https://doi.org/10.26219/heal.aueb.6514

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

SCHOOL OF INFORMATION SCIENCES AND TECHNOLOGY

DEPARTMENT OF STATISTICS
POSTGRADUATE PROGRAM

A simulation study on forecasting seasonal time series under mis-specification of the non-stationary seasonal roots

by

Sofia-Eirini Nikolakakou

M.Sc. Thesis
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfillment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
October 2022

https://doi.org/10.26219/heal.aueb.6514

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ &
ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

Μια μελέτη προσομοίωσης για την πρόβλεψη εποχιακών χρονοσειρών υπό την εσφαλμένη προδιαγραφή των μη στάσιμων εποχιακών ριζών

Σοφία-Ειρήνη Νικολακάκου

Διατριβή
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μετατπυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Οκτώβριος 2022

ACKNOWLEDGEMENTS

I would like to express my appreciation and give my special thanks to my supervisor Prof. Evangelos Ioannidis, for the guidance and support he showed me throughout this dissertation. His constructed feedback and overall insights have made this an inspiring experience for me. I would also like to acknowledge my family and friends for their continuous support and understanding during this challenging process. I could not be more grateful for their love and encouragement.

https://doi.org/10.26219/heal.aueb.6514

ABSTRACT

A simulation study on forecasting seasonal time series under mis-specification of the non-stationary seasonal roots

by Sofia-Eirini Nikolakakou

October 2022

When dealing with stochastic seasonality, especially in the case when seasonal unit roots are present, one well-known class of time series models used to describe them are the so-called Seasonal ARIMA models (SARIMA), which are frequently used as basis for forecasting. One important assumption these types of models make, is that there will be presence of non-stationary roots at all seasonal frequencies. However, since the work of Hylleberg, Engle, Granger and Yoo (1990), it is seen that is not case, since many time series do not forcefully have seasonal unit roots at all their frequencies. The key objective of the present dissertation is to investigate the implications on the models' predictive ability, when imposing in the fitted models more non-stationary or even stationary roots than those present in the data generating process. For the out-of-sample forecast analysis, the comparison of the fitted models' behavior is based on the Mean Squared Prediction Error as well as the variance of the forecast errors. It is generally found that the assumption and estimation of more non-stationary roots, than those actually present, can lead to larger prediction errors, contrasting therefore the main idea around which the SARIMA models are based.

https://doi.org/10.26219/heal.aueb.6514

ΠΕΡΙΛΗΨΗ

Μια μελέτη προσομοίωσης για την πρόβλεψη εποχιακών χρονοσειρών υπό την εσφαλμένη προδιαγραφή των μη στάσιμων εποχιακών ριζών

Σοφία-Ειρήνη Νικολακάκου

Οκτώβριος 2022

Όταν ασχολούμαστε με τη στοχαστική εποχικότητα, ειδικά στην περίπτωση που υπάρχουν εποχιακές μοναδιαίες ρίζες, μια γνωστή κατηγορία μοντέλων χρονοσειρών που χρησιμοποιούνται για την περιγραφή τους είναι τα λεγόμενα εποχιακά μοντέλα ARIMA (SARIMA), τα οποία χρησιμοποιούνται συχνά ως βάση για προβλέψεις. Μια σημαντική υπόθεση που κάνουν αυτοί οι τύποι μοντέλων είναι η παρουσία μη στάσιμων ριζών σε όλες τις εποχιακές συχνότητες. Ωστόσο, από το έργο των Hylleberg, Engle, Granger και Yoo (1990), φαίνεται πως δεν συμβαίνει αυτό, καθώς πολλές χρονολογικές σειρές δεν έχουν αναγκαστικά εποχιακές μοναδιαίες ρίζες σε όλες τις συχνότητές τους. Ο βασικός στόχος της παρούσας διατριβής είναι να διερευνήσει τις επιπτώσεις στην προγνωστική ικανότητα των μοντέλων, όταν επιβάλλονται στα προσαρμοσμένα μοντέλα περισσότερες μη στάσιμες ή ακόμη και στάσιμες ρίζες από αυτές που υπάρχουν στη διαδικασία παραγωγής δεδομένων. Για την ανάλυση πρόβλεψης εκτός δείγματος, η σύγκριση της συμπεριφοράς των προσαρμοσμένων μοντέλων βασίζεται στο Μέσο Τετράγωνο Σφάλμα Πρόβλεψης καθώς και στη διακύμανση των σφαλμάτων πρόβλεψης. Γενικά διαπιστώνεται πως η υπόθεση και η εκτίμηση περισσότερων μη στάσιμων ριζών, από αυτές που υπάρχουν στην πραγματικότητα, μπορεί να οδηγήσει σε μεγαλύτερα σφάλματα πρόβλεψης, αντιπαραβάλλοντας επομένως την κύρια ιδέα γύρω από την οποία βασίζονται τα μοντέλα SARIMA.

https://doi.org/10.26219/heal.aueb.6514

Contents

https://doi.org/10.26219/heal.aueb.6514

Chapter 1

Introduction

When conducting a time series analysis, it is common to take it a step further in order to make predictions. Forecasting stationary time series based on observed past values is already well understood Brockwell & Davis (2002) [6]. Even the implications of an autoregressive root at unity, also known as unit root process, are known and addressed by Brockwell & Davis (2002) [6], as well as Hamilton (1994) [12]. However when seasonality is present and especially seasonal unit roots, the implications on the forecasts are less obvious.

Regarding the seasonal pattern, which is present in many time series, there are three types of seasonality, as presented by Hylleberg, Engle, Granger and Yoo (HEGY) (1990) [13]. Firstly, there is deterministic seasonality, which has a seasonal pattern that repeats year after year. This process is generated by seasonal dummy variables, and can be perfectly forecast while always maintaining its shape.

Besides the deterministic seasonality, if the seasonal pattern randomly varies from one cycle to the next, then stochastic seasonality is present, which can be divided into stationary and non-stationary. Both types of stochastic seasonality can be described by the so-called seasonal ARIMA model which is written in the form of

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D X_t = \theta(B)\Theta(B^S)\varepsilon_t$$

where $\phi(B), \Phi(B^S), \theta(B)$ and $\Theta(B^S)$ are the non-seasonal and seasonal AR and MA polynomials respectively.

As long as the roots of the $\Phi(B^S)$ and $\phi(B)$ polynomials lie outside the unit circle, with some being complex pairs of seasonal periodicities, when referring to the seasonal polynomial, then we refer to stationary stochastic seasonality and the differencing filters are omitted, i.e. $d = D = 0$.

In the case of a non-stationary stochastic seasonality we assume that there exist roots on the unit circle that are allowed only if they are simultaneously present at all seasonal frequencies, which is a usual assumption SARIMA models make.

https://doi.org/10.26219/heal.aueb.6514

Moreover, the roots of the seasonal polynomial $\Phi(B^S)$ that lie outside the unit circle have to be present at all seasonal frequencies simultaneously and are all equally close to the unit circle. This results due to the fact that all roots have the same modulus and their angles differ by multiples of $\frac{2\pi}{S}$.

Using SARIMA models as a basis for forecasting is quite common, however since the work of HEGY [13], it is clear that seasonal time series do not forcefully have non-stationary roots at all seasonal frequencies, which is assumed by SARIMA models. Therefore we want to consider models which allow only some of the seasonal unit roots to be on the unit circle, while ignoring for simplicity the moving-average polynomial, that allow to have roots outside the unit circle or on the unit circle at specific seasonal frequencies. This implies that only some of the roots at certain frequencies will be present while others might not exist at all.

The purpose of this thesis is to better understand how imposing in the fitted model more or less non-stationary roots than those present in the data generating process (DGP), will affect the fitted model's predictive performance.

We are interested in comparing different estimated models and examine how well or not they perform when conducting an out-of-sample forecast analysis, based on specific characteristics of the model used to generate the original time series. We compare their behavior based on the mean squared prediction error, which is used as the evaluation metric. To illustrate the differences between these models, we compare boxplots of their forecast errors (figures 5.1, 5.4) as well as boxplots of the nominal standard error of the forecast as estimated by the R function `predict` divided by the square root of the mean squared prediction error, as estimated by the simulation (figure 5.2, 5.5).

Specifically, for the analysis we defined the model with the differencing filter $(1 - \sqrt{3}B + B^2)$ to generate the desired sample, which corresponds to one non-stationary seasonal root at frequency $2\pi/12$, with a period of 12 months. For the estimation process, four models were used to make comparisons. The first one being the same as the model used in the DGP, the second with 12 identical non-stationary seasonal roots, known as $SARIMA(0,1,0)_{12}$, the third with 12 stationary roots, which are allowed to be as close as desired to the unit circle, namely an $AR(12)$ model and finally a stationary $AR(2)$ model, with the order corresponding to the number of roots present in the DGP.

A second case was taken into account, with a model of the form $(1 - \sqrt{3}B + B^2)(1 - B + B^2)X_t = \varepsilon_t$ being used in the DGP, which has two non-stationary seasonal roots at $2\pi/12$ and $4\pi/12$. The models used in the estimation process were similar to the ones used in the first case, namely, $SARIMA(0,1,0)_{12}$, $AR(4)$ and $AR(12)$ models. In the simulation study, a total of 1000 simulations was performed, where each time a sample of 25 years of monthly observations was split into training and test set, in order to estimate the models which were then used to make predictions for the next two years. The results were stored and portrayed in boxplots of forecast errors and variances, to finally make the comparisons, as shown in Chapter 5.

Based on the results of the analysis, it is noticeable that even though the SARIMA model is often considered as a suitable model, the predictions obtained are not always satisfactory when the original time series is generated by less than 12 non-stationary roots. Interestingly, different results may be obtained depending on which forecast horizon is taken into account, whether it is a multiple of the seasonal period or not.

To sum up, in the second chapter some general concepts of time series are introduced. Chapter 3 refers to Unit Root processes, how they are defined and how we may obtain forecasts when dealing with such non-stationary processes. In Chapter 4 the different types of seasonality are explained and what types of models are used to describe them. Furthermore, we show how the roots of an integrated seasonal process are calculated, along with the frequencies that correspond to each root for monthly time series. In the fifth Chapter, the simulation experiment conducted in this thesis is presented, specifically how the simulation was created and what results were obtained. Finally, Chapter 6 ends with some concluding remarks for the results previously shown.

https://doi.org/10.26219/heal.aueb.6514

Chapter 2

Preliminaries

In this chapter some basic terms of time series analysis and stochastic processes are introduced, such as the concepts of stationarity and the autocovariance and sample autocovariance functions, as defined by Brockwell and Davis (2002) [6].

A **time series** is a set of observations $\{x_t\}$, each one being recorded at a specific time t. A discrete-time time series is one in which the set T_0 of times at which observations are made is a discrete set.

A **time series model** for the observed data $\{x_t\}$ is a specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization.

Important concepts in time series analysis, are the stationarity assumption and the autocovariance function of the series, which are presented by the following definitions:

If $\{X_t\}$ is a time series with $E(X_t^2) < \infty$ and mean function $\mu_t(t) = E(X_t)$, then the **autocovariance function** of this series is

$$\gamma_x(t, t-k) = Cov(X_t, X_{t-k}) = E[(X_t - \mu_x(t))(X_{t-k} - \mu_x(t-k))]$$

for all t and any k.

As for the stationarity condition, a series $\{X_t\}$ is said to be **stationary** if

i. $E(X_t^2) < \infty$ for all t,

ii. $\mu_t(t) = E(X_t)$ is independent of time t,

iii. $\gamma_x(t, t-k)$ is independent of t for each k.

The above definition refers to the weakly or covariance stationarity. Strict stationarity of a time series $\{X_t\}$ is defined by the condition that $(X_1, \ldots, X_n)$ and $(X_{1+h}, \ldots, X_{n+h})$ have the same joint distributions for all integers h and $n > 0$. However for the rest of the thesis, whenever the term stationarity is mentioned it will always refer to the definition stated above.

https://doi.org/10.26219/heal.aueb.6514

Another important definition, is that of the white noise process, which is the basic building block for all the processes considered in the present thesis. A sequence $\{\varepsilon_t\}$ with mean zero and variance σ^2, whose ε's are uncorrelated across time is said to be a **white noise** process.

Next, we define some of the key properties of an important class of linear processes known as ARMA (autoregressive moving average) processes, which are defined by linear difference equations with constant coefficients. These properties are called invertibility and causality.

An ARMA*(p,q)* process defined by $\phi(B)X_t = \theta(B)\varepsilon_t$ is said to be **invertible** if there exists a sequence of constants $\{\pi_j\}$ such that $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad t = 0, \pm 1, \ldots$$

For a process to be invertible, that essentially means that ε_t can be expressed in terms of present and past values of $X_s,\ s \leq t$.

An ARMA*(p,q)* process defined by $\phi(B)X_t = \theta(B)\varepsilon_t$ is said to be a **causal** function of $\{\varepsilon_t\}$ if there exists a sequence of constants $\{\psi_j\}$ such that $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad t = 0, \pm 1, \ldots$$

This means that X_t can be expressed in terms of past and current values of $\varepsilon_s,\ s \leq t$.

In order to understand how to **forecast** non-stationary (seasonal) processes, it is important to first present how forecasts are obtained for ARMA processes, based on section 3.3 of Brockwell & Davis (2002) [6]. Throughout the rest of this chapter, we assume that X_t is stationary and has a known autocovariance function.

We wish to predict the values $X_{n+h}, h > 0$ in terms of the values $\{X_n, \ldots, X_1\}$ up to time n. The goal is to find the best linear combination of $1, X_n, X_{n-1}, \ldots, X_1$ that forecasts X_{n+h} with minimum squared error. We denote the best linear predictor as $P_n X_{n+h}$ which is of the form

$$P_n X_{n+h} = \alpha_0 + \alpha_1 X_n + \cdots + \alpha_n X_1$$

where the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_n$ are obtained by the finding the values that minimize the forecast error variance $S(\alpha_0, \ldots, \alpha_n) = E(X_{n+h} - \hat{X}_{n+h})^2$.

The equations used to solve for the coefficients $\alpha_0, \alpha_1, \ldots, \alpha_n$ are the following:

$$\mathbf{E}\left[X_{n+h} - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_{n+1-i}\right] = 0, \tag{2.1}$$

https://doi.org/10.26219/heal.aueb.6514

$$\mathbf{E}\left[(X_{n+h} - \alpha_0 - \sum_{i=1}^{n} \alpha_i X_{n+1-i})X_{n+1-k}\right] = 0, \qquad k = 1, \ldots, n \tag{2.2}$$

From equation (2.1) if we solve for α_0 we get

$$\alpha_0 = \mu \left(1 - \sum_{i=1}^{n} \alpha_i\right)$$

The form of the best linear predictor is thus,

$$P_n X_{n+h} = \mu + \sum_{i=1}^{n} \alpha_i (X_{n+1-i} - \mu) \tag{2.3}$$

Without loss of generality we can consider the case where $\mu = 0$, in which case $\alpha_0 = 0$.

To forecast a causal ARMA process of the form $\phi(B)X_t = \theta(B)\varepsilon_t, \varepsilon_t \sim WN(0, \sigma^2)$ we can make use of the innovations algorithm, which is a recursive method for forecasting second-order zero-mean processes, regardless of whether they are stationary or not. This algorithm will not be analyzed in this thesis, however details can be found in Brockwell & Davis (2002) [6].

For the ARMA process the innovations algorithm will be applied not directly to $\{X_t\}$ but rather to the transformed process proposed by Ansley (1979) [1], that is

$$\begin{aligned} W_t &= \sigma^{-1} X_t, && t = 1, \ldots, m \\ W_t &= \sigma^{-1} \phi(B) X_t, && t > m \end{aligned} \tag{2.4}$$

where $m = max(p, q)$.

For convenience, we shall define $\theta_0 = 1$, while also assuming that $p \geq 1$ and $q \geq 1$ without loss of generality. We first need to define the autocovariances $\kappa(i, j) = \mathbf{E}(W_i W_j), i, j \geq 1$ which are given from

$$\kappa(i,j) = \begin{cases} \sigma^{-2} \gamma_X(i-j), & 1 \leq j, j \leq m, \\ \sigma^{-2} \left[\gamma_X(i-j) - \sum_{r=1}^{p} \phi_r \gamma_X(r - |i-j|)\right] & \text{min(i,j)} \leq m < max(i,j) \leq 2m, \\ \sum_{r=0}^{q} \theta_r \theta_{r+|i+j|}, & \text{min(i,j)>m}, \\ 0, & \text{otherwise} \end{cases} \tag{2.5}$$

with γ_X being the autocovariance function of $\{X_t\}$.

We can then apply the innovations algorithm to the transformed process $\{W_t\}$ and

https://doi.org/10.26219/heal.aueb.6514

obtain

$$
\begin{aligned}
\hat{W}_{n+1} &= \sum_{j=1}^{n} \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}), && 1 \leq n < m, \\
\hat{W}_{n+1} &= \sum_{j=1}^{q} \theta_{nj}(W_{n+1-j} - \hat{W}_{n+1-j}), && n \geq m \qquad (2.6)
\end{aligned}
$$

where the coefficients θ_{nj} as well as the mean squared errors $\mathbf{E}(W_{n+1} - \hat{W}_{n+1})^2$ can be calculated recursively with the use of the innovations algorithm and the autocovariances defined in equation (2.5).

From equations (2.4) we can observe that the best linear predictor of any random variable Y in terms of $\{1, X_1, \ldots, X_n\}$ is the same as the best linear predictor of Y in terms of $\{1, W_1, \ldots, W_n\}$, which we shall denote as $P_n Y_n$. Using this linearity we conclude that

$$
\begin{aligned}
\hat{W}_t &= \sigma_{-1}\hat{X}_t, && t = 1, \ldots, m, \\
\hat{W}_t &= \sigma_{-1}\left[\hat{X}_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}\right], && t > m \qquad (2.7)
\end{aligned}
$$

which together with (2.4) shows that

$$
X_t - \hat{X}_t = \sigma\left[W_t - \hat{W}_t\right] \qquad \forall t \geq 1 \qquad (2.8)
$$

If we replace $W_j - \hat{W}_j$ by $\sigma^{-1}(X_j - \hat{X}_j)$ in (2.5) and the substitute into (2.6) we can obtain

$$
\hat{X}_{n+1} = \begin{cases} \displaystyle\sum_{j=1}^{n} \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}), & 1 \leq n < m, \\ \displaystyle\phi_1 X_n + \cdots + \phi_p X_{n+1-p} + \sum_{j=1}^{q} \theta_{nj}(X_{n+1-j} - \hat{X}_{n+1-j}), & \mathrm{n} \geq m \end{cases} \qquad (2.9)
$$

from which we can calculate recursively the one-step ahead predictors $\hat{X}_2, \hat{X}_3, \ldots$.

The mean squared error will then be

$$
\mathbf{E}\left(X_{n+1} - \hat{X}_{n+1}\right)^2 = \sigma^2 \mathbf{E}\left(W_{n+1} - \hat{W}_{n+1}\right)^2
$$

For an ARMA*(p,q)* process, after a recursive calculation using the innovations algorithm, we conclude that for $n > m = max(p, q)$ and for all $h \leq 1$, the h-step predictors

https://doi.org/10.26219/heal.aueb.6514

satisfy the following

$$P_n X_{n+h} = \sum_{i=1}^{p} \phi_i P_n X_{n+h-i} + \sum_{j=h}^{q} \theta_{n+h-1,j} \left(X_{n+h-j} - \hat{X}_{n+h-j} \right) \tag{2.10}$$

https://doi.org/10.26219/heal.aueb.6514

Chapter 3

Unit Root Processes

3.1 Unit Roots

In many occasions we are dealing with a set of observations that is not necessarily generated by a stationary time series. One well-known class of such non-stationary time series is the Unit Root process.

A unit root means that either the autoregressive or the moving-average polynomial of an ARMA process has a root on or near the unit circle.

Lets assume we have observed a set of univariate time series data, which show a linear trend in time. The unit root model for such data is defined as:

$$(1-B)X_t = \delta + u_t \tag{3.1}$$

where

$$u_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots = \psi(B)\varepsilon_t \tag{3.2}$$

with $\{\varepsilon_t\} \sim WN(0, \sigma^2)$ and $\psi(1) \neq 0$.

Here B is the backward shift operator,

$$BX_t = X_{t-1}.$$

The best known example of a unit root process is known as a random walk with drift δ, which is obtained by setting $\psi(B) = 1$ in equation (3.1) .

According to Hamilton (1994) [12] we can consider another specification for the unit root process:

$$X_t = \alpha + \delta t + u_t, \tag{3.3}$$

where u_t follows a zero-mean ARMA process:

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)u_t=(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q)\varepsilon_t \tag{3.4}$$

with the moving average polynomial in the right-hand side being invertible. Since we are working with polynomials of order p, it is easier to factor them

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)=(1-\lambda_1 B)(1-\lambda_2 B)\cdots(1-\lambda_p B) \tag{3.5}$$

If all the roots $\lambda_1,\lambda_2,\ldots,\lambda_p$ are inside the unit circle then (3.4) can be written as:

$$u_t=\frac{(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q)}{(1-\lambda_1 B)(1-\lambda_2 B)\cdots(1-\lambda_p B)}\varepsilon_t=\psi(B)\varepsilon_t \tag{3.6}$$

with $\sum_{j\neq 0}^{\infty}|\psi_j|<\infty$ and roots of $\psi(z)=0$ outside the unit circle. If we suppose instead that $\lambda_1=1$ and $|\lambda_i|<1$ for $i=2,3,\ldots,p$, then (3.4) would state that

$$(1-B)(1-\lambda_2 B)(1-\lambda_3 B)\cdots(1-\lambda_p B)u_t=(1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q)\varepsilon_t \tag{3.7}$$

implying that

$$(1-B)u_t=\frac{1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q}{(1-\lambda_2 B)(1-\lambda_3 B)\cdots(1-\lambda_p B)}\varepsilon_t=\psi^*(B)\varepsilon_t \tag{3.8}$$

with $\sum_{j=0}^{\infty}|\psi_j^*|<\infty$ and roots of $\psi^*(z)=0$ outside the unit circle. Thus, if (3.3) is first-differenced, the result is

$$(1-B)X_t=(1-B)\alpha+[\delta t-\delta(t-1)]+(1-B)u_t=0+\delta=\psi^*(B)\varepsilon_t \tag{3.9}$$

which is of the form of the unit root process in equation (3.1).

Equation (3.1) can also be expressed as a process that is integrated of order 1, and is denoted as $X_t\sim I(1)$. The term *"integrated of order d"* (or $I(d)$) is used to describe a series that has d unit roots and that is stationary after being differenced d times.

A general process X_t is called an *autoregressive integrated moving average process* - ARIMA(p,d,q) -, if its dth difference denoted by $(1-B)^d X_t$, results in a stationary ARMA(p,q) process.

https://doi.org/10.26219/heal.aueb.6514

3.2 Forecasting Unit Roots

Suppose we have observed $X_1, X_2, \ldots, X_t$ and want to generate forecasts of future values of an ARIMA(p,d,q) process $\{X_t\}$ for h steps ahead. Since we are referring to unit roots, we assume that the process $\{X_t\}$ satisfies the difference equations

$$(1-B)^d X_t = Y_t, \qquad t = 1, 2, \ldots$$

where $\{Y_t\}$ is a causal ARMA process and $(X_{1-d}, \ldots, X_0)$ is uncorrelated with $Y_t, t > 0$. We can rewrite the above equation in the form

$$X_t = Y_t - \sum_{j=1}^{d} \binom{d}{j} (-1)^j X_{t-j} \tag{3.10}$$

In order to find the best linear predictors $P_n X_{n+h}$ we shall apply the operator P_n, which denotes the best linear prediction in terms of the observations up to time n, to each side of equation (3.10) to obtain

$$P_n X_{n+h} = P_n Y_{n+h} - \sum_{j=1}^{d} \binom{d}{j} (-1)^j P_n X_{n+h-j} \tag{3.11}$$

The predictors $P_n X_{n+h}$ are obtained from (3.11) by noting that $P_n X_{n+1-j} = X_{n+1-j}$ for each $j \geq 1$. The h-step ahead prediction formula, after we express $P_n Y_{n+h}$ in terms of $\{X_j\}$, and based on equation (2.10) is

$$P_n X_{n+h} = \sum_{j=1}^{p+d} \phi_j^* P_n X_{n+h-j} + \sum_{j=h}^{q} \theta_{n+h-1,j} \left(X_{n+h-j} - \hat{X}_{n+h-j} \right) \tag{3.12}$$

where $\phi^*(z) = (1-z)^d \phi(z) = 1 - \phi_1^* z - \cdots - \phi_{p+d} z^{p+d}$.

The mean squared error of the h-step predictor is

$$\sigma_n^2(h) = E(X_{n+h} - P_n X_{n+h})^2 = \sum_{j=0}^{h-1} \left(\sum_{r=0}^{j} \chi_r \theta_{n+h-r-1,j-r} \right)^2 v_{n+h-j-1} \tag{3.13}$$

where $\theta_{n0} = 1$,

$$\chi(z) = \sum_{r=0}^{\infty} \chi_r z^r = (1 - \phi_1^* z - \cdots - \phi_{p+d}^* z^{p+d})^{-1},$$

and

$$v_{n+h-j-1} = E(X_{n+h-j} - \hat{X}_{n+h-j})^2 = E(Y_{n+h-j} - \hat{Y}_{n+h-j})^2.$$

https://doi.org/10.26219/heal.aueb.6514

For large n we can approximate the mean squared error by

$$\sigma_n^2(h) = \sum_{j=0}^{h-1} \psi_j^2 \sigma^2 \tag{3.14}$$

where

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = (\phi^*(z))^{-1} \theta(z).$$

Example: Here we consider the case of the random walk with drift, which is written as

$$(1 - B)X_t = \mathbf{\Delta} X_t = \delta + u_t$$

with $u_t \sim WN(0, \sigma^2)$. Given the observations $\{X_1, \ldots, X_n\}$, the one-step ahead forecast is given by

$$\hat{X}_{n+1} = \mathbf{E}[X_{n+1}|X_n, \ldots, X_n] = \mathbf{E}[\delta + X_n + u_{n+1}|\mathcal{X}_{\mathcal{N}}] = \delta + X_n$$

We compute this recursively, and therefore find that the h-step predictor is

$$\hat{X}_{n+h} = X_n + h\delta$$

To obtain the forecast errors, we can express for convenience X_n as $X_n = n\delta + \sum_{r=1}^{n} u_r$ and for h steps ahead

$$X_{n+h} = X_n + h\delta + \sum_{r=n+1}^{n+h} u_r$$

We are then able to calculate the h-step prediction error which will become

$$\mathbf{E}\left[X_{n+h} - \hat{X}_{n+h}\right]^2 = \mathbf{E}\left[\sum_{r=n+1}^{n+h} u_r\right]^2 = \sigma^2 h$$

For these types of processes, the mean squared prediction error constantly increases linearly with the forecast horizon h, without converging to any fixed value as h goes to infinity.

Chapter 4

Seasonal ARIMA processes

Seasonality is a characteristic of time series where regular and predictable variations occur at specific intervals, e.g. every month, quarter or year. The seasonal component in a time series model repeats itself precisely in the same way cycle after cycle.

As described by Hylleberg et al. (1990) [13], to model seasonality three classes of time series models are most commonly used. These are:

1. Deterministic seasonal processes
2. Stationary seasonal processes
3. Processes with seasonal unit roots

These three types of models are explained in the sections below.

4.1 Deterministic seasonal processes

Deterministic seasonality can be described as the seasonal pattern that repeats itself in the same way cycle after cycle and is due to the unconditional mean of the time series. This type of seasonality can be removed from data by seasonal adjustment procedures and can be expressed by means of seasonal dummy variables.

The seasonal dummy variable representation of seasonality is written as:

$$X_t = \mu + \sum_{k=1}^{S-1} \delta_k D_{kt} + \varepsilon_t, \quad t = 1, 2, \ldots, T \tag{4.1}$$

where S is the period of seasonality, D_{kt} are the seasonal dummy variables which take the value 1 if observation at time t is in season s and 0 otherwise, δ_s are the seasonal dummy variable coefficients and ε_t is some ARMA-type error process. For such processes, the optimal forecast of X_t for season s will be the same for each future point in time for said season, meaning that for example next January's value will be the same no matter the year we are in.

4.2 Stationary seasonal processes

On the other hand, *stochastic seasonality* allows for randomness in the seasonal pattern from one cycle to the next, including both stationary and non-stationary processes.

When a series exhibits seasonal behaviour with known periodicity S, the ARMA model for stationary series can be generalized to include both the regular dependence, which is associated with the sampling intervals of the fixed length, and the seasonal dependence, which is associated with observations separated by S sampling intervals. Both of these dependencies can be expressed by the multiplicative seasonal ARMA process

$$\phi(B)\Phi(B^S)X_t = \theta(B)\Theta(B^S)\varepsilon_t, \qquad \varepsilon_t \sim WN(0,\sigma^2) \tag{4.2}$$

with $\phi(z) = 1-\phi_1 z-\cdots-\phi_p z^p$, $\Phi(z^S) = 1-\Phi_1 z^S-\cdots-\Phi_P z^{PS}$, $\theta(z) = 1+\theta_1 z+\cdots+\theta_q z^q$, $\Theta(z^S) = 1+\Theta_1 z^S+\cdots+\Theta_Q z^{QS}$. The roots of these polynomials are required to lie outside the unit circle to ensure stationarity.

4.3 Non-Stationary seasonal processes

A *non-stationary stochastic* series X_t, observed at S equally spaced time intervals per year, is a seasonally integrated process of order D if it has seasonal unit roots in its autoregressive representation. This term refers to the seasonal differencing so that stationarity may emerge when the differenced series at lag S is modeled by a stationary invertible ARMA process.

The integrated models that display this type of seasonality, are called general multiplicative seasonal ARIMA (SARIMA) models and are formulated by the following definition, according to Brockwell and Davis (2002) [6].

If d and D are nonnegative integers then $\{X_t\}$ is a **seasonal ARIMA**$(\boldsymbol{p,d,q}) \times (\boldsymbol{P,D,Q})_{\boldsymbol{S}}$ **process with period** $\boldsymbol{S}$ if the differenced series $Y_t = (1-B)^d(1-B^S)^D X_t$ is a causal ARMA process defined by

$$\phi(B)\Phi(B^S)Y_t = \theta(B)\Theta(B^S)\varepsilon_t, \qquad \varepsilon_t \sim WN(0,\sigma^2) \tag{4.3}$$

where $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, $\Phi(z) = 1 - \Phi_1 z - \cdots - \Phi_P z^P$, $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$, $\Theta(z) = 1 + \Theta_1 z + \cdots + \Theta_Q z^Q$.

Below are some examples of such processes.

4.3.1 Examples of SARIMA processes

1 Seasonal Random Walk

The simplest case of a seasonally integrated process is the seasonal random walk defined as

$$(1 - B^S)X_t = \varepsilon_t \tag{4.4}$$

or if we consider a drift, in order to allow for a trend behaviour that is usually observed in economic time series, the seasonal random walk can be expressed as

$$(1 - B^S)X_t = \gamma + \varepsilon_t \tag{4.5}$$

with $\gamma \neq 0$ and $\varepsilon_t \sim WN(0, \sigma^2)$.

2 SARIMA$(\mathbf{1}, \mathbf{0}, \mathbf{0}) \times (\mathbf{1}, \mathbf{0}, \mathbf{0})_{\mathbf{12}}$

The process $SARIMA(1, 0, 0) \times (1, 0, 0)_{12}$ includes only non-seasonal and seasonal AR terms of order 1 each and has a seasonal period of S=12. The model is written as

$$(1 - \phi B)(1 - \Phi B^{12})(X_t - \mu) = \varepsilon_t \tag{4.6}$$

or, in its expanded form after we set $Y_t = X_t - \mu$

$$Y_t = \phi Y_{t-1} + \Phi Y_{t-12} - \phi \Phi Y_{t-13} + \varepsilon_t$$

3 SARIMA$(\mathbf{0}, \mathbf{1}, \mathbf{1}) \times (\mathbf{0}, \mathbf{1}, \mathbf{1})_{\mathbf{12}}$

The process $SARIMA(0, 1, 1) \times (0, 1, 1)_{12}$ includes both non-seasonal and seasonal MA terms of order 1 each, the differencing terms d=1 and D=1 and has a seasonal period of S=12. By using the formula in equation (4.3), we obtain the following model

$$(1 - B)(1 - B^{12})X_t = (1 - \theta B)(1 - \Theta B^{12})\varepsilon_t \tag{4.7}$$

or, when expanded

$$X_t = X_{t-1} + X_{t-12} + X_{t-13} + \varepsilon_t + \theta \varepsilon_{t-1} + \Theta \varepsilon_{t-12} + \theta \Theta \varepsilon_{t-13}$$

4.4 Seasonal roots and their frequencies

Any seasonal series has a spectrum with distinct peaks at the frequencies $\omega_j = \frac{2\pi j}{S}, \quad j = 0, 1, \ldots, S-1$ with S being the period of its seasonal component. For $j = 0$ (or equivalently when the frequency is equal to zero) the root is said to be non-seasonal, while the rest are called seasonal roots. To find the frequency that is associated with a specific root we need the value of ω in $e^{\omega i}$, which is the polar representation of the root.

For an integrated seasonal process of order $D = 1$, the characteristic equation will be

$$(1 - B^S) = 0 \qquad \textit{or} \qquad (1 - z^S) = 0$$

Denoting by z_j the roots of this polynomial we obtain

$$z_j = e^{\frac{2\pi j}{S} i} = \cos\left(\frac{2\pi j}{S}\right) + i \sin\left(\frac{2\pi j}{S}\right)$$

with $j = 0, 1, \ldots, S-1$. Each root z_j is associated with a particular frequency ω_j.

For the seasonal unit roots we consider the following seasonal lag polynomial

$$\begin{aligned} S(B) &= (1+B)(1+B^2)(1+B^4+B^8) \\ &= (1+B)(1+B^2)(1-B^2+B^4)(1+B^2+B^4) \\ &= (1+B)(1+B^2)(1-\sqrt{3}B+B^2)(1+\sqrt{3}B+B^2) \\ &\quad \times (1-B+B^2)(1+B+B^2) \end{aligned} \tag{4.8}$$

By seasonal unit roots it is meant that each factor will have its zero(s) on the unit circle in the complex plane.

In the case of monthly observations, where $S = 12$, by using the seasonal differencing operator and the seasonal polynomial in equation (4.8) we can express the time series $\{X_t\}$ as:

$$(1 - B^{12})X_t = (1 - B)S(B)X_t = \varepsilon_t \tag{4.9}$$

For $S = 12$, we present the seasonal unit roots in addition to the non-seasonal one, in the table below:

https://doi.org/10.26219/heal.aueb.6514

Roots	**Frequency**
1	0
-1	π
$\frac{1}{2}(\sqrt{3} \pm i)$	$\frac{\pi}{6}$
$\frac{1}{2}(1 \pm \sqrt{3}i)$	$\frac{\pi}{3}$
$\pm i$	$\frac{\pi}{2}$
$-\frac{1}{2}(1 \pm \sqrt{3}i)$	$\frac{2\pi}{3}$
$-\frac{1}{2}(\sqrt{3} \pm i)$	$\frac{5\pi}{6}$

Table 4.1: Seasonal unit roots of monthly time series

Figure 4.1 shows the disposition of the twelve roots of unity around the unit circle in the complex plane.

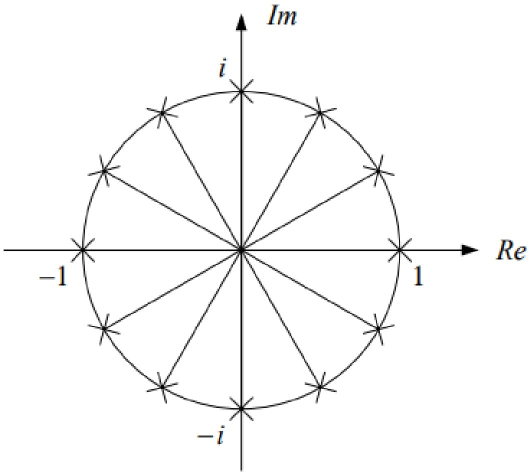

Figure 4.1: Twelve roots of unit resulting from the $(1 - B^{12})$ differencing filter

4.5 Forecasting SARIMA models

In order to forecast SARIMA processes the same procedure is applied as the one used for the forecasting of ARIMA processes, based on section 6.5.1 in Brockwell & Davis (2002) [6].

https://doi.org/10.26219/heal.aueb.6514

Here we assume that the observed process $\{X_t\}$ satisfies the following difference equation

$$(1-B)^d(1-B^S)^D X_t = Y_t, t = 1, 2, \dots$$

with Y_t being a causal ARMA process. Again the goal is to find the best linear predictors in terms of the observed data up to time n, and use them to make forecasts for h steps ahead. By setting the time t as $t = n+h$ we get

$$X_{n+h} = Y_{n+h} + \sum_{j=1}^{d+DS} a_j X_{n+h-j} \tag{4.10}$$

The best linear predictors $P_n X_{n+h}$ of X_{n+h} based on the observations $1, X_{-d-DS}, \dots, X_n$, if we assume the first $d+DS$ observations to be uncorrelated with $\{Y_t, t \geq 1\}$, will be

$$P_n X_{n+h} = P_n Y_{n+h} + \sum_{j=1}^{d+DS} a_j P_n X_{n+h-j} \tag{4.11}$$

We can compute the predictors $P_n X_{n+h}$ recursively from $h = 1, 2, \dots$ from equation (4.11), if we note that $P_n X_{n+1-j} = X_{n+1-j}$ for each $j \geq 1$.

The prediction mean squared error is given as

$$\sigma_n^2(h) = E(X_{n+h} - P_n X_{n+h})^2 = \sum_{j=0}^{h-1} \left(\sum_{r=0}^{j} \chi_r \theta_{n+h-r-1, j-r} \right)^2 v_{n+h-j-1} \tag{4.12}$$

where θ_{nj} and v_n are obtained by applying the innovations algorithm to the differenced series $\{Y_t\}$ and

$$\chi(z) = \sum_{r=0}^{\infty} \chi_r z^r = \left[\phi(z)\Phi(z^s)(1-z)^d(1-z^s)^D\right]^{-1}, |z| < 1 \tag{4.13}$$

The prediction mean squared error can be approximated for large n, considering $\theta(z)\Theta(z^s)$ is nonzero for all $|z| < 1$ as

$$\sigma_n^2(h) = \sum_{j=0}^{h-1} \psi_j^2 \sigma^2 \tag{4.14}$$

where

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)\Theta(z^s)}{\phi(z)\Phi(z^s)(1-z)^d(1-z^s)^D}, |z| < 1$$

https://doi.org/10.26219/heal.aueb.6514

Chapter 5

Methodology

In this section we present the means by which we make comparisons between the estimated model, as well as the process followed for the simulation study, before presenting the results of this research.

5.1 Evaluation criterion and comparisons

To evaluate how the models perform, we need an evaluation metric which will be used to evaluate the out-of-sample forecasts of each estimated model. The Mean Squared Prediction Error (MSPE) is used as the metric for this thesis. The objective is to find the model which corresponds to the lowest MSPE, in order to assess how the specification of certain unit roots affects the models' predictive performance.

To depict the differences between these models we compare boxplots of these forecast errors. Besides the comparison of the forecast errors, an additional comparison is made using the nominal standard error of the forecast as estimated by the R function `predict` divided by the square root of the mean squared prediction error, as estimated by the simulation. For the DGP model (defined in the simulation), this ratio is expected to be around 1, therefore for the rest of the estimated models we base the comparison based on how much this ratio fluctuates around 1, once again for a variety of different forecast horizons.

5.2 Simulation process

In order to produce the sample of interest we follow specific steps. First, it is important to define the model with specific characteristics, based on the sample we wish to produce.

https://doi.org/10.26219/heal.aueb.6514

This means, we define the innovations variance, the number of autoregressive, moving-average and non-stationary roots, as well as the inverse roots of these polynomials whether seasonal or not. The inverse roots are needed in order to calculate the coefficients for each part. Since some of these roots can be real, while others can be complex numbers, we present below how the coefficients are obtained based on the two occasions.

For the complex case, we will use the representation

$$z_i = r_i e^{\pm\omega_i i}, \qquad i = 1, \ldots, C$$

where C is the total number of conjugate pairs (and therefore 2C the total number of complex roots), r_i is the modulus of z_i and ω_i its frequency.

The real inverse roots are denoted as

$$z_i = r_i, i = 1, \ldots, R$$

Depending on what type of root we define, the coefficients are calculated based on the following formulas:

For real roots, the coefficients result from

$$\Phi(B) = 1 - r\cos(\omega)B$$

and for the complex conjugate roots, the characteristic polynomial is

$$\Phi(B) = (1 - re^{-\omega i})(1 - re^{\omega i}) = 1 - 2r\cos(\omega)B + r^2B^2$$

The process of calculating the coefficients continues for as long as there are roots present and each time is it updated until the final coefficients of the each part are calculated. It is important to note that we are only interested in the autoregressive part, since no moving-average roots are taken into account for the purpose of this thesis.

After we have obtained the coefficients of the autoregressive polynomial, we proceed on the simulation of a sample based on the order of the model we defined and the values of its coefficients, whether their are stationary or non-stationary.

In total, 1000 samples are produced, each one of length 300, which corresponds to 25 years of monthly observations. For each iteration, the overall sample is split into two sets, training and test, for which the former is used in the estimation process and the latter in order to make predictions which are then used for the evaluation of the models' performance. For the training set we have chosen to split the data from observation $t = 1$ to $t = 276$, while using as a validation set for the next 24 monthly observations, since we are interested in obtaining forecasts for the next two years.

For each simulation, the prediction errors of each estimated model are calculated and stored, for different forecast horizons up to two years, using the validation set. They are then squared and averaged to obtain the MSPE for all timesteps. Additionally, we obtain the estimated variance of the forecast errors which will then be compared with the true variance of the prediction errors which is calculated during the simulation process. To obtain the forecasts we used the R function `predict` which provides estimated forecasts as well as estimated standard errors of the forecast error of any SARIMA model fit.

Different models will be used, where each time we will generate a sample under a specific "true" model and estimate under the remaining ones, in order to then make predictions and compare them. However only two cases are presented analytically. The goal is to generate samples based on a model with a specific number of non-stationary seasonal root and estimate models with either a larger number of non-stationary roots or an equal or larger number of stationary roots, to see whether their forecasts will lead to same or similar results as if we had obtained forecasts from the original time series.

For the ***first case***, we consider in the DGP a model that has only one non-stationary seasonal root at the frequency $2\pi/12$, with period equal to 12, and no other roots. Specifically, the model of interest is therefore defined as:

$$(1-\sqrt{3}B+B^2)X_t=\varepsilon_t$$

with the innovations variance set equal to 1.

The models used in the prediction process, after being fitted to synthesize the time series simulated by the DGP process, are the following:

1. the DGP model, allowing one non-stationary root at $2\pi/12$
2. $SARIMA(0,1,0)_{12}$ with 12 non-stationary roots (including the root at zero frequency)
3. $AR(12)$ with twelve stationary roots
4. $AR(2)$ with two stationary roots (which correspond to the number of roots of the true model)

Similarly, in the ***second case***, the DGP model has two non-stationary seasonal roots at the $2\pi/12$ and $4\pi/12$ frequencies, defined as

$$(1-\sqrt{3}B+B^2)(1-B+B^2)X_t=\varepsilon_t$$

with the innovations variance set equal to 1. Here the fitted models are:

1. the DGP model, allowing two non-stationary seasonal roots at the $2\pi/12$ and $4\pi/12$
2. $SARIMA(0,1,0)_{12}$ with 12 non-stationary roots (including the root at zero frequency)

3. $AR(4)$ with four stationary roots (which again correspond to the number of roots of the true model)

4. $AR(12)$ with twelve stationary roots

5.3 Results

In this section we will present the results of how each estimated model performs in comparison to the other models, while taking into account a variety of forecast horizons.

Since the analysis is split into two parts, we first begin with the comparison of the forecast errors using boxplots, as shown in figure 5.1. The horizontal axis represents the different models and the vertical axis shows the error values. Predictions are made for different forecast horizons, starting from 3 months ahead up to 24 months, in 3-month intervals. This way it is easier to examine the models' performance during the one year, while accounting for the months that are multiples of their period, as well as analyze the behavior from year to year.

First case

In figure 5.1, where the "true" model, i.e. the one used in the data generating process, has a seasonal unit root at $2\pi/12$, it is noticeable that the SARIMA model makes a forecast error higher than the rest of the models when the steps ahead differ from the period the SARIMA model would induce. However when reaching the 12-step-ahead forecast, which is the period of the SARIMA model, then it seems to provide better forecast errors than the AR(2) and AR(12) models do and interestingly these errors coincide with the errors of the Unit Root model which was used for simulating the samples. This seems to occur because the cycle of both these models is completed over the course of a year (also shown in figure 5.3). Overall the distribution of the forecast errors of the estimated AR(2) and AR(12) models is closer to that of the "true" model, with AR(2) resulting in slightly smaller prediction errors. It can be seen that through the duration of the year, both these models provide a noticeably smaller range of forecasts errors in comparison to the SARIMA model.

In regard to the variances, as shown in figure 5.2, the results are satisfactory for all models in question. For the SARIMA model the ratio remains always around 1 (which is the innovations variance and what we would expect from our "true" model), specifically with the median being above 1 in every step ahead. This would suggest that the variance is accurately estimated by the SARIMA model, even in the cases for which the variance

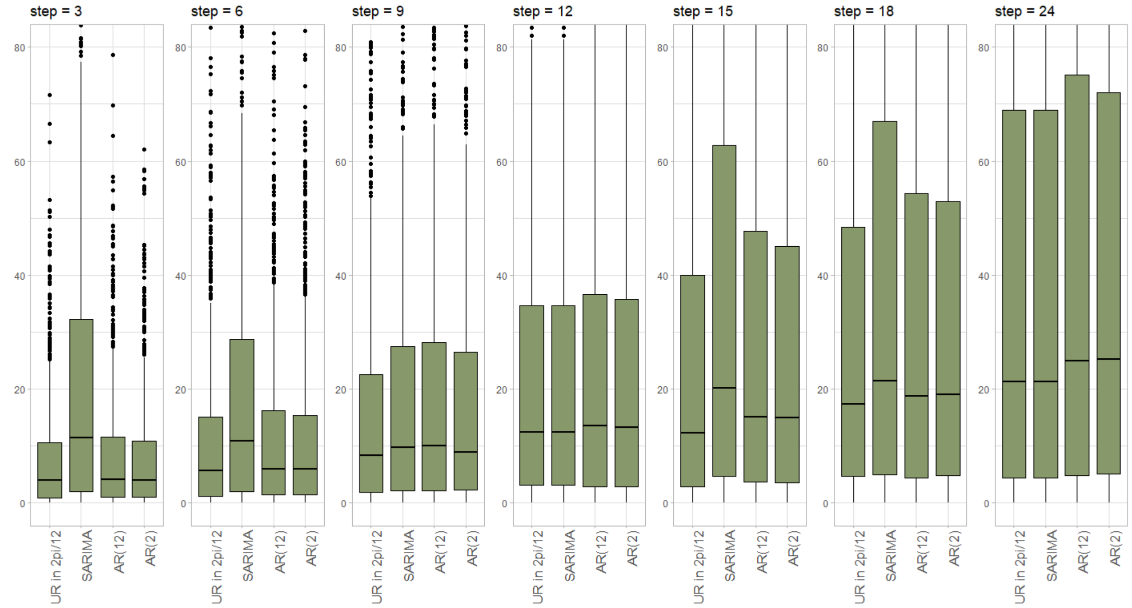

Figure 5.1: Forecast errors when series is generated by a process with a non-stationary seasonal root at $2\pi/12$

of the prediction error is higher than for the DGP model. For the stationary AR(2) and AR(12) models, it can be seen that the estimated variance over the true variance stays under 1, getting lower the more steps further into the future, with the AR(2) model having, however, a less underestimated variance of the prediction error compared to the 12-parameter AR model. Note that the estimated variance that each model calculates is as if the model knows the true parameters, therefore these results may be due to the portion of the variance, in the estimation process of the 2 or 12 parameters, that is unaccounted for when calculating the nominal variance.

To get a better understanding of how these models perform in conjunction with the prediction errors boxplots, a plot of the predictions of each estimated model along with their prediction intervals is depicted in figure 5.3, taken from a random sample, where it can be seen that the AR(2) model produces more similar forecasts to the "true" model (shown in the darker shades, which overlap), while the estimated SARIMA model produces larger intervals that can be quite different at each time point compared to the other two models. It is noticeable however, as is also in figure 5.1, that at specific steps the SARIMA model's predictions overlap with the ones of the "true" model, specifically when the forecast horizon reaches multiples of 12 months, i.e. when the models' cycles are completed.

Second case

In the case where the DGP model has two non-stationary seasonal roots at $2\pi/12$ and

https://doi.org/10.26219/heal.aueb.6514

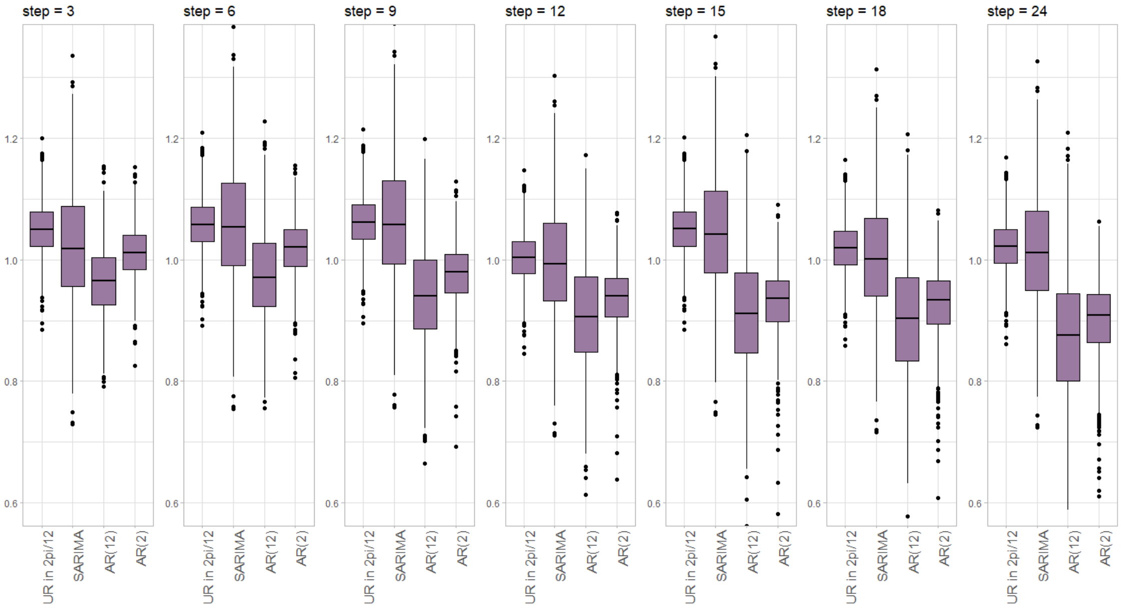

Figure 5.2: Estimated variance of forecast errors divided by true variance of the errors for DGP model with a non-stationary seasonal root at $2\pi/12$

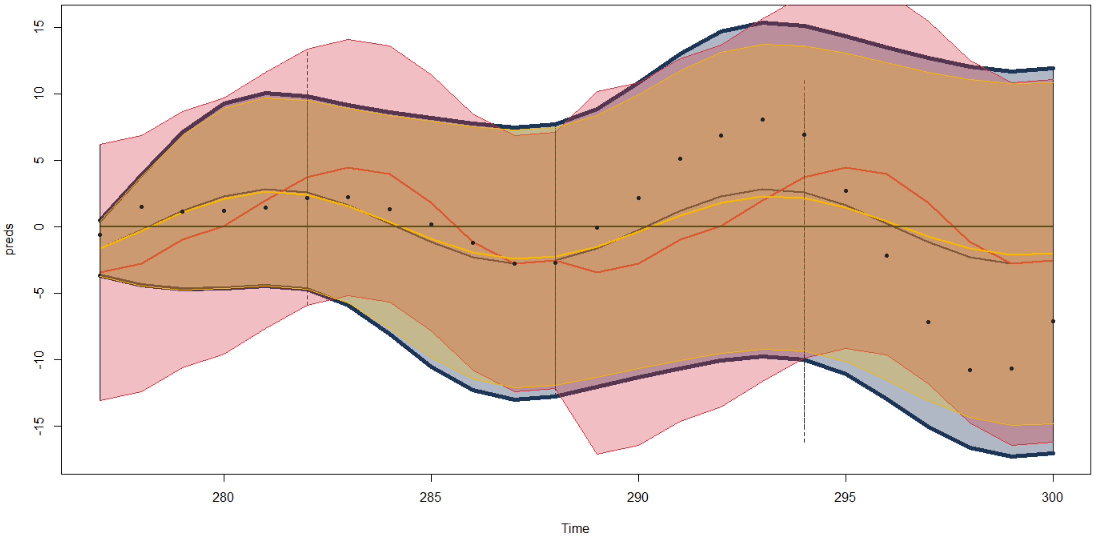

Figure 5.3: Forecasts of the estimated models with their prediction intervals. Blue and yellow shaded ares are the forecast intervals of the UR in $2\pi/12$ and AR(2) models respectively, which overlap (orange shade), and red shaded area is for the SARIMA model.

$4\pi/12$, the results are fairly similar to the ones of the above case. Throughout the first year, the forecast errors of the SARIMA model appear to be higher than the errors of the two other models in question. Only in the 12-step-ahead forecast does the SARIMA model provide better prediction errors than the AR(4) and AR(12) models. Figure 5.6 shows in better detail how each model behaves in terms of the predictions.

As for the variances, the SARIMA models stays around 1 in every step ahead, with

https://doi.org/10.26219/heal.aueb.6514

a bigger range, though the results still suggest, as in the first case, that the variance is being estimated more accurately than what the rest of the models estimate. Once again, the AR(4) model provides a better estimated variance, while still being underestimated, than the AR(12) model.

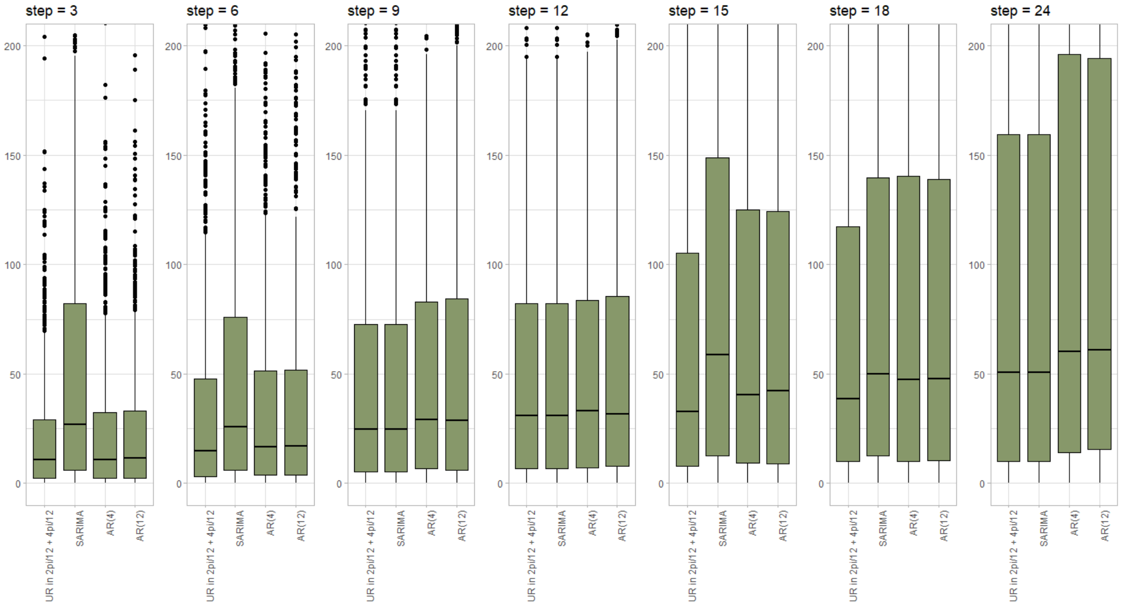

Figure 5.4: Forecast errors when series is generated by a process with two non-stationary seasonal roots at $2\pi/12$ and $4\pi/12$

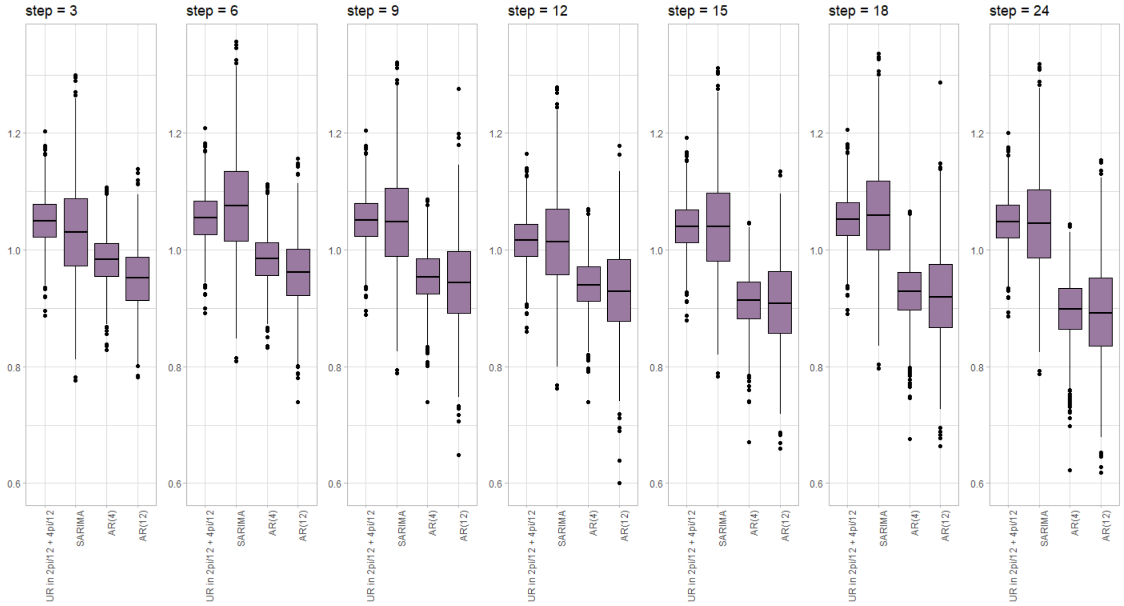

Figure 5.5: Estimated variance of forecast errors divided by true variance of the errors for DGP model with two non-stationary seasonal roots at $2\pi/12$ and $4\pi/12$

https://doi.org/10.26219/heal.aueb.6514

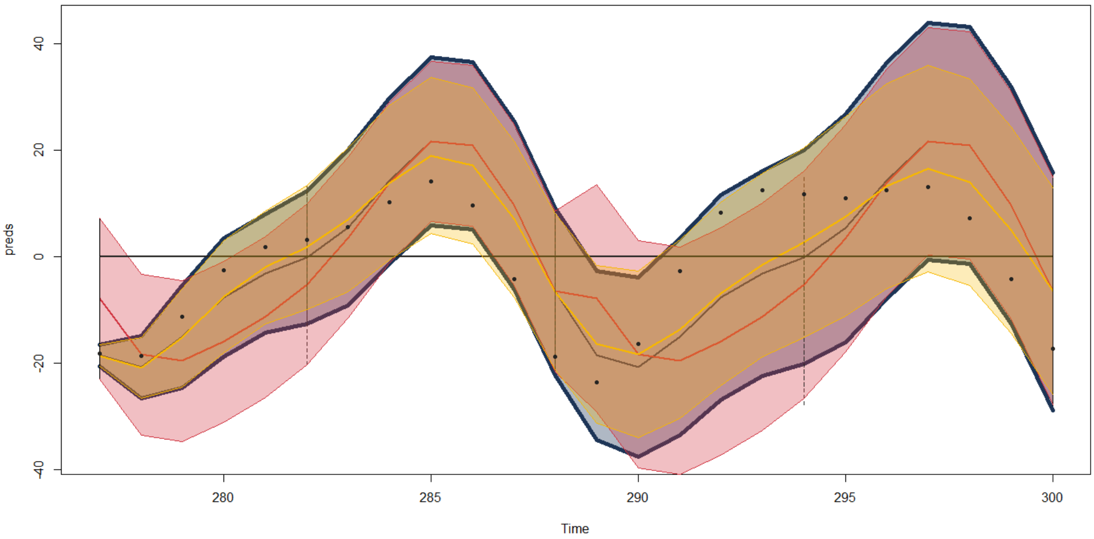

Figure 5.6: Forecasts of the estimated models with their prediction intervals. Blue and yellow shaded ares are the forecast intervals of the UR in $2\pi/12 + 4\pi/12$ and AR(2) models respectively, which overlap (orange shade), and red shaded area is for the SARIMA model.

Different cases of models where also taken into account in the DGP, with the estimated models remaining the same as in the first case. The analysis was made for a model with one non-stationary seasonal root at $4\pi/12$ which has a period of 6 months, as well as a model with a non-stationary root at frequency $8\pi/12$, with a 3-month period . This way we examined how the estimated models with more roots behaved when the original time series was generated by a seasonal unit root at only one specific frequency. The results were similar to the ones shown above for the model with a unit root at $2\pi/12$, with different ranges of forecasts errors, but with the more general picture remaining the same. Once again, the stationary autoregressive models seemed to give smaller errors throughout the year with an underestimated variance the more steps ahead into the future, while on the other hand the SARIMA model performed better when its 12-month cycle was complete, overlapping with the model used in the DGP. The boxplots of the prediction errors and the variances are shown in the Appendix.

Chapter 6

Conclusions

As seen in the previous Chapter, in both cases, all models seem to provide adequate predictions despite the fact that a larger number of roots is introduced during the estimation process, than what was originally assigned in the data generating process. Even though the AR(2)/AR(4) and AR(12) models provided smaller prediction errors, their variances were constantly underestimated, perhaps due to the fact that these models calculate the estimated variance assuming to know what the true parameters are, therefore ending up estimating more parameters. In contrast, the variance of the SARIMA model was more stable and similar as that of the model used in the DGP, no matter the forecast horizon. What is interesting is the fact, that the prediction errors of the SARIMA model seem to fall back to the same level as the "true" model, when the forecast horizon completes one cycle, which in the cases examined is equal to 12 months.

To conclude, it is up to the researcher to decide what type of model should be used in cases where there is presence of non-stationary roots only at specific frequencies, however the assumption that a SARIMA model with all seasonal roots present is the right choice, as often portrayed in the bibliography, should be avoided, since it forces the presence and therefore estimation of more non-stationary roots, which can lead to unnecessary larger prediction errors.

https://doi.org/10.26219/heal.aueb.6514

Bibliography

[1] Ansley, C. [1979], ‘An algorithm for the exact likelihood of a mixed autoregressivemoving-average process’, *Biometrika* **66**, 59–65.

[2] Beaulieu, J. J. and Miron, J. A. [1993], ‘Seasonal unit roots in aggregate u.s. data’, *Journal of Econometrics* **55**(1–2), 305–328.

[3] Boshnakov, G. N. and Halliday, J. [2022], ‘Partial autocorrelation parameterisation of models with unit roots on the unit circle’.

[4] Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. [2016], *Time Series Analysis: Forecasting and Control*, Wiley Series in Probability and Statistics, 5th edn, Wiley.

[5] Brockwell, P. J. and Davis, R. A. [1991], *Time Series: Theory and Methods*, Springer Texts in Statistics, 2nd edn, Springer.

[6] Brockwell, P. J. and Davis, R. A. [2002], *Introduction to Time Series and Forecasting*, Springer Texts in Statistics, 2nd edn, Springer.

[7] Clements, M. P. and Hendry, D. [1997], ‘An empirical study of seasonal unit roots in forecasting’, *International Journal of Forecasting* **13**(3), 341–355.

[8] Franses, P. H. [1991], ‘Seasonality, non-stationarity and the forecasting of monthly time series’, *International Journal of Forecasting* **7**(2), 199–208.

[9] Franses, P. H. and Kawasaki, Y. [2004], ‘Do seasonal unit roots matter for forecasting monthly industrial production?’, *Journal of Forecasting* **23**(3), 77–88.

[10] Ghysels, E. and Osborn, D. [2001], *The Econometric Analysis of Seasonal Time Series*, Cambridge University Press.

[11] Gustavsson, P. and Nordström, J. [2001], ‘The impact of seasonal unit roots and vector arma modelling on forecasting monthly tourism flows’, *Tourism Economics* **7**(2), 117–133.

[12] Hamilton, J. D. [1994], *Time Series Analysis*, Princeton University Press.

https://doi.org/10.26219/heal.aueb.6514

[13] Hylleberg, S., Engle, R., Granger, C. and Yoo, B. [1990], 'Seasonal integration and cointegration', *Journal of Econometrics* **44**(1–2), 215–238.

[14] Maravall, A. [1993], Unobserved components in economic time series.

https://doi.org/10.26219/heal.aueb.6514

Appendix A

Other cases

Below are the boxplots of the prediction errors and the estimated variances, of the two other cases of models used during the simulation process.

The first model has a non-stationary root at frequency $4\pi/12$ and is defined as

$$(1 - B + B^2)X_t = \varepsilon_t$$

with the innovations variance set equal to 1. The results are the following

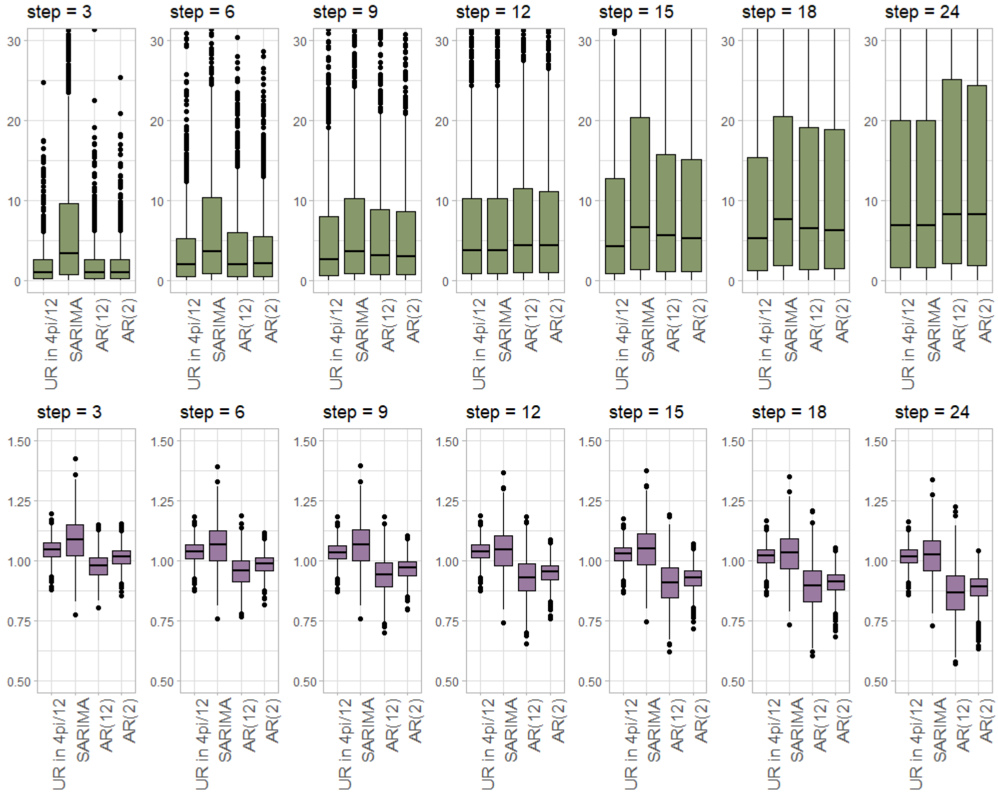

Figure A.1: Forecast errors when series is generated by a process with non-stationary seasonal root at $4\pi/12$

https://doi.org/10.26219/heal.aueb.6514

For the second case, the sample was generated by a model with a a non-stationary root at frequency $8\pi/12$ of the form

$$(1 + B + B^2)X_t = \varepsilon_t$$

The results are shown below

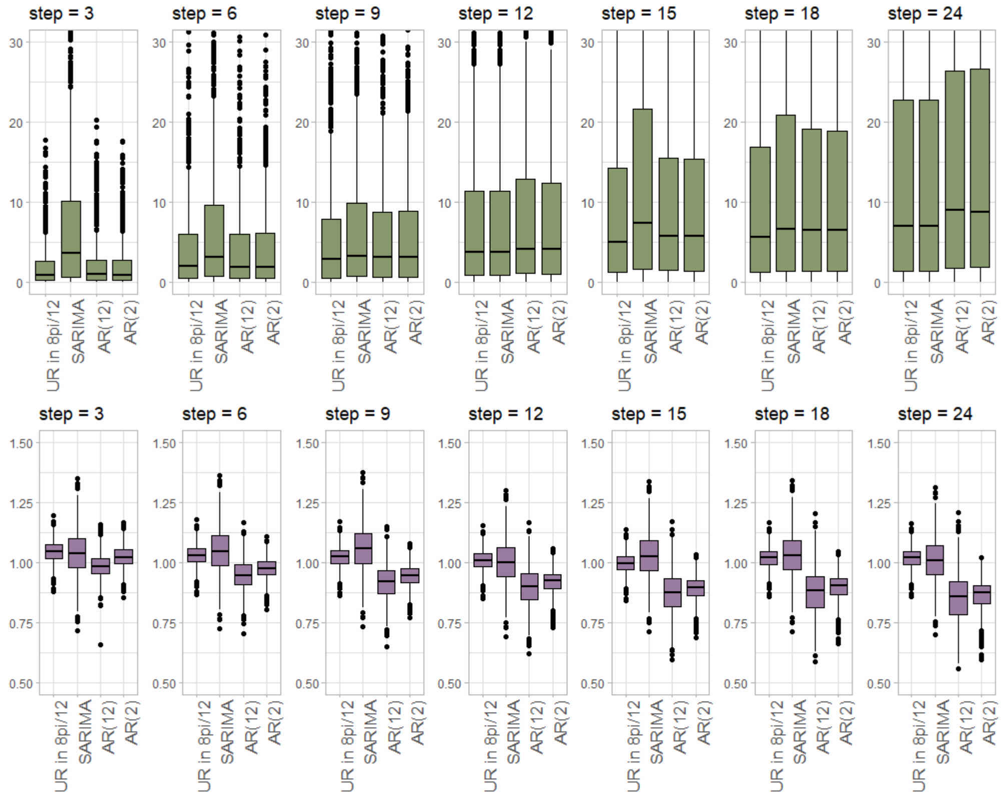

Figure A.2: Forecast errors and estimated variance divided by the true variance of the errors when series is generated by a process with non-stationary seasonal root at $8\pi/12$