Supplementary Information for "The EUPPBench postprocessing benchmark dataset v1.0"

Jonathan Demaeyer[1,2], Jonas Bhend[3], Sebastian Lerch[4], Cristina Primo[5], Bert Van Schaeybroeck[1], Aitor Atencia[6], Zied Ben Bouallègue[7], Jieyu Chen[4], Markus Dabernig[6], Gavin Evans[8], Jana Faganeli Pucer[9], Ben Hooper[8], Nina Horat[4], David Jobst[10], Janko Merše[11], Peter Mlakar[9,11], Annette Möller[12], Olivier Mestre[13], Maxime Taillardat[13], and Stéphane Vannitsem[1,2]

[1]Royal Meteorological Institute of Belgium, Brussels, Belgium
[2]European Meteorological Network (EUMETNET), Brussels, Belgium
[3]Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland
[4]Karlsruhe Institute of Technology, Karlsruhe, Germany
[5]Deutscher Wetterdienst, Offenbach, Germany
[6]GeoSphere Austria, Vienna, Austria
[7]European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom
[8]Met Office, Exeter, United Kingdom
[9]University of Ljubljana, Faculty of Computer and Information Science, Slovenia
[10]University of Hildesheim, Hildesheim, Germany
[11]Slovenian Environment Agency, Ljubljana, Slovenia
[12]Bielefeld University, Bielefeld, Germany
[13]Meteo France, Ecole Nationale de la Meteorologie, Toulouse, France

Correspondence: Jonathan Demaeyer (Jonathan.Demaeyer@meteo.be)

In this supplementary information, we provide an usage example for the dataset, and present additional way to access the data with other programming languages.

1 Dataset usage example with climetlab

In this section, we describe a simple usage example in Python where a set of forecasts and observations is downloaded using a dedicated climetlab plugin. We assume that the reader has installed climetlab.

We point out that other packages are currently being developed in other languages to download data from Zarr datastore (see the next section), which is the format in which the datasets are stored, however we present in this section the only official and mature enough method. Once downloaded in *xarray* dataset, the data can easily be recovered and stored on hard drives using the usual file formats. See the *xarray* documentation for more information.

With climetlab, it is very easy to download the data that you need in the dataset. The simple piece of code below download the forecasts and observations of instantaneous variables for Germany:

Installation with pip (if needed)

[1]:
```
#!pip install matplotlib
#!pip install climetlab
#!pip install climetlab-eumetnet-postprocessing-benchmark
```

Loading climetlab

[2]:
```
import climetlab as cml
```

and matplotlib

[3]:
```
import matplotlib.pyplot as plt
```

Example with deterministic forecasts

Download of deterministic high-resolution forecasts (for the 2 metre temperature) at every nearest grid point of the station points

[4]:
```
ds = cml.load_dataset('EUPPBench-training-data-stations-forecasts-surface',␣
↪"highres", "germany")
```

```
By downloading data from this dataset, you agree to the terms and conditions
defined at

    https://github.com/Climdyn/climetlab-eumetnet-postprocessing-
benchmark/blob/main/DATA_LICENSE

If you do not agree with such terms, do not download the data.
```

The `ds` object returned is a climetlab datasource that can be converted to various format. Note that the data are cached temporarily on your disk so next time you ask for this forecast I might be faster.

For example, conversion to a xarray :

[5]:
```
fcs = ds.to_xarray()
fcs
```

[5]:
```
<xarray.Dataset>
Dimensions:              (station_id: 51, number: 1, time: 730, step: 21,
surface: 1, depthBelowLandLayer: 1)
Coordinates: (12/16)
  * depthBelowLandLayer  (depthBelowLandLayer) float64 0.0
    model_altitude       (station_id) float32 ...
    model_land_usage     (station_id) int8 ...
    model_latitude       (station_id) float64 ...
    model_longitude      (station_id) float64 ...
    model_orography      (station_id) float64 ...
```

```
    ...                      ...
    station_latitude     (station_id) float64 ...
    station_longitude    (station_id) float64 ...
    station_name         (station_id) <U20 ...
  * step                 (step) timedelta64[ns] 0 days 00:00:00 ... 5 days 00...
  * surface              (surface) float64 0.0
  * time                 (time) datetime64[ns] 2017-01-01 ... 2018-12-31
Data variables: (12/15)
    cape                 (station_id, number, time, step, surface) float32 ...
    cin                  (station_id, number, time, step, surface) float32 ...
    sd                   (station_id, number, time, step, surface) float32 ...
    stl1                 (station_id, number, time, step, depthBelowLandLayer)
float32 ...
    swvl1                (station_id, number, time, step, depthBelowLandLayer)
float32 ...
    t2m                  (station_id, number, time, step, surface) float32 ...
    ...                      ...
    u10                  (station_id, number, time, step, surface) float32 ...
    u100                 (station_id, number, time, step, surface) float32 ...
    v10                  (station_id, number, time, step, surface) float32 ...
    v100                 (station_id, number, time, step, surface) float32 ...
    valid_time           (time, step) datetime64[ns] ...
    vis                  (station_id, number, time, step, surface) float32 ...
Attributes:
    Conventions:            CF-1.7
    GRIB_centre:            ecmf
    GRIB_centreDescription: European Centre for Medium-Range Weather Forecasts
    GRIB_edition:           1
    GRIB_subCentre:         0
    history:                2022-07-08T12:53 GRIB to CDM+CF via cfgrib-0.9.1...
    institution:            European Centre for Medium-Range Weather Forecasts
    land usage history:     Retrieved from https://land.copernicus.eu/pan-eu...
    land usage legend:      {1: {'label': '111 - Continuous urban fabric', '...
    land usage source:      European Union, Copernicus Land Monitoring Servi...
    model altitude history: Retrieved from https://land.copernicus.eu/imager...
    model altitude source:  European Union, Copernicus Land Monitoring Servi...
```

Retrieving the observations corresponding to the forecasts (in xarray format)

[6]:
```
obs = ds.get_observations_as_xarray()
obs
```

[6]:
```
<xarray.Dataset>
Dimensions:       (station_id: 51, step: 21, time: 730)
Coordinates:
    altitude      (station_id) float64 ...
    land_usage    (station_id) int8 ...
```

```
    latitude      (station_id) float64 ...
    longitude     (station_id) float64 ...
  * station_id    (station_id) int32 460 662 691 704 953 ... 5705 5839 5871 5906
    station_name  (station_id) <U20 ...
  * step          (step) timedelta64[ns] 0 days 00:00:00 ... 5 days 00:00:00
  * time          (time) datetime64[ns] 2017-01-01 2017-01-02 ... 2018-12-31
Data variables:
    t2m           (time, step, station_id) float64 ...
    tcc           (time, step, station_id) float64 ...
    vis           (time, step, station_id) float64 ...
Attributes:
    full_dataset_metadata:
    history:                Retrieved from https://opendata.dwd.de/climate_en...
    land usage history:     Retrieved from https://land.copernicus.eu/pan-eur...
    land usage legend:      {1: {'label': '111 - Continuous urban fabric', 'n...
    land usage source:      European Union, Copernicus Land Monitoring Servic...
    source:                 DWD, Deutscher Wetterdienst, https://www.dwd.de/
```

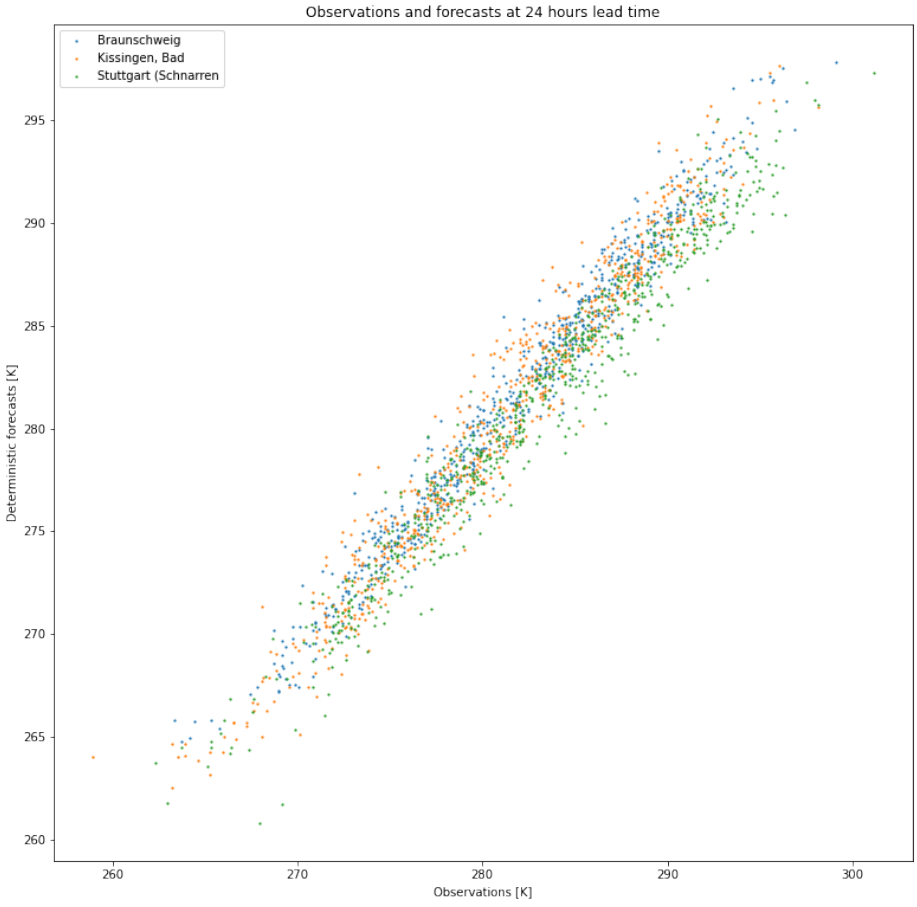
Plotting all the observations and the forecasts after 24 hours of forecast lead time and at 3 different stations

[7]:
```
stations=[1,20,40]

plt.figure(figsize=(13, 13))

for station in stations:
    plt.plot(obs.t2m.isel(station_id=station, step=4).to_numpy(), fcs.t2m.
 ↪isel(station_id=station, surface=0, step=4, number=0).to_numpy(), ls='',␣
 ↪marker='o', ms=1.2, label=obs.station_name.isel(station_id=station).to_numpy())


plt.ylabel('Deterministic forecasts ['+fcs.t2m.units+']')
plt.xlabel('Observations ['+fcs.t2m.units+']')
plt.title('Observations and forecasts at 24 hours lead time')
plt.legend();
```



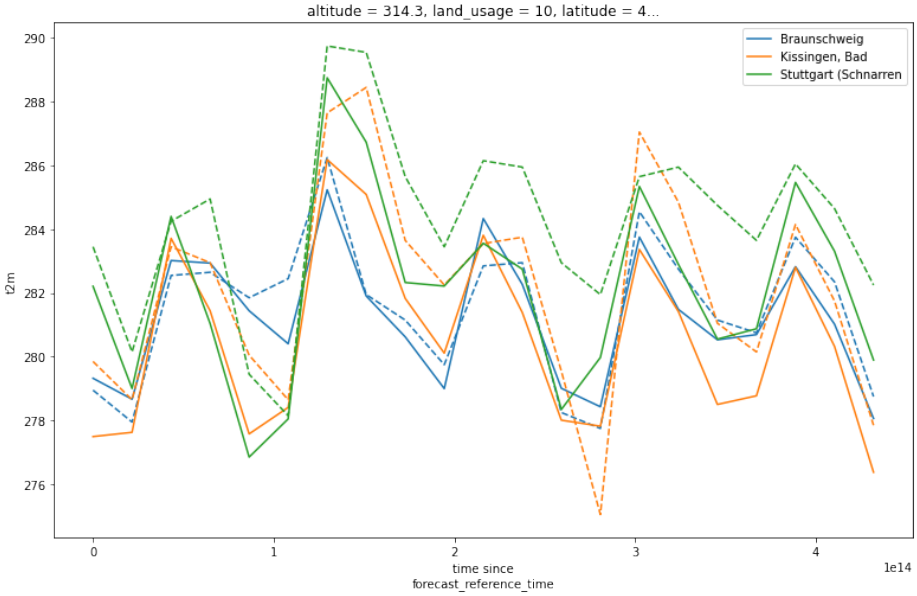
Plotting the observations and a forecast at 3 different stations (forecasts are solid lines, observations are dashed lines)

[8]:
```
stations=[1,20,40]

plt.figure(figsize=(13, 8))

for station in stations:
    line = fcs.t2m.isel(station_id=station, time=100, surface=0, number=0).
 ↪plot(label=obs.station_name.isel(station_id=station).to_numpy())
    obs.t2m.isel(station_id=station, time=100).plot(ls='--', color=line[0].
 ↪get_color())
```

```
plt.legend();
```


altitude = 314.3, land_usage = 10, latitude = 4...
Braunschweig
Kissingen, Bad
Stuttgart (Schnarren
t2m
290
288
286
284
282
280
278
276
0
1
2
3
4
1e14
time since
forecast_reference_time

2 Accessing the dataset with other methods and programming languages

In this section, we document two other ways to access the dataset, one with R and another one with Julia. Both are presently less efficient than using directly the climetlab plugin, but there will probably be quick improvement to Zarr support soon for these two languages.

2.1 Access with Julia

It is possible to access directly the dataset with Julia (Bezanson et al., 2017) using the *Zarr.jl* package (https://github.com/JuliaIO/Zarr.jl). A simple piece of code on the next pages download the forecasts of instantaneous variables for Germany. Please note that this download only the forecasts. Unlike with climetlab, with this method the users need to download the observations separately.

[1]:
```
using Zarr
```

[2]:
```
data = zopen("https://storage.ecmwf.europeanweather.cloud/
 ↪eumetnet-postprocessing-benchmark-1st-phase-training-dataset/data/
 ↪stations_data/stations_ensemble_forecasts_surface_germany.zarr")
```

```
 Warning: In the future, fill values will not be interpreted as missing values
by default.
 Please set the keyword argument `fill_as_missing` to a boolean accordingly.
Setting to `true`
 for now, but in the future `false` will be the default.
 @ Zarr /home/jupyter-jodemaey/.julia/packages/Zarr/T6jFD/src/metadata.jl:182
```

[2]:
```
ZarrGroup at Consolidated
Zarr.HTTPStore("https://storage.ecmwf.europeanweather.cloud/eumetnet-
postprocessing-benchmark-1st-phase-training-
dataset/data/stations_data/stations_ensemble_forecasts_surface_germany.zarr")
and path
Variables: number swvl1 time station_id depthBelowLandLayer stl1 surface tcwv
t2m station_altitude cape u10 v100 station_land_usage step model_orography
station_name valid_time vis tcc model_land_usage v10 model_altitude
model_longitude model_latitude sd tcw u100 station_latitude cin
station_longitude
```

[3]:
```
data.arrays
```

[3]:
```
Dict{String, ZArray} with 31 entries:
  "number"              => ZArray{Int64} of size 51
  "swvl1"               => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "time"                => ZArray{Int64} of size 730
  "station_id"          => ZArray{Int32} of size 51
  "depthBelowLandLayer" => ZArray{Union{Missing, Float64}} of size 1
  "stl1"                => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "surface"             => ZArray{Union{Missing, Float64}} of size 1
  "tcwv"                => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "t2m"                 => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "station_altitude"    => ZArray{Union{Missing, Float64}} of size 51
  "cape"                => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "u10"                 => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
  "v100"                => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
 ↪730...
```

```
  "station_land_usage"  => ZArray{Int8} of size 51
  "step"                => ZArray{Union{Missing, Float64}} of size 21
  "model_orography"     => ZArray{Union{Missing, Float64}} of size 51
  "station_name"        => ZArray{MaxLengthString{20, UInt32}} of size 51
  "valid_time"          => ZArray{Union{Missing, Float64}} of size 21 x 730
  "vis"                 => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
↪730...
  "tcc"                 => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
↪730...
  "model_land_usage"    => ZArray{Int8} of size 51
  "v10"                 => ZArray{Union{Missing, Float32}} of size 1 x 21 x␣
↪730...
  "model_altitude"      => ZArray{Union{Missing, Float32}} of size 51
  "model_longitude"     => ZArray{Union{Missing, Float64}} of size 51
  "model_latitude"      => ZArray{Union{Missing, Float64}} of size 51
                       =>
```

[4]:
```
data.attrs
```

[4]:
```
Dict{String, Any} with 13 entries:
  "history"                => "2022-07-05T05:36 GRIB to CDM+CF via cfgrib-0.9.
↪1...
  "model altitude history" => "Retrieved from https://land.copernicus.eu/
↪imager...
  "coordinates"            => "station_longitude station_land_usage␣
↪model_latit...
  "model altitude source"  => "European Union, Copernicus Land Monitoring␣
↪Servi...
  "GRIB_subCentre"         => 0
  "GRIB_centre"            => "ecmf"
  "Conventions"            => "CF-1.7"
  "GRIB_edition"           => 1
  "land usage history"     => "Retrieved from https://land.copernicus.eu/
↪pan-eu...
  "GRIB_centreDescription" => "European Centre for Medium-Range Weather␣
↪Forecas...
  "land usage source"      => "European Union, Copernicus Land Monitoring␣
↪Servi...
  "land usage legend"      => "{1: {'label': '111 - Continuous urban fabric',␣
↪'...
  "institution"            => "European Centre for Medium-Range Weather␣
↪Forecas...
```

[5]:
```
data["t2m"][1,:,:,1,1]
```

```
[5]: 21×730 reshape(::Array{Union{Missing, Float32}, 5}, 21, 730) with eltype
     Union{Missing, Float32}:
      270.874  271.376  271.056  272.13   ...  271.457  269.911  276.188  277.99
      270.24   271.518  270.594  273.258      270.127  270.396  276.181  278.07
      271.216  273.036  273.406  274.377      273.77   274.027  275.479  281.514
      271.843  270.968  272.291  274.022      272.508  273.511  275.175  279.256
      270.88   266.947  272.759  273.401      271.077  276.466  276.391  275.929
      270.541  266.583  273.681  269.755  ...  272.236  276.533  277.074  276.775
      273.121  272.155  274.59   274.344      275.111  277.453  279.911  278.509
      271.396  271.268  273.967  268.476      275.237  276.013  277.994  278.131
      267.665  271.948  272.801  265.923      277.755  276.444  276.077  276.212
      266.94   273.266  270.569  265.149      277.121  277.435  276.587  274.463
      272.674  274.367  273.443  269.81   ...  278.655  281.141  278.222  277.725
      270.727  274.365  268.27   266.044      276.537  277.421  277.719  274.0
      271.234  272.017  266.085  266.14       276.899  272.958  276.287  269.17
      273.093  270.362  264.669  265.22       277.88   276.353  275.018  268.034
      274.785  275.022  270.547  271.359      281.547  277.771  277.918  276.006
      275.386  267.891  267.632  271.165  ...  275.988  277.374  273.3    272.919
      272.027  266.146  265.711  271.632      271.806  277.746  271.072  273.124
      270.336  265.33   266.036  272.649      273.796  273.537  269.298  270.421
      275.577  271.618  273.295  273.966      277.395  277.513  275.847  276.433
      269.762  267.022  271.885  273.654      276.732  272.56   273.493  274.972
      268.224  264.677  272.047  273.384  ...  276.54   269.284  272.17   274.233
```

2.2 Access with R

It also possible to access directly the dataset with R using the *reticulate* package to load two Python packages, *xarray* (Hoyer and Joseph, 2017) and *fsspec*, and then use them to load the data in R. Again, a simple piece of code on the next pages download the forecasts of instantaneous variables for Germany. Please note that this download only the forecasts. Unlike with climetlab, with this method the users need to download the observations separately.

[1]:
```
library(reticulate)
```

[2]:
```
xarray <- import("xarray")
```

[3]:
```
fsspec <- import("fsspec")
```

[4]:
```
target_fcs <- fsspec$get_mapper("https://storage.ecmwf.europeanweather.cloud/
↪eumetnet-postprocessing-benchmark-1st-phase-training-dataset/data/
↪stations_data/stations_ensemble_forecasts_surface_germany.zarr")
```

[5]:
```
fcs <- xarray$open_zarr(target_fcs)
```

[6]:
```
fcs
```

```
<xarray.Dataset>
Dimensions:                (station_id: 51, number: 51, time: 730, step: 21,
                            surface: 1, depthBelowLandLayer: 1)
Coordinates: (12/16)
  * depthBelowLandLayer    (depthBelowLandLayer) float64 0.0
    model_altitude         (station_id) float32 ...
    model_land_usage       (station_id) int8 ...
    model_latitude         (station_id) float64 ...
    model_longitude        (station_id) float64 ...
    model_orography        (station_id) float64 ...
    ...                     ...
    station_latitude       (station_id) float64 ...
    station_longitude      (station_id) float64 ...
    station_name           (station_id) <U20 ...
  * step                   (step) timedelta64[ns] 0 days 00:00:00 ... 5 days 00...
  * surface                (surface) float64 0.0
  * time                   (time) datetime64[ns] 2017-01-01 ... 2018-12-31
Data variables: (12/15)
    cape                   (station_id, number, time, step, surface) float32 ...
    cin                    (station_id, number, time, step, surface) float32 ...
    sd                     (station_id, number, time, step, surface) float32 ...
    stl1                   (station_id, number, time, step, depthBelowLandLayer)␣
 ↪float32 ...
    swvl1                  (station_id, number, time, step, depthBelowLandLayer)␣
 ↪float32 ...
    t2m                    (station_id, number, time, step, surface) float32 ...
    ...                     ...
    u10                    (station_id, number, time, step, surface) float32 ...
    u100                   (station_id, number, time, step, surface) float32 ...
    v10                    (station_id, number, time, step, surface) float32 ...
    v100                   (station_id, number, time, step, surface) float32 ...
    valid_time             (time, step) datetime64[ns] ...
    vis                    (station_id, number, time, step, surface) float32 ...
```

```
Attributes:
    Conventions:            CF-1.7
    GRIB_centre:            ecmf
    GRIB_centreDescription: European Centre for Medium-Range Weather Forecasts
    GRIB_edition:           1
    GRIB_subCentre:         0
    history:                2022-07-05T05:36 GRIB to CDM+CF via cfgrib-0.9.1...
    institution:            European Centre for Medium-Range Weather Forecasts
    land usage history:     Retrieved from https://land.copernicus.eu/pan-eu...
    land usage legend:      {1: {'label': '111 - Continuous urban fabric', '...
    land usage source:      European Union, Copernicus Land Monitoring Servi...
    model altitude history: Retrieved from https://land.copernicus.eu/imager...
    model altitude source:  European Union, Copernicus Land Monitoring Servi...
```

[7]:
```
fcs$t2m
```

```
<xarray.DataArray 't2m' (station_id: 51, number: 51, time: 730, step: 21,
                         surface: 1)>
[39873330 values with dtype=float32]
Coordinates: (12/15)
    model_altitude      (station_id) float32 ...
    model_land_usage    (station_id) int8 ...
    model_latitude      (station_id) float64 ...
    model_longitude     (station_id) float64 ...
    model_orography     (station_id) float64 ...
  * number              (number) int64 0 1 2 3 4 5 6 7 ... 44 45 46 47 48 49 50
    ...                  ...
    station_latitude    (station_id) float64 ...
    station_longitude   (station_id) float64 ...
    station_name        (station_id) <U20 ...
  * step                (step) timedelta64[ns] 0 days 00:00:00 ... 5 days 00:...
  * surface             (surface) float64 0.0
  * time                (time) datetime64[ns] 2017-01-01 ... 2018-12-31
Attributes:
    GRIB_NV:                                0
    GRIB_Nx:                                93
    GRIB_Ny:                                125
    GRIB_cfName:                            unknown
    GRIB_cfVarName:                         t2m
    GRIB_gridDefinitionDescription:         Latitude/Longitude Grid
    GRIB_gridType:                          regular_ll
    GRIB_iDirectionIncrementInDegrees:      0.25
    GRIB_iScansNegatively:                  0
    GRIB_jDirectionIncrementInDegrees:      0.25
    GRIB_jPointsAreConsecutive:             0
    GRIB_jScansPositively:                  0
    GRIB_latitudeOfFirstGridPointInDegrees: 67.0
    GRIB_latitudeOfLastGridPointInDegrees:  36.0
```

```
GRIB_longitudeOfFirstGridPointInDegrees:  -6.0
GRIB_longitudeOfLastGridPointInDegrees:   17.0
GRIB_missingValue:                        9999
GRIB_name:                                2 metre temperature
GRIB_numberOfPoints:                      11625
GRIB_paramId:                             167
GRIB_shortName:                           2t
GRIB_stepType:                            instant
GRIB_stepUnits:                           1
GRIB_totalNumber:                         51
GRIB_typeOfLevel:                         surface
GRIB_units:                               K
coordinates:                              number time step surface latitu...
long_name:                                2 metre temperature
standard_name:                            unknown
units:                                    K
```

[8]:
```
t2m <- fcs$t2m$to_numpy()
```

[9]:
```
dim(t2m)
```

1. 51 2. 51 3. 730 4. 21 5. 1

2.3 Path to the datasets

The two methods proposed in Sections 2.1 and 2.2 require the users to provide the path to the dataset data that they want to retrieve.

To construct this path, one must use the prefix

https://storage.ecmwf.europeanweather.cloud/eumetnet-postprocessing-benchmark-1st-phase-training-dataset/data/

followed by the local path of the data you wish to access:

- gridded_data/
  - gridded_000z.zarr
  - gridded_ensemble_forecasts_pressure_500.zarr
  - gridded_ensemble_forecasts_pressure_700.zarr
  - gridded_ensemble_forecasts_pressure_850.zarr
  - gridded_ensemble_forecasts_surface.zarr
  - gridded_ensemble_forecasts_surface_postprocessed.zarr
  - gridded_ensemble_reforecasts_pressure_500.zarr
  - gridded_ensemble_reforecasts_pressure_700.zarr
  - gridded_ensemble_reforecasts_pressure_850.zarr
  - gridded_ensemble_reforecasts_surface.zarr
  - gridded_ensemble_reforecasts_surface_postprocessed.zarr
  - gridded_forecasts_efi.zarr
  - gridded_forecasts_observations_pressure_500.zarr
  - gridded_forecasts_observations_pressure_700.zarr
  - gridded_forecasts_observations_pressure_850.zarr
  - gridded_forecasts_observations_surface.zarr
  - gridded_forecasts_observations_surface_postprocessed.zarr
  - gridded_highres_forecasts_pressure_500.zarr
  - gridded_highres_forecasts_pressure_700.zarr
  - gridded_highres_forecasts_pressure_850.zarr
  - gridded_highres_forecasts_surface.zarr

- gridded_highres_forecasts_surface_postprocessed.zarr
- gridded_land_usage.zarr
- gridded_model_terrain_height.zarr
- gridded_reforecasts_observations_pressure_500.zarr
- gridded_reforecasts_observations_pressure_700.zarr
- gridded_reforecasts_observations_pressure_850.zarr
- gridded_reforecasts_observations_surface.zarr
- gridded_reforecasts_observations_surface_postprocessed.zarr

- stations_data/

  - stations_ensemble_forecasts_pressure_500_austria.zarr
  - stations_ensemble_forecasts_pressure_500_belgium.zarr
  - stations_ensemble_forecasts_pressure_500_france.zarr
  - stations_ensemble_forecasts_pressure_500_germany.zarr
  - stations_ensemble_forecasts_pressure_500_netherlands.zarr
  - stations_ensemble_forecasts_pressure_700_austria.zarr
  - stations_ensemble_forecasts_pressure_700_belgium.zarr
  - stations_ensemble_forecasts_pressure_700_france.zarr
  - stations_ensemble_forecasts_pressure_700_germany.zarr
  - stations_ensemble_forecasts_pressure_700_netherlands.zarr
  - stations_ensemble_forecasts_pressure_850_austria.zarr
  - stations_ensemble_forecasts_pressure_850_belgium.zarr
  - stations_ensemble_forecasts_pressure_850_france.zarr
  - stations_ensemble_forecasts_pressure_850_germany.zarr
  - stations_ensemble_forecasts_pressure_850_netherlands.zarr
  - stations_ensemble_forecasts_surface_austria.zarr
  - stations_ensemble_forecasts_surface_belgium.zarr
  - stations_ensemble_forecasts_surface_france.zarr
  - stations_ensemble_forecasts_surface_germany.zarr

- stations_ensemble_forecasts_surface_netherlands.zarr
- stations_ensemble_forecasts_surface_postprocessed_austria.zarr
- stations_ensemble_forecasts_surface_postprocessed_belgium.zarr
- stations_ensemble_forecasts_surface_postprocessed_france.zarr
- stations_ensemble_forecasts_surface_postprocessed_germany.zarr
- stations_ensemble_forecasts_surface_postprocessed_netherlands.zarr
- stations_ensemble_reforecasts_pressure_500_austria.zarr
- stations_ensemble_reforecasts_pressure_500_belgium.zarr
- stations_ensemble_reforecasts_pressure_500_france.zarr
- stations_ensemble_reforecasts_pressure_500_germany.zarr
- stations_ensemble_reforecasts_pressure_500_netherlands.zarr
- stations_ensemble_reforecasts_pressure_700_austria.zarr
- stations_ensemble_reforecasts_pressure_700_belgium.zarr
- stations_ensemble_reforecasts_pressure_700_france.zarr
- stations_ensemble_reforecasts_pressure_700_germany.zarr
- stations_ensemble_reforecasts_pressure_700_netherlands.zarr
- stations_ensemble_reforecasts_pressure_850_austria.zarr
- stations_ensemble_reforecasts_pressure_850_belgium.zarr
- stations_ensemble_reforecasts_pressure_850_france.zarr
- stations_ensemble_reforecasts_pressure_850_germany.zarr
- stations_ensemble_reforecasts_pressure_850_netherlands.zarr
- stations_ensemble_reforecasts_surface_austria.zarr
- stations_ensemble_reforecasts_surface_belgium.zarr
- stations_ensemble_reforecasts_surface_france.zarr
- stations_ensemble_reforecasts_surface_germany.zarr
- stations_ensemble_reforecasts_surface_netherlands.zarr
- stations_ensemble_reforecasts_surface_postprocessed_austria.zarr
- stations_ensemble_reforecasts_surface_postprocessed_belgium.zarr
- stations_ensemble_reforecasts_surface_postprocessed_france.zarr

- stations_ensemble_reforecasts_surface_postprocessed_germany.zarr
- stations_ensemble_reforecasts_surface_postprocessed_netherlands.zarr
- stations_forecasts_efi_austria.zarr
- stations_forecasts_efi_belgium.zarr
- stations_forecasts_efi_france.zarr
- stations_forecasts_efi_germany.zarr
- stations_forecasts_efi_netherlands.zarr
- stations_forecasts_observations_surface_austria.zarr
- stations_forecasts_observations_surface_belgium.zarr
- stations_forecasts_observations_surface_france.zarr
- stations_forecasts_observations_surface_germany.zarr
- stations_forecasts_observations_surface_netherlands.zarr
- stations_forecasts_observations_surface_postprocessed_austria.zarr
- stations_forecasts_observations_surface_postprocessed_belgium.zarr
- stations_forecasts_observations_surface_postprocessed_france.zarr
- stations_forecasts_observations_surface_postprocessed_germany.zarr
- stations_forecasts_observations_surface_postprocessed_netherlands.zarr
- stations_highres_forecasts_pressure_500_austria.zarr
- stations_highres_forecasts_pressure_500_belgium.zarr
- stations_highres_forecasts_pressure_500_france.zarr
- stations_highres_forecasts_pressure_500_germany.zarr
- stations_highres_forecasts_pressure_500_netherlands.zarr
- stations_highres_forecasts_pressure_700_austria.zarr
- stations_highres_forecasts_pressure_700_belgium.zarr
- stations_highres_forecasts_pressure_700_france.zarr
- stations_highres_forecasts_pressure_700_germany.zarr
- stations_highres_forecasts_pressure_700_netherlands.zarr
- stations_highres_forecasts_pressure_850_austria.zarr
- stations_highres_forecasts_pressure_850_belgium.zarr

- stations_highres_forecasts_pressure_850_france.zarr
- stations_highres_forecasts_pressure_850_germany.zarr
- stations_highres_forecasts_pressure_850_netherlands.zarr
- stations_highres_forecasts_surface_austria.zarr
- stations_highres_forecasts_surface_belgium.zarr
- stations_highres_forecasts_surface_france.zarr
- stations_highres_forecasts_surface_germany.zarr
- stations_highres_forecasts_surface_netherlands.zarr
- stations_highres_forecasts_surface_postprocessed_austria.zarr
- stations_highres_forecasts_surface_postprocessed_belgium.zarr
- stations_highres_forecasts_surface_postprocessed_france.zarr
- stations_highres_forecasts_surface_postprocessed_germany.zarr
- stations_highres_forecasts_surface_postprocessed_netherlands.zarr
- stations_reforecasts_observations_surface_austria.zarr
- stations_reforecasts_observations_surface_belgium.zarr
- stations_reforecasts_observations_surface_france.zarr
- stations_reforecasts_observations_surface_germany.zarr
- stations_reforecasts_observations_surface_netherlands.zarr
- stations_reforecasts_observations_surface_postprocessed_austria.zarr
- stations_reforecasts_observations_surface_postprocessed_belgium.zarr
- stations_reforecasts_observations_surface_postprocessed_france.zarr
- stations_reforecasts_observations_surface_postprocessed_germany.zarr
- stations_reforecasts_observations_surface_postprocessed_netherlands.zarr

For instance, the example in the previous sections combine:

"https://storage.ecmwf.europeanweather.cloud/eumetnet-postprocessing-benchmark-1st-phase-training-dataset/data/"

and "stations_data/" and "stations_ensemble_forecasts_surface_germany.zarr" to obtain the forecasts of instantaneous variables for Germany. Data paths are in general explicit, except maybe for the word "postprocessed" which means a variable that has been "processed", i.e. accumulated or filtered (see the article and the documentation where these variable are called *processed*).

References

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B.: Julia: A fresh approach to numerical computing, SIAM review, 59, 65–98, https://doi.org/10.1137/141000671, 2017.

Hoyer, S. and Joseph, H.: xarray: N-D labeled Arrays and Datasets in Python, Journal of Open Research Software, 5, https://doi.org/10.5334/jors.148, 2017.