



Lin Qu¹, Yue Zhou¹, Jiangxin Li², Qiong Yu¹ and Xinguo Jiang^{1,3,4,*}

- ¹ School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, China; truly@my.swjtu.edu.cn (L.Q.); yuezhou@my.swjtu.edu.cn (Y.Z.); qiongyu@my.swjtu.edu.cn (Q.Y.)
- ² School of Computer Science, Fudan University, Shanghai 200433, China; jiangxinli21@m.fudan.edu.cn
- ³ National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu 611756, China
- ⁴ School of Transportation, Fujian University of Technology, Fuzhou 350118, China
- * Correspondence: xjiang@swjtu.edu.cn; Tel.: +86-135-5128-1570

Abstract: Map matching of trajectory data has wide applications in path planning, traffic flow analysis, and intelligent driving. The process of map matching involves matching GPS trajectory points to roads in a roadway network, thereby converting a trajectory sequence into a segment sequence. However, GPS trajectories are frequently incorrectly matched during the map-matching process, leading to matching errors. Considering that few studies have focused on the causes of map-matching errors, as well as the distribution of these errors, the study aims to investigate the spatiotemporal characteristics and the contributing factors that cause map-matching errors. The study employs the Hidden Markov Model (HMM) algorithm to match the trajectories and identifies the four types of map-matching errors by examining the relationship between the matched trajectories and the driving routes. The map-matching errors consist of Off-Road Error (ORE), Wrong-match on Road Error (WRE), Off-Junction Error (OJE), and Wrong-match in Junction Error (WJE). The kernel density method and multinomial logistic model are further exploited to analyze the spatiotemporal patterns of the map-matching errors. The results indicate that the occurrence of map-matching errors substantially varies in time and space, with variation significantly influenced by intersection features and road characteristics. The findings provide a better understanding of the contributing factors associated with map-matching errors and serve to improve the accuracy of map matching for commercial vehicles.

Keywords: map matching; incorrect match trajectories; temporal and spatial distribution; multinomial logistic model; kernel density model

1. Introduction

Vehicle trajectory data have been widely utilized in traffic flow detection, route selection, and traffic violation enforcement, which are vital to transportation management in urban areas [1,2]. The analysis of vehicle trajectories is important for transport industries, such as taxi driving. Trajectory data can assist taxi companies in detecting their drivers' behaviors and understanding their operation patterns, which can be further used to enhance driving safety and transport efficiency [3–5]. For instance, it is possible to evaluate the route planning of taxi drivers with trajectories and optimize their operational routines. Meanwhile, the trajectories can be used to identify traffic incidents encountered by taxi drivers and warn taxi managers if necessary [6,7]. For example, Kan et al. [8] analyzed the conditions of urban traffic congestion using taxi GPS trajectory data and proposed a method for congestion assessment. This approach could help transportation authorities to formulate more effective policies and strategies. In addition to congestion detection, the use of taxi trajectories can reveal the speed characteristics of roadway segments with different control types (e.g., one-way control and road closure) [9–11]. Hereby, the quality



Citation: Qu, L.; Zhou, Y.; Li, J.; Yu, Q.; Jiang, X. HMM-Based Map Matching and Spatiotemporal Analysis for Matching Errors with Taxi Trajectories. *ISPRS Int. J. Geo-Inf.* 2023, *12*, 330. https://doi.org/ 10.3390/ijgi12080330

Academic Editor: Wolfgang Kainz

Received: 23 May 2023 Revised: 1 August 2023 Accepted: 4 August 2023 Published: 7 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). of GPS trajectories should be ensured to successfully implement the above-stated applications.

Although the vehicle trajectories are readily available, the abnormality can be commonly found in most datasets. The trajectory points recorded by the GPS receivers may be positioned away from the actual locations because of the impacts of road environments [12], satellite positioning errors [13], instability of GPS signal transmission [14], and algorithm differences [15]. These deviations lead to inaccurate detection of actual driving lanes and directions and the disappearance of the trajectories from the roadways [16]. To address these issues, studies typically adopt the map-matching algorithm to correct the coordinates of the recorded GPS points and match them to the corresponding roads along the driver's routes.

Different methods and algorithms are used in map matching for the GPS trajectories, which exhibit notable differences in terms of model structure, data inputs, and logical rules [17]. Currently, traditional map-matching algorithms encompass geometric mapmatching, topological map-matching, probabilistic statistical map-matching, and advanced map-matching algorithms [18,19]. Furthermore, various new technologies have recently been introduced to the field of map matching, including low-frequency trajectory data match [20,21]; high-frequency trajectory data match [22]; methods that consider memorized multiple matching candidates [21]; the AMM algorithm for online map matching [23]; deep learning-based models [24,25], including Recurrent Neural Networks (RNNs) [25] or Convolutional Neural Networks (CNNs) [25]; the Python Toolbox (PyTrack) [26]; and the Valhalla solution based on an open-source routing engine [27]. These methods have been successfully used to match the GPS points from various trajectories. However, they often have specific requirements for the trajectory data (e.g., data volume, data frequency, and error distribution) or the scenarios (e.g., online or offline matching). Some methods aim to enhance the matching accuracy by incorporating additional data features or adjusting algorithm parameters. Nonetheless, these approaches often suffer from high memory usage and large time costs during the map-matching process, making it challenging to match large GPS data on complex road networks.

The HMM-based map-matching method effectively addresses many of the drawbacks associated with the aforementioned approaches [28]. Due to the simplicity and Markovian property of the HMM algorithm, it offers significant improvements in terms of computational efficiency, storage efficiency, and broad applicability. It is particularly well-suited to handling errors on complex road networks from large-scale datasets [27,28]. This superiority makes it a valuable tool in the practical applications of trajectory matching. However, these map-matching algorithms may yield incorrect matches because of their complicated trajectories. For instance, past studies have found that curved driving [15], high-speed movement [29], and road network topology [30] can contribute to wrong matches. These map-matching errors could mislead the identification and prediction of vehicle status and behaviors, consequently impairing the ability of systems to monitor the trajectories of running vehicles. Hence, the map-matching errors ought to be specifically identified and addressed.

Currently, although some research has focused on the issue of map-matching errors, there is relatively limited research into the mechanism of map-matching errors, as well as analysis of the spatiotemporal distribution and road characteristics of these errors. For example, Dey et al. [17] proposed a method to automatically identify and detect map-matching errors in the absence of ground truth. Chao et al. [31] found that the density of roads and road segments with curves significantly impacted the quality of map matching. Furthermore, Luo et al. [32] found that intersections and large indoor areas often resulted in significant indoor positioning errors. Different measurements have been developed to detect and deal with these map-matching errors. The application of detection approaches is highly dependent on the availability of data sources and the purposes of matching. The most common method is the rule-based approach that detects a wrong match by examining the relationship between trajectories and roadway networks. For ex-

ample, the method could detect errors that occur when trajectories are out of the roads or a part of the trajectories vanishes from the roads (i.e., deviation) [33]. Another type of method is developed through machine learning, among which supervised learning is commonly used to detect wrong-matching trajectories using training data, such as the use of Support Vector Machines (SVM) [34] and Random Forests (RF) [35]. Comparatively, the Fusion-based method, which generally consists of several modules used to improve judgment accuracy, has also been constructed to detect map-matching errors [36]. Despite the various approaches, the wrongly matched trajectories cannot be fully detected and effectively corrected. One of the critical reasons for this issue is that the mechanism of map-matching errors is not thoroughly uncovered.

As such, investigating how map-matching errors occur could improve the quality of map matching when matching the trajectories to the locations at which errors are prone to appear. To this end, the spatiotemporal analysis could be conducted to reveal the mechanism of the map-matching errors and explore their characteristics. For instance, Santi et al. [37] conducted a spatial-temporal analysis to identify the patterns of taxi travel based on New York City taxi trajectories. Livio et al. [38] combined traffic accidents and GPS trajectory data to identify accident black spots in space and time scales. Likewise, spatiotemporal analysis could be used to identify the distribution of map-matching errors and reveal the location or scenarios that are associated with the occurrence of the errors.

In summary, this study attempts to enrich the knowledge of where, when, and why various types of map-matching error occur by means of spatial-temporal and factor analysis. The study matches the trajectories based on the HMM algorithm and accordingly identifies different types of map-matching errors, as well as investigating the spatial-temporal distributions and contributing factors of the map-matching errors. The main contribution of the study is that we identify the map-matching errors generated via the HMM algorithm, analyze the spatiotemporal distribution patterns of these errors, and explore the relationship between map-matching errors and road environment factors. The conclusions could assist the analysts in understanding the occurrence of map-matching errors when applying the HMM and improving the accuracy of the HMM algorithm.

The rest of the study is organized as follows. The next section describes the study area, GPS trajectory data, and road features. Section 3 provides a detailed introduction of the map-matching method. Section 4 presents the study's results and discussions. Finally, Section 5 summarizes the study's conclusions.

2. Data Description and Pre-Processing

Taxi GPS trajectory data were obtained from Chengdu Municipal Traffic Management Bureau in China. Taxi GPS trajectories were collected in the period 1–14 September 2020 inside the First Ring Road area (Figure 1). The taxi data include three parts: taxi GPS trajectory data, a GIS map of the road network, and the dataset of road features. The taxi GPS trajectories contain approximately 1.4 billion GPS data points in the collection period. To control the GPS data size, 14 million trajectories generated by 500 randomly selected taxis were matched by the HMM algorithm and provided map-matching errors. There were 11 variables in the raw trajectory data recorded using the vehicular GPS recorder, including license plate number, plate color, alarm status, vehicle status, latitude, longitude, direction, speed, satellite time, creation time, and creator (Table 1).

The road feature dataset was manually fetched from Baidu Street View, Version 2020. In addition to road features, time and environmental factors were considered in the study. Table 2 describes all of the key factors of the errors. In addition, features of the roads that were immediately adjacent to the roads/intersections on which a map-matching error occurs in the driver's routine were considered to be factors. The roads adjacent to the error location were named as the previous road and the latter road, respectively (Figure 2).



Figure 1. Research area: the selected road network (marked by the green line).

Table 1. Samp	le of the raw	trajectories.
---------------	---------------	---------------

Field Name	Data Type	Description	Example
plate_number	string	License plate number	川ADT9088
plate_color	bigint	Plate color indicates whether the taxi is an electric vehicle	0
alarm	bigint	Alarm status indicates occasions on which the vehicle loses connection	0
state	bigint	Vehicle status indicates whether the vehicle is occupied by passengers	0
lat	double	Latitude	30.66301
lng	double	Longitude	104.0812
direction	bigint	Direction	214
speed	double	Instant speed	300
position time	bigint	Satellite time	1600457268000
crton	bigint	Creation time	1600457271250
crt_by	string	Time creator	10012



Figure 2. Road relationships: (**a**) when the current road section is segmented; (**b**) when the current road section is an intersection.

Туре	Variables	Descriptions	Mean	Standard Deviation
	Time-weekend	If the error occurs at the weekend, define it as "1"; otherwise, define it as "0".	0.26	0.439
Time	Time-peak hours	If the error occurs in peak hours, define it as "1"; otherwise, define it as "0".	0.28	0.448
	Time-night	If the error occurs at night, define it as "1"; otherwise, define it as "0".	0.37	0.483
	Int-flyover	If the error occurs near the flyover, define it as "1"; otherwise, define it as "0".	0.00	0.046
	Int-crossroad	If the error occurs near the crossroad, define it as "1"; otherwise, define it as "0".	0.25	0.432
	Int-X-junction	If the error occurs near the X-junction, define it as "1"; otherwise, define it as "0".	0.02	0.141
Intersection type	Int-T-junction	If the error occurs near the T-junction, define it as "1"; otherwise, define it as "0".	0.37	0.482
	Whether it is a large intersection	If the error occurs near a large intersection, define it as "1"; otherwise, define it as "0".	0.12	0.323
	Whether it is a medium-sized intersection	If the error occurs near a medium-sized intersection, define it as "1"; otherwise, define it as "0".	0.20	0.397
	Whether it is a small intersection	If the error occurs near a small intersection, define it as "1"; otherwise, define it as "0".	0.25	0.435
	PreRoad-speed limit	If the previous road speed limit is <30 km/h, define it as "1"; If the previous road speed limit is 30–50 km/h, define it as "2"; if the previous road speed limit is >50 km/h, define as "3".	0.55	0.497
	PreRoad-bicycle divider	If the error occurs on the previous road with the bicycle divider, define it as "1"; otherwise, define it as "0".	0.45	0.497
Previous road's characteristics	PreRoad-median divider	If the error occurs on the previous road with the median divider, define it as "1"; otherwise, define it as "0".	0.09	0.287
	PreRoad-roadside parking	If the error occurs on the previous road with roadside parking, define it as "1"; otherwise, define it as "0".	1.96	0.202
	PreRoad-one way	If the error occurs on the previous road that is one way, define it as "1"; if the error occurs on the previous road that is two way, define it as "2"	0.60724981355	4.633130389169
	PreRoad-commercial density	The 20-meter buffer radius POI point commercial density of the previous road.	1.06788664028	2.503877913297
	PreRoad-public service density	The 20-meter buffer radius POI point public service density of the previous road.	3.93085171376	4.339859341936
	PreRoad-resident density	The 20-meter buffer radius POI point resident density of the previous road.	0.26	0.439
	LatRoad-speed limit	"1"; if the latter road speed limit is 30–50 km/h, define as "2"; if the latter road speed limit is >50 km/h, define it as "3"	0.28	0.448
	LatRoad-bicycle divider	If the error occurs on the latter road with the bicycle divider, define it as "1"; otherwise, define it as "0".	0.37	0.483
	LatRoad-median divider	If the error occurs on the latter road with the median divider define it as "1", otherwise define it as " 0 "	0.00	0.046
Latter road's characteristics	LatRoad-roadside parking	If the error occurs on the latter road with roadside parking, define it as "1"; otherwise, define it as "0".	0.25	0.432
	LatRoad-one way	If the error occurs on the latter road that is one way, define it as "1"; if the error occurs on the latter road	0.02	0.141
	LatRoad-commercial density	that is two way, define it as "2". The 20-meter buffer radius POI point commercial density of the latter road	0.37	0.482
	LatRoad-public service	The 20-meter buffer radius POI point public service	0.12	0.323
	LatRoad-resident density	The 20-meter buffer radius POI point resident density of the latter road.	0.20	0.397

Table 2. Definitions and	d values of independent variables.
--------------------------	------------------------------------

Prior to map-matching error detection, there were two pre-processing steps. The first step involved GPS data and road network data pre-processing. In the GPS data pre-processing step, we conducted a cleaning procedure on the raw data, including remov-

ing the missing and duplicate data (e.g., duplicated records of GPS points with the same latitude–longitude and timestamp), as well as filtering out abnormal drift points (e.g., discontinuous jumps in the longitude and latitude data along the trajectory), and the second step was to adjust the coordinate system and restrict the study area by converting the GPS data latitude and longitude coordinates from World Geodetic System 1984 to Xian_1980_3_Degree_GK_CM_105E projected coordinate system for subsequent mapmatching operations. At the same time, we selected the research area within the First Ring Road of Chengdu city from the downloaded road network map for the map-matching process based on the HMM algorithm. This process was conducted by removing all of the roads and intersections beyond this area in the digital map.

Data processing and analysis were sequentially three-fold: raw data collection, data pre-processing, and map matching (Table 3). The data output from the previous process served as the input for the subsequent process.

Stages of Calculations	Field Name	Data Type	Description	Example
	Plate number	string	License plate number	川ADT9088
	lat	double	Latitude	30.67455
	lng	double	Longitude	104.08496
Party Data Stage	Position time	bigint	Satellite time	1600457268000
Raw Data Stage	Shape	string	Road segment type	Line
	0BJECTID	number	The ID of the road	1980
	Name	string	Road name	Section 3 South of First Ring Road
	Shape Leng	string	Road length	0.00751
	plate_number	string	License plate number	川ADT9088
	lat	double	Latitude	30.6746432
	lng	double	Longitude	104.0850109
Data Pre-Processing Stage	position time	bigint	Satellite time	20200919 03:27:48
Data Ties Tocessing Stage	Shape	string	Road segment type	Line
Raw Data Stage Data Pre-Processing Stage Map Matching Completed Stage	0BJECTID	number	The ID of the road	1980
	Name	string	Road name	Section 3 South of First Ring Road
	Shape Leng	string	Road length	0.00751
	plate_number	string	License plate number	川ADT9088
	lat	string	Latitude	3395261.8989394824
	lng	string	Longitude	35412317.51912209
Map Matching Completed	Position time	string	Satellite time	20200919 03:27:48
Stage	Shape	string	Road segment type	Line
	0BJECTID	number	The ID of the road	1980
	Name	string	Road name	Section 3 South of First Ring Road
	Shape Leng	string	Road length	0.00751

Table 3. Data types at various stages.

This study primarily utilized an Inspur server model NF5280M6, which was equipped with an Intel(R) Xeon(R) CPU E5-@ 2.10 GHz processor, 32 GB of RAM, and a 4-terrabite hard disk capacity. The platform runs on Windows Server 2019, and we used software such as ArcGIS, Python, and SPSS. ArcGIS was used for Kernel Density modeling analysis; Python was used for data pre-processing and map matching, incorporating modules like Pandas, ArcpyUtil, and HmmUtil; and SPSS software was used for analyzing the spatiotemporal factors that contributed to map-matching errors using the multinomial logistic regression model.

The experimental results indicated that the CPU time consumption in the data preprocessing stage was approximately 0.02112 s per iteration, memory usage was around 6.21 GB, and disk utilization was 1%. Similarly, in the map-matching error detection and analysis stage, the CPU time consumption was roughly 0.02592 s per iteration, memory usage was about 3.07 GB, and disk utilization stayed constant at 1%.

3. Methodology

In this study, we employed the HMM and Viterbi algorithm to match the GPS trajectories collected from taxis. Then, we defined four types of common map-matching errors. Based on the error dataset, we applied the Kernel Density Estimation model to reveal the spatial-temporal patterns of the map-matching errors. Finally, we adopted a multinomial logistic model to analyze significant factors that contributed to different types of map-matching errors.

3.1. Hidden Markov Model

The study adopted the HMM algorithm to match the GPS trajectories to their corresponding roads using the Chengdu GIS roadway network map. Since the actual position of the GPS point was unknown, HMM decomposed the probability of the real match into a combination of observation probability and transition probability [39]. The former feature represents the probability of the observed variables being on each road, while the latter feature represents the probability of GPS trajectory points transitioning from one road state to another in the road network. Hidden states represented the true locations of GPS points on the roads, while observation states represented the observed location of the GPS data. The HMM algorithm model assumed that the trajectory was generated from a series of hidden states. In the set of observed variables $X = x_1, x_2, \ldots, x_i, \ldots, x_n$, where x_i represents the location information of *i*th GPS trajectory point, i.e., longitude and latitude (lon_i, lat_i) , and in the set of hidden variables $Y = \{y_1, y_2, \dots, y_i, \dots, y_n\}$, where $y_i \in y$ stands for the true location of *i*th GPS trajectory point on a road segment, and n is the index of trajectory point. For the input GPS trajectory points and candidate road segments, we calculated the observation probability of the candidate road segments based on the HMM. Then, for each candidate road segment, we computed its path probability. This probability represented the maximum value obtained via multiplying the path probability of all possible previous states, the transition probability from the previous state to the current state, and the observation probability of the current state. This process selected the previous state with the highest path probability as the current state, thus constructing the optimal hidden state sequence, which corresponded to the best-matched road segments (note that the pseudocode of the HMM-based Map-matching algorithm is shown in Algorithm 1: HMM-based Map Matching).

Based on the transition probability and observation probability, the Viterbi algorithm could compute the most likely sequence of the hidden states, i.e., the best-matched locations of the taxi trajectory. Viterbi algorithm achieved trajectory matching by finding the most likely sequence of hidden states in the Hidden Markov Model. The basic logic of the algorithm was to construct a state path map, where each state represented a hidden state, and each edge on the graph represented the probability of transferring from one state to another and calculated the maximum probability of the initial state relative to that of another state. Finally, the state with the highest probability was selected as the final matching result, that is, the optimal matching path, which was viewed as the true path (Figure 3).



Figure 3. Flowchart of the trajectory matching process based on the HMM and Viterbi algorithm.

Algorithm 1: HMM-based Map Matching
Input : GPS trajectory points $P = (p_n n = 1,, N)$; candidate road segment $R =$
$(r_n \mid n = 1, \ldots, N)$
Output: Matched road segments $Y = (y_n n = 1,, N)$
1: Initialize Viterbi path viterbi[], path probability <i>v</i> [], and state sequence <i>y</i> []
2: For each point <i>p</i> in P
3: Calculate observation probabilities o[] for R based on the HMM
4: For each state <i>s</i> in R
5: Calculate $v[s] = \max_{\{s'\}}(v[s'] * transition(s', s) * o[s])$
6: viterbi[s] = argmax _{{s'} ($v[s']$ * transition(s', s))
7: $y.append(viterbi[argmax_s(v[s])])$
8: For $i = N - 1$ down to 1
9: $y_{i-1} = \text{viterbi}[y_i]$
10: return <i>y</i>

3.2. Matching Error Trajectory Recognition Method

In this research, we identify four types of map-matching errors (Figure 4), which are the most commonly observed errors generated via the HMM algorithm. The errors are as follows:

- Off-Road Error (ORE) refers to a trajectory point or two/three adjacent points (e.g., P5 in Figure 4a) that fall outside of a buffer of a 15-meter radius, while its adjacent GPS points are captured;
- Wrong-match on Road Error (WRE) indicates that a trajectory point or two/three adjacent points (P5 in Figure 4b) are incorrectly matched to a nearby road segment, instead of the road that contains its upstream and downstream trajectory points;
- (iii) Off-Junction Error (OJE) means a trajectory point or two/three adjacent points (P5 in Figure 4c) that ought to be located within an intersection, but the point is not captured within the 15-meter radius buffer;
- (iv) Wrong-match in Junction Error (WJE) represents a trajectory point or two/three adjacent points (P5 in Figure 4d), which are supposed to be located within an intersection, but it is incorrectly matched to a nearby road segment.

Each trajectory of a taxi is evaluated based on the geometric relationship and the corresponding road identifier between adjacent points to extract these four categories of mapmatching errors. All of the extracted errors are combined and included in the error dataset for further analysis.

The specific process of matching error detection is as follows:

- 1. Select the GPS trajectories of a vehicle from the matched Taxi GPS points database based on the plate number and the timestamps.
- 2. Examine the matched road ID of each five or seven adjacent GPS points and check the sequence of the road ID.
- 3. For the errors on road segments, if the first two/three points and the last two/three points share the same road ID, while no road ID is matched to the middle point, i.e., out of the 15-meter buffer zone, it is classified as an "Off Road Error" (ORE). However, if the middle point is matched to a road segment, the road ID of which is different to that of the former two/three and latter two/three points, it is classified as a "Wrong Road Error" (WRE).
- 4. For the errors occurring on intersections, if the former two/three points have the same road ID, and the latter two/three points share another road ID, while no road ID is matched to the middle point, it is classified as an "Off Junction Error" (OJE). However, if the middle point is matched to a road segment, the road ID of which is different to that of the former two/three and latter two/three points, it is classified as a "Wrong Junction Error" (WJE).
- 5. Select the trajectories of another vehicle and repeat steps 1 to 4.



Figure 4. Schematic diagram of four types of error trajectory matching: (**a**) ORE (point P5 is the error trajectory point); (**b**) WRE (point P5 is the error trajectory point); (**c**) OJE (point P5 is the error trajectory point); (**d**) WJE (point P5 is the error trajectory point).

In addition to these four categories, there are other forms of map-matching errors, such as ambiguous lane selection errors, isolated lane selection errors, and route discontinuity errors. However, solutions to these types of map-matching errors are relatively well established [17,40]. Therefore, based on the scope and objectives of our research, we chose to include these four types of errors in the study.

3.3. Spatial-Temporal Characteristics for Trajectory Errors

Kernel Density Estimation (KDE) is a non-parametric statistical method used to estimate the probability density of a variable [41]. In our study, KDE is used to investigate the spatiotemporal patterns of map-matching errors by calculating the spatial density of the errors in a specific spatial-temporal cross-section. The study utilized ArcGIS 10.8 software to conduct Kernel density analysis of map-matching errors. To facilitate the analysis, we transformed the map-matching error dataset from Excel format to a table dataset using the conversion tool in ArcGIS. Then, we converted the table dataset into a layer feature class (LFC) by employing the coordinate notation conversion function in ArcGIS 10.8.

Specifically, each road segment at a specific time of day (e.g., morning, daytime, peak hours, and evening) or during the weekend/weekdays was treated as a spatiotemporal analytical unit in the KDE analysis. The KDE is formulated as follows:

$$f_{\text{hat}(x)} = \left(\frac{1}{nh}\right) \times \sum_{\{i=1\}}^{n} K\left(\frac{x-x_i}{h}\right)$$
(1)

where *n* is the number of variables of interest, *h* is the bandwidth of the kernel function, *K* is the kernel function (e.g., Gaussian or Laplacian), x_i is the *i*th road segment in the data, *x* is the value to be estimated, and $f_{hat(x)}$ provides a smooth estimate of the underlying distribution of the variable *x*.

3.4. Exploration of the Spatial–Temporal Factors

Another primary goal of the study is to explore the contributing factors that affect the occurrence of different kinds of map-matching errors using a multinomial logit model. Let U_{ki} be the utility function [42], so that

$$U_{ki} = \beta_k X_{ki} + \varepsilon_{ki} \tag{2}$$

where β_k is a vector of estimable parameters, and ε_{ki} is the unobserved part that affects the utility and is assumed to follow a Gumbel I extreme value distribution. k = 1, ..., K (K = 4 in our case) represents the categories of the map-matching errors (i.e., ORE, WRE, OJE, and WJE), and X_{ki} represents the contributing factors of the map-matching error *i* in the error category *k*.

Let $P_i(k)$ be the probability of the map-matching error *i* being recognized as category *k*, such that

$$P_i(k) = \frac{\exp(\beta_k X_{ki})}{\sum_{\forall k} \exp(\beta_k X_{ki})}$$
(3)

4. Results

4.1. Outputs of Map Matching

All of the selected taxi trajectories are matched to their nearest road lanes using the HMM map-matching algorithm. Figure 5a,b show the sample outputs of the match on a single segment and a road network, respectively. The red dots represent the raw GPS points from a trajectory, and the green dots represent the matched points. It is observed that the HMM algorithm performs well in matching trajectories on a single segment. All of the trajectory points can be properly matched to the corresponding road lane if the vehicle is moving in a fixed direction. However, map matching on a large-scale road network becomes much more complicated. As shown in Figure 5b, a majority of trajectory points can be correctly matched to the corresponding road lane, but a few points are overtly mismatched when the vehicle is driving on the road segment or passing through the intersection, leading to the occurrence of WREs and WJEs, respectively. In addition, the number of matched trajectory points does not equal the number of original trajectory points, which manifests the existence of OREs or OJEs.



Figure 5. Outputs of map matching: (**a**) matching effect of the single road segment, (**b**) matching effect of the local road network.

Figure 6 illustrates the trajectories before and after map matching, which shows that most of the original trajectories are correctly matched to their corresponding roadways. The success rate of matching reaches 89% for the whole study area, demonstrating a relatively effective matching performance. For the remaining 11% of trajectories with map-

matching errors, we calculate the proportion of the four types of map-matching errors. In Table 4, the total number of OREs, WREs, OJEs, and WJEs in the study area is 175,512, among which there are 113,349 WREs, accounting for 64.6% map-matching errors, followed by OREs (14.1%), OJE (10.8%), and WJE (10.6%). The process of identifying these errors took a total of 4550 s (approximately 75 min).



Figure 6. Matching effect of the large-scale road network: (**a**) before matching in the large-scale road network, (**b**) after matching in the large-scale road network.

Error Matching Trajectory Type	Meaning	Number of Cases	Percentage		
ORE	Road Segment Off	24,702	14.1%		
WRE	Road Segment Error	113,349	64.6%		
OJE	Intersection Off	18,904	10.8%		
WJE	Intersection Error	18,557	10.6%		

Table 4. Statistics of four trajectory error types.

The total time complexity of the HMM map-matching algorithm is $O(N \times M^2)$, and the overall space complexity of the HMM map-matching algorithm is $O(M \times N)$. N represents the number of observed points (trajectory points), and M represents the number of road segments (states) in the road network.

4.2. Temporal and Spatial Distribution of Trajectory Errors

Based on the KDE analysis, Figure 7 illustrates the spatial-temporal distributions of four map-matching errors in different analytical units.

The distributions of OREs (Figure 7(a1–e1)) do not change evidently across times of day and between weekends and weekdays. Specifically, OREs tend to cluster at intersections located in central, eastern, southwestern, and southeastern regions of the road network, indicating that the density of OREs is not consistent in the study area. As for the temporal characteristics, we can observe that daytime and weekdays are associated with more intensive OREs at the above-stated regions, while OREs observed in peak hours, at night, and at the weekend are much sparser. It is also found that the distributions of OJEs (Figure 7(a3–e3)) are almost identical in different time scales, which are clustered in the central, northern, eastern, and southwestern areas of the road network. Similarly, the densities of OJEs in different time scales are analogous to those of OREs.



Figure 7. Density plot of four types of errors.

The clusters of WREs (Figure 7(a2-e2)) are spatially sparser across the road network. It is shown that WREs tend to cluster in the central, southeastern, and southwestern parts of the road network. In the time scale, we identify that WREs in peak hours and weekends are more intensive than those in daytime, at night, and on weekdays. In contrast to WREs, the WJE clusters (Figure 7(a4-e4)) visibly vary across the road network in different time scales. More specifically, WJEs are intensively clustered in the central and southwestern

regions of the road network in the daytime, at peak hours, and on weekdays, while these errors become sparser at night and on weekends. Meanwhile, we reveal that WJEs are prone to cluster in intersections located in the southeastern part at night, on weekdays, and on weekends, but the cluster is not observed in daytime and at peak hours. Furthermore, the densities of WJEs in daytime, peak hours, and weekdays surpasses those at night and on weekends.

4.3. Contributing Factors of Map-Matching Errors

Variance inflation factor (VIF) is used to diagnose multicollinearity among all predictor variables prior to modeling [43]. All factors have a VIF of less than 5, indicating that the model estimates are not explicitly influenced by multicollinearity. A multinomial logistic model is then adopted to explore the contributing factors causing the map-matching errors with the reference category of ORE. Table 5 presents the estimated results. Relative to ORE, there are 26, 23, and 21 factors significantly associated with WRE, OJE, and WJE, respectively.

In terms of time factors, compared to daytime, the probabilities are 8.8% lower, 14.3% lower, and 34.2% higher at night for WRE, OJE, and WJE, respectively. During peak hours, the probability of WRE is 4.0% lower, and the probability of OJE is 5.8% lower. On weekends, the probability of WRE is 5.2% lower than on workdays, while the probability of OJE is 4.8% lower than on workdays.

For intersection types, the probability of WRE is significantly higher if the mapmatching error occurs at a location close to a flyover (Odds Ratio = 225.2%), a crossroad (Odds Ratio = 367.8%), an X-junction (Odds Ratio = 108.7%), or a T-junction (Odds Ratio = 325.1%), as opposed to the errors located outside the vicinity of the intersections. The conclusion is also applicable to WJE (except near the flyover) and OJE (except near the X-junction).

Relative to the previous road section, the probability of WRE is significantly higher if the previous road has bicycle dividers (Odds Ratio = 10.1%), median dividers (Odds Ratio = 53.5%), roadside parking (Odds Ratio = 27.9%), or is one-way controlled (Odds Ratio = 27.8%). The conclusion can be applicable to OJE (except if the previous road has bicycle dividers and median dividers) and WJE (except if the previous road has bicycle dividers, median dividers, and roadside parking). With regard to the speed limit of the previous road, speed limits of <30 km/h and 30–50 km/h reduce the probability of WRE by 64.2% and 10.3%, respectively, compared to roads with a speed limit of \geq 60 km/h, while increasing the probability of OJE by 53.9% and 50.3%, respectively, and reducing the probability of WJE by 30.2% and 32.2%, respectively. Furthermore, an increase of 1 unit per km in resident density raises the possibility of WRE by 4.6%, while a 1-unit-per-km increase in public service density raises the possibility of WRE by 1.9%, while decreasing the possibility of OJE by 3.3%.

Factors related to the latter road show that the probability of WRE is significantly higher if the road has bicycle dividers (Odds Ratio = 63.6%), median dividers (Odds Ratio = 15.6%), and roadside parking (Odds Ratio = 12.5%). Similar findings are observed for OJE and WJE. However, roadside parking decreases the possibility of OJE and WJE. As for speed limit, roads with a speed limit of <30 km/h are associated with significantly reduced probabilities of WRE (65.1%) and OJE (23.1%), but a higher probability of WJE (134.2%), compared to the speed limit of \geq 60 km/h. The speed limit of 30–50 km/h is also linked to a 15.5% reduction in the probability of WRE and a 19.1% reduction in the probability of OJE. Moreover, an increase of 1 unit per km in resident density and public service density is positively associated with higher probabilities of WRE, OJE, and WJE.

		W	RE		OJE				WJE			
Variables	Mean	S.D.	Sig.	OR	Mean	S.D.	Sig.	OR	Mean	S.D.	Sig.	OR
Intercept	-0.078	0.027	0.004		-0.852	0.035	< 0.001		-0.873	0.036	< 0.001	
Time-weekend	-0.054	0.018	0.002	0.948	-0.049	0.023	0.030	0.952	0.042	0.022	0.063	1.042
Time-peak hours	-0.041	0.020	0.035	0.960	-0.060	0.025	0.015	0.942	-0.027	0.026	0.284	0.973
Time-night	-0.092	0.018	< 0.001	0.912	-0.154	0.023	< 0.001	0.857	0.294	0.023	< 0.001	1.342
Time-daytime (set as base)												
Int-close to a flyover	1.179	0.193	< 0.001	3.252	0.703	0.245	0.004	2.021	-2.516	1.017	0.013	0.081
Int road-close to a crossroad	1.543	0.025	< 0.001	4.678	0.780	0.030	< 0.001	2.181	0.826	0.030	< 0.001	2.285
Int-close to an X-junction	0.736	0.063	< 0.001	2.087	0.117	0.080	0.146	1.124	0.336	0.082	< 0.001	1.399
Int-close to a T-junction	1.447	0.019	< 0.001	4.251	0.055	0.024	0.024	1.057	0.287	0.024	< 0.001	1.333
Int-close to a large intersection	-0.044	0.028	0.117	0.957	0.445	0.036	< 0.001	1.560	0.434	0.036	< 0.001	1.543
Int-close to a medium-sized intersection	0.932	0.027	< 0.001	2.540	0.613	0.035	< 0.001	1.846	0.425	0.036	< 0.001	1.529
Int-close to a small intersection	0.985	0.021	< 0.001	2.678	0.353	0.027	< 0.001	1.423	0.242	0.027	< 0.001	1.274
Int-beyond the intersection range (set as base)												
PreRoad-speed limit <30 km/h	-1.027	0.032	< 0.001	0.358	0.431	0.038	< 0.001	1.539	-0.360	0.037	< 0.001	0.698
PreRoad-speed limit 40–50 km/h	-0.108	0.035	0.002	0.897	0.408	0.046	< 0.001	1.503	-0.388	0.047	< 0.001	0.678
PreRoad-speed limit >60 km/h (set as base)												
PreRoad-bicycle divider	0.096	0.029	0.001	1.101	-0.545	0.037	< 0.001	0.580	-0.001	0.036	0.983	0.999
PreRoad-median divider	0.429	0.032	< 0.001	1.535	0.073	0.041	0.079	1.076	-0.437	0.040	< 0.001	0.646
PreRoad-roadside parking	0.246	0.032	< 0.001	1.279	0.179	0.035	< 0.001	1.196	-0.020	0.035	0.573	0.980
PreRoad-one way	0.245	0.041	< 0.001	1.278	0.430	0.043	< 0.001	1.538	0.211	0.046	< 0.001	1.235
PreRoad-commercial density	-0.027	0.002	< 0.001	0.973	-0.008	0.002	< 0.001	0.992	0.004	0.001	0.005	1.004
PreRoad-public service density	0.019	0.004	< 0.001	1.019	-0.033	0.005	< 0.001	0.967	-0.015	0.005	0.002	0.985
PreRoad-resident density	0.045	0.002	< 0.001	1.046	-0.002	0.003	0.466	0.998	-0.011	0.003	< 0.001	0.989
LatRoad-speed limit <30 km/h	-1.051	0.032	< 0.001	0.349	-0.263	0.038	< 0.001	0.769	0.851	0.038	< 0.001	2.342
LatRoad-speed limit 40–50 km/h	-0.168	0.035	< 0.001	0.845	-0.212	0.045	< 0.001	0.809	0.066	0.049	0.174	1.068
LatRoad-speed limit >60 km/h (set as base)												
LatRoad-bicycle divider	0.492	0.029	< 0.001	1.636	0.283	0.036	< 0.001	1.328	-0.351	0.037	< 0.001	0.704
LatRoad-median divider	0.145	0.032	< 0.001	1.156	0.390	0.040	< 0.001	1.477	0.530	0.042	< 0.001	1.699
LatRoad-roadside parking	0.118	0.032	< 0.001	1.125	-0.110	0.037	0.003	0.896	-0.153	0.034	< 0.001	0.858
LatRoad-one way	0.049	0.041	0.233	1.051	-0.018	0.048	0.704	0.982	0.015	0.044	0.738	1.015
LatRoad-commercial density	-0.020	0.002	< 0.001	0.980	-0.001	0.002	0.429	0.999	< 0.001	0.001	0.748	1.000
LatRoad-public service density	0.040	0.004	< 0.001	1.041	0.050	0.005	< 0.001	1.052	0.012	0.005	0.026	1.012
LatRoad-resident density	0.052	0.002	< 0.001	1.053	0.030	0.003	< 0.001	1.031	0.011	0.003	< 0.001	1.011

5. Discussion

It is found that time factors have a significant effect on determining the error types. For instance, WREs and OJEs are more likely to be observed in the daytime and on weekdays. This result is mainly due to the fact that taxis are usually clustered in the urban center during the daytime and on weekdays to seek passengers [44]. Moreover, the city center has a high-density road network, which may be associated with increasing the occurrence of map-matching errors. Specifically, the trajectories of running taxis within this area are more likely to be incorrectly matched to another road (WRE). Also, the taxis have more complex trajectories when they enter the intersections due to the increased traffic volume during the period [45], resulting in more matching losses (OJE). However, WJEs are more likely to be observed at night, which may be interpreted as suggesting that taxis tend to wait or cruise around the city's intersections, hospitals, or transit stations, where the taxi requirements are more intensive [46]. In this case, the trajectories could be incorrectly matched to nearby road segments that are adjacent to intersections, leading to WJEs.

The occurrence of map-matching errors varies across the types and sizes of intersections. Specifically, WREs and OJEs are more likely to occur on flyovers, which is not the case for WJEs. It could be explained by the fact that trajectories are likely to be matched to the ground roads under the flyovers (WRE). Also, the trajectories could be lost on ramps (OJE). However, the trajectories are not likely to be matched to another access (WJE) because of the large size of the flyover. We also found that WRE and WJE are more overtly observed on X- and T-junctions due to the complex movements of vehicles and the difficulty of positioning at these junctions [47,48]. The probabilities of WRE, OJE, and WJE can be interpreted as complex trajectories at medium- and large-sized intersections. Comparatively, the reason for map-matching errors in small-sized intersections may be different. Specifically, the heights and densities of buildings and trees around intersections can affect GPS signal transmission quality [49,50]. As such, the GPS trajectories are more difficult to position within smaller intersections since they are more susceptible to being obstructed by adjacent buildings and trees [51], consequently causing the failure of GPS match (OJE).

As for the characteristics of both previous and latter roads, factors of bicycle dividers, median dividers, and one-way control could increase the possibility of ORE and WRE, which can be interpreted as suggesting that the dividers and one-way control can accelerate the vehicle's speed, meaning that there will be fewer reliable trajectories on the road segment compared to the trajectories generated via low-speed movements. It is also found that the dividers on the latter road could increase the possibility of OJE and WJE because vehicles could easily accelerate or change lanes after they pass the junctions, which causes more OJEs and WJEs. We notice that roadside parking on both previous and latter roads could increase the probability of WREs, which implies that roadside parking could hamper the sight view of the driver and, consequently, encourage them to change lanes or adjust their driving speed. However, the likelihood of junction errors decreases (i.e., OJE and WJE) since vehicles have to slow down if the latter road has roadside parking. For land use, commercial areas lower the possibility of WRE and OJE due to the traffic delay. In contrast, WREs are prone to occur on both previous and latter roads with more public facilities and residents due to the intensive accesses near to the junctions, meaning that the trajectories are likely to be mismatched. This outcome is also the case for OJE and WJE if the latter road has intensive public facilities and residents.

The speed limit of previous and latter road segments is found to significantly influence the map-matching errors. WREs tend to occur on road segments where previous and latter roads have a higher speed limit, which could be explained by the assumption that higher vehicle speeds can lead to deteriorated GPS signal quality [52]. Additionally, vehicles are likely to lose their trajectories (OJE) if they switch from a lower speed limit road to a road with a higher speed limit through a junction. This result occurs because the trajectories on the roads with lower speed limits tend to be more stable, but the vehicle could lose the trajectory signals if it suddenly accelerates and enters a high-speed road. Conversely, the modeling results demonstrate that WJE generally takes place when vehicles move from a road with a high speed limit to a road with a low speed limit. This result may be explained by the fact that vehicles have to slow down in advance before they enter a low speed limit scenario after they enter the junction. Hence, the vehicles could generate a large number of trajectories within the junction, which could be mismatched to other nearby roads rather than being lost.

6. Conclusions

The study identifies four kinds of trajectory map-matching errors (i.e., ORE, WRE, OJE, and WJR) based on the HMM algorithm using taxi trajectories in Chengdu. The study employs temporal Kernel density analysis and a multinomial logistic model to examine the spatial-temporal patterns of the map-matching errors and contributing factors associated with different error likelihoods. Several key findings are offered below:

- 1. The spatial patterns of ORE, WRE, OJE, and WJE overtly vary across the time scales (e.g., time of day and weekday/weekends), signifying that the map-matching errors are not consistently located in the study area.
- Compared to ORE, the probability of WRE and OJE is higher on weekdays, while the probability of WJE is higher on weekends. It is noted that OREs and WJEs are more likely to occur during peak hours and at night.
- 3. WREs, OJEs, and WJEs are more likely to be observed at intersections, especially on a flyover, an X-junction, and a T-junction.
- 4. WREs tend to occur on the road where previous and latter roads simultaneously have bicycle dividers, median dividers, one-way control, and roadside parking, while these factors have mixed impacts on OJEs and WJEs. Also, higher resident and public service density on the latter road could increase the probability of WRE, OJE, and WJE.
- 5. WREs are likely to occur on roads with low-speed limits. OJEs tend to occur when vehicles switch from a low speed limit to a high speed limit road, while the occurrence of WJE has the opposite trend.

There are several limitations to the study. Firstly, we only identified four types of map-matching errors from the taxi trajectories, which may not cover all of the error types in practice. Secondly, we only focus on the effect of time, intersection characteristics, and road features on map-matching errors, but the influence of traffic conditions and drivers' responses are unknown. Thirdly, the primary objective of this paper is to utilize offline GPS data to investigate the spatiotemporal patterns of map-matching errors and examine the effect of road environments on these errors. Therefore, an online map matching system is not considered in the current study.

These three limitations could be overcome if more accurate trajectory data became available. In the future research, we aim to develop advanced algorithms to identify more types of map-matching errors, such as Breakage Error, Ambiguous Match, and Ghost Trajectory Error [17,27]. On the basis of detecting more types of errors, the researchers could gain a more comprehensive understanding of all potential errors that may occur in the scenarios when conducting the map-matching algorithm. Additionally, incorporating real-time traffic data and driver responses will allow us to analyze the influence of traffic conditions and driver-related factors on the distribution of map-matching errors, which can potentially tackle the limitations related to the inadequacy of considering the impact of time, intersection features, and road characteristics and enable comprehensive consideration of traffic situations and driver decision-making processes on the trajectories. We aim to address these challenges and provide more reliable conclusions in a future study.

Author Contributions: Conceptualization, Lin Qu, Yue Zhou, and Xinguo Jiang; methodology, Lin Qu, Yue Zhou, and Jiangxin Li; software, Lin Qu, Yue Zhou, and Jiangxin Li; formal analysis, Lin Qu, Yue Zhou, and Qiong Yu; resources, Yue Zhou and Xinguo Jiang; data curation, Lin Qu, Yue Zhou, and Jiangxin Li; writing—original draft preparation, Lin Qu and Yue Zhou; writing—review and editing, Lin Qu, Yue Zhou, and Xinguo Jiang; supervision, Xinguo Jiang, Yue Zhou, and Jiangxin

Li; project administration, Xinguo Jiang and Yue Zhou; funding acquisition, Lin Qu, Yue Zhou, and Jiangxin Li. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China under Grant NO. 72271207.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Acknowledgments: The authors would like to express their gratitude to the editor and reviewers for their helpful feedback and contributions, as well as for the funding support. Special thanks to Eric Jiang (Vandegrift high school) and his team to help improve the overall language of the paper.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; the collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to publish the results.

References

- 1. Li, L.; Jiang, R.; He, Z.; Chen, X.; Zhou, X. Trajectory Data-Based Traffic Flow Studies: A Revisit. *Transp. Res. Part C Emerg. Technol.* 2020, 114, 225–240. [CrossRef]
- Wang, Y.; Chen, Y.; Li, G.; Lu, Y.; He, Z.; Yu, Z.; Sun, W. City-Scale Holographic Traffic Flow Data Based on Vehicular Trajectory Resampling. Sci. Data 2023, 10, 57. [CrossRef]
- 3. Fan, J.; Li, Y.; Liu, Y.; Zhang, Y.; Ma, C. Analysis of Taxi Driving Behavior and Driving Risk Based on Trajectory Data. *IEEE Intell. Veh. Symp.* **2019**, 2019, 220–225. [CrossRef]
- 4. Dong, X.; Zhang, M.; Zhang, S.; Shen, X.; Hu, B. The Analysis of Urban Taxi Operation Efficiency Based on GPS Trajectory Big Data. *Phys. A Stat. Mech. Its Appl.* **2019**, *528*, 121456. [CrossRef]
- 5. Shahverdy, M.; Fathy, M.; Berangi, R.; Sabokrou, M. Driver Behavior Detection and Classification Using Deep Convolutional Neural Networks. *Expert Syst. Appl.* **2020**, *149*, 113240. [CrossRef]
- 6. Lin, Y.; Li, W.; Qiu, F.; Xu, H. Research on Optimization of Vehicle Routing Problem for Ride-Sharing Taxi. *Procedia Soc. Behav. Sci.* **2012**, *43*, 494–502. [CrossRef]
- Wang, R.; Alazzam, M.B.; Alassery, F.; Almulihi, A.; White, M. Innovative Research of Trajectory Prediction Algorithm Based on Deep Learning in Car Network Collision Detection and Early Warning System. *Mob. Inf. Syst.* 2021, 2021, 3773688. [CrossRef]
- Kan, Z.; Tang, L.; Kwan, M.P.; Ren, C.; Liu, D.; Li, Q. Traffic Congestion Analysis at the Turn Level Using Taxis' GPS Trajectory Data. Comput. Environ. Urban Syst. 2019, 74, 229–243. [CrossRef]
- 9. Kong, X.; Xu, Z.; Shen, G.; Wang, J.; Yang, Q.; Zhang, B. Urban Traffic Congestion Estimation and Prediction Based on Floating Car Trajectory Data. *Futur. Gener. Comput. Syst.* **2016**, *61*, 97–107. [CrossRef]
- 10. Keler, A.; Krisp, J.M.; Ding, L. Detecting Vehicle Traffic Patterns in Urban Environments Using Taxi Trajectory Intersection Points. *Geo-Spatial Inf. Sci.* 2017, 20, 333–344. [CrossRef]
- 11. Qin, J.; Mei, G.; Xiao, L. Building the Traffic Flow Network with Taxi GPS Trajectories and Its Application to Identify Urban Congestion Areas for Traffic Planning. *Sustainability* **2021**, *13*, 266. [CrossRef]
- 12. Qin, H.; Peng, Y.; Zhang, W. Vehicles on RFID: Error-Cognitive Vehicle Localization in GPS-Less Environments. *IEEE Trans. Veh. Technol.* 2017, *66*, 9943–9957. [CrossRef]
- 13. Zhang, Z.; Hu, H.; Yu, Y.; Qian, W.; Shu, K. *Dependency Preserved Raft for Transactions*; Lecture Notes in Computer Science; Springer: Cham, Germany, 2020; Volume 12112, ISBN 978-3-030-59409-1.
- Gunn, L.; Smet, P.; Arbon, E.; McDonnell, M.D. Anomaly Detection in Satellite Communications Systems Using LSTM Networks. In Proceedings of the 2018 Military Communications and Information Systems Conference (MilCIS), Canberra, ACT, Australia, 13–15 November 2018. [CrossRef]
- 15. Koetsier, C.; Fiosina, J.; Gremmel, J.N.; Müller, J.P.; Woisetschläger, D.M.; Sester, M. Detection of Anomalous Vehicle Trajectories Using Federated Learning. *ISPRS Open J. Photogramm. Remote Sens.* **2022**, *4*, 100013. [CrossRef]
- 16. Ranacher, P.; Brunauer, R.; Trutschnig, W.; Van der Spek, S.; Reich, S. GPS Error and Its Effects on Movement Analysis. *Int. J. Geogr. Inf. Sci.* 2015, 30, 316–333. [CrossRef]
- Dey, S.; Tomko, M.; Winter, S. Map-Matching Error Identification in the Absence of Ground Truth. ISPRS Int. J. Geo-Information 2022, 11, 538. [CrossRef]
- Quddus, M.A.; Ochieng, W.Y.; Noland, R.B. Current Map-Matching Algorithms for Transport Applications: State-of-the Art and Future Research Directions. *Transp. Res. Part C Emerg. Technol.* 2007, 15, 312–328. [CrossRef]
- 19. Liu, Z.; Wang, X.D.; Yan, X.Y. Map-Matching Algorithm for GPS Trajectories in Complex Urban Road Networks. *Dianzi Keji Daxue Xuebao J. Univ. Electron. Sci. Technol. China* 2016, 45, 1008–1013. [CrossRef]
- 20. Sharma, K.P.; Poonia, R.C.; Sunda, S. A Novel Map Matching Algorithm for Real-Time Location Using Low Frequency Floating Trajectory Data. *Intell. Paradig.* 2023, 24, 442–455. [CrossRef]
- Li, W.; Wang, Y.; Li, D.; Xu, X. A Robust Map Matching Method by Considering Memorized Multiple Matching Candidates. *Theor. Comput. Sci.* 2023, 941, 104–120. [CrossRef]

- Yu, Q.; Hu, F.; Ye, Z.; Chen, C.; Sun, L.; Luo, Y. High-Frequency Trajectory Map Matching Algorithm Based on Road Network Topology. *IEEE Trans. Intell. Transp. Syst.* 2022, 23, 17530–17545. [CrossRef]
- 23. Hu, H.; Qian, S.; Ouyang, J.; Cao, J.; Han, H.; Wang, J.; Chen, Y. AMM: An Adaptive Online Map Matching Algorithm. *IEEE Trans. Intell. Transp. Syst.* 2023, 24, 5039–5051. [CrossRef]
- 24. Jiang, L.; Chen, C.X.; Chen, C. L2MM: Learning to Map Matching with Deep Models for Low-Quality GPS Trajectory Data. ACM Trans. Knowl. Discov. Data 2023, 17, 39. [CrossRef]
- 25. Feng, J.; Li, Y.; Zhao, K.; Xu, Z.; Xia, T.; Zhang, J.; Jin, D. DeepMM: Deep Learning Based Map Matching With Data Augmentation. *IEEE Trans. Mob. Comput.* 2022, 21, 2372–2384. [CrossRef]
- Tortora, M.; Cordelli, E.; Soda, P. PyTrack: A Map-Matching-Based Python Toolbox for Vehicle Trajectory Reconstruction. *IEEE Access* 2022, 10, 112713–112720. [CrossRef]
- Saki, S.; Hagen, T. A Practical Guide to an Open-Source Map-Matching Approach for Big GPS Data. SN Comput. Sci. 2022, 3, 415. [CrossRef]
- Ma, S.; Lee, H. A Practical HMM-Based Map-Matching Method for Pedestrian Navigation. Int. Conf. Inf. Netw. 2023, 2023, 806–811. [CrossRef]
- 29. Jiang, L.; Chen, C.; Chen, C.; Huang, H.; Guo, B. From Driving Trajectories to Driving Paths: A Survey on Map—Matching Algorithms. *CCF Trans. Pervasive Comput. Interact.* 2022, 4, 252–267. [CrossRef]
- 30. Member, R.S.; Wang, G.; Cheng, Q.; Fu, L.; Chiang, K.; Hsu, L.-T.; Ochieng, W.Y. Improving GPS Code Phase Positioning Accuracy in Urban Environments Using Machine Learning. *IEEE Internet Things J.* **2020**, *8*, 7065–7078. [CrossRef]
- Chao, P.; Xu, Y.; Hua, W.; Zhou, X. A Survey on Map-Matching Algorithms. In Databases Theory and Applications, Proceedings of the 31st Australasian Database Conference, ADC 2020, Melbourne, VIC, Australia, 3–7 February 2020; Lecture Notes in Computer Science; Springer: Cham, Germany, 2020; Volume 12008, pp. 121–133. [CrossRef]
- Luo, S.; Gu, F.; Xu, F.; Shang, J. Effect Evaluation of Spatial Characteristics on Map Matching-Based Indoor Positioning. Sensors 2020, 20, 6698. [CrossRef]
- Wang, J.; Yuan, Y.; Ni, T.; Ma, Y.; Liu, M.; Xu, G. Anomalous Trajectory Detection and Classification Based on Difference and Intersection Set Distance. *IEEE Trans. Veh. Technol.* 2020, 69, 2487–2500. [CrossRef]
- 34. Cervantes, J.; Garcia-lamont, F.; Rodríguez-mazahua, L.; Lopez, A. Neurocomputing A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing* **2020**, *408*, 189–215. [CrossRef]
- Talebi, H.; Peeters, L.J.M.; Otto, A. A Truly Spatial Random Forests Algorithm for Geoscience Data Analysis and Modelling. Math. Geosci. 2022, 54, 1–22. [CrossRef]
- Singh, J.; Singh, S.; Singh, S.; Singh, H. Evaluating the Performance of Map Matching Algorithms for Navigation Systems: An Empirical Study. Spat. Inf. Res. 2019, 27, 63–74. [CrossRef]
- Santi, P.; Resta, G.; Szell, M.; Sobolevsky, S.; Strogatz, S.H.; Ratti, C. Quantifying the Benefits of Vehicle Pooling with Shareability Networks. Proc. Natl. Acad. Sci. USA 2014, 111, 13290–13294. [CrossRef]
- Brühwiler, L.; Fu, C.; Huang, H.; Longhi, L.; Weibel, R. Computers, Environment and Urban Systems Predicting Individuals' Car Accident Risk by Trajectory, Driving Events, and Geographical Context. *Comput. Environ. Urban Syst.* 2022, 93, 101760. [CrossRef]
- Fu, X.; Zhang, J.; Zhang, Y. An Online Map Matching Algorithm Based on Second-Order Hidden Markov Model. J. Adv. Transp. 2021, 2021, 9993860. [CrossRef]
- Hsueh, Y.L.; Chen, H.C. Map Matching for Low-Sampling-Rate GPS Trajectories by Exploring Real-Time Moving Directions. *Inf. Sci.* 2018, 433–434, 55–69. [CrossRef]
- Xue, T.; Zhong, M.; Luo, L.; Li, L.; Ding, S.X. Distributionally Robust Fault Detection by Using Kernel Density Estimation. *IFAC Pap.* 2020, 53, 652–657. [CrossRef]
- 42. Çelik, A.K.; Oktay, E. A Multinomial Logit Analysis of Risk Factors Influencing Road Traffic Injury Severities in the Erzurum and Kars Provinces of Turkey. *Accid. Anal. Prev.* 2014, 72, 66–77. [CrossRef] [PubMed]
- 43. O'Brien, R.M. A Caution Regarding Rules of Thumb for Variance Inflation Factors. Qual. Quant. 2007, 41, 673–690. [CrossRef]
- 44. Wong, R.C.P.; Szeto, W.Y. An Alternative Methodology for Evaluating the Service Quality of Urban Taxis. *Transp. Policy* **2018**, 69, 132–140. [CrossRef]
- 45. Dong, Y.; Xu, J.; Liu, X.; Gao, C.; Ru, H.; Duan, Z. Carbon Emissions and Expressway Traffic Flow Patterns in China. *Sustainability* **2019**, *11*, 2824. [CrossRef]
- Ulak, M.B.; Yazici, A.; Aljarrah, M. Value of Convenience for Taxi Trips in New York City. *Transp. Res. Part A Policy Pract.* 2020, 142, 85–100. [CrossRef]
- Schneider, R.J.; Diogenes, M.C.; Arnold, L.S.; Attaset, V.; Griswold, J.; Ragland, D.R. Association between Roadway Intersection Characteristics and Pedestrian Crash Risk in Alameda County, California. *Transp. Res. Rec.* 2010, 2198, 41–51. [CrossRef]
- Guo, F.; Wang, X.; Abdel-Aty, M.A. Modeling Signalized Intersection Safety with Corridor-Level Spatial Correlations. *Accid. Anal. Prev.* 2010, 42, 84–92. [CrossRef]
- 49. Liu, X.; Sun, L.; Sun, Q.; Gao, G. Spatial Variation of Taxi Demand Using GPS Trajectories and POI Data. J. Adv. Transp. 2020, 2020, 7621576. [CrossRef]

- Zhou, Y.; Fang, Z.; Thill, J.; Li, Q.; Li, Y. Computers, Environment and Urban Systems Functionally Critical Locations in an Urban Transportation Network: Identification and Space—Time Analysis Using Taxi Trajectories. *Comput. Environ. Urban Syst.* 2015, 52, 34–47. [CrossRef]
- Hannover, L.U.; Braunschweig, T.U. Vehicle-to-Vehicle IEEE 802.11p Performance Measurements at Urban Intersections. In Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2022; pp. 7131–7135.
- 52. Deep, S.; Raghavendra, S.; Bharath, B.D. RESEARCH PAPER GPS SNR Prediction in Urban Environment. *Egypt. J. Remote Sens. Sp. Sci.* **2018**, *21*, 83–85. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.