Survey of Protozoa and Metazoa populations in wastewater treatment plants by image analysis and discriminant analysis

A. L. Amaral[1,§], M. da Motta[2,‡], M. N. Pons[2], H. Vivier[2,3], N. Roche[4], M. Mota[1] and E. C. Ferreira[1,*,†]

[1]*Centro de Engenharia Biológica, Universidade do Minho, 4710-057 Braga, Portugal*
[2]*Laboratoire des Sciences du Génie Chimique, CNRS-ENSIC-INPL, 1 rue Grandville, F-54001 Nancy cedex, France*
[3]*Laboratoire Chimie et Environnement, Faculté des Sciences, 23 rue Paul Michelon, F-42023 St Etienne cedex 2, France*
[4]*UT de Marseille, Dept Génie Chimique, Génie des Procédés Laboratoire en Procédés Propres et Environnement—EA 884— Université d'Aix-Marseille 142, traverse C. Susini, BP 157—F-13388 Marseille cedex 13, France*

SUMMARY

The biota present in the activated sludge of a wastewater treatment plant are mainly composed of bacteria, Protozoa and Metazoa. The Protozoa and Metazoa species are important micro-organisms taking part in the ecosystem balance in wastewater treatment plants and are quite sensitive to physical, chemical and operational processes. Several authors have established relationships between the predominance of certain species or groups and some variables of the plant. Such analysis demands the identification and quantification of the different species, which requires skilled workers, specializing in zoology or protozoology, and is time-consuming. To overcome those problems, semi-automatic image analysis and discriminant analysis were carried out for the identification procedures. The overall results were very promising in terms of Protozoa and Metazoa group recognition and of survey of plant conditions. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: activated sludge; Protozoa; Metazoa; image analysis; discriminant analysis

1. INTRODUCTION

The activated sludge wastewater treatment relies on the activity of a bacterial aerobic culture suspended in the plant aerated tank and fed with fresh effluent, which accounts for the biodegradable organic substances needed for its development. The presence of such a bacterial culture allows also the growth of a microfauna consisting mainly of predator organisms such as Protozoa and Metazoa.

---

*Correspondence to: E. C. Ferreira, Centro de Engenharia Biológica, Universidade do Minho, 4710-057 Braga, Portugal.
†E-mail: ecferreira@deb.uminho.pt
‡Present address: Departamento de Engenharia Química, Universidade Federal de Pernambuco UFPE, Cidade Universitária, Recife, Brasil 50.740-521, PE, Brasil.
§Present address: Departamento de Tecnologia Química, Escola Superior de Tecnologia e de Gestão, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Apartado 1038, 5301-854 Bragança, Portugal.

Microscopic observation of both Protozoa and Metazoa is nowadays rather common and can provide precious information on wastewater treatment, final effluent quality and toxic substances. According to Madoni (1994), the following features should be present in order for an efficient treatment to take place: high Protozoa density ($>10^3$ per mL), dominant crawling and sessile forms and a well diversified community, where no species or group of species is overwhelming. When such is not the case, the predominant group or group's knowledge may give some clues for the wastewater treatment plant diagnosis.

Although the exact number of Protozoa species remains unknown, over 50 000 species have been accounted for so far (Jahn *et al.*, 1999; Fenchel, 1999). In wastewater treatment, the three major Protozoa categories are the flagellates, sarcodines and ciliates (Richard, 1989), with a large predominance of ciliate species. The vast majority of the ciliate Protozoa nourish on bacteria, but some feed off other ciliates or flagellates and, according to Madoni (1994), bacteriophage ciliates can be divided into three groups: free swimming (moving freely in the effluent), crawling (grazing and living in the surface of the flocs) and sessiles (attached to sludge flocs by a stalk structure).

The main Metazoa present in a wastewater treatment plant are the rotifers, nematodes (*Gastrotrichia*) and *Oligotricihia* (Richard, 1989).

In the present work the most common Protozoa and Metazoa species and groups in well established wastewater treatment plant were studied—mainly the flagellate, sarcodine and ciliate Protozoa as well as the rotifer, gastrotriches and oligotriches Metazoa. The studied Protozoa and Metazoa species, genera, orders and sub-class are described in Table 1.

Two of the major problems when using Protozoa and Metazoa in the plant's diagnosis stems from the fact that skilled workers specialized in zoology or protozoology are needed, and that the identification process is time-consuming. Image analysis emerges, then, as a potentially alternative

Table 1.   Studied Protozoa and Metazoa in this work

Protozoa	Flagellate		*Peranema* sp.
	Sarcodine		*Arcella* sp.
			Euglypha sp.
	Ciliate	Free swimming Carnivorous	*Trachelophyllum* sp.
			Litonotus sp.
			Suctoria (sub-class)
		Crawling	*Aspidisca cicada*
			Euplotes sp.
			Trithigmostoma sp.
			Trochilia sp.
		Sessile	*Carchesium* sp.
			Epistylis sp.
			Opercularia sp.
			Vorticella (genus)
			V. aquadulcis
			V. microstoma
			Zoothamnium sp.
Metazoa	Rotifer		Digononta (order)
			Monogononta (order)
	Gastrotrichia		Nematoda (sub-class)
	Oligotrichia		*Aelosoma* sp.

tool to overcome such drawbacks. However, there have been few works in this field as yet (Amaral *et al*., 1999; Motta *et al*., 2001; Golz and Lange, 2001).

Therefore, the main objective of this work being the survey of the wastewater treatment plant conditions by Protozoa and Metazoa populations, the focus was set on the development of a semi-automatic program concerning the morphological characterization of each individual organism. Subsequently, the identification of each species or group was carried out by means of the discriminant analysis statistical technique (Einax *et al*., 1997).

2. MATERIALS AND METHODS

The Protozoa and Metazoa species studied in this work were collected from aerated tanks of wastewater treatment plants in Nancy (France) and Braga (Portugal) treating domestic and industrial effluents. In all cases, the period between the sampling, transportation and image acquisition did not exceed 3 h.

For the microscopic observations a drop of sludge was put on a slide and covered with a cover slip in a direct light microscope. The magnifications used for the Protozoa and Metazoa were the following: *Aelosoma* sp. ($25\times$ and $100\times$); Nematoda, ($100\times$ and $250\times$); Digononta, Monogononta, *Arcella* sp. and *Euglypha* sp. ($250\times$ and $400\times$); *A. cicada*, *Carchesium* sp., *Epistylis* sp., *Euplotes* sp., *Litonotus* sp., *Opercularia* sp., *Peranema* sp., Suctoria, *Trachellophyllum* sp., *Trithigmostoma* sp., *Trochilia* sp., *Vorticella*, *V. aquadulcis*, *V. microstoma* and *Zoothamnium* sp. ($400\times$).

Image acquisition in Nancy was achieved by observation in a Leitz Dialux 20 microscope (Leitz, Wetzlar), coupled to a Hitachi CCTV HV-720E(F) gray scale video camera (Hitachi, Tokyo) and grabbed to the computer in 768×576 pixels in 8 bit format, by a Matrox Meteor frame grabber (Matrox, Montreal) using the Visilog 5 commercial software (Noesis S.A., Les Ulis). In Braga, image acquisition was achieved by observation in a Zeiss Axioscop microscope (Zeiss, Oberkochen), coupled to a Sony CCD AVC-D5CE gray scale video camera (Sony, Tokyo) and grabbed to the computer in 768×576 pixels in 8 bit format, by a Data Translation DT 3155 frame grabber (Data Translation, Marlboro) using the Image Pro Plus commercial software (Media Cybernetics, Silver Spring).

3. THE DEVELOPED PROGRAM

A program to process the acquired images and obtain the desired morphological data was developed in Visilog 5.1. On one hand the gray-level image is enhanced by histogram equalization to sharpen the borders of the Protozoa and Metazoa. On the other hand a median filter and local histogram equalization are applied to the gray-level image to diminish pixel-to-pixel differences and resharpen the borders. The resulting images are combined for a better differentiation of the organisms out of the background. Using the mouse, the user defines a polygonal area of interest around the Metazoa(n) or protozoa(n). The image segmentation consists of thresholding the protozoa(n) or Metazoa(n) borders, within the area of interest, by a given gray level (user defined or automatic by intra-variance determination). The obtained binary image (0—background, 1—object) is then subsequently filled, eroded and reconstructed in order to delete small debris. Finally, a closing operation is performed and the object is filled. When Protozoa(n) or Metazoa(n) are touching each other, a separation of the objects of interest is needed. In this case the objects are separated by means of the watershed algorithm (Russ, 1995). As the cilia and/or stalks (see Figure 1) are disconnected from the organism's body in the
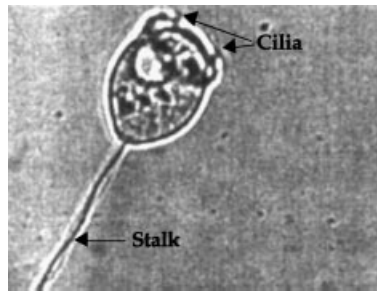
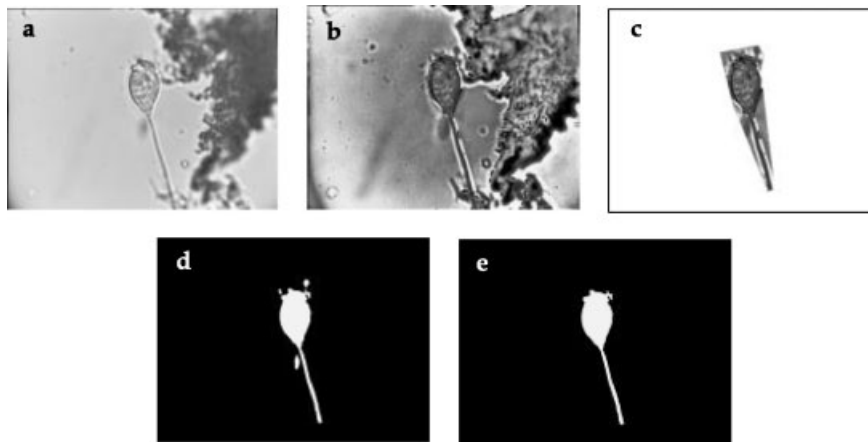Figure 1.   Representation of the stalk and cilia of a sessile Protozoa

Figure 2.   Main steps of the program: original image (a); pre-processed image (b); area of interest (c); thresholded image (d); and final image (e)

watershed stage they should be retrieved. First, all the cilia and stalks from the image prior to the watershed are identified by a logical subtraction between the images before and after the watershed, immediately followed by an erosion and reconstruction step to delete debris.

Keeping in mind the elongated shape of both cilia and stalk, a shaped based procedure is then used to recognize them from remaining debris. This procedure relies on the skeletonization of the objects and determination of the skeleton branch lengths. Branches longer than a given value are then used to reconstruct cilia and stalks. Finally, the cilia and stalks are reconnected to the protozoan or Metazoan, by a logical OR operation. Figure 2 represents the main steps of the image analysis procedure.

4. DETERMINATION OF THE MORPHOLOGICAL DESCRIPTORS

The final stage of the program is the determination of the morphological descriptors for all the Protozoa(n) and Metazoa(n). Except when explicitly indicated, the full definition of these descriptors can be found in Russ (1995): surface (S); equivalent diameter (D); Crofton perimeter (P); length (L)

(e.g. the largest Feret diameter, F_{Max}); width (W) (e.g. the smallest Feret diameter, F_{Min}); mean width ($W_M = S/L$); Feret factor ($F_rF = F_{Max}/F_{Max90}$), where F_{Max90} is the Feret diameter at $90°$ of the largest Feret diameter; eccentricity (Ecc) (Glasbey and Horgan, 1995); form factor ($FF = P^2/(4\pi S)$); largest concavity index (LCI) (Pons *et al.*, 1997); robustness (Rob) (Pons *et al.*, 1997); concavity ratio (CR) (Pons *et al.*, 1997); perimeter factor ($PF = P_P/P$), where P_P is the perimeter after stalk deletion; convexity ($Conv = P_{Conv}/P$), where P_{Conv} is the convex envelope perimeter; compactness ($Comp = D/F_{Max}$); solidity ($Sol = S/S_{Conv}$), where S_{Conv} is the convex envelope surface; fractal dimension (D_{EDM}), calculated using the Euclidian distance map method. Some descriptors, specially designed for Protozoa and Metazoan, are introduced as follows: mean stalk width ($W_{Stk} = S_{Stk}/L_{Stk}$), where S_{Stk} is the stalk surface and L_{Stk} the stalk length, given by:

$$L_{Stk} = \frac{\frac{P_{Stk}}{2} + \sqrt{\left(\frac{P_{Stk}}{2}\right)^2 - 4S_{Stk}}}{2} \tag{1}$$

P_{Stk} being the stalk perimeter.

$$W_S W_{MB} = \frac{W_{Stk}}{W_{MB}} \tag{2}$$

where W_{MB} is the body mean width, and

$$W_M W_B = \frac{W_{MB}}{W_B} \tag{3}$$

where W_B is the body width.

With the exception of perimeter factor, mean stalk width and $W_S W_{MB}$, all the other descriptors were determined for both the full Protozoan or Metazoan and the Protozoan and Metazoan body (without the stalk or stalk alike structures).

4. DATA PROCESSING

4.1. Discriminant analysis

The obtained morphological descriptors where subsequently studied and organized in a manner that allowed the isolation and identification of each species, genus, order or sub-class. Bearing this purpose in mind, the multivariate statistical technique discriminant analysis was performed using the XlStat add-in tool (Thierry Fahmy, Paris) for an Excel spreadsheet (Microsoft Corporation, Redmond).

Unlike in previous works (Amaral *et al.*, 1999), discriminant analysis was used because it increases inter-class variability instead of intra-class variability, as is the case in principal components analysis.

This technique determines new variables (discriminant functions) as linear combinations of the original descriptors, with the goal of increasing the inter-class variability, and thereby obtains a better separation between the studied species and/or groups. Furthermore, in discriminant analysis, the groups or classes of data are modeled with the aim of reclassifying the given object with a low error risk and of classifying new objects using the new discriminant functions (Einax *et al.*, 1997). The object coordinates in the new discriminant functions space are obtained from the original descriptors.

4.2. Division in groups and sub-groups

For identification purposes, due to the large number of studied species, genera, orders and sub-classes we needed to subdivide the observations into several groups where discriminant analysis was carried out. This subdivision was based on morphological similitude as follows: Metazoa (Class IIa), flagellate or fine stalked Protozoa (Class IIb subsequently subdivided into low body width, Class IIIa, or high body width, Class IIIb), prominent ciliated Protozoa (Class IIc), non-prominent ciliated and sarcodine Protozoa (Class IId) and finally thick stalked or elongated Protozoa and Metazoa (Class IIe subsequently subdivided in stalked, Class IIIc, or non-stalked, Class IIId). Additionally, three other groups (Dif) were found to be necessary, accommodating yet poorly identified species, in order to improve the identification. The species, genera, orders or sub-classes belonging to each class are represented in Table 2.

Table 2.   Species, genera, orders or sub-classes of each class

Class I			
	Class IIa		*Aelosoma* sp.
			Nematoda
	Class IIb	Class IIIa	*Peranema* sp.
			V. aquadulcis
			V. microstoma
		Class IIIb	*Carchesium* sp.
			Suctoria
			Vorticella
			Zoothamnium sp.
	Class IIc		*A. cicada*
			Euplotes sp.
	Class IId		*Arcella* sp.
			Euglypha sp.
			Trachelophyllum sp.
			Trithigmostoma sp.
			Trochilia sp.
	Class IIe	Class IIIc	*Epistylis* sp.
			Opercularia sp.
		Class IIId	*Litonotus* sp.
			Monogononta
			Digononta
	Dif 1		*Carchesium* sp.
			Epistylis sp.
			Monogononta
			Trithigmostoma sp.
	Dif 2		*Opercularia* sp.
			V. aquadulcis
			V. microstoma
	Dif 3		*A. cicada*
			Euglypha sp.
			Euplotes sp.
			Litonotus sp.
			Trachelophyllum sp.

4.3. Validation

A total of 50 micro-organisms for each species, genus, order or sub-class (except for *Trochilia* sp. with 37) were used for the identification process by discriminant analysis. Subsequently, the location zones for the new variables (discriminant functions) for each species, genus, order or sub-class were determined with the center as the variable average values and an ellipsoidal location zone as a function of the descriptor standard deviation. The optimum radii of the ellipsoid were determined as

$$r_{kdf} = C_e \times \sigma_{kdf} \tag{4}$$

where r_{kdf} is the radius of the ellipsoid for each group and discriminant function, C_e is the value of the radius coefficient and σ_{kdf} is the value of the standard deviation for each group and discriminant function.

A trial-and-error method was used to determine the optimal radius, in which several values of C_e were proposed and the best differences between the recognition and the error (due to misclassification) percentages were chosen.

In the validation process, another set of 50 micro-organisms was tested for each species, genus, order or sub-class, except for Nematoda, Suctoria and *Trochilia* sp., in which, due to lack of samples, the same organisms were used for the identification process, thus obtaining only estimate values.

5. RESULTS

The best result for the optimal radius trial-and-error method was obtained for $C_e = 6$ and, consequently, the location zones of the species, genera, orders or sub-classes for each axis were set to 6 times the standard deviation.

The correlations among the initial descriptors revealed that for the whole Metazoa or Protozoa (stalk included), high correlations (above 90%) were obtained between *Solidity* and $W_M W_B$ *ratio* and also among all the following descriptors: surface, perimeter, width, length and equivalent diameter. Furthermore, for the body descriptors (without stalk), alongside the above-mentioned, new correlated descriptors were found such as *compactness* with *robustness* and *solidity* with *largest concavity index*.

The variability percentage for the main discriminant functions (axes) in each class is shown in Table 3. The best correlations between the initial descriptors and the main discriminant functions in each class are shown in Table 4.

The values obtained for species, genus, order or sub-class identification are shown in Table 5. The organism's global recognition percentage attained a value of 75%; the misclassification 22% and 3.6% were not identified.

Table 3.   Variability percentage for the main axes in each class

							Class						
Axis	I	IIa	IIb	IIc	IId	IIe	IIIa	IIIb	IIIc	IIId	Dif 1	Dif 2	Dif 3
1	69	100	100	100	60	100	88	94	100	80	48	78	48
2	17				21		12	4		20	36	22	32
3	11				14			1			17		12
4	4				6								7

Table 4. Best correlated initial variables with the main discriminant functions for all the classes

Axis	I	IIa	IIb	IIc	IId	IIe	IIIa	IIIb	IIIc	IIId	Dif 1	Dif 2	Dif 3
								Class					
1	Sol_B FF_B $W_M W_B$	W_M D W_{MB}	W_M D D_B	W_B S S_B	W_M W_{MB} W_B	$Comp_B$ FF_B Rob_B	Rob Rob_B W_B	$Conv$ W P	CR FF_B L	D_B D D_{EDMB}	W_B Ecc_B $Comp_B$	W_M D S_B	$Comp_B$ Ecc $Comp$
2	PF $Comp$ $W_M W_B$				$Conv$ $W_M W_B$ $(W_M W_B)_B$	$Comp_B$ LCl_B L_B	L $(W_M W_B)_B$ D_{EDMB}			Ecc Ecc_B FF_B	$W_M W_B$ $Comp$ Sol	$Comp_B$ Ecc_B FF_B	D D_B S
3	D_{EDM} $Conv$ PF				FF_B FF $Comp$		Ecc $Comp$ W_M				W_{Stk} $Conv_B$ LCl_B		$W_M W_B$ Sol FF
4	$Conv$ W_{Stk} FF				S S_B $W_M W_B$								P FF $Comp_B$

The subscript $_B$ refers to the body descriptors.

Table 5. Recognition and error percentages of the Protozoa and Metazoa species

		Rec. %	Error %
Flagellate	*Peranema* sp.	92	0
Sarcodine	*Arcella* sp.	84	0
	Euglypha sp.	70	10
	Aspidisca cicada	80	27
	Carchesium sp.	57	63
	Epistylis sp.	58	40
	Euplotes sp.	86	16
	Litonotus sp.	88	15
	Opercularia sp.	40	44
	Suctoria	94*	0*
Ciliate	*Trachellophyllum* sp.	92	2
	Trithigmostoma sp.	82	29
	Trochilia sp.	92*	5*
	Vorticella	70	22
	Vorticella aquadulcis	78	28
	Vorticella microstoma	70	44
	Zoothamnium sp.	36	36
Metazoa	*Aelosoma* sp.	81	12
	Monogononta	52	4
	Digononta	71	18
	Nematoda	100*	2*

*Estimate.

The organism's global recognition percentage of the main Protozoa and Metazoa groups attained a value of 93%; the misclassification 3.2% and 3.5% were not identified. For the ciliated Protozoa groups, the organism's global recognition percentage attained a value of 90%, and the misclassification 6.2% and 3.3% were not identified.

In relation to mediocre or bad plant conditions due to high organic load, the organism's global recognition percentage attained a value of 72% and 8% misclassification, for low final effluent quality 79% and 9%, respectively, for low aeration 82% and 8.5%, respectively, and for the presence of nitrification 77% and 2%, respectively.

6. CONCLUSIONS

Analyzing the correlations among the initial morphological descriptors revealed that surface, perimeter, equivalent diameter, length and width were highly correlated (above 90%) and therefore a feature reduction could be performed on these descriptors without loss of significant variability. Such is the case also for Solidity and $W_M W_B$ *ratio*, body compactness and body robustness, and again with body solidity and body largest concavity index, which could be replaced by a single descriptor of each of these pairs.

The correlations between initial descriptors and the new discriminant functions (new axes) have shown that the most important variables within each group (metazoa, flagellate or fine stalked protozoa, prominent ciliated protozoa, non-prominent cilia or sarcodine Protozoa and thick stalked or elongated Protozoa and Metazoa) were the absolute widths and surfaces (or equivalent diameters). However, in order to distinguish amongst the different groups the most important descriptors were dimensionless variables such as solidity, form factor, $W_M W_B$ *ratio* and compactness.

The organism's global recognition percentage for species, genus, order or sub-class identification attained a value of 75%, which could be regarded as a fair recognition percentage due to the fact that some species are, morphologically speaking, hardly distinguishable. Only 4% of the organisms were not recognized by this procedure but the misclassification error attained a value of 22%, meaning that the overall recognition level could only be considered as median.

Regarding the identification of the main Protozoa and Metazoa groups (flagellate protozoa, ciliate protozoa, sarcodine Protozoa and Metazoa) as well as the ciliated Protozoa groups (carnivorous, crawling, free-swimming and sessile), the results could be considered as quite good. This is particularly important since ciliates are crucial for wastewater treatment plant diagnosis.

In relation to mediocre or bad plant conditions due to high organic load, the results proved to be acceptable, for both low effluent quality and low aeration the results could be considered satisfactory and, finally, good results were obtained regarding the occurrence of nitrification in the plant.

As a general conclusion, image analysis coupled with a multivariate statistical technique such as discriminant analysis proved to be a promising tool for assessing and monitoring the Protozoa and Metazoa populations in a wastewater treatment plant.

REFERENCES

Amaral AL, Baptiste C, Pons MN, Nicolau A, Lima N, Ferreira EC, Mota M, Vivier H. 1999. Semi-automated recognition of protozoa by image analysis. *Biotechnology Techniques* **13**: 111–118.

Einax JW, Zwanziger HW, Geiss S. 1997. *Chemometrics in Environmental Analysis*. VCH: Weinheim.

Fenchel T. 1999. *Ecology of Protozoa*. Springer-Verlag: Berlin.

Glasbey CA, Horgan GW. 1995. *Image Analysis for the Biological Sciences*. John Wiley & Sons: Chichester.

Golz C, Lange S. 2001. Analysis of activated sludge using pattern recognition techniques. *Proceedings of the 3rd International Specialized Conference on Microrganisms in Activated Sludge and Biofilm Processes*: Rome.

Jahn TL, Bovee EC, Jahn FF. 1999. *How to know the Protozoa*. Wm. C. Brown Company Publishers: Dubuque.

Madoni P. 1994. A sludge biotic index (SBI) for the evaluation of the biological performance of activated-sludge plants based on the microfauna analysis. *Water Research* **28**: 67–75.

Motta M da, Pons MN, Vivier H, Amaral AL, Ferreira EC, Roche N, Mota M. 2001. Study of protozoa population in wastewater treatment plants by image analysis. *Brazilian Journal of Chemical Engineering* **18**: 103–111.

Pons MN, Vivier H, Dodds J. 1997. Particle shape characterization using morphological descriptors. *Particle System Characterization* **14**: 272–277.

Richard M. 1989. *Activated Sludge Microbiology*. The Water Pollution Control Federation: Alexandria.

Russ CR. 1995. *The Image Processing Handbook*. CRC Press: Boca Raton.