Grid data quality

The impact on energy system planning

MapYourGrid Public webinar

8th January 2025

Open Energy Transition
from the creators of PyPSA meets Earth

PyPSA meets Earth

DATACTIVIST

dynartio

MapYourGrid

Introduction and Agenda

Marina Petkova &
François Lacombe

Ekaterina Fedotova &
Emmanuel Bolarinwa

David Diaz &
Carlos Gaete

1. **MapYourGrid**: Build a digital twin of your power grid with OpenStreetMap. (30 minutes talk and 10 minutes Q&A)
2. **Model Impact:** How these improvements can benefit energy system modelling? Insights from PyPSA-Earth Colombia. (30 minutes talk and 10 minutes Q&A)
3. **Power Plant Data:** How to match similar power generation datasets across different sources with PowerPlantMatching and the new OSM Power Plants features. (30 minutes talk and 10 minutes Q&A)
4. **Closing Words** (5 minutes)

You can ask questions during the presentation via the chat function, which will be answered live in the Chat, or you can ask your question at the end of the presentation.

MapYourGrid

Why? Garbage in, garbage out!

1. In which situations does data quality matter? When is the quality good enough for my use-case?
2. How can data quality be assessed and improved?
3. How can we validate the quality of grid data?
4. Which data governance strategies enable collaborative improvements to data quality?
5. How can I improve the quality and where can I find support in doing so?

OpenStreetMap & MapYourGrid

Build a digital twin of your power grid with OpenStreetMap

What is OpenStreetMap ?

And why should it be your best friend ?

OpenStreetMap is a worldwide **open and collaborative database**.

- ➔ Map real-life static visible objects, including power grids.
- ➔ Everyone can contribute.
- ➔ Everyone can use the data, provided attribution.

The 1st open static cartographic database.

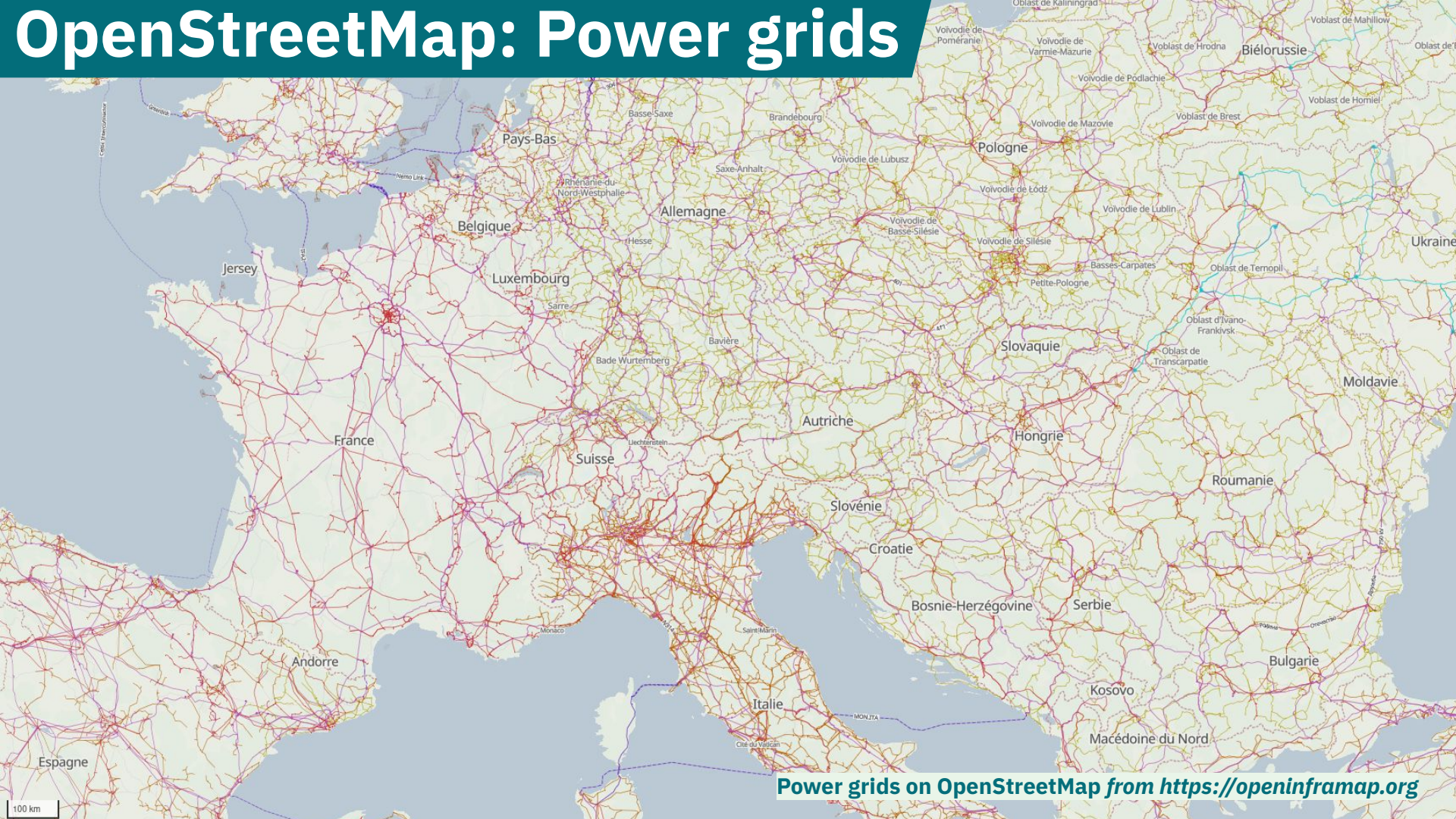
Established in **2004.**

10 050 275 user accounts & **6 000** contributors per day.

MapYourGrid

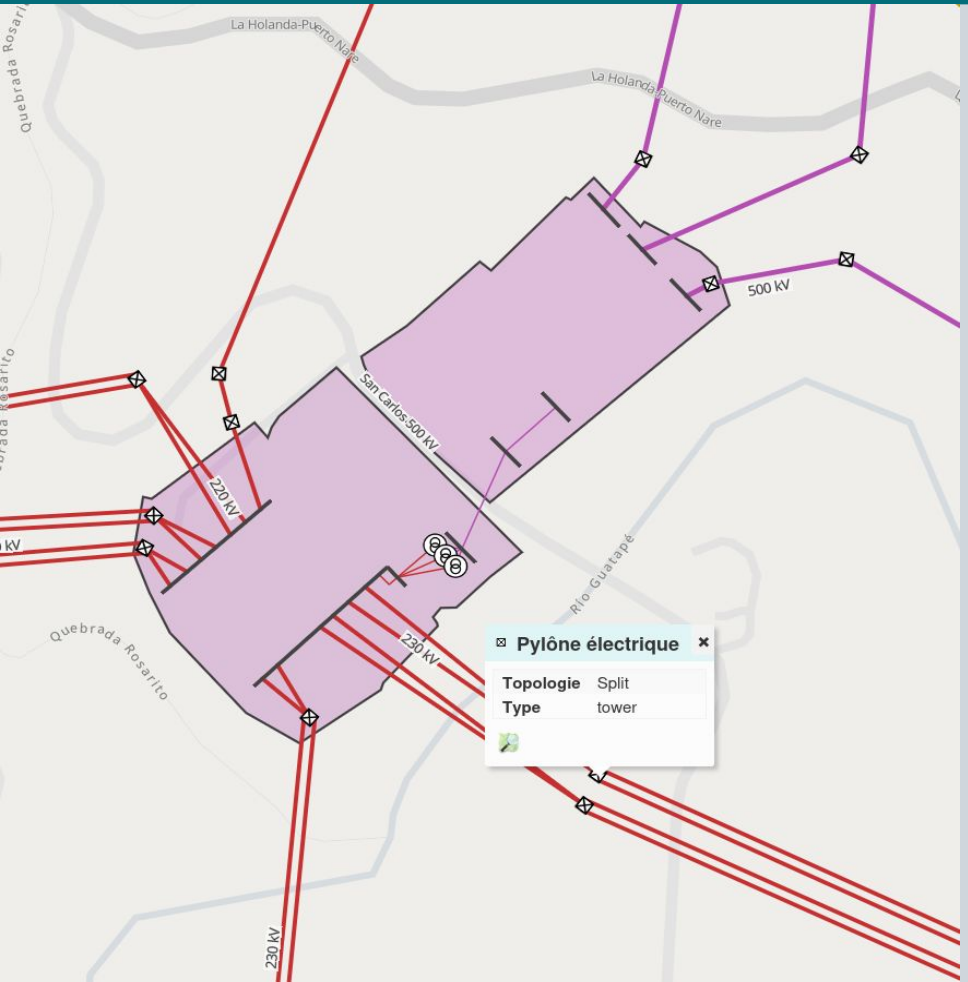
OpenStreetMap: More than a map

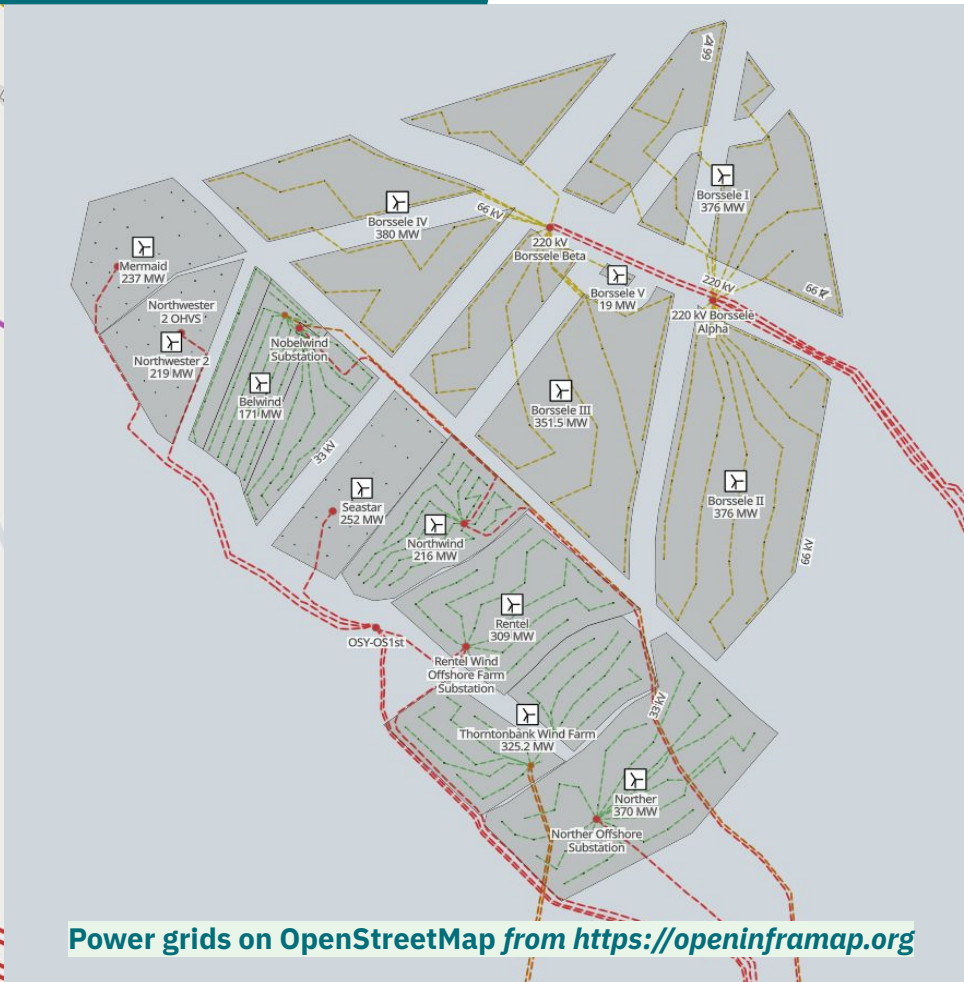
OpenStreetMap: Power grids



Power grids on OpenStreetMap *from https://openinframap.org*

OpenStreetMap: More power grids





Power grids on OpenStreetMap *from https://openinframap.org*

Verifiability and Observation

Lean more online:
https://wiki.openstreetmap.org/wiki/Verifiability

At the core, "verifiability" is that every information you add can be demonstrated to be true or false by other mappers using observable evidence.

Source: OpenStreetMap Wiki

Not all power grid data can be verified in OSM using satellite imagery or on-site surveys.

Can be easily verified on ground or aerial imagery: power tower placement, voltages, the number of cables on a line ...

Needs another suitable source than ground or aerial imagery to be verified: capacity of power plants, sometimes circuits.

Monitoring data quality

About power grids in OpenStreetMap

What is data usability ?

The FAIR principles

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable

YES

BUT

Is usable data always quality data ?

MapYourGrid

Why you need data quality ?

Build a robust digital twin of your grid

Consistency

- ➔ Is the power infrastructure correctly placed on the map ?
- ➔ Are the line management and substation connexions correct ?
- ➔ Are the numbers of circuits and cables consistent across lines ?
- ➔ Are there line segments with different voltages ?

Completeness

There are 75 000 km of power lines in official national data, 85 000 km according to regulatory office, and 100 000 km in OSM.

- ➔ Who is right?
- ➔ Do we meet in the middle?
- ➔ Do the infrastructures match in all datasets?

Up-to-date

- ➔ How many and active are the contributors that maintain the data ?
- ➔ How is feedback taken into account ? How are anomalies reported?
- ➔ Is the feedback verifiable?

MapYourGrid

OpenStreetMap is already fluent at describing power grid digital twins

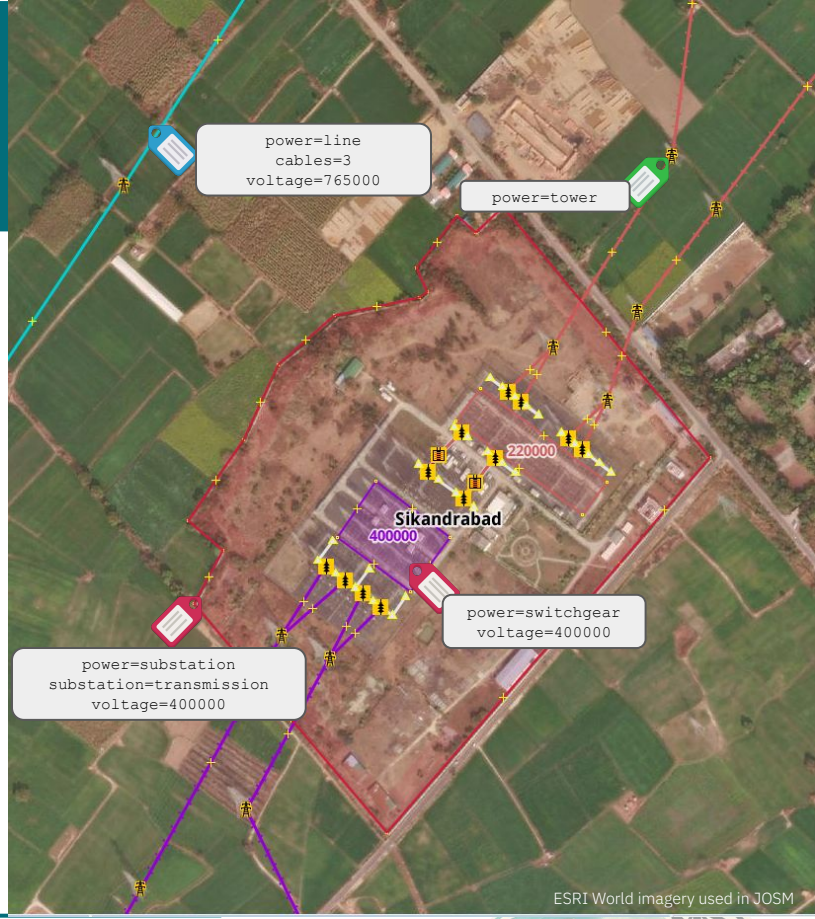
Every OSM object is described by **tags**.

The same tags are **used by all** contributors around the world. This ensures we all **speak the same language**.

Tags can be **refined** by community proposal and voting.

OpenStreetMap tags are a unique worldwide applicable vocabulary for describing power grids elements.

Lean more about **power tagging** in OpenStreetMap:
https://wiki.openstreetmap.org/wiki/Key:power

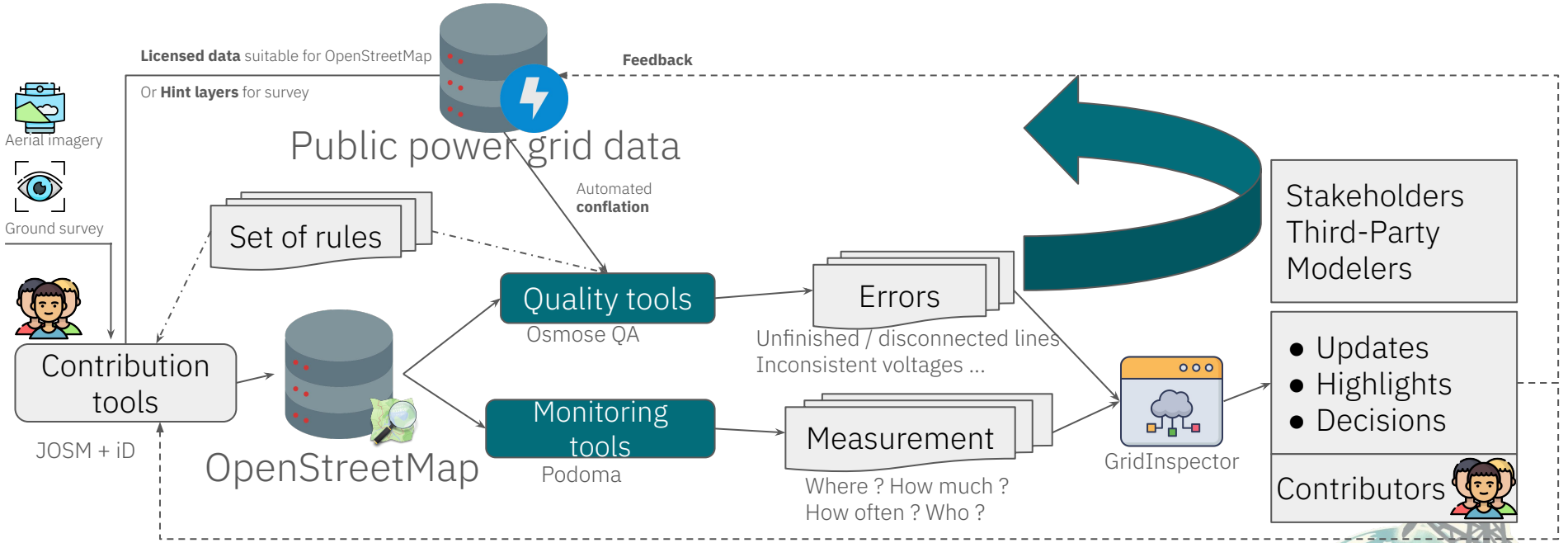


MapYourGrid

Paomedia icons

You too can sustainably build and maintain a digital twin of your static power grid

Step 4: Feedback to contributors and operators



MapYourGrid

Resources to go further

A list of selected references

- ➔ Set of 50+ OSM power rules: https://wiki.openstreetmap.org/wiki/Power_networks/Quality_Assurance
- ➔ Osmose: https://osmose.openstreetmap.fr/en/map/
- ➔ Connectivity maps (e.g. Colombia): https://mapyourgrid.org/countrypages/Colombia/
- ➔ OpenInfrastructureMap: https://openinframap.org
- ➔ GridInspector: https://mapyourgrid.org/gridinspector/ + https://mapyourgrid.dynartio.com/gridinspector/
- ➔ Podoma: https://wiki.openstreetmap.org/wiki/Podoma
- ➔ Progress Dashboard: https://mapyourgrid.infos-reseaux.com/dashboard/
- ➔ Learn grid mapping: https://mapyourgrid.org/starter-kit/
- ➔ Map with JOSM: https://josm.openstreetmap.de/
- ➔ OSM Licensing: https://www.openstreetmap.org/copyright/

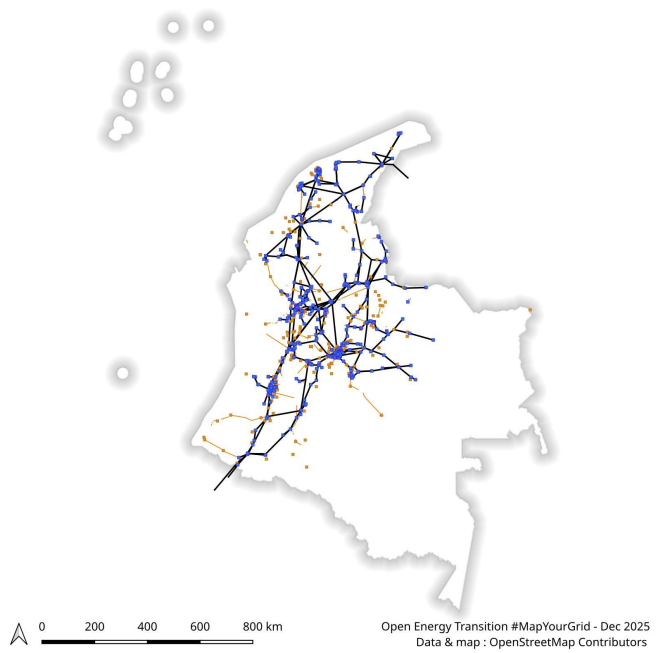
Let’s focus on Colombia

Substation
- Recently mapped
- Existing before 2025

Power line
- Existing before 2025

Modified or added this year
- by #MapYourGrid team
- by other contributors

Recently mapped power tower
- by #MapYourGrid team
- by other contributors



Substation
- Connected
- Not connected

Power line
- Connected
- Not connected

In total, as of 5 Jan. 2026:

2 330 lines segments
712 circuits
323 substations
329 plants

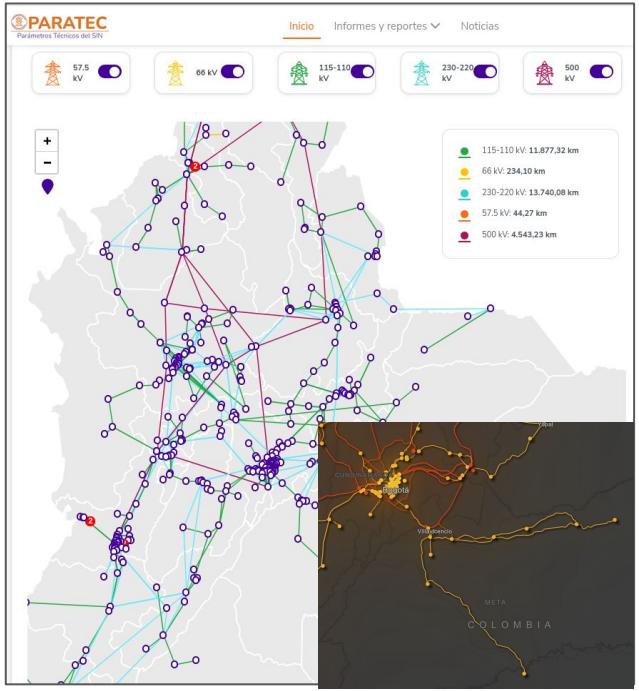
Circuit length (> 50 kV): 28 751 km

Power plant capacity: 15 441 MW

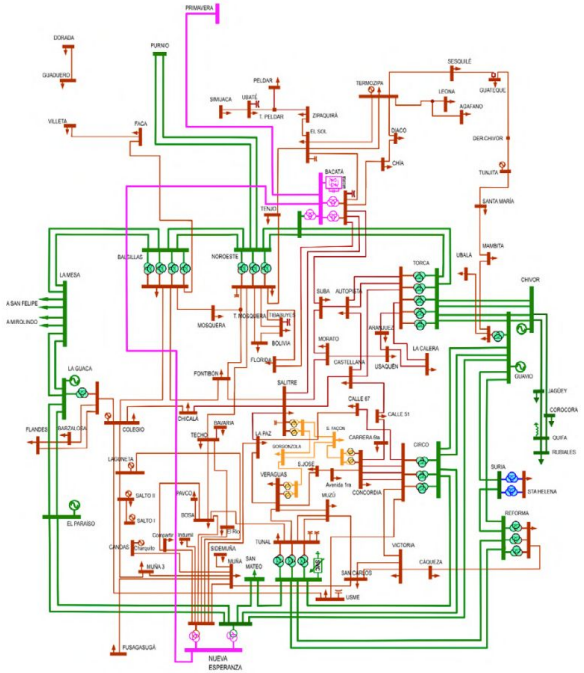
Imagery and observation

Don’t underestimate the importance of combining worldwide and **local imagery**. They can be very useful sources, depending on their updates and resolutions.

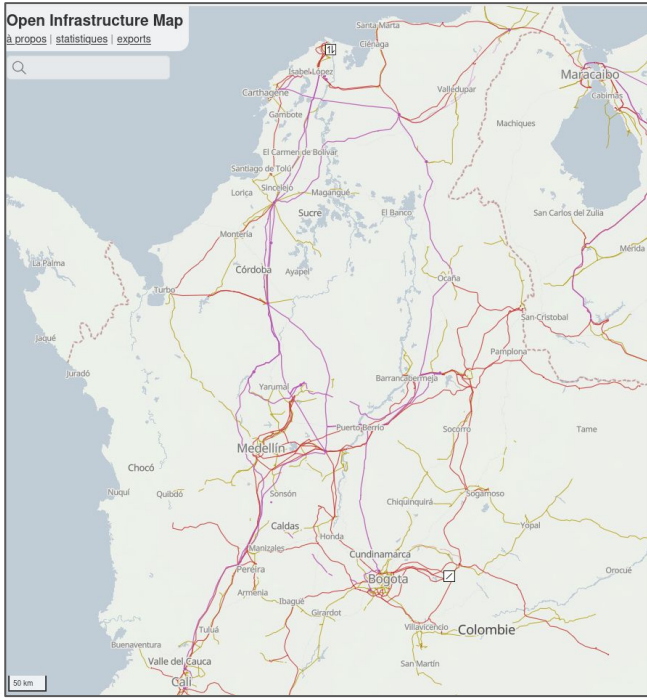
Data Cross-Comparison



XM https://paratec.xm.com.co/mapa
UPME https://www.upme.gov.co/simec/geovisores/



XM & UPME: Plan Maestro for Modernisation and Expansion of Electric Infrastructure, t. 1, p. 156: https://www1.upme.gov.co/siel/Plan_expansin_generacion_transmision/Plan_maestro_modernizacion_Tomo_1.pdf



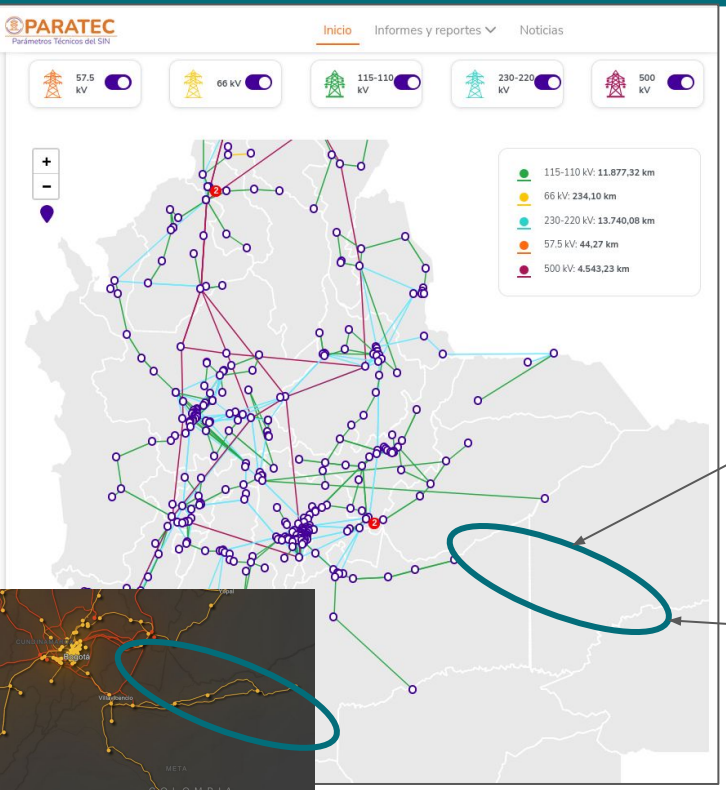
OSM data in OpenInfraMap,
https://openinframap.org/#5.4/3.979/-72.743

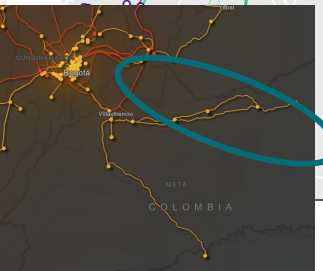
Data Cross-Comparison

What we learnt ?

Sources:

- XM: https://paratec.xm.com.co/mapa
- Plan Maestro for Modernisation and Expansion of Electric Infrastructure, t. 1, p. 156: https://www1.upme.gov.co/siel/Plan_expansin_generacion_transmision/Plan_maestro_modernizacion_Tomo_1.pdf





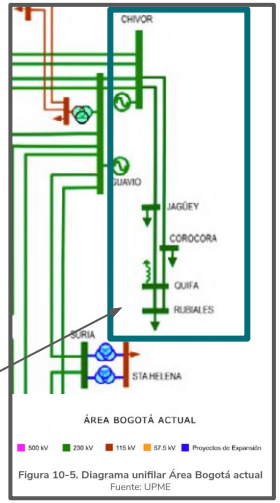
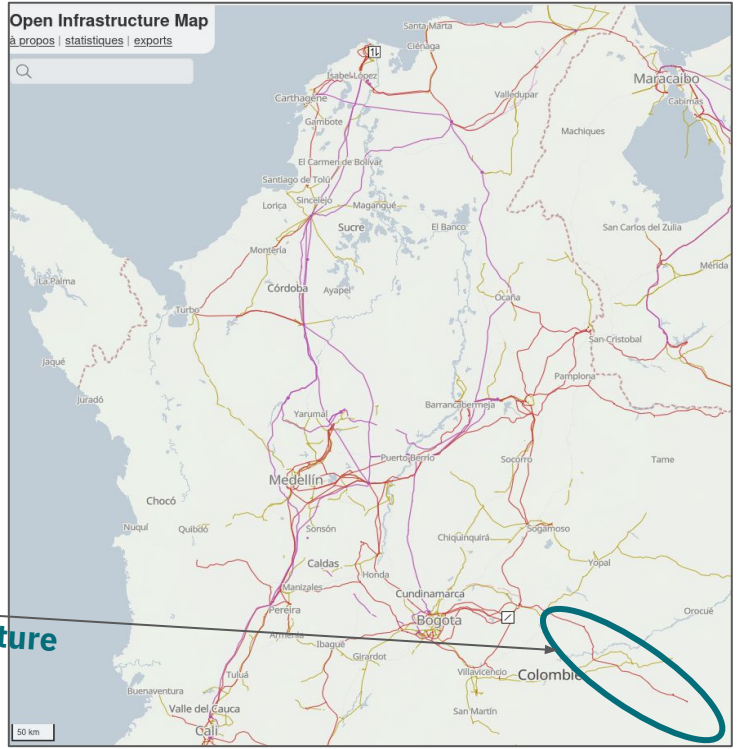


Figura 10-5. Diagrama unifilar Área Bogotá actual
Fuente: UPME

OSM shows some missing 230kV infrastructure



From manual validation to assisted detection

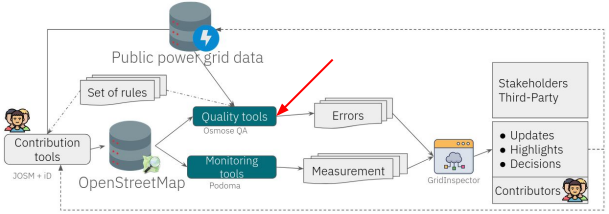
Comparing OSM data with **Earth Observation datasets** like TZ-SAM or Global Renewable Watch.

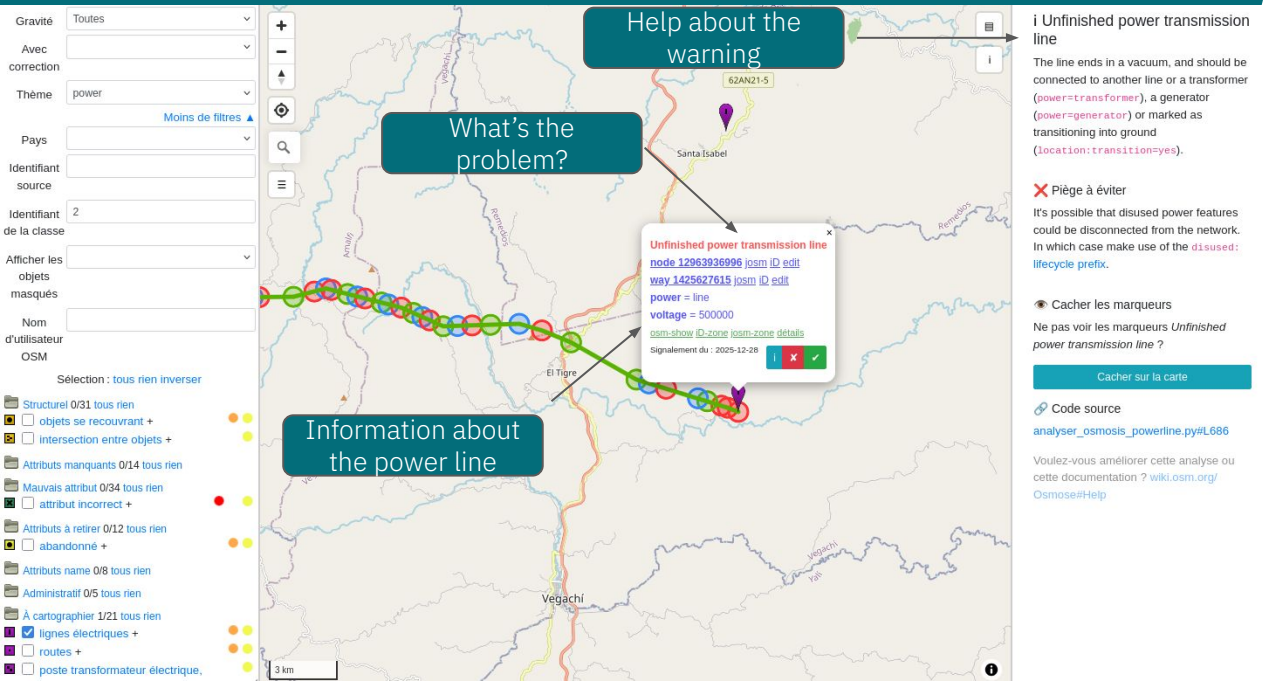
Substation detection: extending and validating existing data in OSM like voltages or location.

Line detection over water bodies: validate accuracy of lines and validate / extend cables and circuits.

A powerful quality assurance tool

Osmose QA https://osmose.openstreetmap.fr





Osmose processes OpenStreetMap features every day in seek of possible inconsistencies according to some rules.

We are using it to challenge our own mapping and to spot the places that need refinement.

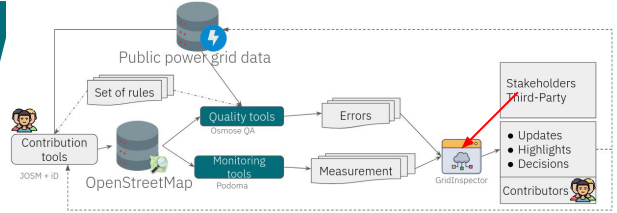
It looks like specific to OpenStreetMap but actually refers to a practice that should be built in any GIS

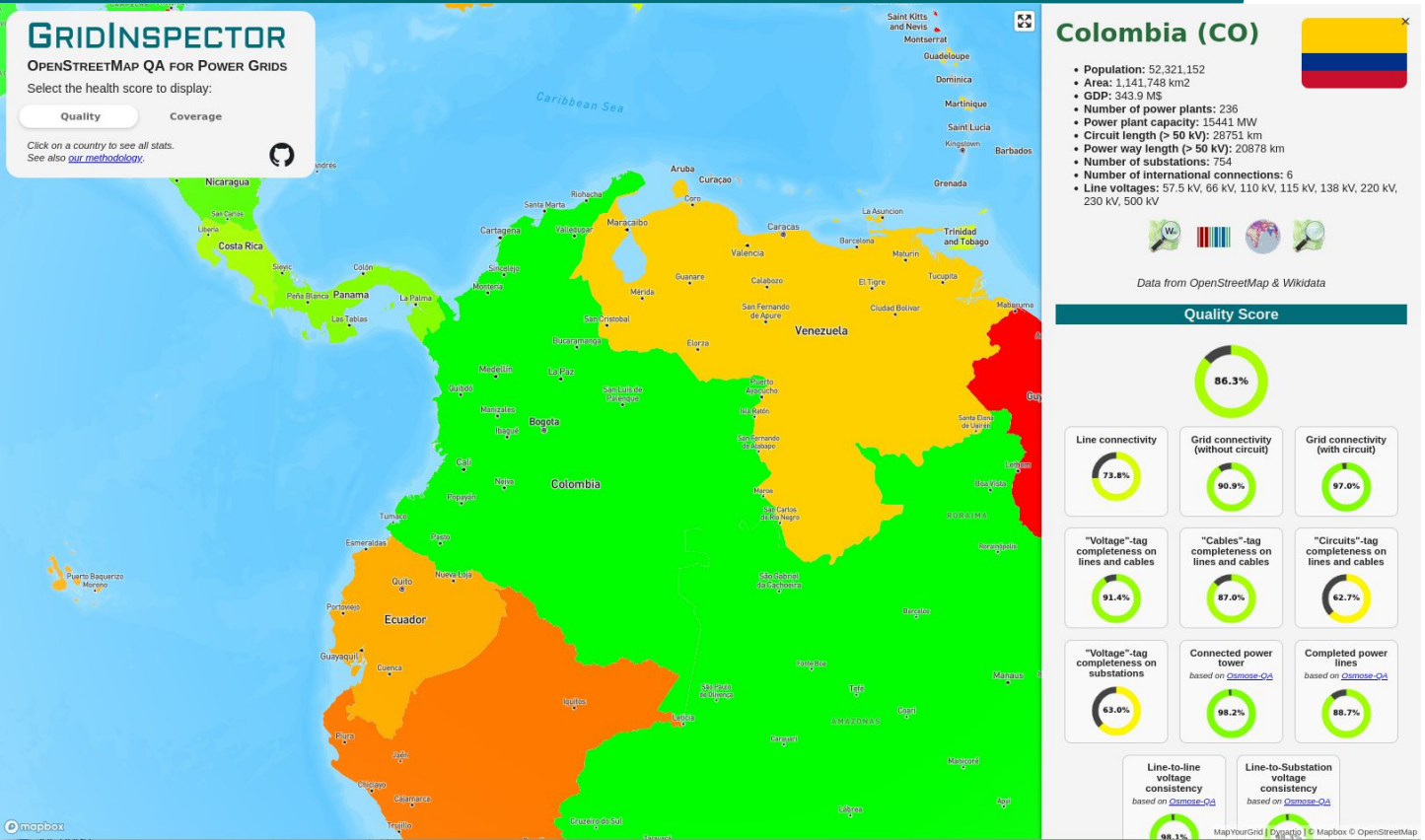
It provides a very rich user interface with contextualised warnings, directions to solve and let also access to data in bulk through an API to load warnings in editors.

MapYourGrid

A comprehensive quality analysis overview

GridInspector - Quality View https://mapyourgrid.org/quality/



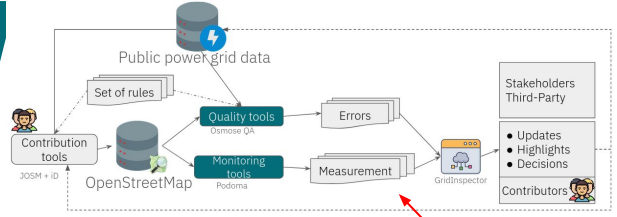


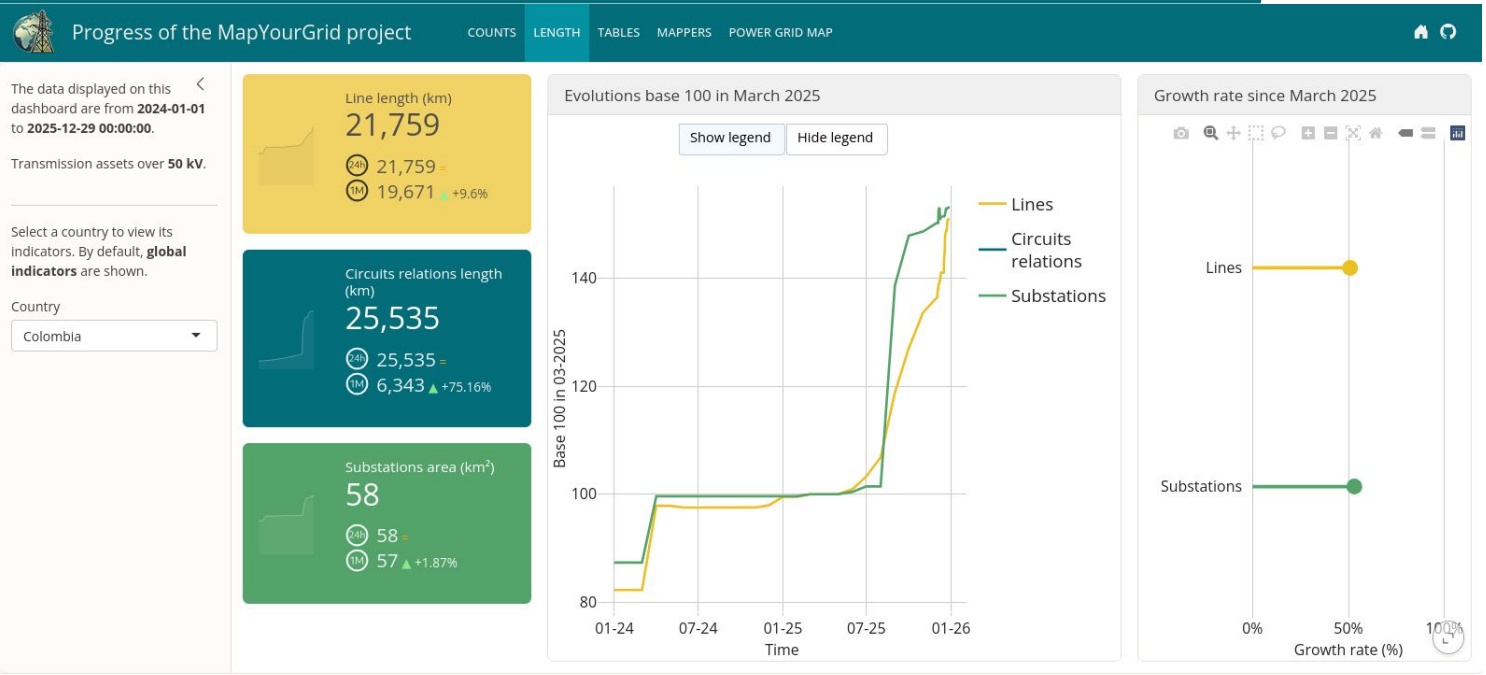
11 rule-based assessment indicators across all countries

An advanced contribution monitoring tool

Podoma
Business specific dashboard

https://mapyourgrid.org/progress/





Contribution monitoring is also a necessary component to establish long term quality.

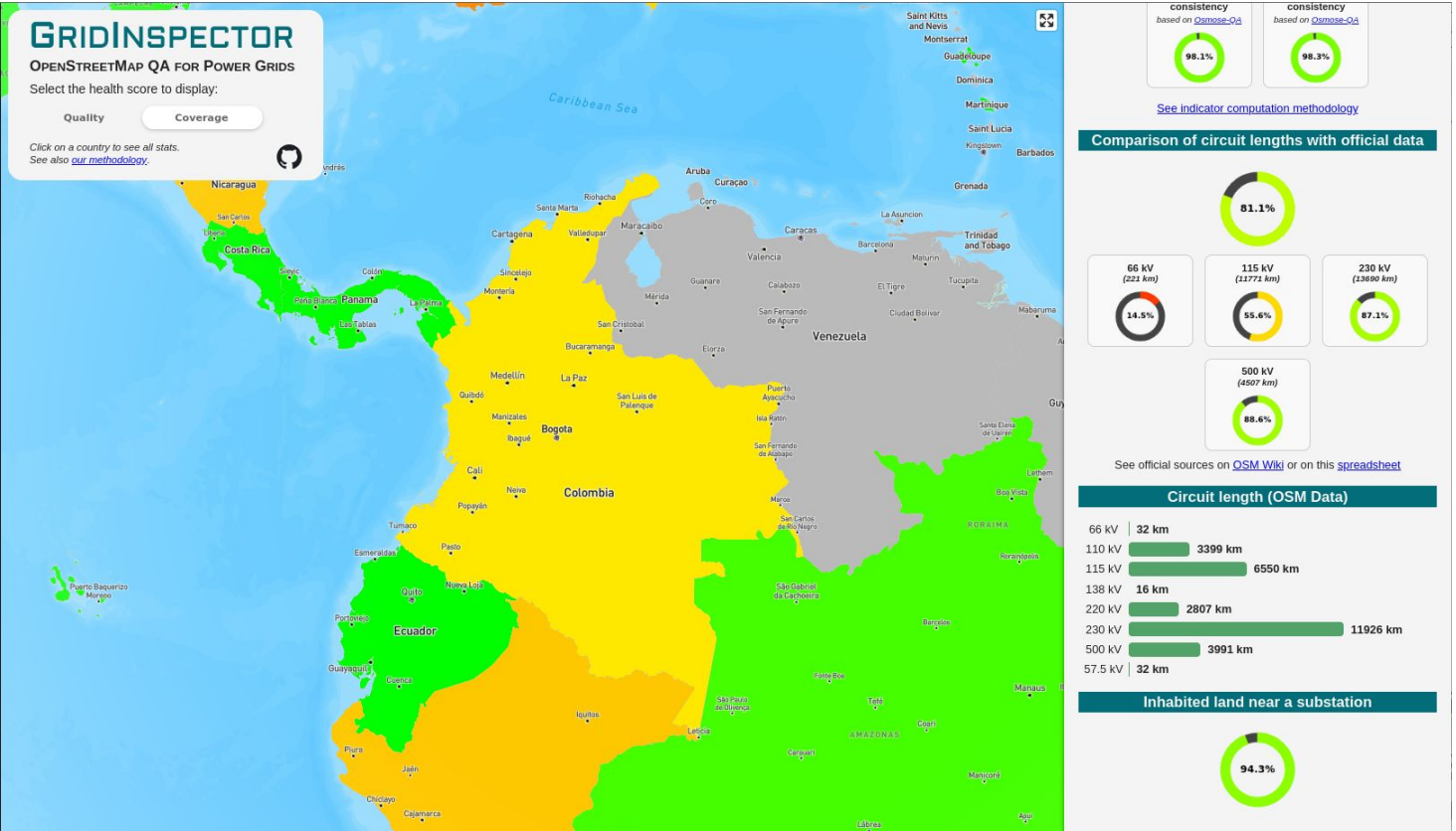
It allows to spot where the contribution and possible large edits occur, on several topics.

It is only a frontend dedicated to power grids to a global tool available in OpenStreetMap community to make such monitoring on any topic possible.

MapYourGrid

A comprehensive grid completeness overview

GridInspector - Coverage View https://mapyourgrid.org/quality/



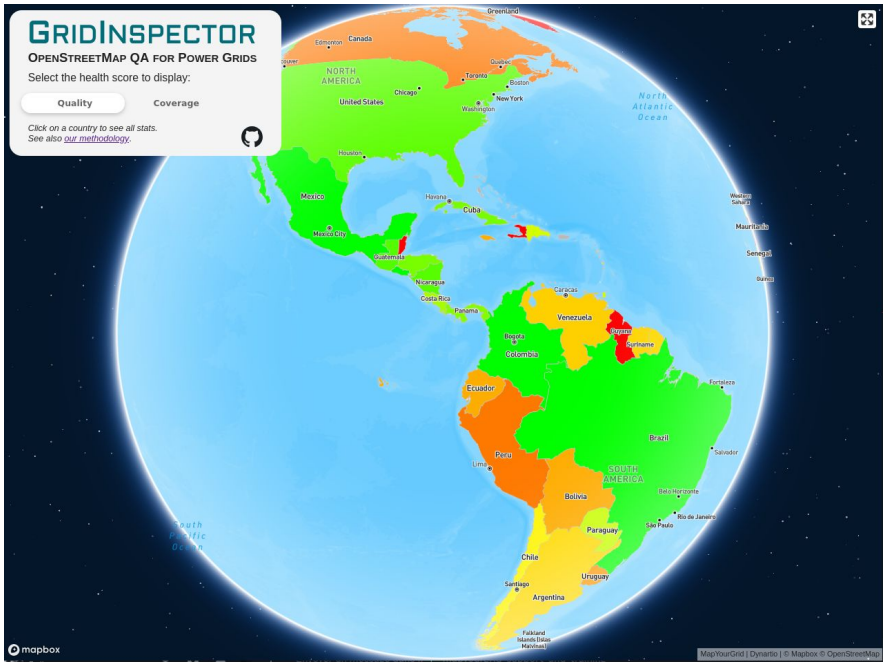
Circuit length comparison between official data and OSM data

And an additional indicator on inhabited land near a substation

Understand your grid at a glance

GridInspector https://mapyourgrid.org/quality/

Colombia's power grid is seen as good quality data in OSM but still needs further contributions or better imagery to be complete.



MapYourGrid

Challenges and opportunities to scale impact

Support what comes next

Quality would improve even more by solving those remaining challenges:

1. **Develop and maintain core tools**
2. **Intensify mapping where data gaps matter most**
3. **Bring the community and institutions together**
4. **Foster alignment with institutional grid data models**
   (e.g. governments, TSO, ENTSOe, CIM)
5. **Improve interoperability with other open databases**
   (e.g. Wikidata)

Let’s tackle them together!

MapYourGrid

Any questions ?

Data Impact Study

Why Energy System Modelling?

- Unsuccessful early PV programme in Bhutan (1980s)
- Shrinking a long-term nuclear power development plan in UK (1980s)
- Canceling plans to build large-scale hydropower plants in Chile: HidroAysén (2008–2012)
- Failure of solar water heating programme in South Africa (2009–2014)
- Repeal of carbon pricing in Australia (2014)

Why Energy System Modelling?

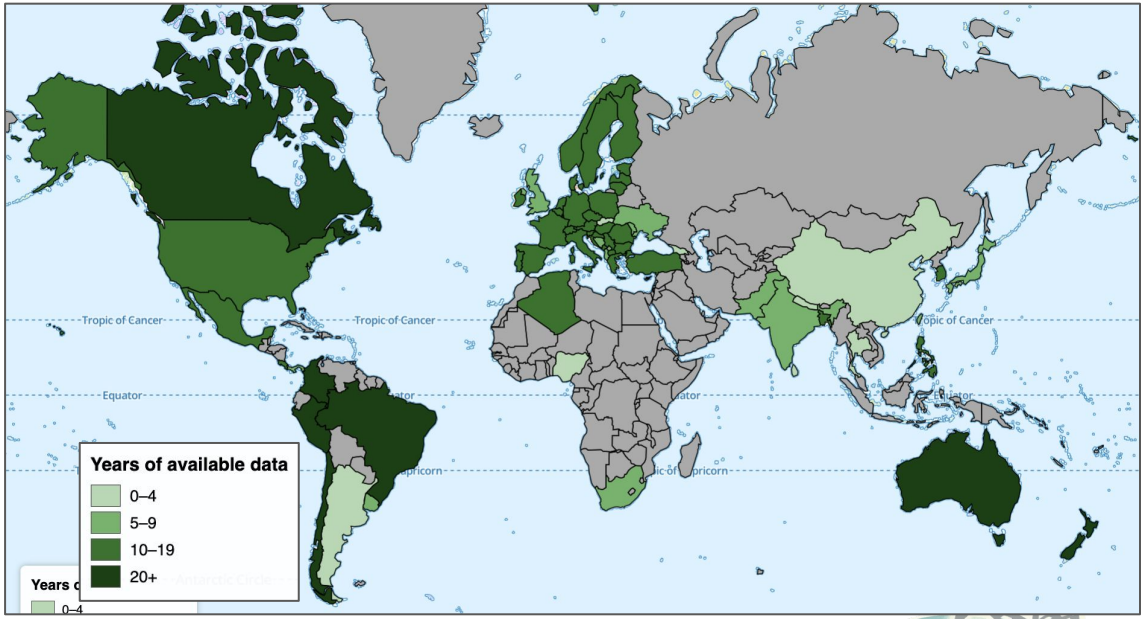
Replace real-world failures with modelling ones

Energy System Models

Rapid development and space for improvement

- Lack of harmonization
- Data availability
- Duplicated modelled efforts
- Uneven geographic coverage

Availability of hourly demand time-series



Source: https://open-energy-transition.github.io/Awesome-Electricity-Demand/

PyPSA-Earth

Independent research and development initiative

https://github.com/pypsa-meets-earth

Motivation
Energy transition knows no borders

Goals
Provide energy planning tools for every part in the world

Applications
academic research, education, policy-relevant studies

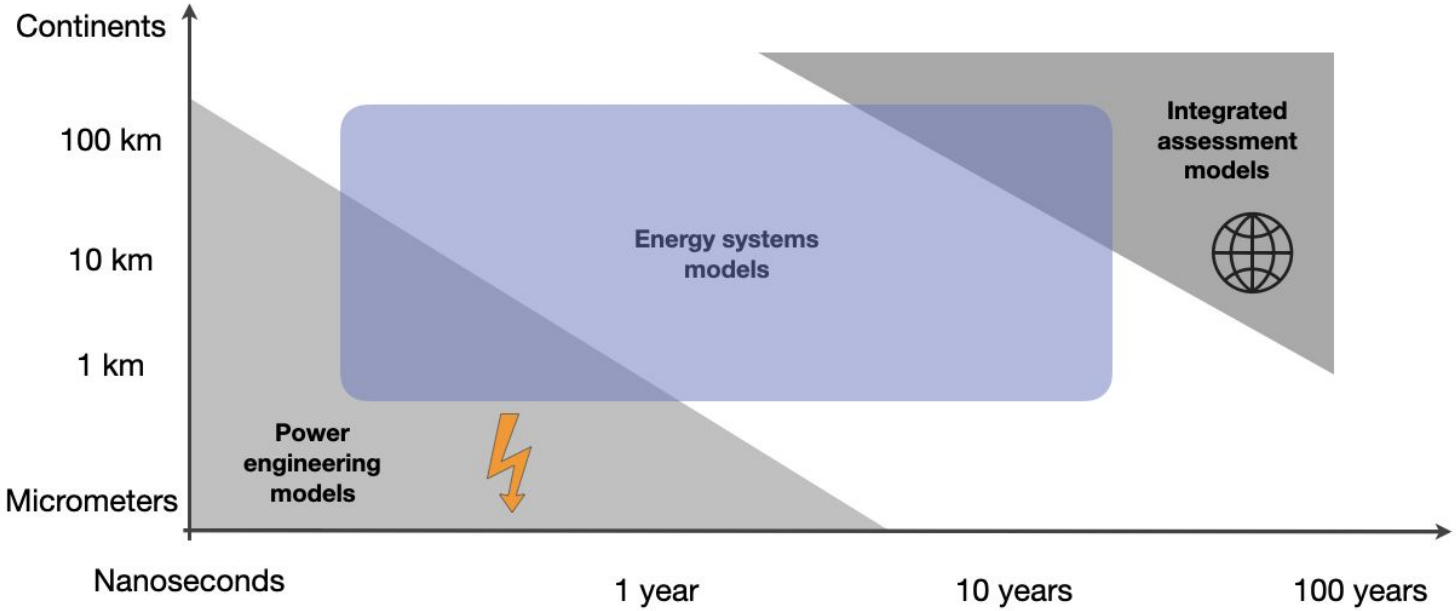
Community
- Suite of open tools integrated with open data
- Transparent governance structure
- User support, open development discussions, outreach, education

PyPSA-Earth

Energy system model to model any country on the Earth

PyPSA meets Earth

https://github.com/pypsa-meets-earth



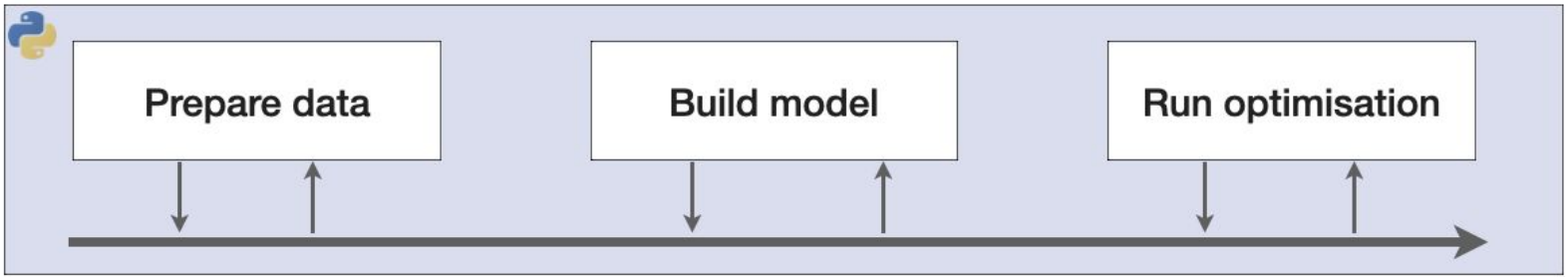
MapYourGrid

Using OSM in modelling

PyPSA-Earth approach

PyPSA-Earth architecture

Going beyond “just one more model”



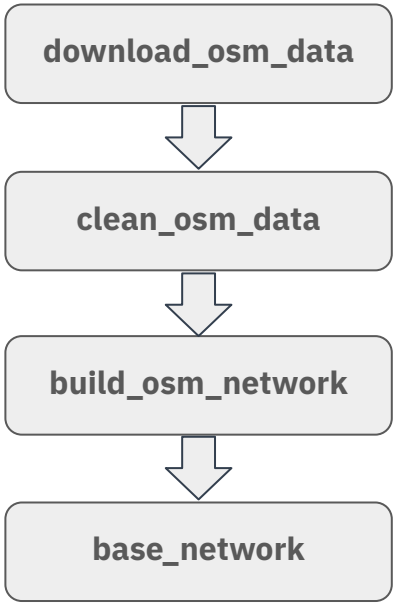
Dependencies management: conda

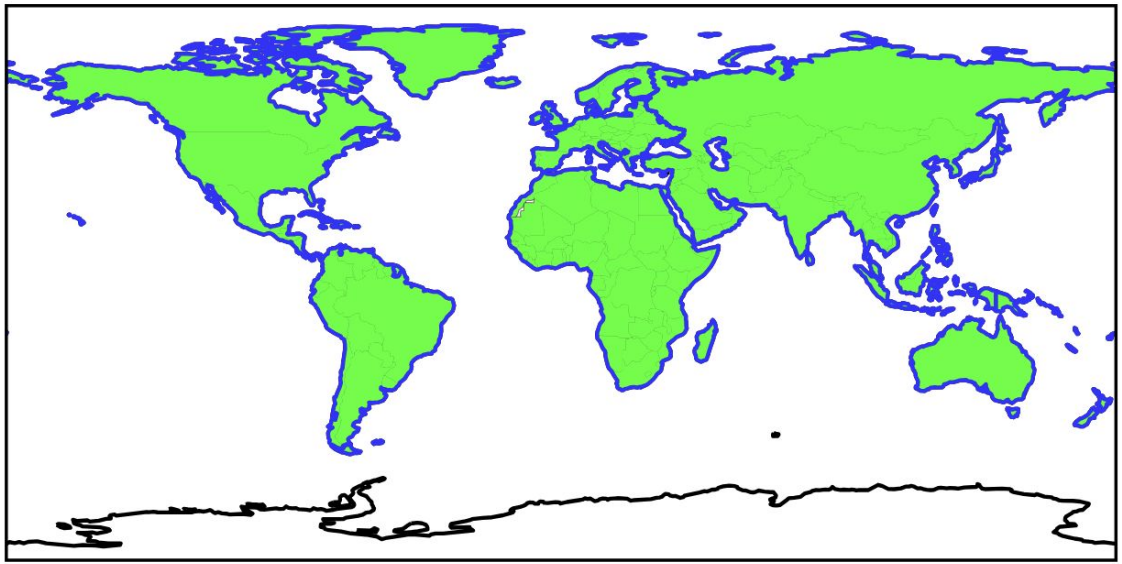
Workflow orchestration: Snakemake

MapYourGrid

PyPSA-Earth: getting a grid model

OpenStreetMap as a game-changer

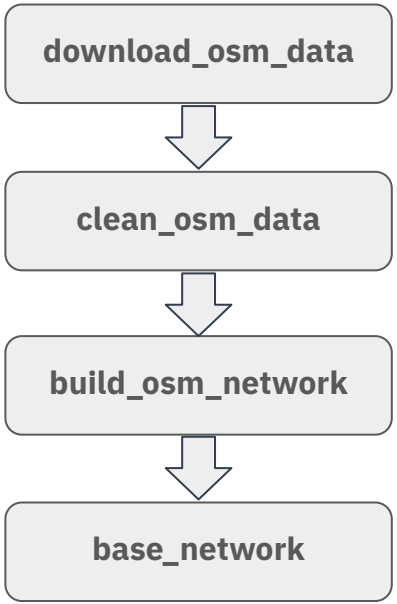


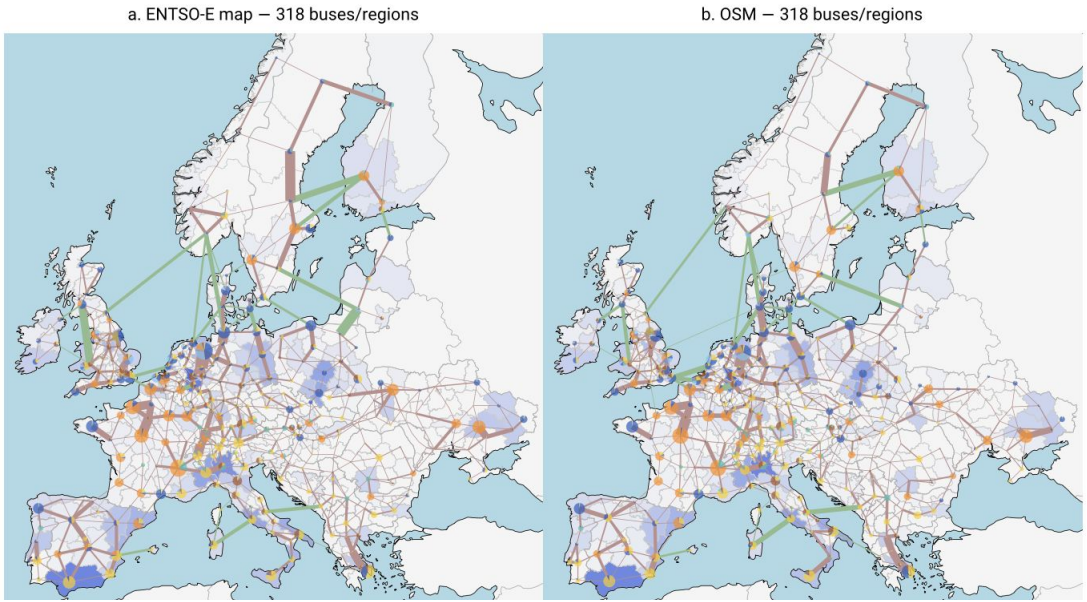


Execution status of PyPSA-Earth (the image by courtesy of Dr. Davide Fioriti)

PyPSA-Earth: getting a grid model right

OpenStreetMap as a game-changer





Validation of OSM-extracted grid against official data for Europe
Source: Xiong, Fioriti et al (2024) Modelling the high-voltage grid using open data for Europe and beyond

MapYourGrid

Interconnecting OSM and ESM

PyPSA-Earth assumptions

1. Each **line type** is mapped into **transmission capacity**
2. Each **voltage level** corresponds to **a single line type** only
3. Two types of grid are considered:
    - **three-phases** alternating current
    - high voltage **direct current**
4. The transmission capacity is determined by **thermal limits**
5. A phase angle is zero, and losses are neglected

Challenges in grid modelling

Global applicability perspective

- ➔ Over-estimation of transmission capacity
- ➔ Lack of reference data
- ➔ Missed parts of grid in OSM database

A solution: custom data

Issues:

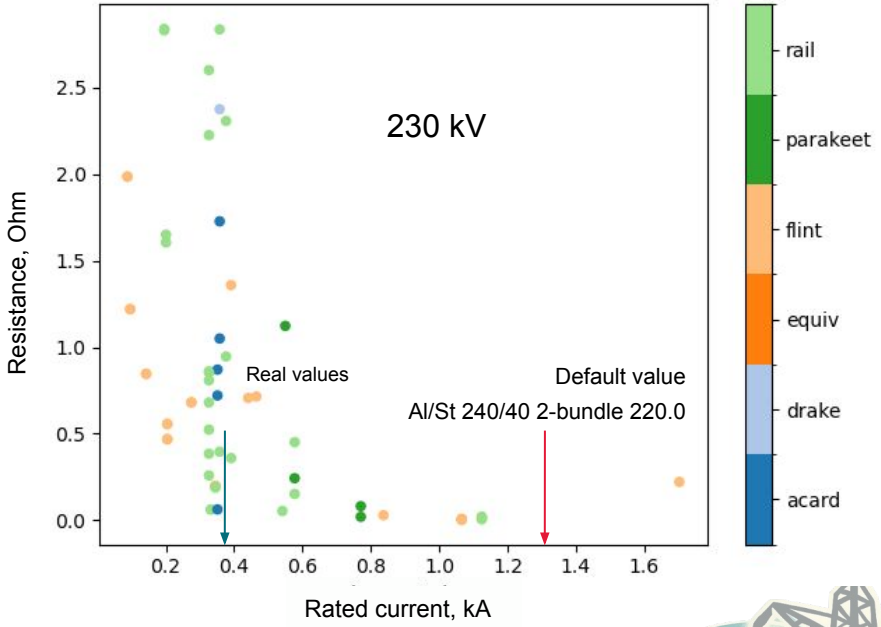
- limited reproducibility
- restricted geographic applicability
- less validation options

MapYourGrid

Reference data are crucial

Bolivia: transmission capacity validation

Data source: Ministerio de Hidrocarburos y Energías
Prepared by David Diaz



Assumptions to cross-check

- **Line types are configurable**
- **PyPSA standard types** are used
- Types parameters correspond to **European standards**

MapYourGrid

Interconnecting OSM and ESM

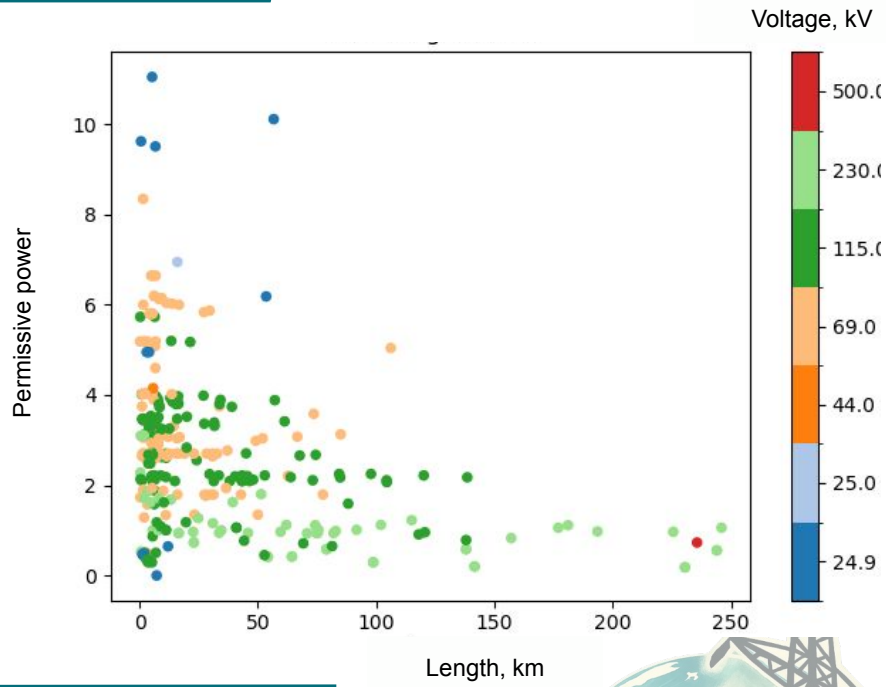
Action points

Improvements needed

- Faiture-safe treatment of missed data
- Regional-specific line types
- Improving a physical model

Joining efforts

- Regional knowledge database
  - **Awesome-Electrical-Grid-Mapping**
- Reproducible validation
  - **PyPSA-Earth-Status**



MapYourGrid

Sensitivity analysis

Translation of grid quality into planning insights

Motivation

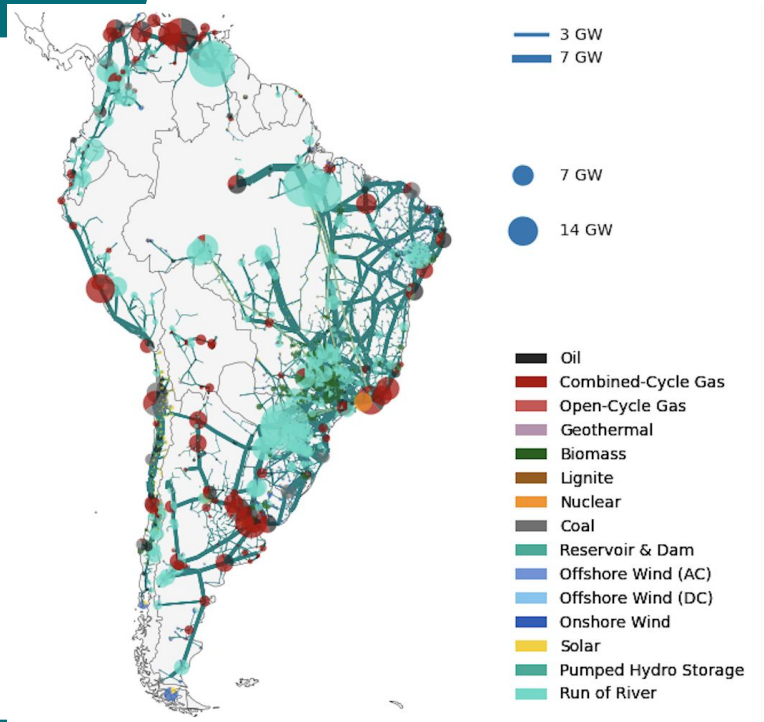
Why did we need this study?

State-of-the-art: grid data may be inaccurate and incomplete

But: energy system models use clustering which implies strong simplifications

Do we really need to bother?

Brownfield capacities extracted by PyPSA-Earth (the image by courtesy of Dr. Davide Fioriti)



MapYourGrid

Design

How was done the study?

Region: Colombia

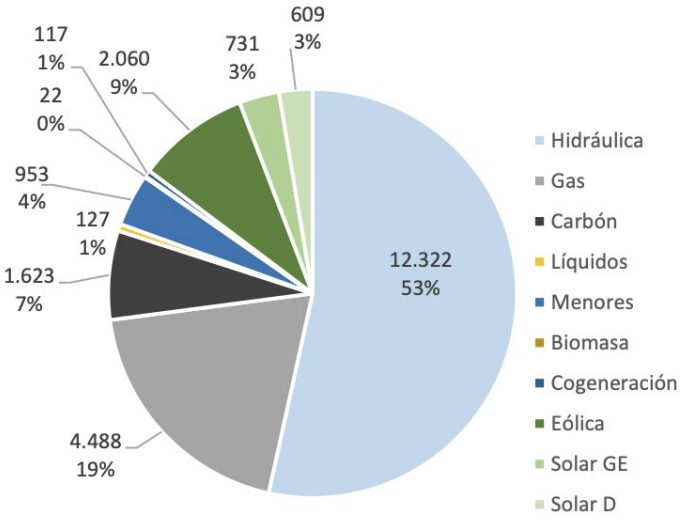
Components:

1. Customisable inputs for grid data
2. PyPSA-Earth workflow
3. Validation routine
4. Sensitivity analysis:
    a. Optimisation run
    b. Operation run

Colombia: capacity expansion study

General setup

1. A high share of hydropower
2. Limited hydro expansion planned
3. Substantial solar and wind potential
4. A projected increase in the electricity demand



Installed capacity
Source: UPME

Development toolset

Using Historical OSM feature for the pipeline

earth-osm package: load power features from OpenStreetMap

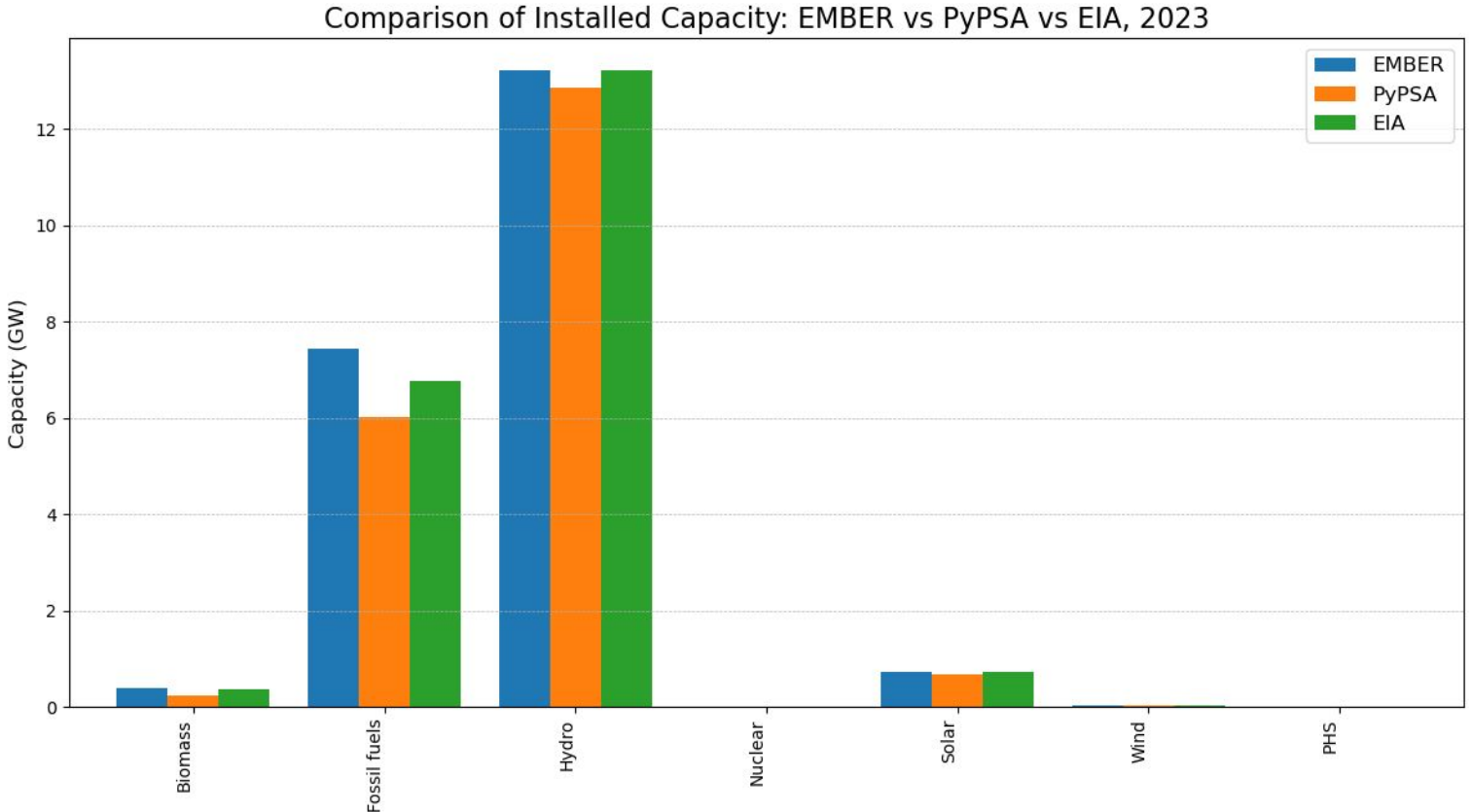
Novelty: added support of historical OSM data

Usage: select a time-stamp of interest for a reproducible sensitivity study

Colombia: sensitivity study

Validation

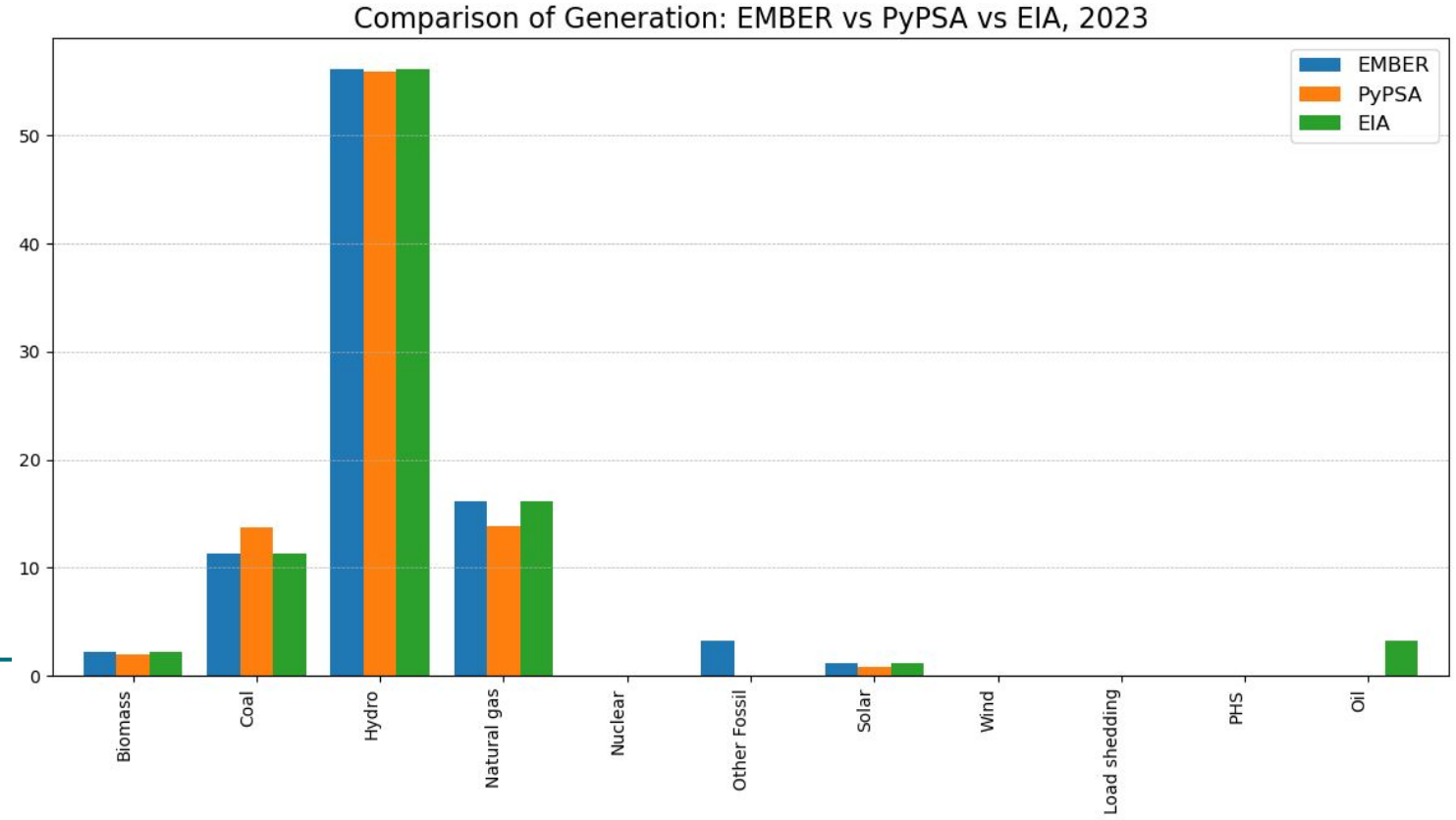
Inputs validation



Colombia: sensitivity study

Validation

Results validation



Colombia: sensitivity study

Scenarios

All scenarios included technologies costs evolution according to PyPSA technology database and projected values of the electricity demand

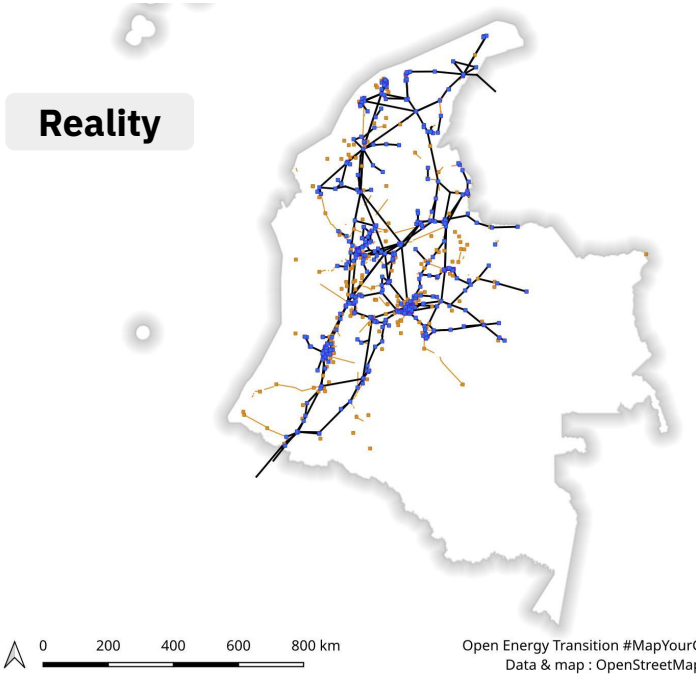
- **Business-as-usual 2030**
  - carbon price
  - expansion allowed for wind, solar and gas-fired generation
  - no transmission expansion
- **GreenPower 2030**
  - net-zero emission constraint
  - expansion allowed for wind and solar
  - 10% linewise constraint on transmission expansion
- **Green Power 2050**
  - net-zero emission constraint
  - expansion allowed for wind and solar
  - optimal transmission expansion

References:
Castaneda et al (2025)
Parra-Rodas et al (2026)
Bernal-del Río et al (2025)
Arango et al (2023)

Colombia: sensitivity study

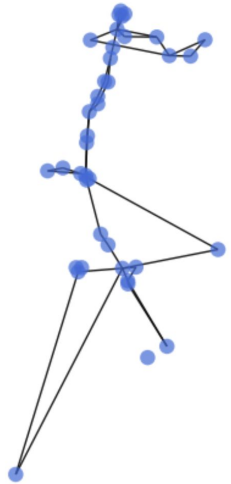
Power grid modelling

Reality

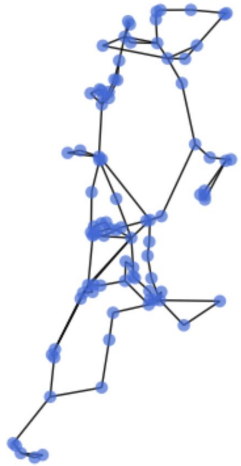


Modelling representation

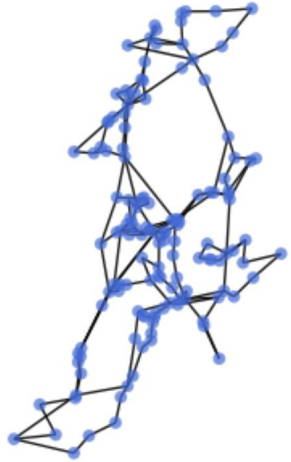
2020



2024



latest



MapYourGrid

Colombia: sensitivity study

Modelling steps

Grid inputs:
(2020, 2024, latest)

- Topology
- Capacity

Capacity expansion model

Capacities: subject to optimisation

Grid model: **2020, 2024, latest**

Objective: cost-optimal design

Capacities

- Generation
- Storage
- Transmission

Operational model

Capacities: optimised for **2020, 2024** and **latest** grid models

Grid: latest

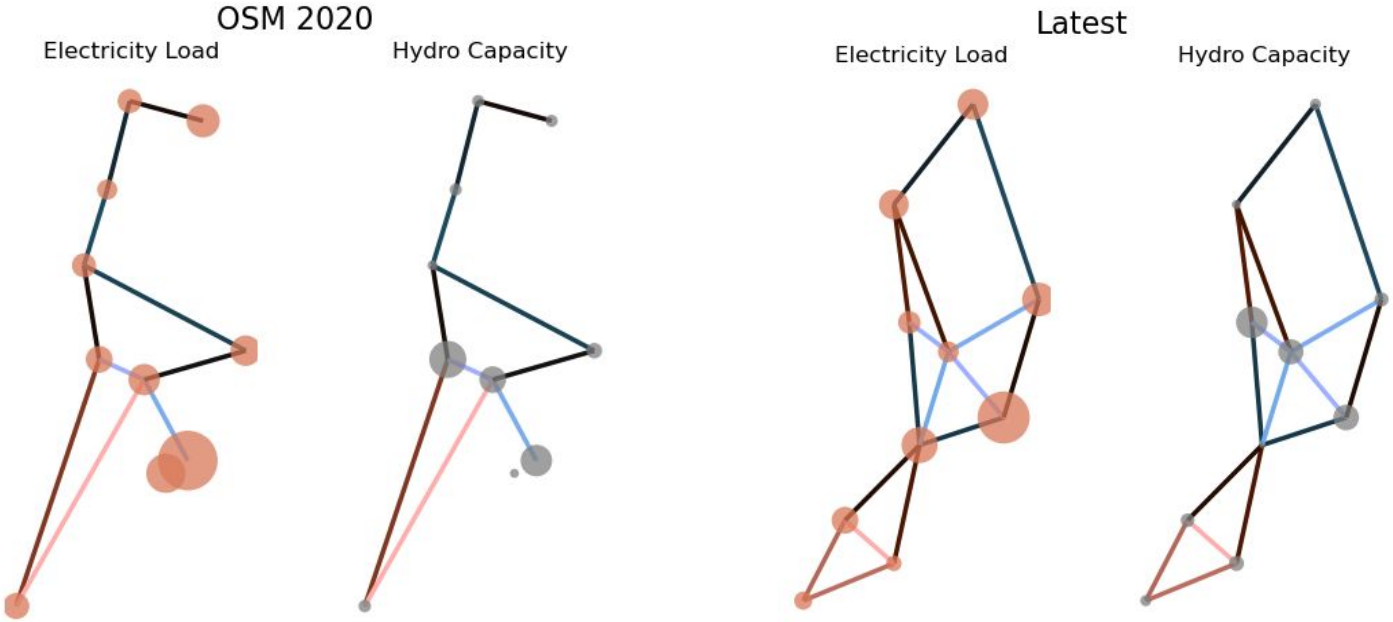
Objective: cost-optimal dispatch of the existing capacities

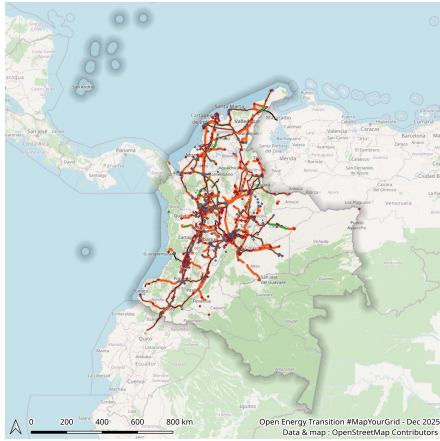
Investment outputs
(or grid models 2020, 2024, latest)

Dispatch outputs:
(for grid models 2020, 2024, latest)

Colombia: sensitivity study

Reproducing structure of the power system



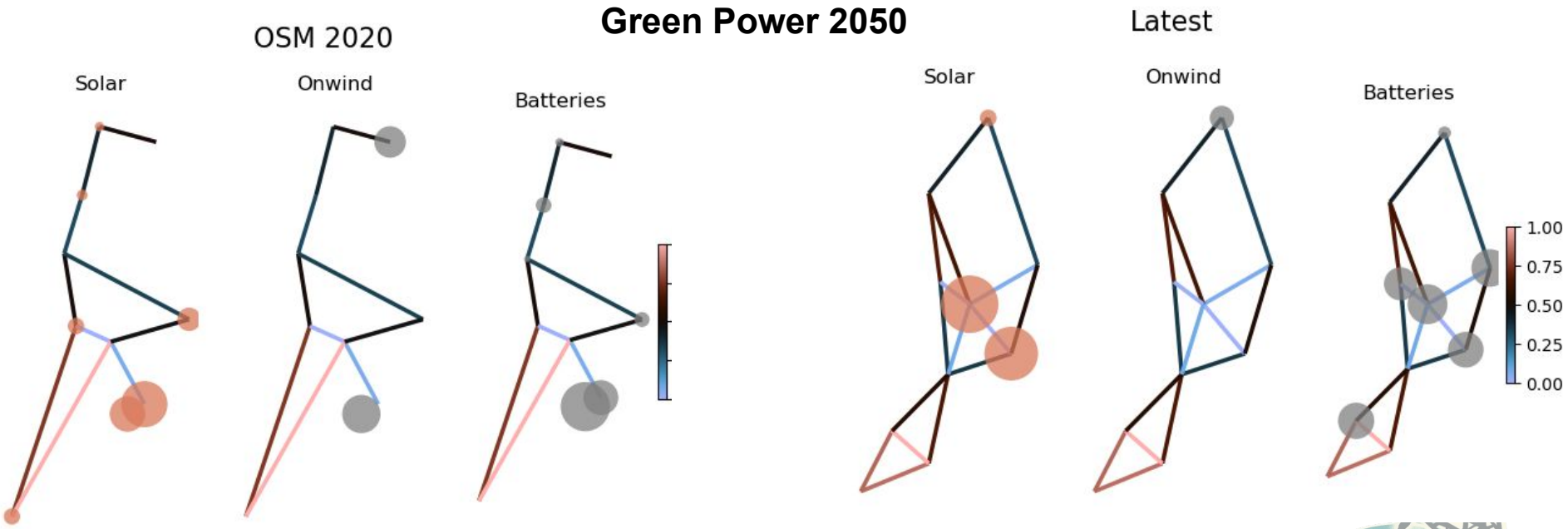


Source: MapYourGrid

MapYourGrid

Colombia: sensitivity study

Capacity expansion outputs



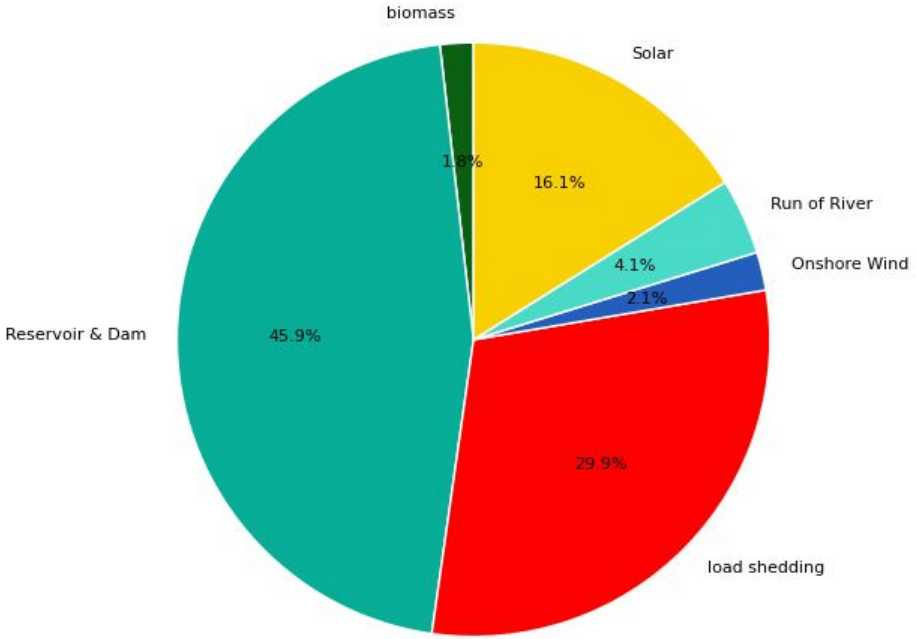
Line colors denote capacity utilisation

Colombia: capacity expansion sensitivity study

"Real-world" operation

Planning based on **inaccurate topology** results in sub-optimal system operation:

- unserved load
- curtailment of wind and solar

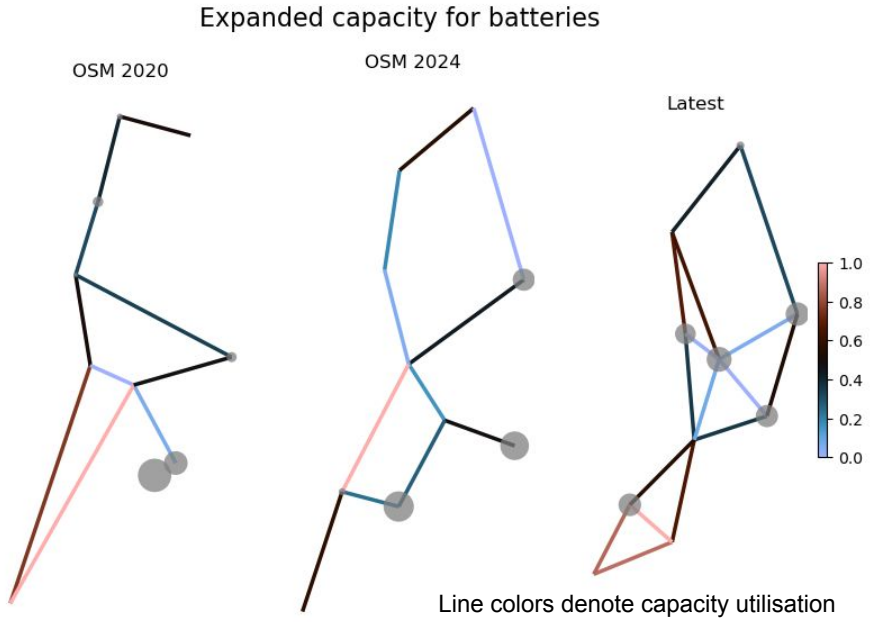


MapYourGrid

Colombia: sensitivity study

Capacity expansion outputs

- **Less accurate data** can lead to **more optimistic** optimisation outputs
- An incomplete topology effects **clustering**
- **Investments are misplaced** as a result of discrepancies in grid inputs



MapYourGrid

OSM & ESM synergies

Paradigm shift in modelling?

Challenges

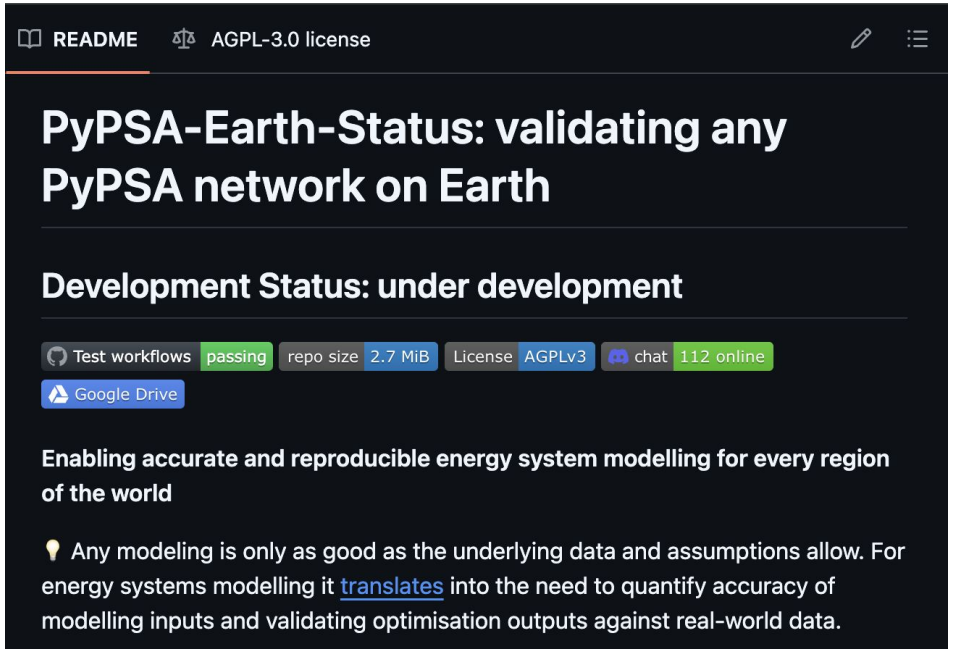
Open questions

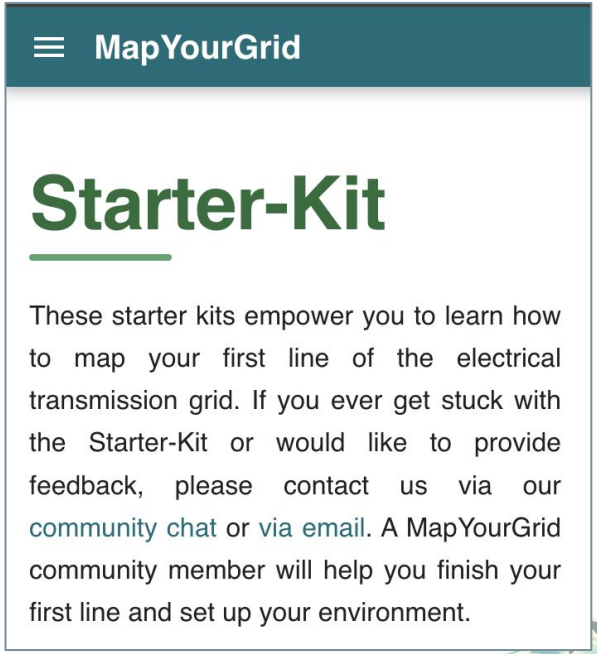
- **Increasing quality** of energy-relevant OSM data
- Regional-focused **improvements of OSM post-processing**
- Advancement of **transmission capacity** data:
  - Maintaining knowledge database on regional standards
  - Parse reference data into modeling-ready outputs
  - Enhance representation of physics behind capacity limitations
- Development of **clustering improvements**

MapYourGrid

Opportunities

Regional expertise is crucial to advance the tools





MapYourGrid

Roadmap

Translating insights into actions

1. **Advance OSM coverage** for power-relevant features
2. **Enhance OSM processing** routines
3. **Crowd-source** inputs insights for **transmission capacity** evaluation
4. Create **a global harmonised dataset** on industrial-quality inputs for power grids

MapYourGrid

Resources to go further

A list of selected references

- PyPSA-meets-Earth (initiative): https://github.com/pypsa-meets-earth
- PyPSA-Earth (model): https://github.com/pypsa-meets-earth/pypsa-earth
- Earth-OSM: https://github.com/pypsa-meets-earth/earth-osm
- Awesome-Electrical-Grid-Mapping (knowledge database): https://github.com/open-energy-transition/Awesome-Electricity-Demand
- PyPSA-Earth-Status (validation): https://github.com/pypsa-meets-earth/pypsa-earth-status

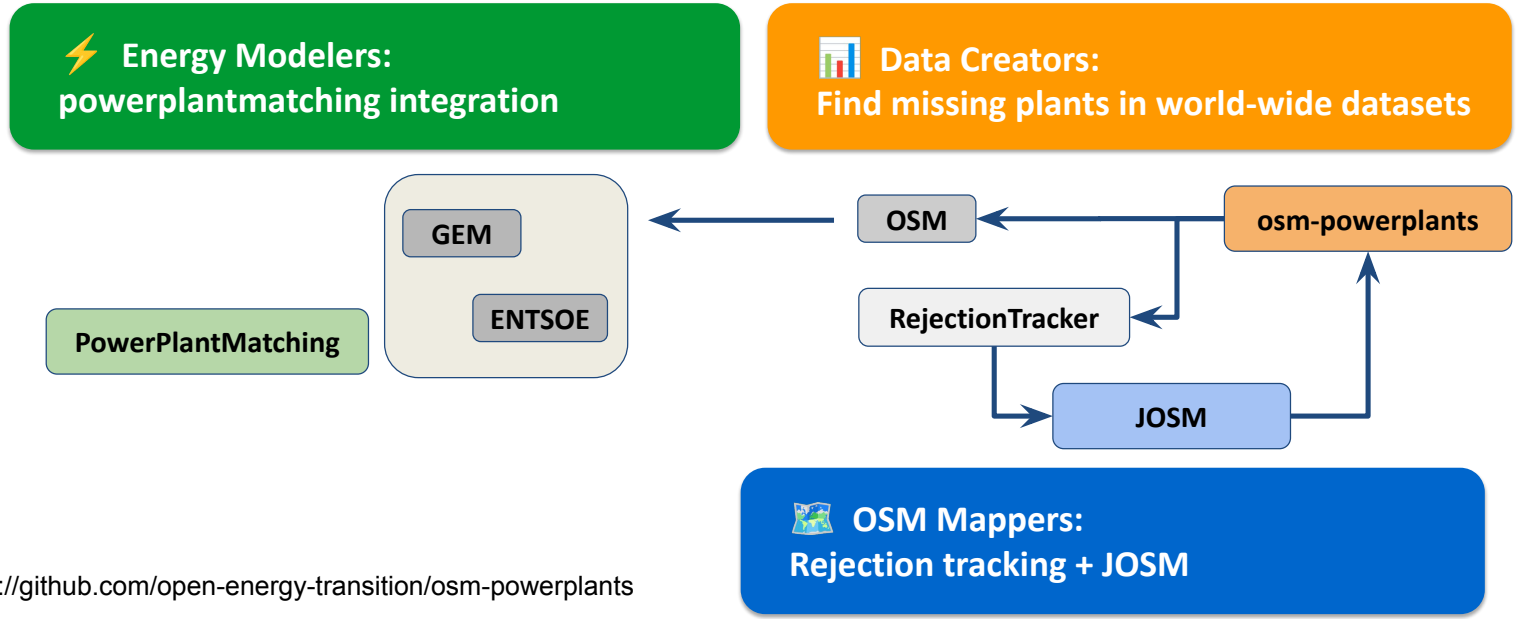
Any questions ?

Powerplantmatching osm-powerplants

Power plant data enhancement with OSM

OSM Enhances Power Plant Data

Addressing the need for: updated datasets, finding new plants, consistent energy data



https://github.com/open-energy-transition/osm-powerplants

https://github.com/PyPSA/powerplantmatching

OSM Mappers

Find and enhance power plant data consistency

1. osm-powerplants

Extract power plant data from OSM via Overpass API

2. Rejection Analysis

27 rejection reasons show exactly what needs fixing

3. Fix in JOSM

Load plants via remote control
Edit tags → Improve OSM

→ **Rejection tracking tells you exactly which tag is missing in each power plant**

⚡ Energy System Modelers

Better data coverage where other datasets have gaps

1. osm-powerplants

12,322 plants in Europe
532.9 GW total capacity

2. powerplantmatching

Combines OSM with GEM, ENTSOE, OPSD sources

3. Enhanced Dataset

11.5 GW unique capacity not in other databases

→ **OSM fills coverage gaps in peripheral countries (Malta, Cyprus, Balkans)**

Data Creators (e.g. GEM, ENTSOe)

Automatically find plants missing from databases

1. Match

powerplantmatching compares other datasets against OSM

2. Diff

Identify plants in OSM that are NOT in e.g. ENTSOE, GEM

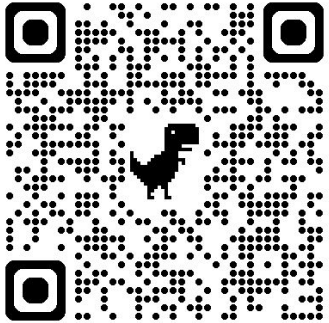
3. Review

Export to GeoJSON/CSV for validation and addition

→ **OSM enhances other datasets — systematically find missing plants**

Demonstration

Colab Notebook



https://open-energy-transition.github.io/osm-powerplants/user-guide/workshop/

Any questions ?

Closing Words on Data Quality

- **OpenStreetMap's power infrastructure is already of a very high standard in most countries.** Most critical gaps for energy system modelling are missing circuits/voltages and voltage tags.
- **Global coverage is ~76% (or ~85% if China is excluded).** Most of the missing data is in Asia, particularly China and India.
- OpenStreetMap is **an excellent source of verifiable data derived from satellite imagery** which is a **crucial input to make energy system modelling accessible worldwide.**
- **Rule of thumb from my personal very extensive mapping experience:**

The higher the voltage → The bigger the infrastructure → The higher the quality of mapping

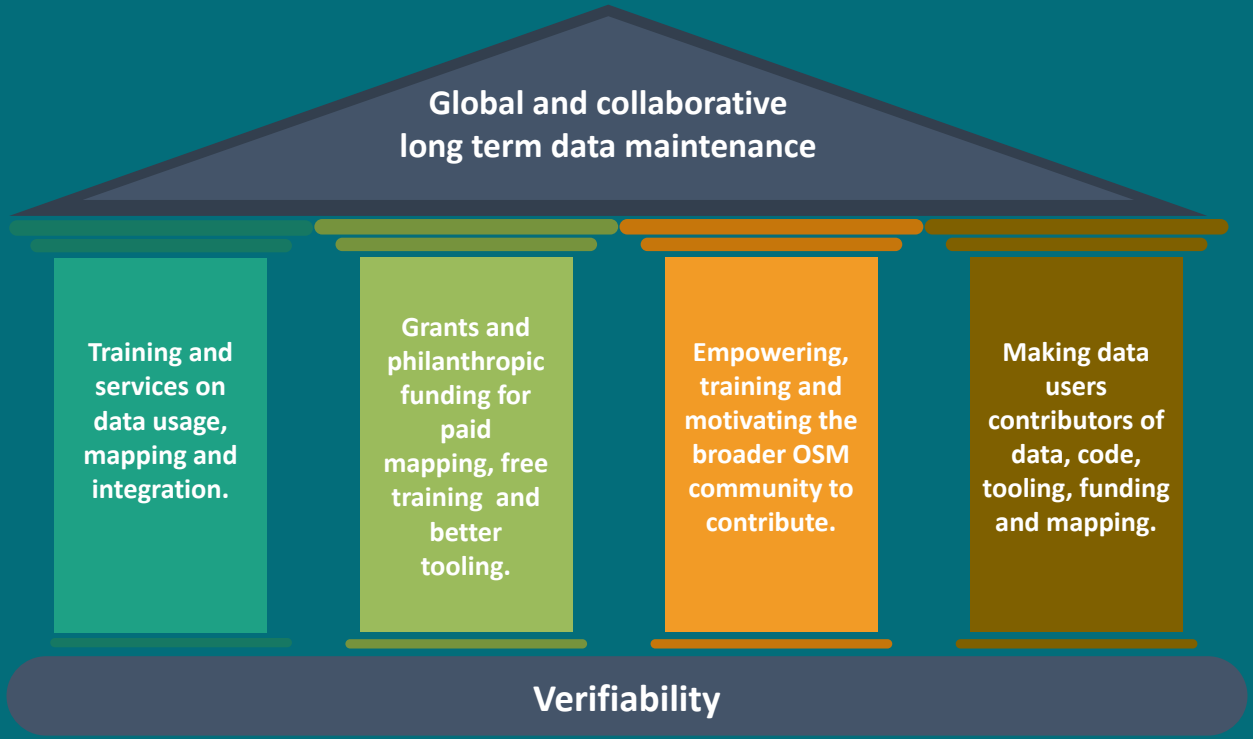
MapYourGrid

Closing Words on Data Quality

- The quality of this **data significantly impacts the modelling of energy systems**, influencing investments return and impacting operation.
- **Validation of power grid inputs** is essential to ensure reliable energy modelling outputs and **regional expertise** is indispensable for that
- The new powerplantmatching function allows users to **validate, extend and enhance various power plant datasets**, such as those from Global Energy Monitor, against OpenStreetMap's power plant data.
- **By 2025, the length of the global transmission grid had grown by around 5.6% in OpenStreetMap**, with almost a half of this growth being achieved by internal MapYourGrid mappers alone. The wider community of OpenStreetMap and MapYourGrid contributors achieved a further 3.4% growth.

MapYourGrid

The Grid Needs You



We empower you to understand and improve the data you're using

1. mapyourgrid.org
2. Join Discord: https://discord.gg/a5znpdFWfD
3. www.youtube.com/@MapYourGrid
4. Join our 2 weekly Mapping Together Meeting.
5. Receive free personal training in grid mapping.