

Meteostat_vs_MeteoCH

October 28, 2023

1 Meteostat vs. MeteoCH

```
[148]: # 'Soft' reset: Only clears your namespace, leaving history intact.
%reset -sf
#[Howto: Indexing DataFrames in Pandas](https://www.dataquest.io/blog/
↳tutorial-indexing-dataframes-in-pandas/)
```

```
[149]: import pandas as pd
import matplotlib.pyplot as plt
from datetime import datetime
from meteostat import Stations
from meteostat import Daily
from IPython.display import Markdown as md
```

1.1 Weather stations at MeteoCH

```
[150]: url = 'https://data.geo.admin.ch'
path = 'ch.meteoschweiz.klima/nbcn-tageswerte'
wsurl = url + '/' + path + '/' + 'liste-download-nbcn-d.csv'
ws = pd.read_csv(wsurl, sep=";", header=0, encoding = "ISO-8859-1").dropna()
pd.options.display.float_format = '{:.2f}'.format
ws.rename(columns={'Station height m. a. sea level': 'elevation'}, inplace=True)
ws.drop(['WIGOS-ID', 'CoordinatesE', 'CoordinatesN', 'URL Previous years',
↳(verified data)',
'URL Current year', 'station/location', 'Canton', 'Climate region'],
↳axis=1)
```

```
[150]:
```

	Station	Data since	elevation	Latitude	Longitude
0	Altdorf	01.01.1864	438.00	46.89	8.62
1	Andermatt	01.01.1864	1438.00	46.63	8.58
2	Basel / Binningen	01.01.1755	316.00	47.54	7.58
3	Bern / Zollikofen	01.01.1864	553.00	46.99	7.46
4	La Chaux-de-Fonds	01.01.1900	1017.00	47.08	6.79
5	Château-d'Oex	01.01.1879	1028.00	46.48	7.14
6	Chaumont	01.01.1864	1136.00	47.05	6.98
7	Davos	01.01.1864	1594.00	46.81	9.84
8	Elm	01.02.1878	958.00	46.92	9.18
9	Engelberg	01.01.1864	1036.00	46.82	8.41

10	Grächen	01.01.1864	1605.00	46.20	7.84
11	Grimsel Hospiz	01.01.1932	1980.00	46.57	8.33
12	Col du Grand St-Bernard	01.01.1818	2472.00	45.87	7.17
13	Genève / Cointrin	01.01.1753	411.00	46.25	6.13
14	Jungfrauojoch	01.01.1933	3571.00	46.55	7.99
15	Lugano	01.01.1864	273.00	46.00	8.96
16	Luzern	01.01.1864	454.00	47.04	8.30
17	Meiringen	01.07.1889	589.00	46.73	8.17
18	Neuchâtel	01.01.1864	485.00	47.00	6.95
19	Locarno / Monti	01.12.1882	367.00	46.17	8.79
20	Payerne	01.08.1964	490.00	46.81	6.94
21	Bad Ragaz	01.06.1870	497.00	47.02	9.50
22	Säntis	01.01.1864	2501.00	47.25	9.34
23	Samedan	01.01.1864	1709.00	46.53	9.88
24	S. Bernardino	01.01.1864	1639.00	46.46	9.18
25	Segl-Maria	01.12.1863	1804.00	46.43	9.76
26	Sion	01.01.1864	482.00	46.22	7.33
27	Zürich / Fluntern	01.01.1864	556.00	47.38	8.57
28	St. Gallen	01.01.1864	776.00	47.43	9.40

```
[151]: # Cell tagged w/ parameters
#
# Calling syntax:
# papermill Meteostat_vs_MeteoCH.ipynb foo.ipynb [-p wsno 42 -p sd 1 -p sm 2 -p sy 2022]
sd = 15
sm = 10
sy = 2023
ed = 25
em = 10
ey = 2023

wsno = -1 # select index number from above table
```

```
[152]: # Compute start and end of period of interest
start = datetime(sy, sm, sd)
end = datetime(ey, em, ed)

# Extract station name
wstation = ws.iloc[wsno]['Station']
df = ws[ws.Station==wstation]
label = df.iloc[0]['station/location'] #.to_string()[::-1][0:3][::-1]
lat = df.iloc[0]['Latitude']
lon = df.iloc[0]['Longitude']
md(f"## Compare data of weather station {wstation} w/ label {label} located at_{lat:.2f}/{lon:.2f}).")
```

[152]:

1.2 Compare data of weather station St. Gallen w/ label STG located at (47.43/9.40).

```
[153]: md("## Observations at MeteoCH for %s"%(wstation))
```

```
[153]:
```

1.3 Observations at MeteoCH for St. Gallen

```
[154]: filenm = "nbcn-daily_"
ext="csv"
# Create urls using above selected label
currurl = url + "/" + path + "/" + filenm + label + "_current." + ext
prevurl = url + "/" + path + "/" + filenm + label + "_previous." + ext
cf = pd.read_csv(currurl, sep=";", index_col='date', converters={'date':pd.
    ↳to_datetime}).drop(['station/location'], axis=1) #, engine='pyarrow')
for col in cf.columns:
    cf[col] = pd.to_numeric(cf[col], errors='coerce')
pf = pd.read_csv(prevurl, sep=";", index_col='date', converters={'date':pd.
    ↳to_datetime}).drop(['station/location'], axis=1) #, engine='pyarrow')
for col in pf.columns:
    pf[col] = pd.to_numeric(pf[col], errors='coerce')
meteochn = pd.concat([pf, cf], axis=0)
daterange = (meteochn.index>=start)&(meteochn.index<=end)
meteochn = meteochn.loc[daterange]
print(meteochn.count())
```

```
gre000d0    13
hto000d0    13
nto000d0     0
prestad0    13
rre150d0    13
sre000d0    13
tre200d0    13
tre200dn    13
tre200dx    13
ure200d0    13
dtype: int64
```

```
[155]: # Rename columns in MeteoCH to allow comparison with Meteostat
meteochn.rename(columns={'sre000d0': 'tsun', 'prestad0': 'pres', 'hto000d0': '
    ↳snow', 'rre150d0': 'prcp',
                        'tre200d0': 'tavg', 'tre200dn': 'tmin', 'tre200dx': 'tmax'},
    ↳inplace=True)
```

```
[156]: md("## Extract weather data from Meteostat using GPS coordinates of_
    ↳%s"%(wstation))
```

```
[156]:
```

1.4 Extract weather data from Meteostat using GPS coordinates of St. Gallen

```
[157]: stations = Stations()
ws = stations.nearby(lat, lon).fetch(3)
ws.drop(['region', 'icao', 'timezone', 'hourly_start', 'hourly_end', '
↪ 'daily_start', 'daily_end', 'monthly_start', 'monthly_end'], axis=1)
```

```
[157]:
```

	name	country	wmo	latitude	longitude	\
id						
06681	Saint Gallen	CH	06681	47.43	9.40	
06690	Altenrhein-Flugplatz	CH	06690	47.48	9.38	
06678	Bischofszell / Sittertal	CH	06678	47.50	9.23	

	elevation	distance
id		
06681	779.00	877.12
06690	398.00	6530.99
06678	470.00	14691.10

1.5 Observations at Meteostat

```
[158]: id = ws.index.values[0]
meteostat = Daily(id, start, end)
meteostat = meteostat.fetch()
print(meteostat.count())
```

```
tavg    13
tmin    13
tmax    13
prcp     0
snow     0
wdir    13
wspd    13
wpgt    13
pres    13
tsun     0
dtype: int64
```

1.6 Differences

```
[159]: md("### Tabular comparison for %s"%(wstation))
```

```
[159]:
```

1.6.1 Tabular comparison for St. Gallen

```
[160]: common = meteoeh.columns.intersection(meteostat.columns)
adiff = meteoeh[common] - meteostat[common]
rdiff = adiff/meteoeh[common]
```

Absolute difference

```
[161]: adiff.describe().map('{:.1f}'.format)
```

```
[161]:      snow  pres prcp tsun  tavg  tmin  tmax
count  0.0   13.0  0.0  0.0   13.0  13.0  13.0
mean   nan  -89.6  nan  nan    0.0  -0.7   1.5
std    nan   1.7  nan  nan    0.9   1.0   1.3
min    nan -92.5  nan  nan   -1.4  -2.4   0.2
25%    nan -91.0  nan  nan   -0.5  -1.1   0.6
50%    nan -89.3  nan  nan    0.0  -0.7   1.1
75%    nan -89.0  nan  nan    0.8  -0.4   1.6
max    nan -86.1  nan  nan    1.2   1.0   4.5
```

Relative difference

```
[162]: rdiff.describe().map('{:.1%}'.format)
```

```
[162]:      snow    pres  prcp  tsun    tavg    tmin    tmax
count  0.0%  1300.0%  0.0%  0.0%  1300.0%  1300.0%  1300.0%
mean   nan%   -9.7%  nan%  nan%   -1.5%  -12.9%    8.4%
std    nan%    0.2%  nan%  nan%    7.7%   16.3%    6.5%
min    nan%  -10.0%  nan%  nan%  -14.5%  -47.4%    1.4%
25%    nan%   -9.8%  nan%  nan%   -9.8%  -19.0%    3.8%
50%    nan%   -9.6%  nan%  nan%    0.0%   -8.5%    7.0%
75%    nan%   -9.5%  nan%  nan%    4.3%   -3.4%   10.9%
max    nan%   -9.5%  nan%  nan%    8.8%   10.8%   22.7%
```

```
[163]: md("### Graphical comparison for %s"%(wstation))
```

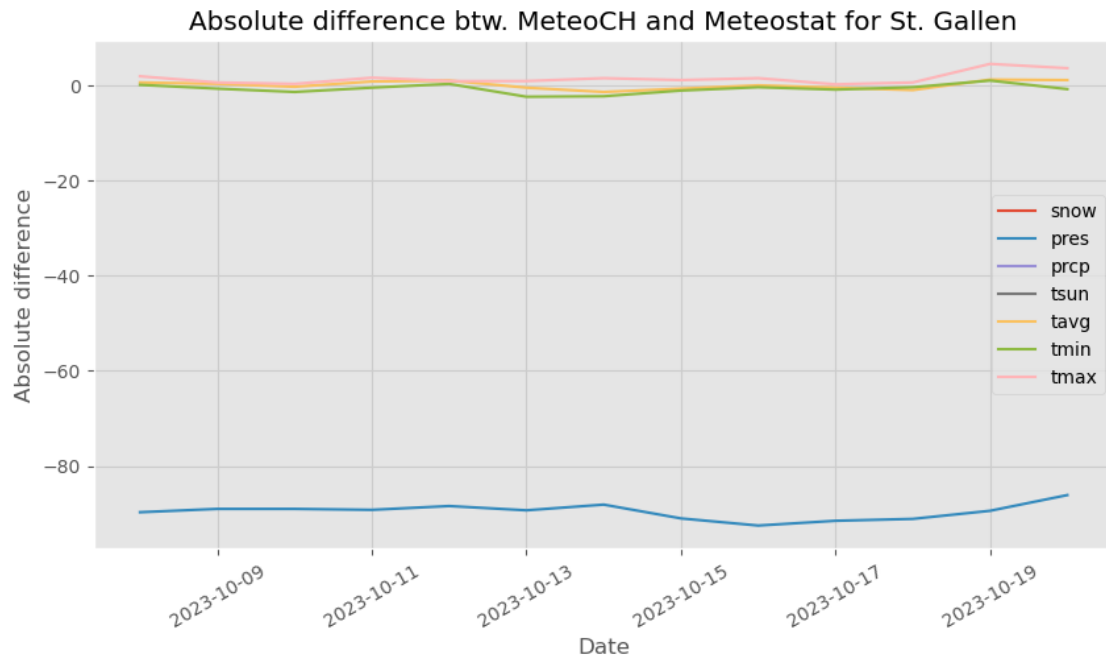
```
[163]:
1.6.2 Graphical comparison for St. Gallen
```

```
[164]: plt.style.use('ggplot')
# print(plt.style.available)
fswidth = 10
fsheight = 5
```

```
[165]: fig, axs = plt.subplots(figsize=(fswidth, fsheight))

axs.plot(adiff)
axs.grid(visible='visible', which='major', color='0.8', linestyle='--')

plt.xlabel('Date')
plt.ylabel('Absolute difference')
plt.title(f'Absolute difference btw. MeteoCH and Meteostat for {wstation}')
plt.xticks(rotation=30)
plt.legend(adiff.columns, loc="center right")
plt.show()
```



```
[166]: fig, axs = plt.subplots(figsize=(fswidth, fsheight))

axs.plot(rdifff)
axs.grid(visible='visible', which='major', color='0.8', linestyle='-')

plt.xlabel('Date')
plt.ylabel('Relative difference')
plt.title(f'Relative difference btw. MeteoCH and Meteostat for {wstation}')
plt.xticks(rotation=30)
plt.legend(rdifff.columns, loc="center right")
plt.show()
```

