

Article

Multi-Scale Road Matching Based on the Summation Product of Orientation and Distance and Shape Descriptors

Ying Sun ¹, Yuefeng Lu ^{1,2,3,*} , Ziqi Ding ¹, Qiao Wen ⁴, Jing Li ¹, Yanru Liu ¹ and Kaizhong Yao ¹

¹ School of Civil Engineering and Geomatics, Shandong University of Technology, Zibo 255049, China; 21407010769@stumail.sdut.edu.cn (Y.S.); 22407010009@stumail.sdut.edu.cn (Z.D.); 21507020777@stumail.sdut.edu.cn (J.L.); 21407010764@stumail.sdut.edu.cn (Y.L.); 21507020796@stumail.sdut.edu.cn (K.Y.)

² State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

³ Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying, Mapping and Remote Sensing, Hunan University of Science and Technology, Xiangtan 411201, China

⁴ Chongqing Road Secondary School, Zibo 255000, China; 14120906066@stumail.sdut.edu.cn

* Correspondence: yflu@sdut.edu.cn; Tel.: +86-0533-278-0964

Abstract: Most commonly used road-based homonymous entity matching algorithms are only applicable to the same scale, and are weak in recognizing the one-to-many and many-to-many types that are common in matching at different scales. This paper explores model matching for multi-scale road data. By considering the sources of various scales and landmark datasets, as well as the spatial relationships between the selected objects and the detailed features of the entities, we propose an improved matching metric, the summation product of orientation and distance (SOD), combined with the shape descriptor based on feature point vectors, the shape area descriptor based on the minimum convex hull, and three other indicators, to establish multiple multi-scale road matching models. Through experiments, the comprehensive road matching model that combines SOD, orientation, distance and length is selected in this paper. When matching the road dataset with a scale of 1:50,000 and 1:10,000, the precision, recall, and F-score of the matching result of this model reached 97.31%, 94.33%, and 95.8%, respectively. In the case that the scale of the two datasets did not differ much, we concluded that the model can be used for matching between large-scale road datasets.

Keywords: SOD; shape descriptor; multi-scale vector spatial data; data matching; landmark extraction



Citation: Sun, Y.; Lu, Y.; Ding, Z.; Wen, Q.; Li, J.; Liu, Y.; Yao, K. Multi-Scale Road Matching Based on the Summation Product of Orientation and Distance and Shape Descriptors. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 457. <https://doi.org/10.3390/ijgi12110457>

Academic Editors: Peng Peng, Shu Wang, Maryam Lotfian, Feng Lu, Yunqiang Zhu and Wolfgang Kainz

Received: 16 September 2023
Revised: 2 November 2023
Accepted: 7 November 2023
Published: 8 November 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/>).

1. Introduction

Map data at different scales express the morphology, structure, and details of geographical spatial phenomena or entities across different levels of detail. The lack of association between differently scaled spatial data is a significant inconvenience to its application [1]. For example, in the application of different scale road network maps, small-scale road networks often lack information from large-scale road networks, or the two cannot accurately correspond, and the lack of correspondence between the two causes difficulties in the applications of the maps. As spatial data are becoming abundant, improving both the reuse and the quality of existing spatial data has become more important than ever. There is now an urgent need to develop approaches for spatial data integration and updating, and object matching has become one of the most productive solutions for data integration and updating between multi-scale maps [2,3]. Object matching is based on the similarity of geometry, topology, and semantics between multi-scale objects [4]. Currently, geometric matching, topological matching, and semantic matching are used as the basis for entity discrimination in vector spatial data matching methods. For the matching of road network vector data, we usually consider line feature matching.

Line features typically include roads and water systems, and are typically matched through geometric features, including distance, shape, and angle.

In terms of road matching using geometric and other similarity metrics: Deng et al. [5] extended the traditional Hausdorff distance to make it more robust and noise resistant, and were able to determine the matching pattern of line elements. The Frechet distance is more in line with people's spatial cognition of curve distance. Generally, the vertex of a polyline is used to approximate the Frechet distance of a curve. Eiter et al. [6] proposed a method for calculating discrete Frechet distances; Alireza et al. [7] also proposed a geometric-based road matching method using genetic algorithms for multi-scale datasets. This method uses real coded genetic algorithms (RCGA) and sensitivity analysis, which eliminates initial dependencies on empirical parameters such as the buffer distance, spatial similarity thresholds, and standard weights. Govindaraj and Sudhakar [8] proposed a feature description scheme based on texture energy measurement. HUYH et al. [9] used proximity with the Hausdorff distance to distinguish matching relationships between faces. Filin et al. [10] proposed performing initial matching on nodes based on characteristic information such as the location of nodes in the road network and the angle of the associated guard, and on this basis, matched the associated guard of nodes. Arkin et al. [11] proposed to initially match nodes based on characteristic information such as the location, topology, and angle of the associated arc segments in the road network, and then match the arc segments associated with the nodes on this basis. A global optimization matching method for multi-representation buildings using road network constraints was proposed [12]. The spatial similarity among the candidate matching objects was calculated and the characteristic similarity weights were determined using the entropy weight method. The matching accuracy was greatly improved compared with the local search strategy. Aiming at the problems of low matching accuracy and slow matching speed of high-frequency trajectory data in complex urban road networks, Wang, H. et al. [13] proposed a matching method based on path increment. Li, B. [14] proposed a novel MM algorithm with semantic fusion from vehicle-borne images (VIS-MM) suited to the parallel road scenes. It can be applied into the fields of unmanned autonomous navigation and crowdsourcing updating of high-definition maps.

In terms of algorithmic improvements for matching data from multiple sources: Wu, J. et al. [15] presented a general approach using the Voronoi diagram for spatial entity matching on multi-scale datasets. Wang, S. et al. [16] analyzed existing algorithms used for vector network matching to develop an improved matching algorithm that can adapt to underground pipeline networks. Zhang, J. et al. [17] proposed an improved probabilistic relaxation method, considering both local and global optimizations for the matching of multi-scale of road networks. Wan, B. et al. [18] developed a particle-swarm optimization (PSO)-based parallel road-network matching method on a graphics-processing unit (GPU). Chehrehghan, A. et al. [19] investigated the efficiency of the most important and applicable spatial distances (13 types of distance methods) in vector datasets with different scales and sources. The results indicated that the short-line median and mean Hausdorff methods achieved higher efficiencies compared to the other employed methods. Guo, Q. et al. [20] proposed a combined stroke-based matching approach of road networks considering the constraints of cartographic generalization for road networks under different scales. Lewandowicz, E. et al. [21] proposed an algorithm for generating the centerline of an elongated polygon based on the transformation of vector data. The proposed method involved the determination of base points denoting the direction of river flow. A new Voronoi diagram-based approach for matching multi-scale road networks (VAMRN) was proposed and demonstrated that the VAMRN outperformed two existing methods in generality and matching quality [22]. Singh, S. et al. [23] briefly explained the category-wise working of map matching algorithms and also provided analytically reviews of the performance of these algorithms. They concluded that for online map matching, the hidden Markov model-based map matching algorithm provided good accuracy in comparison to other considered algorithms.

In terms of matching road network data through systems or network models, etc.: Zhao, L. et al. [24] described how road traffic data from different sources can be integrated into the platform's common data model to enable navigation applications. Xiaorui Yan et al. [25]

proposed a framework for analyzing the match between traffic flux, i.e., the number of individuals driving into or out of a region per unit time, and road resources, using mobile phone data covering approximately 17 million users over one week in Beijing. Ruozhen Cheng et al. [26] proposed a location conversion method for roads, wherein a road ontology was designed to model the interior direction knowledge of the roads, giving a deep learning-based road semantic matching model to support rapid location conversion between roads based on efficient semantic queries. Wang Y. et al. [27] proposed a new multi-scale dynamic matching algorithm based on a hierarchical stroke mesh (HSM) to detect matches between OSM data and professional surveying and mapping data, and to update change information. Li, W. et al. [28] presented Multiple Candidate Matching (MCM) to improve the robustness of map matching by using historical trajectory data. MCM tracks multiple road candidates in the map matching process, while limiting the number of road candidates by excluding the routes whose likelihoods are below a threshold. Because the inertial navigation system does not have the problem of signal hopping or missing, the traditional GPS matching algorithms could not work well as usual. A vector road aided inertial navigation using the ICCP algorithm was proposed, with the features of the vehicle trajectory and the corresponding matching road [29]. Shen, L. et al. [30] proposed a YOLOv3 (You Only Look Once v3)-based method aimed at enhancing the capability of cross-scale detection and focusing on the valuable area. The proposed method filled an urgent need for multi-scale detection, and its individual components will be useful in road object detection. In order to improve the trustworthiness of road condition detection, a real-time artificial intelligence road detection system based on binocular vision sensors was investigated [31]. The system was deployed on the low-power edge computing platform, which can upload the processing results to the cloud through the Internet-of-Things devices. Lei, T. et al. [32] came up with a method of generating high-definition map models based on the geographic information system (GIS). The proposed method provided an efficient way of extracting lane-level information from urban road networks and can be applied for lane-level map matching with good performance. Ma, S. et al. [33] proposed an algorithm to determine the initial probabilities of hidden states using a small number of GPS measurements. It can effectively determine the initial road segment compared with the traditional HMM-based map-matching methods and increase the accuracy of pedestrian map-matching.

To summarize the current state of research, the types of differences in roads with the same geographical entity in different scales can be divided into geometric differences, topological differences, semantic differences and other differences. Among them, other differences include coordinate system differences, data model differences, and data format differences. The properties of different scale data often lack uniquely identifiable information. Thus, the geometric matching method is generally adopted to identify homonymous entities [34,35]. And the geometric differences in road datasets of different scales can be classified into various types of differences such as spatial location type, orientation type, shape type, and length type differences.

In conclusion, in this paper, seven comprehensive metric models based on multiple similarity metrics such as length, direction, angle and shape descriptors are constructed to address the problem of geometric differences in multi-scale scale data. We conducted experiments on multi-scale data to explore the best model for multi-scale data matching, and provide a comparable method for data fusion and updating.

2. Methodology

2.1. The Summation Product of Orientation and Distance

The SOD descriptor is based on the spatial relationship between roads and landmarks for similarity calculation. The calculation consisted of the following two steps:

1. Extract landmarks;
2. Calculate the spatial relationship between landmarks and linear objects.

2.1.1. Extraction of Landmarks

The ROD descriptor [36] identifies dataset landmarks at different scales and sources by calculating the spatial similarity of each intersection. After identifying the dataset landmarks the spatial relationship between the landmarks and the target road is calculated to identify the same geographic entities in the datasets at different scales. The ROD descriptor repeats the calculation of the similarity of the spatial relationship. The improvement of the SOD descriptor proposed in this paper compared with the ROD descriptor is that it simplifies the extraction step of landmarks and the calculation process. We extracted landmarks by the number of road nodes in the specific case of different scale road datasets.

In this paper, two sets of road datasets were experimented with; namely, datasets with scales of 1:250,000 and 1:50,000 and datasets with scales of 1:50,000 and 1:10,000. Datasets with scales of 1:250,000 and 1:50,000 were used as reference datasets. When extracting the landmarks in the dataset with a scale of 1:250,000, road intersections with a number of nodes greater than two were directly selected as landmarks due to the small scale and the small number of road strips. The larger the scale, the more detailed the roads in the dataset. In the dataset with a scale of 1:50,000, we extracted road intersections with a node count greater than three as landmarks in areas with dense roads, and road intersections with a node count greater than two as landmarks in areas with sparse roads. When the scale is large, the roads in the road dataset are mostly main roads. When the scale is small, the roads in the road dataset can be divided into urban and rural roads. In this paper, we defined urban roads as dense roads and rural roads as sparse roads.

2.1.2. Calculation of SOD

The SOD's descriptor is obtained by the product of the direction and distance of the space vector object and landmark. To calculate the relationship between an object and a landmark, the representative node of the line vector object must be extracted to measure the spatial relationship between the object and landmark. The direction and distance of the space vector object and landmark are depicted in Figure 1. The relationship between the vector object $PL_{1,i}$, $PL_{2,j}$ and the landmark $I_{1,i}$, $I_{2,j}$ is represented by the direction and distance; $PL_{1,i}$ and $PL_{2,j}$ represent the candidate object at two scales, while $I_{1,1}$ and $I_{2,1}$ represent the homonymous landmarks at two scales.

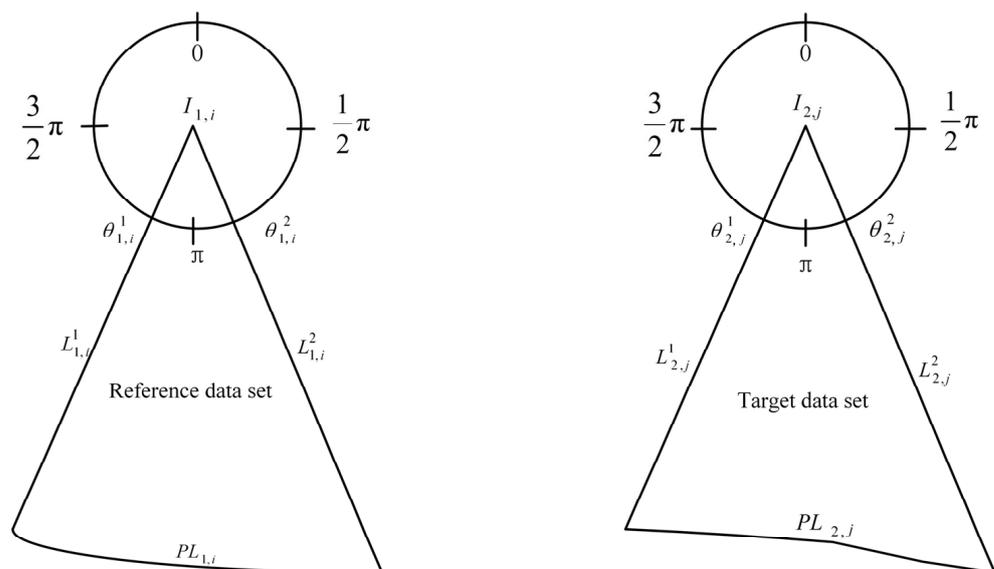


Figure 1. Diagram of the SOD's direction and distance.

As shown in Figure 1, the spatial relationship between candidate objects $PL_{1,i}$ and $PL_{2,j}$ and the landmarks in the two datasets could be calculated by the following formula:

$$F_1 = \sum_{m=1}^n SOD_m(PL_{1,i}, PL_{2,j}) \quad (1)$$

where $SOD_m(PL_{1,i}, PL_{2,j})$ is calculated by Formula (2), F_1 represents the product sum of the direction and distance of the m -th landmark in the two datasets, and n is the total number of landmarks.

$$SOD_m(PL_{1,i}, PL_{2,j}) = dL_i^1 \times d\theta_i^1 + dL_j^2 \times d\theta_j^2 \quad (2)$$

In the formula above, dL_i^1 and dL_j^2 are calculated by Formulas (3) and (4), respectively, while $d\theta_i^1$ and $d\theta_j^2$ are calculated by Formulas (5) and (6), respectively.

$$dL_i^1 = |L_{1,i,m}^1 - L_{1,i,m}^2| \quad (3)$$

$$dL_j^2 = |L_{2,j,m}^1 - L_{2,j,m}^2| \quad (4)$$

$$d\theta_i^1 = \begin{cases} |\theta_{1,i,m}^1 - \theta_{1,i,m}^2|, & \text{if } |\theta_{1,i,m}^1 - \theta_{1,i,m}^2| \leq \pi \\ 2\pi - |\theta_{1,i,m}^1 - \theta_{1,i,m}^2|, & \text{if } |\theta_{1,i,m}^1 - \theta_{1,i,m}^2| > \pi \end{cases} \quad (5)$$

$$d\theta_j^2 = \begin{cases} |\theta_{2,j,m}^1 - \theta_{2,j,m}^2|, & \text{if } |\theta_{2,j,m}^1 - \theta_{2,j,m}^2| \leq \pi \\ 2\pi - |\theta_{2,j,m}^1 - \theta_{2,j,m}^2|, & \text{if } |\theta_{2,j,m}^1 - \theta_{2,j,m}^2| > \pi \end{cases} \quad (6)$$

In the above formulas, $L_{1,i,m}^1$ and $L_{1,i,m}^2$ refer to the distance between the first and second nodes of the line object $PL_{1,i}$ and the m -th landmark in the first dataset, respectively. $L_{2,j,m}^1$ and $L_{2,j,m}^2$ refer to the distance between the first and second nodes of the line object $PL_{2,j}$ and the m -th landmark of the second dataset. $\theta_{1,i,m}^1$ and $\theta_{1,i,m}^2$ refer to the azimuth angle of the first and second nodes, respectively, of the line object $PL_{1,i}$ connected to the m -th landmark in the first dataset. $\theta_{2,j,m}^1$ and $\theta_{2,j,m}^2$ refer to the azimuth angle of the first and second nodes, respectively, of the line object $PL_{2,j}$ that is connected with the m -th landmark in the second dataset.

2.2. Shape Descriptors

Shape matching of geospatial vector graphic elements can not only integrate, fuse, and update multi-source datasets on a regular basis, but also help maintain the integrity and current situation of a spatial database [7]. For the shape area factor, the area of the polygon formed by connecting the head and tail nodes can generally be used as the shape area descriptor for line entity matching [8]. Geospatial data, as a digital carrier in the real world, has different manifestations and features at different scales [9]. The number and location of road feature points at different scales are important considerations for matching calculations.

2.2.1. Shape Area Descriptor Based on the Minimum Convex Hull

In order to improve the accuracy of geometric matching, we considered the spatial relationship between linear object feature points and the minimum convex hull. In addition to using other general standards, a shape area descriptor based on the minimum convex hull was introduced, which described the shape of the road by extracting road feature points at different scales to generate the minimum convex hull of the road segment and

the actual length of the road. For a set X in a real vector space, its convex hull S can be represented by the linear combination (x_1, x_2, \dots, x_n) of all points in X :

$$S = \left\{ \sum_{j=1}^n t_j x_j \mid x_j \in X, \sum_{j=1}^n t_j = 1, t_j \in [0, 1] \right\} \quad (7)$$

As shown in Figure 2, for a road line L , the ordered point set is expressed as $O = \{P_i = (x_i, y_i), i = 1, 2, \dots, n\}$, where n is the feature point number on the line, and there are numerous convex polygons that can completely surround this point set. The convex hull is the smallest convex polygon among them. We found the point P_1 with the smallest ordinate in the convex hull. If the ordinates of the points were consistent, we found the point with the minimum discrete point. Then, using point P_1 as a reference point, we connected P_1 with other points inside the convex hull, and calculated the angle between the generated line segment and the horizontal line in a counterclockwise direction. We sorted the feature points by the angle size, and if the angle size was consistent, we sorted them by distance size. We assumed that the sorted point sequence was P_1, P_2, \dots, P_n , and connected all points in sequence to obtain a polygon. It can be seen that P_1, P_2, \dots, P_n are points on the convex hull boundary. According to the definition of a convex hull put forward by Graham, "each vertex of a convex polygon must be on the same side of any side of the polygon", the non-convex hull vertex in the boundary sequence was deleted, and the minimum convex hull T of the road was finally obtained. The ratio of the minimum convex hull of the road to the length of the road was used as the measurement factor for the shape similarity evaluation of multi-scale road data. We set the linear entities PL_1 and PL_2 , and used the following formula to calculate the difference value between the shape area descriptors of the two linear entities based on the minimum convex hull.

$$F_2 = \Delta_{\text{convex}} = \left| \frac{T_1}{D_1} - \frac{T_2}{D_2} \right| \quad (8)$$

where T_1 and T_2 are the minimum convex hulls constructed from the ordered point sets of PL_1 and PL_2 , respectively, while D_1 and D_2 are the lengths of PL_1 and PL_2 , respectively.

2.2.2. Shape Descriptor Based on the Feature Point Vector

The key to calculating the shape similarity is the method of describing the shape. Due to the complexity of spatial features, the contour boundary description of geographical entities is extremely important for the shape description of overly complex features. For cross scale vector data matching, we proposed a shape descriptor based on feature point vectors to more accurately describe the shape of road vector lines at various scales.

As shown in Figure 3, for a road line L , the vector line was converted into a directional process to extract each feature point of the road. We used an ordered point set to express L as $S = \{P_i = (x_i, y_i), i = 1, 2, \dots, n\}$, where n is the number of feature points on the line, the curve length of the road between any feature point P_i and P_{i+1} is expressed as r_i , and the azimuth angle θ_i , and where P_i is located is defined as the azimuth angle between it and the next adjacent feature point P_{i+1} . The formula is:

$$\theta_i = \pi - \frac{1}{2}\pi \cdot \text{sgn}x(y_i - y_{i+1}) - \arctan((x_i - x_{i+1}) / (y_i - y_{i+1})) \quad (9)$$

The geometric shape curve of the road route was described by a coordinate vector group, as follows:

$$G = \{g_i = (r_i, \theta_i), i = 1, 2, \dots, n\} \quad (10)$$

where n is the number of feature points on the line. For the tail node of any road, the default representation is $g_1 = (0, 0)$.

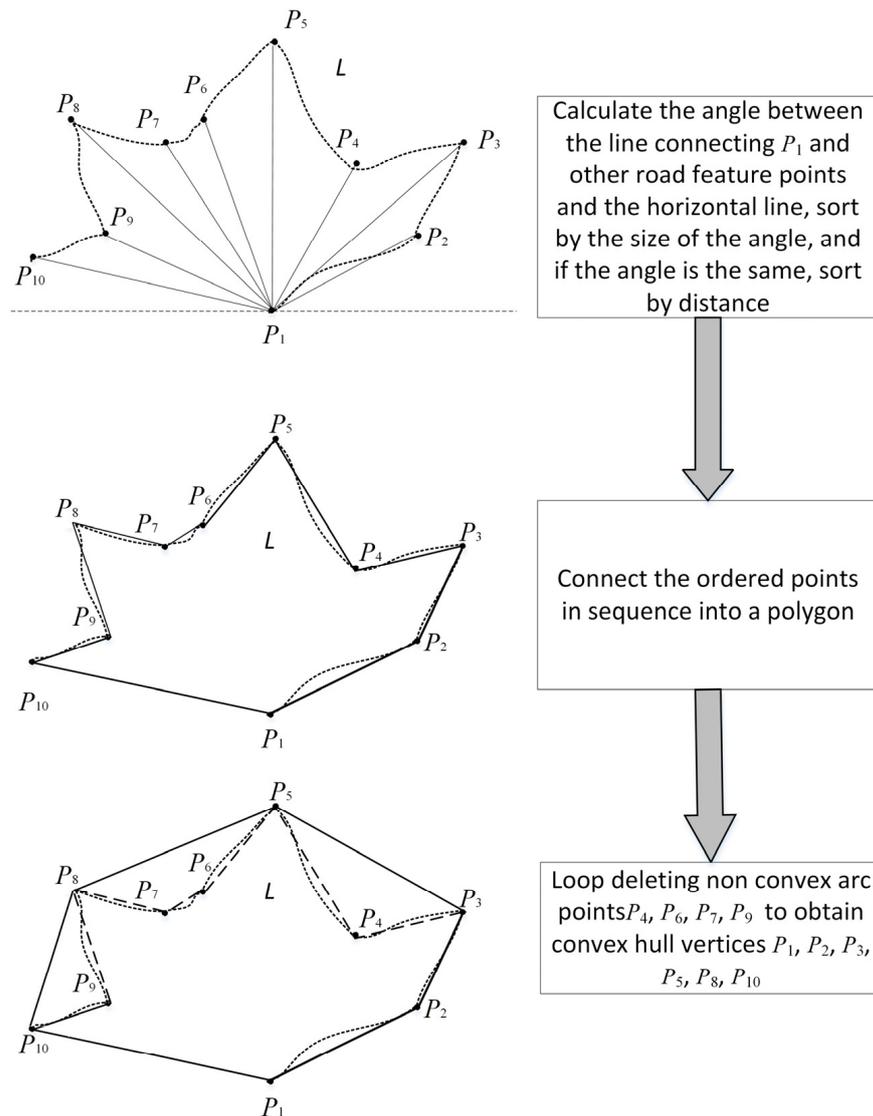


Figure 2. Construction principle of minimum convex hulls for each route.

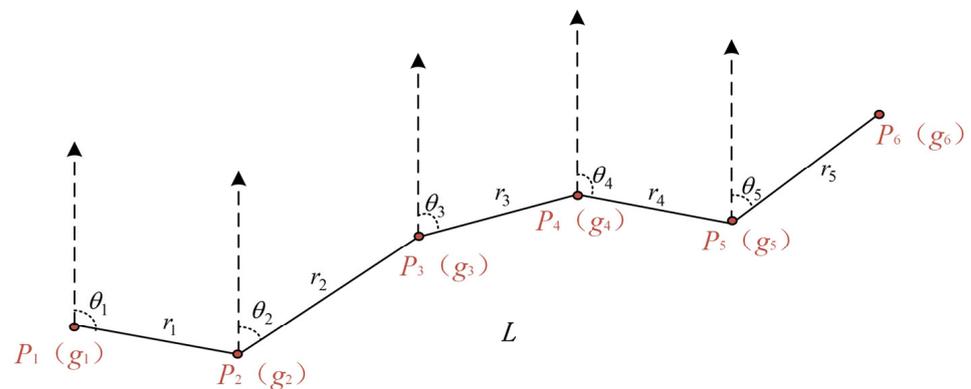


Figure 3. Shape descriptor based on feature point vector.

The similarity of vectors is usually calculated using cosine similarity or Euclidean distance. Due to certain limitations of Euclidean distance, the difference in values between different variables could not be calculated this way. Generally, small differences in variables can lead to large deviations in the results. The Pearson correlation coeffi-

cient is a method for processing different variables, which is an improvement on the square of the Euclidean distance. Assuming that there are two road routes, expressed as $G_1 = \{g_{1i} = (r_{1i}, \theta_{1i}), i = 1, 2, \dots, m\}$ and $G_2 = \{g_{2i} = (r_{2i}, \theta_{2i}), i = 1, 2, \dots, n\}$ using feature point-based shape vectors, the formula for calculating the feature point-based shape vector descriptor was as follows:

$$\left\{ \begin{array}{l} \rho_{G_1 G_2} = 2 \times \frac{\rho_{r_1 r_2} \times \rho_{\theta_1 \theta_2}}{\rho_{r_1 r_2} + \rho_{\theta_1 \theta_2}} \\ \rho_{r_1 r_2} = \frac{N \sum_{i=1}^N r_{1i} r_{2i} - \sum_{i=1}^N r_{1i} \sum_{i=1}^N r_{2i}}{\sqrt{N \sum_{i=1}^N r_{1i}^2 - (\sum_{i=1}^N r_{1i})^2} \sqrt{N \sum_{i=1}^N r_{2i}^2 - (\sum_{i=1}^N r_{2i})^2}} \\ \rho_{\theta_1 \theta_2} = \frac{N \sum_{i=1}^N \theta_{1i} \theta_{2i} - \sum_{i=1}^N \theta_{1i} \sum_{i=1}^N \theta_{2i}}{\sqrt{N \sum_{i=1}^N \theta_{1i}^2 - (\sum_{i=1}^N \theta_{1i})^2} \sqrt{N \sum_{i=1}^N \theta_{2i}^2 - (\sum_{i=1}^N \theta_{2i})^2}} \end{array} \right. \quad (11)$$

$$F_3 = \rho_{G_1 G_2} \quad (12)$$

When the linear correlation between G_1 and G_2 increases, $|\rho_{G_1 G_2}|$ tends to be 1.

2.3. Similarity Indicators for Length, Orientation and Distance

The SOD, shape descriptor based on the feature point vector, shape area descriptor based on the minimum convex hull [37], length [38], short-line median Hausdorff distance [5,39], and orientation [37] were weighted and synthesized in the form of a vector group to build the comprehensive spatial similarity measurement model. Among the six similarity metrics mentioned, length, orientation and distance can be directly calculated by the GIS module. However, the shape descriptor based on the minimum convex hull, the shape descriptor based on the feature point vector, and the SOD descriptor involve more complex computational problems and need to be computed through programming. F_4 , F_5 , and F_6 were measurement indices for length, distance, and orientation, respectively. The calculations for F_4 , F_5 , and F_6 are as follows:

$$F_4 = \sum_{i=1}^{m-1} \left(\sqrt{(x_{P_{i+1}} - x_{P_i})^2 + (y_{P_{i+1}} - y_{P_i})^2} \right) \quad (13)$$

where x_{P_i} and y_{P_i} are the coordinates of the i th node.

$$F_5 = \begin{cases} m(PL_{1,i}, PL_{2,j}), & \text{if } L_{PL_{1,i}} < L_{PL_{2,j}} \\ m(PL_{2,j}, PL_{1,i}), & \text{if } L_{PL_{1,i}} \geq L_{PL_{2,j}} \end{cases} \quad (14)$$

where $L_{PL_{1,i}}$ and $L_{PL_{2,j}}$ are the lengths of two linear objects, $PL_{1,i}$ and $PL_{2,j}$, respectively. $m(PL_{1,i}, PL_{2,j})$ and $m(PL_{2,j}, PL_{1,i})$ are calculated by the following equations:

$$m(PL_{1,i}, PL_{2,j}) = \text{median}_{P_a \in PL_{1,i}} \left\{ \min_{P_b \in PL_{2,j}} P_a - L_b \right\} \quad (15)$$

$$m(PL_{2,j}, PL_{1,i}) = \text{median}_{P_b \in PL_{2,j}} \left\{ \min_{P_a \in PL_{1,i}} P_b - L_a \right\} \quad (16)$$

where L_a and L_b are two arbitrary edges from the linear objects $PL_{1,i}$ and $PL_{2,j}$, respectively; $P_a - L_b$ is the perpendicular distance between a point on object $PL_{1,i}(P_a)$ and one of the

edges of object $PL_{2,j}(L_b)$, and $P_b - L_a$ is the perpendicular distance between a point on object $PL_{2,j}(P_b)$ and one edge of object $PL_{1,i}(L_a)$.

$$F_6 = |\alpha - \beta| = \cos^{-1}\left(\frac{\vec{v}_{PL_{1,i}} \cdot \vec{v}_{PL_{2,j}}}{\left|\vec{v}_{PL_{1,i}}\right| \cdot \left|\vec{v}_{PL_{2,j}}\right|}\right) \quad (17)$$

where $\vec{v}_{PL_{1,i}}$ and $\vec{v}_{PL_{2,j}}$ are the vectors that comprise the starting and ending nodes of the first and second object, respectively, and operator $\|\cdot\|$ is the Euclidean distance between the starting and ending nodes of the considered object.

3. Technical Flow

The given two linear vector datasets were the reference dataset and the target dataset. The matching process is depicted in Figure 4, and the specific matching steps are explained below.

1. Data pre-processing: Remove topological errors from both road datasets and convert them to the same format. If the two datasets have different coordinate systems, they need to be converted to the same coordinate system.
2. Break up roads in the dataset using junctions to facilitate subsequent road matching.
3. Landmarks and nodes are extracted at both ends of the line.
4. Calculate the spatial relationship between landmarks and linear objects.
5. Extract feature points of roads from data at different scales to form the minimum convex hull of the road.
6. Perform directional processing on vector lines and calculate their Pearson coefficients.
7. Each measure is calculated and a comprehensive similarity model is constructed following the methodology described in Section 2.
8. Positive example samples are extracted to derive weight values for each measure of each metric model.
9. After obtaining the optimal weights of each indicator for each model through the positive example samples, matching experiments are conducted on the reference dataset and the dataset to be matched.

To conduct an intuitive comparison and analysis, this study employed precision and recall to quantitatively evaluate the matching results. Precision is the number of correct matches (true positives) divided by the total number of matches found by the algorithm (true positives and false positives). Recall is the number of correct matches divided by the total number of actual true matches (true positives and false negatives). Precision represents the correctness and recall represents the completeness of matching [40]. Nevertheless, the values of precision and recall may have an inverse relationship. The precision value may be high and the recall value may be low, or vice versa. Therefore, for the final assessment, the F-score value, which includes both parameters, was employed [38]. The traditional F-score is the harmonic mean of precision and recall. Formula (18) shows the F-score [38,41,42].

$$\text{F-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \times 100\% \quad (18)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (19)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (20)$$

In the above formulas, TP is the number of matching pairs that are actually detected, FP is the number of matching logarithms that are wrongly detected, and FN is the number of matching pairs that are not detected.

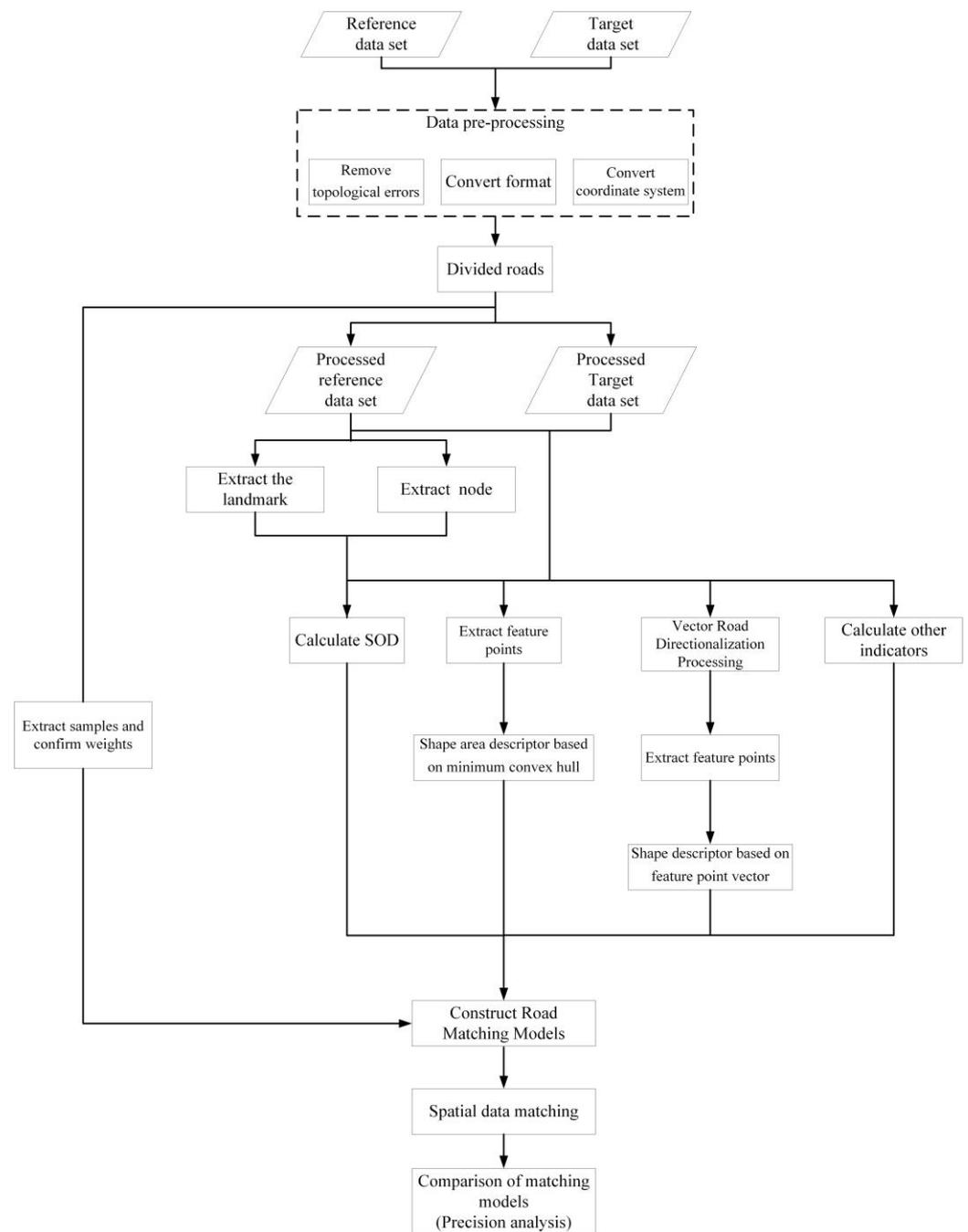


Figure 4. Matching strategy.

4. Experiment and Analysis

4.1. Data

This experiment used road network data (Figure 5) from three different scale datasets of 1:10,000, 1:50,000, and 1:250,000 in Yuyao City, Zhejiang Province in China, from different sources and at different times. Among them, there were 430 road sections in the 1:250,000 road network data (Figure 5a); 4570 road sections in the 1:50,000 road network data (Figure 5b); and 11,550 road sections in the 1:10,000 road network data (Figure 5c). The raw data for the three scales of road data used in the experiments in this paper are different scale vector road datasets obtained from the same publicly available frame library data. Therefore, the raw datasets for the experiments in this paper are the vector road datasets.

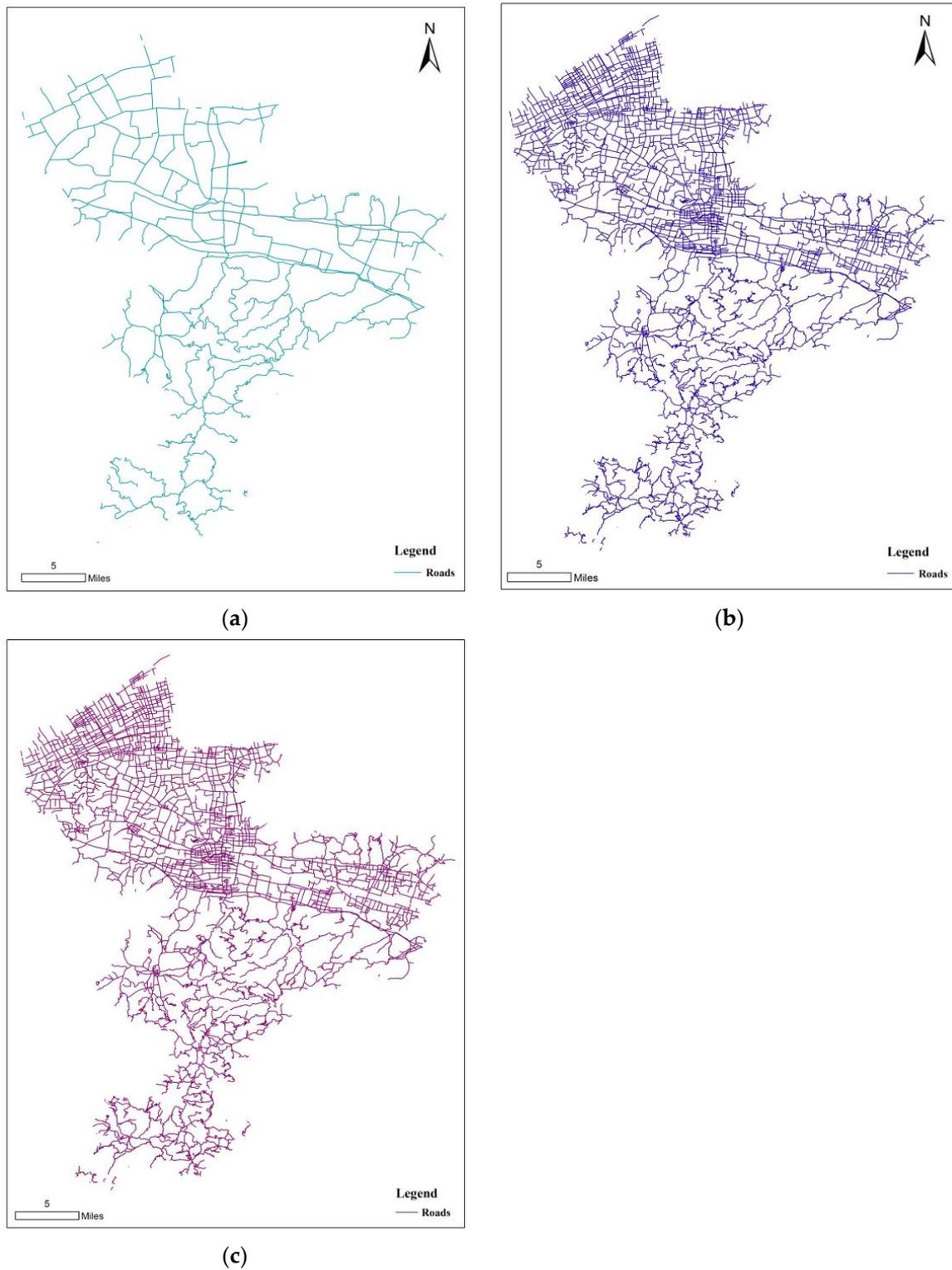


Figure 5. Multi-scale road network data: (a) 1:250,000 road network data; (b) 1:50,000 road network data; (c) 1:10,000 road network data.

4.2. Results

4.2.1. Road Division and Landmark Extraction

Before conducting the road matching experiments, we backed up the original data and then segmented the roads in the dataset using road intersections to facilitate subsequent

road matching. Figure 6 shows the schematic diagram of some of the road segmentations. After the matching experiment, we mapped the matching results to the original data.

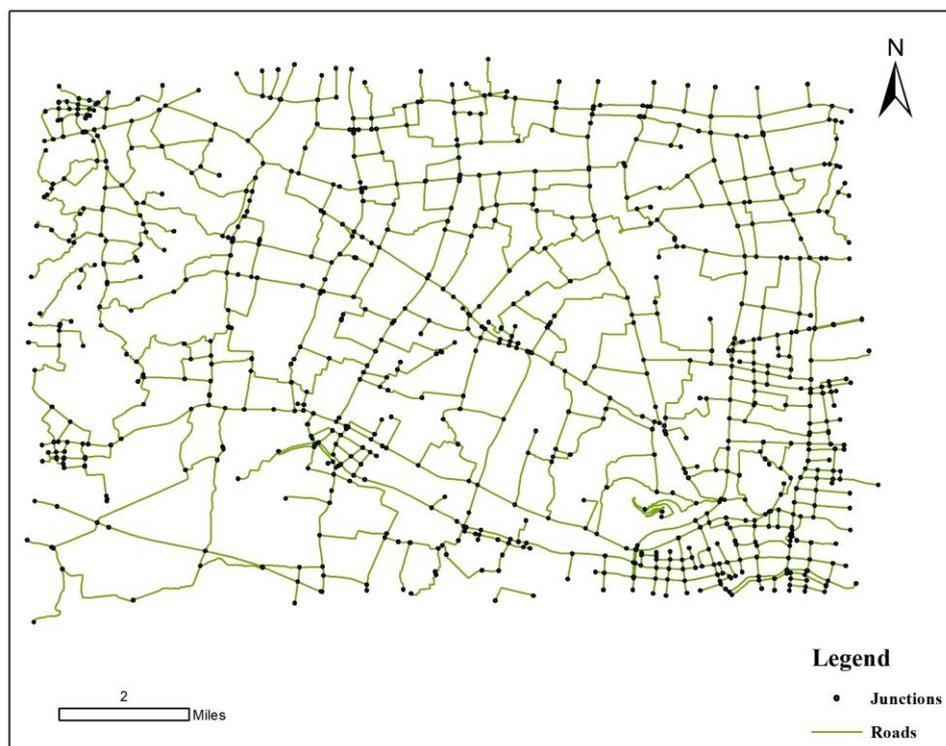


Figure 6. Road dividing schematic.

The junctions in Figure 6 include road intersections, pseudo nodes, and dangling nodes. The landmarks we have selected are road intersections. The number of roads in the road dataset changed after road segmentation was performed. Table 1 shows the change in the number of roads in the dataset before and after road segmentation.

Table 1. Changes in the number of roads before and after road segmentation for different scale datasets.

Number of Roads Map Scale	1:10,000	1:50,000	1:250,000
Before road division	11,550	4570	430
After road division	12,813	5238	482

We performed landmark point selection after completing the road segmentation. The method of selecting landmark points was described in Section 2.1.1. Table 2 shows the number of landmark points for the two scale datasets. Figure 7 shows the results of landmark point selection for the dataset with a scale of 1:250,000. Figure 8 shows the change in the number of road intersections before and after the selection of landmarks for the dataset with a scale of 1:50,000.

Table 2. Number of landmarks in the reference road datasets.

Scale of the Dataset	Number of Landmarks
1:250,000	189
1:50,000	968

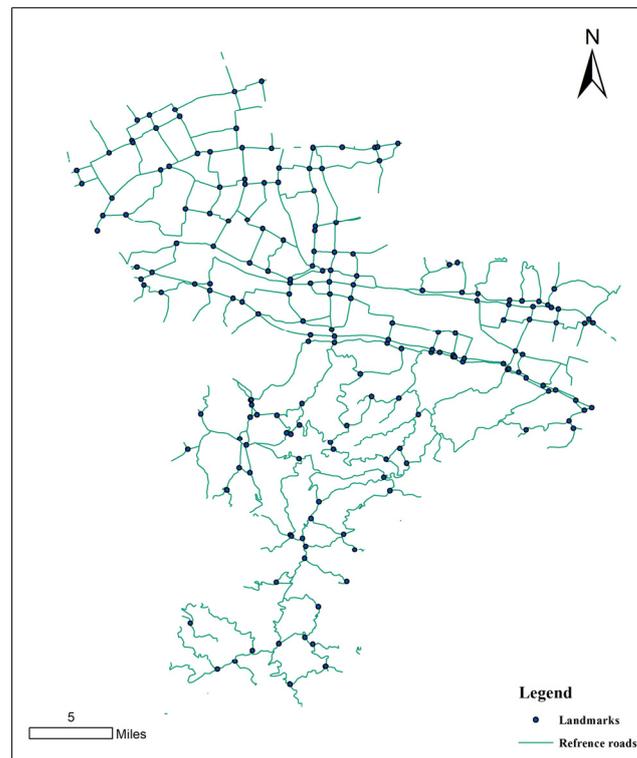


Figure 7. Distribution of landmark points for the dataset with a scale of 1:250,000.

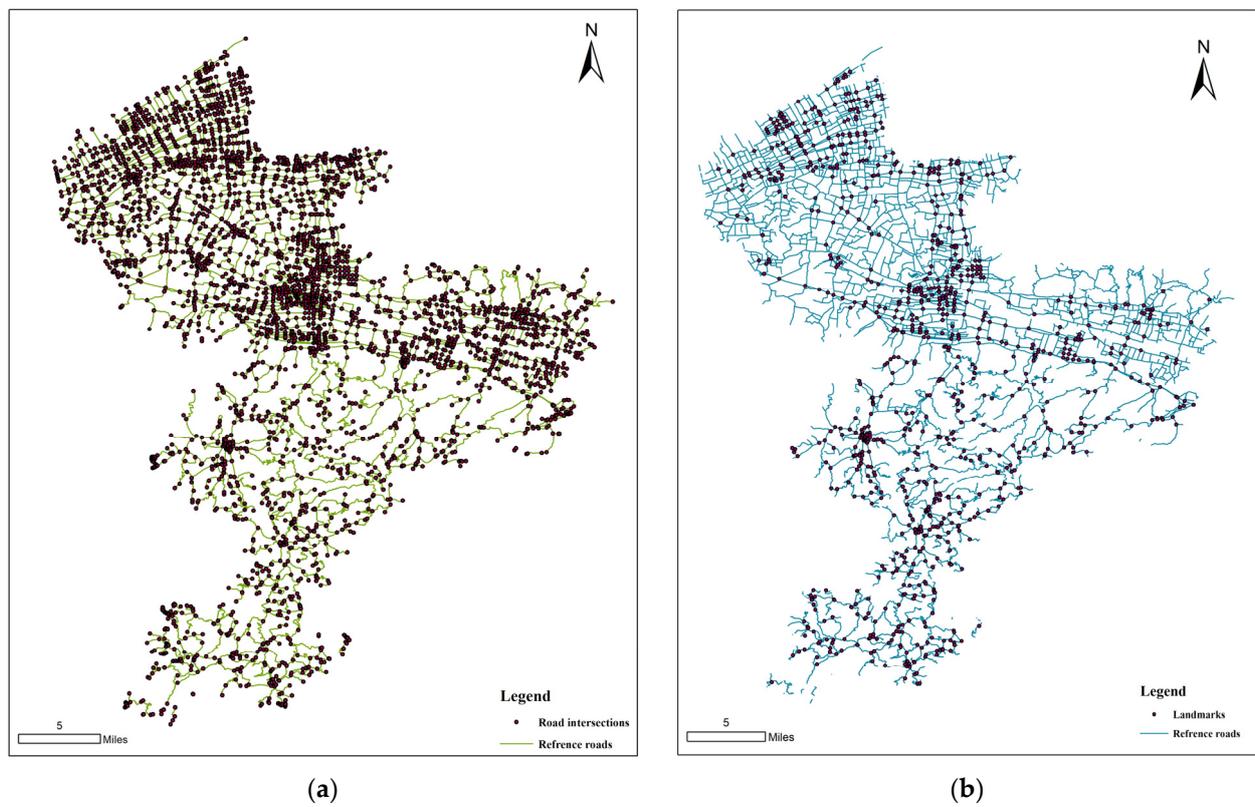


Figure 8. Changes in the distribution of road junctions before and after screening of the dataset at a scale of 1:50,000: (a) distribution of road junctions before landmark selection; (b) landmark distribution.

4.2.2. Results Analysis

We used seven models to match road data at different scales. Due to the large span of the 1:250,000 to 1:10,000 combination scale, the accuracy of the experimental results was low, and the validation effectiveness of the proposed descriptors was low. Therefore, this paper used two scale combinations, namely 1:50,000 to 1:10,000 and 1:250,000 to 1:50,000, respectively. Road datasets with scales of 1:250,000 and 1:50,000 were denoted as Group A. Road datasets with scales of 1:50,000 and 1:10,000 were denoted as Group B. There were significant differences in the amount of data and the level of detail expressed at different scales of data. To improve matching efficiency and accuracy, we used the 1:250,000 and 1:50,000 road data as reference datasets for Groups A and B, respectively. And the 1:50,000 and 1:10,000 road data were used as matching datasets for Groups A and B, respectively. In this study, a total of seven matching models were developed for experiments on two groups of datasets, and the results were summarized and analyzed.

We conducted experiments to determine the weights of each indicator in each of the seven models by extracting positive case samples from each of the two road datasets in Groups A and B. We determined the weights of the indicators in each model by controlling for a single variable. When determining the weights of a model, we first changed the weight of one of the indicators, set the weights of the other indicators and kept them unchanged, then conducted several comparison experiments. The weight corresponding to the best matching result was the weight of the indicator in this model. Subsequently, the weight of the next indicator was changed, the weight of other indicators was kept unchanged, and several comparison experiments are carried out to select the weight corresponding to the best matching results. By analogy, the weights of each indicator can be determined when the model was matched in Groups A and B, respectively. The weights of each indicator in the seven models were normalized. Tables 3 and 4 show the distribution of the indicators for each of the seven models in Groups A and B experiments, respectively.

Table 3. Distribution of weights for each indicator of the seven models in the Group A experiment.

Metrics Model	Similarity Feature					
	Length	Orientation	Short-Line Median Hausdorff Distance	SOD	Shape Area Descriptor Based on Minimum Convex Hull	Shape Descriptor Based on Feature Point Vector
Model 1	0.2	0.2	0.85	0.8		
Model 2	0.2	0.2	0.95		0.7	
Model 3	0.2	0.31	0.8			0.3
Model 4	0.1	0.2	0.9	0.8	0.7	
Model 5	0.1	0.1	0.9	0.9		0.1
Model 6	0.2	0.2	0.9		0.6	0.1
Model 7	0.2	0.3	0.9	0.7	0.6	0.1

Table 4. Distribution of weights for each indicator of the seven models in the Group B experiment.

Metrics Model	Similarity Feature					
	Length	Orientation	Short-Line Median Hausdorff Distance	SOD	Shape Area Descriptor Based on Minimum Convex Hull	Shape Descriptor Based on Feature Point Vector
Model 1	0.3	0.25	0.9	0.9		
Model 2	0.3	0.3	0.75		0.65	
Model 3	0.35	0.4	0.7			0.2

Table 4. Cont.

Metrics Model	Similarity Feature					
	Length	Orientation	Short-Line Median Hausdorff Distance	SOD	Shape Area Descriptor Based on Minimum Convex Hull	Shape Descriptor Based on Feature Point Vector
Model 4	0.2	0.3	0.85	0.8	0.75	
Model 5	0.3	0.3	0.8	0.75		0.2
Model 6	0.25	0.25	0.8		0.5	0.2
Model 7	0.3	0.35	0.8	0.9	0.7	0.15

Figures 9 and 10 show the matching results of Group A and Group B under these seven metric matching models, respectively.

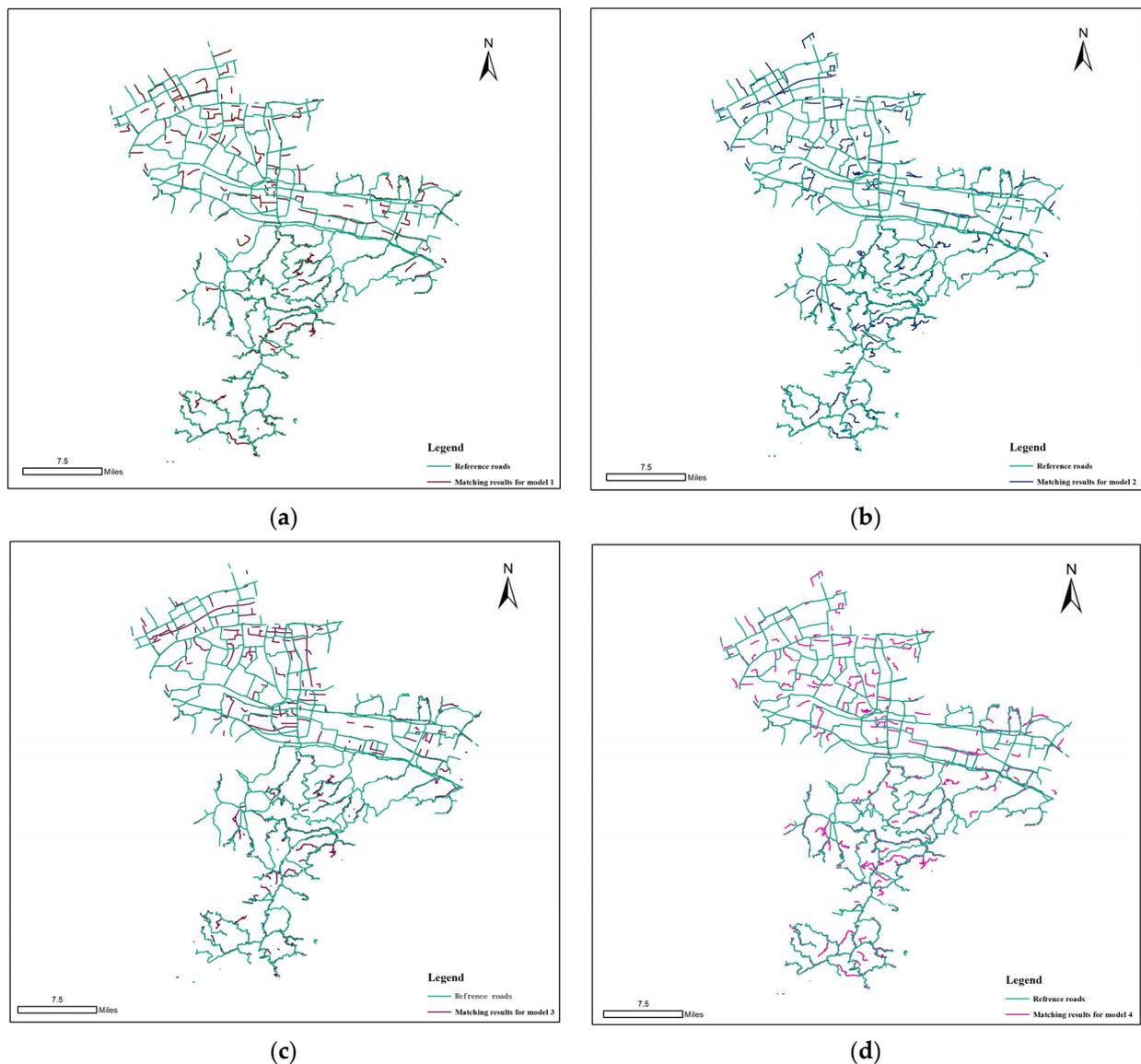


Figure 9. Cont.

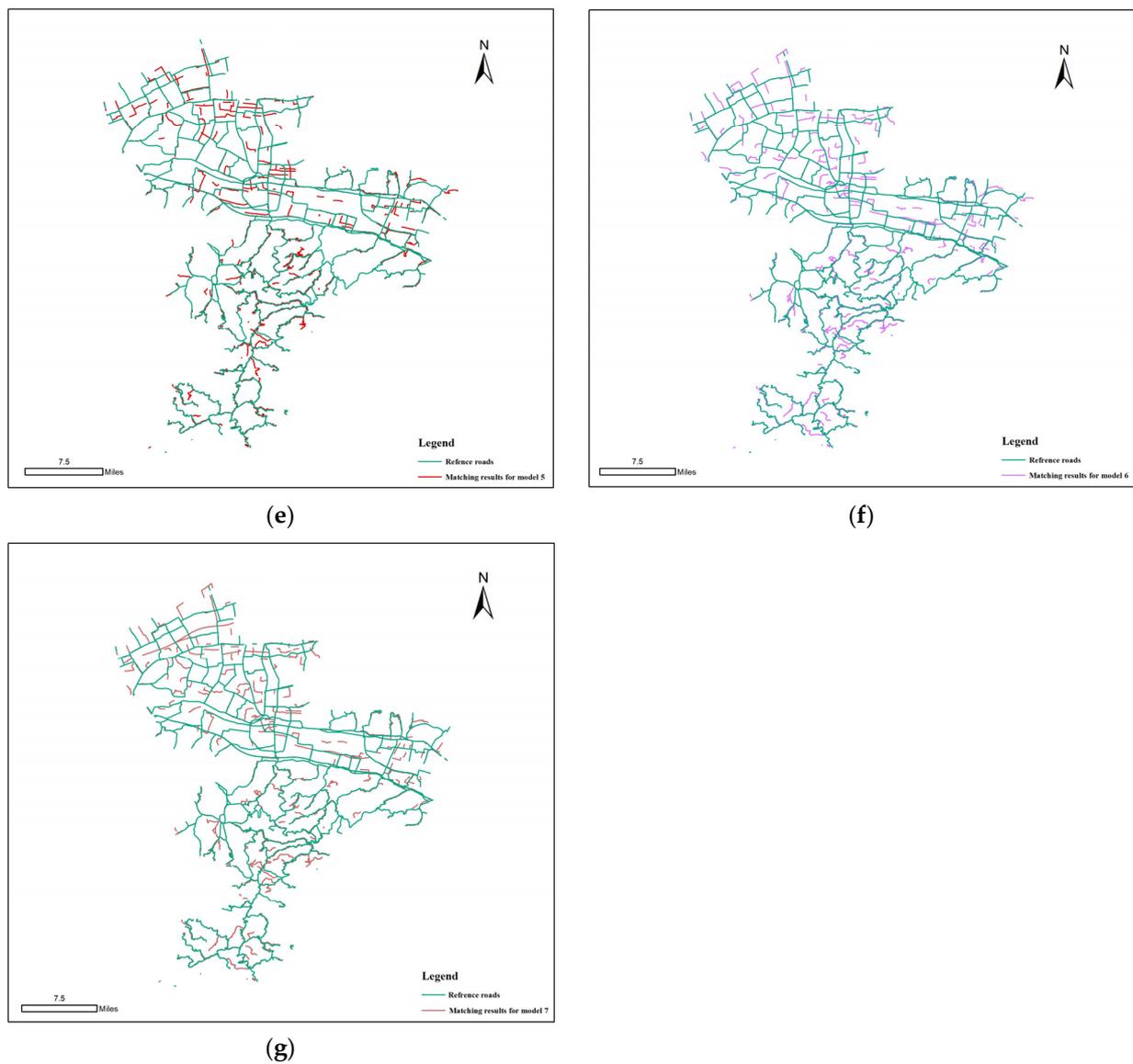


Figure 9. Matching results for Group A: (a–g) show the matching results for metric models 1–7, respectively.

Tables 5 and 6 specifically show the specific results of the two datasets under the seven metric matching models, including the number of missed matches and the number of wrong matches.

Table 5. Performance metrics results for road matching in Group A.

Metrics Model	FN	FP	TP	Precision (%)	Recall (%)	F-Score (%)
Model 1	134	256	1145	81.73	89.52	85.45
Model 2	146	271	1098	80.20	88.26	84.04
Model 3	163	361	1149	76.09	87.58	81.43
Model 4	143	273	1092	80.00	88.42	84.00
Model 5	136	321	1176	78.56	89.63	83.73
Model 6	159	325	1021	75.85	86.53	80.84
Model 7	150	337	1070	76.05	87.70	81.46

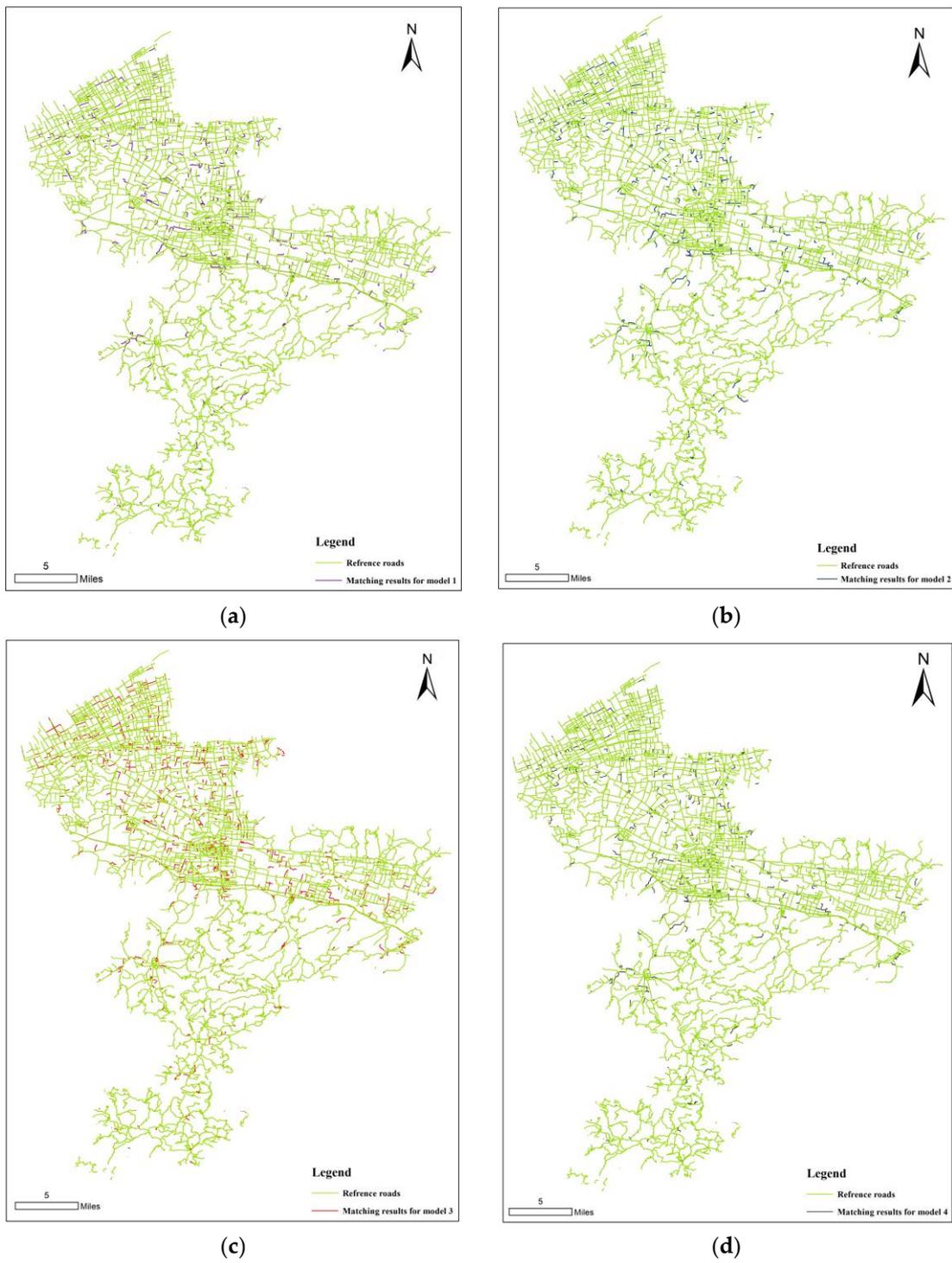


Figure 10. Cont.

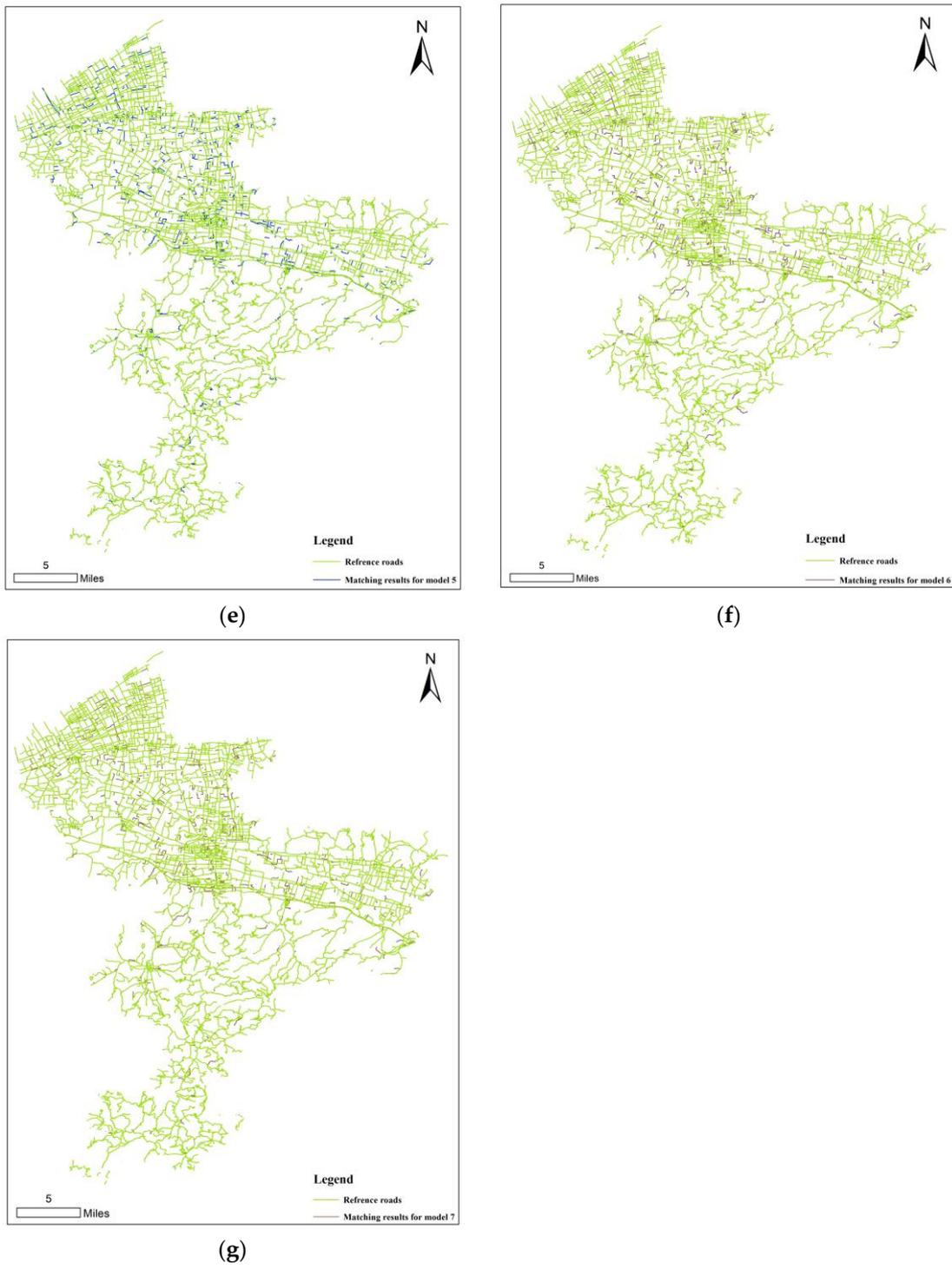
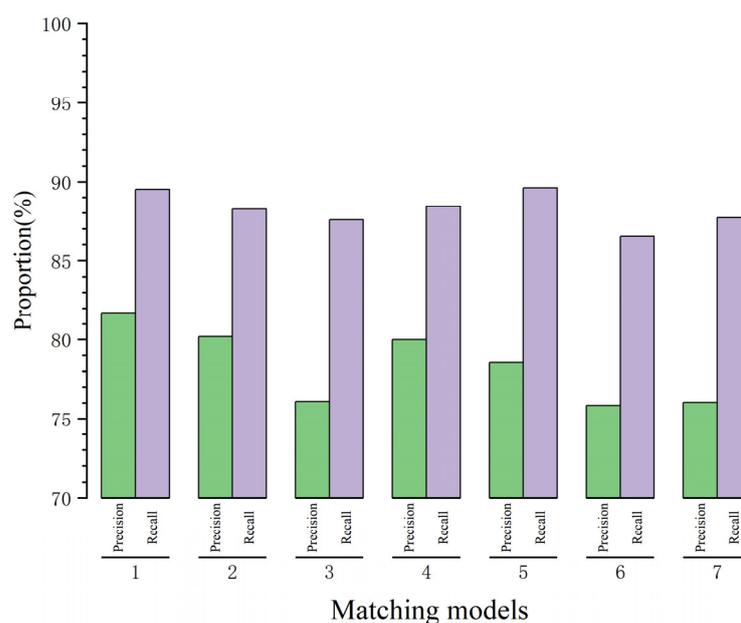


Figure 10. Matching results for Group B: (a–g) show the matching results for metric models 1–7, respectively.

Table 6. Performance metrics results for road matching in Group B.

Metrics Model	FN	FP	TP	Precision (%)	Recall (%)	F-Score (%)
Model 1	749	345	12467	97.31	94.33	95.80
Model 2	810	338	8572	96.21	91.37	93.73
Model 3	1178	856	6742	88.73	85.13	86.89
Model 4	805	355	7448	95.45	90.25	92.78
Model 5	1001	653	7102	91.58	87.65	89.57
Model 6	958	614	7088	92.03	88.09	90.02
Model 7	932	627	7133	91.90	88.44	90.14

First, we analyzed the seven matches of Group A by precision and recall. As can be seen from Table 3, all seven metric matching models used angle, distance, and length. The differences were in the different assignments of SOD, the shape area descriptor based on the minimum convex hull, and the shape descriptor based on the feature point vector. The recall was greater than the precision in the matching results of each model for Group A. The recall was greater than the precision in the matching results of each model. Among models 1–3, model 1 had the highest recall and precision, and model 3 had the lowest. This shows that when adding SOD, the shape area descriptor based on the minimum convex hull, or the shape descriptor based on the feature point vector alone, the matching effect of adding SOD alone works better. When two of the SOD, shape area descriptor based on the minimum convex hull, and shape descriptor based on the feature point vector are added randomly, model 4 achieved higher values for both ratings. With the increase in the number of indicators, the constraints of the model became stronger, so both evaluation values of model 4 were lower than model 1. Model 7 had the strongest constraints due to the addition of all matching indicators, and the two evaluation values were close to the values of model 3. The results of the above analyses showed that the shape descriptor based on the feature point vector, when added alone or in combination with other indicators in the metrics matching model, achieves unsatisfactory matching results, and the two evaluation values of SOD added alone in the matching model were better. Figure 11 show the precision and recall of matching results for Group A.

**Figure 11.** The precision and recall of matching results for Group A.

Secondly, we analyzed the seven matching results of Group B by precision and recall. Distinguishing from the two evaluation values for each matching result of Group A, the precision was higher than the recall for all seven matching results of Group B. The two evaluation values for each matching result of Group B were higher than the recall. In models 1–3, the two evaluation values of model 1 were higher, i.e., the matching effect of adding SOD alone was better. As the number of matching indicators increased, the model became more constrained. In models 4–6, model 4 achieved higher values for both ratings, i.e., better matching with the addition of the two matching metrics of the SOD and shape area descriptor based on the minimum convex hull. Model 7 had the strongest constraints due to the most indicators, and the matching effect was not satisfactory. Figure 12 show the precision and recall of matching results for Group B.

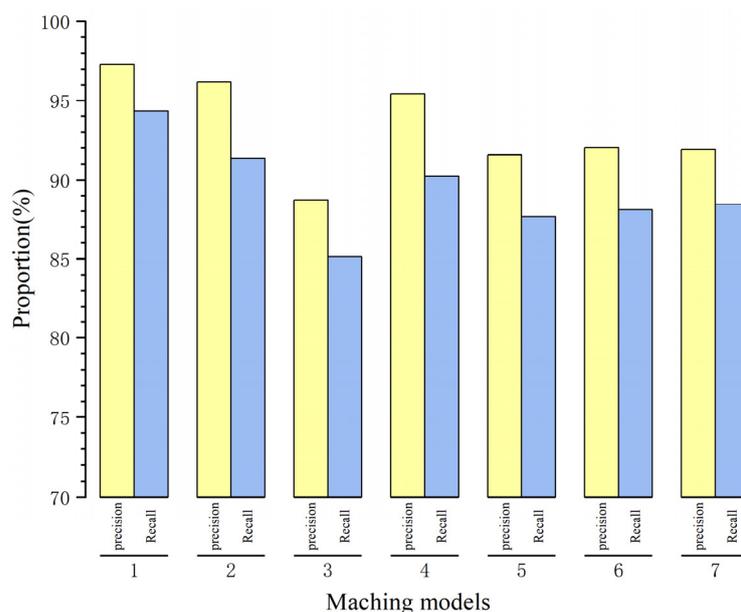


Figure 12. The precision and recall of matching results for Group B.

Finally, in order to prevent large difference between the precision and the recall from affecting the evaluation of the experimental results, we quoted the F-score value for the final evaluation of the experimental results. As shown in the above figure, the F-score values of the seven experiments of Group B were higher than those of Group A. We can preliminarily conclude that the F-score values of the model matching results were higher as the scale increased and the number of roads increased. The F-score value of model 1 was the highest among the seven experimental results of Group A and the seven experimental results of Group B. Model 2 was the second highest, and the F-score value of model 4 was also more satisfactory. Figure 13 show the F-score value of matching results for Group A.

Combining the above data analyses, we can preliminarily draw the following conclusions:

1. In the case that the scale of the two datasets does not differ much, the results of model matching will be better as the scale of the dataset increases.
2. Combining the matching results of the two groups, model 1, model 2, and model 4 re better matched on road datasets with scales of 1:50,000 and 1:10,000, respectively. The precision, recall, and F-score for model 1 were 97.31%, 94.33%, and 95.80%, respectively. The precision, recall, and F-score for model 2 were 96.21%, 91.37%, and 93.73%, respectively. The precision, recall, and F-score for model 4 were 95.45%, 90.25%, and 92.78%, respectively.
3. As far as the experiments in this paper are concerned, as the scale of the dataset increases, the miss-match of model matching increased and the mismatch situations decreased.

4. As the matching indicator increases, the binding of the matching model increases. Therefore, different matching models can be selected for different situations of road matching. For road matching that requires high accuracy, model 1 can be chosen because it has the best road matching results. For road matching situations that require a combination of spatial relationships and shape descriptors, model 4 can be selected, because the matching effect of the model combining SOD and minimum convex hull metrics is the best.

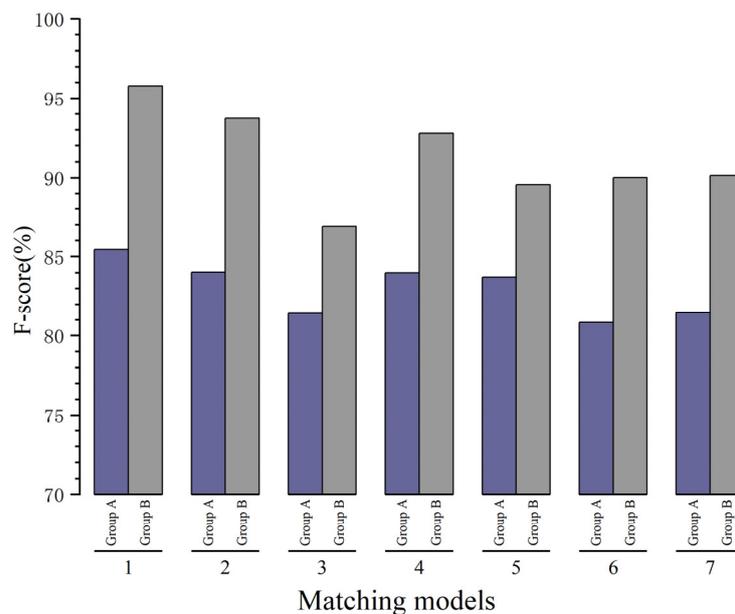


Figure 13. The F-score value of matching results for Group A and Group B.

5. Conclusions

Geographic information spatial data are the basis of any GIS system. Conducting data fusion and rapid updates is one of the major problems in the GIS field. To achieve effective, dynamic, and adaptive integration of different spatial data, research on geospatial data matching methods of different scales is required. For matching linear vector spatial data with different scales, we present here a comprehensive similarity model based on six indices. Through experiments, this study obtained a metric matching model with a better road matching effect, i.e., adding the matching metric of SOD on the basis of the three matching metrics of orientation, length, and distance. This model is more suitable for matching between large-scale road datasets. In this study, the F-score of the matching results of road datasets with scales of 1:250,000 and 1:50,000 based on this model was 85.45%. When the scale was increased, the F-score value of matching results for road datasets with scales of 1:50,000 and 1:10,000 was 95.80%. In comparison, the latter match was improved by 10.35%. Thus, the synthetic similarity model proposed in this paper can be effective in datasets with different scales and phases, as well as achieving high-matching quality and accuracy.

Nevertheless, only two sets of road datasets with different scales were experimented with in this paper, and the conclusion that “the larger the scale, the better the matching effect of the proposed model” needs to be further verified. In future research, we will conduct experiments with datasets of different scales from different cities to further validate our model.

Author Contributions: Conceptualization, Yuefeng Lu; data curation, Ying Sun, Kaizhong Yao, Yanru Liu and Jing Li; methodology, Yuefeng Lu, Ying Sun and Qiao Wen; project administration, Yuefeng Lu; Supervision, Yuefeng Lu; writing—original draft, Ying Sun, Ziqi Ding and Qiao Wen; writing—review and editing, Yuefeng Lu, Ying Sun and Ziqi Ding. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Geological Survey Project of China Geological Survey (No. DD20220956); the Major Project of High Resolution Earth Observation System of China (No. GFZX0404130304); the Open Fund of Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying, Mapping and Remote Sensing, Hunan University of Science and Technology (No. E22201); a grant from the State Key Laboratory of Resources and Environmental Information System; and the Innovation Capability Improvement Project of Scientific and Technological Small and Medium-sized Enterprises in Shandong Province of China (No. 2021TSGC1056).

Data Availability Statement: Restrictions apply to the availability of these data. The raw data supporting the conclusion of this paper will be made available by the authors, without undue reservation. All experimental data were derived from project data. Requests to access these datasets should be directed to corresponding author.

Acknowledgments: The authors are grateful to the editors and the referees for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, Z. Research on Matching and Connection Organization Method of Multi-Scale Road Data. Master's Thesis, Nanjing Normal University, Nanjing, China, 2020.
2. Zhang, X.; Ai, T.; Stoter, J.; Zhao, X. Data matching of building polygons at multiple map scales improved by contextual information and relaxation. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 147–163. [[CrossRef](#)]
3. Abbaspour, R.A.; Chehreghan, A.; Chamani, M. Multi-scale polygons matching using a new geographic context descriptor. *Appl. Geomat.* **2021**, *13*, 885–899. [[CrossRef](#)]
4. Zhao, B.; Deng, M.; Liu, H.; Xu, Z. An Approach to Matching Area Objects and Line Objects of River System in Multi-scale Maps. *Geo-Inf. Sci.* **2011**, *13*, 361–366. [[CrossRef](#)]
5. Deng, M.; Li, Z.; Chen, X. Extended Hausdorff Distance for Spatial Objects in GIS. *Int. J. Geogr. Inf. Sci.* **2007**, *21*, 459–475.
6. Eiter, T.; Mannila, H. Computing discrete Frechet distance. In *Technical Report of Christian Doppler Labor für Expertensysteme*; Technical Report CDTR 94/64; Technical University of Vienna: Vienna, Austria, 25 April 1994.
7. Chehreghan, A.; Ali Abbaspour, R. A geometric-based approach for road matching on multi-scale datasets using a genetic algorithm. *Cartogr. Geogr. Inf. Sci.* **2018**, *45*, 255–269. [[CrossRef](#)]
8. Govindaraj, P.; Sudhakar, M.S. Shape characterization using laws of texture energy measures facilitating retrieval. *Imaging Sci. J.* **2018**, *66*, 98–105. [[CrossRef](#)]
9. Huh, Y.; Yu, K.; Heo, J. Detecting conjugate-point pairs for map alignment between two polygon datasets. *Comput. Environ. Urban Syst.* **2010**, *35*, 250–262. [[CrossRef](#)]
10. Filin, S.; Doytsher, Y. The Detection of Corresponding Objects In A Linear-Based Map Conflation. *Surv. Land Inf. Syst.* **2000**, *60*, 101–108.
11. Arkin, E.; Chew, P.; Huttenlocher, D.; Kedem, K.; Mitchel, J. An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1991**, *13*, 209–216. [[CrossRef](#)]
12. Luo, G.; Qin, K. Global optimisation matching method for multi-representation buildings constrained by road network. *J. Comput. Methods Sci. Eng.* **2023**, *23*, 2413–2424. [[CrossRef](#)]
13. Wang, H.; Liu, Y.; Li, S.; Bo, L.; He, Z. A Path Increment Map Matching Method for High-Frequency Trajectory. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 10948–10962.
14. Li, B.; Wang, M.; Cai, Z.; Su, S.; Kang, M. VIS-MM: A novel map-matching algorithm with semantic fusion from vehicle-borne images. *Int. J. Geogr. Inf. Sci.* **2023**, *37*, 1069–1098. [[CrossRef](#)]
15. Wu, J.; Wan, Y.Y.; Chiang, Y.; Fu, Z.; Deng, M. A Matching Algorithm Based on Voronoi Diagram for Multi-Scale Polygonal Residential Areas. *IEEE Access* **2018**, *6*, 4904–4915. [[CrossRef](#)]
16. Wang, S.; Guo, Q.; Xu, X.; Xie, Y. A Study on a Matching Algorithm for Urban Underground Pipelines. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 352. [[CrossRef](#)]
17. Zhang, J.; Wang, Y.; Zhao, W. An improved probabilistic relaxation method for matching multi-scale road networks. *Int. J. Digit. Earth* **2018**, *11*, 635–655. [[CrossRef](#)]
18. Wan, B.; Yang, L.; Zhou, S.; Wang, R.; Wang, D.; Zhen, W. A Parallel-Computing Approach for Vector Road-Network Matching Using GPU Architecture. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 472. [[CrossRef](#)]

19. Chehrehghan, A.; Ali Abbaspour, R. An assessment of the efficiency of spatial distances in linear object matching on multi-scale, multi-source maps. *Int. J. Image Data Fusion* **2018**, *9*, 114–195. [[CrossRef](#)]
20. Guo, Q.; Xu, X.; Wang, Y.; Liu, J. Combined Matching Approach of Road Networks Under Different Scales Considering Constraints of Cartographic Generalization. *IEEE Access* **2020**, *8*, 944–956. [[CrossRef](#)]
21. Lewandowicz, E.; Flisek, P. A Method for Generating the Centerline of an Elongated Polygon on the Example of a Watercourse. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 304. [[CrossRef](#)]
22. Wu, J.; Zhao, Y.; Yu, M.; Zou, X.; Xiong, J.; Hu, X. A new Voronoi diagram-based approach for matching multi-scale road networks. *J. Geogr. Syst.* **2023**, *25*, 265–289. [[CrossRef](#)]
23. Singh, S.; Singh, J.; Goyal, S.B.; El Barachi, M.; Kumar, M. Analytical Review of Map Matching Algorithms: Analyzing the Performance and Efficiency Using Road Dataset of the Indian Subcontinent. *Arch. Comput. Methods Eng.* **2023**, *30*, 4897–4916. [[CrossRef](#)]
24. Zhao, L.; Liu, Y.; Lu, Y.; Sun, Y.; Li, J.; Yao, K. Research on Vector Road Data Matching Method Based on Deep Learning. *J. Appl. Math. Phys.* **2013**, *11*, 303–315. [[CrossRef](#)]
25. Yan, X.; Song, C.; Pei, T.; Wang, X.; Wu, M.; Liu, T.; Shu, H.; Chen, J. Revealing spatiotemporal matching patterns between traffic flux and road resources using big geodata—A case study of Beijing. *Cities* **2022**, *127*, 103754. [[CrossRef](#)]
26. Cheng, R.; Chen, J. A location conversion method for roads through deep learning-based semantic matching and simplified qualitative direction knowledge representation. *Eng. Appl. Artif. Intell.* **2021**, *104*, 104400. [[CrossRef](#)]
27. Wang, Y.; Yu, B.; Zhu, F.; Zhang, J.; Huang, C. Hierarchical stroke mesh: A new progressive matching method for detecting multi-scale road network changes using OpenStreetMap. *Soft Comput.* **2021**, *25*, 3155–3173. [[CrossRef](#)]
28. Li, W.; Wang, Y.; Li, D.; Xu, X. MCM: A Robust Map Matching Method by Tracking Multiple Road Candidates. In Proceedings of the Algorithmic Aspects in Information and Management, Guangzhou, China, 13–14 August 2022.
29. Li, X.; Liu, W. Research on Vector Road Aided Inertial Navigation by Using ICCP Algorithm. In Proceedings of the Spatial Data and Intelligence, Shenzhen, China, 8–9 May 2020.
30. Shen, L.; Tao, H.; Ni, Y.; Wang, Y.; Stojanovic, V. Improved YOLOv3 model with feature map cropping for multi-scale road object detection. *Meas. Sci. Technol.* **2023**, *34*, 045406. [[CrossRef](#)]
31. Xie, Q.; Hu, X.; Ren, L.; Qi, L.; Sun, Z. A Binocular Vision Application in IoT: Realtime Trustworthy Road Condition Detection System in Passable Area. *IEEE Trans. Ind. Inform.* **2023**, *19*, 973–983. [[CrossRef](#)]
32. Lei, T.; Xiao, G.; Yin, X. Targeting Lane-Level Map Matching for Smart Vehicles: Construction of High-Definition Road Maps Based on GIS. *Appl. Sci.* **2023**, *13*, 862. [[CrossRef](#)]
33. Ma, S.; Lee, H. A Practical HMM-Based Map-Matching Method for Pedestrian Navigation. In Proceedings of the International Conference on Information Networking (ICOIN), Bangkok, Thailand, 11–14 January 2023; pp. 806–811.
34. Zhao, B. A Study on Multi-Scale Vector Map Objects Matching Method and Its application. Ph.D. Thesis, Central South University, Changsha, China, 2011.
35. Xu, F.; Deng, M.; Zhao, B.; Chen, J. A detailed investigation on the methods of object matching. *J. Geo-Inf. Sci.* **2009**, *11*, 657–663. [[CrossRef](#)]
36. Chehrehghan, A.; Ali Abbaspour, R. A new descriptor for improving geometric-based matching of linear objects on multi-scale datasets. *GIScience Remote Sens.* **2017**, *54*, 836–861. [[CrossRef](#)]
37. Zhang, M. Methods and Implementations of Road-Network Matching. Master’s Thesis, Technische Universität München, Munich, Germany, 2009.
38. Wang, Y.; Chen, D.; Zhao, Z.; Ren, F.; Du, Q. A Back-Propagation Neural Network-Based Approach for Multi-Represented Feature Matching in Update Propagation. *Trans. GIS* **2015**, *19*, 964–993. [[CrossRef](#)]
39. Tong, X.; Liang, D.; Jin, Y. A Linear Road Object Matching Method for Conflation Based on Optimization and Logistic Regression. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 824–846. [[CrossRef](#)]
40. Song, W.; Keller, J.M.; Haithcoat, T.L.; Davis, C.H. Relaxation-Based Point Feature Matching for Vector Map Conflation. *Trans. GIS* **2011**, *15*, 43–60. [[CrossRef](#)]
41. Fan, H.; Yang, B.; Zipf, A.; Rousell, A. A Polygon-Based Approach for Matching Openstreetmap Road Networks with Regional Transit Authority Data. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 748–764. [[CrossRef](#)]
42. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Elsevier: Waltham, MA, USA, 2011; pp. 327–391.

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.