UPPSALA
UNIVERSITET

Forecast Comparison of Models Based on SARIMA and the Kalman Filter for Inflation

By Fredrik Nikolaisen Sävås

Independent Thesis Advanced Level
Department of Statistics
Uppsala University

Supervisor: Lars Forsberg

June 5, 2013

Abstract

Inflation is one of the most important macroeconomic variables. It is vital that policy makers receive accurate forecasts of inflation so that they can adjust their monetary policy to attain stability in the economy which has been shown to lead to economic growth. The purpose of this study is to model inflation and evaluate if applying the Kalman filter to SARIMA models lead to higher forecast accuracy compared to just using the SARIMA model. The Box-Jenkins approach to SARIMA modelling is used to obtain well-fitted SARIMA models and then to use a subset of observations to estimate a SARIMA model on which the Kalman filter is applied for the rest of the observations. These models are identified and then estimated with the use of monthly inflation for Luxembourg, Mexico, Portugal and Switzerland with the target to use them for forecasting. The accuracy of the forecasts are then evaluated with the error measures mean squared error (MSE), mean average deviation (MAD), mean average percentage error (MAPE) and the statistic Theil's U. For all countries these measures indicate that the Kalman filtered model yield more accurate forecasts. The significance of these differences are then evaluated with the Diebold-Mariano test for which only the difference in forecast accuracy of Swiss inflation is proven significant. Thus, applying the Kalman filter to SARIMA models with the target to obtain forecasts of monthly inflation seem to lead to higher or at least not lower predictive accuracy for the monthly inflation of these countries.

KEYWORDS: Inflation, SARIMA model, Hyndman-Khandakar algorithm, State-Space models, Kalman filter, Forecast comparison, Diebold-Mariano test.

Table of Contents

List of Figures

List of Tables

1 Introduction

Forecasts of inflation is a central part of the information that policy makers need to base their choice of monetary policy on. To control inflation is very important since it is connected to employment, the public's risk-taking and will to invest. In the long term a stable and low inflation can be said to be connected to growth, efficiency and stability. Shifts in inflation can have a negative effect on the public's belief for future economic stability. It is therefore important that policy makers get good forecasts of inflation. This will ensure that they can adjust their policies to counteract any negative effects concerned with future shifts of inflation and in that way stabilize the output and employment in the economy. (Bernanke and Mishkin, 1997)

The problem in obtaining those forecasts is how to know which model to favour when there exist competing models. The choice of model is usually based on passed accuracy, but how do we evaluate if the differences really are statistically significant? Diebold and Mariano (1995, p. 134) discussed the value of economic agents needing to be more critical when deciding what forecasts that they base their decision on.

Month-to-month inflation measured as the percentage change in the consumer price index (CPI) will be the data of interest for this study. It is assumed that seasonality is a prominent part in monthly inflation, that is that the change in CPI can be seen to generally be bigger between some months compared to others.

In 1970, Box and Jenkins introduced a new class of multiplicative linear autoregressive integrated moving average models, ARIMA. This kind of model were considered path breaking by for example Durbin and Koopman (2001, p.46). The seasonal ARIMA model, called SARIMA, is a special case of this model and is applied to seasonal data, for example weekly or monthly which is the case for this thesis (Brockwell and Davis, 1991, p. 323).

It has been proven difficult to produce forecasts of inflation that can be considered more accurate than what is obtained with a simple autoregressive model (Norman and Richards, 2012). SARIMA models have also clearly been widely accepted as good for forecasting monthly inflation. To prove this is not hard considering that many studies have been made with promising results for the SARIMA model. Examples are Saz (2011); Çatik and Karaçuka (2012) who used it for Turkish inflation, Tarno et al. (2012) for Indonesian inflation, Meyler et al. (1998) for Irish inflation, Omane-Adjepong et al. (2013) for Ghanaian inflation and Junttila (2001) for Finnish inflation.

The Kalman filter was developed by Kalman in 1960 and is applied to models written on the so called state-space form. It was first proven as useful for applications in engineering. In 1977 Morrison and Pike stated that a general method for applying the Kalman filter for statistical forecasting has not yet been developed. However the procedure had at that time started to get its roots into the statistical field, for example Rosenberg (1973), Engle (1979) and Harvey and Phillips (1979) found that the Kalman filter could have an

importance for econometric and statistical applications, for example for time series forecasting of economic properties. These were fields that ARIMA modelling at that point dominated (Harvey, 1989, p. 23).

One example of a clear benefit of the Kalman filter is that only the present state estimate and the next observation is required to update the whole system. This can be compared to models estimated with maximum likelihood or ordinary least squares where updating would imply use of the whole history of data (Morrison and Pike, 1977).

The ARMA model has a clear connection to the recursive procedure developed by Kalman (1960). This can be exemplified by the state-space model first being presented for some simple ARMA models in well-read statistical texts for example by Hamilton (1994), Harvey (1989) and Brockwell and Davis (1991). Of further interest is that the SARIMA model can be equivalently presented as a ARMA model. Thus, the SARIMA model can be written as an ARMA on the so called state-space form which makes Kalman filtering possible.

However work where the Kalman filter is applied on SARIMA models does not seem to be very frequent. One comparable study was performed by Hamilton (1985) who modelled inflation and nominal return on one-period bond purchased at date t and redeemed at $t+1$ with the state being assumed to be expected inflation as a bivariate ARMA model. The model was then put in state-space form and the Kalman filter applied to obtain forecasts. The big difference is that this thesis considers the univariate case to only model inflation. The methodology that is used in this thesis does not seem to be widely applied for predicting inflation but Grosswindhager et al. (2011) use it for forecasting system heat load in district heating networks with promising results.

The reason for the limited use of the Kalman filter on SARIMA models is probably the difficulties with how to specify the prior matrices and doubt about the benefits of the procedure. Hence why not try to evaluate if forecasts could actually get better by applying it?

Thus, the purpose of this thesis is to compare the forecasting performance between a SARIMA model estimated with all of the data and an equally specified SARIMA model with parameters that are first estimated with part of the data, then put in state-space form, and then Kalman filtered with the rest of the data. This study want to evaluate if there might be any benefits associated with applying the Kalman filter on SARIMA models with the purpose of forecasting real data like monthly inflation. Another thing is to try out and possibly get a reason to elevate the simple and clear method for attaining state matrices which puts a SARIMA model in state-space form, that is used in this thesis.

This will be done for the monthly inflation of Luxembourg, Mexico, Portugal and Switzerland. The reason for including a number of countries is to get more evidence regarding the benefits of any given model. The total number of observations used for

estimating the SARIMA and its ratio with the number of observations that are saved for filtering will be different for the countries and it will be interesting to evaluate if this has an effect on the results.

The purpose of this thesis can be summarized with the following question:
Does applying the Kalman filter to a SARIMA model (estimated with part of the data with the remaining observations used for filtering) lead to better forecasts compared to a SARIMA model that is fitted with all of the data?

There are some delimitations to this thesis. The two models are chosen to make the effect of the Kalman filter on the forecasting accuracy as clear as possible. That is for the work in this thesis only the SARIMA and the Kalman filtered SARIMA models are used even though there might exist other and better models for example based on the Holt-Winters approach (Omane-Adjepong et al., 2013), autoregressive fractionally integrated moving average (ARFIMA), fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH), unobserved components models (UCM) and artificial neural network models (ANN) (Çatik and Karaçuka, 2012). In-sample forecast accuracy will not be considered. The reason for this is that it would be difficult to compare these values since the data set will be divided into two parts for the Kalman filtered model.

This study will delimit from the possibility that structural breaks occur in any of the used time series processes. This could be a disputed assumption considering that structural breaks have been proven to sometimes hold relevance for SARIMA modelling of monthly inflation (Junttila, 2001; Saz, 2011). One of the major benefits with the Kalman filter is the possibility to include time dependent coefficients into the model but this will not be done in this thesis (Morrison and Pike, 1977). Another benefit is that it works even if there exist missing data (Petris et al., 2009, p. 59). However the case of missing data and how it would impact the performance of the model forecasts will not be evaluated. The Kalman smoother is closely related to the Kalman framework and it is used to draw inference on the state vector for example for any of its past values (Hamilton, 1994, p. 394). Thus it would be assumed that the state vector is of interest but this will not be the case for this study.

The remainder of this thesis is organized as follows. Section 2 introduces the modelling strategy used for the two models. Section 3 presents theory about the models, describes the chosen tests and how the forecasting accuracy is evaluated and compared. In Section 4 descriptive statistics for the data sets of monthly inflation are presented. Section 5 contains the results for the identification, estimation, diagnostic checking and forecasting of both models. Finally Section 6 concludes the thesis by summarizing the results and by discussing some details.

2 Method

First the methodology related to the SARIMA model is presented, then the method for the Kalman filtered SARIMA model and last how their specific forecast accuracies will be compared.

2.1 SARIMA Model

Box and Jenkins bases the model selection on three stages. Identification, estimation and diagnostic checking, presented as in Figure 1. Thus, the first step is to identify the orders of the SARIMA$(p, d, q) \times (P, D, Q)_s$, the next to estimate the model and then to perform diagnostic checks on the residuals to evaluate if the model is well-fitted. If the model fails these diagnostic checks then the only option is to return to the identification stage with the target of finding a better model. It should be noted that the integration orders d and D are specified to make the series stationary. (Box and Jenkins, 1976, p. 19)

The data sets are first inspected visually in search of structural breaks since such breaks has been shown to make it impossible to obtain stationarity through differencing. Thus only a part of the time series which has a relatively constant mean and appear homoscedastic will be chosen (Harvey, 1989, p. 81). This is clearly a subjective strategy but arguably a quite reasonable one.

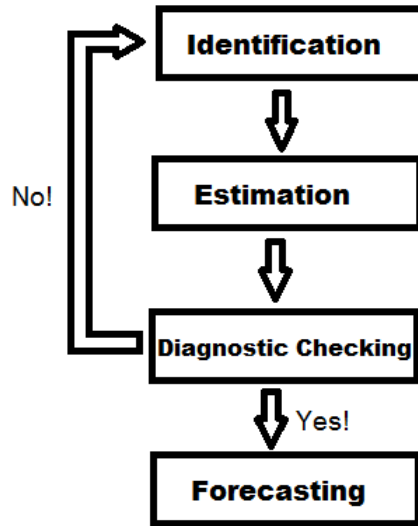Figure 1: The Box-Jenkins procedure (Box and Jenkins, 1976, p. 19)

The identification part then begins by finding the appropriate order of integration in level, that is d. This order will be found with two unit root tests, the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. Kwiatkowski et al. (1992) state that these two tests can be said to complement each other. There are other

tests like the Phillips-Perron test but Davidson and MacKinnon (2003, p. 623) report that it performs worse in finite samples compared to the augmented Dickey-Fuller test.

Another popular test is a unit root test developed by Zivot and Andrews (1992) which is used to test the null hypothesis of a unit root with the alternative of trend stationarity and at least one structural break. The reason for not including this test is that all data included in this study are assumed free of such breaks. Thus, this test would not yield an informative result.

The integration order in season, D, is found with the Canova-Hansen (CH) test and the Hylleberg-Engle-Granger-Yoo (HEGY) test. These tests are used to evaluate the seasonal stationarity of each time series and have been shown to complement each other (Hylleberg, 1995). Further Ghysels et al. (1994) performed a Monte Carlo experiment to compare some seasonal unit root tests and concluded that the HEGY test performs best of all included tests. Both tests are performed with programs integrated in the R package uroot.

The next part is to find find the autoregressive orders p and P and the moving average orders q and Q. The correlogram is first used to make guesses for appropriate orders (Box and Jenkins, 1976, p. 323-324). However this procedure is considered subjective for mixed and seasonal processes. To make the model selection less subjective some frequently used likelihood based information criterions are applied. These are the Akaike information criterion (AIC), the AIC with correction for small samples (AICc) and the Bayes information criterion (BIC). (Harvey, 1989, pp. 80-81)

It has been shown that AIC has a tendency to choose a model that is over-parametrized (Hurvich and Anderson, 1989). Further Burnham and Anderson (2004) suggest that AIC and AICc should be valued over BIC and Brockwell and Davis (1991, p. 273) propose that AICc is most fit for selecting orders of SARIMA models. Thus, the model selected by AICc will be most valued. However finding the model that minimizes each criterion would be very time consuming if every model is compared. Fortunately Hyndman and Khandakar developed an algorithm in 2008 that can be used to speed up this selection process. The HK-algorithm has previously been used by Saz (2011) to identify SARIMA models for Turkish monthly inflation.

It is now assumed that a tentative SARIMA$(p, d, q) \times (P, D, Q)_s$ has been identified. Thus, the next step is to estimate it with maximum likelihood and then perform diagnostic checks on the residuals. For a good fit these residuals should be distributed as Gaussian white noise, that is be random, homoscedastic and normal. The diagnostic checking is first performed visually with the standardized residuals and sample autocorrelation function. (Brockwell and Davis, 1991, pp. 307-309)

Brockwell and Davis (1991, p. 310) further propose the use of statistical tests of the residuals normality and randomness. They further suggest the Ljung and Box (1978) (BL) test for randomness. The Jarque and Bera (1980) (JB) test is chosen to test the normality

of the residuals. It was confirmed as functional by a Monte Carlo study performed by Bai and Ng (2005).

The model should pass all these diagnostic checks to be considered well-fitted and appropriate for forecasting. In this study the first two visual checks will mostly be used to check if the residuals diverges strongly from white noise. The BL and JB tests will hold most weight in deciding if any model should be rejected, implying a return to the identification stage of the Box-Jenkins procedure.

2.2 Kalman Filtered SARIMA Model

The Kalman filter can be applied to any model that is written in state-space form and then uses noisy data to perform recursive updates on an unobserved state vector to minimize its mean squared error, MSE (Kalman, 1960).

Commandeur et al. (2011) suggest that the following programs STAMP, R, MATLAB, REGCMPNT, SAS, EViews, GAUSS, Stata, RATS, gretl, and SsfPack can be applied for state-space modelling with the Kalman filter. They then recommend the use of the DLM package from R which is also the program and package that is used for the work done in this thesis. The Kalman filter algorithm that is integrated in this package ensures numerical stability by using singular value decomposition on the covariance matrices in the way proposed by Zhang and Li (1996).

The method that is used for this second model in this thesis is to use a subset of the data consisting of early observations to estimate the SARIMA that has been identified in the previous part. It should be noted that this subset should consist of at least 50 observations since that is needed for efficient estimation (Box and Jenkins, 1976, p. 18). Thus, it is assumed that the model estimated with use of the whole data set is appropriate also for the case of this subset of observations. The fact that this SARIMA model can be equivalently written as an ARMA model is then used to specify matrices to put the model in state-space form (Box et al., 2008, p. 379). The Kalman filter is then applied to the remaining part of the data and forecasts found. (Hamilton, 1985; Grosswindhager et al., 2011)

The residuals from the recursive Kalman filter procedure are called innovations and should then be extracted and used to diagnostic check the model. Thus this means to evaluate the appropriateness of applying the Kalman filter to this model and data. The innovations are assumed to be distributed as Gaussian white noise, that is they should be serially independent, homoscedastic, and normal (Petris et al., 2009, p. 93). This assumption is the same as for the SARIMA model residuals and the innovations will therefore undergo the same diagnostic checks as those residuals.

2.3   Comparison of Forecast Performance

For this last part it is assumed that both models have passed the diagnostic checks leading to both models being considered fit to use for forecasting. Forecasts are then obtained with the forecast function in the forecast package from R for the SARIMA models (Hyndman and Khandakar, 2008). While for the Kalman filtered model the dlmForecast function from package dlm is used (Petris et al., 2009).

The two models forecast performance are then compared with some frequently used error measures. Petris et al. (2009, pp. 98-99) suggest that examples of such measures are the mean squared error (MSE), the mean average deviation (MAD), the mean average percentage error (MAPE) and the statistic Theil's U. The error measures and Theil's U are used to get an indication of which model that is more accurate but they do not yield significant evidence of any differences in predictive ability (Diebold and Mariano, 1995). The importance of evaluating this is emphasized by Mariano and Preve (2012) who state that when multiple models have been used for forecasting it is important to use tests to evaluate the significance of the model differences in accuracy. They then suggest that the DM test developed by Diebold and Mariano in 1995 is fit to use for that evaluation.

3 Theory

In this section the SARIMA model and the Kalman filtered SARIMA model will be presented and also theory surrounding how these will be identified, estimated, diagnostic checked and used for forecasting.

3.1 The SARIMA Model

The seasonal autoregressive integrated moving average (SARIMA) model is applied to the time series y_t with the following expression (Brockwell and Davis, 1991, p. 323)

$$\Phi(L^s)\phi(L)\Delta^d\Delta_s^D y_t = \theta_0 + \Theta(L^s)\theta(L)\epsilon_t. \tag{1}$$

These models are specified as SARIMA$(p, d, q) \times (P, D, Q)_s$, where s is the seasonal length, for example $s = 12$ for monthly and $s = 4$ for quarterly data, L is the lag operator and ϵ_t is assumed to be a Gaussian white-noise process with mean zero and variance σ^2. The difference operator is Δ^d where d specifies the order of differencing and the seasonal difference operator is Δ_s^D where D is the order of seasonal differencing. The difference operators are applied to transform the observed non-stationary time series y_t to the stationary process y_t^* with the following equation (Brockwell and Davis, 1991, p. 323)

$$y_t^* = (1 - L)^d(1 - L^s)^D y_t. \tag{2}$$

Further $\phi(L)$ and $\theta(L)$ are defined as the following polynomials in the lag operator (Brockwell and Davis, 1991, p. 323)

$$\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p, \tag{3}$$

$$\theta(L) = 1 + \theta_1 L + \ldots - \theta_q L^q. \tag{4}$$

The seasonal polynomials $\Phi(L^s)$ and $\Theta(L^s)$ in the lag operator are specified as follows (Brockwell and Davis, 1991, p. 323)

$$\Phi(L^s) = 1 - \Phi_1 L^s - \ldots - \Phi_p L^{Ps}, \tag{5}$$

$$\Theta(L^s) = 1 + \Theta_1 L^s + \ldots - \Theta_p L^{Qs}. \tag{6}$$

For the state-space model and forecasting of integrated processes the fact that the observed variable y_t can be replaced by the differenced variable y_t^* as in Equation (2) is used. Box et al. (2008, p. 379) propose that the SARIMA$(p, d, q) \times (P, D, Q)_s$ for y_t can be

seen as a special form of the equivalent representation of y_t^* as an $\mathrm{ARMA}(p+sP, q+sQ)$ written as

$$\phi(L)^* y_t^* = \theta_0 + \theta(L)^* \epsilon_t. \tag{7}$$

The AR part $[\phi(L)^*]$ in this model is derived by multiplying the autoregressive lag polynomials $\phi(L)$ and $\Phi(L^s)$ from Equations (3) and (5). Hence it is

$$\phi(L)^* = \phi(L)\Phi(L^s) = (1 - \phi_1 L - \ldots - \phi_p L^p)(1 - \Phi_1 L^s - \ldots - \Phi_p L^{Ps}). \tag{8}$$

For the MA part $[\theta(L)^*]$ the corresponding thing is implied with the moving average lag polynomials $\theta(L)$ and $\Theta(L^s)$ which follow from Equations (4) and (6), hence

$$\theta(L)^* = \theta(L)\Theta(L^s) = (1 + \theta_1 L + \ldots - \theta_q L^q)(1 + \Theta_1 L^s + \ldots - \Theta_p L^{Qs}). \tag{9}$$

3.2 Tests of Level Stationarity

The first part in the identification process is to investigate the stationarity in level to decide the integration order d. This is done with the ADF and KPSS tests and how it is done will be described in this section. It can be noted that the concepts of stationarity and unit roots are explained a bit further in the Appendix.

3.2.1 The ADF test

The Augmented Dickey-Fuller (ADF) test is used to test the null hypothesis of a unit root against the alternative of stationarity and it is based on the following model (Banerjee et al., 1993, p. 108)

$$\Delta y_t = \alpha + \beta t + (\rho - 1)y_{t-1} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \tag{10}$$

where α is a constant, β the coefficient of a simple time trend, ρ is the parameter of interest, Δ is the first difference operator, δ_i are parameters and p the lag order of the autoregressive process. The choice of including the intercept and/or the time trend should be made beforehand. The lagged differenced variables are included to account for possible serial correlation that would otherwise appear in the error term ϵ_t which is assumed to be approximately a white noise process (Banerjee et al., 1993, p. 108). The lag length p is decided in the default way performed by EViews 7 which is to use the model which implies the lowest BIC, while using the maximum lag length of 14 (Schwert, 2009, p. 381).

What is tested is the null hypothesis of a unit root that is that $\rho = 1$ against the alternative hypothesis of stationarity that is $|\rho| < 1$. The test statistic that is used is based on the t-type statistic

$$DF_\tau = \frac{\hat{\rho} - 1}{SE(\hat{\rho})} \tag{11}$$

where the estimated value of the test statistic should be compared to the value of the relevant critical value of the Dickey-Fuller test (Banerjee et al., 1993, p. 108).

3.2.2 The KPSS test

The Kwiatkowski et al. (KPSS) test was developed in 1992 and assumes the following model

$$y_t = \xi t + r_t + \epsilon_t \tag{12}$$

where ξ is the coefficient of a simple time trend, $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ and r_t is a random walk, that is, $r_t = r_{t-1} + u_t$, where u_t is a white noise process with mean zero and variance σ_u^2 and r_0 is considered to be the intercept. It is optional to include or not include the time trend.

The null hypothesis that $\sigma_u^2 = 0$ implies testing that the time series is either level ($\xi = 0$) or trend stationary ($\xi \neq 0$) against the alternative that it is non-stationary. The test statistic is then derived by first fitting y_t depending on only an intercept or an intercept and a trend. The resulting residuals e_t are then used to derive a consistent estimate of the variance

$$s^2(l) = T^{-1} \sum_{i=1}^{T} e_t^2 + 2T^{-1} \sum_{s=1}^{l} w(s, l) + \sum_{t=s+1}^{T} e_t e_{t-s}. \tag{13}$$

In this equation $w(s, l) = 1 - \frac{s}{1+l}$ is the Bartlett window, which is the default in EViews 7, and guarantees that the estimated variance is non-negative. The bandwidth l is decided by the Newey-West automatic with the Bartlett kernel. (Schwert, 2009, p. 383)

The next step is to derive the partial sum series of the residuals, that is

$$S_t = \sum_{t=1}^{T} e_t \tag{14}$$

which is derived for $t = 1, ...T$. The results in Equations (13) and (14) are then used to derive the Lagrange multiplier based KPSS test statistic

$$\eta_\mu = \frac{\eta_\mu}{s^2(l)} = T^{-2} \frac{\sum S_t^2}{s^2(l)} \tag{15}$$

where the critical values can be found in Kwiatkowski et al. (1992).

3.3 Tests of Seasonal Stationarity

The seasonal unit root tests that will be described in this section is the CH test and the HEGY test. These are used to find the order of seasonal integration D.

3.3.1 The CH test

The Canova and Hansen (CH) test was developed in 1992 and is used to test the null hypothesis that the time series process is stationary with deterministic seasonality against the alternative that it has a seasonal unit root. It is closely related to the KPSS test since both are based on the Lagrange Multiplier (LM) statistic. This test assumes the following model

$$y_t = \mu + x_t'\beta + S_t + e_t \tag{16}$$

where y_t is the modelled time series, x_t is a vector of explanatory variables which can be lagged values of y, S_t is a deterministic seasonal component of period $s = 12$ for monthly data and $e_t \sim (0, \sigma^2)$ is white noise and uncorrelated with x_t and S_t. It should be noted that the time series that is used for this test is assumed stationary in level.

If no explanatory variables are included then the error e_t will be the difference between the modelled process y_t and its seasonal component S_t. The requirements of e_t are not strict but it should not appear to have tendencies for serial correlation, heteroskedasticity or seasonal behaviour.

Further the seasonal component can be written like this

$$S_t = d_t'\alpha \tag{17}$$

where d_t is a seasonal dummy indicator for the 12 lags and α is a parameter vector representing the seasonal effects. The seasonal component can then be equivalently written on a trigonometric representation, that is

$$S_i = \sum_{j=1}^{q} f_{jt}'\gamma_j \tag{18}$$

where $q = s/2 = 6$ for monthly data, for $j < q$, $f_{jt} = [cos((j/q)\pi t), sin((j/q)\pi t)]$ and for $j = q$ $f_{jt} = cos(\pi t)$. Thus the following vectors of $s - 1 \times 1$ objects can be specified as follows

$$f_t = \begin{pmatrix} f_{1t} \\ f_{2t} \\ \vdots \\ f_{qt} \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{pmatrix}$$

will then lead to Equation (18) being equivalently written as

$$S_t = f'_t \gamma. \tag{19}$$

Putting Equation (19) in Equation (16) then leads to the model being specified in the following way

$$y_t = \mu + x_t^\top \beta + f_t^\top \gamma + e_t. \tag{20}$$

This representation is good because it presents the seasonal components as cyclical where γ_j is the parameter for the seasonal frequency $j\pi/q$ connected to each cyclical process f_t for each seasonal component S_t. It is also important to note that inclusion of lagged variables of y_t in x_t could potentially lead to seasonal unit roots being captured. This could lead to the null hypothesis not being rejected, since the alternative hypothesis is that the time series has a seasonal unit root. Thus since no included lags could lead to the error term e_t being serially correlated the number must be decided with caution. One way to do this is provided by the R function CH.test in package uroot where the default is to decide the number of lags with the following equation

$$l = trunc(s * (T/100)^0.25) \tag{21}$$

where T is the sample size and s is the seasonal period. The test statistic is then derived by

$$L = \sum_{t=1}^{T} \hat{F}'_t A(A'\hat{\Omega}^f A)^{-1} A' \hat{F}_t \tag{22}$$

where $\hat{F}_t = \sum_{t=1}^{T} f_t \hat{e}_t$ with $\hat{e}_t$ being the residuals from the estimation of the model in Equation (16), Ω^f is the long-run covariance matrix of $f_t e_t$ and A is specified to test for the seasonal unit root at one or a number of seasonal lags. If stationarity is rejected at all frequencies then seasonal differencing should be performed to make the time series stationary.

The test statistic is then showed to asymptotically follow the Von Mises goodness-of-fit distribution with critical values presented in Canova and Hansen (1992) which the estimated value is compared to.

3.3.2 The HEGY test

The HEGY test is an extension on theory from the Dickey-Fuller test to test for seasonal unit roots and was developed by Hylleberg et al. (1990). In their paper they developed the test for quarterly data and an extension to monthly data was later created by Franses (1991) on which this section is based.

The first step is to present the seasonal difference operator Δ_s, where it follows that there should be $s = 12$ roots on the unit circle for monthly data. This can be described by the following equation

$$
\begin{aligned}
\Delta_s = (1 - L^{12}) = (1 - L)(1 + L)(1 - iL)(1 + iL) &\times [1 + (\sqrt{3} + i)L/2][1 + (\sqrt{3} - i)L/2] \times \\
\times [1 - (\sqrt{3} + i)L/2][1 - (\sqrt{3} - i)L/2] &\times [1 + (\sqrt{3} + i)L/2][1 - (\sqrt{3} - i)L/2] \times \\
&\times [1 - (\sqrt{3} + i)L/2][1 + (\sqrt{3} - i)L/2] \quad (23)
\end{aligned}
$$

in this equation L is the lag operator and all polynomials except $(1 - L)$ are connected to seasonal unit roots. The test is then based on the following equation

$$
\begin{aligned}
\zeta^*(L)y_{8,t} = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-1} + \pi_4 y_{3,t-2} + \pi_5 y_{4,t-1} + \pi_6 y_{4,t-2} + \\
\pi_7 y_{5,t-1} + \pi_8 y_{5,t-2} + \pi_9 y_{6,t-1} + \pi_{10} y_{6,t-2} + \pi_{11} y_{7,t-1} + \pi_{12} y_{7,t-2} + \mu_t + \epsilon_t \quad (24)
\end{aligned}
$$

where μ_t is determinstic and specified to include a constant, seasonal dummies and/or a trend, ϵ_t is a white noise process and $\zeta^*(L)$ is a polynomial of L and the significance of the parameters π_i are what's of interest. Further the y's are specified as lagged combinations of the observed time series process y_t as follows

$$
\begin{aligned}
y_{1,t} &= (1 + L)(1 + L^2)(1 + L^4 + L^8)y_t, \\
y_{2,t} &= -(1 - L)(1 + L^2)(1 + L^4 + L^8)y_t, \\
y_{3,t} &= -(1 - L^2)(1 + L^4 + L^8)y_t, \\
y_{4,t} &= -(1 - L^4)(1 - \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t, \\
y_{5,t} &= -(1 - L^4)(1 + \sqrt{3}L + L^2)(1 + L^2 + L^4)y_t, \\
y_{6,t} &= -(1 - L^4)(1 - L^2 + L^4)(1 - L + L^2)y_t, \\
y_{7,t} &= -(1 - L^4)(1 - L^2 + L^4)(1 + L + L^2)y_t, \\
y_{8,t} &= (1 - L^{12})y_t.
\end{aligned}
$$

The model in Equation (24) is then estimated by ordinary least squares with focus on the estimates of the π's. This is done for the specifications of μ_t being considered relevant. The t-test are then used on $\hat{\pi}_1$ and $\hat{\pi}_2$ to test the one-sided null hypothesis of a unit root for $\hat{\pi}_1$ and a seasonal unit root for $\hat{\pi}_2$ against the alternative hypothesis of no unit root. For $\hat{\pi}_i$ with $i > 2$ a seasonal unit root at a specific frequency is only present when it occurs for connected pairs of parameters. Thus, it follows that the F-test is used to the joint two-sided null hypothesis of a unit root on the connected pairs, $(\hat{\pi}_3, \hat{\pi}_4)$, $(\hat{\pi}_5, \hat{\pi}_6)$, $(\hat{\pi}_7, \hat{\pi}_8)$, $(\hat{\pi}_9, \hat{\pi}_{10})$ and $(\hat{\pi}_{11}, \hat{\pi}_{12})$ against the alternative that that there are seasonal unit roots.

Critical values for all these tests can be found in Franses (1991). Non-rejection of the null hypothesis of a seasonal unit root for the test of $\hat{\pi}_2$ and all the connected pairs would imply that seasonal differencing of the series should be performed.

3.4 The Autoregressive and Moving Average Orders

In this section the method of selecting the remaining orders will be described. The autoregressive orders p and P and the moving average orders q and Q will be decided partly by visual inspection of the correlogram and partly by minimizing information criterion's with the use of the HK-algorithm.

3.4.1 Correlogram

The correlogram is used to check the randomness of the data. Box and Jenkins states that the autocorrelation function (ACF) and partial autocorrelation function (PACF) can be used to identify the orders p, q, P, Q for a SARIMA model (Box et al., 2008, p. 378). The ACF at lag τ is given by (Harvey, 1989, p. 50)

$$r_\tau = c_\tau/c_0 \tag{25}$$

where c_τ is the following autocovariance function

$$c_\tau = T^{-1} \sum_{t=\tau+1}^{T} (y_t - \bar{y})(y_{t-\tau} - \bar{y}), \ \tau = 1, 2, 3, ... \tag{26}$$

and c_0 is the variance, derived by

$$c_0 = T^{-1} \sum_{t=1}^{T} (y_t - \bar{y})^2 \tag{27}$$

for the T observations of the process y_t with sample mean $\bar{y}$ (Harvey, 1989, p. 50).

The PACF is the autocorrelation between y_t and its lagged process y_{t+k} while excluding all autocorrelations ranging from y_{t+1} to y_{t+k-1}. These are estimated with the use of the derived autocorrelations r_j where the following function is first needed

$$r_j = \hat{\phi}_{k1}r_{j-1} + \hat{\phi}_{k2}r_{j-2} + \ldots + \hat{\phi}_{k(k-1)}r_{j-k+1} + \hat{\phi}_{kk}r_{j-k} \text{ for } j = 1, 2, ..., k. \tag{28}$$

These target is to extract the parameters $\hat{\phi}_{11}, \hat{\phi}_{22}, ..., \hat{\phi}_{kk}$ where $\hat{\phi}_{jj}$ is the partial autocorrelation for lag j. (Box and Jenkins, 1976, p. 65)

The correlogram includes plots of the sample ACF and the sample PACF both against the time lags τ. For complete randomness the values at all lags should be zero. The most important part of the identification procedure is to look for significant lags meaning the ones lying outside the interval $\pm 2/\sqrt{T}$ (Hamilton, 1994, p. 111). Further the ACF and PACF are used to find appropriate orders for the SARIMA model by using the results in Table 1. It is however assumed that the differencing implied by d and D have been done to the time series, meaning that the SARIMA$(p, d, q) \times (P, D, Q)_s$ of y_t has taken its equivalent representation as a SARMA$(p, q) \times (P, Q)_s$ of the differenced series y_t^* (Box and Jenkins, 1976, p. 79 and 303-304)

Table 1: ACF and PACF to identify the orders of SARMA$(p, q) \times (P, Q)_s$, only positive lags are of interest.

	ACF	PACF
AR(p)	Exponentially decreasing or damped sine wave	Spikes to lag p then zero
MA(q)	Spikes to lag q then zero	Exponentially decreasing or damped sine wave
ARMA(p, q)	Exponentially decreasing or damped sine wave after $q - p$ lags	Exponentially decreasing or damped sine wave after $p - q$ lags
SAR$(P)_s$	Exponentially decreasing or damped sine wave for all lags times s	Spikes for lag Ps then zero
SMA$(Q)_s$	Spikes for lag Qs then zero	Exponentially decreasing or damped sine wave for all lags times s
SARMA$(P, Q)_s$	Exponentially decreasing or damped sine wave for all lags times s after lags $(Q - P)s$	Exponentially decreasing or damped sine wave for all lags times s after lags $(P - Q)s$

3.4.2   Selection with the HK-algorithm

The Hyndman-Khandakar (HK) algorithm was developed by Hyndman and Khandakar (2008) and can be applied in R with the function auto.arima in the forecast package. They suggest an iterative time-saving procedure where the model with the smallest value of some information criterions AIC, AICc or BIC will be found much faster, since it is now found without comparing every possible model.

To derive these information criterions the first thing that is needed is the likelihood function, $L(\tilde{\boldsymbol{\Psi}})$, where $\tilde{\boldsymbol{\Psi}}$ is the maximum likelihood estimates of the parameters for the SARIMA with $n = p + q + P + Q + 1$ parameters and sample size T. The criterions are then derived by the following equations

$$AIC = -2log[L(\tilde{\boldsymbol{\Psi}})] + 2n, \tag{29}$$

$$AICc = AIC + \frac{2n(n+1)}{T - n + 1}, \tag{30}$$

$$BIC = -2log[L(\tilde{\boldsymbol{\Psi}})] + nlog(T). \tag{31}$$

The HK-algorithm then performs an iterative procedure to select the model that mini-

mizes the value of each criterion. It begins with estimation of the following four models

- $\text{SARIMA}(2, d, 2) \times (1, D, 1)_s$

- $\text{SARIMA}(0, d, 0) \times (0, D, 0)_s$

- $\text{SARIMA}(1, d, 0) \times (1, D, 0)_s$

- $\text{SARIMA}(0, d, 1) \times (0, D, 1)_s$

where d and D are assumed to have been found previously and a constant is included in the models if $d + D \leq 1$. The model which attains the smallest value for the chosen information criterion is then selected and the procedure continues with varying the parameters in the following ways

- Let each of p, q, P and Q vary with ± 1.

- Let both p and q vary with ± 1 at the same time.

- Let both P and Q vary with ± 1 at the same time.

- Include the intercept if previously not included otherwise do the opposite.

This step of the procedure will be repeated until none of these variations decreases the value of the criterion.

There are some constraints that follows with the use of this method. These are used to check that the model is reasonable and well-fitted and are the following

- The maximum orders of p and q are five.

- The maximum orders of P and Q are two.

- All non-invertible or non-causal models are rejected. These are found by computing the roots of the lag polynomials $\phi(L)\Phi(L)$ and $\theta(L)\Theta(L)$, if any root is smaller than 1.001 then the model is rejected.

- If errors arise when fitting the model with the non-linear optimization routine then the model is rejected.

At this stage the final model is found and the Box-Jenkins procedure can continue to its second step, meaning estimation.

3.5   Estimation

The second part of the Box-Jenkins methodology for SARIMA modelling is estimation. The method chosen for this is maximum likelihood. It is first assumed that the SARIMA model with parameters $\boldsymbol{\theta}$ have been identified. Further it is assumed that the number of observations must be at least 50 and preferably 100 for efficient estimation (Box and Jenkins, 1976, p. 18).

The first part of the maximum likelihood estimation is to specify the probability density function implied by the chosen model. It is assumed that the error term of the model is distributed as Gaussian white noise. The estimation procedure is then performed in two steps. First the likelihood function is derived and then the value of the parameter vector $\boldsymbol{\theta}$ is specified to maximize the value of that function. The maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ is then interpreted as having the value which maximizes the probability for observing this specific sample of observations (Hamilton, 1994, p. 117).

3.6   Evaluation of the Model Fit

The fit of the model is evaluated by diagnostic checks of the residuals. The residuals should behave like Gaussian white noise, that is appear random, homoscedastic and normal (Box and Jenkins, 1976, p. 324).

The first part is a graphical check of the standardized residuals, meaning the residuals divided with their standard deviation. These should look random and homoscedastic. The number of outliers are also important where a good indication would be that about 95 percent of the residuals lie inside their 95 percent confidence interval ± 1.96 (Brockwell and Davis, 1991, p. 307).

The next step is to evaluate the assumption of randomness by using the sample autocorrelation function of the residuals. The autocorrelations of interest are those that are significantly different from zero, that is those who lie outside the sample size dependent approximately 95 percent confidence interval $\pm 2/\sqrt{T}$ (Hamilton, 1994, p. 111). Those significant lags suggest some kind of inconsistency in the residuals, but there is no reason to worry if only about five percent of the autocorrelations are significant. (Brockwell and Davis, 1991, p. 309)

In this thesis the most important part of the diagnostic checking is the use of tests to possibly acquire statistically significant results which would imply a rejection of the fitted model. The chosen tests are the Box-Ljung test which is used to test the serial independence and the Jarque-Bera test which tests the normality of the residuals.

3.6.1 The Box-Ljung test

The Ljung and Box (BL) test was developed in 1978 and is used to test the randomness of the residuals. For this test the first step is to extract the residuals $\hat{\epsilon}_t$ for the fitted model. The T residuals are then used to derive the sample autocorrelations of the residuals with the following equation

$$\hat{r}_k = \frac{\sum_{t=k+1}^{T} \hat{\epsilon}_t \hat{\epsilon}_{t-k}}{\sum_{t=1}^{T} \hat{\epsilon}_t^2}, \ k = 1, 2, ... \tag{32}$$

This equation is used until a set of autocorrelations $\hat{r}_1, \hat{r}_2, ..., \hat{r}_m$ have been obtained. These are then used to test the null hypothesis of serially independent residuals versus the alternative hypothesis that they are not serially independent with the following test statistic (Ljung and Box, 1978, p. 298)

$$\tilde{Q}(r) = T(T+2) \sum_{k=1}^{m} (T-k)^{-1} r_k^2 \tag{33}$$

which for an appropriate model was shown to be asymptotically distributed as a $\chi_{1-\alpha}^2(m)$ where m is the number of lagged autocorrelations included and α is the selected significance level. Harvey (p. 259, 1990) suggest that the number of lags should be a function of T for example the truncated value of $m = \sqrt{T}$ and that the degrees of freedom should be corrected for SARIMA models to $df = m - p - q - P - Q$. The critical value is included in (Ljung and Box, 1978) and then compared to the value of the test statistic. The null hypothesis of randomness is rejected for large values of the test statistic.

3.6.2 The Jarque-Bera test

The Jarque-Bera test is used to test the normality of the residuals. The null hypothesis of the test is normality and it is tested against the alternative hypothesis of non-normality. The concept of normality is further presented in the Appendix. The statistic that is used is the first part of equation 4 in Jarque and Bera (1980), written in the following way

$$JB = T \left(\frac{S^2}{6} + \frac{K^2 - 3}{24} \right) \tag{34}$$

where T is the number of observations S is the skewness derived by Equation (65) and K is the kurtosis from Equation (66). The statistic is assumed to be distributed as a $\chi_{1-\alpha}^2(2)$ variable and the null hypothesis of normality is rejected for large values of the statistic (Schwert, 2009, p. 466).

3.7 Forecasting with the SARIMA model

In Equation (7) it has been shown that the $\text{SARIMA}(p, d, q) \times (P, D, Q)_s$ for the variable y_t, $t = 1, ..., T$ can be written equivalently as an $\text{ARMA}(p + sP, q + sQ)$ for y_t^* from Equation (2). The forecast function for the assumed stationary variable y_t^* is then written as

$$(\hat{y}_{t+1|t}^* - \mu) = \phi(L)^*(\hat{y}_t^* - \mu) + \theta(L)^*\hat{\epsilon}_t \tag{35}$$

where $\hat{\epsilon}_t = y_t - \hat{y}_{t|t-1}^*$ (Hamilton, 1994, p. 84). The forecast for lead time τ, meaning the time that follows after the last observed information, is then derived by

$$(\hat{y}_{t+s|t}^* - \mu) = \phi(L)^*(\hat{y}_{t+s-1|t}^* - \mu) + \theta(L)^*\hat{\epsilon}_{t+s-1}. \tag{36}$$

That is the forecast for lead time τ will be derived by the previously observed values of y_t^*, previous forecasts of $\hat{y}^*$ and the residuals $\hat{\epsilon}_t$ which have been derived for all time points up to the last observed observation but are equal to zero for the ones where the real values have not yet been observed. (Hamilton, 1994, p. 84)

3.8 State-Space Model

The state-space form of a model makes it possible to apply the Kalman filter (Hamilton, 1994, p. 375). On its general form the state space model consist of two equations, first the observation equation written on the form of a linear regression model

$$\boldsymbol{y}_t = \boldsymbol{A}^\top \boldsymbol{x}_t + \boldsymbol{H}^\top \boldsymbol{\xi}_t + \boldsymbol{w}_t, \ E(\boldsymbol{w}_t \boldsymbol{w}_\tau^\top) = \begin{cases} \boldsymbol{R}, & \text{if } t = \tau \\ \boldsymbol{0}, & \text{if } t \neq \tau \end{cases} \tag{37}$$

second the state equation written on the form of a first order vector autoregressive model (Hamilton, 1994, p. 372)

$$\boldsymbol{\xi}_t = \boldsymbol{F}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t, \ E(\boldsymbol{v}_t \boldsymbol{v}_\tau^\top) = \begin{cases} \boldsymbol{Q}, & \text{if } t = \tau \\ \boldsymbol{0}, & \text{if } t \neq \tau \end{cases}. \tag{38}$$

In these equations $\boldsymbol{y}_t$ is the observed vector of variables, $\boldsymbol{\xi}_t$ is the state vector consisting of unobserved variables, $\boldsymbol{x}_t$ is a vector of predetermined variables that possibly are lagged values of $\boldsymbol{y}_t$. Further the matrices $\boldsymbol{F}$ and $\boldsymbol{H}$ are parameter matrices while $\boldsymbol{w}_t$ and $\boldsymbol{v}_t$ are error terms assumed to be distributed as white noise with covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$. (Hamilton, 1994, p. 373)

All these matrices are specified in different ways depending on the selected model. It should be noted that a specific model can have multiple different specifications that are considered equivalent. Thus it now seems fitting to present how these are specified to

put some frequently used time series models in state-space form. First examplified by an AR(p) process for which the matrices can be specified in the following way (Hamilton, 1994, p. 374)

$$\boldsymbol{\xi}_t = \begin{bmatrix} y_t - \mu \\ y_{t-1} - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix}, \boldsymbol{F}^\top = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_{p-1} & \phi_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$\boldsymbol{v}_t = \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \boldsymbol{Q} = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \boldsymbol{y}_t = y_t, \boldsymbol{A}^\top = \mu, \boldsymbol{x}_t = 1,$$

$$\boldsymbol{H}^\top = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}, \boldsymbol{w}_t = 0, \boldsymbol{R} = 0.$$

For a MA(q) process the matrices can be (Hamilton, 1994, p. 375)

$$\boldsymbol{\xi}_t = \begin{bmatrix} \epsilon_t \\ \epsilon_{t-1} \\ \vdots \\ \epsilon_{t-q} \end{bmatrix}, \boldsymbol{F}^\top = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \boldsymbol{v}_t = \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\boldsymbol{Q} = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \boldsymbol{y}_t = y_t, \boldsymbol{A}^\top = \mu, \boldsymbol{x}_t = 1, \boldsymbol{H}^\top = \begin{bmatrix} 1 & \theta_1 & \dots & \theta_q \end{bmatrix}, \boldsymbol{w}_t = 0,$$

$$\boldsymbol{R} = 0.$$

The next step is to present a way for combining the state-space models for the AR and MA into the form of an ARMA. In Petris et al. (2009, p. 112), they present a way to write the ARMA model in state-space form. This one introduces a slight variation compared to the state-space model in Hamilton (1994) concerning the MA part where the following matrices have been respecified in the following way

$$\boldsymbol{v}_t = \begin{bmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_q \end{bmatrix} \epsilon_t, \boldsymbol{Q} = \sigma^2 \begin{bmatrix} 1 & \theta_1 & \dots & \theta_q \\ \theta_1 & \theta_1^2 & \dots & \theta_1 \theta_q \\ \vdots & \vdots & \dots & \vdots \\ \theta_q & \theta_1 \theta_q & \dots & \theta_q^2 \end{bmatrix}, \boldsymbol{H} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

This leads to an ARMA(p, q), for which $r = max(p, q + 1)$, $\phi_j = 0$ for $j > p$ and $\theta_j = 0$ for $j > q + 1$ being written in state-space form as

$$\boldsymbol{y}_t = \boldsymbol{H}^\top \boldsymbol{\xi}_t \tag{39}$$

$$\boldsymbol{\xi}_t = \boldsymbol{F}\boldsymbol{\xi}_{t-1} + \boldsymbol{v}_t \tag{40}$$

Note that the earlier mentioned matrices $\boldsymbol{A}$, $\boldsymbol{x}_t$, $\boldsymbol{w}_t$ and $\boldsymbol{R}$ have all been set to zero. The rest of the matrices are then specified as follows

$$\boldsymbol{y}_t = y_t, \; \boldsymbol{\xi}_t^\top = \begin{bmatrix} \xi_{1,t} & \xi_{2,t} & \dots & \xi_{r,t} \end{bmatrix}, \; \boldsymbol{H}^\top = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix},$$

$$\boldsymbol{F}^\top = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_{r-1} & \phi_r \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \; \boldsymbol{v}_t = \begin{bmatrix} 1 \\ \theta_1 \\ \dots \\ \theta_{r-1} \end{bmatrix} \epsilon_t,$$

$$\boldsymbol{Q} = \sigma^2 \begin{bmatrix} 1 & \theta_1 & \dots & \theta_{r-1} \\ \theta_1 & \theta_1^2 & \dots & \theta_1\theta_{r-1} \\ \vdots & \vdots & \dots & \vdots \\ \theta_{r-1} & \theta_1\theta_{r-1} & \dots & \theta_{r-1}^2 \end{bmatrix}.$$

Thus, before applying the Kalman filter we need to estimate the matrices $\boldsymbol{F}$, $\boldsymbol{H}$, $\boldsymbol{v}$ and $\boldsymbol{Q}$.

The specification of the state-space model matrices will be simplified by writing the SARIMA model on the equivalent form of an ARMA presented earlier in Equations (8) and (9). Thus, the state-space form of an ARMA model presented above will be sufficient for writing the SARIMA on that form.

3.9   The Kalman Filter

The Kalman filter let the state vector in Equation (38) be updated for each new observation of the possibly multivariate series $\boldsymbol{y}_t$ (Harvey, 1989, p. 105). It consists of linear estimation steps, where $\hat{\boldsymbol{\xi}}_t$ is computed by known values of $\hat{\boldsymbol{\xi}}_{t-1}$ and $\boldsymbol{y}_t$. However the first step is to obtain values for the first state, $\boldsymbol{\xi}_1$. Calculation of this state is based on its unconditional mean (Hamilton, 1994, p. 378)

$$\hat{\boldsymbol{\xi}}_{1|0} = E(\boldsymbol{\xi}_1) = \text{[assumed to be]} = 0. \tag{41}$$

The reason for assuming this is that there are no prior information and it is therefore decided to be zero. Further its variance, MSE, is derived by

$$\boldsymbol{P}_{1|0} = E[(\boldsymbol{\xi}_1 - (E(\boldsymbol{\xi}_1))(\boldsymbol{\xi}_1 - (E(\boldsymbol{\xi}_1))^\top] = \text{[assumed to be]} =$$
$$= vec(\boldsymbol{P}_{1|0}) = [\boldsymbol{I} - (\boldsymbol{F} \otimes \boldsymbol{F})]^{-1} \times vec(\boldsymbol{Q}) \quad (42)$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{F}$ and $\boldsymbol{Q}$ are previously known and $vec(\boldsymbol{P}_{1|0})$ is the column vector which is then easily transformed to the quadratic matrix $\boldsymbol{P}_{1|0}$ (Hamilton, 1994, p. 378).

The prior state that has been obtained can then be updated iteratively with the use of the observations $\boldsymbol{y}_1, ..., \boldsymbol{y}_T$ by recursions on the following equation

$$\hat{\boldsymbol{\xi}}_{t+1|t} = \boldsymbol{F}\hat{\boldsymbol{\xi}}_{t|t-1} + \boldsymbol{F}\boldsymbol{P}_{t|t-1}\boldsymbol{H}(\boldsymbol{H}^\top \boldsymbol{P}_{t|t-1}\boldsymbol{H} + \boldsymbol{R})^{-1}(\boldsymbol{y}_t - \boldsymbol{A}^\top \boldsymbol{x}_t - \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_{t|t-1}). \quad (43)$$

It is assumed that $\hat{\boldsymbol{\xi}}_{t|t-1}$ is the best forecast of the true state $\boldsymbol{\xi}_t$ based on the linear function of the observations $\boldsymbol{y}_1, ..., \boldsymbol{y}_{t-1}$ and deterministic variables $\boldsymbol{x}_1, ..., \boldsymbol{x}_t$ while $\boldsymbol{P}_{t|t-1}$ is the variance, MSE, of this forecast. Note that these assumptions are not for the real mean and variance but for the conditional mean and variance where the estimates at time point t is given by all observations up to $t - 1$. (Hamilton, 1994, p. 380)

Updates are also performed on the forecast variance $\boldsymbol{P}_{t+1|t}$ in the following way

$$\boldsymbol{P}_{t+1|t} = \boldsymbol{F}[\boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{H}(\boldsymbol{H}^\top \boldsymbol{P}_{t|t-1}\boldsymbol{H} + \boldsymbol{R})^{-1})\boldsymbol{H}^\top \boldsymbol{P}_{t|t-1}]\boldsymbol{F}^\top + \boldsymbol{Q} \quad (44)$$

which is done iteratively for each time point up to the last observation at time T (Hamilton, 1994, p. 380). It should be noted that the positive semi-definiteness of $\boldsymbol{Q}$, $\boldsymbol{R}$ and $\boldsymbol{P}_{1|0}$ are guaranteed for the work in this thesis by using singular value decomposition (Petris et al., 2009, p. 56). This ensures that the sequence $\{\boldsymbol{P}_{t+1|t}\}_{t=1}^T$ is monotonically nonincreasing and converges to its steady state $\boldsymbol{P}$. (Hamilton, 1994, p. 390)

The iterative procedure on the state-vector can be related to the forecast of the observed variable $\boldsymbol{y}_{t+1}$ in the following way

$$\hat{\boldsymbol{y}}_{t+1|t} = \hat{E}(\boldsymbol{y}_{t+1}|\boldsymbol{x}_t, \boldsymbol{x}_{t-1}\boldsymbol{x}_1, \boldsymbol{y}_t, \boldsymbol{y}_{t-1}, ..., \boldsymbol{y}_1) = \boldsymbol{A}^\top \boldsymbol{x}_{t+1} + \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_{t+1|t} \quad (45)$$

with MSE (Hamilton, 1994, p. 381)

$$E[(\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t})(\boldsymbol{y}_{t+1} - \hat{\boldsymbol{y}}_{t+1|t})^\top] = \boldsymbol{H}^\top \boldsymbol{P}_{t+1|t}\boldsymbol{H} + \boldsymbol{R}. \quad (46)$$

The s-period ahead forecasts of $\boldsymbol{\xi}_{t+s}$ are obtained by recursive substitution on Equation (38)

$$\boldsymbol{\xi}_{t+s} = \boldsymbol{F}^s \boldsymbol{\xi}_t + \boldsymbol{F}^{s-1} \boldsymbol{v}_{t+1} + \boldsymbol{F}^{s-2} \boldsymbol{v}_{t+2} + \ldots + \boldsymbol{F}^1 \boldsymbol{v}_{t+s-1} + \boldsymbol{v}_{t+s}, \text{ for } s = 1, 2, ..., S \quad (47)$$

it then follows that the estimated forecasts $\hat{\boldsymbol{\xi}}_{t+s|t}$ are seen as the projection of $\boldsymbol{\xi}_{t+s}$ on $\boldsymbol{\xi}_t$ and $\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, ...\boldsymbol{y}_1, \boldsymbol{x}_t, \boldsymbol{x}_{t-1}, ...\boldsymbol{x}_1$. The law of iterated projections then yield

$$\hat{\boldsymbol{\xi}}_{t+s|t} = \hat{E}(\boldsymbol{\xi}_{t+s}|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, ...\boldsymbol{y}_1, \boldsymbol{x}_t, \boldsymbol{x}_{t-1}, ...\boldsymbol{x}_1) = \boldsymbol{F}^s \hat{\boldsymbol{\xi}}_{t|t} \quad (48)$$

while the forecast error is derived by

$$\boldsymbol{\xi}_{t+s} - \hat{\boldsymbol{\xi}}_{t+s|t} = \boldsymbol{F}^s (\boldsymbol{\xi}_t - \hat{\boldsymbol{\xi}}_{t+1|t}) + \boldsymbol{F}^{s-1} \boldsymbol{v}_{t+1} + \boldsymbol{F}^{s-2} \boldsymbol{v}_{t+2} + \ldots + \boldsymbol{F}^1 \boldsymbol{v}_{t+s-1} + \boldsymbol{v}_{t+s} \quad (49)$$

and the MSE is calculated by (Hamilton, 1994, p. 385)

$$\boldsymbol{P}_{t+s|t} = \boldsymbol{F}^s \boldsymbol{P}_{t|t} (\boldsymbol{F}^\top)^s + \boldsymbol{F}^{s-1} \boldsymbol{Q} (\boldsymbol{F}^\top)^{s-1} + \ldots + \boldsymbol{F} \boldsymbol{Q} \boldsymbol{F}^\top + \boldsymbol{Q}. \quad (50)$$

The s-period ahead forecast of the observed variable are then obtained by performing iterations on Equation (37) which leads to the following equation

$$\hat{\boldsymbol{y}}_{t+s} = \hat{E}(\boldsymbol{y}_{t+s}|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, ...\boldsymbol{y}_1, \boldsymbol{x}_t, \boldsymbol{x}_{t-1}, ...\boldsymbol{x}_1) = \boldsymbol{A}^\top \boldsymbol{x}_{t+s} + \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_{t+s|t} \quad (51)$$

with the forecast error derived by

$$\boldsymbol{y}_{t+s} - \hat{\boldsymbol{y}}_{t+s} = (\boldsymbol{A}^\top \boldsymbol{x}_{t+s} + \boldsymbol{H}^\top \boldsymbol{\xi}_{t+s|t} + \boldsymbol{w}_{t+s}) - (\boldsymbol{A}^\top \boldsymbol{x}_{t+s} + \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_{t+s|t}) =$$
$$= \boldsymbol{H}^\top (\boldsymbol{\xi}_{t+s|t} - \hat{\boldsymbol{\xi}}_{t+s|t}) + \boldsymbol{w}_{t+s} \quad (52)$$

and the MSE calculated by (Hamilton, 1994, p. 385)

$$E[(\boldsymbol{y}_{t+s} - \hat{\boldsymbol{y}}_{t+s})(\boldsymbol{y}_{t+s} - \hat{\boldsymbol{y}}_{t+s})^\top] = \boldsymbol{H}^\top \boldsymbol{P}_{t+s|t} \boldsymbol{H} + \boldsymbol{R}. \quad (53)$$

It can further be claimed that the Kalman filter leads to optimal forecasts of $\hat{\boldsymbol{\xi}}_{t|t-1}$ and $\hat{\boldsymbol{y}}_{t|t-1}$ among any functions of $(\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_{t-1}, \boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_t)$. Further if the first state, $\hat{\boldsymbol{\xi}}_1$, and the error processes $\{\boldsymbol{w}_t, \boldsymbol{v}_t\}_{t=1}^T$ are normal then this leads to the following conditional normality (Hamilton, 1994, p. 385)

$$\boldsymbol{y}_t|\boldsymbol{y}_1, ..., \boldsymbol{y}_{t-1}, \boldsymbol{x}_1, ..., \boldsymbol{x}_t \sim N[(\boldsymbol{A}^\top \boldsymbol{x}_t + \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_{t|t-1}), (\boldsymbol{H}^\top \boldsymbol{P}_{t|t-1} \boldsymbol{H} + \boldsymbol{R})]. \quad (54)$$

The fit of applying the Kalman filter to this specific model is evaluated with the so called innovations, $\boldsymbol{i}_t$, which are called that since they are said to represent something that has not been known for earlier observations (Durbin and Koopman, 2001, p. 13). These can be interpreted as the residuals of the process and are derived by (Brockwell and Davis, 1991, p. 476)

$$\boldsymbol{i}_t = \boldsymbol{y}_t - \boldsymbol{H}^\top \hat{\boldsymbol{\xi}}_t + \boldsymbol{w}_t \text{ for } t = 1, 2, ..., T. \tag{55}$$

The innovations are assumed to be distributed as Gaussian white noise. This means that they should be serially independent, homoscedastic, and normal (Commandeur et al., 2011). That is exactly the same as was assumed for the residuals of the SARIMA model.

3.10 Comparing forecasts

The accuracy of the forecasts obtained by the two models will first be compared with some error measures and then also with a test to evaluate if the difference is significant.

3.10.1 Error Measures

Error measures are easily derived and give an indication for which model that is most fit for forecasting. These will be derived by the use of y_t which is the real value, $\hat{y}_t$ which is the forecast value, n which is the length of the forecast horizon and t which goes from one to n. The first error measure that is presented is the mean squared error (MSE) which is estimated by (Petris et al., 2009, p. 98)

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2. \tag{56}$$

A value of zero would imply a perfect forecast and a negative value should not be possible to obtain. The next error measure is the mean absolute deviation (MAD) is derived by (Petris et al., 2009, p. 98)

$$\text{MAD} = \frac{1}{n} \sum_{t=1}^{n} |y_t - \hat{y}_t|. \tag{57}$$

For this measure the value of zero would imply a perfect forecast and a negative value should not be possible. The difference between MAD and MSE is that MSE places relatively greater penalty on large forecast errors. Further the mean absolute percentage error (MAPE) is calculated by (Petris et al., 2009, p. 98)

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right|. \tag{58}$$

A value of zero would imply a perfect forecast and a negative value cannot possibly be obtained. It is also important to state that this error measure is not defined if $y_t = 0$ for any t. The MAPE scales the errors, that is this measure puts relatively more penalty to the forecast error if the true value of the observation is small.

The last of these measures is a statistic called Theil's U. This statistic compares the ratio between the MSE implied by the model forecast and the MSE of forecasts obtained by a naive model which sets the upcoming observation at time $t+1$ as equal to the value observed at time t. It is derived by (Petris et al., 2009, p. 99)

$$U = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - y_{t-1})^2}}. \tag{59}$$

The value of Theil's U will be smaller than one if the model forecast has higher accuracy than the naive model forecast. The estimate will be zero if the forecasts are perfect and it cannot possibly be derived as negative. It is not defined if the naive model perform perfect forecasts (Petris et al., 2009, p. 99).

3.10.2 The Diebold-Mariano test

The Diebold and Mariano (DM) test was developed in 1995 and its first assumed that there are two forecasts $\hat{w}_1, ..., \hat{w}_h$ and $\hat{z}_1, ..., \hat{z}_h$ of the true time series $y_1, ..., y_h$ where h is the forecast horizon. These are then used to derive the forecast errors of the first $e_{1i} = y_i - \hat{w}_i$ and the second model $e_{2i} = y_i - \hat{w}_i$. (Diebold and Mariano, 1995, p. 134)

This test works even if the forecast errors have a non-zero mean are non-Gaussian and are correlated to each other. For the test in R with function dm.test the forecast package with default specifications implies that the errors are used to derive the loss functions $g(e_{1i}) = |e_{1i}|^2$ and $g(e_{2i}) = |e_{2i}|^2$ which lead to the loss-differential series $d_i = \{g(e_{1i}) - g(e_{2i})\}_{i=1}^{h}$ which is assumed normally distributed (Hyndman and Khandakar, 2008). The sample mean is then derived by

$$\bar{d} = \frac{\sum_{i=1}^{h} d_i}{h}. \tag{60}$$

The next step is to derive the auto-covariance at lag 0 for d_i, that is $\hat{\gamma}_0$ which is assumed to be a consistent estimate of the variance of $h\bar{d}$. The test statistic then follows and is derived by

$$DM = \frac{\bar{d}}{\sqrt{\hat{\gamma}_0/h}} \sim N(0,1) \tag{61}$$

Thus, the DM statistic is assumed to be standard normal and the rejection values then follows from that. The null hypothesis of the test is that there is no difference between the accuracy of the two forecasts. The alternative hypothesis is either two-sided to test if either model perform better than the other or one-sided to test if one specific model is more accurate than the other. (Diebold and Mariano, 1995)

4 Data

The data is taken from the website *inflation.eu Worldwide Inflation Data* (2013) which assembles up to date data on current and historic inflation for countries all over the world. The chosen countries are Luxembourg where the data in the website have been assembled from *Service Central de la Statistique et des Etudes Economiques*, for Mexico from *Instituto Nacional de Estadistics y Geografia*, for Portugal from *Instituto Nacional de Estatistica* and for Switzerland from *Bundesamt für Statistik*. Month-to-month inflation based on the consumer price index (CPI) have then been collected for the chosen countries.

The monthly inflation is measured in percentage form with two decimals from January 1980 to December 2012. It is first presented graphically in Figures 2, 3, 4 and 5 where the subjectively chosen well-behaved part of the data is red and blue coloured and the data for 2012 is excluded and used for forecast comparisons. For Luxembourg inflation the chosen part of the data is from January of 1999 to December 2011, for Mexican inflation from January 2001 to December 2011, for Portuguese inflation from January 2003 to December 2011 and for Swiss inflation from January 2002 to December 2011. Descriptive statistics can be found in Table 2 and it should be noted that the null hypothesis of normality for the Jarque-Bera test is rejected for Mexican and Portuguese inflation.

For the specification of the SARIMA models which are put in state-space form it is important to remember that at least 50 observations are needed for efficient modelling. In Figures 2, 3, 4 and 5 the first red coloured part are used for this estimation. Thus the data used for estimation is ranged from January 1999 to December 2004 for Luxembourg, from January 2001 to December 2006 for Mexico, from January 2003 to December 2007 for Portugal and January 2002 to December 2007 for Switzerland. The remaining observations, the blue coloured part, with last observation at December 2011 are then used for Kalman filtering.

Table 2: Descriptive statistics for monthly inflation of each country

	Luxembourg	Mexico	Portugal	Switzerland
N	156	132	108	120
Mean	0.199	0.362	0.180	0.061
Median	0.195	0.395	0.140	0.090
Maximum	1.850	1.140	1.610	1.130 0
Minimum	-1.670	-0.740	-0.640	-1.050
Std	0.682	0.339	0.441	0.424
Skewness	0.064	-0.546	0.728	0.045
Kurtosis	3.186	3.783	3.880	3.138
JB test	0.333	9.927	13.018	0.136
	(0.847)	(0.007)	(0.001)	(0.934)

The JB test for the null of the time series following the normal distribution is presented with the test statistic and the p-value in the parenthesis.
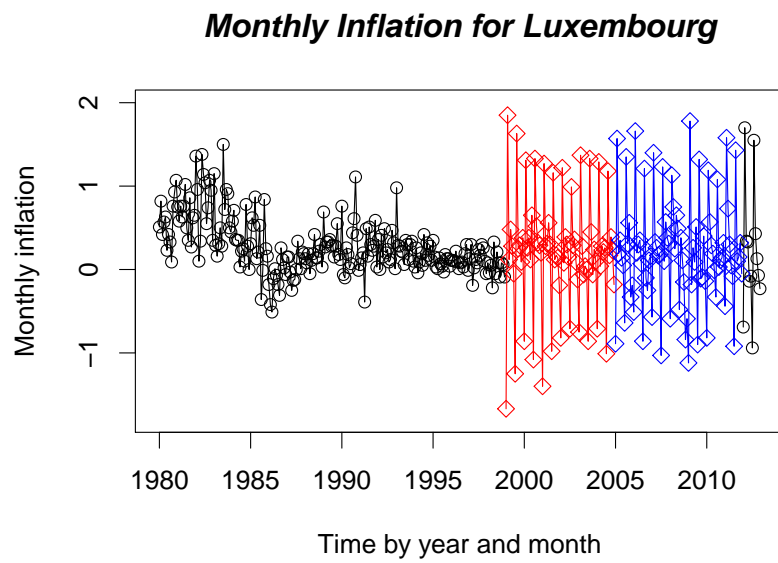
Monthly Inflation for Luxembourg

Figure 2: Luxembourg monthly inflation from January 1980 to December 2012

Monthly Inflation for Mexico
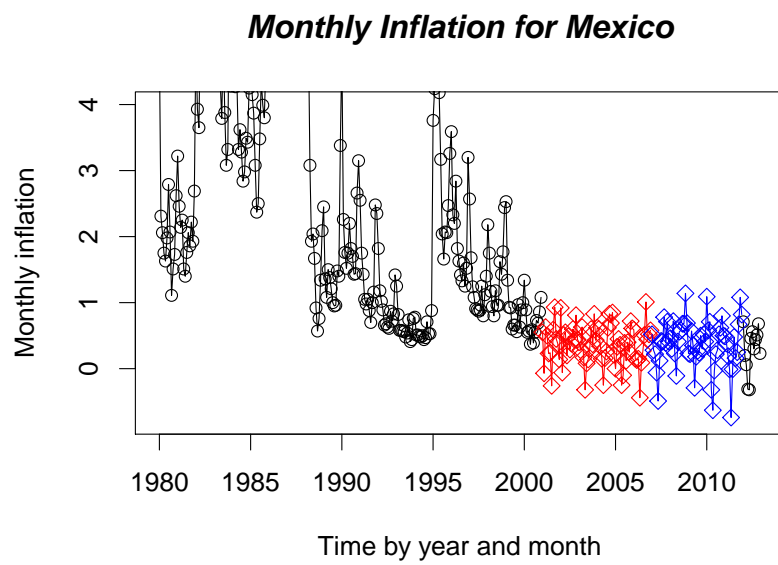
Figure 3: Mexican monthly inflation from January 1980 to December 2012

Monthly Inflation for Portugal
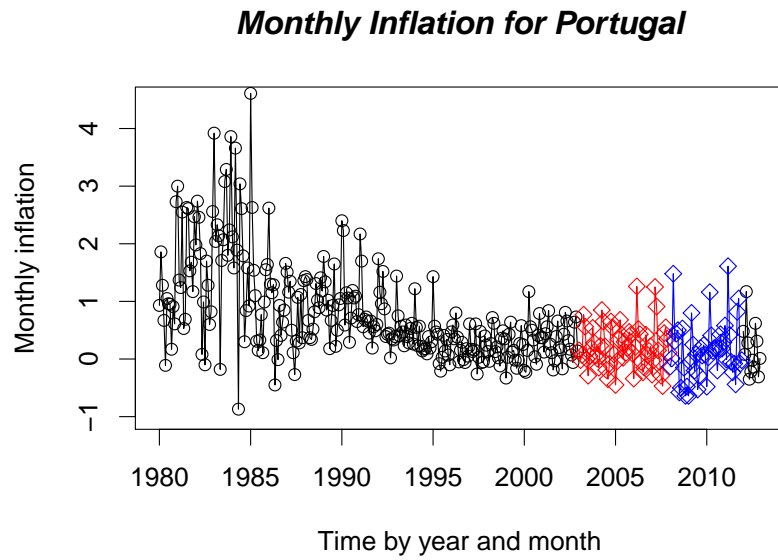
Figure 4: Portuguese monthly inflation from January 1980 to December 2012

Monthly Inflation for Switzerland
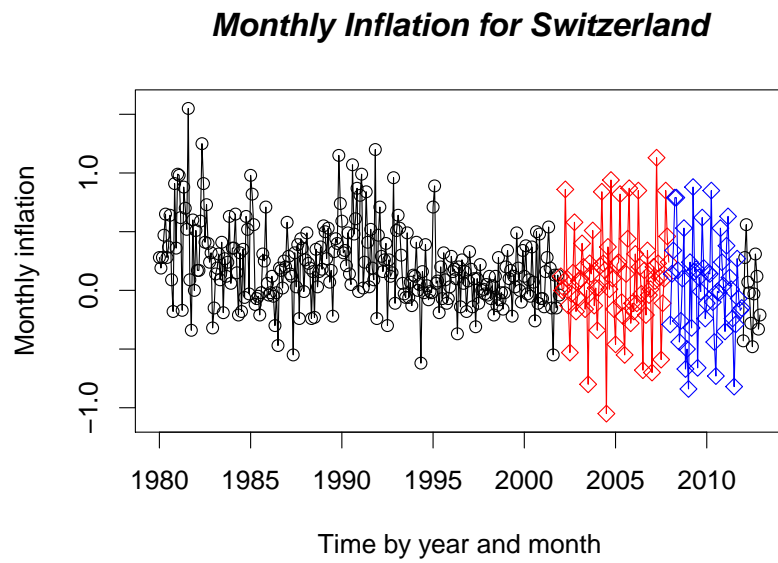
Figure 5: Swiss monthly inflation from January 1980 to December 2012

5 Results

This section contain information on how both models are identified, estimated, diagnostic checked and then used for forecasting for all countries. The differences in forecast accuracy for the two models are then evaluated.

5.1 Stationarity

The graphical presentation of the selected time series which are the coloured parts in Figures 2, 3, 4 and 5 are first observed. There does not seem to be any deterministic trends for any of the time series. Thus, it seem correct to include only an intercept for the unit root tests. The results of the ADF test are presented in Table 3 and for the KPSS test in Table 4. For Luxembourg, Mexican and Swiss monthly inflation the ADF test with the null hypothesis of a unit root is rejected on the five percent significance level. Further for the KPSS test the null of stationarity in not rejected for any of these countries on the five percent significance level. Thus, these three processes appear to be stationary in level.

For Portuguese inflation it is harder to tell since the ADF test leads to not rejecting the null of a unit root while for the KPSS test stationarity cannot be rejected on the five percent significance level. Since it is an important assumption that the time series is stationary differencing the series of Portuguese inflation will still be preferred since stationarity is more certain for that series. Further it might possibly need to be differenced twice but the results of the ADF test for first differenced Portuguese inflation rejects the null of a unit root on the five percent significance level.

Thus, it has been decided that the integration order in level, d, should be equal to zero for every country except for Portugal where taking first difference seem appropriate to make that series stationary in level. Because of this the seasonal unit root tests will be performed for the first differenced series of Portuguese inflation.

The next part is to use tests to evaluate if any of the processes have a seasonal unit root. These tests will be defined with an intercept and seasonal dummies since seasonality is assumed but the trend will not be included following from what was stated for the tests in level. The results of these tests are presented in Table 5 for the Canova-Hansen (CH) test and in Table 6 for the Hylleberg-Engle-Granger-Yoo (HEGY) test.

For the order of seasonal integration, D, the CH test is first used. The results of this test is that the null of stationarity cannot be rejected for any of the frequencies for Luxembourg, Mexican and Swiss inflation on the five percent significance level. The exception of rejection is only true for first differenced Portuguese inflation at frequency $2\pi/3$. Thus there is some reason to doubt if Portuguese inflation should also be seasonally differenced. However for the joint hypothesis of stationarity at all frequencies the null of stationarity cannot be rejected on the five percent significance level.

The second test is the HEGY test for which the null hypothesis of a unit root is tested against the alternative hypothesis of stationarity. The unit root is rejected on the five percent significance level at all frequencies for Luxembourg, Mexican and Swiss inflation. However for differenced Portuguese inflation the null cannot be rejected on the five percent significance level for any of the frequencies except for the test in level (t_{π_1}). This is what is expected for a process that should be differenced in season. The test is then performed on both first and seasonally differenced Portuguese inflation with the result that a unit root is rejected on all frequencies on the five percent significance level. Thus, seasonally differencing first differenced Portuguese inflation seem to have made the time series stationary enough. Hence a seasonal difference operator of order zero seem appropriate for all time series except for first differenced Portuguese inflation where it should be equal to one.

The orders of integration that seem to be needed to make each series stationary enough have now been decided. Thus, the procedure of identifying the SARIMA can continue to the next step which is to determine the autoregressive and moving average orders.

Table 3: ADF test with test-statistic and p-value.

Country	Intercept	p-value
Luxembourg	-3.045	0.033
Mexico	-8.436	0.000
Portugal	-1.383	0.588
Δ(Portugal)	-8.388	0.000
Switzerland	-5.287	0.000

Test of the null of a unit root with test statistic and p-values. The lag length is 11 for Luxembourg and Portugal, 5 for Mexico, 10 for differenced Portugal and 8 for Switzerland.

Table 4: KPSS test with test-statistic and rejection value.

Country	Intercept
Luxembourg	0.153
	(0.463)
Mexico	0.026
	(0.463)
Portugal	0.143
	(0.463)
Switzerland	0.278
	(0.463)

Test of the null of stationarity is rejected for values of test statistic bigger than the rejection value in the parenthesis. The length of the bandwidth was 41 for Luxembourg and Switzerland, 4 for Mexico and 36 for Portugal.

Table 5: CH test with test-statistic and rejection value

Frequency	Luxembourg	Mexico	Δ(Portugal)	Switzerland
$\pi/6$	0.514	0.691	0.357	0.493
	(0.749)	(0.749)	(0.749)	(0.749)
$\pi/3$	0.516	0.613	0.716	0.192
	(0.749)	(0.749)	(0.749)	(0.749)
$\pi/2$	0.535	0.578	0.598	0.242
	(0.749)	(0.749)	(0.749)	(0.749)
$2\pi/3$	0.338	0.273	0.917*	0.397
	(0.749)	(0.749)	(0.749)	(0.749)
$5\pi/6$	0.145	0.613	0.471	0.384
	(0.749)	(0.749)	(0.749)	(0.749)
π	0.346	0.139	0.400	0.227
	(0.470)	(0.470)	(0.470)	(0.470)
ALL	1.584	1.640	1.686	1.544
	(2.750)	(2.750)	(2.750)	(2.750)

Test of the null of stationarity for a specific or all seasonal frequencies is rejected for values of the test statistic bigger than the rejection value in the parentheses. *Significance on the 5% level. The number of truncated lags were 13 for Luxembourg, Mexico and Switzerland, and 12 for first differenced Portuguese inflation.

Table 6: HEGY test with test-statistic and p-value

Statistic	Luxembourg	Mexico	Δ(Portugal)	$\Delta\Delta_s$(Portugal)	Switzerland
t_{π_1}	-3.910	-2.519	-3.596	-2.812	-3.178
	(0.010)	(0.092)	(0.010)	(0.045)	(0.017)
t_{π_2}	-4.586	-3.194	-1.799	-3.081	-2.555
	(0.010)	(0.015)	(0.100)	(0.021)	(0.087)
F_{π_3,π_4}	14.261	6.641	4.779	7.067	13.240
	(0.010)	(0.040)	(0.100)	(0.028)	(0.010)
F_{π_5,π_6}	12.518	9.160	1.412	9.112	8.877
	(0.010)	(0.010)	(0.100)	(0.010)	(0.010)
F_{π_7,π_8}	13.681	10.337	3.187	11.342	6.033
	(0.010)	(0.010)	(0.100)	(0.010)	(0.061)
$F_{\pi_9,\pi_{10}}$	21.578	7.705	4.898	13.624	9.574
	(0.010)	(0.018)	(0.100)	(0.010)	(0.010)
$F_{\pi_{11},\pi_{12}}$	11.057	7.800	5.310	9.643	6.878
	(0.010)	(0.017)	(0.098)	(0.010)	(0.033)

Null of a unit root is tested with the one-sided t-test for some frequencies and the two-sided F-test for pairs of parameters that corresponds to the same frequency. The p-value is found in the parentheses. A constant and seasonal dummies but no trend is specified for each country.

5.2 Autoregressive and Moving Average Orders

In this section appropriate autoregressive orders p and P and moving average orders q and Q will be decided. The first part is to visually check the correlograms for each appropriately differenced process. The result in Table 1 will be used as a guideline to what orders that the shape of the autocorrelation function (ACF) and partial autocorrelation function (PACF) suggests. The results will be presented for each country and then concluded by the orders selected with the HK-algorithm with the information criterion specified as AICc.

The ACF of Luxembourg inflation can be found in Figure 6 and the PACF in Figure 7. The clear significance of both functions at the first lag and the non-significance that follows for the next lags suggest that p and q could possibly be equal but it is also very hard to tell with mixed processes. Further seasonality is clearly indicated since every sixth lag is significant for the autocorrelation function while the sixth and twelfth are significant for the partial one. For the seasonal orders the clear significance of every twelfth lag for the ACF and the twelfth for the PACF suggest that at least one of P and Q should be bigger than one.

The result with the HK-algorithm for Luxembourg can be seen in Table 7 which suggest that a SARIMA$(1, 0, 1) \times (1, 0, 2)_{12}$ should be fitted. This is clearly consistent to the shape of the ACF and PACF.

For Mexican inflation the ACF can be found in Figure 8 and the PACF in 9. The ACF has a clear sine wave pattern with period twelve which is consistent to seasonal data with period twelve. The strongest significance is for lag one and then every sixth lag. For the PACF the sine wave pattern is weaker but significance occurs for the first, sixth and twelfth lag before it tails off. It is hard to tell what orders that should be fitted but the process should clearly be mixed with both autoregressive and moving average orders.

The HK-algorithm with result in Table 7 suggest that a SARIMA$(2, 0, 2) \times (2, 0, 1)_{12}$ should be fitted for Mexican inflation. This seem quite consistent with the result but it is hard to say anything with confidence from the correlogram.

The correlogram for Portuguese differenced inflation is found in Figures 10 and 11. Of interest is that both the ACF and PACF have significant first two lags while seasonality is indicated with the ACF having a twelfth significant lag. This suggest that the seasonal lag orders should be one more for Q compared to P. It is very difficult to say anything about potential orders p and q but there should be at least one value bigger than zero.

For Portugal the HK-algorithm in Table 7 give an indication for a SARIMA$(0, 1, 2) \times (0, 1, 1)_{12}$ being fit for the data. A result that seem acceptable considering the shape of the ACF and PACF.

The ACF and PACF of Swiss inflation are found in Figures 12 and 13. Seasonality is clearly present since every third lag is significant for the ACF while the same but

weaker pattern can be seen for the PACF. The first lags are not significant for any of the functions and it is difficult to draw any conclusions about the lag orders p and q. They might be zero but the process might also be mixed. For the seasonal lag orders P and Q at least some value should be bigger than zero.

The HK-algorithm with result in Table 7 suggest that a SARIMA$(0,0,0) \times (1,0,1)_{12}$ should be fit for Swiss inflation. That does not trail far from what was expected considering the correlogram.

Thus, models have been found for the monthly inflation of each country. The next step is to perform diagnostic checks to evaluate if the fit of each model is good enough.

Table 7: SARIMA$(p,d,q) \times (P,D,Q)_{12}$ chosen by each criterion

Country	AIC	BIC	AICc
Luxembourg	$(1,0,1) \times (1,0,2)$	$(1,0,1) \times (1,0,1)$	$(1,0,1) \times (1,0,2)$
Mexico	$(2,0,2) \times (2,0,1)$	$(2,0,0) \times (2,0,1)$	$(2,0,2) \times (2,0,1)$
Portugal	$(0,1,2) \times (0,1,1)$	$(0,1,1) \times (0,1,1)$	$(0,1,2) \times (0,1,1)$
Switzerland	$(0,0,0) \times (1,0,1)$	$(0,0,0) \times (1,0,1)$	$(0,0,0) \times (1,0,1)$

Selected model for each country and criterion. No intercept is included in any of the models.

ACF of Luxembourg Monthly Inflation
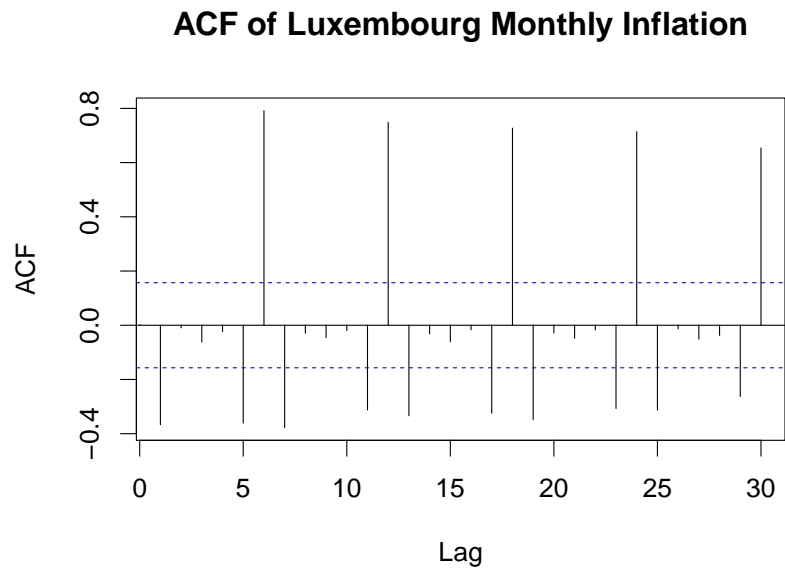
Figure 6: Autocorrelation function for the Luxembourg monthly inflation
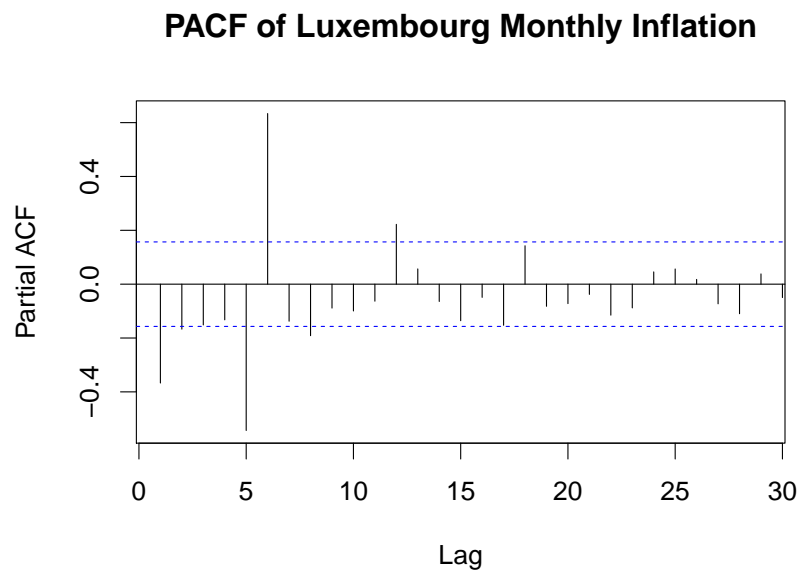
PACF of Luxembourg Monthly Inflation

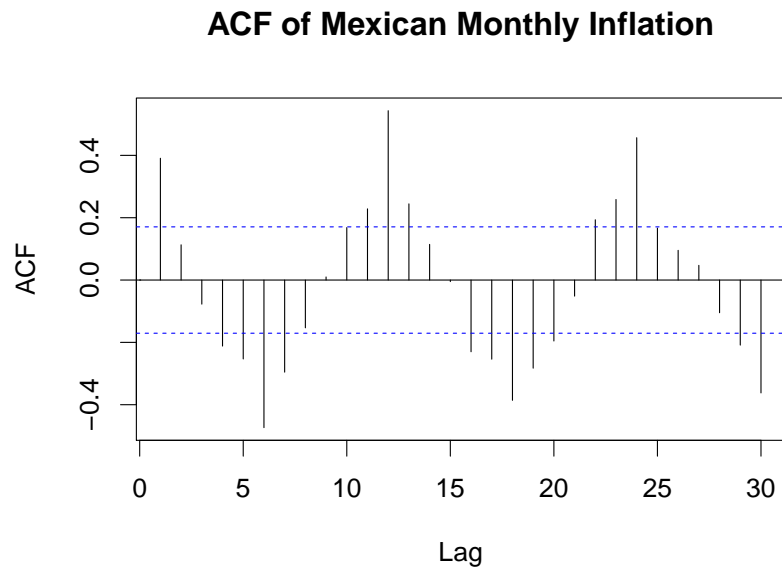Figure 7: Partial autocorrelation function for the Luxembourg monthly inflation

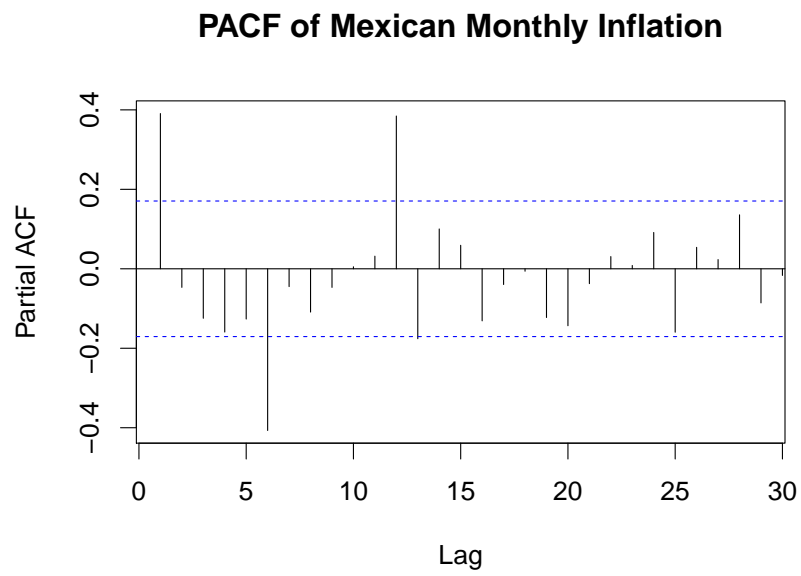Figure 8: Autocorrelation function for the Mexican monthly inflation

Figure 9: Partial autocorrelation function for the Mexican monthly inflation
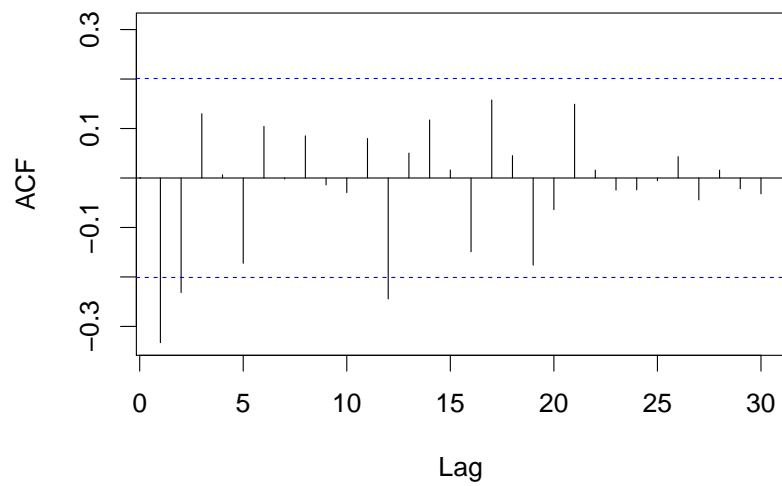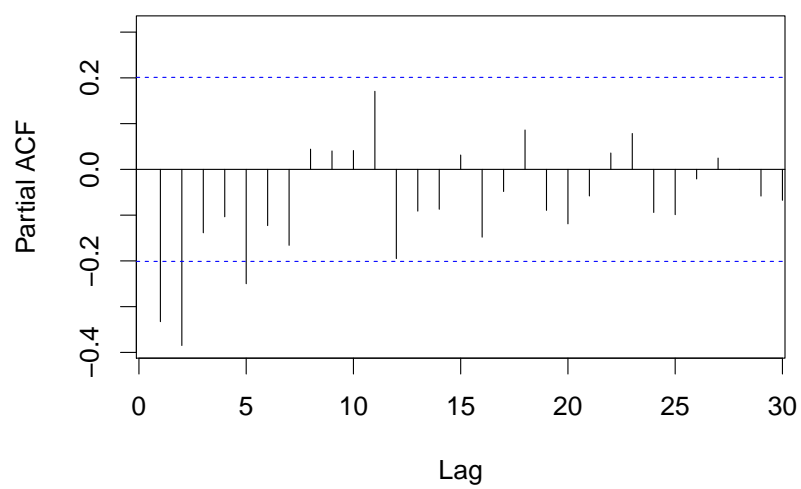
ACF of Differenced Portuguese Monthly Inflation

Figure 10: Autocorrelation function for the Portuguese monthly inflation

PACF of Differenced Portuguese Monthly Inflation

Figure 11: Partial autocorrelation function for the Portuguese monthly inflation
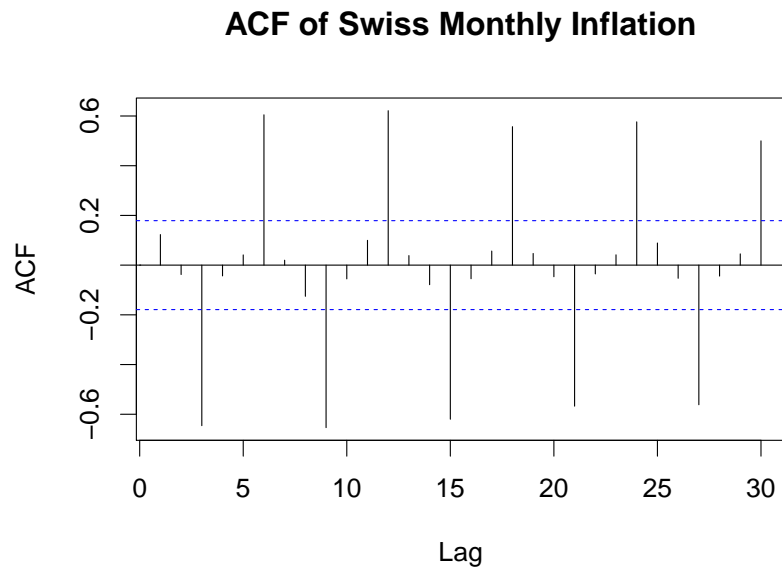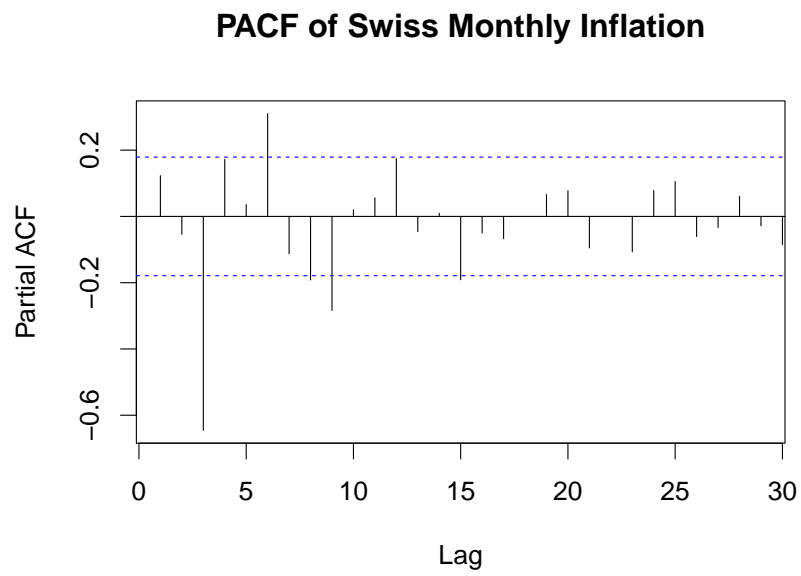
ACF of Swiss Monthly Inflation

Figure 12: Autocorrelation function for the Swiss monthly inflation

PACF of Swiss Monthly Inflation

Figure 13: Partial autocorrelation function for the Swiss monthly inflation

5.3 Estimation and Diagnostic Checking of the SARIMA models

The fitted SARIMA models are found in Tables 8, 9, 10 and 11. The next step is to extract the residuals and first perform a visual evaluation of each models fit and then execute some statistical tests considering their normality and randomness.

Plots of the standardized residuals are found in Figures 14, 15, 16 and 17. For all countries the residuals seem quite homoscedastic and to have a mean close to zero. Further the number of standardized residuals that lie outside the interval ± 1.96 are six (3.85%) for Luxembourg, five (3.79%) for Mexico, four (4.25%) for Portugal and 9 (7.50%) for Switzerland. This result is clearly acceptable for Luxembourg, Mexico and Portugal who all have proportions smaller than five percent that lie outside the interval. However for Switzerland it is slightly too big, further evaluation will be needed to get a more thorough evaluation of that models fit.

The autocorrelation functions for the residuals can be found in Figures 18, 19, 20 and 21. The values for the first 24 lags are presented and the number of significant lags are of interest, meaning those who lie outside the sample size dependent interval $\pm 2/\sqrt{T}$. There are zero significant lags for all models except for Luxembourg who have one significant lag which corresponds to 4.17% of the lags but this is still a smaller proportion than five percent. Thus, the proportion of significant autocorrelations out of the first 24 are under five percent for the residuals extracted from all the models.

The results for the Jarque-Bera tests are presented in Table 12. The null hypothesis of normally distributed residuals cannot be rejected for any of the models on the five percent significance level. This is a good result since the normality assumption cannot be rejected.

The results of the Box-Ljung test can be found in Table 13. This test is performed for each model with the number of autocorrelation lags specified as $\sqrt{T}$ and with degrees of freedom correction for the number of parameters being estimated. The null of randomness cannot be rejected for any of the residuals on the five percent significant level, which is good. Thus, all models seem well-fitted enough and should be suitable for forecasting.

Table 8: SARIMA$(1,0,1) \times (1,0,2)_{12}$ fitted for Luxembourg inflation

Parameter	Estimate	Standard error
$\hat{\phi}_1$	-0.972	0.041
$\hat{\theta}_1$	0.935	0.069
$\hat{\Phi}_1$	0.996	0.003
$\hat{\Theta}_1$	-1.007	0.096
$\hat{\Theta}_2$	0.223	0.098
$\hat{\sigma^2}$	0.071	

For monthly inflation from January 1999 to December 2011, the mean is zero.

Table 9: SARIMA$(2,0,2) \times (2,0,1)_{12}$ fitted for Mexican inflation

Parameter	Estimate	Standard error
$\hat{\phi}_1$	-0.179	0.292
$\hat{\phi}_2$	-0.566	0.255
$\hat{\theta}_1$	0.535	0.217
$\hat{\theta}_2$	0.704	0.328
$\hat{\Phi}_1$	0.999	0.163
$\hat{\Phi}_2$	-0.015	0.155
$\hat{\Theta}_1$	-0.687	0.152
$\hat{\sigma^2}$	0.048	

For monthly inflation from January 2001 to December 2011, the mean is zero.

Table 10: SARIMA$(0,1,2) \times (0,1,1)_{12}$ fitted for Portuguese inflation

Parameter	Estimate	Standard error
$\hat{\theta}_1$	-0.674	0.114
$\hat{\Theta}_1$	-0.230	0.129
$\hat{\Theta}_2$	-0.443	0.117
$\hat{\sigma^2}$	0.107	

For monthly inflation from January 2003 to December 2011, the mean is zero.

Table 11: SARIMA$(0,0,0) \times (1,0,1)_{12}$ fitted for Swiss inflation

Parameter	Estimate	Standard error
$\hat{\Phi}_1$	0.996	0.015
$\hat{\Theta}_1$	-0.876	0.210
$\hat{\sigma^2}$	0.068	

For monthly inflation from January 2002 to December 2011, the mean is zero.
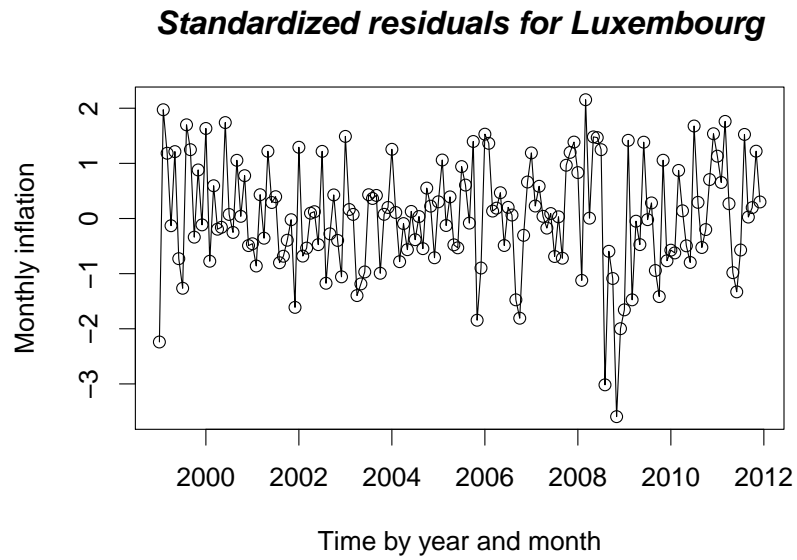
Standardized residuals for Luxembourg

Figure 14: Standardized residuals for Luxembourg from January 1999 to December 2011

Standardized residuals for Mexico

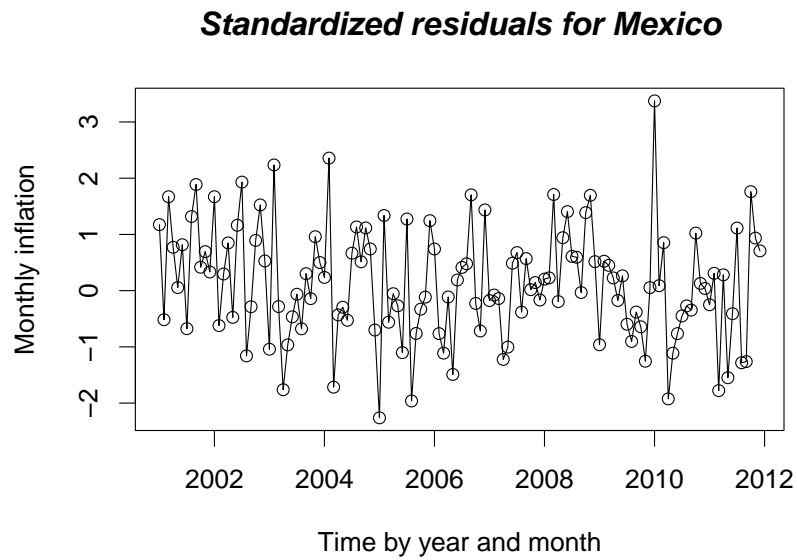
Figure 15: Standardized residuals for Mexico from January 2001 to December 2011
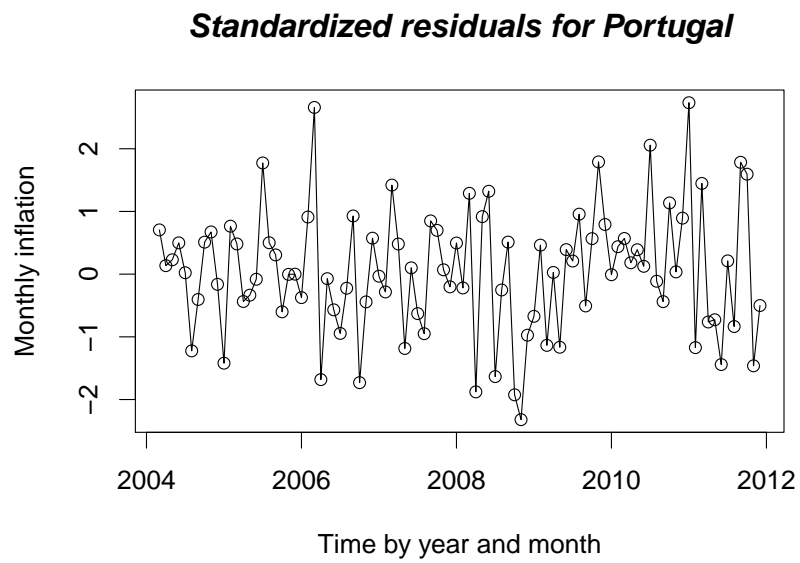
40

Standardized residuals for Portugal

Figure 16: Standardized residuals for Portugal from March 2004 to December 2011

Standardized residuals for Switzerland
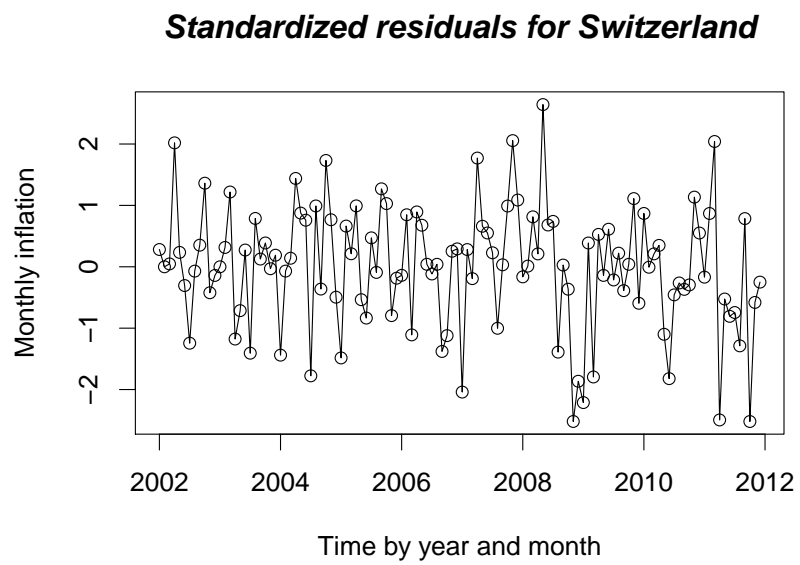
Figure 17: Standardized residuals for Switzerland from January 2002 to December 2011

ACF of Luxembourg Residuals

Figure 18: Autocorrelation function for the Luxembourg residuals
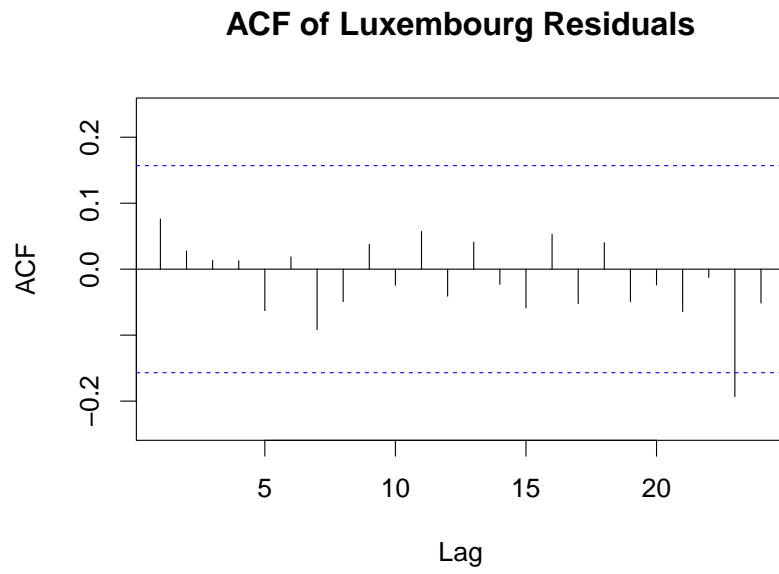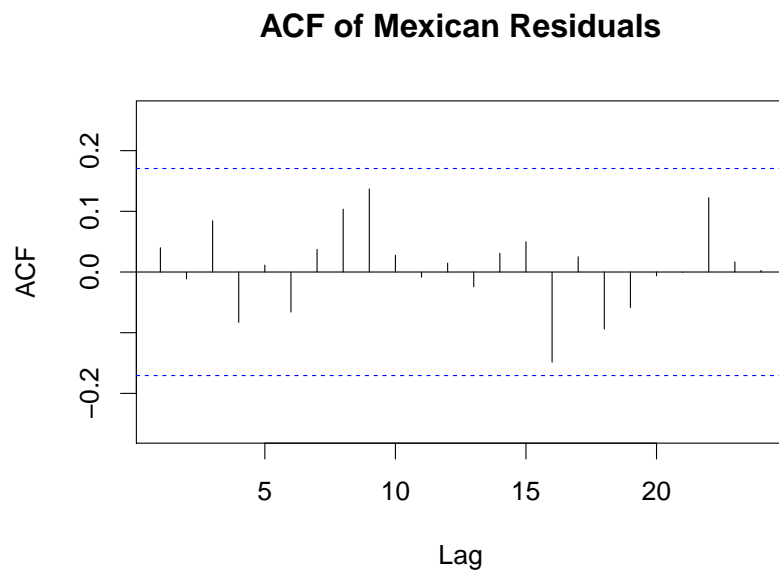

ACF of Mexican Residuals

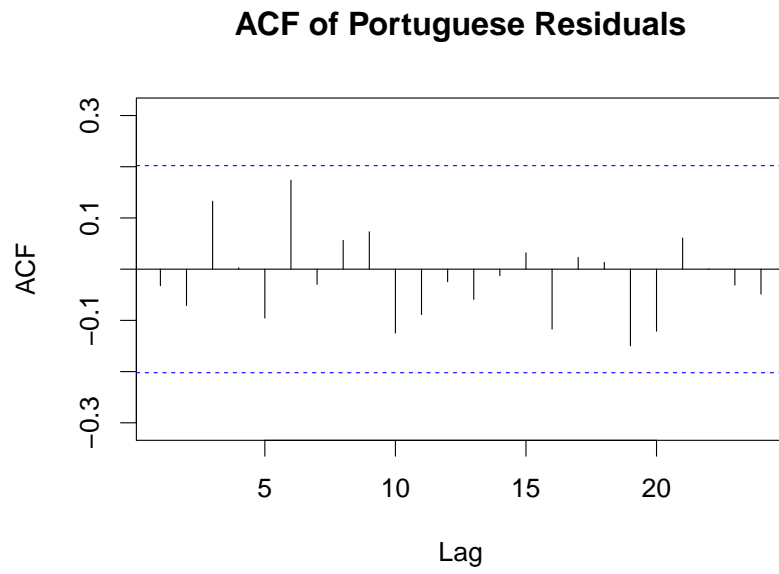Figure 19: Autocorrelation function for the Mexican residuals

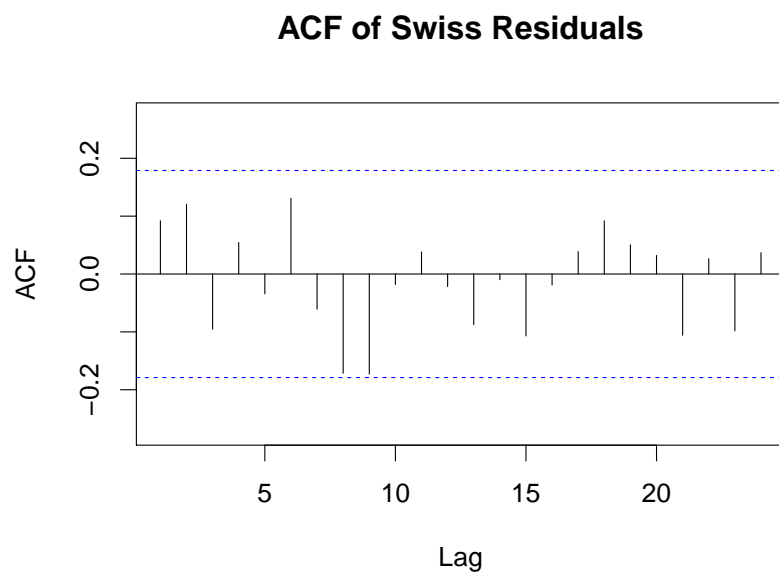Figure 20: Autocorrelation function for the Portuguese residuals

Figure 21: Autocorrelation function for the Swiss residuals

Table 12: Jarque-Bera test for the null of normality for the residuals

	Luxembourg	Mexico	Portugal	Switzerland
JB test	5.385	0.736	0.321	2.024
	(0.068)	(0.692)	(0.852)	(0.363)

Test-statistic with the p-value in the parenthesis for each country.

Table 13: Box-Ljung test of the null of randomness for the residuals

Country	Lag	df	$\tilde{Q}(r)$	p-value
Luxembourg	12	7	4.769	0.688
Mexico	11	4	7.330	0.119
Portugal	10	7	9.029	0.251
Switzerland	11	2	15.210	0.085

5.4 State-Space Models and Application of the Kalman Filter

The first step in this section is to use a subset of the data to estimate the SARIMA model with the same orders as was done in the previous section with all the data. The fitted models can be found in Tables 14, 15, 16 and 17. Thus parameter values have been obtained and the model can be specified in state-space form, that is the model in Equations (39) and (40). For example for Mexican inflation the state-space model is specified by using the relation between the lag polynomials of the ARMA and the SARIMA models as described in Equations (8) and (9). Thus, for this SARIMA$(2,0,2) \times (2,0,1)_{12}$ the matrices are instead specified for an ARMA$(26,14)$ in the following way

$$\boldsymbol{y}_t = y_t,\ \boldsymbol{\xi}_t^\top = \begin{bmatrix} \xi_{1,t} & \xi_{2,t} & \dots & \xi_{26,t} \end{bmatrix},\ \boldsymbol{H}^\top = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix},$$

$$\boldsymbol{F}^\top = \begin{bmatrix} \hat{\phi}_1 & \hat{\phi}_2 & \dots & \hat{\Phi}_1 & -\hat{\Phi}_1\hat{\phi}_1 & -\hat{\Phi}_1\hat{\phi}_2 & \dots & \hat{\Phi}_2 & -\hat{\Phi}_2\hat{\phi}_1 & -\hat{\Phi}_2\hat{\phi}_2 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix},$$

$$\boldsymbol{v}^\top = \begin{bmatrix} 1 & \hat{\theta}_1 & \hat{\theta}_2 & \dots & \hat{\Theta}_1 & \hat{\Theta}_1\hat{\theta}_1 & \hat{\Theta}_1\hat{\theta}_2 & \dots & 0 \end{bmatrix} \epsilon_t,$$

$$\boldsymbol{Q} = \hat{\sigma}^2 \begin{bmatrix}
1 & \hat{\theta}_1 & \hat{\theta}_2 & \ldots & \hat{\Theta}_1 & \hat{\Theta}_1\hat{\theta}_1 & \hat{\Theta}_1\hat{\theta}_2 & \ldots & 0 \\
\hat{\theta}_1 & \hat{\theta}_1^2 & \hat{\theta}_1\hat{\theta}_2 & \ldots & \hat{\Theta}_1\hat{\theta}_1 & \hat{\Theta}_1\hat{\theta}_1^2 & \hat{\Theta}_1\hat{\theta}_1\hat{\theta}_2 & \ldots & 0 \\
\hat{\theta}_2 & \hat{\theta}_1\hat{\theta}_2 & \hat{\theta}_2^2 & \ldots & \hat{\Theta}_1\hat{\theta}_2 & \hat{\Theta}_1\hat{\theta}_1\hat{\theta}_2 & \hat{\Theta}_1\hat{\theta}_2^2 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\hat{\Theta}_1 & \hat{\Theta}_1\hat{\theta}_1 & \hat{\Theta}_1\hat{\theta}_2 & \ldots & \hat{\Theta}_1^2 & \hat{\Theta}_1^2\hat{\theta}_1 & \hat{\Theta}_1^2\hat{\theta}_2 & \ldots & 0 \\
\hat{\Theta}_1\hat{\theta}_1 & \hat{\Theta}_1\hat{\theta}_1^2 & \hat{\Theta}_1\hat{\theta}_1\hat{\theta} & \ldots & \hat{\Theta}_1^2\hat{\theta}_1 & \Theta_1^2\hat{\theta}_1^2 & \hat{\Theta}_1^2\hat{\theta}_1\hat{\theta}_2 & \ldots & 0 \\
\hat{\Theta}_1\hat{\theta}_2 & \hat{\Theta}_1\hat{\theta}_1\hat{\theta}_2 & \hat{\Theta}_1\hat{\theta}_2^2 & \ldots & \hat{\Theta}_1^2\hat{\theta}_2 & \hat{\Theta}_1\hat{\theta}_1\hat{\theta}_2 & \hat{\Theta}_1^2\hat{\theta}_2^2 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0
\end{bmatrix}.$$

With the model in state-space form the next step is to apply the Kalman filter which is specified in Equations (43) and (44) with prior values for the state specified in Equations (41) and (42). The remaining parts of the data meaning from January 2005 to December 2011 for Luxembourg, from January 2007 to December 2011 for Mexico, from January 2008 to December 2011 for Portugal and from January 2008 to December 2011 for Switzerland are then used for filtering.

The innovations derived by Equation (55) are then extracted and assumed to be distributed as Gaussian white noise. Thus diagnostic checking to evaluate the fit of applying the Kalman filter to these state-space models follow.

The standardized innovations can be found in Figures 22, 23, 24 and 25. The number of innovations are quite small but they should appear homoscedastic and have a mean close to zero. The number of standardized residuals that lie outside the interval ± 1.96 are five (5.95%) for Luxembourg, five (8.33%) for Mexico, one (2.86%) for Portugal and two (4.17%) for Switzerland. The result is clearly acceptable for Luxembourg and Portugal with under five percent that lie outside the interval. For Mexico and Switzerland the proportion is a little bit too big, but since the number of innovations is small it would be harsh to reject both models at this point.

The autocorrelation functions for the innovations are found in Figures 26, 27, 28 and 29. The number of significant autocorrelations for the first 24 lags are presented and no more than five percent should lie outside the interval $\pm 2/\sqrt{I}$, where I is the number of innovations. The result is that there are one significant lag (4.17%) for Luxembourg and Mexico, zero (0.00%) for Portugal and two (8.33%) for Switzerland. Thus, the proportion of significant autocorrelations out of the first 24 are smaller than five percent for all sets of innovations except for Switzerland with a slightly too big proportion.

The Jarque-Bera test is used to test the null hypothesis of normally distributed innovations. Results of the tests are presented in Table 18 and the null hypothesis or normality cannot be rejected for any model on the five percent significance level.

The results of the Box-Ljung test can be found in Table 19. The test is performed for each model with the number of autocorrelation lags specified as $\sqrt{I}$ and the result is

that null of randomness cannot be rejected on the five percent significance level for any of the series of innovations.

Thus, the Kalman filter seem to have been consistent to these data sets and specifications and the procedure should be fit for forecasting.

Table 14: SARIMA$(1,0,1) \times (1,0,2)_{12}$ fitted for Luxembourg inflation

Parameter	Estimate	Standard error
$\hat{\phi}_1$	0.167	0.407
$\hat{\theta}_1$	-0.431	0.363
$\hat{\Phi}_1$	0.989	0.010
$\hat{\Theta}_1$	-0.651	0.162
$\hat{\Theta}_2$	0.083	0.172
$\hat{\sigma^2}$	0.054	

For monthly inflation from January 1999 to December 2004, the mean is zero.

Table 15: SARIMA$(2,0,2) \times (2,0,1)_{12}$ fitted for Mexican inflation

Parameter	Estimate	Standard error
$\hat{\phi}_1$	0.502	0.242
$\hat{\phi}_2$	0.498	0.242
$\hat{\theta}_1$	-0.240	0.178
$\hat{\theta}_2$	-0.751	0.177
$\hat{\Phi}_1$	0.266	0.270
$\hat{\Phi}_2$	0.477	0.193
$\hat{\Theta}_1$	0.099	0.308
$\hat{\sigma^2}$	0.050	

For monthly inflation from January 2001 to December 2006, the mean is zero.

Table 16: SARIMA$(0,1,2) \times (0,1,1)_{12}$ fitted for Portuguese inflation

Parameter	Estimate	Standard error
$\hat{\theta}_1$	-0.657	0.188
$\hat{\Theta}_1$	-0.343	0.179
$\hat{\Theta}_2$	-0.180	0.158
$\hat{\sigma^2}$	0.071	

For monthly inflation from January 2003 to December 2007, the mean is zero.

Table 17: SARIMA$(0,0,0) \times (1,0,1)_{12}$ fitted for Swiss inflation

Parameter	Estimate	Standard error
$\hat{\Phi}_1$	0.977	0.028
$\hat{\Theta}_1$	-0.666	0.188
$\hat{\sigma^2}$	0.057	

For monthly inflation from January 2002 to December 2007, the mean is zero.
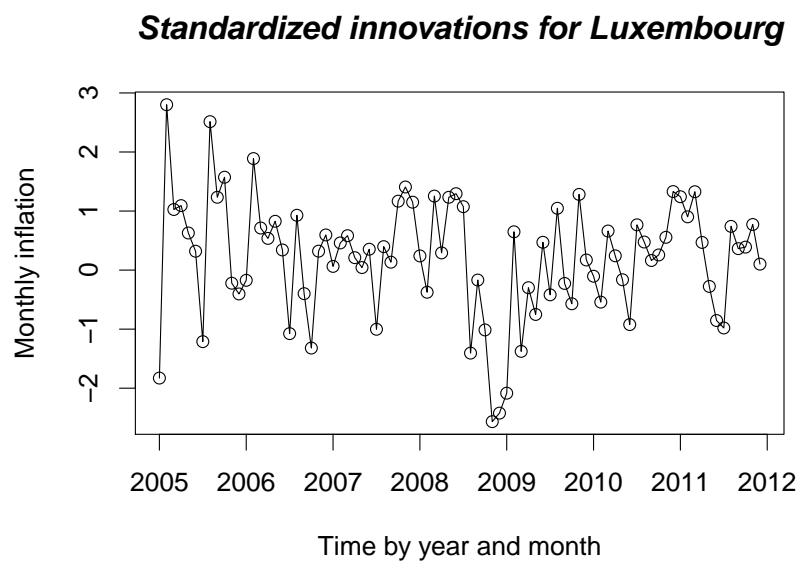
Standardized innovations for Luxembourg

Figure 22: Standardized innovations for Luxembourg from January 2005 to December 2011

Standardized innovations for Mexico
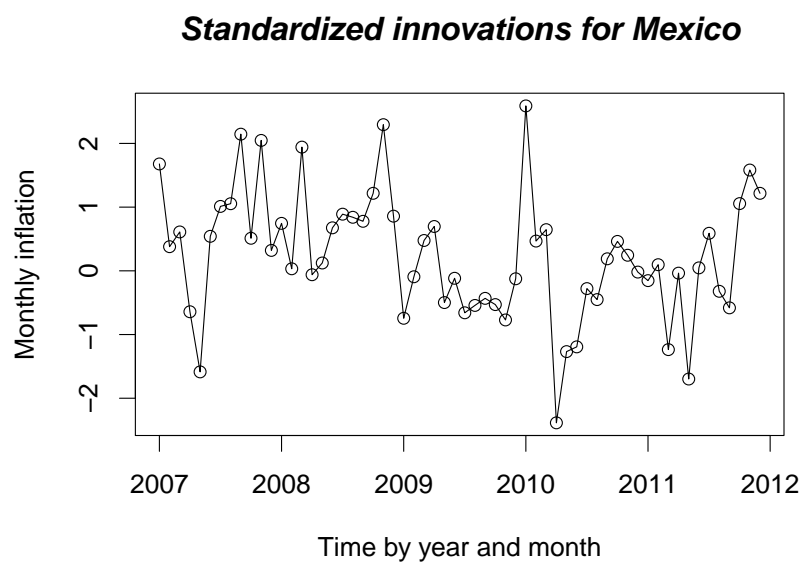
Figure 23: Standardized innovations for Mexico from January 2007 to December 2011
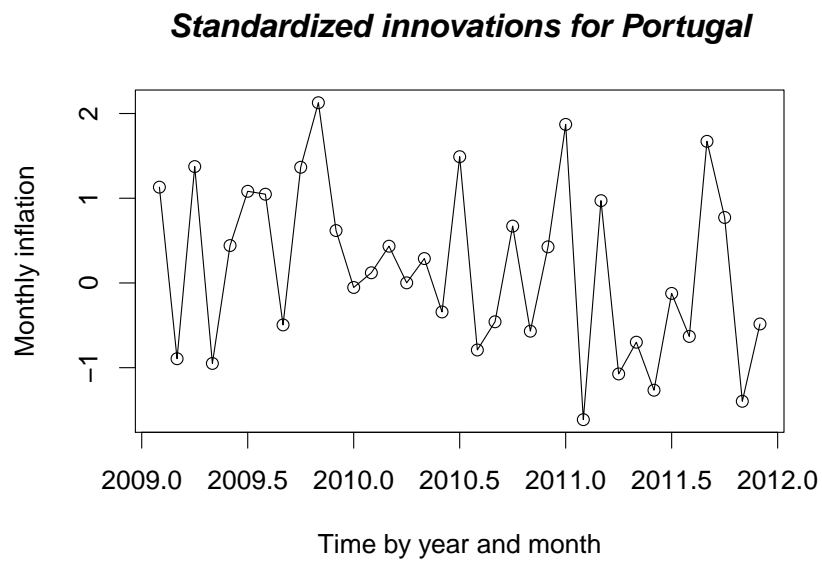
Standardized innovations for Portugal

Figure 24: Standardized innovations for Portugal from February 2009 to December 2011
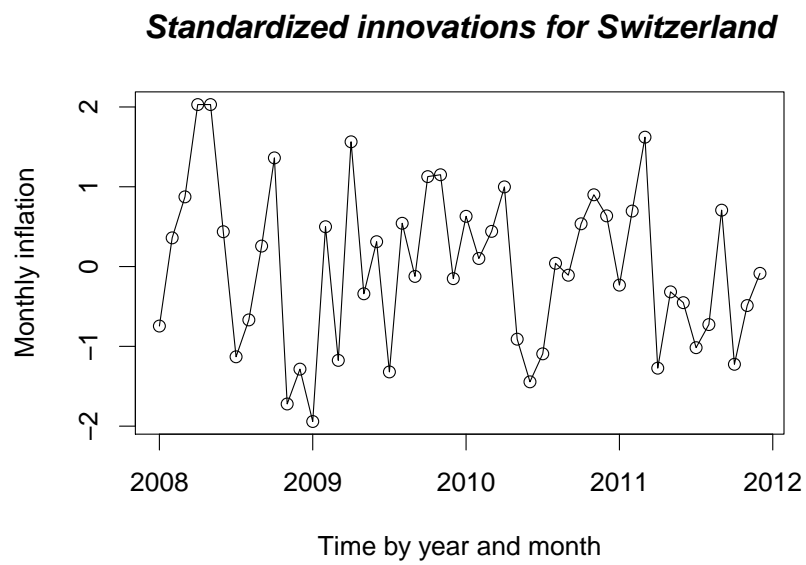

Standardized innovations for Switzerland

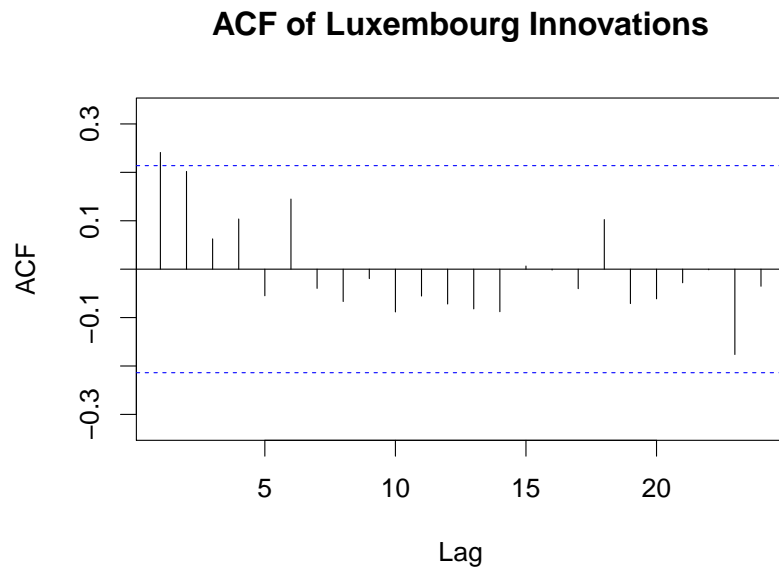Figure 25: Standardized innovations for Switzerland from January 2008 to December 2011

Figure 26: Autocorrelation function for the Luxembourg innovations
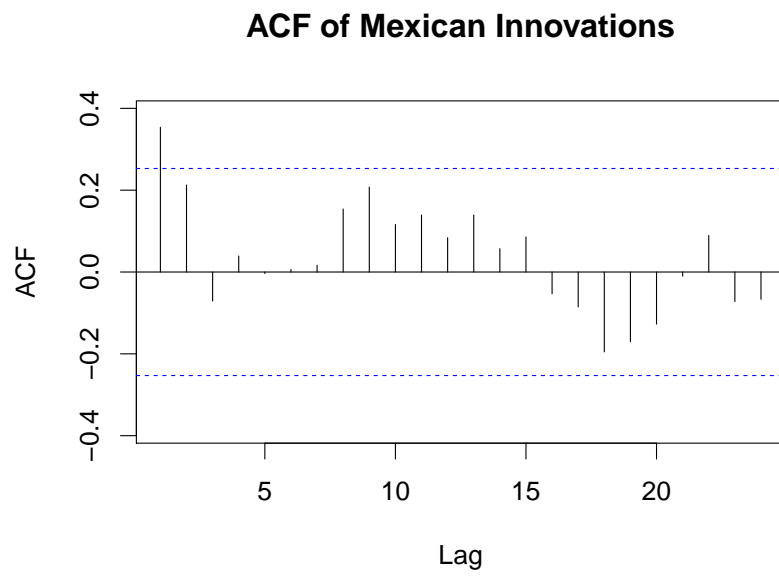

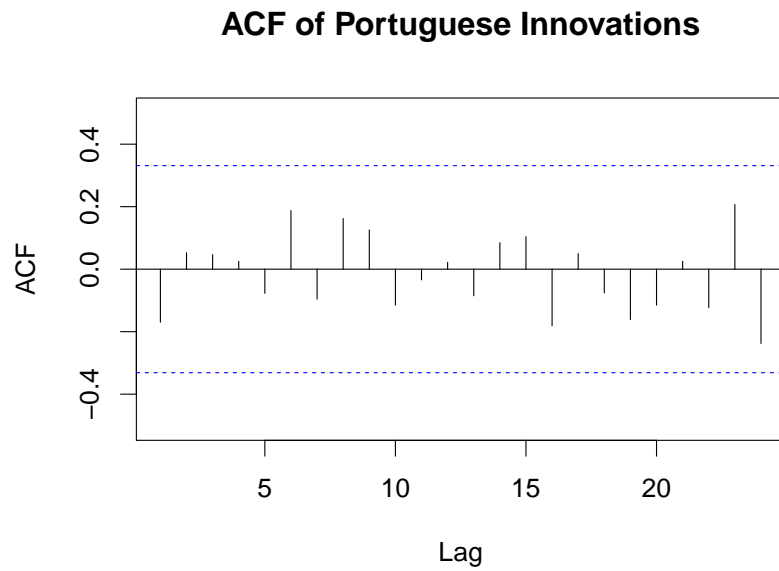Figure 27: Autocorrelation function for the Mexican innovations

ACF of Portuguese Innovations

Figure 28: Autocorrelation function for the Portuguese innovations
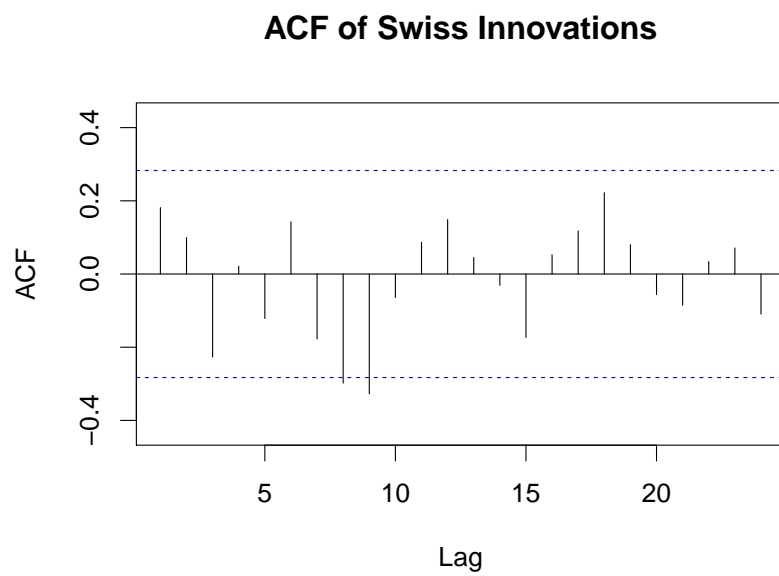

ACF of Swiss Innovations

Figure 29: Autocorrelation function for the Swiss innovations

Table 18: Jarque-Bera test for the null of normality for the innovations

	Luxembourg	Mexico	Portugal	Switzerland
JB test	3.544	0.082	1.510	1.227
	(0.170)	(0.960)	(0.470)	(0.542)

Test-statistic with the p-value in the parenthesis for each country.

Table 19: Box-Ljung test of the null of randomness for the innovations

Country	Lag	df	$\tilde{Q}(r)$	p-value
Luxembourg	9	9	12.791	0.172
Mexico	8	8	12.940	0.114
Portugal	6	6	3.149	0.790
Switzerland	7	7	8.809	0.267

5.5   Comparing Forecasts

The forecasts for all months of 2012 are for the SARIMA models derived as in Equation (36) while for the Kalman filtered models they are obtained with Equation (51).

The model forecasts of monthly inflation for 2012 are plotted against the true values in Figures 30, 31, 32 and 33. All the models appear quite well-fitted, the forecasts follow the true values quite closely.

The error measures for the SARIMA models can be found in Table 20 and for the Kalman filtered models in Table 21. That the models are quite well-fitted is further indicated by the values of the Theil's U statistics which is clearly smaller than one for every model. Further to compare the models it can be seen that the error measures are for all countries slightly smaller for the Kalman filtered models compared to the SARIMA models. The Kalman filtered model have been indicated to be more accurate at forecasting but significant evidence to state that has not yet been obtained.

The Diebold-Mariano test result can be seen in Table 22. All tests are one-sided to test the null hypothesis that the model that implied the smallest MSE, that is the Kalman filtered SARIMA model for each country, has higher predictive ability compared to the other model. The null hypothesis is rejected on the five percent significance level for the difference in forecast accuracy for the models of Swiss inflation. Thus, significant evidence have been obtained to state that the Kalman filtered model is more accurate in forecasting Swiss inflation compared to the SARIMA model. However for the remaining three countries the differences in accuracy are not significant on the five percent significance level.
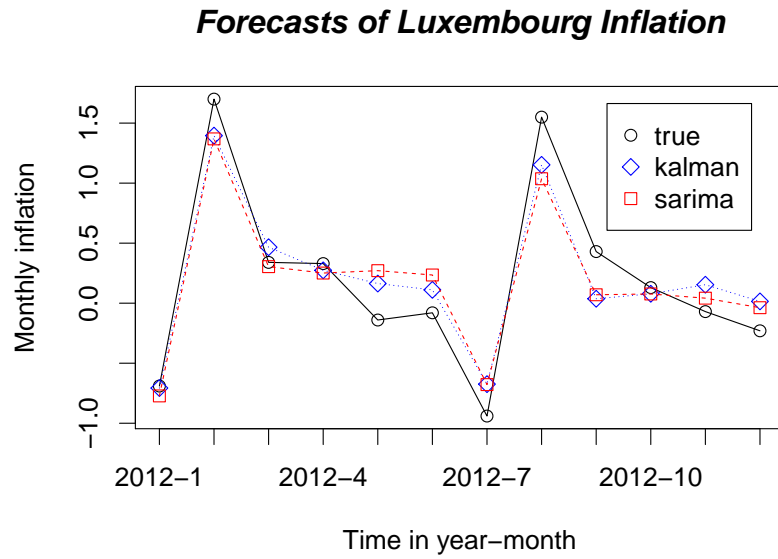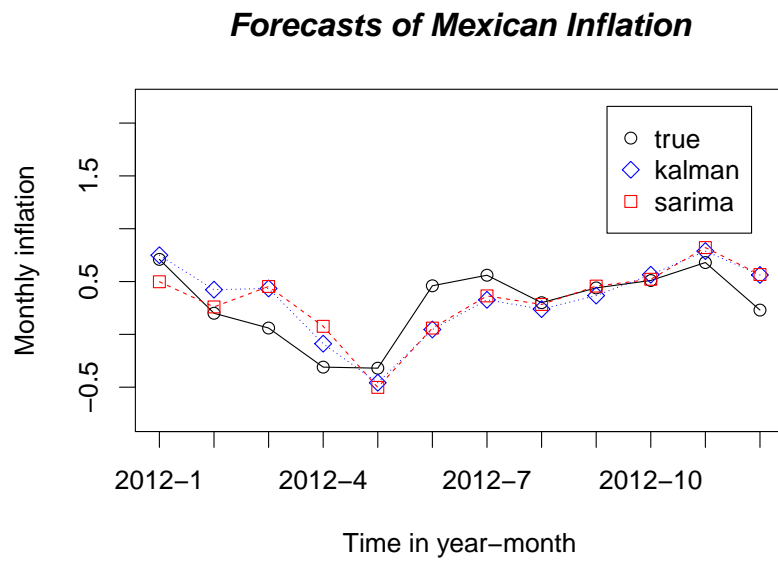
Figure 30: Forecasts and true values of Luxembourg monthly inflation for 2012

Figure 31: Forecasts and true values of Mexican monthly inflation for 2012
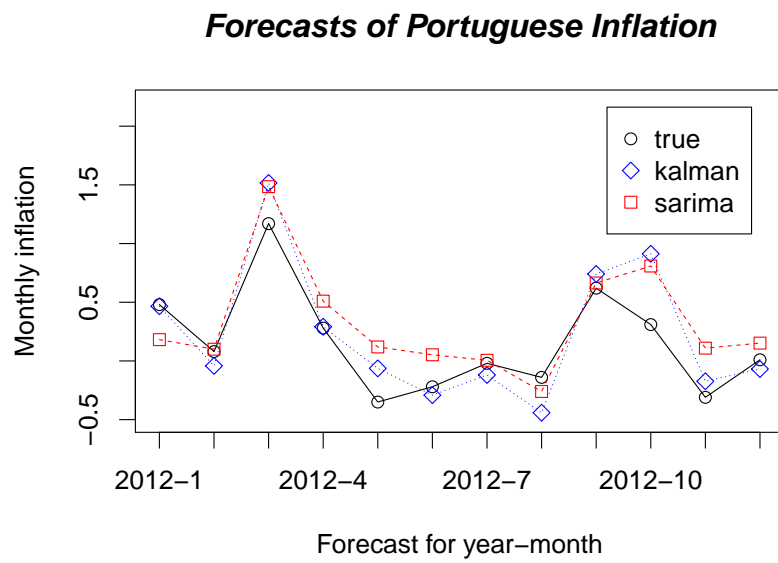
Forecasts of Portuguese Inflation

Figure 32: Forecasts and true values of Portuguese monthly inflation for 2012
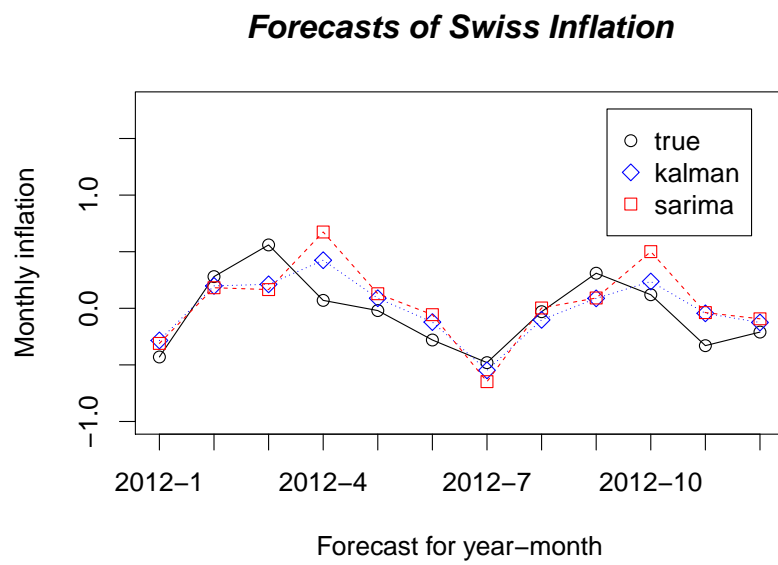
Forecasts of Swiss Inflation

Figure 33: Forecasts and true values of Swiss monthly inflation for 2012

Table 20: Forecast Evaluation with error measures for SARIMA models

Country	MSE	MAD	MAPE	Theil's U
Luxembourg	0.076	0.229	0.985	0.374
Mexico	0.060	0.196	0.998	0.339
Portugal	0.083	0.238	1.991	0.499
Switzerland	0.078	0.233	2.080	0.584

Table 21: Forecast Evaluation with error measures for Kalman filtered SARIMA models

	MSE	MAD	MAPE	Theil's U
Luxembourg	0.061	0.215	0.951	0.337
Mexico	0.052	0.189	0.996	0.315
Portugal	0.060	0.183	1.721	0.425
Switzerland	0.039	0.171	1.486	0.415

Table 22: Diebold-Mariano test for the null of equally accurate forecasts

Country	Lowest MSE	DM	p-value
Luxembourg	Kalman	1.101	0.147
Mexico	Kalman	0.796	0.221
Portugal	Kalman	0.962	0.178
Switzerland	Kalman	1.829	0.047

That the MSE of the Kalman filtered models are smaller compared to the SARIMA models imply that the tests are one-sided with the alternative hypothesis that the Kalman filtered model is more accurate for all countries.

6   Conclusion

In this study SARIMA models and SARIMA models with the Kalman filter applied have been fitted to the monthly inflation of Luxembourg, Mexico, Portugal and Switzerland. All models were proven to be quite well-fitted according to the results of the diagnostic checks.

According to the graphical review of the forecasts both models seem to produce approximately equal forecasts. That is the model being more accurate seem to differ from month to month. However the error measures gave an indication for the Kalman filtered model producing better forecasts. It can be noted that the values of Theil's U is smaller than one for all models. This means that all models performed better compared to the no-change model which just guesses that the next observation will be equal to the present one. That this measure is smaller than one can probably be seen as the minimum requirement for the models being considered to have some kind of predictive ability.

The significance of these differences were then evaluated with the Diebold-Mariano test. The null hypothesis was only rejected for Swiss inflation where the Kalman filtered model was proven to produce better forecasts on the five percent significance level. For the other countries the differences were not significant.

It is difficult to tell how these results should be considered. The forecasts was significantly more accurate for one out of four models. But there does not seem to be any clear pattern between how well the forecasting for each country came out depending on the sample sizes and the different ratios of the number of observations being used for filtering. It seems fitting to now return to the question asked in the beginning of this thesis:

Does applying the Kalman filter to a SARIMA model (estimated with part of the data with the remaining observations used for filtering) lead to better forecasts compared to a SARIMA model that is fitted with all of the data?

The results of the study does not seem clear enough to put the answer as a simple yes or no. The result can probably best be summarized with that the method of applying the Kalman filter to SARIMA models either improve or can be said to at least not hurt the forecasting performance of the model.

References

Bai, J. and Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business and Economic Statistics*, 23(1):49–60.

Banerjee, A., Dolado, J., Galbraith, J., and Hendry, D. (1993). *Cointegration, Error Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford University Press Inc., New York.

Bernanke, B. and Mishkin, F. (1997). Inflation targeting: A new framework for monetary policy? *Journal of Economic Perspectives*, 11(2):97–116.

Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control (Revised Edition).* Holden-Day, Oakland, California, rev. edition.

Box, G., Jenkins, G., and Reinsel, G. (2008). *Time Series Analysis: Forecasting and Control.* Wiley, Oxford, 4 edition.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods.* Springer, New York, 2 edition.

Burnham, K. and Anderson, D. (2004). Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods and Research*, 33(2):261–304.

Canova, F. and Hansen, B. (1992). Are seasonal patterns constant over time? a test for seasonal stability. *Journal of Business and Economic Statistics*, 13(3):237–252.

Çatik, A. and Karaçuka, M. (2012). A comparative analysis of alternative univariate time series models in forecasting turkish inflation. *Journal of Business Economics and Management*, 13(2):275–293.

Commandeur, J., Koopman, S., and Ooms, M. (2011). Statistical software for state space methods. *Economics Letters*, 41(1):1–18.

Davidson, R. and MacKinnon, J. (2003). *Econometric Theory and Methods.* Oxford University Press, USA, New York.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):134–144.

Durbin, J. and Koopman, S. (2001). *Time Series Analysis by State Space Methods.* Oxford University Press (Sd), Oxford, 1 edition.

inflation.eu Worldwide Inflation Data (2013). Accessed February 26th, 2013, <http://www.inflation.eu/>.

Engle, R. (1979). Estimating structural models of seasonality. In Zellner, A., editor, *Seasonal Analysis of Economic Time Series*, pages 281–308. NBER.

Franses, P. (1991). Seasonality, non-stationarity and the forecasting of monthly time series. *International Journal of Forecasting*, 7(2):199–208.

Ghysels, E., Lee, H., and Noh, J. (1994). Testing for unit roots in seasonal time series. *Journal of Econometrics*, 62(2):415–442.

Grosswindhager, S., Voigt, A., and Kozek, M. (2011). Online short-term forecast of system heat load in district heating networks. *In Proceedings of the 31st International Symposium on Forecasting. Prague, Czech Republic.*

Hamilton, J. (1985). Uncovering financial market expectations of inflation. *Journal of Political Economy*, 93(6):1224–1241.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, 1 edition.

Harvey, A. and Phillips, G. (1979). Maximum likelihood estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, 66(1):49–58.

Hurvich, C. and Anderson, T. C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307.

Hylleberg, S. (1995). Tests for seasonal unit roots. general to specific or specific to general? *Journal of Econometrics*, 69:5–25.

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

Hyndman, R. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22.

Jarque, C. and Bera, A. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Junttila, J. (2001). Structural breaks, arima model and finnish inflation forecasts. *International Journal of Forecasting*, 17(2):203–230.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3):159–178.

Ljung, G. and Box, G. (1978). On a measure of a lack of fit in time series models. *Biometrika*, 65(2):297–303.

Mariano, R. and Preve, D. (2012). Statistical tests for multiple forecast comparison. *Journal of Econometrics*, 169(1):123–130.

Meyler, A., Kenny, G., and Quinn, T. (1998). Forecasting irish inflation using arima models. *Economic Analysis, Research and Publications Department, Central Bank of Ireland.*

Morrison, G. and Pike, D. (1977). Kalman filtering applied to statistical forecasting. *Management Science*, 23(7):768–774.

Norman, D. and Richards, A. (2012). The forecasting performance of single equatios models of inflation. *The Economic Record*, 88(208):64–78.

Omane-Adjepong, M., Oduro, F., and Odure, S. (2013). Determining the better approach for short-term forecasting of ghana's inflations: Seasonal-arima vs. holt-winters. *International Journal of Business, Humanities and Technology*, 3(1):69–79.

Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic Linear Models with R*. Springer, New York.

Rosenberg, B. (1973). Random coefficient models: the analysis of a cross-section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement*, 2(4):399–428.

Saz, G. (2011). The efficacy of sarima models for forecasting inflation rates in developing countries: The case for turkey. *International Research Journal of Finance and Economics*, 62:111–142.

Schwert, G. (2009). *EViews 7 User's Guide II*. Quantitative Micro Software, LLC, Irvine, California.

Tarno, Subanar, Rosadi, D., and Suhartono (2012). New procedure for determining order of subset autoregressive integrated moving average (arima) based on over-fitting concept. *International Conference on Statistics in Science, Business and Engineering, ICSSBE, Langkawi, Kedah, Malaysia.*

Zhang, Y. and Li, R. (1996). Fixed-interval smoothing algorithm based on singular value decomposition. *In Proceedings of the 31st International Symposium on Forecasting. Prague, Czech Republic*, page 916–921.

Zivot, E. and Andrews, D. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 10(3):251–270.

Appendix

Some important concepts that hold a huge relevance for this thesis are presented here.

Normality

For a time series process Y_t to follow the normal distribution with mean μ and variance σ^2, that is that $Y_t \sim N(\mu, \sigma^2)$ it should have the following probability density function (Hamilton, 1994, p. 745-746)

$$f(y_t) = \frac{1}{\sqrt{2\pi}\sigma} exp[\frac{-(y_t - \mu)^2}{2\sigma^2}] \tag{62}$$

Further we need that the centered moments of odd order are equal to zero, that is

$$E(Y_t - \mu)^r = 0 \text{ for } r = 1, 3, 5, ... \tag{63}$$

and that the centered fourth moment

$$E(Y_t - \mu)^4 = 3\sigma^4. \tag{64}$$

Two important factors while constructing the JB normality test are the skewness which is the standardized third moment which if positive decides that it is more likely to observe a bigger number compared to the mean

$$S = \frac{E(Y_t - \mu)^3}{[Var(Y_t)]^{3/2}} \tag{65}$$

and the kurtosis which is the standardized fourth moment shows the relationship between the amount of density close to the mean and in the tails. If it is bigger than 3 then we have comparably more density in the tails compared to the normal distribution

$$K = \frac{E(Y_t - \mu)^4}{[Var(Y_t)]^2}. \tag{66}$$

Stationarity

For a time series process Y_t to be stationary, in this thesis this implies covariance-stationarity. Hamilton (1994, p. 45) state that the following two equations should be fulfilled for this to be true

$$E(Y_t) = \mu \text{ for all } t, \tag{67}$$

$$E(Y_t - \mu)(Y_{t-j} - \mu) = \gamma_j \text{ for all } t \text{ and any } j \tag{68}$$

where the mean is μ and γ_j is the autocovariances. That is we need that both the mean and the autocovariances are independent of time.

Unit Root

The concept of unit roots is closely related to stationarity, if a process is said to possess a unit root then it can also be said to be non-stationary. This relationship is often used in tests of stationarity where the null hypothesis instead consists of testing for a unit root. The model that is used to describe a unit root process is written on the following form

$$y_t = \alpha + \delta t + u_t \tag{69}$$

where y_t is the observed process α is an intercept, δ is the parameter for a simple time trend and u_t is assumed to follow a zero mean ARMA process, described by

$$\phi(L)u_t = \theta(L)\epsilon_t \Leftrightarrow (1 - \phi_1 L - \ldots - \phi_p L^p)u_t = (1 + \theta_1 L + \ldots + \theta_p L^p)\epsilon_t \tag{70}$$

where $\phi(L)$ is the autoregressive lag polynomial and $\theta(L)$ is the invertible moving average lag polynomial. The autoregressive lag polynomial is then factored as $\phi(L) = (1 - \theta_1) \ldots (1 - \lambda_p L)$. The process is then assumed to have a unit root if any of the eigenvalues θ_i lies outside the unit circle. It is then assumed that a process that suffers from a unit root can become stationary by first differencing. The seasonal unit root depends on the seasonal period of the data and also implies eigenvalues outside the unit circle. A seasonal unit root of integration order one would imply that we would have to differentiate the time series with its first seasonal lag. (Hamilton, 1994, p. 437)