

Instrumental Temperature Series in Eastern and Central China Back to the 19th Century

**Lijuan Cao, Ping Zhao, Zhongwei Yan,
Phil Jones et al.**

National Meteorological Information center, CMA

Budapest 2014.05.13



国家气象信息中心
National Meteorological Information Center

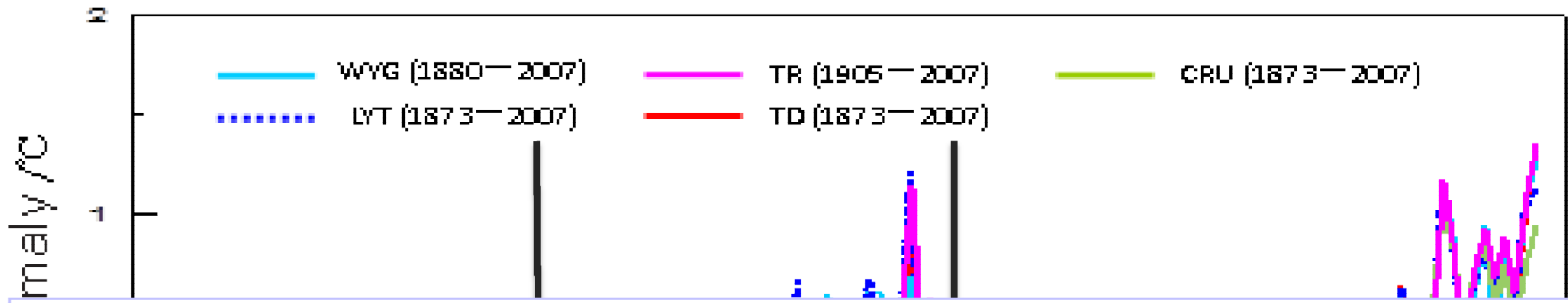


Outline

- **Background**
- Sources and quality control
- Interpolation of missing value
- Homogeneity test
- Summary

Background

- The aim is to construct a new set of long-term instrumental homogenized monthly mean temperature series
- 1995, LYT (1873-1995)(Lin et al., 1995)
- 1998, WYG-10 regions over China(1880-1998) using proxy data (Wang et al., 1998)
- 2005, 2006, TR(1905-2005) and TD(1873-2006) (Tang et al., 2005)
- 2010, LQX(1873-2004)(Li et al., 2010)



| 序列名称 | 1906 — 2005 年 | 1908 — 2007 年 |
|------|---------------|---------------|
| WYG | 0.53 | 0.59 |
| LYT | 0.34 | 0.42 |
| TR | 0.95 | 1.11 |
| TD | 0.86 | 0.96 |
| CRU | 1.20 | 1.27 |

Anomaly series of China (based on 1971–2000)
 (Tang et al., 2009)



Outline

- Background
- **Sources and quality control**
- Interpolation of missing value
- Homogeneity test
- Summary

黑井测候所气象月报表 中華民國三十七年十月份

站址: 25°30' N, 101°50' E, Altitude: 1,250 M.

| 日期 | 气压 | | | 温度 | | | 湿度 | | | 风 | | | 云 | | | 降水 | | 其他 |
|----|------|-----|-----|------|------|------|------|------|----|------|-----|-----|-----|-----|-----|----|---|----|
| | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | |
| 1 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 2 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 3 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 4 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 5 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 6 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |

海日测候所气象月报表 中華民國三十七年十月份

站址: 25°30' N, 101°50' E, Altitude: 1,250 M.

| 日期 | 气压 | | | 温度 | | | 湿度 | | | 风 | | | 云 | | | 降水 | | 其他 |
|----|------|-----|-----|------|------|------|------|------|----|------|-----|-----|-----|-----|-----|----|---|----|
| | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | h | |
| 1 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 2 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 3 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 4 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 5 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |
| 6 | 10.5 | 2.0 | 2.0 | 22.0 | 21.2 | 21.2 | 60.5 | 20.5 | 18 | 14.5 | 2.0 | 2.0 | 1.0 | 1.0 | 1.0 | 0 | 0 | |

中華民國 29 年 1 月 1-6 日昆明气象观测日记

| 日期 | 1月1日 | 1月2日 | 1月3日 | 1月4日 | 1月5日 | 1月6日 |
|------|-------|-------|-------|-------|-------|-------|
| 气压 | 10.0 | 10.0 | 10.0 | 10.4 | 12.0 | 12.0 |
| 最高温度 | 6.8 | 6.2 | 3.6 | 4.4 | 4.6 | 4.4 |
| 最低温度 | 6.8 | 6.2 | 3.6 | 4.4 | 4.6 | 4.4 |
| 湿度 | 62.88 | 62.82 | 72.82 | 72.82 | 69.97 | 62.82 |
| 风速 | 12.2 | 8.5 | 12.5 | 18.3 | 15.0 | 16.4 |
| 风向 | 2.0 | 2.3 | -0.3 | 0.0 | -0.2 | 0.7 |
| 降水 | 6.0 | 4.0 | 1.0 | 1.0 | 0.6 | 0.0 |
| 日照 | 3.9 | 3.5 | 3.2 | 3.2 | 3.2 | 3.2 |
| 雾 | 0 | 0 | 0 | 0 | 0 | 0 |
| 能见度 | 6.05 | 4.57 | 4.58 | 4.30 | 4.48 | 4.58 |
| 其他 | 9.2.0 | 8.9.0 | 9.2.0 | 9.6.0 | 8.8.0 | 8.7.5 |

Digitalized of paper record in early stage

Id: 10.3334/CDIAC/cli.ndp.039

ORNL/CDIAC-47 NDP-039

Two Long-Term Instrumental Climatic Data Bases of the People's Republic of China

Contributed by
Tao Shiyan, Fu Congbin, Zeng Zhaomei, Zhang Qingyun
Institute of Atmospheric Physics
Chinese Academy of Sciences
Beijing, China

Prepared by D. P. Kaiser
Carbon Dioxide Information Analysis Center
Oak Ridge National Laboratory
Oak Ridge, Tennessee, U.S.A.

November 1991



中华人民共和国的两个长期性仪器记录的气候数据库

中国北京
中国科学院大气物理研究所
陶诗言, 符淙斌, 曾昭美, 张庆云供稿

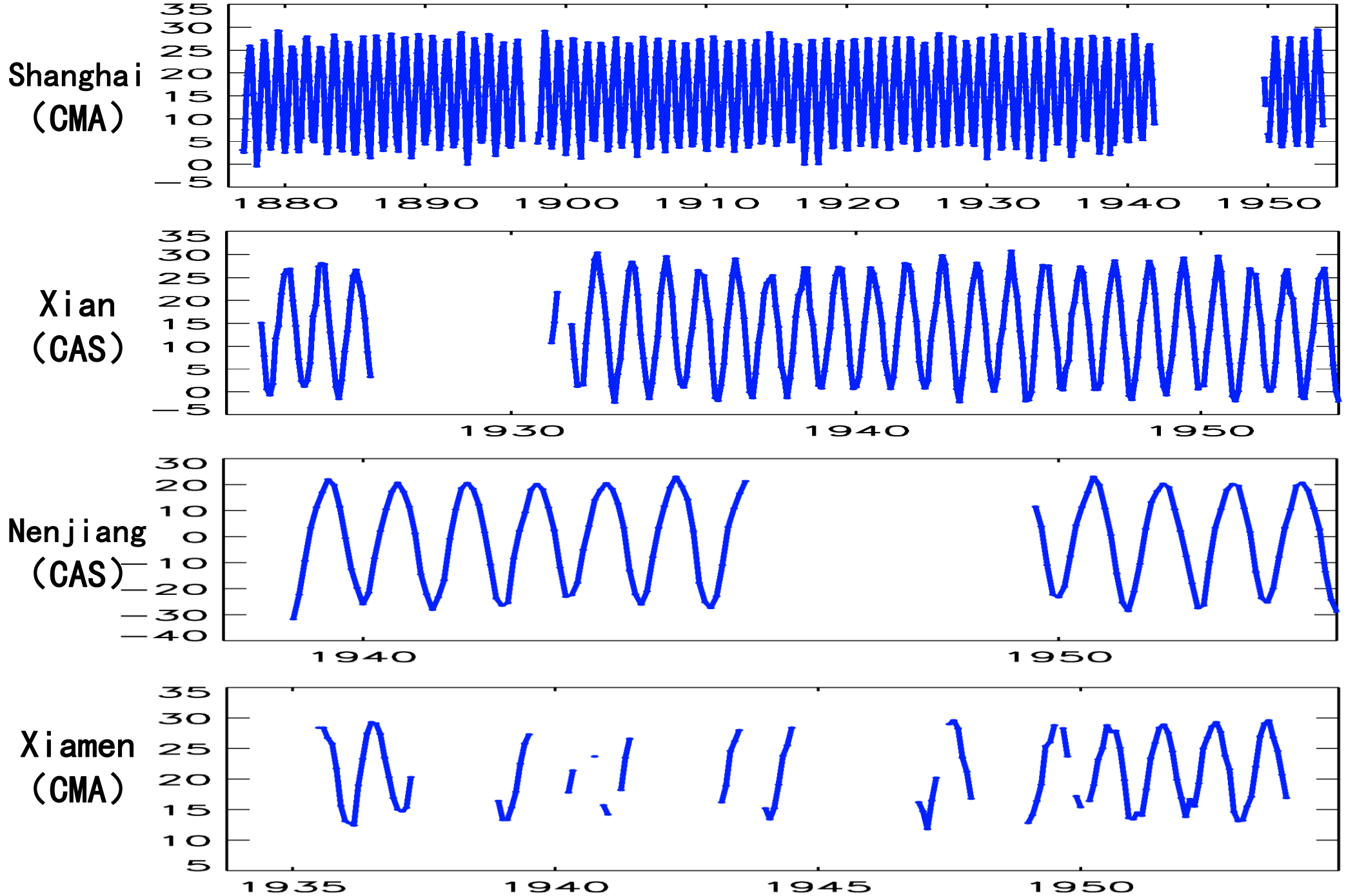
环境科学处 D. P. Kaiser 编写
为 Carbon Dioxide Research Program
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6335 编写

发表日期: 1991年十一月

Problems in long-term series

- Various observation time(Beijing/local time, 2\3\5\6\7\8\10\12\18\19\20\21\24 times a day)
- Missing data in 1940s-1950s(War)
- Relocation of meteorological stations

Monthly mean temperature °C





Two criteria of compiling climatic series from different data sources :

1. Altitude of the observation site---close to

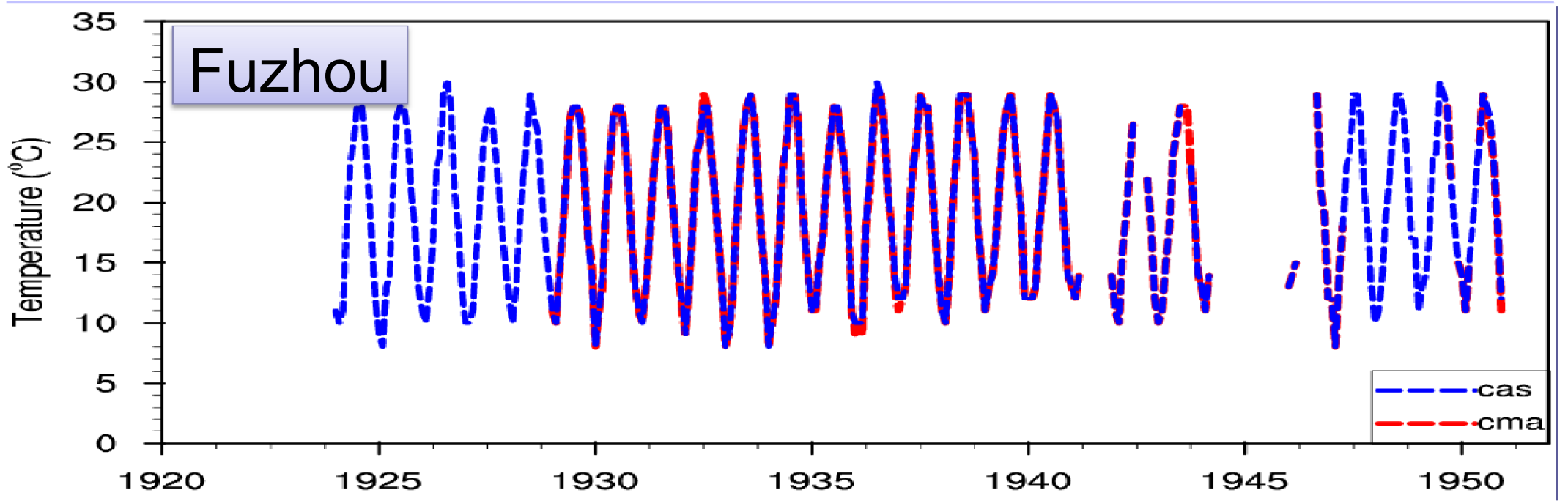
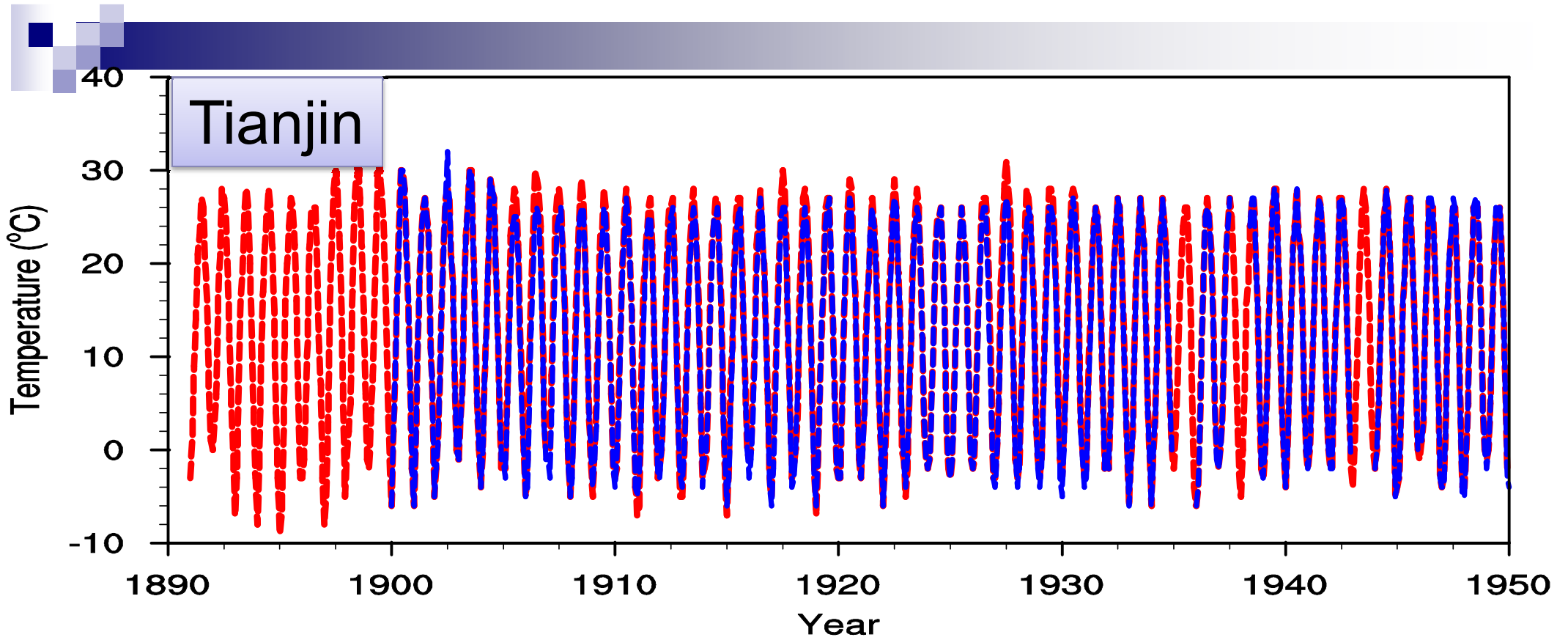
2. Horizontal distance---Close to

Additionally:

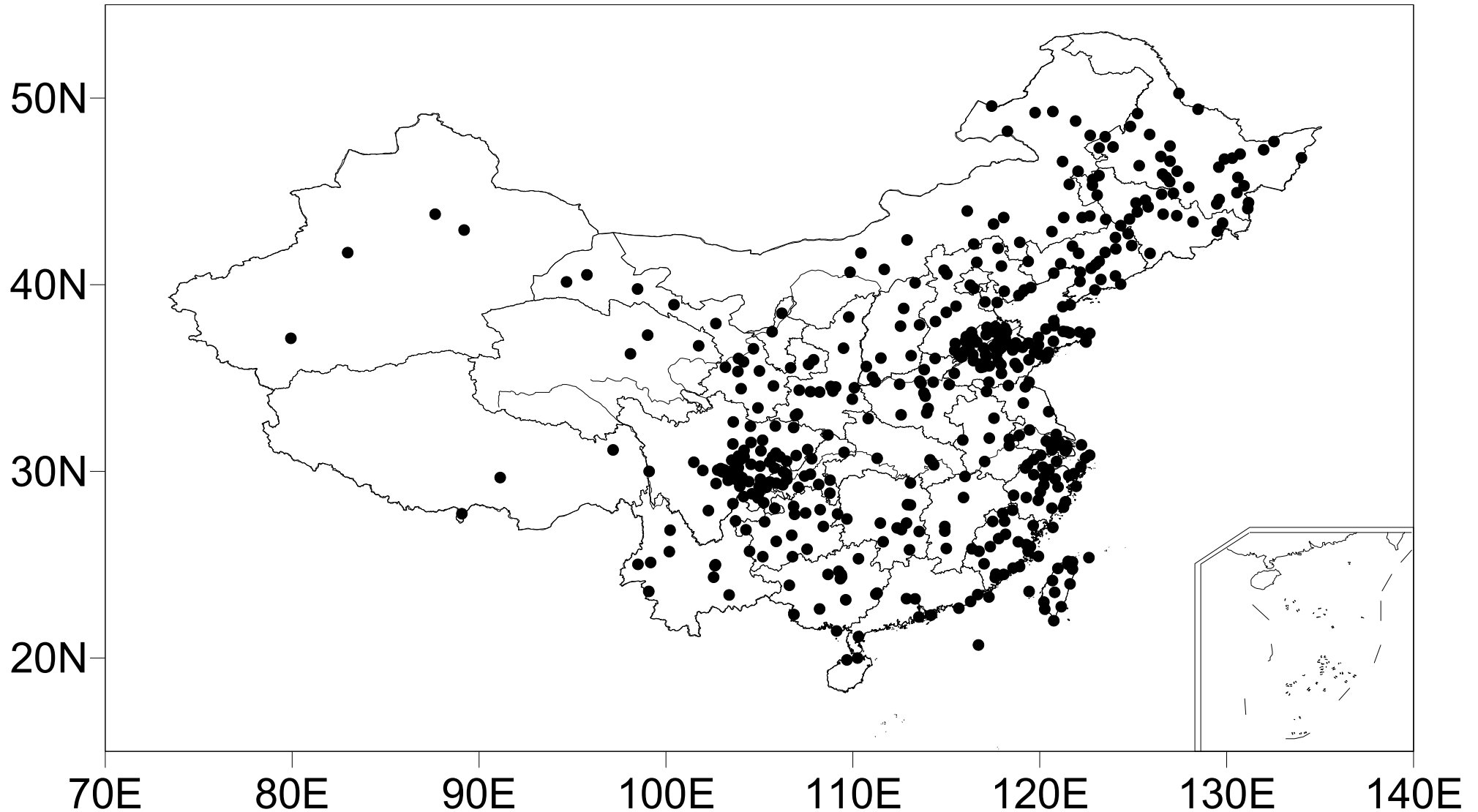
Regression with an average of Tmax and Tmin

Average of multi-times observed Tave series are used as the supplement

Quality control: Outliers, Consistency check



Locations of temperature records



(~1950)

468 stations

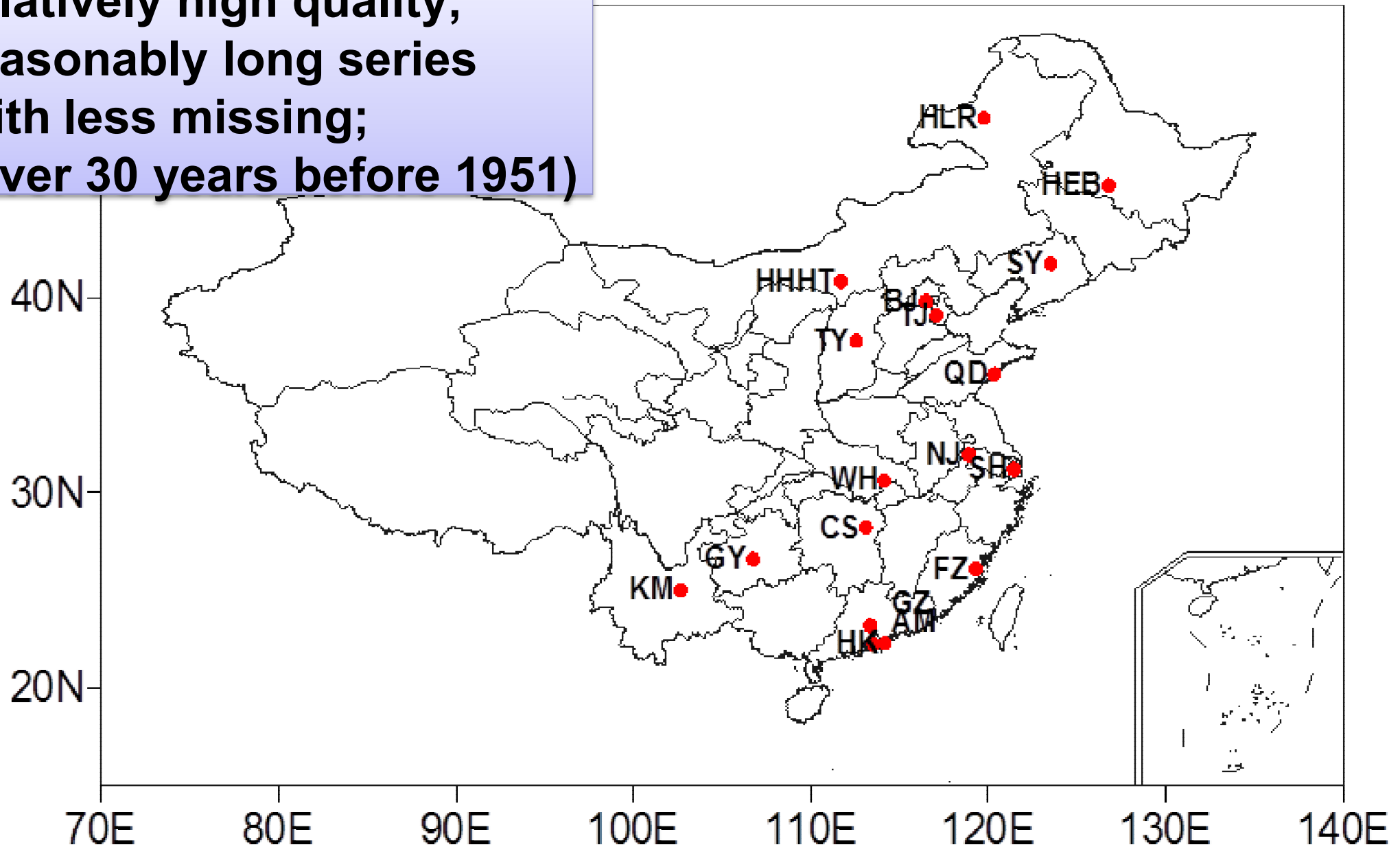
Length of station record (years) From beginning to 1950

| ≥ 50 yr | 40-49yr | 30-39yr | 20-29yr | 10-19yr |
|--------------|---------|---------|---------|---------|
| 7 | 7 | 24 | 26 | 120 |

38 stations over 30 years data records

Location of temperature records (18 stations)

relatively high quality;
reasonably long series
with less missing;
(over 30 years before 1951)





Outline

- Background
- Sources and quality control
- **Interpolation of missing value**
- Homogeneity test
- Summary

Missing percentages of 18 stations(~2010)

| Station index | Station name | Starting time | Length(months) | Missing months | Missing rate(%) |
|---------------|-----------------|----------------|----------------|----------------|-----------------|
| 45005 | Hong Kong | 01/1884 | 1524 | 84 | 5.51 |
| 45011 | Macao | 01/1901 | 1320 | 0 | 0.00 |
| 50527 | Hailar | 01/1909 | 1224 | 126 | 10.29 |
| 50953 | Harbin | 01/1909 | 1224 | 65 | 5.31 |
| 53463 | Hohehot | 01/1915 | 1152 | 55 | 4.77 |
| 53772 | Taiyuan | 01/1916 | 1140 | 55 | 4.83 |
| 54342 | Shenyang | 05/1905 | 1268 | 54 | 4.26 |
| 54511 | Beijing | 09/1890 | 1444 | 131 | 9.07 |
| 54527 | Tianjin | 09/1890 | 1444 | 0 | 0.00 |
| 54857 | Qingdao | 01/1900 | 1332 | 134 | 10.06 |
| 56778 | Kunming | 01/1921 | 1080 | 0 | 0.00 |
| 57494 | Wuhan | 02/1905 | 1271 | 77 | 6.06 |
| 57679 | Changsha | 01/1911 | 1200 | 84 | 7.00 |
| 57816 | Guiyang | 10/1920 | 1083 | 0 | 0.00 |
| 58238 | Nanjing | 01/1905 | 1272 | 98 | 7.70 |
| 58367 | Shanghai | 01/1873 | 1656 | 0 | 0.00 |
| 58847 | Fuzhou | 01/1905 | 1272 | 32 | 2.52 |
| 59287 | Guangzhou | 03/1912 | 1186 | 73 | 6.16 |



Methods of interpolation

Step1: Integration of three methods

1. Standardized method

2. Partial least squares regression

3. Multivariate linear regression model

Step 2:

4. Gradient plus inverse distance square

Step3: same as step 1 with relaxed distance and elevation threshold

Choose reference station:

1. 300km scope

2. $h_0 < 2500\text{m}$, $|h-h_0| \leq 200\text{m}$;

$h_0 \geq 2500\text{m}$, $|h-h_0| \leq 500\text{m}$;

h_0 Height of Candidate Station

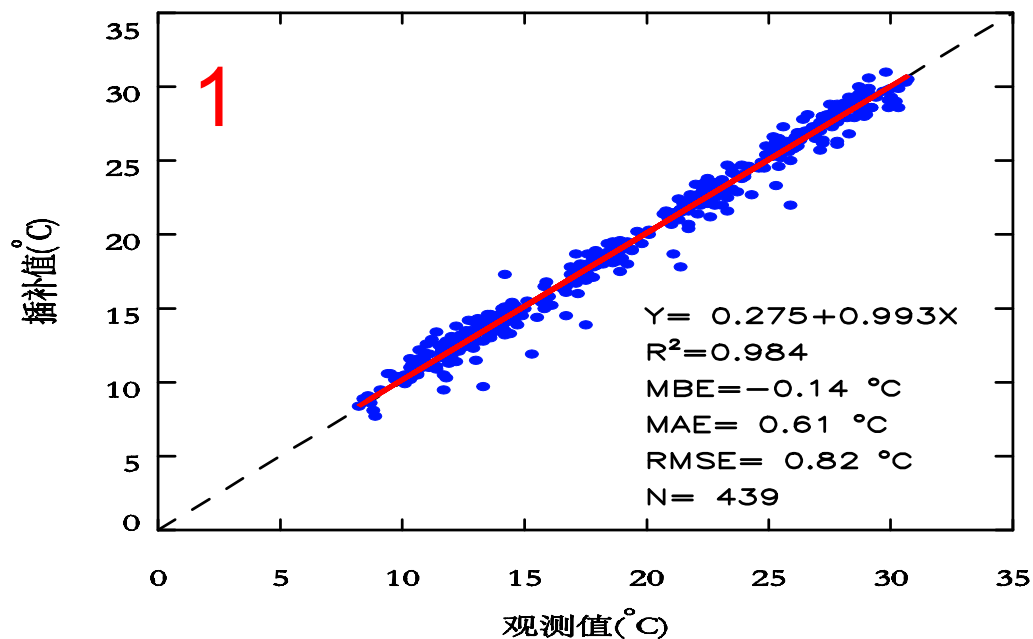
h Height of Reference Station

Best correlation and closest distance are considered

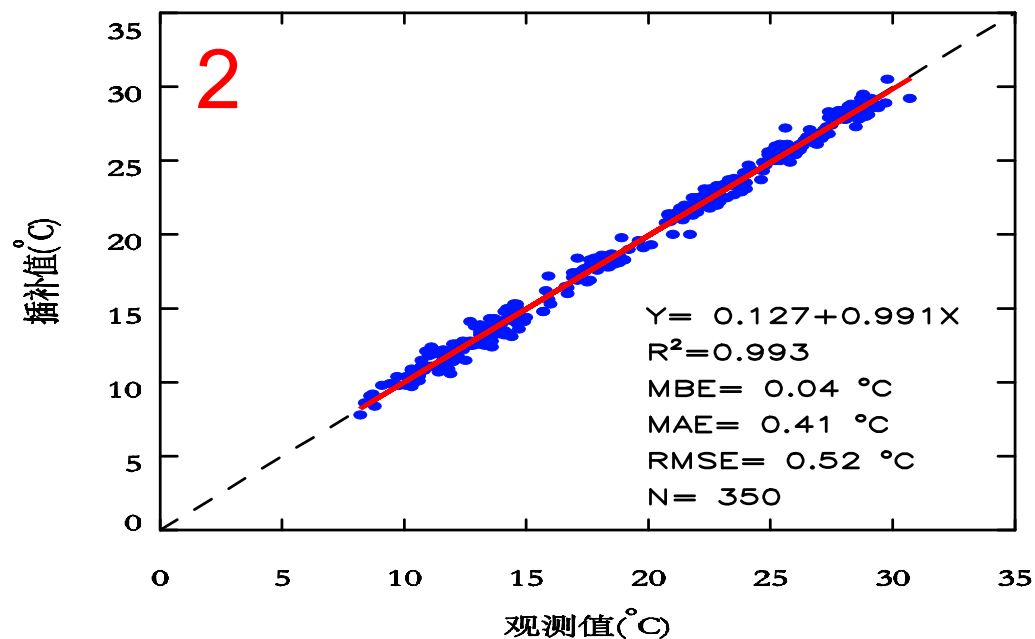
cross validation: comparing interpolated estimates to the actual data values at the candidate station in months when the candidate station had data and the full period of available data could be used to estimate errors.

Exp. Of crosscheck-----Fuzhou station

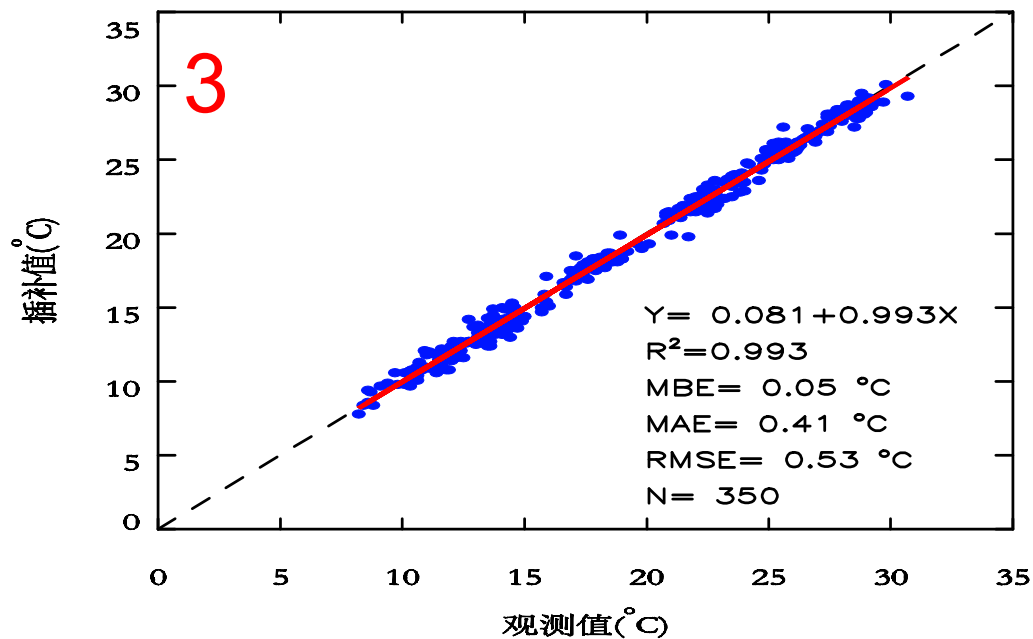
标准序列法



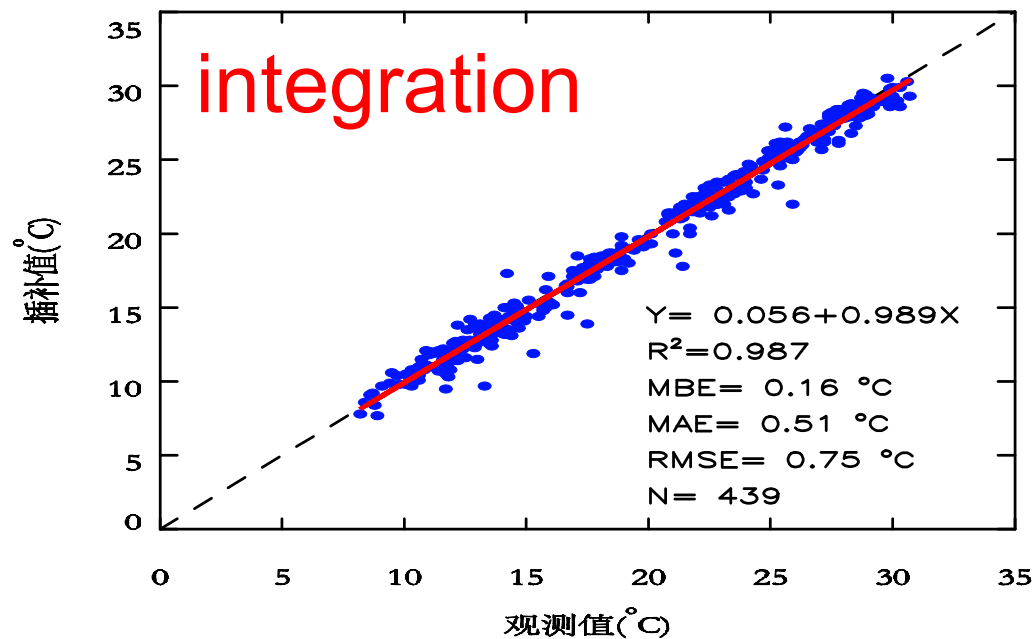
偏最小二乘回归



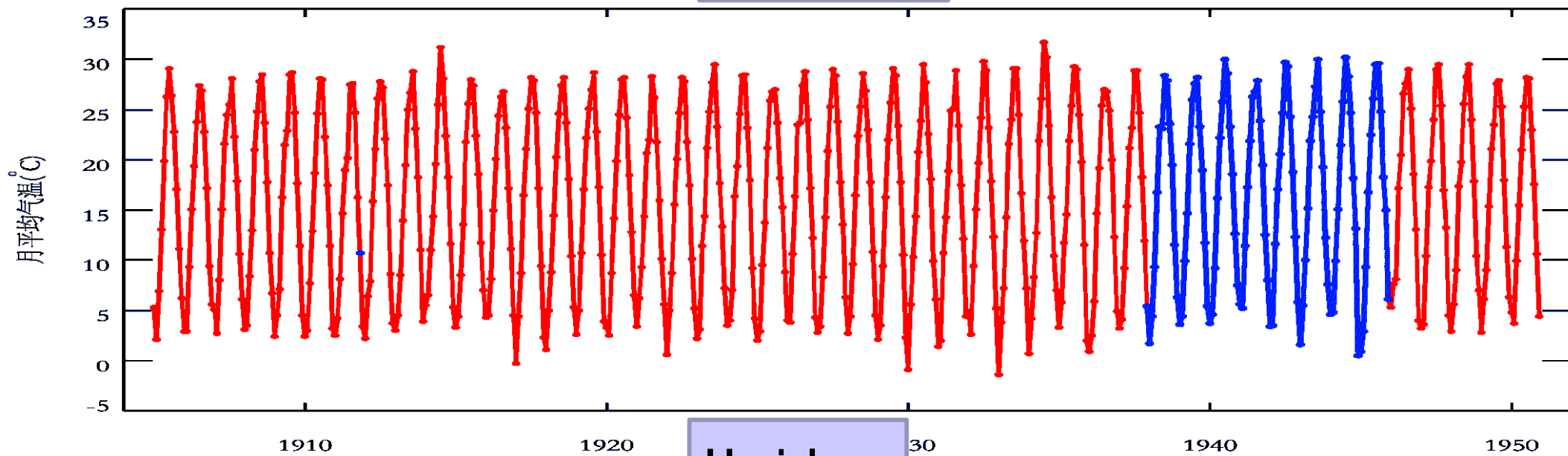
多元线性回归



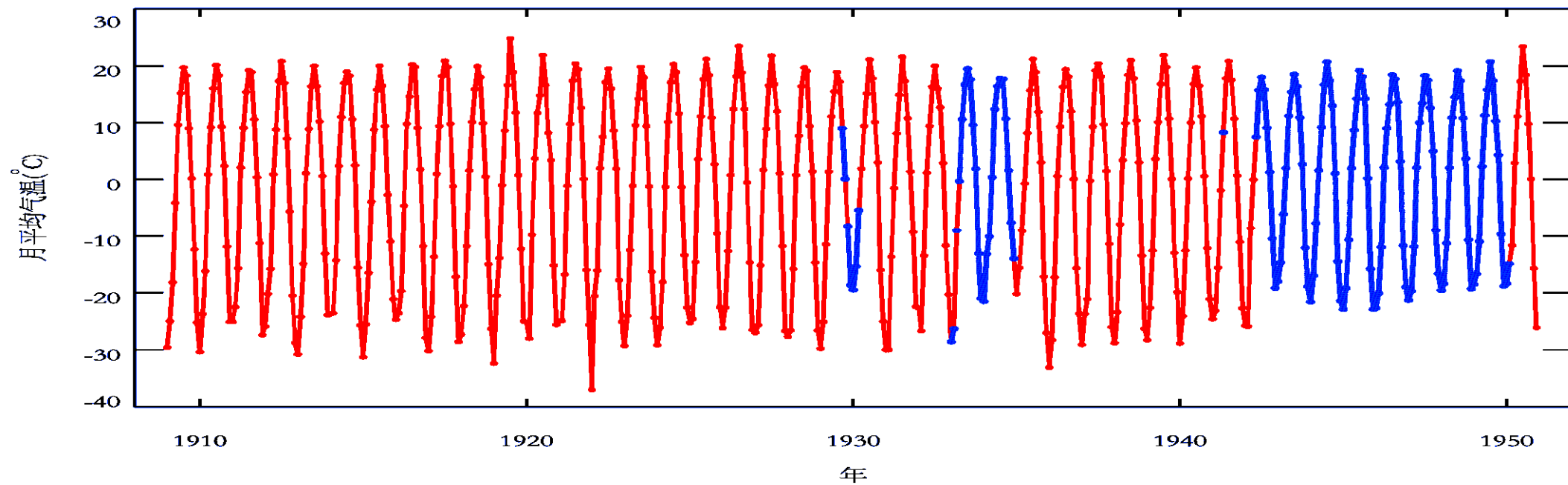
综合插补



NanJing



Hailar



Error estimation of cross check for 18 stations from beginning year to 2010

| St. no. | name | Interpolation time period | | Error | | Interp. Method |
|---------|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|----------|-----------|-------------------|
| | | Time period | missing | MBE (°C) | RMSE (°C) | |
| 45005 | HK | 194001–194612 | 84 | 0.08 | 0.555 | Integrated method |
| 54511 | BJ | 189009–189012, 190003–190304, 190401–190412, 190901–190912, 191201–191312, 191501–191503, 192612–192706, 192710, 192801–192812, 192902–192906, 193709–193812 | 131 | –0.06 | 0.713 | Integrated method |
| 54857 | QD | 191407–191503, 193709–193801, 195801–196012 | 134 | 0.11 | 0.555 | Integrated method |
| 58238 | NJ | 191111, 193712–194512 | 98 | 0.00 | 0.691 | Integrated method |
| 58847 | FZ | 194104–194111, 194207–194209, 194404–194512 | 32 | –0.02 | 0.933 | Integrated method |
| 59287 | GZ | 192207, 193811–193912, 194301–194612, 194705, 194710, 194801, 194806– | 73 | 0.01 | 0.579 | Integrated method |

Error estimation of cross check for 18 stations from beginning year to 2010

| St. no. | name | Interpolation time period | | Error | | Interp. Method |
|---------|------|-----------------------------------------------------------------------------------------------------------------|---------|----------|-----------|----------------|
| | | Time period | missing | MBE (°C) | RMSE (°C) | |
| 50527 | HLR | 192909–192910, 193301–193302, 193304–193312 | 13 | 0.00 | 0.951 | Integrated |
| | | 192911–193003, 193303, 193401–193412, 194105, 194205–194507, 194901–195002 | 72 | –0.28 | 1.134 | Gradient |
| | | 194508–194812 | 41 | –0.02 | 3.161 | Integrated2 |
| 50953 | HEB | 194210–194212, 194702–194804 | 18 | 0.12 | 0.843 | Integrated |
| | | 194308, 194312–194512, 194701, 194805–194812 | 35 | 0.45 | 0.879 | Gradient |
| | | 194601–194612 | 12 | 0.02 | 2.921 | Integrated2 |
| 53463 | HHHT | 193707–193708, 195101, 195301 | 4 | 0.58 | 1.226 | Integrated |
| | | 192406, 193709–193812, 194401–194610 | 51 | 0.00 | 2.529 | Integrated2 |
| 53772 | TY | 193710–193812, 194401–194610, 194701–194702, 195508, 195512 | 51 | –0.01 | 0.796 | Integrated |
| | | 194902–194905 | 4 | 0.02 | 1.864 | Integrated2 |
| 54342 | SY | 194105, 194205–194212, 194308–194309, 194312–194512, 194807–194812 | 42 | 0.02 | 0.497 | Integrated |
| | | 194601–194612 | 12 | 0.03 | 2.676 | Integrated2 |
| 57494 | WH | 193805, 193807–193810, 194102–194104, 194106–194112, 194309–194310, 194312–194407, 194411–194501, 194604–194612 | 49 | –0.08 | 0.608 | Integrated |
| | | 193811–193812, 194301–194308, 194311, 194408–194410, 194502–194603 | 28 | 0.14 | 0.691 | Gradient |
| 57679 | CS | 192301, 193811–193912, 194101–194405, 194904–194909 | 62 | 0.00 | 0.524 | Integrated |
| | | 194406–194603 | 22 | 0.09 | 0.520 | Gradient |



Outline

- Background
- Sources and quality control
- Interpolation of missing value
- **Homogeneity test**
- Summary

Homogeneity test-RHtest

PMT red statistics:

A step - change at c is statistically significant if

$$PT_{\max} = \max_{1 \leq c \leq N-1} [P(c)T(c)] > PT_{\max, \alpha}(\hat{\phi}, N)$$

$P(c)$ is an empirical penalty function, $T(c) = \frac{1}{\hat{\sigma}_c} \left(\frac{c(N-c)}{N} \right)^{\frac{1}{2}} |\bar{X}_1 - \bar{X}_2|$

$$\hat{\sigma}_c^2 = \frac{1}{N-2} \left(\sum_{1 \leq t \leq c} (X_t - \bar{X}_1)^2 + \sum_{(c+1) \leq t \leq N} (X_t - \bar{X}_2)^2 \right)$$

PMFT red statistics:

A step - change at c is statistically significant if

$$PF_{\max} = \max_{1 \leq c \leq N-1} [P(c)F(c)] > PF_{\max, \alpha}(\hat{\phi}, N)$$

$P(c)$ is an empirical penalty function

$$F_c = \frac{(SSE_0 - SSE_A)}{SSE_A / (N - 3)}$$

Version 3 are used

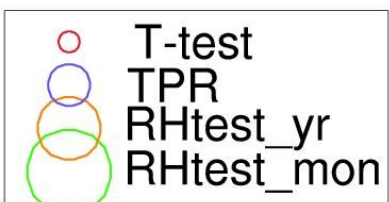
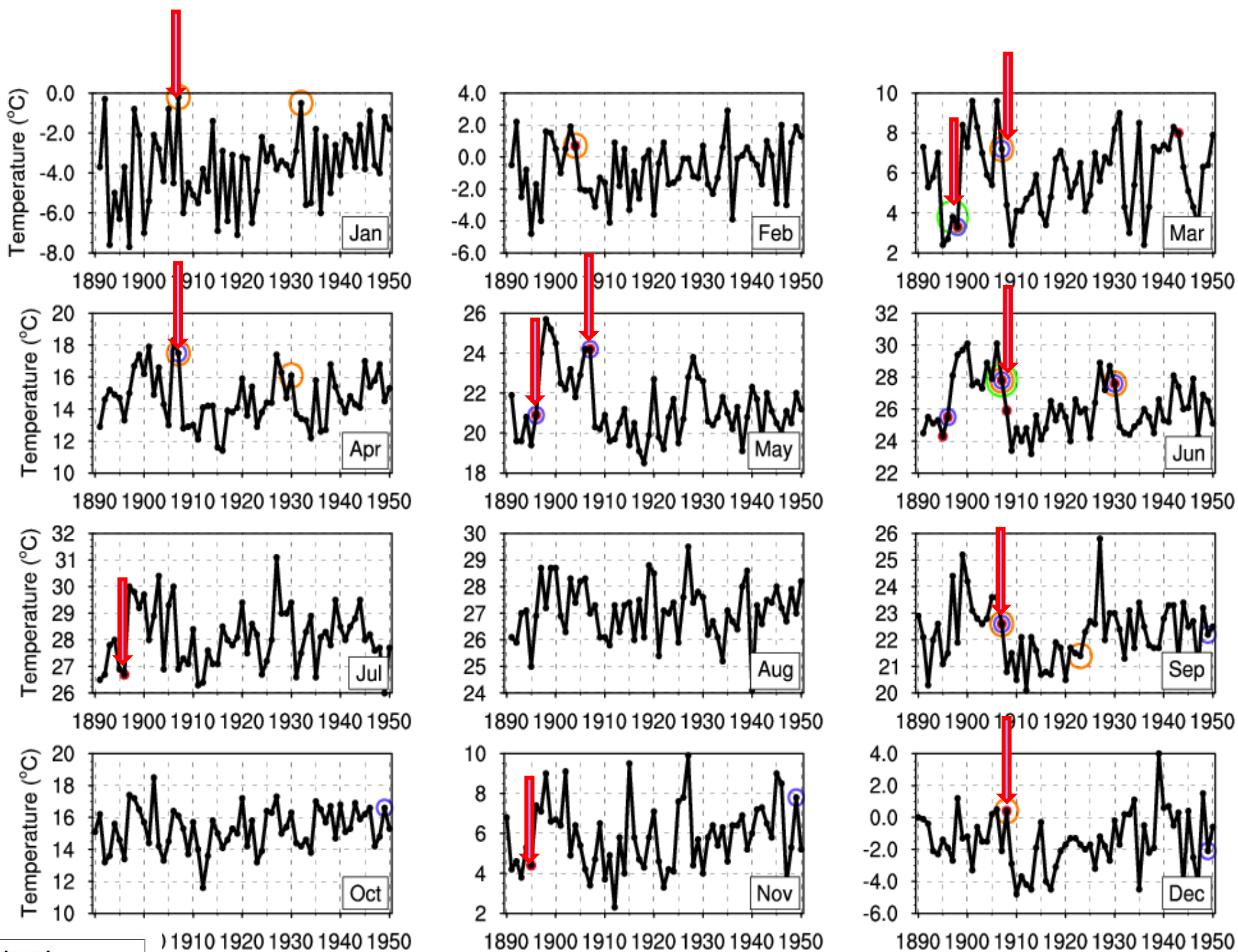
Wang, X. L., and Y. Feng, 2007: RHtestV2 user manual. [Available online at <http://cccma.seos.uvic.ca/ETCCDMI/software.shtml>.]



- **Rhtest+TPR+T:**

- **No more** appropriate reference series to choose
- **Statistical tests** are applied to **each** series
- **Metadata are used** to support the confirmation of the breakpoint
- **Two time period** for homogeneity test :
 - **1. beginning~to~1950**
 - **2. beginning~to~2010**

Tianjin

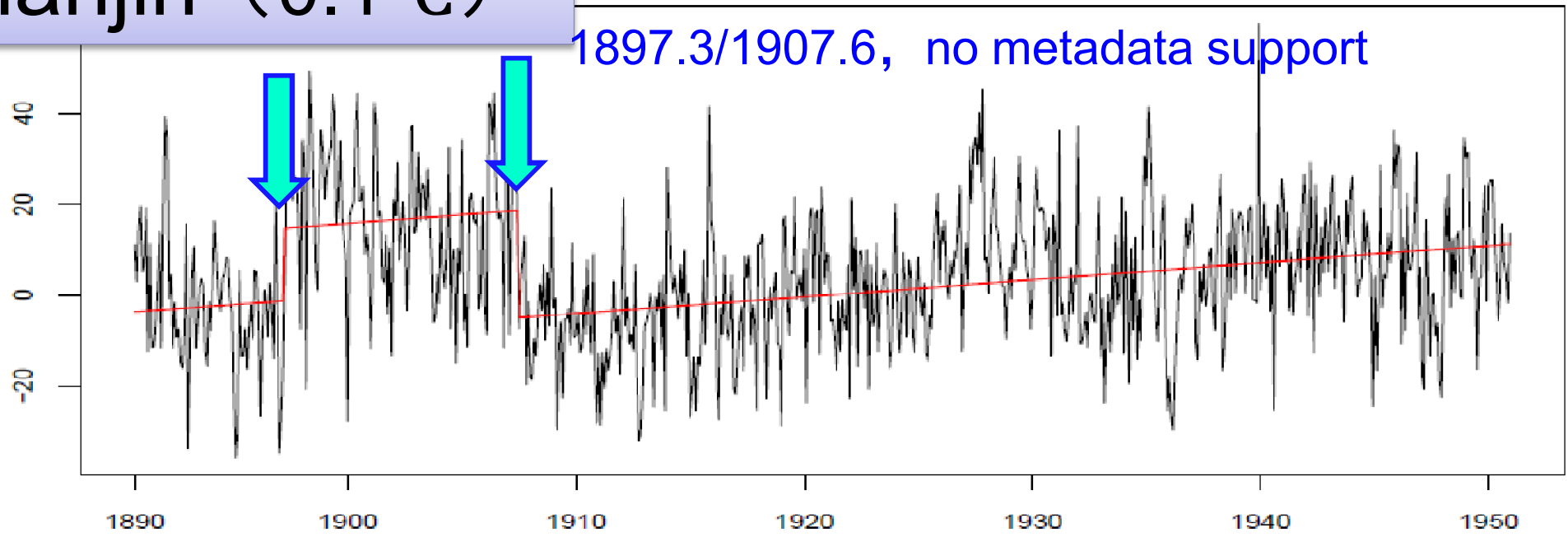


Changepoint: 1897.03 , 1907.06

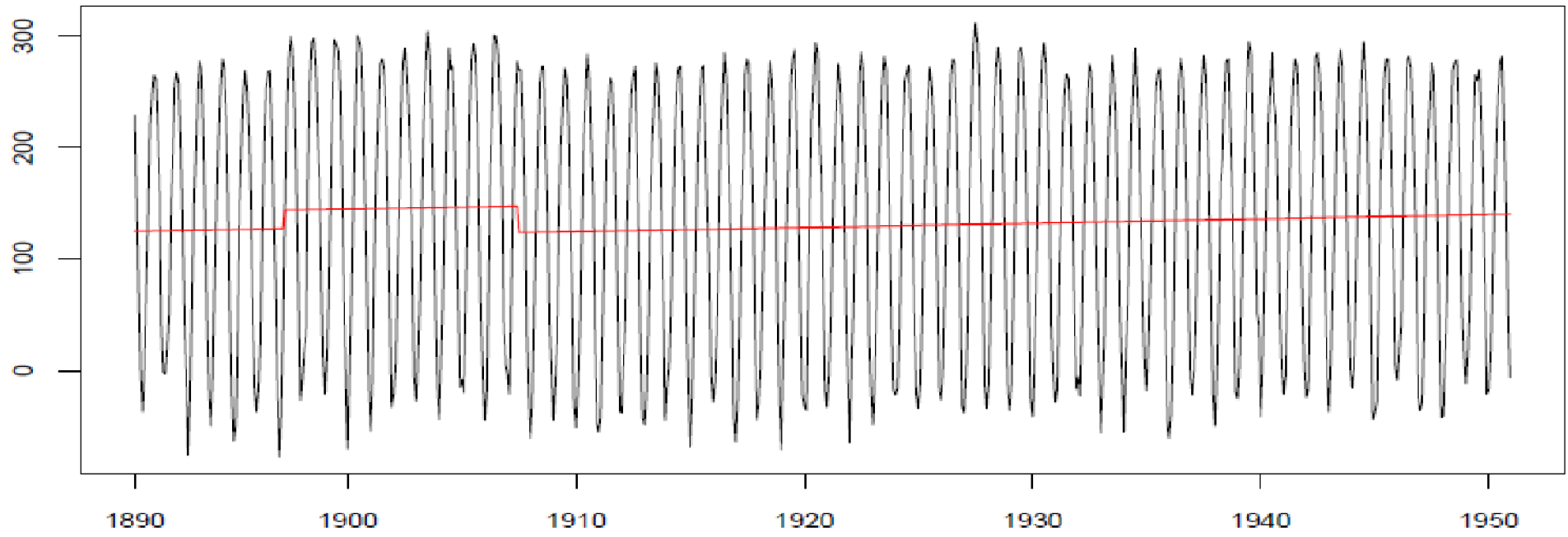
Tianjin (0.1°C)

Base anomaly series and regression fit

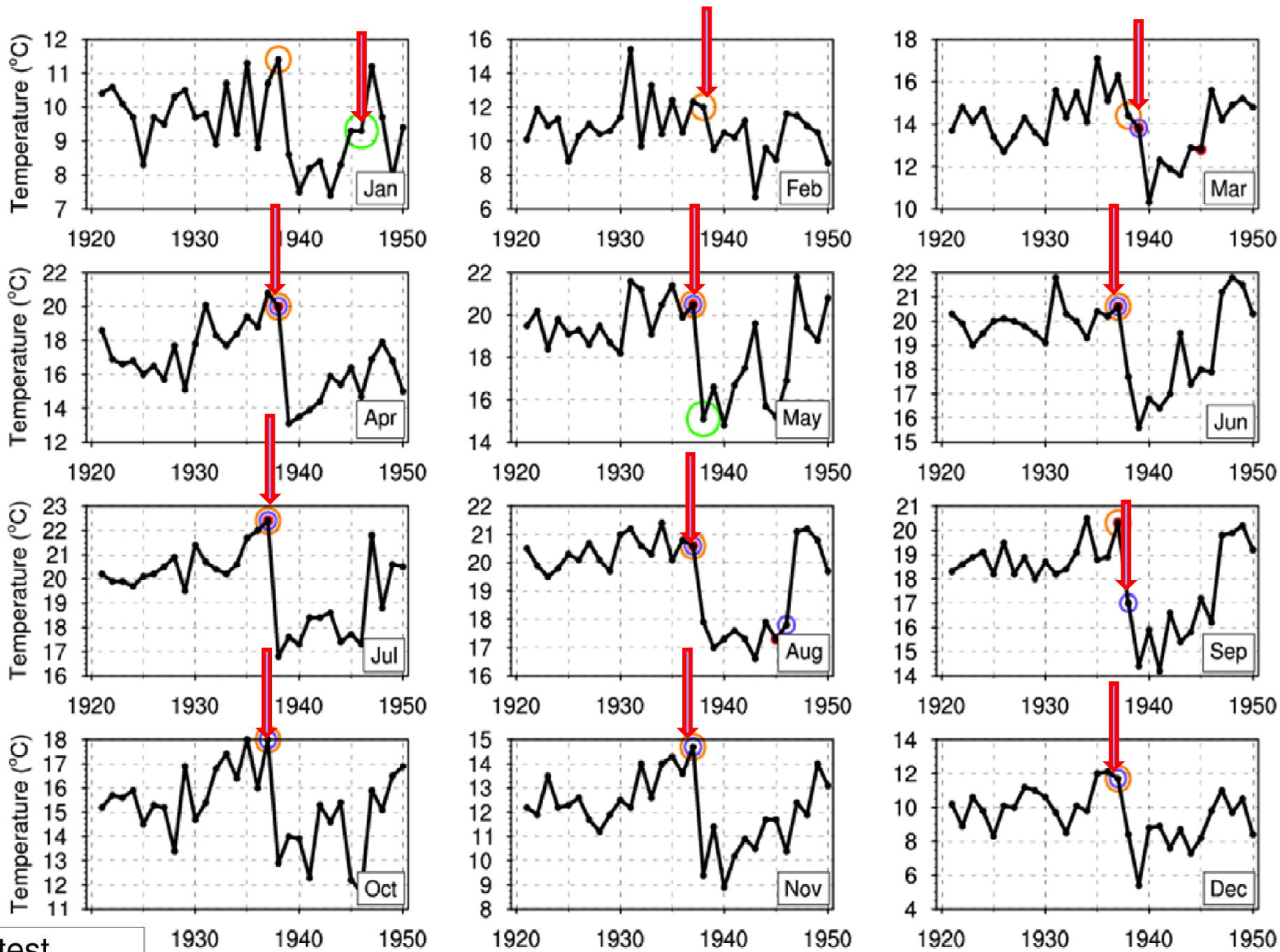
1897.3/1907.6, no metadata support



Base series and regression fit



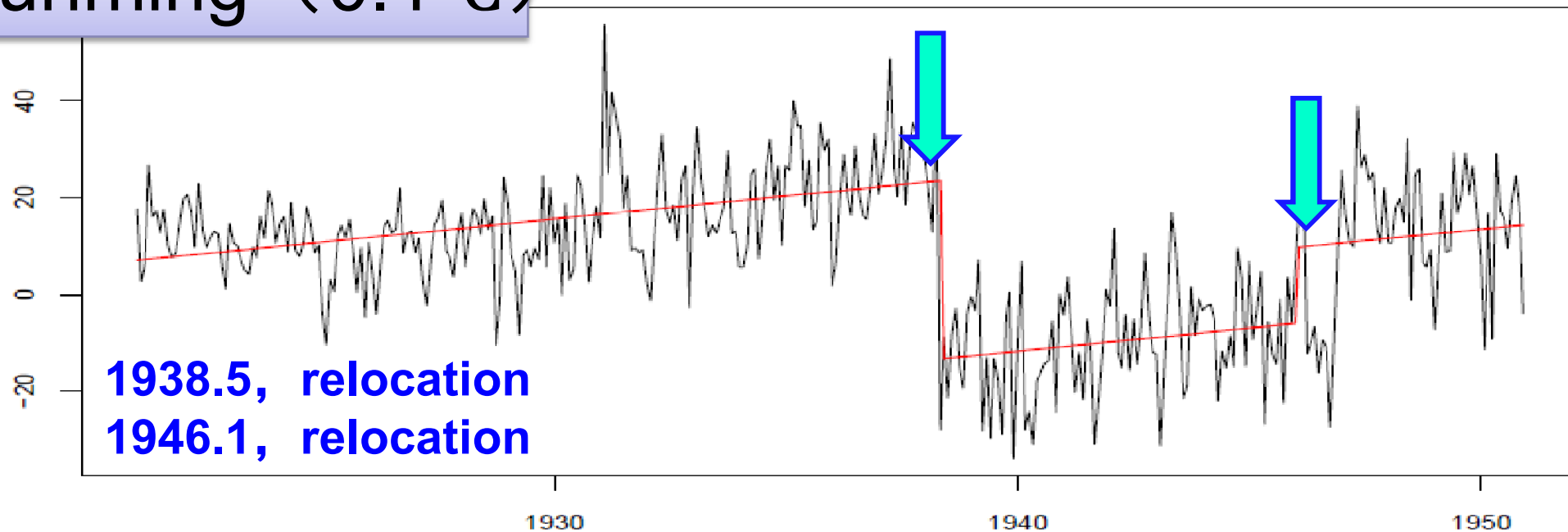
Kunming



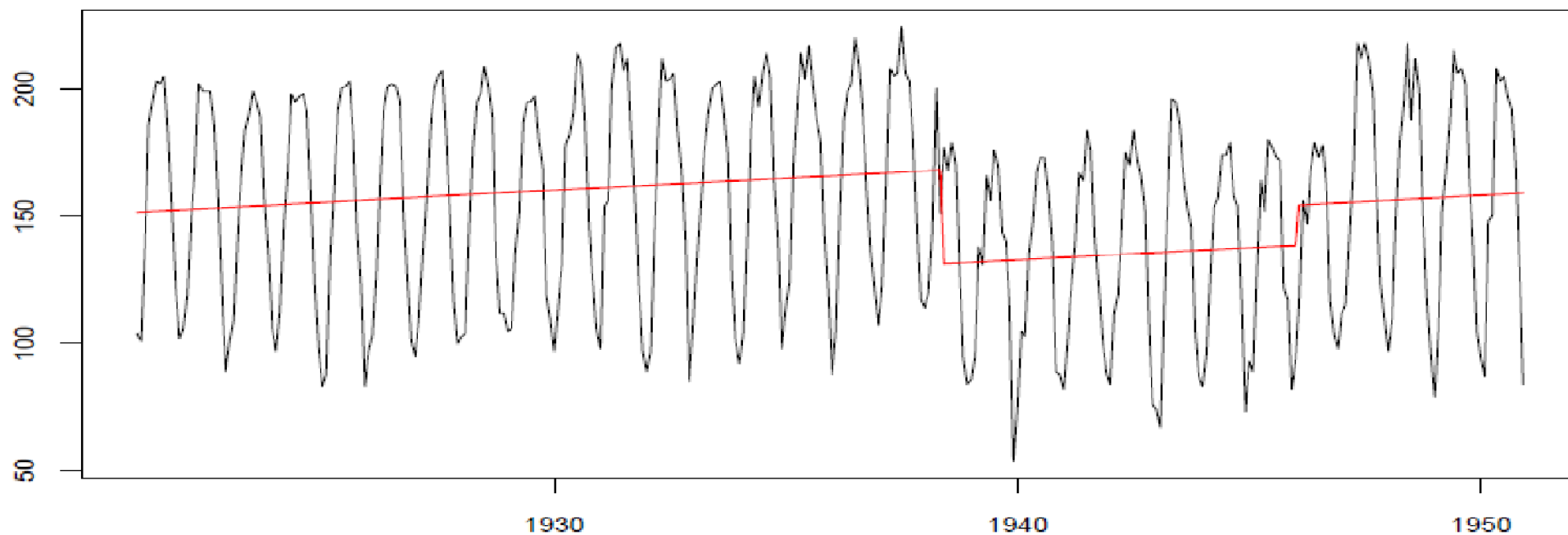
Changepoint: 1938.05, 1946.01

Kunming (0.1°C)

Base anomaly series and regression fit



Base series and regression fit



Homogeneity result (1873~2010)

•1951~2010:

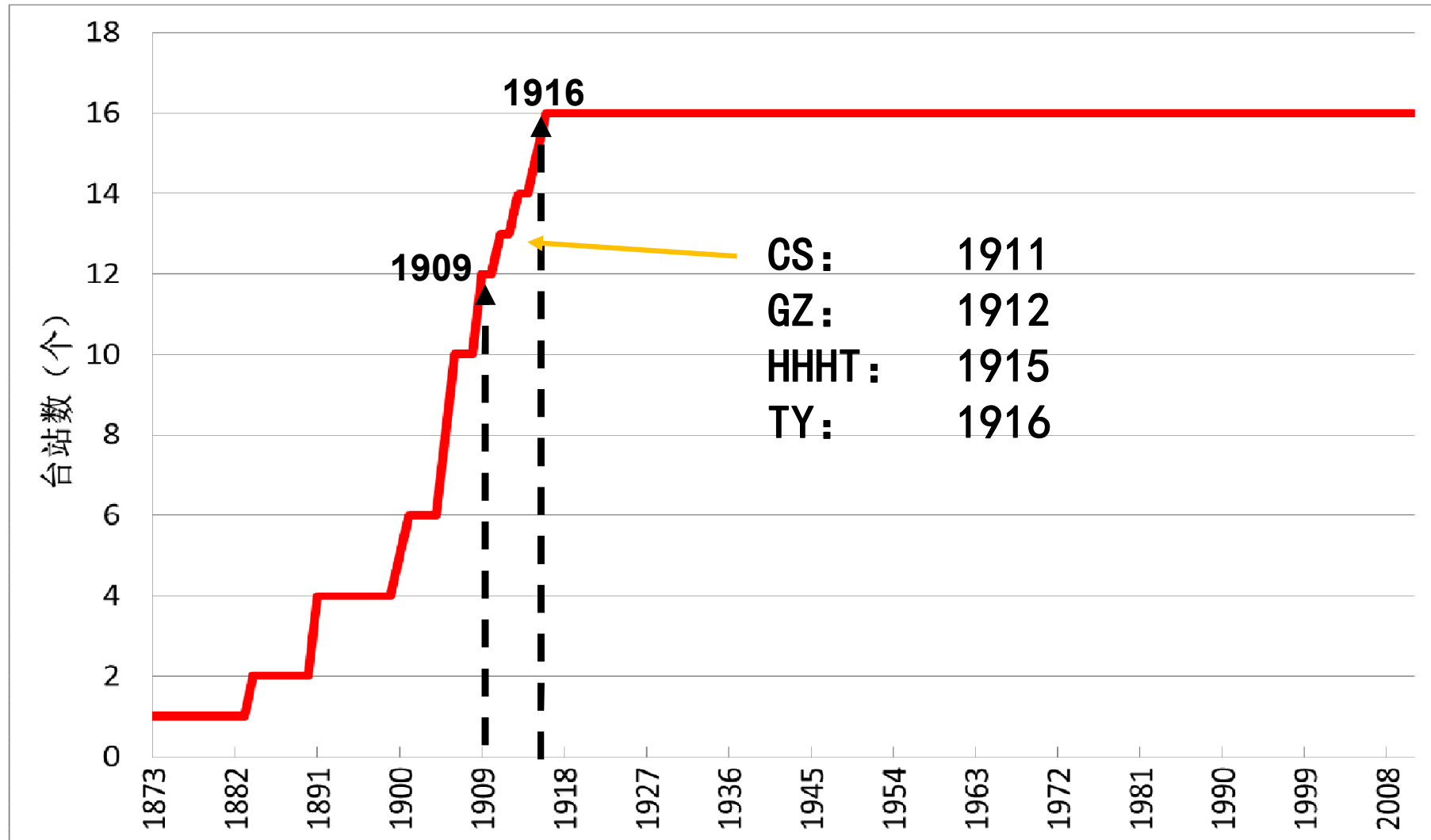
- 28 change points at 11 stations
- 21 change points due to relocation of meteorological station
- 3 change points due to instrument change
- 4 change points without clear reason

•1873~1950:

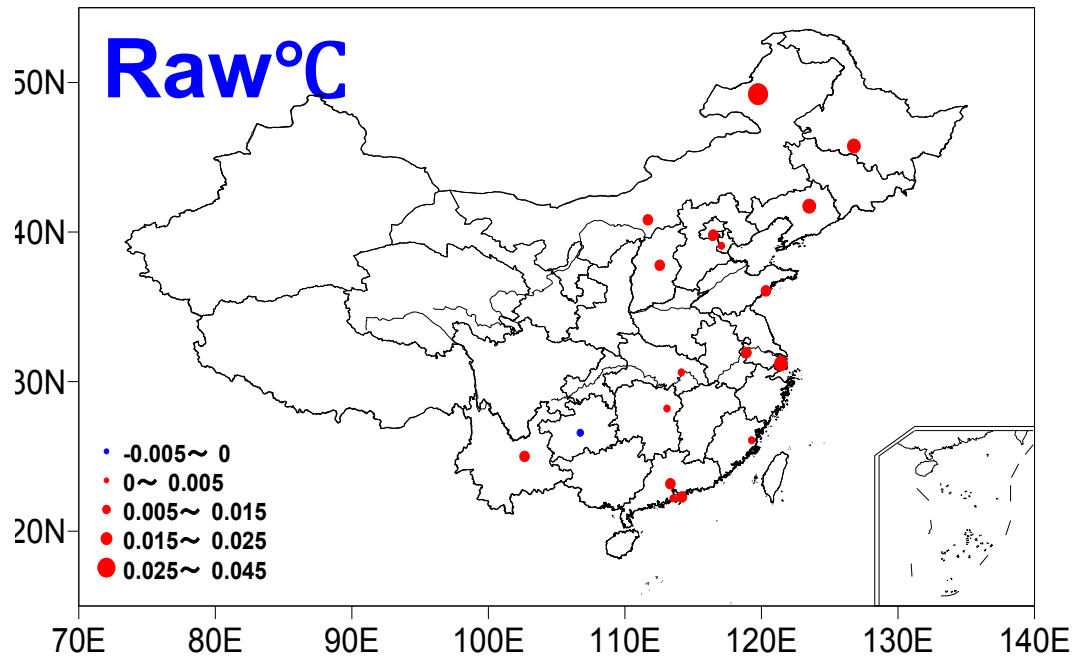
- 5 change points at 3 stations
- 3 change points due to relocation of meteorological station
- 2 change points without clear reason

•Hongkong, Hailar, Hohehot and Qingdao are not adjusted

Station change with year



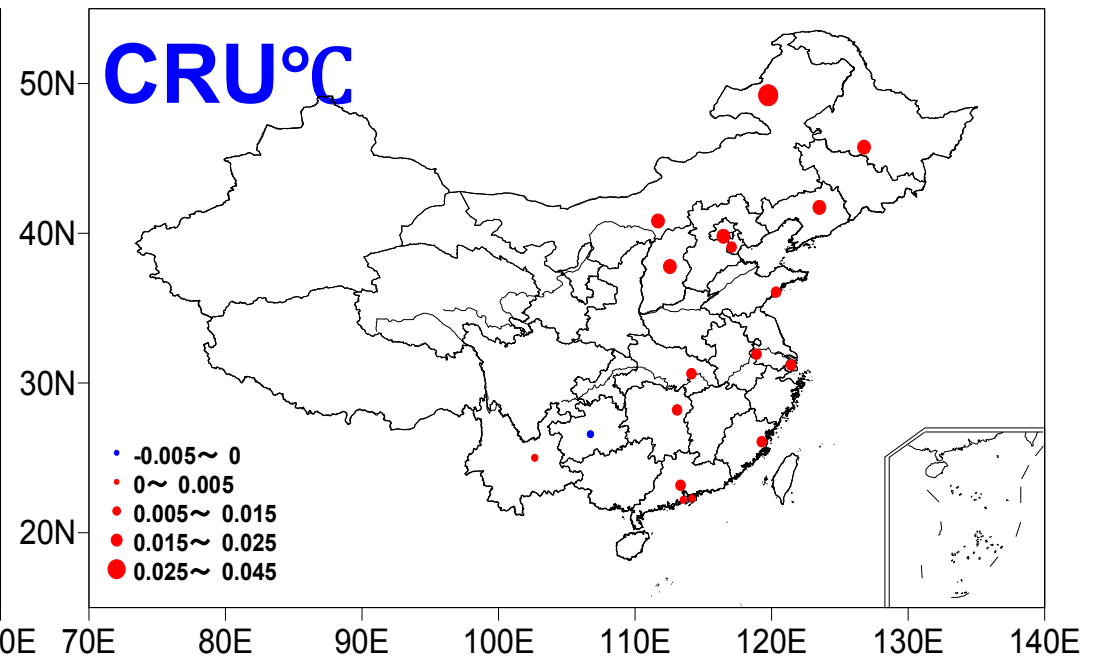
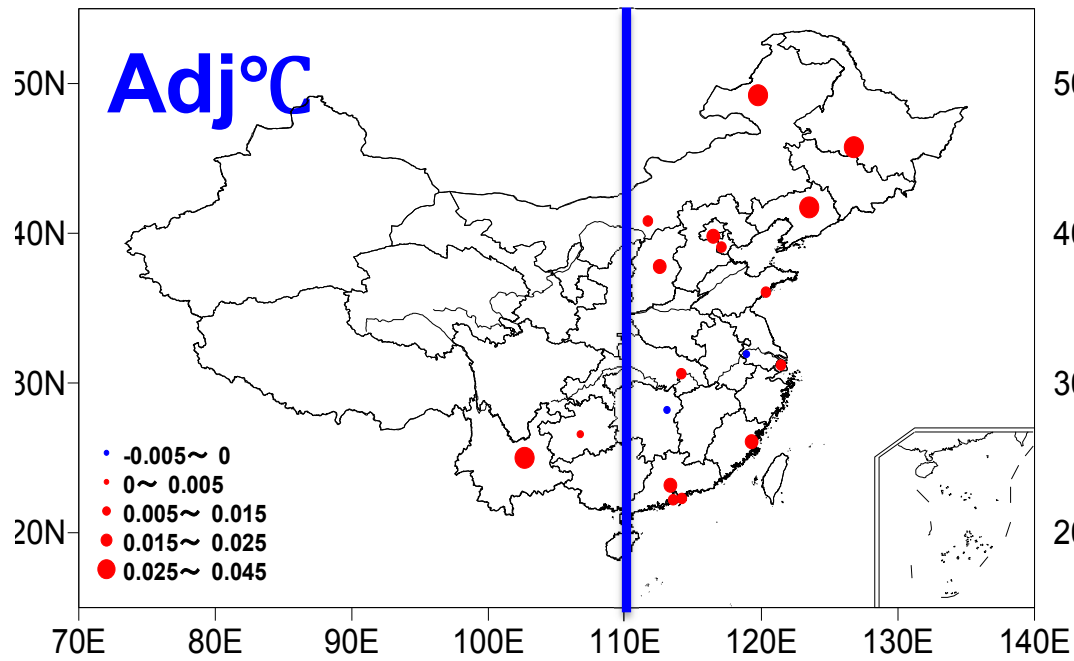
Linear trend of the 18 stations (1909-2010)



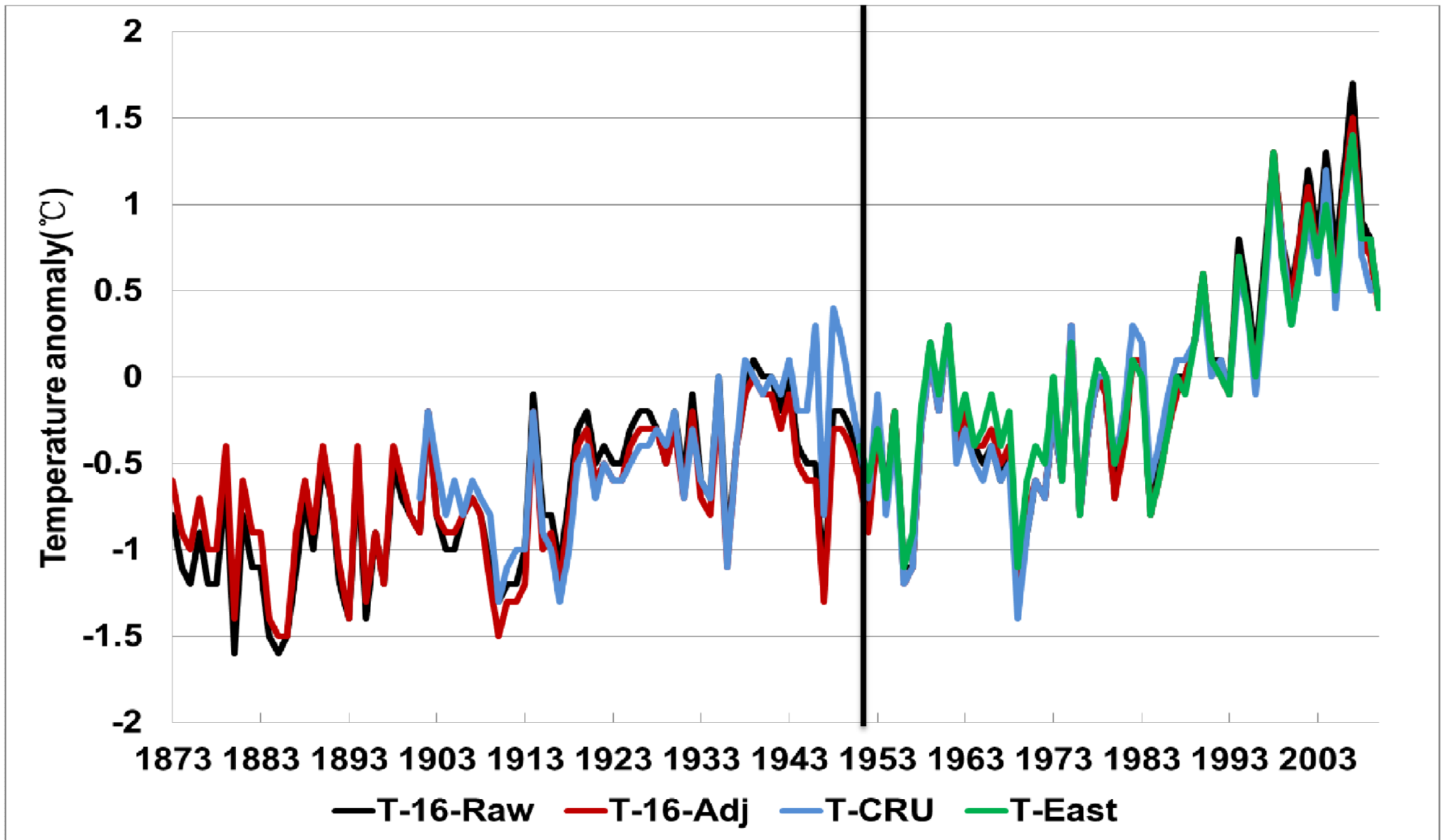
RAW trend=1.29°C/100a

ADJ trend=1.52°C/100a

CRU trend=1.30°C/100a




Annual mean temperature anomalies based on the time period of 1971-2000





Outline

- Background
- Sources and quality control
- Interpolation of missing value
- Homogeneity test
- **Summary**

- 
- (1) A new set of long-term instrumental monthly mean temperature series are reconstructed. Consists of monthly mean surface air temperature of 18 stations in middle-eastern China from 1873 to 2010.
 - (2) To construct long-term series of temperature over western China in the past century by combining limited observational data in western China and some observational data in neighboring countries before 1950.

Questions and comments

Thanks for your attention!

See: Cao L. J., P. Zhao, Z. W. Yan, P. Jones, Y. N. Zhu, Y. Yu, G. L. Tang. Instrumental temperature series in eastern and central China back to the 19th century [J]. *J. Geophys. Res.*, 2013, doi: 10.1002/jgrd.50615.