# Energy Market Time Series Forecasting: Balancing Demand and Supply

**Amrutha Gabbita**
agabbita@purdue.edu

**Ridvan Kanca**
rkanca@purdue.edu

**Sahana Reddy**
reddy47@purdue.edu

**Alexander Scarcelli**
ascarcel@purdue.edu

**Matthew Lanham**
lanhamm@purdue.edu

**PURDUE UNIVERSITY** | Mitchell E. Daniels, Jr. School of Business
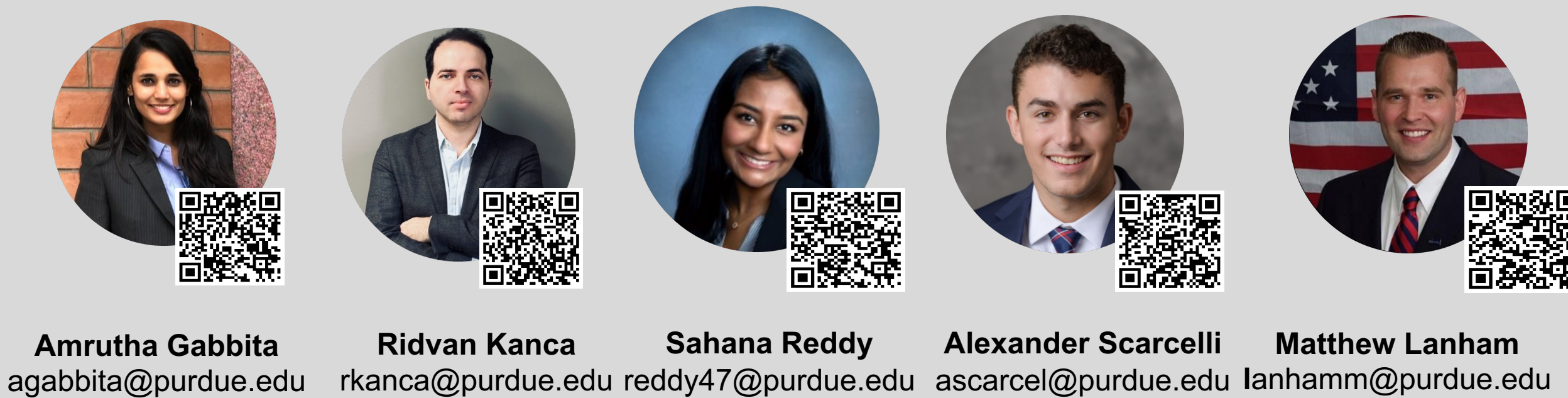
## BUSINESS PROBLEM FRAMING

### Business Problem
Focus on pinpointing locations for grid-scale battery installations to elevate energy distribution across New York.

### Importance
Identifying and filling gaps in generation or storage is critical for improving our competitive edge and optimizing battery usage.

### Business Benefits
Our findings aim to enhance operational efficiency and profitability through informed strategic decisions.

### End Goal
The project seeks to improve how grid-scale batteries are managed, leading to better market positioning, cost savings, and possibly lower energy prices for consumers.

### Purpose
The analysis supports our B2B product and has potential applications in national security by assessing regional market opportunities and identifying gaps.

### Context
Our current focus is on the commercial market, using analytics to assess regional energy fluctuations

### Stakeholders
We target the commercial sector, with an eye toward future national security applications, involving both B2B and possibly B2G relationships.

## ANALYTICS PROBLEM FRAMING

### OBJECTIVES
Analyze hourly energy demand and supply

Evaluate grid infrastructure

Assess renewable energy variability and predictability

Identify regions for optimal grid-scale battery deployment,
.

### STRATEGY
Leverage open-source algorithms like SARIMAX, ARMA, and relevant neural network architectures along with K-Means clustering to forecast hourly energy

**ANALYTICS PROBLEM FRAMING**

1  2  3  4

### ASSUMPTIONS
Assume 2020 data at an hourly level is representative of seasonal hourly trends for 2024 and beyond
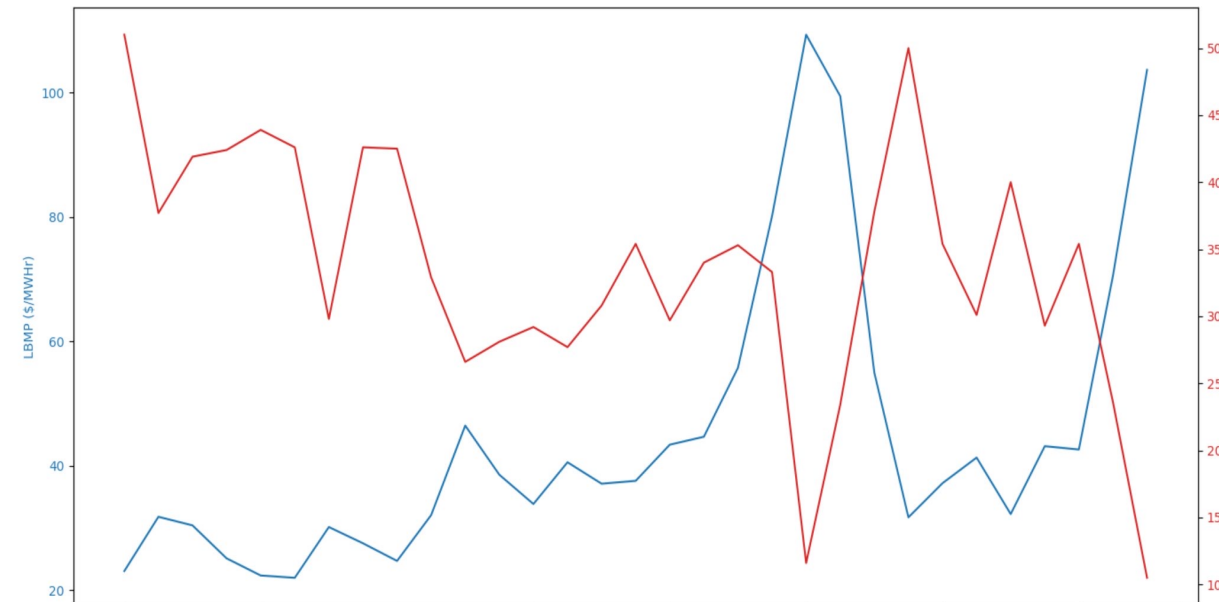
### METRICS
MAPE (Target of 5% or lower deemed successful)

## DATA

Average LBMP by Hour
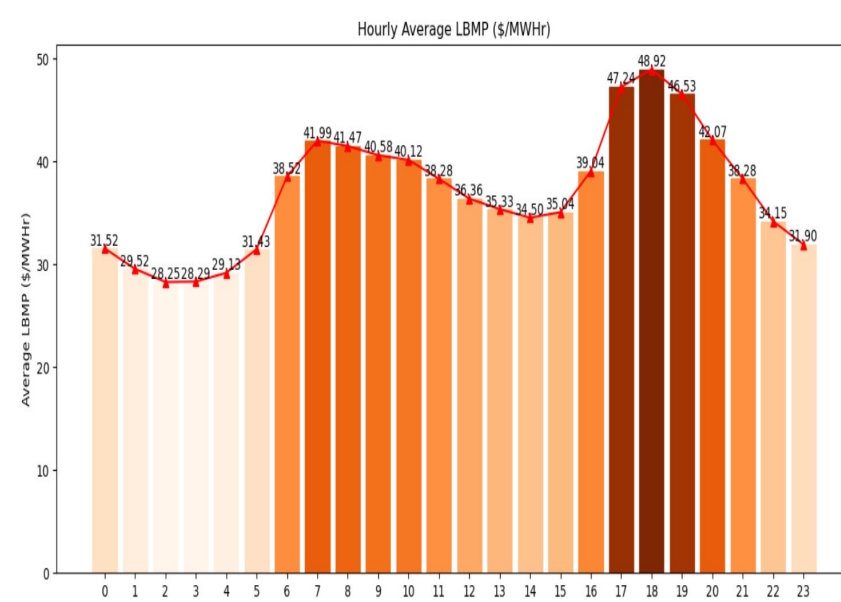
Weather – LBMP relationship

Correlation Matrix
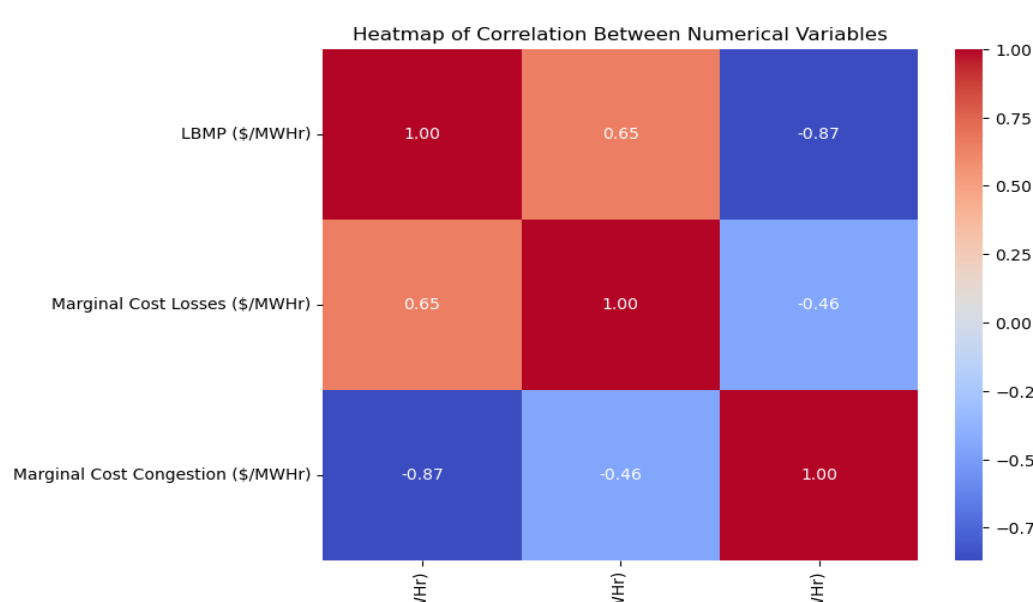
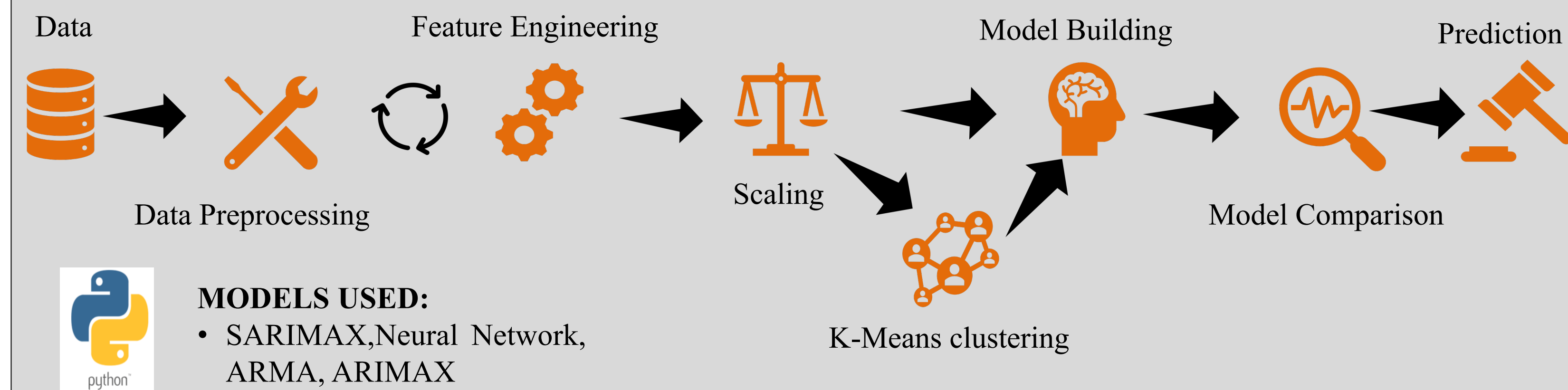| Column Name | Description | Type | Example |
|---|---|---|---|
| Time Stamp | This column records the date and time of a record | String | 1/1/2019 0:00 |
| Name | Street address of the place of interest. | String | 59TH STREET_GT_1 |
| PTID | This column contains unique identifiers assigned to each address. | Int | 25648 |
| LBMP ($/MWHr) | "Locational Based Marginal Pricing". The price of electricity at a specific location on the grid, calculated per megawatt-hour | Float | 25.57 |
| Marginal Cost Losses ($/MWHr) | least bit marginal price. The lowest price at which electricity can be sold in the wholesale market at a given time | Float | 1.07 |
| Marginal Cost Congestion ($/MWHr) | The costs associated with congestion in the power grid. | Float | -14.97 |

### DATA SOURCE
- Primary Data Source: NYISO
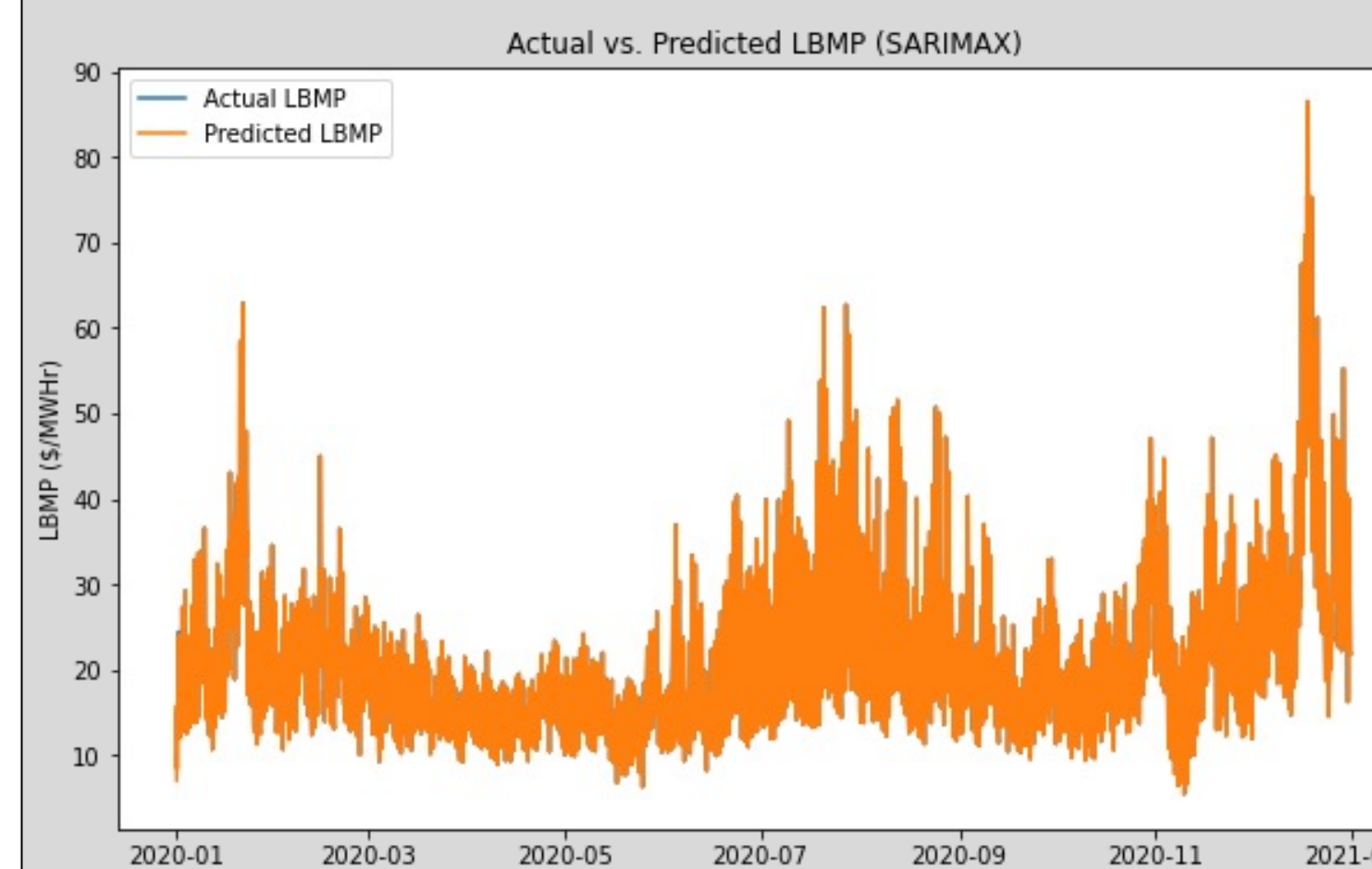- Secondary Data Source: Energy Information Administration

### PREPROCESSING
- Aggregate Timestamp column and filter and impute any null values with forward fill

**New York ISO** Independent System Operator

**eia**

## METHODOLOGY SELECTION

Data → Data Preprocessing → Feature Engineering → Scaling → Model Building → Model Comparison → Prediction

K-Means clustering

**python**

**MODELS USED:**
- SARIMAX, Neural Network, ARMA, ARIMAX

## MODEL BUILDING

Actual vs. Predicted LBMP (SARIMAX)

### SOLUTION ARCHITECTURE
- SARIMAX
- Exogenous Variables: Marginal Cost osses ($/MWHr), Marginal Cost Congestion ($/MWHr), rolling_mean_3 h, LBMP_lag_2h
- Goal MAPE: Less than 5%
- MAPE of Optimal Model: 1.957%

### ENGINEERED FEATURES
- Lagged Variables: The LBMP lagged for the previous 24 hours, 2 days, 1 week
- Rolling Means: Rolling mean of past 3, 12, 24, and 168 hours

### FUTURE DEVELOPMENT
- Improving feature selection techniques (Wrapping, etc.)
- Neural Network training and testing
- Determine optimal model for each season

## DEPLOYMENT AND LIFE CYCLE MANAGEMENT

**Deployment:** SARIMAX-based model to predict hourly LBMP. Includes rigorous validation against MAPE to ensure accuracy.

**Impact:** Improve operational efficiency and cost reduction for energy providers and consumers, and reduce power outages

Time Series Forecasting

Lifecycle management

Impact
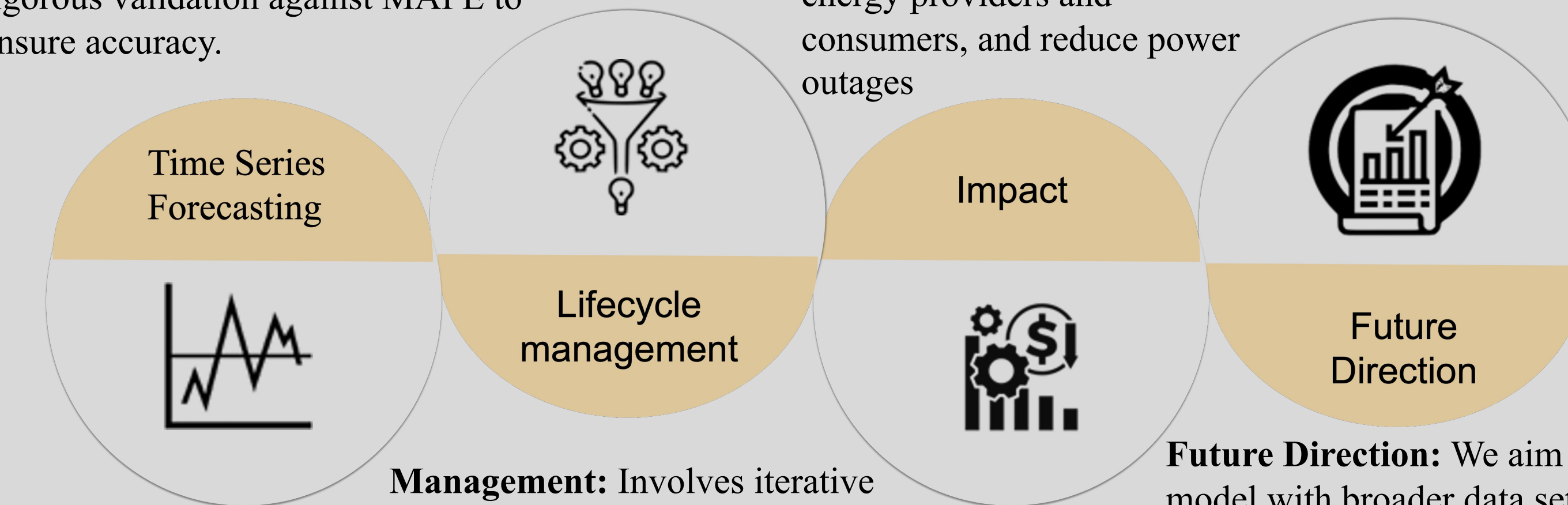
Future Direction

**Management:** Involves iterative refinement, leveraging feedback and cutting-edge research to enhance model precision and adaptability to market changes.

**Future Direction:** We aim to enrich our model with broader data sets and sophisticated machine learning techniques, focusing on renewable energy integration and smart grid advancements .

**BENEFITS:**
- Reduce consumer pricing particularly for time-of-use based plans
- Avoid market failures and blackouts from peaking demand and surging prices

Improved Forecasting Leads to Lower Costs

DIGITAL VERSION