Università
della
Svizzera
italiana

Faculty
of
Informatics

Bachelor Thesis

May 16, 2024

Time Series Forecasting of European Energy Market

European Market and Renewables

Hun Rim

Abstract

The ambitious energy targets, accelerated by the recent energy crisis, are driving the European Union to increase the share of renewable energy in gross energy consumption to 42.5% by 2030 from the current 23%. However, the intermittent and seasonal nature of renewable energy sources presents challenges in predicting their production capacity. The ability to accurately forecast the evolution of renewable energy's stake in the dynamic and ever-evolving energy market is a critical component in the decision making process of policy makers, and market participants alike. This project aims to explore and evaluate the performance of well-established forecasting methods in anticipating the trends of individual renewable energy components, ultimately contributing to the fostering of a balanced, sustainable, and reliable energy market in the EU. The primary focus is to assess auto-regressive forecasting methods and advanced models incorporating moving-average, exploit seasonality of time series data, or those utilising the correlation with exogenous variables. The results are presented for data considering recent history of the most significant enegy component at the European energy markets.

Advisor
Prof. Olaf Schenk
Assistant
Title Juraj Kardos

Advisor's approval (Prof. Olaf Schenk): Date:

1 Introduction

Policymakers and other energy market participants often stress that they need the grid operational statistics faster, particularly the European energy supply including renewable energy components, and consumption. The goal of the project is to develop a framework for forecasting the individual energy components contributing to the overall energy balance. The framework consists of well-established forecasting methods which can be easily applied to the energy market data, compare the forecast of various methods and asses the their reliability.

Like a staple good, we have steady demand for energy in our daily lives unlike luxury goods. When visualized, monthly energy demand reveals distinct patterns, with cyclical peaks occurring during winter and lows in summer, notably in July and August. These observations highlight Europe's heightened energy consumption in colder seasons and highlight the data's **seasonality**, suggesting the possibility of translating these patterns into mathematical models for forecasting future values.

1.1 Auto-Regressive Integrated Moving-Average

ARIMA model is one of the most widely used statistical forecasting model. It is a generalized version of AutoRegressive Moving-Average (ARMA) model and difference between the two models is their ability to handle non-stationary data. Stationarity indicates constancy of statistical property of a dataset such as variance, or mean over time. When data is stationary, it guarantees the stability of underlying patterns in the dataset, and eradicates the concern of the pattern being a result of random fluctuation. ARIMA model is capable of adjusting the data to provide this stability unlike ARMA.

Autoregression is a technique used in time series analysis that assumes temporal trend between the values in the dataset, and uncover a function of order p. In the auto-regression model, the variable regresses against itself. Put simply, the current value is calculated using past values, by estimating the influence of these past values on the present value. However, even if we assume the general trend is captured perfectly, there are inconsistent white-noise in the data which are not replicable. ARIMA model utilizes Moving-average component to estimate the white-noise in the data by studying the temporally closest known white-noises, and capture the short-term fluctuations which are not reflected by only studying the general trend. All these factors are combined to formularize ARIMA model ($ARIMA(p, d, q)$) as the following:

$$\hat{L}_N^{(d)} = C + \sum_{i=1}^{p} \psi_i L_{N-i}^{(d)} + \sum_{j=1}^{q} \varphi_j \epsilon_{N-j} + \epsilon_N \tag{1}$$

1.2 Seasonal ARIMA

As mentioned, EU energy market data often exhibits seasonal patterns, and Seasonal-ARIMA model is an extension of the standard ARIMA model which incorporates the seasonal component, and specifically handles seasonal patterns of the time series data. Essentially, this model does not stop at reflecting lags but extends to reflects on same period of previous years to certain degree if there is an annual periodicity. Similar to the representation of the ARIMA model, SARIMA model is denoted as $ARIMA(p, d, q)(P, D, Q)_m$, and can be formularized as the following:

$$ARIMA(p, d, q)(P, D, Q)_m = \hat{L}_N^{(d)} + \sum_{i=1}^{P} \Phi_i L_{N-im}^{D} + \sum_{j=1}^{Q} \Theta_j \epsilon_{N-jm} \tag{2}$$

Just as (p, d, q) represents the order of autoregression, integration, and moving-average, $(P, D, Q)_m$ represents the components but in a seasonal scale. The m denotes length of season, which in our case will be 12 months.

1.3 Exogenous Variables

The EU power sector, just like any other power sector, experiences volatility due to external factors such as the climate. No matter how accurate the ARIMA model is at forecasting, it has no bypass to account for the forecast errors originating from external factors, and this also applies to SARIMA model. Hence, the framework consists of a statistical model extending the ARIMA called SARIMAX (SARIMA with eXogenous variables) which also is capable of accounting for both seasonality, and errors due to the external factors. In this model, the external variable which has influence on the European power sector such as the economic or climate indicators can be factored in to the appropriate time frame. When forecasting the market movements, instead of soley relying on temporal correlation and statistical factors, the SARIMAX model can additionally reflect on the previous time frames with similar exogenous indicators to produce a better forecast.

$$SARIMAX(p,d,q)(P,D,Q)_m = ARIMA(p,d,q)(P,D,Q)_m + \sum_{i=1}^{k} \beta_i Z_i \tag{3}$$

As it can be seen in the equation above, the SARIMAX model is built on SARIMA model with additionally accounting for k exogenous variables.

1.4 Deep-Learning with Recurrent Neural Network

Along with predictive statistics, deep learning (DL) is frequently used in forecasting. Especially, DL using Recurrent-Neural-Network (RNN) with Long Short-term Memory (LSTM) layers, as shown in figure **??**, is very popular with time series forecasting. RNN repeats the process of aggregation and activation unlike normal neural networks, and LSTM extends the memory of RNNs, and allows the capturing of long-term dependencies. This regression like property of RNN, and the extension of its capability from LSTM layer makes it very effective when building a time series forecasting model. Furthermore, it would work as a good standard to evaluate the performance of statistical methods, as it takes a different approach from them.

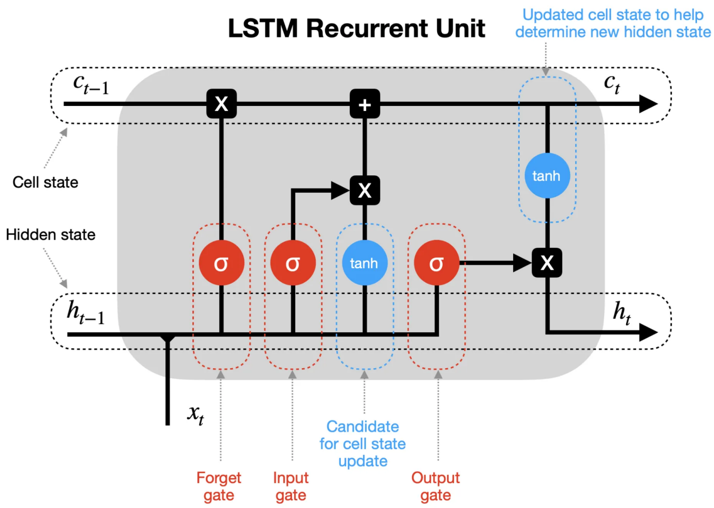

Figure 1. Structure of LSTM Recurrent Unit

2 Exper

sd