**Report Quantitative Analysis FTSE ESG Index**

Author: Ion Adina FMI UNIBUC ANUL 3

Contents

[1.Cleaning Data 1](#_Toc165991639)

[2. Working with data 1](#_Toc165991640)

[Trend analysis 1](#_Toc165991641)

[Lookback period: 1 year or entire history? 3](#_Toc165991642)

[3. Cross-validation and forecasting models 4](#_Toc165991643)

[ARIMA models 4](#_Toc165991644)

[ACF and PACF results. 4](#_Toc165991645)

[Comparing different ARIMA models. 5](#_Toc165991646)

[Normality of error results 6](#_Toc165991647)

[Forecast results ARMA(1,1) 7](#_Toc165991648)

[Regression or autoregression? Which choice is better? 7](#_Toc165991649)

[4. Conclusions 7](#_Toc165991650)

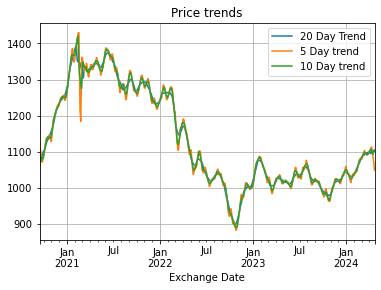
# 1.Cleaning Data

The data were sorted chronologically in ascending order, then the duplicates were removed by making a list from which I removed the element with the index 847 (corresponding to 10th December 2020).  
 Then I removed the fictitious date of February 30 2024, and filled the missing data by linear interpolation.

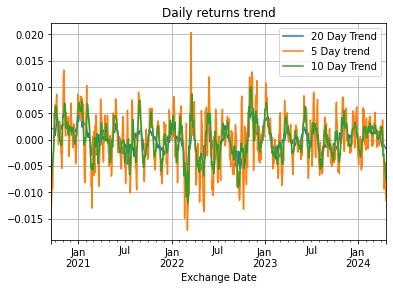
# 2. Working with data

## Trend analysis

*Figure 1: Price trend (Moving averages method)*



*Figure 2: Daily return trends*



As one can see, there is no actual trend in the returns. Therefore it is not useful differencing (e.g. using ARIMA models) for forecasting the daily returns.

**I also propose to use Tau’s kendall statistics to test the trend:**

***Methodology of calculation:***

= number of concordant pairs

= number of discordant pairs.

***Interpretation*:**

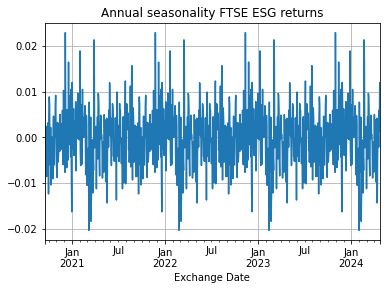
**If is close to , then there is a strong evidence of negative trend in data.**

**If isclose to +1, then there is a strong evidence of positive trend in data.**

## Lookback period: 1 year or entire history?

Before answering the question, I show below the seasonality component by using 1 year moving averages of daily returns, and one can observe a ciclicity of 1 year in the returns.

Considering also the lack of trend, **1 year of data will suffice for analysing and forecasting the 1 day returns for the next 1 week period.**



# 3. Cross-validation and forecasting models

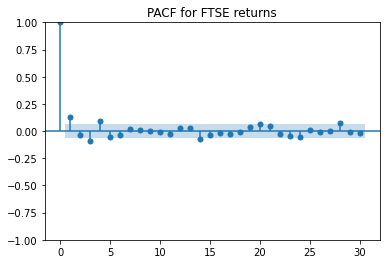
## ARIMA models

In my project, implementing ARMA and ARIMA models could enhance the forecasting capabilities for the FTSE index. These models offer a systematic approach to analyzing historical data, identifying patterns, and predicting future trends. By integrating them into project, we can refine insights, improve decision-making, and potentially optimize investment strategies based on more accurate forecasts of market behavior.

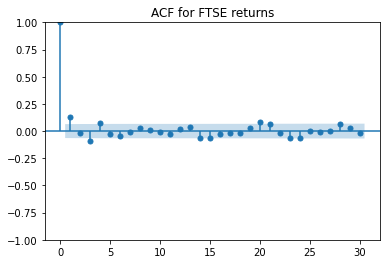
### ACF and PACF results.

I've devised an algorithm that facilitates the determination of optimal parameters for an ARMA(p,q) model by computing autocorrelation and partial autocorrelation functions (ACF and PACF) of the provided time series data. Subsequently, the algorithm employs the Augmented Dickey Fuller test to assess stationarity. If the test result falls below the 99% critical value, implying rejection of the null hypothesis and indicating data stationarity, it ensures accurate parameter selection for effective modeling and forecasting.

Figure 1: Partial Autocorrelation Function for FTSE returns for 30 lags.



*Figure 2: ACF for FTSE returns:*



In the **acf** and **pacf** arrays, each value shows the correlation between the time series and its lagged versions. The ACF indicates correlations at different lags, while the PACF shows correlations at each lag while removing the effects of shorter lags.

The Augmented Dickey-Fuller test (ADF) assesses stationarity in the time series. The low p-value and negative test statistic indicate strong evidence against non-stationarity, suggesting the time series is likely stationary.

### Comparing different ARIMA models.

**Table 1: RMSE = root means square error of prices forecasts for selected models.**

|  |  |  |
| --- | --- | --- |
| **Model** | **RMSE** | **Rank** |
| **ARMA(1,0)** | **17.12055** | **1** |
| **ARMA(2,0)** | **17.12701** | **4** |
| **ARMA(3,0)** | **17.12828** | **6** |
| **ARMA(1,1)** | **17.1266** | **2** |
| **ARMA(2,1)** | **17.12821** | **5** |
| **ARMA(1,2)** | **17.12699** | **3** |
| **ARMA(2,2)** | **17.12935** | **7** |
| **ARIMA(1,1,1)** | **17.1464** | **9** |
| **ARIMA(2,1,1)** | **17.55175** | **10** |
| **ARIMA(2,1,2)** | **17.13841** | **8** |

Table 2: RMSE for returns forecasts for the selected ARIMA models

|  |  |  |
| --- | --- | --- |
| **Model** | **RMSE** | **Rank RMSE** |
| **ARMA(1,0)** | **0.007770** | **9** |
| **ARMA(2,0)** | **0.007758** | **6** |
| **ARMA(3,0)** | **0.007758** | **5** |
| ARMA(1,1) | 0.007749 | 1 |
| **ARMA(2,1)** | **0.007757** | **4** |
| **ARMA(1,2)** | **0.007752** | **2** |
| **ARMA(2,2)** | **0.007757** | **3** |
| **ARIMA(1,1,1)** | **0.007770** | **10** |
| **ARIMA(2,1,1)** | **0.007758** | **7** |
| **ARIMA(2,1,2)** | **0.007769** | **8** |

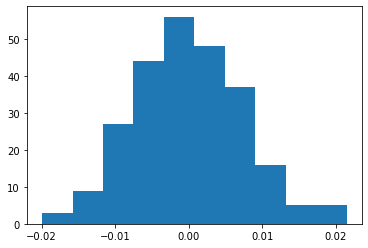
We choose ARMA(1,1) to forecast the returns because its rank is the smallest and this will be our forecasting model.

The fitted ARMA(1,1) model is: where the FTSE ESG return on date

To evaluate the performance of my model in predicting the last 20 days of data, I first use my trained model to make predictions for these days, ensuring that the training data does not include this period. Then, I compare these predicted values with the actual data for the same timeframe, employing evaluation metrics like Mean Absolute Error (MAE) or Mean Squared Error (MSE). I analyze the results to determine my model's effectiveness; close predictions suggest good performance. If the results are unsatisfactory, I consider adjusting my model's hyperparameters or exploring alternative model types for potential improvements.

### Normality of error results

The **standardized errors** of the fitted autoregression ARMA(1,1) can be considered as normally distributed as one can see in the below histogram:



I also have applied the Kolmogorov Smirnov test on the standardized errors:

|  |  |
| --- | --- |
| KS statistic | p-value |
| 0.0274 | 0.9892 |

One cannot reject the null hypothesis which assumes normality of the errors.

### Forecast results ARMA(1,1)

*Table 4: Forecasted returns for the next 5 days*

|  |  |
| --- | --- |
| **ARMA(1,1)** | **Forecasted 1D return** |
| **4/23/2024** | **0.000305** |
| **4/24/2024** | **0.000305** |
| **4/25/2024** | **0.000305** |
| **4/26/2024** | **0.000305** |
| **4/29/2024** | **0.000305** |

## Regression or autoregression? Which choice is better?

METHODOLOGY:

1. I build the regression model where and

The training data consists of returns 270 business days ago up to 20 business days ago.

1. I forecast on the remaining 20 days + extra 5 future days (up to 29th April 2024)
2. I use the estimated coefficients   and the forecasted   days in order to forecast the future (until th of April 2024)

**The RMSE of the linear regression is 0.0070, better than all the autoregressive results.**

The forecasted returns using exogenous FTSE index returns are in the below table:

|  |  |
| --- | --- |
| **Date** | **Forecasted return** |
| **4/23/2024** | **0.0002449** |
| **4/24/2024** | **0.0002449** |
| **4/25/2024** | **0.0002449** |
| **4/26/2024** | **0.0002449** |
| **4/29/2024** | **0.0002449** |

# 4. Conclusions

a. The best autoregressive model for forecasting the returns is ARMA(1,1) model.

b. The RMSE for the best autoregression of daily returns is 0.75% for the last 20 trading days. The RMSE results are however close one to another.

c. The RMSE for the Hybrid regression/autoregression model presented above is better than the pure autoregression model.