

Forecasting House Prices in Norway:

A Univariate Time Series Approach

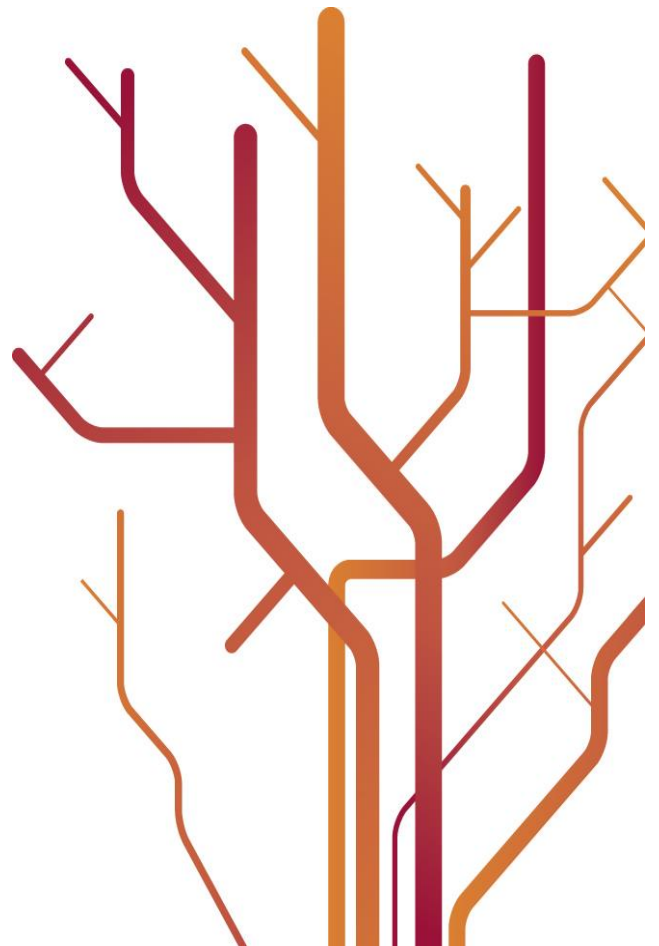


Lars Fiva Skarbøvik

Master's Thesis in Business Administration

Field of Study - Economic Analysis and Finance

May 2013



PREFACE

This thesis concludes my Master's degree at Tromsø University Business School. The thesis is mandatory and constitutes 30 ECTS.

I would like to thank my supervisor Øystein Myrland both for suggesting research topic and for helpful guidance throughout the writing process.

Tromsø, May 2013

Lars Fiva Skarbøvik

ABSTRACT

A wide range of decision makers is interested in forecasts of house prices. The main objective of this thesis is to forecast residential house prices in Norway from April 2013 to March 2014. Three univariate (one variable) time series models are employed in an attempt to find an appropriate fit. The three are an AR-, an ARIMA- and an exponential smoothing state space (ETS) model. The forecast from the three models are also combined, in an effort to improve upon the accuracy of the single “best” forecast. This study implements a weighting scheme based on inverse out-of-sample mean square errors (MSEs). Weights of 0.29, 0.21 and 0.50 are assigned to the AR-, ARIMA- and ETS-models, respectively.

The analysis identifies the forecast from the ETS-model as the most accurate, among the individual models, based on both out-of-sample root mean square error (RMSE) and mean absolute scaled error (MASE). The weighted forecast has a higher RMSE (less accurate), but a lower (more accurate) MASE compared to the ETS. Thus, we cannot reject the idea that a combination of forecast can in fact improve upon the accuracy of the single best forecast, since the two measures give conflicting results.

Residual diagnostics discover some suboptimal attributes of the AR- and ETS-models. Formal tests show that they have significant autocorrelation and heteroscedasticity in their residuals. By examining the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots we find that the residuals appear to be reasonably well behaved. Hence, it does not appear that we need to be too worried about spurious results.

Keywords: *ARIMA, AR process, exponential smoothing, forecasting, forecast combination, house prices*

Table of Contents

| | |
|--|-----|
| PREFACE | ii |
| ABSTRACT | iii |
| List of figures | v |
| List of tables | v |
| 1 Introduction..... | 1 |
| 1.1 Problem formulation..... | 2 |
| 1.2 Thesis structure | 2 |
| 2 Theoretical framework | 3 |
| 2.1 Time series forecasting..... | 3 |
| 2.1.1 AR and ARIMA models..... | 4 |
| 2.1.2 Exponential smoothing and state space models..... | 6 |
| 2.1.3 Other models..... | 9 |
| 2.1.4 Model selection | 9 |
| 2.1.5 Combining forecasts..... | 10 |
| 2.1.6 Forecast evaluation | 12 |
| 2.2 Forecasting house prices..... | 13 |
| 3 Methodology | 16 |
| 3.1 Model estimation | 16 |
| 3.2 Forecast combination..... | 18 |
| 3.3 Diagnostics | 18 |
| 3.4 Accuracy | 20 |
| 4 Data..... | 22 |
| 4.1 The data set..... | 22 |
| 4.2 Data exploration..... | 23 |
| 5 Results..... | 26 |
| 5.1 Forecasts..... | 27 |
| 5.2 Diagnostics | 30 |
| 5.3 Discussion | 33 |
| 6 Concluding remarks..... | 35 |
| 7 References | 36 |
| A. Appendix - The estimated models..... | 41 |
| B. Appendix - Point forecasts and prediction intervals | 42 |
| C. Appendix - The data set..... | 44 |
| D. Appendix - The R code..... | 45 |

List of figures

| | |
|---|----|
| Figure 1.1: Monthly m2 prices (in NOK 1000) in Norway from 2002:1 to 2013:3. | 1 |
| Figure 4.1: First differenced house prices 2002:1-2013:3. | 24 |
| Figure 4.2: Decomposition of the house price series. | 25 |
| Figure 5.1: Forecast for h=12 months ahead with an AR(14) model..... | 27 |
| Figure 5.2: Forecast for h=12 months ahead with an ARIMA(2,0,1)[0,1,1][12] model. | 28 |
| Figure 5.3: Forecast for h=12 months ahead with an ETS(A,A,A) model. | 28 |
| Figure 5.4: Weighted forecast (blue line), 80% (grey) and 95% (red) prediction intervals. .. | 29 |
| Figure 5.5: Residual diagnostics for AR(14). | 31 |
| Figure 5.6: Residual diagnostics for ARIMA(2,0,1)(0,1,1)[12]. | 32 |
| Figure 5.7: Residual diagnostics for ETS(A,A,A) | 32 |

List of tables

| | |
|--|----|
| Table 2.1: Exponential smoothing methods. | 7 |
| Table 4.1: Descriptive statistics for nominal house prices 2002:1-2013:3. | 23 |
| Table 5.1: Forecasted yearly %-price change, month-by-month. | 29 |
| Table 5.2: Forecast accuracy for all models..... | 30 |
| Table 5.3: Residual diagnostics tests..... | 33 |
| Table A.1: The AR(14) with constant and trend. | 41 |
| Table A.2: The ARIMA(2,0,1)(0,1,1)[12] with drift | 41 |
| Table A.3: The ETS(A,A,A) | 41 |
| Table B.1: Point forecast and prediction intervals for 2013:4-2014:3, individual models. | 42 |
| Table B.2: Point forecast and prediction intervals for 2013:4-2014:3, weighted model. | 43 |
| Table C.1: Monthly nominal m2 house prices (in NOK1000) in Norway 2002:1-2013:3. | 44 |

Abbreviations and acronyms

| | |
|------------------|---|
| ACF | Autocorrelation function |
| ADF | Augmented Dickey-Fuller test |
| AIC | Akaike information criterion |
| AIC _c | Akaike information criterion corrected |
| AR | Autoregressive |
| ARCH | Autoregressive conditional heteroskedasticity |
| ARIMA | Autoregressive integrated moving average |
| BIC | Bayesian information criterion |
| ECM | Error correction model |
| ETS | Error trend seasonality (notation for state space models) |
| GARCH | Generalized autoregressive conditional heteroskedasticity |
| IC | Information criterion |
| KL | Kullback-Leibler distance |
| KPSS | Kwiatkowski-Phillips-Schmidt-Shin test |
| m ² | Square meter |
| MASE | Mean absolute scaled error |
| MLE | Maximum likelihood estimation |
| MSE | Mean square error |
| NOK | Norwegian Krone |
| OCSB | Osborn-Chui-Smith-Birchenhall test |
| OLS | Ordinary least square |
| PACF | Partial autocorrelation function |
| RMSE | Root mean square error |
| VAR | Vector autoregressive |

1 Introduction

The forecasting of macroeconomic measures is of great interest to numerous agents in an economy, such as banks, business owners and policymakers. With what seems like daily media coverage, house prices is one such measure that appears to be of particular interest. This interest may stem from the fact that house prices affect such a wide range of decision makers. Movement in house prices affect not only central and commercial banks and their interest rate policies, but also nearly every household.

In recent years, house prices in Norway have increased quite steadily. Figure 1.1 shows nationwide nominal square meter (m²) prices for all types of housing in the period January 2002 to March 2013. Overall prices has increased by 113% in the period. The steady increase is only interrupted by the turmoil in the period from late 2007 to 2009 often referred to as the financial crisis.

The objective of this thesis to guide the reader through a robust analysis of univariate time series forecasting. In the context of statistics, a robust procedure is one that is not heavily dependent on whatever assumptions it makes (Larsen & Marx, 2012). The analysis will culminate in the development of a model constructed to forecast house prices.

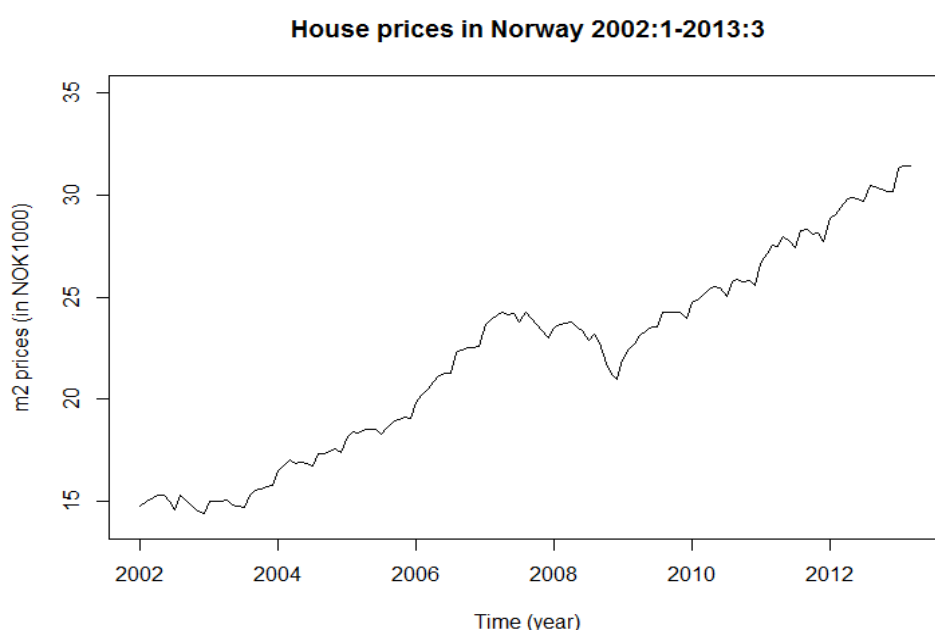


Figure 1.1: Monthly m² prices (in NOK 1000) in Norway from 2002:1 to 2013:3.

The focus of the analysis is forecasting with univariate (one variable) time series models. Especially the exponential smoothing state space models applied here are not widely used in forecasting of house prices. In that sense, this thesis may shed some light on the applicability of exponential smoothing models on house price data.

1.1 Problem formulation

The main goal of this thesis is to forecast residential house prices in Norway for the 12 months from April 2013 to March 2014. Several models are considered in an attempt to find an appropriate fit. I will also examine whether a combination of several forecasts can improve upon the accuracy of the single “best” forecast.

The forecasts are purely based upon univariate time series analysis, i.e. there is no attempt at identifying underlying factors affecting house prices.

1.2 Thesis structure

The thesis is structured as follows. Chapter 2 contains a *literature review* of relevant subjects, mostly focusing on economic time series forecasting. Here all relevant theory is presented. Chapter 3 presents the *methodology* framework used in the analysis. The framework includes procedures on how the models are selected and estimated, as well as how they are combined. Chapter 4 presents the *data* and its properties. Chapter 5 presents the *results* of all the models’ forecasts, in addition to a thorough evaluation. Chapter 6 contains a brief *discussion* of the limitations of the thesis and some *concluding remarks*. Lastly, the reader can find the *bibliography* and an *Appendix* containing the R script file and the data set.

2 Theoretical framework

This chapter contains a review of literature relevant to this thesis. Most of the thesis' theoretical framework is presented here. The first part concerns theory on time series forecasting in general. The second part is a review of literature specifically concerning house prices.

2.1 Time series forecasting

A time series is sequence of successive data points typically taken at equally spaced time intervals. When modelling time series data, the main objective is to develop models that is as close as possible to the true, but unknown data generating process.

An important part of the literature on time series forecasting is the issue of model selection versus combining forecasts. With a plethora of available statistical models (sometimes referred to as experts) to choose from, researchers have the difficult choice of selecting a single expert or combining several that is to make up the final forecast. Model selection is discussed in section 2.1.4 and forecast combination in section 2.1.5.

Another important part of the literature revolves around the decomposition of the time series. The decomposition can be viewed as having the following four components. A trend (T) component that captures the change in the mean of the series, a seasonal (S) component that captures the repetition of a pattern in the series after some fixed time interval, a cycle (C) component that captures a repeating pattern, but with unknown and/or changing periodicity and an error (E) component. The cyclical element is not discussed further in this thesis. Any cyclical component in the data is assumed to be captured in the trend and/or seasonal components, as is consistent with the analysis in Hyndman et al. (2008). Several methods exist to decompose a time series e.g. Holt's method, Holt-Winters' method and other exponential smoothing methods. See section 4.2 for an analysis of the decomposed series of house prices.

After the models are selected they can be used to compute forecasts. Forecasters commonly apply experts such as neural networks, ARCH-, ARIMA- and exponential smoothing models. The focus of this thesis is forecasting using univariate time series models. Next, some of these

univariate models are discussed, as well as model selection, combination and evaluation theory.

2.1.1 AR and ARIMA models

One of the more popular experts among forecasters is the autoregressive integrated moving average (ARIMA) model (e.g. Crawford & Fratantoni, 2003; Dhrymes & Peristiani, 1988; Hein & Spudeck, 1988). An ARIMA model consists of three parts. An autoregressive component (AR) indicating the number of lags of the dependent variable that is to be included, a component indicating order of integration (I) and a moving average (MA) component that captures the effect of lagged values of the error term. It was the early work of among others Yule (1927) that lay the foundation for the development of AR and MA models. Box and Jenkins (1970) integrated the earlier work in this field and developed a three-stage approach for identifying, estimating and verifying ARIMA models. The Box-Jenkins method, as it has come to be known, is still widely used today.

The AR process is by itself a widely used model among forecasters (Bjørnland et al., 2012; Marcellino et al., 2006; Sklarz et al., 1987). It is most common that the pure AR process is included in the literature in a benchmark capacity. That is, if a simple AR model best forecasts a series then there is no point in investing resources in more complex models. The first order autoregressive process, AR(1), is simply the linear relationship between a dependent variable and its own lag. The AR(1) can be expressed as

$$y_t = \lambda + \theta_1 y_{t-1} + \varepsilon_t \quad (2.1)$$

where y_t is an observed time series and ε_t are error terms which we assume to be independent and identically distributed with expectation zero and variance σ^2 . We can then write $\varepsilon_t \sim \text{IID}(0, \sigma^2)$. In (2.1) a constant term λ is included, though this is not always the case. The model parameters, θ_i for $i = (0,1)$, need to be estimated. There are several estimation procedures to choose from such as method of moments, ordinary least square (OLS) and maximum likelihood estimation (MLE).

More generally, a p th order autoregressive process AR(p) is given by

$$y_t = \lambda + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} + \varepsilon_t \quad (2.2)$$

where p is the lag order and $t = (1, 2, \dots, n)$. Similarly, a first order moving average process, MA(1), can be expressed as

$$y_t = \lambda + \varepsilon_t - \phi_1 \varepsilon_{t-1} \quad (2.3)$$

where ε_t are the error terms. And more generally, a q th order moving average process, MA(q) is

$$y_t = \lambda + \varepsilon_t - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} \quad (2.4)$$

where q is the lag order of the error term. By combining the AR(p) and MA(q) models one can express an ARMA(p, q) model as

$$y_t = \lambda + \theta_1 y_{t-1} + \dots + \theta_p y_{t-p} + \varepsilon_t - \phi_1 \varepsilon_{t-1} - \dots - \phi_q \varepsilon_{t-q} \quad (2.5)$$

By following the notation in Hyndman et al. (2008), another way of expressing the ARMA model is

$$\theta(L)y_t = \lambda + \phi(L)\varepsilon_t \quad (2.6)$$

where L is the backshift/lag operator, λ is the intercept and $\theta(z)$ and $\phi(z)$ are polynomials of order p and q respectively. To clarify, the lag operator means $Ly_t = y_{t-1}$. Before moving from the ARMA to the ARIMA model, we need to introduce an important concept in any time series analysis, namely stationarity.

It is common to distinguish between two types of stationarity in time series analysis. First, a series is said to be *difference* (or *weakly*) *stationary* if it has a constant mean and variance, and the covariance between two values in the series depends only on the time separating them, and not on the time they are observed. Normally when we talk about stationarity, we mean difference stationary. Second, a series is said to be *trend stationary* if it becomes stationary after subtracting the deterministic (constant and trend) components (Hill et al., 2008). The label deterministic is used so to distinguish between a constant (deterministic) trend and stochastic trends. An alternative to de-trending the series, is keeping both the constant and trend in the model. Using the AR(1) to illustrate, we can express the process with a constant and a deterministic trend as

$$y_t = \lambda + \delta t + \theta_1 y_{t-1} + \varepsilon_t \quad (2.7)$$

where t is time and δ the coefficient of the trend term. Simply by looking at the house prices in Figure 1.1, it is easy to see that the series does not have a constant mean. This indicates that the house prices is a non-stationary series. We then sometimes say that the series has a unit

root. To formally test for stationarity we can use tests such as the augmented Dickey-Fuller test (Dickey & Fuller, 1981) or the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (Kwiatkowski et al., 1992).

If a series is found to be non-stationary, we can difference the series to make it stationary. If, after take the first difference, the series becomes stationary it is said to be integrated of order one, often denoted as I(1). In general, the order of integration d is the number of differences needed to make the series stationary. In practice, the order of integration is usually either one or two.

This means that an ARIMA(p,d,q) model fitted to a stationary series is simply an ARMA(p,q) model as defined by (2.5) and (2.6). Continuing in the notation of Hyndman et al. (2008), a non-seasonal ARIMA(p,d,q) can be expressed as

$$\theta(L)(1-L)^d y_t = \lambda + \phi(L)\varepsilon_t . \quad (2.8)$$

Note that $Ly_t = y_{t-1} \Rightarrow (1-L)y_t = y_t - y_{t-1}$. A seasonal ARIMA model can be denoted by ARIMA(p,d,q)(P,D,Q)[m] and is expressed as

$$\Theta(L^m)\theta(L)(1-L^m)^D(1-L)^d y_t = \lambda + \Phi(L^m)\phi(L)\varepsilon_t \quad (2.9)$$

where $\Theta(z)$ is the seasonal autoregressive polynomial of order P and $\Phi(z)$ is the seasonal moving average polynomial of order Q, D is the number of seasonal differences and m is the length of seasonality (number of months in a year etc.). To formally test for seasonal integration we can use a Canova-Hansen test (Canova & Hansen, 1995) or a Osborn-Chui-Smith-Birchenhall (OCSB) test (Osborn et al., 1988).

2.1.2 Exponential smoothing and state space models

Another common class of forecasting models are exponential smoothing methods. They have been among the workhorses of forecasters for quite some time (e.g. Koehler et al., 2001; Kolassa, 2011; Makridakis & Hibon, 1991). These methods have their origin in works such as R. G. Brown (1963), Holt (1957, reprinted 2004) and Winters (1960). The broad idea of forecasting using exponential smoothing is to assign exponentially decreasing weights to observations as they go back in time.

Pegels (1969) was the first to classify exponential smoothing methods and provided a taxonomy of the trend and seasonal component based on whether they are additive or multiplicative. Pegels' taxonomy was then extended by Gardner (1985), who expanded the classification to include damped trends. This extension was later modified by Hyndman et al. (2002), before Taylor (2003) made a final extension by introducing damped multiplicative trends. This gives the fifteen exponential smoothing methods seen in Table 2.1.

Table 2.1: Exponential smoothing methods.

| Trend component | | Seasonal component | | |
|-----------------|-------------------------|--------------------|-------------------|-----------------------|
| | | N (None) | A (Additive) | M (Multiplicative) |
| N | (None) | N,N | N,A | N,M |
| A | (Additive) | A,N | A,A | A,M |
| A _d | (Additive damped) | A _d ,N | A _d ,A | A _d ,M |
| M | (Multiplicative) | M,N | M,A | M,M |
| M _d | (Multiplicative damped) | M _d ,N | M _d ,A | M _d ,M |

Source: Hyndman and Khandakar (2008)

Some of the methods go under other names e.g. (N,N) is the simple exponential smoothing (SES) method, (A,N) is Holt's linear method while (A,A) and (A,M) are the additive and multiplicative Holt-Winters' methods, respectively (Hyndman & Khandakar, 2008).

To illustrate for an observed time series given by y_t , following the notation in Hyndman et al. (2008), the additive Holt-Winters method (A,A) can be expressed by the following system

$$\text{Level:} \quad l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \quad (2.10)$$

$$\text{Growth:} \quad b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \quad (2.11)$$

$$\text{Seasonal:} \quad s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \quad (2.12)$$

$$\text{Point forecast:} \quad \hat{y}_{t+h|t} = l_t + b_t h + s_{t-m+h_m^+} \quad (2.13)$$

where m is the length of seasonality, l_t is the level, b_t is growth, s_t denotes the seasonal component, $\hat{y}_{t+h|t}$ is the point forecast for h periods ahead and $h_m^+ = [(h-1) \bmod m] + 1$. The smoothing parameters, α , β^* and γ , indicates how fast the level, trend and seasonality, respectively, adapt to new information. Recursive formulae for all fifteen exponential smoothing methods can be found in Hyndman et al. (2008) pp. 18.

Building on the exponential smoothing literature, Snyder (1985) showed that the SES had an underlying state space model. State space representation provides a framework in which any linear time series model can be presented (De Gooijer & Hyndman, 2006).

Hyndman et al. (2002), Ord et al. (1997) and Taylor (2003) extended this and showed that there are two underlying innovation state space model, one with additive errors and one with multiplicative errors, for each of the fifteen models in Table 2.1. Thus, the present classification include 30 potential models. The underlying innovation state space model, e.g. with additive errors, corresponding to the additive Holt-Winters method is

$$l_t = l_{t-1} + b_{t-1} + \alpha \varepsilon_t \quad (2.14)$$

$$b_t = b_{t-1} + \beta \varepsilon_t \quad (2.15)$$

$$s_t = s_{t-m} + \gamma \varepsilon_t \quad (2.16)$$

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t \quad (2.17)$$

where we assume $\varepsilon_t \sim \text{IID}(0, \sigma^2)$ and $\beta = \alpha \beta^*$.

The model (2.14)-(2.17) with an additive error, trend and seasonality can be indicated by ETS(A,A,A). The acronym ETS refers to the error (E), trend (T) and seasonality (S) component. Equations for all 30 innovation state space models is found in Hyndman et al. (2008, pp. 21-22)

Since the distinction between additive and multiplicative errors makes no difference to point forecasts, it has largely been ignored historically. However, when producing prediction intervals the nature of the error component matters (Hyndman & Khandakar, 2008). This distinction does play a role in this thesis, as prediction intervals are reported for all models. The intervals are included as they are an intuitive way of indicating model uncertainty.

Detailed derivations for the prediction intervals of all 30 state space models can be found in Hyndman et al. (2005).

2.1.3 Other models

A plentitude of other time series models exists. In addition to the univariate models discussed so far, the ARCH model (Engle, 1982) and all of its extensions is often applied in forecasting of financial time series, as it accounts for the heteroscedastic (time varying variance) nature of financial data.

As for multivariate models, the vector autoregressive (VAR) model and its cointegrated counterpart the error correction model (ECM) are also common (e.g. Artis and Zhang (1990) and Simkins (1995)). Forecasters also use non-linear models such as regime switching models (Crawford & Fratantoni, 2003) and neural networks (Zhang et al., 1998). For a comprehensive review of time series forecasting literature see e.g. De Gooijer and Hyndman (2006).

2.1.4 Model selection

After the forecaster has chosen which models to use, they need to specify them. For the for the AR and ARIMA models this entails choosing appropriate values for the model orders p , d , q , P , D and Q . For the exponential smoothing class of models this means choosing among the 30 model specifications in the classification.

One common approach to model selection is based on information criteria (IC). Akaike (1973, 1974, 1981) was at the forefront of this research. The Akaike Information Criterion (AIC) is given by

$$AIC = -2\ln(L) + 2k \quad (2.18)$$

where L is the numeric value of the maximum likelihood for the candidate model and k is the number of estimated parameters (normally the model coefficients plus the sample variance).

Akaike (1974) (see e.g. Burnham and Anderson (2002) for a less complex treatment) showed that the AIC can be viewed as the expected Kullback-Leibler (KL) distance (Kullback & Leibler, 1951). The KL information can be interpreted as the distance between a proposed model and the unknown “true” data generating process (DGP) (Kolassa, 2011).

When we are modelling time series, we would naturally like to find models that is as close to the unknown DGP as possible. As such, when the AIC is used for model selection, we choose the model that minimizes the AIC and by extension, the expected distance from the unknown DGP.

There are many variations of the AIC. One of these is the corrected AIC (Hurvich & Tsai, 1989),

$$AIC_c = -2\ln(L) + \frac{2kn}{n-k-1} = AIC + \frac{2k(k+1)}{n-k-1} \quad (2.19)$$

where n is the number of observations in the series. As we see the AIC_c includes a penalty for small sizes. Burnham and Anderson (2002) suggest using the AIC_c when $n/k < 40$.

A similar measure to the AIC is the Bayesian Information Criterion (BIC) (Schwarz, 1978) and is given by

$$BIC = -2\ln(L) + k \ln(n) \quad (2.20)$$

where n is the number of observations in the series.

For AR and ARIMA models, a more informal (albeit integral part of the Box-Jenkins) model selection procedure is to examine the autocorrelation function (ACF) and partial autocorrelation (PACF) plots. These plots is often hard to interpret, and can lead to substantial selection bias if used on their own. However, ACF and PACF plots still provide helpful insight, especially in a complimentary role to more formal selection methods.

Other model selection approaches include F-tests and likelihood ratio tests. These tests are losing ground to the IC approach to model selection, as they have been found less robust and generally inferior (Kletting & Glatting, 2009; Lin & Dayton, 1997; Ludden et al., 1994; Posada & Buckley, 2004).

2.1.5 Combining forecasts

The main idea behind this part of the forecasting literature is to examine whether combining forecasts can improve the forecast accuracy. This approach can be seen as a complimentary approach to selecting the best model based on criteria discussed in section 2.1.4.

Evidence has been found that a combination of forecasts indeed outperforms the single “best” forecast (e.g. Bjørnland et al., 2012; Clemen, 1989; Stock & Watson, 1998, 2003, 2004). Part of the reason for this may be that the generating process behind an observed time series will never be simple with a known functional form, since seasonality may change and trends come and go (Box & Draper, 1987).

With their seminal paper, Bates and Granger (1969) paved the way for further developments in the literature on the combination of forecasts. They suggested deriving weights from inverse out-of-sample mean square errors (MSEs)

$$w_i = \frac{\frac{1}{MSE_i}}{\sum_{j=1}^n \left(\frac{1}{MSE_j} \right)} . \quad (2.21)$$

where MSE_i defined below by equation (2.24). For an application of this weighting scheme see e.g. Bjørnland et al. (2012).

Other more naïve combination methods include simply weighting forecasts by taking their mean or median. Equal weight forecast combinations have in some instances been found to outperform more intricate schemes (Stock & Watson, 2003).

A more recent development in the forecast combination literature is weighting schemes based on information criteria (see e.g. Kapetanios et al., 2008a; Kapetanios et al., 2008b; Kolassa, 2011). Burnham and Anderson (2002) was among the first to generalize the IC approach to forecast combination.

Using the AIC as an example, the weights can be calculated as follows. First, we find the differences in AIC between each model and the AIC-minimizing model. Considering a finite set of models \mathbb{S} . The AIC of a model $M \in \mathbb{S}$ is denoted by $AIC(M)$. The “delta”-AIC for model M is then given by

$$\Delta_{AIC}(M) = AIC(M) - \min_{N \in \mathbb{S}} AIC(N) . \quad (2.22)$$

As a result, the AIC minimizing model will have a delta-AIC of zero. We do this transformation to ensure we are only dealing with positive values and as only the AIC values relative to each other matter, not their values isolated; it does not change their interpretation.

Furthermore, we calculate $\exp\left(-\frac{1}{2}\Delta_{AIC}(M)\right)$ for each model in \mathbb{S} . This transformation is a result from Akaike (1981) and can be interpreted as the relative likelihood between each model and the AIC-minimizing model, conditional on the data (Kolassa, 2011). The Akaike weights is just the relative likelihoods normalized and can be defined as

$$w_{AIC}(M) = \frac{\exp\left(-\frac{1}{2}\Delta_{AIC}(M)\right)}{\sum_{N \in \mathbb{S}} \exp\left(-\frac{1}{2}\Delta_{AIC}(N)\right)} \quad (2.23)$$

where $w_{AIC}(M)$ is the Akaike weight for model M . An identical derivation of weights can be done using other ICs. For extensive surveys on forecast combination see Clemen (1989) and Timmermann (2006).

2.1.6 Forecast evaluation

After we have chosen a model and specified it, we can use it to compute forecasts. Accuracy measures can then be used to evaluate the models' forecast performances. Some of the more common ones include the MSE, mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute scaled error (MASE) to name a few. The forecast error is simply defined as $e_t = y_t - \hat{y}_t$, where y_t is the observation at time t in a series and \hat{y}_t is the forecast of y_t . The MSE, MAE and RMSE is scale dependent measures and consequently, should not be used when comparing models estimated from different data sets. They are defined as

$$MSE = \frac{\sum_{t=1}^N e_t^2}{N} \quad (2.24)$$

$$MAE = \frac{\sum_{t=1}^N |e_t|}{N} \quad (2.25)$$

$$RMSE = \sqrt{MSE} \quad (2.26)$$

where N is the number of observations in the sample.

We can also define a percentage forecast error as $p_t = 100e_t / y_t$. The MAPE is a measure based on percentage error and is given by

$$MAPE = \frac{\sum_{t=1}^N |p_t|}{N} \quad (2.27)$$

where $|p_t|$ is the absolute value of the percentage error.

Measures based on percentage error have the advantage of being scale independent, but are not reliable for series that are changing signs, as they are undefined/infinite for $y_t = 0$.

A third error type, the scaled error, can be defined as

$$q_t = \left(\frac{e_t}{\frac{1}{N-1} \sum_{i=2}^N |y_i - y_{i-1}|} \right) \quad (2.28)$$

where the denominator is the in-sample MAE of a naïve forecast method, i.e. a method where the forecast for the observation time t is simply the observation at $t-1$. The distinction between in-sample and out-of-sample accuracy measures is discussed in section 3.4.

Measures based on scaled errors have the advantage of being neither scale dependent nor implicated by series floating around zero. The MASE is one such measure and is given by

$$MASE = \frac{\sum_{t=1}^N |q_t|}{N} \quad (2.29)$$

In addition to being very flexible, the MASE also has a simple interpretation. If the value is less than one it means the forecast is more accurate than the naïve forecast. The MASE was proposed by Hyndman and Koehler (2006). They also provide a useful guide through the jungle of forecast accuracy measures.

2.2 Forecasting house prices

Before Most forecasting studies employs several model classes in an attempt to find the best fit. The literature on forecasting house prices is no different, although the majority of studies seem to focus on multivariate forecasting methods. Researchers have long tried to identify the underlying factors and use them to model house prices. Some common factors include inflation, income, housing starts and demographic forces. Hedonic regressions is likely the most frequently used approach to modelling house prices (e.g. Clapp & Giaccotto, 2002; Meese & Wallace, 1997). Others implement frameworks specifically designed to include cointegrated variables. E.g. Zhou (1997) finds that sales volume and prices are cointegrated

and applies an ECM to forecast both. Gupta and Miller (2012) also use VAR models and ECMs to forecast prices in Southern California. Although research have been limited, artificial neural networks are also starting to make their way into the field of house prices prediction showing great potential (Limsombunchai, 2004; Wilson et al., 2002).

Regardless, to assess the predictability of house prices, it is pivotal that we first establish the efficient nature of the housing market. Case and Shiller (1989) found that there is inertia in house prices over short-to-intermediate periods. This implies housing market is weak-form inefficient as defined by Fama (1970). Others support this view of weak-form inefficient house prices, such as Gu (2002) and Kuo (1996) both of whom find that house prices in fact can be predicted over short timeframes. Case and Shiller (1990) tested the semi-strong form of the efficient market hypothesis. They found that house prices did not impound public information quickly (e.g. expected income and inflation). Thus, the consensus in the literature seem to be that house prices indeed is weak-form inefficient and that they are “forecastable” at least over shorter timeframes.

Univariate time series models have been found to forecast very well over shorter periods (Crawford & Fratantoni, 2003). Especially ARIMA models have received a lot of attention from forecasters of house prices. Chin and Fan (2005) compare three different ARIMA models in an application on residential prices in Singapore. They find that an ARIMA with dummy variables perform better than an ARIMA with ARCH errors, but only marginally better than the original model. Hepsen and Vatansever (2010) use a standard Box-Jenkins ARIMA approach to forecasting house price trends in Dubai. Tse (1997) examines forecasting of real estate prices in Hong Kong in a similar framework. He finds that the ARIMA model indeed is able to indicate short-term market direction. ARIMA models also do well when compared to other model classes. Crawford and Fratantoni (2003) compares simple ARIMA models to GARCH and regime-switching models. They find that simple ARIMA models generally perform better when comparing out-of-sample forecast accuracy, while the regime-switching model perform better in-sample. Nevertheless, not everyone is as enthusiastic about the forecasting ability of ARIMA models. Stevenson (2007) warns that although ARIMA models are useful in predicting broad market trends, they differ substantial in their forecasts obtained from different model specifications. That is, they are sensitive to model selection biases.

Furthermore, Sklarz et al. (1987) showed that a long lagged AR process produce lower forecast error variance than an ARIMA model in an application to U.S. housing starts data. Although not forecasting house prices explicitly, they point out that the long lagged AR model is well equipped to forecast the housing market in general, which features strong seasonality and slowly changing trends.

Unlike ARIMA models, exponential smoothing methods have received little attention from forecasters of house prices. Nonetheless, Birch and Sunderman (2003) introduce a two-way exponential smoothing system for estimating true movements in residential property prices. Their method appear to still be in its infancy and seem somewhat experimental. They point out that their system seems to overcome some of the problems attached to the more rigid nature of regression modelling. However, they do not offer any conclusive evidence that their model is superior to more common hedonic price models.

Another recent addition to the literature on forecasting house price is the introduction of models that allow for time varying parameters. J. P. Brown et al. (1997) mention that it is limiting to assume the underlying DGP is stable and applying constant parameter methods suffer in terms of parameter instabilities. Guirguis et al. (2005) support this view and find that allowing coefficients to vary over time improve the precision of house price forecasts.

3 Methodology

Based on the literature review, three models are selected. The three are one AR model, one ARIMA model and one exponential smoothing state space model. As is quite common in the forecasting literature, the AR model is include in a benchmark capacity. That is, it is not expected to perform better than the other two, as it is a lot less flexible. However, the AR process work well as a benchmark to assess if it is worth the added effort of using models that are more complex and tedious in their application.

At first glance, it may seem restrictive to limit the analysis to three models. However, considering all the different specifications, the actual number of models considered is quite large. There are 30 exponential smoothing state space models and a similar number of AR models considered. The number of ARIMA models to choose from can be anywhere from a few hundred to several thousand, depending on the range in which we allow the model orders to vary.

It is the developments in computing power and improved automatic modelling algorithms that has allowed us more freedom in the model selection process. In my analysis, all three of the main models are estimated using the *forecast* package (Hyndman, 2009) developed for the R programming language. The package contains fully automated forecasting algorithms for both ARIMA and exponential smoothing state space models. The algorithms are outlined in Hyndman and Khandakar (2008).

Next, the estimation procedure for each model is outlined, as well as some diagnostics tools for the residuals. Lastly, the procedures for calculating accuracy measures is presented.

3.1 Model estimation

Before estimating any models or computing forecasts, we should do some data exploration (see chapter 0). This is to identify any anomalies in the series, which may affect the forecasting procedure further down the line. After ensuring the data is appropriate, we can specify the models.

In the AR model, both constant and deterministic trend terms are included. The lag order p is selected by searching through lags starting with $p = 1$. The p that minimizes the AIC_c ,

defined by equation (2.19), is chosen. To ensure a consistent analysis, the AIC_c is used for model selection for all three models. This particular IC was picked as we are working with a relatively small sample. The rule of thumb mentioned in section 2.1.4 said that we should use the AIC_c over the AIC when $n/k < 40$. Here the n/k -ratio is less than 40 for all three models by a wide margin. As the AIC_c converges to the AIC for large samples, there really is no reason not to use the AIC_c other than the slightly more tedious computation.

For the ARIMA model, the selection procedure is a bit more intricate. We denoted the seasonal model as $ARIMA(p,d,q)(P,D,Q)[m]$. First, we specify the seasonal frequency m and the seasonal differencing ($D=1$ or $D=0$) with a OCSB test. After the seasonal dummy D is chosen, we find the order of integration d with successive the KPSS test. This is a unit root test with null hypothesis of no unit root. After establishing the values of d and D , we can specify the remaining model orders p, q, P and Q . We do this again by minimizing the AIC_c . This ARIMA procedure is fully automated in the *forecast* package. The algorithm is described in Hyndman and Khandakar (2008).

To apply the exponential smoothing state space models we need the values for the initial states $(l_0, b_0, s_0, s_{-1}, \dots, s_{-m+1})$ and the smoothing parameters $(\alpha, \beta, \gamma, \phi)$, both of which are estimated from the data. The procedure for forecasting with these models is as follows. First, apply all the models to the series using the optimized parameters (read: both the initial states and the smoothing parameters) obtained via maximum likelihood estimation (MLE). Then we select the model that minimizes the AIC_c .

The parameters for all three models are estimated using MLE. The general idea behind MLE is to find estimates of the parameters that maximizes the “likelihood” of the sample (Larsen & Marx, 2012). The likelihood is defined by a likelihood function $L(\bullet)$. To illustrate, for a random sample y_1, y_2, \dots, y_n of n observations, the likelihood function is

$$L(\delta) = \prod_{i=1}^n p(y_i; \delta) \quad (2.30)$$

where δ is an unknown parameter and $p(y_i; \delta)$ is the discrete probability density function (pdf). That is, the likelihood function is the product of the pdfs evaluated at all the n y_i 's. If δ_e is the value of the parameter such that $L(\delta_e) \geq L(\delta)$ for all values of δ , then we call δ_e

the maximum likelihood estimate of the unknown parameter. MLE can be quite complicated in its application and especially intricate for ARIMA models. I consider detailed derivations of the MLE procedures for the models to be beyond the scope of this thesis.

Prediction intervals for models with normal forecast distributions, i.e. models with neither multiplicative error, trend nor seasonality, can be calculated as

$$y_h \pm z_{\alpha/2} \sqrt{v_h} \quad (2.31)$$

where y_h is the point forecast h periods ahead, z_q is the q th quantile of the standard normal distribution, α is the level of significance and v_h is the forecast variance of the model. Even for models with non-normal forecast distributions is the prediction intervals generated by (2.31) be reasonably accurate (Hyndman et al., 2005).

3.2 Forecast combination

In an effort to improve the final forecast's performance, the three models are combined. Since the individual models are specified by the AIC_C, it would make sense to continue the information theoretic approach by also using a weighting scheme base on ICs. However, the conditional likelihoods from the ARIMA- and ETS-models are not comparable as their initializations are different. That is, the likelihoods from the ARIMA-models are conditional on the first few values of the time series, while the likelihoods of the ETS-models are conditional on the initial states.

Instead, the Bates and Granger weighting scheme based on inverse out-of-sample MSEs is implemented. The application of the scheme is straightforward. First, we compute point forecasts from each of the three models. Then these are combined with the weights defined by equation (2.21). The formula for the MSE is given by equation (2.24). In section 3.4, the computation of the out-of-sample MSEs is detailed.

3.3 Diagnostics

In section 2.1.1 and 2.1.2, it was mentioned that we assume the errors in the models to be independent and identically distributed (iid). When using the models for forecasting purposes we sometimes call the error terms residuals. In time series forecasting the residuals are simply $e_t = y_t - \hat{y}_t$ where \hat{y}_t is the one step forecast of y_t . It is important to inspect the residuals to

diagnose if they indeed are iid. The most important aspect of residual diagnostics is to ensure that they are uncorrelated and have a mean of zero. If the residuals are correlated, it means there is still information left that the current model is not able to sufficiently capture. A non-zero mean implies that there is a bias in the forecast. To detect residual correlation we can study the ACF and PACF plots. If there are significant correlations, we may want to consider other potentially more appropriate models. Correlation between residuals also affect the weighting scheme outlined above. Bjørnland and Thorsrud (2012) points out that these type of weights will minimize a quadratic loss function, but only if the residuals from the individual models are uncorrelated. The Ljung and Box (1978) test can check for autocorrelation in the residuals and is given by

$$Q = N(N + 2) \sum_{k=1}^m \frac{r_k^2}{N - k} \quad (2.32)$$

where N is sample size and r_k is sample autocorrelation at lag k . $H_0 : r_1, r_2, \dots, r_m = 0$ and $H_1 : r_1, r_2, \dots, r_m \neq 0$. We reject if $Q \geq \chi_{1-\alpha, m-p}^2$ where $\chi_{1-\alpha, m-p}^2$ is the critical value from the chi-square distribution with significance level α and $m - p$ degrees of freedom, m is the lag order and p is the number of estimated model parameters. To capture sufficiently the autocorrelation effect of seasonal data it is common to include lags up to two time the seasonal frequency. Here this means setting $m = 24$.

We can also investigate whether the residuals are homoscedastic, i.e. if they have constant variance, and if they are normally distributed. These properties do not necessarily have to be fulfilled for the models to generate reliable forecasts, but they are useful as they make the calculation of the prediction intervals easier (Hyndman & Athanasopoulos, 2013). To check for homoscedasticity we can use the Ljung-Box test on squared residuals. The normality can be checked with a Jarque and Bera (1980) test is given by

$$JB = \frac{N}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right) \quad (2.33)$$

where N is sample size, S is sample skewness and K is sample kurtosis. $H_0 : e_t$ are normal and $H_1 : e_t$ are non-normal, where e_t are residuals. We reject if $JB \geq \chi_{1-\alpha, 2}^2$. This test simply checks jointly if the residuals have the same skewness and kurtosis as the normal distribution. Results of all the tests described in this section is detailed in section 5.2.

3.4 Accuracy

After diagnosing the residuals, we can finally use the models to compute forecasts. their forecast precision. Different accuracy measures can yield completely opposite conclusions. For that reason, only two measures are employed to assess the performance of the forecasts, namely RMSE and MASE as defined by equation (2.26) and (2.29) respectively.

There are two ways of calculating the forecast measures, in-sample and out-of-sample. The former method assesses the fitted values generated by the model with the actual values over the whole sample. The out-of-sample method entails comparing the forecasts from a model with a holdout sample, which I will call the test set.

There are several ways of obtaining the out-of-sample forecast errors. Here a technique called cross validation is applied. The procedure, as outlined by Hyndman and Athanasopoulos (2013), goes as follows: First, use the observations at time $k + h + i - 1$ for the test set, where k is some subjectively chosen point in the data set, h is the forecast horizon and i can be interpreted as the iteration variable. Then we use the observations at times $1, 2, \dots, k + i - 1$ to generate forecasts. Based on the obtained forecast errors, we can calculate accuracy measures. Then we repeat this for $i = 1, 2, \dots, N - k - h + 1$ where N is the total number of observations. After the iterative process is done, we calculate the average of the forecast accuracy measures based on the errors collected at each iteration. Thus, the average of the measures is what I call the out-of-sample accuracy.

The intuition behind this procedure is that the goodness-of-fit measures will indicate how well the models were able to forecast the next h periods, when comparing with the actual values in the test sets. The length s of the test set is subjective. A rule of thumb says that the holdout period should be approximately 20% of the original sample. Here, I have decided on holding out the last $s = 24$ months (2011:4-2013:3) in the original sample of $N = 135$ observations (2002:1-2013:3). Leaving $N - s = 111$ observations (2002:1-2011:3) that is used as a training set to fit the models. The forecasts are then generated for $h = 12$ months rolling forward in the test set from 2011:4 to 2013:3. The 12-month forecast horizon and the 24-month test set means the procedure consists of 12 iterative steps from which we calculate forecast errors and accuracy measures.

A model can have a good in-sample fit, but it does not mean it will forecast well out-of-sample. Therefore, the convention has been to use out-of-sample measures, when assessing forecast ability of models, although some have a contrary view (Inoue & Kilian, 2005). Consequently, both in-sample and out-of-sample RMSE and MASE are reported in the results for all three models, as well as for the combined forecasts. Although both are reported, the out-of-sample RMSE and MASE are used as the main forecast evaluation measures. The out-of-sample MSEs obtained from the cross validation procedure is also used in the weighting scheme.

4 Data

Here a presentation of the data and its properties is given. The first part describes the raw data and the data gathering process. The second part is an exploration of the statistical properties of the house price series.

4.1 *The data set*

The data set used in this thesis is based on price statistics for residential housing in Norway. Pöyry Management Consulting (Norway) (from now on referred to as Pöyry) develops the statistics on behalf of Eiendomsmeglerforetakenes Forening (Eff) and Norges Eiendomsmeglerforbund (NEF). It is updated monthly and can be found at NEF's website (under NEF in references). Pöyry reports prices in NOK1000 per square meter (m²).

The numbers are based on properties sold through the online market place Finn.no. Thus, it is not a complete statistic of house prices in Norway, but it is the consensus source for reliable house price data in Norway. Statistics Norway (SSB) uses the same raw data as Pöyry. However, SSB's house price index deviates some from the statistics reported by Pöyry in their transformation and weighting schemes (Christensen, 2003).

The statistics identifies three types of residential housing: Apartments (Norwegian: leiligheter), shared houses (delte boliger) and standalone houses (eneboliger). Pöyry gathers prices for each type regionally. To calculate the national average for each type, they weight the prices by the share of houses sold in the candidate region in the previous three years. Furthermore, to calculate the national average for all types of housing, Pöyry weight the prices for each type with their share of total sales the last three years.

Likely, in an attempt to prevent outliers from influencing the statistics, Pöyry has excluded some special cases from the data. These are any type of house exceeding 500 m², standalone and shared houses smaller than 50 m², apartments smaller than 20 m², and any type of house either costing less than NOK2000 pr. m² or exceeding NOK100 000 pr. m².

The price statistic reaches all the way back to 1985. However, before 2002 prices were only reported quarterly while prices are reported monthly from 2002 onward. In addition, Pöyry changed their data gathering methods in 2002. This resulted in a large increase in the reported

number of sold properties after 2002. These changes may result in significant inference problems if we were to use data both before and after 2002. For that reason, only data from January 2002 and onward is included in the data set, which is updated to include prices up to March 2013.

The properties, transformations and weighting schemes described here is detailed in a report published monthly at Eff's website (listed under Pöyry in references). The final data set contains a time series of 135 observations (months) of house prices for all types of residential houses from 2002:1 to 2013:3.

4.2 Data exploration

Before estimating the models, it is of interest to explore the data. This is to prevent any unforeseen outliers from affecting the analysis and for identifying the statistical properties of the data. Table 4.1 gives some descriptive statistics for the house prices series.

Table 4.1: Descriptive statistics for nominal house prices 2002:1-2013:3.

| | N | Mean | Median | St.dev. | Variance | 1stQu. | 3rdQu. | Min. | Max |
|------------|-----|--------|--------|---------|----------|--------|--------|--------|--------|
| Prices (P) | 135 | 22.057 | 22.852 | 4.9217 | 24.223 | 17.43 | 25.43 | 14.373 | 31.454 |

It was mentioned in section 2.1.1 that a visual inspection of the plotted data in Figure 1.1 indicates that the house prices are non-stationary. An augmented Dickey-Fuller (ADF) test can test for this formally. The general ADF test equation is

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{s=1}^m \delta_s \Delta y_{t-s} + v_t \quad (3.1)$$

with $H_0 : \gamma = 0$ and $H_1 : \gamma < 0$. The lag order m was set to five. The DF test statistic

$\tau = \frac{\hat{\gamma}}{se(\hat{\gamma})}$. If $\tau > \tau_c$ we do not reject the null hypothesis. Some critical τ_c values can be

found in Hill et al. (2008, p. 486). The ADF test statistic, tau ($=-2.4648$), and the p-value ($=0.3831$) support our suspicion, as we cannot reject the null hypothesis of non-stationarity.

The first differenced series is plotted in Figure 4.1.

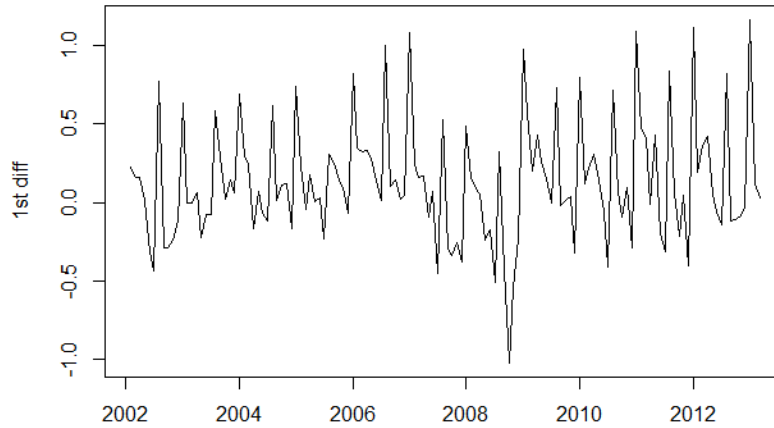


Figure 4.1: First differenced house prices 2002:1-2013:3.

Admittedly, just by “eyeballing” the first difference plot, it is somewhat unclear whether the series is stationary. However, a second ADF test on the first differences indicates the series is in fact integrated of order one ($\tau = -4.5455$, $p\text{-value} = 0.01$). There is a sharp drop around mid-to-late 2008. To prevent any outliers (that does not belong) from affecting the analysis, we should do some data exploration. Upon examining the data further, one finds that there was a drop in prices from September to October 2008 of about -4.5%. This was the largest month-to-month price change in the sample period. Although it may be considered an outlier, nothing indicates it should not be there. The timing of the sharp drop is to be expected as this was during the volatile financial crisis.

Moreover, it was mentioned in section 2.1 that an integral part of time series analysis is the decomposition of the series. Before we estimate the models, it might be interesting to explore the different components of the data a little further. Figure 4.2 exhibits a graphical representation of the decomposed series of house prices. The top panel (*data*) is simply the plotted series equivalent to Figure 1.1. The second panel (*seasonal*) gives us an indication of the seasonality of house prices. There seem to be a repeating pattern every 12 months. This tells us that house prices indeed inherit some seasonal component. More specifically, the decomposition indicate that house prices are at its highest during the early months of the year, while the prices flatten by fall before they start rising again late in the year. The third panel (*trend*) exhibits the trend of the series. It is equal to the original series, but with the seasonality extracted. As we see, there is a distinct upward sloping trend, only interrupted by the financial turmoil of 2008 and 2009. If the models are appropriate for the house price, they should be able to capture the seasonality and trend of the series. The last panel (*remainder*) is

the residuals from the seasonal and trend fits. The large negative spike in the residuals lines up with the timing of the sharp price change drop discussed above.

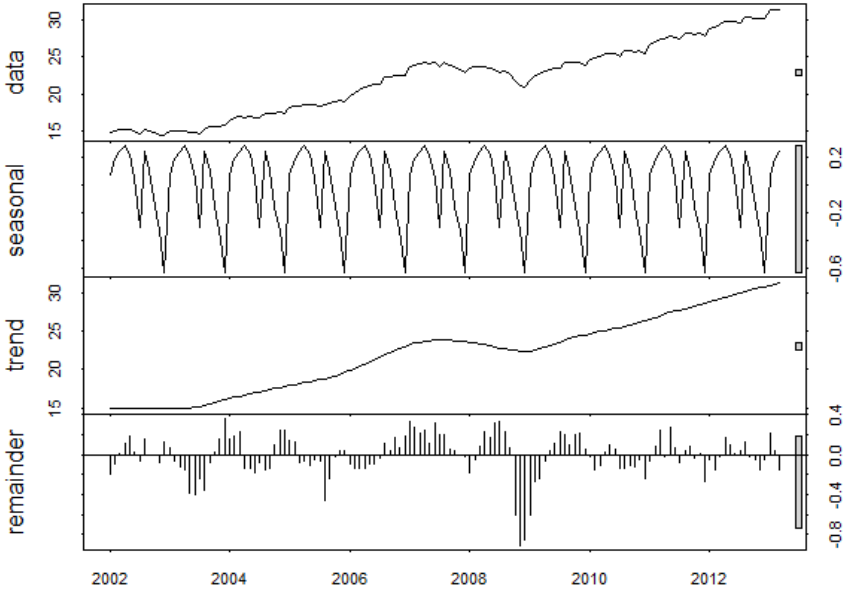


Figure 4.2: *Decomposition of the house price series.*

Lastly, it should be mentioned that it is quite common to perform some kind of data transformation before forecasting. Some popular transformations include taking logarithms and power transformation, such as squaring the series or Box and Cox (1964) transformation. The reason for transforming data is often stabilize variance and to make it more normal to avoid violating a multitude of assumptions when using parametric statistics. In this context, the three models applied here are quite flexible. ARIMA models do assume homoscedasticity (constant variance), but does not require the data to be stationary before estimation. The AR model is just a special case of the ARIMA so it is understood that the same assumptions are made for both. Exponential smoothing state space models are very flexible in that they can be both linear and nonlinear, are non-stationary and robust to heteroscedasticity. In addition, all the model parameters are estimated with MLE, a more robust (although more complex) estimation procedure compared to e.g. OLS. Based on this, I have not found it necessary to perform any transformations of the data. Hence, the forecasts are computed for nominal house prices in levels.

5 Results

In this chapter, all results are presented. The majority of calculations are conducted in the R studio (version 0.97.311) programming environment. Some data manipulations are performed in Microsoft Excel 2013. All three models are estimated using the *forecast* package (Hyndman, 2009) developed for the R programming language (version 4.0). To ensure the replicability of the results, the R code in its entirety is included in appendix D.

Based on the methodology outlined in chapter 3, the following three models are selected: an AR(14) with constant and trend, an ARIMA(2,0,1)(0,1,1)[12] with drift and an ETS(A,A,A). From now on, they are referred to by their acronyms AR, ARIMA and ETS. A few interesting notes about the selected models: The AR-model has a long lag, especially compared to the sample size (135 observations). However, the high model order was necessary to capture the seasonality in the house prices. AR-models with lags shorter than 13 showed significant positive autocorrelation in their residuals every 12th lag, clearly indicating they were not able to capture the seasonal component in the data.

Furthermore, during the data exploration we ran ADF tests that indicated the series was integrated of order one, yet in the ARIMA- model d is selected to be zero. The reason can be inferred from the methodology. There we said that the seasonal order of integration D was specified first, then we used KPSS tests to select d . Since $D = 1$ and $d = 0$ we can derive that after seasonally differencing, the series becomes difference stationary as well. Thus, we cannot reject the null hypothesis of the KPSS tests of no unit roots.

The ETS with both additive error, trend and seasonality is selected. This means that out of all the 30 model specifications a purely linear model is the one that minimizes the IC. In turn, this may indicate the linear and less flexible AR-model actually fits the series well.

The complete estimated models is found in appendix A. Point forecast and numeric values for prediction intervals is given in appendix B. The main objective of this thesis is to forecast house prices in Norway for the 12 months in the period from April 2013 to March 2014. Next, the results of these forecasts are presented. Then the results of several residual diagnostics test are given, before a brief discussion rounds out the chapter.

5.1 Forecasts

Perhaps the most interesting aspect of the forecasts is to examine how the models handle the trend and seasonality components. Figure 5.1, 5.2 and 5.3 show the forecasts from the AR-, ARIMA and ETS-models respectively. The thin blue line indicates the in-sample fitted values, while the forecasts are shown as the thicker blue line. The prediction intervals are illustrated by the light (95%) and dark (80%) grey bands.

By just looking at the plots, it seems that the three models all capture the seasonality in the house prices. The price forecasts appear to flatten from around August until the end of the year. This is consistent with the decomposition discussed above in section 4.2. The AR-model in Figure 5.1 looks to forecast diminishing growth the next 12 months.

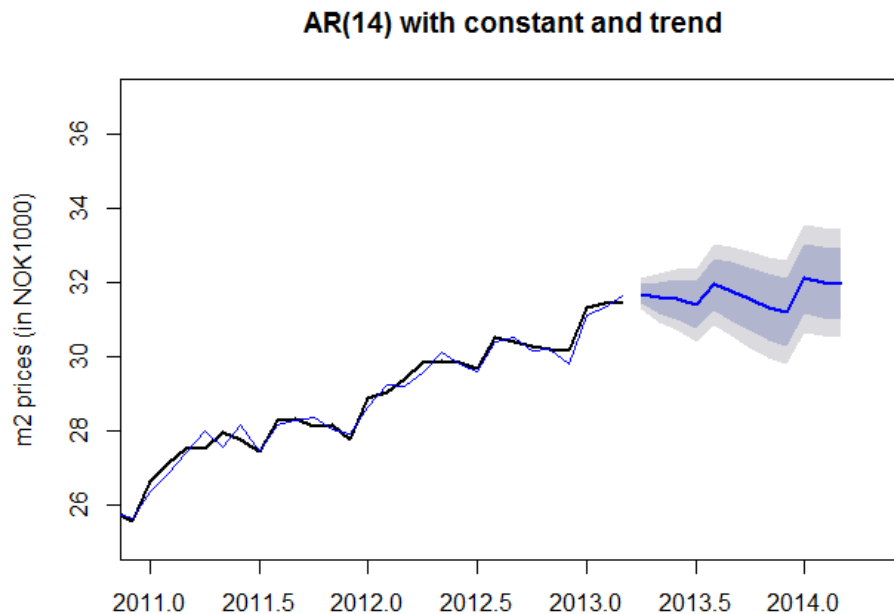


Figure 5.1: Forecast for $h=12$ months ahead with an AR(14) model.

Similarly, the plotted forecast from the ARIMA-model in Figure 5.2 appear to flatten compared to the seemingly linear trend in the house prices in later years. The ARIMA-model appears to differ slightly from the AR-model in that the forecasts at the beginning of 2014 is going upwards compared to the sideways moving forecast of the AR-model.

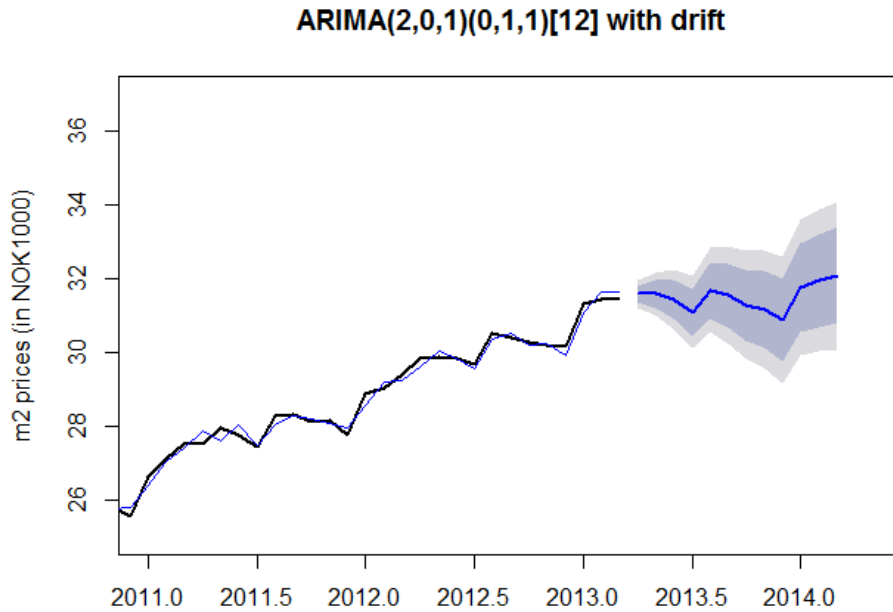


Figure 5.2: Forecast for $h=12$ months ahead with an $ARIMA(2,0,1)[0,1,1][12]$ model.

Interestingly, the ETS-model has the lowest mean square error in the model set by a wide margin. The ETS has an out-of-sample MSE of 0.11 compared to 0.19 and 0.26 for the AR- and ARIMA-models respectively. Thus, the highest weight (0.50) is assigned to the ETS-model. Weights of 0.29 and 0.21 is assigned to the AR- and ARIMA-model, respectively.

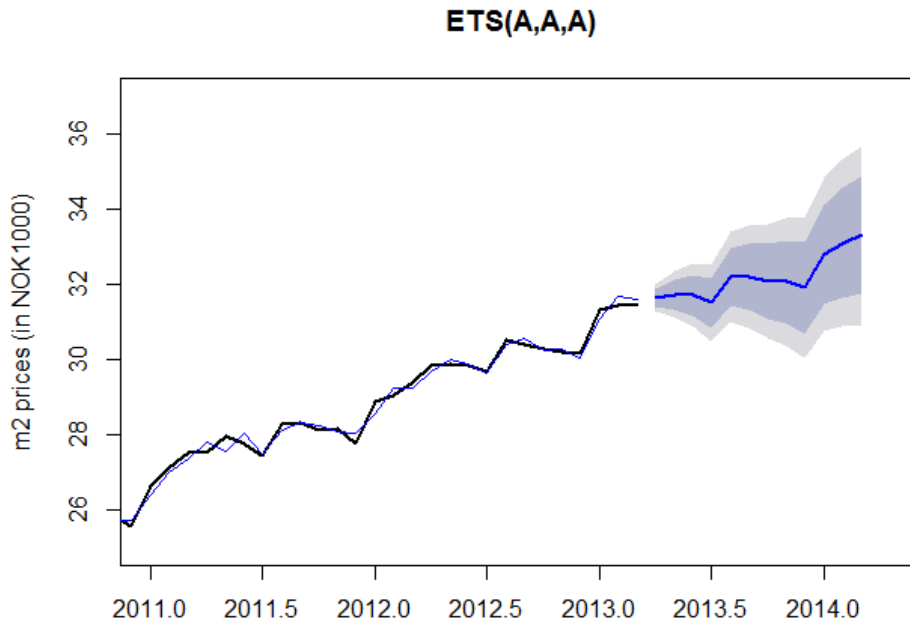


Figure 5.3: Forecast for $h=12$ months ahead with an $ETS(A,A,A)$ model.

Figure 5.4 exhibits the weights, the weighted forecast, as well as 80% and 95% prediction intervals. By diving into the numbers, we can draw a more detailed picture of the models.

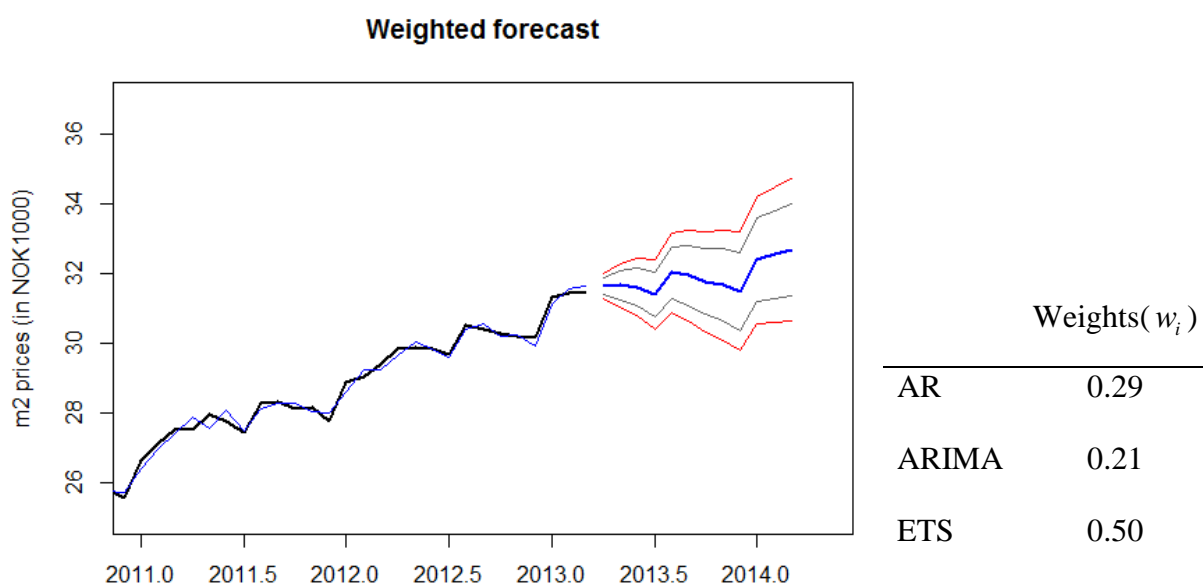


Figure 5.4: Weighted forecast (blue line), 80% (grey) and 95% (red) prediction intervals.

Table 5.1 compares the point forecast with the actual house prices in the corresponding months the previous year. The numbers show that the AR- and ARIMA-models both are forecasting a steady decrease in price growth the next year. Both models start with a forecasted growth from April 2012 to 2013 of about 6 %. By the end of the forecast period, their forecasted growth has dropped to 1.7 % and 2.0 % for the AR and ARIMA, respectively. In contrast, the ETS-model forecasts a steady yearly growth fluctuating around 6 % throughout most of the forecast period.

Table 5.1: Forecasted yearly %-price change, month-by-month.

| Month | AR | | | ARIMA | | | ETS | | | Weighted forecast | | |
|-------|-------------------------------|-----------------------------|------------------------|-------------------------------|-----------------------------|------------------------|-------------------------------|-----------------------------|------------------------|-------------------------------|-----------------------------|------------------------|
| | Forecast 2013:4- 2014:3 | Prices 2012:4- 2013:3 | yearly %- change | Forecast 2013:4- 2014:3 | Prices 2012:4- 2013:3 | yearly %- change | Forecast 2013:4- 2014:3 | Prices 2012:4- 2013:3 | yearly %- change | Forecast 2013:4- 2014:3 | Prices 2012:4- 2013:3 | yearly %- change |
| Apr. | 31.690 | 29.830 | 6.24 % | 31.578 | 29.830 | 5.86 % | 31.631 | 29.830 | 6.04 % | 31.637 | 29.830 | 6.06 % |
| May | 31.583 | 29.881 | 5.70 % | 31.593 | 29.881 | 5.73 % | 31.730 | 29.881 | 6.19 % | 31.659 | 29.881 | 5.95 % |
| Jun. | 31.547 | 29.821 | 5.79 % | 31.435 | 29.821 | 5.41 % | 31.715 | 29.821 | 6.35 % | 31.609 | 29.821 | 6.00 % |
| Jul. | 31.389 | 29.675 | 5.78 % | 31.087 | 29.675 | 4.76 % | 31.507 | 29.675 | 6.17 % | 31.387 | 29.675 | 5.77 % |
| Aug. | 31.939 | 30.497 | 4.73 % | 31.692 | 30.497 | 3.92 % | 32.195 | 30.497 | 5.57 % | 32.019 | 30.497 | 4.99 % |
| Sept. | 31.761 | 30.382 | 4.54 % | 31.550 | 30.382 | 3.84 % | 32.199 | 30.382 | 5.98 % | 31.941 | 30.382 | 5.13 % |
| Oct. | 31.545 | 30.275 | 4.19 % | 31.290 | 30.275 | 3.35 % | 32.087 | 30.275 | 5.99 % | 31.768 | 30.275 | 4.93 % |
| Nov. | 31.330 | 30.186 | 3.79 % | 31.153 | 30.186 | 3.20 % | 32.061 | 30.186 | 6.21 % | 31.665 | 30.186 | 4.90 % |
| Dec. | 31.196 | 30.149 | 3.47 % | 30.868 | 30.149 | 2.38 % | 31.908 | 30.149 | 5.83 % | 31.491 | 30.149 | 4.45 % |
| Jan. | 32.101 | 31.311 | 2.52 % | 31.764 | 31.311 | 1.45 % | 32.805 | 31.311 | 4.77 % | 32.390 | 31.311 | 3.45 % |
| Feb. | 32.010 | 31.428 | 1.85 % | 31.945 | 31.428 | 1.65 % | 33.099 | 31.428 | 5.32 % | 32.550 | 31.428 | 3.57 % |
| Mar. | 31.974 | 31.453 | 1.66 % | 32.088 | 31.453 | 2.02 % | 33.315 | 31.453 | 5.92 % | 32.679 | 31.453 | 3.90 % |

Note: Prices in NOK1000.

Consequentially, the weighted model lands somewhere in between forecasting 3.9 % growth in the 12-month forecast period. This is consistent with what we observed during the visual inspection of the plotted forecasts. The accuracy measures for the three models and the weighted forecast are shown in Table 5.2. The ARIMA-, ETS- and weighted models all have similar in-sample fits, while the AR-model is slightly worse.

Table 5.2: Forecast accuracy for all models.

| | | AR | ARIMA | ETS | Weighted |
|----------------------|------|-------|-------|-------|----------|
| <i>In-sample</i> | RMSE | 0.213 | 0.177 | 0.182 | 0.175 |
| | MASE | 0.583 | 0.477 | 0.511 | 0.487 |
| <i>Out-of-sample</i> | RMSE | 0.502 | 0.554 | 0.319 | 0.348 |
| | MASE | 0.241 | 0.262 | 0.157 | 0.124 |

Note: Green indicates “best” accuracy.

Out-of-sample the AR surprisingly performs better than the ARIMA, when comparing both the RMSE and MASE. However, the ETS-model is the “best” individual model by a wide margin based on out-of-sample RMSE and MASE. It is interesting to note that the exponential smoothing model outperforms the ARIMA model, which is far more frequently used in the house price forecasting literature. When including the weighted model, it is less clear which model is the best. The ETS-model still performs best based on the RMSE, while the weighted model is able to improve upon the MASE of the ETS. Based on these results we cannot reject the notion that a weighting scheme with inverse out-of-sample MSEs, can improve upon the forecast accuracy of the individual best forecast.

5.2 Diagnostics

In the methodology (3.3), the desired properties of the model residuals are detailed. Optimally, we would like the residuals to be white noise, $WN(0, \sigma^2)$. Residual, ACF and PACF plots for the three individual models are shown below. The general interpretation of the ACF and PACF plots is that the autocorrelation is statistically significant if it breaks either the upper or the lower confidence limits (the dotted blue lines). When examining these plots we

should be aware of the fact that on occasion autocorrelations can be significant by pure chance. Particularly this may be true if the autocorrelation is isolated and not at seasonal lags (e.g. the 12th and 24th lag). Repeating autocorrelations at seasonal lags indicate that the models have not been able to capture the seasonality component of the data sufficiently.

Figure 5.5 shows residual plots for the AR-model. There are two significant autocorrelations one at lag 7 and one at lag 12. Although not ideal, the negative autocorrelation at the seasonal 12th lag is not a big concern, as it does not seem to be a repeating pattern. In fact, the autocorrelation at the 24th lag is positive.

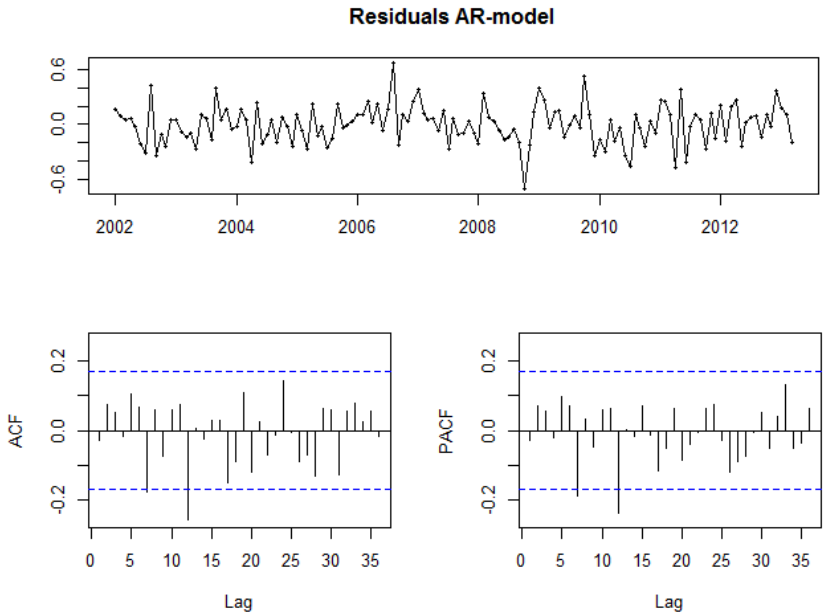


Figure 5.5: Residual diagnostics for AR(14).

The residuals of the ARIMA-model in Figure 5.6 seem very white noisy. Still there is a slightly significant autocorrelation at lag 7. Similarly, the residuals from the ETS-model in Figure 5.7 looks relatively well behaved. However, there is a significant autocorrelation at lag two.

Formal tests for residual autocorrelation, normality and homoscedasticity is presented in Table 5.3. The null hypothesis for all three test are that the residuals are well behaved, meaning there are no autocorrelation, they are normally distributed and homoscedastic (constant variance).

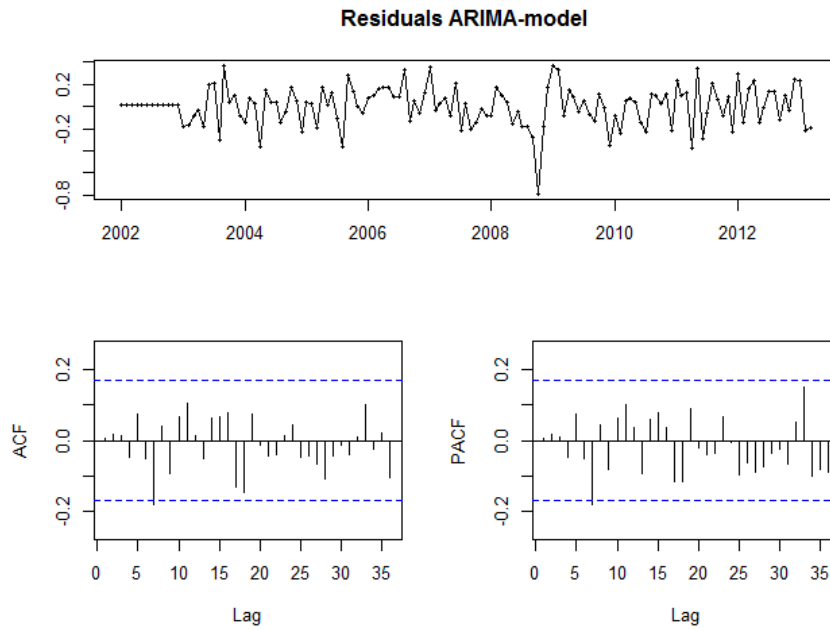


Figure 5.6: Residual diagnostics for $ARIMA(2,0,1)(0,1,1)$ [12].

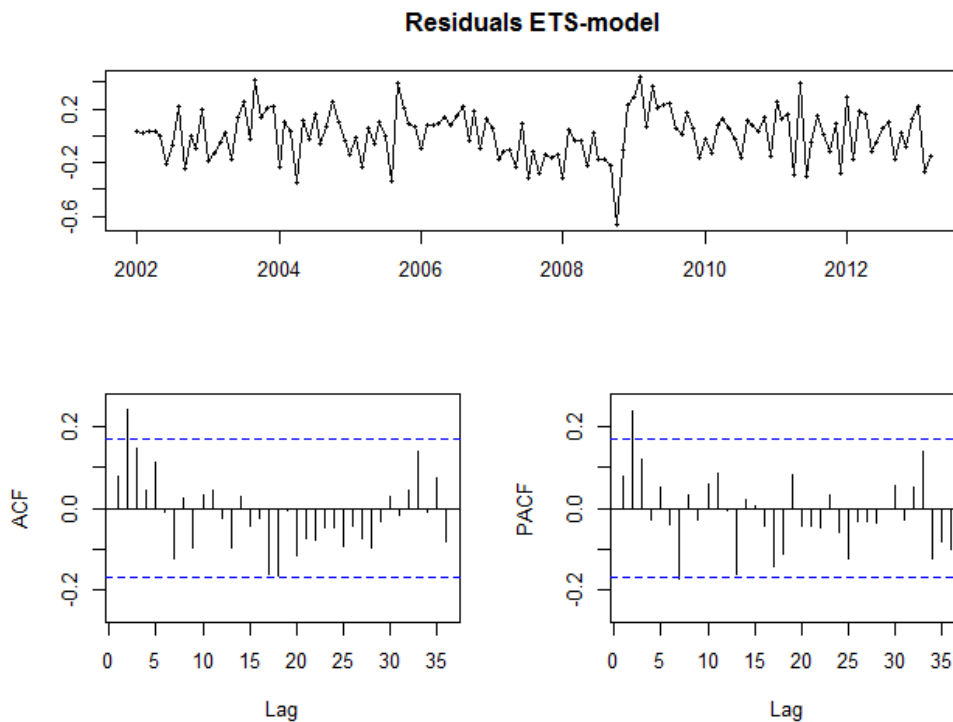


Figure 5.7: Residual diagnostics for $ETS(A,A,A)$

Both the AR- and ETS-model have significant autocorrelation in their residuals. This is not ideal as it means there are still information left in the data, which these models are not able to capture in its entirety. In addition, tests indicate that these two models have normally

distributed, but heteroscedastic residuals. On the other hand, test for the residuals of the ARIMA model show the complete opposite, namely that the model has uncorrelated, non-normal and homoscedastic residuals.

Table 5.3: Residual diagnostics tests

| | AR | ARIMA | ETS |
|---------------------------------------|----------------|----------------|----------------|
| <i>Residual autocorrelation test:</i> | 34.303 | 20.665 | 33.294 |
| Ljung-Box test* (p-values) | (0.000) | (0.297) | (0.000) |
| <i>Residual normality test:</i> | 2.925 | 32.556 | 2.898 |
| Jarque-Bera test** (p-values) | (0.232) | (0.000) | (0.235) |
| <i>Homoscedasticity test:</i> | 18.782 | 15.474 | 20.003 |
| Squared residual | (0.009) | (0.629) | (0.010) |
| Ljung-Box test*** (p-values) | | | |

Note: Red indicates rejection of null hypothesis at 5% and 10 % level.

* H_0 : residuals are uncorrelated, H_1 : residuals are correlated

** H_0 : residuals are normal, H_1 : residuals are non-normal

*** H_0 : residuals are homoscedastic, H_1 : residuals are heteroscedastic

The lack of well-behaved residuals for the AR- and ETS-models based on the Ljung-Box tests are not overly concerning. Particularly for ETS-model, the test result seem less relevant as the ACF/PACF plots showed that the autocorrelation was isolated at a non-seasonal lag. In addition, the ETS-model is as mentioned robust to heteroscedsticity.

5.3 Discussion

Before concluding the thesis, we should take a moment to reflect on the different aspects of the analysis and some alternative approaches that could have been employed. As the residual diagnostics reveal, the models applied here are not a perfect fit. One concern is the large number of estimated parameters in the AR- and ETS-models relative to the sample size. The two models have 17 and 16 estimated parameters (coefficients plus estimate of variance), respectively. The ARIMA-model is by far the most parsimonious in the set with only six estimated parameters. This comes in addition to being the model with the most well behaved residuals. At the same time, the ARIMA model performs the worst out-of-sample. We could

speculate that the other two models include too many parameters, considering the sample size. However, overfitted models normally have a good in-sample fit, but perform worse out-of-sample. The opposite is actually the case here. The AR and ETS have the worst in-sample fit, while they outperform the ARIMA out-of-sample.

Another potential point of critique is the choice of model selection and forecast combination criteria. Here, the AIC_C , a in-sample measure, is used for model selection, while the out-of-sample MSE is applied in the weighting scheme. Hansen (2009) warns against using different criteria for selection and combination, as this may lead to conflicting results.

It is also argued in section 4.2, that no transformations of the data is needed. This an unusual approach, as most forecasters in fact transforms the data before forecasting. Some argue that forecasting non-stationary series could lead to spurious results. Unit roots do not affect the ETS- and ARIMA-models, due to their flexibility. However, for the AR-model the non-stationarity of the series can induce inference problems.

Lastly, the focus of this thesis was forecasting with univariate time series models. Still, it would have been interesting to include multivariate models, such as VAR models, in the model set. Then we would be able to draw a more comprehensive conclusion about which models best forecast house prices in Norway.

6 Concluding remarks

The main objective of this thesis is to forecast residential house prices in Norway from April 2013 to March 2014. Three univariate time series models are employed in an attempt to find an appropriate fit. The three are an AR-, an ARIMA- and an exponential smoothing state space (ETS) model. The forecast from the three models are also combined in an effort to improve upon the accuracy of the single “best” forecast. This study implements a weighting scheme based on inverse out-of-sample mean square errors (MSEs). Weights of 0.29, 0.21 and 0.50 are assigned to the AR-, ARIMA- and ETS-model, respectively.

All three models looks to capture the seasonality in the house prices satisfactory. The analysis identifies the forecast from the ETS-model as the most accurate among the individual models based on both out-of-sample root mean square error (RMSE) and mean absolute scaled error (MASE). The weighted forecast has a higher RMSE (less accurate), but a lower (more accurate) MASE compared to the ETS. Thus, we cannot conclude either way, whether a combination of forecast can in fact improve upon the accuracy of the individually most accurate forecast, since the two measures give conflicting results. However, we can note that although ETS-models are not widely applied in forecasting of house prices, the results here indicate that they should be.

Residual diagnostics discover some suboptimal attributes of the AR- and ETS-models. Formal tests show that they have significant autocorrelation and heteroscedasticity in their residuals. By examining the ACF and PACF plots we find that the residuals appear to be reasonably well behaved. Hence, it does not appear that we have to be too concerned about spurious results.

7 References

- Akaike, Hirotogu. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. *B.N. Petrov, S. Csaki (Eds.), 2nd international symposium on information theory, Akademiai Kiado, Budapest, Hungary*, 267–281.
- Akaike, Hirotogu. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6), 716-723.
- Akaike, Hirotogu. (1981). Likelihood of a model and information criteria. *Journal of Econometrics*, 16(1), 3-14.
- Artis, M. J., & Zhang, W. (1990). BVAR forecasts for the G-7. *International Journal of Forecasting*, 6(3), 349-362.
- Bates, J. M., & Granger, C. W. J. (1969). The Combination of Forecasts. *OR*, 20(4), 451-468.
- Birch, John W. , & Sunderman, Mark A. (2003). Estimating Price Paths for Residential Real Estate. *Journal of Real Estate Research*, 25(3), 277-300.
- Bjørnland, Hilde C., Gerdrup, Karsten, Jore, Anne Sofie, Smith, Christie, & Thorsrud, Leif Anders. (2012). Does Forecast Combination Improve Norges Bank Inflation Forecasts?*. *Oxford Bulletin of Economics and Statistics*, 74(2), 163-179.
- Bjørnland, Hilde C., & Thorsrud, Leif Anders. (2012). CAMP Forecasts 2013. <http://www.bi.edu/cmeFiles/report2013Forecast.pdf>
- Box, George EP, & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2), 211-252.
- Box, George EP, & Draper, Norman R. (1987). *Empirical model-building and response surfaces*. New York: Wiley.
- Box, George EP, & Jenkins, Gailim M. (1970). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.
- Brown, Jane P., Song, Haiyan, & McGillivray, Alan. (1997). Forecasting UK house prices: A time varying coefficient approach. *Economic Modelling*, 14(4), 529-548.
- Brown, Robert G. (1963). *Smoothing, forecasting and prediction of discrete time series*. Englewood Cliffs, NJ: Prentice-Hall.
- Burnham, Kenneth P., & Anderson, David R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer.
- Canova, Fabio, & Hansen, Bruce E. (1995). Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability. *Journal of Business & Economic Statistics*, 13(3), 237-252.
- Case, Karl E., & Shiller, Robert J. (1989). The Efficiency of the Market for Single-Family Homes. *The American Economic Review*, 79(1), 125-137.
- Case, Karl E., & Shiller, Robert J. (1990). Forecasting Prices and Excess Returns in the Housing Market. *Real Estate Economics*, 18(3), 253-273.
- Chin, Lawrence, & Fan, Gang-Zhi. (2005). Autoregressive analysis of Singapore's private residential prices. *Property Management*, 23(4), 257-270.
- Christensen, Thor Herman (2003). Boligprisindeksen - Datagrunnlag og beregningsmetode. Retrieved 08.05.2013, from http://www.ssb.no/a/publikasjoner/pdf/notat_200383/notat_200383.pdf
- Clapp, John, & Giaccotto, Carmelo. (2002). Evaluating House Price Forecasts. *Journal of Real Estate Research*, 24(1), 1-26.
- Clemen, Robert T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559-583.
- Crawford, Gordon W., & Fratantoni, Michael C. (2003). Assessing the Forecasting Performance of Regime-Switching, ARIMA and GARCH Models of House Prices. *Real Estate Economics*, 31(2), 223-243.

- De Gooijer, Jan G., & Hyndman, Rob J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, 22(3), 443-473.
- Dhrymes, Phoebus J., & Peristiani, Stavros C. (1988). A comparison of the forecasting performance of WEFA and ARIMA time series methods. *International Journal of Forecasting*, 4(1), 81-101.
- Dickey, David A., & Fuller, Wayne A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49(4), 1057-1072.
- Engle, Robert F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.
- Fama, Eugene F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
- Gardner, Everette S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1), 1-28.
- Gu, Anthony. (2002). The Predictability of House Prices. *Journal of Real Estate Research*, 24(3), 213-234.
- Guirguis, HanyS, Giannikos, ChristosI, & Anderson, RandyI. (2005). The US Housing Market: Asset Pricing Forecasts Using Time Varying Coefficients. *The Journal of Real Estate Finance and Economics*, 30(1), 33-53.
- Gupta, Rangan, & Miller, StephenM. (2012). The Time-Series Properties of House Prices: A Case Study of the Southern California Market. *The Journal of Real Estate Finance and Economics*, 44(3), 339-361.
- Hansen, Peter R. (2009). In-Sample and Out-of-Sample Fit: Their Joint Distribution and its Implications for Model Selection. *Seminar paper*.
<http://www.fas.nus.edu.sg/ecs/events/seminar/seminar-papers/29Jan09.pdf>
- Hein, Scott E., & Spudeck, Raymond E. (1988). Forecasting the daily federal funds rate. *International Journal of Forecasting*, 4(4), 581-591.
- Hepsen, Ali, & Vatansever, Metin. (2010). Forecasting Future Trends in Dubai Housing Market by Using Box - Jenkins Autoregressive Integrated Moving Average *International Journal of Housing Markets and Analysis*, 4(3), 210-223.
- Hill, R Carter, Griffiths, William E, & Lim, Guay C. (2008). *Principles of econometrics* (4th ed.): Wiley Hoboken, NJ.
- Holt, Charles C. (1957, reprinted 2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5-10.
- Hurvich, Clifford M., & Tsai, Chih-Ling. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, 76(2), 297-307.
- Hyndman, Rob J. . (2009). Forecast: Forecasting functions for time series and linear models (Version 4.04). Retrieved from <http://cran.r-project.org/web/packages/forecast/index.html>
- Hyndman, Rob J. , & Khandakar, Yeasmin (2008). Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, 27(3).
- Hyndman, Rob J., & Athanasopoulos, George (2013). *Forecasting: principles and practice* Retrieved from <http://otexts.com/fpp/>
- Hyndman, Rob J., & Koehler, Anne B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688.
- Hyndman, Rob J., Koehler, Anne B., Ord, J. Keith, & Snyder, Ralph D. (2005). Prediction intervals for exponential smoothing using two new classes of state space models. *Journal of Forecasting*, 24(1), 17-37.
- Hyndman, Rob J., Koehler, Anne B., Ord, Keith, & Snyder, Ralph D. (2008). *Forecasting with exponential smoothing: The State Space Approach*. Berlin: Springer.

- Hyndman, Rob J., Koehler, Anne B., Snyder, Ralph D., & Grose, Simone. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18(3), 439-454.
- Inoue, Atsushi, & Kilian, Lutz. (2005). In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use? *Econometric Reviews*, 23(4), 371-402.
- Jarque, Carlos M., & Bera, Anil K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255-259.
- Kapetanios, George, Labhard, Vincent, & Price, Simon. (2008a). Forecast combination and the Bank of England's suite of statistical forecasting models. *Economic Modelling*, 25(4), 772-792.
- Kapetanios, George, Labhard, Vincent, & Price, Simon. (2008b). Forecasting Using Bayesian and Information-Theoretic Model Averaging. *Journal of Business & Economic Statistics*, 26(1), 33-41.
- Kletting, P., & Glatting, G. (2009). Model selection for time-activity curves: the corrected Akaike information criterion and the F-test. *Z Med Phys*, 19(3), 200-206.
- Koehler, Anne B., Snyder, Ralph D., & Ord, J. Keith. (2001). Forecasting models and prediction intervals for the multiplicative Holt–Winters method. *International Journal of Forecasting*, 17(2), 269-286.
- Kolassa, Stephan. (2011). Combining exponential smoothing forecasts using Akaike weights. *International Journal of Forecasting*, 27(2), 238-251.
- Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79-86.
- Kuo, Chiong-Long. (1996). Serial correlation and seasonality in the real estate market. *The Journal of Real Estate Finance and Economics*, 12(2), 139-162.
- Kwiatkowski, Denis, Phillips, Peter C. B., Schmidt, Peter, & Shin, Yongcheol. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3), 159-178.
- Larsen, Richard J, & Marx, Morris L. (2012). *Introduction to mathematical statistics and its applications* (5th ed.). Boston: Pearson/Prentice Hall.
- Limsombunchai, Visit. (2004). *House price prediction: Hedonic price model vs. artificial neural network*. Paper presented at the New Zealand Agricultural and Resource Economics Society Conference, 25-26 June 2004, Blenheim, New Zealand.
- Lin, Ting Hsiang, & Dayton, C. Mitchell. (1997). Model Selection Information Criteria for Non-Nested Latent Class Models. *Journal of Educational and Behavioral Statistics*, 22(3), 249-264.
- Ljung, Greta M., & Box, George E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.
- Ludden, T. M., Beal, S. L., & Sheiner, L. B. (1994). Comparison of the Akaike Information Criterion, the Schwarz criterion and the F test as guides to model selection. *J Pharmacokinet Biopharm*, 22(5), 431-445.
- Makridakis, Spyros, & Hibon, Michèle. (1991). Exponential smoothing: The effect of initial values and loss functions on post-sample forecasting accuracy. *International Journal of Forecasting*, 7(3), 317-330.
- Marcellino, Massimiliano, Stock, James H., & Watson, Mark W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1-2), 499-526.
- Meese, Richard A., & Wallace, Nancy E. (1997). The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches. *The Journal of Real Estate Finance and Economics*, 14(1-2), 51-73.

- NEF.). Historiske priser 1985-2013. Retrieved 08.05.2013, 2013, from http://www.nef.no/xp/pub/topp/boligprisstatistikk/historiske_priser/index.html
- Ord, J. K., Koehler, A. B., & Snyder, R. D. (1997). Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models. *Journal of the American Statistical Association*, 92(440), 1621-1629.
- Osborn, Denise R., Chui, A. P. L., Smith, Jeremy P., & Birchenhall, C. R. (1988). SEASONALITY AND THE ORDER OF INTEGRATION FOR CONSUMPTION*. *Oxford Bulletin of Economics and Statistics*, 50(4), 361-377.
- Pegels, C. C. (1969). Exponential Forecasting: Some New Variations. *Management Science*, 15(5), 311-315.
- Posada, David, & Buckley, Thomas R. (2004). Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Systematic Biology*, 53(5), 793-808.
- Pöyry. (2013). Eiendomsmeglerbransjens boligprisstatistikk april 2013 Retrieved 08.05.2013, from <http://www.eff.no/filer/04%20Finn%2013%2004apr.pdf>
- Schwarz, Gideon. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461-464.
- Simkins, Scott. (1995). Forecasting with vector autoregressive (VAR) models subject to business cycle restrictions. *International Journal of Forecasting*, 11(4), 569-583.
- Sklarz, Michael A., Miller, Norman G., & Gersch, Will. (1987). Forecasting Using Long-Order Autoregressive Processes: An Example Using Housing Starts. *Real Estate Economics*, 15(4), 374-388.
- Snyder, R. D. (1985). Recursive Estimation of Dynamic Linear Models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(2), 272-276.
- Stevenson, Simon. (2007). A comparison of the forecasting ability of ARIMA models. *Journal of Property Investment & Finance*, 25(3), 223-240.
- Stock, James H., & Watson, Mark W. (1998). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. *National Bureau of Economic Research Working Paper Series, No. 6607*.
- Stock, James H., & Watson, Mark W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, 41(3), 788-829.
- Stock, James H., & Watson, Mark W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6), 405-430.
- Taylor, James W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4), 715-725.
- Timmermann, Allan. (2006). Chapter 4 Forecast Combinations. In C. W. J. G. G. Elliott & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. Volume 1, pp. 135-196): Elsevier.
- Tse, Raymond. (1997). An application of the ARIMA model to real-estate prices in Hong Kong. *Journal of Property Finance*, 8(2), 152-163.
- Wilson, I. D., Paris, S. D., Ware, J. A., & Jenkins, D. H. (2002). Residential property price time series forecasting with neural networks. *Knowledge-Based Systems*, 15(5-6), 335-341.
- Winters, Peter R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, 6(3), 324-342.
- Yule, G. Udny. (1927). On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 226, 267-298.

- Zhang, Guoqiang, Eddy Patuwo, B., & Y. Hu, Michael. (1998). Forecasting with artificial neural networks:: The state of the art. *International Journal of Forecasting*, 14(1), 35-62.
- Zhou, Zhong-guo. (1997). Forecasting Sales and Price for Existing Single-Family Homes: A VAR Model with Error Correction. *Journal of Real Estate Research*, 14(2), 155-167.

A. Appendix - The estimated models

Table A.1: The AR(14) with constant and trend.

| Lag | Coefficient | Std.error | Lag | Coefficient | Std.error |
|--|-------------|-----------|--------|-------------|-----------|
| Constant (c) | 13.8041 | 0.4202 | AR(7) | 0.1096 | 0.0944 |
| Trend (t) | 0.1219 | 0.0054 | AR(8) | -0.0773 | 0.1005 |
| AR(1) | 1.2137 | 0.0849 | AR(9) | 0.0296 | 0.0969 |
| AR(2) | -0.2579 | 0.1080 | AR(10) | -0.1181 | 0.0945 |
| AR(3) | 0.0803 | 0.0927 | AR(11) | 0.1653 | 0.0947 |
| AR(4) | -0.1453 | 0.0915 | AR(12) | 0.6662 | 0.0950 |
| AR(5) | 0.2355 | 0.0928 | AR(13) | -0.9341 | 0.1092 |
| AR(6) | -0.2727 | 0.0936 | AR(14) | 0.2132 | 0.0880 |
| MLE sigma ² | 0.04542 | | | | |
| AIC _C | 20.09 | | | | |
| Next best models based on AIC _C | | | | | |
| AR(15) w/ c and t | 21.27 | | | | |
| AR(13) w/ c and t | 23.14 | | | | |

Table A.2: The ARIMA(2,0,1)(0,1,1)[12] with drift

| Lag | Coefficient | Std.error |
|--|-------------|-----------|
| Drift | 0.1225 | 0.0102 |
| AR(1) | 1.8601 | 0.0733 |
| AR(2) | -0.8772 | 0.0718 |
| MA(1) | -0.6111 | 0.1344 |
| Seasonal MA(1) | -0.7867 | 0.0995 |
| MLE sigma ² | 0.03449 | |
| AIC _C | -36.71 | |
| Next best models based on AIC _C | | |
| ARIMA(2,0,1)(0,1,2)[12] with drift | -34.54 | |
| ARIMA(3,0,1)(0,1,1)[12] with drift | -34.46 | |

Table A.3: The ETS(A,A,A)

| | Smoothing parameters | Initial states | | | | | | | | |
|--|----------------------|----------------|-------------------|---------|-------------------|---------|-------------------|--------|--------------------|--------|
| Alpha | 1.2052 | Level(l) | 14.5911 | | | | | | | |
| Beta | 0.1092 | Growth(b) | 0.0529 | | | | | | | |
| Gamma | 0.0000 | Season(s) | s ₀ = | -0.6738 | s ₋₃ = | 0.0905 | s ₋₆ = | 0.0790 | s ₋₉ = | 0.2597 |
| | | | s ₋₁ = | -0.3630 | s ₋₄ = | 0.2439 | s ₋₇ = | 0.2518 | s ₋₁₀ = | 0.2014 |
| | | | s ₋₂ = | -0.1793 | s ₋₅ = | -0.2865 | s ₋₈ = | 0.3112 | s ₋₁₁ = | 0.0650 |
| | | | | | | | | | | |
| MLE sigma ² | 0.03302 | | | | | | | | | |
| AIC _C | 238.42 | | | | | | | | | |
| Next best models based on AIC _C | | | | | | | | | | |
| ETS(M,A,M) | 239.38 | | | | | | | | | |
| ETS(A,A _d ,A) | 240.02 | | | | | | | | | |

B. Appendix - Point forecasts and prediction intervals

Table B.1: Point forecast and prediction intervals for 2013:4-2014:3, individual models.

| | Month | Point forecast | Lower 95% | Lower 80% | Upper 80% | Upper 95% |
|-----------------------------|----------|----------------|--------------|--------------|--------------|--------------|
| AR(14) | Apr 2013 | 31.690 | 31.272 | 31.417 | 31.963 | 32.108 |
| | May 2013 | 31.583 | 30.926 | 31.153 | 32.012 | 32.239 |
| | Jun 2013 | 31.547 | 30.717 | 31.005 | 32.090 | 32.378 |
| | Jul 2013 | 31.389 | 30.410 | 30.749 | 32.029 | 32.368 |
| | Aug 2013 | 31.939 | 30.849 | 31.227 | 32.652 | 33.029 |
| | Sep 2013 | 31.761 | 30.557 | 30.973 | 32.548 | 32.965 |
| | Oct 2013 | 31.545 | 30.252 | 30.699 | 32.391 | 32.839 |
| | Nov 2013 | 31.330 | 29.966 | 30.438 | 32.221 | 32.693 |
| | Dec 2013 | 31.196 | 29.781 | 30.271 | 32.122 | 32.612 |
| | Jan 2014 | 32.101 | 30.652 | 31.153 | 33.049 | 33.551 |
| | Feb 2014 | 32.010 | 30.544 | 31.051 | 32.969 | 33.477 |
| | Mar 2014 | 31.974 | 30.498 | 31.009 | 32.939 | 33.449 |
| ARIMA(2,0,1) (0,1,1)[12] | Apr 2013 | 31.578 | 31.214 | 31.340 | 31.816 | 31.943 |
| | May 2013 | 31.593 | 31.010 | 31.212 | 31.974 | 32.176 |
| | Jun 2013 | 31.435 | 30.650 | 30.922 | 31.949 | 32.221 |
| | Jul 2013 | 31.087 | 30.110 | 30.448 | 31.726 | 32.064 |
| | Aug 2013 | 31.692 | 30.535 | 30.936 | 32.448 | 32.848 |
| | Sep 2013 | 31.550 | 30.229 | 30.686 | 32.415 | 32.872 |
| | Oct 2013 | 31.290 | 29.817 | 30.327 | 32.253 | 32.762 |
| | Nov 2013 | 31.153 | 29.545 | 30.101 | 32.204 | 32.760 |
| | Dec 2013 | 30.868 | 29.141 | 29.739 | 31.997 | 32.595 |
| | Jan 2014 | 31.764 | 29.933 | 30.567 | 32.961 | 33.595 |
| | Feb 2014 | 31.945 | 30.025 | 30.689 | 33.200 | 33.864 |
| | Mar 2014 | 32.088 | 30.094 | 30.784 | 33.392 | 34.082 |
| ETS(A,A,A) | Apr 2013 | 31.631 | 31.275 | 31.398 | 31.864 | 31.987 |
| | May 2013 | 31.730 | 31.110 | 31.324 | 32.135 | 32.349 |
| | Jun 2013 | 31.715 | 30.889 | 31.175 | 32.255 | 32.541 |
| | Jul 2013 | 31.507 | 30.495 | 30.845 | 32.169 | 32.519 |
| | Aug 2013 | 32.195 | 31.006 | 31.418 | 32.972 | 33.384 |
| | Sep 2013 | 32.199 | 30.838 | 31.309 | 33.089 | 33.561 |
| | Oct 2013 | 32.087 | 30.556 | 31.086 | 33.089 | 33.619 |
| | Nov 2013 | 32.061 | 30.360 | 30.949 | 33.174 | 33.762 |
| | Dec 2013 | 31.908 | 30.037 | 30.685 | 33.132 | 33.779 |
| | Jan 2014 | 32.805 | 30.763 | 31.470 | 34.140 | 34.847 |
| | Feb 2014 | 33.099 | 30.884 | 31.651 | 34.547 | 35.314 |
| | Mar 2014 | 33.315 | 30.926 | 31.753 | 34.878 | 35.705 |

Table B.2: Point forecast and prediction intervals for 2013:4-2014:3, weighted model.

| | Month | Point forecast | Lower 95% | Lower 80% | Upper 80% | Upper 95% |
|----------------------|----------|----------------|--------------|--------------|--------------|--------------|
| Weighted forecast | Apr 2013 | 31.637 | 31.262 | 31.392 | 31.883 | 32.013 |
| | May 2013 | 31.659 | 31.037 | 31.252 | 32.067 | 32.282 |
| | Jun 2013 | 31.609 | 30.791 | 31.074 | 32.145 | 32.428 |
| | Jul 2013 | 31.387 | 30.392 | 30.736 | 32.038 | 32.382 |
| | Aug 2013 | 32.019 | 30.865 | 31.264 | 32.773 | 33.172 |
| | Sep 2013 | 31.941 | 30.633 | 31.085 | 32.796 | 33.249 |
| | Oct 2013 | 31.768 | 30.318 | 30.820 | 32.717 | 33.219 |
| | Nov 2013 | 31.665 | 30.080 | 30.629 | 32.702 | 33.250 |
| | Dec 2013 | 31.491 | 29.781 | 30.373 | 32.610 | 33.202 |
| | Jan 2014 | 32.390 | 30.561 | 31.194 | 33.586 | 34.219 |
| | Feb 2014 | 32.550 | 30.611 | 31.282 | 33.818 | 34.489 |
| | Mar 2014 | 32.679 | 30.633 | 31.341 | 34.016 | 34.725 |

C. Appendix - The data set

Table C.1: Monthly nominal m2 house prices (in NOK1000) in Norway 2002:1-2013:3.

| | 2002-2003 | 2004-2005 | 2006-2007 | 2008-2009 | 2010-2011 | 2012-2013 |
|-----|-----------|-----------|-----------|-----------|-----------|-----------|
| Jan | 14.747 | 16.464 | 19.844 | 23.480 | 24.746 | 28.855 |
| Feb | 14.971 | 16.764 | 20.192 | 23.641 | 24.866 | 29.047 |
| Mar | 15.127 | 17.001 | 20.513 | 23.723 | 25.079 | 29.407 |
| Apr | 15.285 | 16.831 | 20.846 | 23.780 | 25.385 | 29.830 |
| May | 15.297 | 16.902 | 21.123 | 23.539 | 25.527 | 29.881 |
| Jun | 14.980 | 16.830 | 21.259 | 23.364 | 25.467 | 29.821 |
| Jul | 14.542 | 16.715 | 21.270 | 22.852 | 25.058 | 29.675 |
| Aug | 15.309 | 17.331 | 22.272 | 23.173 | 25.775 | 30.497 |
| Sep | 15.017 | 17.344 | 22.376 | 22.744 | 25.842 | 30.382 |
| Oct | 14.731 | 17.457 | 22.520 | 21.723 | 25.746 | 30.275 |
| Nov | 14.488 | 17.574 | 22.539 | 21.180 | 25.837 | 30.186 |
| Des | 14.373 | 17.405 | 22.584 | 20.947 | 25.546 | 30.149 |
| Jan | 15.007 | 18.142 | 23.666 | 21.918 | 26.634 | 31.311 |
| Feb | 15.001 | 18.371 | 23.904 | 22.486 | 27.103 | 31.428 |
| Mar | 14.993 | 18.328 | 24.063 | 22.683 | 27.519 | 31.453 |
| Apr | 15.058 | 18.501 | 24.227 | 23.109 | 27.506 | |
| May | 14.832 | 18.509 | 24.136 | 23.366 | 27.931 | |
| Jun | 14.751 | 18.534 | 24.202 | 23.519 | 27.732 | |
| Jul | 14.676 | 18.302 | 23.749 | 23.514 | 27.422 | |
| Aug | 15.260 | 18.609 | 24.273 | 24.247 | 28.254 | |
| Sep | 15.546 | 18.860 | 23.974 | 24.224 | 28.315 | |
| Oct | 15.567 | 19.003 | 23.637 | 24.240 | 28.103 | |
| Nov | 15.710 | 19.096 | 23.381 | 24.276 | 28.145 | |
| Des | 15.774 | 19.027 | 22.999 | 23.951 | 27.741 | |

D. Appendix - The R code

OBS! Running the code in its entirety may take a while.

```
library(forecast)
library(graphics)
library(tseries)
library(pastecs)

# DATA SET #####
# House prices Jan. 02 - March 13
data <- structure(c(14.747368, 14.971517, 15.127844, 15.285704, 15.297407, 14.980323, 14.542878, 15.309535, 15.017721, 14.731572,
14.488528, 14.373321, 15.007667, 15.00134, 14.993977, 15.058588, 14.832451, 14.751847, 14.67612, 15.260274, 15.546455, 15.567483,
15.710143, 15.774186, 16.464181, 16.764038, 17.001064, 16.831662, 16.902986, 16.830444, 16.715932, 17.331179, 17.344735,
17.45788, 17.574839, 17.405844, 18.14261, 18.371732, 18.328169, 18.501254, 18.509556, 18.534803, 18.302446, 18.609222, 18.860566,
19.00363, 19.096919, 19.027872, 19.844294, 20.192769, 20.512996, 20.846508, 21.123915, 21.259848, 21.270676, 22.272309, 22.375843,
22.520273, 22.539099, 22.58466, 23.666829, 23.904162, 24.063079, 24.227331, 24.136405, 24.202093, 23.749147, 24.273819,
23.97446, 23.637499, 23.381023, 22.999301, 23.480893, 23.64173, 23.723986, 23.780698, 23.539532, 23.364047, 22.852094, 23.173906,
22.744607, 21.723483, 21.18022, 20.94743, 21.918237, 22.48661, 22.683303, 23.1099, 23.36695, 23.519433, 23.513999, 24.247008,
24.224611, 24.240652, 24.276591, 23.951361, 24.746266, 24.866524, 25.079439, 25.385703, 25.526986, 25.46707, 25.058808, 25.775007,
25.842061, 25.746739, 25.83712, 25.546944, 26.634963, 27.103609, 27.519261, 27.506359, 27.931635, 27.7328, 27.422783, 28.254738,
28.315666, 28.10369, 28.14566, 27.741635, 28.855819, 29.047454, 29.407865, 29.830691, 29.8818, 29.821657, 29.675817, 30.497315,
30.382523, 30.275193, 30.186988, 30.149414, 31.311087, 31.428756, 31.453484), .Tsp = c(2002, 2013.16666666667, 12), class = "ts")

diff <- diff(data)

# Data exploration, misc.
plot(data, main = "House prices in Norway 2002:1-2013:3", xlab = "Year", ylab = "m2 prices (in NOK1000)",
      xlim=c(2002,2013+3/12), ylim=c(14,35))
plot(stl(data, s.window=c("periodic"))) #decomposition
stl(data, s.window=c("periodic"))
stat.desc(data) #descriptive statistics
plot(diff, lwd="1", ylab="1st diff", xlab="")

# Unit root tests
adf.test(data, alternative = "s") #augmented Dickey Fuller test
kpss.test(data) #KPSS test

# MODEL ESTIMATIONS - MAIN #####
# ETS
ets <- ets(data, ic=c("aicc"), bounds="admissible")
fets <- forecast(ets, h=12)
plot(fets, lwd="1.5", main="ETS(A,Ad,A)", xlab = "Year", ylab = "m2 prices (in NOK1000)",
      xlim=c(round(2011),round(2014+4/12, digits = 1)), ylim=c(25,37))
lines(ets$fitted, col="blue")

# ARIMA
ima <- auto.arima(data, ic=c("aicc"), approximation=F, stepwise=F, parallel=T)
fima <- forecast(ima, h=12)
plot(fima, lwd="2", main="ARIMA(2,0,1)(0,1,1)[12] with drift", xlab = "Year", ylab = "m2 prices (in NOK1000)",
      xlim=c(2011,2014+4/12), ylim=c(25,37))
lines(fima$fitted, col="blue")

# AR
arm <- Arima(data, order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
farm <- forecast(arm, h=12)
plot(farm, lwd="2", main="AR(14) with intercept and drift", xlab = "Year", ylab = "m2 prices (in NOK1000)",
      xlim=c(2011,2014+4/12), ylim=c(25,37))
lines(farm$fitted, col="blue")

# RESIDUAL DIAGNOSTICS #####
mean(ets$residuals)
mean(ima$residuals)
mean(arm$residuals)

tsdisplay(ets$residuals, main="Residuals ETS-model") #Residual, ACF and PACF plots
tsdisplay(ima$residuals, main="Residuals ARIMA-model")
tsdisplay(arm$residuals, main="Residuals AR-model")

jarque.bera.test(ets$residuals) #Jarque-Bera test - residual normality
jarque.bera.test(ima$residuals)
jarque.bera.test(arm$residuals)

logLik(ets) #Degrees of freedom determination
```

```

logLik(ima)
logLik(arm)

Box.test(ets$residuals, lag = 24, type = c("Ljung-Box"), fitdf = 16)      #Ljung-Box test - residual autocorrelation
Box.test(ima$residuals, lag = 24, type = c("Ljung-Box"), fitdf = 6)
Box.test(arm$residuals, lag = 24, type = c("Ljung-Box"), fitdf = 17)

Box.test(ets$residuals^2, lag = 24, type = c("Ljung-Box"), fitdf = 16)   #Ljung-Box test - homoscedasticity
Box.test(ima$residuals^2, lag = 24, type = c("Ljung-Box"), fitdf = 6)
Box.test(arm$residuals^2, lag = 24, type = c("Ljung-Box"), fitdf = 17)

# Histograms
bins=seq(-2,2,by=0.1)

hist(ets$residuals,breaks=bins, prob=T)
lines(density(ets$residuals))

hist(ima$residuals,breaks=bins)
lines(density(ima$residuals))

hist(arm$residuals,breaks=bins, prob=T)
lines(density(arm$residuals))

hist(diff,breaks=bins, prob=T)
lines(density(diff))

# ACCURACY #####
#In-sample
accets <- accuracy(ets)
accima <- accuracy(ima)
accarm <- accuracy(arm)

#Out-of-sample - Cross validation
#ETS
newets1 <- ets(window(data, start=2002+0/12, end=2011+2/12), ic=c("aicc"), bounds="usual")
newets2 <- ets(window(data, start=2002+1/12, end=2011+3/12), ic=c("aicc"),bounds="usual")
newets3 <- ets(window(data, start=2002+2/12, end=2011+4/12), ic=c("aicc"),bounds="usual")
newets4 <- ets(window(data, start=2002+3/12, end=2011+5/12), ic=c("aicc"),bounds="usual")
newets5 <- ets(window(data, start=2002+4/12, end=2011+6/12), ic=c("aicc"), bounds="usual")
newets6 <- ets(window(data, start=2002+5/12, end=2011+7/12), ic=c("aicc"),bounds="usual")
newets7 <- ets(window(data, start=2002+6/12, end=2011+8/12), ic=c("aicc"),bounds="usual")
newets8 <- ets(window(data, start=2002+7/12, end=2011+9/12), ic=c("aicc"),bounds="usual")
newets9 <- ets(window(data, start=2002+8/12, end=2011+10/12), ic=c("aicc"),bounds="usual")
newets10 <- ets(window(data, start=2002+9/12, end=2011+11/12), ic=c("aicc"),bounds="usual")
newets11 <- ets(window(data, start=2002+10/12, end=2011+12/12), ic=c("aicc"),bounds="usual")
newets12 <- ets(window(data, start=2002+11/12, end=2012+1/12), ic=c("aicc"),bounds="usual")

acets1 <- accuracy(forecast(newets1,h=12), window(data, start=2011+3/12))
acets2 <- accuracy(forecast(newets2,h=12), window(data, start=2011+4/12))
acets3 <- accuracy(forecast(newets3,h=12), window(data, start=2011+5/12))
acets4 <- accuracy(forecast(newets4,h=12), window(data, start=2011+6/12))
acets5 <- accuracy(forecast(newets5,h=12), window(data, start=2011+7/12))
acets6 <- accuracy(forecast(newets6,h=12), window(data, start=2011+8/12))
acets7 <- accuracy(forecast(newets7,h=12), window(data, start=2011+9/12))
acets8 <- accuracy(forecast(newets8,h=12), window(data, start=2011+10/12))
acets9 <- accuracy(forecast(newets9,h=12), window(data, start=2011+11/12))
acets10 <- accuracy(forecast(newets10,h=12), window(data, start=2011+12/12))
acets11 <- accuracy(forecast(newets11,h=12), window(data, start=2012+1/12))
acets12 <- accuracy(forecast(newets12,h=12), window(data, start=2012+2/12))

MSEets <- (acets1[1]^2+acets2[1]^2+acets3[1]^2+acets4[1]^2+acets5[1]^2+acets6[1]^2+
acets7[1]^2+acets8[1]^2+acets9[1]^2+acets10[1]^2+acets11[1]^2+acets12[1]^2)/12      #average MSE
RMSEets <- (acets1[2]+acets2[2]+acets3[2]+acets4[2]+acets5[2]+acets6[2]+
acets7[2]+acets8[2]+acets9[2]+acets10[2]+acets11[2]+acets12[2])/12      #average RMSE
MASEets <- (acets1[6]+acets2[6]+acets3[6]+acets4[6]+acets5[6]+acets6[6]+
acets7[6]+acets8[6]+acets9[6]+acets10[6]+acets11[6]+acets12[6])/12      #average MASE

# ARIMA
newima1 <- auto.arima(window(data, start=2002+0/12, end=2011+2/12), approximation=F, stepwise=F, parallel=T)
newima2 <- auto.arima(window(data, start=2002+1/12, end=2011+3/12), approximation=F, stepwise=F, parallel=T)
newima3 <- auto.arima(window(data, start=2002+2/12, end=2011+4/12), approximation=F, stepwise=F, parallel=T)
newima4 <- auto.arima(window(data, start=2002+3/12, end=2011+5/12), approximation=F, stepwise=F, parallel=T)
newima5 <- auto.arima(window(data, start=2002+4/12, end=2011+6/12), approximation=F, stepwise=F, parallel=T)
newima6 <- auto.arima(window(data, start=2002+5/12, end=2011+7/12), approximation=F, stepwise=F, parallel=T)
newima7 <- auto.arima(window(data, start=2002+6/12, end=2011+8/12), approximation=F, stepwise=F, parallel=T)

```

```

newima8 <- auto.arima(window(data, start=2002+7/12, end=2011+9/12), approximation=F, stepwise=F, parallel=T)
newima9 <- auto.arima(window(data, start=2002+8/12, end=2011+10/12), approximation=F, stepwise=F, parallel=T)
newima10 <- auto.arima(window(data, start=2002+9/12, end=2011+11/12), approximation=F, stepwise=F, parallel=T)
newima11 <- auto.arima(window(data, start=2002+10/12, end=2011+12/12), approximation=F, stepwise=F, parallel=T)
newima12 <- auto.arima(window(data, start=2002+11/12, end=2012+1/12), approximation=F, stepwise=F, parallel=T)

acima1 <- accuracy(forecast(newima1,h=12), window(data, start=2011+3/12))
acima2 <- accuracy(forecast(newima2,h=12), window(data, start=2011+4/12))
acima3 <- accuracy(forecast(newima3,h=12), window(data, start=2011+5/12))
acima4 <- accuracy(forecast(newima4,h=12), window(data, start=2011+6/12))
acima5 <- accuracy(forecast(newima5,h=12), window(data, start=2011+7/12))
acima6 <- accuracy(forecast(newima6,h=12), window(data, start=2011+8/12))
acima7 <- accuracy(forecast(newima7,h=12), window(data, start=2011+9/12))
acima8 <- accuracy(forecast(newima8,h=12), window(data, start=2011+10/12))
acima9 <- accuracy(forecast(newima9,h=12), window(data, start=2011+11/12))
acima10 <- accuracy(forecast(newima10,h=12), window(data, start=2011+12/12))
acima11 <- accuracy(forecast(newima11,h=12), window(data, start=2012+1/12))
acima12 <- accuracy(forecast(newima12,h=12), window(data, start=2012+2/12))

MSEima <- (acima1[1]^2+acima2[1]^2+acima3[1]^2+acima4[1]^2+acima5[1]^2+acima6[1]^2+      #average MSE
           acima7[1]^2+acima8[1]^2+acima9[1]^2+acima10[1]^2+acima11[1]^2+acima12[1]^2)/12
RMSEima <- (acima1[2]+acima2[2]+acima3[2]+acima4[2]+acima5[2]+acima6[2]+      #average RMSE
           acima7[2]+acima8[2]+acima9[2]+acima10[2]+acima11[2]+acima12[2])/12
MASEima <- (acima1[6]+acima2[6]+acima3[6]+acima4[6]+acima5[6]+acima6[6]+      #average MASE
           acima7[6]+acima8[6]+acima9[6]+acima10[6]+acima11[6]+acima12[6])/12

# AR
newarm1 <- Arima(window(data, start=2002+0/12, end=2011+2/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm2 <- Arima(window(data, start=2002+1/12, end=2011+3/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm3 <- Arima(window(data, start=2002+2/12, end=2011+4/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm4 <- Arima(window(data, start=2002+3/12, end=2011+5/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm5 <- Arima(window(data, start=2002+4/12, end=2011+6/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm6 <- Arima(window(data, start=2002+5/12, end=2011+7/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm7 <- Arima(window(data, start=2002+6/12, end=2011+8/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm8 <- Arima(window(data, start=2002+7/12, end=2011+9/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm9 <- Arima(window(data, start=2002+8/12, end=2011+10/12), order=c(14,0,0), method=c("ML"), include.mean=T, include.drift=T)
newarm10 <- Arima(window(data, start=2002+9/12, end=2011+11/12), order=c(14,0,0), method=c("ML"), include.mean=T,
include.drift=T)
newarm11 <- Arima(window(data, start=2002+10/12, end=2011+12/12), order=c(14,0,0), method=c("ML"), include.mean=T,
include.drift=T)
newarm12 <- Arima(window(data, start=2002+11/12, end=2012+1/12), order=c(14,0,0), method=c("ML"), include.mean=T,
include.drift=T)
acarm1 <- accuracy(forecast(newarm1,h=12), window(data, start=2011+3/12))
acarm2 <- accuracy(forecast(newarm2,h=12), window(data, start=2011+4/12))
acarm3 <- accuracy(forecast(newarm3,h=12), window(data, start=2011+5/12))
acarm4 <- accuracy(forecast(newarm4,h=12), window(data, start=2011+6/12))
acarm5 <- accuracy(forecast(newarm5,h=12), window(data, start=2011+7/12))
acarm6 <- accuracy(forecast(newarm6,h=12), window(data, start=2011+8/12))
acarm7 <- accuracy(forecast(newarm7,h=12), window(data, start=2011+9/12))
acarm8 <- accuracy(forecast(newarm8,h=12), window(data, start=2011+10/12))
acarm9 <- accuracy(forecast(newarm9,h=12), window(data, start=2011+11/12))
acarm10 <- accuracy(forecast(newarm10,h=12), window(data, start=2011+12/12))
acarm11 <- accuracy(forecast(newarm11,h=12), window(data, start=2012+1/12))
acarm12 <- accuracy(forecast(newarm12,h=12), window(data, start=2012+2/12))

MSEarm <- (acarm1[1]^2+acarm2[1]^2+acarm3[1]^2+acarm4[1]^2+acarm5[1]^2+acarm6[1]^2+      #average MSE
           acarm7[1]^2+acarm8[1]^2+acarm9[1]^2+acarm10[1]^2+acarm11[1]^2+acarm12[1]^2)/12
RMSEarm <- (acarm1[2]+acarm2[2]+acarm3[2]+acarm4[2]+acarm5[2]+acarm6[2]+      #average RMSE
           acarm7[2]+acarm8[2]+acarm9[2]+acarm10[2]+acarm11[2]+acarm12[2])/12
MASEarm <- (acarm1[6]+acarm2[6]+acarm3[6]+acarm4[6]+acarm5[6]+acarm6[6]+      #average MASE
           acarm7[6]+acarm8[6]+acarm9[6]+acarm10[6]+acarm11[6]+acarm12[6])/12

# FORECAST COMBINATION #####
#Inverse MSE
invets <- 1/MSEets
invima <- 1/MSEima
invarm <- 1/MSEarm
invsum = invets+invima+invarm

#The weights
wets <- invets/invsum
wima <- invima/invsum
warm <- invarm/invsum

# Individual point forecasts

```

```

fitets <- fets$mean
fitima <- fima$mean
fitarm <- farm$mean

#The combined point forecast
fcast <- wets*fitets+wima*fitima+warm*fitarm           #weighted forecasts
fcast

fitcombo <- wets*fets$fitted+wima*fima$fitted+warm*farm$fitted   #weighted fitted values
fitcombo

#Constructing 80% prediction interval
etslo80<-fets$lower[,1]                                     #lower limit 80%
imalo80<-fima$lower[,1]
armlo80<-farm$lower[,1]
lo80<- wets*etslo80+wima*imalo80+warm*armlo80
low80<-ts(lo80, start=2013+3/12, frequency=12)

etshi80<-fets$upper[,1]                                     #upper limit 95%
imahi80<-fima$upper[,1]
armhi80<-farm$upper[,1]
hi80<- wets*etshi80+wima*imahi80+warm*armhi80
high80<-ts(hi80,start=2013+3/12, frequency=12)

###Constructing 95% prediction interval
etslo95<-fets$lower[,2]                                     #lower limit 95%
imalo95<-fima$lower[,2]
armlo95<-farm$lower[,2]
lo95<- wets*etslo95+wima*imalo95+warm*armlo95
low95<-ts(lo95, start=2013+3/12, frequency=12)

etshi95<-fets$upper[,2]                                     #upper limit 95%
imahi95<-fima$upper[,2]
armhi95<-farm$upper[,2]
hi95<- wets*etshi95+wima*imahi95+warm*armhi95
high95<-ts(hi95,start=2013+3/12, frequency=12)

#Plot of combined forecast with 80% and 95% prediction intervals
plot(data, lwd="2",xlab="", ylab="m2 prices (in NOK1000)", main="Combined forecast",
      xlim=c(2011,2014+4/12), ylim=c(25,37))
lines(fcast, col="blue", lwd="2")
lines(fitcombo, col="blue", lwd="1")
lines(low80, col="grey", lwd="1")
lines(high80, col="grey", lwd="1")
lines(low95, col="red", lwd="1")
lines(high95, col="red", lwd="1")

#In-sample accuracy, weighted forecast
fitcombo <- wets*fets$fitted+wima*fima$fitted+warm*farm$fitted
comboerror <- fitcombo-data

inRMSEcombo <- mean(comboerror^2)^0.5
inMASEcombo <- sum(abs(comboerror))/((135/134)*sum(abs(data-lag(data))))

#In-sample RMSE, weighted forecast
#In-sample MASE, weighted forecast

#Out-of-sample accuracy, weighted forecast
etscast1 <- forecast(newets1,h=12)
etscast2 <- forecast(newets2,h=12)
etscast3 <- forecast(newets3,h=12)
etscast4 <- forecast(newets4,h=12)
etscast5 <- forecast(newets5,h=12)
etscast6 <- forecast(newets6,h=12)
etscast7 <- forecast(newets7,h=12)
etscast8 <- forecast(newets8,h=12)
etscast9 <- forecast(newets9,h=12)
etscast10 <- forecast(newets10,h=12)
etscast11 <- forecast(newets11,h=12)
etscast12 <- forecast(newets12,h=12)

imacast1 <- forecast(newima1,h=12)
imacast2 <- forecast(newima2,h=12)
imacast3 <- forecast(newima3,h=12)
imacast4 <- forecast(newima4,h=12)
imacast5 <- forecast(newima5,h=12)
imacast6 <- forecast(newima6,h=12)
imacast7 <- forecast(newima7,h=12)

```

```

imacast8 <- forecast(newwima8,h=12)
imacast9 <- forecast(newwima9,h=12)
imacast10 <- forecast(newwima10,h=12)
imacast11 <- forecast(newwima11,h=12)
imacast12 <- forecast(newwima12,h=12)

```

```

armcast1 <- forecast(newwarm1,h=12)
armcast2 <- forecast(newwarm2,h=12)
armcast3 <- forecast(newwarm3,h=12)
armcast4 <- forecast(newwarm4,h=12)
armcast5 <- forecast(newwarm5,h=12)
armcast6 <- forecast(newwarm6,h=12)
armcast7 <- forecast(newwarm7,h=12)
armcast8 <- forecast(newwarm8,h=12)
armcast9 <- forecast(newwarm9,h=12)
armcast10 <- forecast(newwarm10,h=12)
armcast11 <- forecast(newwarm11,h=12)
armcast12 <- forecast(newwarm12,h=12)

```

```

fit1 <- wets*etscast1$mean+wima*imacast1$mean+warm*armcast1$mean
fit2 <- wets*etscast2$mean+wima*imacast2$mean+warm*armcast2$mean
fit3 <- wets*etscast3$mean+wima*imacast3$mean+warm*armcast3$mean
fit4 <- wets*etscast4$mean+wima*imacast4$mean+warm*armcast4$mean
fit5 <- wets*etscast5$mean+wima*imacast5$mean+warm*armcast5$mean
fit6 <- wets*etscast6$mean+wima*imacast6$mean+warm*armcast6$mean
fit7 <- wets*etscast7$mean+wima*imacast7$mean+warm*armcast7$mean
fit8 <- wets*etscast8$mean+wima*imacast8$mean+warm*armcast8$mean
fit9 <- wets*etscast9$mean+wima*imacast9$mean+warm*armcast9$mean
fit10 <- wets*etscast10$mean+wima*imacast10$mean+warm*armcast10$mean
fit11 <- wets*etscast11$mean+wima*imacast11$mean+warm*armcast11$mean
fit12 <- wets*etscast12$mean+wima*imacast12$mean+warm*armcast12$mean

```

#Weighted forecasts in the test set

```

err1 <- sum((fit1-window(data, start=2011+3/12,end=2012+2/12))^2)/12
err2 <- sum((fit2-window(data, start=2011+4/12,end=2012+3/12))^2)/12
err3 <- sum((fit3-window(data, start=2011+5/12,end=2012+4/12))^2)/12
err4 <- sum((fit4-window(data, start=2011+6/12,end=2012+5/12))^2)/12
err5 <- sum((fit5-window(data, start=2011+7/12,end=2012+6/12))^2)/12
err6 <- sum((fit6-window(data, start=2011+8/12,end=2012+7/12))^2)/12
err7 <- sum((fit7-window(data, start=2011+9/12,end=2012+8/12))^2)/12
err8 <- sum((fit8-window(data, start=2011+10/12,end=2012+9/12))^2)/12
err9 <- sum((fit9-window(data, start=2011+11/12,end=2012+10/12))^2)/12
err10 <- sum((fit10-window(data, start=2011+12/12,end=2012+11/12))^2)/12
err11 <- sum((fit11-window(data, start=2012+1/12,end=2012+12/12))^2)/12
err12 <- sum((fit12-window(data, start=2012+2/12,end=2013+1/12))^2)/12

```

#RMSE calculation, weighted forecast

```

RMSEcombo <- (err1^0.5+err2^0.5+err3^0.5+err4^0.5+err5^0.5+err6^0.5+
err7^0.5+err8^0.5+err9^0.5+err10^0.5+err11^0.5+err12^0.5)/12

```

```

abserr1 <- sum(abs(fit1-window(data, start=2011+3/12,end=2012+2/12)))
abserr2 <- sum(abs(fit2-window(data, start=2011+4/12,end=2012+3/12)))
abserr3 <- sum(abs(fit3-window(data, start=2011+5/12,end=2012+4/12)))
abserr4 <- sum(abs(fit4-window(data, start=2011+6/12,end=2012+5/12)))
abserr5 <- sum(abs(fit5-window(data, start=2011+7/12,end=2012+6/12)))
abserr6 <- sum(abs(fit6-window(data, start=2011+8/12,end=2012+7/12)))
abserr7 <- sum(abs(fit7-window(data, start=2011+9/12,end=2012+8/12)))
abserr8 <- sum(abs(fit8-window(data, start=2011+10/12,end=2012+9/12)))
abserr9 <- sum(abs(fit9-window(data, start=2011+11/12,end=2012+10/12)))
abserr10 <- sum(abs(fit10-window(data, start=2011+12/12,end=2012+11/12)))
abserr11 <- sum(abs(fit11-window(data, start=2012+1/12,end=2012+12/12)))
abserr12 <- sum(abs(fit12-window(data, start=2012+2/12,end=2013+1/12)))

```

#MASE calculation, weighted forecast

```

MAS1 <- abserr1/((111/(111-1))*sum(abs(diff(window(data, start=2002+0/12, end=2011+2/12))))))
MAS2 <- abserr2/((111/(111-1))*sum(abs(diff(window(data, start=2002+1/12, end=2011+3/12))))))
MAS3 <- abserr3/((111/(111-1))*sum(abs(diff(window(data, start=2002+2/12, end=2011+4/12))))))
MAS4 <- abserr4/((111/(111-1))*sum(abs(diff(window(data, start=2002+3/12, end=2011+5/12))))))
MAS5 <- abserr5/((111/(111-1))*sum(abs(diff(window(data, start=2002+4/12, end=2011+6/12))))))
MAS6 <- abserr6/((111/(111-1))*sum(abs(diff(window(data, start=2002+5/12, end=2011+7/12))))))
MAS7 <- abserr7/((111/(111-1))*sum(abs(diff(window(data, start=2002+6/12, end=2011+8/12))))))
MAS8 <- abserr8/((111/(111-1))*sum(abs(diff(window(data, start=2002+7/12, end=2011+9/12))))))
MAS9 <- abserr9/((111/(111-1))*sum(abs(diff(window(data, start=2002+8/12, end=2011+10/12))))))
MAS10 <- abserr10/((111/(111-1))*sum(abs(diff(window(data, start=2002+9/12, end=2011+11/12))))))
MAS11 <- abserr11/((111/(111-1))*sum(abs(diff(window(data, start=2002+10/12, end=2011+12/12))))))
MAS12 <- abserr12/((111/(111-1))*sum(abs(diff(window(data, start=2002+11/12, end=2012+1/12))))))

```

```

MASEcombo <- (MAS1+MAS2+MAS3+MAS4+MAS5+MAS6+MAS7+MAS8+MAS9+MAS10+MAS11+MAS12)/12

```