Project in Economics and Business Administration: Forecasting residential electricity consumption in Aarhus

Number of characters ≈ 53.324

GitHub repo: github.com/Timonflogo/Economics-Project

Student: Timon Florian Godt

School of Business and Social Sciences, Aarhus University

Supervisor: Yana Petrova

Department of Economics and Business Economics, Aarhus University

(Dated: January 3, 2022)

Europe is headed towards a major energy crisis which is very likely to affect Industries, Businesses, and certainly also households, as energy prices are rising and temperatures are falling. This paper investigates the applicability and potential of advanced forecasting techniques such as ARIMA, SARIMA and their regression versions of including exogenous variables, ARIMAX and SARIMAX, as well as Recurrent Neural networks (RNNs) such as the infamous Long-Short-Term-Memory (LSTM), on the electricity consumption of one single student household in Aarhus. The Electricity consumption series was tested for correlations with meteorological features from the region and no significant relationships were able to be established for hourly frequency data. The reasons for this are assumed to be the fluctuations in the timetables of the residents which adds certain noise to the dependent series as well as the representation of a single household as opposed to an aggregate of multiple households in the same area. The electricity consumption was subsequently forecasted with hourly, daily, and monthly frequency, where it was found that ARIMAX, SARIMA, and SARIMAX models performed superior to other more complex models such as the LSTM.

1. INTRODUCTION

Europe is headed towards a major energy crisis which is very likely to affect Industries, Businesses, and certainly also households, as energy prices are rising and temperatures are falling. According to Bloomberg, Natural Gas storages in Europe are significantly lower than last year and uncertainty about potential supply crunches is further driving up prices for natural gas, which subsequently elevates the electricity bills of end consumers (Chadwick, 2021). While some countries in the euro area, especially in Scandinavia, are more progressive regarding the green transition, almost every country has adopted green policies during the recent COP26 meeting. Russia, the largest supplier of natural gas to Europe, on the contrary is arguably not delighted about the recent developments about Europe wanting to go all green, and some voices have been raised about Putin's passive behaviour in the recent developments (Ng, 2021).

According to Rachel Morison from Bloomberg, with the European winter not even having reached its peak, energy prices are still rising and breaking records while energy supply across Europe takes further hits. While Europe is again the epicentre of the COVID-19 pandemic, countries such as the UK have seen a major price spike which forced some industrial companies to cut productions, increased tensions with neighboring states through moving of supplies, and households should brave for potential blackouts or being asked to use less energy (Morison, 2021). Mrs. Morison further states that France is hit by delays in the maintenance of nuclear power generators due to the pandemic. Furthermore, Russia claims to export as much as possible to Europe, which still does not seem to be enough.

Since prices are continuing to rise and the future of the energy situation in Europe is rather uncertain, one might want to investigate their personal household energy consumption to assess the current demand patterns and predict their future consumption. Energy prices are heavily fluctuant at the moment, therefore one might further want to use such forecasts to prepare for rising costs and the eventuality of a blackout, all which are dependent on the consumption itself. If personal consumption can be somewhat accurately forecasted, households have a better understanding about how to change their consumption behaviour and become more conscious about efficient use of energy. An Hourly forecast could be used to adjust the usage of certain appliances which use a lot of energy, while Daily and Monthly forecasts could be used to investigate the current trend of energy consumption and spark efforts to reduce the future consumption.

While energy consumption in a household is composed of a variety of sources, the primary source is arguably the electricity consumption. Data about electricity consumption is furthermore readily available from most electricity suppliers and can be retrieved either through APIs or bulk downloads from the suppliers website.

Multiple traditional and advanced time series analysis and forecasting algorithms could be applied to this problem. This paper focuses on three algorithms, the traditional ARIMA(X), the SARIMA(X), and Long-Short-Term-Memory (LSTM), to answer the following research question:

Which Machine/Deep Learning algorithms perform best on forecasting a single residential household electricity consumption?

2. LITERATURE REVIEW

The recent developments call for action in terms of optimising the energy that we have at hand. Since electricity is a special commodity by nature and cannot be stored cheaply in large quantities, optimising supply and demand on an ongoing basis is essential. These optimization problems are primarily time series related. Recent literature shows that certain Machine/Deep learning methods are becoming inherently superior to traditional forecasting methods such as AR, MA, ARIMA, Holt-Winters, etc.. Furthermore, most Machine Learning as well as Deep Learning algorithms allow for the inclusion of multiple features such as temperature, wind speed, and precipitation which have previously been shown to be correlated with the amount of electricity consumption (Chikobvu & Sigauke, 2012). These features are readily available from country-specific meteorological institutions such as DMI in Denmark and can be easily extracted by the average user.

Chikobvu & Sigauke (2012), used a SARIMA model and a SARIMA-Regression otherwise called as SARIMA with exogenous variables (SARIMAX) model to forecast daily peak demand for electricity up to seven days into the future. The data used in the analysis showed strong seasonality and a trend component which was further deemed significant through the use of the Augmented Dickey Fuller methodology. Subsequently, they found that the SARIMA model itself produces more accurate short-term forecasts, while the SARIMAX captured important features for electricity demand which could enable decision makers and other stakeholders in further analysis and scheduling of electricity (Chikobvu & Sigauke, 2012).

While most papers focus on daily predictions, hourly predictions have recently been investigated more deeply by the use of more complex model specifications such as SARIMAX with interactions (Elamin & Fukushige, 2018). Elamin & Fukushige (2018), analysed a dataset of electricity load for a region in Japan paired with multiple exogenous variables such as meteorological features, social events, and seasonal patterns (daily, weekly, monthly, yearly). They found that while a normal SARIMAX of order (5,1,4)(0,1,0,24) model already produced relatively accurate predictions, the accuracy was even further improved by including interaction terms of the exogenous variables.

Another set of advanced algorithms which have gained in prominence within the last decade are deep learning algorithms, also known as neural networks. There are a variety of neural networks available for different types of problems such as classification, regression, NLP (Natural Language Processing), and time series. In the realm of time series forecasting, which essentially is the forecasting of a sequence, Recurrent Neural Networks (RNN) which are primarily used for sequence processing such as NLP have also shown to be very effective. A recent paper about the application of the Long-Short-Term-Memory (LSTM) algorithm has shown lower error rates than SARIMA(X) models on electricity demand forecasts. The paper by Dubey et al. (2021) investigated aggregated daily electricity consumption of more than 5000 households in London and found the LSTM to be superior to its more traditional counterparts.

While Machine/Deep Learning algorithms can be incredibly powerful, another paper has found that hyperparameter tuning for deep neural networks such as LSTM is essential to boost performance even further (Atef & Eltawil, 2020).

3. METHODS

The next section describes the methods used in this paper and is already deemed as a pre-analysis of the data and an investigation into its characteristics.

3.1. Data Extraction

The electricity consumption data was extracted from NRGi's online service. NRGi is an electricity provider in Aarhus, Denmark. The researcher is a customer of NRGi and was therefore able to retrieve his personal household consumption data through his own account. The data however, does not represent the consumption of a single person household as the researcher lives in a shared apartment with another individual. The consumption data therefore represents the electricity consumption of a two person household. Moreover, data for a period of more than three years was extracted, however the researcher moved into the apartment in October 2020. The other tenant, a young male, let's call him tenant2 moved in around july, 2018. Before the researcher moved in, a young girl (tenant3) was a resident in the household for more than 2 years. All tenants are working students, meaning they are bound to university timetables as well as changes in their working schedule.

The extracted electricity consumption data is of hourly format and values of each time step represent the consumption in kWh (kilowatt hour) for the last hour. Meteorological data was extracted from the Danish Meteorological Institute (DMI). DMI has a large variety of meteorological parameters measured across the country, which are freely available to download through their open access REST API (DMI, 2021). DMI has a weather station in southern Aarhus which lies in close proximity to the researcher's household. It was therefore evaluated as being sensible to be used in combination with the households electricity consumption data. Data was extracted as bulk download for a period of more than 3 years for every weather station in Denmark, including all 29 parameters made available by DMI. Due to differences in data frequency and redundancy some of the extracted parameters were deemed not relevant for this study and therefore excluded from the dataset.

3.2. Data Processing

After extracting data from NRGi and DMI, both datasets have been processed individually to ensure correctness of data types, frequency, and completeness. Firstly, an empty dataframe (df_s) has been constructed using the Pandas package in Python, simulating hourly datetime observations for the specified period from 2018-10-01 to 2021-11-08. The reasoning for this step was to ensure that no sequence gaps due to missing values in the extracted data are present during the modelling process. Initially the data from NRGi (df1) already had a good format, however, the datetime indicator from NRGi was a string format in danish. This had to be changed to datetime format to make a merge of the data frame with the simulated data frame and the following weather data frame possible. Furthermore, duplicate values were present in some instances. These happened to appear quite arbitrarily and were removed from the dataset. When merging df1 into df_s, no missing values have been detected, indicating the data is clean.

The bulk data extract from DMI was incredibly large as it included high frequency data for all weather stations in Denmark. One file for each month for the previously specified period was downloaded, adding up to a total of around 55 gigabyte of weather data. The tool used for processing and shaping the data frame was Microsoft Excel PowerQuery. During the processing phase, the data was grouped by hour from 10-minute frequency to hourly frequency using the average of the last 6 observations. The final weather data frame (df2) contained observations where each represents numerical values for the last hour. This step was necessary to make an accurate merge of the data possible. After merging df2 into df_s, missing values have been detected in all variables as seen in table 1. Due to the share of missing values per category being very large for cloud_height, it was decided to remove the variable from the dataset. All other missing values have been addressed using the data imputation method called Last Observation Carried Forward (LOCF). Although this data imputation method is primarily used in Longitudinal Data Analysis, the researcher found it to be a valid method for imputing data for the given weather features in this setting (Liu, 2015, ch. 14.2.2). The reasoning stems from longitudinal data being similar to time series data, as well as the high frequency of the data being hourly. It is assumed that dramatic changes in the weather condition on an hourly basis are not impossible, but rather unlikely.

TABLE I: Missing values for features in the weather data frame (df2). After preprocessing the data, the time period of the data series was adjusted from 01.11.2018 to 01.11.2021 to include exactly three years worth of data.

Feature	Count of missing values
pressure_at_sea	8
precip_dur_past10min	2440
wind_dir	550
wind_speed	550
temp_dew	7
pressure	10
visib_mean_last10min	2391
temp_dry	7
cloud_height	46752
humidity	1879
cloud_cover	532
visibility	1118

3.3. Target and Feature Engineering

A fundamental part of Machine Learning is correctly carried out Target and Feature engineering. In some cases, the methods deployed can have a significant impact on the model performance of forecasting algorithms. Especially in cases where the data shows some kind of seasonality which has to be addressed to ensure more accurate forecasting.

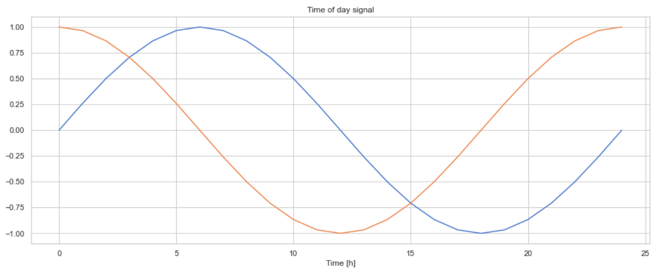

FIG. 1: Hourly signal based on Sine and Cosine Functions (source: (Tensorflow, 2021))

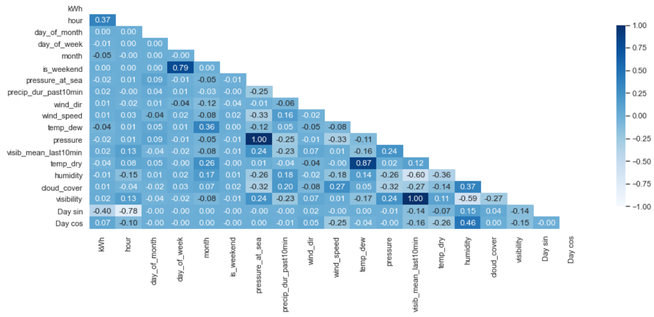

FIG. 2: Correlation heatmap of Target and Feature dataframe (source: own)

3.3.1. Creating new features

Intuitively, it could be assumed that some seasonality for the month, weekdays, hour of day, as well as weekends is present. Hence, it was decided to create the following features to be further analysed and potentially included in the modeling of advanced forecasting algorithms.

New features created:
Hour - The hour of each observation (0-23, where 0 is 00:00)
Weekday - the weekday each observation (0-6, where 0 is Monday)
Month - the month of each observation (1-12, where 1 is January)
Is_weekend - whether an observation is on a weekend (binary, where 0 is no weekend)
Day sin - reflecting hourly signal in the series
Day cos - reflecting hourly signal in the series

Furthermore, to better represent the hourly signal in the electricity consumption pattern, two features based on sine and cosine functions one and two, respectively, were created as seen in figure 1. The features are not as important for advanced time series algorithms such as SARIMAX, however, according to Tensorflow time series forecasting online documentation, these signals are of importance to Neural Networks, to clearly represent seasonality in the data (Tensorflow, 2021). The Network would otherwise get confused by the numerical distance of 0 and 23. These two numbers are closer to each other, when interpreting them as hours, rather than far away, as the numerical distance would suggest. The added features will be further analysed with correlation heatmaps to assess whether or not to be included in the modeling process.

3.3.2. Removing uncorrelated features

After creating six new features in 3.3.1 and combining them with the dataset from 3.2 the new dataset has a total of 19 features. To assess the usability of the given features in the dataset, a correlation heatmap was created. The correlation heatmap is based on the Pearson Correlation coefficient, which states that any value lower than 0.8 or greater than -0.8 is insignificant, as it is the most commonly used coefficient in statistics (Fernando, 2021). After inspecting the correlation heatmap in figure 2, a significant independence of the target variable kWh from the weather and most of the newly created features from section 3.3.1 becomes apparent. According to the Pearson Correlation coefficient, none of the 18 variables in the dataset are significantly correlated to the electricity consumption of the residents. This could partly be explained by the different utilities used during winter and summer. One might assume that electricity consumption in winter would be higher than in summer since days are shorter and lack of natural light needs to be compensated by the use of lamps and other light sources. However, in summer when temperatures are high, the residents are more inclined to use ventilation systems to cool down the room. These ventilation systems are electricity based and potentially cancel out the increase in electricity consumption due to lack of light. The argument stems from the knowledge that the residents use gas based heating, which is a different energy source not included in the dataset. Further comments on this remark can be found in section 5.0.

The decision was made to drop most of the features previously processed and added. It is acknowledged that when inspecting figure 3 the hour and the similar Day sin feature could be assumed to have an influence on the target. Even though, both features are statistically insignificant, according to the Pearson correlation coefficient mentioned earlier, based on the potential influence on the target, a potential increase in performance could

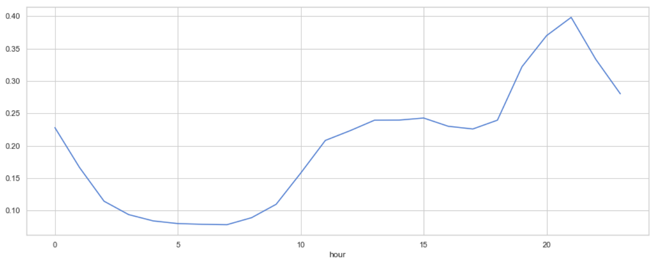

FIG. 3: Electricity consumption in kWh grouped by hour of day (source: own)

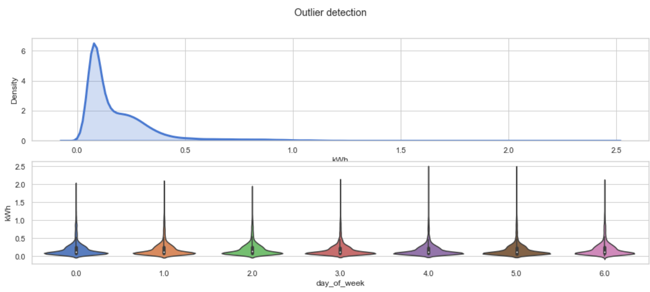

FIG. 4: Density plot (top) of the kWh series and distribution of kWh series by day (bottom) (source: own)

be achieved by including them when modeling the data with advanced algorithms and the hour feature was kept in the dataset.

3.3.3. Outliers

When inspecting the distribution of the target variable kWh in figure 4 it becomes apparent that some outliers are present in the data. These outliers seem to be rather rare and do not occur very often. kWh over the entire period has a mean of around 0.20 and some deviation. However, for the given period 1878 observations have a kWh value of 0.5 or higher, 370 observations have a value of 1.0 or higher and 7 observations have a value of 2.0 or higher. The extreme outliers ($kWh > 2.0$) seem to occur rather arbitrarily, while moderate outliers seem to occur more often in the autumn and winter months.

None of the outliers were removed since kWh is the target variable and the data should be as authentic as possible to model the natural behaviour of electricity consumption in the residential home.

3.4. Frequency adjustments

The initial hourly frequency of the data will further be used to assess whether accurate forecasts can be achieved

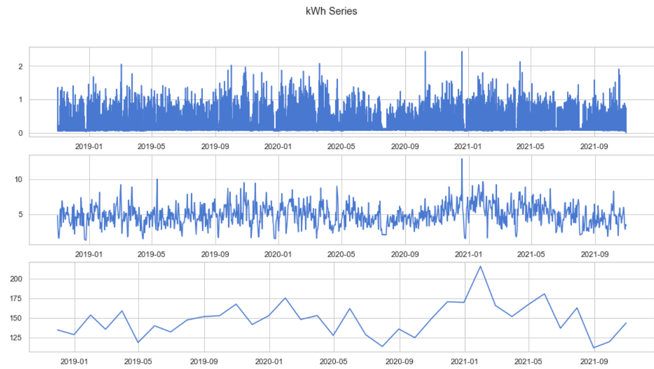

FIG. 5: kWh series over the specified period from 01.11.2018 to 31.10.2021 on hourly (top) daily (middle) and monthly (bottom) frequency (source: own)

using advanced algorithms. Additionally, the reduced series, only including the target feature kWh, was resampled and all values were aggregated on a daily basis and monthly basis. This step will make a monthly forecast for electricity consumption significantly easier as the series becomes less noisy with decreasing frequency. All three series objects can be inspected in figure 5.

For the other two aggregated series, a seasonal pattern such as the one in figure 3 were detected and dummy encoded arrays were created to be used as exogenous variables in the further application of ARIMAX and SARIMAX models in section 4.

3.5. Algorithms and forecasting approaches

The following section will elaborate on the definition and functionality of the algorithms used in the analysis part of this paper. This section will be rather short, since it is assumed that most methods are widely known mathematically and well understood.

3.5.1. ARIMA(X)

The so-called AutoRegressive Integrated Moving Average (ARIMA) model is a combination of an Autoregressive model of order p and a moving average model of order q with a differencing component d. This method is also referred to as the Box-Jenkins Methodology (Wichern & Hanke, 2013, ch. 8). The order of the ARIMA model (p,d,q) can be easily specified with the auto_arima function in the python package pmdarima which is based on the auto.arima equivalent of the programming language R (pmdarima, 2021). This functinality will be further discussed in the Analysis section 4.0.

3.5.2. SARIMA(X)

The Seasonal ARIMA (SARIMA) is an extension of the previously described ARIMA model. It adds a further autoregressive model of order P, a moving-average model of order Q and a differencing component D for the seasonal nature of a series. Furthermore, the seasonality is defined as m by specifying the period which could be e.g. daily, weekly, monthly, or yearly. The final specification has the following components (p,d,q)(P,D,Q,m).

A further extension of the model is the SARIMAX which is the addition of exogenous variables to the SARIMA model. This model allows for the inclusion of exogenous variables such as meteorological data or further engineered features such as the hour feature from section 3.2.2 of this report. These terms are not autoregressed on and are simply just added to the model specification (Phosgene89, 2021). The exogenous variable at time t influences our series at time t. In the same logic, the exogenous variable at any time n of future predictions of our series will have to be present such as to estimate e.g. kWh at t50 we will need the value of the exogenous variable for that particular future observation. This might pose some issues if the exogenous variables used are features which carry future uncertainty such as meteorological data, since an estimation of future values would need to be used based on an auxiliary forecast. Since no significant correlation was found between the meteorological features and our target feature, this will not be further investigated in this paper.

3.5.3. LSTM

The Long-Short-Term-Memory (LSTM) is a type of RNN which is designed to overcome the so-called vanishing gradient problem which makes it difficult and sometimes even impossible to train neural networks if the number of layers increases. Furthermore, the vanishing gradient problem significantly hampers the learning of long-term dependencies of a sequence. The LSTM was designed by Hochreiter, and Schmidhuber in 1997 and is similar to a simple RNN, but adds a way to carry information in memory for many timesteps (Hochreiter & Schmidhuber, 1997). In the LSTM, information can be carried on parallel to the sequence and reintroduced at timestamps in the far future similar to a conveyor belt which runs parallel to the sequence, where information can jump on and off at any time (Chollet, 2018, p.202).

A simplified explanation of the LSTM anatomy would be to view the upper sequence c in figure 6 over many timesteps as the conveyor belt which retains information depending on new input i. At a given timestep t, new information is added to the network through an input gate, while at the same time, multiple forget gates determine which information should be retained and put on the conveyor belt to be carried on to the next step. The forget gates are equipped with sigmoid functions to squish values between 0 and 1. The closer a value is to 0 means forget whereas 1 means retain (Phi, 2018).

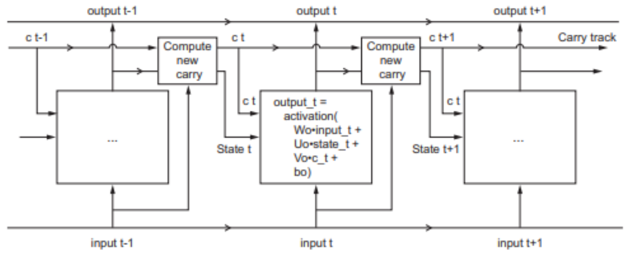

FIG. 6: LSTM Anatomy (Source: (Chollet, 2018, p.204))

4. ANALYSIS

The order of the ARIMA, SARIMA, and SARIMAX models in the following section is determined by the auto_arima function of the pmdarima package in python using the Akaike Information Criteria (AIC) (pmdarima, 2021). The LSTM models are determined by manual tuning of some hyperparameters and the tuning process will be mentioned and commented on when necessary. All results are evaluated using the Root Mean Squared Error (RMSE) and summarised in table 2 at the end of this section.

4.1. Hourly Forecasts

The kWh series on an hourly basis was found to be stationary by using the Augmented Dickey Fuller (ADF) method and receiving a p-value of 1.538983e-25. The forecasting horizon was set to be 168 observations into the future, which represents exactly 7 days into the future. The forecasting horizon was chosen to be 168 instead of 24 as the researcher is more interested in long term forecast performance and getting insights into energy consumption further ahead in time.

Since no correlation of the series kWh with weather features were significant, the first algorithm to be investigated and benchmarked against is a simple ARIMA model. The optimal model specification for this problem is an ARIMA of order (2,0,0) and yields an RMSE of 0.2628. Since a decomposition of the time series indicates a seasonal component, the better model specification is given by a SARIMA of order (2,0,0)(2,0,0,24)

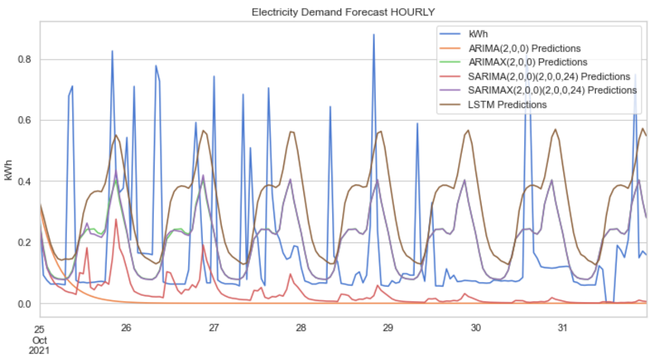

FIG. 7: Hourly forecast performance of kWh series over 168 hours (source: own)

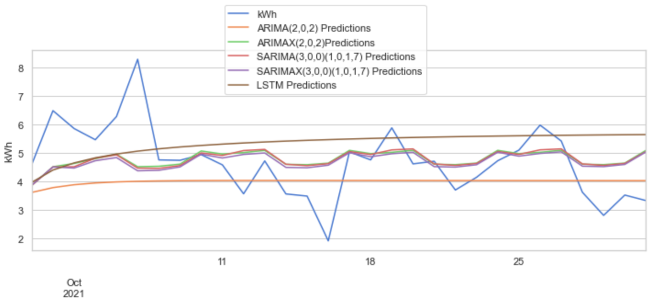

FIG. 8: Daily forecast performance of kWh series over 30 days (source: own)

which yields a reduction to 0.2439. Further introducing an array of dummy encoded hourly features to both models yields an ARIMAX and a SARIMAX with very similar RMSE of 0.2123 and 0.2128, respectively.

The LSTM model was trained with a batch size of 10 and a single layer of 16 neurons. The model was initiated with 30 epochs to reduce training time. Before training, the data was scaled using a min max scaler and the received predictions were inversely transformed back to represent real values of the kWh series predictions. This relatively simple LSTM model still took around 45 minutes to train on an average computer and the RMSE on hourly predictions was not competitive with the other much simpler methods such as the ARIMAX(2,0,0).

A visual comparison of the performance of the algorithms used on these data can be found in figure 7.

4.2. Daily Forecasts

For the daily forecast, the aggregated series kWh was found to be stationary and does not contain a unit root according to the ADF methodology as the p-value of 2.781346e-18 is near zero. The forecasting horizon was set to 30 days as it allows the researcher to get insights into the energy consumption of the next month, as electricity is billed on a monthly basis.

Without specifying seasonality, the optimal order to forecast the daily electricity consumption series is an ARIMA model of order (2,0,2) and yields an RMSE of 1.4365. However, after decomposing the series some seasonal pattern recurring on a weekly basis is visible. No trend component was detected. This seasonality will have to be captured by the forecasting method. Hence the new optimal model specification is a SARIMA of order (3,0,0)(1,0,1,7). Introducing this seasonal pattern into the model yields a reduction in the test error to 1.2478. When introducing a vector of exogenous variables, representing the weekly pattern of energy consumption, to the SARIMAX model, the error remains very similar with 1.2475. However, the same reduction in RMSE also occurs in the simpler ARIMAX(2,0,2) model with an RMSE of 1.2451.

After some manual tuning, The LSTM was trained with a batch size of 10 and one single layer with 8 neurons. The number of neurons was reduced to prevent overfitting. The model was trained for a total of 100 epochs and started to converge around epoch 70. The RMSE error of 1.5781 for the daily test prediction was still relatively high and not competitive with the other models. The LSTM model took significantly longer to train than the other algorithms and still seems to be overfitting on the data.

A visual comparison of the test set performance of this forecasting problem can be found in figure 8.

4.3. Monthly Forecasts

When using the ADF test on the monthly aggregated kWh series. The series was, similar to the previous versions of the kWh series, found to be stationary with a p-value of 0.000574. Hence also this series ARIMA specification will have no integrated component and no differencing is needed. The forecasting horizon in this problem was set to 12 month to allow for long term insights into the residential electricity consumption.

The optimal model specification without taking into account the yearly seasonality of the series is an ARIMA of order (1,0,0) and yields a RMSE of 30.5666. Decomposition of the series however revealed a seasonal component which when specified in a SARIMA model of order (1,0,0)(0,1,0,12) reduces the test error to 28.3448. When including a dummy encoded array, representing the month of the year, the ARIMAX test error increased, while the SARIMAX error stayed the same as the SARIMA. Hence the inclusion of exogenous variables does not seem to yield any benefits and only increases

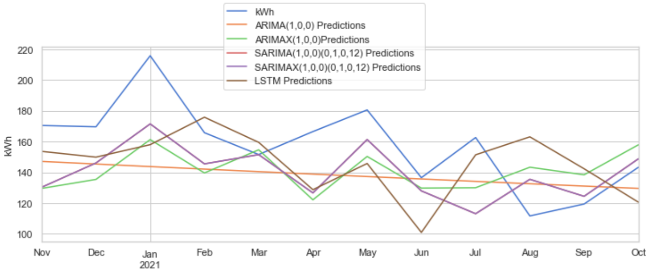

FIG. 9: Monthly forecast performance of kWh series over 12 months (source: own)

TABLE II: RMSE results of the different forecasting models on three different frequencies of the kWh series. The best performing algorithms are highlighted in bold.

Model	Hourly	Daily	Monthly
ARIMA	0.2628	1.4365	30.5666
ARIMAX	**0.2123**	**1.2451**	31.7122
SARIMA	0.2439	1.2478	**28.3448**
SARIMAX	0.2128	1.2475	**28.3448**
LSTM	0.2713	1.5781	31.4659

model complexity.

Due to a minimal amount of observations, the batch size for the LSTM model was set to 1 and the model was trained with a single layer of only 4 neurons. The model was initiated for a total of 1000 epochs and began to converge at around epoch 650. The training time was, similar to the previous problems, a lot longer than the simpler algorithms and again did not yield any benefits in terms of RMSE reduction as the RMSE of 31.4659 was higher than that of the simple ARIMA model.

An overview of the model performance on the test data can be found in figure 9.

5. DISCUSSION

As established in the introduction of this paper, due to the current fluctuations and uncertainty regarding the current energy situation in Europe, it's very beneficial for residential households to shed light into their electricity consumption patterns and forecast them into the future to increase awareness about conscious and more efficient use of electricity.

Initially an assumption was made, based on previous research, that the weather has a significant influence on electricity demand patterns of a single household. After retrieving such meteorological data from DMI for the given location of the household, the researchers found that no significant correlation with any weather features were present in the data. In addition to the reasoning provided in section 3.3.2, there could be multiple further reasons for this being the case.

Firstly, the series to be forecasted is the electricity consumption of only one single household and not an aggregate of a multiple of households in one specific location as is the case in all papers introduced in section 2.0, the literature review, of this paper. It could be argued that the so-called sample size in this case is only one, since only the electricity consumption of one single student household was analysed and forecasted. The reader should therefore be careful to assume that there is no relationship between electricity consumption and the weather for all residential households in Aarhus. It could very well be the case, that an aggregate of more than 5000 households shows similar significant relationships with the weather as have been found by Dubey et al. in London, or Chikobvu & Sigauke in South Africa (Dubey et al., 2021) (Chikobvu & Sigauke, 2012).

Secondly, the series showed a significantly stronger dependency on the seasonal pattern which is directly related to either daily, weekly, or yearly seasonality. This makes sense as the residents of the household are both working student individuals and therefore mostly not using any electricity at their home from 8 am until 4 pm. This however is not always the case. Since both individuals are students, there is a lot of fluctuation in the timetable and a distinct seasonal pattern on hourly as well as weekly basis might be distorted. It could be assumed that this would change when analysing similar data from a normal adult working household. This random component could very well be assumed to distort the relationship with the weather features.

Thirdly, if the residents are on vacation, Ill, or otherwise absent, there is only minimal electricity consumption for e.g. the fridge, which might distort the potential relationship with the weather. If more households would have been sampled, the absence of residents due to the previously mentioned potential cases would not be as high of an outlier as it is in the current data.

In terms of Forecasting performance, the results are somewhat surprising to the researcher as in the hourly and daily forecasting problems, a simple ARIMAX performed equally good as the more complicated and computationally intensive SARIMAX algorithm. The inclusion of an array of dummy encoded features to model the related seasonal dependency of the data seems to yield better results than only including a seasonal component as in the SARIMA. Furthermore, in the hourly forecast, as seen in Figure 7, the short long term predictions of the SARIMA become increasingly inaccurate and the seasonal pattern decays rather quickly, while with the inclusion of the exogenous array in the ARIMAX and SARIMAX, the seasonal pattern prevails. However, the predictions in this case are not very accurate as they very often overestimate or underestimate the actual electricity consumption at a given point. It seems that the simpler SARIMA model was already able to capture the seasonality of the monthly series and is therefore the preferred model, due to less complexity compared to the SARIMAX.

The incredibly computationally intensive LSTM performed worst on all forecasting problems in section 4. Even though in the hourly forecasting problem, the model was able to detect the seasonal pattern of the series kWh it performed worse than all other algorithms which are significantly less computational intensive. It is acknowledged that only some manual hyperparameter tuning was conducted and no grid search of features was carried out by the researcher. As described in the literature, such hyperparameter tuning for LSTMs on similar problems can significantly increase the model performance (Atef & Eltawil, 2020). Therefore, there is almost certainly room for improvement.

Furthermore, It is assumed that there are other components, which have not been investigated in this paper, that would have to be accounted for to further improve the forecasting accuracy. For the hourly forecast, one might want to look into adding geolocation data from the residential mobile phone location to further analyse and improve the accuracy of the forecasts. The inclusion of such a variable would adjust for the fluctuations in the time table of the students. A more simplistic solution would be to add more dummy encoded exogenous variables for whether the person is at home or not. This simple 1 or 0 variable would be able to account for both cases, being at work, school, or on vacation. Illness would be more difficult to account for, as it poses difficulties to predict such a scenario in the future, although certain periods such as autmn are well known for e.g. flu outbreaks.

Going back to section 3.3.3 showed some outliers in the dependent series and the data is heavily skewed to the right. It is acknowledge that no outliers were removed and no transformation, such as a logarithmic or min-max scaling, of the series was used when creating ARIMA(X) and SARIMA(X) models. Such a transformation could have resulted in different results in model performance.

In terms of the daily and monthly forecasts, the researcher has not tested for significant relationships with the retrieved weather features due to lack of time. However, it is acknowledged that the seasonal component of the monthly series investigated in section 4.3 could show a significant relationship with resampled and averaged monthly series of the weather features, as the electricity consumption seems to have risen in the winter of 2020. However, the winter of 2020 was also a period of lockdown in Aarhus, Denmark, due to the COVID-19 pandemic, which forced the residents to stay at home for both university related activities, as well as their student jobs. Hence, this relationship with the weather might be less significant if present and more dependent on the fact that there has been a lockdown. The reasoning stems from the observation in figure 5, that the seasonal monthly electricity consumption pattern was also given in the winter of 2019 but significantly less extreme. Furthermore, the previous resident was not a heavy computer user such as the researcher who moved into the household in October 2020. This could further be an explanation for the spike in winter 2020.

6. CONCLUSIONS

In this paper, a comprehensive analysis of the performance of multiple advanced algorithms on a single residential student household's electricity consumption in Aarhus was conducted. As of the researchers' knowledge, this is the first work given the previously established data characteristics. It was found that in the given case of a student household, no significant relationship with meteorological features such as the weather could be established in the case of hourly frequency of all series. This insignificant relationship could result out of the fact that only one single household is represented in the data, the residents are students and have very fluctuate timetables which increases the difficulty of establishing a seasonal daily daily electricity consumption pattern, as well as periods of vacation, illness, or other reasons of absence, where no consumption occurs, all of which potentially distort the relationship with the weather.

The best performing algorithms on the hourly and daily forecasting challenges with a forecasting horizon of 168 and 30 observations, respectively, was a simple ARIMAX model which is a ARIMA Regression model, using an exogenous array of dummy encoded variables, representing the seasonal pattern of hour and day_of_week. The inclusion of such arrays significantly improved the long term forecasts in the case of hourly predictions of 168 observations, while only including a seasonal component in a SARIMA specification signaled a significant decay of the seasonality after only a couple of observations. In the monthly forecasting of the kWh series with a forecasting horizon of 12, the ARIMAX model performed worse than its more complex variants SARIMA and SARIMAX.

In all cases the LSTM model was the most difficult and computationally expensive to train, with training times that took far more time than the less complex models, and performed worse in terms of RMSE on the test set predictions. However, this does not invalidate the application of the LSTM on these problems. Previous works have shown that correctly carried out hyperparameter tuning can significantly increase the performance of LSTM models on similar data.

In general, forecasting a single residential electricity consumption series is rather difficult to achieve with low error, as most achieved forecasts on the test set were significantly over-underestimating the actual series.. However in some cases, the seasonal pattern was able to be modeled by the ARIMAX, SARIMAX, and LSTM while in other cases such as the monthly forecasting in section 4.3 the general decreasing trend from winter months to summer months was detected.

This paper could be used by other researchers and individuals to gain insights into the performance related to multiple advanced time series methods on similar data and it is believed to spark further research interests such as the inclusion of more exogenous features, perhaps based on geolocation, which account for the absence of the students at a given point in the future, as it is believed to significantly improve the forecasting performance of the so far best performing ARIMAX and SARIMAX models. Furthermore, the forecasts achieved could be paired with auxiliary forecasts for the price of electricity to calculate the the future costs that households will have to carry to keep their lights on, therefore, enabling them to reduce uncertainty about their future cost of living.

7. REFERENCES

Atef, S., & Eltawil, A. B. (2020, October).
Assessment of stacked unidirectional and bidirectional long short -term memory networks for electricity load forecasting. Electric Power Systems Research, 187(NA), NA. https://doi.org/10.1016/j.epsr.2020.106489

Chadwick, L. (2021, October 25).
Europe's energy crisis: Five charts to explain why your bills might go up this winter. Euronews. Retrieved December 27, 2021, from https://www.euronews.com/2021/10/25/europe-s-energy-crisis-five-charts-to-explain-why-your-bills-might-go-up-this-winter

Chikobvu, D., & Sigauke, C. (2012).
Regression-SARIMA modelling of daily peak electricity demand in South Africa. Journal of Energy in Southern Africa, 23(3).

Choi, E., Cho, S., & Kim, K. (2020, February 4).
Power Demand Forecasting using Long Short-Term Memory (LSTM) Deep-Learning Model for Monitoring Energy Sustainability. MDPI. Retrieved December 27, 2021, from https://www.mdpi.com/2071-1050/12/3/1109/htm

Chollet, F. (2018).
Deep Learning with Python. Manning.

DMI. (2021, June 29).
Danish Meteorological Institute - Open Data. Confluence. Retrieved December 1, 2021, from https://confluence.govcloud.dk/display/FDAPI

Dubey, A. K., Kumar, A., García-Díaz, V., Sharma, A. K., & Kanhaiya, K. (2021).
Study and analysis of SARIMA and LSTM in forecasting time series data. Sustainable Energy Technologies and Assessments, 47. Sciencedirect. https://doi.org/10.1016/j.seta.2021.101474.

Elamin, N., & Fukushige, M. (2018).
Modeling and forecasting hourly electricity demand by SARIMAX with interactions. Energy, 165(B), 257-268. https://doi.org/10.1016/j.energy.2018.09.157.

Fernando, J. (2021, 10 05).
Correlation Coefficient Definition. Investopedia. Retrieved December 14, 2021, from https://www.investopedia.com/terms/c/ correlationcoefficient.asp

Hochreiter, S., & Schmidhuber, J. (1997).
Long Short-term Memory. Neural computation, 9(1735-80). 10.1162/neco.1997.9.8.1735.

Liu, X. (2015).
Methods and Applications of Longitudinal Data Analysis. Elsevier Science. https://www.sciencedirect.com/science/article/pii/B9780128013427000149

Morison, R. (2021, November 28).
Europe's Energy Crisis Is About to Get Worse as Winter Arrives. Bloomberg.com. Retrieved December 27, 2021, from https://www.bloomberg.com/news/articles/2021-11-28/europe-s-energy-crisis-is-about-to-get-worse-as-winter-arrives

Ng, A. (2021, November 11).
U.S. official says Russia 'took advantage' of Europe's

energy crisis. CNBC. Retrieved December 27, 2021, from https://www.cnbc.com/2021/11/11/amos-hochstein-russia-took-advantage-of-europes-energy-crisis.html

Phi, M. (2018, 09 24). Illustrated Guide to LSTM's and GRU's: A step by step explanation. Towards Data Science. Retrieved December 27, 2021, from https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

Phosgene89. (2021). From AR to SARIMAX: Mathematical Definitions of Time Series Models. From AR to SARIMAX: Mathematical Definitions of Time Series Models. Retrieved December 26, 2021, from https://phosgene89.github.io/sarima.html

pmdarima. (2021). pmdarima.arima.auto_arima — pmdarima 1.8.4 documentation. alkaline-ml. Retrieved December 26, 2021, from https://alkaline-ml.com/pmdarima/modules/generated/pmdarima.arima.auto_arima.html

Tensorflow. (2021, November 11). Time series forecasting. TensorFlow. Retrieved December 14, 2021, from https://www.tensorflow.org/tutorials/structured_data/time_series

Wichern, D. W., & Hanke, J. E. (2013). Business Forecasting (9th ed.). Pearson.