Code Assignment 1

Vyacheslav Kim 20182011

September 25, 2022

1. The link to the downloaded data can be founded here. The data chosen for the assignment 1 describes the changes of various commodity price over the period from 2000 to these days. Among of all models, the wheat price model was chosen as time series data for this assignment. In a CSV file, the price of wheat is given in EURO per bushel. For simplicity, in this assignment only the time period from 2016 and 2022 is being considered. In addition, a wheat price was recorded per each working day from Monday to Friday. Considering the changes of wheat prices for the past 6 years, we can notice that time series shows a quadratic trend with steady increase from 2016 (470 EUR/bushel) to the mid of 2020 (580 EUR/bushel) followed by a rapid rise until the first quarter of 2022 (800 EUR/bushel) and with a drastic boom up to 1400 EUR/bushel in the same time period. Taking in mind only this particular time period, we might conclude that time series has a quadratic trend, however, after the boom the price got back to 800 EUR/bushel and, thus, the trend might not fit into our model. My interest is to see which trend will properly fit into this time series and how well it describes its data.

2. Since cosine trend has the least MSE among other trends, it can be chosen as a deterministic trend for our model (Figure 1). The parameters for our trend are following:

$$f = \frac{0.001}{2\pi}$$
$$\mathbf{w} = [-178.1091206, 1.50949334, -1359.93352793, 604.47127268] \tag{1}$$

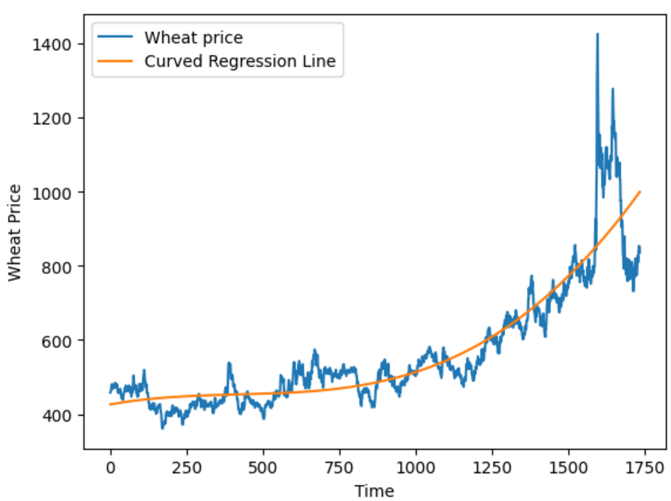

Figure 1: Cosine Trend on Time Series

You might notice that our trend doesn't exactly fit into our model especially after 1600 timestamp, thus, extreme datapoints are not properly described by the chosen trend.

3. Figure 2 shows standardized residuals for our time series. We can see that values before 1600's timestamp are distributed randomly with equal variance. However, after 1600's timestamp the residual curve acquires a different shape reaching to maximum and minimal extremes in a short time period. It's also worth to notice that our time series data shown in thew histogram (Figure 3) is not distributed exactly like a normal distribution but datapoints are rather concentrated around mean.

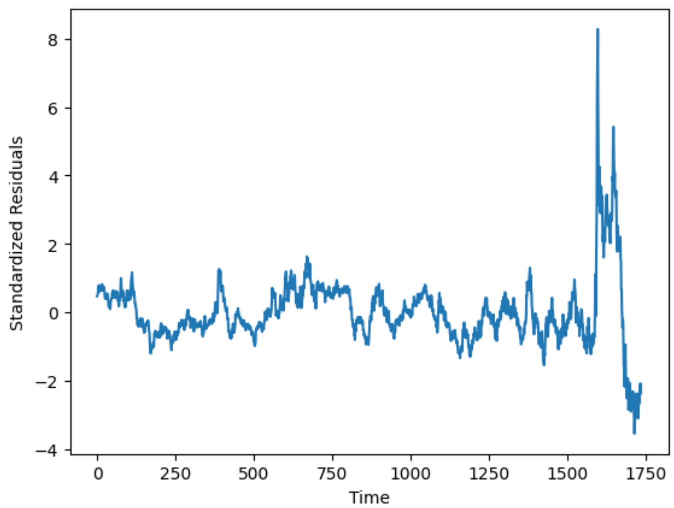

Figure 2: Standardized Residuals

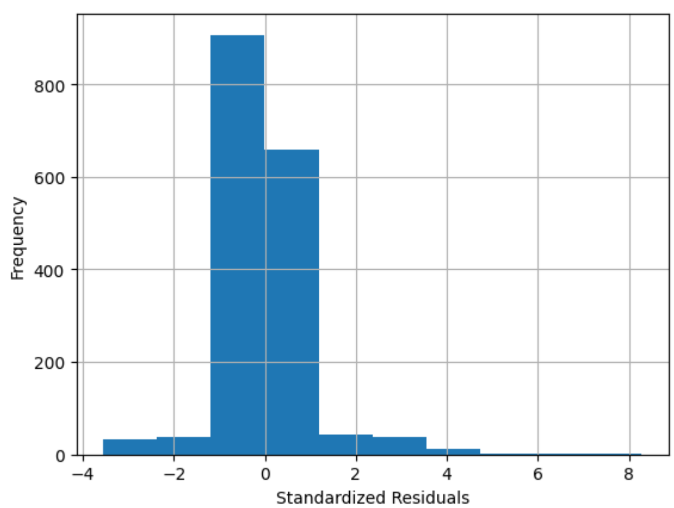

Figure 3: Histogram of Standardized Residuals

In the Q-Q plot (Figure 4) we can see that lower values of residuals appropriately fit into theoretical quantities with minor deviations but with higher values the points diverges stronger. In addition, autocorrelation function shows a strong dependence between consecutive time series datapoints which gradually decreases as lags are increasing. Thus, the chosen trend does not fit into our time series data since there is still a strong correlation between consecutive datapoints and, hence, residuals are not independent and there still exists an undefined trend.

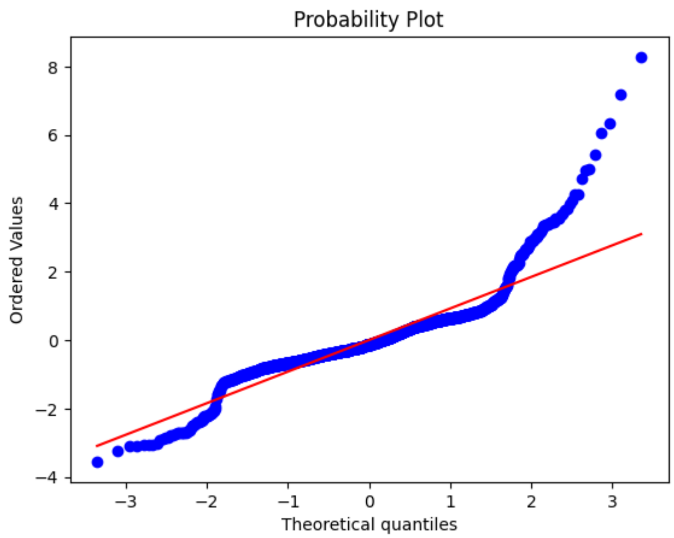

Figure 4: Q-Q plot: Standardized Residuals

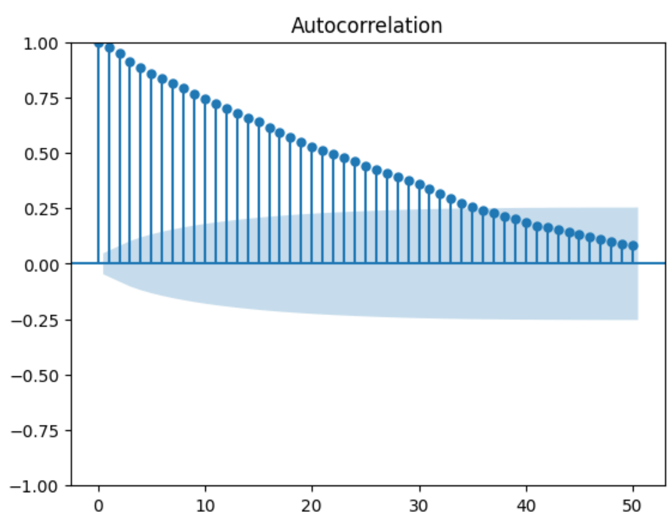

Figure 5: AFC of Residuals