On Bitcoin markets (in)efficiency and its evolution

Ladislav Kristoufek

PII: S0378-4371(18)30241-3
DOI: https://doi.org/10.1016/j.physa.2018.02.161
Reference: PHYSA 19281

To appear in: *Physica A*

Please cite this article as: L. Kristoufek, On Bitcoin markets (in)efficiency and its evolution, *Physica A* (2018), https://doi.org/10.1016/j.physa.2018.02.161

Highlights:

- Efficiency of the USD and CNY Bitcoin markets is studied.
- Efficiency index based on long-range dependence, fractal dimension, and entropy is used.
- We find strong evidence of both Bitcoin markets remaining mostly inefficient between 2010 and 2017.
- Markets are efficient only during cooling-downs after bubble-like price surges.

On Bitcoin markets (in)efficiency and its evolution

Ladislav Kristoufek[a]

[a]*Institute of Economic Studies, Faculty of Social Sciences, Charles University Prague, Opletalova 26, 110 00, Prague, Czech Republic, EU*

Abstract

We study efficiency of two Bitcoin markets (with respect to the US dollar and Chinese yuan) and its evolution in time. As inefficiency can manifest through various channels, we utilize the Efficiency Index of Kristoufek & Vosvrda (2013, *Physica A 392, pp. 184-193*) which can cover different types of (in)efficiency measures. We find strong evidence of both Bitcoin markets remaining mostly inefficient between 2010 and 2017 with exceptions of several periods directly connected to cooling down after the bubble-like price surges.

Keywords: Bitcoin, efficient market hypothesis, efficiency index, long-range dependence, fractal dimension, entropy

1. Introduction

Bitcoin [1] as the most popular cryptocurrency with the highest historical capitalization of around \$175 billion (as of the end of November 2017)[1] has gone a long path from its rather controversial beginnings [2] to its current status. Even though its legal and institutional state has not been properly solved yet, its popularity and completely unprecedented price growth[2] have attracted attention of big institutional players as well as small (amateur) investors looking for "easy profits".

Early research studies of Bitcoin focused primarily on its security, legal, and technical issues [3–6]. Since 2013, studies focusing on the financial aspects of the cryptocurrency have started to emerge as well [7–11]. As Bitcoin started becoming more known to the financial community, the topics of interest have moved closer to the traditional economics and financial issues. Studying whether the Bitcoin markets can be considered as standard financial markets with some relationship to the efficient market hypothesis (EMH) [12, 13] has been one of these important questions. Urquhart [14] studies the Bitcoin market from its beginnings in 2010 to mid-2016 and suggests that the market is inefficient but it is moving closer towards efficiency in time. Nadarajah & Chu [15] dispute these results and conclude that the market is in fact efficient. Bariviera *et al.* [16, 17] study the dynamics of

Email address: `ladislav.kristoufek@fsv.cuni.cz` (Ladislav Kristoufek)

[1]According to `https://coinmarketcap.com`.

[2]Even if we do not consider the very beginnings of Bitcoin, its price rocketed from around \$13 in 2013 to around \$10,000 as of end of November 2017 accounting to a 76,700% growth over mere 5 years.

long-range dependence properties of the Bitcoin price and find a trend towards efficiency. Similar results are found by Alvarez-Ramirez *et al.* [18]. These results serve as a basic motivation for the current study.

We focus on the efficiency of two Bitcoin markets (with respect to the US dollar and Chinese yuan as the most prominent markets) and its evolution in time. As inefficiency can manifest through various channels, we utilize the Efficiency Index [21] which can cover different types of (in)efficiency measures. In the following section, we introduce the index together with its components – specifically, long-range dependence, fractal dimension, and entropy – and we discuss its statistical properties proposing a novel way of approaching the issue. Following section describes the dataset and last section provides results with some further discussion. We find strong evidence that both Bitcoin markets have remained mostly inefficient between 2010 and 2017 with exceptions of several periods directly connected to cooling down after the bubble-like price surges.

2. Methods

When constructing tests for the market efficiency, it is essential to consider statistical and dynamic implications of the theory. The two historical building blocks of the current EMH lead to two different processes of the efficient market. Based on Fama [19], the efficient market follows a random whereas Samuelson [20] argues that a martingale is an appropriate description of such market. Of the two, we use the less restrictive one, i.e. the martingale specification of the efficient market (log-)price process, which implies that the (log-)returns series are serially uncorrelated (there is no statistically significant autocorrelation structure). The random walk specification would require the returns to have the Gaussian distribution as well which we consider unnecessarily restrictive. Such a simple description of the efficient market in the EMH logic gives direct implications of expected values of some useful statistics and parameters of dynamic series that can be used to test and measure efficiency of a capital market. We use these to construct the Efficiency Index which takes into consideration more measures of market efficiency.

2.1. Capital market efficiency measure

Kristoufek & Vosvrda [21–24] define the Efficiency Index (EI) as

$$EI = \sqrt{\sum_{i=1}^{n} \left(\frac{\widehat{M_i} - M_i^*}{R_i} \right)^2}, \tag{1}$$

where M_i is the ith measure of efficiency, $\widehat{M_i}$ is an estimate of the ith measure, M_i^* is an expected value of the ith measure for the efficient market and R_i is a range of the ith measure. The index is then a distance from the efficient market situation and it can be based on a combination of efficiency measures as long as these measures are bounded. This turns out to be a rather restrictive condition, yet still there are various choices of methods that can be utilized for the index construction. We stick to the original articles [21–24] and use measures of long-range dependence, fractal dimension and entropy to construct the index. Long-range dependence parameter Hurst exponent H has an expected value of 0.5 for the efficient market ($M_H^* = 0.5$), fractal dimension D has an expected value of 1.5 ($M_D^* = 1.5$), and the approximate entropy has an expected value of 1 ($M_{AE}^* = 1$). All three parameters have a restricted range which, however, is not the same for all three. As the approximate entropy range is higher, we need to rescale its effect and we have $R_{AE} = 2$ and $R_D = R_H = 1$.

2.2. Long-range dependence and its estimators

Long-range dependent (long-term correlated) processes are defined both in time and frequency domain. In the time domain, their autocorrelation function $\rho(k)$ with time lag k decays as $\rho(k) \propto k^{2H-2}$ for $k \to +\infty$. In the frequency domain, their spectrum $f(\lambda)$ with frequency λ scales as $f(\lambda) \propto \lambda^{1-2H}$ for $\lambda \to 0+$ [25–27]. Hurst exponent H is the characteristic parameter of the long-term correlated processes. The anti-persistent processes with $H < 0.5$ switch their sign more frequently than uncorrelated processes. Processes with no long-range dependence have $H = 0.5$, and the persistent processes have $H > 0.5$. Stationary processes have $H < 1$. As the efficient market has no non-zero correlation structure, it has also no long-term memory. Therefore, the expected value of the Hurst exponent for the efficient market is $H = 0.5$. There are many estimators of Hurst exponent with different properties and sensitivities to different features of examined series [26, 28–33]. We utilize two estimators that are suitable for short time series with possible short-term memory – the local Whittle estimator and the GPH estimator [25, 27–29, 34].

2.3. Fractal dimension

Contrary to long-range dependence, which is a global description of the correlation structure of the series, fractal dimension D is usually interpreted as a measure of local correlation structure of the series as it describes roughness of the series [21]. For univariate series, fractal dimension ranges between $1 < D \leq 2$ and the central point $D = 1.5$ covers the serially uncorrelated processes. This gives the value of D for the efficient market Low levels of fractal dimension suggest lower roughness and thus local positive auto-correlation dynamics. Then high fractal dimension is attached to rougher series which are locally negatively auto-correlated. As parts of the Efficiency Index, we use two estimators that have desirable statistical properties for short time series – the Hall-Wood and Genton estimators [35, 36].

2.4. Approximate entropy

In the time series analysis, entropy is a complexity measure. Series with high entropy have little or no information in the system connected to high uncertainty. Low entropy series can be seen as deterministic, i.e. predictable [37]. From the perspective of an efficient market, maximum entropy yields efficiency as the high entropy series are serially uncorrelated. With the lower entropy, markets become less efficient. In the Efficiency Index, we use the approximate entropy which is bounded (contrary to other versions of entropy) [38].

2.5. Statistical inference and moving window estimation

The analysis is based on estimating the Efficiency Index with respect to Eq. 1. However, statistical properties or liming distribution under the null hypothesis have not been developed mainly due to the fact that the index values are dependent on the specific efficiency measures that are incorporated. The limiting distribution should thus be studied for each specific case. We approach this issue by constructing its distribution under the null hypothesis of an efficient market using bootstrapping. Specifically, we follow these steps:

1. Estimate EI for the original series.
2. Use the original returns series to construct bootstrapped (with replacement) series with the same number of observations.
3. Estimate EI for the bootstrapped series.
4. Repeat Steps 2 and 3 many times (333 in our specific case).
5. Find critical values of the EI under the null hypothesis of an efficient market as quantiles of the bootstrapped estimates for selected significance level (in our case, we work with 90% level, i.e. we get the 5th and 95th quantiles of the distribution of the bootstrapped estimates).
6. Compare the original estimated EI (in Step 1) with the critical values from the previous step. If the original estimate is within the critical values, the hypothesis of an efficient market is not rejected. If the original estimate is outside, the null hypothesis is rejected and the studied series is considered as inefficient.

Logic behind the procedure is straightforward. The bootstrapping produces serially uncorrelated (independent, in fact) series with the same distributional properties as the original one. The distribution of the original series and its possible influence on specific estimators is thus taken into consideration.

We are interested in studying evolution of the market efficiency. To do so, we examine the series on the moving window. As the Bitcoin markets are opened 24/7, we study a window of a size of 365 trading days. For this window, the above-mentioned procedure is performed and we obtain the estimate, critical values and p-value. The window is then moved by 7 days (1 week) and the procedure is repeated until the end of the analyzed period. Eventually, we obtain the time series of the estimates, critical values and p-values so that we are able to comment on the efficiency evolution over the analyzed period.

3. Data

We examine two Bitcoin price indices constructed by the CoinDesk platform (coindesk.com) – the USD-based one and the CNY-based one. The indices are based on average prices over the major Bitcoin exchanges for the given currency. For the USD, the index is based on the Bitstamp, Coinbase, itBit and Bitfinex exchanges. The CNY index uses the

OKCoin exchange as it is the only one meeting the criteria set by CoinDesk[3].

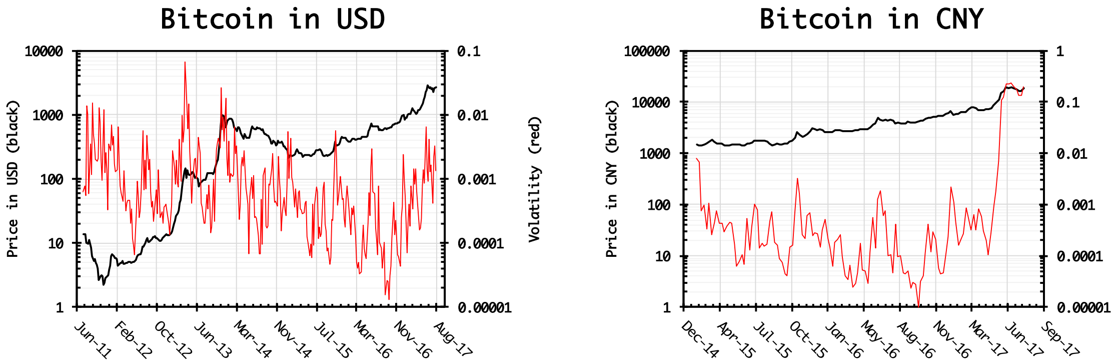

Figure 1: **Bitcoin price index.** Price evolution of Bitcoin price indices (black curve) for USD (left) and CNY (right) based on `https://www.coindesk.com/price/bitcoin-price-index/` is shown here together with volatility (red curve) utilizing range-based estimators.

The Chinese yuan Bitcoin price index is available since 1 February 2014 and we study its evolution till the end of July 2017 (1277 observations). For the US dollar Bitcoin price index, it was founded on 18 July 2010 and we study its behavior from the very beginning (2571 observations). Evolution of the indices is summarized in Fig. 1 together with the volatility evolution[4]. The prices shown in the logarithmic scale and they show an unprecedented growth from approximately \$10 to around \$3,000 at the end this period[5]. Together with a strong positive trend in prices, we observe decreasing levels of volatility. Until the end of 2016, volatility was in a rather stable decreasing trend but it started booming in 2017 (and pretty much exploding for the CNY market).

4. Results and discussion

Efficiency of two Bitcoin price indices (USD and CNY) is studied via the Efficiency Index, which combines effects of long-range dependence, fractal dimension, and approximate entropy. We focus on the efficiency evolution from the foundation of the given indices till the end of July 2017. To see how the Efficiency Index as a measure of efficiency evolves in time, we study its estimates on the moving window. Specifically, we estimate the index on 365 daily observations and we also calculate critical values for the null hypothesis of an efficient market and connected p-value (based on 333 bootstrapped estimates). Resulting evolution of the Efficiency Index for both USD and CNY market is illustrated in Fig. 2.

The results are quite straightforward. Starting with the USD market, we observe that there are only two longer periods of time when the market can be considered as efficient –

[3]More details can be found on the CoinDesk website, specifically `https://www.coindesk.com/price/bitcoin-price-index/`.

[4]Volatility is based on the range-based estimators and it is presented mostly for informative purposes.

[5]Note that between performing the analysis and finalizing the text, the price started attacking the level of \$10,000 at the end of November 2017.

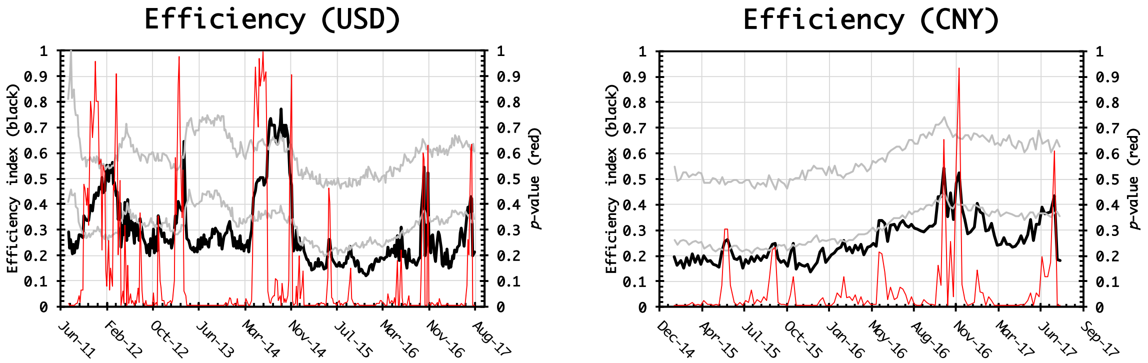

Figure 2: **Evolution of the Efficiency Index.** The Efficinecy Index (black curve) based on Eq. 1. is estimated on 365 daily observations. The critical values (gray curves) for the null hypothesis of an efficient market are based on 333 bootstrapped repetitions and they represent 5% and 95% quantiles. p-value (red curve) for the null hypothesis of an efficient market is based on the whole distribution of the bootstrapped estimates. Low p-values (with respect to the selected confidence level) thus represent inefficiency. The rolling window is presented with a step of 7 days.

from the middle of 2011 to the middle of 2012, and between March and November 2014. Both these periods follow after rapid price increases and they are characteristic by a rather mild dynamics of the Bitcoin price. There are few short-lived efficient periods following other slowdowns, namely at the end of 2012 and the second half of 2016. Apart from these periods, the market efficiency is rejected regularly. The results are not as strong for the CNY market as the examination period misses some very important bubble-like dynamics before 2014. The only efficient periods are thus the ones of "cooling off" after price surges. Apart from these, the Bitcoin markets are evidently inefficient and thus predictable.

To see the contributions of separate parts of the Efficient Index towards inefficiency, we present Fig. 3. There we can see a division between long-range dependence, fractal dimension and entropy, which can be translated into the contribution of global correlations, local correlations, and complex correlations. For the USD market and thus the longer analyzed period, we observe some changes in the structure of inefficiency. Until 2014, entropy and thus complex correlations played an important role which diminished in the more recent years. As for the global correlations, their contribution remains rather stable over the whole analyzed periods, and the original contribution of entropy has been overtaken by fractal dimension. In the most recent years, the inefficiency is driven mostly by the local correlations, i.e. short-term booms and busts. The later years of the USD market are closely followed by the CNY market.

The evidence of inefficiency is strong and visible from the presented results. How is it then possible that such inefficiencies remain and are not mined out by investors. This is very likely connected to the Bitcoin market (and cryptocurrency markets in general) characteristics. The market is still rather shallow and as such, it does not attract big institutional players. This might change soon as some exchanges are getting close to opening Bitcoin-based derivatives and assets copying Bitcoin price. Such change can have an ambiguous effect on the price as the big players will increase demand for Bitcoin which

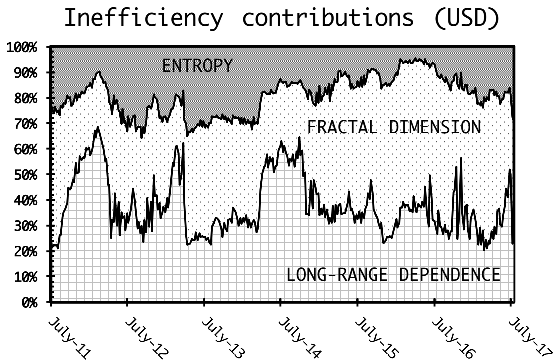

Figure 3: **Contributions to the Efficiency Index.** For each time step of the rolling window, proportions of the Efficiency Index are shown for long-range dependence (bottom), fractal dimension (middle), and entropy (top).

cannot increase its supply above the algorithm-given amount. From the other side, the deepening of the market usually leads to its increased efficiency and thus price stability in the Bitcoin price. At this point, the relatively low liquidity (compared to the stock or FOREX markets) does not ensure that an investor would be able to sell (or purchase) a large amount of the cryptocurrency for the given price. In addition, the Bitcoin price is still strongly driven by exogenous factors connected to its legal and/or security issues. The beauty of decentralization and low or no regulations can be pricy as well as conditions at different exchanges can vary strongly being it e.g. high transaction fees (not necessarily trading fees, but deposit and withdrawal fees) or (in)ability to withdraw the funds and transfer them to the fiat currency. The difference between being "in-the-money" and making an actual profit can be appreciable. Another complication in calculating actual profits arises when it comes to taxation. Country policies vary strongly as well and the discussion about the value-added tax (VAT) and income tax with respect to the crypto-world is still ongoing. Investors "cashing out" their bitcoins can then be on the edge of legality. Either way, the future evolution of Bitcoin and its price remains an exciting topic which will likely keep attracting attention both in the research community and in the investors community being it big institutional players or amateur enthusiasts.

Acknowledgments

The research leading to these results was supported by the European Union's Horizon 2020 Research and Innovation Staff Exchange programme under the Marie Sklodowska-Curie grant agreement No 681228. The author further acknowledges financial support from the Czech Science Foundation (grant number 16-00027S).

References

[1] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. http://bitcoin.org/bitcoin.pdf.

[2] M. J. Barratt. Silk road: Ebay for drugs. *Addiction*, 107:683, 2012.

[3] S. Barber, X. Bayen, E. Shi, and E. Uzun. *Financial Cryptography and Data Security*, chapter Bitter to Better – How to Make Bitcoin a Better Currency, pages 399–414. Springer; Berlin Heidelberg, Germany, 2012.

[4] J. Clark and A. Essex. *Financial Cryptography and Data Security*, chapter CommitCoin: Carbon Dating Commitments with Bitcoin, pages 390–398. Springer; Berlin Heidelberg, Germany, 2012.

[5] F. Reid and M Harrigan. *Security and Privacy in Social Networks*, chapter An Analysis of Anonymity in the Bitcoin System, pages 197–223. Springer; New York, NY, 2013.

[6] F. R. Velde. Bitcoin: A primer. *Chicago Fed Letter*, 317:1–4, 2013.

[7] L. Kristoufek. BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, 3:3415, 2013.

[8] S. Bornholdt and K. Sneppen. Do Bitcoins make the world go around? on the dynamics of competing crypto-currencies. *arXiv:1403.6378*, pages 1–5, 2014.

[9] D. Kondor, M. Posfai, I. Csabai, and G. Vattay. Do the rich get richer? an empirical analysis of the Bitcoin transaction network. *PLoS ONE*, 9(2):e86197, 2014.

[10] D. Garcia, C. Tessone, P. Mavrodiev, and N. Perony. The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface*, 11:2014623, 2014.

[11] L. Kristoufek. What are the main drivers of the Bitcoin price? evidence from wavelet coherence analysis. *PLoS ONE*, 10:e0123923, 2015.

[12] E. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417, 1970.

[13] E. Fama. Efficient capital markets: Ii. *Journal of Finance*, 46(5):1575–1617, 1991.

[14] A. Urquhart. The inefficiency of Bitcoin. *Economics Letters*, 148:80–82, 2016.

[15] S. Nadarajah and J. Chu. On the inefficiency of Bitcoin. *Economics Letters*, 150:6–9, 2017.

[16] A.F. Bariviera, M.J. Basgall, and W. Hasperué. Some stylized facts of the Bitcoin market. *Physica A*, 484:82–90, 2017.

[17] A.F. Bariviera. The inefficiency of Bitcoin revisited: A dynamic approach. *Economics Letters*, 161:1–4, 2017.

[18] J. Alvarez-Ramirez, E. Rodriguez, and C. Ibarra-Valdez. Long-range correlations and asymmetry in the Bitcoin market. *Physica A*, https://doi.org/10.1016/j.physa.2017.11.025, 2017.

[19] E. Fama. Random walks in stock market prices. *Financial Analysts Journal*, 21(5):55–59, 1965.

[20] P. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6:41–49, 1965.

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

[22] L. Kristoufek and M. Vosvrda. Commodity futures and market efficiency. *Energy Economics*, 42(March):50–57, 2014.

[23] L. Kristoufek and M. Vosvrda. Measuring capital market efficiency: long-term memory, fractal dimension and approximate entropy. *European Physical Journal B*, 87:162, 2014.

[24] L. Kristoufek and M. Vosvrda. Gold, currencies and market efficiency. *Physica A*, 449:27–34, 2016.

[25] J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238, 1983.

[26] J. Beran. *Statistics for Long-Memory Processes*, volume 61 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York, 1994.

[27] P.M. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23(5):1630–1661, 1995.

[28] M. Taqqu, W. Teverosky, and W. Willinger. Estimators for long-range dependence: an empirical study. *Fractals*, 3(4):785–798, 1995.

[29] M. Taqqu and V. Teverovsky. On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance Time Series. In *A Practical Guide To Heavy Tails: Statistical Techniques and Applications*, 1996.

[30] T. Di Matteo, T. Aste, and M. Dacorogna. Scaling behaviors in differently developed markets. *Physica A*, 324:183–188, 2003.

[31] T. Di Matteo. Multi-scaling in finance. *Quantitative Finance*, 7(1):21–36, 2007.

[32] J. Barunik and L. Kristoufek. On hurst exponent estimation under heavy-tailed distributions. *Physica A*, 389(18):3844–3855, 2010.

[33] V. Teverovsky, M. Taqqu, and W. Willinger. A critical look at lo's modified r/s statistic. *Journal of Statistical Planning and Inference*, 80(1-2):211–227, 1999.

[34] P. C. B. Phillips and K. Shimotsu. Local whittle estimation in nonstationary and unit root cases. *The Annals of Statistics*, 32(2):659–692, 2004.

[35] T. Gneiting and M. Schlather. Stochastic Models That Separate Fractal Dimension and the Hurst Effect. *SIAM Review*, 46(2):269–282, 2004.

[36] T. Gneiting, H. Sevcikova, and D.B. Percival. Estimators of fractal dimension: Assessing the roughness of time series and spatial data. Technical report, Department of Statistics, University of Washington, 2010.

[37] S. Pincus and R.E. Kalman. Irregularity, volatility, risk, and financial market time series. *Proceedings of the National Academy of Sciences of the United States of America*, 101(38):13709–13714, 2004.

[38] S. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 88:2297–2301, 1991.