



# Kvantifikovanie kauzálnych vzťahov v distribuovanom systéme

Bc. Anton Vančo  
vedúci: doc. Ing. Igor Farkaš, PhD.

# obsah

- kauzalita v komplexných sústavách
- formy analýzy časových radov
- metóda Grangerovej kauzality
  - jednorozmerné dáta (univariate)
  - viacrozmerné dáta (multivariate)
  - autonómia a emergencia
- ukážka analýzy – GCCA toolbox

# kauzalita

- vplyv jedného prvku systému na iný
- (ne)činnosť jedného elementu priamo podmieňuje činnosť druhého
- *intra-level kauzalita*
- *inter-level kauzalita* - top-down, bottom-up

# formy analýzy

- identifikácia *orientovanej* funkčnej konektivity
- analýza fMRI, EEG, MEG údajov
- 3 perspektívy:
  - efektívna konektivita
  - funkčná konektivita
  - kauzálna konektivita

# kauzálna konektivita

- nutne nemodeluje fyziologické vlastnosti
- pokus o identifikáciu *dynamických vzorov* a vzt'ahov
- 3 dôvody:
  - *same brain, different task*
  - *don't tell me everything ... just the important stuff*
  - *they keep messing up my neural connections!*

# meranie zložitosti

- kvantifikovanie správania sa distribuovaného systému
- 3 miery:
  - *neurálna komplexnosť* – segregácia vs. integrácia
  - *integrácia informácií* – podobná NK, sleduje orientované kauzálne interakcie – *effective information*
  - *kauzálna hustota* – zachytáva kauzálne vzťahy naprieč celým systémom

# Grangerova kauzalita

- premenná  $Y$  je kauzálna (*Granger-causes*) na premennú  $X$ , ak sa predpoveď hodnôt  $X$  (pomocou jej minulých hodnôt) zlepší zakomponovaním minulých hodnôt  $Y$ .
- zlepšenie = štatisticky významné informácie v  $Y$
- autoregresné modely – OLS, Yule-Walker

- pôvodný model

$$X_t = A \cdot \mathbf{X}_{t-1}^{(m)} + \varepsilon_t$$

- rozšírený model

$$X_t = A' \cdot (\mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Y}_{t-1}^{(m)}) + \varepsilon'_t$$

- zlepšenie, ak rozptyl  $\varepsilon'$  je významne menší ako  $\varepsilon$

$$\mathcal{F}_{Y \rightarrow X} = \ln \frac{\text{var}(\varepsilon_t)}{\text{var}(\varepsilon'_t)}$$

# určenie stupňa AR modelu

- $m$  – počet minulých meraní premennej zahrnutých v regresii
- *Akaike information criterion (AIK)*
  - $AIK = 2k - 2 \ln(L)$ 
    - $k$  – počet parametrov AR modelu
    - $L$  – maximalizovaná likelihood funkcia AR modelu
- *Bayesian information criterion (BIK)*
  - $BIK = -2 \ln(L) + k \ln(n)$
- zvolený je model s **najnižším AIC/BIC**



# Grangerova kauzalita

- *podmienená kauzalita* – prítomnosť tretej premennej

$$X_t = A \cdot \left( \mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Z}_{t-1}^{(m)} \right) + \varepsilon_t$$

$$X_t = A' \cdot \left( \mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Y}_{t-1}^{(m)} \oplus \mathbf{Z}_{t-1}^{(m)} \right) + \varepsilon'_t$$

- opäť porovnanie rozptylu – **Geweke (1982)**

$$\mathcal{F}_{Y \rightarrow X|Z} = \ln \frac{\text{var}(\varepsilon_t)}{\text{var}(\varepsilon'_t)}$$

# Grangerova kauzalita

- viacrozmerne dáta

$$X_t = A \cdot X_{t-1}^{(m)} + \varepsilon_t$$

$$\mathcal{F}_{Y \rightarrow X|Z} = \ln \left\{ \frac{\text{tr}[\sum \varepsilon_t]}{\text{tr}[\sum \varepsilon'_t]} \right\}$$

- stopa (trace) kovariančnej matice (Ladroue et al.)

$$\mathcal{F}_{Y \rightarrow X|Z} = \ln \left\{ \frac{|\sum \varepsilon_t|}{|\sum \varepsilon'_t|} \right\}$$

- *determinant* kovariančnej matice (Geweke)
- *objem*, v ktorom leží  $\varepsilon$

# kauzálna hustota

- miera kauzálne signifikantných vzťahov v systéme
- *normalizovaná* – hodnoty v  $[0, 1]$

$$C_d(X) = \frac{\alpha}{N(N-1)}$$

- *nenormalizovaná (váhovaná)* – hodnoty v celom  $R$

$$C_{dw}(X) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N \mathcal{F}_{j \rightarrow i}$$

- vysoké hodnoty znamenajú globálnu koordinovanosť prvkov a zároveň ich dynamickú odlišnosť

# spektrálna kauzalita

- kauzálny vplyv frekvencií

$$\mathcal{F}_{Y \rightarrow X}(\omega) = \ln \frac{S_{XX}(\omega)}{\tilde{H}_{XX}(\omega) * \text{var}(\varepsilon'_t) * \tilde{H}_{XX}^*(\omega)}$$

- $\omega$  – frekvencia
  - $S$  – autospektrum  $X_t$
  - $\tilde{H}_{XX}(\omega)$  - transferová matica
  - $\tilde{H}_{XX}^*(\omega)$  - complex conjugate
- umožňuje odhaliť **sprostredkované** kauzálne vzťahy

# G-autonómia

- nezávislosť premennej od iných premenných v systéme

$$X_t = A \cdot \mathbf{Z}_{t-1}^{(m)} + \varepsilon_t$$

$$X_t = A' \cdot (\mathbf{X}_{t-1}^{(m)} \oplus \mathbf{Z}_{t-1}^{(m)}) + \varepsilon'_t$$

- $\mathbf{Z}$  je množina externých premenných

$$\mathcal{A}_{X|Z} = \ln \left\{ \frac{|\sum \varepsilon_t|}{|\sum \varepsilon'_t|} \right\}$$

# G-emergencia

- *makro-premenná **M** je emergentá z mikro-premennýc **m** ( $m_1, m_2, \dots m_n$ ) ak je zároveň*
- ***autonómna** od svojich mikro-zložiek*
  - *minulé hodnoty **M** vedia predpovedať jej budúce lepšie, ako predpovede len na základe hodnôt **m***
- ***závislá** od svojich mikro-zložiek*
  - *hodnoty **m** zlepšujú predpovede budúcich hodnôt **M**, oproti predpovediam len na základe minulých hodnôt **M***

# G-emergencia

- miera *kombinujúca* G-kauzalitu a G-autonómiu

$$ge_{M|m} = \mathcal{A}_{M|m} \left( \frac{1}{N} \sum_{i=1}^N \mathcal{F}_{m_i \rightarrow M} \right)$$

- G-emergencia bude nulová ak
  - **M** je nezávislé od zložiek **m**
  - **M** sa dá úplne predpovedať len pomocou **m**

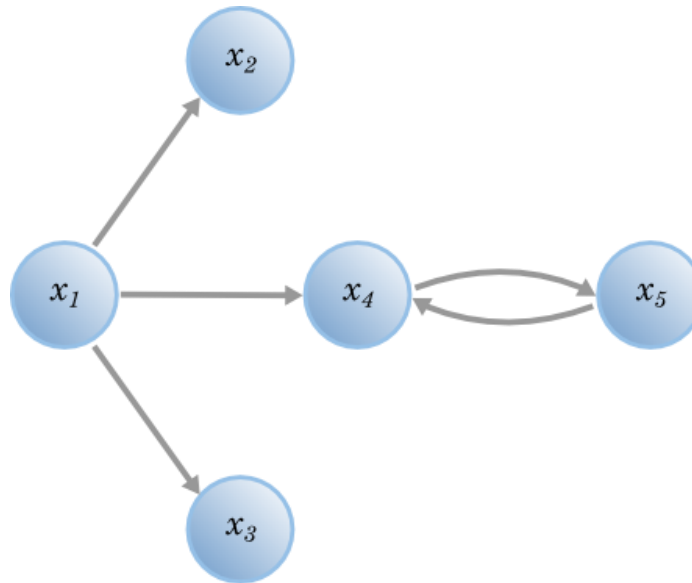
# d'alsie pojmy

- *causal flow* – rozdiel medzi počtom vstupných a výstupných interakcií na danom prvku
  - *causal source* – zdroj, vysoký *flow*
  - *causal sink* – záporný *flow*
- *causal hub* – prvok s veľkou kauzálnou hustotou

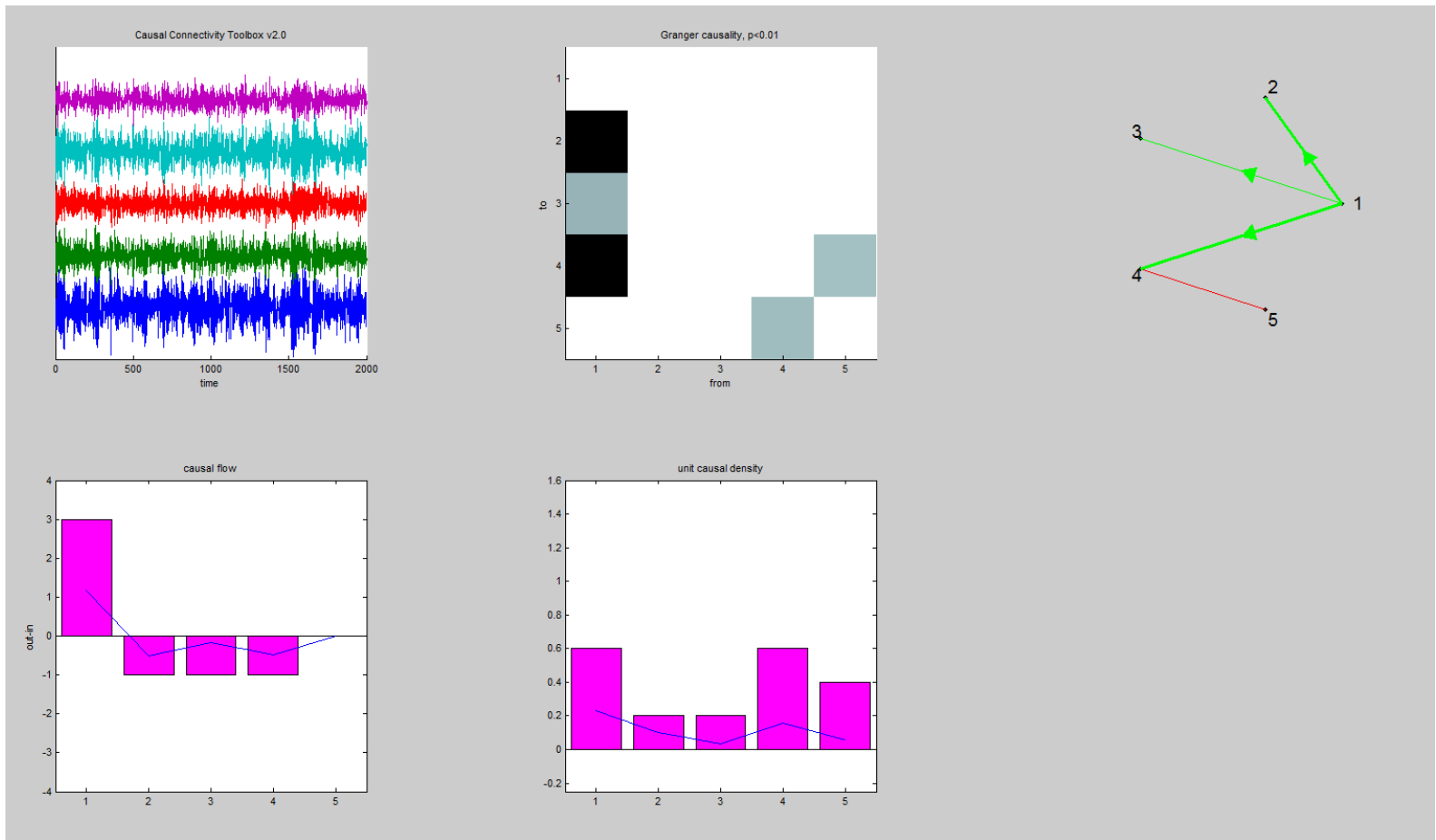


# ukážka analýzy

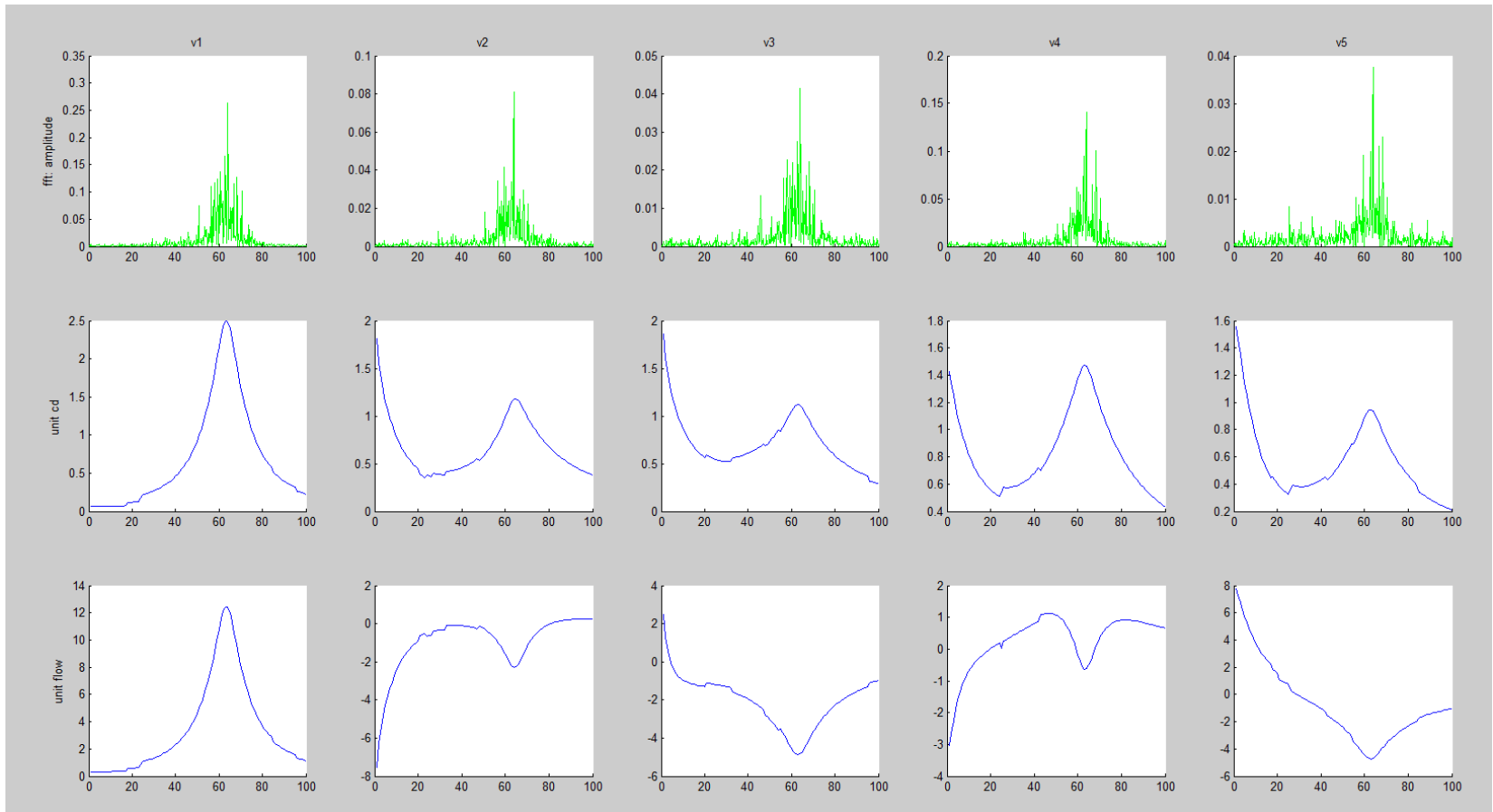
- 5 premenných (**Baccala & Sameshima, 2001**)
- $x_1(n) = 0.95\sqrt{2}x_1(n-1) - 0.9025x_1(n-2) + w_1(n)$   
 $x_2(n) = 0.5x_1(n-2) + w_2(n)$   
 $x_3(n) = -0.4x_1(n-3) + w_3(n)$   
 $x_4(n) = -0.5x_1(n-2) + 0.25\sqrt{2}x_4(n-1) + 0.25\sqrt{2}x_5(n-1) + w_4(n)$   
 $x_5(n) = -0.25\sqrt{2}x_4(n-1) + 0.25\sqrt{2}x_5(n-1) + w_5(n)$
- graf



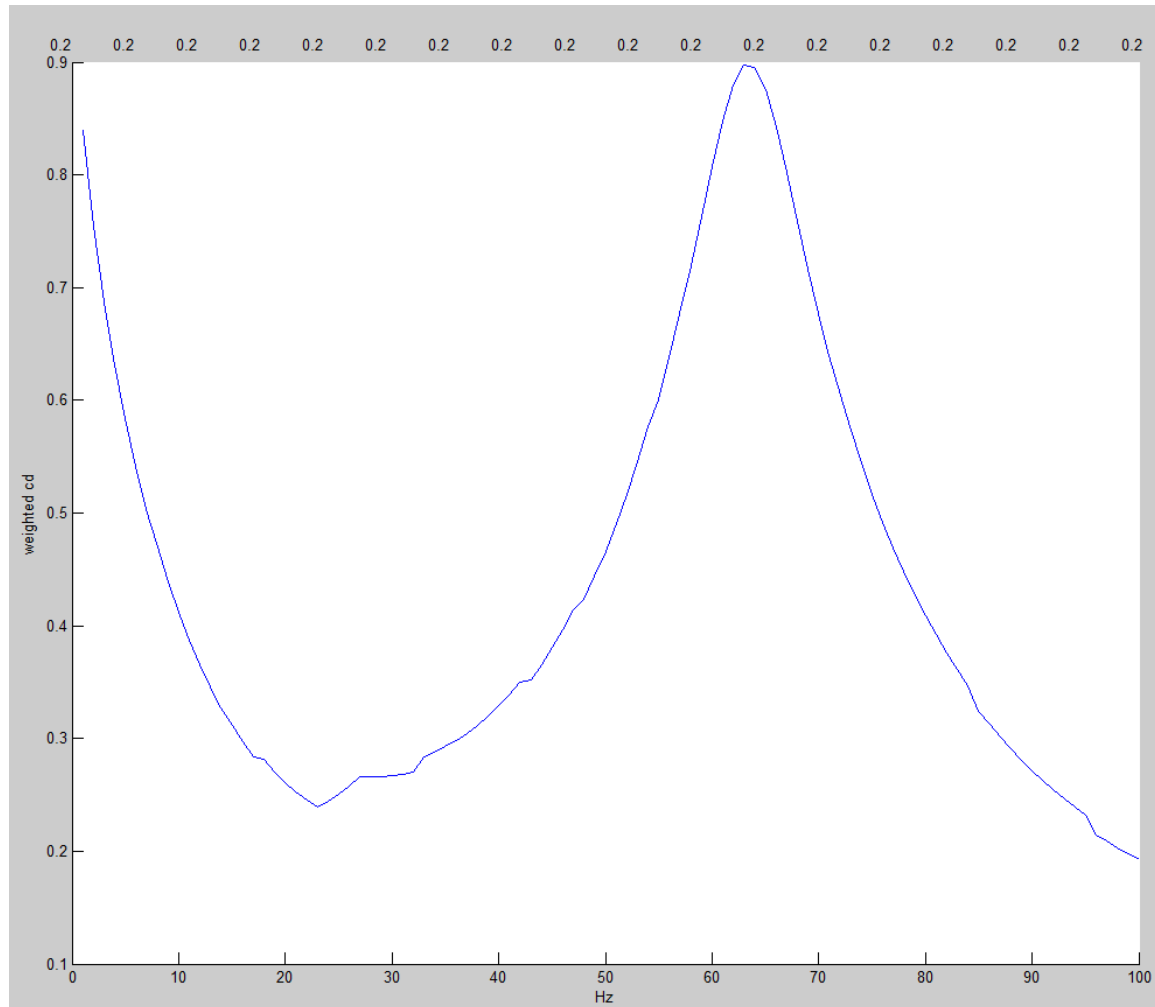
# time-domain analýza



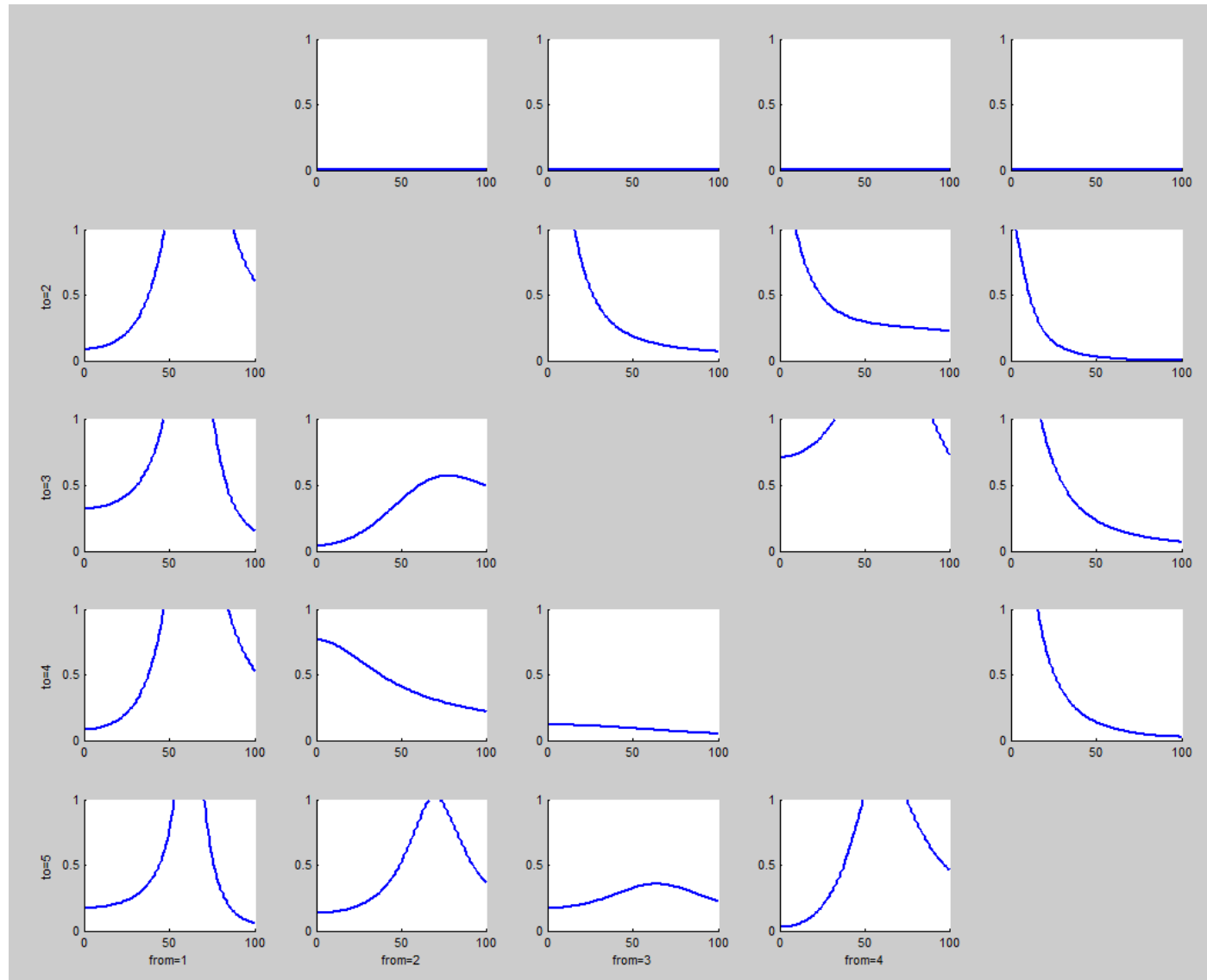
# spektrálna analýza I.



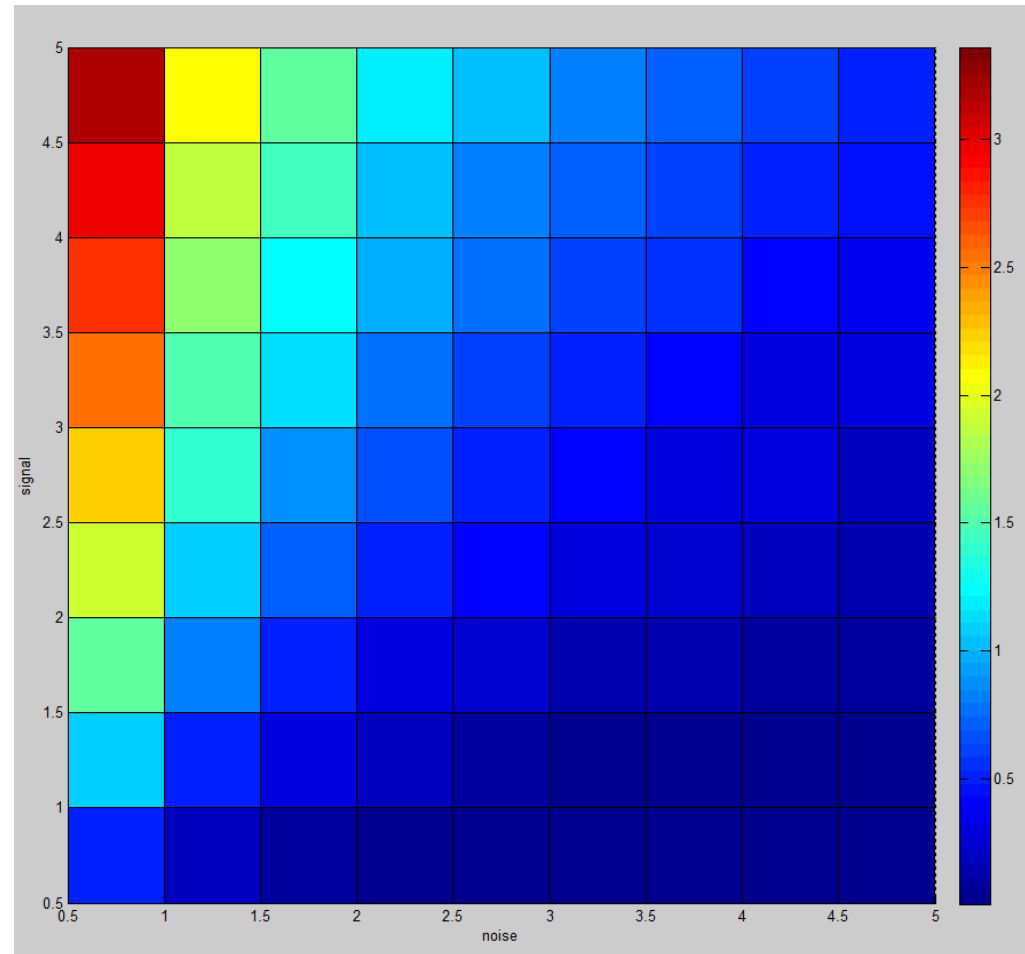
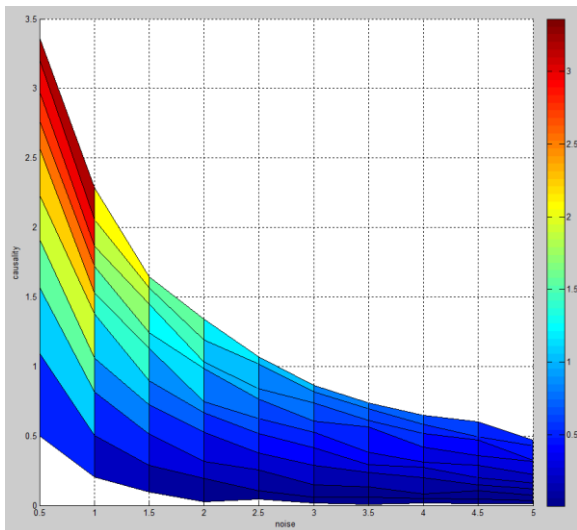
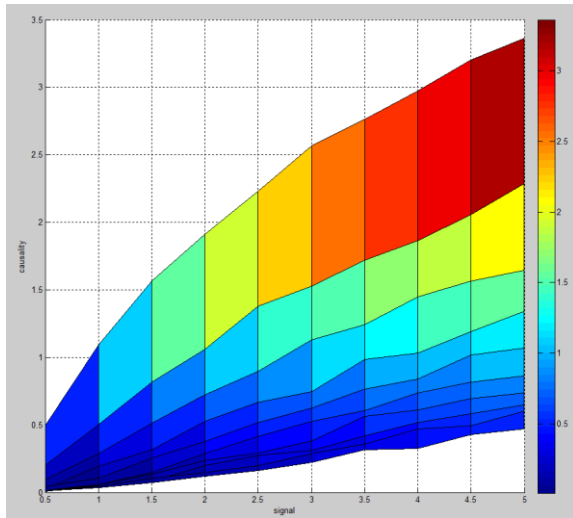
# spektrálna analýza II.



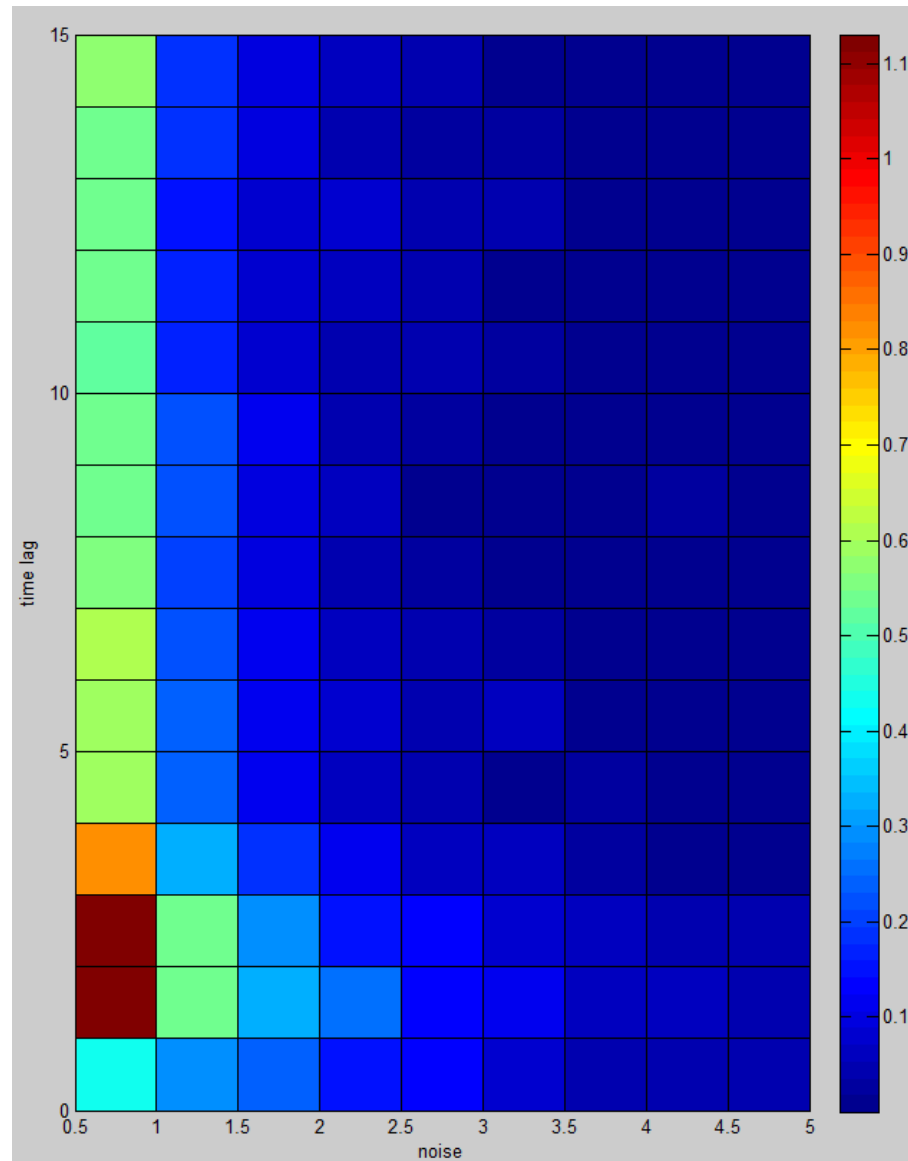
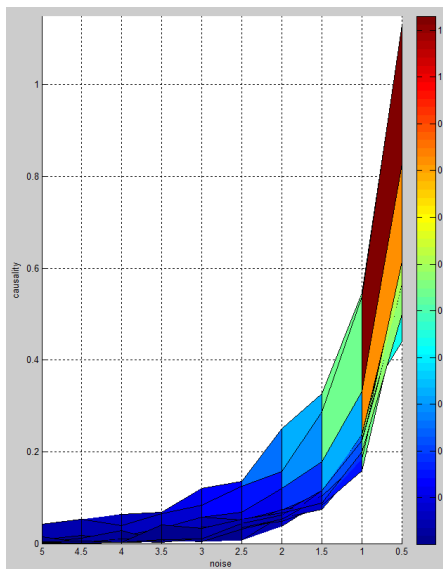
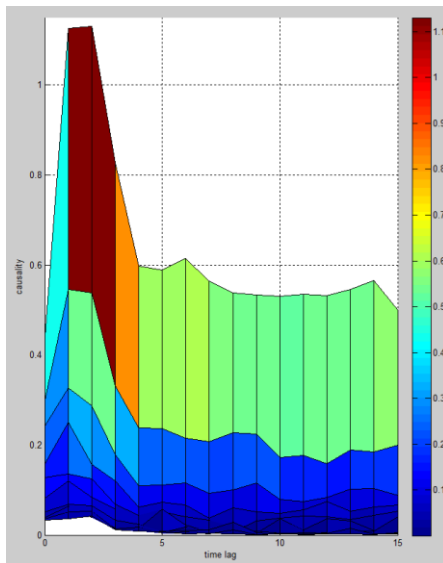
# spektrálna analýza III.



# závislost' kauzality – signál, šum



# závislost' kauzality – šum, čas





*Ďakujem za pozornosť*