New Journal of Physics
The open-access journal for physics

The soundscape dynamics of human agglomeration

Haroldo V Ribeiro[1,2,3], **Rodolfo T de Souza**[1], **Ervin K Lenzi**[1,2], **Renio S Mendes**[1,2] **and Luiz R Evangelista**[1]

[1] Departamento de Física, Universidade Estadual de Maringá, Av. Colombo 5790, 87020-900 Maringá, PR, Brazil
[2] National Institute of Science and Technology for Complex Systems, CNPq, Rua Xavier Sigaud 150, 22290-180 Rio de Janeiro, RJ, Brazil
E-mail: hvr@dfi.uem.br

New Journal of Physics **13** (2011) 023028 (8pp)
Received 5 November 2010
Published 11 February 2011
Online at http://www.njp.org/
doi:10.1088/1367-2630/13/2/023028

Abstract. We report on a statistical analysis of the people agglomeration soundscape. Specifically, we investigate the normalized sound amplitudes and intensities that emerge from human collective meetings. Our findings support the existence of non-trivial dynamics characterized by heavy tail distributions in the sound amplitudes, long-range correlations in the sound intensity and non-exponential distributions in the return interval distributions. Additionally, motivated by the time-dependent behavior present in the volatility/variance series, we compare the observational data with those obtained from a minimalist autoregressive stochastic model, namely the generalized autoregressive conditional heteroskedastic process (the GARCH process), and find that there is good agreement.

Contents

[3] Author to whom any correspondence should be addressed.

1. Introduction

Physicists are now addressing problems very far from their traditional domain. Even social phenomena are now ubiquitous in the research carried out by statistical physicists [1]. In particular, the general framework of Statistical Physics has been successfully applied to diverse interdisciplinary fields, ranging from finance [2], genetics [3] and biology [4], to religion [5], tournaments [6], culinary [7] and music [8].

Naturally, in social phenomena, the basic constituent of the system is humans. Humans are known to have non-trivial collective dynamics, much more complicated than idealized physical interacting systems. Moreover, even individual aspects related to social agents may not be available. This complex scenario is reflected, in some sense, in several human activities. Elections [9], collaborations between actors [10] and also between scientists [11], phone text-message [12], mail [13, 14] or e-mail [14, 15] communication, human travel [16, 17] and collective listening habits [18, 19] are just a few examples where complex structures have been found.

Most of the previous investigations deal with record data obtained directly or indirectly from the system, trying to extract some patterns or regularities about the system dynamics. This approach has been a trend towards investigating social phenomena and also complex systems in general [20]–[24]. Within this framework, the most diversified data were used as the sound. 'Listen to' the system dynamics may be both a simple task and a minimally invasive measurement. In this direction, several studies focusing on the sound time series have been performed. To mention just a few, research works on the acoustic emission from crumpled paper [25, 26], from paper fracture [27] and from fractures in general [28, 29] show several features related to critical phenomena, the power spectrum of music and speech sounds presents $1/f$-like spectra [30] and the normalized sound amplitude shows non-Gaussian features [31], traffic flows have been investigated using the sound noise revealing scaling and memory [32], and avalanche-like dynamics were found in the sound of popping bubbles in foams [33] and also in lung sounds [34, 35].

In this work, we present an investigation of a very common situation related to human collective activities: people agglomeration. Human being agglomerations can emerge at various places for different reasons: for example, people having lunch in a restaurant, parties and workplace meetings. In all of these situations, a common and notorious feature is perceptible: the resulting sound noise from these agglomerations. In this paper, our main goal is to show that non-trivial dynamics emerge on analyzing this kind of time series. In addition, employing a minimalist model, we are able to reproduce the statistical aspects of the empirical data. In the following, we present the details of the data acquisition, the statistical analysis of the data and our minimalist model, and finally we end with a summary.

2. Data presentation

The observational data were obtained by recording the soundscape of people agglomeration during recreation time at our university. The meeting point is an open place where the students spend time until the next class. All of the measurements were carried out by using a condenser microphone (Shure Microflex MX202W/N) positioned in the central part of the agglomeration. We employed a sampling rate of 44.1 kHz in order to cover the full human audible range (between approximately 20 Hz and 20 kHz). Additionally, the measurements were carried out

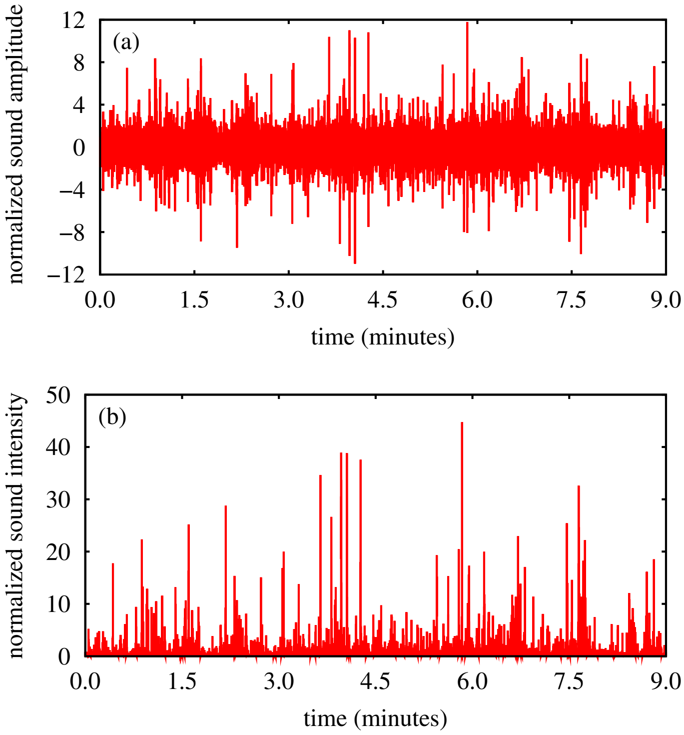

Figure 1. A representative record sound signal: (a) the normalized sound amplitude and (b) the normalized sound intensity.

during different periods in 9 days totaling 16 records. The number of people during the recordings ranged between approximately 100 and 200, and these variations did not significantly change the statistical results. The typical recording time was about 10 min, and during the recording the number of people was approximately constant. We also analyzed ten recordings from a web sound database[4], and found similar results as compared with our measurements.

Figure 1(a) shows a representative record signal where we employed the normalized sound amplitude A_t, i.e. the sound amplitude subtracted by its mean value and divided by its standard deviation. Figure 1(b) presents the sound intensity, A_t^2, divided by its standard deviation. From these two figures, we can observe the existence of some bursts where the sound amplitude and the sound intensity exceed by values much larger than their standard deviations. Qualitatively, the origin of these extreme events may be, for instance, related to the fact that people want to be heard, and if their neighbors are talking loudly they have to increase their sound intensity.

3. Statistical analysis

One of the most direct ways of characterizing the sound amplitude is by evaluating its probability density function (pdf). We show this analysis in figure 2(a) for three typical recordings, where we also plot one Gaussian distribution with zero mean and unitary variance (dashed line). Quite similar behavior has been found for all other realizations of the experiment

[4] freesound.org.

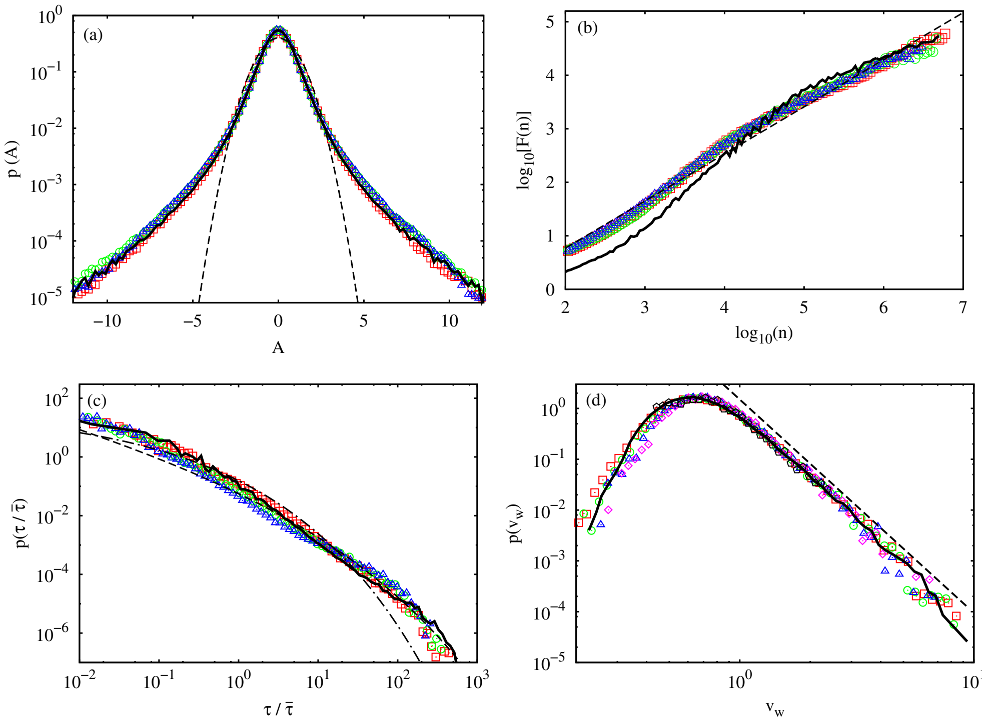

Figure 2. (a) Probability distribution of the normalized sound amplitude A for three realizations of the experiment (squares, circles and triangles) confronted with the standard Gaussian pdf (dashed line) and with the GARCH model (continuous line). (b) Detrended fluctuation analysis (DFA) when considering the same three previous realizations for the normalized sound intensity: $\log_{10}[F(n)]$ versus $\log_{10}(n)$ in comparison with a linear fit (dashed line), where we found $F(n) \propto n^h$ with $h \approx 0.88$, and with the GARCH model (continuous line). Here, n is in units of $1/44.1k$ seconds. (c) Return interval distributions take into account one realization of the experiment for three threshold values: $q = 1$ (squares), $q = 2$ (circles) and $q = 5$ (triangles) compared with the stretched exponential (dashed-doted line) and the Weilbull distribution (dashed line) of equation (1) with $\gamma = 2(1-h) = 0.24$, and also with the GARCH model (continuous line). (d) Volatility distribution for one realization of the experiment and considering five window sizes: $w = 1$ (squares), $w = 2$ (circles), $w = 5$ (triangles), $w = 10$ (diamonds) and $w = 100$ (pentagons) hundredths of a second. The dashed line is a power law with $p(v) \propto v^{-4.1}$ and the continuous line is the GARCH prediction.

and also for the web recordings (at least in the central part of the distribution). The empirical distribution clearly differs from the Gaussian one, especially for larger values of the sound amplitude ($|A|$ greater than four standard deviations). Naturally, this heavy tail behavior reflects the presence of extreme events that we qualitatively see in figure 1.

One possible way to investigate the dynamics of these extreme events is by evaluating the time interval between them. These time intervals can be obtained by considering a threshold value q and storing all of the initial times t_i for which the normalized sound intensity is above this edge. The difference between two consecutive times $\tau_i = t_{i+1} - t_i$ is the so-called return interval. For Gaussian uncorrelated (or weak correlated) random variables, the distribution of τ_i is well known to follow an exponential distribution $p(\tau) \sim e^{-\tau/\bar{\tau}_q}$, where $\bar{\tau}_q$ is the average value of τ_i when considering the threshold value q. Additionally, empirical results have shown that, in the presence of power law correlations in the data, the distribution is well adjusted by a stretched exponential [36]–[38] or by a Weibull distribution [39], i.e.

$$p(\tau) \sim e^{-A(\tau/\bar{\tau}_q)^\gamma} \quad \text{or} \quad p(\tau) \sim (\tau/\bar{\tau}_q)^{\gamma-1} e^{-B(\tau/\bar{\tau}_q)^\gamma}, \tag{1}$$

where A and B are constants and γ is the exponent of the power law autocorrelation function. These distributions also emerge in the analytical framework of Santhanam and Kantz [40] when considering a long-range correlated noise with Gaussian pdf. Note that all of these distributions are dependent on q, but if we employ the scaled variable $\tau_i/\bar{\tau}$, this dependence is eliminated.

Before we investigate the return intervals, let us address the correlation question by using the DFA [41]. This technique basically considers the root mean square fluctuation function $F(n)$ (see for instance [42]) for the integrated and detrended time series for different values of the time scale n. When the data present scale-invariance properties, $F(n)$ follows a power law $F(n) \sim n^h$, where h indicates the degree of correlation in the time series: if $h = 0.5$, the series is uncorrelated, and if $h > 0.5$, the series is long-range correlated. Figure 2(b) shows the fluctuation function versus n for the same three recordings of figure 2(a) where we found $h \approx 0.88$, indicating that long-range correlations are present in the data. Note that the three curves are practically identical. This fact is evidenced by evaluating the mean value of h ($\bar{h}$) and its standard deviation (σ_h) over the 16 recordings and finding that $\bar{h} = 0.88$ and $\sigma_h = 0.001$. When considering the web recordings, these values remain close: $\bar{h} = 0.89$ and $\sigma_h = 0.01$.

Now, advancing with the return interval distribution, it is interesting to emphasize that the exponents h and γ are related via $\gamma = 2(1-h)$. Moreover, since the distribution of $\tau_i/\bar{\tau}$ should be normalized and also have unitary mean, the only fit parameter is γ, which can be obtained from h, leading to $\gamma \approx 0.24$. Figure 2(c) shows this distribution for three values of q where we can observe a reasonable data collapse but not so good agreement with the distributions of equation (1). A similar situation has been observed recently when considering the non-Gaussian distributions related to water boiling [43].

We can also investigate the bursts observed in figure 1(a) by evaluating the volatility of the normalized sound amplitude. This time series refers to the local standard deviation of $A(t)$ estimated over a time window $w = n\Delta t$, i.e.

$$v_w^2(t) = \frac{1}{n-1} \sum_{t'=t}^{t+n-1} (A(t') - \langle A(t) \rangle_w)^2, \tag{2}$$

where $\langle A(t) \rangle_w = (1/n) \sum_{t'=t}^{t+n-1} A(t')$, n is an integer and Δt is the sampling time interval. Figure 2(d) shows the volatility distribution of our empirical data for time windows ranging from 1/100 to 1 second. Note that we found a good collapse of data and that this distribution has an asymptotic power law decay characterized by an exponent $\eta = 4.1$. The mean value and the standard deviation of η calculated over the 16 realizations are, respectively, $\bar{\eta} = 4.29$ and $\sigma_\eta = 0.35$ ($\bar{\eta} = 4.90$ and $\sigma_\eta = 1.10$ for the web recordings).

4. Modeling

Our starting point to model the data behavior is the non-stationary aspect of the volatility. Figure 2(d) supports the conclusion that the volatility of the sound amplitude is a time-dependent stochastic process and figure 2(b) indicates that long-term memory is present in sound intensity series. This feature is very common in financial data where the volatility (or risk) is one of the most essential ingredients in the price dynamics. In this scenario, much work has been performed [44] and consequently a large number of models are available. From a qualitative point of view, the interactions (competitions) among people present in financial markets seem to be similar to those existing in our social system. This picture motivates us to employ a typical financial model to our data.

One of these models is the generalized autoregressive conditional heteroskedastic process (the GARCH process). This model was proposed [45] (at least in part) to take into account the long memory typically found in financial data. It is defined in its most general form, GARCH(p, q), by

$$\begin{aligned} x_t &= \sigma_t \, \xi_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_q \sigma_{t-q}^2, \end{aligned} \tag{3}$$

where α_i and β_i are positive control parameters and ξ_t is an uncorrelated random variable with zero mean and unitary variance. Thus, the GARCH process is uncorrelated in x_t but correlated in the variance. Also note that for $\alpha_i = 0$, the GARCH recovers the so-called ARCH process [46].

Here, for simplicity and also for satisfactoriness, we will focus on the GARCH(1, 1) process,

$$\begin{aligned} x_t &= \sigma_t \, \xi_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned} \tag{4}$$

for which we choose the distribution of ξ_t to follow the standard Gaussian. After this simplification, the model has three parameters, α_0, α_1 and β_1. However, since the sound amplitude is scaled to a unitary variance, we can eliminate one of these parameters by using the expected variance of the GARCH(1, 1) process x_t,

$$\sigma_x^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}. \tag{5}$$

In this manner, we now have two parameters that we incrementally update to minimize, via the method of least squares, the difference between the simulated values of sound amplitude and the observational ones. The best values for the parameters are $\alpha_1 = 0.011$ and $\beta_1 = 0.9889$, leading to $\alpha_0 = 0.0001$ since $\sigma_x = 1$. The comparison with the empirical data is shown in figure 2, where the GARCH(1, 1) predictions are indicated by continuous lines. We can see that there is very good agreement between the data and the GARCH(1, 1).

Concerning figure 2(b), where we compare the DFA analysis, we have to remark that the autocorrelation function of the variable x_t^2 is not really long-range correlated. In fact, it has an exponential decay [45], i.e. $\langle x_t^2 \, x_{t+\tau}^2 \rangle \sim \exp(-t/\tau_c)$, where $\tau_c = |\ln(\alpha_1 + \beta_1)|^{-1}$. However, the GARCH(1, 1) process can mimic the long-range decay for large values of the characteristic time τ_c. This feature can be achieved by choosing the sum $\alpha_1 + \beta_1$ closer to unity. In our case, $\alpha_1 + \beta_1 = 0.9999$, leading to characteristic time $\tau_c \sim 10^4$ s, which is very large, mimicking at least in part the long-range correlations. Note that the empirical data also present deviations from the straight line, suggesting that correlations present in the data may have a kind of exponential cutoff.

5. Summary

In this work, we have investigated some statistical aspects of the collective sound emitted by people when they have agglomerated in a meeting place. Empirical evidence showed that (i) the normalized sound amplitude is not Gaussian distributed, (ii) the sound intensity presents long-range correlations, (iii) the return interval distribution of the sound intensity is not exponential and (iv) the volatility of the sound amplitude is non-stationary, having a power law tail in its distribution. Motivated by the time dependence of the volatility, we compared the observational quantities with the predictions of the GARCH(1, 1) model, and found that there is good agreement among them.

Before concluding, we would like to point out some possible mechanisms responsible for the presence of heavy tail distributions and long-term correlations in the data. The first one is related to the fact that humans already have an intrinsic complex behavior, which may manifest itself in our measurements. Secondly, these individuals form small interacting groups, adding more complexity to the system. On a third level, there is also emergence of interactions between groups. Naturally, more detailed measurements and models should be considered, in comparison with those presented here, to obtain a broad understanding of this system.

Acknowledgments

We thank CNPq/CAPES (Brazilian agencies) for financial support and CENAPAD-SP for computational support. LRE and RTS also thank CNPq/INCT-FCX for financial support.

References

[1] Castellano C, Fortunato S and Severo V 2009 *Rev. Mod. Phys.* **81** 591
[2] Vandewalle N, Ausloos M, Boveroux P and Minguet A 1998 *Eur. Phys. J.* B **4** 139
[3] Peng C K, Buldyrev S V, Goldberger A L, Havlin A, Sciortino F, Simons M and Stanley H E 1992 *Nature* **356** 168
[4] Berg H C 1993 *Random Walks in Biology* (Princeton, NJ: Princeton University Press)
[5] Picoli Jr S and Mendes R S 2008 *Phys. Rev.* E **77** 036105
[6] Ribeiro H V, Mendes R S, Malacarne L C, Picoli Jr S and Santoro P A 2010 *Eur. Phys. J.* B **75** 327
[7] Kinouchi O, Diez-Garcia R W, Holanda A J, Zambianchi P and Roque A C 2008 *New J. Phys.* **10** 073020
[8] Correa D C, Saito J H and Costa L F 2010 *New J. Phys.* **12** 053030
[9] Fortunato S and Castellano C 2007 *Phys. Rev. Lett.* **99** 138701
[10] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[11] Newman M E J 2001 *Phys. Rev.* E **64** 016131
[12] Zhou T, Kiet H A T, Kim B J, Wang B H and Holme P 2008 *Europhys. Lett.* **82** 28002
[13] Oliveira J G and Barabási A L 2005 *Nature* **437** 1251
[14] Vazquez A 2007 *Physica A* **373** 747
[15] Barabási A L 2005 *Nature* **435** 207
[16] Brockmann D, Hufnagel L and Geisel T 2006 *Nature* **439** 462
[17] González M C, Hidalgo C A and Barabási A L 2008 *Nature* **453** 779
[18] Lambiotte R and Ausloos M 2005 *Phys. Rev.* E **72** 066107
[19] Lambiotte R and Ausloos M 2006 *Eur. Phys. J.* B **50** 183
[20] Auyang S Y 1998 *Foundations of Complex-Systems* (Cambridge: Cambridge University Press)
[21] Jensen H J 1998 *Self-Organized Criticality* (Cambridge: Cambridge University Press)

[22] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
[23] Boccara N 2004 *Modeling Complex Systems* (New York: Springer)
[24] Sornette D 2006 *Critical Phenomena in Natural Sciences* (Berlin: Springer)
[25] Kramer E M and Lobkovsky A E 1996 *Phys. Rev.* E **53** 1465
[26] Mendes R S, Malacarne L C, Santos R P B, Ribeiro H V and Picoli Jr S 2010 *Europhys. Lett.* **92** 29001
[27] Salminen L I, Tolvanen A I and Alava M J 2002 *Phys. Rev. Lett.* **89** 185503
[28] Sethna J P, Dahmen K A and Myers C R 2001 *Nature* **410** 242
[29] Minozzi M, Caldarelli G, Pietronero L and Zapperi S 2003 *Eur. Phys. J.* B **36** 203
[30] Voss R F and Clarke J 1975 *Nature* **258** 317
[31] Mendes R S, Ribeiro H V, Freire F C M, Tateishi A A and Lenzi E K 2011 *Phys. Rev.* E **83** 017101
[32] Skagerstam B S K and Hansen A 2005 *Europhys. Lett.* **72** 513
[33] Vandewalle N, Lentz J F, Dorbolo S and Brisbois F 2001 *Phys. Rev. Lett.* **86** 179
[34] Alencar A M, Hantos Z, Peták F, Tolnai J, Asztalos T, Zapperi S, Andrade J S, Buldyrev S V, Stanley H E and Suki B 1999 *Phys. Rev.* E **60** 4659
[35] Alencar A M, Buldyrev S V, Majumdar A, Stanley H E and Suki B 2001 *Phys. Rev. Lett.* **87** 088101
[36] Bunde A, Eichner J F, Kantelhardt J W and Havlin S 2005 *Phys. Rev. Lett.* **94** 048701
[37] Yamasaki K, Muchnik L, Havlin S, Bunde A and Stanley H E 2005 *Proc. Natl Acad. Sci. USA* **102** 9424
[38] Wang F, Yamasaki K, Havlin A and Stanley H E 2006 *Phys. Rev.* E **73** 026117
[39] Blender R, Fraedrich K and Sienz F 2008 *Nonlinear Process. Geophys.* **15** 557
[40] Santhanam M S and Kantz H 2008 *Phys. Rev.* E **78** 051113
[41] Peng C K, Buldyrev S V, Havlin S, Simons M, Stanley H E and Goldberger A L 1994 *Phys. Rev.* E **49** 1685
[42] Kantelhardt J W, Koscielny-Bunde E, Rego H H A, Havlin S and Bunde A 2001 *Physica* A **295** 441
[43] Ribeiro H V, Mendes R S, Lenzi E K, Belancon M P and Malacarne L C 2011 On the bubbles clouds dynamics in boiling water *Chaos Solitons Fractals* accepted arXiv:1101.4408v1
[44] Mantegna R N and Stanley H E 1999 *An Introduction to Econophysics* (Cambridge: Cambridge University Press)
[45] Bollerslev T 1986 *Econometrics* **31** 307
[46] Engle R F 1982 *Econometrica* **50** 987