

# 262004 UE Introduction to Computational Meteorology - Exercises (2024W)

Aiko Voigt

November 5, 2024

## 0.1 Logistics

We meet each Tuesday at 9:30 in UZA2 2F513. Note, however, that we do not meet on Oct 15, so the next meeting is on Oct 22. Please use the week of Oct 15 to work through the material on netCDF and xarray given below, and to read Jerez et al. (2015) and Buster et al. (2024).

## 0.2 Content of the exercises

We will use ERA5 reanalysis data to compute potential solar energy production and its evolution from about 1950 to today. To this end, we will use hourly data from ERA5 for surface downward solar radiation, 2m temperature and near-surface wind speed. To compute the potential solar energy production, we will apply Eqs. 1, 2 and 3 of Jerez et al. (2015).

ERA5 has a horizontal resolution of 31km. We might also experiment with using ERA5-Land, which has a finer resolution of 9km. Moreover, we will study how coarse-graining the input fields to a coarser resolution of 100 km and to daily values affects the calculated potential solar energy production. The coarse graining is interesting because it will allow us to link to the type of data that is typically provided by global climate models. Some background on why this is interesting is provided in Buster et al. (2024).

## 0.3 TeachingHub

We will work with the JupyterHub of the Department of Meteorology and Geophysics. You can access it via the Moodle page of the exercise course.

I will make sure that the ERA5 data is available on the JupyterHub.

## 0.4 Getting started with data analysis

Below are links to tutorials and videos to help you get started with geoscientific data analysis in Python. The two most important concepts to make yourself familiar with are the netCDF file format and the xarray package. We will make extensive use of both.

For netCDF, I suggest you watch <https://www.youtube.com/watch?v=UvNBnjiTXa0> (Figure 1) and <https://www.youtube.com/watch?v=699jkjLJGyM> (Figure 2). As you will see in the second video, netCDF files can easily be opened with xarray.



Figure 1: Youtube video on netCDF files by Adrian Tompkins aka Climate Unboxed. <https://www.youtube.com/watch?v=UvNBnjiTXa0>.

For xarray, I suggest you work through <http://gallery.pangeo.io/repos/pangeo-data/pangeo-tutorial-gallery/xarray.html>. You can also work interactively through the tutorial by clicking on the **launch mybinder** button in the left-top corner (Figure 3). This will start a JupyterLab session for the xarray.ipynb notebook. No registration or data transfer is needed, you can work with the notebook right away. See Figure 4.



Figure 2: Youtube video on netCDF files by “Luke Data Manager”. <https://www.youtube.com/watch?v=699jkjLJGyM>.

**PANGEO GALLERY** Pangeo Tutorial Gallery / Xarray Tutorial

[pangeo-data/pangeo-tutorial-gallery/xarray.ipynb](#)  
[launch](#) [binder](#)

- Contributor Guide
- Gallery for CESM LENS on AWS
- NCAR Gallery
- Pangeo & Dask Gateway.
- NASA CCMP Winds Gallery
- Cloud Storage Benchmarks
- Glaciology
- Landsat 8 Tutorial

## XARRAY TUTORIAL

### OVERVIEW

- **teaching:** 30 minutes
- **exercises:** 0
- **questions:**
  - What is xarray designed to do?
  - How do I create an xarray Dataset?
  - How does indexing work with xarray?
  - How do I run computations on xarray datasets?
  - What are ways to play xarray datasets?

### TABLE OF CONTENTS

Figure 3: Click on the grey-blue button to launch the tutorial in a JupyterLab session.

The screenshot shows a web browser window with the URL `https://hub.ovh2.mybinder.org/user/pangeo-data-pan-utorial-gallery-mg9k1xm0/lab`. The JupyterLab interface includes a left sidebar with a file explorer showing the directory structure of the tutorial gallery, including files like `binder-gallery.yaml`, `dask.ipynb`, `geopandas.ipynb`, `intake.ipynb`, `LICENSE`, `Readme.md`, `thumbnail.png`, and `xarray.ipynb`. The main area displays the 'Xarray Tutorial' with an 'Overview' section. This section lists the tutorial's duration (30 minutes), the number of exercises (0), and a list of questions to be answered. Below the overview is a 'Table of contents' with seven items: 'Xarray primer', 'Creating data', 'Loading data', 'Selecting data', 'Basic computations', 'Advanced computations', and 'ENSO exercise'. The 'Xarray primer' section is currently selected and visible, starting with a paragraph about `Pandas` and `Geopandas` and introducing `Xarray` for multidimensional data. It includes the mathematical notation  $T(x, y, t)$  for Earth's surface temperature.

**Xarray Tutorial**

**Overview**

- **teaching:** 30 minutes
- **exercises:** 0
- **questions:**
  - What is xarray designed to do?
  - How do I create an xarray Dataset?
  - How does indexing work with xarray?
  - How do I run computations on xarray datasets?
  - What are ways to play xarray datasets?

**Table of contents**

1. [Xarray primer](#)
2. [Creating data](#)
3. [Loading data](#)
4. [Selecting data](#)
5. [Basic computations](#)
6. [Advanced computations](#)
7. [ENSO exercise](#)

**Xarray primer**

We've seen that `Pandas` and `Geopandas` are excellent libraries for analyzing tabular "labeled data". `Xarray` is designed to make it easier to work with with *labeled multidimensional data*, data with many independent dimensions or axes. For example, we might represent Earth's surface temperature  $T$  as a three dimensional variable

$$T(x, y, t)$$

where  $x$  and  $y$  are spatial dimensions and  $t$  is time. By *labeled*, we mean data that has metadata associated with it describing the names and relationships between the variables. The temperature and precipitation sharing the same three dimensions, plus longitude and latitude as auxiliary coordinates.

Figure 4: JupyterLab session on mybinder.org.

# 1 Project topics

- Monthly climatology of PVpot over Europe
- How much area of the European continent would need to be covered with Pv to satisfy Europe's energy need, and where would the PV be best located?
- Compute the PVpot change between 1950-1959 and 2000-2009, what is the climate change impact on PVpot?
- Can we see the impact of extreme weather events on PVpot? Can we identify an example of an extreme event and its impact on PVpot? What is an extreme event in the impact of PVpot?
- If we have time: how does the PVpot estimate change if we had not hourly data but only daily mean data for temperature, radiation and winds? Could we reverse engineer the short-term fluctuation of weather with ML to correct for the lack of hourly data?

## References

- Buster, Grant, Brandon Benton, Andrew Glaws, and Ryan King. 2024. "High-Resolution Meteorology with Climate Change Impacts from Global Climate Model Data Using Generative Machine Learning." *Nature Energy* 9 (7): 894–906. <https://doi.org/10.1038/s41560-024-01507-9>.
- Jerez, Sonia, Isabelle Tobin, Robert Vautard, Juan Pedro Montávez, Jose María López-Romero, Françoise Thais, Blanka Bartok, et al. 2015. "The impact of climate change on photovoltaic power generation in Europe." *Nat. Commun.* 6 (1): 10014. <https://doi.org/10.1038/ncomms10014>.