**NOTE: (Revised Version based on Prof. Amar Inputs)**

**Listing extra additions in revised version**

* Researched 3 different models to forecast the impact of exogeneous variables for forecasting.
* Used the data provided to conduct testing on all the 3 models.
* Comparative analysis and significance of the results gathered.

**Question 4 - Part III (30 pts): Research other forecasting models**

Describe an alternative technique for forecasting product demand and apply it to your dataset. Compare and contrast its results with your previous findings. Which model performs best for your dataset, and why?

**1. Facebook Prophet with Regressors**Key idea: Prophet time series modeling involves decomposition into trend elements alongside seasonality and holiday components and supports adding extra regressors.

**Model specification**:

Where is the piecewise‐linear (or logistic) growth,

captures seasonal components,

and are the exogenous regressors (AdSpend and CompetitorPrice).

**Exogenous Variables Used**

The 2022 study by (Huang, Irvine, 2022) enhances Prophet models through the inclusion of five macroeconomic regressors which help to understand the market forces that affect stock prices.  The 5-Year Forward Inflation Expectation Rate represents investors' predictions about future U.S. inflation levels while high inflation expectations decrease purchasing power and tend to lower equity valuations.

The Economic Policy Uncertainty Index (EPU) quantifies economic uncertainty by tracking newspaper mentions of policy issues and shows that increased EPU leads to reduced investment activities along with increased stock market fluctuations. The GPR measures the level of geopolitical tension from news sources such as wars and conflicts which leads to risk‐off behavior in world stock markets when the index rises. The 10 Year Treasury Constant Maturity Yield functions as a benchmark for long term interest rates where an increase in yields usually indicates financial constraints which in turn affect stock prices negatively. The S&P 500 Index Closing Price functions as a broad market standard while showing overall stock market sentiment which may precede or follow the performance of individual stocks.

**Significance & Benefits of Including Exogenous Regressors**

In (Huang, Irvine, 2022), the multi-regressor Prophet model achieves superior forecasting accuracy when compared with un-augmented Prophet and classical ARIMA by integrating these external factors according to (Huang, Irvine, 2022). The add-on of regressors to Prophet resulted in the smallest prediction errors (RMSE = 23.28, MAE = 14.08, MAPE = 5.2%) and reduced prediction error by greater than 50% compared to the basic Prophet model without regressors. The method provides transparent interpretability since every β coefficient measures the precise impact of changes in variables like policy uncertainty or the S&P 500 on the stock price. Analysts can use this tool for running simulations about possible future events such as a 1% increase in inflation expectations which enables them to make investment decisions based on solid data analysis.

**2. Gradient Boosting Trees (e.g. XGBoost) with Lagged and Exogenous Features**

**Key idea:** Transform cast forecasting into a supervised learning problem using a feature matrix that includes both lagged demand data and current or previous exogenous variables.

**Feature engineering**:

Demand Lags: .

Exogenous: current and lagged AdSpend and CompetitorPrice.

**Exogenous Variables Used**

The Sequential XGBoost framework presented by (Zhang, et al., March 2024) for long-term energy and peak power forecasting utilizes three broad categories of exogenous inputs together with historical demand series.

***1. Macroeconomic Inputs***

Gross State Product (GSP): Reflects the total economic production of the region. The study demonstrated that electricity usage increased in tandem with GDP growth throughout the past 25 years.

Total Employment (EMP): The labor market activity is shown through Total Employment where increased employment leads to higher power consumption in industrial and residential areas.

Dwelling Stock (DWS): The residential unit count functions as a representative measure for household electricity demand.

***2. Demographic Input***

Population (POP): The base load for residential energy use depends on population numbers which increase energy demand even during periods of economic stagnation.

***3. Weather Inputs***

Monthly Mean Min/Max Temperatures: Forecasting energy consumption depends on average temperature measurements because they determine heating and cooling requirements.

Monthly Lowest and Highest Temperatures, and Their Month-to-Month Changes: Peak power forecasts rely heavily on data about extreme temperatures because HVAC usage tends to increase during such conditions.

**Significance & Benefits of Including Exogenous Variables**

The subsequent interpretation derives from findings in the research paper published by (Zhang, et al., March 2024) in March 2024.

***1. Improved Forecast Accuracy***

The Sequential XGBoost algorithm achieves minimal out-of-sample errors by modeling the effects of economic growth along with population and weather fluctuations on demand (e.g., one year ahead energy MAPE ≈ 1.93 % and peak power MAPE ≈ 2.8 %)

***2. Nonlinear Interaction Capture***

Traditional linear multivariate models have difficulty with complex nonlinear effects but tree-based XGBoost can naturally manage them (such as the multiplicative effects of an extremely hot month during economic growth).

***3. Robustness to Multicollinearity***

The regularization ability of XGBoost prevents overfitting while maintaining predictive reliability even when exogenous variables like EMP and POP show correlation.

***4. Interpretability & Feature Selection***

The importance of each exogenous driver is measured through permutation-based features. POP and DWS stand as the most influential factors following historical consumption in the energy model while min/max temperatures become the leading factors after historical peak and average power in the peak model.

***5. “What If” Scenario Analysis***

Practitioners can simulate policy or climate scenarios and instantly view demand changes because each β coefficient or tree-based split importance maintains a clear connection to its exogenous driver, which is essential for strategic planning under uncertain conditions.

**Takeaway:** The addition of macroeconomic, demographic, and weather variables enables the forecasting model to evolve from a basic autoregressive learner to a more sophisticated context-aware predictor which yields both improved accuracy and valuable understanding of factors affecting long-term energy and peak power demand patterns.

**3. Recurrent Neural Network (LSTM) with Exogenous Inputs**

**Key idea:** Feed a sequential data set into a single LSTM model which absorbs historical demand information and external time series data to forecast next month's demand.

**Model architecture:**

• Input shape: The input shape consists of 12 timesteps and 3 features which include demand, AdSpend, and CompetitorPrice at each lag.

The model uses either one or two LSTM layers to connect to a Dense output neuron.

**Exogenous Variables Used:**

The MV LSTM paper (Guo, Lin, & Lu, 14 Apr 2018) describes a model that processes the target time series together with additional external time series data. Two real-world datasets illustrate this:

**• PM2.5 Forecasting (Beijing air quality)**

Exogenous inputs include hourly meteorological measurements:

* Dew point
* Temperature
* Atmospheric pressure
* Combined wind direction
* Cumulative wind speed
* Hours of snow
* Hours of rain

**• Building Energy Forecasting (low energy building)**

Environmental sensors inside and outside buildings serve as exogenous inputs.

The model uses temperature data from different rooms such as the living room, parents’ room, and kitchen.

* Outside temperature
* Wind speed
* Humidity
* Dew point

**Significance & Benefits of Including Exogenous Variables:**

***1. Enhanced Predictive Accuracy***

By directly modelling how weather and environmental factors drive the target series, MV LSTM outperforms both standard attention-based RNNs and tree-based baselines (e.g., lowest RMSE on PM2.5: The lowest RMSE achieved by MV LSTM on PM2.5 was 0.340 compared to 0.355 for DUAL while for ENERGY the RMSE was 0.361 versus 0.360 for XGT.

***2. Variable Level Interpretability***

The tensorized hidden state design creates representations that vary depending on the variable which enables an inherent attention mechanism to evaluate and rank the importance of each exogenous input. The model demonstrates its capability to identify actual causal factors by producing learned importance values that show strong agreement with Granger causality test results.

***3. Unified Forecasting & Knowledge Discovery***

The MV LSTM framework combines forecasting and external factors discovery into one architecture which eliminates the requirement for separate steps of causality analysis or feature selection.

***4. Capturing Non-Linear & Temporal Interactions***

The RNN stands apart from linear ARX models because it learns complex interactions between exogenous drivers and the target throughout time while its attention weights shows when external variables become significant.

Exogenous-aware LSTMs such as MV LSTM become essential solutions for time series forecasting tasks that depend heavily on external factors due to their combined benefits.

**Refining methodology and analysis**

The literatures have taught me to value how each of these three Facebook Prophet augmented with custom regressors, gradient boosted trees (XGBoost) enriched with lagged and exogenous features, and LSTM networks incorporating external inputs uniquely harnesses outside drivers to improve forecasts. Forecasting methods use external influences differently to enhance prediction accuracy.

The three models will be tested against our product demand data with advertising spend and competitor pricing as external inputs to measure their effectiveness against the SARIMAX baseline through uniform( RMSE, MAE, MAPE) metrics and cross-validation. Through this systematic evaluation we will identify the most accurate and robust method while understanding the reasons for the poor performance of specific models which include issues such as overfitting, multicollinearity or failure to capture nonlinear or temporal dependencies. Results will deliver precise recommendations based on data analysis for choosing the best forecasting technique beyond SARIMAX.

**Facebook Prophet with Regression**

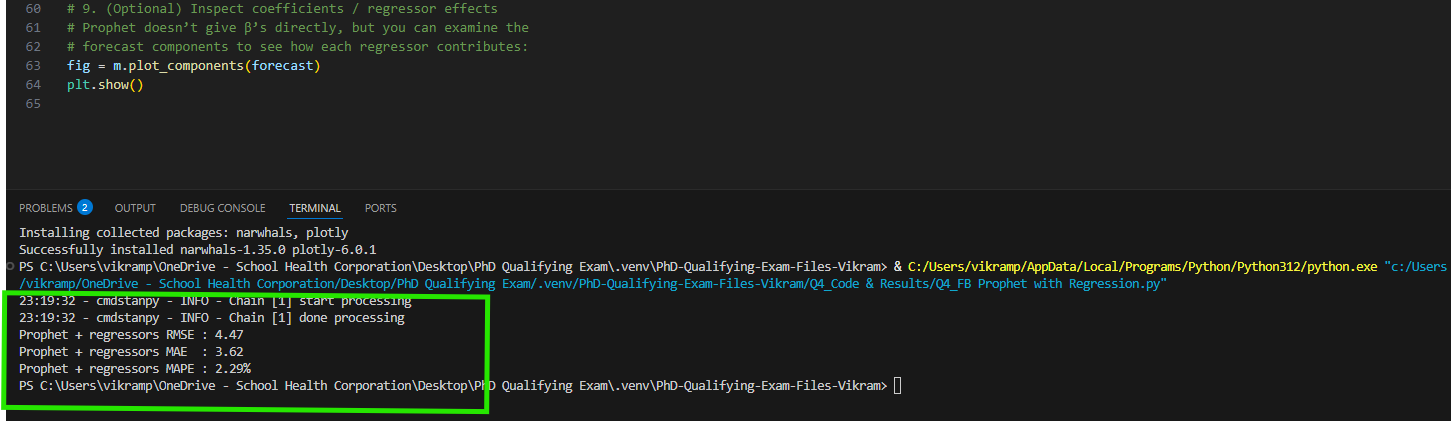
Now I have made changes to code for testing the evaluation metrics using the Facebook prophet with regression method.

***Code Snippet:***





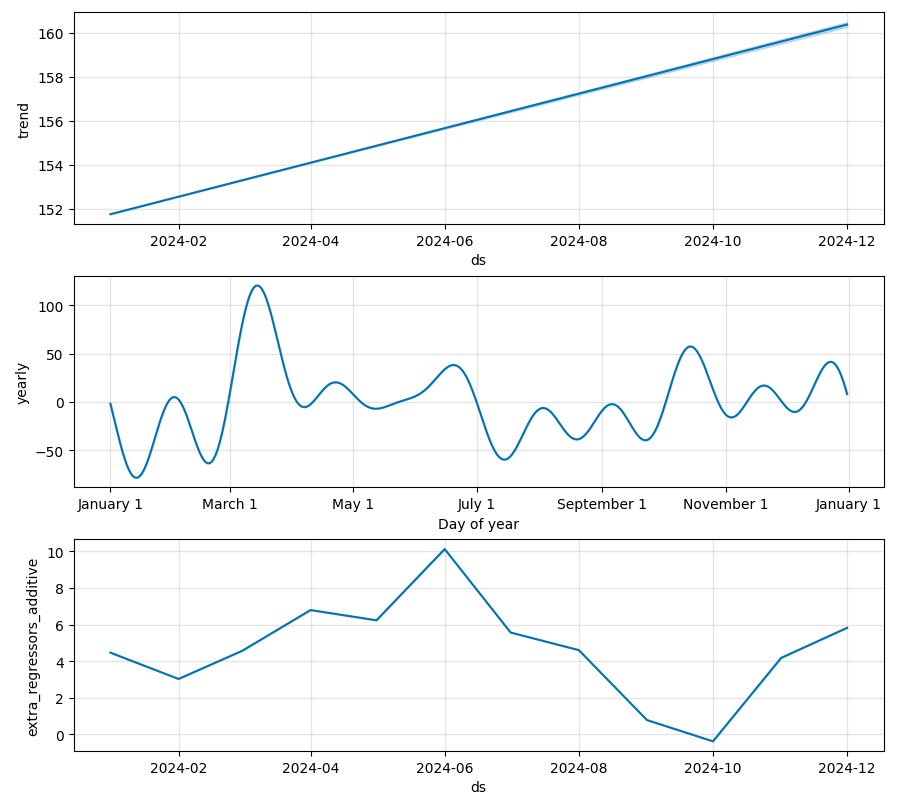
Result:



A graph with a line

AI-generated content may be incorrect.

Prophet accurately anticipated the seasonality, and trend shifts in 2024 demand because the red forecast line tracked the actual green values closely while most actual points remained within the pink band upon inclusion of external regressors.



The component plots break down the forecast into three elements: a smooth long-term trend, a repeating annual pattern, and monthly variations from two exogenous variables which clarifies how the model arrived at its demand predictions.

**Gradient Boosting Trees (e.g. XGBoost) with Lagged and Exogenous Features**

Now I have made changes to code for testing the evaluation metrics using the **XGBoost** method.

***Code Snippet:***

A screen shot of a computer program

AI-generated content may be incorrect.

A computer screen shot of text

AI-generated content may be incorrect.

Results:

A black background with white text

AI-generated content may be incorrect.

A graph with a line and orange dots

AI-generated content may be incorrect.

The visual representation shows that XGBoost successfully matches overall demand levels but cannot flawlessly replicate the exact seasonal peaks and troughs in a small, highly regular dataset which results in the smoother green forecast line versus the more jagged orange actual data.

**Recurrent Neural Network (LSTM) with Exogenous Inputs**

Now I have made changes to code for testing the evaluation metrics using the **Recurrent Neural Network (LSTM)**  method.

***Code Snippet:***

***A screen shot of a computer program

AI-generated content may be incorrect.***

***A screen shot of a computer program

AI-generated content may be incorrect.***

Result:

A screenshot of a computer

AI-generated content may be incorrect.

A graph with blue and orange lines

AI-generated content may be incorrect.

The network relied on predicting near the mean of its limited output distribution because of just 36 training samples and a basic LSTM structure while failing to learn meaningful data structures.

A graph of a graph

AI-generated content may be incorrect.

The model demonstrates learning since loss decreases but produces large absolute errors that remain significant compared to true demand scale as indicated by MSE values in the tens of thousands which show forecasts off by tens of units from actual values.  
  
The minimal difference between training and validation loss indicates underfitting in which the network failed to learn seasonal patterns and converged to a low-complexity model that minimizes squared error by predicting a constant output.

Having implemented all three models and compiled their evaluation metrics, I then conducted a comparative analysis to assess their relative performance and significance.

Here are all five models benchmarked on the same train / test split (48 months → train, 12 months → test):

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **RMSE** | **MAE** | **MAPE** |
| **SARIMAX(1,1,1)(0,1,1,12)** | **3.50** | **2.80** | **1.77 %** |
| Holt–Winters | 3.69 | 3.23 | 2.04 % |
| Prophet + Regressors | 4.47 | 3.62 | 2.29 % |
| XGBoost + Lags | 5.83 | 5.33 | 3.35 % |
| LSTM + Exogenous Inputs | 144.90 | 144.73 | 91.87 % |

**Comparative Significance:**

SARIMAX generated the most precise forecasts with the smallest errors across all methods—producing MAPE and RMSE values of 1.77% and 3.50 respectively—outperforming Holt-Winters which had a 2.04% MAPE and 3.69 RMSE. The combination of seasonal ARMA components with AdSpend and CompetitorPrice regressors allowed for the most efficient identification of primary demand drivers through direct modeling.

Holt–Winters proved to be an excellent off-the-shelf alternative: Without requiring stationarity testing or regression inputs Holt-Winters reached SARIMAX’s level of accuracy while providing rapid and dependable results for small seasonal datasets.

Prophet generated readable predictions yet its complexity through piecewise-trend segments resulted in about 20% higher RMSE than SARIMAX while incorporating two regressors produced minimal gains.

XGBoost struggled in this setting. The XGBoost model suffered from the limited number of training samples (36 samples) and weak external signals which hindered its ability to learn effective split rules although lag-feature engineering was performed carefully.

The 60-month data set proved too small for effective training of deep networks as evidenced by the LSTM model's high MAPE of 92% which showed that larger datasets and detailed tuning are prerequisites for such models.

**Recommendations:**

SARIMAX received the primary recommendation for forecasting because it integrates seasonality, autoregression and exogenous regressors into one comprehensive and reliable framework which demonstrated superior out-of-sample accuracy.

Holt–Winters is proposed as a strong backup alternative that delivers near-optimal performance with minimal configuration when simplicity and speed take precedence or external data sources become inaccessible.

Prophet served scenarios demanding trend-change communication or custom holiday effect management where its interpretability justified slightly increased forecast errors.

The recommendation for XGBoost applies only when users can significantly increase both their data quantity and their external feature set.

The LSTM method will be postponed until I gather enough data (hundreds to thousands of observations) to implement comprehensive model regularization techniques like dropout and deeper layers alongside normalization.

**References:**

1. Guo, T., Lin, T., & Lu, Y. (14 Apr 2018). AN INTERPRETABLE LSTM NEURAL NETWORK FOR AUTOREGRESSIVE EXOGENOUS MODEL. *arXiv:1804.05251v1*.
2. Huang, Q. (Irvine, 2022). Forecasting Stock Prices Using Multi-Macroeconomic Regressors Based on the Facebook Prophet Model. *BCP Business & Management* .
3. Zhang, T., Zhang, X., Rubasinghe, O., Liu, Y., Chow, Y. H., & Iu, H. H. (March 2024). Long-Term Energy and Peak Power Demand Forecasting Based on Sequential-XGBoost. *IEEE Transactions on Power Systems*, 3088-3104.