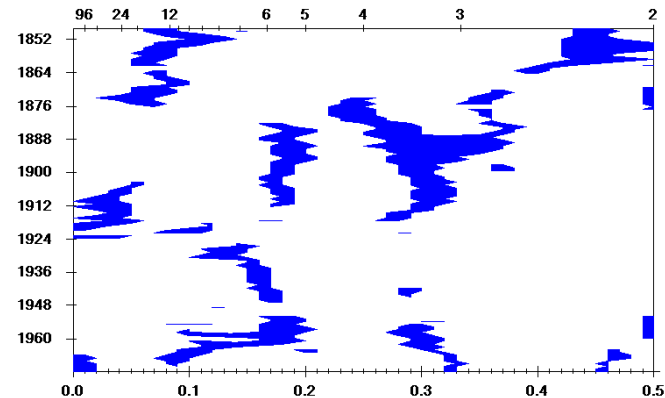
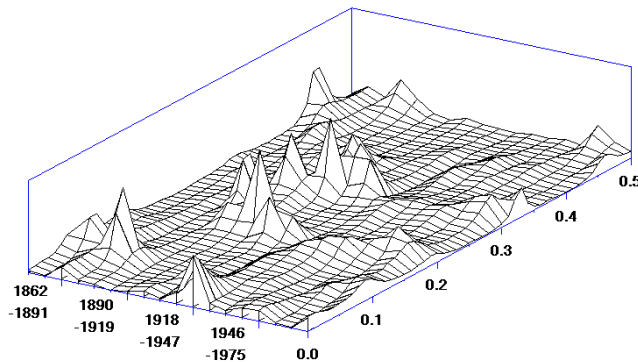
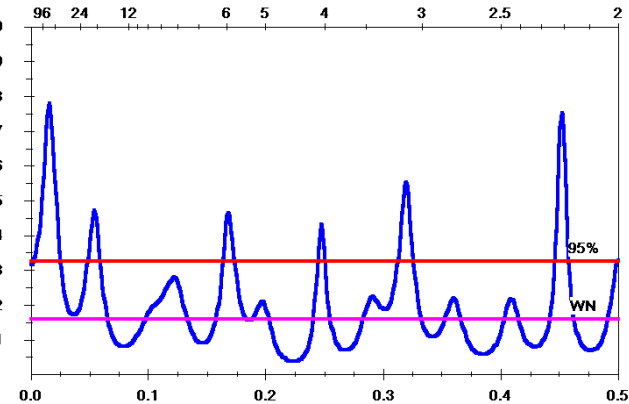
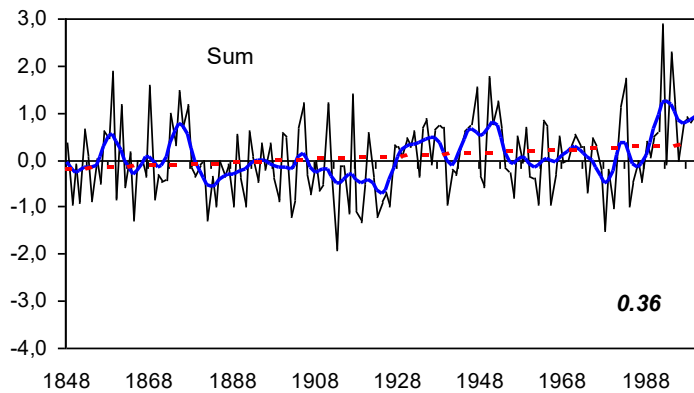


# 3. Metody analýzy časových řad v klimatologii



## 3.1 Periodicita a cykličnost

Klima je vyjádřeno různými prvky (např. teplota vzduchu, srážky, indexy), proměnlivými v prostoru a čase:

$$\{a, b, c, \dots\} = f(x, y, z, t),$$

$a, b, c, \dots$  jsou prvky klimatu,  $x, y, z$  jsou prostorové souřadnice a  $t$  je čas.

Při studiu časových změn platí pro dané místo

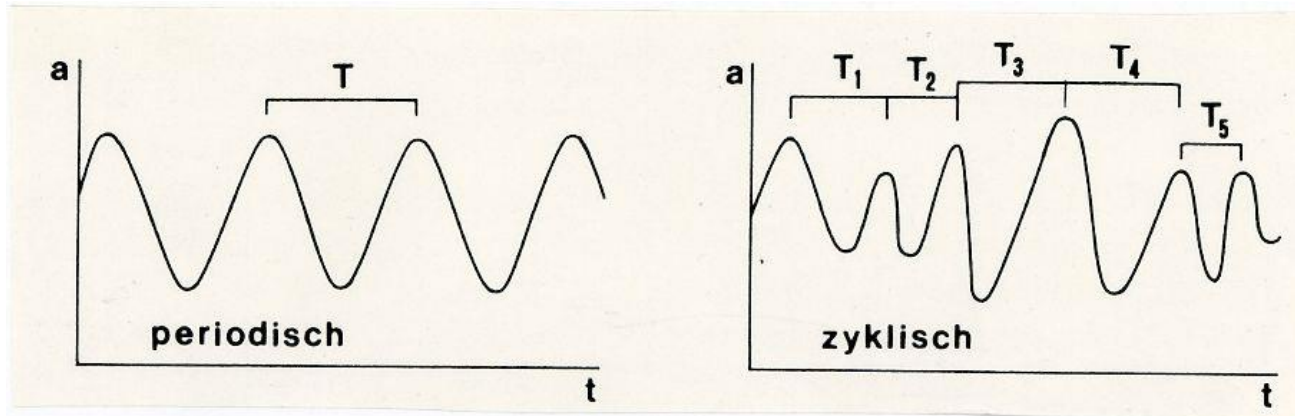
$$\{a, b, c, \dots\} = f(t).$$

Informace o časových změnách je získávána v podobě časové řady

$$\{a_i, b_i, c_i, \dots\} = f(t_i) \text{ s } t_{i+1} - t_i = \Delta t = \text{konst.}, i = 1, 2, \dots, n$$

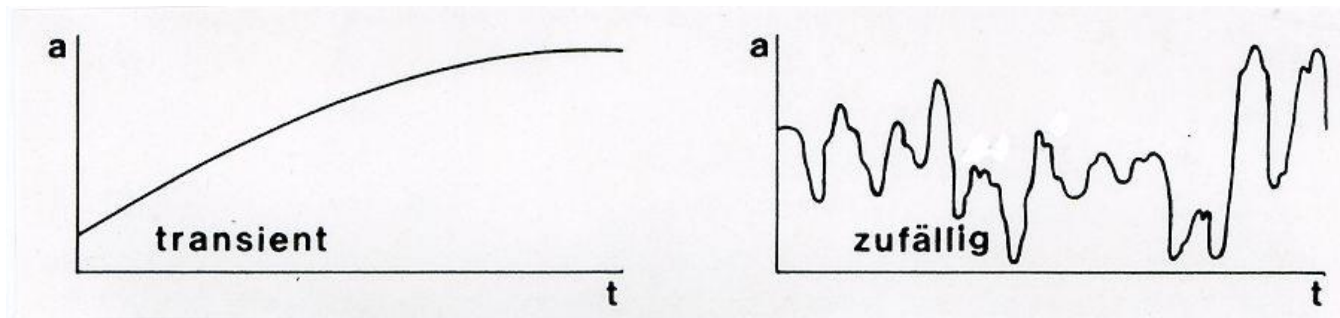
$\Delta t$  je časový interval (např. den, měsíc, rok) a  $n$  je délka řady.

# Typy kolísání v časových řadách



$$a(t_i) = a(t_i + T)$$

$$a(t_i) \approx a(t_i + \bar{T})$$



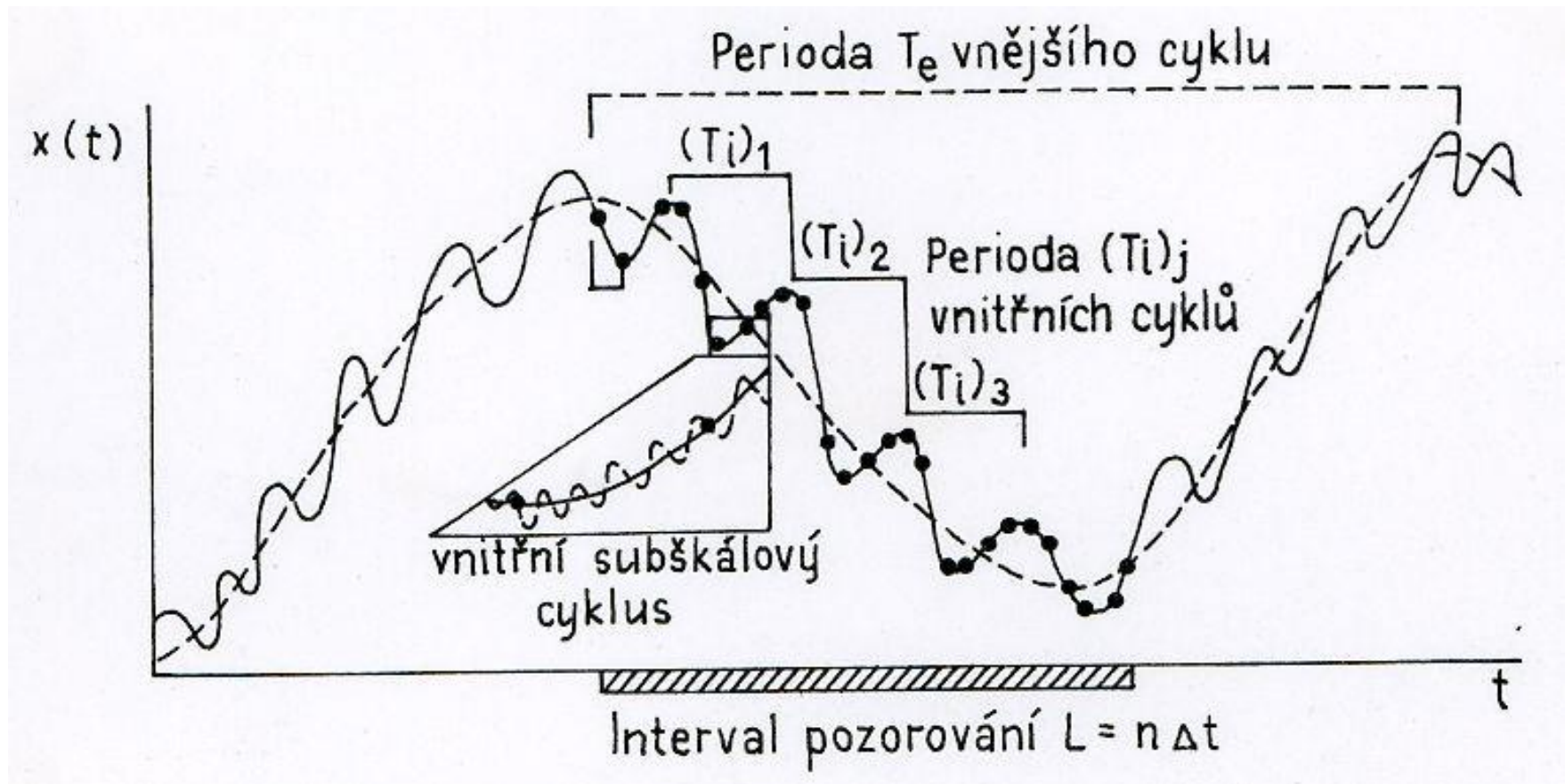
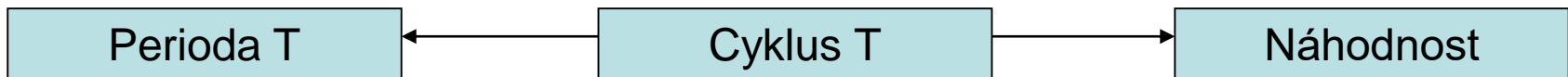
**Periodická:** s periodou  $T$

**Transientní:** lineární nebo nelineární trend

**Cyklická (kvaziperiodická, rytmická):**

s periodou  $\bar{T}$

**Náhodná (stochastická)**



Cyklus s periodou  $T_p$ :  $\Delta t \ll T_p \ll n \Delta t$

Vnitřní cykly:  $T_p < \Delta t$

Vnější cykly:  $T_p > n \Delta t$

## Podmínky aplikovatelnosti statistických metod

- **nezávislost** hodnot časové řady
- **stacionarita** řady, tj. statistické charakteristiky řady  $S_j$  ( $j = 1, 2, \dots, J$ ) jsou časově nezávislé, tedy  $\partial S_j / \partial t = \text{konst.}$
- **ergodicita**, tj. průměr dané řady (výběrový soubor) s jejím prodlužováním v časovém intervalu  $L$  se blíží střední hodnotě celé řady (základního souboru)

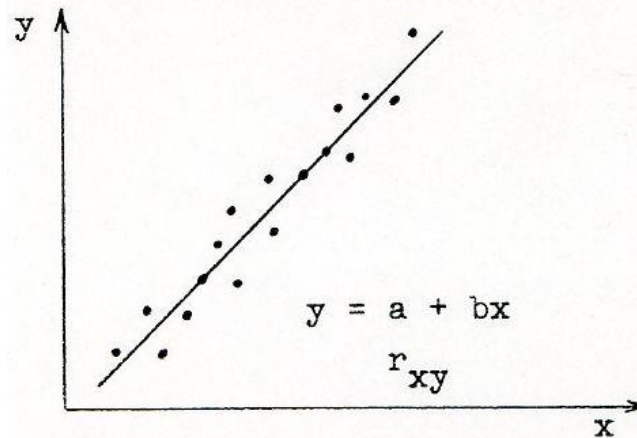
**Nesplnění podmínek:** výsledky analýzy mají platnost pouze pro danou realizaci časové řady (tj. daný výběr), je třeba testovat výsledky se zřetelem na vlastnosti základního souboru.

**Nestacionarita** – přirozená vlastnost klimatických řad (změny a kolísání klimatu).

## 3.2 Korelační analýza

### Korelační koeficient

Podmínka použití:  
hodnoty obou řad jsou  
nezávislé a mají  
přibližně normální  
rozdělení.



$$r_{xy} = \frac{1}{n-1} \frac{\sum_{i=1}^n x'_i \cdot y'_i}{s_x \cdot s_y}$$

$$x'_i = x_i - \bar{x}, \quad y'_i = y_i - \bar{y}$$

$s_x, s_y$  - standard deviation

$r_{xy}^2$  = explained variance / total variance

Test criterion of significance of  $r_{xy}$ :

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

$\hat{r}_{xy}$ -transformation (Fisher) for non-normally distributed series:

$$\hat{r}_{xy} = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}}$$

$$u_p = (\hat{r}_{xy} - z/\sqrt{n-3}; \hat{r}_{xy} + z/\sqrt{n-3})$$

$$\hat{z} = \hat{r}_{xy} \sqrt{n-3}$$

# Autokorelační koeficient

Autokovariance

$$V_x(\tau) = \frac{1}{n - \tau - 1} \sum_{i=1}^{n-\tau} x_i' \cdot x_{i+\tau}'$$

Autokorelace

$$r_x(\tau) = V_x(\tau) / s_1 \cdot s_2$$

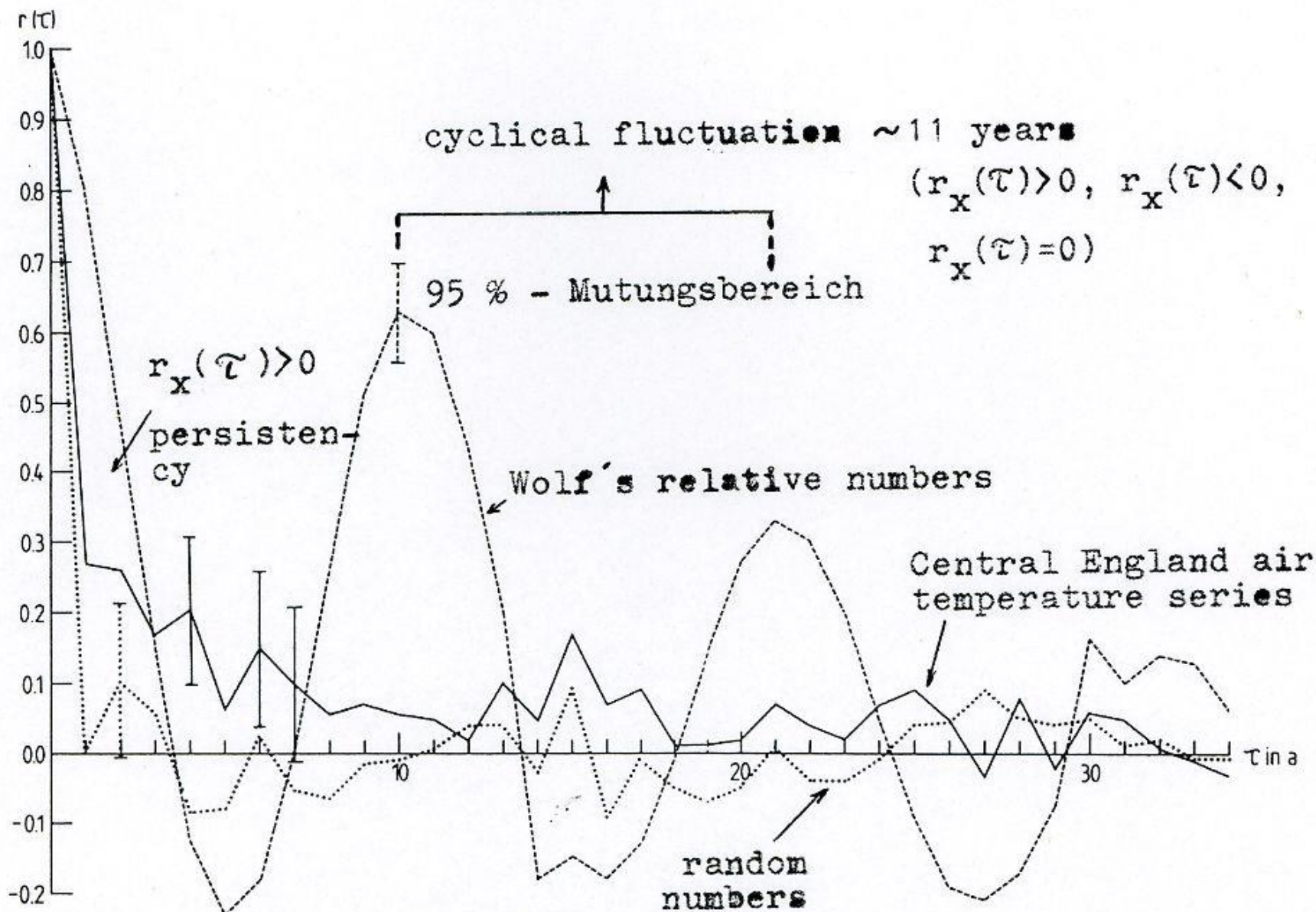
$$\tau = 0, 1, 2, \dots, M < n/2 - \text{time lag, } M - \text{maximum lag}$$

A diagram of the computation:

I.	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	....	$\tau = 0 \cdot \Delta t$	$r_x(0) = 1$
II.	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	....	$\tau = 1 \cdot \Delta t$	$r_x(1)$
III.	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	....	$\tau = 2 \cdot \Delta t$	$r_x(2)$
IV.	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	....	$\tau = 3 \cdot \Delta t$	$r_x(3)$ etc.

Confidence levels of  $r_{xy}$  by Andersson:

$$r_x(\tau) = \frac{-1 \pm z(p) \sqrt{n - \tau - 2}}{n - \tau - 1}$$





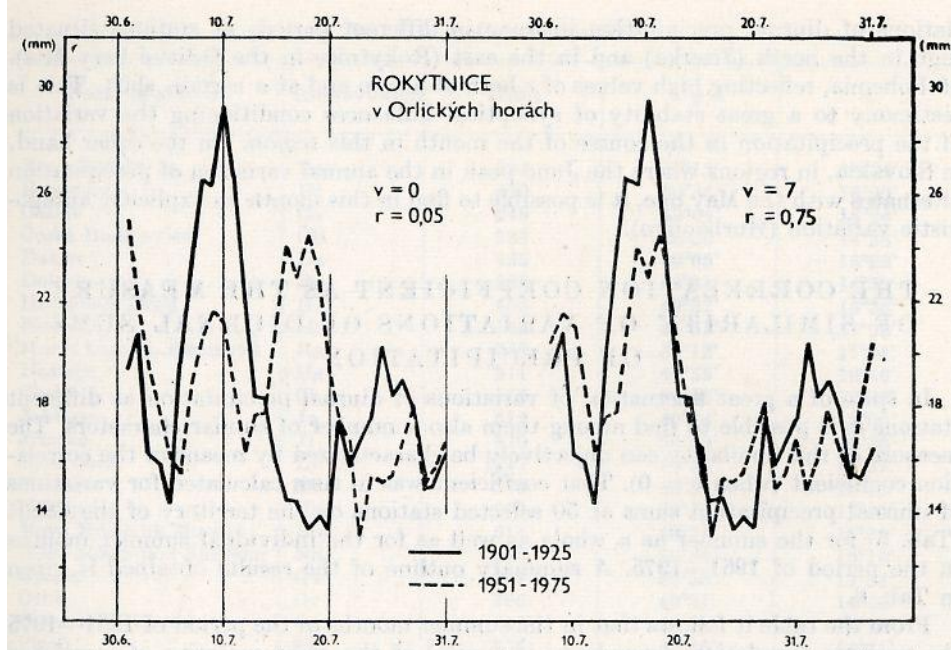
# Křížově-korelační koeficient

series X:  $x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad \dots$

series Y:  $y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6 \quad y_7 \quad \dots$  with possible shift  $\tau$

$$r_{Cxy} = \frac{\sum_{i=1}^{n-\tau} x'_i \cdot y'_{i+\tau}}{\sqrt{\sum_{i=1+\tau}^n x_i'^2 \sum_{i=1}^{n-\tau} y_i'^2}}$$

$$-1 \leq r_{Cxy} \leq 1$$



Denní úhrny srážek

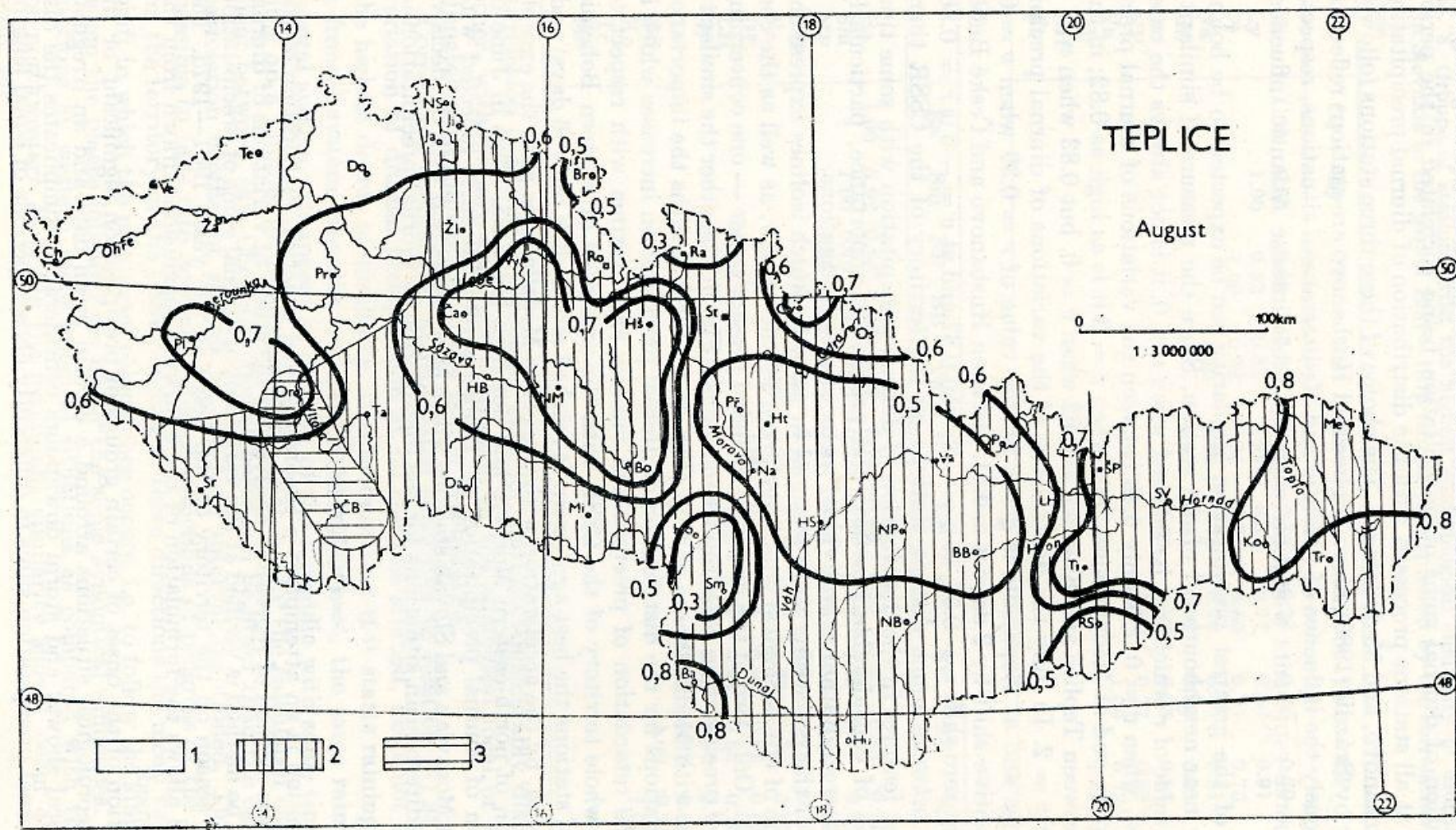
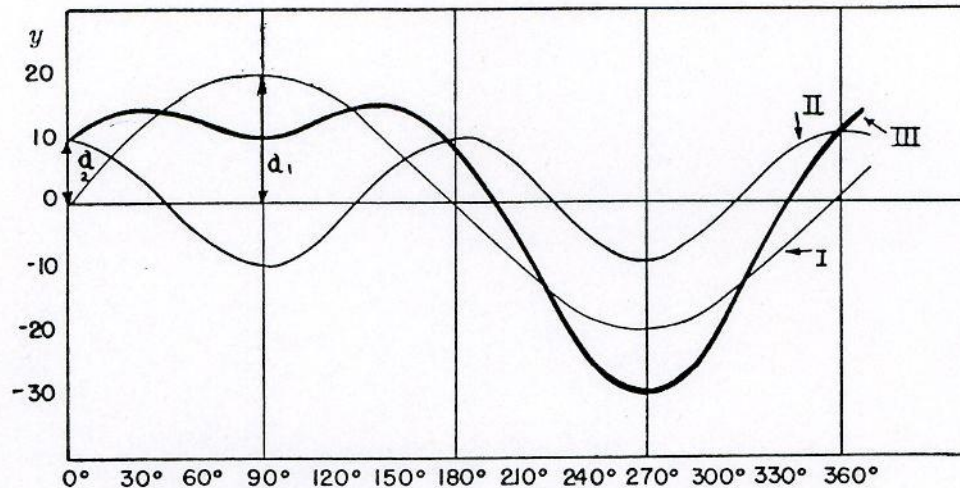


Fig. 3. Isotherms of the maximum degree of climatological similarity of variations of diurnal sums of the precipitation of August with respect to station Teplice and the corresponding values of shift (1 — zero shift, 2 — positive shift, 3 — negative shift)

## 3.3 Harmonická analýza



Example of superposition of two waves.

$$\begin{aligned} \text{I: } y_1 &= a_1 \sin(x + A_1), & A_1 &= 0^\circ; \\ \text{II: } y_2 &= a_2 \sin(2x + A_2), & A_2 &= 90^\circ; \\ \text{III: } y &= y_1 + y_2. \end{aligned}$$

**Řada spojitá, nekonečná a periodická – superpozice konečného počtu sinových a kosinových funkcí – Fourierova analýza – Fourierova řada:**

$$x(t) = A_0/2 + \sum_{i=1}^n A_i \sin \omega t + \sum_{i=1}^n B_i \cos \omega t, \quad \omega = 2\pi/T = 2\pi f$$

$$A_i = \frac{2}{T} \int_0^T x(t) \sin i\omega t dt$$

$$B_i = \frac{2}{T} \int_0^T x(t) \cos i\omega t dt$$

Pro klimatologické řady, nesplňující uvedené podmínky, lze nahradit Fourierovy koeficienty Besselovou formulí, aplikovatelnou pro přibližně periodickou řadu:

$$x_j(t_j) = \bar{x} + \sum_{i=1}^{N/2} \left[ A_i \sin\left(\frac{360^\circ}{P} it\right) + B_i \cos\left(\frac{360^\circ}{P} it\right) \right]$$

$$A_i = \frac{2}{N} \sum_{j=1}^N x_j \sin\left(\frac{360^\circ}{P} it_j\right) \quad i = 1, 2, \dots, N/2 - 1$$

$$B_i = \frac{2}{N} \sum_{j=1}^N x_j \cos\left(\frac{360^\circ}{P} it_j\right) \quad i = 1, 2, \dots, N/2$$

$1/P_{\min}$  = Nyquist's frequency      Počet harmonik  $h_i = P/i$  je  $N/2$ ,  $h_{\min} = 2\Delta t$ ,  $h_{\max} = P$ .

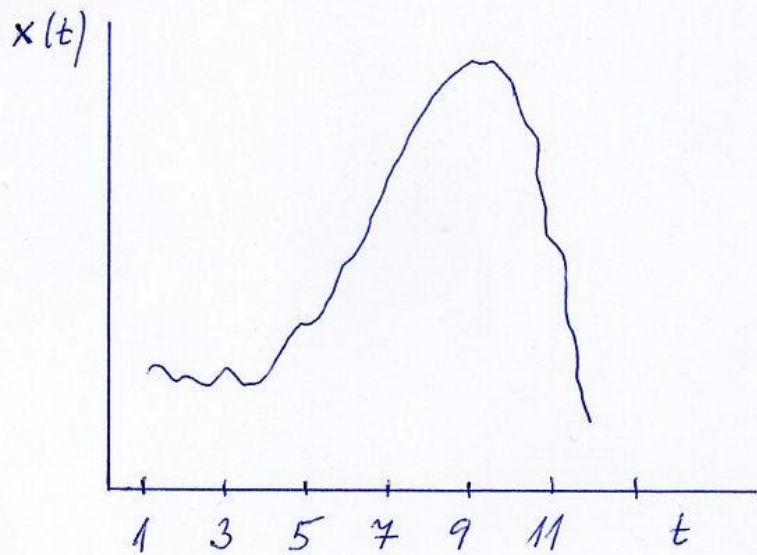
Explained variance through  $h_i$ :

for  $i < N/2$        $C_i^2/2s^2$       with  $C_i = \sqrt{A_i^2 + B_i^2}$

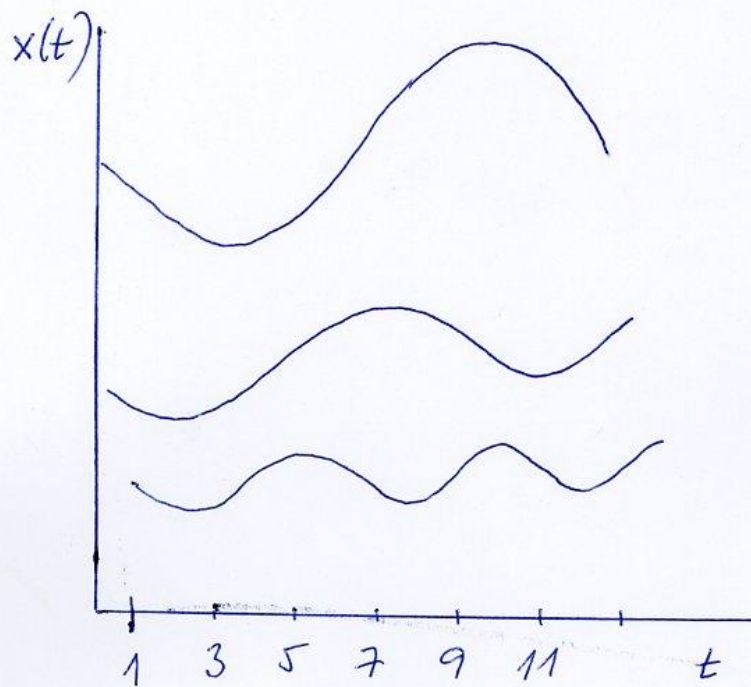
for  $i = N/2$        $C_{N/2}^2/s^2$       with  $C_{N/2} = \sqrt{B_{N/2}^2}$

Time of maximum of  $h_i$ :

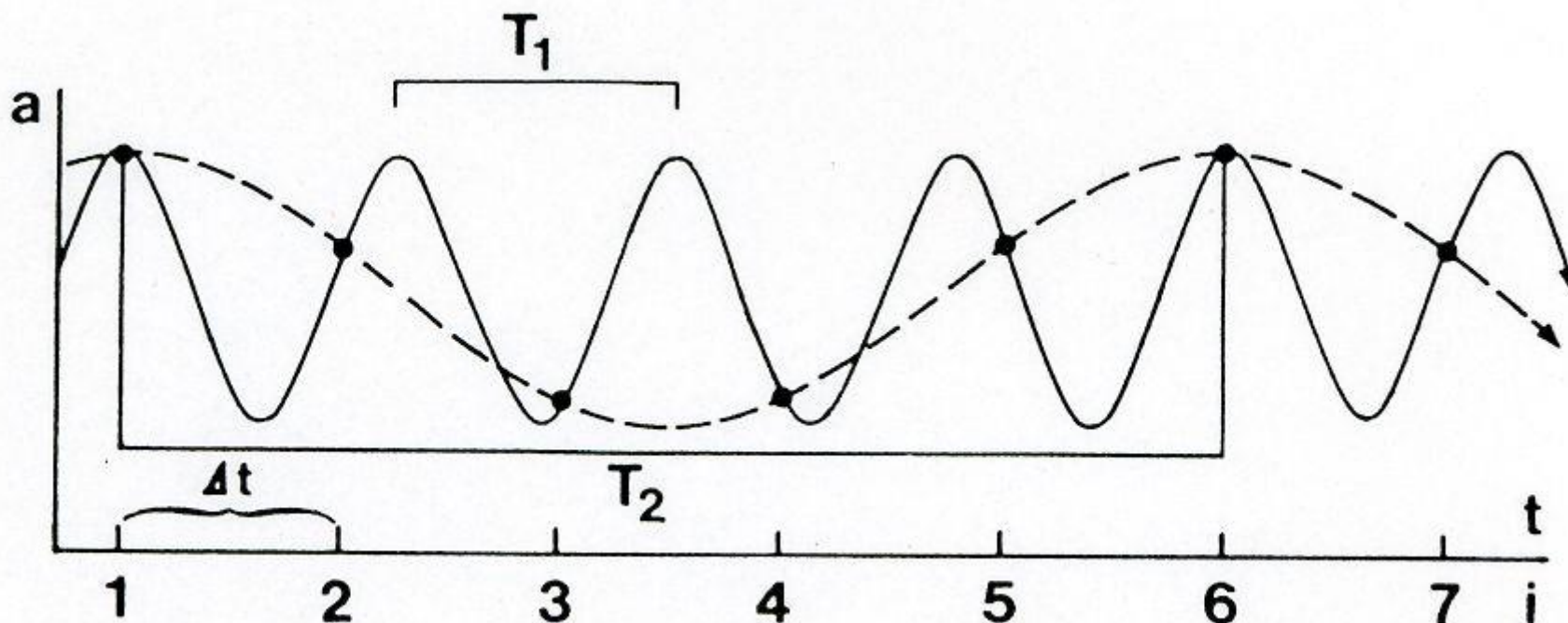
$$t_i = \frac{P}{360^\circ i} \arctan(A_i/B_i)$$



$N = 12$  (délka řady)  
 $P = 12$  (přehledná perioda)  
 $\Delta t = 1$  (časový krok)



$h_1 = 12/1 = 12$  ( $c_1, t_1$ )  
 $h_2 = 12/2 = 6$  ( $c_2, t_2$ )  
 $h_3 = 12/3 = 4$   
 $\vdots$   
 $h_6 = 12/6 = 2$



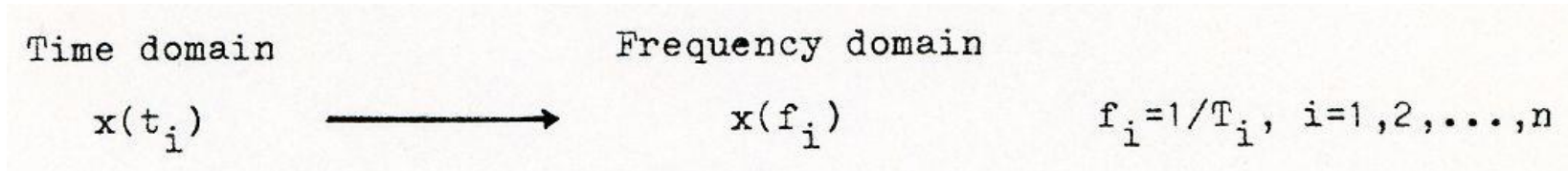
**Falešný efekt** (aliasing effect) – skutečná kolísání s periodou  $T_1$  ( $5/4\Delta T$ ) mohou být se zřetelem na interval měření  $\Delta t$  chybně interpretována periodou  $T_2$  ( $5\Delta T$ ), kde  $T_2 > T_1$

## Harmonická analýza ročních úhrnů trvání slunečního svitu

	1st Harmonic			2nd Harmonic		
	Ampl	Var	day of max	Ampl	Var	day of max
ATHENS	120.6 h	90%	Jul 14		1%	
BUCURESTI	121.6	90	Jul 11		1	
BUDAPEST	122.2	90	Jul 4		1	
DEBRECEN	122.1	90	Jul 4		1	
ESSEN	83.7	91	Jun 27		0	
GEISENHEIM	98.8	90	Jun 25		1	
HURBANOVO	113.6	90	Jul 2		1	
KARLSRUHE	102.2	90	Jun 30		1	
KASSEL	90.8	91	Jun 25		0	
KLAGENFURT	96.2	87	Jun 26	20.9 h	4%	Feb 24
KLATOVY	95.7	90	Jul 1		1	
CRACOW	93.6	91	Jul 2		0	
KRDZHALI	115.0	89	Jul 14	19.8 h	3%	Feb 18
KOSICE	108.9	89	Jul 3		2	
LIPTOVSKY	81.7	87	Jul 2	17.2 h	4%	Mar 9
OBRASCOV	123.3	90	Jul 12		1	
POTSDAM	104.3	91	Jun 29		0	
PRAGUE	105.0	90	Jun 30		1	
SNEZKA	53.9	89	Jun 30		0	
SONNBLICK	27.7	71	Jul 22	12.3 h	14%	Mar 15
TCHIRPAN	118.1	89	Jul 12		2	
THESSALONIKI	117.8	89	Jul 10		2	
VIENNA	106.1	90	Jul 3		1	
ZAGREB	107.8	89	Jul 6		2	
	$C_1$		$t_1$	$C_2$		$t_2$

## 3.4 Spektrální analýza (Power Spectrum Analysis)

- zobecněná harmonická analýza



Postup výpočtu:

- Výpočet autokorelačních (autokovariančních) koeficientů pro posun od 0 do  $M$ , kde  $M < n$ ,  $n$  – délka řady
- Výpočet kosinové transformace  $M+1$  autokorelačních koeficientů (analogie Fourierovy transformace pro spojité hodnoty)
- Kosinovou transformací se získá  $M+1$  „hrubých odhadů“ variančního spektra (jde o vyjádření podílu period blízkých  $i$ -té harmonice  $h$  na celkovém rozptylu dané řady)
- Shlazení „hrubých odhadů“ k získání finálního spektra pro  $M+1$  diskrétních hodnot



## Spectral analysis according to Blackman and Tukey:

Raw spectral estimates of  $S_x(h)$ :

$$S_x(0) = 1/2M \left[ V_x^2(0) + \sum_{k=1}^{M-1} D(k) V_x^2(k) \right] \quad \text{for } h = 0$$

$$S_x(h) = 1/2M \left[ V_x^2(0) + \sum_{k=1}^{M-1} D(k) V_x^2(k) \cos \frac{\pi hk}{M} \right] \quad \text{for } 0 < h < M$$

$$S_x(M) = 1/2M \left[ V_x^2(0) + \sum_{k=1}^{M-1} D(k) V_x^2(k) (-1)^k \right] \quad \text{for } h = M.$$

$$k = 0, 1, 2, \dots, M; \quad D(k) = \begin{cases} 1/2(1 + \cos \pi k/M) & \text{for } 0 < k < M \\ 0 & \text{for } -\infty < k < 0 \text{ and } M < k < +\infty \end{cases}$$

$h$  – harmonika (0 – spektrální odhad odpovídající trendu,  $M$  – odpovídá  $2\Delta t$ )

$D(k)$  – filtrační funkce umožňující použití Fourierovy transformace na konečnou řadu

### Smoothed spectral estimates:

#### Hamming window:

$$S_x(0) = 0.54S_x(0) + 0.46S_x(1)$$

$$S_x(h) = 0.23S_x(h-1) + 0.54S_x(h) + 0.23S_x(h+1) \text{ for } 0 < h < M$$

$$S_x(M) = 0.54S_x(M) + 0.46S_x(M-1)$$

#### Hanning window:

$$S_x(0) = 1/2 (S_x(0) + S_x(1))$$

$$S_x(h) = 1/4 (S_x(h-1) + 2S_x(h) + S_x(h+1)) \text{ for } 0 < h < M$$

$$S_x(M) = 1/2 (S_x(M-1) + S_x(M))$$

Omezení spektra:

$$f_{\min} = 0 \rightarrow T_{\max} = \infty$$

$$f_{\max} = 1/2\Delta t \rightarrow T_{\min} = 2\Delta t$$

## Interpretace spektra

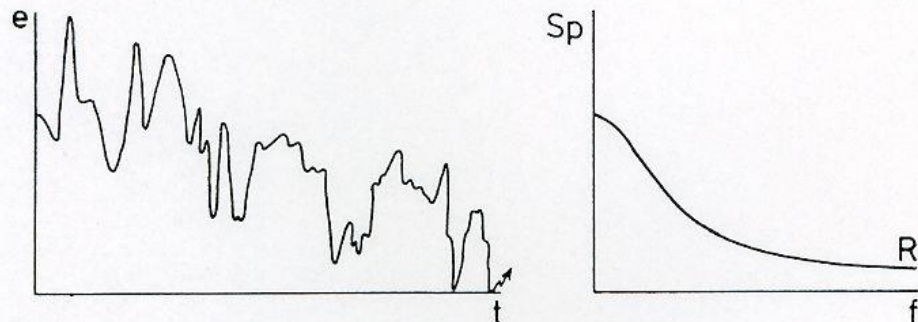
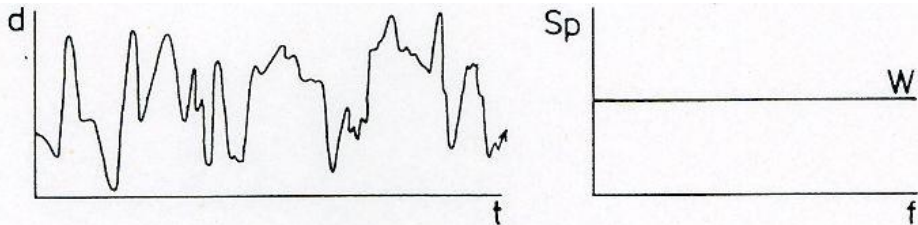
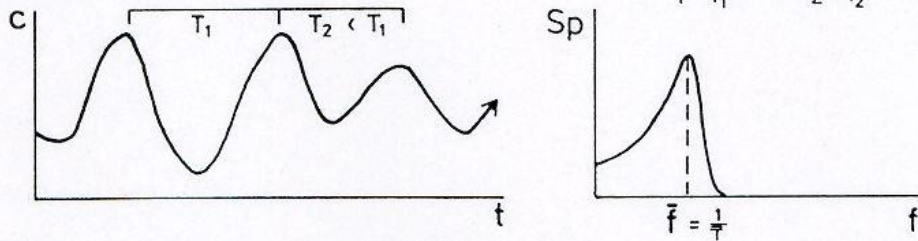
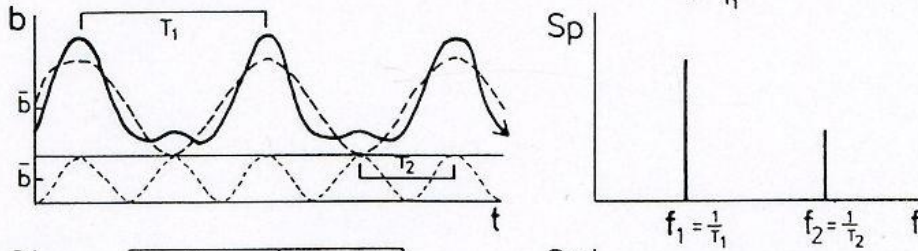
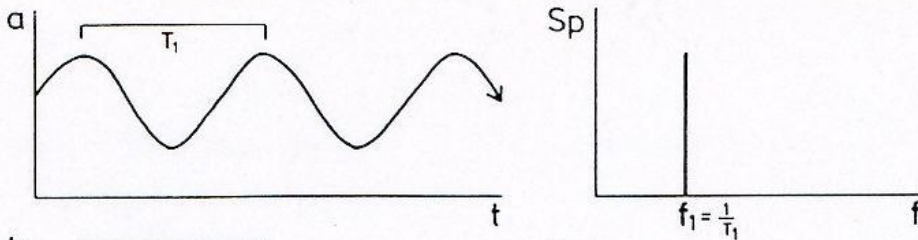
a) Periodická funkce s periodou  $T_1$

b) Periodická funkce s periodami  $T_1$  a  $T_2$

c) Cyklická kolísání se střední periodou  $\bar{T}$

d) Náhodná čísla (tzv. bílý šum nebo bílé spektrum)

e) Náhodná čísla s trendem (tzv. rudý šum nebo rudé spektrum)



Testování významnosti získaného výběrového variančního spektra se provádí s ohledem na bílé, resp. rudé spektrum.

Markov's time series ( $r_x(1)$  differs significantly from zero and further  $r_x(2) \approx [r_x(1)]^2$ ,  $r_x(3) \approx [r_x(1)]^3$ , etc. has a red spectrum:

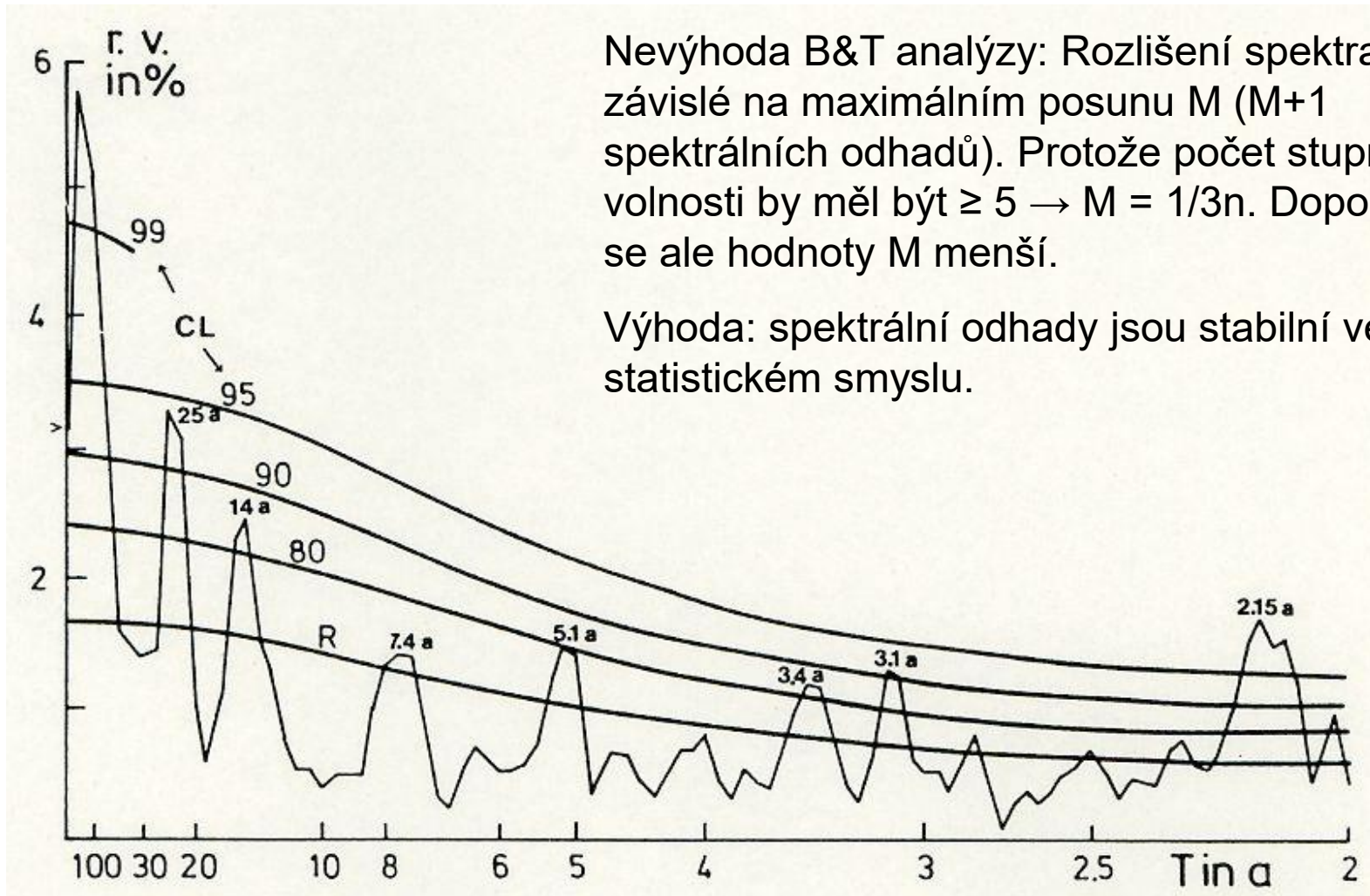
$$RN(h) = \frac{1 - [r_x(1)]^2}{1 + [r_x(1)]^2 - 2r_x(1) \cos \frac{h\pi}{M}} \cdot WN$$

where white spectrum WN

$$WN = \frac{1}{M} \left[ (1/2 S_x(0) + S_x(M)) + \sum_{h=1}^{M-1} S_x(h) \right]$$

Confidence limits:

$$G_\alpha = RN(h) \frac{\chi_{\nu, \alpha}^2}{\nu} \quad \nu = \frac{2n - M/2}{M}$$



Nevýhoda B&T analýzy: Rozlišení spektra je závislé na maximálním posunu  $M$  ( $M+1$  spektrálních odhadů). Protože počet stupňů volnosti by měl být  $\geq 5 \rightarrow M = 1/3n$ . Doporučují se ale hodnoty  $M$  menší.

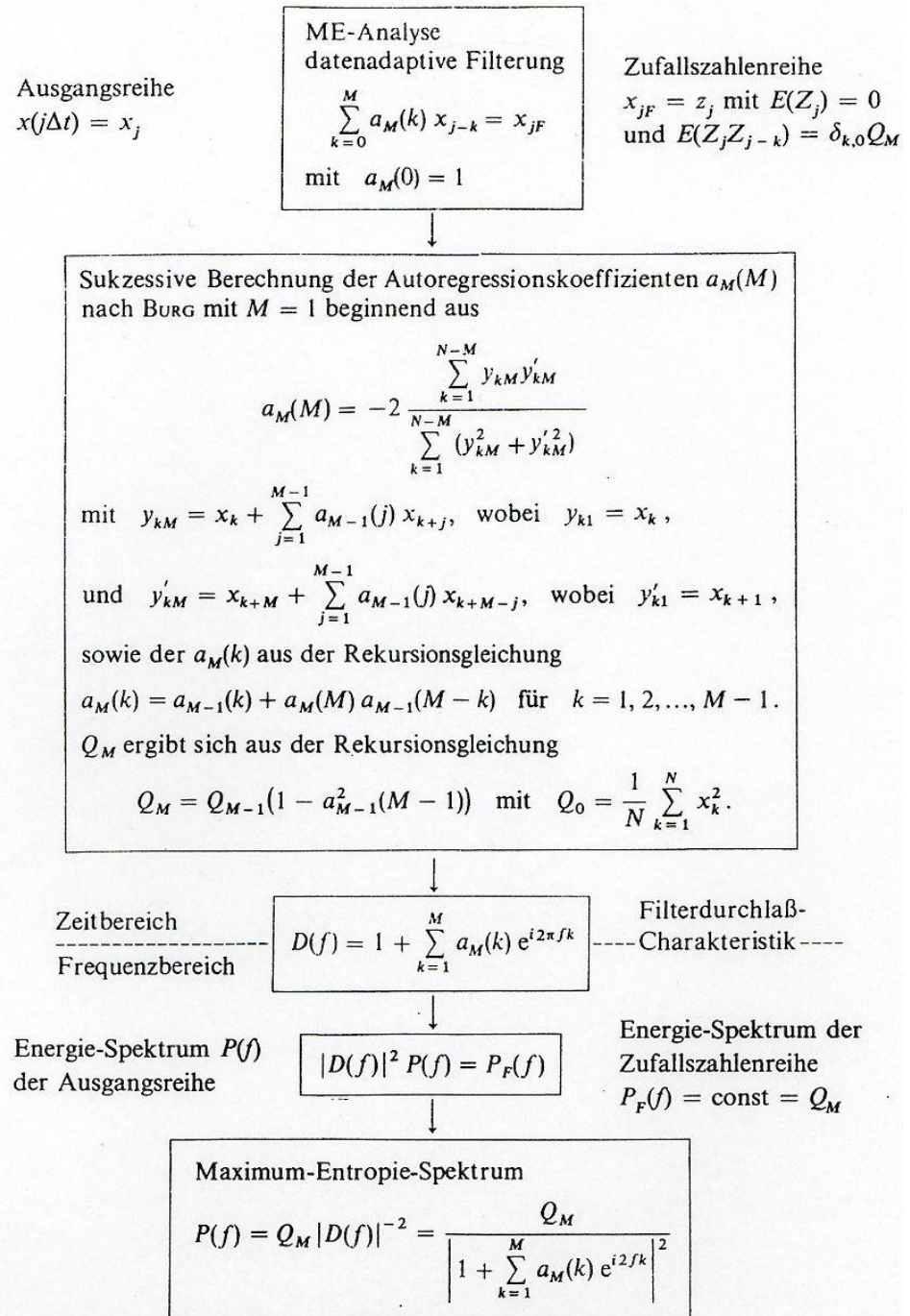
Výhoda: spektrální odhady jsou stabilní ve statistickém smyslu.

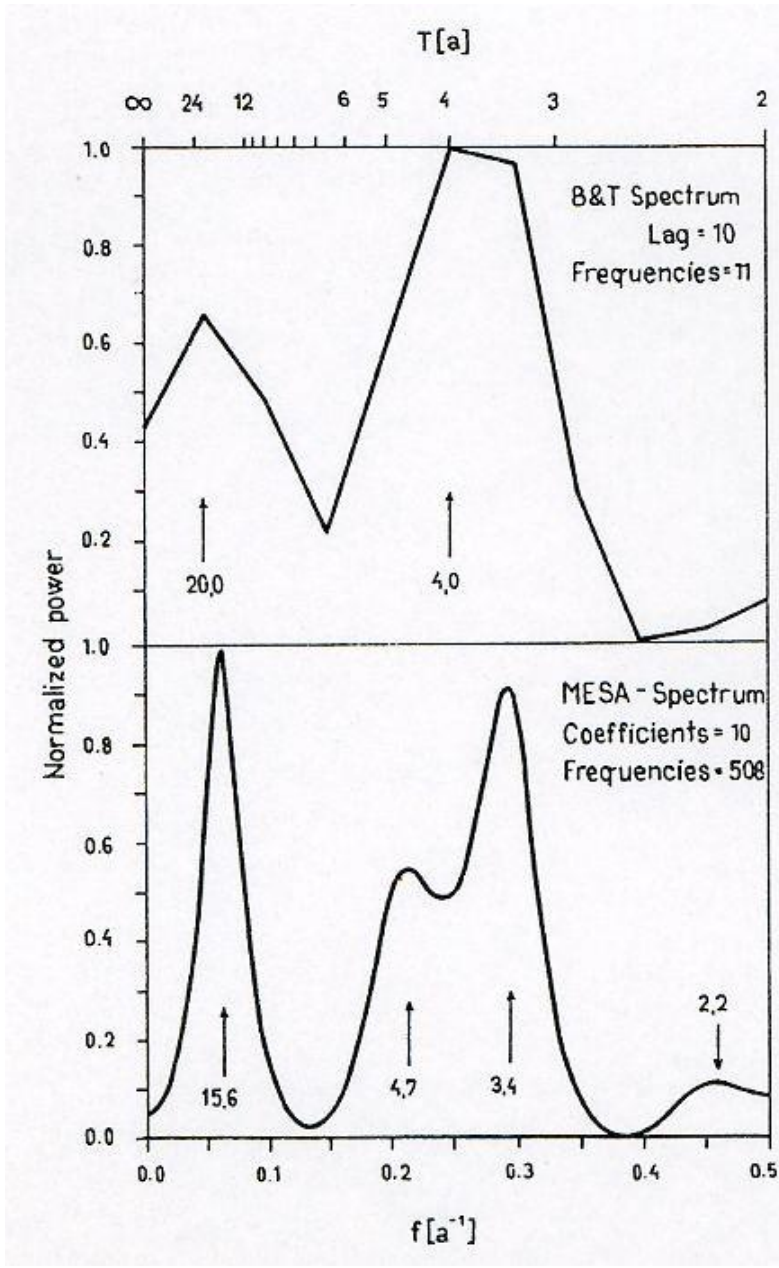
Spektrální analýza Blackmann-Tukey řady ročních průměrných teplot vzduchu Střední Anglie 1660-1969 ( $M = 100$ ). R – rudé spektrum, CL – hladiny významnosti, r. v. – relativní rozptyl

## Nevýhody spektrální analýzy:

- nestabilita spektra
- nestacionarita řad
- „falešný efekt“
- malé frekvenční rozlišení (odstraněno použitím MESA)

## Schéma spektrální analýzy maximální entropie (Maximum Entropy Spectral Analysis – MESA)





Normovaná varianční spektra ročních úhrnů srážek - Maďarsko

## 3.5 Křížově-spektrální a koherenční analýza

Cross-spectrum function:  $P_{xy}(h) = C_{xy}(h) + i Q_{xy}(h)$

Real part of cross-spectrum  $C_{xy}(h)$  (Kospektrum, Wirksamspektrum):

$$C_{xy}(h) = \Delta t / \pi \sum_{p=0}^M \epsilon_p (R_{xy}(p) + R_{yx}(p)) \cos hp\pi/M$$

with  $\epsilon_p = \begin{cases} 1/2 & \text{for } p = 0, M \\ 1 & \text{for } 0 < p < M \end{cases}$   $C_{xy}(h)$  vyjadřuje podíl oscilací různých frekvencí na celkové kovarianci při nulovém posunu.

Imaginary part of cross-spectrum  $Q_{xy}(h)$  (Quadraturespektrum, Blindspektrum):

$$Q_{xy}(h) = \Delta t / \pi \sum_{p=0}^M \epsilon_p (R_{xy}(p) - R_{yx}(p)) \sin hp\pi/M$$

with  $\epsilon_p = \begin{cases} 1/2 & \text{for } p = 0, M \\ 1 & \text{for } 0 < p < M \end{cases}$   $Q_{xy}(h)$  vyjadřuje podíl různých harmonik na celkové kovarianci (harmoniky řady x jsou posunuty o  $1/4$  periody zpět a řada y je beze změny).

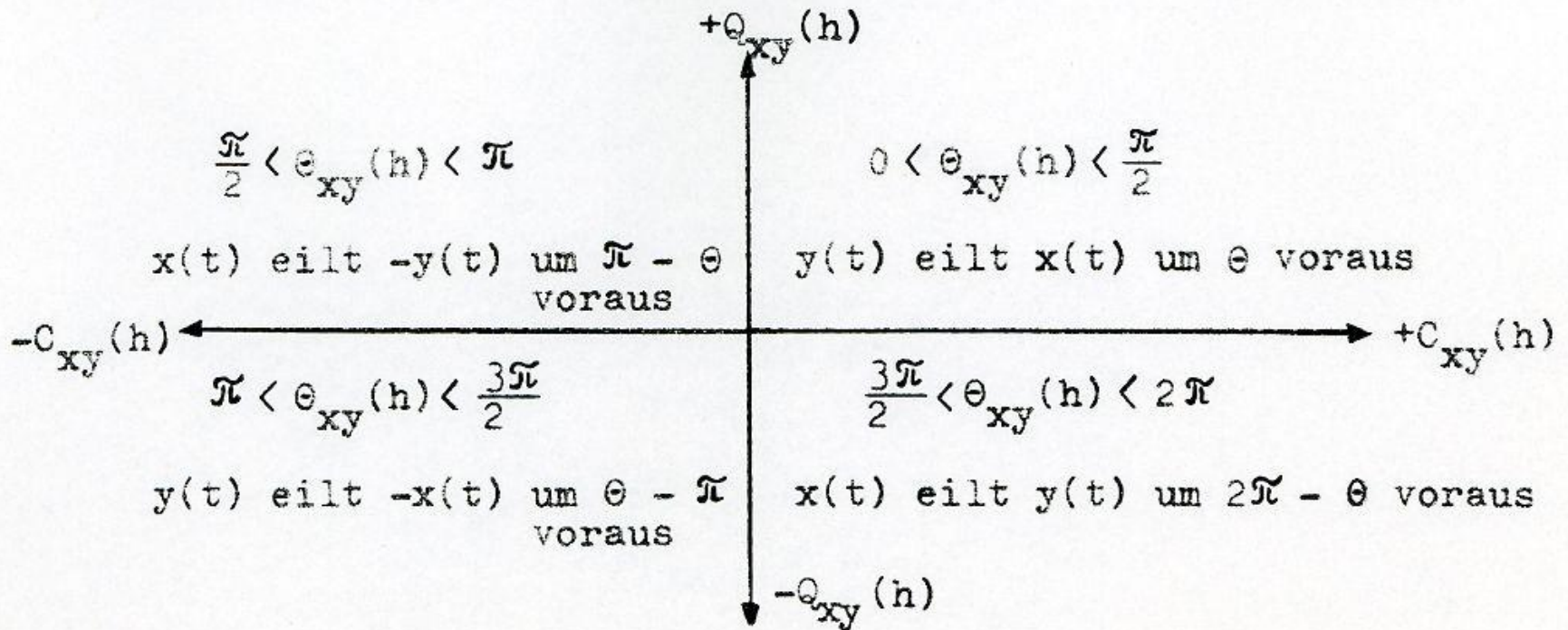


Other possible expression of the cross-spectrum function:

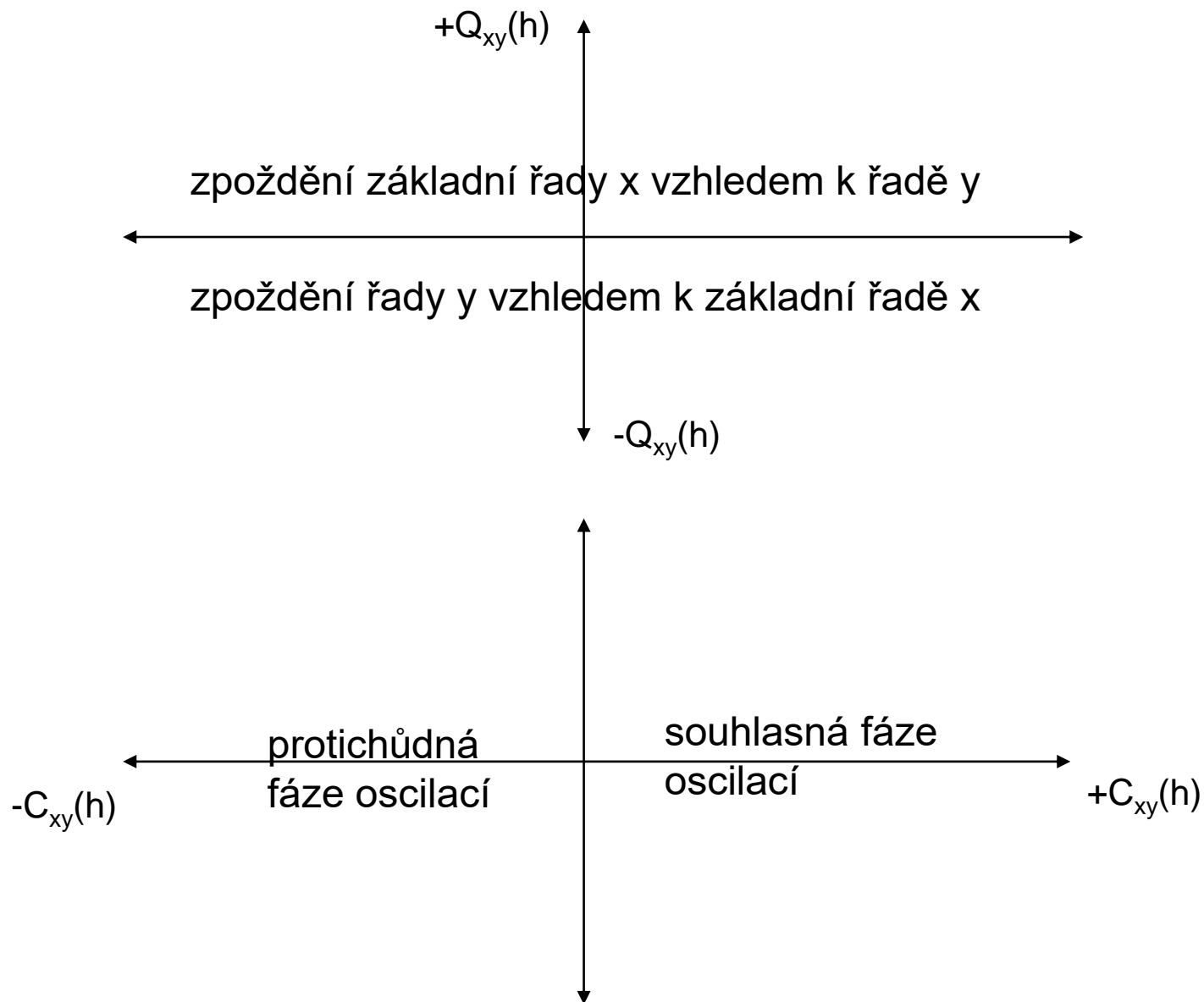
$$P_{xy}(h) = |A_{xy}(h)| e^{-i\theta_{xy}(h)}$$

Amplitude:  $A_{xy}(h) = \sqrt{SC_{xy}^2(h) + SQ_{xy}^2(h)} \quad h=0,1,\dots,M$

Phase angle:  $\theta_{xy}(h) = \tan^{-1}(SQ_{xy}(h)/SC_{xy}(h)) \quad h=2,3,\dots,M-2$



# Interpretace křížově-spektrální analýzy



Coherency:

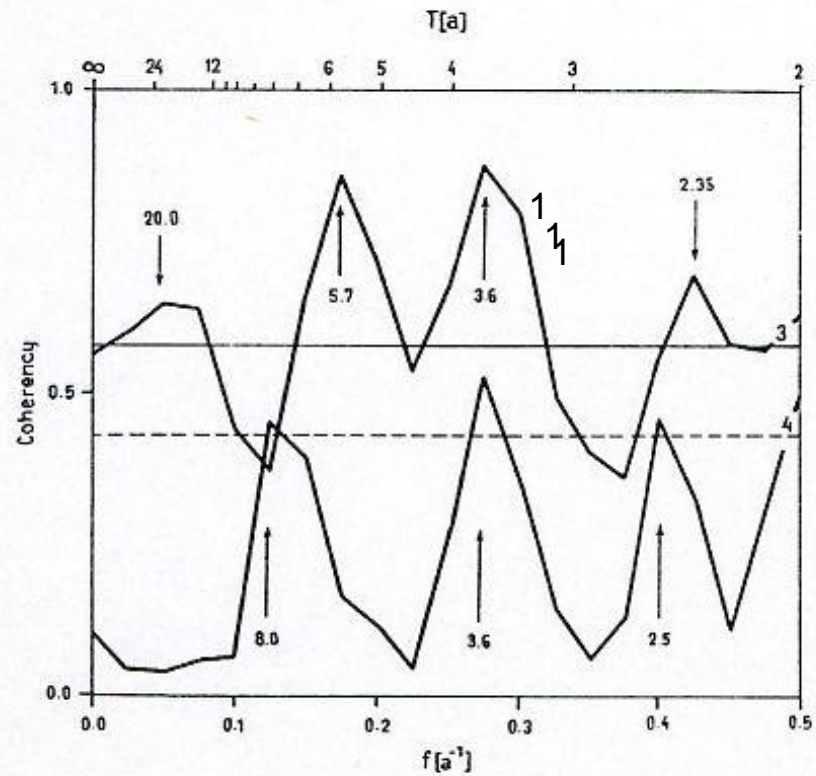
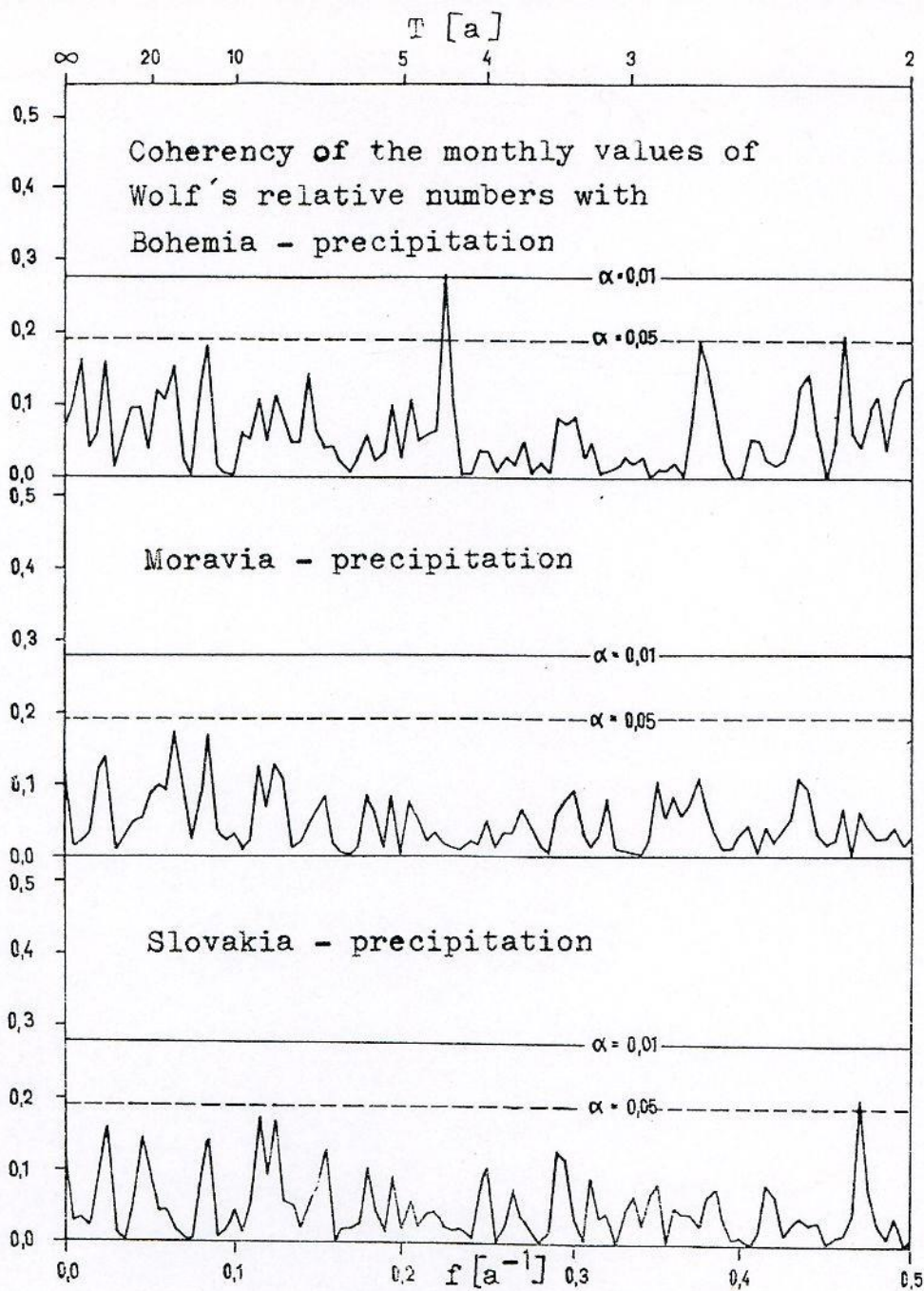
$$K_{xy}^2(h) = \frac{SC_{xy}^2(h) + SQ_{xy}^2(h)}{SP_x(h) + SP_y(h)} \quad h = 1, 2, \dots, M$$

$$0 \leq K_{xy}^2(h) \leq 1$$

Confidence level:

$$\beta = \sqrt{1 - \alpha^{\frac{1}{\nu-1}}} \quad \nu = \frac{2n - M/2}{M}$$

- koherence řad náhodných čísel se významně neliší od nuly, tj. řady jsou navzájem nezávislé
- jestliže je koherence rovna jedné pro všechny frekvence, je řada  $x(t)$  plně závislá na řadě  $y(t)$
- koherence je analogická  $r_{xy}^2$ , ale v závislosti na frekvenci



Koherence ročních úhrnů srážek:

3) Slovensko – Maďarsko

4) Polsko – Maďarsko

Plně  $\alpha = 0,01$ , čárkovaně  $\alpha = 0,05$

## 3.6. Filtrace

- představuje shlazení časové řady, v níž některé cykly (periody) jsou potlačeny a jiné zesíleny

$$x_{iF} = \sum_{k=-n}^m w_k \cdot x_{i+k}$$

$w_k = w_{-n}, w_{-n+1}, \dots, w_{-1}, w_0, w_1, \dots, w_{m-1}, w_m$  - filter weights

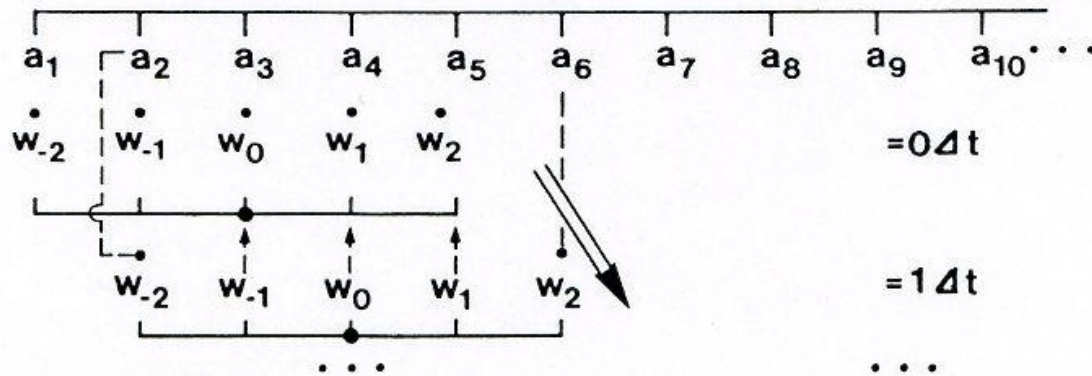
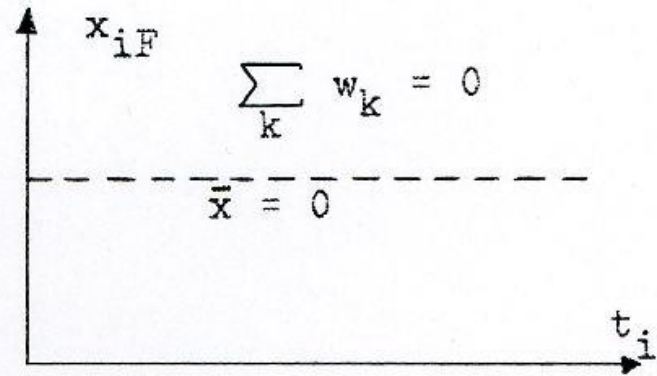
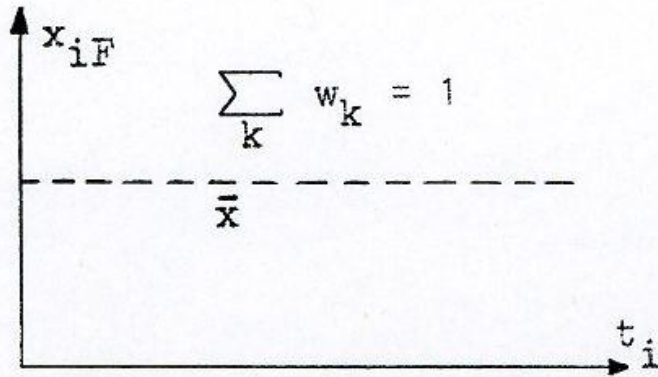


Schéma filtrace časové řady



$n = m$ ,  $w_k = w_{-k}$ ,  $k = 0, 1, 2, \dots$  - symmetrical filter weights

- hledá se kompromis mezi největším  $m$  a zkrácením řady

Response function:

$$R(f) = w_0 + 2 \sum_{k=1}^n w_k \cos 2\pi f k \quad 0 \leq R(f) \leq 1$$

$R(f) = 0$  amplitude completely suppressed

$R(f) = 1$  amplitude without change

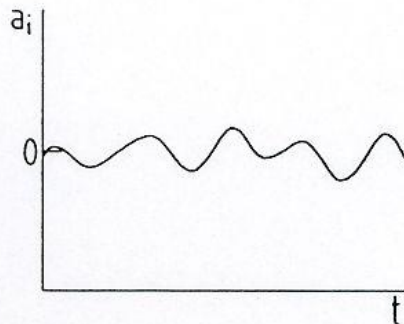
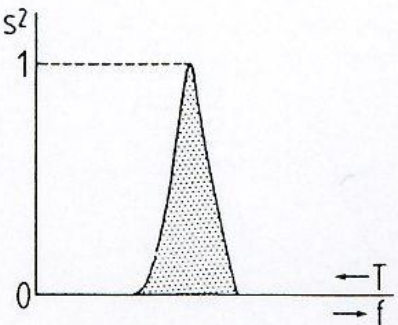
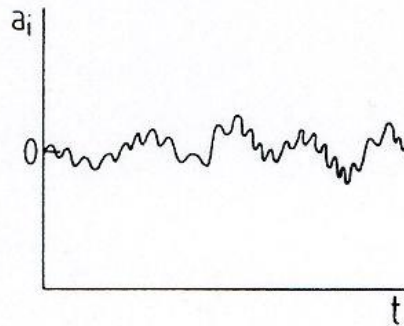
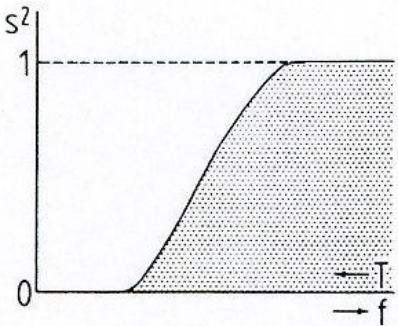
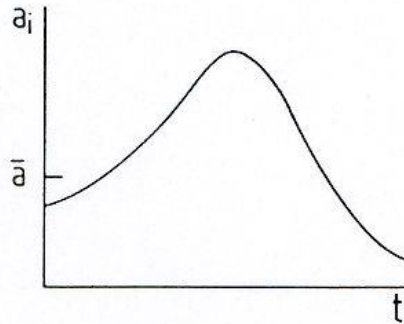
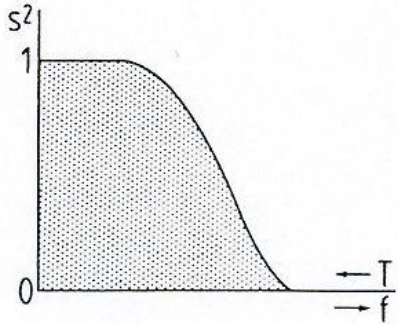
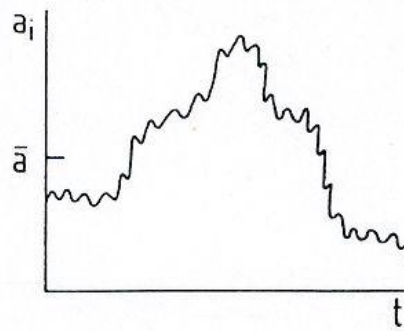
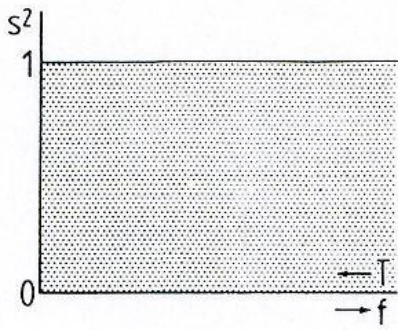
# Filtrační funkce a časová řada

Bez filtrace – výchozí údaje

Nízkofrekvenční filtrace

Vysokofrekvenční filtrace

Pásmová filtrace



Moving (running) averages:

$$R(f) = \text{sin} (\pi f l) / \pi f l$$

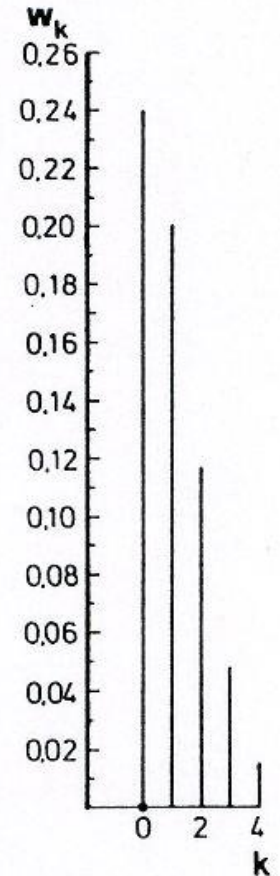
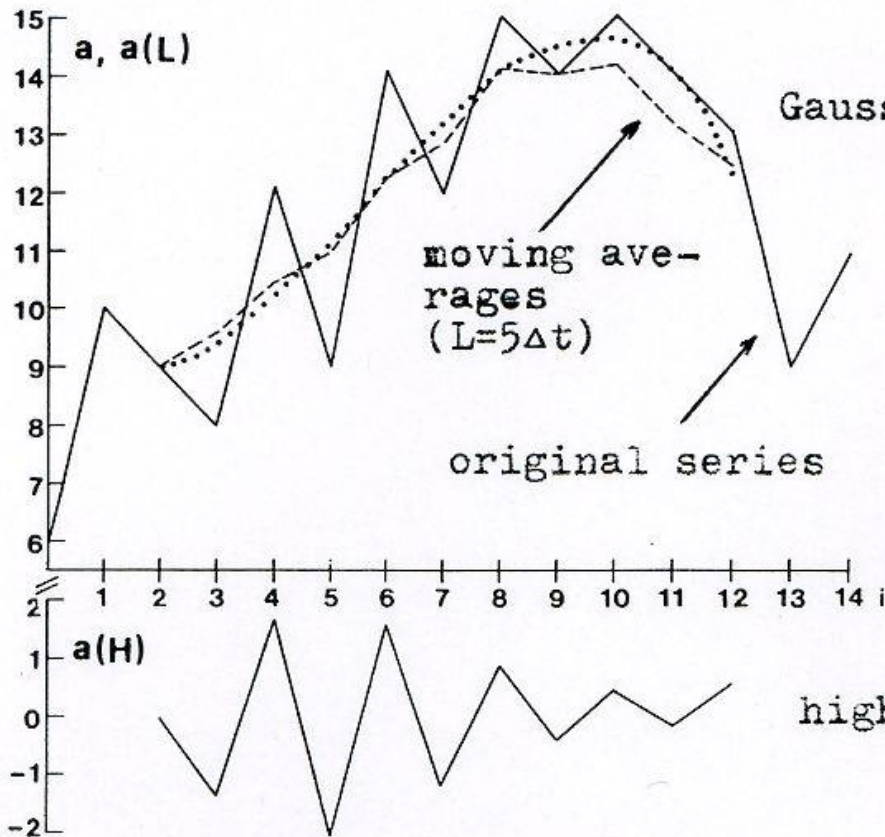
$$w_k = 1/L$$

$$L = l \cdot \Delta t$$

Gauss filtering:

$$R(f) = \text{exp} \left( -\frac{1}{3} \pi^2 f^2 \right)$$

$$w_k = z \left( k \cdot \frac{6}{T^*} \right)$$



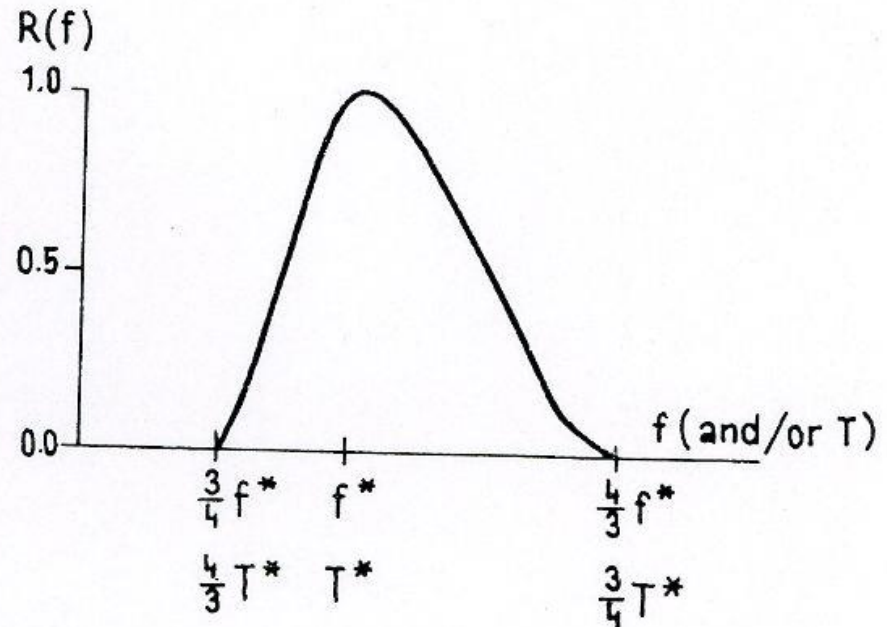
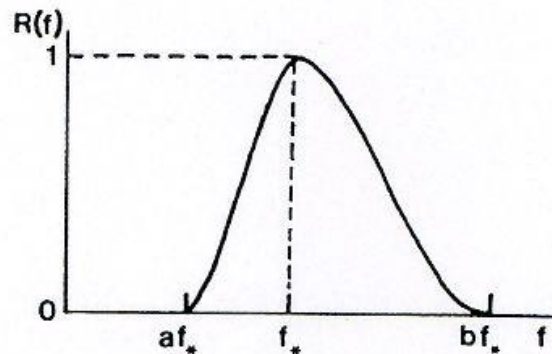
Gauss filter weights



## Pásmový filtr

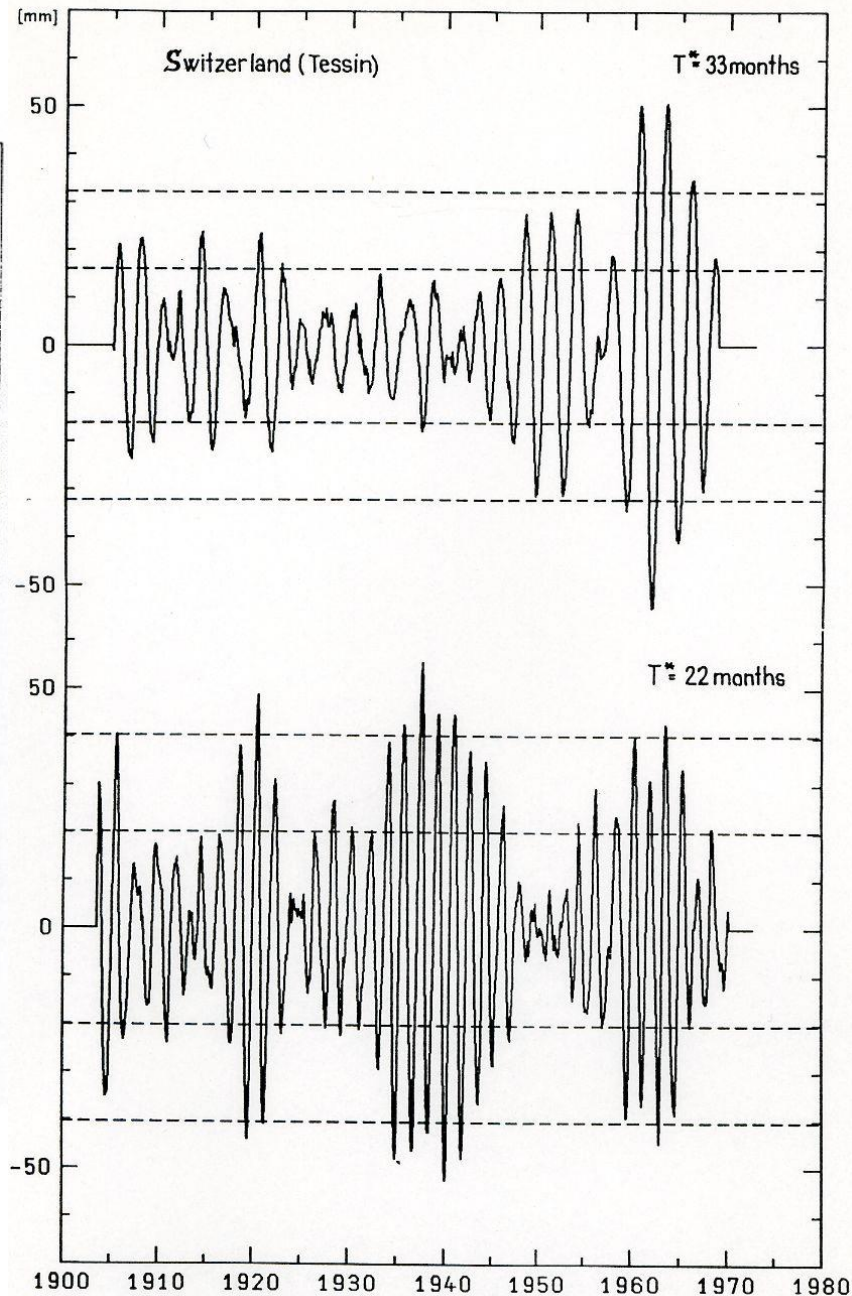
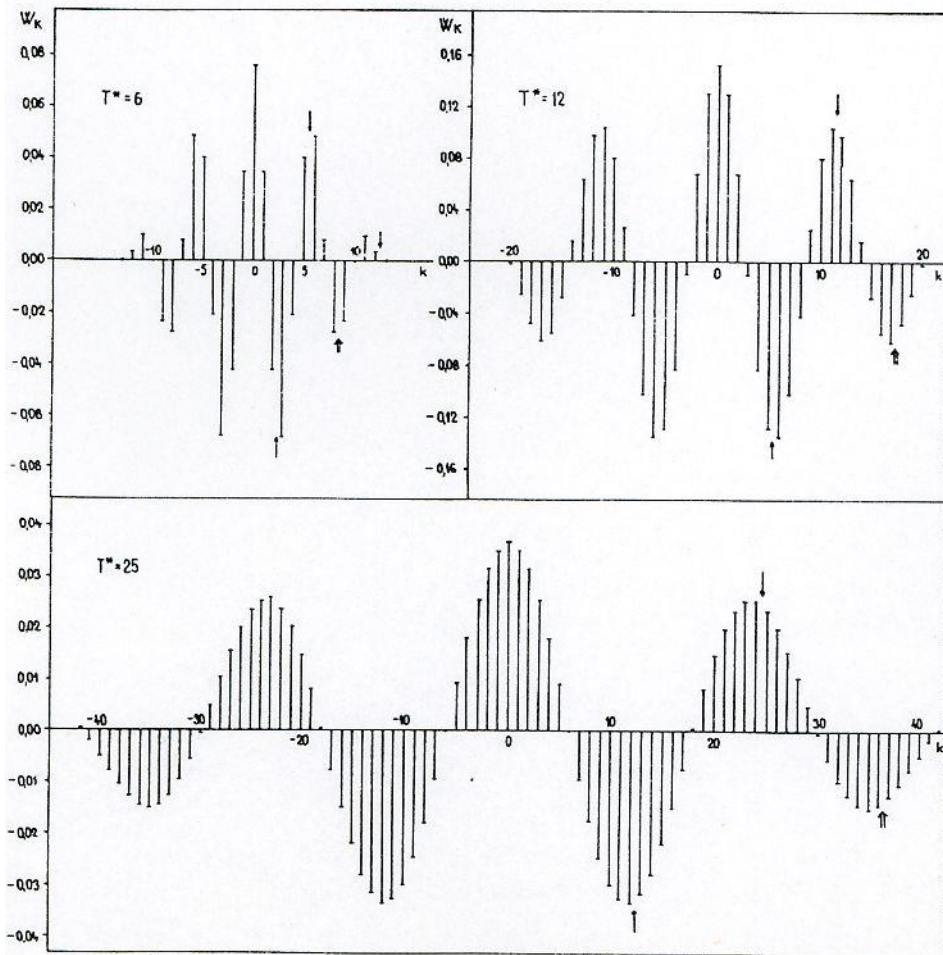
$$\begin{aligned}
 R(f) &= 0 && \text{for } \frac{1}{T_n} < f < af^* && \text{and/or } T_n > T > bT^* \\
 2R(f) &= 1 + \cos(2T^*\pi f) && \text{for } af^* < f < f^* && \text{and/or } bT^* > T > T^* \\
 2R(f) &= 1 - \cos(T^*\pi f) && \text{for } f^* < f < bf^* && \text{and/or } T^* > T > aT^* \\
 R(f) &= 0 && \text{for } bf^* < f < \frac{1}{2\Delta t} && \text{and/or } aT^* > T > 2\Delta t
 \end{aligned}$$

with  $b > a$ ,  $T_n = n \cdot \Delta t$



- pásmová filtrace není  
důkazem cykličnosti řady

# Filtrační koeficienty



Filtrace měsíčních úhrnů srážek pásmovým filtrem pro  $T = 33$ , resp. 22 měsíců

## 3.7 Metody klouzavé analýzy

- podmínka stacionarity časové řady může být snadněji splněna v jejich kratších realizacích → aplikace metod klouzavé analýzy

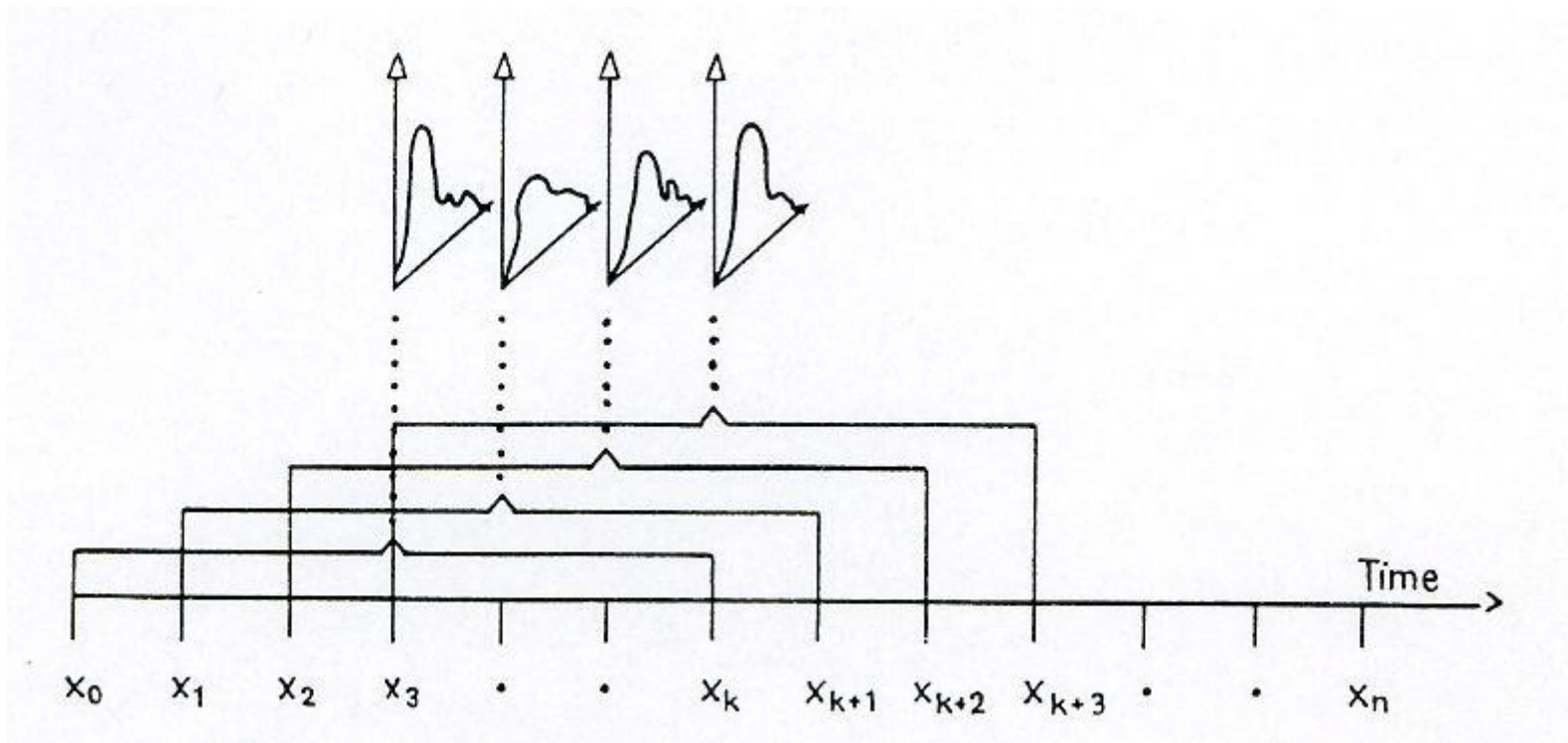


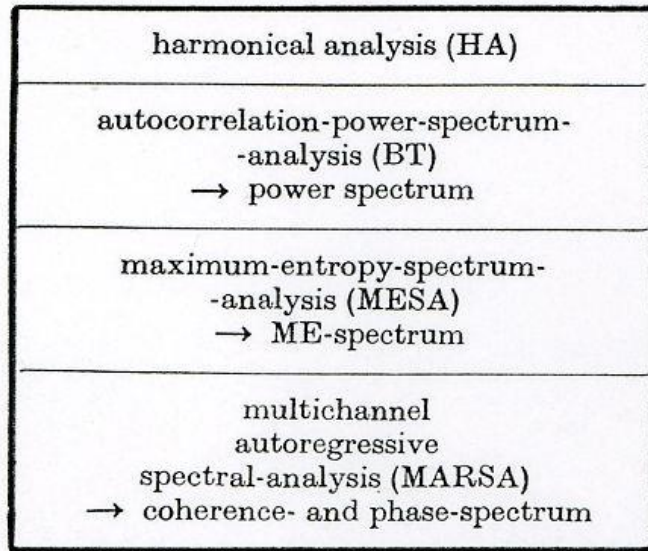
Schéma výpočtu dynamické (klouzavé) MESA

Methods to detect time dependent changes  
in the spectral behaviour of time series

trend analysis

stationarity:

analysis-operator applied to the whole data series

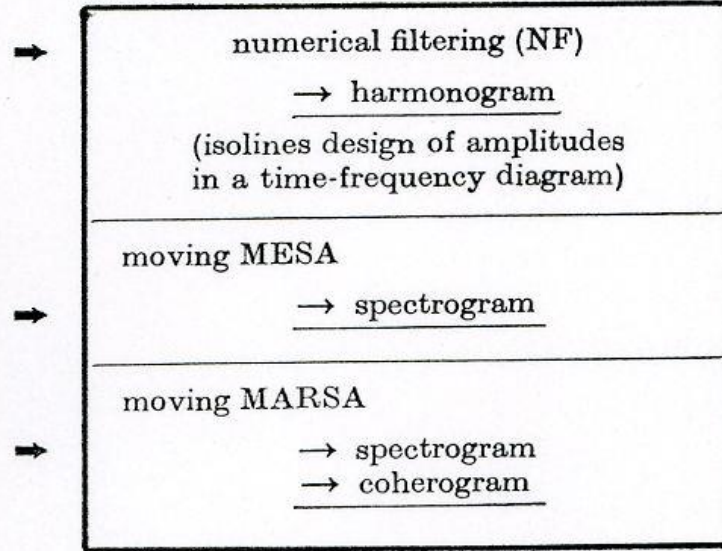


“integral” informations (regarding the whole time series) about:

amplitudes, phases, spectra with  
“peaks”, stochastic relationships  
in several frequency bands

quasi – stationarity:

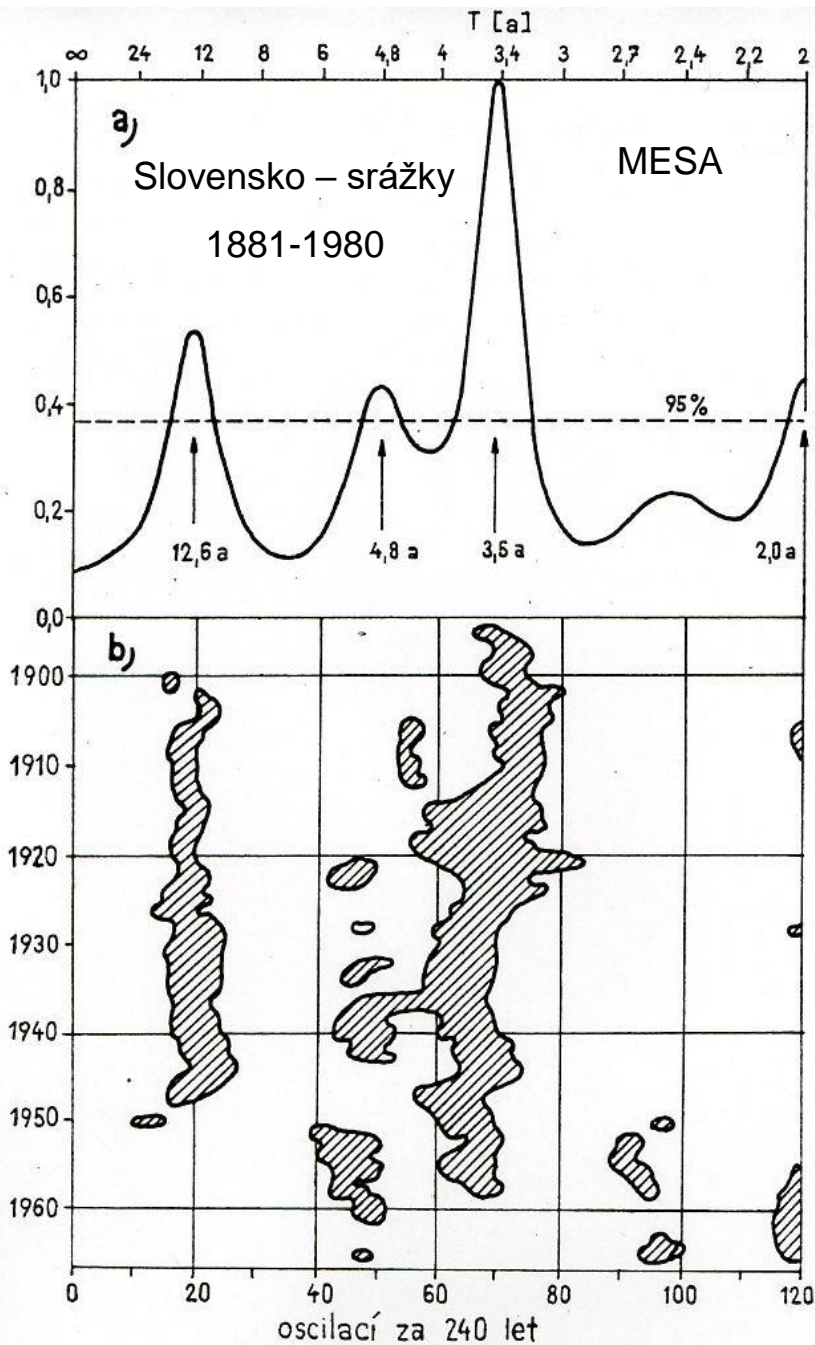
analysis-operator moving\*) applied to subintervals



informations on separate time intervals  
of the data series:

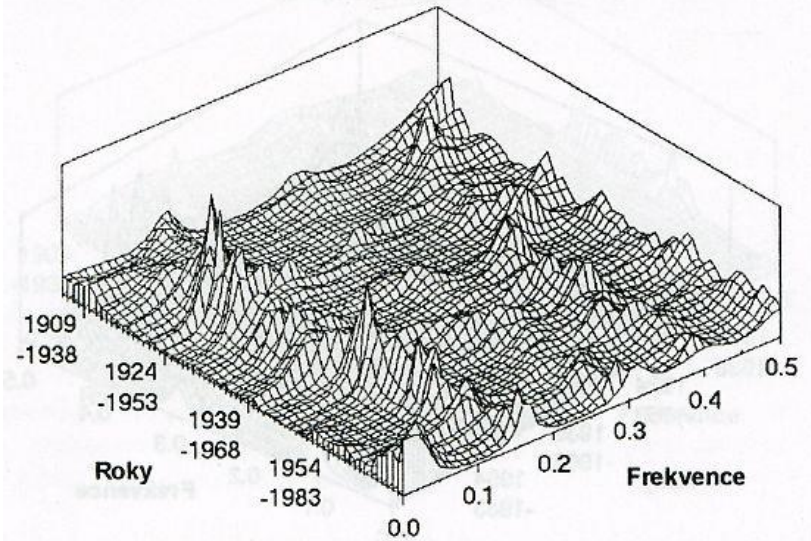
different development of an oscillation  
in the course of time, time dependent  
changes of the spectral behaviour  
and the coherence of two time series

\*) moving, or sliding, or adaptive, or dynamic method.

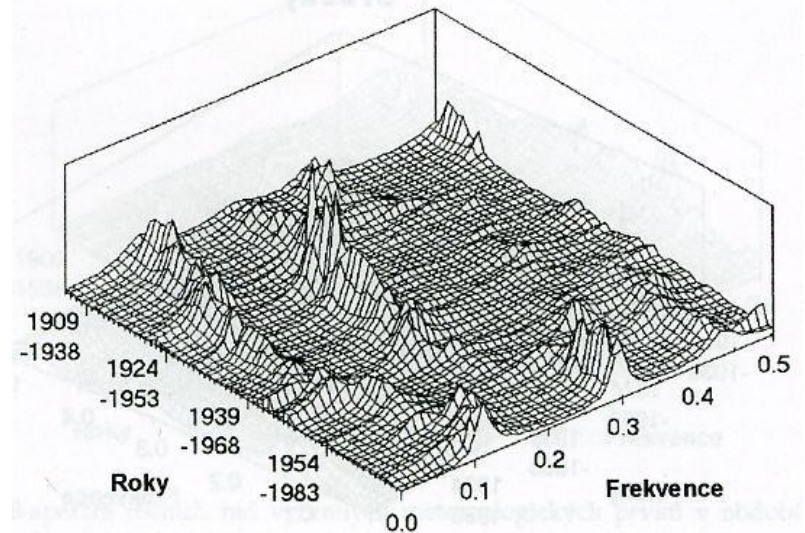


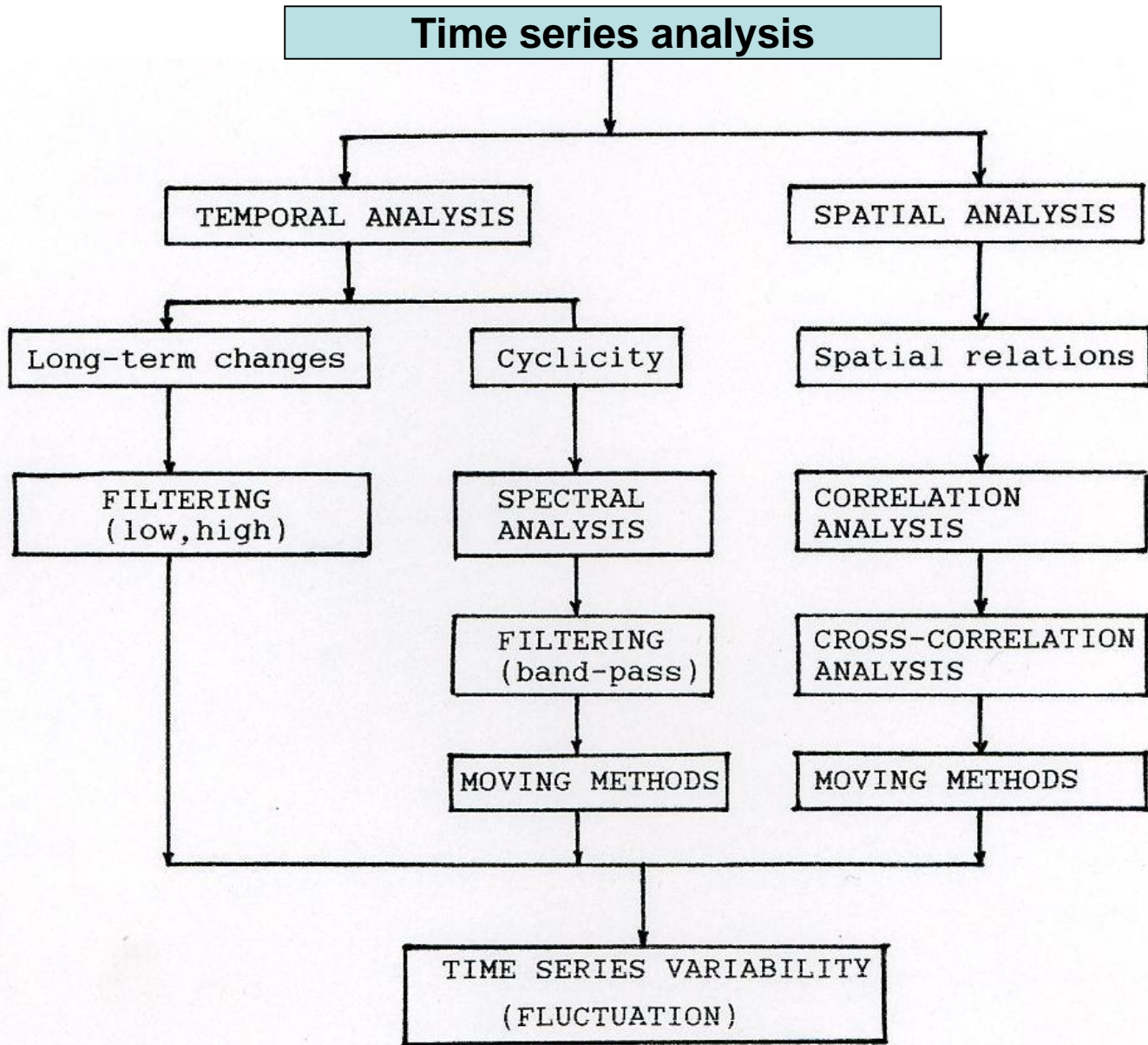
## Milešovka 1905-1994

### Tlak vodní páry



### Relativní vlhkost vzduchu





## 3.8 Wavelet analýza

- časově-frekvenční analýza signálu v nestacionárních řadách

Wavelet transformace  $W_x^\psi$  časové řady  $x(t)$  je definována výrazem (1):

$$W_x^\psi(\tau, \lambda) = \frac{1}{\sqrt{|\lambda|}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t - \tau}{\lambda} \right) dt, \quad (1)$$

kde  $\psi^{(\tau, \lambda)}(t)$  je tzv. wavelet (wavelet – česky vlnka). Wavelet si můžeme představit jako okénko, pomocí kterého se časová řada  $x(t)$  postupně zkoumá. Parametr  $\tau$  (translace) určuje část signálu  $x(t)$ , kde je právě wavelet lokalizován a druhý parametr  $\lambda$  (tzv. škála) vystupuje v definici (1) jako frekvenční parametr, hvězdička \* zde značí proměnnou komplexně sdruženou. Wavelet transformace je v podstatě konvolucí původního signálu  $x(t)$  s novou funkcí  $\psi^{(\tau, \lambda)}(t)$ .

(Konvoluce je matematická operace kombinující dva signály tak, aby vznikl třetí.)

Wavelet je funkce oscilující kolem  $x$ -ové osy, která dále musí splňovat některé podmínky, jako například, že její integrál je roven nule (2) anebo, že integrál kvadrátu je roven jedné (3) (podrobněji viz například [4]).

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (2)$$

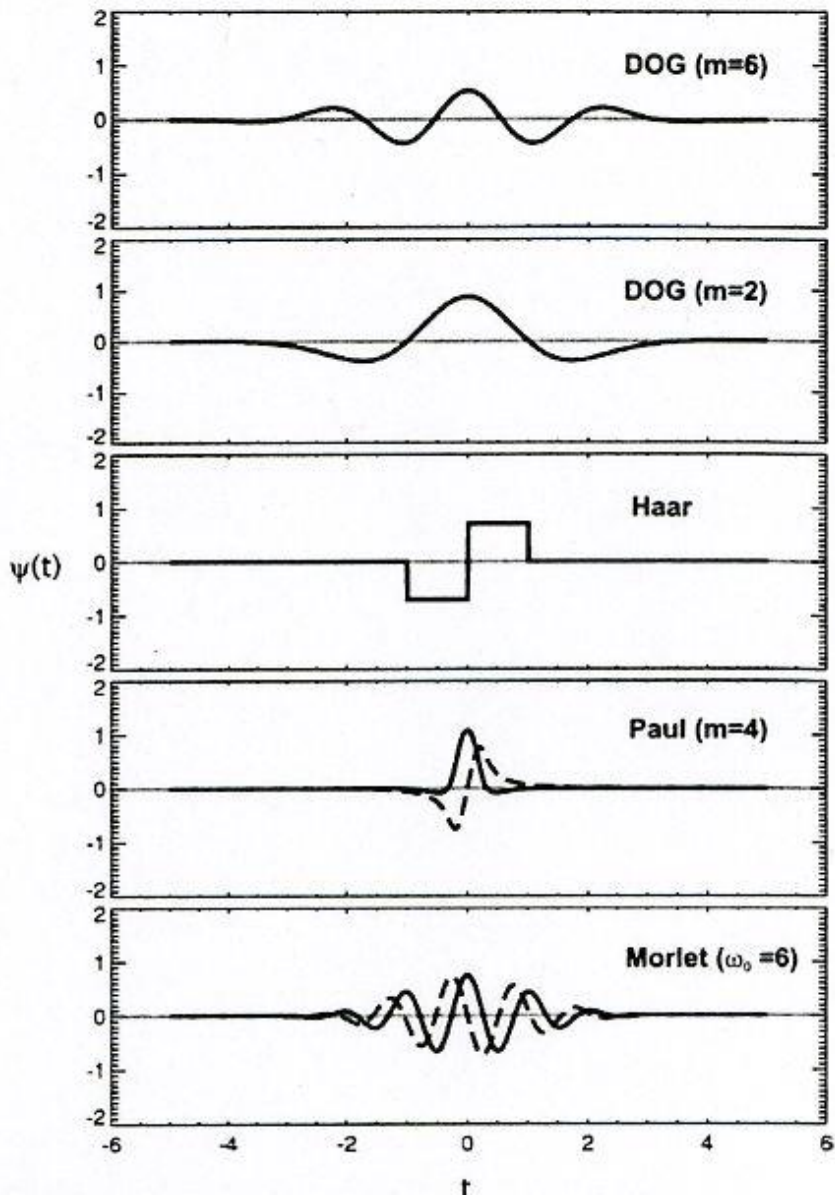
$$\int_{-\infty}^{\infty} \psi^2(t) dt = 1 \quad (3)$$

Všechny wavelety závisí na tzv. mateřském waveletu (mother wavelet), ze kterého se generují další tzv. dceřiné wavelety (daughter wavelets). V tab. 1 jsou uvedeny definice tří často používaných waveletů a grafy na obr. 1 pak ukazují, jak tyto mateřské wavelety vypadají. Symboly v tab. 1 označují parametry mateřských waveletů ( $m$ ,  $\omega_0$ ), gamma funkci ( $\Gamma$ ), komplexní jednotku ( $i$ ) a početní konstantu ( $C$ ). Grafy na obr. 2 ilustrují způsob, jakým jsou dále jednotlivé wavelety generovány – jedná se o postupnou translaci a kontrakci daného waveletu podél signálu v závislosti na parametrech  $\tau$  a  $\lambda$ .

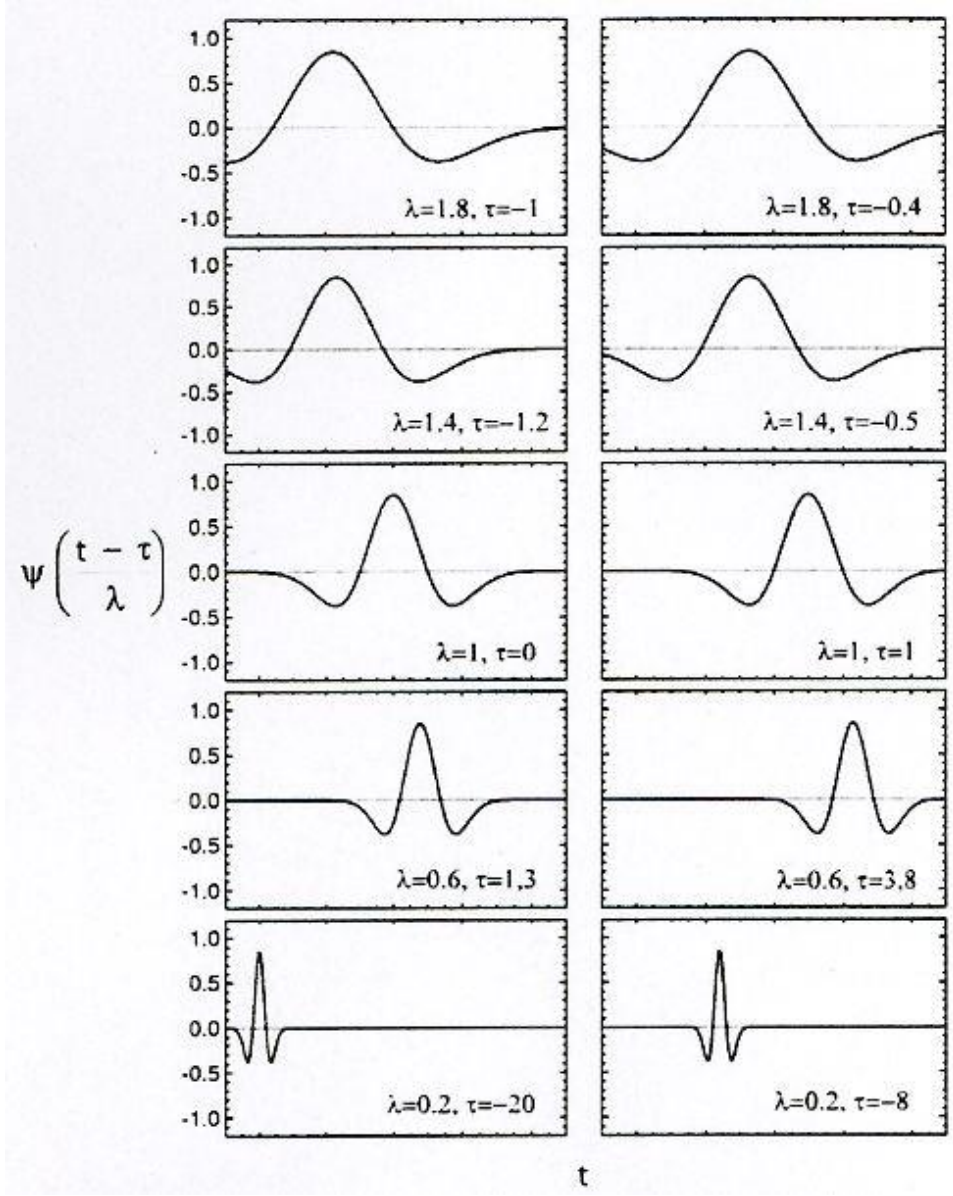
DOG	$\frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m + \frac{1}{2}\right)}} \frac{d^m}{dt^m} e^{-\frac{t^2}{2}}$
Paul	$\frac{2^m i^m m!}{\sqrt{\pi(2m)!}} (1 - it)^{-(m+1)}$
Haar	$\psi(t) = \begin{cases} -1/\sqrt{2}, & t \in (-1; 0) \\ 0, & t \in (-\infty; -1) \cup (1; \infty) \\ 1/\sqrt{2}, & t \in (0; 1) \end{cases}$
Morlet	$C e^{-i\omega_0 t} \left( e^{-\frac{t^2}{2}} - \sqrt{2} e^{-\frac{\omega_0^2}{4}} e^{-t^2} \right)$

Tab. 1. Mateřské wavelety DOG, Paul, Haar, Morlet

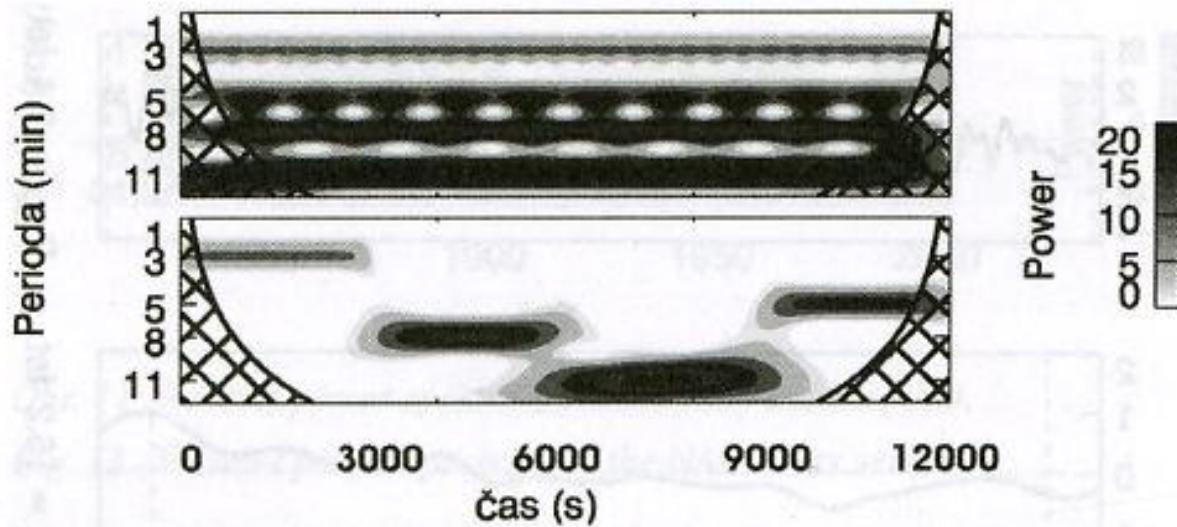




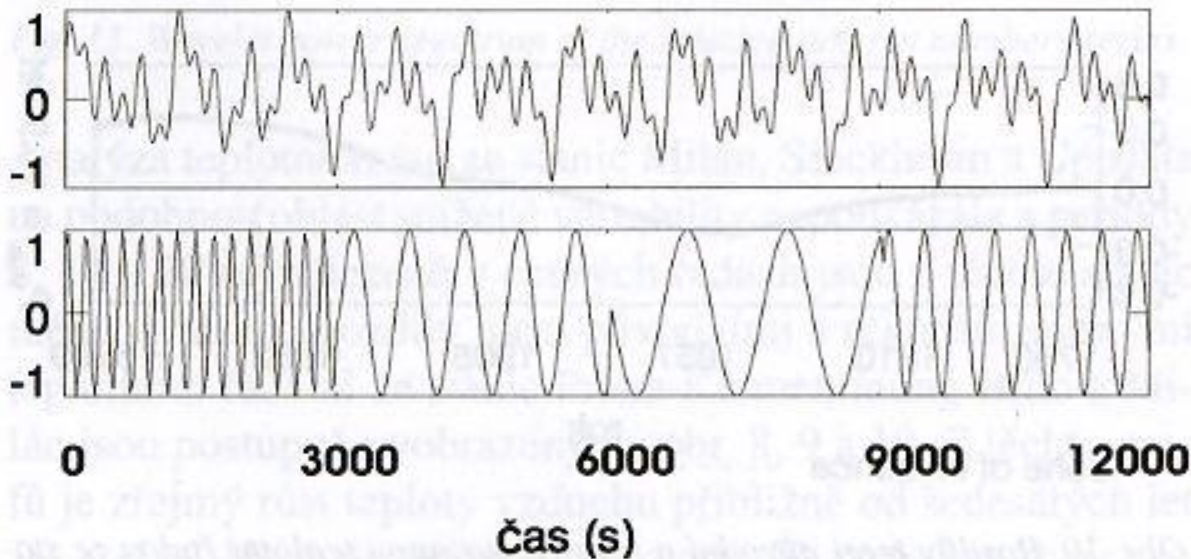
Obr. 1. Reálná (plně) a imaginární (přerušované) části mateřských waveletů



Obr. 2. Generace waveletů – posun a kontrakce mateřského waveletu pro různé  $\tau$  a  $\lambda$



Obr. 3 Spojitá wavelet transformace stacionárního (nahore) a nestacionárního (dole) signálu.



Obr. 4 Stacionární (nahore) a nestacionární (dole) testovací signál.

## Výběr mateřského waveletu

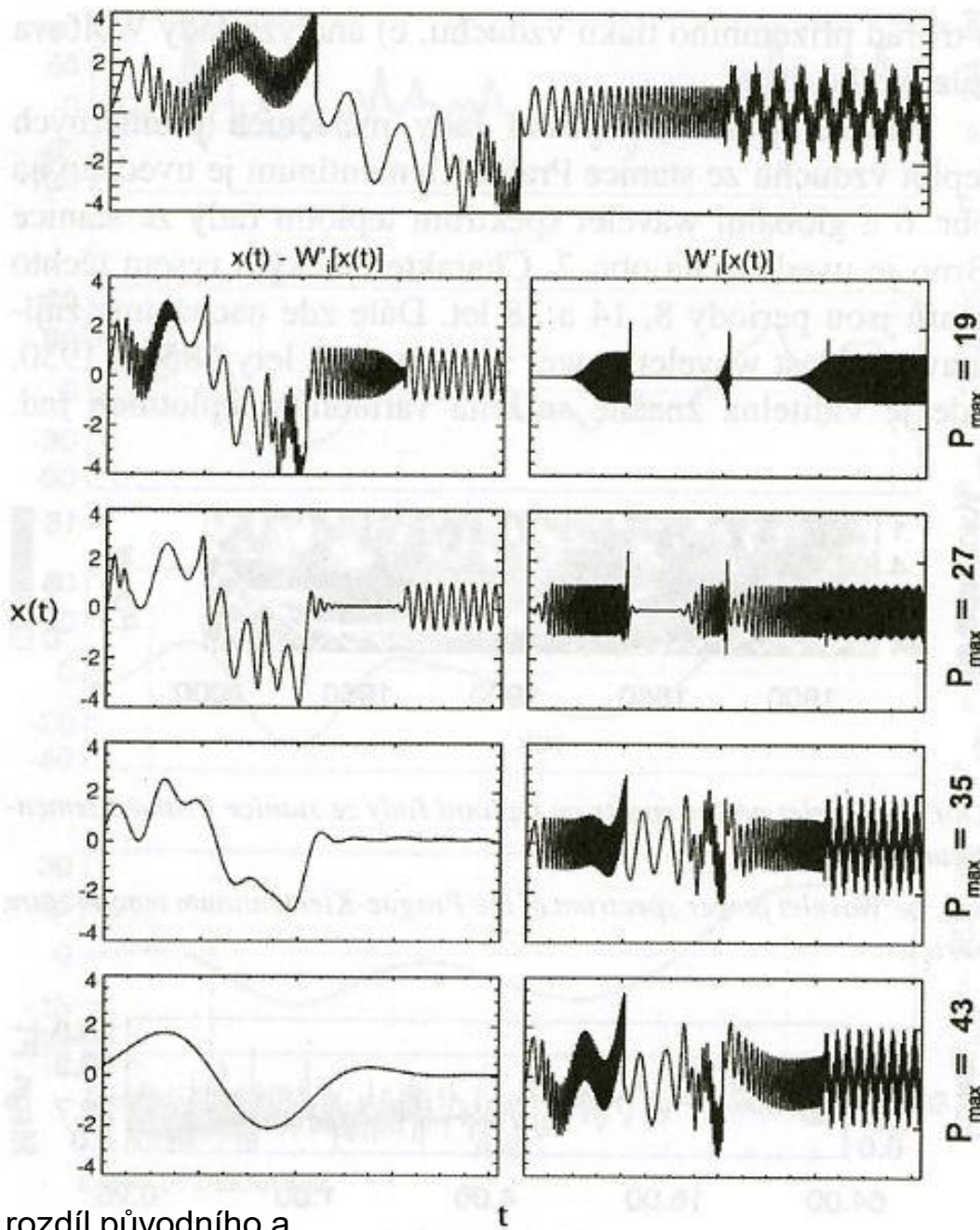
- Morlet – dobré časové a frekvenční rozlišení
- Paul – malý „cone of influence“ (nekorektní výsledky)

## Zobrazení výsledků

- wavelet power spektrum – kvadrát absolutní hodnoty wavelet transformace  $|W_x^\psi|^2$
- globální wavelet spektrum – integrál wavelet power spektra „podél časové osy“  $\frac{1}{t} \int_0^t |W_x^\psi|^2 dt$

# Rekonstrukce testovacího signálu

← původní signál  $x(t)$



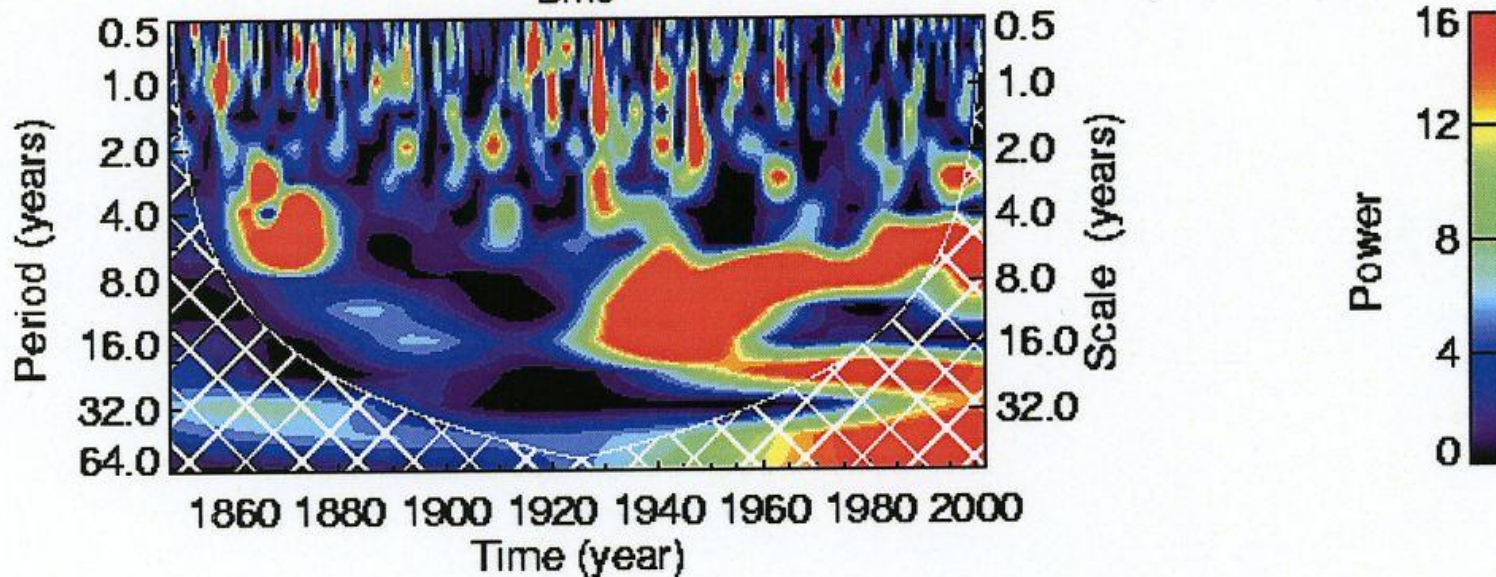
$P_{\max}$  určuje hranice největší (vpravo), resp. nejmenší (vlevo) detekované periody.

rozdíl původního a rekonstruovaného signálu

rekonstrukce signálu (inverzní wavelet transformace)

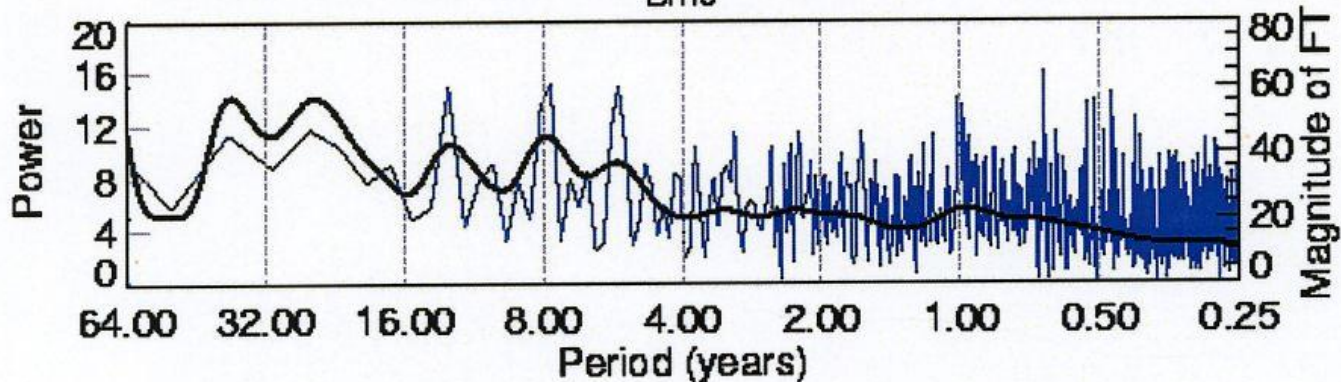
# Wavelet power spectrum

Brno



# Global wavelet spectrum and Fourier transformation

Brno



Fourier transformation \_\_\_\_\_

Global wavelet spectrum \_\_\_\_\_

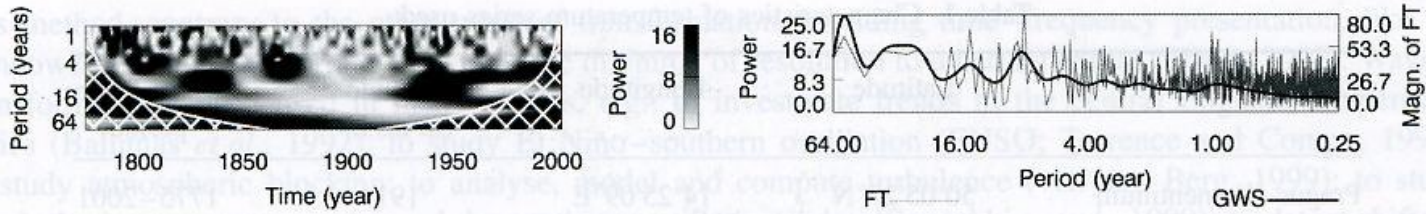


Figure 1. Continuous wavelet spectrum (left) and global wavelet spectrum (GWS; right) of the Prague-Klementinum temperature series

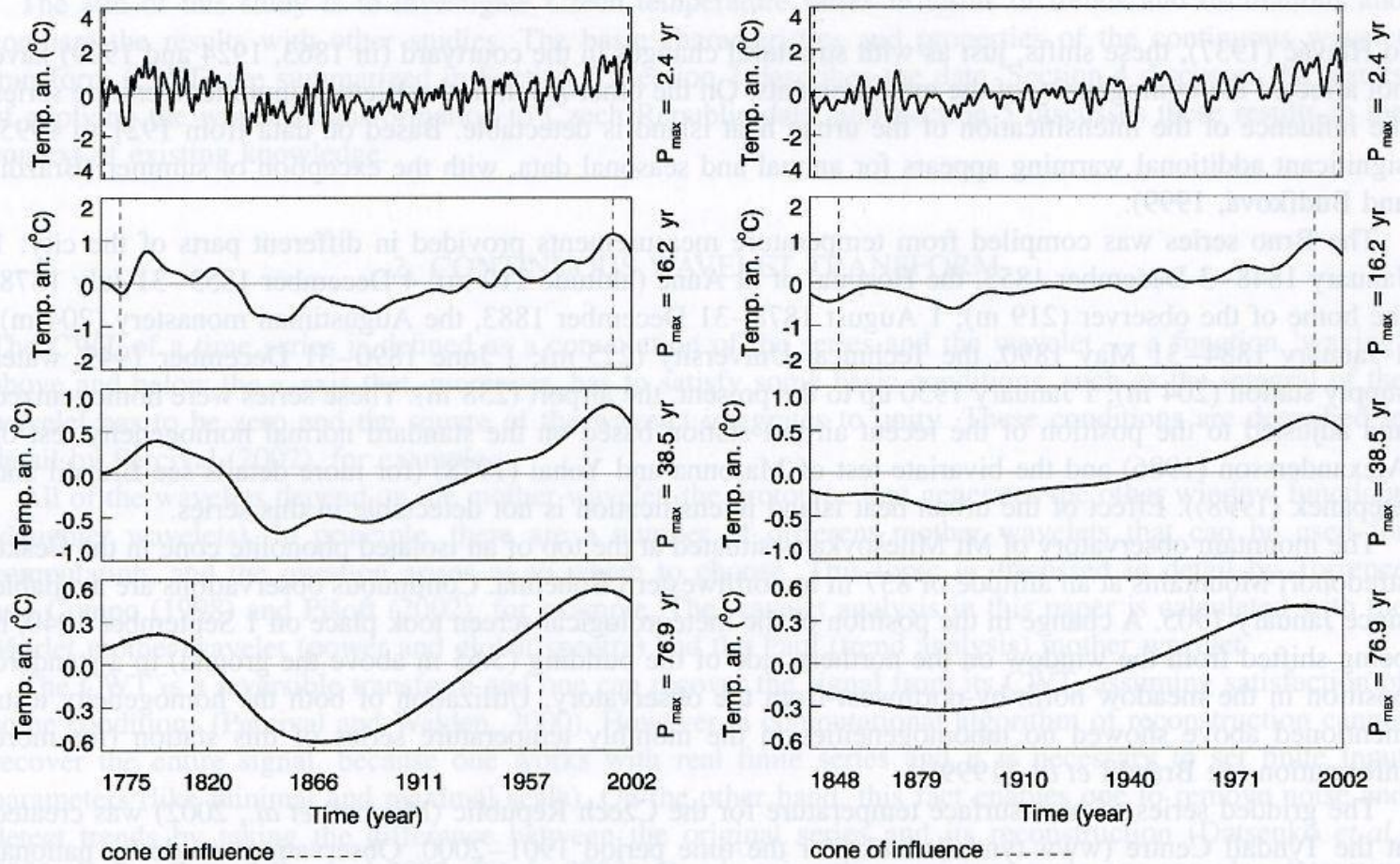
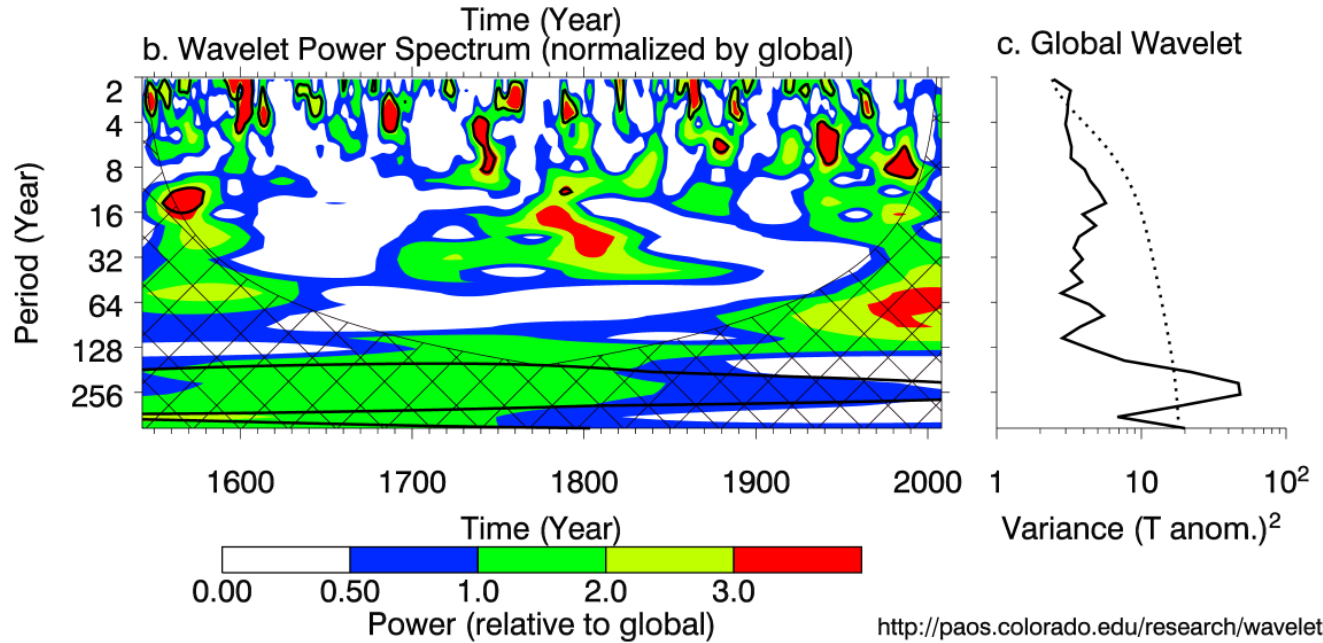


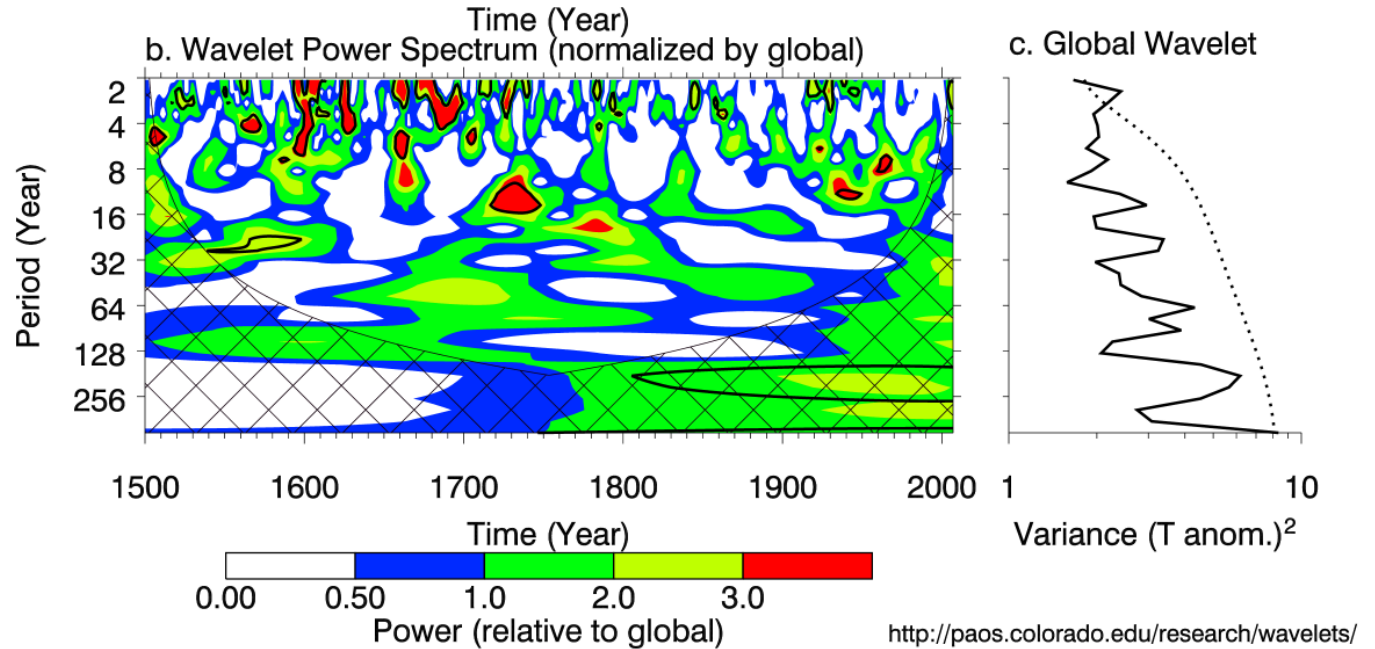
Figure 2. Differences between Prague-Klementinum (left) and Brno (right) temperature series and their inverse CWTs

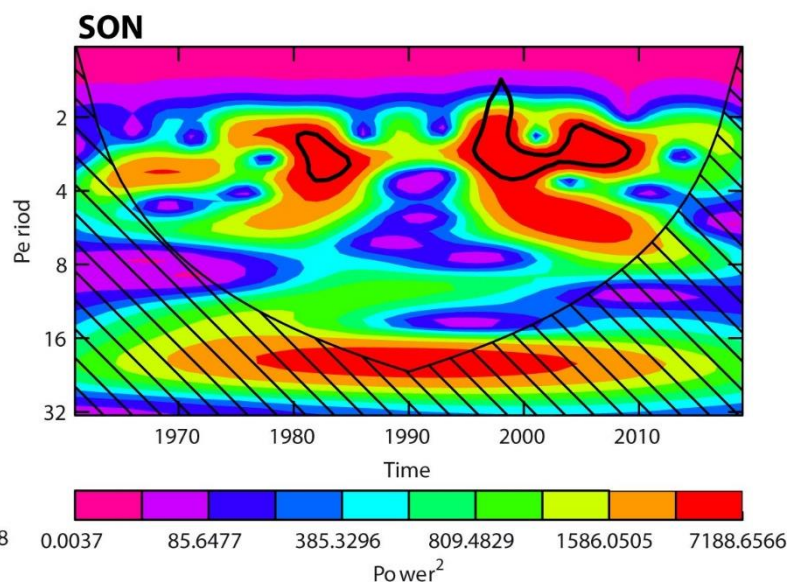
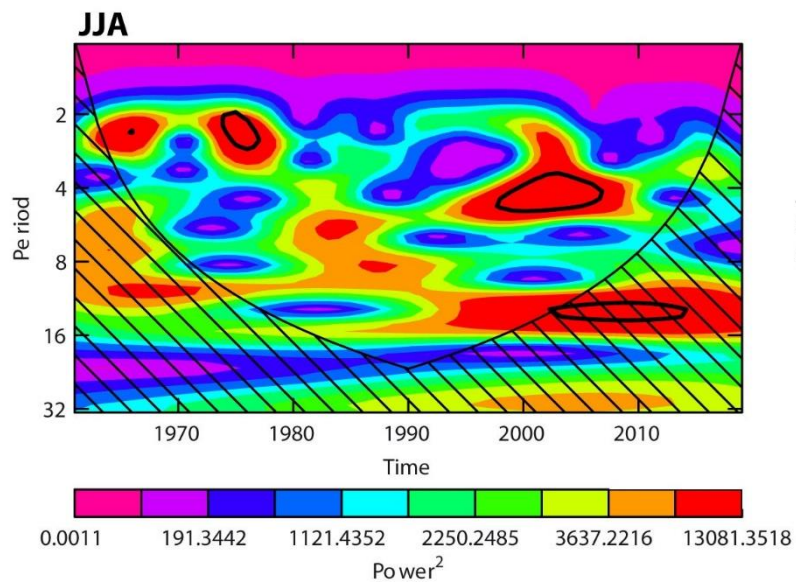
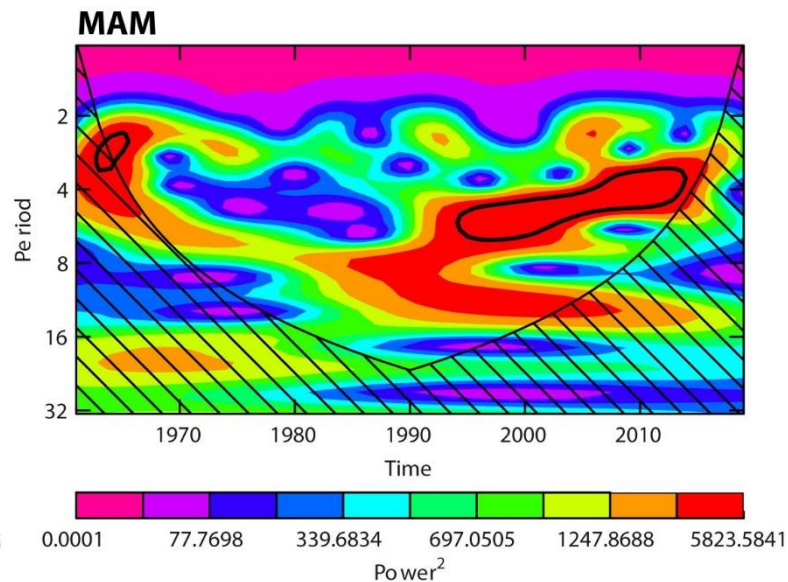
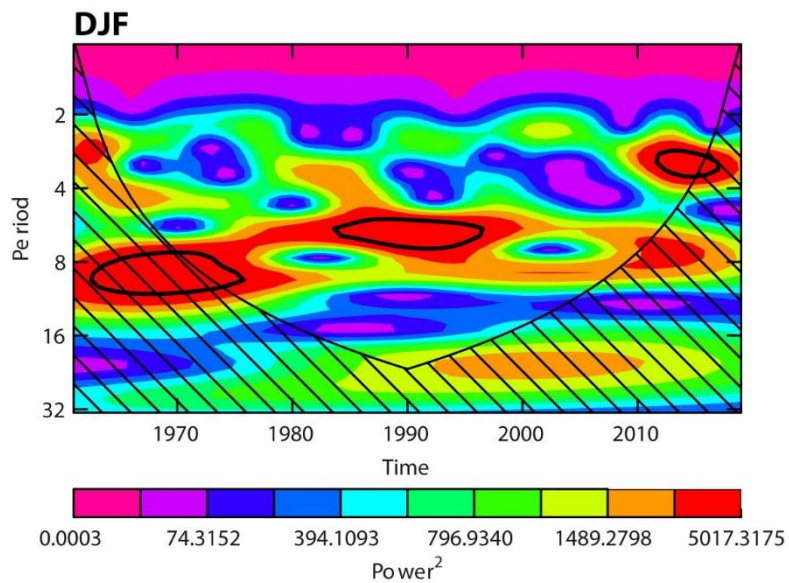
JFMA  
Stockholm  
1543-2008



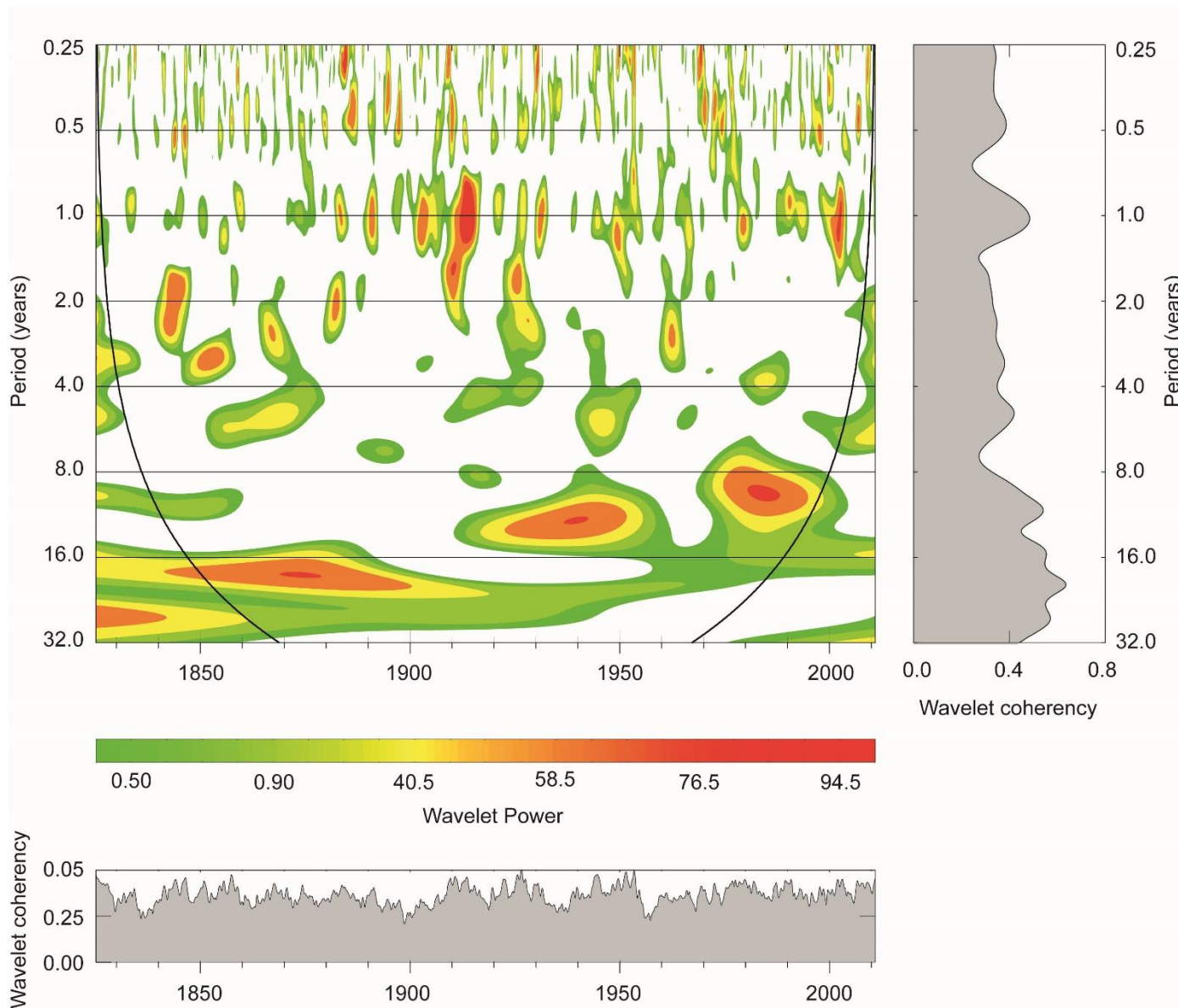
B.

Teplota JFMA  
Střední Evropa  
1500-2007





Wavelet analýza sezonních měsíčních úhrnů srážek v České republice v období 1961–2019 indikuje existenci statisticky významných cyklů hlavně kolem 4-8 let (Brázdil et al., IJC, 2021)

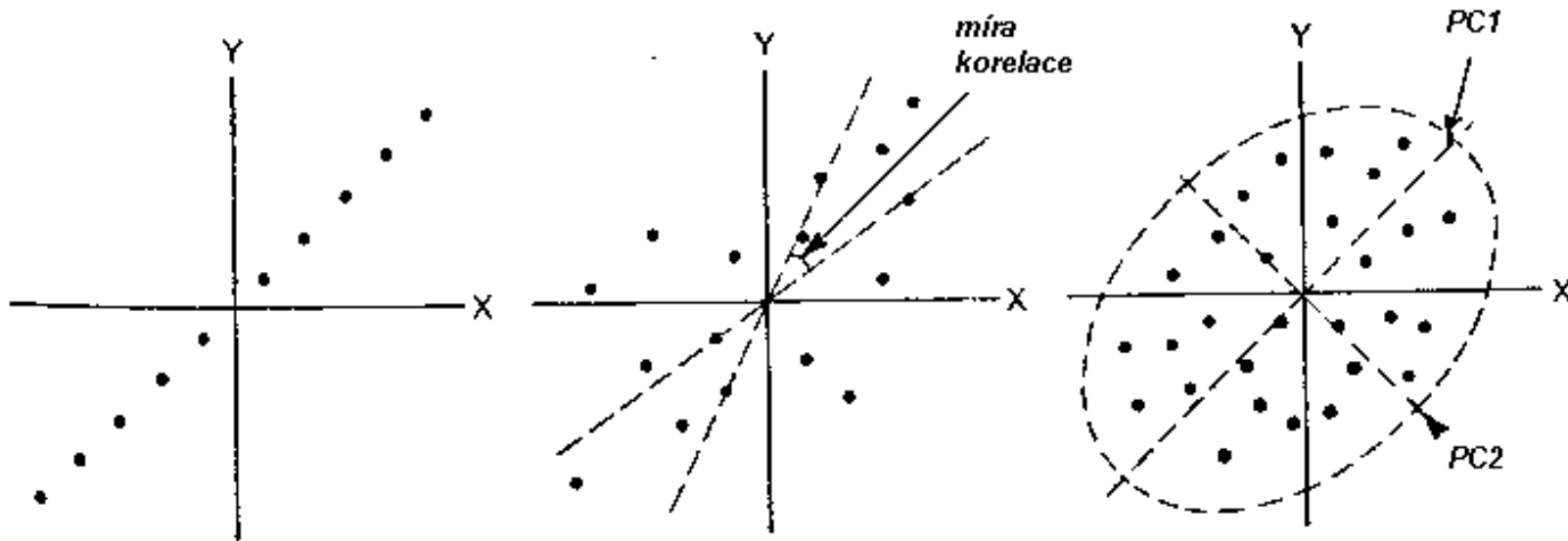


Koherenční wavelet analýza – srážky Brno - NAOI



## 3.9 Analýza hlavních komponent (Principal Component Analysis)

Cílem metody je **eliminování duplicit** a **zhuštění informace** obsažené v původních proměnných do menšího počtu vzájemně nekorelovaných proměnných. Tyto nové proměnné (**hlavní komponenty**) popisují soubor syntetičtěji a úsporněji.



# Redukce dimenzionality – příliš mnoho proměnných

Příklady:

- jedna meteorologická proměnná měřená v  $n$  případech na  $p$  místech
- $p$  meteorologických proměnných měřených na  $n$  stanicích v jednom případě
- $p$  meteorologických proměnných měřených na jedné stanici v  $n$  případech.

Je-li  $p$  proměnných vysoce korelováno, lze vybrat  $m$  proměnných ( $m \ll p$ ), které poskytují veškerou originální informaci o  $p$  proměnných.

Strategie výběru  $m$  proměnných:

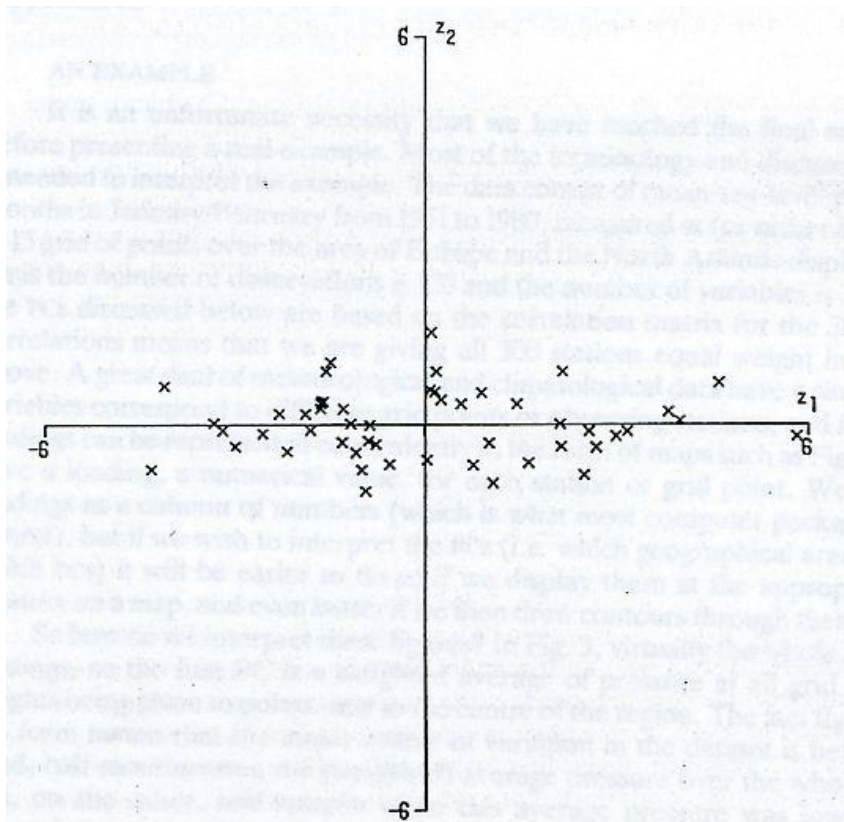
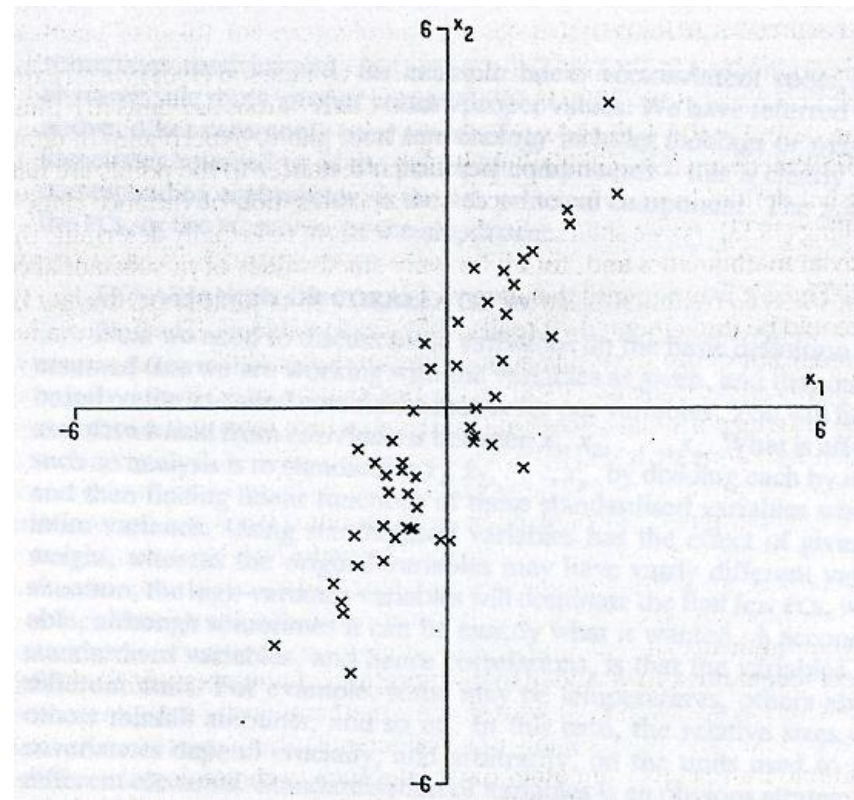
- podsoubor z originálních  $p$  proměnných
- vytvoření nových proměnných z  $p$  původních, které jsou od nich typově odlišné (při téže ztrátě informace lze získat větší redukci dimenzionality)

# Analýza hlavních komponent

- $p$  proměnných  $x_1, x_2, \dots, x_p$
- lineární funkce  $p$  proměnných:  $z = a_1x_1 + a_2x_2 + \dots + a_px_p$ , kde  $a_1, a_2, \dots, a_p$  jsou konstanty
- změní-li se konstanty, dostaneme odlišnou lineární funkci a lze spočítat její rozptyl
- první hlavní komponenta (PC1) je lineární funkcí, která má maximální možný rozptyl; druhá hlavní komponenta (PC2) je lineární funkcí s maximálně možným rozptylem nekorelovatelná s PC1; třetí hlavní komponenta PC3 je lineární funkcí s maximálně možným rozptylem nekorelovatelná s PC1 a PC2, atd.
- PCA nabízí  $m$ -dimenzionální prezentaci údajů pro  $m = 1, 2, \dots, p - 1$ , tedy pro každé  $m = 1, 2, \dots, p - 1$  dostáváme  $m$  lineárních funkcí proměnných  $x_1, x_2, \dots, x_p$  ukazující maximálně možný podíl originálních variací

## Příklad (dvourozměrný):

- 50 měření teploty půdy ( $x_1$ ) a teploty vzduchu ( $x_2$ )
- lineární funkce  $x_1$  a  $x_2$  objasňuje více kolísání než každá proměnná samostatně



- lineární funkce, která maximalizuje rozptyl, je první PC –  $z_1$ , tj. všechna kolísání mohou být nyní vyjádřena s ohledem na  $z_1$
- obrázky se shodují s tím, že došlo k rotaci os – druhá PC ( $z_2$ ) jde pak kolmo na  $z_1$
- PC minimalizuje sumu čtverců podle vzdáleností kolmých k  $z_1$  (předchozí podle vertikálních vzdáleností)

# Vlastní hodnoty a vlastní vektory (eigenvalues and eigenvectors)

- $k$ -tá hlavní komponenta je dána vztahem  $z_k = a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kp}x_p$  pro  $k = 1, 2, \dots, p$
- **vlastní vektor**: první vlastní vektor – koeficienty  $a_{11}, a_{12}, \dots, a_{1p}$ ; další vlastní vektory jsou tvořeny koeficienty u proměnných  $x_1, x_2, \dots, x_p$
- **vlastní hodnota** – rozptyl první hlavní komponenty (míra významnosti objasněných kolísání), atd.

## Vstupní matice výpočtu a standardizace hodnot

		Variables			
		1	2	...	m
Cases	1				
	2				
	3				
	...				
	n				

- standardizací se přisuzuje proměnným stejná váha (např. proměnné v různých jednotkách)

$$t_i = \frac{x_i - \mu}{\sigma}$$

# Příklad – typický výstup PCA (I)

No.	Eigenvalue	Individual Cumulative		Scree Plot
		Percent	Percent	
1	3.487151	38.75	38.75	
2	2.130173	23.67	62.41	
3	1.098958	12.21	74.63	
4	0.994483	11.05	85.68	
5	0.543218	6.04	91.71	
6	0.383428	4.26	95.97	
7	0.225754	2.51	98.48	
8	0.136790	1.52	100.00	
9	0.000046	0.00	100.00	

- pořadové číslo nové proměnné (PC - **hlavní komponenty**)
- **vlastní hodnota** – část z celkového rozptylu původních dat vysvětlená každou z nových komponent
- procentuální vyjádření množství **rozptylu vysvětleného** komponentou
- **kumulativní** hodnota procentuálního podílu vysvětleného příslušnými komponentami (např. první 4 komponenty vysvětlují 85,68 % celkové variability původních dat)
- tzv. **sutinový graf** sloužící k určení počtu významných komponent

## Příklad – typický výstup PCA (II)

Variables	Factor1	Factor2	Factor3	Factor4
AGR	0.523791	0.053594	0.048674	0.028793
MIN	0.001323	0.617807	-0.201100	0.064085
MAN	-0.347495	0.355054	-0.150463	-0.346088
PS	-0.255716	0.261096	-0.561083	0.393309
CON	-0.325179	0.051288	0.153321	-0.668324
SER	-0.378920	-0.350172	-0.115096	-0.050157
FIN	-0.074374	-0.453698	-0.587361	-0.051567
SPS	-0.387409	-0.221521	0.311904	0.412230
TC	-0.366823	0.202592	0.375106	0.314372

Tzv. **zátěže** (loadings) - představují míru korelace mezi původními a novými proměnnými.

# Objektivní klasifikace cirkulačních typů ve dnech s vichřicemi

(P. Dobrovolný)

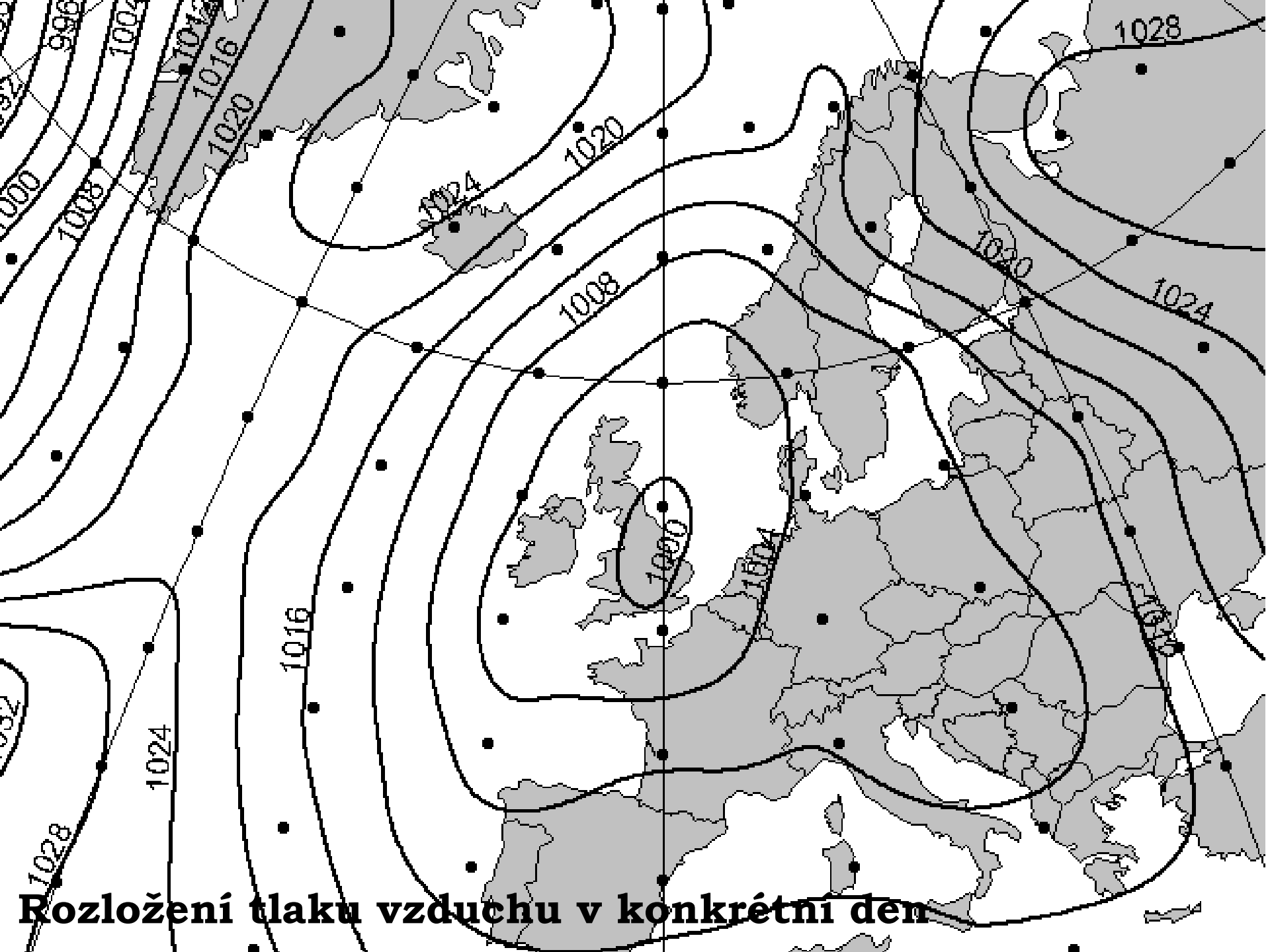


- **98 dnů s vichřicemi z 20. století**
- **Každý případ vichřice charakterizován 121 body s hodnotou popisující přízemní pole tlaku vzduchu**
- **Vstupní matice o rozměru 98 sloupců (proměnných) a 121 řádků (případů)**
- **Jakým způsobem popsat typické a společné rysy cirkulace ve dnech s vichřicemi?**





**121 bodů popisujících pole tlaku vzduchu**



**Rozložení tlaku vzduchu v konkrétní den**

Microsoft Excel - ZITR

Soubor Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda

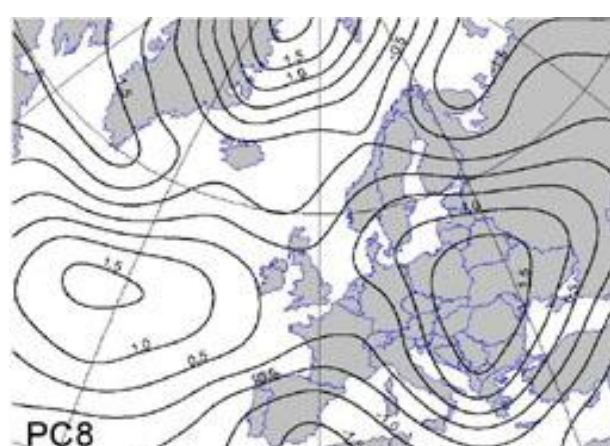
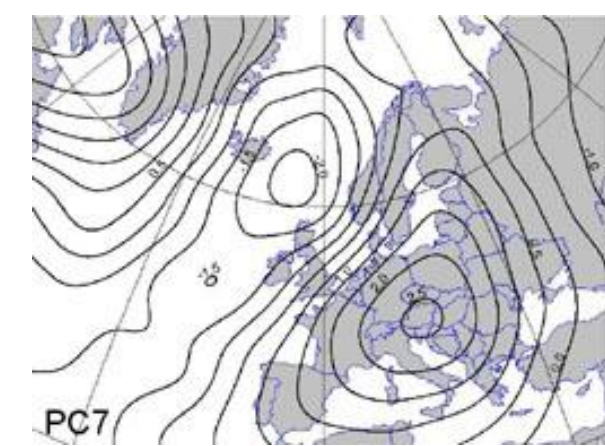
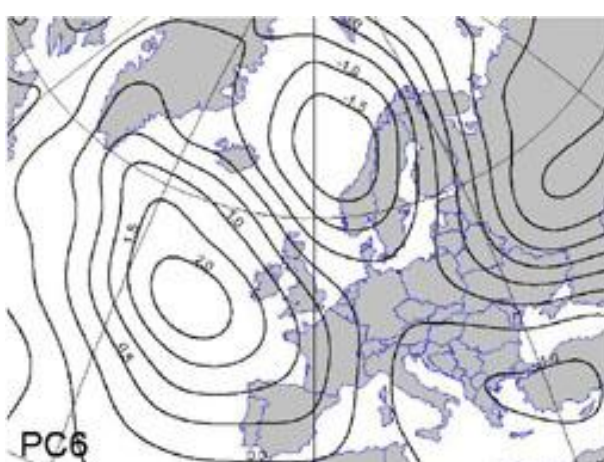
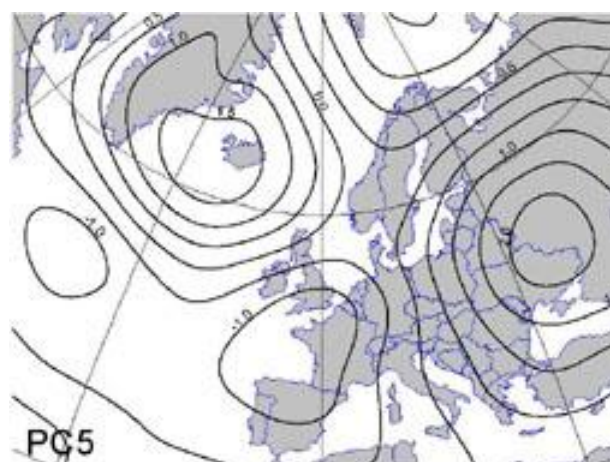
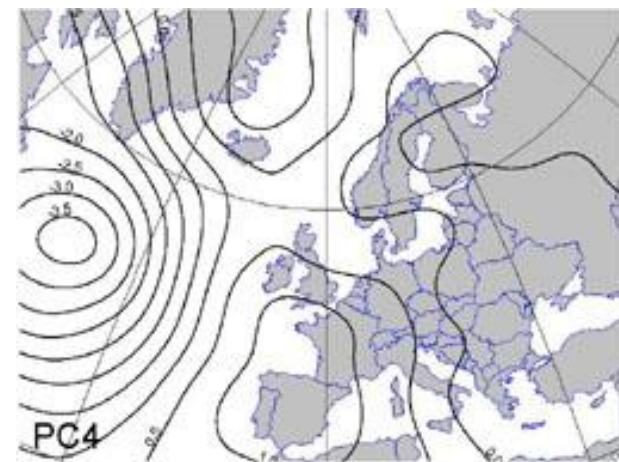
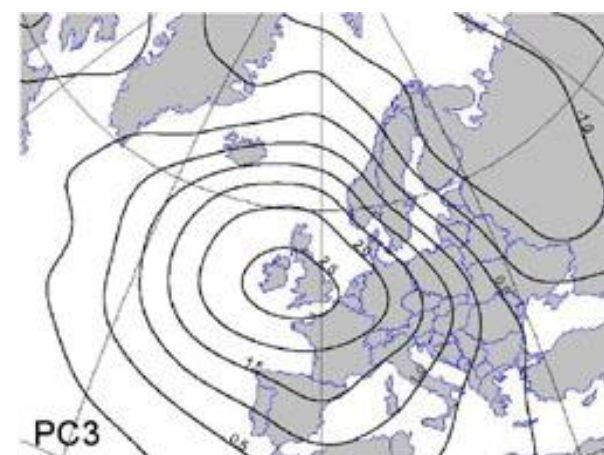
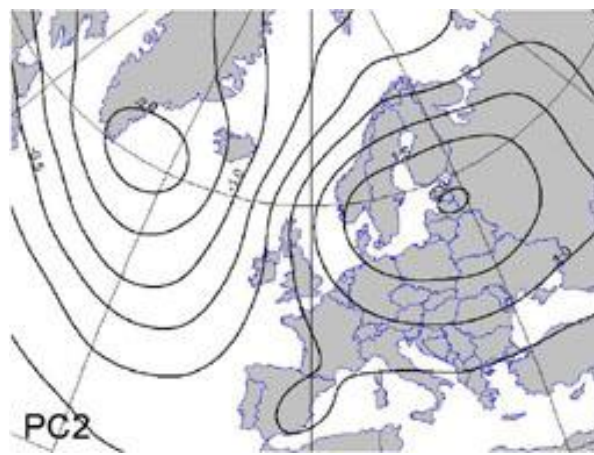
A1 = ZSZD

	A	B	C	D	E	F	G	H	I	J
1	ZSZD	2731900	6121900	21121900	2811901	1611902	31121904	17101907	27121907	721908
2	80/-50	1036,5	1017,71	1005,49	1044,8	1013,2	1002,83	1012,25	1015,9	1025,5
3	80/-40	1035,81	1017,36	1003,14	1044,03	1011,9	1004,01	1011,31	1017,06	1022,3
4	80/-30	1033,32	1015,11	1000,89	1040,21	1009,6	1007,73	1011,41	1018,8	1019,95
5	80/-20	1030,83	1012,87	998,64	1036,4	1007,3	1011,46	1011,51	1020,54	1017,6
6	80/-10	1029,13	1010,96	996,74	1032,71	1003,69	1011,49	1010,96	1020,11	1014,35
7	80/0	1027,43	1009,06	994,84	1029,03	1000,09	1011,52	1010,41	1019,68	1011,1
8	80/10	1026,14	1007,89	994,61	1026,15	997,29	1007,77	1010,33	1017,93	1006,65
9	80/20	1024,85	1006,72	994,38	1023,27	994,49	1004,02	1010,25	1016,18	1002,2
10	80/30	1024,04	1006,79	995,93	1021,68	994,44	997,52	1011,48	1014,71	998,55
11	80/40	1023,24	1006,86	997,48	1020,1	994,39	991,03	1012,71	1013,24	994,9
12	80/50	1022,54	1007,86	999,98	1021,76	997,59	990,62	1015,86	1012,78	995
13	75/-50	1030,38	1012,88	1002,72	1042,12	1014,36	1004,72	1005,11	1009,21	1021,95
14	75/-40	1028,49	1012,97	999,9	1041,57	1012,31	1000,04	1007,21	1010,96	1014,9
15	75/-30	1027,01	1011,68	997,15	1037,03	1009,4	1005,24	1008,16	1015,19	1008,7
16	75/-20	1025,53	1010,4	994,4	1032,5	1006,5	1010,45	1009,11	1019,42	1002,5
17	75/-10	1022,83	1007,34	990,35	1025,34	1003,85	1013,85	1008,28	1021,35	1005,3
18	75/0	1020,13	1004,28	986,3	1018,17	1001,2	1017,25	1007,45	1023,28	1008,1
19	75/10	1020,33	1002,28	985,22	1012,2	992,79	1015,53	1005,97	1019,63	1004,6
20	75/20	1020,53	1000,28	984,14	1006,23	984,39	1013,82	1004,5	1015,98	1001,1
21	75/30	1021,38	1001,89	989,24	1006,6	987,84	1009,37	1006,17	1013,06	992,9
22	75/40	1022,24	1003,5	994,34	1006,97	991,29	1004,93	1007,85	1010,14	984,7
23	75/50	1020,59	1006,68	998,41	1012,43	997,26	1000,95	1013,25	1009,18	987,2
24	70/-50	1025,02	1008,45	1001,32	1039,65	1015,93	1006,32	998,72	1003,31	1018,95
25	70/-40	1021,59	1009,34	998,3	1038,57	1013,45	995,35	1003,35	1006,32	1007,7
26	70/-30	1021,61	1009,46	994,45	1034,1	1009,57	1003	1005,13	1012,47	998,1
27	70/-20	1021,63	1009,58	990,6	1029,63	1005,7	1010,65	1006,91	1018,62	988,5
28	70/-10	1017,91	1004,89	984,11	1019,43	1004,95	1017,27	1006,1	1023,2	997,3

**Vstupní datová matice – řádky: dny s vichřicí,  
sloupce: body (z.š/z.d)**

	A	B	C	D	E	F	G	H	I	J	K	L
1	zs	zd	pc1	pc2	pc3	pc4	pc5	pc6	pc7	pc8	pc9	pc10
2	80	-50	0,49641	-1,26968	-0,97844	-1,20321	0,25347	0,34454	-0,02946	0,36648	-2,01765	-0,58551
3	80	-40	0,44259	-1,15376	-1,04104	-1,25714	0,22378	0,51242	-0,31990	0,49676	-2,12153	-0,50242
4	80	-30	0,45007	-0,96742	-1,10244	-1,19649	-0,04356	0,55165	-0,64280	0,83826	-1,84061	-0,30543
5	80	-20	0,49263	-0,76128	-1,16791	-1,13847	-0,35802	0,56849	-0,94568	1,23672	-1,42215	-0,19927
6	80	-10	0,33909	-0,64057	-1,21051	-0,91347	-0,65665	0,63094	-0,95714	1,35824	-0,98649	-0,27135
7	80	0	0,15609	-0,54253	-1,24405	-0,70223	-1,02297	0,69526	-0,98457	1,34961	-0,58355	-0,45432
8	80	10	-0,02666	-0,44998	-1,30244	-0,50670	-1,38087	0,56990	-0,83424	1,23338	-0,19049	-0,60767
9	80	20	-0,14818	-0,35432	-1,38750	-0,31766	-1,77164	0,35894	-0,63110	1,25735	0,36841	-0,81937
10	80	30	-0,16835	-0,23761	-1,39729	-0,32350	-2,05840	-0,01173	-0,36850	1,32622	0,80405	-1,01546
11	80	40	-0,14224	-0,16428	-1,40846	-0,36104	-2,33727	-0,36343	-0,08894	1,35228	1,18631	-1,27496
12	80	50	-0,11328	-0,09576	-1,14483	-0,49322	-2,41665	-0,82250	0,12016	1,15822	1,37977	-1,77888
13	75	-50	0,19626	-1,52734	-0,80925	-0,75774	0,56541	0,53084	0,86160	-0,31943	-1,53332	-0,47219
14	75	-40	0,21570	-1,37643	-0,92990	-0,97709	0,81336	0,55162	0,43141	0,23666	-1,47759	-0,19451
15	75	-30	0,11872	-1,20165	-0,91264	-1,14941	0,54995	0,59682	-0,17611	0,60603	-1,28434	0,12490
16	75	-20	0,04708	-0,94324	-0,88373	-1,23264	0,34134	0,70482	-0,87384	1,01795	-0,97470	0,52007
17	75	-10	-0,26389	-0,66883	-0,70424	-1,06808	-0,12449	0,90656	-1,06235	0,86620	-0,50917	0,28732
18	75	0	-0,64758	-0,36238	-0,56486	-0,70219	-0,57685	1,17435	-1,19422	0,51703	-0,24730	-0,04912
19	75	10	-0,81903	-0,10226	-0,74788	-0,33904	-1,14793	1,11159	-0,75533	0,12974	-0,03517	0,24812
20	75	20	-0,91858	0,08163	-0,96098	-0,04251	-1,70897	0,88998	-0,27249	-0,13483	0,26759	0,53922
21	75	30	-0,88591	0,11404	-1,13472	0,04166	-1,92237	0,22883	0,02557	-0,14744	0,74328	0,20313
22	75	40	-0,80525	0,12883	-1,29503	0,11665	-2,15667	-0,42896	0,30703	-0,02845	1,21537	-0,15874
23	75	50	-0,62454	0,14485	-1,14549	-0,02609	-2,18546	-1,06746	0,35265	0,11504	1,65076	-0,85658
24	70	-50	-0,08945	-1,75521	-0,59060	-0,25667	0,72134	0,57015	1,84058	-0,80573	-1,36616	-0,89145
25	70	-40	-0,10724	-1,61699	-0,75403	-0,68736	1,04015	0,33472	1,19871	-0,18184	-0,55179	-0,40478
26	70	-30	-0,34680	-1,40414	-0,70265	-1,07341	0,94410	0,36428	0,03641	0,10686	-0,13581	0,32515
27	70	-20	-0,47143	-1,15150	-0,62850	-1,50930	1,05786	0,47126	-1,12661	0,50122	0,51069	0,88931
28	70	-10	-1,06252	-0,68625	-0,19690	-1,22745	0,62181	0,70377	-1,51784	-0,13854	0,39285	0,63733

**Výstupní matice komponentních skóre – řádky:  
dny s vichřicí, sloupce: body (z.š/z.d)**

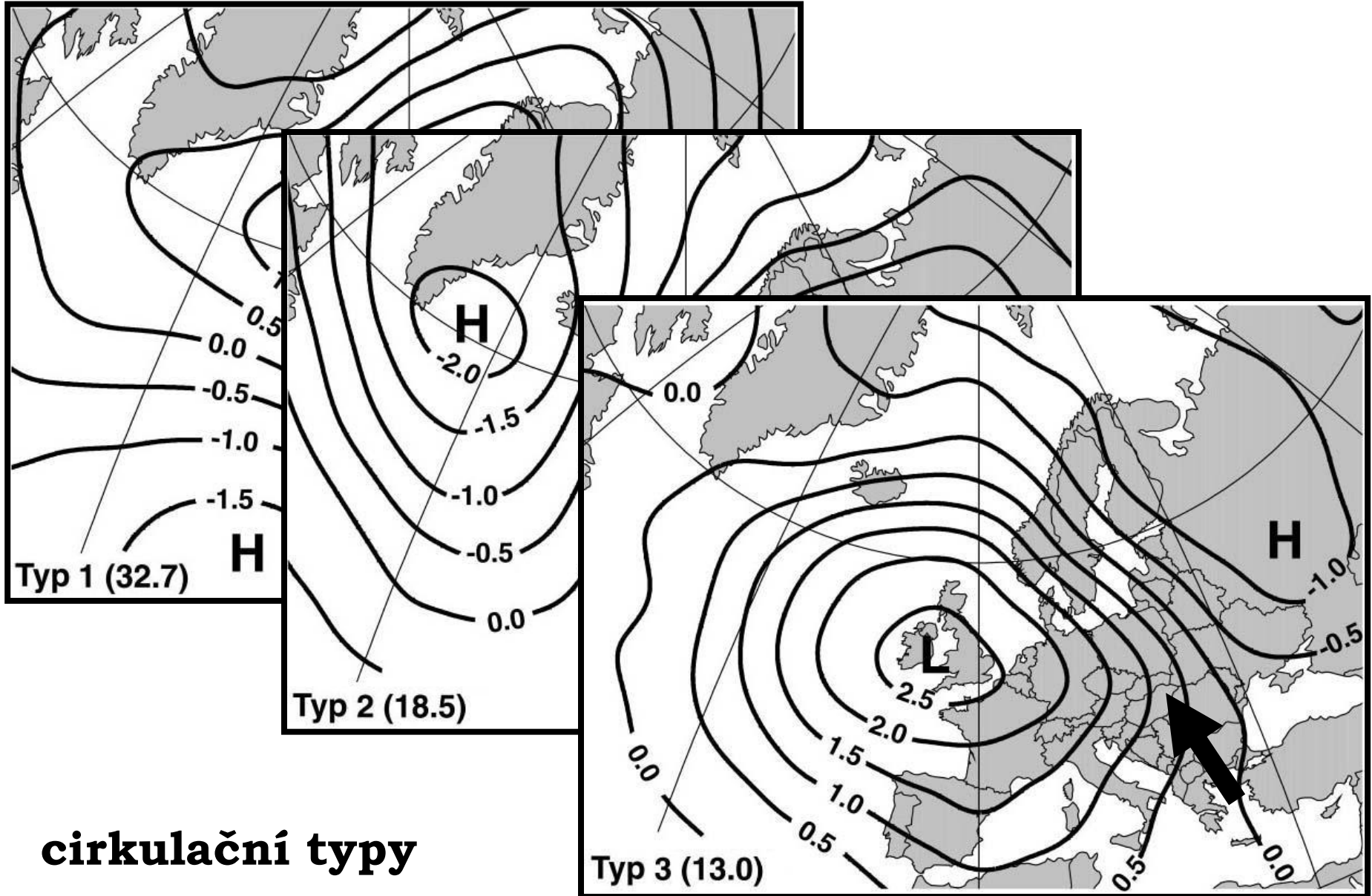


**Pole přizemního  
tlaku vzduchu  
hlavní  
komponenty**

A1		= Číslo hlavní komponenty				
	A	B	C	D	E	
1	Číslo hlavní komponenty	Vlastní čísla (eigenvalues)	% celkového rozptylu	Kumulované hodnoty vlastních čísel	Kumulovaná procenta celkového rozptylu	
2	1	31,19173	34,65748	31,19173	34,6575	
3	2	16,82658	18,69620	48,01831	53,3537	
4	3	11,08150	12,31278	59,09981	65,6665	
5	4	5,86098	6,51220	64,96079	72,1787	
6	5	4,96513	5,51681	69,92593	77,6955	
7	6	3,15952	3,51058	73,08545	81,2061	
8	7	2,96912	3,29902	76,05456	84,5051	
9	8	2,13167	2,36853	78,18624	86,8736	
10	9	1,79902	1,99891	79,98525	88,8725	
11	10	1,37242	1,52491	81,35768	90,3974	
12	11	1,09692	1,21880	82,45460	91,6162	
13	12	0,99723	1,10804	83,45183	92,7243	
14	13	0,80865	0,89850	84,26048	93,6228	
15	14	0,64909	0,72121	84,90957	94,3440	
16	15	0,57632	0,64036	85,48589	94,9843	
17	16	0,53258	0,59175	86,01847	95,5761	
18	17	0,35990	0,39989	86,37837	95,9760	
19	18	0,33623	0,37359	86,71460	96,3496	
20	19	0,30663	0,34070	87,02123	96,6903	
21	20	0,25210	0,28011	87,27333	96,9704	
22	21	0,23443	0,26048	87,50776	97,2308	
23	22	0,22489	0,24987	87,73265	97,4807	
24	23	0,19590	0,21767	87,92855	97,6984	
25	24	0,18218	0,20243	88,11073	97,8808	
26	25	0,16288	0,18098	88,27361	98,0818	
27	26	0,14513	0,15912	88,41673	98,2408	

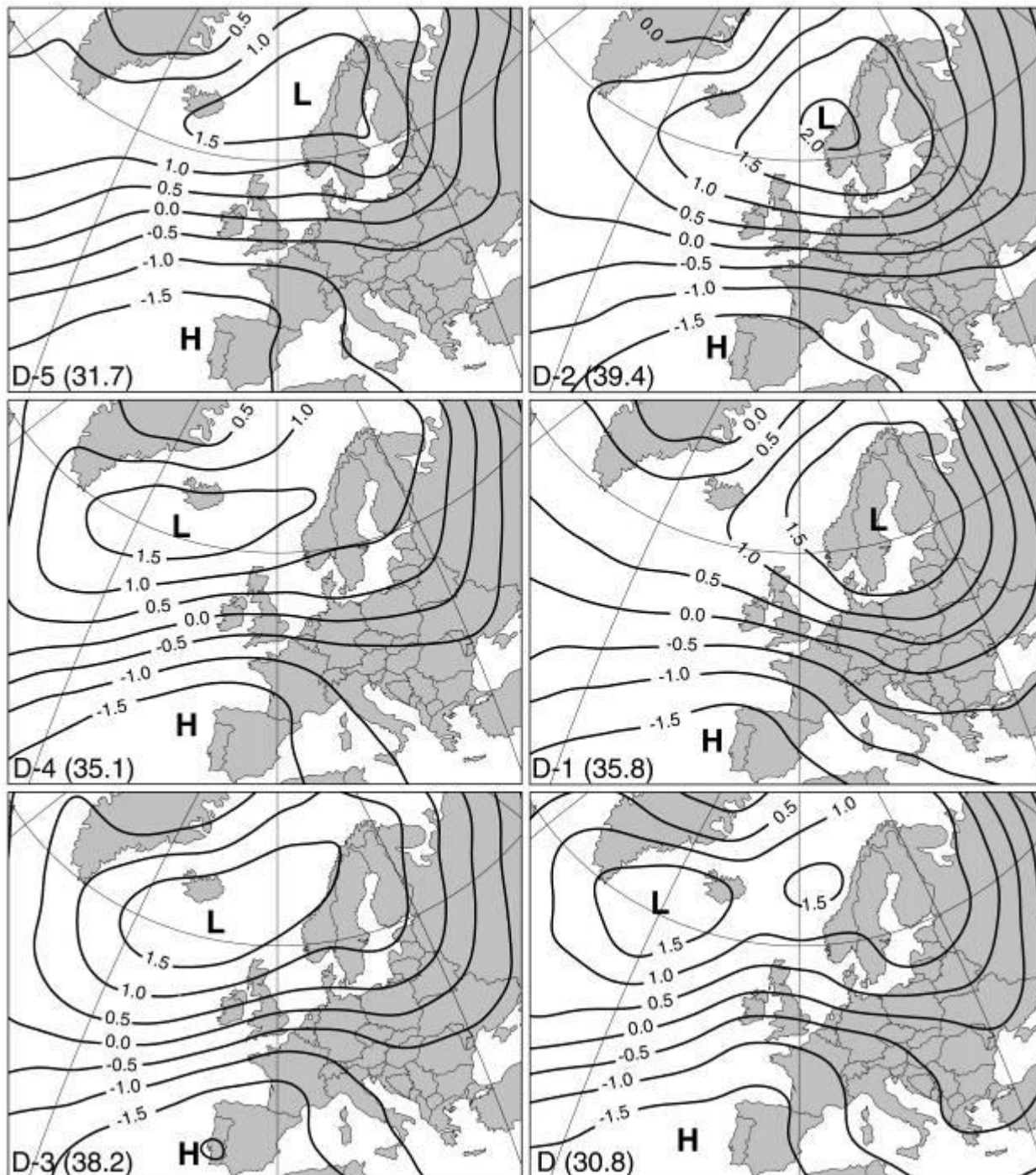
**Vypočtené hlavní komponenty, jejich vlastní čísla a procento vysvětlující rozptyl původních dat**

# ... a jejich interpretace jako hlavních cirkulačních typů

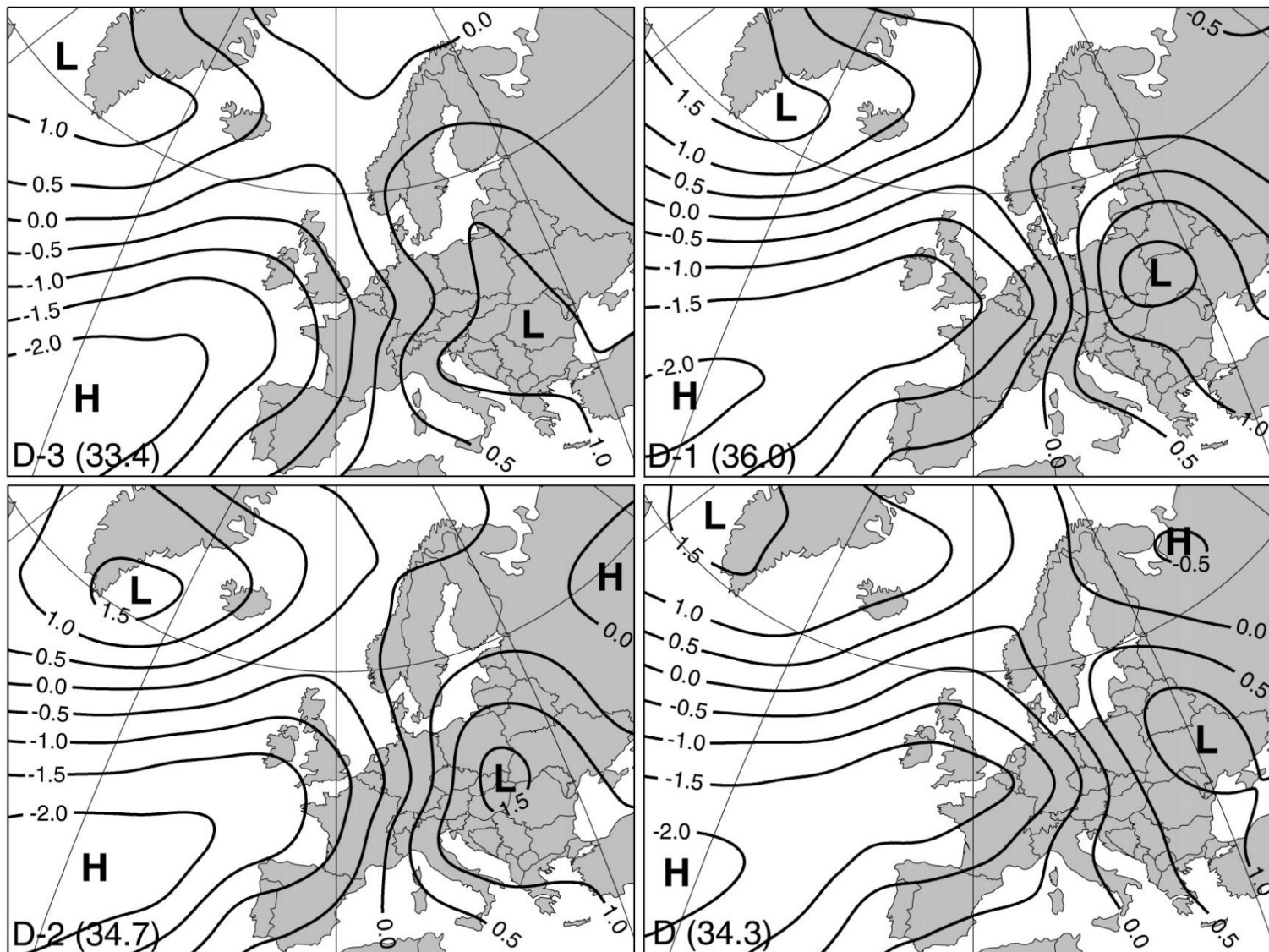


# Meteorologické příčiny povodní

Komponentní skóre první hlavní komponenty pole přízemního tlaku vzduchu ve dnech D-5 až D vypočítané metodou hlavních komponent pro 37 povodní (1881-2000) zimního synoptického typu na řece Vltavě v Praze (v závorce objasněný rozptyl v %) (Brázdil et al., 2005)







Komponentní skóre první hlavní komponenty pole přízemního tlaku vzduchu ve dnech D-3 až D vypočítané metodou hlavních komponent pro 40 povodní (1896-2000) letního synoptického typu na řece Odře v Bohumíně (v závorce objasněný rozptyl v %)(Brázdil et al., 2005)

## Doporučená literatura:

Schönwiese, C.-D. (1985): Praktische Statistik für Meteorologen und Geowissenschaftler. Gebrüder Borntraeger, Berlin, Stuttgart, 231 s.

Brázdil, R. (1991): Kolísání vybraných meteorologických prvků ve střední Evropě v období přístrojových pozorování. Národní klimatický program ČSFR, sv. 2, Praha, 56 s.

Storch, H. von, Zwiers, F. W. (1999): Statistical Analysis in Climate Research. Cambridge University Press, Cambridge, 484 s.

Pišoft, P., Kalvová, J. (2005): Wavelet analýza v meteorologii: teorie a přehled dosavadních výsledků. Meteorologické zprávy, 58, č. 1, s. 1-6.

Brázdil, R., Bělínová, M., Dobrovolný, P., Mikšovský, J., Pišoft, P., Řezníčková, L., Štěpánek, P., Valášek, H., Zahradníček, P. (2012): Temperature and Precipitation Fluctuations in the Czech Lands During the Instrumental Period. Masaryk University, Brno, 236 s.

Mudelsee, M. (2014): Climate Time Series Analysis. Classical Statistical and Bootstrap Methods. Springer, Cham, Heidelberg, New York, Dordrecht, London, 454 s.