Review of: "Exploring the Impact of Future Land Uses on Flood Risks and Ecosystem Services, With Limited Data: Coupling a Cellular Automata Markov (CAM) Model, With Hydraulic and Spatial Valuation Models"

Hamidou Kassogué

Potential competing interests: No potential competing interests to declare.

The paper explores the use of CAM combining GIS and Python codes to predict land use change (2026 to 2051). The obtained future land uses are coupled with hydraulic and spatial valuation models to evaluate flood risks and ecosystem services, which constitutes a novelty. The paper is well written, and the methodology is well presented. Data and codes are available for readers for their own implementation.

I think this paper is good for publication. However, addressing some replies to the following remarks/suggestions and/or questions could enhance the paper quality.

Remarks/suggestions

- **Title:** The title is too long. Contract it if possible.
- **Keywords**: Sort the keywords alphabetically and reduce the number if possible.
- **Data and codes availability:** At line 72 of futurelands.py, the author defines a variable for 2016 land use with the file “cedar16reclas1.tif,” which does not exist in the data folder. Is it for the reader to make this file from files in the cedar2016lu folder? If yes, how can a reader do that? The author could give more practical details. The same observations are valid for lines 133-134, 218-223, and 301-306. Although the reader could replace them with his own values, I think it will be interesting to make available the author's intermediary generated data for the reader in order to check directly and/or quickly the approach.
- **Figures**: It is difficult to read Figure 1. The author could use separated images with better resolution for each land use and put them as sub-figures on a 2x2 grid. The legend could be placed at the bottom.
  The same remark is valid for Fig. 2. The author could use extracted images with better resolution instead of full screenshots of the GIS displaying the concerned data.
  In Fig. 3, it is difficult to see differences in 2026, 2031, 2036, 2041, 2046, and 2051 land use. It is necessary to find mechanisms to see difference since one of the paper goals is predicting future land use change. The author can dispose these land use changes in a 3x2 grid with a larger size and, if possible, zoom on some critical zones to better show how land use changes from one year to another one.
  Same remark on Figs. 5 and 7.
- **Page 12, text after Fig. 5**: Check the figures referenced in the sentence “The simulation results are presented in

Figure 5 and the zonal statistics for each case are shown in Figure 5."

- **Page 2, paragraph 2:** The author refers to (Aburas et al., 2019) to support the sentence "The most common technique to explore future land cover is the Cellular Automata (CA), based on Markov chain modelling, called Cellular Automata Markov (CAM) models," while the paper's goal is not on CAM. The paper concludes from the literature that (page 204) "the CA model is the most common dynamic technique used in the last 10 years." The same assertion is made in the author note based on CAM (page 1 of Alamanos, A. (2023)) while referring to (Aburas et al., 2019). I think the author must remain faithful to the conclusion of (Aburas et al., 2019) or cite a paper that gives such a conclusion.
- **Page 2, paragraph 2**: The author does not mention boundary conditions and cells factors/attributes in the sentence "They are based on transition rules (e.g., changes over a period of time), and initial conditions (i.e., compared to an initial base-year)." I think it is useful to identify the necessary ingredients (cells structuration, attributes/factors, state, neighborhood, transition rules, initial condition, boundary conditions) on which CA evolution depends. One can read further, in page 6, section 3.2, such ingredients in this sentence "By discretizing the terrain into a grid, the software calculates water depth, velocity, and discharge in each cell while considering factors like channel geometry (from the DEM), boundary conditions (e.g., outlet), and the effects of the rainfall."
- **Section 3.1:** In the idea of the above remark, the author can describe more precisely the CAM model by addressing the lattice or cells structuration (square, triangular, hexagonal?), the type of neighborhood, cells attributes/factors, the states set, … See, for example, the section 3 of this paper (https://doi.org/10.1016/j.apm.2016.10.034).
- **Recent works on LULC and/or CAM**: There are several recent works on LULC that the author could cite to highlight his novelty:
  https://doi.org/10.3390/rs15041162
  https://doi.org/10.3390/land12091788
  https://doi.org/10.1016/j.rsase.2022.100774
  https://doi.org/10.3390/su15043572
  https://doi.org/10.1007/s42452-023-05520-7
  https://www.mdpi.com/2072-4292/13/16/3337
  https://link.springer.com/article/10.1007/s10661-023-11224-7
  This research link (https://www.connectedpapers.com/search?q=lulc%20cam) gives a connected papers graph from which I picked the above.

Questions:

1. Did the author use limited data (five land use classes) in his approach? Is the approach valid for many land use classes?
2. Has the author evaluated the model performance (time and memory use)? This can provide assurance when the lattice consists of cells with small resolution (a few meters order).
3. Equation 3 (Overall Accuracy) is for measuring the validation of the predicted land use maps. How does the author perform it practically? Is, for example, the author predicts the 2011 land use change from the 2006 one and performs this accuracy? If yes, how does this approach assure future land use change validation (from 2021 to 2051 by 5 years

as a time step)?

4. In addition, the author has only 2006, 2011, 2016, and 2021 land use maps (although the 2006 one is not present in the provided data). Are these 5 land use maps sufficient to determine efficiently both the transition probability matrix and overall accuracy? In the Python code, I can read the transition probability matrix (by direct input) for 2006 to 2011, 2011 to 2016, and 2016 to 2021. However, the matrix has not been used in the rest of the code. In lines 75-81, another transition probabilities matrix has been given directly for the 2006 to 2021 prediction. How did the author obtain this matrix, and how did he use it concretely in the code?
5. The questions on Eq.3 are available for Eqs. 4 and 5.