SC475 - Time Series Analysis Project

Vrund Leuva (202201199)
Meet Chauhan (202201262)
Time Series Analysis (SC475)
Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, India
Email: 202201199@daiict.ac.in, 202201262@daiict.ac.in

Under the guidance of: Prof. Mukesh Tiwari

***Abstract*—In this project, we have performed time series analysis on the Food Price Index(FPI) given by the FAO. The FAO Food Price Index (FFPI) is a measure of the monthly change in international prices of a basket of food commodities. It consists of the average of five commodity group price indices weighted by the average export shares of each of the groups over 2014-2016. Our objective is to do data visualization, decomposition, transformations, detrending, and deseasonalizing.**

I. Introduction

This dataset contains monthly values of the Food Price Index from the year 1990 to February 2025.

II. Food Price Index Dataset Description

Monthly Food Price Index (1990–2025)

Description:
This dataset presents monthly Food Price Index values spanning from 1990 to February 2025. The data reveals a distinct long-term upward trend along with seasonal variations, indicating both an overall increase over time and recurring monthly patterns.

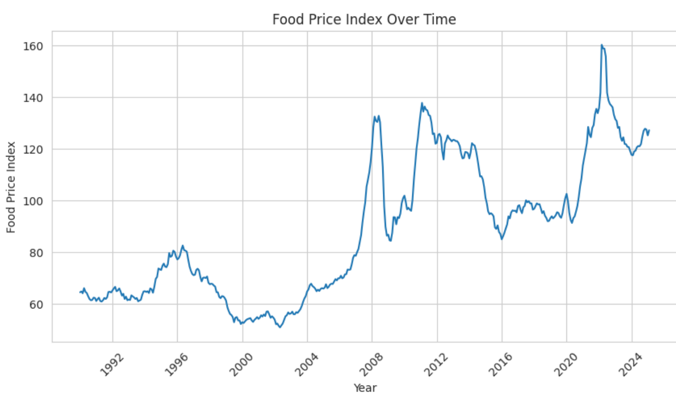

Fig. 1: Food Price Index Over Time (Line Plot)

Conclusion

There is a presence of strong seasonality in the dataset and there is a clear long-term increase in FPI over the years. There are certain months when the FPI remains high, such as the pre-harvesting period, as stocks might have run low which leads to high demand which leads to increase in FPI.

III. Rolling Statistics for Stationarity Check

In this section, the 12-month rolling mean smooths short-term noise, while the rolling variance highlights periods of high price volatility. These patterns indicate strong trends and fluctuations, making the data suitable for time series forecasting.

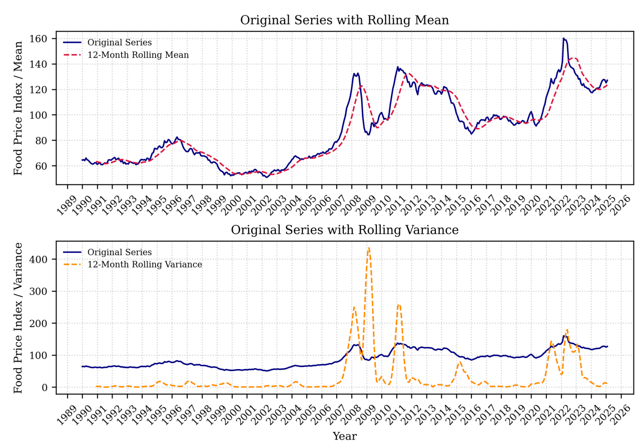

Fig. 2: Rolling Mean and Rolling Variance (Window=12) over the Original Food Price Index.

IV. Observations from Time Series Plots

- The series exhibits a clear upward trend, reflecting a steady rise in global food prices over the observed period.
- There is some evidence of seasonality, but it is less pronounced compared to the overall trend and major price spikes.
- The amplitude of fluctuations increases notably during specific periods (e.g., 2008, 2011, 2022), indicating that both the magnitude of price changes have grown during global crises.

A. Rolling Mean and Rolling Variance

- The **12-month rolling mean** (red dashed line) effectively smoothens the fluctuations.
- The **rolling variance** (orange dashed line) grows with time, confirming increasing variability in the data.
- These patterns indicate that the series is **non-stationary**, as both mean and variance change over time.

B. Conclusion

From the rolling plots:

- The rolling mean (red dashed line) generally trends upward over the years, indicating a non-stationary mean.
- The rolling variance (orange dashed line) exhibits significant fluctuations, especially during major price shocks, confirming that the variance is not constant over time.

V. Year-wise Monthly Trends (Mini Subplots)

In figure 3, We generated a grid of compact subplots, each representing monthly Food Price Index trends for a particular year from 1990 to February 2025.

- Each subplot shows the Food Price Index trend for a specific year, highlighting how values fluctuate across months.
- Many years display a recurring seasonal pattern, with noticeable rises around the early months (e.g., March–April) and slight declines mid-year (e.g., June–July).

Conclusion: The yearly subplot grid effectively reveals both seasonal fluctuations and a steadily rising trend in the Food Price Index. These insights reinforce the need for seasonality-aware and trend-capturing modeling techniques, while also confirming that the data is non-stationary.

VI. Rolling Variances for Various Transformations

This set of plots(Figure 4) visualizes the rolling variance for different transformations of the Food Price Index data, each with a 12-month rolling window.

- **Square Root Transformation**: The rolling variance of the square root-transformed Food Price Index is shown in purple (data) and dark orange (variance). This transformation moderately stabilizes the variance, though some fluctuations remain visible, especially in later years.
- **Logarithmic Transformation**: The log-transformed data is plotted in blue, with its rolling variance in orange. This transformation reduces variability more effectively, showing smoother and more consistent variance over time.
- **Box-Cox Transformation**: The Box-Cox transformed series is displayed in teal, with its rolling variance in gold. This method provides one of the most stable variance patterns, suggesting it's effective at variance stabilization.
- **Seasonal Differencing**: The plot shows the seasonally differenced data in crimson, with the rolling variance in black. Despite removing seasonality, this method results in sharp spikes in variance, especially around major global events, indicating instability in some periods.

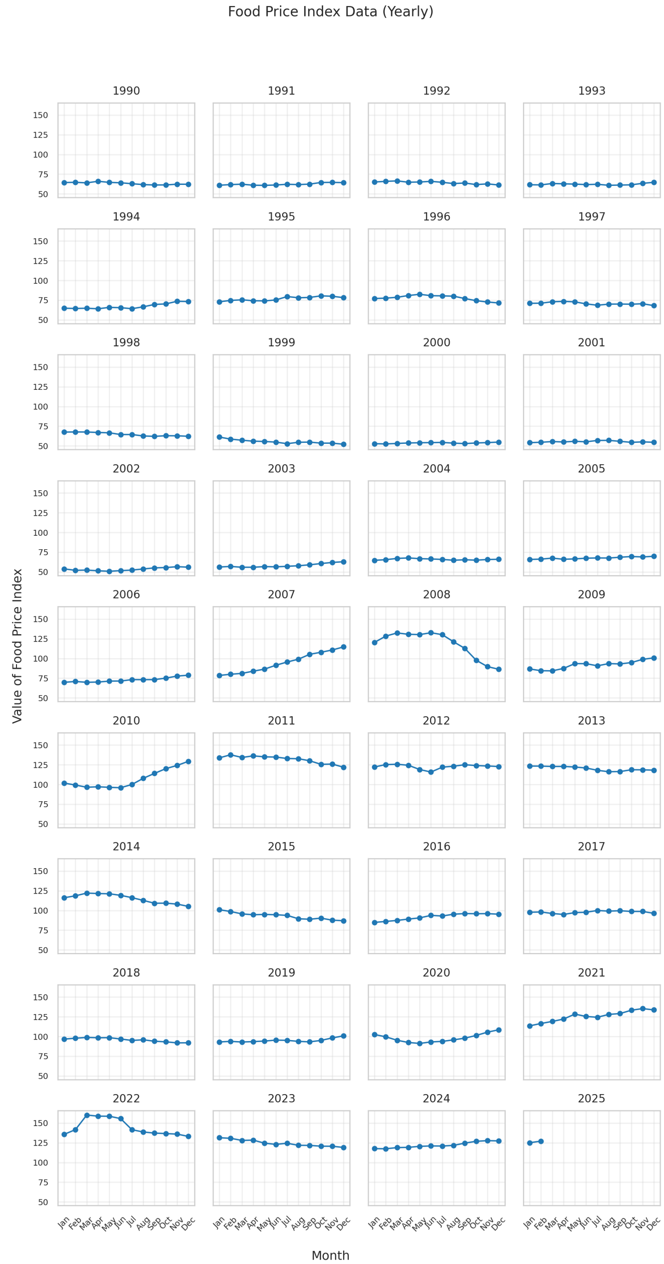

Fig. 3: Monthly Food Price Index Trends for Each Year (1990–2025)

Conclusion: The logarithmic transformation and box-cox suits the best for stabilizing the variance from the other two approaches. Both transformations almost give a similar output.

VII. Rolling Variance Comparison (Log vs. Box-Cox)

This plot(Figure 5) compares the 12-month rolling variance of the log-transformed and Box-Cox-transformed Food Price Index data.

- **Log Transformation**: The rolling variance for the log-transformed data is depicted by the blue dashed line.

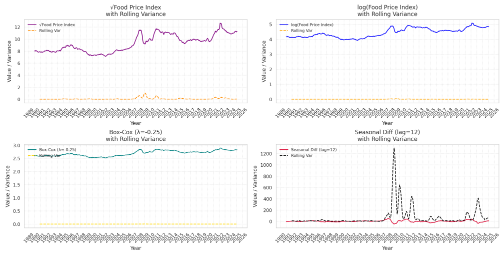

Fig. 4: Rolling Variance for Various Transformations of the FPI

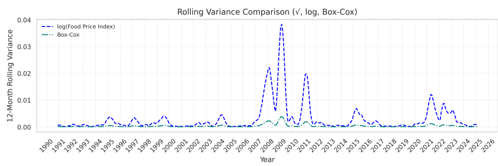

Fig. 5: Comparison of 12-Month Rolling Variance for Log and Box-Cox Transformed Food Price Index

While it captures major fluctuations, it also shows significant spikes in variance during volatile periods (e.g., 2008–2011 and 2021–2023).

- **Box-Cox Transformation**: The rolling variance of the Box-Cox transformed data is shown with a green dash-dot line. Compared to the log transformation, the Box-Cox approach maintains a more stable and smoother variance profile across most of the time span.

Conclusion: The Box-Cox transformation demonstrates greater stability in rolling variance compared to the logarithmic transformation, particularly during turbulent periods. While both methods reduce variance fluctuations, Box-Cox appears more effective in stabilizing variance, suggesting it may be the better choice for further modeling.

VIII. Analysis of Time Series Transformations

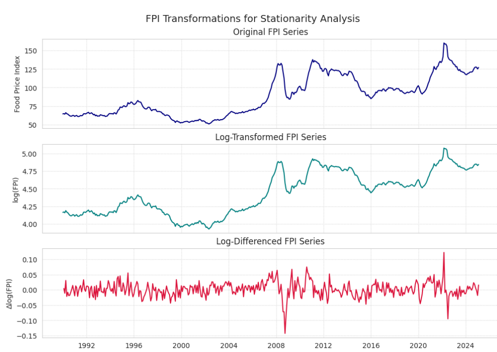

Fig. 6: Data Transformations Plots

This figure shows the effects of log-transformation on the FPI data:

- **Original FPI Data (Top)**:
  - Strong upward trend, increasing seasonal fluctuations, and clear heteroscedasticity.
- **Log-Transformed FPI Data (Middle)**:
  - Linearized trend with stabilized variance and consistent seasonal fluctuations.
- **Log-Transformed FPI Series (Bottom)**:
  - Removes both trend and changing variance, making the series more stationary.
  - Residual fluctuations are more consistent and suit able for ARIMA(1,1,3) modeling.

ADF Test Results: To evaluate stationarity, the Augmented Dickey-Fuller (ADF) test was conducted on the transformed Food Price Index series:

ADF Test Results on Log-Transformed Food Price Index:

- **Before Differencing:**

$$\text{ADF Statistic} = -1.7052328520654347$$
$$\text{p-value} = 0.4284183908772829$$

Interpretation: The p-value is greater than 0.05, so we fail to reject the null hypothesis. The series is likely non-stationary.

- **After Differencing:**

$$\text{ADF Statistic} = -6.895812813292611$$
$$\text{p-value} = 1.319829175805629e-09$$

Interpretation: The p-value is less than 0.05, and the ADF statistic is far below critical values. This confirms that the differenced log-transformed series is stationary.

IX. QQ-Plot Analysis

Fig. 7: QQ-plot of Food Price Index And Log Difference of Food Price Index

Right Plot (Log-Differenced Food Price Index Series): After applying a logarithmic transformation followed by differencing, the data points align more closely with the theoretical quantile line. This suggests that the transformed series approximately follows a normal distribution, satisfying one of the key assumptions for ARIMA(1,1,3) modeling.

X. ACF PACF Analysis

1. ACF and PACF of Original Food Price Index Series (Top Row):

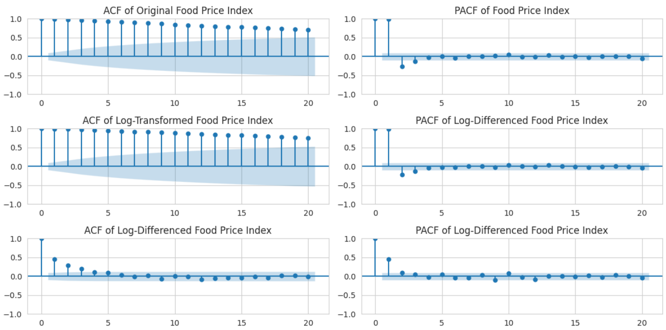

Fig. 8: ACF And PACF Plot

- **ACF:** The autocorrelation function shows a very slow exponential decay, suggesting the presence of strong trend and non-stationarity in the original Food Price Index data.
- **PACF:** The partial autocorrelation function has a significant spike at lag 1 and then cuts off more quickly. This indicates that while some short-term dependency exists, the series is still non-stationary.
- **Conclusion:** The original series is non-stationary and not suitable for ARIMA(1,1,3) modeling in its raw form.

2. ACF and PACF of Log-Transformed Food Price Index Series (Middle Row):

- **ACF:** The ACF still shows strong autocorrelation across many lags, though the pattern is slightly more dampened than in the original series.
- **PACF:** Significant spikes in the first few lags are still visible, suggesting persistent autocorrelation and continued non-stationarity.
- **Conclusion:** Log transformation alone is insufficient to achieve stationarity; differencing is still required.

3. ACF and PACF of Log-Differenced Food Price Index Series (Bottom Row):

- **ACF:** ACF cuts off quickly after lag 3, which is characteristic of a moving average (MA) process.
- **PACF:** PACF also drops sharply after lag 1, indicating a possible autoregressive (AR) component at lag 1.
- **Conclusion:** The log-differenced series appears to be stationary, as the autocorrelations are much weaker and short-lived. This transformation prepares the data well for ARIMA(1,1,3) modeling.

XI. DECOMPOSITION OF LOG(FOOD PRICE INDEX)

- Any time series consist of the components **trend** , **seasonality** and **Residual**

A. Choice of Decomposition Model

There are two primary types of time series decomposition:

1) **Additive Decomposition**
2) **Multiplicative Decomposition**

In additive decomposition, the components of a time series are assumed to combine linearly:

$$Y_t = T_t + S_t + R_t$$

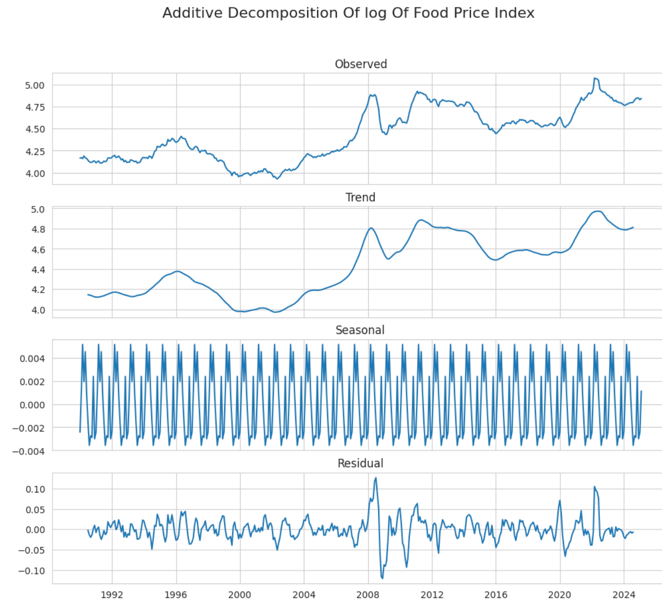

Fig. 9: Trend , Seasonal Residual Decompose

In our case, the trend component is not proportional to the seasonal component, and the seasonal variation remains relatively constant over time. Hence, **additive decomposition** is more suitable for analyzing the log-transformed Food Price Index data.

B. Additive Decomposition of Log-Transformed Food Price Index

Figure 9 presents the additive decomposition of the log-transformed Food Price Index time series into its primary components: *Observed*, *Trend*, *Seasonal*, and *Residual*.

- **Observed:** The log-transformed food price index demonstrates noticeable fluctuations over time, indicating the presence of both long-term trends and recurring seasonal behavior. Spikes and dips are evident around 2008, 2011, and 2021–2022, suggesting the impact of global economic or supply chain disruptions.
- **Trend:** The trend component reveals the underlying long-term trajectory in food prices. A significant upward movement is seen from around 2005 to 2011, with another rise post-2020. These shifts may be associated with major events like financial crises, climate-related impacts on agriculture, or disruptions due to the COVID-19 pandemic and geopolitical tensions.
- **Seasonal:** A clear and consistent seasonal pattern is visible across the entire time span. The amplitude and regularity suggest a strong annual cycle, likely tied to agricultural production seasons, harvest cycles, and global demand fluctuations. This regularity justifies the inclusion of seasonality in predictive models.
- **Residual:** The residual component captures the irregularities and random noise remaining after removing trend and seasonality. While most fluctuations hover close to

zero, pronounced anomalies are observed around 2008 and 2022, possibly reflecting shocks such as economic crashes or conflict-related supply issues. Overall, the residuals appear relatively stationary, which is favorable for time series forecasting.

XII. Detrending Of log(Food Price Index)

Detrending Methods:

- Differencing the log-transformed series.
- Moving average of the log-transformed series.
- Using In-Built Function
- Regression Method
- **Detrended (First Difference of log(Food Price Index))** :
  - After applying first-order differencing to the time series, the trend component is effectively removed, resulting in a stationary series suitable for further time series modeling.

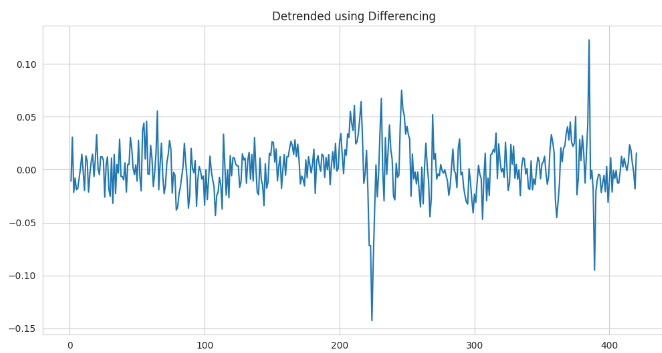

Fig. 10: Detrended (First Difference of log(Food Price Index))

- **Detrended (Food Price Index) - Moving Average)** :
  - Moving average stabilizes the fluctuations, and the variance shows less volatility after removing the trend using a moving average.

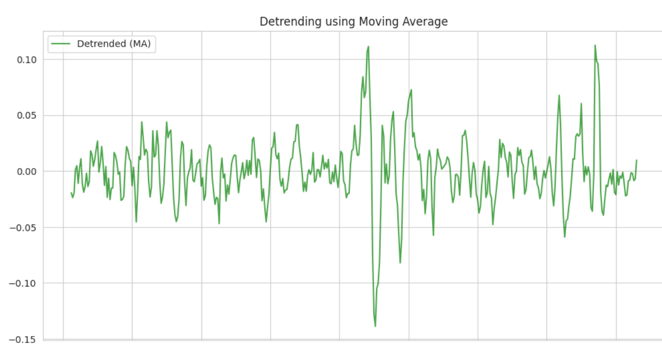

Fig. 11: Detrended (log(Food Price Index) - Moving Average)

- **Detrended (Food Price Index) - Using In-Built Function)** :
  - To better understand and isolate the components of the Food Price Index time series, the built-in `seasonal_decompose` function from the `statsmodels` library was employed using an additive model. The decomposition separates the time series into trend, seasonal, and residual components.

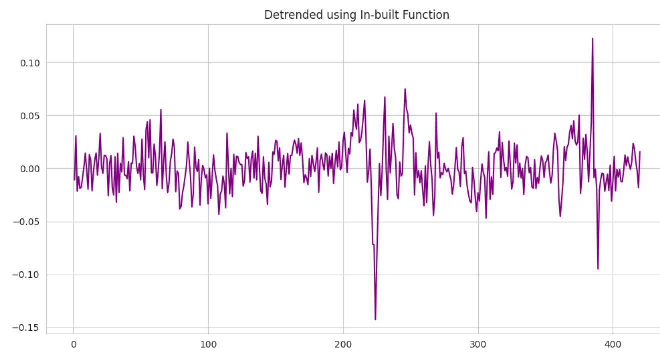

Fig. 12: Detrended (log(Food Price Index) - Decomposition Trend)

- **Detrended (Food Price Index) - Regression Method)** :
  - In this method, an in-built polynomial regression function of degree 1(Linear) is applied to model the underlying trend in the original Food Price Index data.
  - The fitted polynomial curve (red dashed line) represents the estimated trend, which is subtracted from the original data to obtain the detrended series (blue line).

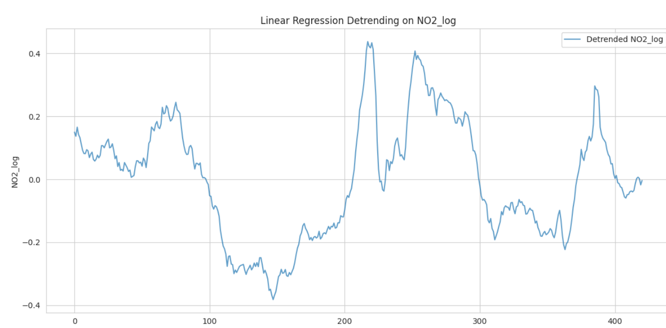

Fig. 13: Detrended (Food Price Index) -Regression Method)

Conclusion: The rolling mean and variance of the detrended data help assess the effect of various detrending methods. The results indicate stabilization of the trend, with reduced volatility in the moving averages after the trend removal process.

XIII. Deseasonalized Of log(Food Price Index)

Deseasonalized Methods:

- Local Trend Method.
- Moving average of the log-transformed series.
- Lag Operator Method

- **Deseasonalized (Local Trend Method)** :
  - The Local Trend Method is a smoothing-based technique that separates the time series into trend, seasonal, and residual components.
  - It identifies and removes seasonal effects by estimating local trends using moving averages or locally weighted regression (LOESS).
  - This method adapts to short-term fluctuations, making it suitable for time series with non-constant seasonal patterns.
  - As shown in the graph, once the seasonal component is removed, the resulting series (in orange) captures the underlying trend and irregular variations without periodic seasonality.

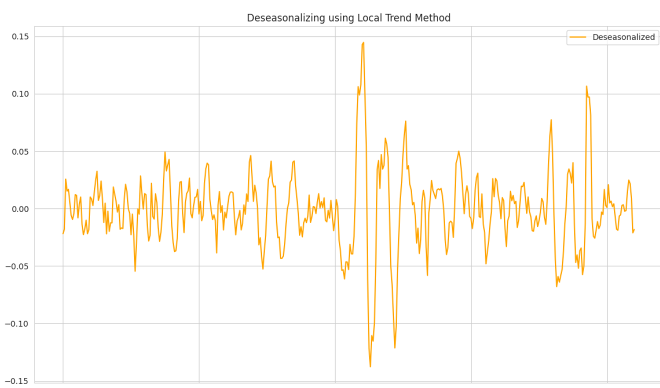

Fig. 14: Deseasonalized (Local Trend Method)

- **Deseasonalized (Food Price Index) - Moving Average)** :
  - The **original series** (shown in blue) includes both trend and seasonal variations.
  - A **centered moving average** (orange dashed line) is calculated to estimate the seasonal component by smoothing out short-term fluctuations.
  - The **deseasonalized series** (green line) is obtained by subtracting the estimated seasonal component from the original series:

$$Y_{\text{deseasonalized}} = Y_{\text{original}} - Y_{\text{seasonal}}$$

  - This helps isolate the trend and irregular components, making further modeling or analysis more effective.

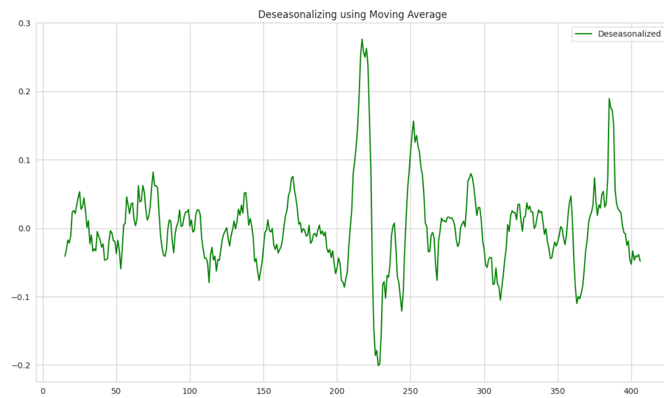

Fig. 15: Deseasonalized (log(Food Price Index) - Moving Average)

- **Deseasonalized (log(Food Price Index) - Lag Operator Method)** :
  - The **original series** (blue line) contains both seasonal and trend components.
  - We apply a lag difference, typically using a lag corresponding to the seasonal period (e.g., 12 for monthly data):

$$Y_t^{\text{deseasonalized}} = Y_t - Y_{t-s}$$

    where s is the seasonal lag.
  - The resulting **deseasonalized series** (purple line) effectively removes the repeating seasonal structure, highlighting the trend and irregular components.

Conclusion: Deseasonalization is essential for analyzing the underlying patterns in a time series by removing seasonal effects. The Moving Average method smooths the data to extract seasonal components, while the Local Trend method isolates and removes both local trend and seasonality using smoothing filters. The Lag Operator method eliminates seasonality by differencing the series at a fixed seasonal interval. Each method has its strengths, with selection depending on the nature of the data and analysis goals. All approaches ultimately help in revealing the true structure of the series for better forecasting and modeling.

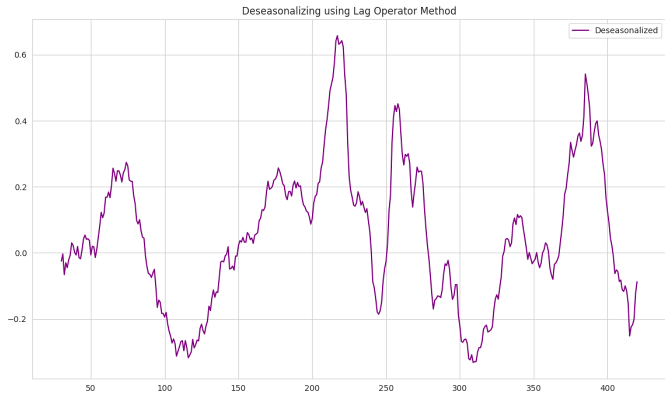

Fig. 16: Deseasonalized (log(Food Price Index) - Lag Operator Method)

XIV. References

References

[1] Peixeiro, Marco. Time Series Forecasting in Python. Simon and Schuster, 2022.
[2] Hyndman, R.J., Athanasopoulos, G. (2018). Forecasting: Principles and Practice.
[3] Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press.