Physica A 583 (2021) 126327

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Is Brazilian music getting more predictable? A statistical physics approach for different music genres

Paulo Ferreira [a,b,c,*], Derick Quintino [d], Bruna Wundervald [e], Andreia Dionísio [c], Faheem Aslam [f], Ana Cantarinha [g,h,i]

[a] VALORIZA - Research Center for Endogenous Resource Valorization, Portalegre, Portugal
[b] Instituto Politécnico de Portalegre, Portalegre, Portugal
[c] CEFAGE-UE, IIFA, Universidade de Évora, Largo dos Colegiais 2, 7000 Évora, Portugal
[d] Department of Economics, Administration and Sociology, University of São Paulo, 13418-900 Piracicaba, SP, Brazil
[e] Hamilton Institute & Department of Mathematics and Statistics, Maynooth University, Maynooth, Kildare, Ireland
[f] Management Sciences, Comsats University, Islamabad 45550, Pakistan
[g] Instituto Politécnico de Beja, Portugal
[h] CIMA-Center for Research on Mathematics and its Applications, University of Évora, Portugal
[i] CMA-Center of Mathematics and Applications, FCTNOVA, NOVA University of Lisbon, Portugal

ARTICLE INFO

Article history:
Received 12 February 2021
Received in revised form 9 July 2021
Available online 9 August 2021

Keywords:
Detrended Fluctuation Analysis
Music genres
Predictability
Shuffle

ABSTRACT

Music is an important part of most people's lives and also of the culture of a country. Moreover, the different characteristics of songs, such as genre and the chord sequences, could have different impacts on individual behaviours. Even considering just seven chords and the respective variations, originality can be a crucial element of a song's success. Considering this, and in the context of Brazilian music, we employed the Detrended Fluctuation Analysis to analyse the possible predictability of eight different music genres. On these genres, we found that Reggae and Pop seem to be the least random considering the sequenced use of chords. With a sliding windows approach, we found that the predictability of chord sequences of Pop decreased over time. Applying the same methodology after shuffling the original series of music, the results point to a randomness of those shuffled series, demonstrating the robustness of our approach.

1. Introduction

Music is an important element in building a nation's cultural identity, with a strong impact on the population, not only due to its power to delight or entertain, but also the influence it can have on human behaviour. Recent research has sought to understand how music and its characteristics, such as genre and chord sequences, can affect consumer demand, by understanding its effects on behaviour in restaurants, in increasing or decreasing a specific demand for goods or services [1,2], for example. Consequently, music has become one of the major economic sectors today and has shown strong growth in various regions of the globe.

According to de Melo et al. [3], a certain regional pattern is observed in preferences for Brazilian musical genres. There is a preponderance of Forró in Recife and Fortaleza (both Northeast), Pop and Techno in Florianópolis and Curitiba (South), Samba and Pagode in Rio de Janeiro (Southeast), and Backcountry in Goiânia and Cuiabá (Midwest) and Rio Branco and Boa Vista (North). More specifically, Wundervald [4] studied the features that characterize Brazilian genres and found that

* Corresponding author.
E-mail addresses: pferreira@ipportalegre.pt (P. Ferreira), derickdq@usp.br (D. Quintino), bruna.wundervald@mu.ie (B. Wundervald).

https://doi.org/10.1016/j.physa.2021.126327
0378-4371/

harmonic differences are of great importance here, which might also correlate to particular cultural elements of genre preferences.

Regardless of the musical genre, according to Serra et al. [5], the key to the music's acceptability lies in a balance between predictability and surprise. Thus, songs tend to have both predictable and unpredictable structures, seeking to balance this dual character.

One way to study music predictability is through analysis of long-range dependence. Here, $1/f$ noise is also referred to as long range dependence, as it is an intermediate stage between white noise and random walk. This is a particular case of a more general form of $1/f^{\beta}$, because when $\beta = 1$, $1/f$ noise is also defined as pink noise. In music, the spectrum $1/f$ ($1/f^{\beta}$ with $\beta = 1$) was interpreted as a trade-off between predictability and surprise: if β tends to zero (white noise), the temporal sequence of notes is highly unrelated (so, very unpredictable) and tends to be unpleasant for the average consumer; on the other hand, if β is too high, the music tends to become monotonous (highly predictable due to greater long-range correlation). According to González-Espinoza et al. [6], values have been found in empirical studies and we can relate the Hurst exponent α with the power spectrum $P(f) \sim f^{\beta}$, through the Wiener–Khinchin theorem [7], where $\beta = 2\alpha - 1$.

Throughout history, mathematicians and philosophers, in addition to composers, have investigated the relationship between mathematical laws and musical composition, in order to discover their underlying mathematical structures [8]. Despite intuitive knowledge, there is little quantitative evidence on the hypothesis of dosing predictability and unpredictability. Therefore, music has a wide range of properties that make it an interesting object of study for different areas of knowledge.

The study of music through statistical physics techniques has been of great interest over the past years [6]. Among others, there is the work of Lima et al. [9], who studied musical opus through entropy, pseudo phase plane and multidimensional scaling, finding relevant differences between musical genres.

Machado et al. [10] analysed musical opus from entropy and multidimensional scaling, focusing on three musical genres (Classical, Jazz and Pop-Rock), and found significant differences in these styles. Levitin et al. [11], verified how classical music rhythms are predictable and how they vary according to genre, and found that Beethoven's rhythms were among the most predictable, whereas Mozart's were less so.

Through the cumulative distribution function of 5 compositions, Liu et al. [12] studied classical composers, namely, Bach, Mozart, Beethoven, Mendelssohn and Chopin, and found different power-law distribution behaviour for each composer, which implies different levels of autocorrelation of musical notes. Xin et al. [13] used complex networks to analyse the compositions of Mozart, Beethoven and Chopin based on note sequences and length, and found relevant differences among composers' network edges.

Lopes and Tenreiro Machado [14] used various mathematical techniques, such as complexity, dimensionality-reduction, clustering and visualization techniques, to analyse the music of four contemporary artists (Frank Sinatra, Rolling Stones, Johnny Hallyday and Julio Iglesias) and the results show that evolution of complexity is correlated with the artists' musical career. Moss et al. [15] used Markov tables to evaluate the musical notes transition in Beethoven string quartets and, among the results found, observed that the distribution of notes obeys a power law, i.e., few chords are responsible for a reasonable proportion of the compositions. Other studies used Detrended Fluctuation Analysis (DFA), such as Wu et al. [16], González-Espinoza et al. [6] and Colley and Dean [17], as well as further developments such as the multifractal DFA (MF-DFA) by Jafari et al. [18] and Telesca and Lovallo [19].

The DFA has the particularity of identifying the existence of power-laws. Wu et al. [16] found that consonance patterns in music followed a free-scale characteristic, therefore compatible with that observed by power-law. González-Espinoza et al. [6] also found that different musical genres showed different patterns of autocorrelation according to the composer, and Colley and Dean [17] showed that DFA is useful to distinguish patterns according to the musical time series.

As well as contributing to the empirical evidence regarding the statistical properties of music time series, the musical databases currently available are valuable for quantitative studies, mainly because they present a vast amount of serial and cross-sectional data [20].

With this in mind, the article intends to analyse a wide range of Brazilian songs, divided in genres and sequenced according to their release date, aiming to assess the randomness of the use of the different chords in their harmonic structures. In particular, we intend to identify which genres could be considered most predictable, and how such structures have evolved over time for the different genres.

The remainder of the article is organized as follows: Section 2 presents the data and methodology used in this study; Section 3 presents the results and Section 4 concludes.

2. Data and methodology

Our main objective is to assess whether Brazilian music, i.e., composed and/or sung, by Brazilian musicians, is becoming more predictable or if some musical genres are more random than others in terms of the construction of the music structure. For this purpose, we use a sample of a user-inputted (or crowd-sourced) music chords dataset, which was first presented in the chords package [21] for the R statistical software [22]. In total, 8 Brazilian music genres are available in the dataset: Reggae, Pop, Forro, Bossa Nova, Sertanejo, MPB, Rock and Samba, with the chords data from 8339 songs by 106 different artists. For each musical piece, the full chord sequences and song keys are provided, as well as a few

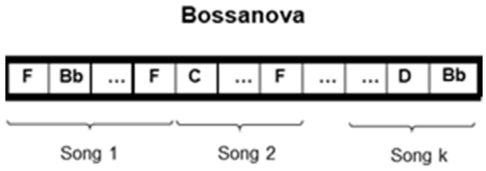

Fig. 1. Example of the construction of the chord series.

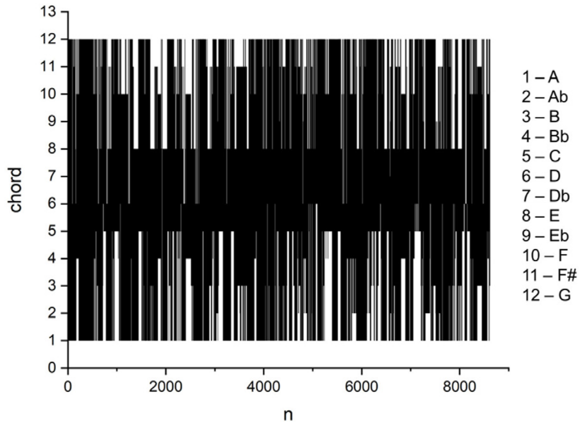

Fig. 2. Example of the time series of the sequence of chords for the Forro genre. Letters A to G refer to the chords from La to Sol; b refers to bemol and # to sharp chords.

auxiliary features such as the moment of the song's release and its popularity, obtained with the aid of the Spotify API. For more details, Wundervald and Zeviani [23] have produced an extensive analysis of the full dataset and its properties for music genre prediction tasks in the Brazilian context, in particular the capacity to correctly classify different songs in different music genres, which is helpful for our analysis. Moreover, the use of a wide set of songs from different authors, enlarging the data sample, is also crucial for the application of statistical physics methodologies.

Considering the richness of the set of available songs, we sequenced them according to the chronological release date, as depicted in Fig. 1, in the case of Bossanova. As mentioned, based on the release date, we pick up the chord sequence of the first song, followed by the chords of the second song and so on, ending with the kth song. The chords were then numbered, in order to be considered as useable with the DFA, a similar rationale being presented by Lopes and Tenreiro Machado [14]. It should be noted that the time series obtained for each musical genre has some artificiality, but all of them respect the date of the songs' release. Moreover, during the study, we took measures to identify the robustness of the analysis. Fig. 2 shows the example of the built time series for the Forro genre. The x-axis represents the sequence of the chords as explained before and the y-axis represents the respective chord.

Fig. 3 gives a graphical representation of the percentage distribution of the chords for the different genres, considering the chords in their simplified version, with no modifications or added notes. Note that we cannot spot significant differences in the percentage use of chords across the different musical genres. In fact, a simpler analysis points to a possible relative uniform distribution of chords in all music genres under analysis.

Considering the created time series for the chords, we used the Detrended Fluctuation Analysis (DFA) to analyse the behaviour of the chord sequence. Proposed by Peng et al. [24], DFA is a methodology able to identify the dependence pattern of time series, even in the case of non-stationarity. Originally used to analyse DNA behaviour, DFA was later used in many other research areas such as finance (see, for example, [25–27], among many others), biology [28] or [29], climatology [30] or [31], criminology [32,33], car traffic [34,35] or even to explain sport results [36], showing its ability to explain and describe different phenomena. Moreover, in this last case, the constructed time series is also non-continuous, as it was used for the sequence of results in that sport, for the teams under analysis.

Considering a given time series x_i of length N, the first step of the DFA consists of calculating the profile $X_t = \sum_{i=1}^{t}(x_i - \langle x \rangle)$, with $\langle x \rangle$ being the average of the original time series. This profile is divided in boxes of length n and for each box the local trend $\tilde{X}_t$ is calculated with ordinary least squares, which is used to detrend the profile X_t and to obtain the fluctuation function given by $F(n) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(X_t - \tilde{X}_t\right)^2}$. The process is repeated for all the boxes of length n and a log–log regression is applied between $F(n)$ and the respective n, resulting in a power-law given by $F(n) \propto n^{\alpha}$, with the α as the exponent of the DFA. This α value could be used to measure the randomness of a given time series. In fact, if $\alpha = 0.5$, the time series is described by a random walk and it has an unpredictable pattern. In this context of using

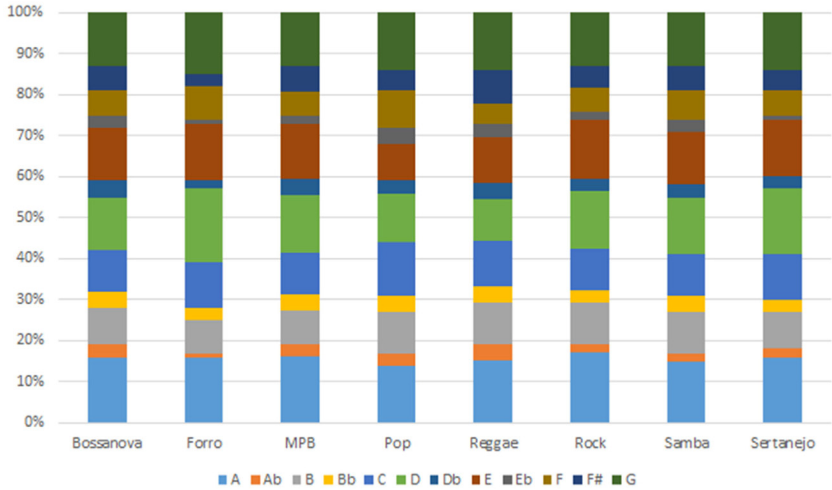

Fig. 3. Distribution of the chords for the different musical genres, considering the following total number of chords by genre: Bossanova (35206), Forró (8621), MPB (103039), Pop (7228), Reggae (2852), Rock (103148), Samba (71029), Sertanejo (152868). Letters A to G refer to the chords from La to Sol; b refers to bemol and # to sharp chords.

the DFA, this situation means that the repetition of chords in a song is not likely. If we have $\alpha > 0.5$, on the other hand, the time series is persistent, i.e., it has a positive long-range dependence and a given pattern is likely to be repeated in the future. In the context of our work, this now means that some predictability could be found in the chord sequence of the songs. Finally, in the case of $\alpha < 0.5$, we have anti-persistent dependence, meaning some possibility of predictability (although with a different pattern).

To analyse the evolution of the dependence pattern of the chords of the different musical genres, we calculated not only the general DFA but also the DFA based on a sliding windows approach. Considering sequential windows of 1000 observations, we calculated evolution of the DFA over time, which could give us some information about the evolution of the randomness of the different music genres.

Application of the sliding windows approach in the DFA is common in different research areas (see, for example, the papers by Cajueiro and Tabak [37,38,39] or Filho et al. [33]). In order to identify the index with the most random behaviour over time, we use an adaptation of the Efficiency Index proposed by Kristoufek and Vosvrda [40]. In his work, Kristoufek and Vosvrda [40] used that index to analyse the randomness of financial assets. Here, we adapt that index, in order to identify the predictability of music genres. We also adapt the index, because the variables have different numbers of observations. So, our Predictable Index (PI) is given by $PI = \sqrt{\frac{\sum_{i=1}^{N}(M_i-0.5)^2}{N}}$, with M_i the estimation for each DFA window, with the numerator assessing the distance to the random level (0.5). The adaptation consists of dividing the squared sum by the number of calculated exponents, since the number of observations is not the same for all musical genres. The PI will allow us to compare directly the different music genres, as higher PI levels imply higher predictability of the chord's series. In this research, PI is a complementary measure of DFA, since it gives the level of predictability in terms of evolution of the DFA exponents over time.

To ensure robustness, after calculating the DFA exponents for the created time series of the chord sequence, according to the chronological release of the songs, we made similar calculations considering a shuffling process for all-time series. For the shuffling procedure, we performed 10.000 random permutations of each series, with different random number generator seeds in each step. The final shuffled time series has the same number of chords as the original time series, but organized in a different form, expected to be random. Then, the DFA was calculated for the shuffled time series.

3. Results

We started our statistical analysis by applying DFA to the generated series considering different polynomial orders, as reported in Fig. 4, and following, for example, the procedures of Telesca and Lovallo [41] and Telesca et al. [42], using the MFDFA R package. There, we can see that in some musical genres the differences are negligible (see, for example, MPB, Rock or Sertanejo), although in others the linear fit is different. Comparing the fluctuation curves from polynomials of order 3, 4 and 5 it is possible to conclude on their similarity, suggesting that possible non-stationarities affect the data up to the 3rd order. For this reason, we decided to apply the DFA considering polynomials of order 3, with the results presented in Table 1. The second column identifies the Hurst exponent and the respective standard deviation while the third column shows the order, considering a descendent degree of persistence.

Fig. 5 is the graphical representation of the fluctuation functions from the DFA analysis, representing the log–log relationship used to estimate the DFA exponents, with the dashed lines showing the linear fit. As seen by the R-squares,

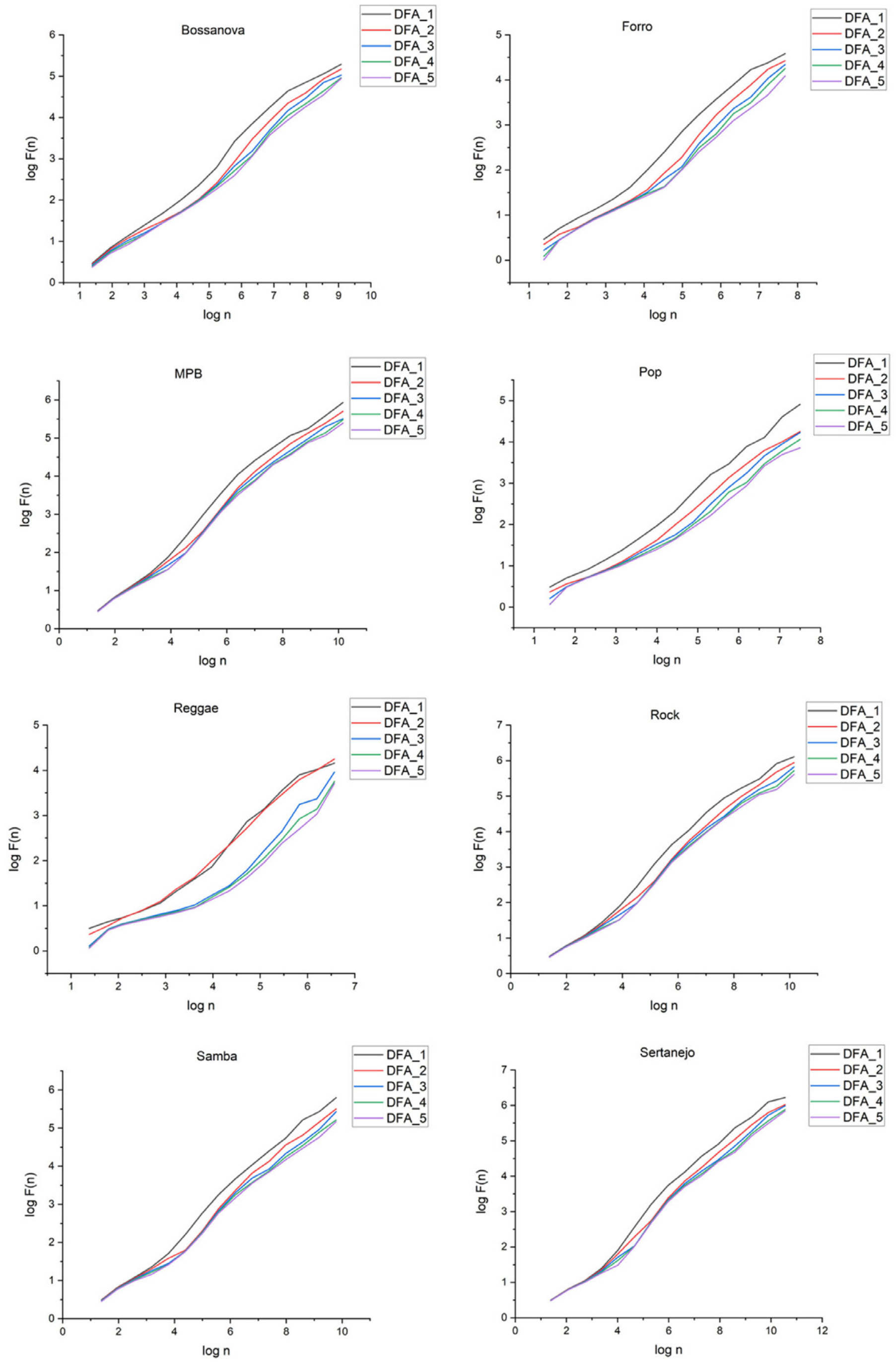

Fig. 4. DFA applied to the whole time series using polynomials from order 1 (DFA-1) to order 5 (DFA-5).

the fit quality is always very high, meaning that the DFA exponent estimation is very good. The musical genre with the lowest R-square is Reggae, which could be related to the reduced sample of that genre.

Regarding the results, we can see that the musical genre with highest dependence is Reggae, followed by Pop music. Samba, MPB and Bossanova have the lowest DFA exponents, meaning they could be considered less likely to be predictable. This agrees with results obtained by Wundervald and Zeviani [23], where it was observed that Samba, MPB and Bossanova have a much greater variety of chords used. According to these authors, over the years the greater use of different chords

Table 1
DFA results for the dependence of chords of the different music genres under study, considering a polynomial of order 3.

	Hurst exponent ± std. dev.	Order
Bossanova	0.625 ± 0.019	# 6
Forro	0.660 ± 0.025	# 3
MPB	0.615 ± 0.016	# 7
Pop	0.667 ± 0.025	# 2
Reggae	0.692 ± 0.053	# 1
Rock	0.649 ± 0.016	# 4
Samba	0.608 ± 0.016	# 8
Sertanejo	0.639 ± 0.015	# 5

Table 2
Predictability index for the evolution of the dependence of the different music genres under study.

	PI	Order
Bossanova	0.0629	# 7
Forro	0.0637	# 6
MPB	0.0808	# 3
Pop	0.0948	# 1
Reggae	0.0675	# 4
Rock	0.0901	# 2
Samba	0.0536	# 8
Sertanejo	0.0651	# 5

Table 3
DFA results for shuffled series.

	Hurst exponent ± std. dev.
Bossanova	0.512 ± 0.006
Forro	0.520 ± 0.006
MPB	0.507 ± 0.005
Pop	0.493 ± 0.011
Reggae	0.531 ± 0.005
Rock	0.497 ± 0.004
Samba	0.498 ± 0.004
Sertanejo	0.504 ± 0.006

in these genres is found, while others seem to have more uniform harmonies. Similarly, in a genre prediction situation, these 3 genres were most likely to be confused with each other, highlighting their harmonic similarities.

Considering the possibility that the fluctuation curves could have non-uniform scaling, which seems to happen, for example, in some musical genres from Fig. 5, meaning there could be different scaling regions, we followed the procedures presented by Telesca and Lovallo [41,41]. Considering the first and last 3000 observations of each series (except for Reggae, where we just considered the first 3000 due to the dimension of that time series), we obtained the scaling exponent $\chi(n) \equiv d[\ln F(n)]/[\ln(n)]$, i.e., the local derivative of the log–log plot of the fluctuation function. If the $\chi(n)$ is not stable, the scaling behaviour deviates from the uniform power-law (see [43]). For this purpose, we calculate the scaling instability index ρ as the difference between maximum and minimum values of the local scaling exponents $\chi(n)$, meaning that higher ρ imply greater deviations from a uniform power-law. As proposed by Telesca and Lovallo [41], we define an upper threshold, corresponding to 5% of the ρ values, and any value above the threshold could be viewed as anomalous. Also in all cases, we considered the DFA calculations based on a polynomial of 3rd order.

Figs. 6 and 7 depict the information for the different musical genres, with the horizontal line representing the defined threshold. This reveals that anomalous values are scarce and there is no regular pattern, as expected considering the way the time series are obtained, even in the case of Reggae, where the fluctuation function seems to have some non-linear patterns. These relatively low values of the instability index are a good indicator for the method which was applied.

As we intended to analyse the evolution of those patterns over time, we developed a sliding windows approach, with the results presented in Fig. 8, showing the different DFA exponents over the different samples of consecutive windows of 1000 chords (represented in the x-axis). There, we can see interesting features. For example, Bossanova seems to have a behaviour centred on the 0.5 level, although with some peaks. This means that over time, this musical genre presents some randomness in the evolution of chords. Looking at the other musical genres, generally they are constantly above the 0.5 level, which implies that the use of chords is persistent, i.e., seemingly more predictable. We can also see that pop music clearly reduced the value of the Hurst exponent over time, meaning it has become a richer musical genre, in that its harmonic complexity is increasing over time.

Based on the PI, presented in Table 2, we can reinforce the conclusions obtained in the previous analysis. In fact, Bossanova and Samba seem to be the most innovative and unpredictable musical genres, i.e., with more random patterns,

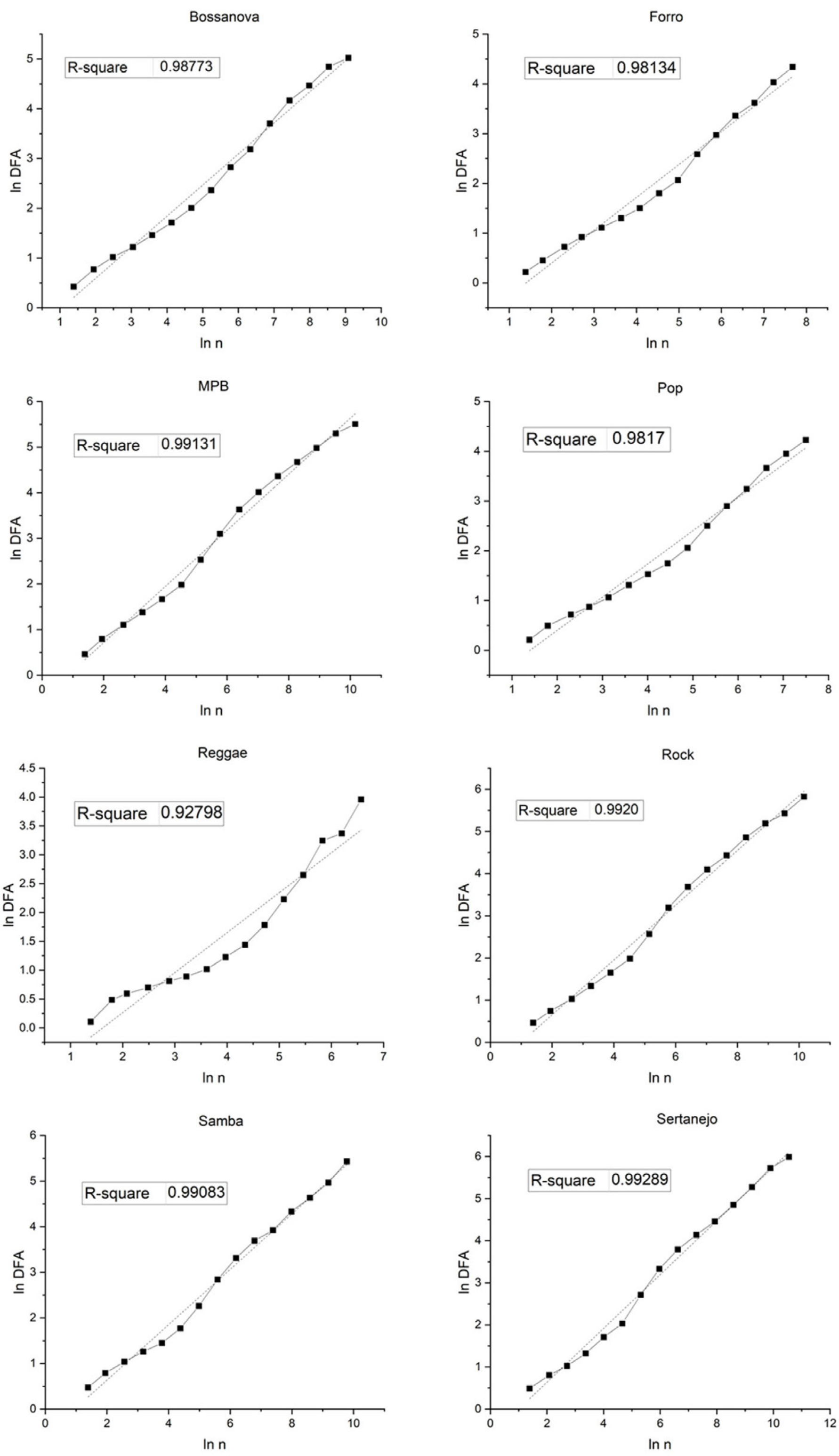

Fig. 5. Fluctuation function of the DFA for each musical genre, considering the polynomial of order 3.

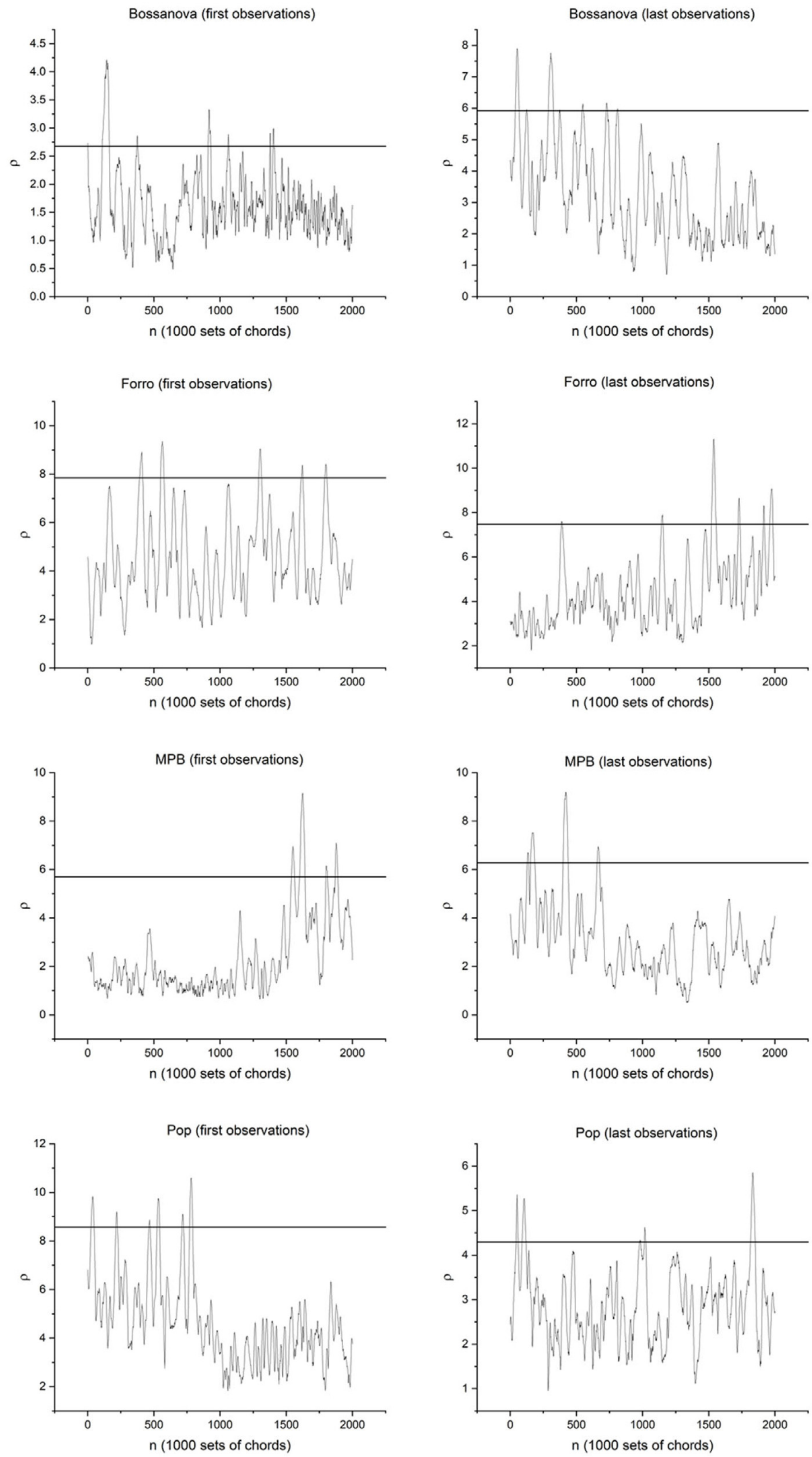

Fig. 6. Time variation of the scaling instability index ρ for Bossanova, Forro, MPB and Pop.

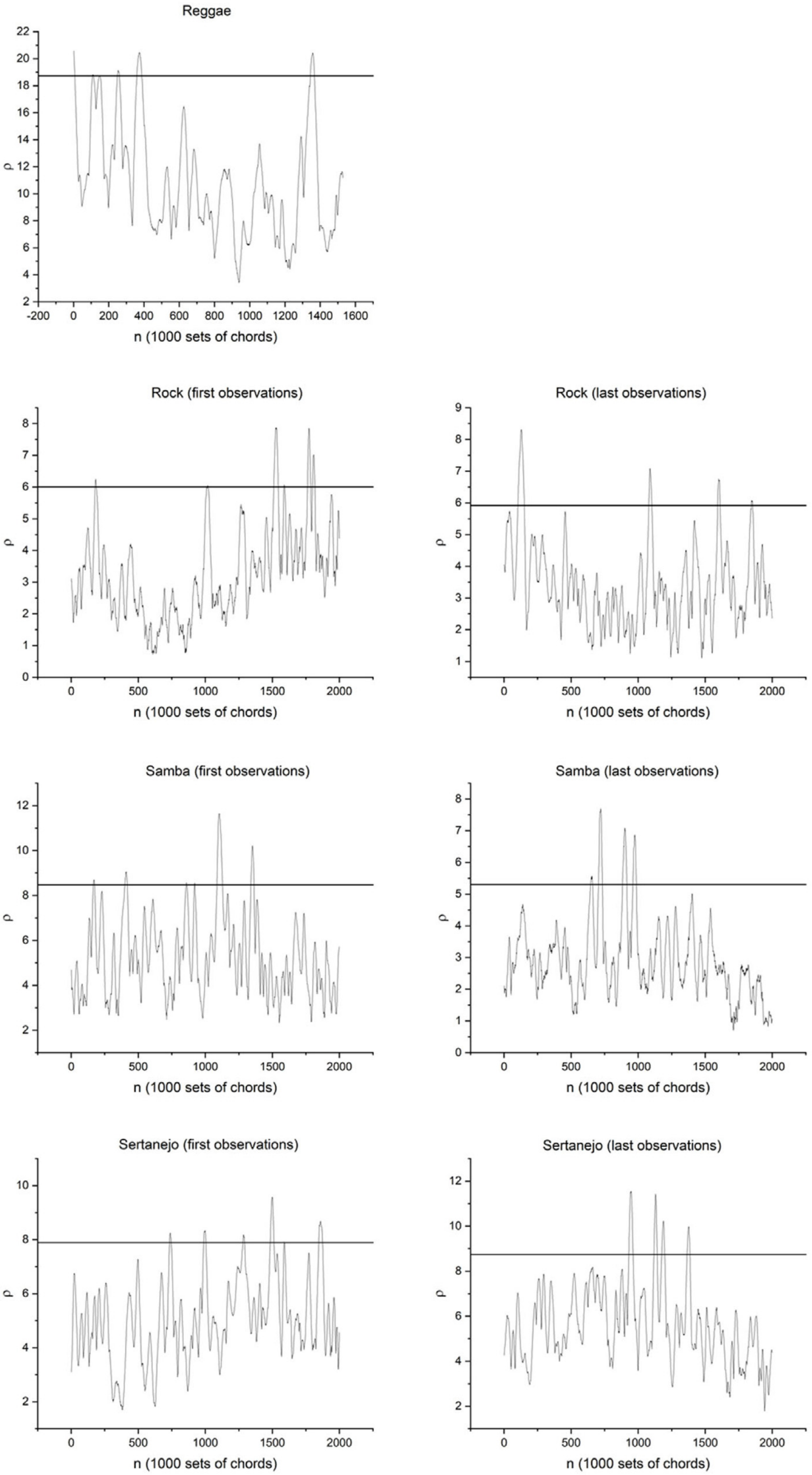

Fig. 7. Time variation of the scaling instability index ρ for Reggae, Rock, Samba and Sertanejo.

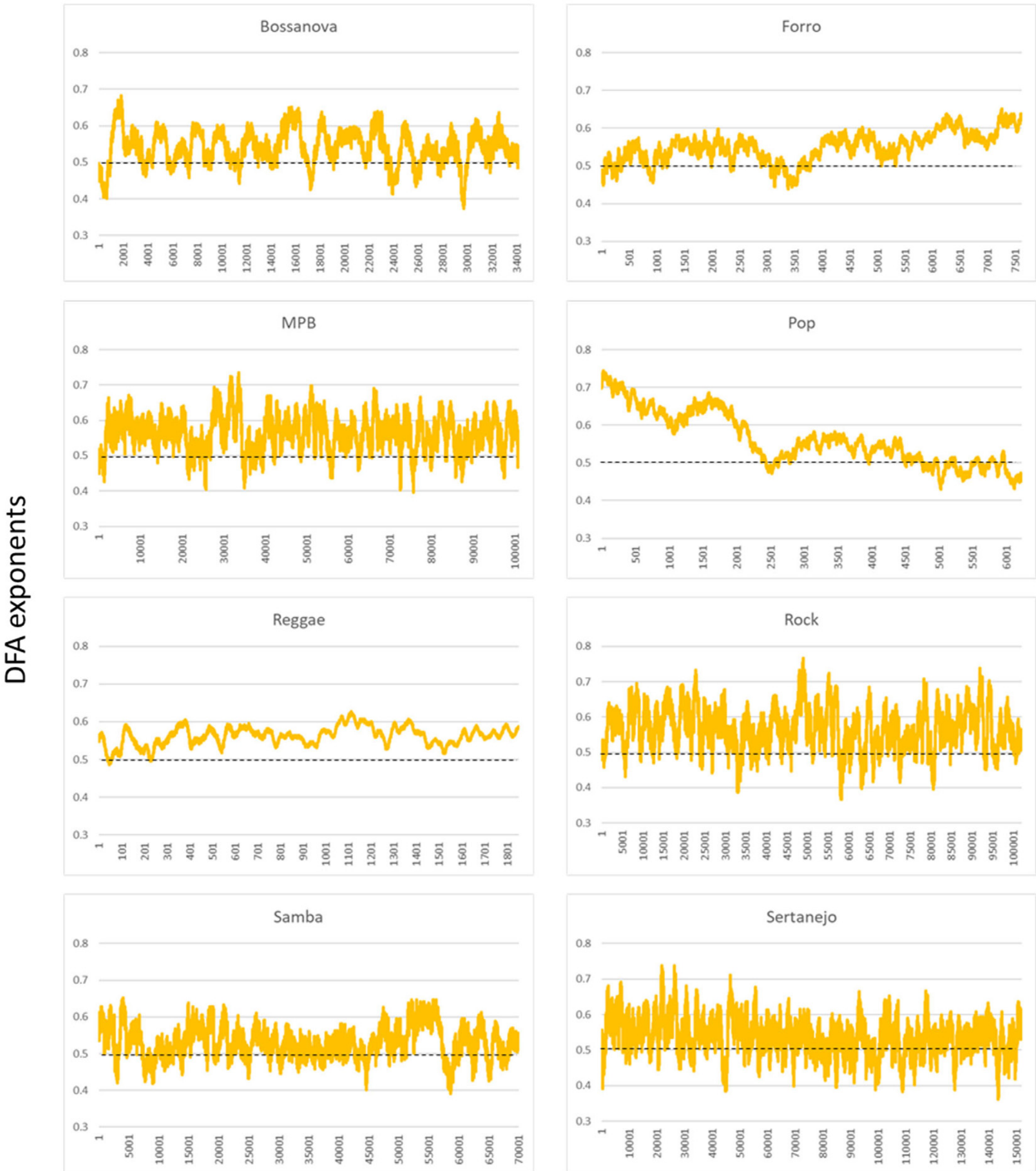

Fig. 8. Evolution of the DFA exponents, based on windows of 1000 observations (chords). The *x*-axis represents the order of the DFA exponent estimated through sliding windows (1 is the first DFA exponent, from the window from n = 1, . . . , 1000; 2 the second, from the window from n = 2, . . . , 1001; and so on).

being the least predictable genres. Reggae is the most predictable music genre as a whole (#1 in Table 1) but is in the middle of the table when considering the predictability over time (#4 in Table 2). Considering the predictability index, Pop and Rock seem to be the most predictable musical genres.

Nevertheless, if we analyse the evolution of the PI itself, based on sliding windows of the same 1000 observations, i.e., each PI accounting for samples of 1000 DFA exponents with sliding windows, we obtain interesting behaviours, shown in detail in Fig. 9. In particular, we have evidence that Pop music is in fact becoming more unpredictable (i.e., harmonically more complex), while the other musical genres do not show these patterns.

As previously mentioned, we shuffled the time series, for robustness purposes. After 10.000 shuffle movements, the final time series is expected to have random behaviour. So, the DFA exponent of those shuffled time series should be near to a random walk. The results of the DFA applied to the shuffled series are presented in Table 3, showing that all the estimated Hurst exponents are near the 0.5 level. The musical genre with the highest Hurst exponent after shuffling is Reggae, which is in this case related to the lower number of observations for this genre, but even so the exponent is

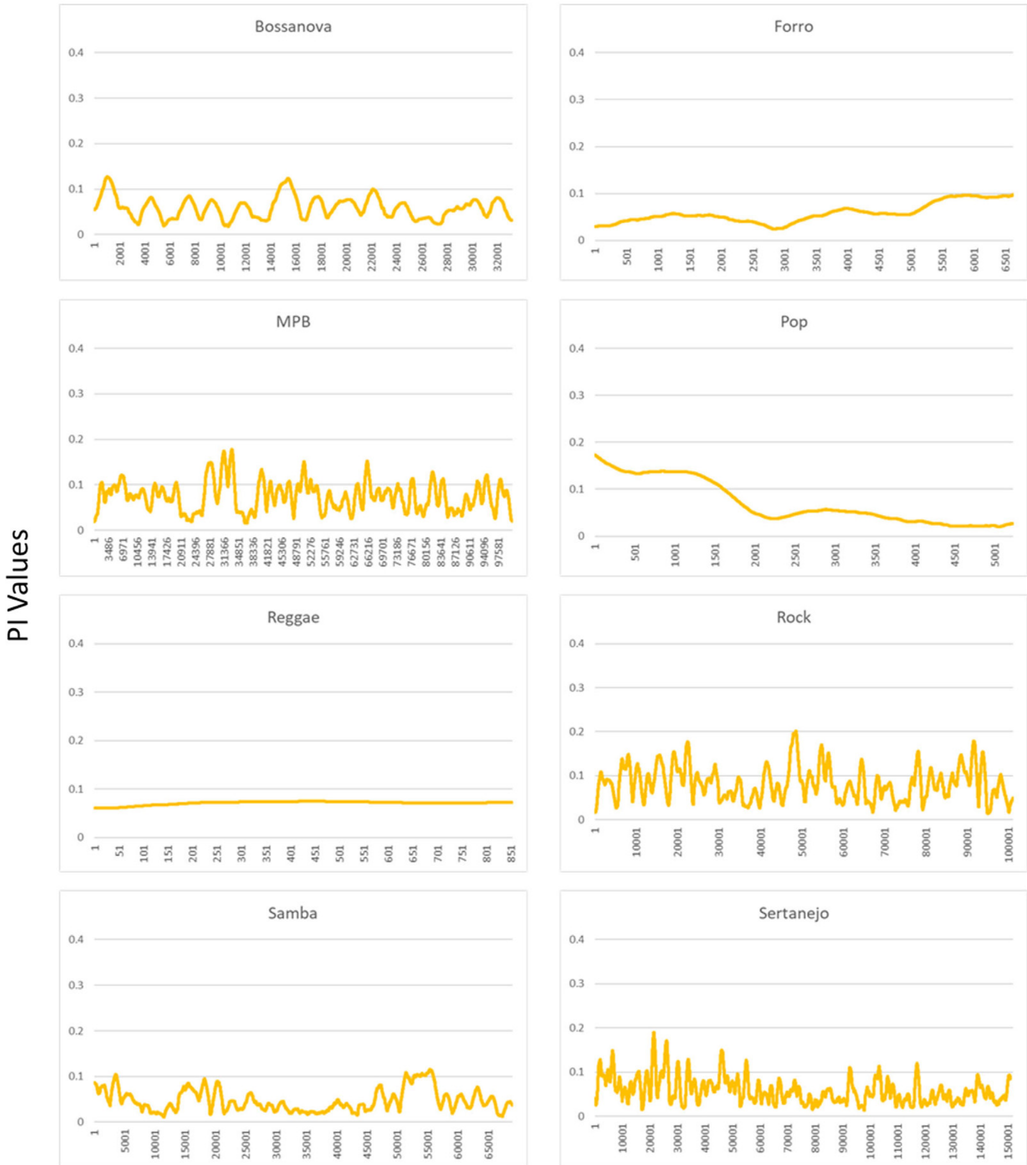

Fig. 9. Evolution of the Predictable index over time.

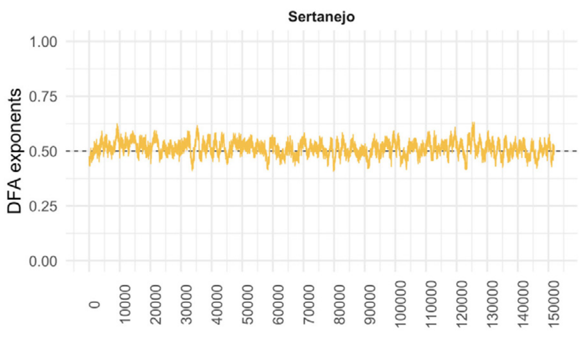

Fig. 10. Evolution of the DFA exponents for the shuffled series of Sertanejo, based on windows of 1000 observations (chords).

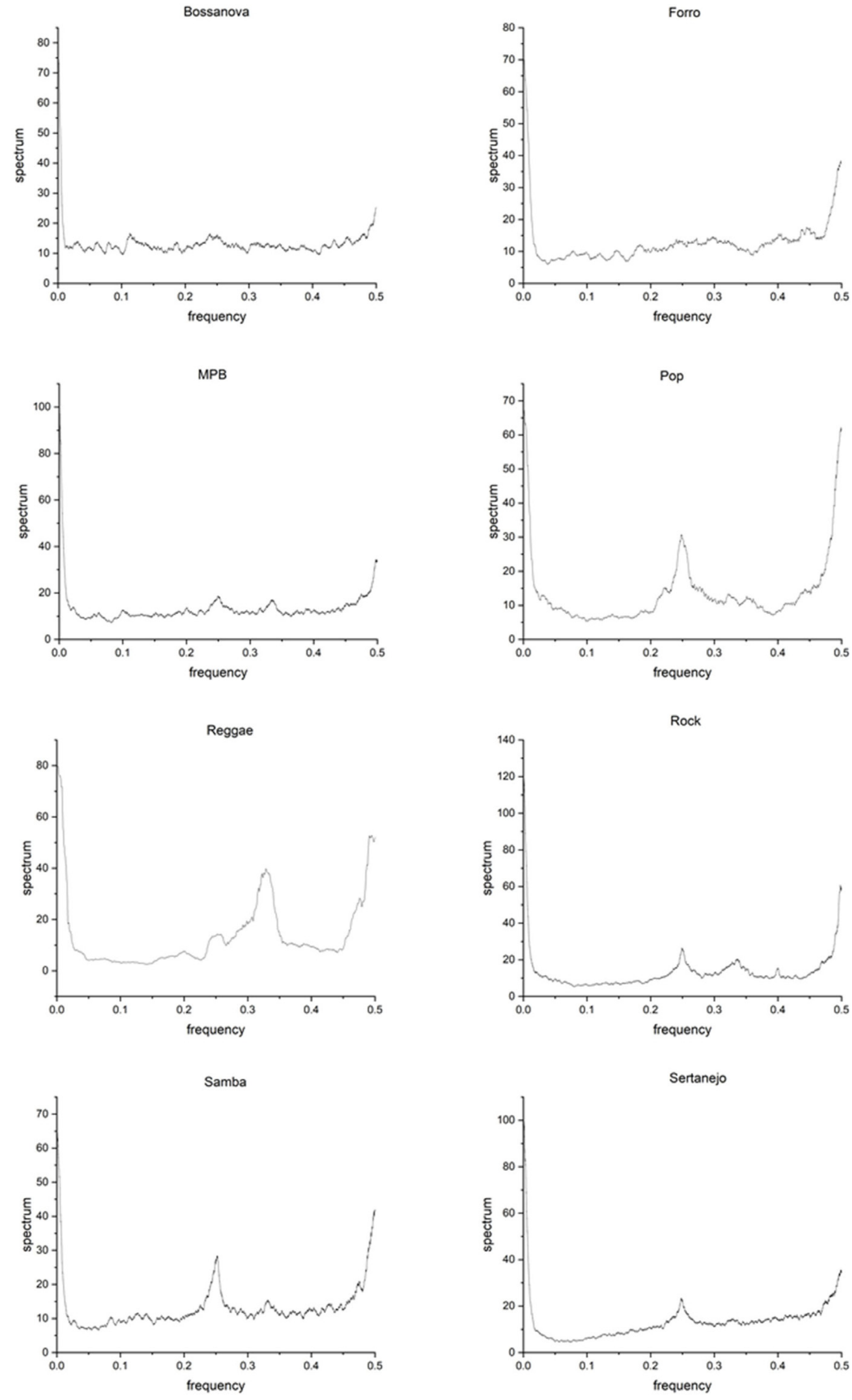

Fig. 11. Spectral analysis of the chords for the different music genres.

not too different from what is expected for a random level. The use of sliding windows confirms this robustness, with Fig. 10 showing the behaviour of Sertanejo with a variation over the 0.5 level and with a PI index of 0.0341, about half that observed in the original series.

To complement our analysis, we performed a spectral analysis, with the results presented in Fig. 11. In a random series, the spectrum analysis should have a constant pattern. None of the genres shows that pattern, although Bossanova and Forro have less fluctuation in their spectral analysis. In general, we can see some peaks in the figures, which are more pronounced in Reggae and Pop, consistent with the results of the DFA for the whole samples, complementing our findings. On the other hand, we have the results of Bossanova and MPB which seem to be the genres with a more constant pattern in their spectral analysis, reflecting a possible wide use of different chords which, according to Wundervald and Zeviani [23], could lead to difficulty in distinguishing between the genres, when predicting which genre a song belongs to.

4. Conclusions

The main goal of this research is to analyse a wide range of Brazilian songs from different genres (both traditional Brazilian and international music genres), sequenced according to their release date moment, so as to assess the randomness of the use of the different chords in their harmonic structures. More specifically, we intend to identify which genres could be considered most predictable, and how such structures evolved over time for the different genres.

Our results point globally to the conclusion that Reggae is the most predictable genre, considering a static approach, while from a dynamic point of view, Pop and Rock seem to have some more predictability than the other musical genres. On the other hand, Bossanova and Samba seem to be more innovative, more unpredictable and consequently richer in terms of structure.

Direct comparison between these musical genres could be difficult, especially because some are globally known and "used" and some others are more "local", such as Bossanova or Sertanejo. However, our results seem to indicate the existence of genres in which innovation and unpredictability are more significant, and this innovation capacity might be an important determinant for the respective demand.

It is also possible that innovation, related to the lack of predictability, could be connected to the commercial demand for some types of songs. For example, some information about the most popular music genres in Brazil appears in the Statista website.[1] Note that a possible direct relationship cannot be made, as in that list we have other music genres including both Brazilian and international music (while our analysis is just of Brazilian music). However, some of the most popular musical genres are also those showing the highest levels of novelty and innovation (see, for example, Samba). On the other hand, Reggae is not listed as one of the preferred genres, which could be related to the evidence that this genre is becoming more predictable. Therefore, future research could be directed to understanding the relationship between general predictability and musical preferences at a more global level.

To some extent, our results corroborate previous findings. As in jazz, Bossanova makes extensive use of chord changes, although with less harmonic alteration, since in Bossanova there is a maximum of two chords per measure, whereas in jazz they can reach four without difficulty [44]. Furthermore, Abezer et al. [45] point out that the syncopate, a technique that results in breaking the listener's expectations, is frequently used in Latin American musical genres such as salsa and Bossanova.

Regarding Reggae in general, and not only the Brazilian form, it is relatively simple in terms of melodic complexity, so that it can be played or sung without major difficulties, consisting, in general, of two or three chords in terms of harmonic complexity [46].

It is important to highlight that the data used in this research is the result of a time series concatenation of music from the same genre. The music is organized in chronological terms by its release date, although the formation of the database could imply some deviations from the actual situation. Therefore, we must also mention that the shuffling results corroborate the results obtained with the original time series, which gives robustness to the analysis. Besides this, the key of each song could be a relevant factor in explaining chord sequences. Considering that it was not possible to incorporate this particular information in our models, we consider this an important future research topic.

Finally, and to answer the main question of this research, of whether Brazilian music is becoming more predictable, our results show that some music genres, namely more traditional Brazilian ones, are becoming more innovative, and probably more complex in terms of structure.

CRediT authorship contribution statement

Paulo Ferreira: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Derick Quintino:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Bruna Wundervald:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Andreia Dionísio:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Faheem Aslam:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Ana Cantarinha:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing.

[1] See the preferences of 2018 in https://www.statista.com/statistics/966242/most-popular-music-genres-brazil/.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Paulo Ferreira acknowledges the financial support of Fundação para a Ciência e a Tecnologia (grants UIDB/05064/2020 and UIDB/04007/2020). Andreia Dionísio is pleased to acknowledge financial support from Fundação para a Ciência e a Tecnologia (grant UIDB/04007/2020). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – FinanceCode 001. Derick Quintino wishes to acknowledge the CAPES for funding support. Bruna Wundervald acknowledges the financial support of Science Foundation Ireland through the Career Development Award grant number: 17/CDA/4695.

References

[1] A. Michel, C. Baumann, L. Gayer, Thank you for the music–or not? The effects of in-store music in service settings, J. Retail. Consum. Serv. 36 (2017) 21–32.

[2] C.H. Cho, A. Mattila, P. Bordi, E. Kwon, It tastes better when Bach meets red: the role of music and plate color on food evaluation, Br. Food J. 122 (1) (2019) 114–125.

[3] G.B.V. de Melo, A.F. Machado, L.R. de Carvalho, Music consumption in Brazil: an analysis of streaming reproductions, PragMATIZES-Revista Latino-Americana de Estudos Em Cultura 10 (19) (2020) 141–169.

[4] B. Wundervald, Feature engineering for genre characterization in Brazilian Music, in: 13th International Workshop on Machine Learning and Music, 2020.

[5] J. Serra, H. Kantz, X. Serra, R.G. Andrzejak, Predictability of music descriptor time series and its application to cover song detection, IEEE Trans. Audio Speech Lang. Process. 20 (2) (2011) 514–525.

[6] A. González-Espinoza, H. Larralde, G. Martínez-Mekler, M. Müller, Multiple scaling behaviour and nonlinear traits in music scores, R. Soc. Open Sci. 4 (12) (2017) 171282.

[7] N. Wiener, Generalized harmonic analysis, Actamathematica 55 (1930) 117–258.

[8] J.T. Machado, And I say to myself: "What a fractional world!", Fract. Calc. Appl. Anal. 14 (4) (2011) 635–654.

[9] M.F. Lima, J.A. Machado, A.C. Costa, A multidimensional scaling analysis of musical sounds based on pseudo phase plane, Abstr. Appl. Anal. 2012 (2012) 436108.

[10] J.T. Machado, A.C. Costa, M.F. Lima, Dynamical analysis of compositions, Nonlinear Dynam. 65 (4) (2011) 399–412.

[11] D.J. Levitin, P. Chordia, V. Menon, Musical rhythm spectra from Bach to Joplin obey a 1/f power law, Proc. Natl. Acad. Sci. 109 (10) (2012) 3716–3720.

[12] L. Liu, J. Wei, H. Zhang, J. Xin, J. Huang, A statistical physics view of pitch fluctuations in the classical music from Bach to Chopin: Evidence for scaling, PLoS One 8 (3) (2013) e58710.

[13] C. Xin, H. Zhang, J. Huang, Complex network approach to classifying classical piano compositions, Europhys. Lett. 116 (1) (2016) 18008.

[14] A.M. Lopes, J.A. Tenreiro Machado, On the complexity analysis and visualization of musical information, Entropy 21 (7) (2019) 669.

[15] F.C. Moss, M. Neuwirth, D. Harasim, M. Rohrmeier, Statistical characteristics of tonal harmony: A corpus study of Beethoven's string quartets, PLoS One 14 (6) (2019) e0217242.

[16] D. Wu, K.M. Kendrick, D.J. Levitin, C. Li, D. Yao, Bach is the father of harmony: revealed by a 1/f fluctuation analysis across musical genres, PLoS One 10 (11) (2015) e0142431.

[17] I.D. Colley, R.T. Dean, Origins of 1/f noise in human music performance from short-range autocorrelations related to rhythmic structures, PlosOne 14 (5) (2019) e0216088.

[18] G.R. Jafari, P. Pedram, L. Hedayatifar, Long-range correlation and multifractality in Bach's inventions pitches, J. Stat. Mech. Theory Exp. (2007) P04012.

[19] L. Telesca, M. Lovallo, Revealing competitive behaviours in music by means of the multifractal detrended fluctuation analysis: application to Bach's Symphonies, Proc. R. Soc. A 467 (2134) (2011) 3022–3032.

[20] H.V. Ribeiro, L. Zunino, R.S. Mendes, E.K. Lenzi, Complexity–entropy causality plane: A useful approach for distinguishing songs, Physica A 391 (7) (2012) 2421–2428.

[21] B. Wundervald, The chords package for extraction of music chords data in R, 2018, https://github.com/r-music/chorrrds.

[22] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2020, URL: https://www.R-project.org/.

[23] B.D. Wundervald, W.M. Zeviani, Machine learning and chord based feature engineering for genre prediction in popular Brazilian music, 2019, arXiv preprint arXiv:1902.03283.

[24] C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E Stat. 49 (1994) 1685.

[25] G. Cao, M. Zhang, Extreme values in the Chinese and American stock markets based on detrended fluctuation analysis, Physica A 436 (15) (2015) 25–35.

[26] P. Anagnostidis, C. Varsakelis, C. Emmanouilides, Has the 2008 financial crisis affected stock market efficiency? The case of eurozone, Physica A 447 (2016) 116–128.

[27] D. Quintino, J. Campoli, H. Burnquist, P. Ferreira, Efficiency of the Brazilian bitcoin: A DFA approach, Int. J. Financial Stud. 8 (2) (2020) 25.

[28] K. Wan, R. Goldstein, Rhythmicity, recurrence, and recovery of flagellar beating, Phys. Rev. Lett. 113 (2014) 238103.

[29] Y.A. Djawad, D. Attwood, J. Kiely, R. Luxton, The application of detrended fluctuation analysis to assess physical characteristics of the human cell line ECV304 following toxic challenges, Sens. Bio-Sensing Res. 23 (2019) 100269.

[30] K. Fraedrich, R. Blender, Scaling of atmosphere and ocean temperature correlations in observations and climate models, Phys. Rev. Lett. 90 (2003) 108501.

[31] S. Adarsh, D. Kumar, B. Deepthi, G. Gayathri, S. Aswathy, S. Bhagyasree, Multifractal characterization of meteorological drought in India using detrended fluctuation analysis, Int. J. Climatol. 39 (11) (2019) 4234–4255.

[32] A. Filho, M. Silva, G. Zebende, Autocorrelation and cross-correlation in time series of homicide and attempted homicide, Physica A 400 (2014) 12–19.

[33] A. Filho, G. Zebende, E. Guedes, Analysis of intentional lethal violent crimes: A sliding windows approach, Physica A 576 (2021) 125653.

[34] G. Zebende, A. Filho, Cross-correlation between time series of vehicles and passengers, Physica A 388 (23) (2009) 4863–4866.

[35] G. Zebende, P. Silva, A. Filho, Study of cross-correlation in a self-affine time series of taxi accidents, Physica A 2011 (2011) 1677–1683.

[36] P. Ferreira, What detrended fluctuation analysis can tell us about NBA results, Physica A 500 (2018) 92–96.

[37] D. Cajueiro, B. Tabak, The Hurst exponent over time: Testing the assertion that emerging markets are becoming more efficient, Physica A 336 (3–4) (2004) 521–537.

[38] D. Cajueiro, B. Tabak, Evidence of long range dependence in Asian equity markets: The role of liquidity and market restrictions, Physica A 342 (3–4) (2004) 656–664.

[39] D. Cajueiro, B. Tabak, Testing for predictability in equity returns for European transition markets, Econ. Syst. 30 (1) (2006) 56–78.

[40] L. Kristoufek, M. Vosvrda, Measuring capital market efficiency: Global and local correlations structure, Physica A 392 (2013) 184–193.

[41] L. Telesca, M. Lovallo, Non-uniform scaling features in central Italy seismicity: A non-linear approach in investigating seismic patterns and detection of possible earthquake precursors, Geophys. Res. Lett. 36 (2009) L01308.

[42] L. Telesca, M. Lovallo, A. Ramirez-Rojas, F. Angulo-Brown, Scaling instability in self-potential earthquake-related signals, Physica A 388 (2009) 1181–1186.

[43] G. Viswanathan, C. Peng, H. Stanley, A. Goldberger, Deviations from uniform power law scaling in nonstationary time series, Phys. Rev. E 55 (1997) 845.

[44] S.A. Reily, Tom Jobim and the bossa nova era, Pop. Music 15 (1) (1996) 1–16.

[45] J. Abezer, H. Lukashevich, P. Bräuer, Classification of music genres based on repetitive basslines, J. New Music Res. 41 (3) (2012) 239–257.

[46] M. Bitz, Teaching Improvisation outside of Jazz Settings: Musical genres that lend themselves to improvisations by beginning student musicians include bluegrass, blues, ska, reggae, rap, klezmer, and rock, Music Educ. J. 84 (4) (1998) 21–41.