**Time Series Forecasting of Commodity Prices: An Empirical Analysis Using ARIMA and Exponential Smoothing Models**

Abstract: This report explores the application of ARIMA and Exponential Smoothing time series models to predict future commodity prices. Employing a dataset of monthly closing prices from January 2018 to November 2023, we conduct extensive exploratory data analysis (EDA), model fitting, and tuning to forecast the prices of Gold, Silver, Copper, Platinum, and Palladium for a three-month horizon. The models are evaluated primarily using the Mean Absolute Percentage Error (MAPE), offering a precise, percentage-based measure of prediction accuracy.

**1. Introduction:**

Forecasting commodity prices represents a fundamental and intricate task for a broad spectrum of stakeholders, including financial analysts, investors, and policymakers. The ability to predict future prices of commodities like Gold, Silver, Copper, Platinum, and Palladium is a matter of economic interest and strategic importance. Commodity price fluctuations can significantly impact global financial markets, trade balances, and even monetary policies.

The complexity in forecasting arises due to the dynamic nature of commodity markets, influenced by many factors, including geopolitical events, supply and demand dynamics, market sentiment, and macroeconomic trends. This research taps into the domain of time series analysis, a statistical approach that considers the temporal dimension of data, offering valuable insights into future trends and patterns.

**Data Analysis**

**1. Introduction:**

Forecasting commodity prices represents a fundamental and intricate task for a broad spectrum of stakeholders, including financial analysts, investors, and policymakers. The ability to predict future prices of commodities like Gold, Silver, Copper, Platinum, and Palladium is a matter of economic interest and strategic importance. Commodity price fluctuations can significantly impact global financial markets, trade balances, and even monetary policies.

The complexity in forecasting arises due to the dynamic nature of commodity markets, influenced by many factors, including geopolitical events, supply and demand dynamics, market sentiment, and macroeconomic trends. This research taps into the domain of time series analysis, a statistical approach that considers the temporal dimension of data, offering valuable insights into future trends and patterns.

**2. Exploratory Data Analysis (EDA) and Data Preprocessing:**

Before delving into the predictive models, a rigorous EDA was conducted. This step was crucial in understanding the underlying structure and characteristics of the time series data. We aimed to uncover any underlying patterns, trends, seasonality, and anomalies in the commodity price data through EDA.

* **Trend Analysis:** We plotted the time series to inspect for long-term trends visually. For instance, a steady increase in Gold prices over a period could suggest an underlying upward trend.
* **Seasonality Detection:** We used autocorrelation and partial autocorrelation function (ACF and PACF) plots to examine repeating patterns or cycles at regular intervals, such as seasonal increases in Silver prices due to industrial demand.
* **Volatility Assessment:** By observing periods of significant price fluctuations, we could identify volatility clustering, often addressed in modeling through variance-stabilizing transformations or models like GARCH.
* **Anomaly Detection:** Outliers or unusual spikes in prices, potentially due to geopolitical events or market crashes, were noted for further analysis or potential removal to avoid skewing the model training.

For preprocessing, the data was cleansed, normalized, and transformed to ensure stationarity and better meet the assumptions of the subsequent time series models.

**3. Methodology:**

**3.1 Stationarity Checks:**

Stationarity in time series data is a critical assumption for ARIMA modeling. A stationary time series has properties that do not depend on the time the series is observed. It has constant mean and variance over time; covariances are independent of time. To test for stationarity, we implemented the Augmented Dickey-Fuller (ADF) test, a type of unit root test that provides a formal statistical test for stationarity. If the series was found to be non-stationary, differencing techniques were applied until stationarity was achieved.

**3.2 Model Selection and Fitting:**

* **ARIMA Model Selection:**
  + **Autoregressive (AR) Component (p):** This component allowed the model to learn the feedback mechanism within the time series data. It was determined by examining significant spikes in the PACF plot up to a certain lag, which indicates the number of lag observations included in the model.
  + **Differencing (I) Component (d):** The differencing steps were applied to make the time series stationary. The order of differencing was chosen based on the number of transformations required to stabilize the series' mean.
  + **Moving Average (MA) Component (q):** This component was included to model the noise or shock effects in the time series, which was observed in the ACF plot as significant spikes.
* **Exponential Smoothing Model Selection:**
  + **Trend Component:** When the data exhibited a consistent upward or downward slope, an additive or multiplicative trend was applied, respectively. The additive model was used when the trend was linear, and the multiplicative model was used when the trend increased or decreased at a non-linear rate.
  + **Seasonal Component:** Seasonality in the data was addressed by incorporating seasonal factors into the model. The additive approach was suitable when the series's seasonal variations were roughly constant. In contrast, the multiplicative approach was used when seasonal variations changed proportionally to the level of the series.

**3.3 Model Tuning:**

* **ARIMA Tuning:** ARIMA models' (p, d, q) parameters were fine-tuned using a grid search approach across a defined parameter space to identify the optimal combination that minimized the AIC.
* **Exponential Smoothing Tuning:** The smoothing parameters, including the level, trend, and seasonal components, were optimized by minimizing the mean squared error (MSE) on a validation set.

**4. Modelling:**

The data was split into training and validation sets. The ARIMA and Exponential Smoothing models were then fitted to the training data.

* **Differencing for ARIMA:** Differencing was applied to the ARIMA models to remove the non-stationary part of the series, transforming it into a stationary series. For instance, a non-seasonal ARIMA model is often denoted ARIMA(p,d,q), where parameters p, d, and q are non-negative integers.
* **Smoothing for Exponential Smoothing:** The Exponential Smoothing models applied weighted averages where the weights decreased exponentially over time, hence 'smoothing' out the series and highlighting different components of the time series.

**5. Model Tuning and Validation:**

The tuning and validation of time series forecasting models require careful consideration due to the inherent temporal structure of the data. We employed a rolling forecast origin approach time series cross-validation for model tuning and validation. This section expands on the rationale behind these techniques and their application in our study.

**5.1 Rolling Forecast Origin Approach:** This approach is particularly well-suited for time series data for several reasons:

* **Temporal Consistency:** Traditional random train-test splits are not appropriate for time series data as they can disrupt the time-dependent structure. The rolling forecast origin approach maintains the temporal order of observations, ensuring that the model is never trained on future data.
* **Simulating Real-World Scenarios:** This method mirrors real-world forecasting scenarios where a model is continuously updated as new data becomes available. It involves using an initial training period and incrementally expanding it one time step at a time, forecasting the next time step.
* **Robust Validation:** By shifting the train-test split point through the series, the model's performance is tested across different data intervals, providing a more comprehensive assessment of its predictive capabilities.

**5.2 Model Tuning:** The model tuning process was designed to optimize key parameters, which included:

* **Parameter Optimization:**
  + For ARIMA, we tuned the (p, d, q) parameters, representing the number of autoregressive terms, the number of differences needed to achieve stationarity, and the number of lagged forecast errors in the prediction equation.
  + For Exponential Smoothing, we focused on optimizing the smoothing parameters for the level, trend, and seasonal components.
* **Grid Search:** We implemented a grid search methodology, systematically working through multiple combinations of parameter values and selecting combinations that minimized a predetermined loss function, such as the AIC for ARIMA models or the MSE for Exponential Smoothing models.
* **Avoiding Overfitting:** By evaluating various parameter settings and their performance on the validation set, we mitigated the risk of overfitting, where a model may perform exceptionally well on training data but poorly on unseen data.

**5.3 Validation Process:** The validation process was critical in assessing the generalizability and effectiveness of the models:

* **Walk-Forward Validation:** This involved using the rolling forecast origin approach to make predictions one step ahead, then comparing these predictions against the actual observed values. The model's parameters were re-estimated each time new data was added to the training set.
* **Performance Metrics:** The primary metric used for evaluating model performance was MAPE. This percentage error measure provided an intuitive understanding of the model's accuracy, making it easier to compare across different models and commodities.
* **Aggregation of Results:** By aggregating the performance metrics across multiple train-test splits, we obtained a more reliable and stable estimate of the model's forecasting ability.

**6. Results and Discussion:**

Model performance evaluation was centered on using the Mean Absolute Percentage Error (MAPE). This metric is handy in time series forecasting, providing a precise, intuitive measure of prediction accuracy relative to the observed values.

* **Advantages of MAPE:**
  + **Scale Independence:** One of the critical strengths of MAPE is its independence from the scale of the data, which allows for direct comparison across different commodities regardless of their price scales.
  + **Interpretability:** MAPE expresses error as a percentage, making it easily interpretable for technical and non-technical stakeholders. For instance, a MAPE of 5% indicates that the forecast deviates from the actual value by 5% on average.
  + **Application to Our Data:** In our analysis, MAPE values varied across commodities, reflecting each market's unique characteristics and volatility. For example, lower MAPE values for Gold indicated more accurate forecasts than those with higher MAPE values like Palladium.
* **Limitations and Considerations:**
  + **Sensitivity to Zero Values:** MAPE can be susceptible to zero values in the series, which might not be applicable in our case but is a general metric limitation.
  + **Not Capturing Direction of Error:** MAPE does not distinguish between overpredictions and underpredictions, which could be critical in specific analytical contexts.
* **Comparative Analysis:**
  + The comparative analysis across different commodities using MAPE allowed us to identify which models performed best for each commodity and understand their predictive abilities' nuances.

**7. Forecasting:**

The forecasts were generated for three months, from December 2023 to February 2024. This short-term forecasting was instrumental in providing timely insights for decision-making.

* **ARIMA Model Forecasts:**
  + **Prediction Intervals:** A significant advantage of the ARIMA models was their ability to provide prediction intervals and point forecasts. These intervals are crucial in risk management and decision-making processes as they quantify the uncertainty inherent in the forecasts.
  + **Interpretation of Intervals:** For instance, a 95% prediction interval implies that we expect future observations to fall within this range 95% of the time, assuming the model is correctly specified, and the residuals are normally distributed.
* **Exponential Smoothing Forecasts:**
  + While Exponential Smoothing models provided effective point forecasts, especially for data with trends and seasonality, they did not natively offer prediction intervals. However, these intervals can be estimated through simulation methods such as bootstrapping.
* **Practical Implications:**
  + The forecasts have practical implications for stakeholders, including investors, market analysts, and policymakers, who rely on accurate price forecasts for planning and decision-making.

**8. Conclusion:**

Applying ARIMA and Exponential Smoothing models to commodity price forecasting showed the effectiveness of time series modeling in predicting future prices. MAPE provided a precise measure of forecast accuracy, which was instrumental in model selection and validation. Future work could include the development of hybrid models that incorporate elements of ARIMA and Exponential Smoothing or applying machine learning techniques to capture non-linear patterns and relationships within the data.