

Measuring the Emergence of Urban Cohesion in Early Settlements: A Spatial Agent-Based Model

Luca Pasquino^{1,†}, Niccolò Kadera^{1,†}, Leonardo Mascagni^{1,2}, Andrea Monoli^{1,2} and Francesco Bertolotti^{1,3,*,†}

¹*Intelligence, Complexity, and Technology Lab (ICT Lab), University Cattaneo – LIUC, Italy*

²*School of Industrial Engineering, Cattaneo University – LIUC, Italy*

³*Università Cattolica di Milano, Department of Philosophy, L. Gemelli 1 - 20123 Milano, Italy*

Abstract

The transition from dispersed agricultural settlements to cohesive urban agglomerations represents a fundamental phase in the evolution of any human socio-spatial organization. This paper investigates such a transition through a spatially explicit Agent-Based Model designed to simulate early agricultural societies, one of the simple case study of these phenomena. In the model, autonomous households manage agricultural fields, satisfy metabolic needs, expand their productive infrastructure, and undergo demographic scission when internal population thresholds are exceeded. We introduce an Urban Cohesion Index, a normalized measure that combines Euclidean distances among households with their topological proximity rankings to capture the relational structure of settlements and allows the identification of phase transitions from dispersed rural patterns to cohesive urban cores. Results show that urbanization does not emerge as a gradual consequence of demographic growth, but rather as a non-linear transition triggered by spatial saturation, ecological constraints, and cooperation requirements. Moreover, results indicate that small household capacity, constrained settlement radius, and broader cooperation distance strongly favor spatial clustering, while oversized households and unconstrained demographic growth tend to produce fragmentation, resource depletion, and systemic collapse.

Keywords

Agent-Based Modeling (ABM), Computational Simulation, Urban Emergence, Spatial Dynamics, Complex Systems, Cooperative Behavior,

1. Introduction

The transition from dispersed agricultural settlements to dense urban agglomerations represents a crucial turning point in the evolution of human spatial and social organization [1]. This profound reorganization is not merely the result of linear demographic growth, but emerges from the interaction between environmental constraints, resources management, and the dynamics of social cooperation [2, 3, 4]. Historically, the need to optimize agricultural labor and manage the ecological lifecycle of arable land has confronted individual family units with a fundamental dilemma: to maintain absolute spatial and productive autonomy, facing the risks of soil depletion alone, or to converge toward cooperative strategies [5, 6]. Spatial convergence fosters the sharing of caloric surplus and risk mitigation, effectively triggering the first rudimentary processes of urbanization [7, 8].

The fundamental motivation for this work is to bridge the analytical gap between microscopic household behaviors and macroscopic settlement patterns. Traditional equation-based models often aggregate populations, thereby obscuring the crucial spatial heterogeneity and the localized, non-linear feedbacks between human actions and ecological degradation [9]. To fully understand the mechanics of this transition, it is necessary to move beyond these traditional aggregated approaches in favor of a

WOA 2026, the 27th Workshop From Objects to Agents, June 15–17, 2026, Salerno, Italy

*Corresponding author.

†These authors contributed equally.

✉ lu28.pasquino@stud.liuc.it (L. Pasquino); ni26.kadera@stud.liuc.it (N. Kadera); le21.mascagni@stud.liuc.it (L. Mascagni); an26.monoli@stud.liuc.it (A. Monoli); francesco.bertolotti@unicatt.it (F. Bertolotti)

ORCID 0009-0001-8075-3971 (L. Pasquino); 0000-0003-1274-9628 (F. Bertolotti)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

bottom-up perspective [10]. Agent-Based Models (ABMs) are particularly suitable for the problem addressed because they do not impose top-down equilibrium states; rather, they explicitly simulate the spatial and temporal interactions of autonomous entities over a restricted landscape [11, 3, 12]. This paradigm allows researchers to observe how structured macroscopic patterns, such as urban density, spontaneously emerge from decentralized, micro-founded decisions [13]. As extensively demonstrated in the literature for modeling social dilemmas [14], collective choices [15], epidemiological dynamics [16], and socio-economical systems [17, 12], the computational approach of ABMs captures local interactions that analytical models struggle to describe [18]. Applied to spatial economics and computational archaeology, this method enables the explicit mapping of the topological evolution of settlements [19, 20].

To address these complex dynamics, this paper introduces a spatially explicit ABM [21, 11] specifically designed to simulate the socio-ecological evolution of early agricultural societies [19]. In our simulated environment, autonomous family units act as the primary actors. They are tasked with managing the life cycle of agricultural fields, continuously evaluating their collective caloric needs, and expanding their agricultural infrastructure. In the model, spatial dispersion is a deterministic consequence of demographic pressure: when a family reaches its maximum structural capacity due to population growth, it is forced to undergo fission, expelling members to establish a new, independent settlement. A central challenge in studying these simulated spatial dynamics is to find a generative process [22] that results in a urban-like settlement, and measuring according. Traditional metrics, such as simple aggregate population density, often fail to capture the true topological structure and the degree of clustering within a settlement [23, 24].

To overcome this analytical limitation and accurately evaluate the transition from dispersed farming to agglomerated living, we propose the Urban Cohesion Index (ω). While established spatial metrics—such as Moran’s I [25] for spatial autocorrelation or Ripley’s K function [26] for point pattern analysis—are excellent for identifying general clustering trends, they often struggle to capture the relative, relational topology of dynamically expanding settlements. The Urban Cohesion Index addresses this gap by cross-referencing absolute physical distances with topological density—specifically, the ordinal rank of distances between households. This hybrid approach provides a normalized, highly sensitive measure of spatial clustering. This allows to quantitatively specify the conditions—such as variations in labor productivity or maximum cooperation distance—that trigger the structural collapse of isolated, demographically-driven households into cohesive, centralized urban cores.

The structure of this paper is organized as follows. Section 2 details the model’s methodological architecture, describing the core entities, the metabolic and ecological mechanics, the demographic fission constraints, and the mathematical formalization of the Urban Cohesion Index. Section 3 outlines the experimental setup, detailing the parameter space and the batch simulation protocols utilized for data generation. Section 4 presents the results of the macroscopic data analysis, highlighting the phase transitions in spatial agglomeration and the systemic triggers of urbanization. Finally, Section 5 discusses the implications of these findings for broader theories of urban development and suggests potential avenues for future research.

2. Materials and methods

The overall logical architecture and the execution phases of the ABM are detailed in Algorithm 1

Algorithm 1: Core Execution Engine (Single Time Step)

Input: Parameters par , Global Variables gv , Agents List al

```
1 Increment time step ( $gv.t \leftarrow gv.t + 1$ );
2 Reset step-level state (clear dead entities, set all agents to available);
3  $Groups \leftarrow$  Find spatial cooperation clusters ( $gv, par$ );
4 for each Group in Groups do
5   Identify active families and available agents in Group;
6   if active families > 1 then
7     Set cooperation status to True;
8   end
9   Phase A: Allocate agents to infrastructure construction;
10  Phase B: Calculate aggregate caloric demand and vital debt;
11  Phase C: Allocate agents to farming and extract resources;
12  Phase D: Resolve metabolism (handle starvation mortality or surplus storage);
13  if Group survived and agents are still available then
14    Phase E: Manage territorial expansion (new fields creation);
15  end
16 end
17 Update ecological lifecycle of agricultural fields;
18 Process demographic reproduction and fission events;
19 Collect step data (Micro and Macro);
20 if data buffer exceeds threshold then
21   Flush data to persistent storage;
22 end
23 Perform lazy deletion cleanup for inactive agents and families;
```

2.1. Environment and Core Entities

The computational model is built upon a spatially explicit, discrete two-dimensional squared grid that represents the physical landscape where the simulation unfolds. The primary demographic units within this environment are autonomous agents A_i that stand for the people working in that environment. Agents A_i are organized into collective households defined as H_i . Within the simulation, the household H_i operates as the central socio-economic decision-making unit, managing both labor allocation and spatial positioning.

The physical presence of a household on the grid is anchored by a primary residential structure, defined as the house S_i . This structure occupies a single, specific coordinate cell (x_i, y_i) and serves as the spatial reference point for all distance calculations, social interactions, and agricultural expansions. To sustain themselves, households H_i must develop agricultural infrastructure, represented by fields F . Each field F_k is a dynamic spatial entity characterized by a specific, localized fertility level, which dictates its potential caloric yield.

A family, its central house, and its continuously evolving network of associated fields collectively constitute a Farm Φ . Each farm Φ_i acts as the aggregate agricultural portfolio of the household. The spatial grid itself is managed dynamically by a terrain system that enforces strict topological constraints, ensuring that each discrete cell can be occupied by only one spatially-defined entity at a time.

2.2. Metabolic and Ecological Dynamics

The survival of each family unit is strictly dependent on its ability to satisfy the aggregate metabolic requirements of its members. At each discrete time step, every agent generates a caloric demand α , which the household H_i must meet by extracting resources from its agricultural assets. To fulfill this demand, available agents are deployed as agricultural labor β on the family's operational fields. The extraction of calories is governed by a local ecological threshold, so that each field possesses a finite fertility level γ that determines its maximum potential energy yield δ . As labor is applied and calories are harvested, the field's fertility degrades proportionally to the extracted resources.

This continuous exploitation inevitably leads to soil exhaustion. When a field's fertility γ is completely depleted, it transitions into a fallow state, rendering it temporarily unproductive. The fallow field must then undergo a strictly defined ecological recovery period τ_r before it can be cleared and eventually reclaimed for agricultural use. This ecological constraint forces families to constantly balance their immediate nutritional needs with long-term land management. Furthermore, the system imposes severe consequences for metabolic failure. If a household fails to gather or store sufficient food to cover its vital caloric debt (whether due to an oversized population, labor shortages, or widespread soil exhaustion) the family faces starvation. In such critical events, the entire household H_i is removed from the simulation as well as their associated infrastructures, freeing the spatial cells for future occupation.

2.3. Spatial Settlement and Expansion Mechanisms

To sustain their growing internal population n_H , households H_i must continuously expand their agricultural footprint. This expansion is defined by a deterministic spatial algorithm that performs a concentric search, starting from the primary residential structure, to locate the nearest available grid cell for establishing a new field. This ensures both explainability and that agricultural infrastructure remains as contiguous and accessible as possible. However, as the population naturally increases, the household eventually reaches its maximum structural capacity. This demographic pressure triggers a mandatory scission event, forcing a subset of the family members to migrate and establish a new, independent settlement. The placement of this new household is not random across the global landscape, but is controlled by a proximity parameter η , which defines a specific maximum exploration radius around the parent settlement. The mechanism identifies all unoccupied spatial cells within this radius and allocates the new house accordingly.

2.4. The Urban Cohesion Index

To quantitatively assess the emergence of urban structures from the previously described expansion mechanics, we introduce the Urban Cohesion Index (ω). Traditional density metrics often fail to capture the relational topology of settlements, treating uniformly dispersed populations similarly to clustered ones. Therefore, ω is specifically designed to evaluate spatial agglomeration by cross-referencing absolute physical distance with topological density, defined as distance rank.

Let $P = \{p_1, p_2, \dots, p_n\}$ be the set of coordinates for n active households within the grid. We define d_{ij} as the Euclidean distance between household i and j , and $r_{ij} \in \{0, \dots, n-1\}$ as the ordinal rank of household j in terms of distance from i (with $r_{ii} = 0$). The global cohesion index is formulated as follows:

$$\omega = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n \left(1 - \frac{r_{ij}}{n-1} \cdot \frac{d_{ij}}{d_{max}} \right) \quad (1)$$

where d_{max} represents the maximum Euclidean distance observed between any two households in the settlement. The index is rigorously normalized in the range $[0, 1]$. A score approaching 1.0 indicates maximum centralized density, where all units are perfectly clustered, whereas lower scores reflect high spatial fragmentation.

The computational procedure, optimized for large-scale simulations via matrix operations, is detailed in Algorithm 1. By tracking the trajectory of this index over the simulation steps, the model allows the identification of phase transitions from dispersed rural patterns to cohesive urban agglomerations.

Algorithm 2: Urban Cohesion Index (ω) Calculation

Input: List of household positions P , population size n

Output: Normalized Urban Cohesion Index ω

```

1 if  $n < 2$  then
2   | return 0.0;
3 end
4 Compute Distance Matrix  $D_{n \times n}$  where  $D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$ ;
5  $d_{max} \leftarrow \max(D)$ ;
6 if  $d_{max} = 0$  then
7   | return 1.0 ;                                // Case where all households overlap
8 end
9 Compute Rank Matrix  $R_{n \times n}$  by sorting  $D$  row-wise;
10 Initialize  $S \leftarrow 0$ ;
11 for  $i \leftarrow 1$  to  $n$  do
12   | for  $j \leftarrow 1$  to  $n$  do
13     | if  $i \neq j$  then
14       |  $C_{ij} = 1.0 - \left( \frac{r_{ij}}{n-1} \cdot \frac{d_{ij}}{d_{max}} \right)$ ;
15       |  $S \leftarrow S + C_{ij}$ ;
16     | end
17   | end
18 end
19 return  $S / (n(n-1))$ ;

```

3. Experimental Setup and Scenario Analysis

To robustly evaluate the model and observe the macroscopic emergence of urban agglomerations, we conducted an extensive campaign of batch simulations. The experimental setup systematically explored the parameter space to identify phase transitions in the spatial behavior of the households. The analysis specifically focused on varying key parameters, as in 1.

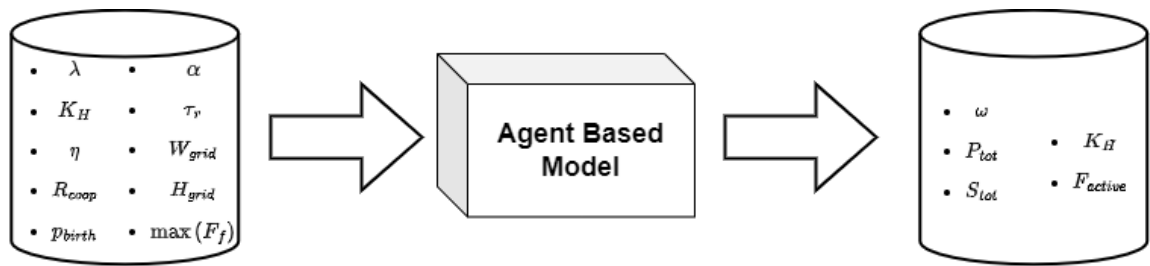


Figure 1: Black box representation of the experimental setup

A comprehensive overview of the primary parameters manipulated during the batch simulation phase is summarized in Table 1.

Table 1: Key experimental parameters manipulated during the batch simulation runs, detailing their specific functional definitions and systemic impacts on the model's macro-dynamics.

Parameter	Value/Range	Description	Macro-Dynamic Influence
λ	5000	<i>Labor Productivity</i> : Caloric output generated by a single agent per work cycle.	Dictates baseline metabolic survival and global food storage capacity.
K_H	$\{2, 3, \dots, 39, 40\}$	<i>Max People per Family</i> : Structural capacity limit of a single house before mandatory fission.	Acts as a demographic bottleneck, triggering spatial expansion when low.
η	6	<i>Proximity Parameter</i> : Maximum exploration radius allowed for placing a new house during a fission event.	Constrains spatial dispersion, forcing local density and cluster formation.
R_{coop}	4	<i>Cooperation Dist.</i> : Maximum topological radius within which different families share labor and resources.	Accelerates the Urban Cohesion Index by incentivizing spatial convergence.
p_{birth}	$\{0.1, 0.2, \dots, 0.9, 1.0\}$	<i>Birth Probability</i> : Baseline probability of a new agent being born in a household per time step.	Drives demographic pressure and accelerates household mortality when extreme.
α	2250	<i>Caloric Demand</i> : Fixed amount of calories consumed by each agent per time step.	Determines the aggregate metabolic pressure on the agricultural infrastructure.
τ_r	5	<i>Recovery Period</i> : Number of time steps required for a completely depleted fallow field to restore its fertility.	Forces families to constantly balance immediate nutritional needs with long-term spatial expansion.
γ_{max}	40000	<i>Max Fertility</i> : The absolute maximum limit of fertility a field can possess or recover to.	Defines the baseline environmental richness and the long-term carrying capacity of the simulated grid.
W_{grid}, H_{grid}	$\{10 \times 10, 20 \times 20, \dots, 190 \times 190, 200 \times 200\}$	<i>Terrain Dimensions</i> : Total width and height of the simulated spatial grid (available cultivable/buildable cells).	Defines the absolute carrying capacity of the environment and dictates the onset of global spatial saturation.

Table 2: Dynamic demographic and ecological variables utilized during the simulation runs.

Variable	Description	Macro-Dynamic Influence
β	<i>Agricultural Labor</i> : Number of available agents deployed as active labor on the family’s operational fields.	Determines the rate of resource extraction and the pace of local soil degradation.
γ	<i>Soil Fertility Level</i> : Dynamic ecological variable representing the current productive capacity of a specific field.	Dictates the potential caloric yield and governs the transition into the fallow state.

The baseline values for critical metabolic and ecological parameters—such as caloric demand ($\alpha = 2250$), labor productivity ($\lambda = 5000$), field maximum fertility ($\gamma_{max} = 40000$), and the ecological recovery period ($\tau_r = 5$)—were established through exploratory parameter tuning rather than direct derivation from specific empirical historical datasets. These values were selected for computational convenience to ensure that the baseline simulation produces a viable, balanced metabolic cycle where agents can survive, reproduce, and face ecological pressures without immediate systemic extinction. While highly abstracted, the ratio between caloric demand and labor productivity loosely mimics the subsistence margins of early agrarian societies, allowing the model to explore the

general mechanisms of spatial phase transitions rather than replicate a specific archaeological case study.

Data acquisition was designed to capture both collective phenomena and individual trajectories. At the macroscopic level, the simulation continuously recorded aggregate variables at each time step, including total population size, global food storage, the number of active agricultural fields, and the Urban Cohesion Index. Concurrently, microscopic data streams logged the precise life cycle, spatial coordinates, and resource accumulation of individual families, providing a high-resolution map of the demographic and spatial evolution. To ensure scientific reproducibility, each simulation run was initialized with a specifically recorded random seed, allowing for the exact reconstruction of stochastic events and the isolated observation of specific parametric effects. This massive generation of simulated longitudinal data provided the necessary statistical foundation to correlate localized micro-behaviors with macro-spatial urban outcomes.

The computational model, as well as the execution of all batch simulations, were fully implemented in Python (version 3.13.1). To ensure high computational efficiency and reproducibility, the core mathematical operations and spatial distance matrices were optimized using *NumPy* (v. 2.4.4) and *SciPy* (v. 1.17.1). Data manipulation, logging, and macroscopic analysis were conducted via *pandas* (v. 3.0.2). The visual rendering of the simulation environment and the generation of macroscopic plots were managed through *Pygame* (v. 2.6.1), *Pillow* (v. 12.2.0), and *Matplotlib* (v. 3.10.8), while *tqdm* (v. 4.67.3) was utilized for iterative process tracking. Finally, all simulation batches were executed on a local workstation equipped with an Intel Core i5-L16G7 processor and 8 GB of RAM, running on the Windows 11 Pro operating system.

4. Results

4.1. Population and Resource Dynamics

In scenarios characterized by low labor productivity or highly constrained initial fertility, the system frequently experiences a premature demographic collapse. In these instances, isolated households are unable to overcome the vital caloric debt, leading to early starvation events that prevent any meaningful settlement expansion.

However, even when initial conditions allow for successful colonization and massive demographic growth, the environment remains highly vulnerable to ecological thresholds. If the rapid population expansion outpaces the recovery capacity of the agricultural infrastructure, the system falls into a severe ‘overshoot and collapse’ trajectory. A spatial visualization of this specific dynamic—where the population completely saturates the grid before ultimately suffering a systemic starvation event and total extinction is illustrated in Figure 2.



Figure 2: Visual progression of an “overshoot and collapse” scenario within the spatial grid. From left to right: (1) Initial colonization with abundant fertile land; (2) Intermediate expansion and structural clustering; (3) Peak demographic saturation and extreme metabolic stress; (4) Systemic starvation event resulting in total population collapse and a landscape of exhausted fallow fields.

Conversely, when baseline agricultural parameters allow sustainable resource extraction, the global population exhibits rapid initial growth followed by a stabilization phase. During this stabilization, the aggregate food storage acts as a buffer against localized soil exhaustion. This demographic sustainability is the prerequisite for spatial complexity, as only environments that support a critical mass of active households generate the spatial pressure necessary for structural reorganization.

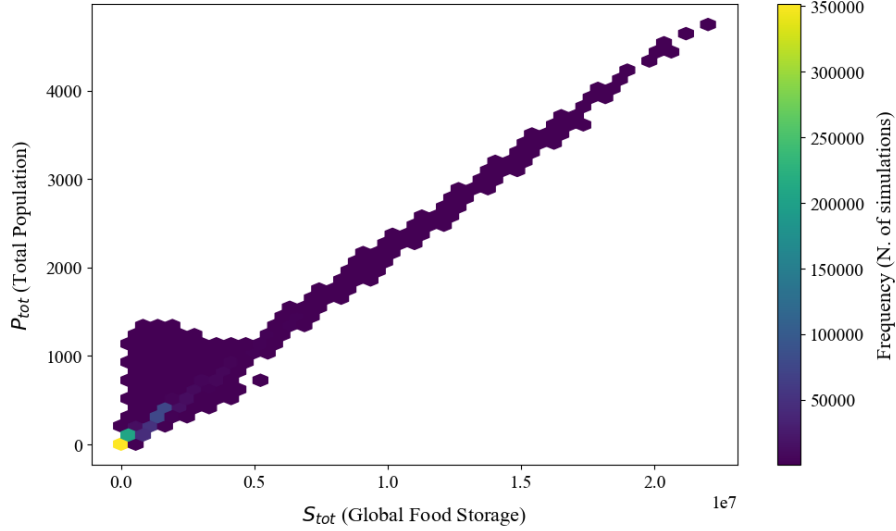


Figure 3: Total population versus global food storage. A threshold at 1,500 units marks a phase transition, from a first vulnerable, high-variance resource accumulation phase to a second highly resilient, standardized linear scaling phase.

The structural relationship between demographic growth and resource accumulation is further shown by mapping total population P_{tot} against global food storage S_{tot} . While a strict, positive linear correlation dominates the macro-scale—indicating a standardized metabolic requirement—the system exhibits significant instability at lower demographic levels. For populations under 1,500 units, the data reveals a high-variance cluster with food reserves falling well below the main trendline. This variance reflects the metabolic vulnerability and relative inefficiency of small, isolated rural households. However, once the population surpasses this critical threshold of 1,500, the variance collapses and the data aligns perfectly along the primary accumulation axis. This phenomenon highlights a clear systemic resilience achieved through critical demographic mass, where aggregate food storage scales up robustly to the order of tens of millions, buffering the urbanizing population against localized ecological exhaustion.

4.2. Phase Transition in Urban Cohesion

The most significant findings emerge from the temporal analysis of the Urban Cohesion Index ω . The data demonstrate that urbanization within the model does not occur as a gradual, linear progression, but exhibits clear phase transitions. In the early stages of simulation, the index remains remarkably low, reflecting a highly fragmented landscape of isolated farms driven by the demographic fission mechanism. However, as the spatial grid reaches a critical saturation point—where the distance between newly founded settlements and existing infrastructure minimizes— ω undergoes a sudden and exponential increase. This spike quantitatively marks the structural collapse of the dispersed rural pattern and the spontaneous self-organization of dense, centralized urban cores.

To further evaluate the structural role of environmental scale, we analyzed the direct impact of grid dimensions ($W_{grid} \times H_{grid}$) on these macro-dynamics. Smaller grid sizes (from 10×10 to 50×50) significantly accelerate the onset of spatial saturation, compressing the time required to trigger the exponential increase in ω . However, these constrained environments severely limit the maximum

carrying capacity of the system, making them highly prone to early ecological overshoot and systemic collapse. Conversely, expanding the grid dimensions up to 200×200 delays the non-linear regime shift of the Urban Cohesion Index, as households can exploit large unoccupied areas before encountering spatial competition. Nonetheless, once saturation eventually occurs in larger grids, the system stabilizes into highly resilient urban cores capable of sustaining a much higher critical demographic mass, confirming that environmental space acts as a fundamental boundary condition for the timing and scale of urban emergence.

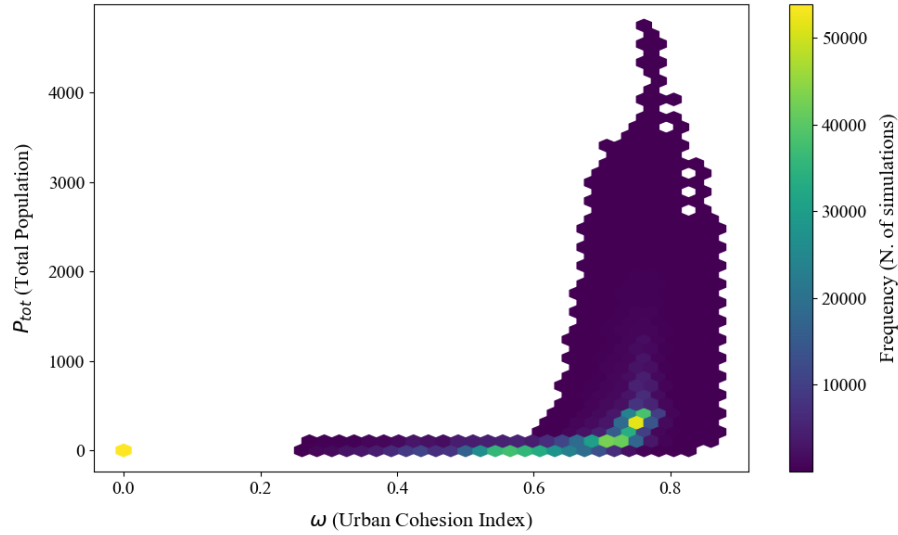


Figure 4: Scatter plot of P_{tot} against ω . A critical spatial threshold at 0.6 triggers a non-linear regime shift, unlocking massive demographic growth that peaks between 0.7 and 0.8 before reaching an upper metabolic limit at extreme densities.

Furthermore, the scatter plot analysis relating total population to ω reveals a distinct non-linear regime shift. The data identifies a critical spatial threshold around a cohesion value of 0.6; below this point, populations remain universally constrained to low numbers (under 500 individuals). However, crossing this 0.6 threshold unlocks massive demographic variability, allowing the system to sustain populations up to 5000 units, peaking between cohesion $0.7 < \omega < 0.8$. Interestingly, extreme cohesion ($\omega > 0.8$) does not correlate with further population growth, suggesting an upper metabolic limit to extreme urban density.

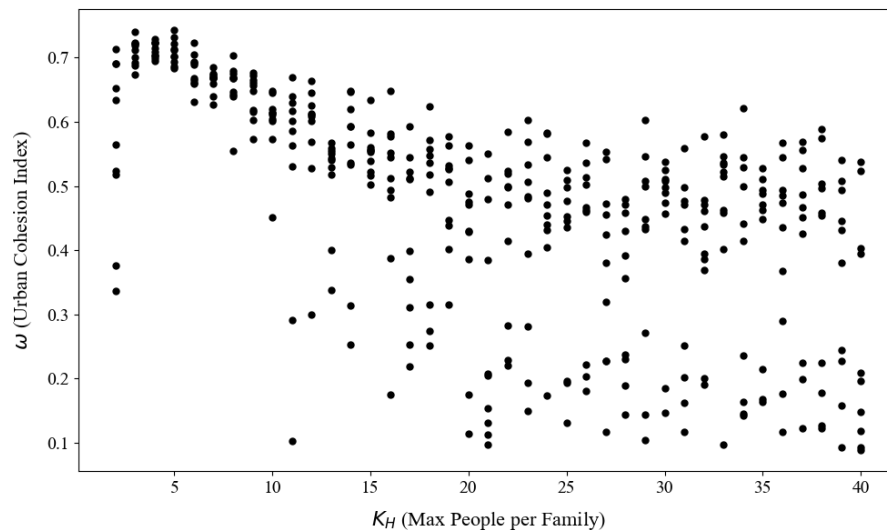


Figure 5: Urban Cohesion Index ω versus maximum household capacity. The data reveals optimal spatial clustering for small households (3-4 individuals) and extreme spatial fragmentation for limits exceeding 15 individuals.

The underlying mechanics of this demographic bottleneck are fully shown when correlating the maximum household capacity directly with ω . The distribution of these data points clearly presents two distinct regimes. In the high-cohesion regime—characterized by strict capacity limits between 2 and 15 individuals—the system exhibits remarkable spatial order. Cohesion peaks optimally ($\omega > 0.7$) for small family units of 3 to 4 individuals. This strict limitation on household size continuously forces proximity-constrained fission events, which organically generate the dense spatial clustering necessary for macro-level urban aggregation. Conversely, allowing a capacity greater than 15 individuals shifts the system into a regime of high structural entropy. In this state, the cohesion index fragments into a highly dispersed, unpredictable cloud, frequently collapsing to minimal values ($0.1 < \omega < 0.3$). This spatial disintegration confirms that oversized households suppress the cooperative clustering required to optimize resource extraction, leading to a fragmented and fragile landscape.

4.3. Triggers of Agglomeration

The batch scenario analysis further isolates the specific parameters that accelerate or delay this phase transition. The maximum cooperation distance plays a pivotal role; scenarios with a broader cooperation radius consistently exhibit earlier and more pronounced spikes in ω , as households H_i actively cluster to deal with collective labor and share caloric surpluses. Furthermore, the households split radius acts as a strong spatial constraint. When the split radius is tightly restricted, newly formed H_i are forced to settle in immediate proximity to their parent structures, accelerating local density. The relationship between these parameters confirms that early urbanization, as captured by the cohesion index, is an emergent response at a system level driven by the need to balance metabolic survival with extreme spatial competition.

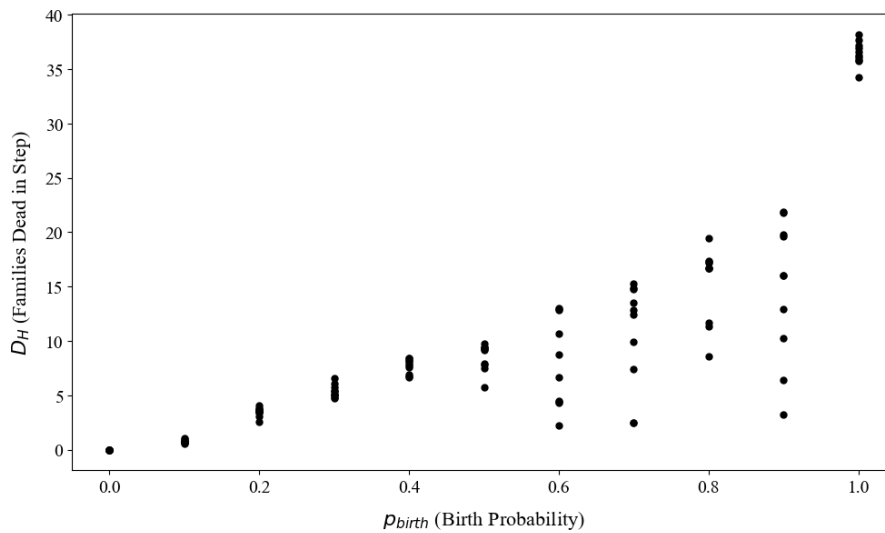


Figure 6: Impact of birth probability on household mortality. Unconstrained natality triggers a non-linear spike in family failures, highlighting the strict ecological limits and the self-regulating nature of the simulated ecosystem.

The analysis of specific demographic parameters further highlights the ecological limits of the simulated environment. By plotting the baseline birth probability against the number of household failures per time step, a clear non-linear relationship emerges. At lower birth probabilities p_{birth} , family mortality remains predictably stable. However, as the birth probability increases, the system exhibits severe heteroskedasticity, with the variance in household failure rates expanding drastically. This trend

culminates in a qualitative breaking point at the maximum probability threshold $\max p_{birth}$ (which is 1.0), where household mortality spikes abruptly to nearly double the rate observed at 0.9. Taken together with the previous macro-dynamics, these results portray a highly complex, self-regulating ecosystem: while spatial cohesion unlocks population growth and demographic mass stabilizes resource storage, unconstrained natality inevitably triggers accelerated household mortality due to extreme spatial and metabolic saturation.

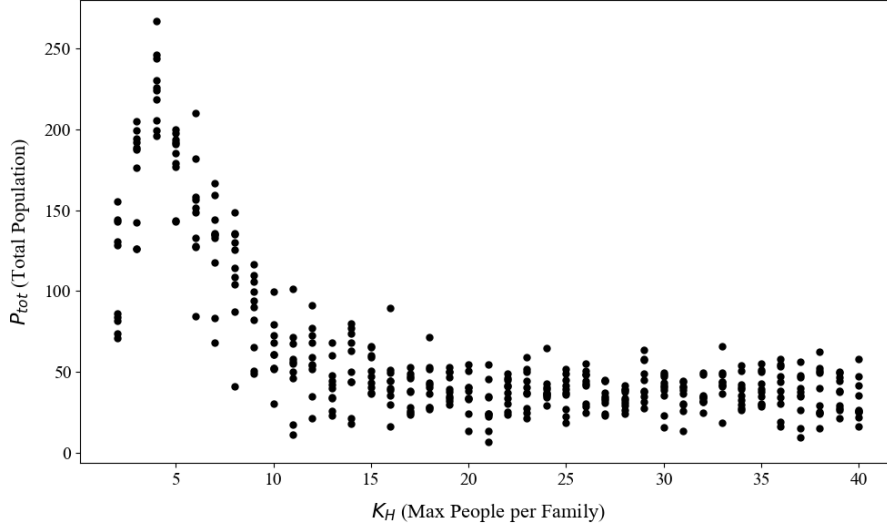


Figure 7: Relationship between maximum family size and global population. The data reveals a counterintuitive dynamic where small, highly constrained households (4-5 members) maximize total carrying capacity, while larger family structures prove inherently inefficient for systemic demographic growth.

Further parametric analysis reveals a highly counterintuitive dynamic regarding household population capacity. By correlating the maximum allowed persons per family with the total sustained population, a severe bottleneck effect becomes evident. The data demonstrates a peak systemic efficiency when the household limit is highly constrained, specifically between 4 and 5 individuals, allowing the global population to maximize its carrying capacity. However, as the maximum family size exceeds this narrow optimum, the system experiences drastic diminishing returns; at a limit of 10 persons, the maximum achievable global population P_{tot} is effectively halved, and for limits between 15 and 40, the global population stabilizes at a severely depressed plateau. This phenomenon indicates that oversized households are inherently inefficient within the model’s ecological constraints. Large families concentrate caloric demand in localized areas, rapidly exhausting nearby soil fertility γ and increasing the risk of massive starvation events. Conversely, artificially restricting household size acts as a catalyst for continuous demographic fission, forcing an aggressive spatial expansion that optimizes global resource extraction and ensures greater systemic resilience.

5. Discussion and Conclusions

The computational experiments presented in this study provide an exploratory, abstract framework demonstrating how macro-level urban agglomerations might emerge from micro-founded, decentralized household decisions. Consistent with phenomena observed in established computational archaeology models, such as the Artificial Anasazi [19], and broader theories of urban scaling [8], our results highlight the critical role of metabolic thresholds in driving spatial self-organization. By utilizing the Urban Cohesion Index ω , we were able to observe that the transition from a dispersed agricultural landscape to a centralized urban structure is not a gradual, linear consequence of population growth.

Rather, it manifests as a sudden phase transition triggered by spatial saturation and metabolic constraints. When the availability of highly fertile, unoccupied land dwindles due to continuous demographic fission and localized ecological degradation, spatial clustering ceases to be merely an option and becomes a systemic necessity for survival. Also, the demographic sustainability appears to be a prerequisite for spatial complexity. Finally, the model illustrates that early urbanization is fundamentally an adaptive response to environmental pressure, where proximity and cooperation optimize resource extraction and mitigate the risks of starvation.

These findings carry deep theoretical implications for the understanding of early spatial economies and urban settings, reinforcing the perspective that the first permanent settlements were complex, self-organizing adaptive systems. From a theoretical standpoint, these results challenge linear models of urban growth, suggesting that archaeological and historical research should look for signatures of sudden phase transitions rather than gradual accumulation. The dynamics observed in our simulations suggest that physical agglomeration is deeply intertwined with the life cycle of the surrounding agricultural infrastructure. The necessity to balance the ecological recovery of fallow fields with immediate caloric demands forces populations to negotiate their spatial distribution, ultimately driving them toward dense, cooperative hubs when spatial expansion is no longer viable. Practically, the newly introduced Urban Cohesion Index provides a transportable, robust metric that can be applied to empirical archaeological datasets or even modern rural-to-urban transitions to detect early warning signals of critical spatial shifts. Future studies can leverage this metric to compare simulated settlement trajectories with actual historical footprints, grounding computational hypotheses in empirical reality.

Despite these insights, the current model possesses several limitations that offer clear avenues for future research. Primarily, the simulated landscape is topologically isotropic, lacking geographical variables such as rivers, terrain morphology, or variations in baseline soil quality, which historically dictated the ecological realism and placement of early cities. Additionally, the agents in our model rely on highly homogeneous rules. All agents consume the same caloric amount and generate identical labor productivity, ignoring demographic variability such as age, health status, or physical capacity, factors that significantly affect household resilience during ecological crises. Real-world social systems are also characterized by the eventual emergence of hierarchical structures, social differentiation, and specialized, non-agricultural labor, which are central to the development of early urban societies but remain absent here. Also, there is a single source of calories, and no meteorological or seasonal events are included. Finally, households are never in conflict, and the urban settlements are a purely emergent result of metabolic needs, while literature suggests it could have depended also from the need for a common and more organized defense [1].

Future extensions of this framework should introduce environmental heterogeneity and socio-economic stratification to observe how unequal resource distribution affects ω . Furthermore, incorporating technological shocks, such as innovations in farming tools that alter labor productivity mid-simulation, could provide deeper insights into the resilience and evolution of these early urban centers. Finally, LLM could be employed as cognitive engines of agents [27].

Code Availability

The complete source code of the simulation, along with the full experimental parameters and baseline tests necessary to replicate the findings, is publicly available on GitHub at <https://github.com/lucapasquino/farm-project>.

Declaration on Generative AI

During the preparation of this work, the authors used OpenAI Codex in order to:

- Drafting content,
- Paraphrase and reword,
- Improve writing style,
- Grammar and spelling check.

After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

- [1] P. Turchin, et al., Ultrasociety how 10,000 years of war made humans the greatest cooperators of earth (2016).
- [2] M. Burtsev, P. Turchin, Evolution of cooperative strategies from first principles, *Nature* 440 (2006) 1041–1044.
- [3] F. Bertolotti, S. Roman, The evolution of risk sensitivity in a sustainability game: an agent-based model., in: WOA, 2022, pp. 101–115.
- [4] F. Bertolotti, S. Roman, Balancing long-term and short-term strategies in a sustainability game, *Iscience* 27 (2024).
- [5] G. Hardin, Extensions of” the tragedy of the commons”, *Science* 280 (1998) 682–683.
- [6] E. Lecoutere, L. Jassogne, Fairness and efficiency in smallholder farming: The relation with intrahousehold decision-making, *The Journal of Development Studies* 55 (2019) 57–82.
- [7] G. B. West, J. H. Brown, Life's universal scaling laws, *Physics today* 57 (2004) 36–42.
- [8] L. M. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, G. B. West, Growth, innovation, scaling, and the pace of life in cities, *Proceedings of the national academy of sciences* 104 (2007) 7301–7306.
- [9] F. Bertolotti, L. Mari, A methodology for limit cycle detection in simulation models, in: EUROSIM Congress, Springer Nature Switzerland Cham, 2023, pp. 317–331.
- [10] C. Berceanu, F. Bertolotti, N. Arshad, M. Patrascu, Understanding the mechanisms of infodemics: Equation-based vs. agent-based models, *Plos one* 20 (2025) e0338614.
- [11] F. Bertolotti, R. Occa, “roads? where we're going we don't need roads.” using agent-based modeling to analyze the economic impact of hyperloop introduction on a supply chain, in: European Conference on Multi-Agent Systems, Springer, 2020, pp. 493–500.
- [12] F. Bertolotti, S. Roman, Risk sensitive scheduling strategies of production studios on the us movie market: An agent-based simulation, *Intelligenza Artificiale* 16 (2022) 81–92.
- [13] R. Occa, F. Bertolotti, et al., Understanding the effect of iot adoption on the behavior of firms: An agent-based model, in: CS & IT Conference Proceedings, volume 12, CS & IT Conference Proceedings, 2022.
- [14] R. M. Dawes, Social dilemmas., *Annual review of psychology* (1980).
- [15] M. Marsili, D. Challet, R. Zecchina, Exact solution of a modified el farol's bar problem: Efficiency and the role of market impact, *Physica A: Statistical Mechanics and its Applications* 280 (2000) 522–553.
- [16] F. Bertolotti, N. Kadera, L. Pasquino, L. Mari, An epidemiological extension of the el farol bar problem, *Frontiers in Big Data* 8 (2025) 1519369.
- [17] F. Bertolotti, S. Roman, Risk sensitivity of production studios on the us movie market: an agent-based simulation., in: WOA, 2021, pp. 210–223.
- [18] H. Rahmandad, J. Sterman, Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models, *Management science* 54 (2008) 998–1014.
- [19] R. L. Axtell, J. M. Epstein, J. S. Dean, G. J. Gumerman, A. C. Swedlund, J. Harburger, S. Chakravarty, R. Hammond, J. Parker, M. Parker, Population growth and collapse in a multiagent model of the kayenta anasazi in long house valley, *Proceedings of the National Academy of Sciences* 99 (2002) 7275–7279.
- [20] A. Piras, F. Bertolotti, How risk preferences shape city-state success: An agent-based model of resource management., in: WOA, 2024, pp. 283–297.

- [21] T. Filatova, P. H. Verburg, D. C. Parker, C. A. Stannard, Spatial agent-based models for socio-ecological systems: Challenges and prospects, *Environmental modelling & software* 45 (2013) 1–7.
- [22] J. M. Epstein, *Generative social science: Studies in agent-based computational modeling*, Princeton University Press, 2012.
- [23] Y. Song, Y. Long, P. Wu, X. Wang, Are all cities with similar urban form or not? redefining cities with ubiquitous points of interest and evaluating them with indicators at city and block levels in china, *International Journal of Geographical Information Science* 32 (2018) 2447–2476.
- [24] G. Dumedah, E. K. Garsonu, Characterising the structural pattern of urban road networks in ghana using geometric and topological measures, *Geo: Geography and Environment* 8 (2021) e00095.
- [25] P. A. P. Moran, Notes on continuous stochastic phenomena, *Biometrika* 37 (1950) 17–23.
- [26] B. D. Ripley, The second-order analysis of stationary point processes, *Journal of Applied Probability* 13 (1976) 255–266.
- [27] F. Bertolotti, S. Roman, F. Carucci, G. Buonanno, L. Mari, An llm-enhanced agent-based model of a sustainability game, in: *Proceedings of the 26th Workshop From Objects to Agents (WOA2025)*, 2025, pp. 02–05.