Data Science for Energy System Modelling
Lecture 2: Time Series Analysis

Dr.-Ing. Fabian Neumann

f.neumann@tu-berlin.de
Department of Digital Transformation in Energy Systems
Technical University of Berlin, Germany

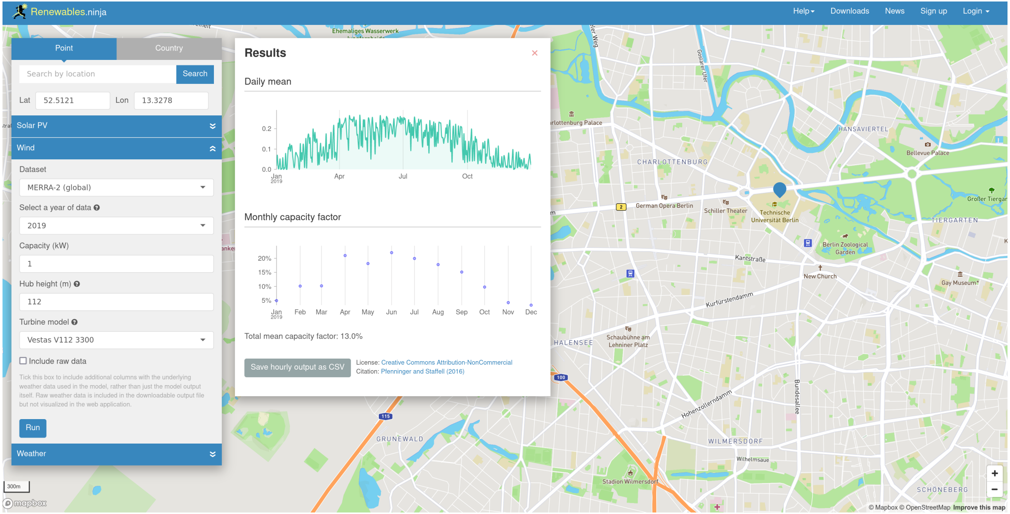
Compiled on 2025/10/23 at 15:10:37

Hybrid Format Information: Zoom and Recordings

The recordings will be **publicly available** in this ISIS course and may also be published on YouTube or other platforms.

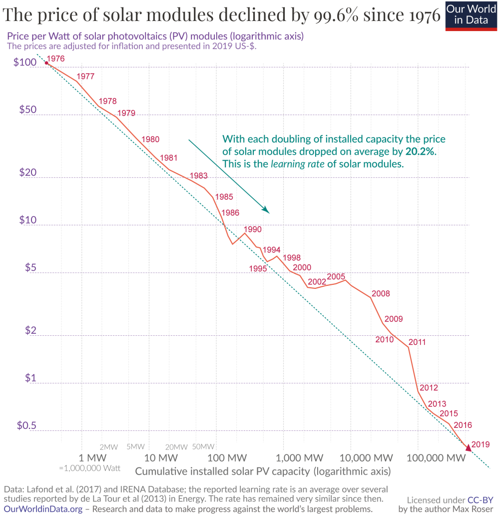
- The camera and microphone are directed **only at the lecturer**.
- Student contributions are **not recorded**.
  When interaction takes place, the recording will be **paused**.
- If you **voluntarily speak** while the recording is running, you **consent** to the recording and publication of your contribution.
- You may **withdraw** this consent at any time for the future; that part will then be deleted or edited.
- Participants who wish to remain anonymous may join under a **pseudonym** and/or use the **chat or Q&A** function instead of speaking.

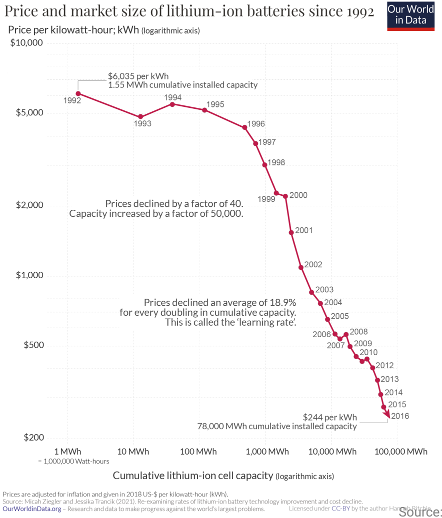
Example 1: Capacity factors of wind and solar power



Source: `https://www.renewables.ninja`

Example 2: Declining costs of solar photovoltaics and batteries





Source:
https://ourworldindata.org/cheap-renewables-growth,
https://ourworldindata.org/battery-price-decline

Time series at different scales

Hourly:

- typical data: electricity consumption, electricity production (esp. wind and solar), wholesale electricity market prices
- typical analyses: recurring daily, weekly, seasonal patterns, extreme events

Annually:

- typical data: technology cost reductions, additions to generation capacity, total electricity/energy consumption, trade balances
- typical analyses: trends, growth rates, decadal variations

In this lecture, we will focus mostly on **hourly** time series, perform **descriptive** analysis, and look at how this data can be **visualised** to yield meaningful insights.

Why do we care about time series in energy system modelling?

Build your own zero-emission electricity supply

Introduction

This tool calculates the cost of meeting a constant electricity demand from a combination of wind power, solar power and storage for different regions of the world.

First choose your location to determine the weather data for the wind and solar generation. Then choose your cost and technology assumptions to find the solution with least cost. Storage options are batteries and hydrogen from electrolysis of water.

Fun things to try out:

- remove technologies with the checkboxes, e.g. hydrogen gas storage or wind, and see system costs rise
- set solar or battery costs very low, to simulate breakthroughs in manufacturing

See also this Twitter thread for an overview of the model's features and capabilities.

This is a **toy model** with a **strongly simplified** setup. Please read the warnings below.

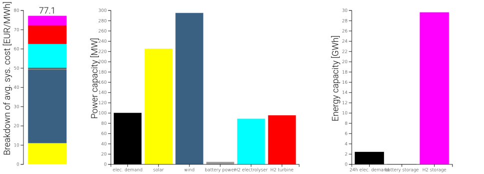
Step 1: Select location and weather year

- Select country
- Select state or province (for Australia, China, Germany, India, Russia and United States)
- Select point, rectangle or polygon, using the toolbox that appears at top-right

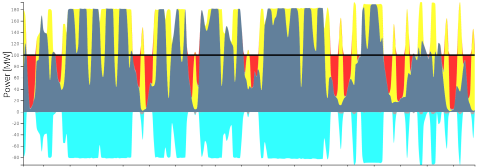
Average system cost [EUR/MWh]: 77.1

Cost is per unit of energy delivered in 2015 euros. For comparison, household electricity rates (including taxes and grid charges) averaged 211 EUR/MWh in the European Union in 2018 & 132 USD/MWh in the United States in 2019

Average marginal price of hydrogen [EUR/MWh LHV]: 78.8, [EUR/kg]: 2.60



Dispatch over a year to meet the constant demand (electricity demand is the black line; you can zoom and pan to see the details; negative values correspond to storage consuming electricity):



Source: https://model.energy/

Technische Universität Berlin

Example time series for this lecture

Consider some hourly load, wind, solar and price time series for Germany in 2015.

Download	load [GW]	onshore wind [per unit]	offshore wind [per unit]	solar [per unit]	prices [€/MWh]
⋮	⋮	⋮	⋮	⋮	⋮
2015-01-05 07:00:00	59.86	0.33	0.37	0.00	37.63
2015-01-05 08:00:00	61.18	0.33	0.35	0.06	39.09
2015-01-05 09:00:00	62.10	0.33	0.33	0.13	41.51
2015-01-05 10:00:00	63.02	0.31	0.31	0.16	40.46
2015-01-05 11:00:00	62.95	0.29	0.30	0.18	40.97
2015-01-05 12:00:00	62.11	0.25	0.28	0.18	40.72
2015-01-05 13:00:00	61.18	0.20	0.24	0.15	41.92
2015-01-05 14:00:00	60.45	0.18	0.20	0.09	42.00
⋮	⋮	⋮	⋮	⋮	⋮

Per-unit [pu] time series are normalized, where values are scaled relative to a specified base or peak value (e.g. installed capacity).

Source: Load: Open Power System Data Platform; Renewables: renewables.ninja; Prices: SMARD Platform

Technische Universität Berlin

As a first step, let's retrieve some general statistics

Mean, minimum, maximum, standard deviation, quartiles, number of values...

	load [GW]	onshore wind [per unit]	offshore wind [per unit]	solar [per unit]	prices [€/MWh]
count	8760.00	8760.00	8760.00	8760.00	8663.00
mean	54.74	0.21	0.36	0.12	31.84
std	9.89	0.19	0.29	0.18	12.48
min	32.42	0.00	0.00	0.00	-79.94
25%	46.38	0.07	0.10	0.00	25.02
50%	54.38	0.14	0.30	0.00	30.68
75%	63.87	0.28	0.59	0.20	39.92
max	75.82	0.97	0.99	0.81	99.77

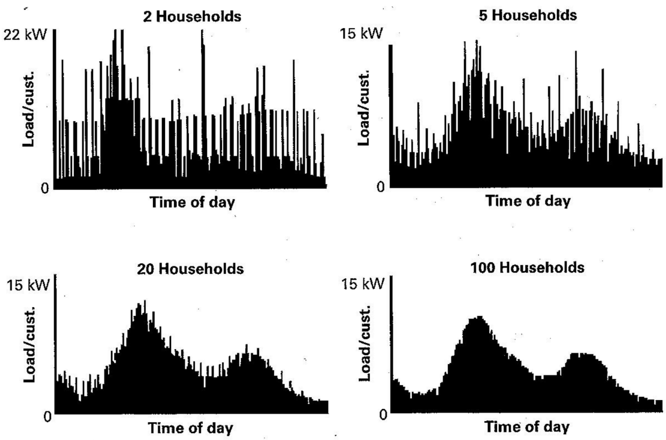
Python packages like `pandas` provide functionality to calculate these statistics in a single function call $\longrightarrow$ **see workshop next week**

Technische Universität Berlin

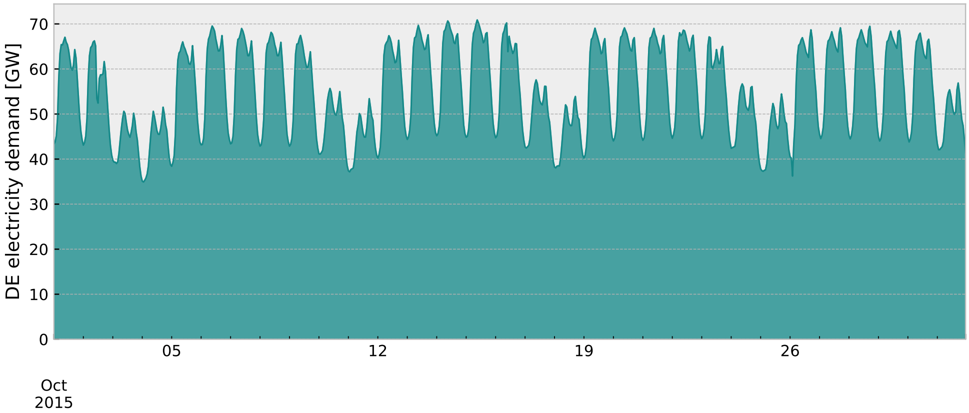
Consumers Aggregation

The load time series in the dataset have undergone **heavy aggregation**.

Actions of <u>individual consumers</u> **smooth out** if we aggregate over many consumers.

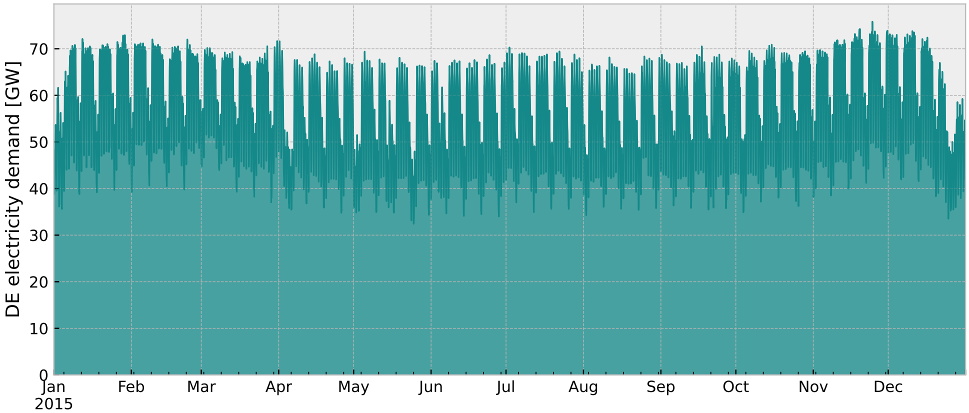


Load curve properties for a month



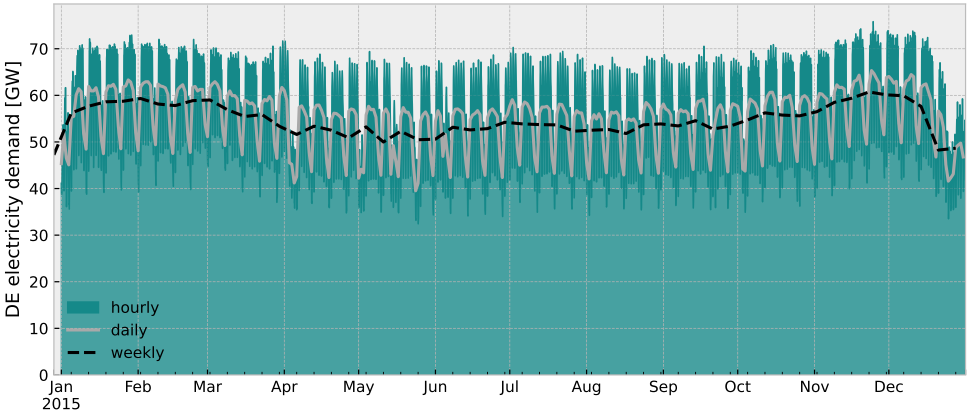
German load curve shows **daily**, **weekly** and **seasonal** patterns; holidays & weekdays are visible as well.

Load curve properties for a year



German load curve shows **daily**, **weekly** and **seasonal** patterns;
holidays & weekdays are visible as well.

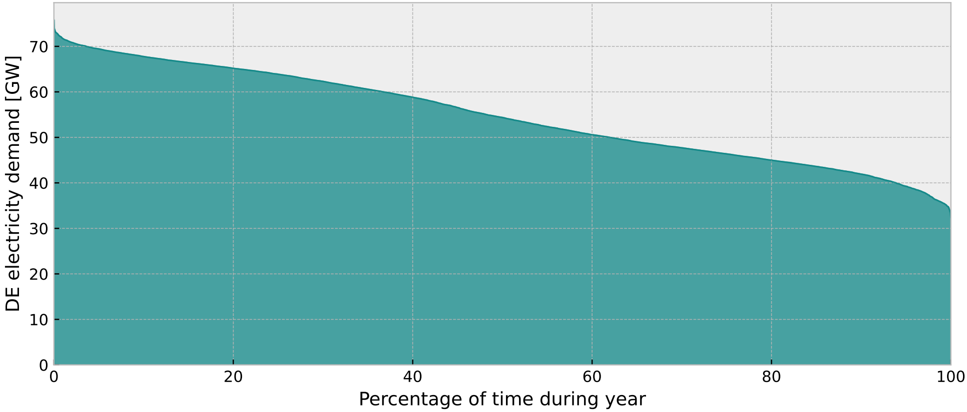
Load curve properties for a year (resampled)



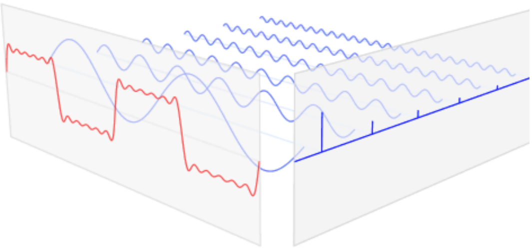
German load curve shows **daily**, **weekly** and **seasonal** patterns;
holidays & weekdays are visible as well.

Load duration curve

Construct a **duration curve** by stacking the hourly values from highest to lowest.



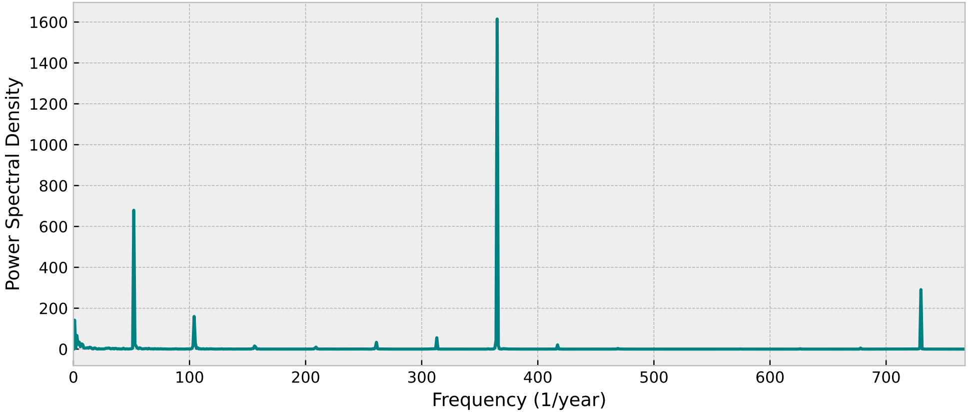
Fourier transform to see spectrum → periodic patterns



- Fourier analysis decomposes a **periodic** signal into simpler sine waves
- Every periodic signal can be broken down into a sum of sine waves with different **frequencies** and **amplitudes**
- the **Fast Fourier Transform** (FFT) is an algorithm for discrete signals, which is implemented in scientific computing libraries like `numpy` → **see workshop next week**
- Watch Youtube video for visual explanation!

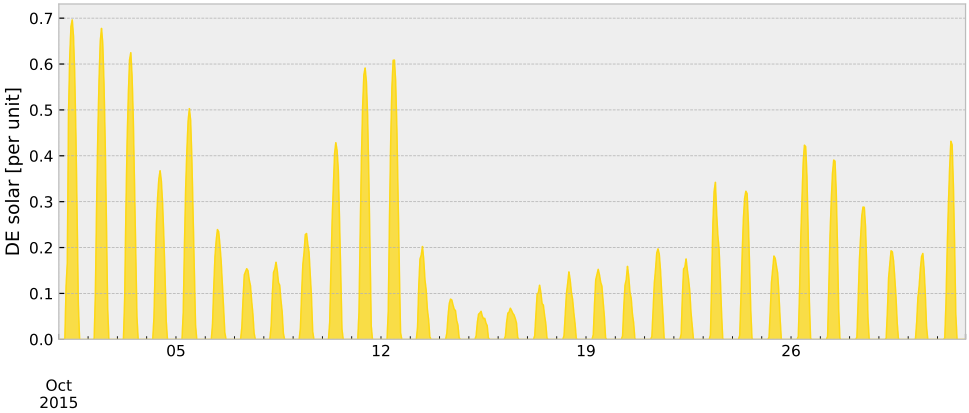
Source: https://www.ritchievink.com/blog/2017/04/23/understanding-the-fourier-transform-by-example/, https://www.youtube.com/watch?v=spUNpyF58BY

Load spectrum using Fourier transformation



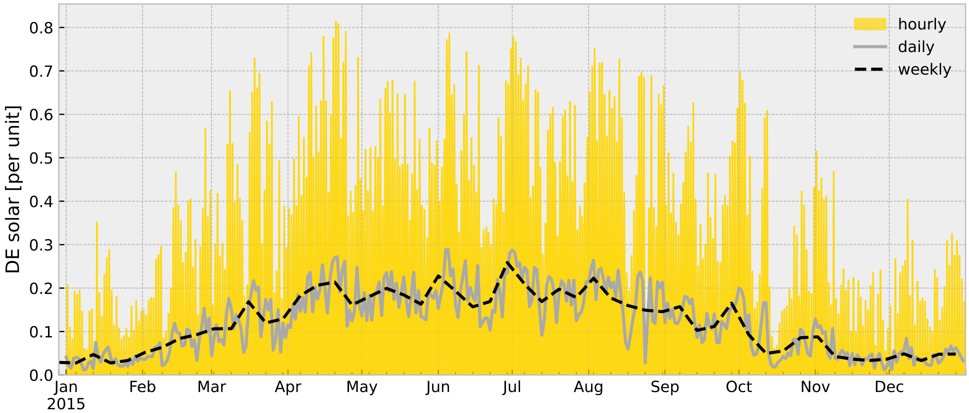
With Fourier transformation, **seasonal**, **weekly** and **daily** frequencies are clearly visible.

Solar photovoltaics time series for a month



Unlike load, solar feed-in drops to zero and never reaches full output (if aggregated to DE).

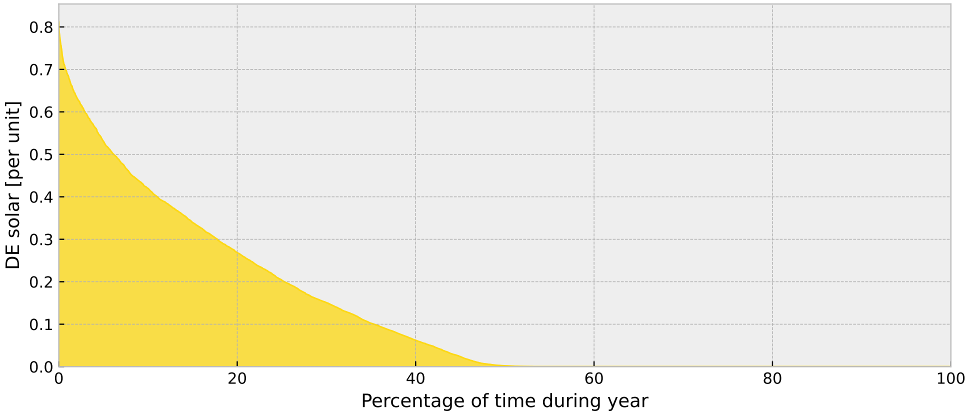
Solar photovoltaics time series: daily and weekly across a year



Daily and weekly averages underline higher solar feed-in in the summer (as expected).

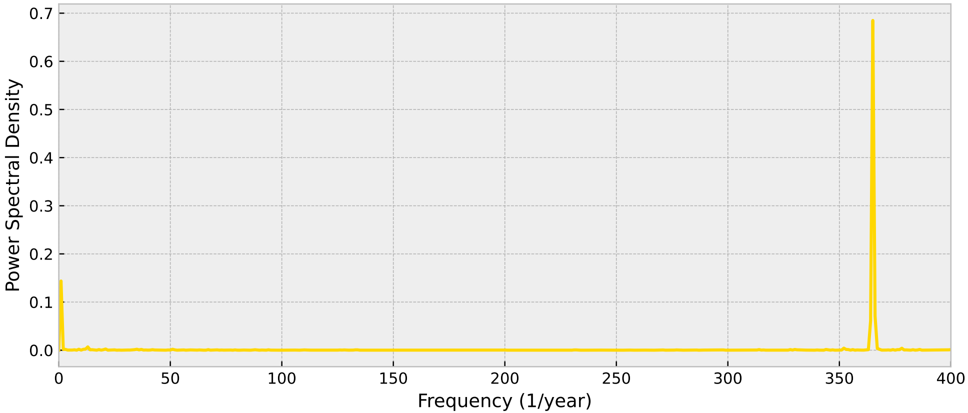
Technische Universität Berlin

Solar photovoltaics duration curve



PV duration curve nicely illustrates how it is night for half of the year.

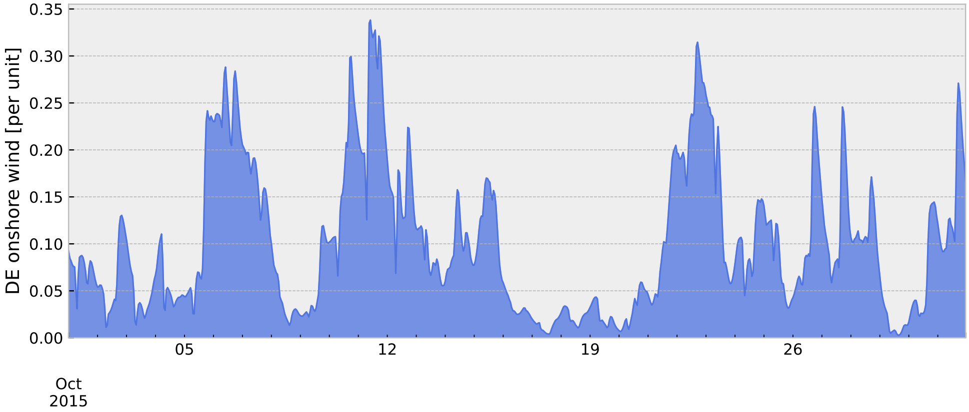
Solar photovoltaics spectrum



If we apply Fourier transformation, **seasonal** and **daily** patterns become visible.

Onshore wind time series for a month

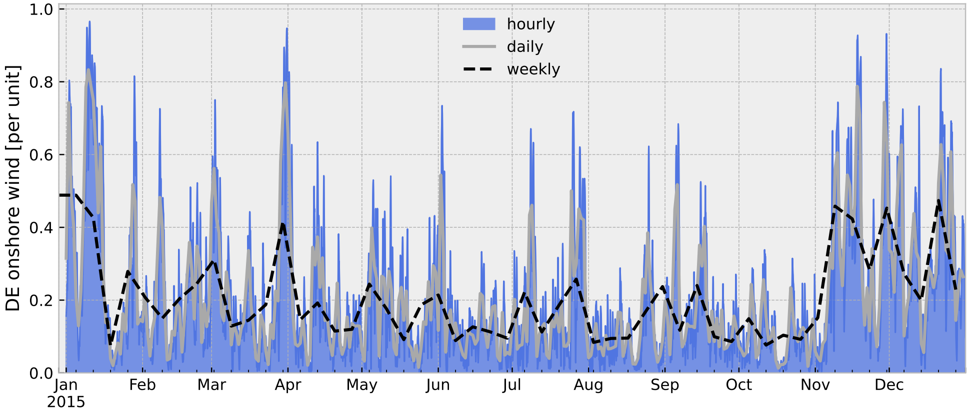
Wind is variable, like solar, but the **variations are on different time scales**.
It drops close to zero and rarely reaches full output (if aggregated for Germany).



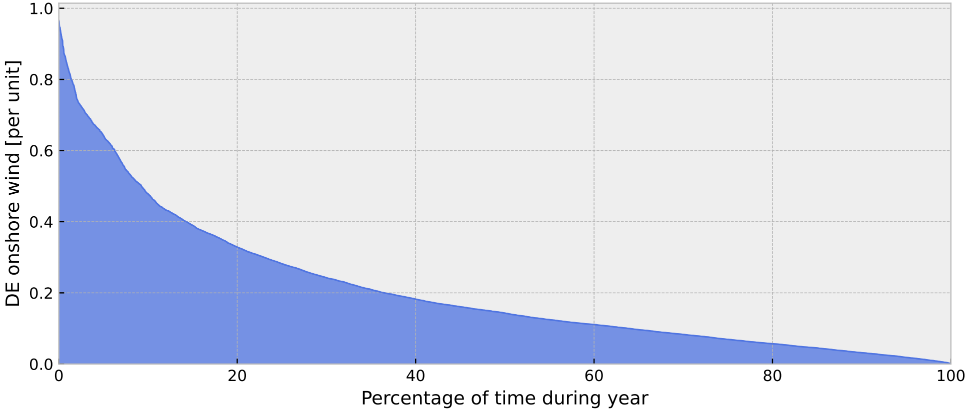
Technische Universität Berlin

Wind time series: daily and weekly patterns across a year

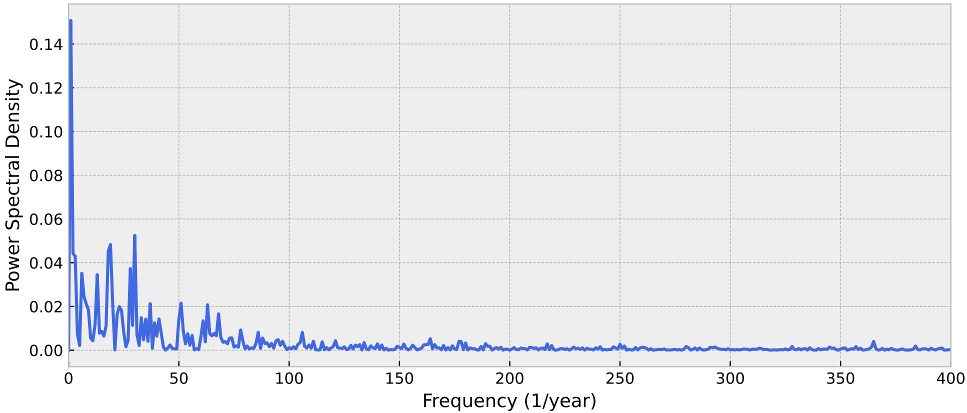
If we take a weekly average we see higher wind in the winter and some periodic patterns over 2-3 weeks (**weekly or 'synoptic' scale**).



Onshore wind duration curve



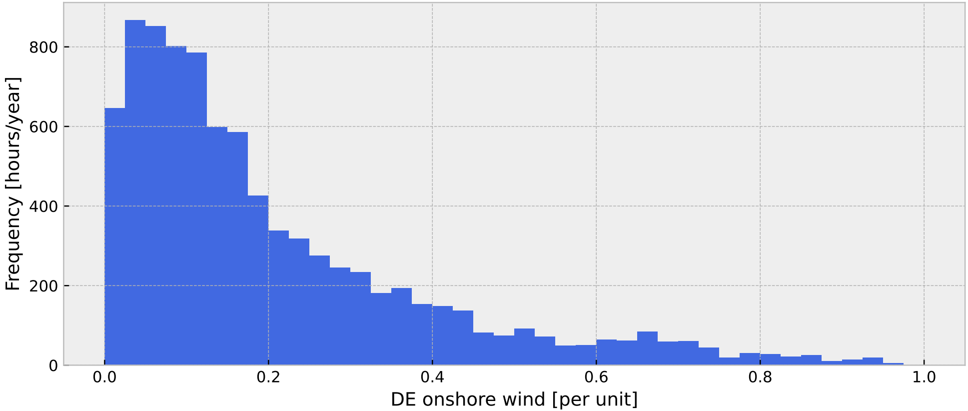
Wind power spectrum



With Fourier transform, **seasonal**, **synoptic** (2 weeks) and **daily** patterns become visible.

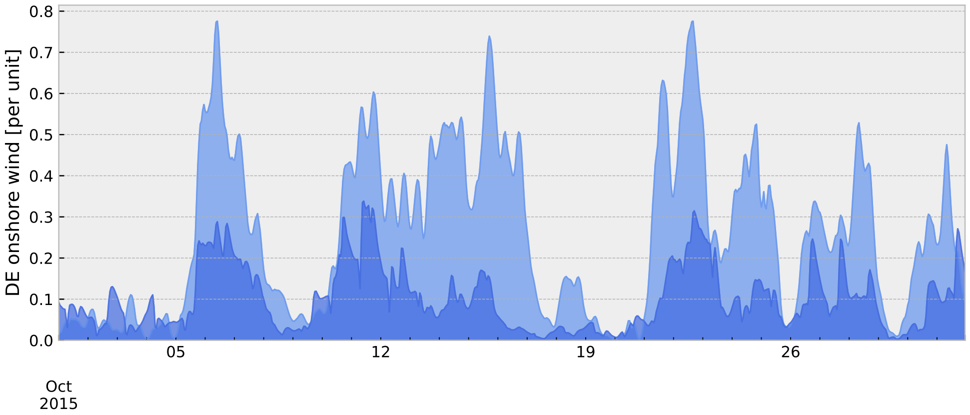
Onshore wind output histogram

The information from price duration curves can also be visualised as a **histogram**.



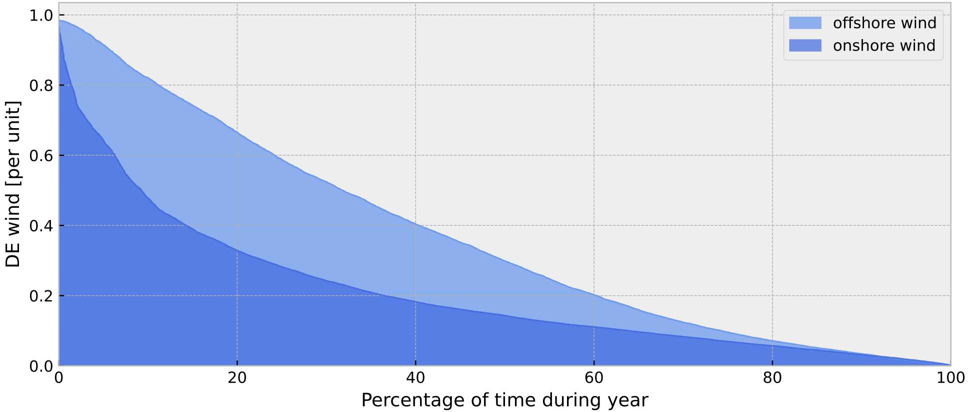
Offshore wind

Compared to onshore wind, offshore wind achieves **higher and steadier** capacity factors.



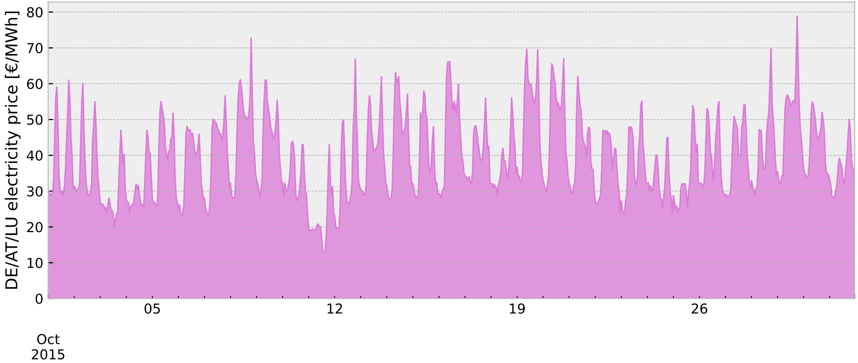
Offshore wind duration curve

Compared to onshore wind, offshore wind achieves **higher and steadier** capacity factors.



Price time series for a month

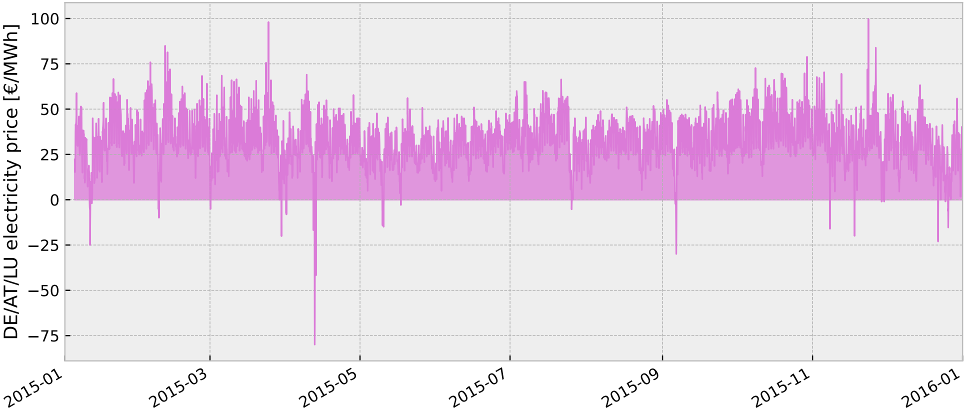
Example of **day-ahead electricity spot market prices** for the bidding zone DE/AT/LU in 2015. On the previous day until 12 PM, electricity was traded for each hour of following day.



Swings between day/night and morning/evening peaks are typical.

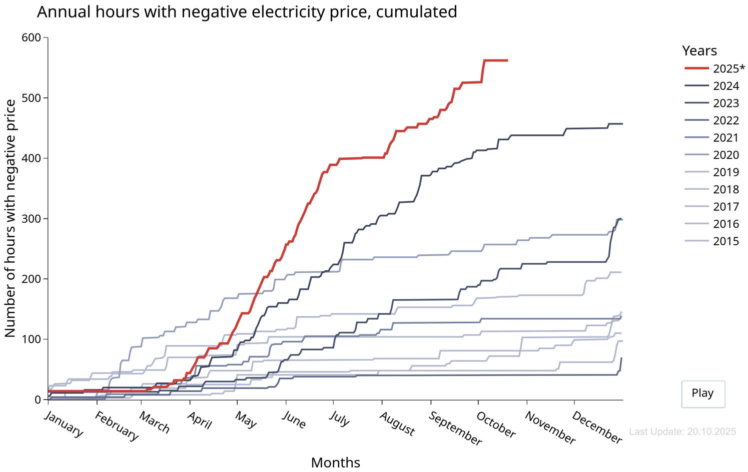
Price time series for a year

There were also a few periods in 2015, where electricity prices were **negative**.



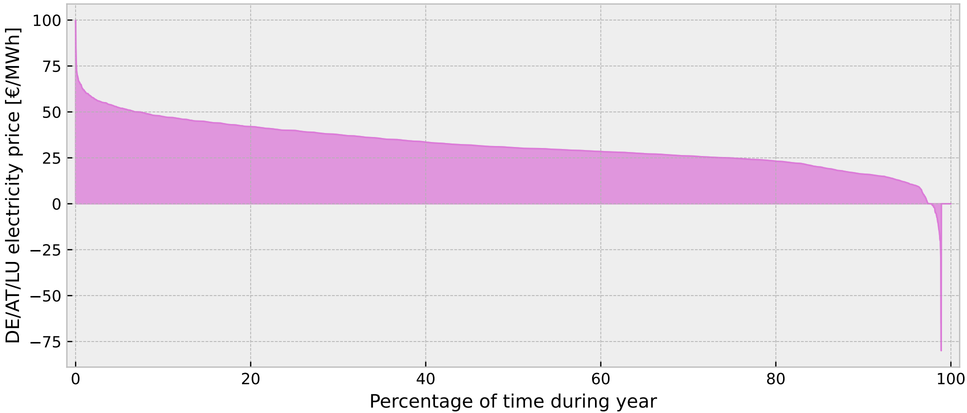
Negative prices in Germany

In 2025, these negative prices were considerably more frequent!



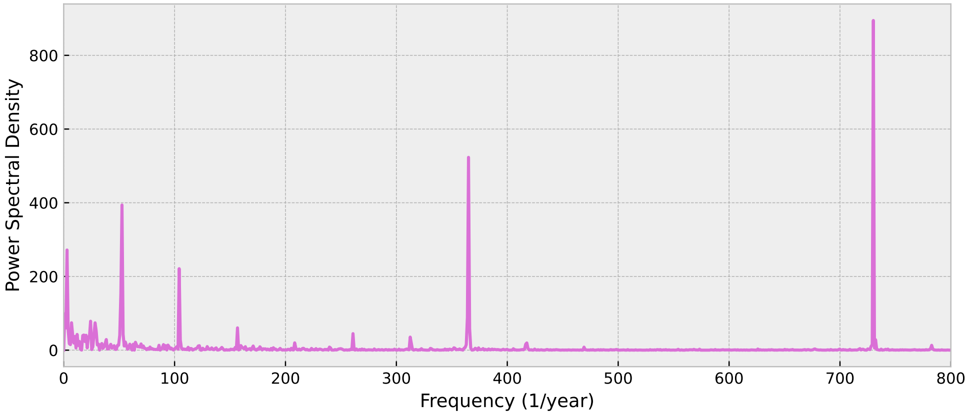
Source: https://openenergytracker.org/en/docs/germany/prices

Price duration curve



Compared to previous time series, the **extremes are more pronounced**. We also have cases of **missing data** at the very left end.

Price spectrum



With Fourier transform, **daily**, **weekly**, **seasonal** and **twice-a-day** patterns become visible.

Correlation between time series

So far, we looked at each time series independently.

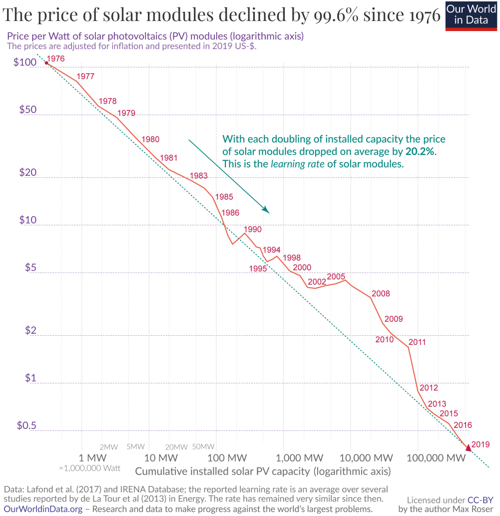
We can also measure how they relate, e.g. with the **Pearson correlation coefficient**:

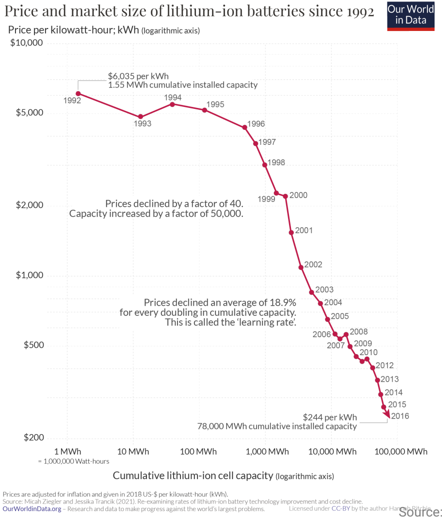
$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where n is the sample size, x_i, y_i are the sample points and $\bar{x}, \bar{y}$ are the sample means.

	load [GW]	onwind [per unit]	offwind [per unit]	solar [per unit]	prices [€/MWh]
load [GW]	1.00	0.10	0.09	0.35	0.56
onwind [per unit]	0.10	1.00	0.80	-0.14	-0.42
offwind [per unit]	0.09	0.80	1.00	-0.17	-0.36
solar [per unit]	0.35	-0.14	-0.17	1.00	0.01
prices [€/MWh]	0.56	-0.42	-0.36	0.01	1.00

Example: Declining costs of solar photovoltaics and batteries

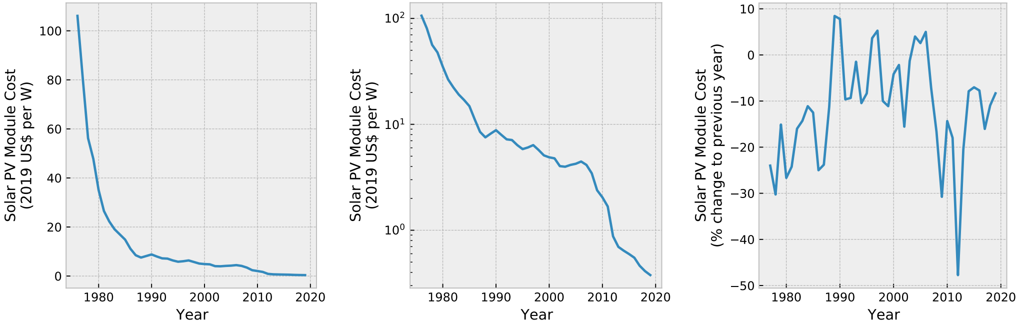




Source:
`https://ourworldindata.org/cheap-renewables-growth`,
`https://ourworldindata.org/battery-price-decline`

Visualisations of rapidly declining PV module costs

For non-stationary time series, we are typically interested in **trends** and **growth rates**.



Log-axes are typical to show near-exponential decline of costs.

Costs have reduced by **on average 11.5% per year** and now cost only 0.3% of late 70s.

Source: https://ourworldindata.org/grapher/solar-pv-prices

Up next

Next slot: Introduction to `Python`

1. Introduction to general concepts and basic functions.
2. Support with installation / setup.

Next week, slot 1: Calculation of wind and solar potentials

1. How to translate weather data to wind and solar capacity factors?
2. What land area is available to develop wind and solar parks? How to compute?

Next week, slot 2: Introduction to `numpy` and `matplotlib`.