World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS OF DAILY BIKE-SHARING DEMAND CURVES IN SEOUL: ASSOCIATIONS WITH SEASON, WEEKENDS AND HOLIDAYS

Mohammed Ashrafur Rahman and Md. Rabiul Haque *

Department of Mathematics, Faculty of Science, University of Rajshahi, Rajshahi-6205, Bangladesh.
* Corresponding Author

ORCID Details
Mohammed Ashrafur Rahman: https://orcid.org/0009-0003-1921-448X
Md. Rabiul Haque: https://orcid.org/0000-0002-8883-1831

World Journal of Advanced Research and Reviews, 2026, 32(01), 1809–1818

Publication history: Received on 12 September 2026; revised on 24 October 2026; accepted on 26 October 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.32.1.2735

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Bike-sharing demand is usually modeled as hourly counts, which hides the shape of the daily profile. We aimed to identify which daily-curve components differ by season, holiday and weekend, and whether a holiday effect remains after adjustment for weekend, temperature and rain. We analyzed 8,760 hourly Seoul rental records (1 December 2017 to 30 November 2018); removing 13 outage days left 352 daily curves. Curves were smoothed with a Fourier basis and summarized by functional principal component analysis (FPCA). Scores on the first three components (PC1 to PC3) were compared with rank-based tests and adjusted linear models with heteroscedasticity-consistent standard errors. Sensitivity analyses included a corrected holiday calendar and log(1 + count) curves. Three components carried 95.5% of the variance (PC1 82.7%). PC1 measured overall level, differed by season ($\epsilon^2 = 0.509$) and rose with temperature (Spearman $\rho = 0.72$). PC2 contrasted commute peaks with an afternoon hump and separated weekends from weekdays (rank-biserial correlation 0.89). After adjustment, holidays were associated with a shift towards the weekend shape on PC2 (+698.5 score units, p = 0.0011; +920.7 with the corrected holiday flag). This was not detected on log(1 + count) curves, so it describes absolute and not proportional curve shape. The holiday effect on level was not established under the corrected flag. In one city over one year, demand level was associated with season and temperature, and demand shape with weekends and holidays. These are associations, not causal effects.

Keywords: Functional principal component analysis; Functional data analysis; Bike-sharing demand; Daily demand curves; Holiday and weekend effects.

1 INTRODUCTION

Bike-sharing demand depends on weather, season and the calendar. A review of recent literature reported that use is higher in warm months, that weekday peaks fall at about 7 to 9 am and 4 to 6 pm, and that weekends show a midday peak [1]. In Washington, DC, cold, rain and high humidity reduced both the chance of a trip and its length [2]. In Oslo, temperature raised use in every trip group. This effect was largest on workday afternoons and evenings, rain lowered

use [3]. In Daejeon, weather and non-working-day effects differed by hour of day, and use fell above 30 °C [4]. Weather and calendar effects have also been examined in Brisbane [5]. The factors affecting demand have been reviewed [6].

Studies that use the Seoul data mostly forecast demand and report accuracy, not interpretable daily patterns. Sathishkumar, Park and Cho predicted Seoul trip duration from trip records and weather, and random forest performed best [7]. Sathishkumar and Cho modeled hourly demand in Seoul with random forest and other models, and variable importance changed by season [8]. Leem et al. produced online forecasts on this dataset, from one hour to one day ahead [9]. Gao and Chen studied Seoul rentals in 2020, a different data period, and random forest and support vector machines performed best [10].

Functional data analysis treats each day as one curve. A daily cycle is a natural unit for this. Movements in Lyon were mostly cyclostationary over the week [11]. Functional principal component analysis (FPCA) is a core technique of functional data analysis [12,13]. It summarizes variation among curves with a few components that can be read as shapes. Bouveyron et al. smoothed hourly station loading curves with a Fourier basis and clustered them [14]. Piter et al. treated the number of bikes at each Helsinki station as a function of time of day in a spatiotemporal model, and weekend patterns shifted towards the afternoon [15]. Among the studies we read, none used FPCA scores of daily system-wide curves as the response.

Our objective was to identify which daily-curve components differ by season, holiday and weekend. We also asked whether a holiday effect remains after adjusting for weekend, temperature and rain. Section 2 describes the data and methods. Section 3 reports the results; Section 4 discusses the results and limitations. Section 5 concludes.

2 METHODS

2.1 Data

The data were from the Seoul Bike Sharing Demand file from the UCI Machine Learning Repository [16]. The file holds 8,760 hourly records from 1 December 2017 to 30 November 2018, covering 365 days. Each record gives the hourly rental count, eight weather variables, hour, season, a holiday flag and a Functioning Day flag (Table 1).

Table 1: Numeric variables in the Seoul Bike Sharing Demand file (n = 8,760 hourly records). Min, mean and max are over all hourly records.

Variable	Unit	Min	Mean	Max	Audit note
Rented bike count	bikes per hour	0	705	3,556	295 zero counts, all during outages
Temperature	°C	−17.8	12.88	39.4	
Humidity	%	0	58.23	98	17 zeros, treated as invalid
Wind speed	m/s	0	1.72	7.4	
Visibility	10 m	27	1,437	2,000	Capped at 2,000 in 25.6% of rows
Dew point temperature	°C	−30.6	4.07	27.2	
Solar radiation	MJ/m^2	0	0.57	3.5	
Rainfall	mm	0	0.15	35	Non-zero in 528 rows
Snowfall	cm	0	0.08	8.8	Non-zero in 443 rows

No values were missing. All 295 zero counts fell in hours with Functioning Day = No. These comprised 12 complete days and 7 hours on 6 October 2018. We treated them as outages, not as zero demand, and excluded the 13 affected days, leaving 352 usable days (Figure 1). The audit also found 17 humidity values recorded as zero and visibility capped at 2,000 (10 m units) in 25.6% of rows. Neither variable entered the adjusted models. Hourly temperature and dew point were strongly collinear (variance inflation factors 87 and 116), and daily mean temperature was strongly related to season ($R^2 = 0.80$).

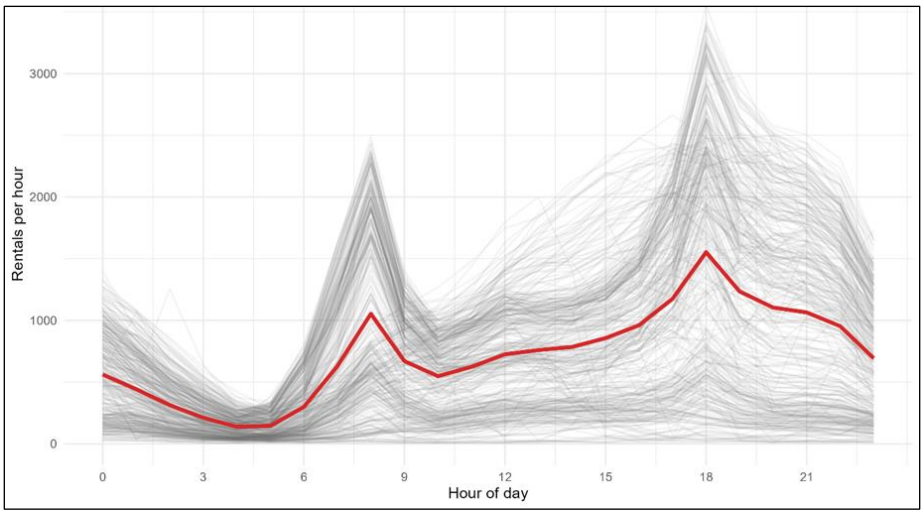

Figure 1: Daily rental curves for the 352 usable days (grey) and the mean curve (red).

We checked the dataset's Holiday flag against published 2018 South Korean public-holiday calendars. The flag marked 18 days as holidays, 17 of them on usable days (13 weekdays and 4 weekend days). It omitted 5 May, 7 May, 13 June and 26 September 2018 and flagged 1 May. It also flagged 22, 24 and 31 December 2017, which could not be verified against the calendar used; these were kept as flagged. The primary analysis used the dataset's own flag. A corrected flag (1 May removed and the four omitted dates added; 20 holiday days, 15 on weekdays) served as a sensitivity analysis (Section 2.6).

2.2 Curves and smoothing

Each usable day was treated as one curve, with the 24 hourly counts (hours 0 to 23) as its observations. Counts were smoothed by least squares on a periodic Fourier basis of 11 functions (a constant and 5 sine-cosine pairs) over t in [0, 24), so that 23:00 joins 00:00 smoothly [12] (Equation 1). The fit gave $R^2 = 0.974$.

$$x_i(t) = \sum_{(k=1)}^{11} c_{(ik)} \cdot \phi_k(t) \tag{1}$$

(where ϕ is a periodic Fourier basis with 5 harmonics on $t \in [0,24)$)

$$c_i = (\Phi'\Phi)^{(-1)}\Phi' y_i$$

(with y_i the 24 hourly counts of day i)

2.3 Functional principal component analysis

Functional principal component analysis (FPCA) decomposed the smoothed curves around the mean curve $\mu(t)$ (Equation 2) [12,13,17]. The eigenfunctions ψ_j and eigenvalues λ_j came from the sample covariance of the curves. The score $\xi_{(ij)}$ measured how far day i lay along component j, and PVE_j is the proportion of variance carried by component j. A component was displayed as $\mu(t) \pm 2 \cdot sqrt(\lambda_j) \cdot \psi_j(t)$. We retained the first three components, which together carried 95.5% of the variance.

$$x_i(t) = \mu(t) + \sum_j \xi_{(ij)} \cdot \psi_j(t)$$
$$\xi_{(ij)} = \int (x_i(t) - \mu(t)) \cdot \psi_j(t)\, dt \tag{2}$$
$$(PVE)_j = \frac{(\lambda_j)}{(\sum \lambda)}$$

2.4 Group comparisons

Daily totals were non-normal (Shapiro-Wilk W = 0.925) and their variance differed by season (Levene F = 37.7), so we used rank-based tests. Scores were compared across the four seasons with the Kruskal-Wallis test, followed by Dunn

tests with Holm adjustment. Holiday and weekend differences used the Mann-Whitney test. Weekend was Saturday or Sunday, derived from the date. Holiday was also compared with other weekdays alone (13 holiday weekdays against 238 other weekdays) to separate holiday from weekend effects. Effect sizes were epsilon-squared for Kruskal-Wallis and the rank-biserial correlation for Mann-Whitney (Equation 3); the correlation is positive when holiday or weekend had higher ranks. Daily mean temperature was related to PC1 by Spearman correlation, overall and within season. P values were Holm-adjusted across the three components where stated.

$$\epsilon^2 = \frac{(H)}{(n-1)} \quad \text{(for Kruskal-Wallis)}$$
$$r = 1 - \frac{(2U)}{(n_1 \cdot n_2)} \quad \text{(for Mann-Whitney)} \qquad (3)$$

2.5 Adjusted models

Each score was modeled by ordinary least squares (Equation 4), fitted separately for components 1 to 3. Winter was the reference season. TempC was the daily mean temperature centered within season, which separates the within-season temperature association from the season contrast. Rain was the daily total rainfall in mm. Standard errors were heteroscedasticity-consistent (HC3). Because daily PC1 had a lag-1 autocorrelation of 0.744, Newey-West standard errors with lag 7 were computed as a check. Log daily total was modeled with the same covariates as a baseline for comparison with the scores.

$$\xi_{(ij)} = b_0 + (Season)_i + b_1 \cdot (Holiday)_i + b_2 \cdot (Weekend)_i + b_3 \cdot (TempC)_i$$
$$+b_4 \cdot (Rain)_i + e_i \qquad (4)$$

(with HC3 standard errors and Newey-West (lag 7) as a check; fitted for $j = 1,2,3$)

2.6 Robustness checks and software

The holiday effect on PC2 was re-estimated with 3 and 8 harmonics, with unsmoothed hourly counts, with all 365 days, after dropping the four weekend holidays, with the corrected holiday flag, and with log(1 + count) curves, which tests whether the holiday result holds for proportional curve shape.

Analyses were run in R 4.5.2 (robust standard errors with the sandwich and lmtest packages).

3 RESULTS

3.1 Curve shape and variance

Daily totals averaged 17,490 rentals (SD 9,951) over the 352 usable days and differed most by season: winter 5,413, spring 17,910, summer 24,820 and autumn 22,170 (Table 2). Holidays (n = 17) averaged 12,700 against 17,730 on other days, and weekdays 17,950 against 16,330 on weekends.

Table 2: Daily rentals by season, holiday status and weekend status (352 usable days). Values are rentals per day, rounded to four significant digits.

Group	Days	Mean	SD	Median
All days	352	17,490	9,951	18,560
Winter	90	5,413	1,808	5,498
Spring	90	17,910	8,357	17,590
Summer	92	24,820	7,297	25,570
Autumn	80	22,170	6,718	23,410
Holiday	17	12,700	10,500	7,184
Non-holiday	335	17,730	9,877	19,130
Weekday	251	17,950	9,960	19,880
Weekend	101	16,330	9,883	17,390

Hourly profiles differed more in shape than in total (Figure 2). The overall mean curve peaked at 18:00 (1,554 rentals per hour). Non-holiday weekdays peaked at 08:00 (1,354) and 18:00 (1,766), weekend days showed one broad afternoon-to-evening hump, and holidays peaked once, at 17:00 (879). Every season peaked at 18:00, and winter was lowest at all hours.

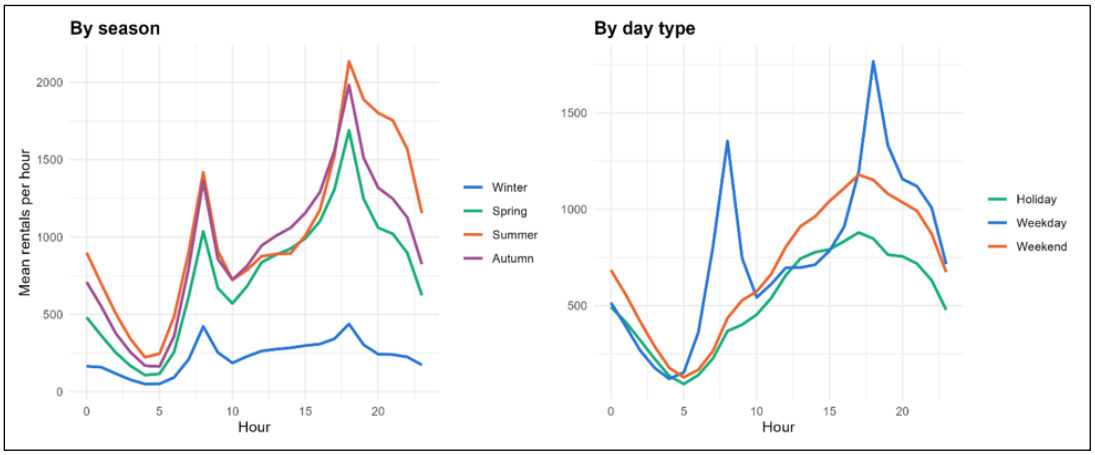

Figure 2: Mean hourly rentals by season (left) and by day type (right), 352 usable days.

The first three components carried 95.54% of the variance: PC1 82.66%, PC2 8.42% and PC3 4.47% (Table 3). The PC1 share changed little with the smoothing choice (81.35% to 84.28%).

Table 3: Percentage of variance explained by the first five components (352 daily curves). Variance and cumulative variance are for the main analysis (5 harmonics); the last three columns repeat the analysis with 3 harmonics, 8 harmonics and no smoothing.

Component	Variance (%)	Cumulative (%)	3 harmonics (%)	8 harmonics (%)	No smoothing (%)
PC1	82.66	82.66	84.28	81.79	81.35
PC2	8.42	91.07	7.66	8.70	8.86
PC3	4.47	95.54	4.29	4.51	4.52
PC4	2.66	98.20	2.57	2.70	2.74
PC5	0.85	99.05	0.82	0.87	0.88

PC1 had positive weights at every hour, largest from 17:00 to 20:00, so it measured overall level with extra weight on the evening (Figure 3). PC2 had negative weights from 06:00 to 09:00 and from 18:00 onwards and positive weights from 10:00 to 17:00, peaking at 15:00; high scores therefore marked one afternoon hump and low scores marked commute peaks. PC3 had positive weights from 06:00 to 17:00 (peak at 08:00) and negative weights at night and from 18:00, contrasting morning and midday with evening. Adding or subtracting two standard deviations confirmed these readings (Figure 4): the PC1 curve at +2 SD exceeded 3,000 rentals per hour at 18:00, PC2 exchanged twin peaks for a single afternoon hump, and PC3 moved demand from evening to morning and midday.

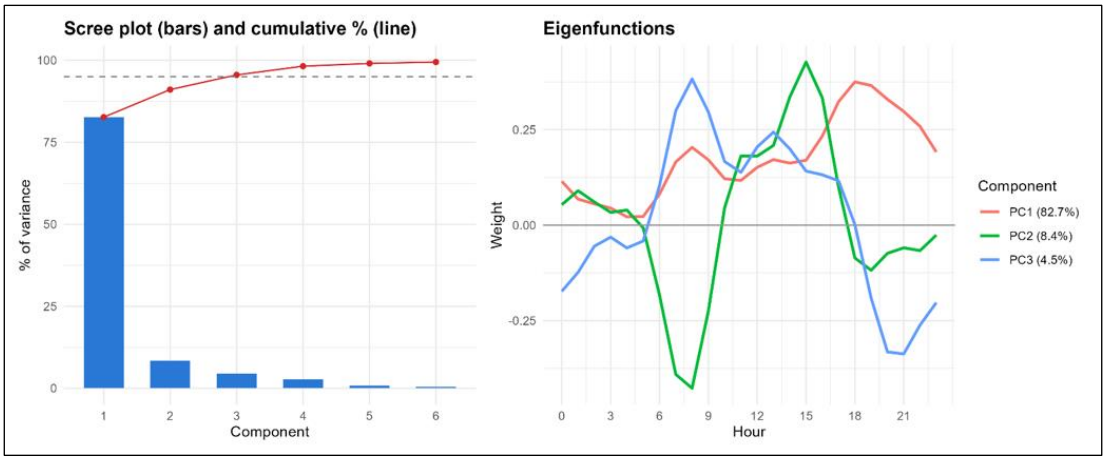

Figure 3: Scree plot (bars, percent of variance; line, cumulative percent) and the first three eigenfunctions against hour of day (weights), n = 352 daily curves.

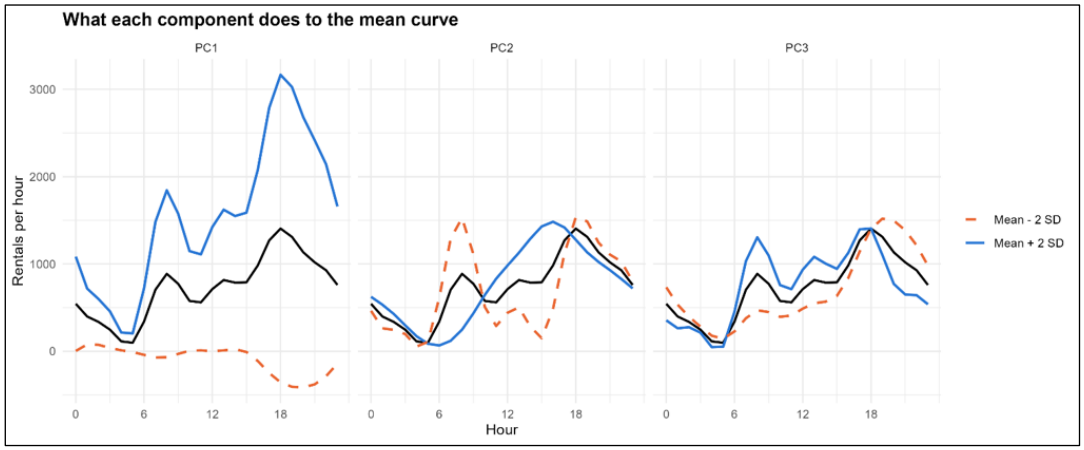

Figure 4: Mean curve plus and minus two standard deviations of each component, $\mu(t) \pm 2 \cdot \sqrt{(\lambda_j)} \cdot \psi_j(t)$, n = 352 daily curves.

3.2 Group differences

Day type separated the days mainly on PC2 (Table 4). Weekend days scored higher than weekdays (rank-biserial r = 0.89, p = 6.5e-39), with smaller differences on PC3 (r = −0.30, p = 1.1e-5) and PC1 (r = −0.15, p = 0.03). Holidays scored higher than other days on PC2 (r = 0.57, p = 6.4e-5; Holm p = 1.9e-4), and the contrast persisted on weekdays alone (13 holiday weekdays against 238 other weekdays: r = 0.91, p = 3.4e-8). Holidays scored lower on PC1 (r = −0.34, p = 0.017; Holm p = 0.034), but on weekdays alone this was weaker (r = −0.28, p = 0.090). In the PC1-PC2 plane, most holidays lay among weekend days (Figure 5).

Table 4: Effect sizes (unadjusted p in brackets) for comparisons of FPCA scores (352 days). ε^2, epsilon-squared (Kruskal-Wallis); r, rank-biserial correlation (Mann-Whitney), positive when the first-named group scored higher; ρ, Spearman correlation. Holm-adjusted p values are given in the text.

Comparison (effect size)	PC1	PC2	PC3
Season (ε^2)	0.509 (<0.001)	0.083 (<0.001)	0.327 (<0.001)
Weekend vs weekday (r)	−0.15 (0.030)	0.89 (<0.001)	−0.30 (<0.001)
Holiday vs other days (r)	−0.34 (0.017)	0.57 (<0.001)	−0.27 (0.061)
Holiday vs other weekdays (r)	−0.28 (0.090)	0.91 (<0.001)	−0.28 (0.086)
Temperature, all days (ρ)	0.72 (<0.001)	−0.20 (<0.001)	−0.30 (<0.001)
Temperature, winter (ρ)	0.39 (<0.001)	0.35 (<0.001)	0.23 (0.032)

Temperature, spring (ρ)	0.63 (<0.001)	−0.05 (0.615)	−0.16 (0.143)
Temperature, summer (ρ)	−0.38 (<0.001)	−0.41 (<0.001)	−0.39 (<0.001)
Temperature, autumn (ρ)	0.64 (<0.001)	0.02 (0.837)	−0.37 (<0.001)

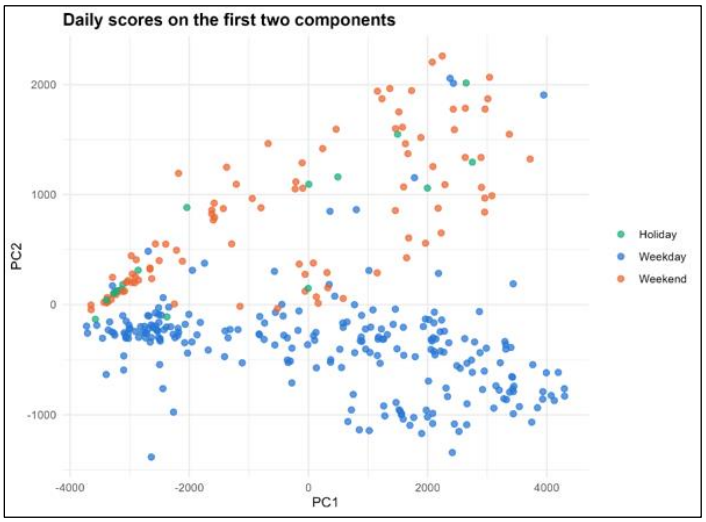

Figure 5: Daily scores on PC1 against PC2 by day type (weekday, weekend, holiday), n = 352 days.

Season differed on all three components. The effect was largest for PC1 (Kruskal-Wallis H = 178.8, ϵ^2 = 0.509, p = 1.6e-38), followed by PC3 (ϵ^2 = 0.327, p = 1.1e-24) and PC2 (ϵ^2 = 0.083, p = 2.0e-6). On PC1, winter differed from every other season (Dunn z = −8.2 to −12.5, Holm p < 1e-15), and spring differed from summer (z = −4.2, Holm p = 7.9e-5); spring against autumn (z = −2.2, Holm p = 0.053) and summer against autumn (z = 1.8, Holm p = 0.065) did not. PC1 rose with daily mean temperature (Spearman ρ = 0.72, p = 6.3e-57). Within season, ρ was 0.64 in autumn, 0.63 in spring and 0.39 in winter, but −0.38 in summer (p = 1.9e-4) (Figure 6).

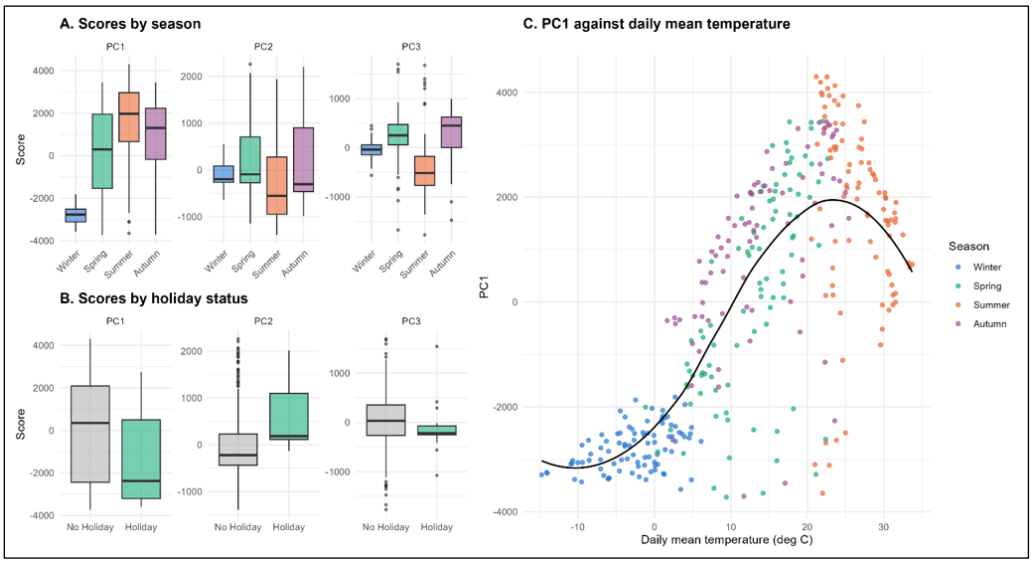

Figure 6: (A) Scores by season and (B) by holiday status for PC1 to PC3; (C) PC1 against daily mean temperature with a loess curve. n = 352 days.

3.3 Adjusted models and robustness

After adjustment (Equation 4), holidays remained associated with PC2 (Table 5): +698.5 score units (HC3 SE 211.5, p = 0.0011; Newey-West p = 0.0017), against +1,105 for weekends (SE 70.1, p = 1.7e-42; R^2 = 0.549). Temperature (−4.6 per °C, p = 0.37) and rain were not associated with PC2. For PC1 (R^2 = 0.756), the holiday coefficient was −746.9 (SE 292.6, p = 0.011; Newey-West p = 0.008), weekend −703.9 (SE 142.2, p = 1.2e-6), temperature +114.5 per °C (SE 12.1,

p = 5.2e-19) and rain −75.2 per mm (SE 8.2, p = 4.1e-18). For PC3 (R^2 = 0.339), weekend was −289 (SE 51.8, p = 4.8e-8) and holiday −148 (SE 121, p = 0.22).

Table 5: Adjusted linear models for the FPCA scores (352 days). Estimates are in score units with heteroscedasticity-consistent (HC3) standard errors. Temperature is the daily mean centered within season; p values are unadjusted across components. Intercepts are omitted.

Term	PC1 estimate (SE)	PC1 p	PC2 estimate (SE)	PC2 p	PC3 estimate (SE)	PC3 p
Spring (vs winter)	3,166 (149.3)	<0.001	363.9 (67.3)	<0.001	255.5 (57.7)	<0.001
Summer (vs winter)	4,820 (181.3)	<0.001	−108.0 (72.8)	0.139	−381.1 (61.4)	<0.001
Autumn (vs winter)	3,914 (117.0)	<0.001	292.1 (77.5)	<0.001	308.1 (51.6)	<0.001
Holiday (vs non-holiday)	−746.9 (292.6)	0.011	698.5 (211.5)	0.0011	−148.3 (121.0)	0.221
Weekend (vs weekday)	−703.9 (142.2)	<0.001	1,105 (70.1)	<0.001	−289.2 (51.8)	<0.001
Temperature (per °C)	114.5 (12.1)	<0.001	−4.6 (5.1)	0.369	−15.2 (4.3)	<0.001
Rainfall (per mm per day)	−75.2 (8.2)	<0.001	−0.5 (3.2)	0.869	−2.1 (4.3)	0.631
R^2	0.756		0.549		0.339	

The PC2 holiday coefficient was stable on the count scale (Table 6), from 616.5 with 3 harmonics (p = 0.0019) to 748.9 without smoothing (p = 8.4e-4). It was 669 on all 365 days (p = 1.0e-3) and 1,039 after dropping the four weekend holidays (p = 3.1e-8). On log(1 + count) curves, where PC1 carried 73.55% of the variance, the holiday coefficient was +0.115 (p = 0.81). Holidays therefore resembled weekends in absolute curve shape, but the pattern was not established for proportional curve shape.

Table 6: Sensitivity of the holiday coefficient on PC2 (adjusted model, HC3 standard errors).

Variant	Days	Holiday coefficient on PC2	p
Main analysis (5 harmonics)	352	698.5	0.0011
3 harmonics	352	616.5	0.0019
8 harmonics	352	732.7	<0.001
No smoothing (hourly counts)	352	748.9	<0.001
All 365 days (outage days kept)	365	669.0	0.0010
Weekend holidays dropped	348	1,039	<0.001
Corrected holiday flag (20 holiday days)	352	920.7	<0.001
log(1 + count) curves*	352	0.115	0.806

*On log(1 + count) curves the coefficient is in log units.

For level, the baseline model for log daily total (R^2 = 0.833) gave a holiday coefficient of −0.309 (SE 0.081, p = 1.5e-4), about 27% lower. With the corrected flag (20 holiday days, 15 on weekdays), the PC2 coefficient was unchanged in direction and stronger (+920.7, SE 227.0, p = 6.2e-5), whereas the PC1 coefficient (−410.3, SE 315.5, p = 0.19) was not distinguishable from zero, and log daily total was about 20% lower (−0.219, SE 0.086, p = 0.012).

4 DISCUSSION

The three components separated level from shape. Season acted mainly on level (PC1, ϵ^2 = 0.509) and on the morning-midday versus evening contrast (PC3, 0.327). Day type acted on shape: weekend PC2 scores differed strongly (r = 0.89) although daily totals differed little (17,950 on weekdays, 16,330 on weekends), a difference that totals alone would have hidden.

Holidays took the weekend shape. Holiday curves lacked the weekday peaks at 08:00 and 18:00 and peaked once, at 17:00 (879). The adjusted PC2 coefficient agreed in direction under the corrected flag and strengthened when weekend holidays were dropped. This applies to absolute counts. On log(1 + count) curves we did not detect it, so we do not claim that holidays and weekends share relative curve shape.

The holiday effect on level was modest and depended on the measure. Under the corrected flag the PC1 coefficient was not distinguishable from zero, whereas log daily total was about 20% lower. With 15 to 17 holiday days the estimates were imprecise.

Within summer, PC1 fell as temperature rose (ρ = −0.38), and the loess curve in Figure 6C rose to a maximum at about 23 °C and declined at higher temperatures. This is a descriptive feature of the smoother, not an estimated threshold, and the within-summer decline is an association within one summer. Temperature was collinear with season (R^2 = 0.80), and the linear temperature term in Equation 4 averages over this curvature.

Our results agree in direction with earlier work. In Washington, DC, cold, rain and high humidity reduced both the chance of a trip and its length [2], and in Oslo rain lowered use [3]; our negative rain coefficient on PC1 (−75.2 per mm of daily rainfall) is in the same direction. In Daejeon, weather and non-working-day effects differed by hour of day [4], as did the day-type differences in our curves (Figure 2). Use there also fell above 30 °C [4]; PC1 fell with temperature within summer in the same direction, but we do not claim the same threshold, because ours is a daily mean temperature within one summer. In Lyon, movements were mostly cyclostationary over the week [11], which supports treating each day as one repetition of a 24-hour cycle.

4.1 Limitations

The data covered one city and one year, so seasonal and holiday estimates rest on a single annual cycle. The analysis was observational and supports associations only. The Holiday flag contained errors, and only 17 holidays fell on usable days, 13 of them on weekdays, so holiday estimates were imprecise and depended on the flag, as the PC1 result showed. Temperature was collinear with season (R^2 = 0.80; largest VIF 7.5, for summer). Daily scores were serially dependent (lag-1 autocorrelation of PC1 0.744), and HC3 and Newey-West standard errors were used to allow for this. Visibility and humidity had data-quality problems (Section 2.1), so neither entered the models. The data held no spatial or station-level information. Finally, the PC2 holiday result was scale-dependent: it held on counts but was not detected on log(1 + count) curves, so it should not be read as a finding about proportional curve shape.

5 CONCLUSION

Functional principal component analysis of 352 daily rental curves gave three components that carried 95.5% of the variance, and the first measured overall level and was associated with season (ϵ^2 = 0.509) and temperature (ρ = 0.72). After adjustment for weekend, season, temperature and rain, holidays were associated with a shift towards the weekend shape on PC2 (+698.5, p = 0.0011; stronger with the corrected holiday flag), but this was not detected on log(1 + count) curves, so it describes absolute and not proportional curve shape. A holiday effect on level was not established, the data covered one city and one year, and multi-year, station-level data would test whether these patterns hold.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The authors gratefully acknowledge the Dean, Faculty of Science, University of Rajshahi, for supporting this research.

Funding Source

This work was financially supported by a research grant from the Dean, Faculty of Science, University of Rajshahi (Grant No. A-156/5/52/RU./Science-3/2025-2026).

Disclosure of conflict of interest

The corresponding author declares no conflict of interest on behalf of all authors.

Statement of ethical approval

This study used a publicly available, anonymized dataset (CC BY 4.0).

Data availability

The data analyzed in this study are publicly available from the UCI Machine Learning Repository (Seoul Bike Sharing Demand, https://doi.org/10.24432/C5F62R) [16].

REFERENCES

[1] Fishman E. Bikeshare: A review of recent literature. Transp Rev. 2016;36(1):92-113. Available from: https://doi.org/10.1080/01441647.2015.1033036

[2] Gebhart K, Noland RB. The impact of weather conditions on bikeshare trips in Washington, DC. Transp. 2014;41(6):1205-1225. Available from: https://doi.org/10.1007/s11116-014-9540-7

[3] Bakke EL, Tørset T. How does the weather affect the shared bicycle usage? Dan J Transp Res. 2019;1:13-29. Available from: https://doi.org/10.5278/ojs.djtr.v1i0.3560

[4] Kim K. Investigation on the effects of weather and calendar events on bike-sharing according to the trip patterns of bike rentals of stations. J Transp Geogr. 2018;66:309-320. Available from: https://doi.org/10.1016/j.jtrangeo.2018.01.001

[5] Corcoran J, Li T, Rohde D, Charles-Edwards E, Mateo-Babiano D. Spatio-temporal patterns of a Public Bicycle Sharing Program: The effect of weather and calendar events. J Transp Geogr. 2014;41:292-305. Available from: https://doi.org/10.1016/j.jtrangeo.2014.09.003

[6] Eren E, Uz VE. A review on bike-sharing: The factors affecting bike-sharing demand. Sustain Cities Soc. 2020;54:101882. Available from: https://doi.org/10.1016/j.scs.2019.101882

[7] Sathishkumar VE, Park J, Cho Y. Seoul bike trip duration prediction using data mining techniques. IET Intell Transp Syst. 2020;14(11):1465-1474. Available from: https://doi.org/10.1049/iet-its.2019.0796

[8] Sathishkumar VE, Cho Y. Season wise bike sharing demand analysis using random forest algorithm. Comput Intell. 2024;40(1):e12287. Available from: https://doi.org/10.1111/coin.12287

[9] Leem S, Oh J, Moon J, Kim M, Rho S. Enhancing multistep-ahead bike-sharing demand prediction with a two-stage online learning-based time-series model: insight from Seoul. J Supercomput. 2024;80(3):4049-4082. Available from: https://doi.org/10.1007/s11227-023-05593-6

[10] Gao C, Chen Y. Using machine learning methods to predict demand for bike sharing. In: Stienmetz JL, Ferrer-Rosell B, Massimo D, editors. Information and communication technologies in tourism 2022. Cham: Springer; 2022. p. 282-296. Available from: https://doi.org/10.1007/978-3-030-94751-4_25

[11] Borgnat P, Abry P, Flandrin P, Robardet C, Rouquier JB, Fleury E. Shared bicycles in a city: A signal processing and data analysis perspective. Adv Complex Syst. 2011;14(3):415-438. Available from: https://doi.org/10.1142/S0219525911002950

[12] Ramsay JO, Silverman BW. Functional data analysis. 2nd ed. New York: Springer; 2005.

[13] Wang JL, Chiou JM, Müller HG. Functional data analysis. Annu Rev Stat Appl. 2016;3:257-295. Available from: https://arxiv.org/abs/1507.05135

[14] Bouveyron C, Côme E, Jacques J. The discriminative functional mixture model for a comparative analysis of bike sharing systems. Ann Appl Stat. 2015;9(4):1726-1760. Available from: https://doi.org/10.1214/15-AOAS861

[15] Piter A, Otto P, Alkhatib H. The Helsinki bike-sharing system: insights gained from a spatiotemporal functional model. J R Stat Soc Ser A Stat Soc. 2022;185(3):1294-1318. Available from: https://doi.org/10.1111/rssa.12834

[16] Seoul Bike Sharing Demand [Dataset]. UCI Machine Learning Repository; 2020. Available from: https://doi.org/10.24432/C5F62R

[17] Hall P, Müller HG, Wang JL. Properties of principal component methods for functional and longitudinal data analysis. Ann Stat. 2006;34(3):1493-1517. Available from: https://doi.org/10.1214/009053606000000272