Efficient Calculation of Fractal Properties via the Higuchi Method

J. A. Wanliss[1] and Grace E. Wanliss[1]

[1]Department of Physics, Presbyterian College, 503 S. Broad St., Clinton, SC, 29325.

Corresponding author: James Wanliss (jawanliss@presby.edu)

Key Points:

- Fractal dimension calculation via the Higuchi method is shown to depend on the tuning parameter k_{max} and also the length of the time series.
- We derive a best-fit to provide researchers with an efficient method of estimating and appropriate k_{max}, given their particular dataset.

Data Availability:

- The data that support the findings of this study are available from the corresponding author, JAW, upon reasonable request.

Disclosure of potential conflicts of interest—ethical and financial:

- The authors declare that they have no conflict of interest.

Abstract

Higuchi's method of determining fractal dimension (*HFD*) is an important, well-used, research tool that, compared to many other methods, gives rapid, efficient, and robust estimations for the range of possible fractal dimensions. One major shortcoming in applying the method is the correct choice of tuning parameter (k_{max}); a poor choice can generate spurious results, and there is no agreed upon methodology to solve this issue. We analyze multiple instances of synthetic fractal signals to minimize an error metric. This allows us to offer a new and general method that allows determination, *a priori*, of the best value for the tuning parameter, for a particular length data set. We demonstrate its use on physical data, by calculating fractal dimensions for a shell model of the nonlinear dynamics of MHD turbulence, and severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1 from the family *Coronaviridae*.

1. Introduction

Since the seminal work of Mandelbrot & Van Ness (1968) the characterization of data in terms of fractal properties has found near ubiquitous and enduring use in diverse research areas, including research within the fields of engineering (Yang et al., 2021), hydrology (Zuo et al., 2009; Koutsoyiannis, 2019), geology (Turcotte, 1992; Ranguelov & Ivanov, 2017), physics (Wang et al., 2018), space science (Wanliss & Reynolds, 2003; Cersosimo & Wanliss, 2007), medicine (Mitsutake et al., 2004; Grizzi et al., 2019), economics (Fama et al., 2021), financial markets (Wątorek et al., 2021) and many more. Fractal properties in nature and human dynamics arguably have served to yield increased understanding and improvement on human society.

Higuchi's method (Higuchi, 1988) is a widely applied time-domain technique to determine fractal properties of complex non-periodic, nonstationary physical data (Esteller et al., 2001; Salazar-Varas & Vazquez, 2018; Yilmaz & Unal, 2020). That is, the method can accurately calculate the fractal dimension of time series. Higuchi initially developed it to study large-scale turbulent fluctuations of the interplanetary magnetic field. It is a modification to the method of Burlaga & Klein (1986) in which fluctuation properties of turbulent space plasmas can be studied beyond the inertial range. It is simple to implement, efficient, and can rapidly achieve accurate and stable values of fractal dimension, even in noisy, nonstationary data (Liehr & Massopust, 2020). The fractal dimension calculated via the Higuchi method is called the Higuchi fractal dimension (*HFD*). Since its initial development the Higuchi method has been applied to numerous fields of research. In medicine, for instance, it has found widespread use to detect and classify epileptic EEG signals (Lu et al., 2021), human locomotion (Santuz & Akay, 2020), and in engineering it has been used to detect faults in rolling bearings (Yang et al., 2021). One difficulty in using the Higuchi method is that certain parameters must be applied to the method, and inappropriate parameter selection results in spurious calculation of fractal properties. Although the method has been used for decades, and is widely employed at present, there is an absence of consensus of the appropriate method to determine the parameters that must be input. In this paper we expose this weakness of the Highuchi method so that there is wider appreciation of its limits, and suggest how to solve the drawbacks of this method when applied to different types of scientific data.

The *HFD* computed depends on the length of the time series, and an internal tuning factor k_{max}. Higuchi's original paper did not elaborate on the selection of the tuning factor but illustrated the method with $k_{max}=2^{11}$ for time series having length $N=2^{17}$. Subsequent authors used similar values for the tuning factor but we will show that the tuning factor plays a crucial

role in estimation the *HFD*. Higuchi's method, if applied appropriately, can reliably find the time series fractal dimension. However, if the tuning factor is incorrectly selected, the method is compromised from the outset.

How is the researcher to determine the appropriate tuning factor for their study that will optimize the calculation of a stable *HFD*, if it exists? In addition, how does the selection of the factor influence the value of the computed *HFD*? The literature is vague in answering these questions, and to do so is the main thrust of our research. Multiple studies have addressed the issue of proper selection of tuning factor k_{max}. Accardo et al. (1997) applied the method in their study of electroencephalograms, and sought the most suitable pair of (k_{max}, N). They experimented with k_{max}=3-10 on time series with lengths from N=50-1000, and settled on an optimum k_{max}=6. Some papers recommend plotting the *HFD* versus a range of k_{max}, and then selecting the appropriate k_{max} at the location where the calculated *HFD* approaches a local maximum or asymptote, which can be considered a saturation point (Doyle et al., 2004; Wajnsztejn et al., 2016). However, there is no reason that in every instance the *HFD* will reach a saturation point. Paramanathan and Uthayakumar (2008) proposed to determine the tuning factor k_{max} based on a size-measure relationship that employed a recursive length of the signal from different scales of measurement. Gomolka et al. (2018) selected k_{max} on the basis of statistical tests that allowed the best discrimination between already known diabetic and healthy subjects. But in the absence of such additional data between systems in different dynamic states (e.g. health or pathology), how can one select the correct tuning parameter?

In this paper we will try to answer these questions in a general way that is helpful to the community of researchers who utilize the Higuchi method. The organization of the paper is as follows. We will generate artificial time series with well-specified fractal dimension, then compare the *HFD* computed from these data for different values of the tuning parameter k_{max}. We will demonstrate the results on several examples of physical data.

2. Data and Method

In order to investigate the optimization of the Higuchi method we turn to the generation of synthetic time series with known fractal properties, to see how well the method performs. One difficulty resides in the the production of truly fractal time series of given dimension, which is a non-trivial task (Kijima and Tam, 2013). Therefore studies must concern themselves with the adequacy of the data-generating algorithms in addition to the fractal dimension estimation algorithms. We will consider synthetic time series realizations of processes with perfect and controlled scale invariance, viz. signals that have only a single type of scaling. Many other theoretical data types exist that have been used to analyze signals that lack local scaling regularity, but rather have a regularity which varies in time or space (Lévy Véhel, 2013; Wanliss et al., 2014). There is also a recent effort to generalize the Higuchi method to distinguish monofractal from multifractal dynamics based on relatively short time series (Carrizales-Velazquez et al., 2021).

In this paper we will limit the research to study of well-understood synthetic data with monofractal scaling. To illustrate how a monofractal scaling exponent can be derived we consider fractional Brownian motion (fBm) which is characterized by a single stable fractal dimension and is a continuous-time random process (Mandelbrot and Van Ness, 1968). Next, we research the these data and compare the fractal dimension recovered using the Higuchi algorithm

with the theoretical fractal dimension. The synthetic data time series can be written in terms of stochastic integrals of time integrations of fractional Gaussian noise:

$$\boldsymbol{B_H(t)} = \frac{\mathbf{1}}{\boldsymbol{\Gamma\left(H+\frac{1}{2}\right)}}\left\{\int_{-\infty}^{\mathbf{0}}[\boldsymbol{(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}}}]\boldsymbol{dW(s)} + \int_{\mathbf{0}}^{\boldsymbol{t}}\boldsymbol{(t-s)^{H-\frac{1}{2}}dW(s)}\right\}.$$

Here **W** is a stationary and ergodic random white noise process with zero mean defined on (-∞,∞). In the above equation $\boldsymbol{H} \in (\mathbf{0}, 1)$ is known as the Hurst exponent. The time series Hurst exponent is related to signal roughness averaged over multiple length scales. The higher the value of $\boldsymbol{H}$, the smoother is the time series, and the longer trends tend to continue. For values closer to zero, the time series rapidly fluctuates, as shown in Fig. 1. The covariance function of the noisy signal can be expressed by:

$$\boldsymbol{cov\{B_H(s), B_H(t)\}} = \frac{\mathbf{1}}{\mathbf{2}}\{\boldsymbol{|s|^{2H} + |t|^{2H} - |s-t|^{2H}}\},$$

so that $\boldsymbol{B_H(0) \equiv 0}$ and $\boldsymbol{var\{B_H(t)\} = t^{2H}}$. For **H**=1/2 the white noise process reduces to the well-known random walk. The theoretical relationship between the Hurst exponent, $\boldsymbol{H}$, and the Higuchi fractal dimension, *HFD,* is $\boldsymbol{HFD = 2 - H}$, with values of *HFD* between 1 and 2.

We consider four different method generators of processes having long-range dependence to generate synthetic series with exact fractal dimension. First, we consider an exact wavelet-based method. This is based on a biorthogonal wavelet method proposed by Meyer and Sellan (Abry and Sellan, 1996; Bardet et al., 2003) and implemented in Matlab software and the wfbm calling function. The second is the method of Davies & Harte (1987) whose generation process uses a fast fourier transform basis and embeds the covariance matrix of the increments of the fractional Brownian motion in a circulant matrix. The third category of synthetic simulated data is produced using the Wood-Chan circulant matrix method (Wood and Chan, 1994), which is a generalization of the previous method (Coeurjolly, 2001). The fourth set of data are simulated using the Hosking method (Hosking, 1984), also known as the Durbin or Levinson method (Levinson, 1947), which utilizes the well-known conditional distribution of the multivariate Gaussian distribution on a recursive scheme to generate samples based on the explicit covariance structure. All these methods of producing simulated data are considered exact methods because they completely completely capture the covariance structure and produce a true realization of series with a single scaling parameter.

Fig. 1 shows various examples of time series produced via the Davies and Harte (1987) method. The smoothest curve corresponds to $\boldsymbol{H}$=0.9, which implies high probability to observe long periods with increments of same sign. The roughest curve corresponds to $\boldsymbol{H}$=0.1, which is sub-diffusive, with high probability that increments feature long sequences of oscillating sign. The curves show data for Hurst exponents $\boldsymbol{H}$=0.3,0.5,0.7,0.9, from top to bottom.

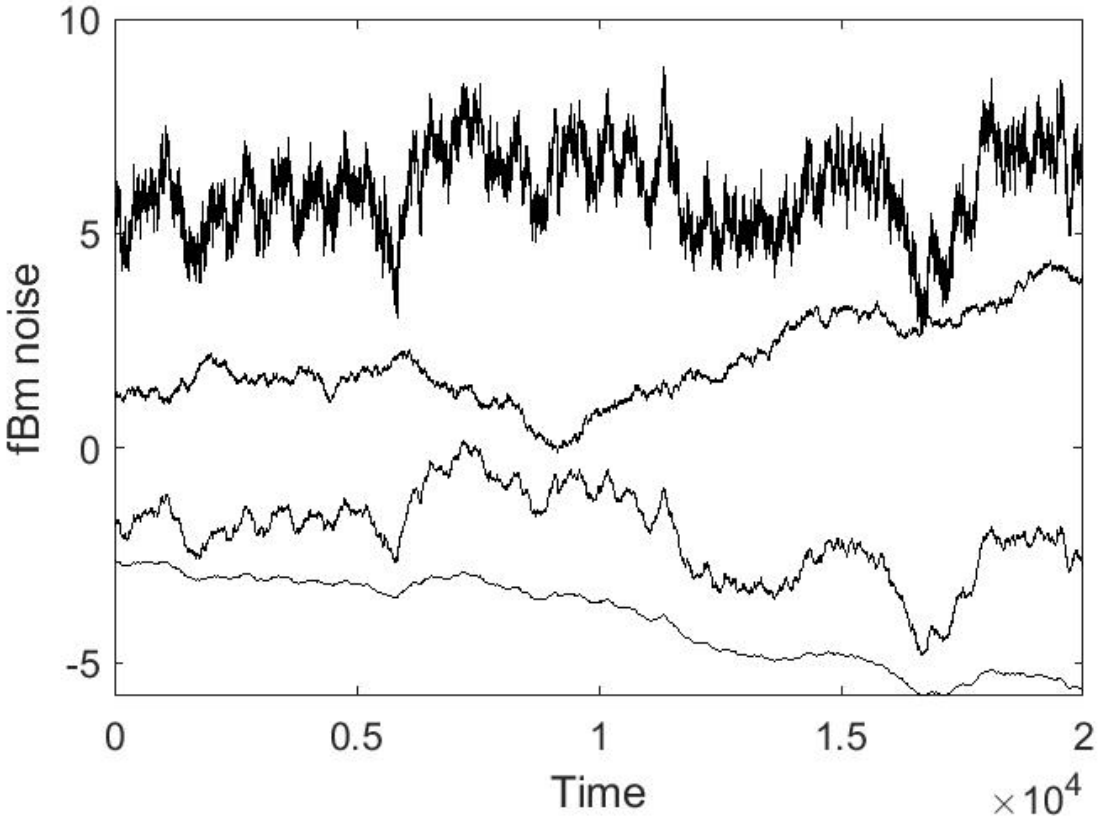


Figure 1. Examples of synthetic time series from the Davies and Harte (1987) method, characterized by Hurst exponent **H**=0.3,0.5,0.7,0.9 (top to bottom).

For each of the four data generating methods we create 100 unique time series, of differing lengths up to maximum length 500,000 data points, for Hurst exponents $\boldsymbol{H}$=0.1,0.3,0.5,0.7,0.9. Thus, for each time series length *N* we have 500 unique simulations of fBm for each method. This produces 44,000 data sets in total, for experimentation. We next apply the Higuchi method to each of these time series with an exact fractal dimension (FD)to determine how well the Higuchi method is able to accurately recover the theoretical value compared to the derived *HFD*.

Next we describe the Higuchi method. The Higuchi method takes a signal, discretized into the form of a time series, $\boldsymbol{x(1), x(2), \dots, x(N)}$ and, from this series, derives a new time series, $\boldsymbol{X_k^m}$, defined as:

$$\boldsymbol{X_k^m: x(m), x(m+k), x(m+2k), \dots, x\left(m + \left[\frac{N-k}{k}\right] \cdot k\right)},$$

Here [] represents the integer part of the enclosed value. The integer $\boldsymbol{m = 1, 2, \dots, k}$ is the start time and $\boldsymbol{k}$ is the time interval, with $\boldsymbol{k = 1, \dots, k_{max}; k_{max}}$ is a free tuning parameter. This means that given time interval equal to $\boldsymbol{k}$, spawns $\boldsymbol{k}$-sets of new time series. For instance, if $\boldsymbol{k =}$

156 $\mathbf{10}$ and the time series has length $\boldsymbol{N = 1000}$, the following new time series are derived from the
157 original data:

158

159 $$\boldsymbol{X_{10}^{1}}\boldsymbol{: x(1), x(11), x(21), \dots, x(991)),}$$

160 $$\boldsymbol{X_{10}^{2}}\boldsymbol{: x(2), x(12), x(22), \dots, x(992)),}$$

161 .

162 .

163 .

164 $$\boldsymbol{X_{10}^{10}}\boldsymbol{: x(10), x(20), x(30), \dots, x(1000).}$$

165

166 These curves have lengths defined by:

167

168 $$\boldsymbol{L_m(k) = \frac{\left\{\left(\sum_{i=1}^{\left\lfloor\frac{N-m}{k}\right\rfloor}|x(m+ik) - x(m+(i-1)\cdot k|\right)\frac{N-1}{\left\lfloor\frac{N-m}{k}\right\rfloor \cdot k}\right\}}{k}}.$$

169

170 The final term in the numerator is a normalization factor, $\boldsymbol{N - \frac{1}{\left\lfloor\frac{N-m}{k}\right\rfloor}\cdot k}$. The length of the
171 curve for the time interval $\boldsymbol{k}$ is then defined as the average over the $\boldsymbol{k}$ sets of $\boldsymbol{L_m(k)}$:

172

173 $$\boldsymbol{L(k) = \langle L_m(k)\rangle.}$$

174

175 In cases when this equation scales according to the rule $\boldsymbol{L(k) \propto k^{-HFD}}$, we consider the
176 time series to behave as a fractal with dimension *HFD*. Thus, the *HFD* is the slope of the straight
177 line that fits the curve of *ln(L(k))* versus *ln(1/k)*. Fig. 2 shows the L(k) curve from simulated data
178 for the fractal dimension *FD*=1.7 (corresponding to **H**=0.3) time series data in Fig. 1. The
179 corresponding curve of *HFD*(k_{max}) is shown in Fig. 3.

180

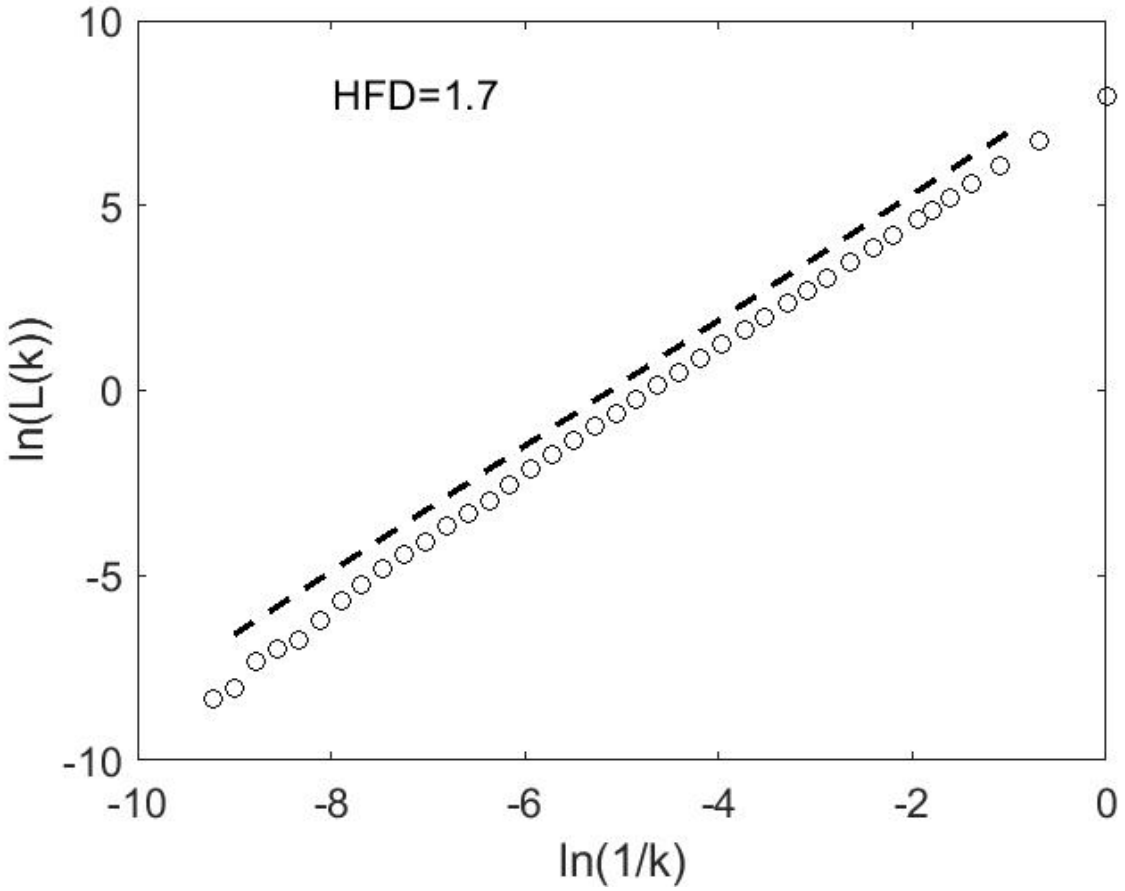


Figure 2. Average curve length versus scale size, k, for the time series with HFD=1.7.

We now turn to finding the best tuning parameter, k_{max}, for the set of data we have simulated. As discussed in the Introduction, a common way to determine the tuning parameter relies upon finding the location, in plots like Fig. 3, of *HFD* versus a range of k_{max}, where the calculated *HFD* approaches a local maximum or asymptote (Doyle et al., 2004; Wajnsztejn et al., 2016). We will call this a tuning curve. In Fig. 3, which is for the time series with *HFD*=1.7, there is only one local maximum which is located at k_{max}=7 which produces a negligible error of 0.5%. There are three places where the Higuchi method finds a best value is achieved for this simulation, viz. k_{max} =4, 14, 727. In this particular instantiation of a fBm the most effective tuning parameter would thus be k_{max}=4, 14, or 727. The easiest method would be to use the smallest k_{max} since this results in the least computational effort. However, in this case, using the local maxima method yields an acceptable estimate result with little additional effort.

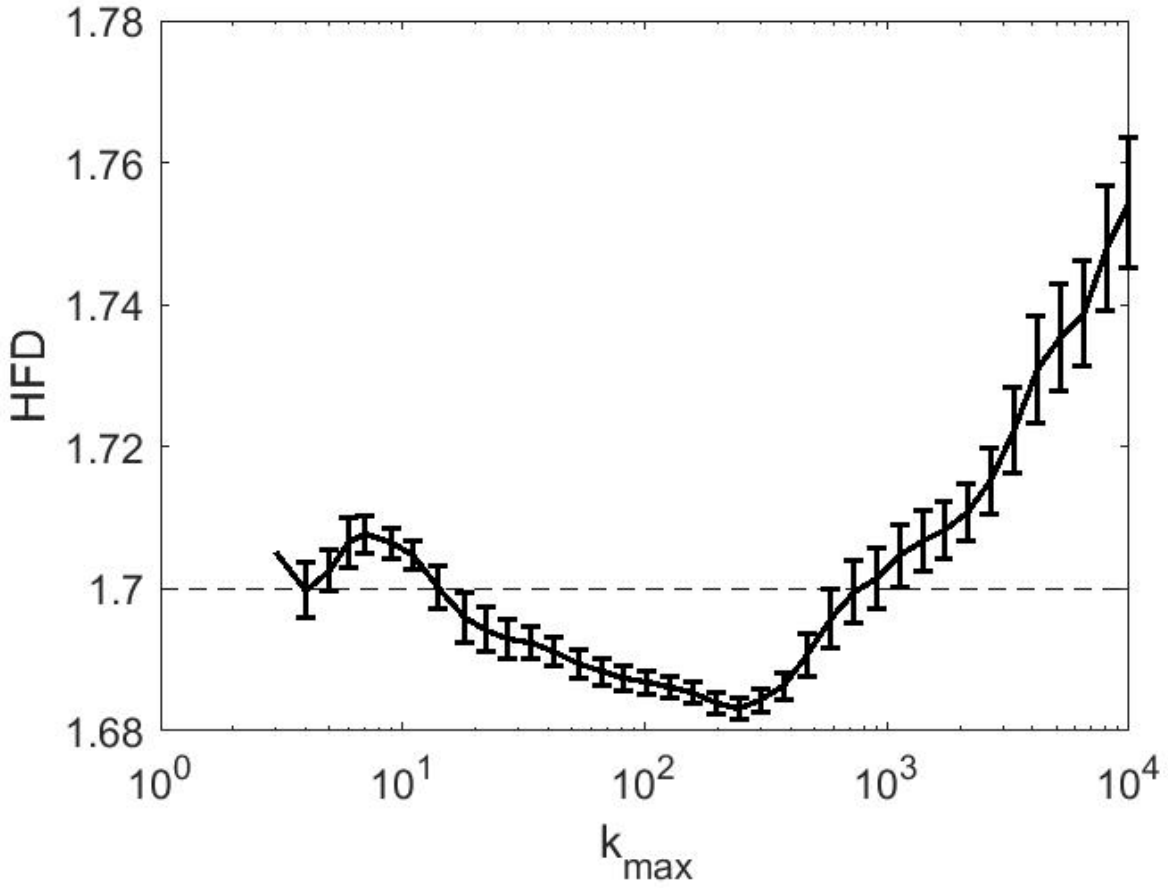


Figure 3. Curve showing the relationship between *HFD* and k_{max} for the *FD*=1.7 time series shown in Fig. 1.

3 Results

There is no reason to expect that a local maxima exists in every case in a tuning curve and is therefore searching through these curves for asymptotes is not a general or practical method to determine the best tuning parameter k_{max}. For instance, Fig. 4 shows the tuning curves for *HFD*=1.9, 1.5, 1.3, 1.1 computed from the simulated data of Fig. 1. The black horizontal dashed line in each subplot shows the theoretical value of the fractal dimension. There is not always a local maximum or an asymptotic convergence to a set value of *HFD*. For *HFD*=1.9 a peak occurs but only near k_{max} ~5000; the region of the plateau is found at the tuning parameter that yields the largest error in fractal dimension. This indicates that in this fBm realization a much smaller k_{max} would be appropriate.

We now turn to analyzing the simulated realizations of fBm. The smallest time series length we select has *N*=1,000, and the largest has *N*=500,000 data points, and compute the *HFD* for each of these series, as a function of tuning parameter k_{max}. We use values between k*max* =2 and k_{max} =*N*/2. This gives a new data set comprised of *HFD* values as a function of the time

series length, and the tuning parameter, yielding $HFD=HFD(N,k_{max})$. The error to be minimized is written by:

$$E(N, k_{max}) = 100 * \frac{[HFD - FD_{theory}]}{FD_{theory}}.$$

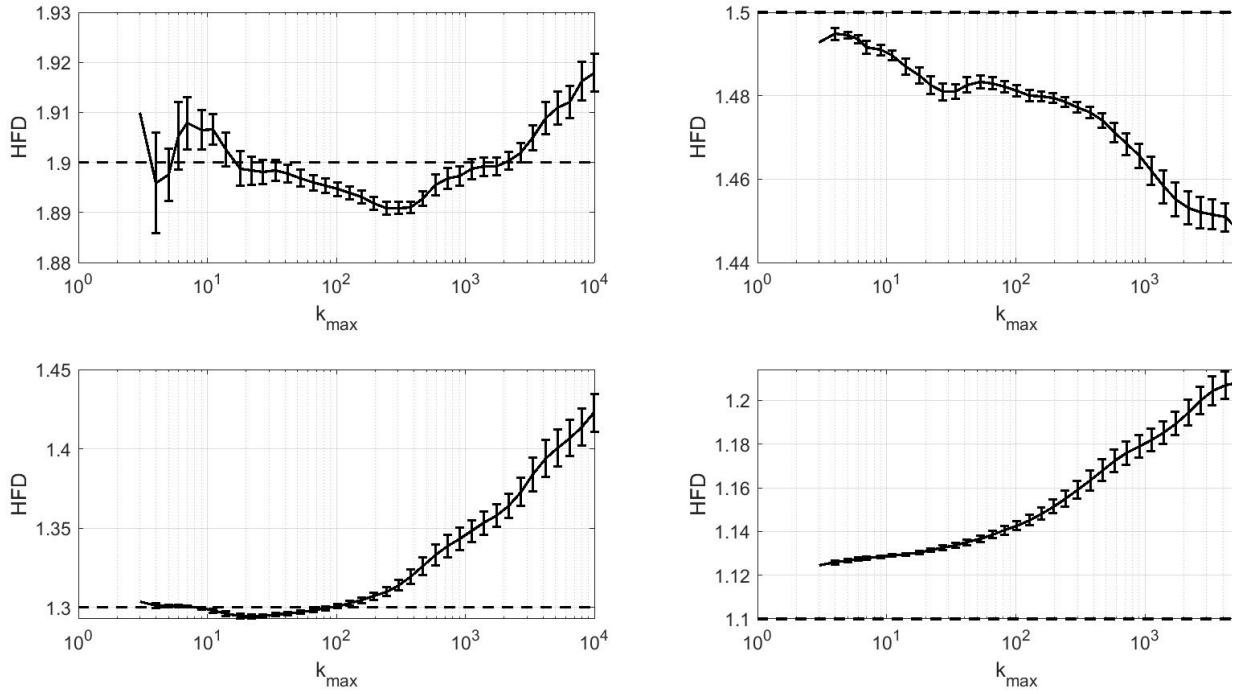


Figure 4. Curves showing the relationship between *HFD* and k_{max} for the *HFD*=1.9,1.5,1.3,1.1 time series shown in Fig. 1. The dashed horizontal curves show the theoretical value for the *HFD*.

The previous equation gives the percentage error to be averaged over all synthetic time series simulations to yield a general result for all simulation data considered. As researchers do not generally know *a priori* which method of simulating artificial data most closely follows the statistics of any particular physical or research data set it is appropriate to use a range of synthetic simulated time series with known fractal dimension, as an average result gives the most general answer. Fig. 5 shows surface plots comparing the percentage error *HFD* versus the tuning parameter k_{max} and time series length, *N* for the theoretical *HFD*=1.7 for each of the four simulation methods described in the previous section. The curve of least error is shown as a thick grey line.

Each method of simulation yields a different curve of least error. Fig. 5b, which is the curve for the Wood-Chan circulant matrix method (Wood and Chan, 1994), has the lowest overall error. The Hosking (1984) method yields *HFD* values with the greatest errors (Fig. 5d). Overall, the location of the minimum error curve varies widely depending on the generation algorithm for the synthetic data.

By taking a geometric mean of these minimum error curves for all *HFD* values we derive a best-fit curve using a sum of sines function since this gave a simple function with few terms and a fit with small sum squared error. Fig. 6 shows the relationship between the time series length and the tuning parameter, for different *HFD* values, and the dashed curve shows the best fit, given by the following equation:

$$k_{max} = [A_1 \sin(B_1 * N + C_1) + A_2 \sin(B_2 * N + C_2)]. \qquad (*)$$

Here [] represents the integer part of the enclosed function value. Table 1 shows the parameter values for the best-fit. Fig. 6 shows that for short time series the use of a plateau criterion to select the k_{max} tuning parameter will result in the use of values smaller than those proposed by this generalized study. For example, in Fig. 1 a time series of length N=20,000 is used. Fig. 3 shows the curve of *HFD* versus k_{max}. Our fitting function yields k_{max}=47 for this length data set.

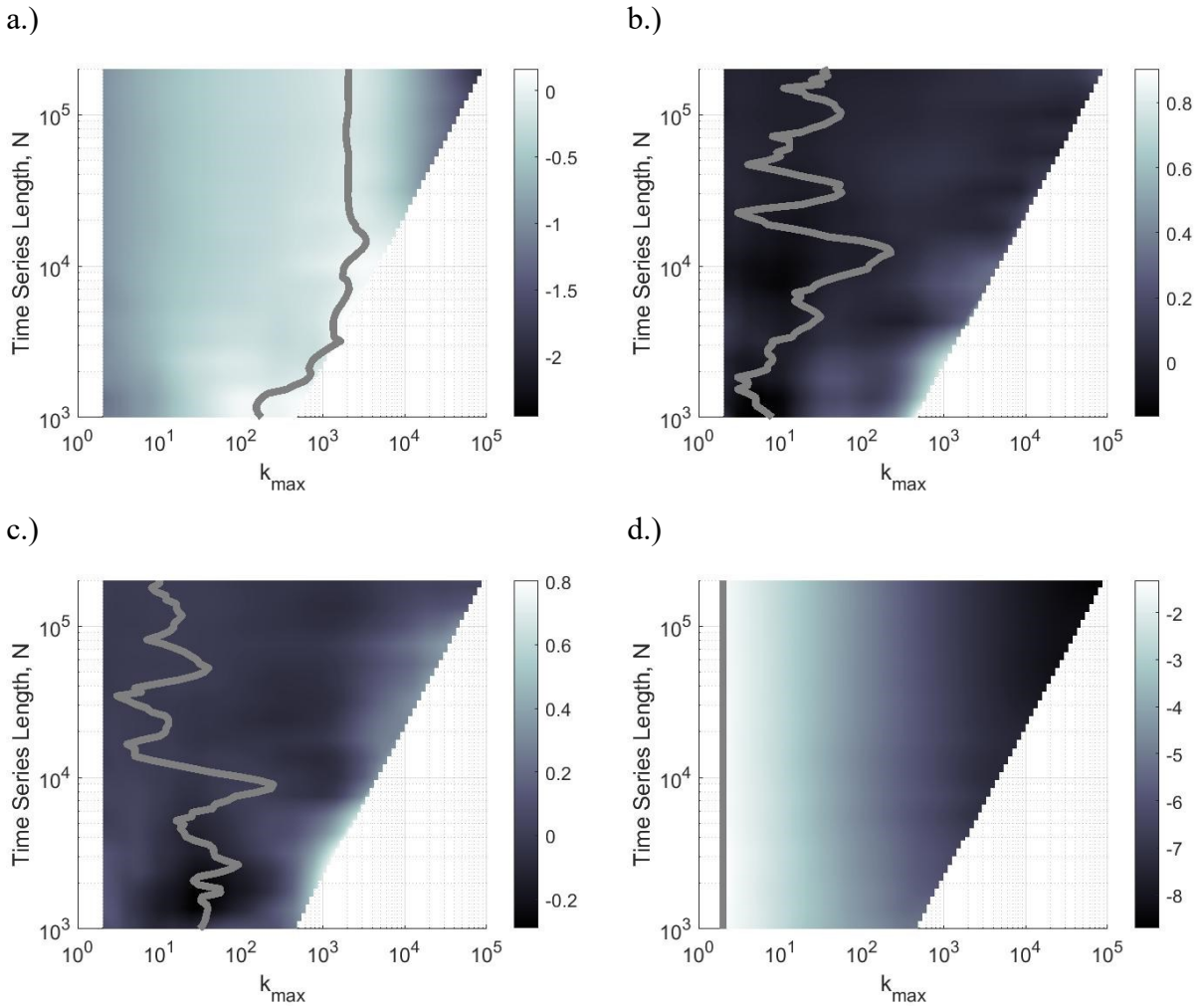


Figure 5. Surface showing the average percentage error between the Higuchi method fractal dimension and theoretical *FD*=1.7 averaged over 100 datasets of different length, N. The curve

of least error is shown (a.) Wavelet generation method, (b.) Wood-Chan method, (c.) Davies-Harte method, and (d.) Hosking method. The curve of least error is shown as a thick grey line.

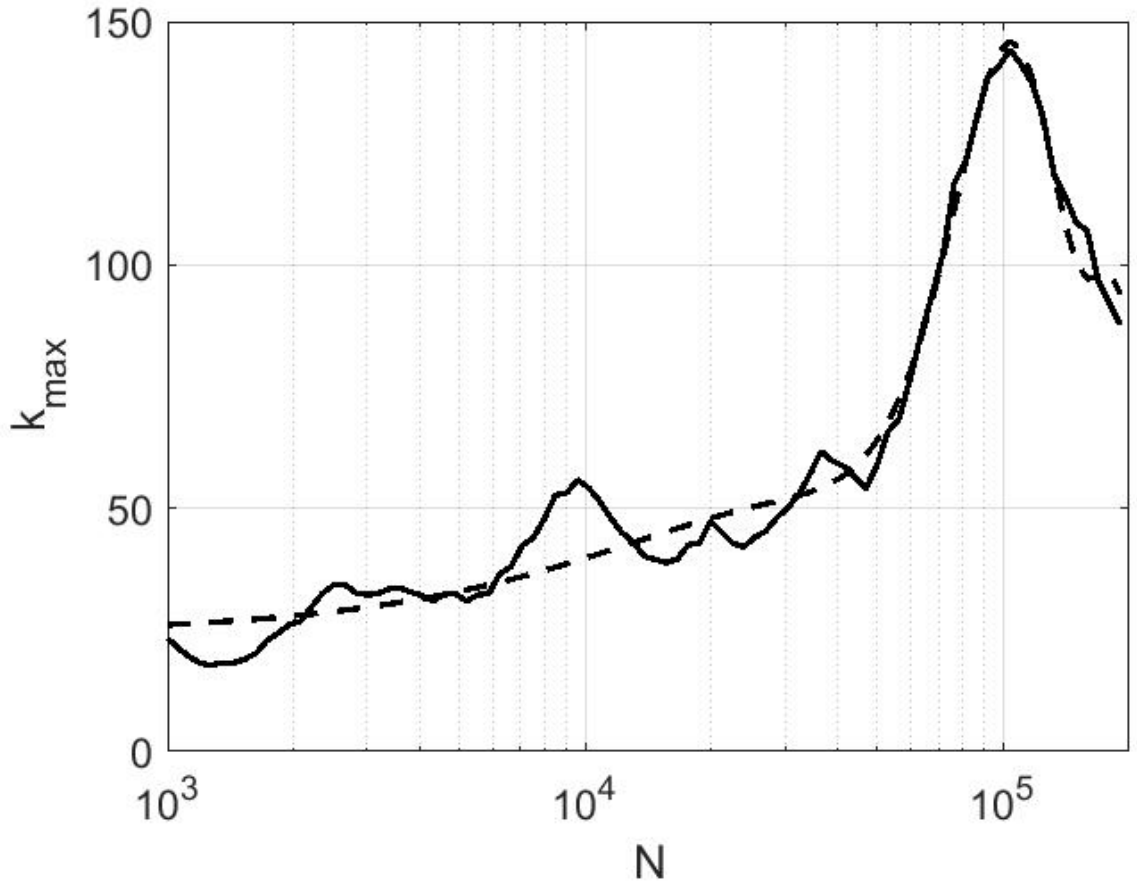


Figure 6. Comparison of the average minimum error curve (solid) and the best fit sum of sines function (dashed).

Table 1. Fitting parameters for best-fit sum of sin function, $k_{max} = [A_1\, sin(B_1 * N + C_1) + A_2\, sin(B_2 * N + C_2)]$.

A_1	129.8±3.0
B_1	(1.292±0.045)×10^{-5}
C_1	0.04488±0.0255
A_2	18.82±2.56
B_2	(6.488±0.280)×10^{-5}
C_2	1.332±0.220

4 Applications

In this section we present two applications of the Higuchi method with the corrections applied to determing the appropriate tuning parameter. The first is a shell model of the nonlinear dynamics of MHD turbulence. We effect this via simplified approximations of the Navier–Stokes fluid equations (Obukhov, 1971; Gledzer, 1973; Yamada and Ohkitani, 1988). We use the MHD Gledzer–Ohkitani–Yamada (GOY) shell model, which captures the intermittent dynamics of the energy cascade in MHD turbulence (Lepreti et al., 2004) as it moves along through the shells in a front-like manner.

Shell models of MHD turbulence are an example of dynamical systems incorporating simplified versions of the Navier–Stokes or MHD turbulence equations. They attempt to conserve some of the invariants in the limit of no dissipation.We use the SHELL-ATM code (Buchlin & Velli, 2006) to produce a time series of length N=500,000 of the magnetic energy dissipation rate (ϵ_b) as a function of time obtained in the MHD shell model (Fig. 7a). The model is described in detail in Lepreti et al. (2004). In short, the SHELL-ATM model makes it possible to obtain rapid simulations of MHD turbulence in volumes in which a longitudinal magnetic field dominates. Model construction begins via division of the wave-vector space (k-space) into a number, *N*, of discrete shells with known radius $k_n = k_0 2^n$ (*n*=0,1,…,*N*) (Giuliani & Carbone, 1998). Each shell is then assigned complex dynamical Elsässer-like fields $u_n(t)$ and $b_n(t)$, which represent longitudinal velocity increments and magnetic field increments. The magnetic energy disspation rate is defined by

$$\epsilon_b(t) = \eta \sum_{n=1}^{N} k_n^2 |b_n^2|$$

where η is the kinematic resistivity. To find the solutions to the above equations we solve the equations

$$\frac{db_n}{dt} = -\eta k_n^2 b_n + \frac{1}{6} i k_n (u_{n+1} b_{n+2} - b_{n+1} u_{n+2}) - \frac{1}{6} i k_n [(u_{n-1} b_{n+1} - b_{n-1} u_{n+1}) + (u_{n-2} b_{n-1} - b_{n-2} u_{n-1})]^* + f_n$$

and

$$\frac{du_n}{dt} = -\nu k_n^2 u_n + i k_n (u_{n+1} u_{n+2} - b_{n+1} b_{n+2}) - \frac{1}{4} i k_n \left[(u_{n-1} u_{n+1} - b_{n-1} b_{n+1}) + \frac{(u_{n-2} b_{n-1} - b_{n-2} u_{n-1})}{2} \right]^* + g_n$$

where ν is the kinematic viscosity and (f_n, g_n) are forcing terms operating on the magnetic and velocity increments. The symbol * represents a complex conjugate. The forcing terms are calculated from the Langevin equation driven by a Gaussian white noise.

These data in Fig. 7a display clear intermittent bursts of dissipated energy. Fig. 7b shows average curve length versus scale size, k, for the time series. Fig. 7c shows the relationship between *HFD* and k_{max}. There is no asymptote which may indicate an appropriate value of k_{max}. We now use Eqn. (*) to select the appropriate tuning parameter k_{max} determined from our prior analysis for data featuring a single fractal scaling, for varying lengths, N, of the time series. Fig. 7d shows the computed *HFD* selected. There is a variation in the fractal dimension with values being estimated as smaller from shorter lengths of the time series, and overall *HFD*~1.04-1.13.

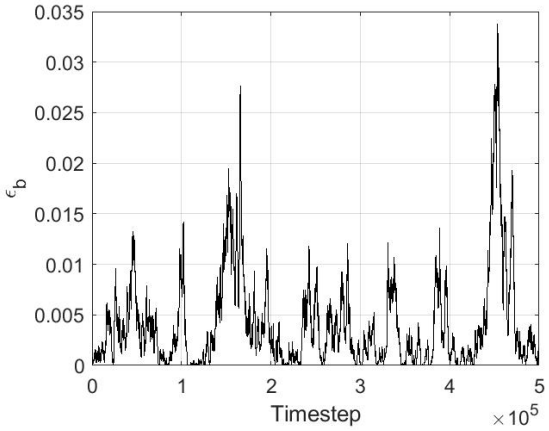


Fig. 7a. Magnetic energy dissipation rate for the GOY shell model.

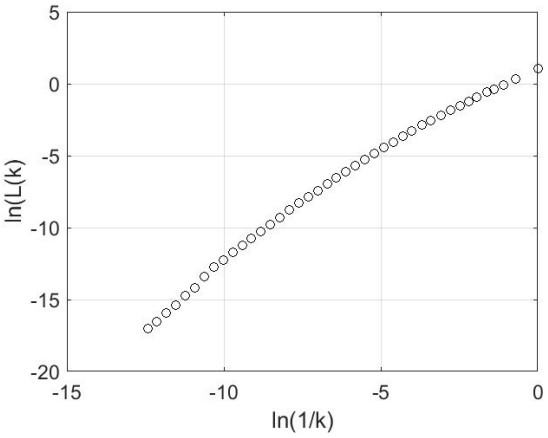
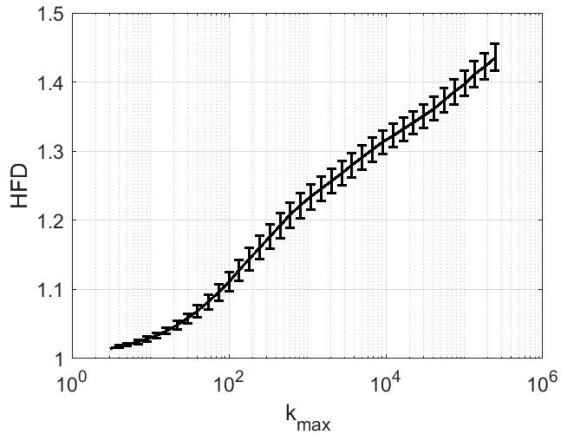
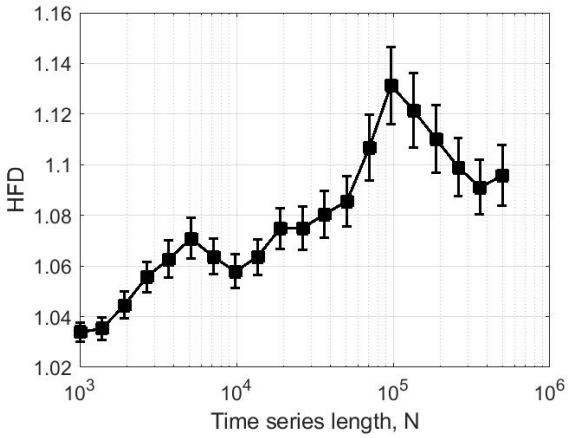


Fig. 7b. Average curve length versus scale size, k.





The second data example is that of the severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1. Wu et al. (2020) reported on the identification of the novel RNA virus strain from the family *Coronaviridae*, which is designated here 'WH-Human-1' coronavirus. We obtained these data from the National Center for Biotechnology Information (NCBI), which is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health (NIH).

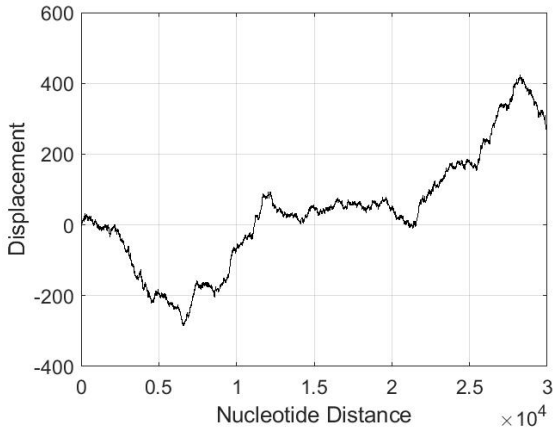


Fig. 8a. WH-Human-1 complete genome represented by the Peng (1992) method.

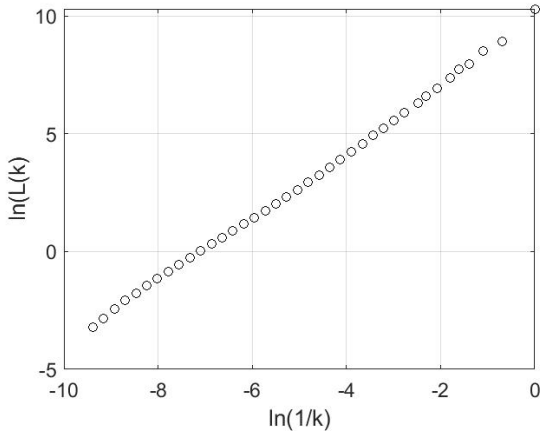
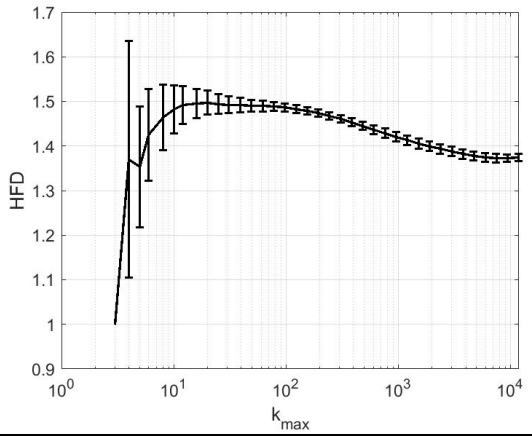
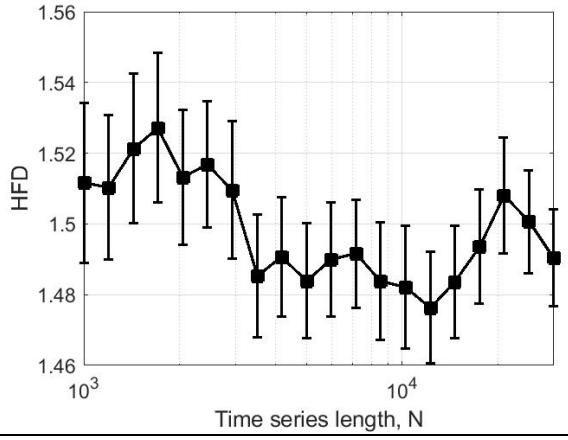


Fig. 7b. Average curve length versus scale size, k.





To analyze the fractal patterns in the genome one must convert the nucleotide sequence from a symbolic sequence, meaning A,G,C,T into a time series. We follow the Peng (1992) method in which DNA is represented as a "random walk" with two parameters ruling the direction of the "walk" and the resulting dynamics. We start with the first nucleotide. If it is a pyrimidine base, we move up one position. Every subsequent pyrimidine base moves up one position. When a purine base is encountered in the series the walk steps down one position. The nucleotide distance from the first nucleotide is then plotted versus the displacement, as in Fig. 8a. Fig. 8b shows average curve length versus scale size, k, for the time series. Fig. 8c shows the relationship shows the computed *HFD* against tuning parameter k_{max} from the whole time series of length N=29,903. In this case there is a distinct asymptote at k_{max}=20, which yields *HFD*=1.497. To test our method we again use Eqn. (*) to select the appropriate tuning parameter k_{max}, for varying lengths, N, of the time series. Fig. 8d shows the computed *HFD* selected. There is no statistically significant variation in the fractal dimension with values being estimated at HFD~1.5.

Our analysis shows that the fractal dimension of WH-Human-1 coronavirus genome is different from its fractal dimension computed from electron microscopic and atomic force

microscopic images of 40 coronaviruses (CoV), as reported by Swapna et al. (2021) who found a scale-invariant dimension of 1.820. This indicated that the images of the virus feature higher complexity and greater roughness than the pattern we have detected in the genome.

5 Conclusions

Higuchi's method to compute the fractal dimension of physical signals is widely used in research. However, a major difficulty in applying the method is the correct choice of tuning parameter (k_{max}) to compute the most accurate results. Poor selection of k_{max} can result in values of the fractal dimension that are spurious, and this can result in potentially invalid interpretations of data. In the past researchers have used various ad hoc methods to determine the appropriate tuning parameter for their particular data. We have shown that a method such as seeking a convergence of the computed *HFD* to a plateau is not in general a valid procedure as not every data instance shows the *HFD* estimate reaches a plateau.

In this paper we have sought to find a more general method of determining, a priori, the optimum tuning parameter k_{max} for a time series of length N. To study this problem we generated synthetic time series of known *HFD* and applied the Higuchi method to each, averaging results over the different fbm within *HFD*=[1.9,1.7,1.5,1.3,1.1] categories. These data allow the calculation of curves showing where in (N, k_{max})-space the most appropriate tuning parameter should be selected. We found that fractal dimension calculation via the Higuchi method is sensitive to both the tuning parameter k_{max} and also the length of the time series. We derive a best-fit curve fitting the location of the average minimum *HFD* error to provide researchers with an efficient method of estimating and appropriate k_{max}, given their particular dataset.

We applied the modified method to two physical cases, one from physics and one from bioinformatics. In the latter case we considered the *Coronaviridae* genome of the severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, first reported by Wu et al. (2020). Our analysis of this data showed strong evidence of monofractality (Fig. 8b) with *HFD*~1.5 (Fig. 8d).

In the former case we computed the magnetic energy dissipation rate from a shell model of the nonlinear dynamics of of MHD turbulence. We used simplified approximations of the Navier–Stokes fluid equations (Obukhov, 1971; Gledzer, 1973; Yamada and Ohkitani, 1988), in particular the MHD Gledzer–Ohkitani–Yamada (GOY) shell model, and found *HFD*~1.10 (Fig. 7d). These data have been reported to feature a multifractal scaling (Pisarenko et al., 1993) and this is consistent with our results in Fig. 8b which show evidence of nonlinear behaviour, which is possibly a reason why there is about a 10 percent variation in the HFD calculation (Fig. 7d).

It is clear that accurate calculation of fractal dimension can be a delicate process and is influenced not only by the method used, but also by the nature of the data. Studies must therefore concern themselves not only with the type of data, but also with the adequacy of the data-generating algorithms, and fractal estimation algorithms. We considered only synthetic time series realizations of processes with perfect and controlled scale invariance, viz. signals that have only a single type of scaling. However, many other theoretical data types exist. For instance, numerous geophysical signals do not have local scaling regularity, but rather have a regularity which varies in time or space (Dobias & Wanliss, 2009; Lévy Véhel, 2013). Data that are multifractal require a variety of scaling exponents to fully describe the dynamics, and methods to generalize the Higuchi method to these more complex data types are going forward at present (Carrizales-Velazquez et al., 2021).

Acknowledgment

National Science Foundation Award AGS-2053689 and SC-INBRE.

References

Abry, P., F. Sellan (1996), *The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation*, Appl. and Comp. Harmonic Anal., 3(4), 377–383.

Accardo, A. et al (1997), *Use of the fractal dimension for the analysis of electroencephalographic time series*, *Biol. Cybern.* 77 (5) (1997) 339–350.

Bardet, J.-M., G. Lang, G. Oppenheim, A. Philippe, S. Stoev, M.S. Taqqu (2003), *Generators of long-range dependence processes: a survey*, Theory and applications of long-range dependence, Birkhäuser, 579–623.

Burlaga, L. F. and Klein, L. W. (1986), *Fractal Structure of the Interplanetary Magnetic Field*, J. Geophys. Res., 91, A1, 347–350.

Carrizales-Velazquez, C., Donner, R. V., & Guzmán-Vargas, L. (2021), Generalization of Higuchi's fractal dimension for multifractal analysis of time series with limited length, arXiv preprint arXiv:2105.11055.

Cersosimo, D. O., and J. A. Wanliss (2007), *Initial studies of high latitude magnetic field data during different magnetospheric conditions*, Earth Planets Space, 59(1), 39– 43.J.-F. Coeurjolly (2000), *Simulation and identification of the fractional Brownian motion: a bibliographic and comparative study*, Journal of Statistical Software, Vol. 5, 7, 1-53, doi:// 10.18637/jss.v005.i07

Davies, Robert B., and D. S. Harte. (1987), *Tests for Hurst effect*, Biometrika 74, no. 1 (1987): 95-101.

Dobias, P. and Wanliss, J.A. (2009), *Intermittency of Storms and Substorms: Is it Related to the Critical Behaviour?*, Annales Geophysicae, 27, 5, 2011-2018, https://doi.org/10.5194/angeo-27-2011-2009.

Doyle, T. L. A. (2004), *Discriminating between elderly and young using a fractal dimension analysis of centre of pressure*, Int. J. Med. Sci. 2004 1(1): 11-20.

Esteller, R., G. Vachtsevanos, J. Echauz, and B. Litt (2001), *A comparison of waveform fractal dimension algorithms*, IEEE Trans Circ Syst I Fundam Theory Appl, 48:177–183.

Fama, Eugene F., Cochrane, John H. and Moskowitz, Tobias J.. *The Fama Portfolio: Selected Papers of Eugene F. Fama*, Chicago: University of Chicago Press, 2021. https://doi.org/10.7208/9780226426983

Giuliani, P. and Carbone, V. (1998). *A note on shell models for MHD turbulence*. Europhys. Lett. 43, 527-532, https://doi.org/10.1209/epl/i1998-00386-y.

Gledzer, E. B. (1973), *System of Hydrodynamic Type Allowing 2 Quadratic Integrals of Motion*, Sov. Phys. Dokl. SSSR, 18, 216–217, 1973.

Gomolka R.S. et al (2018), *Higuchi Fractal Dimension of Heart Rate Variability During Percutaneous Auricular Vagus Nerve Stimulation in Healthy and Diabetic Subjects*, Front. Physiol. 9:1162. doi: 10.3389/fphys.2018.01162

Grizzi, F., Castello, A., Qehajaj, D., Russo, C., & Lopci, E. (2019). *The complexity and fractal geometry of nuclear medicine images*. Molecular Imaging and Biology, 21(3), 401-409.

Higuchi, T. (1988), *Approach to an irregular time series on the basis of the fractal theory*, Phys D 31:277–283.

Hosking, Jonathan RM. (1984), *Modeling persistence in hydrological time series using fractional differencing*, Water resources research 20, no. 12 (1984): 1898-1908.

Kijima, M. and Chun Ming Tam (2013), *Fractional Brownian Motions in Financial Models and Their Monte Carlo Simulation, Theory and Applications of Monte Carlo Simulations*, Victor (Wai Kin) Chan, IntechOpen, DOI: 10.5772/53568.

Koutsoyiannis, D. *Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes*. Hydrol. Sci. J. 2019, 64, 1013–1037.

Lepreti, F., Carbone, V., Giuliani, P., Sorriso-Valvo, L., and Veltri, P. (2004), *Statistical Properties of Dissipation Bursts within Turbulence: Solar Flares and Geomagnetic Activity*, Planet. Space Sci., 52, 957–962, https://doi.org/10.1016/j.pss.2004.03.001, 2004.

Levinson, N. (1947), *The wiener rms error criterion in filter design and prediction*, Journal of Mathematics and Physics, Vol 25 (1947) 261–278.

Lévy Véhel, J.*: Beyond multifractional Brownian motion: new stochastic models for geophysical modelling*, Nonlin. Processes Geophys., 20, 643–655, https://doi.org/10.5194/npg-20-643-2013, 2013.

Liehr, Lukas, Massopust, Peter (2020), *On the mathematical validity of the Higuchi method*, Physica D: Nonlinear Phenomena, 402, 132265, doi:10.1016/j.physd.2019.132265.

Lu, XiaoJie, JiQian Zhang, ShouFang Huang, Jun Lu, MingQuan Ye, MaoSheng Wang, *Detection and classification of epileptic EEG signals by the methods of nonlinear dynamics*, Chaos, Solitons & Fractals, Volume 151, 2021, 111032, doi:10.1016/j.chaos.2021.111032.

Mandelbrot, B. and Van Ness, J.W. (1968), *Fractional Brownian motions: Fractional noises and applications*, SIAM Rev. 10, No. 4, 422–437.

Mitsutake G, Otsuka K, Oinuma S, Ferguson I, Cornélissen G, Wanliss J, Halberg F (2004), *Does exposure to an artificial ULF magnetic field affect blood pressure, heart rate variability and mood?* Biomed Pharmacother 58:S20–S27. doi:10.1016/S0753-3322(04)80004-0

Nobukawa, Sou, Yamanishi, Teruya, Nishimura, Haruhiko, Wada, Yuji, Kikuchi, Mitsuru, and Takahashi, Tetsuya (February 2019), *Atypical temporal-scale-specific fractal changes in Alzheimer's disease EEG and their relevance to cognitive decline*, Cognitive Neurodynamics, 13 (1): 1–11, doi:10.1007/s11571-018-9509-x

Obukhov, A. M. (1971), *Some General Properties of Equations Describing The Dynamics of the Atmosphere*, Akad. Nauk. SSSR, Izv. Serria Fiz. Atmos. Okeana, 7, 695–704, 1971.

Paramanathan, P. and Uthayakumar, R. (2008), *Application of fractal theory in analysis of human electroencephalographic Signals*, Comput. Biol. Med. 38 (3) (2008) 372–378.

Pisarenko, D., Biferale, L., Courvoisier, D., Frisch, U., & Vergassola, M. (1993), *Further results on multifractality in shell models*, Physics of Fluids A: Fluid Dynamics, 5(10), 2533-2538.

Ranguelov, B., & Ivanov, Y. (2017). *Fractal properties of the elements of Plate tectonics*. Journal of mining and Geological Sciences, 60(1), 83-89.

Salazar-Varas, R. and R.A. Vazquez (2018), *Time-Invariant EEG Classification Based on the Fractal Dimension*, In: Castro F., Miranda-Jiménez S., González-Mendoza M. (eds) *Advances in Computational Intelligence.* MICAI 2017. Lecture Notes in Computer Science, vol 10633. Springer, Cham. https://doi.org/10.1007/978-3-030-02840-4_26

Santuz, Alessandro and Turgay Akay, *Fractal analysis of muscle activity patterns during locomotion: pitfalls and how to avoid them*, Journal of Neurophysiology, 2020, 124:4, 1083-1091, doi: 10.1152/jn.00360.2020

Swapna, M.S., Sreejyothi, S., Raj, V. et al. (2021), *Is SARS CoV-2 a Multifractal?—Unveiling the Fractality and Fractal Structure*, Braz J Phys 51, 731–737,doi:10.1007/s13538-020-00844-w

Wajnsztejn, R. et al (2016), *Higuchi fractal dimension applied to RR intervals in children with attention deficit hyperactivity disorder*, J. Hum. Growth Dev. 26, 147–153. doi: 10.7322/jhgd.119256

Wang, W., Moore, M. A., & Katzgraber, H. G. (2018). *Fractal dimension of interfaces in Edwards-Anderson spin glasses for up to six space dimensions*. Physical Review E, 97(3), 032104.

Wanliss, J. A., and M. A. Reynolds (2003), *Measurement of the stochasticity of low-latitude geomagnetic temporal variations*, Ann. Geophys., 21, 2025.

Wanliss, J. A., Shiokawa, K., & Yumoto, K. (2014). *Latitudinal variation of stochastic properties of the geomagnetic field*. Nonlinear Processes in Geophysics, 21(2), 347-356.

Wanliss, J., Arriaza, R. H., Wanliss, G., and Gordon, S. (2021). *Optimization of the Higuchi Method*. International Journal of Research -GRANTHAALAYAH, 9(11), 202-213. doi:10.29121/granthaalayah.v9.i11.2021.4393

Wątorek, M., Drożdż, S., Kwapień, J., Minati, L., Oświęcimka, P., & Stanuszek, M. (2021), *Multiscale characteristics of the emerging global cryptocurrency market*, Physics Reports, 901, 2021, 1-82, https://doi.org/10.1016/j.physrep.2020.10.005.

Wood, Andrew T.A., and Grace Chan (1994), *Simulation of stationary Gaussian processes in [0, 1]d*, Journal of computational and graphical statistics 3, no. 4 (1994): 409-432.

Wu, F., Zhao, S., Yu, B. et al. (2020), *A new coronavirus associated with human respiratory disease in China*, Nature 579, 265–269, https://doi.org/10.1038/s41586-020-2008-3

Yamada, M. and Ohkitani, K. (1988), *Lyapunov Spectrum of a Model of Two-Dimensional Turbulence*, Phys. Rev. Lett., 60, 983–986, https://doi.org/10.1103/PhysRevLett.60.983, 1988.

Yang X, Xiang Y, Jiang B., *On multi-fault detection of rolling bearing through probabilistic principal component analysis denoising and Higuchi fractal dimension transformation*, Journal of Vibration and Control, February 2021, doi:10.1177/1077546321989527

Yilmaz, A., Unal, G. (2020), *Multiscale Higuchi's fractal dimension method*, *Nonlinear Dyn* 101, 1441–1455 (2020), https://doi.org/10.1007/s11071-020-05826-w

Zuo, Renguang, Cheng, Qiuming, Agterberg, F.P., Xia, Qinglin (2009) *Application of singularity mapping technique to identify local anomalies using stream sediment geochemical data, a case study from Gangdese, Tibet, western China*. Journal of Geochemical Exploration, 101 (3). 225-235 doi:10.1016/j.gexplo.2008.08.003