

v čem spočívá síla jazyka R?

Petr Štěpánek

E-mail: petr.stepanek.2@cvut.cz

LinuxDays 2018

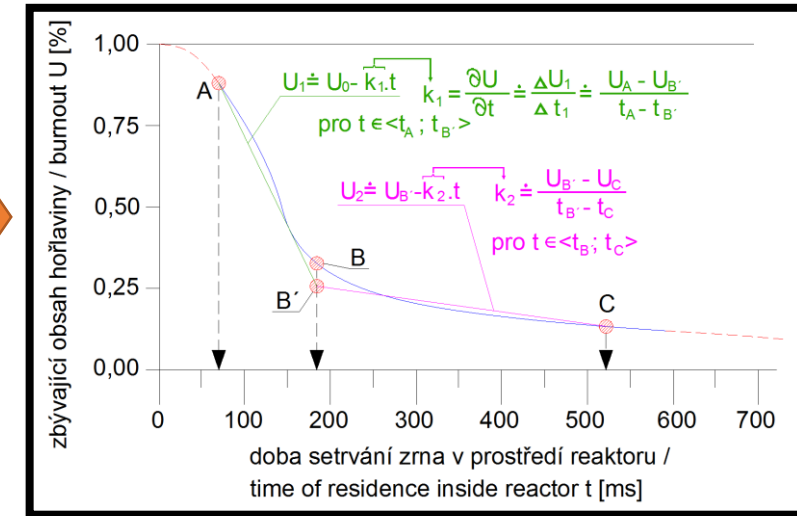
2018/10/07

Možnosti zpracování dat

```
Setup_20120119_122939.csv - Notepad
[Table No.1]
...Cal Kit,Compen Kit
...Open,Short,Load,Open,Short
Freq [Hz],Ave,OSC [dmm],G [S],C [F],R [ohm],L [H],Rs [ohm],Ls [H],G [S],C [F],R [ohm],L [H]
1E+006,1,-13,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,1,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,2,0,8,2E-014,0,0,50,0,0,0,0,0
3.5E+006,10,-10,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,3,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,4,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,5,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,6,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,7,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,7,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,8,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,9,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-14,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-14,0,8,2E-014,0,0,50,0,0,0,0,0
Meas Time,Mode 3,,R_DC,,Offset Delay,,R_DC
...Open:G [S],0,Open [sec],0,Open:G [S],0
...Short:R [ohm],0,Short [sec],0,Short:R [ohm],0
...Load:R [ohm],50,Load [sec],0
[Table No.2]
...Cal Kit,Compen Kit
...Open,Short,Load,Open,Short
Freq [Hz],Ave,OSC [dmm],G [S],C [F],R [ohm],L [H],Rs [ohm],Ls [H],G [S],C [F],R [ohm],L [H]
1E+006,5,-13,0,8,2E-014,0,0,50,0,0,0,0,0
2E+006,5,-13,1,0,8,2E-014,0,0,50,0,0,0,0,0
3E+006,5,-13,2,0,8,2E-014,0,0,50,0,0,0,0,0
```



**ANALÝZA
DAT**



V závislosti na požadavcích:

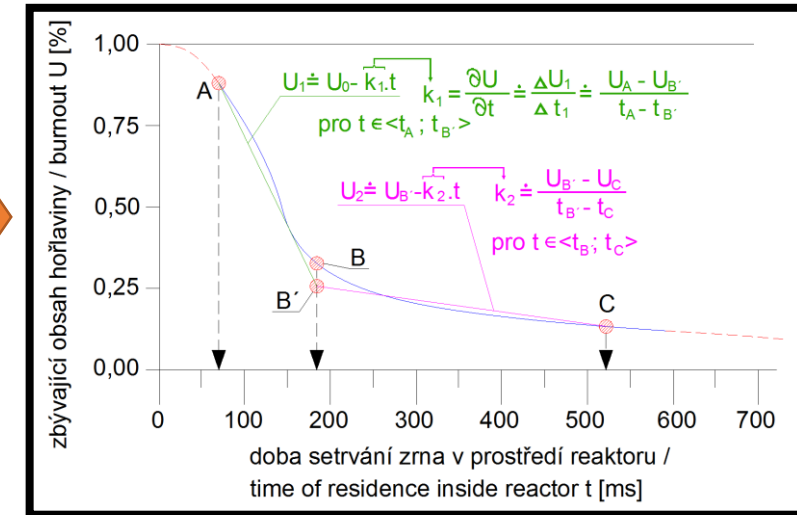
- MS Excel v kombinaci s VBA, Matlab, Mathematica (preferováno v rámci výuky na FS ČVUT),
- Python

Možnosti zpracování dat

```
Setup_20120119_122939.csv - Notepad
[Table No.1]
...Cal Kit,Compen Kit
...Open,Short,Load,Open,Short
Freq [Hz],Ave,OSC [dmm],G [S],C [F],R [ohm],L [H],Rs [ohm],Ls [H],G [S],C [F],R [ohm],L [H]
1E+006,1,-13,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,1,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,2,0,8,2E-014,0,0,50,0,0,0,0,0
3,5E+006,10,-10,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,3,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,4,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,5,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,6,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,7,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,7,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,8,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-13,9,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-14,0,8,2E-014,0,0,50,0,0,0,0,0
1E+006,1,-14,0,8,2E-014,0,0,50,0,0,0,0,0
Meas Time,Mode 3,,R_DC,,Offset Delay,,R_DC
...Open:G [S],0,Open [sec],0,Open:G [S],0
...Short:R [ohm],0,Short [sec],0,Short:R [ohm],0
...Load:R [ohm],50,Load [sec],0
[Table No.2]
...Cal Kit,Compen Kit
...Open,Short,Load,Open,Short
Freq [Hz],Ave,OSC [dmm],G [S],C [F],R [ohm],L [H],Rs [ohm],Ls [H],G [S],C [F],R [ohm],L [H]
1E+006,5,-13,0,8,2E-014,0,0,50,0,0,0,0,0
2E+006,5,-13,1,0,8,2E-014,0,0,50,0,0,0,0,0
3E+006,5,-13,2,0,8,2E-014,0,0,50,0,0,0,0,0
```



**ANALÝZA
DAT**



V závislosti na požadavcích:

- MS Excel v kombinaci s VBA, Matlab, Mathematica (preferováno v rámci výuky na FS ČVUT),
- Python,
- **R.**

Co je R?

- R = interpretovaný, multiplatformní OOP jazyk, šířený pod licencí GNU General Public Licence
- Po instalaci základního R-core obsahuje knihovny všech základních nástrojů pro statistické hodnocení dat „jednoduššího charakterů“ a grafické nástroje pro prezentaci výsledků
- Má početnou komunitu aktivních uživatelů => uživatelé vytváří a sdílí svoje vlastní knihovny metod

Co je R?

- R = interpretovaný, multiplatformní OOP jazyk, šířený pod licencí GNU General Public Licence
- Po instalaci základního R-core obsahuje knihovny všech základních nástrojů pro statistické hodnocení dat „jednoduššího charakterů“ a grafické nástroje pro prezentaci výsledků
- Má **početnou komunitu aktivních uživatelů** => uživatelé vytváří a sdílí svoje vlastní knihovny metod

Zdroj nejčastějších chyb

Defaultně není robustní. V případě, že např. metodě napsané pro vektor dáte jako parametr časovou řadu dojde k situaci:

- A) zobrazí exception nebo error, nebo
- B) metoda „nějak proběhne“ a vy na první pohled nepoznáte, že výsledek je chybný.

Zdroj nejčastějších chyb

Defaultně není robustní. V případě, že např. metodě napsané pro vektor dáte jako parametr časovou řadu dojde k situaci:

- A) zobrazí exception nebo error, nebo
- B) metoda „nějak proběhne“ a vy na první pohled nepoznáte, že výsledek je chybný.

Pro zvýšení robustnosti kódu je vhodné:

- číst manuály metod – zjistit zda lze danou metodu použít pro zamýšlený datový typ,
- ačkoliv to R nevyžaduje, tak definovat datové typy,
- psát kód v try-catch blocích.

Výhody RStudio

- R lze spustit v terminálu,
- nebo v GUI – v současné době nejpopulárnější je **RStudio**,
- RStudio je multiplatformní – ovládací prvky vypadají identicky pod Windows jako pod Linuxem.

Výhody RStudio

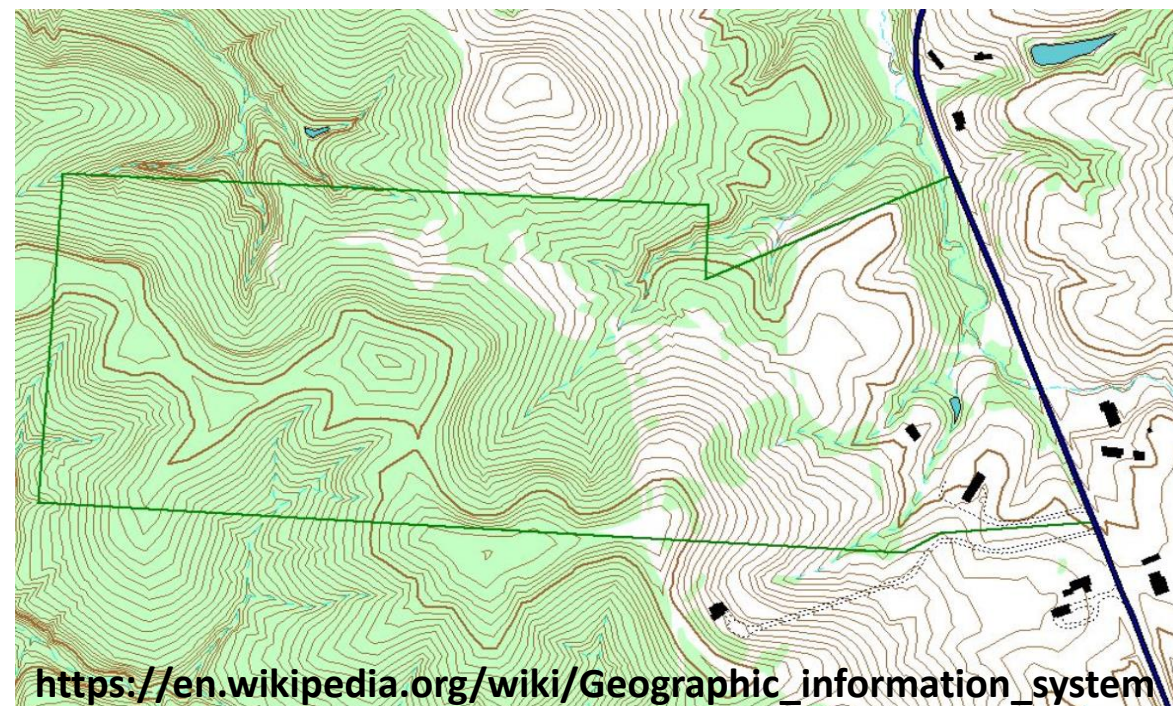
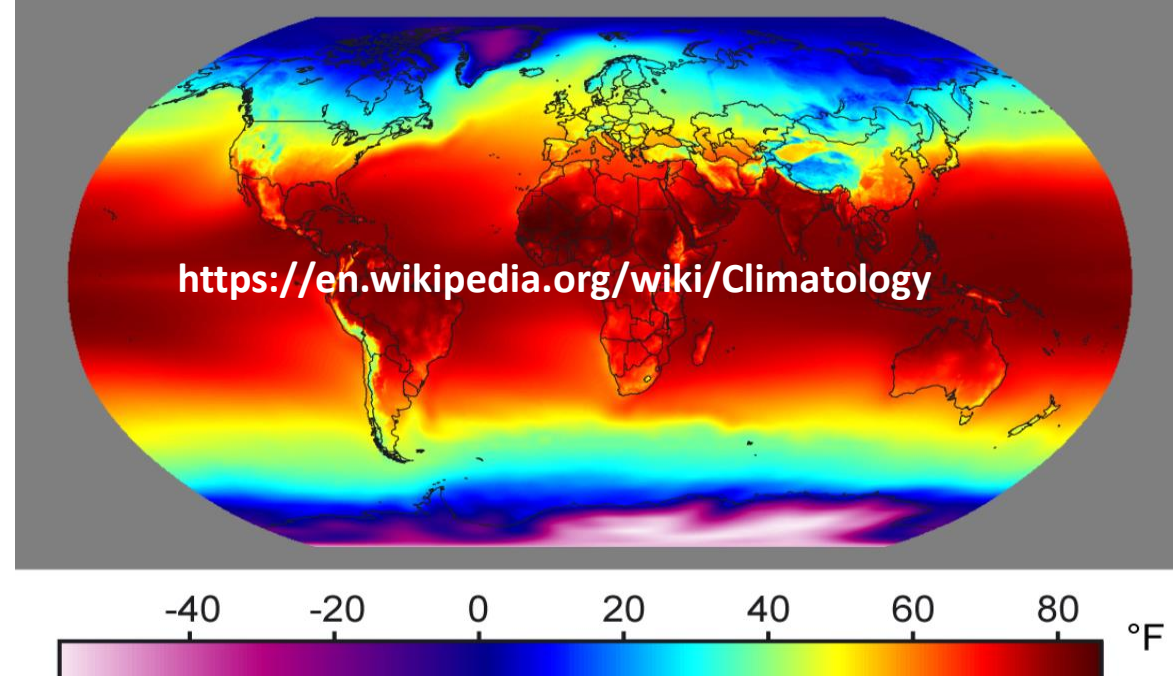
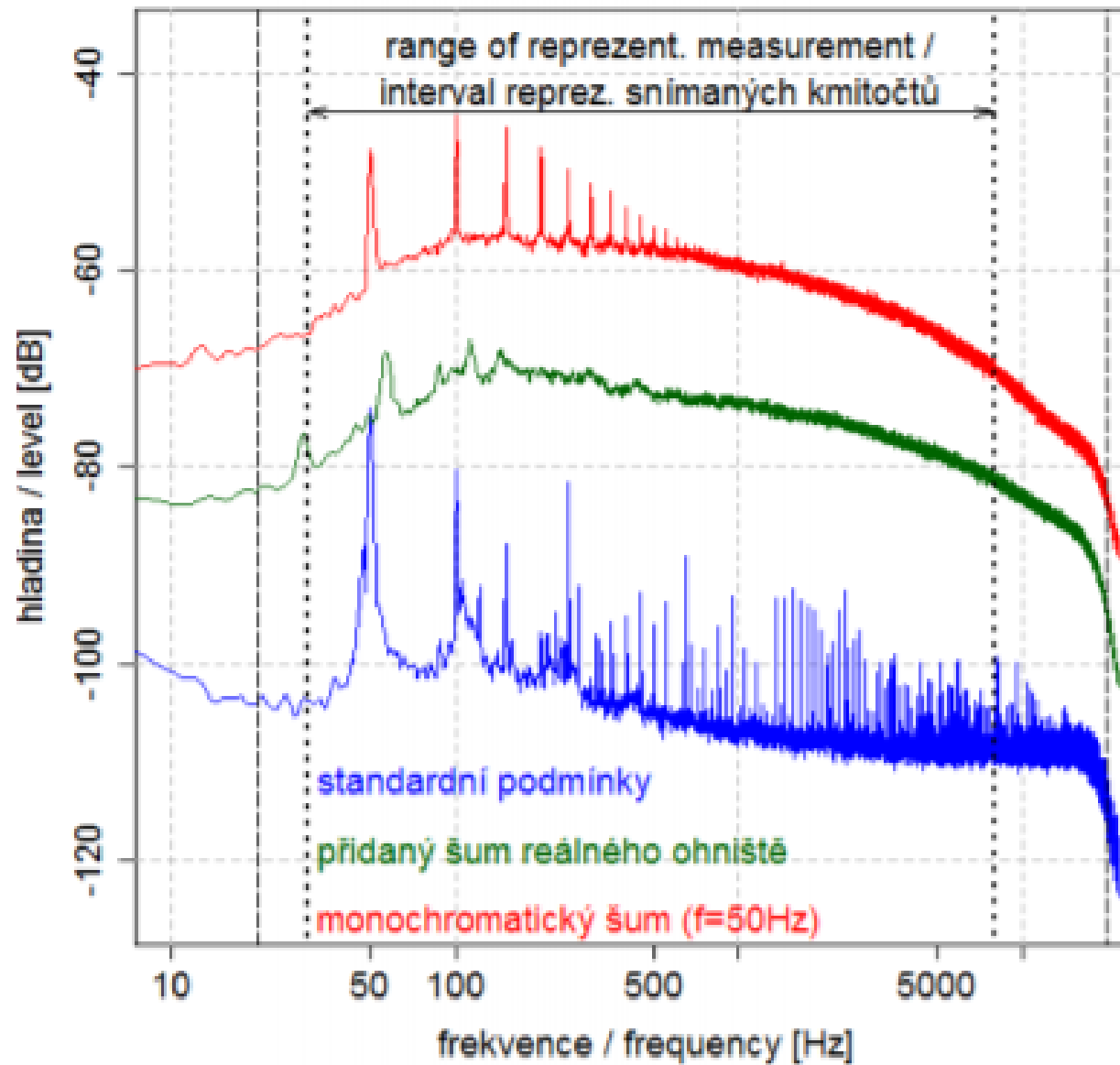
Vstup dat:

- snazší práce se soubory mnoha formátů (CSV, XLS, Rdata, RDS, RDA..)
- pomocí driverů DBI, ODBC.. lze připojit k různým RDBMS (**Postgres**, **MySQL**, Oracle, MS SQL Server, SQLite), detaily viz: <https://db.rstudio.com/odbc/>

Výhody RStudio

Výstup dat:

- pomocí různých nástrojů: **RMarkdown, Sweave, Knitr..** implementovatelných do RStudio lze získat výstup ve formě TEX, HTML, PDF, DOC..
- lze vytvořit webovou aplikaci (**Shiny app**) – vyžaduje server na němž je nainstalované R. Pro méně náročné appky lze využít (<https://www.shinyapps.io/>)



Jednoduchá ukázka webové R aplikace

(Shiny app)

<https://petrstepanek.shinyapps.io/dvojka/>

Možnosti využití R

Průmyslové linky:

- diagnostika provozu na základě měření, tvorba reportů pro management..

Energetika:

- kontrola „čerpání životnosti“ parovodů – podklad pro M&R,
- prokazování nežádoucích provozních stavů způsobených nekázní obsluhy zařízení,
- prokazování nedodržení parametrů provozu a hledání příčin.

Další možné aplikace:

- pojišťovnictví, bankovníctví, technika zařízení budov..
- přírodní vědy: klimatologie, meteorologie, výskyt hmyzu (např. kůrovců)..

Díky za pozornost.

E-mail: petr.stepanek.2@cvut.cz

FAQ

Q: Podporuje R multi-thread?

Q: Je v rámci programu napsaném v R možné zavolat metodu napsanou v jiném jazyce?

Q: Má RStudio implementovaný nástroj pro správu verzí?