Summer 2012

Sieve bootstrap based prediction intervals and unit root tests for time series

Maduka Rupasinghe

Department: Mathematics and Statistics

SIEVE BOOTSTRAP BASED PREDICTION INTERVALS AND UNIT ROOT TESTS FOR TIME SERIES

by

MADUKA NILANGA RUPASINGHE

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS & STATISTICS

2012

Approved by:

Dr. V.A. Samaranayake, Advisor
Dr. Robert Paige
Dr. Xuerong Wen
Dr. Akim Adekpedjou
Dr. Gregory Gelles

DEDICATION

I would like to dedicate this Doctoral dissertation to my wife, Hiroshi Abhayawickrama and to my parents, Sunitha Perera and Chandrasiri Rupasinghe. There is no doubt in my mind that without their continued support and counsel I could not have completed this process.

PUBLICATION DISSERTATION OPTION

This dissertation has been prepared in the styles utilized by the Statistics and Probability Letters, Journal of Statistical Computations and Simulations, Journal of Statistical Planning and Inference and Journal of Business and Economic Statistics. Pages 17-34 and 83-94 will be submitted to the Statistics and Probability Letters; Pages 35-58 to the Journal of Statistical Computations and Simulations; Pages 59-82 to Journal of Statistical Planning and Inference; Pages 95-123 to Journal of Business and Economic Statistics. Appendices A and B and the Bibliography have been added for purposes normal to dissertation writing.

ABSTRACT

The application of the sieve bootstrap procedure, which resamples residuals obtained by fitting a finite autoregressvie (AR) approximation to empirical time series, to obtaining prediction intervals for integrated, long-memory, and seasonal time series as well as constructing a test for seasonal unit roots, is considered. The advantage of this resampling method is that it does not require knowledge about the underlying process generating a given time series and has been shown to work well for $ARMA$ processes.

We extend the application of the sieve bootstrap to $ARIMA$ and $FARIMA$ processes. The asymptotic properties of the sieve bootstrap prediction intervals for such processes are established, and the finite sample properties are examined by employing Monte Carlo simulations. The Monte Carlo simulation study shows that the proposed method works well for both $ARIMA$ and $FARIMA$ processes.

Following the existing sieve bootstrap frame-work for testing unit roots for non-seasonal processes, we propose new bootstrap-based unit root tests for seasonal time series. In this procedure, the bootstrap distributions of the well known Dickey-Hasza-Fuller (DHF) seasonal test statistics are obtained and utilized to determine the critical points for the test. The asymptotic properties of the proposed method are established and a Monte Carlo simulation study is employed to demonstrate that the proposed unit root tests yield higher powers compared to the DHF test. Also, a sieve bootstrap method is implemented to obtaining prediction intervals for time series with seasonal unit roots. The asymptotic properties of the proposed prediction intervals are established and a Monte Carlo simulation study is carried out to examine the finite sample validity.

Finally, we derive expressions for the asymptotic distributions of the Dickey-Fuller (DHF) type test statistics, under weakly dependent errors and show that they can be expressed as functional of the standard Brownian motions. Currently, the asymptotic results are available only for non-seasonal time series.

ACKNOWLEDGMENTS

I would like to acknowledge the support and assistance given me by my advisor, Prof. V.A. Samaranayake and my graduate committee. I am thankful to the faculty members of the department of Mathematics and Statistics, Missouri University of Science and Technology, for their help and advise in many ways to bring my academic pursuits to a success. Finally, I would like to thank my wife, Hiroshi, for her support and encouragement. I could not have completed this effort without her assistance, tolerance, and enthusiasm.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

Time series analysis is an important area of specialization in the field of Statistics. It also plays a prominent role in economics and finance. Other areas where time series methodology is utilized include Geophysical Sciences and branches of electrical engineering, such as signal processing. In all these areas, time series methodologies are used not only to model empirical data observed over time, but also to forecast future values of such processes. In this dissertation, both these aspects would be studied, with major emphasis giving to the use of a re-sampling technique know as the Sieve Bootstrap, to obtain estimates for the distribution of certain statistics as well as to obtain prediction intervals for future values of a time series. The proposed sieve bootstrap procedures provide improvements to existing methodologies for obtaining prediction intervals for specific types of processes and testing for the presence of a certain seasonal structure in the underlying time series model.

In order to discuss the specific classes of time series for which prediction intervals or statistical hypothesis tests are proposed, a basic background in the terminology and models used in time series analysis is required. As such, preliminary concepts and definitions are given in the remainder of this section.

1.1. INTRODUCTION TO TIME SERIES ANALYSIS

The followings are some fundamental terms used in time series analysis (see Brockwell and Davis (1991)). We begin with the definition of a time series and then go onto introduce various concepts such as stationarity and specific time series models that will be considered in the following sections.

Definition 1.1. *(Stochastic Process). A stochastic process is a family of random variables* $\{X_t, t \in T\}$ *defined on a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, where* T *denotes an index set, which is usually a subset of* $\mathbb{R}$*.*

When T denotes a set of points in time, $\{X_t, t \in T\}$ is called a time series. If in addition, $T \subseteq \mathbb{Z}$, it is called a discrete time series.

Definition 1.2. *(Realization of a Stochastic Process). The functional values* $\{X_\cdot(\omega), \omega \in \Omega\}$ *on* T *are known as a realizations or sample-paths of the process* $\{X_t, t \in T\}$.

The term "time series" is used to mean both the data and the process of which it is a realization. The most important behavioral characteristic of a given time series is its dependent structure, which is defined below.

From here on we shall assume that $T = \mathbb{Z}$ or $T = \mathbb{N}_0$. Also note that we shall use the notations $\{X_t, t \in T\}$ and $\{X_t\}_{t\in T}$ interchangeably in the reminder of this thesis.

Definition 1.3. *(The Covariance Function). If* $\{X_t, t \in T\}$ *is a time series such that* $Var(X_t) < \infty$ *for each* $t \in T$*, then the covariance function* $\gamma_X(\cdot, \cdot)$ *of* $\{X_t\}$ *is defined by*

$$\gamma_X(r, s) := Cov(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)], r, s \in T.$$

Definition 1.4. *(Stationarity). The time series* $\{X_t, t \in T\}$*, with index set* $\mathbb{Z} = \{0, \pm1, \pm2, ...\}$*, is said to be stationary if*

(i) $E|X_t|^2 < \infty$ *for all* $t \in \mathbb{Z}$,

(ii) $EX_t = m$ *for all* $t \in \mathbb{Z}$*, and*

(iii) $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ *for all* $r, s, t \in \mathbb{Z}$.

Note that the stationarity defined above is sometimes referred as weak stationarity, covariance stationarity, stationarity in the wide sense, or second-order stationarity. The covariance function of a stationary time series , known as "autocovariance function" (ACVF), is reduced to a function of a single variable as given below.

$$\gamma_X(h) := \gamma_X(h, 0) = Cov(X_{t+h}, X_t) \text{ for all } t, h \in \mathbb{Z}.$$

This autocovariance function has the following elementary properties.

Proposition 1.5. *If $\gamma_X(\cdot)$ is the autocovariance function of a stationary processes $\{X_t, t \in \mathbb{Z}\}$, then*

$$\gamma_X(0) \geq 0,$$
$$|\gamma_X(h)| \leq \gamma_X(0) \text{ for all } h \in \mathbb{Z}, \text{ and,}$$
$$\gamma_X(h) = \gamma_X(-h) \text{ for all } h \in \mathbb{Z}.$$

Proof. See Brockwell and Davis (1991), page 26. □

The autocorrelation function (ACF) of a stationary process $\{X_t, t \in \mathbb{Z}\}$ is then defined as,

$$\rho_X(h) := \frac{\gamma_X(h)}{\gamma_X(0)} \text{ for all } h \in \mathbb{Z}.$$

The characterization of the autocovariance function is given in the following theorem. The proof can be found in Brockwell and Davis (1991), page 27.

Theorem 1.6. *A real-valued function defined on the integers is the autocovariance function of a stationary time series if and only if it is even and non-negative definite.*

In practice, these unknown autocovariance and autocorrelation functions are estimated based on the empirically observed data $\{x_1, x_2, ..., x_n\}$ using the following definitions.

Definition 1.7. *(The Sample Autocovariance Function). The sample autocovariance function of $\{x_1, x_2, ..., x_n\}$ is defined by*

$$\hat{\gamma}(h) := n^{-1} \sum_{j=1}^{n-h} (x_{j+h} - \bar{x})(x_j - \bar{x}), \qquad 0 \leq h < n,$$

and $\hat{\gamma}(h) = \hat{\gamma}(-h), -n < h \leq 0$, where $\bar{x}$ is the sample mean determined by $\bar{x} := n^{-1} \sum_{j-1}^{n} x_j$.

The following defines the White Noise process, which is assumed for the innovations (errors) in certain types of time series models.

Definition 1.8. *(White Noise Process). The process* $\{X_t\}$ *is said to be white noise if,*

$$\gamma_X(h) = \begin{cases} \sigma^2 & \text{for } h = 0. \\ 0 & \text{otherwise.} \end{cases}$$

In the following, time series models that are extensively used in the literature are presented. These are sometimes called Box and Jenkins models because of their introduction to the wider audience of empirical time series analysts by Box and Jenkins (1976).

1.2. *ARMA* PROCESSES

Definition 1.9. *(ARMA Process). A real-valued process* $\{X_t\}_{t\in\mathbb{Z}}$ *is said to be an Autoregressive Moving Average (ARMA(p,q)) process if it is stationary and satisfies*

$$\alpha(B)(X_t - \mu) = \theta(B)\epsilon_t, t \in \mathbb{Z},$$

where $\alpha(z) = 1 - \alpha_1 z - ... - \alpha_p z^p$ *and* $\theta(z) = 1 + \theta_1 z + ... + \theta(q)z^q$ *represent autoregressive and moving average polynomials of degrees p and q respectively. The innovations (or the error terms),* $\{\epsilon_t\}$*, are assumed to be zero-mean white noise with variance* σ^2*. The mean of the process is* $\mu = E[x_t]$ *for all t. The back-shift operator, B, is defined by* $B^k x_t = x_{t-k}$ *for* $k \in \mathbb{N}$*.*

When the order of autoregressive polynomial p is equal to 0, we obtain the class of Moving Average Processes. Similarly, the class of Autoregressive Processes is obtained when $q = 0$.

Sometimes, it is necessary to represent a given $ARMA$ time series as an infinite order moving average of the current and past innovations. When a time series can be represented in this manner it is called a causal process.

Definition 1.10. *(Causality). An* $ARMA(p, q)$ *process defined in Definition 1.9 is said to be causal if there exists a sequence of constants* $\{\psi_j\}_{j\in\mathbb{N}_0}$ *such that* $\sum_{j=0}^{\infty} |\psi_j| < \infty$ *and*

$$X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \qquad \text{for } t \in \mathbb{Z}. \tag{1.1}$$

The following theorem from Brockwell and Davis (1991, page 85) gives necessary and sufficient conditions for an $ARMA$ process to be causal.

Theorem 1.11. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an $ARMA(p,q)$ process for which the polynomials $\alpha(.)$ and $\theta(.)$ have no common zeroes. Then $\{X_t\}_{t\in\mathbb{Z}}$ is causal if and only if $\alpha(z)\neq 0$ for all $z\in\mathbb{C}$ such that $|z|\leq 1$. The coefficients $\{\psi_j\}_{j\in\mathbb{N}_0}$ in (1.1) are determined by the relation*

$$\psi(z)=\sum_{j=0}^{\infty}\psi_j z^j=\frac{\theta(z)}{\alpha(z)}, \qquad \text{for } |z|\leq 1, \text{ with } \psi_0=1. \tag{1.2}$$

Similar to the case where a causal $ARMA$ process is expressed in terms of the past innovations, under certain conditions it can also be written as an infinite sum of the past realizations. This concept is called invertibility.

Definition 1.12. *(Invertibility). An $ARMA(p,q)$ process defined in Definition 1.9 is said to be invertible if there exists a sequence of constants $\{\phi_j\}_{j\in\mathbb{N}}$ such that $\sum_{j=1}^{\infty}|\phi_j|<\infty$ and*

$$X_t=\sum_{j=1}^{\infty}\phi_j X_{t-j}, \qquad \text{for } t\in\mathbb{Z}. \tag{1.3}$$

Similar to the conditions that ensure causality, we have the following theorem (Brockwell and Davis (1991, page 86)) giving the conditions needed for invertibility.

Theorem 1.13. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be an $ARMA(p,q)$ process for which the polynomials $\alpha(.)$ and $\theta(.)$ have no common zeroes. Then $\{X_t\}_{t\in\mathbb{Z}}$ is invertible if and only if $\theta(z)\neq 0$ for all $z\in\mathbb{C}$ such that $|z|\leq 1$. The coefficients $\{\phi_j\}_{j\in\mathbb{N}}$ in (1.3) are determined by the relation*

$$\phi(z)=\sum_{j=0}^{\infty}\phi_j z^j=\frac{\alpha(z)}{\theta(z)}, \qquad |z|\leq 1, \text{ with } \phi_0=1. \tag{1.4}$$

Observe that the autocovariance function $\gamma_X(h)$ of an $ARMA(p,q)$ process decays exponentially to zero as $h \to \infty$. This exhibits the short-rage dependence among the values of the time series. Therefore, $ARMA$ processes are employed to model time series that show short-memory behavior. The short memory property of a time series is usually recognized by inspecting the sample autocovariance function of the empirical series (Box and Jenkins 1970).

1.3. *ARIMA* PROCESSES

The $ARMA(p,q)$ processes discussed in the previous section can be generalized to include non-stationary behavior. The traditional generalization of $ARMA$ models leads to the $ARIMA(p,d,q)$ process, which is defined below.

Definition 1.14. *(ARIMA process). A real-valued process $\{X_t\}_{t\in\mathbb{Z}}$ is said to be an Autoregressive Integrated Moving Average (ARIMA(p,d,q)) process if the process $\{Y_t\}_{t\in\mathbb{Z}}$ with $Y_t := \bigtriangledown^d(x_t - \mu)$, is a causal $ARMA(p,q)$ process, where $\bigtriangledown = 1 - B$ and $d \in \mathbb{N}$.*

Observe that an $ARIMA(p,d,q)$ process can be written as $\alpha^*(B)(X_t - \mu) = \theta(B)\epsilon_t$, $\epsilon_t \sim WN(0,\sigma^2)$, $t \in \mathbb{Z}$, where $\alpha^*(B) := (1-B)^d\alpha(B)$, $\alpha(z)$ is a polynomial of order p and $\theta(z)$ is a polynomial of order q. Moreover, $\alpha(z) \neq 0$ for $|z| \leq 1$. Such processes are also called unit root processes because the polynomial $\alpha^*(z)$ has d roots equal to unity. In the formulation $(1-B)^d\alpha(B)$, these unit roots are factored out so that $\alpha(z)$ has all roots outside the unit circle (i.e. $|z| > 1$ if $\alpha(z) = 0$). Thus, $\{Y_t\}_{t\in\mathbb{Z}}$ obeys the $ARMA(p,q)$ model $\alpha(B)Y_t = \theta(B)\epsilon_t$, $t \in \mathbb{Z}$ because $Y_t = (1-B)^{-d}X_t$ and $(1-B)^d\alpha(B)X_t = \alpha(B)(1-B)^dX_t = \alpha(B)Y_t$.

Many financial and economics time series are nonstationary and hence the $ARIMA$ processes are extensively used in these areas. Figures 1.1 and 1.2 show the behavior of simulated $ARMA$ and $ARIMA$ time series, respectively. The sample path of the model given in Figure 1.2 is an example of nonstationarity or unit root behavior. It is necessary to difference the original non-stationary repeatedly to render the stationarity; see Brockwell and Davis (1991 Ch. 9) for more details.

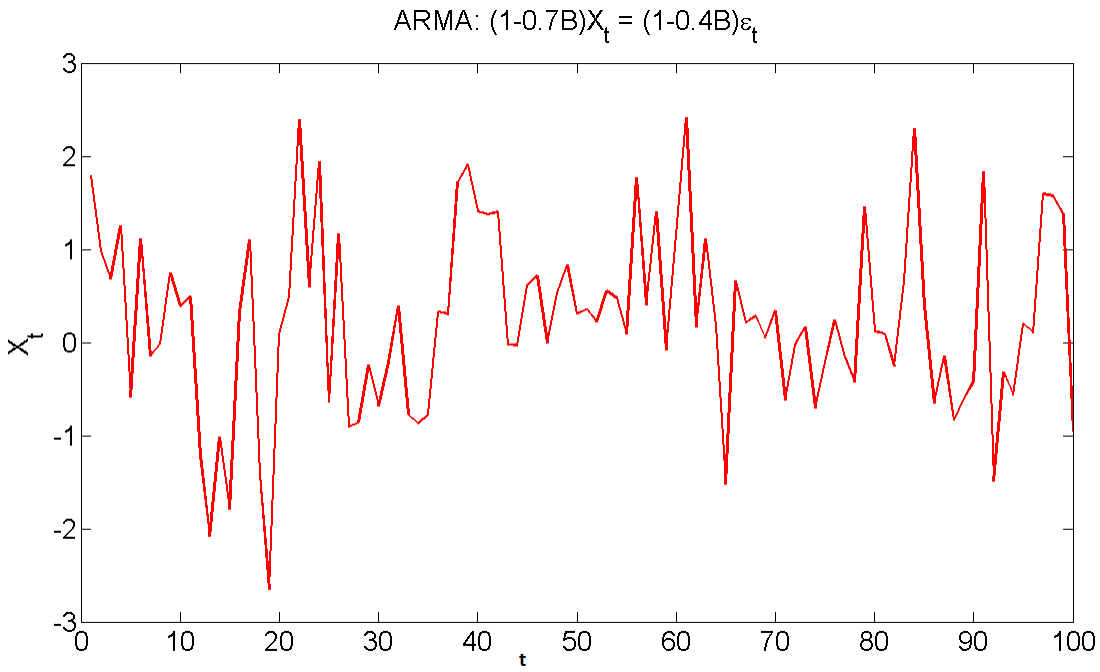

Figure 1.1. Simulated $ARMA(1,1)$ time series

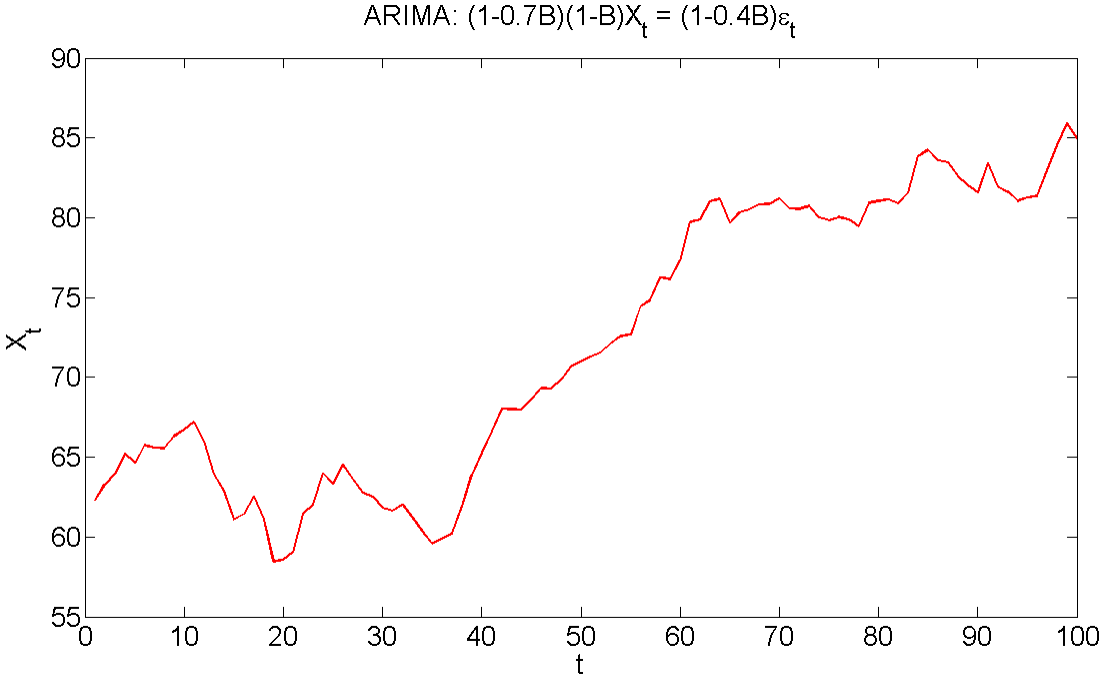

Figure 1.2. Simulated $ARIMA(1,1,1)$ time series

The sample autocorrelation functions of the above simulated $ARMA$ and $ARIMA$ time series are displayed in Figures 1.3 and 1.4, respectively. As mentioned before, the sample ACF of the $ARMA$ time series exponentially decays to zero as $h \to \infty$. Whereas, Figure 1.4 exhibits the typical behavior of the sample ACF of a unit root process.

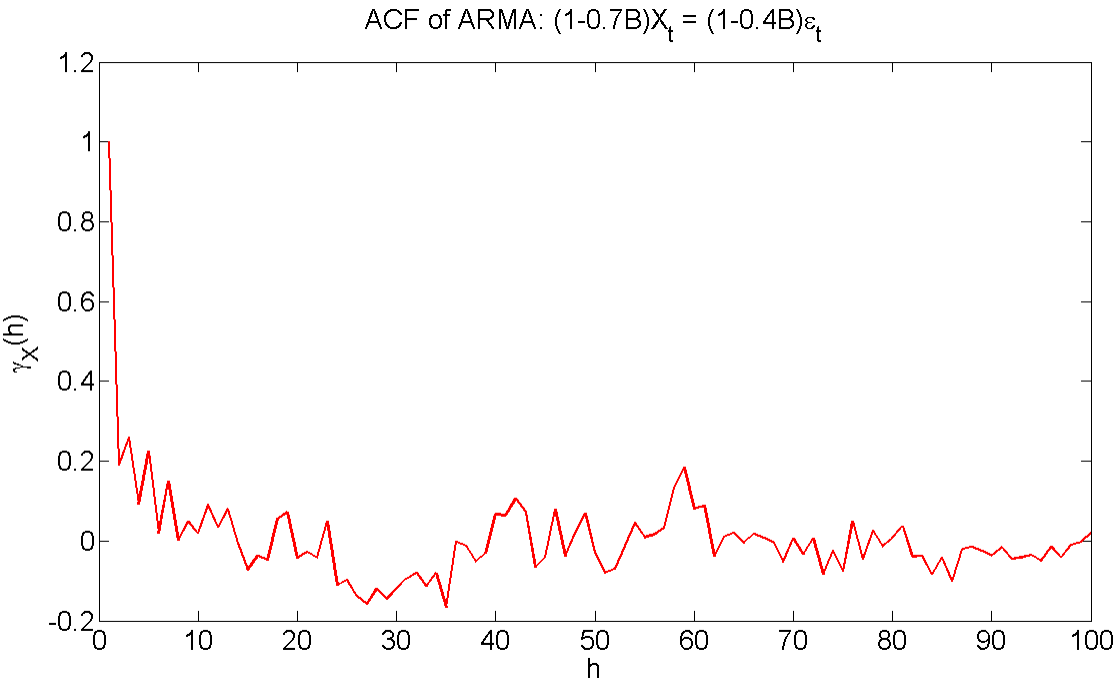

Figure 1.3. The ACF of $ARMA(1,1)$ time series

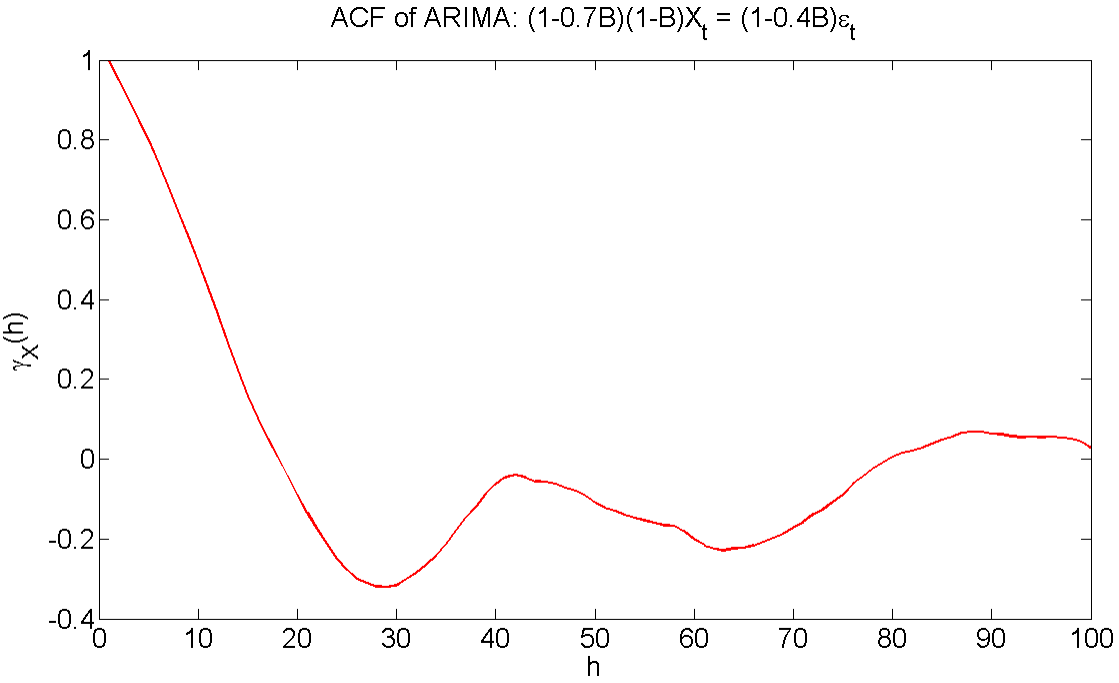

Figure 1.4. The ACF of $ARIMA(1,1,1)$ time series

So far, we have discussed stationary processes with short-memory and unit root processes that exhibit long-memory. A class of time series that exhibit long-memory, but are still stationary, straddles the gap between $ARMA$ and $ARIMA$ processes. These are called *Fractionally Integrated Moving Average* processes.

1.4. *FARIMA* PROCESSES

The long-memory time series that are very common in geophysical sciences, macroeconomics, asset pricing, stock returns and exchange rates can be modeled using *Fractionally Integrated Autoregressive Moving Average* ($FARIMA$ or $ARFIMA$) processes (see, for example, Baillie (1996) and Taqqu et al. (2003)).

Definition 1.15. *($FARIMA$ process). A real-valued process $\{X_t\}_{t\in\mathbb{Z}}$ is said to be a Fractionally Integrated Autoregressive Moving Average (FARIMA(p,d,q)) process if the process $Y_t := \bigtriangledown^d(x_t - \mu)$ is a causal $ARMA(p,q)$ process, where the difference parameter $d \in (-0.5, 0.5)$.*

Instead of an exponential rate of decay for $\gamma_X(h)$ as in the case for $ARMA$ models, the autocovariance function of a $FARIMA(p, d, q)$ process decays to zero at a hyperbolic rate. This is the reason for empirical processes with long-memory are being modeled using $FARIMA$ processes.

A $FARIMA(p, d, q)$ process is stationary if $d < 0.5$. The sample paths of two $FARIMA$ models were simulated for $d = 0.3$ and $d = 0.49$, and displayed in Figures 1.5 and 1.6. As seen in Figure 1.6, the series with $d = 0.49$ behaves almost like an $ARIMA$ process and thus is near non-stationarity. If $d \geq 0.5$, then the process given by Definition 1.15 is not stationary. Note that $FARIMA$ processes with $0 < d < 0.5$ are called long-memory processes while those with $-0.5 < d < 0$ are called intermediate memory.

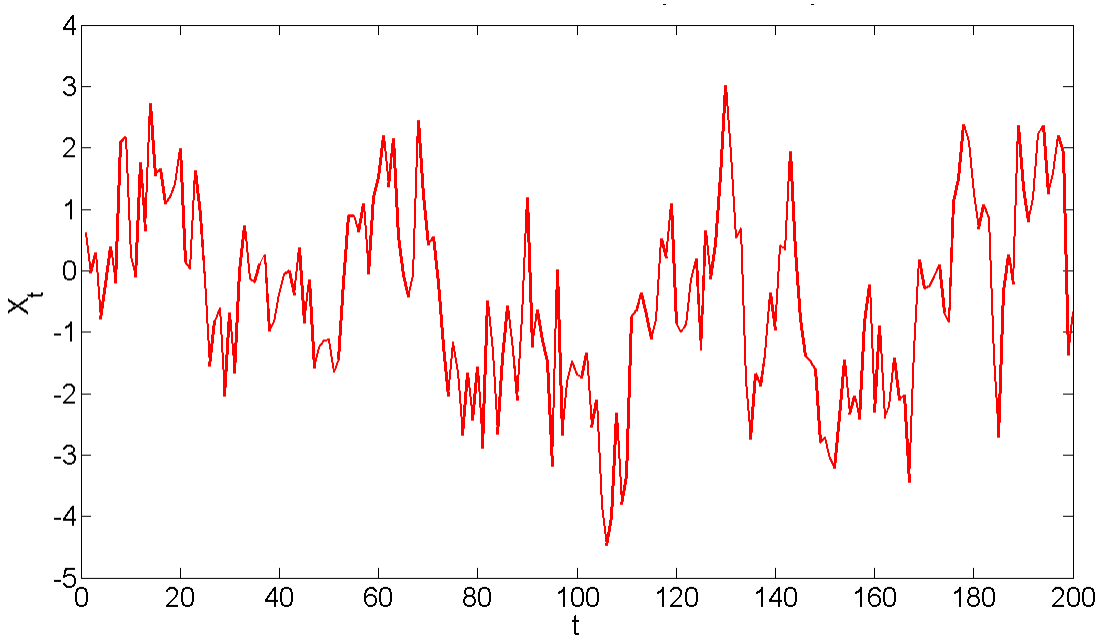

Figure 1.5. Simulated $FARIMA : (1 - 0.7B)(1 - B)^{0.3}X_t = (1 - 0.4B)\epsilon_t$ time series

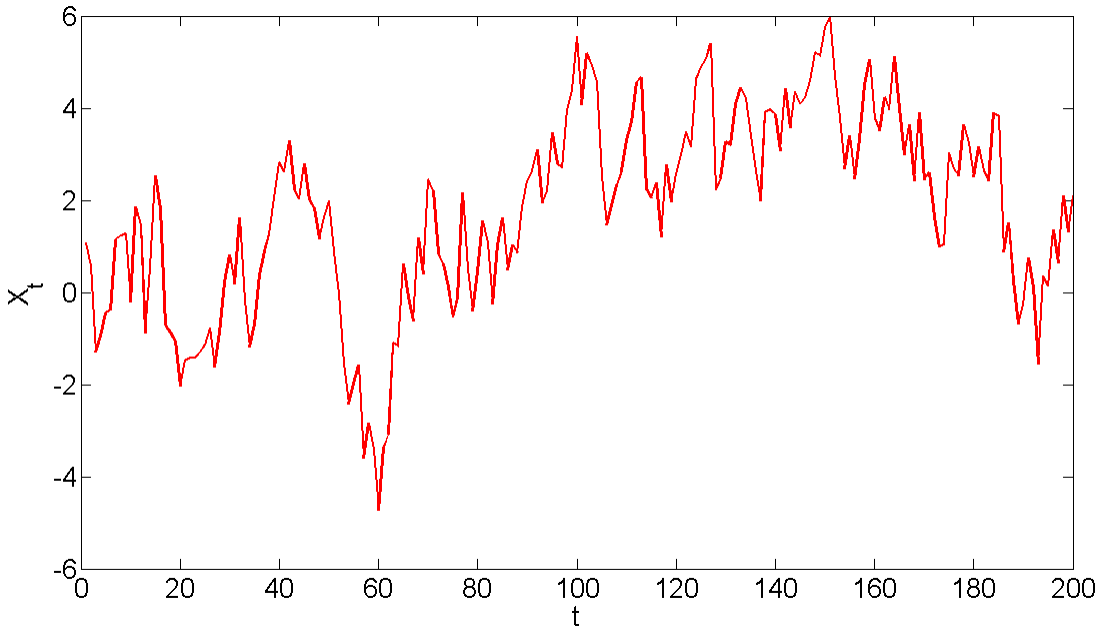

Figure 1.6. Simulated $FARIMA : (1 - 0.7B)(1 - B)^{0.49}X_t = (1 - 0.4B)\epsilon_t$ time series

The sample autocorrelation functions of the above simulated $FARIMA$ processes are displayed in Figures 1.7 and 1.8. Both of the sample ACF exhibit the long-range dependence among the observations. As seen in Figure 1.8, however, the $FARIMA$ process with d close to 0.5 is embedded in a wider long-range dependence structure.

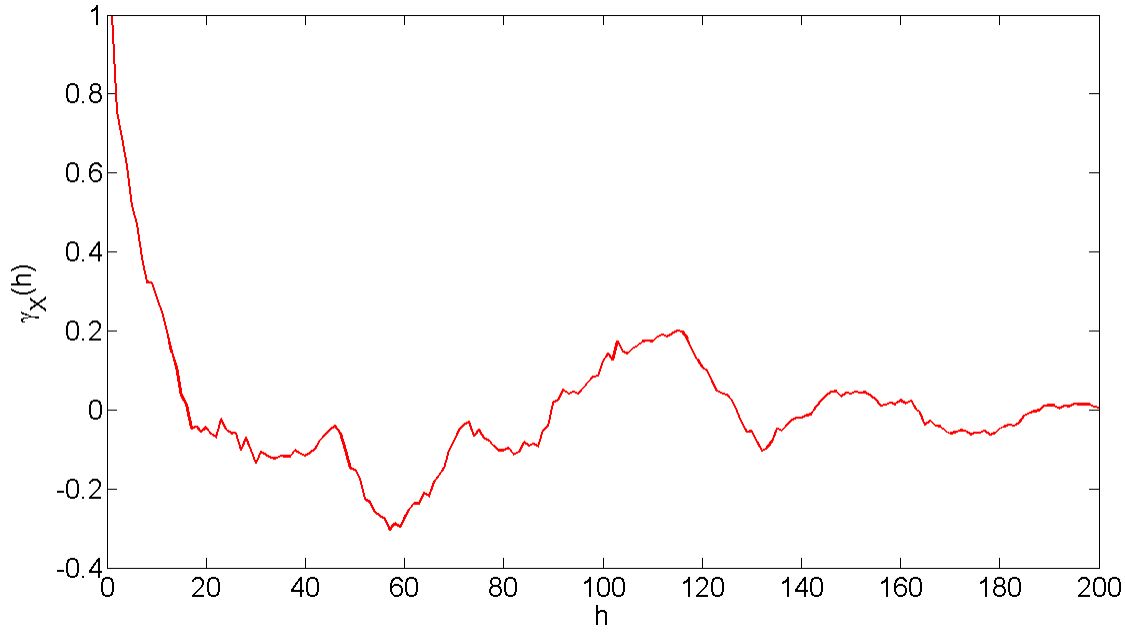

Figure 1.7. The ACF of $FARIMA : (1 - 0.7B)(1 - B)^{0.3}X_t = (1 - 0.4B)\epsilon_t$ time series

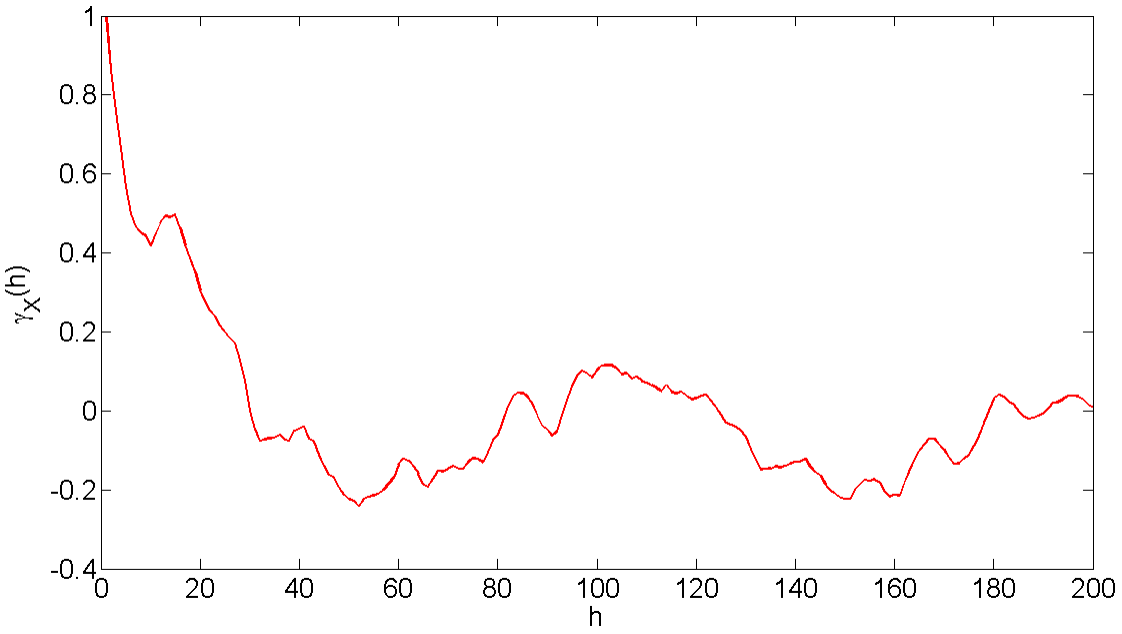

Figure 1.8. The ACF of $FARIMA : (1 - 0.7B)(1 - B)^{0.49}X_t = (1 - 0.4B)\epsilon_t$ time series

1.5. THE SIEVE BOOTSTRAP PROCEDURE

Non-parametric Bootstrap techniques play vital role in statistics as they can be utilized to make asymptotically valid statistical inferences even when the underlying probability distributions are not known. The Sieve Bootstrap method for constructing prediction intervals for invertible processes is based on re-samples of residuals obtained by fitting a finite degree autoregressive approximation to the time series. The advantage of this technique is that it does not require the knowledge of orders associated with the original model. In order to understand the merits of the sieve bootstrap method in relation to other re-sampling techniques in time series, a brief discussion of the adoption of the bootstrap techniques to time series analysis is needed. This is the motivation behind the following discussion.

The first application of bootstrap technique to time series was by Stine (1982, 1987) and Findley (1986), who proposed a bootstrap method to estimate the prediction mean squared error of the estimated linear predictor of an $AR(p)$ processes with unknown order p. They assumed that the error distribution is symmetric and with finite moments. Thombs and Schucany (1990) presented a bootstrap procedure for obtaining forecast intervals that required the ability to express the time series as a linear function of future values (Backward Representation). While this condition is met for stationary AR processes, it is, however, not possible for time series with a moving average component. Thus their method has limited applicability. Cao et al. (1997) studied a computationally faster conditional bootstrap method as an alternative to the procedure of Thombs and Schucany. All the above methods assumed that the order of the process is known and the coverage probabilities can suffer if the order is grossly misspecified. Masarotto (1990) and Grigoletto (1998) were the first to propose a bootstrap method for $AR(p)$ processes with finite but unknown order p. Their method is based on the assumption that some consistent estimator, $\hat{p}$, is available. This approximation is the basis for the sieve bootstrap.

The foundation for this sieve bootstrap approach was laid by Kreiss (1988, 1992), for time series that can be represented by an infinite autoregressive process. Bühlmann (1997) and Bickel and Bühlmann (1999) extended this approach to a more general class of time series that can be written as an infinite order moving average and introduced the

term "sieve bootstrap". The advantage of the sieve bootstrap is that it does not require the knowledge of the orders associated with the underlying process. The class of time series, however, is limited to the processes whose infinite moving average coefficients are absolutely summable. In another seminal paper, Poskitt (2006) discussed the ways of relaxing the assumption of absolutely summability of the infinite moving average coefficients while maintaining the validity of an AR approximation. He achieved this by computing the coefficients of the AR approximation using the Yule-Walker equation. The class of time series considered by Poskitt (2006) were known as regular processes, whose infinite moving average coefficients are square summable. Poskitt (2008) went onto establish the large sample properties of the sieve bootstrap for such processes. He showed that the method of sieves yields asymptotically valid estimator of the distribution of test statistics that satisfy some regularity conditions.

Alonso, Pẽna and Romo (2002) applied the Bühlmann (1997) sieve bootstrap procedure to obtain prediction intervals for processes that have an infinite moving average representation with absolutely summable coefficients (such processes are called linear processes). In 2003, the same authors established the large sample properties of their sieve bootstrap method, and further refined the technique in 2004 by introducing methods to incorporate the variation due to model uncertainty in parameter estimation. The coverage probabilities obtained with this method, however, are liberal in many situations. Mukhopadhyay and Samaranayake (2010) proposed a rescaling factor for residuals to improve the coverage probabilities of the method of Alonso et *al.* (2004). Their adjustment is based on an intuitive observation. The same residual could be resampled many times and as a results, the variance of the bootstrapped residuals may be smaller than that of the original sample. The effect of the rescaling factor introduced by Mukhopadhyay and Samaranayake (2010), however, is asymptotically negligible.

Another important application of the sieve bootstrap is in hypothesis testing. In particular, our interest is on the use of the sieve bootstrap for testing seasonal unit roots. Psaradakis (2001) was the first to introduce the application of the sieve bootstrap to test the presence of non-seasonal unit roots. He assumed that the errors are weakly dependent. Complications with the regular Dickey-Fuller (DF) (1979) and Augmented-

Dickey-Fuller (ADF) (Said, Dickey and Fuller (1984)) unit root tests arise when there is a root near unity in the moving average polynomial of the underlying process. As a solution, Psaradakis suggested bootstrapping the DF test statistics using the method of sieves. Following Psaradakis (2001), the sieve bootstrap versions of the ADF tests for non-seasonal unit roots were suggested by Chang and Park (2003). Palm, Smeekes and Urbain (2008) proposed an alternative way of computing residuals by fitting the DF regression model instead of fitting an $AR(p)$ model to the differenced series. Moreover, Psaradakis (2000) proposed a (nonsieve) bootstrap method to obtain the empirical distribution of the Dickey-Hasza-Fuller (DHF) seasonal unit root test statistics. He, however, assumed that the errors are independently and identically distributed. One of the advantages of the bootstrap based unit root tests is that the critical values need not to be read from standard tables. Instead, the critical values are computed from the percentiles of the empirical (bootstrap) distribution of the test statistics.

The remaining portion of this dissertation is organized the form of a series of papers. Paper I is about establishing the asymptotic properties of the sieve bootstrap based prediction intervals for $FARIMA$ processes utilizing some of the results of Poskitt (2006 and (2008). We generalize the Alonso et *al.* (2003) sieve bootstrap procedure to $FARIMA$ processes by adjusting the order of $p(n)$, where $p = p(n)$ is the order of the AR approximation. In Paper II, we continue our interest on the proposed method by carrying out an extensive Monte Carlo simulation study. Our method is compared with the only existing method for obtaining prediction intervals for $FARIMA$ processes proposed by Bisaglia and Grigoletto (2001). An interesting application of the sieve bootstrap based prediction intervals is implemented in Paper III, where the re-sampling technique is applied to $ARIMA$ processes. The main advantage of the proposed method is that it does not use pre testing for unit roots and select the prediction interval procedure based on the results of these tests. The large sample properties of the intervals obtained using this method are also established. In Paper V, we apply the Psaradakis (2001) and Pam et *al.* sieve bootstrap techniques to seasonal time series with weakly dependent errors. A method for computing prediction intervals for (non)seasonally integrated time series using sieve bootstrap is also proposed with asymptotic validation. Papers II, III and V carry

out extensive Monte Carlo simulation studies to investigate the finite sample properties of the proposed methods.

Paper IV is, actually, little off from the theme of this dissertation but is essential for Paper V in establishing the asymptotic properties of the proposed sieve bootstrap procedure in that paper. The literature, surprisingly, lacks an expression for the asymptotic distribution of the DHF seasonal unit root test statistics under weakly dependent errors. We fill this gap by deriving the asymptotic distributions of the DHF seasonal unit root test statistics as functional of the standard Brownian motions.

PAPER

I. ASYMPTOTIC PROPERTIES OF SIEVE BOOTSTRAP PREDICTION INTERVALS FOR *FARIMA* PROCESSES

ABSTRACT

The Sieve Bootstrap is a model-free re-sampling technique that uses autoregressive approximations to model invertible linear time series and assumes that the order of the autoregressive process, p, goes to infinity with sample size n. The asymptotic properties of sieve bootstrap prediction intervals for stationary invertible linear processes with short-memory, such as autoregressive moving average time series, have been established under conditions that specify the rate of increase of the order p as a function of n. In this paper we extend these results to long memory ($FARIMA$) processes. We show that under certain regularity conditions the sieve bootstrap provides consistent estimators of the conditional distribution of future values of a $FARIMA$ processes, given the observed data.

Keywords: ARFIMA; Forecast Intervals; Fractionally Integrated Time Series; Long Memory Processes; Autoregressive approximations

1. INTRODUCTION

Long memory processes are increasingly used in modeling time series prevalent in many areas, for example, in geophysical sciences, macroeconomics, asset pricing, stock returns and exchange rates (see Baillie (1996) and Taqqu et *al.* (2003)). The *Fractionally Integrated Auto Regressive Moving Average* (*ARFIMA* or *FARIMA*) processes have been used extensively to model such processes (Granger and Joyeux (1980) and Hosking (1981)). In addition to estimation, obtaining forecasts and forecast intervals for such series are an important part of the empirical modeling process. As shown by Alonso (2002, 2003 and 2004) and Mukhopadhyay and Samaranayake (2010), the *Sieve Bootstrap* method works well in obtaining prediction intervals for invertible *ARMA* processes. Therefore, it is worthwhile to investigate its utility in obtaining prediction intervals for *FARIMA* processes as well. The assumptions under which the sieve bootstrap produces asymptotically valid prediction intervals in the *ARMA* case, however, do not hold for *FARIMA* processes. In this paper we present a modified formulation under which the sieve bootstrap based prediction intervals for *FARIMA* processes do achieve nominal coverage probabilities asymptotically. In the following, a brief outline of the origins of the sieve bootstrap is presented, followed by its adaptation to the *FARIMA* case.

Stine (1982, 1987) and Findley (1986) were the first to introduce bootstrap methods to compute prediction mean squared error for time series. These methods assumed that the order of the process is known. The main idea behind the sieve bootstrap, namely approximating the process with a finite order Autoregressive (AR) approximation (which therefore do not require the knowledge of the order of the process), originated with the proposal by Küsnch (1989) for a nonparametric block bootstrap procedure for obtaining the empirical distribution of some test statistics. Bootstrapping approximately *i.i.d.* residuals obtained from an assumed model was presented by Kreisss (1992). Bühlmann (1997) introduced the term *sieve bootstrap* and derived the asymptotic properties of the forecast intervals for a class of linear processes that can be written as an infinite moving average time series. His method used a truncated version of the $AR(\infty)$ representation

of the time series, whose order increases with sample size. Following Bühlmann (1997), Alonso et *al.* (2002, 2003, 2004) further refined this method and presented a sieve bootstrap approach to obtaining prediction intervals for a general class of linear processes that includes *ARMA* processes as a subset. Alonso's method was modified by Mukhopadhyay and Samaranayake (2010) to improve the coverages of the prediction intervals.

2. *FARIMA* PROCESSES

A real-valued process $\{y_t\}_{t\in\mathbb{Z}}$ is said to be a *Fractionally Integrated Autoregressive Moving Average* (*FARIMA*(p,d,q)) process if it is stationary and satisfies

$$\alpha(B) \bigtriangledown^d (y_t - \mu) = \theta(B)\epsilon_t, t \in \mathbb{Z}, \tag{2.1}$$

where $\alpha(z) = 1 - \alpha_1 z.... - \alpha_p z^p$ and $\theta(z) = 1 + \theta_1 z..... + \theta(q)z^q$ represent autoregressive and moving average polynomials of degrees p and q respectively. The mean of the process is $\mu = E[y_t]$ for all t. It is assumed that $\alpha(.)$ and $\theta(.)$ do not share common zeros. The error terms, $\{\epsilon_t\}$, are assumed to be zero-mean white noise[1] with finite variance σ^2. Note that $\bigtriangledown = 1 - B$, where B is the back-shift operator defined by $B^k y_t = y_{t-k}$ for $k \in \mathbb{N}$.

The process exhibits *long-memory* when $0 < d < 0.5$, in which case the autocovariance function, $\gamma_Y(\cdot)$, is not absolutely summable but $\gamma_Y(0) < \infty$. For $-0.5 < d < 0$, the process is said to be *intermediate-memory* as the autocovarice function is absolutely summable. When $0 < d < 0.5$, $\theta(z) \neq 0$ for $|z| \leq 1$, $\{y_t\}$ is invertible and has an $AR(\infty)$ representation, $\sum_{j=0}^{\infty} \phi_j (y_{t-j} - \mu) = \epsilon_t$ with $\sum_{j=0}^{\infty} |\phi_j| < \infty$. Also, when $\alpha(z) \neq 0$ for $|z| \leq 1$, $\{y_t\}$ can be expressed as $y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$ with $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. $FARIMA(p,d,q)$ processes with $0 < d < 0.5$, however, do not satisfy the assumption that $\sum_{j=0}^{\infty} j^r |\psi_j| < \infty$ for some $r \in \mathbb{N}$, crucial for establishing the asymptotic results of Bühlmann (1997) and Alonso et *al.* (2003), because for such time series $\psi_j \sim j^{d-1}$ as $j \longrightarrow \infty$.

[1]The stationary process $\{x_t : t \in \mathbb{Z}\}$ with an autocovariance function (ACVF) $\gamma_X(\cdot)$ is said to be white noise if (i) $Ex_t = 0$ for all $t \in \mathbb{Z}$, and (ii) $\gamma_X(h) = \sigma^2$ for $h = 0$ and $\gamma_X(h) = 0$ for $h \neq 0$.

Poskitt (2006) discussed ways of relaxing this condition and yet maintaining the statistical viability of finite autoregressive approximations to *FARIMA* processes. In a later paper, Poskitt (2007) showed how the sieve bootstrap method can be utilized to compute empirical distributions of specific types of statistics associated with *FARIMA* processes. He also derived the asymptotic properties of these empirical distributions under certain regularity conditions.

While Poskitt (2006 and 2007) did not directly address the issue of determining an asymptotically valid estimator of the distribution of a future observation from a *FARIMA* process, his results provide a theoretical foundation on which such a solution can be constructed. In the following sections we show how the method proposed by Alonso (2003) can be modified, based on insights from Poskitt (2006 and 2007), to obtain sieve bootstrap prediction intervals for a $FARIMA(p, d, q)$ process when $0 < d < 0.5$. While Alonso (2004) and Mukhopadhyay and Samaranayake (2010) provide additional refinements to the original method proposed in Alonso et *al.* (2002, 2003), the 2003 paper by Alonso et *al.* sets the fundamental theoretical framework for the application of the sieve bootstrap for invertible processes. As such, we use it as the platform for our proposed modifications. This would serve as a first, but an important, step in adopting the sieve bootstrap to obtain prediction intervals for *FARIMA* processes.

The rest of this paper is organized as follows. Section 3 introduces the sieve bootstrap procedure for obtaining prediction intervals and Section 4 establishes asymptotic validity of the proposed method.

3. THE PROPOSED SIEVE BOOTSTRAP PROCEDURE

The proposed sieve bootstrap procedure given below is identical to that introduced by Alonso et *al.* (2002, 2003) except for the criterion used in selecting the order of the autoregressive approximation. This change in the order together with a few additional lemmas is sufficient to provide the necessary convergence results.

Let the realization $\{y_t\}_{t=1}^n$ be obtained from the stationary and invertible process given in equation (2.1) with infinite AR representation $y_t - \mu = \sum_{j=0}^{\infty} \phi_j (y_{t-j} - \mu)$, $t \in \mathbb{Z}$. As Alonso et *al.* (2002, 2003, 2004) did, we shall estimate μ by the empirical mean $\bar{y} = n^{-1} \sum_{t=1}^{n} y_t$. Then proceed as follows.

1. Select the order $p = p(n)$ of the autoregressive approximation, following Poskitt (2006, 2007), from among models with $p \in \{1, 2, ..., M_n\}$ with $M_n = c[log(n)]^a$ for some $c > 0$ and $a \geq 1$ by AIC criterion. Alonso et *al.* (2003) preferred AICC over AIC and used $M_n = o\{[log(n)/n]^{1/4}\}$.

2. Calculate the autoregressive coefficients, $\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n}$, of the $AR(p)$ model, $y_t - \bar{y} = \sum_{j=0}^{\infty} \phi_j (y_{t-j} - \bar{y})$ by the Yule-Walker method.

3. Obtain the $(n-p)$ residuals: $\hat{\epsilon}_{t,n} = \sum_{j=0}^{p} \hat{\phi}_{j,p,n} (y_{t-j} - \bar{y})$, $t = p+1, ..., n$ and define the empirical distribution function of the centered residuals, $\tilde{\epsilon}_t = \hat{\epsilon}_{t,n} - \hat{\epsilon}^{(.)}$, where $\hat{\epsilon}^{(.)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n}$, by $\hat{F}_{\tilde{\epsilon}}(x) = (n-p)^{-1} \sum_{t=p+1}^{n} I_{[\tilde{\epsilon}_t \leq x]}$.

4. Draw a resample $\epsilon_{t,n}^*$, $t = p+1, ..., n$ of i.i.d. observations from $\hat{F}_{\tilde{\epsilon}}$.

5. Set $y_t^* = \bar{y}$ for $t = 1, ..., p$ and obtain y_t^* by the recursion: $\sum_{j=0}^{p} \hat{\phi}_{j,p,n} (y_{t-j}^* - \bar{y}) = \epsilon_{t,n}^*$ for $t = p+1, ..., n$.

6. Compute the estimates $(\hat{\phi}_{1,p,n}^*, ..., \hat{\phi}_{p,p,n}^*)'$ as in Step 2, using $\{y_t^*\}_{t=1}^n$.

7. For $h \in \mathbb{N}$, compute the future bootstrap observations by the recursion: $y_{n+h}^* - \bar{y} = \sum_{j=1}^{p} \hat{\phi}_{j,p,n}^* (y_{n+h-j}^* - \bar{y}) + \epsilon_{n+h,n}^*$ where, $y_t^* = y_t$, for $t \leq n$.

8. Obtain a Monte Carlo estimate of the bootstrapped distribution function of y_{n+h}^* by repeating steps 4-7 B times and use this bootstrapped distribution is used to approximate the unknown distribution of y_{n+h} given the observed sample.

9. The $100(1-\alpha)\%$ prediction interval for y_{n+h} is given by $\{Q^*(\frac{\alpha}{2}), Q^*(1-\frac{\alpha}{2})\}$ where, $Q^*(.)$ are the quantiles of the estimated bootstrap distribution.

4. ASYMPTOTIC RESULTS

In order to establish the asymptotic validity of the sieve bootstrap intervals, Alonso et *al.* (2003) first established the convergence of $\hat{\boldsymbol{\phi}}^*_{p,n}$ to $\hat{\boldsymbol{\phi}}_{p,n}$ and then went onto prove the convergence of the conditional distribution of X^*_{n+h} to that of X_{n+h}. We follow the same approach, but modify the proofs to accommodate the changes arising out of the presence of fractional integration.

Some of the results in Bühlmann (1995, 1997) and Alonso et *al.* (2003) can be extended to *regular processes* that include both *FARIMA* and non-invertible time series. Note that the process $\{y_t\}_{t\in\mathbb{Z}}$ is said to be linearly regular if $\{y_t\}_{t\in\mathbb{Z}}$ is covariance stationary with,

$$y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \tag{4.1}$$

where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is a zero mean white noise process with variance σ^2 and the impulse response coefficients $\{\psi_j\}_{j=0}^{\infty}$ satisfy the condition $\psi_0 = 1$ and $\sum_{j\geq 0} \psi_j^2 < \infty$ (Poskitt (2006)). Since $FARIMA$ processes satisfy this condition, we can use the $AR(p)$ approximation suggested by Poskitt (2006) for such time series.

Definition 4.1: Let $\{y_t\}$ satisfy (4.1) and define, for $p < n$, $\{\epsilon_{t,p}\}$ such that $\sum_{j=0}^{p} \phi_{j,p} y_{t-j} = \epsilon_{t,p}$, where the AR coefficients vector $\boldsymbol{\phi}_p = (\phi_{1,p}, ..., \phi_{p,p})'$ is obtained using the Yule-Walker equations, $\boldsymbol{\Gamma}_p \boldsymbol{\phi}_p = -\boldsymbol{\gamma}_p$, $\boldsymbol{\gamma}_p = (\gamma(1), ..., \gamma(p))'$, $\boldsymbol{\Gamma}_p = [\gamma(i-j)]_{i,j=1}^{p}$, with $\gamma(k) = E[y_t y_{t+k}]$ for $k \in \mathbb{N}_0$. Note that Lemma 1 of Poskitt (2006) establishes that $\epsilon_{t,p} \to \epsilon_t$ in mean square as $p \to \infty$.

The following sets of assumptions are required in order to prove our asymptotic results.

A1: Let ξ_t denote the σ-algebra of events determined by ϵ_s, $s \leq t$. Also, assume $\{\epsilon_t\}_{\in \mathbb{Z}}$ are $i.i.d.$[2] and that

$$E[\epsilon_t|\xi_{t-1}] = 0 \qquad and \qquad E[\epsilon_t^2|\xi_{t-1}] = \sigma^2, t \in \mathbb{Z}. \tag{4.2}$$

Furthermore, assume $E[\epsilon_t^4] < \infty$ for $t \in \mathbb{Z}$.

A2: The series y_t is a linearly regular covariance-stationary process with Wold representation $y_t = \sum_{j\geq 0} \psi_j \epsilon_{t-j}$ where $\Psi(z) = m(z)/(1-z)^d$ for $|d| < 0.5$ and $m(z) = \sum_{j=0}^{\infty} \mu_j z^j$ is a causal transfer function with impulse response coefficients satisfying $\sum_{j\geq 0} |\mu_j| < \infty$.

B: Let $p(n) = o\{[n/log(n)]^{1/2-d}\}$ and $\hat{\boldsymbol{\phi}}_{p,n} = (\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n})'$ satisfy the empirical Yule-Walker equations $\hat{\boldsymbol{\Gamma}}_{p,n}\hat{\boldsymbol{\phi}}_{p,n} = -\hat{\boldsymbol{\gamma}}_{p,n}$ where
$\hat{\boldsymbol{\Gamma}}_{p,n} = [\hat{R}(i-j)]_{i,j=1}^{p}, \hat{\boldsymbol{\gamma}}_{p,n} = (\hat{R}(1), ..., \hat{R}(p))'$, and $\hat{R}(j) = n^{-1}\sum_{t=1}^{n-|j|}(y_t - \bar{y})(y_{t+|j|} - \bar{y})$ for $|j| < n$.

Assumptions in A1 imposes a Martingale difference structure on the innovations. The condition on the infinite polynomial, $m(z)$, in A2 is satisfied by the *ARMA* component of *FARIMA* processes.

Note that in the proposed method, $p(n)$ is chosen from among values $\{1, 2, ..., M_n\}$, where $M_n = c[log(n)]^a$ with $c > 0$ and $a \geq 1$. As stated in Poskitt (2006) the above choice of M_n is sufficient to ensure the order for $p(n)$ stated in the Assumption set B. Poskitt (2006) used $c = 1$ and $a = 1.962$ for his simulation studies. Moreover, these values for c and a produced coverage probabilities close to nominal levels in a Monte-Carlo simulation study that we conducted to investigate the finite sample performances of the sieve bootstrap procedure proposed in Section 3. Thus, we recommend using the above values.

Next we present asymptotic properties of the sieve bootstrap method given in Section 3 by generalizing results in Bühlmann (1995, 1997) and Alonso et *al.* (2003) to regular processes. In the following we let, $\| \boldsymbol{x} \|_\infty = \max_{1\leq j\leq m} |x_j|$, $\| \boldsymbol{x}\|_2 = (\sum_{j=1}^{m} x_j^2)^{1/2}$ for $x \in \mathbb{R}^m$, and $\|\boldsymbol{X}\|_{row} = \max_{1\leq i\leq m} \sum_{j=1}^{n} |x_{i,j}|$ for $\boldsymbol{X} \in \mathbb{R}^{m\times n}$.

[2]One can assume $\{\epsilon_t\} \sim WN(0, \sigma^2)$ for all results except Proposition 1 which requires the $i.i.d.$ assumption.

The following lemma is crucial for proving Proposition 1 by providing a way around the conditions the fractional integration imposes on the infinite moving average $MA(\infty)$ representation of $FARIMA$ processes. Alonso et *al.* (2003) uses this $MA(\infty)$ representation to establish the convergence of the bootstrap autoregressive parameter estimators to the original estimators.

Lemma 1: Assume that A1, A2 and B hold. Then,

$$\sum_{j=0}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})^2 = o_{a.s.}\{[log(n)/n]^{1/2-d}\},$$

where $\phi_{j,p}$, $j = 1, 2, ..., p$, $p < n$ are the coefficients given in Definition 4.1.

Proof: We follow the same argument as in An et *al.* (1982, pp. 935, 936). First, observe that from Theorem 1 of Poskitt (2006),

$$\max_{0\leq j\leq p} |\hat{R}(j) - \gamma(j)| = O_{a.s.}\{[log(n)/n]^{1/2-d}\}, \tag{4.3}$$

where $\hat{R}(\cdot)$ is the sample ACVF defined in the Assumption set B. Now consider,

$$\boldsymbol{\Gamma}_p(\hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p) = -(\hat{\boldsymbol{\Gamma}}_{p,n} - \boldsymbol{\Gamma}_p)(\hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p) - (\hat{\boldsymbol{\gamma}}_{p,n} - \boldsymbol{\gamma}_p) - (\hat{\boldsymbol{\Gamma}}_{p,n} - \boldsymbol{\Gamma}_p)\boldsymbol{\phi}_p. \tag{4.4}$$

$$\begin{aligned}\text{Observe that } \| \boldsymbol{\Gamma}_p(\hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p) \|_2^2 &= \sum_{k=1}^{p}[\sum_{j=1}^{p}\gamma(j-k)(\hat{\phi}_{j,p,n} - \phi_{j,p})]^2 \\ &\leq \sum_{k=1}^{p}\sum_{i=1}^{p}\gamma(i-k)^2\sum_{j=1}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})^2.\end{aligned}$$

$$\text{Also, } \| (\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)(\hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p) \|_2^2 \leq \sum_{k=1}^{p}\sum_{i=1}^{p}[\hat{R}(i-k) - \gamma(i-k)]^2 \sum_{j=1}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})^2$$

$$\leq p^2 \max_{1\leq k,j\leq p}[\hat{R}(j-k) - \gamma(j-k)]^2 \| \hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p \|_2^2$$

$$\leq p^2\{O_{a.s.}[[log(n)/n]^{2(1/2-d)}]\} \| \hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p \|_2^2,$$

$$= o_{a.s.}(1) \| \hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p \|_2^2 . \tag{4.5}$$

$$\text{Furthermore, } \| \hat{\boldsymbol{\gamma}}_{p,n} - \boldsymbol{\gamma}_p \|_2^2 = \sum_{j=1}^{p}[\hat{R}(j) - \gamma(j)]^2 = p\{O_{a.s.}[[log(n)/n]^{2(1/2-d)}]\}$$

$$= o_{a.s.}\{[log(n)/n]^{1/2-d}\}, \text{ and}$$

$$\| (\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)\phi_p \|_2^2 \leq \sum_{k=1}^{p}\sum_{i=1}^{p}[\hat{R}(i-k) - \gamma(i-k)]^2 \sum_{j=1}^{p}\phi_{j,p}^2 = o_{a.s.}(1),$$

because $\sum_{j=1}^{p}\phi_{j,p}^2 \leq c < \infty$ for all p as observed by An et *al.* (1982).
Now, post multiplying Equation (4.4) by $\mathbf{\Gamma}_p^{-1}$ [3] and then bringing the first term of the right hand side of (4.4) to the left of the equal sign, we have

$$[\boldsymbol{I}_p + \mathbf{\Gamma}_p^{-1}(\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)](\hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p) = -\mathbf{\Gamma}_p^{-1}(\hat{\boldsymbol{\gamma}}_{p,n} - \boldsymbol{\gamma}_p) - \mathbf{\Gamma}_p^{-1}(\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)\boldsymbol{\phi}_p. \tag{4.6}$$

But $[\boldsymbol{I}_p + \mathbf{\Gamma}_p^{-1}(\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)] = I_p[1 + o_{a.s.}(1)]$ since each element of $(\hat{\mathbf{\Gamma}}_{p,n} - \mathbf{\Gamma}_p)$ is bounded by $\sum_{k=1}^{p}\sum_{j=1}^{p}[\hat{R}(j-k) - \gamma(j-k)]^2 = o_{a.s.}(1)$, and the left hand side of Equation (4.6) is $o_{a.s.}\{[log(n)/n]^{1/2-d}\}$. Therefore, $\{1 + o_{a.s.}(1)\} \| \hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p \|_2^2 = o_{a.s.}\{[log(n)/n]^{1/2-d}\}$ which implies that $\| \hat{\boldsymbol{\phi}}_{p,n} - \boldsymbol{\phi}_p \|_2^2 = o_{a.s.}\{[log(n)/n]^{1/2-d}\}$, as was observed by An et *al.* (1982, pp. 935, 936). □

The following Lemma is analogous to Lemma 5.3 of Bühlmann (1997), and establishes the convergence of the second moment of bootstrap innovations to their theoretical second moment, which is needed to prove Proposition 1.

[3] The proof of Corollary 1 in Poskitt (2006) ensures that the minimum eigenvalue of $\mathbf{\Gamma}_p$ is bounded away from zero.

Lemma 2: Assume that A1, A2 and B hold. Then, for any fixed $t \in \mathbb{Z}$,

$$E^*(\epsilon_{t,n}^{*2}) = E(\epsilon_t^2) + o_p(1). \tag{4.7}$$

Proof: By definition of expectation with respect to the bootstrap distribution,

$$E^*(\epsilon_{t,n}^{*2}) = (n-p)^{-1} \sum_{t=p+1}^{n} (\hat{\epsilon}_{t,n} - \hat{\epsilon}_n^{(.)})^2, \tag{4.8}$$

where $\hat{\epsilon}_n^{(.)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n}$.

We first show that, $\hat{\epsilon}_n^{(.)} = o_p(1)$. Observe that $\hat{\epsilon}_n^{(.)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n} \leq (n-p)^{-1} \sum_{t=p+1}^{n} |\hat{\epsilon}_{t,n} - \epsilon_t| + (n-p)^{-1} \sum_{t=p+1}^{n} |\epsilon_t|$ and that $(n-p)^{-1} \sum_{t=p+1}^{n} |\epsilon_t| = O_p(n^{-1/2})$. Moreover, $(n-p)^{-1} \sum_{t=p+1}^{n} |\hat{\epsilon}_{t,n} - \epsilon_t|$

$\leq (n-p)^{-1} \sum_{t=p+1}^{n} |\hat{\epsilon}_{t,n} - \epsilon_{t,p}| + (n-p)^{-1} \sum_{t=p+1}^{n} |\epsilon_{t,p} - \epsilon_t| = I_1 + I_2.$

Now, $I_2 = \sum_{t=p+1}^{n} (n-p)^{-1} |\epsilon_{t,p} - \epsilon_t| = (n-p)^{-1} o_p(n-p) = o_p(1)$, due to Lemma 1 of Poskitt's (2006) and the fact that mean square convergence implies convergence in probability.

In addition, using Holder's inequality, it can be shown that

$$I_1 = (n-p)^{-1} \sum_{t=p+1}^{n} |\hat{\epsilon}_{t,n} - \epsilon_{t,p}| = (n-p)^{-1} \sum_{t=p+1}^{n} \sum_{j=0}^{p} |(\hat{\phi}_{j,p,n} - \phi_{j,p}) y_{t-j}|$$

$$\leq [\sum_{j=0}^{p} (\hat{\phi}_{j,p,n} - \phi_{j,p})^2]^{1/2} \{(n-p)^{-1} \sum_{t=p+1}^{n} [\sum_{i=0}^{p} |y_{t-i}|^2]^{1/2}\}.$$

But from Lemma 1, $\sum_{j=0}^{p} (\hat{\phi}_{j,p,n} - \phi_{j,p})^2 = o_{a.s.}\{[log(n)/n]^{1-2d}\}$.

$$\text{Therefore, } I_1 = o\{[log(n)/n]^{(1-2d)/2}\} O_p(p^{1/2})$$

$$= o\{[log(n)/n]^{(1-2d)/2}\} O_p\{[log(n)/n]^{-(1-2d)/2}\} = o_p(1).$$

Next we show that, $(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}|^2 = E(\epsilon_t^2) + o_p(1)$.

Observe that

$$\begin{aligned}(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}|^2 &= (n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}-\epsilon_t+\epsilon_t|^2 \\ &\leq (n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}-\epsilon_t|^2+ \\ &\quad 2(n-p)^{-1}\sum_{t=p+1}^{n}|\epsilon_t||\hat{\epsilon}_{t,n}-\epsilon_t| + (n-p)^{-1}\sum_{t=p+1}^{n}\epsilon_t^2.\end{aligned}$$

Employing the same argument as used for I_1, the first term can be shown to be $o_p(1)$. The ergodicity of ϵ_t implies that the last term in the above inequality converges to $E(\epsilon_t^2)$ in probability. Using Holder's inequality, the middle term in the above expression can be bounded above by,

$2[(n-p)^{-1}\sum_{t=p+1}^{n}|\epsilon_t|^2]^{1/2}[(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}-\epsilon_t|^2]^{1/2} = o_p(1)$.

Hence, $(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}|^2 = E(\epsilon_t^2) + o_p(1)$.

Now, expanding the right hand side of (4.8) we complete the proof. □

The next Lemma states asymptotic convergence of bootstrap innovations to theoretical innovations, and is similar to Lemma 5.4 in Buhlmann (1997).

Lemma 3: Assume that assumptions given in A1, A2 and B hold. Then, for each fixed $t \in \mathbb{N}$,

$$\epsilon_{t,n}^{*} \xrightarrow{d^*} \epsilon_t, \text{in probability.}$$

Proof: Let $F_{\epsilon,n}(x) = (n-p)^{-1}\sum_{t=p+1}^{n}1_{[\epsilon_t \leq x]}$, $F_{\epsilon}(x) = \mathbb{P}[\epsilon_t \leq x]$ for $x \in \mathbb{R}$, and denote the Mallows metric by $d_2(.,.)$. Then, from standard results it follows that $d_2(F_{\epsilon,n}, F_{\epsilon}) = o_{a.s.}(1)$. Thus we need to only show that $d_2(\hat{F}_{\epsilon,n}, F_{\epsilon,n}) = o_p(1)$. Let S be uniformly distributed on $\{p+1, ..., n\}$ and let $Z_1 = \epsilon_S$, $Z_2 = \bar{\epsilon}_S$, where $\bar{\epsilon}_{t,n} = \hat{\epsilon}_{t,n} - \hat{\epsilon}_n^{(.)}$. Then, $d_2(\hat{F}_{\epsilon,n}, F_{\epsilon,n})^2 \leq E|Z_1 - Z_2|^2 = (n-p)^{-1}\sum_{t=p+1}^{n}(\bar{\epsilon}_{t,n}-\epsilon_t)^2 = (n-p)^{-1}\sum_{t=p+1}^{n}(\hat{\epsilon}_{t,n}-\hat{\epsilon}_n^{(.)}-\epsilon_t)^2$. From the proof of Lemma 2, $\hat{\epsilon}_n^{(.)} = o_p(1)$ and $(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n}-\epsilon_t| = o_p(1)$.

Hence $d_2(\hat{F}_{\epsilon,n}, F_{\epsilon,n}) = o_p(1)$. $\square$

The following proposition is analogous to Proposition 1 of Alonso et *al.* (2003) and shows that the bootstrap autoregressive coefficients obtained in Step 6 converge to the autoregressive coefficients of the fitted model obtained in Step 2.

Proposition 4.1. *Assume A1, A2 and B hold. Then, in probability*

$$\max_{1\le j\le p(n)} |\hat{\phi}^*_{j,p,n} - \hat{\phi}_{j,p,n}| \xrightarrow{P^*} 0.$$

Proof. The vector $\hat{\phi}^*_p$ is defined by the bootstrap empirical YuleWalker equations.
$\hat{\mathbf{\Gamma}}^*_{p,n}\hat{\boldsymbol{\phi}}^*_{p,n} = -\hat{\boldsymbol{\gamma}}^*_{p,n}$, where $\hat{\mathbf{\Gamma}}^*_{p,n} = [\hat{R}^*(i-j)]^p_{i,j=1}, \hat{\boldsymbol{\gamma}}^*_p = (\hat{R}^*(1), ..., \hat{R}^*(p))'$,
and $\hat{R}^*(j) = n^{-1}\sum_{t=1}^{n-|j|}(y^*_t - \bar{y^*})(y^*_{t+|j|} - \bar{y^*})$.
Therefore,

$$\begin{aligned}\| \hat{\boldsymbol{\phi}}^*_{p,n} - \hat{\boldsymbol{\phi}}_{p,n} \|_\infty =& \| \hat{\mathbf{\Gamma}}^{*-1}_{p,n}\hat{\boldsymbol{\gamma}}^*_{p,n} - \hat{\mathbf{\Gamma}}^{-1}_{p,n}\hat{\boldsymbol{\gamma}}_{p,n} \|_\infty \\ \le& \| \hat{\mathbf{\Gamma}}^{*-1}_{p,n} - \hat{\mathbf{\Gamma}}^{-1}_{p,n} \|_{row} \| \hat{\boldsymbol{\gamma}}^*_{p,n} \|_\infty + \| \hat{\mathbf{\Gamma}}^{-1}_{p,n} \|_{row} \| \hat{\boldsymbol{\gamma}}^*_{p,n} - \hat{\boldsymbol{\gamma}}_{p,n} \|_\infty .\end{aligned}$$

Assumptions A1, A2, B together with Theorem 1 of Poskitt (2006) and results in Hannan and Kavalieris (1986) can be utilized to show that $\| \hat{\mathbf{\Gamma}}_{p,n} \|_{row}$ and $\| \hat{\mathbf{\Gamma}}^{-1}_p \|_{row}$ are uniformly bounded in p. Following Alonso (2003), it is sufficient to show convergence of $\| \hat{\boldsymbol{\gamma}}^*_{p,n} - \hat{\boldsymbol{\gamma}}_{p,n} \|_2$ to zero in probability to establish the convergence of $\| \hat{\boldsymbol{\phi}}^*_{p,n} - \hat{\boldsymbol{\phi}}_{p,n} \|_\infty$. Now,

$$\begin{aligned}\| \hat{\boldsymbol{\gamma}}^*_{p,n} - \hat{\boldsymbol{\gamma}}_{p,n} \|^2_2 &= \sum_{k=1}^{p}(\hat{R}^*(k) - \hat{R}(k))^2 \\ &\le 2\sum_{k=1}^{p}[\hat{R}^*(k) - E^*\hat{R}^*(k)]^2 + 2\sum_{k=1}^{p}[E^*\hat{R}^*(k) - \hat{R}(k)]^2 \\ &= 2(S_1 + S_2),\end{aligned}$$

and

$$
\begin{aligned}
S_2 &= \sum_{k=1}^{p}(E^*[\epsilon_1^{*^2}]\sum_{i=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{i+k,n} - E[\epsilon_1^2]\sum_{i=0}^{\infty}\psi_i\psi_{i+k})^2 \\
&= \sum_{k=1}^{p}(E^*[\epsilon_1^{*^2}]\sum_{i=0}^{\infty}(\hat{\psi}_{i,n}\hat{\psi}_{i+k,n} - \psi_i\psi_{i+k}) + (E^*[\epsilon_1^{*^2}] - E[\epsilon_1^2])\sum_{i=0}^{\infty}\psi_i\psi_{i+k})^2 \\
&\leq 2\sum_{k=1}^{p}\left(E^*[\epsilon_1^{*^2}]\sum_{i=0}^{\infty}(\hat{\psi}_{i,n}\hat{\psi}_{i+k,n} - \psi_i\psi_{i+k})\right)^2 \\
&+ 2\sum_{k=1}^{p}\left((E^*[\epsilon_1^{*^2}] - E[\epsilon_1^2])\sum_{i=0}^{\infty}\psi_i\psi_{i+k}\right)^2 \\
&= I_1 + I_2.
\end{aligned}
$$

The $MA(\infty)$ transfer function of $\hat{\Phi}_{p,n} = 1 + \hat{\phi}_{1,p,n}z + ... + \hat{\phi}_{p,p,n}z^p$ is $\hat{\Psi}_{p,n} = 1/\hat{\Phi}_{p,n}$. Let x_t be the underline process of this $MA(\infty)$ transfer function. Then, $x_t = \sum_{j=0}^{\infty}\hat{\psi}_{j,n}\eta_{t-j}$, where η_t is i.i.d. with $E[\eta_t] = 0$ and $E[\eta^2] = \sigma^2$ for $t \in \mathbb{Z}$. The autocovariace function of x_t is $\gamma_x(k) = \sigma^2\sum_{j=0}^{\infty}\hat{\psi}_{j,n}\hat{\psi}_{j+k,n}$. However, as was observed by Bühlmann (1995) in the proof of Theorem 3.2, $\gamma_x(k) = \hat{R}_y(k)$ for $0 \leq k \leq p$. Observe that

$$
\begin{aligned}
I_1 &= 2E^*[\epsilon_1^{*^2}]^2\sigma^{-4}\sum_{k=1}^{p}(\hat{R}(k) - \gamma(k))^2 \\
&\leq 2E^*[\epsilon_1^{*^2}]^2\sigma^{-4}p\max_{1\leq k\leq p}(\hat{R}(k) - \gamma(k))^2 = O_{a.s.}\{[log(n)/n]^{1/2-d}\}.
\end{aligned}
$$

Moreover, $I_2 = o_p(1)$ since, $E^*[\epsilon_1^{*^2}] - E[\epsilon_1^2] = o_p(1)$ by Lemma 3.3 and
$\sum_{i=0}^{\infty}\psi_i\psi_{i+k} \leq (\sum_{i=0}^{\infty}\psi_i^2)^{1/2}\left(\sum_{i=0}^{\infty}\psi_{i+k}^2\right)^{1/2} \leq \sum_{i=0}^{\infty}\psi_i^2 < \infty$.
Therefore, we have $S_2 = O_{a.s.}\{[log(n)/n]^{1/2-d}\}$.

Next we bound S_1 using the same technique as Alonso (2002).

$$\begin{aligned}
S_1 &= \sum_{k=1}^{p}(n^{-1}\sum_{t=1}^{n-k}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{j,n}\epsilon^*_{t-i}\epsilon^*_{t+k-j} \\
&\quad - \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{j,n}E^*[\epsilon^{*2}_1]\delta_{i+k,j})^2 \\
&= \sum_{k=1}^{p} n^{-2}\sum_{t,s=1}^{n-k}\sum_{i,j=0}^{\infty}\sum_{h,l=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{j,n}\hat{\psi}_{h,n}\hat{\psi}_{l,n} \\
&\quad \times (\epsilon^*_{t-i}\epsilon^*_{t+k-j} - E^*[\epsilon^{*2}_1]\delta_{i+k,j})(\epsilon^*_{s-h}\epsilon^*_{s+k-l} - E^*[\epsilon^{*2}_1]\delta_{h+k,l}),
\end{aligned}$$

where $\delta_{i,j} = 1if i = j$, and 0 otherwise. Taking E^* on S_1,

$$E^*[S_1] = \sum_{k=1}^{p} n^{-2}\sum_{t,s=1}^{n-k}\sum_{i,j=0}^{\infty}\sum_{h,l=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{j,n}\hat{\psi}_{h,n}\hat{\psi}_{l,n} \tag{4.9}$$

$$\times \left(E^*[\epsilon^*_{t-i}\epsilon^*_{t+k-j}\epsilon^*_{s-h}\epsilon^*_{s+k-l}] - (E^*[\epsilon^*_1]^2)^2\delta_{i+k,j}\delta_{h+k,l}\right)$$

The expectation of error terms are given by,

$$E^*[\epsilon^*_{t-i}\epsilon^*_{t+k-j}\epsilon^*_{s-h}\epsilon^*_{s+k-l}] = \begin{cases} E^*[\epsilon^{*4}_1] & \text{if } t-i = t+k-j = s-h = s+k-l \\ E^*[\epsilon^{*2}_1]^2 & \text{if two pairs different indices} \\ 0 & \text{Otherwiswe} \end{cases}$$

and

$$E^*[\epsilon^*_{t-i}\epsilon^*_{t+k-j}\epsilon^*_{s-h}\epsilon^*_{s+k-l}] - (E^*[\epsilon^*_1]^2)^2\delta_{i+k,j}\delta_{h+k,l} =$$

$$\begin{cases} E^*[\epsilon^{*4}_1] - E^*[\epsilon^{*2}_1]^2 & \text{if } t-i = t+k-j = s-h = s+k-l \\ 0 & \text{if } t-i = t+k-j \neq s-h = s+k-l \\ E^*[\epsilon^{*2}_1]^2 & t-i = s-h \neq t+k-j = s+k-l \\ & t-i = s+k-l \neq s-h = t+k-l \\ 0 & \text{Otherwiswe} \end{cases}$$

Since $\sum_{i=0}^{\infty}\hat{\psi}_{i,n} < \infty$ (causality of the model fitted in Step 2 of Section 3), $\sum_{i,j,h,l=0}^{\infty}\hat{\psi}_{i,n}\hat{\psi}_{j,n}\hat{\psi}_{h,n}\hat{\psi}_{l,n} < \infty$ and using the fact that for fixed i, j, h and l, $\sum_{t=1}^{n-k}(\cdot)$

contains at most $n-k$ non-zero summands, we have $E^*[S_1] = O_p(pn^{-1})$. The preceding argument was used by Alonso (2003) in the proof of Proposition 1.

Finally, $\| \hat{\gamma}_p^* - \hat{\gamma}_p \|_2 = O_{a.s.}\{[log(n)/n]^{(1/2-d)/2}\}$, and therefore,

$\max_{1\leq j\leq p(n)} |\hat{\phi}_j^* - \hat{\phi}_j| = p^{1/2} O_{a.s.}\{[log(n)/n]^{(1/2-d)/2}\} = o_{a.s.}(1)$, which completes the proof.

□

We now establish the main result for our approach of obtaining prediction intervals for *FARIMA* processes, which is equivalent to the Theorem 1 of Alonso (2003).

Theorem 4.2. *Assume that A1, A2 and B hold with* $0 < d < 0.5$. *Then, in probability,*

$$y_{n+h}^* \xrightarrow{d^*} y_{n+h}, \text{ for } h = 0, 1, ...$$

Proof. Observe that,

$$y_{n+h} = -\sum_{j=1}^{\infty} \phi_j y_{n+h-j} + \epsilon_{n+h}, \tag{4.10}$$

$$y_{n+h}^* = -\sum_{j=1}^{\infty} \hat{\phi}_{j,p,n} y_{n+h-j}^* + \epsilon_{n+h,n}^*, \tag{4.11}$$

where $y_t^* = y_t$ for $t \leq n$. For simplicity of notation, we prove the theorem for $h = 1$.

From Lemma 3.4, $\epsilon_{n+1}^* \xrightarrow{d^*} \epsilon_{n+1}$ and thus we need only to show that the difference of the first terms on the right hand side of (3.4) and (3.5) converges to zero in probability. Therefore consider,

$$\begin{aligned} &-\sum_{j=1}^{\infty} (\hat{\phi}_{j,p,n} - \phi_j) y_{n+1-j} \\ &= -\sum_{j=1}^{p(n)} (\hat{\phi}_{j,p,n} - \phi_j) y_{n+1-j} + \sum_{j=p(n)+1}^{\infty} \phi_j y_{n+1-j} = S_{1,1} + S_{2,1}. \end{aligned}$$

Now, $E[|S_{2,1}|] \leq E[|y_t|] \sum_{j=p(n)+1}^{\infty} |\phi_j| = o(p^{-1}) = o_p\{[log(n)/n]^{1/2-d}\}$.

Therefore, $S_{2,1} = o_p\{[log(n)/n]^{1/2-d}\}$. To get the convergence of the first term of $S_{1,1}$, we observe that

$$|S_{1,1}| \leq \left| \sum_{j=1}^{p(n)} (\hat{\phi}_{j,p,n} - \phi_{j,p}) y_{n+1-j} \right| + \left| \sum_{j=1}^{p(n)} (\phi_{j,p} - \phi_j) y_{n+1-j} \right| = I_1 + I_2$$

.

Clearly, $I_1 = o_p(1)$ since,

$$\begin{aligned} I_1 &\leq \left(\sum_{j=1}^{p(n)} (\hat{\phi}_{j,p,n} - \phi_{j,p})^2 \right)^{1/2} \left(\sum_{i=1}^{p(n)} y_{n+1-i}^2 \right)^{1/2} \\ &= \{o_{a.s.}[[log(n)/n]^{\frac{1/2-d}{2}}]\}\{O_p[p(n)^{1/2}]\} = o_p(1). \end{aligned}$$

For I_2 we apply the Baxter's inequality which was generalized by Inoue and Kasahara (2006) for long memory processes for $0 < d < 0.5$.

$$\sum_{j=1}^{p(n)} |\phi_{j,p} - \phi_j| \leq b \sum_{j=p(n)+1}^{\infty} |\phi_j|$$

where b is a constant depending on the true structure. Therefore,

$$E[I_2] \leq E|y_t| \sum_{j=1}^{p(n)} |\phi_{j,p} - \phi_j| = O(p^{-1}) = O\{[log(n)/n]^{1/2-d}\}$$

which implies,

$$|S_{1,1}| = O_p\{[log(n)/n]^{1/2-d}\}.$$

Finally, $-\sum_{j=1}^{\infty} \hat{\phi}_{j,p,n} y_{n+1-j} = -\sum_{j=1}^{\infty} \phi_j y_{n+1-j} + O_p\{[log(n)/n]^{1/2-d}\}$.
Then, $y_{n+1}^* \xrightarrow{d^*} y_{n+1}$, in probability.

□

5. CONCLUSION

The sieve bootstrap method currently available for constructing prediction intervals for short-memory linear processes is modified to enable its application to $FARIMA$ processes. The asymptotic validity of this modified sieve bootstrap based prediction intervals under certain regularity conditions is established. The proposed method is based on the version of sieve bootstrap introduced by Alonso et *al.* (2003), and the fundamental work in Poskitt (2006, 2007) provides the theoretical foundation for the modifications to the existing procedure.

6. REFERENCES

[1] Alonso A. M., Pena D., Romo J., 2002. Forecasting time series with sieve bootstrap. Journal of Statistical Planning and Inference. 100, 1-11.

[2] Alonso A. M., Pena D., Romo J., 2003. On sieve bootstrap prediction intervals. Statistics and Probability Letters. 65, 13-20.

[3] Alonso A. M., Pena D., Romo J., 2004. Introducing model uncertainty in time series bootstrap. Statistica Sinica. 14, 155-174.

[4] An, H. Z., Chen, Z. G., Hannan, E.J., 1982. Autocorrelation, autoregression and autoregressive approximations. Ann. Statist. 10, 926-936.

[5] Baillie R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics. 73, 5-59.

[6] Brockwell P.J., Davis R.A., 1991. Time Series: Theory and Methods. Second Edition, Springer.

[7] Bühlmann P., 1995. Moving-average representation of autoregressive approximations. Stochastic Process. Appl. 60, 331342.

[8] Bühlmann P., 1997. Sieve bootstrap for time series. Bernoulli, 3, 123-148.

[9] Findley D.F., 1986. On Bootstrap estimates of forecast mean square errors for autoregressive processes. Computer Science and Statistics: Proceedings for the 17th Symposium on the Interface, 11-17.

[10] Granger, C.W.J., Joyeux, R., 1980. An introduction to long memory time series and fractional differencing. Journal of Time Series Analysis. 1, 15-39.

[11] Hannan, E.J., Kavalieris, L., 1986. Regression, autoregression models. Journal of Time Series Analysis. 7, 2749.

[12] Hosking, J.R.M., 1981. Fractional differencing. Biometrika. 68, 165-176.

[13] Inoue, A., Kasahara Y., 2006. Explicit representation of finite predictor coefficients and its applications. The Annals of Statistics. 34, 973-993.

[14] Künsch H.R., 1989. The jackknife and the bootstrap for general stationary observations. The Annals of Statistics. 17, 1217-1241.

[15] Mukhopadhyay P. and. Samaranayake V.A., 2010. Prediction intervals for time series: A Modified Sieve Bootstrap Approach. Communications in Statistics - Simulation and Computation. 39, 517-538.

[16] Poskitt D. S., 2006. Autoregressive approximation in nonstandard situations: the fractionally integrated and non-invertible cases. Annals of the Institute of Statistical Mathematics. 59, 697725

[17] Poskitt D. S., 2007. Properties of the Sieve bootstrap for fractionally integrated and non-invertible processes. Journal of Time Series Analysis. 29, 224-250.

[18] Stine R.A., 1982. Prediction intervals for time series. Unpublished Ph.D. dissertation, Princeton University, Dept of Statistics.

[19] Stine R.A., 1987. Estimating properties of autoregressive forecasts. Journal of the American Statistical Association. 82, 1072-1078.

[20] Taqqu M., Doukhan P., Oppenheim G., 2003. Theory and Application of Long-Range Dependence. Birkhãuser.

II. OBTAINING PREDICTION INTERVALS FOR *FARIMA* PROCESSES USING SIEVE BOOTSTRAP

ABSTRACT

The Sieve Bootstrap method for constructing prediction intervals for invertible ARMA processes is based on re-samples of residuals obtained by fitting a finite degree autoregressive approximation to the time series. The advantage of this approach is that it does not require the knowledge of the orders, p and q, associated with the ARMA model. The application of this method has been, up to now, limited to ARMA processes whose autoregressive polynomials do not have fractional roots. In this paper, we propose the sieve bootstrap method to obtain prediction intervals for ARFIMA (p, d, q) processes with $0 < d < 0.5$. The proposed procedure is a simpler alternative to an existing method, which requires the estimation of p, d, and q. Monte-Carlo simulation studies, carried out under the assumption of normal, mixture of normals, and exponential distributions for the innovations, show near nominal coverages for short term and long term prediction intervals under all situations. In addition, the proposed method is more precise than the existing method in most cases.

Keywords: Forecasting, Long Memory Processes, Fractionally Integrated Time Series, Model-based Bootstrap, ARFIMA processes.

1. INTRODUCTION

The modelling and forecasting of long-memory processes have become an important aspect for time series analysts. For example, empirical series that exhibit long-memory are quite common in geophysical sciences, macroeconomics, asset pricing, stock returns and exchange rates (see [1,12]). The Fractionally Integrated Auto Regressive Moving Average (ARFIMA or FARIMA) processes have been used extensively to model such processes, see [24, 25]. For example, Liu et *al.* [26] suggests modelling actual web traffic using a FARIMA process. With the prevalence of empirical processes that are well approximated by FARIMA models, there is the corresponding need for methods of obtaining prediction intervals for such processes. In this paper, a relatively simple method for obtaining bootstrap-based prediction intervals for FARIMA processes is presented and compared against an existing method through a Monte Carlo simulation study.

The proposed method is based on the *Sieve Bootstrap* approach of Alonso et *al.* [6,7]. Monte Carlo study shows that in many cases the sieve bootstrap method performs better than the only currently available bootstrap-based method introduced by Bisaglia and Grigoletto [4].

An extensive discussion of research literature on the application of bootstrap techniques to AR and ARMA models are discussed in [5–8]. For bevirity, only a brief discussion of the literature on the sieve bootstrap is presented here.

As the Gaussian-based prediction intervals produce poor coverages when the distributional assumptions are violated, Stine [9,10] and Findley [11] were the first to introduce bootstrap methods to compute prediction mean squared error for time series. Thombs and Schucany [16] presented a bootstrap procedure for obtaining forecast intervals, but their method required the backward representation of the time series, which was not possible for time series with a moving average component. All the above methods also assumed that the order of the process is known. A block bootstrap method for stationary processes with unknown orders to compute the empirical distributions of statistics was first introduced by Künsch [13]. The sieve bootstrap approach, which also do not require the knowledge

of the orders associated with the underline process, was originated by Bühlmann [27]. He achieved this by expressing the process as a truncated infinite order AR model. It should be noted that the use of Akaike's Information Criteria (AIC) to approximate AR models was introduced by Grenander [19] and Geman and Hwang [20]. Alonso et *al.* [6–8] further presented the sieve bootstrap approach to obtain prediction intervals for a general class of linear processes that include ARMA processes as a subset. In the 2004 article, the authors discussed two variations of the sieve bootstrap method; one variation is based on the moving block bootstrap introduced by Künsch [13] and the other uses the information criterion function order distribution. A modified version of the sieve bootstrap method of Alonso et *al.* [8] was implemented and applied to ARMA processes by Mukhopadhyay and Samaranayake [5]. They were able to improve the coverage probabilities of prediction intervals with their modifications.

1.1. FARIMA PROCESSES

A good introduction to the mathematical background of the FARIMA processes is given by Brockwell & Davis [2]. Based on their definition, the process $\{X_t : t = 0, \pm 1, ...\}$ is said to be a FARIMA(0,d,0) process with $d \in (-0.5, 0.5)$ if $\{X_t\}$ is a zero mean stationary solution to the difference equation,

$$\bigtriangledown^d X_t = \epsilon_t, t \in \mathbb{Z}, \tag{1.1}$$

where $\{\epsilon_t\} \sim WN(0, \sigma^2)$.

The process $\{X_t\}$ is also known as fractionally integrated noise or fractional Gaussian noise (fGn). Here, $\bigtriangledown^d = (1-B)^d = \sum_{j=0}^{\infty} \pi_j B^j$, where B is the back-shift operator defined by $B^k X_t = X_{t-k}$ for $k = 1, 2, ...$, and

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{o<k\leq j} \frac{k-1-d}{k}, j = 0, 1, 2... \qquad ,$$

with $\Gamma(\cdot)$ representing the gamma function.

Notice that when $d = 1$, the process $\{X_t\}$ is a random walk, which will not be considered in our discussion. If $d \in (-0.5, 0.5)$ then there is a unique purely non-deterministic stationary solution $\{X_t\}$ of (1.1) given by

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} = \bigtriangledown^{-d} \epsilon_t.$$

where,

$$\psi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} = \prod_{o<k\leq j} \frac{k-1+d}{k}, j = 0, 1, 2... \qquad .$$

The autocorrelation function of $\{X_t\}$ is $\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} = \prod_{o<k\leq h} \frac{k-1+d}{k-d}, k = 1, 2, ...,$ and it can be shown that $\rho(h) \sim h^{(2d-1)}\Gamma(1-d)/\Gamma(d)$ as $h \to \infty$, which implies the long range dependence among $X_t's$ when $d \in (-0.5, 0.5)$.

The process $\{X_t : t = 0, \pm 1, ...\}$ is said to be a FARIMA(p, d, q) process with $d \in (-0.5, 0.5)$ if $\{X_t\}$ is stationary and satisfies the difference equation,

$$\phi(B) \bigtriangledown^d X_t = \theta(B)\epsilon_t,$$

where $\{\epsilon_t\} \sim WN(0, \sigma^2)$ with $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 + ... + \theta_q z^q$ representing polynomials of degrees p, q respectively.

If $d \in (-0.5, 0.5), \phi(z) \neq 0$ for $|z| \leq 1$, and $\phi(\cdot), \theta(\cdot)$ have no common zeros, then there is an infinite order moving average representation of $\{X_t\}$ which can be written as

$$X_t = \psi(B) \bigtriangledown^{-d} \epsilon_t, \tag{1.2}$$

where

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \theta(z)/\phi(z).$$

Note that the relationship between FARIMA processes and self-similar processes is addressed by [14]. The process is said to have long-range dependence or long-memory when $0 < d < 0.5$ and "intermediate memory" when $-0.5 < d < 0$. However, as many physical phenomena indicate long-range dependence [15], our discussion is limited to FARIMA processes with $0 < d < 0.5$.

1.2. PREDICTION & PREDICTION INTERVALS FOR FARIMA PROCESSES

Point predictors for future values of a FARIMA(p, d, q) process based on the innovations algorithm are presented by Brockwell & Davis [2]. Ray [28] finds that forecasts of a FARIMA process can be computed by fitting an AR(p) model, which is a part of the technique discussed in this paper. Other papers, for example, Brodsky and Hurvich [30], Ray [29], Geweke and Porter-Hudack [32] and Eisinga et *al.* [31], continue the discussion on forecasting a FARIMA process. More recently, Gonzaga et *al.* [3] introduces a wavelet based Bayesian estimation for predicting a Generalized FARIMA (p, d, u, q) process.

In spite of the availability of point predictors, there is a dearth of research publications on prediction intervals for a FARIMA(p, d, q) process. Bisaglia and Grigoletto [4] was the first to introduce bootstrap-based prediction intervals for FARIMA processes. The method introduced by these authors will be hereafter referred as B-G method. The B-G method performs quite well, providing near nominal coverages when the sample size is large. Their technique involves jointly estimating the fractional difference parameter, d as well as the AR and MA coefficients using the Whittle approximation [12, 13], which minimizes the variance of the underlying white noise process.

1.2.1. The B-G method.

An outline of the B-G algorithm is as follows. For a given long-memory process $\{X_t\}_{t=0}^{n}$, the B-G method first fits a FARIMA(p, d, q) model using the Bayesian Information Criterion (BIC) and the Whittle approximation, and then uses the residuals of the fitted model to obtain bootstrap replicates $\{X_t^*\}_{t=0}^{n+k}$, where k is the number of ahead-leads to be predicted. The first n replicates $\{X_t^*\}_{t=0}^{n}$ are then used to identify the model and

estimate parameters of the bootstrapped series using the BIC and the Whittle approximation. A finite approximation of the AR representation of the fitted model is used to compute k-step prediction and the prediction error. This step is performed 1,000 times to obtain the bootstrap distribution of prediction errors. Finally, the prediction interval is computed based on the percentiles of this bootstrap distribution of prediction errors.

The Whittle approximation for estimating parameters involved with a stationary Gaussian time series was proposed by Fox and Taqqu [13]. The application of Whittle estimators for FARIMA(p, d, q) time series is computationally demanding as the optimization step takes a long time even with fast computers. This computational demand hurts the B-G method even more as it fits FARIMA(p, d, q) model in each bootstrap run. However, our proposed sieve bootstrap method does not require such optimizations in estimating parameters associated with AR(p) models and hence it is computationally much faster.

2. THE PROPOSED SIEVE BOOTSTRAP METHOD

Let $\{X_t\}_{t=0}^n$ be a zero-mean FARIMA process defined as (1.1). Under the case of invertibility ($0 < d < 0.5$ and $\theta(z) \neq 0$ for $|z| \leq 1$ - see Brockwell & Davis [2] for details), it can be written as an infinite order autoregressive process $\sum_{j=0}^{\infty} \phi_j X_{t-j} = \epsilon_t$ for $t \in Z$ with $\phi_0 = 1$ and $\sum_{j=0}^{\infty} |\phi_j| < \infty$.

A direct application of the sieve bootstrap method proposed by Alonso et *al.* [6–8] is not feasible for FARIMA processes as the MA(∞) coefficients do not satisfy one of the key assumptions, namely $\sum_{j=0}^{\infty} j^r \psi_j < \infty$ for some $r \in \mathbb{N}_0$, made by the above authors.

Poskitt [33] discussed the ways of approximating a class of more general linear processes, which includes FARIMA processes, by finite autoregressive polynomials. In a later paper, Poskitt [34] showed how the sieve bootstrap method can be utilized to compute empirical distributions of statistics associated with FARIMA processes. He also derived the asymptotic properties of these empirical distributions under certain regularity conditions. The coefficients of the $AR(p)$ approximation of X_t, $\boldsymbol{\phi}_p = (\phi_1, ..., \phi_p)'$, are obtained

using the Yule-Walker equations, $\mathbf{\Gamma}_p\boldsymbol{\phi}_p = -\boldsymbol{\gamma}_p$, $\boldsymbol{\gamma}_p = (\gamma(1), ..., \gamma(p))'$, $\mathbf{\Gamma}_p = [\gamma(i-j)]_{i,j=1}^{p}$ with $\gamma(k) = E[(X_t - \mu_X)(X_{t+k} - \mu_X)]$ for $k \in \mathbb{N}_0$ where, $\mu_X = E(X_t)$. Then the errors, $\epsilon_{t,p} = \sum_{j=0}^{p} \phi_j X_{t-j}$, converge to ϵ_t in mean square error as $p \to \infty$ [33].

Rupasinghe and Samaranayake [35] utilized some of results in Poskitt [33, 35] to establish asymptotic properties of the sieve bootstrap prediction intervals for $FARIMA$ processes generalizing Alonso *et al.* [7]. In this paper, we continue our interest in the sieve bootstrap method by carrying out an extensive simulation study based on the method proposed by Rupasinghe and Samaranayake [35].

The optimal order p is selected by the AIC as recommended by Poskitt [35] but Alonso et *al.* [7] preferred the corrected AIC (AICC) for ARMA processes. To find the optimal order using the AIC criteria, one needs a maximum order p_{max} to be specified. Following Poskitt [33,35], p_{max} is set to 20 and 27 for sample sizes 100 and 200 respectively as the long-range dependence has to be captured by fitting a large order AR model. In fact, Poskitt suggested, $p_{max} = [log(n)]^{1.962}$.

The following steps are required to compute prediction intervals for FARIMA processes by Rupasinghe and Samaranayake [35].

1. Given a realization, $\{X_t\}_{t=1}^{n}$ of a FARIMA process, select a maximum order p_{max}. Then, find the optimal order p_{AIC} by the AIC criterion among the values $p = 1, 2, ..., p_{max}$. Based on our initial investigations, we recommend the value of $p_{max} = 20, 27$ for $n = 100, 200$ respectively.

2. Estimate the coefficients, $\hat{\phi}_1, ..., \hat{\phi}_{\hat{p}}$ of the AR($\hat{p}$) process by Yule Walker or least-squares method. Alonso et *al.* [6–8]) uses the Yule Walker method but Mukhopadhyay and Samaranayake [5] recommends the least-squares method. The Yule-Walker method is used in our study as well.

3. Compute the $(n - \hat{p})$ residuals as $\tilde{\epsilon}_t = \sum_{j=0}^{\hat{p}} \hat{\phi}_j (X_{t-j} - \bar{X})$; $\hat{\phi}_0 = 1, t \in (\hat{p}, ..., n)$, where $\bar{X}$ is the mean of $\{X_t\}_{t=1}^{n}$.

4. The residuals need to be centered when using Yule Walker method [16]. These rescaled residuals are denoted by $\hat{\epsilon}_t, t \in (\hat{p}, ..., n)$.

5. Compute the empirical distribution function of the residuals as $\hat{F}_{\hat{\epsilon}}(x) = (n-\hat{p})^{-1}\sum_{t=\hat{p}+1}^{n} I_{(-\infty,x]}(\hat{\epsilon}_t)$.

6. Then resample, with replacement, the bootstrap innovations, ϵ_t^*, for $t = -199, -198, ..., 0, 1, 2, ..., n$, from this distribution.

7. Generate the bootstrapped series $X_t^*, t = -199, -198, ..., 0, 1, 2, ..., n$ based on the recursion $\sum_{j=0}^{\hat{p}} \hat{\phi}_j(X_{t-j}^* - \bar{X}) = \epsilon_t^*$ with $X_t^* = \bar{X}$ for $t = -199, ..., \hat{p}$. The non-positive lags represent "burn-in" observations which need to be dropped to make the effect of initial values negligible.

8. Fit an AR($\hat{p}$) model to X_t^* using the Yule-Walker method and let the estimated AR coefficients be $\hat{\phi}_1^*, \hat{\phi}_2^*, ..., \hat{\phi}_{\hat{p}}^*$. Note that the same order, $\hat{p}$ obtained in the Step 1 is used in this step as well. Alonso et *al.* [8] suggested the use of AICC to find the order, which utilizes to capture uncertainties due to model being different, instead of using the same $\hat{p}$. However, our initial simulation studies show that the use of the same $\hat{p}$ obtained in the Step 1 yields better coverages.

9. Using the new coefficients $\hat{\phi}_1^*, \hat{\phi}_2^*, ..., \hat{\phi}_{\hat{p}}^*$ obtained in the previous step, compute the k-step ahead bootstrap observations by the recursion as follows: $X_{n+k}^* - \bar{X} = -\sum_{j=1}^{\hat{p}} \hat{\phi}_j^*(X_{n+k-j}^* - \bar{X}) + \epsilon_{n+k}^*$, where $k > 0$ and $X_t^* = X_t$ for $t \leq n$. The bootstrap distribution of X_{n+k} should be conditioned on the original observed data rather the bootstrap $\{X_t^*\}_{t=1}^n$ by setting $X_t^* = X_t$ for $t \leq n$ as implemented by Cao et *al.* [23] and Alonso et *al.* [6, 8].

10. Obtain the bootstrap distribution of X_{n+k}, denoted by $\hat{F}^*_{\hat{X}^*_{n+k}}(.)$, by repeating the Steps 6 to 9 B times, where B is set to be 1,000 in the simulation study.

11. A $100(1-\alpha)\%$ prediction interval for X_{n+k} is then computed by $[Q^*(\alpha/2), Q^*(1-\alpha/2)]$ where $Q^*(s) = \hat{F}^{*-1}_{\hat{X}^*_{n+k}}(s)$ is the s^{th} percentile of the estimated bootstrap distribution. Finally, the lower and upper bounds of the prediction interval are obtained by sorting the bootstrapped future values $\hat{X}^*_{n+k}$ and choosing the $(\alpha/2)100^{th}$ and the $(1-\alpha/2)100^{th}$ percentile points.

3. SIMULATION STUDY

In order to investigate the performances of the method proposed in this paper and contrast it with the intervals obtained by the B-G method, a Monte-Carlo simulation study was carried out with three different error distributions and with sample sizes 100 and 200. The coverage, bootstrap length, and the length of the interval theoretically achievable under known order and parameter values were computed for 95% and 99% prediction intervals to asses the performance of the two methods.

The representation of X_t in (1.2) was used to generate FARIMA(p, d, q) processes assuming that the negative lags of ϵ_t are zeros. So, one can write $X_t = \sum_{j=0}^{t} \lambda_j Z_{t-j}$ where $\lambda(z) = \psi(z)(1-z)^{-d} = \sum_{j=0}^{\infty} \psi_j z^j \sum_{k=0}^{\infty} b_k z^k = \sum_{k=0}^{\infty} \lambda_k z^k$ with $(1-z)^{-d} = \sum_{k=0}^{\infty} b_k z^k$.

The models employed in this study are FARIMA(0,d,0), FARIMA(1,d,0), FARIMA(0,d,1) and FARIMA(1,d,1) processes with $d \in \{0.1, 0.25, 0.4, 0.49\}$, $\phi_1 \in \{0, 0.5\}$ and $\theta_1 \in \{0, -0.8\}$. Bisaglia and Grigoletto [4] also studies the same models except FARIMA(1,d,1). The standard normal distribution, exponential distribution with mean 1 centered at zero, and the skewed bimodal mixture distribution $0.9F_1 + 0.1F_2$, where $F_1 \sim N(-1, 1)$ and $F_2 \sim N(9, 1)$, were considered for error distributions for each of the above combinations of d, ϕ_1 and θ_1 but Bisaglia and Grigoletto [4] looked at only normal errors. Prediction intervals for leads $k = 1, 10, 20$ were computed using the both methods. The Matlab (version 2008b) software was used for these simulations. Since, an important part of this paper is to compare our prediction intervals with that of the B-G method, simulations on some of the above combinations of d, ϕ_1 and θ_1 were also run for the B-G method. As was mentioned in [4], the simulation study of the B-G method has been coded in Gauss by the authors. We ran the authors' code in Gauss version 11 to implement the prediction intervals of the B-G method. In order that the two methods are compared using series generated by the same mechanism, we changed their generating algorithm to match ours. Test runs were made to verify that both algorithms produce similar results for the B-G method.

For each combination of model, sample size, nominal coverage and error distribution, $N = 1,000$ independent series were generated and for each of these simulated series, steps 1 to 15 were implemented. To compute the coverage probabilities for each of this N simulations, $R = 1,000$ future observations (X_{n+k}) were generated using the original model.

The proportion of those falling in between the lower and upper bounds of the bootstrap prediction interval was then defined to be the coverage. Thus, the coverage at the i^{th} simulation run is given by $C(i) = R^{-1} \sum_{r=1}^{R} I_A[X_{n+k}^r(i)]$ where $A = [Q^*(\alpha/2), Q^*(1-\alpha/2)]$, $I_A(.)$ is the indicator function of the set A and $X_{n+k}^r(i)$, $r = 1, 2, ...1,000$ are the R future values generated at the i^{th} simulation run. The bootstrap length and theoretical length for the i^{th} simulation run are given by $L_B(i) = Q^*(1 - \alpha/2) - Q^*(\alpha/2)$ and $L_T(i) = X_{n+k}^r(1 - \alpha/2) - X_{n+k}^r(\alpha/2)$ respectively. $L_T(i)$ is the difference between the $100(1-\alpha/2)^{th}$ and $100(\alpha/2)^{th}$ percentile points the empirical distribution of the 1,000 future observations that were generated using the underlying time series model with known order and the true values of the coefficients. Using these statistics, the mean coverage, mean length of bootstrap prediction intervals, mean length of theoretical intervals, and their standard errors are computed as:

Mean Coverage $\bar{C} = N^{-1} \sum_{i=1}^{N} C(i)$

Standard Error of Mean Coverage $SE_{\bar{C}} = \{[N(N-1)]^{-1} \sum_{i=1}^{N} [C(i) - \bar{C}]^2\}^{1/2}$

Mean Length (bootstrap) $\bar{L}_B = N^{-1} \sum_{i=1}^{N} L_B(i)$

Standard Error of Mean Length $SE_{\bar{L}_B} = \{[N(N-1)]^{-1} \sum_{i=1}^{N} [L_B(i) - \bar{L}_B]^2\}^{1/2}$

Mean Theoretical Length $\bar{L}_B = N^{-1} \sum_{i=1}^{N} L_B(i)$

In total 192 different combinations of model type, sample size, nominal coverage probability, and error distributions were investigated in this simulation study. However, due to space limitations, we report only a representative sample of results for 95% intervals, in Table 1 through 7. These tables report the mean coverage, mean interval length, and mean theoretical length, standard error of mean coverage and standard error of mean interval length. Tables 1 to 3, 4 to 5, and 6 to 7 represent coverage probabilities and lengths of computed prediction intervals for normal, exponential and t distribution errors

respectively. The complete results of the simulation study are available upon request from the corresponding author.

To investigate the behaviour of the intervals for each of the 192 combinations, the minimum value, percentiles (25^{th}, 50^{th}, and 75^{th}), and the maximum value of (a) the coverage probabilities, (b) the bootstrap interval bounds (upper and lower), and (c) the theoretical interval bounds (upper and lower), were further computed, based on the 1,000 values generated through simulation, and these statistics are also available upon request.

From Tables 1-7, we can see that the both methods provide coverages closer to the nominal coverage as sample size increases. This is expected since large sample sizes provide more accurate parameter estimates as well as yield more residuals for resampling.

An interesting observation is that the mean coverages of the proposed Sieve Bootstrap (SB) method are very close to but just below the nominal coverage for all the lags while that of the B-G method are more conservative for lags 10 and 20 for normal errors as seen in Tables 1 to 3. Also, it is observed that the mean bootstrap interval lengths of the B-G method are larger than the that of the SB method and the theoretical lengths. For instance, in table 1, the mean coverages for $k = 10$, for the case with 200 observations are 0.9486 and 0.9650 while the prediction intervals' lengths are 4.2493 and 4.5912 for the SB and B-G respectively. In general, for normally distributed errors, the SB method provides coverages marginally below the nominal level while slightly wider intervals with conservative coverages were attained by the B-G method.

For exponential errors, both method yield nearly the same coverages which are close to the nominal level. However, it is very interesting to see that the SB prediction interval lengths are shorter even with high coverages, than the B-G prediction intervals. For example, the SB method outperforms the B-G method when $k = 1$ with sample size 100, with shorter prediction intervals as shown in Table 5. The coverages are 0.9536 and 0.9434 for SB and BG respectively, with SB having a shorter length of 4.0014.

The BG method fails to perform accurately for all the leads when the errors are skewed and bimodal. From Table 6 we can see that the BG method provides very liberal coverages for lead 1 while near nominal coverages are yielded for the leads 10 and 20. Notice that in the cases where near nominal coverages are obtained, the BG intervals

are exceedingly wider than than the SB and theoretical intervals. When the difference parameter, d is close to 0.5, the BG method performs even worst as shown in Table 7. It was unable to provide coverages close to the nominal level even with way wider intervals. The Whittle estimator used in BG procedure, fails to estimate the parameters accurately when the errors are skewed and bimodal, and this causes to produce poor coverages for the prediction intervals computed by the BG method.

Table 8 demonstrates average times taken by SB and BG methods to compute prediction intervals for different various values of difference parameter, d. Both methods took a longer time for larger values of difference parameters. However, the SB method is considerably faster than the BG method as expected earlier. The optimization process of the likelihood function in estimating the parameters causes the BG method to take a long time to compute prediction intervals.

Table 1. Coverage of 95% intervals for $\bigtriangledown^{0.25} X_t = \epsilon_t$ with normal errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9040	0.9395 (0.0025)	3.9010 (0.0338)	0.9488 (0.0098)	4.2147 (0.0364)
	200	3.9430	0.9413 (0.0021)	3.9217 (0.0308)	0.9538 (0.0016)	4.0377 (0.0199)
10	100	4.1691	0.9415 (0.0024)	4.1955 (0.0413)	0.9664 (0.0022)	4.7335 (0.0603)
	200	4.1684	0.9486 (0.0017)	4.2493 (0.0298)	0.9650 (0.0016)	4.5912 (0.0341)
20	100	4.2020	0.9421 (0.0025)	4.2174 (0.0425)	0.9651 (0.0027)	4.7969 (0.0681)
	200	4.1907	0.9480 (0.0020)	4.2647 (0.0337)	0.9662 (0.0016)	4.6592 (0.0385)

Table 2. Coverage of 95% intervals for $\bigtriangledown^{0.49} X_t = (1 - 0.8B)\epsilon_t$ with normal errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9247	0.9379 (0.0025)	3.9693 (0.0415)	0.9281 (0.0157)	4.2868 (0.0355)
	200	3.9003	0.9396 (0.0024)	3.9119 (0.0281)	0.9341 (0.0108)	4.0721 (0.0213)
10	100	4.1318	0.9456 (0.0021)	4.1803 (0.0372)	0.9733 (0.0020)	4.8421 (0.0531)
	200	4.1207	0.9457 (0.0017)	4.1266 (0.0251)	0.9655 (0.0024)	4.5659 (0.0248)
20	100	4.1218	0.9472 (0.0022)	4.1833 (0.0389)	0.9746 (0.0018)	4.8453 (0.0532)
	200	4.1338	0.9477 (0.0018)	4.1521 (0.0275)	0.9690 (0.0011)	4.5780 (0.0251)

Table 3. Coverage of 95% intervals for $(1 - 0.5B) \bigtriangledown^{0.49} X_t = (1 - 0.8B)\epsilon_t$ with normal errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9283	0.9402 (0.0028)	3.9743 (0.0364)	0.9528 (0.0042)	4.1702 (0.0365)
	200	3.9031	0.9441 (0.0020)	3.9681 (0.0240)	0.9515 (0.0015)	4.0182 (0.0187)
10	100	4.0172	0.9554 (0.0025)	4.3014 (0.0517)	0.9608 (0.0020)	4.3922 (0.0512)
	200	4.0146	0.9555 (0.0019)	4.2798 (0.0380)	0.9567 (0.0016)	4.2505 (0.0266)
20	100	4.0658	0.9518 (0.0027)	4.3139 (0.0588)	0.9585 (0.0021)	4.4081 (0.0530)
	200	4.0858	0.9572 (0.0021)	4.3995 (0.0474)	0.9548 (0.0016)	4.2689 (0.0278)

Table 4. Coverage of 95% intervals for $(1-0.5B)\bigtriangledown^{0.4}X_t = (1-0.8B)\epsilon_t$ with exponential errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.6677	0.9437 (0.0074)	3.9224 (0.0812)	0.9437 (0.0055)	4.2696 (0.1329)
	200	3.6367	0.9514 (0.0046)	3.8099 (0.0545)	0.9476 (0.0015)	4.0068 (0.0423)
10	100	3.7210	0.9584 (0.0031)	4.0577 (0.0809)	0.9526 (0.0019)	4.3695 (0.1518)
	200	3.7571	0.9535 (0.0039)	3.9549 (0.0569)	0.9503 (0.0014)	4.0623 (0.0427)
20	100	3.7836	0.9549 (0.0032)	4.0749 (0.0819)	0.9531 (0.0018)	4.3724 (0.1521)
	200	3.7811	0.9520 (0.0035)	3.9135 (0.0509)	0.9487 (0.0014)	4.0672 (0.0429)

Table 5. Coverage of 95% intervals for $\bigtriangledown^{0.49} X_t = (1 - 0.8B)\epsilon_t$ with exponential errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.6385	0.9536 (0.0052)	4.0014 (0.0721)	0.9434 (0.0079)	4.3515 (0.1528)
	200	3.6609	0.9475 (0.0074)	3.8874 (0.0546)	0.9457 (0.0050)	4.0584 (0.0454)
10	100	4.1541	0.9490 (0.0028)	4.2878 (0.0699)	0.9603 (0.0018)	4.9143 (0.1848)
	200	4.1295	0.9467 (0.0026)	4.1884 (0.0542)	0.9549 (0.0014)	4.5200 (0.0540)
20	100	4.1459	0.9503 (0.0027)	4.3202 (0.0713)	0.9601 (0.0020)	4.9189 (0.1850)
	200	4.1658	0.9513 (0.0022)	4.2786 (0.0527)	0.9555 (0.0015)	4.5282 (0.0541)

Table 6. Coverage of 95% intervals for $\bigtriangledown^{0.4} X_t = \epsilon_t$ with Mixture errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9150	0.9397 (0.0024)	3.9312 (0.0315)	0.8879 (0.0115)	13.0698 (0.1740)
	200	3.9149	0.9420 (0.0019)	3.9193 (0.0273)	0.9030 (0.0010)	12.5954 (0.1065)
10	100	4.6746	0.9478 (0.0032)	4.9539 (0.0724)	0.9434 (0.0031)	17.5663 (0.3263)
	200	4.6745	0.9471 (0.0028)	4.8746 (0.0548)	0.9457 (0.0025)	17.4127 (0.2354)
20	100	4.8267	0.9478 (0.0034)	5.1265 (0.0903)	0.9485 (0.0032)	18.4148 (0.3930)
	200	4.8360	0.9472 (0.0028)	5.0685 (0.0670)	0.9535 (0.0024)	18.3427 (0.2867)

Table 7. Coverage of 95% intervals for $\bigtriangledown^{0.49} X_t = (1 - 0.8B)\epsilon_t$ with Mixture errors

Leads	Size	TheoLen	SB		B-G	
			Coverage Mean (SE)	Length Mean (SE)	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9247	0.9379 (0.0025)	3.9693 (0.0415)	0.8958 (0.0116)	13.6464 (0.1829)
	200	3.9003	0.9396 (0.0024)	3.9119 (0.0281)	0.8935 (0.0081)	12.8626 (0.1242)
10	100	4.1318	0.9456 (0.0021)	4.1803 (0.0372)	0.9239 (0.0027)	15.5551 (0.2577)
	200	4.1207	0.9457 (0.0017)	4.1266 (0.0251)	0.9110 (0.0014)	14.3395 (0.1512)
20	100	4.1218	0.9472 (0.0022)	4.1833 (0.0389)	0.9244 (0.0026)	15.5723 (0.2601)
	200	4.1338	0.9477 (0.0018)	4.1521 (0.0275)	0.9130 (0.0012)	14.3678 (0.1508)

Table 8. Time comparison with sample size 200 (in seconds)

d	SB	BG
0.1	41.9	127
0.25	65	204
0.4	88.6	257.3
0.49	128.5	453.25

4. APPLICATION TO A REAL DATA SET

The proposed sieve bootstrap method have been applied to a historical time series of 663 annual minimal water levels of the River Nile, measured at Rodga Gorge (near Cairo) between 622 and 1284 A.D.; the data set is available in Tousson [36]. This time series also has been used by Bisaglia and Grigoletto [4] and they have fitted a $FARIMA(0, d, 0)$ with $d = 0.3842$ for the first 563 observations.

Following Bisaglia and Grigoletto [4], we used the observations for the years 622 through 1184 to build prediction intervals for the subsequent 100 years. The lower and upper bounds of the 95% SB and BG prediction intervals are given in Figure 1. The prediction intervals reported by Bisaglia and Grigoletto [4] were unable to capture the true time series for the years 1202 and 1231, whereas, the sieve bootstrap method failed to capture the true value only for the year 1231. Most beneficially, the sieve bootstrap approach produced more precise prediction intervals compared to BG method.

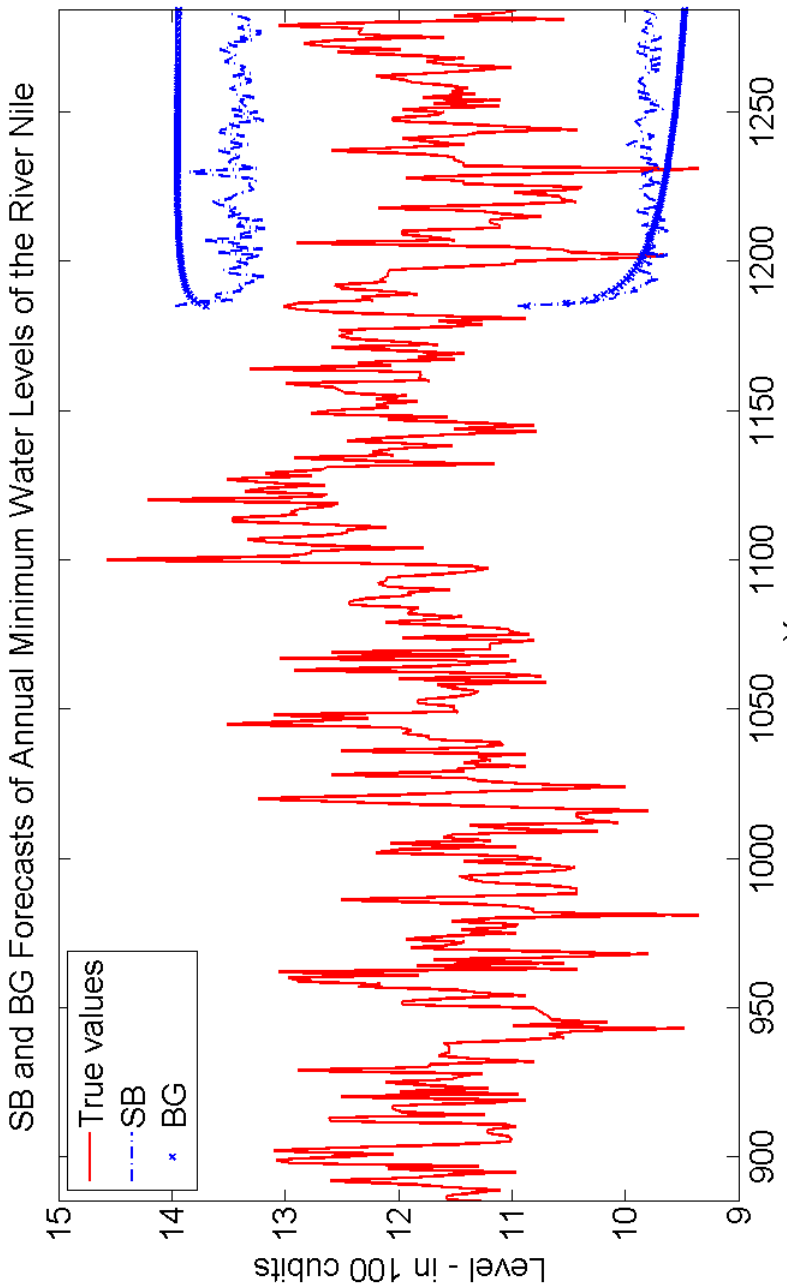

Figure 1. Lower and upper bounds of SB and BG prediction intervals for the Nile River annual minimal water levels

5. CONCLUSION

In this paper we have proposed a sieve bootstrap based prediction intervals for long memory time series that can be modeled using $FARIMA$ processes. The sieve bootstrap method produces coverages close to the nominal level with shorter intervals in all cases. In contrast, the BG method produces slightly better coverages with wider intervals in some cases, the proposed method performs consistently regardless of the error distribution. This was further confirmed by the application to the Nile River data set. In general, we can recommend the sieve bootstrap method over BG for faster, more accurate and precise prediction intervals for long memory processes.

6. REFERENCES

[1] R.T. Baillie, *Long memory processes and fractional integration in econometrics*, Journal of Econometrics. 73 (1996), pp. 5–59.

[2] P.J. Brockwell, and R.A. Davis, *Time Series: Theory and Methods*, Springer, New York, 1991.

[3] A. Gonzaga, and A. Kawanaka, *Wavelet-based Bayesian analysis of generalized long-memory process*, IEE ICASSP (2005).

[4] L. Bisaglia, and M. Grigoletto, *Prediction Intervals for FARIMA processes by bootstrap methods*, J. Statist. Comput. Simul., 68 (2001), pp. 185 - 201.

[5] P. Mukhopadhyay, and V.A. Samaranayake, *Prediction Intervals for Time Series: A Modified Sieve Bootstrap Approach*, Communications in Statistics - Simulation and Computation, 39 (2010), 517-538.

[6] A. M. Alonso, D. Peña, and J. Romo, *Forecasting time series with sieve bootstrap*, Journal of Statistical Planning and Inference, 100 (2002) 100, 1-11.

[7] A. M. Alonso, D. Peña, and J. Romo, *On sieve bootstrap prediction intervals*, Statistics and Probability Letters, 65 (2003), 13-20.

[8] A. M. Alonso, D. Peña, and J. Romo, *Introducing model uncertainty in time series bootstrap*, Statistica Sinica, 14 (2002), 155-174.

[9] R.A. Stine, *Prediction Intervals for Time Series*, Unpublished Ph.D. dissertation, Princeton University, Dept of Statistics.

[10] R.A. Stine, *Estimating Properties of Autoregressive Forecasts*, Journal of the American Statistical Association, (1987), 82, 1072-1078.

[11] D.F Findley, *On Bootstrap Estimates of Forecast Mean Square Errors for Autoregressive Processes*, Computer Science and Statistics: Proceedings for the 17th Symposium on the Interface, (1986), 11-17.

[12] P. Doukhan, G. Oppenheim, and M. Taqqu, *Theory and Application of Long-Range Dependence*, Birkhãuser, Boston, 2003.

[13] R. Fox and M. Taqqu, *Large-sample properties of parameter estimates for strongly dependent stationary Gaussian Time Series*, The Annals of Statistics, 14 (1986)

[14] R. Dahlhaus, *Efficient parameter estimation for self-similar processes*, The Annals of Statistics, 17 (1989), 1749-1766.

[15] A. McLeod, and K. Hipel, *Preservation of the Rescaled Adjusted Range 1. A Reassessment of the Hurst Phenomenon*, Water Resources Research, 14 (1978), 491-508.

[16] L.A. Thombs, and W.R. Schucany, *Bootstrap Prediction Intervals for Autoregression*,Journal of the American Statistical Association, 85 (1990),486-492.

[17] L. Kilian, *Confidence Intervals for impulse responses under departures from normality*,Econometric Reviews, 17 (1998a), 1-29.

[18] H.R. Künsch, *The jackknife and the bootstrap for general stationary observations*, The Annals of Statistics, 17 (1989), 1217-1241.

[19] U. Grenander, *Abstract Inference*, Wiley, New York, 1981.

[20] S. Geman, and C.R. Hwang, *Nonparametric maximum likelihood estimation by the method of sieves*, The Annals of Statistics, 10 (1982), 401-414.

[21] G. Masarotto, *International Journal of Forecasting, 6, 229-239.*, International Journal of Forecasting, 6 (1990), 229-239.

[22] M. Grigoletto, *Bootstrap Prediction Intervals for Autoregression: Some Alternatives*,International Journal of Forecasting, 14 (1998), 447-456.

[23] R. Cao, M. Feberobande, W. Gonzles-Manteiga, J.M. Prada-Snchez, I. Garcia-Juardo, *Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes*,Communications in Statistics: Simulation and Computation, 26 (1997), 961-978.

[24] C. W. J. Granger, and R. Joyeux, *An introduction to long-range time series models and fractionally differencing*, Journal of Time Series Analysis, 1 (1980), 15-30.

[25] J. R. M. Hosking, *Fractional differencing*, Biometrika, 68 (1981), 165-176.

[26] J. Liu, Y. Shu, L. Zhang, F. Xue, and O.W.W. Yang, *Traffic Modeling Based on FARIMA Models*, IEEE Canadian Conference on Electrical and Computer Engineering, Alberta Canada, 1999.

[27] P. Bühlmann, *Sieve bootstrap for time series*. Bernoulli, 3 (1997), 123-148.

[28] B. K. Ray, *Modelling lon-memory processes for optimal long-range prediction*, Journal of Time Series Analysis, 9 (1993a), 215-220.

[29] B. K. Ray, *Long-range forecasting of IBM product revenue using a seasonal fractionally differenced ARMA model*, International Journal of Forecasting, 9 (1993b), 255-296.

[30] J. Brodsky, and C. Hurvich, *Multi-step forecasting for long-memory processes*, Journal of Forecasting, 18 (1999), 59-75.

[31] R. Eisinga, P. Franses, and M. Ooms, *Forecasting long-memory left-right political orientation*, International Journal of Forecasting, 15 (1999), 185-199.

[32] J. Geweke, and S. Porter-Hudak, *The estimation and application of long-memory time series models* , Journal of Time Series Analysis, 4 (1983), 221-237.

[33] D. S. Poskitt, *Autoregressive approximation in nonstandard situations: the fractionally integrated and non-invertible cases*, Annals of the Institute of Statistical Mathematics, 59 (2006), 697-725.

[34] D. S. Poskitt, *Properties of the Sieve bootstrap for fractionally integrated and non-invertible processes*, Journal of Time Series Analysis, 29 (2007), 224-250.

[35] M. Rupasinghe, and V.A. Samaranayake, *Asymptomatic Properties of Sieve Bootstrap Prediction Intervals for FARIMA Processes*, Statistics and Probability Letters, (2012), submitted.

[36] O. Tousson, *Mèmoire sur l'hitoire du Nile*, Mèmoires de l'Institut d'Egypte , 9 (1925).

III. PREDICTION INTERVALS FOR *ARIMA* PROCESSES: A SIEVE BOOTSTRAP APPROACH

ABSTRACT

The sieve bootstrap is a model-free re-sampling method that approximates an invertible linear process with a finite autoregressive model whose order increases with sample size. Prediction intervals based on this approach have been successfully implemented for stationary invertible ARMA processes. The coverage probabilities of sieve bootstrap intervals developed for $ARMA$ models, however, are well below the nominal level in the presence of a unit root in the autoregressive polynomial. An approach that overcomes this drawback is proposed and the asymptotic properties of the proposed method are derived. Monte Carlo simulation results indicate that the proposed method provides near nominal coverage at moderate sample sizes.

Keywords: Unit root processes; Forecast intervals; ARMA; Nonstationarity

1. INTRODUCTION

Many financial and economic time series are non-stationary, and *Autoregressive Integrated Moving Average* (ARIMA) processes are often used to model such empirical process. For the practitioner, one of the main goals of empirical time series modeling is to obtain forecasts based on its past values. Standard parametric point and interval forecasts are quite accurate under normally distributed innovations. As noted by Stine (1987) and Thombs and Schucany (1990), parametric prediction intervals perform poorly when the normal assumption is violated. Nonparametric bootstrap based prediction intervals, therefore, have been used as an alternative to parametric estimates by time series analysts. While nonparametric approaches have been proposed for stationary processes, a method that provides prediction intervals for the class of $ARIMA$ models with unknown orders p, q is not available. In the following sections, a nonparmetric bootstrap approach to obtain prediction intervals for $ARIMA$ processes with unknown orders is presented.

One drawback of the original bootstrap methods is the requirement of the knowledge of the orders associated with the underlying process. For instance, the bootstrap approach proposed by Stine (1987) assumes the order, p, of the $AR(p)$ process is known. The same is true for methods introduced by Thombs and Schucany (1990), Cao *et al.* (1997) and Pascual *et al.* (2004).

The method proposed in this paper, however, does not require any knowledge of the orders associated with autoregressive and moving average polynomials. Our framework is identical to the Sieve Bootstrap prediction intervals implemented by Alonso, Pena and Romo (2002, 2003 and 2004), which resamples residuals obtained by sequence of $AR(p)$ models with order $p = p(n)$, which increases with the sample size n. The foundation of this sieve bootstrap approach was laid by Kreiss (1988) and (1992), for time series that can be represented by an infinite autoregressive process. Bühlmann (1997) extended this approach to more general class of time series that can be written as an infinite order moving average time series and introduced the term *sieve bootstrap*. Alonso et *al.* (2002, 2003) formalized this sieve bootstrap concept and applied it to obtain prediction intervals

for linear processes. The same authors made further refinements in 2004 by introducing model uncertainty in computing prediction intervals. Alonso's method was modified by Mukhopadhyay and Samaranayake (2010) to improve the coverages of the prediction intervals. They achieved this by introducing a variance inflation factor for bootstrap residuals. These preceding bootstrap methods are, however, limited to stationary linear processes such as $ARMA$ models. Rupasinghe and Samaranayake (2012) extended Alonso's 2003 sieve bootstrap procedure to compute prediction intervals for *long memory* processes ($FARIMA$). In this paper, we extend Alonso's 2003 sieve bootstrap procedure to obtain prediction intervals for $ARIMA$ processes.

1.1. *ARIMA* PROCESSES

A real-valued process $\{x_t\}_{t\in\mathbb{Z}}$ is said to be a *Autoregressive Integrated Moving Average* ($ARIMA(p,d,q)$) process if it is stationary and satisfies

$$\alpha(B) \bigtriangledown^d (x_t - \mu) = \theta(B)\epsilon_t, t \in \mathbb{Z}, \tag{1.1}$$

where $\alpha(z) = 1 - \alpha_1 z - ... - \alpha_p z^p$ and $\theta(z) = 1 + \theta_1 z + ... + \theta(q) z^q$ represent autoregressive and moving average polynomials of degrees p and q respectively. The mean of the process is $\mu = E[x_t]$ for all t. It is assumed that $\alpha(.)$ and $\theta(.)$ do not share common zeros. The error terms, $\{\epsilon_t\}$, are assumed to be zero-mean white noise with finite variance σ^2. Note that $\bigtriangledown = 1 - B$, where B is the back-shift operator defined by $B^k x_t = x_{t-k}$ for $k \in \mathbb{N}$. The difference parameter, d, can take any non-negative integer, but we assume that $d = 1$ or 0 which represents the most common type of $ARIMA$ processes used in empirical modeling.

The literature on methods for obtaining prediction intervals for $ARIMA$ processes is very limited. Kim (2001) extended the forward and backward bootstrap procedure of Thombs and Schucany (1990) to obtain prediction intervals for $AR(p)$ models with unit roots by incorporating a bias correction on the bootstrap estimates of the forward and backward AR coefficients. The backward AR representation is obtained by reversing the forward (usual) $AR(p)$ model. This bias correction was adopted from Kilian (1998a) and

utilized to improve the coverage probabilities in the presence of unit roots. Their method, however, assumes that the process is $AR(p)$ and the order, p, is known, which could be a weakness in situations where the order is unknown. They also assumed normal errors in establishing the asymptotic validity of the method.

In their recent articles, Panichkitkosolkul and Niwitpong (2011, 2012) introduced parametric prediction intervals for Gaussian $AR(p)$ models that may include unit root processes. The prediction intervals are computed following preliminary unit root tests and two different formulations for prediction intervals were used based on the outcome of the initial tests. They used well known Dickey-Fuller (DF) (Dickey and Fuller (1979)), Augmented Dickey-Fuller (ADF) (Said and Dickey (1984)), and SSL (Shin, Sarkar and Lee (1996)) unit root tests. The random walk model is used to obtain point forecasts in case the preliminary test did not reject the null hypothesis that the process has an autoregressive root equal to unity. There are concerns on the use of unit root tests prior to compute prediction intervals, as the power of these tests is small under many situations. See Psaradakis (2001), Chang and Park (2003) and Palm, Smeekes and Urbain (2008).

Our method, however, do not alter the procedure of computing prediction intervals based on results of a unit root test. If the observed series $\{x_t\}$ satisfies $\alpha(B)(1-B)x_t = \theta(B)\epsilon_t$, $\{\epsilon_t\} \sim WN(0, \sigma^2)$, observe that the differenced series $y_t = x_t - x_{t-1}$ is stationary. One can first compute the bootstrap distribution of the future observations, y_{n+h}, of the differenced series and then use it to obtain that of x_{n+h}. This implementation is simple if the underlying process of the original observations is $ARIMA(p, 1, q)$ because $\{y_t\}$ is then both stationary and invertible, but poses a problem if the underlying process is $ARMA(p, q)$. In the latter case, the differenced series is non-invertible since $\alpha(B)y_t = (1-B)\theta(B)\epsilon_t$. To be able to invert the time series, Alonso *et al.* (2003) and Bühlman (1997) required that the moving average polynomial has no roots on or inside the unit circle. This was a key assumption in their sieve bootstrap procedure and was needed in order to approximate the time series by a sequence of AR polynomials.

Poskitt (2006, 2007) discussed ways of relaxing this condition while maintaining the statistical viability of finite order autoregressive approximations to non-invertible processes. Poskitt (2006, 2007) did not show how to obtain an asymptotically valid

estimator of the distribution of a future observation from a non-invertible process, but his results provide a theoretical foundation on which such an estimator can be derived. In the following sections we show how the method proposed by Alonso et *al.* (2003) can be modified, based on insights from Poskitt (2006, 2007), to obtain sieve bootstrap prediction intervals for a non-invertible process. While Alonso et *al.* (2004) and Mukhopadhyay and Samaranayake (2010) provide additional refinements to the original method proposed in Alonso et *al.* (2002, 2003), the 2003 paper by Alonso et *al.* set the fundamental theoretical framework for the application of the sieve bootstrap for invertible processes. As such, we use it as the platform for our proposed modifications even though Alonso et *al.* (2004) and Mukhopadhyay and Samaranayake (2010) give further refinements to the original method. As the Monte Carlo simulation results in Section 4 show, the proposed method provide good finite sample coverage even without additional refinements adopted in the above two papers. Thus, the proposed method can be taken good initial step in adopting the sieve bootstrap to obtain prediction intervals for $ARIMA$ processes.

The rest of this paper is organized as follows. Section 2 introduces the sieve bootstrap procedure for obtaining prediction intervals and Section 3 establishes asymptotic validity of the proposed method. The simulation study along with an application is presented in Sections 4 and 5.

2. THE PROPOSED SIEVE BOOTSTRAP PROCEDURE

The main difference between the sieve bootstrap procedure given below and the procedure introduced by Alonso et *al.* (2002, 2003) is the criterion used in selecting the order of the autoregressive approximation. This change in the order, together with Poskitt's AR approximation to non-invertible processes, are sufficient to establish the convergence results. Also, note that we introduce a differencing step at the beginning of the procedure in order to accommodate $ARIMA$ processes.

Assume that a realization $\{x_t\}_{t=1}^n$ is obtained from $ARIMA(p, d, q)$ process given in Equation (1.1) with $d = 1$ or 0. Define the differenced series, $\{y_t\}$, using $y_t = x_t - x_{t-1}$.

1. Select the order $p = p(n)$ of the autoregressive approximation from among models with $p \in \{1, 2, ..., M_n\}$ with $M_n = o\{[log(n)/n]^{1/2}\}$ by the AIC criterion. Alonso et *al.* (2003) preferred AICC over AIC and used $M_n = o\{[log(n)/n]^{1/4}\}$.

2. Estimate the autoregressive coefficients, $\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n}$, of the $AR(p)$ approximation, $\sum_{j=0}^{p} \phi_{j,p} y_{t-j} = \epsilon_{t,p}$, by the Yule-Walker method.

3. Obtain the $(n-p)$ residuals: $\hat{\epsilon}_{t,n} = \sum_{j=0}^{p} \hat{\phi}_{j,p,n}(y_{t-j} - \bar{y})$, $t = p+1, ..., n$ and define the empirical distribution function of the centered residuals, $\tilde{\epsilon}_t = \hat{\epsilon}_{t,n} - \hat{\epsilon}^{(.)}$, where $\hat{\epsilon}^{(.)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n}$, by $\hat{F}_{\tilde{\epsilon},n}(x) = (n-p)^{-1} \sum_{t=p+1}^{n} I_{[\tilde{\epsilon}_t \leq x]}$.

4. Draw a resample $\epsilon^*_{t,n}$, $t = p+1, ..., n$ of i.i.d. observations from $\hat{F}_{\tilde{\epsilon},n}$.

5. Obtain y^*_t by the recursion: $\sum_{j=0}^{p} \hat{\phi}_{j,p,n}(y^*_{t-j} - \bar{y}) = \epsilon^*_{t,n}$ for $t = p+1, ..., n$ and set $y^*_t = \bar{y}$ for $t = 1, ..., p$.

6. Compute the estimates $\hat{\phi}^*_{1,p,n}, ..., \hat{\phi}^*_{p,p,n}$ as in Step 2, using $\{y^*_t\}_{t=1}^{n}$.

7. For $h > 0$, compute the future bootstrap observations of the differenced series by the recursion: $y^*_{n+h} - \bar{y} = \sum_{j=1}^{p} \hat{\phi}^*_{j,p,n}(y^*_{n+h-j} - \bar{y}) + \epsilon^*_{n+h,n}$ where, $y^*_t = y_t, t \leq n$.

   Up to this point we have followed Alonso *et al.* (2003) sieve bootstrap procedure but the next step is crucial to obtaining bootstrap future observation of the original time series $\{x_t\}$.

8. Compute the future bootstrap observations of the original series by the recursion: $x^*_{n+h} = x^*_{n+h-1} + y^*_{n+h}$ where, $x^*_t = x_t, t \leq n$, $h > 0$.

9. Obtain a Monte Carlo estimate of the bootstrapped distribution function of x^*_{n+h} by repeating steps 4-8 B times.

10. Use the bootstrapped distribution to approximate the unknown distribution of x_{n+h} given the observed sample.

11. The $100(1-\alpha)\%$ prediction interval for x_{n+h} is given by $\{Q^*(\frac{\alpha}{2}), Q^*(1-\frac{\alpha}{2})\}$ where, $Q^*(.)$ are the quantiles of the estimated bootstrap distribution.

3. ASYMPTOTIC RESULTS

Note that if the original process $\{x_t\}$ is indeed an $ARIMA(p,1,q)$ process, then the differenced process $\{y_t\}$ is $ARMA(p,q)$ and the results of Alonso et *al.* (2003) applies directly to the bootstrap distribution of y^*_{n+h}. It then follows by simple argument that the bootstrap distribution of x^*_{n+h} converges to that of x_{n+h}. On the other hand complications arise if $\{x_t\}$ has no unit root. Then $\{y_t\}$ would not be invertible and hence the results of Alonso et *al.* (2003) do not apply. This is where the new order for M_n (Step 1) and results of Poskitt (2006, 2007) come into play. This approach avoids the need to pre-test for unit roots and then select the prediction interval procedure based on the outcome of the test.

In order to establish the asymptotic validity of the sieve bootstrap intervals, Alonso (2003) first established the convergence of $\hat{\boldsymbol{\phi}}^*_{p,n}$ to $\hat{\boldsymbol{\phi}}_{p,n}$. We follow the same approach, but modify the proofs to accommodate the changes arising out of the possibility that the differenced series is non-invertible. We first establish asymptotic properties of the differenced series $\{y_t\}$ and then move onto proving results for $\{x_t\}$.

Rupasinghe and Samaranayake (2012) extended some of the results in Bühlmann (1995, 1997) and Alonso et *al.* (2003) to *regular processes*, a general class of linear processes that includes both *FARIMA* and non-invertible time series. As stated in Poskitt (2006), the process $\{y_t\}_{t\in\mathbb{Z}}$ is said to be linearly regular if $\{y_t\}_{t\in\mathbb{Z}}$ is covariance stationary with,

$$y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \tag{3.1}$$

where $\{\epsilon_t\}_{t\in\mathbb{Z}}$, is a zero mean white noise process with finite variance σ^2 and the impulse response coefficients $\{\psi_j\}_{j=0}^{\infty}$ satisfy the condition $\psi_0 = 1$ and $\sum_{j\geq 0}\psi_j^2 < \infty$.

In the following derivations, we will use the $AR(p)$ approximation suggested by Poskitt (2006) for such time series.

Definition 3.1. *Let* $\{y_t\}$ *satisfy equation (3.1) and define, for* $p < n$, $\{\epsilon_{t,p}\}$ *such that* $\sum_{j=0}^{p} \phi_{j,p} y_{t-j} = \epsilon_{t,p}$, *where the AR coefficients vector* $\boldsymbol{\phi}_p = (\phi_{1,p}, ..., \phi_{p,p})'$ *is obtained using the Yule-Walker equations,* $\boldsymbol{\Gamma}_p \boldsymbol{\phi}_p = -\boldsymbol{\gamma}_p$, $\boldsymbol{\gamma}_p = (\gamma(1), ..., \gamma(p))'$, $\boldsymbol{\Gamma}_p = [\gamma(i-j)]_{i,j=1}^{p}$, *with* $\gamma(k) = E[y_t y_{t+k}]$ *for* $k \in \mathbb{N}_0$.

This definition provides us the AR approximation to $\{y_t\}$ even in the case where the series is not invertible and thus cannot be written as an infinite autoregressive process. It should be noted that Lemma 1 of Poskitt (2006) establishes that $\epsilon_{t,p} \to \epsilon_t$ in mean square as $p \to \infty$.

The following sets of assumptions are required in order to prove our asymptotic results.

A1: Let ξ_t denote the σ-algebra of events determined by ϵ_s, $s \leq t$. Also, assume ϵ_t is $i.i.d.$ and that

$$E[\epsilon_t | \xi_{t-1}] = 0 \qquad and \qquad E[\epsilon_t^2 | \xi_{t-1}] = \sigma^2, t \in \mathbb{Z}.$$

Furthermore, assume $E[\epsilon_t^4] < \infty$ for $t \in \mathbb{Z}$.

A2: The series y_t is a linearly regular covariance-stationary process with Wold representation $y_t = \sum_{j \geq 0} \psi_j \epsilon_{t-j}$ with $\sum_{j \geq 0} |\psi_j^2| < \infty$.

B: Let $p(n) = o\{[n/log(n)]^{1/2}\}$ and $\hat{\boldsymbol{\phi}}_{p,n} = (\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n})'$ satisfy the empirical Yule-Walker equations $\hat{\boldsymbol{\Gamma}}_{p,n} \hat{\boldsymbol{\phi}}_{p,n} = -\hat{\boldsymbol{\gamma}}_{p,n}$, where
$\hat{\boldsymbol{\Gamma}}_{p,n} = [\hat{R}(i-j)]_{i,j=1}^{p}, \hat{\boldsymbol{\gamma}}_{p,n} = (\hat{R}(1), ..., \hat{R}(p))'$, and $\hat{R}(j) = n^{-1} \sum_{t=1}^{n-|j|} (y_t - \bar{y})(y_{t+|j|} - \bar{y})$ for $|j| < n$.

Assumptions in A1 imposes a Martingale difference structure on the innovations. Since the sieve bootstrap scheme draws resamples independently and identically, it is unable to capture the correlation structure of the innovations if they are correlated. Therefore, in Proposition 1 we assume $i.i.d.$ innovations for the underlying processes.

The order of p in Assumption B is slightly different from that of Rupasinghe and Samaranayake (2012). They assumed that $p(n) = o\{[n/log(n)]^{1/2-d}\}$, where d is the difference parameter taking fractional values from -0.5 to 0.5. The value for d is set to

zero throughout since we are only interested in $ARIMA(p, d, q)$ models with $d = 0$ or 1 and differencing removes the unit root, if present.

Next we present asymptotic properties of the sieve bootstrap method given in Section 2 by adopting some results form Rupasinghe and Samaranayake (2012).

The following results follows from the same arguments use in Lemmas 1, 2 and 3 of Rupasinghe and Samaranayake (2012). Therefore, they are stated without proof.

Lemma 3.2. *Assume that A1, A2 and B hold. Then,*

$$\sum_{j=0}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})^2 = o_{a.s.}\{[log(n)/n]^{1/2}\},$$

where $\phi_{j,p}$, $j = 1, 2, ..., p$, $p < n$ are the coefficients given in Definition 3.1.

Lemma 3.3. *Assume that A1, A2 and B hold. Then, for any fixed $t \in \mathbb{Z}$,*

$$E^*(\epsilon_{t,n}^{*2}) = E(\epsilon_t^2) + o_p(1).$$

The next Lemma states asymptotic convergence of bootstrap innovations to theoretical innovations, and is similar to Lemma 5.4 in Buhlmann (1997).

Lemma 3.4. *Assume that assumptions given in A1, A2 and B hold. Then, for each fixed $t \in \mathbb{N}$,*

$$\epsilon_{t,n}^* \xrightarrow{d^*} \epsilon_t, \textit{ in probability.}$$

The following proposition is analogous to Proposition 1 of Alonso et *al.* (2003) and shows that the bootstrap autoregressive coefficients obtained in Step 6 converge to the autoregressive coefficients of the fitted model obtained in Step 2.

Proposition 3.5. *Assume A1, A2 and B hold. Then,*

$$\max_{1\leq j\leq p(n)} |\hat{\phi}_{j,p,n}^* - \hat{\phi}_{j,p,n}| \xrightarrow{P^*} 0, \textit{ in probability.} \tag{3.2}$$

The $AR(p)$ approximation described in Definition 3.1 is used to establish the convergence of future bootstrap values of the differenced series. It is, therefore, essential to show the convergence of bootstrap innovations, $\epsilon^*_{t,n}$, to the approximated errors, $\epsilon_{t,p}$. This is a strategic feature proposed to overcome issues raised in generalizing Alonso *et al.* (2003) results for regular processes.

Lemma 3.6. *Assume that assumptions given in A1, A2 and B hold. Then, for each fixed* $t \in \mathbb{N}$,

$$\epsilon^*_{t,n} \xrightarrow{d^*} \epsilon_{t,p}, \textit{ in probability.}$$

Proof. Let $F_{\epsilon,n}(x) = (n-p)^{-1}\sum_{t=p+1}^{n} 1_{[\epsilon_{t,p}\leq x]}$, $F_{\epsilon,p}(x) = \mathbb{P}[\epsilon_{t,p} \leq x]$ for $x \in \mathbb{R}$, and denote the Mallows metric by $d_2(.,.)$. Then, from standard results it follows that $d_2(F_{\epsilon,n}, F_{\epsilon,p}) = o_{a.s.}(1)$. Thus we need to only show that $d_2(\hat{F}_{\tilde{\epsilon},n}, F_{\epsilon,n}) = o_p(1)$. Let S be uniformly distributed on $\{p+1, ..., n\}$ and let $Z_1 = \epsilon_S$, $Z_2 = \bar{\epsilon}_S$, where $\bar{\epsilon}_{t,n} = \hat{\epsilon}_{t,n} - \hat{\epsilon}^{(.)}_n$. Then, $d_2(\hat{F}_{\tilde{\epsilon},n}, F_{\epsilon,n})^2 \leq E|Z_1 - Z_2|^2 = (n-p)^{-1}\sum_{t=p+1}^{n}(\bar{\epsilon}_{t,n} - \epsilon_{t,p})^2 = (n-p)^{-1}\sum_{t=p+1}^{n}(\hat{\epsilon}_{t,n} - \hat{\epsilon}^{(.)}_n - \epsilon_{t,p})^2$. From the proof of Lemma 2 in Rupasinghe and Samaranayake (2012), $\hat{\epsilon}^{(.)}_n = o_p(1)$ and $(n-p)^{-1}\sum_{t=p+1}^{n}|\hat{\epsilon}_{t,n} - \epsilon_{t,p}| = o_p(1)$. Hence $d_2(\hat{F}_{\epsilon,n}, F_{\epsilon,n}) = o_p(1)$. □

Now we establish the convergence of the bootstrap differenced series.

Theorem 3.7. *Assume that A1, A2 and B hold. Then, in probability, as* $n \to \infty$,

$$y^*_{n+h} \xrightarrow{d^*} y_{n+h}, \textit{ for fixed } h \in \mathbb{N} \tag{3.3}$$

Proof.

$$\text{Observe that, } y_{n+h} = -\sum_{j=1}^{p}\phi_{j,p}y_{n+h-j} + \epsilon_{n+h,p} \tag{3.4}$$

$$\text{and } y^*_{n+h} = -\sum_{j=1}^{p}\hat{\phi}_{j,p,n}y^*_{n+h-j} + \epsilon^*_{n+h,n}, \tag{3.5}$$

where $y^*_t = y_t$ for $t \leq n$. For beverity, we prove the theorem for $h = 1$.

From Lemma 3.6, $\epsilon^*_{n+1,n} \xrightarrow{d^*} \epsilon_{n+1,p}$ and thus we need only to show that the difference of the first terms on the right hand side of (3.4) and (3.5) converges to zero in probability. Therefore consider,

$$-\sum_{j=1}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})y_{n+1-j} \leq \left(\sum_{j=1}^{p}(\hat{\phi}_{j,p,n} - \phi_{j,p})^2\right)^{1/2}\left(\sum_{j=1}^{p(n)} y^2_{n+1-j}\right)^{1/2}$$
$$= \{o_{a.s.}[[log(n)/n]^{\frac{1/2}{2}}]\}\{O_p[p^{1/2}]\} = o_p(1).$$

Thus, $y^*_{n+1} \xrightarrow{d^*} y_{n+1}$, in probability. □

Finally, we establish the large sample validity of sieve bootstrap prediction intervals for $ARIMA(p, d, q)$ processes with $d = 0$ or 1 by proving the convergence of the future bootstrap values of the original time series, obtained in Step 8.

Theorem 3.8. *Assume that A1, A2 and B hold. Then, in probability, as* $n \to \infty$,

$$x^*_{n+h} \xrightarrow{d^*} x_{n+h}, \text{ for } h = 0, 1, ... \tag{3.6}$$

Proof. The future values of the originally observed time series, $\{x_{n+h}\}$ can be written as $x_{n+h} = x_{n+h-1} + y_{n+h}$. Then the bootstrap one-step ahead value exhibit the following property:
$x^*_{n+1} = x_n + y^*_{n+1} \xrightarrow{d^*} x_n + y_{n+1} = x_{n+1}$. For $h > 1$, the result can be proven using the mathematical induction.

□

4. SIMULATION STUDY

In order to investigate the finite sample performances of the method proposed in this paper, a Monte-Carlo simulation study, using a series of models given in Table 1, was carried out with three different error distributions and sample sizes 100 and 200. The coverage, bootstrap length, and the length of the interval theoretically achievable under

known order and parameter values were computed for 95% and 99% prediction intervals to asses the performance of the proposed method. Results are reported in Tables 2 through 7.

Table 1. Models considered in the simulation study

Nomenclature	Model	AR roots	MA roots
$M1$	$(1-0.75B+0.5B^2)X_t=\epsilon_t$	1.414, 1.414	-
$IM1$	$(1-0.75B+0.5B^2)(1-B)X_t=\epsilon_t$	1, 1.414, 1.414	-
$M2$	$X_t=(1-0.9B)\epsilon_t$	-	$1.\bar{1}$
$IM2$	$(1-B)X_t=(1-0.9B)\epsilon_t$	1	$1.\bar{1}$
$M3$	$X_t=(1-0.3B+0.7B^2)\epsilon_t$	-	1.195, 1.195
$IM3$	$X_t=(1-0.3B+0.7B^2)\epsilon_t$	1	1.195, 1.195
$M4$	$(1-0.7B)X_t=(1-0.3B)\epsilon_t$	1.428	$3.\bar{3}$
$IM4$	$(1-0.7B)(1-B)X_t=(1-0.3B)\epsilon_t$	1, 1.428	$3.\bar{3}$
$M5$	$(1-0.95B)X_t=(1-0.3B)\epsilon_t$	1.05	$3.\bar{3}$
$IM5$	$(1-0.7B)(1-B)X_t=(1-0.3B)\epsilon_t$	1, 1.05	$3.\bar{3}$

Note that the models employed in the study are the same $ARMA$ models studied by Mukhopadhyay and Samaranakaye (2010) and Alonso *et al.* (2004). We also considered corresponding $ARIMA$ models (begin with I) since we are interested in unit root processes. The standard normal distribution, t-distribution with 3 degrees of freedom, and exponential (1) distribution centered at zero, were considered for error distributions. Prediction intervals for leads $h=1,2,3$ were computed. The Matlab (Version 2011a) software was used for these simulations.

For each combination of model, sample size, nominal coverage and error distribution, $N=1,000$ independent series were generated and for each of these simulated series, steps 1 to 15 were implemented. To compute the coverage probabilities for each of this N

simulations, $R = 1,000$ future observations (x_{n+h}) were generated using the original model.

The proportion of those falling in between the lower and upper bounds of the bootstrap prediction interval was then defined to be the coverage. Thus, the coverage at the i^{th} simulation run is given by $C(i) = R^{-1}\sum_{r=1}^{R} I_A[x_{n+h}^r(i)]$ where $A = [Q^*(\alpha/2), Q^*(1-\alpha/2)]$, $I_A(.)$ is the indicator function of the set A and $x_{n+h}^r(i)$, $r = 1, 2, ...1,000$ are the R future values generated at the i^{th} simulation run. The bootstrap length and theoretical length for the i^{th} simulation run are given by $L_B(i) = Q^*(1-\alpha/2) - Q^*(\alpha/2)$ and $L_T(i) = x_{n+h}^r(1-\alpha/2) - x_{n+h}^r(\alpha/2)$ respectively. The theoretical length $L_T(i)$ is the difference between the $100(1-\alpha/2)^{th}$ and $100(\alpha/2)^{th}$ percentile points the empirical distribution of the 1,000 future observations that were generated using the underlying time series model with known order and the true values of the coefficients. Using these statistics, the mean coverage, mean length of bootstrap prediction intervals, mean length of theoretical intervals, and their standard errors were computed as:

Mean Coverage $\bar{C} = N^{-1}\sum_{i=1}^{N} C(i)$

Standard Error of Mean Coverage $SE_{\bar{C}} = \{[N(N-1)]^{-1}\sum_{i=1}^{N}[C(i)-\bar{C}]^2\}^{1/2}$

Mean Length (bootstrap) $\bar{L}_B = N^{-1}\sum_{i=1}^{N} L_B(i)$

Standard Error of Mean Length $SE_{\bar{L}_B} = \{[N(N-1)]^{-1}\sum_{i=1}^{N}[L_B(i)-\bar{L}_B]^2\}^{1/2}$

Mean theoretical Length $\bar{L}_B = N^{-1}\sum_{i=1}^{N} L_B(i)$

In total 120 different combinations of model type, sample size, nominal coverage probability, and error distributions were investigated in this simulation study. However, due to space limitations, we report only a representative sample of results for 95% intervals, in Table 2 through 7. These tables report the mean coverage, mean interval length, and mean theoretical length, standard error of mean coverage and standard error of mean interval length. The complete results of the simulation study are available upon request from the corresponding author.

To investigate the behaviour of the intervals for each of the 120 combinations, the minimum value, percentiles (25^{th}, 50^{th}, and 75^{th}), and the maximum value of (a) the coverage probabilities, (b) the bootstrap interval bounds (upper and lower), and (c) the

theoretical interval bounds (upper and lower), were further computed, based on the 1,000 values generated through simulation, and these statistics are also available upon request.

From Tables 2-7, we can see that our method provides coverages closer to the nominal level as sample size increases for both $ARMA$ and $ARIMA$ models. This is expected since large sample sizes provide more accurate parameter estimates as well as yield more residuals for resampling. Furthermore, the mean coverages of the proposed sieve bootstrap method are very close to the nominal coverage for all the leads regardless of presence or absence of a unit root and of the nature their error distribution. Also, it is seen that the mean bootstrap interval lengths are much closer to the theoretical lengths.

It is interesting how the proposed sieve bootstrap procedure performs for models $M5$ and $IM5$ in which the AR root is close to unity. In practice, many parametric and nonparametric prediction intervals produce very liberal coverages when the AR polynomial has a root close to unity (see Alonso *et al.* (2002, 2004)). However, from Tables 4, 6 and 7, we can see that our proposed method is capable of producing accurate prediction intervals for time series with an AR root close to one.

Table 2. Coverage of 95% intervals for Models $M1$ & $IM1$ with normal errors

Leads	Size	Model $M1$			Model $IM1$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9040	0.9548 (0.0026)	4.2994 (0.0444)	3.9339	0.9561 (0.0029)	4.4168 (0.0492)
	200	3.9178	0.9503 (0.0024)	4.2013 (0.0298)	3.9153	0.9598 (0.0019)	4.3081 (0.0403)
2	100	6.7753	0.9500 (0.0031)	7.5408 (0.0901)	10.3208	0.9464 (0.0034)	11.1515 (0.1263)
	200	6.7719	0.9484 (0.0031)	7.4664 (0.0619)	10.2422	0.9539 (0.0023)	10.9454 (0.1049)
3	100	8.9874	0.9456 (0.0041)	10.1540 (0.1308)	18.5608	0.9365 (0.0040)	19.5478 (0.2382)
	200	8.9473	0.9476 (0.0036)	10.1857 (0.0961)	18.3911	0.9491 (0.0027)	19.3903 (0.1969)

Table 3. Coverage of 95% intervals for Models $M4$ & $IM4$ with normal errors

Leads	Size	Model $M4$			Model $IM4$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.8980	0.9463 (0.0031)	4.2205 (0.0275)	3.9305	0.9545 (0.0020)	4.1908 (0.0380)
	200	3.9102	0.9438 (0.0037)	4.2945 (0.0462)	3.9267	0.9579 (0.0012)	4.1649 (0.0218)
2	100	4.2136	0.9425 (0.0045)	4.6845 (0.0329)	6.7384	0.9528 (0.0023)	7.2182 (0.0729)
	200	4.2094	0.9512 (0.0043)	4.9479 (0.0550)	6.7303	0.9573 (0.0015)	7.1610 (0.0472)
3	100	4.3746	0.9400 (0.0063)	5.0388 (0.0374)	9.4048	0.9492 (0.0028)	9.9920 (0.1182)
	200	4.3675	0.9523 (0.0044)	5.3092 (0.0704)	9.4416	0.9568 (0.0018)	10.0501 (0.0803)

Table 4. Coverage of 95% intervals for Models $M5$ & $IM5$ with normal errors

Leads	Size	Model $M5$			Model $IM5$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9134	0.9396 (0.0029)	3.9704 (0.0369)	3.9281	0.9439 (0.0026)	4.0540 (0.0368)
	200	3.9051	0.9466 (0.0021)	4.0106 (0.0313)	3.9068	0.9481 (0.0020)	3.9901 (0.0293)
2	100	4.6468	0.9431 (0.0029)	4.8376 (0.0469)	7.5649	0.9485 (0.0024)	8.1119 (0.0854)
	200	4.6862	0.9465 (0.0021)	4.8470 (0.0378)	7.5790	0.9525 (0.0019)	7.9298 (0.0573)
3	100	5.2358	0.9468 (0.0032)	5.6353 (0.0595)	11.6686	0.9480 (0.0031)	12.7837 (0.1745)
	200	5.2679	0.9488 (0.0025)	5.5933 (0.0537)	11.6951	0.9523 (0.0022)	12.3978 (0.1048)

Table 5. Coverage of 95% intervals for Models $M4$ & $IM4$ with exponential errors

Leads	Size	Model $M4$			Model $IM4$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.6257	0.9522 (0.0081)	4.2604 (0.0626)	3.6652	0.9583 (0.0045)	3.8065 (0.0502)
	200	3.6535	0.9610 (0.0055)	4.2893 (0.0743)	3.6562	0.9694 (0.0020)	4.1766 (0.0766)
2	100	4.0041	0.9499 (0.0075)	4.7753 (0.0665)	6.4765	0.9520 (0.0032)	6.6051 (0.0841)
	200	4.0081	0.9617 (0.0038)	4.8360 (0.0813)	6.4841	0.9662 (0.0025)	7.1918 (0.1256)
3	100	4.2433	0.9498 (0.0080)	5.1285 (0.0688)	9.1766	0.9468 (0.0031)	9.2601 (0.1174)
	200	4.2318	0.9553 (0.0059)	5.1947 (0.0974)	9.1801	0.9592 (0.0032)	10.0295 (0.1871)

Table 6. Coverage of 95% intervals for Models $M5$ & $IM5$ with exponential errors

Leads	Size	Model $M5$			Model $IM5$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.6596	0.9637 (0.0041)	4.0219 (0.0525)	3.6823	0.9520 (0.0048)	4.0145 (0.0725)
	200	3.6978	0.9613 (0.0036)	4.0387 (0.0723)	3.6831	0.9570 (0.0043)	3.8241 (0.0558)
2	100	4.4705	0.9481 (0.0051)	4.8064 (0.0528)	7.2723	0.9479 (0.0044)	8.0177 (0.1331)
	200	4.4917	0.9493 (0.0041)	4.8302 (0.0790)	7.2880	0.9510 (0.0034)	7.4932 (0.0941)
3	100	5.1322	0.9463 (0.0052)	5.5740 (0.0641)	11.3299	0.9462 (0.0046)	12.6384 (0.2140)
	200	5.1352	0.9503 (0.0044)	5.6875 (0.0920)	11.3583	0.9497 (0.0029)	11.7789 (0.1525)

Table 7. Coverage of 95% intervals for $M5$ & $IM5$ with t-dist errors

Leads	Size	$M5$			$IM5$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	6.3676	0.9367 (0.0029)	6.3982 (0.1329)	6.3954	0.9442 (0.0026)	6.8128 (0.1451)
	200	6.3828	0.9428 (0.0022)	6.5108 (0.0989)	6.3827	0.9445 (0.0019)	6.4212 (0.0884)
2	100	7.7655	0.9379 (0.0031)	8.2689 (0.1965)	12.5904	0.9470 (0.0026)	13.9975 (0.3207)
	200	7.7979	0.9441 (0.0022)	8.1410 (0.1388)	12.5435	0.9446 (0.0020)	12.7702 (0.1609)
3	100	8.7704	0.9405 (0.0034)	9.8278 (0.3086)	19.5889	0.9450 (0.0032)	22.6252 (0.6393)
	200	8.8182	0.9443 (0.0026)	9.4723 (0.1757)	19.3835	0.9442 (0.0022)	20.1380 (0.3077)

5. APPLICATION TO A REAL DATA SET

The proposed sieve bootstrap method was applied to the daily highest Yahoo stock prices from February 4, 2009 to March 31, 2011; 544 observations in total. The data set can be found at `http://finance.yahoo.com/q/hp?s=YHOO`. The time series is displayed in Figure 1 and clearly exhibits a unit root behavior. The first 535 observations were used to compute 95% prediction intervals for the next consecutive 10 days using the proposed method. The dashed lines in Figure 2 show the upper and lower bounds of the computed prediction intervals. The sieve bootstrap method was able to capture the true future values of this empirical time series precisely and accurately confirming the results in the simulation study.

Figure 1. Daily Highest Yahoo Stock Prices

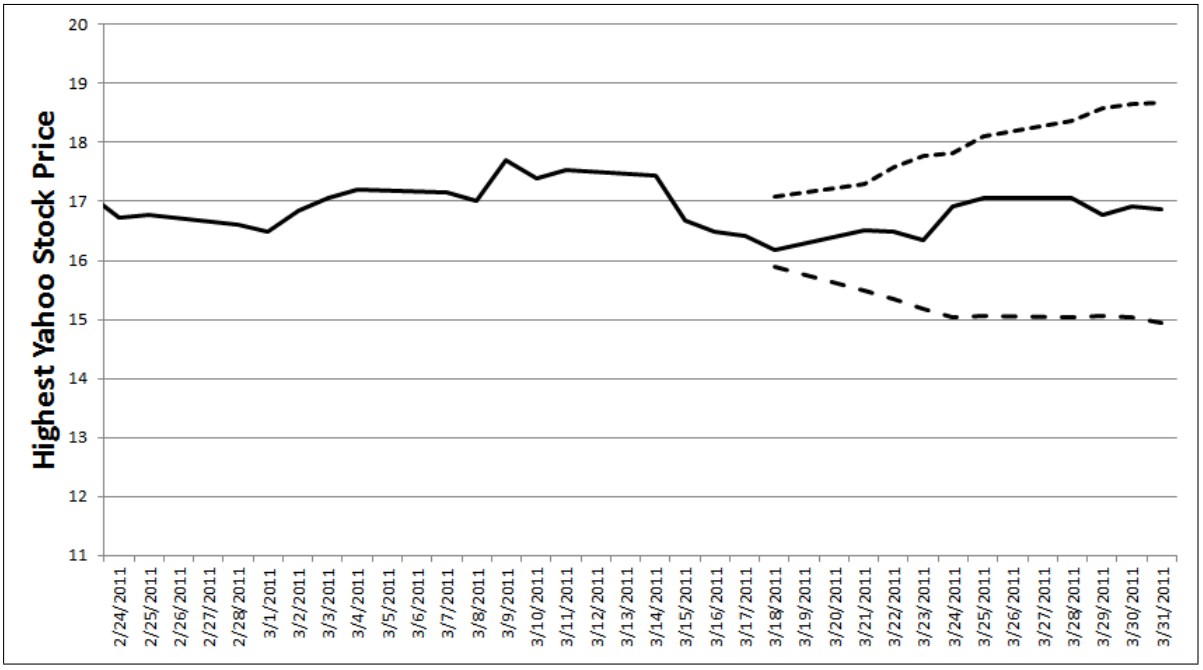

Figure 2. 95% SB Prediction Bands for Yahoo Stock Prices in dashed lines; Only a segment of Figure 1 is displayed

6. CONCLUSION

In this paper, we proposed a sieve bootstrap based prediction intervals for unit root ($ARIMA$) processes that provides proper coverage without altering the computational steps based on the results of a unit root test. Large sample properties are established for the proposed method and a Monte-Carlo simulation study was carried out. The Monte-Carlo study indicates that the procedure works very well under normal, exponential and t distributed errors. Most importantly, the method is stable even when the AR polynomial of the underlying process has a root close to unity.

7. REFERENCES

[1] Alonso A. M., Pena D., Romo J., 2002. Forecasting time series with sieve bootstrap. Journal of Statistical Planning and Inference. 100, 1-11.

[2] Alonso A. M., Pena D., Romo J., 2003. On sieve bootstrap prediction intervals. Statistics and Probability Letters. 65, 13-20.

[3] Alonso A. M., Pena D., Romo J., 2004. Introducing model uncertainty in time series bootstrap. Statistica Sinica. 14, 155-174.

[4] Bühlmann P., 1995. Moving-average representation of autoregressive approximations. Stochastic Process. Appl. 60, 331342.

[5] Bühlmann P., 1997. Sieve bootstrap for time series. Bernoulli, 3, 123-148.

[6] Cao, R., Febrero-Bande, M., Gonzalez-Manteiga, W., Prada-Sanchez, J.M., Garcia-Jurado, I., 1997. Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. Comm. Statist. Simulation Comput. 26, 961-978.

[7] Chang Y.,Park Y., 2003. A sieve bootstrap for the test of a unit root. Journal of Time Series Analysis. 24, 379 - 400.

[8] Dickey D.A.,Fuller W.A., 1979. Distribution of estimators for autoregressive time series with a unit root. Journal of the American Statistical Association. 74, 427 - 431.

[9] Kilian L. 1998(a). Small-sample Condence Intervals for Impulse Response Functions. Review of Economics and Statistics. 80(2), 218-230.

[10] Kim J., 2001. Bootstrap-After-Bootstrap Prediction Intervals for Autoregressive Models. Journal of Business and Economic Statistics. 19, 117-128.

[11] Kreiss,J.P., 1988. Asymptotic Statistical Inference for a Class of Stochastic Processes. Habilitationsschrift, Universiti at Hamburg.

[12] Kreiss J.P., 1992. Bootstrap procedures for AR(∞)-processes. In: JEockel, K.H., Rothe, G., Sendler, W. (Eds.), Bootstrapping and Related Techniques. Springer, Heidelberg, pp. 107-113.

[13] Künsch H.R., 1989. The jackknife and the bootstrap for general stationary observations. The Annals of Statistics. 17, 1217-1241.

[14] Mukhopadhyay P., Samaranayake V.A., 2010. Prediction intervals for time series: A Modified Sieve Bootstrap Approach. Communications in Statistics - Simulation and Computation. 39, 517-538.

[15] Palm F.C., Smeekes S., Urbain J., 2007. Bootstrap unit-root tests: comparison and extensions. Journal of Time Series Analysis. 20, 371-401.

[16] Panichkitkosolkul W., Niwitpong S., 2011. On Multistep-Ahead Prediction Intervals Following Unit Root Tests for a Gaussian AR(1) Process with Additive Outliers. Applied Mathematical Sciences. 47, 2297-2316.

[17] Panichkitkosolkul W., Niwitpong S., 2011. Prediction intervals for the Gaussian autoregressive processes following the unit root tests. Model Assisted Statistics and Applications. 7, 1-15.

[18] Pascual, L., Romo, J., Ruiz, E., 2004. Bootstrap predictive inference for ARIMA processes. Journal of Time Series Analysis. 25 , 449-465.

[19] Poskitt D. S., 2006. Autoregressive approximation in nonstandard situations: the fractionally integrated and non-invertible cases. Annals of the Institute of Statistical Mathematics. 59, 697-725

[20] Poskitt D. S., 2008. Properties of the Sieve bootstrap for fractionally integrated and non-invertible processes. Journal of Time Series Analysis. 29, 224-250.

[21] Psaradakis Z., 2000. Bootstrap test for unit roots in seasonal autoregressive models. Statistics and Probability Letters. 22, 389-395.

[22] Psaradakis Z., 2001. Bootstrap test for an autoregressive unit root in the presence of weakly dependent errors. Journal of Time Series Analysis. 22, 577-594.

[23] Rupasinghe M., Samaranayake V.A., 2012. Asymptomatic Properties of Sieve Bootstrap Prediction Intervals for FARIMA Processes. Statistics and Probability Letters. Submitted.

[24] Said E., Dickey D., 1984. Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order. Biometrika. 71, 599607

[25] Shin D., Sarkar S., Lee J., 1996. Unit root tests for time series with outliers. Statistics & Probability Letter. 30, 189-197.

[26] Stine, R.A., 1987. Estimating properties of autoregressive forecasts. J. Amer. Statist. Assoc. 82, 1072-1078.

[27] Thombs, L.A., Schucany, W.R., 1990. Bootstrap prediction intervals for autoregression. J. Amer. Statist. Assoc. 85, 486-492.

IV. THE ASYMPTOTIC DISTRIBUTIONS OF THE DICKEY-HASZA-FULLER SEASONAL UNIT ROOT TESTS UNDER WEAKLY DEPENDENT ERRORS

ABSTRACT

The Dickey-Hasza-Fuller (DHF) test is frequently used by applied time series analysts to determine whether or not a seasonal unit root is present in the model underlying an observed process. The asymptotic distributions of the DHF test statistics have been derived, as functional of the standard Brownian motion, under the assumption that the time series can be represented by an autoregressive (AR) model that consists of only a seasonal factor, and independent and identically distributed innovations. In this paper, the asymptotic distribution of DHF type test statistics are derived under the assumption of weakly dependent innovations. Autoregressive Moving Average time series with a more general dependent structure than a purely seasonal AR model satisfy this assumption and thus the asymptotic results presented here in provides a theoretical framework for the use of the DHF tests for $ARMA$ processes.

Keywords: Unit root processes; Nonstationarity; Wiener Process; Seasonal Integration

1. INTRODUCTION

Seasonal time series models are extensively used in analyzing empirical data, such as monthly sales, that show annual cycles. Seasonal models that have an autoregressive moving average ($ARMA$) structure with unit roots in the autoregressive polynomial were introduced by Box and Jenkins (1970). Following their formulation we let the seasonal time series, $\{x_t : t \in \mathbb{Z}\}$, be defined by

$$(1 - \rho B^s)\alpha(B)x_t = \theta(B)\epsilon_t, t \in \mathbb{Z}, \tag{1.1}$$

where $\alpha(z) = 1 - \alpha_1 z - .... - \alpha_p z^p$ and $\theta(z) = 1 + \theta_1 z + ... + \theta(q)z^q$ represent autoregressive (AR) and moving average (MA) polynomials of degrees p and q respectively. It is assumed that $\alpha(\cdot)$ and $\theta(\cdot)$ do not share common zeros and that $\alpha(z) \neq 0$ for $|z| \leq 1$. The error terms, $\{\epsilon_t\}$, are assumed to be zero-mean white noise with finite variance σ^2. When $s \geq 2$ and $\rho = 1$, we have a seasonal unit root in the $AR(p+s)$ polynomial $(1 - B^s)\alpha(B)$ and the time series will exhibit cyclical behavior. For example, $s = 2$, $s = 4$ and $s = 12$ indicate that the underlying process follows a cycle with a one year period when the data are gathered biannual, quarterly and monthly intervals. By inspecting the empirical series, a practitioner could identify the period of the cyclical behavior and therefore, it is reasonable to assume that the value of s is known.

2. PRELIMINARIES

To place the asymptotic distribution of the DHF test under weakly dependent errors (that will be derived later in the paper), it is important to give a brief overview of the unit root tests for non-seasonal time series. Dickey and Fuller (1979) were the first to introduce a procedure for testing the null hypothesis of $\rho = 1$ against the alternative

$|\rho| < 1$ for first order autoregressive ($AR(1)$) processes $\{x_t\}$ satisfying

$$x_t = \rho x_{t-1} + \epsilon_t, t \in \mathbb{Z}, \tag{2.1}$$

where $\epsilon_t \sim i.i.d.N(0, \sigma^2)$. Observe that Model (1.1) can also be written in a form similar to (2.1) by letting $u_t = [\alpha(B)]^{-1}\theta(B)\epsilon_t$, so that

$$x_t = \rho x_{t-s} + u_t, \qquad t \in \mathbb{Z}. \tag{2.2}$$

2.1. THE DF TEST FOR NON-SEASONAL ($S = 1$) TIME SERIES

The widely used procedure for testing the null hypothesis $H_0 : \rho = 1$ versus $H_a : |\rho| < 1$ under the model formulation given in (2.1) is the Dickey-Fuller (DF) test introduced by Dickey, and Fuller (1979). For an observed realization, $\{x_t : t = 1, 2, ..., n\}$, that follows Model (2.2) and $u_t \sim i.i.d.(0, \sigma^2)$, the DF test statistics are:

$M_n := n(\hat{\rho}_n - 1)$ and $N_n := \dfrac{(\hat{\rho}_n - 1)}{\hat{\tau}_n}$,

where $\hat{\rho}_n$ is the least-squares estimator of ρ, given by $\hat{\rho}_n := (\sum_{t=1}^{n} x_{t-1}^2)^{-1} \sum_{t=1}^{n} x_{t-1}x_t$, and $\hat{\tau}_n$ is the standard error of $\hat{\rho}_n$.

The asymptotic distributions of the above test statistics under $H_0 : \rho = 1$, for the data generated by the process defined in (2.1), are well known (see Dickey and Fuller (1979)). The same statistics M_n and N_n can be computed even if the underlying process is given by Equation (2.2), with $s = 1$, where the innovation are not assumed to be $i.i.d.(0, \sigma^2)$. Hamilton (1994) provides the asymptotic distribution of M_n and N_n, for the case $s = 1$, under the following assumptions about the process $\{u_t\}$ in Equation (2.2).

(C1) $\{u_t\}_{t\in\mathbb{Z}}$ is the linear process,

$$u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \tag{2.3}$$

where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ are $i.i.d$ random variables with $E[\epsilon_t] = 0$, $E[\epsilon_t^2] = \sigma^2$ and $E[\epsilon_t^4] < \infty$ for all $t \in \mathbb{Z}$.

(C2) The sequence of constants $\{\psi_j\}_{j\in\mathbb{N}_0}$ in the condition (C1) is such that $\sum_{j=0}^{\infty} j|\psi_j| < \infty$, $\sum_{j=0}^{\infty}\psi_j \neq 0$, and $\sum_{j=0}^{\infty}\psi_j z^j \neq 0$ in $\{z \in \mathbb{C} : |z| \leq 1\}$.

We first begin with the non-seasonal case to provide a perspective for the proposed method and then move onto seasonal time series. In the following, let $\mathscr{L}[Y]$ denote the law of Y. Observe that Equation (2.2), with $s = 1$, under the above assumptions generalizes the *i.i.d.* $N(0, \sigma^2)$ condition imposed by Equation (2.1). Under these more general assumptions, Hamilton (1994) shows that

$$\mathscr{L}[M_n] \xrightarrow{w} \mathscr{L}\left[\left\{\int_0^1 \mathbb{W}_0^2(r)dr\right\}^{-1}\left\{\int_0^1 \mathbb{W}_0(r)d\mathbb{W}_0(r) + \frac{\lambda-\lambda_u}{2\lambda}\right\}\right],$$

and

$$\mathscr{L}[N_n] \xrightarrow{w} \mathscr{L}\left[\left\{\int_0^1 \mathbb{W}_0^2(r)dr\right\}^{-1/2}\left\{\int_0^1 \mathbb{W}_0(r)d\mathbb{W}_0(r)) + \frac{\lambda-\lambda_u}{2\lambda}\right\}\right],$$

where $\mathbb{W}_0 = \{\mathbb{W}_0(r) : r \in [0,1]\}$ is the standard (one-dimensional) Brownian motion, and
$\lambda_u := var[u_t] = \sigma^2\sum_{j=0}^{\infty}\psi_j^2$,
$\lambda := \lim_{n\to\infty} var[n^{-1/2}\sum_{t=1}^{n} u_t] = \sigma^2(\sum_{j=0}^{\infty}\psi_j)^2$.

2.2. THE ADF TEST FOR NON-SEASONAL TIME SERIES

Said and Dickey (1984) generalized the DF tests, by what is known as the Augmented Dickey-Fuller test (ADF), to accommodate higher order $ARMA$ models. Given an $ARMA(p,q)$ process with unknown orders p and q, they approimated it with a k^{th} order AR process, where it is assumed that there exists $c > 0$, $r > 0$ such that $ck > n^{1/r}$ and $n^{-1/3}k \to 0$ as $n \to \infty$. They then fitted the regression model

$$x_t = \rho x_{t-1} + \sum_{j=1}^{k}\gamma_j \bigtriangledown x_{t-j} + \epsilon_t \tag{2.4}$$

using least squares, where $\bigtriangledown x_{t-j} = x_{t-j} - x_{t-j-1}$. As was done by Dickey and Fuller (1979), they assumed that the $\{\epsilon_t\}$ are independent and identically distributed.

The ADF test statistics under the above formulation are $P_n := n(\hat{\rho}_n - 1)$ and $Q_n := \dfrac{(\hat{\rho}_n - 1)}{\hat{\tau}_n}$, where $\hat{\rho}_n$ is the least squares estimator of ρ, obtained from the fitting of Equation (2.4). Observe that unlike estimating ρ in Equation (2.2) with by fitting an $AR(1)$ model, here ρ is estimated by fitting a model that has x_{t-1} as well as $\bigtriangledown x_{t-j}$, $j = 1, 2, ..., k$. Moreover, it is assumed that the innovations $\{\epsilon_t\}$ are $i.i.d.(0, \sigma^2)$ rather than weakly dependent. Again, the asymptotic distributions of these test statistics, as functional of the standard Brownian motion (see Hamilton (1994)), are:

$$\mathscr{L}[P_n] \xrightarrow{w} \mathscr{L}\left[\left\{\int_0^1 \mathbb{W}_0^2(r)dr\right\}^{-1}\left\{\int_0^1 \mathbb{W}_0(r)d\mathbb{W}_0(r)\right\}\right],$$

and

$$\mathscr{L}[Q_n] \xrightarrow{w} \mathscr{L}\left[\left\{\int_0^1 \mathbb{W}_0^2(r)dr\right\}^{-1/2}\left\{\int_0^1 \mathbb{W}_0(r)d\mathbb{W}_0(r))\right\}\right].$$

2.3. THE DHF TEST FOR SEASONAL TIME SERIES

The Dickey-Hasza-Fuller (DHF) seasonal test is widely used for testing the null hypothesis $H_0 : \rho = 1$ under the formulation given in Model (2.2) with $s \geq 2$. Introduced by Dickey, Hasza and Fuller (1984), the test assumes that $\{u_t\}$ are $i.i.d.(0, \sigma^2)$. The DHF test statistics are:

$K_n := n(\hat{\rho}_n - 1)$ and $T_n := \dfrac{(\hat{\rho}_n - 1)}{\hat{\tau}_n}$,

where $\hat{\rho}_n$ is the least-squares estimator of ρ, $\hat{\rho}_n := (\sum_{t=1}^n x_{t-s}^2)^{-1} \sum_{t=1}^n x_{t-s}x_t$, and $\hat{\tau}_n$ is the standard error of $\hat{\rho}_n$.

Under the null hypothesis $H_0 : \rho = 1$, the asymptotic distributions of the above test statistics, for the process defined in (2.2) with *i.i.d.* errors, have been established as functional of the standard Brownian motion. For example, in Ghysels et *al.* (2000), states that

$$\mathscr{L}[K_n] \xrightarrow{w} \mathscr{L}\left[\left\{\sum_{m=1}^{s}\int_0^1 \mathbb{W}_m^2(r)dr\right\}^{-1}\left\{s\sum_{m=1}^{s}\left[\int_0^1 \mathbb{W}_m(r)d\mathbb{W}_m(r)\right]\right\}\right],$$

and

$$\mathscr{L}[T_n] \xrightarrow{w} \mathscr{L}\left[\left\{\sum_{m=1}^{s}\int_0^1 \mathbb{W}_m^2(r)dr\right\}^{-1/2}\left\{\sum_{m=1}^{s}\left[\int_0^1 \mathbb{W}_m(r)d\mathbb{W}_m(r)\right]\right\}\right],$$

where $\mathbb{W}_m = \{\mathbb{W}_m(r) : r \in [0,1]\}$ for $m = 1, 2, ..., s$ are independent standard (one-dimensional) Brownian motions.

The implication of the assumption that $\{u_t\}$ in Model (2.2) are $i.i.d.(0, \sigma^2)$ is that the DHF test is applicable to only purely seasonal time series. In otherwords, it only applies to time series whose underlying model satisfies Equation (1.1) with $\alpha(z) = \theta(z) = 1$. Unlike in the case of non-seasonal unit root testing, there is no "Augmented" extension of the DHF seasonal unit root test. More recenlty, Castro, Osburn and Taylor (2011), however, expressed the distributions of the HEGY test statistics (Hylleberg et *al.* (1990)) under serially correlated errors but did not consider the DHF test. The objective of this paper is to fill this gap by deriving the asymptotic distributions of the DHF test statistics under the relaxed assumption of weakly dependent errors.

3. THE ASYMPTOTIC DISTRIBUTIONS

In this section, we derive the asymptotic distributions of the test statistics K_n and T_n, under the null hypothesis $H_0 : \rho = 1$ for the model given in (2.2). The following proposition plays the key roll in obtaining the desired results, and an equivalent result for non-seasonal time series can be found in Hamilton (1994).

Proposition 3.1. *If conditions (C1)-(C2) hold for Model (2.2), then for* $i \in \mathbb{N}$ *and* $m = 0, 1, ..., (s-1)$,

$$u_m + u_{s+m} + ... + u_{s(i-1)+m} = \sum_{j=0}^{s-1} [\psi_j(1) \sum_{k=1}^{i} \epsilon_{s(k-1)+m-j}] + \eta_{i,m} - \eta_{0,m},$$

where $\psi_j(1) = \sum_{q=0}^{\infty} \psi_{sq+j}$, $\eta_{i,m} = \sum_{q=1}^{\infty} \sum_{j=0}^{s-1} \alpha_{q,j} \epsilon_{s(i-q)+m-j}$, $\alpha_{q,j} = -(\psi_{sq+j} + \psi_{s(q+1)+j} + \psi_{s(q+2)+j} + ...)$, *and* $\sum_{q=1}^{\infty} \sum_{j=0}^{s-1} |\alpha_{q,j}| < \infty$.

Proof. Observe that, for $i \in \mathbb{N}$,

$$\begin{aligned}
&\sum_{k=1}^{i} u_{s(k-1)+m} = \sum_{k=1}^{i} \sum_{j=0}^{\infty} \psi_j \epsilon_{s(k-1)+m-j} \\
&= \{\psi_0 \epsilon_{s(i-1)+m} + \psi_1 \epsilon_{s(i-1)+m-1} + ... + \psi_s \epsilon_{s(i-2)+m} + ...\} \\
&+ \{\psi_0 \epsilon_{s(i-2)+m} + \psi_1 \epsilon_{s(i-2)+m-1} + ... + \psi_s \epsilon_{s(i-3)+m} + ...\} \\
&+ ... + \{\psi_0 \epsilon_m + \psi_1 \epsilon_{m-1} + ... + \psi_s \epsilon_{-s+m} + ...\} \\
&= \psi_0 \epsilon_{s(i-1)+m} + (\psi_0 + \psi_s) \epsilon_{s(i-2)+m} + ... + (\psi_0 + \psi_s + ... + \psi_{s(i-1)}) \epsilon_m + ... \\
&+ \psi_1 \epsilon_{s(i-1)+m-1} + (\psi_1 + \psi_{s+1}) \epsilon_{s(i-2)+m-1} + ... \\
&+ (\psi_1 + \psi_{s+1} + ... + \psi_{s(i-1)+1}) \epsilon_{m-1} + ... \\
&+ \psi_{s-1} \epsilon_{s(i-1)+m-(s-1)} + (\psi_{s-1} + \psi_{s+s-1}) \epsilon_{s(i-2)+m-(s-1)} + ... \\
&+ (\psi_{s-1} + \psi_{s+s-1} + ... + \psi_{s(i-1)+s-1}) \epsilon_{m-(s-1)} + ... \\
&+ (\psi_s + ... + \psi_{s(i-1)} + \psi_{si}) \epsilon_{-s+m} \\
&+ (\psi_{s+1} + ... + \psi_{s(i-1)+1} + \psi_{si+1}) \epsilon_{-s+m-1} + ...
\end{aligned}$$

$$
\begin{aligned}
&= (\psi_0 + \psi_s + \psi_{2s} + ...)\epsilon_{s(i-1)+m} - (\psi_s + \psi_{2s} + ...)\epsilon_{s(i-1)+m} \\
&+ (\psi_0 + \psi_s + \psi_{2s} + ...)\epsilon_{s(i-2)+m} - (\psi_{2s} + \psi_{3s} + ...)\epsilon_{s(i-2)+m} + ... \\
&+ (\psi_1 + \psi_{s+1} + \psi_{2s+1} + ...)\epsilon_{s(i-1)+m-1} - (\psi_{s+1} + \psi_{2s+1} + ...)\epsilon_{s(i-1)+m-1} \\
&+ (\psi_1 + \psi_{s+1} + \psi_{2s+1} + ...)\epsilon_{s(i-2)+m-1} - (\psi_{2s+1} + \psi_{3s+1} + ...)\epsilon_{s(i-2)+m-1} + ... \\
&+ ... \\
&+ (\psi_{s-1} + \psi_{2s-1} + \psi_{3s-1} + ...)\epsilon_{s(i-1)+m-(s-1)} - (\psi_{2s-1} + \psi_{3s-1} + ...)\epsilon_{s(i-1)+m-(s-1)} \\
&+ (\psi_{s-1} + \psi_{2s-1} + \psi_{3s-1} + ...)\epsilon_{m-(s-1)} - (\psi_{3s-1} + \psi_{4s-1} + ...)\epsilon_{m-(s-1)} + ... \\
&+ (\psi_s + ... + \psi_{s(i-1)} + \psi_{si} + ...)\epsilon_{-s+m} - (\psi_{s(i+1)} + ... + \psi_{s(i+2)} + ...)\epsilon_{-s+m} \\
&+ (\psi_{s+1} + ... + \psi_{s(i-1)+1} + \psi_{si+1} + ...)\epsilon_{-s+m-1} \\
&- (\psi_{s(i+1)+1} + ... + \psi_{s(i+2)+1} + ...)\epsilon_{-s+m-1} \\
&+ ... \qquad\qquad .
\end{aligned}
$$

Hence, we can write

$$\sum_{k=1}^{i} u_{s(k-1)+m} = \sum_{j=0}^{s-1}[\psi_j(1)\sum_{k=1}^{i}\epsilon_{s(k-1)+m-j}] + \eta_{i,m} - \eta_{0,m},$$

where

$$
\begin{aligned}
\eta_{i,m} &= -(\psi_s + \psi_{2s} + ...)\epsilon_{s(i-1)+m} - (\psi_{2s} + \psi_{3s} + ...)\epsilon_{s(i-2)+m} + ... \\
&- (\psi_{s+1} + \psi_{2s+1} + ...)\epsilon_{s(i-1)+m-1} - (\psi_{2s+1} + \psi_{3s+1} + ...)\epsilon_{s(i-2)+m-1} + ... \\
&- (\psi_{2s-1} + \psi_{3s-1} + ...)\epsilon_{s(i-1)+m-(s-1)} - (\psi_{3s-1} + \psi_{4s-1} + ...)\epsilon_{s(i-2)+m-1} + ... \\
\eta_{0,m} &= -(\psi_s + ... + \psi_{s(i-1)} + \psi_{si} + ...)\epsilon_{-s+m} \\
&- (\psi_{s+1} + ... + \psi_{s(i-1)+1} + \psi_{si+1} + ...)\epsilon_{-s+m-1} + ...
\end{aligned}
$$

Notice that $\eta_{i,m} = \sum_{q=1}^{\infty}\sum_{j=0}^{s-1}\alpha_{q,j}\epsilon_{s(i-q)+m-j}$, where $\alpha_{q,j} = -(\psi_{sq+j} + \psi_{s(q+1)+j} + \psi_{s(q+2)+j} + ...)$, with

$$\sum_{q=1}^{\infty}\sum_{j=0}^{s-1}|\alpha_{q,j}| \leq \sum_{j=0}^{\infty} j|\psi_j| < \infty,$$

which establishes the proposition. □

The following proposition shows the convergence results of important summations that appear in the test statistics. Here we denote independent standard Brownian motions by $\mathbb{W}_{m,j} = \{\mathbb{W}_{m,j}(r) : r \in [0,1]\}$.

Proposition 3.2. *If conditions (C1)-(C2) hold with $\rho = 1$ in the model (2.2) and $n = s(N-1)+m$. Then, as $n \to \infty$,*

(a) $\mathscr{L}[n^{-1/2}\sum_{m=1}^{s}\sum_{k=1}^{N}u_{s(k-1)+m}] \xrightarrow{w}$
$\mathscr{L}\left[s^{-1/2}\sigma\sum_{j=0}^{s-1}[\psi_j(1)\sum_{m=1}^{s}\mathbb{W}_{m,j}(1)]\right]$,

(b) $\mathscr{L}[n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}x_{s(k-2)+m}\epsilon_{s(k-1)+m}] \xrightarrow{w}$
$\mathscr{L}\left[(1/2)s^{-1}\sum_{m=1}^{s}\sum_{j=0}^{s-1}\psi_j(1)[\sigma^2\{\mathbb{W}_{m,j}^2(1)-1\}]\right]$,

(c) $\mathscr{L}[n^{-2}\sum_{m=1}^{s}\sum_{k=1}^{N}x_{s(k-1)+m}^2] \xrightarrow{w}$
$\mathscr{L}\left[s^{-2}\sigma^2\sum_{j=0}^{s-1}[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1\mathbb{W}_{m,j}^2(r)]dr\right]$, *and*

(d) $\mathscr{L}[n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}x_{s(k-2)+m}u_{s(k-1)+m}] \xrightarrow{w}$
$\mathscr{L}\left[(1/2)\sigma^2\left[s^{-1}\sum_{j=0}^{s-1}\left(\psi_j^2(1)\sum_{m=1}^{s}\mathbb{W}_{m,j}^2(1)\right)-\sum_{j=0}^{\infty}\psi_j^2\right]\right]$.

Proof. This result follows immediately from Proposition 3.1 and results in Hamilton (1994, Chap. 17) with the appropriate modifications. For example, for part (b), using Proposition 3.1,

$$
\begin{aligned}
& n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N} x_{s(k-2)+m}\epsilon_{s(k-1)+m} \\
& = n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}\{\sum_{j=0}^{s-1}[\psi_j(1)\sum_{i=1}^{k-1}\epsilon_{s(i-1)+m-j}] + \eta_{k-1,m} - \eta_{0,m}\}\epsilon_{s(k-1)+m} \\
& = n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}\{\sum_{j=0}^{s-1}[\psi_j(1)\sum_{i=1}^{k-1}\epsilon_{s(i-1)+m-j}]\epsilon_{s(k-1)+m}\} \\
& + n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}[\eta_{k-1,m} - \eta_{0,m}]\epsilon_{s(k-1)+m} \\
& = s^{-1}\sum_{m=1}^{s}\sum_{j=0}^{s-1}\psi_j(1)\{N^{-1}\sum_{k=1}^{N}(\sum_{i=1}^{k-1}\epsilon_{s(i-1)+m-j})\epsilon_{s(k-1)+m}\} \\
& + s^{-1}\sum_{m=1}^{s}\{N^{-1}\sum_{k=1}^{N}[\eta_{k-1,m} - \eta_{0,m}]\epsilon_{s(k-1)+m}\}.
\end{aligned}
$$

From Hamilton (1994, Chap. 17), we have

$$
\mathscr{L}[N^{-1}\sum_{k=1}^{N}(\sum_{i=1}^{k-1}\epsilon_{s(i-1)+m-j})\epsilon_{s(k-1)+m}\}] \xrightarrow{w} \mathscr{L}[(1/2)\sigma^2\{\mathbb{W}^2_{m,j}(1) - 1\}],
$$

and for each $m = 1, ..., s$,

$$
N^{-1}\sum_{k=1}^{N}[\eta_{k-1,m} - \eta_{0,m}]\epsilon_{s(k-1)+m} \xrightarrow{P} 0.
$$

Hence Part (b) is proven.

□

Remark 3.3. *All the other asymptotic results stated on page 507 in Hamilton (1994) can be established for seasonally integrated processes with weakly dependent errors, but are omitted to save space here.*

We finally derive the asymptotic distributions of the DHF test statistics under weakly dependent errors.

Theorem 3.4. *Under the null hypothesis $H_0 : \rho = 1$, with conditions (C1)-(C2) holding for Model (2.2), as $n \longrightarrow \infty$,*

$$\mathscr{L}[K_n] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{s\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s^2/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]}\right],$$

and

$$\mathscr{L}[T_n] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\left\{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]\right\}^{-1/2}}\right].$$

Proof. Note that

$$\begin{aligned} K_n &= n[(\sum_{t=1}^{n} x_{t-s}^2)^{-1}\sum_{t=1}^{n} x_{t-s}x_t - 1] \\ &= n[(\sum_{t=1}^{n} x_{t-s}^2)^{-1}\sum_{t=1}^{n} x_{t-s}u_t] \end{aligned}$$

and therefore, with the aid of Proposition 3.2 and the fact that $\int_0^1 \mathbb{W}_{m,j} = \frac{1}{2}\{\mathbb{W}_{m,j}^2(1) - 1\}$, the convergence is established.

Similar proof can be obtained for T_n utilizing Proposition 3.2 as in Hamilton (1994, Chap. 17). □

Remark 3.5. *One could obtain the asymptotic distributions of the DF non-seasonal unit root test statistics by setting $s = 1$ in the above theorem. Moreover, the asymptotic distribution for the seasonal case with i.i.d errors can be obtained by letting $\psi_0 = 1$ and $\psi_j = 0$ for $j \in \mathbb{N}$.*

4. CONCLUSION

In this paper, the functional of standard Brownian motion are obtained for the Dickey-Hasza-Fuller seasonal unit root test statistics under weakly dependent errors. This is a generalization of the standard Dickey-Hasza-Fuller tests for seasonal unit roots and is applicable to a wider class of seasonal models.

5. REFERENCES

[1] Box G.P.E.,Jenkins G.M., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, CA.

[2] Castro T.,Osburn D.,Taylor A.M., 2011. On augmented HEGY tests for seasonal unit roots. The University of Manchester Discussion Paper Series. EDP-1121.

[3] Dickey D.A., Fuller W.A., 1979. Distribution of estimators for autoregressive time series with a unit root. Journal of the American Statistical Association. 74, 427 - 431.

[4] Dickey D.A., Hasza D.P., Fuller W.A. 1984. Testing for unit roots in seasonal time series. Journal of the American Statistical Association. 79, 355 - 367.

[5] Ghysels E. ,Osborn D.R, Rodrigues P.M.M., 2000. Seasonal Nonstationarity and Near-Nonstationarity, in: B.H. Baltagi (Ed.), A Companion to Theoretical Econometrics, Blackwell Publishers, Oxford.

[6] Hamilton, J. D. (1994) Time Series Analysis. Princeton: Princeton University Press.

[7] Hylleberg, S., Engle, R. F., Granger, C. W. J., Yoo, B. S., 1990. Seasonal integration and cointegration. Journal of Econometrics. 44, 215-238.

[8] Said E., Dickey D., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrica. 71, 599-607.

V. SIEVE BOOTSTRAP FOR SEASONAL TIME SERIES: UNIT ROOT TESTS AND PREDICTION INTERVALS

ABSTRACT

The sieve bootstrap, which obtains residuals for re-sampling by fitting finite order autoregressive models to time series, can be utilized to obtain prediction intervals as well as approximate distribution of statistics of interest. While this re-sampling method has been used to obtain prediction intervals for $ARMA$ processes and test for non-seasonal unit roots, it has not been adopted to obtain prediction intervals for seasonal time series or to test for seasonal unit roots. In this paper, conditions under which the sieve bootstrap can be applied to time series with a seasonal unit root are derived. In particular, its application for obtaining prediction intervals and for conducting Dickey-Hasza-Fuller (DHF) type tests for a seasonal unit root are considered. The asymptotic properties of the proposed prediction intervals and the unit root tests are derived and finite sample properties of these procedures are studied using Monte Carlo simulations. The simulation results indicate that both the prediction intervals and the unit root test based on the sieve bootstrap have good finite sample properties.

Keywords: Unit root processes; Forecast intervals; Seasonal unit roots; Nonstationarity; Seasonal integration

1. INTRODUCTION

In the past decades seasonal time series models have been extensively used in modeling financial and economics data. These seasonal models were originated by Box and Jenkins (1970) and further studied by many researchers; see Ghysels and Osburn (2001). The specific type of seasonal time series, $\{x_t : t \in \mathbb{Z}\}$, that will be the focus of this paper is the process given by

$$(1 - \rho B^s)\alpha(B)x_t = \theta(B)\epsilon_t, t \in \mathbb{Z}, \tag{1.1}$$

where $\alpha(z) = 1-\alpha_1 z-...-\alpha_p z^p$ and $\theta(z) = 1+\theta_1 z+...+\theta(q)z^q$ represent autoregressive and moving average polynomials of degrees p and q respectively, and B defines the "backshift operator" given by $B^k x_t = x_{t-k}$ for $k \in \mathbb{N}_0$. It is assumed that $\alpha(.)$ and $\theta(.)$ do not share common zeros. The error terms, $\{\epsilon_t\}$, are assumed to be zero-mean white noise with finite variance σ^2. Note that $s \geq 2$ is used to model the seasonality. For example, $s = 2$, $s = 4$ and $s = 12$ indicate that the underlying process follows semi-annual, quarterly and monthly seasonal behaviors respectively. By inspecting the empirical data, the practitioner could identify if the process has a cyclical behavior of a certain period and therefore, it is reasonable to assume that the value of s is known.

For the practitioner, presence of the term $(1 - \rho B^s)$ signifies the existence of a seasonal component of period s. When ρ in that expression equals unity, one obtains a seasonal unit root process, where the effect of identical seasons in previous years have on the corresponding season in the current year do not decay with time. Under this formulation, testing for a seasonal unit root boils down to deciding between the null hypothesis that $\rho = 1$ versus the alternative hypothesis that $|\rho| < 1$. Such a test is important not only to determine if the underlying process is stationary, but also to provide the applied time series analyst valuable information as to the seasonal nature of the underlying process.

Moreover, when $|\rho| < 1$, Equation (1.1) yields a stationary $ARMA$ process and existing procedures for obtaining sieve-bootstrap-based prediction intervals can be directly

applied to the time series. On the other hand, when $\rho = 1$, currently available sieve-bootstrap–based prediction intervals fails to provide asymptotically valid coverage and the intervals provide very liberal coverage for finite sample sizes.

Traditional, and well known, Dickey-Hasza-Fuller test (DHF) (Dickey, Hasza and Fuller (1984)) is a prominent tool for testing whether or not the underlying process of a given seasonal time series has a unit root in the autoregressive polynomial. The DHF test is an extension of the Dickey-Fuller (DF) (Dickey and Fuller (1979)) non-seasonal unit root test. Since the introduction of the DF test, the unit root tests have attracted the attention of many researchers. Another seminal regression-based seasonal unit root test, known as HEGY, was developed by Hylleberg et *al.* (1990). The HEGY test has both $t-$ and $F-$ test statistics, and competitively, it has potential for testing unit roots at the zero, Nyquist and annual (harmonic) frequencies for a quarterly observed series.

The DF-type unit root tests for both seasonal and non-seasonal time series do not perform well under weakly dependent errors and alternatively, bootstrap methods are used in the literature. Psaradakis (2000) implemented a bootstrap method for pure seasonal time series with independent errors and their bootstrap tests have higher powers than the DHF tests. Psaradakis (2001) was the first to introduce the sieve bootstrap to unit root testing for non-seasonal time series with weakly dependent errors. Following Psaradakis (2001), the sieve bootstrap versions of the Augmented Dickey-Fuller tests (ADF) (1984) for non-seasonal unit roots were suggested by Chang and Park (2003). Palm, Smeekes and Urbain (2008) proposed an alternative way of computing residuals by fitting the DF regression model instead of fitting an $AR(p)$ model to the differenced series. Palm et *al.* (2008), therefore, named their proposed method as residual based and Psaradakis (2001) method as difference based. In this paper, we adapt Psaradakis's (2001) difference based and Palm et *al.* (2008) residual based unit root tests for seasonal time series with weakly dependent errors.

The main feature of the sieve bootstrap is the autoregressive approximation introduced by Kreiss (1998) and further refined by Bühlmann (1997). In their formulation, the order of the autoregressive approximation is assumed to be increasing with the sample size thus yielding finer sieves for the underlying infinite-process. The sieve bootstrap

procedure is considered as a model-free and nonparametric resampling method and its current formulation requires the assumption that the process is invertible and has an $AR(\infty)$ representation; see Philips and Solo (1992) and Bühlmann (1997) for more details. Poskitt (2006, 2007) relaxed this condition and showed that the sieve bootstrap can still be applied to a class of regular processes that includes non-invertible and long memory (*Fractionally Integrated Autoregressive Moving Average - FARIMA*) processes.

As the second goal of this paper, we propose the sieve bootstrap method introduced by Rupasinghe and Samaranayake (2012c) for obtaining prediction intervals for seasonal time series. The literature is short on research on computing forecast intervals for seasonal time series and our method would lay the foundation for the future research in this area.

The sieve bootstrap procedure seems to be a promising technique for computing prediction intervals for many types of time series. Alonso, Pẽna and Romo (2002, 2003) utilized the Bühlmann (1997) sieve bootstrap method in computing prediction intervals for a class of linear processes that have an infinite order moving average representation. They also established the asymptotic justification for this procedure. Note that $ARMA$ models are examples of the linear processes. Alonso et *al.* (2004) further refined their method by introducing the sampling uncertainty in parameter estimation and Mukhopadhyay and Samaranayake (2010) introduced a rescaling factor for the residuals to improve the coverage of the sieve bootstrap based prediction intervals. The method of Alonso et *al.* (2003) was extended to $FARIMA$ processes by Rupasinghe and Samaranayake (2012a, 2012b). Following the results in Poskitt (2006, 2007), they establish the large sample validity of the sieve bootstrap based prediction intervals for $FARIMA$ processes by relaxing the invertibility condition. In a seperate paper, Rupasinghe and Samaranayake (2012c) proposed the sieve bootstrap procedure for obtaining prediction intervals for non-stationary *Autoregressive Integrated Moving Average (ARIMA)* processes and showed that their method works even if the underlying process is $ARMA$. In contrast all the other methods currently available to compute prediction intervals either used two different formulations based on the outcome of preliminary unit root tests or assumed the process has a unit root in the AR polynomial. See for example, Panichkitkosolkul and Niwitpong (2011, 2012), Kim (2001) and Pascual, Romo, and Ruiz (2004). As the second goal of

this paper, we propose a sieve bootstrap procedure to obtain prediction intervals for time series with possible seasonal unit roots.

The structure of this article is as follows. Section 2 introduces the sieve bootstrap schemes for testing seasonal unit roots. The sieve bootstrap procedure for prediction intervals is implemented in Section 3 along with establishing the asymptotic properties and a Monte-Carlo simulation study.

2. SEASONAL UNIT ROOT TESTS

Note that the Model (1.1) can be written as,

$$x_t = \rho x_{t-s} + u_t, t \in \mathbb{Z}, \tag{2.1}$$

where $\{u_t\}_{t\in\mathbb{Z}}$ is a stationary stochastic process with zero mean given by $u_t = [\alpha(B)]^{-1}\theta(B)\epsilon_t$, $t \in \mathbb{Z}$. We shall assume that $\{u_t\}$ satisfies the following two conditions.

(C1) $\{u_t\}_{t\in\mathbb{Z}}$ is the linear process,

$$u_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}, \tag{2.2}$$

where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ are *i.i.d.* random variables with $E[\epsilon_t] = 0$, $E[\epsilon_t^2] = \sigma^2$ and $E[\epsilon_t^4] < \infty$ for all $t \in \mathbb{Z}$, and

(C2) the sequence of constants $\{\psi_j\}_{j\in\mathbb{N}_0}$ in the condition (C1) is such that $\sum_{j=0}^{\infty} j|\psi_j| < \infty$, $\sum_{j=0}^{\infty} \psi_j \neq 0$, and $\sum_{j=0}^{\infty} \psi_j z^j \neq 0$ in $\{z \in \mathbb{C} : |z| \leq 1\}$.

The condition (C2) implies that u_t is invertible and has an $AR(\infty)$ representations which is one of the assumptions of Bühlmann (1997).

In this section, we focus on testing the null hypothesis $H_0 : \rho = 1$ versus the alternative hypothesis $H_a : |\rho| < 1$, where ρ is the parameter associated with the factor $(1 - \rho B^s)$ in the autoregressive polynomial in Model (1.1). The process is nonstationary

if $\rho = 1$ and consequently, it should be seasonally differenced to acquire the stationarity. As discussed in the introduction, the widely used procedure for testing this hypothesis is the Dickey-Hasza-Fuller (DHF) seasonal unit root test (Dickey, Hasza and Fuller (1984)). Their test, however, is for pure seasonal models with $\alpha(z) = 1$ and $\theta(z) = 1$ for all $z \in \mathbb{R}$ in Model (1.1). For an observed realization, $\{x_t : t = 1, 2, ..., n\}$, that follows Model (1.1), the DHF test statistics are:

$$K_n := n(\hat{\rho}_n - 1) \text{ and } T_n := \frac{(\hat{\rho}_n - 1)}{\hat{\tau}_n},$$

where $\hat{\rho}_n$ is the least-squares estimator of ρ, $\hat{\rho}_n := (\sum_{t=1}^{n} x_{t-s}^2)^{-1} \sum_{t=1}^{n} x_{t-s} x_t$, and $\hat{\tau}_n$ is the standard error of $\hat{\rho}_n$.

Under the null hypothesis $H_0 : \rho = 1$, the asymptotic distributions of the above test statistics, for the process defined in (2.1) with $\{u_t\} \sim i.i.d.(0, \sigma^2)$ errors, have been established as a functional of standard Brownian motions. For example, Ghysels et *al.* (2000) states that when $\{u_t\} \sim i.i.d.(0, \sigma^2)$,

$$\mathscr{L}[K_n] \xrightarrow{w} \mathscr{L}\left[\left\{\sum_{m=1}^{s} \int_0^1 \mathbb{W}_m^2(r)dr\right\}^{-1} \left\{s \sum_{m=1}^{s} \left[\int_0^1 \mathbb{W}_m(r)d\mathbb{W}_m(r)\right]\right\}\right]$$

and

$$\mathscr{L}[T_n] \xrightarrow{w} \mathscr{L}\left[\left\{\sum_{m=1}^{s} \int_0^1 \mathbb{W}_m^2(r)dr\right\}^{-1/2} \left\{\sum_{m=1}^{s} \left[\int_0^1 \mathbb{W}_m(r)d\mathbb{W}_m(r)\right]\right\}\right],$$

where $\mathbb{W}_0 = \{\mathbb{W}_0(r) : r \in [0, 1]\}$ is the standard (one-dimensional) Brownian motion.

If the errors $\{u_t\}$ in Model (2.1) are weakly dependent and satisfies conditions (C1)-(C2), Rupasinghe and Samaranayake (2012d) showed that, under $H_0 : \rho = 1$, as $n \longrightarrow \infty$,

$$\mathscr{L}[K_n] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{s\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s^2/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]}\right]$$

and

$$\mathscr{L}[T_n] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\left\{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]\right\}^{-1/2}}\right],$$

where $\psi_j(1) = \sum_{q=0}^{\infty}\psi_{sq+j}$, $\mathbb{W}_{m,j} = \{\mathbb{W}_{m,j}(r) : r \in [0,1]\}$ are independent standard (one-dimensional) Brownian motions.

In the next two subsections, we present two sieve bootstrap procedures, difference-based and residual-based, to test the null hypothesis $H_0 : \rho = 1$ under the model given in (2.1) with weakly dependent errors $\{u_t\}$.

2.1. DIFFERENCE-BASED DHF SEASONAL UNIT ROOT TEST

The sieve bootstrap procedure given below is identical to that introduced by Psaradakis (2001) except for the seasonal differencing. Assume that a realization $\{x_t\}_{t=1}^n$ is obtained from the model given in Equation (2.1). Compute the seasonally differenced series, $u_t = x_t - x_{t-s}$ and proceed with the following steps.

1. Obtain the Yule-Walker estimates, $\hat{\phi}_{1,n}, ..., \hat{\phi}_{p,n}$ of the coefficients of the $AR(p)$ model, $\sum_{j=0}^{p}\phi_{j,n}u_{t-j} = \epsilon_{t,n}$, where $\phi_{0,n} := 0$ and p is chosen so that $p \to \infty$ and $p/n \to 0$ as $n \to \infty$.

2. Obtain the $(n-p)$ residuals: $\hat{\epsilon}_{t,n} = \sum_{j=0}^{p} \hat{\phi}_{j,p,n} u_{t-j}$, $t = p+1, ..., n$ and define the empirical distribution function of the centered residuals, $\tilde{\epsilon}_t = \hat{\epsilon}_{t,n} - \hat{\epsilon}^{(\cdot)}$, where $\hat{\epsilon}^{(\cdot)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n}$, by $\hat{F}_{\tilde{\epsilon},n}(x) = (n-p)^{-1} \sum_{t=p+1}^{n} I_{[\tilde{\epsilon}_t \leq x]}$.

3. Draw a resample $\epsilon^*_{t,n}$, $t = p+1, ..., n$ of i.i.d. observations from $\hat{F}_{\tilde{\epsilon}}$.

4. Obtain $u^*_{t,n}$ by the recursion: $\sum_{j=0}^{p} \hat{\phi}_{j,p,n} u^*_{t-j} = \epsilon^*_{t,n}$ for $t = p+1, ..., n$ and set $u^*_{t,n} = \bar{y}$ for $t = 1, ..., p$.

5. Using $u^*_{t,n}$, construct the bootstrap replicates $\{x^*_t\}_{t=1}^{n}$ according to $x^*_t = x^*_{t-s} + u^*_t$ for $t = 1, ..., n$, where $x^*_{-n} = ... = x^*_0 = 0$. In practice, $n + 200$ observations are generated and then the first 200 of them are discarded for the "burn-in" period.

6. The bootstrap versions of K_n and T_n are obtained by computing corresponding DHF test statistics. That is, $K^{*D}_n := n(\hat{\rho}^*_n - 1)$ and $T^{*D}_n := (\hat{\rho}^*_n - 1)/\tau^*_n$, where $\hat{\rho}^*_n$ and τ^*_n are, respectively, the least-squares estimators of ρ^* and the standard deviation of $\hat{\rho}^*_n$ obtained from in the regression equation $x^*_t = \rho^* x^*_{t-s} + u^*_t$.

7. Steps 3 through 6 are repeated B (sufficiently large) number of times to obtain the bootstrap empirical distribution of K_n and T_n.

8. The null hypothesis $H_0 : \rho = 1$ is rejected at $\gamma\%$ significance level if $K_n < K^*_{\gamma,n}$ and $T_n < T^*_{\gamma,n}$ where $K^*_{\gamma,n}$ and $T^*_{\gamma,n}$ are the γ^{th} percentiles of the corresponding bootstrap distributions.

Whenever it is required, we shall assume that $x_t = y_t = 0$ and for $t \leq 0$.

2.2. RESIDUAL-BASED DHF SEASONAL UNIT ROOT TEST

As suggested by Pam et *al.* (2008), one could compute the residuals directly from the Dickey-Fuller regression model,

$$\hat{\epsilon}_n = x_t - \hat{\rho}_n x_{t-s} - \sum_{j=1}^{p} \hat{\phi}_{j,n} u_{t-j}, \qquad t = p+1, ..., n. \tag{2.3}$$

The remaining steps are the same as the difference-based seasonal unit root test given in Subsection 2.1. We denote the corresponding residual-based test statistics by K_n^{*R} and T_n^{*R}.

2.3. ASYMPTOTIC RESULTS

In this section, the asymptotic convergence results for the difference-based DHF seasonal unit root test statistics are established relying on the conditions (C1) and (C2). In addition, the following two conditions are also required.

(C3) The order p of the autoregressive approximation is such that $p = p(n) \to \infty$ as $n \to \infty$ with $p(n) = o[(n/logn)^{1/4}]$.

(C4) $\hat{\boldsymbol{\phi}}_{p,n} = (\hat{\phi}_{1,n}, ..., \hat{\phi}_{p,n})'$ satisfy the empirical Yule-Walker equations $\hat{\boldsymbol{\Gamma}}_{p,n}\hat{\boldsymbol{\phi}}_{p,n} = -\hat{\boldsymbol{\gamma}}_{p,n}$ where
$\hat{\boldsymbol{\Gamma}}_{p,n} = [\hat{R}(i-j)]_{i,j=1}^{p}, \hat{\boldsymbol{\gamma}}_{p,n} = (\hat{R}(1), ..., \hat{R}(p))'$, and $\hat{R}(j) = n^{-1}\sum_{t=1}^{n-|j|} u_t u_{t+|j|}$ for $|j| < n$.

Condition (C3) was imposed by Bühlmann (1995,1997) to establish the convergence of the infinite moving average representation of the AR approximation to that of theoretical representation given in Condition (C1). In Condition (C4), the Yule-Walker equations are used to ensure that the bootstrap process $\{u_t^*\}_{t\in\mathbb{Z}}$, defined in the recursion Step 4, admits the one-sided $MA(\infty)$ representation

$$u_{t,n}^* = \sum_{j=0}^{\infty} \hat{\psi}_{j,n}\epsilon_{t-j,n}^* \text{ with } \sum_{j=0}^{\infty} |\hat{\psi}_{j,n}| < \infty, \tag{2.4}$$

where $\{\hat{\psi}_{j,n}\}_{j\in\mathbb{N}_0}$ are determined through the relation
$(\sum_{i=0}^{p} \hat{\phi}_{i,n} z^i)(\sum_{j=0}^{p} \hat{\psi}_{j,n} z^j) = 1$ $(|z| \leq 1)$.

The following Lemma is from Psaradakis (2001) and states the properties of the sieve bootstrap on the process $\{u_t^*\}_{t\in\mathbb{Z}}$, induced by the autoregressive approximation.

Lemma 2.1. *Let $\{x_t\}_{t\in\mathbb{N}}$ satisfy Model (2.1) with $\rho = 1$ and suppose that the condition (C1)-(C4) hold. Then, as $n \to \infty$,*

(a) there exists a random variable n_0 such that $\sup_{n\geq n_0} \sum_{j=0}^{\infty} j|\hat{\psi}_{j,n}| < \infty$ a.s.,

(b) $\sup_{0\leq j<\infty} |\hat{\psi}_{j,n} - \psi_j| = o(1)$ a.s.,

(c) $var^*[u^*_{t,n}] - \sigma_u^2 = o_p(1)$, where $\sigma_u^2 = \sigma \sum_{j=0}^{\infty} \psi_j^2$, and

(d) $var^*[n^{-1/2} \sum_{t=1}^{n} u^*_{t,n}] - \sigma^2 = o_p(1)$.

These results are helpful in establishing a weak invariance principle for the sequence of partial sums of $\{u^*_t\}_{t\in\mathbb{Z}}$. Assuming, for convenience, that $n/s =: N \in \mathbb{N}$, define s stochastic processes $\mathbb{U}^*_{N,m} = \{\mathbb{U}^*_{N,m}(r) : r \in [0,1]\}$, $m = 1, 2, ..., s$, with trajectories in the càdlàg-space $D[0,1]$ by defining

$$\mathbb{U}^*_{N,m}(r) = \begin{cases} 0 & : r \in [0, N^{-1}) \\ N^{-1/2} \sum_{i=1}^{\lfloor Nr \rfloor} \upsilon^*_{m,i} & : r \in [N^{-1}, 1] \end{cases}$$

where $\lfloor \cdot \rfloor$ indicates the floor function, and $\upsilon^*_{m,i} := u^*_{s(i-1)+m,n}$. The asymptotic behavior of $\{\mathbb{U}^*_{N,m}\}$ is presented in the following Lemma.

Lemma 2.2. *If the conditions of Lemma 1 hold, then for $m = 1, ..., s$,*
$\mathscr{L}^*[\mathbb{U}^*_{N,m}] \xrightarrow{w} \mathscr{L}[\sigma \sum_{j=0}^{s-1} \psi_j(1)\mathbb{W}_{m,j}]$ in probability in $D[0,1]$.

Proof. Following the proof of Lemma 3 in Psaradakis (2001), it needs to be shown that $\ell(\mathscr{L}^*[\mathbb{U}^*_{N,m}], \mathscr{L}[\sigma \sum_{j=0}^{s-1} \psi_j(1)\mathbb{W}_{m,j}]) = o_p(1)$ as $n \to \infty$, for any metric ℓ metrizing weak convergence of laws on $D[0,1]$. Let $\{\mathbb{U}^*_{N_k,m}\}$ be any subsequence of $\{\mathbb{U}^*_{N,m}\}$. Then it is sufficient to show that there exists a further subsequence $\{\mathbb{U}^*_{N_{k(i)},m}\}$ such that

$$\ell(\mathscr{L}^*[\mathbb{U}^*_{N_{k(i)},m}], \mathscr{L}[\sigma \sum_{j=0}^{s-1} \psi_j(1)\mathbb{W}_{m,j}]) = o(1) a.s. \qquad \text{for } i \to \infty. \tag{2.5}$$

Since $\{\epsilon^*_{t,n}\}$ is an *i.i.d.* sample from $\hat{F}_{\tilde{\epsilon},n}$, with $E[\epsilon^*_{t,n}] = 0$, Lemma 2.1 (a) implies that any subsequence $\{N_k\}$ of $\mathbb{N}$ contains a further subsequence $\{N_{k(i)}\}$ along which the conditions

of Proposition 3.1 in Rupasinghe and Samaranayake (2012d) are satisfied. Therefore, we have

$$u^*_{m,n} + u^*_{s+m,n} + ... + u^*_{s(N_{k(i)}-1)+m,n} = \sum_{j=0}^{s-1}[\hat{\psi}_{j,n}(1) \sum_{l=1}^{N_{k(i)}} \epsilon^*_{s(l-1)+m-j,n}] + \eta^*_{N_{k(i)},m,n} - \eta^*_{0,m,n}.$$

By utilizing the arguments followed by Proposition 17.2 in Hamilton (1994) and Lemma 5.3 in Bühlmann (1997), it is possible to establish Equation (2.5), which completes the proof. □

The following Lemma is analogous to Lemma A1 of Psaradakis (2001) but incorporates seasonal periodicity.

Lemma 2.3. *Let* $\chi^*_{0,m,n} := 0$ *and* $\chi^*_{i,m,n} := \sum_{j=1}^{i} \upsilon^*_{m,j}$ *for* $m = 1, ..., s$. *Then, if the conditions of Lemma 1 hold, as* $n \to \infty$,

(a) $\mathscr{L}^*[n^{-3/2} \sum_{m=1}^{s} \sum_{k=1}^{N} \chi^*_{k-1,m,n}] \xrightarrow{w} \mathscr{L}[s^{-3/2}\sigma \sum_{j=0}^{s-1} \psi_j(1) \sum_{m=1}^{s} \int_0^1 \mathbb{W}_{m,j}(r)dr]$ *in probability,*

(b) $\mathscr{L}^*[n^{-2} \sum_{m=1}^{s} \sum_{k=1}^{N} \chi^{*2}_{k-1,m,n}] \xrightarrow{w} \mathscr{L}[s^{-2}\sigma^2 \sum_{j=0}^{s-1} \psi^2_j(1) \sum_{m=1}^{s} \int_0^1 \mathbb{W}^2_{m,j}(r)dr]$ *in probability,*

(c) $\mathscr{L}^*[n^{-5/2} \sum_{m=1}^{s} \sum_{k=1}^{N} t\chi^*_{k-1,m,n}] \xrightarrow{w} \mathscr{L}[s^{-5/2}\sigma \sum_{j=0}^{s-1} \psi_j(1) \sum_{m=1}^{s} \int_0^1 r\mathbb{W}_{m,j}(r)dr]$ *in probability,*

(d) $\mathscr{L}^*[n^{-1} \sum_{m=1}^{s} \sum_{k=1}^{N} \chi^*_{k-1,m,n}\upsilon^*_{m,k}] \xrightarrow{w} \mathscr{L}[s^{-1}\sigma^2\{\sum_{j=0}^{s-1} \psi^2_j(1) \sum_{m=1}^{s} \mathbb{W}^2_{m,j}(1) - \sum_{j=0}^{\infty} \psi^2_j\}]$ *in probability, and*

(e) $\mathscr{L}^*[n^{-3/2} \sum_{m=1}^{s} \sum_{k=1}^{N} [s(k-1)+m]\upsilon^*_{m,k}] \xrightarrow{w} \mathscr{L}[s^{-3/2}\sigma \sum_{j=0}^{s-1} \psi_j(1) \sum_{m=1}^{s} \int_0^1 rd\mathbb{W}_{m,j}(r)]$ *in probability;*

Proof. Following the same analysis as in the proof of Lemma A.1 in Psaradakis (2001) but adjusting for the seasonality, we observe that

$$\mathbb{U}^*_{N,m}(r) = N^{-1/2}\chi^*_{\lfloor Nr \rfloor,m,n} = N^{-1/2}\chi^*_{t-1,m,n}, \qquad \frac{t-1}{N} \leq r < \frac{t}{N} \text{ for } t = 1, 2, ..., N,$$

and

$$\int_{(t-1)/N}^{t/n} \mathbb{U}^{*b}_{N,m}(r)dr = N^{-1}\mathbb{U}^{*b}_{N,m}\left(\frac{t-1}{N}\right) \text{ for } b = 1, 2.$$

Consequently,

$$\begin{aligned}
n^{-3/2}\sum_{m=1}^{s}\sum_{k=1}^{N}\chi^*_{k-1,m,n} &= s^{-3/2}\sum_{m=1}^{s}\{N^{-1}\sum_{k=1}^{N}\mathbb{U}^*_{N,m}\left(\frac{k-1}{N}\right)\} \\
&= s^{-3/2}\sum_{m=1}^{s}\int_0^1 \mathbb{U}^*_{N,m}(r)dr; \\
n^{-2}\sum_{m=1}^{s}\sum_{k=1}^{N}\chi^{*2}_{k-1,m,n} &= s^{-2}\sum_{m=1}^{s}\{n^{-1}\sum_{k=1}^{N}\mathbb{U}^{*2}_{N,m}\left(\frac{k-1}{N}\right)\} \\
&= s^{-2}\sum_{m=1}^{s}\int_0^1 \mathbb{U}^{*2}_{N,m}(r)dr; \\
n^{-5/2}\sum_{m=1}^{s}\sum_{k=1}^{N}[(s(k-1)+m]\chi^*_{k-1,m,n} &= s^{-5/2}\sum_{m=1}^{s}\left\{N^{-1}\sum_{k=1}^{N}\left(\frac{t}{N}\right)\mathbb{U}^*_{N,m}\left(\frac{k-1}{N}\right)\right\} \\
&= s^{-5/2}\sum_{m=1}^{s}\int_0^1 \left(\frac{1+\lfloor Nr \rfloor}{N}\right)\mathbb{U}^*_{N,m}(r)dr; \\
n^{-1}\sum_{m=1}^{s}\sum_{k=1}^{N}\chi^*_{k-1,m,n}\upsilon^*_{m,k} &= s^{-1}\sum_{m=1}^{s}\left\{(2N)^{-1}\chi^*_{N,m,n} - (2N)^{-1}\sum_{k=1}^{N}\upsilon^{*2}_{m,k}\right\} \\
&= s^{-1}\sum_{m=1}^{s}\left\{\frac{1}{2}\mathbb{U}^{*2}_{N,m}(1) - (2N)^{-1}\sum_{k=1}^{N}\upsilon^{*2}_{m,k}\right\}; \\
n^{-3/2}\sum_{m=1}^{s}\sum_{k=1}^{N}[(s(k-1)+m]\upsilon^*_{m,k} &= s^{-3/2}\sum_{m=1}^{s}\left\{N^{-1/2}\chi^*_{N,m,n} - N^{-3/2}\sum_{k=1}^{N}\chi^*_{k-1,m,n}\right\} \\
&= s^{-3/2}\sum_{m=1}^{s}\left\{\mathbb{U}^*_{N,m}(1) - \int_0^1 \mathbb{U}^*_{N,m}(r)dr\right\}.
\end{aligned}$$

The desired results are obtained using the fact that $\{\upsilon^{*2}_{m,k}\}$ obeys the bootstrap weak law of large numbers, the bootstrap weak invariance principle in Lemma 2.2 and the continuous mapping theorem. □

Using this bootstrap invariant property, the large sample consistency of the bootstrap distribution of the difference-based DHF test statistics is established below.

Theorem 2.4. *If the conditions C1 through C4 hold, then, as* $n \to \infty$,

$$\mathscr{L}[K_n^{*D}] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{s\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s^2/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]}\right]$$

and

$$\mathscr{L}[T_n^{*D}] \xrightarrow{w}$$
$$\mathscr{L}\left[\frac{\sum_{j=0}^{s-1}\sum_{m=1}^{s}\psi_j^2(1)\int_0^1 \mathbb{W}_{m,j}(r)d\mathbb{W}_{m,j}(r) + (s/2)(\sum_{j=0}^{s-1}\psi_j^2(1) - \sum_{j=0}^{\infty}\psi_j^2)}{\left\{\sum_{j=0}^{s-1}\left[\psi_j^2(1)\sum_{m=1}^{s}\int_0^1 \mathbb{W}_{m,j}^2(r)dr\right]\right\}^{-1/2}}\right].$$

Proof. The proof immediately follows from Lemma 2.3. □

Note that the asymptotic properties of the residual-based DHF tests statistics can also be established with minor modifications to the results of the difference-based DHF test statistics.

2.4. SIMULATION STUDY

To investigate the finite sample performances of the proposed difference-based and residual based DHF seasonal unit roots tests, different combinations of the values of ρ, ϕ and θ in the following model were considered:

$$(1 - \rho B^4)(1 - \phi B)x_t = (1 + \theta B)\epsilon_t. \tag{2.6}$$

The sets of values studied are: $\rho \in \{1, 0.95, 0.9, 0.8\}$, $\phi \in \{0, 0.4, 0.8\}$ and $\theta \in \{0, -0.5, -0.9\}$. We also looked at the performances of the tests at significance levels $\gamma = 0.05$ and 0.1 under standard normal and exponential (with mean 1 but centered at

zero) errors $\{\epsilon_t\}$. Sample sizes of 80 and 200 were employed in the study. These two sample sizes are chosen because of the availability of the DHF critical values. Using each combination of parameters, error distributions, sample sizes and significance levels, 1,000 time series were generated and the frequency of rejections of the null hypothesis $H_0 : \rho = 1$ by the regular DHF, difference-based bootstrap DHF and residual-based bootstrap DHF tests were recorded to study the power and the Type I error of each test. For the regular DHF test, the critical values were obtained from Tables 2 and 3 in Dickey, Hasza and Fuller (1984), and for the bootstrap tests, critical values were computed using 1,000 bootstrap replicates of K_n^{*D}, T_n^{*D}, K_n^{*R} and T_n^{*R}.

From Table 1 through 3, we can see that the powers of all the test statistics increase with the sample size. It is interesting to observe that the coefficient test statistics of both difference-based and residual-based bootstrap methods (K_n^{*D} and K_n^{*R}) yield the highest powers in each case. The regular DHF t-statistic, however, has the highest competitive powers when $|\rho| < 1$ only for sample size of 200 for the model presented in Table 1. Note that this model has a moving average close to unity and the sieve bootstrap based coefficient test statistics have the greatest power for the sample size of 80 than that of the regular DHF tests.

Table 1. Empirical rejection frequencies of $H_0 : \rho = 1$ for the model, $(1 - \rho B^4)x_t = (1 - 0.9B)\epsilon_t$ with normal errors

		ρ				ρ			
n		1	0.95	0.9	0.8	1	0.95	0.9	0.8
		$\gamma = 0.05$				$\gamma = 0.10$			
80	K_n	0.03	0.23	0.55	0.94	0.08	0.45	0.81	0.99
	T_n	0.03	0.47	0.64	0.94	0.08	0.62	0.87	0.97
	K_n^{*D}	0.04	0.61	0.9	0.99	0.09	0.78	0.95	1
	T_n^{*D}	0.03	0.4	0.54	0.92	0.08	0.56	0.8	0.97
	K_n^{*R}	0.04	0.61	0.91	1	0.1	0.76	0.95	1
	T_n^{*R}	0.03	0.39	0.56	0.9	0.09	0.54	0.77	0.97
200	K_n	0.07	0.74	0.98	1	0.13	0.84	1	1
	T_n	0.08	0.85	0.99	1	0.17	0.9	1	1
	K_n^{*D}	0.07	0.82	0.98	1	0.16	0.9	0.99	1
	T_n^{*D}	0.05	0.63	0.94	1	0.12	0.85	0.97	1
	K_n^{*R}	0.07	0.84	0.98	1	0.16	0.91	0.99	1
	T_n^{*R}	0.05	0.67	0.93	1	0.14	0.86	0.97	1

The sieve bootstrap based coefficient tests are uniformly most powerful for the model presented in Table 2. It is interesting to see that the sieve bootstrap based t-statistics and the regular DHF t-statistic have very close rejection frequencies when $|\rho| < 1$.

Table 2. Empirical rejection frequencies of $H_0 : \rho = 1$ for the model, $(1 - \rho B^4)(1 - 0.4B)x_t = (1 - 0.5B)\epsilon_t$ with normal errors

		ρ				ρ			
n		1	0.95	0.9	0.8	1	0.95	0.9	0.8
		$\gamma = 0.05$				$\gamma = 0.10$			
80	K_n	0.04	0.17	0.39	0.97	0.1	0.31	0.76	0.99
	T_n	0.04	0.33	0.63	0.95	0.09	0.47	0.84	0.99
	K_n^{*D}	0.05	0.67	0.91	1	0.12	0.77	0.95	1
	T_n^{*D}	0.05	0.35	0.65	0.94	0.13	0.52	0.83	0.98
	K_n^{*R}	0.08	0.66	0.92	1	0.15	0.79	0.97	1
	T_n^{*R}	0.06	0.33	0.65	0.96	0.12	0.49	0.84	0.97
200	K_n	0.07	0.71	0.99	1	0.12	0.9	1	1
	T_n	0.07	0.8	1	1	0.13	0.92	1	1
	K_n^{*D}	0.09	0.95	1	1	0.2	0.99	1	1
	T_n^{*D}	0.07	0.81	1	1	0.15	0.92	1	1
	K_n^{*R}	0.08	0.94	1	1	0.22	0.99	1	1
	T_n^{*R}	0.08	0.81	1	1	0.17	0.92	1	1

Table 3 represents the performances of the regular DHF and sieve bootstrap based tests under exponential errors. Our sieve bootstrap based coefficient test statistics are most powerful in the cases considered, and this was the case with all the other models as well.

However, it should be noted that the bootstrap test statistics have more size distortions compare to the asymptotic test statistics. Overall, the sieve test bootstrap statistics are more powerful than the DHF test statistics with some size distortions.

Table 3. Empirical rejection frequencies of $H_0 : \rho = 1$ for the model, $(1 - \rho B^4)(1 - 0.5B)x_t = \epsilon_t$ with exponential errors

		ρ				ρ			
n		1	0.95	0.9	0.8	1	0.95	0.9	0.8
		$\gamma = 0.05$				$\gamma = 0.10$			
80	K_n	0.04	0.14	0.51	0.91	0.09	0.22	0.72	0.95
	T_n	0.05	0.3	0.65	0.93	0.1	0.46	0.8	0.95
	K_n^{*D}	0.04	0.58	0.88	0.98	0.13	0.79	0.91	1
	T_n^{*D}	0.03	0.24	0.53	0.89	0.11	0.37	0.75	0.94
	K_n^{*R}	0.06	0.61	0.88	0.98	0.13	0.79	0.91	1
	T_n^{*R}	0.04	0.24	0.54	0.91	0.09	0.42	0.77	0.93
200	K_n	0.05	0.63	0.96	1	0.08	0.79	0.97	1
	T_n	0.05	0.76	0.97	1	0.07	0.9	0.97	1
	K_n^{*D}	0.06	0.91	0.97	1	0.08	0.97	0.97	1
	T_n^{*D}	0.06	0.67	0.96	1	0.06	0.87	0.97	1
	K_n^{*R}	0.06	0.93	0.97	1	0.07	0.96	0.97	1
	T_n^{*R}	0.05	0.69	0.95	1	0.06	0.84	0.97	1

3. PREDICTION INTERVALS

In this section, a sieve bootstrap method to obtain prediction intervals for the seasonal time series defined in Model (1.1) with $|\rho| < 1$ or $\rho = 1$ is discussed. The following procedure is identical to the that of Rupasinghe and Samaranayake (2012c) except for the seasonal differencing. As mentioned before, the proposed method is applicable to both seasonally integrated ($\rho = 1$ in Equation (1.1)) and non-integrated processes without altering the computational steps.

Assume that a realization $\{x_t\}_{t=1}^n$ is obtained from Model (1.1) regardless of $\rho = 1$ or $|\rho| < 1$. Define the differenced series, $\{y_t\}$, using $y_t = x_t - x_{t-s}$. Observe that $\{y_t\}$ is $ARMA$ if $\{x_t\}$ is seasonally integrated. Then, the methods proposed in Alonso et *al.* (2001, 2003, 2004) can be directly applied to $\{y_t\}$ and estimate the distribution of y_{n+h}. Complications arise, however, if $\{x_t\}$ has no seasonal unit root. In this case $\{y_t\}$ becomes non-invertible, and this violates the assumption necessary for the application of the existing methodology. This drawback can be overcome by using the following modified sieve bootstrap procedure.

1. Select the order $p = p(n)$ of the autoregressive approximation from among models with $p \in \{1, 2, ..., M_n\}$ with $M_n = o\{[log(n)/n]^{1/2}\}$ by the AIC criterion. Alonso et *al.* (2003) preferred AICC over AIC and used $M_n = o\{[log(n)/n]^{1/4}\}$.

2. Estimate the autoregressive coefficients, $\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n}$, of the $AR(p)$ approximation, $\sum_{j=0}^{p} \phi_{j,p} y_{t-j} = \epsilon_{t,p}$, by the Yule-Walker method.

3. Obtain the $(n-p)$ residuals: $\hat{\epsilon}_{t,n} = \sum_{j=0}^{p} \hat{\phi}_{j,p,n}(y_{t-j} - \bar{y})$, $t = p+1, ..., n$ and define the empirical distribution function of the centered residuals, $\tilde{\epsilon}_t = \hat{\epsilon}_{t,n} - \hat{\epsilon}^{(\cdot)}$, where $\hat{\epsilon}^{(\cdot)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\epsilon}_{t,n}$, by $\hat{F}_{\tilde{\epsilon},n}(x) = (n-p)^{-1} \sum_{t=p+1}^{n} I_{[\tilde{\epsilon}_t \leq x]}$.

4. Draw a resample $\epsilon^*_{t,n}$, $t = p+1, ..., n$ of i.i.d. observations from $\hat{F}_{\tilde{\epsilon},n}$.

5. Obtain y^*_t by the recursion: $\sum_{j=0}^{p} \hat{\phi}_{j,p,n}(y^*_{t-j} - \bar{y}) = \epsilon^*_{t,n}$ for $t = p+1, ..., n$ and set $y^*_t = \bar{y}$ for $t = 1, ..., p$.

6. Compute the estimates $\hat{\phi}^*_{1,p,n}, ..., \hat{\phi}^*_{p,p,n}$ as in Step 2, using $\{y^*_t\}_{t=1}^n$.

7. For $h > 0$, compute the future bootstrap observations of the differenced series by the recursion. $y^*_{n+h} - \bar{y} = \sum_{j=1}^{p} \hat{\phi}^*_{j,p,n}(y^*_{n+h-j} - \bar{y}) + \epsilon^*_{n+h,n}$, where $y^*_t = y_t$ for $t \leq n$.

   The following step is a modification to the method proposed by Rupasinghe and Samaranayake (2012c) adopted to accommodate seasonal differencing.

8. Compute the future bootstrap observations of the original series by the recursion. $x^*_{n+h} = x^*_{n+h-s} + y^*_{n+h}$ where, $x^*_t = x_t, t \leq n$, $h > 0$.

9. Obtain a Monte Carlo estimate of the bootstrapped distribution function of x^*_{n+h} by repeating steps 4-8 B times.

10. Use the bootstrapped distribution to approximate the unknown distribution of x_{n+h} given the observed sample.

11. The $100(1-\alpha)\%$ prediction interval for x_{n+h} is given by $\{Q^*(\frac{\alpha}{2}), Q^*(1-\frac{\alpha}{2})\}$ where, $Q^*(.)$ are the quantiles of the estimated bootstrap distribution.

3.1. ASYMPTOTIC RESULTS

It is clear that if the original process $\{x_t\}$ is seasonally integrated, then the differenced series $\{y_t\}$ is $ARMA(p,q)$ and that otherwise it is not invertible. Following Rupasinghe and Samaranayake (2012c), the assumptions below are necessary to establish the asymptotic results irrespective of the true value of ρ. Here onwards we assume that the negative lag values of $\{x_t\}$ and $\{y_t\}$ are zero for $t \leq 0$.

A1: Let ξ_t denote the σ-algebra of events determined by ϵ_s, $s \leq t$. Also, assume ϵ_t are $i.i.d.$ and that

$$E[\epsilon_t|\xi_{t-1}] = 0 \qquad and \qquad E[\epsilon_t^2|, t \in \mathbb{Z}.\xi_{t-1}] = \sigma^2.$$

Furthermore, assume $E[\epsilon_t^4] < \infty$ for $t \in \mathbb{Z}$.

A2: The series y_t has infinite moving average representation $y_t = \sum_{j\geq 0} \psi_j \epsilon_{t-j}$ with $\sum_{j\geq 0} |\psi_j^2| < \infty$.

B: Let $p(n) = o\{[n/log(n)]^{1/2}\}$ and $\hat{\boldsymbol{\phi}}_{p,n} = (\hat{\phi}_{1,p,n}, ..., \hat{\phi}_{p,p,n})'$ satisfy the empirical Yule-Walker equations $\hat{\boldsymbol{\Gamma}}_{p,n}\hat{\boldsymbol{\phi}}_{p,n} = -\hat{\boldsymbol{\gamma}}_{p,n}$, where
$\hat{\boldsymbol{\Gamma}}_{p,n} = [\hat{R}(i-j)]_{i,j=1}^{p}, \hat{\boldsymbol{\gamma}}_{p,n} = (\hat{R}(1), ..., \hat{R}(p))'$, and $\hat{R}(j) = n^{-1}\sum_{t=1}^{n-|j|}(y_t - \bar{y})(y_{t+|j|} - \bar{y})$ for $|j| < n$.

Note that Assumption A2 is satisfied by the time series defined in Model (1.1) under both $\rho = 1$ and $|\rho < 1$. The following theorem holds since the process $\{y_t\}$ is assumed to have the same conditions as Theorem 3.1 in Rupasinghe and Samaranayake (2012c).

Theorem 3.1. *Assume that A1, A2 and B hold. Then, in probability, as* $n \to \infty$,

$$y^*_{n+h} \xrightarrow{d^*} y_{n+h}, \text{ for fixed } h \in \mathbb{N}$$

Using Theorem 3.1, we can establish the large sample validity of the proposed sieve bootstrap prediction intervals.

Theorem 3.2. *Assume that A1, A2 and B hold. Then, in probability, as* $n \to \infty$,

$$x^*_{n+h} \xrightarrow{d^*} x_{n+h}, \text{ for } h = 0, 1, ... \tag{3.1}$$

Proof. The future values of the originally observed time series, $\{x_{n+h}\}$ can be written as $x_{n+h} = x_{n+h-s} + y_{n+h}$. For $h \leq s$, we have
$x^*_{n+h} = x_{n+h-s} + y^*_{n+h} \xrightarrow{d^*} x_{n+h-s} + y_{n+h} = x_{n+h}$.
Now, for $h = s+1$, $x^*_{n+h} = x^*_{n+(s+1)-s} + y^*_{n+h} \xrightarrow{d^*} x_{n+1} + y_{n+h} = x_{n+h}$. For $h > s+1$, the result can be proven using the mathematical induction.

□

3.2. SIMULATION STUDY

The finite sample performances of the method proposed in this paper were investigated using a Monte-Carlo simulation study. A sequence of time series was simulated from the model, $(1 - \rho B^4)(1 - \phi B)x_t = (1 + \theta B)\epsilon_t$, with different combinations of the parameters, error distributions and sample sizes. The values considered for the parameters are: $\rho \in \{0.6, 1\}$, $\phi \in \{0, 0.7\}$ and $\theta \in \{0, -0.3\}$. The standard normal distribution, exponential (1) centered at zero and $t(3)$ distributed errors are chosen for sample sizes 100 and 200. The coverage, bootstrap length, and the length of the interval theoretically achievable under known order and parameter values were computed for 95% and 99% prediction intervals to asses the performance of the two methods.

For each combination of model, sample size, nominal coverage and error distribution, $N = 1,000$ independent series were generated and for each of these simulated series, steps 1 to 11 were implemented. To compute the coverage probabilities for each of this N

simulations, $R = 1,000$ future observations (x_{n+h}) were generated using the original model.

The proportion of those falling in between the lower and upper bounds of the bootstrap prediction interval was then defined to be the coverage. Thus, the coverage at the i^{th} simulation run is given by $C(i) = R^{-1}\sum_{r=1}^{R} I_A[x_{n+h}^r(i)]$ where $A = [Q^*(\alpha/2), Q^*(1-\alpha/2)]$, $I_A(.)$ is the indicator function of the set A and $x_{n+h}^r(i)$, $r = 1, 2, ...1,000$ are the R future values generated at the i^{th} simulation run. The bootstrap length and theoretical length for the i^{th} simulation run are given by $L_B(i) = Q^*(1-\alpha/2) - Q^*(\alpha/2)$ and $L_T(i) = x_{n+h}^r(1-\alpha/2) - x_{n+h}^r(\alpha/2)$ respectively.. The theoretical length $L_T(i)$ is the difference between the $100(1-\alpha/2)^{th}$ and $100(\alpha/2)^{th}$ percentile points the empirical distribution of the 1,000 future observations that were generated using the underlying time series model with known order and the true values of the coefficients. Using these statistics, the mean coverage, mean length of bootstrap prediction intervals, mean length of theoretical intervals, and their standard errors were computed as:

Mean Coverage $\bar{C} = N^{-1}\sum_{i=1}^{N} C(i)$

Standard Error of Mean Coverage $SE_{\bar{C}} = \{[N(N-1)]^{-1}\sum_{i=1}^{N}[C(i) - \bar{C}]^2\}^{1/2}$

Mean Length (bootstrap) $\bar{L}_B = N^{-1}\sum_{i=1}^{N} L_B(i)$

Standard Error of Mean Length $SE_{\bar{L}_B} = \{[N(N-1)]^{-1}\sum_{i=1}^{N}[L_B(i) - \bar{L}_B]^2\}^{1/2}$

Mean Theoretical Length $\bar{L}_B = N^{-1}\sum_{i=1}^{N} L_B(i)$

In total 96 different combinations of model type, sample size, nominal coverage probability, and error distributions were investigated in this simulation study. However, due to space limitations, we report only a representative sample of results for 95% intervals, in Tables 4 through 7. These tables report the mean coverage, mean interval length, and mean theoretical length, standard error of mean coverage and standard error of mean interval length. The complete results of the simulation study are available upon request from the corresponding author.

The representative models are labelled as follows:

$M1$: $(1 - 0.7B)x_t = \epsilon_t$ (non-integrated)

$IM1$: $(1 - B^4)(1 - 0.7B)x_t = \epsilon_t$ (integrated)

$M2$: $(1 - 0.7B)x_t = (1 - 0.3B)\epsilon_t$ (non-integrated)

$IM2$: $(1 - B^4)(1 - 0.7B)x_t = (1 - 0.3B)\epsilon_t$ (integrated)

$M3$: $x_t = \epsilon_t$ (non-integrated)

$IM3$: $(1 - B^4)x_t = \epsilon_t$ (integrated)

To investigate the behaviour of the intervals for each of the 96 combinations, the minimum value, percentiles (25^{th}, 50^{th}, and 75^{th}), and the maximum value of (a) the coverage probabilities, (b) the bootstrap interval bounds (upper and lower), and (c) the theoretical interval bounds (upper and lower), were further computed, based on the 1,000 values generated through simulation, and these statistics are also available upon request.

The coverage probabilities of the sieve bootstrap prediction intervals get closer to the nominal level as the sample size increase for both seasonally integrated and non-integrated time series. Large samples help to obtain more residuals and yield accurate parameter estimates. Thus, this phenomenon is expected. Furthermore, the mean coverages of the proposed sieve bootstrap method are very close to the nominal coverage for all the leads regardless of the presence or absence of a unit root and of the nature their error distributions. Only for the Model $IM3$ with t-distributed errors (Table 6), however, our procedure yields slightly liberal coverages for the small sample sizes. However, it is seen that the mean bootstrap interval lengths are much close to the theoretical lengths in all the situations.

Table 4. Coverage of 95% intervals for Models $M1$ & $IM1$ with normal errors

Leads	Size	$M1$			$IM1$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9009	0.9428 (0.0040)	4.3907 (0.0384)	3.9318	0.9519 (0.0021)	4.1224 (0.0415)
	200	3.9116	0.9451 (0.0039)	4.3411 (0.0291)	3.9234	0.9531 (0.0017)	4.0502 (0.0299)
2	100	4.7719	0.9483 (0.0057)	5.1595 (0.0479)	4.7988	0.9449 (0.0024)	4.9080 (0.0517)
	200	4.7742	0.9490 (0.0036)	5.2026 (0.0342)	4.8001	0.9471 (0.0020)	4.8539 (0.0383)
3	100	5.1522	0.9462 (0.0044)	5.5051 (0.0581)	5.1785	0.9418 (0.0027)	5.2404 (0.0601)
	200	5.1553	0.9464 (0.0047)	5.5513 (0.0437)	5.1681	0.9459 (0.0021)	5.2010 (0.0440)

Table 5. Coverage of 95% intervals for Models $M2$ & $IM2$ with normal errors

Leads	Size	$M2$			$IM2$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.9226	0.9484 (0.0040)	4.4733 (0.0389)	3.9174	0.9471 (0.0025)	4.1413 (0.0392)
	200	3.9118	0.9410 (0.0052)	4.3499 (0.0316)	3.9190	0.9452 (0.0021)	4.0331 (0.0267)
2	100	4.2324	0.9610 (0.0040)	5.3116 (0.0562)	4.1984	0.9693 (0.0019)	4.9256 (0.0472)
	200	4.2357	0.9579 (0.0057)	5.2374 (0.0374)	4.2201	0.9693 (0.0014)	4.8169 (0.0337)
3	100	4.3564	0.9584 (0.0047)	5.5997 (0.0656)	4.3658	0.9718 (0.0017)	5.2172 (0.0527)
	200	4.3597	0.9628 (0.0038)	5.5654 (0.0436)	4.3506	0.9747 (0.0014)	5.1779 (0.0425)

Table 6. Coverage of 95% intervals for Models $M3$ & $IM3$ with t-dist errors

Leads	Size	$M3$			$IM3$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	6.4154	0.9446 (0.0038)	7.4010 (0.1622)	6.3989	0.9354 (0.0024)	6.3541 (0.1223)
	200	6.3542	0.9474 (0.0023)	7.0369 (0.1036)	6.3906	0.9416 (0.0020)	6.3745 (0.0854)
2	100	6.4156	0.9449 (0.0042)	7.4803 (0.1742)	6.3615	0.9388 (0.0024)	6.4068 (0.1256)
	200	6.3971	0.9455 (0.0066)	7.1497 (0.1114)	6.3637	0.9415 (0.0019)	6.2822 (0.0748)
3	100	6.4098	0.9467 (0.0036)	7.5347 (0.1721)	6.3793	0.9392 (0.0025)	6.4577 (0.1237)
	200	6.4130	0.9385 (0.0062)	7.2097 (0.1099)	6.4048	0.9422 (0.0019)	6.3646 (0.0850)

Table 7. Coverage of 95% intervals for Models $M2$ & $IM2$ with exp errors

Leads	Size	$M2$			$IM2$		
		Theo. Length	Coverage Mean (SE)	Length Mean (SE)	Theo. Length	Coverage Mean (SE)	Length Mean (SE)
1	100	3.6588	0.9486 (0.0085)	4.6587 (0.0779)	3.6628	0.9540 (0.0054)	4.1082 (0.0655)
	200	3.6869	0.9441 (0.0085)	4.3682 (0.0546)	3.6430	0.9388 (0.0097)	3.9832 (0.0486)
2	100	4.0538	0.9614 (0.0068)	5.4505 (0.0917)	4.0700	0.9753 (0.0018)	4.8933 (0.0776)
	200	4.0306	0.9559 (0.0070)	5.1899 (0.0606)	4.0280	0.9768 (0.0017)	4.8270 (0.0558)
3	100	4.2169	0.9604 (0.0074)	5.7912 (0.1013)	4.2228	0.9776 (0.0016)	5.2121 (0.0830)
	200	4.2358	0.9622 (0.0058)	5.5229 (0.0651)	4.1880	0.9804 (0.0012)	5.1613 (0.0578)

4. CONCLUSION

In this paper, sieve bootstrap procedures for testing seasonal unit roots and building prediction intervals for seasonal time series were proposed. The large sample validities of the proposed methods were then established

We proposed two sieve bootstrap versions (difference-based and residual based) of the Dickey-Hasza-Fuller seasonal unit root test. The bootstrap test statistics, however, have size distortions in some cases. Nevertheless, the simulation study indicate that the sieve bootstrap tests are more powerful than the asymptotic DHF test in most of the cases.

The proposed sieve bootstrap based prediction intervals perform fairly accurately in all the cases, and the coverages get closer to the nominal level as the sample size increase. Moreover, the bootstrap prediction interval lengths are very close to the theoretical lengths. Interestingly, our method can be applied to both seasonally integrated and non-integrated processes without altering the procedure. Overall, the sieve bootstrap perform accurately in unit root testing and in prediction intervals for seasonal time series.

5. REFERENCES

[1] Alonso A. M., Pena D., and Romo J. (2002), "Forecasting time series with sieve bootstrap," *Journal of Statistical Planning and Inference*, 100, 1-11.

[2] Alonso A. M., Pena D., and Romo J. (2003), "On sieve bootstrap prediction intervals," *Statistics and Probability Letters*, 65, 13-20.

[3] Alonso A. M., Pena D., and Romo J. (2004), "Introducing model uncertainty in time series bootstrap," *Statistica Sinica*, 14, 155-174.

[4] Box G., and Jenkins G. (1970), "Time Series Analysis: Forecasting & Control," Prentice Hall.

[5] Bühlmann P. (1995), "Moving-average representation of autoregressive approximations," *Stochastic Process. Appl.*, 60, 331342.

[6] Bühlmann P. (1997), "Sieve bootstrap for time series," *Bernoulli*, 3, 123-148.

[7] Dickey D.A., and Fuller W.A. (1979), "Distribution of estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427 - 431.

[8] Dickey D.A., Hasza D.P., and Fuller W.A. (1984), "Testing for unit roots in seasonal time series," *Journal of the American Statistical Association*, 79, 355 - 367.

[9] Ghysels E., and Osborn D.R (2001), "The econometric analysis of seasonal time series," New York: Cambridge University Press.

[10] Ghysels E., Osborn D.R, Rodrigues P.M.M. (2000), "Seasonal Nonstationarity and Near-Nonstationarity," in: B.H. Baltagi (Ed.), A Companion to Theoretical Econometrics, Blackwell Publishers, Oxford.

[11] Hamilton J. D. (1994), "Time Series Analysis," Princeton: Princeton University Press.

[12] Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S. (1990), "Seasonal integration and cointegration," *Journal of Econometrics*, 44, 215-238.

[13] Kilian L. (1998a), "Small-sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics,* 80, 218-230.

[14] Kim J. (2001), "Bootstrap-After-Bootstrap Prediction Intervals for Autoregressive Models," *Journal of Business and Economic Statistics*, 19, 117-128.

[15] Kreiss, J.P. (1998), "Asymptotic Statistical Inference for a Class of Stochastic Processes," Habilitationsschrift, Universiti at Hamburg.

[16] Kreiss J.P. (1992), "Bootstrap procedures for AR(∞)-processes," In: JEockel, K.H., Rothe, G., Sendler, W. (Eds.), Bootstrapping and Related Techniques, Springer, Heidelberg, pp. 107-113.

[17] Mukhopadhyay P., and Samaranayake V.A. (2010), "Prediction intervals for time series: A Modified Sieve Bootstrap Approach," *Communications in Statistics - Simulation and Computation*, 39, 517-538.

[18] Palm F., Smeekes S., Urbain J. (2008), "Bootstrap unit-root tests: comparison and extension," *Journal of Time Series Analysis*, 29, 371-401.

[19] Panichkitkosolkul W.,and Niwitpong S. (2011), "On Multistep-Ahead Prediction Intervals Following Unit Root Tests for a Gaussian AR(1) Process with Additive Outliers," *Applied Mathematical Sciences*, 47, 2297-2316.

[20] Panichkitkosolkul W., and Niwitpong S. (2011), "Prediction intervals for the Gaussian autoregressive processes following the unit root tests," *Model Assisted Statistics and Applications*, 7, 1-15.

[21] Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for linear processes," *Annals of Statistics*, 20, 9711001.

[22] Poskitt D. S. (2006), "Autoregressive approximation in nonstandard situations: the fractionally integrated and non-invertible cases," *Annals of the Institute of Statistical Mathematics*, 59, 697-725

[23] Poskitt D. S. (2007), "Properties of the Sieve bootstrap for fractionally integrated and non-invertible processes," *Journal of Time Series Analysis*, 29, 224-250.

[24] Psaradakis Z. (2001), "Bootstrap test for an autoregressive unit root in the presence of weakly dependent errors," *Journal of Time Series Analysis*, 22, 577-594.

[25] Rupasinghe M., and Samaranayake V.A. (2012a), "Asymptomatic Properties of Sieve Bootstrap Prediction Intervals for FARIMA Processes," *Statistics and Probability Letters*, Submitted.

[26] Rupasinghe M., and Samaranayake V.A. (2012b), "Obtaining Prediction Intervals for FARIMA Processes using sieve bootstrap," Preprint.

[27] Rupasinghe M., and Samaranayake V.A. (2012c), "Prediction intervals for ARIMA processes: A sieve bootstrap approach," Preprint.

[28] Rupasinghe M., and Samaranayake V.A. (2012d), "The asymptotic distributions of the Dickey-Hasza-Fuller seasonal unit root tests under weakly dependent errors," Preprint.

[29] Said E., and Dickey D. (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika*, 71, 599607.

SECTION

2. CONCLUSION

A modified version of the currently available sieve bootstrap procedure is developed and applied to $ARIMA$ and $FARIMA$ process in order to obtain asymptotically accurate prediction intervals. In addition, the asymptotic distributions of the Dickey-Hasza-Fuller test statistics are derived under the assumption of weakly dependent innovations of the underlying seasonal model. Furthermore, a sieve-bootstrap based test statistics are derived for making inference about the presence of a seasonal unit root. These results were presented in five papers as indicated below.

In Papers I and II, a relatively simple bootstrap procedure for obtaining prediction intervals for $FARIMA$ processes was implemented. The asymptotic properties of the proposed method were established in Paper I, and the simulation results given in Paper II indicate that the proposed method performs competitively well compared to an existing method that was introduced by Bisaglia and Grigoletto (2001), and outperforms this method when the errors are skewed or bimodal.

In Paper III, sieve-bootstrap-based prediction intervals were proposed for $ARIMA(p, d, q)$ processes irrespective of whether $d = 0$ or 1. The advantage is the lack of the need to test for a unit root prior to deciding on a prediction method. We established the asymptotic properties of the sieve-bootstrap-based prediction intervals and our prediction intervals yield coverage close to the nominal level in all the cases as shown by a Monte Carlo simulation study.

The asymptotic distributions of the Dickey-Hasza-Fuller seasonal unit root tests were derived in Paper IV as functional of standard Brownian motions under weakly dependent errors. A key proposition was developed to establish these results.

In Paper V, we showed how to employ the sieve bootstrap to approximate the distributions of the Dickey-Hasza-Fuller seasonal unit root test statistics. The proposed

sieve bootstrap seasonal unit root tests (difference-based and residual-based) were more powerful than the asymptotic DHF tests under weakly dependent errors in most of the situations. The asymptotic properties of the difference-based sieve bootstrap DHF test were established. Note that the seasonal models considered in this study assumed zero mean and no trend. The asymptotic results, however, for these case can be easily established with minor adjustments. A procedure to obtain prediction intervals for seasonally integrated time series using sieve bootstrap was also proposed. The asymptotic properties of the proposed method were established and a Monte-Carlo simulation study was carried out to examine the finite sample validity.

APPENDIX A

MATLAB ALGORITHM FOR SIEVE BOOTSTRAP PREDICTION INTERVALS

```
function EDF = ESeivesBootstrap2
(sigmahat,XSeries,YWPhi,Residuals,PHat,MaxLag,B,XSize,PMax)

XMean = mean(XSeries);
WSeries = XSeries - XMean;
WResiduals = Residuals;
RRESVec = zeros(B,1);
PhatVec = zeros(B,1);
WYRESVec = zeros(B,1);
EDF = zeros(MaxLag,B);

meanWResiduals = mean(WResiduals);
WResiduals = (WResiduals - mean(WResiduals))/std(WResiduals);

for J = 1: B
      %Randomize residuals for resampling
           RandomIndex = unidrnd(XSize,XSize+200+MaxLag,1);
           RResiduals = WResiduals(RandomIndex);

%STEP 9

WSeries =RResiduals;

WSeries(1:PHat)= XSeries(unidrnd(XSize - PHat, PHat,1) + PHat);

for I = PHat +1: XSize +200
    for Ip =2 :PHat + 1
            WSeries(I) =WSeries(I) -YWPhi(Ip)*WSeries(I-(Ip-1));
    end
```

```
end

%Compute bootstrap AR coefficients
[WYWPhi,YWResiduals] = aryule(WSeries(201:XSize+200),PHatMod);
WW = zeros(XSize+PHatMod,1);
WW(1:PHatMod) = WSeries(XSize+200 -PHatMod+1:XSize +200);
WW(PHatMod+1:XSize+PHatMod) = WSeries(201:XSize+200);

for ri = PHatMod + 1:XSize + PHatMod
    YWResiduals(ri) = WW(ri) +
     WYWPhi(2:PHatMod+1)*fliplr(WW(ri-PHatMod:ri-1)')';
end
YWResiduals(1:PHatMod) = [];

WYWPhi = WYWPhi';
WYRESVec(J) = std(YWResiduals);
%New subroutine

%Calculating Mean of the newly created WSeries
WMean = mean(WSeries(201:XSize+200));

%CONDITIONING on the past values

WSeries(201:XSize+200) = XSeries-XMean;

%PREDICTION STEP 12

for I = XSize +201:XSize+200+MaxLag
     WSeries(I) = RResiduals(I);
     for Ip = 2:PHatMod +1
```

```
                WSeries(I)= WSeries(I)- WYWPhi(Ip)*WSeries(I-(Ip-1));
     end
end

EDF(1:MaxLag, J)= WSeries(XSize+201:XSize+200+MaxLag) + XMean; %+ WMean;

%Reinitializing the RResidual Array
RResiduals(:) = 0;

end
end

%-----------------------------------------------------
function f = ACV(X) % Computes the sample ACF

n = length(X);
f = zeros(n,1);
xmean = mean(X);
for k = 0:n-1
    sm = 0;
    for i = 1:n-k
        sm = sm + (X(i)-xmean)*(X(i+k)-xmean);
    end
    f(k+1) = sm/n;
end
%-----------------------------------------------------
function PHat = AICCSelection(XSeries,ACVector,PMax)

XSize = length(XSeries);
RXSize = XSize;
```

```
%Durbin Levinson Algorithm

PPhi = 0;
PPhi(1,1) = ACVector(2)/ACVector(1);
VarianceVector(1)= ACVector(1)*(1.0- PPhi(1, 1)^2);

for I = 2:PMax
     VWork = 0;
     for J = 1: I-1
           VWork = VWork + PPhi( I-1, J)*ACVector(I-J+1);
     end
     PPhi(I, I)=(ACVector(I+1) - VWork)/VarianceVector(I-1);
     for J = 1: I-1
           PPhi( I,J) = PPhi(I-1, J) - PPhi(I,I)*PPhi(I-1, I-J);
     end
     VarianceVector(I) = VarianceVector(I-1)*(1.00 - PPhi(I,I)^2);
end

I = 0;
MinimumAIC = log(ACVector(1))+ 2.0*(I)/(RXSize);

WorkP =0;
for I = 1: PMax
     WorkAIC = log(VarianceVector(I)) + 2.0*(I)/(RXSize);
     % trying to get WorkAIC printed out
     WorkAICAr(I)= WorkAIC;

    if ( WorkAIC < MinimumAIC)
        MinimumAIC = WorkAIC;
        WorkP = I;
```

```
    end
end
PHat = WorkP;


end
%-------------------------------------------------------------
```

APPENDIX B

MATLAB ALGORITHM FOR SIEVE BOOTSTRAP DIFFERENCE-BASED SEASONAL DHF TEST

```
function [H1,H2] = SBDifftest(X,PMax,s,B,alpha)
nx = length(X);
U = X(s+1:nx) - X(1:nx-s);
nu = length(U);

%DF test par estimation
[bx,sebx,rx] = getDFCoef(X,s);
bxtst = nx*(bx(1)-1);
txtst = (bx(1)-1)/sebx(1);

ACVector = ACV(U);
PHat = AICSelection(U,ACVector,PMax);

%SB procedure begins here
[bu,sigu] = aryule(U,PHat);
ru = U; %compute residuals
for i = PHat+1:nu
    ru(i)= bu*U(i:-1:i-PHat);
end
ru(1:PHat) = [];
ru = ru - mean(ru);% center the residuals
ru = [zeros(PHat,1);ru];
%bootstrap
bbxtst = zeros(B,1);
tbxtst = zeros(B,1);

for bi = 1:B

    indx = unidrnd(nu-PHat,nu+200,1) + PHat;
```

```
    bru = ru(indx);
    bru = [zeros(PHat,1);bru];

    BU = U; % The bootstrap differenced series
    for i = PHat+1:nu+200
        BU(i) = -bu(2:PHat+1)*BU(i-1:-1:i-PHat) + bru(i);
    end
    BX = X; %The bootstrap series
    for i = s+1:nu+200
        BX(i) = BX(i-s) + BU(i);
    end
    BX(1:200) = []; %delete the first 200 bootstrap observations
    [bbx,sebbx,rx] = getDFCoef(BX,s);
    bbxtst(bi) = nx*(bbx(1)-1); %bootstrap test statistics
    tbxtst(bi) = (bbx(1)-1)/sebbx(1);
end
%find critical values
bperc = prctile(bbxtst,alpha*100);
tperc = prctile(tbxtst,alpha*100);

%Decision of the test
H1 =0; H2 = 0;
if bxtst < bperc
    H1 = 1;
end

if txtst < tperc
    H2 = 1;
end
%--------------------------------------------------
```

```
function [b,seb,r] = getDFCoef(U,s)
nu = length(U);
UY = U(s+1:nu);
UX = U(1:nu-s);

[b,bint,r] = regress(UY,UX);
sigmahat = ((UY-UX*b).^2)/(nu-s-1);
seb = sigmahat*inv(UX'*UX);
seb = sqrt(seb);
r = [mean(r)*ones(s,1);r];
end

%------------------------------------------------------
function f = ACV(X) % Computes the sample ACF

n = length(X);
f = zeros(n,1);
xmean = mean(X);
for k = 0:n-1
    sm = 0;
    for i = 1:n-k
        sm = sm + (X(i)-xmean)*(X(i+k)-xmean);
    end
    f(k+1) = sm/n;
end
%------------------------------------------------------
function PHat = AICCSelection(XSeries,ACVector,PMax)

XSize = length(XSeries);
```

```
RXSize = XSize;
%Durbin Levinson Algorithm

PPhi = 0;
PPhi(1,1) = ACVector(2)/ACVector(1);
VarianceVector(1)= ACVector(1)*(1.0- PPhi(1, 1)^2);

for I = 2:PMax
     VWork = 0;
     for J = 1: I-1
           VWork = VWork + PPhi( I-1, J)*ACVector(I-J+1);
     end
     PPhi(I, I)=(ACVector(I+1) - VWork)/VarianceVector(I-1);
     for J = 1: I-1
           PPhi( I,J) = PPhi(I-1, J) - PPhi(I,I)*PPhi(I-1, I-J);
     end
     VarianceVector(I) = VarianceVector(I-1)*(1.00 - PPhi(I,I)^2);
end

I = 0;
MinimumAIC = log(ACVector(1))+ 2.0*(I)/(RXSize);

WorkP =0;
for I = 1: PMax
     WorkAIC = log(VarianceVector(I)) + 2.0*(I)/(RXSize);
     % trying to get WorkAIC printed out
     WorkAICAr(I)= WorkAIC;

    if ( WorkAIC < MinimumAIC)
        MinimumAIC = WorkAIC;
```

```
        WorkP = I;
    end
end
PHat = WorkP;


end
%-------------------------------------------------------------
```

BIBLIOGRAPHY

[1] Alonso A. M., Pena D., and Romo J. (2002), "Forecasting time series with sieve bootstrap," *Journal of Statistical Planning and Inference*, 100, 1-11.

[2] Alonso A. M., Pena D., and Romo J. (2003), "On sieve bootstrap prediction intervals," *Statistics and Probability Letters*, 65, 13-20.

[3] Alonso A. M., Pena D., and Romo J. (2004), "Introducing model uncertainty in time series bootstrap," *Statistica Sinica*, 14, 155-174.

[4] An, H. Z., Chen, Z. G., Hannan, E.J., (1982), "Autocorrelation, autoregression and autoregressive approximations," *Ann. Statist*, 10, 926936.

[5] Box G., and Jenkins G. (1970), "Time Series Analysis: Forecasting & Control," Prentice Hall.

[6] Bühlmann P. (1995), "Moving-average representation of autoregressive approximations," *Stochastic Process. Appl.*, 60, 331342.

[7] Bühlmann P. (1997), "Sieve bootstrap for time series," *Bernoulli*, 3, 123-148.

[8] Cao, R., Febrero-Bande, M., Gonzalez-Manteiga, W., Prada-Sanchez, J.M., and Garcia-Jurado, I., (1997), "Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes," *Comm. Statist. Simulation Comput.*, 26, 961-978.

[9] Chang Y., and Park Y., (2003), "A sieve bootstrap for the test of a unit root," *Journal of Time Series Analysis*, 24, 379 - 400.

[10] Dickey D.A., and Fuller W.A. (1979), "Distribution of estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427 - 431.

[11] Dickey D.A., Hasza D.P., and Fuller W.A. (1984), "Testing for unit roots in seasonal time series," *Journal of the American Statistical Association*, 79, 355 - 367.

[12] Ghysels E. , and Osborn D.R (2001), "The econometric analysis of seasonal time series," New York: Cambridge University Press.

[13] Ghysels E. , Osborn D.R, Rodrigues P.M.M. (2000), "Seasonal Nonstationarity and Near-Nonstationarity," in: B.H. Baltagi (Ed.), A Companion to Theoretical Econometrics, Blackwell Publishers, Oxford.

[14] Hamilton J. D. (1994), "Time Series Analysis," Princeton: Princeton University Press.

[15] Granger, C.W.J., Joyeux, R., (1980), "An introduction to long memory time series and fractional differencing," *Journal of Time Series Analysis*, 1, 15-39.

[16] Hannan, E.J., Kavalieris, L., (1986), "Regression, autoregression models," *Journal of Time Series Analysis*, 7, 27-49.

[17] Hosking, J.R.M., (1981), "Fractional differencing," *Biometrika*, 68, 165-176.

[18] Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S. (1990), "Seasonal integration and cointegration," *Journal of Econometrics,* 44, 215-238.

[19] Inoue, A., Kasahara Y., (2006),"Explicit representation of finite predictor coefficients and its applications," *The Annals of Statistics*, 34, 973-993.

[20] Kilian L. (1998a), "Small-sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics,* 80, 218-230.

[21] Kim J. (2001), "Bootstrap-After-Bootstrap Prediction Intervals for Autoregressive Models," *Journal of Business and Economic Statistics*, 19, 117-128.

[22] Kreiss, J.P. (1998), "Asymptotic Statistical Inference for a Class of Stochastic Processes," Habilitationsschrift, Universiti at Hamburg.

[23] Kreiss J.P. (1992), "Bootstrap procedures for AR(∞)-processes," In: JEockel, K.H., Rothe, G., Sendler, W. (Eds.), Bootstrapping and Related Techniques, Springer, Heidelberg, pp. 107-113.

[24] Mukhopadhyay P., and Samaranayake V.A. (2010), "Prediction intervals for time series: A Modified Sieve Bootstrap Approach," *Communications in Statistics - Simulation and Computation*, 39, 517-538.

[25] Palm F., Smeekes S., Urbain J. (2008), "Bootstrap unit-root tests: comparison and extension," *Journal of Time Series Analysis*, 29, 371-401.

[26] Panichkitkosolkul W.,and Niwitpong S. (2011), "On Multistep-Ahead Prediction Intervals Following Unit Root Tests for a Gaussian AR(1) Process with Additive Outliers," *Applied Mathematical Sciences*, 47, 2297-2316.

[27] Panichkitkosolkul W., and Niwitpong S. (2011), "Prediction intervals for the Gaussian autoregressive processes following the unit root tests," *Model Assisted Statistics and Applications*, 7, 1-15.

[28] Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for linear processes," *Annals of Statistics*, 20, 9711001.

[29] Poskitt D. S. (2006), "Autoregressive approximation in nonstandard situations: the fractionally integrated and non-invertible cases," *Annals of the Institute of Statistical Mathematics*, 59, 697-725

[30] Poskitt D. S. (2007), "Properties of the Sieve bootstrap for fractionally integrated and non-invertible processes," *Journal of Time Series Analysis*, 29, 224-250.

[31] Psaradakis Z. (2001), "Bootstrap test for an autoregressive unit root in the presence of weakly dependent errors," *Journal of Time Series Analysis*, 22, 577-594.

[32] Rupasinghe M., and Samaranayake V.A. (2012a), "Asymptomatic Properties of Sieve Bootstrap Prediction Intervals for FARIMA Processes," *Statistics and Probability Letters*, Submitted.

[33] Rupasinghe M., and Samaranayake V.A. (2012b), "Obtaining Prediction Intervals for FARIMA Processes using sieve bootstrap," Preprint.

[34] Rupasinghe M., and Samaranayake V.A. (2012c), "Prediction intervals for ARIMA processes: A sieve bootstrap approach," Preprint.

[35] Rupasinghe M., and Samaranayake V.A. (2012d), "The asymptotic distributions of the Dickey-Hasza-Fuller seasonal unit root tests under weakly dependent errors," Preprint.

[36] Said E., and Dickey D. (1984), "Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order," *Biometrika*, 71, 599607.

VITA

Maduka Nilanga Rupasinghe attended Royal College, Colombo Sri Lanka from 1987 to 2000. In 2001 he entered University of Colombo, Sri Lanka and completed his bachelor's degree in Business, Finance and Computational Mathematics in 2005. Maduka entered Missouri University of Science and Technology for graduate studies in 2007. He received his MS degree in Applied Mathematics with Statistics emphasis in 2010 and further entered to the PhD program at the same.