OPEN ACCESS

Int. J. Creative Res. Publl. *Peer-Reviewed Journal* *Volume 1 Issue 4 [Jul-Aug] 2026*

International Journal of Creative Research and Publications
Double Blind Peer Reviewed Journal

Research Article

Uncertainty-Aware Cellular Automata with Memory for Annual Land-Cover Transition Forecasting: Hindcast Validation Using ESA CCI Maps

Tathagata Satapathy [1*], Deepak Kumar Sahoo [2]

[1] Assistant Professor, Department of Basic Science and Humanities, NIIS Institute of Engineering and Technology, Bhubaneswar, Odisha, India

[2] Assistant Professor, Department of Electrical Engineering, NIIS Institute of Engineering and Technology, Bhubaneswar, Odisha, India

Corresponding Author: * Tathagata Satapathy

DOI: https://doi.org/10.5281/zenodo.21806481

Abstract

Land-cover change is one of the most consequential and most poorly constrained inputs to Earth system, biodiversity, and carbon-accounting models. Most operational cellular automata (CA) forecasting frameworks couple a Markov transition matrix with a neighbourhood-based spatial allocation rule, an approach that is computationally efficient but structurally memoryless: transition probabilities are re-estimated from only the two most recent maps, and no explicit account is taken of the confidence with which any given cell's future state can be known. This paper develops and specifies, in full methodological detail, an Uncertainty-Aware Cellular Automata with Memory (UA-CAM) framework for annual, cell-level land-cover transition forecasting. The framework augments the classical CA-Markov formulation with (i) a temporal memory kernel that weights multiple historical transitions rather than a single one-step-back transition, (ii) a persistence/inertia term calibrated from the length of each cell's most recent land-cover spell, and (iii) an explicit uncertainty model that propagates class-confusion probabilities, neighbourhood heterogeneity, and transition-rarity into a per-cell, per-class predictive distribution rather than a single deterministic label. The framework is designed for hindcast validation against the European Space Agency Climate Change Initiative (ESA CCI) global annual land-cover maps (1992–2020, 300 m resolution), using a train-on-the-past/predict-the-present design in which the model is calibrated on maps up to year t and validated against the independently observed map at year t+1, ..., t+k. We present the complete mathematical formulation of the memory-weighted transition kernel, the neighbourhood suitability function, the stochastic cell-allocation procedure, and a Dempster–Shafer-based uncertainty-combination rule; we give parallel, ready-to-adapt Python and R implementations of the full workflow, from ESA CCI ingestion through hindcast scoring; and we specify a validation protocol built on the three-map Figure of Merit (FoM), quantity/allocation disagreement, fuzzy Kappa, and a novel Uncertainty Calibration Score (UCS) that scores whether the model's stated confidence matches its empirical hit rate. Because no hindcast run has yet been executed under this exact specification, all numerical results reported in the Expected Results section are explicitly labelled as anticipated, literature-informed estimates rather than empirical findings; they are intended to give a reader a calibrated sense of the plausible range of outcomes and to serve as pre-registered benchmarks against which a future empirical run can be compared. We close with a discussion of the framework's limitations — particularly the coarse thematic resolution of ESA CCI relative to local drivers of change, the computational cost of ensemble uncertainty propagation at global scale, and the difficulty of validating rare-class transitions — and outline a future-work agenda spanning driver-informed suitability layers, learned memory kernels, and multi-resolution hierarchical hindcasting.

Manuscript Information

- **ISSN No:** 3139-6313
- **Received:** 13-06-2026
- **Accepted:** 26-07-2026
- **Published:** 05-08-2026
- **IJCRP:**1(4); 2026: 62-77
- **©2026, All Rights Reserved**
- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Article

Satapathy T, Sahoo D K. Uncertainty-Aware Cellular Automata with Memory for Annual Land-Cover Transition Forecasting: Hindcast Validation Using ESA CCI Maps. Int. J. Creative Res. Publ. 2026;1(4):62-77.

Access this Article Online

www.jcrp.in

KEYWORDS: cellular automata; land-cover change; Markov chain; memory kernel; uncertainty quantification; Dempster–Shafer theory; ESA CCI; hindcast validation; Figure of Merit; fuzzy Kappa; spatiotemporal forecasting; remote sensing

1. INTRODUCTION

Land-cover change (LCC) is a first-order control on terrestrial carbon fluxes, biodiversity loss, water balance, and regional climate feedbacks, and forecasting it credibly at annual, cell-level resolution remains one of the more stubborn problems in land-system science. Two properties of LCC make it difficult to forecast well. First, it is spatially contagious: whether a given 300 m cell converts from cropland to built-up land in a given year depends strongly on what its neighbours are doing, which is precisely the property that cellular automata (CA) were designed to represent. Second, it is not simply Markovian in time: a cell's probability of converting in year t+1 depends not only on its state in year t, but on how long it has occupied that state, how many times it has previously oscillated between states, and on multi-year regional trends that a single one-step transition matrix cannot see. Most operationalised CA forecasting tools — CA-Markov as implemented in IDRISI/TerrSet, and its many regional derivatives — address the first property but not the second: they recompute a transition probability matrix from exactly two time points and then allocate change spatially using a contiguity filter, discarding all deeper temporal history at each re-calibration step.

A second, largely separate gap is uncertainty. Land-cover maps are themselves noisy classifications with known, spatially and thematically heterogeneous confusion characteristics; a CA model trained on them and used to forecast further maps compounds classification error with model-structural error, transition-rarity error (rare conversions are estimated from very few samples and are correspondingly noisy), and spatial-allocation error. Almost all applied CA-Markov studies report a single deterministic future map along with an aggregate accuracy statistic (typically Kappa or Per cent Correctly Classified) computed after the fact, but they do not produce — and cannot, by construction, produce — a per-cell statement of how confident the model should be in its own forecast for that cell. This is a serious limitation for any downstream use of the forecast (e.g., feeding a forecast land-cover map into a carbon or hydrology model), because the user has no principled way to know where the forecast can be trusted and where it cannot.

This paper addresses both gaps together. We specify Uncertainty-Aware Cellular Automata with Memory (UA-CAM), a CA framework in which (i) the transition kernel driving each cell's evolution is a weighted function of several past transitions rather than a single one, with the weighting governed by an explicit memory/decay parameter and a cell-level persistence (spell-length) covariate, and (ii) every forecast is accompanied by a per-cell, per-class belief distribution derived from a Dempster–Shafer combination of classification confusion evidence, neighbourhood heterogeneity evidence, and transition-rarity evidence. We specify the framework for use with the ESA CCI annual global land-cover time series (1992–2020), which is, to our knowledge, the longest globally consistent annual land-cover record publicly available and therefore the only dataset that permits genuine multi-year hindcast validation (as opposed to two-point calibration and single-point testing, which is what most published CA-Markov validations actually do).

The remainder of the paper is organised as follows. Section 2 reviews CA-Markov and its temporal and uncertainty-aware extensions. Section 3 lays out the overall study framework and design philosophy. Section 4 describes the ESA CCI materials used. Section 5 gives the complete methodology: the memory-weighted transition kernel, the neighbourhood suitability function, the stochastic allocation algorithm, and the Dempster–Shafer uncertainty model, together with full Python and R workflows. Section 6 specifies the hindcast validation protocol, and Section 7 the evaluation metrics. Section 8 presents Expected Results, explicitly labelled as anticipated rather than empirically observed. Sections 9–12 discuss limitations, propose future work, and conclude.

2. LITERATURE REVIEW

Cellular automata for land-use/cover simulation. CA-based land-use simulation traces its conceptual lineage to von Neumann and Burks' (1966) work on self-reproducing automata and to Tobler's (1979) proposal that geographic cellular models could represent spatial processes through local transition rules — an idea encapsulated in Tobler's First Law of Geography, that near things are more related than distant things. Batty (1997) and Batty and Xie (1994) formalised urban CA as a tool for simulating city growth from simple local rules, and White and Engelen (2000) extended constrained CA to represent land-use competition among multiple classes under exogenous demand. Clarke, Hoppen, and Gaydos (1997) introduced SLEUTH, one of the most widely applied urban-growth CA models, calibrated via brute-force parameter search against historical growth. Couclelis (1997) provided an early critical assessment of CA's promise and limits as a representation of urban dynamics.

Coupling CA with Markov chains. The dominant applied paradigm — CA-Markov — pairs a Markov transition-probability matrix, estimated from two historical land-cover maps, with CA-based spatial allocation, typically implemented via a contiguity/suitability filter (Eastman, 2003; implemented operationally in IDRISI/TerrSet). This is the design used in the large majority of applied studies, including CA-Markov chain modeling of urban expansion in the Ethiopian Rift (Hyandye & Martz, 2017), forecasting of LULC in Lebanon using a modified CA-Markov chain that decomposes built-up density classes (Al-Shaar et al., 2021), CA-Markov prediction of land-use in Ternate City (Latue & Rakuasa, 2023), and integration of CA-Markov with artificial neural networks to incorporate driving-force covariates such as slope, road distance, and population density (reviewed in the PMC study on ANN-CA-MC integration). Reviews of this literature (e.g., the geospatial CA-Markov review indexed by ResearchGate) consistently note that the Markov component's key limitation is that it is a memoryless, aspatial accounting device: it predicts *how much* change of each type will occur but says nothing about *where*, which is why it must be paired with CA or another spatial-allocation mechanism; conversely, the CA component's key limitation, absent modification, is that it inherits the Markov chain's assumption that the future depends only on the present state and not on trajectory. Markov chain is a technique widely used to predict changes in vegetation based on the rate of previous changes, but its main disadvantage is its lack of spatial dimension, which is why it is typically combined with cellular automata that change state based on the previous condition of a cell and its neighbours under a specific rule.

Temporal-dependence-aware extensions. A recent and directly relevant line of work has begun to address the memorylessness of CA-Markov by embedding recurrent neural architectures inside the CA loop. Previous CA-based methods that treated land-use-change dynamics as Markov processes were intrinsically unable to capture long-term temporal dependency, motivating the application of long short-term memory (LSTM) networks within a CA framework to solve the temporal-dependence problem. Building on this, hybrid architectures combining convolutional networks (for spatial heterogeneity) with LSTM (for temporal dependence) inside a CA loop — including CNN-LSTM-CA, DST-CA, and KCLP-CA variants that additionally couple with the Patch-generating Land Use Simulation (PLUS) model — have been proposed specifically because prior CA studies overlooked the significant temporal dependence and spatial heterogeneity associated with land-use change. A parallel deep-learning approach, ConvLSTM-CA (DST-CA), integrates a 3D convolutional network to capture local short-term spatiotemporal features together with LSTM to extract long-term chronological dependencies, aiming to more comprehensively represent the nonlinear spatiotemporal characteristics of land-use/cover change, and was demonstrated on Guangdong Province. While these deep architectures demonstrably improve on plain CA-Markov, they require large volumes of training imagery, substantial computational infrastructure, and offer limited interpretability of the memory mechanism itself — motivating, for many applied and policy contexts, a lighter-weight, explicitly parameterized memory-kernel approach of the kind developed here, in which the "memory" is a transparent weighted combination of past transition matrices and spell lengths rather than a learned latent state.

Uncertainty in CA-based land-change models. Uncertainty has been studied primarily as a post-hoc sensitivity question rather than as an integral part of the forecast. Uncertainties arising from the spatial resolution and proportional error of input maps, and from the number of CA iterations, have been shown to affect simulation accuracy as quantified via Kappa statistics and confusion matrices, with input-data uncertainty exerting a larger influence on simulation results than CA-parameter uncertainty. This literature establishes that classification error in the input maps is often the dominant uncertainty source — a finding directly relevant to any CA model built on ESA CCI, which itself carries documented classification uncertainty via its quality-flag layers. To our knowledge, however, no widely cited applied CA-Markov study propagates this input uncertainty forward into an explicit, per-cell predictive confidence for the forecast map itself; uncertainty is typically summarised as a single global accuracy number computed after the simulation, not carried through the simulation as a state variable — the gap this paper's uncertainty model is designed to close.

Global annual land-cover data. The ESA CCI Land Cover project is the primary data source enabling genuine multi-year hindcast validation. The ESA CCI medium-resolution land cover dataset, operationalised within the EU Copernicus Climate Change Service, provides the longest consistent land-cover climate data record, with annual maps from 1992 to 2020 at 300 m spatial resolution, describing the land surface in 22 land-cover classes according to the UN Land Cover Classification System and 13 land-cover change types consistent with IPCC land categories. The dataset is produced using a combination of multi-year, multi-sensor strategies incorporating AVHRR, SPOT-Vegetation, ENVISAT-MERIS, and PROBA-Vegetation imagery, and is distributed in GeoTIFF and NetCDF formats together with four quality flags documenting classification reliability and change detection. A new map is produced annually roughly nine months after the start of each year and validated within a further three months of release, giving the series both the temporal depth and the documented uncertainty metadata needed for the uncertainty-aware hindcast design proposed here.

Validation of land-change models. The methodological literature on validating simulated land-change maps, led by Pontius and colleagues, establishes that naive two-map Kappa or per cent-correct comparisons are misleading because a two-map comparison cannot distinguish correctly simulated change from correctly simulated persistence, and when persistence dominates a landscape, such comparisons yield large per cent-correct and Kappa values regardless of whether the model simulates change correctly. The recommended remedy is a three-map comparison — reference at the start of the validation interval, simulation at the end of the interval, and reference at the end of the interval — from which the Figure of Merit (FoM) is computed, a metric popularised by Pontius et al. (2008, 2011, 2018). The FoM ranges from zero, indicating no intersection between simulated and reference change, to one, indicating perfect intersection, although it has limited diagnostic power on its own because it combines quantity and allocation information into a single number, which is why FoM is normally decomposed into hit/miss/false-alarm components (Chen & Pontius, 2010) and supplemented with quantity and allocation disagreement (Pontius & Millones, 2011). Complementary approaches include fuzzy Kappa (Hagen, 2003), which relaxes the strict cell-by-cell matching requirement to account for near-miss spatial allocation, and the Null Model comparison (Pontius & Malanson, 2005), which compares the agreement between the base map and the reference map against the agreement between the simulated map and the reference map, such that a modeling exercise performs poorly if persistence alone predicts the reference map better than the model does. Recent large-scale applications of PLUS-based CA report validation statistics in this tradition — for example, a global 1 km land-use/cover dataset for 2020–2100 built on an improved CA model (PLUS) reported Kappa = 0.94, overall accuracy = 0.97, and FoM = 0.10 at the water-basin scale, illustrating the now-common empirical pattern that FoM values are substantially lower than Kappa or overall-accuracy values because FoM specifically penalizes missed and misallocated change rather than being inflated by correctly simulated persistence.

The present paper's contribution relative to this literature is threefold: (1) it replaces the two-map Markov transition estimate with an explicit, parameterized multi-year memory kernel, while remaining lightweight and interpretable relative to deep-learning temporal-CA hybrids; (2) it embeds a Dempster–Shafer uncertainty-combination model directly into the cell-allocation step, producing per-cell predictive distributions rather than a single deterministic label with a post-hoc global accuracy score; and (3) it specifies a genuine multi-year hindcast protocol against ESA CCI, rather than the single train-test split typical of most applied CA-Markov studies, and pre-registers the evaluation metrics (three-map FoM, quantity/allocation disagreement, fuzzy Kappa, and a novel Uncertainty Calibration Score) before any run is executed.

3. STUDY FRAMEWORK

The study is organised around four design principles.

(a) Memory over Markov. Instead of estimating a single transition matrix **P**_{t→t+1} from the two most recent maps, UA-CAM estimates a memory-weighted transition matrix **P̃**_{t+1} as an exponentially decayed combination of the L most recent one-step transition matrices, further modulated by each cell's current spell length (time since its last class change). This allows the model to represent both short-run momentum (a class that has been converting rapidly in the last few years is expected to keep doing so) and long-run inertia (a cell that has been stable for decades is harder to convert than one that changed class last year), which a two-point Markov estimate cannot represent.

(b) Cell-level uncertainty as a first-class output. Every forecast cell carries not a single predicted class but a belief mass distribution over the full class set, derived by Dempster–Shafer combination of three independent evidence sources: (i) the ESA CCI classification-confusion evidence for the cell's current class, (ii) neighborhood heterogeneity evidence (mixed neighborhoods are harder to forecast than homogeneous ones), and (iii) transition-rarity evidence (rare transition types are estimated from few historical samples and are therefore less reliable). The combined belief function is used both to select the most probable class stochastically (rather than deterministically assigning the arg-max class to every cell) and to report a calibrated confidence interval for every forecast.

(c) Genuine hindcast validation, not two-point calibration. The model is calibrated using only maps available up to a cutoff year *t*₀, and its forecasts for *t*₀+1 through *t*₀+*k* are compared against the independently observed ESA CCI maps for those same years, which were withheld from calibration. This design directly tests the value added by the memory kernel, since a plain CA-Markov baseline calibrated on the same withheld-data design can be run side by side.

(d) Reproducibility across two open-source ecosystems. Because applied land-system science is split fairly evenly between Python (rasterio/xarray/numpy) and R (terra/raster) tooling, the full workflow is specified and implemented in both languages, using the same equations and the same random-seed convention, so that either community can adopt, audit, and extend the framework without needing to reimplement the mathematics.

Figure 1 (described rather than rendered here) would show the framework as a pipeline: ESA CCI ingestion → per-cell spell-length and transition-history extraction → memory-weighted transition kernel estimation → neighbourhood suitability computation → Dempster–Shafer uncertainty combination → stochastic cell allocation with a global-quantity constraint → iterate one year at a time up to the forecast horizon → hindcast comparison against withheld ESA CCI maps.

4. MATERIALS

Primary dataset. The ESA Climate Change Initiative (CCI) Land Cover time series is used as the sole primary dataset, for both calibration and hindcast validation. The dataset provides consistent global annual land-cover maps at 300 m spatial resolution from 1992 to 2020, with land-cover classes defined using the UN Food and Agriculture Organisation's Land Cover Classification System (LCCS), and covers the period 1992–2022 at approximately 300 m (0.002778°) resolution, generated from the entire multi-mission Earth observation archive. Four quality flags are delivered with the maps, including the number of land-cover changes detected over the full period, documenting classification reliability; these quality flags are the direct empirical basis for the classification-confusion evidence layer in the uncertainty model (Section 5.3). A user tool is provided by the CCI Land Cover team for sub-setting, re-projecting, and re-sampling the products to formats suitable for downstream climate and land models, and the maps are also distributed as Cloud-Optimized GeoTIFFs and via the Copernicus Climate Data Store, both of which are used here for programmatic access in the Python and R workflows.

Legend and class aggregation. The CCI maps use a "level 1" legend of ten broad classes (values 10, 20, 30, etc.) and a more detailed "level 2" legend of regionally refined classes (values such as 11, 12, 61, 62), with the level-1 legend chosen to be globally consistent. For the CA-with-memory framework, we recommend calibrating and forecasting at the level-1 (broad, ~10–13 class) resolution, since transition-matrix estimation for rare level-2 classes over a 29-year record already strains sample sizes; level-2 detail can be reintroduced as a post-hoc disaggregation layer using the fractional Plant Functional Type products described below.

Ancillary products. Two CCI-derived ancillary products are used to construct static and quasi-static suitability covariates: (i) the CCI Plant Functional Type (PFT) fractional composition product, which delivers 19 global PFT maps at 300 m resolution for each year from 1992 to 2020, including broadleaved/needleleaved evergreen/deciduous trees and shrubs, C3/C4 natural and managed grasses, built-up areas, bare areas, and snow/ice, used here as a sub-pixel heterogeneity covariate; and (ii) elevation, slope, distance-to-road, and distance-to-existing-urban-center layers of the kind used in ANN-augmented CA-Markov studies, used as static suitability covariates in the neighborhood function (Section 5).

No additional proprietary or restricted datasets are required; the entire pipeline can be run from public ESA CCI and standard topographic layers, which is a deliberate design choice to maximise reproducibility.

Study extent. The methodology is resolution- and extent-agnostic and is written to be applied to any bounding box or national extent within the ESA CCI global grid; for hindcast demonstration purposes we recommend a moderate-size study window (e.g., a single level-2 administrative region or a 500 km × 500 km tile) large enough to contain a representative mix of stable and dynamic land-cover classes, but small enough that the full 29-year annual stack and its memory-kernel history fit comfortably in memory on a single workstation.

5. COMPLETE METHODOLOGY

5.1 Notation

Let the study area be discretised into a raster of cells indexed by $i \in \{1, ..., N\}$. Let $\mathcal{S} = \{s_1, ..., s\_K\}$ be the set of K land-cover classes. Let $X\_i(t) \in \mathcal{S}$ denote the observed class of cell i in year t, drawn from the ESA CCI archive for $t = 1992, ..., 2020$. Let t_0 denote the calibration cutoff year, and let the forecast horizon be $h = 1, ..., H$, so that the model produces forecasts $\hat{X}\_i(t_0+h)$ to be compared against the withheld observations $X\_i(t_0+h)$.

5.2 Cellular Automata with Memory: algorithmic overview

The core forecasting loop advances one year at a time. At each step from year t to t+1, four operations are performed for every cell i:

1. **Memory-weighted transition estimation.** Compute a per-class-pair transition probability that blends the L most recent observed one-step transitions, exponentially discounting older transitions, and further conditioned on the cell's current spell length (Eq. 1–3).
2. **Neighbourhood suitability.** Compute, for each candidate class s_k, a suitability score reflecting the composition of the cell's neighbourhood, static covariates (slope, distance-to-road, distance-to-existing class), and the PFT fractional layer (Eq. 4).
3. **Combined transition potential and Dempster–Shafer uncertainty combination.** Combine the memory-weighted transition probability and the neighbourhood suitability into a transition-potential score, and combine three independent uncertainty evidence sources (classification confusion, neighbourhood heterogeneity, transition rarity) via Dempster–Shafer combination to obtain a belief mass function over $\mathcal{S}$ for the cell (Eq. 5–9).
4. **Stochastic, demand-constrained allocation.** Sample the forecast class for each cell from its belief-mass distribution, subject to a global constraint that the total number of cells allocated to each class in year t+1 matches the quantity implied by the memory-weighted transition matrix aggregated over the whole study area (a "iterative proportional fitting"-style rebalancing step, consistent with the demand-allocation logic used in CA-Markov and CLUE-S-style frameworks).

This loop is repeated for $h = 1, ..., H$, using the model's own *previous* forecast (not ground truth) as the basis for subsequent years' spell-length and transition-history updates, which is the correct procedure for genuine multi-year forecasting (as opposed to one-step-ahead forecasting re-initialised from ground truth every year).

5.3 Mathematical formulation

(1) One-step empirical transition matrix. For each historical year pair (t−1, t) within the calibration window, define the empirical one-step transition probability from class a to class b as

P_t (a → b) = N_t(a → b) / N_{t−1}(a)

where $N\_t(a \to b)$ is the count of cells that were class a at t−1 and class b at t, and $N\_\{t-1\}$ (a) is the total count of cells in class a at t−1.

(2) Memory kernel. The memory-weighted transition matrix used to drive the forecast from t to t+1 is

$$\tilde{P}(a \to b) = \Sigma\_\{\tau=1\}\ \hat{}\{L\}\ w\_\tau \cdot P\_\{t-\tau+1\}\ (a \to b)\ ,\ \Sigma\_\tau\ w\_\tau = 1$$

with exponential decay weights

$$w\_\tau = (1 - \lambda) \cdot \lambda\hat{}\{\tau-1\} / \Sigma\_\{j=1\}\hat{}\{L\}\ (1 - \lambda) \cdot \lambda\hat{}\{j-1\}$$

where $\lambda \in (0, 1)$ is the memory-decay parameter ($\lambda \to 0$ recovers the classical two-point Markov estimate; $\lambda \to 1$ approaches an unweighted average over the full available history) and L is the memory window length, both calibrated by cross-validated grid search (Section 5.6).

(3) Spell-length modulation (inertia term). Let *spell_i(t)* be the number of consecutive years cell i has held its current class as of year t. The cell-specific, inertia-adjusted persistence probability is

$$\tilde{P}\_i(a \to a) = \tilde{P}(a \to a) + (1 - \tilde{P}(a \to a)) \cdot [1 - \exp(-spell\_i(t) / \theta\_a)]$$

where θ_a is a class-specific inertia time-constant estimated from the empirical distribution of historical spell lengths for class a (long-lived classes such as permanent water or dense forest are expected to have large θ_a; volatile classes such as shifting cropland have small θ_a). The off-diagonal transition probabilities $\tilde{P}\_i(a \to b)$, $b \neq a$, are then rescaled proportionally so that $\Sigma\_b\ \tilde{P}\_i(a \to b) = 1$.

(4) Neighbourhood suitability. For cell i and candidate class s_k, define a neighbourhood suitability score

$$Suit\_i(k) = \alpha_1 \cdot (n\_i(k) / n\_i\hat{}\{total\}) + \alpha_2 \cdot f(covariates\_i, k) + \alpha_3 \cdot PFT\_i(k)$$

where *n_i(k)* is the count of cells of class k within a defined neighborhood window (e.g., 5×5 or 7×7 Moore neighborhood) around cell i, *n_i^{total}* is the total neighborhood cell count, $f(\cdot)$ is a logistic function of static covariates (slope, distance-to-road, distance-to-urban-center) fitted by multinomial logistic regression against historical conversions (in the spirit of the ANN/logistic-regression-augmented CA-Markov literature), *PFT_i(k)* is the relevant plant-functional-type fractional cover for class k at cell i, and α_1, α_2, α_3 are weights that sum to one and are calibrated by maximizing historical one-step forecast likelihood.

(5) Combined transition potential. The overall potential for cell i to occupy class k at t+1, given its current class $a = X\_i(t)$, is

TP_i(k) = P̃_i(a → k) · Suit_i(k) / Σ_{k'} P̃_i(a → k') · Suit_i(k')

i.e., a normalised product of the memory-weighted transition probability and the neighbourhood suitability, following the standard CA-Markov "transition potential" construction but substituting the memory-weighted, inertia-adjusted probability for the plain two-point Markov probability.

(6)–(9) Dempster–Shafer uncertainty combination. Three independent bodies of evidence are defined over the frame of discernment $\mathcal{S}$:

- m_1 (classification-confusion evidence): derived from the ESA CCI quality flags and any available reference/validation confusion matrix for the region, assigning mass to the observed class a in proportion to its documented producer's/user's accuracy and distributing the residual mass across classes historically confused with a.
- m_2 (neighbourhood-heterogeneity evidence): assigns higher mass to a small set of dominant neighbourhood classes when the neighbourhood is homogeneous, and spreads mass more broadly (higher inherent uncertainty) when Shannon entropy of the neighbourhood class composition is high.
- m_3 (transition-rarity evidence): assigns mass to candidate classes in inverse proportion to the estimation variance of P̃_i(a → k), computed from the effective historical sample size $N\_\{t−τ\}(a)$ underlying the memory-weighted average; rare transitions (small historical sample) receive wide, low-confidence mass; common transitions receive narrow, high-confidence mass.

These are combined pairwise using Dempster's rule of combination,

$(m_1 \oplus m_2)(C) = [\ \Sigma\_\{A \cap B = C\}\ m_1(A) \cdot m_2(B)\ ] / [\ 1 - \Sigma\_\{A \cap B = \emptyset\}\ m□1(A) \cdot m_2(B)\ ],\ \ C \neq \emptyset$

applied first to (m_1, m_2) and then to the result combined with m_3, yielding a final combined belief mass function $m(\cdot)$ over $\mathcal{S}$ for cell i. From $m(\cdot)$ we derive the belief and plausibility bounds, Bel(k) = Σ_{A ⊆ {k}} m(A) and Pl(k) = Σ_{A ∩ {k} ≠ ∅} m(A), which together define the per-cell, per-class confidence interval reported alongside the forecast. The final class-selection probability used for stochastic sampling is the pignistic transformation of $m(\cdot)$,

BetP(k) = Σ_{A ⊆ $\mathcal{S}$, k ∈ A} m(A) / |A| .

The transition potential TP_i(k) from Eq. (5) is itself folded in as a fourth evidence source (m_4) before combination, so that the spatial-allocation signal and the three uncertainty sources jointly determine the sampling distribution — this is the point at which "uncertainty-aware" and "cellular automata with memory" are formally unified into a single per-cell distribution.

(10) Demand-constrained stochastic allocation. For each year, the aggregate number of cells expected to occupy each class k, Demand(k), is computed by applying the whole-area memory-weighted transition matrix to the current global class totals. Cells are then sampled class-by-class in descending order of BetP(k) confidence margin, with a running check against Demand(k), and any residual imbalance is corrected by a final iterative proportional fitting (IPF) adjustment applied to the lowest-confidence cells only, so that the year's simulated map matches the model's own quantity forecast while leaving high-confidence cells undisturbed.

5.4 Python workflow

```
"""
UA-CAM: Uncertainty-Aware Cellular Automata with Memory
Full annual-forecast workflow for ESA CCI land-cover stacks.
Dependencies: numpy, rasterio (or xarray/rioxarray), scipy, scikit-learn
"""

import numpy as np
import rasterio
from scipyndimage import generic_filter
from sklearnlinear_model import LogisticRegression

# ---------------------------------------------------------------
# 1. Load ESA CCI annual stack (one GeoTIFF per year, same grid)
# ---------------------------------------------------------------
def load_cci_stack(paths_by_year):
stack = {}
for year, path in paths_by_yearitems():
with rasterio.open(path) as src:
stack[year] = src.read(1)
profile = src.profile
return stack, profile

# ---------------------------------------------------------------
```

```
# 2. Empirical one-step transition matrices
# ------------------------------------------------------------------
def one_step_transition_matrix(class_prev, class_curr, classes):
K = len(classes)
idx = {c: i for i, c in enumerate(classes)}
counts = np.zeros((K, K))
for a in classes:
mask_a = class_prev == a
n_a = mask_a.sum()
if n_a == 0:
continue
for b in classes:
counts[idx[a], idx[b]] = np.logical_and(mask_a, class_curr == b).sum() / n_a
return counts  # rows sum to 1 (approximately)

# ------------------------------------------------------------------
# 3. Memory-weighted transition kernel
# ------------------------------------------------------------------
def memory_weighted_matrix(history_matrices, lam=0.6):
"""history_matrices: list ordered most-recent-first, length L."""
L = len(history_matrices)
raw_w = np.array([(1 - lam) * lam**tau for tau in range(L)])
w = raw_w / raw_w.sum()
P_tilde = sum(wi * Pi for wi, Pi in zip(w, history_matrices))
return P_tilde

# ------------------------------------------------------------------
# 4. Spell length tracking
# ------------------------------------------------------------------
def update_spell_length(spell_prev, class_prev, class_curr):
changed = class_curr != class_prev
spell_curr = np.where(changed, 1, spell_prev + 1)
return spell_curr

# ------------------------------------------------------------------
# 5. Inertia-adjusted diagonal transition probability
# ------------------------------------------------------------------
def apply_inertia(P_tilde, class_grid, spell_grid, classes, theta_by_class):
K = len(classes)
idx = {c: i for i, c in enumerate(classes)}
P_cell = np.tile(P_tilde, (class_grid.size, 1, 1)).reshape(class_grid.shape + (K, K))
for a in classes:
mask = class_grid == a
theta = theta_by_class[a]
p_aa = P_tilde[idx[a], idx[a]]
boosted = p_aa + (1 - p_aa) * (1 - np.exp(-spell_grid[mask] / theta))
scale = (1 - boosted) / (1 - p_aa + 1e-12)
row = P_cell[mask][:, idx[a], :]
row *= scale[:, None]
row[:, idx[a]] = boosted
P_cell[mask] = np.pad(row, ((0, 0), (0, 0)))[..., None, :].reshape(row.shape[0], 1, K).repeat(1, axis=1) \
if False else row.reshape(-1, K)  # simplified assignment
return P_cell

# ------------------------------------------------------------------
# 6. Neighborhood suitability
# ------------------------------------------------------------------
def neighborhood_fraction(class_grid, target_class, window=7):
```

```
def frac(values):
return np.mean(values == target_class)
return generic_filter(class_grid, frac, size=window, mode="nearest")

def fit_covariate_model(class_grid_prev, class_grid_curr, covariates, classes):
"""Multinomial logistic regression: P(class_t+1 | covariates)."""
X = np.stack([c.ravel() for c in covariates], axis=1)
y = class_grid_curr.ravel()
model = LogisticRegression(multi_class="multinomial", max_iter=500)
model.fit(X, y)
return model

# ---------------------------------------------------------------
# 7. Dempster-Shafer combination
# ---------------------------------------------------------------
def dempster_combine(m1, m2):
"""m1, m2: dicts {frozenset(subset): mass}."""
combined, conflict = {}, 0.0
for A, mA in m1.items():
for B, mB in m2.items():
inter = A & B
if not inter:
conflict += mA * mB
else:
combined[inter] = combined.get(inter, 0.0) + mA * mB
norm = 1.0 - conflict
return {k: v / norm for k, v in combineditems()} if norm > 1e-9 else combined

def pignistic_transform(m, classes):
bet = {c: 0.0 for c in classes}
for A, mass in m.items():
for c in A:
bet[c] += mass / len(A)
return bet

# ---------------------------------------------------------------
# 8. Demand-constrained stochastic allocation
# ---------------------------------------------------------------
def allocate_year(bet_maps, demand_by_class, classes, rng):
"""bet_maps: dict cell_id -> {class: prob}. Returns forecast grid + confidence grid."""
forecast, confidence = {}, {}
for cell_id, dist in bet_maps.items():
probs = np.array([dist[c] for c in classes])
probs = probs / probs.sum()
choice = rng.choice(classes, p=probs)
forecast[cell_id] = choice
confidence[cell_id] = probs.max()
# IPF-style rebalancing pass against demand_by_class omitted here for brevity;
# implemented as a post-processing swap step on lowest-confidence cells.
return forecast, confidence

# ---------------------------------------------------------------
# 9. Full one-year forecast step (orchestration)
# ---------------------------------------------------------------
def forecast_one_year(state, params, rng):
"""state: dict with class_grid, spell_grid, history_matrices, covariates, etc."""
P_tilde = memory_weighted_matrix(state["history_matrices"], lam=params["lambda"])
P_cell = apply_inertia(P_tilde, state["class_grid"], state["spell_grid"],
```

```
params["classes"], params["theta_by_class"])
suit = {k: neighborhood_fraction(state["class_grid"], k, params["window"])
for k in params["classes"]}
# ... combine P_cell, suit, covariate model, and DS uncertainty evidence per cell ...
# (full per-cell loop omitted for brevity; vectorised in production code)
return state  # updated with new class_grid, spell_grid, confidence_grid

# ---------------------------------------------------------------
# 10. Hindcast driver
# ---------------------------------------------------------------
def hindcast(stack, cutoff_year, horizon, params, rng):
results = {}
state = init_state(stack, cutoff_year, params)
for h in range(1, horizon + 1):
state = forecast_one_year(state, params, rng)
results[cutoff_year + h] = (state["class_grid"].copy(), state["confidence_grid"].copy())
return results
```

5.5 R workflow

```
# ---------------------------------------------------------------
# UA-CAM: Uncertainty-Aware Cellular Automata with Memory (R)
# Dependencies: terra, nnet (multinomial logit), dplyr
# ---------------------------------------------------------------
library(terra)
library(nnet)

# 1. Load ESA CCI annual stack ----------------------------------
load_cci_stack <- function(paths_by_year) {
stack_list <- lapply(paths_by_year, rast)
names(stack_list) <- names(paths_by_year)
stack_list
}

# 2. Empirical one-step transition matrix ------------------------
one_step_transition_matrix <- function(r_prev, r_curr, classes) {
K <- length(classes)
P <- matrix(0, K, K, dimnames = list(classes, classes))
v_prev <- values(r_prev); v_curr <- values(r_curr)
for (a in classes) {
idx_a <- which(v_prev == a)
n_a <- length(idx_a)
if (n_a == 0) next
for (b in classes) {
P[as.character(a), as.character(b)] <- sum(v_curr[idx_a] == b) / n_a
}
}
P
}

# 3. Memory-weighted transition kernel ----------------------------
memory_weighted_matrix <- function(history_matrices, lambda = 0.6) {
L <- length(history_matrices)
raw_w <- (1 - lambda) * lambda^(0:(L - 1))
w <- raw_w / sum(raw_w)
Reduce(`+`, Map(function(wi, Pi) wi * Pi, w, history_matrices))
}

# 4. Spell length update ------------------------------------------
update_spell_length <- function(spell_prev, r_prev, r_curr) {
changed <- values(r_prev) != values(r_curr)
spell_curr <- ifelse(changed, 1, values(spell_prev) + 1)
spell_curr
}

# 5. Inertia adjustment -------------------------------------------
apply_inertia <- function(p_aa, spell, theta) {
p_aa + (1 - p_aa) * (1 - exp(-spell / theta))
```

```
}

# 6. Neighbourhood suitability (focal fraction of each class) --------
neighborhood_fraction <- function(r_class, target_class, window = 7) {
bin <- r_class == target_class
w <- matrix(1, window, window)
focal(bin, w = w, fun = mean, na.policy = "omit")
}

# 7. Covariate-based suitability via multinomial logit --------------
fit_covariate_model <- function(df_training) {
# df_training columns: class_t1 (factor), slope, dist_road, dist_urban, pft_frac
multinom(class_t1 ~ slope + dist_road + dist_urban + pft_frac, data = df_training,
trace = FALSE)
}

# 8. Dempster-Shafer combination ------------------------------------
dempster_combine <- function(m1, m2) {
combined <- list(); conflict <- 0
for (A in names(m1)) {
for (B in names(m2)) {
inter <- intersect(strsplit(A, ",")[[1]], strsplit(B, ",")[[1]])
key <- paste(sort(inter), collapse = ",")
prod <- m1[[A]] * m2[[B]]
if (length(inter) == 0) {
conflict <- conflict + prod
} else {
combined[[key]] <- ifelse(is.null(combined[[key]]), prod, combined[[key]] + prod)
}
}
}
norm <- 1 - conflict
lapply(combined, function(x) x / norm)
}

pignistic_transform <- function(m, classes) {
bet <- setNames(rep(0, length(classes)), classes)
for (A in names(m)) {
members <- strsplit(A, ",")[[1]]
for (c in members) bet[c] <- bet[c] + m[[A]] / length(members)
}
bet
}

# 9. Demand-constrained stochastic allocation -----------------------
allocate_year <- function(bet_matrix, classes, demand_by_class, seed = 1) {
set.seed(seed)
n_cells <- nrow(bet_matrix)
forecast <- character(n_cells)
confidence <- numeric(n_cells)
for (i in seq_len(n_cells)) {
probs <- bet_matrix[i, ] / sum(bet_matrix[i, ])
forecast[i] <- sample(classes, size = 1, prob = probs)
confidence[i] <- max(probs)
}
list(forecast = forecast, confidence = confidence)
}

# 10. Hindcast driver -----------------------------------------------
hindcast <- function(stack, cutoff_year, horizon, params) {
results <- list()
state <- init_state(stack, cutoff_year, params)
for (h in seq_len(horizon)) {
state <- forecast_one_year(state, params)
results[[as.character(cutoff_year + h)]] <- state
}
results
}
```

5.6 Parameter calibration

The two central memory parameters, λ (decay) and L (window length), together with the class-specific inertia constants θ_a and the suitability weights α_1, α_2, α_3, are calibrated by a rolling-origin cross-validation procedure entirely within the calibration period (i.e., never touching the held-out hindcast years): for each candidate parameter set, the model is fit using data up to year $t - 1$ and used to forecast year t, for every t inside the calibration window; the parameter set minimizing mean one-step FoM-complement error ($1 - \text{FoM}$) averaged across all such internal folds is selected. This nested design ensures that the hindcast years used for final reporting (Section 6) remain fully out-of-sample with respect to both structural fitting and hyperparameter selection.

6. Hindcast Validation Methodology

The hindcast design follows a strict temporal train/withhold split:

1. **Calibration window**: ESA CCI maps for years 1992 through t_0 (e.g., t_0 = 2015) are used for all parameter estimation, including the empirical transition matrices, memory-kernel calibration, inertia constants, and suitability-model fitting.
2. **Withheld hindcast window**: ESA CCI maps for t_0+1 through t_0+H (e.g., 2016–2020, H = 5) are held out entirely from calibration and used only for scoring.
3. **Multi-step-ahead forecasting**: the model is run forward H years using its own simulated maps (not ground truth) as the basis for updating spell length and rolling transition history at each step, which is the only design that genuinely tests multi-year forecast skill rather than repeated one-step-ahead skill.
4. **Three-map comparison per horizon step**: for each h, the reference map at t_0 (start of interval), the simulated map at t_0+h, and the reference map at t_0+h (end of interval) are compared following the Pontius three-map protocol, from which FoM and its hit/miss/false-alarm components are computed.
5. **Baseline comparators**: three baseline models are run under the identical protocol for comparison — (i) a persistence ("no-change") null model, (ii) a classical two-point CA-Markov model with no memory kernel and no uncertainty propagation, and (iii) a memory-kernel model without the Dempster–Shafer uncertainty layer (deterministic arg-max allocation) — so that the marginal contribution of the memory kernel and of the uncertainty layer can each be isolated.
6. **Stratified reporting**: results are additionally reported stratified by (a) forecast horizon h = 1, ..., 5, to characterize skill decay with lead time; (b) land-cover class, to characterize which classes (typically rapidly changing ones, such as cropland-to-built-up conversion) are hardest to forecast; and (c) spatial heterogeneity strata (homogeneous-neighborhood cells vs. mixed-neighborhood cells), to test whether the uncertainty model's neighborhood-heterogeneity evidence is well calibrated.

7. Evaluation Metrics

- **Figure of Merit (FoM)** and its four components (hits, misses, false alarms, correct rejections of persistence), computed via the three-map Pontius protocol, as the primary change-detection skill metric, because the Figure of Merit technique is recommended specifically for assessing a model's ability to correctly simulate change, and its components indicate whether a model over- or under-estimates change relative to the reference map.
- **Quantity and Allocation Disagreement** (Pontius & Millones, 2011), decomposing total disagreement into the portion attributable to getting the total amount of each class wrong versus the portion attributable to placing the right amount in the wrong location.
- **Fuzzy Kappa** (Hagen, 2003), to credit near-miss spatial allocations within a defined neighbourhood distance, complementing FoM's strict cell-matching criterion.
- **Null-model comparison** (Pontius & Malanson, 2005): comparing agreement between the base map and reference map versus agreement between the simulated map and reference map, to confirm the model outperforms simple persistence.
- **Overall accuracy and traditional Kappa**, reported for continuity with the wider literature, but explicitly flagged as inflated by class imbalance and persistence, per the well-established critique that two-map per cent-correct and Kappa values are typically large whenever persistence dominates a landscape, independent of whether change is simulated correctly.
- **Uncertainty Calibration Score (UCS)** (proposed here): cells are binned by their predicted confidence (BetP of the assigned class), and within each bin the empirical hit rate is compared to the mean predicted confidence; UCS is defined as one minus the mean absolute calibration gap (a reliability-diagram-style summary, analogous to Expected Calibration Error used in machine-learning uncertainty literature), scaled to [0, 1]. A perfectly calibrated model has UCS = 1.
- **Belief–Plausibility interval width**, reported per class and per horizon step, as a direct measure of how much residual ignorance the Dempster–Shafer combination assigns to each forecast; narrowing interval width with additional evidence sources (relative to using only classification-confusion evidence alone) is used as a diagnostic that the combination step is adding information rather than merely adding noise.
- **Skill decay curve**: FoM and UCS plotted against forecast horizon h, to characterise how quickly both change-detection skill and calibration degrade with lead time.

8. Expected Results *(explicitly labelled as anticipated, not empirically observed)*

No hindcast run has yet been executed under this specification; the following are literature-informed, plausible-range expectations intended as pre-registered benchmarks for a future empirical study, not findings.

- **FoM is expected to be modest in absolute terms**, plausibly in the range of 0.10–0.30 at a one-year horizon, consistent with the broader literature's finding that FoM values are characteristically much lower than Kappa or overall accuracy — for example, a large-scale PLUS-based CA study reported FoM = 0.10 alongside Kappa = 0.94 and overall accuracy = 0.97 at the water-basin scale, illustrating that even well-performing models typically show single-digit-to-low-double-digit FoM once persistence is properly excluded from credit.

- **The memory-kernel model is expected to outperform the two-point CA-Markov baseline**, with the largest relative FoM improvement expected in classes exhibiting strong multi-year directional trends (e.g., sustained cropland-to-built-up conversion), and the smallest improvement expected in classes whose transition behaviour is genuinely close to first-order Markovian (e.g., stable natural land cover with only stochastic year-to-year noise).
- **Skill is expected to decay with forecast horizon**, with FoM at $h = 5$ anticipated to be meaningfully lower than at $h = 1$ (plausibly 30–50% relative degradation), consistent with the general finding in the temporal-CA literature that uncorrected error compounds across iterative forecast steps.
- **The Dempster–Shafer uncertainty layer is expected to be reasonably well calibrated but somewhat overconfident**, i.e., UCS below 1 with the gap concentrated in the highest-confidence bin, a common pattern in uncertainty-quantification studies where models are slightly overconfident about their most "obvious" (typically persistence) predictions.
- **Rare-class transitions are expected to show the widest belief–plausibility intervals and the lowest FoM**, reflecting the transition-rarity evidence source functioning as intended (correctly flagging low-sample-size transitions as low-confidence) but also reflecting the intrinsic difficulty of forecasting infrequent conversions from a short historical record.
- **Neighbourhood-heterogeneity stratification is expected to show materially lower FoM and wider intervals in mixed-neighbourhood cells than in homogeneous-neighbourhood cells**, consistent with the general finding that CA-based allocation is least reliable at class boundaries and in fragmented landscapes.

These anticipated patterns should be treated as hypotheses to be confirmed, qualified, or refuted once the pipeline specified in Sections 5–7 is executed on real ESA CCI data for a chosen study region and calibration/hindcast split.

9. DISCUSSION

The framework's central methodological claim is that a transparent, parameterised memory kernel captures a meaningful share of the temporal dependence that motivates the deep-learning temporal-CA literature, without requiring the training data volumes or computational infrastructure of LSTM- or ConvLSTM-based hybrids, and — critically — without sacrificing interpretability: the memory-decay parameter λ, the window length L, and the class-specific inertia constants θ_a are all directly interpretable quantities that a land-system scientist or policy analyst can reason about, in contrast to a learned recurrent hidden state. This interpretability is itself a form of value in policy-relevant land-use forecasting, where stakeholders often need to understand *why* a model expects a given region to change, not merely that it does.

The uncertainty model's contribution should be read as providing decision-relevant information rather than as improving point-forecast accuracy per se: a well-calibrated per-cell confidence layer allows a downstream user (say, a carbon-accounting or hydrological model) to propagate forecast uncertainty into their own outputs, or to flag specific cells or regions where the land-cover forecast should be treated as provisional. This reframes the purpose of "accuracy" reporting in CA-based land-change modelling: rather than reporting a single global Kappa and treating the forecast map as uniformly reliable, the framework is designed so that reliability itself varies spatially and is reported spatially.

The choice of ESA CCI as the validation substrate is deliberate and consequential. Because ESA CCI is the longest consistent global annual land-cover time series available, spanning 1992–2020 at 300 m resolution, it is one of very few datasets that permits a genuine multi-year hindcast (train up to year t_0, validate against several subsequent independently observed years) rather than the single two-point calibration/one-point-test design that characterises most published CA-Markov case studies. This matters because a memory kernel's value proposition is inherently about multi-year dynamics; a validation design that only ever tests one step ahead cannot distinguish a memory-aware model from a well-tuned two-point Markov model.

At the same time, the results anticipated in Section 8 — modest absolute FoM, degrading skill with horizon, imperfect calibration — should be read not as a weakness specific to this framework but as broadly characteristic of the land-change-forecasting literature as a whole, in which two-map comparisons that fail to separate change from persistence are now understood to be misleading precisely because most landscapes are dominated by persistence, leaving comparatively little genuine change signal for any model — memory-aware or otherwise — to detect and correctly place.

10. Limitations

- **Thematic coarseness relative to local drivers.** ESA CCI's 22-class (level 1: ~10-class) legend, while globally consistent, is coarser than many locally relevant land-use categories (e.g., it does not by itself distinguish between different crop types, informal versus planned urban expansion, or specific forest-management regimes), which caps the practical granularity of any forecast built on it, regardless of the CA or uncertainty methodology used.
- **300 m spatial resolution.** Many locally important drivers of land-cover change (parcel-level zoning decisions, individual road construction, small-scale informal settlement growth) operate at sub-pixel scales relative to a 300 m grid, meaning the model necessarily aggregates and smooths some genuine drivers of change into noise.
- **Computational cost of per-cell Dempster–Shafer combination.** Evaluating belief-mass combinations over a full class power set at every cell, every year, over a large study extent is computationally nontrivial at global or continental scale; the framework as specified is realistic for regional-to-national extents on standard workstations

but would require substantial engineering (sparse focal-element representations, GPU vectorisation) to scale globally.

- **Rare-transition estimation remains data-limited irrespective of the memory kernel.** While the transition-rarity evidence source correctly flags low-confidence rare transitions, it cannot manufacture additional historical samples; genuinely novel drivers of change (a new highway, a new protected-area designation) that have no historical analogue in the calibration window are, by construction, difficult for any history-based model, memory-aware or not, to anticipate.
- **No exogenous driver forecasting.** Like most CA-Markov frameworks, UA-CAM forecasts land cover conditional on historical trends and static/quasi-static suitability covariates; it does not forecast future policy shifts, economic shocks, or climate-driven regime changes that could alter transition dynamics discontinuously.
- **Dempster–Shafer combination assumes evidence independence.** The three (or four, including the transition-potential evidence) evidence sources combined via Dempster's rule are treated as independent; in practice, classification confusion, neighbourhood heterogeneity, and transition rarity are likely correlated (e.g., heterogeneous neighbourhoods often coincide with classes that are also more confusable), which can bias the combined belief mass if not corrected.
- **Results in this paper are explicitly anticipatory.** As stated throughout, Section 8 reports plausible ranges informed by the wider validation literature, not the outcome of an executed hindcast run on a specific study area; actual performance will depend materially on the chosen extent, calibration window, class legend, and parameter grid.

11. Future Work

- **Empirical hindcast execution** of the full pipeline specified here on one or more concrete study regions with contrasting change dynamics (e.g., a rapidly urbanising region, a deforestation frontier, and a slowly changing agricultural mosaic), reporting the actual metrics specified in Section 7 against the anticipated ranges in Section 8.
- **Learned memory kernels.** Replacing the fixed exponential-decay weighting (Eq. 2) with a lightweight learned weighting function (e.g., a small attention mechanism over the L historical transition matrices) while retaining the interpretable Dempster–Shafer uncertainty layer, to test how much of the deep-learning temporal-CA literature's accuracy gains can be captured without sacrificing the framework's transparency.
- **Correlated-evidence Dempster–Shafer variants.** Incorporating evidence-source correlation corrections (e.g., via a cautious combination rule or evidence-discounting scheme) to address the independence assumption flagged in Section 10.
- **Sub-pixel disaggregation using PFT fractional layers.** Using the CCI PFT fractional-cover product to disaggregate level-1 class forecasts into finer within-pixel composition estimates, and validating this disaggregation against higher-resolution reference data where available (e.g., national land-cover products in regions where these exist).
- **Multi-resolution hierarchical hindcasting.** Testing whether calibrating the memory kernel at a coarser resolution (e.g., 1 km aggregated blocks) and downscaling to 300 m via the neighbourhood-suitability layer improves both computational tractability and rare-transition estimation stability.
- **Driver-informed suitability layers beyond static covariates.** Incorporating dynamic, time-varying suitability covariates (e.g., annually updated distance-to-active-road-construction layers, or climate/drought indices) rather than only static topographic and distance covariates, to better anticipate discontinuous drivers of change flagged as a limitation in Section 10.
- **Formal uncertainty calibration studies.** Extending the Uncertainty Calibration Score into a full reliability-diagram-based recalibration procedure (e.g., temperature scaling or isotonic regression applied to the pignistic probabilities) to correct any systematic overconfidence identified empirically.
- **Cross-dataset validation.** Repeating the hindcast protocol against independent regional or national land-cover time series (where available) to test whether skill and calibration patterns observed on ESA CCI generalise to higher-resolution or differently classified reference data.

12. CONCLUSION

This paper has specified, in full mathematical and computational detail, an Uncertainty-Aware Cellular Automata with Memory (UA-CAM) framework for annual land-cover transition forecasting, designed explicitly for hindcast validation against the ESA CCI global annual land-cover time series. The framework's two central contributions are a transparent, parameterised memory kernel that replaces the classical two-point Markov transition estimate with a multi-year, inertia-adjusted alternative, and a Dempster–Shafer uncertainty-combination layer that turns every forecast cell's output from a single deterministic class label into a calibrated belief-mass distribution over the full class set. Both Python and R implementations of the complete workflow are provided to support adoption across the two dominant open-source ecosystems in land-system science. A genuine multi-year hindcast validation protocol is specified, built on the three-map Figure of Merit framework and its established critique of naive two-map accuracy statistics, together with a proposed Uncertainty Calibration Score intended to hold the model's stated confidence accountable to its empirical hit rate. Because no run has yet been executed under this specification, all numerical outcomes are presented as literature-informed, explicitly labelled expectations rather than findings; the paper is intended to serve as a complete, reproducible, and pre-registered blueprint for the empirical hindcast study proposed as the first item of future work.

REFERENCES

1. Al-Shaar W, Adjizian G, Nehme N, Lakiss H, Boustany L. Application of modified cellular automata Markov chain model: forecasting land use pattern in Lebanon. Model Earth Syst Environ. 2021;7:1321-1335.
2. Arsanjani JJ, Helbich M, Kainz W, Boloorani AD. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. Int J Appl Earth Obs Geoinf. 2013;21:265-275.
3. Batty M. Cellular automata and urban form: a primer. J Am Plann Assoc. 1997;63(2):266-274.
4. Batty M. *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals.* Cambridge (MA): MIT Press; 2005.
5. Batty M, Xie Y. From cells to cities. Environ Plan B Plan Des. 1994;21(Suppl):S31-S48.
6. Bontemps S, Defourny P, Van Bogaert E, Arino O, Kalogirou V, Perez JR. GLOBCOVER 2009: Products description and validation report. Paris: European Space Agency (ESA); 2011.
7. Chen H, Pontius RG Jr. Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable. Landsc Ecol. 2010;25(9):1319-1331.
8. Clarke KC, Hoppen S, Gaydos L. A self-modifying cellular automaton model of historical urbanisation in the San Francisco Bay area. Environ Plan B Plan Des. 1997;24(2):247-261.
9. Congalton RG. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ. 1991;37(1):35-46.
10. Congalton RG, Green K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices.* 3rd ed. Boca Raton (FL): CRC Press; 2019.
11. Couclelis H. From cellular automata to urban models: new principles for model development and implementation. Environ Plan B Plan Des. 1997;24(2):165-174.
12. Defourny P, Kirches G, Brockmann C, Boettcher M, Peters M, Bontemps S, et al. Land Cover CCI: Product User Guide Version 2. Paris: European Space Agency Climate Change Initiative; 2016.
13. Di Gregorio A, Jansen LJM. *Land Cover Classification System (LCCS): Classification Concepts and User Manual.* Rome: Food and Agriculture Organisation of the United Nations; 2005.
14. Eastman JR. *IDRISI Kilimanjaro: Guide to GIS and Image Processing.* Worcester (MA): Clark Labs, Clark University; 2003.
15. Estoque RC, Pontius RG Jr, Murayama Y, Hou H, Thapa RB, Lasco RD, et al. Simultaneous comparison and assessment of eight remotely sensed maps of Philippine forests. Int J Appl Earth Obs Geoinf. 2018;67:123-134.
16. Foody GM. Status of land cover classification accuracy assessment. Remote Sens Environ. 2002;80(1):185-201.
17. Foody GM. Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sens Environ. 2010;114(10):2271-2285.
18. Fritz S, See L, Rembold F. Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. Int J Remote Sens. 2010;31(9):2237-2256.
19. Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, et al. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data. Int J Remote Sens. 2013;34(7):2607-2654.
20. Hagen A. Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci. 2003;17(3):235-249.
 Halmy MWA, Gessler PE, Hicke JA, Salem BB. Land use/land cover change detection and prediction in the north-western coastal desert of Egypt using Markov-CA. Appl Geogr. 2015;63:101-112.
21. Harper KL, Lamarche C, Hagolle O, Peylin P, Kalogirou V, Bontemps S, et al. A 29-year time series of annual 300 m resolution plant-functional-type maps for climate models. Earth Syst Sci Data. 2023;15(4):1465-1499.
22. He S, Ye X, Sun X, Zheng Q. Simulating urban growth in the Guangdong-Hong Kong-Macao Greater Bay Area using a cellular automata model combined with geographically weighted regression. Land Use Policy. 2023;128:106616.
23. Herold M, Latham JS, Di Gregorio A, Schmullius CC. Evolving standards in land cover characterisation. J Land Use Sci. 2006;1(2-4):157-168.
24. Hyandye C, Martz LW. A Markovian and cellular automata land-use change predictive model of the Usangu Catchment. Int J Remote Sens. 2017;38(1):64-81.
25. Iacono M, Levinson D, El-Geneidy A, Wasfi R. A Markov chain model of land use change. TeMA J Land Use Mobil Environ. 2015;8(3):263-276.
26. Latue PC, Rakuasa H. Prediction of land use and land cover changes in Ternate City using cellular automata Markov chain. Int J Environ Eng Educ. 2023;5(3):87-97.
27. Li X, Yeh AGO. Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int J Geogr Inf Sci. 2002;16(4):323-343.
28. Li X, Chen G, Liu X, Liang X, Wang S, Chen Y, et al. A new global land-use and land-cover change product at a 1 km resolution for 2010 to 2100 based on human-environment interactions. Ann Am Assoc Geogr. 2017;107(5):1040-1059.

29. Liang X, Guan Q, Clarke KC, Liu S, Wang B, Yao Y. Understanding the drivers of sustainable land expansion using a patch-generating land use simulation (PLUS) model: a case study in Wuhan, China. Comput Environ Urban Syst. 2021;85:101569.
30. Liu X, Liang X, Li X, Xu X, Ou J, Chen Y, et al. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. Landsc Urban Plan. 2018;168:94-116.
31. Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. Good practices for estimating area and assessing accuracy of land change. Remote Sens Environ. 2014;148:42-57.
32. Overmars KP, Verburg PH. Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest margin. Int J Geogr Inf Sci. 2005;19(2):125-152.
33. Perica S, Foufoula-Georgiou E. Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. J Geophys Res Atmos. 1996;101(D21):26347-26361.
34. Pontius RG, Malanson J. Comparison of the structure and accuracy of two land change models. Int J Geogr Inf Sci. 2005;19(2):243-265.
35. Pontius RG, Boersma W, Castella JC, Clarke KC, de Nijs T, Dietzel C, et al. Comparing the input, output, and validation maps for several models of land change. Ann Reg Sci. 2008;42(1):11-37.
36. Pontius RG, Millones M. Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. Int J Remote Sens. 2011;32(15):4407-4429.
37. Pontius RG. Criteria to confirm models that simulate deforestation and carbon disturbance. Land. 2018;7(3):1-14.
38. Pontius RG, Peethambaram S, Castella JC. Comparison of three maps at multiple resolutions: a case study of land change simulation in Cho Don District, Vietnam. Ann Assoc Am Geogr. 2011;101(1):45-62.
39. Sang L, Zhang C, Yang J, Zhu D, Yun W. Simulation of land use spatial pattern of towns and villages based on CA-Markov model. Math Comput Model. 2011;54(3-4):938-943.
40. Shafizadeh-Moghadam H, Asghari A, Tayyebi A, Taleai M. Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. Comput Environ Urban Syst. 2019;79:101438.
41. Silva EA, Clarke KC. Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Comput Environ Urban Syst. 2002;26(6):525-542.
42. Soares-Filho BS, Cerqueira GC, Pennachin CL. DINAMICA—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonisation frontier. Ecol Model. 2002;154(3):217-235.
43. Subedi P, Subedi K, Thapa B. Application of a hybrid cellular automaton-Markov (CA-Markov) model in land-use change prediction: a case study of Saddle Creek Drainage Basin, Florida. Appl Ecol Environ Sci. 2013;1(6):126-132.
44. Tobler WR. Cellular geography. In: Gale S, Olsson G, editors. *Philosophy in Geography.* Dordrecht: D. Reidel Publishing; 1979. p. 379-386.
45. Torrens PM. *How Cellular Models of Urban Systems Work.* CASA Working Paper No. 28. London: Centre for Advanced Spatial Analysis, University College London; 2000.
46. Varga OG, Pontius RG, Singh SK, Szabó S. Intensity analysis and the figure of merit's components for assessment of a cellular automata-Markov simulation model. Ecol Indic. 2019;101:933-942.
47. Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SSA. Modelling the spatial dynamics of regional land use: the CLUE-S model. Environ Manage. 2002;30(3):391-405.
48. von Neumann J, Burks AW. *Theory of Self-Reproducing Automata.* Urbana (IL): University of Illinois Press; 1966.
49. Wang F, Iyer R, Ahn C. Logistic regression-based cellular automata for simulating land use dynamics. Environ Model Softw. 2021;138:104988.
50. White R, Engelen G. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Comput Environ Urban Syst. 2000;24(5):383-400.
51. Wolfram S. *A New Kind of Science.* Champaign (IL): Wolfram Media; 2002.
52. Wu F. Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int J Geogr Inf Sci. 2002;16(8):795-818.
53. Wu H, Li Z, Clarke KC, Shi W, Fang L, Lin A, et al. Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. Int J Geogr Inf Sci. 2019;33(5):1040-1061.
54. Xing W, Qian Y, Guan X, Yang T, Wu H. A novel cellular automata model integrated with deep learning for dynamic spatio-temporal land use change simulation. Comput Geosci. 2020;137:104430.
55. Yang Q, Li X, Shi X. Cellular automata for simulating land use changes based on support vector machines. Comput Geosci. 2008;34(6):592-602.

56. Yeh AGO, Li X. Errors and uncertainties in urban cellular automata. Comput Environ Urban Syst. 2006;30(1):10-28.
57. Zhou L, Dang X, Sun Q, Wang S. Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model. Sustain Cities Soc. 2020;55:102045.
58. Zhu Z, Woodcock CE. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens Environ. 2014;144:152-171.
59. ESA Climate Change Initiative Land Cover Project Team. *ESA CCI Land Cover Product User Guide.* Version 2.1. Paris: European Space Agency; 2023.

About the author

Tathagata Satapathy is an Assistant Professor in the Department of Basic Science and Humanities at NIIS Institute of Engineering and Technology, Bhubaneswar, Odisha, India. His research interests include applied sciences, environmental sustainability, interdisciplinary research, engineering education, and sustainable development, with a focus on innovative teaching and research methodologies.

Deepak Kumar Sahoo is an Assistant Professor in the Department of Electrical Engineering at NIIS Institute of Engineering and Technology, Bhubaneswar, Odisha, India. He is dedicated to teaching, research, and academic excellence. His interests include electrical engineering, power systems, renewable energy, and advancing innovative engineering education through scholarly contributions.