Efficient Calculation of Fractal Properties via the Higuchi Method

James Wanliss (✉ jawanliss@presby.edu)

Presbyterian College <https://orcid.org/0000-0002-3291-6529>

Grace Wanliss

Presbyterian Eye Ear and Throat Charity Hospital: Greater Baltimore Medical Center

Research Article

Keywords: Higuchi's method of determining fractal dimension (HFD), tuning parameter (kmax)

Posted Date: August 3rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-756319/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Efficient Calculation of Fractal Properties via the Higuchi Method

J. A. Wanliss¹ and Grace E. Wanliss¹

¹Department 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 research tool that, compared to many other methods, gives rapid, efficient, and accurate estimations for the range of possible fractal dimensions. One major difficulty in applying the method is the correct choice of tuning parameter (k_{max}) to compute the most accurate results as a poor choice can generate spurious results. We analyze synthetic fractional Brownian motion to offer a general equation that allows determination, *a priori*, of the best value for the tuning parameter, given a particular length data set.

1. Introduction

Higuchi's method (Higuchi, 1988) is one of the most widely applied time-domain techniques to determine fractal properties of complex non-periodic, nonstationary physical data (Esteller et al., 2001; Salazar-Varas, 2018; Yilmaz & Unal, 2020). Higuchi 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. The fractal dimension calculated via the Higuchi method is called the Higuchi fractal dimension (*HFD*).

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 it has been discovered 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 address 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; Wajnsztein 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?

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 exhibiting fractional Brownian motion (fBm) with a stable fractal dimension, and then apply what we have learned to actual physical data.

The synthetic data we experiment on are theoretically derived with pre-specified fractal dimension, which we recover using the Higuchi algorithm. Mandelbrot and Van Ness (1968) proposed the fBm as a continuous-time random process. A time series displaying fBm can be written in terms of stochastic integrals of time integrations of fractional Gaussian noise:

$$\mathbf{B}_H(t) = \frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^0 [(t-s)^{H-1/2} - (-s)^{H-1/2}] d\mathbf{W}(s) + \int_0^t (t-s)^{H-1/2} d\mathbf{W}(s) \right\}$$

Here $\mathbf{W}$ is a stationary and ergodic random white noise process with zero mean defined on $(-\infty, \infty)$. In the above equation $H \in (0, 1)$ 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 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 Figure 1. The covariance function of the noisy signal can be expressed by:

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

so that $\mathbf{B}_H(0) \equiv \mathbf{0}$ and $\text{var}\{\mathbf{B}_H(t)\} = t^{2H}$. For $H=1/2$ the white noise fBm process reduces to the well-known random walk. The theoretical relationship between the Hurst exponent, H , and the Higuchi fractal dimension, HFD , is $HFD = 2 - H$, with values of HFD between 1 and 2. We will use the fBm process as our data source with a well-defined HFD in order to determine how well the Higuchi method, when appropriately optimized, is able to accurately recover the theoretical value.

We consider four different method generators of processes having long-range dependence to generate synthetic with exact fractal dimension. First, we consider an exact wavelet-based method to synthesize fBm. 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 capture the covariance structure and produce a true realization of fBm.

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

Figure 1. Examples of fBm 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 $H=0.1, 0.3, 0.5, 0.7, 0.9$. 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, and use the fBm process as our data source with a well-defined *HFD* in order to determine how well the Higuchi method is able to accurately recover the theoretical value.

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

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

Here $[\]$ represents the integer part of the enclosed value. The integer $m = 1, 2, \dots, k$ is the start time and k is the time interval, with $k = 1, \dots, k_{max}$; k_{max} is a free tuning parameter. This means that given time interval equal to k , spawns k -sets of new time series. For instance, if $k = 10$ and the time series has length $N = 1000$, the following new time series are derived from the original data:

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

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

.

.

.

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

These curves have lengths defined by:

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

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

$$L(k) = \langle L_m(k) \rangle.$$

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

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

149

150

151

152

153

154

155

156

157

158

159

160

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 Figure 3, of HFD versus a range of $k_{\max}$, where the calculated HFD approaches a local maximum or asymptote (Doyle et al., 2004; Wajnsztein et al., 2016). We will call this a tuning curve. In Figure 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 Figure 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, Figure 4 shows the tuning curves for $HFD=1.9, 1.5, 1.3, 1.1$ computed from the simulated data of Figure 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 $k_{\max} \sim 5000$; 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 Figure 1. The dashed horizontal curves show the theoretical value for the HFD .

The previous equation gives the percentage error to be averaged over all fBm simulations to yield a general result for all fBm simulations considered. As researchers do not generally know *a priori* which method of simulating fBm 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. Figure 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 fBm simulation methods. The curve of least error is shown as a thick grey line.

Each method of simulation yields a different curve of least error. Figure 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 (Figure 5d). Overall, the location of the minimum error curve varies widely depending on the fBm generation algorithm.

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. Figure 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:

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

Here $[]$ represents the integer part of the enclosed function value. Table 1 shows the parameter values for the best-fit. Figure 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 Figure 1 a time series of length $N=20,000$ is used. Figure 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.

207

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

208

209

210

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	$(1.292 \pm 0.045) \times 10^{-5}$
C_1	0.04488 ± 0.0255
A_2	18.82 ± 2.56
B_2	$(6.488 \pm 0.280) \times 10^{-5}$
C_2	1.332 ± 0.220

211

212

213

5 Conclusions

Higuchi's method to compute fractal properties of complex nonlinear 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 methods such as seeking a convergence of the computed *HFD* to a plateau is not in general a valid procedure and that not every case 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 fBm 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 depends on 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.

Acknowledgment

National Science Foundation Award AGS-2053689

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.
- 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.
- 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.

- 254 Esteller, R. et al (2001), *A comparison of waveform fractal dimension algorithms*, IEEE Trans
255 Circ Syst I Fundam Theory Appl, 48:177–183.
- 256 Gomolka R.S. et al (2018), *Higuchi Fractal Dimension of Heart Rate Variability During*
257 *Percutaneous Auricular Vagus Nerve Stimulation in Healthy and Diabetic Subjects*, Front.
258 Physiol. 9:1162. doi: 10.3389/fphys.2018.01162
- 259 Higuchi, T. (1988), *Approach to an irregular time series on the basis of the fractal theory*, Phys
260 D 31:277–283.
- 261 Hosking, Jonathan RM. (1984), *Modeling persistence in hydrological time series using*
262 *fractional differencing*, Water resources research 20, no. 12 (1984): 1898-1908.
- 263 Kijima, M. and Chun Ming Tam (2013), *Fractional Brownian Motions in Financial Models and*
264 *Their Monte Carlo Simulation, Theory and Applications of Monte Carlo Simulations*, Victor
265 (Wai Kin) Chan, IntechOpen, DOI: 10.5772/53568.
- 266 Levinson, N. (1947), *The wiener rms error criterion in filter design and prediction*, Journal of
267 Mathematics and Physics, Vol 25 (1947) 261–278.
- 268 Mandelbrot, B. and Van Ness, J.W. (1968), *Fractional Brownian motions: Fractional noises*
269 *and applications*, SIAM Rev. 10, No. 4, 422–437.
- 270 Paramanathan, P. and Uthayakumar, R. (2008), *Application of fractal theory in analysis of*
271 *human electroencephalographic Signals*, Comput. Biol. Med. 38 (3) (2008) 372–378.
- 272 R. Salazar-Varas, and R.A. Vazquez (2018), *Time-Invariant EEG Classification Based on the*
273 *Fractal Dimension*, In: Castro F., Miranda-Jiménez S., González-Mendoza M. (eds) *Advances in*
274 *Computational Intelligence. MICAI 2017. Lecture Notes in Computer Science*, vol 10633.
275 Springer, Cham. https://doi.org/10.1007/978-3-030-02840-4_26
- 276 Wood, Andrew T.A., and Grace Chan (1994), *Simulation of stationary Gaussian processes in [0,*
277 *1]*, Journal of computational and graphical statistics 3, no. 4 (1994): 409-432.
- 278 Wajnsztein, R. et al (2016), *Higuchi fractal dimension applied to RR intervals in children with*
279 *attention deficit hyperactivity disorder*, J. Hum. Growth Dev. 26, 147–153. doi:
280 10.7322/jhgd.119256
- 281 Yilmaz, A., Unal, G. (2020), *Multiscale Higuchi’s fractal dimension method*, *Nonlinear Dyn*
282 101, 1441–1455 (2020), <https://doi.org/10.1007/s11071-020-05826-w>