Computational Macrohistory: Axiomatic Foundations for a Quantitative Science of Large-Scale Social Systems

Stefano Angeli
Foundations Institute of Computational Social Science (FICSS)
stefano.angeli@ficss.institute

Version 3.0, July 2026

Abstract

In December 1989, few analysts predicted the Soviet collapse within two years. The failure was not one of intelligence but of framework: available models did not distinguish structural forecasts (rising probability of instability) from event-level predictions (the exact date). This distinction, which Computational Macrohistory makes axiomatic, is the central contribution of this paper. We establish eight axioms (A1 through A8) as admissibility constraints for valid CMH models, and we extend the foundational treatment of the preceding versions with four formal theorems. Theorem 1 derives predictive decay from endogenous indeterminacy and isolates the linear divergence regime from the saturating accuracy baseline. Theorem 2 formalises the aggregation condition into an explicit sample-size threshold under intra-class correlation, yielding the CMH Herding Threshold: a correlation ceiling beyond which aggregation fails regardless of population size. Theorem 3 applies the Banach fixed-point theorem to reflexive feedback under the Wasserstein-1 metric, establishing conditions for self-consistent predictions and an empirical identifiability strategy via regression discontinuity. Theorem 4 anchors the falsifiability requirement to the Murphy decomposition of the Brier score. A formal bistable model with noise-induced switching is developed in Appendix C, with Feller boundary analysis confirming domain invariance. The axioms do not describe the world; they define the conditions under which historical systems become scientifically tractable.

Keywords: computational macrohistory, cliodynamics, historical prediction, complex systems, social dynamics, quantitative history, axiomatics, structural causality, nonlinear dynamics

JEL Classification: C02, C63, N00, D74, Z13

Suggested citation: Angeli, S. (2026). *Computational Macrohistory: Axiomatic Foundations for a Quantitative Science of Large-Scale Social Systems* (FICSS Working Paper No. WP-2026-001, v3.0). Foundations Institute of Computational Social Science. https://doi.org/10.5281/zenodo.21178550

1 Introduction

1.1 The challenge of historical science

In December 1989, the Soviet Union appeared to most Western analysts as a stable, if sclerotic, superpower. The Berlin Wall had fallen the previous month, but the internal cohesion of the USSR itself was not widely questioned. By December 1991, the Soviet state had ceased to exist. The failure to anticipate this outcome was not primarily a failure of intelligence gathering or analytical competence. It was a structural failure of the predictive framework itself. The models available to analysts (informal mental models, scenario-based geopolitical assessments, qualitative expert judgment) lacked the machinery to separate two questions that Computational Macrohistory treats as axiomatically distinct: whether a system is approaching a structural tipping point, and when exactly it will cross.

The distinction is not academic. A framework that can estimate that the probability of severe instability in a given country has risen from fifteen to fifty-five per cent over a decade provides actionable structural warning, even though it cannot name the month or the spark. A framework that promises the date and fails is worse than none. The difference between the two (probabilistic structural forecasting and deterministic event prediction) is the organising principle of the present paper.

The aspiration to identify regularities in human history is ancient. From Ibn Khaldun's cyclical theory of civilizations in the fourteenth century to Marx's historical materialism and Spengler's morphology of cultures, thinkers have sought patterns beneath the apparent chaos of historical events. Yet history has remained largely resistant to the kind of systematic, predictive science that characterises physics, chemistry, or biology.

This resistance stems from legitimate concerns. Human agency, cultural uniqueness, the apparent non-repeatability of historical events, and the reflexive nature of social knowledge all seem to preclude the kind of prediction available in the natural sciences. As the philosopher of history R. G. Collingwood argued, historical events are expressions of human thought, and understanding them requires interpretive methods fundamentally different from natural-scientific explanation.

The past two decades have nonetheless seen significant progress in quantitative historical analysis. The cliodynamics program initiated by Peter Turchin and colleagues (Turchin, 2003; Turchin & Nefedov, 2009) has shown that mathematical models can capture recurrent patterns in historical dynamics, including secular cycles of political instability, elite overproduction crises, and episodes of state fiscal collapse, and that these patterns exhibit statistical regularities across diverse societies and periods. The question is no longer whether quantitative historical analysis is possible, but under what conditions it is valid and what its inherent limits are.

1.2 The need for axiomatic foundations

Cliodynamics has produced substantial results, among them the identification of secular cycles of the order of 200 to 300 years in agrarian societies, the development of structural-demographic theory linking population dynamics to political instability, and a forecast of rising political instability in the United States (Turchin, 2010) that the events of 2020 and 2021 broadly bore out, and a structural-demographic analysis of American history demonstrating the framework's applicability to modern industrial societies (Turchin, 2016). The field has nonetheless operated without an explicit axiomatic foundation specifying the scope conditions under which quantitative historical analysis is valid, the epistemological limits of its predictive capability, the criteria by which models are to be evaluated and falsified, and the relationship between individual agency and aggregate dynamics. Without such foundations the field risks both overclaiming, by making predictions beyond its valid domain, and underclaiming, by failing to recognise its genuine achievements. An axiomatic framework addresses both risks by rendering its assumptions explicit and testable.

1.3 Purpose and scope of this document

This paper proposes eight foundational axioms for Computational Macrohistory, defined as a quantitative science that studies the temporal evolution of large-scale complex social systems through dynamic, probabilistic, and computational mathematical models. The axioms serve three functions at once. They are definitional, specifying what CMH is and what falls within its domain; prescriptive, establishing the standards that valid CMH models and predictions must meet; and delimiting, making explicit what CMH cannot do.

The present version extends the foundational treatment with four formal theorems that derive operational consequences directly from the axioms. Theorem 1 (Predictive Decay) formalises the logical relationship between continuous dynamics, endogenous indeterminacy, and accuracy decay in the stochastic-flow framework of Arnold (1998) and Khasminskii (2011). Theorem 2 (Aggregative Convergence) translates the qualitative aggregation condition into a quantitative threshold incorporating intra-class correlation and effective sample size. Theorem 3 (Reflexive Fixed Point) identifies the mathematical conditions under which predictions can be self-consistent rather than self-defeating, using the Wasserstein-1 metric. Theorem 4 (Calibration Decomposition) anchors the falsifiability requirement to the standard probabilistic scoring framework. Together the theorems demonstrate that the axioms are not merely a classification scheme but a functional structure that produces testable mathematical consequences.

The contribution of CMH is therefore not the introduction of new mathematical techniques, but the integration of established results from dynamical systems, probability theory, fixed-point theory, and forecast verification into a single coherent axiomatic framework. Within that framework, four fundamental limitations of quantitative historical prediction emerge as formal mathematical consequences rather than independent assumptions.

This is the first paper in a series, and it sets out the axiomatic foundations. A second develops the operational framework, comprising the state space, the variables, and the evolution equations. Subsequent papers apply the framework to empirical cases.

2 Methodological premise

2.1 What axioms do and do not do

The axioms presented here do not describe the world as it is. They are not empirical claims about social reality that could be directly tested and potentially falsified. They define the conditions under which historical systems become scientifically tractable within the CMH framework. This follows the tradition of axiomatic foundations in mathematics and physics. Euclidean geometry does not claim that physical space is Euclidean; it defines the rules under which Euclidean theorems hold. Newtonian mechanics does not claim that $F = ma$ is universally applicable; it defines a framework valid at certain scales and velocities. In the same way, the CMH axioms do not claim that all historical processes are predictable; they define when and how prediction is meaningful.

The practical implication is twofold. When the axioms are satisfied, CMH methods apply and predictions are meaningful within stated uncertainty bounds. When any axiom is violated, the system falls outside the framework's domain of validity, and CMH predictions should not be trusted.

A notational convention applies throughout. All stochastic differential equations in this paper are interpreted in the Itô sense, in accordance with the CMH notation standard; the informal forcing term $\varepsilon(t)$ of the master equation denotes the corresponding Itô stochastic component, written explicitly as $\sigma\, dW$ wherever the diffusion structure is used.

2.2 Admissibility axioms versus logical primitives

A clarification is needed regarding the logical status of the axioms. In formal mathematics, axioms are typically primitive propositions from which theorems are derived, statements that cannot themselves be proven within the system. The CMH axioms function differently. Some of them, particularly A5 and A7, encode necessary consequences of broader mathematical classes, specifically properties of nonlinear dynamical systems that are well established in mathematics and physics. A reviewer familiar with dynamical systems theory will observe that predictive decay (A7) follows from sensitive dependence on initial conditions (A5) combined with finite measurement resolution. The observation is correct, and the present version makes the derivation explicit in Theorem 1. Making the derivation explicit does not diminish the role of the axioms: it demonstrates how they connect, and thereby sharpens the conditions under which each one is operative.

The CMH axioms are best understood as admissibility constraints: mandatory conditions that any model claiming to be a valid CMH model must satisfy. They function as a filter that excludes models claiming deterministic prediction of chaotic systems, in violation of A5; asserting unlimited predictive horizons, in violation of A7; ignoring reflexive feedback, in violation of A6; or resisting empirical testing, in violation of A8. A model might be internally consistent and mathematically sophisticated while still failing to qualify as CMH if it violates these constraints. The axioms are therefore prescriptive rather than merely descriptive: they establish standards, not just descriptions. This approach has precedent in applied mathematics and engineering, where admissibility conditions routinely constrain the class of acceptable solutions, as with stability conditions in control theory or realizability conditions in signal processing.

2.3 Conditions for scientific tractability

For a domain of phenomena to be scientifically tractable, several conditions must hold. The phenomena must exhibit patterns rather than pure randomness; the relevant variables must be observable and measurable; distinct instances must be comparable within some common framework; and future states must depend on past and present states in ways that can be modelled. The eight CMH axioms operationalise these general requirements for the specific domain of large-scale historical processes. They specify the scale of analysis that is appropriate (A1), the conditions under which historical comparison is valid (A2), the kind of explanation that is possible (A3), the mathematical tools that apply (A4), the limits that prediction faces (A5, A6, A7), and the standards that predictions must meet (A8).

2.4 The role of formalization

Several considerations motivate stating these assumptions as axioms rather than leaving them implicit. Explicit axioms can be examined, debated, and refined, whereas implicit assumptions resist scrutiny. They generate testable expectations about when models should and should not work: a model that fails where the axioms hold is wrong, while one that fails where they are violated has simply been applied outside its domain, and the failure is then expected and informative. Explicit foundations also support cumulative progress, since shared standards allow researchers to build on one another's work with a clear understanding of what is being assumed. Finally, stating the limits openly guards against overclaiming, for a researcher who acknowledges that predictions decay beyond certain horizons (A7) is more credible than one who claims unlimited foresight.

3 The eight foundational axioms

3.1 Axiom A1: Statistical aggregation

Statement. There exists a critical demographic scale threshold N_c such that, for any social system with population $N \geq N_c$, aggregate collective behaviour is describable as a regular statistical process.

Formalization. Let $X_N(t)$ be an aggregate variable defined over a population of size N. Under independent or weakly correlated individual contributions, the variance of the aggregate mean satisfies

$$\lim_{N\to\infty} \mathrm{Var}[\bar{X}_N(t)] = 0$$

for all finite t and for all classes of individual micro-fluctuations with bounded variance and intra-class correlation ρ_N decaying at least as $1/N$. When ρ does not decay (as in tightly coupled social networks, information cascades, or synchronised collective behaviour), Theorem 2 shows that the variance converges to a non-zero floor $\sigma^2\rho$ set by the structural correlation, and A1's guarantee is limited to the precision achievable within that floor rather than to arbitrary precision. The axiom thus identifies the regime in which aggregation yields statistical regularity (sufficiently large N with bounded dependence) and the regime in which collective correlation imposes a hard precision limit regardless of scale.

When populations are sufficiently large, individual idiosyncrasies average out. The law of large numbers ensures that aggregate behaviour becomes statistically predictable even when individual behaviour is not, in the same way that thermodynamics describes the behaviour of a gas without tracking individual molecules, or epidemiology predicts the spread of disease without modelling each person's contacts. The key point is that unpredictability at the micro level is compatible with predictability at the macro level. We cannot predict whether a specific individual will join a protest, but we can estimate the proportion of a population likely to protest under given structural conditions. Revolutionary participation illustrates the principle clearly: individual decisions depend on personal beliefs, social networks, risk tolerance, and countless idiosyncratic factors, yet aggregate participation rates correlate strongly with structural variables such as youth unemployment, inequality, and regime legitimacy. In populations of millions, the statistical relationship emerges clearly despite individual unpredictability.

Unlike the earlier formulation, which treated N_c as a qualitative empirical hyperparameter, the present version supplies the explicit probabilistic bound derived in Theorem 2. The threshold admits a quantitative estimate incorporating intra-class correlation ρ: for desired precision ϵ and confidence level $1-\alpha$, $N_c = \sigma^2(1-\rho)/(\epsilon^2\alpha - \sigma^2\rho)$, which reduces to the classical Chebyshev threshold $\sigma^2/(\epsilon^2\alpha)$ under independence ($\rho = 0$) and diverges as correlation approaches the Herding Threshold $\rho^* = \epsilon^2\alpha/\sigma^2$. A further consequence of Theorem 2 is that even for arbitrarily large N, structural correlation $\rho > 0$ imposes a variance floor $\sigma^2\rho$ on achievable precision: A1 guarantees statistical regularity, not arbitrarily high precision, and the attainable accuracy is bounded by the degree of collective coupling. In practical terms, CMH applies most reliably to systems with large populations and low intra-class behavioural correlation. If $N < N_c$ or $\rho \geq \rho^*$ for the variable of interest, the aggregation conditions are not met and the model is being applied outside its valid domain.

3.2 Axiom A2: Conditioned historical ergodicity

Statement. Given a set of stable structural constraints, the historical trajectories of comparable systems belong to the same statistical class.

Formalization. There exists a distribution $P(X)$ such that, for a class of systems $\mathcal{S}$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(X(t))\,dt = \mathbb{E}[f(X)|\theta]$$

for every macro-observable f, conditional on the stability of the structural parameters.

History does not repeat exactly, but it rhymes. Systems facing similar structural conditions, with comparable demographics, economic arrangements, political institutions, and technological constraints, tend to follow similar statistical patterns. This is not because history is deterministic but because similar causes produce similar probability distributions of effects. The qualifier "conditioned" carries the weight of the axiom: ergodicity holds only while the structural parameters remain stable, and fundamental structural breaks, such as the transition from agrarian to industrial economies or from the pre-nuclear to the nuclear age, can invalidate comparisons across the break.

Pre-revolutionary France in the 1780s and pre-revolutionary Russia in the 1910s differed enormously in culture, geography, and specific circumstance. Both nonetheless exhibited the structural-demographic syndrome identified by Goldstone (1991) and Turchin (2003): fiscal crisis, elite overproduction, popular immiseration, and declining state capacity. The structural similarity made them statistically comparable, and both produced revolutionary outcomes, not by coincidence but through the operation of similar causal mechanisms under similar conditions.

The axiom grounds several methodological commitments. The past informs the future under comparable structural conditions, systematic comparison across time and space is legitimate, and retrospective validation against historical data is a valid form of testing. Historical analogies acquire a statistical rather than merely narrative foundation. At the same time, changes in fundamental parameters must be monitored, because when they shift the historical baseline may cease to apply. If structural parameters change fundamentally, historical comparisons become invalid, and a model trained on agrarian societies cannot predict industrial dynamics without recalibration. More radically, if every historical situation were genuinely unique, with no comparable cases, historical science would be impossible. The empirical success of cliodynamics suggests that this is not the case.

3.3 Axiom A3: Structural causality

Statement. Macro-historical events are not primitive entities but emerge from causal relationships among observable structural variables.

Formalization. Every critical event E satisfies

$$P(E \mid X) \neq P(E)$$

for at least one structural variable X_i belonging to the state space. In the language of structural causal models (Pearl, 2009), this dependence is represented by assuming the existence of a directed acyclic graph $\mathcal{G}$, in which macro-events are downstream nodes with identifiable parent variables.

Historical events such as revolutions, wars, and state collapses do not occur at random. They have causes, measurable preconditions that alter their probability. We cannot predict the specific triggering event, the assassination of an archduke or a fruit vendor's self-immolation, but we can identify the structural conditions that make certain classes of event more or less likely. The axiom rejects two extreme positions. One is pure contingency, the view that history is merely one thing after another; the other is strict determinism, the view that outcomes are uniquely fixed once initial conditions are given. CMH adopts an intermediate position of probabilistic structural causation, in which structures shape the probability distribution of events without uniquely determining any specific outcome.

The French Revolution was not a random occurrence. Fiscal bankruptcy following expensive wars, a proliferating nobility competing for limited positions, poor harvests depressing popular welfare, and Enlightenment ideas delegitimising traditional authority together created conditions in which revolution became probable. The specific spark, the calling of the Estates-General or the storming of the Bastille, was contingent; the underlying structural vulnerability was not. A counterfactual France without fiscal crisis, without noble overproduction, and without popular immiseration would have faced a much lower probability of revolution regardless of any particular trigger.

The axiom licenses the use of directed acyclic graphs and structural equation models, directs research toward identifying which structural variables affect which outcomes, and brings the do-calculus (Pearl, 2009) to bear on questions about interventions. Confounders must be identified and controlled, and causal pathways, not merely correlations, should be specified. Were macro-events truly random, uncorrelated with any observable structural variable, prediction would be impossible in principle. The axiom asserts, and the empirical evidence supports, that this is not so: major historical events have structural preconditions that can in principle be measured and modelled.

3.4 Axiom A4: Continuous temporal dynamics

Statement. The evolution of macro-historical systems can be approximated as a continuous dynamic process in time.

Formalization. There exists a state vector $X(t) \in \mathbb{R}^n$ such that

$$\frac{dX}{dt} = F(X, \theta, \varepsilon(t))$$

where F is a locally Lipschitz function, θ represents the parameters, and $\varepsilon(t)$ is a stochastic process representing shocks.

Social change, although it sometimes appears sudden, emerges from continuous underlying processes. Populations grow gradually, economic conditions shift incrementally, political tensions accumulate over time, and legitimacy erodes slowly before collapsing rapidly. Apparent discontinuities such as revolutions, regime collapses, and market crashes are often phase transitions in continuous systems, tipping points at which gradual change produces sudden transformation. The mathematics of continuous dynamics, including bifurcation theory (Nicolis & Prigogine, 1977) and catastrophe theory, can model such transitions.

The collapse of the Soviet Union between 1989 and 1991 appeared sudden, but the underlying pressures had accumulated over decades: economic stagnation evident in declining growth rates from the 1970s, the erosion of ideological legitimacy, elite fragmentation that glasnost allowed to surface, and fiscal strain from military spending and falling oil revenues. The continuous evolution of these variables produced a discontinuous outcome once the system crossed a critical threshold. The suddenness was real but emerged from continuous processes.

The requirement that F be locally Lipschitz is the formal content of the axiom: it ensures existence and uniqueness of solutions for the deterministic component, prevents instantaneous discontinuities in the drift, and bounds the local rate of change. Without this condition, the system could undergo genuine mathematical discontinuities, state jumps without passage through intermediate states, which would invalidate the continuous approximation on which the entire dynamical apparatus rests. Because the dynamics are continuous, ordinary differential equations are legitimate modelling instruments, stability can be examined through Jacobian matrices and eigenvalue analysis, critical transitions can be detected mathematically, discrete events such as wars and pandemics can be represented as impulses within a continuous system, and standard numerical solvers apply. If social dynamics were instead fundamentally discrete at the macro level, rather than merely appearing so at phase transitions, differential equation models would be inappropriate.

3.5 Axiom A5: Endogenous indeterminacy

Statement. Trajectory-level prediction of macro-historical systems is intrinsically probabilistic even under a perfect model.

Formalization. The uncertainty admits two canonical mathematical forms, either of which satisfies the axiom:

(i) *Deterministic chaos.* There exist systems with maximum Lyapunov exponent $\lambda_{\max} > 0$, implying $|\delta X(t)| \sim |\delta X(0)|\, e^{\lambda_{\max} t}$ and exponential divergence of nearby trajectories.

(ii) *Noise-induced indeterminacy.* In multistable systems subject to finite-variance stochastic forcing, the probability of barrier crossing (switching from one basin of attraction to another under the influence of noise) is governed by the Kramers escape rate $r \sim e^{-\Delta V/\sigma^2}$, where ΔV is the potential barrier height and σ^2 the noise intensity. The crossing time is exponentially distributed with mean $1/r$, making the exact moment of transition intrinsically unpredictable despite deterministic dynamics within each basin.

In both regimes, point predictions of the system's exact state at a future time are impossible in principle, and predictions must take the form of probability distributions.

Social systems exhibit both forms of indeterminacy. The atmospheric analogy captures regime (i): even with perfect knowledge of the governing equations, sensitive dependence on initial conditions limits weather prediction to roughly ten to fourteen days. Regime (ii) is arguably the more prevalent form in macropolitical dynamics: a country may drift within a stable basin for decades, accumulating structural stress, and then cross a tipping point whose precise timing is governed by the stochastic amplification of small fluctuations near the barrier: the self-immolation of a street vendor, a disputed election result, a sudden capital flight. The structural vulnerability is measurable (the SSI of WP-2026-003 and WP-2026-004 estimates the height of the barrier), but the spark is not.

The axiom requires that predictions take the form of probability distributions rather than point forecasts, that ensemble methods such as Monte Carlo simulation with parameter uncertainty be used, and that confidence or credible intervals always accompany a prediction. Predictions must acknowledge their temporal limits, and overconfident forecasts indicate methodological failure rather than scientific achievement. Uncertainty here is structural, not merely epistemic: better data and better models can sharpen probability estimates but cannot eliminate the underlying indeterminacy of which regime the system will occupy at a future time.

3.6 Axiom A6: Bounded reflexivity

Statement. Knowledge of predictions modifies the observed system, but this effect is modelable as bounded endogenous feedback.

Formalization. Let $P(E \mid X)$ be a published prediction. Then

$$X(t^+) = X(t^-) + G(P(E))$$

where G is a bounded, non-arbitrary feedback function.

Social predictions can alter the reality they describe. This is the phenomenon of self-fulfilling and self-defeating prophecy. A credible prediction of bank failure can cause the depositor panic that produces the failure, while a credible prediction of a pandemic can trigger the preparation that prevents or mitigates it. This reflexivity distinguishes social science from natural science: planets do not change their orbits because astronomers predict eclipses, but people may change their behaviour because economists predict a recession. The central claim of A6 is that this reflexivity is bounded rather than unlimited. The feedback effects are themselves subject to structural constraints, so that not every prediction can produce any arbitrary outcome, and reflexive dynamics can in principle be modelled within the system rather than treated as an external invalidation.

Consider a prediction that a given country has a seventy per cent probability of revolution within five years. If the regime believes it, it might intensify repression, which could accelerate revolution, or implement reforms, which could forestall it. If opposition groups believe it, they might mobilise more aggressively or, expecting an inevitable victory, wait. If foreign actors believe it, they might withdraw investment or intervene to support the regime. These effects are real but not arbitrary: they depend on existing power structures, resource constraints, and behavioural patterns, all of which can be modelled. The prediction does not create unlimited optionality; it perturbs a structured system in structured ways.

The practical response is to seek predictions that are stable under reflexive feedback. Theorem 3 formalises the condition: when the reflexive feedback function G is a contraction under the Wasserstein-1 metric, a unique self-consistent prediction P^* exists and is the attractor of iterative belief dynamics. The contraction condition provides an operational diagnostic: if the Lipschitz constant κ of G is estimated or bounded above by one, the prediction is reflexively stable; if $\kappa \geq 1$, the system is reflexively unstable and the prediction domain is compromised. The theorem thus transforms A6 from a qualitative caution into a quantitative condition. Were reflexivity unbounded, so that any prediction could produce any outcome through belief alone, social prediction would be meaningless. The axiom asserts that reflexive effects, while real and important, operate within modelable constraints, and the evidence supports this, since many social predictions, including most economic forecasts and demographic projections, do not dramatically alter the outcomes they describe.

3.7 Axiom A7: Predictive decay

Statement. Predictive accuracy decreases monotonically with the temporal horizon for trajectory-level forecasts.

Formalization. There exists $\lambda > 0$ and a baseline accuracy $\mathcal{A}_\infty \geq 0$ such that, for trajectory-level predictions of fixed resolution,

$$\mathcal{A}(\tau) = \mathcal{A}_\infty + (\mathcal{A}_0 - \mathcal{A}_\infty)\, e^{-\lambda\tau}$$

where $\mathcal{A}_0$ is the initial accuracy and $\mathcal{A}_\infty$ is the irreducible baseline set by the climatological or ergodic uncertainty. The exponential form governs the trajectory-level prediction class (event prediction); regime-probability forecasts decay with a substantially smaller effective λ, and spectral-structure predictions (invariant modes of the dynamics) are not subject to this decay, a hierarchy formalised in WP-2026-005 and tested empirically in WP-2026-006. This axiom follows from A5, combined with finite measurement resolution and bounded computational resources. It is stated separately because of its practical importance: making it explicit ensures that every CMH-admissible model incorporates appropriate humility about temporal horizons. Theorem 1 derives the exponential form in the linear divergence regime and identifies the transition to saturation at the attractor diameter.

Near-term predictions are more reliable than long-term ones. This follows mathematically from endogenous indeterminacy but deserves explicit axiomatic status because of its methodological weight. The decay rate λ varies with the system, since some systems are more chaotic than others; with the variable, since some variables are more stable than others; and with the resolution, since coarse predictions decay more slowly than precise ones. The monotonic decrease itself is universal for complex systems, and no method achieves stable accuracy over arbitrary horizons. The approximate relationship between horizon and achievable accuracy is summarised in Table 1.

Table 1: Illustrative relationship between prediction horizon and achievable accuracy.

Horizon	**Type of prediction**	**Typical accuracy**	**Character**
1–2 years	Binary events (revolution yes or no)	70–80%	Quantitative
3–5 years	Probabilistic scenarios	60–70%	Semi-quantitative
5–10 years	Directional trends	50–60%	Qualitative
10–20 years	Cyclical phase	40–50%	Very qualitative
Over 20 years	General tendencies	Below 40%	Scenario planning only

These figures are illustrative, but the pattern is consistent across domains. In practical terms, every prediction must state its valid temporal range, error bars must widen with the horizon, long-term analysis should rely on scenarios rather than point predictions, and models should be tested on the horizons where meaningful accuracy is possible. A claim of predictive accuracy that does not decay falls outside CMH, since to predict political dynamics in 2050 with the same confidence as in 2027 is either to trade on unfalsifiable vagueness or to make claims that complex-systems science cannot support. The decay rate can be estimated empirically, but Theorem 1 proves that it cannot be zero.

3.8 Axiom A8: Computational falsifiability

Statement. Every CMH model must generate predictive distributions quantitatively comparable with observed data.

Formalization. Given a model M, the predictive distribution

$$P_M(X(t + \Delta t) \mid X(t))$$

must be testable through probabilistic scoring metrics. This axiom is methodological rather than ontological: it prescribes how CMH science should be conducted rather than describing a property of social systems. It belongs among the axioms because adherence to falsifiability is a necessary condition for admissibility, and a model that resists empirical testing, however sophisticated, falls outside the framework.

CMH is a science, not a narrative or a philosophical position, and to be scientific its models must make predictions that can fail. A model that cannot be wrong is not scientific. The difficulty in social science is that predictions are probabilistic, so that a model predicting a sixty per cent probability of instability is not falsified by a single case of stability or instability. Falsifiability is therefore assessed through calibration: across many predictions, the stated probabilities should match the observed frequencies. This preserves the Popperian criterion of falsifiability while accommodating the inherently probabilistic character of social prediction. A statement that the model assigns sixty per cent probability to significant instability in a given country within five years is falsifiable, because across many such predictions roughly sixty per cent should materialise, and a model that consistently assigns sixty per cent to events occurring thirty per cent of the time is poorly calibrated and thereby refuted. A statement that instability is merely possible under certain conditions cannot fail, since instability is always possible, and is therefore not a scientific prediction.

Theorem 4 formalises the falsifiability apparatus by anchoring it to the Murphy (1973) decomposition of the Brier score. The Brier score of a probabilistic forecast decomposes into three additive terms:

$$\mathrm{BS} = \mathrm{REL} - \mathrm{RES} + \mathrm{UNC}$$

where REL (reliability) measures calibration, RES (resolution) measures the forecast's ability to discriminate among outcomes, and UNC (uncertainty) is the irreducible variance of the outcome variable. This decomposition, standard in meteorological forecast verification since the 1970s, provides CMH with a complete formal apparatus for evaluating A8 compliance: a well-calibrated model minimises REL, a discriminating model maximises RES, and the irreducible baseline UNC sets the ceiling on achievable skill. The axiom carries concrete methodological obligations. Quantitative metrics such as the Brier score, log-likelihood, and AUC-ROC must be reported; predicted probabilities must be checked against observed frequencies; standard backtesting procedures must be followed; failed predictions must be reported alongside successes; code and data must be available for replication; and predictions should ideally be registered before outcomes are known. Models that cannot be tested against data are not CMH models. They may be useful heuristics, interesting narratives, or stimulating thought experiments, but they lack scientific status within the framework. The axiom functions as a demarcation criterion, marking the boundary between science and non-science in historical analysis.

4 Axiomatic closure

4.1 Non-redundancy and functional distinctness

The eight axioms are conceptually non-redundant and functionally distinct: each addresses an aspect of the tractability problem that the others do not fully cover. Table 2 records the question that each one answers.

Table 2: The domain and guiding question of each axiom.

Axiom	Domain	Question addressed
A1	Scale	How large must systems be for statistical analysis?
A2	Comparison	When are historical cases comparable?
A3	Explanation	What kind of causation operates?
A4	Mathematics	What formal tools are appropriate?
A5	Limits	Why is prediction inherently uncertain?
A6	Reflexivity	How do predictions affect outcomes?
A7	Horizons	How far ahead can we predict?
A8	Science	What makes predictions scientific?

Some logical dependencies do exist among the axioms. Predictive decay (A7) follows from indeterminacy (A5) combined with finite resolution, now formalised in Theorem 1, and falsifiability (A8) is methodological where A1 through A7 are primarily ontological or epistemological. These dependencies do not make any axiom redundant in practice. Each highlights a distinct constraint that practitioners must respect, and removing any one would leave an important requirement implicit rather than explicit. The aim of the set is practical guidance for the construction and evaluation of models, not a minimal logical basis in the formal sense.

4.2 Sufficiency of the axiom set

The eight axioms are jointly sufficient to define the foundations of CMH. Together they specify the domain of applicability (A1, A2), the type of models permitted (A3, A4), the nature of the predictions generated (A5, A6, A7), and the standards for scientific validity (A8). No further axioms appear necessary to establish these foundations. Subsequent papers build operational frameworks, with state spaces, equations, and variables, on this base, but the base itself is complete for its intended purpose. To show that the axioms are jointly satisfiable, that a model meeting all eight constraints can exist, Appendix C develops a bistable model with multiplicative noise that satisfies them under conditions richer and more realistic than those of the earlier minimal construction.

4.3 Logical dependencies formalized

This section presents four theorems that make explicit the logical relationships among the axioms. The derivations highlight the essential structure; technical conditions (integrability, domain regularity, measure-theoretic completeness) are standard in the cited literature and are noted in the assumptions blocks where they bear on applicability. The theorems demonstrate that the axiom set is not a flat list of desiderata but a structure in which certain axioms follow mathematically from others, and in which specific operational criteria can be derived.

Theorem 1 (Predictive Decay)**. Assumptions.** Let the system satisfy Axiom A4, so that the evolution of the state vector $X \in \mathbb{R}^n$ is governed by the stochastic differential equation

$$\frac{dX}{dt} = F(X, \theta, \varepsilon(t))$$

with F locally Lipschitz in X and satisfying a linear growth condition $\|F(X)\| \leq C(1 + \|X\|)$, which guarantees non-explosion of solutions on finite horizons. When the system is cast in Itô SDE form $dX = F(X, \theta)\,dt + \sigma(X, t)\,dW(t)$ (the CMH Form 3 of the notation standard), the diffusion coefficient σ is assumed uniformly Lipschitz and non-degenerate on compacta, ensuring that the associated stochastic flow of diffeomorphisms is well-defined and admits a top Lyapunov exponent in the sense of Arnold (1998) and Khasminskii (2011). Let the system further satisfy Axiom A5, so that the maximum Lyapunov exponent of the stochastic flow, defined as

$$\lambda_{\max} = \lim_{t \to \infty} \frac{1}{t} \ln \|\Phi(t, \omega)\|$$

where $\Phi(t, \omega)$ is the linearisation (variational equation) of the stochastic flow along a trajectory, satisfies $\lambda_{\max} > 0$. Let $\hat{X}(t + \tau)$ be any prediction at horizon τ based on information available at time t, subject to initial measurement error $\|\delta X(t)\| \leq \delta_0$.

Statement. Then for horizons τ within the linear divergence regime ($\tau \lesssim \tau_{\text{sat}}$, where $\tau_{\text{sat}} \approx \lambda_{\max}^{-1} \ln(K/\delta_0)$ for a system bounded in a domain of diameter K),

$$\mathbb{E}\|X(t + \tau) - \hat{X}(t + \tau)\| \geq c\,\delta_0\, e^{\lambda_{\max}\tau}$$

provided that $\lambda_{\max}$ is simple and isolated. At saturation ($\tau \gtrsim \tau_{\text{sat}}$), the Euclidean error is bounded by the attractor diameter and the predictive accuracy, defined as the probability of assigning the system to the correct basin, decays monotonically toward the climatological baseline, establishing the saturating form of A7, $\mathcal{A}(\tau) = \mathcal{A}_\infty + (\mathcal{A}_0 - \mathcal{A}_\infty)e^{-\lambda\tau}$, with $\lambda \approx \lambda_{\max}$ in the exponential window. At saturation ($\tau \gtrsim \tau_{\text{sat}}$), the Euclidean error is bounded by the attractor diameter and the predictive accuracy, defined as the probability of assigning the system to the correct basin, decays monotonically toward the climatological baseline, establishing the saturating form of A7, $\mathcal{A}(\tau) = \mathcal{A}_\infty + (\mathcal{A}_0 - \mathcal{A}_\infty)e^{-\lambda\tau}$, with $\lambda \approx \lambda_{\max}$ in the exponential window.

Derivation. Under the stated regularity conditions, the stochastic flow admits a multiplicative ergodic theorem in the sense of Arnold (1998, Ch. 3–4): the tangent linear propagator $\Phi(t + \tau, t) = \partial X(t + \tau)/\partial X(t)$, obtained from the variational SDE associated with the original system, decomposes into Lyapunov subspaces with well-defined exponential growth rates $\lambda_1 \geq \lambda_2 \geq \ldots$. Let $v_{\max}$ be the unit vector in the Oseledets subspace corresponding to $\lambda_{\max}$. Under the simplicity and isolation hypotheses, for sufficiently small δ_0 the projection of the initial error onto $v_{\max}$ is non-zero with probability one, retaining a fraction $c \in (0, 1)$ of the initial magnitude δ_0; this constant c is the prefactor in the lower bound stated above. The linearised flow along $v_{\max}$ satisfies $\|\delta X(t + \tau)\| \approx \|\delta X(t)\|\, e^{\lambda_{\max}\tau}$ for τ within the linear regime, defined as the horizon over which $\|\delta X\| \leq \eta$ where η depends on the Lipschitz constant of F.

Remark on scope and on the cubic drift. This theorem covers the deterministic-chaos regime of A5, clause (i): it requires a strictly positive top Lyapunov exponent of the stochastic flow. The bistable witness model of Appendix C does not fall under this theorem, since its deterministic skeleton is multistable rather than chaotic and its top Lyapunov exponent is not positive; that model satisfies A5 through clause (ii), noise-induced barrier crossing, and its predictive decay is derived separately, from the Kramers escape rate, in Appendix C.5. A technical caution applies to any super-linear drift such as the cubic $f(X) = \alpha X(1 - X/K)(X - X_c) + \beta$: global existence of solutions and finiteness of Lyapunov exponents are not automatic, so the analysis is localised to a compact domain $\mathcal{D}$ containing the attractor, with a stopping time $\tau_{\mathcal{D}}$ at first exit. For the empirically meaningful forecast windows $\tau < \tau_{\mathcal{D}}$ the linear-growth bound holds on $\mathcal{D}$ and the SDE is well-posed; escape to infinity occurs at most on a measure-zero set and does not affect the finite-horizon bound.

No prediction algorithm can circumvent this divergence, because it is a property of the system, not of the estimator. The lower bound holds in the linear regime, for $\tau \lesssim \tau_{\text{sat}}$, where the error has not yet reached the attractor diameter. Beyond saturation the Euclidean error is bounded by the attractor diameter and ceases to grow; what continues to decay is probabilistic skill, the accuracy of assigning the system to the correct basin, which relaxes monotonically toward the climatological baseline $\mathcal{A}_\infty$. Assuming that predictive accuracy is a monotone decreasing function of trajectory error, the exponential divergence of nearby trajectories induces an exponential decay of forecast skill throughout the linear regime. Combining the exponential growth of the linear regime with the bounded post-saturation regime yields the saturating accuracy law $\mathcal{A}(\tau) = \mathcal{A}_\infty + (\mathcal{A}_0 - \mathcal{A}_\infty)e^{-\lambda\tau}$, with $\lambda \approx \lambda_{\max}$ inside the exponential window. This establishes the saturating form of A7 as a consequence of A4 and A5, not an unbounded exponential decay to zero, which a confined system cannot exhibit. $\square$

The theorem clarifies why A7 is stated as an independent axiom despite being derivable from A4 and A5. The derivation depends on the simplicity of the dominant Lyapunov exponent, on the availability of linear regime approximations, on the stochastic-flow regularity conditions, and on the localisation to compact domains when the drift is super-linear. When any of these conditions fails (for instance, when $\lambda_{\max}$ is degenerate or the measurement error is large enough to exit the linear regime immediately) A7 still holds as a phenomenological constraint, but the exponential form may not be derivable from A4 and A5 alone. The axiom is a safety condition: even in regimes where the formal derivation does not carry through, the requirement of monotonic predictive decay is non-negotiable within CMH.

Theorem 2 (Aggregative Convergence)**. Assumptions.** Let X_i for $i = 1, \ldots, N$ be the individual contributions to an aggregate variable, with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$ for all i. Define the sample mean $\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$. The individuals are not assumed independent; let ρ denote the average intra-class (within-group) correlation, $\rho = \mathrm{Cov}(X_i, X_j)/\sigma^2$ for $i \neq j$, satisfying $\rho \geq 0$ (non-negative dependence is the empirically relevant case in social systems, where imitation, network effects, and common shocks produce positive co-movement).

Statement. Then for any $\epsilon > 0$,

$$P(|\bar{X}_N - \mu| > \epsilon) \leq \frac{\sigma^2[1 + (N-1)\rho]}{N\epsilon^2}$$

and consequently, for a desired precision ϵ and confidence level $1 - \alpha$, the critical population threshold of Axiom A1 satisfies

$$N_c = \frac{\sigma^2}{\epsilon^2\alpha} \cdot [1 + (N_c - 1)\rho].$$

Solving the implicit equation for N_c yields the explicit form

$$N_c = \frac{\sigma^2(1-\rho)}{\epsilon^2\alpha - \sigma^2\rho}.$$

The condition $\epsilon^2\alpha - \sigma^2\rho > 0$, or equivalently $\rho < \epsilon^2\alpha/\sigma^2$, is the feasibility constraint: if correlation is too high relative to the desired precision, the denominator vanishes or becomes negative, meaning no finite sample size suffices, and A1's aggregation guarantee fails. The critical value $\rho^* = \epsilon^2\alpha/\sigma^2$ defines what we term the **CMH Herding Threshold**. When intra-class correlation exceeds ρ^*, the system enters a regime of structural herding (polarisation cascades, information monocultures, or synchronised collective behaviour) in which individual fluctuations no longer average out because macroscopic dependence is structural rather than incidental. In such a regime, increasing the population does not improve aggregate predictability; it may amplify collective swings. The Herding Threshold thus identifies a boundary beyond which CMH's aggregation axiom (A1) is inapplicable, regardless of demographic scale.

Derivation. The variance of the sample mean under arbitrary dependence is $\mathrm{Var}(\bar{X}_N) = \frac{1}{N^2}\sum_{i,j}\mathrm{Cov}(X_i, X_j) = \frac{\sigma^2}{N} + \frac{N-1}{N}\sigma^2\rho$. Substituting into the Chebyshev inequality $P(|\bar{X}_N - \mu| > \epsilon) \leq \mathrm{Var}(\bar{X}_N)/\epsilon^2$ gives the first bound. Setting the right-hand side equal to α and solving for N yields the implicit equation $N_c = \frac{\sigma^2}{\epsilon^2\alpha}[1 + (N_c - 1)\rho]$; the explicit form follows by collecting terms in N_c. In the limit $\rho \to 0$ (independent individuals), the formula reduces to the classical Chebyshev threshold $N_c = \sigma^2/(\epsilon^2\alpha)$. For $\rho > 0$, N_c is strictly larger, and as $\rho \to \epsilon^2\alpha/\sigma^2$ from below, $N_c \to \infty$: aggregation becomes impossible. $\square$

This result is stronger than the independence-only formulation and makes a substantive prediction. For a macropolitical index normalised to $\sigma = 0.5$, with a realistically achievable precision of $\epsilon = 0.2$ (20% of the variable's typical range is already ambitious for historical data at $\alpha = 0.05$), the independence threshold is $N_c = 125$. The feasibility constraint yields $\rho^* = \epsilon^2\alpha/\sigma^2 = 0.008$, meaning the CMH Herding Threshold sits at approximately 0.8% average intra-class correlation. At $\rho = 0.001$ (mild social imitation), N_c rises to approximately 143; at $\rho = 0.004$ (moderate network coupling), $N_c \approx 249$; and as $\rho \to 0.008$ from below, $N_c \to \infty$. These values are conservative, since the Chebyshev bound is the weakest concentration inequality and the true thresholds under sharper bounds (Hoeffding, Bernstein) would be larger, but the structural message is robust: any non-zero intra-class correlation substantially increases the required population size, and beyond a well-defined correlation ceiling, aggregation fails regardless of N.

A subtler insight emerges from the derivation. With $\rho > 0$ constant, $\mathrm{Var}(\bar{X}_N) \to \sigma^2\rho$ as $N \to \infty$ rather than to zero: structural correlation imposes a variance floor on achievable precision that no increase in population can breach. In this regime, A1 does not guarantee arbitrarily precise aggregation; it guarantees that the aggregate distribution stabilises at a well-defined limit whose precision is bounded below by the structural coupling. CMH thus applies where this floor is below the precision required by the analysis, a condition that can be checked empirically from estimated ρ. The effective sample size $N_{\text{eff}} = N/[1 + (N-1)\rho]$ formalises the trade-off: large N compensates for moderate ρ, but past the Herding Threshold no compensation is possible. Empirically, ρ can be estimated from survey data with cluster-sampling designs or from social network density measures.

Theorem 3 (Reflexive Fixed Point)**. Assumptions.** The reflexive feedback of Axiom A6 displaces the state in response to a published prediction; this displacement induces a map on the predictive distribution itself, since shifting the state changes the law of the future event. Theorem 3 operates on that induced distributional map, written $G : \mathcal{P}_1 \to \mathcal{P}_1$, which carries a published distribution P to the predictive distribution $G(P)$ that results once beliefs and behaviour have responded to P; in the conditional notation of A6 this is the operator $\Psi(P) = P(E \mid P)$ analysed below. Let G act on the space $\mathcal{P}_1$ of probability distributions with finite first moment, equipped with the Wasserstein-1 metric W_1 (Earth Mover's Distance). The restriction to $\mathcal{P}_1$ is required because W_1 is defined only on measures with finite expectation; G is assumed to preserve this property, i.e. $\int_\Omega |x|\, d[G(P)](x) < \infty$ for all $P \in \mathcal{P}_1$, excluding reflexive feedback so extreme that it generates infinite-mean or heavy-tailed distributions (a runaway panic whose behavioural response has undefined expectation). W_1 is chosen over the Total Variation distance because it is topologically weaker (contraction is easier to satisfy), metrises weak convergence, and is the natural geometry for belief dynamics in which small shifts in predicted probabilities produce proportionally small shifts in behavioural response, a physically meaningful condition absent in the discrete-jump regime that the Total Variation metric encodes. The space $(\mathcal{P}_1, W_1)$ is complete. Let $\kappa < 1$ be the Lipschitz constant of G under W_1, so that $W_1(G(P), G(Q)) \le \kappa\, W_1(P, Q)$ for all $P, Q \in \mathcal{P}_1$.

Statement. Define the prediction operator $\Psi : \mathcal{P}_1 \to \mathcal{P}_1$ by $\Psi(P) = P(E \mid P)$. Then Ψ has a unique fixed point $P^* \in \mathcal{P}_1$ satisfying $\Psi(P^*) = P^*$, and for any initial prediction P_0, the iterative sequence $P_{n+1} = \Psi(P_n)$ satisfies

$$W_1(P_n, P^*) \le \frac{\kappa^n}{1-\kappa}\, W_1(P_1, P_0).$$

If $\kappa \ge 1$, the iteration may fail to converge, and the system is reflexively unstable. More generally, if G is a local contraction, $\kappa < 1$ on some W_1-ball $B_\delta(P_0)$, and P_0 lies within its basin of attraction, the same conclusion holds locally; this relaxation accommodates the threshold effects (e.g. panic cascades) that make global contraction unrealistic in social systems.

Derivation. The prediction operator $\Psi : \mathcal{P}_1 \to \mathcal{P}_1$ defined by $\Psi(P) = P(E \mid P)$ is, under the global contraction hypothesis, a contraction with constant κ on the complete metric space $(\mathcal{P}_1, W_1)$. The Banach fixed-point theorem guarantees the existence, uniqueness, and convergence rate of the fixed point P^*. The explicit geometric bound follows from the standard estimate for contraction iterations. For the local version, one restricts the space to the W_1-ball $B_\delta(P_0)$ within which $\kappa < 1$ and verifies that Ψ maps $B_\delta(P_0)$ into itself; the Banach theorem then applies on this invariant subset. The instability claim for $\kappa \geq 1$ follows from the possibility of period-two cycles, divergence, or chaotic iteration in the absence of the contraction property; the Banach theorem provides no guarantee and counterexamples are readily constructed. □

Theorem 3 transforms A6 from a qualitative acknowledgment of reflexivity into a quantitative diagnostic. The contraction constant κ under W_1 can be bounded empirically: if a published prediction elicits a behavioural response that alters the measured state variables by an amount proportional to the prediction's shift from baseline, the proportionality constant provides an upper bound for κ. When $\kappa < 1$, the system admits a self-consistent prediction, a fixed point of belief dynamics, and the CMH analyst can report this prediction with the caveat that it assumes reflexive convergence. When $\kappa \geq 1$, predictions themselves become drivers of instability, and the CMH framework does not license point predictions, though it may still estimate the probability of entering such a regime.

Empirical identifiability. The contraction constant κ is not a free theoretical parameter. It can be estimated from observational data using a regression discontinuity design (RDD) or a natural experiment in which predictive information was released exogenously. For example, if an international organisation publishes a country risk rating at a known date, and the affected state variables (capital flows, policy changes, protest frequency) respond with a measurable discontinuity at that date, the magnitude of the response relative to the prediction shift estimates $\partial X/\partial P$, which bounds κ under mild smoothness assumptions. The existence of such estimable quantities ensures that Theorem 3 is not merely formal but falsifiable, in keeping with Axiom A8.

Theorem 4 (Calibration Decomposition)**.** For a probabilistic forecast f of a binary outcome $o \in \{0, 1\}$, with forecasts grouped into M bins and f_i the mean forecast probability in bin i, let $\bar{o}_i$ be the observed relative frequency of events in bin i, n_i the number of forecasts in bin i, and $\bar{o}$ the unconditional event frequency. Then the Brier score decomposes as

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{M} n_i(f_i - \bar{o}_i)^2 \; - \; \frac{1}{N}\sum_{i=1}^{M} n_i(\bar{o}_i - \bar{o})^2 \; + \; \bar{o}(1 - \bar{o})$$

where $N = \sum_i n_i$. The first term (REL) measures reliability, the second term enters with a negative sign and measures resolution (RES), and the third (UNC) is the irreducible uncertainty. This algebraic decomposition, due to Murphy (1973), clarifies why a discriminating model achieves a lower Brier score: larger RES directly subtracts from the total, penalising forecasts that fail to separate outcome classes. The decomposition establishes the formal apparatus for evaluating Axiom A8 compliance: a CMH-admissible model minimises REL, maximises RES, and reports both alongside the baseline UNC.

Derivation. The derivation is algebraic and standard in the forecast verification literature (Murphy, 1973; Wilks, 2011). Expand $\mathrm{BS} = \frac{1}{N}\sum_{k=1}^{N}(f_k - o_k)^2$ by adding and subtracting $\bar{o}_i$ for each forecast in bin i, then group by bin. The cross-term vanishes by construction of $\bar{o}_i$ as the bin-conditional mean. The result is the three-term decomposition stated. Each term has a direct interpretation: REL is zero for a perfectly calibrated forecast, RES is large for a forecast that separates outcome classes, and UNC is fixed by the data. □

Theorem 4 serves two purposes. First, it demonstrates that the falsifiability requirement of A8 inherits a complete, standardised validation apparatus from the probabilistic forecasting literature, anchoring CMH's methodological commitments to established practice in meteorology, finance, and machine learning. Second, it provides a concrete operational checklist: a CMH model's A8 compliance is assessed by reporting REL, RES, UNC, and the aggregate BS, alongside standard auxiliary metrics (AUC-ROC, log-likelihood). A model whose REL term dominates its RES term is well-calibrated but non-discriminating; a model with high RES but poor REL is overconfident. Only a model that is both calibrated and discriminating satisfies the full intent of A8.

4.4 Consequences of axiom violation

Each axiom, when violated, produces a characteristic failure mode, summarised in Table 3.

Table 3: Failure modes associated with the violation of each axiom.

Axiom	**Violation scenario**	**Consequence**	**Diagnostic**
A1	Population too small ($N < N_c$)	Statistical regularities absent; individual variance dominates	N_c bound from Theorem 2
A2	Structural parameters changed	Historical comparisons invalid; past does not inform future	Structural break test
A3	Events truly random	Prediction impossible in principle; no causal structure	Conditional independence test
A4	Discontinuous dynamics (F not Lipschitz)	Differential equations inappropriate; discrete models needed	Lipschitz constant estimation
A5	Ignored or denied	Overconfident predictions; claims exceed the valid domain	Lyapunov exponent estimation (chaotic regime) or Kramers escape-rate estimation (noise-induced regime)
A6	Unbounded reflexivity ($\kappa \geq 1$)	Predictions meaningless; any outcome possible	Contraction constant from Theorem 3
A7	Ignored or denied	Unjustified long-term claims; false precision	Exponential bound from Theorem 1
A8	Not implemented	Not science; unfalsifiable narrative	Brier score decomposition (Theorem 4)

The essential point is that axiom violation does not indicate model failure but scope violation. A model is not wrong when applied outside its valid domain; it is being misused. Recognising the limits of scope is part of responsible scientific practice. The diagnostic column, new in this version, supplies the specific quantitative test associated with each axiom via the four theorems.

5 Relationship to existing frameworks

5.1 Cliodynamics and structural-demographic theory

CMH stands in a relationship of continuity with cliodynamics, the quantitative study of historical dynamics initiated by Turchin (2003) and developed with collaborators over the following two decades. It draws on the substantive achievements of that programme while remaining a distinct framework with its own foundations.

Cliodynamics established that the mathematical modelling of historical dynamics is feasible, identified secular cycles of roughly two to three centuries in agrarian societies, advanced elite overproduction as a key driver of political instability, and showed that quantitative historical databases such as Seshat (Turchin et al., 2015; Turchin et al., 2018) can be used to test theoretical predictions. Structural-demographic theory (Goldstone, 1991; Turchin & Nefedov, 2009) supplied a framework linking population dynamics, elite competition, and state fiscal health to instability, together with empirically validated models of pre-modern state breakdown. CMH takes these insights as core theoretical content.

What CMH adds is foundational rather than substantive. It supplies the explicit axiomatic structure set out in this paper, a clear specification of scope conditions and limits, formalised standards for prediction and validation, an orientation toward computational implementation, and, in the present version, formal theorems that derive operational consequences from the axiomatic base. Much of what cliodynamic work assumes implicitly, that large-scale societies exhibit statistical regularities (A1), that historical comparison is valid (A2), and that structural variables carry causal weight (A3), is here made explicit and testable. The distinction matters. CMH is not a relabelling of cliodynamics, nor a sub-field within it, but a framework whose formal axiomatic structure, institutionalised falsifiability, and demonstrated logical closure give it a different standing, even as it builds directly on cliodynamic results.

5.2 Complexity science and nonlinear dynamics

CMH draws on complexity science for much of its mathematical toolkit and conceptual vocabulary. The relevant concepts include nonlinear dynamics, with its threshold effects, feedback loops, and emergent behaviour (A4); deterministic chaos, in which sensitive dependence on initial conditions limits predictability (A5); phase transitions, through which continuous processes generate discontinuous outcomes (A4); self-organisation, in which macro-patterns emerge from micro-interactions (A1); and the behaviour of dissipative structures far from equilibrium (Nicolis & Prigogine, 1977). The mathematical tools borrowed from this tradition include differential equations and dynamical systems theory, bifurcation and stability analysis, Lyapunov exponents and other measures of chaos, agent-based modelling for micro-foundations, and network theory for social structure. CMH differs from general complexity science in its specific focus on historical prediction and in its explicit treatment of reflexivity (A6), a feature largely absent from physical complex systems.

5.3 Structural causality and causal inference

The commitment to structural causality in A3 aligns with recent advances in causal inference, in particular the structural causal model framework of Pearl (2009) and collaborators. Directed acyclic graphs represent the causal structure among variables, the do-calculus provides formal tools for reasoning about interventions, counterfactual reasoning supports the analysis of "what if" questions, and systematic methods exist for identifying confounders. CMH extends these tools to historical analysis, where interventions are rarely possible and reliance falls on natural experiments, where confounding is pervasive because many variables interact, and where temporal precedence aids causal identification. Combining structural causal models with dynamical systems theory yields a framework for historical analysis that neither approach offers on its own.

5.4 Bayesian epistemology and probabilistic prediction

The probabilistic framing of CMH predictions aligns with Bayesian epistemology. Historical data informs prior probability estimates (A2), new evidence revises predictions through Bayes' rule, uncertainty is quantified by treating predictions as distributions rather than points (A5), and calibration requires that stated probabilities match observed frequencies (A8). CMH predictions are best understood as posterior distributions,

$$P(\text{Event} \mid \text{Data}, \text{Model}) = \frac{P(\text{Data} \mid \text{Event}, \text{Model}) \cdot P(\text{Event} \mid \text{Model})}{P(\text{Data} \mid \text{Model})}$$

a framing that incorporates prior knowledge from historical analysis, updates in light of new evidence, quantifies uncertainty explicitly, and supports principled model comparison.

5.5 Philosophy of social science

CMH engages several longstanding debates in the philosophy of social science. On the question of explanation versus understanding it takes a naturalistic position, holding that social phenomena can be explained through causal mechanisms and not only interpreted through the reconstruction of meaning; this does not deny the importance of meaning and intention but maintains that aggregate dynamics exhibit law-like regularities open to scientific analysis. On the question of methodological individualism versus holism it occupies a middle position: individual behaviour is not modelled directly (A1), but aggregate patterns are understood as emerging from individual interactions, and the aggregation process is itself a theoretical object. On prediction versus explanation it values both while distinguishing them, since good explanations need not yield good predictions when systems are chaotic, and good predictions need not supply satisfying explanations when mechanisms remain unclear; both are scientific achievements, and neither suffices alone. On value-neutrality it aspires to descriptive and predictive accuracy rather than normative judgment, able to analyse revolutions without endorsing or condemning them, while acknowledging that the choice of what to study and how to communicate findings carries ethical dimensions (Cartwright, 1999).

6 Conclusion and next steps

6.1 Summary of contributions

This paper has established eight foundational axioms for Computational Macrohistory, summarised in Table 4.

Table 4: The eight foundational axioms of Computational Macrohistory.

Axiom	Name	Core claim
A1	Statistical aggregation	Large populations enable statistical regularity
A2	Conditioned ergodicity	Comparable structures yield comparable dynamics
A3	Structural causality	Events have observable structural causes
A4	Continuous dynamics	Social change follows continuous processes
A5	Endogenous indeterminacy	Prediction is intrinsically probabilistic
A6	Bounded reflexivity	Reflexive effects are limited and modelable
A7	Predictive decay	Accuracy declines with temporal horizon
A8	Computational falsifiability	Models must generate testable predictions

Together these axioms define what CMH is, a quantitative science of large-scale social dynamics; when it applies, in the presence of large populations, stable structures, and observable variables; what it can do, namely probabilistic prediction within finite horizons; what it cannot do, including deterministic prediction, individual forecasting, and prediction over unlimited horizons; and how it qualifies as scientific, through falsifiable predictions and calibration metrics.

The present version advances beyond the earlier formulation in one central respect: the axioms are no longer merely stated but shown to produce formal consequences. Theorem 1 derives predictive decay from continuous dynamics and endogenous indeterminacy in the stochastic-flow framework. Theorem 2 quantifies the aggregation threshold incorporating intra-class correlation, yielding a criterion that is strictly more informative than the independence-only bound. Theorem 3 identifies the fixed-point conditions under which reflexive predictions are self-consistent and the instability condition where they are not, using the Wasserstein-1 metric. Theorem 4 anchors the falsifiability requirement to the standard probabilistic scoring apparatus. The four theorems together demonstrate that the CMH axiom set is not a flat classification scheme but a functional structure whose elements constrain one another and generate testable mathematical consequences.

6.2 The CMH research program

The axiomatic foundation supports a systematic research program. On the theoretical side, operational models can be derived from the axioms, state spaces and evolution equations specified, and domain-specific variants developed. The theorems proved here open specific avenues: the Lyapunov bound of Theorem 1 can be estimated empirically and compared across historical systems (acknowledging the practical challenges of finite, noisy historical records, which may require ensemble methods or synthetic data calibration as bridges); the aggregation threshold of Theorem 2 can be calibrated to specific variables and datasets, with intra-class correlation ρ estimated from survey or network data; the contraction diagnostic of Theorem 3 can be tested against historical cases where published predictions demonstrably altered the outcomes they described; and the Murphy decomposition of Theorem 4 provides the standardised reporting template for all CMH empirical work. On the empirical side, predictions can be tested against historical data, probability estimates calibrated, and scope conditions identified empirically. Methodologically, the program calls for standardised data protocols, benchmarking datasets, and established validation procedures. In application, it points toward early-warning systems for political instability, scenario analysis for policy planning, and risk assessment in international relations.

6.3 Transition to the operational framework

The axioms establish what CMH can legitimately do; they do not specify how to do it. Subsequent papers supply the operational content. The operational framework will specify the multi-dimensional state space, spanning demographic, economic, political, social, and psychological variables, together with the evolution equations linking them, the procedures for parameter estimation, and the simulation protocols. The methodological framework will address data sources and quality standards, the handling of missing data, model selection criteria, backtesting protocols, and calibration metrics. The case studies will demonstrate application to historical episodes, assess retrospective predictive performance, and draw lessons for prospective use.

6.4 Invitation to critique

These axioms are proposed, not proclaimed. Scientific progress depends on critical engagement, and readers are invited to test the set against several questions. Are any of the axioms fully derivable from the others, and can the set be reduced further? Are additional axioms required, and what questions remain unanswered? Are any of the axioms unnecessary for the goals of CMH? Do they correctly characterise the conditions for historical prediction, and what alternative axiomatisations might prove superior? The theorems proved here sharpen several of these questions: the derivation of A7 from A4 and A5 suggests that the axiom count might be reduced to seven under certain conditions, and the contraction condition of Theorem 3 may prove too restrictive for some applications, motivating a relaxation of A6. The purpose is not to fix dogma but to enable cumulative progress through explicit and debatable foundations. CMH will improve through critique rather than deference.

References

Arnold, L. (1998). *Random Dynamical Systems.* Springer.

Cartwright, N. (1999). *The Dappled World: A Study of the Boundaries of Science.* Cambridge University Press.

Collingwood, R. G. (1946). *The Idea of History.* Oxford University Press.

Feller, W. (1952). The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics*, 55(3), 468–519.

Goldstone, J. A. (1991). *Revolution and Rebellion in the Early Modern World.* University of California Press.
Ibn Khaldun. (1377/1967). *The Muqaddimah: An Introduction to History* (F. Rosenthal, Trans.). Princeton University Press.
Karatzas, I., & Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus* (2nd ed.). Springer.
Khasminskii, R. (2011). *Stochastic Stability of Differential Equations* (2nd ed.). Springer.
Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, 12(4), 595–600.
Nicolis, G., & Prigogine, I. (1977). *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations.* Wiley.
Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press.
Popper, K. (1959). *The Logic of Scientific Discovery.* Hutchinson.
Turchin, P. (2003). *Historical Dynamics: Why States Rise and Fall.* Princeton University Press.
Turchin, P. (2010). Political instability may be a contributor in the coming decade. *Nature*, 463(7281), 608.
Turchin, P. (2016). *Ages of Discord: A Structural-Demographic Analysis of American History.* Beresta Books.
Turchin, P., & Nefedov, S. A. (2009). *Secular Cycles.* Princeton University Press.
Turchin, P., Brennan, R., Currie, T. E., et al. (2015). Seshat: The Global History Databank. *Cliodynamics*, 6(1), 77–107.
Turchin, P., Currie, T. E., Whitehouse, H., et al. (2018). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences*, 115(2), E144–E151.
Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences* (3rd ed.). Academic Press.

A Mathematical notation summary

Symbol	**Meaning**
N	Population size
N_c	Critical population threshold for statistical aggregation
t	Time
$X(t)$	State vector at time t
$X_N(t)$	Aggregate variable for a population of size N
$\mathbb{R}^n$	n-dimensional real vector space
$P(E)$	Probability of event E
$P(E \mid X)$	Conditional probability of E given state X
$F(X, \theta, \varepsilon)$	Evolution function with parameters θ and stochastic term ε
θ	Model parameters
$\varepsilon(t)$	Stochastic process (noise, shocks)
$\lambda_{\max}$	Maximum Lyapunov exponent
δX	Perturbation to state X
$\mathcal{A}(t)$	Predictive accuracy at horizon t
$\mathcal{A}_0$	Initial predictive accuracy
$\mathcal{A}_\infty$	No-skill baseline accuracy (climatological / ergodic)
$\mathcal{S}$	Class of comparable systems

Symbol	Meaning
$\mathbb{E}_P[f]$	Expected value of function f under distribution P
$\mathrm{Var}[X]$	Variance of random variable X
$G(\cdot)$	Reflexive feedback function
$P_M(\cdot)$	Predictive distribution generated by model M
P^*	Fixed-point prediction
$\mathcal{G}$	Directed acyclic graph (causal structure)
$\Phi(t+\tau, t)$	Tangent linear propagator (state transition)
κ	Lipschitz constant of reflexive feedback
W_1	Wasserstein-1 distance on probability space
ρ	Intra-class correlation (average pairwise correlation of individual contributions)
N_{eff}	Effective sample size under correlated observations
BS	Brier score
REL	Reliability component of Brier score
RES	Resolution component of Brier score
UNC	Uncertainty component of Brier score
σ	Standard deviation of individual fluctuations (A1, Theorem 2, C.2) or diffusion coefficient (A4, Appendix C)
$dW(t)$	Wiener process increment
α	Bifurcation parameter in Appendix C model
K	Carrying capacity in Appendix C model
X_c	Critical threshold in Appendix C model
β	Baseline drift in Appendix C model
r	Kramers escape rate (Appendix C)
ΔV	Potential barrier height (Appendix C)
τ_{escape}	Mean first-passage time for barrier crossing (Appendix C)
τ_{sat}	Saturation time at which error reaches attractor diameter (Theorem 1, Appendix C)

B Glossary of key terms

Admissibility axiom. A constraint that any model must satisfy to qualify as valid within a given framework. Unlike logical primitives, admissibility axioms may encode consequences of other principles.

Axiom. In this paper, a foundational constraint defining the conditions under which CMH is valid. CMH axioms are admissibility conditions, not logical primitives.

Backtesting. Retrospective validation of a predictive model using historical data that was not used in model development.

Bifurcation. A qualitative change in system behaviour as parameters cross critical thresholds.

Calibration. The correspondence between predicted probabilities and observed frequencies. A well-calibrated model assigns sixty per cent probability to events that occur sixty per cent of the time.

Cliodynamics. The scientific study of historical dynamics using mathematical models and quantitative data, initiated by Peter Turchin.

Computational Macrohistory (CMH). A quantitative science studying the temporal evolution of large-scale complex social systems through dynamic, probabilistic, and computational mathematical models.

Deterministic chaos. The behaviour of a dynamical system that is deterministic, governed by fixed equations, yet unpredictable over long horizons because of sensitive dependence on initial conditions.

Directed acyclic graph (DAG). A graph with directed edges and no cycles, used to represent causal relationships among variables.

Effective sample size (N_{eff}). The sample size that an equally precise independent sample would require to achieve the same estimation precision as the actual correlated sample. Under intra-class correlation ρ, $N_{\text{eff}} = N/[1 + (N-1)\rho]$.

Ensemble statistics. Statistical properties such as mean and variance computed across multiple realisations or trajectories of a stochastic system.

Ergodicity. The property whereby time averages equal ensemble averages. In CMH, conditioned ergodicity means that historical trajectories under similar conditions belong to the same statistical class.

Falsifiability. The property of a scientific claim that it can, in principle, be proven false by evidence. A criterion of scientific status proposed by Karl Popper.

Lyapunov exponent. A measure of the rate at which nearby trajectories in a dynamical system diverge. Positive Lyapunov exponents indicate chaotic dynamics.

Oseledets decomposition. The multiplicative ergodic theorem that decomposes the tangent linear propagator of a dynamical system into Lyapunov subspaces with well-defined exponential growth rates.

Phase transition. A discontinuous change in system state emerging from continuous underlying dynamics, as when water freezes or a revolution erupts.

Prediction horizon. The temporal limit beyond which forecast uncertainty exceeds useful thresholds.

Reflexivity. The property whereby knowledge of a prediction can alter the system being predicted, potentially affecting the accuracy of the prediction.

Secular cycle. A long-term oscillation, of the order of 200 to 300 years, in political stability, population, and elite dynamics observed in agrarian societies.

Self-defeating prophecy. A prediction that triggers actions preventing the predicted outcome.

Self-fulfilling prophecy. A prediction that triggers actions producing the predicted outcome.

State space. The set of all possible states a system can occupy, represented mathematically as a vector space.

Stationary distribution. A probability distribution that remains unchanged under the dynamics of the system.

Structural causal model. A mathematical framework representing causal relationships through directed graphs and structural equations.

Structural-demographic theory. A theoretical framework linking population dynamics, elite competition, and state fiscal health to political instability.

Wasserstein-1 distance (W_1). A metric on the space of probability distributions with finite first moment, defined as the minimum cost of transporting one distribution into another under the cost function $c(x, y) = |x - y|$. Also called Earth Mover's Distance. Topologically weaker than the Total Variation metric and metrises weak convergence, making it the more appropriate geometry for belief-dynamics models where small shifts in predicted probabilities produce proportionally small behavioural responses.

C A bistable model with multiplicative noise consistent with the CMH axioms

This appendix presents a formal model that is richer than the minimal schematic toy of earlier versions while remaining analytically tractable. It demonstrates that the CMH axioms define a non-empty, internally consistent class of models capable of exhibiting the phenomena (critical transitions, noise-induced switching, reflexive feedback) that the framework is designed to address.

C.1 Model structure

Consider a one-dimensional aggregate variable $X(t) \in \mathbb{R}_+$ representing, for instance, a normalised systemic stress index. Its evolution is governed by the stochastic differential equation

$$dX = \left[\alpha X \left(1 - \frac{X}{K}\right)(X - X_c) + \beta\right] dt + \sigma X \, dW(t)$$

where all parameters are positive: $\alpha > 0$ scales the nonlinear drift, $K > 0$ is a saturation scale, $X_c \in (0, K)$ is the unstable equilibrium of the deterministic skeleton separating the two stable equilibria at $X = 0$ and $X = K$, $\beta \geq 0$ is a constant baseline drift representing exogenous pressure, $\sigma > 0$ controls the multiplicative noise intensity, and $W(t)$ is a standard Wiener process. The multiplicative structure $\sigma X \, dW$ ensures that fluctuations scale with the current stress level: a system near equilibrium experiences small perturbations, while a system already in transition experiences amplified volatility.

The drift term $f(X) = \alpha X(1 - X/K)(X - X_c) + \beta$ is a cubic polynomial and therefore locally Lipschitz, satisfying A4. For β sufficiently small relative to the potential barrier height $\Delta V = \alpha X_c^3(2K - X_c)/(12K)$, the deterministic flow has two stable fixed points near 0 and K, separated by an unstable fixed point near X_c. The system is bistable: trajectories relax to one of the two basins and can switch basins only through noise-induced transitions.

Positivity and domain invariance. The state variable X is a stress index and must remain in the physically meaningful domain $\mathbb{R}_+$. The multiplicative noise $\sigma X \, dW(t)$ vanishes at $X = 0$, but a Wiener process can in principle drive X negative. Feller's boundary classification test (Feller, 1952; Karatzas & Shreve, 1991, §5.5) resolves this. For a one-dimensional SDE $dX = \mu(X)dt + \varsigma(X)dW(t)$, the boundary at $X = 0$ is classified by the scale function $s(x) = \int^x \exp\left(-2\int^y \frac{\mu(z)}{\varsigma^2(z)}\, dz\right) dy$ and the speed measure. For the present model, $\mu(X) = \alpha X(1 - X/K)(X - X_c) + \beta$ and $\varsigma(X) = \sigma X$. As $X \to 0^+$, $\mu(X) \to \beta > 0$ (the baseline drift is strictly positive) while $\varsigma^2(X) = \sigma^2 X^2 \to 0$ faster than $\mu(X)$. The scale function integral $\int^y \frac{\mu(z)}{\varsigma^2(z)}\, dz \sim \int^y \frac{\beta}{\sigma^2 z^2}\, dz$ diverges to $-\infty$ as $y \to 0^+$, placing $X = 0$ in the *entrance* boundary class, since the scale function diverges while the speed measure remains finite: the boundary is inaccessible from the interior, and no trajectory starting at $X_0 > 0$ reaches zero in finite time.

The domain $\mathbb{R}_+$ is therefore stochastically invariant. The calculation also confirms that $X = 0$ is not an absorbing state, which is physically appropriate: a dormant stress index can be reactivated by the baseline drift β.

C.2 Statistical aggregation (A1–A2)

The variable X is an aggregate descriptor. Its validity under A1 requires that the population size N exceeds the threshold $N_c = \sigma^2(1-\rho)/(\epsilon^2\alpha - \sigma^2\rho)$ given by Theorem 2, where σ^2 is the variance of individual-level fluctuations, ρ is the intra-class correlation, ϵ the desired precision, and α the confidence level (not to be confused with the drift parameter α, which carries a distinct role). For fixed parameters $\theta = (\alpha, K, X_c, \beta, \sigma)$, the system's stationary distribution $P_\theta(X)$ is well-defined and independent of initial conditions, satisfying the conditioned ergodicity of A2. Time averages of any integrable observable $g(X)$ converge to the ensemble average $\mathbb{E}_{P_\theta}[g]$ for almost every trajectory. Structural breaks, a shift from, say, K_1 to K_2, produce a new stationary distribution, and comparisons across the break require recalibration, exactly as A2 prescribes.

C.3 Structural causality and continuous dynamics (A3–A4)

The model embodies structural causality (A3) because the probability of crossing the critical threshold X_{crit} depends on the current state:

$$P(X(t+\Delta t) > X_{\text{crit}} \mid X(t), \theta) \neq P(X(t+\Delta t) > X_{\text{crit}})$$

for any X_{crit} not degenerate with the stationary distribution. The causal drivers are the structural parameters θ, the current state $X(t)$, and the noise realisation. The locally Lipschitz drift guarantees existence and uniqueness of strong solutions (A4).

C.4 Endogenous indeterminacy (A5)

The deterministic skeleton of this model is not chaotic but multistable: trajectories relax to one of the two stable equilibria depending on initial conditions, and once settled in a basin they remain there indefinitely in the absence of noise. Indeterminacy enters through the stochastic term $\sigma X\, dW(t)$. For $\sigma > 0$, the system undergoes noise-induced barrier crossing: a trajectory in the lower basin ($X \approx 0$) may be driven by a sequence of favourable Wiener increments across the unstable equilibrium X_c into the upper basin ($X \approx K$), and vice versa. For this multiplicative-noise model, the Kramers escape rate is governed by the effective potential barrier:

$$\tau_{\text{escape}} \sim \exp\left(\frac{\Delta V}{\sigma^2 X_c^2}\right)$$

where $\Delta V = \alpha X_c^3(2K - X_c)/(12K)$ is the height of the potential barrier separating the two basins. (This reduces to the additive-noise expression $\exp(\Delta V/\sigma^2)$ only in the approximation where the diffusion coefficient is treated as constant near the barrier, i.e. $\sigma X \approx \sigma X_c$.) The actual crossing time for any single trajectory is exponentially distributed with mean τ_{escape}: it is intrinsically unpredictable, in the precise sense that even with perfect knowledge of the current state and the governing equations, the moment of transition cannot be forecast. Two trajectories initialised at the same point X_0 diverge because they experience independent Wiener increments, not because of sensitive dependence on the initial condition (the deterministic skeleton is contracting within each basin) but because the stochastic forcing is irreducibly unpredictable.

This is the regime of A5, clause (ii): noise-induced indeterminacy. The model satisfies A5 not through deterministic chaos ($\lambda_{\max} > 0$) but through irreducible stochastic barrier crossing. The operational consequence is identical: point predictions of the exact state at a future time are impossible, and the only meaningful prediction is the occupation probability, the likelihood that the system will be in the upper basin at horizon τ, which can be estimated from the Kramers rate and the current basin.

C.5 Predictive decay (A7)

The predictive accuracy of the bistable model follows the saturating form of A7. For a symmetric two-state process with switching rate r, the probability that the system remains in the original basin is

$$P_{\text{same basin}}(\tau) = \tfrac{1}{2}(1 + e^{-2r\tau})$$

where $r = 1/\tau_{\text{escape}}$ is the escape rate. At short horizons ($\tau \ll \tau_{\text{escape}}$), the system is almost certainly still in the original basin and prediction accuracy is high. At intermediate horizons the accuracy decays exponentially toward the ergodic baseline (the stationary basin-occupation probability, approximately 1/2 for a symmetric potential). At long horizons, only the stationary distribution is knowable, exactly the saturating form $\mathcal{A}(\tau) = \mathcal{A}_\infty + (\mathcal{A}_0 - \mathcal{A}_\infty)e^{-r\tau}$ with $\mathcal{A}_\infty \approx 1/2$ that A7 prescribes. Unlike the deterministic-chaos regime, where decay is governed by $\lambda_{\max}$, here the decay rate r is set by the Kramers escape time, which can be much slower (years to decades for macropolitical systems with high barriers ΔV relative to the noise intensity σ^2). The horizon of useful prediction is correspondingly longer, consistent with the empirical finding that structural stress indices discriminate among regimes over 5–15 year windows (WP-2026-004).

C.6 Bounded reflexivity (A6)

Reflexive feedback is modelled by allowing the baseline drift β and the nonlinearity α to respond to a published prediction P:

$$\beta = \beta_0 + \gamma_\beta P, \quad \alpha = \alpha_0 + \gamma_\alpha P$$

where $\gamma_\beta, \gamma_\alpha$ capture the intensity of the behavioural response. A prediction of high instability (P large) might increase external pressure (β rises, destabilising) or strengthen internal cohesion (α rises, potentially stabilising or destabilising depending on the sign of $X - X_c$). The combined feedback function $G(P) = \left(\gamma_\beta + \gamma_\alpha \frac{\partial f}{\partial \alpha}\right) P$ is linear in P and therefore a contraction provided $\left|\gamma_\beta + \gamma_\alpha \frac{\partial f}{\partial \alpha}\right| < 1$, satisfying the condition of Theorem 3. When this bound holds, a unique self-consistent prediction P^* exists and the iterative belief dynamics converge. When the bound is violated, the system is reflexively unstable, and CMH predictions are not licensed; this is the operational demarcation that Theorem 3 provides.

C.7 Falsifiability (A8)

The model generates predictive distributions $P_M(X(t+\Delta t) > X_{\text{crit}} \mid X(t), \theta)$ that can be estimated via Monte Carlo simulation and tested against historical data. For a set of N retrospective predictions, the Brier score decomposes as in Theorem 4, and a CMH-compliant validation report includes REL, RES, UNC, and the aggregate BS. The model is thus falsifiable in the precise sense of A8.

C.8 Interpretive status

The model is not proposed as a calibrated description of any specific historical case. It is an existence proof that the CMH axioms define a mathematically coherent class of stochastic dynamical systems that exhibit critical transitions, noise-induced regime switching, bounded reflexivity, and falsifiable probabilistic predictions, the full set of properties the framework requires.

Document I of the Computational Macrohistory Series

Foundations Institute of Computational Social Science
`https://ficss.institute`

Contact: stefano.angeli@ficss.institute