

Towards an uncertainty reduction framework for land-cover change prediction using possibility theory

Ahlem Ferchichi¹ · Wadii Boulila^{1,2} · Imed Riadh Farah^{1,2}

Received: 30 April 2016 / Accepted: 3 October 2016 / Published online: 18 October 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This paper presents an approach for reducing uncertainty related to the process of land-cover change (LCC) prediction. LCC prediction models have, almost, two sources of uncertainty which are the uncertainty related to model parameters and the uncertainty related to model structure. These uncertainties have a big impact on decisions of the prediction model. To deal with these problems, the proposed approach is divided into three main steps: (1) an uncertainty propagation step based on possibility theory is used as a tool to evaluate the performance of the model; (2) a sensitivity analysis step based on Hartley-like measure is then used to find the most important sources of uncertainty; and (3) a knowledge base based on machine learning algorithm is built to identify the reduction factors of all uncertainty sources of parameters and to reshape their values to reduce in a significant way the uncertainty about future changes of land cover. In this study, the present and future growths of two case studies were anticipated using multi-temporal Spot-4 and Landsat satellite images. These data are used for the preparation of prediction map of year 2025. The results show that our approach based on possibility theory has a potential for reducing uncertainty in LCC prediction modeling.

Keywords LCC prediction · Parameter uncertainty · Structural uncertainty · Possibility theory · Sensitivity analysis

1 Introduction

LCC is a central issue in the sustainability debate because of its wide range of environmental impacts. Models of LCC start with an initial land-cover situation for a given case study area. Then, they use an inferred transition function, representing the processes of change, to simulate the expansion and contraction of a predefined set of land-cover types over a given period. LCC models help to improve our understanding of the land system by establishing cause-effect relations and testing them on historic data. They help to identify the drivers of LCC and their relative importance. In addition, LCC models can be used to explore future land-cover pathways for different scenarios. However, the performance of the LCC prediction models is affected by different types of uncertainties (i.e., aleatory or/and epistemic uncertainties). These uncertainties can be subdivided into two sources: parameter uncertainty (adequate values of model parameters) [1,2] and structural uncertainty (ability of the model to describe the catchment's response) [3]. These sources contribute with different levels to the uncertainty associated with the predictive model. It is important to quantify the uncertainty due to uncertain model parameter, but methods for quantifying uncertainty due to uncertainty in model structure are less well developed. For quantifying, probability theory is generally used. Moreover, numerous authors conclude that there are limitations in using probability theory in this context. So far, several alternative frameworks based on non-probabilistic theories have been proposed in the literature. By no means do the promoters of theories pretend to replace probability theory; they just

✉ Ahlem Ferchichi
ferchichi.ahlem@gmail.com

Wadii Boulila
wadii.boulila@riadi.rnu.tn

Imed Riadh Farah
riadh.farah@ensi.rnu.tn

¹ RIADI Laboratory, National School of Computer Sciences, University of Manouba, Manouba, Tunisia

² ITI Department, Telecom-Bretagne, Brest, France

present different levels of expressiveness that leave room for properly representing the lack of background knowledge [4]. The most common theories that are used from these alternatives are imprecise probabilities [5], random sets [6], belief function theory [7], fuzzy sets [8], and possibility theory [9]. In our context of continuous measurements, the possibility theory is more adapted, because it generalises interval analysis and provides a bridge with probability theory by its ability to represent a family of probability distributions. In summary, the possibility distribution has the ability to handle both aleatory and epistemic uncertainty of pixel detection through a possibility and a necessity measures. In this framework, the possibility distributions of the model outputs are used to derive the prediction uncertainty bounds.

Understanding the impact of parameter and structural uncertainty on LCC prediction models outcomes is crucial to the successful use of these models. On the other hand, model optimization with multiple uncertainty sources is complex and very time-consuming task. However, the sensitivity analysis has been proved to be efficient and robust to find the most important sources of uncertainty that have effect on LCC prediction models output [1, 9, 10]. Parameter sensitivity analysis allows to examine effects of model parameter on results, whereas structural sensitivity analysis allows to modify the structure of the model and to identify the possible structural factors that affect the robustness of the results (vary structure of model and see impact on results and trade-offs between choices). Several sensitivity analysis methods exist, including screening method [11], differential analysis [12], variance-based methods [13], sampling-based methods [14], and a relative entropy-based method [15]. However, all these require specific probability distribution in modeling both model parameters and model structure. In the literature, previous non-probabilistic methods of sensitivity analysis are developed [16, 17]. Several studies have confirmed the robustness of use of Hartley-like measure to apply sensitivity analysis in fuzzy theory framework in numerous fields [33–35]. Minimum value to Hartley-like measure of the model output is considered to be the most sensitive source.

Based on possibilistic approach, this study proposes an approach for reducing parameter and structural uncertainty in LCC prediction modeling. The proposed approach is divided into three main steps: (1) an uncertainty propagation step based on possibility theory is used as a tool to evaluate the performance of the model; (2) a sensitivity analysis step based on Hartley-like measure is used to find the most important sources of uncertainty; and (3) a knowledge base based on machine learning algorithm is built to identify the reduction factors of all uncertainty sources of parameters. Then, values of these parameters are reshaped to improve decisions about future changes of land cover in Saint-Denis city, Reunion Island and Cairo region, Egypt.

The rest of this paper is organized as follows: Sect. 2 presents a description of the proposed approach for reducing uncertainty throughout the model of LCC prediction. Results are given and described in Sect. 3. Finally, conclusion and future works are outlined in Sect. 4.

2 Proposed approach

Modeling LCC helps analyzing causes and consequences of land change to support land-cover planning and policy. In the literature, previous models are proposed for predicting LCC [18–23]. In this study, we use the LCC prediction model described by Boulila et al. in [18]. This model exploits machine learning tools to build predictions and decisions for several remote sensing fields. It takes into account uncertainty related to the spatiotemporal mining process to provide more reliable and accurate information about LCC in satellite images.

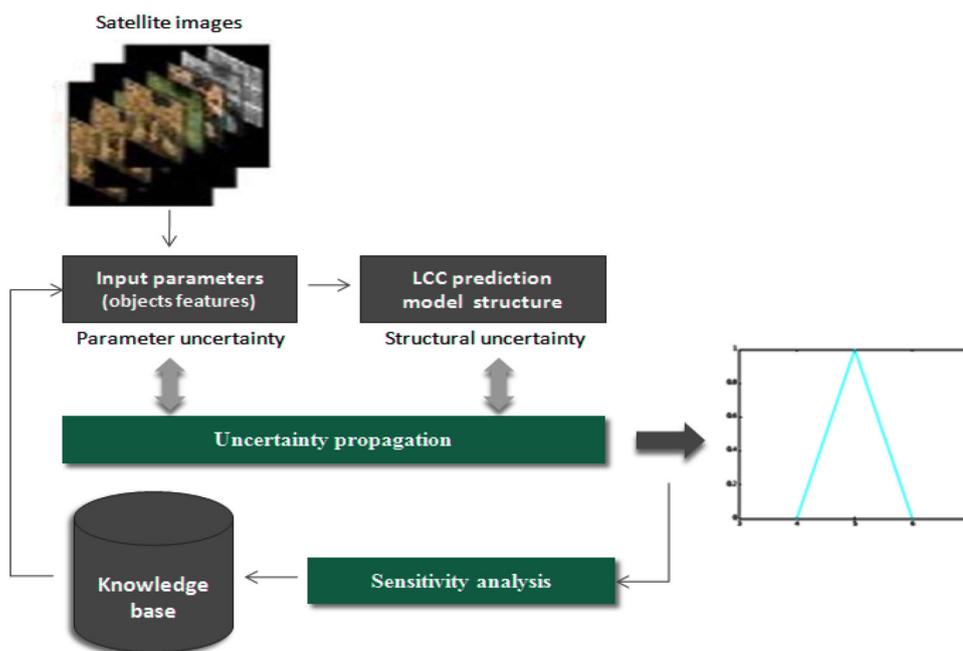
In this paper, the proposed approach for reducing parameter and structural uncertainty is applied to model presented in [18] and it has the following steps (Fig. 1): (1) identifying uncertainty related to parameters and model structure; (2) propagating the uncertainty through the LCC prediction model using the possibility theory; (3) performing a sensitivity analysis using the Hartley-like measure; and (4) constructing knowledge base using machine learning algorithm to improve parameters' quality.

2.1 Step 1: identifying parameters and structure of LCC prediction model

2.1.1 Choice of parameters

Input parameters of LCC prediction model describe the objects' features extracted from satellite images which are the subject of studying changes. In this study, we consider 26 features: ten spectral, five texture, seven shape, one vegetation, and three climate features. Spectral features are: mean values and standard deviation values of green (MG, SDG), red (MR, SDR), NIR (MN, SDN), SWIR (MS, SDS), and monospectral (MM, SDM) bands for each image object. Texture features are: homogeneity (Hom), contrast (Ctr), entropy (Ent), standard deviation (SD), and correlation (Cor) generated from gray-level co-occurrence matrix (GLCM). Shape and spatial relationship features are: area (A), length/width (LW), shape index (SI), roundness (R), density (D), metric relations (MR), and direction relations (DR). Vegetation feature is: Normalized Difference Vegetation Index (NDVI) that is the ratio of the difference between NIR and red reflectance. Finally, climate features are: temperature (Tem), humidity (Hum), and pressure (Pre). These features are selected based on previous results, as reported in [18], and are considered as input parameters to the LCC model.

Fig. 1 General modeling proposed framework



Uncertainties related to these input parameters are very numerous and affect model outputs. In general, these uncertainties can be of two types: epistemic and aleatory. The type of uncertainty of each parameter depends on sources of its uncertainty. Therefore, it is necessary to identify uncertainty sources related to input parameters:

- *Uncertainty sources of spectral parameters* Several studies investigated effects of spectral parameters [28]. Among these effects, we list: spectral reflectance of the surface (S1), sensor calibration (S2), effect of mixed pixels (S3), effect of a shift in the channel location (S4), pixel registration between several spectral channels (S5), atmospheric temperature and moisture profile (S6), effect of haze particles (S7), instrument’s operation conditions (S8), atmospheric conditions (S9), as well as by the stability of the instrument itself characteristics (S10).
- *Uncertainty sources of texture parameters* Among these sources, we list: the spatial interaction between the size of the object in the scene and the spatial resolution of the sensor (S11), a border effect (S12), and ambiguity in the object/background distinction (S13).
- *Uncertainty sources of shape parameters* Uncertainty related to shape parameters can rely to the following factors [28]: accounting for the seasonal position of the sun with respect to the Earth (S14), conditions in which the image was acquired changes in the scene’s illumination (S15), atmospheric conditions (S16), and observation geometry (S17).
- *Uncertainty sources of NDVI* Among factors that affect NDVI, we can list: variation in the brightness of soil

background (S18), red and NIR bands (S19), atmospheric perturbations (S20), and variability in the sub-pixel structure (S21).

- *Uncertainty sources of climate parameters* According to [29], uncertainty sources related to climate parameters can be: atmospheric correction (S22), noise of the sensor (S23), land surface emissivity (S24), aerosols and other gaseous absorbers (S25), angular effects (S26), wavelength uncertainty (S27), full-width half-maximum of the sensor (S28), and bandpass effects (S29).

2.1.2 Description of model structure

In this study, we use the LCC prediction model described in [18]. This model is divided into three main steps. It starts by a similarity measurement step to find similar states (in the object database) to a query state (representing the query object at a given date). Here, a state is a set of attributes describing an object at a given data. The second step is composed by three substeps: (1) finding the corresponding model for the state; (2) finding all forthcoming states in the model (states having dates superior to the date of the retrieved state); and (3) for each forthcoming date, build the spatiotemporal change tree for the retrieved state. The third step is to construct the spatiotemporal changes for the query state. Each of these steps is based on a number of assumptions as follows:

- *Similarity measure step*: Distance between states ($d(S_t, S_{t_1}) \geq 0.9$ indicates a higher similarity between the query and the retrieved states). In addition, similarity measure between states is based on time assumption.

– Spatiotemporal change tree building step: The aim of this step is to determine the confidence degrees and the percentage of changes of the model between two dates and for different land-cover types. The confidence degree of changes is achieved by a fuzzy decision tree (fuzzy ID3). This method is based on a number of assumptions such as: the proportion of a data set of land-cover type, the size of a data set, etc. The percentage of changes is achieved by computing the distances between two states and the centroid of the classes.

In this study, we consider structural uncertainty as uncertainty associated with assumptions of model structure, including distance between states, time assumption for similarity measure, assumptions of fuzzy ID3, and distance between states and centroid for changes percentage.

2.2 Step 2: propagating the uncertainty

In this step, we focus on how to propagate parameter and structural uncertainty through the LCC prediction model described in [18] via the possibility theory.

2.2.1 Basics of possibility theory

The possibility theory developed by Dubois and Prade [30] handles uncertainty in a qualitative way, but encodes it in the interval $[0, 1]$ called possibilistic scale. The basic building block in the possibility theory is named possibility distribution. A possibility distribution is defined as a mapping $\pi : \Omega \rightarrow [0, 1]$. It is formally equivalent to the fuzzy set $\mu(x) = \pi(x)$. Distribution π describes the more or less plausible values of some uncertain variable X . A possibility distribution is associated with two measures, namely, the possibility (Π) and necessity (N) measures, which are represented by Eq. (1):

$$\Pi(A) = \sup_{x \in A} \pi(x), \quad N(A) = \inf_{x \notin A} (1 - \pi(x)). \quad (1)$$

The possibility measure indicates to which extent event A is plausible, while the necessity measure indicates to which extent it is certain. They are dual, in the sense that $\Pi(A) = 1 - N(\bar{A})$, with \bar{A} the complement of A . They obey the following axioms:

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \quad (2)$$

$$N(A \cap B) = \min(N(A), N(B)) \quad (3)$$

An α cut of π is the interval $[\underline{x}_\alpha, \bar{x}_\alpha] = \{x, \pi(x) \geq \alpha\}$. The degree of certainty that $[\underline{x}_\alpha, \bar{x}_\alpha]$ contains the true value of X is $N([\underline{x}_\alpha, \bar{x}_\alpha]) = 1 - \alpha$. Conversely, a collection of nested sets A_i with (lower) confidence levels λ_i can be modeled as a possibility distribution, since the α cut of a (continuous)

possibility distribution can be understood as the probabilistic constraint $P(X \in [\underline{x}_\alpha, \bar{x}_\alpha]) \geq 1 - \alpha$. In this setting, necessity degrees are equated to lower probability bounds and possibility degrees to upper probability bounds.

2.2.2 Propagation of parameter uncertainty

In this section, the procedures of propagating unified structures dealing with parameter uncertainty of LCC prediction model will be addressed. Let us denote by $Y = f(X) = f(X_1, X_2, \dots, X_j, \dots, X_n)$ the model for LCC prediction with n uncertain parameters $X_j, j = 1, 2, \dots, n$, that are possibilistic, i.e., their uncertainties are described by possibility distributions $\pi_{X_1}(x_1), \pi_{X_2}(x_2), \dots, \pi_{X_j}(x_j), \dots, \pi_{X_n}(x_n)$. In more detail, the operative steps of the procedure are the following:

1. Set $\alpha = 0$.
2. Select the α cuts $A_\alpha^{X_1}, A_\alpha^{X_2}, \dots, A_\alpha^{X_j}, \dots, A_\alpha^{X_n}$ of the possibility distributions $\pi_{X_1}(x_1), \pi_{X_2}(x_2), \dots, \pi_{X_j}(x_j), \dots, \pi_{X_n}(x_n)$ of the possibilistic parameters $X_j, j = 1, 2, \dots, n$, as intervals of possible values $[\underline{x}_{j,\alpha}, \bar{x}_{j,\alpha}] j = 1, 2, \dots, n$.
3. Calculate the smallest and largest values of Y , denoted by \underline{y}_α and \bar{y}_α , respectively, letting variables X_j range within the intervals $[\underline{x}_{j,\alpha}, \bar{x}_{j,\alpha}] j = 1, 2, \dots, n$; in particular, $\underline{y}_\alpha = \inf_{j, X_j \in [\underline{x}_{j,\alpha}, \bar{x}_{j,\alpha}]} f(X_1, X_2, \dots, X_j, \dots, X_n)$ and $\bar{y}_\alpha = \sup_{j, X_j \in [\underline{x}_{j,\alpha}, \bar{x}_{j,\alpha}]} f(X_1, X_2, \dots, X_j, \dots, X_n)$.
4. Take the values \underline{y}_α and \bar{y}_α found in step 3 as the lower and upper limits of the α cut A_α^Y of Y ;
5. If $\alpha < 1$, then set $\alpha = \alpha + \Delta\alpha$ and return to step 2; otherwise, stop the algorithm. The possibility distribution $\pi_Y(y)$ of $Y = f(X_1, X_2, \dots, X_n)$ is constructed as the collection of the values \underline{y}_α and \bar{y}_α for each α cut.

2.2.3 Propagation of structural uncertainty

The propagation of structural uncertainty is implemented in combination with the propagation of parameter uncertainty. In this section, as parameter uncertainty is modeled by possibility theory, we use this method in this framework.

Suppose that a set of model structures $M_k, 1 \leq k \leq K$ represents the uncertainty related to the choice of model. For each model M_k , parameter uncertainty is propagated through this model. Consequently, the output indicator Y is characterized by a set of uncertainty representations according to each model structure. Thus, for all model structures $M_k, 1 \leq k \leq K$, we have a set of possibility distributions for output variable Y , noted $\pi_{Y_1}(y), \pi_{Y_2}(y), \dots, \pi_{Y_K}(y)$. The difference between these representations reflects the variation associated with structural uncertainty of LCC prediction model. These different representations $\pi_{Y_i}(y), 1 \leq i \leq K$

can be combined into a single representation. Therefore, the final uncertainty representation of output variable Y can be obtained by the following formulas:

$$y_{\alpha}^* = \inf_{i, Y_i \in [y_{i,\alpha}, \bar{y}_{i,\alpha}]} f(Y_1, Y_2, \dots, Y_i, \dots, Y_K) \tag{4}$$

$$\bar{y}_{\alpha}^* = \sup_{i, Y_i \in [y_{i,\alpha}, \bar{y}_{i,\alpha}]} f(Y_1, Y_2, \dots, Y_i, \dots, Y_K). \tag{5}$$

The possibility distribution $\pi_Y(y)$ of $Y = f(Y_1, Y_2, \dots, Y_K)$ is constructed as the collection of the values y_{α}^* and \bar{y}_{α}^* for each α cut. This distribution takes into account both parameter and structural uncertainty in the final output results of the prediction model.

2.3 Step 3: performing the sensitivity analysis

Based on Hartley-like measure, the third step consists to test impact of parameter and structural uncertainties on LCC prediction model output. The Hartley-like measure quantifies the most fundamental type of uncertainty (i.e., aleatory and epistemic uncertainty). This measure is generalized to fuzzy set by Higashi and Klir [31,32]. How to perform sensitivity analysis of both uncertainty sources in the possibilistic framework? The generalized measure H for any non-empty possibility distribution A defined on a finite universal set X has the following form:

$$H(A) = \frac{1}{h(A)} \int_0^{h(A)} \log_2 |A_{\alpha}| d\alpha, \tag{6}$$

where A_{α} denotes the cardinality of the α cuts of the possibility distributions A and $h(A)$ the height of A . For possibilistic intervals or numbers on the real line, the Hartley-like measure is defined as

$$HL(A) = \int_0^1 \log_2(1 + \lambda(A_{\alpha})) d\alpha, \tag{7}$$

where $\lambda(A_{\alpha})$ is the Lebesgue measure of A_{α} [31]. Mathematically, for a possibilistic number $A = [a_L, a_m, a_R]$ given by the possibility distribution

$$\pi_A(x) = \begin{cases} \frac{x-a_L}{a_m-a_L}, & \text{if } a_L \leq x \leq a_m \\ \frac{x-a_R}{a_m-a_R}, & \text{if } a_m \leq x \leq a_R \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

the Hartley-like measure is given by the expression as follows:

$$HL(A) = \frac{1}{(a_L - a_R) \ln(2)} \times ([1 + (a_R - a_L)] \ln[1 + (a_R - a_L)] - (a_R - a_L)). \tag{9}$$

The minimum value of Hartley-like measure of the model output with respect to fixing a particular parameter to the

most likely value, for a particular point of observation, leads to finding the most sensitive parameter. We can use the same measure to perform structural sensitivity analysis.

2.4 Step 4: constructing the knowledge base

After performing parameter and structural sensitivity analysis, the main purpose of this step is to identify reduction approaches of all uncertainty sources of LCC model parameters. In general, the knowledge base stores the embedded knowledge in the system and the rules defined by an expert. In this study, we used an inductive learning technique to automatically build a knowledge base. Two main steps are proposed which are training and decision tree generation. The learning step provides examples of concepts to be learned. The second step is the decision tree generation. This step generates the first decision trees from the training data. These decision trees are then transformed into production rules. Then, our knowledge base that contains all uncertainty sources and their reduction approaches is presented in Fig. 2. This knowledge base is used to improve data quality and then reduce in a significant way the uncertainty about future changes of land cover.

3 Experimental results

The aim of this section is to validate and to evaluate the performance of the proposed approach through two case studies for reducing parameter and structural uncertainty in LCC prediction modeling.

3.1 Case study 1

3.1.1 Description of the study area and data

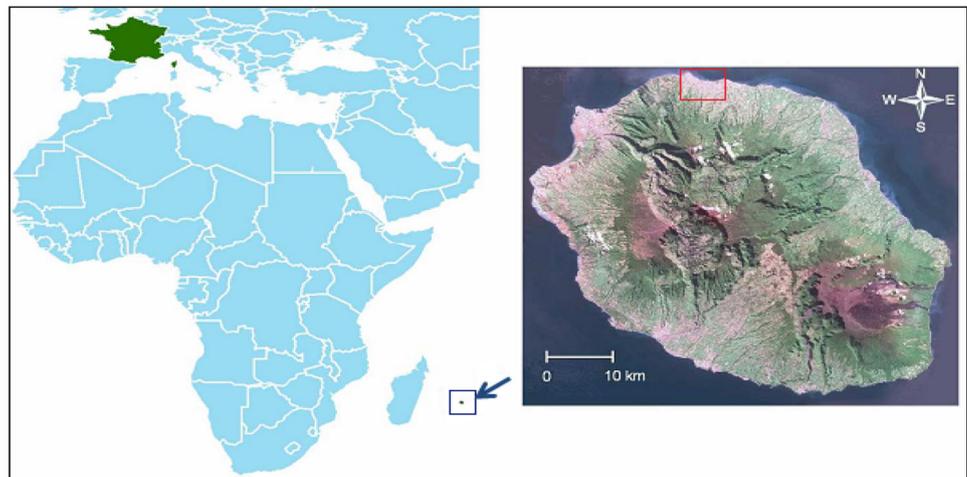
Reunion Island is a French territory of 2500 km² located in the Indian Ocean, 200 km South-West of Mauritius and 700 km to the East of Madagascar (Fig. 3). Mean annual temperatures decrease from 24 °C in the lowlands to 12 °C at ca 2000 m. Mean annual precipitation ranges from 3 m on the eastern windward coast, up to 8 m in the mountains and down to 1 m along the south western coast. Vegetation is most clearly structured along gradients of altitude and rainfall [27].

Reunion Island has a strong growth in a limited area with an estimated population of 833,000 in 2010 that will probably be more than 1 million in 2030 [24]. It has been significant changes, putting pressure on agricultural and natural areas. The urban areas expanded by 189 % over the period from 1989 to 2002 [25] and available land became a rare and coveted resource. The landscapes are now expected to fulfil multiple functions, i.e., urbanization, agriculture production,

Fig. 2 Production rules generated from uncertainty sources of input parameters

<p>If spectral parameters</p> <ul style="list-style-type: none"> Source = S1 and S2 : 'strict requirements for the instrument's design' and 'envisaging of appropriate procedures for on-board calibration' and 'choosing appropriate algorithms for radiometric correction' Source = S3 and S4 and S5 : 'reducing the wavelength range of the irradiance or spectral response measurement' and 'reducing the cloud shadows and cloud contamination effects' Source = S6 and S9 : 'choosing appropriate algorithms for atmospheric correction' Source = S7 and S8 and S10 : 'reducing errors of sensor system itself' <p>If texture parameters</p> <ul style="list-style-type: none"> Source = S11 : 'using high spatial resolution' Source = S12 and S13 : 'choosing appropriate methods for segmentation' <p>If shape parameters</p> <ul style="list-style-type: none"> Source = S14 and S15 : 'improving the platforms' stability and the carrier's velocity and 'technological enhancement of the sensors themselves' and 'reducing effects of sun angles' Source = S16 : 'reducing the effects of atmospheric conditions' Source = S17 : 'improving the overall segmentation quality' and 'reducing the number of bad pixels and the size of bad areas and 'improvement the imperfection of pixels' response' <p>If NDVI parameter</p> <ul style="list-style-type: none"> Source = S18 and S20 : 'choosing appropriate algorithms for atmospheric correction' Source = S19 : 'reducing errors in surface measurements for the NIR and red bands' Source = S21 : 'reducing the temporal variations effects in the solar zenith and azimuth angles and 'reducing the sun angle effects and noise contamination' <p>If climate parameters</p> <ul style="list-style-type: none"> Source = S22 : 'choosing appropriate algorithms for atmospheric correction' Source = S23 : 'reducing errors of sensor system itself' Source = S24 : 'reducing the emissivity variations' Source = S25 and S26 : 'reducing the sun angle effects and solar heating' Source = S27 and S28 and S29 : 'reducing the errors of radiometer calibration and the errors of radiation' and 'reducing errors of spatial and temporal variability of clouds'
--

Fig. 3 Studied area for case study 1



and ecosystem conservation, and this causes conflicts among stakeholders about their planning and management [26].

Saint Denis is the capital of Reunion Island and the city with the most inhabitants on the island (Fig. 3). It hosts all the important administrative offices, and it is also a cultural center with numerous museums. Saint-Denis is also the largest city in all the French Overseas Departments. Available remote sensing data for this research include classified images of land over of Saint Denis from SPOT-4 images for the years 2006 and 2011 (Fig. 4). For this case, satellite data are classified after initial corrections and processing to prepare the data for extracting useful information. Spec-

tral, geometric, and atmospheric corrections of images are conducted to make features manifest, to increase the quality of images, and to eliminate the adverse effects of light and atmosphere. According to the study objective, five categories, including water, urban, forest, bare soil, and vegetation, are identified and classified.

3.1.2 Results of uncertainty propagation

As mentioned perviously, the model parameter and model structure of LCC prediction are marred by uncertainty. Ignoring each of these sources can affect the results of uncertainty

Fig. 4 Land-cover maps

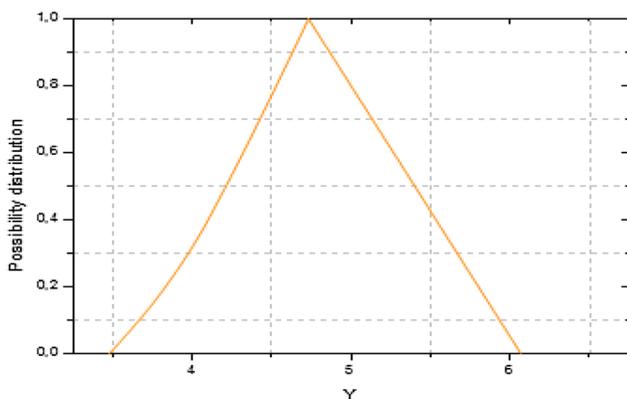
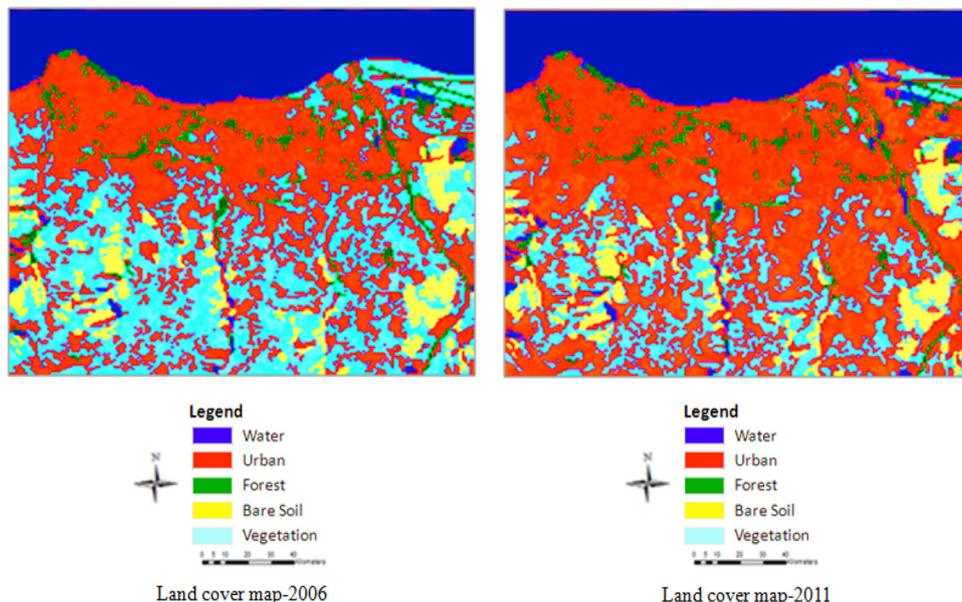


Fig. 5 Possibility distribution of LCC prediction model output for only parameter uncertainty

propagation. To illustrate the importance of propagating uncertainty related to model parameter and model structure through the LCC prediction model, the analysis with pure parameter uncertainty assumption is conducted. In this case, the possibility distribution of output representing only parameter uncertainty is obtained via possibility theory. Figure 5 shows this distribution based on 10,000 samples. With uncertainty in model parameter, there is uncertainty in model structure. Therefore, it is also important to illustrate the importance of structural uncertainty in LCC prediction modeling by the proposed approach. This is the reason behind using the LCC prediction model described in [18] with three different structures. Then, we obtain three different models (M_1 , M_2 , and M_3) with different assumptions. To take into account structural uncertainty in the final result, uncertainty related to parameters is first propagated and this for each prediction model.

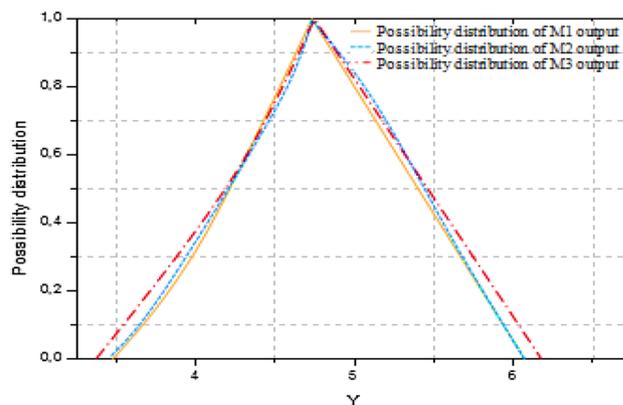


Fig. 6 Possibility distributions of LCC for three different prediction model structures

Figure 5 shows the possibility distribution of the LCC prediction model output, where only parameter uncertainty is propagated.

After propagating uncertainty of parameters through three different model structures, we obtain three uncertain representations of LCC, which are shown in Fig. 6. The difference between these three representations illustrates the impact of structural uncertainty. Compared with the result of the original LCC prediction model (M_1), we can see that there is an important difference between them.

Figure 7 shows possibility distribution representing integrated parameter and structural uncertainty through the LCC prediction modeling. Note that combining parameter and structural uncertainty can be crucially important to enhance the accuracy of the LCC prediction model.

3.1.3 Results of sensitivity analysis

In this paper, the sensitivity analysis based on Hartley-like measure is implemented to estimate the effect of 26 uncertain parameters through three different LCC prediction model structures. Results of the sensitivity analysis are shown in Fig. 8.

The different heights of the bars reveal the various levels of sensitivity, and a long bar indicates high sensitivity

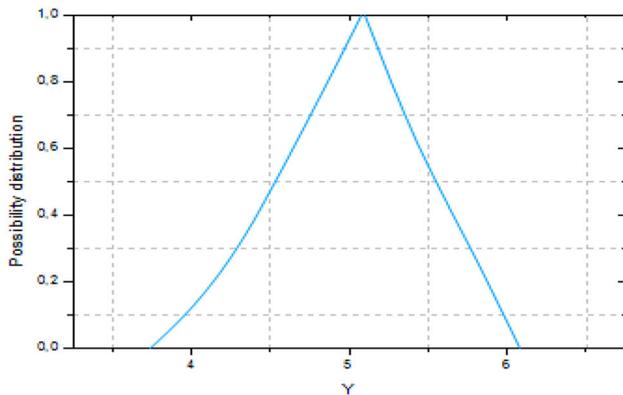


Fig. 7 Possibility distribution of the combined parameter and structural uncertainty of LCC prediction model output

parameter. Parameter variations are illustrated individually for each of the three model structures M_1 , M_2 , and M_3 . The most complex model structure generally shows a higher sensitivity of parameters. M_1 and M_2 have given, almost, the same results. On the other hand, parameters in M_3 are relatively sensitive compared to M_1 and M_2 . According to these differences, structural uncertainty plays an important role in the sensitivity analysis and should not be overlooked as part of overall uncertainty reductions. The overall contribution of spectral, shape, and NDVI parameters to the LCC prediction model, which are the highest and the indicative of the most sensitive for the three model structures. After applying the sensitivity analysis process, we will only consider these parameters for preprocessing based on the knowledge base and for optimal parameter estimation. Then, the uncertainty propagation based on possibility theory method is applied to reduce the parameter and structural uncertainty of the LCC prediction model.

3.1.4 Results of LCC prediction maps

LCC prediction maps are validated based on temporal series of multispectral SPOT images. First, the 2011 LCC was simulated using the 2006 data sets. Then, the simulated changes are compared with the real LCC in 2011 to evaluate the accu-

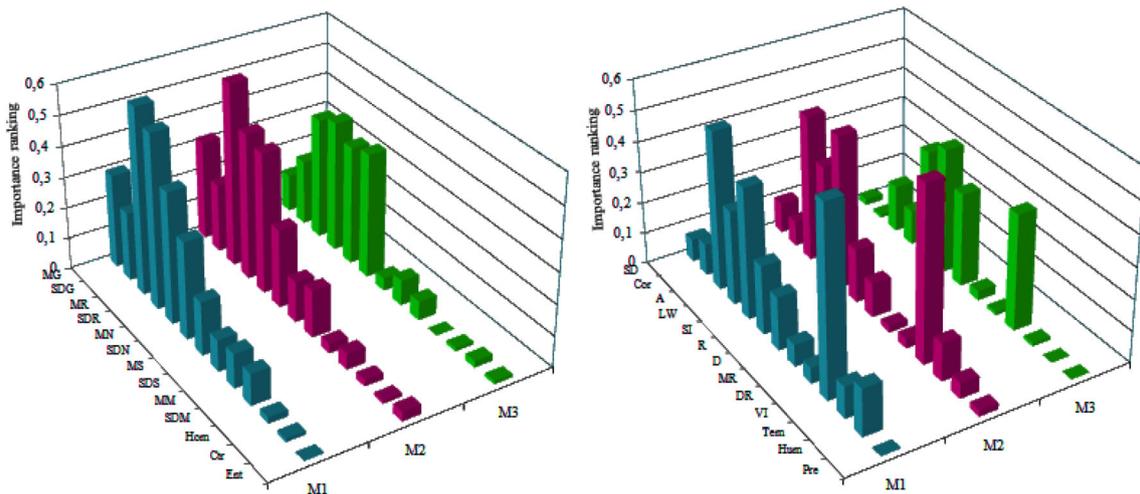


Fig. 8 Comparison between the sensitivity of uncertain parameters in three different LCC prediction model structures based on Hartley-like measure

Table 1 Percentages of LCC of the actual and simulated LCC

	Water (%)	Urban (%)	Forest (%)	Bare soil (%)	Vegetation (%)
Predicted changes in 2025	1.9	37.4	39.31	26.95	26.7
Output of proposed model	1.5	23.18	35.97	22.87	20.08
Real changes in 2011	1.7	21.4	36.1	24.1	16.7

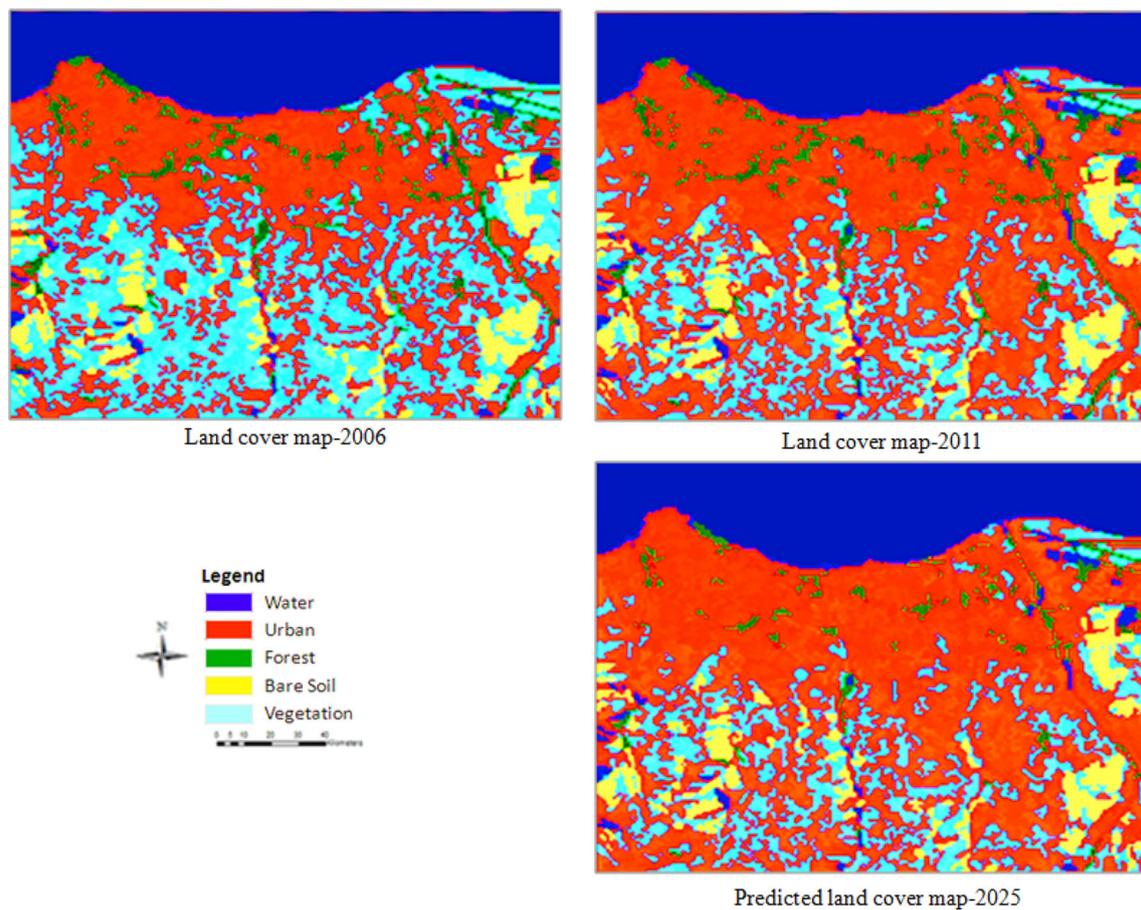


Fig. 9 Comparison between the land-cover maps for years 2006 and 2011 and the predicted land-cover map for 2025

racy and the performance of the proposed approach. Second, the process of LCC is conducted to predict land-cover distributions for forthcoming dates.

Table 1 illustrates a comparison between the actual and simulated percentages occupied by the different land-cover types (water, urban, forest, bare soil, and vegetation) between 2006 and 2011. It shows that the modeled changes generally matched that of the actual changes. These results confirm that the proposed approach can simulate the prediction of LCC with an acceptable accuracy.

After the validation, the next step is to simulate the LCC in 2025, assuming the changes between 2006 and 2011. In this simulation, the LCC and the parameters acquired in 2011 are used as input to simulate the LCC in 2025.

Table 1 shows the simulated changes between 2006 and 2025. Urban expansion is the dominant change process. This can be attributed to the increase in population by increased demands for residential land. There have been significant LCC, where urban land covered 21.4 % of simulated changes in 2011 and 37.4 % in 2025. From these results, it can be found the replacing of the land natural cover (forest and veg-

etation lands) in the study area by residential land (urban land).

Figure 9 depicts the simulated future changes compared with land-cover maps for the years 2006 and 2011.

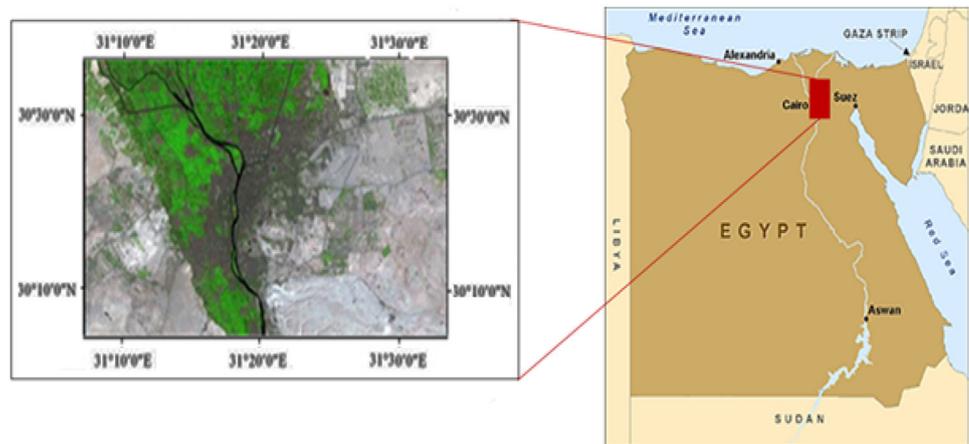
3.1.5 Evaluation of the proposed approach

To evaluate the proposed approach in improving LCC prediction, we apply the proposed uncertainty propagation approach on the LCC model described by Qiang and Lam in [40] to the Saint-Denis city, Reunion Island. The LCC prediction model proposed in [40] uses the Artificial Neural Network (ANN) to derive the LCC rules and then applies the Cellular Automate (CA) model to simulate future scenarios.

Table 2 depicts the percentages of change of the five land-cover types (water, urban, bare soil, forest, and non-dense vegetation). It shows the difference between real changes, predicted changes of the proposed approach, and changes made by the proposed approach applied to model described in [40].

Table 2 Comparison between real changes, predicted changes of the proposed approach, and changes made by the proposed approach applied to model described in [40]

	Water (%)	Urban (%)	Forest (%)	Bare soil (%)	Vegetation (%)
Proposed approach	1.5	23.18	35.97	22.87	20.08
Approach applied to model in [40]	1.5	25.32	34.98	20.03	16.24
Real changes in 2011	1.7	21.4	36.1	24.1	16.7

Fig. 10 Location of the study area for the case study 2

3.2 Case study 2

3.2.1 Description of the study area and data

Cairo, the capital of Egypt, is one of the most crowded cities in Egypt (Fig. 10) and is considered as a world megacity. Mapping LCC is important to understand and analyze the relationships between the geomorphology (highlands and deserts), natural resources (agricultural lands and the Nile River), and human activities. Agricultural lands around Cairo have witnessed severe encroachment practices due to the accelerated population growth. However, adjacent desert plains have also witnessed urbanization practices to encompass the intensive population growth. Different studies have previously been carried out for LCC detection and modeling in the Cairo Region [36–39]. Population of Cairo (Cairo city and Giza) increased from about 6.4 millions in 1976 [36] to about 12.5 million in 2006 according to the Egyptian Central Agency for Public Mobilization and Statistics. The importance of Cairo arises from its location in the mid-way between the Nile Valley and the delta. Main government facilities and services occur at Cairo.

In this case, two Landsat TM5 satellite images are obtained from the United States Geological Survey (USGS) database online resources. These two images acquired in 6 April 1987 and 15 March 2014, respectively, are classified into four land-cover types which are urban, agriculture,

desert, and water to produce LCC maps (Fig. 11). During this time period, Cairo population has increased from an estimated 7 million in 1987 to over 15 million in 2014. The recent population growth has caused the city and its associated urban areas to expand into the surrounding desert, as seen in the right image in Fig. 11. Within the main Nile River Valley, these two images also show an overall increase in developed urban area (red) versus agricultural land (green). As new urban and agricultural areas are being developed in the desert, they require diversion of water supplies from the main Nile River Valley.

In this case study, satellite data are classified after initial corrections and processing to prepare the data for extracting useful information. Spectral, geometric, and atmospheric corrections of images are conducted to make features manifest, to increase the quality of images, and to eliminate the adverse effects of light and atmosphere.

3.2.2 Results of uncertainty propagation

As we mentioned in the first case study, it is necessary to study the effect of both uncertainty sources through LCC prediction model. Figure 12 shows the possibility distribution of output representing only parameter uncertainty based on 10,000 samples. Therefore, it is also important to illustrate the importance of structural uncertainty in LCC prediction modeling by proposed approach. Figure 13 shows

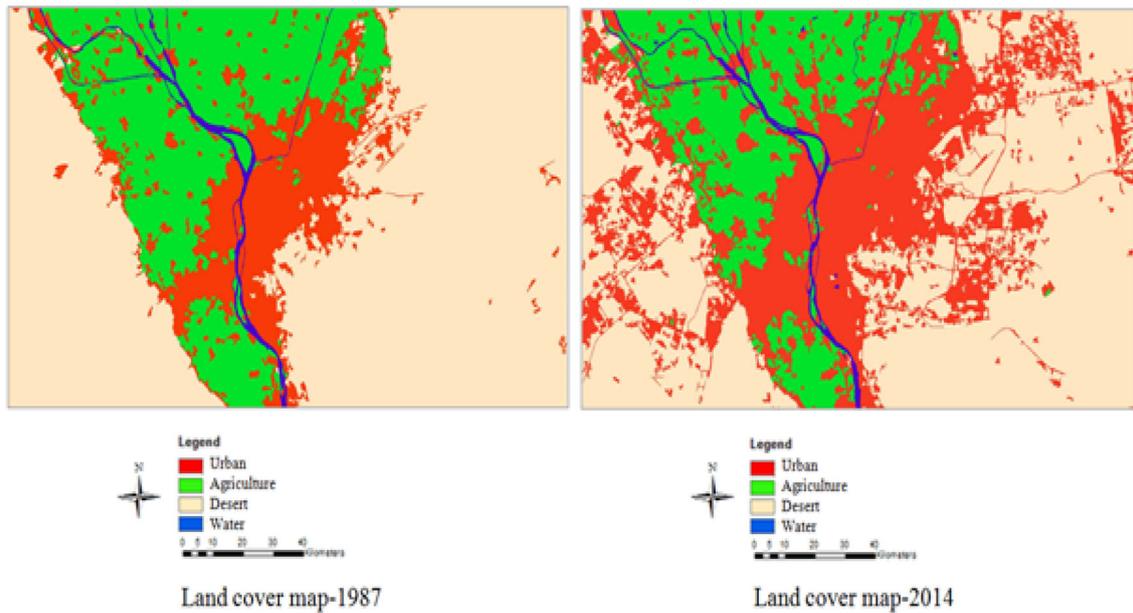


Fig. 11 Land-cover maps

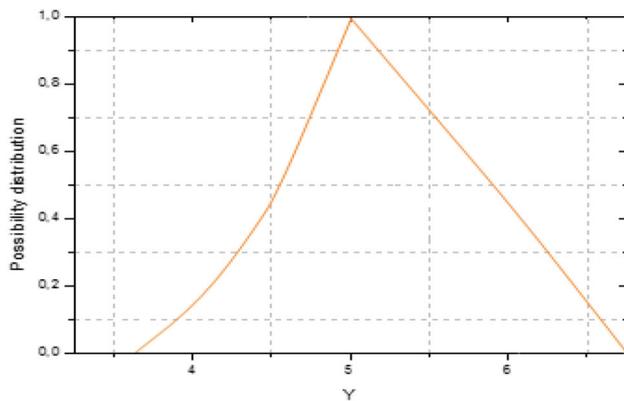


Fig. 12 Possibility distribution of LCC prediction model output for only parameter uncertainty

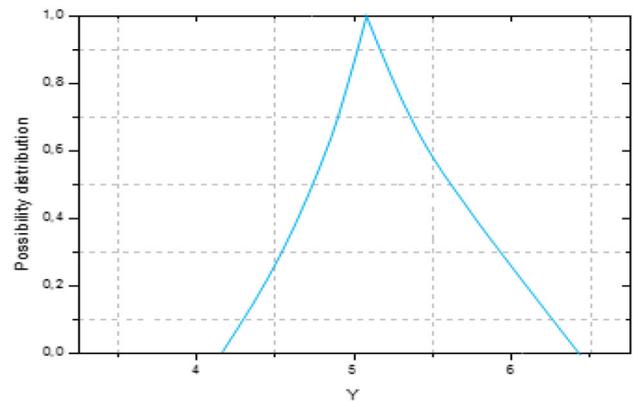


Fig. 14 Possibility distribution of the combined parameter and structural uncertainties of LCC prediction model output

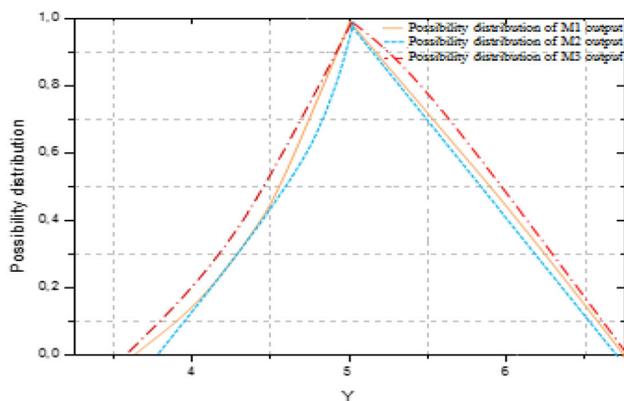


Fig. 13 Possibility distributions of LCCs for three different prediction model structures

the possibility distributions of output representing parameter uncertainty through three different model structures. The difference between these three representations presents the impact of structural uncertainty. In this case study, results of the uncertainty propagation of both parameter and structural uncertainties are shown in Fig. 14.

3.2.3 Results of sensitivity analysis

In this case study, we have also used Hartley-like measure to estimate the 26 uncertain parameters through three different LCC prediction model structures. The main objective is to test the impact of parameter and structural uncertainties. Results of the sensitivity analysis are shown in Fig. 15. In this case, parameters in M_3 are highly sensitive com-

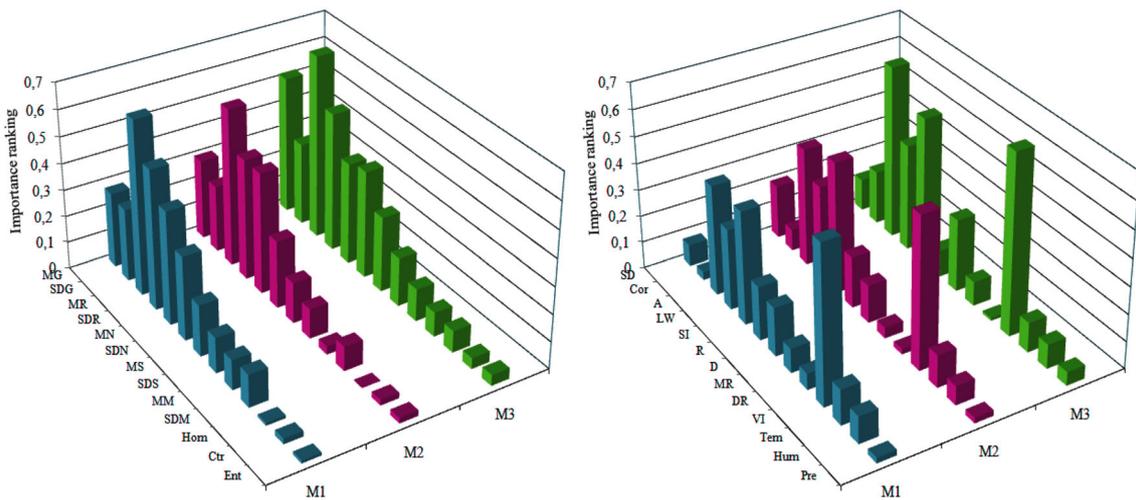


Fig. 15 Comparison between the sensitivity of uncertain parameters in three different LCC prediction model structures based on Hartley-like measure

Table 3 Output of the proposed LCC prediction model in comparison with real changes between 1987 and 2014 in Cairo region

	Urban (%)	Agriculture (%)	Water (%)	Desert (%)
Output of proposed model	15.63	13.80	0.01	4.03
Real changes in 2014	17.32	13.00	0.02	5.00

Table 4 Output of the proposed LCC prediction model of the predicted LCC between 2014 and 2025 in Cairo region

	Urban (%)	Agriculture (%)	Water (%)	Desert (%)
Predicted changes in 2025	20.16	14.72	0.03	6.11
Real changes in 2014	17.32	13.00	0.02	5.00

pared to M_1 and M_2 . On the other hand, as in the first case study, the overall contribution of spectral, shape, and NDVI parameters to the LCC prediction model is the highest and represents the most sensitive parameters for the three model structures.

3.2.4 Results of LCC prediction maps

The validation of LCC prediction maps consists of two phases. First, the 2014 LCC is simulated using the 1987 data sets, which is then compared with the real LCC in 2014 to evaluate the accuracy and the performance of the proposed approach. Second, future changes are simulated using the real 2014 data sets.

To check the accuracy of our approach, Table 3 compares actual and simulated percentages occupied by the different land-cover types (urban, agriculture, water, and desert) between 1987 and 2014. According to the proposed model output, the most significant changes in this period are the transitions from agriculture and desert to urban areas (Fig. 11). Over 27 years, from 1987 to 2014, agriculture lost 12 % to urban areas. In addition, 4 % of desert areas became urban between 1987 and 2014, which is equivalent to 24,687

hectares. This percentage results from the application of desert reconstruction strategies to build new communities outside the Nile Valley. The obtained results depict that the proposed approach gives an accurate prediction with about 3.96 % of error through a comparison with the real changes in Cairo region. These results confirm that the proposed LCC prediction model is able to describe the LCC. The proposed approach can simulate the prediction of LCC with an acceptable accuracy. After the validation of the proposed approach, the next step is to simulate the LCC in 2025, assuming that the changes between 1987 and 2014 will continue during the next 11 years. In this simulation, the LCC and the input parameters acquired in 2014 are used as input to simulate the LCC in 2025.

Table 4 shows the simulated changes between 2014 and 2025. There have been significant LCC, where urban land covered 15.63 % of simulated changes in 2014 and 20.16 % in 2025. This could be attributed to the increase in population by increased demands for residential land. The resulting effect is the decrease in desert land. From these results, we note that the desert land cover in the study area is replaced by residential land (urban land). Knowing the current and estimated urbanization situation will help decision makers to

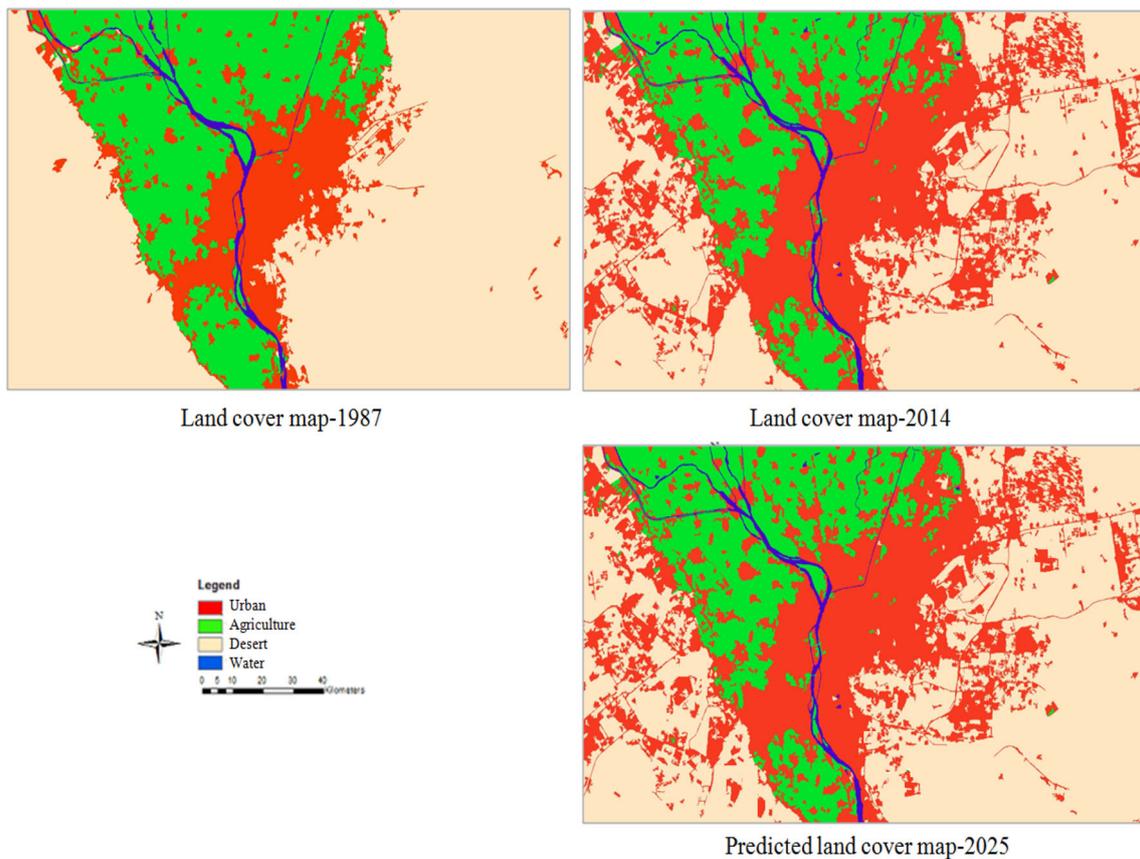


Fig. 16 Comparison between the land-cover maps for years 1987 and 2014 and the predicted land-cover map for 2025

Table 5 Comparison between real changes and changes prediction for the proposed approach and the proposed approach applied to model described in [40]

	Urban (%)	Agriculture (%)	Water (%)	Desert (%)
Proposed approach	15.63	13.80	0.01	4.03
Approach applied to model in [40]	14.93	13.61	0.01	5.92
Real changes in 2014	17.32	13.00	0.02	5.00

adjust and develop new plans to achieve a sustainable development of urban areas and to protect the historical locations.

Figure 16 depicts the simulated future changes compared with land-cover maps for the years 1987 and 2014. These results indicate usefulness and applicability of the proposed approach in predicting the LCC.

3.2.5 Evaluation of the proposed approach

In this case study, we also apply the proposed uncertainty propagation approach on the LCC model described by Qiang and Lam in [40] to the Cairo region, Egypt.

Table 5 depicts the percentages of change of the four land-cover types (urban, agriculture, water, and desert). This table shows the difference between real changes and

changes prediction for the proposed approach and the proposed approach applied to model described in [40].

4 Conclusion

This study has proposed an approach for reducing parameter and structural uncertainty in LCC prediction modeling. The proposed approach herein quantifies uncertainty based on possibility theory. Subsequently, the Hartley-like measure is used to perform the sensitivity of the LCC prediction model parameter and structure. Using the sensitivity analysis, we are able to quantify precisely the effect of each LCC prediction model parameter, and also the effect of model structure. This analysis yields that the spectral, shape, and vegetation parameters are the most sensitive parameters in three different model structures.

To validate the proposed approach, we choose two case studies which are: Saint-Denis city, Reunion Island, and Cairo region, Egypt. We study spectral parameters, texture parameters, shape parameters, vegetation parameter, and climate parameters for three different model structures to simulate forthcoming LCC. Results show that the urban expansion in the two case studies is rapid and should be monitored in the future.

As future work, we propose to put online a tool for uncertainty propagation and sensitivity analysis based on possibility theory. This tool will help researchers to improve the performance of their models. It has also as input parameters of a considered model and as output which of these input parameters that most influence the model output.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided 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.

References

- Boulila, W., Bouatay, A., Farah, I.R.: A probabilistic collocation method for the imperfection propagation: application to land cover change prediction. *J. Multimed. Process. Technol.* **5**(1), 12–32 (2014)
- Liu, R., Sun, J., Wang, J., Li, X.: Study of remote sensing based parameter uncertainty in production efficiency models. In: *IEEE International Geoscience and Remote Sensing Symposium*, pp. 3303–3306 (2010)
- Jacquín, A.P.: Possibilistic uncertainty analysis of a conceptual model of snowmelt runoff. *Hydrol. Earth Syst. Sci.* **14**, 1681–1695 (2010)
- Dubois, D., Prade, H.: Formal representations of uncertainty. In: *Decision-Making Process: Concepts and Methods*, ch. 3, pp. 85–156. ISTE and Wiley, London (2010)
- Caselton, W.F., Luo, W.: Decision making with imprecise probabilities: Dempster–Shafer theory and application. *Water Resour. Res.* **28**(12), 3071–3083 (1992)
- Zhao, X., Stein, A., Chen, X.: Application of random sets to model uncertainties of natural entities extracted from remote sensing images. *Stoch. Environ. Res. Risk Assess.* **24**(5), 713–723 (2009)
- Hgarat-Masclé, S.L., Bloch, I., Vidal-Madjar, D.: Application of Dempster–Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Trans. Geosci. Remote Sens.* **35**(4), 1018–1031 (1997)
- Gomez, D., Javier, M.: Fuzzy sets in remote sensing classification. *Soft Comput.* **12**(3), 243–249 (2008)
- Ferchichi, A., Boulila, W., Farah, I.R.: An intelligent possibilistic approach to reduce the effect of the imperfection propagation on land cover change prediction. *Int. Conf. Comput. Collect. Intell.* **9330**, 520–529 (2015)
- Mondal, M.S., Garg, P.K., Sharma, N., Kappas, M.: Cellular automata (ca) markov modeling of LULC change and sensitivity analysis to identify sensitive parameter(s). In: *Proceedings of the 27th International Cartographic Conference*, vol. 38(818) (2015)
- Sanchez-Canales, M., Benito, A.L., Passuello, A., Terrado, M., Ziv, G., Acuna, V., Schuhmacher, M., Elorza, F.J.: Sensitivity analysis of ecosystem service valuation in a Mediterranean watershed. *Sci. Total Environ.* **440**, 140–153 (2012)
- Bettemier, O.H.: Error estimation of orthorectification of small satellite images by differential sensitivity analysis. *J. Aeronaut. Space Technol.* **4**(4), 65–74 (2010)
- Zielinskaa, A.L., Sunb, L.: Applying time-dependent variance-based global sensitivity analysis to represent the dynamics of an agent-based model of land use change. *Int. J. Geogr. Inf. Sci.* **24**(12), 1829–1850 (2010)
- Helton, J.C., Johnson, J.D., Sallaberry, C.J., Storlie, C.B.: Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliab. Eng. Syst. Saf.* **91**(10–11), 1175–1209 (2006)
- Li, C., Wang, W., Xiong, J., Chen, P.: Sensitivity analysis for urban drainage modeling using mutual information. *Entropy* **16**, 5738–5752 (2014)
- Sengupta, A., Pal, T.K.: Theory and methodology: on comparing interval numbers. *Eur. J. Oper. Res.* **127**, 28–43 (2000)
- Ferson, S., Tucker, W.T.: Sensitivity analysis using probability bounding. *Reliab. Eng. Syst. Saf.* **91**(1011), 1435–1442 (2006)
- Boulila, W., Farah, I.R., Ettabaa, K.S., Solaiman, B., Ben Ghzala, H.: A data mining based approach to predict spatio-temporal changes in satellite images. *Int. J. Appl. Earth Obs. Geoinf.* **13**(3), 386–395 (2011)
- Ballestores, F., Jr., Qiu, Z., Nedorezova, B.N., Nedorezov, L.V., Ferrarini, A., Ramathilaga, A., Ackah, M.: An integrated parcel-based land use change model using cellular automata and decision tree. In: *Proceedings of the International Academy of Ecology and Environmental Sciences*, vol. 2(2), pp. 53–69 (2012)
- Razavi, B.S.: Predicting the trend of land use changes using artificial neural network and markov chain model (Case Study: Kermanshah City). *Res. J. Environ. Earth Sci.* **6**(4), 215–226 (2014)
- Tayyebi, A., Perry, P.C., Tayyebi, A.H.: Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster/vector routines with remote sensing and GIS. *Int. J. Geogr. Inf. Sci.* **28**(4), 1–21 (2013)
- Ralha, C.G., Abreu, C.G., Coelho, C.G., Zaghetto, A., Macchiavello, B., Machado, R.B.: A multi-agent model system for land-use change simulation. *Environ. Model. Softw.* **42**, 30–46 (2013)
- Li, X., Yeh, A.G.O.: Neural-network-based cellular automata for simulating multiple land use changes using GIS. *Int. J. Geogr. Inf. Sci.* **16**(4), 323–343 (2002)
- INSEE Reunion: Bilan démographique 2009, Resultats (40) (2011)
- Durieux, L., Lagabrielle, E., Andrew, N.: A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS J. Photogramm.* **63**, 399–408 (2008)
- van der Valk, A.: The Dutch planning experience. *Landsc. Urban Plan.* **58**(2), 201–210 (2002)
- Cadet, T.: La végétation de l'île de La reunion, tude phytocologique et phytosociologique. Ph.D. Thesis, University of Aix Marseille (1980)
- Atanassov, V., Jeleu, G., Kraveva, L.: Some peculiarities of the pre-processing of spectral data and images. *J. Ship. Ocean Eng.* **3**, 55–60 (2013)
- Jimenez-Munoz, J.C., Sobrino, J.A.: Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. *Int. J. Remote Sens.* **27**(5), 999–1014 (2006)
- Dubois, D., Prade, H.: When upper probabilities are possibility measures. *Fuzzy Sets Syst.* **49**, 65–74 (1992)
- Klir, G.J., Wiermann, M.J.: *Uncertainty Based Information. Elements of Generalised Information Theory*. Physica-Verlag, Heidelberg (1998)

32. Hagashi, M., Klir, G.J.: Measure of uncertainty and information based on possibility distribution. *Int. J. Gen. Syst.* **9**, 43–58 (1983)
33. Chutta, R., Mahanta, S., Datta, D.: Sensitivity analysis of atmospheric dispersion model-Rimpuff using Hartley-like measure. *J. Appl. Math. Inform.* **31**(1–2), 99–110 (2013)
34. Datta D.: Measures of sensitivity and uncertainty with the fuzzy model of solute transport through groundwater. In: International Conference on Communication and Industrial Application, pp. 1–6 (2011)
35. Ivanov, L.M., Tokmakian, R.T.: Sensitivity analysis of nonlinear models to parameter perturbations for small size ensembles of model outputs. *Int. J. Bifurc. Chaos* **21**(12), 3589–3609 (2011)
36. Sutton, K., Fahmi, W.: Cairo's urban growth and strategic master plans in the light of Egypt's 1996 population census results. *Cities* **18**(3), 135–149 (2001)
37. Yin, Z.Y., Stewart, D.J., Bullard, S., Maclachlan, J.T.: Changes in urban built-up surface and population distribution patterns during 1986–1999: a case study of Cairo, Egypt. *Comput. Environ. Urban Syst.* **29**, 595–616 (2005)
38. Hereher, M.E.: Analysis of urban growth at Cairo, Egypt using remote sensing and GIS. *Nat. Sci.* **4**, 355–361 (2012)
39. de Noronha, V.E., Caetano, M., Nijkamp, P.: Trapped between antiquity and urbanism—a multi-criteria assessment model of the Greater Cairo Metropolitan Area. *J. Land Use Sci.* **6**, 283–299 (2011)
40. Qiang, Y., Lam, N.S.N.: Modeling land use and land cover changes in a vulnerable coastal region using artificial neural networks and cellular automata. *Environ. Monit. Assess.* **187**(57), 1–16 (2015)