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

Angelos Alamanos

ABSTRACT

Land use changes can majorly affect many parameters that are directly or indirectly interlinked to various human-environmental systems, including hydrological processes and flood risks. The knowledge of future land cover changes is crucial for better managing human-environmental interactions and addressing potential environmental challenges, such as floods. In this work, the impact of future land cover changes in flood inundation is assessed, using a case study in northeast Indiana, US. A Cellular Automata Markov (CAM) model is applied, combining Geographic Information Systems (GIS) and Python, to predict land changes and provide future land cover maps, along with statistical validation measures. The land use map outputs are then used in a HEC-RAS hydraulic model, to test the different flooding impacts under a design storm, using the rain-on-grid routine. The results indicate that even slightly more urbanized and deforested areas can increase the potential flood extent. Furthermore, the impacts of these forecasted land cover changes are quantified in monetary terms, based on a spatial Ecosystem Services Valuation (ESV) model. The findings indicate that as certain land uses (mainly wetlands, followed by forests) give their place to build-up areas, barren land, or even agricultural lands, the ‘lost’ value due can reach 1.5 million USD in 2051. The novelty of this study lies in its integrated character, combining for the first time to our knowledge land cover forecast with hydrologic-hydraulic modelling and spatial ESV, showing thus the future changes, risks, and potential economic losses, respectively. This application uses the minimum necessary input data to perform the analyses, and all data and codes are publicly available, contributing thus to the transferability and reproducibility of the approach.

Keywords: Land uses; Cellular Automata Markov model; GIS; Python; Hydraulic modelling; Rain-on-Grid; Flood; HEC-RAS; HEC-HMS; Ecosystem Services; Valuation.

1. INTRODUCTION

Studying the future alterations in land cover within specific areas is essential, given their intricate connections with various factors dynamically influencing human-environmental systems. Land cover changes exert a profound impact on urban planning, environmental sustainability, resource management, and overall quality of life (Liu et al., 2022; Hassan and Nazem, 2016). Understanding these changes is critical for making well-informed decisions that influence population growth, resource availability and utilization, infrastructure development, and the conservation of natural capital and biodiversity, among other aspects (McDermott et al., 2022). Knowledge of shifts in land use patterns empowers policymakers...

https://doi.org/10.32388/JJWWBD
to confront challenges like urban sprawl, deforestation, water flow dynamics - potentially exacerbating the risks of floods, and habitat loss (Hassan and Nazem, 2016; McDermott et al., 2022).

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 (Aburas et al., 2019). The logic of CA models is to simulate land cover changes by considering the local interactions between cells (geographic units in spatial datasets). 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). Markov chain models rely on the assumption that future land cover depends on the current state (base year map) and is independent of past states. Thus, CAM models use historical data, usually in the form of maps, to derive transition rules, namely transition probability matrices among the land use categories, and generate future maps by applying these matrices as rules, iteratively to the historic data (Corner et al., 2013). The field is fast developing, and more complex methodologies occur, such as combinations of CAM and Geographic Information Systems (GIS) with Remote Sensing observations (Islam and Ahmed, 2012), machine learning techniques (Xing et al., 2020; Zambrano-Asanza et al., 2023), and Multi-Criteria Analysis techniques (Addae and Dragičević, 2022), for improved prediction accuracies and/or the consideration of more factors in the analyses. The validation of the projections is performed statistically, comparing the historic data with the predicted ones, for the same year(s), and the most commonly used measure is the Kappa statistics, assessing the accuracy of the projections (Saputra and Lee, 2019; Liu et al., 2021). There have been several CAM modelling applications, for urbanization projections (Ulloa-Espíndola and Martín-Fernández, 2021; Mansour et al., 2020), for the evaluation of different development scenarios (Han et al., 2015), agriculture and biodiversity management (Halmy et al., 2015), the impact of climatic parameters (Tariq and Shu, 2020; Al Kafy et al., 2021), etc. Moreover, future land uses can be necessary for a variety of other analyses, such as soil and hydrological assessments (Anard et al., 2018; Dai et al., 2023), wetland management (Ansari and Golabi, 2019; Alamanos and Papaioannou, 2020), urban and rural development (Agustina et al., 2022; Alamanos et al., 2022a; 2022b), flood risk assessments (Roy et al., 2020; Papaioannou et al., 2023), ecological assessments (Qin and Fu, 2019), optimal agricultural management (Garcia and Alamanos 2022; 2023), management of transboundary environmental and economic assets (Englezos et al., 2023; Mendoza-Poce et al., 2021), and many more.

However, the impact of land cover changes in potential flood risks is still poorly understood (Rogger et al., 2017), and the need for models simulating such impacts has been long recognized (Tollan, 2002). There are studies exploring this topic, but mainly from the perspective of different land use scenarios or specific management practices (e.g. focusing on agricultural land uses), rather than land use predictions (Hounkpè et al., 2019; Saghafian et al., 2008). There are also very few studies examining the effect of multiple factors in future flooding, such as land use, climate, topography etc., as for example in the paper by Avand et al. (2021). There are also very limited applications where actual land use prediction models have been applied to investigate their impacts on flood risks (Roy et al., 2020), with the exception of the paper by Adnan et al. (2020) who developed a CAM model to predict future land uses and explore the associated future flooding implications. So far, the main tools for assessing floods in response to altered land uses have been SWAT, WaSiM, or other custom approaches, including machine learning (Hounkpè et al., 2019; Schilling et al., 2013; Avand et al., 2021). However, to our knowledge there has not been any application where predicted land uses are analyzed as part of a hydraulic modelling approach to showcase potential flood risks. This is one gap that this work aims to fill.
Another gap this work aims to bridge is the mapping of the monetary value of these land cover changes, in the future, based on the Ecosystem Services (ES) they can provide. Environmental valuation studies are based on the concept of Total Economic Values, considering the use values (direct, indirect, and option values) and nonuse values (existence and bequest values), of environmental assets (Koundouri et al., 2023). Practically, Ecosystem Services Valuation (ESV) assigns monetary values to environmental impacts/changes, reflecting the value of ES such as provisioning services (food, raw materials); supporting services (life-cycle maintenance for flora, fauna, biodiversity); regulating services (climate, carbon sequestration and storage, erosion prevention, nutrient treatment, moderation of extreme events); and cultural services (tourism, recreational, aesthetic, and spiritual benefits) (Koundouri et al., 2022). ES are closely related to land use changes, so linking them is critical for better land use planning and sustainable provision of ES (Fu et al., 2015). The importance of mapping the ES has long been recognized (Troy and Wilson, 2006), and value transfer methods are increasingly utilized, to increase the reproducibility of such studies (Tammi et al., 2017; Xue and Luo, 2015). Value transfer practically is the application of economic values derived from one study area to another context, often to estimate the economic value of ES (Koundouri et al., 2022). The ability of providing integrated assessments considering the ESV (Alamanos, 2021), especially as related to land cover changes, and their future evolution is rare in the literature, but crucial for strategic planning, decision prioritization and optimal investment allocation (Rajsic et al., 2023; Alamanos and Brouwer, 2020). In contrast to previous studies considering land cover changes trends and management scenarios (Schirpke et al., 2020), in this work, the CAM model’s forecasts are used as the basis for the ESV. There is a handful of applications considering CAM models to estimate ESV, including the paper by Zhang et al. (2021) focusing on urbanization effects, the study by Gao et al. (2021) exploring past alterations, the paper by Zhong et al. (2022) showing how CAM can be combined with InVEST Models, and Zhang et al. (2023) who combined CAM with system dynamics modelling to forecast the ES values under alternative scenarios.

In this paper, a CAM model is presented, as a combination of processes in GIS environment and open-source coding, using Python. A hydraulic model is developed in HEC-RAS software to produce flood inundation maps under the different future land uses, for a watershed that has received limited attention with respect to floods. The ESV is performed based on a value transfer method from established values found in the literature. All the future values (land cover, flood risks, ESV) refer to 2051. Showing information of the impacts of land use changes in future exposure to natural phenomena (e.g. flooding), together with their evolution of economic value (e.g. ESV) is presented for the first time to our knowledge, and is a comparative advantage to previous studies, as the results provide significant insights for informed and holistic decision-making regarding future planning and sustainable landscape management.

Another contribution of the presented approach is that it can work effectively with limited data: the CAM tool uses the minimum number of inputs, namely historic land use maps, and generates future land use maps. The future flooding estimations are based also on the minimum necessary input, considering a design-storm under a rain-on-grid approach. Finally, the value transfer approach considers the values obtained from the existing literature. Thus, the overall framework can be easily applied in other case studies, even with limited data.
2. STUDY AREA

The Cedar Creek Watershed (CCW) in Indiana, US, is used as a case study. CCW in northeastern Indiana is an area within the St. Joseph River watershed, with a diverse landscape of farms, urban areas with cities and settlements, and notable geological features (CCW Management Plan, 2005). The land uses have been traditionally agricultural, with a slight increase of urban areas, while forests and water bodies are also present (Figure 1). The degree of urbanization in the CCW has slightly increased over the past 15 years, as rural areas transform into suburban or urban spaces, accompanied by residential areas and transportation networks.

Figure 1. The land uses of CCW from 2006 to 2021, classified in five main categories.

Positioned just north of Fort Wayne, Indiana’s second-largest city, Cedar Creek flows into the St. Joseph River, where Fort Wayne sources its drinking water downstream. CCW covers approximately 700 km²; it has a gentle sloping topography and is primarily used for agriculture (mainly corn, soybeans, and other crops) (Pignotti et al., 2017). The region experiences an average temperature range of -1 to 28°C with an
annual precipitation of 940 mm (Wallace et al., 2018). The greater CCW area has experienced some flood events, in the past. The flood of 1982 (Glatfelter and Chin, 1987), and the flood of 2009 (Fowler, 2017) are the most significant ones (Bassett et al., 2009).

The CCW has been studied extensively from a hydrologic and soil assessment point of view focusing on streamflows and sediment (Larose et al., 2007; Jiang et al., 2008; Kumar and Merwade, 2009; Kang and Merwade, 2011; Pathak and Kalra, 2015), with HEC-HMS and SWAT models. However, the area has received little attention regarding its land use evolution, as well as flood risks, and to our knowledge this is the first paper presenting a land use prediction model, and a hydraulic model for CCW.

3. METHODOLOGY

3.1 The Cellular Automata Markov (CAM) model

Creating a Cellular Automata Markov (CAM) model for land use change prediction involves the estimation of the transition probability matrix, and simulation of land use changes over time. The CAM model can be mathematically represented as follows (Equation 1):

\[ L_{t+1} = P_{ij} \cdot L_t \], for land use types \( i, j = 1,2,\ldots,n \)  

(1)

Where \( L_t \) and \( L_{t+1} \) are the land use maps at the year \( t \) and \( t+1 \) respectively, and \( P_{ij} \) is the transition probability matrix expressing the probability of each cell (pixel) to change from the land use type \( i \) in the year \( t \) to the land use type \( j \) in year \( t+1 \). So, this matrix can be expressed as shown in Equation 2 below:

\[
P_{ij} = \begin{bmatrix}
P_{11} & P_{12} & \cdots & P_{1n} \\
P_{21} & P_{22} & \cdots & P_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
P_{n1} & P_{n2} & \cdots & P_{nn}
\end{bmatrix}, \text{ with } 0 \leq P_{ij} \leq 1, \text{ and } \sum_{i,j=1}^{n} P_{ij} = 1
\]

(2)

The transition probability matrix is estimated by assessing the changes among land uses from different years \( (t, t+1) \). This process can be done in a spatial analysis software first by determining the number of pixels in a land-use class \( i \) that changed into another class \( j \) during that time.

The input data of land uses were obtained from the United States Geological Survey (USGS) website (USGS, 2021), as shape files for the years 2006, 2011, 2016 and 2021, following the National Land Cover Database (NLCD) categorization. The main land use categories were grouped for simplicity in classes of: 'Water': 1, 'Urban': 2, 'Barren Land': 3, 'Forest': 4 and 'Crops': 5. This (indicative for this example) approach significantly reduces the computational load and effort, and allows to handle easier the land use changes from year to year. The CAM model was used to generate the future land use maps for every five years until 2051. The validation of the predicted land use maps was achieved by the following statistical tests:

- **Percentage of Accurate Results (Overall Accuracy):** This is a straightforward measure that calculates the percentage of correctly classified pixels compared to the total number of pixels, according to the relation expressed in Equation 3:

\[
Accuracy = \frac{\text{Number of Correctly Classified Data Points}}{\text{Total Number of Data Points}} \times 100\% 
\]

(3)
• Mean Absolute Error (MAE): Expressing the magnitude of errors between predicted and actual land use values (Equation 4):

\[
MAE = \frac{1}{Total\ Number\ of\ Data\ Points} \sum_{point\ 1}^{point\ N} |actual - predicted|
\]  

(4)

• Root Mean Square Error (RMSE): This measures the square root of the average squared difference between predicted and observed values, penalizing thus larger errors more heavily (Equation 5):

\[
RMSE = \sqrt{\frac{1}{Total\ Number\ of\ Data\ Points} \sum_{point\ 1}^{point\ N} (predicted - actual)^2}
\]  

(5)

• Kappa (κ): Kappa coefficient measures the level of agreement between two rasters or classifiers, often used in the context of classification tasks like land use mapping. It quantifies the agreement between the predicted and true labels while taking into account the possibility of agreement occurring by chance (Equation 6):

\[
\kappa = \frac{Po - Pe}{1 - Pe}
\]  

(6)

Po is the observed agreement between the predicted and true land use categories. It represents the proportion of instances where the predicted and true labels match. Pe is the expected agreement between the predicted and true land use categories, if their agreement were purely due to random chance.

3.2 The Hydraulic Model

For the hydraulic model, the Digital Elevation Model (DEM) of CCW is required to delineate the watershed into sub-basins and define all the necessary elements for the spatial analysis of the data. This process can be done in a GIS software, e.g. with tools such as ArcHydro (ESRI, 2014) or GeoHEC-HMS (HEC, 2023). However, the recent versions of HEC-HMS and HEC-RAS (HEC, 2022) provide GIS tools for doing such analyses within the software. In this case, HEC-HMS was used to create the basin model and delineate it, processing the drainage, producing the reaches, junctions and outlet (streams) (Figure 2A). The DEM of the area was retrieved from the Open Topography website (OpenTopography, 2023). The outputs of the HEC-HMS can be exported and used as inputs in the HEC-RAS software, where further refining can be done within the RAS-Mapper environment. The model with its ‘Geometry’ is developed there (Figure 2B), with the 2D Flow Area grid and its computational mesh (Figure 2C). A rain-on-grid approach was followed for the analysis as a way to easily show the impact of rainfall in the gridded terrain. The method leverages numerical methods to solve the Saint-Venant equations, which describe the conservation of mass and momentum in open channel flows, solving them in every cell of the computational mesh (HEC, 2022). 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.
Figure 2. The development of the hydraulic model: A) Data preprocessing, sub-basins and streams in HEC-HMS. B) Input data in HEC-RAS. C) Development of the computational mesh (50x50) for the 2D Flow Area. D) Inserting the land use map layer in the model.

The land use maps can be also inserted in the model as a map layer (Figure 2D), allowing the model to assign varying Manning’s roughness coefficients (n) per land use category, and also taking this into account for the calculations. There are publicly available tables providing values for the Manning’s roughness coefficients (n) based on the NLCD for the USA (HEC, 2016; Chow, 1959; Barnes, 1987), which were used in this study.

For the rain-on-grid simulation an indicative 24-hour design-storm was used for a return period T=50 years, to make more straightforward the examination the effect of different land use maps, under common precipitation conditions. The NOAA Atlas 14 (NOAA, 2022) provides information on precipitation frequency and depth-duration-frequency curves for US locations, including the CCW study area (Table 1).
### Table 1. PDS-based point precipitation frequency estimates for the CCW, with 90% confidence intervals (in inches).


<table>
<thead>
<tr>
<th>Duration</th>
<th>Average recurrence interval (years)</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>25</th>
<th>50</th>
<th>100</th>
<th>200</th>
<th>500</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-min</td>
<td>(0.347, 0.426)</td>
<td>0.384</td>
<td>0.456</td>
<td>0.545</td>
<td>0.617</td>
<td>0.708</td>
<td>0.782</td>
<td>0.853</td>
<td>0.927</td>
<td>1.03</td>
<td>1.10</td>
</tr>
<tr>
<td>10-min</td>
<td>(0.359, 0.426)</td>
<td>0.597</td>
<td>0.713</td>
<td>0.848</td>
<td>0.952</td>
<td>1.08</td>
<td>1.18</td>
<td>1.26</td>
<td>1.38</td>
<td>1.51</td>
<td>1.60</td>
</tr>
<tr>
<td>15-min</td>
<td>(0.661, 0.011)</td>
<td>0.732</td>
<td>0.871</td>
<td>1.04</td>
<td>1.17</td>
<td>1.34</td>
<td>1.47</td>
<td>1.60</td>
<td>1.72</td>
<td>1.88</td>
<td>2.00</td>
</tr>
<tr>
<td>30-min</td>
<td>(0.875, 0.107)</td>
<td>0.968</td>
<td>1.17</td>
<td>1.42</td>
<td>1.63</td>
<td>1.89</td>
<td>2.10</td>
<td>2.30</td>
<td>2.51</td>
<td>2.79</td>
<td>3.00</td>
</tr>
<tr>
<td>60-min</td>
<td>(1.07, 1.31)</td>
<td>1.18</td>
<td>1.43</td>
<td>1.79</td>
<td>2.07</td>
<td>2.45</td>
<td>2.76</td>
<td>3.08</td>
<td>3.41</td>
<td>3.86</td>
<td>4.22</td>
</tr>
<tr>
<td>2-hr</td>
<td>(1.26, 1.54)</td>
<td>1.39</td>
<td>1.68</td>
<td>2.10</td>
<td>2.46</td>
<td>2.94</td>
<td>3.34</td>
<td>3.77</td>
<td>4.21</td>
<td>4.84</td>
<td>5.36</td>
</tr>
<tr>
<td>3-hr</td>
<td>(1.35, 1.64)</td>
<td>1.47</td>
<td>1.78</td>
<td>2.24</td>
<td>2.62</td>
<td>3.15</td>
<td>3.60</td>
<td>4.07</td>
<td>4.68</td>
<td>5.30</td>
<td>5.99</td>
</tr>
<tr>
<td>6-hr</td>
<td>(1.56, 1.94)</td>
<td>1.74</td>
<td>2.11</td>
<td>2.66</td>
<td>3.11</td>
<td>3.79</td>
<td>4.31</td>
<td>4.90</td>
<td>5.52</td>
<td>6.3</td>
<td>7.18</td>
</tr>
<tr>
<td>12-hr</td>
<td>(1.89, 2.28)</td>
<td>2.07</td>
<td>2.49</td>
<td>3.10</td>
<td>3.66</td>
<td>4.30</td>
<td>4.88</td>
<td>5.49</td>
<td>6.13</td>
<td>7.03</td>
<td>7.77</td>
</tr>
<tr>
<td>24-hr</td>
<td>(2.26, 2.62)</td>
<td>2.44</td>
<td>2.93</td>
<td>3.61</td>
<td>4.14</td>
<td>4.86</td>
<td>5.64</td>
<td>6.30</td>
<td>6.63</td>
<td>7.46</td>
<td>8.11</td>
</tr>
<tr>
<td>2-day</td>
<td>(2.66, 3.05)</td>
<td>2.85</td>
<td>3.41</td>
<td>4.17</td>
<td>4.76</td>
<td>5.55</td>
<td>6.18</td>
<td>6.62</td>
<td>7.47</td>
<td>8.36</td>
<td>9.04</td>
</tr>
<tr>
<td>3-day</td>
<td>(2.86, 3.25)</td>
<td>3.04</td>
<td>3.64</td>
<td>4.42</td>
<td>5.02</td>
<td>5.85</td>
<td>6.50</td>
<td>7.15</td>
<td>7.82</td>
<td>8.72</td>
<td>9.43</td>
</tr>
<tr>
<td>4-day</td>
<td>(3.05, 3.40)</td>
<td>3.24</td>
<td>3.87</td>
<td>4.67</td>
<td>5.30</td>
<td>6.15</td>
<td>6.81</td>
<td>7.49</td>
<td>8.17</td>
<td>8.99</td>
<td>9.81</td>
</tr>
<tr>
<td>7-day</td>
<td>(3.64, 4.09)</td>
<td>3.84</td>
<td>4.56</td>
<td>5.47</td>
<td>6.19</td>
<td>7.16</td>
<td>7.93</td>
<td>8.71</td>
<td>9.50</td>
<td>10.16</td>
<td>11.4</td>
</tr>
<tr>
<td>10-day</td>
<td>(4.26, 4.69)</td>
<td>4.38</td>
<td>5.19</td>
<td>6.19</td>
<td>6.99</td>
<td>7.95</td>
<td>8.69</td>
<td>9.76</td>
<td>10.5</td>
<td>11.8</td>
<td>12.7</td>
</tr>
<tr>
<td>20-day</td>
<td>(5.05, 5.34)</td>
<td>5.98</td>
<td>7.06</td>
<td>8.31</td>
<td>9.36</td>
<td>10.6</td>
<td>11.6</td>
<td>12.8</td>
<td>14.0</td>
<td>15.8</td>
<td>16.9</td>
</tr>
<tr>
<td>30-day</td>
<td>(5.97, 7.79)</td>
<td>7.38</td>
<td>8.67</td>
<td>10.1</td>
<td>11.2</td>
<td>12.6</td>
<td>13.7</td>
<td>14.7</td>
<td>17.1</td>
<td>18.0</td>
<td>18.0</td>
</tr>
<tr>
<td>45-day</td>
<td>(8.85, 9.84)</td>
<td>9.32</td>
<td>10.9</td>
<td>12.6</td>
<td>13.8</td>
<td>15.5</td>
<td>16.7</td>
<td>17.8</td>
<td>20.3</td>
<td>21.4</td>
<td>21.4</td>
</tr>
<tr>
<td>60-day</td>
<td>(10.6, 11.8)</td>
<td>11.2</td>
<td>13.1</td>
<td>15.0</td>
<td>16.4</td>
<td>18.2</td>
<td>19.6</td>
<td>20.8</td>
<td>22.1</td>
<td>23.6</td>
<td>24.7</td>
</tr>
</tbody>
</table>

### 3.3 Ecosystem Services Valuation and Mapping

The valuation of Ecosystem Services (ESV) is a tool increasingly found in several environmental studies. There are many techniques for assigning monetary values to ES, mainly survey-based, which however are not easy to perform every time in different contexts. Thus, value transfer methods can be used for that purpose. Costanza et al. (1997) developed a robust method for ESV, estimating global economic values for ES based on existing literature and original calculations considering 17 ES from 16 biomes. Several papers valuating land cover changes have used this method (e.g. Chuai et al., 2016; Tolessa et al., 2017; Xue and Luo, 2015; Rahman and Szabó, 2021). The most straightforward way to apply this method is to compare the 16 biomes identified by Constanza et al. (1997) and assign their ESV coefficients (as estimated by Constanza et al.) to the respective land cover category (Rahman and Szabó, 2021) (Table 2).
Table 2. Biome equivalents for the five land-use categories and their corresponding ecosystem values (Constanza et al., 1997; Rahman and Szabó, 2021).

<table>
<thead>
<tr>
<th>Land-use types considered for CCW</th>
<th>Equivalent Biome</th>
<th>ESV Coefficient (2022 USD/ha/yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>Water Bodies (lakes/ wetlands/ rivers)</td>
<td>29191.90</td>
</tr>
<tr>
<td>Urban</td>
<td>Build-up areas</td>
<td>0.00</td>
</tr>
<tr>
<td>Barren Land</td>
<td>Bare land (ice or rock)</td>
<td>0.00</td>
</tr>
<tr>
<td>Forest</td>
<td>Vegetation (forest and trees)</td>
<td>3962.67</td>
</tr>
<tr>
<td>Crops</td>
<td>Agriculture land (crop land)</td>
<td>181.65</td>
</tr>
</tbody>
</table>

The ESV can be then estimated spatially by multiplying the areas of each category ($A_c$) with their respective ESV coefficients from Table 2 ($ESV_{coef}$), as expressed in Equation 7, for a given year.

\[
ESV = \sum (A_c \times ESV_{coef})
\]  

In this paper, this was estimated for each historic year (2011, 2016, 2021), and each predicted year (2026, 2031, 2036, 2041, 2046, 2051), to assess the ESV changes. The differences between the estimated ESV for each land use category ($ESV_{dif}$), between two specific years (e.g. a ‘final year’, and a ‘starting year’) can then be estimated with the formula of Equation 8 (Xue and Luo, 2015; Rahman and Szabó, 2021).

\[
ESV_{dif} = \frac{ESV_{final\ year} - ESV_{starting\ year}}{ESV_{starting\ year}} \times \frac{1}{T}
\]  

Where $T$ is the study period.

4. RESULTS AND DISCUSSION

An open-source Python model (script in Spyder, Anaconda) was used to process the historic land use maps, estimate the transition probability matrix, and generate the predicted maps (Alamanos, 2023). The script allows the user to import the results of a GIS software, namely the land use change matrices and it returns the transition probability matrices for each year studied. The model can also read the spatial data (land use maps) in order to apply the CAM model and its validation, as expressed in Equations 1-6. Finally, it provides directly a map with all predicted years, so the user does not need to do this through GIS (Figure 3).
The validation statistics resulted as follows: Accuracy: 99.63%; MAE: 0.0094; RMSE: 0.1613; Kappa: 99.25%. These high values of accuracy and Kappa, as well as the MAE and RMSE values that are close to zero, indicate that the model achieved a satisfactory performance. The predicted land uses show an increase of urban areas over crops and some forest areas, while water and barren land remain almost stable during the predicted years (Figure 4).
Figure 4. The predicted land use areas as number of pixels predicted by the CAM model for CCW.

The hydraulic model was formulated in HEC-RAS and the necessary data preprocessing was performed in HEC-HMS, which is often combined with other tools for hydrological data analysis and hydraulic modelling (Pathak and Kalra, 2015; Alamanos and Papaioannou, 2022). The rain-on-grid simulation run with the Full Momentum method, which is more detailed compared to the Diffusion Wave method. The Full Momentum method takes into account the conservation of both momentum and energy as water flows through the channels, while it considers the influence of all the other input factors such as channel shape, roughness, and slope.
The simulation results are presented in Figure 5 and the zonal statistics for each case are shown in Figure 5. The flood extent and water depth differences are not easily evident in Figure 5 because of the relatively small scale that these changes occurred. However, from the results’ statistics, one can see a slight increase in the flood extent as well as minor increases in the water depths (Figure 6).
Jiang et al. (2008) had studied the historic land use changes in CCW and observed substantial changes for the period 2000 to 2004, where at least 49% of land cover types changed into other types. However, this behaviour was not the case in the subsequent years, and the present study found that the historic and predicted land uses did not significantly vary. Jiang et al. report that more heavy rainfall does not always mean more runoff because the combination of different land cover types always modifies the runoff coefficients. Although in this work the same design storm was considered in all scenarios, the mixed effects in the flood extent were also observed in certain parts of the watershed. This finding is in line with similar behaviours reported in the literature of land use change impacts in runoff. For example, Hounkpè et al. (2019) find that the expansion of agricultural and pasture lands leads to a slight increase in flooding, while Schilling et al. (2013) find that the greatest flood risk reduction is achieved in perennial vegetation areas. In the present study it was also observed that deforestation can cause increased flooding.

The mapping of ESV is shown in Figure 7, where the lower value ‘white’ and ‘orange’ areas very slowly and gradually replace the ‘blue’ areas of higher value.
In order to estimate the evolution of the ES value, the calculated areas per category and year were converted into hectares (Table 3 - upper), since the ESV coefficients are in USD/ha/year, and then Equations 7 and 8 were applied (Table 3 - lower).
Table 3. The evolution of the different land use types (ha), their ESV (million USD, 2022 values), and the estimated yearly changes of ESV.

<table>
<thead>
<tr>
<th>Land Use</th>
<th>Areas (Ha)</th>
<th>2011</th>
<th>2016</th>
<th>2021</th>
<th>2026</th>
<th>2031</th>
<th>2036</th>
<th>2041</th>
<th>2046</th>
<th>2051</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>5866.92</td>
<td>5859.99</td>
<td>5873.04</td>
<td>5862.24</td>
<td>5871.96</td>
<td>5874.03</td>
<td>5877.18</td>
<td>5880.51</td>
<td>5874.66</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>7988.76</td>
<td>8133.03</td>
<td>8265.96</td>
<td>8428.59</td>
<td>8583.48</td>
<td>8739.36</td>
<td>8890.38</td>
<td>9054.63</td>
<td>9054.63</td>
<td></td>
</tr>
<tr>
<td>Barren Land</td>
<td>39.42</td>
<td>40.23</td>
<td>114.84</td>
<td>171.18</td>
<td>165.42</td>
<td>159.48</td>
<td>154.62</td>
<td>152.73</td>
<td>158.40</td>
<td></td>
</tr>
<tr>
<td>Forest</td>
<td>8006.94</td>
<td>7952.04</td>
<td>7920.27</td>
<td>7886.25</td>
<td>7820.10</td>
<td>7753.86</td>
<td>7686.00</td>
<td>7615.80</td>
<td>7607.70</td>
<td></td>
</tr>
<tr>
<td>Crops</td>
<td>47884.68</td>
<td>47801.43</td>
<td>47612.61</td>
<td>47438.46</td>
<td>47259.99</td>
<td>47178.54</td>
<td>47083.05</td>
<td>47091.33</td>
<td>47091.33</td>
<td></td>
</tr>
<tr>
<td>Total Area</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td>69786.72</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ESV (million USD)</th>
<th>2011</th>
<th>2016</th>
<th>2021</th>
<th>2026</th>
<th>2031</th>
<th>2036</th>
<th>2041</th>
<th>2046</th>
<th>2051</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>171.27</td>
<td>171.06</td>
<td>171.45</td>
<td>171.13</td>
<td>171.41</td>
<td>171.57</td>
<td>171.66</td>
<td>171.49</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Barren Land</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Forest</td>
<td>31.73</td>
<td>31.51</td>
<td>31.39</td>
<td>31.25</td>
<td>30.99</td>
<td>30.73</td>
<td>30.46</td>
<td>30.18</td>
<td>30.15</td>
</tr>
<tr>
<td>Crops</td>
<td>8.70</td>
<td>8.68</td>
<td>8.65</td>
<td>8.62</td>
<td>8.60</td>
<td>8.58</td>
<td>8.57</td>
<td>8.55</td>
<td>8.55</td>
</tr>
<tr>
<td>Total</td>
<td>211.69</td>
<td>211.26</td>
<td>211.48</td>
<td>211.00</td>
<td>211.00</td>
<td>210.78</td>
<td>210.59</td>
<td>210.39</td>
<td>210.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ESV Changes (million USD)</th>
<th>2011-2016</th>
<th>2016-2021</th>
<th>2021-2026</th>
<th>2026-2031</th>
<th>2031-2036</th>
<th>2036-2041</th>
<th>2041-2046</th>
<th>2046-2051</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>-0.202</td>
<td>0.381</td>
<td>-0.315</td>
<td>0.284</td>
<td>0.060</td>
<td>0.092</td>
<td>0.097</td>
<td>-0.171</td>
</tr>
<tr>
<td>Urban</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Barren Land</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Forest</td>
<td>-0.218</td>
<td>-0.126</td>
<td>-0.135</td>
<td>-0.262</td>
<td>-0.262</td>
<td>-0.269</td>
<td>-0.278</td>
<td>-0.032</td>
</tr>
<tr>
<td>Crops</td>
<td>-0.015</td>
<td>-0.034</td>
<td>-0.032</td>
<td>-0.017</td>
<td>-0.016</td>
<td>-0.015</td>
<td>-0.017</td>
<td>0.002</td>
</tr>
<tr>
<td>Total</td>
<td>-0.435</td>
<td>0.221</td>
<td>-0.482</td>
<td>0.005</td>
<td>-0.218</td>
<td>-0.192</td>
<td>-0.198</td>
<td>-0.201</td>
</tr>
</tbody>
</table>

The total estimated ESV of CCW in 2011 was 211.69 mil.USD, and the projected value in 2051 is 210.19 mil.USD, which is 1.5 mil.USD less. The loss of forests and wetlands might be the reason. The water bodies have been accounted for approximately of the 81.27% of the total CCW’s ESV, on average, every year. Thus, water bodies have been the greatest contributor to ESV, given the multiple and diverse services they can provide.

CONCLUSIONS

This exercise showcased how a CAM model using GIS and Python can be applied to future land use predictions, as well as their validation. As explained in the introductory section, this analysis can find multiple applications in a variety of studies on human-environmental systems. Such an example is the examination of the impacts of future land use changes to flood extent, and the evolution of their monetary value according to their ES, which are overlooked applications. This paper examined these impacts, working with the minimum necessary inputs for the CAM, the hydraulic, and the ESV models, in order to
showcase a simple way to assess them in data-scarce areas. Although this approach is often necessary in data-scarce areas, it comes with certain limitations.

It is important to keep in mind that the behavior of a CAM model always depends on the specific rules and parameters set for the model. For example, one could expect a concentrated urbanization around the existing urban centers rather than a more scattered picture. The CAM model works on a pixel-by-pixel basis and does not inherently consider the spatial distribution of land use changes unless explicitly programmed to do so. It uses factors such as the current land use category of a pixel, its neighboring pixels and the transition probabilities (according to the estimated matrices). For more refined changes, factors like proximity to existing urban centers, can be included in the code. In this case, we kept the example simple, given the purpose to show how a limited-data approach would work, as proximity to existing urban centers would require additional data (e.g. shape files or point data of urban centers), and extra analysis (e.g. raster-distance tools, using the proximity as an additional layer, and incorporate it in the CAM). Machine Learning can also assist in improving the prediction accuracy. In any case, the validation over historic land use observations helps ensure a degree of accuracy of the model and make improvements by further adjusting it (mainly through the transition probability matrices). In this example, the validation showed a good performance of the model. Moreover, the results are in line with the general perception that more urbanized and deforested areas can increase the potential extent and severity of floods. Limited input data is often a practical and modelling consideration, so in this example, a design storm is used for the hydraulic model. Although this is a common approach, it is a synthetic event that might not capture the characteristics of real storms, and it also simplifies the modelling process. However, it allows us to examine the sensitivity of flooding solely based on the land use changes, under the same storm conditions.

The relationship between land use changes, flooding, and ESV is complex and context-specific, depending on factors such as topography, land uses, hydrological conditions, storm characteristics, and benefits people derive from ecosystems. Future analyses can consider more data to provide more detailed assessments considering more refined land use predictions with more classes, improved adjustments of the transition probability matrices, or considering the proximity to neighboring cells, as well as more thorough hydrological models and real storms. Also, future research could also further explore the distribution of ESV per land use type, breaking them down to more specific ES, e.g. provisioning (production, supply, raw materials), regulating (water regulation, waste treatment, erosion control, climate regulation, biological control, gas regulation, disturbance regulation), supporting (nutrient cycling, pollination, soil formation, habitat/refugia), and cultural (recreation, cultural). The approach followed in this example, indicates that the loss of the major ESV contributor (i.e. water bodies), can be a factor highly related to the increasing flood risks, as water bodies and in particular wetlands, exhibit important regulatory services, of which climate regulation and water retention are directly connected with flooding.

The importance of having reliable land use change models is widely recognized, and the provision of user-friendly tools is crucial. However, there are very few available free and publicly accessible tools for such processes. This work contributes to the provision of tools for land use prediction (data and models used are publicly available), facilitating the integration of land use changes in a variety of studies within the broader context of human-environmental systems management. The examples of flooding and ESV prove the multi-layered nature of such studies, which can be highly insightful for policymakers and land use planners for integrated assessments and decisions on the optimum land use development patterns, ensuring resilience and sustainable provision of ES.
Data and Code Availability:

All data used for the analyses can be retrieved from publicly available sources, which are cited in the paper. The software used in this paper are also available (QGIS, HEC-HMS, HEC-RAS).

The GIS guide to perform the necessary analyses and the Python script for the CAM model and its validation are accessible at: https://github.com/Alamanos11/Land_uses_prediction. The Python model developed provides also more tests as validation options, such as Confusion Matrix, and Confusion Matrix Classification Report, which were not presented in the paper to keep it concise.

REFERENCES


Han, H., Yang, C., & Song, J. (2015). Scenario Simulation and the Prediction of Land Use and Land Cover Change in Beijing, China. Sustainability, 7(4), Article 4. [https://doi.org/10.3390/su7044260](https://doi.org/10.3390/su7044260)


