1 fractalRegression: An R package for multiscale regression and fractal analyses

2 Aaron D. Likens[1] & Travis J. Wiltshire[2]

3 [1] Department of Biomechanics, University of Nebraska at Omaha

4 [2] Department of Cognitive Science & Artificial Intelligence, Tilburg University

5 Author Note

₁₃   Abstract

₁₄   Time series data from scientific fields as diverse as astrophysics, economics, human

₁₅ movement science, and neuroscience all exhibit fractal properties. That is, these time series

₁₆ often exhibit self-similarity and long-range correlations. This `fractalRegression` package

₁₇ implements a number of univariate and bivariate time series tools appropriate for analyzing

₁₈ noisy data exhibiting these properties. These methods, especially the bivariate tools

₁₉ (Kristoufek, 2015; Likens, Amazeen, West, & Gibbons, 2019) have yet to be implemented

₂₀ in an open source and complete package for the R Statistical Software environment. As

₂₁ both practitioners and developers of these methods, we expect these tools will be of

₂₂ interest to a wide audience of scientists who use R, especially those from fields such as the

₂₃ human movement, cognitive, and other behavioral sciences. The algorithms have been

₂₄ developed in C++ using the popular Rcpp (Eddelbuettel & Francois, 2011) and

₂₅ RcppArmadillo (Eddelbuettel & Sanderson, 2014) packages. The result is a collection of

₂₆ efficient functions that perform well even on long time series (e.g., $\geq 10,000$ data points).

₂₇ In this work, we motivate introduce the package, each of the functions, and give examples

₂₈ of their use as well as issues to consider to correctly use these methods.

₂₉   *Keywords:* long range correlation, fractal, multiscale, dynamics

₃₀   Word count: X

31   fractalRegression: An R package for multiscale regression and fractal analyses

Introduction

33   Fractal analysis, in its many forms, has become an important framework in virtually

34   every area of science, often serving as an indicator of system health (Goldberger et al.,

35   2002), adaptability (Bak, Tang, & Wiesenfeld, 1987), control (Likens, Fine, Amazeen, &

36   Amazeen, 2015), cognitive function (Euler, Wiltshire, Niermeyer, & Butner, 2016), and

37   multi-scale interactions (Kelty-Stephen, 2017). In particular, various methods related to

38   Detrended Fluctuation Analysis (DFA) (Peng et al., 1994) have rose to prominence due to

39   their ease of understanding and broad applicability to stationary and nonstationary time

40   series, alike. The basic DFA algorithm has been implemented in numerous packages and

41   software programs. However, advanced methods such as Multifractal Detrended

42   Fluctuation Analysis (MFDFA) (Jan W. Kantelhardt et al., 2002), Detrended Cross

43   Correlation (DCCA) (Podobnik & Stanley, 2008; Zebende, 2011), and, in particular, fractal

44   regression techniques such as Multiscale Regression Analysis (MRA) (Kristoufek, 2015;

45   Likens, Amazeen, West, & Gibbons, 2019) have not yet been implemented in a

46   comprehensive CRAN Package for the R Statistical Software Environment. Thus, there is a

47   clear need for a package that incorporates this functionality in order to advance theoretical

48   research focused on understanding the time varying properties of natural phenomena and

49   applied research that uses those insights in important areas such as healthcare (Cavanaugh,

50   Kelty-Stephen, & Stergiou, 2017) and education (Snow, Likens, Allen, & McNamara, 2016).

Package Overview

52   Some foundational efforts in fractal analyses, which partially overlap with the

53   functionality of this package, have been implemented elsewhere. For example, a number of

54   univariate fractal and multifractal analyses have been implemented in the 'fracLab' library

55   for MATLAB (Legrand & Véhel, 2003) and other toolboxes that are mainly targeted at

⁵⁶ multifractal analysis (E. A. F. E. A. F. I. Ihlen, 2012; E. A. F. Ihlen & Vereijken, 2010). In

⁵⁷ terms of open access packages, there are other packages that implement some, but not all

⁵⁸ of the same functions such as the `fathon` package (Bianchi, 2020) that has been

⁵⁹ implemented in Python as well as the R packages: `fractal` [@], `nonlinearTseries`

⁶⁰ (Garcia, 2020), and `MFDFA` (Laib, Golay, Telesca, & Kanevski, 2018). However, none of the

⁶¹ above packages incorporate univariate monofractal and multifractal DFA with bivariate

⁶² DCCA and MRA nor do they run on a C++ architecture. Our `fractalRegression`

⁶³ package is unique in this combination of analyses and efficiency. For instance, we are not

⁶⁴ aware of any other packages that feature MRA and Multiscale Lagged Regression (MLRA).

Methodological Details and Examples

⁶⁶ In order to demonstrate the methods within the 'fractalRegression' package, we

⁶⁷ group this into univariate (DFA, MFDFA) and bivariate methods (DCCA, MRA, MRLA).

⁶⁸ For each method, we 1) highlight the key question(s) that can be answered with that

⁶⁹ method, 2) briefly describe the algorithm with sources for additional details, 3) describe

⁷⁰ some key consideration for appropriately applying the algorithm, and demonstrate the use

⁷¹ of the functions on a 4) simulated and 5) empirical application of the function.

⁷² Table 1. Overview of package functions, objectives, and output.

Univariate Methods

⁷⁴ **Detrended Fluctuation Analysis.** The key questions that can be answered by

⁷⁵ Detrended Fluctuation Analysis (DFA) (Peng et al., 1994) is INSERT KEY QUESTION

⁷⁶ HERE. DFA has been described extensively elsewhere (Jan W. Kantelhardt,

⁷⁷ Koscielny-Bunde, Rego, Havlin, & Bunde, 2001) and visualized nicely (Kelty-Stephen,

⁷⁸ Stirling, & Lipsitz, 2016); so we provide a brief summary here. DFA entails splitting a time

⁷⁹ series into several small bins (e.g., 16). In each bin, the least squares regression is fit and

₈₀ subtracted within each window. Residuals are squared and averaged within each window.

₈₁ Then, the square root is taken of the average squared residual across all windows of a given

₈₂ size. This process repeats for larger window sizes, growing by, say a power of 2, up to N/4,

₈₃ where N is the length of the series. In a final step, the logarithm of those scaled root mean

₈₄ squared residuals (i.e., fluctuations) is regressed on the logarithm of window sizes. The

₈₅ slope of this line is termed α and it provides a measure of the long range correlation.

₈₆ Conventional interpretation of α is:

₈₇   • $\alpha < 0.5$ = anti-correlated

₈₈   • $\alpha = 0.5$ = uncorrelated, white noise

₈₉   • $\alpha > 0.5$ = temporally correlated

₉₀   • $\alpha = 1$ = 1/f-noise, pink noise

₉₁   • $\alpha > 1$ = non-stationary and unbounded

₉₂   • $\alpha = 1.5$ = fractional brownian motion

₉₃   ***DFA Examples.***

₉₄   INCLUDE EXAMPLES HERE

₉₅   ***DFA Considerations.***

₉₆   We recommend a few points of consideration here in using this function. One is to be

₉₇ sure to verify there are not cross-over points in the logScale-logFluctuation plots (Peng et

₉₈ al., 1995; Perakakis et al ., 2009). Cross-over points (or a visible change in the slope as a

₉₉ function of of scale) indicate that a mono-fractal characterization does not sufficiently

₁₀₀ characterize the data. If cross-over points are evident, we recommend proceeding to using

₁₀₁ the 'mfdfa()' to estimate the multi-fractal fluctuation dynamics across scales.

₁₀₂   While it is common to use only linear detrending with DFA, it is important to

₁₀₃ inspect the trends in the data to determine if it would be more appropriate to use a higher

104 order polynomial for detrending, and/or compare the DFA output for different polynomial

105 orders (see Kantelhardt et al., 2001).

106   General recommendations for choosing the min and max scale are an sc_min = 10

107 and sc_max = (N/4), where N is the number of observations. See Eke et al. (2002) and

108 Gulich and Zunino (2014) for additional considerations. - Key Question and method

109 description. - Simulated data: fGN - Empirical data: EPICLE Movement Data? - MFDFA

110 - Key Question(s) - Simulated data: mfbrownian motion from Ihlen matlab (Aaron might

111 have R port) see mbm_mgn for R from aaron - Empirical data: EPICLE Movement Data?

112   **Multifractal Detrended Fluctuation Analysis.**   Multifractal Detrended

113 Fluctuation Analysis (MFDFA; Kantelhardt et al., 2002) is an extension of DFA by

114 generalizing the fluctuation function to a range of exponents of the qth order.

Bivariate Methods

116 - DCCA

117   – Key Question: What the timescale(s) of coordination?

118   – Simulated data: MC-ARFIMA

119   – Empirical data: EPICLE Movement Data?

120 - MRA

121   – Key Question

122   – Simulated data:

123   – Empirical data: FNIRS from Aaron?

124 - MLRA

125   – Key Question

126   – Simulated data: Equation from Aaron from grant on MLRA

127   – Empirical data: FNIRS from Aaron?

Surrogate Methods (and 'full' data analysis)

Methods are ranked in terms of increasing levels of rigor.

- Randomization - Estimates should be different.
- IAAFT - Estimates should be different.
- Model based surrogate (Simulated exponents) - See Likens 2019 paper with model of postural sway/control, taking an educated guess about the data generating process underlying the time series. Estimates should not be different. See Roume et al 2018 windowed detrended CCA
- Can we incorporate lags into MC-ARFIMA?

General Discussion

- General value of methods and the types of questions
- Practical consideration of univariate methods
- Practical consideration of bivariate methods
- Unique contribution of the methods

Acknowledgements

References

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, *59*(4), 381–384. https://doi.org/10.1103/PhysRevLett.59.381

Bianchi, S. (2020). Fathon: A python package for a fast computation of detrendend fluctuation analysis and related algorithms. *Journal of Open Source Software*, *5*(45), 1828.

Cavanaugh, J. T., Kelty-Stephen, D. G., & Stergiou, N. (2017). Multifractality, Interactivity, and the Adaptive Capacity of the Human Movement System: A Perspective for Advancing the Conceptual Basis of Neurologic Physical Therapy. Retrieved from https: //www.ingentaconnect.com/content/wk/npt/2017/00000041/00000004/art00007

Eddelbuettel, D., & Francois, R. (2011). Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software*, *40*(1), 1–18. https://doi.org/10.18637/jss.v040.i08

Eddelbuettel, D., & Sanderson, C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics & Data Analysis*, *71*, 1054–1063. https://doi.org/10.1016/j.csda.2013.02.005

Euler, M. J., Wiltshire, T. J., Niermeyer, M. A., & Butner, J. E. (2016). Working memory performance inversely predicts spontaneous delta and theta-band scaling relations. *Brain Research*, *1637*, 22–33. https://doi.org/10.1016/j.brainres.2016.02.008

Garcia, C. A. (2020). nonlinearTseries: Nonlinear Time Series Analysis. Retrieved from https://CRAN.R-project.org/package=nonlinearTseries

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. C., Peng, C.-K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences*, *99*(suppl 1), 2466–2472. https://doi.org/10.1073/pnas.012579499

Ihlen, E. A. F. E. A. F. I. (2012). Introduction to Multifractal Detrended Fluctuation Analysis in Matlab. *Frontiers in Physiology*, *3*. https://doi.org/10.3389/fphys.2012.00141

Ihlen, E. A. F., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond 1/? fluctuation. *Journal of Experimental Psychology: General*, *139*(3), 436–463. https://doi.org/10.1037/a0019098

Kantelhardt, Jan W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., & Bunde, A. (2001). Detecting long-range correlations with detrended fluctuation analysis. *Physica A: Statistical Mechanics and Its Applications*, *295*(3), 441–454. https://doi.org/10.1016/S0378-4371(01)00144-3

Kantelhardt, Jan W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., & Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and Its Applications*, *316*(1), 87–114. https://doi.org/10.1016/S0378-4371(02)01383-3

Kelty-Stephen, D. G. (2017). Threading a multifractal social psychology through within-organism coordination to within-group interactions: A tale of coordination in three acts. *Chaos, Solitons & Fractals*, *104*, 363–370. https://doi.org/10.1016/j.chaos.2017.08.037

Kelty-Stephen, D. G., Stirling, L. A., & Lipsitz, L. A. (2016). Multifractal temporal correlations in circle-tracing behaviors are associated with the executive function of rule-switching assessed by the Trail Making Test. *Psychological Assessment*, *28*(2), 171–180. https://doi.org/10.1037/pas0000177

Kristoufek, L. (2015). Detrended fluctuation analysis as a regression framework: Estimating dependence at different scales. *Physical Review E*, *91*(2), 022802. https://doi.org/10.1103/PhysRevE.91.022802

Laib, M., Golay, J., Telesca, L., & Kanevski, M. (2018). Multifractal analysis of the time series of daily means of wind speed in complex regions. *Chaos, Solitons & Fractals*, *109*, 118–127. https://doi.org/10.1016/j.chaos.2018.02.024

Legrand, P., & Véhel, J. L. (2003). Signal and image processing with FracLab. *Thinking in Patterns: Fractals and Related Phenomena in Nature*, 321322.

Likens, A. D., Amazeen, P. G., West, S. G., & Gibbons, C. T. (2019). Statistical properties of Multiscale Regression Analysis: Simulation and application to human postural control. *Physica A: Statistical Mechanics and Its Applications*, *532*, 121580. https://doi.org/10.1016/j.physa.2019.121580

Likens, A. D., Fine, J. M., Amazeen, E. L., & Amazeen, P. G. (2015). Experimental control of scaling behavior: What is not fractal? *Experimental Brain Research*, *233*(10), 2813–2821. https://doi.org/10.1007/s00221-015-4351-4

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, *49*(2), 1685–1689. https://doi.org/10.1103/PhysRevE.49.1685

Podobnik, B., & Stanley, H. E. (2008). Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. *Physical Review Letters*, *100*(8), 084102. https://doi.org/10.1103/PhysRevLett.100.084102

Snow, E. L., Likens, A. D., Allen, L. K., & McNamara, D. S. (2016). Taking Control: Stealth Assessment of Deterministic Behaviors Within a Game-Based System. *International Journal of Artificial Intelligence in Education*, *26*(4), 1011–1032. https://doi.org/10.1007/s40593-015-0085-5

[220]   Zebende, G. F. (2011). DCCA cross-correlation coefficient: Quantifying level of

[221]   cross-correlation. *Physica A: Statistical Mechanics and Its Applications*, *390*(4),

[222]   614–618. https://doi.org/10.1016/j.physa.2010.10.022

Function	Objective	Output