Sales Forecasting Using Time Series Decomposition and Predictive Modeling in MATLAB

1. Problem Identification:

Time series analysis plays a critical role in engineering, especially for tasks such as capacity planning, demand forecasting, predictive maintenance, and production optimization. In predictive maintenance, it allows engineers to monitor sensor data such as vibration and temperature to foresee equipment failures and schedule timely interventions, reducing downtime and costs. In manufacturing and supply chain operations, demand forecasting through time series helps optimize inventory, streamline production, and allocate workforce efficiently. For energy management, utilities rely on time series models to anticipate electricity and gas consumption, improving generation and distribution strategies. Civil engineers use it to monitor structural stress over time, aiding in the early detection of deterioration in infrastructure like bridges and pipelines, thus prioritizing maintenance. In quality control, time series-based statistical process control detects anomalies in product quality, guiding adjustments in production parameters. Furthermore, in environmental engineering, forecasting pollution levels, water usage, and climate patterns through time series supports sustainable resource planning and environmental protection.

The given time series forecasting challenge involves analyzing 36 monthly observations that exhibit four essential components. First, the trend reflects a long-term upward or downward movement in the data, such as consistent sales growth over time. Second, seasonality captures repeating patterns that occur every 12 months, like increased sales during holiday seasons. Third, cyclical variations represent medium-term fluctuations that span more than a year, often influenced by broader economic cycles. Finally, random noise consists of unpredictable and irregular variations caused by external events, such as sudden supply chain disruptions. The objective of this analysis is to decompose the series into these individual components, understand their interactions, and develop a reliable model to forecast sales for the next six months using a decomposition-based approach.

Mathematical models play a vital role in transforming raw time series data into actionable insights by systematically analyzing its underlying components. Through decomposition, the series is broken down into trend (T_t), seasonality (S_t), cyclical variations (C_t), and residual noise (R_t). This can be represented using either a multiplicative model $Y_t = T_t*S_t*C_t$ or an additive model $Y_t= T_t+S_t+C_t$, depending on the nature of the data. Decomposition helps uncover hidden patterns, such as quantifying the impact of holiday sales or promotional campaigns. Moving beyond decomposition, forecasting mathematical techniques such as SARIMA (Seasonal ARIMA) and Exponential Smoothing (ETS) can be employed to extrapolate trends and seasonal effects, while also capturing autocorrelation structures**.** Classical regression models also play a crucial role by establishing linear relationships between time-dependent variables and external predictors, offering interpretable insights for forecasting. Additionally, advanced machine learning models like XGBoost, RNN (Recurrent Neural Networks), and LSTM (Long Short-Term Memory) networks are capable of modeling complex, nonlinear relationships within the data. The insights derived from these forecasts support strategic decision-making across various domains, guiding inventory management, marketing budget allocation, and workforce planning.

2. MATLAB Implementation

The MATLAB based implementation file is attached as "*script.m*". Below are some plots generated based on implementation of this script:

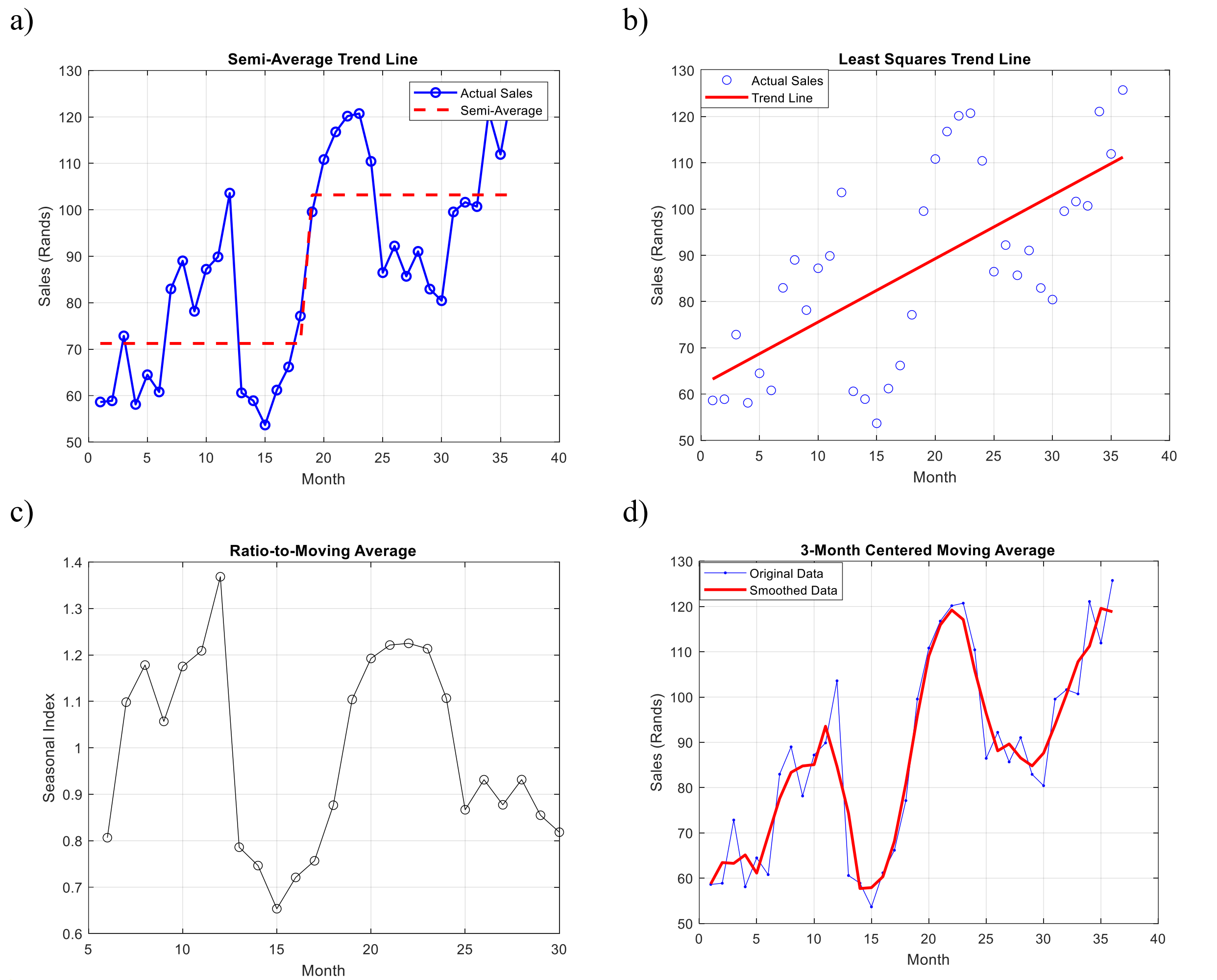

Figure 1 Components of given Time Series

3. Analysis and Evaluation:

The overall quality of the models applied to the sales data demonstrates a solid understanding of trend and seasonality analysis. The semi-average trend line (Figure 1a) divides the data into two clear phases and captures a general shift in sales behavior, but it oversimplifies the trend by using only two averages. The least squares trend line (Figure 1b) provides a more accurate linear representation of the upward trend over time, effectively modeling the long-term direction of sales growth. However, this linear model does not account for repeating seasonal fluctuations that are evident in the data. The ratio-to-moving average plot (Figure 1c) highlights strong seasonal effects, with consistent peaks and troughs recurring across fixed intervals, confirming the presence of seasonality in the dataset. The 3-month centered moving average (Figure 1d) further smooths short-term irregularities and exposes the cyclical nature of the data, making it easier to identify underlying patterns. The residual scatter plot (Figure 2) supports this finding, rather than being randomly scattered around zero, the residuals follow a wave-like pattern. This non-random structure in the residuals indicates that the model is missing key seasonal or cyclical components. In conclusion, while the trend models successfully capture the overall direction of the data, the residual analysis suggests that a more advanced model incorporating both trend and seasonality would better represent the true structure of the sales series and improve forecast accuracy. To achieve this, a non-linear model—such as seasonal exponential smoothing, Holt-Winters method, or SARIMA, should be considered to capture the complex dynamics present in the data.

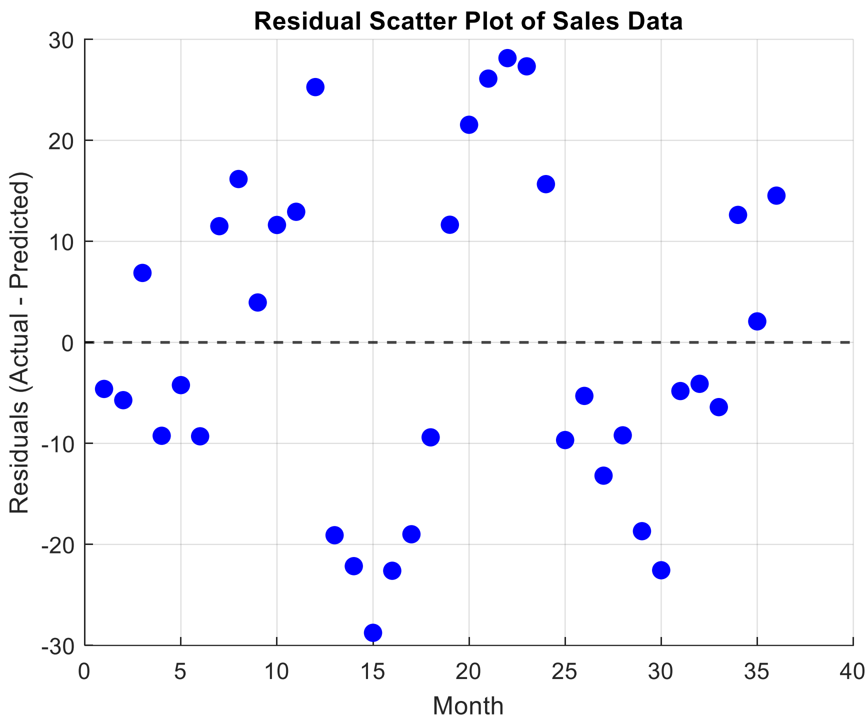

Figure 2 Residual Scatter Plot

4. Forecasting and Interpretation:

The forecasted sales for the next 6 months have been generated using a least-squares linear trend model, as shown in the Figure 3. The blue line represents the historical sales data over the past 36 months, the green dashed line indicates the fitted linear trend, and the red points illustrate the forecasted sales for months 37 to 42. The trend line follows a steady upward slope, and the forecast extends that pattern, predicting a gradual increase in sales over the next half-year period. This model reflects the overall positive growth in sales and is particularly useful for understanding long-term directional trends. However, as the historical sales data clearly exhibits seasonal fluctuations, the trend-only model does not fully capture these repeating patterns. The forecasted points lack the peaks and troughs seen in the actual data, indicating that the model oversimplifies the sales behavior by ignoring cyclical and seasonal components.

To improve forecast accuracy and capture the real dynamics of sales, more advanced models should be employed. Techniques such as Holt-Winters exponential smoothing, SARIMA, or even machine learning models (like XGBoost with time-based features) can incorporate both trend and seasonality. These models can better adapt to changing sales behavior, especially when consistent seasonal effects are present. In conclusion, while the trend model provides a good starting point for forecasting, future analysis should integrate non-linear models with seasonal adjustment to provide more reliable and insightful sales predictions.

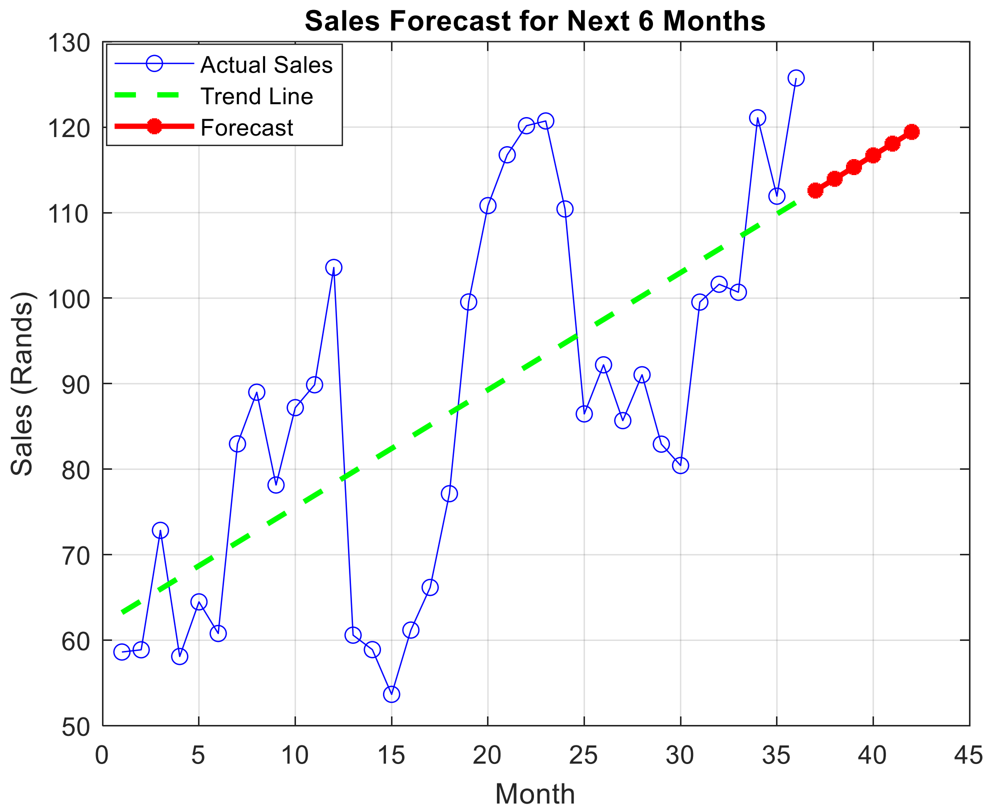

Figure 3 Sales Forecast (Next 6-months)

---THE END---