

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,000

Open access books available

148,000

International authors and editors

185M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Chaos Analysis Framework: How to Safely Identify and Quantify Time-Series Dynamics

Markus Vogl

Abstract

Within this chapter, a practical introduction to a nonlinear analysis framework tailored for time-series data is provided, enabling the safe quantification of underlying evolutionary dynamics, which describe the referring empirical data generating process. Furthermore, its application provides the possibility to distinct between underlying chaotic versus stochastic dynamics. In addition, an optional combination with (strange) attractor reconstruction algorithms to visualize the denoted system's dynamics is possible. Since the framework builds upon a large variety of algorithms and methods, its application is by far trivial, especially, in hindsight of reconstruction algorithms for (strange) attractors. Therefore, a general implementation and application guideline for the correct algorithm specifications and avoidance of pitfalls or other unfavorable settings is proposed and respective (graphical) empirical examples are shown. It is intended to provide the readers the possibility to incorporate the proposed analysis framework themselves and to conduct the analyses and reconstructions properly with correct specifications and to be knowledgeable about misleading propositions or parameter choices. Finally, concluding remarks, future avenues of research and future refinements of the framework are proposed.

Keywords: nonlinear dynamics, attractor reconstruction, time-series quantification, chaos analysis framework, financial markets

1. Introduction

Following current estimates, the predictive analytics market is expected to grow from around 10.5 Bn. USD at the end of 2021 to around 28–30 Bn. USD until 2030, thus, stating the immense relevance of successful forecasting capabilities for the technological advancement in our digitalized, fully connected and global economy [1–3]. Therefore, respective fields of applications for predictive analytics (or related methodologies) can be represented by any real-world system interacting with practitioners or researchers alike [4]. For example, climate, hydrological cycles, ecosystems, the human brain, neuroscientific observations, the universe, engineering applications, economic systems or financial markets can be seen as such real-world systems [5]. Nonetheless, previously denoted examples are all classified as complex systems [5].

The meaning of referring to complex systems renders itself obvious once one regards contemplation of real-life confrontations, in which similar scenarios tend to evolve similarly and occur repeatedly [6]. The latter similar repetition of scenarios leads to the association of a predefined level of determinism in said real-life systems due to the timely evolution of memory and experience effects [6, 7]. Henceforth, quantitative modeling via deterministic differential equations proposes itself as a suitable methodology to cope with these kind of systems, since the respective entireties can be characterized by equivalent mathematical differential equations [6, 8]. Presupposing that the initial conditions of the referring systems are exactly disclosed, the respective differential equations enable the predictions of the systems' final states to an indefinite level of precision and time span due to the deterministic characteristics of the systems [6]. In terms of predictive analytics and forecasting attempts of such systems, presupposing said deterministic scenario, would illustrate the prerequisite of the future evolutions of the systems to be completely explicable via the current states, principally indicating a 'plainness' in terms of the predictability of such systems [9, 10].

However, scrutinizing one of the previously denoted examples as a representation of such systems with deterministic real-world characteristics leads to the emergence of unexpectedly drastic insights with vast practical implications [11]. The before mentioned real-world systems, such as financial markets, viewed as complex systems and driven by inherent or underlying empirical properties (i.e. stylized facts¹), result in a contrastingly challenging effort in terms of predictability and mathematical modeling in comparison to the before assumed 'plainness' of deterministic forecasts [6, 11]. Consequently, the determination of the true data generating process (DGP) of time-series, which are empirical observations of the underlying complex systems such as a stock price series for financial markets, with respect to stylized facts and other innovations is advantageous for the systems' observers, researchers or other involved entities [20, 21]. Under the presupposition of complex dynamical systems, seemingly conceptual differences are the basis for the discussion on the underlying nature and essential functioning of the emerging dynamics of, for example, financial markets or other defined real-world systems [6]. A deeper understanding of such assumed underlying laws of dynamical motions would facilitate the thorough application of chaos analysis in such real-world systems [22].

Substantial literature about testing underlying systems' dynamics and chaos in such real-world systems (e.g. financial markets) provides strong evidence of nonlinearity and as a consequence, a special class of models, namely chaos models, arose [11, 23]. Chaos institutes a deeper rationale for the above-mentioned essential characteristics and the underlying nature of the evolutionary processes driving a complex (real-world) system, which is affected by nonlinearities [10]. The first property, or distinctive feature of chaotic dynamical systems, is that even though deterministic, these systems characterize themselves via sensitivity to initial conditions², imply-

¹ Stylized facts, in particular, on financial markets can be volatility dynamics (e.g. [12]), nonlinearity (e.g. [13]), asymmetry (e.g. [3]), long memory (e.g. [14]), multifractality (e.g. [15]) and momentum driven trend characteristics, which clearly contradict the efficient market hypothesis [16]. Furthermore, studying stylized facts requires considering the heterogeneity of actors (e.g. [17]), resulting in multifractal timescales and behavioral patterns (e.g. [18]). All these properties occur at different timescales simultaneously, indicating the existence of stated nonlinearities (e.g. [19]) within the complex system of financial markets. Note that other complex systems may yield a similar variety of empirical characteristics to be regarded in respective predictive endeavors.

² Deviations from a trajectory of the system's phase (or state) space.

ing slight fluctuations or even marginal perturbations of the initial conditions to be capable to render precise predictions on a long time scale meaningless and futile in their totality due to exponentially increasing error terms [9, 22]. In addition, data measurement limitations³ in regard to current initial conditions, state an upper bound of the predictability, even if the model is completely disclosed [22]. Second to elaborate on, is the recurrence property, reflecting upon the dynamical behavior of such systems, which can potentially be exploited for the characterization of underlying dynamical evolutionary rules (or empirical DGPs) as presented later [6, 8].

Recent trends within chaotic dynamical analysis have led to a proliferation of publications, stating structural nonlinear models to be capable of displaying instabilities and chaos to be able to mimic empirical time-series properties⁴ [22]. Henceforth, a crucial pillar in nonlinear forecasting for over 40 years is the revelation of whether the considered time-series data sets are generated via deterministic or stochastic⁵ dynamical systems since their respective mathematical operations differ noticeable (see the bibliometric review of Vogl [25]) [23, 24, 26]. Speaking in a mathematical sense, a chaotic dynamical system has a dense collection of points with periodic orbits, sensitivity to initial conditions and topological transitivity, which is discussed in Eckmann and Ruelle [27, Devaney [28], BenSaïda and Litimi [29]. Chaos further refers to bounded steady-state behavior, which neither represents an equilibrium point, nor a quasi-periodic nor a periodic point nor indicates that nearby points separate exponentially in finite time, resulting in those chaotic systems revealing very complex and seemingly random evolutions out of the view of standard statistical tests [22].

Hence, chaos reveals apparent randomness of (chaotic) complex system realizations, yielding underlying patterns, interconnectedness, feedback loops, recurrence, self-similarity (fractality) and self-organization capabilities [30–32]. For example, in financial systems, hyperchaotic⁶ phenomena potentially evolve into crises, denying any form of system control [37]. Referring to scientific literature, the first tests of chaotic behavior for complex time-series systems were executed following the Brock-Dechert-Scheinkman (BDS) test of Brock et al. [38], yet, revealed its omnipotence, since it is unable to differentiate, whether the revealed nonlinearities originate from stochastic or chaotic dynamics [39]. Unfortunately, even comparisons between the most powerful tests (e.g. close-return test, BDS test and Lyapunov exponent⁷) do not result in conclusive findings [39]. In point of fact, several propositions toward a more conclusive solution in the scientific literature were brought to light, with no further positive indications [39].

The former statement is an allegory for the vast dilemma concerning the determination of the true, mostly unknown nature of complex dynamical (real-world) systems – whether it be stochasticity or chaoticity [40]. These systems are almost graphically similar and cannot be differentiated by respective statistical standard tests [29, 41]. Following Aguirre and Billings [41], a verification of strong noise influence on the identifiability of chaotic dynamics is provided, leading to misspecifications

³ In terms of measurement errors, sampling frequency and data accuracy, among others.

⁴ Thus, vast disseminations of literature about deterministic chaotic behavior and the design of (economic) models in the regime of chaotic behavior from a theoretical view arose [24].

⁵ Originating from pure randomness.

⁶ Hyperchaos is considered, if more than two positive Lyapunov exponents exist (e.g. [33–35]). If a discrete nonlinear system is dissipative (spontaneously symmetry breaking), a positive maximum Lyapunov exponent is an indication of chaotic dynamics within the system under regard [36].

⁷ A positive maximum Lyapunov exponent can occur even in non-chaotic series, due to inadequate application on noisy data sets [39].

of chaotic dynamics as stochastic dynamics due to noise-disturbance, rendering the discovery of evolutionary chaotic processes very difficult [42]. The majority of empirical time-series data is often small and noisy in comparison to its laboratory-based ‘physics’ data counterparts, suggesting a preclusion of dynamical identification if the noise levels are greater than a predefined critical threshold value [10, 41, 43]. Therefore, the great controversy of the nonlinear empirical literature as stated above is whether a complex system is characterizable via (low-dimensional) deterministic chaos or generated via stochastic dynamics and if those chaotic complex systems are controllable [10, 43].

To solve this ongoing debate, Vogl and Rötzel [40] and Vogl [44] successfully proposed distinct analysis frameworks, enabling the clear and safe quantification and determination of the underlying (empirical) DGPs of time-series data sets. Nonetheless, due to publication technical reasons, Vogl and Rötzel [40] present a framework specification tailored toward solely stationary time-series, while Vogl [44] supposes additional customization for non-stationary data. However, these specifications originate from one original, singular and holistic chaos analysis framework for nonlinear time-series, which will be presented in its totality hereinafters. Furthermore, since the determination of the (empirical) DGP by the proposed framework is by far trivial in its application, owing to a large variety of advanced algorithms to be implemented, this chapter is purposed to provide a clear and distinct practical guideline on how to successfully implement the denoted analysis framework. Particularly, the determination of ‘scaling regions’ via correlation sum and correlation dimensional schemes (refer to [45–47]) and the reconstruction algorithms of (strange, fractal) attractors of complex dynamical systems with chaotic traits (see [10, 48]) represent one of the main emphases, among others, since many erroneous conceptions are possible and are widely dispersed throughout the scientific literature [25, 40, 45]. The distinct goal and aim of this chapter are to provide the researcher and practitioner with an empirical-practical guide on how to implement the presented chaos analysis framework successfully, thus, determining the (empirical) DGP and reconstructing potentially existing system attractors out of a scalar time-series given [40, 44]. Moreover, the insights provided are mostly independent of the framework, generalizable and abstractable to any other kind of subsequent or related empirical analyses conducted.

Therefore, given in Section 2, the framework will be introduced completely and its inherent parts briefly reconciled, while the example data and subsequent correct specifications and selections to conduct a correct analysis are proposed in Section 3. Furthermore, Section 4 levels around the avoidance of pitfalls and the empirical results of misspecifications via practical examples, before concluding remarks and future avenues are discussed in Section 5. Please note that mathematical definitions or formulas are neglected and the reader is referred to the stated literature instead. If no further explication is granted, the literature is seen as a prerequisite for arguments and propositions, since the focus is purely on practical applicability in a theoretical-scientific context.

2. Framework overview

Before elaborating on the analysis framework itself, it is relevant to reconcile the contribution and relevance of the propagated inherent paradigm shift in quantitative modeling, namely, the previous determination of the empirical DGP and its

characteristics before selecting referring mathematical procedures and models [40, 44]. Referring back to the introductory mentioned expected market increase by around 20 Bn. USD in predictive analytics, it is crucial for successful forecasting to be informed about the underlying evolutionary building mechanics of the to be analyzed time-series data, before deploying cost-intensive predictive applications. One may imagine deploying planned-out long-term predictive solutions on time-series systems, which however are chaotic and, thus, only predictable over a short time scale. This would either lead to disastrous outcomes and very poor predictive accuracy or render model performance and the totality of forecasts futile due to exponentially increasing error terms owing to said chaotic mechanics [49, 50]. Instead, the proposed paradigm shift underlying the framework states the initial determination of the empirical DGP with inherent empirical characteristics, leading to exploitable knowledge about the predictive time horizons, hidden system properties and, therefore, the minimum model capability requirements, before practical implementation and roll-outs are conducted [40]. The model selections, thus, follow the insights of the determined empirical DGP [44]. Regarding the scientific side of the framework, existing literature and research do not execute sufficient theoretical precaution within respective applications and interpretations, leading to fragmentation and dispersion of methodology and modeling, thereby representing the rationale for the framework's creation [40, 45].

Hereinafter, the chaos analysis framework is presented in detail. In Section 2.1, the framework in general and its components are elucidated, while sections 2.2 through 2.7 level around the contents of each analysis step, while also introducing the inherently applied algorithms and methods in a sparse and nonmathematical as well as practical-error avoidant oriented manner.

2.1 Chaos analysis framework

The holistic chaos analysis framework presented in **Figure 1**, consists of six steps and will be elucidated hereinafter. Before elaborating on the steps in detail, the brief course of analysis will be outlined. First, it is mandatory to analyze given noise contamination and its respective levels and the nature of the potential noise [43]. Noise is capable of disturbing the identification of the underlying dynamics and, thus, is regarded as analysis destructive [45]. Furthermore, it is deemed favorable to gather basic statistical insights from the (denoised) datasets under analysis via determination of standard statistical tests, which incorporate tests for stationarity, nonlinearity and correlations, among others [40, 45]. It is possible to determine the applicability of reconstructions solely based on these insights. Second, several chaos measures and nonlinear metrics are calculated such as the sample entropy (see [51]), Lyapunov exponents (refer to [52, 53]) or the Hurst exponent (see [54]). These insights are relevant to determining the nature of the underlying dynamical system based upon mathematical procedures. Third, if applicable, (strange) attractor reconstruction algorithms can be implemented to reconstruct the system's attractor visually. Fourth, an independent recurrence quantification analysis (RQA) paired with discrete wavelet transformations (DWT, refer to [55]) can be applied to (a) determine the existence of various sub-dynamics and (b) exploit denoted recurrence properties mathematically as well as visually [6, 8, 56]. This reveals hidden characteristics of the analyzed datasets. Fifth, spectral characteristics, especially exploitable in forecasting by applying fractional calculus, are analyzed via wavelet-based multiresolution analysis (MRA) [57–59].

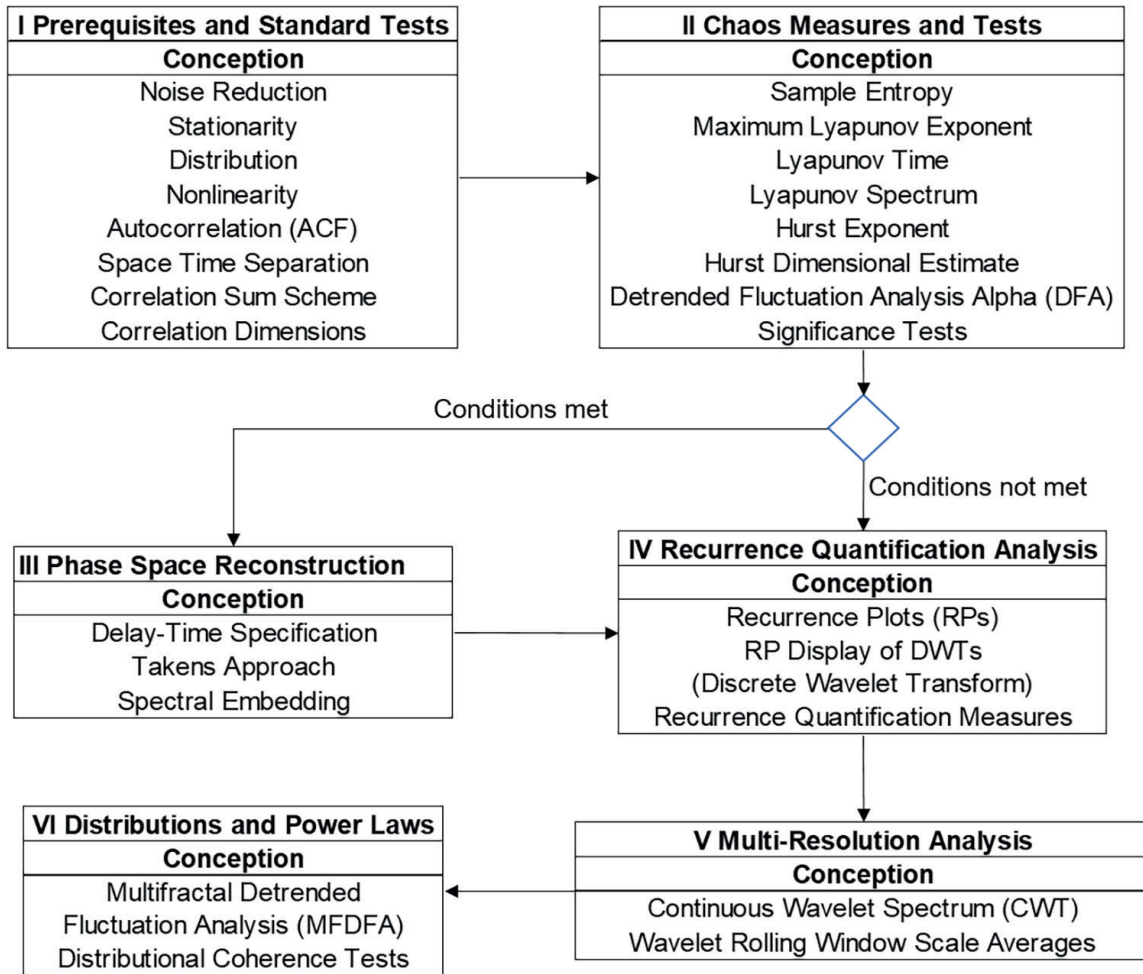


Figure 1. Generalized chaos analysis framework for the determination of the empirical DGP and underlying empirical system characteristics based on scalar time-series data taken with permission of [40, 44]. Step I (prerequisites and standard tests) consists of testing prerequisites, which are required to conduct a nonlinear dynamical analysis. Therefore, noise reduction is mandatory, followed by tests for stationarity, Gaussianity (distribution in general), nonlinearity, and space-time separations, which can prevent an analysis. Step II (chaos measures and tests) encompasses a collection of effective nonlinear dynamical or chaotic measures. First, a correlation sum scheme is applied to determine and test significant ‘scaling regions’. Moreover, the dimensionalities and properties of the system are tested (e.g. correlation dimension, Lyapunov exponents). Furthermore, information content via sample entropy is analyzed, among others. Step III (phase space reconstruction) involves the proper reconstruction of the system and a graphical representation using embedding theorems such as the traditional Takens’ embedding. Step IV (recurrence quantification analysis) is an independent confirmation of the previous steps, namely, the ability to describe and quantitatively measure the characteristics of the underlying dynamics, optionally, with the application of rolling window scale averages and a subsequent discrete wavelet transformation (DWT) application to determine the potential existence of sub-dynamics within the data. Furthermore, the quantification is not dependent on the prerequisites of steps I–III. Step V (multiresolution analysis) elaborates on the spectral properties of the data and is elucidated via continuous wavelet power spectra (CWT). Step VI (distributions and power-Laws) is to determine power-law characteristics via multifractal detrended fluctuation analysis (MFDFA) and distributional coherence tests.

Finally, (multi)scaling and (multi) fractal characteristics are elaborated on via the conduction of a multifractal analysis. The multifractal analysis includes a multifractal detrended fluctuation analysis (MFDFA, see [60]) as basis, which renders (locally) minimum and maximum Hurst exponents graphically visible, while subsequently providing inherent scaling coefficients. In particular, generalized Hurst exponents, multifractal scaling exponents and the multifractal scaling spectrum can be derived, thereafter. Moreover, distributional coherence tests can be applied to validate, the ‘less worst’ distributional fit and, whether a power law is present in the data (refer to

[61]). In total, holistic insight into the underlying empirical DGP and inherent characteristics is obtained, with which one may select appropriate models thereafter.

2.2 Prerequisites and standard tests

Initially, the prerequisites and standard tests are presented briefly. The time-series has mandatorily to be denoised properly to ensure a descent beyond a predefined threshold level, best via nonlinear filter schemes [43, 45, 58]. Most time-series are contaminated by noise due to measurement errors and microstructure noise occurrences [39, 62]. Following Aguirre and Billings [41], noise exerts a mentionable (negative) influence on the identifiability of processes inherited via chaotic dynamics. Henceforth, if a certain level of noise is exceeded, accurate estimations of dynamic models and subsequent analyses are voided in their entirety⁸ [41]. The only feasible approach, therefore, is the drastic reduction of noise levels to ‘workable levels’, since contaminating noise in evolutionary dynamics may be dynamical noise in either additive or multiplicative specification, thus, disrupting the dynamical identification on several even small scales [26, 63]. Regarding the nonlinear filter structures, two criteria have to be met, namely, (1) the applied filters are required to be unbiased and (2) the residual variance of the filters levels the noise variance [41]. Please note that some nonlinear filters such as the median filter Introduce (artificial) autocorrelations in the data, which should be avoided, thus, wavelet filters are deemed to be favorable for nonlinear denoising (refer to [55, 65, 66]).

Once the time-series is successfully denoised, standard statistics can be applied to elaborate on primal insights into the underlying mechanics [40]. Within the framework, the first and – destructive of reconstruction algorithms if missing – property is stationarity⁹ [40]. One has to exert special strictness in terms of stationarity, thus, proposing a 1% significance level for two successive tests is deemed favorable [44, 45]. For the framework, the augmented Dickey-Fuller (ADF) test and the more powerful Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are executed (refer to [67, 68]), which both have to concur to be regarded as valid results in terms of stationarity. To elaborate on the initial test of distributional characteristics, a Kolmogorov-Smirnov (KS) test for a Gaussian specification (refer to [69]) is conducted, yet, other specifications are possible. Nonetheless, a 1% level of significance is recommended to adhere to the strictness of the presuppositions of the analysis. Moreover, to test for the existence of nonlinearity, which is a prerequisite for the existence of chaotic dynamics, the BDS test (refer to [38]) is executed. Please note that due to its stated omnipotence, it is only applied to identify nonlinearity in general and specifically not to distinguish stochasticity versus chaoticity [39]. Further note that sufficiently enough embedding dimensions have to be selected for the BDS test as well as subsequent methodologies to meet practical insights.

Lastly, correlation structures have to be elucidated, beginning with the calculation of autocorrelation functions (ACFs, see [70]) with sufficiently large lags (e.g. 100–300). The ACFs serve as the basis for the validation of potential reconstructions (see Section 2.4) and indicate, whether analysis disturbance is given. Moreover, following Kantz

⁸ For example, too much noise, leads to test rejections, disrupts the Grassberger-Procaccia algorithm (see [46, 47]), thus, the correlation dimensional estimates and alters the Lyapunov exponent calculations [63, 64].

⁹ Even if scientifically debatable, ‘brute-force’ methods such as logarithmic distances will provide sufficient results, since the sole purpose is analysis not forecasting, thus, no drawbacks are to be expected.

and Schreiber [45], nonzero autocorrelations are deemed problematic owing to trajectory vectors being closely located in phase space due to continuously evolving time, which is also known as temporal correlation. To determine a relevant ‘scaling region’ by the application of correlation sum schemes (see [46, 47]), the absence of temporal correlations is mandatory due to fitting issues in regional curve shapes and the lack of invariance of said correlation sums as depicted in [45]. Hence, dynamically correlated time-series violate the estimation requirements and if sufficiently large or worst oscillating, the analysis is futile [45].

To analyze temporal correlations, Provenzale et al. [71] propose estimates of the correlation time by applying time separation plots, presupposing pairs of points in phase space to be dependent on threshold distance and, additionally, on elapsed time in between respective measurements. Henceforth, the contour curves of said plots have to saturate and remain at an acceptable and non-oscillating boundary level [71]. The existence of a sufficient ‘scaling region’ is the premise for successful reconstructions [45]. Building upon the correlation sum scheme, determining the slopes of each correlation sum curve per selected embedding dimension results in an estimate of the (fractal) correlation dimension¹⁰, which is plotted by itself and has to saturate as well (refer to [45]). Novelty within the framework to determine the validity of the underlying ‘scaling region’ is a step difference test proposed by Vogl and Rötzel [40], which tests step differences of the correlation sum curves in a Student’s t-test against zero and graphically examines the resulting p-value heatmap. One can select the minimum embedding dimension by selecting the one, which has no p-value above 1% significance. Note that the existence of an ongoing ‘scaling region’ is also directly visible in the heatmap.

2.3 Chaos measures and tests

In addition to the prerequisites, several singular chaos metrics are worth determining to gather more initial insights into the potential underlying nature of the time-series dynamics under analysis [40]. First, the sample entropy as proposed in Richman and Moorman [51] is calculated, reflecting information content and self-similarity characteristics, thus, delivering insights into the presence of fractality within the data.

Furthermore, various Lyapunov exponents are determined, namely, (1) the maximum Lyapunov exponent, (2) the Lyapunov spectrum and (3) the Lyapunov time. Lyapunov exponents measure chaotic strength in a dynamical system by measuring the exponential convergence or divergence of nearby trajectories in phase space [45, 73]. It is possible to calculate Lyapunov exponents equaling the number of phase space dimensions, i.e. the number of the estimated embedding dimension, leading to the Lyapunov spectrum, which indicates the nature of the underlying dynamical systems, whether it be conservative or dissipative [74]. The largest exponent is labeled as the maximum Lyapunov exponent, depicting the exponential divergence or convergence of close trajectories and can be determined via the algorithm of Rosenstein et al. [75]. Note that a positive maximum Lyapunov exponent in combination with a negative Lyapunov spectrum sum is mostly seen as a sign of chaos, yet, is critiqued by the lack of distributional tests [11]. Therefore, a distributional rationale in form of the Bask-Gençay bootstrapping test is favorable, since it provides

¹⁰ In finite scalars like time-series, according to Ramsey et al. [72], correlation dimensional estimates tend to return artificially smaller values than the theoretically assumable fractal dimension.

a significance indication, particularly, in cases of small positive, beyond zero maximum Lyapunov exponents at a sufficient level of significance [76]. Please note that several ten to a hundred thousand of bootstrapping steps are advisable to obtain reliable results. Moreover, the Lyapunov time represents the inverse of the maximum Lyapunov exponent and, thus, implies the time-span, the system requires to render itself chaotic and non-predictable, i.e. the time in which the exponential growing errors remain in a 'forecastable' range, before diverging too far [53]. The Lyapunov time is interpretable in either time-series units or in SI units [seconds] for real-world applications [53].

Finally, the Hurst exponent or in the case of non-stationary data, the detrended fluctuation analysis (DFA) alpha value is calculated to obtain in-depth information about the evolutionary nature of the dynamical system [54, 77, 78]. In an ongoing debate, the interpretation of the Hurst exponent and its initial interpretation by Benoit Mandelbrot (see [79, 80]) is challenged [16, 25, 44].

The Hurst exponent is interpreted as follows, namely, (1) the system is representing a Wiener process¹¹, should the Hurst exponent equal exactly 0.5, (2) the system is revealing long memory effects if it is exceeding 0.5 and (3) is being a mean-reverting system, should the exponent value be below 0.5 [16]. Nonetheless, recent empirical studies (refer to [16, 44]) state that the exceedance of 0.5 by the Hurst exponent reveals measurable fractal trends (or trending characteristics), which are an explicative rationale for momentum effects on financial markets. Within the setting of this analysis, the latter, novel indication is more suitable. The exceedance of 0.5 indicates persistency and the existence of a power-law, resulting in the denoted fractal characteristics [16, 44, 81]. Additionally, the Hurst exponent can be applied to determine the fractal dimensionality estimation (2-H) [82]. Finally, as an additional novelty, is adapting the Bask-Gençay test to the Hurst exponent as depicted in Vogl [44], ensuring the said exponent to be tested on significance [83]. In total, the second step enables the elucidation of the dynamical systems' properties directly, thus, providing a solid indication of its underlying evolutionary nature.

2.4 Phase space reconstruction

An important step toward the conduction of successful predictions of nonlinear time-series systems is the method of attractor reconstruction, leading back to the 1920s (refer to [84]) and the ideas of Packard et al. [85], Ruelle [86] and Takens [48], which represent the calculation of various invariant quantities required to characterize the underlying system [87]. This is mostly the presupposition for the nonlinear dynamical analysis of a time-series and state space model implementations [88]. The main contribution of reconstructions is given by the reconstruction of phase space, which is capable of preserving geometrical invariants (e.g. eigenvalues, fixed points or fractal dimension) of referring system attractors, including the Lyapunov exponent of according trajectories [88]. To phrase it differently, attractor reconstruction can be seen as a method to recreate the full deterministic state space based upon a lower dimensional time-series (i.e. a scalar) [87]. Thence, state space reconstruction is the generation of a multidimensional, deterministic state space out of the underlying, sampled time-series data [88]. Furthermore, embedding is, thus, the mathematical

¹¹ Only in this scenario the efficient market hypothesis taken out of quantitative finance holds and is violated else.

process by which an attractor is reconstructable presupposing a given set of scalar measurements, i.e. time-series datasets owing to dimensional preservation characteristics [43].

The resulting accuracy of the attractor reconstruction is directly dependent on the methodology applied to the reconstruction process and also influences the Lyapunov spectrum [87]. Therefore, several problems may occur, since the Lyapunov exponent cannot be labeled as invariant toward initial conditions, thus, stating a dependence on sample size within the reconstruction of time-series trajectories in phase space [11]. Following Nichols and Nichols [87], several methods for delay-time and embedding dimension selection exist for the standardized delay coordinate reconstruction, namely, the comparison between ACFs and the probabilistic concept of mutual information, while false nearest neighbor approaches are feasible to minimize said delay vectors. Nonetheless, the most common procedure is the delay-time reconstruction in combination with various embedding dimensions [89]. A non-exhaustive overview is proposed in **Table 1**. To be more detailed, the delay-time is defined as the time-span between two neighboring points applied to reconstruct the attractor, while the referring embedding dimension represents an estimate of the true dimension of the assumed phase space, which is intended to be reconstructed [97]. To refer back to the denoising presupposition given in Section 2.2, the underlying scientific theory requires noise-free data, on which natural processes timely evolve, which else leads to difficulties in state variable estimations [87].

According to Takens [48], in absence of noise contaminants, it is always feasible to embed a scalar time-series into a state space. Assuming the existence of noise, two trajectories of the same initial condition, potentially evolve differently and converge to different asymptotic behavior, thus, even the exact knowledge of said initial conditions does not guarantee the predictability of the system's final state [88, 98]. Therefore, noise has to be treated as an influential source of unpredictability, which cannot be fully disclosed via the deployment of conventional methodologies of nonlinear dynamical analysis such as exit bases or uncertainty exponents [98].

Algorithm	Short Description	Reference
TE	Takens' delay-time embedding, implies shifting a timely delayed comb through the data to generate 3D coordinates	[48]
SE	Spectral embedding in combination with a k-nearest neighbor algorithm, principal component analysis and Laplacian Eigenmaps	[10]
LRNN & CRBP	Locally recurrent neural networks with casual recursive backpropagation learning by applying algebraic observability through Takens' theorem	[90]
HH	The reconstruction is conducted by a combination of hyperhelices	[91]
ESN	The reconstruction is conducted by echo state networks	[92]
SVM	The reconstruction is conducted by support vector machines	[93]
HT	State space reconstruction by Hilbert transformations	[94]
MI	Mutual information in the probabilistic method	[95]
TM	Trajectory matrix for singular system analysis	[96]

Table 1. Overview of existing reconstruction algorithms within the scientific literature.

To revert to practical implementations, it is relevant to determine the choice of delay- times (τ), which directly influences the success and accuracy of reconstruction algorithms [87]. Hence, a too small selection results in vectors to be very near and almost identical, thus, carrying redundant information and leading the attractor to collapse onto the 45° line in state space [87]. In contrast, a too large selection will produce uncorrelated (unrelated) coordinates owing to exponentially growing errors in chaotic regimes, resulting in decorrelated vectors in hindsight of the underlying time-series [87]. Henceforth, the two boundary scenarios have to be well-balanced to receive a proper reconstruction, yielding maximal independence, while preserving dynamically related coordinate properties [87]. The most commonly applied variant is the ACF delay, with several possibilities, namely, (1) the first zero crossing, (2) crossing of 0.1, or 0.5 and (3) not exceeding $1/e$ [87]. Please note that ACFs propose linear time evolutionary calculations and may, thus, be misleading [87]. Due to experiments, the most accurate representations by the author were achieved by selecting variant (1), i.e. the first zero crossing or in modification, the first zero crossing, while subsequent coefficients additionally stay insignificant. Moreover, the embedding dimension as shown by Sauer et al. [99] has to be topologically equivalent to the true attractor, if the embedding dimension is chosen to be larger than two times the fractal dimension of said attractor. Note that once the embedding dimension is selected sufficiently high, a reconstruction resembles almost always an embedding, independent of parameter selections [88]. Mostly, delay-coordinates are selected, yet, there exist the families of derivatives and principal component reconstructions, as depicted combinatorial in the spectral embedding (see [10]) [88]. Within practical applications, the author deems a combination of (1) Takens delay-time embedding, which, unfortunately, resembles a ‘spaghetti monster’ in most cases and (2) the more sophisticated variant by Song et al. [10], applying a spectral embedding in combination with a k-nearest neighbors algorithm (k-NN), principal component analysis (PCA) and Laplacian Eigenmaps as very suitable. In the author’s empirical experiments, the PCA components are best selected to equal the embedding dimension, while the number of neighbors for the k-NN can be best determined by the following heuristic, namely, $0.01\text{len}(\text{data}) * 1.5\tau$ [40]. Moreover, to receive a correct reconstruction the properties stated in **Table 2** are strict to be adhered to.

Parameter	Favorable	Negative	Impact
Stationarity	stationary	non-stationary	very high
Nonlinearity	nonlinear	linear	medium
ACFs	not significant	significant	high
Space-Time Separation	very low, non-oscillating	high, oscillating	high
Correlation Sum	significant ‘scaling region’	no ‘scaling region’	very high
Correlation Dimension	saturating	not saturating	high
Maximum Lyapunov Exponent	significant and positive	negative	high
Lyapunov Spectrum Sum	negative	positive	high

Table 2.

Overview of parameter selections for attractor reconstruction specifications.

2.5 Recurrence quantification analysis

Independently from previous attractor reconstruction and other prerequisites, the RQA bases itself upon the introductory denoted exploitation of the recurrence property¹² of dynamical systems, thus, is applicable to any time-series data (e.g. [6]) [100]. The RQA is conducted by quantifying the recurrence plot (RP) as introduced in Eckmann et al. [101]. The RP and RQA analysis benefit from the preservation¹³ of time-ordering information contents given in the analyzed data as well as contained spatial structure [22]. With the RQA one may detect fundamentally given characteristics underlying a dynamical system, namely, the recurrence states, resulting in a ‘robust to noise and data limitations’ method of quantifying and identifying (chaotic) dynamics [22]. Thus, respective trajectories and transitions are rendered visible, in combination with the degree of complexity, i.e. the fractal structures, which may be inherent in the analyzed data [8]. To determine the RQA a threshold level has to be decided on, which determines whether nearby points are counted as recurrent or not [7]. Following van den Hoorn et al. [102], propose several threshold determination methods, yet, traditionally, according to Koebbe and Mayer-Kress [103] as well as Zbilut and Webber Jr. [104], the threshold value should not exceed 10%. Additionally, the threshold value should not be lower than five times the sample noise [6]. Furthermore, it is common to exclude the identity line (proportional to the maximum Lyapunov exponent) of the RP from the analysis [105]. First, the RPs can be interpreted visually, which is presented in Marwan et al. [6], p. 251. Second, there exist two different types of measures taken out of a RQA, namely, minima-dependent versus single-value measures [6]. One may plot the minima dependency for several selections and choose an appropriate value to quantify the RPs. The length of diagonal lines represents the duration of trajectory local evolutions, while vertical (horizontal) lines mark time durations, in which the underlying dynamics are trapped (labeled as intermittency or laminar state) [7]. The commonly applied measures are depicted in Vogl and Rötzel [40], **Table 3**. Finally, to distinguish the results from stochastic, chaotic or other systems, one may either apply a Wiener process realization, a mathematical chaotic system realization or respective surrogate¹⁴ datasets (see [45]). Paired with the conceptions taken out of Section 2.6, signal theoretical decompositions can be applied to identify potential hidden sub-dynamics [56]. To gather and obtain the most information out of the analyzed data signal, wavelets with better localization properties are commonly proposed in form of a DWT filter bank [44]. The resulting low pass and high pass decompositions can then be applied as novel datasets to the RQA analysis and sub-RPs can be created and quantified to demonstrate potential sub-dynamics [56]. Note that the process can be repeated as often as required, should more than residual noise remain after one respective decomposition or iteration.

2.6 Multi-resolution analysis

To be very brief, time-series data are localized in the time domain, yet, may also yield exploitable frequency components, which in case of non-stationarity or

¹² The recurrence property originates from a topological approach and is given by the Poincaré recurrence theorem.

¹³ Presupposing the existence of a low-dimensional attractor, presence of dependence on initial conditions and the manifestation of said recurrence property.

¹⁴ I.e. destroying given determinism by shuffling via FTs. Then, comparison with original data.

Parameters	Generic	Mars	S&P500	Implication
Denoising	no	yes	yes	Denoising with level 6 cascade wavelet filter and 'Bior 2.8' wavelet, since 'biorthogonal' states the best localization characteristics for non-tailored wavelet functions.
Stationarity	yes	yes	yes	ADF and KPSS (c, ct) tests with 1% significance each
Gaussianity	no	no	no	KS test with 1% significance
Nonlinearity	yes	yes	no	BDS test for embedding dimension of five and 1% significance
Correlation Sum	scaling	no scaling	no scaling	Build upon correlation sum graphs and step-test with 1% significance
Correlation Dimension	saturating	dropping	not saturating	Based upon different correlation dimensions for embedding dimension
ACFs	very low	depending on lag	very low	ACFs with 300 lags and 1% significance
Temporal Correlations	none	very strong	oscillating	Graphical via space-time separation plots for min. 100 steps
Sample Entropy	very low	low	very low	Very low entropy is seen as a sign of self-similarity in terms of information contents
Maximum Lyapunov	low positive	low positive	low positive	Tested for significance via Bask-Gençay test to propose the distributional theory
Lyapunov sum	negative	negative	negative	Negative sum indicates dissipative system in combination with maximum Lyapunov exponent positivity
Hurst Exponent	trending	trending	mean reverting	Trending: Hurst >0.5; mean reverting: Hurst <0.5
Result	positive	negative	negative	Reconstruction invalid for real-world datasets, even if (non)linear dissipative systems

Table 3.
 Overview of steps one and two for all datasets with implication.

non-periodicity or other unfavorable characteristics will not be extractable via classical Fourier transformations (FT) [13, 106, 107]. Therefore, wavelets (i.e. tailored or bi-orthogonal) applied in a multi-resolution analysis (MRA) are well suited to extract underlying frequency information by retaining as much time localization information as possible [13, 16]. Thus, for filtering or denoising activities of time-series, discrete cascade filter banks with wavelet shrinkage (see [55, 108]) are applicable at length for various scales [40]. Moreover, to obtain insights into time-frequency localizations of to-be-analyzed datasets, one may apply a continuous wavelet transformation (CWT), resulting in a spectrum [57].

2.7 Distributions and power-laws

To elaborate on power-law characteristics, it is important to denote the interconnections between chaotic dynamics, strange attractors, fractals and power-laws. In short, a dissipative (chaotic) dynamical system will reveal its phase space over timely

evolution to deflate onto its own strange attractor, which is characterized via a fractal set [109, 110]. Generally, a fractal set yields a non-integer (non-Euclidean, thus, generalized) dimension, namely, the Hausdorff-Besicovitch dimension and is further characterized via self-similarity, i.e. multi-scaling, in addition to irregularities, non-differentiability and recursiveness [111]. Henceforth, a (multi) fractal system requires a local power-law contributing to the mentioned scaling properties [111]. Therefore, a power-law is defined as the scalar relationship between two quantities and, thus, is characterized via scale invariance [112]. A fractal system with one scaling exponent is labeled monofractal, yet, multifractal systems require a singularity spectrum of exponents [111]. Referring back to a dissipative dynamical system, which deflates onto its strange attractor, thus, is represented by a fractal set. The fractal set of a strange attractor is rendered visible via its Poincaré sections, which show intersections of said strange attractor [110, 111]. To be more detailed, the intersections of strange attractors are fractal sets, which are described via multi-scaling and, thus, via powerlaws [111].

Analyzing time-series enables not only the reconstruction of potential (strange) attractors, yet, opens the way to mathematically determine given power-laws (i.e. the multi-scaling characteristics) of its underlying (multi) fractal properties [113]. Following Yuan et al. [114], state two rationales for time-series multifractality, namely, (1) the existence of fat-tailed probability distributions and (2) nonlinear temporal correlations. To draw out the multifractal spectrum, one may apply a multifractal analysis, which is built upon the MFDFA. The MFDFA visually depicts the scaling properties, as well as the (local) maximum and minimum Hurst exponents, also supporting the fractal trending interpretation discussed earlier [60]. Moreover, the generalized Hurst exponents, multifractal scaling exponents and the previously denoted multifractal scaling spectrum can be derived from the MFDFA output quantities [60]. In addition, calculating complementary cumulative distribution functions (CCDFs) and comparing them with power-law or other potential theoretical distribution types, enables the more or less save determination of power-law or other distribution fits [61]. However, as a word of absolute caution, the determination of coherence tests for various distributions has to be interpreted very carefully. The coherence tests are calculated via paired distributional fitting comparisons based upon log-likelihood measures, alongside other parameters [61]. These serve the purpose of achieving insights into suitable distributions, which may describe the datasets best, or to phrase it realistically, which at least represent the 'less worst' fit [61]. The coherence tests, thus, represent a comparison and no goodness of fit, which as indicated requires the reader to exert special care with the interpretation. It is advisable to fall back on graphical displays on log-log plots, which revealed as a useful guide for the practical implementations of the author. Concluding the powerlaws, the analysis is complete and the interpretation can carefully be exerted.

3. Correct empirical specifications

For each step of the analysis several algorithms are to be determined and a larger variety yield graphical insights, which can be either quantified or applied as a visual aid to deduce further insights and implications. Since a complete analysis as shown in Vogl and Rötzel [40] or Vogl [44] would vastly exceed the page limitations of this guide, the didactics of the practical display are as follows. First, this section will provide an idealistic outcome of a generic and mathematically tailored time-series,

based upon the Lorenz system (refer to [115]) to demonstrate the ‘best-case’ scenario as the generalized point of reference, while two additional real-world datasets are presented as a comparison, namely, (1) the change rate of wind speed of Mars and (2) S&P500 1-minute tick return data. Second, the real-world datasets will be elaborated on in Section 4, since some hindrances are given and, thus, require analysis of potential misspecifications. The preliminary elaborations via steps one and two are depicted in **Table 3**, while **Figure 2** presents the correlation sum, the correlation dimensional scheme as well as the correlation structures for the generic dataset. Note that for attractor reconstructions, **Table 2** already proposes the favorable characteristics to enable the correct implementation of reconstruction algorithms. Steps three

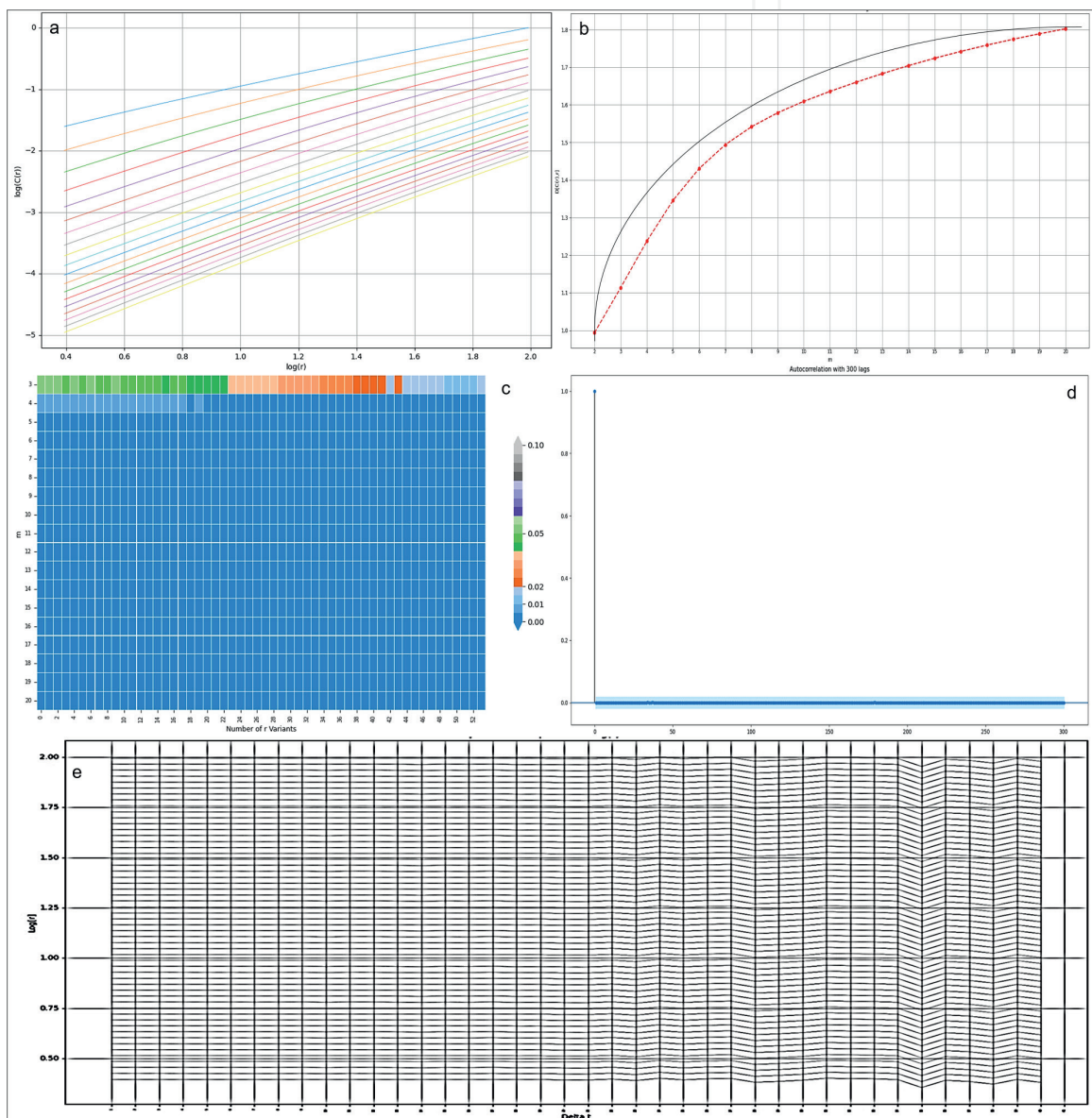


Figure 2. Generalized point of orientation via the generic dataset for determination of reconstruction possibility. (a) Shows the correlation sum plots, which visually depict ‘scaling regions’, (b) shows the correlation dimensional plot, which ideally saturates as shown, (c) states the heatmap of p -values for the step-test of the lines of (a), (d) states the ACFs for 300 lags, which are insignificant and (ideally) stay that way and (e) shows the space-time separation plot for 100 steps, which is very low and non-oscillating. Note that for step two of the analysis, the minimum embedding dimension can be taken out of the heatmap, namely, by picking the first row with only 1% significance or lower p -values. During the existing experiments of the author, the heuristic of two times the fractal dimension as stated in the main text is also given by applying this selection method.

to six are only stated in Section 4 for the real-world examples. In general, the stricter the interpretation and analysis, the better the results of the reconstruction, correct specification of underlying empirical DGPs and subsequent modeling.

Moreover, while steps one and two as presented above (in addition to potential reconstructions of step three), suffice to quantify the nature of the underlying system (i.e. whether it is dissipative, reconstructable or potentially chaotic), regarding the analysis of steps four through six, provide the exact quantified details of the system's characteristics and serve as double confirmation procedure. Mainly, the RQA measure quantification provides exact details about the underlying empirical DGP, namely, a percentage comparison with surrogate data or pure stochastic (e.g. a Wiener process) or pure deterministic chaotic systems (e.g. a Lorenz system) (refer to [6, 40]). Thence, one is capable of pinpointing whether the underlying system is pure chaotic, pure stochastic or a mixture of both and in which margins [40]. Subsequently, deriving frequency information, information about sub-dynamics, the existence of power-laws and multifractal spectra enables the correct model selection as final outcome of the quantification. Nonetheless, a proper differentiation can be achieved after step four already. Since the RQA is too spacious it is neglected for this display, thus, the reader is referred to [6, 8, 40, 56].

4. Empirical negative examples

Continuing the previous section, hereinafter, the results for the two real-world datasets are presented in **Figures 3** and **4**, while the remaining indications are provided in **Table 4**. The invalid reconstruction algorithm results via Takens delay-time embedding (refer to [48]) and via spectral embedding (refer to [10]) are proposed in **Figure 5** to clarify the relevance of proper prerequisites analysis. Regarding the Mars wind speed change rates, clear deterministic traits and sub dynamics are observable by the RQA, yet, a clear identity as a chaotic system as well as a distinct reconstruction is not possible. This is illustrated by the lack of a clear 'scaling region', dropping correlation dimensions, high ACFs and temporal correlations, which render this analysis step invalid. Furthermore, no scale independent multifractal scaling spectrum is visible and a nested power-law-exponential distribution is proposed as 'less worst' distribution via the coherence tests (see [61]). In addition, no frequency information is determinable via CWT. Thus, the only insight generated is that it is a potentially chaotic, deterministic and dissipative system, while the exact modeling metrics are extractable out of the RQA quantification tables (refer to [6]). Regarding the (invalid) Takens reconstruction may suggest a non-chaotic attractor, since the results resemble a valid one in parts, yet, this is an invalid approach nonetheless. Spurious chaotic measure results are obtained by the S&P500 1-minute return series, since according to the BDS test nonlinearity is excluded, while the Hurst exponent indicates a clear mean-reversion. Furthermore, no 'scaling region' by correlation sums and a non-saturating correlation dimension in combination with oscillating temporal correlation voids any other step of the analysis or reconstruction. Regarding the reconstruction by Takens, the linear nature is determinable. Moreover, the system has frequency information, yet, no power-law nor multifractal scaling characteristics (in agreement with the Hurst exponent indication of mean-reversion). Following the RQA, sub dynamics and low levels of determinism are given, while vast stochastic characteristics are dominant. A final concluding remark at this point, considers

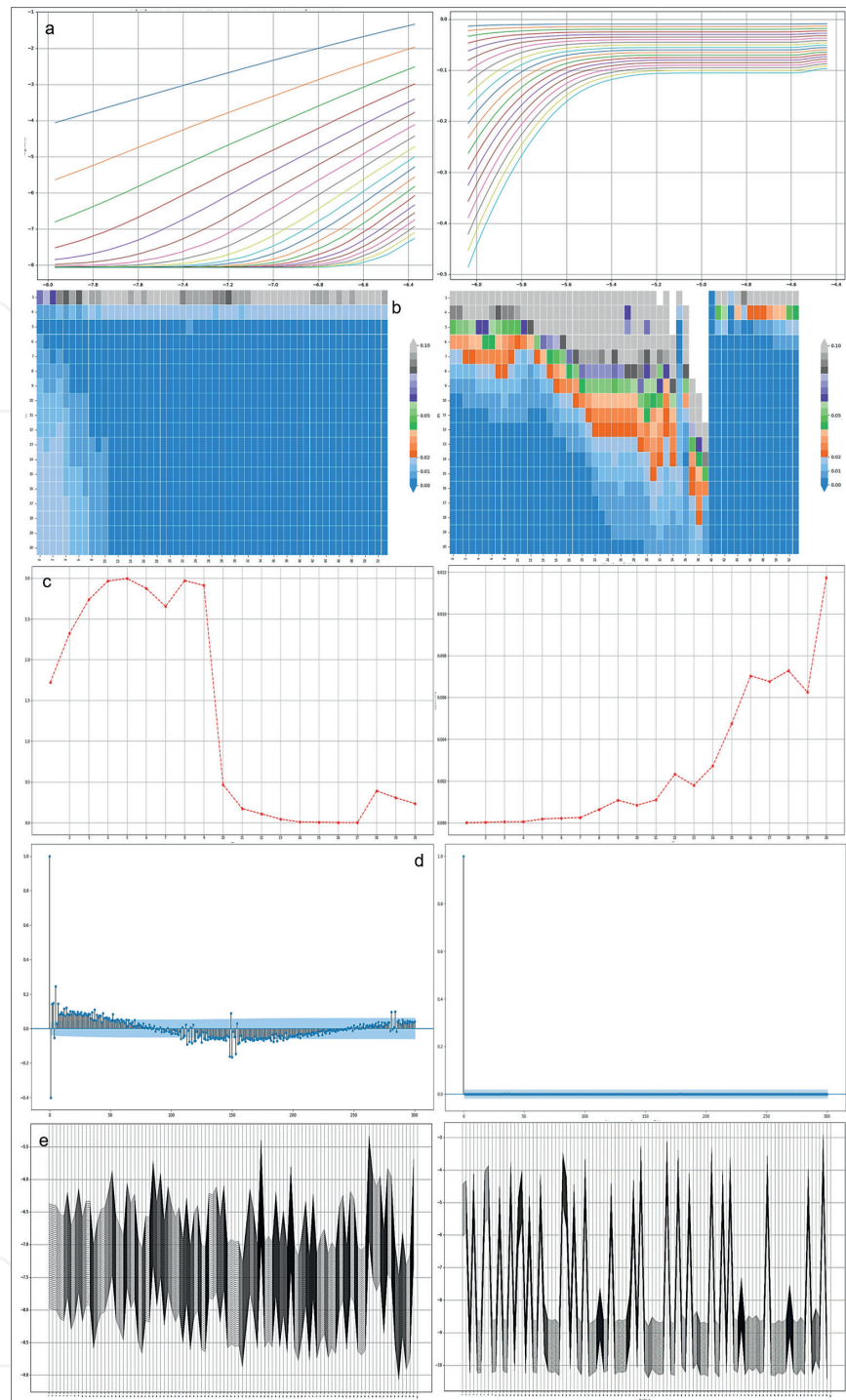


Figure 3. Results for the dataset wind speed change rate Mars (left) and S&P500 1-minute ticks (right). (a) States the correlation sum scheme, (b) the p -value heatmap, (c) the correlation dimensions, (d) the ACFs for 300 lags and (e) the space-time separation plots with 100 steps. Comparing with the generic datasets visually already reveals the conceptual differences and problems inherent in the analyzed data.

the frequency of the data samples, namely, Vogl and Rötzel [40] observed chaos in daily S&P500 returns, while in S&P500 1-minute tick return data, mean-reversion is present, leading to insights proposed by BenSaïda [39], namely, that the same system at different frequency levels, may propose different dynamics, revealing a scale dependence of the underlying empirical DGP.

Referring back to step three, namely, the attractor reconstruction, one may see various outcomes based upon false pretenses in the reconstruction results. In terms

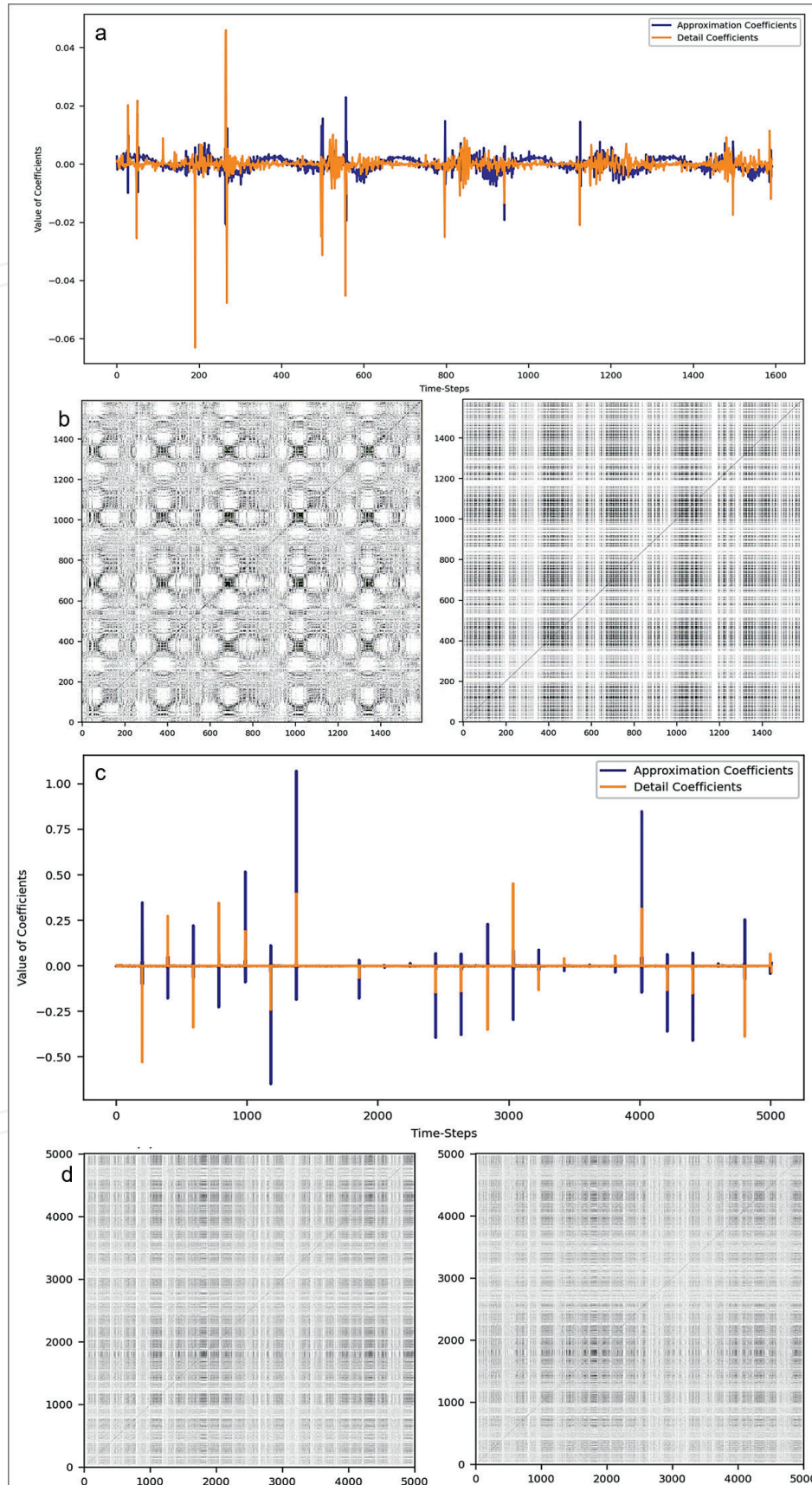


Figure 4. Overview of RQA-DWT results for the display of RPs and sub-dynamics. (a) Is the DWT for the wind speed change rates of Mars and (c) for the SeP500 1-minute return ticks. In addition, (b) represents the approximation (left) and detail (right) coefficients for (a), while (d) represents the same for (c). It is denotable that both series consists of sub-dynamics. Note that the Mars detail coefficients may resemble a hidden chaotic subsystem, which can be separately analyzed.

Parameters	Generic	Mars	S&P500	Implication
RQA	strong determinism, chaotic	determinism	low determinism, stochastic	RQA measures versus comparative data provide characteristics for quantitative modeling
Sub-Dynamics	no	yes	yes	DWT with one iteration based upon a 'Bior 2.8' wavelet filter bank, the resulting high and low pass data are inserted into RQA
Frequency	yes, distinct	no	yes, recurrent	Based upon a CWT with a 'Shannon' wavelet, 1024 scales
Multifractal	yes, clear	scale-dependent	no	Based upon a MFDFA analysis
Power law	yes	nested exponential with fat tails	no	Based upon CCDFs versus theoretical power-law backed with coherence tests on 1% significance
Result	Chaotic, Multifractal System	Stochastic, Deterministic System	Mean Reverting system	Very careful interpretation advisable

Table 4.
 Overview of steps four to six for all datasets with implications.

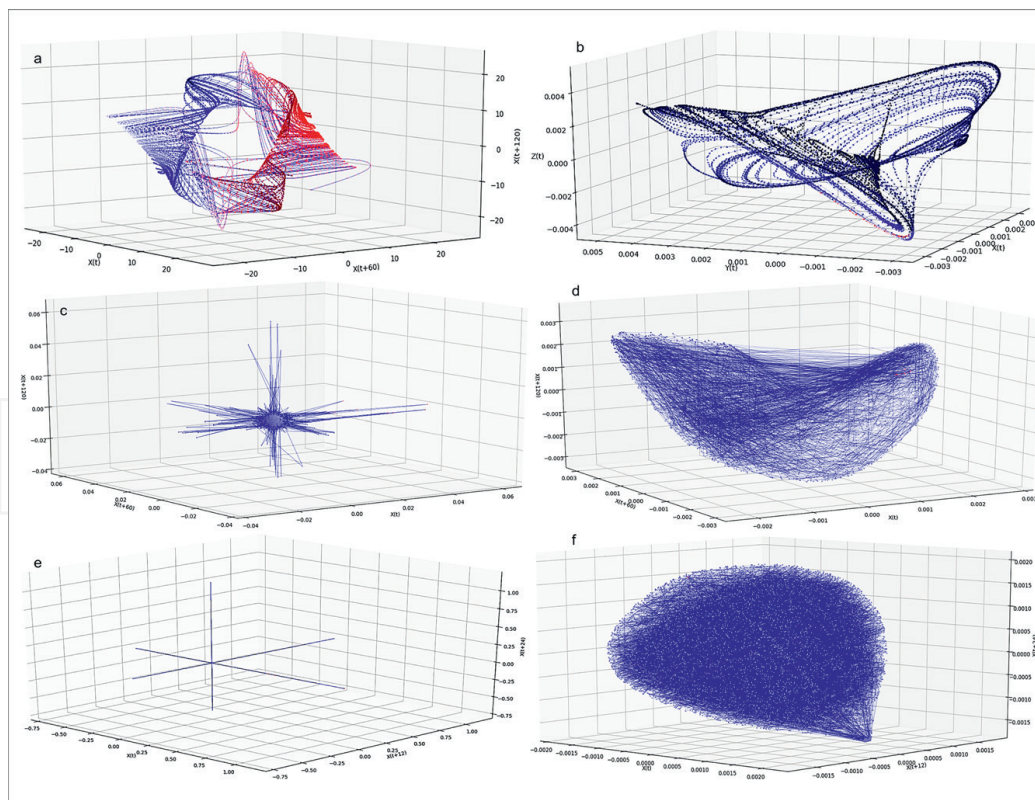


Figure 5.
 Display of attractor reconstructions for the generic dataset based upon a Lorenz set (a, b), the wind speed change rates of Mars (c, d) and the S&P500 1-minute return ticks (e, f). (a), (c) and (e) represent the Takens delay-time embedding, yet, (c) and (e) are proven to be not reconstructable. (b), (d) and (f) display the spectral embedding in combination with a k-NN algorithm and a PCA with Laplacian Eigenmaps. As with (c) and (e), note that the analysis shows that (d) and (f) are not to be reconstructable. It is visible that a violation of reconstruction prerequisites results in very poor reconstructions since those are not to be conducted in the first place.

of extreme high ACFs or temporal correlations, the attractor is dispersed and flattened, while a lack of scaling characteristics results in singular ‘spaghetti like’ lines. Furthermore, as stated in Nichols and Nichols [87], a deflation or stretching on the 45° line of the 3D space is also possible. Note that in the case of a linear system such as the S&P500 1-minute tick returns, the Takens embedding only states straight lines, which clearly indicates the absence of nonlinearity. A proper reconstruction shows a closed and dense area and visible attractor-like structures. For reference, as stated in Vogl and Rötzel [40], a pure stochastic system such as a Wiener process will end up resembling a ‘ball’ with no trajectory structure. In regards to time-series data with higher dimensional estimates, exceeding 3D spaces, the reconstructed graphical display may appear ‘deformed’, owing to the lack of degrees of freedom in the visual display. On a final note, the exertion of particular care regarding the prerequisites of the reconstruction is highly advised, since violations result in poor representations and false characteristics, which will build the groundwork for subsequent quantitative modeling attempts. Furthermore, it is advisable to alter the delay-times and dimension estimates in several iterations to be sure to hit the most ‘representable’ form of the time-series system under analysis, especially, in more complex applications such as spectral embedding. Finally, the framework only provides the most basic intuitions or the minimum set of knowledge for analysis to be possible at all, refinements are always encouraged. Taken together, the stated insights can be abstracted into a minimum set of requirements, which have to be fulfilled by potential model selections. Furthermore, one may reapply the whole analysis on the DWT sub-dynamics series to elaborate on potential hidden (strange) attractors.

5. Concluding remarks

Within this chapter, a practical guideline for the complete implementation of a combinatory, chaos analysis framework separately proposed in Vogl and Rötzel [40] for stationary and in Vogl [44] for non-stationary data is presented in its entirety. The framework is proposed as an integrated, holistic approach to analyzing the empirical DGP of nonlinear time-series data and provides the possibility to distinguish chaoticity from stochasticity while referring to underlying evolutionary dynamics. The analysis steps are elucidated, potential pitfalls and theoretical rationales stated and prerequisites discussed in detail. Moreover, an ‘idealistic’ versus ‘negative’ case is empirically and graphically introduced and debated based upon real-world time-series sets¹⁵. With this guide, the reader should be able to conduct the analysis themselves, without being prone to misspecifications and common errors present in the scientific literature.

Lastly, concluding remarks and current frontiers in the elaborated context are briefly to be stated. Current gaps in research and frontiers on the reconstruction of attractors is vastly seen in the application of neural network, evolutionary algorithms and other reconstruction methodologies to obtain sufficient and high-quality reconstructions and analysis insights (see [116, 117]). Nonetheless, the research field of time-series reconstruction and quantification of empirical DGPs is scarce and defined as a current gap in research, particularly, in hindsight of novel technological advancements such as artificial intelligence solutions. To conclude, Nieto et al. [98], states

¹⁵ Even if not displayed in this chapter, during the preparation period, several different time-series have been analyzed, e.g. flood and river discharge series, wind power, energy prices, tweet-frequencies, nonlinear fluids and fundamental economic indicators, among others.

unpredictability as a ‘fundamental topic’ in the nonlinear scientific domain, owing to its consequences being rooted in the existence of sensitivity to initial conditions as the main trait of chaotic dynamics. Furthermore, no common understanding of unpredictability exists, since differing definitions may be applicable, e.g. problems in predicting trajectory evolutions may not be seen as a problem in hindsight of scattering problems, which only level around asymptotic behavior, thus, define problems only in the prediction of final system states [98]. Furthermore, predictability in subsequently implemented models is a vast topic, which is neglected for the discussion of this chapter, yet, deemed of utmost relevance to it.


The stated framework can be enhanced further and shows several limitations, namely, it is computationally expensive and consists of many algorithms and methods, which are time intensive. Moreover, the selected methods are chosen due to their vast application in the scientific literature and not on performance. Hence, no optimization has been conducted yet, owing to the goal dependence of the analysis framework, even if applicability to various time-series is given. Furthermore, there exists no way to resolve attractor reconstructions given the existence of high ACFs and high or oscillating temporal correlations. Moreover, the framework is graphically reliant, which is seen as a potential hindrance in terms of future automatization and application on larger data pools and automated decision rule generations. Nonetheless, to conclude, the presented framework is seen as the fundamental basis or minimal building block for future research, i.e. as the provision of a stepping stone toward more advanced, transparent and reliable insights originating from the scientific nonlinear dynamics community. The enablement to safely distinguish chaoticity from stochasticity paired with the detailed characterization of the empirical time-series DGP, in general, is expected to have a positive influence on the quantification, modeling and the future prospects of the field, solving a 40-year-old debate. Resolving stated debate, hopefully, opens the way to more coherent insights and persistent knowledge about time-series systems and the quantification of the real-world in various disciplines such as medicine, hydrology, economics and physics. The inherent paradigm shift is also expected to make model selection easier and more self-explanatory in the future of time-series predictions.

Author details

Markus Vogl
University of Applied Sciences Aschaffenburg (TH Aschaffenburg), Wiesbaden,
Germany

*Address all correspondence to: markus.vogl@vogl-datascience.de

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Gaul V and Kumar V. Allied Market Research. 2020. [Online]. Available: <https://www.alliedmarketresearch.com/predictive-analytics-market>. [Accessed May 27, 2022].
- [2] Markets and Markets. Markets and Markets. 2021. [Online]. Available: <https://www.marketsandmarkets.com/Market-Reports/predictive-analytics-market-1181.html>. [Accessed May 27, 2022].
- [3] Aguilar-Rivera R, Valenzuela-Rendón M, Rodríguez-Ortiz J. Genetic algorithms and Darwinian approaches in financial applications: A survey. *Expert Systems with Applications*. 2015;**42**:7684-7697
- [4] Poornima S, Pushpalatha M. A survey on various applications of prescriptive analytics. *International Journal of Intelligent Networks*. 2020;**1**:76-84
- [5] Dudkowski D, Jafari S, Kapitaniak T, Kuznetsov N, Lenonov G, Prasad A. Hidden attractors in dynamical systems. *Physics Reports*. 2016;**637**:1-50
- [6] Marwan N, Romano MC, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. *Physics Reports*. 2007;**438**:237-329
- [7] Marwan N, Wessel N, Meyerfeldt A, Schirdewan A, Kurths J. Recurrence plot based measures of complexity and its application to heart rate variability data. *Physical Reviews E*. 2002;**66**(2):026702
- [8] Marwan N, Kurths J. Line structures in recurrence plots. *Physical Letters A*. 2005;**336**(4-5):349-357
- [9] Guégan D, Leroux J. Forecasting chaotic systems: The role of local Lyapunov exponents. *Chaos, Solitons and Fractals*. 2009;**41**:2401-2404
- [10] Song X, Niu D, Zhang Y. The chaotic attractor analysis of DJIA based on manifold embedding and Laplacian Eigenmaps. *Mathematical Problems in Engineering*. 2016;**4**:1-10
- [11] Fernández-Rodríguez F, Sosvilla-Rivero S, Andrada-Félix J. Testing chaotic dynamics via Lyapunov exponents. *Journal of Applied Econometrics*. 2005;**20**:911-930
- [12] Adams Z, Füss R, Glück T. Are correlations constant? Empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*. 2017;**84**:9-24
- [13] Alexandridis AK, Kampouridis M, Cramer S. A comparison of wavelet networks and genetic programming in the context of temperature derivatives. *International Journal of Forecasting*. 2017;**33**:21-47
- [14] Shi Y, Ho K-Y. Long memory and regime switching: A simulation study on the Markov regime-switching ARFIMA model. *Journal of Banking & Finance*. 2015;**61**:189-204
- [15] Kristoufek L. Multifractal height cross-correlation analysis: A new method for analyzing long-range cross correlation. *Europhysics Letters*. 2011;**95**(6):68001
- [16] Berghorn W. Trend Momentum. *Quantitative Finance*. 2015;**15**:261-284
- [17] Ramiah V, Xu X, Moosa IA. Neoclassical finance, behavioural finance and noise traders: A review and assessment of the literature. *International Review of Financial Analysis*. 2015;**41**:89-100
- [18] Celeste V, Corbet S, Gurdgiev C. Fractal dynamics and wavelet analysis:

Deep volatility and return properties of bitcoin, Ethereum and ripple. The Quarterly Review of Economics and Finance. 2019;**76**:310-324

[19] De Luca G, Dominique G, Giorgia R. Assessing tail risk for nonlinear dependence of MSCI sector indices: A copula three-stage approach. Finance Research Letters. 2019;**30**:327-333

[20] Beltratti A, Stulz RM. Why is contagion asymmetric during the European sovereign crisis? Journal of International Money and Finance. 2019;(99-C):102081

[21] Charfeddine L. True or spurious long memory in volatility: Further evidence on the energy futures markets. Energy Policy. 2014;(71-C):76-93

[22] Barkoulas JT, Chakraborty A, Ouandlous A. A metric and topological analysis of determinism in the crude oil spot market. Energy Economics. 2012;**34**:584-591

[23] Matilla-García M, Queralt R, Sanz P, Vázquez F. A generalized BDS statistic. Computational Economics. 2004;**24**:277-300

[24] Sandubete JE, Escot L. Chaotic signals inside some tick-by-tick financial time series. Chaos, Solitons and Fractals. 2020;**137**:109852

[25] Vogl M. Controversy in financial chaos research and nonlinear dynamics: A short literature review. Chaos, Solitons and Fractals. 2022;**162**:112444

[26] Çoban G, Büyüklü AH. Deterministic flow in phase space of exchange rates: Evidence of chaos in filtered series of Turkish lira-Dollar daily growth rates. Chaos, Solitons and Fractals. 2009;**42**(2):1062-1067

[27] Eckmann J, Ruelle D. Ergodic theory of chaos and strange attractors.

Reviews of Modern Physics. 1985;**57**(3):617-656

[28] Devaney R. An Introduction to Chaotic Dynamical Systems. Cambridge: Addison Wesley; 1989

[29] BenSaïda A, Litimi H. High level chaos in the exchange and index markets. Chaos, Solitons and Fractals. 2013;**54**:90-95

[30] Abarbanel H, Brown R, Sidorowich J, Tsimring L. The analysis of observed chaotic data in physical systems. Reviews of Modern Physics. 1993;**65**:1331

[31] Fuh C-C, Tsai H-H, Yao W-H. Combining a feedback linearization controller with a disturbance observer to control a chaotic system under external excitation. Communications in Nonlinear Science and Numerical Simulation. 2012;**17**:1423-1429

[32] Sornette D. Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools. Heidelberg: Springer Verlag; 2004

[33] Rössler O. An equation for hyperchaos. Physics Letters A. 1979;**71**:155-157

[34] Ma C, Wang X. Hopf bifurcation and topological horseshoe of a novel finance chaotic system. Communications in Nonlinear Science and Numerical Simulation. 2012;**17**:721-730

[35] Gao Q, Ma J. Chaos and Hopf bifurcation of a finance system. Nonlinear Dynamics. 2009;**58**:209

[36] Dechert WD, Gençay. The topological invariance of Lyapunov exponents in embedded dynamics. Physica D. 1996;**90**:40-55

[37] Jahanshahi H, Yousefpour A, Wei Z, Alcaraz R, Bekiros S. A financial

- hyperchaotic system with coexisting attractors: Dynamic investigation, entropy analysis, control and synchronization. *Chaos, Solitons and Fractals*. 2019a;**126**:66-77
- [38] Brock W, Dechert W, Scheinkman J, LeBaron B. A test for independence based on the correlation dimension. *Econometric Reviews*. 1996;**15**:197-235
- [39] BenSaïda. Noisy chaos in intraday financial data: Evidence from the American index. *Applied Mathematics and Computation*. 2014;**226**:258-265
- [40] Vogl M, Rötzel PG. Chaoticity versus stochasticity in financial markets: Are Daily S&P 500 return dynamics chaotic? *Communications in Nonlinear Science and Numerical Simulation*. 2022;**108**:106218
- [41] Aguirre LA, Billings S. Identification of models for chaotic systems from noisy data: Implications for performance and nonlinear filtering. *Physica D*. 1995;**85**:239-258
- [42] Kyrtsou C, Labys WC, Terraza M. Noisy chaotic dynamics in commodity markets. *Empirical Economics*. 2004;**29**:489-502
- [43] Kostelich EJ. The analysis of chaotic time-series data. *Systems & Control Letters*. 1997;**31**:313-319
- [44] Vogl M. Hurst Exponent Dynamics of S&P 500 Returns: Implications for Market Efficiency, Long Memory, Multifractality and Financial Crises Predictability by Application of a Nonlinear Dynamics Analysis Framework. Working Paper SSRN, Under Review. 2022. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3838850
- [45] Kantz H, Schreiber T. *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press; 2003
- [46] Grassberger P, Procaccia I. Characterization of strange attractors. *Physica Review Letters*. 1983a;**50**:346-394
- [47] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. *Physica 9D*. 1983;**9**(1-2):189-208
- [48] Takens F. Detecting strange attractors in fluid turbulence. In: Rand D, Young L-S, editors. *Dynamical Systems and Turbulence*. Berlin: Springer; 1981. pp. 366-381
- [49] Cencini M, Cecconi E, Vulpiani A. *Chaos from simple models to complex systems*. World Scientific. 2010;**17**:1-480
- [50] Tirandaz H, Aminabadi S, Tavakoli H. Chaos synchronization and parameter identification of a finance chaotic system with unknown parameters, a linear feedback controller. *Alexandria Engineering Journal*. 2018;**57**:1519-1524
- [51] Richman J, Moorman J. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology - Heart and Circulatory Physiology*. 2000;**278**(6):H2039-H2049
- [52] Park JY, Whang Y-J. Random walk or chaos: A formal test on the Lyapunov exponent. *Journal of Econometrics*. 2012;**169**:61-74
- [53] Shevchenko II. Lyapunov and diffusion timescales in the solar neighborhood. Working Paper with arXiv-ID: 1012. 2018;**3606v2**:1-22
- [54] Hurst H. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*. 1951;**116**:770
- [55] Sundararajan D. *Discrete Wavelet Transform - a Signal Processing Approach*, India. Singapore: John Wiley & Sons; 2015

- [56] Chen Y, Yang H. Multiscale recurrence analysis of long-term nonlinear and nonstationary time series. *Chaos, Solitons and Fractals*. 2012;**45**(7):978-987
- [57] Mallat S. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1989;**11**(7):674-693
- [58] Mitra S. A wavelet filtering based analysis of macroeconomic indicators: The Indian evidence. *Applied Mathematics and Computation*. 2006;**175**:1055-1079
- [59] López de Prado M. *Advances in Financial Machine Learning*. Hoboken: John Wiley & Sons Inc.; 2018
- [60] Fan Q, Liu S, Wang K. Multiscale multifractal detrended fluctuation analysis of multivariate time series. *Physica A*. 2019;**532**:121864
- [61] Alstott J, Bullmore E, Plenz D. Powerlaw: A python package for analysis of heavy-tailed distributions. *PloSONE*. 2014;**9**(4):e95816
- [62] Bao D, Yang Z. Intelligent stock trading system by turning point confirming and probabilistic reasoning. *Expert Systems with Applications*. 2008;**34**:620-627
- [63] Altan A, Karasu S, Bekiros S. Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos, Solitons and Fractals*. 2019;**126**:325-336
- [64] Argyris J, Andreadis I. The influence of noise on the correlation dimension of chaotic attractors. *Chaos, Solitons and Fractals*. 1998;**9**(3):343-361
- [65] Donoho D. Denoising by soft threshold. *IEEE Trans on Information Theory*. 1995;**41**(3):613-627
- [66] Donoho D, Johnstone I. Ideal spatial adaption by wavelet shrinkage. *Biometrika*. 1994;**81**(3):425-455
- [67] MacKinnon J. Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business and Economic Statistics*. 1994;**12**:167-176
- [68] Kwiatkowski D, Phillips P, Schmidt P, Shin Y. Testing the null hypothesis of stationary against the alternative of a unit root. *Journal of Econometrics*. 1992;**54**:159-178
- [69] Massey FJ Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*. 2012;**46**(253):68-78
- [70] Andrews D. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*. 1991;**59**:817-858
- [71] Provenzale A, Smith L, Vio R, Murante G. Distinguishing between low-dimensional dynamics and randomness in measured time series. *Physica D*. 1992;**58**:31-49
- [72] Ramsey J, Sayers C, Rothman P. The statistical properties of dimension calculations using small data sets: Some economic applications. *International Economic Review*. 1990;**4**:991-1020
- [73] Bajo-Rubio O, Fernandez-Rodriguez F, Sosvilla-Rivero S. Chaotic behaviour in exchange-rate series: First results for the peseta-U.S. Dollar case. *Economics Letters*. 1992;**39**:207-211
- [74] Darbyshire A, Broomhead D. Robust estimation of tangent maps and Lyapunov spectra. *Physica D*. 1996;**89**:287
- [75] Rosenstein M, Collins J, De Luca C. A practical method for calculating largest

- Lyapunov exponents from small data sets. *Physica D*. 1993;**65**:117-134
- [76] Bask M, Gençay R. Testing chaotic dynamics via Lyapunov exponents. *Physica D*. 1998;**114**:1-2
- [77] Peng C-K, Buldyrev S, Havlin S, Simons M, Stanley H, Goldberger L. Mosaic organization of DNA nucleotides. *Physical Review E*. 1994;**49**(2):1685
- [78] Hardstone R, Poil S-S, Schiavone G, Jansen R, Nikulin V, Mansvelder H, et al. Detrended fluctuation analysis: A scale-free view on neuronal oscillations. *Frontiers in Physiology*. 2012;**3**:450
- [79] Mandelbrot BB, van Ness J. Fractional Brownian motions, fractional noises and applications. *SIAM Review*. 1968;**10**(4):422-437
- [80] Mandelbrot BB, Wallis JR. Some long-run properties of geophysical records. *Water Resources Research*. 1969;**5**(2):321-340
- [81] Opong K, Mulholland G, Fox A, Farahmand K. The behaviour of some UK equity indices: An application of Hurst and BDS tests. *Journal of Empirical Finance*. 1999;**6**(3):267-282
- [82] Mandelbrot BB. *Fractals and Chaos*. New York: Springer; 2004
- [83] Norouzzadeh P, Jafari G. Application of multifractal measures to Teheran price index. *Physica A*. 2005;**356**:609-627
- [84] Yule G. On a method of investigating periodicities in disturbed series with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London Series A*. 1927;**226**:267-298
- [85] Packard N, Crutchfield J, Farmer J, Shaw R. Geometry from a time series. *Physical Review Letters*. 1980;**45**:712-716
- [86] Ruelle D. *Chaotic Evolution and Strange Attractors*. Cambridge: Cambridge University Press; 1989
- [87] Nichols J, Nichols J. Attractor reconstruction for non-linear systems: A methodological note. *Mathematical Biosciences*. 2001;**171**:21-32
- [88] Gibson J, Farmer J, Casdagli M, Eubank S. An analytic approach to practical state space reconstruction. *Physica D*. 1992;**57**:1-30
- [89] Cao L. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D*. 1997;**110**(1-2):43-50
- [90] Cannas B, Cincotti S. Neural reconstruction of Lorenz attractors by an observable. *Chaos, Solitons and Fractals*. 2002;**14**:81-86
- [91] Toledo-Suárez C. Meta-chaos: Reconstructing chaotic attractors from the separation of nearby initial conditions on hyperhelices. *Communications in Nonlinear Dynamics and Numerical Simulation*. 2010;**15**:2249-2253
- [92] Yeo K. Data-driven reconstruction of nonlinear dynamics from sparse observation. *Journal of Computational Physics*. 2019;**395**:671-689
- [93] Asefa T, Kembrowski M, Lall U, Urroz G. Support vector machines for nonlinear state space reconstruction: Application to the great salt Lake time series. *Water Resources Research*. 2005;**41**:W12422
- [94] Ma H-G, Zhang C-L, Li F. State space reconstruction of nonstationary time-series. *Journal of Computational*

and Nonlinear Dynamics.
2017;12:031009-031001

[95] Fraser A, Swinney H. Independent coordinates for strange attractors from mutual information. *Physical Review A*. 1986;33(2):1134-1140

[96] Broomhead D, King G. Extracting qualitative dynamics from experimental data. *Physica D*. 1986;20:217-236

[97] Rüdüsüli M, Schildhauer T, Biollaz S, Van Ommen J. Measurement, monitoring and control of fluidized bed combustion and gasification. In: *Fluidized Bed Technologies for Near-Zero Emission Combustion and Gasification*. UK: Woodhead Publishing; 2013. pp. 813-864

[98] Nieto A, Seoane J, Sanjuán M. Final state sensitivity in noisy chaotic scattering. *Chaos, Solitons and Fractals*. 2021;150:111181

[99] Sauer T, Yorke J, Casdagli M. Embedology. *Journal of Statistical Physics*. 1991;65:579

[100] Pentari A, Tzagkarakis G, Tsakalides P, Simos P, Bertsiadis G, Kavroulakis E, et al. Changes in resting-state functional connectivity in neuropsychiatric lupus: A dynamic approach based on recurrence quantification analysis. *Biomedical Signal Processing and Control*. 2022;72:103285

[101] Eckmann J-P, Kamphorst S, Ruelle D. Recurrence plots of dynamical systems. *Europhysics Letters*. 1987;5:973-977

[102] Van den Hoorn W, Hodges P, van Dieen J, Kerr G. Reliability of recurrence quantification analysis of postural sway data. A comparison of two methods to determine recurrence threshold. *Journal of Biomechanics*. 2020;107:109793

[103] Koebbe M, Mayer-Kress G. Use of recurrence plots in the analysis of time-series data. In: Casdagli M, Eubank S, editors. *Proceedings of SFI Studies in the Science of Complexity*, Vol. XXI, Redwood City, 1992. Reading, MA: Addison-Wesley; 1992. pp. 361-378

[104] Zbilut J, Webber C Jr. Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters A*. 1992;171(3-4):199-203

[105] Theiler J. Spurious dimension from correlation algorithms applied to limited time-series data. *Physical Reviews A*. 1986;34(3):2427-2432

[106] Cariolaro G. *Unified Signal Theory*. London: Springer; 2011

[107] Wojtaszczyk P. *A Mathematical Introduction to Wavelets*. Cambridge: Cambridge University Press; 1997

[108] Chang S, Grace S, Yu B, Vetterli M. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions*. 2000;9(9):1532-1546

[109] Strogatz S. *Nonlinear Dynamics and Chaos*. Colorado: Westview Press; 2014

[110] Katok A, Hasselblatt B. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge: Cambridge University Press; 1995

[111] Mandelbrot BB. *The Fractal Geometry of Nature*. USA: Freeman; 1977

[112] Cao G, He L-Y, Cao J. *Multifractal Detrended Analysis Method and its Application in Financial Markets*. Singapore: Springer; 2018

[113] Barunik J, Aste T, Di Matteo T, Liu R. Understanding the source of multifractality in financial markets. *Physica A*. 2012;391:4234-4251

[114] Yuan Y, Zhuang X-T, Jin X. Measuring multifractality of stock price fluctuation using multifractal detrended fluctuation analysis. *Physica A*. 2009;**388**:2189-2197

[115] Lorenz E. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*. 1963;**20**:130-141

[116] Zelinka I, Chadli M, Davendra D, Senkerik R, Jasek R. An investigation on evolutionary reconstruction of continuous chaotic systems. *Mathematical and Computer Modelling*. 2013;**57**:2-15

[117] Jokar M, Salarieh H, Alasty A. On the existence of proper stochastic Markov models for statistical reconstruction and prediction of chaotic time series. *Chaos, Solitons and Fractals*. 2019;**123**:373-382