

Chapter 15

Uncertainty challenge in Geospatial analysis. An approximation from the Land Use Cover Change Modelling perspective

David García Álvarez¹, Hedwig Van Delden², María Teresa Camacho Olmedo¹, Martin Paegelow³

¹Departamento de Análisis Geográfico Regional y Geografía Física, Universidad de Granada, Granada, Spain dagaral@ugr.es camacho@ugr.es

²Research Institute for Knowledge Systems, Maastricht, The Netherlands hvdelden@riks.nl

³GEODE, University of Toulouse Jean Jaurés, Toulouse, France; Département de Géographie, Aménagement et Environment, Université de Toulouse Jean Jaurés, Toulouse, France paegelow@univ-tlse2.fr

Abstract

All data and geospatial analysis come with uncertainty. Although its importance has been widely recognized, uncertainty issues are still not correctly addressed in most of the current geospatial research. This chapter aims to provide an overview of the concepts, sources and tools to manage the uncertainty in geospatial analysis. To this end, we intend to increase the awareness about the importance of uncertainty for all geospatial data and analyses. Due to time and chapter length considerations, we address this topic from the Land Use Cover Change Modelling perspective.

15.1 Introduction

Geospatial data and analyses have been gaining an increasing importance in the last decades, achieving a prominent position in the study, management and solution of many of the problems that our society copes with. Geospatial analysis is used to study many current and future challenges of our society, such as climate change, urban sprawl and environmental degradation. However, there are still unresolved questions about the limits of those studies and, specifically, about the uncertainty that their conclusions convey.

Although uncertainty has always been present in any type of geographical research, it is more relevant nowadays because of the wider use of geospatial technologies and the false idea of precision that this quantitative branch of geography embraces. Geography is about synthesis, being the maps the main language through which geographical knowledge is spread. They, as an abstraction of the real world, always come with uncertainty. We can therefore trace this uncertainty back to the beginning of mapping and geography as a knowledge of synthesis. Every study and analysis, even the manually ones made before the emergence of geospatial technologies, is therefore equally uncertain.

Despite of the importance of uncertainty in any geospatial research, especially after the advent of geospatial technologies, little attention has been paid to all sources of uncertainty that particular geospatial studies convey. They are important and affect the meaningfulness of the results user and audience get.

Users are still not aware of the uncertainties associated to the use of geospatial tools and methods and how their decisions affect those. Neither is the audience nor the agents, for whom uncertainty acts as a barrier for the adoption of the conclusions from any geospatial analysis in their decisions. That is why more and better information about uncertainty is needed. In this regard, Goodchild (1991) and Hunter (1999) laid out an uncertainty research agenda to achieve better methods for uncertainty management and hence a better understanding of the problem. This agenda has not been met yet. In addition, despite of the increasing awareness and research about uncertainty, the uncertainty recognition is not widely spread across most of the geospatial studies and analyses. Accordingly, we present here the uncertainty issue as one of the key geospatial challenges of the twenty-first century.

To remark the importance of this topic and make users aware about the extent of the issue, this chapter gives a general overview about concepts, sources and methods to deal with uncertainty in geospatial analysis. We address this topic from a Land Use Cover Change Modelling (LUCCM) perspective. LUCCM integrates many geospatial tools, processes and data in just one wide analysis. Accordingly, this perspective covers most of the uncertainty issues that any user can face.

15.2 What does uncertainty mean?

There are many definitions of uncertainty, which range from generic approaches to those that focus on specific fields of knowledge. In general, all authors agree in referring to the difference between the perfect representation of any feature or process and how it is really addressed through data and geospatial tools when talking about uncertainty. Given this difference, authors talk about doubt, reliability, lack of knowledge, degree of distrust, etc to define uncertainty as the user perception about how that difference affects their studies. In other cases, uncertainty also refers to the confidence that an agent has when using geospatial data to make decisions. All in all, in simple terms, uncertainty can be defined as the lack or the degree of

certainty about any data or geospatial analysis due to the difference between reality and its representation through geospatial data or tools.

Many concepts arise when talking about uncertainty, like ignorance, error, accuracy, vagueness or ambiguity. They refer to specific aspects of uncertainty or issues that this term does not cover. Their definition is therefore essential to clarify the conceptual framework of uncertainty, which we propose in Figure 15.1.

Whereas uncertainty implies that the user is aware about the limitation of his knowledge, ignorance exists when there is not such awareness. It is therefore related to the unknown unknowns, that is, those aspects that we do not even know we do not know (Recker 2015).

Error is objective and, hence, quantitative. It is a measure of the distance between reality and our representation. Accordingly, it just refers to those features which can be measured and quantified. For land cover data, as a forest density map, error could be measured from the difference between the forest density in a pixel and the real forest density measured on the ground. Accuracy is the closeness of observations to the truth (Kemp 2008). As an ideal, complete truth is usually unachievable. The data that we consider the truth comes with uncertainty, which the accuracy assessment, unlike the uncertainty analysis, does not consider. When assessing a simulated map to a reference map for the same area, we are checking the accuracy of our simulation. That reference map is not completely certain and, therefore, contains some uncertainty that also transfers to the accuracy assessment analysis.

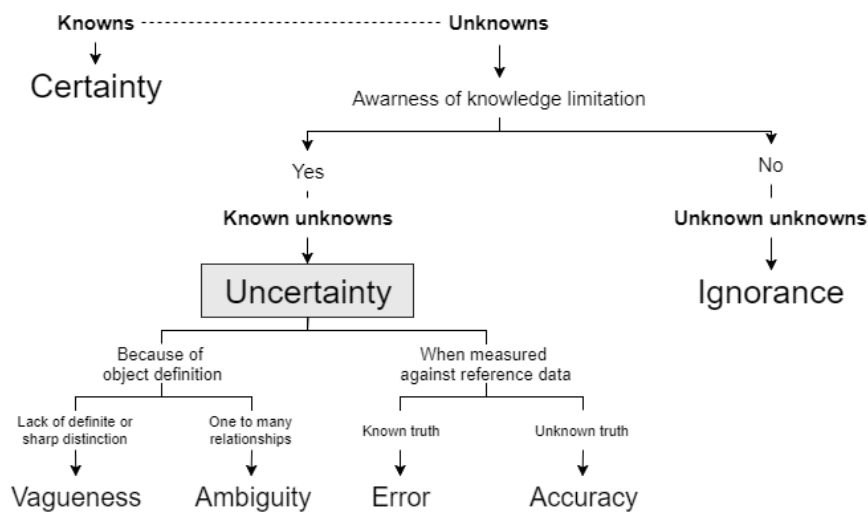


Fig. 15.1. Conceptual chart of uncertainty and related concepts. Source: partially based on Klir and Wierman (1999) and Gómez Delgado and Barredo (2006)

On the other hand, vagueness and ambiguity are qualitative types of uncertainty. They refer to the certainty at which objects or phenomena are delimited. Vagueness refers to the poor or precise definition of classes or objects, whereas ambiguity

arises when the same object can be classified as part of different groups or assets. The former is more about the object definition and ambiguity about the object classification. For land use data, vagueness arises when we are not completely sure about what the label of a class, as urban, covers. Is a village where most of the land is dedicated to agricultural uses urban or not? On the other hand, when upscaling a land use map, ambiguity arises when we are not sure if a landscape previously classified as several patches of cropland and tree cover is then agroforestry or any other category.

15.3 The three dimensions of uncertainty

Following the proposal made by Refsgaard et al. (2013), we differentiate three dimensions of uncertainty as follows: source, level and nature (Fig. 15.2).

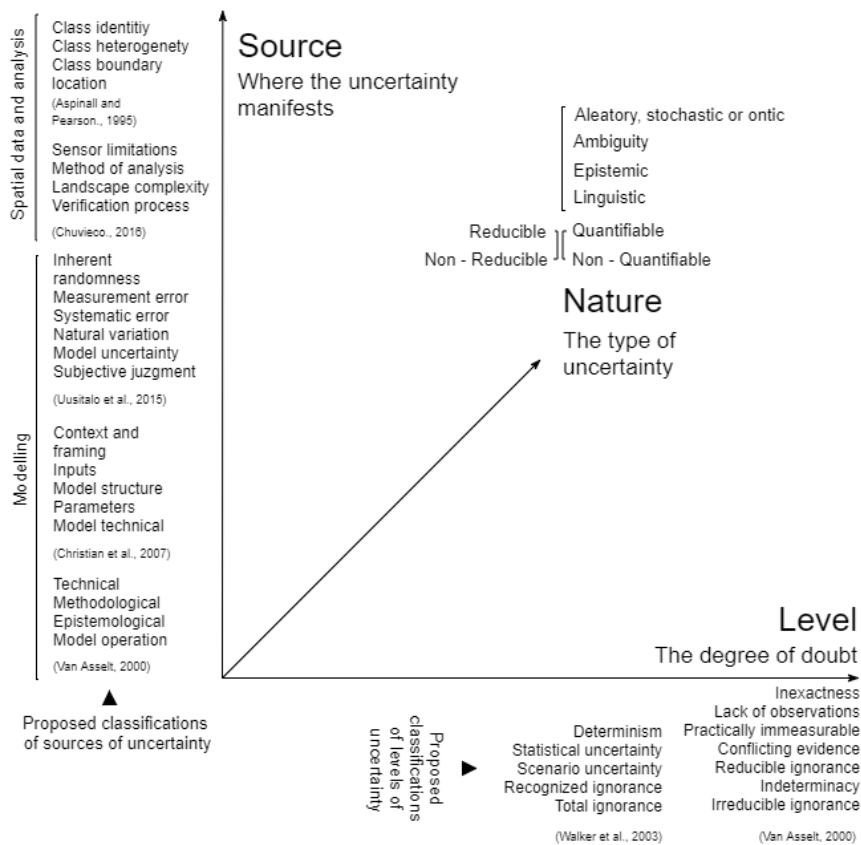


Fig. 15.2. The three dimensions of uncertainty. Source: based on Walker et al. (2003)

- The source is the dimension which gives information where the uncertainty manifests itself. Several studies have proposed a generic classification of sources of uncertainty for modelling environments (Van Asselt 2000; Walker et al. 2003; Refsgaard et al. 2007, 2013; Klein Goldewijk and Verburg 2013; Uusitalo et al. 2015) or even for categorical maps (Aspinall and Pearson 1995) and spatial data obtained through remote sensing analysis (Congalton et al. 1991; Chuvieco 2016).
- The level of uncertainty is the degree of doubt, reliability or lack of knowledge. It moves from complete knowledge and, consequently, absence of uncertainty, to total ignorance.
- The nature of uncertainty is the type of uncertainty we are working with. Literature usually distinguishes to this end between uncertainty due to imperfect knowledge and uncertainty because of natural phenomena variability, being the last one identified as knowledge or epistemic uncertainty and the first one as aleatory, stochastic or ontic uncertainty. Ascough et al. (2008) distinguish different components of the epistemic uncertainty: natural, human, institutional and technological, depending on the source of the variability. Uusitalo et al. (2015) added to the previous two the linguistic uncertainty, which arises from language issues. Refsgaard et al. (2013) considered the ambiguity as a third nature of uncertainty. It refers to the different possible understanding of the same system. Finally, other distinctions exist that differentiate between reducible and non-reducible uncertainties and those measurable and, therefore, quantifiable, and those which are not.

15.4 Sources of Uncertainty in Geospatial Analysis. An approximation from the Land Use Cover Change Modelling perspective

Many sources of uncertainty can be pointed out when studying uncertainty in LUCCM or any geospatial analysis. Accordingly, several papers have tried to propose a classification, as pointed out in the previous section and reviewed by Matott et al. (2009). However, those classifications, as the one proposed here (Fig. 15.3), are just theoretical frameworks aiding in the comprehension and management of uncertainty. The user must be aware of the complex interactions between all sources of uncertainty and the difficulty of individualizing any of them. All together, they are known as output uncertainty (Refsgaard et al. 2007) or the uncertainty cascade (Refsgaard et al. 2013), that is, the combination of all uncertainties that the results from a geospatial analysis convey.

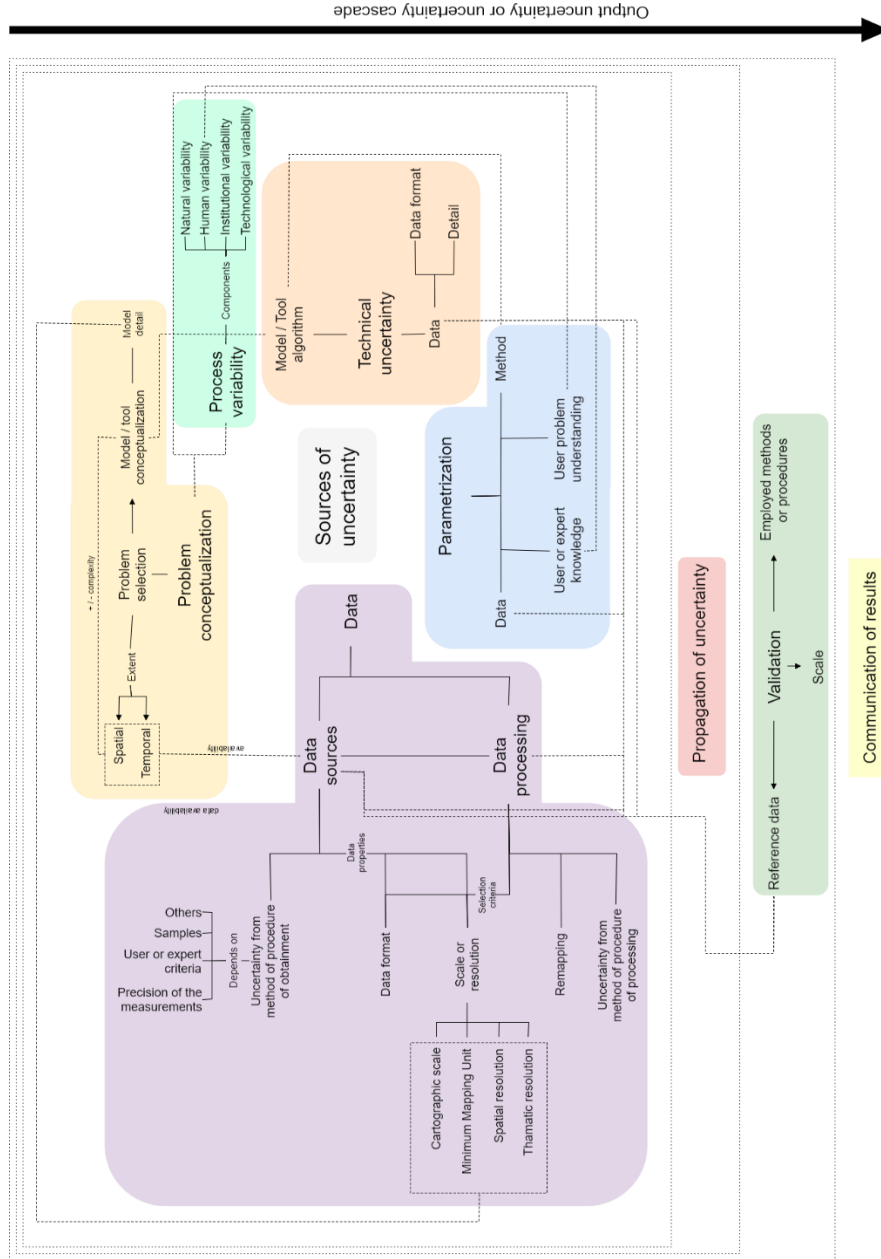


Fig. 15.3. Author's proposal of sources of uncertainty in Geospatial Analysis from a LUCCM perspective. It represents the main sources of uncertainty as well as the interactions between them (dotted lines)

15.4.1 Problem conceptualization

The problem conceptualization can be considered as one of the most important sources of uncertainty, since it comes from the very initial decisions of any analysis, which will shape the study to be carried out. Geospatial data and analysis, as LUCCM, means simplifying the real world and phenomena. When doing this, the user conceptualizes them. The coherence of that conceptualization and how it fits with the problem to be studied or resolved, determine the goodness of the managed concepts and the level of uncertainty that the study conveys.

The very first decision comes from the selection of the problem to be studied. Problems are usually complex and related to other processes. Accordingly, their selection is usually difficult and, often, not very well reasoned. There are no widespread methods to deal with this issue, though Dunn (2001) proposed a context validation framework which can serve as basis to this end. For modelling environments, several papers have addressed, as part of general guides for setting up models, these conceptual questions, pointing out, among other recommendations, the need of engaging agents and final users in those decisions (Van Delden et al. 2011; Elsayah et al. 2017).

Once the problem to be analysed is chosen, it is necessary to define its boundaries. These are not usually real, given the continuation of the earth processes and features. Choosing different temporal or spatial extents (Fig. 15.4) for the same problem may result in different outputs, as showed by papers studying the impact of spatial (extent) (Verburg and Veldkamp 2004) and temporal scale (Pontius Jr and Spencer 2005; Rosa et al. 2015) in LUCCM.

Regarding the temporal extent, the further the time horizon the LUCCM simulation has, the higher the uncertainty becomes, given the unknown variability of real-world processes (epistemic uncertainty). As most LUCC models rely on historic data for their calibration, the temporal resolution of these data as well as the temporal extent of the calibration period, also affects the obtained results (Burnicki et al. 2010). The duration of the calibration period must be long enough to avoid noise and the influence of one-off events in the process comprehension. However, as the aim of the calibration is to simulate future land use developments, it is important to focus on processes that are relevant for simulating future dynamics, as not all past developments would be relevant for this. In some cases, more recent development might be more representative than a long historic period. Analysing and understanding the processes and dynamics of the area under study and placing them in context is therefore essential.

Regarding the spatial extent, every spatial extent must be linked to a conceptualization level of the system under study: when working at continental scale, the system's conceptualization should not be as detailed as a national or regional scales. Therefore, thematic, temporal or spatial resolutions, that is, model detail, should be chosen considering these criteria.

The model conceptualization is then, for LUCCM, the critical step in the problem definition. Model conceptualization means the simplification of the system to be modelled into terms that are understandable, as simple as possible and, especially, possible to be implemented through a software. That conceptualization will be affected by the modeller experience and preferences (Klein Goldewijk and Verburg 2013), but also by the data available and the feasibility of its implementation in a specific algorithm. Accordingly, there is a close link between problem conceptualization uncertainty and the one arising from data and technical issues.

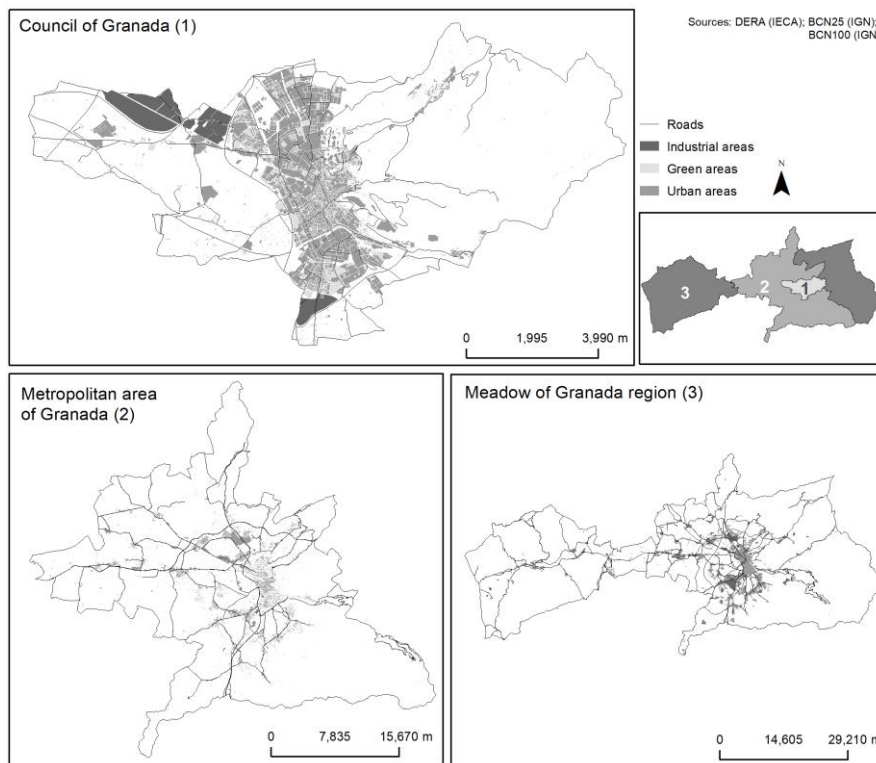


Fig. 15.4. Different spatial extents and approaches for the study of the Granada area (Spain). Every extent matches with a specific level of spatial and thematic detail. The extents refer to different type of boundaries (political, planning and agricultural management)

When conceptualizing the model, one must consider that more complexity does not always mean less uncertainty, but can also lead to the opposite (Van Asselt 2000). Complex models are not easy to understand, making the engagement of agents and final users difficult. Complex models also require many parameters and complex algorithms, making the technical and parameter uncertainties very high.

As previously stated, when managing the complexity, the user must keep in mind the thematic, temporal and spatial resolution of the data and the model, since they

are the main source of model detail. Finally, the model conceptualization must ensure that all processes and interactions of interest are in the model, that is, the model completeness.

15.4.2 Data sources and processing

A big part of the uncertainty of any analysis comes from data. The utility of the analysis will rely on the uncertainties of these data and how data are used. In addition, depending on the study purpose, data uncertainties will play a major or minor role. LUCCM usually rely on data to set up the starting conditions of the model and to parametrize (calibrate) and validate it. Accordingly, data plays a key role in the obtained results. The importance of the availability and quality of the data will increase with the dependence of the model on data, and therefore will be particularly relevant for data-driven models.

We differentiate between those uncertainties that data comes with, which are unavoidable, and those caused by data processing. Altogether, for LUCCM they are also known as the input uncertainty (Refsgaard et al. 2007).

15.4.2.1 Data sources

All data comes with uncertainty, even in the case of raw data obtained by measurement or surveys. Because it is unavoidable, the user needs to have as much information as possible about it, ergo they can understand the limits that these data put on their analysis. Sometimes accuracy or uncertainty reports are attached to the data. However, this information is uncertain as well, because of the uncertainty of the validation process.

For categorical data, as Land Use Land Cover (LULC) maps, we can broadly distinguish between two main sources of uncertainty: the categorical or thematic uncertainty, that comes from the attribute definition, and the positional uncertainty, which comes from the object position (Castilla and Hay 2007). Numerous causes explain for every case the extent and importance of these sources of uncertainty. Expert knowledge is one main reason in the case of data obtained by photointerpretation. The data format (raster vs vector) is another factor determining the accuracy of data, especially regarding the position of elements. However, the scale or detail of data is maybe the key element explaining the uncertainty. The smaller or coarser the scale, the bigger the abstraction and, therefore, the larger the uncertainty. Notwithstanding, at finer scales, the definition of elements must be more detailed, making the chance of error larger. Including much detail can also make the analysis too complex. The comparison of several datasets at different scales for the same study area (Fig. 15.5) has proven the influence of scale in data uncertainty (Waser and Schwarz 2006; García Martínez et al. 2015).

The main input of LUCC models is LULC maps, from which modellers analyse LULC change. Changes often represent a small portion of the maps, which usually have an overall accuracy around 80% (Pontius Jr. and Lippitt 2006). Consequently, identified changes can be affected by large uncertainties. These uncertainties can propagate to the quantity of change estimation or the model parametrization. Therefore, providing information about the uncertainty of the measured changes would be very relevant (Pontius Jr. and Lippitt 2006), including the distinction between technical and real changes (Verburg et al. 2011), that is between those changes that really happened and those ones that are caused by atmospheric or radiance noise in the case of data obtained by remote sensing classification techniques or by variable photointerpretation criteria in case of data obtained by photointerpretation.

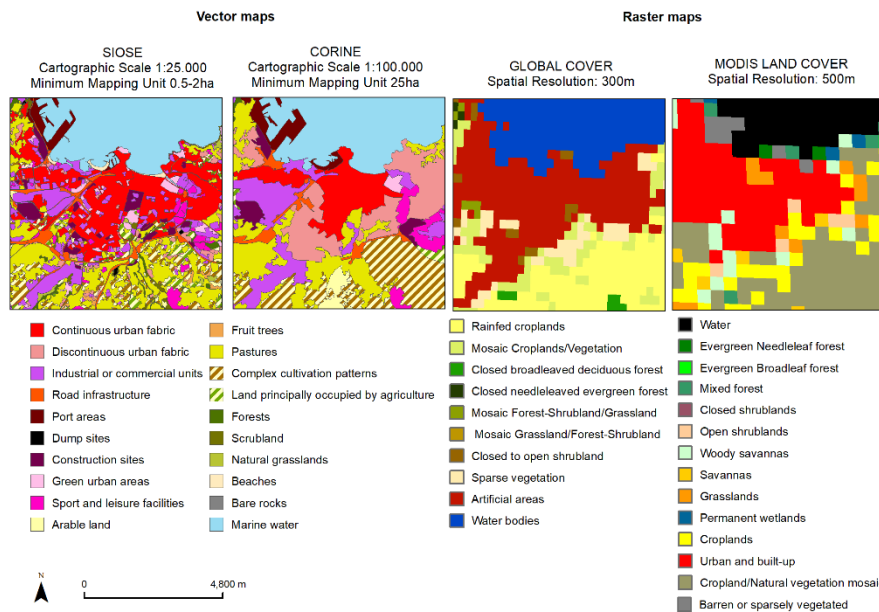


Fig. 15.5. Comparison of different LULC datasets with different formats and at different spatial and thematic resolutions for a test area (Gijón, Spain) in 2005

According to Grinblat et al. (2016), maps obtained through remote sensing classification techniques are far from the requirements of LUCCM, because of the uncertainty of the classification methods. Traditional photointerpretation provides better results, with the uncertainty dependent on the knowledge of the expert. In both cases it is difficult to find consistent temporal series of maps without technical changes that allow the user an adequate analysis and comprehension of LULC changes (Verburg et al. 2011). The recent changes in the method used in the production of CORINE for some countries (García-Álvarez and Camacho Olmedo 2017) are, in this regard, a perfect example of those difficulties.

In addition, most data are usually produced for specific purposes, so their use for other aims, as LUCCM, can introduce new uncertainties. All in all, it would be extremely useful to increase efforts in the uncertainty analysis of data. For LUCCM, there is still a great need of sources suitable to the needs and requirements of this type of analysis.

15.4.2. Data processing

For complex geospatial analysis, as LUCCM, there are strong data requirements, not just in terms of quantity of data, but also in terms of format or resolution. In consequence, it is usually necessary to perform data treatments to fit those data with the characteristics required by the employed tool. As any geospatial analysis, these transformations introduce new sources of uncertainty, which must be considered.

Among the usual transformations performed, we can mention vector to raster and raster to vector conversions (Fig. 15.6) (Congalton 1997) or raster resampling techniques (Fig. 15.7) (Dendoncker et al. 2008). Scale issues, as the thematic (Aldwaik et al. 2015) or spatial resolution (Díaz-Pacheco et al. 2018), and the minimum mapping unit (García-Álvarez 2018), are also key decisions that introduce great variations in data. Regarding the thematic resolution, for LUCCM, some classes, as construction sites, might need to be remapped to fit with the model conceptualization. The way this treatment is made also introduces new sources of uncertainty.

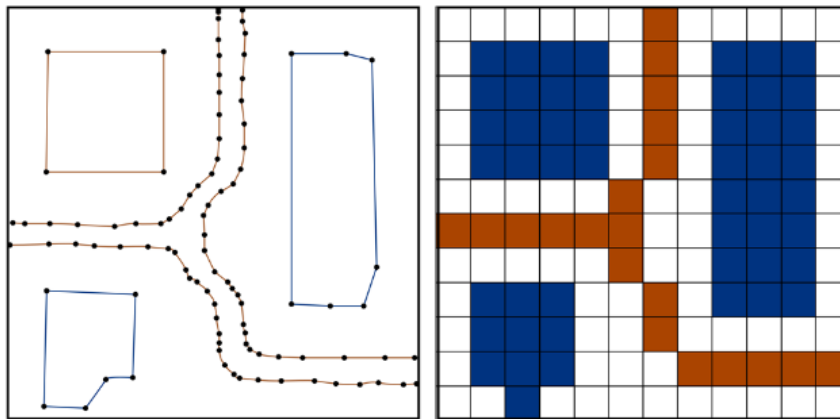


Fig. 15.6. Differences in the representation of the same features under vector (right) and raster (left) formats. Source: taken from Olaya (2014)

Although this processing can be seen as a new source of uncertainty, it also provides a way of dealing with data, hence it becomes more suitable for the intended analysis purpose, aiding in the management of the parameter uncertainty. To this end, several studies have addressed the scale influence in LUCCM (Ménard and Marceau

2005; Conway 2009) or the general effects that data transformations, in the case of driving forces, have in LUCCM (Houet et al. 2015).

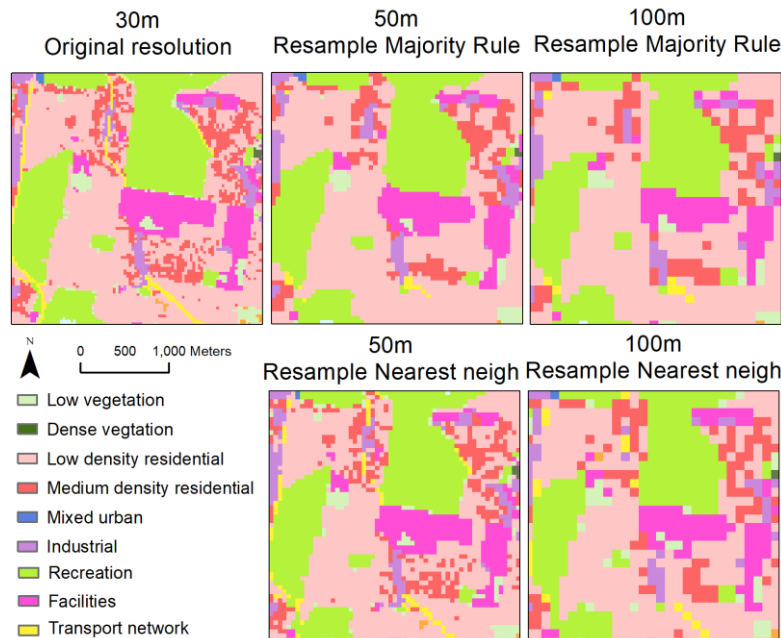


Fig. 15.7. Influence of different resampling methods and spatial resolutions in the pattern and proportions of a LULC map for the Kensington suburb (Sydney)

15.4.3 Technical uncertainty

It refers to the uncertainty that arises from the computer implementation of tools and data, that is, from the translation of concepts to computer entities and tools. It is therefore related to the problem of conceptualization uncertainty, since the possibilities of addressing a problem will be given by the technical limits of the computer.

Regarding data, this uncertainty comes from the format and data detail, taking part of the data uncertainty (Section 15.4.2). There is ample information about pros and cons of different data formats in most GIS manuals, as well as about scale and resolution issues (Quattrochi and Goodchild 1997; Lloyd 2014). For LUCCM, technical uncertainty mostly comes from the model algorithm.

Some model approaches are designed for specific applications, whereas general approaches, as CA_Markov or Metronamica (Fig. 15.8), can be applied to a wide range of problems (Torrens 2011). Whereas the first group provides an adjusted solution for a particular problem, allowing the reduction of the technical uncertainty to the minimum, the second group allows more research about their limits and disadvantages, and about the uncertainty that they convey. That is why some authors

appeal to put the effort in the improvement of these general models (Hewitt et al. 2014).

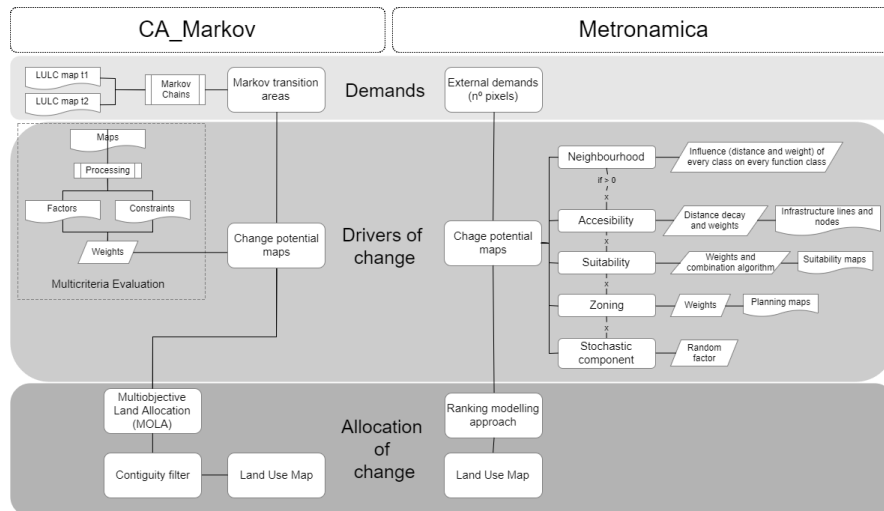


Fig. 15.8. Comparison of the structure of two common LUCCM approaches: CA_Markov and Metronamic

When selecting or developing a computer model, one must consider the desired complexity. Following Lee (1973), the user must strike a balance between model complexity and data. That is, the model must not be more complex than the data on which it is based. Neither should the complexity avoid agent engagement or user comprehension.

Several model approaches can be distinguished (Yeh and Li 2003; National Research Council 2014), which go from data-driven to knowledge-driven models. Data-driven models incorporate statistical or automatic methods for finding the relation between the dependent and independent variables, that is, to determine the model parameters. In the case of expert-driven approaches these parameters are decided by expert judgment. Each method introduces a different type of uncertainty in the model parametrization, as we will see in the next section.

15.4.4 Parametrization

The way a geospatial tool, as a LUC model, is parametrized (i.e. calibrated) determines the utility of the results. Poor parametrizations end in meaningless conclusions. Accordingly, attention must be paid to the uncertainties of this process.

Parameter uncertainty is affected by model structure (technical uncertainty) and data, but also by the problem conceptualization and the user understanding of all

the analyses and concepts. Open or poorly defined systems are more complex to understand and, ergo, more difficult to parametrize. User knowledge and understanding of the system (real and conceptualized) is vital to get a correct parametrization. Notwithstanding, the importance of user knowledge in the parameter uncertainty will depend on the role that they play in the model (technical uncertainty).

In statistical or automatic approaches, most of the uncertainty comes from the method used. Some of them are described as black boxes. The user cannot understand how they find the relation between drivers and changes, which makes the obtained result very uncertain. These approaches are also affected by methodological issues as multicollinearity, autocorrelation, the Modifiable Areal Unit Problem, the Ecological Fallacy Problem or the Category Aggregation Problem (Pontius Jr and Spencer 2005). In addition, if the method needs to be technically parametrized through modeller criteria, uncertainty can be also introduced because of human subjectivity (Jafarnezhad et al. 2012).

In knowledge-driven approaches, uncertainty mostly comes from the user understanding of both the process and the conceptual model. Some authors stated that user knowledge is case-specific and, therefore, uncertain (Botterweg 1995). Statistical or automatic procedures guarantee repeatability over study cases, which is not possible under knowledge-driven approaches. However, practical experience of the authors showed that, when working with a meaningful group of agents, there is a common consensus and, therefore, a way of repeatability.

In both approaches, data uncertainty plays a decisive role. Most parametrizations, at least for LUCCM, are based on historical data and, therefore, they will be as certain as the data are (Batisani and Yarnal 2009). Notwithstanding, whereas data uncertainties can be dealt with through user intervention, this is not possible if the method is completely statistical or automatic.

Scale and resolution can be used as tools to reduce the uncertainty of the parametrization. By changing the temporal (Bolliger et al. 2017), spatial (Ménard and Marceau 2005) or thematic resolution (Conway 2009), the user can discard information which he or she cannot explain and, therefore, for which parameters are uncertain. Finer resolutions, that is, more complexity, do not necessarily mean less uncertainty. Thus, complexity must be decided according to the parameters and issues that can be explained by knowledge or data.

15.4.5 Validation and propagation of uncertainty

Once the analysis is performed, we usually validate the obtained results. These come with an additional uncertainty to those referred to previously: the one that arises from the propagation of the other uncertainties when data and processes, and their uncertainties, interact. Because of the origin of this uncertainty, it is very difficult to study and usually takes part of the called unknown unknowns, as we are not usually aware about their existence and extent. That is why its study has been

pointed out as an important research need (Gómez Delgado and Bosque Sendra 2004; Hunter 2005).

The larger the amount of data and components, the more important the propagation of uncertainty, being especially meaningful for integrated modelling (Van Delden et al. 2011). Where possible the user should avoid data or methods which introduce great sources of uncertainty, trying to keep an equilibrium between all sources of uncertainty that will take part of the analysis. Important drivers can however not be discarded and when they come with high uncertainty, solutions (participatory approaches, exploratory scenarios, etc) need to be sought to deal with them.

Several studies have tried to quantify the propagation of uncertainty (Tayyebi et al. 2014; Ferchichi et al. 2017). From a conceptual point of view, Yeh and Li (2006) tried to study this topic for a Cellular Automata model.

The validation of the analysis results, including the above-mentioned source of uncertainty, is usually performed against reference data. However, there is a great deal of methods and tools to perform this validation, which come with different sources of uncertainty as well. When using reference data, the uncertainty that they convey also affects the uncertainty of the validation.

When validating data, as LULC maps, the sample strategy in case of field validation data or the reference data uncertainty in the case of data comparison validation are important sources of uncertainty. Regarding the methods, there is a vast amount of literature about validation tools and techniques acknowledging their limits. For LUCCM, there are noticeable the critics to the kappa statistics (Pontius Jr. and Millones 2011), which have been answered by the development of new kappa measures that account for the LULC changes (Van Vliet et al. 2011). Different concepts and methods as the null resolution, fuzzy similarity or integrated validation tools have been proposed in the research about the validation uncertainties and limits (Hagen 2003; Pontius Jr. et al. 2004; Bradley et al. 2016).

In many cases, uncertainty or accuracy measures are provided for the whole dataset or analysis. However, these uncertainties can be different depending on the area or component considered. For data, uncertainties are usually contrasted between areas or, in the case of categorical data, between classes (Castilla and Hay 2007). Therefore, scale also plays a key role in the study of those uncertainties. For specific analysis, as LUCCM, some of the components can gather most of the uncertainty.

While there is a large number of methods available for the validation of those aspects of the analysis which are quantifiable or more evident (Paegelow et al. 2014), as the final results, there is less research about those qualitative uncertainties, regarding the problem conceptualization (Van Asselt 2000) or the role played by the user's knowledge and perspectives (Brown 2004). Nevertheless, when validating, the user must acknowledge that perfect validation is usually an utopia, especially for some analysis, as LUCCM, due to the inherent uncertainty of the complex systems which LULC models try to replicate (Van Asselt 2000).

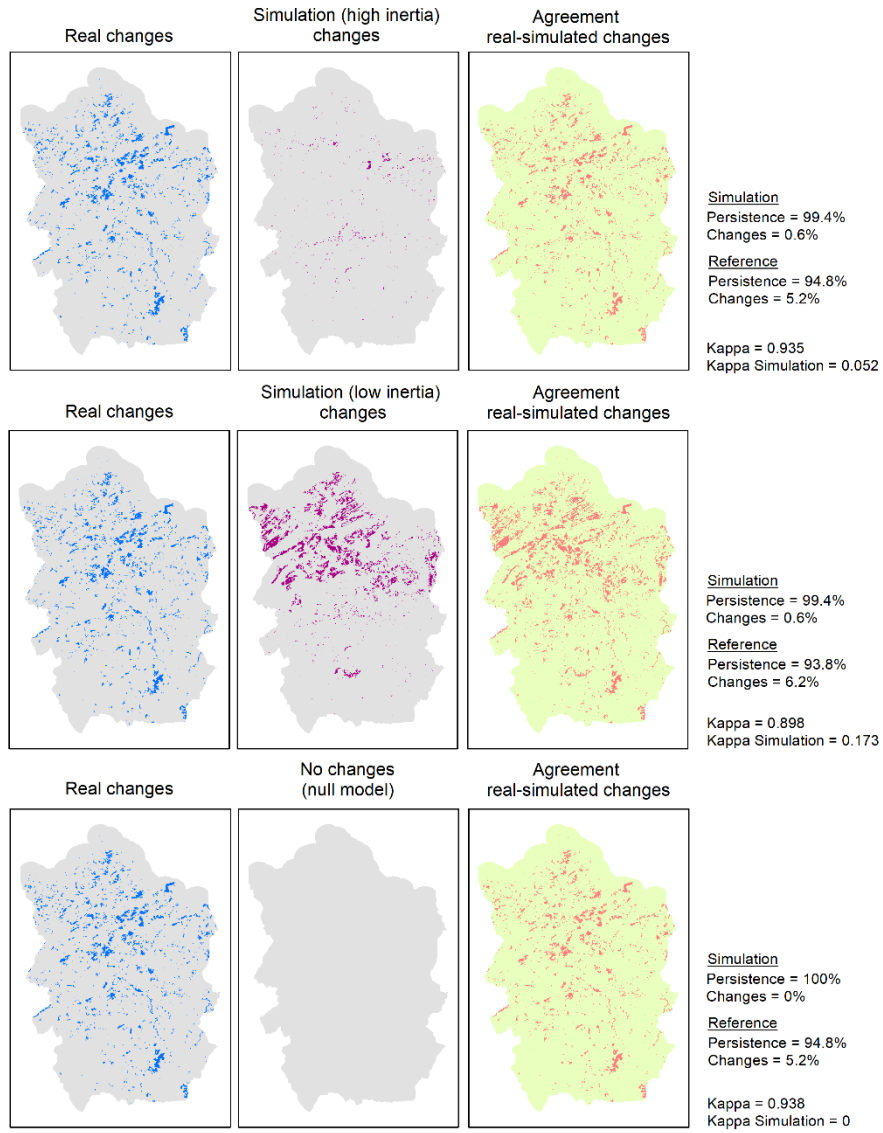


Fig. 15.9. Uncertainties of the kappa validation measures. Kappa accounts well for persistence, but not for change. Conversely, Kappa Simulation accounts well for change, but not for persistence. The null model only accounts for persistence, serving as reference

15.4.6 Process variability and communication of results

Regarding the epistemic uncertainty, we cannot know how the system or processes studied will evolve in the future (Fig. 15.10). This evolution will depend on a wide variety of factors, as environmental, socioeconomic or political ones, whose behaviour we cannot know with certainty in advance.

Accounting for this uncertainty is one of the main challenges, since we are not aware about all the possibilities of change that the analysed system can experience. In this regard, it is useful referring to the known unknowns and the unknown unknowns, that is, those issues that we know that we cannot know, but also those ones that we do not even know that we do not know (Recker 2015). Keeping in mind these ideas helps in the dealing and recognition of this uncertainty.

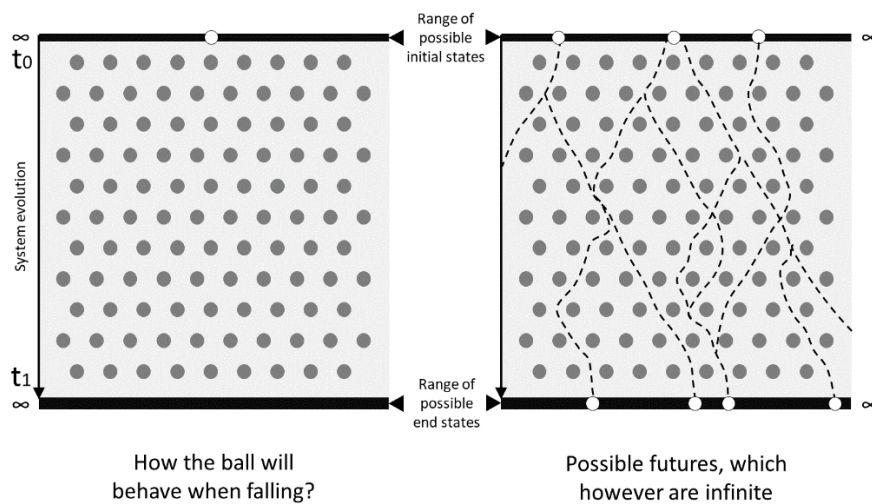


Fig. 15.10. Authors' view on the different ways a system can evolve in the future. Like a ball when falling through a labyrinth, we cannot be sure about the circumstances that the process will face. The end state of the system will depend on those, but also on its initial state, as the ball in our example

When communicating the results of a LUC model, accounting for those possible system evolutions is essential to transmit the audience one of the main uncertainties that the results of our analysis convey. In this regard, how we communicate these also introduces important uncertainties that must be conveniently addressed.

In the same manner we simplify the real world to represent it through data and geospatial tools, when we communicate the results of our analysis, we also simplify the outputs. If our analysis is stochastic, we must communicate how this stochasticity has affected the results (Fig. 15.11). In a similar vein, if the obtained maps are too detailed to be represented without any modifications, information about differences between maps showed and maps actually obtained should be provided.

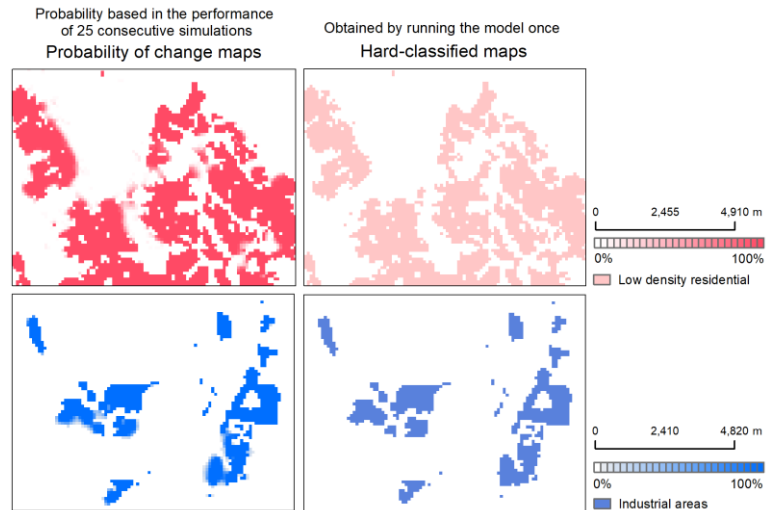


Fig. 15.11. Two different ways of communicating LUCCM results, which are especially contrasted in the case of high stochastic models, unlike the example included here. It shows part of the results of a model for the Great Sydney Area

15.5 How to deal with uncertainty?

When dealing with uncertainty one must be aware that it is unavoidable. Accordingly, the user can just act to reduce it and to provide as much information as possible about it. To do so, the user must first locate and determine the extent of the uncertainty.

15.5.1 Transparency and communication

Every analysis contains uncertainty and, consequently, all the achieved conclusions are uncertain as well. The utility of those conclusions and studies will depend on their level of uncertainty and on the users understanding of this uncertainty as well as on the limitations that it imposes.

To help in the understanding of the uncertainty, providing as much information as possible about it is an essential step. As every decision that we take, from the very selection of the problem to be studied, comes with uncertainty, providing information about uncertainty means spreading all the information and details of the analysis carried out, that is, all the decisions taken and their justification. For LUCCM, problem, level of detail (spatial, temporal and thematic resolution) and

software choices are important decisions impacting on the uncertainty of the analysis. Therefore, appropriate justification must be given. Thus, transparency of the way the study has been carried out is a required step for communicating uncertainty.

In addition to transparency, we must provide reports and measures about uncertainties, but always warn about the limitations and uncertainties of this information. In this regard, scale issues must always be considered when communicating this information, since uncertainty will be different depending on the area, class or component considered.

According to the public for which the analysis is intended, the language used in the communication of the uncertainty will be different, but always transparent, i.e. clear and concise. This avoids confusion, misinterpretation and lack of understanding, that is, the translational uncertainties (Van Asselt 2000). Moreover, when delivering the information, it must be clearly hierarchized in order to get the audience's interest. A summary of the most important uncertainties should first be presented, going from there into the details. This strategy, which has been called "progressive disclosure of information" (Kloprogge et al. 2007), favours the audience attention and, ergo, the audience awareness.

According to the previous ideas, when delivering conclusions that will form a base for decision-making, giving information about the uncertainty of the possible decisions that will be taken is important as well. That is, providing the trust interval over which the results and conclusions could be used to take decisions. In this regard, Openshaw (1989) stated that, instead of mere estimations of error, what is really needed is the certainty that decisions could be taken from the results of geospatial analysis. In a similar vein, Brown (2004) talks about educating uncertainty in the decision-making process and Hunter (1999) points out the need of knowing the uncertainty of data and analysis from a legal point of view.

15.5.2 Methods and strategies

Different strategies have been laid out for dealing with uncertainty. They come from different perspectives in the uncertainty management, which go from positivism (uncertainty can be studied objectively) to constructivism (uncertainty always depends on the context in which the analysis has been carried out, including an important subjectivity) (Wardekker et al. 2008). The positivist perspective has dominated most of the uncertainty research over the last decades. Main attention has focused in the quantification and reduction of uncertainty. In fact, Klir and Wierman (1999) stated that for three centuries uncertainty analysis was interpreted as probability theory.

The research community admits today the existence of qualitative and quantitative uncertainties (Warmink et al. 2010), which ergo are studied by qualitative or quantitative approaches (Van Asselt 2000). There is not a single method that can analyse all types of uncertainty and, therefore, different approaches are needed for

its complete study. Nevertheless, even in that case, a whole inventory and analysis of uncertainties is not possible. That is why every uncertainty analysis is at the same time uncertain. Thus, the study of uncertainty just offers an overview of the limitations of our analysis, arising the awareness about the possible interpretations of its conclusions.

Table 15.1. Common procedures for uncertainty analysis in Modelling environments. Quantitative (QN) and qualitative (QL) approaches

Method	Description	Studied uncertainties
Expert elicitation (QN/QL)	Collection of expert opinions through an established method about one of several aspects of the performed analysis to account for the analysis uncertainty.	Very flexible approach. Most of the uncertainties
Model comparison (QN)	Comparison of the same application over several model structures to evaluate the influence of the model structure on the achieved results.	Problem conceptualization and technical uncertainties
Monte Carlo analysis (QN)	It gives output statistics from the repetition of the same analysis under random variations of input data and parameters.	Data and parameter uncertainty
NUSAP (QN/QL)	It is a method that provides a comprehensive assessment of uncertainty for policy purposes. It characterizes the uncertainty according to five categories: Numeral, Unit, Spread (QN), assessment and pedigree (QL).	Data and parameter uncertainties
Participatory approaches (QL)	Accounting for the opinion of stakeholders or policy makers (practical knowledge) to study the parameters and assumptions made.	Very flexible approach. Most of the uncertainties
Scenario analysis (QN/QL)	It explores the possible future states of the studied system through the consideration of different assumptions about its evolution, which translates in changes in the parameters (QN) or trajectories (QL).	Process variability
Sensitivity analysis (QN)	It studies how changes in the parameters or the initial values of the analysis affect the outcomes.	Data and parameter uncertainties
Uncertainty matrix (QL)	Identification of types and location of the analysis uncertainties.	Identification of all sources of uncertainty
Validation techniques (QN/QL)	Evaluation of the agreement between the model outcome and reference data. It can comprise quantitative methods, but also expert assessment and participatory approaches.	Output uncertainty

Christian et al. (2007), Harbin and Jianguo (2006), Klir and Wierman (1999), Matott et al. (2009), Uusitalo et al. (2015) and Van Asselt (2000) give an overview of different methods and strategies for the management of uncertainty. We summarize some of the most common ones in the Table 15.1.

Although the quantitative approaches are common, we want to stress here the importance of qualitative ones. One of the main objectives of LUCCM is assessing the possible future states of a system under different management policies. Scenario analysis provides to this end a consistent tool to deal with the future system variability. Scenarios can act as qualitative tools, as an image or representation of possible future narratives (storylines), or as quantitative ones, when those narratives are translated to parameters. Some research has gone in depth in this translation (Van Delden and Hagen-Zanker 2009). In addition, Mahmoud et al. (2009) provide a framework for the development of scenarios in environmental decision-making.

The Uncertainty Matrix proposed by Walker et al. (2003) offers a simple and easily understandable approach for uncertainty identification. It locates the different uncertainties of an analysis, giving information about their level and nature (Fig. 15.12). It is therefore a good tool for transmitting directly and in a summarized way all the uncertainties that an analysis conveys. The main problem of this method comes from the difficulty of individualizing every source of uncertainty, given the complex interactions between them. Warmink et al. (2010) have proposed a guideline to solve this issue, so every source of uncertainty can be classified in the matrix just once.

LOCATION		LEVEL			NATURE	
		Statistical uncertainty	Scenario uncertainty	Recognized ignorance	Epistemic uncertainty	Stochastic uncertainty
Context	Natural, technological, economic, social and political					
Model	Model structure					
	Technical					
Inputs	Driving forces					
	System data					
	Parameters					
	Outputs					

Fig. 15.12. The uncertainty matrix. Source: based on Walker et al. (2003)

Finally, participatory approaches, both from experts and agents, provide a useful and easy way of accounting for most of the uncertainties at the same time that the user makes the audience aware about the limitations and problems of the analysis. For LUCCM, the study developed by Hewitt et al. (2014) can be used as reference. Moreover, Reed et al. (2017) and Van Delden et al. (2011) give some useful tips and frameworks for the inclusion of participation in the analysis.

15.6 Concluding remarks

Uncertainty, as an unavoidable attribute of any geospatial data or analysis, is a complex issue which deserves special attention in order to increase the reliability of our studies. This can be meaningful as well in the search for the daily applicability of our analysis and theories in policy processes.

Although substantial research has been done in this field, there are still plenty of unresolved issues which need to be addressed. It is especially important to focus on the analysis and management of qualitative uncertainties, which greatly determine the achieved results and whose communication is still a pending matter. How uncertainties are communicated is, at the same time, another key point that has not received much attention. Efforts should be put in the increment of uncertainty communication awareness among the research community. Many studies are still presented today without proper information and justification of key decisions, as resolution or extent. Just by changing this common practise we can really account for the uncertainty problem and translate those practices to the practical works delivered by the same community.

Acknowledgments

This work has been supported by project SIGEOMOD_2020 · BIA2013-43462-P (Spanish Ministry of Economy and Competitiveness and the Feder European Regional Development Fund). The first author is also grateful to the Spanish Ministry of Economy and Competitiveness and the European Social Fund for the funding of his research activity (Ayudas para contratos predoctorales para la formación de doctores 2014).

References

- Aldwaik SZ, Onsted JA, Pontius Jr. RG (2015) Behavior-based aggregation of land categories for temporal change analysis. *Int J Appl Earth Obs Geoinf* 35:229–238 . doi: 10.1016/j.jag.2014.09.007
- Ascough JC, Maier HR, Ravalico JK, Strudley MW (2008) Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecol Modell* 219:383–399 . doi: 10.1016/j.ecolmodel.2008.07.015
- Aspinall RJ, Pearson DM (1995) Describing and managing uncertainty of categorical maps in GIS. In: Fisher P (ed) *Innovations In GIS 2*. Taylor & Francis, London, Bristol, pp 71–83
- Batisani N, Yarnal B (2009) Uncertainty awareness in urban sprawl simulations: Lessons from a small US metropolitan region. *Land use policy* 26:178–185 . doi: 10.1016/j.landusepol.2008.01.013
- Bolliger J, Schmatz D, Pazúr R, et al (2017) Reconstructing forest-cover change in the Swiss Alps between 1880 and 2010 using ensemble modelling. *Reg Environ Chang* 17:2265–2277 . doi: 10.1007/s10113-016-1090-4

- Botterweg P (1995) The user's influence on model calibration results: an example of the model SOIL, independently calibrated by two users. *Ecol Modell* 81:71–81 . doi: 10.1016/0304-3800(94)00161-A
- Bradley A V., Rosa IMD, Pontius Jr. RG, et al (2016) SimiVal, a multi-criteria map comparison tool for land-change model projections. *Environ Model Softw* 82:229–240 . doi: 10.1016/j.envsoft.2016.04.016
- Brown JD (2004) Knowledge, uncertainty and physical geography: towards the development of methodologies for questioning belief. *Trans Inst Br Geogr* 24:367–381 . doi: 10.1111/j.0020-2754.2004.00342.x
- Burnicki AC, Brown DG, Goovaerts P, et al (2010) Propagating error in land-cover-change analyses: impact of temporal dependence under increased thematic complexity. *Int J Geogr Inf Sci* 24:1043–1060 . doi: 10.1080/13658810903279008
- Castilla G, Hay GJ (2007) Uncertainties in land use data. *Hydrol Earth Syst Sci* 11:1857–1868 . doi: 10.5194/hess-11-1857-2007
- Chuvieco E (2016) *Fundamentals of Satellite Remote Sensing. An environmental approach*, 2 edition. CRC Press, Boca Raton
- Congalton RG (1997) Exploring and Evaluating the Consequences of Vector-to-Raster and Raster-to-Vector Conversion. *Photogramm Eng Remote Sens* 63:425–434
- Congalton RG, Fenstermaker LK, John R, Mcgwire KC (1991) Remote Sensing and Geographic Information System Data Integration: Error Sources and Research Issues. *Photogramm Eng Remote Sens* 57:677–687
- Conway TM (2009) The impact of class resolution in land use change models. *Comput Environ Urban Syst* 33:269–277 . doi: 10.1016/j.compenvurbsys.2009.02.001
- Dendoncker N, Schmit C, Rounsevell M (2008) Exploring spatial data uncertainties in land-use change scenarios. *Int J Geogr Inf Sci* 22:1013–1030 . doi: 10.1080/13658810701812836
- Díaz-Pacheco J, Van Delden H, Hewitt R (2018) The importance of scale in land use models: experiments in data conversion, data resampling, resolution and neighbourhood extent. In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) *Geomatic Approaches for Modeling Land Change Scenarios*. Springer, Cham, Switzerland, pp 163–186
- Dunn WN (2001) Using the method of context validation to mitigate type III error in environmental policy analysis. In: Hisschemöller M, Hoppe R, Dunn WN, Ravetz JR (eds) *Knowledge, power and participation in environmental policy analysis*. Taylor & Francis, New York, pp 417–436
- Elsawah S, Pierce SA, Hamilton SH, et al (2017) An overview of the system dynamics process for integrated modelling of socio-ecological systems: Lessons on good modelling practice from five case studies. *Environ Model Softw* 93:127–145 . doi: 10.1016/j.envsoft.2017.03.001
- Ferchichi A, Boulila W, Farah IR (2017) Reducing uncertainties in land cover change models using sensitivity analysis. *Knowl Inf Syst*. doi: 10.1007/s10115-017-1102-9
- García-Álvarez D (2018) The influence of scale in LULC modelling. A comparison between two different LULC maps (SIOSE and CORINE). In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F (eds) *Geomatic Approaches for Modeling Land Change Scenarios*. Springer, Cham, Switzerland, pp 187–213
- 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 . doi: 10.1016/j.jag.2017.07.001

- García Martínez ED, Chas-Amil ML, Touza J (2015) Assessment of the Spanish land cover information to estimate forest area in Galicia. *Boletín la Asoc Geógrafos Españoles* 69:333–350
- Gómez Delgado M, Barredo JI (2006) *Sistemas de información geográfica y evaluación multicriterio en la ordenación del territorio*, 2 edition. Ra-Ma, Madrid
- Gómez Delgado M, Bosque Sendra J (2004) Aplicación de análisis de incertidumbre como método de validación y control del riesgo en la toma de decisiones. *GeoFocus* 4:179–208
- Goodchild MF (1991) Issues of quality and uncertainty. In: Muller J-C (ed) *Advances in Cartography*. International Cartographic Association, Elsevier Applied Science, London, New York, pp 17–42
- Grinblat Y, Gilichinsky M, Benenson I (2016) Cellular Automata Modeling of Land-Use/Land-Cover Dynamics: Questioning the Reliability of Data Sources and Classification Methods. *Ann Am Assoc Geogr* 106:1299–1320 . doi: 10.1080/24694452.2016.1213154
- Hagen A (2003) Fuzzy set approach to assessing similarity of categorical maps. *Int J Geogr Inf Sci* 17:235–249 . doi: 10.1080/13658810210157822
- Hewitt R, Van Delden H, Escobar F (2014) Participatory land use modelling, pathways to an integrated approach. *Environ Model Softw* 52:149–165 . doi: 10.1016/j.envsoft.2013.10.019
- Houet T, Vacquie L, Sheeren D (2015) Evaluating the spatial uncertainty of future land abandonment in a mountain valley (Videssos, Pyrenees - France): Insights from model parameterization and experiments. *J Mt Sci* 12:1095–1112 . doi: 10.1007/s11629-014-3404-7
- Hunter G (2005) Managing uncertainty in GIS. In: Longley PA, Goodchild MF, Maguire DJ, Rhind DW (eds) *Geographical Information Systems: Principles, Techniques, Management and Applications*. Wiley, pp 633–641
- Jafarnejhad J, Salmanmahiny A, Sakieh Y (2012) Subjectivity versus Objectivity: Comparative Study between Brute Force Method and Genetic Algorithm for Calibrating the SLEUTH Urban Growth Model. *J Urban Plan Dev* 142:1–12 . doi: 10.1061/(ASCE)UP.1943-5444.0000307.
- Kemp KK (ed) (2008) *Encyclopedia of geographic information science*. SAGE Publications, Waimea, Hawaii
- Klein Goldewijk K, Verburg PH (2013) Uncertainties in global-scale reconstructions of historical land use: An illustration using the HYDE data set. *Landsc Ecol* 28:861–877 . doi: 10.1007/s10980-013-9877-x
- Klir G, Wierman M (1999) *Uncertainty-based information: elements of generalized information theory*, 2 edition. Springer, Berlin, Heidelberg
- Klopprogge P, Van der Sluijs JP, Wardekker JA (2007) *Uncertainty Communication. Issues and good practice*. Utrecht
- Lee DB (1973) Requiem for Large-Scale Models. *J. Am. Inst. Plann.* 39:163–178
- Li H, Wu J (2006) Uncertainty analysis in ecological studies: an overview. In: Wu J, Jones KB, Li H, Loucks OL (eds) *Scaling and Uncertainty Analysis in Ecology: Methods and Applications*. Springer, Dordrecht, pp 45–66
- Lloyd CD (2014) *Exploring Spatial Scale in Geography*. Wiley, Chichester
- Mahmoud M, Liu Y, Hartmann H, et al (2009) A formal framework for scenario development in support of environmental decision-making. *Environ Model Softw* 24:798–808 . doi: 10.1016/j.envsoft.2008.11.010
- Matott LS, Babendreier JE, Purucker ST (2009) Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resour Res* 45:1–14 . doi: 10.1029/2008WR007301

- Ménard A, Marceau DJ (2005) Exploration of spatial scale sensitivity in geographic cellular automata. *Environ Plan B Plan Des* 32:693–714 . doi: 10.1068/b31163
- National Research Council (2014) *Advancing Land Change Modeling: Opportunities and Research Requirements*. National Academies Press, Washington, D.C.
- Olaya V (2014) *Sistemas de Información Geográfica*
- Openshaw S (1989) Learning to live with errors in spatial databases. In: Goodchild MF, Gopal S (eds) *Accuracy of spatial databases*. Taylor & Francis, London, pp 263–276
- Paegelow M, Camacho Olmedo MT, Mas J-F, Houet T (2014) Benchmarking of LUCC modelling tools by various validation techniques and error analysis. *Cybergeo*. doi: 10.4000/cybergeo.26610
- Pontius Jr. RG, Huffaker D, Denman K (2004) Useful techniques of validation for spatially explicit land-change models. *Ecol Modell* 179:445–461 . doi: 10.1016/j.ecolmodel.2004.05.010
- Pontius Jr. RG, Lippitt CD (2006) Can Error Explain Map Differences Over Time? *Cartogr Geogr Inf Sci* 33:159–171 . doi: 10.1559/152304006777681706
- Pontius Jr. RG, Millones M (2011) Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int J Remote Sens* 32:4407–4429 . doi: 10.1080/01431161.2011.552923
- Pontius Jr RG, Spencer J (2005) Uncertainty in extrapolations of predictive land change models. *Environ Plan B Plan Des* 32:211–230
- Quattrochi DA, Goodchild MF (1997) *Scale in remote sensing and GIS*. CRC Press, Boca Raton
- Recker J (2015) Research on Conceptual Modelling: Less Known Knowns and More Unknown Unknowns, Please. In: *Proceedings of the 11th Asia-Pacific Conference on Conceptual Modelling*, Australian Computer Society, Sydney, pp 3–7
- Reed MS, Challies E, de Vente J, et al (2017) A theory of participation: what makes stakeholder and public engagement in environmental management work? *Restor Ecol* 1–19 . doi: 10.1111/j.1399-0012.2012.01641.x
- Refsgaard JC, Drews M, Jeppesen E, et al (2013) The role of uncertainty in climate change adaptation strategies — A Danish water management example. *Mitig Adapt Strateg Glob Chang* 18:337–359 . doi: 10.1007/s11027-012-9366-6
- Refsgaard JC, Van der Sluijs JP, Højberg AL, Vanrolleghem PA (2007) Uncertainty in the environmental modelling process - A framework and guidance. *Environ Model Softw* 22:1543–1556 . doi: 10.1016/j.envsoft.2007.02.004
- Rosa IMD, Purves D, Carreiras JMB, Ewers RM (2015) Modelling land cover change in the Brazilian Amazon: temporal changes in drivers and calibration issues. *Reg Environ Chang* 15:123–137 . doi: 10.1007/s10113-014-0614-z
- Tayyebi AH, Tayyebi A, Khanna N (2014) Assessing uncertainty dimensions in land-use change models: Using swap and multiplicative error models for injecting attribute and positional errors in spatial data. *Int J Remote Sens* 35:149–170 . doi: 10.1080/01431161.2013.866293
- Torrens PM (2011) Calibrating and Validating Cellular Automata Models of Urbanization. In: Yang X (ed) *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*. Wiley, pp 335–345
- Uusitalo L, Lehtikoinen A, Helle I, Myrberg K (2015) An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environ Model Softw* 63:24–31 . doi: 10.1016/j.envsoft.2014.09.017
- Van Asselt MBA (2000) *Perspectives on Uncertainty and Risk - the PRIMA Approach to Decision*

- Support. Kluwer Academic Publishers, Boston, Dordrecht, London
- Van Delden H, Hagen-Zanker A (2009) New Ways of Supporting Decision Making: Linking Qualitative Storylines with Quantitative Modelling. In: Geertman S, Stillwell J (eds) *Planning Support Systems Best Practice and New Methods*. Springer, pp 347–367
- Van Delden H, Seppelt R, White R, Jakeman AJ (2011) A methodology for the design and development of integrated models for policy support. *Environ Model Softw* 26:266–279 . doi: 10.1016/j.envsoft.2010.03.021
- 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 Modell* 222:1367–1375 . doi: 10.1016/j.ecolmodel.2011.01.017
- Verburg PH, Neumann K, Nol L (2011) Challenges in using land use and land cover data for global change studies. *Glob Chang Biol* 17:974–989 . doi: 10.1111/j.1365-2486.2010.02307.x
- Verburg PH, Veldkamp A (2004) Projecting land use transitions at forest in the Philippines at two spatial scales. *Landsc Ecol* 19:77–98 . doi: 10.1023/B:LAND.0000018370.57457.58
- Walker WE, Harremoës P, Rotmans J, et al (2003) Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integr Assess* 4:5–17 . doi: <https://doi.org/10.1076/1aij.4.1.5.16466>
- Wardekker JA, Van der Sluijs JP, Janssen PHM, et al (2008) Uncertainty communication in environmental assessments: views from the Dutch science-policy interface. *Environ Sci Policy* 11:627–641 . doi: 10.1016/j.envsci.2008.05.005
- Warmink JJ, Janssen JAEB, Booij MJ, Krol MS (2010) Identification and classification of uncertainties in the application of environmental models. *Environ Model Softw* 25:1518–1527 . doi: 10.1016/j.envsoft.2010.04.011
- Waser LT, Schwarz M (2006) Comparison of large-area land cover products with national forest inventories and CORINE land cover in the European Alps. *Int J Appl Earth Obs Geoinf* 8:196–207 . doi: 10.1016/j.jag.2005.10.001
- Yeh AG-O, Li X (2003) Uncertainties In Urban Simulation Using Cellular Automata And Gis. In: *Proceedings of the 7th International Conference on GeoComputation*. Southampton
- Yeh AG-O, Li X (2006) Errors and uncertainties in urban cellular automata. *Comput Environ Urban Syst* 30:10–28 . doi: 10.1016/j.compenvurbysys.2004.05.007