Analyzing Global Warming through GHCN-Daily Data

Jinchen Gong, Lanxi Zhang, Yaling Hong, Shuangyu Wang

December 7, 2023

- Current situation: global warming
- Significant impacts of global warming
- Global response and policies

Our Question

How does historical climate data from GHCN-Daily contribute to our understanding and analysis of **global warming trends**?

- GHCN (Global Historical Climatology Network)-Daily is a comprehensive database providing essential historical records of daily **temperature**, **precipitation**, and **snow** over global land areas. It combines climate data from various sources, all undergoing rigorous quality checks.
- ghcnd-inventory is a txt file containing the information of stations, latitude, longitude and recorded variables together with record time range.

GHCN-daily encompasses basic geographic information on meteorological stations over 40 meteorological elements, the following variables are our main focus.

Table: Variables

Variable	Description	Units
PRCP	Precipitation	Inches
TMAX	Maximum temperature	°F to tenths
TMIN	Minimum temperature	°F to tenths

We select stations which have records of the above three variables based on the information provided by ghcnd-inventory for further analysis.

Comparison of global weather station data

- Average annual precipitation - the trend cannot be accurately described.
- Average annual temperature - there may be an upward trend in temperature over the period 2000-2023.

Mann-Kendall Test

H_0: There isn't an increasing trend in the time sequence.
H_1: There is an increasing trend in the time sequence

- $S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(X_j - X_k)$
- $Var(S) = \frac{1}{18}[n(n-1)(2n+5) - \sum_{k=1}^{q} q_k(q_k-1)(2q_k+5)]$
- $Z_{MK} = \frac{S-1}{Var(S)}\mathbf{I}(S > 0) + \frac{S+1}{Var(S)}\mathbf{I}(S < 0)$
- Reject H_0 if $Z_{MK} \geq U_{1-\alpha}$

Two-sample Kolmogorov-Smirnov Test

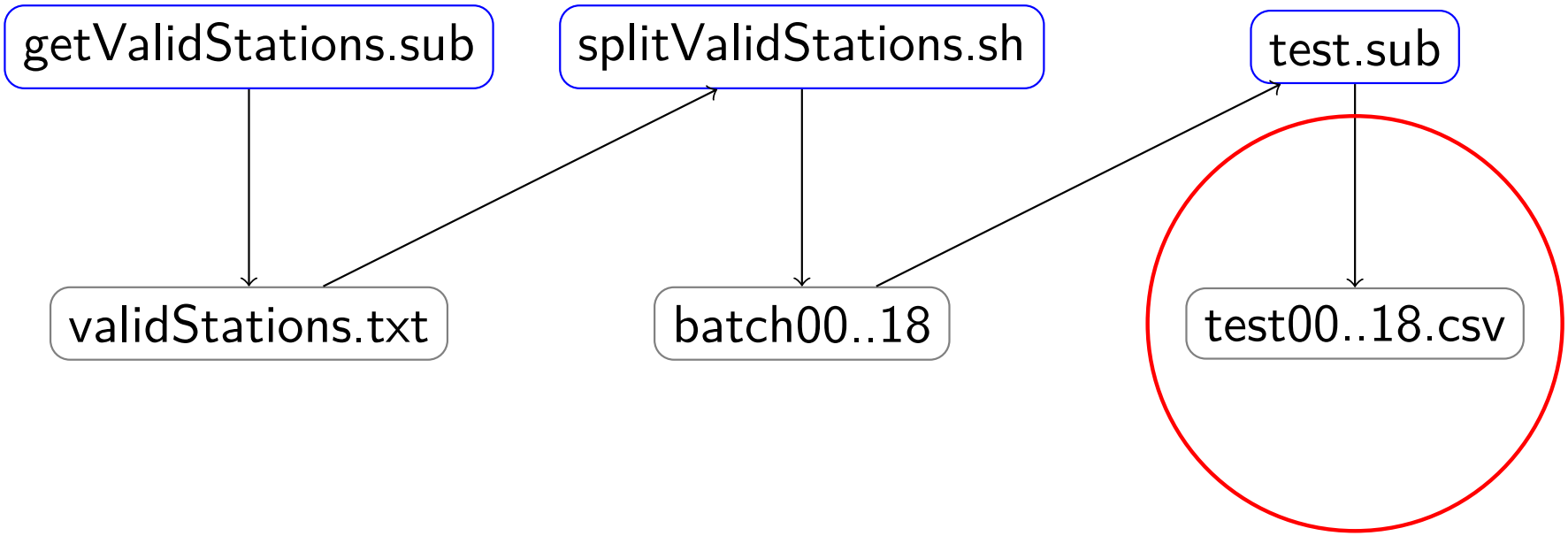
$X_1, ..., X_n \sim F_1$ and $Y_1, ..., Y_m \sim F_2$
$H_0 : F_1 = F_2 \leftrightarrow H_1 : F_1 \neq F_2$

- $\tilde{D}_{n,m} = sup_x |F_{1,n}(x) - F_{2,m}(x)|$
- Reject H_0 if $\tilde{D}_{n,m} > c(\alpha)\sqrt{\frac{n+m}{nm}}$

Batch Change Detection

$H_0 : X_i \sim F_0$
$H_1 : X_i \sim F_0, i \leq k$ and $X_i \sim F_1, i > k$

- $D_n = max_k D_{k,n-k}, k = 2, ..., n-1$
- Reject H_0 if $D_n > h_n$
- Change point location: $\hat{\tau} = argmax_k D_{k,n-k}$



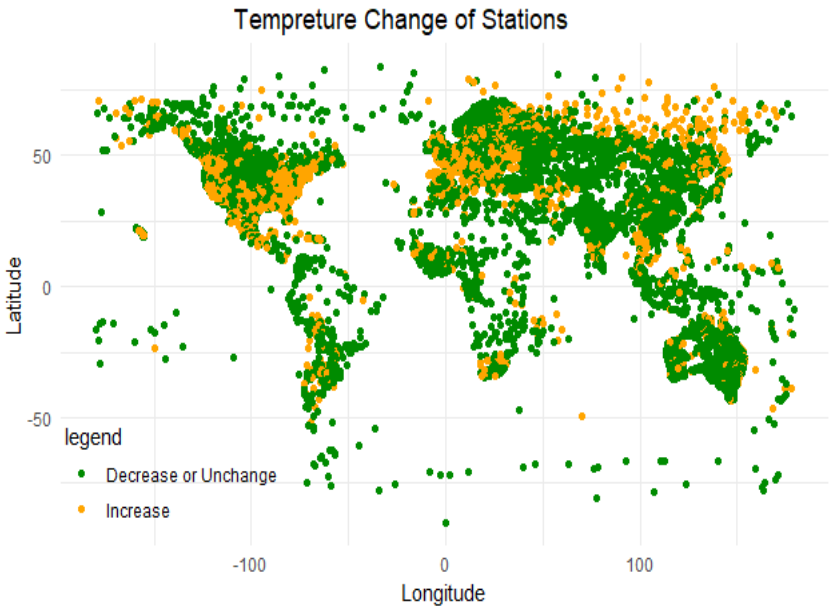
We get 19 output files containing p.values of MK test and change points of CPM test for valid stations.

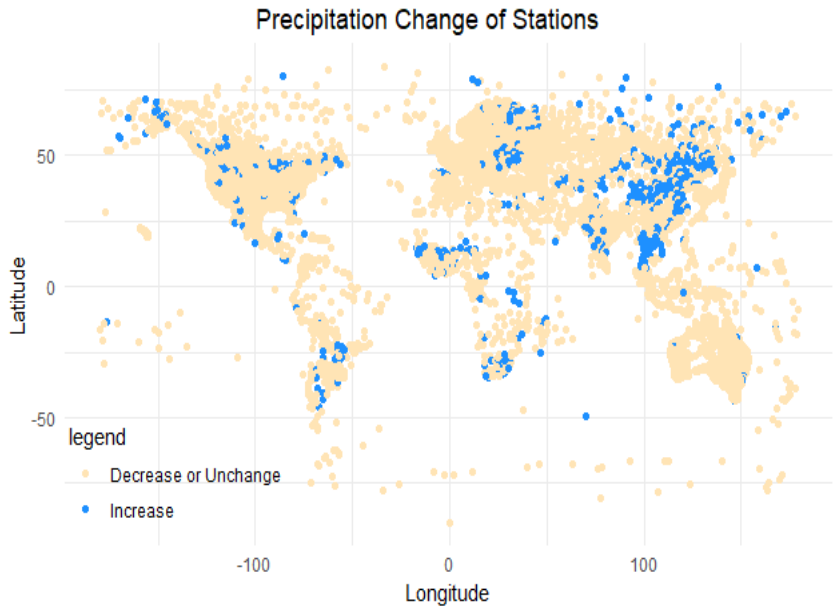
test.dag

```
job 1 getValidStations.sub
job 2 test.sub
script post 1 splitValidStations.sh
parent 1 child 2
```

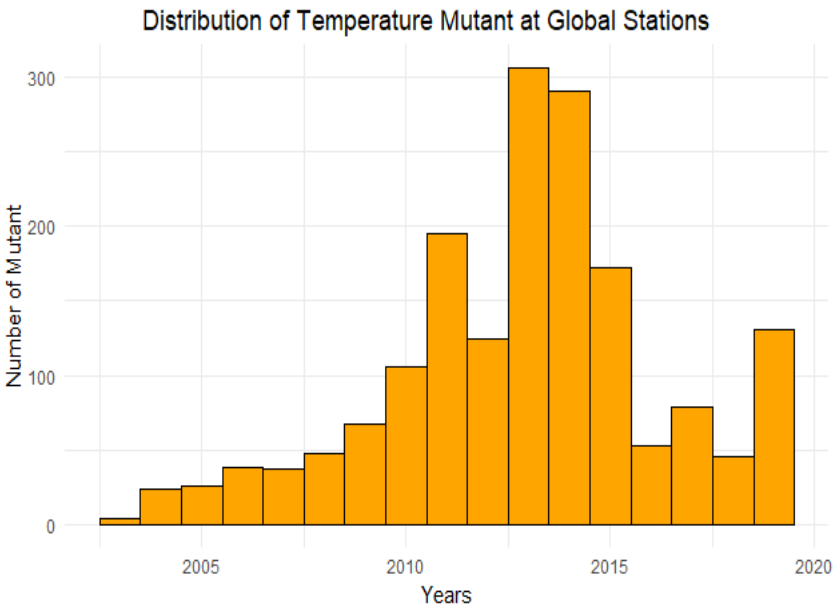
Table: Jobs Information

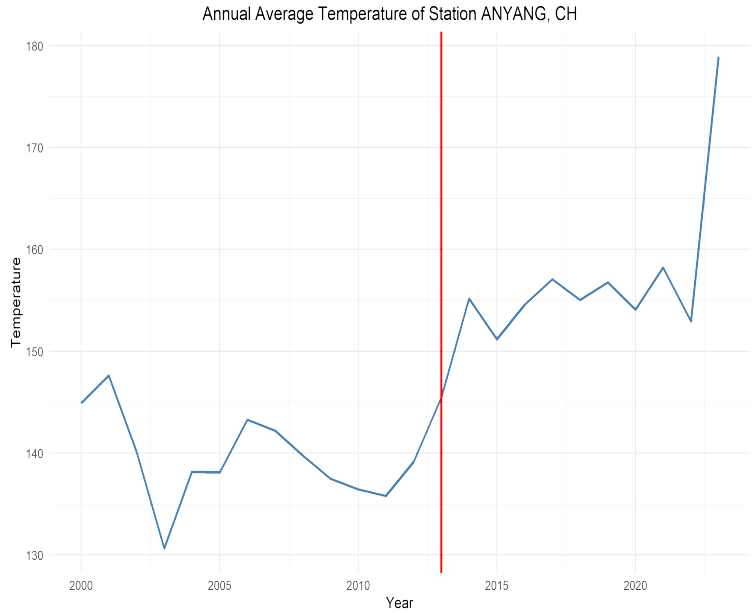
JobName	JobNumbers	Time	Memory	Disk
getValidStations	1	10min	1GB	20MB
test	19	40min	1GB	1GB



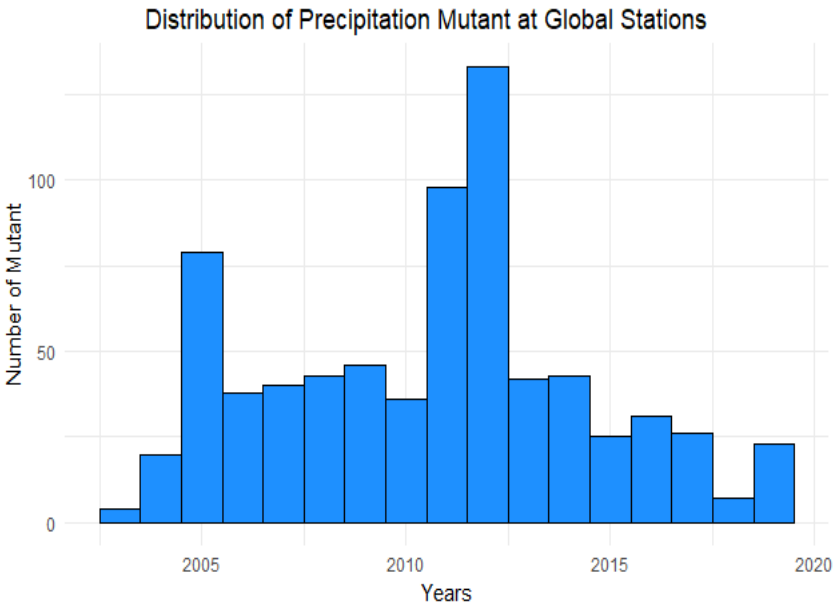


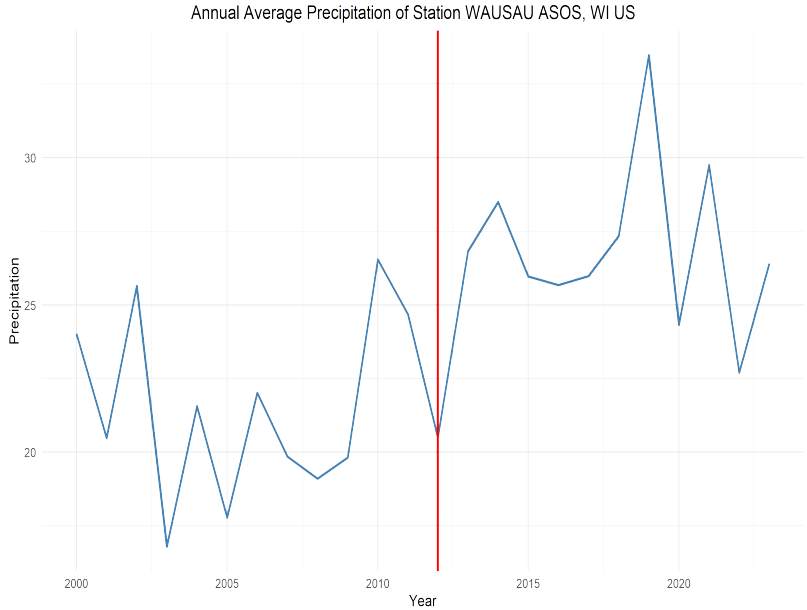
3,231 stations experienced a significant increase in annual average temperature, and 1,025 stations experienced a significant increase in annual average precipitation.





The mutant years of annual average temperature are concentrated in 2013 and 2014. Areas with abnormally high temperatures included most of Eurasia.





The mutant years of annual average precipitation are concentrated in 2011 and 2012. In 2012, the United States experienced its worst drought since 2000.

Drawbacks

1. Limitations of Mann-Kendall Test
   - Cannot provide a quantitative measure or rate of change.
   - Lack of sensitivity for small datasets.
2. Limited variable selection
   - Our research didn't inculde other key variables such as snowfall and cloud cover.
3. Constrained Time Range
   - The time span from 2000 to 2023 is too brief to fully represent long-term climate trends.
4. Lack of Seasonal Analysis

Future Work

1. Incorporating more variables.
2. Expanding the time frame.
3. Considering seasonal analysis.
4. Increasing sensitivity for small datasets.

Thank you !