Emergency Department Overcrowding Detection by a Multifractal Analysis

Sid'Ahmed Emine, Guillaume Bouleux, H Haouba, Eric Marcon

Emergency Department Overcrowding Detection by a Multifractal Analysis

S.A. Emine *, **G. Bouleux** *, **H. Haouba** **, and **E. Marcon** *

* *Univ Lyon, INSA Lyon, UJM-Saint Etienne, DISP, EA 4570, F-69621 Villeurbanne, France (e-mail: {sidahmed.emine,guillaume.bouleux,eric.marcon}@insa-lyon.fr).*
** *Nouakchott University, Mauritania*

Abstract
The observation of data which counts the arrivals at the ED in scales of hours or days reveals that these arrivals are characterized by the phenomenon of burstiness. The burstiness is a phenomenon that appears in the majority of Emergencies due to a batches of arrivals coming to the ED in a small time interval over a wide scales. The prediction of congestion in ED caused by the batched arrivals seems important to the medical staff. The proper modeling data like arrivals process that has this scaling property is relevant with the self similar time series. The goal of our study is to establish that the time series of patient flow displays fractal behavior, whose quantitative characteristics vary with time. To discover whether a Patients flow has begun to congest, we analyze the time series of patients arrivals data with a Multi- Fractal Detrended Fluctuation (MF-DFA) algorithm.

Keywords: Fractals Systems, Decision Support Systems, Detection, Signal Analysis, Data Flow Analysis

1. INTRODUCTION

Overcrowding in the Emergency Department (ED) is a worldwide problem impairing the ability of hospitals to offer emergency care within a reasonable time frame Lin et al. (2014); Bouleux et al. (2015). It is a reflection of several internal or external effects such as the capacity of the ED or the stream of arrivals. During the last several years, many research papers have been devoted to the topic of patient flow or stream of arrivals in the ED. One can see for example, the following manuscript and the references that are quoted in it Armony et al. (2015). The work proposed in Monte et al. (2002) analyze the ED admissions time series and search the distribution that best models it. In her thesis, Jlassi (2009) improved the performance of patient logistics flows in a hospital emergency department by applying to the hospital the analysis methods and the resolution tools from the manufacturing field. By using the time series, specifically an ARIMA model, Kadri et al. (2014) modeled the daily flows of patients for the pediatric emergency of Lille, France, Hospital. Finally, the most commonly encountered model for the hospital emergency arrivals is definitely the Poisson process. This modeling is for example introduced in Lin et al. (2014) where the authors have estimated the waiting time of multi-priority emergency patients using the Fast track system and a queueing model. They modeled the arrival process of demand accessing the ED by a Poisson process so. In Koizumi et al. (2005), the authors divide the whole day into several segments such that the arrival rate in each segment can be assumed to be constant and affirme that the arrival process of emergency patients in most time segments can be modeled as a Poisson process with varying rates. Nevertheless, these approaches turn out not to work for ED; the pattern of arrivals found by the staff of the emergency ward service suggested that this model was inadequate Monte et al. (2002). Indeed, the time series of arrivals in ED exhibit burstiness over a wide range of time scales due to pikes of arrivals process batched over hours, days or weeks.

A process which takes into account features related to bursts is known to be a self-similar process with the Long Range Dependence (LRD) property Radev et al. (2010). Such a process has a statistical structure that repeats itself on subintervals or time scales Espen (2012). This statistical characteristic has shown its good performances in network traffic congestion prediction Kim et al. (2006) and some works initiated the reflexion on ED admissions modeling by self-similar processes Monte et al. (2002). If the self-similarity propose a global scaling description of the process, the fractality and multifractily are properties which focus on local scaling behavior and appears as key features for congestion detection as well Feng et al. (2018). Consequently, we propose in this paper a Multifractal analysis of the ED arrivals time series with the aim of being able to detect the burstiness and congestion caused by this phenomenon.

We present next the basic notions of our study and recall the fundamental of the self-similarity properties for a time series. Section 3 is dedicated to the local study of time series exhibiting scaling by the multifractal analysis. We illustrate then this analysis on real ED data admissions flow, in view to depict the congestion caused by the burst of arrivals at the ED. Finally, a conclusion is drawn in the last section.

2. FRACTAL ANALYSIS

2.1 Notion Of Self-Similarity

A time series $X_t, t \in \mathbb{N}$ with mean $\mu(t)$, variance $\sigma^2(t)$ and autocorrelation function $\rho(k), k \geq 0$ is called a stationary process if

$$E[X_t] = \mu(t) = \mu \; \forall \, t \in \mathbb{N} \tag{1}$$

with $E[.]$ the mathematical expectation, and

$$\text{cov}(X_t, X_{t-j}) = E[(X_t - \mu)(X_{t-j} - \mu)] = \gamma_j, \forall \, t, \forall \, j. \tag{2}$$

For each $m = 1, 2, \ldots$, constructs from the covariance stationary process X_t a new covariance stationary time series $X^{(m)} = \{X_k^{(m)} : k = 1, 2, ...\}$ with corresponding autocorrelation function $\rho^{(m)}$, given by

$$X^{(m)} := \frac{1}{m}(X_{km-m+1} + ... + X_{km}), k \geq 1. \tag{3}$$

X_t is said to exhibit self-similar (or asymptotically self similar) property, if for all $t = 0, 1, 2, ...,$

$$X_t \stackrel{d}{=} m^{(1-H)} X^{(m)} \tag{4}$$

holds $\forall m$ (or as $m \to \infty$), where $\frac{1}{2} < H < 1$ and $\stackrel{d}{=}$ is the finite dimensional law equality. X_t is said to exhibit (exactly) second-order self-similar property with self-similarity parameter H if $m^{(1-H)} X_t^{(m)}$ has the same autocorrelation function as X_t, $\forall \, m$ and $\forall \, t$.

$$\text{cov}(X^{(m)}) = \sigma^2 m^{2(H-1)} \; and \; \rho^{(m)}(k) = \rho(k), \;\; k \geq 0. \tag{5}$$

In other words, X_t is exactly second-order self-similar if the aggregated processes $X(m)$ are indistinguishable from X_t with respect to their first and second order properties Strzałka et al. (2012). In addition, X_t is said to exhibit asymptotically second order self-similar with self-similarity parameter H (H-a.s.o.s.s.) if $\forall \, k$,

$$\lim_{m \longrightarrow \infty} \rho^{(m)}(k) = \rho(k) \tag{6}$$

2.2 Notion of Long Range Dependence

The Long Range Dependence (LRD) as a phenomenon was first observed by Hurst Hurst (1956) when he studied the flow of water in the Nile river. Many works have emerged since but Mandelbrot (1965) is certainly the reference papers. Mathematically, a stationary process X_t is said to exhibit short range dependence if its autocorrelation function $\rho(k)$ satisfies

$$\sum_{k=1}^{\infty} \rho(k) < \infty. \tag{7}$$

Conversely, if the autocorrelation function $\rho(k)$ is not summable, the process X_t is said to exhibit Long Range Dependence such as

$$\sum_{k \geq 0} |\rho(k)| = \infty. \tag{8}$$

An equivalent definition is often proposed regarding whether the autocorrelation behaves like a power function decaying to zero hyperbolically

$$\rho(k) \sim L_\rho(k) k^{-\alpha} \;\; 0 < \alpha < 1 \tag{9}$$

with $L_\rho(k)$ a slowly varying function Bingham et al. (1989). Lastly, the Karmata definition of this class of functions is the set of functions such that the following equation holds

$$\lim_{x \longrightarrow \infty} \frac{L_\rho(k)(ax)}{L_\rho(k)(x)} = 1. \tag{10}$$

2.3 The connection of Self similarity and LRD

The connection between Self-similarity and long memory phenomena was studied extensively by Samorodnitsky (2006). It was shown that asymptotically second order self-similarity property implies LRD propertyLazarou et al. (2009); Park and Willinger (2000). The autocorrelation of a Self Similar process with Hurst exponent H and finite second order momentEmbrechts and Maejima (2000) is then given by

$$\rho(k) = \frac{1}{2}((k+1)^{2H} + (k-1)^{2H} - 2k^{2H}). \tag{11}$$

From this identity we straightforwardly obtain

$$\rho(k) \stackrel{\infty}{\approx} H(2H-1)k^{-2(1-H)} \tag{12}$$

in asymptotic sense so. If $0.5 < H < 1$, then the process admits LRD, we say it is persistent. On contrary, when $0 < H < 0.5$ it is short range dependence and is qualified as antipersistent Park and Willinger (2000). The Hurst exponent H is also understood in that context as a measure of the decay speed of the autocorrelation function tail. It captures then both, the intensity of LRD Vemuri (2005) and the degree of self-similarity Cervantes-De la Torre et al. (2013).

2.4 The Fractal Analysis and the ED arrivals process

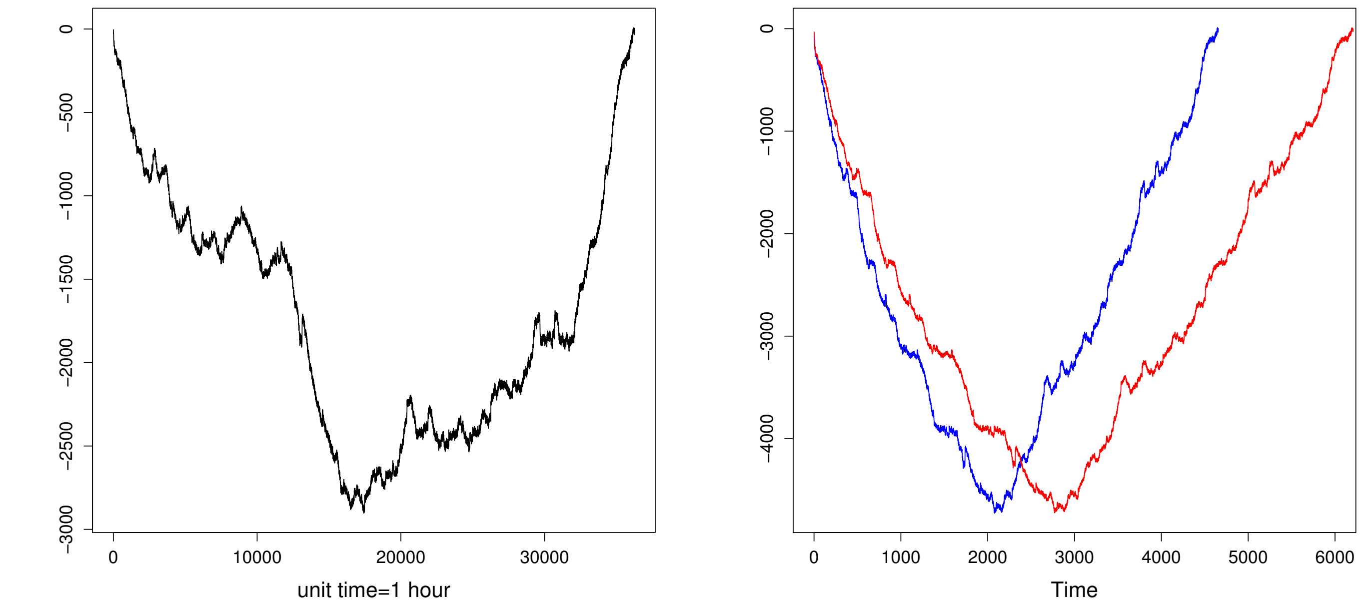
The burstiness phenomenon we have mentioned in the introduction is very much related to the Hurst exponent. To describe quantitatively the burstiness due to the batches of arrivals processes, which depend on the time scale, Gusella Gusella (1991) propose the Index of Dispersion for Counts (IDC). It is defined as

$$IDC(N) := \frac{var(\sum_{t=1}^{N} X_t)}{E(\sum_{t=1}^{N} X_t)} \tag{13}$$

where X_t is the self-similar arrivals process at the ED. In the case of persistence for X_t, the IDC varies exponentially with the interval length and follows the relation

$$IDC(N) \sim c \, N^{2H-1} \tag{14}$$

$$\log(IDC) \sim (2H-1)\log(N) + const. \tag{15}$$



(a) (b)

Figure 1. The cumulated time series of Arrivals, (a) time series aggregated over 1 hour (black), (b) time series aggregated over 6 hours (red) and 8 hours (blue).

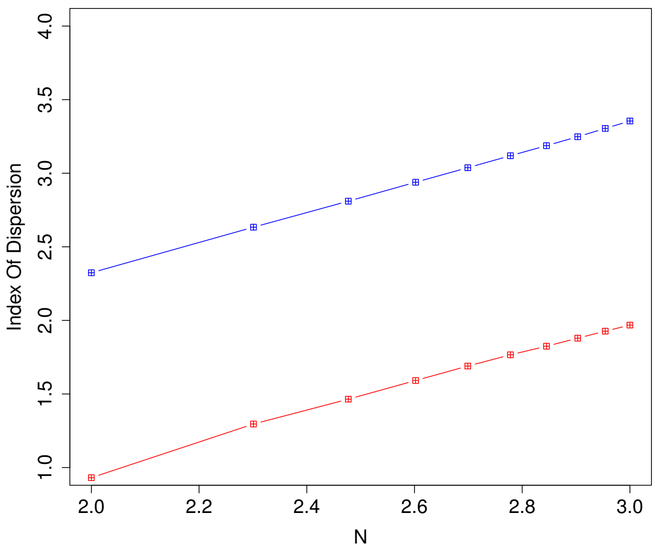


Figure 2. Plot in logarithmic scales of the Index of Dispersion of Counts (IDC) for the ED arrivals process with respect to hours (red) and days (blue) scales.The Hurst exponent of both is H=0.54

The IDC can also served as a statistical method for testing the self-similarity as illustrated by Figure 2. This plot gives the picture of IDC for the ED arrivals for different aggregation sizes N. The Figure indicates that the hurst exponent for the ED arrivals process is around $H \sim 0.54$.

Another popular approach adopted for self-similarity characterization is the Detrended Fluctuation Analysis (DFA), originally introduced by Peng Peng et al. (1995) and has

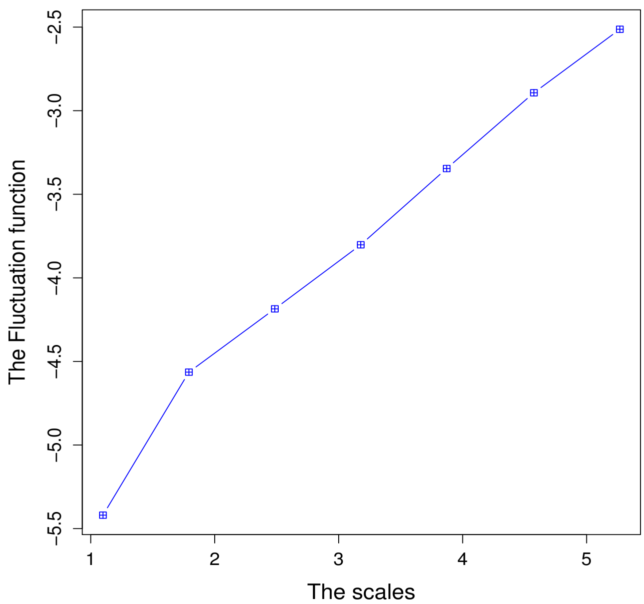


Figure 3. The DFA-Method to estimate the Hurst exponent of ED arrivals stream.

been established as an important method to reliably detect long range correlations. We performed such an analysis on ED admissions and the results are depicted along with the Figure 3. This Figure shows a small detrend due to the long range correlations and the approximated slope of the line allows us to consider that the ED arrivals time series displays fractal behavior, whose quantitative characteristics vary with time. Such behavior is called multifractal.

3. THE MULTIFRACTAL ANALYSIS

The multifractal framework has become a major concept and tool involved in practical analysis of scaling in data. The multifractal (MF) framework relates the scaling properties to the local regularity properties Lashermes et al. (2008). In general, multifractality in time series is due to different LRD of the small and large fluctuations in the time seriesKantelhardt et al. (2002).

A self-similar stochastic process $\{X_t : t \in [0,T]\}$ is multifractal Thompson and Wilson (2016) if it satisfies

$$E[|X_t|]^q = c(q)t^{\tau(q)+1}, \quad \forall\, t \in \mathcal{G} \text{ and } q \in \mathcal{Q} \tag{16}$$

where $0 < T < \infty$, $\mathcal{G}$ and $\mathcal{Q}$ are open intervals on $[0,T] \subset G$ and $[0,1] \subset \mathcal{Q}$. The function $\tau(q)$ is called the scaling function Calvet and Fisher (2002) of the multifractal process.

The Multifractal Formalism is based on the calculation of two sets of coefficients: the Holder exponents h which quantifies the local regularity of a process X_t and the MultiFractal Spectrum that quantifies the multifractality of the process Xiong et al. (2012).

3.1 The Holder Exponent and the Spectrum of Singularities

The local regularity of the sample path of stochastic process X_t is commonly studied via the notion of pointwise Holder exponent. Recall that X_t is assumed to be locally bounded, *i.e* it belongs to $C^\alpha(t_0)$, $t_0 \in \mathbb{R}$ with $\alpha \geq 0$, if there exists a constant $C > 0$ and a polynomial P satisfying $deg(P) < \alpha$ and such that, in a neighborhood of t_0

$$|X_t - P(t-t_0)| \leq C|t-t_0|^\alpha. \tag{17}$$

The Holder exponent of X_t at t_0 is therefore defined as

$$h(t_0) = \sup\{\alpha : X \in C^\alpha(t_0)\}. \tag{18}$$

The value of the Holder exponent is interpreted as follows: the closer to 0 $h(t_0)$ is, the more irregular (or singular) at point t_0 the function is. In contrast, a larger value of $h(t_0)$ is related to a smoother (more regular) behavior at t_0 Lashermes et al. (2008). The points t_0 associated with a specific Holder exponent value h are distributed on interwoven fractal subsets $E(h)$ such as

$$\mathbb{E}(h) = \{\alpha : h(t_0) = \alpha\}. \tag{19}$$

The fluctuations, along time t of the Holder exponent h are usually described through the singularity (or multifractal) spectrum, labelled $D(h)$ and defined as the Hausdorff dimension Przytycki and Urbański (1989) of the set of points $E(h)$ where the Holder exponent takes the value h, this leads then to

$$D(h) = \dim_H(\mathbb{E}(h)). \tag{20}$$

The MultiFractal analysis aims at characterizing or derive the spatial distribution of Holder exponents with singularity spectrum D(h) Lashermes et al. (2008). Some relations between the two sets of multifractal scaling exponent have been established.

$$\tau(q) = qh(q) - 1 \tag{21}$$

with $h(q)$ the generalized Hurst exponent or the q-order Hurst exponent ($F_q(s) \sim s^{h(q)}$) and $\tau(q)$ the scaling exponent and ($Z_q \sim s^{\tau(q)}$) the generalized multifractal dimensions. We obtained consequently the following relation

$$D(q) = \frac{\tau(q)}{q-1} = \frac{qh(q)-1}{q-1}. \tag{22}$$

Next, the spectrum of singularity is related to the scaling exponent via the so called Legendre transform

$$\alpha = \tau'(q) = h(q) + qh(q), \tag{23}$$

and

$$f(\alpha) = q\alpha - \tau(q) = q[\alpha h(q)] + 1 \tag{24}$$

where α is the Holder exponent, while $f(\alpha)$ denotes the dimension of the subset of the series that is characterized by α. When Holder exponent takes a unique value H at every point t_0, the multifractal process becomes a monofractal one and we obtain

$$f(\alpha) = \begin{cases} 1 \; if \;\; \alpha = H \\ 0 \; if \;\; \alpha \neq H \;\; and \;\; \alpha > 0 \end{cases} \tag{25}$$

3.2 The MF-DFA Method

We propose here to introduce the Multifractal Detrend Fluctuation Analysis, originaly introduced by Kantelhardt Kantelhardt et al. (2002) as a generalization of the DFA method. It has a practical interest in giving singularities information, reason why we give its implementation principles in what follows. Suppose that $X_t, t = 1, \ldots T$ is a series of the arrivals by days with mean μ. Timely integrate the detrended series X_t to obtain

$$Y_t = \sum_{k=1}^{t}(X_k - \mu) \tag{26}$$

Next, divide the profile Y_t into $N_s = \lfloor N/s \rfloor$ (with $\lfloor . \rfloor$ the integer part) non-overlapping segments of equal length s. Then, calculate the local trend for each of the $2N_s$ segments by a least-square fit of the series and determine the variance with

$$F^2(\nu, s) := \frac{1}{s}\sum_{i=1}^{s}\{Y[(\nu-1)s+i] - y_\nu(i)\}^2, \qquad \nu = 1, ..., N_s$$

and

$$F^2(s, \nu) := \frac{1}{s}\sum_{i=1}^{s} Y[N-(\nu-N_s)s+i] - y_\nu(i)^2, \qquad \nu = N_s+1, ..., 2N_s$$

where $y_\nu(i)$ is the fitting polynomial in segment ν. Compute the average over all segments ν to obtain the q-th order fluctuation function

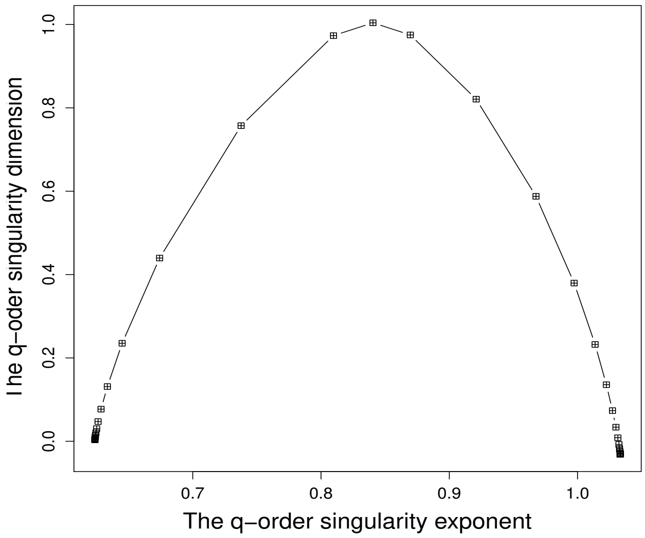


Figure 4. The MF spectrum of The time series arrivals in ED ever days estimated by the MF-DFA.

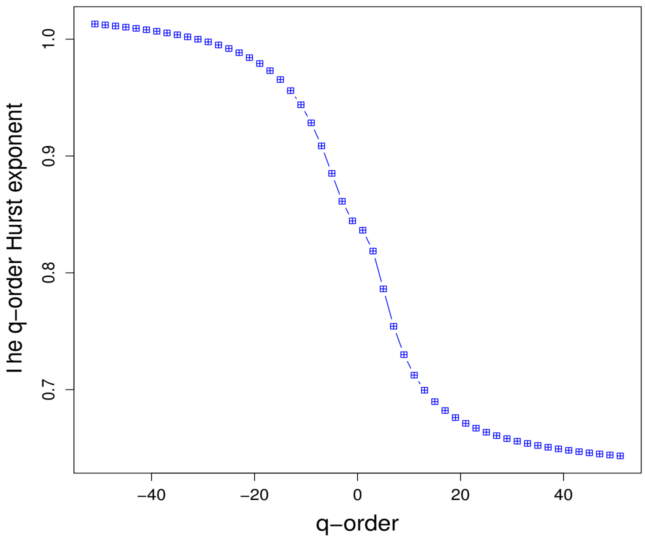


Figure 5. The q-order Hurst exponent of Arrival in ED estimated by the MF-DFA.

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{\nu=1}^{2N_s} (F^2(s,\nu))^{\frac{q}{2}} \right\}^{\frac{1}{q}}. \tag{27}$$

If the time series X_t has a long-range power-law correlated, then $F_q(s)$ increases for large values of s as a power-law which means that $F_q(s) \sim s^{h(q)}$.

3.3 The Arrivals Process and the MultiFractal Analysis

In this section we analyse the time series of ED arrivals of Saint Etienne University Hospital (CHU), France, from 06/01/2013 to 08/31/2017. To estimate the q-order Hurst exponent of Figure 5, the local Hurst exponent Figure 6 and the the MultiFractal Spectrum (figure 4), we use

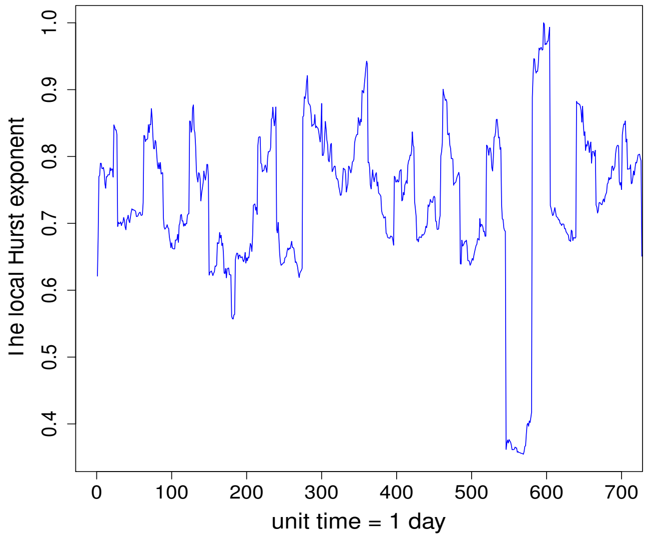


Figure 6. The local Hurst exponent of the arrivals by day estimated by the MF-DFA.

the MF-DFA method Kantelhardt et al. (2002). The MultiFractal Spectrum (MFS) of the time series is plotted in Figure 4. The MFS is a downward concave function with maximum value 1 and minimum value around 0.45. This particular shape justifies the multifractal approches. To depict the congestion in the stream, we observe the local Holder exponent through Figure 6. A cursory glance at the Holder exponent plot which depicts the temporal dynamics, shows that the arrivals time series admits a high variability of corresponding so to fractal behavior. However, it is not possible to distinguish a regularly regime of tensions. The average of arrivals at ED is 136 patients a day. Overall, the local Hurst exponent oscillates between 0.6 and 0.8 and when it exceeds 0.8 we notice that the arrivals are globally below the daily average of arrivals. On the other hand, when it is down 0.6, the system registers peaks of arrivals more than average.

4. CONCLUSION

Led by recent statistical analysis taking into account burstiness effects in view to early predict congestion on computers networks and traffic networks, we have proposed in this work to apply the underlying fractal methods to the stream of patients' arrivals at an Emergency Department (ED). The stream of arrivals at ED is characterized by a high variability not only on the dynamics of the process but also on strong irregularities of its scaling exponent. To fully characterize the arrivals flow, the fractal methods appeared to be insufficient and MultiFractal Analysis (MFA) is thus recommended. MFA only based on patients' arrivals at the ED gives information on the regimes of the arrivals, but it does not reveal all the system characteristics, particularly the burstiness induced by capacity constraints. Thus, MFA could be giving more information if the Length Of Stay (LOS) of patients at the ED were considered and used in conjonction with the patients' arrivals.

REFERENCES

Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y.N., Tseytlin, Y., and Yom-Tov, G.B. (2015). On patient flow in hospitals: A data-based queueing-science perspective. an extended version. *Stochastic Systems*, 5(1), 146–194.

Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1989). *Regular variation*, volume 27. Cambridge university press.

Bouleux, G., Marcon, E., and Mory, O. (2015). Early index for detection of pediatric emergency department crowding. *IEEE journal of biomedical and health informatics*, 19(6), 1929–1936.

Calvet, L. and Fisher, A. (2002). Multifractality in asset returns: theory and evidence. *The review of economics and Statistics*, 84(3), 381–406.

Cervantes-De la Torre, F., Gonzalez-Trejo, J., Real-Ramirez, C., and Hoyos-Reyes, L. (2013). Fractal dimension algorithms and their application to time series associated with natural phenomena. In *Journal of Physics: Conference Series*, volume 475, 012002. IOP Publishing.

Embrechts, P. and Maejima, M. (2000). An introduction to the theory of self-similar stochastic processes. *International Journal of Modern Physics B*, 14(12n13), 1399–1420.

Espen, I. (2012). Introduction to multifractal detrended fluctuation analysis in matlab. *Frontiers in Physiology*, 3, 141. doi:10.3389/fphys.2012.00141.

Feng, S., Wang, X., Sun, H., Zhang, Y., and Li, L. (2018). A better understanding of long-range temporal dependence of traffic flow time series. *Physica A: Statistical Mechanics and its Applications*, 492, 639 – 650. doi:https://doi.org/10.1016/j.physa.2017.10.006.

Gusella, R. (1991). Characterizing the variability of arrival processes with indexes of dispersion. *IEEE Journal on Selected Areas in Communications*, 9(2), 203–211.

Hurst, H.E. (1956). Methods of using long-term storage in reservoirs. *Proceedings of the Institution of Civil Engineers*, 5(5), 519–543.

Jlassi, J. (2009). *Amélioration de la performance par la modélisation des flux logistiques des patients dans un service d'urgence hospitalier*. Theses, Université Paris VIII Vincennes-Saint Denis.

Kadri, F., Chaabane, S., Harrou, F., and Tahon, C. (2014). Modélisation et prévision des flux quotidiens des patients aux urgences hospitalières en utilisant l'analyse de séries chronologiques. In *7ème conférence de Gestion et Ingénierie des Systèmes Hospitaliers (GISEH)*, 1–8. Liège, Belgium.

Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H.E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1), 87–114.

Kim, Y.G., Shiravi, A., and Min, P.S. (2006). Congestion prediction of self-similar network through parameter estimation. In *2006 IEEE/IFIP Network Operations and Management Symposium NOMS 2006*, 1–4. doi:10.1109/NOMS.2006.1687669.

Koizumi, N., Kuno, E., and Smith, T.E. (2005). Modeling patient flows using a queuing network with blocking. *Health care management science*, 8(1), 49–60.

Lashermes, B., Roux, S.G., Abry, P., and Jaffard, S. (2008). Comprehensive multifractal analysis of turbulent velocity using the wavelet leaders. *The European Physical Journal B-Condensed Matter and Complex Systems*, 61(2), 201–215.

Lazarou, G.Y., Baca, J., Frost, V.S., and Evans, J.B. (2009). Describing network traffic using the index of variability. *IEEE/ACM Transactions on Networking (TON)*, 17(5), 1672–1683.

Lin, D., Patrick, J., and Labeau, F. (2014). Estimating the waiting time of multi-priority emergency patients with downstream blocking. *Health care management science*, 17(1), 88–99.

Mandelbrot, B. (1965). Une classe de processus stochastiques homothetiques a soi-application a la loi climatologique de hurst. *Comptes rendus de l'Academie des Sciences*, 260(12), 3274–+.

Monte, E., Roca, J., and Vilardell, L. (2002). On the self-similar distribution of the emergency ward arrivals time series. *Fractals*, 10(04), 413–427.

Park, K. and Willinger, W. (2000). *Self-similar network traffic and performance evaluation.* Wiley Online Library.

Peng, C.K., Havlin, S., Stanley, H.E., and Goldberger, A.L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5(1), 82–87.

Przytycki, F. and Urbański, M. (1989). On the hausdorff dimension of some fractal sets. *Studia Mathematica*, 93(2), 155–186.

Radev, D., Lokshina, I., and Radeva, S. (2010). Modeling and simulation of self-similar traffic in wireless ip networks. In *Networking and Telecommunications: Concepts, Methodologies, Tools, and Applications*, 1631–1647. IGI Global.

Samorodnitsky, G. (2006). Long memory and self-similar processes. In *Annales-Faculte des Sciences de Toulouse Mathematiques*, volume 15, 107. Université Paul Sabatier.

Strzałka, B., Mazurek, M., and Strzałka, D. (2012). Queue performance in presence of long-range dependencies–an empirical study. *International Journal of Information Science*, 2(4), 47–53.

Thompson, J.R. and Wilson, J.R. (2016). Multifractal detrended fluctuation analysis: Practical applications to financial time series. *Mathematics and Computers in Simulation*, 126, 63–88.

Vemuri, V.R. (2005). *Enhancing Computer Security with Smart Technology.* CRC Press.

Xiong, G., Zhang, S., and Liu, Q. (2012). The time-singularity multifractal spectrum distribution. *Physica A: Statistical Mechanics and its Applications*, 391(20), 4727–4739.