**Modelling and forecasting of day-ahead electricity price in Indian energy exchange**

# **Introduction**

Electricity demand and price forecasts have become increasingly important since the rise of the competitive energy markets. This is a key component of risk management strategies adopted by the market participants. For a utility, price forecast plays an important role in the day-to-day operations, and it is typically used for optimising system operation and scheduling of hydro units and other peaking plants, such as gas turbines. The objective of the operators is to minimise variable costs without jeopardising the electric system to power failures. Similarly, once a good next day and a week ahead price forecast is available, a consumer can derive a plan to maximise its own utility using the electricity purchased from the pool. Because of the non-storable nature of electricity coupled with significant seasonal and diurnal variations of demand; supply constraints at peak hours; and transmission bottlenecks, electricity prices often exhibit extreme volatility.

# **Literature Survey**

Time series techniques have become the most widely used method for short and medium-term forecasts in practice. In power systems, autoregressive integrated moving average (ARIMA) models have been used for short-term price forecasting with good results. Contreras et al. (2003) proposes two ARIMA models to predict hourly prices in the electricity markets of Spain and California. Conejo et al. (2005) analyses the performances of three families of forecasting techniques namely time series, neural networks and wavelets to predict the 24 market-clearing prices of a day-ahead electric energy in Pennsylvania–New Jersey–Maryland (PJM) interconnection, where time series techniques reveal themselves as more efficacious than wavelet-transform or neural network techniques. Among time series techniques, the dynamic regression and transfer function algorithms are found to be more effective than ARIMA models. It has been observed that multiplicative double-seasonal ARIMA and exponential smoothing (ES) models outperform ANN forecasts of electricity demand for lead times up to a day ahead for the state of Rio de Janeiro in Brazil and for England and Wales (Taylor et al., 2006).

Volatility has been regarded as one of the most intrinsic features of electricity prices. According to Johnson and Barz (1999), Hadsell et al. (2004) and Escribano et al. (2011), electricity prices are seen to exhibit more volatility as compared to other commodity prices traded in financial markets. This behaviour is largely attributed to the non-storable nature of electricity, highly variable demand, real-time demand-supply match, market power, and the diversity of plant costs among other things.

In India, long run (yearly basis) power demand forecasting is conducted by Central Electricity Authority (CEA) and Planning Commission of India. To the best of my knowledge, there has been no official report on monthly energy and/or peak demand forecasting in India, let alone weekly demand (and price) forecasting.

This study forecasts day-ahead electricity price of IEX using a multivariate model with the objective of modelling time-varying volatility present in the time series data. Keeping in mind the importance of understanding electricity price volatility and the fact that the day-ahead electricity market is a relatively new phenomenon in the Indian context, a thorough investigation of price forecasting and volatility modelling is useful contributions to the subject.

# **Data Preparation and Preprocessing**

## IEX Market Clearing Price data:

* + Historical as well as Real time data on electricity price (Rs per MwH) in IEX have been collected from IEX.
  + Parameters taken:

1. Market Clearing Price(MCP)
2. Cleared Volume
3. Purchase Bid
4. Sell Bid
5. Difference of purchase and sell bid

## Weather Data:

* + Real time week ahead data is available through Meteoblue source.
  + Reanalysis historical data was collected which was available till Aug-01,2022. The data from Aug to Nov, 2022 is replicated from the same dates of 2021.
  + Parameters taken:

1. Temperature
2. Apparent/Felt temperature
3. Relative Humidity
4. Precipitation Rate

The above parameters are taken for 16 cities:

| Delhi | Raipur | Visakhapatnam | Jaipur |
| --- | --- | --- | --- |
| Chennai | Chandigarh | Lucknow | Ahmedabad |
| Patna | Kolkata | Dehradun | Mumbai |
| Bangalore | Ranchi | Hyderabad | Guwahati |

## 

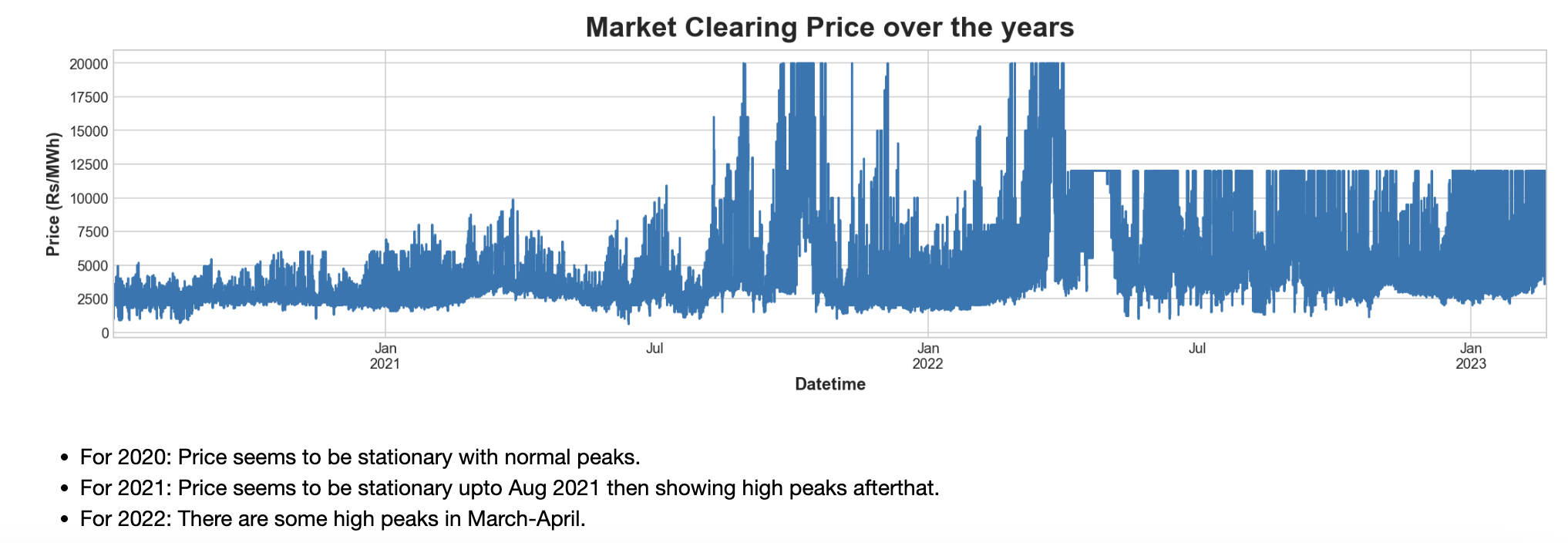
## Power Stations Data:

* + Rainfall data for hydro stations, wind data for wind stations and GHI data for top 5 solar stations are collected.

# **Exploratory Data Analysis**

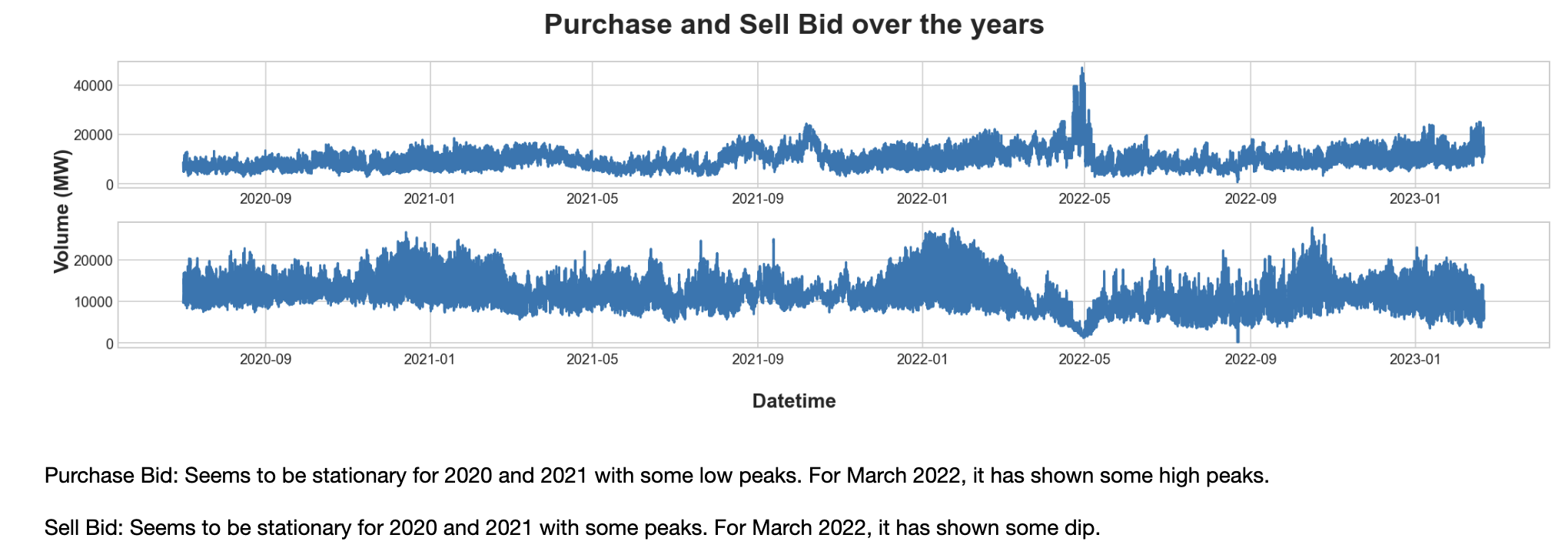
### **Overall Data**

1. **Market Clearing Price (MCP) over the years**

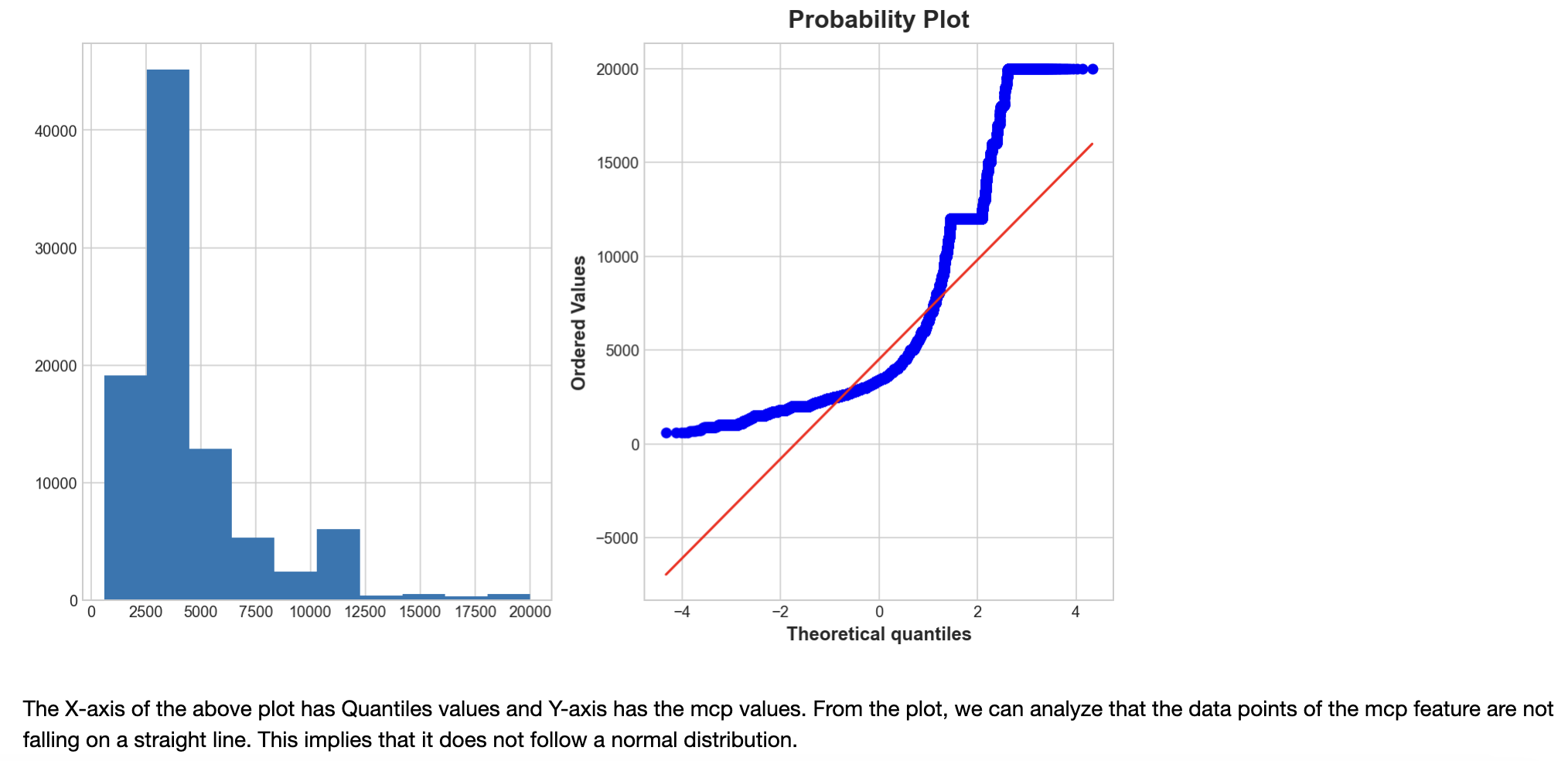


From the above plot, it is evident that from April 2022 onwards, MCP is capped at a price of 12000 Rs/MWh.

1. **Purchase and Sell Bid over the years**

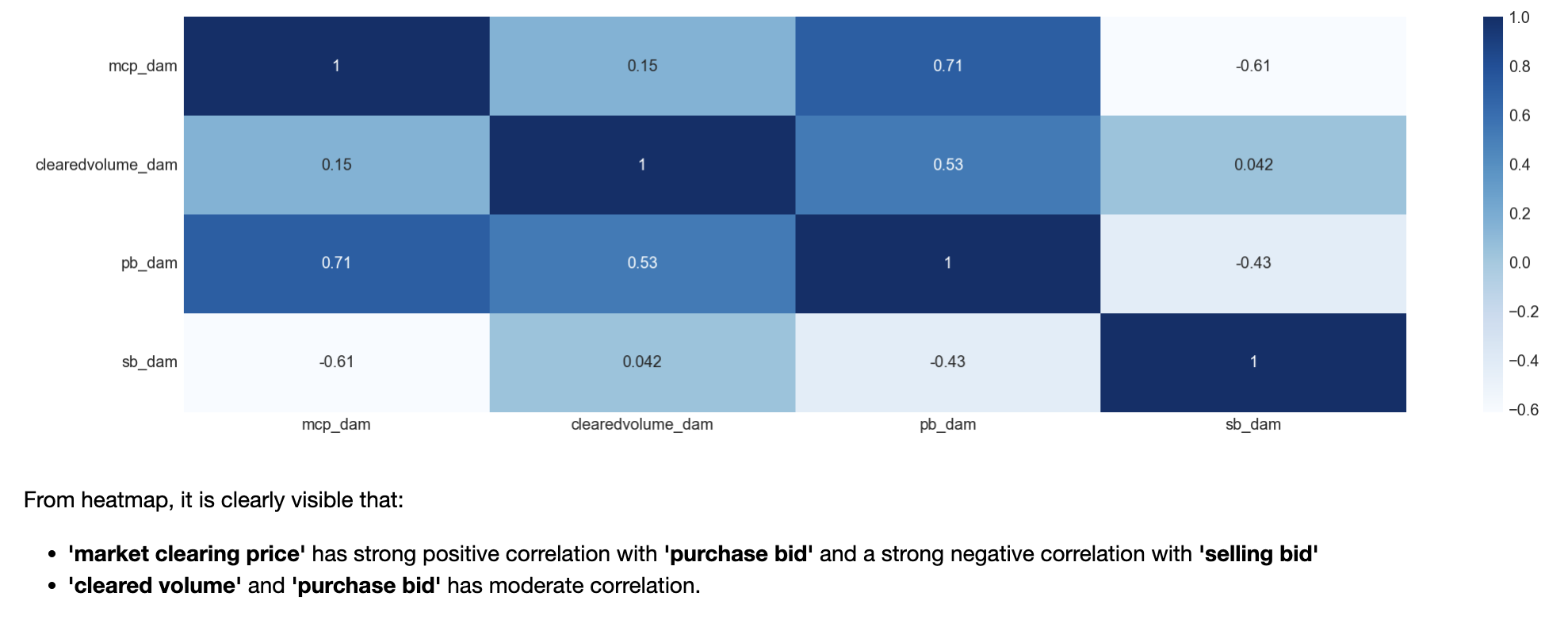
****

1. **Skewness**

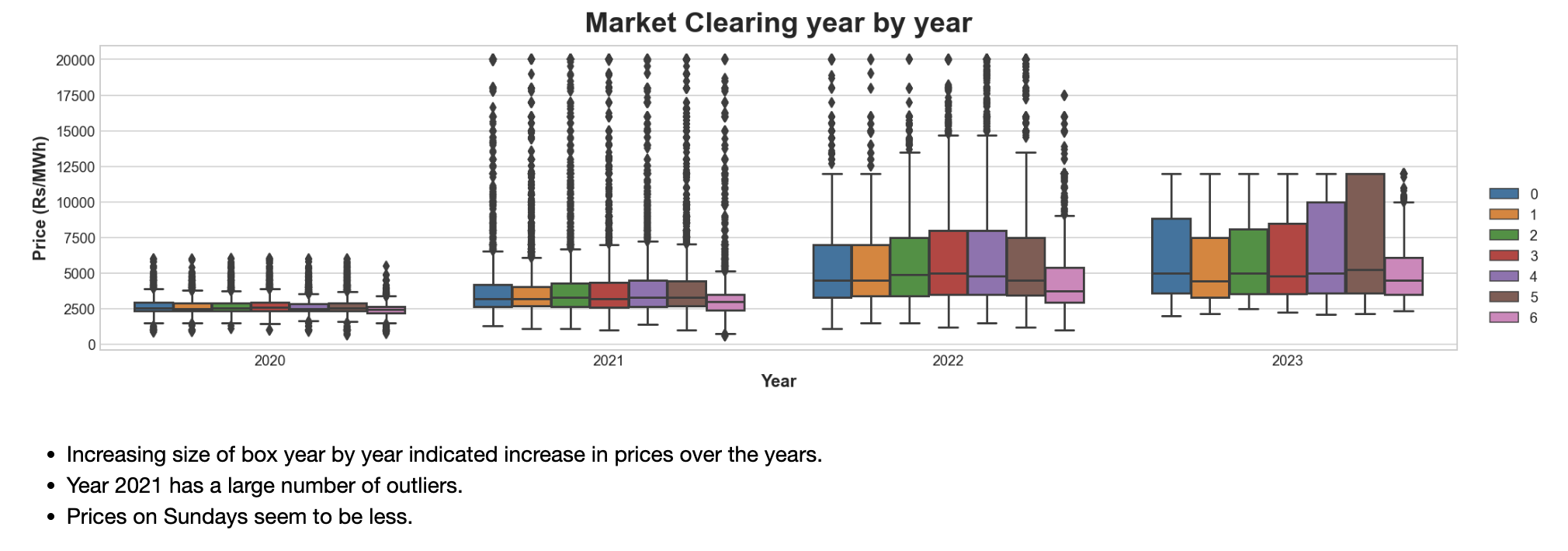
****

MCP seems to be right skewed. From the Probability plot, we can analyze that MCP is not falling on a straight line. This implies that it does not follow a normal distribution.

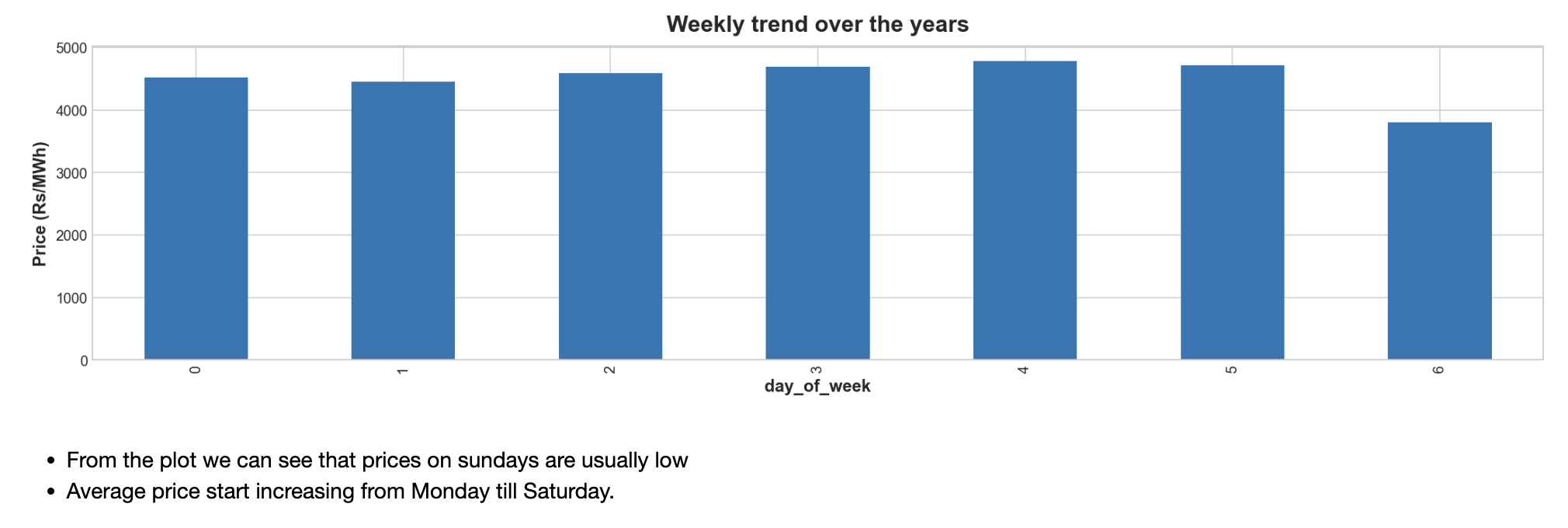
1. **Correlations**

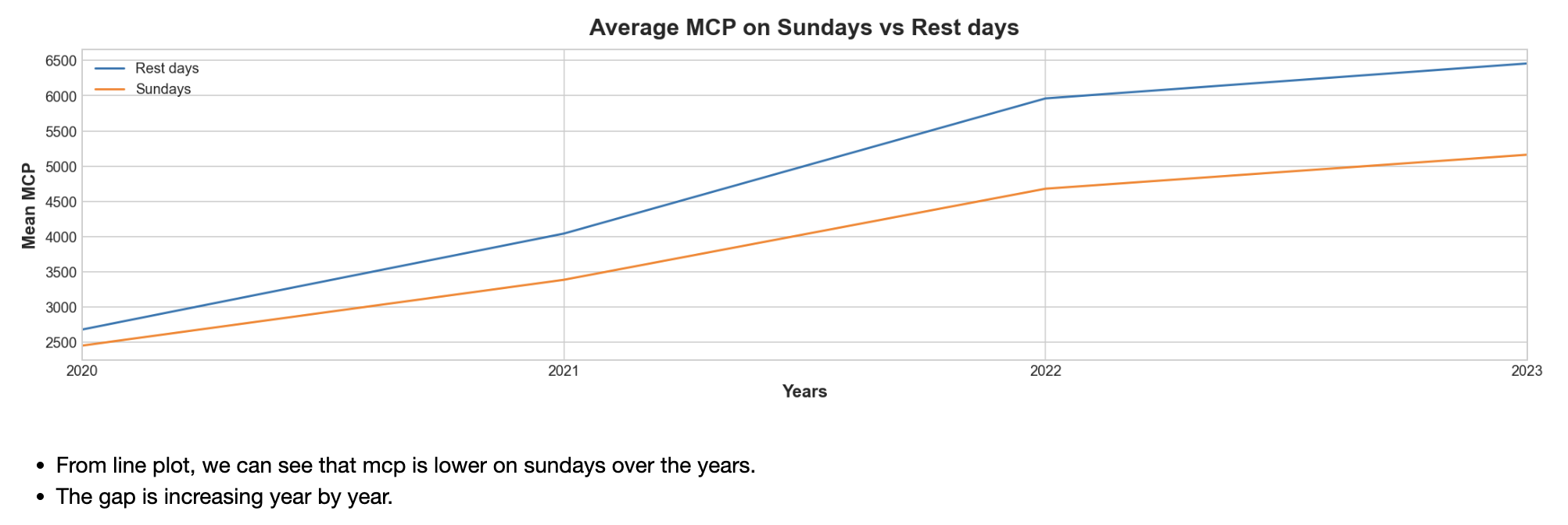
****

1. **Yearly Trend over the years**

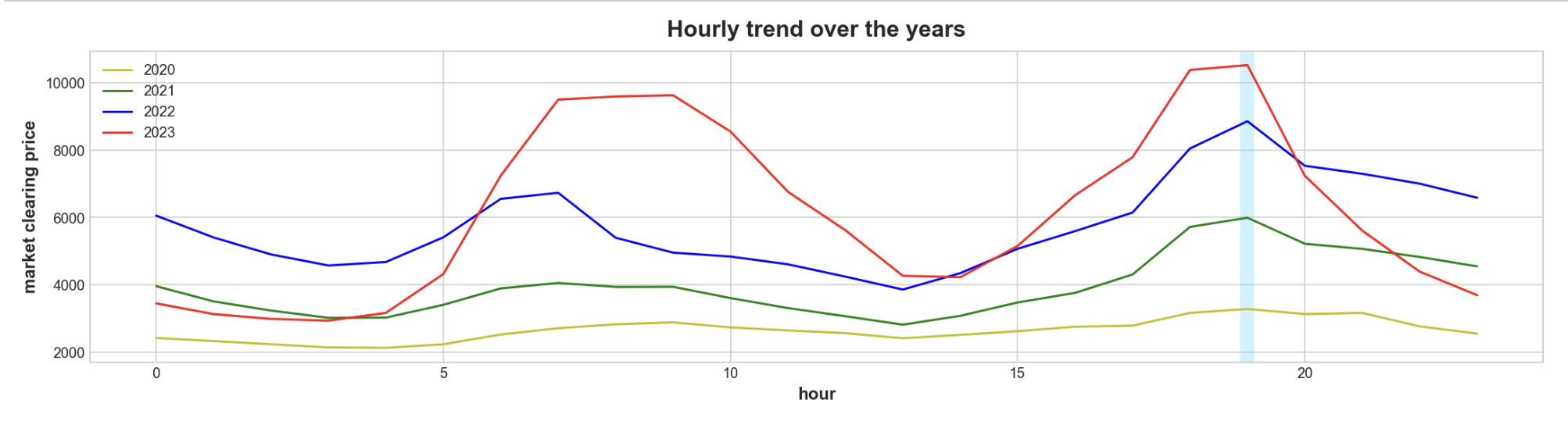
****

1. **Weekly Trend over the years**

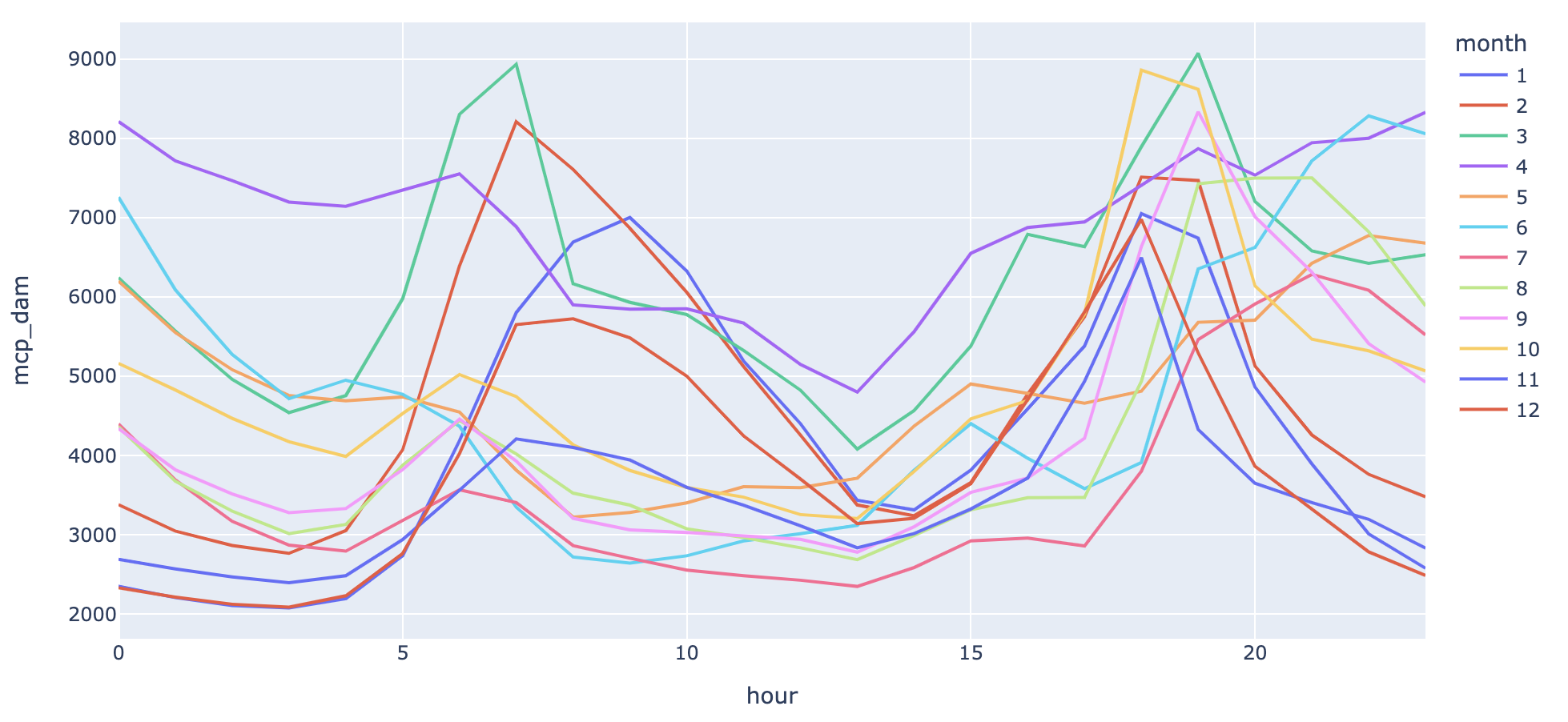
****

****

1. **Hourly Trend over the years**

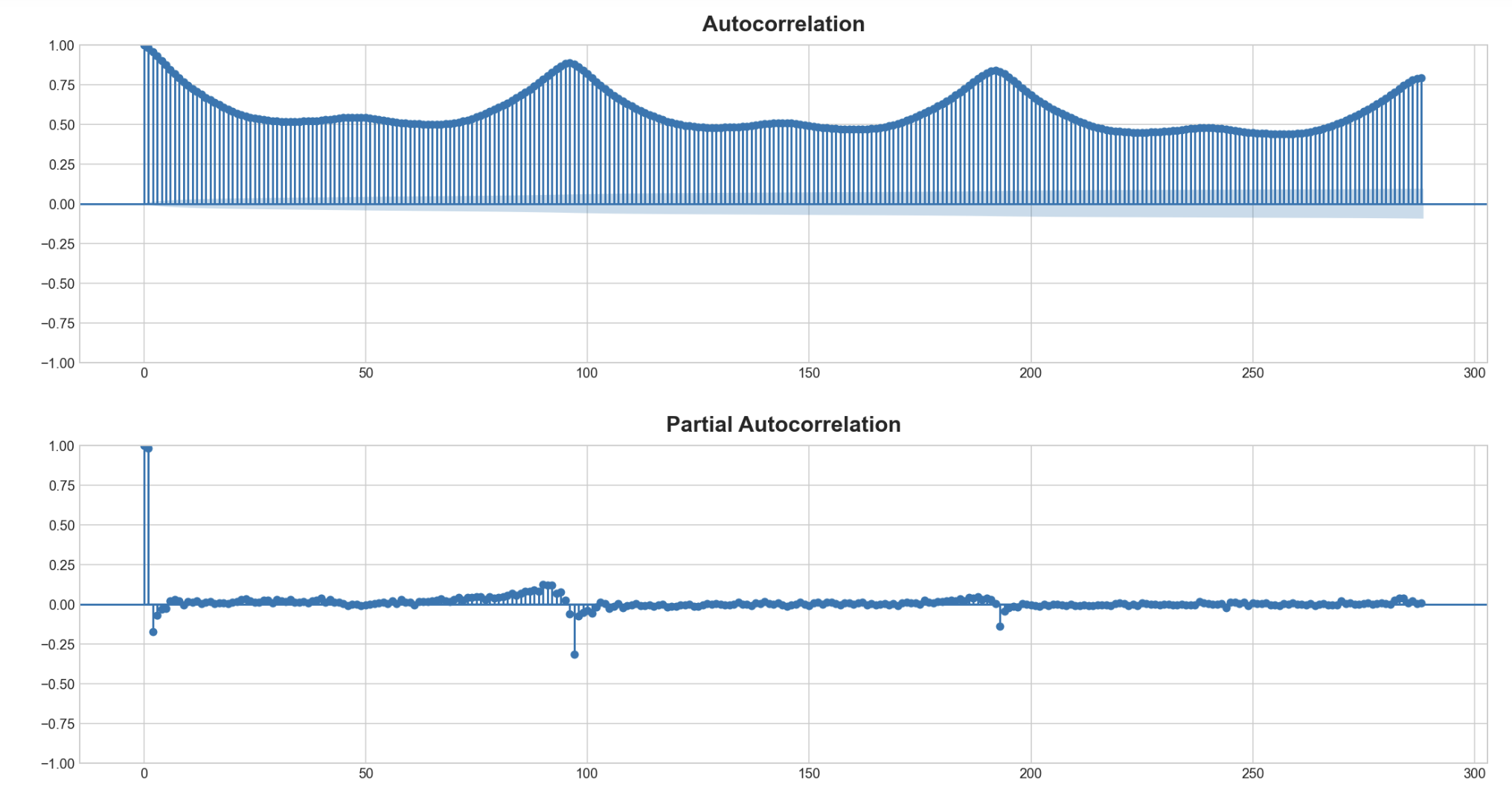
****

Hourly trend has increased over the years. The average maximum MCP seems to occur around 6-7pm.



Based on hourly trends over the months, we later segregated the data into 4 seasons:

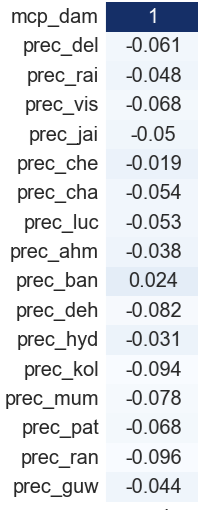
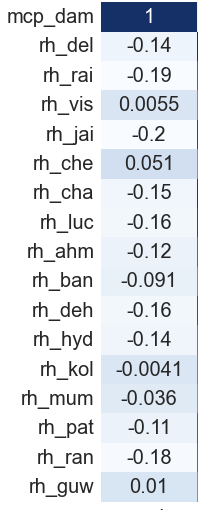
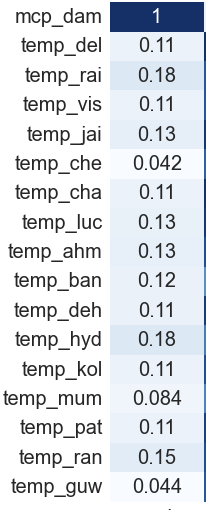
1. Autumn: Oct, Nov
2. Monsoon: July, Aug, Sep
3. Summer: Mar, April, May, June
4. Winter: Dec, Jan, Feb
5. **ACF-PACF plots**

****

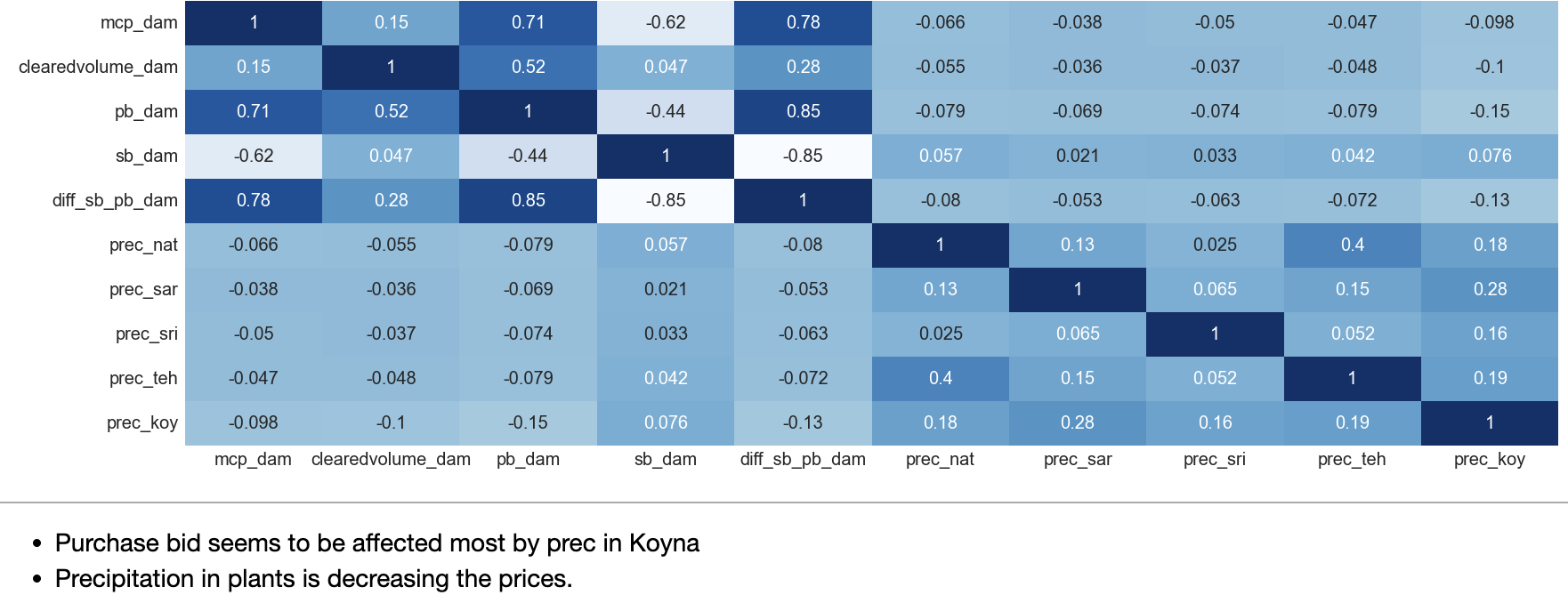
We can see from the ACF plot that there is autocorrelation.

### 

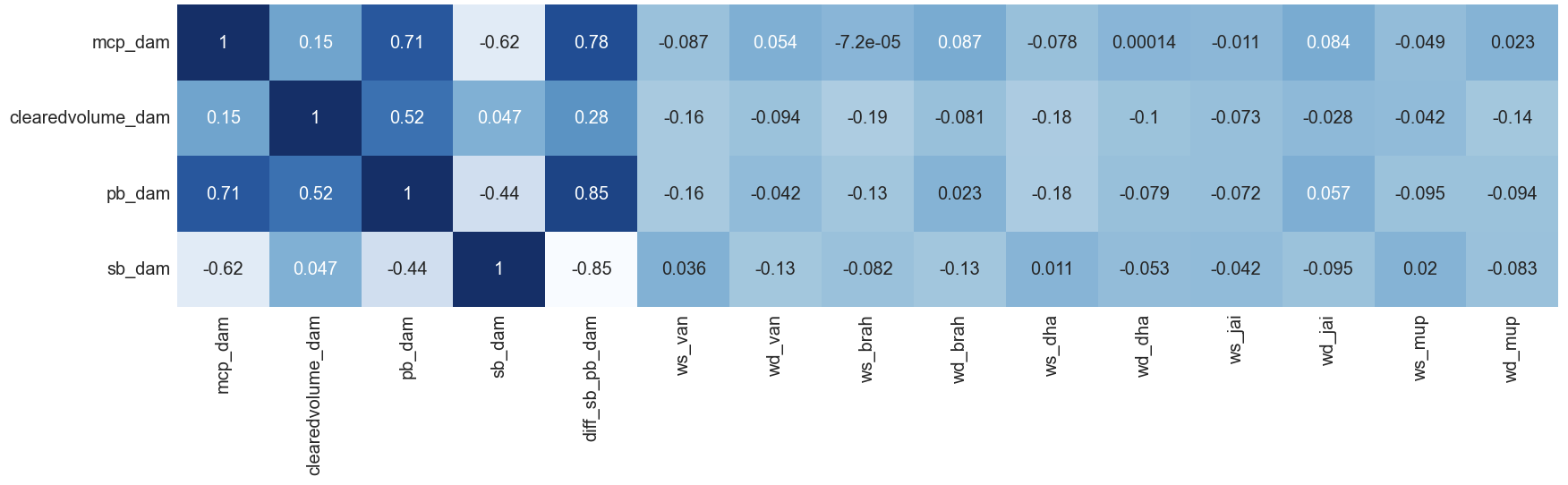
1. **Correlation of MCP with weather parameters**

****

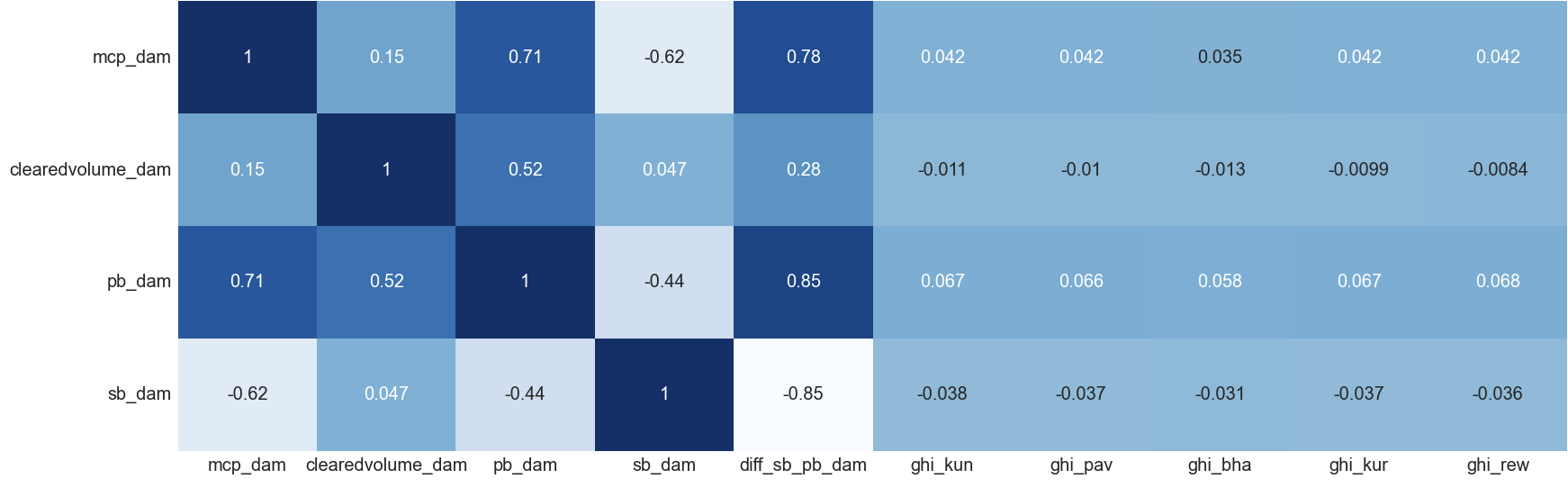
1. **Correlation with precipitation over hydro stations data**

****

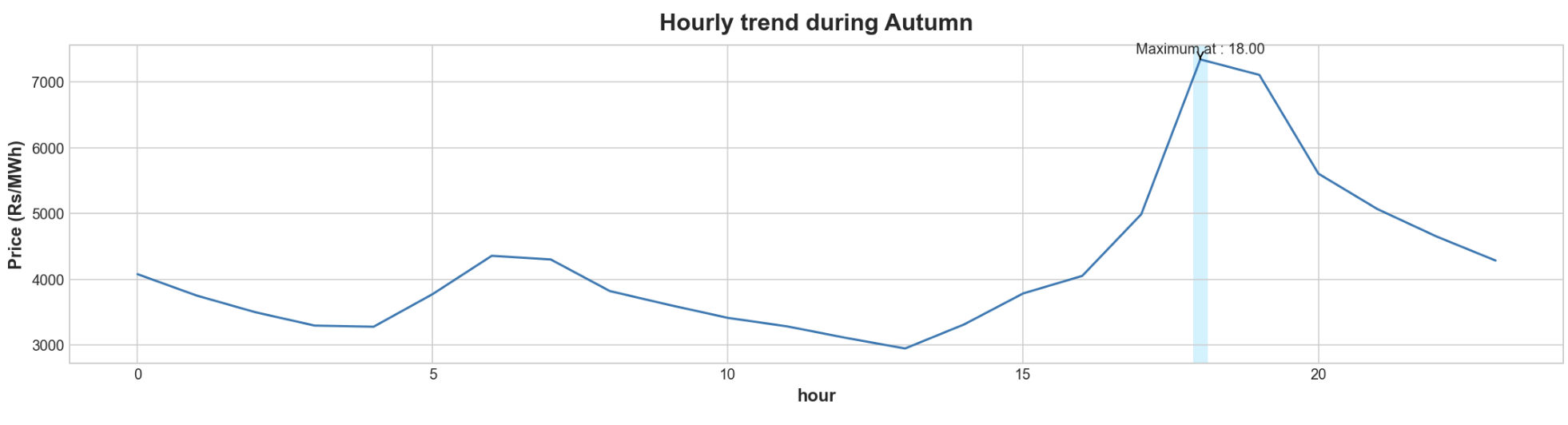
1. **Correlation with wind data of wind stations**

****

1. **Correlation with GHI data over solar stations**

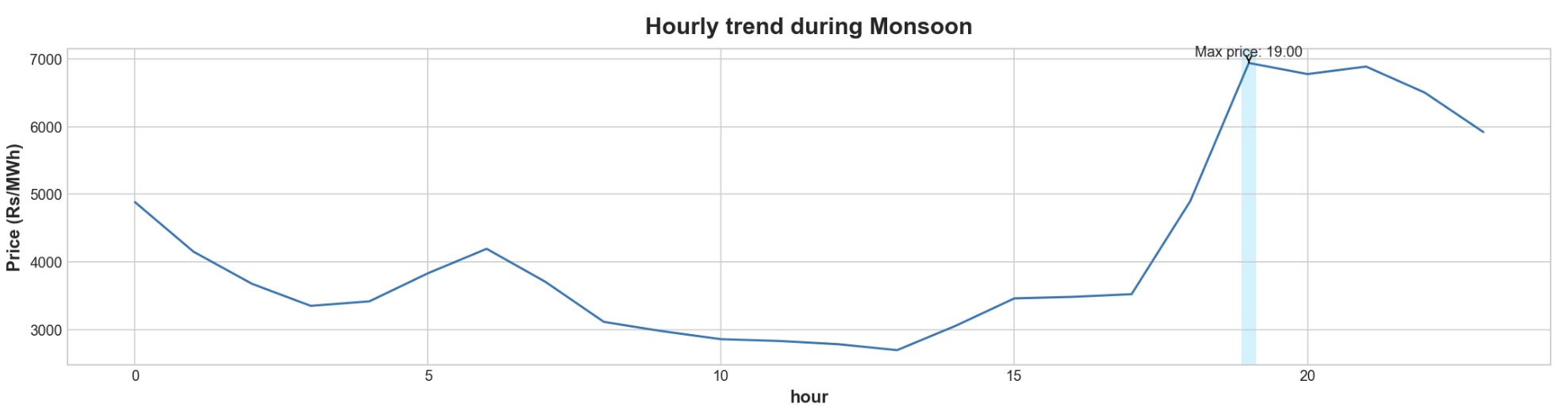
****

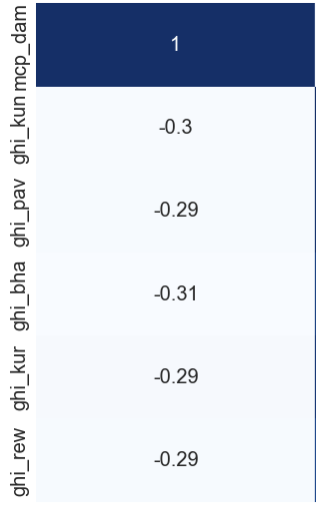
### **B. Autumn Data**



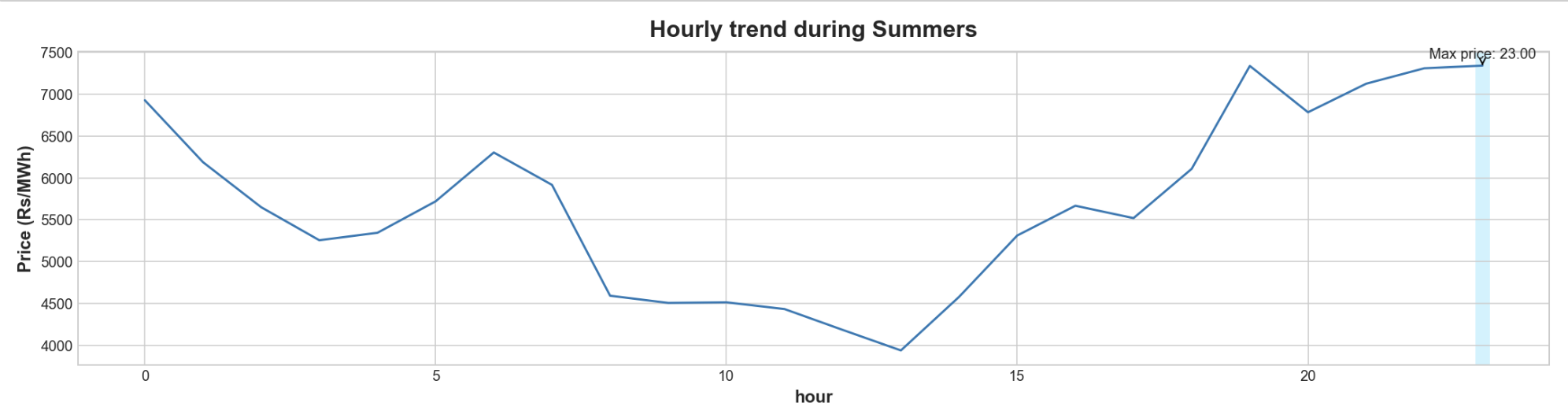
### 

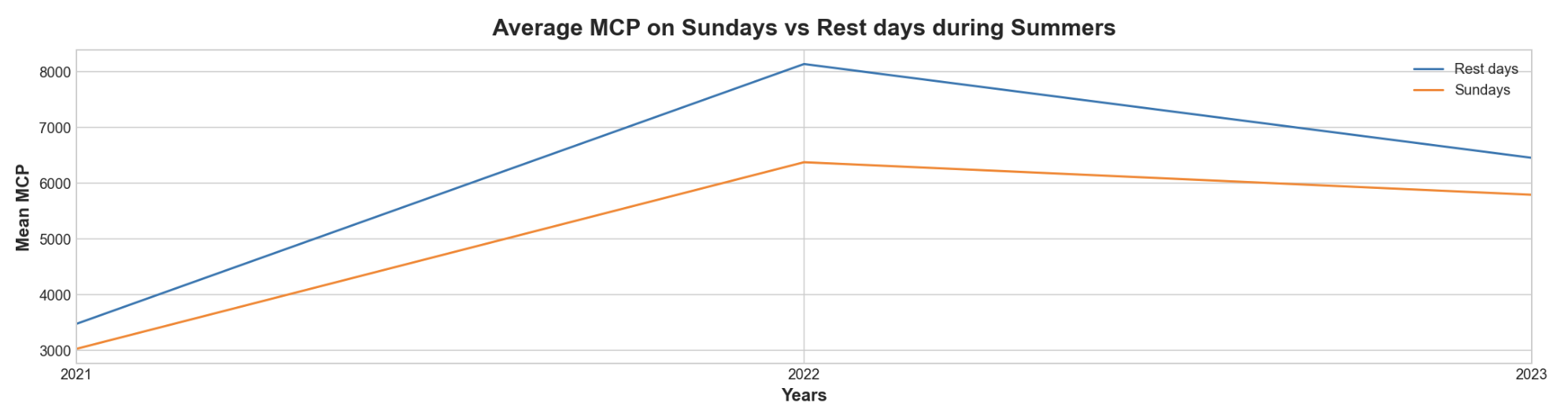
### **C. Monsoon Data**

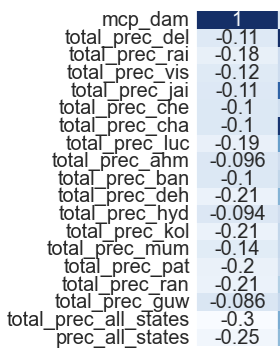




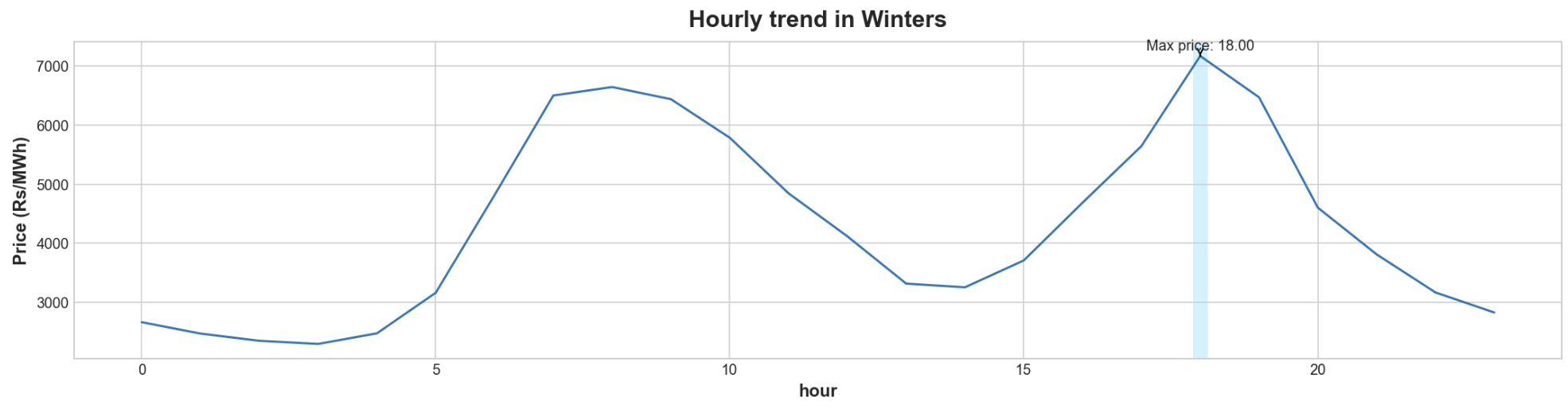
### **D. Summer Data**

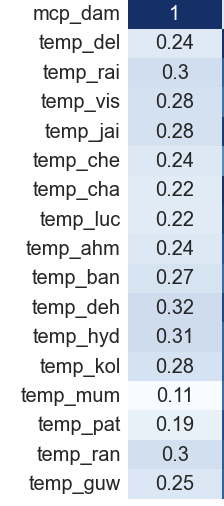
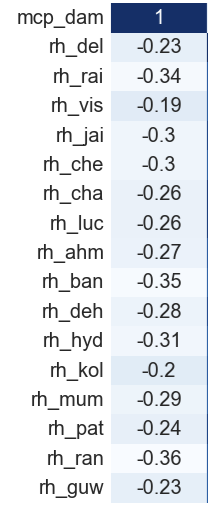






### **E. Winter Data**



# **Feature Engineering**

### **Time and Season Related Features:**

Features for date, year, month, hour, day of month(dom), day of week(dow), week of year(woy), time block,

Time of day: morning, day, evening, night,

Season of the year: summer, monsoon, autumn, winter.

### **Interaction Features:**

Certain interaction features are created by multiplying price with time block, with different time of day, with all weather parameters, multiplying hour by day of week etc.

### **Differencing:**

Differences between current day and previous several days Market Clearing Price, Cleared Volume, Purchase Bid and Sell Bid.

### **Lags:**

Hours and day lags for Market Clearing Price, Cleared Volume, Purchase Bid and Sell Bid.

### **Exponential Moving Averages(EMA):**

EMA for Market Clearing Price, Cleared Volume, Purchase Bid and Sell Bid.

### **Hourly Mean:**

Hourly mean for Market Clearing Price, Cleared Volume, Purchase Bid and Sell Bid.

### **Daily Min, Max and Mean:**

Daily min, max and mean for Market Clearing Price, Cleared Volume, Purchase Bid and Sell Bid.

### **Averages:**

Last 2, 3 and 5 days averages at the same time block for Market Clearing Price, Cleared Volume,

Purchase Bid and Sell Bid.

### **Cyclic Time Indicators:**

Cyclic Indicators like sine and cosine of time block, hour, day of week, day of year.

### **Other Features:**

Next day Sunday or Holiday, Covid first and second wave

### **Peak Value Feature:**

Peak value for Market Clearing Price each day.

### **Mean of Weather Parameters:**

Mean value for temperature, apparent temperature, relative humidity, precipitation rate for the day.

### **Difference of Weather Parameters:**

1-day difference as well as percentage change in weather parameters are created.

### **Precipitation Features:**

Total precipitation for each city for whole day, Sum of total precipitation at each time block,

Sum of total precipitation for the whole day.

# **Model Development**

### **Feature Selection:**

Firstly, SelectKBest feature selection method is used which selects the most relevant features by considering the correlation with the target variable and removing redundant features.

Secondly, ‘feature importances’ method is used which is available for LightGBM and CatBoost.

### **Description of Models:**

**Model1(LightGBM):** Multivariate analysis using above mentioned features and feature selection methods and using LightGBM Regressor.

**Model2(CatBoost):** Multivariate analysis using above mentioned features and feature selection methods and using CatBoost Regressor.

**Model3(LightGBM+PCA):** Multivariate analysis using above mentioned features and using PCA as a feature selection method.

**Model4(t2-t1):** Multivariate analysis using above mentioned features and feature selection methods, LightGBM Regressor as an algorithm and taking target variable as difference of current and previous day.

**Model5(pct):** Multivariate analysis using above mentioned features and feature selection methods, LightGBM Regressor as an algorithm and taking target variable as percentage change of current and previous day.

### **Results:**

All the models are tested for dates between 14-Feb to 18-Feb.

| **Models** | **Date** | **14/02/23** | **15/02/23** | **16/02/23** | **17/02/23** | **18/02/23** | **Average** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model-1** | **MAE** | 523.16 | 424.96 | 376.42 | 327.37 | 510.27 | 432.436 |
| **MAPE** | 8.62 | 7.03 | 5.98 | 4.95 | 7.57 | 6.83 |
|  |  |  |  |  |  |  |  |
| **Model-2** | **MAE** | 478.33 | 595.85 | 392.6 | 541.14 | 645.82 | 530.75 |
| **MAPE** | 7.48 | 10.24 | 6.28 | 7.04 | 9.49 | 8.11 |
|  |  |  |  |  |  |  |  |
| **Model-3** | **MAE** | 392.73 | 581.83 | 435.67 | 486.15 | 632.18 | 505.71 |
| **MAPE** | 5.88 | 9.25 | 6.96 | 6.24 | 9.83 | 7.63 |
|  |  |  |  |  |  |  |  |
| **Model-4** | **MAE** | 474.98 | 411.81 | 351.1 | 389.66 | 662.52 | 458.01 |
| **MAPE** | 7.34 | 6.62 | 5.05 | 5.18 | 9.2 | 6.98 |
|  |  |  |  |  |  |  |  |
| **Model-5** | **MAE** | 457.68 | 709.13 | 478.86 | 613.04 | 646.37 | 581.02 |
| **MAPE** | 7.17 | 11.01 | 6.71 | 7.52 | 9.41 | 8.36 |

# **Conclusions**

The present study establishes that the day-ahead electricity price in IEX exhibits nonstationarity, seasonality and time-varying volatility. To model time-varying volatility, the LightGBM model has been employed. The findings of this study may have implications on the way buyers and sellers anticipate and/or model electricity price volatility and thus their bids in the day and week-ahead market.

In terms of MAE, the LightGBM model has performed the best whereas if we consider MAPE as our evaluation metrics, t2-t1 model has performed slightly better than the LightGBM model.