

# Analýza funkcionálních dat v praxi

Jan Kolářček

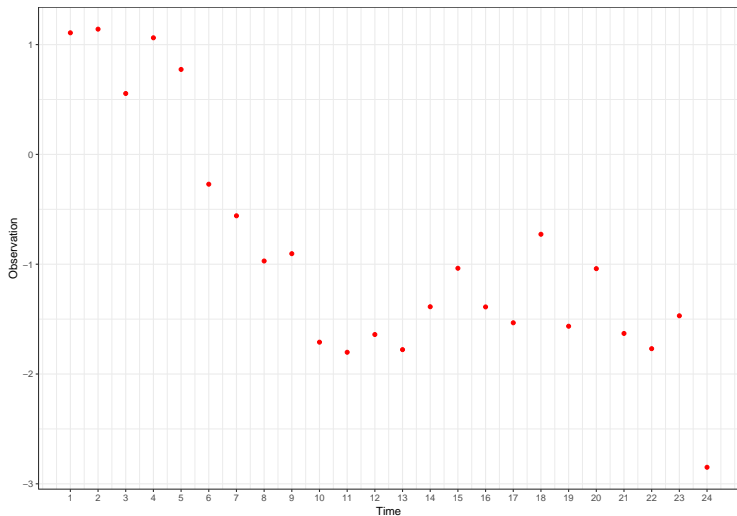
Ústav matematiky a statistiky, Masarykova Univerzita, Brno

Brno 2018

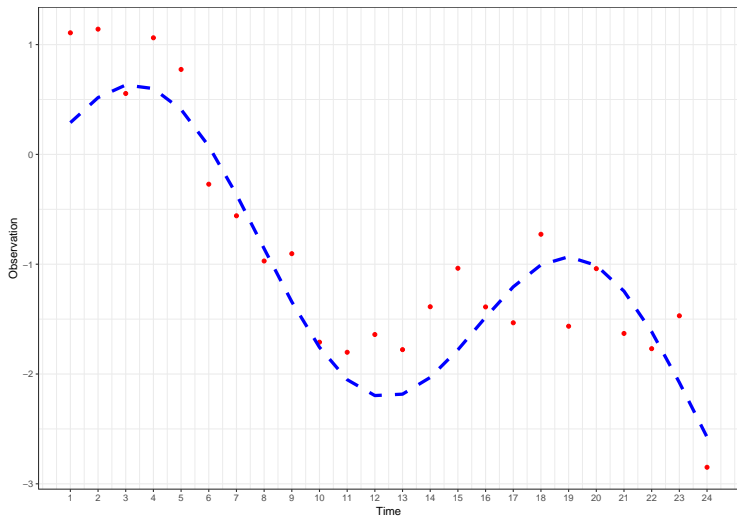
Analýza funkcionálních dat – FDA

Functional Data Analysis

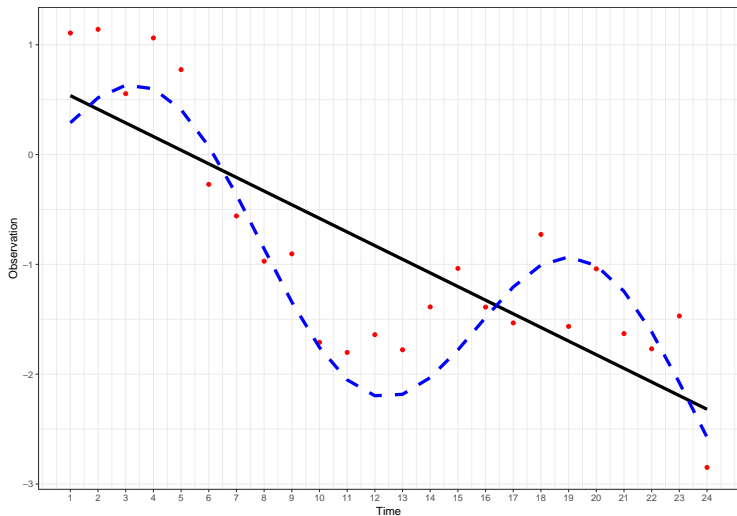
Proč **funkcionální** data a co to znamená?



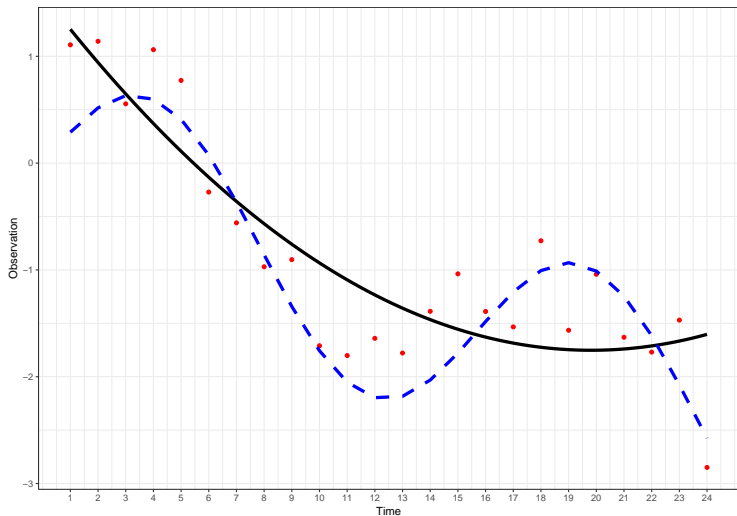
Data



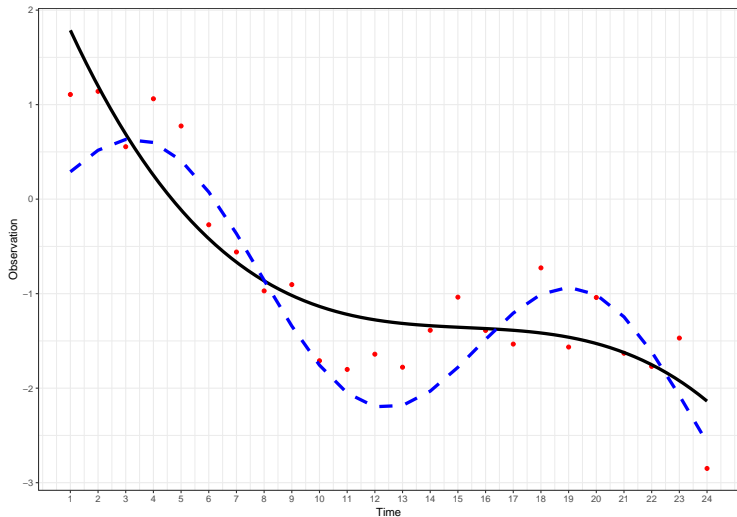
Data



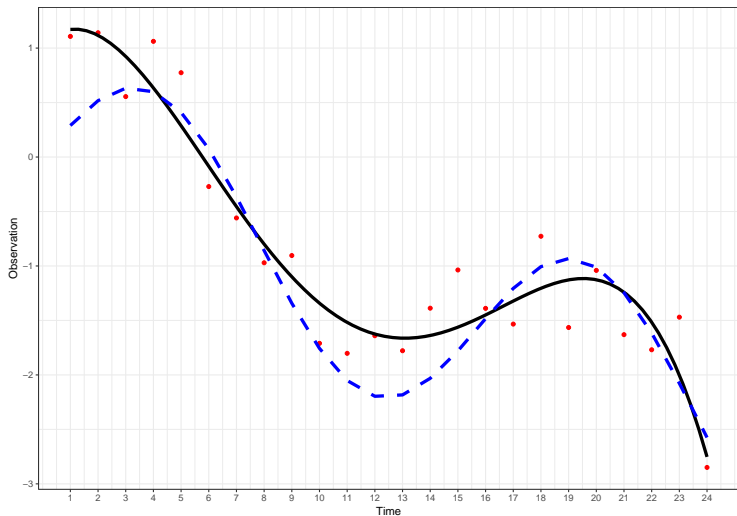
Lineární model – přímka



Lineární model – parabola

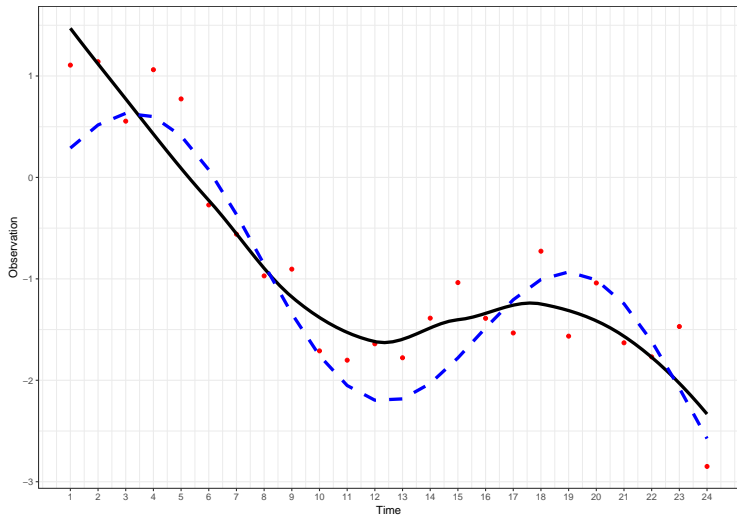


Lineární model – polynom st. 3

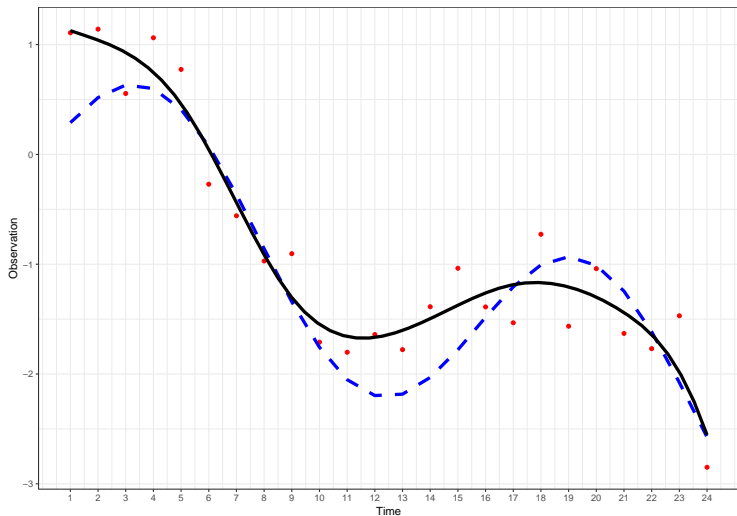


Lineární model – polynom st. 4





Neparametrický model – vyhlazování



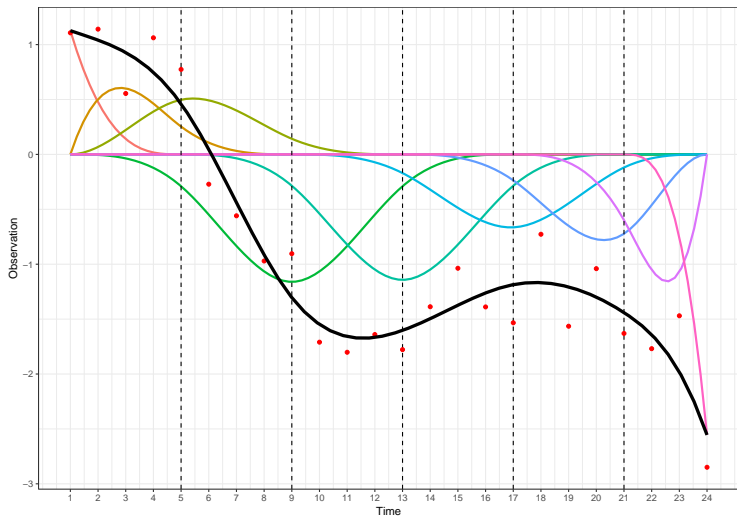
FDA - odhad

# B - splines



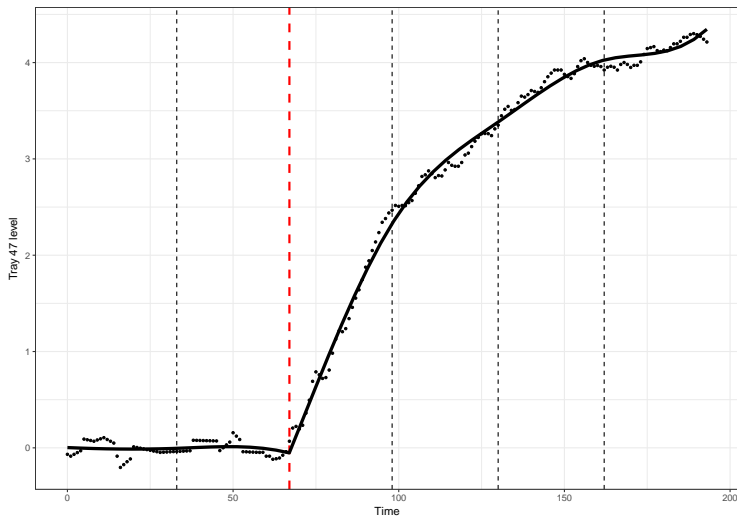
B - splines

# B - splines



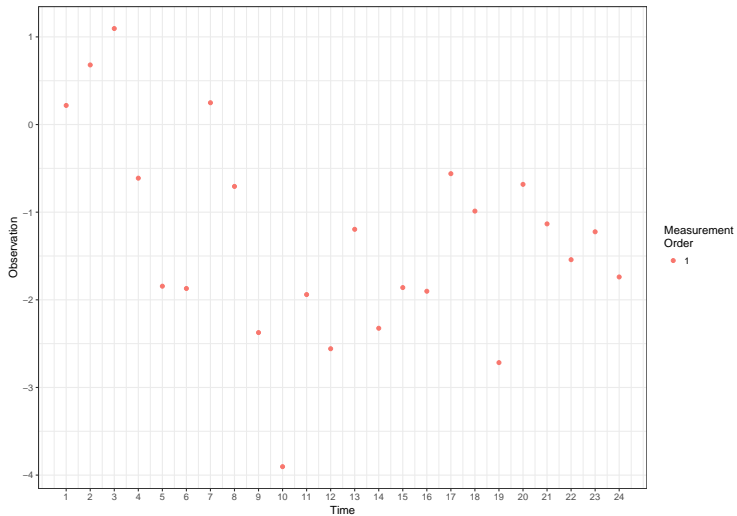
Konstrukce

## B - splines

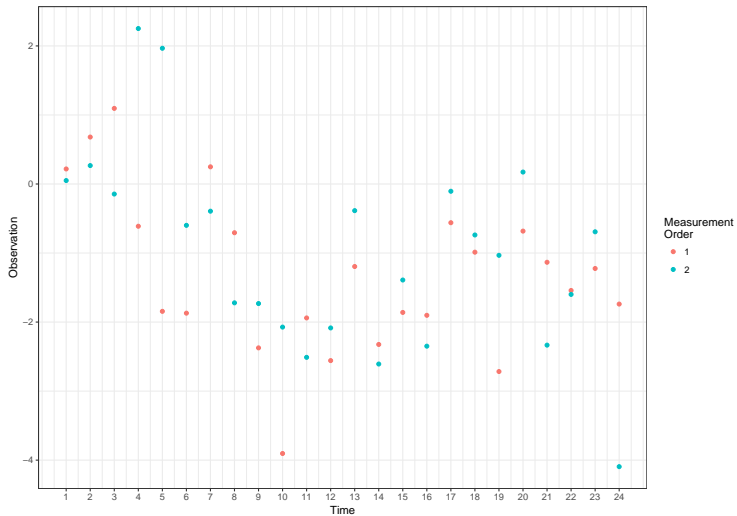


Vícenásobné uzly

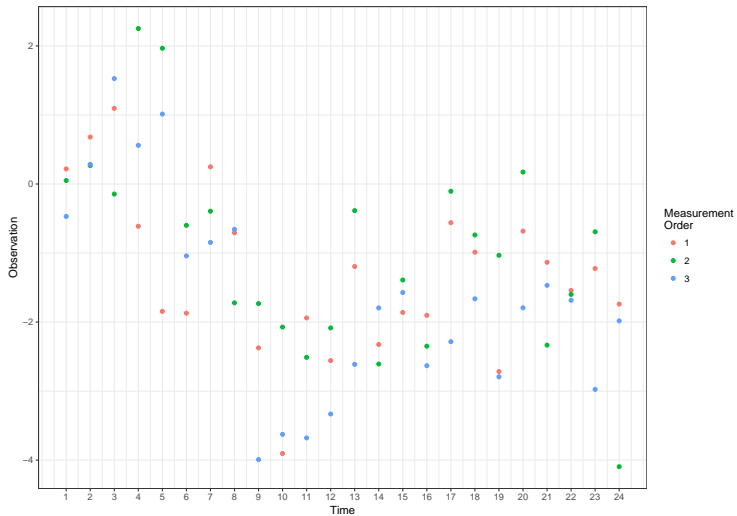
# Opakovaná měření



# Opakovaná měření

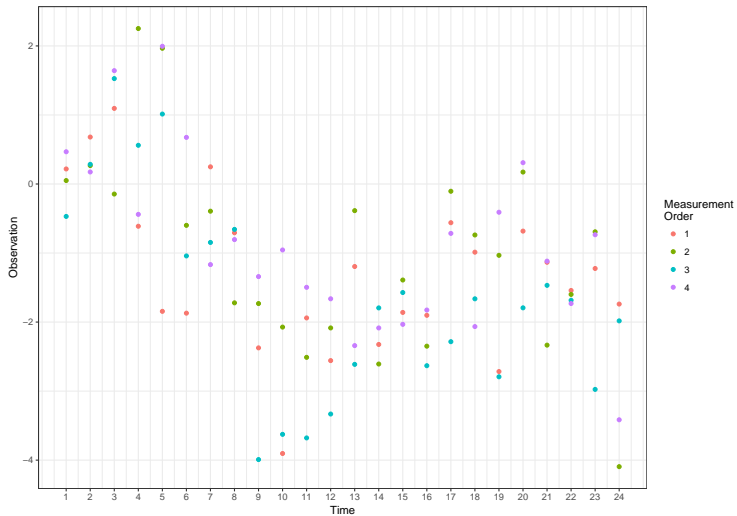


# Opakovaná měření

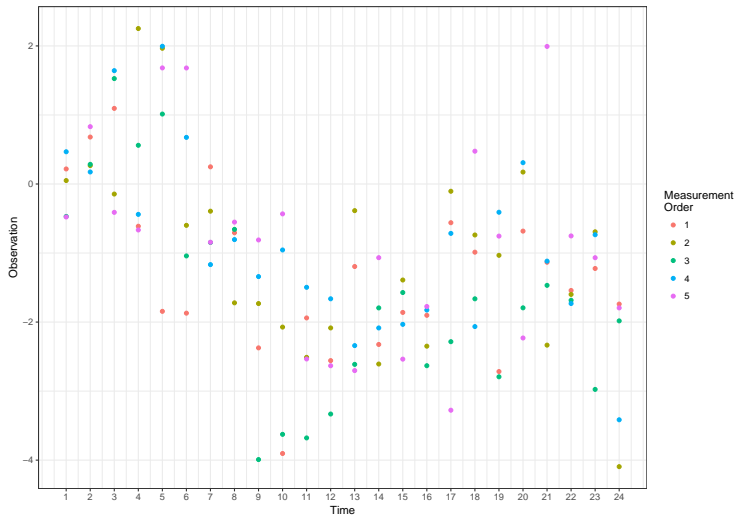




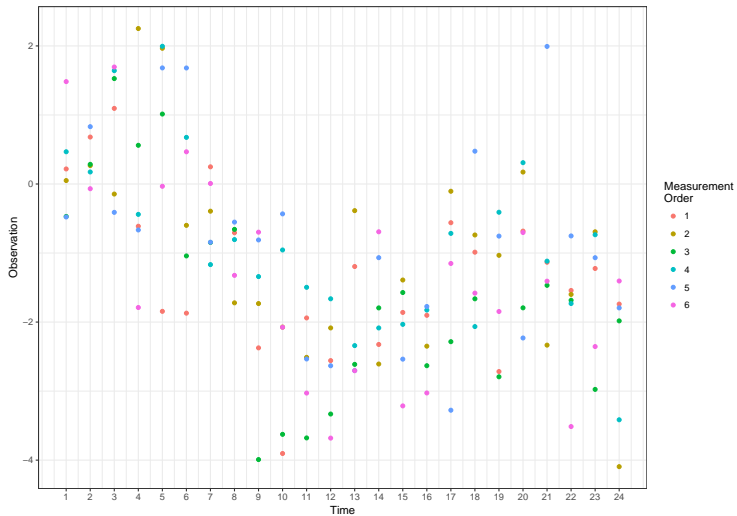
# Opakovaná měření



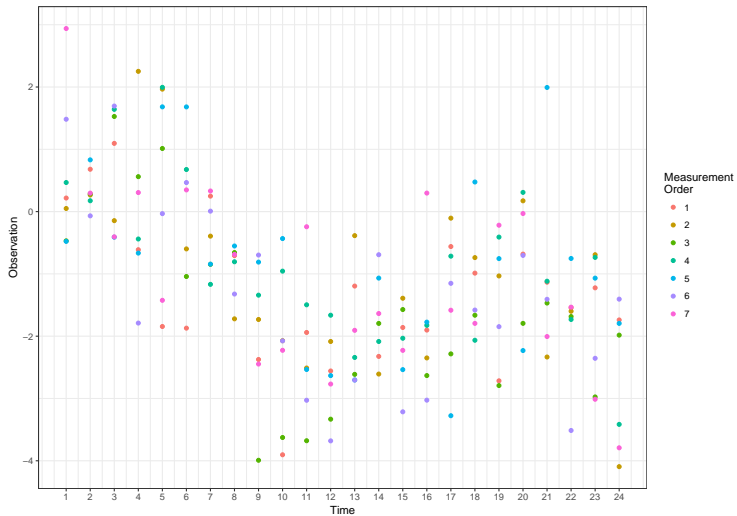
# Opakovaná měření



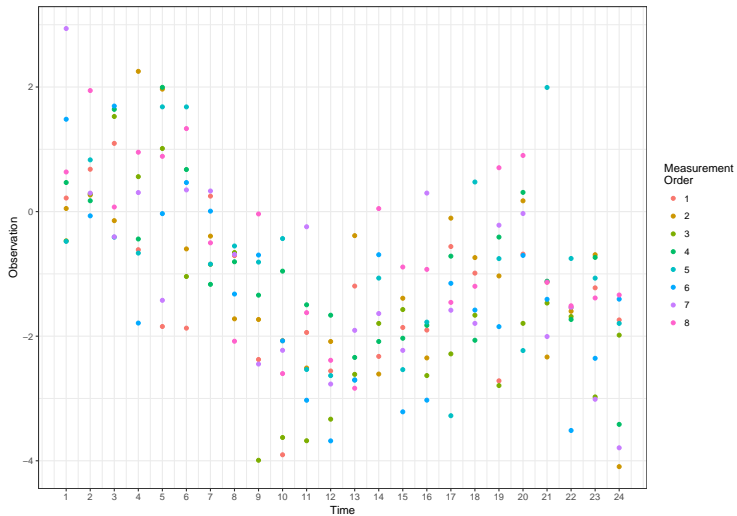
# Opakovaná měření



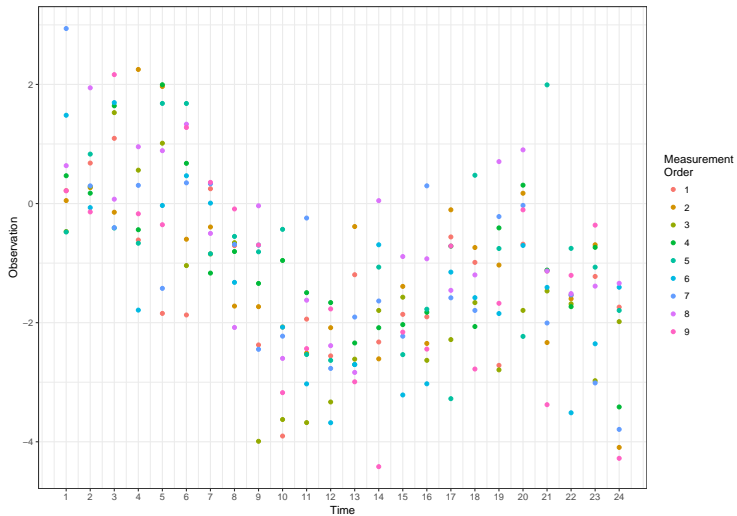
# Opakovaná měření



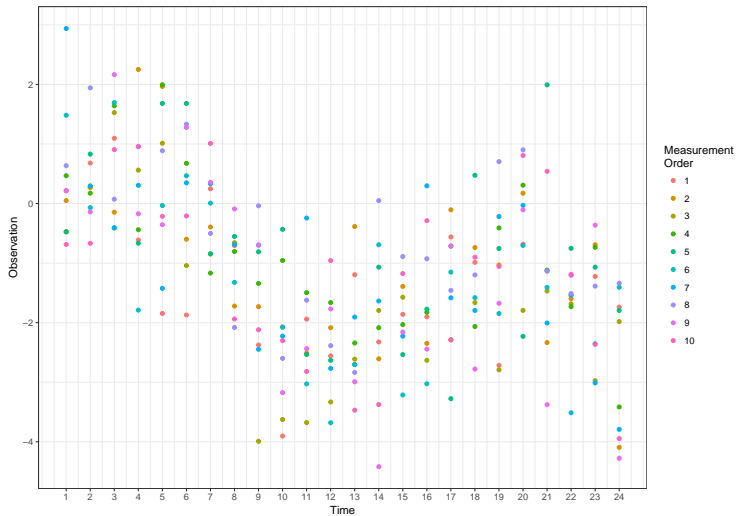
# Opakovaná měření



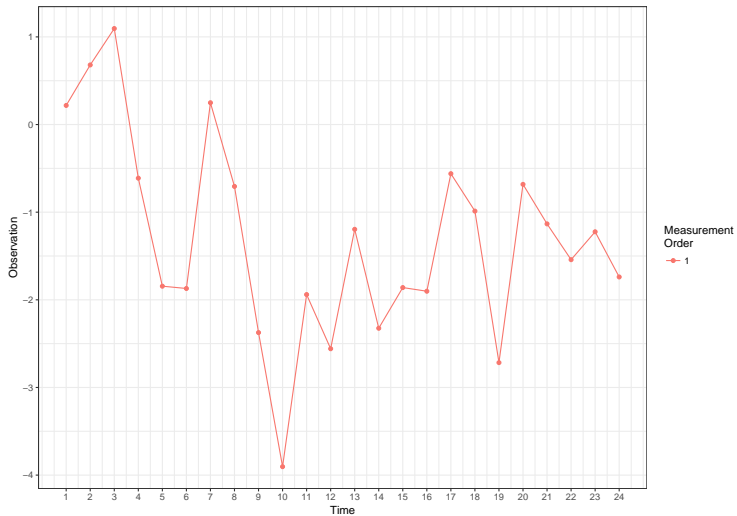
# Opakovaná měření



# Opakovaná měření

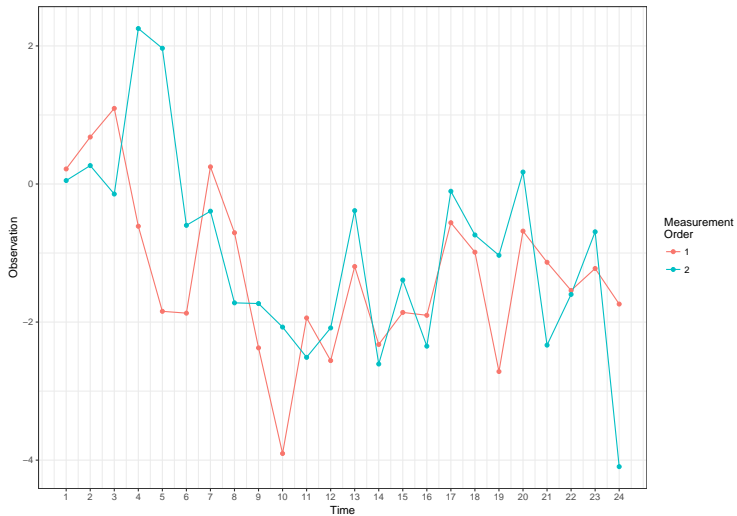


# Opakovaná měření

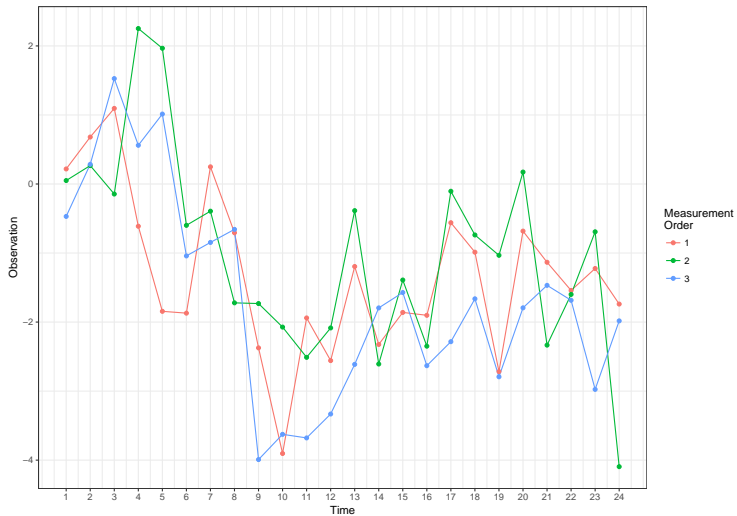




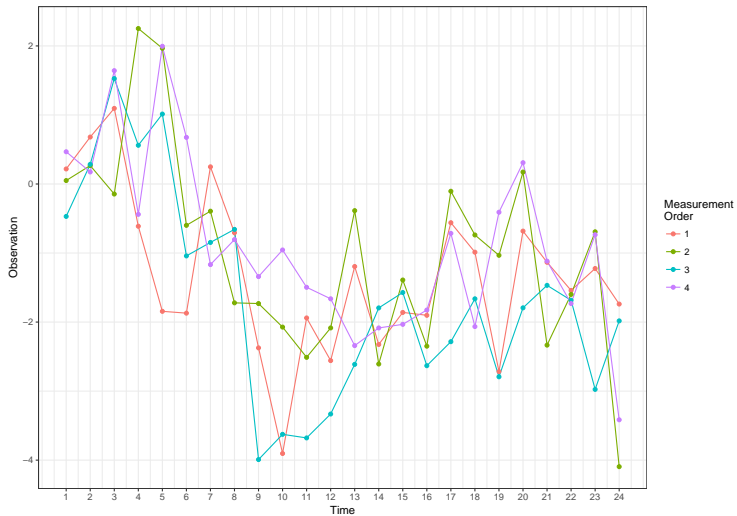
# Opakovaná měření



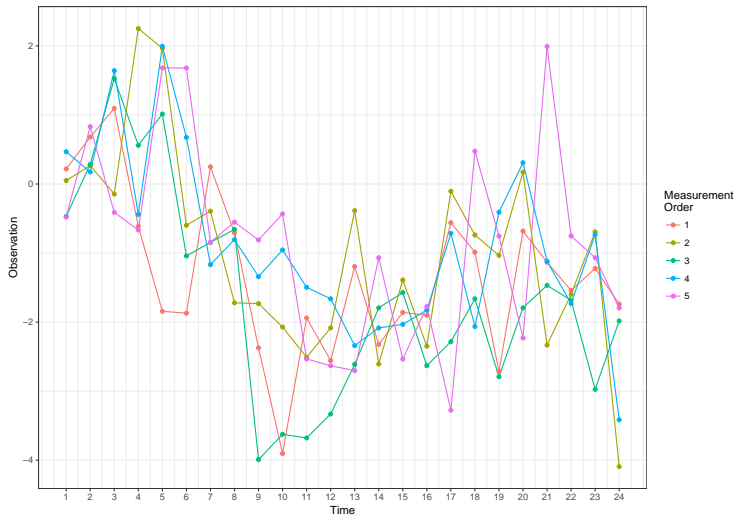
# Opakovaná měření



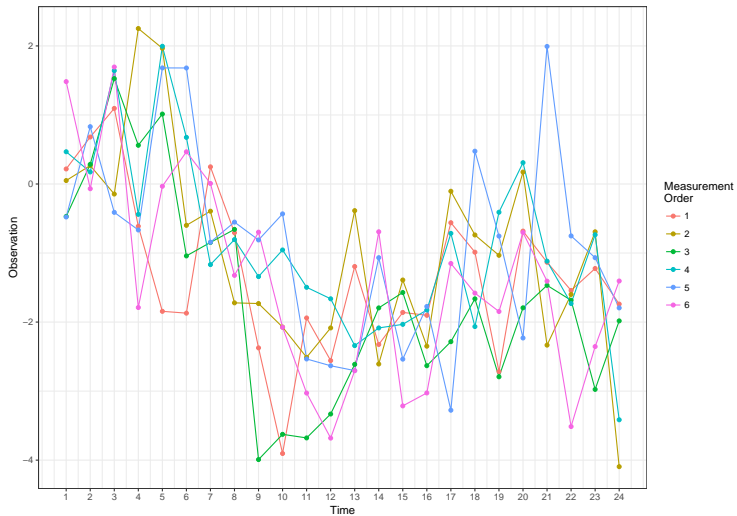
# Opakovaná měření



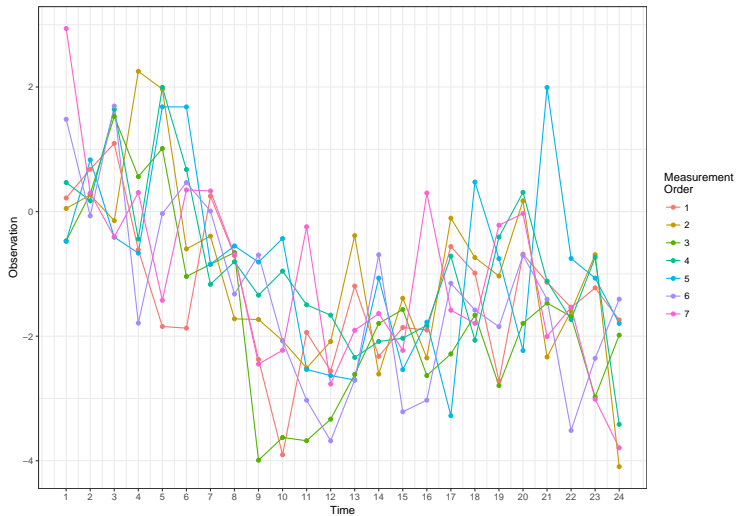
# Opakovaná měření



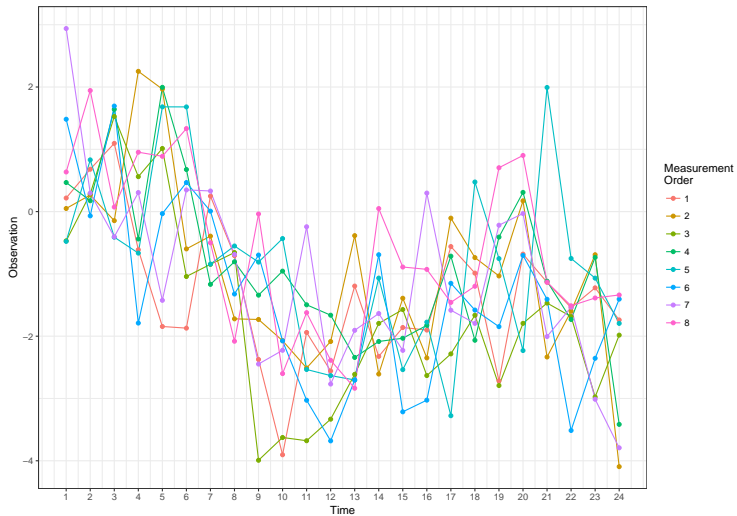
# Opakovaná měření



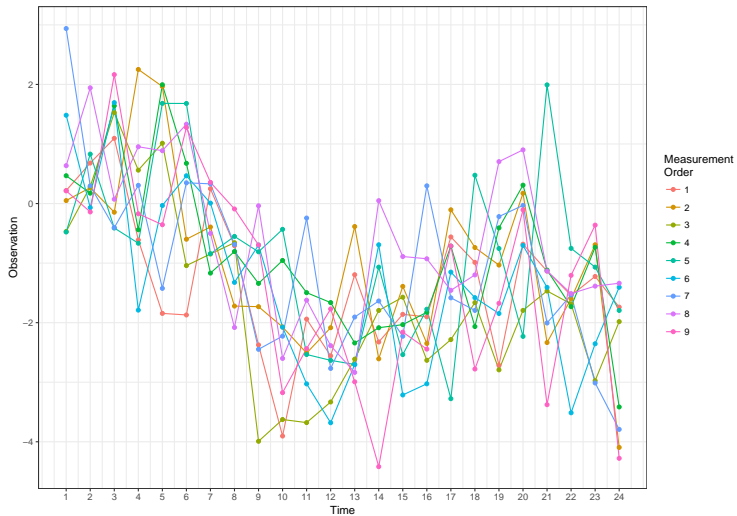
# Opakovaná měření



# Opakovaná měření

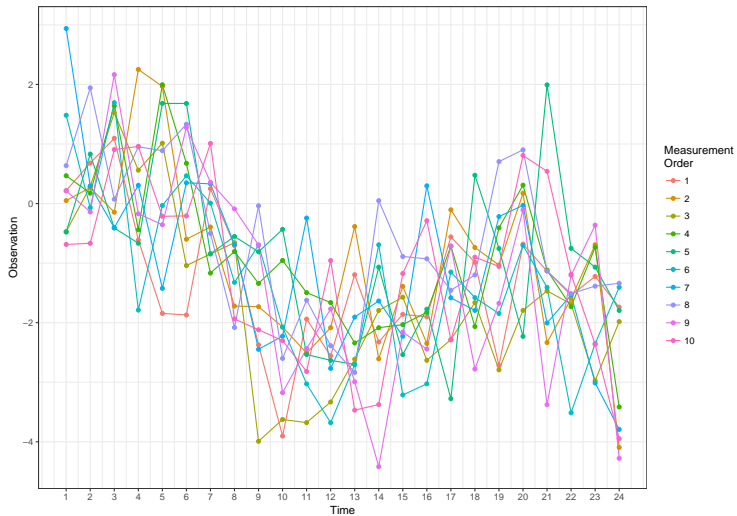


# Opakovaná měření

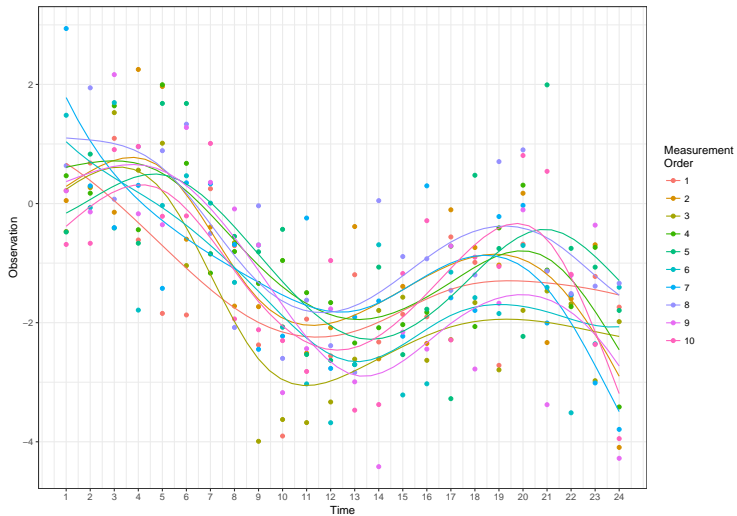




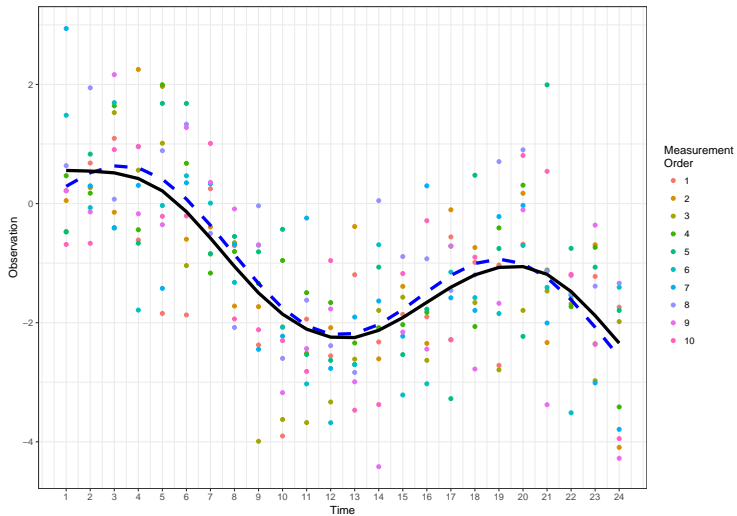
# Opakovaná měření



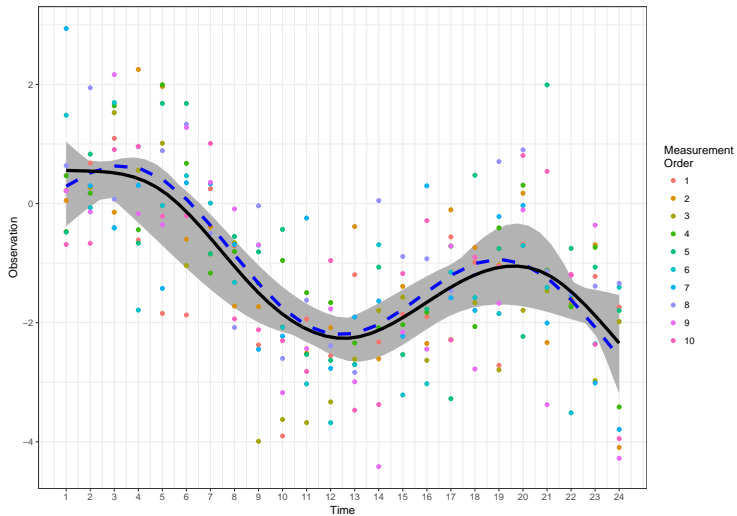
# Opakovaná měření



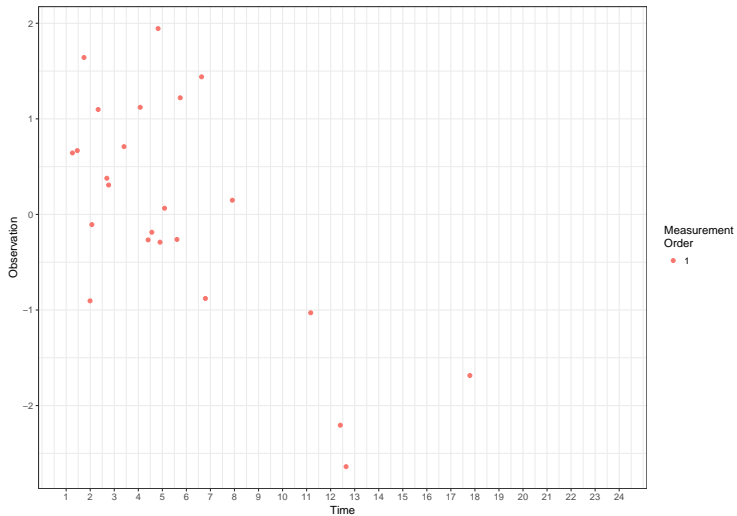
# Opakovaná měření



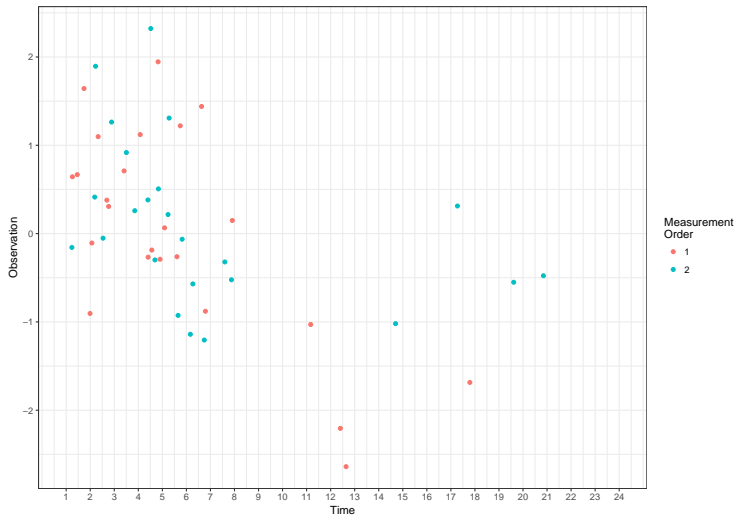
# Opakovaná měření



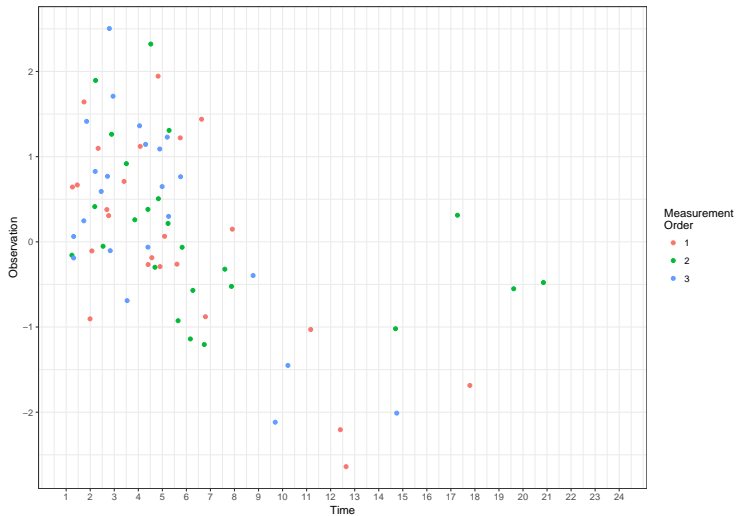
# Řídká data



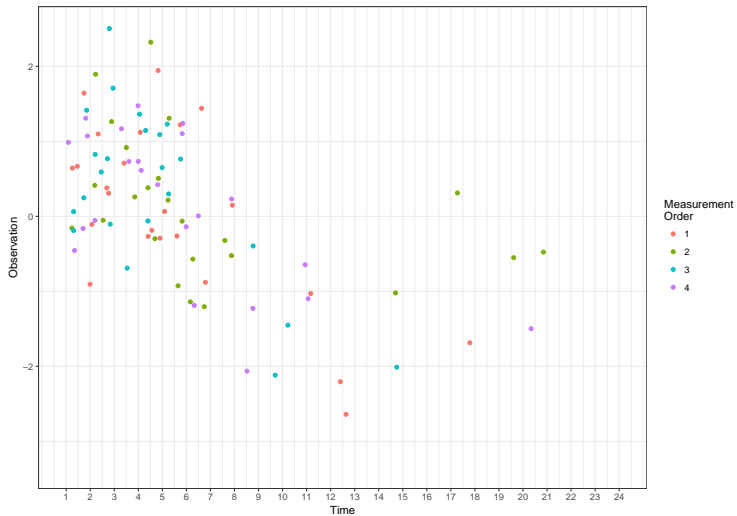
# Řídká data



# Řídká data

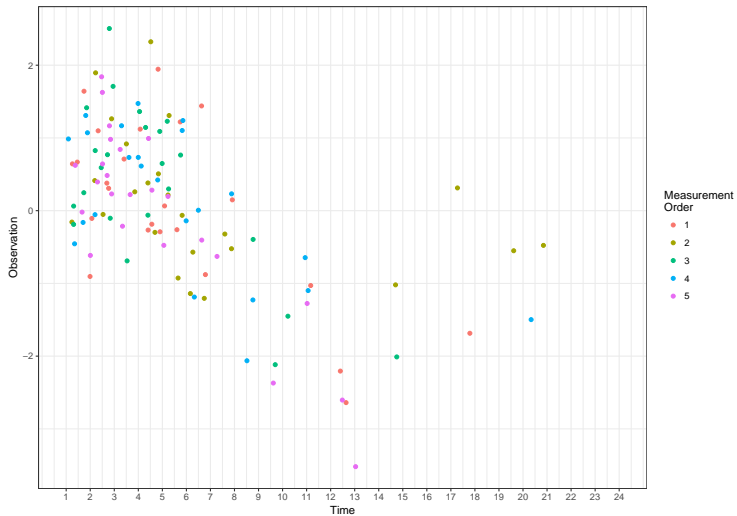


# Řídká data

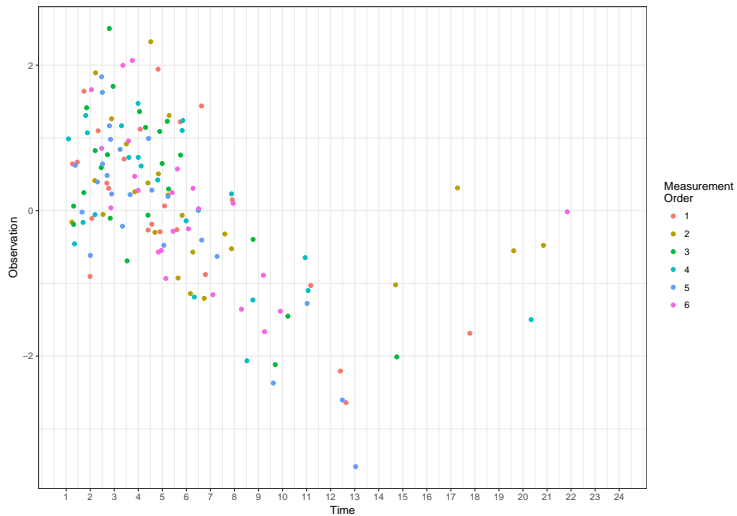




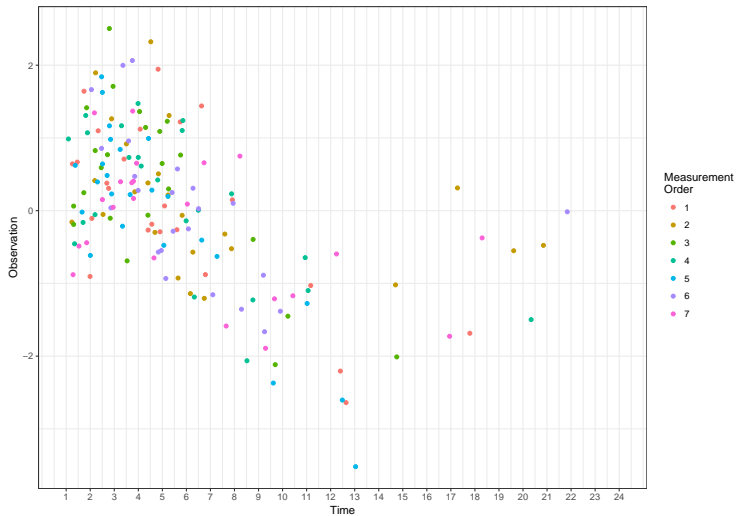
# Řídká data



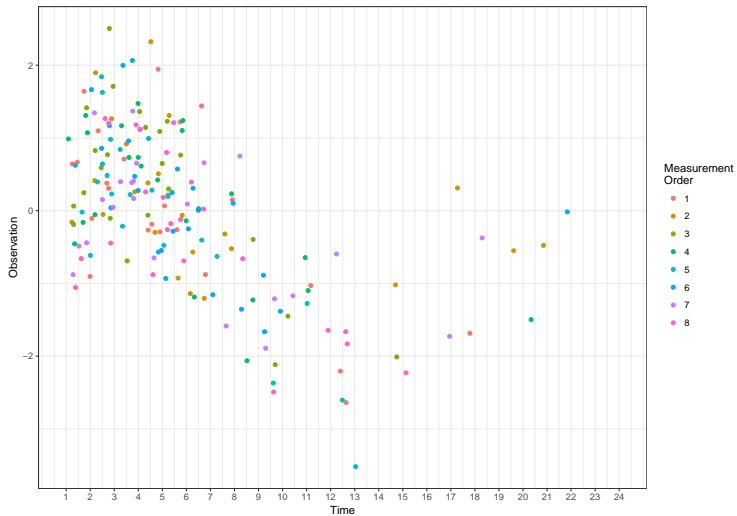
# Řídká data



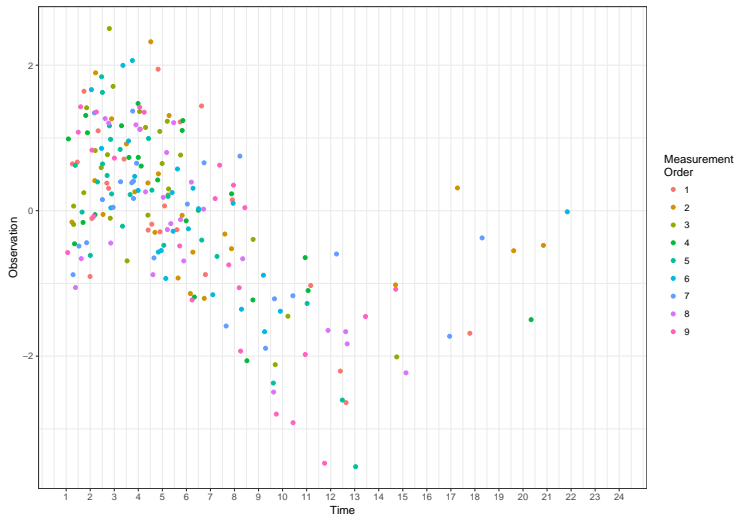
# Řídká data



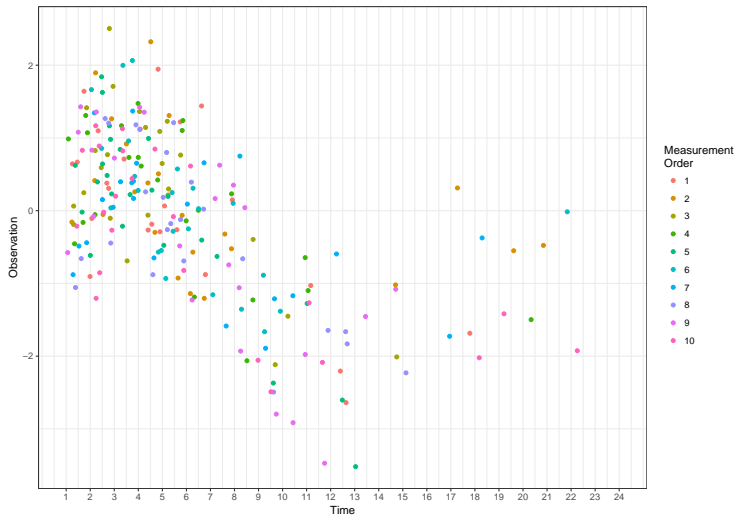
# Řídká data



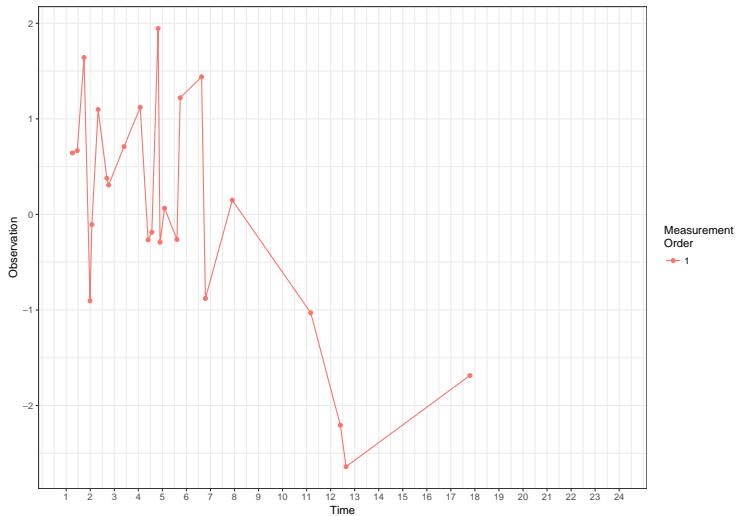
# Řídká data



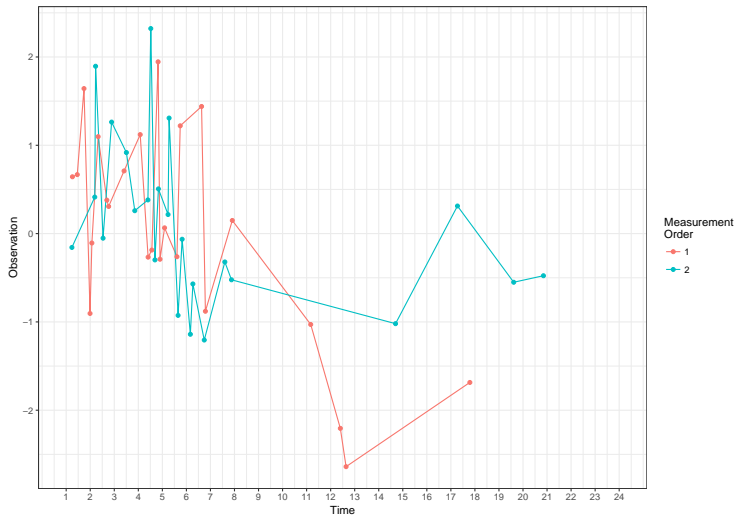
# Řídká data



# Řídká data

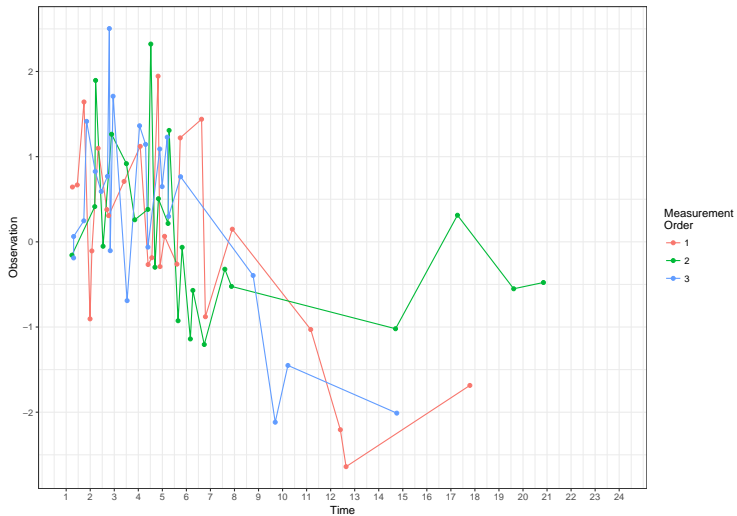


# Řídká data

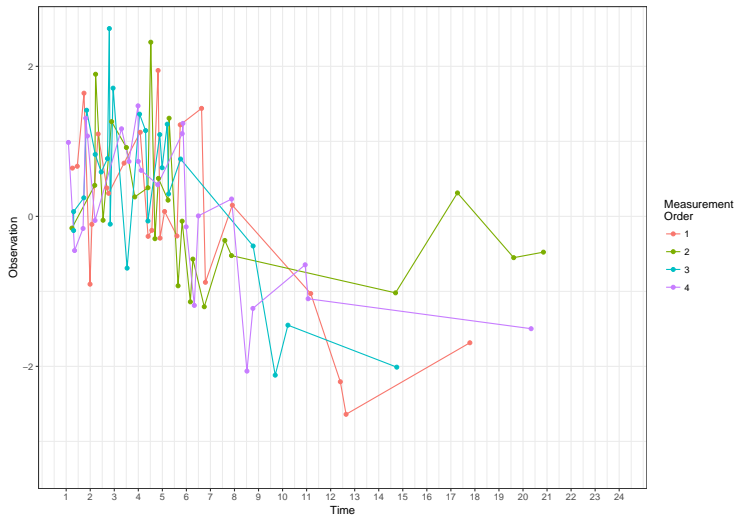




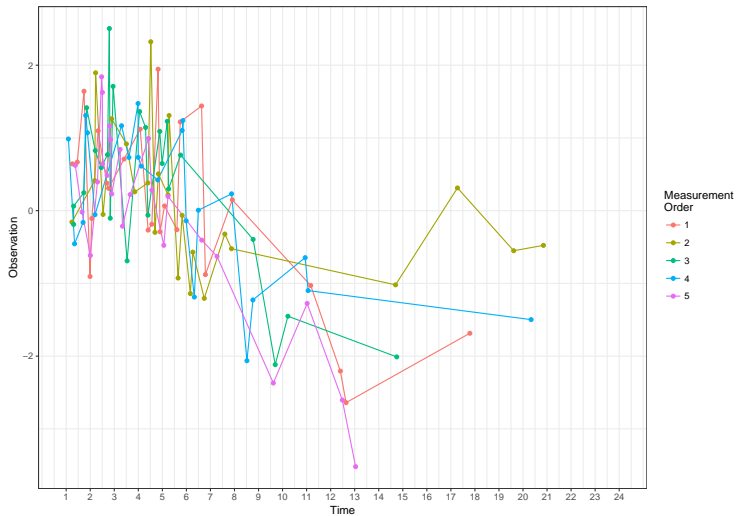
# Řídká data



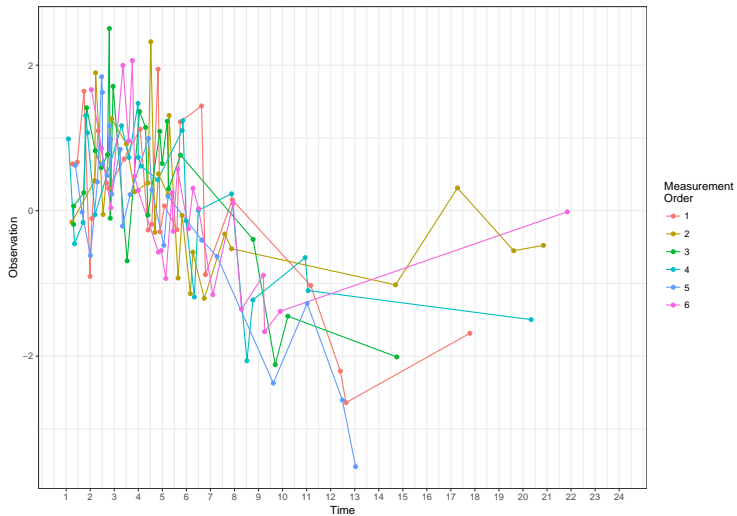
# Řídká data



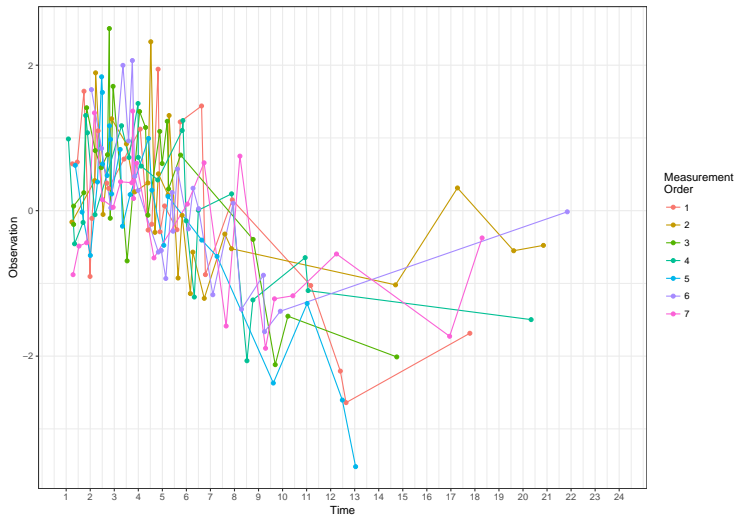
# Řídká data



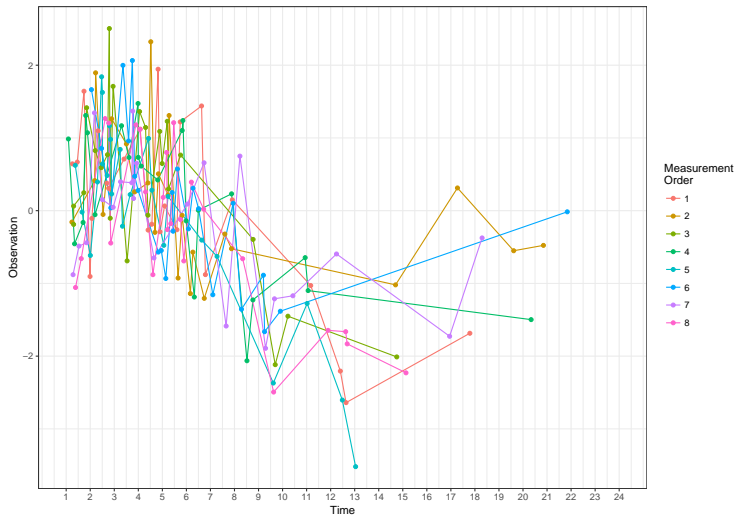
# Řídká data



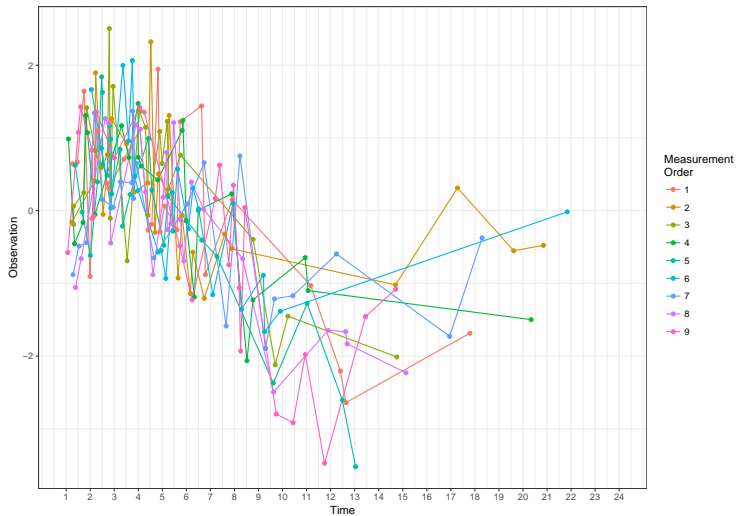
# Řídká data



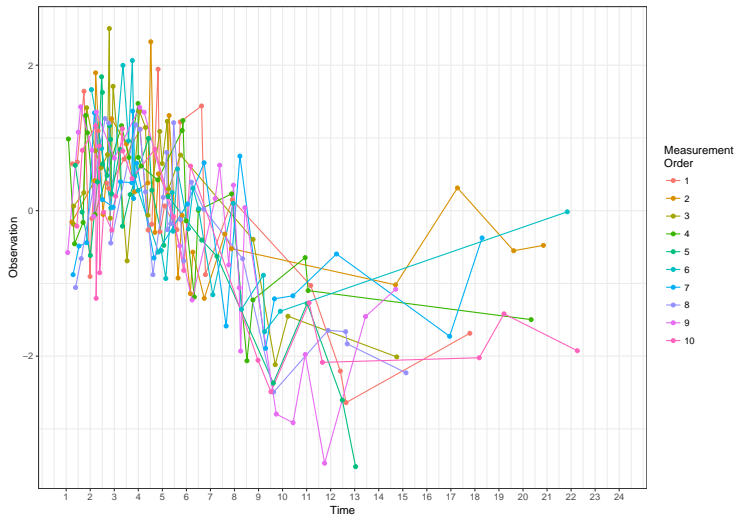
# Řídká data



# Řídká data

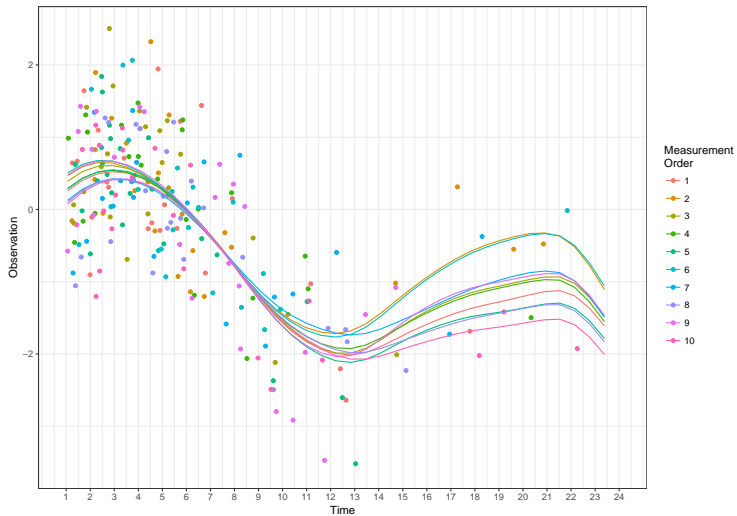


# Řídká data

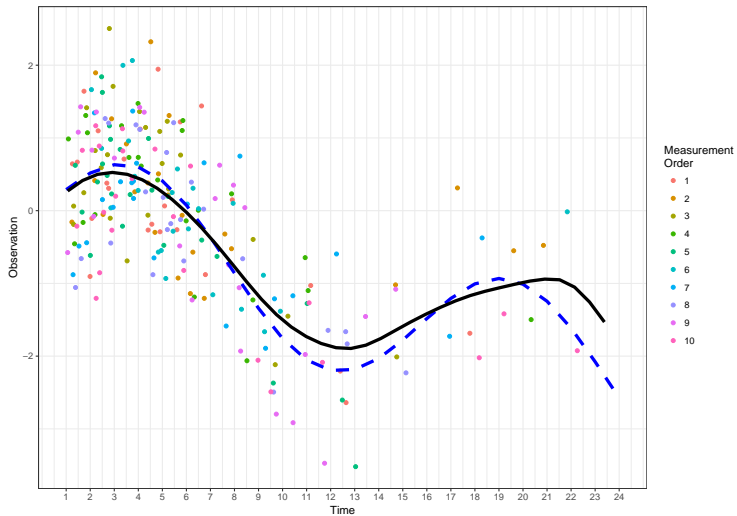




# Řídká data



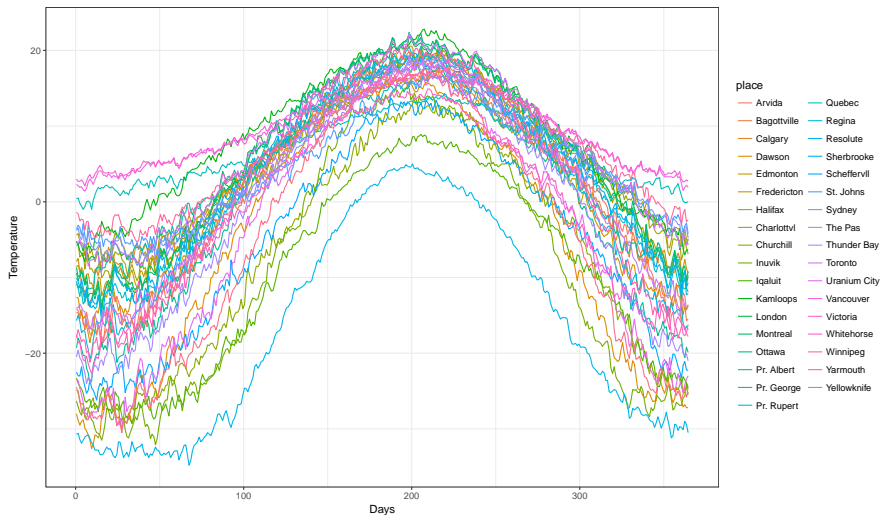
# Řídká data



## Počasí v Kanadě

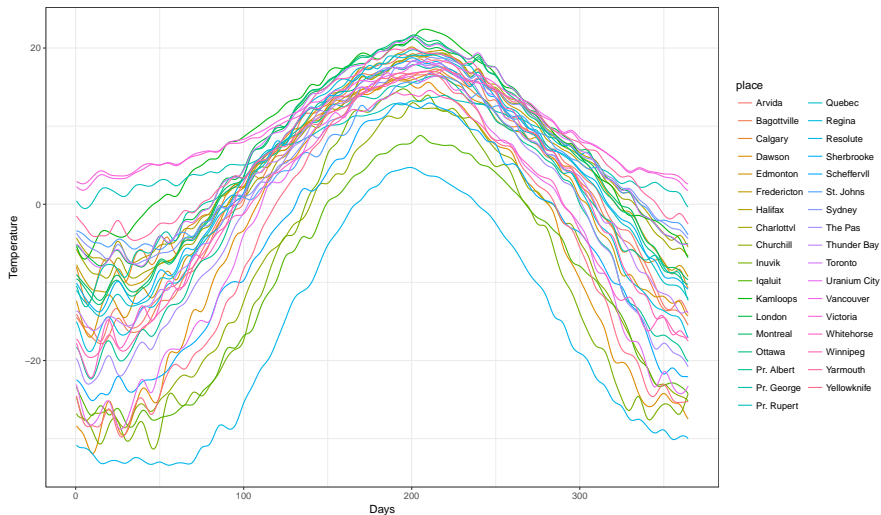
- roční srážky a teploty v Kanadě
- průměrné denní hodnoty za období 1960 – 1994
- měřeno v 35 různých lokalitách

# Canadian Weather



Teploty

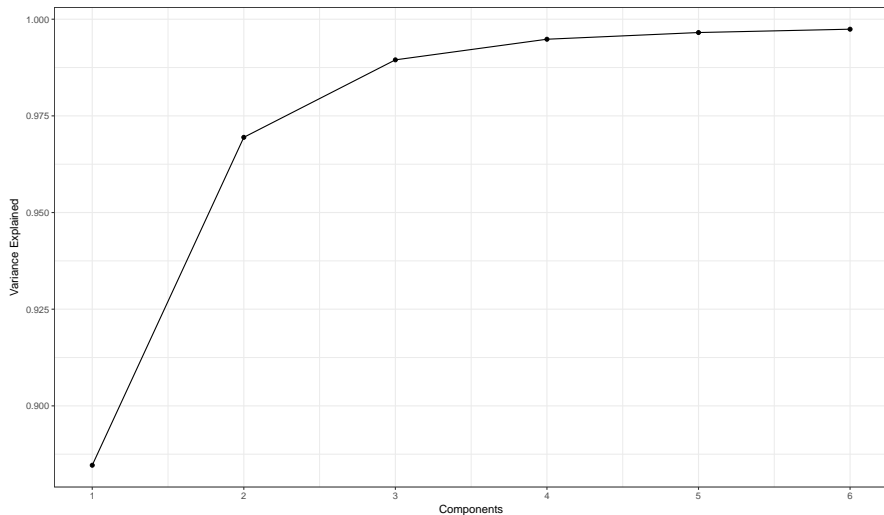
# Canadian Weather



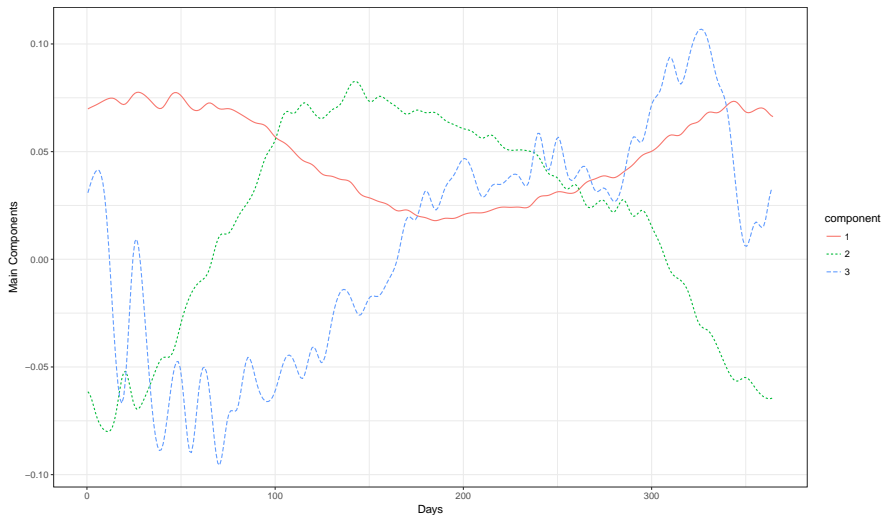
Vyhlazení

## Funkcionální analýza hlavních komponent – **FPCA**

- každá křivka je reprezentována konečným počtem „jiných“ křivek
- „jiné“ křivky se nazývají **komponenty**
- kolik komponent?

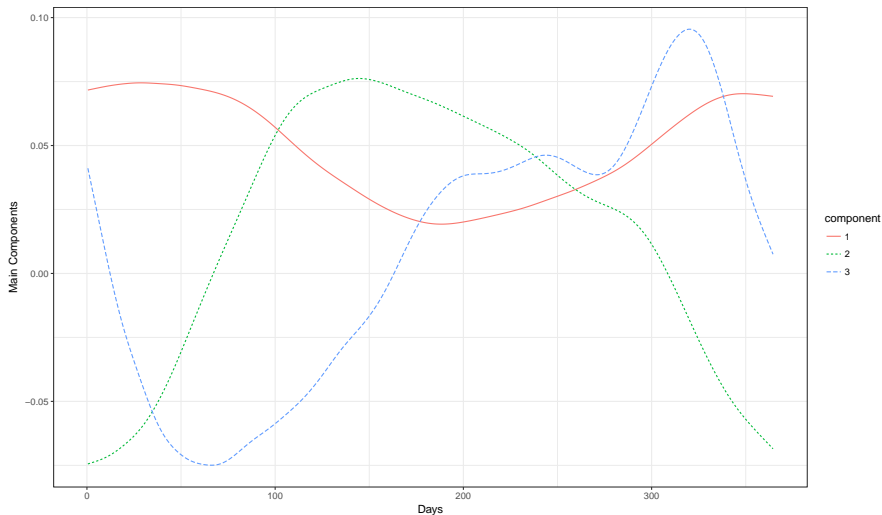


Podíl komponent na vysvětlení rozptylu

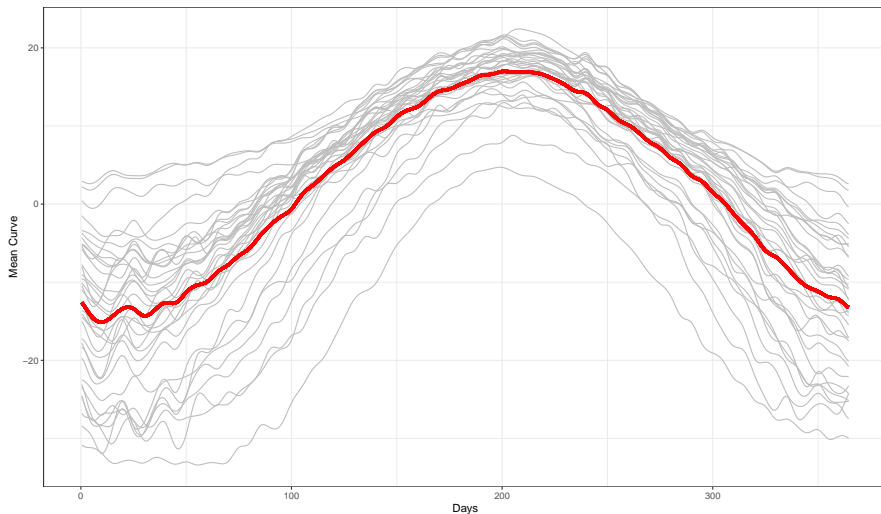


První 3 hlavní komponenty



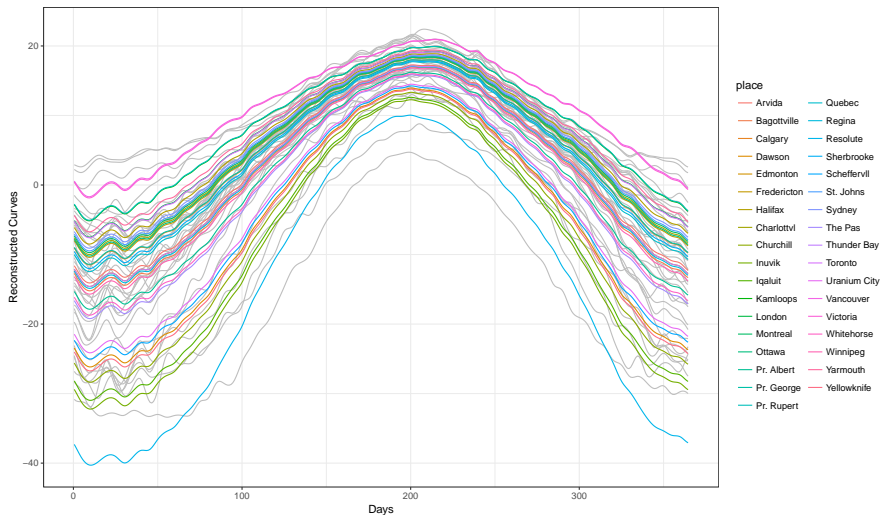


První 3 hlavní komponenty – vyhlazené



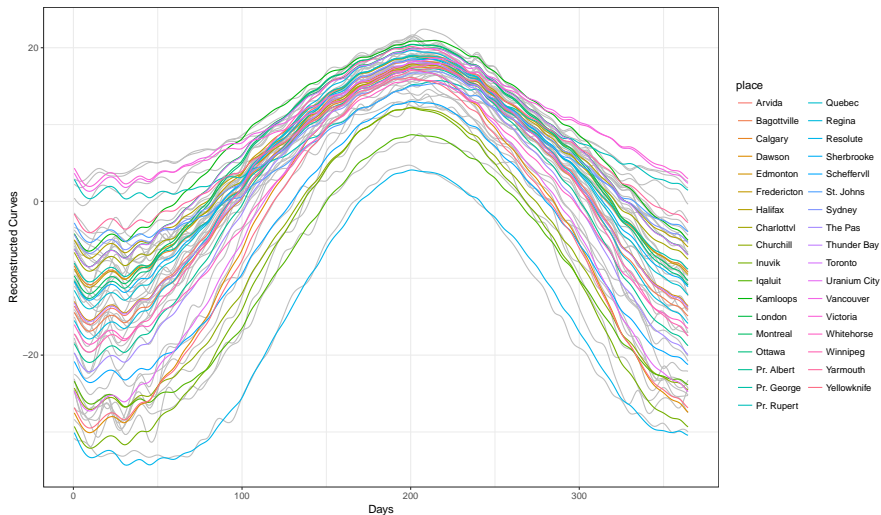
Průměrná křivka

# Canadian Weather



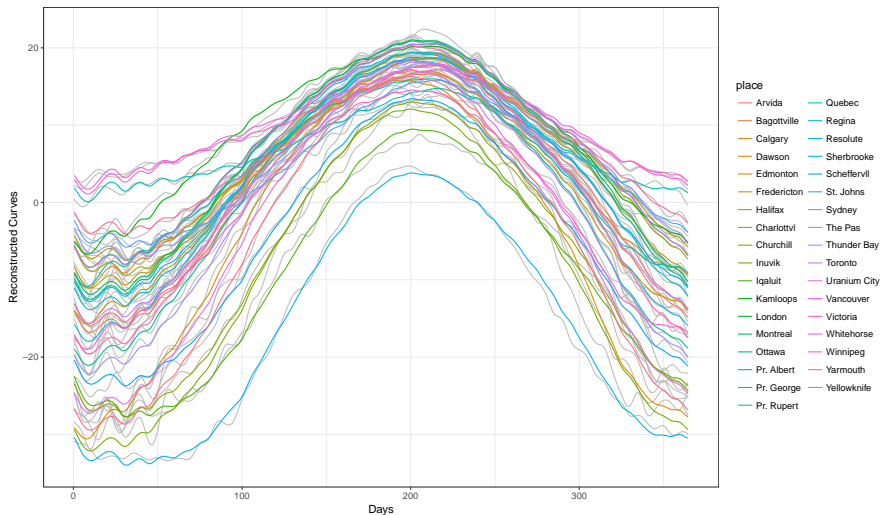
Rekonstrukce – 1 hlavní komponenta

# Canadian Weather

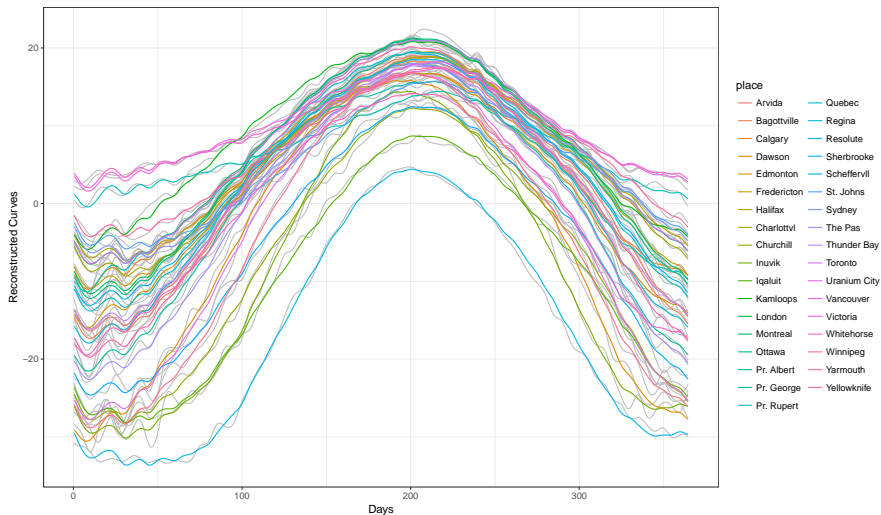


Rekonstrukce – 2 hlavní komponenty

# Canadian Weather

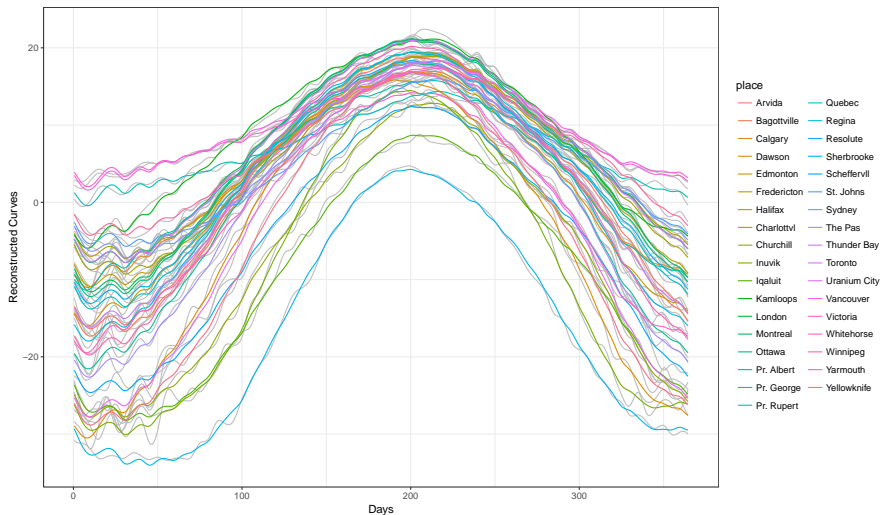


Rekonstrukce – 3 hlavní komponenty



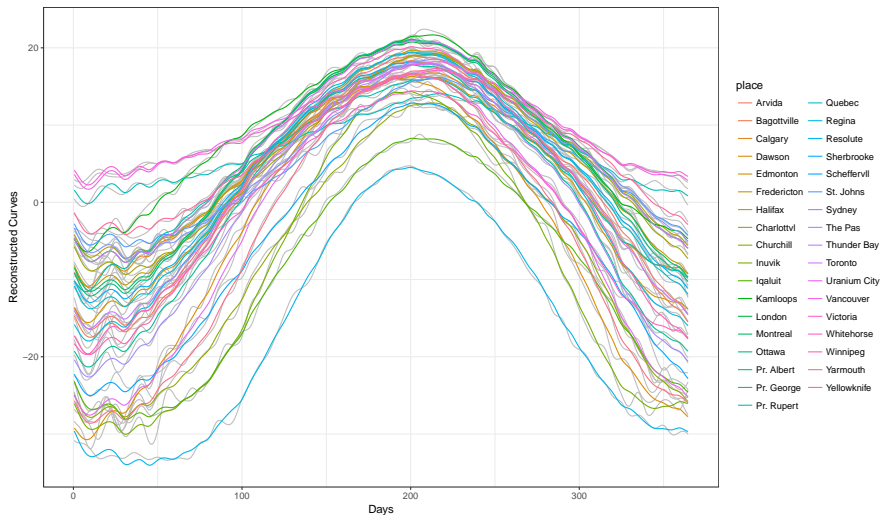
Rekonstrukce – 4 hlavní komponenty

# Canadian Weather



Rekonstrukce – 5 hlavních komponent

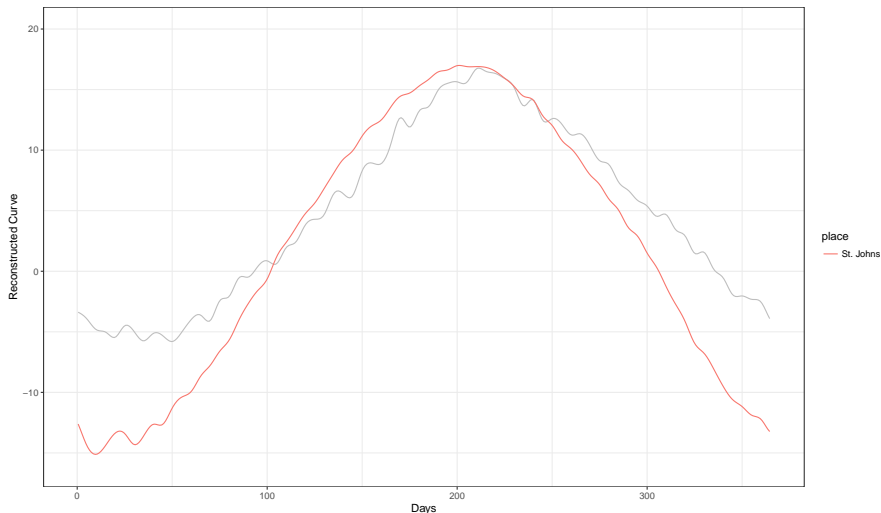
# Canadian Weather



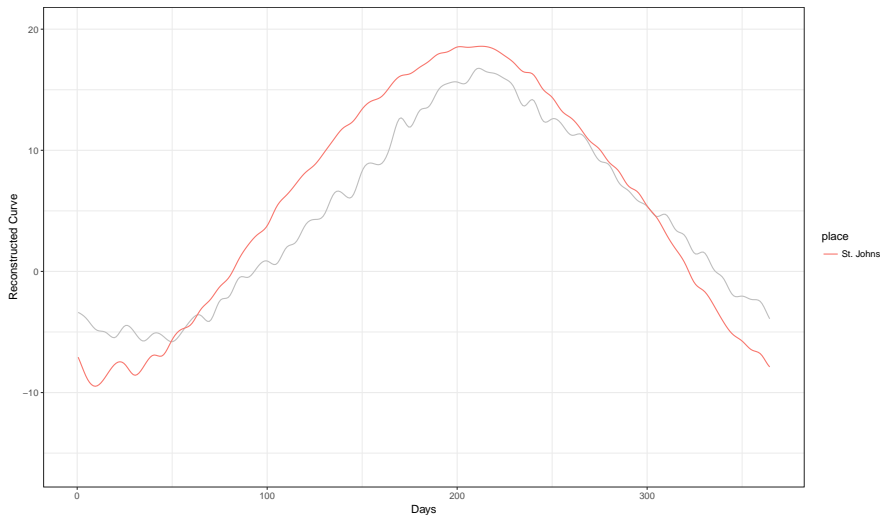
Rekonstrukce – 6 hlavních komponent



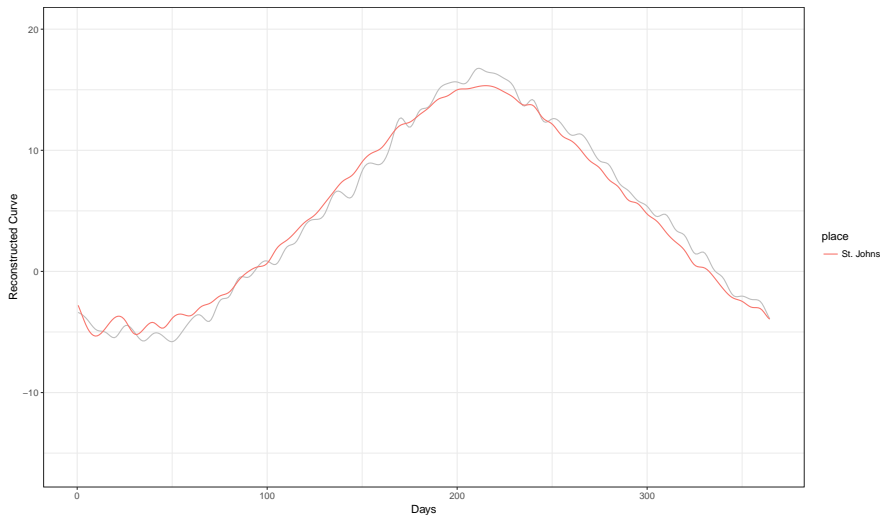
# Canadian Weather



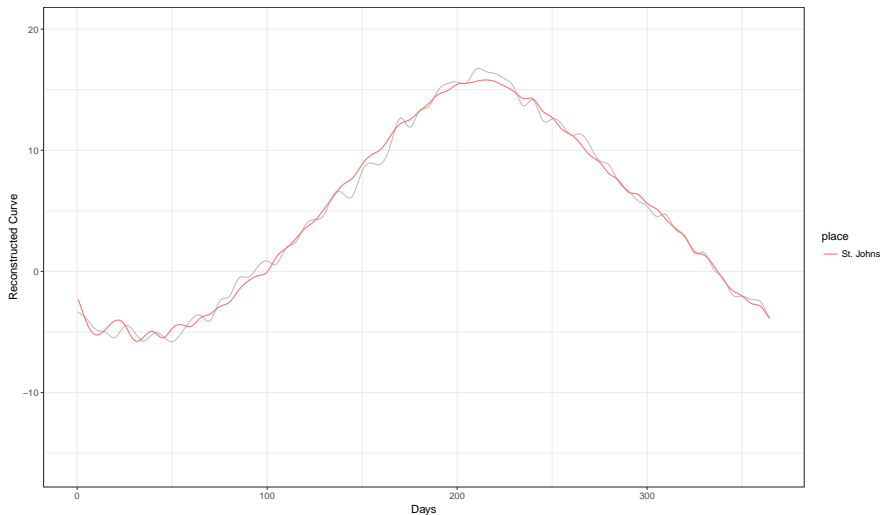
Průměrná křivka



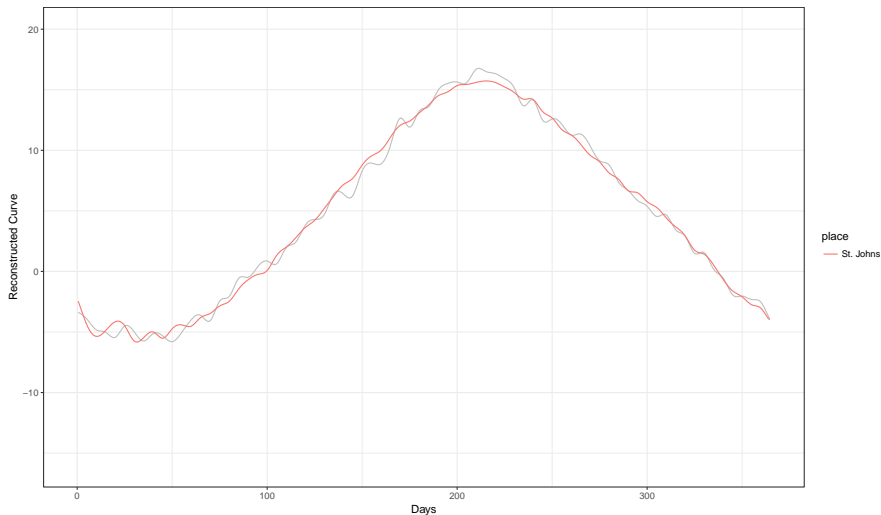
Rekonstrukce – 1 hlavní komponenta



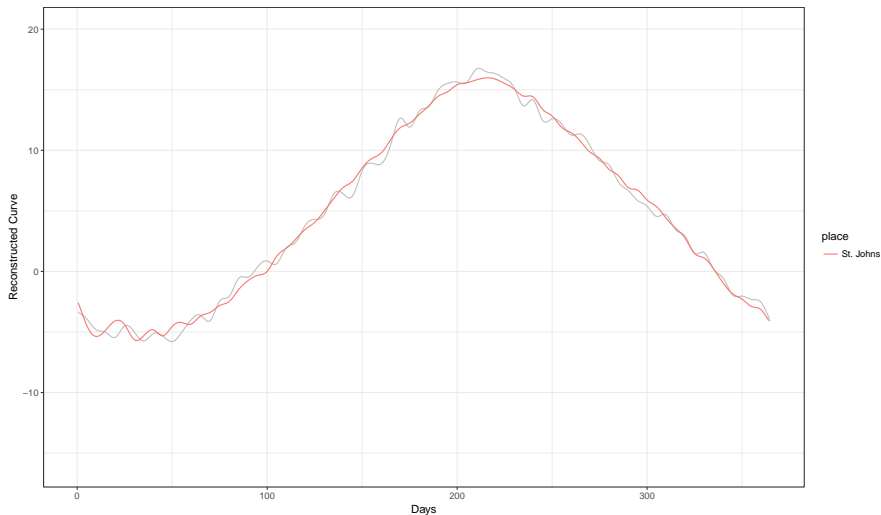
Rekonstrukce – 2 hlavní komponenty



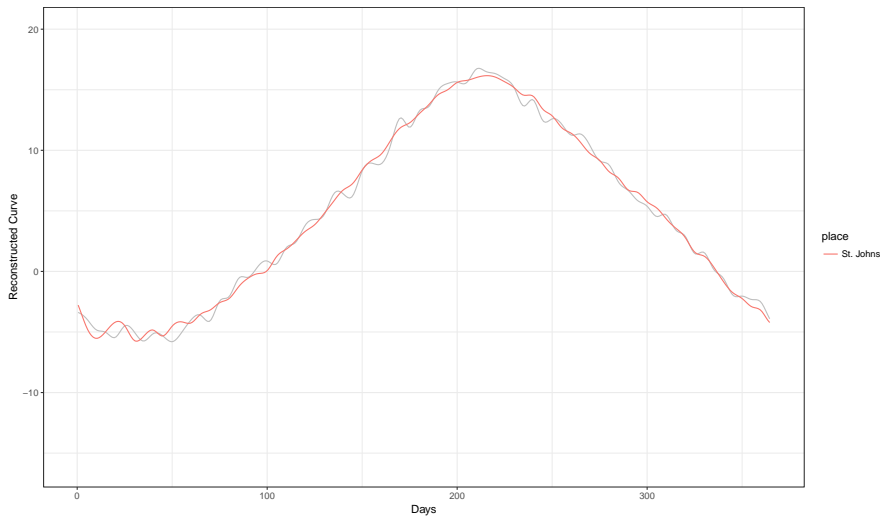
Rekonstrukce – 3 hlavní komponenty



Rekonstrukce – 4 hlavní komponenty

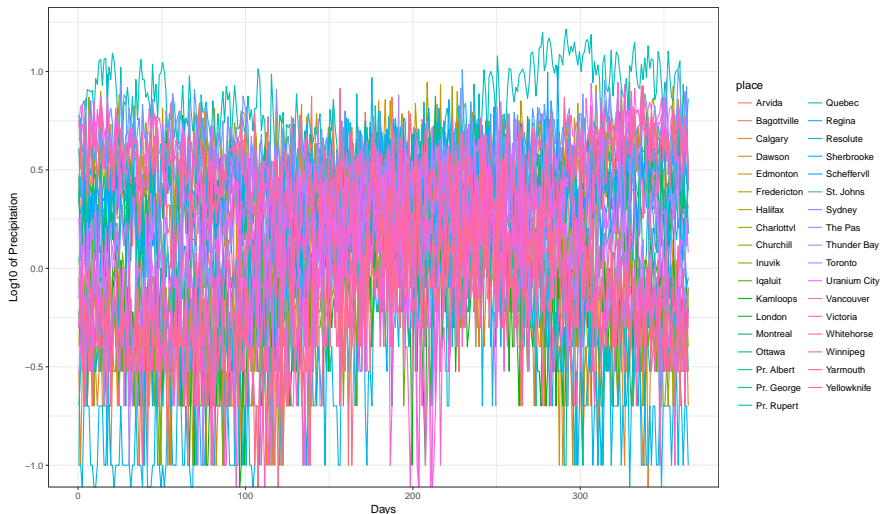


Rekonstrukce – 5 hlavních komponent



Rekonstrukce – 6 hlavních komponent

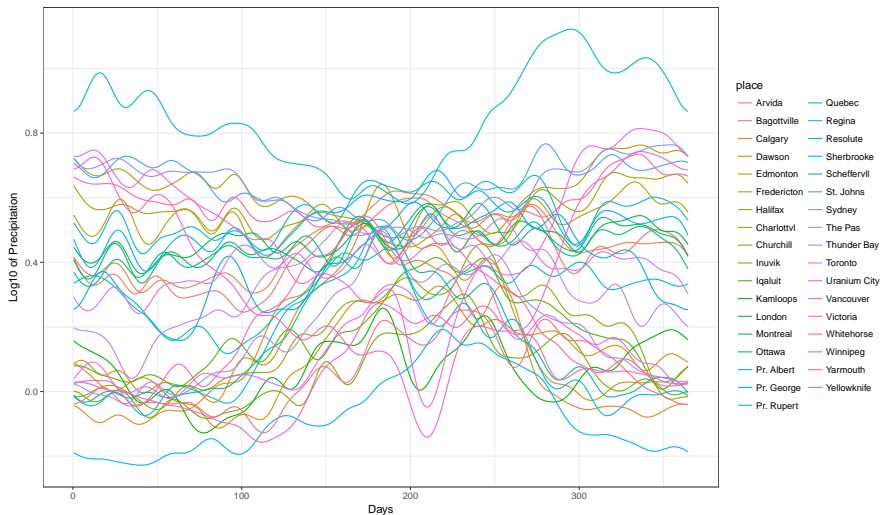
# Canadian Weather



Srážky

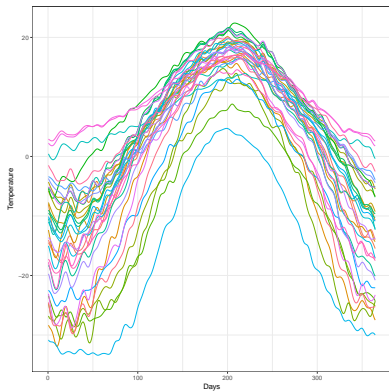


# Canadian Weather



Vyhlazená data

# Canadian Weather

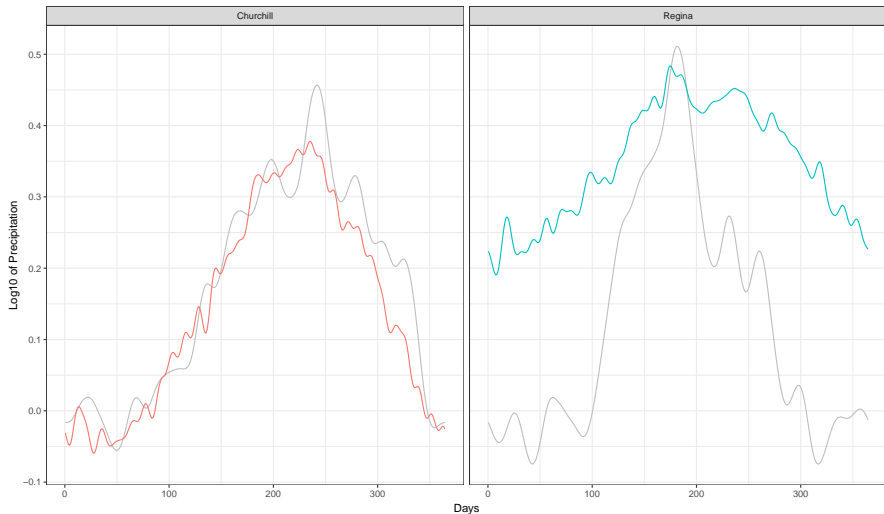


$\beta_0, \beta_1$   
 $\longrightarrow$



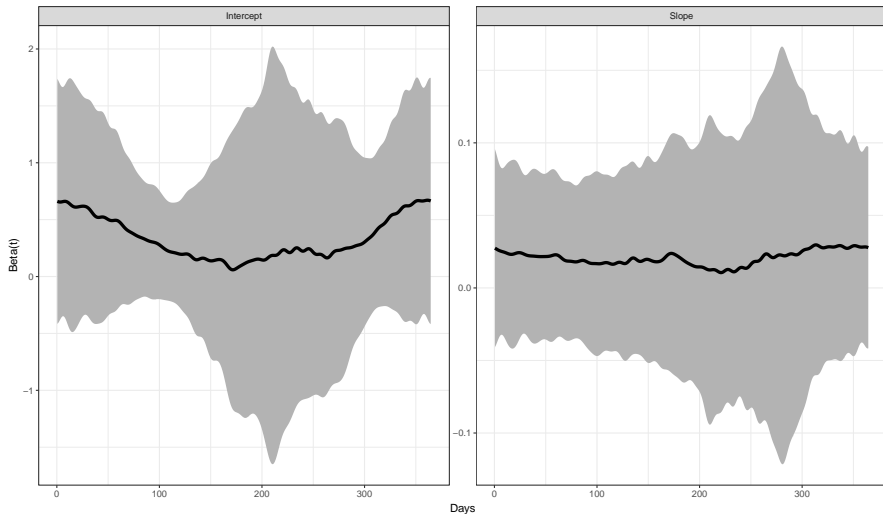
**Regresní model:** srážky  $\sim$  teploty

# Canadian Weather



Odhady srážek

# Canadian Weather

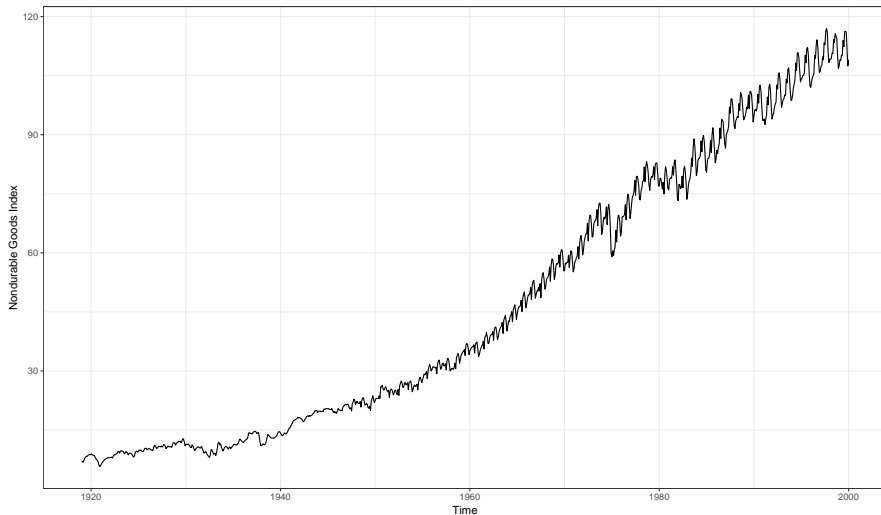


Odhady parametrů modelu,  $\hat{\beta}_0(t)$ ,  $\hat{\beta}_1(t)$

## Index cen zboží krátkodobé spotřeby v USA (1919 – 2000)

- produkty, které se opotřebí (vyčerpají) do 2 let (potraviny, tabák, oblečení, papírové výrobky, služby, pohonné hmoty, atd.)
- v indexu se odráží ekonomika každodenního života
- exponenciální růst v minulém století
- patrná sezónní složka (např. zahrnuje vánoční dárky, začátek školního roku – nové šaty, prázdniny – dovolené)
- zajímavé sledovat vývoj indexu v „normálních“ časech a v době krize

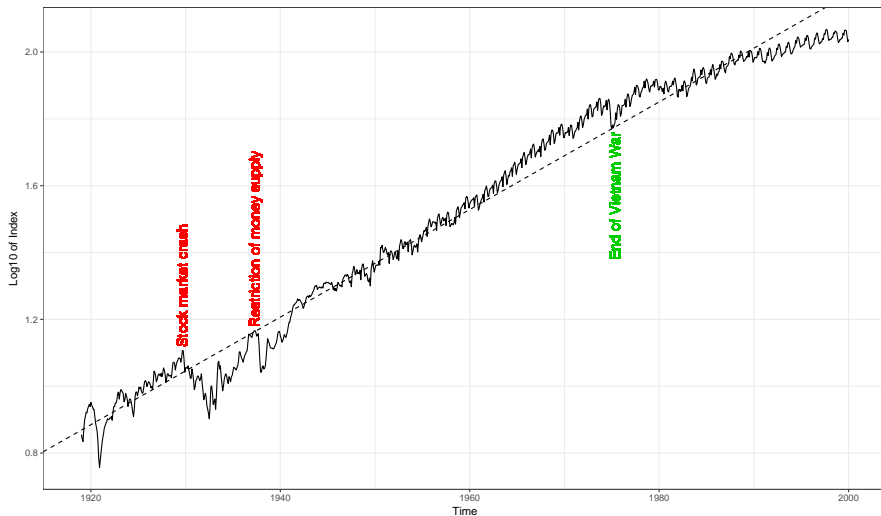
# Nondurable Goods Index



Index cen zboží krátkodobé spotřeby v USA (1919 – 2000)

- exponenciální růst, roste i volatilita
- logaritmická transformace → lineární růst s indexem 1,6 %
- 1960 – 1975 „baby boom“

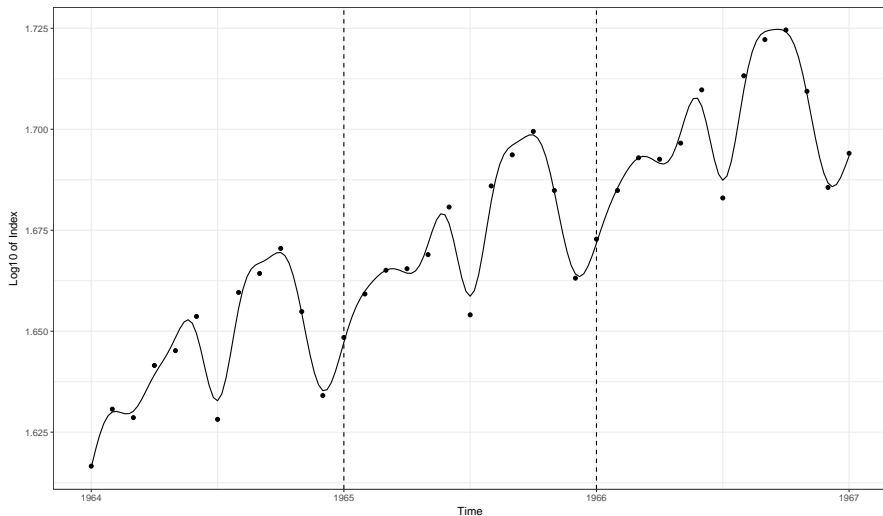
# Nondurable Goods Index



Zlogaritmovaná data



# Nondurable Goods Index

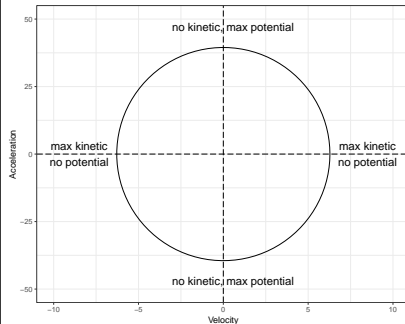
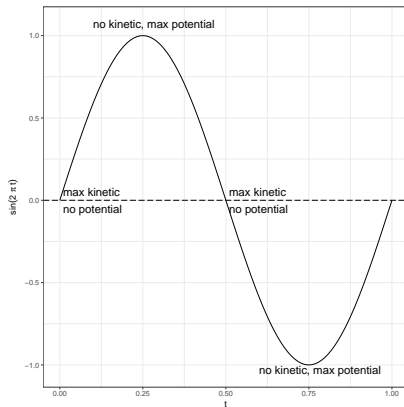


Zlogaritmovaná data 1964 – 1967

Fázový diagram (phase – plane plot)

- 1. derivace – rychlost růstu (velocity)
- 2. derivace – moment zrychlení (acceleration)
- Př.  $\sin(2\pi t)$  – harmonický proces, pružina oscilující mezi dvěma stavy (potenciální, kinetický)
- kinetická energie – výrobní proces v plném proudu a množství vyprodukovaných výrobků
- potenciální energie – volný kapitál, lidské zdroje, suroviny a ostatní zdroje, které jsou aktuálně dostupné pro růst ekonomiky

# Nondurable Goods Index

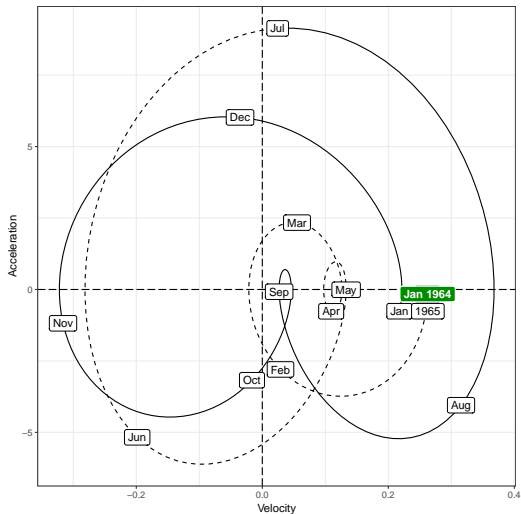


Phase – plane plot pro  $\sin(2\pi t)$

V diagramu sledujeme:

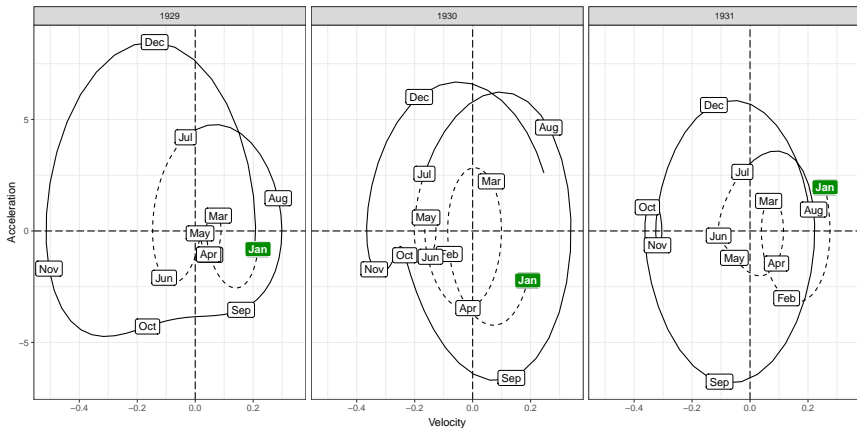
- podstatné cykly
- poloměr cyklu
- horizontální polohu středu
- vertikální polohu středu
- změny tvaru v čase

# Nondurable Goods Index



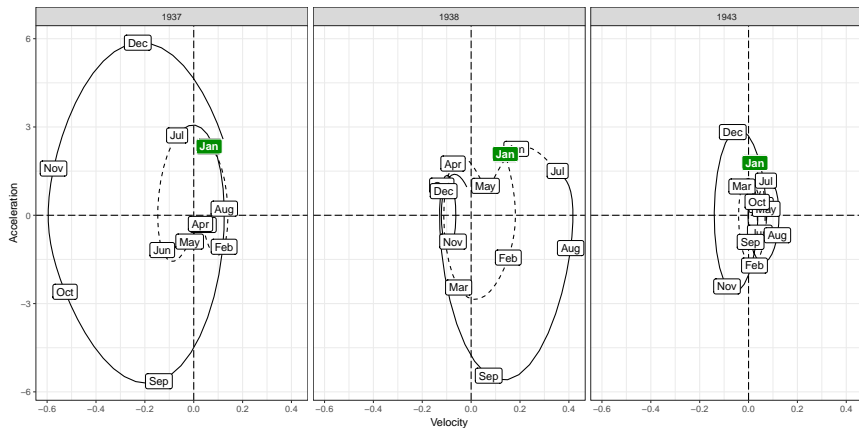
„Klidný“ rok 1964

# Nondurable Goods Index



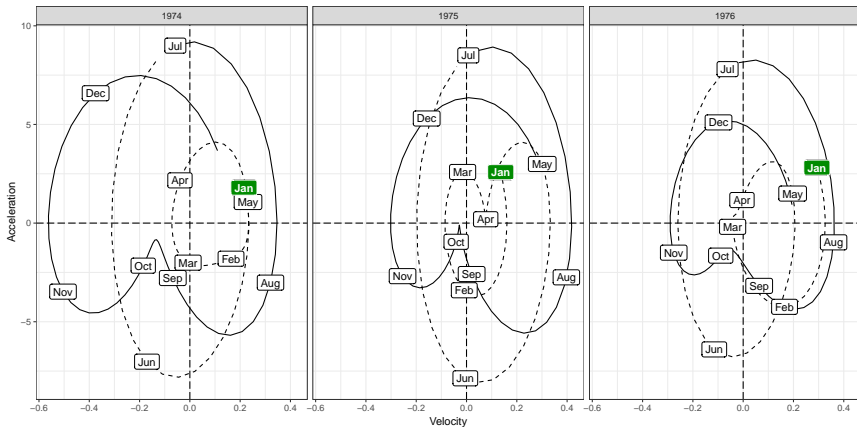
Nov 1929 – krach na burze

# Nondurable Goods Index



Předválečné období, válka

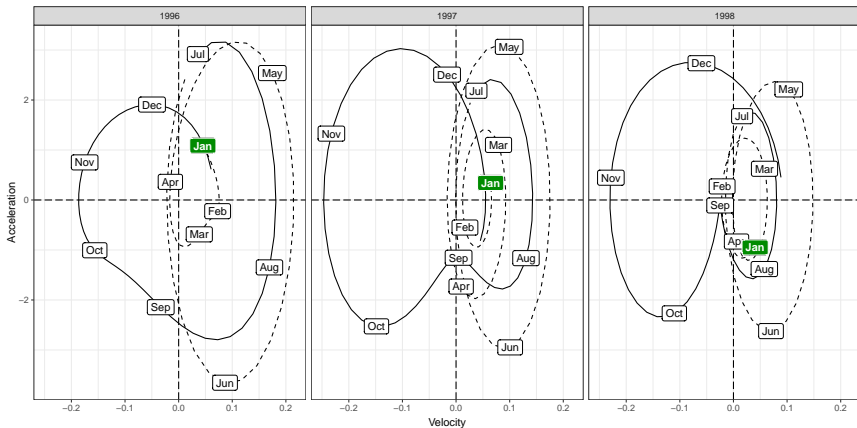
# Nondurable Goods Index



Apr 1975 – konec války ve Vietnamu



# Nondurable Goods Index



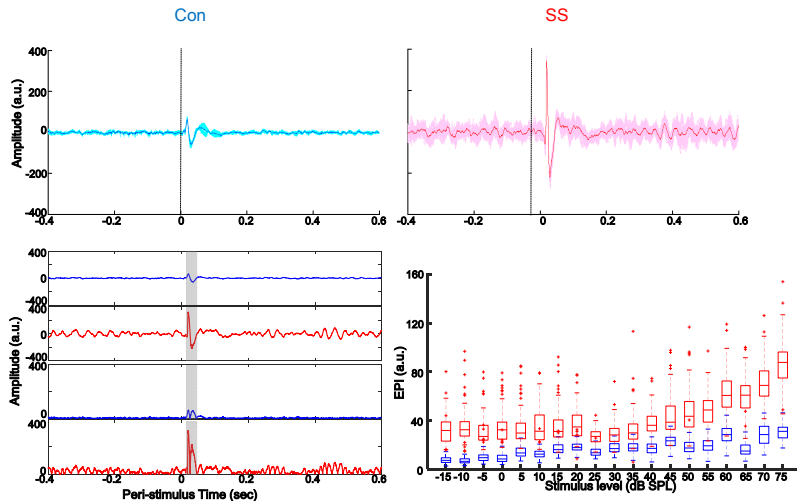
Ekonomika 90. let

## Neuronová aktivita při výskytu tinnitu

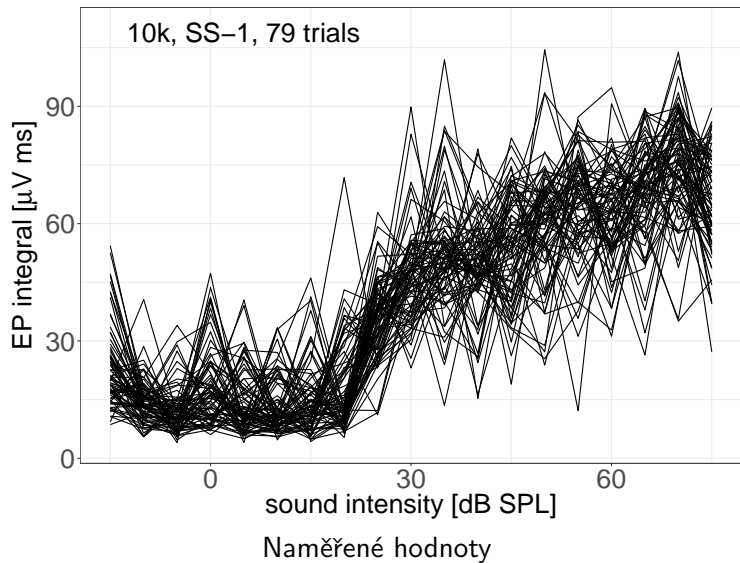
- měření sluchově evokovaného potenciálu **EPI** na kryse při různých intenzitách zvuku
- 19 intenzit zvuku: -15 dB, -10 dB, . . . , 70 dB, 75 dB
- 4 typy zvuku: tóny 1 kHz, 10 kHz, 16 kHz a „click“ (směs)
- 79 měření každý den
- 6 kontrolních dní (**con**)
- 8 dní po aplikaci salycilátu sodného (**SS**)

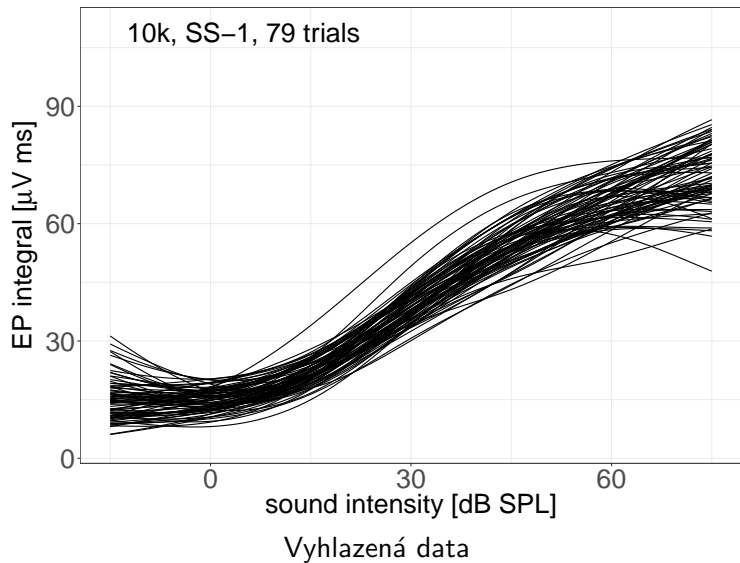
*Zdroj: Department of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan*

# Auditory EPI

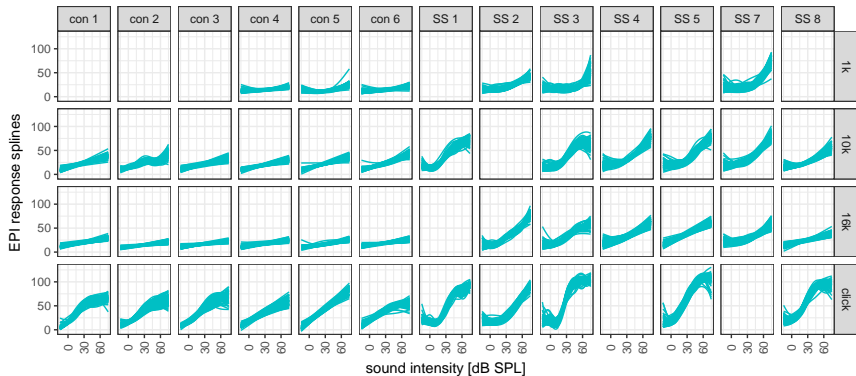


Měření neuronové aktivity



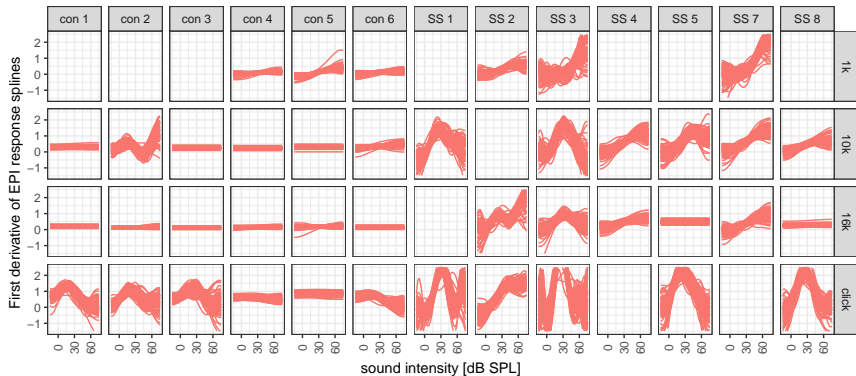


# Auditory EPI



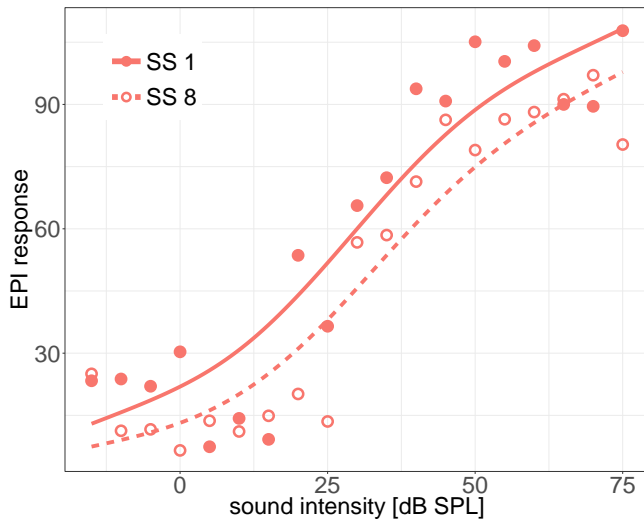
Vyhlazená data

# Auditory EPI



Vyhlazená data – 1. derivace

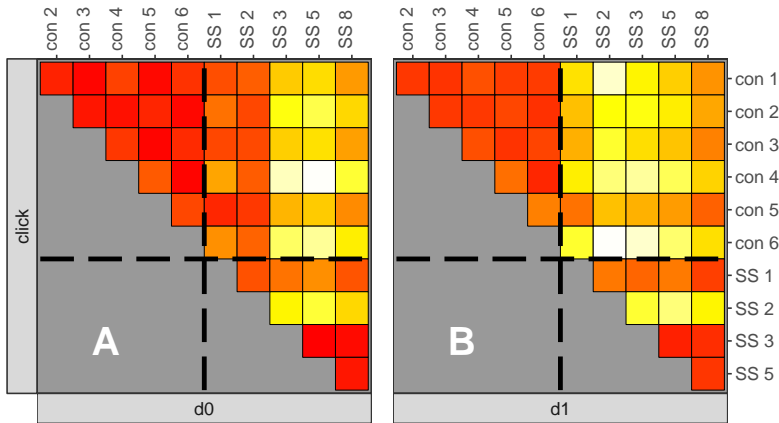
# Auditory EPI



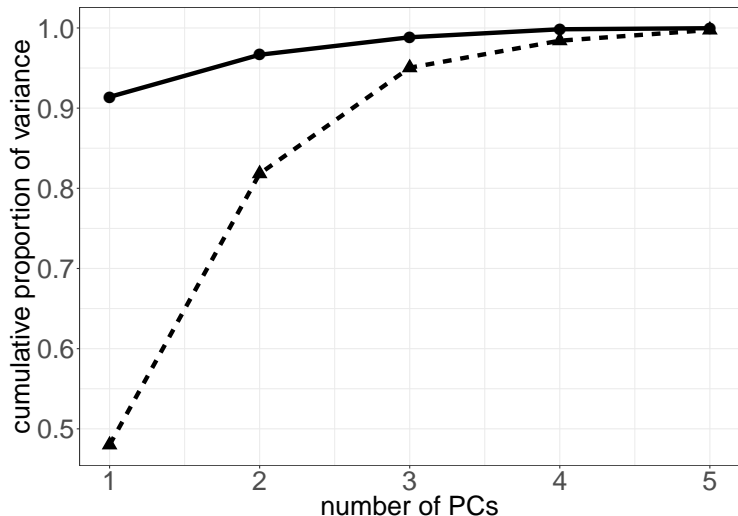
(Semi)metriky:  $d_0 = 112,8$ ;  $d_1 = 1,58$



# Auditory EPI

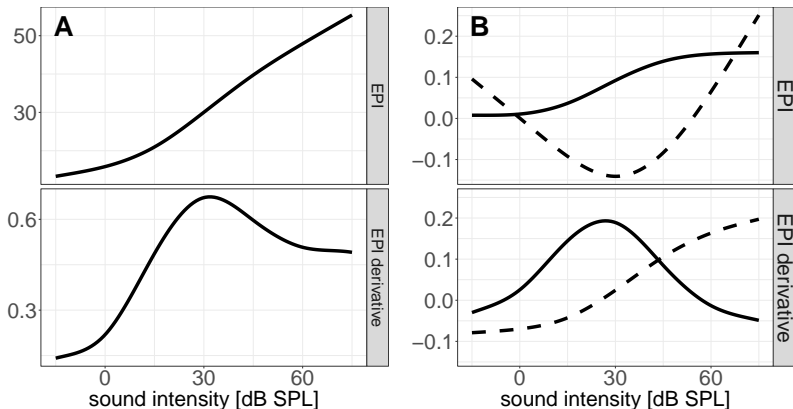


A – medián  $d_0$ , B – medián  $d_1$



Podíl komponent na vysvětlení rozptylu

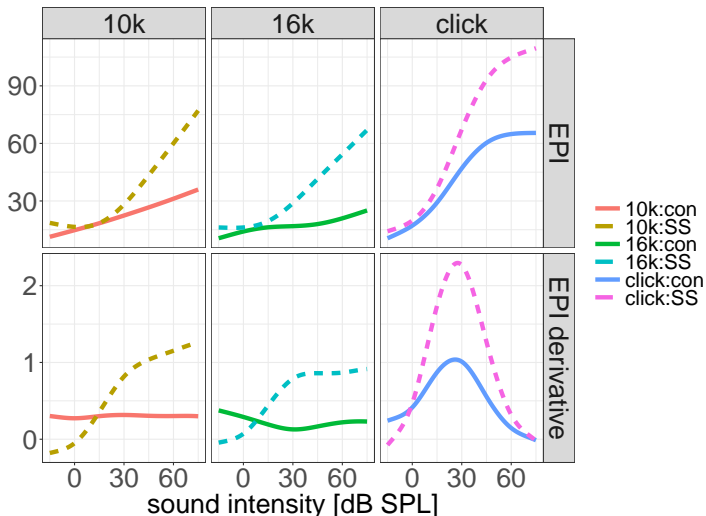
# Auditory EPI



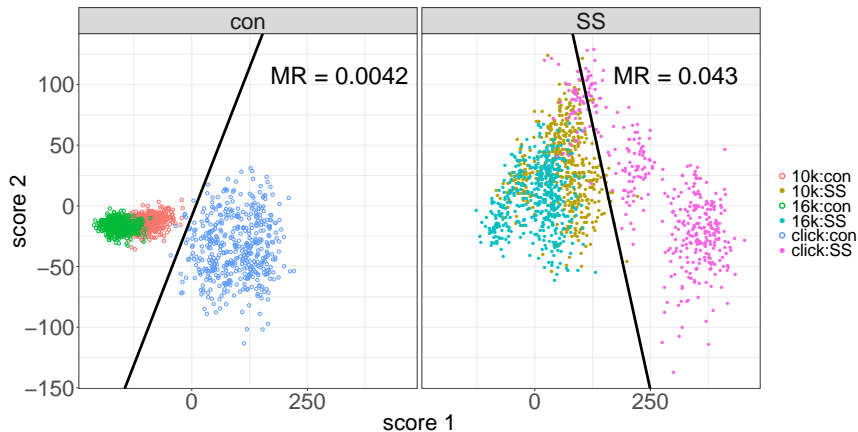
A – průměrné křivky,

B – 1. komponenta (plná), 2. komponenta (čárkovaně)

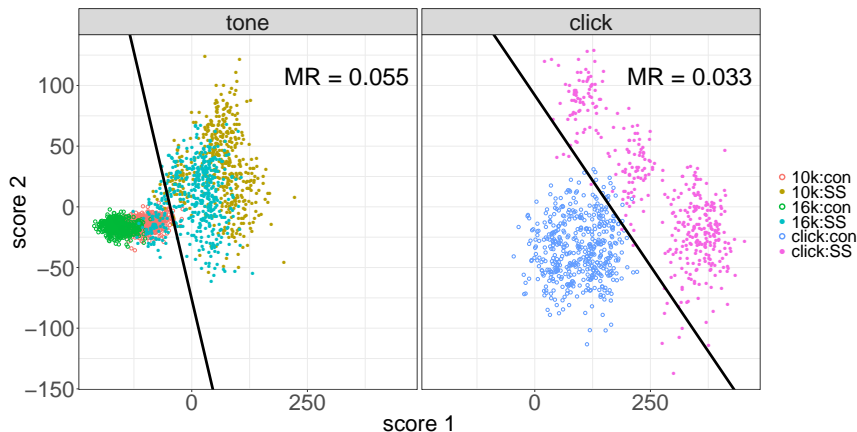
# Auditory EPI



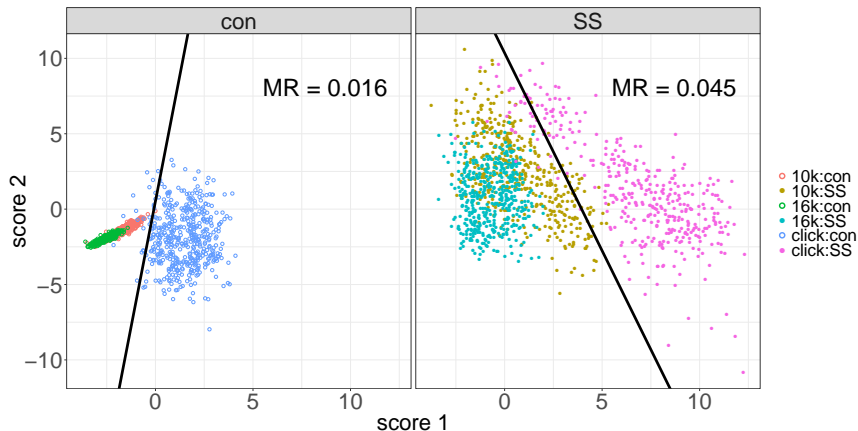
ML reprezentanti pro každou skupinu



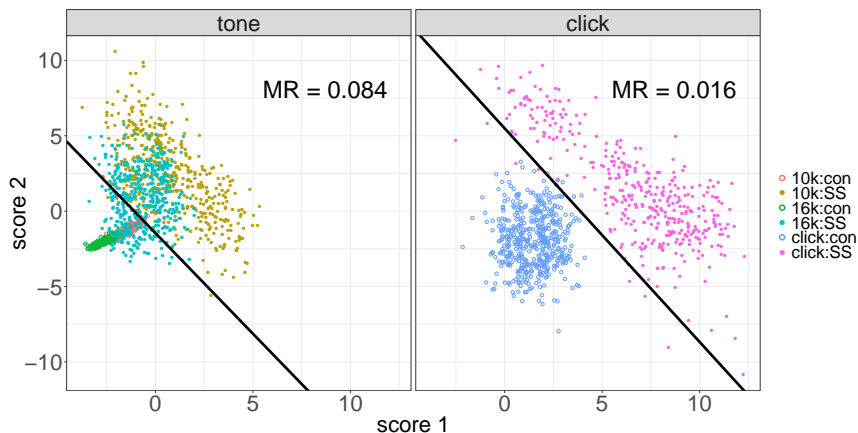
Koeficienty 1. a 2. komponenty pro funkce



Koeficienty 1. a 2. komponenty pro funkce



Koeficienty 1. a 2. komponenty pro 1. derivace



Koeficienty 1. a 2. komponenty pro 1. derivace



- Ferraty, F., Vieu, P., 2006. *Nonparametric functional data analysis: theory and practice*. Springer Science & Business Media.
- Kolářček, J., Pokora, O., Kuruczová, D., Chiu T.-W., 2018. *Benefits of functional PCA in the analysis of single-trial auditory evoked potentials*. Computational Statistics, accepted.
- Pokora, O., Kolářček, J., Chiu, T.-W., Qiu, W., 2017. *Functional data analysis of single-trial auditory evoked potentials recorded in the awake rat*. Biosystems 161, 67–75.
- Ramsay, J. O., Silverman, B. W., 2005. *Functional data analysis*, 2nd Edition. Springer, New York.
- Ramsay, J. O., Silverman, B. W., 2007. *Applied functional data analysis: methods and case studies*. Springer.
- Ramsay, J. O., Wickham, H., Graves, S., Hooker, G., 2017. fda: Functional Data Analysis. R package version 2.4.7.  
<https://CRAN.R-project.org/package=fda>