

# Zpracování “Japan Radiation Open Data”

Ondřej Tichý

November 11, 2013

## 1 Získání a obsah dat

K havárii elektrárny Fukushima I došlo dne 11. března 2011 po zasažení Japonska zemětřesením a vlnou tsunami. Od té doby se zpracování dat z měřících stanic v Japonsku stal důležitý úkol. Data, o kterých pojednává tato zpráva, jsou získána z 218ti stanic SPEEDI (“System for Prediction of Environment Emergency Dose Information”) rozmístěných po celém Japonsku, viz Obrázek 1. Aktuální prohlížení dat je možné zde:

<http://www.bousai.ne.jp/eng/index.html>

Stahování aktualizovaných dat je možné ze stránky Mariana Steinbacha z následující stránky:

<http://www.sendung.de/japan-radiation-open-data/>

Odsud je možno stahovat taktéž seznam jednotlivých stanic s přesnými souřadnicemi a dalšími podrobnostmi. Data z měření jsou získávána každých 10 minut a to od 1.3.2011 do současnosti (s krátkým výpadkem 10.-12.8.2011, což je ale výpadek ve sběru dat, nikoliv v jejich měření). Jednotlivé řádky obsahují následující informace:

**datetime\_utc** čas v UTC (Universal Time Coordinated)

**station\_id** identifikační číslo jednotlivých stanic souhlasící s identifikačním číslem v seznamu stanic

**sa** měřená radiační dávka v nGy/h (občas obsahuje záporné hodnoty, ani autor sesbíraných dat si není jist, co znamenají, v následující zprávě je nulují při vykreslování)

**ra** informace o dešti v milimetrech (o záporných hodnotách platí to samé co pro parametr sa)

Jeden řádek dat tedy může vypadat např. takto:

```
1170000001 2011-03-01 00:00:00 46 0
```

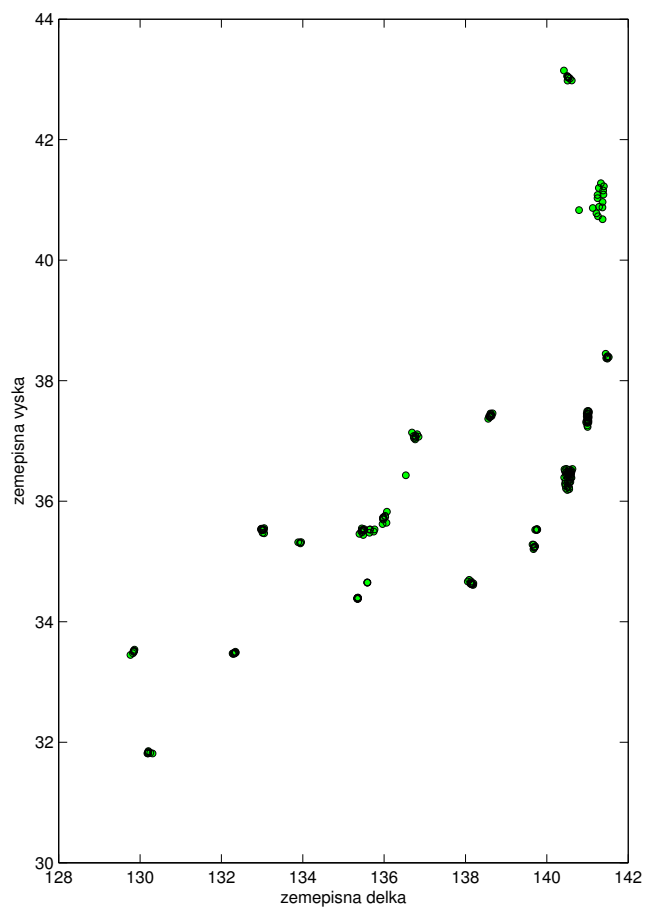


Figure 1: Rozložení měřících stanic.

## 2 Zpracování do podoby pro použití

Problém s používáním popsaných dat je v tom, že jejich otevření je pro běžný počítač nemožné, neboť jejich velikost již přesahuje GB textových dat. Do ruky se mi však dostal jednoduchý separátor pro rozdělení těchto dat do souborů pro jednotlivé dny. Tyto soubory lze již bez problémů načítat a pracovat s nimi (velikost cca 1,2 MB).

Pro efektivní práci je však možné data komprimovat do matic, jež mohou obsahovat kompletní informace z jedné stanice, tedy matice  $S \in \mathbf{R}^{n \times d}$  a  $R \in \mathbf{R}^{n \times d}$ , přičemž  $n$  je rozměr určující počet časových kroků za jeden den, tedy  $n = 6 \cdot 24 = 144$  a  $d$  označuje počet dnů měření. Tyto dvě matice mohou být vyrobeny pro každou stanici a práce s daty se značně zrychlí. Komprimace je bezztrátová, protože pozice čísla v matici určuje jednoznačně den i čas, kdy měření vzniklo. Celková velikost takto extrahovaných dat pro všechny stanice je cca 10 MB.

### 2.1 Zpracování v Matlabu 2008a

```
eval(['M = importdata('' cesta_k_csv 'station_data.' dat_str '.csv'');']);
% zpracovani nacteneho data M
% M.data(:,1) - sa
% M.data(:,2) - ra
% M.textdata(1,:) - popisek
% M.textdata(:,1) - station_id
% M.textdata(:,2) - datum a cas po deseti minutach
n = size(M.textdata,1);

for i = 2:n % prvni radek je popisek
    st_id = str2double(cell2mat(M.textdata(i,1)));
    cas = cell2mat(M.textdata(i,2));
    min = cas(15:16);
    hod = cas(12:13);
    sa = M.data(i-1,1);
    ra = M.data(i-1,2);
    transf_cas = str2double(hod)*6 + str2double(min)/10 + 1;
    D_sa(transf_cas,find(stations == st_id)) = sa;
    D_ra(transf_cas,find(stations == st_id)) = ra;
end
% ulozeni dat do jednotlivych stanic
for i = 1:size(stations,1)
    eval(['load ' cesta_ke_stanicim 'd_' num2str(stations(i)) '.mat']);
    % dostanu S a R ze souboru pro danou matici
    S(:,den_no) = D_sa(:,i);
    R(:,den_no) = D_ra(:,i);
    eval(['save ' cesta_ke_stanicim 'd_' num2str(stations(i)) '.mat S R']);
end
```

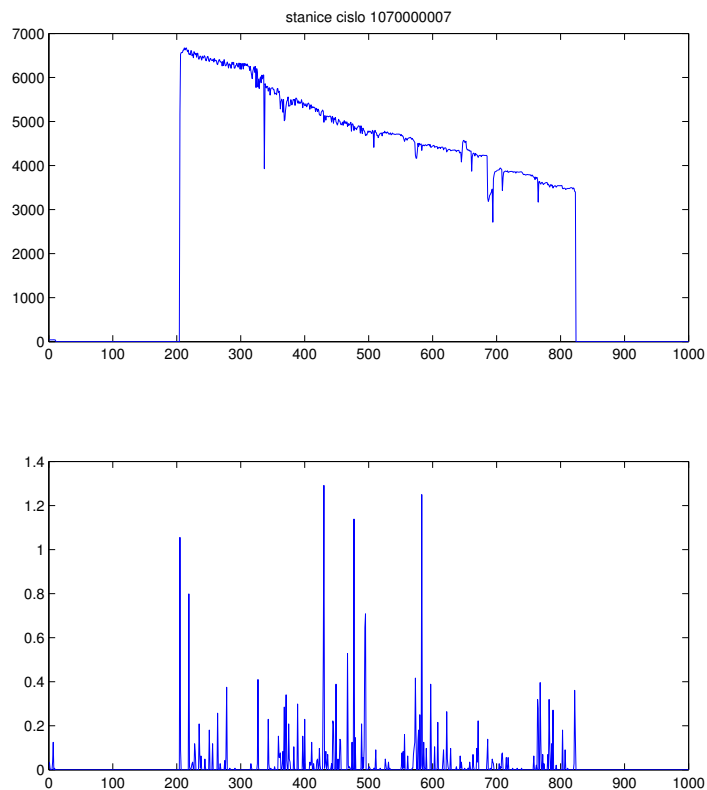


Figure 2: Ukázka grafů pro vybranou stanici po dnech.

### 3 Ukázka výsledků

Se získanými maticemi pro jednotlivé stanice lze pak velice rychle a efektivně pracovat. Pro ukázkou uvedeme grafy z měřicí stanice 1070000007. První graf zobrazuje průměrnou denní radiační dávku, tedy zprůměrovanou aktivitu z jednotlivých měření z daného dne s vynulovanými zápornými hodnotami. Druhý graf pak ukazuje celkové denní srážky v milimetrech. Stojí za povšimnutí, že výpadek měřicí stanice (z prefektury Fukushima) nastal 11tý den měření (počítáno od 1.3.2011) a měření bylo obnoveno opět 205tý den. Obnovení měření se samozřejmě může lišit stanici od stanice.