

ENTSO-E Data Availability, Quality and Licensing

Tom Brown

Frankfurt Institute for Advanced Studies (FIAS), University of Frankfurt

7th Open Energy Modelling Initiative Workshop, Munich, 13th October 2017

1. Who Uses Energy Data?
2. Grid Data (TYNDP)
3. Market Data (Transparency Platform)
4. Licensing
5. We're openmod and we're here to help

Who Uses Energy Data?

Energy data collected by ENTSO-E is used by many **third parties**:

- **Market participants** to make informed decisions for efficient operation;
- **Researchers** to investigate the different pathways for transforming the energy system;
- **NGOs and private citizens** to be empowered to participate in the energy transition.

This work is **complementary** to the work TSOs and ENTSO-E are already doing, such as the Ten Year Network Development Plan (TYNDP).

Problem: Missing Grid Data

Researchers **cannot use** ENTSO-E's official TYNDP 2016 dataset because of:

- **Missing geo-coordinates:** We don't know where substations are (important for e.g. weather-dependent generation); security concerns can be addressed with e.g. fuzzy coordinates or aggregation
- **Cryptic substation names:** 1000s of shortened names, e.g. 'MAISA' = ?
- **Missing generator data:** No individual generator information, e.g. capacity, node, fuel type, efficiency, unit commitment info, reserve requirements
- **Missing renewables data:** Capacities, time series, reservoir capacities for hydro
- **Missing spatially resolved load:** I.e. not just per country time series, but per node
- **Missing grid data:** Norway, Sweden, Finland, Cyprus, Iceland
- **Restrictive licence:** Re-use of data is strictly prohibited

Bad Solution: Using Maps

Researchers are **forced to use maps**, such as the ENTSO-E grid map or OpenStreetMap, which is bad because:

- **Maps are inaccurate:** Contain 'artistic' distortions for readability
- **Maps don't include substation data:** Such as busbars, transformers, switches, reactive power compensation etc.
- **Maps have no load data:** No information where load and generation is connected, leading to inaccuracy.
- **Maps have no information on conducting**



ENTSO-E Transparency Platform: Missing and Strange Data

Country: D 20.11.2015 CET (UTC+1) / CEST (UTC+2)

MTU	Netherlands (NL)					
	Solar		Wind Offshore		Wind Onshore	
	Actual Aggregated	Actual Consumption	Actual Aggregated	Actual Consumption	Actual Aggregated	Actual Consumption
	[MW]	[MW]	[MW]	[MW]	[MW]	[MW]
	D	D	D	D	D	D
00:00 - 00:15	N/A	N/A	293		1299	
00:15 - 00:30	N/A	N/A	295		1284	
00:30 - 00:45	N/A	N/A	295		1243	
00:45 - 01:00	N/A	N/A	296		1293	
01:00 - 01:15	N/A	N/A	296		1309	
01:15 - 01:30	N/A	N/A	296		1393	
01:30 - 01:45	N/A	N/A	294		1301	
01:45 - 02:00	N/A	N/A	296		1347	
02:00 - 02:15	N/A	N/A	296		1336	
02:15 - 02:30	N/A	N/A	296		1496	
02:30 - 02:45	N/A	N/A	296		1592	
02:45 - 03:00	N/A	N/A	296		1542	
03:00 - 03:15	N/A	N/A	296		1554	
03:15 - 03:30	N/A	N/A	295		1535	
03:30 - 03:45	N/A	N/A	296		1456	
03:45 - 04:00	N/A	N/A	296		1446	
04:00 - 04:15	N/A	N/A	296		1517	

- A lot of data on the platform is **missing**, e.g. almost 1/4 of wind and solar data points
- Almost all time series have around 0.5% missing data (summer time & random omissions)
- Not clear what data is there and what is not
- No public interface for **reporting quality issues**
- Missing **metadata** (e.g. data source)
- Data only required by law going back five years
- **Restrictive licence:** Re-use of data requires checking with Primary Data Owner

ENTSO-E Transparency Platform Licensing

- EU Regulation No 543/2013 requires that the 'central information transparency platform shall be **available to the public free of charge through the internet**' but does **not** specify the terms and conditions of this availability.
- Current Transparency Platform Terms and Conditions do not provide a proper licence; data users must **check with all relevant primary data owners** in order to re-use data. You cannot process the data, combine it with other sources, use it in research papers or publish visualisations of the data.
- **Our view:** This is an unnecessary obstruction and **goes against the spirit** of Regulation 543/2013 to make the data 'available to the public'. Open data licences are **critical for legal certainty** when using and re-publishing data.

The merits of open data

Open refers to model source code and energy system datasets that can be **freely** used and studied, improved and distributed with **legal certainty**, provided by clear **licences**.

The benefits are clear:

- Enables market participants to make **sound decisions**
- Enables **good scientific practice** (transparency, reproducibility, credibility)
- Reduces wasteful **double-work** collecting data, enables collaboration
- Ensures **legitimacy** of modelling results (e.g. TYNDP)

This is **not a big step**, but simply gives a **legally clear, common-sense** interpretation to Regulation 543/2013's 'available to the public'.

Examples of Open Data

- **Standard open data licences** include the Open Database Licence (ODbL) and the Creative Commons Attribution 4.0 International Licence (CC BY 4.0). Standard licences have the benefit of being legally tested and compatibility with EU Directive 96/9/EC on database copyright.
- For it's new **Electricity Market Data Portal SMARD**, the Germany Federal Network Agency (BNetzA) allows re-use with attribution with CC BY 4.0.
- **Eurostat** explicitly allows re-use of datasets, unless otherwise stated.
- **Commission Decision 2011/833/EU** provides that 'Commission documents [be made] generally available for reuse without the need for individual applications, through open reuse licences or simple disclaimers' and is the basis for the European Union Open Data Portal.
- **Commission Recommendation 2012/417/EU** advised that 'publicly funded research should be widely disseminated through open access publication of scientific data and paper' and 'enable the use and reuse of scientific research results'. They maintain an Open Science Portal.
- Openness is also central to the **Digital Agenda for Europe**.

ENTSO-E is **ideally placed** to coordinate the data collection and management and solve these issues for the **common good** (including for inter-TSO cooperation). Solutions would require:

- **A full grid model:** Should contain geo-coordinates and proper names for all substations, full generator information, more load and weather snapshots, today's grid topology
- **Governance structure for data quality:** Incentivise an efficient process to correct data errors or missing data on the ENTSO-E Transparency Platform, perhaps leveraging third-party resources
- **Open data licence:** Clear, permissive licence, so everybody can re-use the data in legal safety

We're here to help

openmod **open energy modelling initiative**

The **Open Energy Modelling Initiative** is a grass roots community from universities, research institutions and the interested public (nearly 400 members). We can help:

- Understand different types of **licences**
- Help with **software**, **data formats** and **meta-data** (cf. OPSD and OEP)
- Liaise with members of the **energy modelling community**, e.g. to identify and visualise missing data (perhaps over our forum?), complement ENTSO-E's modelling work
- Help identify useful **compromises**, e.g. addressing security concerns by only releasing approximate substation coordinates, or aggregated networks









Forum for data quality questions

As a temporary measure, we could dedicate a topic in our forum to ENTSO-E TP data quality:

Secure | <https://forum.openmod-initiative.org>

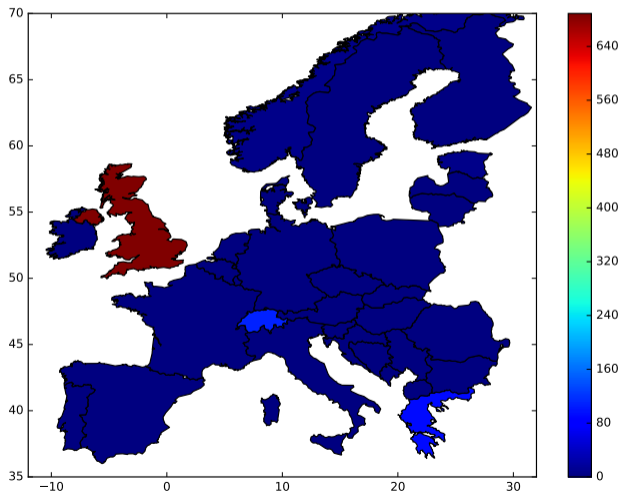
openmod open energy modelling initiative

all categories ▾ all tags ▾ **Categories** Latest New Unread Top + New Topic ▾

Category	Topics	Latest
Open data For discussing open data issues as data publishing (where and how), data acquisition (from where and how) and data processing	7	 Breakout Group on Integrating gas in energy system models 7 ■ Modelling breakout-group, frankfurt-2017 3h
Workshop Discussions covering upcoming openmod workshops . Previous events, such as the April 2017 Frankfurt workshop , may also be labeled with this category for historical reasons. ■ Breakout Groups	3	 Location for Openmod Workshop in Autumn 2017 5 ■ Workshop 5d
Modelling Discussion, tips, and announcements on the models used and maintained by openmoders. This includes design of modules, discussions on standardisation, and opening models up for wider use. ■ Module design ■ Open modelling	8	 Projects and publications regarding model evaluation 0 ■ Modelling 6d
Management Discussion of how the Openmod Initiative is managed, e.g. decisions about workshop locations, control of the websites, how and when the Openmod name can be used, etc.	4	 Breakout group about metadata 6 ■ Open data breakout-group, frankfurt-2017 6d
Site Feedback	9	 Category "breakout group" 6 ■ Site Feedback 0 votes 11d
		 Energy Modelling Platform — Europe May 2017 meeting 1 ■ Open data 11d
		 Break-out Group: Open Network Data Validation 0 ■ Open data breakout-group, frankfurt-2017 11d
		 Opening GAMS-based energy system models 0

Traffic light system for data quality

For example, number of missing data points for hourly load for 2010-2015:



Unless otherwise stated, the graphics and text are Copyright ©Tom Brown, 2017.

The graphics and text for which no other attribution are given are licensed under a Creative Commons Attribution 4.0 International Licence.

