Wavelet procedures applied in ecosystem models

Vladimir Balan and Simona Mihaela Bibic

Abstract. The Haar wavelet decomposition of the numerical solution for the Truscott-Brindley 3-dimensional SODE produces sets of coefficients which contain essential information on the local oscillating patterns of the population evolution. The segmentation of the initial sample finite series gives besides an increase of speed of the decomposition process ([3]), via the analysis of partial energies and Shannon entropy, information on the prevalence of wavelet-type evolution periodic behavior.

M.S.C. 2000: 92D40, 42C40, 65T20, 92C15, 92D25.
Key words: Haar wavelets, nonlinear dynamical systems, differential operators, entropy, numerical approximation.

1 Introduction

Consider the $\mathcal{C}^1$ vector field $X \in \mathcal{X}(D)$, $(D \subset \mathbb{R}^n)$ and let $x_0 \in \mathbb{R}^n$. Then the SODE Cauchy problem

$$\dot{x} = X(x), \;\; x(0) = x_0, \tag{1.1}$$

can have its numerical solution $\{x_k\}_{k=\overline{1,N}}$ approximated by the associated Haar wavelet associated truncated time-series. The resulting wavelet coefficients exhibit relevant features of the SODE (1.1), such as: fluctuations, trends and jumps, providing essential data for the analysis and classification of the SODE (e.g., as in [1, 9, 10]).

Except for periodic signals, the time-series are characterized by local changes and local properties, which recommend the use of wavelets ([9, 10, 11]). We perform the segmented analysis of (1.1) following the subsequent typical steps (see e.g., [3]):

- find the numerical approximation of the solution of the Cauchy problem (1.1), via the 4-th order Runge Kutta method;
- perform the segmentation of the time series of length $N = 2^M$ into a set of σ short subsets of samples of equal length $p = \frac{N}{\sigma} = 2^m$, $(\sigma = 2^{M-m}$, $m = \overline{0, M})$; on each segment we apply the wavelet transform.

Proceedings of The 4-th International Colloquium "Mathematics in Engineering and Numerical Physics" October 6-8 , 2006, Bucharest, Romania, pp. 12-18.

Since the correlation of different segments of data are independent of the choice of used the wavelet family, we use the Haar wavelets for fast computation reasons ([1, 2, 7]); the results still preserve the generality regarding the freedom of wavelet basis choice.

In Section 3, we apply this method - both in direct and segmented approach, to the biological system introduced by Truscott and Brindley ([13]), used as model for the evolution of plankton ecosystems. Features of the data like, e.g., the concentration of energy will be described ([6, 11]) using the wavelet entropy built using the Shannon information function and the wavelet energy.

2 Energies and entropy for windowed Haar DWT

We shall further denote by $X = \{X_i\}_{i=\overline{0,N-1}}$ be the numerical solution of (1.1), at the moments $\{t_i = \frac{i}{N}\}_{i=\overline{0,N-1}} \subset [0,1)$, where $N = 2^M, \ \ M \in \mathbb{N}$. We split X into $\sigma = \frac{N}{p}$ subsets (segments) of $p = 2^m$ ($m = \overline{0,M}$) samples. Then, the windowed (reduced) p-parameter discrete Haar wavelet transform is given by ([3]):

$$W^{p,\sigma}X = \{W^pX^s\}_{s=\overline{0,\sigma-1}}, X = \{X^s\}_{s=\overline{0,\sigma-1}}, \quad X^s = \{X_{sp+\alpha}\}_{\alpha=\overline{0,p-1}}, \tag{2.1}$$

where $s \in \overline{0,\sigma-1}$, $\sigma = \frac{N}{p}$, p is fixed, and W^p is the discrete Haar wavelet linear transform ([1]): $W^pX = \{\alpha, \{\beta_k^j\}_{j=\overline{0,m-1},k=\overline{0,2^j-1}}\}$, where $2^m = p$. The transform W^p can be defined by recursive matrix factorization ([1, 2]):

$$W^p = \prod_{l=1}^{m} [\,\mathrm{diag}\,(P_{2^l}, I_{2^m-2^l}) \cdot \mathrm{diag}\,(H_{2^l}, I_{2^m-2^l})], \tag{2.2}$$

where the matrix $I_n \in M_n(\mathbb{R})$ is the identity matrix, $I_0 \equiv \varnothing$, P_n is the $n \times n$ permutation matrix that moves the odd place components of the vector X into the first half positions and the even place components into the second half.

For the vector $X = \{X_i\}_{i=\overline{0,2^M-1}}$ ($N = 2^M$), the Haar wavelet transform is $W^NX = W^{2^M}X$, while the segmented transform uses the matices $W^{p,\sigma} = \{W^{2^m,2^n}\}_{m+n=M}$, providing

$$W^{p,\sigma}X^{(s)} = \{\alpha_0^{0(s)}, \{\beta_k^{j(s)}\}_{j=\overline{0,m-1},k=\overline{0,2^{m-1}-1}}\},$$

where $N = 2^M$, $p = 2^m$, $m = \overline{0,M}$, $\sigma = \frac{N}{p} = 2^{M-m} = 2^n$, $s = \overline{0,\sigma-1}$ and $m = \overline{0,M}$, $M \in \mathbb{N}$. In particular, for $\sigma = 1$, we have $p = N$ and $W^{p,\sigma}X = W^NX$.

Each segment of the time-series is characterized by an index, and the corresponding wavelet coefficients depend on three indices:

- the cardinal number of the segment $s \in \overline{0,\sigma-1}$;
- the scale parameter $j \in \overline{0,m-1}$, which relates to the wavelet amplitude $2^{j/2}$ and the domain width $1/(2^j\sigma)$;
- the location $k \in \overline{0,2^j-1}$ of the wavelet definition domain inside the segment domain.

In our case, the used data conversion matrices are

$$W^8 = \begin{pmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}. \tag{2.3}$$

and $W^{4,2} = \text{diag}\,(W^4, W^4$. All in all, applying W^8 and alternatively $W^{4,2}$ for a given particular vector $X = \{X_i \mid i = \overline{0,7}\}$, we have two representations:

- the discrete Haar wavelet transform, giving a single vector of 8-components W^8X with W^8 given by (2.3);
- the reduced Haar wavelet transform giving two 4-component vectors $W^{4,2}X = \{W^4X^{(s)}\}_{s=\overline{0,\sigma-1}}$, where $N = 8,\ p = 4,\ \sigma = \frac{N}{p} = 2,\ m = \log_2 p = 2$, with $X^{(0)} = \{X_i \mid i = \overline{0,3}\}$, and $X^{(1)} = \{X_i \mid i = \overline{4,7}\}$.

The rate of change of the energy along the time-series distribution gives useful information for classifying dynamical systems and characterizing chaotic motions, while the $L^2(\mathbb{R})$-energy of a function $x(t)$ is $E(x) = \int_{-\infty}^{\infty} \mid x(t) \mid^2 dt$. As shown in [3], the following functions of $X = \{X_i\}_{i=\overline{0,2^M-1}}$ can provide useful information on the discrete time-series:

a) the (total) *energy* $E(X) = (\alpha_0^0)^2 + \sum\limits_{j=0}^{M-1} e^j$, where $e^j = \sum\limits_{k=0}^{2^j-1} (\beta_k^j)^2$ $(j = \overline{0, M-1})$ is the *scale energy*.

b) The *segment-energy* which can describe better the variation of the energy in X:

$$E(X^{(s)}) = (\alpha_0^{0(s)})^2 + \sum_{j=0}^{m-1} e^{j(s)}, \tag{2.4}$$

where $e^{j(s)} = \sum\limits_{k=0}^{2^j-1} (\beta_k^{j(s)})^2$, $j = \overline{0, m-1}$, $m = \overline{0, M}$, $s = \overline{0, \sigma-1}$.

c) The *average energy* value is $\bar{E}(X) = \frac{p}{N} \sum\limits_{s=0}^{\sigma-1} E(X^{(s)})$, where $N = 2^M$, $\sigma = \frac{N}{p}$, $p = 2^m$, $m = \overline{0, M}$, and the *relative energy of a segment* $X^{(s)}$ is $E_r(X^{(s)}) = \frac{p \cdot E(X^{(s)})}{2^M \cdot \bar{E}(X)}$, which shows the distribution of energy along X.

d) The *local energy* has the coefficients $S = \frac{(\alpha_0^0)^2}{E(X)}$ and $S_k^j = \frac{(\beta_k^j)^2}{E(X)}$, and shows the role of the *trend* α_0^0 and of the *detail variations* β_k^j.

e) The *local energy of the wavelets within the segment* s of X:

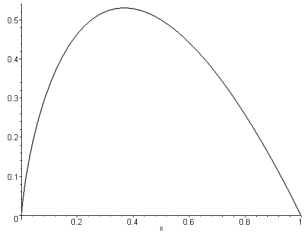
$$S^{(s)} = \frac{(\alpha_0^{0(s)})^2}{E(X^{(s)})}, \quad S_k^{j(s)} = \frac{(\beta_k^{j(s)})^2}{E(X^{(s)})},\ s = \overline{0, \sigma-1}, \tag{2.5}$$

which shows the influence of each wavelet coefficient on the segment energy. Higher values of S reveal a significant influence of the corresponding coefficient in the energy.

A useful tool in data-analysis is the *Shannon* X-dependent information function ([9, 11]):

$$\epsilon(X) = \varphi(S) + \sum_{j=0}^{M-1} \sum_{k=0}^{2^j-1} \varphi(S_k^j), \tag{2.6}$$

where $\varphi(u) = \begin{cases} -u \cdot \log_2 u, & \text{for } u \neq 0 \\ 0, & \text{for } u = 0 \end{cases}$ (see the figure). This function measures the weight of each wavelet coefficient in X, and is non-negative, since $\varphi([0,1]) \subset \mathbb{R}_+$.



If all the coefficients but one are zero, the entropy vanishes, reaching its minimal value, and low values of ϵ shows regularity (some particular wavelet prevalence within all segments) of the motion.

Contrarily, if all coefficients are equally important $\alpha_0^0 = \beta_k^j = \frac{1}{\sqrt{N}}$, then ϵ get its maximum value $\epsilon = M = \log_2 N$. Hence the entropy measures the data concentration, i.e., if the data of X are nearly equal, then the entropy is at the minimum, while if the data shows some sharp jumps, then the entropy is at the maximum ([3]). We note that low or high values of the variable in the Shannon constructive function φ provide low output, which justifies these considerations.

In the case of segmentation, the variation of the entropy along the time-series can be studied, i.e.,

$$\epsilon(X^{(s)}) = \varphi(S^{(s)}) + \sum_{j=0}^{M-1} \sum_{k=0}^{2^j-1} \varphi(S_k^{j(s)}). \tag{2.7}$$

3 The Truscott and Brindley ecological model

We shall further analyse the time-series obtained via the 4-th Runge-Kutta numerical approximation of the solution of the nonlinear Truscott and Brindley SODE ([13]):

$$\begin{cases} \frac{dn}{dt} = \delta \cdot (n_0 - n) - f(n) \cdot p + r \cdot p \\ \frac{dp}{dt} = p \cdot [f(n) - r - k] - R \cdot z \cdot \frac{p^2}{p^2 + \nu^2} \\ \frac{dz}{dt} = \gamma \cdot z \cdot \left(R \cdot \frac{p^2}{p^2 + \nu^2} - \omega \right), \end{cases} \tag{3.1}$$

- $n(t)$ - non-dimensional general nutrient concentration (macro-nutrient: nitrate, phosphate, silicate, etc, or micto-nutrient: iron, etc) limiting plankton growth;
- $p(t)$ - quantity of phytoplankton (plankton);
- $z(t)$ - quantity of zooplankton;
- γ - the assimilation efficiency of zooplankton;
- $\delta(n_0 - n)$ - the mixture of nutrients from a weak base;
- n_0 - the nutrient concentration below the mixed layer (nutrient content in the deep reservoir);
- ω - relates to zooplankton growth efficiency;
- $f(n)p$ is related to the extraction of nutrients by the phytoplankton;
- r - the death decay rate of the phytoplankton;
- k - the reduction rate of the phytoplankton,

where $f(u) = \frac{V \cdot u}{a+u}$, $\frac{u}{a+u} \sim 1$, $a \in (0,1)$, $a \ll 1$, $V = \frac{125}{129}$, $\delta = 0.05$, $\gamma = 0.25$, $R = 0.6$, $r = 0.04$, $k = 0.04$, $\upsilon = 0.035$, $h = V - r - k$, $h = 0.0528992$, for $t \in [0, 1)$.

Renoting $x = (x^1, x^2, x^3) = (n, p, z)$, in order to obtain information on the SODE (3.1), we have used Maple 9.5 software code, with initial data $x_* = (1, 1, 1)$, Runge-Kutta step $h = 0.0528992, x_0 = 1, \omega = 2.5, N = 8$ and $a = 0.01$. As output, we get:

- the 8−data samples for n, p, z, as matrix;

$$\begin{bmatrix} 1. & 1. & 1. \\ -.5888841725 & -.08375984222 & -.5000070920 \\ .1351255792 & .2204844390 & .2617053509 \\ -.07100902535 & -.05389988346 & -.1318192653 \\ .09768220094 & .01948460189 & .07205561568 \\ .03767325748 & -.001753573982 & -.04469035190 \\ .05133427443 & -.0005314593319 & .02936720623 \\ .05016933379 & -.0002120785834 & -.01930849883 \end{bmatrix}$$

- the Haar decomposition samples for the W^8 and $W^{4,2}$ cases, respectively:

$$\begin{bmatrix} .2517623458 & .3888423334 & .2359272257 \\ -.08427769379 & -.3768303637 & -.2094644819 \\ -.1734996369 & -.3748278011 & -.1850534112 \\ -.01692592509 & -.009237282912 & -.008653278185 \\ -1.123510773 & -.7663339334 & -1.060665186 \\ -.1457591767 & -.1940190150 & -.2782639245 \\ -.04243273084 & -.01501765818 & -.08255186533 \\ -.0008237374260 & .0002258362931 & -.03441892113 \end{bmatrix} \begin{bmatrix} .2376161907 & .5414123567 & .3149394968 \\ -.1734996369 & -.3748278011 & -.1850534112 \\ -1.123510773 & -.7663339334 & -1.060665186 \\ -.1457591767 & -.1940190150 & -.2782639245 \\ .1184295333 & .008493744996 & .01871198560 \\ -.01692592509 & -.009237282912 & -.008653278185 \\ -.04243273084 & -.01501765818 & -.08255186533 \\ -.0008237374260 & .0002258362931 & -.03441892113 \end{bmatrix}$$

- the relative energies of the wavelet coefficients in the W^8 and $W^{4,2}$ cases:

$$\begin{bmatrix} .04572523671 & .1427857997 & .04140575438 \\ .005123888793 & .1341002897 & .03263813719 \\ .02171558582 & .1326787976 & .02547409534 \\ .0002066708561 & .00008057984484 & .00005570137435 \\ .9106026117 & .5545925741 & .8368761839 \\ .01532661410 & .03554892953 & .05759945731 \\ .001298901968 & .0002129817382 & .005069421027 \\ .4894992225 \cdot 10^{-6} & .4816431743 \cdot 10^{-7} & .0008812499114 \end{bmatrix} \begin{bmatrix} .04121016012 & .2769181321 & .07424873402 \\ .02197097773 & .1327268101 & .02563474401 \\ .9213119951 & .5547932646 & .8421538216 \\ .01550686681 & .03556179363 & .05796269989 \\ .8704357414 & .1883391135 & .04156196937 \\ .01777958046 & .2227565974 & .008888279440 \\ .1117425673 & .5887711426 & .8089285267 \\ .00004211087609 & .0001331464399 & .1406212245 \end{bmatrix}$$

- the total energy and the relative averages of energies (weights) of segment data within the entire data-set:

$$\begin{bmatrix} .6930995145 & .5294586740 & .6721487960 \end{bmatrix} \begin{bmatrix} .9883759435 & .9996382605 & .9937331660 \\ .01162405698 & .0003617394312 & .006266834280 \end{bmatrix}$$

- the Shannon entropies for both cases:

$$\begin{bmatrix} .5928924115 & 1.822793178 & .9868116854 \end{bmatrix} \begin{bmatrix} .5127691596 & 1.542417152 & .8609185702 \\ .6315308558 & 1.387891545 & .8967127196 \end{bmatrix}$$

For the first component of x renoted $n = x^1$, in the case W^8, we note the prevalence of the Haar wavelet $\psi_0^{2(0)}$ defined on the first quarter of the interval $[0, 1]$, which emphasizes the crisp jump in the first data samples of $n \equiv x^1$ (see the data series below).

As well, on $[0, 2^{-1})$ one may infer from the amplitudes provided by the transformed series and from the distribution of relative energies, the prevalence of the wavelet $\psi_0^{1(0)}$ in the case $W^{4,2}$. These lead to relatively low Shannon entropies compared to the other coordinates p, z, which emphasize a sharp jump in the nutrient close to the starting moment $t = 0$ - compared to the later evolution.

On the other hand, it should be noted that the segmentation induces a loss of information while sensing the jumps between segments. E.g., in our example, the W^8 transform provides extra information on a (relatively) sharp jump between the two (averaged) segments on x^1 and x^3, while the same information is absent in the $W^{4,2}$-transform; however, the segmented approach can recollect this information from the difference of the averages (first coefficients of the x_1 and x^3-transforms).

Finally, we note that a more uniform approach than the usage of the Haar ($D2$) approach (which half-splits the segments while building the corresponding wavelet basis but assumes non-ciclicity of data-samples) is the usage of Daubechies ($D4$) basis, which emphasizes the rolled 2-differences and 2-averages within the single data-segment. The $D4$ Daubechies transfer matrix is

$$W_D^8 = \begin{bmatrix} s_0 & s_1 & s_2 & s_3 & 0 & 0 & 0 & 0 \\ w_0 & w_1 & w_2 & w_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & s_0 & s_1 & s_2 & s_3 & 0 & 0 \\ 0 & 0 & w_0 & w_1 & w_2 & w_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & s_0 & s_1 & s_2 & s_3 \\ 0 & 0 & 0 & 0 & w_0 & w_1 & w_2 & w_3 \\ s_2 & s_3 & 0 & 0 & 0 & 0 & s_0 & s_1 \\ w_2 & w_3 & 0 & 0 & 0 & 0 & w_0 & w_1 \end{bmatrix}$$

where $s_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}$, $s_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}$, $s_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}$, $s_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$ and $w_0 = s_3, w_1 = -s_2$, $w_2 = s_1, w_3 = -s_0$. Then the $D4$-transform $W_D^8(X)$ of the data sequence will be

$$\begin{bmatrix} .02982851594 & .4692918327 & .1404171465 \\ .1499147293 & .09983520513 & .2652486102 \\ .02288002404 & .06599195893 & .03805919561 \\ .06194766840 & .0006946407089 & .07753879527 \\ .08370504772 & .007851768501 & .006497164890 \\ -.00237320675 & -.002470588480 & .03458382114 \\ .3671111039 & .2345491068 & .2868809444 \\ 1.103037226 & .8770855130 & 1.078528671 \end{bmatrix}$$

which still emphasize the crisp jump in the beginning of the first data-sequence, but points out by its last large coefficients in each of the three variables, that 2-term-mediated ciclicity of the data at segment-ends is not likely to happen.

Acknowledgement. The present work was partially supported by the Grant CNCSIS A1478.

References

[1] C. Cattani, *Reduced Haar wavelet spline analysis*, Pharos, 8, 1 (2001), 47-62.

[2] C. Cattani, *Wavelets solving variational problems*, Proc. of the VIII Erouguine Readings Int. Conf., 20-23 My 2002, Brest 2002, 79-85.

[3] C. Cattani, *Wavelet analysis of nonlinear dynamical systems*, Bulletin of Academy of Sciences of Moldova, Mathematics 1, 41 (2003), 58-69.

[4] C. Cattani and A. Ciancio, *On the Haar wavelet analysis of jumps*, Proc. of The 3-rd Int. Coll. "Mathematics in Engineering and Numerical Physics" (MENP-3), October 7-9, 2004, BSG Proceedings 12, Geometry Balkan Press, Bucharest 2005, 72-81.

[5] C. Cattani and A. Ciancio, *Wavelet clustering in time series analysis*, Balkan Journal of Geometry and Its Applications 10, 2 (2005), 33-44.

[6] R. Coifman and M.V. Wickerhauser, *Entropy-based algorithms for best basis selection*, IEEE Transactions on Information Theory 38, 2 (1992), 713-718.

[7] I. Daubechies, *Ten lectures on wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[8] Y. Dai, *The time-frequency analysis approach of electric noise based on the wavelet transform*, Solid State Electronics 44 (2002), 2147-2153.

[9] E. Hulata, R. Segev and E. Ben-Jacob, *A method for spike sorting and detection based on wavelet packets and Shannon's mutual information*, Journal of Neuroscience Methods 117 (2002), 1-12.

[10] D.B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, 2000.

[11] O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schürmann and E. Basar, *Wavelet entropy: a new tool for analysis of short duration brain electrical signals*, Journal of Neuroscience Methods 105 (2000), 65-75.

[12] O. Stănăşilă, *The Mathematical Analysis of Signals and Wavelets* (Romanian), Matrix Rom, Bucharest 1997.

[13] J.E. Truscott and J. Brindley, *Equilibriua, stability and excitability in a general class of plankton population models*, Phyl. Trans. Royal Soc. A347 (1994).

Authors' address:

Vladimir Balan and Simona Mihaela Bibic,
University Pohlitehnica of Bucharest, Faculty of Applied Sciences,
Departments of Mathematics I and III, Splaiul Independentei 313,
RO-060042, Bucharest, Romania.
e-mail: vbalan@mathem.pub.ro and simonabibic@yahoo.com