

# Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rate

Andrew Hencic\*  
Christian Gourieroux†

first version: January 28, 2014

revised: March 17, 2014

This paper introduces a noncausal autoregressive process with Cauchy errors in application to the exchange rates of the Bitcoin electronic currency against the US Dollar. The dynamics of the daily Bitcoin/USD exchange rate series displays episodes of local trends, which can be modelled and interpreted as speculative bubbles. The bubbles may result from the speculative component in the on-line trading. The Bitcoin/USD exchange rates are modelled and predicted.

**Keywords:** Electronic Currency, Bubble, Noncausal Autoregressive Process, Global and Local Trends.

JEL number: C14, G32, G23.

---

\*York University, Canada, *e-mail:* ahencic@econ.yorku.ca.

†University of Toronto and CREST, *e-mail:* gourierou@ensae.fr.

# 1 Introduction

In recent months digital currencies (sometimes referred to as crypto-currencies) and their standard bearer, Bitcoin, have been garnering more public attention[1]. This can likely be attributed to two factors. Public adoption of the digital currency is beginning to become more commonplace [2] and its more nefarious uses are slowly being exposed [3].

A prime example of the first point is the University of Nicosia in Cyprus. The University is the largest private university in Cyprus and is beginning to accept bitcoins as tuition payment. The university's reasoning is that they wish to be at the forefront of global commerce, but there may be other reasons at play. More recently Cyprus has gone through significant financial stress and many of the country's depositors will likely face significant losses [4]. Mistrust of the established financial system may lead institutions to begin accepting alternative means of payment.

As for the nefarious uses of bitcoins, the most recent story about the raid on the website The Silk Road can speak to the dark side of digital and anonymous currency. In October of 2013 the FBI shut down The Silk Road for allegedly selling illegal drugs and charged its owner with a whole host of offenses. Critics of digital currencies say that the anonymity provided to their users is dangerous and should be further regulated. The government of the United States has responded to these concerns by implementing rules to attempt to curb the use of digital currencies in money laundering [5]. With the market capitalization of bitcoin surpassing \$12 Billion USD [6] and its ever increasing adoption further study of the uses, threats and mechanisms that govern digital currencies is needed.

The objective of this paper is to examine the dynamics of the Bitcoin/USD exchange rate and to predict its future evolution. The dynamics of the series are characterized by the presence of local trends and short-lived episodes of soaring Bitcoin/USD rates, followed by sudden almost vertical declines. These patterns are referred to as bubbles. In economics, bubbles in asset prices have been introduced in the context of the rational expectation hypothesis in the seventies and as a result of the speculative behavior of traders. The bubbles in the Bitcoin/USD rate may originate from a) the fact that the bitcoin market is still an emerging market with a lot of speculative trading, b) the asymmetric information and crowd phenomena [see, e.g. Bunnermeier (2001) for the analogous on Nasdaq], c) the lack of a centralized management and control of exchange rate volatil-

ity, d) the deterministic supply of bitcoins and the evolution of the volume over time. As the volume of bitcoins available on the market is exogenously determined, this enhances the bitcoin price and exchange rate volatility.

Because of the presence of local explosive trends, depicted as bubbles, the Bitcoin/USD exchange rate cannot be modelled by any traditional ARIMA or ARCH models [see, e.g. Andrews et al (2009)]. In this paper, we use the mixed causal-noncausal autoregressive process with Cauchy errors [Gourieroux, Zakoian (2013)] to estimate and predict the Bitcoin/USD exchange rate.

The structure of the paper is as follows. In Section 2, we describe the bitcoin as an electronic currency, and we explain the mechanisms of bitcoin trading and storage. Next, we describe the data and the period of interest that includes a bubble burst and crash. A speculative bubble is a nonlinear dynamic feature that can be accommodated by the aforementioned noncausal autoregressive process. In Section 3, we review the properties of noncausal processes and introduce the associated inference and prediction methods. The application to the Bitcoin/ US Dollar exchange rates recorded on the Mt. Gox <sup>1</sup> exchange market is presented in Section 4. The noncausal model is used to predict the occurrence of the bubble in the Bitcoin/USD exchange rate. Section 5 concludes the paper.

## 2 The Bitcoin/USD exchange rate

### 2.1 Bitcoin currency

Bitcoin (BTC) is an electronic currency originally created by a developer under the pseudonym Satoshi Nakamoto in 2009[7]. The electronic currency is distributed on a peer-to-peer network anonymously between any two accounts. There is no formal denomination or name for units of the currency other than 1.00 BTC being referred to as a bitcoin and the smallest possible denomination, 0.00000001 BTC, being a "satoshi".

The bitcoin can be purchased on a virtual exchange market, such as the [www.mtgox.com](http://www.mtgox.com) against the US Dollar or other currencies <sup>2</sup>. Users of the currency store it on a private

---

<sup>1</sup>Formerly Magic: The Gathering Online Exchange.

<sup>2</sup>The transactions on this market have been suspended on February 25, 2014. The reason is yet to be revealed, but an attack by hackers has been declared.

digital "wallet". This wallet has no personal identification with an individual and is comprised of three components: an address, a private key, and a public key. There is nothing that connects a wallet to an individual. This level of anonymity has been one of the driving forces behind the currency's popularity. The bitcoin can be used to purchase a number of goods and services that are listed on the Bitcoin website.

Three types of wallets exist: the software wallet, the mobile wallet and the web wallet. Software wallets are installed directly on a computer and allow the user complete control over the wallet. Mobile wallets are installed on mobile devices and operate the same way. Web wallets host an individual's bitcoins online. All of these wallets can be accessed with just the private key assigned to the address. Again, there is nothing to associate a physical human being with a Bitcoin address other than if the person owns the hardware on which the wallet is installed.

As of December 2, 2013 the total market capitalization of bitcoin is approximately \$12,326,035,159 USD [8]. Bitcoin is traded 24 hours a day on various exchanges, the largest of which include Mt. Gox (based in Japan) <sup>3</sup> and BTC China (recently the world's largest BTC exchange [9]). The former is a real time exchange whereas BTC China is a fixed rate exchange[10]. Bitcoins are denominated in USD on Mt. Gox and in Renminbi on BTC China. After a clarification by the People's Bank of China on Bitcoin's status at the beginning of December 2013, the exchanges on BTC China can only be done in Chinese Yuan, and the users have to now provide their identity using, for example, a passport number. Trading and use of Bitcoin is forbidden in Thailand.

The trading volume of bitcoin on Mt. Gox has slowly increased over time as adoption of the currency has increased. On its first day of trading on Mt. Gox, the total volume of bitcoin traded was 20 units. Obviously this is a very small number in comparison to the 3,436,900 bitcoins in circulation at that time. However, trading volume has gradually increased since then as Bitcoin has become more generally accepted and garnered more attention. Trading volume reached an all-time high on April 15, 2013 with 572, 185.7 bitcoins changing hands on Mt. Gox. At the time there were approximately 11,027,700 units in existence, meaning that on this day approximately 5 % of all bitcoins in circulation were traded on Mt. Gox.

The long term supply of BTC will never exceed 21,000,000 units. However, the daily

---

<sup>3</sup>It represented 12% of the trades before it collapsed.

volume traded on the platforms can be much smaller. The traded volume was 31 800 BTC on Mt. Gox on December 8, 2013. The evolution of the traded volume of bitcoins between February and July 2013 is displayed in Figures 1 and 2.

[Insert Figure 1: Bitcoin volume, Feb-Jul 2013]

[Insert Figure 2: Bitcoin transactions, Feb-Jul 2013]

Figure 1 provides the daily volume exchanged against USD while Figure 2 provides the daily volume of bitcoins used for real transactions that is for the sale and purchase of goods and services offered in bitcoin. The daily volumes are small compared to the capitalization of bitcoin, showing that this emerging market may encounter liquidity problems.

Bitcoins are produced in such a way that the volume of new bitcoins produced will be halved every four years until the volume of new coins produced decays to zero. At this point the final supply of bitcoins will be fixed (the exact amount of units varies depending on rounding, but it will be less than 21 million units) [11]. Bitcoins are produced in a process referred to as "mining". Computers on the Bitcoin network solve complex mathematical problems and are rewarded for their work with a predetermined amount of bitcoins, referred to as a "block reward", and a transaction fee. The current block reward is 25 bitcoins [12]. In order to control the supply of bitcoins being produced the difficulty of these problems is automatically adjusted so that the time between solutions averages 10 minutes.

## **2.2 Bitcoin transactions**

To ensure the security of transactions, the Bitcoin system uses public key cryptography. Each individual has one or more addresses with an associated private and public key. The system is totally anonymous and balances are only associated with an address and its keys. Only the user with the private key can sign a transfer of bitcoins to another party, whereas anybody in the network can validate the signature and transaction using the user's public key [13]. When a transaction occurs, one user sends another an amount of bitcoins and signs the transaction with their private key. The user who sends the bitcoins announces a public key and it falls on the network to verify the signature. The user then broadcasts the transaction on the Bitcoin network. In order to prevent double spending the details about a transaction are sent to as many other computers on the network as

possible in a block. Each computer on this network has a registry of these blocks called a "block chain". In order for the newest block to be accepted into the chain, it must be valid and must include proof of work (the solution to the aforementioned math problem). When a block is announced the miners work to verify the transaction by solving the math problem. When a solution is reached it is verified by the rest of the network. This allows for the tracking of the life of every individual bitcoin produced.

Thus for any individual to double spend their Bitcoins, their computing power would have to exceed the combined computing power of all other Bitcoin computers.

Alternatives to Bitcoin have already begun to spring up. The largest competitor is Litecoin, which as of December 2, 2013 has a market capitalization of \$695,376,891 USD [14]. Litecoin seeks to be an improvement over Bitcoin by attempting to overcome some of the more technical issues facing Bitcoin.

## 2.3 The data

In our empirical study, we consider the Bitcoin/USD exchange rate from the first part of year 2013 that includes a bubble, which bursted on April 10, 2013.

[Insert Figure 3: Bitcoin/USD exchange rate, Feb-Jul 2013]

More specifically, the sample consists of 150 observations on the daily closing values of the Bitcoin/USD exchange rate over the period February 20 -July 20, 2013. The dynamics of the data is displayed in Figure 3. We observe a nonlinear trend as well as the bubble that peaked at the virtual time  $t=50$ . The sample median, interquartile range and total range <sup>4</sup> are 103.27, 46.69 and 208.21, respectively. For comparison, the sample mean and variance are 96.98 and 1327.63, respectively. These standard summary statistics can be misleading, since their usual interpretation assumes the stationarity of the process. This assumption is clearly not satisfied for the Bitcoin/USD exchange rate. Figure 4 shows the histogram and a kernel-based density estimate of the sample marginal density.

[Insert Figure 4: Bitcoin histogram]

Both estimates display fat tails, as suggested by the fact that the total range is five times greater than the interquartile range. Also, the histogram indicates a discontinuity

---

<sup>4</sup>The difference between the sample max and min

in the left tail, which shows as an almost bimodal pattern in the kernel-smoothed density estimate.

### 3 The model

This section presents the mixed causal-noncausal autoregressive process, explains how this process accommodates the bubble effects observed in the Bitcoin/USD exchange rate series and discusses the estimation and inference.

#### 3.1 The noncausal and mixed autoregressive process

A mixed (causal-noncausal) autoregressive process is a stochastic process  $\{y_t; t = 0, \pm 1, \pm 2, \dots\}$ , defined by:

$$\Psi(L^{-1})\Phi(L)y_t = e_t, \quad (3.1)$$

where  $\Psi(L^{-1})$  and  $\Phi(L)$  are polynomials in the negative (resp. positive) powers of the lag operator  $L$ , such that  $\Psi(L^{-1}) = 1 - \psi_1 L^{-1} - \dots - \psi_s L^{-s}$  and  $\Phi(L) = 1 - \phi_1 L - \dots - \phi_r L^r$ . The roots of both polynomials are assumed to lie outside the unit circle, and error terms  $e_t$  are identically and independently distributed. When  $\phi_1 = \dots = \phi_r = 0$ , model (3.1) defines a pure noncausal autoregressive process of order  $s$ , while for  $\psi_1 = \dots = \psi_s = 0$ , the process  $y_t$  is the traditional pure causal AR( $r$ ) process. When some of the coefficients of both polynomials are non-zero, we obtain a mixed process that contains both the lags and leads of  $y_t$ . Under the above assumptions, there exists a unique stationary solution to equation (3.1). This solution admits a strong, two-sided moving average representation:

$$y_t = \sum_{j=-\infty}^{\infty} \xi_j e_{t-j},$$

where the  $\xi_j$ 's are the coefficients of an infinite order polynomial in positive and negative powers of the lag operator  $L$  and such that:  $\Xi(z) = \sum_{j=-\infty}^{\infty} \xi_j z^j = [\Psi(z^{-1})]^{-1}[\Phi(z)]^{-1}$ .

When errors  $e_t$  are normally distributed, the causal and noncausal components of the dynamics cannot be distinguished, and model (3.1) is not identifiable. However, the causal and noncausal autoregressive coefficients are identifiable when the process  $(e_t)$

is not Gaussian <sup>5</sup>. For example, Lanne, Saikkonen (2010) consider t-Student distributed errors, while Gouriéroux and Zakoian (2013) discuss the properties of the purely noncausal autoregressive process ( $r=0$ ,  $s=1$ ) with Cauchy distributed errors. In particular, the density of a Cauchy distributed random variable  $X$  with location  $\mu$  and scale  $\gamma$  is:

$$f(e_t) = \frac{1}{\pi} \left[ \frac{\gamma}{(x - \mu)^2 + \gamma^2} \right]$$

In Section 3, we will assume that  $e_t \sim Cauchy(0, \gamma)$ . A particular feature of the Cauchy distribution is that the expected value as well as all population moments of any higher order do not exist.

### 3.2 The bubble effect

The trajectory of the Bitcoin/USD exchange rate displays repetitive episodes of upward trends, followed by instantaneous drops, which are called bubbles. In general, a bubble has two phases: 1) a phase of fast upward (or downward) departure from the stationary path that resembles an explosive pattern and displays an exponential rate of growth, followed by 2) a phase of sudden almost vertical drop (or upspring) back to the underlying fundamental path. There exist several definitions of a bubble in the economic literature. The first definition was introduced by Blanchard (1979) in the framework of rational expectation models. The formal definition by Blanchard as well as the later definitions by Blanchard and Watson (1982) and Evans (1991) all assume a nonlinear dynamic models of  $x_t$  (say) with two components, one of which depicts the fundamental path of  $x_t$ , while the second one represents the bubble effect. The economic explanation of this phenomenon is as follows: a bubble results from the departure of a price of an asset from its fundamental value. In the context of the Bitcoin/USD exchange rate, the bubbles may result from the speculative trading that makes the rate deviate quickly above its trend, although it is hard to say if the trend is representative of the fundamental value of the bitcoin. Indeed, the bitcoin is a virtual currency, which is backed neither on a real asset, nor on the performance of a firm or a national economy.

---

<sup>5</sup>See e.g. Cheng (1992), Rosenblatt (2000), Theorem 1.3.1. for errors with finite variance, Breidt, Davis (1992) for errors with finite expectation and infinite variance, Gouriéroux, Zakoian (2013)b for errors without finite expectation, as the Cauchy errors considered in the application.



So far, the bubbles were considered in the time series literature as nonstationary phenomena and treated similarly to the explosive, stochastic trends due to unit roots. In fact, the existing tests for the presence of a bubble are essentially tests of a breakpoint in the general explosive stochastic trend of a nonstationary process [see e.g. Phillips, Wu and Yu (2011) Phillips, Shi and Yu (2012)].

Gourieroux and Zakoian (2013a) propose a different approach and assume that the bubbles are rather short-lived explosive patterns caused by extreme valued shocks in a noncausal, stationary process. Formally, that process is a noncausal AR(1) model with Cauchy distributed errors. The approach in reverse time, based on a noncausal model allows for accommodating the asymmetric pattern of the bubble. The merit of the Cauchy distributed errors is in replicating the sudden spike in the reverse time trajectory that is observed as a bubble burst from the calendar time perspective. Such a noncausal or mixed process has to be examined conditionally on the information of the current and past rates. It is known that a noncausal, linear autoregressive process also has a nonlinear causal autoregressive dynamics, except in the Gaussian case. This is the special nonlinear feature, which makes it suitable for modelling the bubbles in Bitcoin/USD exchange rate. Moreover the noncausal autoregressive model allows for forecasting the occurrence of a future bubble and the time of bubble burst. The methodology of forecasting is discussed in Section 3.4 and illustrated in the application in Section 4.

### 3.3 Estimation and Inference

The traditional approach to the estimation of causal time series models relies on the Box-Jenkins methodology that consists of three steps: identification, estimation and diagnostics. In application to noncausal and mixed processes, most of the traditional Box-Jenkins tools of analysis need to be interpreted with caution. The reason is that most of the traditional estimators are based on the first- and second-order sample moments of the process and rely on the Gaussian approximation of its density, while the noncausal processes need to be non-Gaussian to solve the aforementioned identification purpose and may have infinite moments of order one and/or two.

#### a) *Identification*

The autocorrelation function (ACF) is the basic tool for detecting temporal depen-

dence. By construction, the ACF estimators rely on an implicit normality assumption, as they are computed from the sample moments up to order two. Due to the aforementioned nonidentifiability problem, the ACF cannot reveal whether a time series is causal or not, as it yields identical results in either case. It remains however a valid tool for detecting serial dependence in variables with infinite variances [see Andrews, Davis (2013)]. In particular Andrews and Davis show that the total autoregressive order  $p = r + s$  can be inferred from the autocorrelation function, while  $r$  and  $s$  need to be inferred from the estimated models by comparing their fit criteria, computed from the sample.

For variables with infinite variance, Davis and Resnick (1985) established the asymptotic properties of the sample autocorrelation  $\hat{\rho}$  at lag  $l$  defined as:

$$\hat{\rho}(l) = \frac{\hat{\gamma}(l)}{\hat{\gamma}(0)}, \quad \hat{\gamma}(l) = \frac{1}{T} \sum_{t=1}^{T-l} (y_t - \bar{y})(y_{t+l} - \bar{y}), \quad l > 0,$$

where  $T$  is the sample size and  $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ .

In the presence of Cauchy errors, the standard confidence intervals of the ACF are no longer valid as the sample ACF is no longer asymptotically normally distributed and has a nonstandard speed of convergence. By using the results of Davis and Resnick (1985), Gouriéroux and Zakoian (2013a) (Proposition 6) show that the sample autocorrelations of a noncausal AR(1) process with Cauchy errors and autoregressive coefficient  $\rho$  have a limiting stable distribution and a rate of convergence that is different from the standard  $\sqrt{T}$  rate. More specifically, let us denote the vector of sample autocorrelations up to lag  $M$ : by  $\hat{\rho}_T = (\hat{\rho}_T(1), \dots, \hat{\rho}_T(M))'$  and consider the true values  $\rho = (\rho, \dots, \rho^M)'$ . Then,

$$\frac{T}{\log T} (\hat{\rho}_T - \rho) \xrightarrow{d} Z = (Z_1, \dots, Z_M)',$$

where for  $l = 1, \dots, M$ ,  $Z_l = \sum_{j=1}^{\infty} [\rho^{j+l} - \rho^{j-l}] S_j / S_0$ , and  $S_1, S_2, \dots$  is an i.i.d. sequence of symmetric 1-stable random variables independent of the positive 1/2 stable random variable  $S_0$ . The limiting true values can be interpreted as pseudo-autocorrelations, as the autocorrelations themselves do not exist in a process with infinite variance.

#### b) *Estimation*

The standard Gaussian quasi-maximum likelihood approach can no longer be used to estimate the autoregressive parameters due to the Gaussian-specific identification prob-

lem. However, when the distribution of the errors is non-Gaussian, the estimation of the parameters in noncausal and mixed processes can be based on the maximum likelihood estimator, which preserves its speed of convergence and asymptotic normality [see, Lanne, Saikkonen (2010)]. The maximum likelihood method differs slightly from that used in causal processes. It is called the "approximate maximum likelihood" for the reason that the sample used in the approximate likelihood is reduced to  $T - (r + s)$  observations<sup>6</sup>. Indeed, the first error to be included in the likelihood function that can be written without a value of  $y_t$  prior to the sample is  $e_{r+1}$ . To see that, assume  $\psi_1 = \dots = \psi_s = 0$  and write:

$$e_{r+1} = y_t - \phi_1 y_{t-1} - \dots - \phi_r y_{t-r}.$$

Suppose now that  $\phi_1 = \dots = \phi_r = 0$ . The last error in the sample to be included in the likelihood function that can be written without the values of  $y_t$  posterior to the sample is

$$e_{T-s-1} = y_T - \psi_1 y_{T+1} - \dots - \psi_s y_{T+s}$$

The Approximate Maximum Likelihood (AML) is defined as:

$$(\hat{\Psi}, \hat{\Phi}, \hat{\theta}) = \underset{\Psi, \Phi, \theta}{\operatorname{Argmax}} \sum_{t=r+1}^{T-s} \ln g[\Psi(L^{-1})\Phi(L)y_t; \theta], \quad (3.2)$$

where  $g[.; \theta]$  denotes the probability density function of  $e_t$ .

Lanne, Saikkonen (2010) show that the traditional Wald tests and other inference methods, like the AIC and SBC fit criteria based on the approximated likelihood remain valid. The fit criteria are essentially used for determining the autoregressive orders  $r$  and  $s$  of the process. The model with the  $r$  and  $s$  that minimizes one of the fit criteria is selected at this "ex post" identification stage, analogously to the choice of orders  $p$  and  $q$  in an ARMA( $p, q$ ) process.

### c) *Diagnostics*

The diagnostic checking consists in testing if the estimated shocks  $\hat{e}_t = \hat{\Psi}(L^{-1})\hat{\Phi}(L)y_t$  of the model are strong white noise. The asymptotic distribution of the sample autocorrelation of the residuals is different from the standard one derived for processes with

---

<sup>6</sup>The approximate likelihood disregards the first  $r$  state variables that summarize the effect of shocks before time  $r$  and the last  $s$  state variables that summarize the effect of shocks after time  $T - s$  [Gourieroux, Jasiak (2013)] and is therefore constructed from shocks  $e_{r+1}, \dots, e_{T-s-1}$  only.

finite variance. For instance, for a noncausal Cauchy autoregressive process of order 1, the limiting distribution of the residual autocorrelation estimator at lag 1 is:

$$\frac{T}{\ln T} \hat{r}_T(1) \xrightarrow{d} \rho^*(1 + 2\rho^*)S_1/S_0,$$

where  $\rho^*$  is the noncausal autoregressive coefficient of process  $Y$ . Contrary to the standard process with finite variance, the limiting distribution depends on  $\rho^*$ .

### 3.4 Forecasting

Due to the different dynamics of non-Gaussian processes in the calendar and the reverse times, the "backcasting" algorithm in the spirit of Newbold [Biometrika (1973)] is no longer valid. Nevertheless, it is possible to extend the concept of the Kalman filter and make it applicable to noncausal and mixed processes. The approach consists of three steps [see e.g. Gouriéroux, Jasiak (2014)]:

*step 1:* Filter shocks  $e_t$  for  $t = 1, \dots, T$  and the causal and noncausal state components of the process.

*step 2:* Estimate the predictive distribution of the noncausal component of the process by a look-ahead estimator.

*step 3:* Simulate the future noncausal components of the process by a sampling importance resampling (SIR) algorithm and infer the simulated future values of the process.

This methodology is used in the next section to derive the prediction intervals.

## 4 Application

### 4.1 ACF Analysis

The traditional Box-Jenkins approach starts from the analysis of the sample autocorrelation function (ACF). The ACF provides information on the possible linear serial dependence in the series, but its interpretation can be rather misleading in the case of extreme events.

The standard confidence interval for testing the statistical significance of the autocorrelations is based on the approximate limiting standard normal distribution of the

autocorrelation estimator at a given lag, under the null hypothesis that the true value of that autocorrelation is zero. Hence, with the sample size of 150, the statistically significant autocorrelations exceed 0.16 in absolute value.

In order to establish the confidence interval for Cauchy distributed errors, we approximate the limiting distribution of the pseudo-autocorrelation estimator given in Section 3.3 by simulations. We draw independent standard normals  $e_1, e_2, e_3$  and build the ratio

$$Z = \frac{S_1}{S_0} = \frac{e_1 e_3^2}{e_2}$$

where  $S_1 = \frac{e_1}{e_2}$  is a symmetric 1-stable random variable and  $S_0 = \frac{1}{e_3^2}$  is a symmetric 0.5 stable random variable. From the 25th and 975th order statistics from a sample of 1000 values of  $Z$  multiplied by  $\frac{\ln(150)}{150}$ , we obtain the confidence interval  $[-0.36, 0.36]$ . Under the null hypothesis of zero pseudo-autocorrelation at lag  $l$ , the statistically significant autocorrelation at lag  $l$  is less than 0.36 in absolute value with the asymptotic probability of 95%.

In Figure 5, we plot the ACF of the data with the standard confidence interval and the interval adjusted for infinite variance.

[Insert Figure 5: ACF Bitcoin]

The ACF displays slow, linear decay, which resembles the patterns observed in unit root processes. Moreover, the Dickey-Fuller and the Augmented Dickey-Fuller ADF(4) tests accept the null hypothesis of a unit root in the data with p-values 0.4 and 0.6, respectively. However, it is easily checked that the standard procedure of transforming the data into first differences and estimating a stationary ARMA cannot accommodate the nonlinear features of the series.

## 4.2 Global and Local Trends

In the Bitcoin/USD exchange rate series, it is important to disentangle the fundamental and the bubble components. The fundamental component is modelled as a nonlinear deterministic trend <sup>7</sup> and the bubble component as a noncausal autoregressive process with Cauchy errors. Accordingly, we define the Bitcoin/USD rate as:

---

<sup>7</sup>Alternatively, it can be represented by a model with a stochastic trend, assumed independent of the shocks that create the speculative bubble.

$$rate_t = trend_t + y_t,$$

a) *Estimation of the trend and detrended series*

In order to remove the trend, we fit a nonlinear function of time by regressing the data on a 3rd degree polynomial in time. The detrended series, obtained as the following series of residuals:

$$y_t = rate_t + 3.045 - 3.854t + 3.499t^2 - 0.866t^3$$

is calculated and plotted in Figure 6.

[Insert Figure 6: Detrended series]

The marginal density of  $y_t$  is shown in Figure 7.

[Insert Figure 7: Detrended series, histogram]

We observe that the detrended series no longer displays the bimodal pattern, while it preserves the peaked and long-tailed shape of the density of the Bitcoin/USD rate. The ACF function of the detrended series, given in Figure 8, shows considerably less persistence than the original series and indicates short linear memory.

[Insert Figure 8: Detrended series, ACF]

b) *Noncausal analysis of the detrended series*

Next, the detrended series is modelled as a noncausal autoregressive process.

Let us first consider a noncausal Cauchy AR(1) process:

$$y_t = \psi y_{t+1} + e_t, \tag{4.3}$$

where  $e_t$  are independent and Cauchy distributed with location 0 and scale  $\gamma$ ,  $e_t \sim Cauchy(0, \gamma)$ . At this point, it is interesting to compare the trajectory of  $y_t$  with the simulated path of a noncausal AR(1) with the autoregressive coefficient 0.9, as displayed in Gouriou, Zakoian (2013a), Figure 4. It is clear that the dynamics of the transformed Bitcoin/USD rate and of the simulated series resemble one another. The model is

estimated by maximizing the approximated log-likelihood function, based on the Cauchy density function:

$$\ln L(\psi, \gamma) = (T - 1)[- \ln(\pi) + \ln(\gamma)] - \sum_{t=1}^{T-1} [\ln((y_t - \psi y_{t+1})^2 + \gamma^2)]. \quad (4.4)$$

The parameter estimates and the estimated standard errors are given in Table 1:

Table 1: AR(1) Parameter Estimates

	Parameter	standard error	t-ratio
$\psi$	0.9122	0.024	37.025
$\gamma$	2.734	0.113	8.833
$\ln L$	469.165		

The residuals are plotted in Figure 9.

[Insert Figure 9: Residuals, noncausal AR(1)]

The model has not only removed the serial correlation, as shown in Figure 10, but has also removed the asymmetry due to the bubble (see Figure 7).

[Insert Figure 10: ACF residuals, noncausal AR(1)]

The speculative subperiod is just characterized by a rather standard volatility clustering.

An additional autoregressive term can be introduced. We consider a noncausal AR(2) model:

$$y_t = \psi_1 y_{t+1} + \psi_2 y_{t+2} + e_t, \quad (4.5)$$

and estimate it by the approximated maximum likelihood.

Table 2: AR(2) Parameter Estimates

	Parameter	standard error	t-ratio
$\psi_1$	1.316	0.077	17.046
$\psi_2$	-0.401	0.065	-6.166
$\gamma$	2.433	0.112	7.894
$\ln L$	478.709		

The roots of the noncausal polynomial are 1.194 and 2.084 and the noncausal AR(2) is stationary. The residuals of the noncausal AR(2) satisfy also white noise features, as their autocorrelations are not statistically significant. The noncausal AR(2) provides good fit to the data and the value of its log likelihood function at the maximum is very close to that of the next model considered with the same number of parameters.

The next specification considered is a mixed autoregressive model MAR(1,1) with both causal and noncausal orders equal to 1. The estimated model is:

$$(1 - \phi L)(1 - \psi L^{-1})y_t = e_t. \quad (4.6)$$

The parameter estimates are provided in Table 3. Both roots of the polynomials lie outside the unit circle.

Table 3: MAR(1,1) Parameter Estimates

	Parameter	standard error	t-ratio
$\psi$	0.678	0.028	23.864
$\phi$	0.717	0.023	30.507
$\gamma$	2.559	0.109	8.552
$\ln L$	479.402		

The noncausal AR(2) process and the MAR(1,1) process are not equivalent from the modeling point of view. Noncausal parameters  $\psi_j$  (resp. causal parameters  $\phi_j$ ) have a significant impact on the rate of increase of the bubble (resp. decrease of the bubble). The noncausal AR(2) model is able to fit bubbles with two possible rates of increase, corresponding to  $\psi_1$  and  $\psi_2$ , but with sharp decrease due to the absence of causal autoregressive parameter. The mixed MAR(1,1) model is flexible enough to fit any asymmetric bubbles. For illustration purpose, we focus below on the mixed MAR(1,1) model.

### 4.3 The Causal and Noncausal Components

The MAR(1,1) process can be decomposed into a "causal" and a "noncausal" components [see e.g. Gouriéroux, Jasiak (2014)]:

$$y_t = \frac{1}{1 - \phi\psi}(u_t + \phi v_{t-1}), \quad (4.7)$$



where the noncausal component is defined by

$$u_t - \psi u_{t-1} = e_t, \quad (4.8)$$

and the causal component by

$$v_t - \phi v_{t-1} = e_t. \quad (4.9)$$

The causal component  $v_t$  (resp. the noncausal component  $u_t$ ) is a combination of current and lagged values (resp. of the current and future values) of the noise  $e_t$ . This explains the terminology used above.

In Figure 11, we provide the filtered noise and the filtered causal and noncausal components of the detrended series.

[Insert Figure 11: Component series of the MAR(1,1)]

The series of filtered  $e_t$  is close to a series of independent random variables, as suggests the ACF given in Figure 12. From the series of filtered  $e_t$ , we can see the dates of extreme shocks.

[Insert Figure 12: ACF of the MAR(1,1) components]

The component series determine jointly the bubble patterns. In particular, the component  $u$  with parameter  $\psi$  determines the growth phase of the bubble, while the  $v$  component and parameter  $\phi$  determine the bubble burst.

## 4.4 Prediction

Let us now consider the prediction performance of the MAR(1,1) with nonlinear deterministic trend. In particular, we will study its ability of predicting the future bubbles.

For this purpose we need to predict the future path of detrended process  $y$  at some horizon  $H$ , that is  $y_{T+1}, \dots, y_{T+H}$ , given the available information  $y_1, \dots, y_T$ . We are interested in nonlinear prediction of this path represented by the predictive density of  $y_{T+1}, \dots, y_{T+H}$ . This analysis depends on horizon  $H$ . It becomes more complex when  $H$  increases, but also more informative concerning the possible future downturns. There exist consistent approximations of this joint predictive density, based on a look-ahead estimator, which

admit closed form expressions. They can be used for small  $H$  to display the predictive densities and in general to compute any moment of the type:

$$E(a(y_{T+1}, \dots, y_{T+H})|y_1, \dots, y_T),$$

by simulation or numerical integration.

For  $s = 1$ , the predictive distribution at horizon  $H$  is defined as [Gourieroux, Jasiak(2014)]:

$$\begin{aligned} & \hat{\Pi}(u_{T+1}, \dots, u_{T+H}|\hat{u}_T) \\ &= \left\{ \hat{g}(\hat{u}_T - \hat{\psi}u_{T+1})\hat{g}(u_{T+1} - \hat{\psi}u_{T+2}) \dots \hat{g}(u_{T+H-1} - \hat{\psi}u_{T+H}) \sum_{t=1}^{T-1} \hat{g}(u_{T+H} - \hat{\psi}\hat{u}_t) \right\} \\ & \quad \left\{ \sum_{t=1}^{T-1} \hat{g}(\hat{u}_T - \hat{\psi}\hat{u}_t) \right\}^{-1} \end{aligned} \tag{4.10}$$

We display in Figure 13 the predictive density at horizon 1, in Figure 14 the joint predictive density at horizon 2, and its contour plot in Figure 15.

[Insert Figure 13: Predictive density at horizon 1]

[Insert Figure 14: Joint predictive density at horizon 2]

[Insert Figure 15: Contour plot of the joint predictive density at horizon 2]

We observe some asymmetry in the predictive density of  $y_{T+1}$ , which would not have been detected with a standard Gaussian ARMA model.

The joint predictive density at horizon 2 is much more informative. For the MAR(1,1) process, all the predictive densities depend on the information by means of two state variables only, which are  $y_T, u_T$ , or equivalently by  $y_T, y_{T-1}$ . The values of the state variables for the Bitcoin data are :  $u_T = 2.87$ ,  $y_{T-1}=9.64$ ,  $y_T = 12.27$ . Thus at the end of the observation period we are in an increasing phase of the detrended series. From the knowledge of the joint predictive density, we can infer the type of pattern that will follow that increasing phase. The series can continue to increase (the upper North-East semi-orthant in Figure 15 from the top point with coordinates of about  $(y_T, y_T) =$

(12.27, 12.27)), increase and then slowly decrease (the bottom North-East semi-orthant), or immediately decrease sharply (the bottom South-West semi-orthant) and so on.

The joint predictive density has a complicated pattern, far from Gaussian, with a strong dependence in extreme future risks in some directions. Its associated copula is close to an extreme value copula [see e.g. Balrema, Embrechts and Nolde (2013) for examples of extreme value copulas]. By considering Figure 15, we see that the probability of a continuing increase of  $y$  (North-East orthant) is rather high, but so is the probability of a sharp downturn at date  $T + 1$  (South-West orthant). However, the probability of a downturn at date  $T + 2$  (South-East orthant) is small. Thus the joint predictive density can be used to recognize the future pattern of  $y$  by comparing the likelihood of the different scenarios, in particular to evaluate the probability of the downturn at dates  $T+1$ ,  $T+2$ , etc. The above discussion based on the graphical representation is limited to horizon 2. However, The probabilities of the different types of future patterns can be evaluated numerically for larger  $H$ .

[Insert Figure 16: Predicted Dynamics]

(to be completed)

## 5 Conclusion

The causal-noncausal autoregressive models have been proposed as nonlinear dynamic models that are able to fit speculative bubbles. We applied this methodology to analyse the Bitcoin/USD exchange rates over the period February-July 2013. Indeed, speculative bubbles appeared in that period and could be used to calibrate the parameters of the model. We have considered a mixed model with both causal and noncausal orders equal to 1, estimated the parameters by the Approximated Maximum Likelihood, filtered the underlying components of the process to better understand the type of existing bubbles. Next, we built the joint predictive density of the future path. This joint density was used to predict the future patterns of the process, a kind of model-based chartist approach, in particular to evaluate the likelihood of the future dates of downturn.

The series of Bitcoin/USD exchange rates has been used as a playground for analyzing the relevance of the causal-noncausal modeling to capture bubble phenomena. It was

typically an example of highly speculative emerging market. Recently, several exchange platforms have closed, temporarily <sup>8</sup> or definitely. Other platforms were submitted to regulations. There is clearly a need of supervision to better protect the investors in bitcoins against the theft of their bitcoins, but also against the speculative behavior of large bitcoin holders. This supervision will likely make disappear the previously observed speculative bubbles and perhaps the market for this electronic currency itself.

However, there will still exist a large number of financial markets, not necessarily emerging, with frequently appearing bubbles. Examples are the markets for commodity futures, and the markets with high frequency trading. These are potential applications for the causal-noncausal model presented in this paper.

---

<sup>8</sup>The French platform Bitcoin-Central has been closed for five months in 2013 due to hackers attack. Nevertheless the customers had still the possibility to withdraw their bitcoins.

## REFERENCES

- Andrews, B., Calder, M. and R. Davis (2009): "Maximum Likelihood Estimation for  $\alpha$ -Stable Autoregressive Processes", *Annals of Statistics*, 37, 1946-1982.
- Andrews, B. and R. Davis (2013): "Model Identification for Infinite Variance Autoregressive Processes", *Journal of Econometrics*, 172, 222-234.
- Balkema, G., Embrechts, P. and N. Nolde (2013): "The Shape of Asymptotic Dependence" in *Springer Proceedings in Mathematics and Statistics*, special volume, "Prokhorov and Contemporary Probability Theory", eds. A. Shirayev, S. Varadhan and E. Presman, 33, 43-67.
- Blanchard, O. (1979): "Speculative Bubbles: Crashes and Rational Expectations", *Economic Letters*, 3, 387-389.
- Blanchard, O. and M. Watson, (1982): "Bubbles, Rational Expectations and Financial Markets" in *Crisis in the Economic and Financial Structure*, Ed. P. Wachtel, 295-315, Lexington, MA.
- Breidt, F. and R. Davis (1992): "Time Reversibility, Identifiability and Independence of Innovations for Stationary Time Series", *Journal of Time Series Analysis*, 13, 273-390
- Breidt, F., Davis, R., Lii, K. and M. Rosenblatt (1991): "Maximum Likelihood Estimation for Noncausal Autoregressive Processes", *Journal of Multivariate Analysis*, 36, 175-198
- Brunnermeier, M. (2001): "Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis and Herding", Oxford University Press, Oxford.
- Cheng, Q. (1992): "On the Unique Representation of Non-Gaussian Linear Processes", *Annals of Statistics*, 20, 1143-1145.
- Davis, R. and S. Resnick (1985): "Limit Theory for Moving Averages of Random Variables with Regularly Varying Tail Probabilities", *The Annals of Probabilities*, 13, 179-195.
- Davis, R. and S. Resnick (1986): "Limit Theory for the Sample Covariance and Correlation Functions of Moving Averages", *The Annals of Statistics*, 14, 533-558.
- Davis, R. and L. Song (2012): "Noncausal Vector AR Processes with Application to Economic Time Series", DP Columbia University

Evans, G. (1991): "Pitfalls in Testing for Explosive Bubbles in Asset Prices", *American Economic Review*, 81, 922-930.

Gourieroux, C. and J. Jasiak (2014): "Filtering, Prediction and Estimation of Noncausal Processes", CREST DP.

Gourieroux, C. and J.M. Zakoian (2013a): "Explosive Bubble Modelling by Noncausal Cauchy Autoregressive Process", CREST, DP

Gourieroux, C. and J.M. Zakoian (2013b): "On Uniqueness of Moving Average Representation of Heavy Tailed Stationary Processes", CREST, DP

Lanne, M. and P. Saikkonen (2011): "Noncausal Autoregressions for Economic Time Series", *Journal of Time Series Econometrics*, 3 (3), Article 2.

Lanne, M., Luoto, J. and P. Saikkonen (2010) "Optimal Forecasting of Nonlinear Autoregressive Time Series", *International Journal of Forecasting*, 28, 623-631.

Lanne, M. and P. Saikkonen (2013): "Noncausal Vector Autoregression", *Econometric Theory*, 29, 447-481.

Muth, J. (1961): "Rational Expectations and the Theory of Price Movements", *Econometrica*, 29, 315-335.

Phillips, P., Shi, S. and J. Yu (2012): "Testing for Multiple Bubbles", DP Cowles Foundation, 1843.

Phillips, P., Wu, Y. and J.. Yu (2011): "Explosive Behavior in the 1990s Nasdaq: When Did Exuberance Escalate Asset Values?", *International Economic Review*, 52, 201-226.

Rosenblatt, M. (2000): "Gaussian and Non-Gaussian Linear Time Series and Random Fields", Springer Verlag, New York.

## Website References

- [1] Velde, F.R., (2013): Bitcoin A primer, Chicago Fed Letter. Retrieved from [http://www.chicagofed.org/digital\\_assets/publications/chicago\\_fed\\_letter/2013/cfldecember2013\\_317.pdf](http://www.chicagofed.org/digital_assets/publications/chicago_fed_letter/2013/cfldecember2013_317.pdf)
- [2] Li. Shan, (2013): Bitcoin now accepted as tuition payment at a Cyprus University, Los Angeles Times. Retrieved from <http://www.latimes.com/business/money/la-fi-mo-cyprus-university-bitcoin-20131120,0,3194094.story\#axzz2mXKIff7E>
- [3] Flitter, E. (2013): FBI shuts alleged online drug marketplace, Silk Road, Reuters. Retrieved from <http://www.reuters.com/article/2013/10/02/us-crime-silkroad-raid-idUSBRE9910TR20131002>
- [4] Tagaris. K (2013): Cyprus details heavy losses for major bank customers, Reuters. Retrieved from <http://www.reuters.com/article/2013/03/30/us-cyprus-parliament-idUSBRE92G03I20130330>
- [5] Sparshott, J. (2013): Web Money Gets Laundering Rule, The Wall Street Journal. Retrieved from <http://online.wsj.com/news/articles/SB10001424127887324373204578374611351125202>
- [6] Bitcoin Market Capitalization. Retrieved from Bitcoin Block Explorer - Blockchain [http://blockchain.info/charts/market\\_cap](http://blockchain.info/charts/market_cap)
- [7] [8.] Davis. J (2011): The Crypto-Currency, The New Yorker. Retrieved from [http://www.newyorker.com/reporting/2011/10/10/111010fa\\_fact\\_davis](http://www.newyorker.com/reporting/2011/10/10/111010fa_fact_davis)
- [8] Bitcoin Market Capitalization. Retrieved from Bitcoin Block Explorer - Blockchain [http://blockchain.info/charts/market\\_cap](http://blockchain.info/charts/market_cap)
- [9] Liu, J. (2013): BTC China the world's largest Bitcoin trading platform, ZD Net. Retrieved from <http://www.zdnet.com/btc-china-the-worlds-largest-bitcoin-trading-platform-7000023316/>

- [10] Trade (2013, December 4). Bitcoin Wiki. Retrieved December 4, 2013, from <https://en.bitcoin.it/wiki/Trade>
- [11] Controlled Supply (2013, December 4). Bitcoin Wiki. Retrieved December 4, 2013, from [https://en.bitcoin.it/wiki/Controlled\\_Currency\\_Supply](https://en.bitcoin.it/wiki/Controlled_Currency_Supply)
- [12] Introduction (2013, December 4). Bitcoin Wiki. Retrieved December 4, 2013, from <https://en.bitcoin.it/wiki/Introduction>
- [13] Introduction (2013, December 4). Bitcoin Wiki. Retrieved December 4, 2013, from <https://en.bitcoin.it/wiki/Introduction>
- [14] Litecoin Block Explorer Charts (2013, December 4). Litecoin Block Explorer. Retrieved December 4, 2013, from <http://ltc.block-explorer.com/charts>



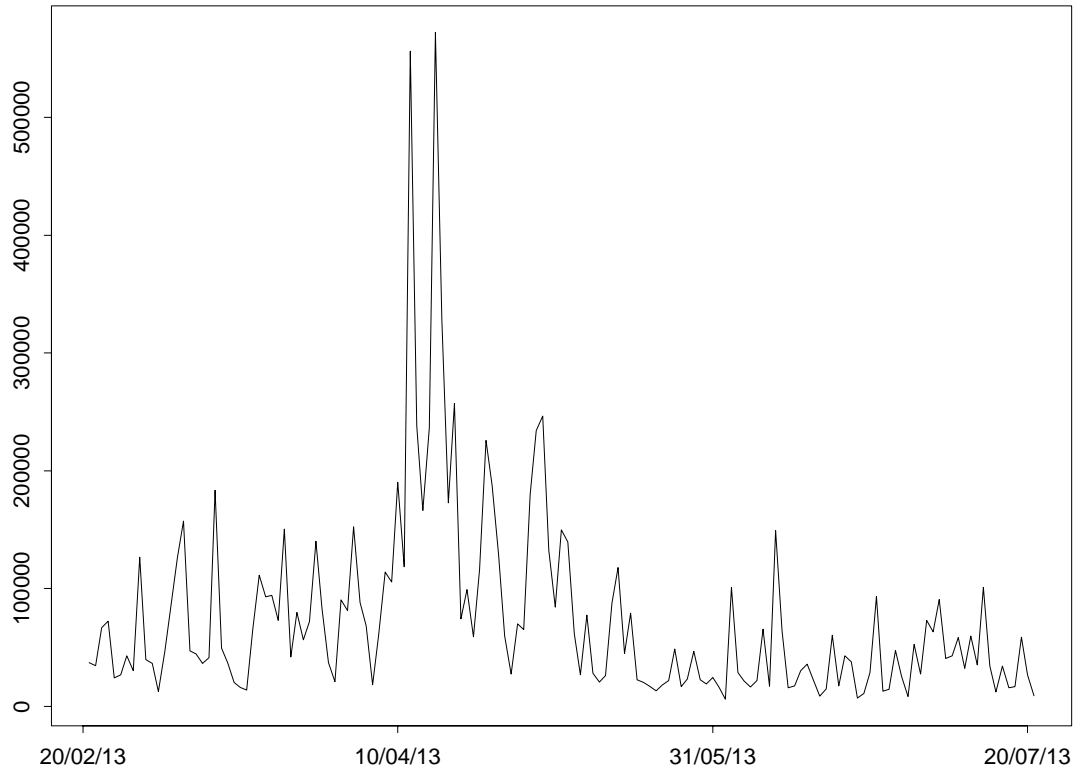


Figure 1: Bitcoin volume, Feb-July 2013

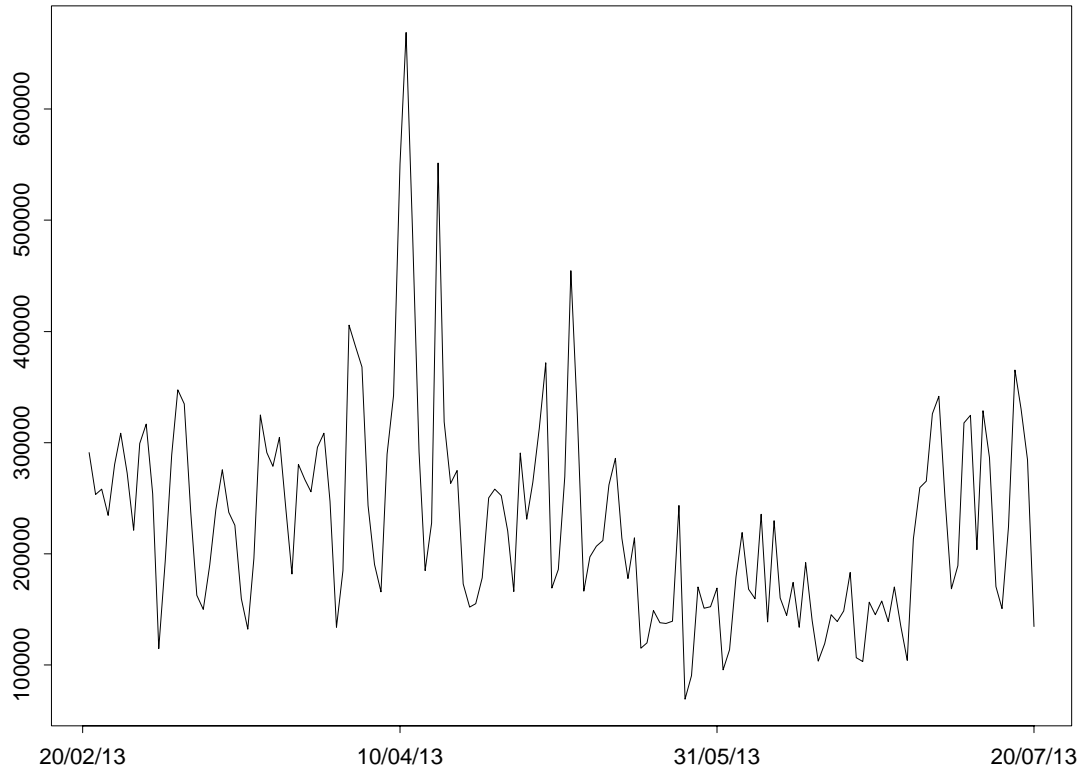


Figure 2: Bitcoin transactions, Feb-July 2013

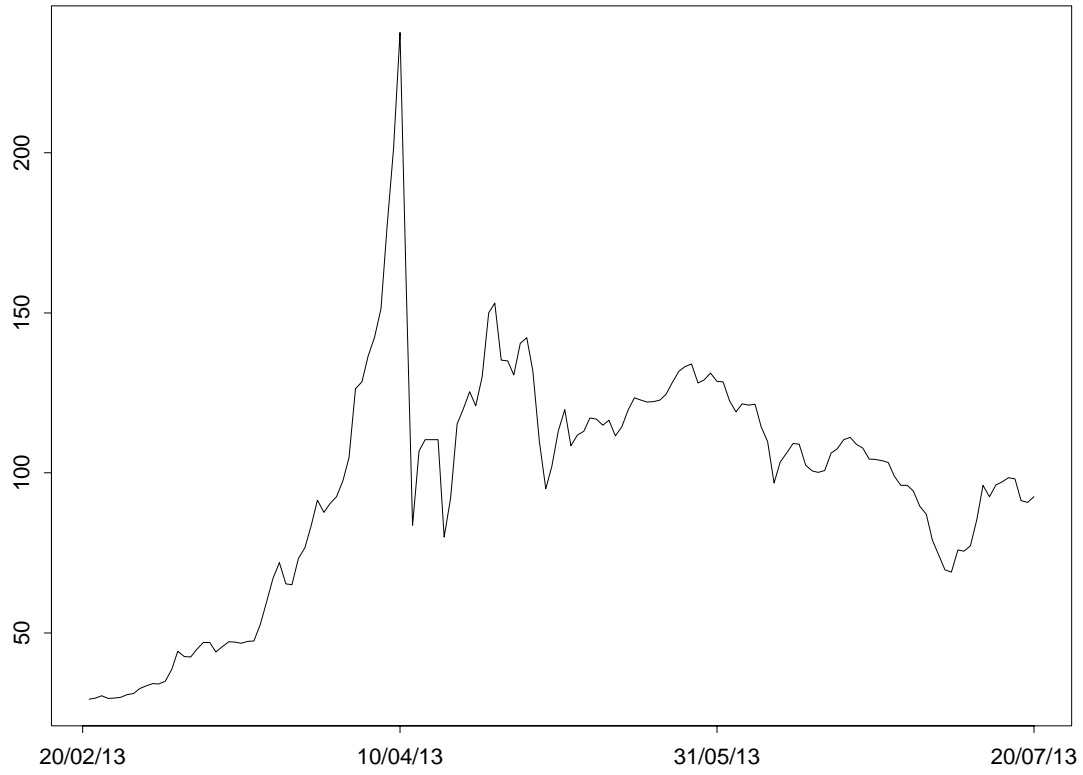


Figure 3: Bitcoin/USD exchange rate, Feb-July 2013

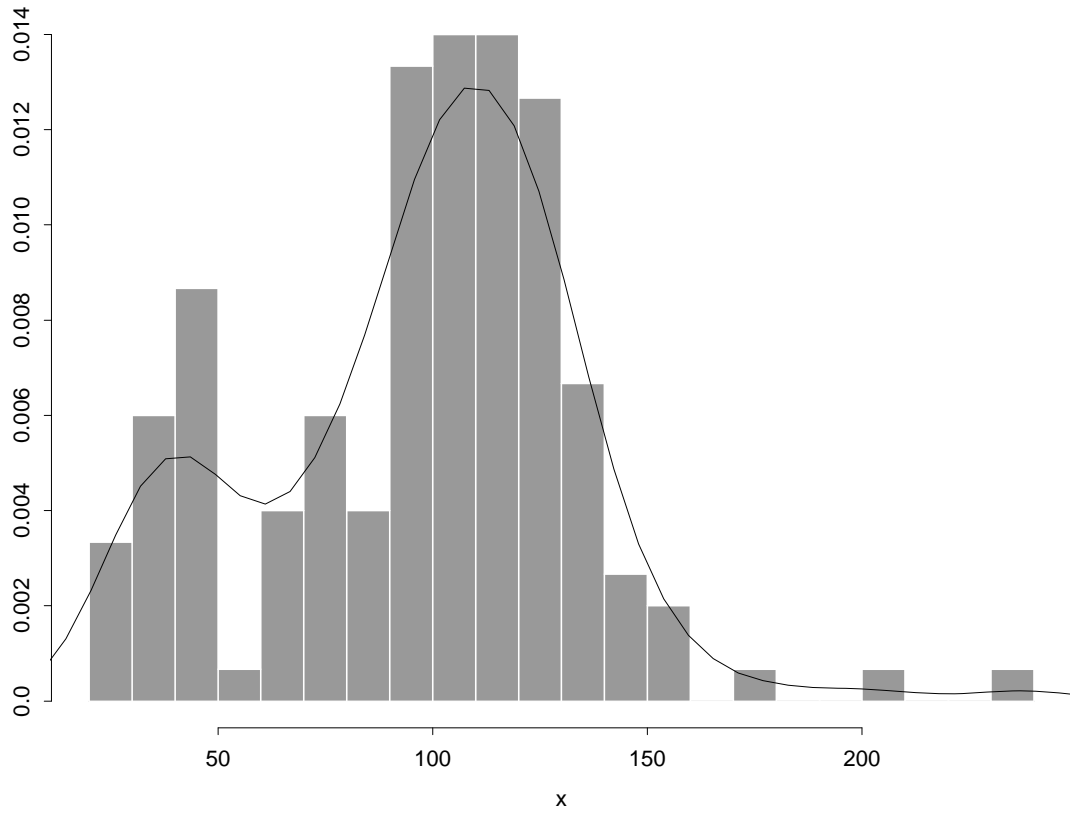


Figure 4: Bitcoin, histogram

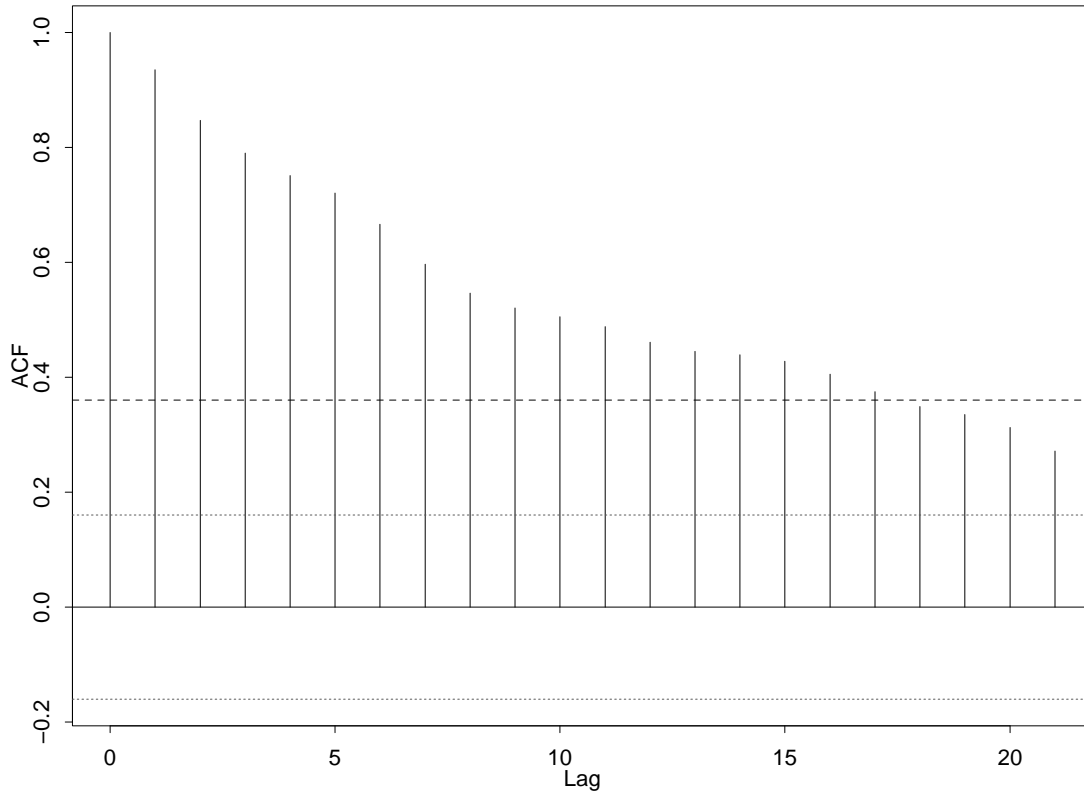


Figure 5: ACF, Bitcoin

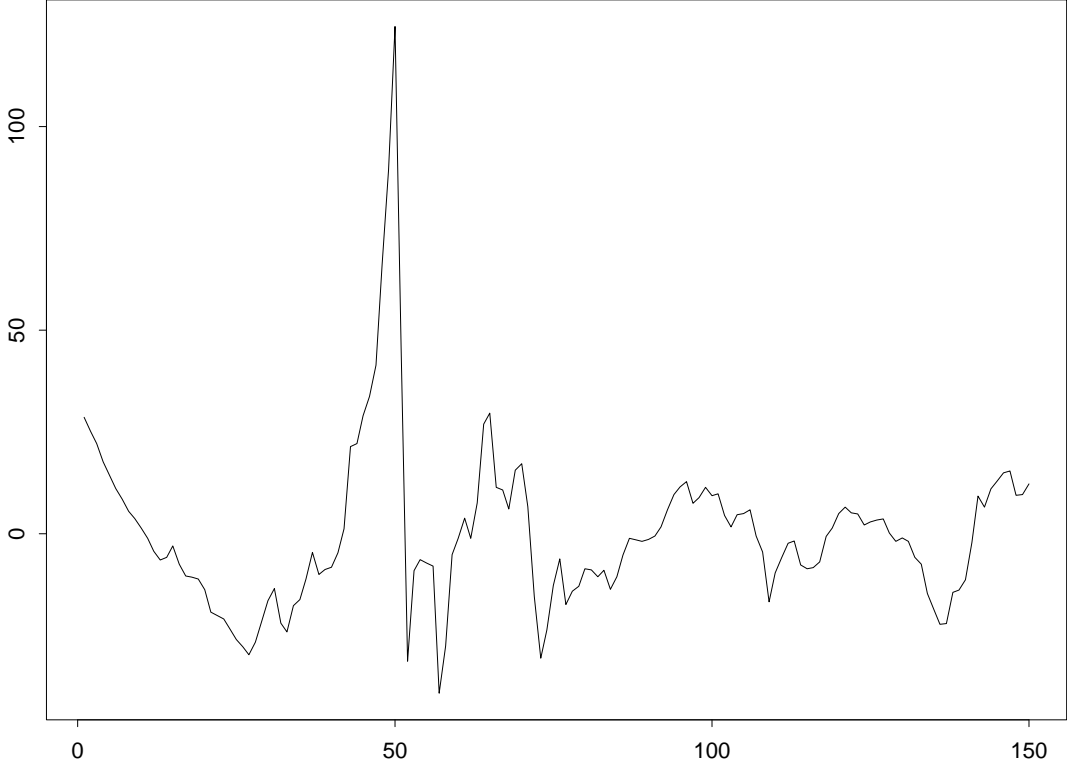


Figure 6: Detrended series

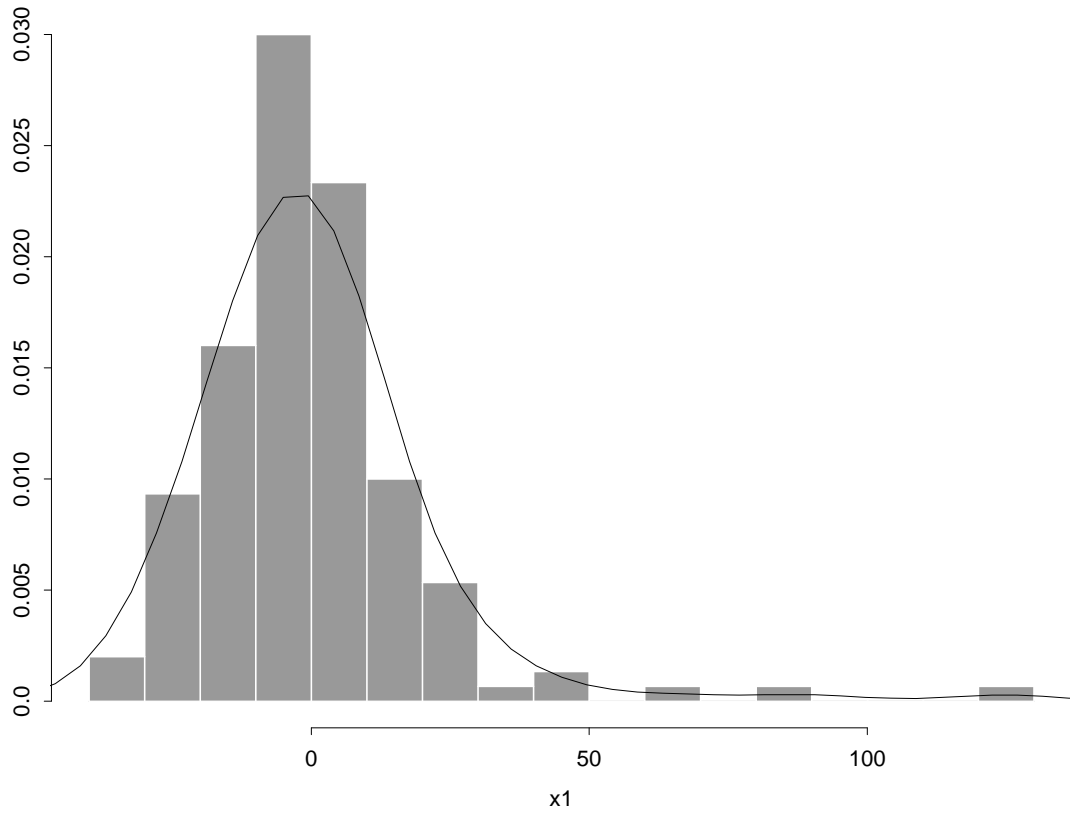


Figure 7: Detrended series, histogram

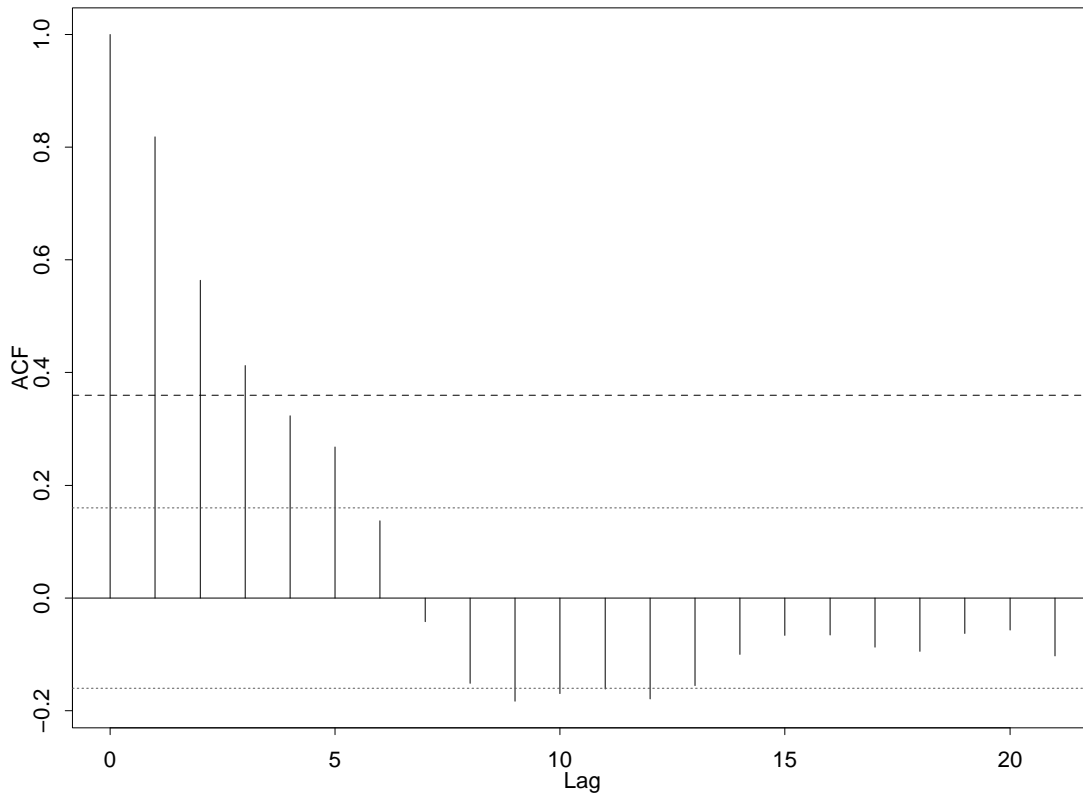


Figure 8: Detrended series, ACF



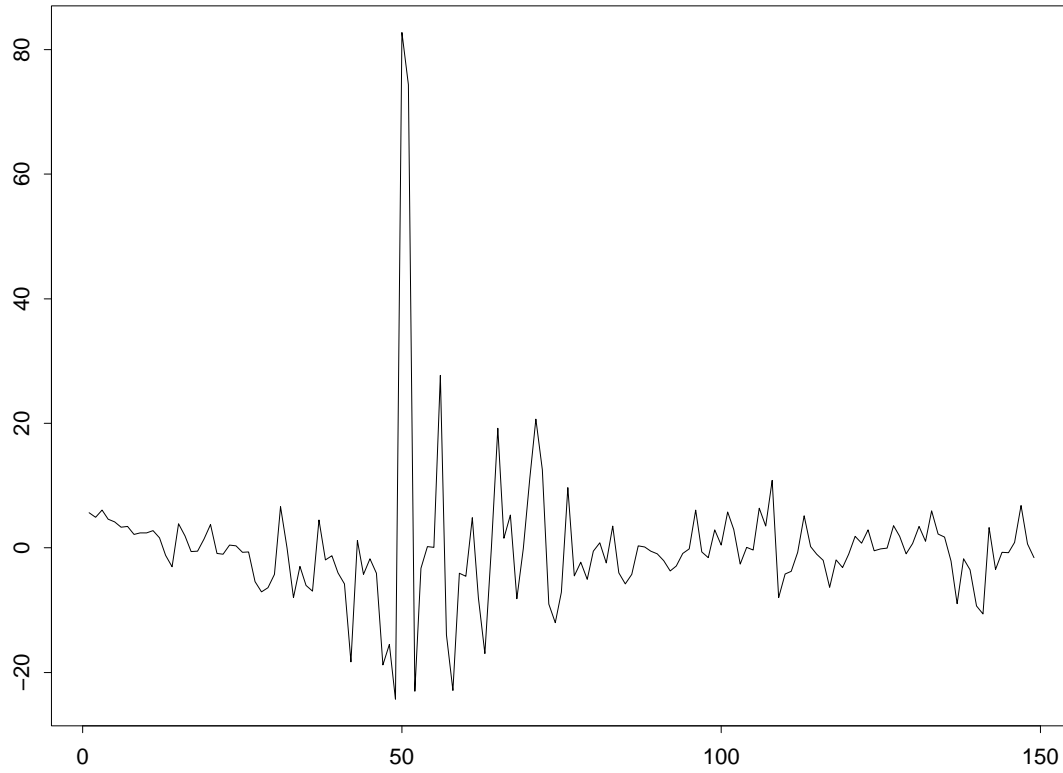


Figure 9: Residuals, noncausal AR(1)

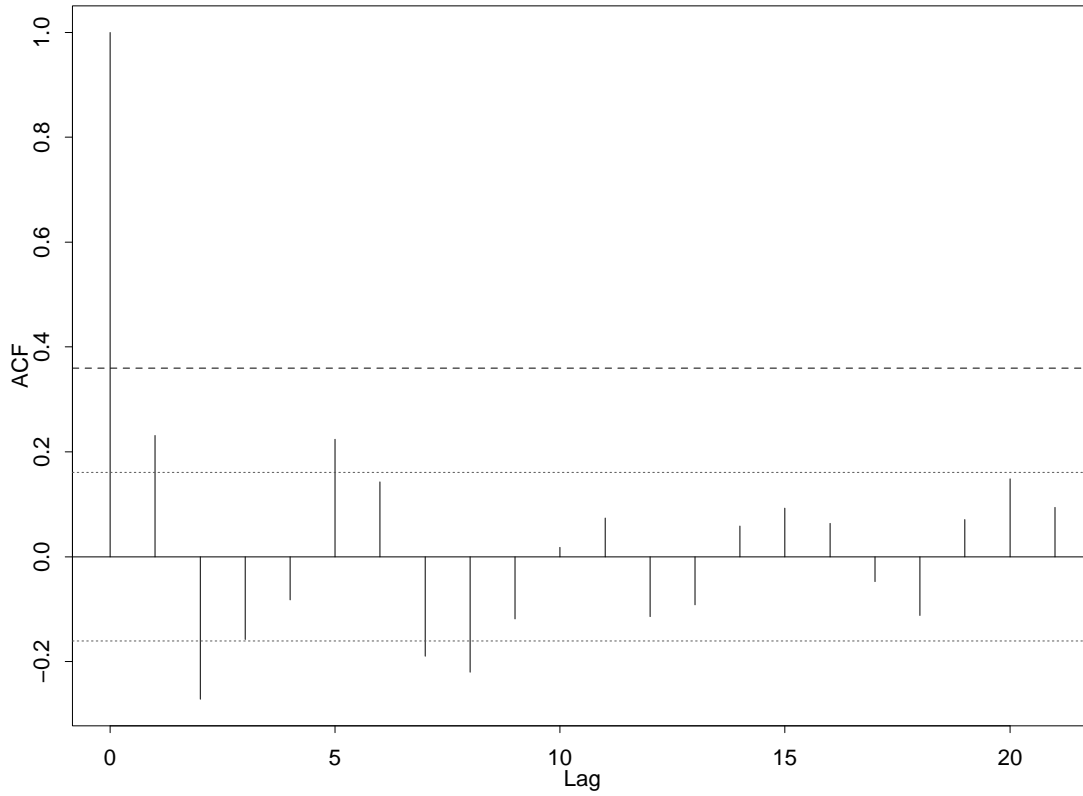


Figure 10: ACF residuals, noncausal AR(1)

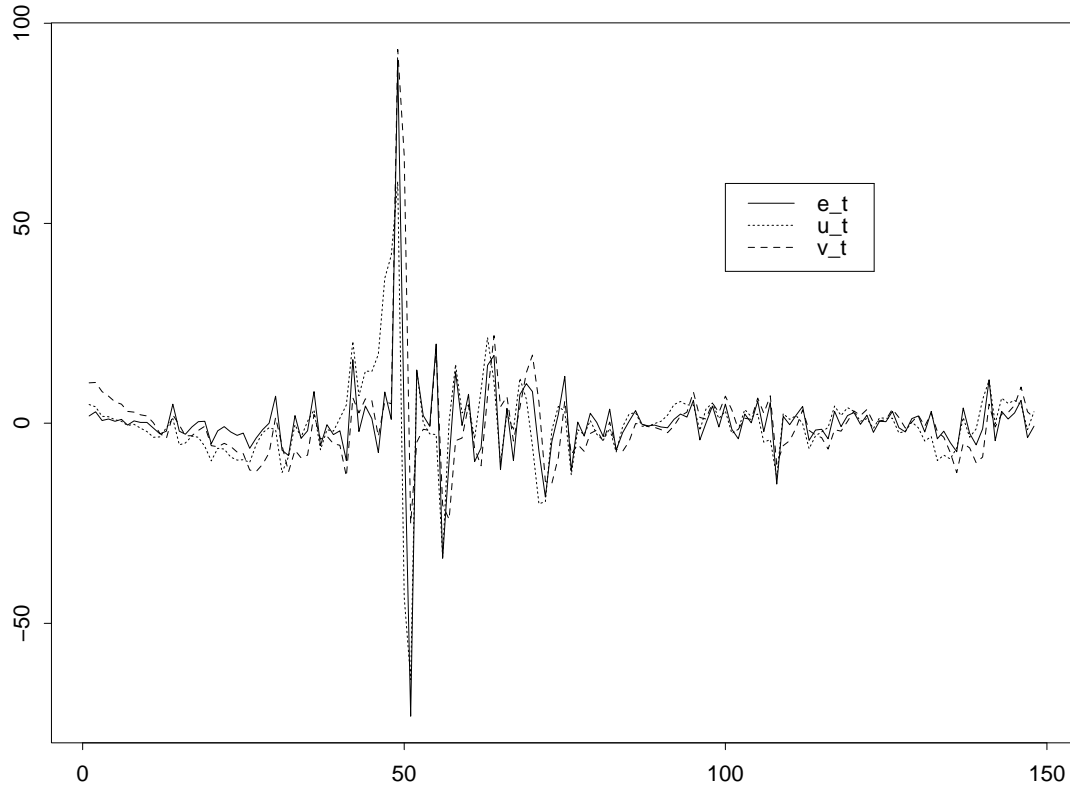


Figure 11: Components series of the MAR(1,1)

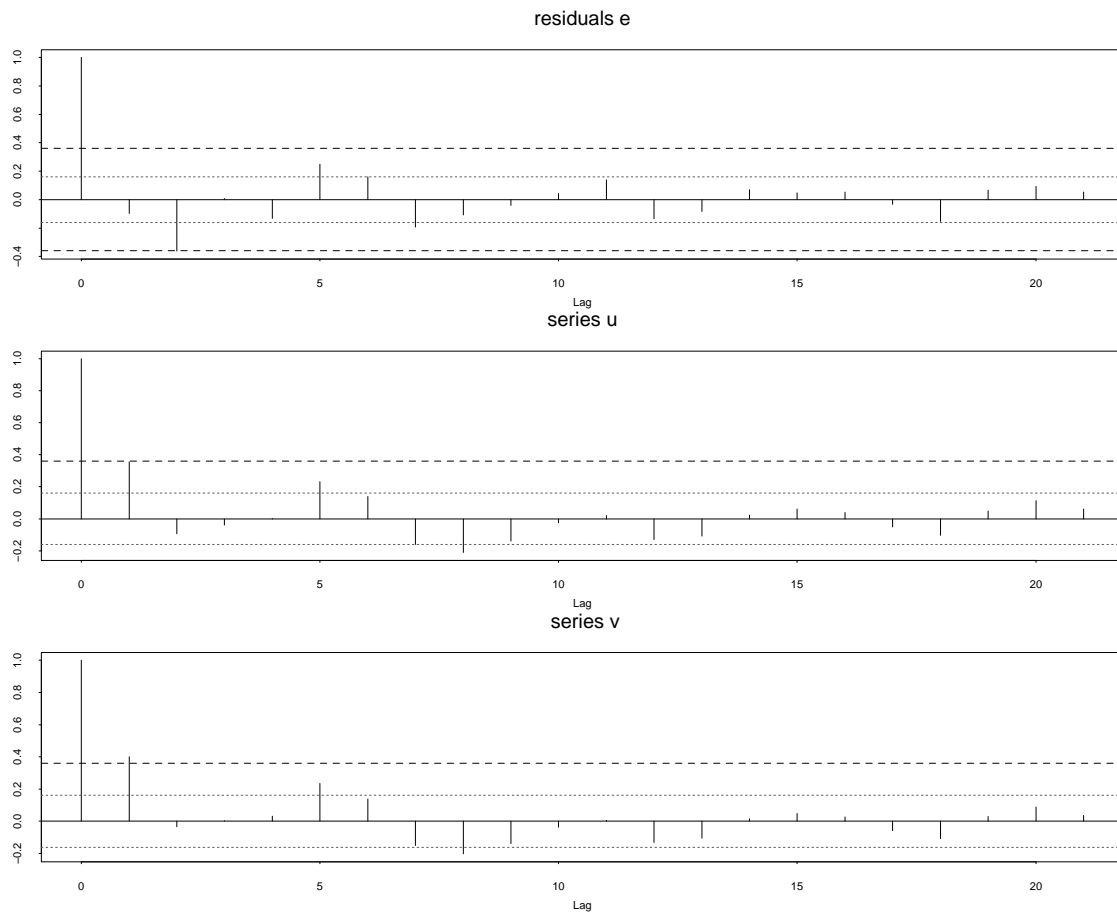


Figure 12: ACF of the MAR(1,1) components

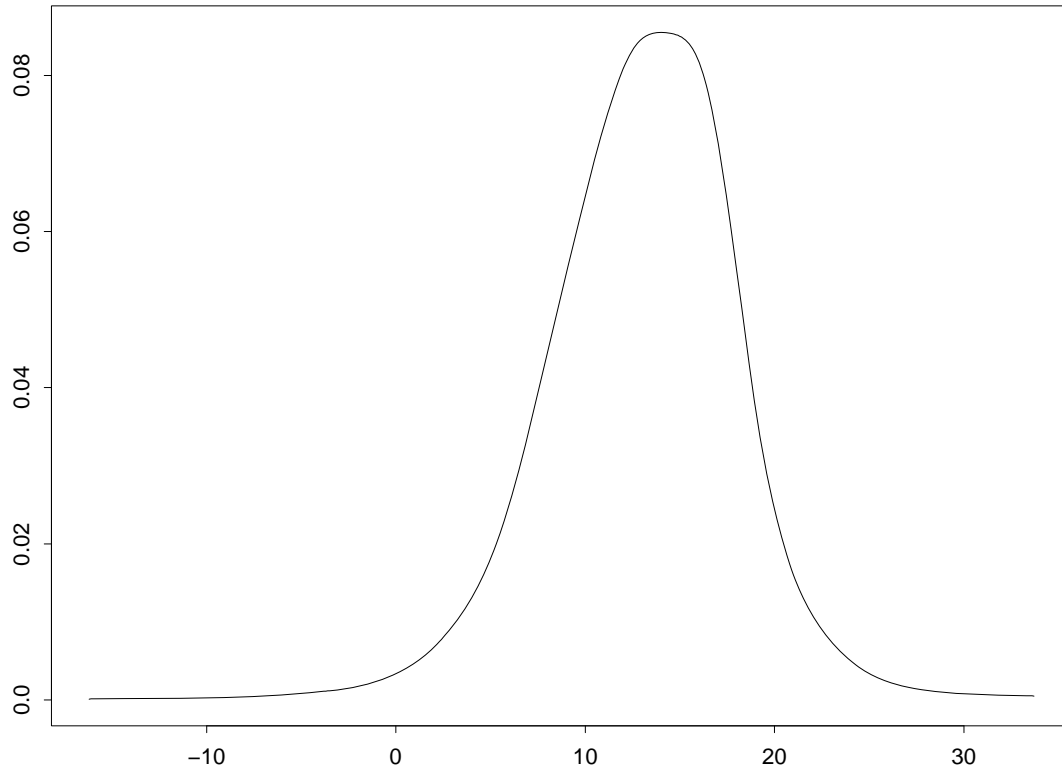


Figure 13: Predictive density at horizon 1

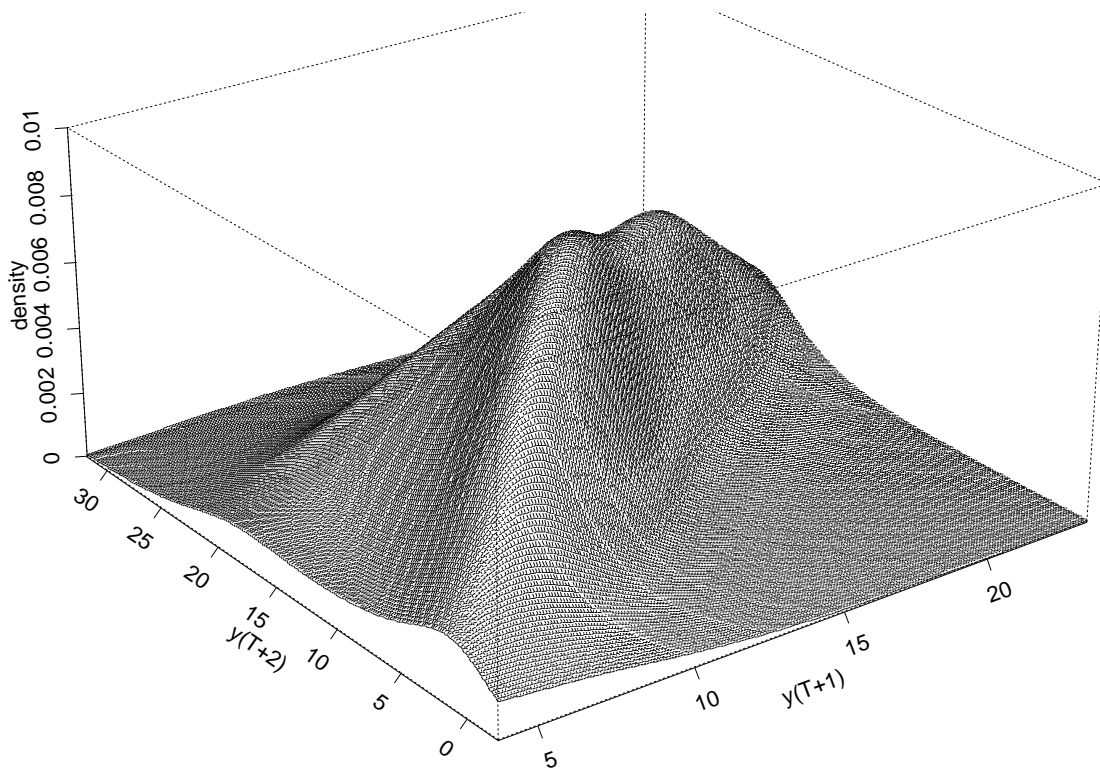


Figure 14: Joint predictive density at horizons 1 and 2

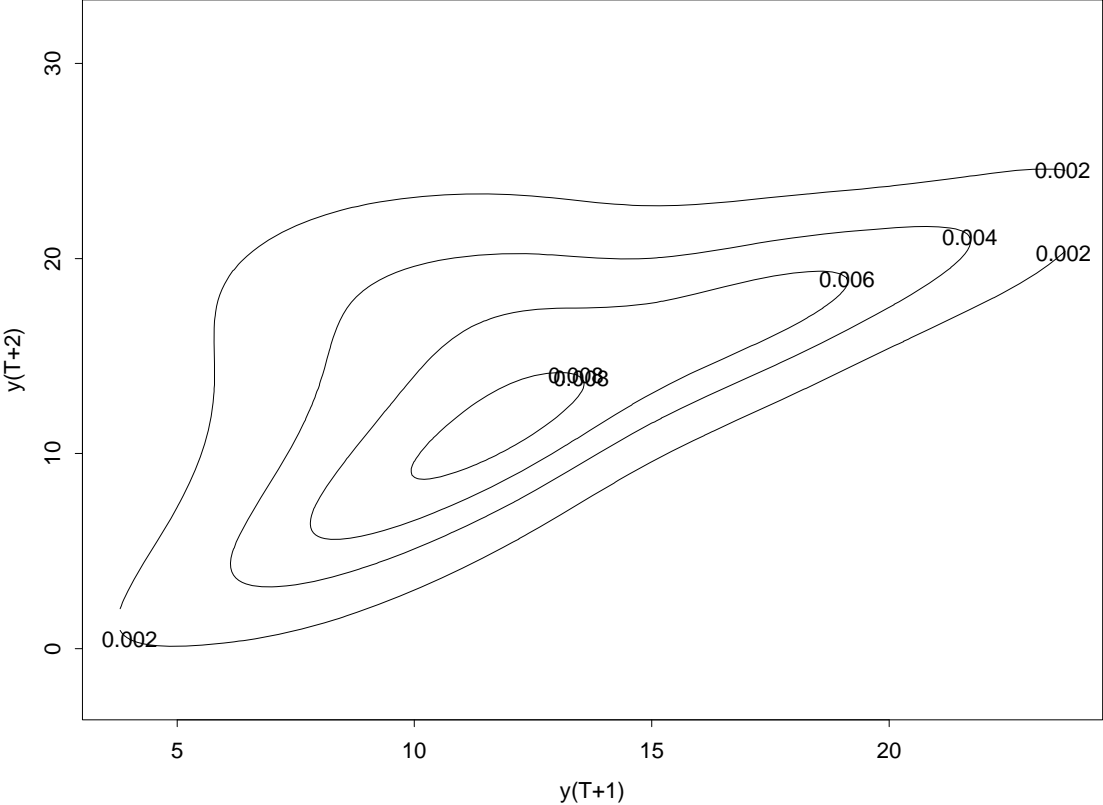


Figure 15: Contour plot of predictive density

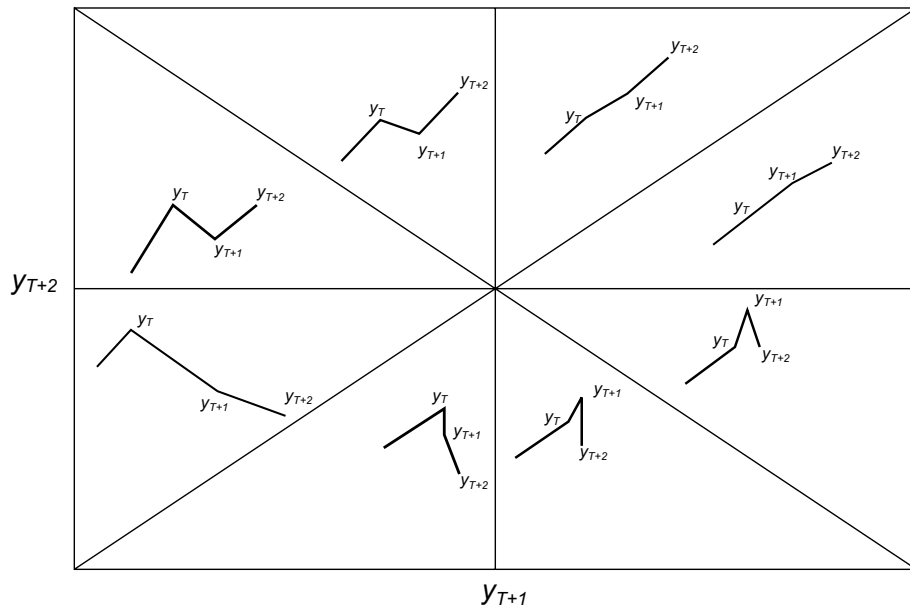


Figure 16: Predicted Dynamics