Statistics with Meteorological Applications
Homework II

Group 6 林之然 B11209005 蔡知諺 B11209014

March 2, 2024

Abstract

In this homework we dealt with a set of historical Surface Temperature T_s data at Taipei (46692) and Taitung (46766) stations. We mainly focus on the change of T_s diurnal cycle in July between 1960's and 2000's.

1 Introduction

1.1 Data Structure

According to Figure 1 from given folder, the surface temperature T_s data at Taipei (46692) and Taitung (46766) for 1960's and 2000's is stored in 4 24*3720 sized .txt files. The 0th dimension is the 24 hours of each day and the 1st dimension is the day within 10 years where each month is expanded to 31 days (the excess days are filled by NaN).

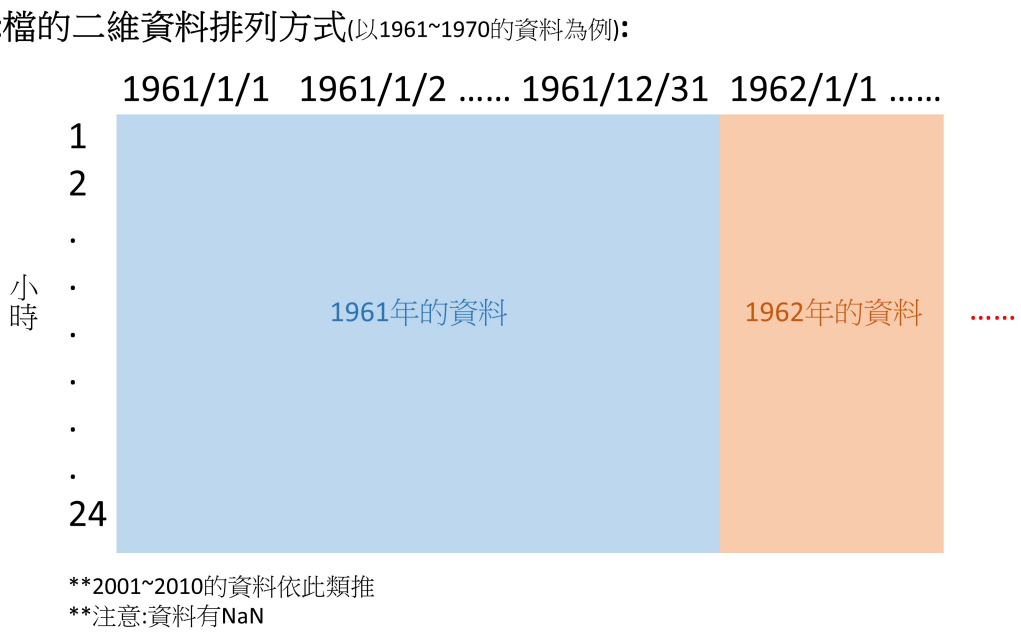

Figure 1: The data structure

1.2 Data Processing

To focus on July each year, we use some loops to read the specific regions of data and concatenate them together, so the 4 files becomes 4 arrays of T_s in July within 10 years. The code for data processing plotting is in the attached file "hw2.ipynb".

1.3 Problem Encountered

It's necessary to mention a problem we encountered during data processing. The problem is that the T_s data at Taitung, July 1961 is entirely empty (filled by NaN), meaning that the subsequent statistics at Taitung for 1960's base on only 9 years of data.

2 Output Data and Visualization (if needed)

2.1 Mean T_s for 1960's and 2000's

Table 1: Mean Surface Temperature T_s (°C)

Station	1960's	2000's
Taipei	28.71	30.00
Taitung	28.54	29.06

2.2 Mean Maximum T_s and Mean Minimum T_s for 1960's and 2000's

Mean Maximum and Mean Minimum Temperature sound confusing, but its simply the mean of daily Maximum (31*10 numbers for each case) or Minimum. Table 2 gives us a concept about "normally how hot/cold it could be in a day in July within the decade".

Table 2: Mean Maximum and Mean Minimum T_s (°C)

Station		1960's	2000's
Taipei	Max.	33.46	34.15
	Min.	25.12	26.95
Taitung	Max.	31.94	32.17
	Min.	25.26	26.33

2.3 Mean DTR (Diurnal Temperature Range) and Domain Variance s^2

- The diurnal temperature range is defined as:

$$\text{DTR} = \text{Max.}(T) - \text{Min.}(T)$$

  so we calculate daily DTR by subtracting daily maximum with daily minimum and find the decade mean to be mean DTR.

- According to the updated explanation of the question by the TAs, the Variance should base on the hourly data in our domain (each July within 1960's), so to be clear, let's call it "Domain" Variance.

Table 3: Mean DTR and Domain Variance of T_s at two stations

Station		1960's	2000's
Taipei	DTR (°C)	8.34	7.20
	Var. ($°C^2$)	9.53	6.89
Taitung	DTR (°C)	6.67	5.85
	Var. ($°C^2$)	6.92	5.22

Mean of Daily Variance This is what we did before the command was updated, and we consider this still useful to evaluate the temperature variability within a normal day in July, so Table 4 is left here as a reference.

Table 4: Mean of Daily Variance of T_s at two stations

Station		1960's	2000's
Teipei	Mean Daily Var. ($^\circ C^2$)	8.62	5.58
Taitung	Mean Daily Var. ($^\circ C^2$)	5.60	3.74

2.3.1 Discussion

As we take a deeper look into the variability of the data. We observe that a larger DTR is associated with a larger Variance. It makes much sense because both measurements give us an aspect of how closely the data is distributed, and the sinusoidal characteristic of temperature curve states further support the relevance (see Figure 2). However, it's important to note that the two aspects are not equivalent in both statistical meaning and how we feel.

From a statistical perspective, the "range" of data is simply the sample maximum minus sample minimum, which can be easily influenced by outliers and therefore reduce its reliability. Meanwhile, the use of variance may provide a more accurate statistical result.

Interestingly, from the perspective of our daily life, the DTR is apparently more useful as it indicates the temperature difference between day and night, helping us whether to bring an extra jacket in our bag. In contrast, the concept of variance is less intuitive. While We can definitely find the correlation between variance and the speed of heating/cooling, in most cases, variance can be replaced by other more intuitive measures. In conclusion, we think that both measurements are important, with DTR being more common and variance being more statistically meaningful.

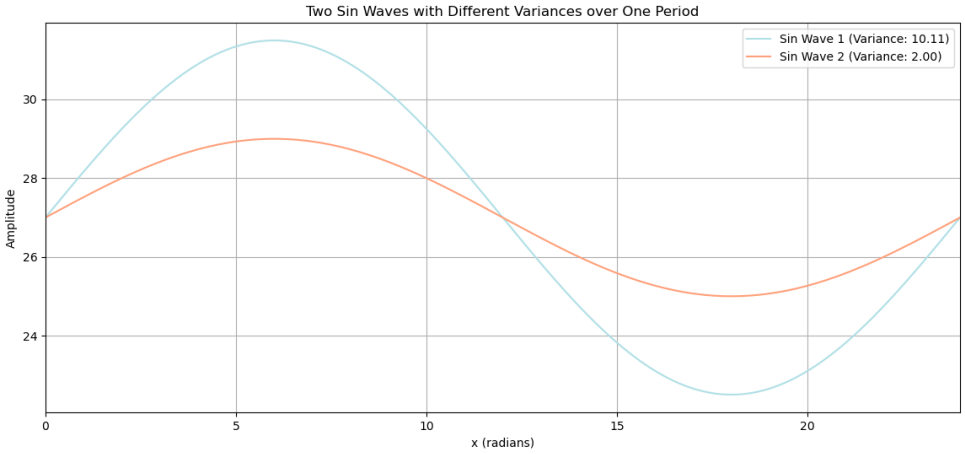

Figure 2: The Relationship Between Range (Amplitude) and Variance in Sinusoidal Waves

2.4 Average Diurnal Cycle for 1960's and 2000's

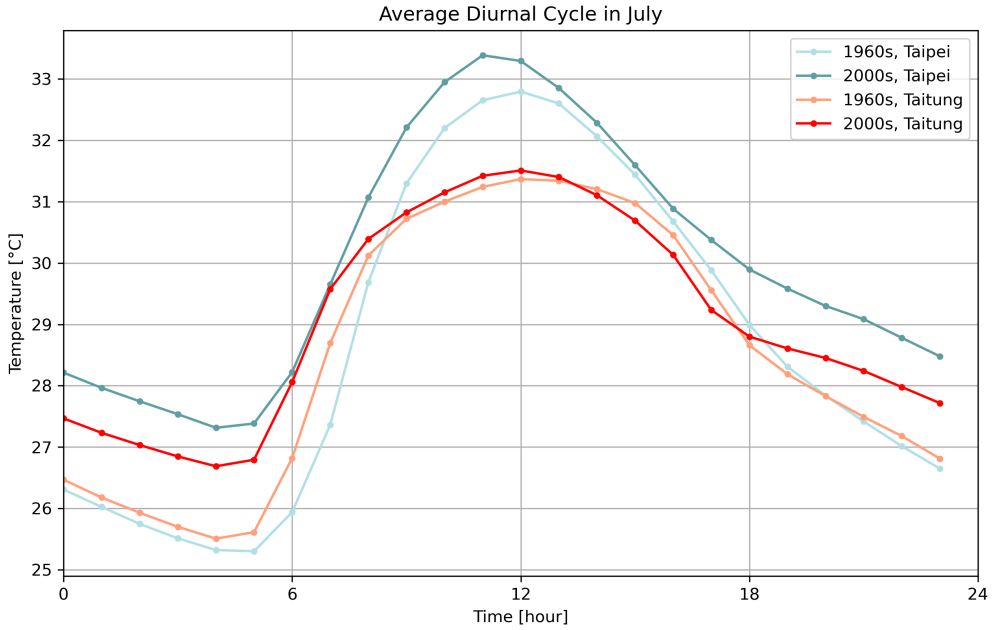

Figure 3: Average Diurnal Cycle in July for 1960's and 2000's at Taipei/Taitung

2.4.1 Discussion

In the diurnal cycle of Figure 3, we can identify 3 major differences between Taipei and Taitung as well as 2 significant changes over time.

Major Differences Between Stations

1. **Curve shapes**: As the four curves exhibit a pattern similar to a typical temperature curve in Taiwan, it's notable that the curves for Taipei show a larger Diurnal Temperature Range (DTR) and significantly higher maximum T_s compared to Taitung, despite Taitung being located further south. At first glance, we associates it with the famous "urban heat island" (UHI) effect, which will be discussed further in the following points.

2. **Higher maximum T_s**: The most significant difference is the high temperature of Taipei at day regardless of which era it's in. This difference may be attributed to variations in land use and subsequent lower albedo, resulting in higher temperatures at noon. These factors are highly related to the UHI effect.

3. **Faster heating/cooling in the morning/evening**: The rate heating and cooling in Taipei appears to be more rapid. In addition to albedo, this phenomenon may also be linked to the higher heat capacity of land in Taipei, as Taitung's proximity to the sea allows more effective temperature regulation.

Significant Changes Over Time

1. **T_s shift at night**: The T_s difference between two stations at night has not been discussed yet because the difference between eras is more dominant. We suggest that another crucial part of UHI plays an important role in the shift. Tall buildings have increased in the late 20th century, which constitute an obstruction to radiative cooling. And the growing urban area has removed vegetation that previously contributed much to the nighttime cooling through evapotranspiration.

2. **The effects of global warming**: We all know that the earth had warmed at least more than 0.5 °C from 1960 to 2010. The effect is not very significant in the diurnal cycle as the regional effect is associated with much internal factors, but the statistics has certainly showed the clear signal of warming. The mean T_s has increased from 28.71 to 30.00 in Taipei and from 28.54 to 29.06 in Taitung.

2.5 Summation

From sections 2.1 to 2.3, we've analyzed some statistics of the dataset and have briefly discussed the relationship and differences between DTR and Variance. In section 2.4, we plotted the diurnal cycle and had been able to discuss the change of those statistics in a science way. While we've already been accustomed to the scientific and plotting approach for years, both the scientific and statistical way are important. The former helps us construct the model, while the latter reveals hidden truths.