Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-315-RC3, 2017

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Evaluation of statistical methods for quantifying fractal scaling in water quality time series with irregular sampling" *by* Qian Zhang et al.

Anonymous Referee #1

Received and published: 26 September 2017

Thank you for your time and diligence in preparing your responses. Albeit the well articulated arguments, which deserved my full consideration, unfortunately they do not yet solve some of the fundamental scientific concerns that the paper raises.

Before discussing such concerns in detail, I would like to note that I have great consideration for the technical merits of the work undergone in the manuscript, and for the problems tackled by the authors in the present quest. Fractal scaling in irregular time series is a worthy venture with important relevance in many technical fields, and it has well-proven scientific merits in various disciplines as the reference list in point 5 of this report documents.

This work shall also become more relevant on a scientific standpoint once it is complemented by a sound scientific basis advancing the understanding of hydrologic functioning, namely on the fundamental principles that explain the fundamental nature of the detected signatures and the implemented approaches beyond the already presented statistical and geometric considerations.

In detail, this report discusses the following key concerns:

1) On the hydrological insights or lack thereof:

With all due respect, and albeit the arguments put forward by the authors, the paper is still fundamentally devoid of any hydrological insights. The well-known fractal scaling approaches discussed in the manuscript have purely descriptive merits and are inherently grounded on statistical geometry. The underlying physical understanding is thus entirely missing.

The ability of fractal scaling to elicit trends is again a purely descriptive merit, adding nothing to the fundamental understanding of hydrological functioning. To reach that understanding, one must fundamentally address the questions: Why? What are the process reasons behind an observed statistical and/or geometric signature? What is the physical meaning of a fractional scaling exponent? Of a fractional law? And what do these inform about the "hows" and "whys" in hydrology?

2) On the long history of fractal theories and scaling - centuries before hydrology:

The science underlying fractal analysis, geometry and scaling has a long scientific history before the empirical work of Hurst. While he contributed to the field from its empirical side, the fundamental contributions date back to the 17th century calculus. You may have learnt at the university about the many contributions of Leibniz: among many contributions, he introduced the first solid concepts of fractional exponents and laid the foundations of fractional calculus (aside from making fundamental contributions to mainstream calculus as well). We owe Leibniz - not Hurst - the fundamentals on fractal scaling.

Soon these 17th century concepts were linked to physical laws and a new branch of analytical and statistical mechanics was born. Among the many illustrious users of fractional mechanics was Einstein, which provided systematic rigorous physical grounds to what predecessors had only described with geometry and statistics (the physics behind the random walks in Brownian motion). This is actually what landed him the Nobel prize in Physics (instead of his more famous contributions on relativity and quantum mechanics).

3) On the descriptive nature of the hydrological work - science still elusive

In its current state, hydrology is still an applied discipline, and there is nothing wrong about that in principle. The service that the discipline plays to society is undeniable and is ultimately why all of us gather here in this forum to advance the field for the benefit of all. While some important laws have been formulated, the discipline still lacks a consistent fundamental theory, and all theoretical formulations are imported from elsewhere (e.g. Darcy's law is just an import of the Ohm's law to hydrology).

Living off empiricisms and imports without any fundamental theoretical explanation may be good in statistical and engineering hydrology but will not lead us anywhere in real science. In fact, the inherent empiricism of many hydrological literature is understandable in an engineering setting where all that matters is to get some number that ensures portability of measured features from one scale to another for design and decision support. However, that brings no understanding about the real functioning of the hydrologic system.

The work presented in this manuscript has a good place in an applied statistics or engineering setting, for use by practitioners in engineering hydrology that have no time or scientific background to study, understand or build on the real scientific literature.

However, as a candidate for scientific paper, there is, at this point, no science to be

learnt in the present study. Venturing into approaches without complementing them with the supporting physical principles brings little benefit to hydrological understanding, since their descriptive aptitude is not accompanied by any physically related insight.

4) On the (ir)relevance of comparing various methods that in essence are more of the same

The different approaches compared in the study belong to the same class of methodological equivalence of naïve statistical geometry, and do not necessarily represent the best in the field. Therefore, comparing the analysed approaches is of little methodological relevance since it is not taking a comprehensive and useful up-to-date selection.

In other words, there little relevance in performing a comparison among different methods that are no longer up to date, and even less so when they belong to the same methodological class of equivalence. An illustrative view to get the idea: publishing a study comparing the various models of chariot transport will add nothing relevant in the age of the automobile, unless we are interested in the history of science and technology. At best, we should compare methods that are fundamentally different from each other rather than variants of the very same concept.

5) On the "onus of proof"

It is the duty of the authors to unequivocally demonstrate a significant degree of scientific innovation and novel hydrological insights. So far, from the manuscript and responses, that unequivocal proof is still missing.

While it is not the referee's duty to demonstrate the vacuity of the study - but rather the authors' duty to demonstrate its substance - I will gladly provide elements to help the authors find a wealth of literature on studies that cover the same problems and solutions discussed in the present study.

A careful and thorough literature review can aptly demonstrated that existing studies actually perform in due terms what the present article claims was missing in the literature, thus effectively deflating the innovation claims.

For instance, the quantification of fractal scaling in irregularly sampled time series is a well studied problem and the experts are well aware about which methods perform better under which circumstances, therefore there is no gap in that area that would support any claim of innovation and relevance.

The references are found at the end of this report, at point 5.

4) On the ways forward to improve the manuscript:

The application of more comprehensive and effective fractal scaling techniques to your particular hydrological problems must be accompanied with real insights on hydrological functioning, rather than simply statistical description of results (e.g. a "trend" is hardly an insight, for it has descriptive but no explanatory power as noted above. Moreover, there is far more to fractal scaling than what is argued in the paper - again, the physics that would help elicit the "whys" in hydrologic functioning are missing).

In sum, my recommendation is to: a) bring real hydrological understanding rather than vague generic considerations and descriptive statements without quantitative physical reasoning to substantiate any explanatory claim; b) interpret the fractal metrics from a physical point of view that will shed light on the science behind the metric; c) provide if and how hydrologic series differ from others where fractal scaling over irregular series has already been extensively applied to justify separate publication.

At this stage, I leave the authors with some of the relevant literature references (apologies for non-uniform formatting):

5) REFERENCES:

A) Fractal approaches in irregular time series analysis:

Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal the-

ory. Physica D: Nonlinear Phenomena. 31. 277-283. 10.1016/0167-2789(88)90081-4.

P. Paramanathan, R. Uthayakumar (2007): ICCIMA '07 Proc. Intl. Conf. on Computational Intelligence and Multimedia Applications (ICCIMA 2007), 02, 323-327

Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos. 1995;5(1):82-7.

Stanley HE, Amaral LA, Goldberger AL, Havlin S, Ivanov PCh, Peng CK. Statistical physics and physiology: monofractal and multifractal approaches. Physica A. 1999 Aug 1;270(1-2):309-24.

Viswanathan GM1, Peng CK, Stanley HE, Goldberger AL. (1997): Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. 1997 Jan;55(1):845-9. Deviations from uniform power law scaling in nonstationary time series.

L. Telesca, V. Cuomo, V. Lapenna, M. Macchiato, C. Serio (1999): Detecting Stochastic Behaviour and Scaling Laws in Time Series of Geomagnetic Daily Means Pure appl. geophys. 156 (1999) 487–501, doi.org/10.1007/s000240050309.

Sergio Cerutti, Carlo Marchesi (2011): Advanced Methods of Biomedical Signal Processing. John Wiley & Sons.

M. Malik (1998): Clinical Guide to Cardiac Autonomic Tests. Springer Science & Business Media.

F Cervantes-De la Torre et al (2013): Fractal dimension algorithms and their application to time series associated with natural phenomena. J. Phys.: Conf. Ser. 475 012002.

B1) Analysis of irregularly sampled time series:

Roberto Vio et al (2000): PASP 112 74. https://doi.org/10.1086/316495

Piet M.T.Broersen (2005): Time series analysis for irregularly sampled data. IFAC Proc. Vol., 38, 1, 2005, 154-159 (http://www.sciencedirect.com/science/article/pii/S1474667016360384)

J.Durbin, S.J.Koopman (2012): Time Series Analysis by State Space Methods, 2nd edition, 2012. [See e.g. sec. 4.10, on statistical modeling under missing observations]

Adolf Mathias, Florian Grond, Ramon Guardans, Detlef Seese, Miguel Canela, Hans H. Diebner (2004): Algorithms for Spectral Analysis of Irregularly Sampled Time Series. Journal of Statistical Software. 10.18637/jss.v011.i02.

P Stoica, N Sandgren (2006): Spectral analysis of irregularly-sampled data: Paralleling the regularly-sampled data approaches. Digital Signal Processing, 2006, Elsevier.

Dilmaghani, Shabnam, et al. (2007): Harmonic analysis of environmental time series with missing data or irregular sample spacing. Environmental science & technology 41.20 (2007): 7030-7038.

Babu, Prabhu. Spectral analysis of nonuniformly sampled data and applications. Diss. Uppsala universitet, 2012.

B2) Analysis of unequally spaced data

JONES, R. H. (1985) Time series analysis with unequally spaced data. In Handbook of Statistics, Volume 5: Time Series in the Time Domain, ed. E. J. Hannan, P. R. Krishnaiah and M. M. Rao. North-Holland, Amsterdam.

Jones, R. (1986). Time Series Regression with Unequally Spaced Data. Journal of Applied Probability, 23, 89-98. doi:10.2307/3214345

C Hackman and T E Parker (1996): Noise analysis of unevenly spaced time series data. Metrologia 33 457

Eckner, Andreas. (2017). A Framework for the Analysis of Unevenly Spaced Time Series Data.

Doan TK, Haslett J, Parnell AC (2015): Joint inference of misaligned irregular time series with application to Greenland ice core data

Nieto-Barajas LE, Sinha T (2014): Bayesian interpolation of unequally spaced time series

Piet M T Broersen (2008): Time series models for spectral analysis of irregular data far beyond the mean data rate. Meas. Sci. Technol. 19 015103. https://doi.org/10.1088/0957-0233/19/1/015103

P.M.T. Broersen ; R. Bos (2006): Estimating time-series models from irregularly spaced data. IEEE Transactions on Instrumentation and Measurement

Piet M. T. Broersen, "Five Separate Bias Contributions in Time Series Models for Equidistantly Resampled Irregular Data", Instrumentation and Measurement IEEE Transactions on, vol. 58, pp. 1370-1379, 2009, ISSN 0018-9456.

Piet M.T. Broersen, "Bias Contributions in Time Series Models for Resampled Irregular Data", Instrumentation and Measurement Technology Conference Proceedings 2008. IMTC 2008. IEEE, pp. 882-889, 2008, ISSN 1091-5281.

Piet M. T. Broersen, "The Removal of Spurious Spectral Peaks From Autoregressive Models for Irregularly Sampled Data", Instrumentation and Measurement IEEE Transactions on, vol. 59, pp. 205-214, 2010, ISSN 0018-9456.

Piet M. T. Broersen, "Autoregressive Order Selection for Irregularly Sampled Data", Instrumentation and Measurement Technology Conference 2006. IMTC 2006. Proceedings of the IEEE, pp. 1004-1009, 2006, ISSN 1091-5281.

Piet M.T. Broersen, "Spectral Estimation from Irregularly Sampled Data for Frequencies Far Above the Mean Data Rate", Instrumentation and Measurement Technology Conference Proceedings 2007. IMTC 2007. IEEE, pp. 1-6, 2007, ISSN 1091-5281.

R. H. Jones, "Fitting multivariate models to unequally spaced data" in Time Series

Analysis of Irregularly Spaced Data, New York:Springer-Verlag, pp. 158-188, 1983.

E. K. Larsson, T. Söderström, "Identification of continuous-time AR processes from unevenly sampled data", Automatica, vol. 38, no. 4, pp. 709-718, Apr. 2002.

E. K. Larsson, E. G. Larsson, "The CRB for parameter estimation in irregularly sampled continuous-time ARMA systems", IEEE Signal Process. Lett., vol. 11, no. 2, pp. 197-200, Feb. 2002.

E. Lahalle, G. Fleury, A. Rivoira, "Continuous ARMA spectral estimation from irregularly sampled observations", Proc. IEEE/IMTC Conf., pp. 923-927, 2004.

S. de Waele, P. M. T. Broersen, "Error measures for resampled irregular data", IEEE Trans. Instrum. Meas., vol. 49, no. 2, pp. 216-222, Apr. 2000.

R. Bos, S. de Waele, P. M. T. Broersen, "Autoregressive spectral estimation by application of the Burg algorithm to irregularly sampled data", IEEE Trans. Instrum. Meas., vol. 51, no. 6, pp. 1289-1294, Dec. 2002.

R. H. Jones, "Maximum likelihood fitting of ARMA models to time series with missing observations", Technometrics, vol. 22, no. 3, pp. 389-395, 1980.

P. M. T. Broersen, S. de Waele, R. Bos, "Autoregressive spectral analysis when observations are missing", Automatica, vol. 40, no. 9, pp. 1495-1504, 2004.

P. M. T. Broersen, S. de Waele, R. Bos, "Application of autoregressive spectral analysis to missing data problems", IEEE Trans. Instrum. Meas., vol. 53, no. 4, pp. 981-986, Aug. 2004.

W. K. Harteveld, R. F. Mudde, H. E. A. van den Akker, "Estimation of turbulence power spectra for bubbly flows from laser Doppler anemometry signals", Chem. Eng. Sci., vol. 60, pp. 6160-6168, 2005.

P. M. T. Broersen; R. Bos (2006): Time-series analysis if data are randomly missing. IEEE Transactions on Instrumentation and Measurement, 55, 1, 79-84.

P. M. T. Broersen; S. de Waele; R. Bos (2003): Estimation of autoregressive spectra with randomly missing data Proc. 20th IEEE Instrumentation Technology Conf. (Cat. No.03CH37412), 2, 1154-1159.

R. Bos; S. de Waele; P. M. T. Broersen (2001): AR spectral estimation by application of the Burg algorithm to irregularly sampled data IMTC 2001. Proc. 18th IEEE Instrumentation and Measurement Technology Conf. Rediscovering Measurement in the Age of Informatics (Cat. No.01CH 37188), 2, 1208-1213.

C) Robust analysis of sparsely distributed data

Trendafilov, Nickolay; Kleinsteuber, Martin and Zou, Hui (2014). Sparse matrices in data analysis. Computational Statistics, 29(3-4) pp. 403–405.

Pires, C.A.L.; Perdigão, R.A.P. (2013): Minimum Mutual Information and Non-Gaussianity through the Maximum Entropy Method: Estimation from Finite Samples. Entropy 2013, 15, 721-752.

Peter Melchior, Andy D. Goulding (2016): Filling the gaps: Gaussian mixture models from noisy, truncated or incomplete samples. arXiv:1611.05806v1

Testa T., et al. (2016): Sparse representation of signals: from astrophysics to real-time data analysis for fusion plasmas and system optimization analysis for ITER and TCV. Plasma Physics and Controlled Fusion, Volume 58, Number 12. [review paper]

D) Astrophysical studies

Feigelson, E. D. & Babu, G. J. (1992): Improving the Statistical Methodology of Astronomical Data Analysis. Astronomical Data Analysis Software and Systems I, A.S.P. Conference Series, Vol. 25, 1992, Diana M. Worrall, Chris Biemesderfer, and Jeannette Barnes, eds., p. 237.

Bourguignon, S. and Carfantan, H. (2015): Line Spectra Estimation for Irregularly Sampled Signals in Astrophysics, in Regularization and Bayesian Methods for Inverse Problems in Signal and Image Processing (eds J.-F. Giovannelli and J. Idier), John Wiley & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/9781118827253.ch6

Jean-Luc Starck (2016): Sparsity and inverse problems in astrophysics. J. Phys.: Conf. Ser. 699 012010

E) Geo/environmental studies

K. Rehfeld and J. Kurths (2014): Similarity estimators for irregular and age-uncertain time series. Clim. Past, 10, 107–122, 2014. doi:10.5194/cp-10-107-2014

Björg Ólafsdóttir, Kristín; Schulz, Michael; Mudelsee, Manfred (2016): REDFIT-X: Cross-spectral analysis of unevenly spaced paleoclimate time series Computers and Geosciences, Volume 91, p. 11-18. 10.1016/j.cageo.2016.03.001

David J. Thomson: Time–Series Analysis of Paleoclimate Data. Encyclopedia of Paleoclimatology and Ancient Environments Part of the series Encyclopedia of Earth Sciences Series pp 949-959

I hope this helps.

Best wishes.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-315, 2017.