Estimating the Economic Value of Grassland Tourism Services Based on Mobile Phone Data

Na Liu
China Center for Agricultural Policy, School of Advanced Agricultural Sciences
Peking University, Beijing 100871, China
Email: xeo_liu@163.com

Xinxin Lv
China Center for Agricultural Policy, School of Advanced Agricultural Sciences
Peking University, Beijing 100871, China
Email: 2306196041@pke.edu.cn

Pengfei Liu
Associate Professor
Department of Environmental and Natural Resource Economics
University of Rhode Island, Rhode Island 02881-2020, USA
Email: pengfei_liu@uri.edu

Lingling Hou*
Associate Professor
China Center for Agricultural Policy, School of Advanced Agricultural Sciences
Peking University
No. 5 Yiheyuan Road, Haidian District, Beijing 100871, China
Email: llhou.ccap@pku.edu.cn
*Correspondence author

doi:10.3368/le.102.3.012726-0016

Abstract

This study evaluates the economic value of grassland tourism in Inner Mongolia, China, using mobile phone signaling data to track tourist flows from 333 cities to 103 counties in 2020. A zonal travel cost model, both parametric and non-parametric, was employed. Results indicate the tourism value ranged from 585.9 to 598.4 billion CNY (US $84.9–86.7 billion), about 2.5 times the tourism revenue, constituting 29% of regional GDP. In the absence of COVID-19, the value would have reached 936.5 billion CNY. The study reveals spatial disparities, with grassland quality, infrastructure, and regional development driving tourism value.

Estimating the Economic Value of Grassland Tourism Services Based on Mobile Phone Data

1. Introduction

Assessing the economic value of ecosystems at scale is essential for informing conservation investments and sustainable land management. Ecosystem services, and the natural capital that underpins them, are integral to Earth's life-support systems and contribute substantially to human well-being, representing a significant share of global economic value (Costanza et al., 1997; Foley et al., 2005). These services are rarely reflected in market transactions and systematically undervalued in policy processes, creating risks for long-term sustainability (Costanza et al., 1997; Taye et al., 2021; Börger et al., 2023). Several studies have attempted to estimate the global or regional economic value of ecosystem services (Costanza et al., 1997; Costanza et al., 2014; Xie et al., 2017; Taye et al., 2021), but most rely on coarse or approximate parameters, which can lead to biased estimates. For instance, Costanza et al. (2014) used a basic benefit transfer approach assuming constant per-hectare values. Such limitations highlight the need for more precise and context-specific parameters to improve the accuracy of global and regional ecosystem service valuation.

Tourism constitutes a major component of ecosystem economic value, classified as a use value of ecosystem services (Peng et al., 2014; Emiru, 2017). Tourism reflects the tangible economic benefits generated by related resources and environments. In welfare economics, use value is typically measured by the willingness of visitors to pay for services that satisfy their needs (Peng et al., 2014). Accurate assessment of this value is essential for policymakers when designing conservation and land-use policies. Beyond direct revenues, tourism also contributes more broadly to human welfare and regional economic development (Zhang et al., 2006). Welfare theory suggests that consumer surplus, derived from the tourism demand curve, should be added to tourism revenues to capture the full value of tourism. Accordingly, studies that rely solely on revenue measures are likely to underestimate tourism's true economic importance.

The travel cost method (TCM) is a well-established tool for estimating the recreational and tourism value of ecosystems (Freeman III et al., 2014; Champ et al., 2017), yet its application to large-scale grassland systems remains limited. Most empirical studies have focused on coastal, forest, or wetland ecosystems. For example, Parsons et al. (2013) estimated that a 75% reduction in Delaware beach width would lower visitor welfare by approximately $5 per day, while Houngbeme et al. (2021) assessed the recreational value of Fidjrosse Beach in Benin. In forests, Getzner and Meyerhoff (2020) valued annual hiking and wildlife benefits at roughly €500 per person, and Liu et al. (2021) estimated recreation value in Taiwan's Alishan National Forest at 1,703 New Taiwan Dollars per visitor. Wetland-focused studies include Xu and He (2022) and Gimenez and Mas (2020). By contrast, applications to grasslands are rare and largely confined to small-scale case studies, such as Lv et al. (2006) on the Siziwang Banner grasslands in Inner Mongolia of China, and Wang et al. (2022) on visitor use of Tibetan Plateau grasslands. This limited evidence underscores a critical research gap due to the absence of large-scale assessments of grassland tourism value, despite their ecological and economic importance.

Applying the travel cost method (TCM) to large-scale grassland ecosystems poses significant challenges because most studies rely on visitor surveys with inherent limitations. For example, Parsons et al. (2021) surveyed visitors at seven U.S. national park entrances, but such data typically capture tourist information only on specific days, overlooking temporal variation in travel demand (Pascoe, 2019). Surveys also tend to focus on individual scenic sites, thereby underestimating the broader regional tourism value of ecosystems. For grasslands, which cover vast areas but often attract relatively few visitors, surveys are logistically difficult and prohibitively costly.

Long-term monitoring data, such as mobile phone signals, provide a promising alternative for large-scale tourism valuation by offering broad coverage and greater reliability (Loomis et al., 2000). Owing to their large volume, extensive spatial coverage, and high locational accuracy, mobile phone data are increasingly used to study individual travel behaviors (Goodchild, 2013; Chen & Zhang, 2014; Salganik, 2019). Compared with survey or official statistical data, they yield richer insights into accessibility of scenic sites (Lin et al., 2021), overnight trips (Nyns & Schmitz, 2022), points of interest (Liang et al., 2022), and other tourism activities (Raun et al., 2016). Recent research has further demonstrated the potential of integrating mobile phone data into ecosystem service valuation. For example, Kubo et al. (2020) combined mobile phone data with the TCM to estimate welfare impacts of climate change on Japanese beach tourism, while Jaung and Carrasco (2020) used mobile phone signals to assess recreational ecosystem services in Singapore's parks and protected areas.

We use Inner Mongolia as a case study to estimate the economic value of grassland tourism by applying both parametric and non-parametric travel cost methods to mobile phone signaling data. Specifically, we constructed a comprehensive dataset tracking weekly tourist flows in 2020 from all 333 Chinese cities to all 103 counties in Inner Mongolia. The dataset records the number of tourists traveling from each city to each county, along with the distribution of transportation modes, gender, and age groups. Using these data, we estimated tourism demand functions and calculated the tourism value of Inner Mongolia's grasslands in 2020. We also examined spatial variations in tourism value across the region and assessed the contributions of different parts of the country to Inner Mongolia's tourism economy. Finally, by analyzing the determinants of grassland tourism value, we identify the key drivers and constraints shaping tourism development in the region.

Our study yields three key findings on the tourism value of Inner Mongolia's grasslands. First, the total tourism value was estimated at 585.9 billion CNY (U.S. $84.9 billion) using the parametric method and 598.4 billion CNY (U.S. $86.7 billion) using the non-parametric method. In the absence of COVID-19, the predicted tourism value would have reached 936.5 billion CNY (U.S. $135.7 billion). These results highlight both the resilience of grassland tourism and the significant disruptions caused by the pandemic.

Second, the spatial distribution of tourism value across Inner Mongolia and its origin cities exhibits substantial heterogeneity. Southern cities, such as Hohhot and Ordos, generate higher tourism values than northern cities, likely reflecting more developed transportation infrastructure and tourism facilities. Similarly, eastern cities, including Hulunbuir, record higher tourism values than western cities, largely due to dominant vegetation types—such as temperate meadow steppes—that enhance scenic quality. Moreover, contributions to Inner Mongolia's overall tourism value decline with distance: cities within the region account for 43% of total tourism value, underscoring the importance of geographic proximity for tourism demand.

Third, grassland quality, local infrastructure, and regional economic development are all significantly correlated with both visitor numbers and tourism revenue. Grassland quality, measured by the Normalized Difference Vegetation Index (NDVI), demonstrates the strongest correlation across all model specifications. Causal analysis further reveals a progressively positive effect of grassland quality on tourism outcomes, with impacts increasing from medium to high NDVI levels. High-grade scenic areas (4A and 5A) and well-developed accommodation facilities are consistently correlated with higher visitor numbers and tourism value, underscoring their contribution to destination attractiveness. Transportation infrastructure, particularly the presence of high-speed rail stations, exhibits a strong correlation with visitor numbers. Regional economic development, as measured by GDP and population, is positively correlated with tourism value, although its association with visitor numbers varies across specifications.

This study makes two primary contributions to the literature. First, this paper uses cellphone signal data with the travel cost method to estimate ecosystem values. While the travel cost method has been widely used with survey data (Parsons et al., 2013; Getzner & Meyerhoff, 2020), survey approaches are limited by small-scale coverage and temporal constraints (Pascoe, 2019). Recent studies have begun incorporating cellphone data for broader insights (Kubo et al., 2020; Jaung & Carrasco, 2020), but these remain focused on small-scale or specific regional ecosystems. This paper extends this approach to a large-scale grassland ecosystem, providing more comprehensive value estimation.

Second, this paper complements the existing literature by focusing on grassland ecosystems.

Previous travel cost studies have predominantly examined coastal, forest, and wetland ecosystems (Parsons et al., 2013; Houngbeme et al., 2021; Getzner and Meyerhoff, 2020; Liu et al., 2021; Xu and He, 2022; Gimenez and Mas, 2020), while grassland studies remain limited in scale (Lv et al., 2006; Wang et al., 2022). This study assesses Inner Mongolia's grasslands—one of the world's largest grassland ecosystems—using both parametric and non-parametric methods to provide more accurate tourism value estimation.

The remainder of the paper is structured as follows. Section 2 describes the study area, data sources, key variables, and provides descriptive statistics. Section 3 outlines the zonal travel cost method and the estimation of the recreation demand function. Section 4 presents the empirical results of tourism value, and Section 5 discusses the findings and their broader implications.

2. Background and data

Study area: Inner Mongolia in China

China's grasslands cover a total of 265 million hectares, representing 27.6% of the national territory and making them the country's second-largest terrestrial ecosystem after forests (Announcement of the Main Data from the Third National Land Survey, 2021). The six provinces with the largest grassland areas—Tibet, Inner Mongolia, Xinjiang, Qinghai, Gansu, and Sichuan—together account for 94% of the national total. These vast ecosystems support over 15,000 plant species and more than 2,000 wild animal species (Ding et al., 2020) and are home to approximately 70% of China's ethnic minority population, who maintain vibrant nomadic cultures. This combination of ecological richness and cultural heritage provides substantial potential for tourism development.

Despite this potential, grassland tourism in China remains underdeveloped due to limited infrastructure, insufficient investment, and other constraints (Li et al., 2021). In 2021, the combined number of tourists visiting Tibet, Inner Mongolia, Xinjiang, Qinghai, and Gansu reached 837 million—roughly equivalent to the number visiting Shanxi Province, which ranks sixth nationwide (China Statistical Yearbook, 2022). These figures underscore the need to explore sustainable development models and promote the grassland tourism industry to support the long-term development of China's grasslands.

Inner Mongolia, located in northern China, is the country's largest pastoral region and the second-largest grassland province, with abundant grassland resources (Figure 1). Its grasslands account for 20.06% of China's total grassland area (China Statistical Yearbook, 2022). Natural grasslands in the region cover 79 million hectares, representing 66.7% of the autonomous region's total land area (China Statistical Yearbook, 2020). Administratively, Inner Mongolia is divided into eight grassland sub-regions and features three main grassland types: temperate meadow grasslands, temperate typical grasslands, and temperate desert grasslands. Summers are generally cool, and variations in temperature, precipitation, and seasonality create diverse grassland ecosystems across the region.

In recent years, Inner Mongolia's tourism industry has experienced rapid growth. In 2011, the region received 51.8 million domestic tourists, generating 84.7 billion RMB in tourism revenue. By 2019, these figures had increased dramatically to 193.2 million domestic tourists and 455.9 billion RMB in revenue—nearly a fourfold increase in visitors and a fivefold increase in revenue compared with 2011. This growth underscores Inner Mongolia's rising importance as a national tourism destination and highlights the substantial potential for further development of grassland-based ecotourism.

[Insert Figure 1 Here]

Mobile phone signal data

We constructed a comprehensive weekly departure-to-destination dataset using mobile phone signaling data. This dataset records weekly tourist flows from all 333 Chinese cities to 103 counties in Inner Mongolia between May and October 2020, including information on transportation modes. Mobile phone data enable identification of travelers' residence, destination, and transportation mode based on the locations of connected base stations. Specifically, when a user's phone is turned on, it continuously scans for nearby base stations and connects to the one with the strongest signal. The base stations log signal strength and location data for network optimization and location services (Calabrese et al., 2014; Yang et al., 2021). This process allows for real-time tracking of users when they enter a base station's coverage area, which typically extends about 500 meters (Ratti et al., 2006; Xu et al., 2017).

We define tourists as individuals who temporarily leave their usual residence or workplace to engage in sightseeing, leisure, vacation, or other tourism-related activities at domestic destinations. This inclusive definition encompasses travel for family visits, health and recuperation, field investigations, conference participation, and tourism components embedded within business, scientific, cultural, educational, or religious trips.

We obtained mobile phone data from China Unicom, one of China's three major operators. In 2021, China Unicom had 319 million users, representing approximately 20% of the national market, and operated 2.17 million base stations covering all cities in China (Communication Industry Statistical Report, 2021). Due to confidentiality and privacy regulations, direct access to individual-level raw data is not possible. We established a collaborative framework with China Unicom's technical team to obtain aggregated mobile signaling data that satisfied our research needs while ensuring compliance with data privacy standards.

To identify tourists, we applied a multi-stage filtering procedure to ensure the accuracy and relevance of the sample. First, we conducted temporal and spatial screening by extracting mobile signaling records for users whose devices appeared in Inner Mongolia between May 1 and October 31, 2020 (hereafter the study period). Second, we defined tourism activity as travel more than 10 kilometers from an individual's primary residence for at least six hours, consistent with the official definition of the Ministry of Culture and Tourism of China. For these travelers, we determined their origin cities and destination counties. Third, we excluded non-tourist trips by removing cases where the destination county coincided with the individual's registered residence or workplace, thereby eliminating routine commuting and business travel. Fourth, we conducted scenic area verification using China Unicom's scenic area base station interface data, including only users who connected to base stations within designated tourist attractions.

Through this systematic filtering process, we obtained the final tourist sample, providing weekly tourist flow data from origin cities to destination counties for China Unicom subscribers during the study period. The original signaling records allow tracking of tourists' movements across multiple counties. For instance, if a traveler from Guangzhou first visits County *A* and then proceeds to County *B*, the dataset records two distinct flows: one from Guangzhou to County *A* and another from County A to County B. Each flow is treated as an independent observation and is used to calculate the corresponding tourism value.

Transportation modes were assigned using a spatial intersection analysis between destination and transportation infrastructure base stations. We classified travel into four modes: airplane, train, bus, and private vehicle. Specifically, when a mobile phone connected to a base station that spatially intersected with an airport, railway station, or bus station, the corresponding transportation mode was assigned. If no intersection occurred with these transportation hubs, the travel mode was classified as a private vehicle. Given that urban base stations typically have a service radius of approximately 500 meters and that transportation terminals are generally located at substantial distances from highways, this approach provides high accuracy in identifying travelers' transportation modes.

Key variables and data sources

Travel cost refers to the expenses incurred by a representative individual from the place of residence to the destination, comprising transportation cost and time cost. Following Jaung & Carrasco (2020), the standard formula is:

$$TC_{ij} = \sum_{N=1}^{n} w_{ijn}\, cost_{ijn} + \sum_{N=1}^{n} w_{ijn}\, Time_{ijn} \times \frac{1}{3} wage_i \qquad [1]$$

where TC_{ij} represents the travel cost from residence i to destination j, $cost_{ijn}$ is the cost of transportation mode n that a tourist chooses for travel from residence i to destination j (i.e., ticket price). w_{ijn} is the proportion of tourists choosing transportation mode n for travel from residence i to destination j. $Time_{ijn}$ is the travel time (in hours) from residence i to destination j using transportation mode n, and $wage_i$ is the average hourly wage at residence i. As the value of leisure time is often lower than the value of working time, one-third of the average salary is used as the value of leisure time, following previous literature (Armbrecht, 2014; Jaung & Carrasco, 2020). For robustness checks, we conduct sensitivity analysis using alternative values of 0, 1/2, 3/4s, and 1 time the average wage, with results presented in Appendix A, Table A3.

To estimate the travel costs, we compile a database that includes transportation expenses and travel time associated with different transportation modes. Firstly, we obtained flight data from VariFlight, a flight information querying app (https://flightadsb.vari-flight.com/). Flight data, including full fare and flight time for economy class, was collected for all connections from the origin to the destination cities. If the original city only has one airport, we assume all travelers use that airport. If there are multiple airports, an equal probability of selection is considered. The average cost and time of all flights from all airports in the original city to the destination were computed as the monetary expenditure and duration for air travelers. If the original city does not have an airport, we assume tourists choose the nearest airport to the municipal government. We calculated the average cost and time of all flights from the closest airport to the destination and the cost and time of driving from the government office to the nearest airport. These costs and times represent the monetary expenditure and duration for tourists opting for air travel in the city. Similarly, the described procedure is applied to destination airports. Ultimately, each route from origin to destination has a corresponding monetary expenditure and flight time associated with the travel.

Secondly, we collect train fares and travel time from the official website of China Railway's ticketing service (https://www.12306.cn/index/). Trains are categorized into high-speed and standard trains. High-speed trains generally have 85% second-class and 15% first-class seats, while standard trains have 70% seats and 30% beds. We use train fare and travel time of second-class seats or seats during calculation. In cases where there is no direct train from the origin to the destination, travelers can opt for train transfer routes with the shortest duration. Additionally, the time interval between each transfer is assumed to be one hour based on the operations of connection trains. Travelers choose train stations similarly to airports. If there is more than one train station in the city, tourists choose one of the stations with equal probability. When there is no train station in the city, visitors drive or take a bus to the nearest train station to the city's government. The total duration and monetary expenditure encompass the sum of driving and taking the train.

Thirdly, bus fare and travel time information are collected from the bus fare website (https://www.piaojia.cn/changtu/). In cases where there is no direct bus for each origin-destination route, the transfer route is chosen based on the shortest time, and the interval between each transfer is assumed to be one hour. City visitors with multiple bus stations select one as the departure station with equal probability. In cities without bus stations, we believe visitors choose the nearest bus station to the city's government.

Lastly, car costs are computed based on travel distances and tolls. The driving distance (km) from the residence to the destination and the toll are obtained from Amap (https://gaode.com/, similar to Google Maps). This navigation software plans routes from origin to destination based on driver preferences. We assume drivers prefer to choose the route that minimizes the travel time, and Amap provides information on the route's duration, distance, and tolls. According to data the Ministry of Industry and Information Technology disclosed, we collected the average fuel consumption per kilometer for passenger cars in China in 2020 and the average gasoline prices in each province from May to October 2020. The one-way monetary expenditure is calculated as "distance × average fuel consumption of passenger vehicles × average gasoline price + toll fees".

Time cost refers to the opportunity cost of the time spent traveling from the point of departure to the destination, primarily composed of the duration of the journey and the price per unit of time for the traveler. Travel decisions are typically constrained by predetermined budgets for money and time, which reflect the long-term trade-offs individuals make between labor and leisure (Lupi et al., 2020). Consequently, the opportunity cost of travel time is a function of income. Following the approach of Armbrecht (2014) and Jaung & Carrasco (2020), we use one-third of the hourly wage rate in the tourist's province of residence as the price per unit of travel time. The hourly wage rate varies significantly among different provinces. China Statistical Yearbook (2021) provides employment and wage data across provinces, distinguishing between non-private and private sectors. We calculate sector employment proportions and use them as weights to determine average salaries. These annual figures are converted to weekly rates and, considering China's average weekly working hours (46.9, China Statistical Yearbook, 2021), we derive hourly wage rates for each province using the weekly wage rate divided by 46.9. Existing literature also suggests that the cost of on-site time can be ignored when valuing recreational trips because on-site time is considered a complementary good (English, 2020; Fezzi et al., 2014).

The visitor rate is another key variable, calculated using the following formula for each destination *j*:

$$VR_i^j = \frac{V_i^j}{N_i}, \quad \text{for each j} = 1 \ldots \text{J} \qquad [2]$$

where VR_i^j represents the visitor rate from residence i to destination j, V_i^j is the number of visitors from residence i to destination *j*; Our data came from China Unicom's tourism visitor numbers between May and October 2020. We adjusted these figures using city-specific market share data from China Unicom to account for regional differences in operator market penetration. To annualize the 6-month data, we further divided by 0.5353, reflecting the proportion of annual tourism activity during the peak season from May to October (53.53%, Statistical Bulletin on the Development of Culture and Tourism in Inner Mongolia, 2021). N_i represents the population of residence *i*, derived from the data of the seventh national population census in 2020. Generally, users face no external constraints when selecting mobile service providers, and there is no evidence of significant differences in the user bases of the three major carriers. Since December 2019, China's mobile number portability policy has allowed users to switch carriers freely without changing phone numbers. Therefore, we assume that users of different providers do not exhibit significantly different traveling patterns.

Descriptive analysis

We depict the distribution of tourists in different regions and examine visitors' preferences for various modes of transportation. The number of visitors varies among counties in Inner Mongolia, as illustrated in Figure A1.[1] This variation indicates that eastern counties attract more tourists than their western counterparts. The east region of Inner Mongolia is characterized by temperate meadow steppes, which offer diverse natural landscapes with high scenic value. Meanwhile, southern counties receive more visitors than northern counties, possibly due to the more developed transportation infrastructure and better overall facilities in cities closer to the inland areas. In 2020, the county with the highest number of tourists in Inner Mongolia was Saihan District, with a total of 7.31 million visitors, while the county with the fewest tourists was Baiyun Ebo Mining District, with fewer than 51 thousand visitors.

Figure A2 illustrates the number of tourists visiting Inner Mongolia from each city across the country.[2] The color intensity reflects the volume of visitors originating from each city. Notably, the number of visitors to Inner Mongolia generally decreases as the distance from the city to Inner Mongolia increases. In 2020, Hohhot, the regional capital of Inner Mongolia, emerged as the primary source market, contributing 10.57 million visitors. In contrast, Sansha City in Hainan Province accounted for the smallest number of visitors, with only 142 recorded visitors.

3. Zonal travel cost method and recreation demand function

We employ the zonal travel cost method, better suited to align with mobile phone signal data. Given the restrictions imposed by the Chinese government and China Unicom on data collection, mobile phone data in this study could only be accessed at an aggregate level.

Zonal travel cost function

The zonal travel cost function for each destination j is represented by the following equation:

$$VR_i^j = f\left(TC_i^j\right), \quad \text{for each j} = 1 \ldots \text{J} \qquad [3]$$

where VR_i^j is the visitor rate from origin i to destination j, and TC_i^j is the travel cost from origin i to destination j.

We employ both parametric and non-parametric methods to estimate the travel cost function. Existing research has mostly relied on parametric estimation methods, such as log-linear forms and inverse proportion function forms, to model the travel cost function (Dai et al., 2022; Jaung & Carrasco, 2020). Parametric methods may be prone to estimation errors due to potential model misspecification. Non-parametric estimation does not rely on predefined assumptions regarding the model's functional form; thus, it may reduce errors stemming from incorrect distributional assumptions. By using both approaches, we aim to enhance the robustness and accuracy of our travel cost function estimates.

We use standard model specifications for the parametric methods with linear, log-linear, linear-log, and log-log forms. The resulting demand function represents the recreation demand function for each destination. For each destination j, the demand function equations for travel costs and visitor rates are:

$$Linear: VR_i = \gamma + \delta TC_i \quad [4]$$

$$Log - linear: \ln(VR_i) = \gamma + \delta TC_i \quad [5]$$

$$Linear - log: VR_i = \gamma + \delta\, ln(TC_i) \quad [6]$$

$$Log - log: \ln(VR_i) = \gamma + \delta\, ln(TC_i) \quad [7]$$

where VR_i represents the visitor rate from each city i to a given destination county in Inner Mongolia, and TC_i represents the corresponding travel costs. The equation adopts the standard linear-log form (McAleer, 1994; Ong, 1995).

We compare the performance of different functional forms based on statistical criteria. Appendix A, Table A1, columns (1)-(4) report the estimation results for the four equations.

Following the approach of Jaung & Carrasco (2020) and Dai et al. (2022), we used adjusted R^2 and AIC to compare the relative fit of the models. The results show that the Linear-Log model is the most suitable for estimation.

The demand specification is as follows:

$$VR_{ijr} = \gamma + \delta\, ln\left(TC_{ij}\right) + Region_r + \varepsilon_{ij} \qquad [8]$$

where VR_{ij} represents the visitor rate from each city *i* to destination county *j* in Inner Mongolia, TC_{ij} represents the corresponding travel costs. The departure cities' provinces are classified into seven geographic regions: East China (Shandong, Jiangsu, Anhui, Zhejiang, Fujian, and Shanghai); South China (Guangdong, Guangxi, and Hainan); Central China (Hubei, Hunan, Henan, and Jiangxi); North China (Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia); Northwest China (Ningxia, Xinjiang, Qinghai, Shaanxi, and Gansu); Southwest China (Sichuan, Yunnan, Guizhou, Tibet, and Chongqing); and Northeast China (Liaoning, Jilin, and Heilongjiang). $Region_r$ controls for regional fixed effects corresponding to the origin city *j*'s location, and ε_{ij} is the error term. The regression results are presented in Appendix A, Table A1.

Regarding the issue of spatial autocorrelation, we first tested for spatial dependence in the model residuals using the global Moran's I statistic. The test yielded a coefficient of 0.0396, indicating modest but statistically significant positive spatial dependence among neighboring districts. To further address potential correlation in error terms among tourists from the same origin region, we re-estimated our specifications with standard errors clustered at the origin-province level, thereby accounting for shared unobserved preferences and similar travel constraints within regions. These spatial adjustments confirm the robustness of our baseline results.

Non-parametric methods include a range of techniques such as non-parametric kernel regression, multivariate kernel regression, k-nearest neighbor regression (KNN), and local linear regression, among others. The KNN regression method does not impose strict requirements on the data distribution and exhibits low sensitivity to outliers, demonstrating strong adaptability to different data structures (Altman, 1992; Kamarol et al., 2017). We apply the KNN regression as a non-parametric approach to estimate the recreation demand function, leveraging the target values of the k nearest neighbors to a given query point (Kamel et al., 2008). Specifically, the KNN method calculates the average visitor rate of the k nearest neighbors, where k is a predefined parameter determining the number of neighbors considered in the estimation.[3] The KNN estimate for the visitor rate (VR) at a given travel cost level TC_0 is expressed as:

$$\widehat{VR}(TC_0) = \frac{1}{k}\sum_{i=1}^{n} 1\{TC_i \epsilon N_k(TC_0)\} \cdot VR_i \qquad [9]$$

where $\widehat{VR}(TC_0)$ represents the predicted visitor rate at the target travel cost TC_0. The term $1\{TC_i \epsilon N_k(TC_0)\}$ is an indicator function that identifies whether the travel cost TC_i belongs to the set of *k* nearest neighbors of TC_0, denoted as $N_k(TC_0)$. For example, if *k* = 5 and TC_0 = 100, the function $1\{TC_i \epsilon N_k(TC_0)\}$ will take the value 1 for the 5 observations with travel costs closest to 100, and 0 for all other observations. The visitor rate VR_i corresponds to the observed visitor rate from the *i*th city to the destination.

Consumer surplus and tourism value

Based on the recreation demand function, we use the following equation to calculate the consumer surplus:

$$CS_{ij} = N_i \times \int_{TC_{ij}^0}^{TC_{ij}^m} f(TC_{ij}) dTC_{ij} \quad [10]$$

where CS_{ij} represents the consumer surplus from origin i to destination j, N_i denotes the population of origin i, TC_{ij}^0 is the actual travel cost from origin i to destination j, TC_{ij}^m is the maximum travel cost from origin i to destination j as the marginal effect approaches zero. The function $f(.)$ is a function of travel cost and the visitor rate for each destination j. In the calculation process, if the destination j is the same, the demand function from origin i to destination j is identical for different origins i and the same function $f(.)$ is used.

For each destination j, the tourism value can be obtained from the sum of the consumer surplus and travel expenses across all origin cities i, as shown in formula (11):

$$REV_j = \sum_{i=1}^{I} CS_{ij} + \sum_{i=1}^{I} TC_{ij} \times V_{ij} + \varepsilon_j \quad [11]$$

where REV_j is the annual tourism value of destination j, CS_{ij} is consumer surplus from origin i to destination j, TC_{ij} represents the travel cost from origin i to destination j, V_{ij} is the number of visitors from residence i to destination j.

Furthermore, we calculated the annual tourism value of grasslands for the whole Inner Mongolia:

$$REV = \sum_{j=1}^{J} REV_j \quad [12]$$

where REV is the annual tourism value of the whole Inner Mongolia, REV_j is the annual tourism value of destination j.

Influencing factors

After estimating the tourism value for each county, we further examine the factors influencing both tourist numbers and tourism value.

The formula below is employed to regress county features on tourism value and tourist numbers:

$$\ln(Y_{ij}) = \beta_0 + \beta_1 NDVI_j + \beta_2 \ln(GDP_j) + \beta_3 SH_j + \beta_4 SL_j + \beta_5 HH_j + \beta_6 HL_j + \beta_7 \ln(Area_j) + \beta_8 Types_j + \beta_9 Pop_j + \beta_{10} Rail_j + City_i + \varepsilon_{ij} \quad [13]$$

where Y_{ij} represents the number of tourists or the tourism value from origin city *i* to destination county *j* in Inner Mongolia in 2020; $NDVI_j$ denotes the maximum grassland NDVI value in destination county *j* in 2020. GDP_j represents the GDP in destination county *j*. SH_j indicates the number of 4A and 5A scenic spots in destination county *j*. SL_j indicates the number of 2A and 3A scenic spots in destination county *j*. HH_j and HL_jrepresents the number of three-star and above hotels and three-star and below hotels in destination county *j*, respectively. $Area_j$ indicates the area of main grassland types suitable for grassland tourism in the destination county *j*. $Types_j$ suitable for grassland tourism in destination county *j*. Pop_j denotes the permanent resident population (10 thousand people) in destination county *j*. $Railway_j$ is a binary variable indicating whether destination county *j* has a high-speed rail station (1=Yes, 0=No). $City_i$ controls for the origin city fixed effects. ε_{ij} represents the error term.

Table 1 presents descriptive statistics for all variables included in our analysis. Maximum grassland NDVI data are sourced from the "China Regional 250m Normalized Difference Vegetation Index Dataset (2000-2024)" from the National Tibetan Plateau Scientific Data Center (https://data.tpdc.ac.cn/home). [4] This dataset provides vegetation coverage at grid scale and includes boundary information for vegetation types (grassland, forest, etc.). Using 2000 grassland boundaries, we extracted grassland-specific NDVI data at grid scale, ensuring the index reflects only grassland vegetation, excluding other vegetation types. This variable represents natural resources. Additionally, grassland type area covers lowland meadow, alpine meadow, mountain meadow, temperate meadow steppe, and temperate steppe suitable for tourism, obtained from the National Ecosystem Science Data Center (https://www.nesdc.org.cn/).

To represent economic development, GDP and population data from the 2020 Inner Mongolia Statistical Yearbook are utilized (Inner Mongolia Autonomous Region Bureau of Statistics, 2021). Data on scenic spots and hotels, sourced from the Chinese Ministry of Culture and Tourism (2020), serve as indicators of tourism infrastructure. Information on high-speed rail stations, sourced from Amap (https://gaode.com/), is used to measure transportation accessibility.

The preceding analysis is restricted to correlations. Given that the dataset spans only a single year and key covariates, such as infrastructure, are observed exclusively at the annual frequency, identifying causal effects from annual aggregates is not feasible. To alleviate this constraint, we partition the data into three intra-annual periods, thereby enabling a more credible assessment of the causal relationship with NDVI. Specifically, we divide the months from May to October into three periods. Time 1 (May–June), Time 2 (July–August) and Time 3 (September–October).

The formula is specified as follows:

$$\ln\left(Y_{ijt}\right) = \beta_0 + \beta_1 NDVI\_M_{jt} + \beta_2 NDVI\_H_{jt} + City_i + County_j + \varepsilon_{ijt} \qquad [14]$$

where Y_{ijt} represents the number of tourists or the tourism value from origin city *i* to destination county j in Inner Mongolia at time *t*; $NDVI\_M_{jt}$ and $NDVI\_H_{jt}$ are dummy variables representing the medium and high tertiles of the maximum grassland NDVI value in destination county j at time t, respectively, with the low tertile serving as the reference category; $City_i$ controls for the origin city fixed effects; $County_j$ controls for the destination county fixed effects; ε_{ij} represents the error term.

[Insert Table 1 Here]

4. Results

The tourism value of grasslands

Figure 2 illustrates the spatial distribution of tourism values across cities in Inner Mongolia in 2020. The corresponding numerical estimates are reported in Table A2 in Appendix A. Using the parametric method, the tourism value of Inner Mongolia is estimated at 585.9 billion CNY (U.S. $84.9 billion), about 2.4 times its tourism revenue (240 billion CNY) and accounting for 28.6% of the region's GDP. The non-parametric method yields a comparable estimate of 598.4 billion CNY (U.S. $86.7 billion), which is 2.5 times the tourism revenue and equivalent to 29.4% of regional GDP.

To assess the reliability of our estimates, we conduct several robustness checks. First, we test alternative control variable specifications, including demographic controls for age and gender. Moreover, we estimate the demand curve separately for departure cities grouped by per capita GDP levels. Second, we compare estimates using annual versus time-period data. The results, presented in Table 2, demonstrate that our findings are consistent across these different specifications, confirming the robustness of the estimated tourism values.

[Insert Table 2 Here]

Figure 2 highlights substantial inter-city disparities in tourism value. In general, southern cities exhibit higher values than their northern counterparts. For example, Hohhot and Ordos, both located in the south, rank first and second, respectively. The tourism value of Hohhot is estimated at 90.7–102.4 billion CNY (U.S. $13.1–14.8 billion), representing 29–33% of local GDP, while Ordos is estimated at 66.0–73.3 billion CNY (U.S. $9.6–10.6 billion), or 14–15.5% of GDP. By contrast, northern cities report markedly lower values. Bayannur, for instance, ranks near the bottom, with an estimated tourism value of 27.3–31.2 billion CNY (U.S. $4.0–4.5 billion), equivalent to 28–32% of its local GDP. This north–south disparity likely reflects the relatively more developed transportation infrastructure and tourism facilities in southern cities, which are closer to major inland population centers and thus more accessible to visitors.

Cities in the eastern region generally exhibit higher grassland tourism values than those in the west. For instance, Hulunbuir, located in the east, is estimated at 50.9–54.2 billion CNY (U.S. $7.4–7.9 billion), accounting for 38–40% of local GDP. By contrast, Wuhai in the west records a substantially lower value of 16.0–17.3 billion CNY (U.S. $2.3–2.5 billion), or 22–24% of GDP. This east–west disparity is consistent with the distribution of tourist flows and can largely be attributed to the dominant vegetation type in the east—temperate meadow steppes—which provide more diverse landscapes and greater scenic value, thereby attracting larger numbers of visitors.

[Insert Figure 2 Here]

To further assess the contributions of origin cities across China to Inner Mongolia’s tourism value, Figure A3 presents the spatial distribution of tourism value by origin city in 2020. The results indicate that a substantial share of Inner Mongolia’s grassland tourism value originates from cities within the region itself, accounting for approximately 19.4% of the total. Among these, Hohhot emerges as the largest contributor, with an estimated value of 21.3–43.9 billion CNY (U.S. $3.1–6.4 billion), equivalent to roughly 3.6–7.3% of Inner Mongolia's total tourism value.

Contributions to Inner Mongolia's tourism value generally decline with increasing distance from the region, likely reflecting differences in transportation accessibility, as proximate cities enable more frequent and convenient travel. In addition, cities in the southeastern coastal areas contribute more than those in central China, a pattern likely driven by their higher GDP levels and greater capacity for tourism-related expenditures.

Estimating tourism value in the absence of COVID-19

The outbreak of COVID-19 in early 2020 led to extensive domestic and international travel restrictions, resulting in a sharp decline in tourist numbers. Consequently, the tourism value of Inner Mongolia's grasslands in 2020 may be underestimated. In 2019, Inner Mongolia received 193.2 million visitors, generating tourism revenue of 455.9 billion CNY (U.S. $66.1 billion). In contrast, tourist arrivals in 2020 fell to 124.9 million, or 64.7% of the 2019 level, and tourism revenue declined to 240.4 billion CNY (U.S. $18.1 billion). To provide a more accurate assessment, we further analyze the tourism value of Inner Mongolia's grasslands by estimating a counterfactual scenario that excludes the impact of the COVID-19 pandemic.

Figure 3 presents the projected number of tourists and tourism revenue in Inner Mongolia in 2020, excluding the impact of COVID-19. Using data from 2000 to 2019 and assuming that trends in 2020 would have followed the historical pattern, we applied a polynomial trend extrapolation. Specifically, a third-order polynomial function was fitted to both tourist numbers and tourism revenue. The resulting functions achieved high goodness-of-fit, with R^2 values of 0.97 and 0.99 for tourist numbers and tourism revenue, respectively, indicating an excellent fit to the historical data.

In Figure 3, the blue bars represent observed values, while the orange dots and curves indicate the estimated values and trends. In the absence of COVID-19, the projected number of visitors to Inner Mongolia in 2020 is 201.26 million, and projected tourism revenue is 544.9 billion CNY (U.S. $78.9 billion). The estimated visitor number is 1.61 times the actual count, and estimated revenue is 2.27 times the observed figure. These results imply that the COVID-19 pandemic reduced visitor numbers by approximately 40% and tourism revenue by 56%.

Using the projected visitor numbers and tourism revenue, we estimated the tourism value of Inner Mongolia's grassland tourism in the absence of COVID-19. We assumed that the change in visitor numbers along each route in the origin–destination matrices is proportional to the overall increase in total visitors. Accordingly, visitor numbers on all routes were multiplied by 1.61, while all other calculation procedures remained unchanged. As shown in Table A5, the resulting tourism value, estimated using the parametric method, is 936.5 billion CNY (U.S. $135.7 billion), or 1.6 times the observed value. This finding suggests that the COVID-19 pandemic reduced the tourism value by approximately 59.8%.

[Insert Figure 3 Here]

Factors influencing visitor numbers and tourism value

Given the uneven spatial distribution of tourism value across Inner Mongolia, we investigate the key factors influencing visitor numbers and tourism value. Our analysis employs both annual data to explore correlations and time-period data to examine potential causal relationships, particularly focusing on NDVI effects.

Table 3 presents regression results using annual data, revealing that grassland characteristics, local infrastructure, and regional economic development are significantly associated with both visitor numbers and tourism value. Grassland NDVI emerges as a particularly strong predictor across all specifications. Among infrastructure variables, high-grade scenic areas (4A and 5A) and accommodation facilities show consistently positive and significant effects, while high-speed rail stations demonstrate substantial impact on visitor numbers. Regional GDP and population are positively associated with tourism value, though the relationship with visitor numbers varies across specifications.

[Insert Table 3 Here]

To establish causal relationships, Table 4 presents results from time-period analysis focusing on grassland NDVI levels. The analysis categorizes NDVI into low, medium, and high levels, using low level as the reference category. Results demonstrate a clear positive causal relationship between grassland quality and tourism outcomes, with effects increasing progressively from medium to high NDVI levels. Both parametric and non-parametric approaches yield consistent findings, confirming the robustness of the causal effect of grassland quality on visitor numbers and tourism revenue.

[Insert Table 4 Here]

To further explore the spillover effects of infrastructure and grassland quality, we run two additional regression models. The first model examinesOne is to answer whether the positive effects from of local infrastructure is are enhanced by the improvement in grassland quality in the neighboring counties. Specifically, we create a dummy variable indicating whether the county has local infrastructure (4A/5A attractions or high-speed rail stations). We then interact it with a dummy variable indicating whether at least one neighboring county has high NDVI values, based on median classification. As shown in Appendix Table A6, the interaction effects between local infrastructure and neighboring high NDVI are positive and significant across all specifications. This suggests that infrastructure-equipped counties may capture additional tourism benefits when located near high-quality grasslands.

The second model examines whether counties with high grassland quality benefit from infrastructure in neighboring counties. We interact a dummy variable for local high NDVI (above median) with dummy variables for neighboring infrastructure. As shown in Appendix Table A7, the results reveal differentiated effects by infrastructure type. The interaction effects between local high NDVI and neighboring 4A/5A attractions are positive and significant, indicating regional tourism synergies where attractions and ecological resources complement each other. However, the interaction effects between local high NDVI and neighboring high-speed rail stations are negative and significant.

5. Discussion and conclusion

This study integrates mobile signaling data with the travel cost method to estimate the tourism value of Inner Mongolia's grasslands. The results indicate a total tourism value of 585.9–598.4 billion CNY (USD 84.9–86.7 billion), equivalent to approximately 2.5 times the official tourism revenue and about 29% of the region's GDP. Simulation analysis suggests that, in the absence of the COVID-19 pandemic, the 2020 tourism value would have reached approximately 936.5 billion CNY (USD 135.7 billion). Spatial analysis reveals pronounced heterogeneity: tourism value declines with distance from Inner Mongolia, with intra-regional cities contributing 43% of the total, while southern and eastern cities exhibit higher values due to superior infrastructure and vegetation quality. Grassland characteristics, local infrastructure, and regional economic development are all significantly associated with visitor numbers and tourism revenue. Causal analysis further demonstrates that high-quality grasslands play a decisive role in driving both visitation and the overall economic value of grassland tourism.

Future research should address several limitations and extend the present analysis. First, the study is constrained by the multi-destination nature of tourism in Inner Mongolia. In our dataset, multi-destination trips are recorded as sequential flows: a traveler from Guangzhou visiting County A and then County B in Inner Mongolia generates two observations—one from Guangzhou to County A, and another from County A to County B. Travel costs are calculated based on the recorded origin of each leg. While this avoids double-counting of initial travel expenses, it may underestimate true travel costs for long-distance tourists in subsequent legs, effectively treating them as near-origin visitors and potentially leading to downward bias in tourism value estimates for secondary destinations. Although adjacency-based spatial adjustments were employed to partially mitigate this issue, future research combining mobile phone data with GPS trajectory or survey information would enable more accurate cost attribution across multi-destination itineraries.

Second, this study is limited by traveler heterogeneity. The city–county level aggregation obscures individual-level variation in characteristics such as age, education, and trip purpose, all of which may influence visitation behavior. Although per capita GDP, gender, and age proportion of origin cities were included as a proxy for economic heterogeneity, the observable control variables may not fully capture the diversity of individual preferences and travel motivations. Consequently, the results should be interpreted as representing aggregate behavioral tendencies rather than individual-level responses. Future research integrating mobile signaling data with micro-level survey or trajectory information would enable more precise modeling of tourism demand and welfare effects.

Finally, the empirical analysis relies on data from 2020, a year coinciding with the onset of the COVID-19 pandemic. During this period, travel restrictions, public health concerns, and heterogeneous reopening policies across regions likely altered tourism flows and visitor behavior in ways not representative of typical conditions. Consequently, the estimated demand elasticities and tourism values may not fully reflect pre-pandemic patterns or long-term behavioral dynamics. Future research incorporating multi-year datasets—spanning both pre- and post-pandemic periods—would allow for a more comprehensive assessment of the temporal stability and external validity of the results.

These findings offer three policy implications centered on a single goal: making sure the economic value of nature is reflected in policy decisions. First, total tourism value—not just ticket sales—should be used to design ecological compensation. We found a massive gap between the estimated tourism value (585.9 billion CNY) and actual revenue (240 billion CNY). This gap represents a "consumer surplus," where visitors receive a high-value experience for a low cost. Currently, counties that produce high ecological value but collect little revenue are essentially giving away their natural assets for free. To fix this, these counties should receive higher conservation payments to make up for the wealth they provide to the public.

Second, protecting grasslands should be viewed as a smart investment in tourism infrastructure. Our data shows that as the greenery of the land (NDVI) improves, the number of visitors and the money they bring in both increases. In fact, this effect becomes even stronger as the land gets healthier. This means that conservation and economic growth work together: if the land is over-developed and damaged, its appeal to tourists disappears. We recommend that planners include these expected tourism gains when deciding whether to fund restoration projects. Additionally, we should set limits on visitor numbers based on the land's health (NDVI) to ensure that tourism does not destroy the very nature that attracts people in the first place.

Third, regional coordination is essential when planning infrastructure development and ecological conservation. Our study reveals that infrastructure-equipped “hub” counties benefit significantly from proximity to high-quality grasslands. However, ecological “destination” counties show mixed responses: they benefit from nearby attractions through regional synergies, but lose visitors when neighboring counties have high-speed rail stations that divert tourists elsewhere. These findings suggest that hub counties should develop infrastructure to capture ecological spillovers, while destination counties should prioritize grassland conservation and leverage connections with nearby attractions. Policy-makers must coordinate across county lines to build complementary regional tourism networks.

Acknowledgements

This work acknowledges the financial support of the National Natural Science Foundation of China (Grant no. 72322008, W2412012 and 72173004).

References

Altman, N. S. (1992). An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *The American Statistician*, *46*(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

Armbrecht, J. (2014). Use value of cultural experiences: A comparison of contingent valuation and travel cost. *Tourism Management*, *42*, 141–148. https://doi.org/10.1016/j.tourman.2013.11.010

Börger T, Maccagnan A, White MP, et al (2023) Was the trip worth it? Consistency between decision and experienced utility assessments of recreational nature visits. *American Journal of Agricultural Economics* 105:525–545. https://doi.org/10.1111/ajae.12328

Calabrese, F., Ferrari, L., & Blondel, V. D. (2014). Urban Sensing Using Mobile Phone Network Data: A Survey of Research. *ACM Comput. Surv.*, *47*(2), 25:1-25:20. https://doi.org/10.1145/2655691

Champ, P. A., Boyle, K. J., & Brown, T. C. (2017). *A Primer on Nonmarket Valuation*. Springer. https://books.google.com/books?hl=zh-CN&lr=&id=Rg8bDgAAQBAJ&oi=fnd&pg=PR5&ots=PLB3R_Fcq3&sig=nM-eJRUsr6kjlwUvg2l9VFdN1Hg#v=onepage&q&f=false

Chinese Ministry of Culture and Tourism, ed. (2020). Chinese Cultural Relics and Tourism Statistical Yearbook. Beijing: National Library Press.

Chinese Ministry of Industry and Information Technology of the People's Republic of China (2021). Communication Industry Statistical Report. Retrieved from China Government

Website.

Chinese National Bureau of Statistics (2020). China Statistical Yearbook. Beijing: China Statistics Press.

Chinese National Bureau of Statistics (2021). China Statistical Yearbook. Beijing: China Statistics Press.

Chinese National Bureau of Statistics (2022). China Statistical Yearbook. Beijing: China Statistics Press.

Costanza R, d'Arge R, de Groot R, et al (1997) The value of the world's ecosystem services and natural capital. *Nature* 387:253–260. https://doi.org/10.1038/387253a0

Costanza R, de Groot R, Sutton P, et al (2014) Changes in the global value of ecosystem services. Global *Environmental Change* 26:152 – 158. https://doi.org/10.1016/j.gloenvcha.2014.04.002

Dai, P., Zhang, S., Gong, Y., Zhou, Y., & Hou, H. (2022). A crowd-sourced valuation of recreational ecosystem services using mobile signal data applied to a restored wetland in China. *Ecological Economics*, *192*, 107249. https://doi.org/10.1016/j.ecolecon.2021.107249

Ding, Y., Chun, L., Sun, J., Wu, Z., Yun, X., & Lai, Y. (2020). China's Grasslands. *Forests and People*, *365*, 20-39+12-19 (Chinese).

Emiru, R., & Gemechu, A. (2017). Valuing the benefits of recreational wetland ecosystem: An application of contingent valuation and travel cost methods: The case of Boye recreational wetland, Jimma zone, Oromia National Regional State, Ethiopia. *Journal of Resources Development and Management*, 29, 78–99.

English, E. (2020). Time in Consumer Theory. *Unpublished Manuscript*.

Fezzi, C., Bateman, I. J., & Ferrini, S. (2014). Using revealed preferences to estimate the Value of Travel Time to recreation sites. *Journal of Environmental Economics and Management*, *67*(1), 58–70. https://doi.org/10.1016/j.jeem.2013.10.003

Foley JA, DeFries R, Asner GP, et al (2005) Global consequences of land use. *Science* 309:570–574. https://doi.org/10.1126/science.1111772

Freeman III, A. M., Herriges, J. A., & Kling, C. L. (2014). *The Measurement of Environmental and Resource Values: Theory and Methods* (3rd ed.). Routledge. https://doi.org/10.4324/9781315780917

Getzner, M., & Meyerhoff, J. (2020). The Benefits of Local Forest Recreation in Austria and Its Dependence on Naturalness and Quietude. *Forests*, *11*(3), Article 3. https://doi.org/10.3390/f11030326

Goodchild, M. F. (2013). The quality of big (geo)data. *Dialogues in Human Geography*, *3*(3), 280–284. https://doi.org/10.1177/2043820613513392

Houngbeme, D. J.-L., Igue, C. B., & Cloquet, I. (2021). Estimating the value of beach recreation in Benin. *Tourism Recreation Research*, *46*(3), 390–402. https://doi.org/10.1080/02508281.2020.1777052

Inner Mongolia Autonomous Region Bureau of Statistics (2021). Inner Mongolia Statistical Yearbook 2021. Beijing: China Statistics Press.

Jaung, W., & Carrasco, L. R. (2020). Travel cost analysis of an urban protected area and parks in Singapore: A mobile phone data application. *Journal of Environmental Management*, *261*, 110238. https://doi.org/10.1016/j.jenvman.2020.110238

Kamarol, S. K. A., Jaward, M. H., Kälviäinen, H., Parkkinen, J., & Parthiban, R. (2017). Joint facial expression recognition and intensity estimation based on weighted votes of image sequences. *Pattern Recognition Letters*, *92*, 25–32. https://doi.org/10.1016/j.patrec.2017.04.003

Kamel, N., Atiya, A. F., EI Gayar, N., & El-Shishiny, H. (2008). Tourism demand foreacsting using machine learning methods. *ICGST International Journal on Artificial Intelligence and Machine Learning*, *8*, 1–7.

Kubo, T., Uryu, S., Yamano, H., Tsuge, T., Yamakita, T., & Shirayama, Y. (2020). Mobile phone network data reveal nationwide economic value of coastal tourism under climate change. *Tourism Management*, *77*, 104010. https://doi.org/10.1016/j.tourman.2019.104010

Liang, Y., Yin, J., Pan, B., Lin, M. S., Miller, L., Taff, B. D., & Chi, G. (2022). Assessing the validity of mobile device data for estimating visitor demographics and visitation patterns in Yellowstone National Park. *Journal of Environmental Management*, *317*, 115410. https://doi.org/10.1016/j.jenvman.2022.115410

Li, W., Gao, H., Lv, J., & Li, L. (2021). Thoughts on the transformation and upgrading of the grassland tourism industry and the improvement of herders' livelihoods in Inner Mongolia. *Northern Economy,* No. 406, 57–59. (Chinese)

Lin, Y., Zhou, Y., Lin, M., Wu, S., & Li, B. (2021). Exploring the disparities in park accessibility through mobile phone data: Evidence from Fuzhou of China. *Journal of Environmental Management*, *281*, 111849. https://doi.org/10.1016/j.jenvman.2020.111849

Liu, W.-Y., Fang, B.-S., & Hsieh, C.-M. (2021). Evaluating the Recreation Value of Alishan National Forest Recreation Area in Taiwan. *Forests*, *12*(9), Article 9. https://doi.org/10.3390/f12091245

Loomis, J., Englin, J., McDonald, J., Hilger, J., & Gonzalez-Caban, A. (2000). Testing Transferability of Forest Recreation Demand in the Three Intermountain States with an Application to Forest Fire Effects. *Western Region-Western Extension Directors Association (WEDA)*.

Lupi, F., Phaneuf, D. J., & von Haefen, R. H. (2020). Best Practices for Implementing Recreation Demand Models. *Review of Environmental Economics and Policy*, *14*(2), 302–323. https://doi.org/10.1093/reep/reaa007

Lv, J., Wang, Y., & Liu, L. (2006). Assessment of the recreation value of grassland ecosystems: A case study of Siziwang Banner in Inner Mongolia Autonomous Region. *Tourism Tribune*, 69-74(Chinese).

McAleer, M. (1994). Sherlock Holmes and the Search for Truth: A Diagnostic Tale. *Journal of Economic Surveys*, *8*(4), 317–370. https://doi.org/10.1111/j.1467-6419.1994.tb00106.x

Nyns, S., & Schmitz, S. (2022). Using mobile data to evaluate unobserved tourist overnight stays. *Tourism Management*, *89*, 104453. https://doi.org/10.1016/j.tourman.2021.104453

Office of the Leading Group for the Third National Land Survey of the State Council, Ministry of Natural Resources, National Bureau of Statistics (2021). Announcement of the Main Data from the Third National Land Survey. Beijing: National Bureau of Statistics.

Ong, C. (1995). Tourism demand models: A critique. *Mathematics and Computers in Simulation*, *39*(3–4), 367–372. https://doi.org/10.1016/0378-4754(95)00085-1

Parsons, G., Leggett, C. G., Herriges, J., Boyle, K., Bockstael, N., & Chen, Z. (2021). A Site-Portfolio Model for Multiple-Destination Recreation Trips: Valuing Trips to National Parks in the Southwestern United States. *Journal of the Association of Environmental and Resource Economists*, *8*(1), 1–25. https://doi.org/10.1086/710714

Parsons, G. R., Chen, Z., Hidrue, M. K., Standing, N., & Lilley, J. (2013). Valuing Beach Width for Recreational Use: Combining Revealed and Stated Preference Data. *Marine Resource Economics*, *28*(3), 221–241. https://doi.org/10.5950/0738-1360-28.3.221

Pascoe, S. (2019). Recreational beach use values with multiple activities. *Ecological Economics*, *160*, 137–144. https://doi.org/10.1016/j.ecolecon.2019.02.018

Peng, W., Yao, S., & Feng, Y. (2014). Recreational value assessment by TCIA and CVM: A case study of Taibai Mountain National Forest Park. *Economic Geography*, 34, 186–192.

Philip Chen CL, Zhang C-Y (2014) Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences 275:314 – 347. https://doi.org/10.1016/j.ins.2014.01.015

Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. *Environment and Planning B: Planning and Design*, *33*(5), 727–748. https://doi.org/10.1068/b32047

Raun, J., Ahas, R., & Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, *57*, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006

Salganik, M. J. (2019). *Bit by Bit: Social Research in the Digital Age*. Princeton University Press.

Taye FA, Folkersen MV, Fleming CM, et al (2021) The economic values of global forest ecosystem services: A meta-analysis. *Ecological Economics* 189:107145. https://doi.org/10.1016/j.ecolecon.2021.107145

The Culture and Tourism Department of Inner Mongolia Autonomous Region (2021). Statistical Bulletin on the Development of Culture and Tourism in Inner Mongolia. Hohhot: The Culture and Tourism Department of Inner Mongolia Autonomous Region.

Vidal Gimenez, F., & Ruiz Mas, C. (2020). The Valuation of Recreational Use of Wetlands and the Impact of the Economic Crisis. *International Journal of Environmental Research and Public Health*, *17*(9), Article 9. https://doi.org/10.3390/ijerph17093228

Wang, X., Chen, Y., & Wang, L. (2022). Assessment of the ecotourism and recreational value of grasslands on the Tibetan Plateau—Based on the modified travel cost interval analysis method. *Journal of Arid Land Resources and Environment*, *36*, 192-200(Chinese).

Xie G, Zhang C, Zhen L, Zhang L (2017) Dynamic changes in the value of China's ecosystem services. *Ecosystem Services* 26:146–154. https://doi.org/10.1016/j.ecoser.2017.06.010

Xu, M., Xin, J., Su, S., Weng, M., & Cai, Z. (2017). Social inequalities of park accessibility in Shenzhen, China: The role of park quality, transport modes, and hierarchical socioeconomic characteristics. *Journal of Transport Geography*, *62*, 38–50. https://doi.org/10.1016/j.jtrangeo.2017.05.010

Xu, S., & He, X. (2022). Estimating the recreational value of a coastal wetland park: Application of the choice experiment method and travel cost interval analysis. *Journal of Environmental Management*, *304*, 114225. https://doi.org/10.1016/j.jenvman.2021.114225

Yang, Y., Xiong, C., Zhuo, J., & Cai, M. (2021). Detecting Home and Work Locations from Mobile Phone Cellular Signaling Data. *Mobile Information Systems*, *2021*(1), 5546329. https://doi.org/10.1155/2021/5546329

Zhang, H.-X., Su, Q., & Wang, Q. (2006). A summary of overseas research on the recreational value evaluation of tourism resources. *Tourism Tribune*, 21, 31–35.

Table 1. Descriptive statistics

Variable	Description	Mean	SD	Min	Max	N
Dependent variables						
Ln (county visitor)	Log of tourist numbers in Inner Mongolia destination counties	13.14	0.83	10.22	15.18	35712
Ln (county tourism value)	Log of tourism value in Inner Mongolia destination counties based on parametric method	15.27	1.67	9.10	21.67	35712
	Log of tourism value in Inner Mongolia destination counties based on non-parametric method	14.65	1.72	9.96	22.69	35712
Independent variables						
Grassland characteristics						
Grassland NDVI	Maximum daily grassland NDVI value in the current period	0.94	0.06	0.58	1.00	35712
Ln (Main grassland area)	Log of area of grassland types suitable for grassland tourism, including lowland meadow, alpine meadow, mountain meadow, temperate meadow steppe, and temperate steppe	21.57	1.51	16.06	24.53	29079
Number of main grassland types	Number of grassland types suitable for grassland tourism (same types as above)	2.63	0.86	1.00	4.00	29079
Local infrastructure						
Number of 4A and 5A scenic areas	Number of 4A and 5A rated tourist attractions in Inner Mongolia destination counties	1.50	1.58	0.00	9.00	35712
Number of 2A and 3A scenic areas	Number of 2A and 3A rated tourist attractions in Inner Mongolia destination counties	2.43	1.93	0.00	8.00	35712
Number of high-end hotels	Number of three-star and above hotels in Inner Mongolia destination counties	0.50	1.15	0.00	6.00	35712
Number of budget hotels	Number of three-star and below hotels in Inner Mongolia destination counties	1.71	1.94	0.00	10.00	35712
High-speed rail station	Whether Inner Mongolia destination counties have high-speed rail stations in 2020 (1=Yes, 0=No)	0.12	0.32	0.00	1.00	35712
Regional economic development						
Ln (GDP)	Log of GDP in Inner Mongolia destination counties	7.32	1.17	3.43	10.56	35712
Population	Permanent resident population in Inner Mongolia destination counties (10 thousand people)	24.02	15.92	1.50	84.29	35712

Notes: This table provides descriptive statistics for variables used in the analysis of tourism in Inner Mongolia counties.

Table 2. Robustness analysis of tourism value estimation

Scenario 1: Control variables and subgroup analysis		
	Tourism value (billion CNY)	
Control Variables		
Region Fixed effects	585.89	
Region Fixed effects, Age, and Gender	686.54	
Group Estimates		
	Each group	Total
Low departure city per capita GDP group	175.54	
Medium departure city per capita GDP group	198.92	590.70
High departure city per capita GDP group	216.24	
Scenario 2: Data frequency analysis		
	Tourism value (billion CNY)	
	Parametric	Non-parametric
Based on annual data	585.89	598.36
Based on time-period data	593.94	640.40

Notes: The table reports robustness checks of tourism value estimates under alternative model specifications and scenarios. *Scenario 1* presents results using different control variable sets (baseline: region fixed effects; extended: adding age and gender controls) and subgroup analyses by departure city per capita GDP levels (low, medium, high). *Scenario 2* compares tourism value estimates based on annual data and time-period data.

Table 3. Factors influencing visitor numbers and tourism value

	Ln (county visitor)	Ln (county tourism value)	
		Parametric	Non-parametric
	(1)	(2)	(3)
Grassland characteristics			
Grassland NDVI	0.412***	0.843***	0.212*
	(0.057)	(0.138)	(0.113)
Ln (Main grassland area)	0.007**	0.072***	0.007
	(0.003)	(0.006)	(0.005)
Number of main grassland types	-0.119***	-0.199***	-0.072***
	(0.004)	(0.011)	(0.009)
Local infrastructure			
Number of 4A and 5A scenic areas	0.136***	0.153***	0.143***
	(0.002)	(0.005)	(0.004)
Number of 2A and 3A scenic areas	-0.012***	0.007*	0.068***
	(0.002)	(0.004)	(0.003)
Number of high-end hotels	0.056***	0.016**	0.163***
	(0.003)	(0.008)	(0.006)
Number of budget hotels	0.071***	0.072***	0.075***
	(0.002)	(0.004)	(0.003)
High-speed rail station	0.222***	0.030	0.093***
	(0.009)	(0.020)	(0.017)
Regional economic development			
Ln (GDP)	-0.016	0.919***	1.196***
	(0.016)	(0.039)	(0.027)
Population	0.027***	0.031***	0.017***
	(0.000)	(0.001)	(0.000)
Constant	12.010***	5.681***	4.587***
	(0.158)	(0.415)	(0.279)
City fixed effects	Yes	Yes	Yes
N	29,079	29,079	29,079
R-squared	0.531	0.594	0.731

Notes: The table presents regression results examining factors influencing visitor numbers and tourism value using annual data. Column (1) presents regression results for visitor numbers based on annual data, while columns (2)-(3) show regression results for tourism value based on annual data using parametric and non-parametric methods respectively. The numbers in parentheses represent the estimated standard errors. Asterisks denote the level of statistical significance, *** indicates $p<0.01$, ** indicates $p<0.05$, and * indicates $p<0.1$.

Table 4. The effects of grassland quality on tourism

	Ln (county visitor)	Ln (county tourism value)	
		Parametric	Non-parametric
	(1)	(2)	(3)
The level of grassland NDVI (Based on low level)			
Medium level of grassland NDVI	0.204***	0.240***	0.533***
	(0.009)	(0.011)	(0.014)
High level of of grassland NDVI	0.461***	0.410***	1.276***
	(0.010)	(0.013)	(0.016)
Constant	10.529***	18.406***	18.025***
	(0.051)	(0.053)	(0.067)
City fixed effects	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes
N	97,731	97,731	97,731
R-squared	0.738	0.547	0.404

Notes: The table presents analytical results regarding NDVI effects using time-period data. Column (1) presents regression results for visitor numbers based on time-period data, while columns (2)-(3) show regression results for tourism value based on time-period data using parametric and non-parametric methods respectively. The numbers in parentheses represent the estimated standard errors. Asterisks denote the level of statistical significance, *** indicates $p<0.01$, ** indicates $p<0.05$, and * indicates $p<0.1$.

Figure 1. Geographic location of the Inner Mongolia Autonomous Region, China, and the distribution of grasslands within Inner Mongolia in 2020.

[Insert Liu Figure 1.png Here]

Notes: The left panel shows the geographic location of the Inner Mongolia Autonomous Region within China, while the right panel illustrates the spatial distribution of grassland quality indicated by NDVI in Inner Mongolia in 2020.

Figure 2. Distribution of tourism value across cities in Inner Mongolia, China in 2020

(a) Based on the parametric method

[Insert Liu Figure 2a.png Here]

(b) Based on a non-parametric method

[Insert Liu Figure 2b.png Here]

Notes: The figure presents the distribution of tourism value across cities in Inner Mongolia, China, in 2020. Panel (a) shows the tourism value of each city calculated using the parametric method, while Panel (b) shows the tourism value of each city calculated using the non-parametric method.

Figure 3. Visitor number and tourism income in Inner Mongolia, China, from 2000 to 2020

(a) Visitor number

[Insert Liu Figure 3a.png Here]

(b) Tourism income

[Insert Liu Figure 3b.png Here]

Notes: The figure shows the number of tourists and tourism income in Inner Mongolia from 2000 to 2019, with 2020 trends extrapolated using a third-order polynomial method to exclude COVID-19 impacts. The blue bars represent actual values, while the orange dots and curves indicate estimated values and trends.

Footnotes

[1] When creating Figure A1, we divide the number of visitors into five equal groups.

[2] When creating Figure A2, we divide the number of visitors into five equal groups.

[3] For k-value selection in KNN regression, our main results are based on an adaptive optimal k-value selection method. Specifically, we test different k values and employ 5-fold cross-validation to evaluate the prediction performance of each k value. We select the optimal k that minimizes prediction errors based on both mean training MSE and mean cross-validation MSE. Additionally, to test the robustness of our results, we conduct sensitivity analysis using fixed k values (k=5,7,10). Detailed evaluation results are presented in Appendix A, Table A4.

[4] Data source: Gao, J., Shi, Y., Zhang, H., Chen, X., Zhang, W., Shen, W., Xiao, T., Zhang, Y. (2022). China regional 250m fractional vegetation cover data set (2000-2024). National Tibetan Plateau / Third Pole Environment Data Center. https://doi.org/10.11888/Terre.tpdc.300330. https://cstr.cn/18406.11.Terre.tpdc.300330.