NETLAKE toolbox for the analysis of high-frequency data from lakes

Factsheet #5

High frequency data treatment and visualization with ECDSOFT and OnLineMonitor

Dario Omanović and Ivanka Pižeta

Objective

After repeatedly collecting a series of numerical x,y pairs of data, and before further treatment, it might be useful to visualize them, to check with "*an expert eye*" whether the data are in the expected frames and/or to decide about subsequent steps such as smoothing, background subtraction, and determination of signal heights and positions. 2D, or even better, 3D visualization with rotation feasibility could reveal useful information.

Getting rid of high-frequency noise that is superimposed on a signal of interest helps to recognize the useful parts of a signal. However, it is important to visually check that the signals are not distorted when smoothing, because peak features will be extracted from these smoothed series, and the so-called secondary curves will be constructed, and these will be checked against expected models (e.g. linear, exponential, etc.). So, visualization of smoothing effects while choosing algorithms and their parameters is necessary, and is welcome and appreciated in any step of data treatment.

The basic software tool is ElectroChemistry Data SOFTware (ECDSOFT) (Omanović and Branica 1998) designed for treatment of data obtained by the electrochemical method, but it is capable of accepting any set of signals that matches the required format. The software itself has been continuously upgraded and is open for further improvements on request. The subsequently developed software, sharing the main data treatment features of ECDSOFT, intended to automatically analyse sets of such data and present them in a near-real time domain is OnLineMonitor.

Specific application

ECDSOFT is designed specifically for the treatment of voltammetric signals, i.e. current vs. voltage peak-shaped curves (Figure 1). Such signals usually consist of only a few well-defined peaks (two peaks are shown in Figures 1a and 1c), but there could be a high-frequency noise superimposed on the signal and there could be a background level that is not related to the analyte expressed as a peak, and which preferably should be removed (e.g. subtracted). Still, any such set of curves originating from different methods (e.g. lake temperature vs. depth and time), could benefit from the software, as it is handy for visualization, smoothing, background

subtraction, and peak height and position determination, and can perform all this automatically once the parameters have been determined.

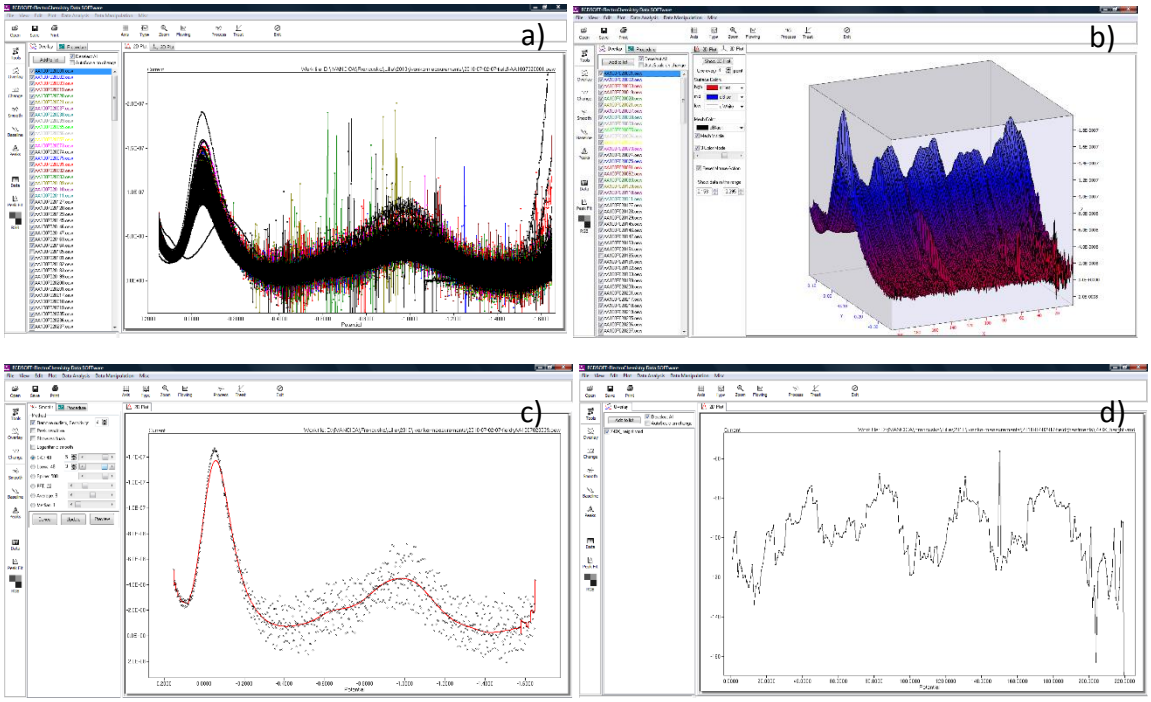

Figure 1. a) an example of raw data from voltammetric automatic measurements of dissolved oxygen in a small lake for several days at 1 hour intervals; b) 3D plot of the data from a) (only first peak on the left); c) S-G smoothing of one curve from a) – distortion of peak height is shown in red (with S-G: 40) (indicating that another parameter of S-G smoothing should be chosen, i.e. S-G: 18); d) result of automatic oxygen peak height determination for all series of signals from b) (showing day – night variations of oxygen concentration for five days).

OnLineMonitor software is designed to track specific folders in a defined time period (e.g. every 5 mins) for new files (e.g. voltammograms, spectra, etc.) and analyse them using predefined parameters for specific peak shaped signals. Up to 20 different parameters can be monitored. The software also allows simple correlation analysis of data (Figure 2).

Background

Both software packages are free of charge and could be downloaded at https://sites.google.com/site/daromasoft/ (ECDSOFT) or could be obtained on request (OnLineMonitor; the program is still in the early development stage and thus is not yet available for download over the web-page). Opening and presentation of files (of predefined format, see below) is as simple as any other document, while the use of data treatment routines needs knowledge on general signal characteristics. The user is responsible for organizing the data in repeatable sequences if a 3D presentation is required. Otherwise it can be any sequence of high-frequency data (e.g. time series of temperature, wind, pressure...).

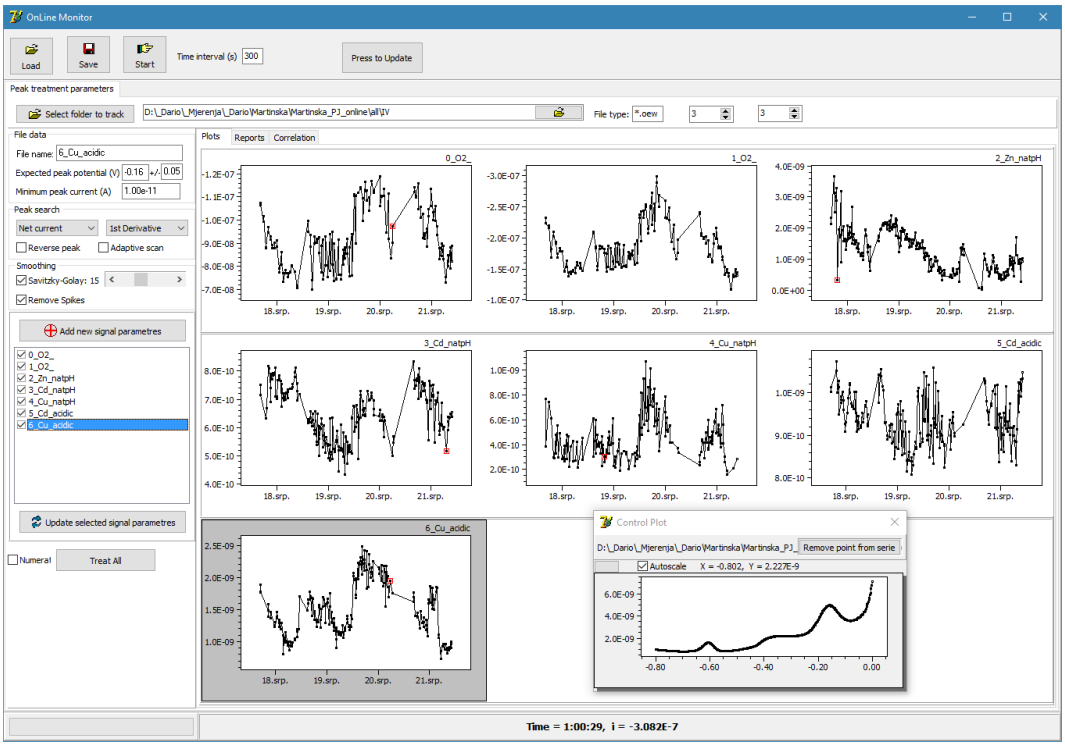

Figure 2. OnLineMonitor program showing automatic analysis and presentation of signal data (O_2, Zn, Cd, Cu) obtained during ~3.5 days of automated voltammetric measurements in the Krka River estuary, Croatia.

Type of data and requirements

Any standard laptop would do to install the software and to use it. The data should be stored in the format "x TAB y" or "x SPACE y" under its name in .vmd files (e.g., name.vmd) (Figure 3). Up to 500 files (XY datasets) could be loaded at the same time and graphically presented, as shown in Figure 1a. Any XY data-pair (e.g. selected in the Excel table) could be plotted in ECDSOFT by using "copy-paste" formalism. In this case, for further processing the dataset should be saved. Also, vice versa, all datasets that have been entered can be copied and transferred to another program (e.g., Excel) by copying all data to the clipboard.

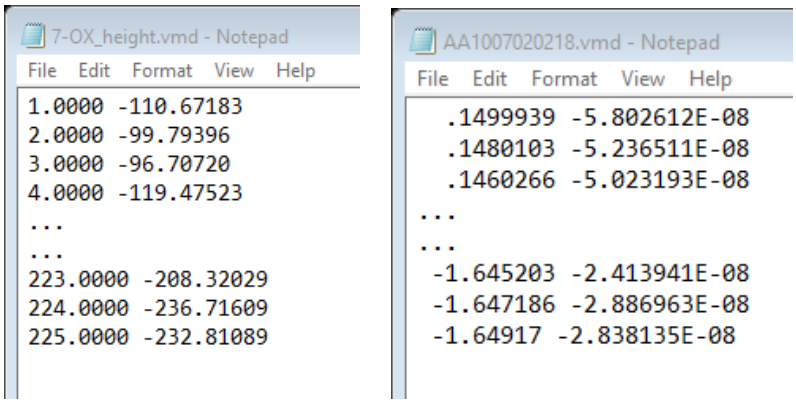

Figure 3. Examples of format files ready to be used by ECDSOFT.

Basic procedures

GENERAL

Once data are inserted they are displayed, as seen in Figure 1a. The name of each file is listed on the left side of the screen. Each file can be excluded or included by clicking the checkmark to the left of the file name. It is also possible to zoom-in on any part of the dataset. A 3D presentation is available by selecting the **3D Plot** box (in Figure 1b the 3D presentation of zoomed signals from Figure 1a is shown). The figure can easily be rotated in any direction to enable the best view by left-clicking on the plot and dragging it with the mouse.

SMOOTHING

For smoothing, select *Smooth* on the left toolbar. Choose the method of smoothing from *Savitzky-Golay (S-G), Loess, Spline, FFT, Average* or *Median.* The choice of the method would depend on the character of the signal and noise spectrum, while its effects are automatically refreshed and easily visually checked (the user is advised to get informed about the basic principles of the methods). Each method has a scrollbar to adjust for additional parameters. In Figure 1c, the S-G method was chosen and the scrollbar was adjusted to 40. The red line is a *Preview* of the results of smoothing. After pressing the *Update* button, a new smoothed curve is ready for further treatment. It can be saved under a different name. The smoothing could be repeated with different methods by pressing *Resume* on the *Plot/Resume* menu (F5).

PEAK FEATURES

For peak height and position determination of a whole set of curves, select *Data Management/Automatic processing* and adjust the parameters, selecting from many options. Generally, it is better to consider *peak feature*, because except for peak height, the signal can also be quantified by a peak area and/or by determination of a peak 1^{st}, 2^{nd} or 4^{th} derivative. Selected peak features are automatically determined and saved in a file that can be uploaded and viewed in the same program (Figure 1d) or in another program (e.g. Excel). The OnLineMonitor program (Figure 2) provides similar final analysis results as those presented in Figure 1d, but for up to 20 measuring parameters. It is advised to first decide about parameters of signal treatment in ECDSOFT, and then apply them in OnLineMonitor, where more parameters could be followed in the same time.

Pitfalls and tips

ECDSOFT and OnLineMonitor, primarily designed for voltammetric data sets, have many specific features for this type of signal, but could analyse “peak-shaped” signals of any kind. The software is full of useful details and is very handy once the user becomes familiar with it. An intuitive trial-and-error approach is advised when first approaching this software. ECDSOFT is more intuitive and intended for “In-depth” analysis and visualisation, while OnLineMonitor is more appropriate for tracking changes of selected parameters in (e.g. time) sequence.

Further reading

For examples on the use and usefulness of this software see:

Omanović, D., Branica, M. 1998. Automation of voltammetric measurements by polarographic analyser PAR 384B. *Croatica Chemica Acta* 71: 421-433.

Superville, P.J., Louis, Y., Billon, G., Prygiel, J., Omanović, D., Pižeta, I. 2011. An adaptable automatic trace metal monitoring system for on line measuring in natural waters. *Talanta* 87: 85-92.

Superville, P.J., Pižeta, I., Omanović, D., Billon, G. 2013. Identification and on-line monitoring of reduced sulphur species (RSS) by voltammetry in oxic waters. *Talanta* 112: 55-62.

Code

The ECDSOFT program can be downloaded from the following link (source code is not available):
https://sites.google.com/site/daromasoft/home/ecdsoft

Contact details

Ivanka Pižeta. Ruđer Bošković Institute, Zagreb, Croatia.
pizeta@irb.hr

Dario Omanović. Ruđer Bošković Institute (Zagreb, Croatia)
omanovic@irb.hr

Suggested citation

Omanović, D. and Pižeta, I. 2016. High Frequency data treatment and visualization with ECDSOFT and OnLineMonitor. In Obrador, B., Jones, I.D. and Jennings, E. (Eds.) *NETLAKE toolbox for the analysis of high-frequency data from lakes* (Factsheet 5). Technical report. NETLAKE COST Action ES1201. pp. 23-27. http://eprints.dkit.ie/id/eprint/536.

Acknowledgement

This factsheet is based upon work from the NETLAKE COST Action ES1201, supported by COST (European Cooperation in Science and Technology).