Introduction to Earth System

Working with data

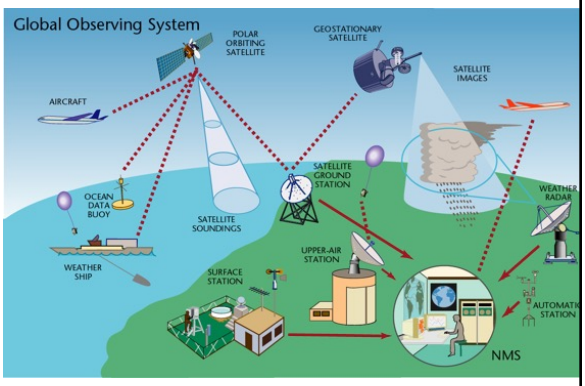
1

Data from direct Observation & Remote Sensing

Meteorological observation network

- WMO coordinate a network of 191 NHMS members
- WMO Global Observing System (GOS):
  - Surface observations
  - Upper-Air obs.
  - Marine observations
  - Aircraft-based obs.
  - Satellite observations
  - Weather radars
  - Other obs. platforms

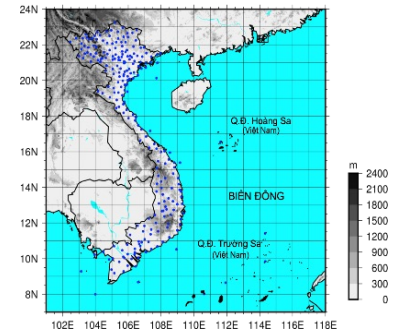
e.g. solar radiation obs., lightning detection, tide-gauge measurements, wind profilers



http://www.wmo.int

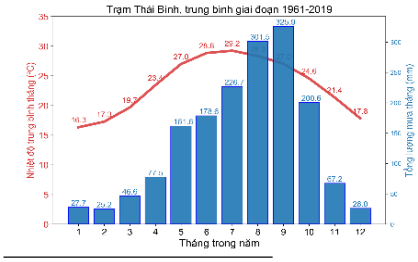
2

E.g. Station data



Location of meteorological stations in Vietnam

NhaTrang meteorological station



3

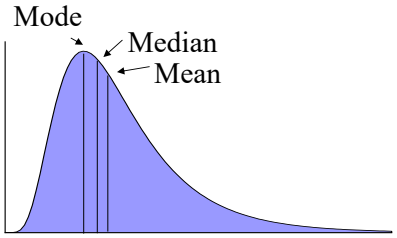
1	0.0	0.0	0.0	0.0	1.9	0.0	45.2	0.2	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.1	0.0	20.6	0.0	0.0	0.9	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	22.3	26.6	0.0	1.2	0.0	0.0
4	0.0	0.8	0.0	0.0	6.2	22.0	0.0	0.1	30.9	5.1	0.0	0.0
5	0.0	1.7	1.9	0.0	0.0	0.0	0.0	0.0	76.1	0.0	0.0	0.0
6	0.0	0.6	1.3	0.4	0.0	0.0	0.8	2.3	0.5	7.6	0.0	0.0
7	0.0	0.1	0.7	3.9	10.1	0.0	0.0	16.5	0.7	0.0	0.0	0.0
8	0.0	0.0	0.2	0.0	10.9	0.0	0.0	20.4	0.0	0.0	0.0	1.2
9	0.0	0.6	2.0	0.0	9.2	7.0	0.0	0.0	57.3	0.6	0.0	0.0
10	0.0	0.3	0.2	0.0	1.2	8.6	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	2.7	0.0	0.9	0.0	0.0	0.0	0.0	8.0	0.0	0.0
12	0.0	1.0	1.3	0.0	0.0	0.0	3.4	0.0	0.0	0.1	1.3	0.0
13	0.1	0.1	1.9	0.0	3.5	0.0	13.5	7.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0
15	0.0	0.0	0.0	0.0	0.0	5.8	0.0	14.3	0.0	0.0	0.0	0.0
16	0.0	0.1	0.0	0.0	0.0	0.0	0.0	14.8	0.0	0.0	0.0	0.0
17	0.0	1.7	0.2	0.0	0.0	13.5	14.8	53.7	0.0	0.0	0.0	0.0
18	0.0	7.2	1.0	0.0	0.0	13.9	1.9	99.3	0.0	0.0	0.0	0.0
19	0.0	0.2	3.0	6.7	0.0	0.0	0.1	11.2	0.0	0.0	0.0	0.0
20	0.0	0.0	3.1	0.0	0.0	0.0	0.1	58.1	0.0	0.0	86.0	0.0
21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	23.3	0.0
22	0.0	3.9	0.5	1.3	4.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
23	0.0	0.0	0.2	0.0	28.6	0.0	0.0	0.0	0.0	0.0	4.8	0.0
24	0.0	0.0	6.3	0.0	8.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0
25	0.0	0.3	2.0	0.0	2.2	0.0	0.1	0.0	7.2	3.6	0.0	0.0
26	0.0	0.9	0.5	0.0	6.2	0.0	0.0	3.5	10.3	0.0	0.0	0.0
27	0.0	2.4	2.1	0.0	38.3	14.2	0.0	2.8	0.1	0.0	0.0	0.0
28	0.0	3.2	0.0	3.5	1.8	0.2	8.3	0.0	0.0	0.0	0.0	0.0
29	0.0	-99.0	0.0	2.1	2.6	0.7	72.1	8.6	0.0	0.0	0.0	0.0
30	0.0	-99.0	0.0	0.0	3.2	10.9	21.3	13.3	0.0	0.0	0.8	0.0
31	0.0	-99.0	0.0	-99.0	0.1	-99.0	22.3	1.0	-99.0	0.0	-99.0	0.0

Example Hanoi station
Daily Rainfall 2012 (mm/d)

4

Basic statistics

- **Measures of Central Tendency:** mean, median, mode (single mode, bimodal, multi-modal) → the "middle region" of the sample



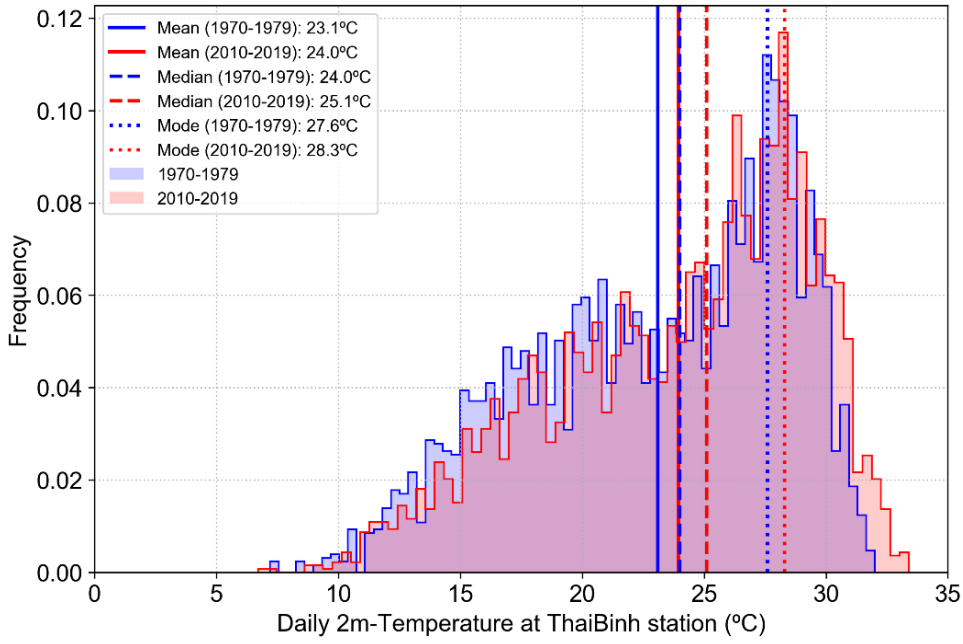
- **Measures of Dispersion: range, variance, standard deviation** → the spread or range of variability

Example

Rainfall event 1 (mm/d): 48, 49, 50, 51, 52

Rainfall event 2 (mm/d): 5, 15, 50, 80, 100

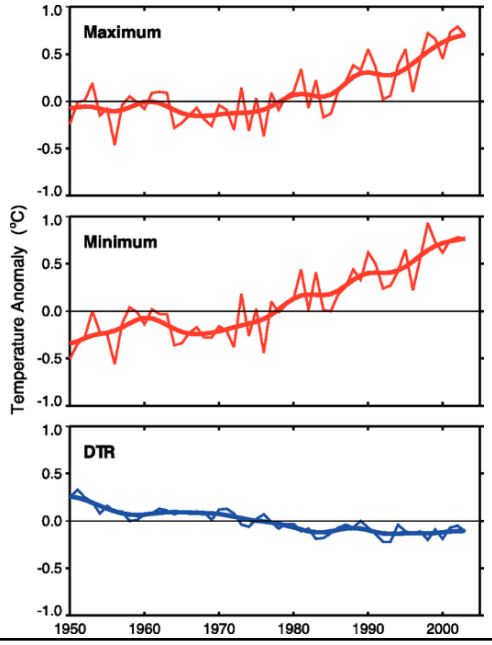
5



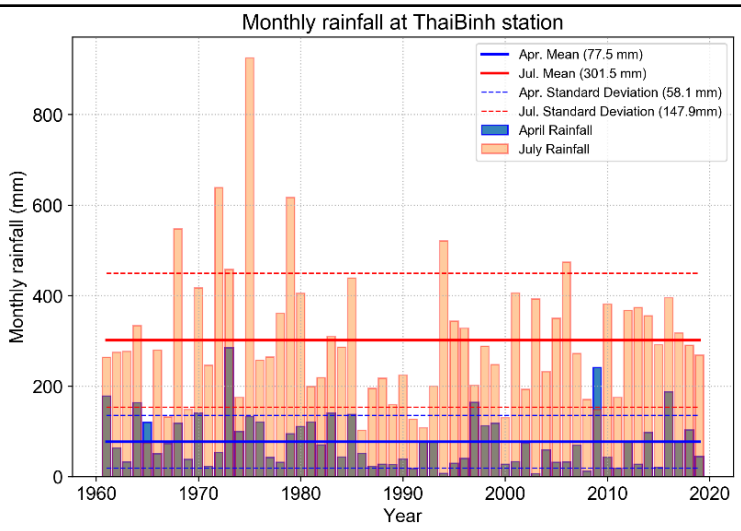
6

Diurnal temperature range (DTR)

- Tmax: 0.14˚C/decade
- Tmin: 0.2˚C/decade
- DTR: -0.07˚C/decade



7



Can standard deviation be comparable?

→ **Coefficient of Variation (CV):** $CV = \frac{s_y}{\bar{y}} \quad (\%)$

Q: CV_{Jul}? CV_{Apr}?

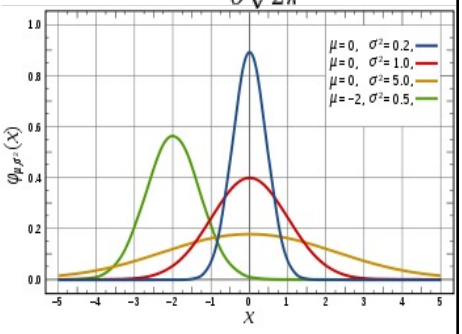
8

Normal Distribution

- When carrying out measurements → variations in the measured results.
- Results will be centred on **a mean with a degree of scatter** either side
- Many variables in nature form a bell-shaped distribution

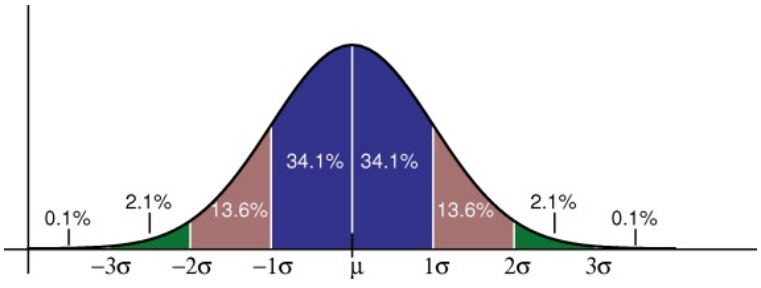
→ **normal distribution** or Gaussian curve

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



$\mu = 0$ and $\sigma^2 = 1$ → the **standard normal distribution** or the **unit normal distribution**

σ-rule



The Empirical Rule: Given a distribution of measurements drawn from a population that is approximately bell-shaped, the interval:

- $\bar{x} \pm \sigma$ contains ~68% of the measurements
- $\bar{x} \pm 2\sigma$ contains ~95% of the measurements
- $\bar{x} \pm 3\sigma$ contains ~99.7% of the measurements

Example for the empirical σ-rule

- An and Nam score 8.0 and 9.0 in Maths and Physics, respectively. Who did the best?

→ Not sure.

Now, giving the average scores of the class are 6.5 and 6.0 for Maths and Physics, respectively.

→ Both did above average, but still need more info …

We know the standard deviations:

- SD_{Maths} = 0.5
- $SD_{Physics}$ = 1.5

→ So An's score is 3 SDs above the mean (8.0-6.5)/0.5

– Only 0.15% of the students could be better

- For Nam, his score is 2 SDs above the mean (9.0-6.0)/1.5

– 2.5% of the students could be better

Exercise #1

- Annual average of daily minimum temperature (Tmin) in Hanoi is 22ºC. The std of Tmin is 5ºC.
- Annual average of daily minimum temperature (Tmin) in Tokyo is 8ºC. The std of Tmin is 5ºC.

Question

- In a cold event in 2015, the lowest Tmin values in Hanoi and in Tokyo were 7ºC and -5ºC, respectively. Which value is more extreme?

Exercise #2

- A meteorologist observer wrote the observed temperature value in the notebook. However, instead of writing 28.5ºC, he mistakenly wrote 285ºC. How to point out this error without doing manual check?

- How about the case of rainfall data?

→ Before assessing climate change signals
❑ Data collection
❑ Data quality control

Practice Python #9: Station meteorological data

1	0.0	0.0	0.0	0.0	1.9	0.0	45.2	0.2	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.1	0.0	20.6	0.0	0.0	0.9	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	22.3	26.6	0.0	1.2	0.0	0.0
4	0.0	0.8	0.0	0.0	6.2	22.0	0.0	0.1	30.9	5.1	0.0	0.0
5	0.0	1.7	1.9	0.0	0.0	0.0	0.0	0.0	76.1	0.0	0.0	0.0
6	0.0	0.6	1.3	0.4	0.0	0.0	0.8	2.3	0.5	7.6	0.0	0.0
7	0.0	0.1	0.7	3.9	10.1	0.0	0.0	16.5	0.7	0.0	0.0	0.0
8	0.0	0.0	0.2	0.0	10.9	0.0	0.0	20.4	0.0	0.0	0.0	1.2
9	0.0	0.6	2.0	0.0	9.2	7.0	0.0	0.0	57.3	0.6	0.0	0.0
10	0.0	0.3	0.2	0.0	1.2	8.6	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	2.7	0.0	0.9	0.0	0.0	0.0	0.0	8.0	0.0	0.0
12	0.0	1.0	1.3	0.0	0.0	0.0	3.4	0.0	0.0	0.1	1.3	0.0
13	0.1	0.1	1.9	0.0	3.5	0.0	13.5	7.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0
15	0.0	0.0	0.0	0.0	0.0	5.8	0.0	14.3	0.0	0.0	0.0	0.0
16	0.0	0.1	0.0	0.0	0.0	0.0	0.0	14.8	0.0	0.0	0.0	0.0
17	0.0	1.7	0.2	0.0	0.0	13.5	14.8	53.7	0.0	0.0	0.0	0.0
18	0.0	7.2	1.0	0.0	0.0	13.9	1.9	99.3	0.0	0.0	0.0	0.0
19	0.0	0.2	3.0	6.7	0.0	0.0	0.1	11.2	0.0	0.0	0.0	0.0
20	0.0	0.0	3.1	0.0	0.0	0.0	0.1	58.1	0.0	0.0	86.0	0.0
21	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	23.3	0.0
22	0.0	3.9	0.5	1.3	4.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
23	0.0	0.0	0.2	0.0	28.6	0.0	0.0	0.0	0.0	0.0	4.8	0.0
24	0.0	0.0	6.3	0.0	8.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0
25	0.0	0.3	2.0	0.0	2.2	0.0	0.1	0.0	7.2	3.6	0.0	0.0
26	0.0	0.9	0.5	0.0	6.2	0.0	0.0	3.5	10.3	0.0	0.0	0.0
27	0.0	2.4	2.1	0.0	38.3	14.2	0.0	2.8	0.1	0.0	0.0	0.0
28	0.0	3.2	0.0	3.5	1.8	0.2	8.3	0.0	0.0	0.0	0.0	0.0
29	0.0	-99.0	0.0	2.1	2.6	0.7	72.1	8.6	0.0	0.0	0.0	0.0
30	0.0	-99.0	0.0	0.0	3.2	10.9	21.3	13.3	0.0	0.0	0.8	0.0
31	0.0	-99.0	0.0	-99.0	0.1	-99.0	22.3	1.0	-99.0	0.0	-99.0	0.0

Hanoi station
Daily Rainfall 2012
(mm/d)

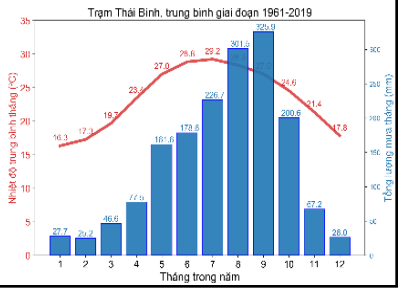
Practice Python #9

I. The daily 2m-temperature data at VIETTRI station can be downloaded at:
http://remosat.usth.edu.vn/~thanhnd/Download/T2m_VIETTRI_1961_2019.txt

1. Write a python program to read the daily temperature data at VIETTRI station
2. Estimate monthly values for each year
3. Plot monthly-mean climatology at VIETTRI station

II. Download the daily rainfall data at VIETTRI station:
http://remosat.usth.edu.vn/~thanhnd/Download/R_VIETTRI_1961_2019.txt

4. Repeat the above tasks for rainfall



15

```python
# To Read daily T2mfall data at ThaiBinh station
#*********************************************
import numpy as np
import matplotlib.pyplot as plt
# To read the T2m data from the input file
r2T2m=np.loadtxt("T2m_THAIBINH_1961_2019.txt",comments="#")
nyr=len(r2T2m[:,1])/31     # to count the number of the years
nyr=int(nyr)
print ("number of the years", nyr)
# reshape to a 3D array: r3T2m[year,dayofmonth,month]
r3T2m=np.reshape(r2T2m[:,1:13],(nyr,31,12))
#*********************************************
#1. Estimate monthly average values for each year
r2T2m_mn=np.zeros(shape=(nyr,12))
for iyr in range(0,nyr):
    for imn in range(0,12):
        ndy=0
        for idy in range(0,31):
            if r3T2m[iyr,idy,imn]>=-10.:   # to avoid missing data -99.
                ndy=ndy+1
                r2T2m_mn[iyr,imn]=r2T2m_mn[iyr,imn]+r3T2m[iyr,idy,imn]
        r2T2m_mn[iyr,imn]=round(r2T2m_mn[iyr,imn]/float(ndy),2)
    print ("Year:",iyr+1961,"  ",r2T2m_mn[iyr])
#*********************************************************************
# 2. plot monthly mean climatology at ThaiBinh station
# from the above code --> we have obtained monthly values for each year;
# now, we will compute the monthly mean climatology
r1T2m_clim=np.zeros(12)
r1T2m_clim=np.mean(r2T2m_mn,axis=0)
# To plot Figure
months=np.arange(1,12.5,1)
plt.plot(months,r1T2m_clim,'blue',linewidth=3)
plt.xlabel('Month')
plt.ylabel('Average monthly T2m (deg.C)')
plt.title('ThaiBinh 2m-temperature Climatology 1961-2014')
plt.savefig("fig_ThaiBinh_T2m_Clim.png")
```


16