Article https://www.korseaj.org

파크골프장의 효율적 관리를 위한 군집 기반 유형화 연구

이부건[1], 송윤아[2], 김영선[3], 윤상후[2,4*]

Typological Classification of Park Golf Courses Using Cluster Analysis for Efficient Management

Bugeon Lee[1], Yunah Song[2], Young-Sun Kim[3] and Sanghoo Yoon[2,4*]

[1]전남대학교 수학/통계학과, [2]전남대학교 통계학과, [3]대구대학교 스마트원예학과, [4]전남대학교 G-LAMP사업단
[1]Department of Mathematics and Statistics, Chonnam National University, Gwangju, 61186, Korea, [2]Department of Statistics, Chonnam National University, Gwangju, 61186, Korea, [3]Department of Horticultural Science, Daegu University, Gyeongsan 38453, Korea, [4]G-LAMP research institute, Chonnam National University, Gwangju, 61186, Korea
* Correspondence: statstar77@jnu.ac.kr

Abstract: This study conducted a comprehensive analysis of the multi-dimensional characteristics of park golf courses across South Korea to establish a typology system and provide a foundation for tailored management and policy development. A total of 424 park golf courses were analyzed using data provided by the Korea Park Golf Association. The variables considered included climate favorability (based on the Korea Tourism Climate Index, K-TCI), surrounding environmental features (such as proximity to rivers, parks, and land cover types), accessibility indicators (population density and age structure), and facility size (number of holes). The 12 K-TCI indicators were reduced to three principal components—Peak Season Comfort, Winter Limitation, Summer Polarization—through Logratio Principal Component Analysis (log ratio PCA), which explained 88.2% of the cumulative variance. These three components, along with six additional variables, were integrated into a Gaussian Mixture Model (GMM)-based cluster analysis. Based on the Bayesian Information Criterion (BIC), six distinct clusters were identified, each representing a unique typological group of park golf courses.

Keywords: Cluster analysis, Gaussian Mixture Model, Park golf, Principal component analysis, Typology

https://doi.org/10.5338/KJEA.2025.44.38

Agric. Environ. Sci. 2025, 44, 386-395

Received: October 14, 2025
Revised: October 27, 2025
Accepted: November 12, 2025
Published: December 8, 2025

Online ISSN: 2233-4173
Print ISSN: 1225-3537

Check for updates

ⓒ The Korean Society of Environmental Agriculture 2025

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

서 론

파크골프(park golf)는 1983년 일본 홋카이도의 소도시에서 시작되어, 국내에는 1998년 보광 휘닉스파크 부대시설로 도입된 뒤 2003년 대한파크골프협회 창립을 계기로 함께 보급이 본격화된 생활 스포츠이다. 파크골프는 정규 골프와 다르게 단일 클럽을 사용하여 비교적 짧은 코스를 플레이하는 종목으로 정규 골프에 비해 진입 장벽과 비용부담이 낮다. 쉽고 간단히 즐길 수 있는 파크골프장은 세대 간 교류 및 지역 커뮤니티 형성 등 공동체 활동에 효과적이다[1].

파크골프에 관한 연구를 살펴보면 파크골프 참여가 노년층의 여가만족과 생활만족, 자기효능감 및 삶의 질 향상과 유의하게 연관됨을 보이고 있다[2-4]. 최근에는 파크골프가 노년층의 사회자본 축적 및 성공적 노화에 기여하는 선행연구가 등장하여 많은 관심이 생기는 추세이다[5]. 이러한 추세로 인해 전국 파크골프장의 수도 늘어나고 있다. 대한파크골프협회(www.kpga7330.com)에 따르면 2025년 9월 기준 전국에 총 424개의 파크골프장이 운영 중이다. Chang[6]은 노인 인구의 증가로 파크골프장이 앞으로 꾸준히 신설 또는 증설된다고 주장했다. 정책, 수요 측면에서도 등록 선수 증가와 저연령층 참여 확대가 관측되어 지자체의 시설 조성 및 확충이 적극화되고 있다. 뉴스 키워드 분석을 통한 미디어, 여론 측면에서도 파크골프 보도의 양적 증가와 함께 '복지', '건강' 프레이밍이 강화되는 등 사회적 관심의 확산이 확인된다[7]. 그러나 이러한 양적 팽창은 전국 파크골프장들의 이질성(heterogeneity)을 심화시켜 체계적인 관리에 어려움이 존재한다. 표본조사 이론에서 이질적인 모집단은 특성이 유사한 동질적 집단(층, strata)으로 구분하여 접근할 때 대표성과 분석 정밀도를 높일 수 있듯이[8], 전국 단위의 파크골프장 역시 효과적인 관리를 위해서는 유형별 분류가 필요하다.

이에 본 연구는 전국의 파크골프장을 분석 단위로 하여, 다차원적 특성을 반영하는 변수들을 종합적으로 고려한 군집분석을 수행하고자 한다. 분석에는 기후 우호도(한국형 관광기후지수), 주변 환경(인근 하천, 공원 유무, 토지피복), 접근성(인구밀도, 연령구조) 그리고 규모 변수를 활용하였다. 본 연구는 군집분석을 통해 전국 단위 모니터링의 대표성과 효율성을 검증하고, 시설 관리와 환경 리스크 완화 등 실무적 의사결정을 지원하는 것을 목적으로 한다. 아울러 수요 증가에 대응하여 신규 골프장 개설을 위한 후보 입지 선별과 적정 규모(홀 수) 산정, 수요–공급 격차 진단에 활용 가능한 근거를 제시하며, 국가 단위 관리 체계의 구체화에 기여하고자 한다.

결과 및 고찰

파크골프장 유형화를 위한 군집분석의 첫 단계로 K-TCI 변수의 차원 축소를 수행하였다. K-TCI는 12개의 계절별 범주 비율로 구성되며, 이 중 '여름_불리함'과 '가을_불리함'은 대부분의 관측에서 값이 0에 가까워 변별력이 낮다고 판단해 분석에서 제외하였다. 최종적으로 10개 변수에 대해, 합이 1인 조성자료(compositional data)의 기하학적 특성에 부합하도록 로그비율 변환 기반 주성분분석(log-ratio PCA)을 적용하였다. 이때 10개 변수 내 소량의 0값은 수치적 안정을 위해 소규모 양의 값으로 대체하였다. 이를 통해 탐색 단계에서는 clr(centred log-ratio) 변환을 사용하였으며, 최종 차원 축소는 정규직교 기반의 ILR (isometric log-ratio) 좌표계에서 PCA로 추정하였다. 이는 조성자료 분석에서 발생할 수 있는 거짓 상관(spurious correlation)을 피하고 비율 간 상대적 크기 정보를 보존하는 방식으로 변동 구조를 식별하기 위함이다. 주성분 개수는 scree plot에 의한 시각적 판단과 고유값 1 이상을 기준으로 하는 Kaiser 기준을 종합하여 3개로 결정하였다[9,10]. 분석 결과(Table 1), 상위 3개 주성분의

Table 1. The result of principal components analysis

K-TCI	PC1	PC2	PC3
spring_good	0.44	0.12	-0.04
spring_normal	0.22	0.3	-0.59
spring_bad	-0.4	-0.21	0.29
summer_good	0.37	0.15	0.48
summer_normal	-0.37	-0.15	-0.48
autumn_good	0.34	-0.35	0.01
autumn_normal	-0.34	0.35	-0.01
winter_good	-0.09	0.52	0.11
winter_normal	-0.01	-0.52	0.01
winter_bad	0.27	-0.16	-0.32
SS loading	2.087	1.791	1.123
Proportion Var	0.436	0.321	0.126
Cumulative Var	0.436	0.757	0.883

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

누적 설명분산은 88.2%로 나타나 원자료의 변동을 충분히 설명하는 것으로 판단하였다.

PC1은 봄, 여름, 가을의 '좋음' 비중이 높고 같은 계절의 '보통', '나쁨' 비중이 낮을수록 값이 커져, 주요 활동 시즌의 전반적 쾌적성을 나타낸다. 이에 "성수기 쾌적성(Peak Season Comfort)"으로 명명하였다. PC2는 겨울철 '나쁨'(+)과 '좋음'(–)가 대칭적으로 크게 작용하여 겨울 날씨의 취약도를 나타낸다. 이에 "겨울철 취약성(Winter Limitation)"으로 명명하였다. 마지막으로 PC3는 봄, 여름의 '보통' 비중이 높을수록 값이 낮아지고, 여름철 '좋음' 비중이 높을수록 값이 커지는 뚜렷한 대비를 보인다. 이에 여름 성과의 특화 정도를 의미하는 "여름 성과 극화(Summer Polarization)"로 명명하였다.

각 요인이 설명하는 분산의 크기는 성수기 쾌적성 (PC1)이 43.6%로 가장 높았으며, 겨울철 취약성(PC2, 32.1%)과 여름 성과 극화(PC3, 12.6%)가 그 뒤를 이었다. 이는 파크골프장의 기후 환경 특성을 구분하는 데 있어 성수기의 날씨가 가장 중요한 요인임을 보여준다.

최종적으로 주성분 분석을 통해 변환된 3개의 회전된 주성분 점수를 새로운 기후 환경 변수로 사용했으며, 군집분석에 사용될 최종 데이터는 Table 2와 같다.

Table 2. The final data for cluster analysis

Region	Park golf name	Peak Season Comfort	Winter Limitation	Summer Polarization	River	Park	Cropland _rate	Pop_ total	Senior _rate	Hole
Gyeongnam	Gagok	0.082	-3.878	0.003	1	0	0.288	63,886	0.722	18
Gyeongnam	Gasan Waterfront Park	-2.826	-1.886	-0.708	0	0	0.045	225,625	0.587	36
Gyeongbuk	Gasan	0.042	-1.034	0.464	1	0	0.000	4,980	0.580	9
Chungnam	Gayagok	1.422	0.950	0.252	0	0	0.290	8,041	0.680	9
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Daegu	Hyoryeong	0.381	-0.712	-0.371	1	0	0.426	3,334	0.818	9

최종 구축된 변수들을 이용하여 파크골프장의 유형화를 위해 GMM 기반의 군집분석을 실시했다. 최적의 군집 수를 결정하기 위해 군집 개수를 2개부터 10개까지 변화시키며 모델을 생성하고 베이지안 정보 기준(Bayesian Information Criterion, BIC)을 비교했다. 군집 수가 6개일 때 BIC가 –5480.8로 가장 높은 값을 보여 최적 군집 수를 6개로 결정하였다. 군집별 평균 점수를 구하면 Table 3이다.

Table 3. Average of the park golf course clusters

Cluster	Peak season comfort	Winter Limitation	Summer Polarization	River	Park	Cropland rate	Pop total	Senior rate	Hole
1	0.194	-2.298	0.381	0.623	0	0.361	107971	0.686	24.5
2	-0.475	0.565	-0.03	0	0	0.124	133373	0.643	18.3
3	-4.188	1.101	-1.573	0.125	0.375	0.040	51282	0.625	15.8
4	0.064	0.482	0.025	1	0	0.313	139980	0.659	20.3
5	0.409	0.2	-0.031	0.647	1	0.143	195530	0.621	17.3
6	1.58	0.343	-0.668	0.706	0.353	0.084	1397336	0.578	10.7

군집 1은 성수기 쾌적성(PC1)이 0.194로 평균 수준이고, 겨울 취약성(PC2)은 -2.298로 매우 낮아 겨울에도 비교적 우호적이다. 여름 특화(PC3)는 0.381로 약한 강점이 보인다. 입지는 하천 300m 이내 비율 0.623로 강변 성격이 뚜렷하고, 공원 인접은 없다(Park=0). 경작지 인접도는 0.361로 중간 이상, 인구 수는 107,971명으로 중·대형 수요권, 평균 24.5홀로 규모도 중대형이다. 이에 본 군집은 "온화기후 강변, 대형형(Mild-Riverfront, Large-scale, MRL)"으로 해석되며 69개의 파크골프장이 속한다.

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

군집 2는 하천과 공원 어느 쪽에도 입지하지 않는 독립적인 입지 특성이 가장 뚜렷하게 보인다. 기후적으로는 성수기 쾌적성(PC1)이 -0.475로 낮은 편에 속하며, 겨울철 취약성(PC2)은 0.565로 (군집 1의 -2.298에 비해) 높아 겨울이 온화하지 않고 다소 혹독한 기후 특성을 보인다. 여름 성과 극화(PC3)는 -0.03으로 거의 0에 근접해 계절적 특화가 두드러지지 않는다. 규모는 평균 18.3홀로 중간 규모를 가지며 40세 이상 비율은 0.643으로 높은 수준을 보인다. 이에 본 군집은 "내륙 독립형(Inland-Standalone, IS)으로 해석되며 106개의 파크골프장이 속한다.

군집 3은 성수기 쾌적성(PC1)이 -4.188로 모든 군집 중 압도적으로 낮고, 겨울철 취약성(PC2)은 1.101로 가장 높아 기후적으로 가장 불리한 특성을 명확하게 보인다. 또한 여름 성과 극화(PC3) 수치도 -1.573으로 매우 낮아, 뚜렷한 계절 특성 없이 전반적으로 계절적 특화가 보이지 않는다. 입지적으로는 공원 내 비율이 37.5%로 일정 부분 공원에 입지하는 특성을 보이며, 농경지 비율은 4%로 매우 낮다. 특히 인구 수가 51,282명으로 모든 군집 중 가장 낮고, 코스 규모역시 평균 15.8홀로 가장 작은 편에 속한다. 이에 본 군집은 "기후 최약, 소규모형(Climate Vulnerable, Small scale, CVS)으로 해석되며, 8개의 파크골프장이 속한다.

군집 4는 군집 내 모든 파크골프장이 하천 300m 이내에 입지(River=1)하고 공원 내에는 전혀 입지하지 않는(Park=0) 전형적인 강변 입지 특성을 보인다. 기후적으로 성수기 쾌적성(PC1)과 여름 성과 극화(PC3)는 모두 0에 근접해 뚜렷한 강점이 없으며, 겨울철 취약성(PC2)은 0.482로 다소 높아 겨울 기후가 약간 불리한 경향을 보인다. 농경지 비율은 31.3%로 높은 편이고, 코스 규모는 평균 20.3홀로 비교적 크며, 인구 수와 40세 이상 비율도 준수한 수준이다. 이에 본 군집은 100% 강변 입지라는 특성을 강조하여 “전형적 강변형(Typical Riverfront, TR)”으로 해석되며, 139개의 파크골프장이 속한다.

군집 5는 군집 내 모든 파크골프장이 공원 내에 입지(Park=1)하는 전형적인 공원 입지 특성을 보인다. 또한 64.7%가 하천 300m 이내에 위치하여, 하천을 끼고 조성된 도시 공원의 양상을 띤다. 기후적으로 성수기 쾌적성(PC1)은 0.409로 다소 높아 전반적으로 쾌적하며, 겨울철 취약성(PC2)은 0.200으로 0에 근접해 겨울 조건이 비교적 온화하다. 인구 수는 195,530명으로 두 번째로 높아 수요가 풍부한 도시 지역에 위치함을 시사한다. 코스 규모는 평균 17.3홀로 중간 수준이다. 이에 본 군집은 100% 공원 입지와 풍부한 배후 인구라는 특성을 반영하여 “도시 공원형(Urban Park, UP)”으로 해석되며, 85개의 파크골프장이 속한다.

마지막으로 군집 6은 인구 수가 1,397,336명으로 6개 군집 중 압도적으로 높아 대도시 도심권의 특성이 가장 두드러진다. 또한 70.6%가 하천 300m 이내에 위치하여 하천 접근성이 매우 우수하다. 기후적으로 성수기 쾌적성(PC1)은 1.58로 모든 군집 중 가장 높아 전반적으로 매우 쾌적하고, 겨울철 취약성(PC2)은 0.343으로 0에 근접해 겨울 조건도 비교적 양호하다. 평균

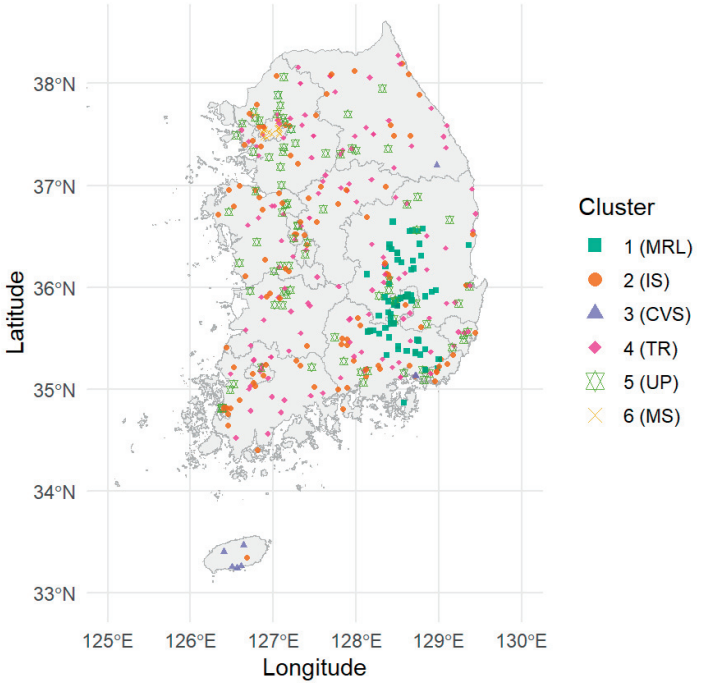

Fig. 1. Spatial distribution of park golf course clusters.

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

홀 수는 10.7홀로 가장 작고, 40세 이상 비율은 0.578로 가장 낮아 상대적으로 젊은층 비중이 높은 대도시적 인구 구성을 시사한다. 이에 본 군집은 압도적인 배후 인구와 작은 규모의 특성을 반영하여 “대도시 소규모형(Metro, Small scale, MS)”으로 해석되며, 17개의 파크골프장이 속한다.

이렇게 정리된 군집을 바탕으로 우리나라 지도에 공간적 분포를 그리면 Fig 1과 같으며, 광역자치단체별로 어떻게 분포하는지 정리한 결과는 Table 4와 같다.

Table 4. Frequency of cluster type by region

Region	N	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Gangwon	37	0	9	1	21	6	0
Gyeonggi	43	0	9	0	12	21	1
Gyeongsangnam-do	66	21	17	2	18	8	0
Gyeongsangbuk-do	62	23	9	0	19	11	0
Gwangju	9	0	4	0	4	1	0
Daegu	34	24	2	0	4	4	0
Daejeon	4	0	1	0	1	2	0
Busan	11	1	7	0	1	2	0
Seoul	25	0	2	0	7	0	16
Sejong	7	0	3	0	1	3	0
Ulsan	7	0	1	0	3	3	0
Incheon	5	0	2	0	2	1	0
Jeollanam-do	36	0	15	0	15	6	0
Jeollabuk-do	22	0	4	0	12	6	0
Jeju	6	0	1	5	0	0	0
Chungcheongnam-do	32	0	13	0	9	10	0
Chungcheongbuk-do	18	0	7	0	10	1	0
Total	424	69	106	8	139	85	17

파크골프장 유형의 공간적 분포는 군집별로 뚜렷한 지역적 편중을 보였다. 군집 1(온난기 강변·대형형, MRL)은 경북, 경남, 대구 등 영남권에 밀집하여 비수도권에서 대규모 거점 시설로 기능하는 경향을 보이며, 군집 2(내륙 독립형, IS)는 특정 지역 편중 없이 전국 모든 시·도에 고르게 분포되는 경향을 보인다. 군집 3(기후 최약·소규모형, CVS)은 주로 제주도에 집중되어 기후 입지 제약 하의 소규모 공급 양상을 드러냈다. 군집 4(전형적 강변형, TR)는 군집 2와 유사하게 전국적 분포를 보이며, 하천을 따라 형성되는 전형적 입지 특성을 반영한다. 군집 5(도시 공원형, UP)는 수도권과 주요 대도시에 집중되어 도시공원 체계와의 결합을 통해 공급되는 특성이 확인되었으며, 군집 6(대도시 소규모형, MS)은 서울에 중심적으로 분포되는 경향을 보인다.

이러한 공간적 분포 패턴은 파크골프장의 유형이 무작위가 아니라, 해당 지역의 도시화 수준, 개발 역사, 자연환경과 밀접한 관련을 맺고 형성되었음을 명확하게 보여준다.

결 론

본 연구는 전국 424개 파크골프장을 대상으로 기후, 주변 환경, 접근성, 규모를 나타내는 9개 변수를 활용하여 GMM 기반 군집분석을 수행하고, 관리 관점의 유형화 체계를 구축하고자 하였다. 분석 결과, 전국 파크골프장은 “온화기후 강변, 대형형(MRL)”, “내륙 독립형(IS)”, “기후 최약, 소규모형(CVS)”, “전형적 강변형(TR)”, “도시 공원형(UP)”, “대도시, 소규모형(MS)”으로

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

총 여섯 가지 대표 유형이 성공적으로 분류되었다.

각 유형의 핵심 특성을 요약하면 다음과 같다. “온화기후 강변, 대형형”은 우수한 성수기 기후와 강변 입지, 큰 규모를 특징으로 하는 고령층 중심 시설이었다. “내륙 독립형”은 파크골프장 인근 하천과 공원이 입지하지 않아 독립적이고 계절의 특화가 없는 시설이다. “기후 최약, 소규모형”은 기후의 취약성을 가지고 작은 규모를 특징으로 하는 시설이다. “전형적 강변형”은 모든 파크골프장 인근 하천이 존재하고 겨울의 기후가 불리한 경향을 보이는 시설이다. “도시, 공원형”은 경기도의 신도시를 중심으로 접근성이 뛰어난 공원 내 시설의 특징을 보였다. 마지막으로 “대도시, 소규모형”은 서울에 집중된 소규모 시설로, 불리한 겨울 기후에도 불구하고 압도적인 인구 수를 가진 유형으로 나타났다.

이러한 유형화 결과는 파크골프장의 획일적인 관리에서 벗어나 특성에 맞는 차별화된 정책을 수립하는 데 중요한 근거를 제공한다. 예를 들어, 하천에 인접한 ‘강변’ 유형들(온화기후 강변, 대형형, 전형적 강변형)에 대해서는 배수 시설 점검 및 약제 유출 방지를 위한 환경 관리 가이드를 우선적으로 강화할 수 있다. 또한, 기후 특성에 따라 ‘성수기 우호형’ 유형은 이용객 집중 관리 방안을, ‘겨울 비우호형’ 유형은 동계 대체 프로그램 개발을 고려하는 등 효율적인 계절별 운영 전략을 수립하는 데 직접적으로 활용될 수 있다.

본 연구의 유형화 및 시각화는 전국 파크골프장의 체계적 관리, 환경 영향 저감, 수요 기반 운영 전략 및 대민 정보 서비스 강화 등 실무 전반에 활용될 수 있는 기초 자료를 제공했다는 점에서 의의가 있다. 향후 연구에서는 침수 취약도나 배수 인프라와 같은 수문 환경 변수와 실제 이용 데이터를 추가하여 정책 활용성을 더욱 제고하고, 패널 데이터 분석을 통해 기후 변화와 같은 장기적인 환경 변화에 대응하는 동적인 관리 체계를 모색할 필요가 있다.

재료 및 방법

한국형 관광기후지수

본 연구는 파크골프장의 기후 환경을 정량화하기 위해 한국형 관광기후지수(Korean Tourism Climate Index, K-TCI)를 사용했다. K-TCI는 Mieczkowski[11]가 제안한 관광기후지수(Tourism Climate Index, TCI)의 가중치를 한국의 계절 특성에 맞게 조정한 지수이다[12]. K-TCI는 총 7개의 기상 변수를 5개의 하위 요소로 조합하여 계산한다. 5개의 요소는 주간 열쾌적성(Cd), 야간 열쾌적성(Ca), 강수(P), 풍속(W), 일조시간(S)으로 구성된다. 계절별 날씨의 중요도를 다르게 반영하기 위해 각 요소의 가중치를 다음 식과 같이 나타냈다.

$$K\text{-}TCI_{\text{Spring}} = 2(2.58Cd + 1.93Ca + 3.04P + 1.31W + 1.14S),$$
$$K\text{-}TCI_{\text{Summer}} = 2(3.07Cd + 1.90Ca + 3.27P + 0.90W + 0.86S),$$
$$K\text{-}TCI_{\text{Autumn}} = 2(2.37Cd + 2.18Ca + 3.43P + 1.04W + 0.98S),$$
$$K\text{-}TCI_{\text{Winter}} = 2(2.27Cd + 2.39Ca + 3.07P + 1.46W + 0.81S),$$

여기서 열쾌적성 지표(Cd, Ca)는 유효온도(Effective Temperature, ET)를 이용하여 다음 식과 같이 계산된다[13].

$$ET_{C_d} = T_{\max} - 0.4(T_{\max} - 10)\left(1 - \frac{RH_{\min}}{100}\right),$$
$$ET_{C_a} = T_{\text{mean}} - 0.4(T_{mean} - 10)\left(1 - \frac{RH_{mean}}{100}\right),$$

여기서 T는 기온, RH는 습도를 의미한다.

최종적으로 K-TCI 산출에는 최고기온, 최저 상대습도, 평균기온, 평균 상대습도, 총 강수량, 평균 풍속, 총 일조시간의 7가지 기상 관측 변수가 사용된다. 본 연구에서는 봄(3~5월), 여름(6~8월), 가을(9~11월), 겨울(12~2월)의 계절 구분에 따라 산출된 K-TCI 점수를 분석에 활용하였다.

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

가우시안 혼합 모델 기반 군집분석

본 연구는 파크골프장을 기준으로 다차원적 특성을 반영하는 변수를 이용하여 군집화하기 위해 가우시안 혼합 모델(Gaussian Mixture Model, GMM)을 사용하였다. GMM은 데이터가 여러 개의 가우시안 분포가 혼합된 형태로 나타난다고 가정한다. 군집화를 위해 전체 데이터 분포를 k개의 하위 정규 분포의 선형 결합으로 모델링하여 복잡한 다변량 구조를 유연하게 반영한다[14]. GMM의 모수는 EM 알고리즘을 이용한 최대우도추정법으로 추정된다[15]. GMM의 모델 구조는 다음과 같다.

$$p(x) = \sum_{k=1}^{K} \pi_k N(x \mid \mu_k, \sum_k), \pi_k \geq 0, \sum_{k=1}^{K} \pi_k = 1,$$

여기서 K는 혼합 모델을 구성하는 가우시안 분포의 개수(군집의 수)이며, π_k는 전체 데이터 중 k번째 군집이 차지하는 비율이다. $N(x|\mu, \sum_k)$는 k번째 가우시안 분포의 확률 밀도 함수이다. μ_k와 $\sum_k$는 각각 k번째 가우시안 분포의 평균 벡터와 공분산 행렬로 군집의 중심 위치와 형태, 그리고 방향성을 결정한다[16]. 분석은 R 프로그램의 mclust 패키지를 이용하여 수행했으며, 최적 군집 수는 BIC 기준을 활용하였으며, mclust 패키지에서는 BIC 값이 가장 큰 모델을 최적으로 판단한다[17].

변수선정 및 구성근거

본 연구는 전국 파크골프장의 데이터 기반 유형화를 위해, 서비스 시설 입지 분석의 핵심 개념인 공급-수요-환경적 맥락을 포괄하는 분석 틀을 설정하였다[18,19]. 구체적으로 기후환경과 주변환경은 환경적 맥락을, 접근성은 수요를 그리고 규모는 공급을 대표하는 요인으로 구성했다. 각 요인을 구성하는 다변량 자료와 데이터 출처는 Table 5와 같다.

Table 5. Data composition for park golf course typology

Group	Variable	Source
Climatic environment	Spring (Good, Normal, Bad) Summer (Good, Normal, Bad) Autumn (Good, Normal, Bad) Winter (Good, Normal, Bad)	Korea Meteorological Administration (KMA)
Site environment	River Park Cropland_rate	National Spatial Data Infrastructure (NSDI) Public Data Portal Worldcover
Accessibility	Total_pop Senior_ratio	Worldpop
Facility size	Hole	Korea Parkgolf Association

기후환경은 야외 레저 활동 참여와 만족도를 결정하는 핵심 요인이다. 특히 파크골프와 같은 실외 스포츠는 기온, 강수, 일조 등 기후 조건에 직접적인 영향을 받으므로 잠재적 이용 가능일을 측정하는 것은 수요 예측의 중요한 기준이 된다[20]. 본 연구에서는 계절별 기후의 우호도를 '좋음', '보통', '나쁨'의 세 범주로 나누어 파크골프 활동에 적합한 시기와 제약 요인을 변수로 구성하였다.

주변 환경은 시설의 공간적 맥락과 관리활동의 잠재적 환경영향을 함께 고려하여 선정하였다. 잔디 관리를 위한 비료나 약제 살포시 이 과정에서 인접 환경에 미치는 잠재적 영향이 지속적으로 논의되어 왔으며[21], 하천 인접 여부(River)는 잠재적 환경 리스크를 가늠하는 핵심 지표가 된다[22]. 또한, 공원(Park) 내에 위치한 시설은 다른 여가 활동과의 시너지를 창출한다. 농작지 비율(Cropland rate)도 약제 살포시 농작지에 미치는 잠재적 영향을 평가하기 위해 사용되었으며, 간접적으로 농촌 지역 특성을 반영하는 변수로 활용하였다.

접근성은 시설의 잠재적 수요 기반을 측정하는 지표이다. 레크레이션 시설의 이용률은 기본적으로 배후 인구 규모(Total

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

pop)에 비례하며, 주요 이용자 계층의 인구학적 특성에 따라 크게 달라진다[23]. 파크골프가 주로 중장년 및 노년층에서 높은 인기를 보이는 점을 고려하여, 40세 이상 인구 비율(Senior ratio)을 함께 변수로 구성하였다.

마지막으로 시설 규모는 서비스 수용력과 관리 범위를 간접적으로 나타내는 지표이다. 시설의 홀(Hole) 수는 수용 가능한 이용객 수를 결정하고, 유지관리에 필요한 자원의 규모를 가늠하게 하는 핵심적인 물리적 특성이다.

변수 측정 방법

앞서 선정한 각 요인의 변수들은 Table 6 및 Table 7과 같이 수집되었다.

Table 6. Climate environment variables

Region	Park golf name	Spring _good	Spring _normal	Spring _bad	Summer _good	Summer _normal	Summer _bad	Autumn _good	Autumn _normal	Autumn _bad	Winter _good	Winter _normal	Winter _bad
Gyeongnam	Gagok	0.579	0.416	0.004	0.265	0.735	0.000	0.502	0.498	0.000	0.260	0.740	0.000
Gyeongnam	Gasan Waterfront Park	0.427	0.561	0.012	0.157	0.843	0.000	0.410	0.590	0.000	0.262	0.734	0.003
Gyeongbuk	Gasan	0.627	0.368	0.004	0.277	0.723	0.000	0.495	0.505	0.000	0.110	0.885	0.006
Chungnam	Gayagok	0.598	0.401	0.001	0.342	0.658	0.000	0.471	0.524	0.004	0.110	0.872	0.019
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Daegu	Hyoryeong	0.604	0.393	0.002	0.245	0.755	0.000	0.493	0.505	0.002	0.127	0.866	0.007

Table 7. Environmental, Accessibility, and Scale Variables

Region	Park golf name	River	Park	Cropland_rate	Pop_total	Senior_rate	Hole
Gyeongnam	Gagok	1	0	0.288	63886	0.722	18
Gyeongnam	Gasan Waterfront Park	0	0	0.045	225625	0.587	36
Gyeongbuk	Gasan	1	0	0.000	4980	0.580	9
Chungnam	Gayagok	0	0	0.290	8041	0.680	9
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Daegu	Hyoryeong	1	0	0.426	3334	0.818	9

기후환경 변수는 계절별 K-TCI 범주 비율로 산출되었다. 먼저 기상청 ASOS 관측소의 일별 기상자료(2015~2024년, 3,653일)를 이용해 각 관측소의 일일 K-TCI점수를 산출하였다. 이어서 파크골프장별 일일 K-TCI를 얻기 위해 하루 단위로 크리깅(Kriging) 공간보간하여 전국 424개의 파크골프장의 일일 K-TCI를 추정하였다. 이렇게 추정된 점수는 '좋음(70점 이상)', '보통(40~70점 미만)', '나쁨(40점 미만)'의 세 범주로 분류 후 계절별 범주가 차지하는 비율을 변수로 구축했다.

주변 환경 변수는 공간 분석으로 구축했다. 하천(River)과 공원(Park)은 파크골프장 중심 300m 반경 내 하천 또는 공원의 존재 여부를 나타내는 더미 변수(존재 시 1, 그 외 0)이다. 하천의 300m 기준은 수질 영향 범위[24]와 위치 오차를 함께 고려한 것이며, 공원의 기준은 세계보건기구(World Health Organization; WHO)의 2017년 정책적 권고안을 따랐다. 농작지 비율(Cropland rate)은 약제 살포의 핵심 영향권인 100 m 반경 내 농경지 면적 비율로 계산했다[25]. 이때 사용된 하천 자료는 국가 공간 정보 자료[26]를 이용했으며, 공원 자료는 공공데이터포털에서 제공하는 전국 공원 자료를 이용했다. 또한 토지피복 자료는 Worldcover에서 제공하는 우리나라 자료를 이용했다[27].

접근성 변수는 잠재 수요를 반영하도록 구성했다. 차량 기반 생활권을 가정하여 파크골프장 중심 반경 5 km 내 총인구(Pop total)를 산출했다[28]. 또한, 파크골프의 주 이용층이 중장년층이라는 점과 최근의 이용 연령대 확대 추세를 반영하여 동일 반경 내 40세 이상 인구 비율(Senior rate)을 계산했다. 이때 인구밀도 자료는 Worldpop에서 제공하는 우리나라 인구밀도 자료를 이용했다.

https://doi.org/10.5338/KJEA.2025.44.38
Agric. Environ. Sci. 2025, 44, 386-395

시설 규모 변수인 홀 수(Hole)는 대한파크골프협회 자료를 바탕으로 각 시설의 총 홀 수를 직접 사용하였으며, 이는 서비스 수용력과 유지관리 범위를 간접적으로 나타내는 지표로 사용했다.

Data Availability: All data are available in the main text or in the Supplementary Information.

Author Contributions: Conceptualization, Yoon and Kim; Methodology, Yoon and Lee; Data curation, Yoon; Formal analysis Song and Lee; Writing – original draft Lee, Writing – review & editing, Yoon, Kim and Lee, Visualization Lee and Song, Supervision Yoon.

Notes: The authors declare no conflict of interest.

Acknowledgments: This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research, No. 512020091-3674) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF).

Additional Information:
Supplementary information The online version contains supplementary material available at https://doi.org/10.5338/KJEA.2025.44.38
Correspondence and requests for materials should be addressed to Sanghoo Yoon.
Peer review information Korean Journal of Environmental Agriculture thanks the anonymous reviewers for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.korseaj.org

References

1. Sim CS, Seong NH (2004) The effect and introduction background of park golf as sport for all. Korea Sport Research, 15(5), 551-562.
2. Shim CH, Huh JY (2006) The relationship between leisure satisfaction and life satisfaction on park golf participation in old. The Korean Journal of Physical Education, 45(2), 189-196.
3. Park JL, An, MJ (2015). The effect of the elderly park golf participant's participation motivation on the leisure satisfaction and life satisfaction. Journal of Golf Studies, 9(2), 1-12.
4. Kim OJ, Seo JR (2023) A study on the structural relationship between life stress, self-efficacy, and quality of life among the elderly participating in park golf. Journal of Convergence Consilience, 6(3), 20-35. https://doi.org/10.33090/SFCC.6.4.2
5. Lee K, Lee KE, Lee HJ, Min JW, Kang C, Heo J (2024) Social capital and successful aging among senior park golfers. Journal of Leisure Research, 56(5), 701-721. https://doi.org/10.1080/00222216.2024.2389054.
6. Chang SW (2024) Current status of park golf course and turfgrass management practices in South Korea. Weed and Turfgrass Science, 13(3), 125-134. https://doi.org/10.5660/WTS.2024.13.3.3.
7. Chi DC, Jang HY, Jung JS (2023) Analysis of keywords for news related to park golf for the elderly sports policy. The Journal of Sports and Entertainment Law, 26(2), 117-131. http://doi.org/10.19051/kasel.2023.26.2.117.
8. Lohr SL (2021). Sampling: design and analysis. Chapman and Hall/CRC.
9. Kaiser, H. F. (1960). The application of electronic computers to factor analysis Educational and psychological measurement,20(1), 141-151.
10. Cattell RB (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10.
11. Mieczkowski Z (1985) The tourism climatic index: a method of evaluating world climates for tourism. Canadian Geographer/Le Géographe Canadien, 29(3), 220-233. https://doi.org/10.1111/j.1541-0064.1985.tb00365.x.
12. Kim NJ, Kim ST (2014) Development of Korea Tourism Climate Index (KTCI). Journal of Tourism Sciences, 38(6), 253-275.
13. Houghton FC, Yaglou CP (1923) Determining equal comfort lines. J ashve,29, 165-176.
14. Reynolds D (2009) Gaussian mixture models. In: Li SZ, Jain A (eds) Encyclopedia of Biometrics. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-73003-5_196
15. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B (methodological), 39(1), 1-22.

16. Fraley C, Raftery AE (2002) Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association, 97(458), 611-631. https://doi.org/10.1198/016214502760047131.
17. Fraley C, Raftery AE (1998) MCLUST: Software for model-based cluster and discriminant analysis. Department of Statistics, University of Washington: Technical Report, 342, 1312.
18. Goodchild MF (1984) ILACS: A location-allocation model for retail site selection. Journal of Retailing, 60(1), 84-100.
19. Malczewski J (2004) GIS-based land-use suitability analysis: a critical overview. Progress in Planning, 62(1), 3-65. https://doi.org/10.1016/j.progress.2003.09.002.
20. Scott D, McBoyle G (2007) Climate change adaptation in the ski industry. Mitigation and Adaptation Strategies for Global Change, 12(8), 1411-1431. https://doi.org/10.1007/s11027-006-9071-4.
21. Kim CM, Chung IM (2023) The structural relationship between local residents' perception of tourism impact on park golf course development, satisfaction, and supporting intention. Journal of Golf Studies, 17(3), 203-212.
22. King KW, Balogh JC, Harmel RD (2007) Nutrient flux in storm water runoff and baseflow from managed turf. Environmental Pollution, 150(3), 321-328. https://doi.org/10.1016/j.envpol.2007.01.038.
23. Maroko, A. R., Maantay, J. A., Sohler, N. L., Grady, K. L., & Arno, P. S. (2009). The complexities of measuring access to parks and physical activity sites in New York City: a quantitative and qualitative approach. International Journal of Health Geographics,8(1), 34. https://doi.org/10.1186/1476-072X-8-34
24. Jordan TE, Weller DE, Pelc CE (2018) Effects of local watershed land use on water quality in mid-Atlantic coastal bays and subestuaries of the Chesapeake Bay. Estuaries and Coasts, 41 (Suppl 1), 38-53. https://doi.org/10.1007/s12237-017-0303-5.
25. Fuchs MD, Gebler S, Lorke A (2025) Landscape-level assessment of spray drift–A virtual experiment using the Droplet and Atmospheric Dispersion drift (DAD-drift) model. Environmental Modelling and Software,189, 106455. https://doi.org/10.1016/j.envsoft.2025.106455.
26. South A, Michael S, Massicotte P (2025). rnaturalearthhires: High resolution world vector map data from natural earth used in 'rnaturalearth'. R package version 1.0.0.9000
27. Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, Wevers J, Grosu A, Paccini A, et al. (2021) ESA WorldCover 10 m 2020 v100
28. Tu X, Huang G, Wu J, Guo X (2020) How do travel distance and park size influence urban park visits? Urban Forestry and Urban Greening, 52, 126689. https://doi.org/10.1016/j.ufug.2020.126689.