



Validation of Land Use Cover Maps: A Guideline

María Teresa Camacho Olmedo, David García-Álvarez, Marta Gallardo, Jean-François Mas, Martin Paegelow, Miguel Ángel Castillo-Santiago, and Ramón Molinero-Parejo

Abstract

This chapter offers a general overview of the available tools and strategies for validating Land Use Cover (LUC) data—specifically LUC maps—and Land Use Cover Change Modelling (LUCCM) exercises. We give readers some guidelines according to the type of maps they want to validate: single LUC maps (Sect. 3), time series of LUC maps (Sect. 4) or the results of LUCCM exercises (Sect. 5). Despite the fact that some of the available methods are applicable to all these maps, each type of validation exercise has its own particularities which must be taken into account. Each section of this chapter starts with a brief introduction about the specific type of maps (single, time series or modelling exercises) and the reference data needed to validate them. We also present the validation methods/functions and the corresponding exercises developed in Part III of this book. To this end, we address, in this order, the tools for validating Land

Use Cover data based on basic and Multiple-Resolution Cross-Tabulation (see chapter “[Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps](#)”), metrics based on the Cross-Tabulation matrix (see chapter “[Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps](#)”), Pontius Jr. methods based on the Cross-Tabulation matrix (see chapter “[Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps](#)”), validation practices with soft maps produced by Land Use Cover models (see chapter “[Validation of Soft Maps Produced by a Land Use Cover Change Model](#)”), spatial metrics (see chapter “[Spatial Metrics to Validate Land Use Cover Maps](#)”), advanced pattern analysis (see chapter “[Advanced Pattern Analysis to Validate Land Use Cover Maps](#)”) and geographically weighted methods (see chapter “[Geographically Weighted Methods to Validate Land Use Cover Maps](#)”).

M. T. Camacho Olmedo (✉)
Departamento de Análisis Geográfico Regional y Geografía
Física, Universidad de Granada, Granada, Spain
e-mail: camacho@ugr.es

D. García-Álvarez · R. Molinero-Parejo
Departamento de Geología, Geografía y Medio Ambiente,
Universidad de Alcalá, Alcalá de Henares, Spain

M. Gallardo
Departamento de Geografía, Universidad Nacional de Educación a
Distancia, Madrid, Spain

J.-F. Mas
Laboratorio de Análisis Espacial, Centro de Investigaciones en
Geografía Ambiental, Universidad Nacional Autónoma de
México, Morelia, Mexico

M. Paegelow
Département de Géographie, Aménagement et Environnement,
Université de Toulouse Jean Jaurès, Toulouse, France

M. Á. Castillo-Santiago
Departamento de Observación y Estudio de la Tierra, la Atmósfera
y el Océano, El Colegio de la Frontera Sur, San Cristóbal de las
Casas, Mexico

Keywords

Land Use Cover • Land Use Cover Change Modelling
exercises • Validation

1 Introduction

Validation is a required step prior to the effective use of any Land Use Cover (LUC) dataset or of the results of a Land Use Cover Change Modelling (LUCCM) exercise. We need to understand to what extent these datasets and results are uncertain in order to be able to assess the limits that these uncertainties may impose on the conclusions of our analyses and studies.

There are many methods, tools and strategies currently available for validating LUC data and LUCCM exercises. However, comprehensive guidelines providing users with clear instructions and recommendations about how to carry

out this validation are scarce. Olofsson et al. (2013, 2014) review the validation of land change maps and offer a series of recommendations as to how to perform a credible scientific validation, accepting that other recommendations or good practice guidelines could be equally valid and perhaps even more so. Paegelow et al. (2014, 2018) propose a variety of validation techniques and error analysis which can be used to validate different LUCCM exercises.

In this chapter, we aim to provide readers with a general overview of the available tools and strategies for validating LUC data—specifically LUC maps—and LUCCM exercises. We give readers different guidelines according to the type of maps they want to validate: single LUC maps (Sect. 3), time series of LUC maps (Sect. 4) and results of LUCCM exercises (Sect. 5). Although some of the available methods and tools can be applied to all these maps, each type of validation exercise has its own specific aspects that users must bear in mind. For example, the results of LUCCM exercises include soft and hard LUC maps. The hard outputs of a model—hard maps—are very similar to input LUC maps, while the soft outputs—soft maps—are continuous and ranked. We therefore also present some validation methods that focus specifically on soft maps.

Before presenting these validation methods and functions, it is important to make clear that visual inspection is an essential part of any validation exercise. It can provide a great deal of information about the uncertainties of the data being evaluated, which are not detected by the quantitative methods reviewed in this book. Visual inspection should be conducted during all validation exercises, at the beginning, at the end and throughout the entire process.

2 Validation Methods/Functions and Exercises Presented in Part III of This Book

This chapter is intended as a presentation of Part III of this book. Figure 1 shows the validation methods/functions and the corresponding exercises presented in the chapters and sections of Part III. With this in mind, in this chapter we address, in this order: the available tools for validating Land Use Cover data related with basic and Multiple Resolution Cross-Tabulation (see chapter “[Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps](#)”), metrics derived from the Cross-Tabulation matrix (see chapter “[Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps](#)”), methods proposed by Pontius Jr. based on the Cross-Tabulation matrix (see chapter “[Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps](#)”), validation practices with soft maps produced by Land Use Cover Change models (see chapter “[Validation of Soft Maps](#)

Produced by a Land Use Cover Change Model”), spatial metrics (see chapter “[Spatial Metrics to Validate Land Use Cover Maps](#)”), advanced pattern analysis (see chapter “[Advanced Pattern Analysis to Validate Land Use Cover Maps](#)”) and geographically weighted methods (see chapter “[Geographically Weighted Methods to Validate Land Use Cover Maps](#)”).

The exercises presented in Part III have been applied using the Quantum GIS (QGIS) software and R scripts. To homogenize the exercises across the different chapters, they have the same standard objectives: *to validate a map (t_1) against reference data/map (t_1)* (single LUC map); *to validate a series of maps with two or more time points (t_0, t_1, t_2, \dots)* (LUC maps series/ LUC changes); and, for results from LUCCM exercise, *to validate soft maps produced by the model against a reference map of changes ($t_0 - t_1$)* (soft LUC maps), *to validate a simulation (T_1) against a reference map (t_1)* (single LUC map - hard LUC maps) and *to validate simulated changes ($t_0 - T_1$) against a reference map of changes ($t_0 - t_1$)* (LUC maps series / LUC changes - hard LUC maps). However, in certain specific cases, additions have been made to these standard titles. In addition to the applications of each method/function implemented in the practical exercises in this book, the cells shaded in grey in Fig. 1 indicate that the method has other potential applications that are not described here.

3 Validation of Single Land Use Cover Maps

The validation of single LUC maps is the most widespread practice of all those addressed in this book. Foody (2002) concludes that there is no single universally acceptable measure of accuracy but rather a variety of indices, each sensitive to different features. Creating a single, all-purpose measure of classification accuracy would therefore seem an almost impossible goal. However, accuracy assessment must follow certain guidelines and principles in order to guarantee scientifically defensible assessment of map accuracy (Stehman 1999; Stehman and Czaplewski 1998).

Users have been validating their maps since the advent of digital remote sensing and the first classifications of digital imagery, as a means of assessing to what extent the classified images resemble the real LUC on the ground. Now, several decades later, the validation of single LUC maps is a very common practice, and although new methods and tools have been developed over the years, the original ones remain popular. These are based above all on the comparison of the assessed LUC map with reference datasets through cross-tabulation (Foody 2002; Strahler et al. 2006). In recent years, the use of pattern analysis and other validation methods has become increasingly common.

Validation methods/functions		Single LUC map	LUC maps series/ LUC changes	LUCC modelling exercises		
				Soft LUC maps	Hard LUC maps	
		<i>To validate a map (t) against reference data/map (t)</i> ¹	<i>To validate a series of maps with two or more time points (t₀, t₁, t₂...)</i>	<i>To validate soft maps produced by the model against a reference map of changes (t₀ - t)</i>	<i>To validate a simulation (T_i) against a reference map (t)</i>	<i>To validate simulated changes (t₀ - T_j) against a reference map of changes (t₀ - t)</i>
Book chapters	Chapter sections					
Basic and Multiple-Resolution Cross Tabulation	1. Basic Cross Tabulation	Exercise 1	Exercise 4	Exercise 2	Exercise 3	
	2. Multiple-Resolution Cross Tabulation	Exercise 1		Exercise 2	Exercise 3	
Metrics based on a Cross Tabulation Matrix	1. Change statistics		Exercise 1			
	2. Areal and spatial agreement metrics	Exercise 1			Exercise 2	Exercise 3
	3. Kappa Indices	Exercise 1			Exercises 2, 3	
	4. Agreement between maps at global and stratum level					Exercise 1
	5. Accuracy assessment statistics	Exercise 1				
Pontius Jr. methods based on a Cross Tabulation matrix	1. Null Model		Exercise 1			Exercise 1
	2. LUC Budget		Exercise 1			
	3. Quantity and allocation disagreement		Exercise 1			
	4. Figure of Merit (FoM) and complementary Producer's and User's accuracy					Exercises 1, 2, 3
	5. Incidents and States		Exercise 1			
	6. Intensity analysis		Exercise 1			
	7. Flow matrix		Exercise 1			
Validation of soft maps produced by a Land Use Cover Change model	1. Correlation			Exercise 1		
	2. Receiver Operating Characteristic (ROC)			Exercise 1		
	3. Difference in Potential (DiP)			Exercises 1, 2		
	4. Total uncertainty, quantity uncertainty, allocation uncertainty			Exercise 1 ²		
Spatial metrics	1. Spatial metrics	Exercise 1	Exercises 4, 5, 6		Exercise 2	Exercise 3
Advanced pattern analysis	1. Map Curves	Exercise 1	Exercise 4		Exercise 2	Exercise 3
	2. Change on pattern borders		Exercise 1			
	3. Allocation distance error				Exercises 1, 2	
Geographically weighted methods	1. Overall, user's and producer's accuracy through GWR	Exercise 1				

¹ Titles of exercises included in the corresponding sections.

² This validation technique is not calculated against a reference map of changes.

Fig. 1 Validation methods/functions and corresponding exercises presented in Part III of this book for single LUC maps, LUC maps series/LUC changes and LUCCM exercises. The grey cells highlight the possible applications of each method/function

The reference datasets for validating single LUC maps may be obtained from different sources of LUC data. These can be classified into two main groups: ground samples and reference LUC maps. However, in the validation exercises, other reference spatial data can also be used, such as the raw imagery used in the classification process or the soft maps obtained as a result.

The ground samples collected through field surveys provide highly accurate, detailed data. However, this information is very expensive to obtain and fieldwork is not an option when working with large study areas. This is why most reference LUC samples are obtained by photointerpretation or classification of satellite imagery. The data obtained via photointerpretation must be of higher quality than the data being validated. This usually involves careful interpretation of a set of samples using imagery with a higher spatial resolution than the images used to create the map. Another option is photointerpretation of the same imagery used to obtain the dataset, applying a different workflow and methods or techniques that guarantee better quality.

Those using these methods to obtain LUC samples for validation purposes should provide information about their accuracy or uncertainty. When obtaining reference data by field surveys or photointerpretation, users must take particular care when selecting the sampling strategy they will apply during the collection of this information, as it can have an important impact on the results of the validation exercise

and on their validity (see chapter “[Visualization and Communication of LUC Data](#)”).

LUC maps can also be validated against other LUC maps. In these cases, the reference LUC map must have a higher spatial resolution and greater detail than the map being assessed. They must also be of proven quality, i.e. maps or datasets with verified accuracy and uncertainty. Although less precise, validation exercises carried out by comparing the evaluated map with other LUC maps are quick and very cheap, hence their popularity. This also allows a wider set of methods and techniques to be used compared to the possibilities offered by reference datasets other than maps.

Users can also validate their LUC maps against additional sources of information other than reference datasets, in order to characterize the maps in more detail and gain a clearer picture of their uncertainty. Such sources include raw imagery, which is often used in the classification process, or the soft maps obtained from it, which are used to assess the characteristics of the pixels that make up each class. Raw imagery can be used to evaluate the reflectance value for all the pixels belonging to a particular class and how close it is to the reference reflectance value used in the classification process. When available, users can also compare each category pixel with soft maps showing the percentage of each pixel belonging to each of the LUC categories under consideration. Similar insights into the accuracy of LUC maps can be obtained by comparing them with continuous LUC

data (reference data), such as the Vegetation Continuous Fields (VCF) products.

If we focus on validation tools (Fig. 1), the agreement between the reference data/map (t_1) and the LUC map under evaluation (t_2)—the two maps should have the same date t_1 —can be assessed using the **cross-tabulation matrix**¹ (see Sect. 1 in chapter “Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps”). This is also referred to in the literature as the confusion or error matrix, or as the contingency table. Cross tabulation is usually the first step in any validation exercise, as the raw matrix provides plenty of information regarding the spatial agreement between the LUC map being validated and the reference dataset.

In some cases, the level of agreement may vary at different levels of spatial detail. For example, when spatially aggregated and simplified, the LUC map being evaluated may show more agreement with the reference dataset. The choice of spatial resolution is therefore a source of uncertainty. To account for this uncertainty, we can cross-tabulate the assessed and reference datasets at **multiple spatial resolutions** (see Sect. 2 in chapter “Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps”), i.e. the original resolution and other coarser ones.

Different metrics are calculated from the confusion matrix (see chapters “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps” and “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). These metrics summarize the agreement between reference and validated datasets in a single value and are therefore very easy to interpret. As a result, they have been widely used in LUC validation.

The most common metrics are the **accuracy assessment statistics** (see Sect. 5 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) and the **Kappa Indices** (see Sect. 3 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). The accuracy assessment statistics are standard metrics that provide information about the similarity between two georeferenced data. They are obtained from the cross-tabulation matrix and enable the extraction of specific information contained in the matrix. They include, among others, the overall, producer’s and user’s accuracy metrics. They are usually supplied with the cross-tabulation matrix, providing extra information in addition to that provided by the matrix itself (e.g. category area adjusted by the error level, confidence intervals...).

Of all these metrics, the most commonly used in validation exercises is probably Overall accuracy. There has been great debate in the literature about the threshold above which the Overall accuracy of a map can be considered acceptable. The 85% threshold proposed by Anderson (1971) was the common reference for many years and continues to be applied by a lot of users nowadays (Wulder et al. 2006; Foody 2008). However, there is no specific accuracy threshold regarded as valid for all study cases and datasets. The acceptable level of accuracy will depend on the intended application of the dataset and the characteristics of the area being mapped. As regards different scales and spatial resolution, we cannot compare the accuracy of global or supra-national LUC maps with that of regional and local ones, which are not subject to the same level of simplification or abstraction as the global or supra-national maps.

The overall accuracy metric does not provide information about the accuracy at which each category on the LUC map is mapped. Important differences are often identified in terms of the relative accuracy of the different categories. Mixed LUC categories do not usually show the same accuracy as spectrally pure categories. At high levels of thematic detail, very similar LUC categories can be easily confused and will, therefore, have lower levels of accuracy. Users must take these differences at the category level into account and report the accuracy values for each category. The general approach for **agreement between maps at global and stratum level** may be useful to this end (see Sect. 4 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). Some authors talk specifically about Overall and Individual Spatial Agreement, proposing different metrics for these purposes (Yang et al. 2017; Islam et al. 2019) (see **Areal and spatial agreement metrics** in Sect. 2 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”).

It is also important to remember that the accuracy of a LUC map is not usually the same across the entire mapped area and considerable spatial variations are possible. The bigger the area being mapped, the more likely it is for there to be spatial differences in accuracy levels across the mapped area. The cross-tabulation matrix does not provide information about these spatial differences. When mapping large study areas made up of different, clearly distinguishable regions, each region can be validated independently, producing a specific cross-tabulation matrix in each case. The global analysis would cover the entire map, while specific areas of the map (e.g. a region, a municipality...) could also be analysed at the stratum level.

Overall Accuracy is highly correlated with the Kappa Index (Olofsson et al. 2014), which explains why both metrics provide similar information. One difference is that Kappa takes into account the agreement expected by chance,

¹ The methods/functions presented in the corresponding chapters in Part III of this book are highlighted in **bold**.

a factor that is not considered in Overall Accuracy. The Kappa Index (see Sect. 3 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) has been criticized by a range of authors, who claim that it can sometimes be misleading (Pontius and Millones 2011; Olofsson et al. 2014). Moreover, standard indices such as overall, producer’s and user’s accuracy have the advantage that they can be interpreted as measures of the probability of encountering pixels, patches, etc. that have been allocated to the correct category (Stehman 1997).

The methods mentioned above do not employ fuzzy logic and, instead, apply a binary logic when calculating agreement, i.e. the two elements agree or don’t agree. Partial agreements are not considered. However, there are some tools for calculating map agreement that incorporate fuzzy logic, such as the Fuzzy Kappa or the Fuzzy Kappa Simulation (Woodcock and Gopal 2000).

Other metrics, similar to Kappa, have also been proposed. Usually they aim to outperform Kappa and correct some of its associated problems. These include, among others, the F-Score (Pérez-Hoyos et al. 2020), Scott’s pi statistic (Gwet 2002) and Krippendorff’s α -coefficient (Kerr et al. 2015). These metrics are not widely used and they provide similar information to Kappa, which is why we do not recommend that they be used in a standard LUC validation exercise.

Extensive research by Pontius Jr. has given rise to other metrics based on the cross-tabulation matrix which can be used to validate a single LUC map against a reference map (see chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). **Quantity & allocation disagreement** (see Sect. 3 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (Pontius and Millones 2011) compares the agreement between maps regarding the proportions allocated to the different categories and regarding the way they are allocated, i.e. differences in the quantities allocated to each category and differences in their location. These metrics complement the cross-tabulation table, so enabling users to take full advantage of the information it provides. Quantity and Allocation disagreement is a very good method for validating a single map against a reference map (García-Álvarez and Camacho Olmedo 2017).

Users can also specifically assess the pattern of the map they want to validate to find out how much its pattern coincides with that of the reference map. Pattern agreement can be assessed using **Spatial metrics** (see Sect. 1 in chapter “Spatial Metrics to Validate Land Use Cover Maps”) and the **Map Curves** method (see Sect. 1 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”). Spatial metrics allow us to characterize different aspects of the map’s pattern in detail, such as its fragmentation, the proportion allocated to each category, the complexity of the patches... (Botequilha et al. 2006; Forman 1995). Initially

developed within the field of landscape ecology, these metrics are also widely used for characterizing the pattern of categorical maps. For its part, Mapcurves (Hargrove et al. 2006) provides a single value summarizing the pattern agreement between two maps. In both cases, we should always compare maps drawn at the same spatial and thematic resolution, as any changes in resolution would severely alter the pattern of the map, so rendering the comparison uninformative.

Geographic weighting methods (GWR) (see chapter “Geographically Weighted Methods to Validate Land Use Cover Maps”) can also be used to study the spatial distribution of LUC accuracy measures. The overall, user’s and producer’s accuracy metrics mentioned above are derived from the cross-tabulation matrix and are therefore not spatial metrics, i.e. they provide overall information for the entire area, without assessing the spatial distribution of error and accuracy. The application of **Overall, user’s and producer’s accuracy metrics through GWR** (see Sect. 1 in chapter “Geographically Weighted Methods to Validate Land Use Cover Maps”) can help the user to assess the suitability of the LUC data and to observe local variations in accuracy and error on the map (Comber 2013). In some cases, local assessments may be necessary because they can uncover possible clusters of errors in the LUC data. By adapting logistic Geographically Weighted Regression (GWR) (Brunsdon et al. 1996), the spatial variations in Boolean LUC (classified data) and fuzzy LUC (reference data) can be modelled, providing maps that show the distribution of the overall, user’s and producer’s accuracy metrics.

4 Validation of Land Use Cover Maps Series/Land Use Cover Changes

There is no common practice or set of methods for validating or evaluating the uncertainty of a LUC map series with two or more time points (t_0, t_1, t_2, \dots). Most of the exercises for the validation of LUC data only refer to single LUC maps, without focusing specifically on the LUC change studied through a series of LUC maps.

One of the facets that users most demand from LUC data is the ability to study and display LUC changes over time. We therefore need methods and tools to assess the uncertainty of the changes that are measured from LUC maps. It is worth noting that the individual accuracy of two LUC maps involved in a post-classification comparison offers few clues as to the accuracy of change, because the relation between the errors in the two maps is unknown. As pointed out by Olofsson et al. (2013), even when both maps are highly accurate, it is possible that the change map accuracy will be low and the estimated change area strongly biased.

One of the main limitations when it comes to validating LUC changes and LUC map series is the lack of reference data. We could obtain reference datasets via photointerpretation or field surveys. However, it is difficult to guess where the LUC changes will take place, as they may happen at different places and with different intensities and patterns over space and time. In addition, there is a clear lack of LUC map series showing accurate, validated LUC change that could be used as reference data. Another option would be to validate the LUC changes against other types of reference data. This could be done for example by comparing the LUC changes measured over a time series of LUC maps against the difference in reflectance between two satellite images for the same time period. This is because when LUC change takes place, there is a significant change in the reflectance value registered by the satellite capturing the images.

Nevertheless, as commented earlier, the most common situation is that there are no reference datasets available. In these cases, the uncertainty of the LUC map series must be assessed by evaluating the consistency and the logic of the measured LUC change. The tools and techniques recommended here provide a great deal of information to the user. However, the final interpretation of the measured LUC change will be subjective, based on the user's expertise and understanding of the study area. In this situation, visual inspection can be very useful for quickly understanding many of the uncertainties in the time series of LUC maps that cannot be measured using quantitative metrics. This is why we recommend visual inspection as a first essential step prior to the validation of any LUC map or LUC modelling exercise.

Users must be aware that LUC change usually represents a very small portion of the mapped area. For a specific, not very large landscape, we would only expect a few features to change over a short period of time. In addition, the same area would not normally be expected to be affected by various successive changes. On the contrary, when an area changes, the new land use or cover tends to remain unchanged over time. In addition, there are some LUC transitions that make less sense than others. For example, one would not expect an artificial area to change to vegetation or agricultural land. These general assumptions may be adapted in line with the particular characteristics of the study area and also within the context of each element being analysed.

The same validation techniques reviewed above for single LUC maps (Sect. 3) can also be applied when comparing measured and reference changes or just for evaluating the consistency and logic of measured LUC change. However, some tools are specific to time series (Fig. 1).

The **cross-tabulation matrix** (see Sect. 1 in chapter "Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps") is the tool that provides most information about the change happening between two LUC

maps. For a time series, we can compare each pair of LUC maps to find out the changes that take place at each date and the area they cover, for the map as a whole and at category level. We can summarize the main processes of change in our study area, such as, for example, the artificialization or deforestation rates for each time period. This gives us an overview of the change that has taken place over our map series and makes it easier to interpret some of the inconsistencies in measured change. Some authors also propose making a summary of all the transitions taking place, associating some of them with a default degree of uncertainty (Gómez et al. 2016; Hao and Gen-Suo 2014). For example, a transition from artificial surfaces to agricultural areas is not expected and could therefore be assigned a high degree of uncertainty.

Multi-resolution cross-tabulation (see Sect. 2 in chapter "Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps") offers a means of checking whether some of the errors, inconsistencies or uncertainties we detect at the original resolution are not detected at coarser resolutions. When this happens, the errors and inconsistencies probably arise due to the level of detail at which the dataset was created.

The cross-tabulation matrix is an excellent source of information, which we can easily summarize using other tools and metrics. As commented in Sect. 3, **Areal and spatial agreement metrics** (see Sect. 2 in chapter "Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps") and **Kappa Indices** (see Sect. 3 in chapter "Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps") are used to assess the agreement between two maps. Despite their limitations, these metrics can be used to chart, in a generic way, the persistence or changes between two dates. If two maps in a series undergo the normal rate of change that we associate with any landscape, the differences between them should be slight, which means that the Kappa and agreement metrics should reflect high levels of coincidence between the maps being compared.

The **Agreement between maps at global and stratum level** (see Sect. 4 in chapter "Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps") analysis could provide additional specific information about the agreement in a time series of LUC maps at whole map level, or for a given stratum, i.e. a smaller area or a specific LUC category. **Accuracy assessment statistics** can also be calculated for a LUC map series, either globally (see Sect. 5 in chapter "Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps") or locally (Sect. 1 in chapter "Geographically Weighted Methods to Validate Land Use Cover Maps"). For example, when the LUC map series is obtained using a base map that is progressively updated, the first stage is to validate the base map

of the series using the same procedure described earlier for validating single LUC maps. Once this has been done, we can validate the changes against a reference dataset of changes through cross-tabulation, obtaining from the resulting table the overall, producer's and user's accuracy metrics. Pouliot and Latifovic (2013) coined the term Update Accuracy (UA) to refer to the accuracy of the measured changes. They refer to the accuracy of the base map as the Base Map Accuracy (BMA). They also propose a metric called Time Series Accuracy (TSA) as the mean accuracy of all the LUC maps that make up the series, individually validated through a specific reference LUC dataset for each case.

Change statistics (see Sect. 1 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (FAO 1995; Puyravaud 2003) are widely used to assess land use and cover changes. These indices measure, for example, relative change or rates of change and allow us to compare the change between regions of different sizes. These indices can be complemented by the change matrix obtained from cross-tabulation. They are calculated from the map series itself, rather than from the cross-tabulation matrix.

Robert Gilmore Pontius Jr. has made major contributions to the family of validation techniques based on the cross-tabulation matrix (chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). The **LUCC budget** (see Sect. 2 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (Pontius et al. 2004) provides more information about the changes that take place between pairs of maps. It differentiates between net and gross changes, therefore, allowing us to gain a clearer understanding of the transitions and swaps between categories, providing useful additional information to identify category confusion over time. Category confusion arises when the same area is mapped as different, albeit similar, categories at different points in time, when no change has actually taken place.

Quantity and allocation disagreement (see Sect. 3 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) show, at overall and category level, differences between pairs of maps in terms of category proportions due to the different allocation of the categories. Few changes are expected in a time series of maps. This means that quantity and allocation disagreement should be low and should centre on the most dynamic categories.

The number of **incidents and states** (see Sect. 5 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (Pontius et al. 2017) also provides information that can help identify errors. This technique allows us to identify those areas that are more dynamic than expected, i.e. those that change a lot over a

short period of time, always transitioning between the same categories. **Intensity analysis** (see Sect. 6 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (Aldwaik and Pontius 2012) compares the rates of LUC change between periods, categories, and transitions. Based on the assumption that a category or area is expected to change at similar levels of intensity over time, this analysis enables us to identify those categories that do not comply with this assumption. The **Flow matrix** (see Sect. 7 in chapter “Pontius Jr. Methods Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) (Runfola and Pontius 2013) measures the instability of annual land use change over different time intervals, so as to identify anomalies relative to the amount of change over the whole time series.

Spatial metrics (see Sect. 1 in chapter “”) and **Map curves** (see Sect. 1 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) enable us to characterize the pattern of each LUC map in the series. We do not expect the pattern of the map to vary significantly over the time period being analysed. This means that only smooth changes should be observed when comparing the spatial metrics for each of the periods analysed.

Spatial metrics that specifically measure the areas that change between pairs of maps may also be useful. In the case of a pair of maps or a time series, the detection of **change on pattern borders** (see Sect. 2 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) (Paegelow et al. 2014) enables us to identify data errors resulting from different data sources, different classifiers or spectral responses. For example, the noise or error shown by a time series of LUC maps often arises due to border areas between categories being interpreted differently each year. Users can specifically analyse the changes that take place in these border patches, often elongated and less than 1 or 2 pixels wide, so helping them to identify potential errors. These patches can also be characterized through the calculation of spatial metrics.

5 Validation of Land Use Cover Change Modelling Exercises

Validating a LUCC modelling exercise is a complex task. In this case, we are not validating a single LUC map or a series of LUC maps, but a model application made up of multiple inputs, which interact to deliver new results. When validating LUCC modelling exercises, users tend to focus exclusively on the validation of the model's hard maps, i.e. maps with a categorical legend similar to the input LUC maps (Camacho Olmedo et al. 2018). These hard maps are the main final output of any modelling exercise, but not the only one. To properly validate a LUCC modelling exercise we

should focus not only on the scenario generated by the model, but also on the other outputs and inputs.

Given the nature of this book, we will be dealing exclusively with the validation of LUC maps associated with LUCC modelling exercises: input LUC maps, output soft LUC maps and output hard LUC maps. Users must bear in mind that other sources of data can be used in LUCC modelling exercises and can be validated via complementary methods.

Modellers can begin a modelling exercise by evaluating the uncertainty of the input LUC maps used in the model and their changes according to the guidelines set out in Sects. 3 and 4 above. This is because the quality of the input LUC maps can have a significant effect on the performance of the model. When setting up LUCC models, it is essential to understand the changes that take place in the set of input and reference maps. An assessment of the uncertainty of these LUC changes is therefore vital for determining and characterizing the uncertainty of the LUCC modelling exercise.

In the following subsections, we present the validation tools for output LUC maps, i.e. the products obtained by the model, differentiating between soft and hard LUC maps.

5.1 Soft LUC Maps

Soft LUC maps, also referred to as suitability, change potential or change probability maps, are produced by the model to express the propensity to change over space, that is, the potential of each pixel to become a specific category in the future (Camacho Olmedo et al. 2018). Modellers can assess the internal behaviour and coherence of the model they are building by comparing the model's soft maps with the maps of simulated changes. They can also find out to what extent the changes simulated by the model coincide with the areas of highest potential in the respective maps for each modelled category. In addition, they can compare the soft maps obtained by different models and assess their level of agreement.

Soft LUC maps are usually validated against a reference map of changes ($t_0 - t_1$), and there are various methods for carrying out this analysis (see chapter "Validation of Soft Maps Produced by a Land Use Cover Change Model"). The Pearson and Spearman **correlation** (see Sect. 1 in chapter "Validation of Soft Maps Produced by a Land Use Cover Change Model") is appropriate for a quick assessment of the soft map, by computing it against the map of observed change (Bonham-Carter 1994; Camacho Olmedo et al. 2013). The **Receiver Operating Characteristic (ROC)** (see Sect. 2 in chapter "Validation of Soft Maps Produced by a Land Use Cover Change Model") (Pontius and Parmentier 2014) is used to assess soft maps by comparing them with the observed binary event map. A highly predictive model

produces a soft map in which the highly ranked values coincide with the actual event. In soft maps, the **Difference in Potential (DiP)** proposed by Eastman et al. (2005) (see Sect. 3 in chapter "Validation of Soft Maps Produced by a Land Use Cover Change Model") compares the relative weight of values allocated to changed areas, in other words the difference between the mean potential in the areas of change and the mean potential in the areas of no change (Pérez-Vega et al. 2012).

In short, the previous three methods evaluate the relationship between the observed changed area and the soft LUC map, assuming that a good model output allocates the highest change probability values to the areas that did actually change, and the lowest change probability values to the areas that did not change. Unlike the previous methods, the **total uncertainty, quantity uncertainty and allocation uncertainty** indices (see Sect. 4 in chapter "Validation of Soft Maps Produced by a Land Use Cover Change Model") (Krüger and Lakes 2016) are not calculated against a reference map of changes, and instead estimate uncertainty by adding together misses and false alarms based on soft prediction score levels.

In addition to these specific indices for soft LUC maps, validation can also be conducted after reclassifying the original soft maps, so transforming continuous, ranked maps (soft) into categorical maps (hard) (see Sects. 1 and 2 in chapter "Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps"). This preliminary step enables most of the validation tools presented in this chapter to be applied for this purpose.

5.2 Hard LUC Maps

The second output obtained by the model is the hard LUC map. Also known as prospective LUC maps, these are simulated LUC maps with an identical categorical legend to the input LUC maps (Camacho Olmedo et al. 2018). The hard maps must be validated in order to understand more about the behaviour of the model and how well it simulates changes. These maps provide a clearer picture of the characteristics of the simulated changes and how they resemble our reference data.

5.2.1 Single LUC Maps

The simulation (T_1) can only be validated against a single LUC map (t_1) if both maps correspond to the same year. This will also enable users to apply the panoply of tools presented in Sect. 3. The **Accuracy assessment statistics**, computed either globally (see Sect. 5 in chapter "Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps") or locally (see Sect. 1 in chapter "Geographically Weighted Methods to Validate Land Use Cover Maps")

could also be applied to validate the simulation against other LUC data such as ground points.

In addition to this generic list of tools, some metrics are specifically used for validating the hard LUC maps obtained from LUCCM exercises. **Allocation distance error** (see Sect. 3 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) (Paegelow et al. 2014) measures the relevance of simulation errors by computing the distance between a false positive (commission) and the closest object in the reference map, considering the minimum distance or the centroids of the area in question.

5.2.2 LUC Maps Series/LUC Changes

The most appropriate, most complete validation procedure for hard maps must include three different maps: the simulation (T_1), a reference LUC map for the same year (t_1) and the base map over which the simulation is executed (t_0). In other words, if our modelling exercise starts in the year 2010, we will need a base map for 2010 to establish the initial landscape on which the simulation will be calculated. Then, if we run a simulation for the year 2020, we will also need a reference map for 2020 in order to be able to compare how well our model simulates change. By comparing the simulation and the reference map we can understand to what extent the simulation matches the reference data. The changes that take place on the reference map and the simulation can be extracted by comparing them with the base map. The changes extracted from the two maps can then be compared so as to find out how well the simulated changes agree with the changes that took place on the reference maps.

There are many tools for validating and understanding the errors and uncertainties of simulated changes. In fact, all the methods and strategies explained in Sect. 4 can be applied in LUCC modelling. In this case, however, the main purpose is to achieve the best possible fit between the results of the model and the reference data.

The majority of metrics are obtained from the **cross-tabulation matrix** (see Sect. 1 in chapter “Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps”). The cross-tabulation matrix offers a detailed picture of the changes that were simulated (by cross-tabulating the simulation with the base map), the changes we used as a reference (by cross-tabulating the reference map with the base map) and the agreement and disagreement between the simulation and the reference map (by cross-tabulating the simulation with the reference map). The cross-tabulation matrix can also be used to summarize simulated and reference change in a series covering the main processes of change (artificialization, deforestation...). This enables us to quickly identify the changes that have taken place in our simulation and to spot potential change patterns that do not make sense.

Cross tabulation can be carried out at **multiple resolutions** (see Sect. 2 in chapter “Basic and Multiple-Resolution Cross Tabulation to Validate Land Use Cover Maps”) (the original and coarser ones), to find out at which resolution there is the greatest agreement. Sometimes, the simulation and the reference landscape do not agree on the details but show high consistency at coarser scales. This implies that the model is unable to simulate the precise location of the changes, but it does simulate the main patterns of change correctly.

Different metrics have been proposed for summarizing the agreement between the simulation and the reference maps that the cross-tabulation matrix shows in raw (see chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”). The **Areal and spatial agreement metrics** (see Sect. 2 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) could be applied to summarize the agreement between two maps of changes, the simulated and the reference change maps, overall or per category. **Kappa** (see Sect. 3 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) also summarizes the overall agreement between two maps. However, it has been widely criticized because it assesses the similarity between the simulation and the reference map, but does not distinguish between the areas that change between the two dates and those that do not. Therefore, in maps that simulate permanence correctly, the Kappa metric will be high. Accordingly, we only recommend Kappa for assessing how well permanence is simulated, and it should not be used for a detailed assessment of the accuracy of simulated changes. The Kappa Simulation proposed by Van Vliet et al. (2011) takes the standard Kappa flaws regarding LUCC modelling into account. It focuses on the agreement between the changes in the simulation and the changes in the reference map with regard to the initial map used as a base for the simulation.

The **Agreement between maps at global and stratum level** (see Sect. 4 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”) analysis can assess for a specific LUC transition, for example, whether the agreement between an observed (reference map) and a simulated transition varies or not for several distance classes resulting from a driver (e.g. distance to roads). Other metrics, such as **change statistics** (see Sect. 1 in chapter “Metrics Based on a Cross-Tabulation Matrix to Validate Land Use Cover Maps”), are widely used for characterizing the simulated changes, providing extra information that may be helpful for their validation.

Pontius proposes several metrics for validating simulated change (see chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”). Some of them can also be used to validate time series of

LUC maps and were therefore described in Sect. 2. The **LUC budget** (see Sect. 2 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) technique helps users to understand the changes that take place between the simulation and the base map and between the reference and the base maps. This tool calculates the gross and net changes, overall and per category, as well as the category swaps, in both the simulated and the reference landscapes. This enables us to assess in detail whether the changes we simulated are similar to the changes that take place on the reference maps and follow the same trends.

Quantity & allocation disagreement (see Sect. 3 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) differentiates, at an overall level and per category, between the (dis)agreement between two maps in terms of the proportion of the map occupied by each category (quantities) and the (dis)agreement due to the allocation of the categories in the same/different places on the map (allocation). It is therefore useful for assessing how much of the disagreement is due to the way the model simulates quantities and how much is due to its incorrect allocation of categories. By making the analysis at the category level, it also allows us to assess where (i.e. in which categories) the errors and uncertainties arise.

If a chronological series of simulations (more than two-time points) is available, **Incidents and States** (see Sect. 5 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) may also be employed. This metric helps identify pixels that follow illogical transition patterns, with changes at successive time intervals between the same pair of categories (e.g. from agricultural to urban fabric and then back to agricultural).

Intensity analysis (see Sect. 6 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) compares the different intensities of change per category in simulations and reference maps over at least three points in time. In this way we can assess whether our model correctly simulated the change trend displayed by the reference data. The **flow matrix** (see Sect. 7 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) could also be applied to validate simulated changes in a generic way, assessing the stability and instability of the real and simulated changes over time.

The **Null model** (Pontius and Malanson 2005) (see Sect. 1 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) compares the agreement between the base map for the simulation and the reference map versus the agreement between the simulation and the reference map. If the former

is higher than the latter, our modelling exercise could be judged to have performed poorly, in that the accuracy of the obtained simulation is lower than that for a reference map in which no change takes place. This assertion may be clarified by using other validation tools to obtain a clearer understanding of the logic and pattern of the simulated change. The null model is also a valuable tool for evaluating how well the model simulates permanence.

The **Figure of Merit** (Pontius et al. 2008) and **complementary Producer’s and User’s accuracy**, (see Sect. 4 in chapter “Pontius Jr. Methods Based on a Cross Tabulation Matrix to Validate Land Use Cover Maps”) also measure the agreement between simulated changes and changes in the reference map. The Figure of Merit technique is recommended when trying to assess the model’s ability to correctly simulate change. The different components of the Figure of Merit can be used to discover whether the model estimates more or less change than the reference map. It is also highly recommended for evaluating the congruence of model outputs and model robustness. This is a form of validation that evaluates the agreement between simulations obtained using different models or using the same model parametrized in different ways (Paegelow et al. 2014; Camacho Olmedo et al. 2015).

None of the above tools assesses the accuracy of the pattern of LUC change in the simulation. This aspect is important because even if the quantities simulated are wrong and the categories are not allocated in the same positions as in the reference maps, the pattern of LUC change may have been simulated correctly. Pattern can be validated using **Spatial metrics** (see Sect. 1 in chapter “Spatial Metrics to Validate Land Use Cover Maps”) and the **Map Curves** (see Sect. 1 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) method, which compare the pattern of the simulation with the pattern of the reference landscape.

Spatial metrics characterize many different elements of the landscape: fragmentation, shape complexity, category proportions, diversity.... They can be calculated specifically for the simulated and reference changes, so allowing users to identify the specific pattern characteristics of the features that changed during the simulation period. In this way we can understand the size and shape of the simulated changes, inferring from this information how logical or uncertain they may be.

The MapCurves method gives a summary figure for the pattern agreement between two maps, and is therefore much easier to interpret. However, it does not provide all the complex detail that can be revealed by applying the different spatial metrics.

We can also analyse the changes that take place on the borders of existing patches and the changes that result in the appearance of new patches. This distinction may be useful for identifying errors or inconsistencies. The detection of

change on pattern borders (see Sect. 2 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) enables us to evaluate and identify errors in the simulations, which may be due to different parameters being applied in the model allocation procedure, such as, for example, the use of a contiguity filter. The **Allocation distance error** (see Sect. 3 in chapter “Advanced Pattern Analysis to Validate Land Use Cover Maps”) calculates the distance between wrongly simulated patches and reference patches, so as to gain a better picture of how well the patches are simulated. In this sense, a model that wrongly allocates change close to areas that actually change on the ground would be considered to have performed better than a model that allocates them further away.

References

- Aldwaik SZ, Pontius RG (2012) Intensity analysis to unify measurements of size and stationarity of land changes by interval, category and transition. *Landsc Urban Plan* 106:103–114
- Anderson JR (1971) Land-use classification schemes used in selected recent geographic applications of remote sensing. *Photogramm Eng Remote Sens* 37:379–387
- Bonham-Carter GF (1994) Tools for map analysis: map pairs. In: Bonham-Carter GF (ed) *Geographic information systems for geoscientists*, pp 221–266. Pergamon. ISBN 9780080418674. <https://doi.org/10.1016/B978-0-08-041867-4.50013-8>
- Botequilha Leitao A, Miller J, Ahern J, McGarigal K (2006) *Measuring landscapes: a planner's handbook*. Island Press, Washington, Covelo, London
- Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geogr Anal* 28(4):281–298. <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
- Camacho Olmedo MT, Paegelow M, Mas J-F (2013) Interest in intermediate soft-classified maps in land change model validation: suitability versus transition potential. *Int J Geogr Inf Sci* 27(12):2343–2361
- Camacho Olmedo MT, Pontius RG Jr, Paegelow M, Mas JF (2015) Comparison of simulation models in terms of quantity and allocation of land change. *Environ Model Softw* 69(2015):214–221. <https://doi.org/10.1016/j.envsoft.2015.03.003>
- Camacho Olmedo MT., Mas JF, Paegelow M (2018) The Simulation Stage in LUCC modeling. In: Camacho Olmedo M, Paegelow M, Mas JF, Escobar F (eds) *Geomatic approaches for modeling land change scenarios*. Lecture Notes in Geoinformation and Cartography. Springer, Cham, pp 27–51. Publisher Name Springer, Cham Print ISBN 978-3-319-60800-6 Online ISBN 978-3-319-60801-3 eBook Packages Earth and Environmental Science. https://doi.org/10.1007/978-3-319-60801-3_3
- Comber AJ (2013) Geographically weighted methods for estimating local surfaces of overall, user and producer accuracies. *Remote Sens Lett* 4(4):373–380. <https://doi.org/10.1080/2150704X.2012.736694>
- Eastman JR, Van Fossen ME, Solarzano LA (2005) Transition potential modelling for land cover change. In: Maguire D, Goodchild M, Batty M (eds), *GIS, Spatial analysis and modeling*. ESRI Press, Redlands, California
- FAO (1995) *Forest resources assessment 1990*. Global synthesis. FAO, Rome
- Foody GM (2002) Status of land cover classification accuracy assessment. *Remote Sensing Environ* 80(1):185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Foody GM (2008) Harshness in image classification accuracy assessment. *Int J Remote Sens* 29(11):3137–3158
- Forman RTT (1995) *Land mosaics: the ecology of landscapes and regions*. Cambridge University Press, Cambridge, United Kingdom
- García-Álvarez D, Camacho Olmedo MT (2017) Changes in the methodology used in the production of the Spanish CORINE: Uncertainty analysis of the new maps. *Int J Appl Earth Obs Geoinf* 63:55–67. <https://doi.org/10.1016/j.jag.2017.07.001>
- Gwet K (2002) Kappa statistic is not satisfactory for assessing the extent of agreement between raters. Series: statistical methods for inter-rater reliability assessment
- Gómez C, White JC, Wulder MA (2016) Optical remotely sensed time series data for land cover classification: A review. *ISPRS J Photogramm Remote Sens* 116:55–72. <https://doi.org/10.1016/j.isprsjprs.2016.03.008>
- Hao G, Gen-Suo J (2014) Assessing MODIS land cover products over China with probability of interannual change. *Atmosph Ocean Sci Lett* 7(6):564–570. <https://doi.org/10.1080/16742834.2014.11447225>
- Hargrove WW, Hoffman FM, Hessburg PF (2006) Mapcurves: a quantitative method for comparing categorical maps. *J Geogr Syst* 8:187–208. <https://doi.org/10.1007/s10109-006-0025-x>
- Islam S, Zhang M, Yang H, Ma M (2019) Assessing inconsistency in global land cover products and synthesis of studies on land use and land cover dynamics during 2001 to 2017 in the southeastern region of Bangladesh. *J Appl Remote Sens* 13(04):1. <https://doi.org/10.1117/1.JRS.13.048501>
- Kerr GHG, Fischer C, Reulke R (2015) Reliability assessment for remote sensing data: beyond Cohen's kappa. In: *International geoscience and remote sensing symposium (IGARSS)*, pp 4995–4998. <https://doi.org/10.1109/IGARSS.2015.7326954>
- Krüger C, Lakes T (2016) Revealing uncertainties in land change modeling using probabilities. *Trans GIS* 20(4):526–546. <https://doi.org/10.1111/tgis.12161>
- Olofsson P, Foody GM, Stehman SV, Woodcock CE (2013) Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens Environ* 129(2013):122–131. <https://doi.org/10.1016/j.rse.2012.10.031>
- Olofsson P, Foody GM, Herold M et al (2014) Good practices for estimating area and assessing accuracy of land change. *Remote Sens Environ* 148:42–57. <https://doi.org/10.1016/j.rse.2014.02.015>
- Paegelow M, Camacho Olmedo MT, Mas JF and Houet T (2014) Benchmarking of LUCC modelling tools by various validation techniques and error analysis. *Cybergeo Eur J Geogr [En ligne] Systèmes, Modélisation, Géostatistiques*, document 701, mis en ligne le 22 décembre 2014. ISSN: 1278-3366. CNRS-UMR Géographie-cités 8504. <https://doi.org/10.4000/cybergeo.26610>
- Paegelow M, Camacho Olmedo MT, Mas JF (2018) Techniques for the validation of LUCC modeling outputs. In: Camacho Olmedo M, Paegelow M, Mas JF, Escobar F (eds) *Geomatic approaches for modeling land change scenarios*. Lecture Notes in Geoinformation and Cartography. Springer, Cham, pp 53–80. Publisher Name Springer, Cham Print ISBN 978-3-319-60800-6 Online ISBN 978-3-319-60801-3 eBook Packages Earth and Environmental Science
- Pérez-Hoyos A, Udías A, Rembold F (2020) Integrating multiple land cover maps through a multi-criteria analysis to improve agricultural monitoring in Africa. *Int J Appl Earth Obs Geoinf* 88:102064. <https://doi.org/10.1016/j.jag.2020.102064>
- Pérez-Vega A, Mas JF, Ligmann-Zielinska A (2012) Comparing two approaches to land use/cover change modeling and their

- implications for the assessment of biodiversity loss in a deciduous tropical forest. *Environ Model Softw* 29(1):11–23
- Pontius RG Jr, Malanson J (2005) Comparison of the structure and accuracy of two land change models. *Int J Geogr Inf Sci* 19:243–265. <https://doi.org/10.1080/13658810410001713434>
- Pontius RG Jr, Millones M (2011) Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int J Remote Sens* 32:4407–4429. <https://doi.org/10.1080/01431161.2011.552923>
- Pontius RG Jr, Parmentier B (2014) Recommendations for using the relative operating characteristic (ROC). *Landscape Ecol* 29(3):367–382
- Pontius RG Jr, Shusas E, McEachern M (2004) Detecting important categorical land changes while accounting for persistence. *Agr Ecosyst Environ* 101:251–268
- Pontius Jr RG, Boersma W, Castella JC, Clarke K, de Nijs T, Dietzel C, Duan Z, Fotsing E, Goldstein N, Kok K, Koomen E, Lippitt CD, McConnell W, MohdSood A, Pijanowski B, Pithadia S, Sweeney S, Trung TN, Veldkamp AT, Verburg PH (2008) Comparing input, output, and validation maps for several models of land change. *Ann Reg Sci* 42(1):11e47
- Pontius RG Jr, Krithivasan R, Sauls L et al (2017) Methods to summarize change among land categories across time intervals. *J Land Use Sci* 12:218–230. <https://doi.org/10.1080/1747423X.2017.1338768>
- Pouliot D, Latifovic R (2013) Accuracy assessment of annual land cover time series derived from change-based updating. In: Proceedings of the 7th International workshop on the analysis of multi-temporal remote sensing images: “our dynamic environment” (MultiTemp 2013), pp 1–3. <https://doi.org/10.1109/Multi-Temp.2013.6866005>
- Puyravaud J-P (2003) Standardizing the calculation of the annual rate of deforestation. *For Ecol Manage* 177(1–3):593–596
- Runfola DSM, Pontius RG (2013) Measuring the temporal instability of land change using the Flow matrix. *Int J Geogr Inf Sci* 27:1696–1716. <https://doi.org/10.1080/13658816.2013.792344>
- Stehman SV, Czaplewski RL (1998) Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote Sens Environ* 64:331–344
- Stehman SV (1997) Selecting and interpreting measures of thematic classification accuracy. *Remote Sens Environ* 62(1):77–89. [https://doi.org/10.1016/S0034-4257\(97\)00083-7](https://doi.org/10.1016/S0034-4257(97)00083-7)
- Stehman SV (1999) Basic probability sampling designs for thematic map accuracy assessment. *Int J Remote Sens* 20(12):2423–2441. <https://doi.org/10.1080/014311699212100>
- Strahler AH, Boschetti L, Foody GM, Friedl MA, Hansen MC, Herold M, Mayaux P, Morissette JT, Stehman SV, Woodcock CE (2006) Global land cover validation: recommendations for evaluation and accuracy assessment of global land cover maps. Office for official publications of the European communities. GOCF-GOLD Report No 25. Luxembourg
- Van Vliet J, Bregt AK, Hagen-Zanker A (2011) Revisiting Kappa to account for change in the accuracy assessment of land-use change models. *Ecol Model* 222(8):1367–1375. <https://doi.org/10.1016/j.ecolmodel.2011.01.017>
- Woodcock CE, Gopal S (2000) Fuzzy set theory and thematic maps: accuracy assessment and area estimation. *Int J Geogr Inf Sci* 14(2):153–172. <https://doi.org/10.1080/136588100240895>
- Wulder MA, Franklin SE, White JC, Linke J, Magnussen S (2006) An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data. *Int J Remote Sens* 27(4):663–683. <https://doi.org/10.1080/0143116050185284>
- Yang Y, Xiao P, Feng X, Li H (2017) Accuracy assessment of seven global land cover datasets over China. *ISPRS J Photogramm Remote Sens* 125:156–173. <https://doi.org/10.1016/j.isprsjprs.2017.01.016>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.



The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.