GLOBÁLNÍ OTEPLOVÁNÍ VE SVĚTLE ANALÝZY KLEMENTINSKÉ TEPLOTNÍ ŘADY

JAROMÍR ANTOCH

ENERGY DAYS 2023
PRAHA
11. LISTOPADU 2023

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

KLEMENTINUM

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

DATA DATA DATA

KLEMENTINUM

JOSEPH STEPLING (1716 – 1778)

- Česko-německý jezuitský kněz, fyzik, astronom, matematik a meteorolog. Vzdělání se mu dostalo pouze na jezuitských školách.
- Roku 1748 odmítl veřejně vyučovat aristotelovskou filozofii, protože nehodlal jiným vnucovat nauku, která byla proti jeho přesvědčení.
- Na jeho žádost mu povolili přednášet pro své spolubratry matematiku a fyziku v newtonovském duchu, a jmenovali jej ředitelem matematicko-fyzikálních studií na pražské univerzitě.
- V roce 1751 se významně zasloužil o vybudování klementinské hvězdárny a stal se jejím prvním ředitelem.
- V roce 1753 se stal státním direktorem na filozofické fakultě v Praze. Díky zásluhám pro rozvoj vzdělávání mu byla ponechána profesura **i po zrušení jezuitského řádu**.
- Z dědictví daroval na zakoupení přístrojů pro klementinskou hvězdárnu částku 4 000 zlatých, desetkrát víc, než činil příspěvek českých stavů. Před svou smrtí odkázal univerzitní knihovně 600 svazků knih.

JOSEPH STEPLING

V roce 1748 požádala Berlínská akademie věd Univerzitu Karlovu o změření zeměpisné délky Prahy. A ta tím „pověřila" Steplinga.

- Při měření používal dvou metod,
  - zatmění Měsíce;
  - zatmění satelitů Jupitera.

  Aby zpřesnil měření, doplnil údaje o vstup a výstup některých měsíčných kráterů do zemského stínu. Měření se snažil co nejvíce zpřesňovat, proto v padesátých letech použil sedmi současných měření zatmění měsíce v Praze a ve Vídni.
- Stepling uvedl pražský polednik jako $32°11'15''$
  - Tím určil délku Klementina na $12°11'15''$ vychodně od Paříže
  - Protože Paříž má v dnešnim počítání souřadnice $2°20'11,5''E$, mělo by Klementinum délku $14°3'27''E$, čili o $0.10765°$ více na východ.
  - Stepling tedy umístil Klementinum jen asi o 7.68 km východněji (do dnešnich Malešic), což byla s tehdejšími přístroji to byla úžasná přesnost!!!

PŘESNOST STANOVENÍ

- Stepling vztahuje svou informaci k poledníku, který vedl přes nejmenší a nejzápadnější ostrov Kanárského souostroví.
- Připomeňme, že v letech 1634 – 1884 byl za nultý poledník považován poledník procházející mysem Orchilla na ostrově El Hierro, nejzápadnějším ostrovem Kanárských ostrovů. Ten byl „odjakživa“ považován za západní „konec světa“, protože již Ptolemaios se domníval, že tam končí svět, takže by se geografické délky měly měřit od tamního poledníku.
- Konkrétně jej jako nultý poledník ustanovil francouzský král Ludvík XIII, a to v roce 1634 královským rozhodnutím a výnosem, že délka pařížské hvězdárny je přesně $20^\circ E$.
- Nultých poledníků existovalo v dějinách astronomie několik. Ke sjednocení na greenwichskou observatoř v jihovýchodním Londýně došlo až v říjnu 1884 ve Washingtonu.
- JA dospěl podle Google maps k poloze Prahy $32^\circ 34' 12'' E$. Problém leží v nepřesném odhadu astronomů Ludvíka XIII.

MORE ABOUT KLEMENTINUM

History

- Observatory founded (by Jesuits): 1752
- Location: $50°5'11.8''N$, $14°24'56.4''E$, 197.0 m above sea level
- Measures: Temperature, humidity, precipitations, sunshine etc.
- Continuous measurement of temperature: since 1775

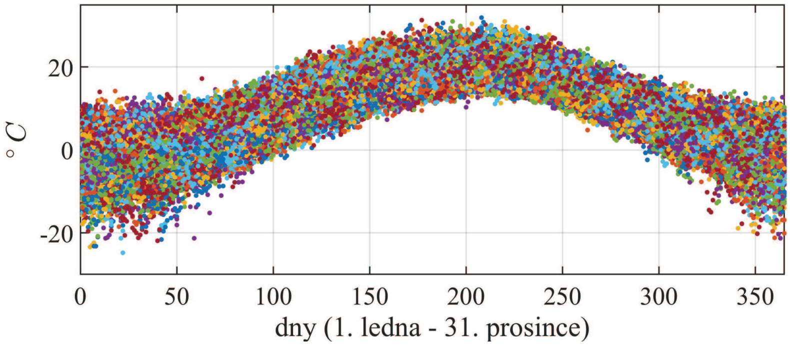
Year means

- Mean year temperature: 9.74°C
- Mean year precipitation: 464.5 mm
- Mean year sunshine: 1632.7 hours

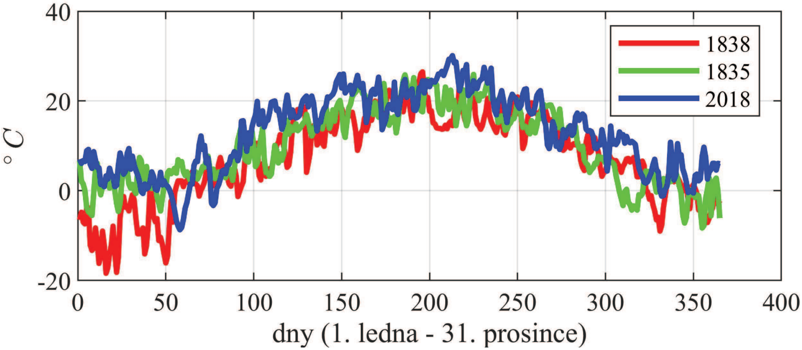
Daily extremes

- Maximal temperature: 37.9°C (July 1, 2019)
  37.8°C (July 27, 1983 & July ??, 2013)
- Minimal temperature: -27.6°C (March 1, 1785)
- Maximal precipitations: 90.0 mm (July 19, 1981)

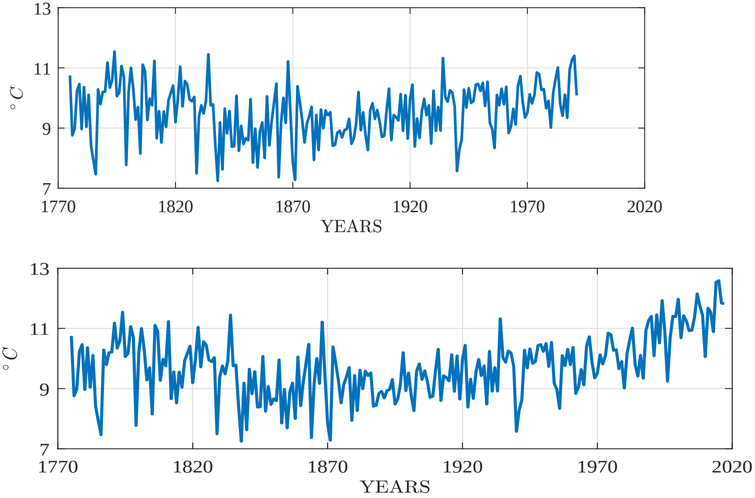
MEAN DAILY TEMPERATURES 1775 – 2018



COURSE OF MEAN DAILY TEMPERATURES SELECTED YEARS

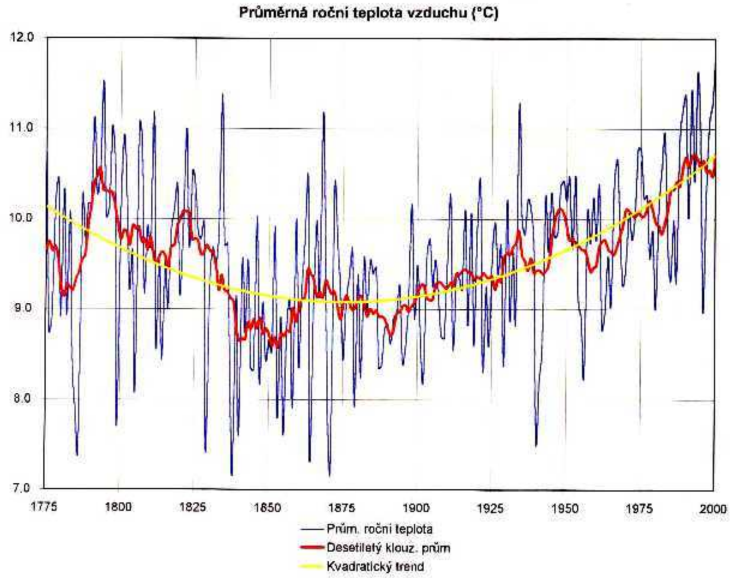


MEAN YEAR TEMPERATURES

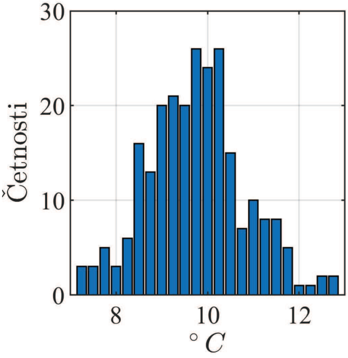


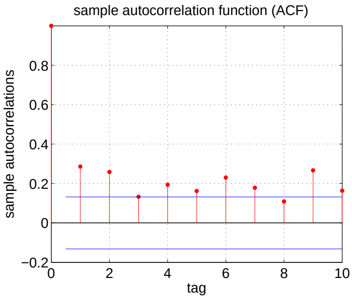
Mean year temperatures in Praha-Klementinum

METEOROLOGISTS' SMOOTHING



DISTRIBUTION OF DATA





min=7.24 (1838), max=12.87 (2019)

mean=9.74, skewness=0.19, kurtosis=3.05

DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA

OBECNÝ POVZDECH

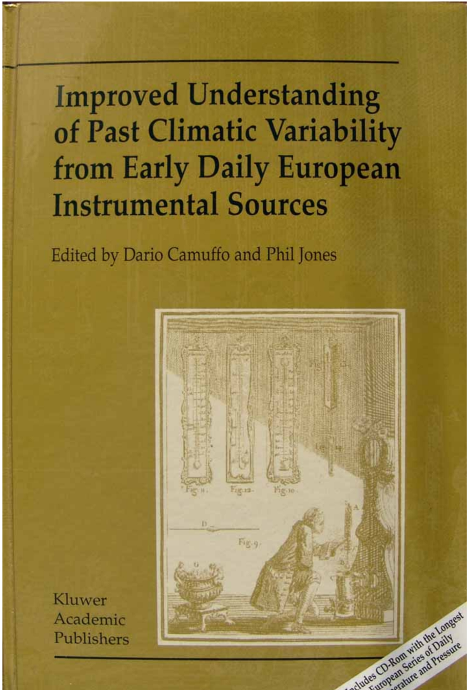
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA

DATA SOURCES

ČHMÚ Český hydrometeorologický ústav

KLEIN TANK A.M.G. et al. (2002). Daily datasets of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International J. of Climatology*, 22, 1441–1453.
Data and metadata are available at `http://www.ecad.eu/`
For Klementinum data see files:
`TG_STAID000027.txt`, `TN_STAID000027.txt`, `TX_STAID000027.txt` and `RR_STAID000027.txt`

TRENDS 93: A compendium of data on global change
January 2003, DOI: 10.2172/10106351
T.A. Boden, Dale P. Kaiser, R. Sepanski, G.M. Logsdon


Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources
Edited by Dario Camuffo and Phil Jones
Kluwer Academic Publishers
Includes CD-Rom with the Longest
European Series of Daily

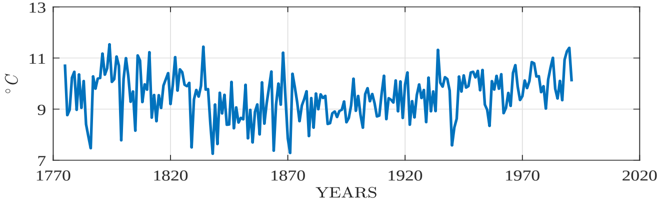
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA

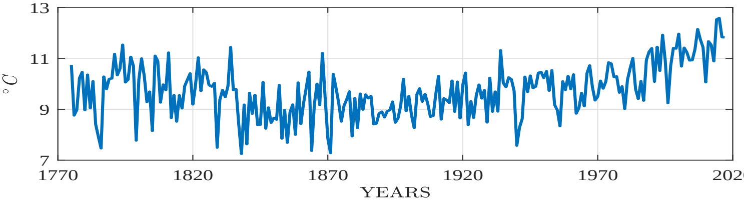
SEGMENTATION

DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA
DATA DATA DATA

GOAL

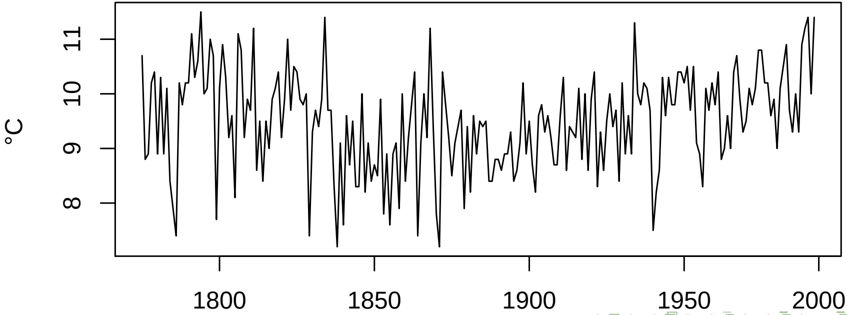
Assume observations $Y_1, \ldots, Y_n$ obtained in ordered time moments $t_1 < \ldots < t_n$. Our goal is to split them into smaller – more "homogeneous" – parts. We will call this procedure a segmentation.





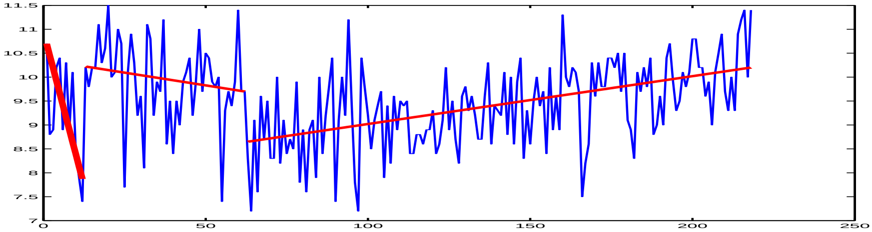
SELECTED BASIC QUESTIONS

- What model is appropriate?
- Is there **statistically detectable change** in the data?
- Is there just one change or are there more changes?
- What is the number of changes and how to estimate them?
- Which optimality criteria to use for both testing for a change and estimating its location(s)?



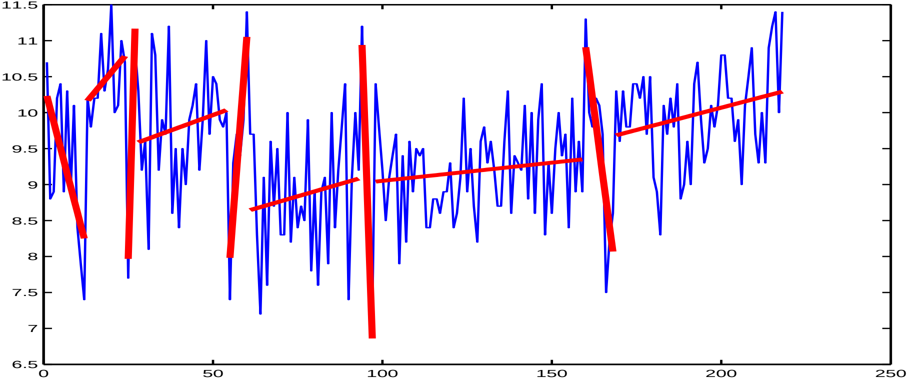
ONE POSSIBLE SOLUTION

Assume observations $Y_1, \ldots, Y_n$ obtained in ordered time moments $t_1 < \ldots < t_n$. Our goal is to split them into smaller and more "homogeneous" parts. We will call this procedure a segmentation.



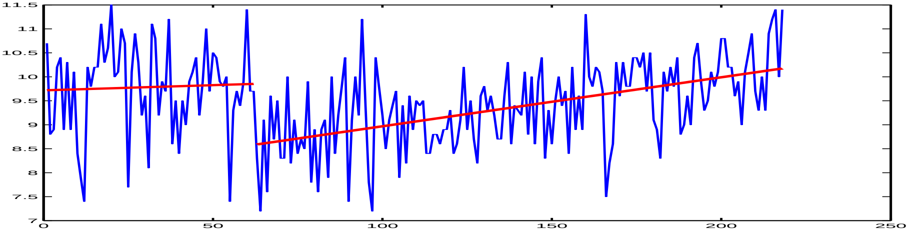
Notice that our goal IS NOT to analyze some specific data. On the contrary, we will speak about selected methods and their properties and use data (well known from the literature) as the illustration.

"OPTIMAL" SOLUTION BASED ON INFORMATION CRITERIA

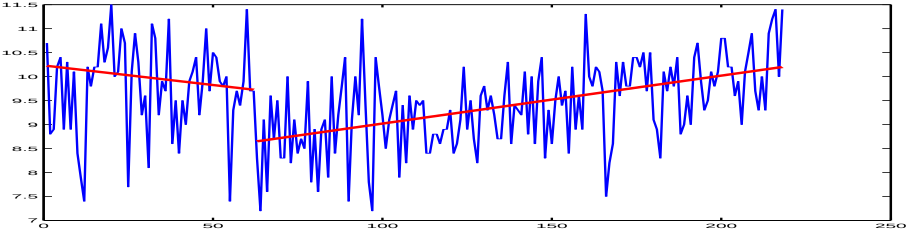


Best multi-phase linear regression with 9 jumps suggested by the information criteria (both BIC and AIC), L_2-norm criteria function used.

ANOTHER POSSIBLE SOLUTION

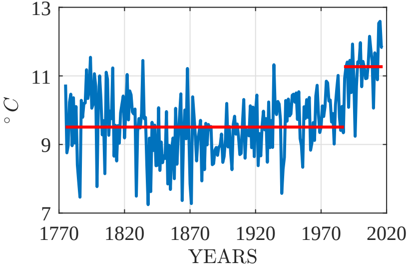


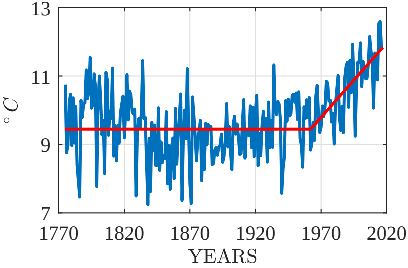
L_2-norm



L_1-norm

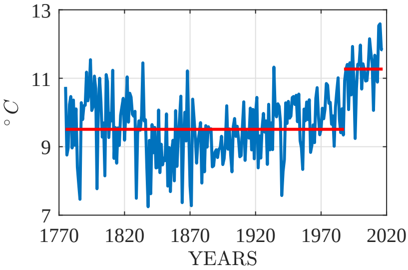
ONE CHANGE ONLY

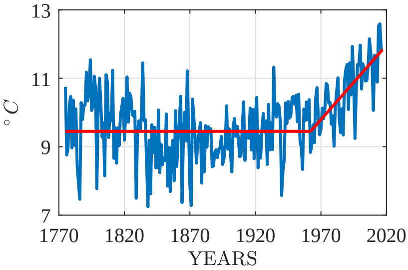




$\mathbf{(1987 \mp 2)}$ $\mathbf{(1962 \mp 11)}$

ONE CHANGE ONLY

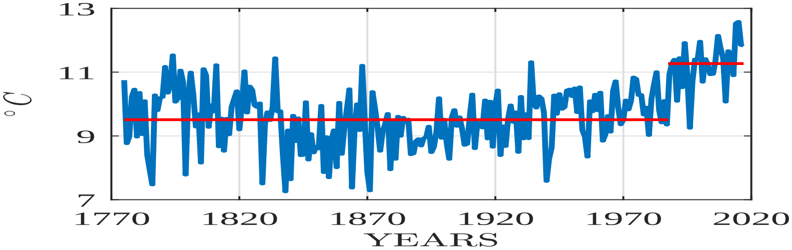




$(\mathbf{1987} \mp \mathbf{2})$ $(\mathbf{1962} \mp \mathbf{11})$

Danger! Setting and fitting the model results in estimated parameters (and models), plots etc. However, this says nothing about the appropriateness of the model.

“OPTIMAL” ABRUPT-SHIFT MODEL (1987∓2)



$$P\left(\frac{\delta^2}{\sigma^2}\left(\widehat{k}-k\right) \le x\right) \approx P(V \le x), \quad x \in \mathcal{R}^1$$

$$V = \arg\max \left\{W(s) - |s|/2\,;\ s \in \mathcal{R}^1\right\}$$

$\left\{W(s); s \in \mathcal{R}^1\right\}$ is two-sided standard Wiener process

$$W(s) = \begin{cases} W_1(-s) & s < 0 \\ W_2(s) & s > 0 \end{cases}$$

"OPTIMAL" ABRUPT-SHIFT MODEL (1987∓2)

$\{W(s); s \in \mathcal{R}^1\}$ is two-sided standard Wiener process

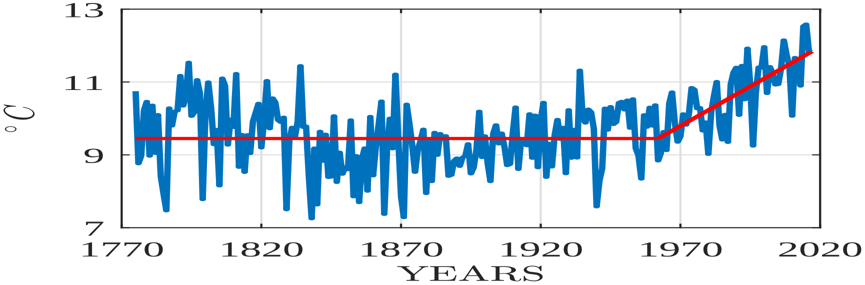
$$W(s) = \begin{cases} W_1(-s) & s < 0 \\ W_2(s) & s \geq 0 \end{cases}$$

$\{W_1(t);\ t \in [0, \infty)\}$ and $\{W_2(t);\ t \in [0, \infty)\}$ are independent standard Wiener processes.

- This approximation is "distribution free", which means that it does not depend on the distribution of error terms
- Distribution of V was independently derived by several authors.

$$P(V \leq x) = \begin{cases} 1 + \sqrt{\frac{x}{2\pi}} e^{-\frac{x}{8}} - \frac{x+5}{2} \Phi(-\frac{1}{2}\sqrt{x}) + \frac{3e^x}{2} \Phi\big(-\frac{3\sqrt{x}}{2}\big) & x \geq 0 \\ 1 - P(V \leq -x) & x < 0 \end{cases}$$

"OPTIMAL" ICE-HOCKEY-STICK MODEL (1962∓11)

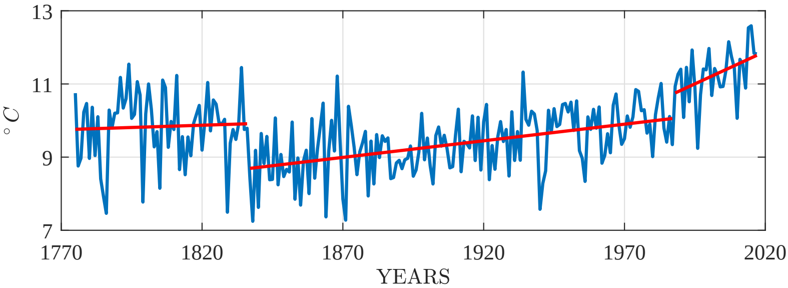


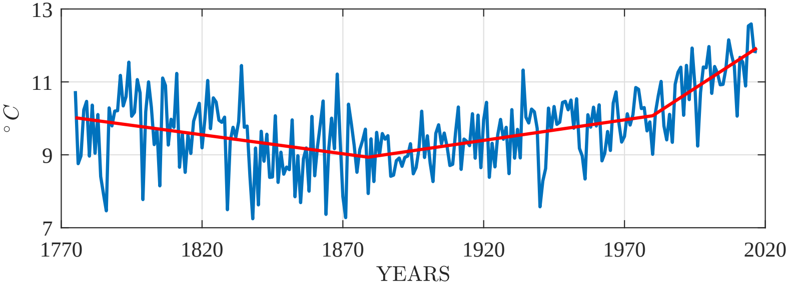
Using approach of Feder (1975), asymptotic distribution of change point $\tau^\star = k^\star/n$ satisfies

$$\sqrt{n}\Big(\frac{\widehat{k}^\star}{n} - \tau^\star\Big) \sim N\Big(0, \frac{1}{\widehat{b}^2}\,\frac{(1+3\tau^\star)}{(1-\tau^\star)\tau^\star}\,\widehat{\sigma}^2\Big).$$

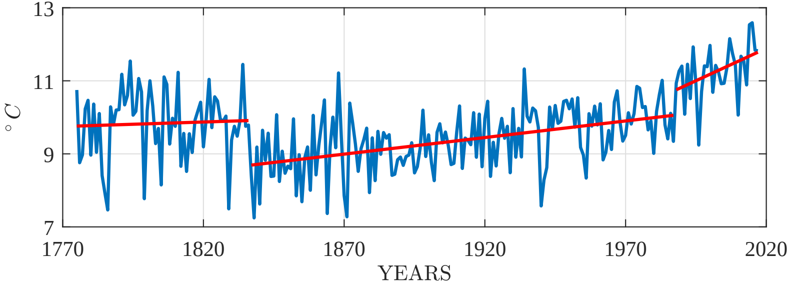
leading to the approximate 95% confidence interval 1962 ∓ 11.

MULTIPLE CHANGES



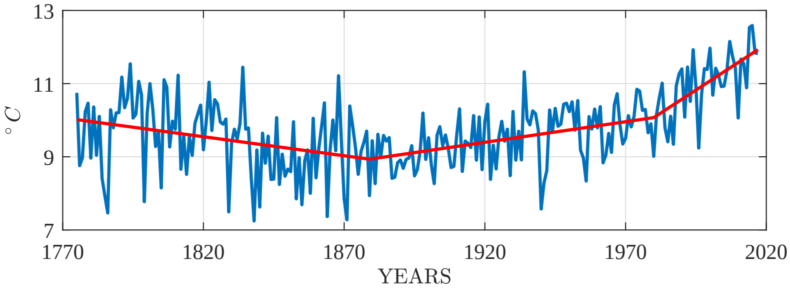


"OPTIMAL" PIECEWISE LINEAR MODEL (1836, 1988)



- Slope parameters for 2^{nd} segment : **0.0091**
  growth of more than two degrees during 1836 – 1987,
- Slope parameters for 3^{rd} segment : **0.0355**
  dramatically quadrupling during last three decades !!!

"OPTIMAL" MODEL WITH TWO GRADUAL CHANGES (1879, 1980)



- Slope parameters for 2^{nd} segment : **0.0217**
  growth of more than two degrees during 1879 – 1980,
- Slope parameters for 3^{rd} segment : **0.0389**
  dramatically doubling during last four decades !!!

SUMMARY

Optimal piecewise linear model with two changes estimated in 1836 and 1988

- Slope parameters for 2^{nd} segment : **0.0091**
  growth of more than two degrees during 1836 – 1987,
- Slope parameters for 3^{rd} segment : **0.0355**
  dramatically quadrupling during last three decades !!!

Optimal model with two gradual changes estimated in 1879 and 1980

- Slope parameters for 2^{nd} segment : **0.0217**
  growth of more than two degrees during 1879 – 1980,
- Slope parameters for 3^{rd} segment : **0.0389**
  dramatically doubling during last four decades !!!

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

MCMC MCMC MCMC

CONSIDERED MODELS

1. Model with a piecewise constant expected value
2. Two-phase linear model with a jump
3. More-phase linear model with jumps
4. Two-phase linear model with gradual change
5. More-phase linear model with gradual changes

In the case of multi-phase segmentation (splitting the data into smaller and more "homogeneous" parts) people usually look for:

- Globally optimal model obtained through the exhaustive search over all possible solutions
- Suboptimal solution obtained typically through recursive splitting of the data

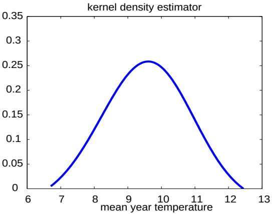
Danger! Setting and fitting the model results in estimated parameters (and models), plots etc. However, this says nothing about the appropriateness of the model.

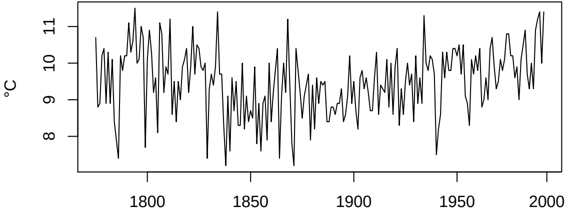
BAYESIAN MODELLING

We assume that observations $Y_1, \ldots, Y_n$ are independent, normally distributed rv's with random parameters and following models with random parameters were assumed.

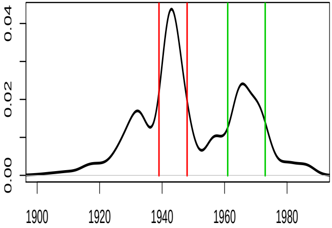
1. Model with a piecewise constant expected value
2. Two-phase linear model with a jump
3. Two-phase linear model with gradual change

Models with more than one change have been also considered.

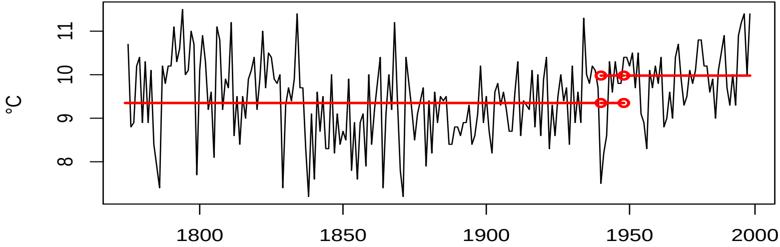




RESULTS FOR MODEL 1 (1775 – 1992)



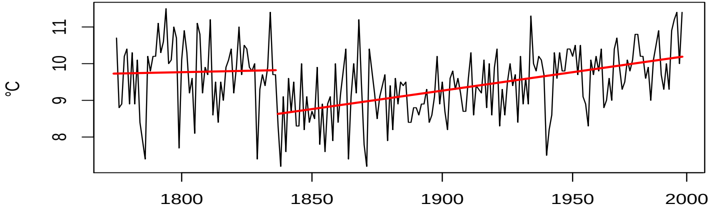
Kernel density estimator of the parameter *r*



RESULTS FOR MODEL 2 (1775 – 1992)

r	60	61	62	63	64	65	else
year	1834	1835	1836	1837	1838	1839	
frequency	566	2868	4870	338	0	75	283
rel. frequency	0.063	0.319	0.541	0.038	0	0.008	0.031

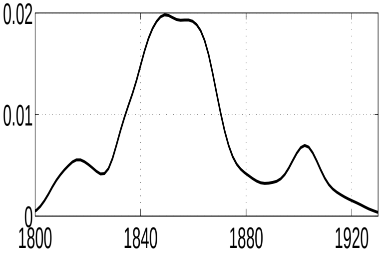
Frequency table of sequence of $r^{(i)}$



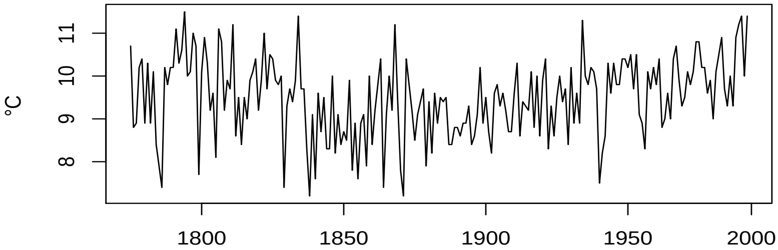
$\widehat{\alpha}_1 \sim 9.717,\ \widehat{\alpha}_2 \sim 8.58,\ \widehat{\beta}_1 \sim 0.0022,\ \widehat{\beta}_2 \sim 0.0102,\ \widehat{\gamma} \sim 1.539$

0.0022 °C per year vs 0.01 °C per year

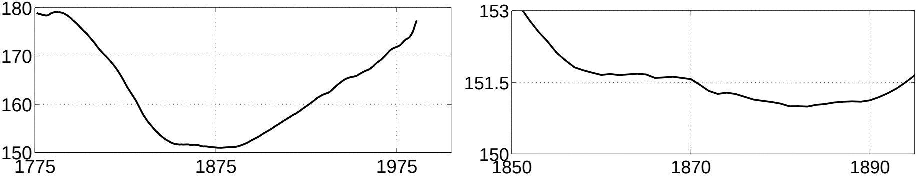
RESULTS FOR MODEL 3 (1775 – 1992)



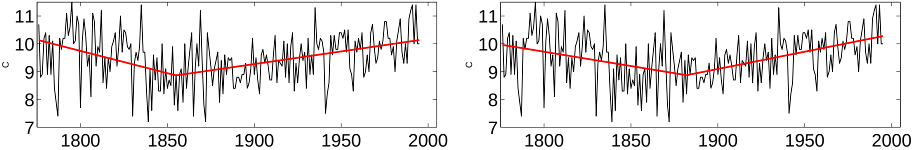
Kernel density estimator of the parameter r



RESULTS FOR MODEL 3 (1775 – 1992)



RSS for all possible changes and its detail for the years 1850 – 1895.



Data and estimated Model 3 for $r = 1855$ (left) and $r = 1882$ (right)

ADDITIONAL QUESTIONS

- Klementinum annual mean temperature series is not stationary
- Applying the same statistical tests, the null hypothesis on stationarity was rejected also for all twelve monthly series
- Supposing that the monthly means follow either an abrupt-shift or an ice-hockey-stick model, we estimated a change point for all individual months

  The results hint that an increase in temperature **starts earlier** for the winter months than for the summer months
- As the monthly mean series are not stationary because they almost certainly exhibit increasing trend in the most recent few decades, a natural question may arise whether increase in temperature has the same character across all calendar days or whether in some period(s) of the year the trend is larger than in others.

  Has the mean annual profile a change in the form of a shift only or is the change more complex?

HOW TO TEST FOR ABRUPT-SHIFT IN THE MEAN OF RANDOM VECTOR?

Suppose that we observe sequence of random vectors $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})^\top, \; i = 1, \ldots, n$, obeying

$$X_{ij} = \mu_{ij} + e_{ij}, \quad i = 1, \ldots, n, \, j = 1, \ldots, p,$$

- $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{ip})^\top$ are vectors of some unknown constants
- $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{ip})^\top$ are zero mean iid rv's with covariance matrix $\boldsymbol{\Sigma}$

We consider a hypothesis testing problem for testing the null hypothesis H_0 against the alternative A

$$\begin{aligned} H_0 &: \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_n \\ A &: \exists\, k \in \{1, \ldots, n-1\} \\ &\quad \boldsymbol{\mu}_1 = \ldots = \boldsymbol{\mu}_k \neq \boldsymbol{\mu}_{k+1} = \ldots = \boldsymbol{\mu}_n \end{aligned}$$

NATURAL TEST STATISTIC

- Assume first situation with k fixed

$$\chi^2(k) = \frac{k(n-k)}{n} \left(\overline{\boldsymbol{X}}(k) - \overline{\boldsymbol{X}}^\star(k)\right)^\top \boldsymbol{\Sigma}^{-1} \left(\overline{\boldsymbol{X}}(k) - \overline{\boldsymbol{X}}^\star(k)\right),$$

  - $\overline{\boldsymbol{X}}(k) = \left(\overline{X}_1(k), \ldots, \overline{X}_p(k)\right)^\top$ is vector of averages over first k years
  - $\overline{\boldsymbol{X}}^\star(k) = \left(\overline{X}_1^\star(k), \ldots, \overline{X}_p^\star(k)\right)^\top$ is vector of averages over last $(n-k)$ years

- Assume next situation with k uknown

$$T = \max_{\lfloor \beta\, n \rfloor \le k \le \lfloor (1-\beta)\, n \rfloor} \chi^2(k)$$

$$TW = \max_{1 \le k \le n} \frac{k(n-k)}{n^2} \chi^2(k)$$

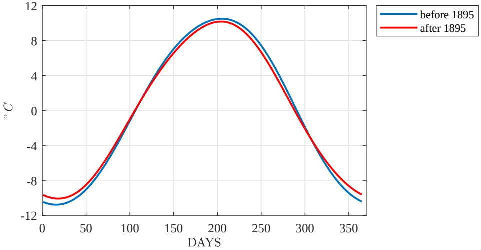
$$MW = \frac{1}{n} \sum_{k=1}^{n} \frac{k(n-k)}{n^2} \chi^2(k)$$

- If the matrix $\boldsymbol{\Sigma}$ is unknown, it may be replaced by any consistent estimate

AND THE RESULTs?

All tests rejected the null hypothesis claiming that an increase of temperature is the same in all days during a calendar year

- The results suggested that increase is larger in winter days than in summer days



Estimated mean annual cycles before and after the year 1895.

- Cycle has mostly been changed in winter periods where the increase in temperature is larger than in the summer periods
- We see again that estimated temperature difference after the year 1895 is smaller than the range before this year

SUMMARY

ZÁVĚR SI JISTĚ KAŽDÝ UDĚLÁ SÁM PODLE SVÉHO POHLEDU NA SVĚT

KDO CHCE ZMĚNU VIDĚT, TAK JI VIDÍ I TAM KDE NENÍ

KDO NECHCE, TAK JI NEVIDÍ I KDYBY TAM BYLA

A TOMU MÉMU ZÁVĚRU BYSTE ASI STEJNĚ NEVĚŘILI