as of February 2026 (updated)



Source: Own illustration based on OSM data extracted using Overpass Turbo API
https://openstreetmap.org

Technische Universität Berlin

Methodology

Process diagram

retrieve_osm_data.py

- overpass-turbo API requests to retrieve power features of OpenStreetMap per country

clean_osm_data.py

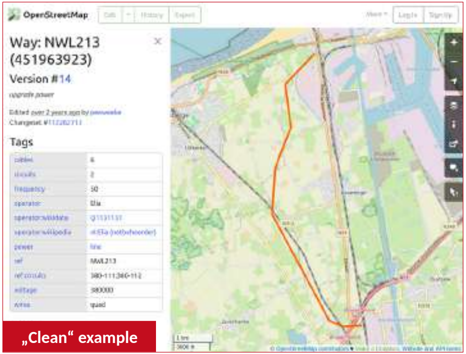
- Heuristic filtering + cleaning process to obtain single buses and lines (voltage, AC/DC, length, and number of circuits)

build_osm_network.py
(from PyPSA-Earth)

- Building the network on lines/links and aggregated buses;
- Constructing transformers + converters

base_network.py
(remaining PyPSA-EUR workflow)

- Creating PyPSA-ready base network base.nc
- TYNDP and NEP projects being added



- Updated workflow (02/2026) includes optional coverage for assets under construction as well as an updated interactive map.
- Prebuilt network v0.7 released on zenodo `https://zenodo.org/records/14144752`, interactive map demo `bxio.ng/assets/osm`

Source: Own illustration based on [12].

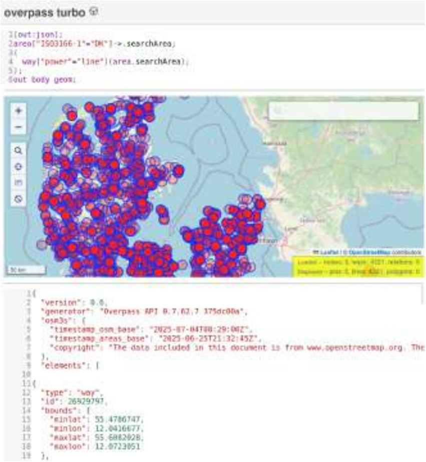
Technische Universität Berlin

Steps — 1. Data retrieval

Grid element	**Overpass turbo queries**
AC lines & cables	`way["power"="line"]`, `way["power"="cable"]`
Power routes & circuits	`relation ["route"="power"]`, `relation ["power"="circuit"]`
Substations	`way["power"="substation"]`, `relation ["power"="substation"]`

Table: Overpass turbo API queries*

*Each element is also queried with `construction:power=*` and `power=construction` variants to capture planned infrastructure.



Source: Screenshot taken from https://overpass-turbo.eu

Technische Universität Berlin

Steps — 2. Data cleaning (before)

line id	**cables**	**circuits**	**frequency** (Hz)	**type**	**voltage** (V)
way/1		double	50	cable	380 kv
relation/1	3		50	cable	380 000 abc
way/3	9	1;2	50.0001	line	380 000 / 220 000
way/4	9	3	50; 50	line	380 000! 220 000
way/5	8		50	line	110 000; 220 000
way/6			50	cable	300 000

Table: Illustrative example of AC lines and cables input data.

Source: Own illustration.

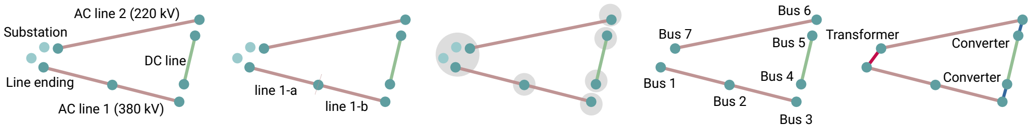
Steps — 2. Data cleaning (after)

line id	circuits	frequency (Hz)	type	voltage (V)
way/1	2	50	cable	380 000
relation/1	1	50	cable	380 000
way/3-1	1	50	line	380 000
way/3-2	2	50	line	220 000
way/4-1	1	50	line	380 000
way/4-2	1	50	line	220 000
~~way/5-1~~	~~1~~	~~50~~	~~line~~	~~110 000~~
way/5-2	1	50	line	220 000
way/6	1	50	cable	300 000

Table: Illustrative example of AC lines and cables after cleaning. Changes highlighted in yellow.

Source: Own illustration.

Steps — 3. Network building: Creating a connected topology



a. Line endings are added to set of buses

b. Lines overpassing nodes are split into segments

c. IDs are assigned to buses within tolerance of 500 m

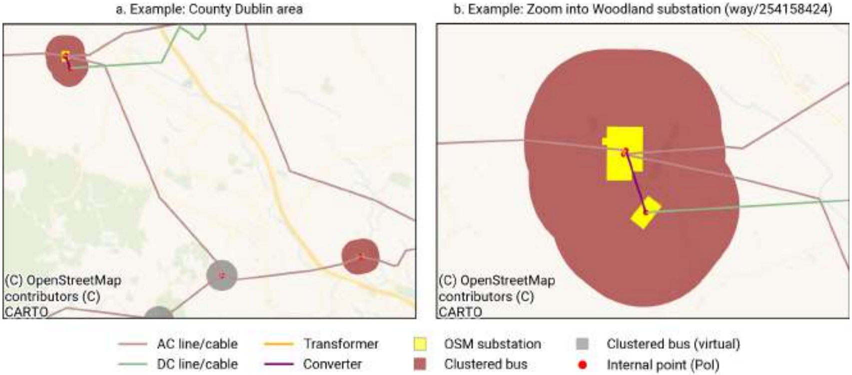
d. Buses are aggregated by voltage and AC/DC type

e. Transformers (red) and converters (blue) are added

Source: Own illustration.

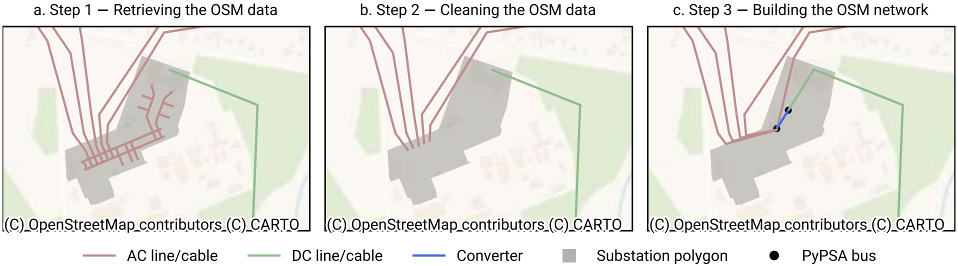
Technische Universität Berlin

Steps — 3. Network building: Example zoom-in



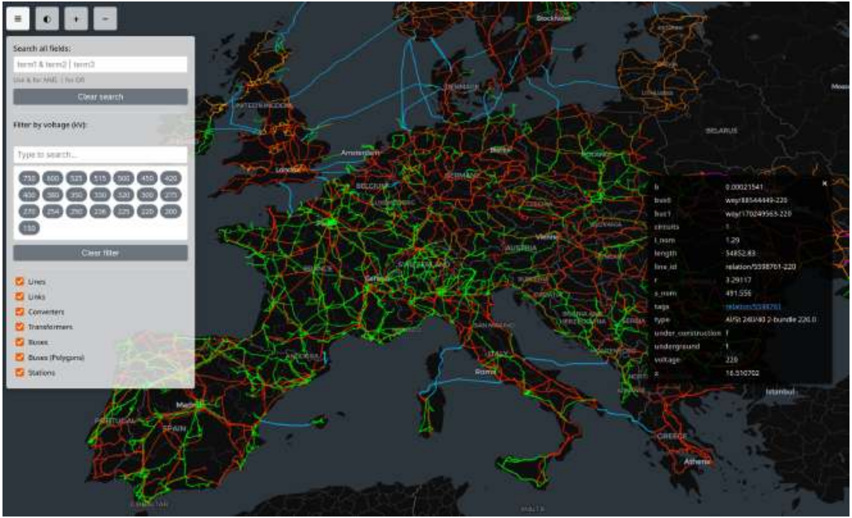
Source: Own illustration.

Example of a completed section



Source: Own illustration.

Interactive demo

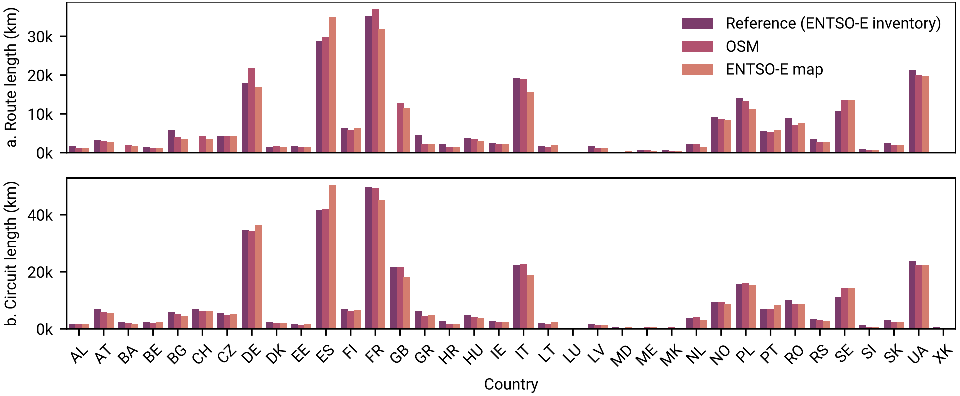


Source: Own illustration. `bxio.ng/assets/osm`

Results & Validation

Technische Universität Berlin

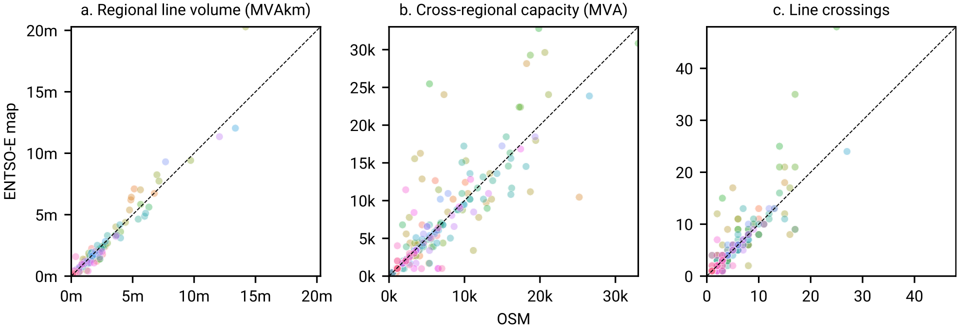
Route and circuit lengths



→ Improvement from ($\rho_{routes} = 0.9489$, $\rho_{circuits} = 0.9862$) to ($\rho_{routes} = 0.9575$, $\rho_{circuits} = 0.9980$)

Source: Own illustration.

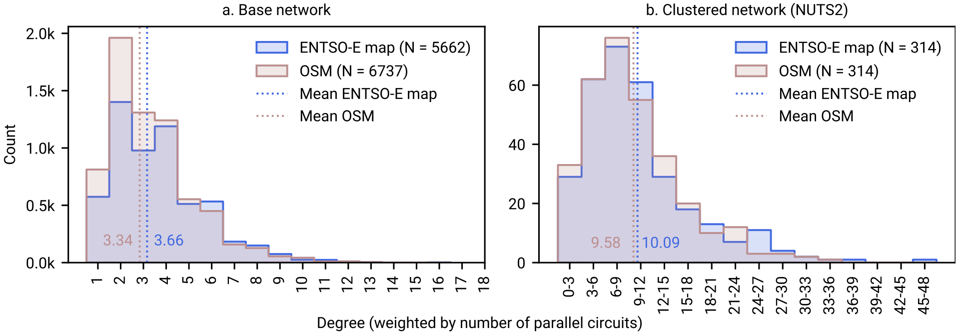
Line metrics based on original PyPSA paper [2]



→ At NUTS1, generally high correlation: $\rho_{MVAkm} = 0.9674$, $\rho_{MVA} = 0.8441$ $\rho_{crossings} = 0.8575$

Source: Own illustration.

Graph-theoretic metrics, e.g. weighted degree distribution

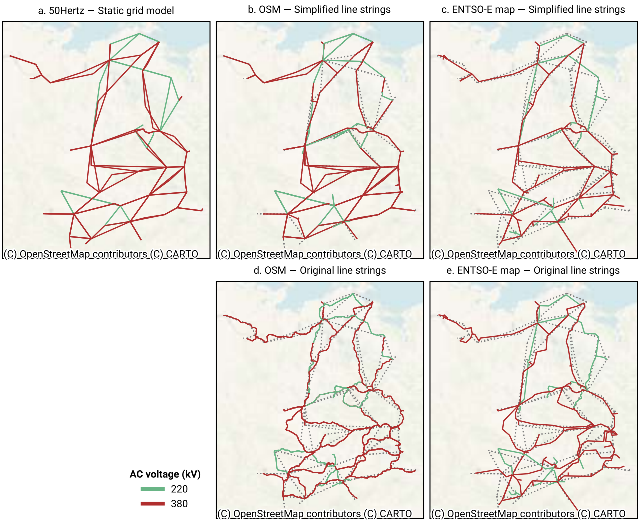


In the base network, number of circuits can be tracked differently, i.e., (1×3) vs. (3×1) vs. $(2 \times 1 + 1)$. Hence, a clustered topology is better suited for comparing the degree distribution, as it captures the total number of circuits between two regions regardless of how they are tracked.

Source: Own illustration.

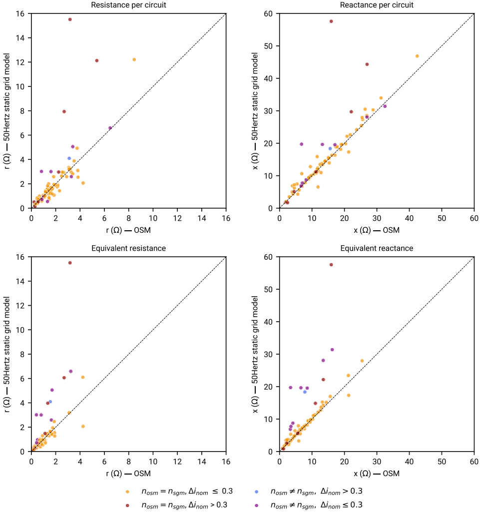
Technische Universität Berlin

50Hertz SGM comparison — Topology



Source: Own illustration.

50Hertz SGM comparison — Electrical parameters



- Scatter plot generated by mapping AC lines and cables of the OSM-based transmission grid to the 50Hertz static grid model (SGM) [1] using OSM tags and SGM names (right join)
- Data explains 4475 km of 5126 km in route length, as not all lines could mapped
- For 79 % of the data, using pandapower's standard line types [11] for calculating the resistance and reactance comes close to official data in the SGM (orange)
- Purple data points show a discrepance primarily due to unequal number of parallel circuits in both datasets (SGM data larger by factor 2). Red and blue data points indicate that underlying line types are entirely different. This is the case for some lines in SGM which are of a newer/stronger 380 kV (allowing higher currents) or weaker 220 kV line type.

Source: Own illustration.

Usage notes

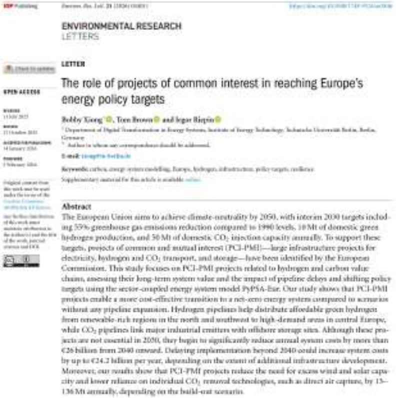
- To reproduce the transmission network, use the example config `config/examples/config.osm-release.yaml` and run `snakemake osm_release -j 4`
- Key settings in the config:
  - `base_network: osm` — builds the network from OSM data
  - `electricity.voltages` — list of voltage levels to include (HV only by default)
- For an experimental **distribution grid** (down to 63 kV), use `config/examples/config.distribution-grid-experimental.yaml` instead, which extends voltages to `[63, 66, 90, 110, 132, 150, 220, ...]`
- Per default, buses within a 500 m radius are merged. This can be changed in the script, though it may affect topological connectedness.
- Networks can be built for specific countries or subsets of PyPSA-Eur countries by adjusting the `countries` list in the config.

Code availability

The code to replicate the entire workflow and dataset is provided as part of PyPSA-Eur and released as free software under the MIT licence. Different licences and terms of use may apply to the underlying input data.

- PyPSA-Eur [6] on GitHub: https://github.com/pypsa/pypsa-eur
- Version 0.7 of the prebuilt network based on OSM data can be retrieved via the Zenodo repository [13]. This link will also point to future updates: https://doi.org/10.5281/zenodo.14144752
- An interactive map is bundled with the dataset on Zenodo[13]. It can also be directly accessed via https://bxio.ng/assets/osm

Other work: Role of EU cross-border CO_2 and H_2 projects



ENVIRONMENTAL RESEARCH
LETTERS

OPEN ACCESS

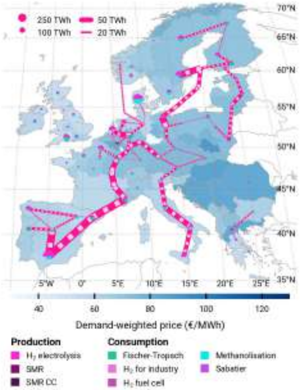
LETTER

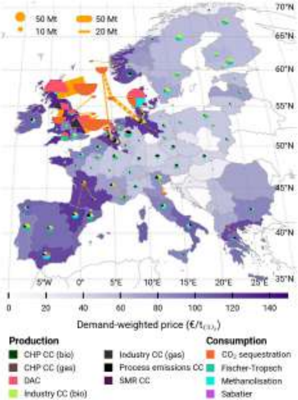
The role of projects of common interest in reaching Europe's energy policy targets

Bobby Xiong, Tom Brown and Iegor Riepin

Abstract

The European Union aims to achieve climate-neutrality by 2050, with interim 2030 targets including 55% greenhouse gas emissions reduction compared to 1990 levels, 10 Mt of domestic green hydrogen production, and 50 Mt of domestic CO_2 injection capacity annually. To support these targets, projects of common and mutual interest (PCI-PMI)—large infrastructure projects for electricity, hydrogen and CO_2 transport, and storage—have been identified by the European Commission. This study focuses on PCI-PMI projects related to hydrogen and carbon value chains, assessing their long-term system value and the impact of pipeline delays and shifting policy targets using the sector-coupled energy system model PyPSA-Eur. Our study shows that PCI-PMI projects enable a more cost-effective transition to a net-zero energy system compared to scenarios without any pipeline expansion. Hydrogen pipelines help distribute affordable green hydrogen from renewable-rich regions in the north and southwest to high-demand areas in central Europe, while CO_2 pipelines link major industrial emitters with offshore storage sites. Although these projects are not essential in 2030, they begin to significantly reduce annual system costs by more than €26 billion from 2040 onward. Delaying implementation beyond 2040 could increase system costs by up to €24.2 billion per year, depending on the extent of additional infrastructure development. Moreover, our results show that PCI-PMI projects reduce the need for excess wind and solar capacity and lower reliance on individual CO_2 removal technologies, such as direct air capture, by 13–136 Mt annually, depending on the build-out scenario.





Source: B. Xiong, T. Brown, I. Riepin (2026), https://iopscience.iop.org/article/10.1088/1748-9326/ae3846

Technische Universität Berlin

References (excerpt)

[1] 50Hertz. *Static Grid Model*. Apr. 2022. (Visited on 07/31/2024).

[2] Thomas Brown, Jonas Hörsch, and David Schlachtberger. "PyPSA: Python for Power System Analysis". In: *Journal of Open Research Software* 6.1 (Jan. 2018). ISSN: 2049-9647. DOI: 10.5334/jors.188. (Visited on 05/30/2024).

[3] Jonas Egerer et al. *Electricity Sector Data for Policy-Relevant Modeling: Data Documentation and Applications to the German and European Electricity Markets*. Research Report 72. DIW Data Documentation, 2014. (Visited on 07/31/2024).

[4] ENTSO-E. *ENTSO-E Transmission System Map*. https://www.entsoe.eu/data/map/. (Visited on 06/28/2024).

[5] Russ Garrett. *Open Infrastructure Map*. https://openinframap.org. 2024. (Visited on 07/08/2024).

[6] Jonas Hörsch et al. "PyPSA-Eur: An Open Optimisation Model of the European Transmission System". In: *Energy Strategy Reviews* 22 (Nov. 2018), pp. 207–215. ISSN: 2211-467X. DOI: 10.1016/j.esr.2018.08.012. (Visited on 03/20/2024).

[7] Neil Hutcheon and Janusz W. Bialek. "Updated and Validated Power Flow Model of the Main Continental European Transmission Network". In: *2013 IEEE Grenoble Conference*. June 2013, pp. 1–5. DOI: 10.1109/PTC.2013.6652178. (Visited on 07/30/2024).

[8] JAO. *Static Grid Model*. https://www.jao.eu/static-grid-model. Sept. 2023. (Visited on 07/30/2024).

[9] Wided Medjroubi et al. "Open Data in Power Grid Modelling: New Approaches Towards Transparent Grid Models". In: *Energy Reports* 3 (Nov. 2017), pp. 14–21. ISSN: 2352-4847. DOI: 10.1016/j.egyr.2016.12.001. (Visited on 06/06/2024).

[10] Martin Raifer. *Overpass Turbo*. https://overpass-turbo.eu/. 2024. (Visited on 07/01/2024).

[11] Leon Thurner et al. "Pandapower—An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems". In: *IEEE Transactions on Power Systems* 33.6 (Nov. 2018), pp. 6510–6521. ISSN: 1558-0679. DOI: 10.1109/TPWRS.2018.2829021. (Visited on 07/05/2024).

[12] Bobby Xiong et al. "Modelling the High-Voltage Grid Using Open Data for Europe and Beyond". In: *Scientific Data* 12.1 (Feb. 2025), p. 277. ISSN: 2052-4463. DOI: 10.1038/s41597-025-04550-7. (Visited on 02/19/2025).

[13] Bobby Xiong et al. *Prebuilt Electricity Network for PyPSA-Eur Based on OpenStreetMap Data. Version 0.6*. Nov. 2024. DOI: 10.5281/zenodo.14144752. (Visited on 11/13/2024).

Thank you!

Results from solving the electricity network model

	System costs (bn. €)	CAPEX (bn. €)	OPEX (bn. €)	Curt. (TWh)	Gen. (TWh)
GridKit	319.28	283.47	35.81	2749.41	3119.82
OSM	318.19	283.09	35.10	2742.26	3118.50
Delta (%)	-0.34	-0.13	-1.99	-0.26	-0.04

Bottlenecks in both model runs are located in the same regions, contributing to an annual curtailment in the range of 2742 TW h to 2149 TW h. More prominent differences in line utilisation are visible in Poland, Sweden, northern France, in the western region of Ukraine, southern and central Spain around the Madrid area, as well as southern Italy.

Setup

- PyPSA-Eur electricity sector only for validation
- Year 2030, hourly resolution, clustered at NUTS2 (318 regions)
- No transmission line expansion allowed
- Generation and storage technologies expansion allowed
- CO_2 price set to 100 €/tCO_2

Source: Own illustration.