


Article

Nonlinear Time Series Modeling: A Unified Perspective, Algorithm and Application

Subhadeep Mukhopadhyay ^{1,*}  and Emanuel Parzen ^{2,†}

¹ Department of Statistical Science, Temple University, Philadelphia, PA 19122, USA

² Department of Statistics, Texas A&M University, College Station, TX 77843, USA; eparzen@tamu.edu

* Correspondence: deep@temple.edu

† Shortly after finishing the first draft of this paper, Manny Parzen passed away. Deceased 6 February 2016.

Received: 4 June 2018; Accepted: 3 July 2018; Published: 6 July 2018



Abstract: A new comprehensive approach to nonlinear time series analysis and modeling is developed in the present paper. We introduce novel data-specific mid-distribution-based Legendre Polynomial (LP)-like nonlinear transformations of the original time series $\{Y(t)\}$ that enable us to adapt all the existing stationary linear Gaussian time series modeling strategies and make them applicable to non-Gaussian and nonlinear processes in a robust fashion. The emphasis of the present paper is on empirical time series modeling via the algorithm LPTIME. We demonstrate the effectiveness of our theoretical framework using daily S&P 500 return data between 2 January 1963 and 31 December 2009. Our proposed LPTIME algorithm systematically discovers all the ‘stylized facts’ of the financial time series automatically, all at once, which were previously noted by many researchers one at a time.

Keywords: nonparametric time series modeling; nonlinearity; unified time series algorithm; exploratory diagnostics

1. Introduction

When one observes a sample $Y(t), t = 1, \dots, T$, of a (discrete parameter) time series $Y(t)$, one seeks to nonparametrically learn from the data a stochastic model with two purposes: (a1) scientific understanding; (a2) forecasting (predict future values of the time series under the assumption that the future obeys the same laws as the past). Our prime focus in this paper is on developing a nonparametric empirical modeling technique for nonlinear (stationary) time series that can be used by data scientists as a practical tool for obtaining insights into (i) the temporal dynamic patterns and (ii) the internal data generating mechanism; a crucial step for achieving (a1) and (a2).

Under the assumption that the time series is stationary (which can be extended to asymptotically stationary), the distribution of $Y(t)$ is identical for all t , and the joint distribution of $Y(t)$ and $Y(t+h)$ depends only on lag h . Typical estimation goals are as follows:

- (1) **Marginal modeling:** The identification of the marginal probability law (in particular, the heavy-tailed marginal densities) of a time series plays a vital role in financial econometrics. Notations: common quantile Q , inverse of the distribution function F , respectively denoted $Q(u; Y), 0 < u < 1$ and $F(y; Y)$. The mid-distribution is defined as $F^{\text{mid}}(y; Y) = F(y; Y) - 0.5 \Pr(Y(t) = y)$.
- (2) **Correlation modeling:** Covariance function (defined for positive and negative lag h) $R(h; Y) = \text{Cov}[Y(t), Y(t+h)]$. $R(0; Y) = \text{Var}[Y(t)]$, $\mu = \mathbb{E}[Y(t)]$ assumed zero in our prediction theory. Correlation function $\rho(h) = \text{Cor}[Y(t), Y(t+h)] = R(h; Y) / R(0; Y)$.

- (3) Frequency-domain modeling: When covariance is absolutely summable, define spectral density function $f(\omega; Y) = \sum R(h; Y) e^{-2\pi i \omega h}$, $-1/2 < \omega < 1/2$.
- (4) Time-domain modeling: The time domain model is a linear filter relating $Y(t)$ to white noise $\epsilon(t)$, $\mathcal{N}(0, 1)$ independent random variables. Autoregressive scheme of order m , a predominant linear time series technique for modeling conditional mean, is defined as (assuming $\mathbb{E}[Y(t)] = 0$):

$$Y(t) - a(1; m)Y(t-1) - \dots - a(m; m)Y(t-m) = \sigma_m \epsilon(t), \quad (1)$$

with the spectral density function given by:

$$f(\omega; Y) = \frac{\sigma_m^2}{|1 - \sum_{k=1}^m a(k; m) e^{2\pi i \omega k}|^2}. \quad (2)$$

To fit an AR model, compute the linear predictor of $Y(t)$ given $Y(t-j)$, $j = 1, \dots, m$ by:

$$Y^{\mu, m}[t] = \mathbb{E}[Y(t) | Y(t-1), \dots, Y(t-m)] = a(1; m)Y(t-1) + \dots + a(m; m)Y(t-m). \quad (3)$$

Verify that the prediction error $Y[t] - Y^{\mu, m}[t]$ is white noise. The best fitting AR order is identified by the Akaike criterion (AIC) (or Schwarz's criterion, BIC) as the value of m minimizes:

$$\text{AIC}(m) = 2 \log \sigma_m + 2m/n.$$

In what follows, we aim to develop a parallel modeling framework for nonlinear time series.

2. From Linear to Nonlinear Modeling

Our approach to nonlinear modeling, called LPTime¹, is via approximate calculation of conditional expectation $\mathbb{E}[Y(t) | Y(t-1), \dots, Y(t-m)]$. Because with probability one $Q(F(Y)) = Y$, one can prove that the conditional expectation of $Y(t)$ given past values $Y(t-j)$ is equal to (with probability one) the conditional expectation of $Y(t)$ given past values $F^{\text{mid}}(Y(t-j))$, which can be approximated by linear orthogonal series expansion in score functions $T_k[F^{\text{mid}}(Y(t-j))]$ constructed by Gram–Schmidt orthonormalization of powers of:

$$T_1 = \frac{F^{\text{mid}}(Y(t); Y) - 0.5}{\sigma[F^{\text{mid}}(Y(t); Y)]}, \quad (4)$$

where $\sigma[F^{\text{mid}}(Y(t); Y)]$ is the standard deviation of the mid-distribution transform random variable given by $\sqrt{(1 - \sum_y p^3(y))/12}$ and $p(y)$ denotes the probability mass function of Y . This score polynomial allows us to simultaneously tackle the discrete (say count-valued) and continuous time series. Note that for Y continuous, T_1 reduces to:

$$T_1 = \sqrt{12}(F(Y(t)) - 0.5). \quad (5)$$

and all the higher order polynomials T_j can be compactly expressed as $\text{Leg}_j[F(Y)]$, where $\text{Leg}_j(u)$, $0 < u < 1$ denotes orthonormal Legendre polynomials. It is worthwhile to note that T_j are orthonormal polynomials of mid-rank (instead of polynomials of the original y 's), which inject robustness into our

¹ The LP nomenclature: In nonparametric statistics, the letter L plays a special role to denote robust methods based on ranks and order statistics such as quantile-domain methods. With the same motivation, we use the letter L. On the other hand, P simply stands for Polynomials. Our custom-constructed basis functions are orthonormal polynomials of mid-rank transform instead of raw y -values; for more details see [Mukhopadhyay and Parzen \(2014\)](#).

analysis while allowing us to capture nonlinear patterns. Having constructed score functions of y denoted by T_j , we transform it into a unit interval by letting $y = Q(u; Y)$ and defining:

$$S_j(u; Y) = T_j[F^{\text{mid}}(Q(u; Y))], \quad T_j(y; Y) = S_j[F^{\text{mid}}(Y(t))]. \quad (6)$$

In general, our score functions are custom constructed for each distribution function F , which can be discrete or continuous.

3. Nonparametric LPTIME Analysis

Our LPTIME empirical time series modeling strategy for nonlinear modeling of a univariate time series $Y(t)$ is based on linear modeling of the multivariate time series:

$$\text{Vec}(\text{YS})(t) = [Y_{S1}(t), \dots, Y_{Sk}(t)]^T, \quad (7)$$

where $Y_{Sk}(t) = T_k[F^{\text{mid}}(Y(t))]$, our tailor-made orthonormal mid-rank-based nonlinear transformed series. We summarize below the main steps of the algorithm LPTIME. To better understand the functionality and applicability of LPTIME, we break it into several inter-connected steps, each of which highlights:

- The algorithmic modeling aspect (how it works).
- The required theoretical ideas and notions (why it works).
- The application to daily S&P 500 return data between 2 January 1963 and 31 December 2009 (empirical proof-of-work).

3.1. The Data and LP-Transformation

The data used in this paper are daily S&P 500 return data between 2 January 1963 and 31 December 2009 (defined as $\log(P_t / P_{t-1})$, where P_t is the closing price on trading day t). We begin our modeling process by transforming the given univariate time series $\{Y(t)\}$ into multiple (robust) time series by means of special data-analytic construction rules described in Equations (4)–(6) and (7). We display the original “normalized” time series $Z(Y(t)) = (Y(t) - \mathbb{E}[Y(t)]) / \sigma[Y(t)]$ and the transformed time series $Y_{S1}(t), \dots, Y_{Sk}(t)$ on a single plot.

Figure 1 shows the first look at the transformed S&P 500 return data between October 1986 and October 1988. These newly-constructed time series work as a universal preprocessor for any time series modeling in contrast with other ad hoc power transformations. In the next sections, we will describe how the temporal patterns of these multivariate LP-transformed series $\text{Vec}(\text{YS})(t) = \{Y_{S1}(t), \dots, Y_{Sk}(t)\}$ generate various insights for the time series $\{Y(t)\}$ in an organized fashion.

3.2. Marginal Modeling

Our time series modeling starts with the nonparametric identification of probability distributions.

Non-Normality Diagnosis

Does the normal probability distribution provide a good fit to the S&P 500 return data? Figure 2a clearly indicates that the distribution of daily return is certainly non-normal. At this point, the natural question is how the distribution is different from the assumed normal one? A quick insight into this question can be gained by looking at the distribution of the random variable $U = G(Y)$, called the comparison density (Mukhopadhyay 2017; Parzen 1997), given by:

$$d(u; G, F) = \frac{f(Q(u; G))}{g(Q(u; G))}, \quad 0 \leq u \leq 1, \quad (8)$$

where $Q(u; G) = \inf\{x : G(x) \geq u\}$ is the quantile function. The flat uniform shape of the estimated comparison density provides a quick graphical diagnostic to test the fit of the parametric G to the true unknown distribution F . The Legendre polynomial-based orthogonal series comparison density estimator is given by:

$$d(u; G, F) = 1 + \sum_j \text{LP}[j; G, F] \text{Leg}_j(u), \quad 0 < u < 1 \quad (9)$$

where the Fourier coefficients $\text{LP}[j; G, F] = \mathbb{E}[\text{Leg}_j \circ G(Y)]$.

For $G = \Phi$, Figure 2b displays the histogram of $U_i = \Phi(Y_i)$ for $i = 1, \dots, n$. The corresponding comparison density estimate $\hat{d}(u; G, F) = 1 - 0.271 \text{Leg}_2(u) - 0.021 \text{Leg}_3(u) + 0.193 \text{Leg}_4(u)$ is shown with the blue curve, which reflects the fact that the distribution of daily return (i) has a sharp peaked (inverted “U” shape) and (ii) is negatively skewed with (iii) fatter tails than the Gaussian distribution. We can carry out a similar analysis by asking whether the t-distribution with two degrees of freedom provides a better fit. Figure 2c demonstrates the full analysis, where the estimated comparison density $\hat{d}(u; G, F) = 1 - 0.492 \text{Leg}_2(u) - 0.015 \text{Leg}_3(u) + 0.084 \text{Leg}_4(u)$ indicates that (iv) the t-distribution fits the data better than normal, especially in the tails, although not a fully-adequate model.

The shape of the comparison density (along with the histogram of $U_i = G(Y_i)$, $i = 1, \dots, n$) captures and exposes the adequacy of the assumed model G for the true unknown F ; thus acting as an exploratory, as well as confirmatory tool.

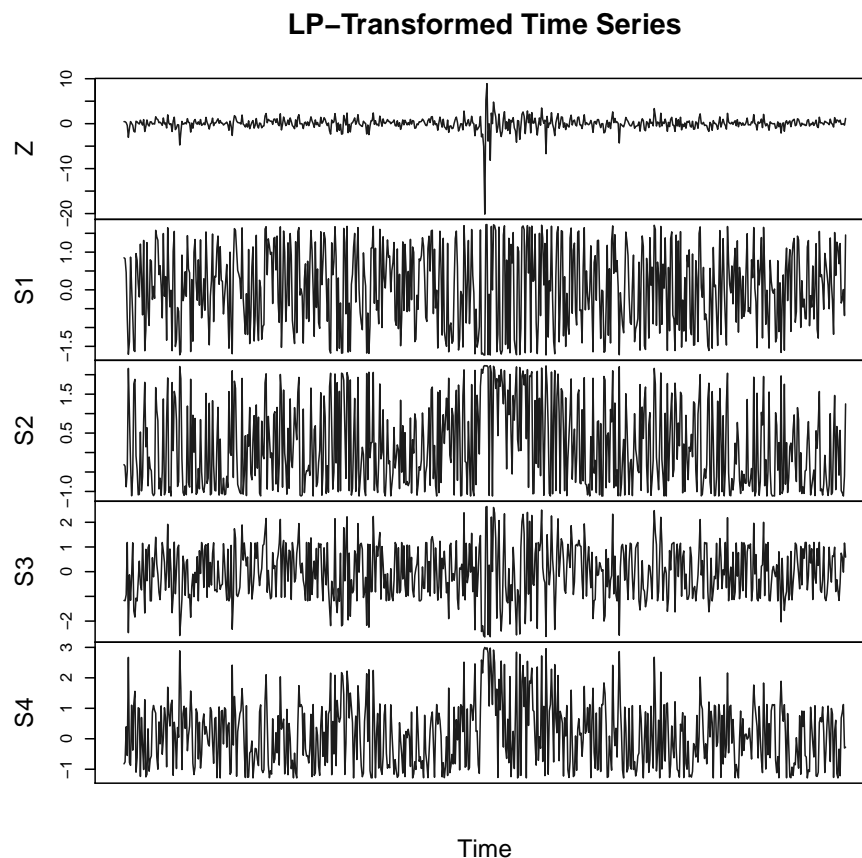


Figure 1. LP-transformed S&P 500 daily stock returns between October 1986 and October 1988. This is just a small part of the full time series from 2 January 1963–31 December 2009 (cf. Section 3.1).

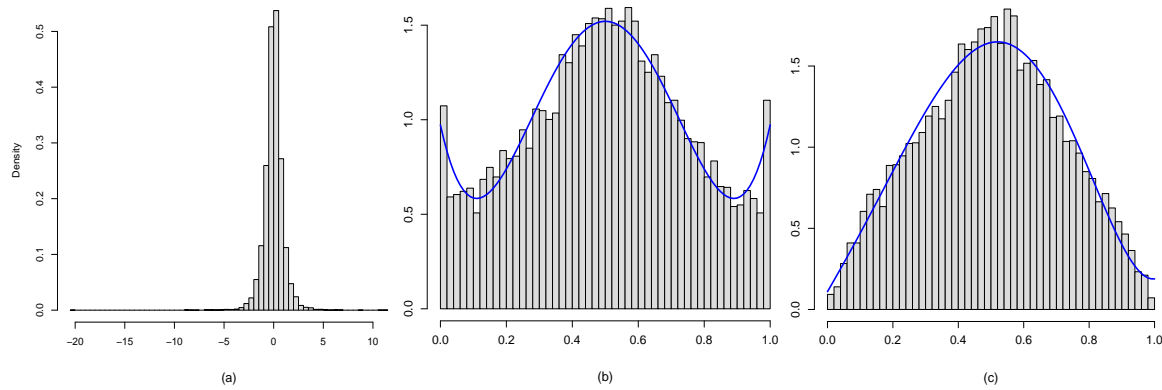


Figure 2. (a) The marginal distribution of daily returns; (b) plots the histogram of $\Phi(y_i)$ and display the LP-estimated comparison density curve. and (c) shows the associated comparison density estimate with G as t-distribution with 2 degrees of freedom.

3.3. Copula Dependence Modeling

Distinguishing uncorrelatedness and independence by properly quantifying association is an essential task in empirical nonlinear time series modeling.

3.3.1. Nonparametric Serial Copula

We display the nonparametrically-estimated smooth serial copula density $\text{cop}(u, v; Y(t), Y(t+h))$ to get a much finer understanding of the lagged interdependence structure of a stationary time series. For a continuous distribution, define the copula density for the pair $(Y(t), Y(t+h))$ as the joint density of $U = F(Y(t))$ and $V = F(Y(t+h))$, which is estimated by sample mid-distribution transform $\tilde{U} = \tilde{F}^{\text{mid}}(Y(t))$, $\tilde{V} = \tilde{F}^{\text{mid}}(Y(t+h))$. Following [Mukhopadhyay and Parzen \(2014\)](#) and [Parzen and Mukhopadhyay \(2012\)](#), we expand the copula density (square integrable) in a orthogonal series of product LP-basis functions as:

$$\text{cop}(u, v; Y(t), Y(t+h)) - 1 = \sum_{j,k} \text{LP}[j, k; Y(t), Y(t+h)] S_j(u; Y(t)) S_k(v; Y(t+h)), \quad (10)$$

where $S_j(u; Y(t)) = \text{YS}_j(Q(u; Y(t)); Y(t))$. Equation (10) allows us to pictorially represent the information present in the LP-comoment matrix via copula density. The various “shapes” of the copula density give insight into the structure and dynamics of the time series.

Now, we apply this nonparametric copula estimation theory to model the temporal dependence structure of S&P return data. The copula density estimate $\widehat{\text{cop}}(u, v; Y(t), Y(t+1))$ based on the smooth LP-comoments is displayed in Figure 3. The shape of the copula density shows strong evidence of asymmetric tail dependence. Note that the dependence is only present in the extreme quantiles, another well-known stylized fact of economic and financial time series.

3.3.2. LP-Comoment of Lag h

Here, we will introduce the concept of the LP-comoment to get a complete understanding of the nature of the serial dependence present in the data. The LP-comoment of lag h is defined as the joint covariance of $\text{Vec}(\text{YS})(t)$ and $\text{Vec}(\text{YS})(t+h)$.

The lag one LP-comoment matrix for S&P 500 return data is displayed below:

$$\text{LP}[Y(t), Y(t+1)] = \begin{bmatrix} 0.0705^* & -0.0617^* & 0.0199 & 0.0113 \\ 0.0074 & 0.1542^* & 0.0077 & 0.0652^* \\ -0.0104 & -0.0071 & 0.0262 & -0.0355 \\ 0.0166 & 0.0438^* & 0.0113 & 0.0698^* \end{bmatrix} \quad (11)$$

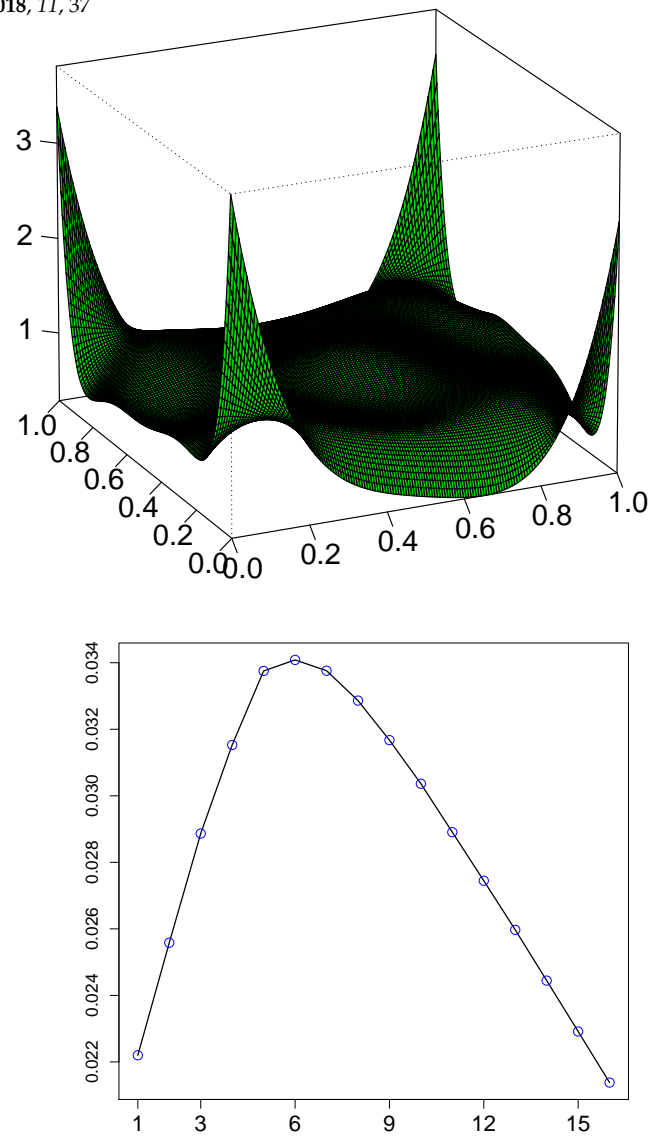


Figure 3. Top: Nonparametric smooth serial copula density (lag one) estimate of S&P return data. Bottom: BIC plot to select the significant LP-comoments computed in Equation (11).

To identify the significant elements, we first rank order the squared LP-comoments. Then, we take the penalized cumulative sum of m comoments using BIC criterion $2m \log(n)/n$, where n is the sample size, and choose the m for which the BIC is maximum. The complete BIC path for S&P 500 data is shown in Figure 3, which selects the top six comments, also denoted by $*$ in the LP-comoment matrix display (Equation (11)). By making all those uninteresting “small” comoments equal to zero, we get the “smooth” LP-comoment matrix denoted by \widehat{LP} . The linear auto-correlation is captured by the $LP[1, 1; Y(t), Y(t+1)] = \mathbb{E}[YS_1(t) YS_1(t+1)]$ term. The presence of higher order significant terms in the LP-comoment matrix indicates the possible nonlinearity. Another interesting point to note is that $CORR[Y(t), Y(t+1)] = 0.027$, whereas the auto-correlation between the mid-rank transformed data $CORR[F^{\text{mid}}(Y(t)), F^{\text{mid}}(Y(t+1))] = 0.071$, considerably larger and picked by the BIC criterion. This is an interesting fact as it indicates that the rank-transform time series ($YS_1(t)$) is much more predictable than the original raw time series $Y(t)$.

3.3.3. LP-Correlogram, Evidence and Source of Nonlinearity

We provide a nonparametric exploratory test for (non)linearity (the spectral domain test is given in Section 3.6). Plot the correlogram of $YS_1(t), \dots, YS_4(t)$: (a) diagnose possible nonlinearity; and (b) identify possible sources. This constitutes an important building block for methods of

model identification. The LP-correlogram generalizes the classical sample Autocorrelation Function (ACF). Applying the `acf()` R function on $\text{Vec}(\text{YS})(t)$ generates the graphical display of our proposed LP-correlogram plot.

Figure 4 shows the LP-correlogram of S&P stock return data. Panel A shows the absence of linear autocorrelation, which is known as an efficient market hypothesis in finance literature. A prominent auto-correlation pattern for the series $\text{YS}_2(t)$ (top right panel of Figure 4) is the source of nonlinearity. This fact is known as “volatility clustering”, which says that a large price fluctuation is more likely to be followed by large price fluctuations. Furthermore, the slow decay of the autocorrelation of the series $\text{YS}_2(t)$ can be interpreted as an indication of the long-memory volatility structure.

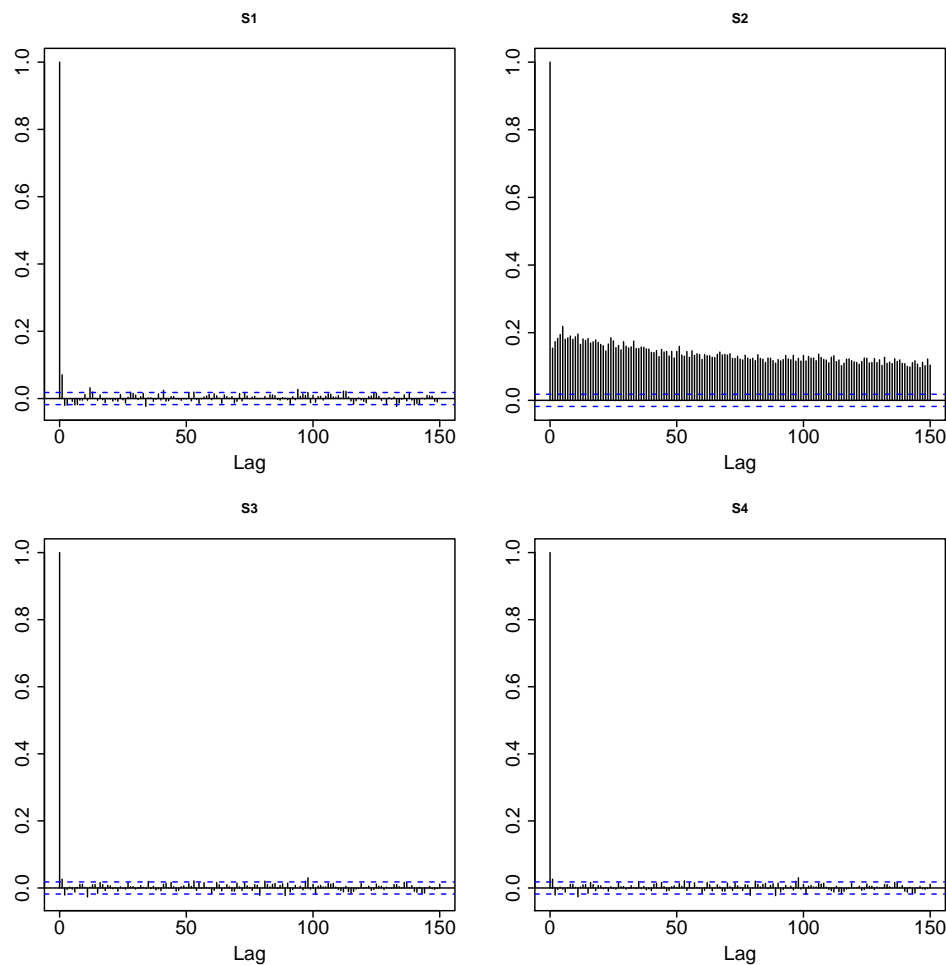


Figure 4. LP-correlogram: Sample autocorrelations of LP-transformed time series. The decay rate of the sample autocorrelations of $\text{YS}_2(t)$ appears to be much slower than the exponential decay of the ARMA process, implying possible long-memory behavior.

3.3.4. AutoLPinfor: Nonlinear Correlation Measure

We display the sample AutoLPinforplot, a diagnostic tool for nonlinear autocorrelation. We define the lag h AutoLPinfor as the squared Frobenius norm of the smooth-LP-comoment matrix of lag h ,

$$\text{AutoLPinfor}(h) = \sum_{j,k} \left| \text{LP}[j,k; Y(t), Y(t+h)] \right|^2, \quad (12)$$

where the sum is over BIC selected j, k for which LP-comoments are significantly non-zero.

Our robust nonparametric measure can be viewed as capturing the deviation of the copula density from uniformity:

$$\text{AutoLPinfor}(h) = \iint \text{cop}^2[u, v; Y(t), Y(t+h)] \, du \, dv - 1, \quad (13)$$

which is closely related to the entropy measure of association proposed in Granger and Lin (1994):

$$\text{Granger-Lin}(h) = \iint \text{cop}[u, v; Y(t), Y(t+h)] \log \text{cop}[u, v; Y(t), Y(t+h)] \, du \, dv. \quad (14)$$

It can be shown using Taylor series expansion that asymptotically:

$$\text{AutoLPinfor}(h) \approx 2 \times \text{Granger-Lin}(h). \quad (15)$$

An excellent discussion of the role of information theory methods for unified time series analysis is given in Parzen (1992) and Brillinger (2004). For an extensive survey of tests of independence for nonlinear processes, see Chapter 7.7 of Terasvirta et al. (2010). AutoLPinfor is a new information theoretic nonlinear autocorrelation measure, which detects generic association and serial dependence present in a time series. Contrast the AutoLPinfor plot for S&P 500 return data shown in Figure 5 with the ACF plot (left panel). This underlies the need for building a nonlinear time series model, which we will be discussing next.

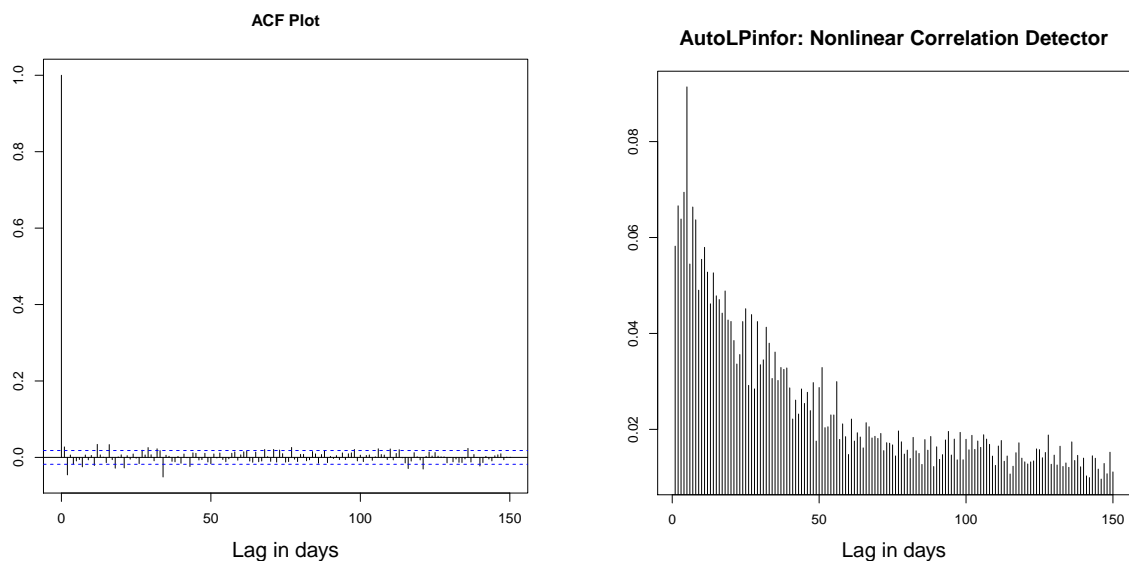


Figure 5. Left: ACF plot of S&P 500 data. Right: AutoLPinforPlot up to lag 150.

3.3.5. Nonparametric Estimation of Blomqvist's Beta

Estimate the Blomqvist's β (also known as the medial correlation coefficient) of lag h by using the LP-copula estimate in the following equation,

$$\hat{\beta}_{\text{LP}}(h; Y(t)) := -1 + 4 \int_0^{1/2} \int_0^{1/2} \widehat{\text{cop}}[u, v; Y(t), Y(t+h)] \, du \, dv \quad (16)$$

The β values $-1, 0$ and 1 are interpreted as reverse correlation, independence and perfect correlation, respectively. Note that,

- Blomqvist's β : Normalized distance of copula distribution $\text{Cop}(u, v)$ from independence copula uv
- AutoLPinfor : Distance of copula density $\text{cop}(u, v)$ from uniformity 1.

For S&P 500 return data, we compute the following dependence numbers,

$$\begin{aligned}\hat{\beta}_{LP}(1; Y(t)) &= 0.0528 \\ \hat{\beta}_{LP}(1; YS_1(t)) &= 0.0528 \\ \hat{\beta}_{LP}(1; YS_2(t)) &= 0.0729 \\ \hat{\beta}_{LP}(1; YS_3(t)) &= 0.0 \\ \hat{\beta}_{LP}(1; YS_4(t)) &= 0.003.\end{aligned}$$

3.3.6. Nonstationarity Diagnosis, LP-Comoment Approach

Viewing the time index $T = 1, \dots, n$ as the covariate, we propose a nonstationarity diagnosis based on LP-comoments of $Y(t)$ and the time index variable T . Our treatment has the ability to detect the time-varying nature of mean, variance, skewness, and so on, represented by various custom-made LP-transformed time series.

For S&P data, we computed the following LP-comoment matrix to investigate the nonstationarity:

$$LP[T, Y(t)] = \begin{bmatrix} 0.012 & 0.180^* & -0.010 & 0.058^* \\ -0.005 & -0.034 & -0.036 & 0.080^* \\ -0.016 & 0.115^* & 0.001 & -0.001 \\ 0.024 & -0.040 & -0.010 & 0.049^* \end{bmatrix} \quad (17)$$

This indicates the presence of the slight non-stationarity behavior of variance or volatility ($YS_2(t)$) and the kurtosis of tail-thickness ($YS_4(t)$). Similar to AutoLPinfor, we propose the following statistic for detecting nonstationarity:

$$LPinfor[Y(t), T] = \sum_{j,k} |LP[j, k; T, Y(t)]|^2. \quad (18)$$

We can also generate the corresponding smooth copula density of $(T, Y(t))$ based on the smooth LP $[T, Y(t)]$ matrix to visualize the time-varying information as in Figure 6.

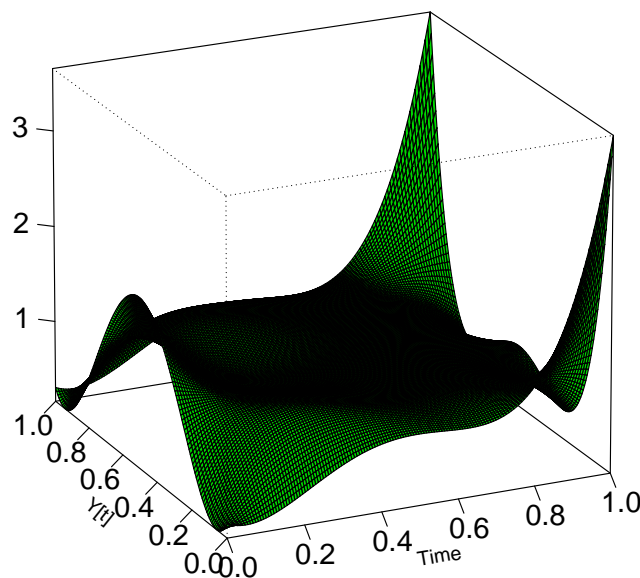


Figure 6. LP copula diagnostic for detecting non-stationarity in S&P 500 return data.

3.4. Local Dependence Modeling

3.4.1. Quantile Correlation Plot and Test for Asymmetry

We display the quantile correlation plot, a copula distribution-based graphical diagnostic to visually examine the asymmetry of dependence. The goal is to get more insight into the nature of tail-correlation.

Motivated by the concept of the lower and upper tail dependence coefficient, we define the quantile correlation function (QCF) as the following in terms of the copula distribution function of $(Y(t), Y(t+h))$ denoted by $\text{Cop}(u, v; Y(t), Y(t+h)) := \text{Cop}(u, v; h)$,

$$\lambda[u; Y(t), Y(t+h)] := \frac{\text{Cop}(u, u; h)}{u} \mathbb{I}_{\{u \leq 0.5\}} + \frac{1 - 2u + \text{Cop}(u, u; h)}{1 - u} \mathbb{I}_{\{u > 0.5\}}, \quad (19)$$

Our nonparametric estimate of the quantile correlation function is based on the LP-copula density, which we denote as $\hat{\lambda}_{\text{LP}}[u; Y(t), Y(t+h)]$. Figure 7 shows the corresponding quantile correlation plot for S&P 500 data. The dotted line represents QCF under the independence assumption. Deviation from this line helps us to better understand the nature of asymmetry. We compute $\hat{\lambda}_{\text{G}}[u; Y(t), Y(t+h)]$ using the fitted Gaussian copula:

$$\widehat{\text{Cop}}_{\text{G}}(u, v; Y(t), Y(t+h)) = \Phi(\Phi^{-1}(u), \Phi^{-1}(v); \hat{\Sigma} = S) \quad (20)$$

where S is the sample covariance matrix. The dark green line in Figure 7 shows the corresponding curve, which is almost identical to the “no dependence” curve, albeit misleading. The reason is the Gaussian copula is characterized by linear correlation, while S&P data are highly nonlinear in nature. As the linear auto-correlation of a stock return is almost zero, we have approximately $\Phi(\Phi^{-1}(u), \Phi^{-1}(u); \hat{\Sigma} = S) \approx \Phi(\Phi^{-1}(u))\Phi(\Phi^{-1}(u)) = u^2$. Similar to the Gaussian copula, there are several other parametric copula families, which can give similar misleading conclusions. This simple illustration reminds us of the pernicious effect of not “looking into the data”.

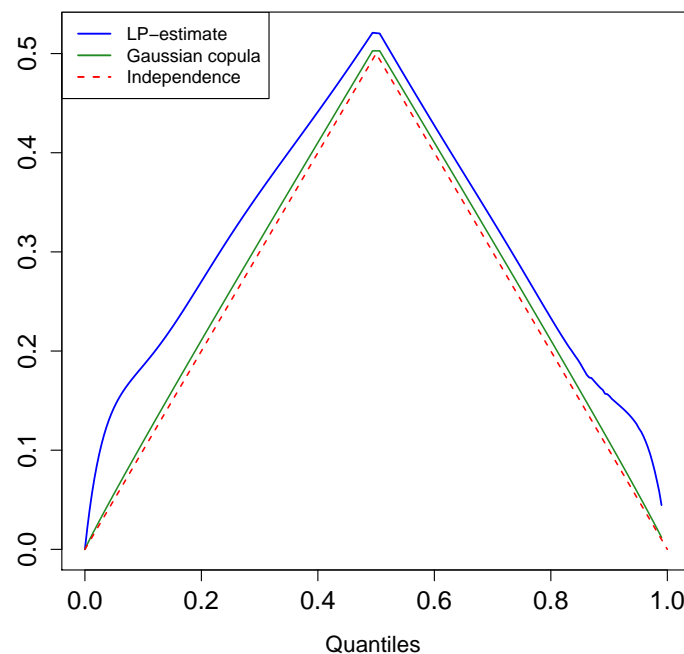


Figure 7. Estimated Quantile Correlation Function (QCF) $\hat{\lambda}_{\text{LP}}[u; Y(t), Y(t+1)]$. It detects asymmetry in the tail dependence between the lower-left quadrant and upper-right quadrant for S&P 500 return data. The red dotted line denotes the quantile correlation function under dependence. The dark green line shows the quantile correlation curve for the fitted Gaussian copula.

3.4.2. Conditional LPinfor Dependence Measure

For more transparent and clear insight into the asymmetric nature of the tail dependence, we need to introduce the concept of conditional dependence. In what follows, we propose a conditional LPinfor function $\text{LPinfor}(Y(t+h)|Y(t) = Q(u; Y(t)))$, a quantile-based diagnostic for tracking how the dependence of $Y(t+h)$ on $Y(t)$ changes at various quantiles.

To quantify the conditional dependence, we seek to estimate $f(y; Y(t+h)|Y(t))/f(y; Y(t+h))$. A brute force approach estimates separately the conditional distribution and the unconditional distribution and takes the ratio to estimate this arbitrary function. An alternative elegant way is to recognize that by “going to the quantile domain” (i.e., $Y(t+h) = Q(v; Y(t+h))$ and $Y(t) = Q(u; Y(t))$), we can interpret the ratio as “slices” of the copula density, which we call the conditional comparison density:

$$d[v; Y(t+h), Y(t+h)|Y(t) = Q(u; Y(t))] = 1 + \sum_j \text{LP}[j; h, u] S_j(v; Y(t+h)), \quad (21)$$

where the LP-Fourier orthogonal coefficients $\text{LP}[j; h]$ are given by:

$$\text{LP}[j; h, u] = \sum_k \text{LP}[j, k; Y(t), Y(t+h)] S_k(u; Y(t)).$$

Define the conditional LPinfor as:

$$\text{LPinfor}[Y(t+h)|Y(t) = Q(u; Y(t))] = \sum_j |\text{LP}[j; h, u]|^2. \quad (22)$$

We use this theory to investigate the conditional dependency structure of S&P 500 return data. Figure 8a traces out the complete path of the estimated $\text{LPinfor}[Y(t+h)|Y(t) = Q(u; Y(t))]$ function, which indicates the high asymmetric tail correlation. These conditional correlation curves can be viewed as a “local” dependence measure. An excellent discussion on this topic is given in Section 3.3.8 of Terasvirta et al. (2010).

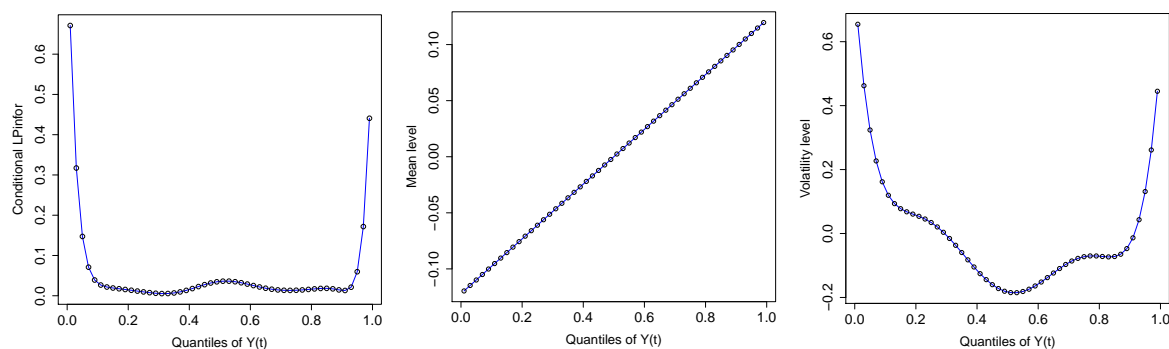


Figure 8. (a) The conditional LPinfor curve is shown for the pair $[Y(t), Y(t+1)]$. The asymmetric dependence in the tails is clearly shown, and almost nothing is going on in between. (b,c) Display of how the mean and volatility levels of conditional distribution $f[y; Y(t+1)|Y(t) = Q(u; Y(t))]$ change with respect to the unconditional marginal distribution $f(y; Y(t))$ at different quantiles.

At this point, we can legitimately ask: What aspects of the conditional distributions are changing most? Figure 8b,c displays only the two coefficients $\text{LP}[1; h, u]$ and $\text{LP}[2; h, u]$ for the S&P 500 return data for the pairs $(Y(t), Y(t+1))$. These two coefficients represent how the mean and the volatility levels of the conditional density change with respect to the unconditional reference distribution. The typical asymmetric shape of conditional volatility shown in the right panel of Figure 8b,c indicates what is known as the “leverage effect”; future stock volatility negatively correlated with past stock return, i.e., stock volatility tends to increase when stock prices drop.

3.5. Non-Crossing Conditional Quantile Modeling

We display the nonparametrically-estimated conditional quantile curves of $Y(t+h)$ given $Y(t)$. Our new modeling approach uses the estimated conditional comparison density $\hat{d}(v; h, u)$ to simulate from $F[y; Y(t+h)|Y(t) = Q(u; Y(t))]$ by utilizing the given sample $\tilde{Q}(u; Y(t))$ via an accept-reject rule to arrive at the “smooth” nonparametric model for $\hat{Q}[v; Y(t+h)|Y(t) = Q(u; Y(t))]$. See [Parzen and Mukhopadhyay \(2013b\)](#) for details about the method. Our proposed algorithm generates “large” additional simulated samples from the conditional distribution, which allows us to accurately estimate the conditional quantiles (especially the extreme quantiles). By construction, our method is guaranteed to produce non-crossing quantile curves; thus tackling a challenging practical problem.

For S&P 500 data, we first nonparametrically estimate the conditional comparison densities $\hat{d}(v; h, u)$ shown in the left panel of Figure 9 for $F(y; Y(t)) = 0.01, 0.5$ and 0.99 , which can be thought of as a “weighting function” for an unconditional marginal distribution to produce the conditional distributions:

$$\hat{f}[y; Y(t+h)|Y(t) = Q(u; Y(t))] = f(y; Y(t)) \times \hat{d}[F(y; Y(t+h)); h, u]. \quad (23)$$

This density estimation technique belongs to the skew-G modeling class ([Mukhopadhyay 2016](#)). We simulate $n = 10,000$ samples from $\hat{f}(y; Y(t+h)|Y(t))$ by accept-reject sampling from $\hat{d}(v; h, u)$, $u = \{0.01, 0.5, 0.99\}$. The histograms and the smooth conditional densities are shown in the right panel of Figure 9. It shows some typical shapes in terms of long-tailedness.

Next, we proceed to estimate the nonparametric conditional quantiles $\hat{Q}(v; Y(t+h)|Y(t))$, for $v = 0.001, 0.25, 0.5, 0.75, 0.999$, from the simulated data. Figure 10 shows the estimated conditional quantiles. The extreme conditional quantiles have a special significance in the context of financial time series. They are sometimes popularly known as Conditional Value at Risk (CoVaR), currently the most popular quantitative risk management tool (see [Adrian and Brunnermeier \(2011\)](#); [Engle and Manganelli \(2004\)](#)). The red solid line in Figure 10 is $\hat{Q}[0.001; Y(t+1) | Y(t) = Q(u; Y(t))]$, which is known as the 0.1% CoVaR function for a one-day holding period for S&P 500 daily return data. Although the upper conditional quantile curve $\hat{Q}(0.999; Y(t+1)|Y(t))$ (blue solid line) shows symmetric behavior around $F(y; Y(t)) = 0.5$, the lower quantile has a prominent asymmetric shape. These conditional quantiles give the ultimate description of the auto-regressive dependence of S&P 500 return movement in the tail region.

3.6. Nonlinear Spectrum Analysis

Here, we extend the concept of spectral density for nonlinear processes. We display the LPSpectrum -Autoregressive (AR) spectral density estimates of $YS_1(t), \dots, YS_4(t)$. The spectral density for each LP-transformed series is defined as:

$$\begin{aligned} f(\omega; YS_j) &= \sum_h LP[j, j; Y(t), Y(t+h)] e^{-i2\pi h\omega}, \quad -1/2 < \omega < 1/2 \\ &= \sum_h \text{Cov}[YS_j(t), YS_j(t+h)] e^{-i2\pi h\omega}, \quad -1/2 < \omega < 1/2. \end{aligned} \quad (24)$$

We separately fit the univariate AR model for the components of $\text{Vec}(YS)(t)$ and use the BIC order selection criterion to select the “best” parsimonious parametrization using the Burg method.

Finally, we use the estimated model coefficients to produce the “smooth” estimate of the spectral density function (see Equation (2)). The copula spectral density is defined as:

$$f(\omega; u, v) = \sum_h \text{cop}(u, v; h) e^{-i2\pi h\omega}, \quad -1/2 < \omega < 1/2. \quad (25)$$

To estimate the copula spectral density, we use the LP-comoment-based nonparametric copula density estimate. Note that both the serial copula (3.12) and the corresponding spectral density (3.25) capture the same amount of information for the serial dependence of $\{Y(t)\}$. For that reason, we recommend computing AutoLPinfor as a general dependence measure for non-Gaussian nonlinear processes.

The application of our LPSpectral tool on S&P 500 return data is shown in Figure 11. A few interesting observations are: (i) the conventional spectral density (black solid line) provides no insight into the (complex) serial dependency present in the data; (ii) the nonlinearity in the series is captured by the interesting shapes of our specially-designed times series $YS_2(t)$ and $YS_4(t)$, which classical (linear) correlogram-based spectra cannot account for; (iii) the shape of the spectra of $Z(Y(t))$ and the rank-transformed time series $YS_1(t)$ look very similar; and a (iv) pronounced singularity near zero of the spectrum of $YS_2(t)$ hints at some kind of “long-memory” behavior. This phenomena is also known as regular variation representation at frequency $\omega = 0$ (Granger and Joyeux 1980).

A quick diagnostic measure for screening significant spectrums can be computed via the information number $2 \int_0^{1/2} \log \hat{f}(\omega; S_j) d\omega$. The LPSpectrum methodology is highly robust and, thus, can tackle the heavy-tailed S&P data quite successfully.

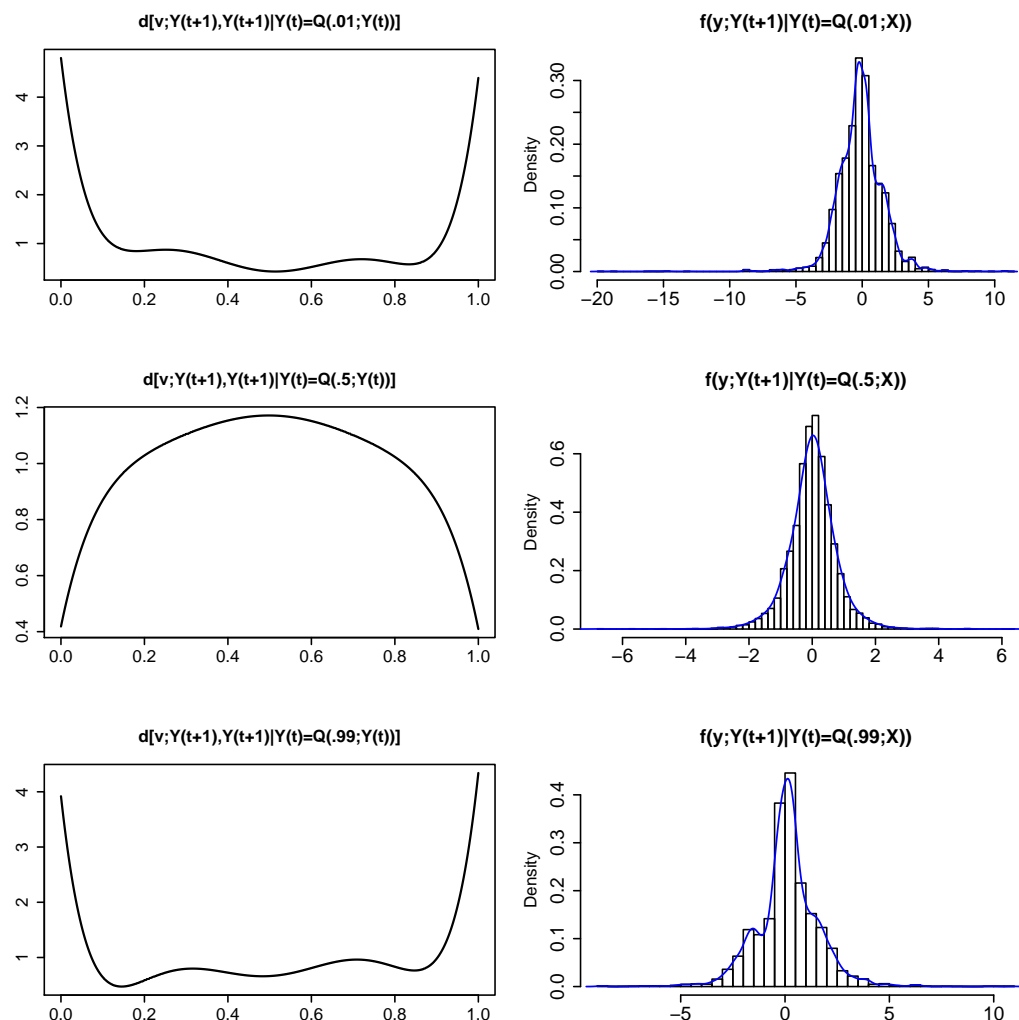


Figure 9. Each row displays the estimated conditional comparison density and the corresponding conditional distribution for $u = 0.01, 0.5, 0.99$.

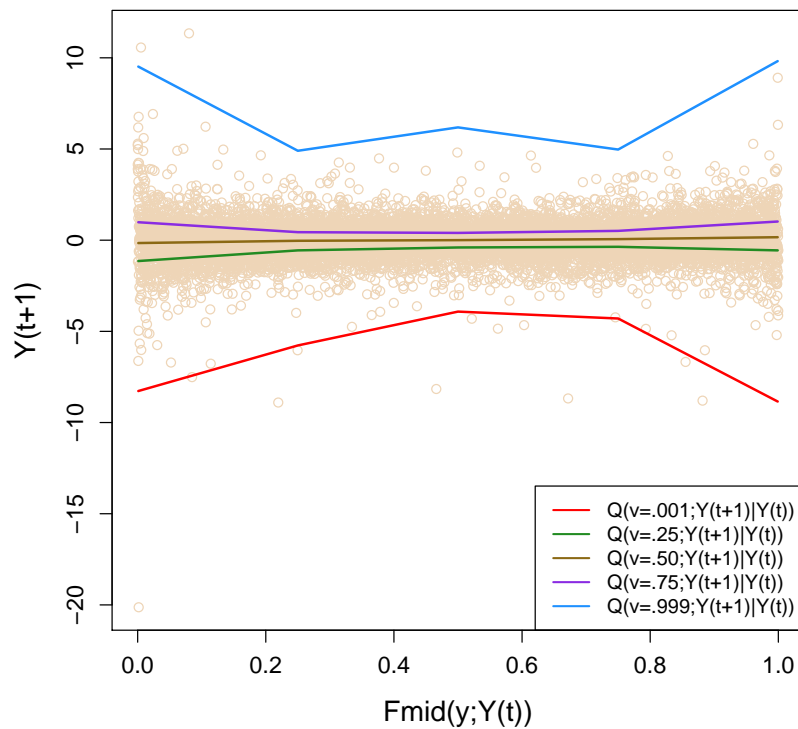


Figure 10. The figure shows estimated non-parametric conditional quantile curves for S&p 500 return data. The red solid line, which represents $\hat{Q}(0.001; Y(t+1)|Y(t))$, is popularly known as the one-day 0.1% Conditional Value at Risk measure (CoVaR).

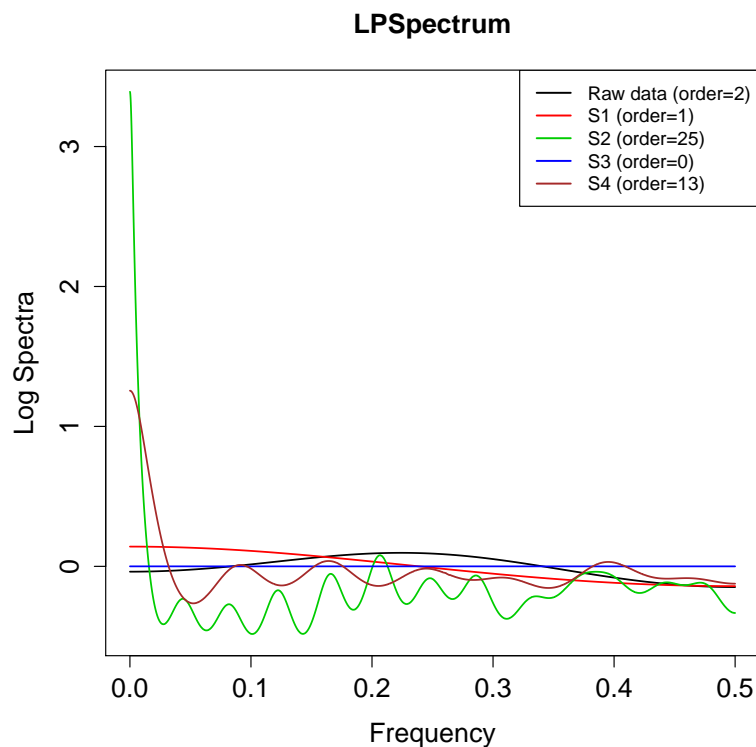


Figure 11. LPSpectrum: AR spectral density estimate for S&P 500 return data. Order selected by the BIC method. This provides a diagnostic tool for providing evidence of hidden periodicities in non-Gaussian nonlinear time series.

3.7. Nonparametric Model Specification

The ultimate goal of empirical time series analysis is nonparametric model identification. To model the univariate stationary nonlinear process, we specify the multiple autoregressive model based on $\text{Vec}(\text{YS})(t) = [\text{YS}_1(t), \dots, \text{YS}_k(t)]^T$ of the form:

$$\text{Vec}(\text{YS})(t) = \sum_{k=1}^m A(k; m) \text{Vec}(\text{YS})(t-k) + \epsilon(t). \quad (26)$$

where $\epsilon(t)$ is multivariate mean zero Gaussian white noise with covariance Σ_m . This system of equations jointly describes the dynamics of the nonlinear process and how it evolves over time. We use the BIC criterion to select the model order m , which minimizes:

$$\text{BIC}(m) = \log |\hat{\Sigma}_m| + mk^2 \frac{\log T}{T}. \quad (27)$$

We carry out this step for our S&P 500 return data. We estimate our multiple AR model based on $\text{Vec}(\text{YS})(t) = [\text{YS}_1(t), \text{YS}_2(t), \text{YS}_4(t)]^T$. We discard $\text{YS}_3(t)$ due to its flat spectrum (see Figure 11). BIC selects “best” order eight. Although the complete description of the estimated model is clearly cumbersome, we provide below the approximate structure by selecting a few large coefficients from the actual matrix equation. The goal is to interpret the coefficients (statistical parameters) of the estimated model and relate them to economic theory (scientific parameters/theory). This multiple AR LP-model (LPVAR) is given by:

$$\begin{aligned} \text{YS}_1(t) &\approx 0.071 \text{YS}_1(t-1) - 0.024 \text{YS}_1(t-2) + \epsilon_1(t) \\ \text{YS}_2(t) &\approx -0.063 \text{YS}_1(t-1) - 0.075 \text{YS}_1(t-2) + 0.06 \text{YS}_2(t-2) + 0.123 \text{YS}_2(t-5) + 0.04 \text{YS}_4(t-2) + \epsilon_2(t) \\ \text{YS}_4(t) &\approx 0.04 \text{YS}_4(t-1) + 0.038 \text{YS}_4(t-2) + 0.04 \text{YS}_2(t-3) + \epsilon_4(t). \end{aligned} \quad (28)$$

and the residual covariance matrix is:

$$\hat{\Sigma}_8 = \begin{bmatrix} 0.993 & -0.001 & -0.002 \\ -0.001 & 0.853 & -0.058 \\ -0.002 & -0.058 & 0.964 \end{bmatrix}$$

The autoregressive model of $\text{YS}_2(t)$ can be considered as a robust stock return volatility model (LPVolatility modeling), which is less affected by unusually large extreme events. The model for $\text{YS}_2(t)$ automatically discovers many known facts: (a) the sign of the coefficient linking volatility and return is negative, confirming the “leverage effect”; (b) $\text{YS}_2(t)$ is positively autocorrelated, known as volatility clustering; (c) the positive interaction with lagged $\text{YS}_4(t)$ accounts for the “excess kurtosis”.

4. Conclusions

This article provides a pragmatic and comprehensive framework for nonlinear time series modeling that is easier to use, more versatile and has a strong theoretical foundation based on the recently-developed theory of unified algorithms of data science via LP modeling (Mukhopadhyay 2016, 2017; Mukhopadhyay and Fletcher 2018; Mukhopadhyay and Parzen 2014; Parzen and Mukhopadhyay 2012, 2013a, 2013b). The summary and broader implications of the proposed research are:

- From the theoretical standpoint, the unique aspect of our proposal lies in its ability to simultaneously embrace and employ the spectral domain, time domain, quantile domain and information domain analyses for enhanced insights, which to the best of our knowledge has not appeared in the nonlinear time series literature before.
- From a practical angle, the novelty of our technique is that it permits us to use the techniques from linear Gaussian time series to create non-Gaussian nonlinear time series models with highly interpretable parameters. This aspect makes LPTIME computationally extremely attractive for data

scientists, as they can now borrow all the standard time series analysis machinery from R libraries for implementation purposes.

- From the pedagogical side, we believe that these concepts and methods can easily be augmented with the standard time series analysis course to modernize the current curriculum so that students can handle complex time series modeling problems (McNeil et al. 2010) using the tools with which they are already familiar.

The main thrust of this article is to describe and interpret the steps of LPTIME technology to create a realistic general-purpose algorithm for empirical time series modeling. In addition, many new theoretical results and diagnostic measures were presented, which laid the foundation for the algorithmic implementation of LPTIME. We showed how LPTIME can systematically explore the data to discover empirical facts hidden in time series. For example, LPTIME empirical modeling of S&P 500 return data reproduces the ‘stylized facts’—(a) heavy tails; (b) non-Gaussian; (c) nonlinear serial dependence; (d) tail correlation; (e) asymmetric dependence; (f) volatility clustering; (g) long-memory volatility structure; (h) efficient market hypothesis; (i) leverage effect; (j) excess kurtosis—in a coherent manner under a single general unified framework. We have emphasized how the statistical parameters of our model can be interpreted in light of established economic theory.

We have recently applied this theory for large-scale eye-movement pattern discovery problem, which came out as the winner (among 82 competing algorithms) of the 2014 IEEE International Biometric Eye Movements Verification and Identification Competition (Mukhopadhyay and Nandi 2017). The proposed algorithm is implemented in the R package LPTIME (Mukhopadhyay and Nandi 2015), which is available on CRAN.

We conclude with some general references: a few popular articles: Brillinger (1977, 2004); Engle (1982); Granger and Lin (1994); Granger (1993, 2003); Parzen (1967, 1979); Salmon (2012); Tukey (1980); books: Guo et al. (2017); Terasvirta et al. (2010); Tsay (2010); Woodward et al. (2011); and review articles: Granger (1998); Hendry (2011).

Author Contributions: Conceptualization, S.M. and E.P.; Methodology, S.M. and E.P.; Formal Analysis, S.M.; Computation, S.M.; Writing-Review & Editing, S.M.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Adrian, Tobias, and Markus K Brunnermeier. 2011. *Covar*. Technical Report. Cambridge: National Bureau of Economic Research.
- Brillinger, David R. 1977. The identification of a particular nonlinear time series system. *Biometrika* 64: 509–15. [CrossRef]
- Brillinger, David R. 2004. Some data analyses using mutual information. *Brazilian Journal of Probability and Statistics* 18: 163–83.
- Engle, Robert F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society* 50: 987–1007. [CrossRef]
- Engle, Robert F., and Simone Manganelli. 2004. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics* 22: 367–81.
- Granger, Clive, and Jin-Lung Lin. 1994. Using the mutual information coefficient to identify lags in nonlinear models. *Journal of Time Series Analysis* 15: 371–84. [CrossRef]
- Granger, Clive W. J. 1993. Strategies for modelling nonlinear time-series relationships. *Economic Record* 69: 233–38. [CrossRef]
- Granger, Clive W. J. 1998. Overview of nonlinear time series specification in economics. Paper presented at the NSF Symposium on Nonlinear Time Series Models, University of California, Berkeley, CA, USA, 22 May 1998.

- Granger, Clive W. J. 2003. Time series concepts for conditional distributions. *Oxford Bulletin of Economics and Statistics* 65: 689–701. [[CrossRef](#)]
- Granger, Clive W. J., and Roselyne Joyeux. 1980. An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1: 15–29. [[CrossRef](#)]
- Guo, Xin, Howard Shek, Tze Leung Lai, and Samuel Po-Shing Wong. 2017. *Quantitative Trading: Algorithms, Analytics, Data, Models, Optimization*. New York: Chapman and Hall/CRC.
- Hendry, David F. 2011. Empirical economic model discovery and theory evaluation. *Rationality, Markets and Morals* 2: 115–45.
- McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts. 2010. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton: Princeton University Press.
- Mukhopadhyay, Subhadeep. 2016. Large scale signal detection: A unifying view. *Biometrics* 72: 325–34. [[CrossRef](#)] [[PubMed](#)]
- Mukhopadhyay, Subhadeep. 2017. Large-scale mode identification and data-driven sciences. *Electronic Journal of Statistics* 11: 215–40, doi:10.1214/17-EJS1229. [[CrossRef](#)]
- Mukhopadhyay, Subhadeep, and Douglas Fletcher. 2018. Generalized Empirical Bayes Modeling via Frequentist Goodness-of-Fit. *Nature Scientific Reports* 8: 1–15. [[CrossRef](#)] [[PubMed](#)]
- Mukhopadhyay, Subhadeep, and Shinjini Nandi. 2015. *LPTIME: LP Nonparametric Approach to Non-Gaussian Non-Linear Time Series Modelling*. CRAN, R Package Version 1.0-2. Ithaca: Cornell University Library.
- Mukhopadhyay, Subhadeep, and Shinjini Nandi. 2017. LPiTrack: Eye movement pattern recognition algorithm and application to biometric identification. *Machine Learning*. [[CrossRef](#)]
- Mukhopadhyay, Subhadeep, and Emanuel Parzen. 2014. LP approach to statistical modeling. *arXiv*. arXiv:1405.2601.
- Parzen, Emanuel. 1967. On empirical multiple time series analysis. In *Statistics, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, vol. 1, pp. 305–40.
- Parzen, Emanuel. 1979. Nonparametric statistical data modeling (with discussion). *Journal of the American Statistical Association* 74: 105–31. [[CrossRef](#)]
- Parzen, Emanuel. 1992. Time series, statistics, and information. In *New Directions in Time Series Analysis*. Edited by Emanuel Parzen, Murad Taqqu, David R. Brillinger, Peter Caines, John Geweke and Murray Rosenblatt. New York: Springer Verlag, pp. 265–86.
- Parzen, Emanuel. 1997. Comparison distributions and quantile limit theorems. Paper presented at the International Conference on Asymptotic Methods in Probability and Statistics, Carleton University, Ottawa, ON, Canada, July 8–13.
- Parzen, Emanuel, and Subhadeep Mukhopadhyay. 2012. Modeling, Dependence, Classification, United Statistical Science, Many Cultures. *arXiv*. arXiv:1204.4699.
- Parzen, Emanuel, and Subhadeep Mukhopadhyay. 2013a. United Statistical Algorithms, LP comoment, Copula Density, Nonparametric Modeling. Paper presented at the 59th ISI World Statistics Congress (WSC) of the International Statistical Institute, Hong Kong, China, August 25–30.
- Parzen, Emanuel, and Subhadeep Mukhopadhyay. 2013b. United Statistical Algorithms, Small and Big Data, Future of Statisticians. *arXiv*. arXiv:1308.0641.
- Salmon, Felix. 2012. The formula that killed wall street. *Significance* 9: 16–20. [[CrossRef](#)]
- Terasvirta, Timo, Dag Tjøstheim, and Clive W. J. Granger. 2010. *Modelling Nonlinear Economic Time Series*. Kettering: OUP Catalogue.
- Tsay, Ruey S. 2010. *Analysis of Financial Time Series*. Hoboken: Wiley.
- Tukey, John. 1980. Can we predict where “time series” should go next. In *Directions in Times Series*. Hayward: IMS, pp. 1–31.
- Woodward, Wayne A., Henry L. Gray, and Alan C. Elliott. 2011. *Applied Time Series Analysis*. Boca Raton: CRC Press.

