Journal of Statistical and Econometric Methods, Vol.9, No.1, 2020, 27-45
ISSN: 1792-6602 (print), 1792-6939 (online)
Scientific Press International Limited, 2020

Forecasting Electricity Prices Using Ensemble Kalman Filter

Emmanuel Kipchumba Korir[1], Jane Aduda[2] and Thomas Mageto[3]

Abstract

Forecasting is a key feature in the analysis of statistical data. It entails estimating the current state using observations and the prior knowledge of the state. The nonlinear and non stationary nature of the electricity data calls for a better forecasting method. In this study the States space models under Kalman filter method and Ensemble Kalman filter are used. This is a variance minimizing where for each step it minimizes the variance of the estimation errors resulting to an optimal estimate. Having an ensemble forecast is of interest in the study to check how effective it is compared to a single forecast. This paper also gives the mathematical approach for each method. Electricity price data for a five year period obtained from Nordpool (United Kingdom set) was used.

Keywords: Forecasting; States space model; Kalman filter (KL); Ensemble Kalman filter (EnKF); Ensembles

[1] Pan African University Institute for Basic Sciences, Technology and Innovation. E-mail: sekoinn@gmail.com

[2] Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology. E-mail: jaduda@jkuat.ac.ke

[3] Jomo Kenyatta University of Agriculture and Technology. E-mail: tttmageto@gmail.com

Article Info: *Received* : August 26, 2019. *Revised* : September 12, 2019.
Published online : January 5, 2020

1 Introduction

Energy is a conservative quantity which must be transferred to an object for it to produce heat and perform functions. It's a day-to-day commodity that is used almost everywhere. The production, conversion, measurement and usage all have to be analyzed in such a way that it best suits the customers' demands. The growth of energy markets has resulted to it trading globally. It's market is increasingly becoming interdependent and expanded at a faster rate, this is for both developing and developed markets [1].

Electricity pricing is governed by complex dynamics driven by factors such as, day-to-day and seasonal variation in demand, seasonal variation in temperature, availability of electricity from neighboring regions, and cascade effects when plants are shut down [4]. This commodity is produced when actually demanded since it cannot be stored and must be consumed immediately. This non-storability property and the presence of a wide scope of demand and supply calls for analysis of how prices relates with the underlying drivers [5]. A better pricing mechanism has been studied through a number of structural models.

Modeling of the prices has been an interesting field for most industries and academicians. Modeling causes of commodity prices including market tightness, demand flexibility, weather in the spot markets has been focused. The hourly time series data of electricity spot prices differ from time series of equities and other commodities since of its non-non-non-storability feature [9]. These prices are usually considered multivariate time series because the spot prices are simultaneously set a day before for each 24 intra-day [6]. They have focused mainly in hedging effectiveness, cash flow at risk analysis and volatility forecasting [11]. Several models can be used to test for volatility, seasonality, price jumps and conditional heteroscedasticity.

Forecasting has a huge application in time series analysis and other fields too. A forecasting model that is more accurate is required to facilitate all of the stakeholders, where the future electricity prices is a major factor to business of to the wholesalers, traders and retailers [10]. In their paper they stated that forecasting electricity price problem is related to electricity load forecasting yet distinct. The nature of electricity prices has always lead to generation of different models. The use of multi-factor forecasting methods has not been effective to accurately forecast prices leading to lots of researchers making use of time series forecasting methods.

To solve the above-mentioned problems,the introduction of state space models has provided a better approach in time series analysis. The Kalman filter updates the state variable often when there is a new data point. The Kalman filter however is limited to linear systems. To deal with non linearity of the data, Extended Kalman filter was introduced. This however has a high computational cost since it converts the non-linear state to linear. [12] states that these models can be updated recursively and their flexibility to model nonlinearity is obtained by using ensemble Kalman filter as the algorithm for estimating the state space models.The paper states that ensemble approximation techniques reduces the computational cost significantly and this makes it possible to assimilate data into the system that are too large for previous methods.

The estimates from Ensemble Kalman filter was found to give improved forecast of the day ahead ozone concentration maxima [2]. It is seen as an improved method of Kalman filter, where distribution of the state is expressed by a sample or "ensemble" from the distribution. Forward calculation of the ensemble and its updates obtained over a given period is observed. Ensemble representation gives a dimension reduction technique, leading to fast computational systems with high-dimensional systems. The ensemble size is selected that which will statistically represent the whole model.

This study aims at determining an optimal ensemble size given a data set. This size will be obtained at a size having the lowest root mean squared error. The ensemble size will be used in the Ensemble Kalman filter method. Secondly we discuss the mathematical background of the states space models and lastly see how well the models predicts.

2 Methodology

2.1 Ensembles

Definition 2.1. *An ensemble is an idealization of a large number of virtual copies of a system, taken as a whole, with each representing a possibility of* *state at which the system might be. It is a probability distribution for the state of the system [8].*

We denote the ensemble by

$$x_t = (x_t^1, x_t^2, ..., x_t^k) \tag{1}$$

which represents the forecast ensemble members at time t. At each particular time the electricity prices, x produces k ensemble forecasts. These k forecasts represents the different realizations of the prices at that particular time. This can be illustrated as below

Time	Observations	for k = 2	k = 5	k = k
1	x_1	x_1^1,x_1^2	$x_1^1, x_1^2, x_1^3, x_1^4, x_1^5$	$x_1^1, x_1^2, x_1^3, ..., x_1^k$
2	x_2	x_2^1, x_2^2	$x_2^1, x_2^2, x_2^3, x_2^4, x_2^5$	$x_2^1, x_2^2, x_2^3, ..., x_2^k$
3	x_3	x_3^1, x_3^2	$x_3^1, x_3^2, x_3^3, x_3^4, x_3^5$	$x_3^1, x_3^2, x_3^3, ..., x_3^k$
...				
$t-1$	x_{t-1}	x_{t-1}^1, x_{t-1}^2	$x_{t-1}^1, x_{t-1}^2, x_{t-1}^3, x_{t-1}^4, x_{t-1}^5$	$x_{t-1}^1, x_{t-1}^2, x_{t-1}^3, ..., x_{t-1}^k$
t	x_t	x_t^1, x_t^2	$x_t^1, x_t^2, x_t^3, x_t^4, x_t^5$	$x_t^1, x_t^2, x_t^3, ..., x_t^k$

The ensemble mean and covariance is defined by

$$\bar{x}_t = \frac{1}{k}\sum_{i=1}^{k} x_i \tag{2}$$

$$\Sigma_t = \frac{1}{k-1}\sum_{i=1}^{k} (x_i - \bar{x})(x_i - \bar{x})^T \tag{3}$$

Σ_t is summed over $k-1$ to ensure that Σ_t is an unbiased estimate of the covariance. We take the mean of the ensemble as optimal guess estimate and the error variance from the spread. Ensemble members smoothness determines its covariance [2]. The performance of EnKF relates to the size used in the analysis [2].

Results of the EnKF analysis are influenced by the size. The smaller the size the forecasts is affected through the underestimation of the error covariance. We calculate the root mean squared error of different ensemble sizes and select the one with the least value. The root mean squared error is given by

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{N}\left(X_i^k - X_i^{obs}\right)^2} \tag{4}$$

where, X_i^{obs} represents the i^{th} observation variable, X_i^k corresponds to the ensemble set of size k of the i^{th} observation and N is the total observations.

2.2 State Space

State space models provide a wide view of data assimilation and handling of missing values [13]. This model is expressed as:

$$y_t = x_t + e_t, \qquad e_t \sim N(0, R_t) \tag{5}$$

$$x_t = x_{t-1} + \eta_t, \qquad \eta_t \sim N(0, Q_t) \tag{6}$$

Where $\{e_t\}$and $\{\eta_t\}$ are independent white noise . x_t has initial state x_0 and is random walk and y_t is an observed form of x_t with noise e_t. x_t is the trend or the state of the system at time t and y_t as the observations with noise e_t. Let $x_{t|t} = E(x_t|F_t)$, as the conditional mean and $\Sigma_{t|t} = Var(x_t|F_t)$ the variance of x_t given F_t, where F_t is information available upto time t (inclusive). $y_{t|t}$ represents the conditional mean of y_t with F_t. The 1-step ahead forecast error and variance is given as $s_t = y_t - y_{t|t-1}$ and $S_t = Var(s_t|F_{t-1})$ respectively.

$$\begin{aligned} E(s_t) &= E[E(s_t|F_{t-1})] \\ &= E[E((y_t - y_{t|t-1})|F_{t-1})] \\ &= E[y_{t|t-1} - y_{t|t-1}] = 0 \end{aligned}$$

$$\begin{aligned} cov(s_t, y_t) &= E(s_t, y_t) \\ &= E[E(s_t, y_t|F_{t-1})] \\ &= E[y_t E(s_t|F_{t-1})] = 0 \end{aligned}$$

This shows the independence (they are uncorrelated) between the forecast errors and the observations.

2.3 Kalman Filter

The Kalman filter was first presented in 1960. It elaborates a recursive solution of the linear filtering discrete data problem [7]. The state and observation equations are given as

$$x_t = M_t x_{t-1} + \eta_t \tag{7}$$

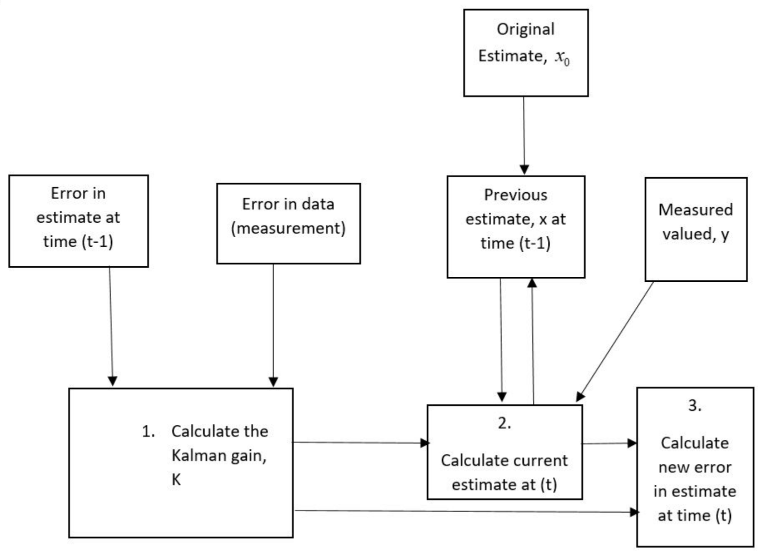

Figure 1: Flow Chart of the Kalman filter Process

where M_t is a linear operator relating the previous time state at step x_{t-1} to the present step x_t and η_t is the process noise.

$$y_t = H_t x_t + e_t \tag{8}$$

$H = (1, 0, ..., 0)_{1\times m}$, is a linear observation operator that provides a mapping from model space x_t to observation space y_t.

Assumptions of the filter

1. The state and the observation process are linear
2. The state noise η_t and observation noise e_t are sequences of independent Gaussian white noise with zero mean.
3. x_t, η_t and e_t are uncorrelated for $t > 0$. This means
$E[\eta_t \eta_j^T] = Q_t$ for $t = j$ and 0 for $t \neq j$
$E[e_t e_j^T] = R_t$ for $t = j$ and 0 for $t \neq j$
$E[\eta_t e_j] = 0$ for all t and j
4. x_0 which is given or follows a known distribution.

The Figure 1 shows the steps taken under the Kalman filter modeling.

The labels 1,2 and 3 are the major process in the calculation. The steps are elaborated below.

1. $Kalmangain = \frac{Estimateerror}{Estimateerror+measurementerror}$
2. $Esimate_t = Estimate_{t-1} + K[Measurement(y) - Estimate_{t-1}]$
3. $Errorestimate_t = [I - K](Errorestimate_{t-1})$

Where K is the Kalman gain and, $0 \leq K \geq 1$. Once we have the new (current) estimate and current error they become the previous estimate and error respectively in the next iteration.

Let $x_t^f \in \Re^n$ as the prior state estimate at t with given information to period t and x_t^a as the posterior estimate at period t with measurements y_t. During forecasting, suppose the original forecast time is t and we want to predict x_{t+j} for $j = 1, ..., h$, where $h > 0$. The $j-$step ahead forecast $x_t(j) = E(x_{t+j}|F_t)$, where F_t is the past information.
A 1-step forecast is obtained as

$$x_t(1) = E(x_{t+1}|F_t) \tag{9}$$

The forecast error from above is $s_t(1) = x_{t+1} - x_t(1)$
Each iteration of the KF is started with a prior estimate x_t^f, which is the expected value of the state just before assimilating the measurement.
The state error for the prior and posterior estimates are obtained as

$$w_t^f = x_t - x_t^f \qquad\qquad w_t^a = x_t - x_t^a \tag{10}$$

The corresponding prior and posterior estimate error covariances are given as

$$\Sigma_t^f = E[w_t^f (w_t^f)^T]$$
$$\Sigma_t^a = E[w_t^a (w_t^a)^T]$$

The estimation error is assumed having a mean zero with covariance matrix Σ_t. The prior state estimate from equation 7 can be expressed as

$$\begin{aligned} x_t^f &= E[x_t|F_{t-1}] \\ &= E(M_t x_{t-1} + \eta_t|F_{t-1}) \\ &= E(M_t x_{t-1}|F_{t-1}) + E(\eta_t|F_{t-1}) \\ &= M_t x_{t-1}^a \end{aligned}$$

The corresponding error covariance matrix is

$$\begin{aligned}\Sigma_t^f &= E[w_t^f(w_t^f)^T] \\ &= E[(M_t w_{t-1}^a + \eta_t)(M_t w_{t-1}^a + \eta_t)^T] \\ &= M_t \Sigma_{t-1}^a M_t^T + Q_t\end{aligned}$$

In the derivation of the KF equations, we derive an equation that calculates the posterior estimate, x_t^a which is as a linear combination of a previous estimate x_t^f and the difference between actual and a predicted measurement [3]. The recursive KF algorithm is demostrated in equation 11.

$$x_t^a = x_t^f + K_t(y_t - H_t x_t^f) \tag{11}$$

which is the posterior estimate x_t^a, where K_t is the Kalman gain which updates the mean state given observations. The difference $(y_t - H_t x_t^f)$ is refereed to as the measurement shock. The Kalman gain explains the effect of the new shock to the state variable x_t.

From equation 11 the posterior estimate error is given by

$$\begin{aligned}w_t^a &= x_t - x_t^a \\ &= x_t - x_t^f - k_t(y_t - H_t x_t^f) \\ &= x_t - x_t^f - k_t(H_t x_t + e_t - H_t x_t^f) \\ &= (I - k_t H_t)(x_t - x_t^f) - k_t e_t\end{aligned} \tag{12}$$

where I is an $(n \times n)$ identity matrix.
The error covariance matrix in the posterior estimate is obtained as

$$\begin{aligned}E[w_t^a(w_t^a)^T] &= E[[(I - K_t H_t)(x_t - x_t^f) - K_t e_t][(I - K_t H_t)(x_t - x_t^f) - K_t e_t]^T] \\ &= E[[(I - K_t H_t)(x_t - x_t^f) - K_t e_t][(I - K_t H_t)^T(x_t - x_t^f)^T - K_t^T e_t^T]] \\ &= E[(I - K_t H_t)(x_t - x_t^f)(I - K_t H_t)^T(x_t - x_t^f)^T] - K_t e_t K_t^T e_t^T \\ &= (I - K_t H_t)E[(x_t - x_t^f)(x_t - x_t^f)^T](I - K_t H_t)^T - K_t E[e_t e_t^T]K_t^T \\ \Sigma_t^a &= (I - K_t H_t)\Sigma_t^f(I - K_t H_t)^T + K_t R_t K_t^T\end{aligned} \tag{13}$$

We then minimize the sum of the variances equation with respect to K_t. This is achieved by minimizing the trace of the covariance matrix. The least square method is used to achieve this.

$$min(\Sigma_t^a)^T = [(I - K_t H_t)\Sigma_t^f(I - K_t H_t)^T + K_t R_t K_t^T]^T$$

$$\frac{\partial(\sum_t^a)^T}{\partial K_t} = 2(I - K_tH_t)\Sigma_t^f(-H_t^T) + 2K_tR_t$$
$$0 = -\Sigma_t^f H_t^T + K_tH_t\Sigma_t^f H_t^T + K_tR_t$$
$$\Sigma_t^f H_t^T = K_t(H_t\Sigma_t^f H_t^T + R_t)$$

This process gives the Kalman gain equation to be

$$K_t = \frac{\Sigma_t H_t}{H_t\Sigma_t H_t^T + R_t} \tag{14}$$

This means that as R_t tends to zero, the measurement y_t gives a "big trust" while the predicted estimate $H_t x_t^f$ is less trusted. Similarly if Σ_t^f approaches the actual value y_t is less trusted while the estimates $H_t x_t^f$ are trusted more. The error covariance is then updated given new observations by

$$\begin{aligned}\Sigma_t^a &= (I - K_tH_t)\Sigma_t^f(I - K_tH_t)^T + K_tR_tK_t^T \\ &= \Sigma_t - \Sigma_t H_t^T K_t^T - K_tH_t\Sigma_t + K_t^T H_t^T \Sigma_t \\ &= (I - K_tH_t)\Sigma_t \end{aligned} \tag{15}$$

Hence we have,

$$\Sigma_{t+1} = (I - K_tH_t)\Sigma_t \tag{16}$$

The Kalman filter works under assumption of a linear evolution model and observation operator.

2.4 Ensemble Kalman Filter

The method was first studied by Evensen (1994) based on Monte Carlo simulations that makes it possible to use Kalman filter method on nonlinear models. A single state estimate in KL is used to give an analysis while for EnKF the ensemble is used. This step is applied k times resulting to generation of k-member ensemble of state vector estimates. This k ensemble members are taken as a sample drawn from the prior distribution at that time.

The ensemble mean and its covariance matrix used to compute the Kalman gain. The ensemble representation gives a dimension reduction, resulting to easy computation for higher dimensional systems. It gives optimal estimates on parameters of interest from indirect, inaccurate and uncertain observation.

Thus from Equation 2 and 3, we have the best estimate and the error covariance.

The update equation follows where each ensemble member is updated according to the updating equation 11 and calculate

$$x_{t+1}^{(i)} = x_t^{(i)} + k_t \left(y_t - H_t x_t^i \right) \tag{17}$$

where k_t is the Kalman gain given by,

$$k_t = \Sigma_t H_t \left(H_t \Sigma_t H_t^T + R_t \right)^{-1} \tag{18}$$

Each ensemble member according to equation 17 gets to be updated where from this the updated state vectors and error covariance matrix are estimated. This is a recursive process by first setting initial estimates for x_0^a and Σ_0^a and then generating a set of ensemble members.

2.5 Model performance

The measure of forecast accuracy is measured in terms of Mean Absolute Percentage Error(MAPE) and the Root Mean Squared Error (RMSE). They help to validate the methods used on the variability of their accuracy. They are defined by the following equations:

$$MAPE = 100 \times \frac{1}{N} \sum_{i=1}^{N} \frac{\mid x_i - \hat{x}_i \mid}{x_i} \tag{19}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (e_i - \bar{e})^2} \tag{20}$$

where $e_i = \frac{\hat{x}_i - x_i}{\bar{x}}$ and x_i is the actual price, $\hat{x_i}$ the predicted price, $\bar{x}$ is the value of mean price at that time and N is the size of the predicted time.

3 Results

3.1 Ensemble size determination

The model was fit to some simulated data sets. For each observed data set, copies from the same data are simulated with respect to different ensemble size. This is done for each data set in the column. Ensemble sizes of 20, 50, 70, 100, 140, 180, 200, 230, 250 and 300 were tested. For example, at an ensemble size of 20, we have 20 data sets for each observed value generated from them. 10 observations were used for this process. Each line on the graph represents an observation.

The Figure 2 shows the different values of RMSE when the generated set of ensembles follow a standard normal distribution.

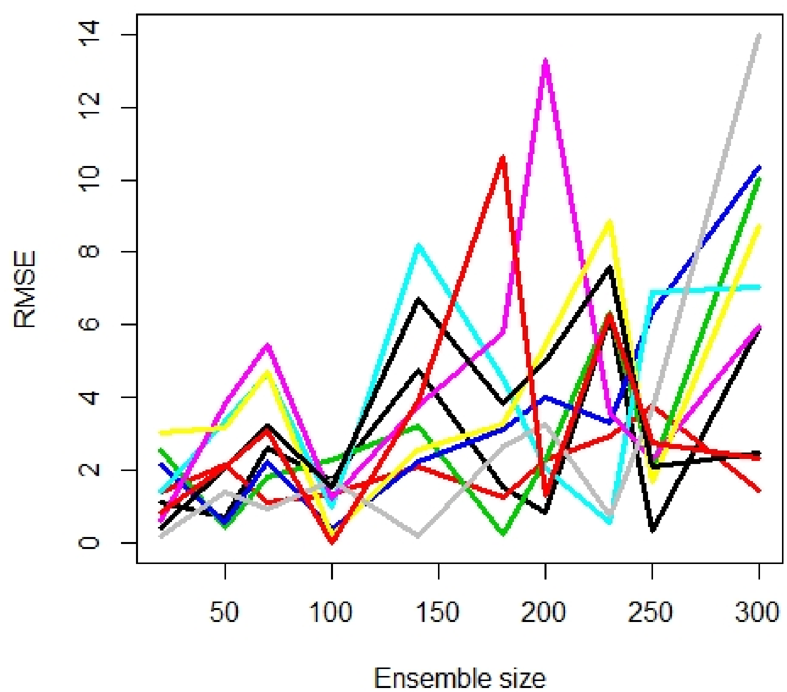

Figure 2: The RMSE values for different ensemble sizes with data N(0,1)

The corresponding RMSE at each ensemble size is

Ensemble size	20	30	50	70	100	140	180	200	250	300
RMSE	1.369	1.966	2.980	1.138	3.747	3.687	3.976	4.647	3.195	6.822

From the table we can see that at ensemble size 70, the RMSE is 1.1384 which is the smallest. Given the above ensemble sizes we select 70 as our optimal size.

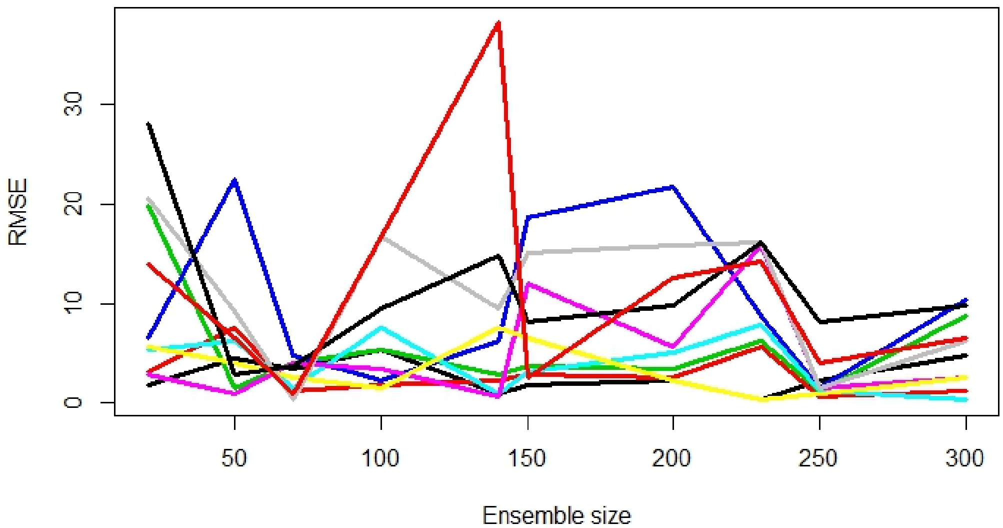

Figure 3: The RMSE values for different ensemble sizes with data, $\lambda = 5$

The same can be done for any distribution of data. The second case involves a poison distribution with $\lambda = 5$. From Figure 3 we conclude at size 70 we have the least value of RMSE hence the optimal size. It means that if we take a sample of 70 for each observation with this distribution and do the prediction, there is a high chance of obtaining a best predicted value with a low variation.

The following Tables 1 and 2 gives a summary for distributions.

Table 1: Optimal selection of the ensemble size from different normally distributed data sets

Distribution	Ensemble size with the least RMSE
N(0,1)	70
N(0,100)	70
N(40,1)	140
N(35,5)	140
N(10,10)	180
N(100,2)	250
N(200,200)	50

Table 2: Optimal selection of the ensemble size from different data sets following a poison distribution

Distribution	Ensemble size with the least RMSE
$\lambda = 1$	100
2	100
3	70
5	70
7	180
99	50
200	100

The theory behind ensembles can be applied in any distribution. Ensemble approach has a greater advantage in minimizing forecast errors than a single forecast [14]. There is always a challenge in having too small and too big ensemble size, the size must be at a point where the RMSE is small. Given a data set one can obtain an ensemble forecast with an optimal ensemble size as obtained above.

3.2 Data Exploration

3.2.1 Data

In this section we explore prices from Nord Pool which is one of the best power market in Europe. It gives day-ahead and intraday market prices to customers. It offers prices for 13 markets (countries), 19 bidding zones and over 300 buyers and sellers place over 2000 orders every day. The data for United Kingdom, N2EX, was used. This data include the prices for the period running from 1^{st} of January 2014 to 31^{st} of December 2018. The analysis was done using R software. The data was obtained from Nordpool website.

With the prices, denoted as p_t, at time t given we calculate daily returns, r_t, as the continuously compounded returns given as

$$r_t = log(p_t) - log(p_{t-1}) = log\left(\frac{p_t}{p_{t-1}}\right) \tag{21}$$

p_t is the present price and p_{t-1} is the previous day's price.

3.2.2 Descriptive Statistics

Table 3 gives the basic statistics of the whole series. This table shows that the series tend to have excess kurtosis, since its value 16.899 > 3 implying the data is not normally distributed.

Table 3: Descriptive statistics for auction prices

Number of observations	1826
Minimum	27.93
Maximum	169.65
1. Quartile	38.705
3. Quartile	50.2
Mean	45.141599
Median	42.745
Sum	82428.56
SE Mean	0.23247
Variance	98.576767
Stdev	9.928583
Skewness	2.313936
Kurtosis	16.895592

3.2.3 Time series plots

The Figure 4 shows that the prices have a trend and are generally not stationary. Figure 5 is a plot of returns which makes the time series data stationary since the returns are close to zero. However there are presence of high spikes and low spikes.

Tables 4 and 5 confirms that the data is stationary (since the $p{-}value < \alpha$) and returns are not normally distributed (since $p - value < \alpha$) for $\alpha = 0.05$.

3.3 Price filter forecast

The filter method gives a forecast of the return data set for a period of 365, which is 20 percent of the entire data set as shown in Figure 6. This

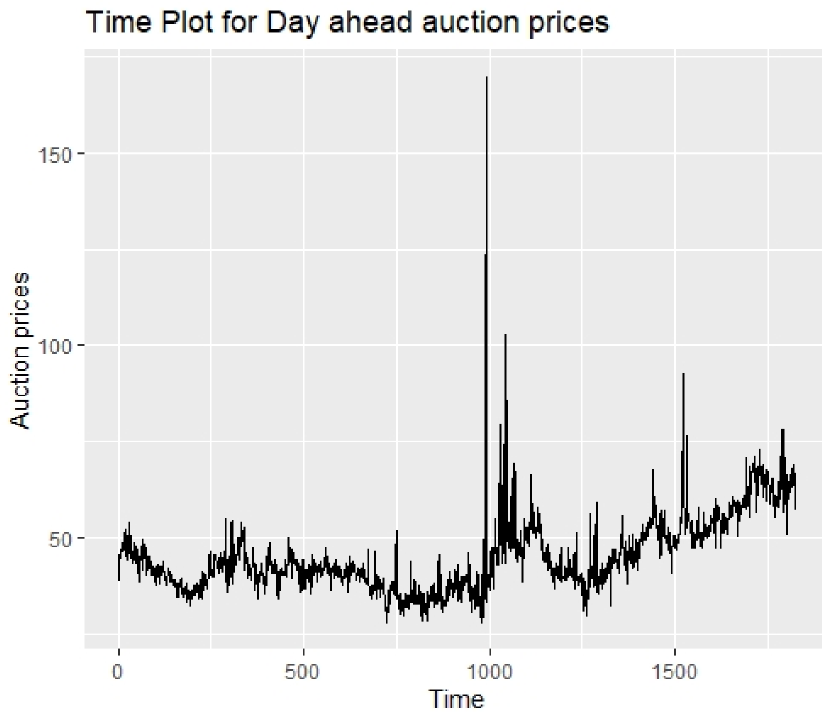

Figure 4: Time series plot for day ahead auction prices

Table 4: Augmented Dickey-Fuller Test

Dickey-Fuller	p-value
-19.181	0.01

Table 5: Shapiro-Wilk normality test

W	p-value
0.70948	2.2e-16

sample was used to check at how the filter returns a one step ahead prediction of states.

From the plot we can see that the filter estimates almost have the same values as the actual prices. This forecast confirms how well it predicts. The Table 6 gives a view of the predicted values versus the actual values

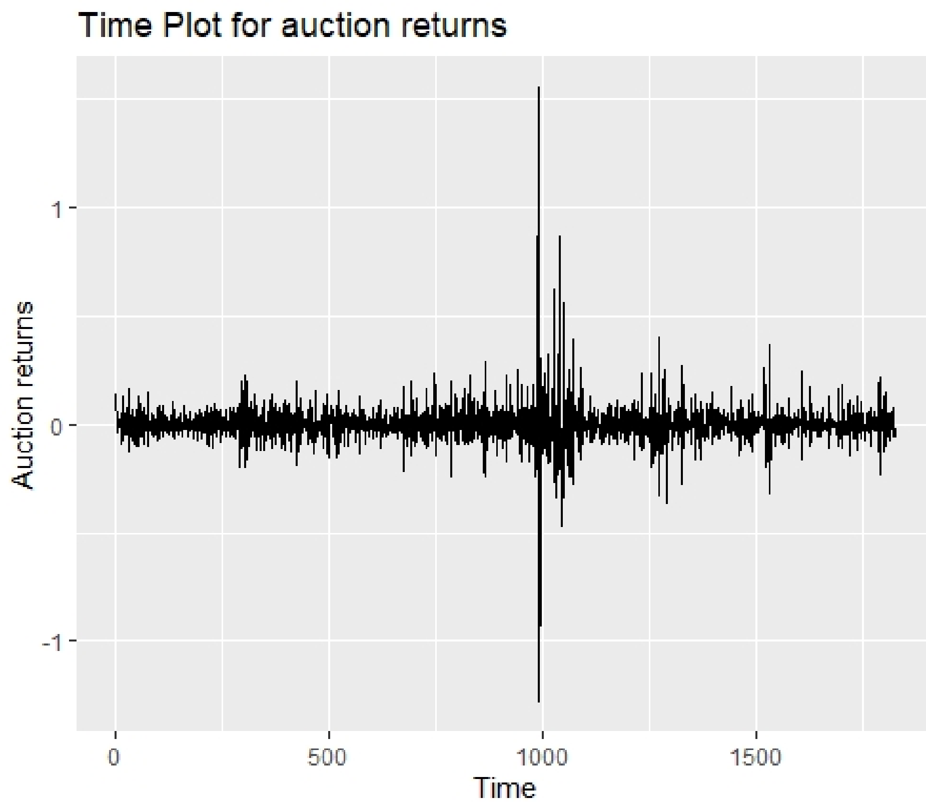

Figure 5: Time series plot for day ahead auction returns

Table 6: Table showing the logprices of the actual and predicted at 80 and 95 confidence intervals

Actual logprices	Predicted logprices	Lo.80	Hi.80	Lo.95	Hi.95
3.8299	3.8299	3.8016	4.0255	3.7424	4.0847
3.9448	3.9065	3.8254	4.0737	3.7597	4.1396
3.9090	3.9081	3.8500	4.1039	3.7828	4.1712
3.97180	3.9475	3.8116	4.0663	3.7441	4.1337
4.0397	4.0045	3.7950	4.0569	3.7257	4.1263
3.91899	3.9517	3.7874	4.055	3.7166	4.1259

3.4 Model Evalution

From the forecast, which is the data from $1^{st} January, 2018$ to $31^{st} December, 2018$ we use MAPE and RMSE to check how well the method performs in forecasting the prices.

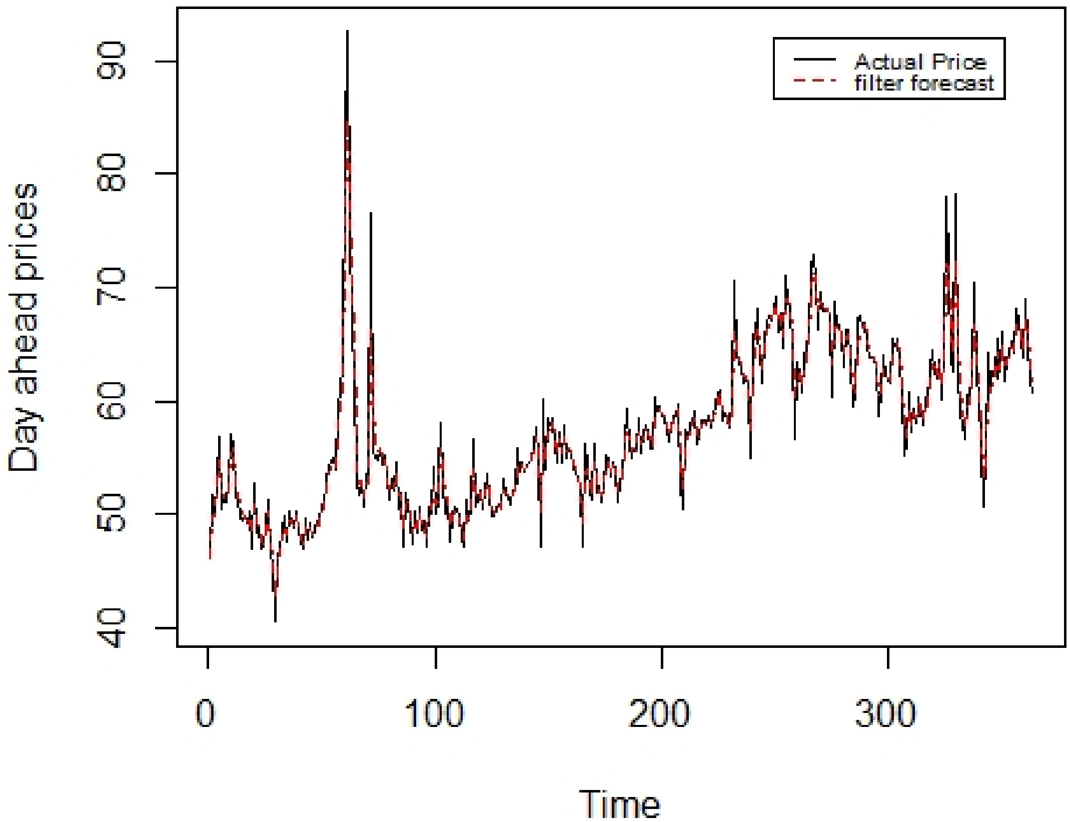

Figure 6: Plot of the filter forecast estimates using 365 data sets

Table 7: Method evaluation using MAPE and RMSE

MAPE	0.005895775
RMSE	0.008985829

The Table 7 gives the MAPE and RMSE output. The values of the RMSE and MAPE are low and this means that the filter predicts the prices well. It also confirms from the literature.

4 Conclusion

The research has shown how state space models performs given a certain data set. The introduction of state space to now the ensemble filters has been helpful in the time series modeling. We discussed the ensemble approach and how to determine an optimal ensemble given different distributions. It was observed that the sizes vary and that the RMSE gives a point at which we select the size. However, the size might not give a perfect forecast but its values have higher degree of accuracy. The mathematical formulation of the models was discussed too. It gives a clear part of how this filters work from their two steps, the forecast state and the update state. Finally the results from the RMSE and MAPE shows that the method gives a better prediction. This study dealt with prices individually, future studies can be done where factors affecting the prices such as weather conditions, fuel prices, etc are incorporated in the model to determine how these will impact on the forecast.

Acknowledgments The appreciation goes to Almighty God for His protection and care and also PAUISTI for their support towards my research. Will like to thank too my supervisors for their full time support throughout my study.

References

[1] Aduda, J.A., Application of multivariate volatility models in estimating optimal dynamic hedge ratios for crack spreads with volatility spillovers in energy markets, 2017.

[2] Anderson, J.L., An ensemble adjustment kalman Filter for data assimilation, *Monthly weather review*, **129**(12), (2001), 2884-2903.

[3] Awashie, G.E., *Pricing Financial Options Using Ensemble Kalman Filter*, PhD thesis, 2012.

[4] Buzoianu, M., Brockwell, A.E., and Seppi, D.J., *A dynamic supplydemand model for electricity prices*. Cerca con Google, (2005).

[5] Carmona, R. and Coulon, M., *A survey of commodity markets and structural models for electricity prices. In Quantitative Energy Finance*, Springer, pages 41-83, 2014.

[6] Huisman, R., Huurman, C., and Mahieu, R., Hourly electricity prices in day-ahead markets, *Energy Economics*, **29**(2), (2007), 240-248.

[7] Kalman, R.E. A new approach to linear Filtering and prediction problems, *Journal of basic Engineering*, **82**(1), (1960), 35-45.

[8] Katzfuss, M., Stroud, J.R., and Wikle, C.K., Understanding the ensemble Kalman filter, *The American Statistician*, **70**(4), (2016), 350-357.

[9] Knittel, C.R. and Roberts, M.R., An empirical examination of restructured electricity prices, *Energy Economics*, **27**(5), (2005), 791-817.

[10] Neupane, B., Woon, W.L. and Aung, Z., Ensemble prediction model with expert selection for electricity price forecasting, *Energies*, **10**(1), (2017), 77.

[11] Spodniak, P. and Bertsch, V., Determinants of power spreads in electricity futures markets: A multinational analysis. *Technical report, ESRI Working Paper*, (2017).

[12] Takeda, H. et al., Electricity load forecasting: Ensemble approach, (2016).

[13] Tsay, R.S., *Analysis of Financial time series*, vol. **543**, John Wiley & Sons, 2005.

[14] Robert Zwanzig, Ensemble method in the theory of irreversibility, *The Journal of Chemical Physics*, **33**(5), (1960), 1338-1341.