GRAPHICAL USER INTERFACE FOR ENVIRONMENTAL SIGNAL PROCESSING

M. Slavík, M. Mudrová and A. Procházka

Prague University of Chemical Technology,
Department of Computing and Control Engineering

Abstract

The air pollution measuring systems produce large quantities of data every day allowing to monitor selected pollutants and to obtain cumulative information about evolution of pollution sources. The paper presents selected methods of statistical signal processing, digital filtering of its noise components and analysis of evolution of pollution. Processing of measured pollution was implemented in Matlab graphical user interface (GUI) allowing simple and fast data processing.

1 Introduction

History of air pollution protection began in Czechoslovakia in 1967, under the leadership of Mr. Dr. B. Böhm, as a part of the Hydrometerological institute (HMU). In 1970, a new department of emissions was established. Special attention has been always payed to monitoring of air pollution. In early 70s, the measurements were carried out using discontinual manual devices. In 1971, a new information system for processing of air pollution data was started up. At present, the core of the air pollution measuring system is constituted of about 100 Automatic Immision Monitoring stations (AIMs), which measure the content of NO, NO_2, SO_2 and dust particles in the air with the sampling period of 30 minutes. 33 AIMs measure additionally content of ozone and carbon monoooxide [1]. All these monitoring stations located within the Czech Republic are precisely defined by their longitude and latitude.

Real data processing consists of several steps discussed in the paper.

The first step includes the elimination of substantial measurement errors and their substitution by values interpolated from their neighbours. This problem is closely related to two types of error sources including:
(i) unmeasured values caused by the time limited defect of the measuring system
(ii) fatal errors.

The second step includes the basic analysis of signals observed using the fast Fourier transform (FFT) for estimation of periodicities in given values.

The third step of signal processing consists, in this case, of signal filtering using the low-pass finite impulse response (FIR) filters. This filter excludes signals with higher frequencies. As a result of the FIR filter application on the data smoother time series were obtained.

2 Mathematical Methods

Presented GUI provides selected mathematical methods of signal processing including interpolation, Fourier analysis and signal filtering. Methods of linear interpolation are used for compensation of missing data in the given real series $x(n)$ where $n = 0, 1, ..N - 1$. Algorithms of the fast Fourier transform [5] implemented in Matlab as 'fft' command [2] have been used for signal analysis to obtain the signal amplitude spectrum $|X(k)|$ where

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{i2\pi nk}{N}} \quad (1)$$

The sampling period is assumed to be normalized to $\Delta t = 1$.

The lowpass filters have been implemented to simple noise rejection. There are used Finite Impluse Response (FIR) Filters. It is possible obtain their impulse characteristic $h(k)$ where $k = 0, 1, ...M - 1$ by 'fir1' Matlab command with cutting frequency ω_c. Resulting signal $y(n)$ can be constituted as a convolution of original signal and selected filter

$$y(n) = \sum_{k=0}^{M-1} h(k)x(n-k) \tag{2}$$

3 Specification of Data Processed

All data kindly provided by Czech Hydrometeorological Institute (CHMI) were measured at 108 stations every 30 minutes. The determination of content of dust particles is carried out by the radiometric method based on absorption of β-rays. A beam of radiation is split into two beams. The first goes trough a clean filtration material and the second through another one covered with dust particles. Concentration of dust particles is calculated from the difference of luminous intensity of beams.

Data are available from January till December 2000. Number of data from each station is 48·366 (number of measurements per day · number of days).

3.1 Import of the Data Into the Matlab Environment

Measured data stored in the DBF format can be imported into the Matlab environment using MS Excel. The first column of data contains identification numbers of AIMs, in the second and the third columns there are latitude and longitude respectively. Next three columns contain the name of the station, the operator of the station and the region of its placement. The seventh column informs about data measured (PM10 means dust particles up to 10 μm). The next three columns contain the year, the month and the day of the measurements respectively. Following columns contain data measured from midnight till 23:30 with the period of 30 minutes with data sorted according to stations and then according to dates. Export of the given values from MS Excel to Matlab environment is then achieved via the Dynamic Data Exchange (DDE) channel. Data transfer from EXCEL to MATLAB can then be achieved by the following commands:

```
chan=ddeinit('excel','032000.xls');
souradnice=ddereq(chan,'r2c2:r3349c3');
```

The first command initiates the DDE channel to Excel file 032000.xls containing data observed in March. The second command forms the request for transfer of locations of AIM stations which are in the second and the third column. After the transfer of all other values the variable PM10 is defined having the following structure:

```
PM10 =
     identifikatory: [108x1 double]
              jmena: [108x15 char]
         souradnice: [108x2 double]
               data: [1x12 struct]
              verse: 2
    datum_vytvoreni: '27-Aug-2001'
            typ_dat: {'Zdrojova data bez uprav'}
           poznamky: ''
```

Variable PM10.identifikatory shows IDs of the measuring stations, while variable PM10.jmena shows names of all stations, variable PM10.souradnice contains information about longitude and latitude of measuring stations and PM10.data contains for each month 3 dimensional matrix with dimensions representing the number of stations, number of measured values per day and number of days in month. Variables PM10.verse, PM10.datum_vytvoreni, PM10.typ_dat and PM10.poznamky are used as an informative part of structured variable PM10.

3.2 Data Analysis and Preprocessing

Owing to errors of the measuring system some values are missing. *Unmeasured values* are represented by values of -1. To allow simple statistical signal processing it is necessary to find these values at first. To achieve this goal a new variable with information about excluded values (either unmeasured or measured but with substantial errors) was created. This row vector named "info_data" points by values 1 to excluded values.

To find *measured values with substantial errors* the value of standard deviation was used. After the subtraction of the signal mean, all values with their absolute values grater then 3 standard deviations were considered as error observations using the following commands:

```
prumer = mean(data(1,find(info_data~=1)));
data(1, find(info_data==1))=prumer;
data = data-prumer;
odchylka = std(data(1,:));
info_data(1,find(data>3.*odchylka | data<-3.*odchylka))=1;
```

Assuming that all data are in a variable "data" the average from data without unmeasured values is computed. Then unmeasured data are substituted by average values mentioned above. This average value is then subtracted from each value of data measured. In this way the vector of data with average of 0 is obtained. In the fourth line standard deviation is computed and in the last line all data with substantial errors are marked.

Values with substantial errors can be then substituted. *Linear interpolation of excluded data* forming the simplest option is described by the following program segment:

```
pocet_dat = length(data);
for i = 1:pocet_dat
  if info_data(1,i)==1&(i~=1&i~=pocet_dat)
    for k =i+1:pocet_dat
      if info_data(1,k)~=1
        data(1,i)=(data(1,i-1)+(data(1,k)-data(1,i-1))/(k-(i-1)));
        break
      end
    end
  end
end
```

When the excluded value is found (line 3) then the search of the first not excluded value begins (line 4) to apply interpolation (line 6).

After the preprocessing stage, it is possible to analyze *spectral components using the Fourier transform.* Resulting plot presents the dependency of the absolute value of FT on frequency (amplitude spectrum) pointing to the dominant periodicity of one day.

3.3 Exclusion of High-Frequency Components

Exclusion of high-frequency signal components was carried out by the Finite impulse response filters. Matlab's Signal processing toolbox allows a simple evaluation of FIR filter coefficients. Computing of filter coefficients and application of filter on data shows following program segment:

```
 coefficients = fir1(number_of_coefficients, cutoff_freq);
 data = filter(coefficients,1,data);
```

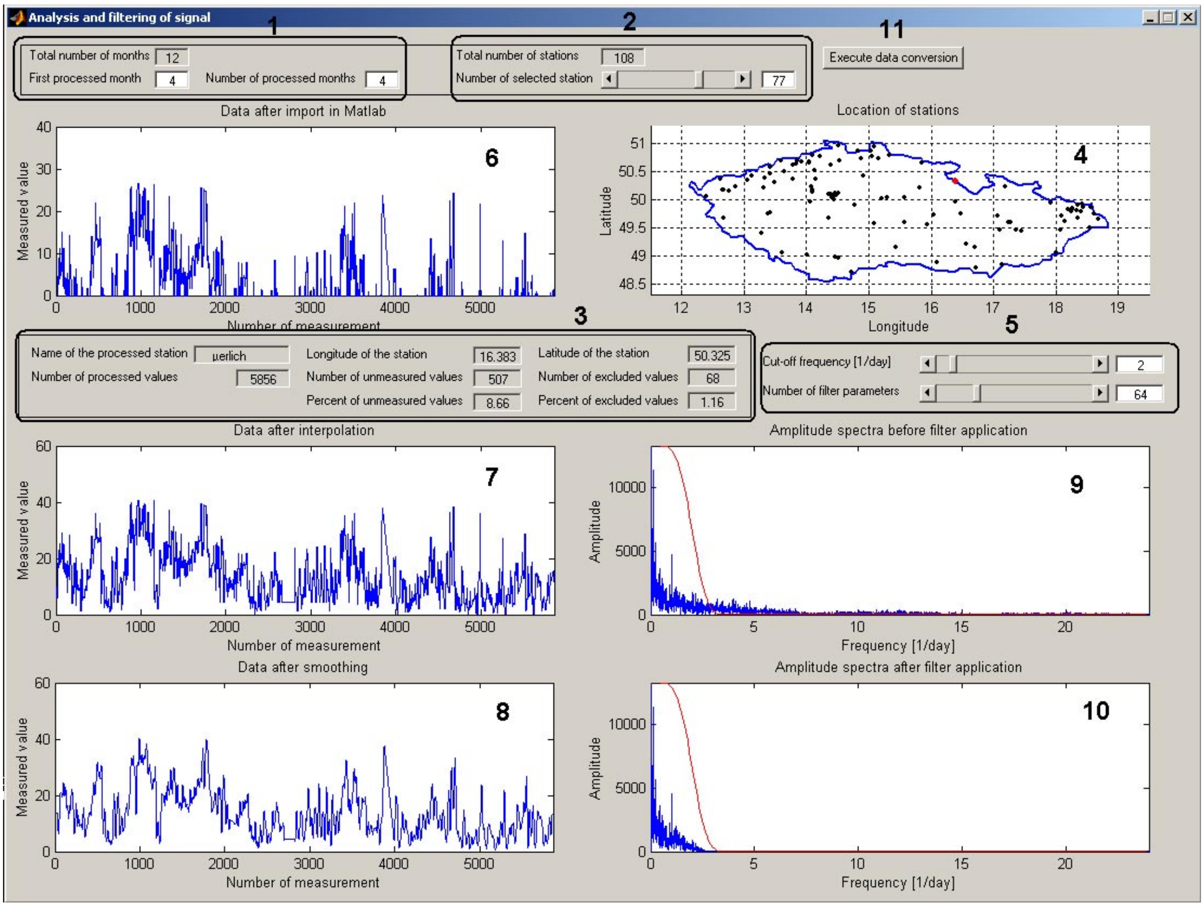
In the first line selected *number_of_coefficients* of the FIR filter with the selected cut-off frequency are evaluated defining its frequency response. For filtering of processed data FIR filters with cut-off frequencies in range $\langle 0/24, 24/24 \rangle$ 1/day could be used [4]. In the second line the FIR filter is applied.

4 Graphical User Interface in Processing of Air Pollution Data

Graphical user interface (simply GUI) in Matlab release 12 is slightly different from that one in previous version. GUI consists of two files - *fig* file that contains information about GUI appearance and *m* file that contains code necessary for GUI function. The *m* file consists of three parts: one launches GUI, the second one that launches callbacks and the callbacks forms the third part. Callback is a function that is launched when the user change state or value of user interface control object. There are two user interface control objects in Matlab, that cannot have callback. These are static text and frame.

Layout editor helps in creation of GUI. With the aid of this editor all user interface controls and axes can be placed. Layout editor also creates first two parts of m file and all function headers according to name of user interface control objects that have callback. Creating of GUI is quite simple with this aid.

All steps of processing in chapter "Data processing" except data import in Matlab environment were included in the GUI. Because of proper GUI function callbacks for all UI control objects were added. The GUI is showed in Figure 1.



Analysis and filtering of signal

Total number of months 12
First processed month 4
Number of processed months 4

Total number of stations 108
Number of selected station 77

Execute data conversion

Name of the processed station µerlich
Number of processed values 5856
Longitude of the station 16.383
Number of unmeasured values 507
Percent of unmeasured values 8.66
Latitude of the station 50.325
Number of excluded values 68
Percent of excluded values 1.16

Cut-off frequency [1/day] 2
Number of filter parameters 64

Figure 1: GUI for air pollution data processing

The GUI showed in Figure 1 consists of several parts. User interface controls (in next part of the paper is used term UI control) labeled 1 informs about number of months that could be processed and allows the user to set up which month he wants to process. UI controls labeled 2 allows to select number of processed station. Informative frame labeled 3 informs about selected station and about number of processed and excluded values. Informative graph labeled 4 shows location of selected station. Set-up of FIR filter - cut-off frequency and number of filter coefficients are labeled 5. Procedure of process illustrate graphs 6-10.Graph labeled 6 contains original data. Graph 7 contains data after exclusion and interpolation and graph 8 contains data after filter application. Graphs 9 and 10 show frequency characteristics for data displayed on graphs 7 and 8 respectively. Graphs 9 and 10 also show frequency characteristics of FIR filter according UI controls labeled 5. Button labeled with number 11 executes data processing that results in new variable PM10.

As an example one of the callbacks is showed below. This callback is used for handle with slider for selection of station (labeled 2 on figure 1)

```
function varargout = Number_of_station_slider_Callback...
    (h, eventdata, handles, varargin)
set(handles.Number_of_station_slider,'Value',round(...
    get(handles.Number_of_station_slider,'Value')));
set(handles.Number_of_station,'String',num2str(round(...
    get(handles.Number_of_station_slider,'Value'))));
guidata(h,handles);
Info_about_station(h,[],guidata(h));
Data_processing(h,[],handles);
```

In the first row of function there is a value that was set-up by the slider rounded to integer. In the second row the edit box value for station selection is set and in the third row are all changed data stored in the variable handles. Finally, functions for updating of GUI for chosen station are executed.

5 Conclusion

FIR filtering is one of the common methods for digital signal Processing. These facts were used in the program for signal de-noising and for estimation of air pollution in given regions of the Czech Republic.

Further studies will be devoted to two dimensional data interpolation, detection of air pollution sources and prediction of concentration of air pollution in selected regions.

6 Acknowledgments

We express thanks to the Czech Hyrometeorolgical Institute which provided all processed data.

7 References

[1] J. Fiala, J. Ostatnická, and collective. *Znečištění ovzduší na území České republiky v roce 1998*. Český hydrometerologický ústav, Praha, first edition, 1999.

[2] The MathWorks Inc. *Matlab relase 12 program help*. The MathWorks Inc., 2001.

[3] J. W. Tukey J. W. Cooley. An Algorithm for the Machine Computation of the Complex Fourier Series. *Mathematics of Computation*, 19:297–301, April 1965.

[4] T. Krauss, L. Shure, and J. Little. *Signal processing toolbox*. The MathWorks Inc., Massatchussets, USA, third edition, 1994.

[5] V. Čížek. *Diskrétní Fourierova transformace a její použití*. SNTL, Praha, 1981.

Institute of Chemical Technology, Prague,
Department of Computing and Control Engineering,
Technická 1905, 166 28 Prague 6, Phone.: 00420-2-2435 4198, Fax: 00420-2-2435 5053,
E-mail: {Martin.Slavik, Martina.Mudrova, Ales.Prochazka}@vscht.cz