

Article

Validation of Land Cover Products Using Reliability Evaluation Methods

Wenzhong Shi ¹, Xiaokang Zhang ^{2,*}, Ming Hao ³, Pan Shao ², Liping Cai ³ and Xuzhe Lyu ²

¹ Joint Research Laboratory on Spatial Information, The Hong Kong Polytechnic University and Wuhan University, Hong Kong and Wuhan 430079, China; E-Mail: lswzshi@polyu.edu.hk

² School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; E-Mails: shaopan19861223@126.com (P.S.); lxz112@126.com (X.L.)

³ School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou 221116, China; E-Mails: haomingcumt@gmail.com (M.H.); cumtcailp@126.com (L.C.)

* Author to whom correspondence should be addressed; E-Mail: natezhangxk@gmail.com; Tel.: +86-27-6877-8546.

Academic Editors: Ioannis Gitas and Prasad S. Thenkabil

Received: 19 March 2015 / Accepted: 22 May 2015 / Published: 16 June 2015

Abstract: Validation of land cover products is a fundamental task prior to data applications. Current validation schemes and methods are, however, suited only for assessing classification accuracy and disregard the reliability of land cover products. The reliability evaluation of land cover products should be undertaken to provide reliable land cover information. In addition, the lack of high-quality reference data often constrains validation and affects the reliability results of land cover products. This study proposes a validation schema to evaluate the reliability of land cover products, including two methods, namely, result reliability evaluation and process reliability evaluation. Result reliability evaluation computes the reliability of land cover products using seven reliability indicators. Process reliability evaluation analyzes the reliability propagation in the data production process to obtain the reliability of land cover products. Fuzzy fault tree analysis is introduced and improved in the reliability analysis of a data production process. Research results show that the proposed reliability evaluation scheme is reasonable and can be applied to validate land cover products. Through the analysis of the seven indicators of result reliability evaluation, more information on land cover can be obtained for strategic decision-making and planning, compared with traditional accuracy assessment methods. Process reliability

evaluation without the need for reference data can facilitate the validation and reflect the change trends of reliabilities to some extent.

Keywords: land cover; validation; reliability evaluation; fuzzy fault tree

1. Introduction

Land cover is one of the most critical variables of the earth system and affects various parts of the human and physical environments. It has a major role in the exchanges of energy, heat, and momentum between continents and the atmosphere [1]. Therefore, information on the state of land cover and land cover dynamics is much needed on a regular basis to support major scientific and policy applications. Land cover products that focus on characterizing different vegetation types have been developed via remotely sensed datasets. To provide reliable and accurate land cover information, validating land cover data should be a fundamental task prior to data application.

Efforts to improve the validation methods and schemes of land cover products were realized by focusing on sampling design, response design, and analysis [2]. Sampling designs differ in their suitability to achieve different objectives in the validation of land cover [3]. Mayaux, *et al.* [4] used a two-stage stratified clustered sampling to validate the Global Land Cover 2000 Map. Stehman, *et al.* [5] constructed a stratified sampling design incorporating class-level stratification to estimate the accuracy of land cover maps. Considerable research on response design focused on reducing the effect of error in ground reference data to enable more accurate estimation in the validation of land cover products. A virtual field reference database was designed to support detailed assessments of land cover products by providing a robust database that characterizes representative cover types [6]. Foody *et al.* [7] used two sources of volunteered data to illustrate the potential of amateur activity in validating the forest cover representation provided by the GlobCover map of the European Space Agency. Foody *et al.* also studied the impact of ground reference data error on the accuracy of estimates of the extent of change and on the accuracy of change detection in [8]. Wulder *et al.* [9] used the collection of airborne videos used for validating sustainable development of forests' land cover maps. Estimation and analysis are commonly based on a confusion matrix that summarizes the key information obtained from sampling and response designs. The confusion matrix is mainly used to provide a basic description of land cover product accuracy and to compare accuracies [10]. Indices such as overall, users' and producers' accuracies derived from the confusion matrix are used to assess the accuracy in validating land cover products [4,11,12].

However, these validations only provide information on the classification accuracy of land cover data as a whole; they cannot be used to characterize the reliability of the data, substantially limiting the value of data in use. In the field of system engineering, reliability is defined as the probability that an item can perform its specified function for a certain interval under given conditions [13]. Reliability of land cover products refers to the function that the data can correctly, effectively, and completely reflect the actual land cover conditions under certain spatio-temporal conditions. Reliability reflects the characteristics more than the classification accuracy of land cover data, and thus it has numerous impacts in applying the data and affects strategic decision and planning. Reliability evaluation is widely applied in various

fields, such as power system, network, software, and pipeline [14–17]. In the field of geographical information science and remote sensing, Zhang *et al.* [18] proposed a fusion approach of several typical change-detection algorithms based on the reliability of individual change-detection algorithms to generate reliable land cover change information. Based on the research of uncertainties in spatial analysis [19], Shi *et al.* [20] introduced reliability theory into spatial analysis to obtain more accurate and reliable spatial analysis results. Considerable attention should be directed toward the study of reliability evaluation of land cover products.

Validation is also often constrained by the lack of high-quality ground reference data [7]. Furthermore, the reference data are expensive and logistically challenging to collect for large area land cover products [21]. Reliability analysis on the data production process to obtain the reliability of data products can solve this problem to some extent because this analysis does not require the reference data, and the reliability propagation in the data production process determines the reliability of land cover products.

Thus, this paper proposes a new validation schema to evaluate the reliability of land cover data. The schema includes two validation methods, namely, result reliability evaluation and process reliability evaluation. Result reliability evaluation addresses a few reliability indicators to evaluate the reliability of land cover products. Process reliability evaluation obtains the reliability of land cover data through the reliability analysis on the process of data production without the need for reference data. Section 2 describes the proposed validation schema. Section 3 tests the validity of the proposed schema by providing applications in six typical study areas and gives the validation results. Discussions and conclusions are presented in the last section.

2. Methodology

The general framework of our validation schema is shown in Figure 1. The validation schema includes two methods: result reliability evaluation and process reliability evaluation. Sampling design should be conducted when evaluating the result reliabilities of a group of land cover maps or a certain land cover type in the group. The reliability of the sample unit (one map or certain land cover type in the map) can then be obtained using seven reliability indicators and a weighted computation method. Process reliability evaluation is based on the process reliability analysis model to evaluate the reliability of one map. Given that the analysis does not need reference data, evaluating all maps in the group is relatively easy without the need for sampling design.

2.1. Result Reliability Evaluation

Result reliability evaluation uses reliability indicators to obtain the reliability of land cover data. Shi *et al.* [20] presented reliability indicators such as correctness, integrity, consistency, robustness, and applicability in reliable spatial analysis. Errors affecting the reliability of land cover data can be caused by confusion between land cover types (wrong label or missing types) or can be spatial errors (wrong position of the boundary between types or disappearance of small patches). Other factors such as acquisition scale, human factors, and acquisition time can also affect data reliability. Based on these factors and the characteristics of land cover data, we propose seven quantized indicators (Figure 1) and a weighted computation method to evaluate the reliability of land cover data. In the hierarchical

classification system, the multilevel reliabilities of one map or a certain land cover type are supported. These reliability indicators are defined as follows:

(1) *Classification correctness* (C) refers to the probability of incorrectly classified area in the land cover data. When computing multilevel reliabilities of the data, error level should be judged according to the feature type and actual land cover type to compute the incorrect area. Let A_{c_error} be the incorrectly classified area of the data, and A_{total} be the total area of the data. The following formula can thus be obtained

$$C = A_{c_error} / A_{total} \quad (1)$$

(2) *Scale reasonableness* (S) represents the scale effect to the reliability of the data. Let $A_{current_scale}$ be the area of the data at current acquisition scale, and A_{actual} be the actual area of the data. Thus, scale reasonableness can be computed by

$$S = A_{current_scale} / A_{actual} \quad (2)$$

(3) *Integrity* (I) reflects the effect of missing or extra features and feature types on the reliability of the land cover data. I can be defined as

$$I = w_a(A_{a_error} / A_{a_total}) + w_c(1 - N_{t_error} / N_{t_total}) \quad (3)$$

where $w_a, w_c > 0$ are the weighting parameters, N_{t_total} is the total number of feature types, N_{t_error} is the number of missing or extra feature types, A_{a_total} is the total area of the data, and A_{a_error} is the area of missing or extra features.

(4) *Robustness* (B) refers to the stability of the land cover data to maintain its reliability when evaluated by different evaluation professionals, who can be affected by several factors, such as evaluating experience, evaluating condition, spirit state, and psychological quality. Based on the variance value δ of the evaluation results obtained by different experts, B is expressed as

$$B = 1 - 0.4 \times \delta / \delta_0 \quad (4)$$

where δ_0 is a constant normally set to 10. If B is less than zero, then we assign zero to B .

(5) *Consistency* (K) refers to the level of congruence between land cover data and actual land cover condition. This paper adopts Cohen's kappa coefficient for the consistency measurement. Cohen's kappa coefficient is often applied as an index of classification accuracy to evaluate the consistency between classification result and imagery data [22]. Derived from a confusion matrix, K is expressed as:

$$K = \frac{N \times \sum_{i=1}^r N_{ii} - \sum_{i=1}^r (N_{i+} N_{+i})}{N^2 - \sum_{i=1}^r (N_{i+} N_{+i})} \quad (5)$$

where r denotes the number of land cover types, N is the total number of features, N_{ii} is the main diagonal of the matrix, N_{i+} is the sum of the i th row of the matrix, and N_{+i} is the sum of the i th column of the matrix.

(6) *Currency* (T) measures the extent that the land cover data acquired before can reflect the actual land cover condition at the current time. Given the actual land cover condition changes over time, the reliability of the data will decrease. Let r_{change} be the change ratio of the land cover condition from the data acquisition time to the data evaluation time. Large r_{change} value represents a low reliability of the data. Thus, T is defined as

$$T = 1 - r_{change} \tag{6}$$

(7) *Position Precision (P)* presents the position offsets of the features in the land cover data according to the actual land cover condition. *P* can be computed according to the features with geometry displacement error and overedge error, *i.e.*,

$$P = w_g(N_{g_error}/N_{total}) + w_o(N_{o_error}/N_{total}) \tag{7}$$

where $w_g, w_o > 0$ are the weighting parameters, N_{g_error} and N_{o_error} represent the number of features with geometry displacement error and overedge error respectively, and N_{total} denotes the total number of features of the data.

Eventually, the reliability of the sample unit can be computed by

$$R = \sum w_i R_i \tag{8}$$

where *R* is the reliability of the sample unit, R_i is the *i*th aforementioned reliability indicator, and $w_i > 0$ and $\sum w_i = 1$ are the weights of the indicator determined by the data application and the professional knowledge.

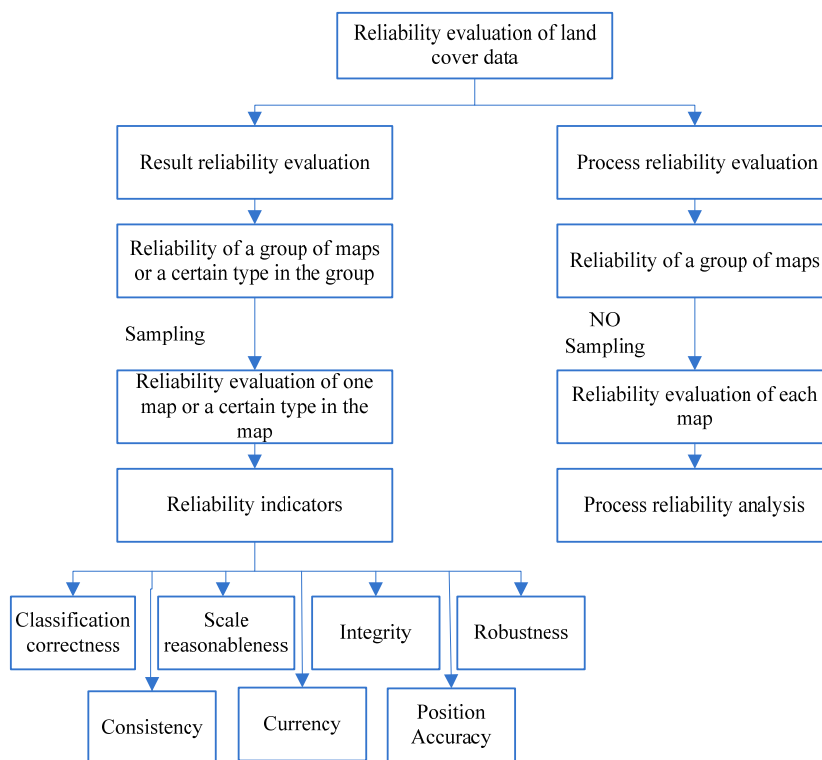


Figure 1. General framework of the reliability evaluation of land cover data.

2.2. Process Reliability Evaluation

Process reliability evaluation obtains the reliability of land cover products through the analysis on reliability propagation in the data production process without the need for reference data. High uncertainty and variability associated with the process of data production pose major challenges in the analysis on reliability propagation. In this paper, we introduced the fuzzy fault tree analysis (FFTA) to analyze the reliability propagation in the production process of land cover data.

Fault tree analysis (FTA) is a logical and diagrammatic approach to represent the sequences and combinations of possible events occurring in a system that lead to the top undesired event [23]. This model is widely used to evaluate the reliability of a complex system in various fields, such as pipelines, aerospace and petrochemical industry [24]. FFTA combines fuzzy set theories for the uncertainty analysis in FTA model [25]. The probabilities of events are treated as a fuzzy number, which can be obtained by theory of fuzzy set [26]. In the process reliability analysis model of land cover data, the FFTA was used to describe the relationships among various events in the data production system and the probability of the top event was regarded as a function of the reliability probability of the system.

2.2.1. Construction of Process Reliability Analysis Model

Two methods to produce land cover products are commonly used [27]. One involves using visual interpretation based on professional knowledge with the foundation geographical data and interpretation symbols as the auxiliary information. The reliability of visual interpretation is affected by operation staff and data sources, including digital orthophoto images and foundation geographical data. Another method of computer classification was applied to certain land cover types that have homogeneous color and texture, such as water bodies, built-up land, and sandy land. The reliabilities of classification algorithms and digital orthophoto images affect the reliability of the computer classification. Furthermore, the reliability of the digital orthophoto image is affected by the image source (*i.e.*, spatial resolution, spectral type and currency) and the image pre-processing (*i.e.*, plane precision and overedge precision).

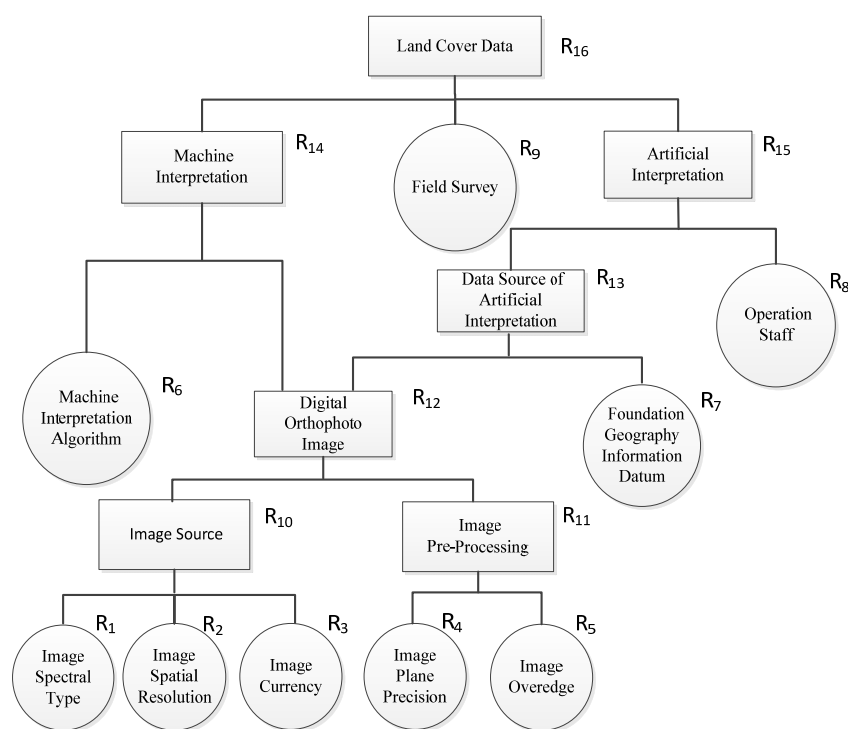


Figure 2. The process reliability analysis model.

Field survey is the auxiliary process of data production and will be performed to identify the land cover type in the region where the type cannot be determined by visual interpretation and computer classification. Therefore, land cover data are derived from machine interpretation, visual interpretation, and field survey, and its reliability depends on the reliabilities of the three processes.

Based on the production process of land cover data and the FFTA, the process reliability analysis model was constructed to evaluate the reliability of land cover products (Figure 2). In the figure, circular frames denote the basic events and square frames denote the intermediate and top events. The evaluation model involves nine basic events, five intermediate events, and one top event. The reliability of land cover products was defined as the top event in this model. The reliabilities of basic events can be obtained by professional knowledge or quantitative methods. Through the reliabilities of basic events and the relationships among events, the reliability of intermediate events and the top event can be computed.

2.2.2. Reliabilities of Basic Events

To evaluate the reliability of the top event in the FFTA model, probabilities of the basic events, as the parameters of the model, must be known in advance [24]. Normally, probabilities of the basic events can be obtained according to expert knowledge and experience [28]. Based on the research on the data production process in the production department, the following preliminary quantitative methods are presented to compute these probabilities:

(1) *Image Spectral Type* (R_1) includes panchromatic and multispectral. Multispectral image contains more adequate spectral information than panchromatic image. Thus, the reliability of multispectral image is more reliable than that of panchromatic image in theory. In this paper, *Image Spectral Type* is set to 0.7 and 0.9 when the type is panchromatic and multispectral, respectively.

(2) *Image Spatial Resolution* (R_2). A high resolution image will theoretically provide more adequate and reliable land cover information. Setting the reliability of the image with spatial resolution of 2 m to 0.7 m and considering the image with spatial resolution more than 10 m as unreliable, the possibility of *Image Spatial Resolution* can be computed as

$$R_2 = \begin{cases} 0.7 + \frac{2-r}{2} \times 0.3 & r < 2 \\ 0.7 \times \frac{10-r}{8} & 2 \leq r < 10 \\ 0 & r \geq 10 \end{cases} \quad (9)$$

where r is the resolution of the image.

(3) *Image Currency* (R_3) measures the extent that the image acquired before can reflect the current land cover condition considering the land cover changes. Setting the reliability of the image acquired at the earliest data-acquired time demanded by the design of the data production to 0.6 and assuming that the reliability of the image will decrease linearly as time passes by, the possibility of *Image Currency* can be computed as

$$R_3 = \begin{cases} 0.6 + \frac{t}{t_0} \times 0.4 & 0 < t < t_0 \\ 0 & t < 0 \end{cases} \quad (10)$$

where t is the month span between the actual data-acquired time and the earliest data-acquired time demanded by the design of the data production, and t_0 is the month span between the actual

data-acquired time and the data-evaluated time. If the image is acquired before the earliest data-acquired time demanded by the design of the data production, then t is less than 0 and R_3 is set to 0.

(4) *Image Plane Precision* (R_4) influences the reliability of the image pre-processing and is defined as

$$R_4 = \begin{cases} 0.6 + \frac{0.4}{0.7 \times m_0} \times (m_0 - m); & m > 0.3m_0 \\ 1; & m \leq 0.3m_0 \end{cases} \quad (11)$$

where m_0 is a constant denoting the limit value of the mean square error of the image plane precision, and m is the mean square error of the image plane precision.

(5) *Image Overedge* (R_5) influences the reliability of the image pre-processing and is defined as

$$R_5 = \begin{cases} 0.6 + \frac{0.4}{0.7 \times l_0} \times (l_0 - l); & l > 0.3l_0 \\ 1; & l \leq 0.3l_0 \end{cases} \quad (12)$$

where l_0 is a constant denoting the limit value of the mean square error of the image overedge, and l is the mean square error of the image overedge.

(6) *Machine Interpretation Algorithm* (R_6). Different classification algorithms will yield different classification results. Let p_j be the maximum posteriori possibility of the j th pixel or image object, and n be the total number of pixels or image objects. Thus, R_6 can be denoted as

$$R_6 = \frac{1}{n} \sum_{i=1}^n p_i \quad (13)$$

(7) *Foundation Geography Information Datum* (R_7) as the auxiliary datum in the artificial interpretation is one of the data sources of artificial interpretation. R_7 can refer to the previous accuracy assessment result of the data.

(8) *Operation Staff* (R_8) reflects the ability of the data production operator in artificial interpretation. R_8 can refer to the previous average accuracy of the land cover data produced by the operator.

(9) *Field Survey* (R_9). The reliability of land cover data will increase after field survey. Thus, R_9 can be set to a high possibility value according to expert experience.

2.2.3. Reliability Interval of Top Event

Given all the probabilities of the basic events, the reliability of the top event can be obtained through the relationships among various events which are normally represented by logical AND and OR gates [29]. In the process reliability analysis model, we improved the reliability computation methods of intermediate events and top event according to the relationships among the events in the production process of land cover data. In addition, we use the possibility intervals instead of the exact possibility values of intermediate events and top event to obtain more flexible and reliable results.

Image source (R_{10}) is affected by image spectral type (R_1), image spatial resolution (R_2), and image currency (R_3). Thus, the left value of the possibility interval can be set to the minimum value of the three event possibilities, and the right value of the interval can be set to the maximum value as follows:

$$R_{10} = [\min(R_1, R_2, R_3), \max(R_1, R_2, R_3)] \quad (14)$$

where the operators $\min ()$ and $\max ()$ find the minimum and maximum values of the input values, respectively.

Given that the image pre-processing (R_{11}) accuracy depends on image plane precision (R_4) and image overedge (R_5), and plane position errors and overedge errors can happen at the same places, the reliability possibility interval can be computed as

$$R_{11} = [R_4 \cdot R_5, \min(R_4, R_5)] \quad (15)$$

Digital orthophoto image (R_{12}) used to produce land cover data depends on image source (R_{10}) and image pre-processing (R_{11}), two events that are independent of each other. Thus, the reliability possibility interval can be obtained by

$$R_{12} = [\text{left}(R_{10}) \cdot \text{left}(R_{11}), \text{right}(R_{10}) \cdot \text{right}(R_{11})] \quad (16)$$

where the operators left () and right () find the left value and right value of the input possibility interval, respectively.

Data source of artificial interpretation (R_{13}) containing foundation geography information datum (R_7) and digital orthophoto image (R_{12}) can be computed by the average reliability possibility of the two data

$$R_{13} = \left[\frac{\text{left}(R_{12}) + R_7}{2}, \frac{\text{right}(R_{12}) + R_7}{2} \right] \quad (17)$$

Machine interpretation (R_{14}) is affected by the reliability of the image and the classification algorithm applied to the images. Machine interpretation algorithm (R_6) and digital orthophoto image (R_{12}) are also independent of each other. Thus, machine interpretation (R_{14}) can be computed as

$$R_{14} = [R_6 \cdot \text{left}(R_{12}), R_6 \cdot \text{right}(R_{12})] \quad (18)$$

Similarly, artificial interpretation (R_{15}) is affected by operation staff (R_8) and data source of artificial interpretation (R_{13}), which are independent of each other. We can thus obtain the artificial interpretation (R_{15}) by

$$R_{15} = [R_8 \cdot \text{left}(R_{13}), R_8 \cdot \text{right}(R_{13})] \quad (19)$$

Considering that land cover products are derived from machine interpretation, visual interpretation, and field survey, the reliability interval of land cover products can be calculated as

$$R_{16} = [w_1 \cdot R_9 + w_2 \cdot \text{left}(R_{14}) + w_3 \cdot \text{left}(R_{15}), w_1 \cdot R_9 + w_2 \cdot \text{right}(R_{14}) + w_3 \cdot \text{right}(R_{15})] \quad (20)$$

where $w_1, w_2, w_3 > 0$ and $\sum w_i = 1$ are the weighting parameters. These parameters are the proportions of land cover data acquired through field survey, machine interpretation, and artificial interpretation.

3. Case Study

3.1. Area and Data

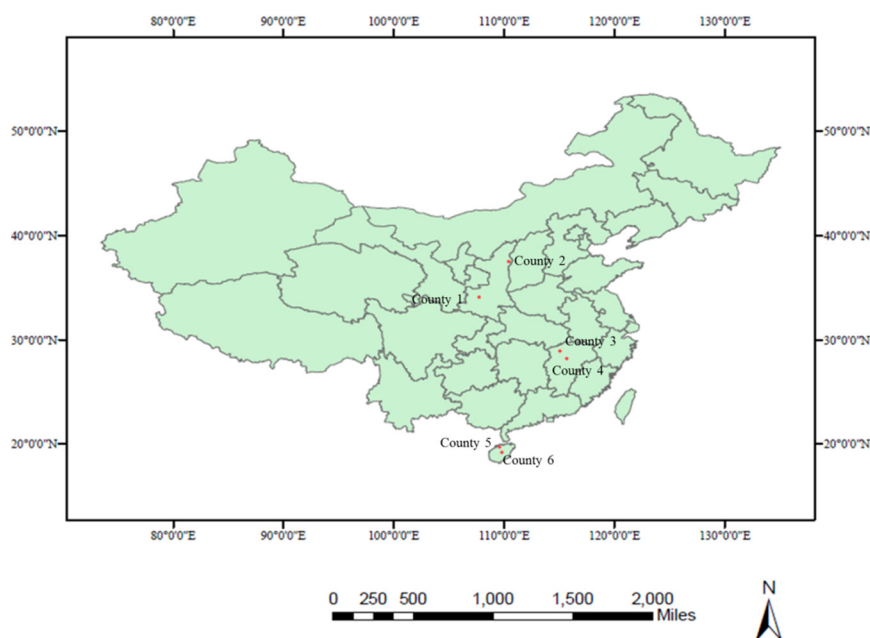
In this paper, we quantitatively analyze the applications of the proposed validation schema using land cover data generated from the project of the National Geographical State Monitoring (NGSM). The project launched by the State Council of China was initiated in 2011. It completely mapped the national land cover in China through remote sensing and geographic information system techniques.

The land cover data of NGSM were generated from multisource remote sensing images, such as aerial images and images acquired by ZY-3, WorldView-2, and QuickBird. NGSM applied a hierarchical classification system with 10 first-level types, 46 second-level types, and 77 third-level types. The hierarchical nature of the classification scheme enables generalization and reporting at high levels of the hierarchy.

We choose land cover data acquired in six typical counties in July 2014 to apply the proposed validation schema, named County 1 and County 2 of Shaanxi Province, County 3 and County 4 of Jiangxi Province, and County 3 and County 4 of Hainan Province. The three provinces of Shaanxi, Jiangxi, and Hainan with different landforms and climates are located in the north, middle, and south of China, respectively. Table 1 displays the detail information of the six study areas. The locations of the six counties are shown as red points in Figure 3. These data are subdivided into map sheets at 1:10,000 scale (in Jiangxi and Hainan Province) or 1:25,000 scale (in Shaanxi Province). The corresponding remote sensing image of one typical land cover map in County 3 is illustrated in Figure 3b.

Table 1. Information of the six study areas.

Study Area	Location	Area (km ²)	Landform
County 1	Midwest of Shaanxi Province	720	Rich types of landforms, such as hilly, river terrace, and loess tableland
County 2	Northeast of Shaanxi Province	428	Mainly ravine regions
County 3	Northwest of Jiangxi Province	1377.49	Mainly mountainous regions
County 4	Middle of Jiangxi Province	2439.33	Mainly hilly regions and river valleys
County 5	Northwest of Hainan Province	1317	Mainly flat regions
County 6	South Central of Hainan Province	1169	Mainly mountains and valleys



(a)

Figure 3. Cont.

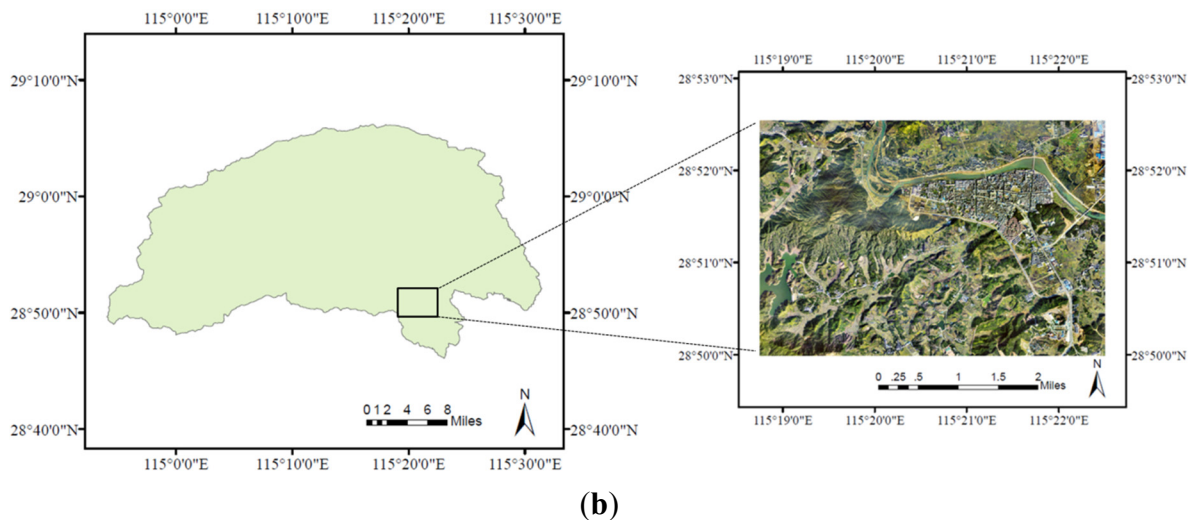


Figure 3. Study areas. (a) Chinese provincial administrative region map; (b) County 3 and the corresponding remote sensing image of one typical land cover map.

3.2. Reliability Evaluation Results

We established a stratified sampling for result reliability evaluation. The stratification was based on the landscape complexity and the production department. The sampling was applied in each study county and based on the sample strategy and the total land cover maps in each county, a total of 42 samples were selected. Reference data were collected by the professionals from the field survey and the high-resolution images used for the data production. For process reliability evaluation, we evaluated the total land cover maps without the sampling design.

Figure 4 shows the result reliabilities and process reliabilities of a total of 42 samples. Considering that NGSM applied a hierarchical classification system with three levels, the reliability of each level of the data could be obtained as shown in Figure 4a. Most data perform the reliabilities higher than 0.98 because these data were produced by the professional production departments and satisfied the accuracy requirements of the NGSM. For example, the percentage of features with first-level classification error in one map should be less than 0.3% according to the accuracy requirements of the NGSM. The reliability of the first level is higher than that of the other two levels because the first-level types are more easily recognized.

Table 2 shows the overall result reliabilities of the six study areas. The reliabilities of County 1, County 2, and County 3 are lower than those of County 4, County 5, and County 6 because of different production departments and landscape complexities. Experienced professionals and high-reliability images in the production department can normally facilitate the data production. On the contrary, complex landscapes in the county can increase the difficulty of the data production.

The reliabilities obtained from the process reliability analysis are presented as the reliability intervals as shown in Figure 4b. System bias existing in the process reliability analysis results in the process reliabilities lower than the result reliabilities. The differences in process reliabilities of land cover data are mainly caused by the reliabilities of basic events (e.g., operation staff, image spatial resolution and image currency) and the proportions of the land cover data acquired through field survey, machine interpretation, and artificial interpretation.

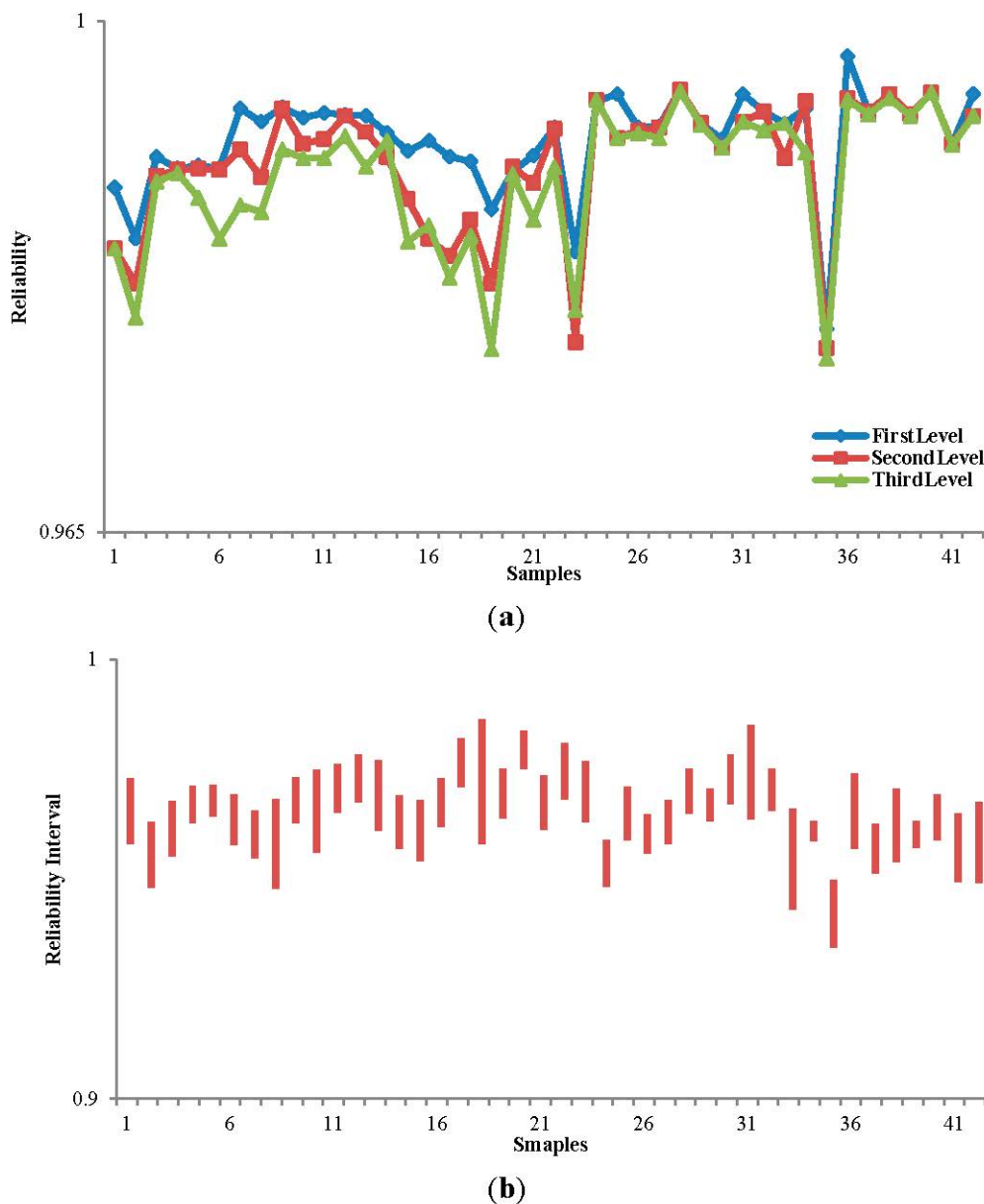


Figure 4. Reliabilities of the samples of six study areas. (a) Result reliabilities; (b) Process reliabilities.

Table 2. Overall result reliabilities of the six study areas.

Reliability Level	County 1	County 2	County 3	County 4	County 5	County 6
First level	0.98813	0.98997	0.98937	0.99307	0.99229	0.99337
Second level	0.98527	0.98983	0.98597	0.99075	0.99216	0.99289
Third level	0.98437	0.98753	0.98447	0.98915	0.9916	0.99289

3.3. Classification Correctness Analysis

Figure 5 shows the classification correctness of first-level land cover types in six cities. If the reliabilities of the three levels of a type are equal, then this means the error features of this type are classified to the wrong first-level types. If the second-level and third-level reliabilities of a type are lower than the first-level reliability, then this means a few error features of this type are classified to the wrong second-level or third-level, but the first-level type of the error features is true.

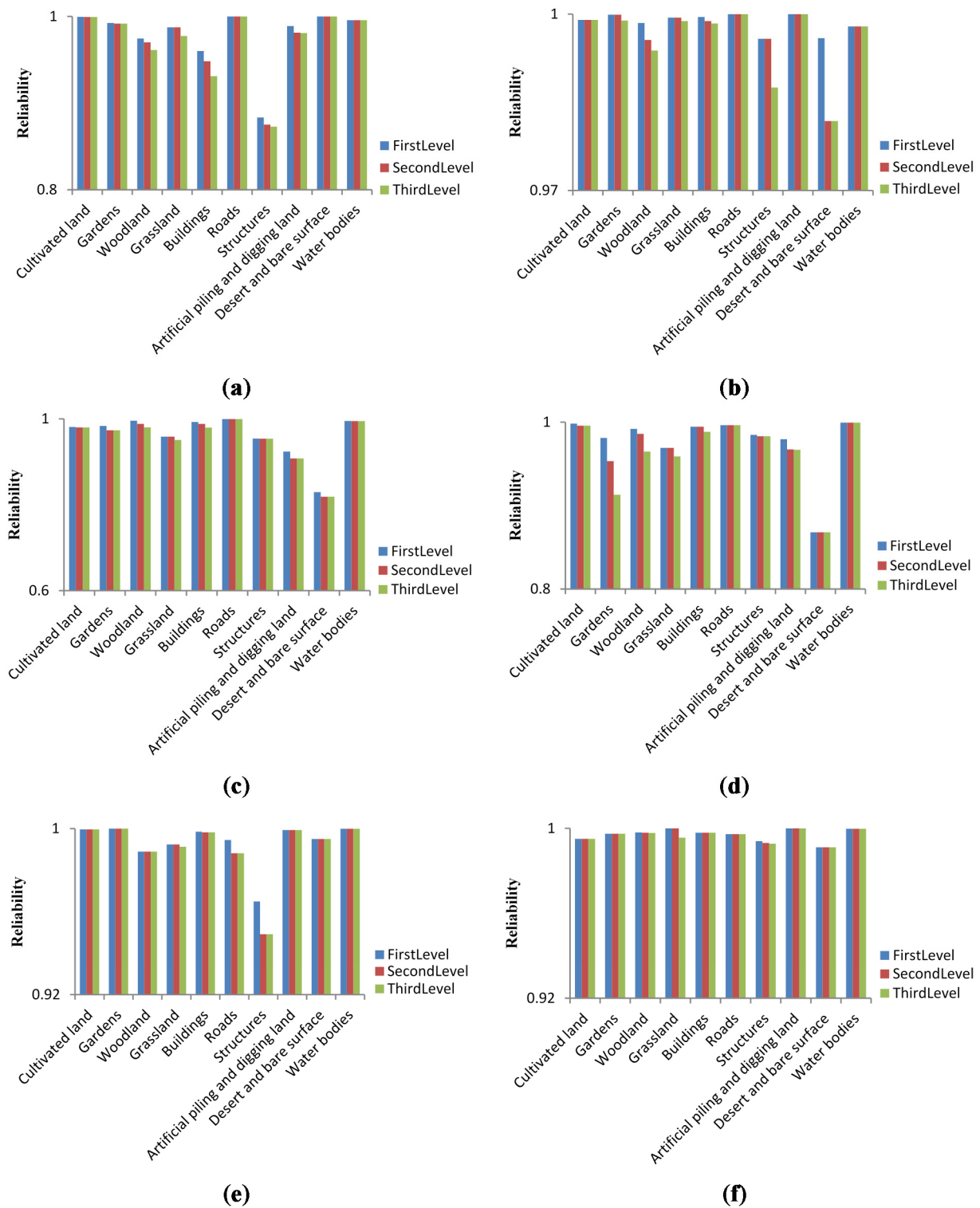


Figure 5. Classification correctness of first-level types in the six study areas: (a) County 1; (b) County 2; (c) County 3; (d) County 4; (e) County 5; and (f) County 6.

The classification correctness of woodland, buildings, and structures are lower than other types in County 1 (Figure 5a). In County 2 (Figure 5b), the second-level and third-level classification correctness of desert and bare surface are lower, whereas the first-level classification correctness is normal. This result is attributed to the difficulty of distinguishing between the second-level and third-level types of desert and bare surface in County 2, which is covered by loess tableland. The

situation is similar to the third-level types of structures. In County 4 (Figure 5d), the third-level classification correctness of gardens is low because it is difficult to distinguish between orchard, tea garden, nursery, and mulberry. Several features of desert and bare surface were wrongly classified to cultivated land, gardens, and artificial piling and digging land in County 4 and County 3 (Figure 5c). The classification correctness of structures in County 5 (Figure 5e) is low because different types of hardening land are difficult to distinguish. The overall classification correctness of first-level types in County 6 (Figure 5f) is high. The analysis of the classification correctness can be the reference for the routes of field surveys in the production process of land cover data.

3.4. Scale Reasonableness Analysis

The acquisition scale of the study data is 400 m², which means that features less than 400 m² were abandoned during data production. Therefore, we analyze the scale reasonableness of the study data at the scale of 200 and 100 m², which are represented as Scale 1 and Scale 2 in Figure 6, respectively, to obtain further details on the land cover information. The figure shows the scale reasonableness of first-level types in the six study areas. The scale reasonableness of buildings is lower than other land cover types in all the six counties because several separate buildings, which have an area of less than 400 m², exist in the land surface, and these buildings were abandoned during data production. The scale reasonableness of water bodies in County 3 (Figure 6c) and County 4 (Figure 6d) is lower than in other counties because many little broken water regions exist in the two counties of Jiangxi than in the counties of other provinces. The analysis on the scale reasonableness can be the reference for further data production to obtain more reliable land cover data. For example, a smaller acquisition scale for the buildings and the water bodies in regions containing many little broken water regions can be defined to obtain more details of actual land cover conditions and reliable land cover information.

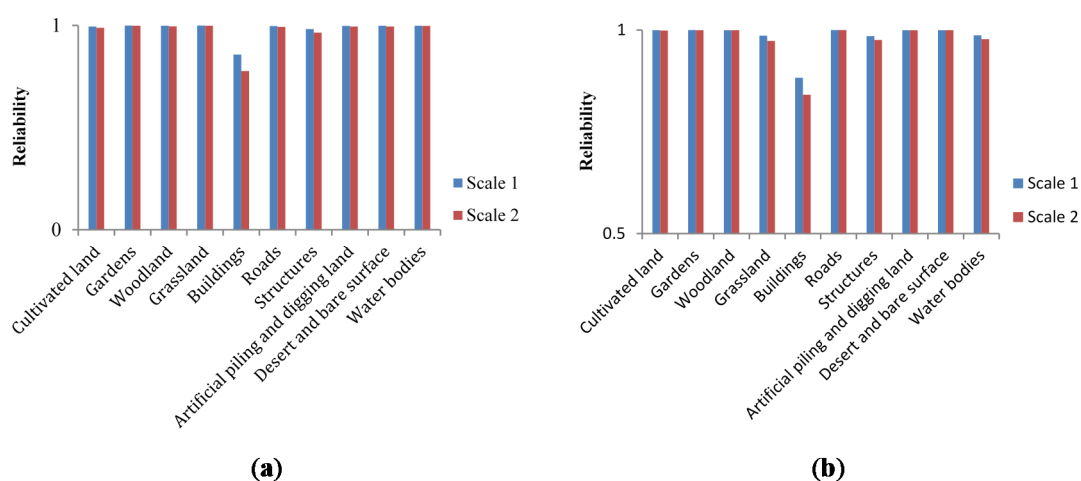


Figure 6. Cont.

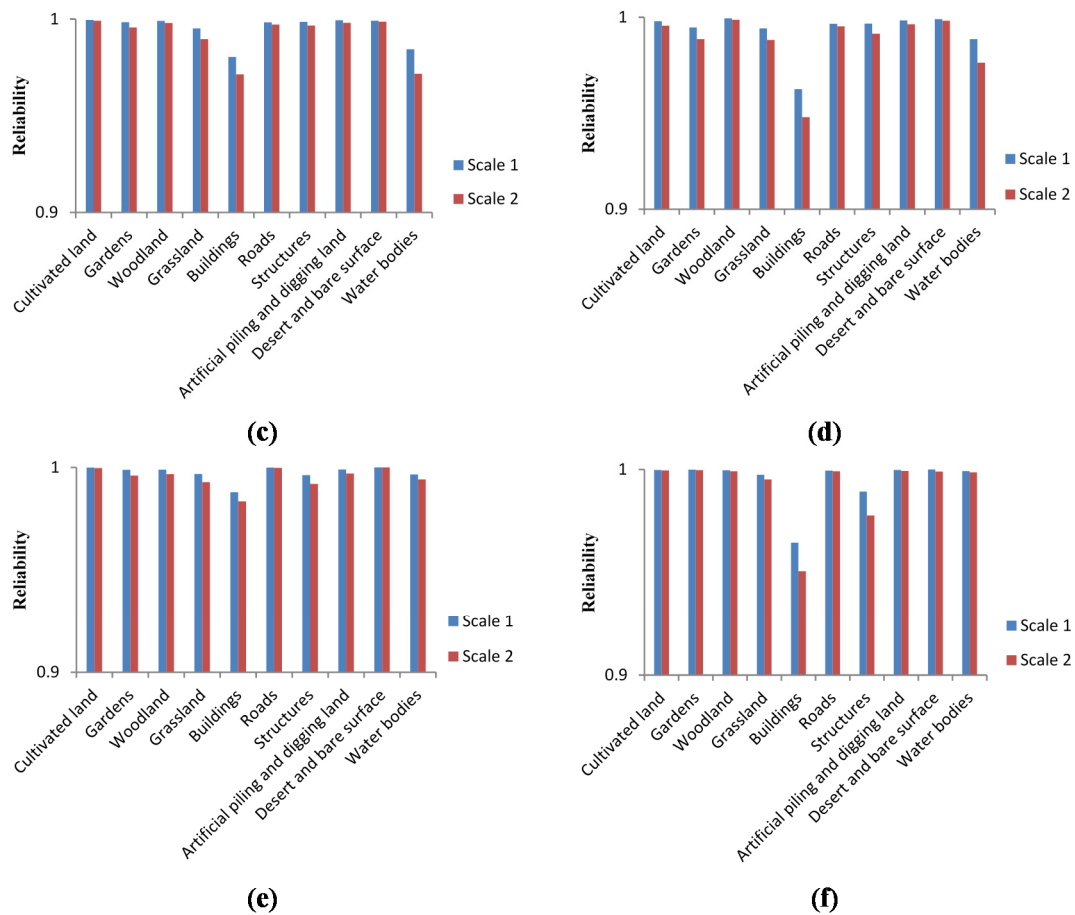


Figure 6. Scale reasonableness of first-level types in the six study areas: (a) County 1; (b) County 2; (c) County 3; (d) County 4; (e) County 5; and (f) County 6.

3.5. Consistency Analysis between Result Reliability and Process Reliability

The comparison between result reliabilities and process reliabilities is illustrated in Figure 7. In the figure, red points denote the first-level result reliabilities of the samples and blue points represent the medians of process reliability intervals. The gradients of reliabilities of the two samples were computed to compare two groups of reliabilities. The sign (positive or negative) of the gradient reflects the change trend of reliabilities of two samples, while the gradient value reflects the change magnitude between the reliabilities of two samples.

The comparison shows that the reliability trends of 34 samples are consistent between the two groups. Only eight of the 42 samples are inconsistent, shown as the points with black circle in Figure 7. The inconsistency happens when the data in the complex landscapes have numerous errors caused by visual interpretation or computer classification leading to low result reliability. Meanwhile, the weight of field survey in the data is too high, resulting in the high process reliability. The opposite case can happen in the homogenous landscapes. The gradient values of process reliabilities in 37 samples are larger than those of the result reliabilities, which means most change magnitudes between the reliabilities of the samples are different in the two groups. The comparison indicates that process reliability evaluation is reasonable and can reflect the change trend of reliabilities to some extent. The

comparison between the second-level or third-level result reliabilities and the process reliabilities can reach a similar conclusion.

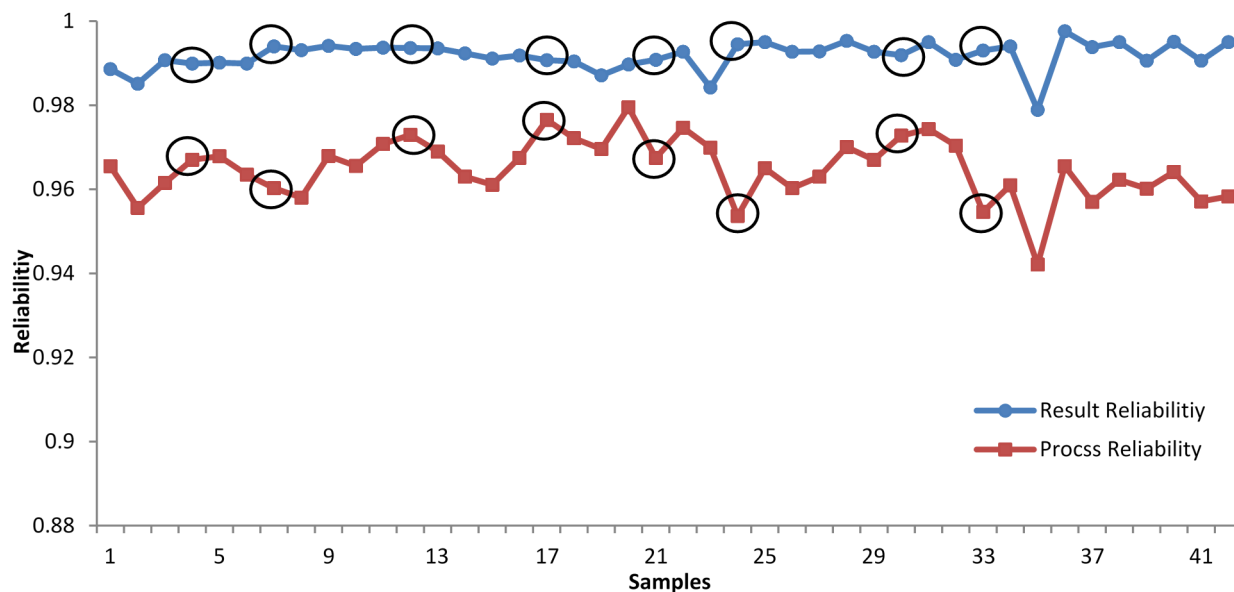


Figure 7. Comparison between first-level result reliabilities and process reliabilities.

4. Discussion

Experimental results show that the proposed reliability evaluation scheme is applicable in the validation of land cover products. Through the analysis on the seven result reliability indicators, multilevel and multitype reliabilities of land cover data can be obtained for different applications and strategic decisions. For example, the routes of field surveys can be established based on the land cover types with low classification correctness, and the analysis on the scale reasonableness can be the reference of further data acquisition scale for the land cover types with low scale reasonableness. In addition, process reliability evaluation with no need for reference data can save costs, facilitate the validation and solve the problem that the validation is often constrained by the lack of high-quality ground reference data to some extent. The consistency analysis between result reliability and process reliability shows that process reliability evaluation can reflect the consistent change trends of reliabilities with the result evaluation method to some extent.

However, limitations exist in the process reliability analysis model. As the production of land cover data is a complicated process with high uncertainty and the process reliability analysis model simplifies this process, system biases exist in the results of process reliability evaluation. Therefore, process reliability evaluation cannot take the place of result reliability evaluation for the validation of land cover products. The construction of the process reliability analysis model, including the definitions of the events and relationships among these events, should be further studied.

5. Conclusions

The reliability of land cover products has numerous impacts on the applications of the data and affects strategic decision and planning. The proposed reliability evaluation schema includes two reliability

evaluation methods, namely, result reliability evaluation and process reliability evaluation. Result reliability evaluation computes the reliability of land cover data by using seven reliability indicators. Process reliability evaluation obtains the reliability through the reliability analysis on the data production process without the need for reference data. Experimental results using land cover data in six typical counties show that the proposed reliability evaluation scheme is reasonable and can be applied in the validation of land cover products. More reliability information of land cover can be obtained for the strategic decision and planning through the analysis on the seven indicators of result reliability evaluation, compared with traditional accuracy assessment methods. Process reliability evaluation, which relatively saves costs and does not need reference data compared with result reliability evaluation, can facilitate the validation, reflect the change trends of reliabilities and solve the problem that the validation is often constrained by the lack of high-quality ground reference data to some extent. However, the process reliability analysis model should be improved in further studies.

Acknowledgments

This work was supported by the Ministry of Science and Technology of China (Project No.: 2012BAJ15B04).

Author Contributions

Wenzhong Shi had the original idea for the study and, with all co-authors carried out the design. Xiaokang Zhang, Ming Hao and Pan Shao were responsible for recruitment and follow-up of study participants. Xuzhe Lyu was responsible for data cleaning and Liping Cai carried out the analyses. Xiaokang Zhang drafted the manuscript, which was revised by all authors. All authors read and approved the final manuscript

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Sellers, P.; Dickinson, R.; Randall, D.; Betts, A.; Hall, F.; Berry, J.; Collatz, G.; Denning, A.; Mooney, H.; Nobre, C. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science* **1997**, *275*, 502–509.
2. Stehman, S.V.; Czaplewski, R.L. Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sens. Environ.* **1998**, *64*, 331–344.
3. Stehman, S.V. Sampling designs for accuracy assessment of land cover. *Int. J. Remote Sens.* **2009**, *30*, 5243–5272.
4. Mayaux, P.; Eva, H.; Gallego, J.; Strahler, A.H.; Herold, M.; Agrawal, S.; Naumov, S.; de Miranda, E.E.; Di Bella, C.M.; Ordoyne, C. Validation of the global land cover 2000 map. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1728–1739.

5. Stehman, S.V.; Olofsson, P.; Woodcock, C.E.; Herold, M.; Friedl, M.A. A global land-cover validation data set, II: Augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *Int. J. Remote Sens.* **2012**, *33*, 6975–6993.
6. Lunetta, R.S.; Iiames, J.; Knight, J.; Congalton, R.G.; Mace, T.H. An assessment of reference data variability using a virtual field reference database. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 707–715.
7. Foody, G.M.; Boyd, D.S. Using volunteered data in land cover map validation: Mapping west african forests. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1305–1312.
8. Foody, G.M. The impact of imperfect ground reference data on the accuracy of land cover change estimation. *Int. J. Remote Sens.* **2009**, *30*, 3275–3281.
9. Wulder, M.A.; White, J.C.; Magnussen, S.; McDonald, S. Validation of a large area land cover product using purpose-acquired airborne video. *Remote Sens. Environ.* **2007**, *106*, 480–491.
10. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201.
11. Wickham, J.; Stehman, S.; Fry, J.; Smith, J.; Homer, C. Thematic accuracy of the NLCD 2001 land cover for the conterminous united states. *Remote Sens. Environ.* **2010**, *114*, 1286–1296.
12. Wickham, J.D.; Stehman, S.V.; Gass, L.; Dewitz, J.; Fry, J.A.; Wade, T.G. Accuracy assessment of NLCD 2006 land cover and impervious surface. *Remote Sens. Environ.* **2013**, *130*, 294–304.
13. Elsayed, E.A. *Reliability Engineering*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
14. Amirat, A.; Mohamed-Chateaufneuf, A.; Chaoui, K. Reliability assessment of underground pipelines under the combined effect of active corrosion and residual stress. *Int. J. Press. Vessel. Pip.* **2006**, *83*, 107–117.
15. Goševa-Popstojanova, K.; Trivedi, K.S. Architecture-based approach to reliability assessment of software systems. *Perform. Eval.* **2001**, *45*, 179–204.
16. Soszynska, J. Reliability and risk evaluation of a port oil pipeline transportation system in variable operation conditions. *Int. J. Press. Vessel. Pip.* **2010**, *87*, 81–87.
17. Zhang, P.; Li, W.; Li, S.; Wang, Y.; Xiao, W. Reliability assessment of photovoltaic power systems: Review of current status and future perspectives. *Appl. Energy* **2013**, *104*, 822–833.
18. Zhang, P.; Shi, W.; Wong, M.S.; Chen, J. A reliability-based multi-algorithm fusion technique in detecting changes in land cover. *Remote Sens.* **2013**, *5*, 1134–1151.
19. Shi, W. *Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses*; CRC Press: Boca Raton, FL, USA, 2009.
20. Shi, W.; Chen, J.; Zhan, Q.; Shu, H. Reliable spatial analysis. *Geomat. Inf. Sci. Wuhan Univ.* **2012**, *37*, 883–887. (In Chinese)
21. Cihlar, J. Land cover mapping of large areas from satellites: Status and research priorities. *Int. J. Remote Sens.* **2000**, *21*, 1093–1114.
22. Smits, P.; Dellepiane, S.; Schowengerdt, R. Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *Int. J. Remote Sens.* **1999**, *20*, 1461–1486.
23. Chang, J.-R.; Chang, K.-H.; Liao, S.-H.; Cheng, C.-H. The reliability of general vague fault-tree analysis on weapon systems fault diagnosis. *Soft Comput.* **2006**, *10*, 531–542.
24. Yuhua, D.; Datao, Y. Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis. *J. Loss Prev. Process Ind.* **2005**, *18*, 83–88.

25. Suresh, P.; Babar, A.; Raj, V.V. Uncertainty in fault tree analysis: A fuzzy approach. *Fuzzy Sets Syst.* **1996**, *83*, 135–141.
26. Singer, D. A fuzzy set approach to fault tree and reliability analysis. *Fuzzy Sets Syst.* **1990**, *34*, 145–155.
27. Zhang, Z.; Wang, X.; Zhao, X.; Liu, B.; Yi, L.; Zuo, L.; Wen, Q.; Liu, F.; Xu, J.; Hu, S. A 2010 update of national land use/cover database of China at 1:100000 scale using medium spatial resolution satellite images. *Remote Sens. Environ.* **2014**, *149*, 142–154.
28. Haarla, L.; Pulkkinen, U.; Koskinen, M.; Jyrinsalo, J. A method for analysing the reliability of a transmission grid. *Reliab. Eng. Syst. Saf.* **2008**, *93*, 277–287.
29. Bobbio, A.; Portinale, L.; Minichino, M.; Ciancamerla, E. Improving the analysis of dependable systems by mapping fault trees into bayesian networks. *Reliab. Eng. Syst. Saf.* **2001**, *71*, 249–260.

© 2015 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 license (<http://creativecommons.org/licenses/by/4.0/>).