Human Societies as Complex Adaptive Meta-Systems:

Introducing the CAMS Framework

Kari Freyr McKern

Independent Researcher, Sydney, Australia • neuralnations.org

April 29, 2026

Abstract

This paper presents the Complex Adaptive Model of Societies (CAMS) as a methodological framework for treating human societies as dissipative meta-organisms—complex adaptive systems that maintain coherence by importing free energy and exporting social entropy. The framework comprises eight institutional nodes (Helm, Shield, Lore, Archive, Craft, Flow, Stewards, Hands) each scored on four orthogonal metrics (Coherence, Capacity, Stress, Abstraction) via blind multi-LLM ensemble, yielding a 32-dimensional state-space trajectory. We present the formal architecture and illustrate its operationalisation through five longitudinal datasets: Germany (1880–2026), United Kingdom (1880–2026), Russia (1800–2026), Argentina (1950–2026), and Thailand (1850–2025), comprising approximately 634 society-years and 5,072 node-year scoring records (the full Neural Nations repository covers 38 societies and approximately 39,351 node-year records). The datasets demonstrate internal consistency and generate falsifiable predictions; independent empirical testing against external benchmarks is developed in companion papers. The primary structural finding—a universal Stress–Capacity anti-correlation (ranging from 0.351 to 0.797 across the five datasets)—is robust across all scorer combinations and constitutes the framework's central falsifiable prediction. We present six specific falsification criteria and a retrospective validation sketch for Germany 1923–1945. CAMS is positioned as a methodological development: a formal language for civilisational dynamics rather than a confirmed empirical theory, with the architecture, scoring protocol, and initial evidence presented here as the foundation for subsequent independent testing.

Keywords: complex adaptive systems, CAMS, thermodynamics, social entropy, institutional dynamics, dissipative structures, ensemble AI scoring, civilisational analysis

1 Introduction

Human societies are not merely collections of individuals pursuing self-interest. They are coherent meta-organisms—living, dissipative structures that extract energy from their environment, process information, maintain internal coupling, and contend with entropy in ways that are both predictable and potentially quantifiable. This paper introduces the Complex Adaptive Model of Societies (CAMS): a thermodynamic framework that treats every human society, from Palaeolithic band to modern nation-state, as an integrated entity whose dynamics are governed by the same physics that governs all complex adaptive systems.

The framework rests on eight institutional nodes—Helm, Shield, Lore, Archive, Craft, Flow, Stewards, and Hands—each scored on four orthogonal metrics: Coherence (C), Capacity (K), Stress (S), and Abstraction (A). Scores are produced by blind multi-LLM ensemble, yielding a 32-dimensional state-space trajectory for any society across any time period for which documentary evidence exists. The scoring ensemble generates not only mean estimates but standard deviation envelopes that serve as direct measures of historical epistemic uncertainty.

We present this paper explicitly as a methodological development rather than a completed empirical study. The framework, scoring protocol, and five illustrative datasets are offered as a formal system whose architecture is internally consistent, whose predictions are falsifiable, and whose initial results are sufficiently structured to warrant independent replication. The circularity concern inherent in any LLM-scored framework—that theoretical priors may be embedded in prompts, generating scores that confirm the theory—is real and is addressed directly in Section 5. Our claim is that the five datasets demonstrate internal consistency and generate falsifiable structure; the work of independent empirical confirmation is reserved for subsequent papers.

The practical case for the framework is best introduced through data. Consider Germany 1880–2026 in the CAMS5 dataset. In 1932—the Depression trough—mean Node Value across all eight nodes is 1.60. In 1933, following the Nazi consolidation of power, mean Node Value jumps to 9.69 in a single year: the signature of a coercive re-coordination event that suppresses entropy indicators without resolving underlying structural stress. By 1940 the system reaches its measured peak of 15.08, driven by extreme Helm and Shield scores (V = 17.6 and 20.1 respectively) reflecting the wartime command economy at maximum coercive output. In 1945 the system collapses to a mean Node Value of 4.53—the only year in the five-country dataset in which all eight nodes simultaneously record negative Node Values, and the most negative system mean across all 634 scored society-years; every node records Stress scores of 8.8 or higher and Capacity scores of 2.4 or lower, the thermodynamic signature of total entropic overload. The post-war Wirtschaftswunder is a separate trajectory beginning in the late 1940s and stabilising around 13.8 in the 1960s. These are not historian's judgments overlaid on the model; they are the model's own output from blind scoring.

The remainder of this paper proceeds as follows. Section 2 presents the eight nodes and their functional grounding. Section 3 details the four metrics. Section 4 develops the mathematical formalism. Section 5 describes the ensemble scoring protocol, including epistemic fidelity considerations. Section 6 presents empirical results from five longitudinal datasets. Section 7 addresses geographic clustering and scale-covariant structure. Section 8 presents six specific falsification criteria and a retrospective validation sketch. Section 9 concludes.

2 The Eight Nodes

2.1 Architecture and Functional Grounding

The eight nodes are derived from a rate-separation argument: eight is the minimal number of functionally distinct coordination sub-systems that permits full coverage of the functional space of any complex social entity while maintaining near-optimal communication efficiency across the network (McKern 2026c). The formal derivation uses communication-cost and redundancy arguments from network information theory; here we note that this is not an architectural choice but a mathematical result, with the implication that any stable complex social entity will exhibit all eight functional roles, whether or not they are institutionally differentiated.

Each node has documented precursors in non-human primate and mammalian social organisation (McKern & Foulkes 2026b; Boehm 2012; Wrangham & Peterson 1996; Tomasello 2019; McComb et al. 2001). The coordination architecture CAMS measures is phylogenetically continuous: human institutional complexity scaled up from functional sub-systems that already existed in our primate ancestors. Helm traces to alpha-dominance and coalition brokering in chimpanzees (de Waal 1982; Boehm 2012). Shield traces to territorial boundary patrol and inter-group threat management across all social primates (Wrangham & Peterson 1996). Lore traces to social learning transmission and ritual behaviour (Whiten et al. 1999; Tomasello 2019). Archive traces to spatial memory and inter-generational knowledge transfer in elephants and cetaceans (McComb et al. 2001). Craft traces to tool manufacture in great apes (Boesch & Boesch 1990). Flow traces to reciprocal exchange networks in primates (Trivers 1971; Silk 2007). Stewards traces to food hoarding and resource governance in matriarchal mammalian groups (Moss 2000). Hands traces to collective foraging and coordinated labour in social mammals (Wilson 1975; Henrich 2015).

2.2 Node Descriptions

- **Helm**: executive coordination, strategic direction, and authoritative decision-making. In modern societies: government, executive leadership, and political authority.
- **Shield**: security, defence, and the management of external and internal threat. Territorial and policing functions at all scales.
- **Lore**: normative and symbolic systems—religion, ideology, law, shared narrative, and the meaning-making apparatus through which the meta-organism understands itself.
- **Archive**: institutional memory—accumulated knowledge, records, and transmitted expertise. Libraries, educational systems, legal codification, and all forms of institutionalised recall.
- **Craft**: productive and technological capacity. The ability to transform raw materials and energy into usable outputs. Distinct from Hands in representing the design and organisation of production, not its physical execution.
- **Flow**: economic circulation, trade, and the distribution of resources and information within and beyond the system. The meta-organism's circulatory sub-system.
- **Stewards**: governance of property, land, capital, and material assets over medium-to-long timescales. Responsible for the sustainability of the resource base.
- **Hands**: the direct labour and productive capacity of the population. The most populous node and the one most directly connected to the material welfare of the majority.

2.3 Fast Loop and Slow Loop

Nodes partition into two loops with different characteristic timescales, derived empirically from the validation datasets and confirmed in the v3.2-R specification. The Slow Loop (Archive, Lore, Stewards) operates over decades to centuries and governs structural coherence and long-horizon memory.

The Fast Loop (Helm, Shield, Craft, Flow, Hands) operates over months to years and governs entropy management, material throughput, and tactical response.

This is a 3–5 partition, not a 4–4 partition. The critical insight from validation is that Helm behaves as a fast-loop reactive node under acute stress—responding to immediate crisis signals rather than maintaining long-horizon strategic coherence. The 1933 Germany data illustrates this precisely: Helm Coherence rises from 6.4 in 1933 to 8.4 by 1935, tracking rapid coercive restructuring rather than deep institutional change. Archive and Lore, by contrast, show multi-decade degradation trajectories during the Weimar period and only partial recovery in the post-war decades. The Scissors Effect occurs when fast-loop Coherence rises while slow-loop Coherence decays, producing divergent trajectories that indicate coercively maintained surface cohesion decoupled from genuine structural renewal—a signature observed during the 1938–1940 coercive peak in this dataset, rather than as a consistent 5–15 year early-warning signal.

3 The Four Metrics

Each node is scored on four metrics (0–10 scale), representing the minimal state variables required to characterise a dissipative institutional sub-system:

Metric	Symbol	Definition	Thermodynamic Analogue	Direction
Coherence	C	Internal consistency and institutional integrity of the node	Structural order; low internal entropy	Higher = healthier
Capacity	K	Functional throughput; ability to translate intention into outcome	Available free energy for coordinated work	Higher = healthier
Stress	S	Entropy load: conflicting demands, resource pressure, informational disorder	Entropy production rate	Lower = healthier
Abstraction	A	Symbolic and conceptual reach; degree of operation via abstract rather than material mechanisms	Information-processing complexity; cognitive overhead	Double-edged

Table 1: The four CAMS metrics. Abstraction is double-edged: it amplifies effective cognition when C and K are strong, and amplifies fragility when they are weak—the Abstraction Paradox.

The four metrics are not interchangeable with traditional social-science measures but are more fundamental: they capture the thermodynamic state of the institutional sub-system rather than proxies derived from that state. GDP approximates Craft and Flow Capacity aggregated but misses distributional dynamics, coherence failures, and the abstraction overhead that makes advanced systems simultaneously more powerful and more fragile. Regime type approximates Helm Coherence but predicts near-zero variance in systemic health once node-level Capacity and Stress are controlled (McKern 2026b).

4 Mathematical Formalism

4.1 State Space

At time t, a society occupies a position in the 32-dimensional CAMS state space:

$$\Psi(t) = \{C_i(t), K_i(t), S_i(t), A_i(t)\} \quad \text{for } i = 1 \ldots 8$$

Scores are integers on $[0, 10]$. Ensemble averaging over five independent scoring passes produces continuous means in $[0.0, 10.0]$ with associated standard deviations. Treating ensemble means as real numbers in derived quantities is justified because averaging over five scorers effectively maps the integer-scale inputs to a continuous distribution; the precision of derived quantities is bounded by the ensemble standard deviation, not by integer discretisation.

4.2 Node Value

The primary scalar derived from raw scores is Node Value, the canonical CAMS v3.2-R specification:

$$V_i(t) = C_i(t) + K_i(t) - S_i(t) + 0.5 \times A_i(t)$$

Node Value is the net institutional output of node i—the sum of coherent capacity minus entropy load, with a 0.5 weighting on Abstraction reflecting its double-edged contribution. It is defined on $(-\infty, +\infty)$ in principle but empirically ranges from approximately 8 to +22 in the current datasets. Negative Node Values (which occur only under extreme entropic overload, e.g. Germany 1945: Helm $V = -6.0$, Flow $V = -5.7$) signify nodes in which Stress so exceeds Coherence and Capacity that the institution is net-negative to system coordination.

Verification against published dataset: Germany 1880 Helm — $C = 6.8$, $K = 7.2$, $S = 2.8$, $A = 6.2$. $V = 6.8 + 7.2 - 2.8 + 0.5(6.2) = 14.3$. CSV value: 14.3

System health is $H(t) = \sum_i V_i(t)/8$. The Cognitive Signature, capturing the net cognitive-energetic capacity of the meta-organism, is:

$$\Gamma(t) = \bar{A}(t) \times \bar{C}(t) \times (\bar{K}(t) - \bar{S}(t))$$

where overbars denote cross-node means. $\Gamma(t)$ is positive when net Capacity exceeds Stress—the system has energy for coherent abstract action—and negative when the system is under entropic overload.

4.3 Bond Strength

Bond Strength $B_i(t)$ is the inter-node coupling strength of node i: the degree to which node i is dynamically integrated with the overall system, reflecting its structural connectivity and contribution to system-level coordination. Bond Strength is always non-negative, because coupling is a structural property independent of the direction of a node's trajectory—a node can be degraded and still maintain coupling with other nodes (as in Germany 1945, where Bond Strength values of 1.2–2.1 persist despite all Node Values being negative).

Bond Strength is computed by the CAMS5 ensemble pipeline from the cross-node co-movement structure of the scoring passes; its formal derivation is specified in the CAMS v3.2-R technical documentation. The published datasets include the Bond Strength column as a directly reportable empirical quantity. In the current datasets, Bond Strength ranges from approximately 1.2 (Germany 1945, catastrophic entropic overload) to approximately 44 (Germany 1940, wartime coercive peak on Shield), with healthy systems typically showing values of 25–40 per node.

Formally, $B_i(t)$ is derived from the algebraic connectivity structure of the inter-node coupling graph, estimated from rolling 5-year score windows:

$$B_i(t) = \lambda_2(L(t)) \times w_i(t)$$

where λ_2 is the Fiedler eigenvalue of the normalised inter-node Laplacian $L(t)$ and w_i is the coupling weight of node i in the Fiedler eigenvector decomposition. This formulation guarantees $B_i(t) \geq 0$ by construction: coupling strength is a structural property independent of the direction of a node's trajectory. The formula presented in Paper 1 v1.0 ($B_{ij} = \rho(V_i, V_j) \times \frac{1}{2}(V_i + V_j)$) was inconsistent with the empirical CSV values and has been superseded by this specification.

4.4 Criticality Index and Phase Transitions

The Criticality Index $\chi(t) = \bar{B}(t)/\omega(t)$, the ratio of mean Bond Strength to institutional disorder frequency $\omega(t)$, functions as the system's order parameter, where:

$$\omega(t) = \mathrm{std}_i(\Delta S_i/\Delta t)$$

is the cross-node dispersion of stress-change rates—how asynchronously the eight nodes are responding to the same forcing conditions. Low ω indicates synchronised absorption (ordered); high ω indicates divergent stress dynamics (disordered). Consequently, low χ signals danger: coupling has weakened relative to the disorder the system is generating. The v3.2-R specification calibrates four escalating alert tiers from retrospective analysis of France 1789, Germany 1933, and USA 2020, using non-overlapping ranges:

Tier	**χ range**
WATCH	$0.42 \leq \chi < 0.57$
WARNING	$0.35 \leq \chi < 0.42$
CRITICAL	$0.30 \leq \chi < 0.35$
EXTREME	$\chi < 0.30$

Table 2: Criticality Index tier thresholds (CAMS v3.2-R). Ranges are non-overlapping and escalate as χ decreases. Calibrated retrospectively against France 1789, Germany 1933, and USA 2020.

Note on order parameters: The Criticality Index $\chi = \bar{B}/\omega$ (low = danger) is the inverse of the disorder/coupling ratio $\kappa = D_\Psi/\Lambda$ used in some companion derivations. Both formulations identify the same phase-transition threshold; readers should treat low χ and high κ as equivalent danger signals. A full reconciliation of the two parameterisations is provided in McKern & Foulkes (2026b).

4.5 Free Energy Potential and Forward Model

The system's macro-state is characterised by a Free Energy potential:

$$F(t) = -\alpha\bar{V}(t) - \beta\Lambda(t) + \gamma D\Psi(t)$$

where $\bar{V}$ is mean Node Value (health), Λ is mean Bond Strength (coupling), D_Ψ is the variance of the state vector (internal disorder), and α, β, γ are empirically estimated weights. The system tends toward the minimum of F; a transition event (revolution, collapse, re-consolidation) corresponds to a bifurcation in this potential landscape.

The forward model captures state evolution as:

$$\Psi_i(t+1) = \Psi_i(t) + \Delta_{\text{drift}}(t) + \Delta_{\text{shock}}(t) + \varepsilon(t)$$

In this expression, $\Psi_i(t)$ is the 4-vector state of node i: $[C_i, K_i, S_i, A_i]^\top$; Δ_{drift} and Δ_{shock} are correspondingly node-specific 4-vector updates. Δ_{drift} is the smooth institutional drift driven by structural path dependence, Δ_{shock} captures discrete event-driven disruptions (wars, pandemics, policy shocks), and ε is the residual scoring uncertainty. The drift and shock terms are estimated from the rolling ensemble data; the forward model is used for the prospective accuracy predictions specified in Section 8.

5 The Ensemble Scoring Protocol and Epistemic Fidelity

5.1 Protocol Design

CAMS operationalises a 32-dimensional institutional assessment across societies and centuries through blind multi-LLM ensemble scoring. Each scoring pass presents an LLM with the historical record for a specific society-year, without revealing the CAMS framework, prior scores, or adjacent-year scores. The LLM scores each of the eight nodes on each of the four metrics (0–10) with chain-of-thought justification. Five independent passes produce an ensemble mean and standard deviation at the node-metric level.

The blind protocol insulates results from three sources of systematic bias: single-LLM conceptual priors, framework confirmation bias (if the scorer knows the expected result), and temporal autocorrelation (if adjacent years are scored in sequence). The standard deviation envelope is not noise but an epistemic fidelity indicator: it shows where the historical record is genuinely ambiguous and where additional archival research would most improve scoring confidence.

5.2 Circularity and the LLM Prior Problem

The primary methodological concern with LLM-based scoring is that theoretical priors may be embedded in the scoring prompts, generating scores that confirm the theory by construction. This concern is legitimate and cannot be fully dismissed on the basis of internal consistency alone. We address it through three design features. First, the scoring prompt does not mention CAMS, thermodynamics, or any theoretical framework: it asks the LLM to rate institutional coherence, capacity, stress, and abstraction on a purely descriptive basis. Second, the cross-LLM concordance results (Section 5.3) show that agreement is higher on well-documented periods and lower on ambiguous ones—the pattern expected from genuine extraction of historical signal, not theoretical imposition. Third, retrospective predictions (Section 8) derived from the scoring data match known historical outcomes without post-hoc calibration.

Independent replication—scoring the same societies with the same protocol by research teams with no prior knowledge of CAMS—is the definitive test and is explicitly called for in the falsification criteria in Section 8.

5.3 Inter-Scorer Reliability

Inter-scorer reliability was assessed using Intraclass Correlation Coefficients (ICC), as appropriate for five-pass ensemble data. On the USA 1900–2026 validation pipeline, ICC(2,k) = 0.973 (average-rater absolute agreement) and ICC(2,1) = 0.199 (single-rater agreement). The high average-rater ICC confirms that the five-pass ensemble mean is a reliable composite measure; the low single-rater ICC confirms that individual passes show substantial variance, justifying the ensemble rather than single-pass approach.

Across the five longitudinal datasets presented here, Pearson concordance between scoring pass means on Node Value and Stress ranges from 0.66 to 0.86. This concordance is highest on data-rich modern periods (post-1900) and lowest on data-sparse pre-modern or pre-archival periods—a theoretically expected and empirically reassuring pattern.

6 Empirical Results: Five Longitudinal Datasets

6.1 Dataset Overview

Society	**Period**	**Years**	**Node-years**	**S–K ρ**
Germany	1880–2026	147	1,176	0.715
United Kingdom	1880–2026	147	1,176	0.593
Russia	1800–2026	227	1,816	0.428
Argentina	1950–2026	77	616	0.797
Thailand	1850–2025	36	288	0.351
TOTAL	1800–2026	**634 soc.-yrs**	**5,072**	**0.577 (mean)**

Table 3: Summary of the five core validation datasets. 'Node-years' = society-years × 8 nodes. Thailand is scored at 5-year intervals (36 observations spanning 175 years), accounting for the smaller society-year count relative to the period length. The full Neural Nations repository covers 38 societies and approximately 39,351 node-year records.

6.2 The Universal Stress–Capacity Anti-Correlation

The framework's primary thermodynamic prediction is a universal negative correlation between Stress and Capacity across nodes and time—a structural consequence of the dissipative-structure model: sustained high Stress overwhelms dissipative Capacity, and maintaining high Capacity under extreme Stress requires structural buffers that most institutions cannot sustain indefinitely. The correlation is recovered across all five datasets (Table 3), with a cross-dataset mean of = 0.577.

The variation in correlation magnitude is itself theoretically interpretable. Argentina (= 0.797) shows the tightest coupling: the chronic institutional fragility of a commodity-export-dependent peripheral economy propagates external shocks across all nodes with minimal buffering, producing near-proportional Stress–Capacity co-movement. Germany (= 0.715) shows strong coupling interrupted by the 1933 discontinuity—the coercive Nazi re-coordination suppresses Stress indicators while driving Capacity in selected nodes, producing a brief decoupling visible as a local break in the correlation structure.

Russia (= 0.428) shows the weakest coupling in the dataset. This is not a measurement anomaly but the thermodynamic signature of the continental-defensive civilisational type: Russian institutions have evolved under conditions of chronic extreme external threat to maintain Capacity under Stress levels that would produce rapid Capacity decline in more tightly coupled systems. Shield and Helm sustain high Capacity even when Stress is extreme, driven by existential threat mobilisation. This is the quantitative basis for what comparative historians describe as 'Russian resilience under adversity'—measurable and structurally explicable, not cultural mysticism.

Thailand (= 0.351) shows an analogously weak coupling driven by a different mechanism: the tight Helm–Lore–Shield coupling characteristic of Thailand's monarchical-Buddhist governance architecture buffers the Stress–Capacity relationship by coordinating the three nodes that most directly manage entropy, decoupling the system's response to stress from what would otherwise be a proportional Capacity decline.

6.3 Germany 1880–2026: A Retrospective Validation Sketch

The Germany dataset provides the most structurally rich validation case. Key phase signatures and their thermodynamic interpretations:

- **1880–1913 (Wilhelmine consolidation)**: Mean NV = 12.9–13.3. Helm and Shield at high Capacity with moderate Stress. Pre-war signature: Shield Stress begins rising from 1905, while system mean Bond Strength declines from approximately 40.4 (1900) to 25.1 (1914) and Shield Bond Strength from 41.5 to 29.0—consistent with a system approaching a criticality threshold through arms-race dynamics (note: exact 1880 values not in current CSV extract).

- **1923 (hyperinflation)**: Mean NV = 1.1. Stewards V = 5.0, Flow V = 4.1. The entropic overload is concentrated in the material nodes, precisely as the currency collapse mechanism would predict: Craft and Hands show partial survival (V = 1.4 and 0.4) while Flow and Stewards—the distribution and property nodes—collapse first.

- **1932 (Depression trough)**: Mean NV = 1.6. System has not recovered from 1923 structural damage; Lore coherence declining as ideological fragmentation accelerates.

- **1933 (Nazi consolidation)**: Mean NV jumps from 1.6 to 9.7 in a single year. This is the largest single-year Node Value increase in the 147-year series. It is the CAMS signature of a coercive re-coordination event: suppression of entropy indicators through political force rather than genuine institutional renewal.

- **1940 (wartime peak)**: Mean NV = 15.08. Shield V = 20.1 (C = 9.0, K = 9.6, S = 2.6)—the militarised state at maximal coercive output. Lore V = 9.2, the lowest node—the Nazi normative system shows low coherence consistent with propaganda-based, forced ideological conformity rather than genuine symbolic integration. The system's 1940 peak is structurally brittle: entirely dependent on continued military success and coercive suppression.

- **1945 (total collapse)**: Mean NV = 4.53. All eight nodes record negative Node Values—Helm V = 6.0, Flow V = 5.7, Stewards V = 5.4. All Stress scores are 8.8 or higher; all Capacity scores are 2.4 or lower. 1923 and 1946 also have negative system means (1.1 and 0.51 respectively), but 1945 is uniquely distinguished as the only year across all 634 scored society-years in which every node simultaneously posts a negative Node Value. The 1945 signature is the defining empirical case for what complete entropic overload looks like in CAMS: not a gradual decline but a sudden transition from coercively maintained high output (1940) to total system failure (1945).
- **1946–1970 (Wirtschaftswunder)**: Recovery trajectory from 0.51 (1946) through positive mean by 1947, stabilising at approximately 13.8 by 1960. The recovery is driven primarily by Craft, Flow, and Hands—the material fast-loop nodes—with Helm and Lore trailing, as expected from a genuine institutional reconstruction rather than coercive re-coordination.

7 Scale-Covariant Structure and Geographic Clustering

7.1 Scale-Covariance

The eight-node architecture exhibits scale-covariant structure: the same functional roles recur across social entities at different scales of aggregation—households, firms, civil society organisations, and nation-states—when the appropriate timescales are applied. This is a qualified claim: scale-covariance holds where functional differentiation is durable and non-redundant. The CAMS v3.2-R Compression Theorem specifies that the eight-node partition is a regime-dependent coarse-grained basis, not a universal upper bound; decomposition is warranted only where stable-period sub-node relations show persistent non-redundancy. A household operating with genuinely undifferentiated Archive and Shield functions may require a six-node representation; the claim is that eight nodes is the empirically recoverable optimum for the social entities examined to date.

The strongest evidence for scale-covariance comes from blind scoring of entities at the extremes of the organisational spectrum. Pre-contact Indigenous Australian societies (Palaeolithic-scale, population 300,000–500,000) and SpaceX (2015–2025) independently recover the same high-value lower-left attractor in phase space across all four LLM scorers—both showing high Node Values and low Stress across all eight nodes. This unexpected clustering across 60,000 years and every technological development imaginable is difficult to explain as a modelling artefact.

7.2 Geographic Civilisational Types

The full 38-society dataset recovers four major civilisational types, not imposed but emergent from hierarchical clustering on node-metric profiles. Riverine/centralised societies (China, Mesopotamia, Egypt) show high Lore–Archive coupling and stable Helm dominance; maritime/modular societies (British Empire, Greece, Singapore) show high Flow Capacity and competitive elite fragmentation; desert/distributed societies (Persian, Arabian) show high Lore coherence with decentralised Helm; continental/defensive societies (Russia, central European) show high Shield Capacity with characteristic weak Stress–Capacity coupling. Thailand illustrates a fifth sub-type—the buffer-state coupled configuration—in which tight Helm–Lore–Shield coordination produces exceptional stability through political turbulence, precisely because political-surface volatility is decoupled from deep institutional continuity.

8 Falsifiability and Validation Criteria

8.1 Six Falsification Criteria

A framework claiming thermodynamic grounding must be falsifiable in specific, testable terms. The following six criteria constitute the empirical contract for CAMS:

1. **Stress–Capacity anti-correlation**: (S, K) < 0.50 must be recoverable across all node-year observations in any independently scored dataset of 50 or more society-years, using any scoring protocol that produces Stress and Capacity estimates. A dataset in which (S, K) 0.30 would falsify the thermodynamic model.

2. **Bond–Health coupling**: r($\bar{B}$, H) > 0.60 in any independently scored longitudinal series of 30 or more years. Bond Strength and System Health should co-move; a series showing r < 0.40 sustained over 30 years would falsify the coupling claim.

3. **Shield ranking test**: Shield node must show elevated Stress rank (top 3 of 8 nodes by Stress) in 75% or more of independently identified pre-collapse windows (defined as 5–15 years before documented systemic collapse events). This is the specific prediction of the entropy-export mechanism: Shield absorbs or redirects system entropy under crisis conditions before it dissipates across all nodes.

4. **Algebraic connectivity decline**: Mean Bond Strength decline of 30% or more sustained over 10 years must predict a coordination phase transition (as independently documented by historians) within 15 years in 70% or more of test cases drawn from the 38-society dataset. This is the Criticality Index prediction.

5. **Cross-LLM concordance**: ICC(2,k) > 0.70 on Stress and Node Value between any two LLM scorers (from different developers, using the published prompt protocol) applied to any unseen test society not in the training corpus. The USA pipeline validation yielded ICC(2,k) = 0.973; the criterion requires a lower bound of 0.70 to account for less data-rich test cases.

6. **Prospective accuracy**: 2026–2028 trajectory predictions for the five societies in Table 3—derived from the 2025 ensemble means and the forward model in Section 4.5—must achieve 65% or greater directional accuracy (correct sign of mean Node Value movement) when evaluated against independently scored 2029 data. This criterion has a fixed evaluation date and cannot be post-hoc adjusted.

8.2 Limitations

Known limitations are as follows. Scoring quality degrades for pre-modern periods where written records are sparse; the envelope (SD) datasets accurately reflect this uncertainty but cannot resolve it without additional archival evidence. The four metrics, while orthogonal in construction, may not be fully independent in LLM measurement; structural equation modelling of metric inter-relationships across the full dataset is warranted. The five datasets do not yet include African, Mesoamerican, or South Asian longitudinal series; geographic universality requires these additions. CAMS currently operates at annual resolution; higher-frequency scoring for well-documented recent periods would sharpen crisis dynamics detection. Finally, the LLM prior problem (Section 5.2) cannot be fully addressed without independent replication by research teams external to the framework's development.

9 Conclusion

This paper has introduced CAMS as a formal methodological framework for treating human societies as dissipative meta-organisms, and illustrated its operationalisation through five longitudinal datasets spanning 226 years of scored history. The primary structural finding—a universal Stress–Capacity anti-correlation recovered across all five datasets, with magnitude varying predictably by civilisational type—is robust across scoring methodologies and constitutes the framework's central empirical claim.

The framework is presented as a methodological development, not a confirmed empirical theory. The architecture is internally consistent, the predictions are falsifiable, and the initial evidence is structured enough to warrant independent replication. The six falsification criteria in Section 8 define precisely what independent testing should attempt to establish or refute. The prospective accuracy criterion (Section 8.1, criterion 6) provides a fixed evaluation point that cannot be adjusted after the fact.

The datasets, ensemble scoring rubrics, envelope files, and phase-space visualisation tools are published openly at neuralnations.org and github.com/KaliBond/wintermute. The invitation is to replicate, extend, and challenge.

References

[1] Boesch, C. & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica*, 54(1–2), 86–99.

[2] Boehm, C. (2012). *Moral Origins: The Evolution of Virtue, Altruism, and Shame*. Basic Books.

[3] Dakos, V., Carpenter, S.R., van Nes, E.H. & Scheffer, M. (2012). Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B*, 367, 196–204.

[4] de Waal, F.B.M. (1982). *Chimpanzee Politics: Power and Sex Among Apes*. Jonathan Cape.

[5] Dunbar, R.I.M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6(5), 178–190.

[6] Gell-Mann, M. (1994). *The Quark and the Jaguar*. W.H. Freeman.

[7] Henrich, J. (2015). *The Secret of Our Success*. Princeton University Press.

[8] Holland, J.H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley.

[9] Holling, C.S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1–23.

[10] Kauffman, S.A. (1993). *The Origins of Order*. Oxford University Press.

[11] Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.

[12] Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. Springer.

[13] McComb, K., Moss, C., Durant, S.M., Baker, L. & Sayialel, S. (2001). Matriarchs as repositories of social knowledge in African elephants. *Science*, 292(5516), 491–494.

[14] McKern, K.F. (2026b). A taxonomy of human meta-systems: geographic and evolutionary clustering via thermodynamic signatures. Preprint, neuralnations.org.

[15] McKern, K.F. (2026c). Why eight nodes? A rate-separation and coupling argument for the minimal coordination architecture of complex societies. Preprint, neuralnations.org.

[16] McKern, K.F. & Foulkes, D.W.A. (2026a). CAMS and the modern evolutionary synthesis. Preprint, neuralnations.org.

[17] McKern, K.F. & Foulkes, D.W.A. (2026b). The adaptive cycle in civilisational phase space. Preprint, neuralnations.org.

[18] Moss, C. (2000). *Elephant Memories*. University of Chicago Press.

[19] Prigogine, I. & Stengers, I. (1984). *Order Out of Chaos*. Bantam Books.

[20] Scheffer, M., Carpenter, S.R., Lenton, T.M., Bascompte, J., Brock, W., Dakos, V., ... & Vandermeer, J. (2012). Anticipating critical transitions. *Science*, 338(6105), 344–348.

[21] Silk, J.B. (2007). The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B*, 362, 539–559.

[22] Strogatz, S.H. (2003). *Sync: The Emerging Science of Spontaneous Order*. Theia.

[23] Tainter, J.A. (1988). *The Collapse of Complex Societies*. Cambridge University Press.

[24] Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.

[25] Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35–57.

[26] Turchin, P. (2003). *Historical Dynamics: Why States Rise and Fall*. Princeton University Press.

[27] Turchin, P. (2023). *End Times*. Penguin Press.

[28] Whiten, A. et al. (1999). Cultures in chimpanzees. *Nature*, 399, 682–685.

[29] Wilson, D.S. & Wilson, E.O. (2007). Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology*, 82(4), 327–348.

[30] Wilson, E.O. (1975). *Sociobiology: The New Synthesis*. Harvard University Press.

[31] Wrangham, R. & Peterson, D. (1996). *Demonic Males*. Houghton Mifflin.