DTD 5
ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Ecological Modelling xxx (2004) xxx–xxx

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

Identification of step pattern in ordered data sets using the Walsh transform algorithm

Ben-hamadou Radhouan[a,b,*], Ibanez Frédéric[b], Picheral Marc[b], Gorsky Gabriel[b]

[a] *Institut National des Sciences et Technologies de la Mer (INSTM), 28 Rue 2 mars 1934, 2025 Salammbô, Tunisie*
[b] *Laboratoire d'Océanographie de Villefranche (LOV), UMR 7093, Station Zoologique, BP 28, F-06230 Villefranche-sur-mer, France*

Received 7 May 2003; received in revised form 8 June 2004; accepted 6 July 2004

Abstract

This work is an attempt to recognize discontinuities within ordered data series by the use of the Walsh functions. Walsh functions are a system of orthogonal functions used in showing how the energy in a given signal is distributed among these rectangular wave components revealing the existent boundaries. This paper presents the capacity of Walsh transform on determining the statistical significance of discontinuities within ordered marine ecological data. Marine ecosystems present high vertical patterns; dataseries on particulate matter concentrations (measured by the Underwater Video Profiler (UVP)) were used to recognize homogeneity limits in water columns. Firstly, Walsh functions are generated in a compact form. Then, a spectral decomposition is performed on the ecological signal to obtain a stepped Walsh version; the step width at this stage is unvarying, picking constant vertical sections. To avoid this limitation, the estimated series is then smoothed to recognize changing step widths by merging successive blocks presenting statistical non-significant difference; here the nonparametric Kolmogorov–Smirnov test is used. This Walsh transform algorithm provides a fast, simple yet accurate means of separating ordered ecological data into groups of observations corresponding to different marine water masses. This method is applied to four profiles measured at the same site during 2 days. The temporal evolution of depth transitions are first discussed and then used for the validation of the transform model.

Keywords: Ordered data discontinuities; Boundary detection; Walsh function; Spectral decomposition; Underwater Video Profiler; Water masses; Nonparametric; Classification; POMME cruise

1. Introduction

Ecologists (e.g., Livingston, 1903; Clements, 1905) have recognized long, the existence of boundaries, discontinuities, or transition zones between relatively homogeneous ecosystem units. The interpretation of their ecological significance has been quite varied (Whittaker, 1956; Margalef, 1979). Most traditional methods for analyzing ecological data have concentrated on comparisons between homogeneous units, rather than on the boundaries between them (Cornelius and Reynolds, 1991). In order to detect, characterize,

* Corresponding author. Tel.: +216 22 780 469; fax: +216 71 732 622.
E-mail address: hamadou@obs-vlfr.fr (B.-h. Radhouan).

0304-3800/$ – see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.ecolmodel.2004.07.008

and classify boundaries, advance development of improved numerical methods are needed (Holland, 1988; Van der Maarel, 1990). Van der Maarel (1976) related that phytosociologists used quantitative methods for locating discontinuities in ordered data since 1930s. Ludwig and Cornelius (1987) listed the commonly used techniques for (1) searching patterns in one variable along a one-dimensional transect, e.g., spectral analysis, correlograms, periodograms (Ludwig and Goodall, 1978; Legendre and Legendre, 1998); (2) predicting, smoothing, or interpolating patterns of one variable across a two-dimensional grid, e.g., smoothing splines, kriging (Souissi et al., 2001); and (3) classifying community data that repeat due to multivariate sampling along random transects, e.g., cluster analysis, association analysis (Legendre and Legendre, 1998; Hamadou et al., 2001).

Analyses of ecological datasets in general and on marine systems especially, usually focus on the homogeneous communities that are presumed to occur in the study area (Int Panis and Verheyen, 1995). In order to identify and isolate their boundaries, researchers have used classification analysis or ordination technique to recognize different communities within the sampled area. Advances in marine exploration techniques are leading to an increasing number of data series and frequency acquisition, which calls for improved data analysis methods.

In ocean systems, the vertical speed of currents and the mixing of layers are of limited importance with respect to horizontal dynamics. High discontinuity events are then encountered along a vertical transect on ocean water columns. These stratifications and boundaries between homogeneous masses of water are, in spite of their importance, often undetected and no appropriate numerical analysis exists to reveal them.

To overcome these shortfalls, a method using the set of Walsh functions (Walsh, 1923) has been developed, which enhances transitions in signal level and combines information from several profiles in selecting a set of boundaries or clusters. The Walsh functions are a set of rectangular waveforms that have discrete transitions and assume only the values +1 and −1 on the interval [0, 1] (Lanning and Johnson, 1983). Decomposition of the original data series by the Walsh functions returns a stepped version of the original signal. The number of Walsh functions used in this filter procedure determines the numbers of resulting clusters within the ordered data. The use of a high number of Walsh functions allows a high resolution on picking boundaries, even when large clusters are desired a limited number of Walsh functions are used. A major problem encountered with the use of the Walsh transform procedure is that the resulting clusters present the same number of observation and so the same length. To circumvent this limitation, we used a nonparametric statistical method that combined successive clusters from a Walsh version.

This algorithm resulting from this method was applied to four particulate matter profiles collected from 0 to 500 m with a special video system at a single site in the Northeast Atlantic during the cruise Programme d'Océanographie Multidisciplinaire à Méso-Echelle (POMME3).

The objectives of this study included: (1) identification of the boundaries between homogenous successive water columns; (2) dynamics of these water columns during the 2-day sampling period; (3) the development of a tool for a prospective analysis with multiple biogeochemical parameters measured under the same conditions.

This paper describes a mathematical approach to detect discontinuities in the water column. The obtained results are a partition of each profile into successive water layers regarding their particulate matter abundances.

The advantage and possible applications of the method are discussed.

2. Materials and methods

2.1. Study areas and data acquisition

Using the Underwater Video Profiler (UVP) (Gorsky et al., 1992; Stemmann et al., 2000) profiles were performed during the POMME cruise between February and October 2001 in the Northeast Atlantic (Fig. 1). The UVP is a waterproof, vertically deployed multi-instrument array for in situ image and environmental data acquisition and of a system allowing the analysis of images and the treatment of data (Gorsky et al., 2000). The lowering speed of the UVP is 1 m/s, the maximum depth is 1000 m and the acquisition frequency is 12 images per second.

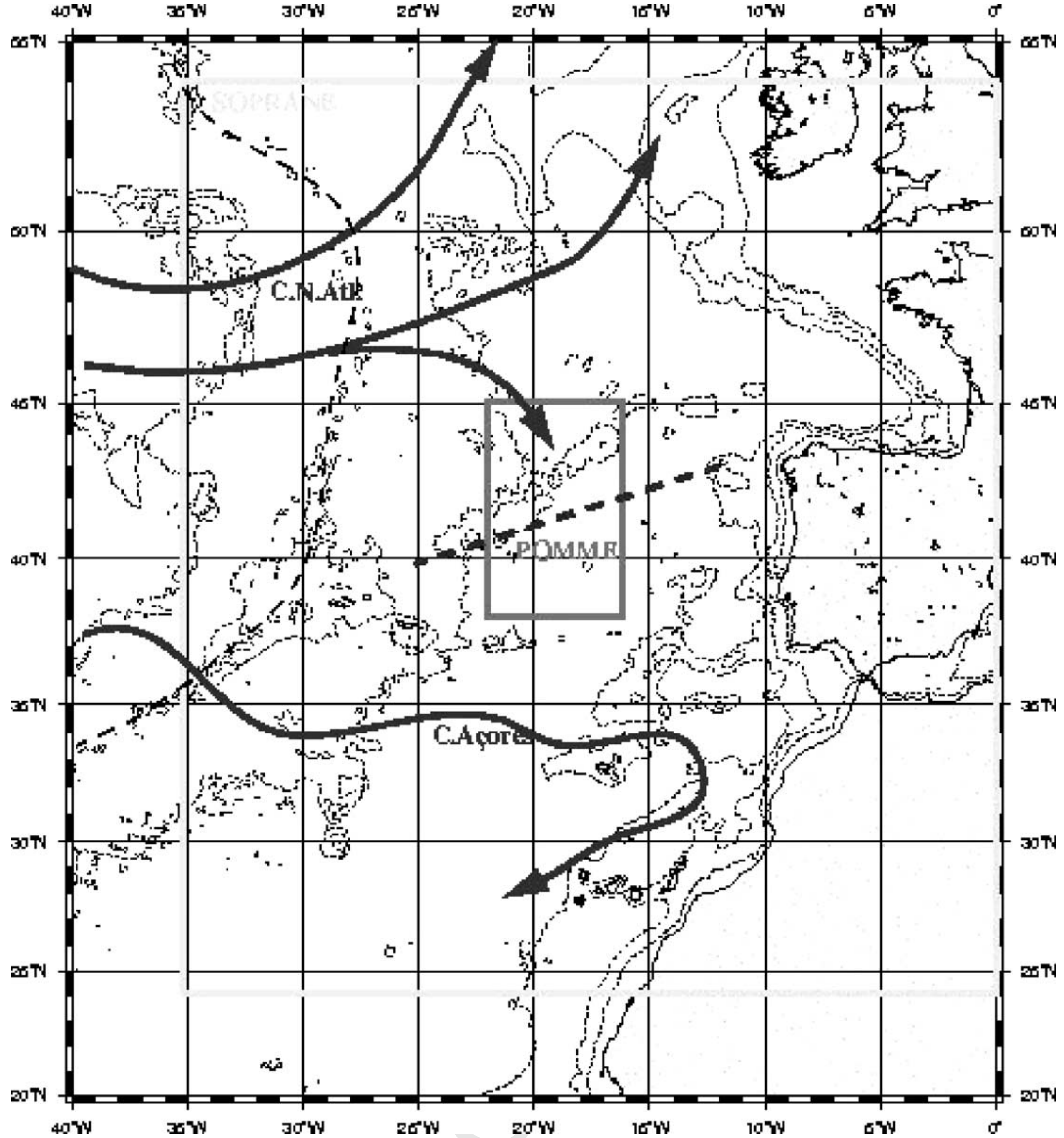


Fig. 1. POMME site. The large frame (SOPRANE) indicates the geographic limits of the SOPRANE hydrodynamic model. The small frame delimits the POMME sampling site. Dotted line shows the subduction zone of modal waters coming form the North Atlantic. North Atlantic and Azores currents are also presented.

Data on number and size of large particulate matter (LPM larger than 0.1 mm) were then extracted from numerical video sequences (Echevarria et al., 2002; Stemmann et al., 2002). In this study, application is made on the number of particle profiles. The number of LPM is computed in each image of the profile record. Then the median of each 4 m (48 images) is computed to avoid high resolution and noisy signal errors (Fig. 2).

In this paper, the case study will be limited to four profiles performed for a single site, throughout 2 days in September 2001. The 0–500 m sampled water column is situated in the center of an anticyclonic gyre.

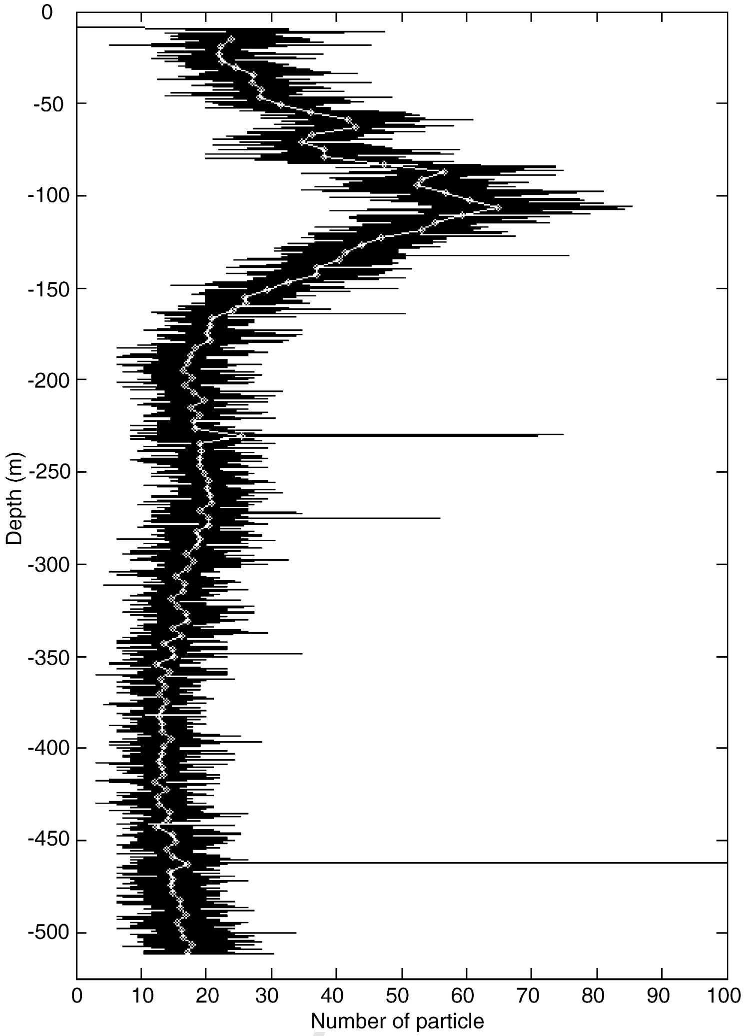


Fig. 2. Original signal in black continuous line corresponds to the number of particles in each image from UVP records. In superimposed dotted white line is the resulted median for each 4 m of records.

3. Data analysis

Typically and due to the high frequency of data acquisition using the UVP, the original dataset presents a large variance comparing to the mean signal (Fig. 2). Nevertheless, some parts of the profile are quite uniform and can be considered to be unvarying. These terms urge to decompose the dataset to an equivalent less "noisy" dataset formed by successive blocks or clusters. Each cluster can then be interpreted separately towards the environmental settings. Variations between clusters are then the result of the variation in these hydrologic and hydrodynamic conditions. This clustering must be sequential, involving the use of constrained statistical methods because of the necessity to gather only successive observations.

A method using the set of Walsh functions has been developed; analogous to the Fourier transform analysis, the Walsh transform performs a spectral decomposition (Morettin, 1974). In this way, the ordered data series can be represented by a set of component functions permitting the reproduction of the original data. When suitably used, the Walsh transform generates smoothed series arranged in steps but also provides the informa-

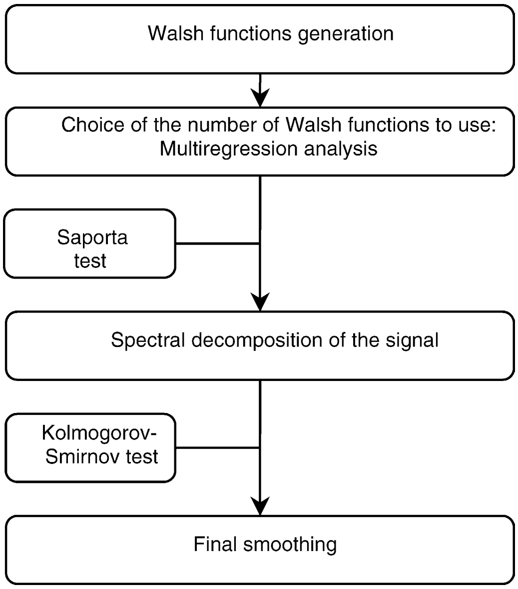


Fig. 3. Organogram presenting the different steps of the numerical analysis.

tion and variance quantity, eliminated by the smoothing procedure. Fig. 3 presents the different steps to perform the Walsh transform.

4. Walsh functions

The Walsh functions form an orthogonal and square-wave set of functions originally presented by Walsh (1923). The sets of functions are a set of rectangular waveforms, which assume only the values +1 and −1. There are several method used to generate the subset of Walsh functions, recursive relation, from Rademacher function Rademacher (1922), through the Hadamard matrix, etc., the method used in this assignment is the discrete version given in a compact form by Brown (1977) and Sen (1982) as,

$$\mathrm{WAL}(0, t) = 1 \quad \text{for } t = 1, 2, \ldots, T,$$

$$\mathrm{WAL}(1, t) = \begin{cases} 1 & \text{for } t = 1, 2, \ldots, \frac{1}{2}T \\ -1 & \text{for } t = (\frac{1}{2}T) + 1, (\frac{1}{2}T) + 2, \ldots, T, \end{cases}$$

where T depicts the total number of orthogonal functions in the set; and $[n/2]$ is the integer part of the argument, n indicates which functions of all the set we are referring to, it is its index or sequence number and t is the time duration of the signal. Fig. 4 shows the first eight Walsh functions. Even numbered Walsh functions are symmetric and odd numbered ones are antisymmetric with respect to the interval mid-point (see Fig. 4). This is analogous to the relationship of the trigonometric cosine and sine functions (Lanning and Johnson, 1983). Otherwise, a general property enables a symmetry relationship as,

$$\mathrm{WAL}(i, j) = \mathrm{WAL}(j, i) \qquad (2)$$

$$\mathrm{WAL}(n, t) = \mathrm{WAL}(t, n) \qquad (3)$$

The practical importance of this is that the Walsh transforms and their inverse represent the same mathematical operation (Sen, 1982).

5. Spectral decomposition

The concept of spectral decomposition is that any data series can be represented by a set of component functions, when combined, recovers the original data (Lanning and Johnson, 1983). The goal is to find the magnitude and sign of the weights of these components. These weights depend on the nature of the data and the characteristics of functions used in decomposition. In order to represent a continuous limited series X_i (i = 1, 2, ..., N) completely by the Walsh functions, it is necessary that the number of data points in the series be equal to the minimum sequence order, N. A necessary requirement for a successful application of Walsh functions is that $N = 2^q$, where q is any convenient positive integer power satisfying the precedent property. Since the Walsh function forms a basis this implies that exists a representation of the series of the form:

$$X_i = \sum_{j=0}^{N} C_j \cdot \mathrm{WAL}(j, i) \qquad (4)$$

where the coefficients C_j are given as a result of the Walsh transform:

$$\mathrm{WAL}(n, t) = \mathrm{WAL}([\tfrac{1}{2}n], 2t) \cdot \mathrm{WAL}(n - 2[\tfrac{1}{2}n], t) \qquad (1)$$

$$C_j = \frac{1}{N} \sum_{j=0}^{N} X_j \cdot \mathrm{WAL}(j, i) \qquad (5)$$

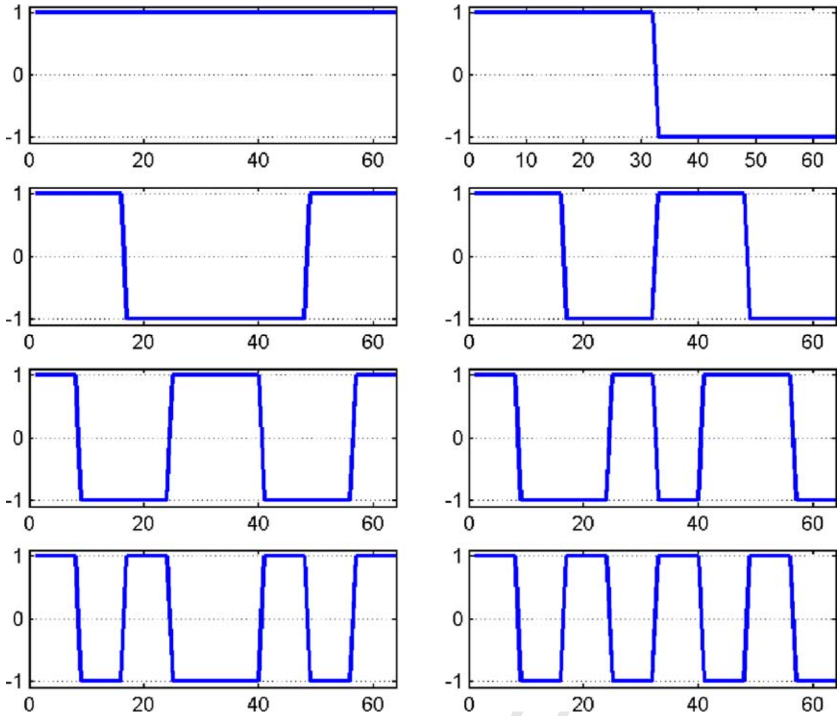


Fig. 4. Walsh functions for $N = 8$. Each function assumes only the values +1 or −1 with discrete transitions between the two.

The coefficients C_j are to be chosen in a way to minimize the mean-square approximation error (Sen, 1982). For ordered data and especially in time series analyses, the Walsh functions are capable of depicting the periodic component with minimum effort of computation and great accuracy.

As previously suggested the number of Walsh functions N must be a power of 2, but the choice remains arbitrary in the literature. Points worthy of notice are that the number of Walsh functions will be the same than the number of blocks issued from the Walsh transform. This fact involves the possibility to change the number of used Walsh functions to reveal either macro-scale or micro-scale changes throughout the data series.

To make the selection of this number objective, a sequential multiple regression adapted from the stepwise multiple regression (Legendre and Legendre, 1998; Hakanson and Boulion, 2002) is used. First the R^2-like coefficient is computed for the multiple regression between the original data series and the independent (and uncorrelated) variables being each set of Walsh functions computed for $N = 4, 8, 16, 32, 64$, etc.

Then an estimator presented by Saporta (1990) and modified here to give more weight to choices with less Walsh functions is computed:

$$\hat{\sigma}^2 = \frac{n}{n-k-1}(1-R^2)s_y^2$$

The modified criterion is:

$$\hat{\sigma}^2 = \left(\frac{n}{n-k-1}\right)^{3/2}(1-R^2)s_y^2$$

This modification is retained following an analysis of sensibility.

The choice of the number of Walsh functions N is admitted for the one that minimize the $\hat{\sigma}^2$ criterion (Fig. 5).

This selection procedure and the stepwise one differ from one another in that there is no test made between additions of new explanatory series (Walsh functions). In that way, new Walsh functions are added in each loop of the sequential multiple regression to the pre-existing dataset with respect to their original position in the entire set of Walsh functions. In each step, not only the following Walsh function is added but also the entire subset of Walsh functions needed

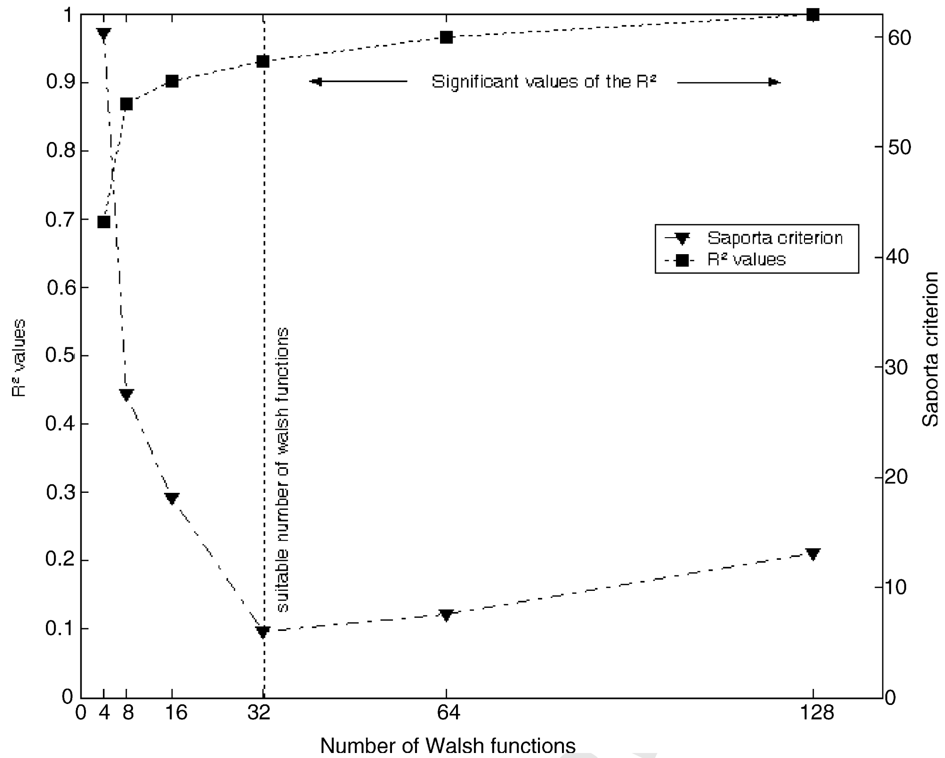


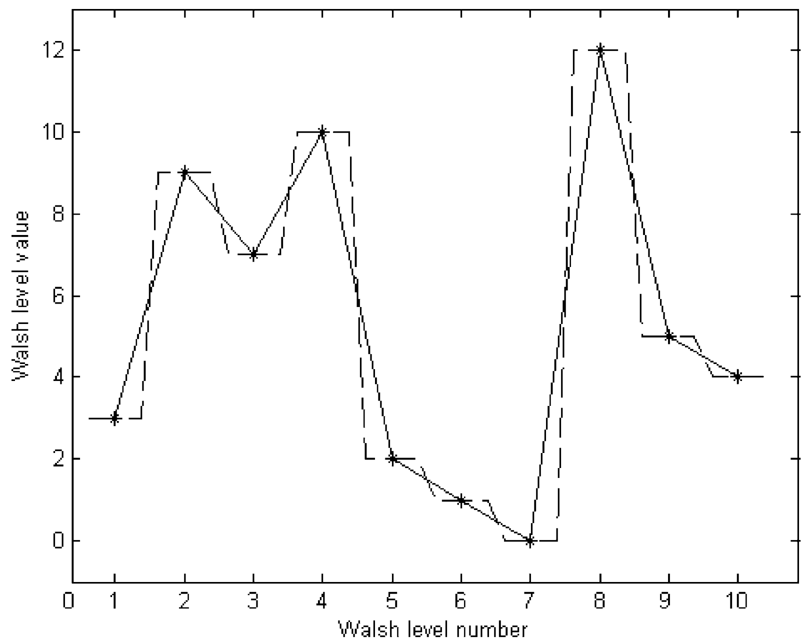
Fig. 5. Growth of the R^2 and $\hat{\sigma}^2$ statistics depending on adding new Walsh functions. The suitable number of Walsh function is picked for the minimum Saporta criterion $\hat{\sigma}^2$ value.

to reach the following power of 2 ($N = 2^q$) as number of total Walsh functions used for the multiple regression.

6. Smoothing the Walsh versions

At this stage, the Walsh version consists in N parts resulting from the decomposition of the original data series. The level's gaps between successive blocks or clusters are more or less important. Regarding the noisy nature of the data series and the high level of acquisition frequency, an additional smoothing is applied. Consisting in merging clusters presenting non-significant differences, a Kolmogorov–Smirnov test is used to test the significance of the difference between the successive magnitudes of the Walsh version (Siegel, 1956; Legendre and Legendre, 1998). To compute this test a contrast function is exploited. This contrast function consists of the alternation of the observed maximum magnitude followed by the observed minimum one, then continued by the second maximum value and the second minimum one and so on; from this contrast function, the successive differences are computed. These differences present the maximum disparity that can be observed between the whole levels considered from the Walsh version; an explanatory model is presented in the (Fig. 6).

The nonparametric Kolmogorov–Smirnov test is computed between the observed difference series (6-C) and the hypothetical difference series (6-D) to determine if these independent random series can be assimilated to be drawn from the same underlying sampling. The Kolmogorov–Smirnov test, to be a nonparametric random (signed-ranks free) test, does not take into account the ranks of the two tested series values, this fact allows to test unsorted series vectors to identify if the empirical distribution functions from these sample vectors are equals.



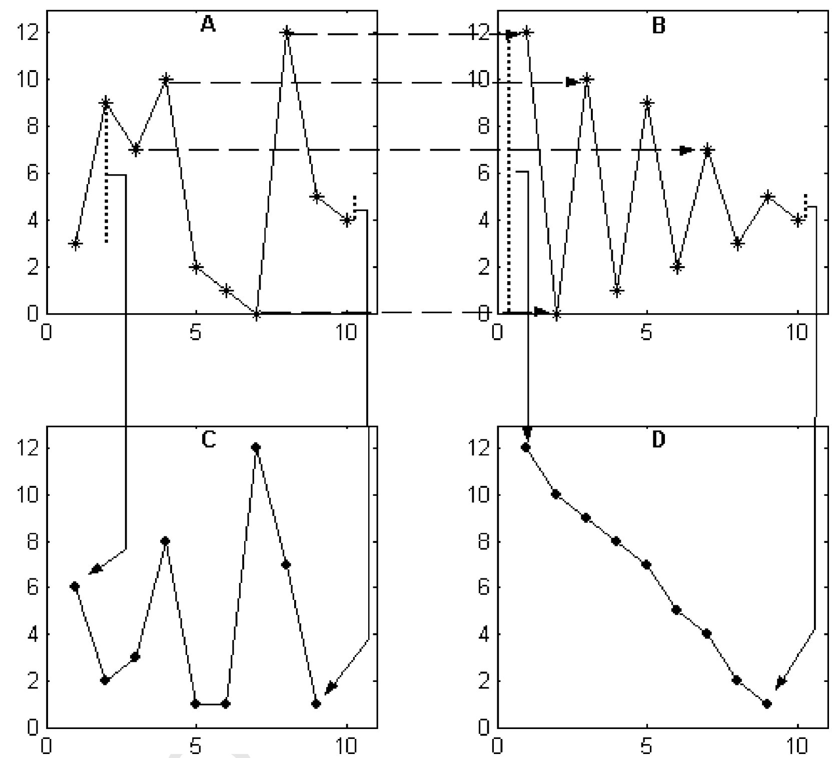


Fig. 6. Explanatory example for the computation of the K–S test. (A) original signal, (B) the resulting contrast function by sorting the original signal, (C) observed differences between successive original signal values, (D) differences between successive values of the contrast function.

If the null hypothesis H_0 that the two series are analogous is rejected then the two clusters separated by the minimum gap, are merged together and the Kolmogorov–Smirnov test is computed one more time between the new observed difference series and the new theoretical one. The amalgamation of the successive clusters presenting low differences stops when hypothesis H_0 that there is no difference between the two series is retained. The hypothesis H_0 is retained when the threshold value of the Kolmogorov–Smirnov test is reached at level α = 0.05. Fig. 7A presents the Kolmogorov–Smirnov test probabilities values func-

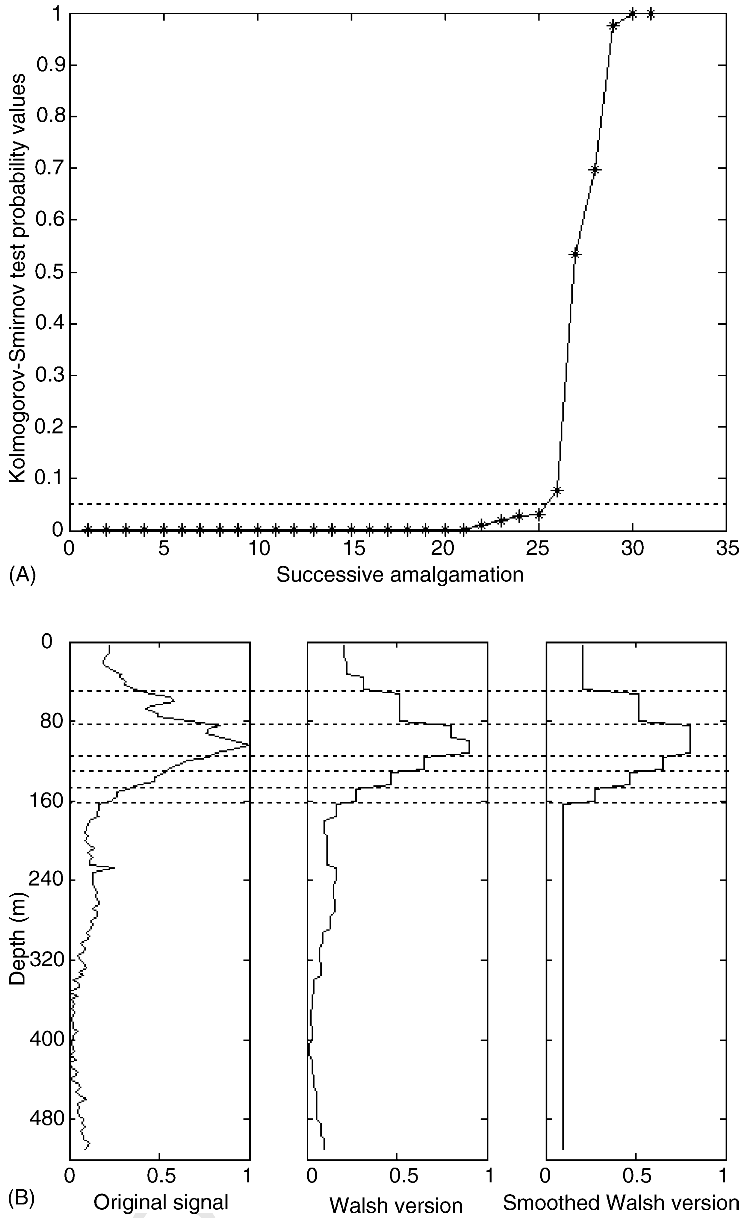


Fig. 7. (A) Kolomogorov–Smirnov test probabilities function of the amalgamation level. Suitable level is picked when probability reaches 0.05. (B) Stages from the original signal to the final smoothed Walsh version. Scales are normalized to 1.

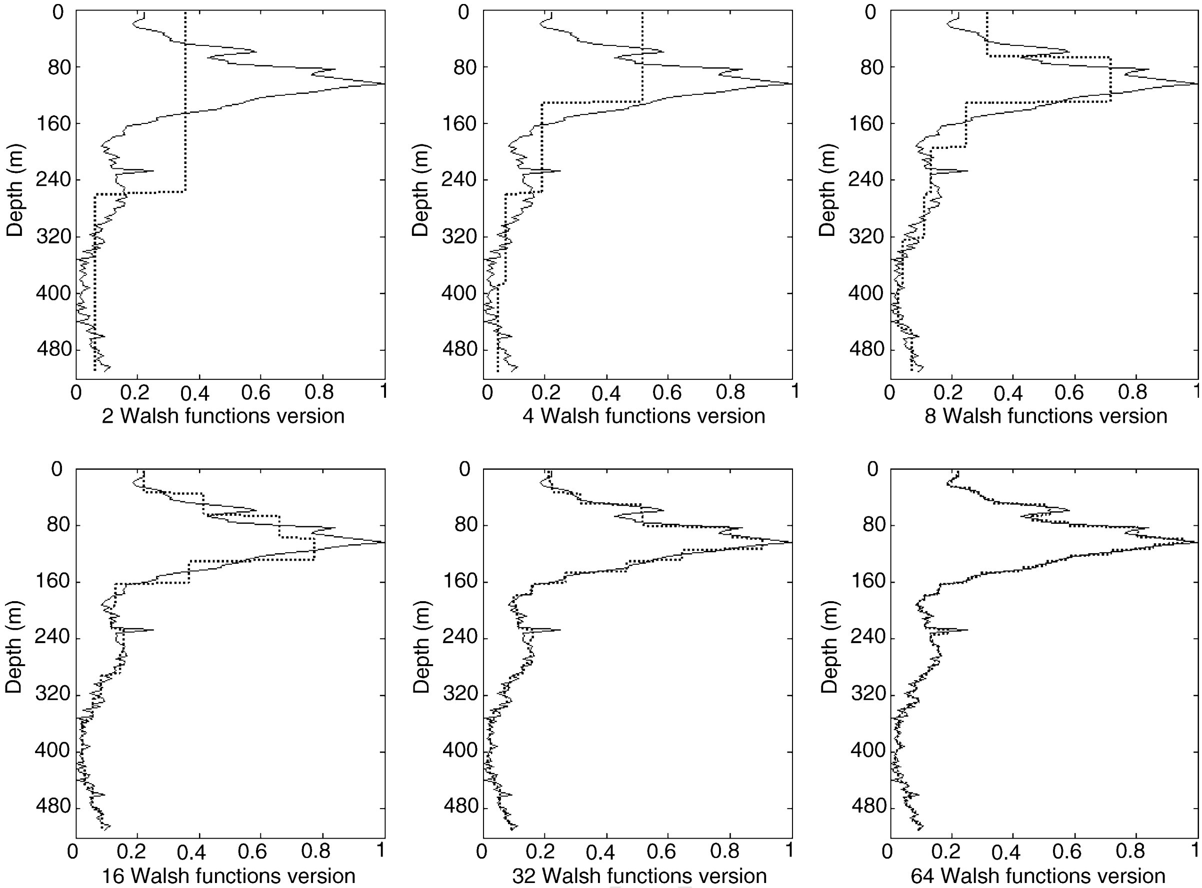


Fig. 8. Walsh functions for the same signal at $N = 2, 4, 8, 16, 32, 64$. Original signal in continuous line, Walsh function in dotted line. Notice that according to the sequential multiple regression the 32 Walsh version was retained.

tion of the amalgamation level; the threshold at 0.05 suggests the number of the amalgamation level to be reached, Fig. 7B presents the final resulting smoothing of the original Walsh version. This smoothing allows the amalgamation of two successive levels on the Walsh version until obtaining a well contrasted Walsh version.

7. A case study

The discontinuities picking procedure is used for a UVP data series recorded in the Atlantic Ocean in which several discontinuities could be found. Due to numerical constraints, only the first 128 ($N = 2^7$) measures were retained, this corresponds to a depth of 480 m. Fig. 8 presents the Walsh versions for different numbers of Walsh functions applied to the same original data series. It can be seen that in the Walsh versions the step width is a constant throughout the duration of the signal and their number is equal to the considered number of Walsh functions. As specified previously, a sequential multiple regression is applied to recognize the suitable number of Walsh functions to be used. Fig. 5 presents the evolution of the R^2 value as a function of the number of used Walsh functions, the suitable number is 32 functions according to the $\hat{\sigma}^2$ criterion.

The Kolmogorov–Smirnov test is then used to determine, which of the different preformed clusters will be

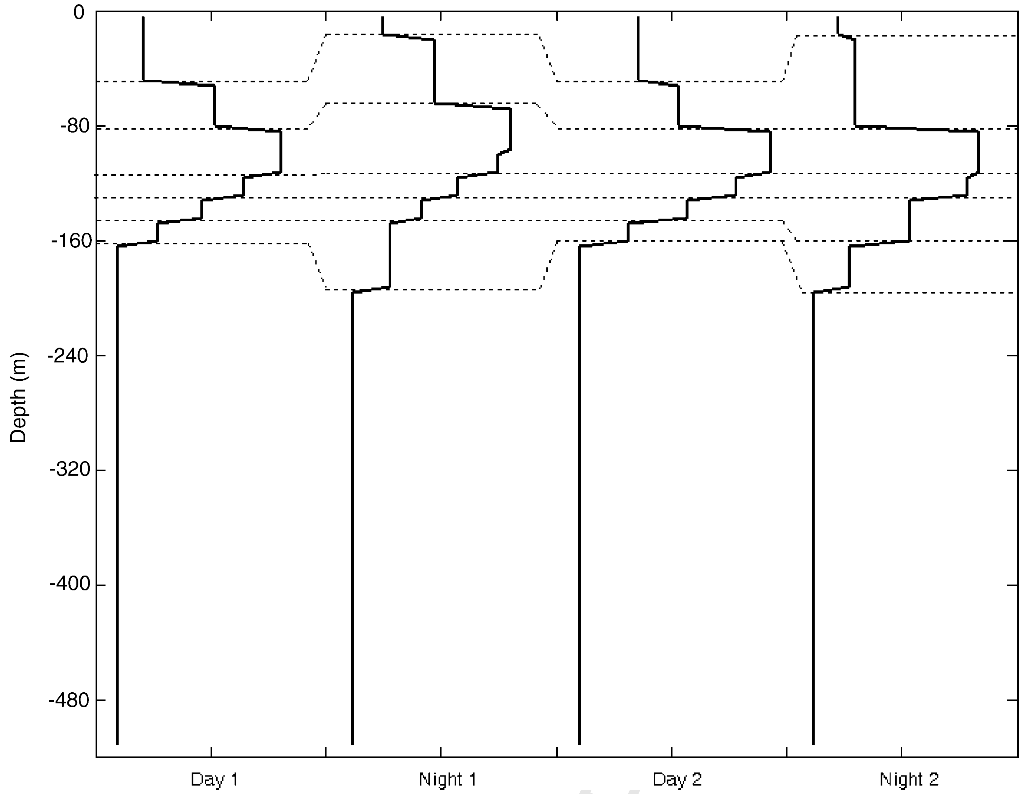


Fig. 9. Walsh versions for successive four UVP profiles recorded in the same site during 48 h. The first 32 Walsh functions were used for the four profiles. Similar depths for discontinuities picking suggest (1) limited changes in water masses allocation during the sampling period (2) validation of the picking discontinuities procedure.

gathered together. Fig. 7B presents the final amalgamation to be considered revealing the different homogenous water layers to be considered in the ecological analysis of the data.

This whole procedure is repeated for three additional profiles recorded with the first one in the same site during 2 days. The results illustrated in (Fig. 9) present the different clustering and the final Walsh version made for each of the profiles considered.

Considering the continuity of the ocean water properties and the relative smallness of the time intervals between profiles (4–12 h), a homogenous stratification of the water column at almost permanent depths is demonstrated suggesting a relative homogeneity of the water columns during the sampling time in terms of relative abundance of particles.

8. Discussion

The main purpose of this paper was to develop an algorithm to classify ordered observations in the marine ecosystems. Methods used for smoothing such as the running means and the running medians are liable to perform a clustering, but the choice of the width of the smoothing-window remains arbitrary. Constrained or conditional classification techniques maintaining the order of samples in the formation of clusters (Gordon, 1973; Legendre et al., 1985) as well as association analysis with contiguity constraint (Williams and Lambert, 1959; Legendre and Legendre, 1998; Hamadou et al., 2001) can also be used when the final number of clusters is specific or predetermined. Another method used in finding discontinuities in time series is the split moving window (SMW), originally described by Whittaker

(1960). The SMW-method is a quantitative method by which relative discontinuities might be objectively revealed from multivariate ordered data (Cornelius and Reynolds, 1991). The method compare distances computed between the two halves of all sequential windows of specific sizes. This technique is in current use in geophysics for identifying soil-boundaries (Webster, 1973), in separating vegetation zones (Wierenga et al., 1987) and is also used for boundary detection in data series from benthic communities (Int Panis and Verheyen, 1995). But this technique seems to be mostly adapted to multivariate series and presents limitations in delimiting homogeneous clusters for univariate datasets, such as the use of statistical means for the computation of dissimilarity between the two halves of the window (Ludwig and Cornelius, 1987). Momen et al. (1996) used cluster analysis to detect both step (abrupt) and monotonic (gradual) changes in time and space patterns. The D^2 to the centre method (Ibanez, 1981) is also inadequate because it allows us to detect one-off discontinuities in a multivariate record but does not allow us to find successive homogeneous periods.

The Walsh transform has been presented as an appropriate tool for use in the analysis of ordered data clustering. Stoffer (1991) presents the statistical applications of this transform comparing to the Fourier one; the method is then presented as suitable to handle discrete or categorical-valued series and to recognize sharp discontinuities.

The cutting out of the data series into multiple boundaries is based on the decomposition of the original signal according to multiple square-wave functions. If measurements made on systems, which oscillate harmonically should be analyzed using Fourier techniques then Walsh analysis is more appropriate for data taken from systems that have discrete or very sudden changes from one state to another (Lanning and Johnson, 1983). This fact presents an important issue for using Walsh analysis on marine data series resulting in the irregular marine structure due to multiple causes (e.g., summer stratification, atmosphere interactions, coastal pollution, complex hydrodynamics, etc.). Such a possibility is encouraged by the lecture of Brillinger (1981) who gives an application of the Walsh transform for the Nile River Discharge data series and Morettin (1981) who provides an excellent review article on the statistical approach of Walsh spectral analysis.

To illustrate this detection analysis of clusters into ordered data, profiles of particulate matter abundances were used. First, Walsh functions are generated to compute a spectral decomposition on data series. Then a smoothing method performs an amalgamation of the resulting clusters to obtain an informative decomposition. This amalgamation is stopped when a statistical threshold is reached (according to the Kolmogorov–Smirnov test). The resulting classification suggests a stratification of the water column (0–500 m) into six layers. This result is reached for the four data series used for the method illustration. High number of significant discontinuities is observed on the upper layers of the water column due to almost certainly to a homogenization at deeper depths coupled to a relative stratification on the surface. Treating time ranged profiles from the same location permits the recognition of homogenous mass of water permanently distributed according to their depth position and their identified LPM abundances. The detection of discontinuities at the same depth for time-scale dependent data series suggests that the Walsh analysis is appropriate for such data series.

Another advantage in using Walsh analysis is the switch between small clusters to large ones allowing different space-scale interpretations. First interpretations can be made according to unsmoothed Walsh version allowing the recognition of small distribution on particulate matters. During the smoothing procedure one can also make use of intermediate clustering for ecological interpretation.

Analyzing Fig. 9, considering the continuity of the ocean water properties and the relative smallness of the time intervals between profiles (4–12 h), a homogenous stratification of the water column at almost permanent depths is demonstrated suggesting a relative homogeneity of the water columns during the sampling time in terms of relative abundance of particles. This point of note corresponds appropriately to the assumption that the central water column in an anticyclonic gyre shows a stable stratification maintained almost certainly by a halocline (Semtner, 1976; Rougerie and Rancher, 1994). Beyond this large connotation a small variation is detected between successive diurnal profiles. A larger distribution of large particles can be identified as a fine diel migration of zooplankton in night-time searching more food supplies. In fact, many species migrate from day-time depths below 200 m up to the surface at night where they feed on the phytoplankton. There has been little modelling of this migration (Wroblewski, 1982; Andersen and Nival, 1991; Steele and Henderson, 1998).

In general, the ecological patterns of these layers can be identified when coupled to other biogeochemical distributions recorded for the same site and in the same condition.

In summary, this paper proposes a method for detecting discontinuities within an ordered data series coupled with nonparametric methods for determining the statistical significance of scale-dependent discontinuities. The Walsh analysis successfully decomposes a given data series into linear and simple components. Their mathematical manipulations are based on simple addition and/or subtraction. A full set of Walsh functions is the most suitable transformation for representing the sudden changes in vertical distribution of marine snow. The use of Walsh analysis according to other parameters in oceanography will certainly give enhanced appreciation of its concordance with marine ecology data series.

Acknowledgements

We thank Marsh Youngbluth and Clare Beckhelling for reading and for their constructive comments. We all thank the two anonymous referees for constructive comments on the manuscript. The research was supported by the "Programme National-Environnement Côtier (PNEC art4)". Data collection was enabled by the "Programme Océan Multidisciplinaire Méso Echelle" POMME cruise.

References

Andersen, V., Nival, P., 1991. A model of the diel vertical migration of zooplankton based on euphausiids. J. Mar. Res. 49, 153–175.

Brillinger, D.M., 1981. Time series: Data Analysis and Theory. Holden-Day, San Francisco, p. 540.

Brown, R.D., 1977. A recursive algorithm for sequency-ordered fast Walsh transforms. IEEE Trans. Comp. 26 (8), 879–882.

Clements, F.E., 1905. Research Methods in Ecology. Lincoln, Nebraska, 334.

Cornelius, J.M., Reynolds, J.F., 1991. On determining the statistical significance of discontinuities within ordered ecological data. Ecology 72 (6), 2057–2070.

Echevarria, F., Garcia Lafuente, J., Bruno, M., Gorsky, G., Goutx, M., Gonzalez, N., Garcia, C.M., Gomez, F., Vargas, J.M., Picheral, M., 2002. Physical-biological coupling in the Strait of Gibraltar. Deep Sea Res., II Top. Stud. Oceanogr. 49 (19), 4115–4130.

Gordon, A.D., 1973. Classsification in the presence of constraints. Biometrics 29, 821–827.

Gorsky, G., Aldorf, C., Kage, M., Picheral, M., Garcia, J., Favole, J., 1992. Vertical distribution of suspended aggregates determined by a new Underwater Video Profiler. Ann. Inst. Oceanogr. Paris 68 (1–2), 275–280.

Gorsky, G., Picheral, M., Stemmann, L., 2000. Use of the Underwater Video Profiler for the study of aggregate dynamics in the North Mediterranean. Estuar. Coast. Shelf Sci. 50 (1), 121–128.

Hakanson, L., Boulion, V.V., 2002. Empirical and dynamical models to predict the cover, biomass and production of macrophytes in lakes. Ecol. Model. 151 (2–3), 213–243.

Hamadou, R.B., Ibanez, F., Souissi, S., Cathelineau, A.C., 2001. Spatial analysis of hydrological and phytoplanktonic data of the Bay of Tunis. Multivariate cartography. Medit. Mar. Sci. 2 (2), 67–85.

Holland, M.M., 1988. SCOPE/MAB technical consultations on landscape boundaries. SCOPE/MAB workshop on ecotones, Paris, France. Biol. Int. (special issue) 47–106.

Ibanez, F., 1981. Immediate detection of heterogeneities in continuous multivariate, Oceanographic recordings. Application to time series analysis of changes in the Bay of Villefranche sur Mer. Limnol. Oceanogr. 26 (2), 336–349.

Int Panis, L., Verheyen, R.F., 1995. On the use of split moving window analysis for boundary detection in ordered dataseries from benthic communities. Neth. J. Aquat. Ecol. 29 (1), 49–53.

Lanning, N., Johnson, D.M., 1983. Automated identification of rock boundaries: an application of the Walsh transform to geophysical well-log analysis. Geophysics 48 (2), 197–205.

Legendre, P., Dallot, S., Legendre, L., 1985. Succession of species within a community: chronological clustering, with applications to marine and freshwater zooplankton. Am. Nat. 125, 257–288.

Legendre, P., Legendre, L., 1998. Numerical ecology. Developments in Environmental Modelling 20. Elsevier Science B.V., Amsterdam.

Livingston, B.E., 1903. The distribution of the upland societies of Kent County. MI Bot. Gaz. 35, 36–55.

Ludwig, J.A., Goodall, D., 1978. A comparison of paired- with blocked-quadrat variance methods for the analysis of spatial pattern. Vegetatio 38, 49–59.

Ludwig, J.A., Cornelius, J.M., 1987. Locating discontinuities along ecological gradients. Ecology 68, 448–450.

Margalef, R., 1979. The organization of space. Oikos 33, 152–159.

Momen, B., Eichler, L.W., Boylen, C.W., Zehr, J.P., 1996. Application of multivariate statistics in detecting temporal and spatial patterns of water chemistry in Lake Georges, New York. Ecol. Model., 183–192.

Morettin, P.A., 1974. Walsh spectral analysis. SIAM Rev. 23, 279–291.

Morettin, P.A., 1981. Walsh Fourier transforms. Bull. Int. Stat. Inst. 49, 1211–1230.

Rademacher, H., 1922. Einige Sätze über Reihen von allgemeinen orthogonal Funktionen. Math. Ann. 87, 112–138.

Rougerie, F., Rancher, J., 1994. The Polynesian South Ocean: features and circulation. Mar. Pollut. Bull. 29 (1–3), 14–25.

Saporta, G., 1990. Probabilités. In: Analyse des Données et Statistiques. Editions Technip, Paris, 493.

Semtner Jr., A.J., 1976. Numerical simulation of the Arctic Ocean circulation. J. Phys. Oceanogr. 6 (4), 409–425.

Sen, Z., 1982. Walsh solutions in hydroscience. In: Hydrosciences, A.H., El-Shaarawi, S.R., Esterby (Eds.), Time Series Methods in Hydrosciences. Elsevier Scientific Publishing Company, Amsterdam, pp. 390–404.

Siegel, S., 1956. Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Series in Psychology. Kogakusha Company, Tokyo.

Souissi, S., Ibanez, F., Ben-Hamadou, R., Boucher, J., Cathelineau, A.C., Blanchard, F., Poulard, J-C., 2001. A new multivariate mapping method for studying species assemblages and their habitats: example using bottom trawl surveys in the Bay of Biscay (France). Sarsia 86, 527–542.

Steele, J.H., Henderson, E.W., 1998. Vertical migration of copepods. J. Plankton Res. 20, 787–799.

Stemmann, L., Picheral, M., Gorsky, G., 2000. Diel variation in the vertical distribution of particulate matter (>0.15 mm) in the NW Mediterranean Sea investigated with the Underwater Video Profiler. Deep Sea Res., I Oceanogr. Res. Pap. 47 (3), 505–531.

Stemmann, L., Gorsky, G., Marty, J.-C., Picheral, M., Miquel, J.-C.U., 2002. Four-year study of large-particle vertical distribution (0–1000 m) in the NW Mediterranean in relation to hydrology, phytoplankton, and vertical flux. Deep Sea Res., II Top. Stud. Oceanogr. 49 (11), 2143–2162.

Stoffer, D.S., 1991. Walsh-Fourier analysis and its statistical applications. J. Ame. Stat. Ass. 86, 461–485.

Van der Maarel, E., 1976. On the establishment of plant community boundaries. Berichte der Deutschen Botanischen Gesellschaft 89, 415–443.

Van der Maarel, E., 1990. Ecotone and ecoclines are different. J. Veg. Sci. 1, 135–138.

Walsh, J.L., 1923. A closed set of orthogonal functions. Am. J. Math. 45, 5–24.

Webster, R., 1973. Automatic soil-boundary location from transect data. Math. Geol. 5, 27–37.

Whittaker, R.H., 1956. Vegetation of the Great Smoky Mountains. Ecol. Monogr. 26, 1–80.

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon California. Ecol. Monogr. 30, 279–338.

Wierenga, P.J., Hendrickx, J.M.H., Nash, M.H., Ludwig, J.A., Daugherty, L.A., 1987. Variation of soil and vegetation with distance along a transect in the Chihuahuan Desert. J. Arid Environ. 13, 53–63.

Williams, W.T., Lambert, J.M., 1959. Multivariate methods in plant ecology. I. Association-analysis in plant communities. J. Ecol. 47, 83–101.

Wroblewski, J.S., 1982. Interaction of currents and vertical migration in maintaining *Calanus marshallae* in the Oregon upwelling zone—a simulation. Deep Sea Res. 29, 665–686.