CHEMICAL ENGINEERING TRANSACTIONS
VOL. 57, 2017
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Guest Editors: Sauro Pierucci, Jiří Jaromír Klemeš, Laura Piazza, Serafim Bakalis
Copyright © 2017, AIDIC Servizi S.r.l.
ISBN 978-88-95608- 48-8; **ISSN** 2283-9216
DOI: 10.3303/CET1757202

Stochastic Analysis of Aromatic Amino Acids Chromatographic Pulses

Marco A. Cremasco, Victor B. Moura, Rafael Scatena, Wilson M. Ferrari, Wesley H. Prieto

Process Engineering Department – Chemical Engineering School – University of Campinas
Albert Einstein Avenue, 500 – ZIP CODE: 13083-852 – Campinas - SP – Brazil
cremasco@feq.unicamp.br

The amino acids L-phenylalanine (Phe), L-tyrosine (Tyr) and L-Tryptophan (Trp) play an important role in the human body. L-Phenylalanine and L-tyrosine are precursors of several neurotransmitters, while Trp has been indicated as an aid for schizophrenic patients. These amino acids can be separated by ion exchange chromatography, and, usually, its description adopts differential mass balance equations for the stationary and mobile phases. The stochastic model can be an alternative to the traditional approach. Normally based on a microscopic view of chromatography, the random migration of a molecule is considered from a probabilistic aspect to describe the distribution function of the solute in the elution process. In the present work, the stochastic model in a fixed-bed adsorption has been studied by another point of view. The response of the adsorption column to the injection is measured as concentration vs. time at the exit of the column. Based on this, chromatographic pulses of aromatic amino acids are analyzed by Kolmogorov entropy theory. The rate of generation of information about concentration at the column exit can be identified with certain probability density function (PDF), and with a degree of the rate of generation of information of the system, characterized by the Kolmogorov entropy. The system tested is a fixed-bed, packed with poly-4-vinylpyridine cross-linked with divinyl benzene, and dilute aqueous solutions of NaCl and the aromatic amino acids Phe, Tyr, and Trp in two liquid flow rates. This work shows that is possible to reconstruct the solute effluent history using an appropriate PDF. In order to find the PDF that provides the best description of the experimental data, fourteen functions where fitted by the Least Squares Method. According to Kolmogorov-Smirnov test, Normal, Log-Normal, Logistic, Gamma, Pearson III and Beta models provided a satisfactory fit, and the dimensionless Kolmogorov entropy parameter depends on the affinity between solute and adsorbent, and the flow rate of carrier stream, implying its relation with dispersion phenomena and the nature of solute and adsorbent.

1. Introduction

L-phenylalanine (Phe), L-tyrosine (Tyr) and L-Tryptophan (Trp) are classified as aromatic amino acids due to the presence of a benzene ring in their structure. L-Phenylalanine is essential for many bodily functions and it is one of the few amino acids that can cross the blood-brain barrier and thus directly affect brain chemistry (Cremasco et al., 2001). L-Tyrosine is important to the structure of almost all proteins in the body. It and Phe are also the precursor of several neurotransmitters (Cremasco et al., 2001). L-Tryptophan cannot be synthesized by the organism and therefore must be part of its diet. L-Tryptophan is a precursor for serotonin, melatonin and niacin. Trp has been implicated as a possible cause of schizophrenia in people who cannot metabolize it properly. When improperly metabolized, it creates a waste product in the brain that is toxic, causing hallucinations and delusions. L-Tryptophan has then been indicated as an aid for schizophrenic patients (Cremasco et al., 2009).
Amino acids are separated usually by ion exchange chromatography, which is necessary to know the parameters of the system, such as isotherms, mass transfer parameters and axial dispersion, where the moment analysis of the response peaks is a common technique applied (Schneider and Smith, 1968). On the other hand, stochastic models are appropriate for systems where randomness is inherent to the process or where randomness is an useful simplification of complexity (Gogolek, 1998). Its mathematical models include

Please cite this article as: Cremasco M., Moura V., Scatena R., Ferrari W., Prieto W., 2017, Stochastic analysis of aromatic amino acids chromatographic pulses, Chemical Engineering Transactions, 57, 1207-1212 DOI: 10.3303/CET1757202

equations for the distribution functions of random variables where corresponding deterministic models have equations for the value of the variables. This approach started with Gidding and Eyring (1955), based on a microscopic view of chromatography, where the time of the adsorption and the number of the adsorption–desorption steps were used as parameters (Sepsy et al., 2014). The random migration of a single molecule is considered from a probabilistic way to describe the distribution function of the solute in the elution process (Baeza-Baeza and García-Álvarez-Coque, 2016). This model can be extended for several chromatographic methods such as partition, ion-exchange or size-exclusion chromatography (Sepsy et al., 2014).
In the present work, the stochastic model in a fixed-bed adsorption column has been studied by another point of view. The response of the adsorption column to the injection is measured as concentration vs. time at the exit of the column, which can be defined as RTD. Based on this, in the present work, chromatographic pulses of aromatic amino acids are analyzed by the Kolmogorov entropy view or chaotic approach. The information generation rate about concentration at the column exit can be identified with certain probability density function, and with a degree of the system's rate, characterized by dimensionless Kolmogorov entropy.

2. Stochastic analysis from a chaotic approach

In general, a dynamic system can be described by a phase space diagram, whose trajectories describe the evolution of the dynamical system from some known initial states through time (Lee and Lin, 2008). In the chaotic model it is necessary to indentify the phase space, which the information is immersed and associated, now, with a tracer concentration. The flow behavior through a system is conveniently achieved by determining the age distribution of the elements of the fluid in the exit stream or the residence time distribution (RTD) within the system. Usually, this technique involves the injection of a solute as a tracer at the inlet stream and the corresponding response of tracer concentration or pulse method. This assumption allows to consider the variation of trace concentration with a time step, Δt, resulting the information about local concentration at a certain time (m+1) as

$$c(t) = \{c(t_1), c(t_2), \ldots c(t_{m+1})\} \tag{1}$$

The multidimensional phase portraits can be constructed from the time series, Eq(1), applying the delay embedding theorem, proposed by Takens (1981). This dynamical system, considering c(t), it is embedding into a phase space of embedding dimension m to reconstruct the new coordinate system, f(t), by delay version of c(t) (Cassanelo et al., 1995; Faure and Korn, 1998). Each point f(t) within the phase space of embedding dimension m is (Lee and Lin, 2008)

$$f(t) = \{c(i\Delta t), c(i\Delta t + \tau), \ldots c(i\Delta t + (p-1)\tau)\} \tag{2}$$

with i = 0, 1, {m-[p-1]j}; τ = jΔt; j = 1, 2, 3; p is the dimension of vector of f(t); τ is the time delay. If m and τ are appropriately selected, the reconstructed space state has the same dynamic properties than the original one (Llop et al., 2012). According to Lee et al. (2002), discrete values of c, as shown in Eq(1), may be divided into bins, each with range c(t), and denoted by values c_1, c_2,…, c_n. For any data set, the probabilities of any value of c falling into a specific bin are $E(t_i)$. Hence, a set of probabilities $E(t_1)$, $E(t_2)$,…, $E(t_n)$ can be created from the original set, and the set of probabilities becomes $f(t_1)$, $f(t_2)$,…, $f(t_n)$. The rate of generation of information about concentration at column exit, in a certain time t, can be expressed as (Parmar and Majumder, 2013)

$$E(t) = E(t_0) + Kf(t), \text{ for } t \to \infty \tag{3}$$

E(t) in Eq(3) is the solute distribution with respect to time, where $E(t_0)$ is the information of the solute at initial time; f(t) is an information of the solute at time t; K is identified as dimensioneless Kolmogorov entropy. Then, in the pulse chromatography, a sample is injected into a carrier stream and passes through an adsorbent packed column. The response of the column to the injection is measured as concentration vs. time at the exit (Harlick and Tezel, 2000), which is defined as residence time distribution function, RTD (Mun et al., 2003) as

$$E(t) = \frac{c(t)}{\int_0^\infty c(t)dt} \tag{4}$$

The effluent history at column exit can be seen as E(t) proportional to f(t), given by some probability density function (PDF).The values of E(t) are obtained directly from experimental data, and $E(t_0)$ is assumed as initial baseline description. In this work, in order to find the PDF that provides the best description of the experimental data, fourteen functions where fitted by the Least Squares Method, and the Kolmogorov-Smirnov test was applied on the determination of the best fit.

3. Results and Discussion

The reference experimental system can be found in Cremasco et al. (2001), Cremasco and Wang (2009). The fixed-bed is a 1.5 cm I.D. x 12.5 cm column, packed with poly-4-vinylpyridine (PVP) cross-linked with divinyl benzene, with bed porosity, ε = 0.37, and particle porosity, ε_p = 0.55; average particle diameter, d_p = 0,036 cm. Dilute aqueous solutions, at 25 °C, with NaCl, the amino acids L-phenylalanine (Phe), L-tyrosine (Tyr) and L-Tryptophan (Trp). The liquid flow rates, Q, were 1.5 cm^3/min and 2.5 cm^3/min. The pulse concentration, C_0, and the affinity between solute and adsorbent, related with dimensionless Henry's constant for linear isotherm, k_p, are presented in Table 1.

Table 1: Solute properties (Cremasco et al., 2001; Cremasco and Wang, 2009)

	NaCl	Phe	Tyr	Trp
C_0 (g/L)	3.00	2.00	0.35	0.35
k_p (-)	0.00	1.95	3.23	13.70

The results of the residence time distribution, E(t), were obtained from experimental concentration profile. The following PFD models were tested: Normal, Log-Normal, Weibull, Exponential, Gamma, Pearson III, Beta (Singh, 1998; Walck, 2007), Logistic, Moyal, Gumbel, Cauchy, Chi-square, Rayleigh, Maxwell (Walck, 2007). In order to find the PDF that provides the description of the experimental data, Least Squares Method and the Kolmogorov-Smirnov test were applied to determine the best fit. The method of moments was applied to get initial values to the iterative Least Squares Method. The experimental first, second and third moments are associated with expectation value, variance and skewness given, respectively, by (Hahn and Shapiro, 1967)

$$\mu = \int_0^\infty tE(t)dt \tag{5}$$

$$\sigma^2 = \int_0^\infty (t-\mu)^2 E(t)dt \tag{6}$$

$$\gamma^3 = \int_0^\infty (t-\mu)^3 E(t)dt \tag{7}$$

whose values are presented in Table 2.

Table 2: Moments parameters from E(t) experimental

	NaCl		Phe		Tyr		Trp	
Parameters/Q	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
μ (min)	11.064	6.542	19.367	11.523	25.425	15.241	94.535	53.894
σ^2 (min^2)	4.826	1.753	29.543	14.622	65.647	33.547	875.236	423.139
γ^3(min^3)	8.076	0.500	89.220	42.185	324.267	164.890	10927.687	5083.464

The Kolmogorov-Smirnov test was used, and the maximum deviation (D) among accumulative distributions from each model and experimental data were compared to the critical deviation value, D_C, with 1% as significance level (O' Connor and Kleyner, 2012). The Gumbel, Weibull, Cauchy, Moyal, Rayleigh, Chi-square, Exponential and Maxwell models do not attend the criteria in all conditions studied in this work, while Normal, Log-normal, Logistic, Gamma, Pearson III and Beta distributions attend these criteria, as shown in Table 3.

Table 3: Maximum deviation (D) of the Kolmogorov-Smirnov test

	NaCl		Phe		Tyr		Trp	
Q (cm^3/min)	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
D_C	0.1768	0.2755	0.1261	0.1533	0.1507	0.1811	0.0791	0.0910
Normal	0.0230	0.0082	0.0293	0.0395	0.0335	0.0442	0.0459	0.0489
Log-normal	0.0336	0.0290	0.0173	0.0103	0.0187	0.0115	0.0204	0.0274
Logistic	0.0210	0.0272	0,0452	0.0547	0.0486	0.0567	0.0603	0.0588
Gamma	0.0239	0.0178	0.0050	0.0084	0.0028	0.0074	0.0160	0.0157
Pearson III	0.0239	0.0175	0.0613	0.0218	0.0225	0.0074	0.0160	0.0157
Beta	0.0262	0.0186	0.0953	0.0281	0.0622	0.0315	0.0274	0.0205

From Table 3, it is observed that Logistic for 1.5 cm^3/min and Normal for 2.5 cm^3/min can be used to fit NaCl tests. Gamma function provides the best fit for aromatic amino acids. The Figures 1 up to 4 show the comparison between the experimental results of E(t) and the best fit. For aromatic amino acids, the curves present high level of asymmetry, characterized by skewness (see Table 3 for γ^3), whose values increases with the affinity between solute and adsorbent, as shown in Table 1, considering k_p values on it.

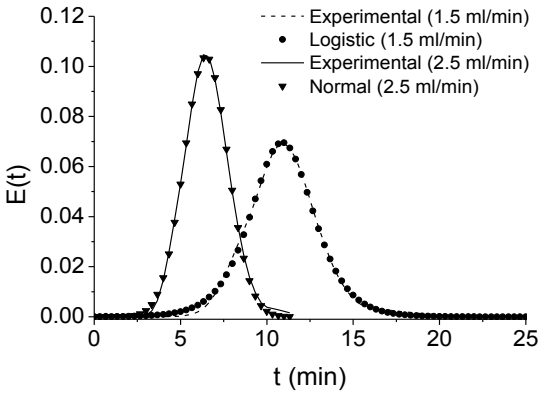

Figure 1: Chromatograms for NaCl

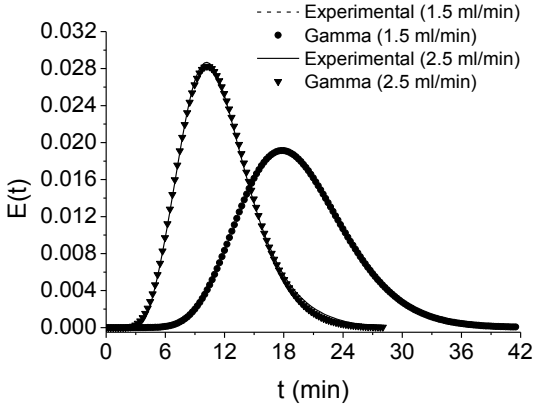

Figure 2: Chromatograms for Phe

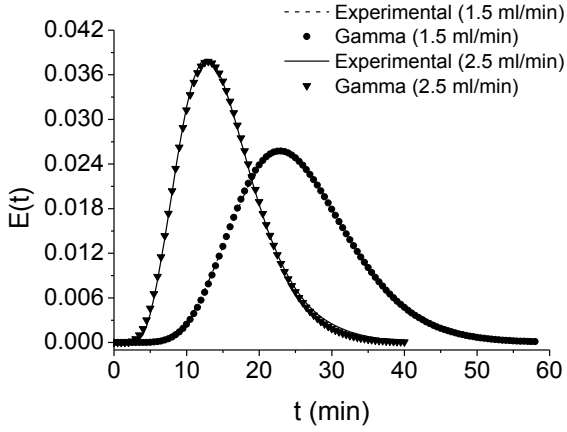

Figure 3: Chromatograms for Tyr

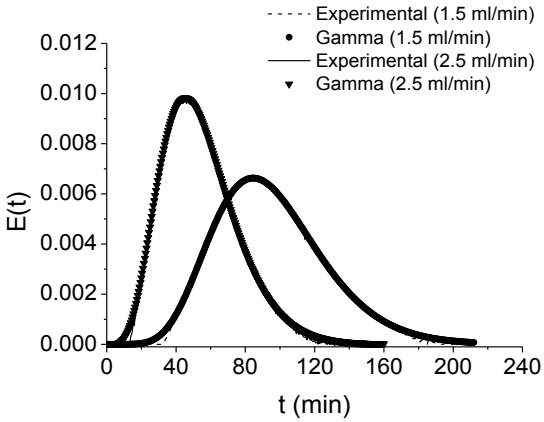

Figure 4: Chromatograms for Trp

Equations 8 up to 10 are the Logistic, Normal, and Gamma distributions, respectively, and the Table 4 presents the parameters for these distributions in which case with best fit.

$$f(t) = \frac{1}{k} exp\left(\frac{t-\alpha}{k}\right)\left(1 + exp\left(\frac{t-\alpha}{k}\right)\right)^{-2} \tag{8}$$

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}\right) \tag{9}$$

$$f(t) = \frac{a(at)^{b-1} exp(-at)}{\Gamma(b)} \tag{10}$$

An essential parameter that was adjusted through Eq(3) to provide the Least Square Deviation is the parameter K, whose values for PDF fit are exposed at the Table 5.

Table 4: Distribution parameters of E(t) modelled

	NaCl		Phe		Tyr		Trp	
Parameters/Q	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
μ (min)	-	6.472	-	-	-	-	-	-
σ^2 (min^2)	-	1.268	-	-	-	-	-	-
α (min)	10.892	-	-	-	-	-	-	-
k (min)	1.204	-	-	-	-	-	-	-
a (min^{-1})	-	-	0.672	0.846	0.390	0.483	0.092	0.110
b (-)	-	-	12.992	9.656	9.909	7.246	8.781	6.021

Table 5: Kolmogorov entropy, K (-)

	NaCl		Phe		Tyr		Trp	
PDF/Q	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
Normal	0.3256	0.3308	0.2464	0.2445	0.4930	0.4863	0.5033	0.4999
Log-normal	0.3265	0.3320	0.2507	0.2518	0.5038	0.5045	0.5149	0.5162
Logistic	0.3356	0.3407	0.2540	0.2525	0.5087	0.5035	0.5189	0.5177
Gamma	0.3260	0.3313	0.2489	0.2488	0.4990	0.4966	0.5092	0.5080
Pearson III	0.3260	0.3325	0.2647	0.2527	0.4892	0.4966	0.5092	0.5080
Beta	0.3245	0.3269	0.2702	0.2438	0.5212	0.4842	0.4975	0.4981

As can be seen from Table 5, the values of K do not provide a significant variation with probability density functions. If one observes in this table, it is possible to see the K value practically remains constant with liquid flow rate variation for certain solute, with a very slightly tendency to decrease with liquid flow increment. This tendency can be notice using the best fit (Gamma model) for aromatic amino acids. This effect can be associated with chromatogram band spread or variance (σ^2), which is larger for lower liquid flow rate, as shown in Table 3. In terms of entropy, it is can be related with Shannon's information. According with Santamaría-Bonfil et al. (2016), its revels a complexity that measures how a system's predictability changes in terms of the PDF parameters. In case of Normal distribution, the Shannon entropy, in bits, is given by (Santamaría-Bonfil et al., 2016)

$$H = \frac{1}{2}\ell og_2\left(2\pi e\sigma^2\right) \tag{11}$$

Then, if one utilizes the values of variance (σ^2) present in Table 2 into Eq(11), it is possible to do a comparison between Shannon entropy and dimensionless Kolmogorov entropy, from Normal model in Table 5, as shown in Table 6. However, NaCl presents opposite tendency for K when compared with amino acids. This behaviour, for entropy K, indicates its dependence on the variance, expectation value and the skewness, denoting the influence of nature of the substances involved in the phenomena, which it is strongly dependent on affinity between solute and adsorbent, and the initial concentration of the solute (see Table 1).

Table 6: Comparison between Shannon entropy (H) and dimensionless Kolmogorov entropy (K)

	NaCl		Phe		Tyr		Trp	
PDF/Q	1.5	2.5	1.5	2.5	1.5	2.5	1.5	2.5
H	1.4352	1.1483	2.1176	1.8429	2.3930	2.1194	3.3521	3.0772
K	0.3256	0.3308	0.2464	0.2445	0.4930	0.4863	0.5033	0.4999

4. Conclusion

This work shows that is possible to reconstruct the solute effluent history at an adsorption column exit, using an adequate PDF model. For NaCl, diluted in the mobile phase, the best fits are symmetric models, such as the Normal and Logistic distributions. However, for the amino acids studied in this work, the Gamma model describes the adsorption, and it presents asymmetry that increases with affinity between solute and adsorbent.

The dimensionless Kolmogorov entropy (K), obtained from different PDF models, seemed to converge basically to a same value, preserving some relationship with Shannon entropy (H) for amino acids. While H depends only on the variance (σ^2), the parameter K depends on the variance, expected value (μ) and the

skewness (γ^3). Those last parameters show how the initial concentration and the solute-adsorbent affinity influence the dimensionless Kolmogorov entropy (K).
Finally, this new approach to chromatographic analysis may be an alternative to classical techniques for the obtainment of the necessary parameters to describe the adsorption phenomena in chromatographic system, and supports the affirmation of Gogolek (1998), that says “stochastic models are appropriate for systems where randomness is an useful simplification of complexity”.

Acknowledgements

The authors would like to thank CAPES and CNPq (Proc. 01/08101-3), and to DEPro/FEQ/UNICAMP for the financial support.

References

Baeza-Baeza, J.J., García-Álvarez-Coque, M.C., 2016, General solution of the extended plate model including diffusion, slow transfer kinetics and extra-column effects for isocratic chromatographic elution, *Separation*, 3, 11,<www.mdpi.com/journal/separations> accessed 14.11.2016.

Cassanello, M., Larachi, F., Marie, M-N., Guy, C., Chaouki, J., 1995, Experimental characterization of the solid phase chaotic dynamics in three-phase fluidization, *Ind. Eng. Chem. Res*., 34, 2971-2980.

Cremasco, M.A., Hritzko, B.J., Xie, Y., Wang, N.-H.L., 2001, Parameters estimation for amino acids in a fixed bed column. *Brazilian J. Chem. Eng*., 18, 2, 181 – 194.

Cremasco, M.A., Starquit, A., Wang, N-H.L., 2009, Separation of L-tryptophan present is an aromatic amino acids mixture in a four-column simulated moving bed: experimental and simulation studies, *Brazilian J. Chem. Eng*., 26, 3, 611 – 618.

Faure, P., Korn, H., 1998, A new method to estimate the Kolmogorov entropy from recurrence plots: its application to neuronal signals, *Physica D*, 122, 265-279.

Giddings, J.C., Eyring, H.J., 1955, A molecular dynamic theory of chromatography, *Phys. Chem.*, 59, 416–421.

Gogolek, P.E.G., 1998, *Mathematical modelling of fluctuations in fluidized bed combustion.* Ph D. Thesis, Queen´s University, 198p.

Hahn, G. J., Shapiro, S. S., 1967, *Statistical models in engineering*, Wiley, New York, 376p.

Harlick, P.J.E., Tezel, F.H., 2000, A novel solution method for interpreting binary adsorption isotherms from concentration pulse chromatography data, *Adsorption*, 6, 3, 293 – 309.

Lee, S.H., Kim, S.D., Park., S.H., 2002, Statistical and deterministic chaos analysis of pressure fluctuation in a fluidized bed of polymer powders. *Korean J. Chem. Eng.*, 19, 6, 1020 – 1025.

Lee, C.-K., Lin, S.-C., 2008, Chaos in air pollutant concentration (APC) time series, *Aerosol and Air Quality Research*, 8, 4, 381-391.

Llop, M.F., Jand, N., Galluci, K., Llauró, F.X., 2012, Characterizing gas-solid fluidization by nonlinear tools: chaotic invariants and dynamic moments, *Chem. Eng. Sci.,* 71, 252 – 263.

Mun, S., Xie, Y., Wang, N.-H., 2003, Residence time distribution in a size-exclusion SMB for insulin purification. *AIChe J.,* 49, 8, 2039 – 2058.

O’Connor, P., Kleyner, A., 2012, P*ractical reliability engineering*. 5th Edition. New York: John Wiley & Sons.

Parmar, S., Majumder, K., 2013, A stochastic analysis of liquid mixing in bubble column, *American J. Fluid Dynamics*, 3, 3, 75 – 79.

Santamaría-Bonfil, G., Fernández, N., Gershenson, C., 2016, Measuring the complexity of continuous distributions, *Entropy*, 18, 72, <www.mdpi.com/journal/entropy> accessed 03.12.2016.

Schneider, P., Smith, J., 1968, Adsorption rate constant from chromatography, J. *AIChE J*., 14, 5, 762 – 771.

Sepsey, A., Bacskay, I., Felinger, A., 2014, Polydispersity in size-exclusion chromatography: A stochastic approach, *J. Chromatogr*. A ,1365,156–163.

Singh, V.P., 1998, *Entropy-based parameter estimation in hydrology*, Dordrecht: Springer. 368p.

Takens, F., 1981, *Detecting strange attractors in turbulence*, in Lecture Notes in Mathematics, Eds. D.A. Rand and L-S. Young, Springer Verlag, 898.

Walck, C., 2007, *Handbook on statistical distributions for experimentalists,* Internal Report. Universitet Stockholm, 188p.