

Article

Measuring Similarity of Deforestation Patterns in Time and Space across Differences in Resolution

Desi Suyamto ^{1,*}, Lilik Prasetyo ², Yudi Setiawan ², Arief Wijaya ³, Kustiyo Kustiyo ⁴, Tatik Kartika ⁴, Hefni Effendi ¹ and Prita Permatasari ¹

¹ Center for Environmental Science of IPB University (PPLH), PPLH Building, Jalan Lingkar Akademik, IPB University, Dramaga, Bogor 16680, Indonesia; hefni.effendi@apps.ipb.ac.id (H.E.); prita07@apps.ipb.ac.id (P.P.)

² Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry, IPB University, Dramaga, Bogor 16680, Indonesia; lbprastdp@apps.ipb.ac.id (L.P.); setiawan.yudi@apps.ipb.ac.id (Y.S.)

³ World Resources Institute Indonesia (WRI Indonesia), Wisma PMI Building 7th Floor, Jalan Wijaya I, No. 63, Kebayoran Baru, Jakarta 12170, Indonesia; arief.wijaya@wri.org

⁴ Center for Remote Sensing Technology and Data, Deputy of Remote Sensing Affairs, Indonesian National Institute of Aeronautics and Space (LAPAN), Jalan LAPAN No. 70, Pekayon, Pasar Rebo, Jakarta 13710, Indonesia; kustiyo@lapan.go.id (K.K.); tatik.kartika@lapan.go.id (T.K.)

* Correspondence: d_suyamto@apps.ipb.ac.id; Tel.: +62-813-688-60965

Abstract: This article demonstrated an easily applicable method for measuring the similarity between a pair of point patterns, which applies to spatial or temporal data sets. Such a measurement was performed using similarity-based pattern analysis as an alternative to conventional approaches, which typically utilize straightforward point-to-point matching. Using our approach, in each point data set, two geometric features (i.e., the distance and angle from the centroid) were calculated and represented as probability density functions (PDFs). The PDF similarity of each geometric feature was measured using nine metrics, with values ranging from zero (very contrasting) to one (exactly the same). The overall similarity was defined as the average of the distance and angle similarities. In terms of sensibility, the method was shown to be capable of measuring, at a human visual sensing level, two pairs of hypothetical patterns, presenting reasonable results. Meanwhile, in terms of the method's sensitivity to both spatial and temporal displacements from the hypothetical origin, the method is also capable of consistently measuring the similarity of spatial and temporal patterns. The application of the method to assess both spatial and temporal pattern similarities between two deforestation data sets with different resolutions was also discussed.

Keywords: deforestation hotspots; generic method; similarity; spatial pattern; temporal pattern



Citation: Suyamto, D.; Prasetyo, L.; Setiawan, Y.; Wijaya, A.; Kustiyo, K.; Kartika, T.; Effendi, H.; Permatasari, P. Measuring Similarity of Deforestation Patterns in Time and Space across Differences in Resolution. *Geomatics* **2021**, *1*, 464–495. <https://doi.org/10.3390/geomatics1040027>

Academic Editors: Joanne N. Halls, Weidong Li and Chuanrong Zhang

Received: 2 October 2021

Accepted: 27 November 2021

Published: 29 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

The similarity between multiple spatial patterns (or their representations) has often been approached by a 'distance' concept, often after normalization in which the range (maximum–minimum) is used as a unitary metric. The dimensions often include both space and time plus other metrics of interest. In a two-dimensional plot, 'lack of fit' can be expressed in terms of distance along either the x - or y -axis (as in standard deviations) or as Euclidean distance (as in correlation coefficients). Beyond two dimensions, direct visualization is more challenging, but Euclidean distances can be readily calculated. A specific form of this problem also occurs in discussions of deforestation maps.

Measurement of the spatial agreement between deforestation maps from various producers is usually carried out based on the point-to-point degree of matching (i.e., with regard to their omission and commission disagreements; see, e.g., [1–5]). Nevertheless, spatial discrepancies due to omission and commission disagreements between deforestation maps from various producers often trigger political debates, regardless of the overall number, the temporal trend, the spatial pattern, or the operational forest definition (threshold)

used, which are mostly counterproductive to efforts for providing salient, legitimate, and credible data on deforestation for relevant policy formulation as guidance on the ground [6]. As suggested by [7], when the problem is situated within a political zone, science should be a mediator.

In fact, point distributions over space from two different data sets can have zero point-to-point degrees of matching, while they actually are similar in the sense that hypotheses of random pattern generation can be rejected. Previously, considering such a possibility, [8] proposed an approach to measure the spatial goodness of fit based on the degree of matching between two spatial patterns at multiple resolutions; that is, using scanning windows of various sizes, ranging from 1 pixel, 2×2 pixels, 3×3 pixels, to maximum sizes of the extent, which are weighted using an exponential decay function of the window size such that degree of matching resulting from larger windows is less weighted. Yet, several significant innovations have been made to methods for measuring the similarity of spatial patterns (i.e., concerning spatial pattern recognition) [9].

Moreover, other than important issues related to the spatial agreement between deforestation maps from various producers, there are also significant issues on the temporal agreement between deforestation time-series data from various producers, as several producers have developed near-real-time deforestation monitoring systems derived from various sources of satellite imageries with different spatial resolutions and using various change-detection algorithms. A discussion of uncertainty in image interpretation at different spatial resolutions relating to metrics relevant for users in a specific location was provided by [10]. Meanwhile, [11,12] provide discussions of uncertainty in various change-detection algorithms and thresholds.

This article summarizes our study, which explored a generic method for measuring spatial and temporal pattern similarities between deforestation data sets with different spatial resolutions based on similarity-based pattern analysis. To achieve our objectives, three steps were carried out consecutively: (i) conceptualizing the method; (ii) evaluating the robustness of the method through sensibility and sensitivity analyses; and (iii) applying the method using two deforestation data sets in Indonesia for 2020, derived from MODIS and Landsat-8 OLI remote-sensing imagery.

The method offers an alternative to conventional approaches to accuracy assessment of deforestation patterns in time and space across differences in resolution, which should apply to various deforestation monitoring data sets, ranging from near-real-time [13,14] to sub-annual and annual [15,16] monitoring systems, and covering local and national scales [4–6] to global scales [2,3,5]. Beyond the deforestation data sets, another potential application of the method is for validating spatio-temporal patterns of simulation results from land-change models, such as [17–20].

2. Materials and Methods

2.1. Methods

This subsection elaborates the conceptual framework of the method and how its robustness was evaluated through sensibility and sensitivity analyses before its application to the measurement of spatial and temporal pattern similarities between two deforestation data sets derived from two satellite imageries with different spatial resolutions (i.e., MODIS and Landsat-8 OLI). For sensibility and sensitivity analyses, the method was assessed using hypothetical data, either spatial or temporal.

2.1.1. Conceptual Framework

To measure the similarity between various patterns in point data sets, some type of approximate point matching should be applied instead of exact-point pattern matching [21]. Moreover, regarding the presence of uncertainty in the data, the location of each point can be represented as a probability density function (PDF) [22–26].

Inspired by [27], in which the similarity between various 3D shapes was sophisticatedly measured based on the similarity of probability density functions (PDFs) and

cumulative distribution functions (CDFs) of some geometric features, we adopted the method for measuring the spatial pattern similarity between various point data sets based on PDF similarity measures of the distance (δ) and angle (θ) from a fixed point to each point of each data set. With regard to the fairness of measuring δ and θ , we defined the fixed point as the centroid of the extent where all point data sets under comparison were located (Figure 1).

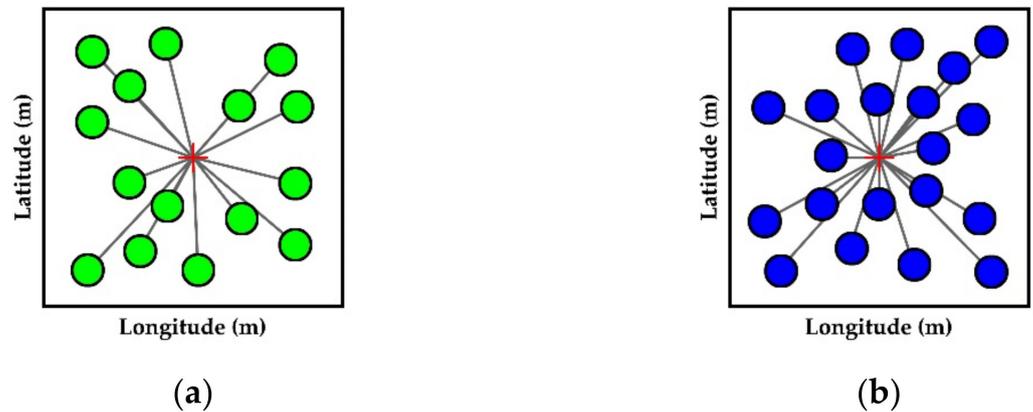


Figure 1. We compared the spatial pattern similarity of two point data sets by measuring PDF similarities for the distance (δ) and angle (θ) from the centroid of the extent (red cross) to each point of each data set (a,b).

As temporal data are, in fact, comparable to spatial data [28], where temporal data are also projected using a particular coordinate system (i.e., Cartesian), with the x -axis referring to time and the y -axis referring to magnitude, in this study, we also applied the same method for measuring temporal pattern similarity between various temporal data sets. Furthermore, we treated the temporal data as spatial data, where the x -axis was considered as the longitude, and the y -axis was considered as the latitude (Figure 2). Similarly, we measured the temporal pattern similarity between various temporal data sets based on the PDF similarities of δ and θ , which, in this case, were measured from the centroid of the graph.

Moreover, to measure the PDF similarity, 9 metrics were selected from the 45 metrics reviewed in [29], where the selected metrics resulted in a similarity of zero for very contrasting pairs of PDFs and a similarity of one for PDFs which were exactly the same (Figure 3). Meanwhile, the other metrics reviewed in [29] resulted in similarity values beyond sensible scales; for example, the metrics of City Blok L1, Czekanowski, and Squared-Chord resulted in similarity values ranging from -1 to 1 . The equations of the nine selected metrics applied in this study are provided in Table 1.

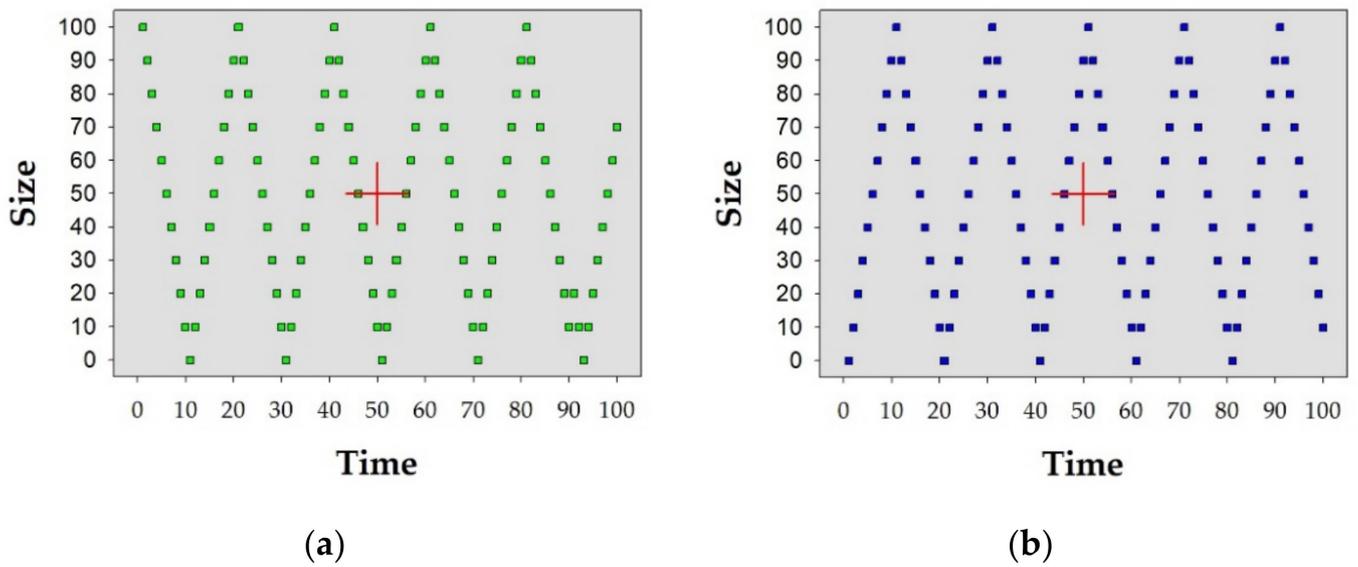


Figure 2. We compared the temporal pattern similarity of two time-series data sets by measuring the similarity of the PDFs of distance (δ) and angle (θ) from the centroid of the graphs where all time-series data sets under comparison were plotted (red cross) to each point within the pair of time series data sets (a,b).

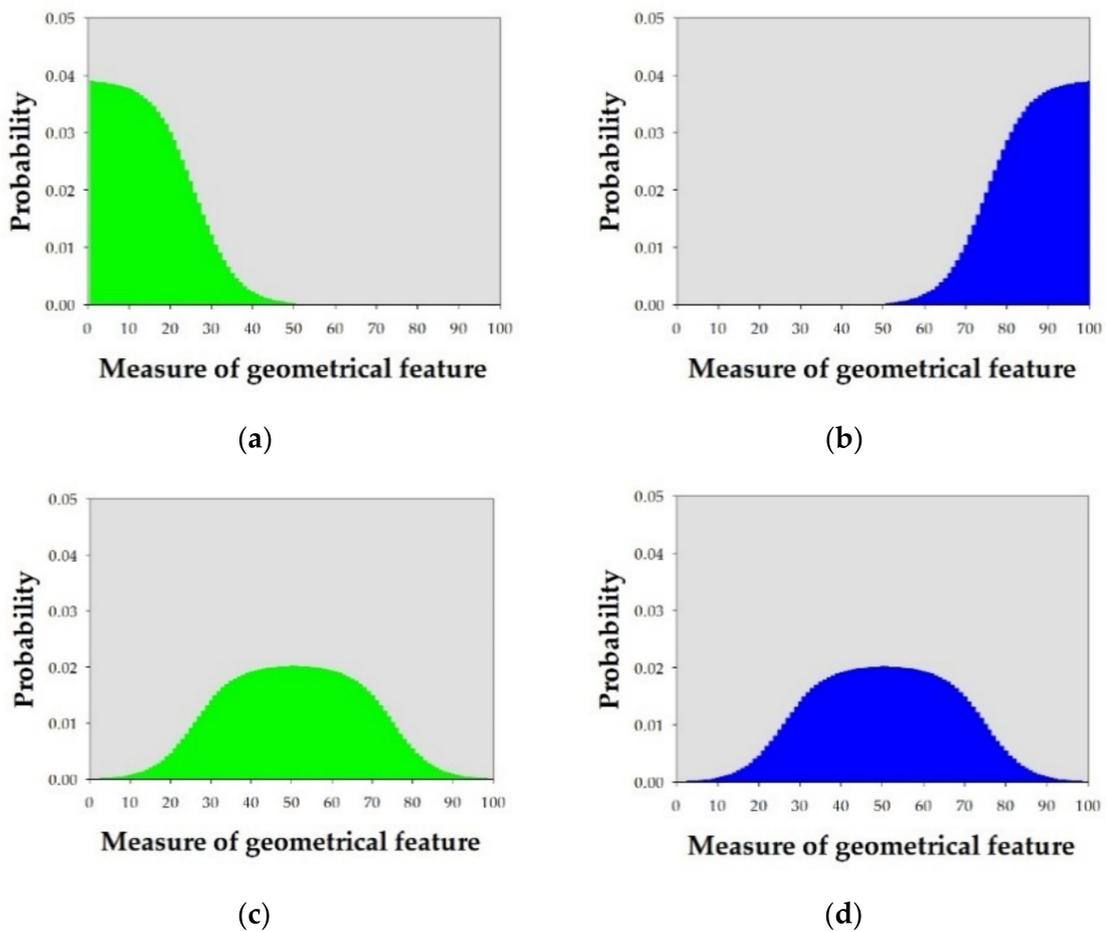


Figure 3. Two pairs of hypothetical PDFs: PDFs (a,b) represent very contrasting pairs of PDFs, while PDFs (c,d) represent exactly matching PDFs. We selected nine metrics out of the 45 metrics reviewed in [29], where the nine selected metrics resulted in a similarity of 0 for (a,b) and a similarity of 1 for (c,d).

Later, PDF similarity was measured for both of the geometric features used in this study—that is, the distance (δ) and angle (θ) from the centroid to each point within the pair of data sets under comparison—which applies to either spatial or temporal data sets. The overall pattern similarity between various spatial or temporal data sets was calculated from the average between the PDF similarities of δ and θ . Figure 4 summarizes the overall workflow of our method.

Table 1. Equations of nine similarity measures (S) between PDF pairs, p and q , with n ordinal classes. We selected these nine PDF similarity measures out of the 45 metrics reviewed by [29], as their values ranged from 0 (for a pair of very contrasting PDFs) to 1 (for a pair of exactly matching PDFs); the other metrics reviewed by [29] resulted in similarity values ranging beyond sensible scales (e.g., the metrics of City Blok L1, Czekanowski, and Squared-Chord metrics had values ranging from -1 to 1).

Metric	Equation
Sørensen	$S = 1 - \frac{\sum_{i=1}^n p_i - q_i }{\sum_{i=1}^n (p_i + q_i)}$
Soergel	$S = 1 - \frac{\sum_{i=1}^n p_i - q_i }{\sum_{i=1}^n \max(p_i, q_i)}$
Intersection	$S = \sum_{i=1}^n \min(p_i, q_i)$
Ruzicka	$S = \frac{\sum_{i=1}^n \min(p_i, q_i)}{\sum_{i=1}^n \max(p_i, q_i)}$
Tanimoto	$S = 1 - \frac{\sum_{i=1}^n (\max(p_i, q_i) - \min(p_i, q_i))}{\sum_{i=1}^n \max(p_i, q_i)}$
Cosine	$S = \frac{\sum_{i=1}^n p_i q_i}{\sqrt{\sum_{i=1}^n p_i^2} \sqrt{\sum_{i=1}^n q_i^2}}$
Jaccard	$S = \frac{\sum_{i=1}^n p_i q_i}{\sum_{i=1}^n p_i^2 + \sum_{i=1}^n q_i^2 - \sum_{i=1}^n p_i q_i}$
Dice	$S = \frac{2 \sum_{i=1}^n p_i q_i}{\sum_{i=1}^n p_i^2 + \sum_{i=1}^n q_i^2}$
Fidelity	$S = \sum_{i=1}^n \sqrt{p_i q_i}$

2.1.2. Sensibility Analyses

Two pairs of hypothetical spatial patterns were used to evaluate the method in terms of its sensibility (Figure 5). The first pair was two images of mirrored eagles with the same size and pattern. The second pair was two images of turtles, having the same pattern but different sizes. All four images were located within the same extent at the centered position, converted into point data sets, and projected using the WGS 1984 World Mercator coordinate system. It is visually obvious that both the eagle and turtle pairs had different spatial distributions across their respective extent. However, concerning the geometric features used in this study—namely, the distance (δ) and angle (θ) from the centroid to each point within each pair—it is also visually obvious that the eagle pair should have a relatively high PDF similarity of δ but relatively low PDF similarity of θ , and vice versa for the turtle pair. Thus, using these hypothetical patterns, we evaluated the method with regard to how it was capable of measuring their spatial pattern similarities, compared to the human visual sense.

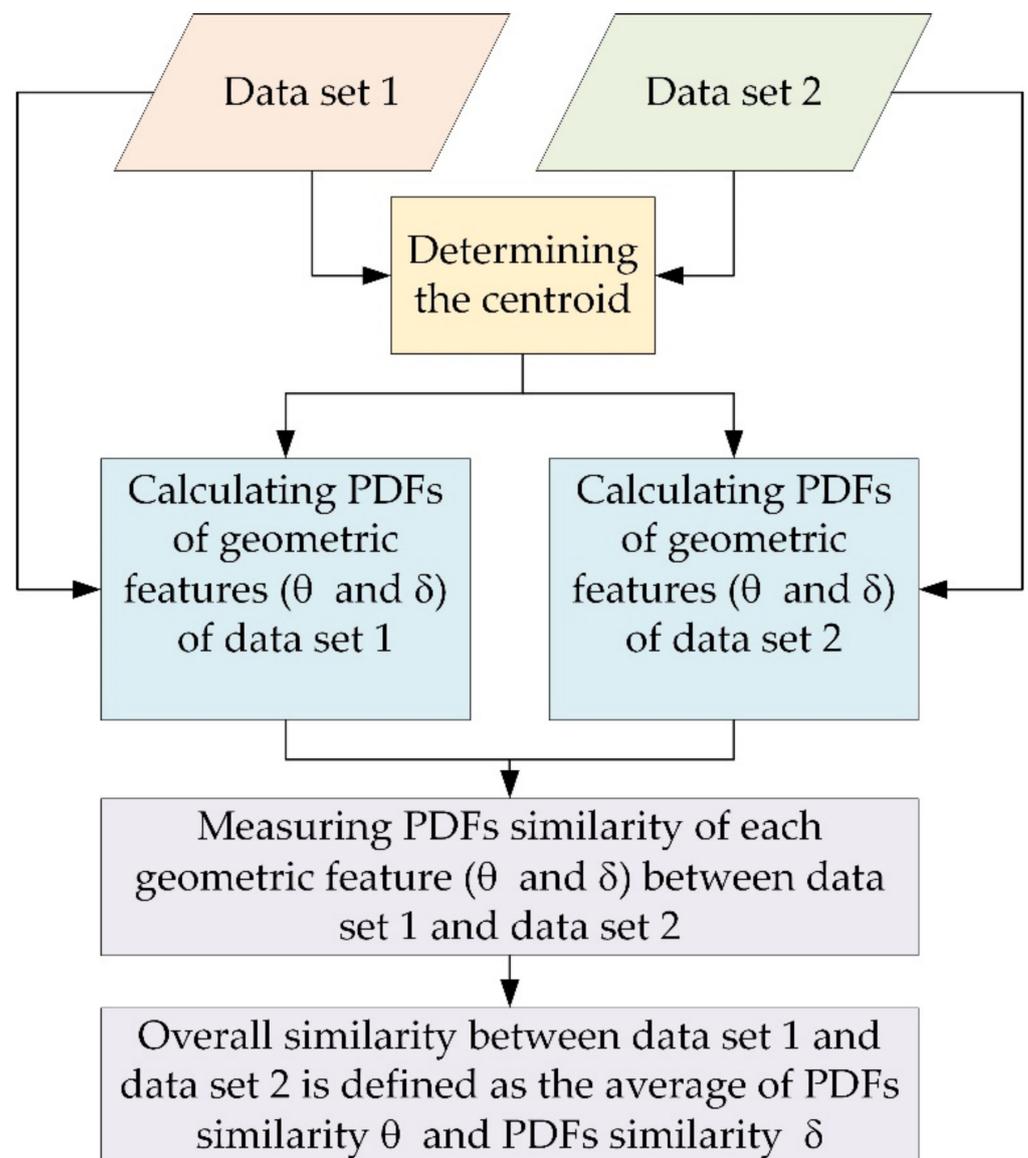


Figure 4. Overall workflow of the proposed method for measuring spatial and temporal pattern similarity between pairs of data sets.

2.1.3. Sensitivity Analyses

First, to evaluate the method's sensitivity due to systematic displacement, sensitivity analyses were carried out using two hypothetical spatial patterns: (i) a circular shape with a diameter of about 1.0567 m and (ii) a square shape with dimensions of about $1.0567 \times 1.0567 \text{ m}^2$. All patterns were projected to the same extent, converted into point data sets, and projected using the WGS 1984 World Mercator coordinate system. Each pattern was then shifted by distances of 0 m, 0.1 m, 0.2 m, ..., 1 m, 1.5 m, 2 m, and 2.5 m, and at angles of 0° , 30° , 60° , ..., 330° . Considering the dimensions of each shape, with a maximum radius of about 1.057 m for the circular shape and that about 1.494 m for the square shape, using the maximum shifting distance of 2 m in our sensitivity analyses should represent no matching at all with the origin, while the original patterns of the shapes were still maintained. Furthermore, we also aimed to evaluate the consistency of the method in measuring the systematic displacement of the shapes at certain distances in all directions, using shifting angles ranging from 0° to 330° . The similarity measures of the method were then evaluated in terms of sensitivity to systematic shifts from the origin. Hence, in this case, the geometric features used in this study (i.e., distance, δ ; and

angle, θ) were measured from the centroid of each pattern of origin, not from the centroid of the extent.

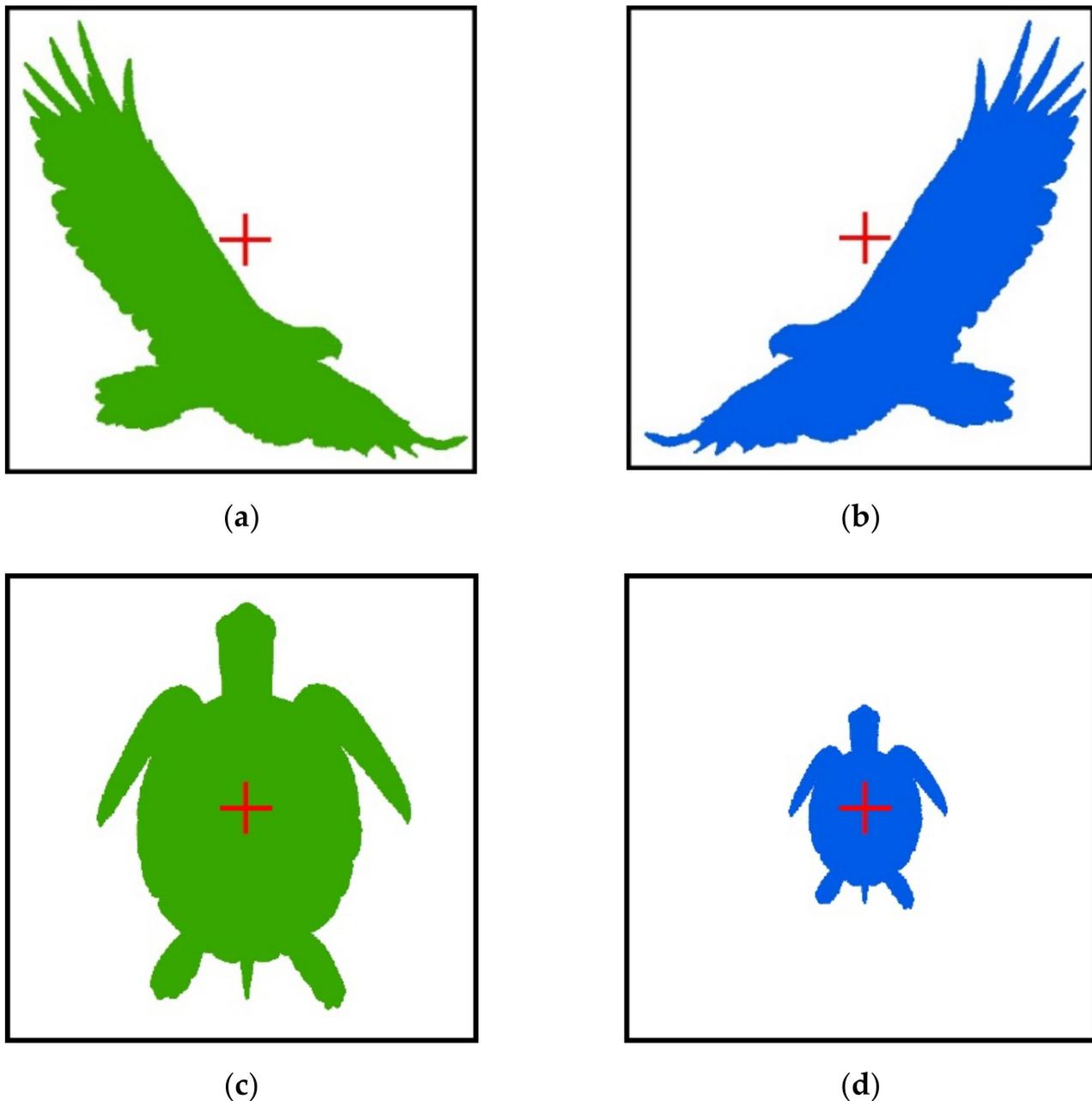


Figure 5. Two pairs of hypothetical spatial patterns used for evaluating the method in terms of its sensibility. With regard to the geometric features used in this study—namely, the distance (δ) and angle (θ) from the centroid to each point within each pair—the method should result in relatively high PDF similarity of δ but relatively low PDF similarity of θ for the eagle pair (a,b), and vice versa for the turtle pair (c,d). The red cross is the centroid of the extent.

Second, to evaluate the sensitivity of the method due to changes in spatial resolution, sensitivity analyses were carried out using two hypothetical spatial patterns of mirrored eagles, as shown in Figure 5, which were resampled at resolutions of $2\times$, $4\times$, $6\times$, \dots , $20\times$ coarser than the resolution of the origin (Figure 6). With regard to the sensitivity of the method due to changes in the spatial resolution of a pair of exactly the same patterns, the spatial pattern similarity between each eagle pattern at its original resolution, as shown in Figure 6a,b, and their resampled patterns at a resolution ranging from $2\times$ to $20\times$ coarser

were measured. Moreover, regarding the sensitivity of the method due to changes in spatial resolution of a pair of contrasting patterns, the spatial pattern similarity between the eagle pattern at its original resolution (Figure 6a) and its mirrored pattern (Figure 6b) at its original resolution, as well as at its resampled resolution ranging from $2\times$ to $20\times$ coarser, were measured.

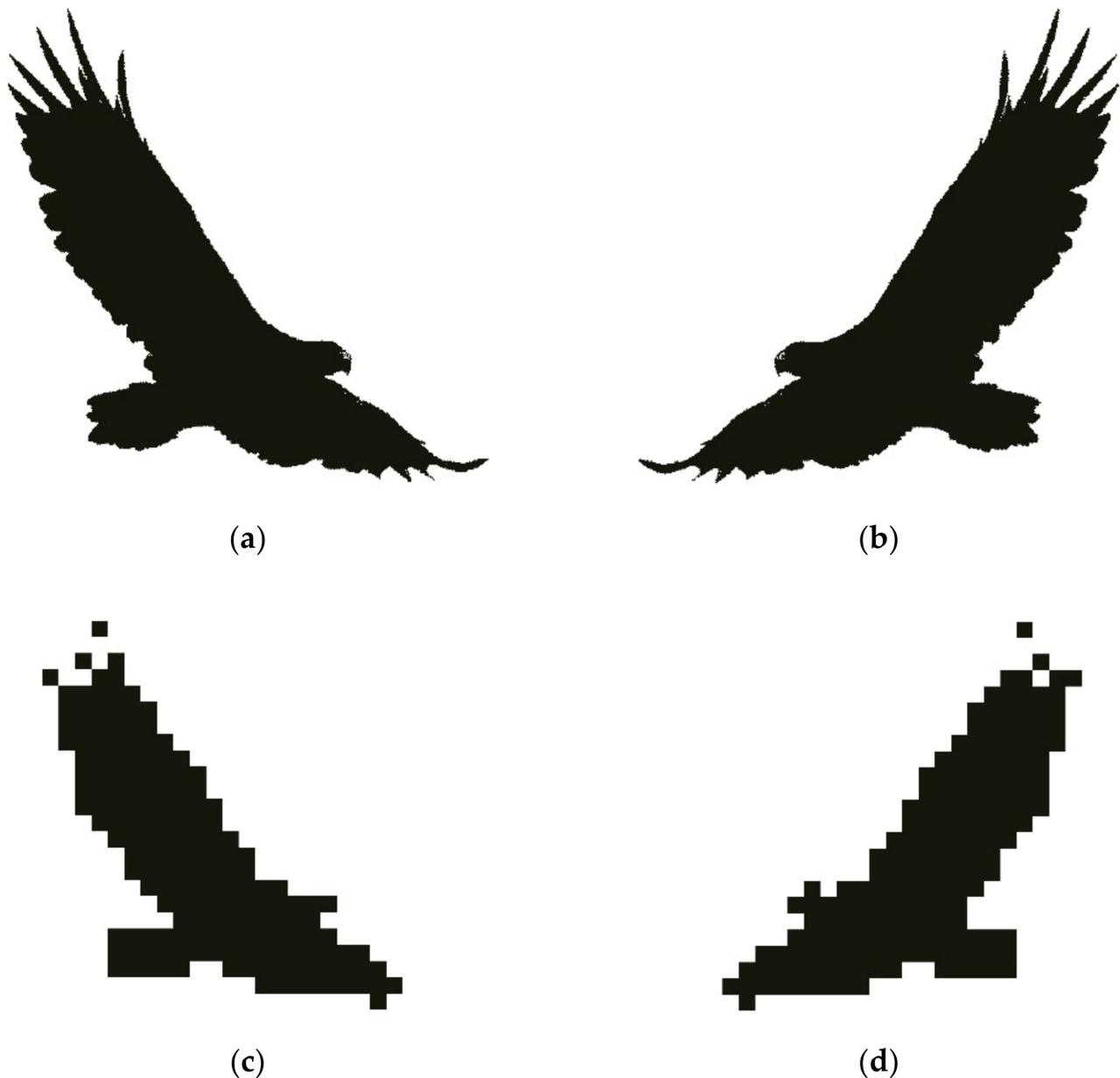


Figure 6. A pair of hypothetical mirrored eagle patterns at original resolution (a,b) and their resampled patterns at a $20\times$ coarser resolution (c,d). These hypothetical data sets were used to evaluate the method in terms of its sensitivity to changes in resolution.

Third, we also performed sensitivity analyses for hypothetical temporal data sets of deforestation by (i) shifting a hypothetical reference deforestation curve to the left from 16 days to 80 days, representing an acceleration of detection (Figure 7); (ii) shifting the reference to the right from 16 days to 80 days, representing a delay of detection; (iii) shifting the reference upward from 1 pixel to 5 pixels, representing an overestimation of detection; and (iv) shifting the reference downward from 1 pixel to 5 pixels, representing an underestimation of detection.

2.2. Deforestation Data Sets

Several deforestation hotspots in Indonesia for 2020 were detected using the 8-day composite of MODIS at a resolution of 500 m using the algorithm developed by [30]. These were compared with higher-spatial-resolution deforestation hotspots in Indonesia for 2020 based on Landsat-8 OLI data at a resolution of 30 m using the classification method described in Appendix A (see Figure 8). In this case, to synchronize the data temporally, deforestation hotspots resulting from the 8-day MODIS composite were aggregated with a 16-day composite time interval.

Deforestation hotspots derived from MODIS were determined by a threshold of change in a composite index, which was calculated using short-wave infrared and near-infrared bands [30]. Meanwhile, deforestation hotspots derived from Landsat-8 OLI were defined as pixels experiencing land clearing, according to a classification method using red, short-wave infrared, and near-infrared bands. Detailed methodology for the deforestation detection method derived from MODIS has been provided elsewhere [30], while a general description of the classification method applied to deforestation detection derived from Landsat-8 OLI is provided in Appendix A.

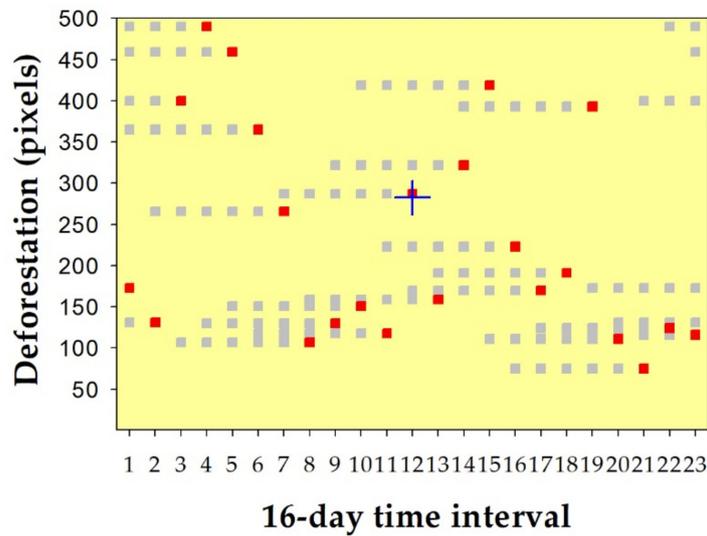


Figure 7. A hypothetical reference deforestation time-series (red squares) was shifted to the left from 16 days to 80 days (grey squares)—representing an acceleration of detection—to perform sensitivity analyses of the proposed method in terms of its robustness to measure the similarity of temporal patterns. The blue cross is the centroid of the graph.

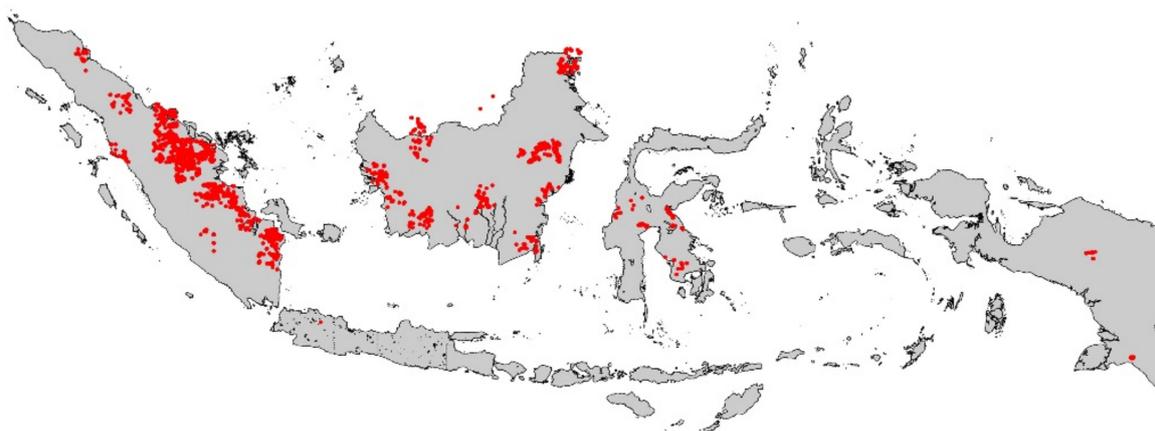


Figure 8. Deforestation hotspots (red) for 2020 derived from Landsat-8 OLI imagery.

Furthermore, within each deforestation pixel derived from MODIS, with a size of $500 \times 500 \text{ m}^2$, the deforestation intensity was calculated based on Landsat-8 OLI with a pixel size of $30 \times 30 \text{ m}^2$, ranging from 0% to 100%. In addition, with regard to the quality of Landsat-8 OLI due to cloud and haze, the clarity of pixels in Landsat-8 was also considered, ranging from 0% to 100%.

In this study, we applied our method to measure the pattern similarity—either spatial or temporal—between deforestation data sets in Indonesia for 2020 derived from MODIS, with a threshold ranging from 50 to 130, and deforestation data sets in Indonesia for 2020 derived from Landsat-8 OLI, considering clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$. All point data sets were projected using the WGS 1984 World Mercator coordinate system. The maps of deforestation data sets used in this study are shown in Figures S1–S7, while the time-series data sets used in this study are shown in Figures S8 and S9.

2.3. Tool

Based on our method described above, we developed a headless tool using the open-source modeling platform NetLogo 6.2.0, which is detailed in [31]. For the input data of the tool, all data sets used in this study were formatted into comma-separated variable (csv) format, with the first line as header, the first column as longitude for spatial data sets or time for temporal data sets, and the second column as latitude for spatial data sets or magnitude for temporal data sets. As we were dealing with combinations of pairs of data sets, either spatial or temporal, we therefore determined the centroids of the spatial and temporal data sets externally (i.e., outside the tool), which were also formatted into the csv files. The tool, including all the input data used in this study, can be downloaded from ipb.link/similarity-tool.

3. Results

3.1. Sensibility of the Method

The PDFs of angle from the centroid to each point (θ) for each pair of hypothetical patterns shown in Figure 5 clearly indicate that pair of eagle patterns had contrasting θ PDFs, while the pair of turtle patterns had exactly similar θ PDFs (Figure 9). On the other hand, Figure 10 shows that the pair of eagle patterns had exactly similar δ PDFs, while the pair of turtle patterns had contrasting δ PDFs.

Furthermore, using the nine metrics listed in Table 1, the PDF similarities of θ and δ for each pair shown in Figures 9 and 10 were measured. In this case, the overall similarity was calculated as the average PDF similarity for θ and δ . Table 2 indicates that the spatial patterns of pair of eagle shapes shown in Figure 5 had relatively low to fair similarity, in terms of θ , and very high similarity in terms of δ , which was consistent across the nine metrics; vice versa, the pair of turtle shapes has very high similarity for θ and relatively low to fair similarity for δ , which was also consistent across the nine metrics.

In addition, although the Soergel, Ruzicka, and Tanimoto metrics use different approaches to measure similarity between PDF pairs (Table 1), they consistently resulted in exactly the same results. Thus, henceforth, we use the term Soergel–Ruzicka–Tanimoto to refer to these three identical metrics.

3.2. Sensitivity of the Method

Figures 11–13 show the sensitivity of the method to the systematic displacement of two hypothetical shapes—a circle with a diameter of about 1.0567 m and a square with dimensions $1.0567 \times 1.0567 \text{ m}^2$ —in terms of how the nine similarity metrics used in our method measure changes in spatial pattern similarity due to shifting from the origin, with regard to the similarity of the angle from the centroids of origins (Figure 11), the similarity of the distance from the centroids of origins (Figure 12), and the overall similarity (Figure 13). It is obvious that the method performed consistently across these three types of similarity: θ , δ , and overall similarity (i.e., the average of the θ and δ similarities). The Fidelity metric produced the most optimistic similarity measure, bordering the upper

edge of the envelopes, while the Soergel–Ruzicka–Tanimoto metric produced the most pessimistic similarity measure, bordering the lower edge of the envelopes.

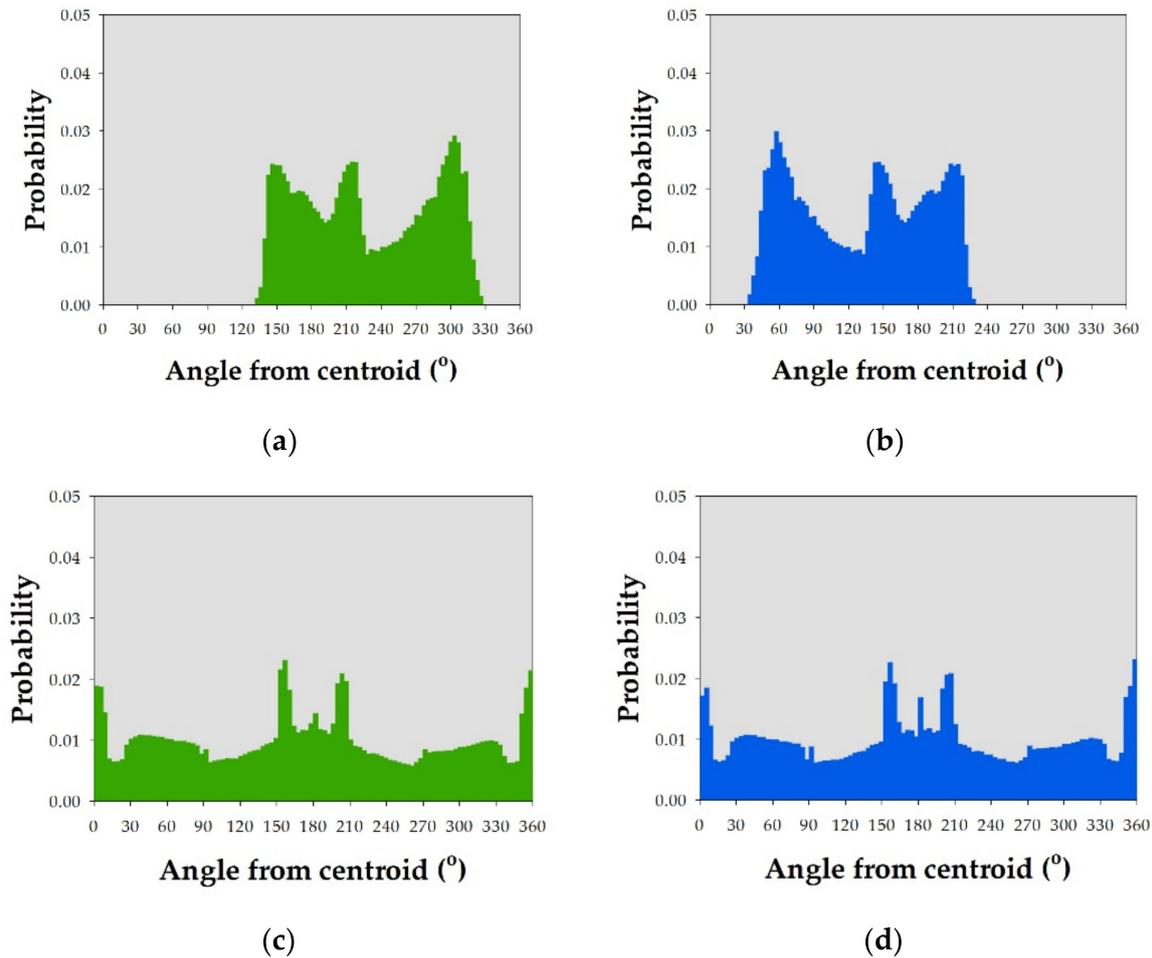


Figure 9. PDFs of angle from the centroid to each point (θ) for each pair of hypothetical patterns shown in Figure 5: (a,b) the pair of eagle patterns and (c,d) the pair of turtle patterns.

Table 2. Spatial pattern similarity between two pairs of hypothetical patterns (shown in Figure 5), measured using nine metrics (as listed in Table 1), in terms of the geometric features used in this study (i.e., angle, θ , and distance, δ , from the centroid of the extent to each point). The overall similarity was calculated as the average of similarity of θ and similarity of δ .

Metric	Pair of Eagle Patterns			Pair of Turtle Patterns		
	θ	δ	Overall	θ	δ	Overall
Sørensen	0.48	1.00	0.74	0.98	0.39	0.68
Soergel	0.31	1.00	0.66	0.95	0.24	0.60
Intersection	0.48	1.00	0.74	0.98	0.39	0.68
Ruzicka	0.31	1.00	0.66	0.95	0.24	0.60
Tanimoto	0.31	1.00	0.66	0.95	0.24	0.60
Cosine	0.53	1.00	0.77	1.00	0.53	0.76
Jaccard	0.36	1.00	0.68	1.00	0.33	0.66
Dice	0.53	1.00	0.77	1.00	0.50	0.75
Fidelity	0.53	1.00	0.76	1.00	0.67	0.83

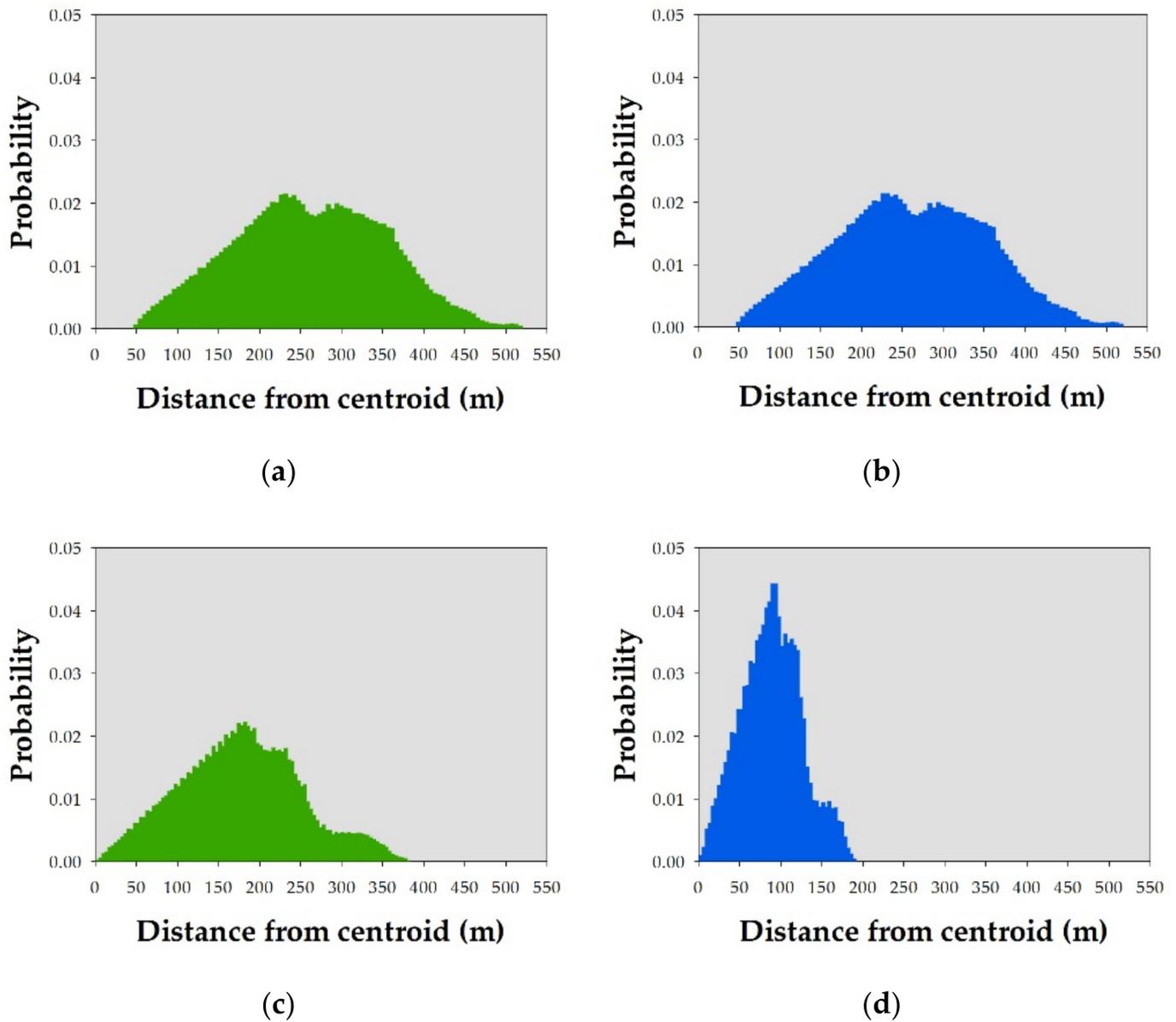


Figure 10. PDFs of distance from the centroid to each point (δ) for each pair of hypothetical patterns shown in Figure 5: (a,b) pair of eagle patterns and (c,d) pair of turtle patterns.

Figure 14 shows the sensitivity of the method to changes in spatial resolution using the number of pairs of hypothetical shapes with different resolutions: (a) the eagle pattern shown in Figure 6a compared with itself at different resolutions, ranging from 1× to 20× coarser; (b) the eagle pattern shown in Figure 6b compared to itself at different resolutions, ranging from 1× to 20× coarser; and (c) the eagle pattern shown in Figure 6a compared with its mirrored pattern shown in Figure 6b at different resolutions, ranging from 1× to 20× coarser. Consistent with the previous results, the Fidelity metric was the most optimistic similarity measure, bordering the upper edge of the envelopes, while the metric of Soergel–Ruzicka–Tanimoto was the most pessimistic similarity measure, bordering the lower edge of the envelopes.

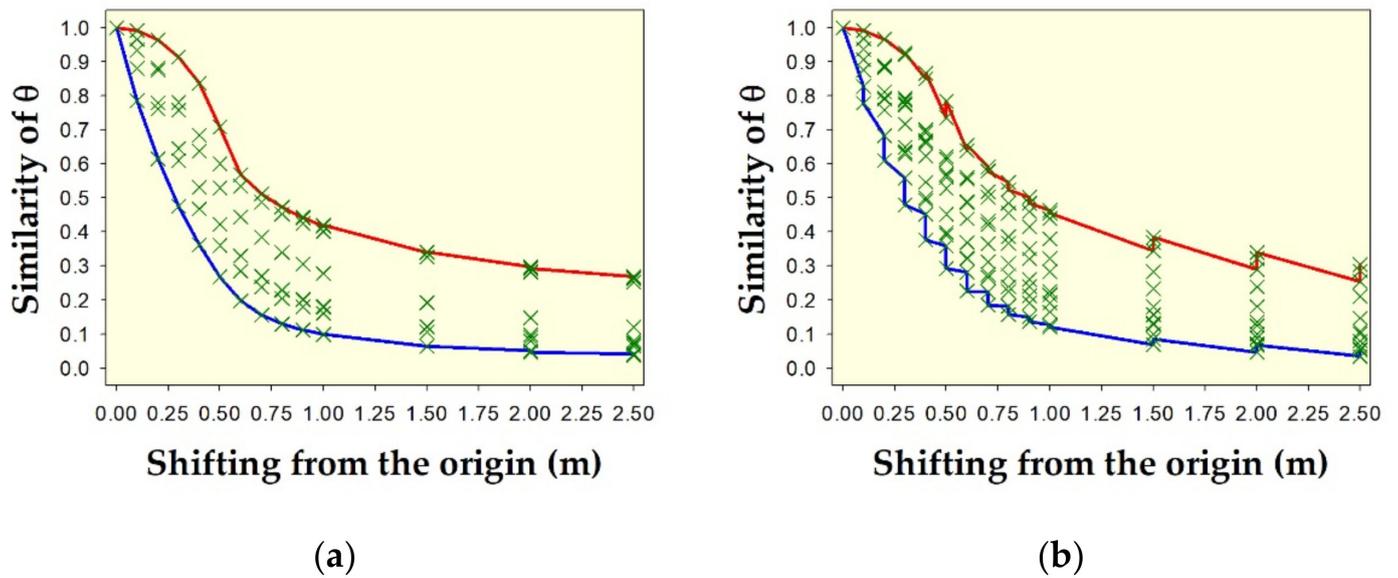


Figure 11. PDF similarity of angle from the centroid of origin to each point (θ) of two hypothetical shapes: (a) circle with diameter of about 1.0567 m, and (b) square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$, due to shifting from the origin in all directions. Green crosses indicate similarity of θ from all nine metrics. Red lines on the upper edge of the envelopes indicate the similarity of θ using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the similarity of θ using the Soergel–Ruzicka–Tanimoto metric.

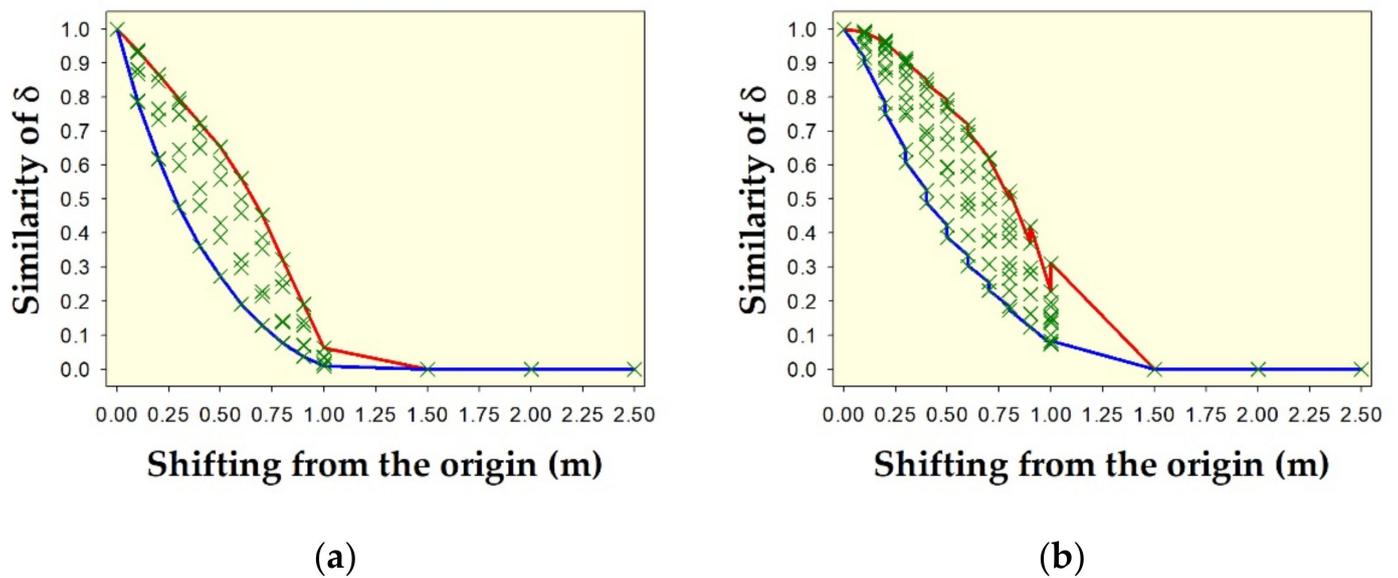


Figure 12. PDF similarity of distance from the centroid of origin to each point (δ) of two hypothetical shapes: (a) circle with diameter of about 1.0567 m, and (b) square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$, due to shifting from the origin in all directions. Green crosses indicate the similarity of δ from all nine metrics. Red lines on the upper edge of the envelopes indicate the similarity of δ using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the similarity of δ using the metric of Soergel–Ruzicka–Tanimoto.

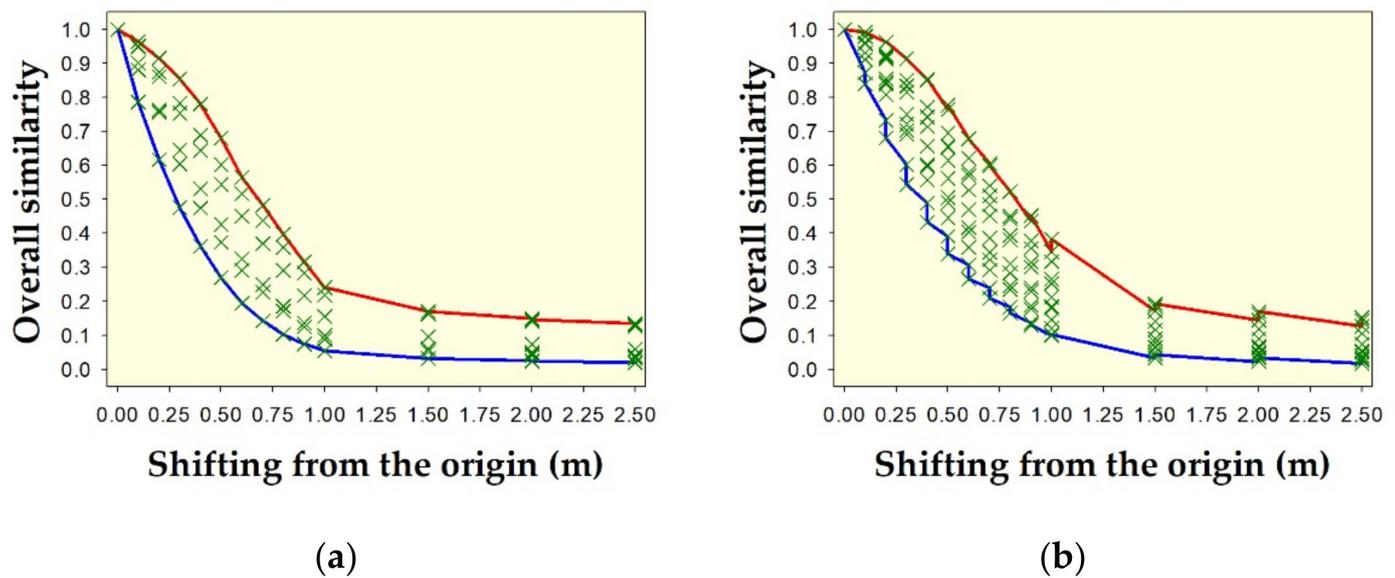


Figure 13. Overall spatial pattern similarity (i.e., the average of similarity of θ and similarity of δ) for two hypothetical shapes: (a) circle with diameter of about 1.0567 m, and (b) square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$, due to shifting from the origin in all directions. Green crosses indicate overall similarity from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

These findings indicate that the similarity measures resulting from the method decreased consistently when the resolutions of the pairs were much coarser than the original image at relatively low decreasing rates: of about -0.007 , -0.009 , and -0.005 per km resolution decrease from the original in Figure 14a–c, respectively. This suggests that, when the method is applied to measure the similarity of pairs of data sets with different resolutions, the result should be calibrated with regard to the resolution difference effect at rates ranging from -0.005 to -0.009 per km resolution difference from the origin.

Figure 15 shows the method's sensitivity to the systematic displacement of hypothetical reference deforestation time-series data due to acceleration, delay, overestimation, and underestimation of detection. The results suggest that the method was less sensitive to acceleration or delay (Δx) but relatively very sensitive to overestimation or underestimation (Δy). The similarity of the two temporal data sets decreased at rates of about -0.0009 per day of departure due to acceleration or delay, while it decreased at rates of about -0.004 per pixel of departure due to overestimation or underestimation. With regard to deforestation alerts coming from different sources, this implies that a relatively high temporal pattern similarity (of about 0.9) between two data sets derived from different monitoring systems can imply a time lag of about 111 days or a magnitude lag of about 25 pixels. Again, still consistent with the previous results, the Fidelity metric was the most optimistic similarity measure, bordering the upper edge of the envelopes, while the metric of Soergel–Ruzicka–Tanimoto was the most pessimistic similarity measure, bordering the lower edge of the envelopes.

3.3. Spatial Pattern Similarity of Two Deforestation Data Sets with Different Spatial Resolutions

Our method yielded relatively high overall similarity (i.e., the average of similarity of θ and similarity of δ) for the spatial pattern between the deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI using the classification method described in Appendix A (at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$), and the deforestation hotspots in Indonesia in 2020 derived from MODIS using the change-detection algorithm developed by [30] (at change-detection thresholds ranging from 50 to 130). The similarity from all nine metrics ranged from about 0.81 to about 1.00 (see Figure 16). Although at the scale

of Figure 16, the upper edge of the envelopes—indicating the Fidelity metric (red line)—seemed relatively insensitive to changes of MODIS-detection thresholds, Table 3 suggests that a threshold of about 100 was optimal. However, the lower edge of the envelopes—indicating the metric of Soergel–Ruzicka–Tanimoto (blue line)—suggests that the most optimal MODIS detection threshold was about 80, with regard to spatial pattern similarity.

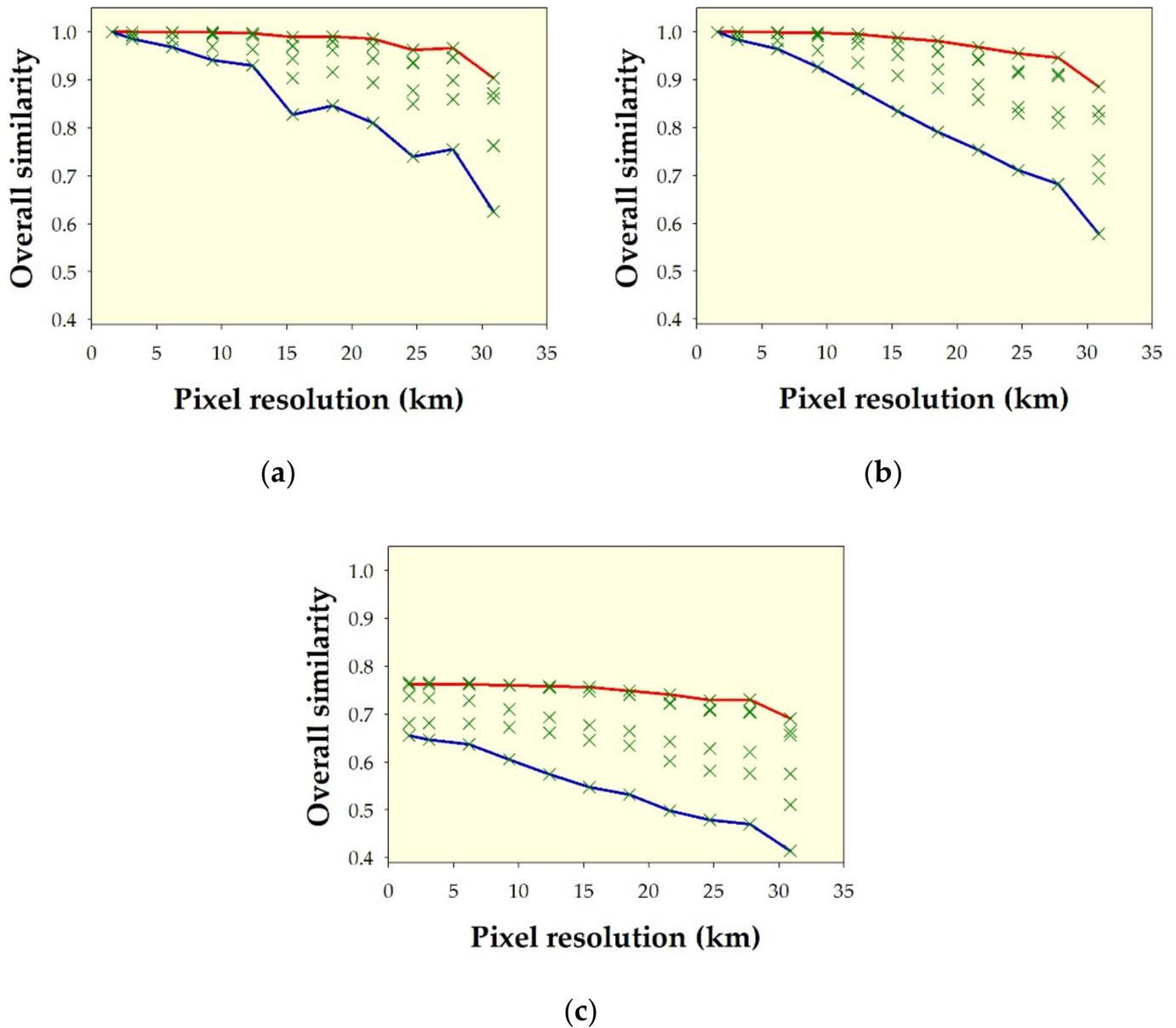


Figure 14. Overall spatial pattern similarity (i.e., the average of similarity of θ and similarity of δ) of various pairs of hypothetical shapes with different resolutions: (a) eagle pattern (shown in Figure 6a) compared with itself at different resolutions, ranging from 1× to 20× coarser; (b) eagle pattern (shown in Figure 6b) compared with itself at different resolutions, ranging from 1× to 20× coarser; and (c) eagle pattern (shown in Figure 6a) compared with its mirrored pattern (shown in Figure 6b) at different resolutions, ranging from 1× to 20× coarser. Green crosses indicate overall similarity from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

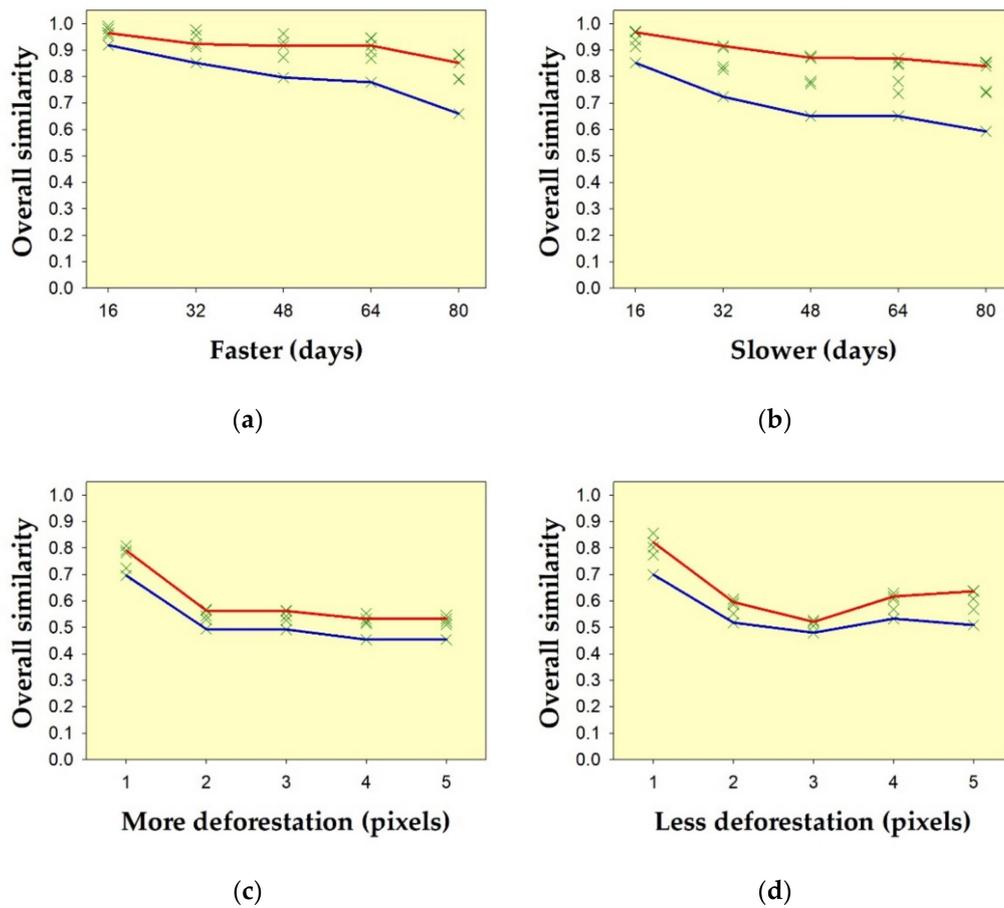


Figure 15. Sensitivity of the method on hypothetical temporal deforestation data sets, due to: (a) acceleration; (b) delay; (c) overestimation; and (d) underestimation of deforestation detection, compared to the reference. Green crosses indicate overall similarity (i.e., the average of similarity of θ and similarity of δ) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

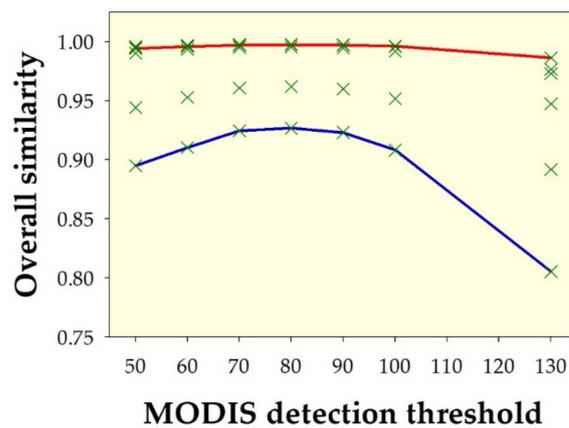


Figure 16. Overall similarity (i.e., the average of θ and δ similarities) of the spatial pattern between deforestation hotspots in Indonesia in 2020, derived from Landsat-8 OLI (at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (at change-detection thresholds ranging from 50 to 130). Green crosses indicate overall similarity from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

Table 3. Overall similarity (i.e., the average of θ and δ similarities) using nine metrics of spatial pattern between deforestation hotspots in Indonesia in 2020, derived from Landsat-8 OLI (at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (at change-detection thresholds ranging from 50 to 130).

Metric	MODIS Change-Detection Threshold						
	≥ 50	≥ 60	≥ 70	≥ 80	≥ 90	≥ 100	≥ 130
Sørensen	0.9444	0.9530	0.9607	0.9620	0.9600	0.9518	0.8920
Soergel	0.8950	0.9104	0.9246	0.9268	0.9232	0.9083	0.8054
Intersection	0.9444	0.9530	0.9607	0.9620	0.9600	0.9518	0.8920
Ruzicka	0.8950	0.9104	0.9246	0.9268	0.9232	0.9083	0.8054
Tanimoto	0.8950	0.9104	0.9246	0.9268	0.9232	0.9083	0.8054
Cosine	0.9960	0.9970	0.9977	0.9978	0.9975	0.9964	0.9765
Jaccard	0.9904	0.9934	0.9953	0.9957	0.9947	0.9920	0.9474
Dice	0.9952	0.9967	0.9977	0.9978	0.9974	0.9960	0.9729
Fidelity	0.9942	0.9960	0.9971	0.9974	0.9971	0.9963	0.9866

3.4. Temporal Pattern Similarity of Two Deforestation Data Sets with Different Spatial Resolutions

Our method resulted in relatively low to fair overall similarity of temporal patterns between deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI using the classification method described in Appendix A, at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$, and deforestation hotspots in Indonesia in 2020, derived from MODIS with the change-detection algorithm developed by [30], at change-detection thresholds ranging from 50 to 130; the similarity from all nine metrics ranged from about 0.25 to about 0.71 (Figure 17 and Table 4). Both the upper edge of the envelopes from the metric of Fidelity (red line) and the lower edge of the envelopes from the metric of Soergel–Ruzicka–Tanimoto (blue line) suggest that increasing MODIS detection threshold above 90 will consistently increase the temporal pattern similarity with deforestation hotspots derived from Landsat-8 OLI.

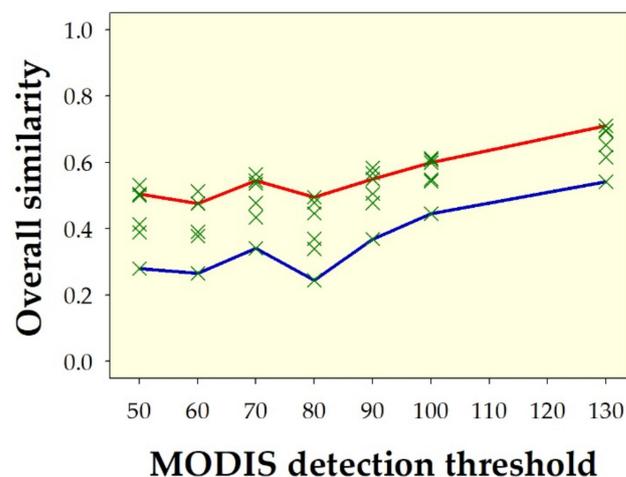


Figure 17. Overall similarity (i.e., the average of similarity of θ and similarity of δ) of temporal pattern between deforestation hotspots in Indonesia in 2020, derived from Landsat-8 OLI at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$, and deforestation hotspots in Indonesia in 2020, derived from MODIS at change-detection thresholds ranging from 50 to 130. Green crosses indicate overall similarity from all nine metrics. Red lines on the upper edge of the envelopes indicate overall similarity using metric of Fidelity. Blue lines on the upper edge of the envelopes indicate overall similarity using metric of Soergel–Ruzicka–Tanimoto.

Table 4. Overall similarity (i.e., the average of similarity of θ and similarity of δ) using nine metrics of temporal pattern between deforestation hotspots in Indonesia in 2020, derived from Landsat-8 OLI at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$, and deforestation hotspots in Indonesia in 2020, derived from MODIS at change-detection thresholds ranging from 50 to 130.

Metric	MODIS Change-Detection Threshold						
	≥ 50	≥ 60	≥ 70	≥ 80	≥ 90	≥ 100	≥ 130
Sørensen	0.4130	0.3913	0.4783	0.3696	0.4783	0.5435	0.6522
Soergel	0.2797	0.2664	0.3417	0.2446	0.3690	0.4456	0.5416
Intersection	0.4130	0.3913	0.4783	0.3696	0.4783	0.5435	0.6522
Ruzicka	0.2797	0.2664	0.3417	0.2446	0.3690	0.4456	0.5416
Tanimoto	0.2797	0.2664	0.3417	0.2446	0.3690	0.4456	0.5416
Cosine	0.5327	0.5127	0.5653	0.4792	0.5841	0.6124	0.6968
Jaccard	0.3896	0.3787	0.4349	0.3395	0.5059	0.5491	0.6151
Dice	0.5005	0.4769	0.5365	0.4464	0.5675	0.6072	0.6950
Fidelity	0.5055	0.4767	0.5454	0.4956	0.5500	0.5998	0.7107

4. Discussion

4.1. Spatial Sensibility

In terms of sensibility, we demonstrated that our method could measure at a human visual level of sensing two pairs of hypothetical patterns (Figure 5), showing reasonable results (Figures 9 and 10 and Table 2). As we described previously in the Methods section, we used two geometric features of point patterns—the distance and angle from the centroid to each point within the extent. For further studies, other geometric features (e.g., the square root of the area of the triangle between three random points from the centroid [27]) may be explored.

4.2. Spatial Sensitivity

In addition, the results of the sensitivity analyses performed in this study suggest that, when the method is applied for measuring the similarity of pairs of spatial data sets with different resolutions, the result should be calibrated, with regard to the resolution difference effect, at rates ranging from -0.005 to -0.009 per km resolution difference from the origin. This implies that, when deforestation hotspots derived from MODIS with a resolution of 500 m are compared to those derived from Landsat-8 OLI with a resolution of 30 m, their similarity, as measured using the method, should be calibrated with a resolution lag of about 0.470 km such that that the calibrated similarity will be around 0.0024 to 0.0042 higher than the measurement.

Moreover, we also found that our method consistently produced envelopes, with upper and lower edges always being bordered by the Fidelity and Soergel–Ruzicka–Tanimoto metrics, respectively. Thus, we suggest a new single metric that considers both optimistic and pessimistic metrics, which is the average of these two abovementioned metrics, which we denote by Ruzicka–Fidelity, as follows:

$$S = \frac{\left(\frac{\sum_{i=1}^n \min(p_i, q_i)}{\sum_{i=1}^n \max(p_i, q_i)} \right) + \left(\sum_{i=1}^n \sqrt{p_i q_i} \right)}{2}, \tag{1}$$

where S denotes the similarity measure for a pair of PDFs (p and q) with n ordinal classes. In this case, we selected Ruzicka among the other two (Soergel and Tanimoto) metrics, as Ruzicka provides the most sophisticated equation compared with Soergel and Tanimoto (see Table 1).

We applied Equation (1) to perform sensitivity analyses of the systematic displacement of two hypothetical shapes: (i) a circle with a diameter of about 1.0567 m and (ii) a

square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$; our previous results regarding the overall similarity (as shown in Figure 13) were enhanced by the Ruzicka–Fidelity metric, as shown in Figure 18. Furthermore, Figure 18 suggests that the Ruzicka–Fidelity metric provides a more sensible measure, in terms of its proportionality, as it resulted in an overall similarity ranging from 0.4 to 0.5 when displacement from the origin took place at about half the size of the maximum radii of the shapes and decreased to relatively low similarity (of about 0.1–0.2) when displacement from the origin took place at about the maximum radii of the shapes. The results also suggest that, when displacement from the origin took place much further than the maximum radii of the shapes, the method still appreciated the pattern similarity between the shifted shapes and their origins, with an overall similarity between 0 and 0.1, which is not recognized in conventional accuracy-assessment methods, where the agreement between pairs of spatial patterns is measured based on their point-to-point degree of matching. Appendix B illustrates how our method, based on pattern-similarity analysis, differs from conventional omission/commission disagreement approaches using a confusion matrix.

With regard to spatial data sets of deforestation hotspots derived from two satellite imageries with different resolutions (e.g., MODIS and Landsat-8 OLI), results suggested that: (i) when the higher resolution data set (Landsat-8 OLI) is considered as the reference, spatial pattern similarity between the two data sets is relatively poor, indicating relatively far departure from the reference such that the change-detection algorithm of the lower-resolution data set (MODIS) should be revisited, but (ii) when both data sets (MODIS and Landsat-8 OLI) are considered as two versions of information on deforestation hotspots, then the classification method used by the higher-resolution data set (Landsat-8 OLI) should also be revisited.

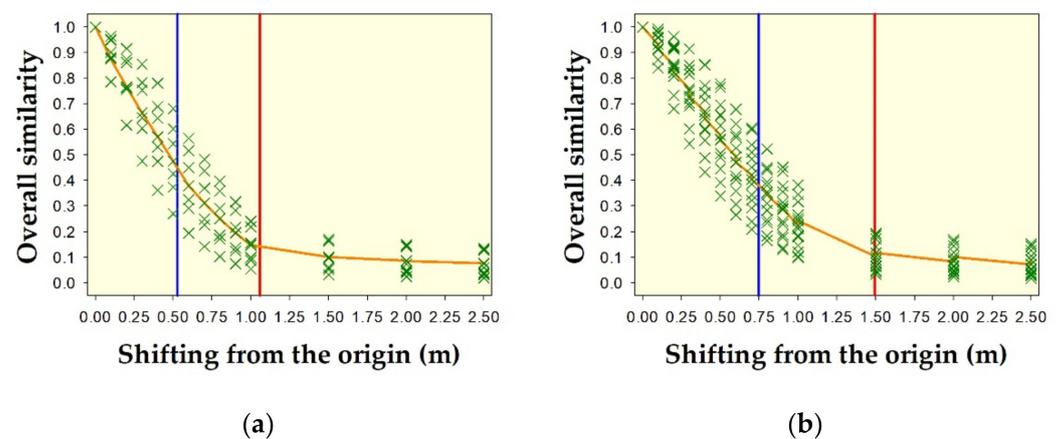


Figure 18. Overall spatial pattern similarity (i.e., the average of similarity of θ and similarity of δ) of two hypothetical shapes: (a) circle with diameter of about 1.0567 m and (b) square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$, due to shifting from the origin in all directions. Green crosses indicate overall similarity from all 10 metrics. Orange lines in the middle of the envelopes indicate overall similarity using the new metric of Ruzicka–Fidelity (Equation (1)). Vertical blue lines indicate half size of the maximum radii of the shapes, while vertical red lines indicate the maximum radii of the shapes.

4.3. Temporal Sensitivity

In terms of the sensitivity of our method to displacement in temporal data (Figure 15), our method was found to be much more sensitive in similarity measures to changes in the y -axis (Δy) due to overestimation or underestimation, rather than changes in the x -axis (Δx) due to acceleration or delay; as such, it may be constrained by different dimensions of the x -axis or y -axis of temporal data. Thus, we performed rescaling of our hypothetical time-series data such that Figure 7 was transformed into Figure 19.

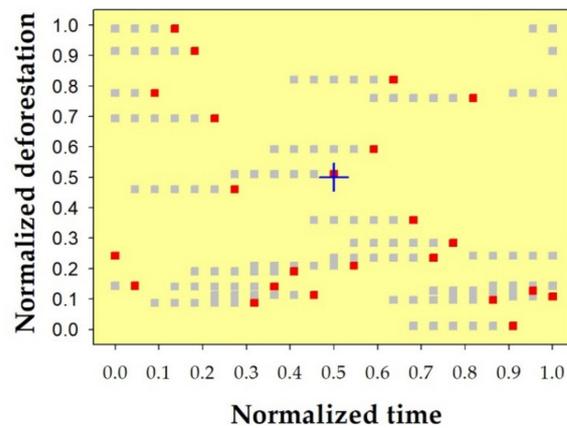


Figure 19. Normalized hypothetical time-series data as shown in Figure 7. Red squares indicate the origin; grey squares indicate left shifts from the origin, representing an acceleration of detection from 16 days to 80 days. The blue cross is the centroid of the graph.

We found that the method still had similar sensitivity as our previous results shown in Figure 15, where the method was more sensitive to changes in y -axis variables than changes in x -axis variables (Figure 20). Thus, we argue that rescaling the x - and y -axes of time-series data before applying our method should not be performed due to the persistence of the sensitivity results. Moreover, we also argue that rescaling time-series data can degrade the fundamental meaning of the results, as shown by Figures 15 and 20, where the level of similarity due to a shift in the x -axis dropped from a relatively high similarity to a relatively low similarity when the data were normalized.

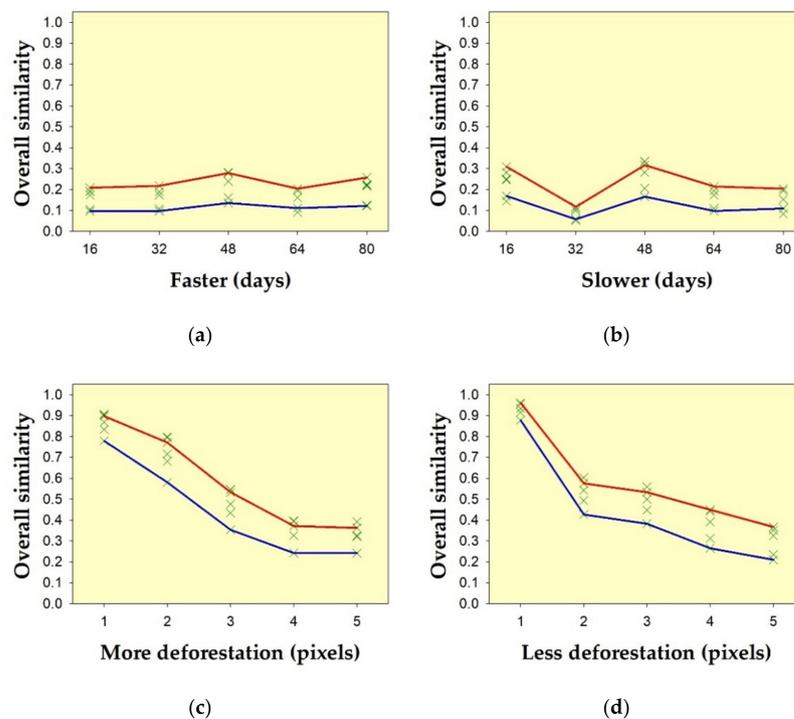


Figure 20. Sensitivity of the method on hypothetical deforestation temporal data sets after normalization, due to possibility of: (a) acceleration, (b) delay, (c) overestimation, and (d) underestimation of deforestation detection, compared to the reference. Green crosses indicate the overall similarity (i.e., the average similarity of θ and similarity of δ) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

We then performed linear interpolation on our hypothetical deforestation time-series data shown in Figure 7 to downscale the temporal resolution from a 16-day time interval to a 1-day, 1-h, and 1-min interval, aimed at assessing the sensitivity of our method to displacement in temporal data with regard to the temporal resolution. Figure 21 shows the interpolated results at a time slice with a 1-day interval for our hypothetical deforestation time-series data, as well as their displacement, in terms of acceleration of detection from 16 days to 80 days (see Figure 7).

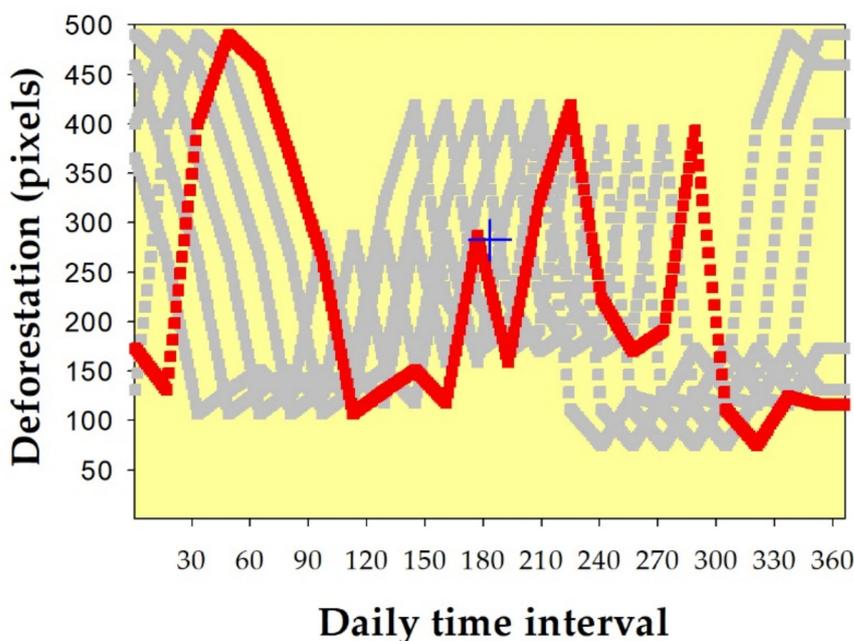


Figure 21. Interpolated hypothetical deforestation time-series data shown in Figure 7 at a time slice of a 1-day interval. Red squares indicate the original time-series; grey squares indicate left shifts from the original, representing an acceleration on detection from 16 days to 80 days. The blue cross is the centroid of the graph.

The sensitivity results to changes in y -axis variables (Δy) due to overestimation or underestimation, as well as changes in x -axis variables (Δx) due to acceleration or delay, using linearly interpolated data at time-slices of 1-day, 1-h, and 1-min intervals are shown in Figures 22–24, respectively, and summarized in Table 5. The results suggested that the most sensible sensitivity of our method to changes in y -axis variables (Δy), due to overestimation or underestimation, as well as changes in x -axis variables (Δx), due to acceleration or delay, was given by temporal data sets at a 1-day temporal resolution.

Table 5. Sensitivity summary of the method on hypothetical temporal data sets on deforestation after interpolation at a time slice of 16 days, 1 day, 1 h, and 1 min, due to the possibility of (a) acceleration, (b) delay, (c) overestimation, and (d) underestimation of deforestation detection, compared to the reference.

Time Interval	Equidistance Interpolation Slice from the Original time Interval	Linear Slope, Indicating Sensitivity Due to			
		Acceleration, from 16 to 80 Days from the Reference	Delay, from 16 to 80 Days from the Reference	Overestimation, 1 Pixel to 5 Pixels from the Reference	Underestimation, 1 Pixel to 5 Pixels from the Reference
16-day	1×	−0.0026	−0.0027	−0.0526	−0.0379
Daily	16×	−0.0026	−0.0045	−0.0230	−0.0214
Hourly	384×	−0.0026	−0.0027	−0.0034	−0.0033
Minutely	34560×	0.0001	0.0000	0.0000	0.0000

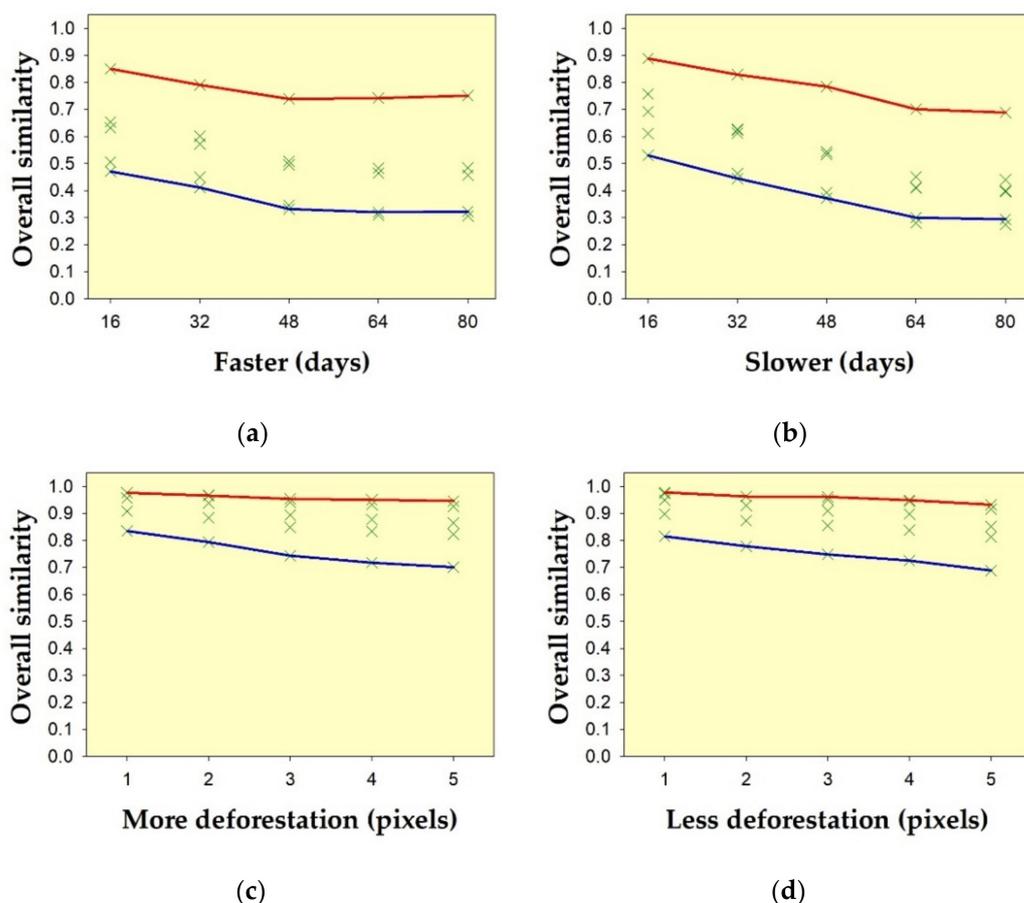


Figure 22. Sensitivity of the method on hypothetical temporal data sets on deforestation after interpolation at a time slice of 1 day, due to the possibility of (a) acceleration, (b) delay, (c) overestimation, and (d) underestimation of deforestation detection, compared to the reference. Green crosses indicate overall similarity (i.e., the average of θ and δ similarities) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

Thus, we applied linear interpolation to temporal deforestation data in Indonesia for 2020 to downscale its temporal resolution from a 16-day interval to a 1-day interval (Figure S9). The overall similarity (i.e., the average of θ and δ similarities) of the temporal pattern between deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI (at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (at change-detection thresholds ranging from 50 to 130), after interpolation with a 1-day interval, is shown in Figure 25.

Moreover, relatively high overall similarity (0.81–1.00) was obtained by our method when it was applied to measuring the spatial pattern similarity between deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI (using the classification method described in Appendix A, at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (using the change-detection algorithm developed by [30], at change-detection thresholds ranging from 50 to 130). Figure 16 suggests that the spatial patterns of the deforestation results from those two data sets were, in fact, similar. This result is confirmed by Figures S1–S7, where the patterns between pairs of data sets are visually similar for all MODIS-detection thresholds. In addition, when the effect of resolution lag between MODIS and Landsat-8 OLI of about 0.470 km was considered, the calibrated overall similarity could still be a bit higher.

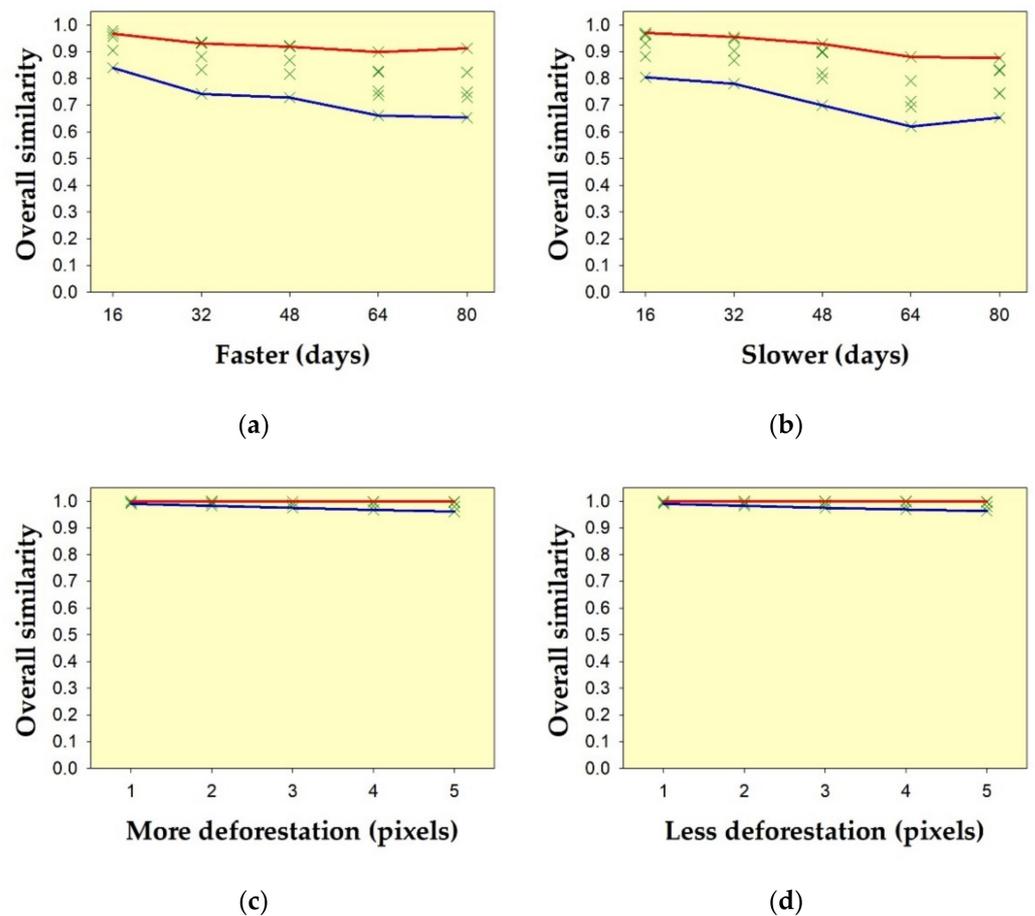


Figure 23. Sensitivity of the method on hypothetical temporal data sets on deforestation after interpolation at a time slice of 1 h, due to the possibility of (a) acceleration, (b) delay, (c) overestimation, and (d) underestimation of deforestation detection, compared to the reference. Green crosses indicate overall similarity (i.e., the average of θ and δ similarities) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

Using Equation (1) to perform similarity analyses for either spatial or temporal patterns of deforestation between MODIS and Landsat-8 OLI, as shown in Figures 16 and 25, we can obtain a clearer result (Figure 26), which suggests that the MODIS-based deforestation monitoring system developed by [30] can still be calibrated—in terms of increasing its change-detection threshold parameter to above 130—to increase its temporal pattern similarity with the reference derived from Landsat-8 OLI without significantly losing spatial pattern similarity.

4.4. Further Development

Finally, other than providing a generic method to measure both spatial and temporal pattern similarities between deforestation data sets derived from various monitoring systems, the method developed in this paper may also be applied beyond the remote-sensing domain, for example, in ecological modeling, to validate the simulated data of an ecological model with the reference data or to compare the similarity between the simulated data of an ecological model and simulated data derived from other models. Our method is efficient in terms of computation and, thus, applies to big data. The method will further be developed by integrating spatial and temporal point data sets in 3D point cloud xyz-coordinate format—where x is the longitude, y is the latitude, and z is time—to obtain a single spatio-temporal pattern similarity measure.

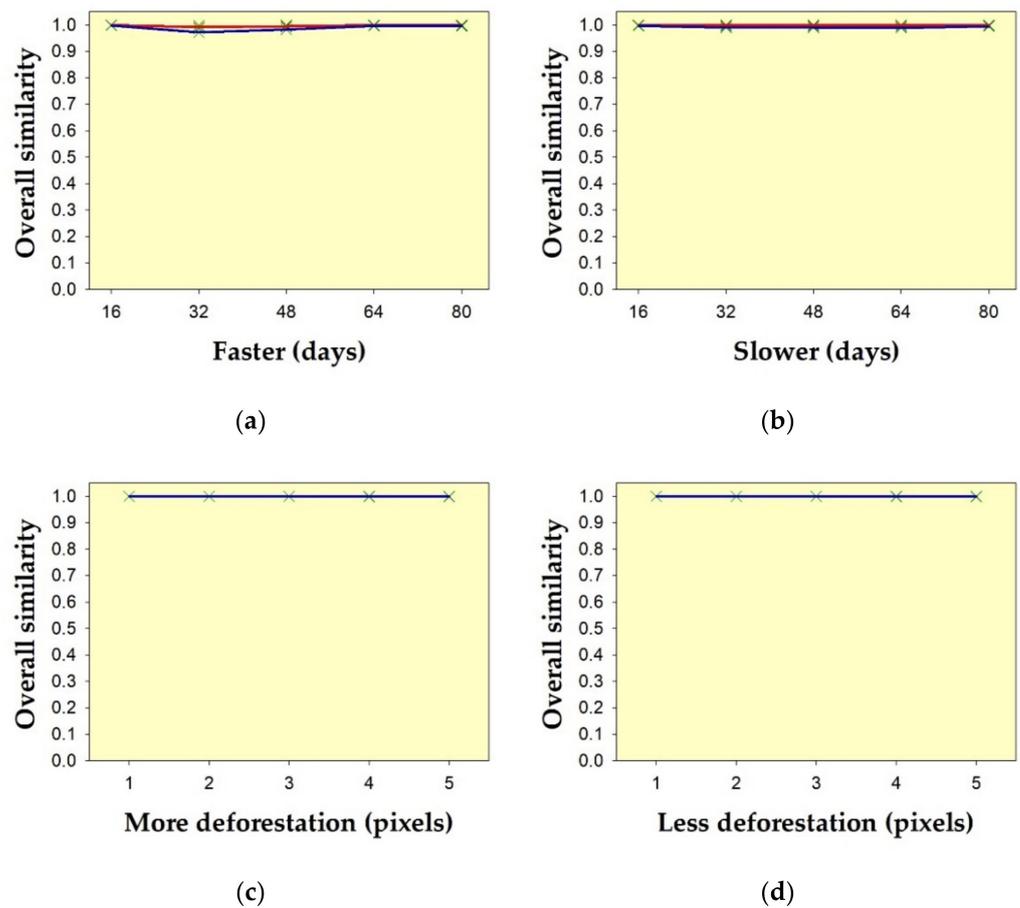


Figure 24. Sensitivity of the method on hypothetical temporal data sets on deforestation after interpolation at a time slice of 1 min, due to the possibility of (a) acceleration, (b) delay, (c) overestimation, and (d) underestimation of deforestation detection, compared to the reference. Green crosses indicate overall similarity (i.e., the average of θ and δ similarities) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

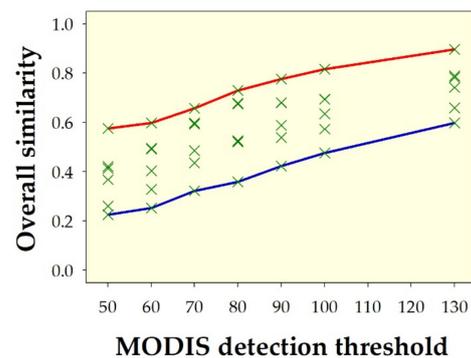


Figure 25. Overall similarity (i.e., the average of θ and δ similarities) of the temporal pattern between deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI (at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (at change-detection thresholds ranging from 50 to 130), after interpolation at a time slice of 1 day. Green crosses indicate overall similarity (i.e., the average of θ and δ similarities) from all nine metrics. Red lines on the upper edge of the envelopes indicate the overall similarity using the Fidelity metric. Blue lines on the lower edge of the envelopes indicate the overall similarity using the metric of Soergel–Ruzicka–Tanimoto.

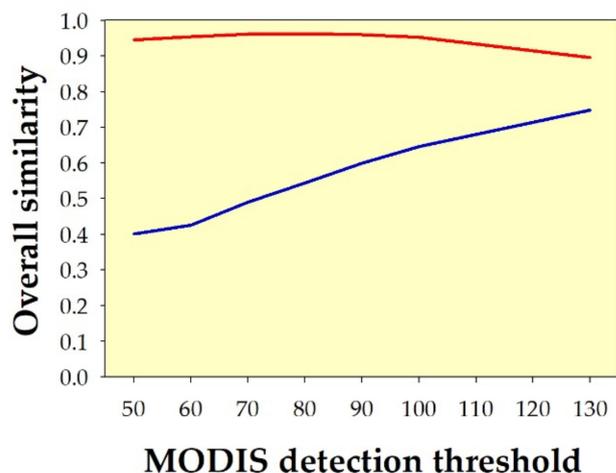


Figure 26. Overall similarity (i.e., the average of θ and δ similarities) of spatial pattern (red line) and temporal pattern (blue line) between deforestation hotspots in Indonesia in 2020 derived from Landsat-8 OLI (using the classification method described in Appendix A, at clarity $\geq 50\%$ and deforestation intensity $\geq 50\%$) and deforestation hotspots in Indonesia in 2020 derived from MODIS (using the change-detection algorithm developed by [30], at change-detection thresholds ranging from 50 to 130). In this case, the overall similarity was calculated using the new Ruzicka–Fidelity metric (see Equation (1)), and time-series data were interpolated at a time slice of 1 day.

5. Conclusions

In this article, we developed a generic method for measuring the spatial and temporal pattern similarity between deforestation data sets with different spatial resolutions based on similarity-based pattern analysis. The robustness of the method was evaluated through sensibility and sensitivity analyses before its application to the measurement of spatial and temporal pattern similarity between two deforestation data sets derived from two satellite imageries with different spatial resolutions (i.e., MODIS and Landsat-8 OLI). For sensibility and sensitivity analyses, the method was assessed using hypothetical data, either spatial or temporal. The highlighted findings of our study are as follows.

Firstly, we demonstrated that our method could measure, at the level of human visual sensing, two pairs of hypothetical patterns with reasonable results. The first pair was two images of mirrored eagles with the same size and pattern. The second pair was two images of turtles with the same pattern but different sizes. The PDFs of angle from the centroid to each point (θ) for each pair of hypothetical patterns indicated that the pair of eagle patterns had contrasting θ PDFs, while the pair of turtle patterns had exactly similar θ PDFs. On the other hand, the pair of eagle patterns had exactly similar δ PDFs, while the pair of turtle patterns had contrasting δ PDFs. Furthermore, using the nine metrics, the PDF similarities of θ and δ for each pair were measured. In this case, the overall similarity was calculated as the average PDF similarity for θ and δ . The spatial patterns of pairs of eagle shapes had relatively low to fair similarity, in terms of θ , with values ranging from 0.31 to 0.53, and very high similarity in terms of δ , with values close to 1, which was consistent across the nine metrics. Vice versa, the pair of turtle shapes had very high similarity for θ , with values ranging from 0.95 to 1, and relatively low to fair similarity for δ , with values ranging from 0.60 to 0.83, which was also consistent across the nine metrics. Moreover, we also found that our method consistently produced envelopes, with upper and lower edges always being bordered by the Fidelity and Soergel–Ruzicka–Tanimoto metrics, respectively. Thus, we suggest a new single metric that considers both optimistic and pessimistic metrics, which is the average of these two abovementioned metrics, which we denote as Ruzicka–Fidelity.

Secondly, in terms of the method's sensitivity to spatial displacement from the hypothetical origin, we demonstrated that our method was capable of consistently measuring

spatial pattern similarity with reasonable results. Based on sensitivity analyses of the systematic displacement of two hypothetical shapes: (i) a circle with a diameter of about 1.0567 m and (ii) a square with dimensions of $1.0567 \times 1.0567 \text{ m}^2$, the results suggested that the method provides a sensible measure, in term of its proportionality, as it resulted in an overall similarity ranging from 0.4 to 0.5 when displacement from the origin took place at about half the size of the maximum radii of the shapes, and decreased to relatively low similarity (of about 0.1–0.2) when displacement from the origin took place at about the maximum radii of the shapes. The results also suggested that, when displacement from the origin took place much further than the maximum radii of the shapes, the method still appreciated the pattern similarity between the shifted shapes and their origins, with the overall similarity between 0 and 0.1, which is not recognized in conventional accuracy assessment methods where the agreement between pairs of spatial patterns is measured based on their point-to-point degree of matching.

Thirdly, we also demonstrated the method's sensitivity to changes in spatial resolution using several pairs of hypothetical shapes with different resolutions, ranging from $1\times$ to $20\times$ coarser. The results indicated that the similarity measures resulting from the method decreased consistently when the resolutions of the pairs were much coarser than the original image at relatively low decreasing rates, ranging from -0.005 to -0.009 per km resolution decrease from the original. This suggests that, when the method is applied to measure the similarity of pairs of data sets with different resolutions, the result should be calibrated with regard to the resolution difference effect at rates ranging from -0.005 to -0.009 per km resolution difference from the origin.

Fourthly, we also performed sensitivity analyses for hypothetical temporal data sets of deforestation by (i) shifting a hypothetical reference deforestation curve to the left from 16 days to 80 days, representing an acceleration of detection; (ii) shifting the reference to the right from 16 days to 80 days, representing a delay of detection; (iii) shifting the reference upward from 1 pixel to 5 pixels, representing an overestimation of detection; and (iv) shifting the reference downward from 1 pixel to 5 pixels, representing an underestimation of detection. Moreover, linear interpolation on our hypothetical deforestation time-series data was performed to downscale the temporal resolution from a 16-day time interval to a 1-day, 1-h, and 1-min interval, aimed at assessing the sensitivity of our method to displacement in temporal data with regard to the temporal resolution. The results suggested that the most sensible sensitivity of our method to changes in y -axis variables (Δy), due to overestimation or underestimation, as well as changes in x -axis variables (Δx), due to acceleration or delay, was given by temporal data sets at a 1-day temporal resolution.

Finally, we performed similarity analyses for spatial and temporal patterns of deforestation data sets in Indonesia in 2020, derived from MODIS and Landsat-8 OLI. The results suggested that the MODIS-based deforestation monitoring system developed by [30] can still be calibrated, in terms of increasing its change-detection threshold parameter to above 130, to increase its temporal pattern similarity with the reference derived from Landsat-8 OLI, without significantly losing spatial pattern similarity.

Overall, we demonstrated that the proposed method offers an alternative to conventional approaches to accuracy assessment of deforestation patterns in time and space across differences in resolution, which should apply to various deforestation-monitoring data sets, ranging from near-real-time [13,14] to sub-annual and annual [15,16] monitoring systems, and covering local and national scales [4–6] to global scales [2,3,5]. In addition, with regard to the urgent need to take action to mitigate climate change through, among other things, reducing deforestation [32], the method should offer an easily applicable approach to evaluate several available deforestation data sets from various producers to saliently provide the most credible data on deforestation for legitimate guidance on the ground to take immediate action. Meanwhile, beyond the deforestation data sets, another potential application of the method is for validating spatio-temporal patterns of simulation results from land-change models, such as [17–20].

However, the current method should further be improved by, among other things: (i) exploring other geometric features, e.g., the square root of the area of the triangle between three random points from the centroid [27]; and (ii) integrating spatial and temporal point data sets in 3D point cloud xyz-coordinate format, where x is the longitude, y is the latitude, and z is time, to obtain a single spatio-temporal-pattern-similarity measure. We developed a tool for this method using the open-source modeling platform NetLogo 6.2.0 and have provided it at ipb.link/similarity-tool. Thus, the tool should be adopted by others looking to utilize and further improve it and the method.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/geomatics1040000/s1>, Figure S1: Deforestation hotspots in Indonesia in 2020 derived from MODIS, using the change-detection algorithm developed by [1] at a change-detection threshold of 50 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S2: Deforestation hotspots in Indonesia in 2020, derived from MODIS using the change-detection algorithm developed by [1] at a change-detection threshold of 60 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S3: Deforestation hotspots in Indonesia in 2020, derived from MODIS using the change-detection algorithm developed by [1] at a change-detection threshold of 70 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S4: Deforestation hotspots in Indonesia in 2020, derived from MODIS using the change-detection algorithm developed by [1] at a change-detection threshold of 80 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S5: Deforestation hotspots in Indonesia in 2020, derived from MODIS using the change-detection algorithm developed by [1] at a change-detection threshold of 90 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S6: Deforestation hotspots in Indonesia in 2020, derived from MODIS using the change-detection algorithm developed by [1] at change-detection threshold of 100 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S7: Deforestation hotspots in Indonesia in 2020, derived from MODIS using change-detection algorithm developed by [1] at a change-detection threshold of 130 (green points). Red points indicate deforestation hotspots confirmed by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid used to measure the spatial pattern similarity between these two deforestation data sets. Figure S8: Green squares indicate deforestation hotspots in Indonesia in 2020, as detected using the 8-day monitoring system based on MODIS developed by [1] at change-detection thresholds of: (a) 50, (b) 60, (c) 70, (d) 80, (e) 90, (f) 100, and (g) 130. Red squares indicate deforestation hotspots detected by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid of the graph, which was used to measure the temporal pattern similarity between these two deforestation data sets. In this case, to synchronize the time interval of MODIS to Landsat-8 OLI, time aggregation was completed for MODIS results at a 16-day time interval. Figure S9: Interpolated deforestation time-series data shown in Figure S8 with 1-day interval. Green squares indicate deforestation hotspots in Indonesia in 2020, as detected using the 8-day monitoring system based on MODIS developed by [1] at change-detection thresholds of: (a) 50, (b) 60, (c) 70, (d) 80, (e) 90, (f) 100, and (g) 130. Red squares indicate deforestation hotspots detected by Landsat-8 OLI at land clarity $\geq 50\%$ and land-clearing intensity $\geq 50\%$. The blue cross is the centroid of the graph, which was used to measure the temporal pattern similarity between these two deforestation data sets.

Author Contributions: Conceptualization, D.S. and L.P.; methodology, D.S., Y.S. and K.K.; software, D.S.; validation, D.S.; formal analysis, D.S.; investigation, D.S.; resources, D.S., Y.S., K.K. and T.K.; data curation, Y.S., K.K. and T.K.; writing—original draft preparation, D.S.; writing—review and editing, D.S., L.P., Y.S. and A.W.; visualization, D.S.; supervision, L.P., Y.S. and A.W.; project administration, Y.S., P.P. and H.E.; funding acquisition, A.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by World Resources Institute Indonesia, through Agreement No. 50016.01-IX-20.

Data Availability Statement: The main data of the MODIS-based deforestation monitoring system are available at <https://lulcc.ipb.ac.id/>, accessed on 8 April 2021. The similarity analysis tool we developed in this study within an open-source platform is also available at ipb.link/similarity-tool.

Acknowledgments: The authors would like to thank all members of the MDPI Geomatics and MDPI Remote Sensing Editorial Boards, including the Special Issue Editors, the Academic Editor, the Assistant Editors, and four anonymous reviewers, for their thorough and critical reviews, which significantly improved this paper. We also would like to express our gratitude to the World Resources Institute Indonesia (WRI Indonesia) for financial support, as well as to the Indonesian National Research and Innovation Agency (BRIN) and the Indonesian Institute of Aeronautics and Space (LAPAN) for data collection and provision supports. Finally, we are grateful to the Finance and Administration Officer of the Department of Forest Resources Conservation and Ecotourism, Faculty of Forestry, IPB University, Irene Carolina Fatmayanti, for her assistance concerning financial and administrative matters.

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

Appendix A

This part describes the classification method (i.e., the knowledge-based tree-classifier algorithm [33]) for deforestation detection derived from Landsat-8 OLI, which was used in this study. Furthermore, deforestation hotspots were defined as pixels experiencing land clearing based on the classification results.

Table A1. Knowledge-based tree-classifier algorithm used in deforestation detection derived from Landsat-8 OLI.

Step	Category	Classifier	Supporting Findings
1	Thick cloud	Red > 200	Use of red band, in combination with blue, green, and TIRS bands thresholding of Landsat-8 OLI for detecting thick cloud [34]
2	High NDVI	Not(thick cloud) and (NDVI > 0.35)	Landsat-8 NDVI of deciduous and coniferous forests ranging from 0.40 to 0.74 [35]
3	Shade	(Not(thick cloud or high NDVI) and (NIR < 100 and SWIR < 50)) and (NIR < 150 and SWIR < 80 and NIR-SWIR < 90)	Use of NIR and SWIR bands thresholding of Landsat-8 OLI, in combination with projective geometry modeling for detecting cloud shadows [36,37]
4	Water	((Not(thick cloud or high NDVI) and (NIR < 100 and SWIR < 50)) and not(shade))	Use of NIR and SWIR bands, in combination with blue or green bands of Landsat-8 OLI, composed into composite indices for detecting water [38,39]
5	Cleared land	Not(thick cloud or high NDVI or shade or water) and (SWIR-max(NIR, Red)) > -50	Use of NIR and SWIR bands of Landsat-8 OLI, composed into a composite index for detecting bare land [40] Use of NIR and SWIR bands, in combination with green band of Landsat-8 OLI, composed into composite indices for detecting bare land [41]
6	Green zone 1	Not(thick cloud or high NDVI or shade or water or cleared land) and (NIR-max(SWIR, Red) > 50)	—
7	Green zone 2	Not(thick cloud or high NDVI or shade or water or cleared land or green zone 1) and (NIR-Red > 20)	—
8	Haze	Not(thick cloud or high NDVI or shade or water or cleared land or green zone 1 or green zone 2)	Use of NIR and SWIR bands of Landsat-8 OLI for detecting haze [42]

Appendix B

This part describes how our method, based on pattern-similarity analysis, differs from conventional omission/commission disagreement approaches using confusion matrices. Pairs of hypothetical circular shapes with various overlapping possibilities were used for the purpose. A hypothetical circular shape with a radius of about 0.417 m (red) and its displacements (green), in which (a), (b), (c), and (d) intersect with the origin in the cardinal directions, from north to west; (e), (f), (g), and (h), touch the origin around the edge in the cardinal directions, from north to west; and (i), (j), (k), and (l) are not connected to the origin in the cardinal directions, from north to west. These hypothetical shapes were used to demonstrate the difference between conventional omission/commission disagreement analyses—based on the point-to-point degree of matching using a confusion matrix—and our method, which is based on pattern-similarity analysis. Tables A2–A4 summarize the results using those two approaches. Once a pair of spatial patterns is disjoint, the conventional approach immediately ignores the pair’s agreement, regardless of whether the pair is, in fact, still relatively close to each other, while our method still appreciates such a possibility.

Table A2. Comparison between omission/commission disagreement approach and pattern-similarity based approach to measure pairs of a hypothetical circular shape and their displacements in cardinal directions intersecting each other.

Approach	Metric	Pair of Hypothetical Circular Shapes			
					
Omission/commission disagreement using confusion matrix	User’s agreement of true category (%)	39.20	40.08	40.08	39.20
	User’s agreement of false category (%)	0.00	0.00	0.00	0.00
	Producer’s agreement of true category (%)	39.20	40.08	40.08	39.20
	Producer’s agreement of false category (%)	0.00	0.00	0.00	0.00
	Overall agreement (%)	24.38	25.06	25.06	24.38
Pattern-based similarity	Sørensen	0.39	0.40	0.40	0.39
	Soergel	0.25	0.25	0.25	0.25
	Intersection	0.39	0.40	0.40	0.39
	Ruzicka	0.25	0.25	0.25	0.25
	Tanimoto	0.25	0.25	0.25	0.25
	Cosine	0.58	0.58	0.58	0.58
	Jaccard	0.35	0.35	0.35	0.35
	Dice	0.51	0.52	0.52	0.51
	Fidelity	0.63	0.65	0.65	0.63
	Ruzicka–Fidelity	0.44	0.45	0.45	0.44

Table A3. Comparison between omission/commission disagreement approach and pattern-similarity-based approach to measure pairs of a hypothetical circular shape and their displacements in cardinal directions that touch at the perimeter.

Approach	Metric	Pair of Hypothetical Circular Shapes			
					
Omission/commission disagreement based on confusion matrix	User's agreement of true category (%)	0.00	0.14	0.14	0.00
	User's agreement of false category (%)	0.00	0.00	0.00	0.00
	Producer's agreement of true category (%)	0.00	0.14	0.14	0.00
	Producer's agreement of false category (%)	0.00	0.00	0.00	0.00
	Overall agreement (%)	0.00	0.07	0.07	0.00
Pattern-based similarity	Sørensen	0.08	0.09	0.09	0.09
	Soergel	0.05	0.05	0.05	0.05
	Intersection	0.08	0.09	0.09	0.09
	Ruzicka	0.05	0.05	0.05	0.05
	Tanimoto	0.05	0.05	0.05	0.05
	Cosine	0.20	0.20	0.20	0.20
	Jaccard	0.08	0.08	0.08	0.08
	Dice	0.13	0.13	0.13	0.13
	Fidelity	0.21	0.21	0.21	0.21
Ruzicka–Fidelity	0.13	0.13	0.13	0.13	

Table A4. Comparison between omission/commission disagreement approach and pattern-similarity based approach to measure pairs of a hypothetical circular shape and their displacements in cardinal directions that are disconnected.

Approach	Metric	Pair of Hypothetical Circular Shapes			
					
Omission/commission disagreement based on confusion matrix	User's agreement of true category (%)	0.00	0.00	0.00	0.00
	User's agreement of false category (%)	0.00	0.00	0.00	0.00
	Producer's agreement of true category (%)	0.00	0.00	0.00	0.00
	Producer's agreement of false category (%)	0.00	0.00	0.00	0.00
	Overall agreement (%)	0.00	0.00	0.00	0.00
Pattern-based similarity	Sørensen	0.07	0.07	0.07	0.07
	Soergel	0.04	0.04	0.04	0.04
	Intersection	0.07	0.07	0.07	0.07
	Ruzicka	0.04	0.04	0.04	0.04
	Tanimoto	0.04	0.04	0.04	0.04
	Cosine	0.17	0.17	0.17	0.17
	Jaccard	0.06	0.06	0.06	0.06
	Dice	0.11	0.11	0.11	0.11
	Fidelity	0.18	0.18	0.18	0.18
Ruzicka–Fidelity	0.11	0.11	0.11	0.11	

References

1. Pontius, R.G., Jr.; Gao, Y.; Giner, N.M.; Kohyama, T.; Osaki, M.; Hirose, K. Design and Interpretation of Intensity Analysis Illustrated by Land Change in Central Kalimantan, Indonesia. *Land* **2013**, *2*, 351–369. [[CrossRef](#)]
2. Feng, M.; Sexton, J.O.; Huang, C.; Anand, A.; Channan, S.; Song, X.-P.; Song, D.-X.; Kim, D.-H.; Noojipady, P.; Townshend, J.R. Earth science data records of global forest cover and change: Assessment of accuracy in 1990, 2000, and 2005 epochs. *Remote Sens. Environ.* **2016**, *184*, 73–85. [[CrossRef](#)]
3. Bos, A.B.; De Sy, V.; Duchelle, A.E.; Herold, M.; Martius, C.; Tsendbazar, N.-E. Global data and tools for local forest cover loss and REDD+ performance assessment: Accuracy, uncertainty, complementarity and impact. *Int. J. Appl. Earth Obs. Geoinform.* **2019**, *80*, 295–311. [[CrossRef](#)]
4. Galiatsatos, N.; Donoghue, D.N.; Watt, P.; Bholanath, P.; Pickering, J.; Hansen, M.C.; Mahmood, A.R. An Assessment of Global Forest Change Datasets for National Forest Monitoring and Reporting. *Remote Sens.* **2020**, *12*, 1790. [[CrossRef](#)]
5. Shimizu, K.; Ota, T.; Mizoue, N. Accuracy Assessments of Local and Global Forest Change Data to Estimate Annual Disturbances in Temperate Forests. *Remote Sens.* **2020**, *12*, 2438. [[CrossRef](#)]
6. Suyamto, D.; Condro, A.; Prasetyo, L.; Setiawan, Y.; Wijayanto, A. Assessing the Agreement between Deforestation Maps of Kalimantan from Various Sources. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *556*, 012011. [[CrossRef](#)]
7. De Vries, B.J.M. Scenarios: Guidance for uncertain and complex world? In *Sustainability or Collapse? An Integrated History and Future of People on Earth*; Costanza, R., Graumlich, L.J., Steffen, W., Eds.; The MIT Press: Cambridge, MA, USA; London, UK, 2007; pp. 379–397.
8. Costanza, R. Model goodness of fit: A multiple resolution procedure. *Ecol. Model.* **1989**, *47*, 199–215. [[CrossRef](#)]
9. Pelillo, M. Introduction: The SIMBAD project. In *Similarity-Based Pattern Analysis and Recognition*; Pelillo, M., Ed.; Springer: London, UK; Heidelberg, Germany; New York, NY, USA; Dordrecht, The Netherlands, 2013; pp. 1–9.
10. Lusiana, B.; van Noordwijk, M.; Johana, F.; Galudra, G.; Suyanto, S.; Cadisch, G. Implications of uncertainty and scale in carbon emission estimates on locally appropriate designs to reduce emissions from deforestation and degradation (REDD+). *Mitig. Adapt. Strat. Glob. Chang.* **2013**, *19*, 757–772. [[CrossRef](#)]
11. Pitkänen, T.P.; Sirro, L.; Häme, L.; Häme, T.; Törmä, M.; Kangas, A. Errors related to the automatized satellite-based change detection of boreal forests in Finland. *Int. J. Appl. Earth Obs. Geoinform.* **2019**, *86*, 102011. [[CrossRef](#)]
12. Wang, W.; Hall-Beyer, M.; Wu, C.; Fang, W.; Nsengiyumva, W. Uncertainty Problems in Image Change Detection. *Sustainability* **2019**, *12*, 274. [[CrossRef](#)]
13. Setiawan, Y.; Kustiyo; Darmawan, A. Development of Near-real Time Forest Monitoring (Phase I: Data Preparation). *Procedia Environ. Sci.* **2015**, *24*, 317–323. [[CrossRef](#)]
14. Ye, S.; Rogan, J.; Zhu, Z.; Eastman, J.R. A near-real-time approach for monitoring forest disturbance using Landsat time series: Stochastic continuous change detection. *Remote Sens. Environ.* **2020**, *252*, 112167. [[CrossRef](#)]
15. Sedano, F.; Lisboa, S.; Duncanson, L.; Ribeiro, N.; Sioe, A.; Sahajpal, R.; Hurtt, G.; Tucker, C. Monitoring intra and inter annual dynamics of forest degradation from charcoal production in Southern Africa with Sentinel—2 imagery. *Int. J. Appl. Earth Obs. Geoinform.* **2020**, *92*, 102184. [[CrossRef](#)]
16. Chen, N.; Tsendbazar, N.-E.; Hamunyela, E.; Verbesselt, J.; Herold, M. Sub-annual tropical forest disturbance monitoring using harmonized Landsat and Sentinel-2 data. *Int. J. Appl. Earth Obs. Geoinform.* **2021**, *102*, 102386. [[CrossRef](#)]
17. Van Noordwijk, M.; Suyamto, D.A.; Lusiana, B.; Ekadinata, A.; Hairiah, K. Facilitating agroforestation of landscapes for sustainable benefits: Tradeoffs between carbon stocks and local development benefits in Indonesia according to the FALLOW model. *Agric. Ecosyst. Environ.* **2008**, *126*, 98–112. [[CrossRef](#)]
18. Lusiana, B.; van Noordwijk, M.; Suyamto, D.; Mulia, R.; Joshi, L.; Cadisch, G. Users' perspectives on validity of a simulation model for natural resource management. *Int. J. Agric. Sustain.* **2011**, *9*, 364–378. [[CrossRef](#)]
19. Suwarno, A.; Van Noordwijk, M.; Weikard, H.-P.; Suyamto, D. Indonesia's forest conversion moratorium assessed with an agent-based model of Land-Use Change and Ecosystem Services (LUCES). *Mitig. Adapt. Strat. Glob. Chang.* **2016**, *23*, 211–229. [[CrossRef](#)]
20. Malek, Ž.; Verburg, P.H. Mapping global patterns of land use decision-making. *Glob. Environ. Chang.* **2020**, *65*, 102170. [[CrossRef](#)]
21. O'Rourke, J.; Toussaint, G.T. Pattern recognition. In *Handbook of Discrete and Computational Geometry*, 3rd ed.; Goodman, J.E., O'Rourke, J., Tóth, C.D., Eds.; Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2018; pp. 1421–1450, ISBN 9781498711395.
22. Schneider, R. Discrete aspects of stochastic geometry. In *Handbook of Discrete and Computational Geometry*, 3rd ed.; Goodman, J.E., O'Rourke, J., Tóth, C.D., Eds.; Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2018; pp. 299–329, ISBN 9781498711395.
23. Agarwal, P.K. Range searching. In *Handbook of Discrete and Computational Geometry*, 3rd ed.; Goodman, J.E., O'Rourke, J., Tóth, C.D., Eds.; Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2018; pp. 1057–1092, ISBN 9781498711395.
24. Chazal, F. High-dimensional topological data analysis. In *Handbook of Discrete and Computational Geometry*, 3rd ed.; Goodman, J.E., O'Rourke, J., Tóth, C.D., Eds.; Taylor & Francis Group, LLC: Boca Raton, FL, USA, 2018; pp. 663–683, ISBN 9781498711395.
25. Duin, R.P.W.; Pełkalska, E.; Loog, M. Non-Euclidean dissimilarities: Causes, embedding and informativeness. In *Similarity-Based Pattern Analysis and Recognition*; Pelillo, M., Ed.; Springer: London, UK; Heidelberg, Germany; New York, NY, USA; Dordrecht, The Netherlands, 2013; pp. 13–44.
26. Lampinen, J.; Laaksonen, J.; Oja, E. Pattern recognition. In *Image Processing and Pattern Recognition*; Leondes, C.T., Ed.; Academic Press: San Diego, CA, USA; London, UK, 1998; pp. 1–59.

27. Osada, R.; Funkhouser, T.; Chazelle, B.; Dobkin, D. Shape distributions. *ACM Trans. Graph.* **2002**, *21*, 807–832. [[CrossRef](#)]
28. Nakamura, T.; Taki, K.; Nomiya, H.; Seki, K.; Uehara, K. A shape-based similarity measure for time series data with ensemble learning. *Pattern Anal. Appl.* **2012**, *16*, 535–548. [[CrossRef](#)]
29. Cha, S.-H. Comprehensive survey on distance/similarity measures between probability density functions. *Int. J. Math. Models Methods Appl. Sci.* **2007**, *4*, 300–3007.
30. Setiawan, Y.; Kustiyo, K.; Darmawan, A. A simple method for developing near real-time nationwide forest monitoring for Indonesia using MODIS near- and shortwave infrared bands. *Remote Sens. Lett.* **2016**, *7*, 318–327. [[CrossRef](#)]
31. Wilensky, U.; Rand, W. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*; The MIT Press: Cambridge, MA, USA, 2015; ISBN 978-0-262-73189-8.
32. Van Kooten, G.C.; Withey, P.; Johnston, C.M. Climate urgency and the timing of carbon fluxes. *Biomass Bioenergy* **2021**, *151*, 106162. [[CrossRef](#)]
33. Verma, N.; Mishra, P.; Purohit, N. Development of a knowledge based decision tree classifier using hybrid polarimetric SAR observables. *Int. J. Remote Sens.* **2019**, *41*, 1302–1320. [[CrossRef](#)]
34. Candra, D.S.; Phinn, S.; Scarth, P. Automated Cloud and Cloud-Shadow Masking for Landsat 8 Using Multitemporal Images in a Variety of Environments. *Remote Sens.* **2019**, *11*, 2060. [[CrossRef](#)]
35. Ke, Y.; Im, J.; Lee, J.; Gong, H.; Ryu, Y. Characteristics of Landsat 8 OLI-derived NDVI by comparison with multiple satellite sensors and in-situ observations. *Remote Sens. Environ.* **2015**, *164*, 298–313. [[CrossRef](#)]
36. Suyanto, D.; Prasetyo, L.; Setiawan, Y.; Wijayanto, A.K. Combining Projective Geometry Modelling and Spectral Thresholding for Automated Cloud Shadow Masking in Landsat 8 Imageries. In Proceedings of the 2017 European Modelling Symposium (EMS), Manchester, UK, 20–21 November 2017; pp. 22–27. [[CrossRef](#)]
37. Sun, L.; Liu, X.; Yang, Y.; Chen, T.; Wang, Q.; Zhou, X. A cloud shadow detection method combined with cloud height iteration and spectral analysis for Landsat 8 OLI data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 193–207. [[CrossRef](#)]
38. Nguyen, U.N.T.; Pham, L.T.H.; Dang, T.D. An automatic water detection approach using Landsat 8 OLI and Google Earth Engine cloud computing to map lakes and reservoirs in New Zealand. *Environ. Monit. Assess.* **2019**, *191*, 235. [[CrossRef](#)] [[PubMed](#)]
39. Acharya, T.D.; Lee, D.H.; Yang, I.T.; Lee, J.K. Identification of Water Bodies in a Landsat 8 OLI Image Using a J48 Decision Tree. *Sensors* **2016**, *16*, 1075. [[CrossRef](#)]
40. Zhou, Y.; Yang, G.; Wang, S.; Wang, L.; Wang, F.; Liu, X. A new index for mapping built-up and bare land areas from Landsat-8 OLI data. *Remote Sens. Lett.* **2014**, *5*, 862–871. [[CrossRef](#)]
41. Rasul, A.; Balzter, H.; Ibrahim, G.R.F.; Hameed, H.M.; Wheeler, J.; Adamu, B.; Ibrahim, S.; Najmaddin, P.M. Applying Built-Up and Bare-Soil Indices from Landsat 8 to Cities in Dry Climates. *Land* **2018**, *7*, 81. [[CrossRef](#)]
42. Kustiyo; Hayati, A.K. Haze Removal in the Visible Bands of Landsat 8 Oli over Shallow Water Area. *Int. J. Remote Sens. Earth Sci.* **2017**, *13*, 151–157. [[CrossRef](#)]