STM JOURNALS

Research & Reviews: Journal of
Statistics

https://journals.stmjournals.com/etsy

ISSN: 2278-2273 (Online)
ISSN: 2348-7909 (Print)
Volume 13, Issue 2, 2024
May—August
DOI (Journal): 10.37591/RRJoST

Research RRJoST

Exploratory Analysis of the Statistical Signatures of Hourly Electricity Demand Data of India from 2022-2024

Rupsha Bhattacharyya[1,2,*]

Abstract

Electricity demand and consumption and its evolution over time is an important indicator of the development status of a country. The time series data pertaining to the electrical demand curves of India for the years 2022, 2023 and 2024 are analyzed in this work. For each month, the data are clustered according to Saturdays, Sundays and weekdays and the cluster average profiles are used for further analysis. Descriptive statistical analyses are first reported, followed by preliminary assessment of the time series characteristics of the available demand data sets. Year-on-year changes in demand curves are also analyzed, and annual demand is seen to have grown by 7.51% between 2022 and 2023 and by 5.58% between 2023 and 2024 respectively. However, demand changes across the same month in different years is very heterogeneous and show variations from -5.68% to +22.10%, due to weather-related factors. The non-stationary nature of the monthly average demand data time series is established due to changing mean and variance throughout the year. The coefficient of variation is seen to vary by month, but for the same months in different years it is much more stable. The temperature sensitivity of demand is found to be 2142-2576 GWh(e)/month/°C change of mean temperature. Higher variability in diurnal demand is observed in winter months than in summer. Z scores are calculated for each monthly demand profile and they are found to lie between ±2 in all cases. These estimates can be useful in forecasting power demand for designing day ahead and short-term electricity markets.

Keywords: Data analysis; demand; electricity; India; power; statistics; time series

INTRODUCTION

Energy is a crucial component of the welfare and development status of a nation [1]. Rapidly developing economies and emerging markets like India are expected to be the major drivers of the energy demand growth at the global level in the near and midterm time frame. In fact, the growth of energy demand in India has been among the highest in the world in recent times and this high growth rate is expected to continue in the near term as well to match its development objects [2].

***Author for Correspondence**
Rupsha Bhattacharyya
E-mail: rupshabhattacharyya1986@gmail.com

[1]Scientific Officer F, Applied Systems Analysis, Homi Bhabha National Institute, Mumbai, India
[2]Upgrading Plant Design and Fabrication Section, BARC, Mumbai, India

Received Date: January 09, 2025
Accepted Date: January 16, 2025
Published Date: May 29, 2025

Citation: Rupsha Bhattacharyya. Exploratory Analysis of the Statistical Signatures of Hourly Electricity Demand Data of India from 2022-2024. Research & Reviews: Journal of Statistics. 2024; 13(2): 35–57p.

Energy consumption in the form of electricity represents one share of the total final energy consumption. Globally this share is about 20% at present, with other final energy consumption forms being refined petroleum products for mobility services, natural gas for industrial and domestic heating services, and so on [3]. As the global clean energy transition is underway to ensure climate change mitigation, increased levels of electrification using all sources of clean electricity generation and a massive build out of electricity infrastructure will continue to play a major role in this deep structural change of the energy system.

Understanding of the changing electricity demand nature is an important aspect of energy analytics studies today, particularly with a view to predict future demand trends and plan the optimal mix of power supply options. This work is motivated by the electricity sector of India and the evolution of gross national electricity demand in the recent past. It uses openly accessible data on India's electricity consumption covering years 2022, 2023 and 2024 and performs exploratory statistical analysis to summarize the trends each year and how they are changing year-on-year, while also looking at correlations of power consumption to important climate variables like temperature.

Prior studies of similar nature but for other regions or nations have been reported in literature; most of them use time series techniques to forecast electricity demand. For example, Idhom et al [4] have performed time series regression analysis to estimate electricity demand in Indonesia. Lee et al apply a number of techniques to forecast energy demand on a Malaysian University Campus, using data from previous six years [5]. Kim et al [6] have used a combination of time series clustering methods and time series forecasting techniques such as autoregressive moving average methods to forecast energy demand in a microgrid setting, considering variations in energy consumption patterns at the household level. Kouakou [7] has used time series techniques and causality tests to map the development of Cote d'Ivoire on to the electricity demand in the country. Yu et al [8] have used clustering techniques to study demand consumption patterns of several households in a community and have used it for future load forecasting. Chowksi et al [9] have deployed clustering techniques with the objective of studying demand response mechanisms associated with smart meter use at the domestic level in India. Domestic load profiles have been analysed by Praveen et al [10] for customer load segmentation in India, using smart meter data.

In these studies, detailed analysis of the statistical nature of previous data for India have not been reported, though some have extracted and used general features of the demand curves in their assessments. The emphasis has primarily been on forecasting future values using different algorithms. This present work has the novelty that it aims to first make a detailed investigation of the statistical features of electricity demand in India as a preparatory activity for forecasting studies, particularly for the use of power exchanges and energy markets, which are still evolving in India. Such data and analysis for India for the years from 2008 up to 2016 have been reported by Grid Controller of India (available at https://posoco.in/en/reports/electricity-demand-pattern-analysis/), but compilation and analyses of more recent power demand data have not been presented by them. This gap area is partly addressed in this study by considering the 3 most recent years (2022, 2023 and 2024), marking the phase of recovery from the economic shocks brought on by the COVID 19 pandemic and a return to the normal electricity demand evolution in India.

DATA DESCRIPTION AND ANALYSES

Data Sets Used

The electrical load data for India are extracted from the IEA Real Time Electricity Tracker tool by selecting the date range of interest (available at https://www.iea.org/data-and-statistics/data-tools/real-time-electricity-tracker?tracker=true&from=2024-9-17&to=2024-10-17&category=demand, last accessed on 1.1.2025).

For each month of the year from 2022-2024, the data are available as hourly loads averaged over the weekdays, Saturdays, and Sundays of the month. Therefore, there are three representative profiles for each month, 12 X 3 = 36 profiles for the entire year and a total of 36 X 3 = 108 profiles for the three years considered in this study. These profiles (and the 108 X 24 = 3672 data points of power demand versus time of day) form the fundamental data set extracted and used in this study. In addition to it, the monthly average, minimum and maximum temperature data of India for these 36 months are also obtained from the Indian Meteorological Department's monthly climate summary bulletins (which are made freely available at https://imdpune.gov.in/cmpg/Product/Monthly_Climate_Summary/, last accessed on 18.1.2025).

ISSN: 2278-2273 (Online), ISSN: 2348-7909 (Print)

Calculations and Analyses Performed

The summary statistics calculations are performed from the load profiles for each month and this includes determination of the following parameters:

1. Determination of minimum, maximum, mean and standard deviation of hourly load curves for Saturdays, Sundays and weekdays of each month; calculation of coefficient of variance and the ratio of maximum to minimum hourly loads from each of these profiles; calculations correlation coefficients between pairs of demand profiles for each month
2. Total energy demand per day by integration of the daily load curve over time
3. Total annual electricity demand and monthly share of the total annual demand
4. Euclidean distance and cosine separation between a pair of time series to understand the extent of similarity between them, ratios of instantaneous demand at same time of same month in different years
5. Monthly average profile for each month of the year derived from the daily average profiles; determination of correlation coefficients between the monthly average profiles in a given year

The Euclidean distance between two time series P_1 and P_2 of equal lengths (in this case the hourly power demand data from two occasions, e.g., that for Saturday and Sunday of a given month) is defined as follows [11]:

$$D_{Euclid}(P_1, P_2) = \sqrt{\sum_{t=0}^{23}\left[\left(P_1(t) - P_2(t)\right)|\square|2\right]} \quad (1)$$

This is a measure of statistical similarity between two equal sized vectors, which in these cases are the time series data sets for any two days.

The cosine separation between two time series P_1 and P_2 of equal lengths is defined as follows [11]:

$$Sep_{cosine}(P_1, P_2) = \frac{\sum_{t=0}^{23} P_1(t)P_2(t)}{\sqrt{\sum_{t=0}^{23} P_1(t)^2} \times \sqrt{\sum_{t=0}^{23} P_2(t)^2}} \quad (2)$$

This too is a measure of statistical similarity between two equal sized vectors of time series data elements, but this measure is dimensionless and is meaningful even when the corresponding elements of the two vectors are of different magnitudes.

The correlation coefficient between two time series P_1 and P_2 (of equal length) shows the extent of linear association between the time series and it is calculated as [12]

$$Corr(P_1, P_2) = \frac{Covariance(P_1, P_2)}{\sqrt{Var(P_1)} \times \sqrt{Var(P_2)}} \quad (3)$$

The ratio between instantaneous power demand at the same time of day but on different occasions (e.g., ratio of power demand at time t of Saturdays in January 2022 and January 2023) is calculated as follows:

$$R_d(t) = \frac{P_2(t)}{P_1(t)} \quad (4)$$

If $P(t)$ describes the value of power demand at the t^{th} hour of the day, total electrical energy demanded during the whole day is calculated as

$$E = \int_{t=0}^{t=23} P(t)dt \quad (5)$$

In this study, 3 metrics are evaluated for a given time series power demand data set to understand its characteristics. The first of these is the autocorrelation function (*ACF*) of a time series $P(t)$ with itself at different time lags given by k = 0, 1, 2, etc. It is defined as [13]

$$ACF(k) = \frac{Covariance(P(t), P(t+k))}{\sqrt{variance}} \tag{6}$$

The partial autocorrelation function (*PACF*) of a time series *P(t)* is actually a conditional correlation coefficient and it is given by [13]

$$PACF(k) = Corr\big(P(t), P(t-k) \vee P(t-1), P(t-2) \ldots P(t-k+1)\big) \tag{7}$$

The ACF and PACF profiles held understand whether the time series data are better represented as auto-regressive or moving average series or whether a combination of these two should be considered in modeling them.

The cross-correlation function (*CCF*) between two time series X_t and Y_t is given by [13]

$$CCF(k) = \sum_{t=k+1}^{n} \left(X_t - \overline{X}\right) \tag{8}$$

RESULTS AND DISCUSSION

Descriptive Statistical Analysis of India’s Electricity Demand Data for 2022-2024

Figures 1a, 1b and 1c show the hourly average load curves of India for each month of the years 2022, 2023 and 2024 respectively. Some important observations and inferences are described next.

The maximum to minimum instantaneous demand ratio in a given day varies between 1.09 to 1.51 in 2022, 1.08 to 1.50 in 2023 and 1.09 to 1.48 in 2024, depending on the season. This is a measure of demand variability and the extent of supply side flexibility that is expected to be required in future as well. The coefficient of variance calculated from a given load curve also provides similar information, showing the spread about the mean demand during a day. It is also seen to have highest values during winter months and lowest values during summer, lying between 3 to 6% in summer months and 10 to 14% in winter months respectively. Thus, summer load curves from March to October are somewhat more uniform than those of the winter months from November to February where higher variability is observed. The change in means and standard deviations of demand from month to month in a year indicate that the demand data time series are not stationery.

The monthly allocation of total annual electrical energy consumption is correspondingly found to be non-uniform, being about 7.5% per month in the winter months, and going over 9% per month in the peak summer months of April and May every year. The total annual electric power consumption was 1485 TWh in 2022, 1597 TWh in 2023 (i.e., 7.51% increase on the year-on-year basis) and 1686 TWh in 2024 respectively (i.e., 5.58% increase on the year-on-year basis).

For any given month, the pairwise correlation coefficients between load curves of Saturday-Sunday, Sunday-weekday and weekday-Saturday do not show any clear trend. They approach unity in many cases but some months show substantially lower coefficient between some or all of these pairs of data sets. Thus, statistically the profiles on weekends and weekdays of a month can be significantly different in different months. From the power generator’s point of view, this reflects a need to have different and dynamically varying supply side strategies for each day as even within the same month (with presumably similar climatic conditions) also, the characteristics of the days are very different from power demand point of view.

In course of a day, the lowest demand is seen to occur between 2 AM to 4 AM, irrespective of the season. The highest demands occur at different times of the day depending upon season, during 10 AM to 3 PM and again during evening hours of 7 PM to 8 PM.

Additional statistical inferences from the power demand data of these 3 years are provided in Tables 1 to 3.

Table 1. Summary statistics of India's electricity demand data for 2022.

Month	**January**			**February**			**March**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	125.92	118.10	119.21	134.08	135.41	134.01	154.41	156.44	157.79
Maximum load (GW)	181.74	170.32	179.43	190.31	183.32	188.99	188.60	184.04	192.93
Maximum to minimum ratio	1.44	1.44	1.51	1.42	1.35	1.41	1.22	1.18	1.22
Mean load (GW)	153.03	145.01	151.63	163.48	157.06	162.35	171.81	170.15	175.57
Standard deviation (GW)	18.44	17.98	20.99	19.53	15.98	19.20	11.00	8.31	11.17
Pair wise linear correlation coefficient	0.99	1.00	0.99	0.98	0.97	0.98	0.97	0.95	0.97
Coefficient of variation (%)	12.05	12.40	13.85	11.94	10.18	11.83	6.40	4.89	6.36
Energy demand/day (GWh)	3672.76	3480.31	3639.08	3923.55	3769.46	3896.41	4123.42	4083.71	4213.75
Total demand for month (GWh)	112186.03			108700.24			129744.77		
Monthly share of annual demand (%)	7.55			7.32			8.74		
Month	**April**			**May**			**June**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	172.94	168.70	172.99	172.81	164.79	171.33	173.11	167.39	174.79
Maximum load (GW)	197.29	184.09	196.20	195.95	185.12	194.35	196.80	189.83	196.40
Maximum to minimum ratio	1.14	1.09	1.13	1.13	1.12	1.13	1.14	1.13	1.12
Mean load (GW)	184.93	177.57	183.67	184.05	177.02	182.31	187.11	179.71	186.10
Standard deviation (GW)	7.90	3.99	6.90	7.74	5.80	7.35	7.52	5.40	7.03
Pair wise linear correlation coefficient	0.71	0.67	0.90	0.64	0.58	0.97	0.76	0.73	0.99
Coefficient of variation (%)	4.27	2.24	3.76	4.20	3.28	4.03	4.02	3.01	3.78
Energy demand/day (GWh)	4438.41	4261.61	4408.06	4417.17	4248.43	4375.48	4490.67	4313.05	4466.38
Total demand for month (GWh)	131807.75			135171.39			133475.24		
Monthly share of annual demand (%)	8.88			9.10			8.99		
Month	**July**			**August**			**September**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	159.79	154.87	161.12	163.14	158.18	161.72	161.36	161.11	163.74
Maximum load (GW)	182.78	173.58	182.20	186.25	176.87	183.88	187.57	179.25	188.17
Maximum to minimum ratio	1.14	1.12	1.13	1.14	1.12	1.14	1.16	1.11	1.15
Mean load (GW)	172.94	163.57	173.47	177.20	170.86	174.66	176.75	168.79	177.67
Standard deviation (GW)	7.47	5.65	6.70	7.34	4.75	6.76	7.88	5.21	7.36
Pair wise linear correlation coefficient	0.69	0.67	0.99	0.21	0.25	0.97	0.78	0.77	1.00
Coefficient of variation (%)	4.32	3.46	3.86	4.14	2.78	3.87	4.46	3.09	4.14
Energy demand/day (GWh)	4150.44	3925.63	4163.20	4252.89	4100.62	4191.92	4241.90	4050.88	4264.14
Total demand for month (GWh)	127807.55			129828.20			126982.20		
Monthly share of annual demand (%)	8.61			8.74			8.55		
Month	**October**			**November**			**December**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	136.58	138.41	135.80	132.75	130.68	132.12	131.64	129.91	129.84
Maximum load (GW)	172.15	161.81	169.13	178.96	171.78	177.73	192.72	183.46	191.44

Maximum to minimum ratio	1.26	1.17	1.25	1.35	1.31	1.35	1.46	1.41	1.47
Mean load (GW)	156.00	150.70	154.38	159.89	151.40	157.84	165.42	157.06	163.94
Standard deviation (GW)	10.89	6.76	9.95	15.72	13.97	15.97	21.91	18.61	21.56
Pair wise linear correlation coefficient	0.98	0.97	1.00	0.96	0.97	0.98	0.99	0.99	1.00
Coefficient of variation (%)	6.98	4.49	6.44	9.83	9.23	10.12	13.24	11.85	13.15
Energy demand/day (GWh)	3744.03	3616.68	3705.08	3837.36	3633.59	3788.06	3970.16	3769.44	3934.49
Total demand for month (GWh)	114610.23			113221.12			121487.34		
Monthly share of annual demand (%)	7.72			7.62			8.18		

Table 2. Summary statistics of India's electricity demand data for 2023.

Month	**January**			**February**			**March**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	136.4	136.92	138.84	152.04	153.87	150.69	159.24	153.69	154.7
Maximum load (GW)	204.05	197.07	202.41	205.62	191.02	204.36	193.99	182.55	193.24
Maximum to minimum ratio	1.50	1.44	1.46	1.35	1.24	1.36	1.22	1.19	1.25
Mean load (GW)	172.30	165.11	170.12	178.09	170.27	177.22	174.91	166.97	173.92
Standard deviation (GW)	23.09	20.03	21.23	18.37	12.86	18.10	11.07	9.06	12.22
Pair wise linear correlation coefficient	0.98	0.98	0.98	0.98	0.97	1.00	0.98	0.97	0.99
Coefficient of variation (%)	13.40	12.13	12.48	10.31	7.55	10.22	6.33	5.43	7.03
Energy demand/day (GWh)	4135.24	3962.72	4082.93	4274.24	4086.51	4253.25	4197.78	4007.33	4173.96
Total demand for month (GWh)	122096.09			118508.00			128821.52		
Monthly share of annual demand (%)	7.65			7.42			8.07		
Month	**April**			**May**			**June**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	174.38	160.56	171.53	169.41	170.81	169.9	182.52	170.34	182.38
Maximum load (GW)	198.91	179.44	194.98	199.58	193.57	197.12	209.16	201.28	208.7
Maximum to minimum ratio	1.14	1.12	1.14	1.18	1.13	1.16	1.15	1.18	1.14
Mean load (GW)	187.12	171.64	185.19	185.83	182.14	184.52	196.43	186.37	195.93
Standard deviation (GW)	7.65	4.63	7.28	9.86	6.61	8.94	8.93	7.56	8.83
Pair wise linear correlation coefficient	0.73	0.69	0.98	0.65	0.68	0.99	0.59	0.59	1.00
Coefficient of variation (%)	4.09	2.70	3.93	5.30	3.63	4.85	4.55	4.06	4.51
Energy demand/day (GWh)	4490.79	4119.30	4444.52	4460.02	4371.36	4428.56	4714.21	4472.84	4702.30
Total demand for month (GWh)	131940.85			137182.40			140198.80		
Monthly share of annual demand (%)	8.26			8.59			8.78		
Month	**July**			**August**			**September**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	176.33	170.26	174.07	188.96	191.1	189.51	185.33	180.59	180.62
Maximum load (GW)	199.29	191.08	197.88	221.56	211.41	219.62	212.56	194.42	208.55
Maximum to minimum ratio	1.13	1.12	1.14	1.17	1.11	1.16	1.15	1.08	1.15
Mean load (GW)	190.07	180.58	188.83	206.29	201.26	205.62	200.19	188.56	196.95
Standard deviation (GW)	7.64	6.83	8.03	10.58	6.72	9.89	8.35	4.89	8.94

Pair wise linear correlation coefficient	0.83	0.83	0.99	0.93	0.92	1.00	0.82	0.81	0.99
Coefficient of variation (%)	4.02	3.78	4.25	5.13	3.34	4.81	4.17	2.59	4.54
Energy demand/day (GWh)	4561.56	4334.03	4532.02	4951.04	4830.29	4934.76	4804.56	4525.51	4726.91
Total demand for month (GWh)	139650.37			152624.80			141389.95		
Monthly share of annual demand (%)	8.75			9.56			8.86		
Month	**October**			**November**			**December**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	171.23	170.2	167.41	147.1	143.61	140.34	135.06	134.83	133.18
Maximum load (GW)	211.17	197.31	205.8	197.01	182.38	188.62	194.31	190.56	196.39
Maximum to minimum ratio	1.23	1.16	1.23	1.34	1.27	1.34	1.44	1.41	1.47
Mean load (GW)	192.50	184.33	188.15	173.49	161.91	167.52	166.52	161.60	166.96
Standard deviation (GW)	13.07	8.60	12.66	16.70	13.20	16.85	20.00	18.66	21.58
Pair wise linear correlation coefficient	0.98	0.98	1.00	0.94	0.92	0.99	0.97	0.98	0.99
Coefficient of variation (%)	6.79	4.67	6.73	9.63	8.15	10.06	12.01	11.55	12.93
Energy demand/day (GWh)	4620.04	4423.84	4515.62	4163.64	3885.76	4020.46	3996.54	3878.42	4006.98
Total demand for month (GWh)	139943.00			120647.72			123521.38		
Monthly share of annual demand (%)	8.77			7.56			7.74		

Table 3. Summary statistics of India's electricity demand data for 2024.

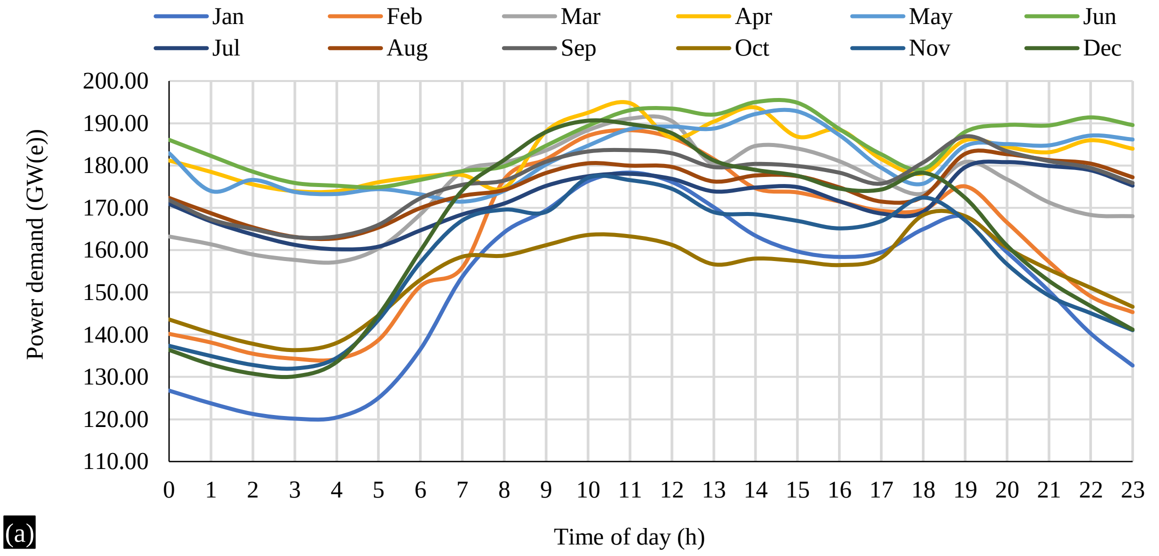
Month	**January**			**February**			**March**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	146.01	145.71	143.01	154.03	154.45	152.63	169.56	165.85	165.74
Maximum load (GW)	215.40	208.77	213.43	219.40	208.55	215.07	216.51	199.38	210.09
Maximum to minimum ratio	1.48	1.43	1.49	1.42	1.35	1.41	1.28	1.20	1.27
Mean load (GW)	184.50	176.55	181.04	186.90	178.82	184.77	191.97	181.07	188.13
Standard deviation (GW)	23.97	20.98	24.37	21.58	17.92	20.69	15.14	10.49	14.26
Pair wise linear correlation coefficient	0.98	0.99	0.99	0.98	0.97	1.00	0.98	0.96	0.99
Coefficient of variation (%)	12.99	11.88	13.46	11.55	10.02	11.20	7.89	5.80	7.58
Energy demand/day (GWh)	4428.02	4237.28	4344.97	4485.55	4291.57	4434.47	4607.21	4345.57	4515.20
Total demand for month (GWh)	134595.51			128232.35			139583.10		
Monthly share of annual demand (%)	7.98			7.61			8.28		
Month	**April**			**May**			**June**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	190.04	183.24	186.84	198.12	187.01	194.94	194.16	190.57	198.59
Maximum load (GW)	210.65	204.65	210.63	230.26	215.44	228.04	226.62	216.26	232.00
Maximum to minimum ratio	1.11	1.12	1.13	1.16	1.15	1.17	1.17	1.13	1.17
Mean load (GW)	202.01	194.38	200.91	213.66	202.71	210.09	211.16	204.20	213.61
Standard deviation (GW)	6.37	4.84	7.58	10.13	7.87	10.33	9.97	7.66	10.30
Pair wise linear correlation coefficient	0.64	0.53	0.98	0.70	0.74	0.98	0.78	0.81	0.99
Coefficient of variation (%)	3.15	2.49	3.78	4.74	3.88	4.92	4.72	3.75	4.82
Energy demand/day (GWh)	4848.27	4665.06	4821.72	5127.83	4864.94	5042.13	5067.95	4900.74	5126.53
Total demand for month (GWh)	144131.16			155940.07			152374.05		

Monthly share of annual demand (%)	8.55			9.25			9.04		
Month	**July**			**August**			**September**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	189.70	186.76	189.20	178.40	176.52	178.17	178.89	178.99	179.74
Maximum load (GW)	211.39	204.31	211.92	205.89	195.05	204.21	208.13	197.76	207.46
Maximum to minimum ratio	1.11	1.09	1.12	1.15	1.10	1.15	1.16	1.10	1.15
Mean load (GW)	202.69	195.83	203.19	195.60	186.79	194.22	196.67	189.53	197.26
Standard deviation (GW)	6.95	5.87	7.74	9.13	6.13	8.49	9.37	6.38	9.26
Pair wise linear correlation coefficient	0.83	0.81	0.98	0.72	0.76	1.00	0.84	0.87	0.99
Coefficient of variation (%)	3.43	3.00	3.81	4.67	3.28	4.37	4.76	3.37	4.70
Energy demand/day (GWh)	4864.60	4699.82	4876.48	4694.33	4482.88	4661.33	4720.03	4548.72	4734.25
Total demand for month (GWh)	150416.72			143952.43			141042.97		
Monthly share of annual demand (%)	8.92			8.54			8.37		
Month	**October**			**November**			**December**		
Type of day	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***	***Sat***	***Sun***	***Week***
Minimum load (GW)	170.38	164.98	171.37	142.06	144.72	144.40	138.50	136.25	138.18
Maximum load (GW)	200.33	195.14	207.33	199.47	185.67	196.36	207.53	199.32	211.95
Maximum to minimum ratio	1.18	1.18	1.21	1.40	1.28	1.36	1.50	1.46	1.53
Mean load (GW)	186.85	180.16	191.58	171.54	166.30	173.97	174.59	166.67	176.97
Standard deviation (GW)	8.40	8.11	10.65	18.45	13.81	18.67	23.77	20.75	25.44
Pair wise linear correlation coefficient	0.96	0.96	1.00	0.99	0.99	0.99	0.99	0.99	1.00
Coefficient of variation (%)	4.49	4.50	5.56	10.75	8.30	10.73	13.62	12.45	14.38
Energy demand/day (GWh)	4484.29	4323.75	4597.88	4117.00	3991.28	4175.33	4190.14	4000.07	4247.16
Total demand for month (GWh)	140983.40			124232.05			130198.43		
Monthly share of annual demand (%)	8.36			7.37			7.72		

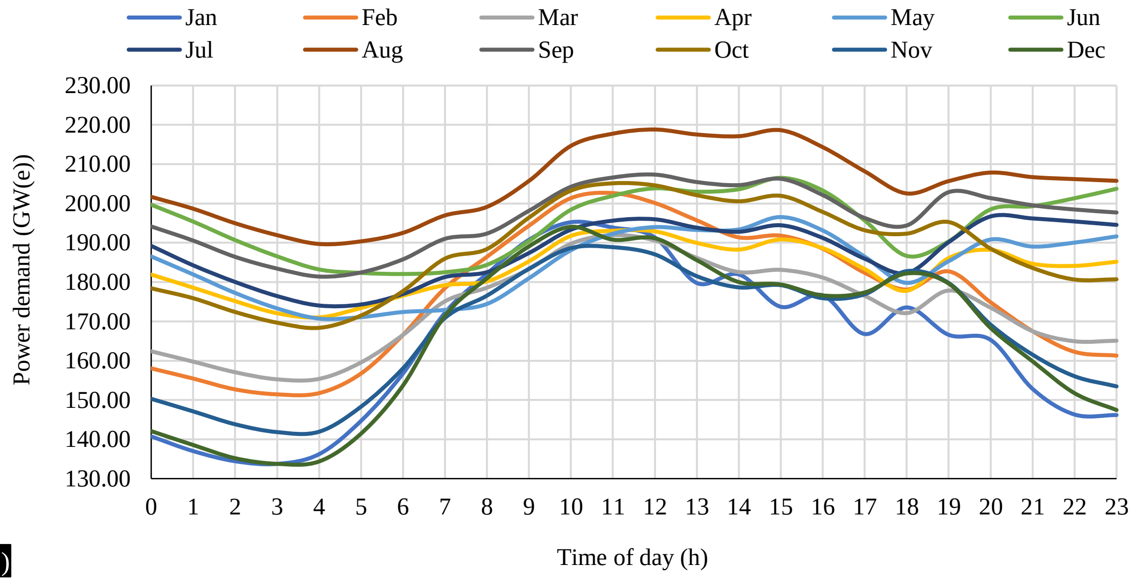
Figures 1d, 1e and 1f show the monthly demand data from the Figures 1a, 1b and 1c respectively in the form of histograms. The distributions are seen to be skewed to the left in all 3 cases (most notably in case of year 2022), indicating that in course of the year, above average demand is experienced on fewer occasions, limited to the summer months.

These distributions are therefore not normal. This also indicates a typical characteristic of energy system design and an important point for energy policy formulation – for a truly reliable energy system, the supply side must take care of the most extreme demand situations, without a loss of load situation. The average values are not the basis of system design and the extreme values are important; as a result, the power generation and transmission infrastructure are generally not always utilized to full capacity.

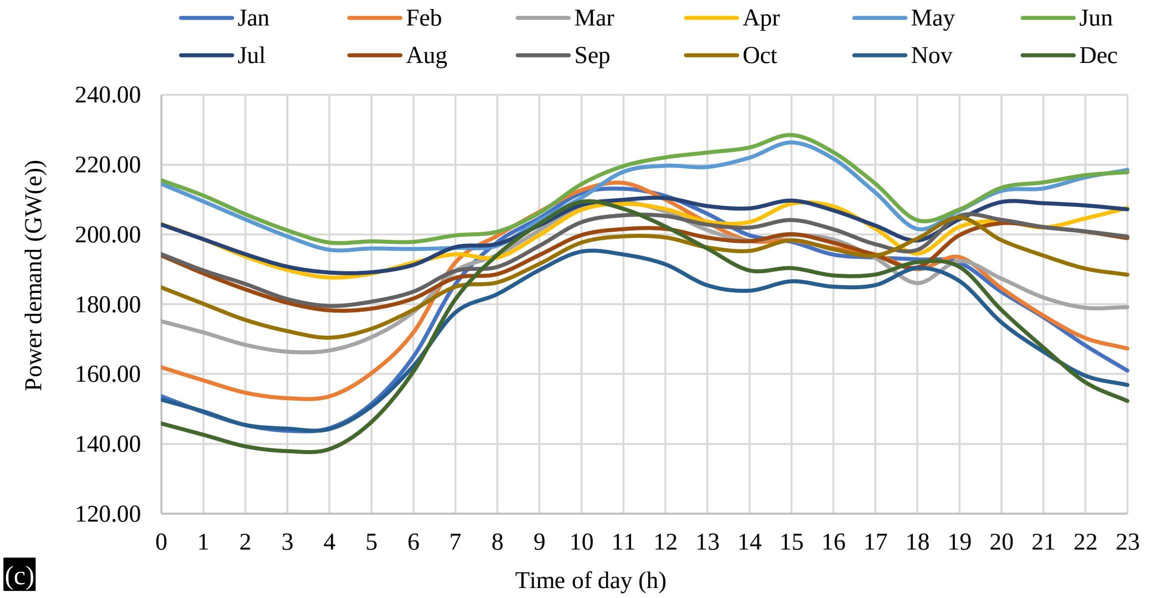
Figures 2a, 2b and 2c show pairwise linear correlation coefficient values between monthly power demand profiles in the form of heat maps, one for 2022, 2023 and 2024 respectively. For example, the correlation between average monthly demand curve of January 2022 with that of February 2022 is 0.987, while that between January 2022 and May 2022 is 0.537. The correlations between adjacent months are fairly high and they drop sharply after about 3 to 4 months in all cases. This is a representation of the very strong seasonality (mainly due to weather i.e., temperature and rainfall effects) on electricity demand in India; long-range correlations are weak and only the short-range ones are substantially high.


Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Power demand (GW(e))
200.00
190.00
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Time of day (h)

(a)


Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Power demand (GW(e))
230.00
220.00
210.00
200.00
190.00
180.00
170.00
160.00
150.00
140.00
130.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Time of day (h)

(b)


Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Power demand (GW(e))
240.00
220.00
200.00
180.00
160.00
140.00
120.00
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Time of day (h)

(c)

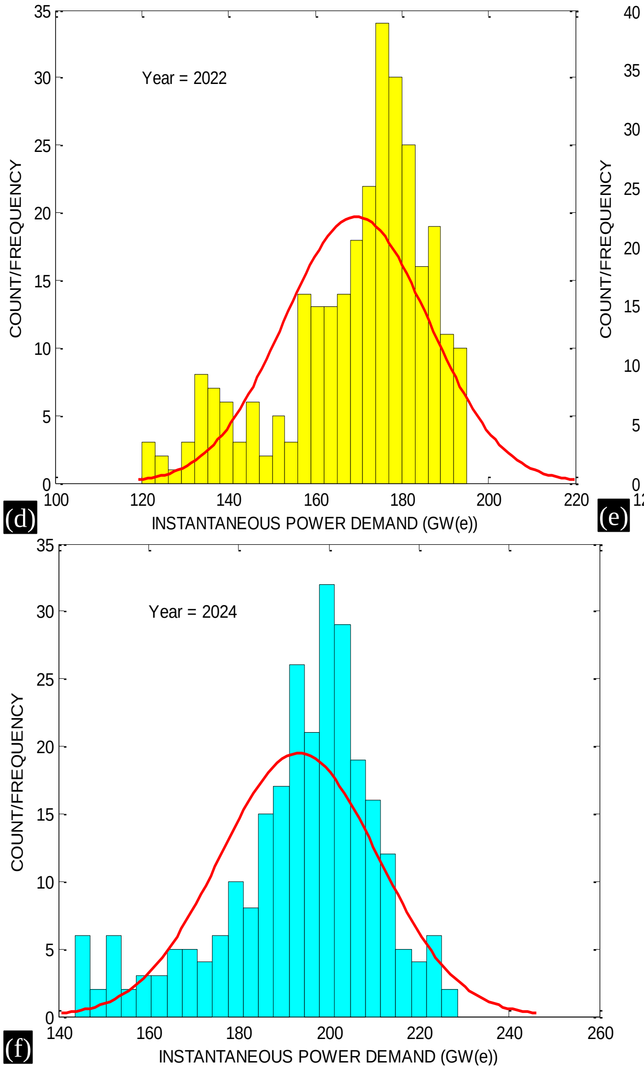


Figure 1. (a) Monthly average power demand curves of India in 2022. (b) Monthly average power demand curves of India in 2023. (c) Monthly average power demand curves of India in 2024. (d) Histogram of power demand data of India in 2022. (e) Histogram of power demand data of India in 2023. (f) Histogram of power demand data of India in 2024.

Figure 2d shows the correlation coefficients of power demand curves in same months of different years (e.g., between January 2022 and January 2023, between January 2023 and January 2024 and so on). These values are seen to be positive and consistently high (greater than 0.9) for all months except August, September and October. The difference may potentially be ascribed to difference in monsoon rainfall experienced in these specific months over the successive years and the resultant changes in the temperature patterns. There is a strong dependence of power demand on temperature (discussed and quantified further in Section 3.2), thus temperature anomalies have affected the power demand somewhat randomly in these months. As rainfall levels and corresponding temperature dips and drop in economic activities which may be hampered by rainfall (e.g., mining, supply chain disruptions) have not shown the same trends year on year, the power demand in the monsoon months of adjacent years has also become less correlated.

The share of annual electricity consumption in each month has been between 7.32-9.10% (with standard deviation of 0.63%) in the year 2022, 7.42-9.56% (with standard deviation of 0.66%) in 2023 and 7.37-9.25% (with 0.58% standard deviation) in 2024, respectively. Higher shares are noted for the summer months each year, showing the impact of growing cooling needs (in the form of refrigeration and air conditioning of buildings, cold supply chains for perishables) on power consumption in India. Looking at the difference in power demand at the monthly level each year, as shown in Figure 2e (e.g., between January 2022 and January 2023, and so on), the values are found to vary from -0.71% to 22.1% for the years 2022 and 2023, and from -5.68% to 13.67% for the years 2023 and 2024 respectively. Thus, demand shifts in corresponding months of different years have shown much greater variability than the average annual demand growth.

Temperature Effects on Electric Power Demand in India

The previous observation about seasonality of demand is further reinforced by Figure 3a which shows mean power demand of each month against the national average temperature, for that month, from 2019 to 2024.

(a)

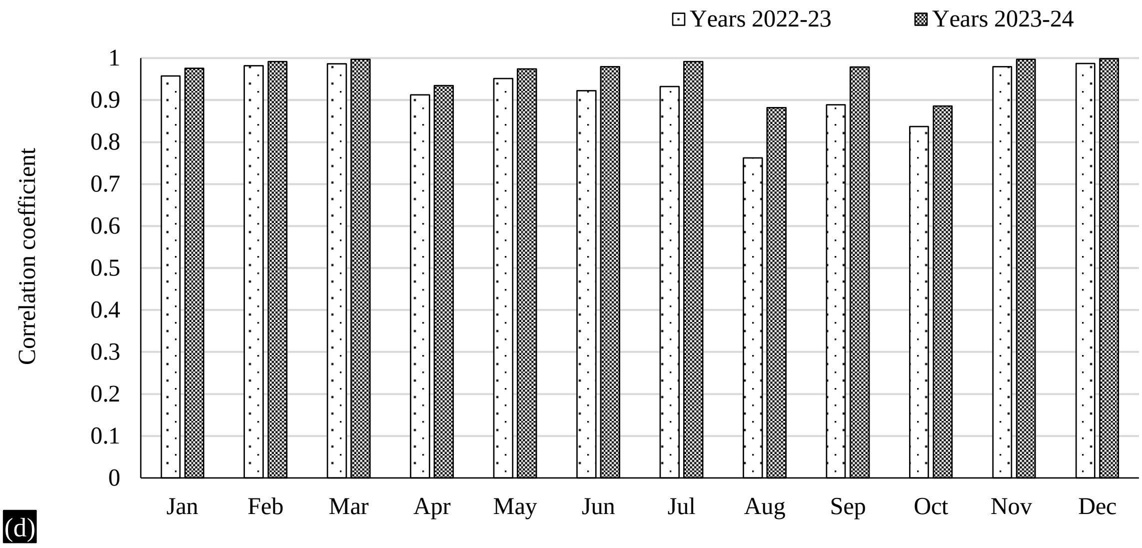
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	1.000											
Feb	0.987	1.000										
Mar	0.961	0.967	1.000									
Apr	0.712	0.736	0.771	1.000								
May	0.537	0.567	0.611	0.867	1.000							
Jun	0.603	0.619	0.678	0.886	0.966	1.000						
Jul	0.697	0.666	0.699	0.770	0.771	0.861	1.000					
Aug	0.754	0.721	0.757	0.772	0.740	0.829	0.988	1.000				
Sep	0.893	0.856	0.860	0.767	0.683	0.755	0.911	0.950	1.000			
Oct	0.935	0.890	0.866	0.584	0.414	0.481	0.679	0.760	0.915	1.000		
Nov	0.961	0.952	0.938	0.611	0.403	0.456	0.526	0.615	0.801	0.929	1.000	
Dec	0.965	0.966	0.951	0.640	0.425	0.477	0.534	0.617	0.790	0.902	0.992	1.000

(b)

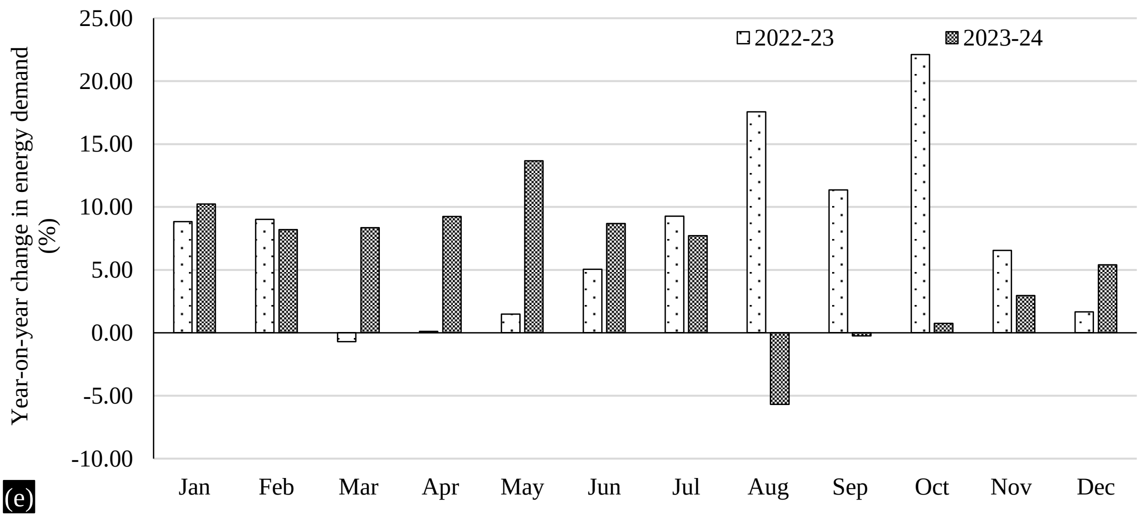
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	1.000											
Feb	0.980	1.000										
Mar	0.969	0.995	1.000									
Apr	0.743	0.821	0.862	1.000								
May	0.439	0.556	0.603	0.898	1.000							
Jun	0.298	0.424	0.475	0.817	0.970	1.000						
Jul	0.483	0.567	0.622	0.916	0.942	0.893	1.000					
Aug	0.724	0.818	0.846	0.961	0.926	0.858	0.877	1.000				
Sep	0.727	0.811	0.845	0.979	0.914	0.816	0.924	0.972	1.000			
Oct	0.926	0.973	0.977	0.891	0.704	0.573	0.686	0.909	0.904	1.000		
Nov	0.967	0.972	0.963	0.776	0.523	0.356	0.551	0.775	0.796	0.961	1.000	
Dec	0.976	0.976	0.966	0.769	0.502	0.338	0.542	0.763	0.784	0.954	0.997	1.000

(c)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	1.000											
Feb	0.995	1.000										
Mar	0.981	0.987	1.000									
Apr	0.602	0.587	0.684	1.000								
May	0.433	0.401	0.516	0.915	1.000							
Jun	0.497	0.466	0.576	0.907	0.993	1.000						
Jul	0.667	0.636	0.708	0.959	0.880	0.881	1.000					
Aug	0.664	0.629	0.688	0.932	0.823	0.818	0.989	1.000				
Sep	0.754	0.722	0.772	0.923	0.808	0.812	0.972	0.984	1.000			
Oct	0.844	0.810	0.819	0.794	0.646	0.661	0.858	0.891	0.949	1.000		
Nov	0.985	0.979	0.953	0.565	0.380	0.438	0.616	0.626	0.730	0.860	1.000	
Dec	0.994	0.993	0.967	0.548	0.354	0.417	0.609	0.615	0.713	0.829	0.992	1.000



(d)

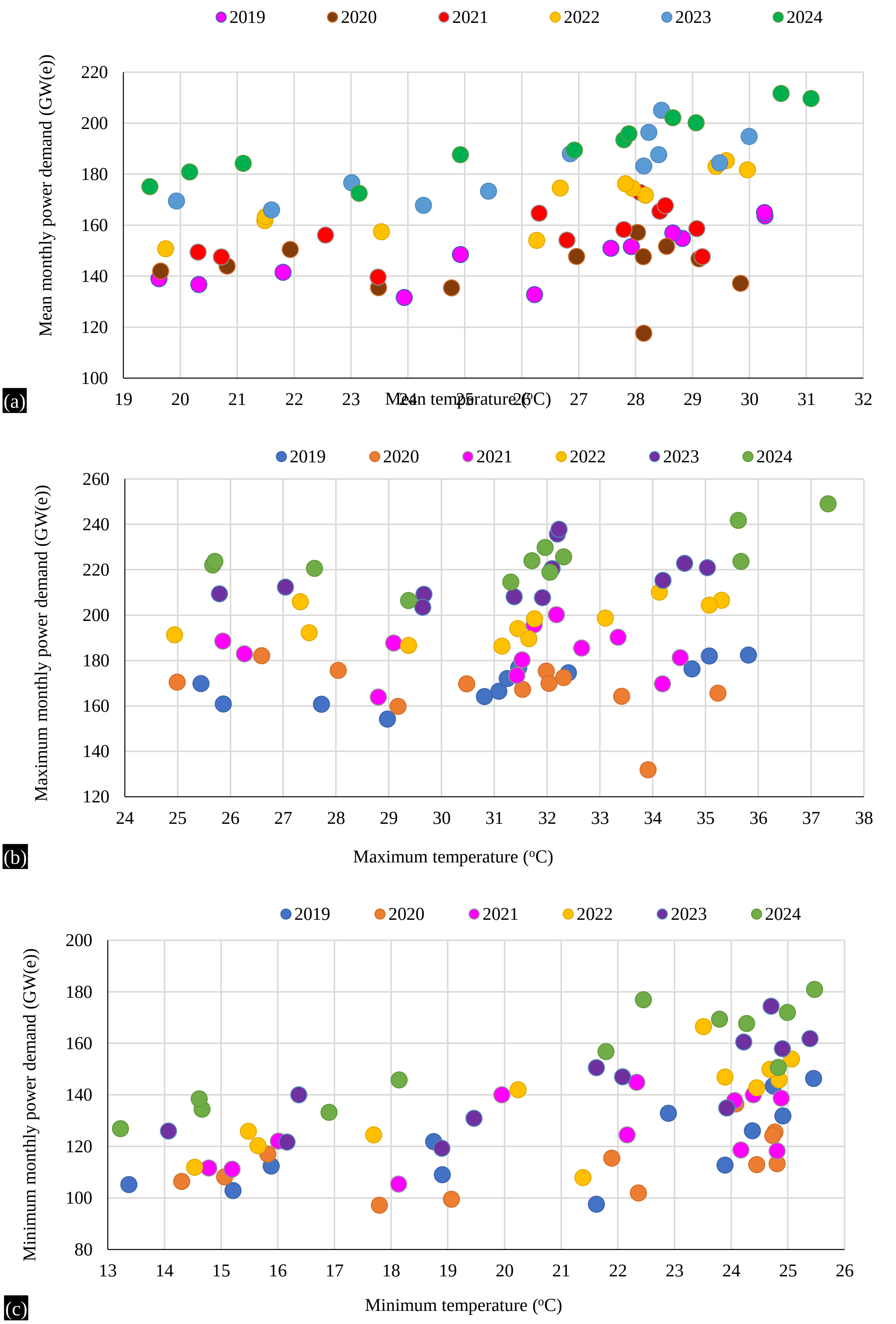


(e)

Figure 2. (a) Heat map of correlation coefficients of monthly average power demand profiles of India in year 2022. (b) Heat map of correlation coefficients of monthly average power demand profiles of India in year 2023. (c) Heat map of correlation coefficients of monthly average power demand profiles of India in year 2024. (d) Linear correlation coefficients of monthly average power demand profiles in respective months across two successive years in India. (e) Year-on-year per cent change of monthly total energy demand in India

The strong positive correlation of mean power demand and mean temperature (linear correlation coefficient values are 0.795, -0.003, 0.562, 0.845, 0.805, and 0.908 for the years 2019, 2020, 2021, 2022, 2023 and 2024 respectively) is seen in every case except for years 2020 and 2021, where COVID 19 pandemic related disruptions severely impacted economic activities as well as the power demand, in India and globally. This shows that while temperature effects influence power demand, there are other exogenous, systemic conditions that also impact it. The sensitivity of power demand to temperature is calculated as the slope of the average demand versus average temperature graphs for each year and it is seen that at the nationwide level on average, the slope values lie between 2.79-3.05 GW(e) per ^{o}C of temperature change, based on the data from years 2022-24.

Figures 3b and 3c also indicate the same general trend or correlation as well as slopes as shown in Figure 3a; however, the numerical values of the correlation coefficients are seen to be lower than that in the case of Figure 3a for most of the years. Therefore, it can be concluded that average temperature is a stronger driver of electricity demand in India than the maximum or minimum temperature.


2019
2020
2021
2022
2023
2024
Mean monthly power demand (GW(e))
Mean temperature (°C)
(a)
Maximum monthly power demand (GW(e))
Maximum temperature (°C)
(b)
Minimum monthly power demand (GW(e))
Minimum temperature (°C)
(c)

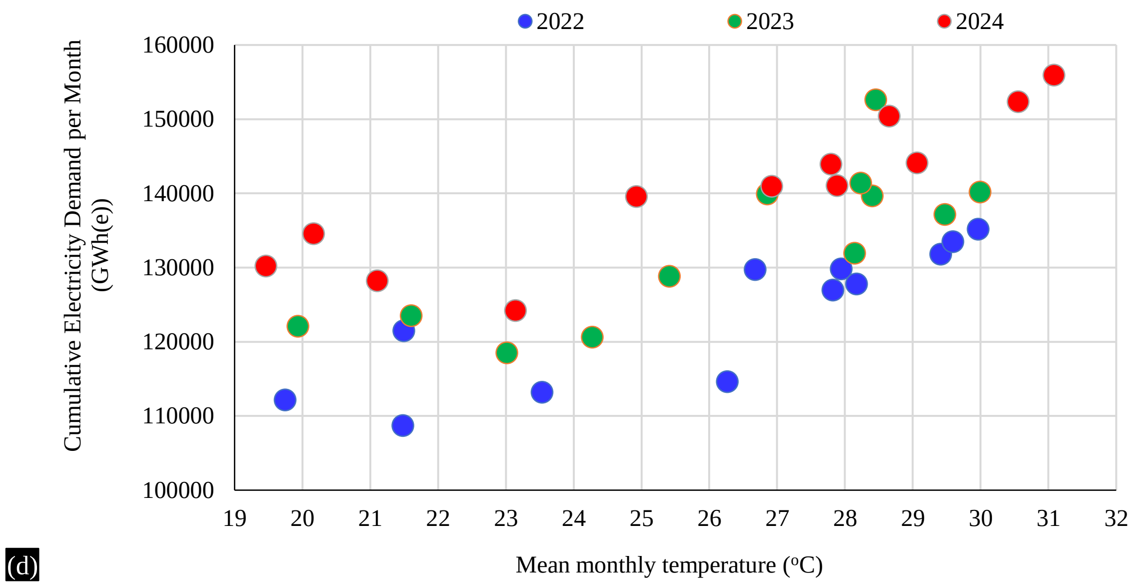


Figure 3. (a) Mean monthly power demand in India as function of mean monthly temperature using data from 2019 to 2024. (b) Maximum monthly power demand in India as function of maximum monthly temperature using data from 2019 to 2024. (c) Minimum monthly power demand in India as function of minimum monthly temperature using data from 2019 to 2024. (d) Cumulative electrical energy demand in India as function of mean temperature for the years 2022, 2023 and 2024.

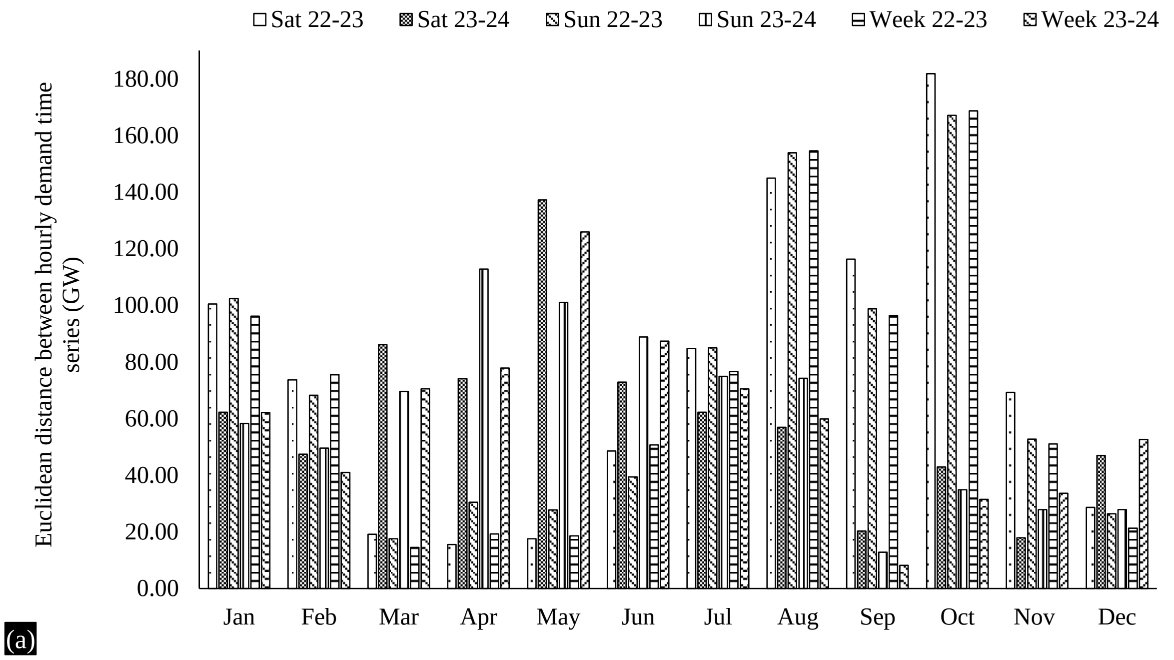
Figure 3d shows the strong positive correlation between total electric energy demand per month and monthly average temperature for the years 2022-2024 (correlation coefficient values being 0.861, 0.806 and 0.884 respectively for the 3 years under consideration in this study). The slope or sensitivity is positive and is calculated to be between 2142-2576 GWh(e)/month/ °C change of mean temperature, based on the demand data for these three years. This may be interpreted as a change or an increment in cumulative monthly demand by about 1.36-2.17% of the cumulative monthly energy demand today per degree Celsius rise in mean temperature. On an annual basis, this implies on average a rise in electricity demand by 25-31 TWh(e) over today's demand per degree rise in mean temperature.

Similarity Assessment of India's Electricity Demand Profiles for 2022-2024

Figure 4a shows the Euclidean distances between hourly power demand time series in different months of the years 2022-23 and 2023-24, separately for Saturdays, Sundays and weekdays, calculated using Eqn. 1. For example, the purple bars in Figure 3a shows the distance between Saturdays of January 2022 and Saturdays of January 2023. This metric is essentially a measure of the magnitude of the difference between two time series, with larger value indicating higher distance or difference between the data sets. There is no clear trend or monotonicity observed from this metric calculated across different months of the same year and in different years.

Figure 4b shows separation cosines of hourly power demand time series in different months of the years 2022-23 and 2023-24, separately for Saturdays, Sundays and weekdays, calculated using Eqn. 2. For example, the purple bars in Figure 4b shows the difference between Saturdays of January 2022 and Saturdays of January 2023.

This is essentially a dimensionless measure of the magnitude of the difference between two time series and its value lies between 0 and 1, as opposed to Euclidean distances which depend on the magnitudes of the variables in question. In this case also, there is no clear trend or monotonicity observed from this metric calculated across different months of the same year and in different years – the values are very close to unity in all cases.



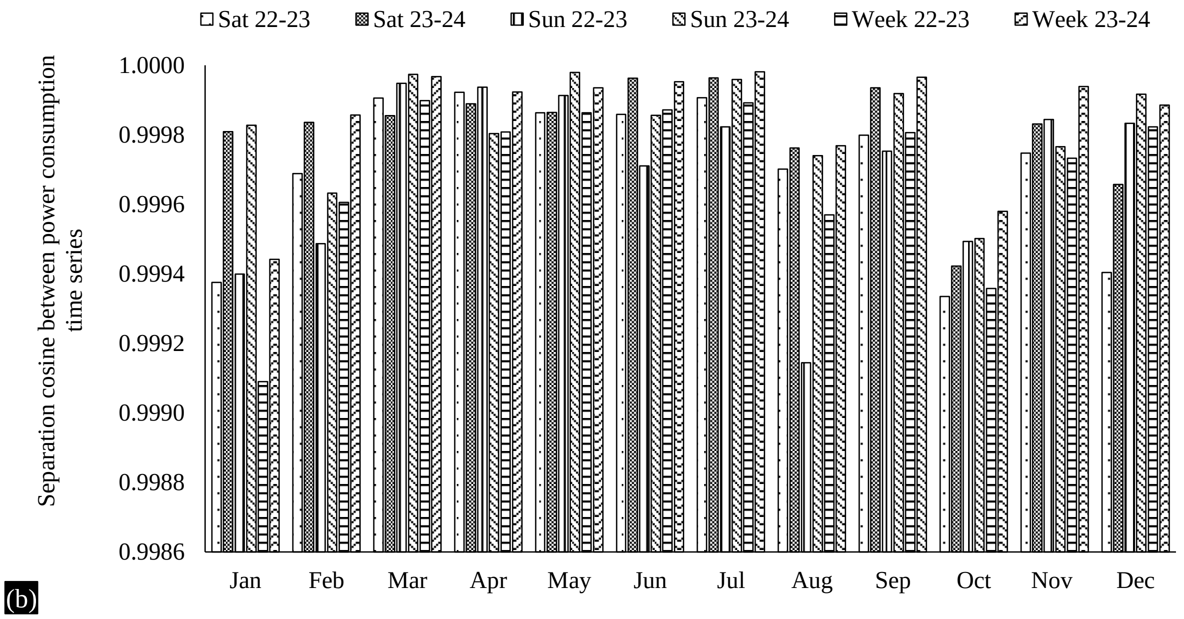


Figure 4. (a) Euclidean distances between hourly power demand profiles of India on Saturdays, Sundays and weekdays of each month in 2022-23 and 2023-24. (b) Separation cosines between hourly power demand profiles of India on Saturdays, Sundays and weekdays of each month in 2022-23 and 2023-24.

Figure 5a shows the minimum values of the ratios and Figure 5b shows the maximum values of the ratios of hourly power demand at the same time in different months of the years 2022-23 and 2023-24, separately for Saturdays, Sundays and weekdays. The ratios are above unity in most cases but they are also less than one in some of the cases. This shows that while overall power demand has increased year on year, at the same hour of each year, the demand may not show a strictly increasing trend due to various factors (e.g., temperature or other weather conditions, whether it is a public holiday or not and so on). This is a factor that makes high accuracy hourly forecasting of power demand extremely difficulty, using data from previous years.

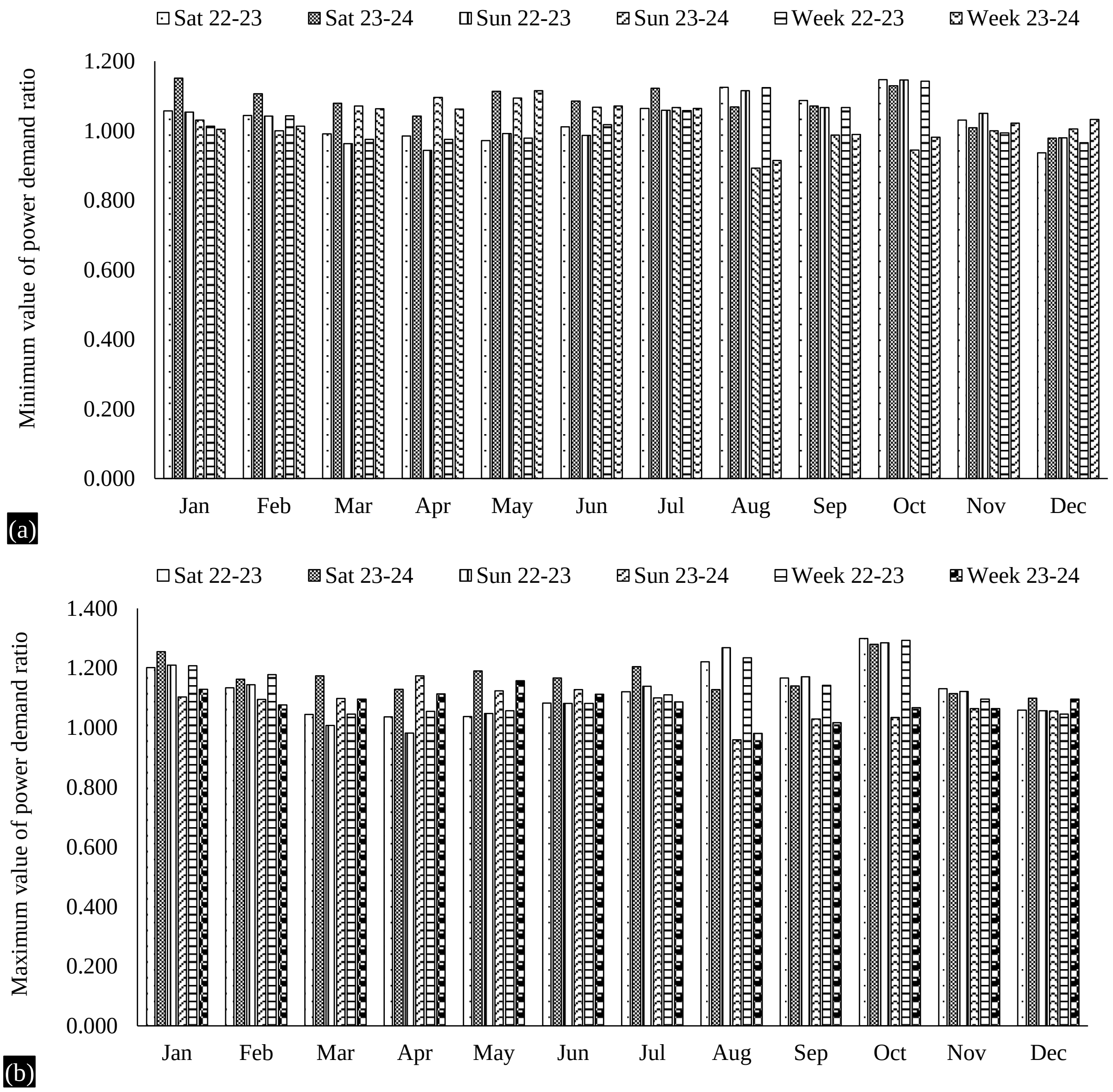


Figure 5. (a) Minimum ratio of power demand at same time of each month in different years.
(b) Maximum ratio of power demand at same time of each month in different years.

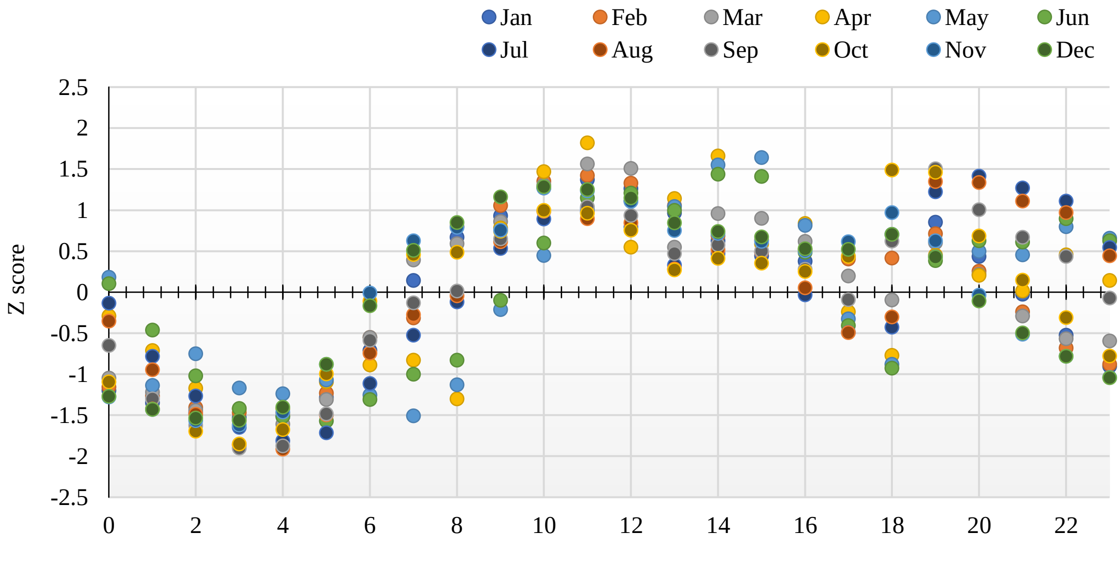
Anomaly Detection in Power Demand Patterns

To check for the presence of anomalies or outliers in the power demand data, the Z-score is calculated for the monthly average time series for each year. The Z-score for each element of a time series data set *X(t)* is generally defined as [12]

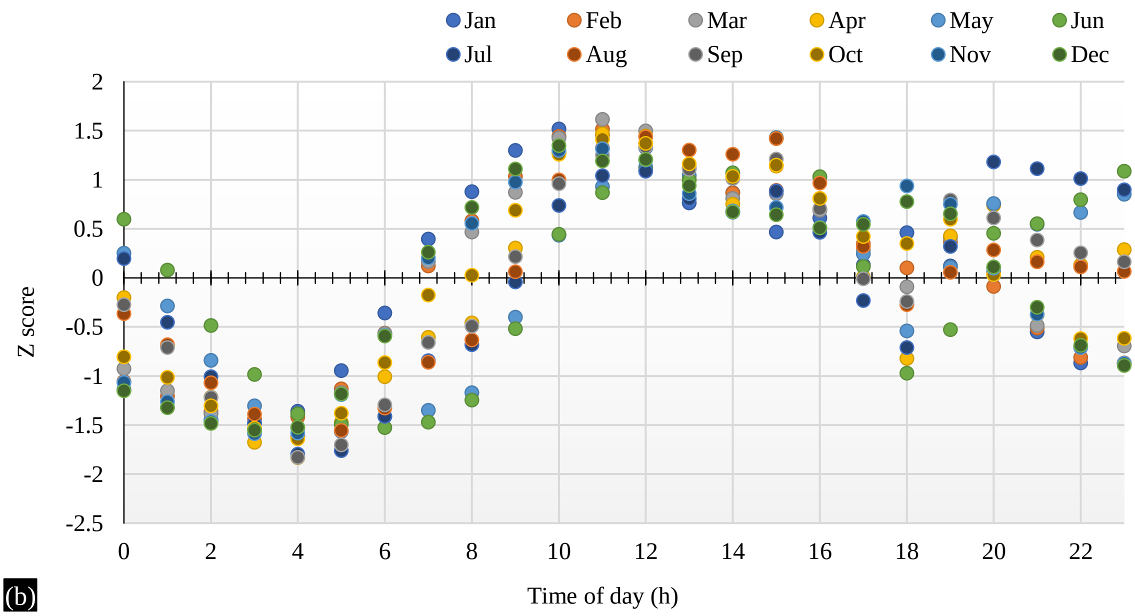
$$Z(t) = \frac{X(t)-\mu}{\sigma} \tag{7}$$

where *X(t)* represents power demand at time *t* in this study, μ represents the mean and σ represents the standard deviation of the values in the concerned time series data set. Therefore, this represents the demand data in non-dimensional and normalized form, on an identical scale for all the time series, permitting better understanding of their features and their intercomparison. These non-dimensional figures are important because power demand has increased every year and demand profiles have also shifted due to annual and seasonal factors affecting the environment and economy. Directly comparison of only the numerical values will not be adequate, but a comparison of the overall statistical nature of any demand curve is possible by using the Z-score.

The results are shown in Figures 6a, 6b and 6c for 2022, 2023 and 2024 respectively using the monthly average of the daily demand curves of each month. The values are seen to lie between -2 and +2 in all the cases. The Z-scores are always seen to be positive between 10 AM and 4 PM, and always negative between 2 AM and 6 AM, irrespective of the month, for all 3 years studied. This shows that power demand is typically higher than average and lower than average respectively, during these hours in every season. In the evenings, beyond 6 PM as well, certain months show above average demand, indicating the increasing need for space cooling after sunset as well. This may be attributed at least partially to the urban heat island effect, becoming stronger with higher levels of urbanization of the country and partially to the service nature of the economy enabled by digitalization, catering to different geographical regions of the world, with work hours spread out throughput the day. Highest positive and negative Z-scores represent daily peaks and troughs. As the calculated Z-scores do not go beyond the limits of ±3 in any of the cases, it may be concluded that monthly average power demand data in India from 2022-24 do not show outliers or anomalous values. Higher level of granularity in the data (e.g., at resolutions higher than one hour, such as 15-minute intervals, daily data instead of monthly average data) is likely to be needed to identify them.



(a)



(b)

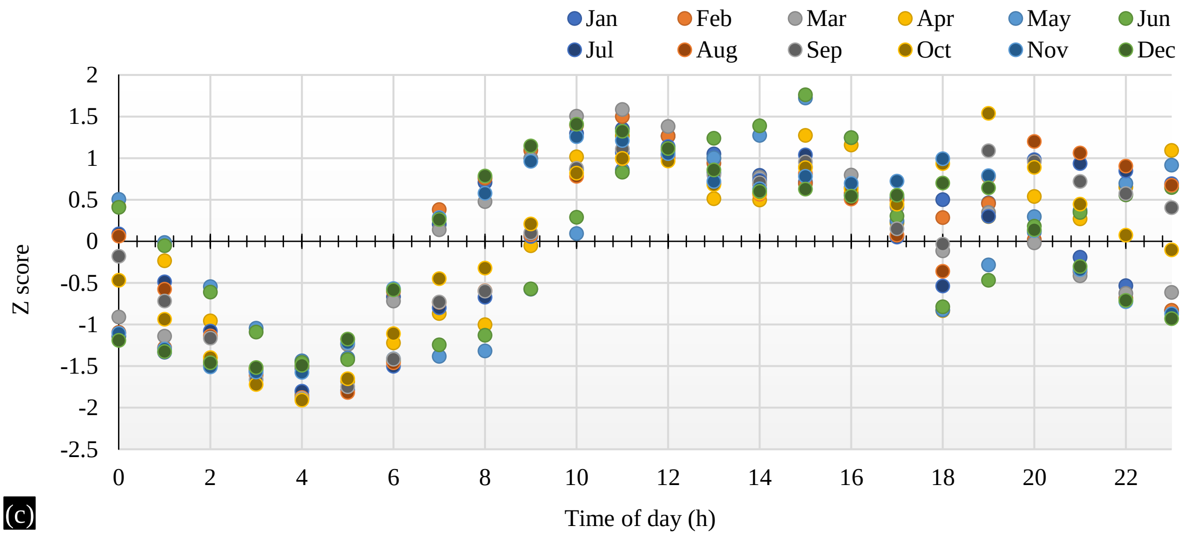


Figure 6. (a) Z-scores of power demand in India for each month in 2022. (b) Z-scores of power demand in India for each month in 2023. (c) Z-scores of power demand in India for each month in 2024.

The choice of the Z-score for anomaly detection may be justified as follows. The limits of ±3 on the values of the Z-score originate in the properties of the Z-distribution i.e., the standard normal distribution function, where 99.7% of all observation values are expected to lie between Z-values of ±3 [12]. Therefore, values lying outside this range are designated as extreme values on either end and their probability of occurrence is very, very low. Hence, those values are designated as outliers, with reference to other values in the series.

Uncovering Features from Time Series Analysis of Power Demand Data

Useful insights about the nature of time series data are provided by the calculated values of the autocorrelation functions, especially for the purpose of forecasting future values of the time series. As an example, the function values have been evaluated for 20 lags, using Eqn. 6 for the power demand data of January 2022 for the Saturday, Sunday and weekday power demand time series data. The corresponding results are shown in Figure 7a. Figure 7a shows that even up to the 4^{th} time lag, the autocorrelation coefficients are significantly different from zero (as the value lies above the horizontal blue line in the Figure, which marks the 95% confidence interval for the coefficient based on the sample drawn). Thus, it is concluded that serial correlation is present in the power demand patterns for India over successive time intervals. Due to the geometric nature of the decline of these values at successive lags, it may be concluded that the time series data show auto-regressive nature.

Figure 7b shows the calculated partial autocorrelation functions for the same month. At higher lags also (i.e., about 10 or 11), the *PACF* values tend to exceed the 95% confidence interval, showing their statistically significant nature. Fig. 7c shows the cross-correlation coefficients for the Saturday-Sunday, Sunday-weekday and weekday-Saturday pairs of time series data for January 2022. Here also the *CCF* values become statistically insignificant after about the fourth lag. Similar plots can be generated for each of the months in the 3 years under study in this work.

For the purpose of comparison, the autocorrelation, partial autocorrelation and cross correlation functions for January 2023 and January 2024 are also plotted in Figures 8a, 8b and 8c and Figures 9a, 9b and 9c respectively. The differences (e.g., in the number and distribution of statistically significant lags in the autocorrelation or partial autocorrelation functions) show that it is not possible to have a single kind of time series model for the entire year or for even the same months in different years for demand forecasting applications. This explains the complicated nature of the forecasting problem; only time series data may not be adequate for such an exercise and other explanatory variables will be required.

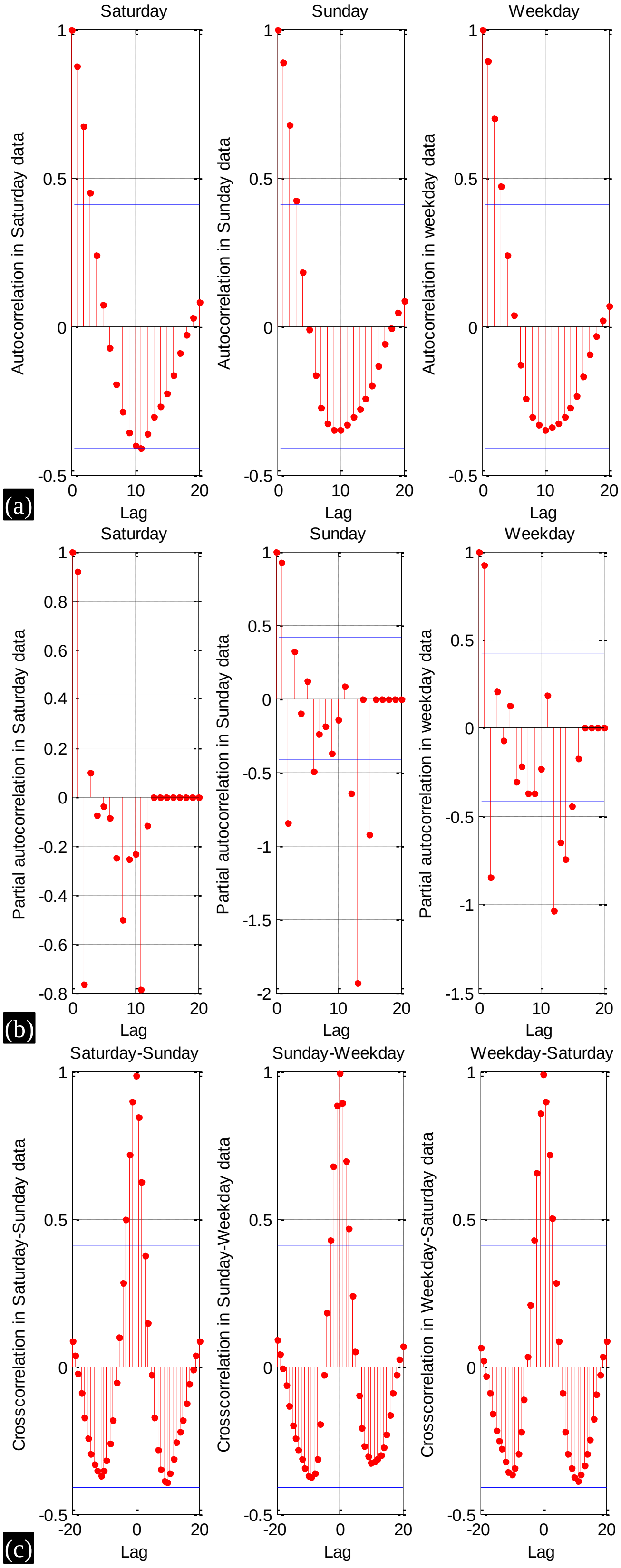


Figure 7. (a) Autocorrelation coefficients of daily power demand data at different time lags (data for January 2022). (b) Partial autocorrelation coefficients of daily power demand data at different time lags (data for January 2022). (c) Cross correlation coefficients of pair-wise daily power demand data at different time lags (data for January 2022).

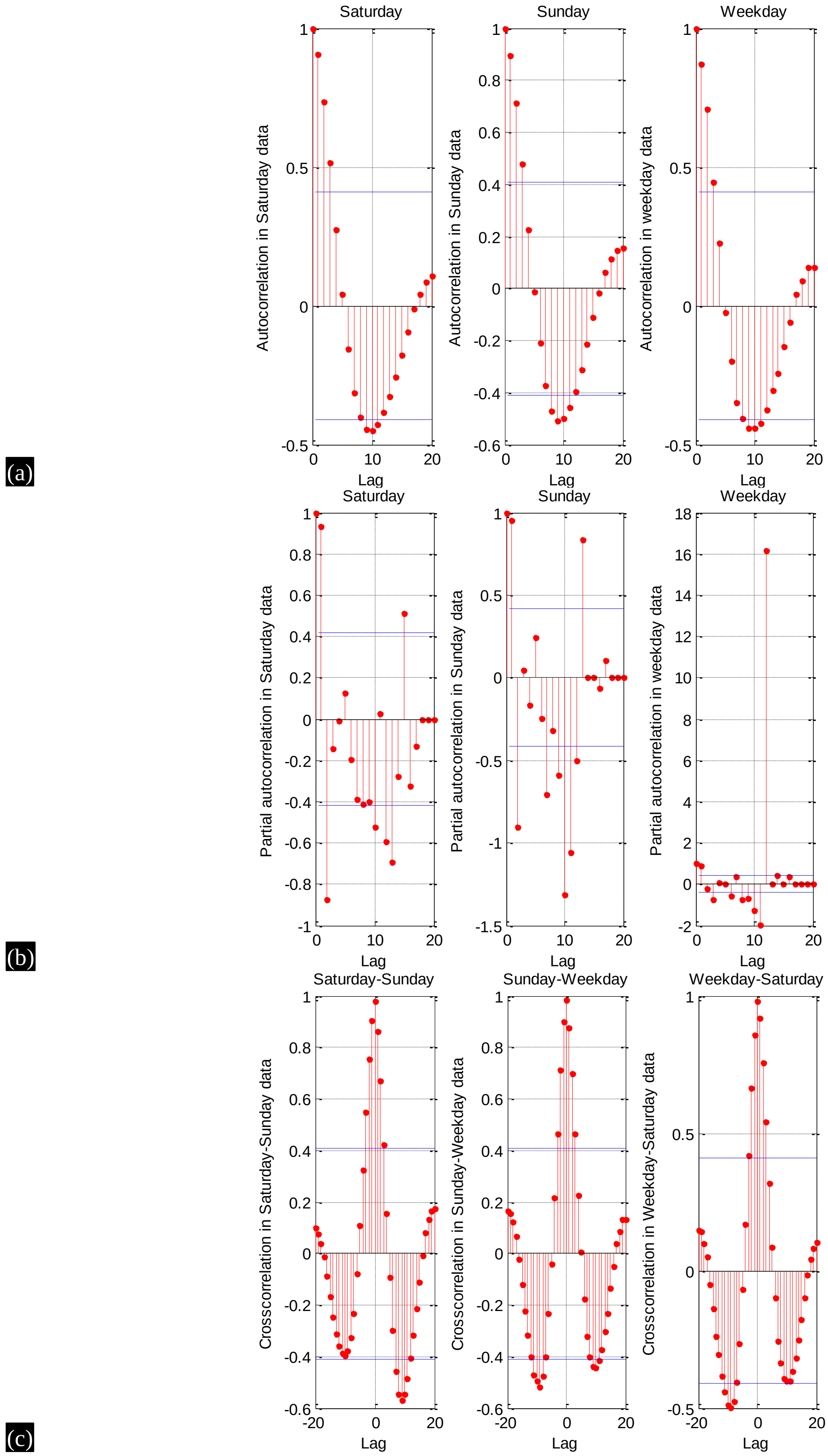


Figure 8. (a) Autocorrelation coefficients of daily power demand data at different time lags (data for January 2023). (b) Partial autocorrelation coefficients of daily power demand data at different time lags (data for January 2023). (c) Cross correlation coefficients of pair-wise daily power demand data at different time lags (data for January 2023).

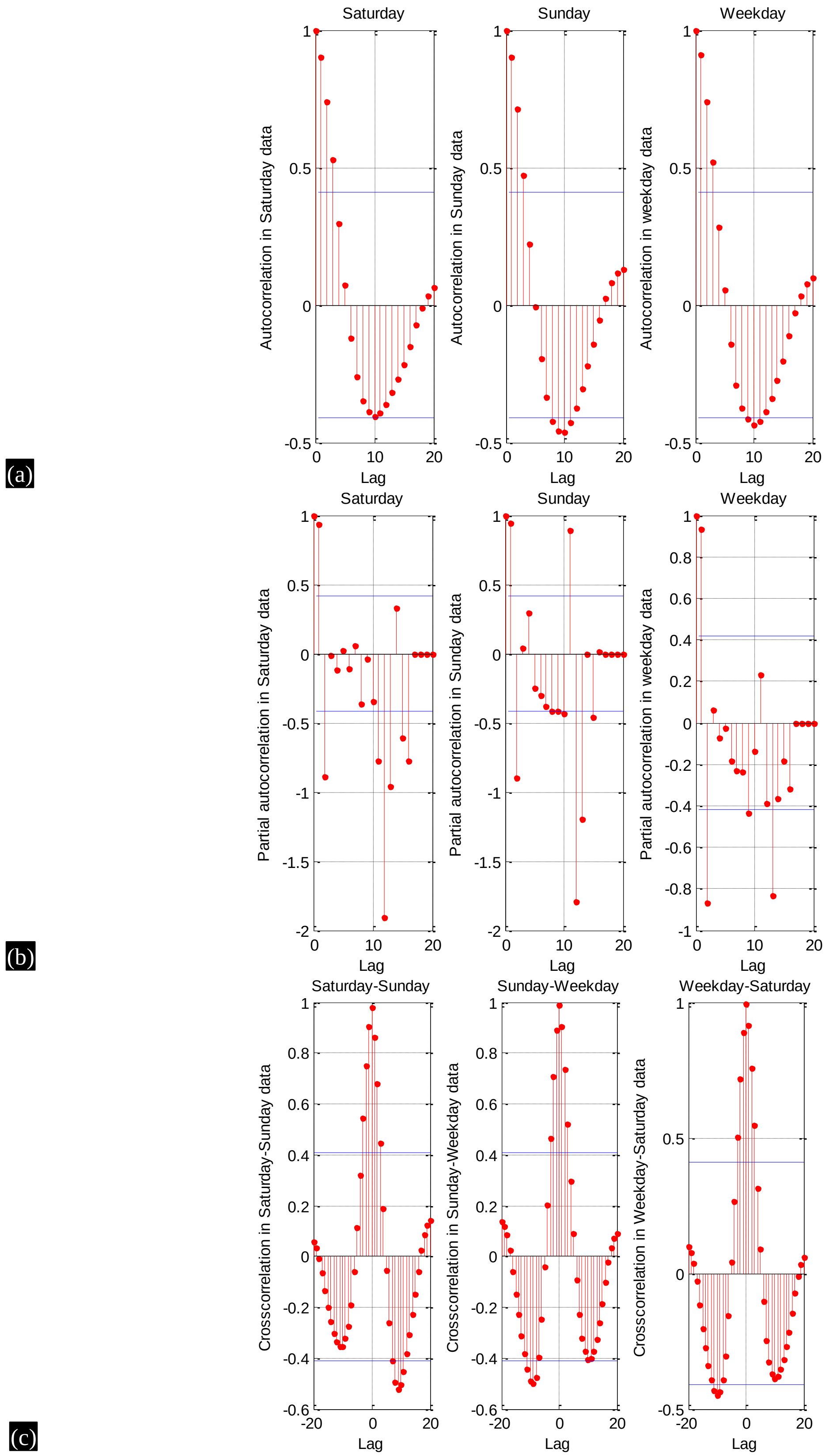


Figure 9. (a) Autocorrelation coefficients of daily power demand data at different time lags (data for January 2024). (b) Partial autocorrelation coefficients of daily power demand data at different time lags (data for January 2024). (c) Cross correlation coefficients of pair-wise daily power demand data at different time lags (data for January 2024).

CONCLUSION

In this study, exploratory statistical data analyses have been performed to understand the trends and features of electric power demand evolution in India over the last 3 years (2022, 2023 and 2024) using publicly available datasets. Descriptive statistics calculations have been performed using the monthly power demand data, available at hourly resolution, to understand average properties, maximum and minimum values, share of annual power consumption consumed in each month of the year. These assessments have enabled quantification of the seasonal and diurnal variation of demand within in a year and across years, the spread and the correlations between demand across different months of the year, identification of any anomalous power demand situation, as well as the temperature dependence of power demand. The evolution of demand curves over time has also been studied using last 3 years' data and the extent of their similarities have been estimated on a month-to-month basis using separation cosines and Euclidean distances.

From the energy planning and policy formulation point of view, these analyses are useful inputs towards the following objectives

1. Developing models for electric power demand forecasting and for implementing demand response measures in the electricity sector, particularly in the context of day ahead and term ahead electricity markets design, for which the power supply profiles from the installed generation mix are also crucial inputs.
2. Planning near term energy capacity addition based on observed rate of power demand increase and deciding on safety margins needed for unexpected demand increase
3. Correlation of weather conditions and electricity demand to implement season specific measures for using an generation mix of electric power generation
4. Daily planning of operation of installed capacity based on known patterns from the Z-scores of patterns of previous years
5. Understanding surplus power availability profiles considering the demand-supply difference at different parts of the day and determining optimal ways to use the surplus instead of curtailing

REFERENCES

1. Arto I, Capellan-Perez I, Lago R, Bueno G, Bermojo R (2016). The energy requirements of a developed world. Energy for Sustainable Development, 33:1-13.
2. MoSPI (2024). Energy Statistics India 2024. National Statistical Office, Government of India.
3. IEA (2020). Key World Energy Statistics 2020, IEA, Paris https://www.iea.org/reports/key-world-energy-statistics-2020 (Last accessed on 13.4.2024)
4. Idhom M, Fauzi A, Trimono T, Riyantoko P (2023). Time Series Regression: Prediction of Electricity Consumption Based on Number of Consumers at National Electricity Supply Company. TEM Journal, 12(3):1575-1581.
5. Lee YW, Tay KG, Choy YY (2018). Forecasting Electricity Consumption Using Time Series Model. International Journal of Engineering and Technology, 7(4.30):218-223.
6. Kim H, Park S, Kim S (2023). Time-series clustering and forecasting household electricity demand using smart meter data. Energy Reports, 9:4111-4121.
7. Kouakou AK (2011). Economic growth and electricity consumption in Cote d'Ivoire: Evidence from time series analysis. Energy Policy, 39(6):3638-3644.
8. Yu K, Cao J, Chen X, Yang Z, Gan L (2023). Residential load forecasting based on electricity consumption pattern clustering. Frontiers in Energy Research, 10:1113733.
9. Choksi K, Kowli A (2023). New Perspectives on Clustering for Demand Response. Energy Informatics: Energy Informatics: Third Energy Informatics Academy Conference, EI.A 2023, Campinas, Brazil, December 6–8, 2023, Proceedings, Part II, Pages 175 – 191.
10. Praveen P, Balachandra P, Pandarasamy A (2024). Enhanced Consumer Segmentation Through Load Profile Analysis Using Autoencoder and K-Shape Clustering. In: Jørgensen, B.N., Ma, Z.G., Wijaya, F.D., Irnawan, R., Sarjiya, S. (eds) Energy Informatics. EI.A 2024. Lecture Notes in Computer Science, vol 15271. Springer, Cham.

11. Assfalg J, Kriegel H, Kroger P, Kunath P, Pryakhin A, Renz M (2006). Time Series Analysis Using the Concept of Adaptable Threshold Similarity. Proceedings of 18th International Conference on Scientific and Statistical Database Management (SSDBM'06), Vienna, Austria, 2006. Available at https://www.dbs.ifi.lmu.de/Publikationen/Papers/ssdbm06.threshold.pdf (last accessed on 8.11.2024)
12. Navidi W (2011). Statistics for Engineers and Scientists, 3rd ed. The McGraw-Hill Companies Inc, New York, USA.
13. Montgomery DC, Jennings CL, Kulahci M (2015). Introduction to Time Series Analysis and Forecasting, 2nd ed., Wiley International, Hoboken, New Jersey, USA.

