ANALYSIS OF BIOREMEDIATION RESPIROMETRIC DATA USING WAVELETS

Vila, M. Cristina; Fiúza, António M.
CIGAR (Geo-Environment and Resources Research Centre) – University of Porto, Engineering Faculty
Rua Roberto Frias, 4200-465 Porto, Portugal, e-mail: mvila@fe.up.pt; afiuza@fe.up.pt

ABSTRACT

The study of biodegradation using respirometry generates an enormous quantity of data, with several millions of registers for each variable. We have been treating this enormous amount of information using several mathematical techniques. The first step is always the filtration of the data in order to eliminate anomalies strange to the process, such as voltage breakages. The length of the data can be reduced using conventional statistical methodologies or by using wavelets or by combination of both. We have been applying wavelet analysis to signals generated by the respirometry of biodegradation with three different purposes: (i) as a method of data filtration or denoising that keeps the inner core structure of the information without aliasing; (ii) as an interpretation tool; (iii) to detect variation patterns at smaller scales. The synthesized signals can be subsequently used to create digital data-driven mathematical models, either single input-single output or multiple input-multiple output, using the tools of the system identification theory.

1. INTRODUCTION

Until recently soil respirometry has not been commonly used for studying the biodegradation of organic contaminants in soils, in opposition to its broad application for similar studies in aquatic environment. A soil respirometer measures continuously the concentrations in oxygen and carbon dioxide, sometimes other gases such as methane, and some environmental variables such as temperature, in the atmosphere in the vicinities of a soil under biodegradation in a reactor.

The respirometer we use in our research is model TR-RM8 Respirometer Multiplexer from Sable Systems International. The system, as it was initially conceived, consists of instruments that capture, direct and control a continuous airflow through a succession of interrelated devices, in order to analyse the composition of the air in the biodegradation reactor, and to communicate this information to a computer for storage and subsequent analysis. The air feed is carefully collected outside the laboratory in a non-polluted zone, it is scrubbed in a drierite column and it is pushed by a pump with a mass flow controller (to maintain a constant volume of air) to the rest of the system. It first enters one of the maximum eight possible soil sample chambers (reactors) selected in the multiplexer, where biodegradation is occurring, and goes through the condenser and another drierite column to lose eventual existent humidity. Then, the dried air enters into the carbon dioxide analyser, passes through a small column of ascarite to remove the CO2 content and is analysed in the oxygen probe, returning the air to the laboratory atmosphere. Although this hardware system includes an analogical/digital converter AD-201 with a Universal Interface in order to record and treat the data using software specially developed for respirometric studies, we connect each analyser directly to the computer using a multi-port serial interface. With this procedure, the data is directly colleted, avoiding two data transformations as well as the data aliasing operated by the original software. The computer receives two records per second from each analyser, which results in an intense amount of data. The soil samples subjected to biodegradation inside the reactors are periodically moisted and nutrients are added once or twice during each experiment.

2. SHORT INTRODUCTION TO WAVELET THEORY

Wavelets theory has its basis on the windowed Fourier transform, although the wave window is significantly different. Between 1960 and 1980, Morlet (seismologist) and Grossman (physician) developed some functions which constitute orthonormal basis and that can be scaled in order to be short in the high frequencies and long in the low frequencies however maintaining its form – to such functions they called wavelets (Mallat, 1998). A real-value function $\psi(t)$ defined over real axis is a wavelet if

1. The integral of $\psi(t)$ is zero: $\int \psi(t)dt = 0;$
2. The integral of $\psi^2(t)$ is unity (called "unit energy property").

Wavelets are then waves with an effective limited duration with a null average. If we compare wavelets with the sine function used in Fourier analysis, one radical difference is immediately perceptible: sines do not a limited duration and they extend from minus to plus infinity. Sines are soft and previewable while wavelets are irregular and asymmetric. As it is known, the Fourier analysis consists in transforming one signal in a set of sinusoidal waves with different

wavelengths. Similarly, the wavelet analysis consists in transforming one signal in a set of scaled and translated versions of the original (or mother) wavelet.
Starting from a mother-wavelet it is possible to cover all time-frequency domain through its successively dilatation and translation, which is mathematically translated by the following equation, where $\psi_{u,s}$ (t) is the mother wavelet, t is time, s is the scale and u is time delay or translation unit:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right)$$

The scaling function brings about notorious advantages, especially when we are interested in details located in the high frequencies and consequently masked by white noise. Scaling function is also the responsible for the large application of wavelets, in digital signal processing, audio processing, compression data and image processing. Wavelet analysis consists then in the decomposition of a signal in a hierarchic set of approximations and details. The hierarchic levels correspond, in the great majority of situations, to those existing in a dyadic scale (this is according to powers of 2). Another analogies with Fourier analysis can be stressed: the wavelet scale has equivalence in frequency and the spectrogram in the windowed Fourier transform has also its equivalent in wavelets transform, which is designed by scalogram (Graps, 1995).
In our research we have been applying wavelet analysis to three different situations;

(i) As a tool for denoising the original signals allowing them to be used in subsequent mathematical studies and simultaneously reducing the amount of data;
(ii) As a mean of detecting the main features of biodegradation through analysis of scalograms;
(iii) A method of detecting variation patterns at small scales;

3. WAVELETS AS A DENOISING TOOL FOR RESPIROMETRIC DATA

Respirometry of biodegradation generates an huge amount of data. As an example 21,7 days of respirometric analysis of bioremediation generates around 3.8 million registers, this only for the oxygen concentration. The same amount of registers are obtained for CO2 and for temperature. These values are normally submitted to a previous classic data filtering operation in order to eliminate data variations extrinsic to the process such as voltage fluctuations or even cut of power supply. The data can be decomposed using a selected wavelet such as the third order Daubechie wavelet. The next problem is to select the order for the decomposition. Two alternatives are possible: (i) a trial and error approach increasing the decomposition order from the second onwards until the residuals are constituted only by white noise or (ii) optimizing the decomposition using a minimizing criterion such as the sum of the logarithms of the entropy of the energy which is an additive function, as required.
As example Fig. 1 shows a comparison between an original signal and the synthesized signal using a Daubechie wavelet.

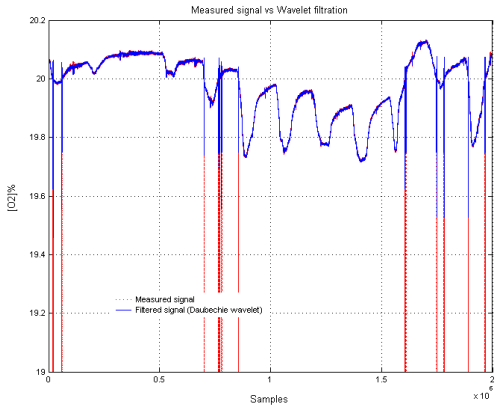

Fig 1 – Original and synthesized using a Daubechie wavelet

Figure 2 shows statistical information concerning the residuals, after a third order decomposition. As one can clearly see, it is constituted only by white noise.

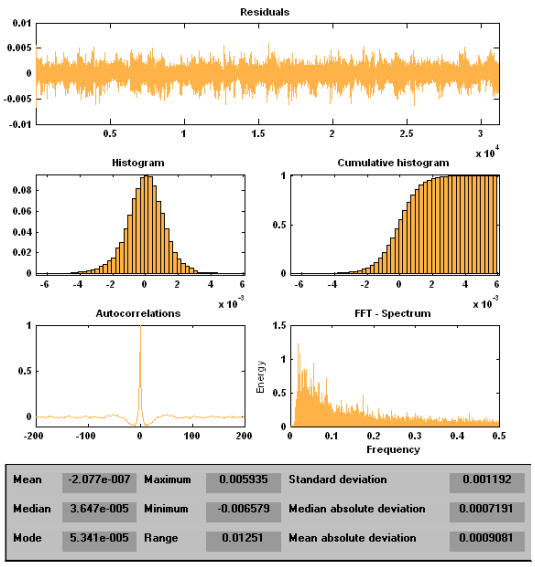

Fig. 2 – Residuals after a third order decomposition

Using this approach the original signal was condensed in a much shorter synthesized signal without loosing any of the main and secondary features of the data and without aliasing.

4. WAVELETS AS AN INTERPRETATION TOOL

Wavelets may be used as an interpretation tool. We will focus on the example shown in Fig.3. The first graph in the figure represents the evolution of oxygen concentration during bioremediation. One can notice the different periods: an initial unstable period corresponding to bacterial adaptation lasting for 10 days (lag period), an active period of biodegradation lasting from the 18th to the 28th day corresponding to the exponential phase of bacterial growth and a period of decreasing biological activity from the 28th day onwards (stationary and death phases). It is also possible to notice that during the 28th day there was an abnormal decrease in oxygen concentration, which was due to external reasons: a visiting day when the laboratory was open to the public. It is also evident that bacteria have a daily cycle of biological activity (the sinusoidal curves) induced by the daily temperature fluctuations.

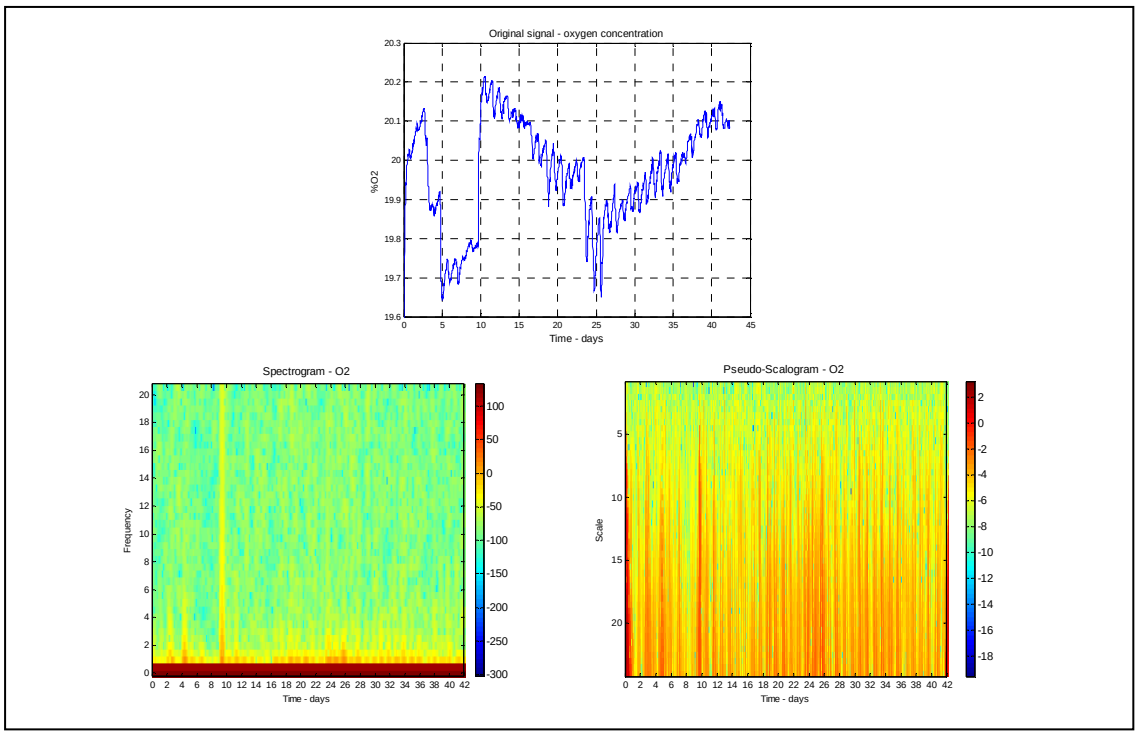

Fig. 3 – Data treatment using wavelets

These events can also be clearly noticed using the wavelets. If we look at the oxygen spectrogram we can see around the tenth day an energy spot spread for all the frequencies; on the 23rd day a concentration of frequencies may be seen; on the pseudo-scalogram the daily cycle of biological activity is evident.

5. DETECTION OF VARIATION PATTERNS AT SMALL SCALES USING WAVELETS

Wavelets also allow the detection of other phenomena that otherwise would remain unnoticed: the zooming of the details of the last selected order in different areas of the map allows verifying if the sinusoid period remains constant. As an example Fig. 4 shows a zoom in the fourth order details; a cyclic phenomenon with a period around 45 minutes is clearly seen although it was unnoticeable using all the other signal treatment techniques.

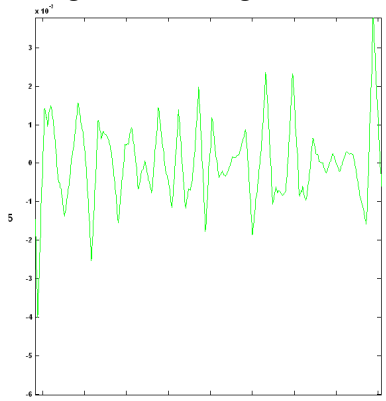

Fig. 4 – Zoom on the map of the 4th order details

6. BIODEGRADATION MODELS USING WAVELET SYNTHESIZED SIGNALS

The synthesized data can then be submitted to signal treatment using several analytical techniques such as directional circular statistics, time series analysis and system identification.

As a final example, the synthesized signals may be used to build data driven quantitative models that predict the time evolution of biodegradation as a function of input variables such as outside oxygen concentration and temperature. Normally we are using system identification models, specially the Box-Jenkins (Fiuza, Vila, 2005). Fig. 5 exemplifies one application of such methodology using a Box-Jenkins multiple input-single output (MISO) model.

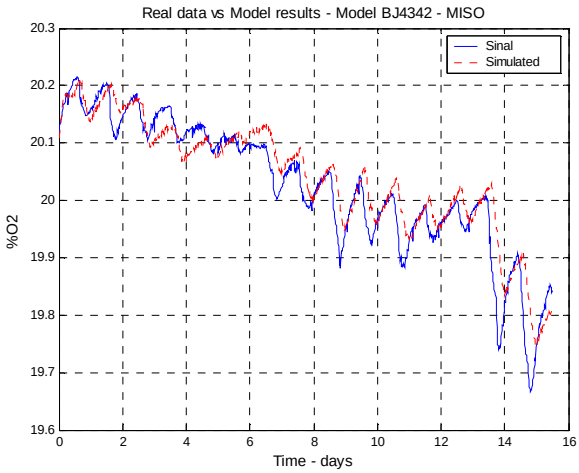

Fig. 5 – Application of a Box-Jenkins model to data synthesized by wavelets

REFERENCES

[1] Fiúza António, Vila Cristina, "An insight into soil remediation through respirometry", *Environment International*, 31 (2005), 179-183

[2] Fiúza, A., "Signal treatment applied to environmental biological systems", *Proceedings of 2005 Consoil*, Gent, Belgium .

[3] Graves, D. A., Lang, C. A., Leavitt, M. E., 1991, "Respirometric analysis of the biodegradation of organic contaminants in soil and water", *Applied Biochemistry and Biotechnology*, vol 28/29, 813-826.

[4] Mallat, S., 1999, *A wavelet tour of signal processing*, 2nd ed. , Academic Press.

[5] Misiti, M., Misiti, Y., Oppenheim, G., Poggi, J-M., 2000, *Wavelet Toolbox User's Guide*, The MathWorks.

[6] Ljung Lennart, 1987. *System Identification Theory for the User*, Prentice Hall.