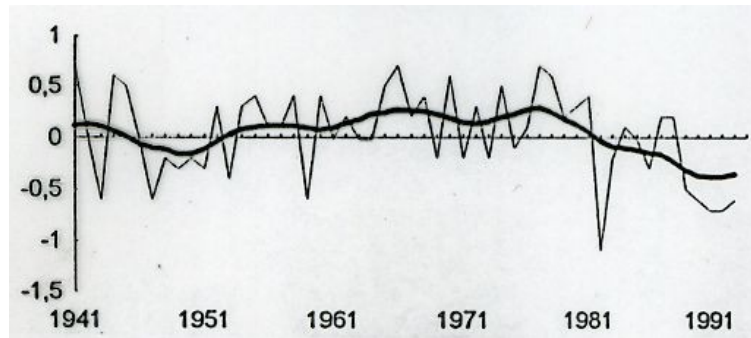
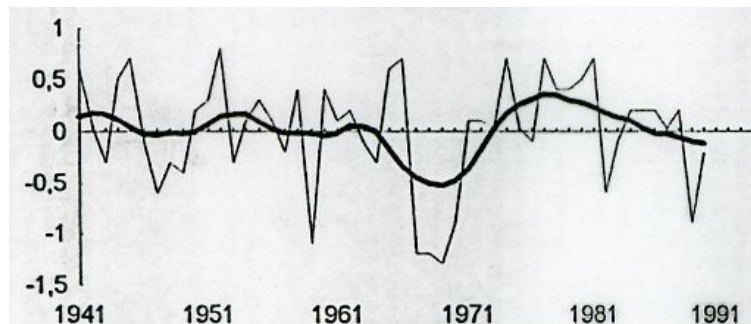


2. Homogenita meteorologických pozorování

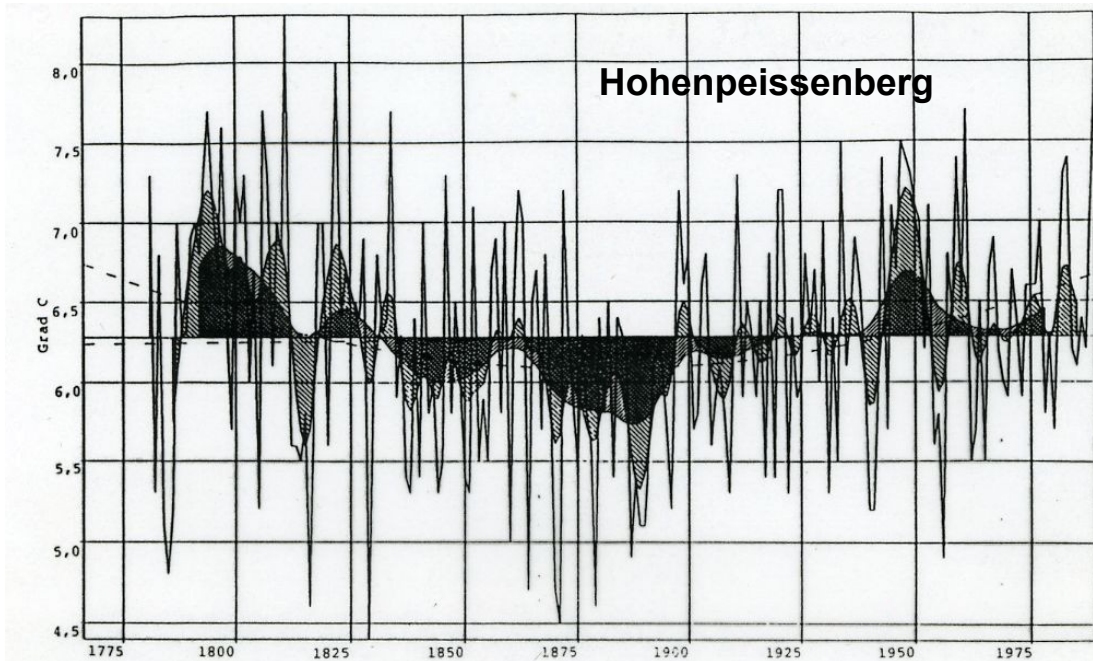
Množství oblačnosti
(desetiny)



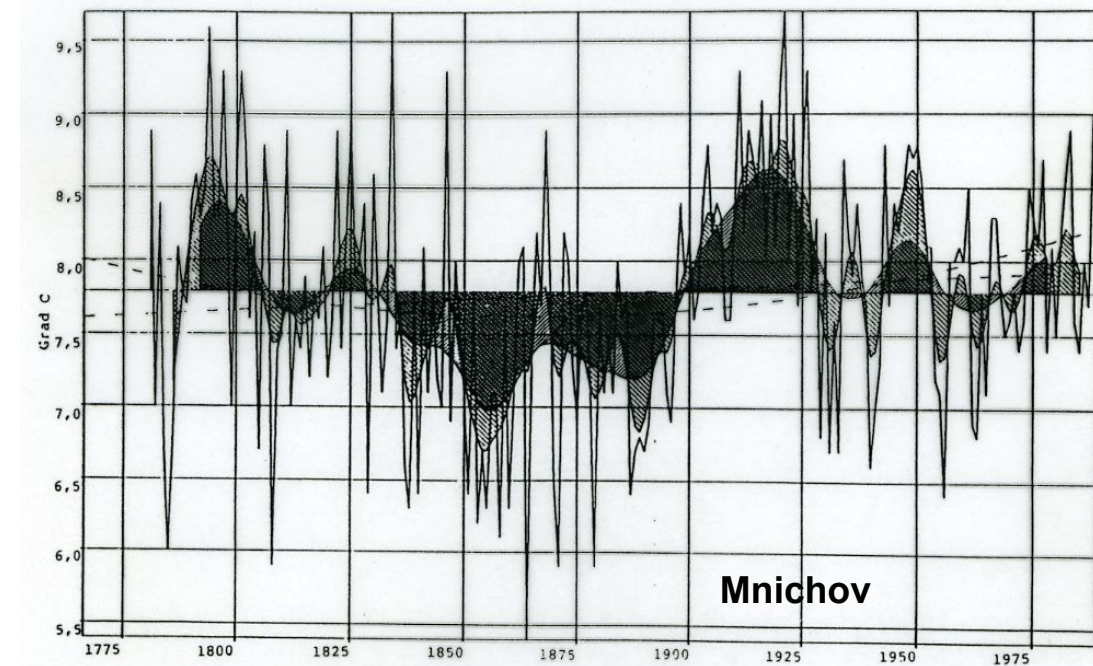
Milešovka
(837 m)



Fichtelberg
(1215 m)

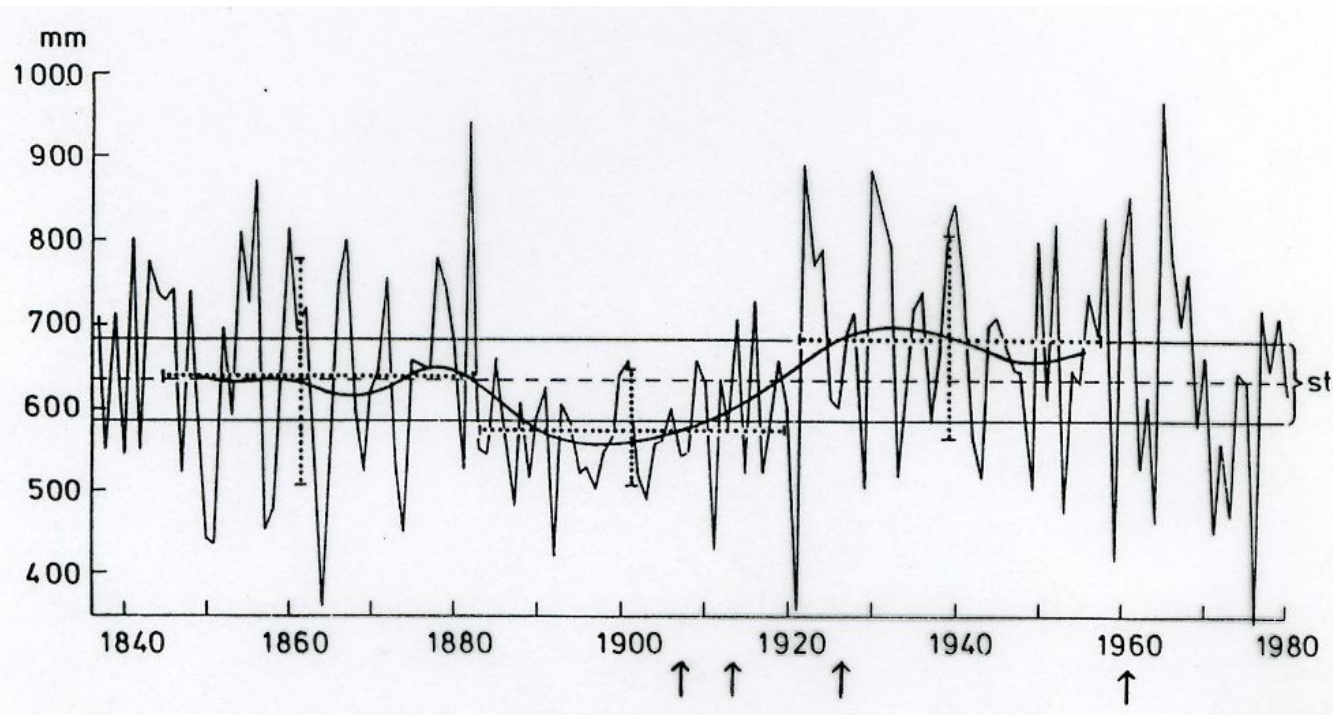


Průměrná roční teplota
vzduchu



G. Müller-Westermeier (1992):
Untersuchungen langer deutscher
Temperaturreihen.

Meteorol. Zeitschrift N.F. 1, č. 3.



Frankfurt nad Mohanem - roční úhrny srážek. Šipky značí roky, v nichž je doloženo přemístění stanice

2.1 Homogenita klimatologických řad

Conrad, W., Pollak, L. W. (1950): Klimatologickou řadu vyjadřující kolísání klimatologického prvku označíme za homogenní, jestliže její kolísání jsou zapříčiněna jen kolísáním počasí a podnebí.

souvisí s měřením a pozorováním na meteorologických stanicích

Zdroje nehomogenit v klimatologických řadách

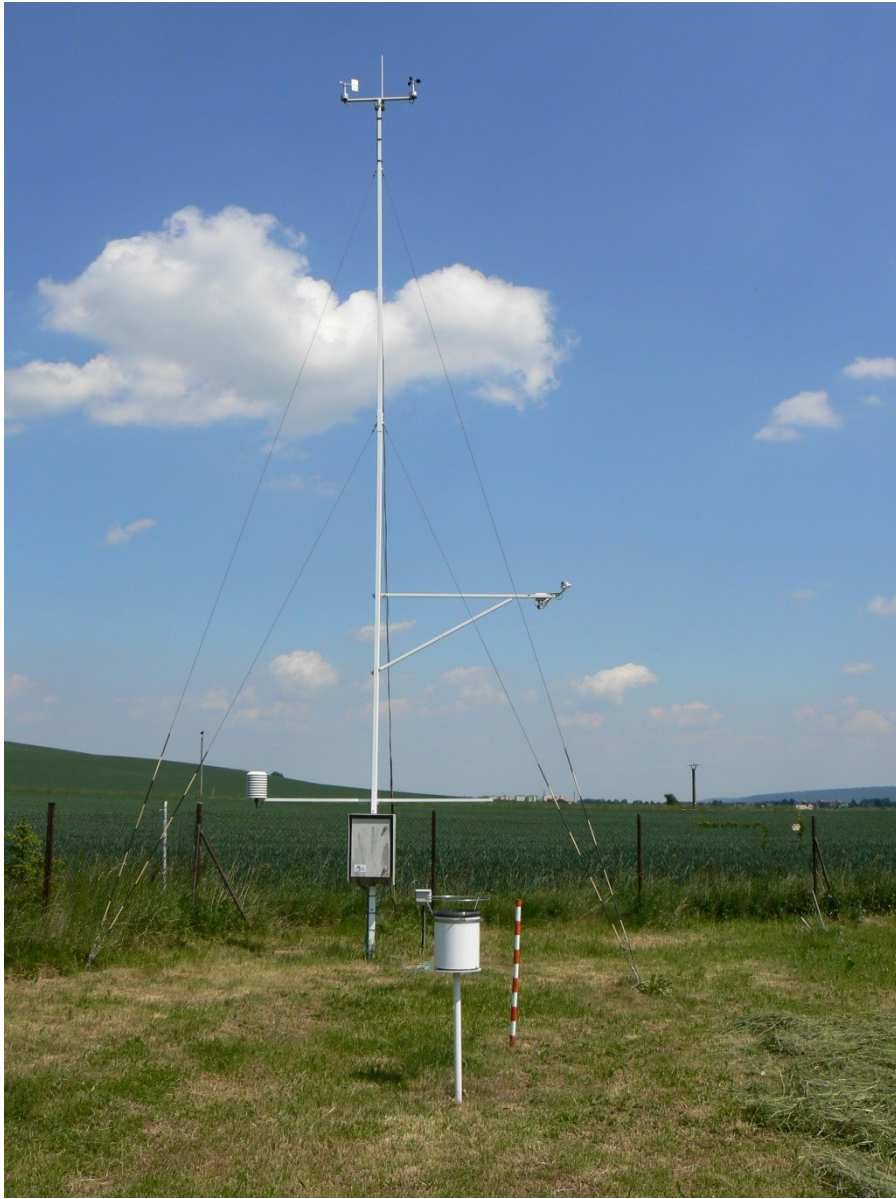
- přemístění stanice
- přístroje a jejich umístění (standardní x automatizovaná měření, meteorologická budka, typy přístrojů)
- metodika a doba pozorování
- okolí stanice
- pozorovatelé

Význam **metadat** (podle dokumentace stanice) pro potvrzení nehomogenit zjištěných testováním.

Náhodné a systematické chyby měření.

Přerov, letiště

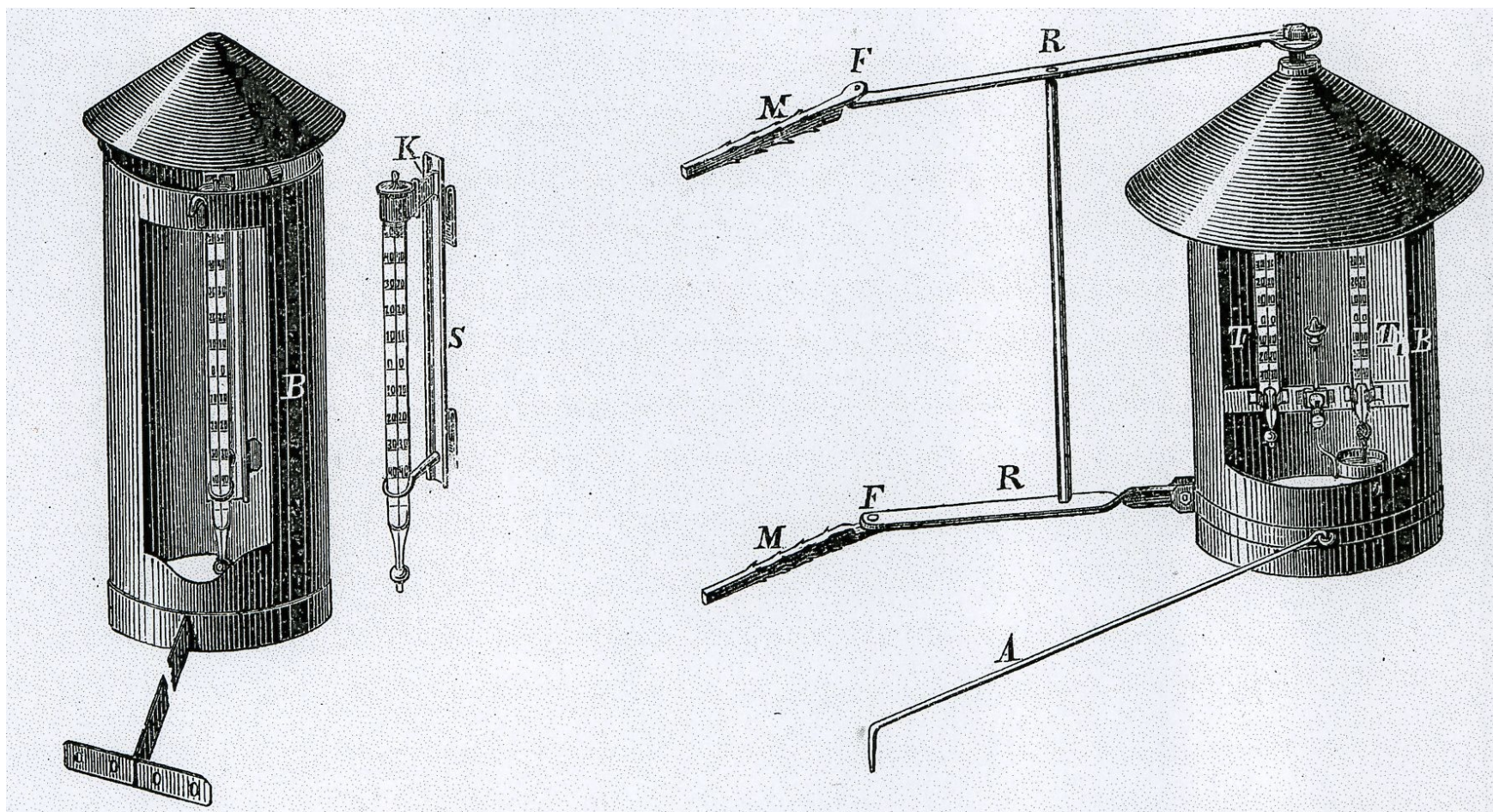




Opava-Otice



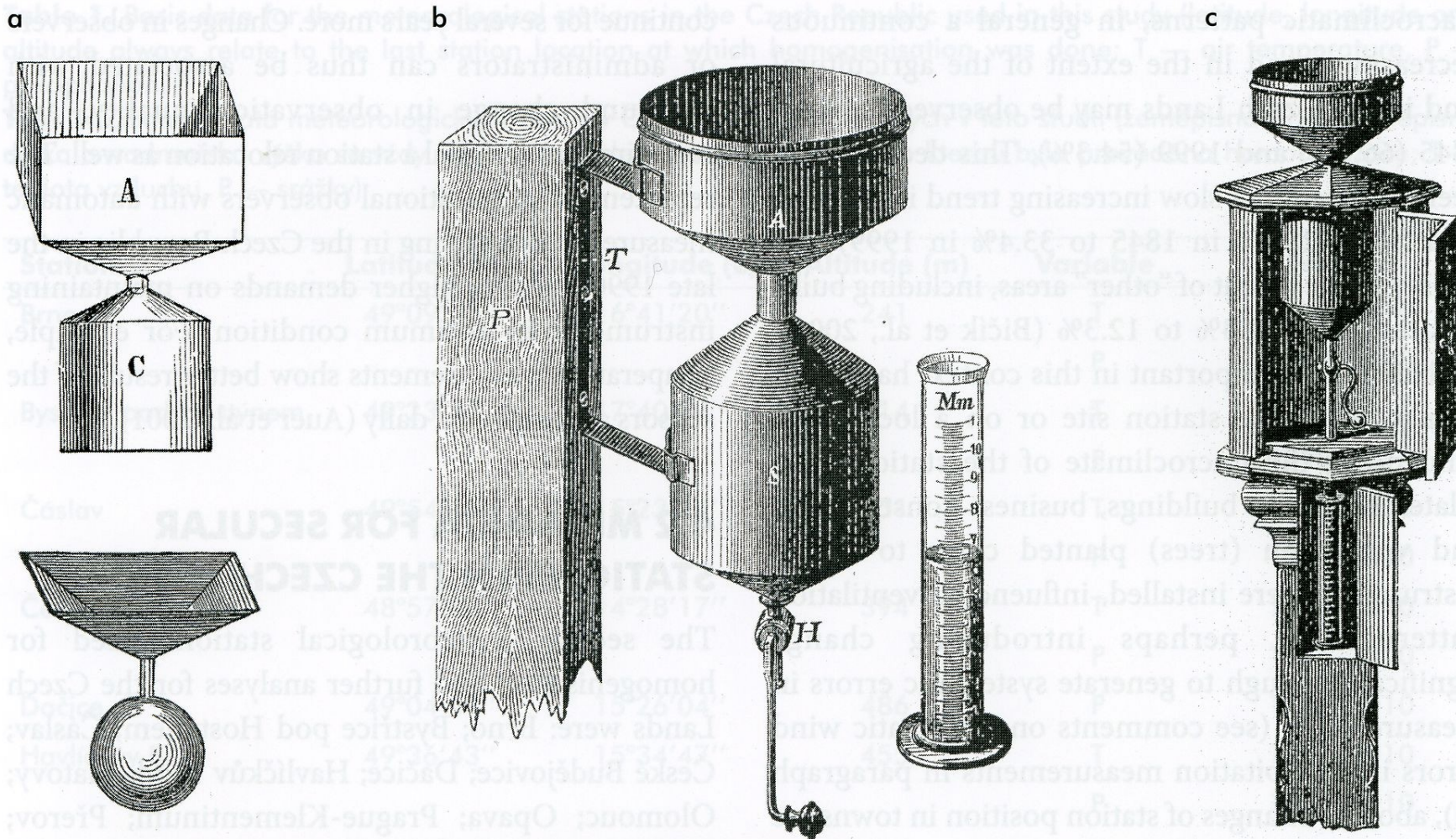
Olomouc-Holice



Doporučená malá a velká kovová budka pro teploměry, umístěvaná na severní stranu objektů, podle rakouské instrukce pro pozorovatele meteorologických stanic (po roce 1851)



Modifikovaná Stevensonova budka používaná na stanicích v síti ČHMÚ



Srážkoměry podle rakouské instrukce pro pozorovatele meteorologických stanic: a) se čtvercovou záchytnou plochou, b) Osnaghi – starší verze a c) v dřevěné skříňce



Srážkoměry používané v síti stanic ČHMÚ: vlevo – Metra 886, uprostřed – ombrograf, vpravo – automatický srážkoměr (foto J. Záruba)

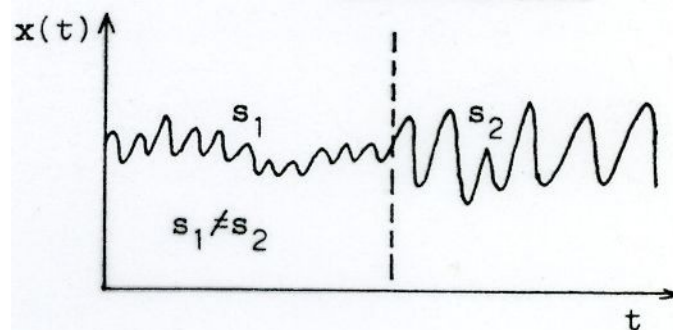
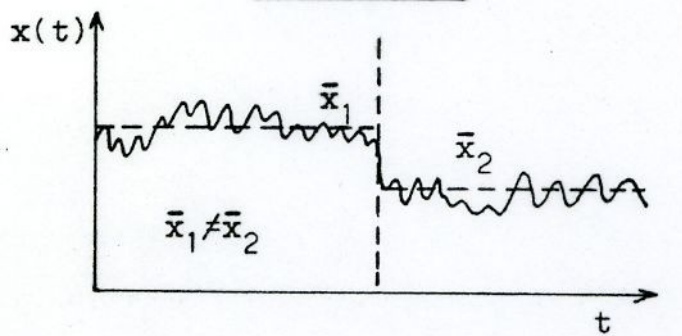


Nestandardní umístění přístrojů na sekulární stanici Praha-Klementinum

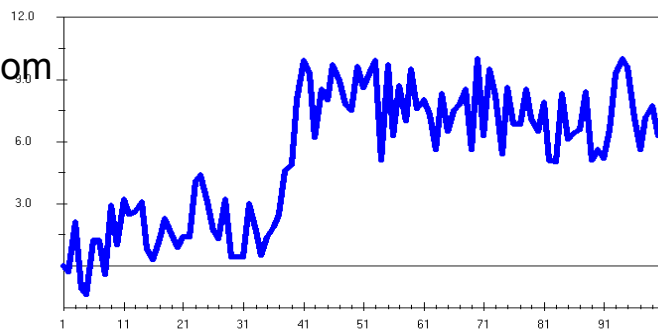
Nehomogenita

v průměru

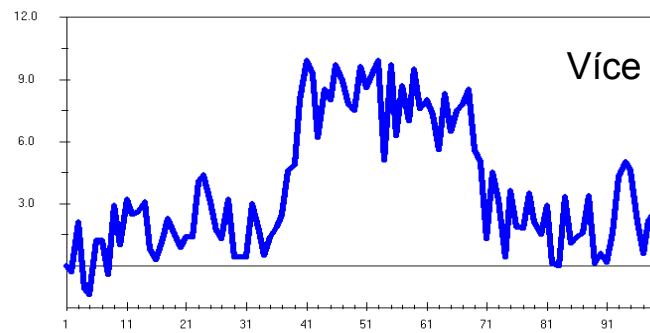
v rozptylu



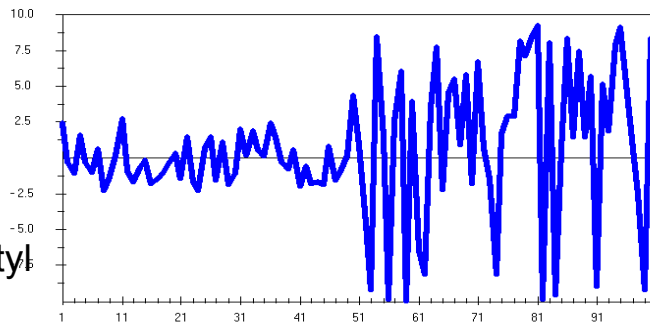
Jeden zlom



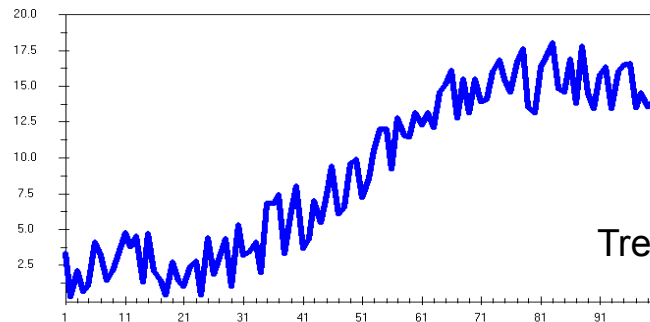
Více zlomů



Rozptyl



Trend



2.1.1 Relativní homogenita

Conrad, W., Pollak, L. W. (1950):

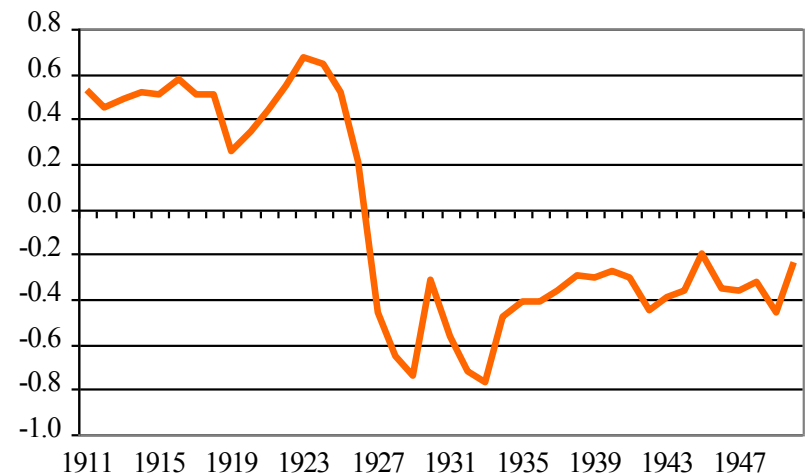
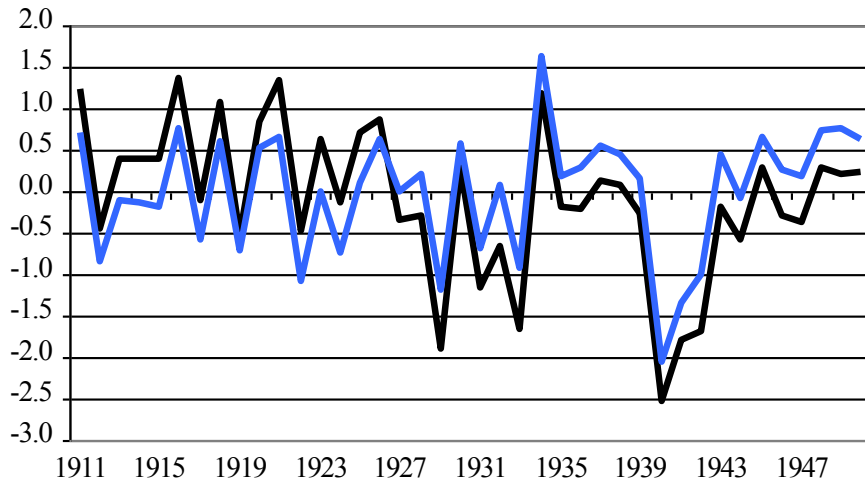
Klimatologická řada je relativně homogenní vzhledem k synchronní řadě jiného místa (homogenní), jestliže difference (teplota vzduchu), popř. podíly (srážky) odpovídajících si dvojic hodnot tvoří řadu náhodných čísel, která vyhovuje zákonu chyb.

$x(t_i)$, $i = 1, 2, \dots, n$ – homogenní srovnávací (referenční) řada

$y(t_i)$, $i = 1, 2, \dots, n$ – testovaná řada

Teplota vzduchu: $d_i = y(t_i) - x(t_i)$

Srážky: $q_i = y(t_i) / x(t_i)$



2.2 Homogenizace klimatologických řad

- výběr referenční řady
- zjištění nehomogenit (testy relativní homogenity)
- homogenizace a doplnění chybějících hodnot

2.2.1 Výběr referenční řady

- dostatečný stupeň podobnosti mezi referenční a testovanou řadou ($r_{xy} > 0,70$)
- referenční řada:
 - a) řada jedné homogenní stanice
 - b) aritmetický průměr z vybraných stanic (tj. ze stanic vybraných podle korelací nebo podle vzdáleností)
 - c) průměrná řada (aritmetický průměr ze všech stanic v dané oblasti)

2.2.2 Testy relativní homogenity

Obecný postup testování:

- podmínky aplikovatelnosti testu (nezávislost hodnot, normální rozdělení atd.)
- nulová a alternativní hypotéza (H_0 , H_1)
- výpočet testovacího kritéria s určitým teoretickým rozdělením (μ , σ)
- nalezení kritické hodnoty
- porovnání testovacího kritéria a kritické hodnoty – přijetí nebo zamítnutí H_0 (H_1) na dané hladině významnosti

- testování náhodnosti hodnot řady diferencí (kvocientů)

Helmert's criterion:

S - number of sequences ++ or --

C - number of sequences -+ or +-
$$-\sqrt{n-1} \leq S - C \leq \sqrt{n-1}$$

Abbe's criterion:

$$A = \sum_{i=1}^n d_i'^2 \quad B = \sum_{i=1}^n (d_i' - d_{i+1}')^2 \quad d_i' = d_i - \bar{d}$$

$$1 - \sqrt{\frac{1}{n}} \leq 2A/B \leq 1 + \sqrt{\frac{1}{n}}$$

- testování náhodnosti hodnot řady diferencí (podílů), majících přibližně normální rozdělení

Autocorrelation coefficient $r_x(1)$:

$r_x(1) > r(p)$ - non-randomness

$$r(p) = \frac{-1 + z(p) \sqrt{n-2}}{n-1}$$

The von Neumann ratio:

$V > V(p)$ - randomness

$$V = \frac{n}{n-1} \cdot \frac{\sum_{i=1}^{n-1} (d_i - d_{i+1})^2}{\sum_{i=1}^n d_i^2 - \frac{1}{n} \left(\sum_{i=1}^n d_i \right)^2}$$

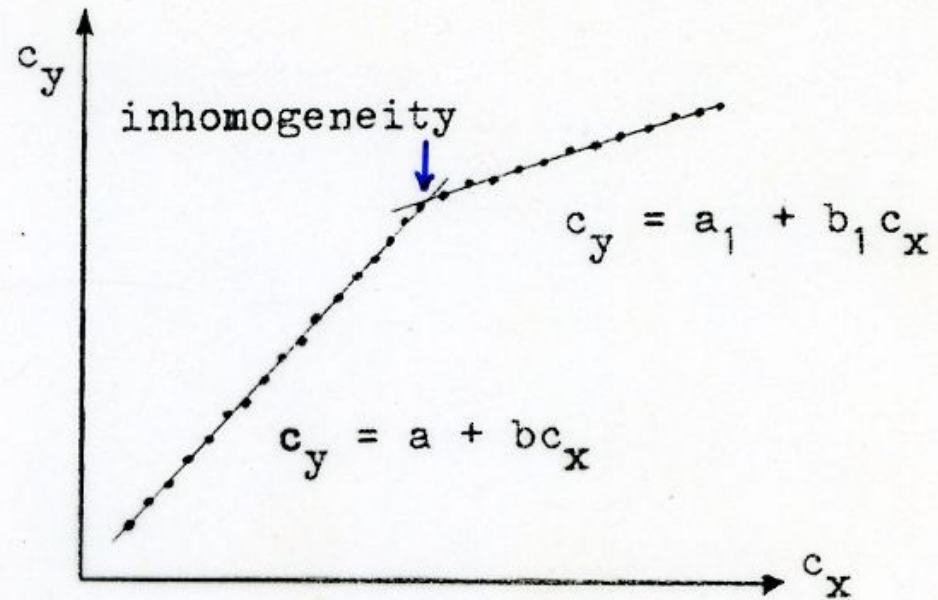
$$V(p) = \frac{2n + 2z(p) \sqrt{n-2}}{n-1}$$

Metoda dvojné součtové čary (double-mass analysis)

$$c_{x_k} = \sum_{k=1}^n x_k$$

$$c_{y_k} = \sum_{k=1}^n y_k$$

$$k = 1, 2, \dots, n$$

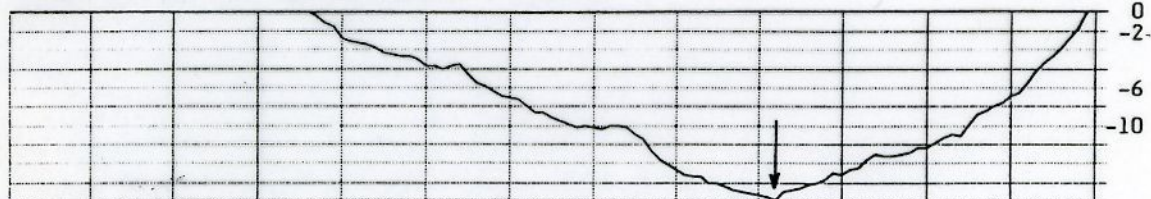


Craddock's test

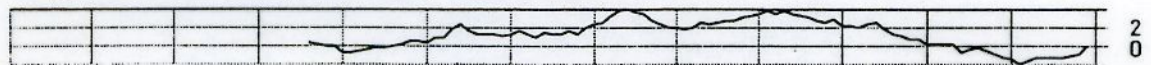
cumulative deviations $s_k = d_1 + d_2 + \dots + d_k, k = 1, 2, \dots, n$

$$d_k = \sum_{i=1}^k (c x_i - y_i) \quad c = \bar{y}/\bar{x}$$

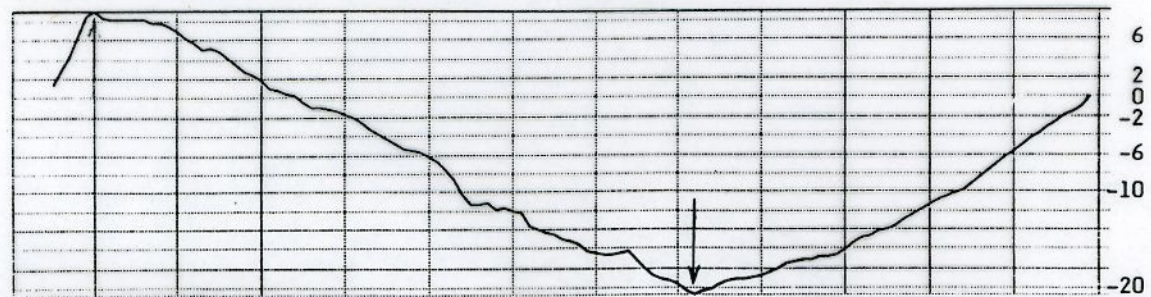
Vergleich GALTÜR (unreduziert) - SÄNTIS



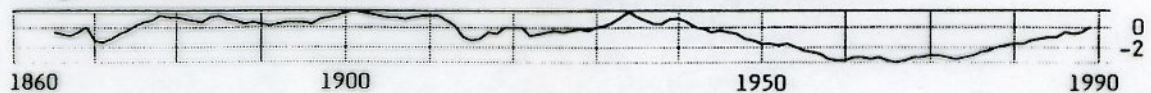
Vergleich GALTÜR (reduziert) - SÄNTIS



Vergleich REICHENAU (unreduziert) - REGION UNTER 1500 M



Vergleich REICHENAU (reduziert) - REGION UNTER 1500 M



1860

1900

1950

1990

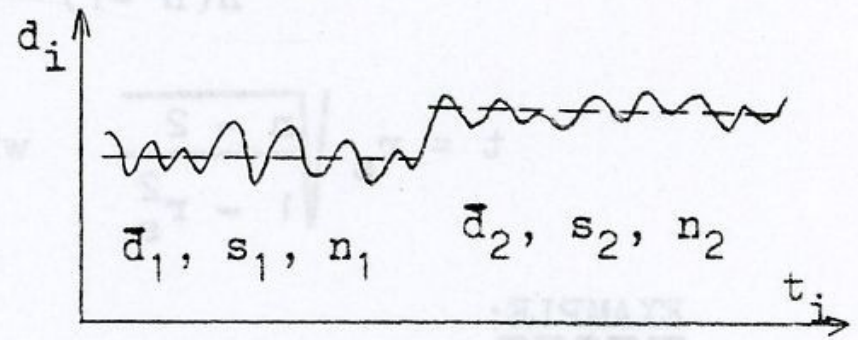
Studentův t-test (testování statistické významnosti průměrů)

- náhlá změna se znalostí roku nehomogenity

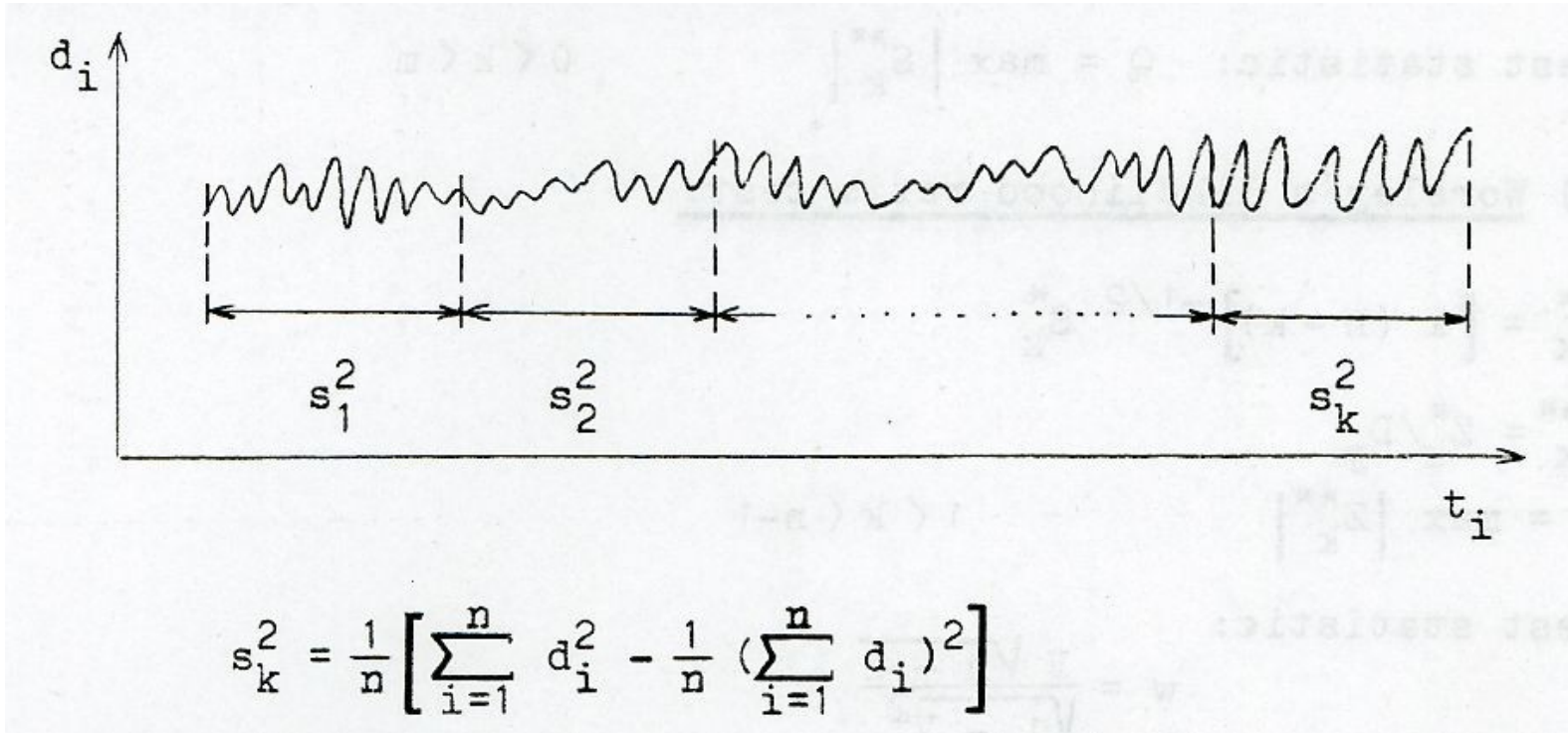
$$t_d = \frac{\bar{d}_1 - \bar{d}_2 - (\mu_1 - \mu_2)}{\left[\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right]^{1/2}}$$

$$n_1 s_1^2 = \sum_{i=1}^{n_1} d_i^2 - \frac{1}{n_1} \left(\sum_{i=1}^{n_1} d_i \right)^2$$

with t-distribution and $\nu =$
 $= n_1 + n_2 - 2$



Bartlettův test – testování homogenity řady v rozptylu



Poměr s_{\max}^2 / s_{\min}^2 se porovnává s tabelovanými kritickými hodnotami.

Testy náhodnosti vzhledem k trendu

Mann-Kendallův test

- difference d_i jsou nahrazeny jejich pořadím k_i ($i = 1, 2, \dots, n$)
- pořadí k_i se porovnává s následujícími pořadími k_{i+1}, \dots, k_n
- počítá se počet následných pořadí n_i , jejichž hodnota přesahuje k_i

$$\tau = \frac{4P}{n(n-1)} - 1$$

$$P = \sum_{i=1}^{n-1} n_i$$

$$\tau(p) = 0 \pm z(p) \sqrt{\frac{4n+10}{9n(n-1)}}$$

EXAMPLE:

Year	i	d_i	k_i	n_i	$c_i = k_i - i$
1935	1	-0.4	6	13	5
	2	1.1	15	5	13
	3	1.9	17	3	14
	4	-1.2	2	15	-2
	5	-1.0	3	14	-2
1940	6	1.0	14	4	8
	7	2.0	18	2	11
	8	0.6	9	7	1
	9	0.2	8	7	-1
	10	-0.9	4	9	-6
1945	11	-0.8	5	8	-6
	12	0.9	12.5	3.5	0.5
	13	0.1	7	6	-6
	14	0.9	12.5	3	-1.5
	15	0.8	11	3	-4
1950	16	2.1	19	1	3
	17	1.2	16	1	-1
	18	0.7	10	1	-8
	19	2.8	20	0	1
1954	20 = N	-1.4	1	-	-19

$P=105.5$

$$\tau = \frac{4 \cdot 105.5}{20 \cdot 19} - 1 = +0.111$$

$$\tau(p) = 0 \pm 1.96 \sqrt{\frac{80 + 10}{180 \cdot 19}} = \pm 0.318$$

$0.111 \in (-0.318; 0.318)$

differences d_i are random

Spearmanův test pořadové statistiky

$$r_s = 1 - \frac{6 \sum_{i=1}^n c_i^2}{n(n^2-1)}$$

$$c_i = k_i - i$$

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}} \quad \text{with } t\text{-distribution and } \nu = n - 2$$

EXAMPLE:

$$r_s = 1 - \frac{6 \cdot 1147.5}{20 \cdot 399} = +0.137$$

$$t = 0.137 \sqrt{\frac{18}{1 - (0.137)^2}} = +0.587$$

$$0.587 \in (-2.101; 2.101)$$

differences d_i are random

Testy skokové změny v průměru

Buishandův test kumulativních odchylek

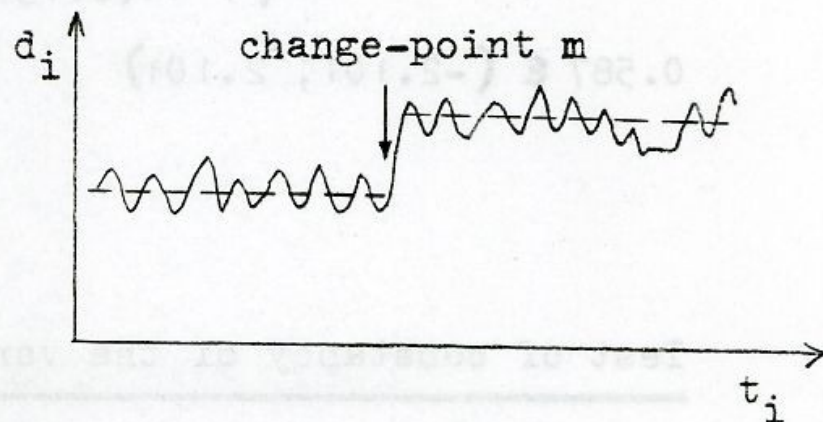
$$S_0^* = 0, S_k^* = \sum_{i=1}^k (d_i - \bar{d})$$

$$k = 1, 2, \dots, n$$

$$D_y^2 = \sum_{i=1}^n (d_i - \bar{d})^2 / n$$

$$S_k^{**} = S_k^* / D_y, k = 0, 1, \dots, n$$

$$\text{Test statistic: } Q = \max |S_k^{**}| \quad 0 < k < m$$



Worsleyho test pravděpodobnostního poměru

$$Z_k^* = [k(n-k)]^{-1/2} S_k^*$$

$$Z_k^{**} = Z_k^* / D_y$$

$$V = \max |Z_k^{**}| \quad 1 < k < n-1$$

Test statistic:

$$w = \frac{V \sqrt{n-2}}{\sqrt{1-V^2}}$$

Poloha maxim $|S_k^{**}|$ nebo $|Z_k^{**}|$
odpovídá bodu změny m

Test Maronny a Yohaie pro detekci posunu v průměru

$H_0: \{x_i, y_i\}$ - the same bivariate normal distribution $N\{\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho\}$

$H_1: \{x_i, y_i\}$ - for $i \leq i_0$ ($0 < i_0 < n$) normal distribution $N\{\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho\}$

- for $i > i_0$ ($0 < i_0 < n$) normal distribution $N\{\mu_x, \mu_y + d, \sigma_x^2, \sigma_y^2, \rho\}$, $d \neq 0$

$$X_i = 1/i \sum_{j=1}^i x_j \quad X_n = \bar{x}$$

$$Y_i = 1/i \sum_{j=1}^i y_j \quad Y_n = \bar{y}$$

$$S_x = \sum_{j=1}^n (x_j - \bar{x})^2 \quad S_y = \sum_{j=1}^n (y_j - \bar{y})^2$$

$$S_{xy} = \sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})$$

$$F_i = S_x - (X_i - \bar{x})^2 ni / (n - i) \quad i < n$$

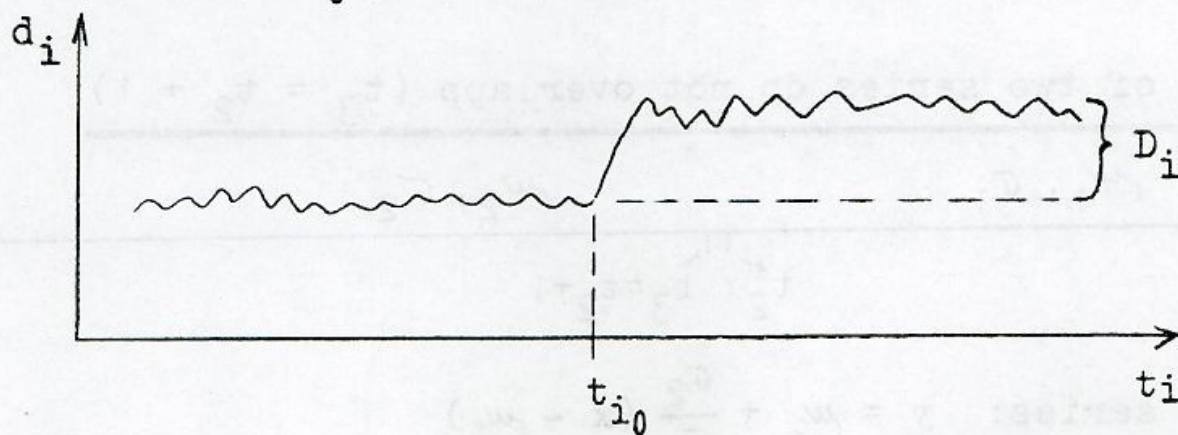
$$D_i = [S_x (\bar{y} - Y_i) - S_{xy} (\bar{x} - X_i)] n / (n - i) F_i$$

$$T_i = [i(n - i) D_i^2 F_i] / (S_x S_y - S_{xy}^2)$$

The test statistic:

$$T_0 = \max_{i < n} \{T_i\}$$

H_0 is reject when T_0 is larger than constant k (tabulated)



Tab. 2 Kritické hodnoty k pro statistiku T_0 , pro rozsah řady n a hladinu významnosti 0,05 [15].

n	10	15	20	30	40	70	100
k	6.8	7.4	7.8	8.2	8.7	9.3	9.3

Alexanderssonův test homogenity pro jednoduchý zlom (Standard Normal Homogeneity Test)

Nejprve se vytvoří řada poměrů (v případě srážek) nebo diferencí (v případě teplot vzduchu) mezi testovanou a referenční řadou (možný způsob konstrukce referenční řady udává např. [1, 2]). Tato řada $\{q_i\}$ je následně standardizována:

$$z_i = \frac{(q_i - \bar{q})}{s_q}$$

kde \bar{q} je hodnota aritmetického průměru poměrů nebo diferencí $\{q_i\}$, s_q je směrodatná odchylka této řady (s váhou $n-1$, což ovlivňuje hodnotu testovacího kritéria a hladinu významnosti).

Nulová hypotéza je definovaná následovně [2]:

$$H_0: z_i \in N(0,1), i \in \{1, \dots, n\}.$$

Alternativní hypotéza:

$$H_1: z_i \in N(\mu_1, 1), i \in \{1, \dots, a\}.$$

$$z_i \in N(\mu_2, 1), i \in \{a+1, \dots, n\}.$$

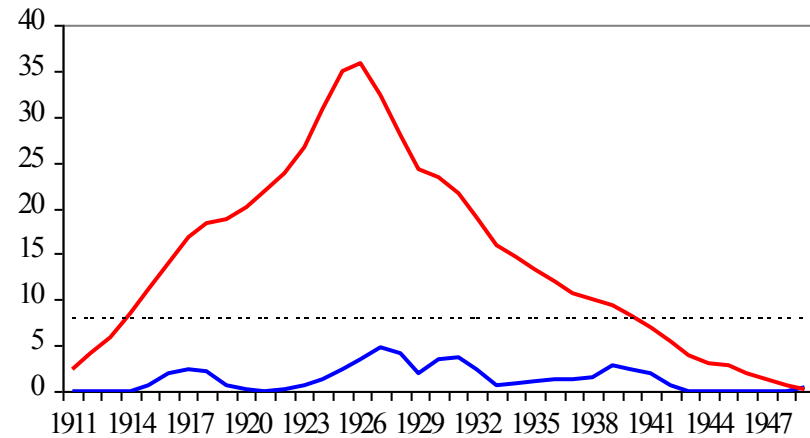
kde $1 \leq a < n$ a $\mu_1 \neq \mu_2$, $z_i \in N(0,1)$ značí, že řada z_i má normální rozdělení s nulovým průměrem a jednotkovou směrodatnou odchylkou.

Testovací kritérium T_0 se získá ze vztahu:

$$T_0 = \max_{1 \leq a < n-1} \{T_a\} = \max_{1 \leq a < n-1} \left\{ a\bar{z}_1^2 + (n-a)\bar{z}_2^2 \right\}$$

kde: $\bar{z}_1 = \frac{1}{a} \sum_{i=1}^a z_i, (\bar{z}_1 \approx \mu_1)$

$$\bar{z}_2 = \frac{1}{n-a} \sum_{i=a+1}^n z_i, (\bar{z}_2 \approx \mu_2).$$



Hodnota a určuje rok, ve kterém s největší pravděpodobností došlo ke zlomu. Přesněji řečeno, je to poslední rok s prvním průměrem μ_1 .

Jestliže hodnota T_0 je větší než určitá kritická hodnota (tab. 1), řadu hodnotíme jako nehomogenní na dané hladině významnosti. Hodnoty T_a můžeme dále vynést do grafu a podle nich potom soudit na charakter řady $\{q_i\}$.

Poměr \bar{q}_2 / \bar{q}_1 , resp. rozdíl $\bar{q}_2 - \bar{q}_1$, kde \bar{q}_1, \bar{q}_2 jsou průměrné hodnoty řady poměrů (diferenci) před a po možném zlomu, udávají hodnotu opravy testované řady.

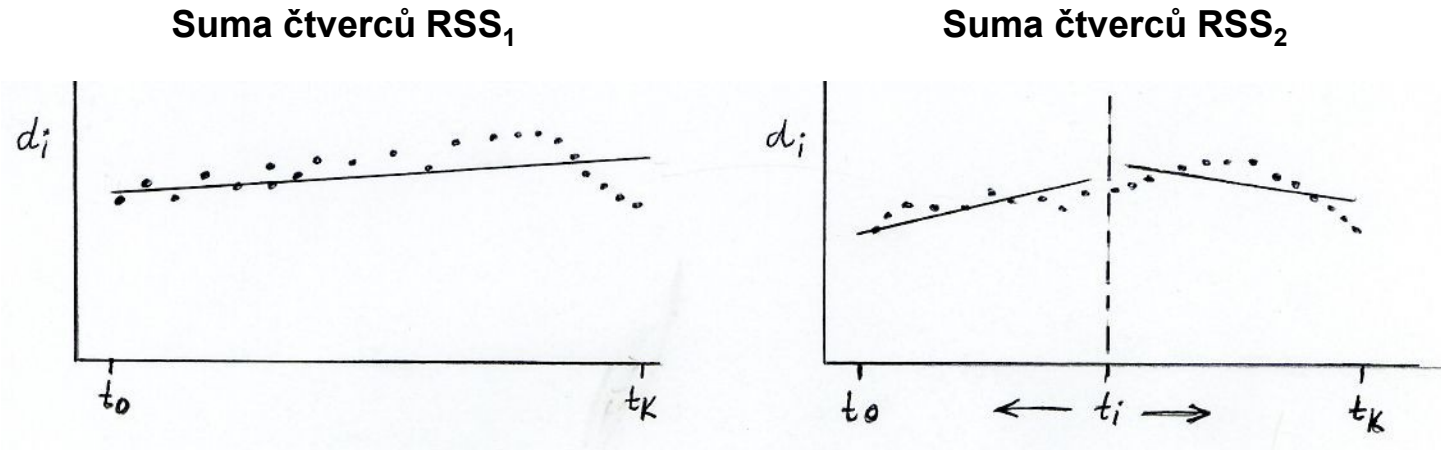
Nevýhodou testu je, že rok nejpravděpodobnější změny má tendenci se vyskytovat blízko začátku a konce řady. V testu není ani specifikováno rozdělení samotných řad. Prostě se předpokládá, že poměry (rozdíly) mají alespoň přibližně normální rozdělení. Je proto dobré řady, které nemají normální rozdělení, na toto rozdělení transformovat.

Uvedený test může být aplikován pouze na jednoduchý zlom (posun v průměru).

Tab. 1 Kritické hodnoty k Alexanderssonova testu pro rozsah řady n a hladinu významnosti 0,05 [1].

n	10	20	40	60	80	100	150	200
k	5.7	7.0	8.1	8.7	9.0	9.2	9.4	9.6

Test Easterlinga a Petersona



místo s nejmenší hodnotou RSS_2 – možná nehomogenita

testovací kritérium:

$$U = [(RSS_1 - RSS_2) / 3] / [RSS_2 / (n - 4)]$$

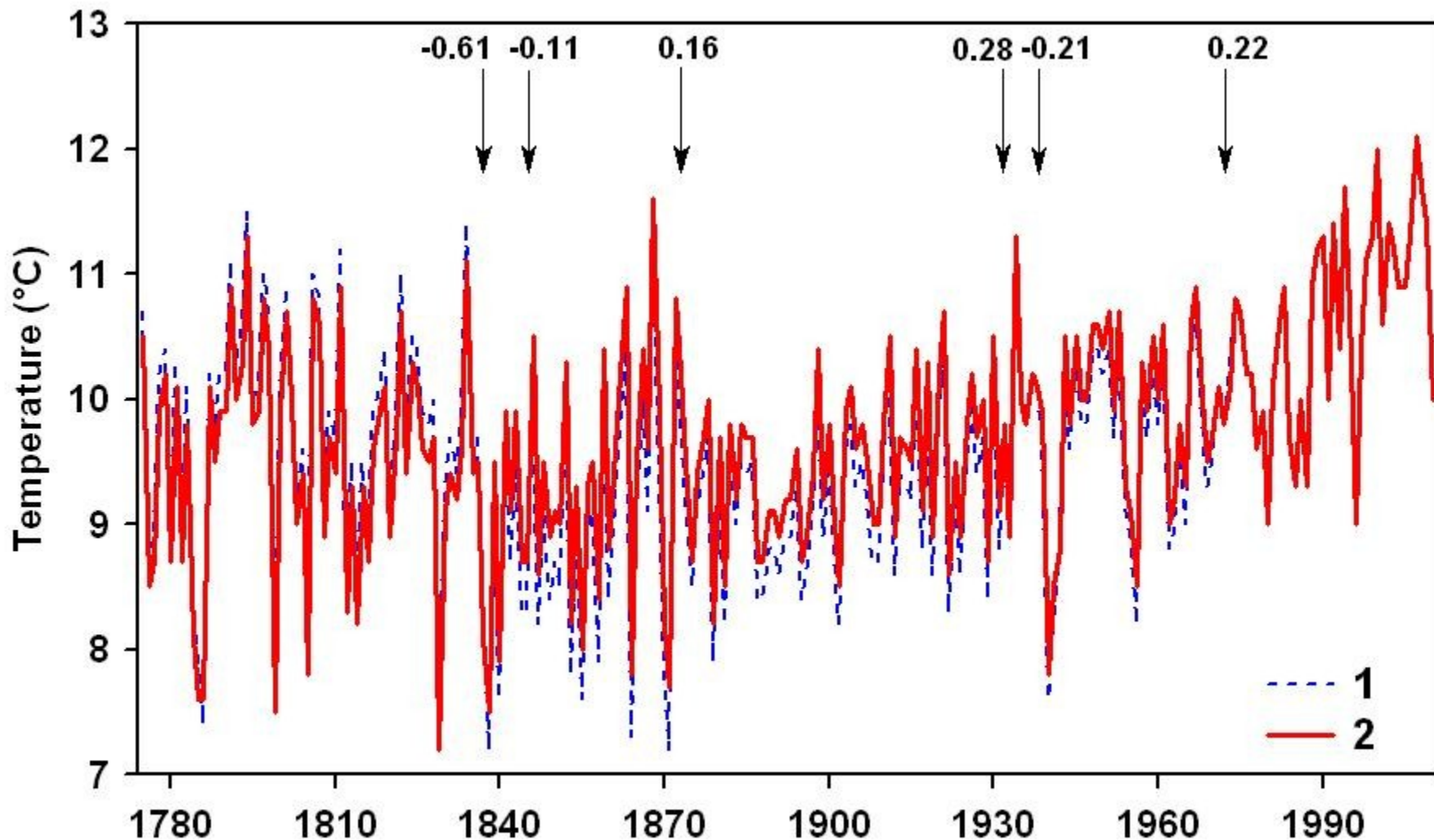
má F rozdělení s 3 a $n - 4$ stupni volnosti

Pomocí t-testu se zjišťuje, zda je statisticky významný rozdíl v průměrech dané části řady před a po možné nehomogenitě

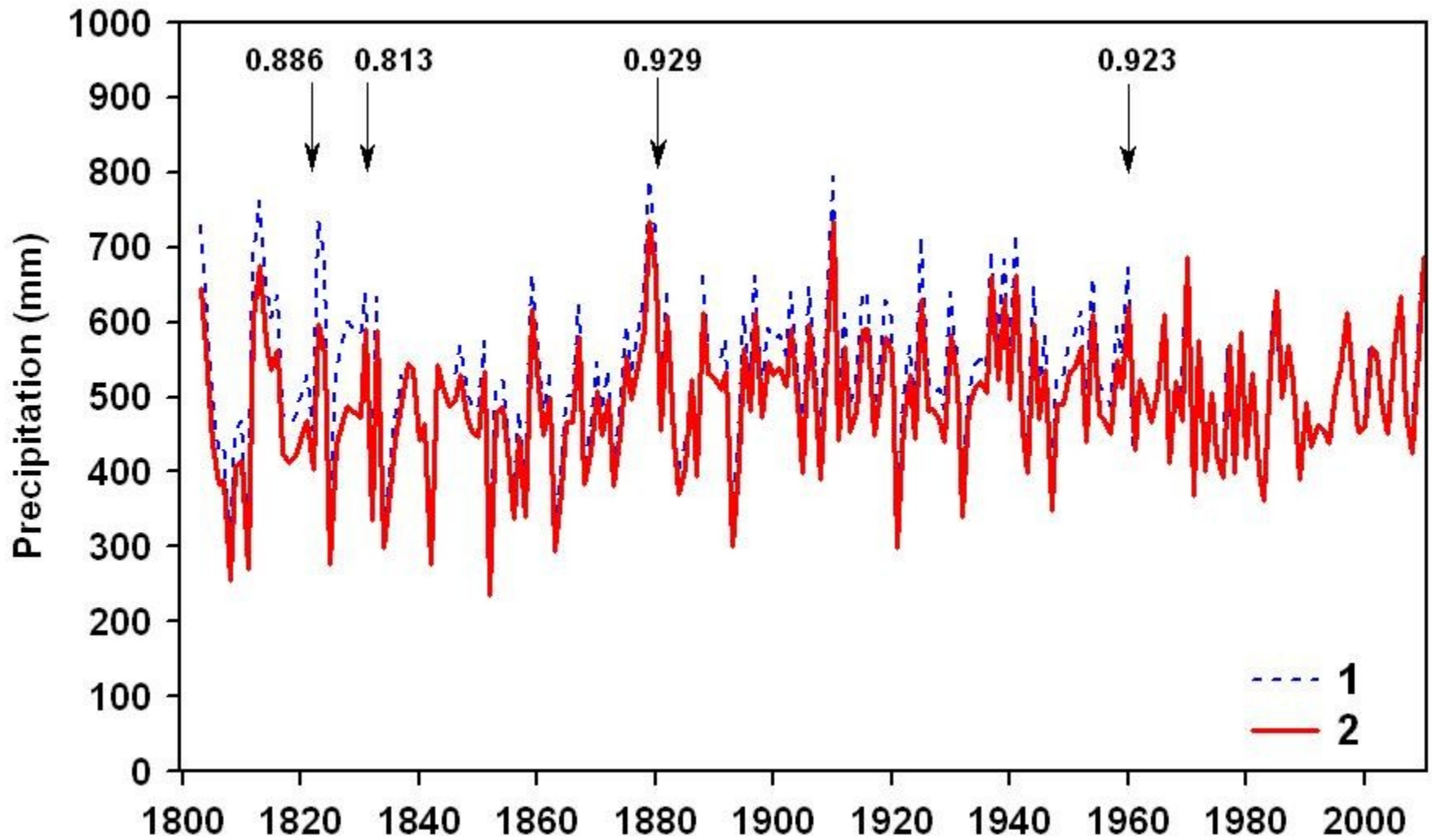
2.2.3 Homogenizace a doplnění chybějících hodnot

- homogenizace řady se provádí na současné období měření (tj. do minulosti)
- opravují se **statisticky významné nehomogenity** vysvětlitelné metadaty, popř. tzv. nezpochybnitelné nehomogenity (projevující se ve větším počtu období a logicky ukazující na stejnou příčinu)
- nehomogenní hodnoty testované řady se opravují podle průměrných diferencí (kvocientů), odvozených ze společného období testované a referenční stanice
- chybějící hodnoty se doplňují podle referenční řady se zřetelem na příslušnou opravu mezi referenční a testovanou řadou

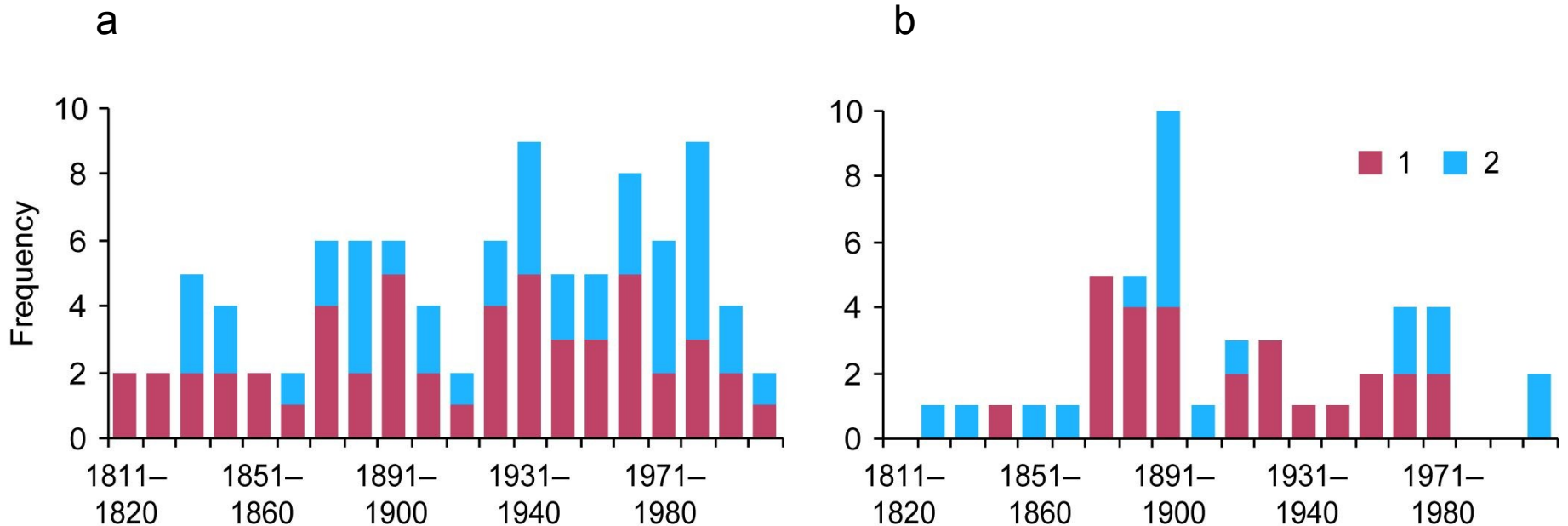
Poznámka: chybějící hodnoty v řadě mohou být doplněny různými technikami před homogenizací (např. regresní analýzou, metodou diferencí či kvocientů), což ale může být zdrojem nehomogenity.



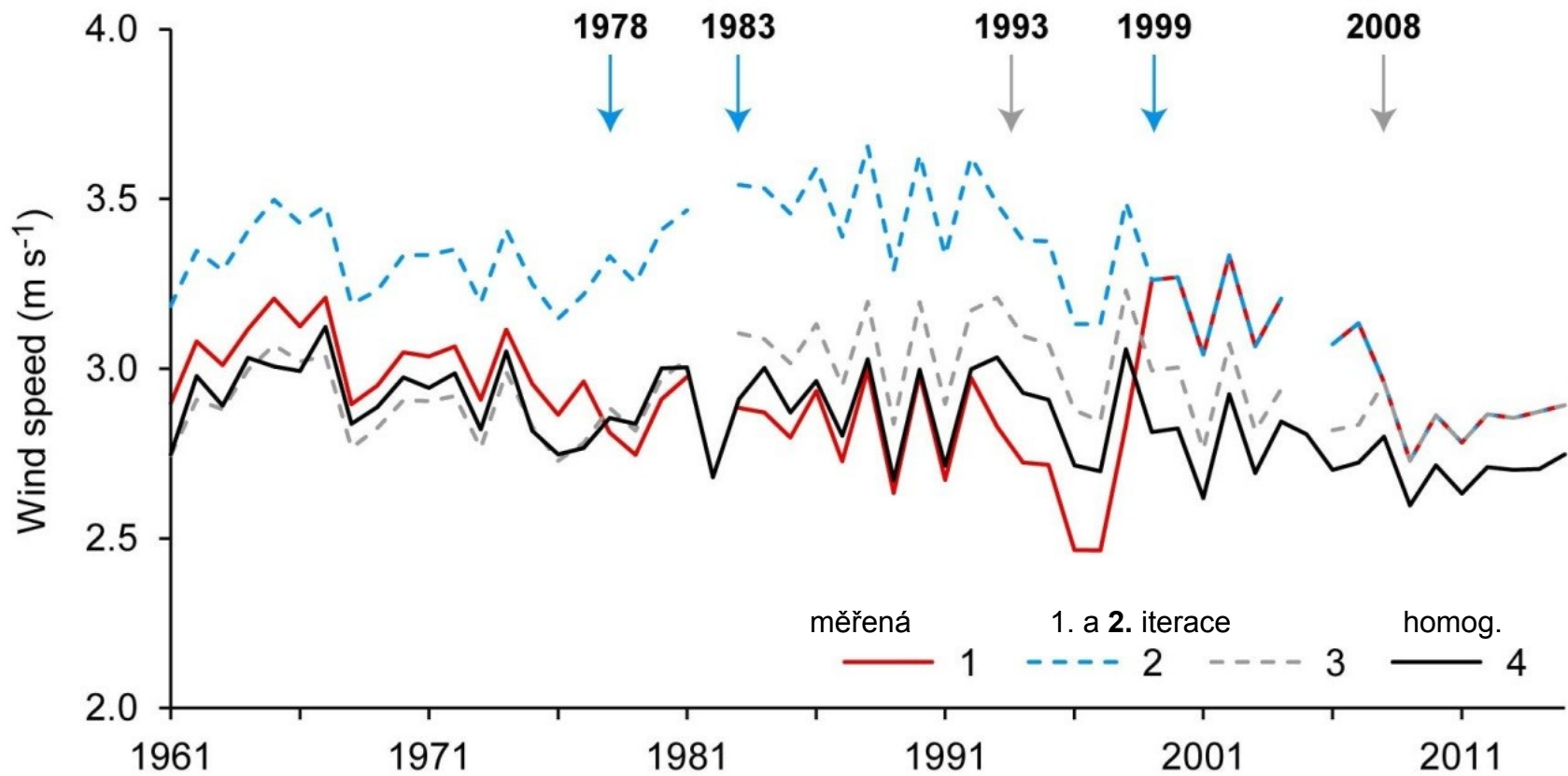
Porovnání originální (1) a homogenizované (2) řady průměrné roční teploty vzduchu Prahy-Klementina v letech 1775–2010. Pro jednotlivé zlomy je uvedena velikost korekce ve °C



Porovnání originální (1) a homogenizované (2) řady ročních úhrnů srážek Brna v období 1803 – 2010. Pro jednotlivé zlomy je uvedena hodnota korekčního faktoru



Dekádové četnosti detekovaných zlomů objasněných (1) a neobjasněných (2) metadaty v sekulárních řadách v České republice: a) teplota vzduchu (10 stanic), b) srážky (14 stanic)



Homogenizace roční řady průměrných denních rychlostí větru na stanici Hradec Králové v období 1961–2015: zlomy 1978, 1983 a **1993** bez metadat (do roku 1998 anemoindikátor, 1999 – anemometr Vaisala WAA151, **2008** – recalibrace stanice)

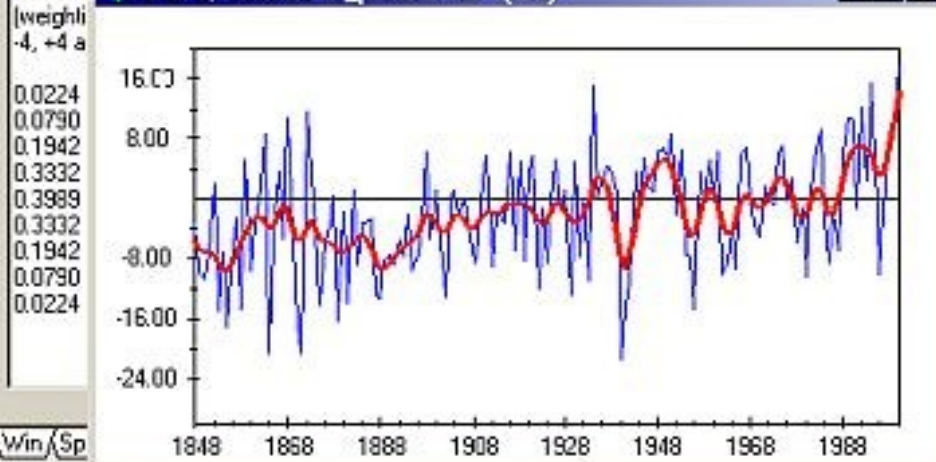




Low-pass Filter: a_prumCR.txt

Low-pass Filter: Gaussian ordinate method

Plots of Filtered a_prumCR.txt (Yea)



PS - MESA: a_prumCR.txt

Power Spectrum - MESA

Frequencies + Values + Period		
0.0000	674.3299	<
0.0042	716.3279	<
0.0083	808.9999	<
0.0125	802.4849	<
0.0167	601.3849	<
0.0208	390.8654	<
0.0250	266.0807	<
0.0292	204.7484	<
0.0333	181.4865	<
0.0375	186.5342	<
0.0417	224.4611	<
0.0458	320.5823	<
0.0500	537.5234	<
0.0542	870.4781	<
0.0583	823.4554	<
0.0625	512.3353	<
0.0667	335.1720	<

M = 30

Estimates related to

Harmonics

Frequencies

Normalize PS % Variance

Plot WN

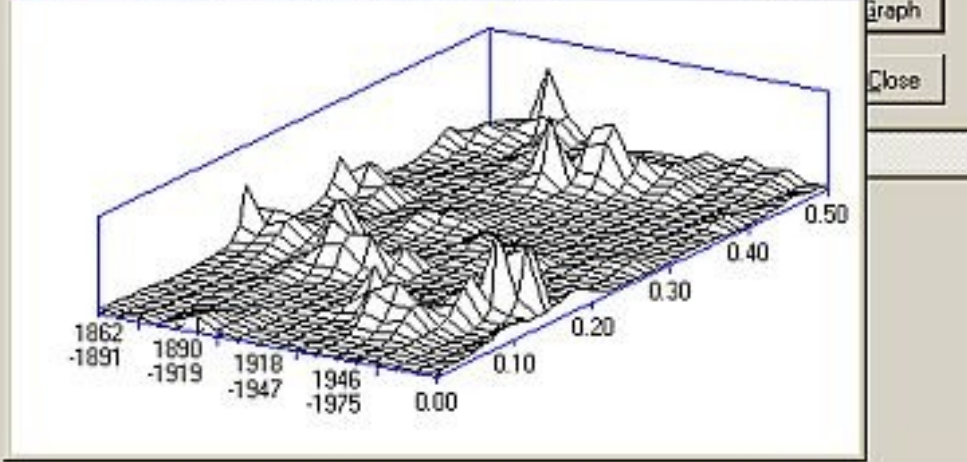
Plot Confidence Limits 95%

Save with Conf. Limits

Graph Save Save All Series Close

Win/Spr/Sum/Aut/Yea/

PS - Dynamic MESA - 3D : a_prumCR.txt



Series Controller

Active File Selection: *Open Files: 9*

D:\...\anom\va_prumCR.txt

Period: 1848 - 2000; 1 Missing Values

Series

Single series

Merged Series of one File

Merged Series of two Files

Analyzing

Simple series

Differences (Temperature)

Ratios (Precipitation)

Open all series of the file Use Seasonal and Annual Averages

Number of Series: 5

METODOLOGIE KONTROLY A HOMOGENIZACE ČASOVÝCH ŘAD V KLIMATOLOGII



Petr Štěpánek, Pavel Zahradníček, Rudolf Brázdil, Radim Tolasz

Doporučená literatura:

Alexandersson, A. (1986): A homogeneity test applied to precipitation data. *Journal of Climatology*, 6, č. 6, s. 661–675.

Brázdil, R., Štěpánek, P. (1998): Kolísání teploty vzduchu v Brně v období 1891–1995. *Geografie – Sborník České geografické společnosti*, 103, č. 1, s. 13-30.

Peterson, T. C. (1998): Homogeneity adjustments of in situ atmospheric climatic data: A review. *International Journal of Climatology*, 18, s. 1493–1517.

Štěpánek, P. (2004): Homogenizace teploty vzduchu na území České republiky v období přístrojových pozorování. *Práce a studie, seš. 32. Český hydrometeorologický ústav, Praha*, 56 s.

Štěpánek, P., Zahradníček, P., Brázdil, R., Tolasz, R. (2011): *Metodologie kontroly a homogenizace časových řad v klimatologii. Český hydrometeorologický ústav, Praha*, 118 s.

Štěpánek, P., Zahradníček, P., Farda, A. (2013): Experiences with data quality control and homogenization of daily records of various meteorological elements in the Czech Republic in the period 1961–2010. *Időjárás*, 117, s. 123–141.

Publikace na stránkách COST-ES0601.

Action COST-ES0601: Advances in
homogenisation methods of climate series:
integrated approach (HOME)

<http://www.homogenisation.org/>

nyní

<http://www.climatol.eu/tt-hom/>

Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C.N., Menne, M., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., Brandsma, T. (2012): Benchmarking homogenization algorithms for monthly data. *Climate of the Past*, 8, 89–115.