Forecasting the electricity consumption of Sweden
using ETS, ARIMA and a LSTM neural network

Final Project in Predictive Analytics

MSc in Data Science
Copenhagen Business School
August 2022

Written by: Erik Konstenius
Professor: Andrea Tafuro

Number of pages: 10
Number of characters: 22,349

Table of Contents

1. Introduction

The energy sector has a crucial role in every nation as the products, whether they are fossil fuels or renewables, are in principle fundamental in a functioning society. Much of the products we extract from the energy sector end up as electricity that powers our homes and businesses. We always need electricity yet how much we need can vary greatly depending on factors such as the population size, household size, living situations (World electricity consumption, 2022) as well as the economic situation, employment rate and weather (Chen, 2017). The main factors that determine the demand of electricity can be broken down into social, economic, technological and policy related (Energy Projections, calculating long term energy demand, 2022).

Although Russian gas only makes up about 1 percent of Swedish energy supply, many speculate that Swedish households risk seeing a doubling in electricity prices in winter as the Russian Gazprom is decreasing its deliveries of gas through the Nord Stream 1 pipeline to a fifth of its full capacity. The extent of the effects of the decreased supply is yet to be fully understood as weather and the political situation in Ukraine are also key factors. A worst-case situation could be a nation-wide blackout as the energy demand surpasses the available supply. Some politicians and spokespeople of the energy sector have emphasised that consumers must decrease the demand in order to reduce the risk of a blackout (Dahl, 2022).

Projecting a society's long-term electricity demand helps determine what capacity is needed for future energy generation. Forecasting is also important for determining the feasibility of energy projects (Energy Projections, calculating long term energy demand, 2022). In present times, forecasting the electricity demand could help policymakers understand the economic effects of Russia's gas cuts and the price development of electricity (Dahl, 2022).

The aim of this paper is to forecast the electricity consumption in Sweden using exponential smoothing, ARIMA and a LSTM neural network.

2. Dataset description

The dataset contains the monthly electricity consumption of Sweden, measured through gigawatt hours (GWh), from all areas of use including both the private and public sector between January 1990 and May 2022. There are in total 389 datapoints meaning that there are no null values. The dataset is considered official statistics of Sweden and has been compiled through a collaboration between Statistics Sweden and the Swedish Energy Agency (Elanvändning, GWh efter användningsområde och månad, 2022).

3. Splitting the data

The data is thereafter split into a train set and a test set. The split is done this early to prevent data leakage in future data explorations and subsequent modelling. The train set includes all observations except the final four years. This means that the test set will include about 12 % of all observations – a bit less than the common rule of thumb of 20 % of the entire dataset. The final forecast will then be performed on the period June 2018 to May 2022.

4. Exploratory data analysis

The first step is to plot the data in order to highlight different features of the time series. These features could be patterns, unusual observations like outliers and changes over time. A good forecast is one that has incorporated many of the features of the time series and reflects the true data generating process (Hyndman & Athanasopoulos, 2014). A set of plots highlighting the features of the time series used in this forecast are shown below (see figure 1).

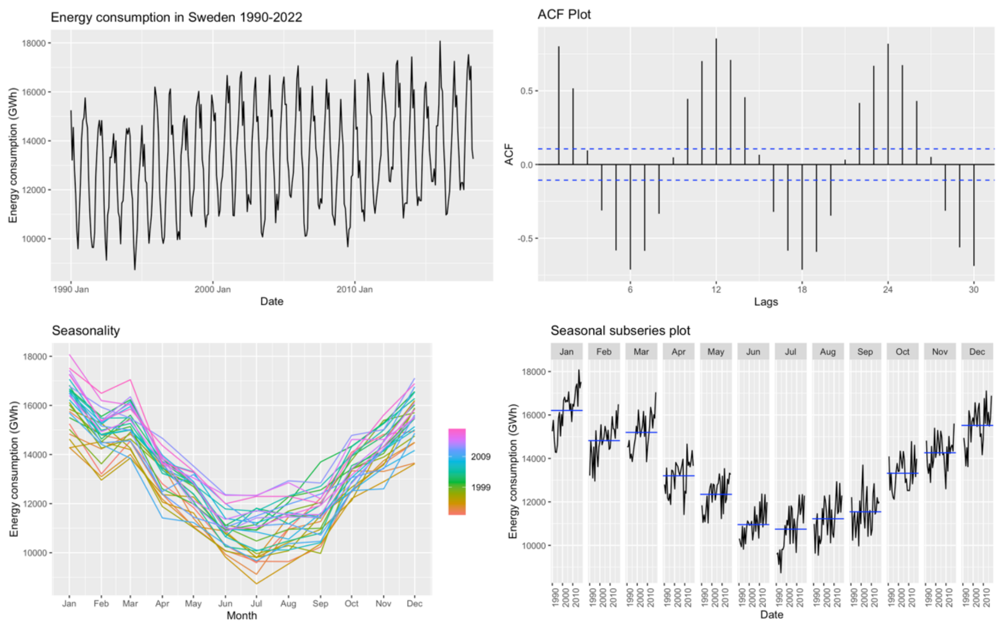


Figure 1. Time plot (top left), ACF plot (top right), Seasonality plot (bottom left), Seasonality subseries plot (bottom right)

The time plot (top left of figure 1) and ACF plot (top right of figure 2) show that the time series has a slight increasing trend with a distinct seasonality where the electricity consumption of one month is heavily correlated with previous observations. An observation is the most positively correlated with the same month one or two years earlier and the most negatively correlated with the months on the other side of the calendar one or two years earlier. A Ljung-Box test returns a p-value smaller than 0.001, thus we can reject the null hypothesis and indicates that the time series contains autocorrelation between lagged values, and thereby that previous observations have a large influence on a particular observation.

The seasonality plot (bottom left of figure 1) reveals that the electricity consumption peaks around November to March and decrease during the other months. January appears to be the month with the highest electricity consumption of about 16,000 GWh and July appears to be the month with the lowest electricity consumption of about 11,000 GWh. Colder months require households to consume more electricity to keep the houses warm. Additionally, the business activities usually drop down during the summer which leads to a decrease of the electricity consumption (Kalle, 2018).

The seasonal subseries plot (bottom right of figure 1) helps the time plot (top left of figure 1) visualize that the electricity consumption has increased slightly over time. It is not entirely clear what the increase in recent years can be attributed to. Perhaps it could be due to the pandemic, colder winters, good years for energy intensive industries or good general economic growth.

There are some cyclical deviations around the long-run trend between the years. The irregular time between each upturn and downturn and that the cycles does not seem to be related to time signals that these are cyclical patterns and not seasonal patterns. These deviations could have been a consequence of the economic environment. The later parts of the 1990s are characterized by the Dot Com-Bubble that had a large effect on both households and businesses in Sweden. The later parts of the 2000s are characterized by the Global Financial Crisis. During recessions household spending tend to decrease because of higher unemployment rates or rapid price increases. The corporate sector can also reduce spending because of higher risk sentiment and lower sales caused by the decreased willingness of households to spend (Rodeck, 2022). Sweden has a large share electricity-intensive industry that is particularly sensitive to the economic environment (Kalle, 2018). It is from the abovementioned factors evident that the energy consumption of Sweden is to some extent affected by the economic environment. There are no major structural breaks in the time series.

5. Data Transformation

It can sometimes be useful to transform the data to simplify the pattern of the series. Transformations are often used when the data shows variation that change with the level of the time series. Some popular transformations are logarithmic transformation and Box-Cox transformation (Hyndman & Athanasopoulos, 2014). The time series in this paper will not be transformed because the variation does not seem to increase or decrease over time and there is no obvious skewness caused by disproportionate high growth rates that would motivate for a transformation like a logarithmic transformation. It is also nice to work with values that have not been transformed when the exploratory data analysis and modelling allow for it.

6. Exponential smoothing

a. Background

Exponential smoothing is a modelling technique that estimates that an observation is a sum of exponentially decaying weighted averages of past observations. Holt-Winters exponential smoothing captures the level, trend, and the seasonal component of the series in in state equations. Additionally, exponential smoothing models also include an error component to capture the remaining noise in the data. The models are able to capture trend, season and error that increase linearly (additive) and exponentially (multiplicative). The components can also be dampened linearly (additive) or exponentially (multiplicative) to produce more conservative forecast over longer periods (Hyndman & Athanasopoulos, 2014).

b. Model creation

The SEATS decomposition reveals that the series displays a weak, if even present, positive linear trend (see figure 2). The seasonality is clearly present, but the variance is not changing throughout the series. It is however difficult to tell if the seasonality is additive or multiplicative in the absent of a trend as the amplitude of seasonality will be constant in both cases (Svetunkov, n.d.). No clear patterns can be detected in the remaining variation. Multiplicative errors are said to be useful when a large degree of heteroscedasticity is present - which does not seem to be the case with this dataset (Hyndman & Athanasopoulos, 2014). With these insights in mind, an ETS(A,N,A), ETS(A,A,A) and ETS(A,Ad,A) will be constructed as well as an automatically constructed model. Three hand-picked models are constructed because the trend is not easily distinguished in the SEATs decomposition. The models are shown below together with the information criterions AIC, AICc and BIC.

Model	AIC	AICc	BIC
ETS(A,N,A)	6331.346	6332.823	6388.824
ETS(A,A,A)	6337.185	6339.080	6402.327
ETS(A,Ad,A)	6338.467	6340.591	6407.441
ETS(A,N,A) (auto)	6331.346	6332.823	6388.824

Figure 2. Comparison of ETS models

The model that had the best performance is the ETS(A,N,A) with additive errors and seasonality but no trend. The remainder of this section will use the ETS(A,N,A) to produce a forecast unless it shows unsatisfactory results – in that case previously mentioned models will also be commented. All models have had their residuals analysed and have been used to create a forecast, but the figures are kept in the appendix for the interested reader. The chosen model is given by the following equation.

$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ where the level equation, ℓ, is given by $\ell = \ell_{t-1} + \alpha\varepsilon_t$ and the season equation, s, is given by $s = s_{t-m} + \gamma\varepsilon_t$ and the parameters α and γ are smoothing parameters. The residuals of the ETS(A,N,A) model are shown below.

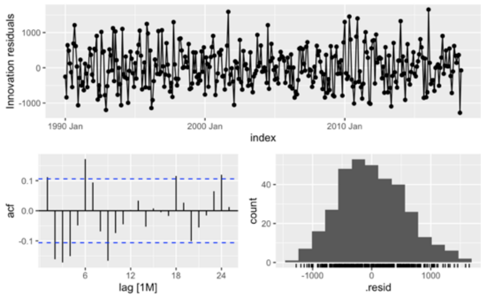


Figure 3. Residuals of ETS(A,N,A)

The residual plots reveal what patterns are left after fitting the model and if the model has been able to capture the key characteristics in the data. If there are patterns left in the residuals, then there are important features that are left out of the model and the uncertainties in the forecasts will increase. Furthermore, the residuals should have a zero mean. If not, forecasts using the model will have a bias, and thus a systematic deviation from the correct value (Hyndman & Athanasopoulos, 2014). The lack of patterns in the residuals indicate that the model appears to be able to capture the key characteristics in the series. The apparent constant variance indicate that the residuals are homoscedastic. The residuals also appear to be normally distributed implying that inference, and in particular the confidence intervals, can be reliable. Results from the Ljung-Box test and the Shapiro-Wilk test show that the residuals are homoscedastic, non-autocorrelated and normally distributed.

c. Forecast using ETS(A,N,A)

It is important to keep in mind that a model that has a good fit to the data is not certain to produce accurate forecasts. A model that fails to produce accurate forecasts can have overfitted the training data or created forecasts during a period that could resemble a structural break. The forecast using the ETS(A,N,A) is shown below.

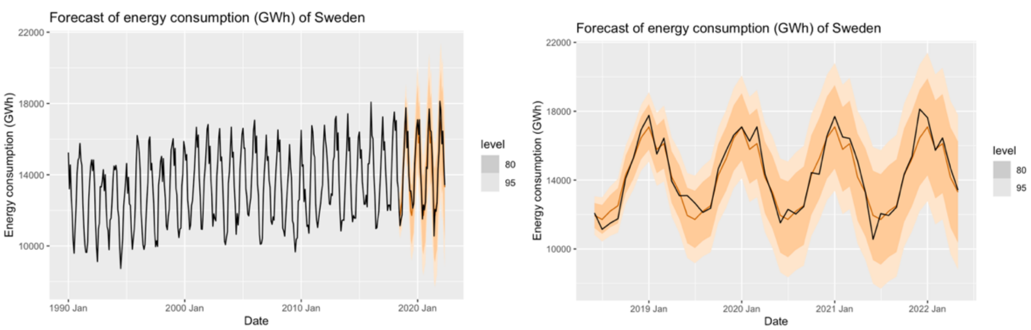


Figure 4. Forecasts with ETS(A,N,A) *on the test set. Left figure shows the entire series including the test set and the right figure shows the forecast only on the test set.*

A summary of the forecasting accuracy of the three different models is shown below.

Model	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
ETS(A,N,A)	143.06	561.73	420.43	0.6818	2.9828	0.5473	0.5949	0.2115
ETS(A,A,A)	354.38	679.41	544.24	2.1583	3.7550	0.7084	0.7196	0.3086
ETS(A,Ad,A)	144.49	562.16	420.68	0.6918	2.9836	0.5476	0.5954	0.2114

Figure 5. Comparison of ETS models

The ETS(A,N,A) model achieves the best performance on all accuracy measures. The difference is small between the ETS(A,N,A) model and the ETS(A,Ad,A) model.

7. ARIMA

a. Background

Autoregressive Integrated Moving Average (ARIMA) models try to model the autocorrelation in the data by an autoregressive and a moving average component. Autoregressive models forecast the value using linear combination of past values of the variable and a moving average model foreacast the value using weighted past forecast errors (Hyndman & Athanasopoulos, 2014). A seasonal ARIMA model can be specified by ARIMA(p,d,q)(P,D,Q)[m], where the lowercase letters relate to the non-seasonal part and the capital letters refer to the seasonal part. The “p” represents the order of the autoregressive part, the “d” represents the degree of first differencing involved and “q “represent the order of the moving average part. The letter “m” denotes the frequency in the seasonality.

b. Stationarity assumption

Many econometric models, including ARIMA, assume that the time series is stationary. A stationary time series is one whose statistical properties do not depend on the time at which the series is observed. Each value is then independent of other values and the overall patterns in the data should remain constant. The persistence of shocks is different in stationary and non-stationary series. While the series reverts to its mean in stationary series, non-stationary time series might exhibit different behaviours after the shock. Models that assume that the series is stationary will produce different forecasts if the data are not stationary. There are however models, such as ETS models, that do not require the data to be stationary in order to produce accurate forecasts. Non-stationarity can be detected visually through a time plot and an ACF plot (Hyndman & Athanasopoulos, 2014). Stationarity can also be tested numerically using unit root tests like the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test and the Augmented Dickey-Fuller (ADF) test. In the KPSS test, we investigate if we can reject the null hypothesis that the series is stationary, whereas the null hypothesis in the ADF test says that there is a unit root in the series and that the series is non-stationary

(Hyndman & Athanasopoulos, 2014). To better understand the data generating process and what is required to make the series stationary, a KPSS test and a ADF test are conducted. Although both tests test similar things, both are used considering the high type 1 error of incorrectly rejecting the null hypothesis of the two tests.

Considering the existence of a weak trend, first step is to conduct the two tests assuming a drift and a trend. The KPSS test returns a t-statistic of 1.0435 which is larger than the critical value at one percent, 0.739, implying that we can reject the null hypothesis and say that the series appears to be non-stationary. From the ADF test, assuming the series follows a random walk with drift and trend, the test statistic that gamma is equal to zero, τ_3, is -3.2302 which is less extreme than the critical value at the five percent significance level (-3.42) implying that we cannot reject the null of the existence of a unit root. The results from the first ADF test show that the series appears to be non-stationary. The test statistic τ_3 that under the null say that a_2 and γ are both zero, that there is a presence of a unit root but no trend, is 5.3304 which is less than the critical value at the five percent significance level (6.30) implying that we cannot reject the null, we can assume that both are equal to zero, and that there is no evidence of a trend in the data. The test which null is that a_0, a_2 and γ are all equal to 0 and that the series displays a unit root without drift and trend, returns a lower value (3.6266) than the critical value at five percent significance level (4.71). We can therefore not reject the null that a_0, a_2 and γ are all equal to 0. The lack of evidence for a trend indicates that it is motivated to perform a ADF test that assumes the series follows a random walk with drift but no trend. The results from the ADF test with drift but no trend shows that we cannot reject that there is a unit root, and that the data appear to be non-stationary (observed test statistic of -1.9891 is less than critical test statistic at five percent -2.87). Furthermore, we cannot reject the null hypothesis of the presence of a unit root but no drift. The series appears to be non-stationary with no drift or trend.

One common way to make a time series stationarity is by differencing. Differencing can eliminate both trend and seasonality (Hyndman & Athanasopoulos, 2014). Below are three ways to differentiate the time series used in this paper.

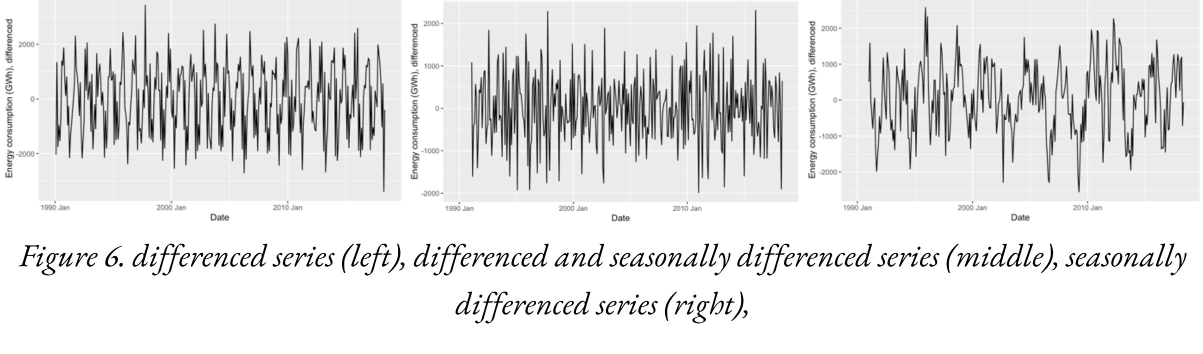


Figure 6. differenced series (left), differenced and seasonally differenced series (middle), seasonally differenced series (right),

The strong seasonality in the data can affect the power in the tests, so the data will be seasonally differenced. Next, the ADF and KPSS test are rerun to understand if it now appears to be stationary. The seasonally differenced series does not show an apparent trend (see right of figure 6), so the KPSS and ADF test will not assume a trend. The results of the ADF test shows that the null hypothesis of the presence of a unit root can be rejected as the observed test static is -6.5673 and is lower than the critical test statistic of -3.44, and the series appears to be stationary at the one percent significance level. Furthermore, the test that tests the null hypothesis that one of a_0, a_2 and γ are equal to zero can be rejected. The null hypothesis in KPSS test with no trend on the seasonally differenced cannot be rejected as the observed test statistic of 0.0386 is smaller than the critical test statistic at one percent, 0.739. Together, the results indicates that the series appears to be stationary and that it may be motivated to assume it has a drift. These results seem plausible considering the characteristics of the series observed in time plot in figure 1.

c. Model creation

The parameter D in the ARIMA is set to 1 as seasonally differentiation of order one has shown to be motivated. The autoregressive and moving average components are used to model the remaining autocorrelation after differencing the time series. To identify a suitable autoregressive component and a moving average component in the time series, we look at the ACF and PACF plots. The plots are shown below.

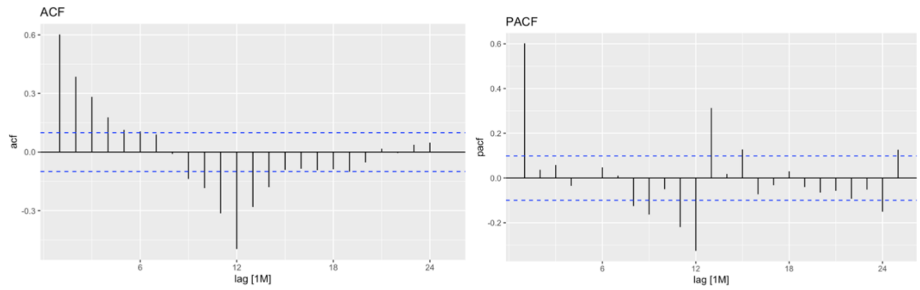


Figure 7. ACF plot, seasonally differenced series (left), PACF plot, seasonally differenced series (right)

The ACF shows a sinusoidal pattern and the PACF shows nine significant lags scattered across the graph. The order of the autoregressive component is sometimes set to the number of significant subsequent lags in the PACF plot if the ACF shows a sinusoidal pattern and the PACF shows non-significant values beyond lag p. The sinusoidal decay observed in the ACF plot signals that a higher order of the autoregressive component is motivate. Also, the scattered significant lags in the PACF plot indicate that a moving average component of one could be motivated. Additionally, the significant spike at the twelfth month indicate that some seasonality is still left. A MA(1) term in the seasonal component will therefore be added. The seasonal lags could display an exponential decay – further indicating that at least a seasonal MA(1) is motivated. Large values for p and q imply that a large number of parameters needs to be estimated and the possibility to overfit the model arises and the confidence intervals in the forecast will be larger. The models will be constructed conservatively, setting p equal to 1 and 2, and q equal to 0 and 1. The handpicked models are ARIMA(1,0,0)(0,1,1)[12], ARIMA(2,0,0)(0,1,1)[12], ARIMA(1,0,1)(0,1,1)[12] and ARIMA(2,0,1)(0,1,1)[12]. Additionally, an Auto ARIMA that is not using a stepwise search, and thereby traverse through the entire model space of the parameters p and q, is created. The Auto ARIMA picked an ARIMA(1,0,2)(0,1,1)[12] model. The comparison of the three models is shown below.

Model	AIC	AICc	BIC
ARIMA(1,0,0)(0,1,1)[12]	5108.51	5108.73	5130.00
ARIMA(2,0,0)(0,1,1)[12]	5108.61	5108.95	5127.75
ARIMA(1,0,1)(0,1,1)[12]	5108.06	5108.25	5127.04
ARIMA(2,0,1)(0,1,1)[12]	5106.50	5106.76	5129.28
ARIMA(1,0,2)(0,1,1)[12] (auto)	5106.22	5106.48	5123.79

Figure 8. Comparison of ARIMA models

The results from figure 8 show that the model with the lowest AIC and AICc is the Auto ARIMA model. The performances are however very similar to each other. The residuals of the ARIMA(1,0,2)(0,1,1)[12] model can be seen below.

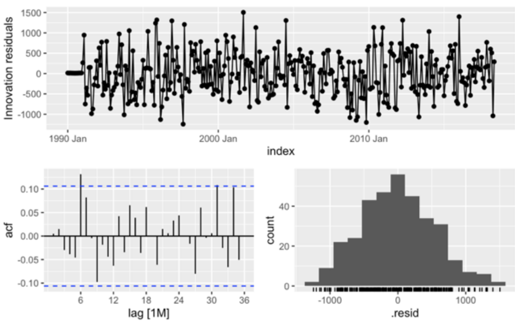
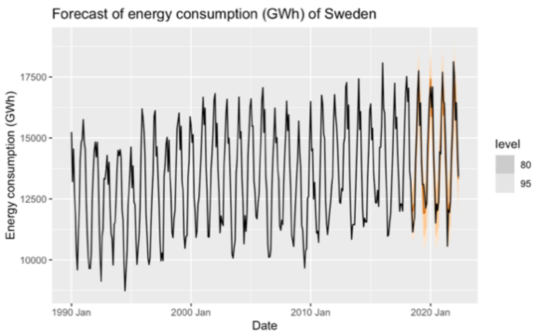


Figure 9. Residuals of ARIMA(1,0,2)(0,1,1)[12]

The residuals do not show any clear correlation between residuals but there are some signs of weak heteroscedasticity, where residuals between the years 1990 and 1995 show more fluctuation compared to between the years 2005 and 2010. The residuals appear to have a zero mean. Results from Ljung-Box test and Shapiro-Wilk test show that the residuals are homoscedastic, non-autocorrelated and normally distributed (see appendix).

d. Forecast using ARIMA(1,0,2)(0,1,1)[12]

The forecast of the ARIMA(1,0,2)(0,1,1)[12] model is shown below. The performance is also compared to the four hand-picked models. Their forecasts can be found in the appendix.



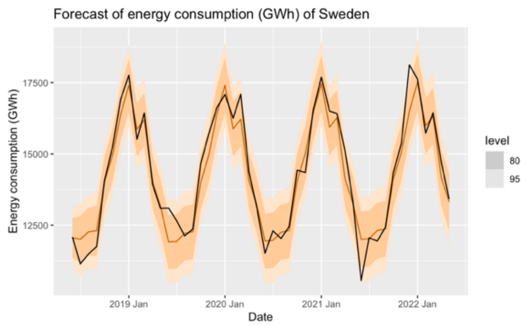


Figure 10. Forecasts with ARIMA(1,0,2)(0,1,1)[12] on the test set. Left figure shows the entire series including the test set and the right figure shows the forecast only on the test set.

Model	ME	RMSE	MAE	MPE	MAPE	MASE	RMSSE	ACF1
ARIMA(1,0,0)(0,1,1)[12]	167.16	549.48	415.08	0.8877	2.9446	0.5403	0.5820	0.1916
ARIMA(2,0,0)(0,1,1)[12]	165.72	548.62	414.73	0.8803	2.9440	0.5398	0.5811	0.1883
ARIMA(1,0,1)(0,1,1)[12]	162.65	548.19	414.00	0.8582	2.9405	0.5389	0.5806	0.1897
ARIMA(2,0,1)(0,1,1)[12]	124.29	542.10	403.66	0.5781	2.8855	0.5254	0.5741	0.2027
ARIMA(1,0,2)(0,1,1)[12] (auto)	129.14	545.00	405.33	0.6090	2.8939	0.5276	0.5772	0.2102

Figure 11. Performance of ARIMA models

The hand-picked ARIMA(2,0,1)(0,1,1)[12] appears to be the model that produces the best forecasts in terms of all accuracy measures. Visually, all ARIMA models produce very similar forecasts.

8. Conclusion

In this paper, a model using exponential smoothing and a model using ARIMA has been constructed in order to forecast electricity consumption of Sweden. The models were trained on a dataset of monthly electricity consumption of Sweden between January 1990 and May 2018 to produce a forecast of the subsequent four years. The best performing models from each category were an ETS(A,N,A) and an ARIMA(2,0,1)(0,1,1)[12]. The overall best performing model in terms of the accuracy measures used in this paper was the ARIMA(2,0,1)(0,1,1)[12].

For the curious reader, a LSTM neural network has been constructed that also provided satisfactory results. You can read the code on my GitHub (https://github.com/ErikKonstenius) or inspect some of the key plots that have been included in the very bottom of the appendix.

9. References

Amanda, D., 2022. *Risk för betydligt högre elpriser i vinter*. [online] DN.SE. Available at: <https://www.dn.se/ekonomi/risk-for-betydligt-hogre-elkostnader-i-vinter/> [Accessed 20 August 2022].

Chen, Y., 2017. The Factors Affecting Electricity Consumption and the Consumption Characteristics in the Residential Sector—A Case Example of Taiwan. *Sustainability*, 9(8), p.1484.

Hyndman, R.J & Athana-sopou-los, G (2014): Forecasting: principles and Practice (HA)

Iaea.org. 2022. *Energy Projections, calculating long term energy demand*. [online] Available at: <https://www.iaea.org/topics/energy-projections> [Accessed 20 August 2022].

Kalle, L., 2018. *Elanvändning*. [online] Energiföretagen. Available at: <https://www.energiforetagen.se/energifakta/elsystemet/energibranschen-viktig-for-svensk-ekonomi/elanvandning/> [Accessed 20 August 2022].

Statista. 2022. *World electricity consumption*. [online] Available at: <https://www.statista.com/statistics/280704/world-power-consumption/> [Accessed 20 August 2022].

Statistikdatabasen. 2022. *Elanvändning, GWh efter användningsområde och månad*. [online] Available at: <https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__EN__EN0108__EN0108A/ElanvM/> [Accessed 20 August 2022].

Svetunkov, I., n.d. 4.1 ETS taxonomy | Forecasting and Analytics with ADAM. [online] Openforecast.org. Available at: <https://www.openforecast.org/adam/ETSTaxonomy.html> [Accessed 20 August 2022].

Rodeck, D., 2022. *What Is A Recession?*. [online] Forbes Advisor. Available at: <https://www.forbes.com/advisor/investing/what-is-a-recession/> [Accessed 20 August 2022].

Appendix

a. Appendix A. Exploratory Data Analysis

```
        Box-Ljung test

data:  df_ts$value
X-squared = 253.39, df = 1, p-value < 2.2e-16
```

Figure a.1. Ljung-Box test

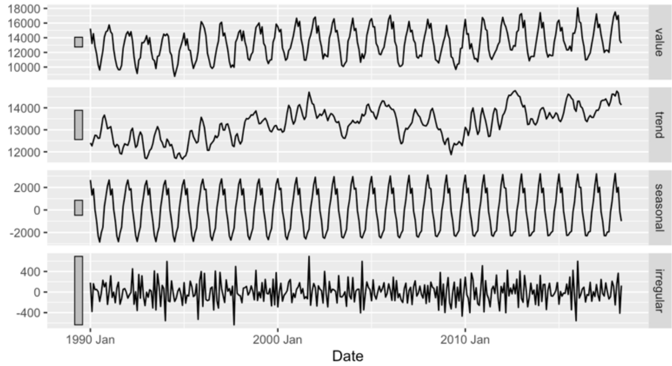


Figure a.2. SEATS decomposition of time series

b. Appendix B. Exponential Smoothing

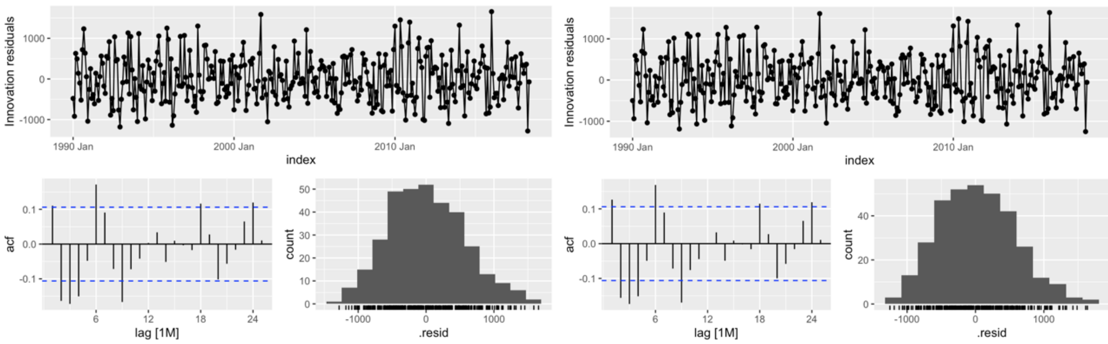


Figure b.1. Residuals of ETS(A,A,A) right, Residuals of ETS(A,Ad,A) left

Model	lb_stat	lb_pvalue
ETS(A,N,A)	20.93039	0.269656
ETS(A,A,A)	25.59411	0.156365
ETS(A,Ad,A)	23.97771	0.136693
ETS(A,N,A) (auto)	20.93039	0.269656

Figure b.2. Ljung-Box test on the residuals of the ETS models

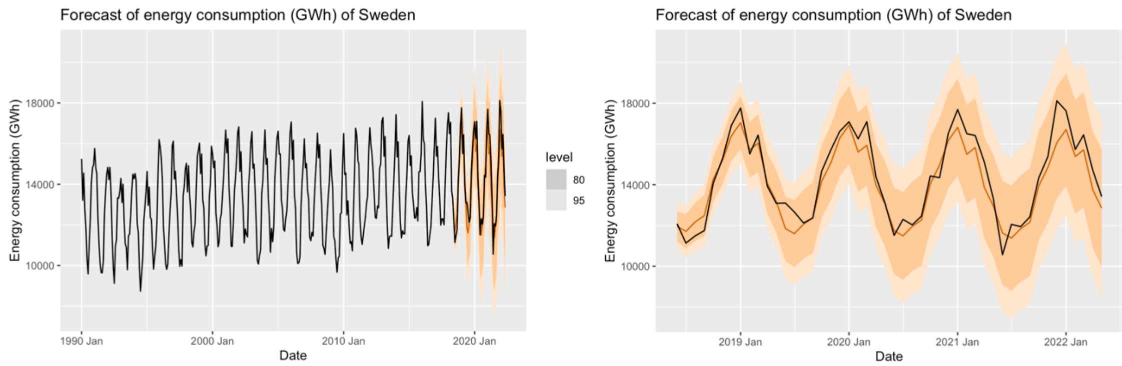


Figure b.3. Forecasts with ETS(A,A,A).

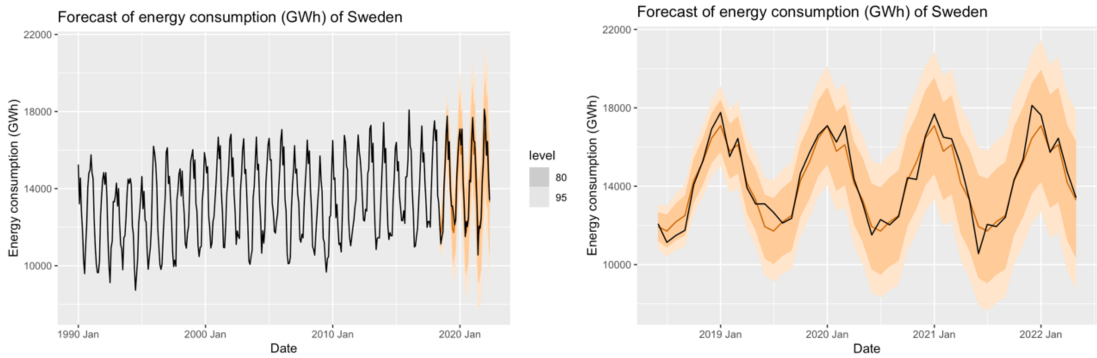


Figure b.4. Forecasts ETS(A,Ad,A)

c. Appendix B. ARIMA

```
Test is of type: tau with 5 lags.

Value of test-statistic is: 0.0374

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.119 0.146  0.176 0.216
```

```
Test is of type: mu with 5 lags.

Value of test-statistic is: 0.6659

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure c.1. KPSS test with trend (left), KPSS test without trend (right)

```
Value of test-statistic is: -3.2302 3.6266 5.3304

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

```
Value of test-statistic is: -1.9891 2.0859

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
phi1  6.47  4.61  3.79
```

```
Value of test-statistic is: 0.369

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

Figure c.2. ADF test, trend and drift (left), ADF test, drift but no trend (middle), ADF test, drift but no trend and no drift (right)

```
Test is of type: mu with 5 lags.

Value of test-statistic is: 0.0386

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Figure c.3. KPSS test without trend on seasonally differenced data

```
Value of test-statistic is: -5.7804 16.7284

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
phi1  6.47  4.61  3.79
```

Figure c.4. ADF test without trend on seasonally differenced data, drift assumed

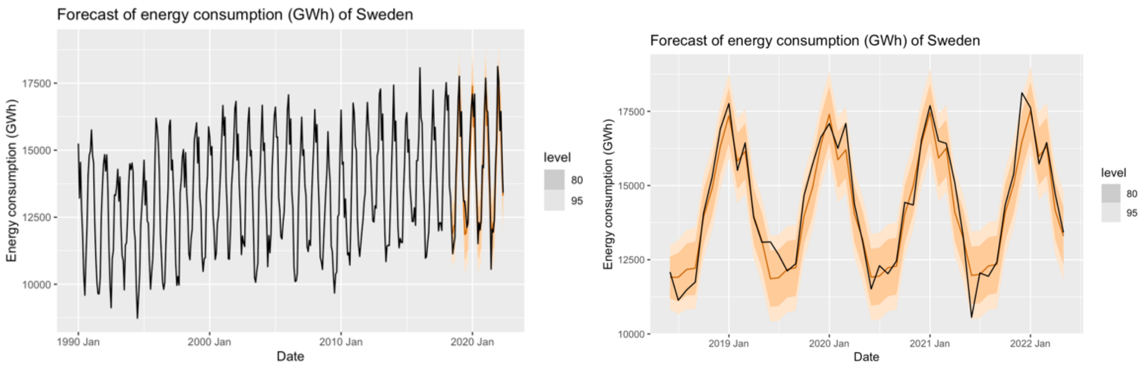


Figure c.5. Forecasts with ARIMA(1,0,0)(0,1,1)[12]

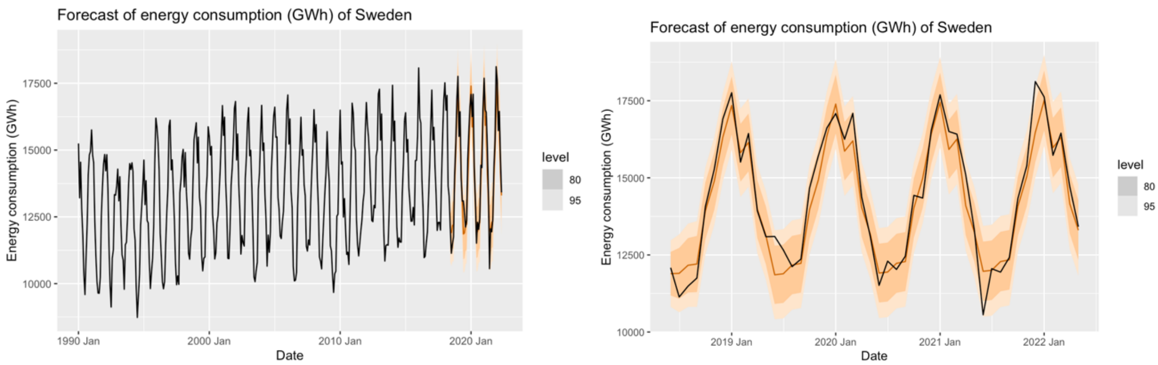


Figure c.6. Forecasts with ARIMA(2,0,0)(0,1,1)[12]

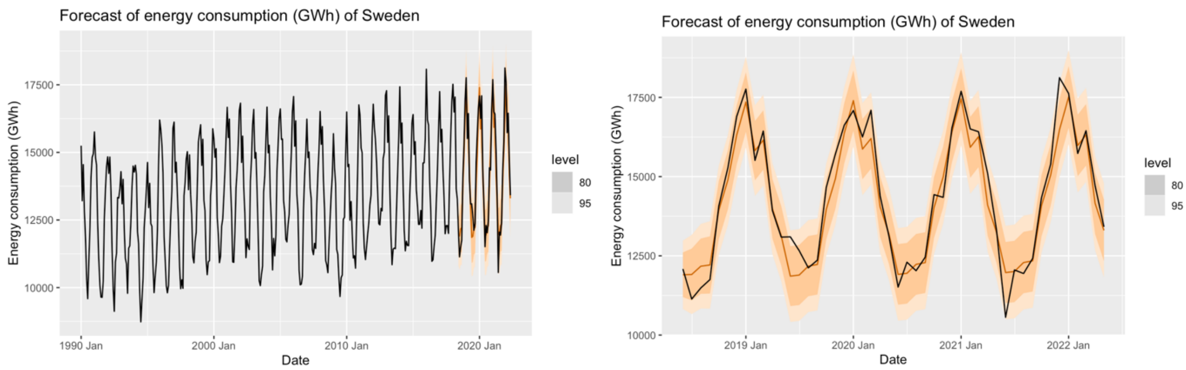


Figure c.7. Forecasts with ARIMA(1,0,1)(0,1,1)[12]

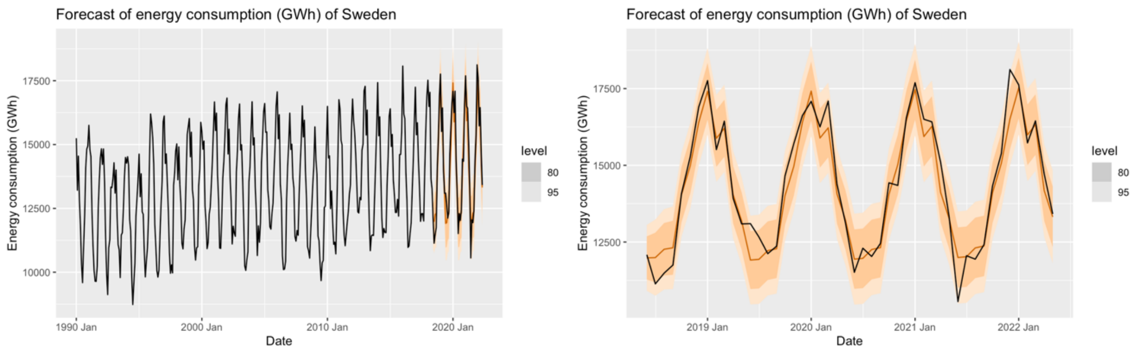


Figure c.8. Forecasts with ARIMA(2,0,1)(0,1,1)[12]

Model	lb_stat	lb_pvalue
ARIMA(1,0,0)(0,1,1)[12]	26.10014	0.1625298
ARIMA(2,0,0)(0,1,1)[12]	25.21938	0.1931669
ARIMA(1,0,1)(0,1,1)[12]	25.24993	0.1920361
ARIMA(2,0,1)(0,1,1)[12]	22.00189	0.3404080
ARIMA(1,0,2)(0,1,1)[12] (auto)	21.96328	0.3425063

Figure c.9. Ljung-Box test on the residuals of the ETS models

d. Appendix C. Long Short-term Memory Neural Network

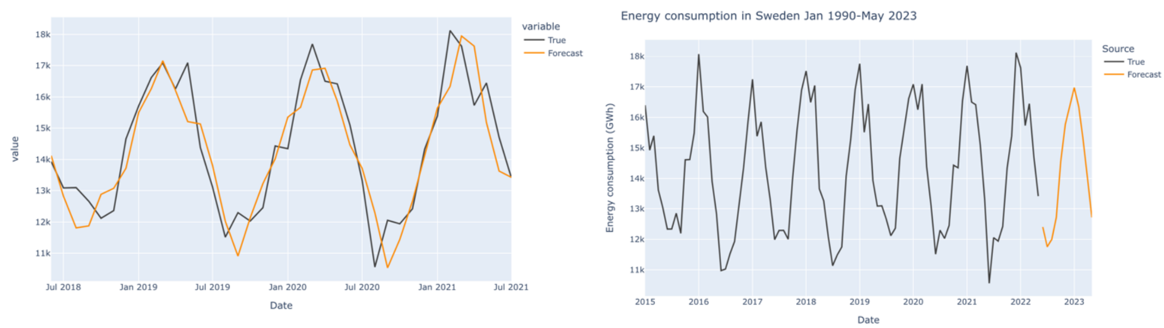


Figure d.1. LSTM NN forecast on test set (left), LSTM NN future forecasts (right)