RESEARCH PROTOCOL · V1.0 · METHODOLOGY ONLY

The Brittle Network

A protocol for testing whether the Eastern Mediterranean collapsed because exogenous shocks struck a trade system optimized past the point of resilience — and whether the Sea Peoples were a *symptom* rather than a cause.

WINDOW	UNIT OF ANALYSIS	DESIGN	HYPOTHESES
1250 – 1100 BCE	The exchange network	Bayesian model comparison	H_0 – H_4, pre-registered

PRIMARY OUTCOME

Differential survival

The collapse of the Late Bronze Age is the best-known systemic failure in the ancient record and one of the worst-served by its own evidence. This document specifies how a multidisciplinary team would test the cascading-failure hypothesis rigorously enough to be wrong.

It is a protocol, not a study. It contains no results, and it is written so that its conclusions could fail: §5.1 lists seven conditions under which the central hypothesis is rejected, and the abstract to be published if it fails is drafted in advance.

METALS	NETWORK	TEXTUAL	FAILURE
Copper, tin, provenance, the supply chain that bound the system	Topology, routes, cascade dynamics, percolation	Amarna, Ḫattuša, Ugarit — claims by interested parties	Destruction, abandonment, and the ferrous transition after

Colour is used as a register throughout this document. It encodes which evidence stream a claim belongs to, and nothing else.

SECTION 00

Scope, Premise Audit, and the Epistemic Contract

Three corrections to the framing are entered before any framework is specified. A protocol that inherits an imprecise premise produces preciselooking answers to the wrong question.

“Prove or disprove” is not an available inferential mode

The collapse is a single, unrepeatable realization with `n = 1` at the system level. No dataset can *prove* a causal architecture for it. The protocol replaces proof with **severity of test**: each hypothesis must generate predictions that would very probably have failed had the hypothesis been false, and the analysis reports posterior model probabilities and Bayes factors rather than verdicts. A hypothesis that cannot be made to generate a differential prediction is declared **empirically inert for this study** and excluded rather than argued about.

> Given the observed spatiotemporal pattern of destruction, abandonment, contraction, and *survival* between 1250 and 1100 BCE, which class of generative model — exogenous-shock-only, network-topology-only, or shock-propagating-through-brittle-topology — assigns the highest likelihood to the observed pattern, and how large is the margin relative to the uncertainty in the evidence?

“1177 BC” is a heuristic anchor, not a datum

1177 BCE corresponds to Year 8 of Ramesses III in the Medinet Habu inscriptions — the best-known Egyptian account of conflict with the coalition conventionally labelled the Sea Peoples. Cline deploys the date as a narrative anchor for a multi-decadal process, not as the year of a discrete event. The protocol treats the collapse as a **transition interval**, provisionally 1225–1130 BCE, whose boundaries are parameters estimated from a Bayesian chronological model rather than assumed.

> CORRECTION ENTERED · THE BRIEF NAMES THE WRONG OBSTACLE
>
> **The Hallstatt plateau is irrelevant to this period.** It spans approximately **800–400 cal BC**: determinations near 2450 BP calibrate across that entire four-century span regardless of measurement precision, because atmospheric ^{14}C change flattens the calibration curve there. It sits two to six centuries *after* the study window.

The genuine chronological hazards at 1250–1100 BCE are different in kind, and §2.3 is built around them:

1. **Calibration-curve structure.** IntCal20 is comparatively well-behaved here, but twelfth–eleventh century wiggles produce multi-modal calibrated distributions routinely spanning 60–120 years at 95.4% for single determinations — *coarser than the entire causal sequence under investigation.*
2. **Regional and growing-season offsets.** Demonstrated offsets between Northern Hemisphere calibration data and Eastern Mediterranean samples run to a few decades: negligible for most prehistory, decisive for ordering events inside a century.
3. **Dating-target ambiguity.** A radiocarbon date dates a plant's death, not a destruction. Old wood, curated heirlooms, and residual charcoal all displace the date from the event.

4. **The Aegean–Levantine interlock.** Mycenaean LH IIIB/IIIC, Cypriot LC IIC/IIIA, Levantine LB IIB/Iron IA, and Egyptian regnal chronology are cross-linked by ceramic synchronisms. An error in one link propagates through the grid, so synchronisms are modelled as **uncertain priors, never fixed pegs**.

The epistemic contract

- **Non-fabrication.** No site, text, tablet, core, ingot, or study is cited unless verifiable in the published record. Where an evidence class is required but no specific dataset is confirmed, the document specifies the evidence *type*, acceptance criteria, and sampling requirement instead, marked `[DATA REQUIREMENT]`.
- **Separation of layers.** Evidence, model, and interpretation stay in separate registers. A modelled quantity is never reported alongside an observed one without an explicit marker.
- **Pre-registration.** The analysis plan, falsification criteria, and stopping rules are fixed before data compilation. Deviations go in an amendment register, not silently into the method.
- **Adversarial review.** Each phase is reviewed by a named scholar on record as sceptical of the hypothesis.

PHASE 1

Theoretical Framework and Literature Synthesis

Establish the trajectory from monocausal invasion narrative to systems explanation — and translate each position into a *formal object* the Phase 3 model can instantiate. Synthesis that does not terminate in a formalizable claim is not admitted.

The historiographical arc

The literature falls into five broad, partially overlapping generations. The protocol characterizes each by its **causal topology**, because that is what determines whether the position can be tested by network methods at all.

GENERATION	CAUSAL TOPOLOGY	FORMAL REPRESENTATION	TESTABLE IMPLICATION
I · Migrationist late 19th – mid 20th c.	Single exogenous agent → many simultaneous failures. Read from Medinet Habu and the Great Karnak Inscription as reportage.	Exogenous node deletion, spatially contiguous and temporally near-simultaneous, on a causally inert graph	Destruction dates cluster along a coherent front; incidence **uncorrelated** with network position once coastal exposure is controlled
II · Systems collapse Renfrew 1979 · Tainter 1988	Internal structure → failure under generic stress. Explanatory weight moves from the external agent to the society's architecture.	State-dependent failure probability rising with an internal complexity parameter	Failure severity scales with pre-collapse specialization and administrative elaboration, and is **insensitive to which shock** is applied
III · Perfect storm Cline 2014/2021, 2024	Conjunction of stressors, each insufficient alone: drought, famine, seismicity, rebellion, invasion, severed trade.	Multiplicative or threshold interaction among stressors	Interaction terms carry significant weight; each stressor's conditional effect materially exceeds its unconditional effect
IV · Formal networks Knappett, Evans & Rivers	Exogenous shock → topologically structured propagation → spatially patterned, temporally lagged failure.	Perturbation of an explicit spatial network model; consequence read from the model, not from destruction layers	Which centres matter, and what happens when one is removed, is predicted *a priori* and checked against the ground
V · Revisionist Millek 2021–2023	The destruction horizon is thinner, later, more staggered and less trade-terminating than the standard narrative asserts.	Not a model but a **measurement correction** applied to the outcome variable	If many “destructions” are misidentified, the pattern to be explained is partly an artefact of excavation and publication practice

Generation III states that stressors interacted; it does not specify the medium of interaction, nor predict which polities fail. Naming that medium as the exchange network, and making the claim quantitative, is where this study begins.

MANDATORY COUNTERWEIGHT · NOT A CITATION BUT A PROCESSING STAGE

Millek's re-examination of claimed destruction layers is the single most important control on this project. A protocol that reads only the collapse literature will overfit to collapse. The re-audit is therefore built into the data pipeline at **§2.4**, before any site-level “failure” enters the model.

The translation: from historiographical claim to mathematical object

This is the pivotal deliverable of Phase 1. Each narrative claim maps to a formal operator so Phase 3 can implement it — and, crucially, so that rival claims produce *different* observable signatures.

HISTORIOGRAPHICAL CLAIM	FORMAL OBJECT	PERTURBATION OPERATOR	DISCRIMINATING OBSERVA
“The Sea Peoples destroyed the palaces”	Exogenous, spatially correlated node deletion on a causally inert graph	Delete coastal nodes over a short interval Δt; topology plays no role	Incidence predicted by coastal exposure alone; **no** residual effe centrality
“Drought caused famine and unrest”	Reduction in node-level agricultural carrying capacity	$\pi_i \to (1-\delta_i)\cdot\pi_i$, with δ_i drawn from the hydroclimate field	Failure severity tracks the *local* hydroclimate gradient, with short
“Earthquake storms shattered the palaces”	Temporally clustered, tectonically constrained node damage	Degrade nodes along active faults on a clustered point process	Damage confined to seismogeni zones; failure order follows fault geometry, not trade rank
“Trade routes were cut”	Edge deletion, targeted or random	Remove edges by weight rank, by geography, or at random	Loss of connectivity precedes lo of settlement; downstream node fail before upstream
“The system was hypercoherent”	High E_{glob}, low redundancy, heavy-tailed degrees	None — a *property of the intact graph*, measured against null models	High efficiency *and* low robustne relative to degree-preserving an spatial nulls
“A cascading failure occurred”	Load-redistribution dynamics on a weighted graph	Motter–Lai: $C_i = (1+\alpha)\cdot L_i(0)$; failures redistribute load	Failure times **ordered along network paths**, lag proportional distance from the seed
“The collapse was a phase transition”	Percolation transition in the supply-satisfaction functional	Sweep removal fraction f; locate the critical point	Sharp drop in the order paramet susceptibility χ peaks; finite-siz scaling consistent
“Interdependence amplified failure”	Coupled multiplex with inter-layer dependency	Buldyrev-type iterative cascade across metals, grain, political layers	Transition is **first-order** rather th continuous — the signature of interdependence
“The Sea Peoples were a symptom”	Endogenous generation of mobile actors by prior stress	Migration intensity is an *output* of the model, not an input	Attestations of raiding **post-date** supply and hydroclimate stress i the source regions

THE WAGER, AND HOW IT FAILS

The last row is the study's central historiographical claim, and it is falsifiable cleanly: **if the earliest robust attestations of Sea Peoples activity precede the earliest robust indicators of network and climate stress, the symptom hypothesis fails** — independently of anything the network model produces.

Foundational positions on the environmental and economic drivers

Five verified anchors, chosen because each supplies a methodologically *independent* proxy stream rather than a restatement of the same evidence.

ANCHOR	STREAM	WHAT IT CONTRIBUTES	HOW THE PROTOCOL USES IT
Langgut, Finkelstein & Litt 2013 *Tel Aviv* 40(2)	Palynology	Sea of Galilee core; dry episode at the Late Bronze–Iron transition in the southern Levant	Primary terrestrial hydroclimate record for the southern Levant
Kaniewski et al. 2013 *PLOS ONE* 8(8)	Palynology	Larnaca Salt Lake, Cyprus; a ~300-year drought episode from the late 13th c. BCE	Cyprus hydroclimate; explicitly frames Sea Peoples as consequence
Drake 2012 *J. Archaeol. Sci.* 39(6)	Marine	Sea-surface temperature and isotope evidence for reduced precipitation	Independent *marine* axis, breaking pollen's land-use circularity
Nur & Cline 2000 *J. Archaeol. Sci.* 27(1)	Seismicity	The "earthquake storm" hypothesis — clustered events releasing strain in sequence	Treated as a **hypothesis to be tested**, not an established chronology (§2.2)
Powell et al. 2022 *Science Advances* 8(48)	Metals	Tin-isotope provenancing of the Uluburun tin: remote sources, small-scale exchange	**Load-bearing.** Lengthens the tin chain *and* polycentrifies its sources — opposing effects the model must resolve

A published response contesting Powell et al. 2022 exists and is cited alongside it. Tin provenance is treated as **contested**, and the contestation is propagated as a prior over source configurations rather than settled by fiat.

Competing hypotheses and their discriminating predictions

The study is a model comparison, so the rivals must be stated in advance with predictions that *differ*. A design in which every hypothesis predicts the same observations is not a test.

INTERACTIVE · PHASE 1 　　6 OF 9 OBSERVABLES SEPARATE THEM

Hypothesis discriminator

H_4 · Shock × brittle topology — Shocks of historically plausible magnitude collapse the empirical topology and not counterfactual topologies of equal size and volume. The study hypothesis.

H_1 · Exogenous shock only — Climate and seismicity fully determine failure; topology is causally inert.

OBSERVABLE	H_4	H_1
Failure tracks network centrality, after controls	+	−
Failure severity tracks the local hydroclimate gradient	+	+
Damage confined to seismogenic zones	+	+
Failure times ordered along network paths	+	−
Aggregate decline is discontinuous (first-order)	+	−
Random perturbation reproduces the pattern equally well	−	−
Same shock collapses the redundancy-augmented counterfactual	−	+
Sea Peoples attestation precedes stress indicators	−	·
Survivors predicted by low coupling + high substitution	+	−

Highlighted rows are the observables that separate the two hypotheses. + predicted present · − predicted absent · · the hypothesis is silent. Where two models predict identically on every row, no amount of evidence distinguishes them and one must be reformulated or dropped.

THE PRIMARY OUTCOME IS SURVIVAL, NOT COLLAPSE

The literature is saturated with explanations of why centres fell; few explain why Egypt contracted without collapsing, why several Cypriot and northern Levantine coastal centres show substantial continuity, or why Assyria's trajectory differs. **A model that predicts everyone's failure equally well predicts nothing.** Differential survival is therefore the discriminating outcome of this study, and a model that predicts collapse well and survival badly has learned the shock rather than the structure.

Phase 1 deliverables

- **D1.1** — Annotated historiographical review (~12,000 words), organized by causal topology rather than chronology.
- **D1.2** — The claim→operator table in machine-readable form, each row bound to the Phase 3 module implementing it.
- **D1.3** — Pre-registered hypothesis set with prior model probabilities elicited by structured expert elicitation (Cooke's classical method, with calibration questions weighting each expert).
- **D1.4** — A register of **inert claims**: assertions in the literature that cannot generate a differential prediction, with reasons. Published, not suppressed.

PHASE 2

Data Architecture and Multiproxy Normalization

Build one uncertainty-bearing evidence base from datasets that differ in physical medium, temporal resolution, spatial support, and error structure — without manufacturing agreement between them.

> **GOVERNING PRINCIPLE**
>
> Proxies are **never** aligned by assumption. Every correlation between two proxy streams must survive a test in which the alignment itself is treated as an unknown with a prior.

The historical universe

The universe is defined by **participation in the exchange system**, not by modern geography. A site enters if it satisfies at least one of: documented possession of imports traceable to another region within the universe; attestation in an inter-polity textual corpus; or production or transhipment of a commodity attested in circulation elsewhere.

Core spheres are the Mycenaean Aegean, Crete, the Hittite empire and its Anatolian and North Syrian dependencies, Cyprus (Alashiya), the Levantine coast and interior, and Egypt. Three extensions are mandatory — without them the tin question cannot be posed at all:

- **Mesopotamia and the Middle Euphrates**, as the overland conduit for eastern tin.
- **Central Asian and Anatolian tin sources** implicated by tin-isotope provenancing.
- **Central Mediterranean nodes** — Sardinia, Sicily, southern Italy — where oxhide ingots and Aegean-type material occur.

WINDOW	INTERVAL	ROLE IN THE DESIGN
Baseline	1400–1250 BCE	Characterizes the network at maturity; supplies the normal-operating distribution against which anomalies are defined. **Not optional** — a brittleness claim is meaningless without a prior state to be brittle relative to.
Transition	1250–1100 BCE	The study window.
Aftermath	1100–1000 BCE	Tests recovery, substitution (notably the ferrous transition), and reconfiguration; supplies the differential-survival outcome.

Evidence tiers

Every datum carries a tier label, and tier is a covariate in every downstream model.

TIER	DEFINITION	EXAMPLE	TREATMENT
A	Directly measured, independently replicable, with published analytical uncertainty	Isotope ratio with stated 2σ; ^{14}C determination with lab code and $\delta^{13}C$	Full weight
B	Directly observed but interpretation-dependent	Destruction layer with published section drawing	Full weight on the observation, modelled uncertainty on the interpretation
C	Reported without full supporting documentation	Destruction asserted in a preliminary report	Down-weighted; flagged for re-audit
D	Textual attestation	Ugaritic letter; Amarna letter	Evidence of *claims made by ancient actors* — never directly event data
E	Synthetic or secondary	Figure redrawn from a synthesis	Excluded from analysis; narrative only

Data streams and acceptance criteria

◆ Stream P — Palynological and terrestrial hydroclimate

A core is admitted only if it has a published age–depth model built from **≥ 5 dated horizons** within or bracketing 1600–900 BCE with a stated method and uncertainty envelope; sampling resolution **≤ 50 years per sample** across the transition; published raw counts by taxon, not only an interpretive curve; and an explicit strategy for discriminating aridity from clearance or abandoned orchard husbandry.

> **MANDATORY HETEROGENEITY TEST**
>
> The Aegean hydroclimate literature reports *divergent* regional patterns. The protocol therefore does not fit a single “Eastern Mediterranean drought.” It estimates a **spatially varying anomaly field with explicit between-record disagreement**, and carries that disagreement into the network model as parameter uncertainty. A drought signal that exists only after averaging away regional divergence is treated as an artefact of averaging.

◆ Stream S — Speleothem, marine, and independent physical proxies

Included to break the circularity in Stream P, since pollen responds to human land use as well as to climate. Requirements: U–Th chronology with published errors, $\delta^{18}O$ and $\delta^{13}C$ series, and a stated hydrological interpretation for the specific cave system. Agreement between streams is a *finding*; disagreement is *data*, not noise to be smoothed.

◆ Stream E — Archaeoseismological

The weakest-evidenced and most over-interpreted stream in the collapse literature, and treated with corresponding severity. Every candidate event is scored on a formal quality index; **admission requires ≥ 6 of 10, with a non-zero score on damage typology and on alternative-cause exclusion.**

CRITERION	SCORE 0	SCORE 1	SCORE 2
Damage typology	Fire or collapse only	Directional wall collapse	Diagnostic: rotated or displaced masonry, tilted walls, chevron fractures, ground-rupture offset
Site-effect control	None	Qualitative	Modelled local amplification / geotechnical assessment
Geological corroboration	None	Regional palaeoseismic literature	Dated palaeoseismic trench or offset feature in the region
Chronological control	Ceramic phase only	One ^{14}C date	Bayesian model with multiple dates
Alternative-cause exclusion	Not addressed	Discussed	Systematically excluded: siege, subsidence, structural failure, post-depositional

Events below threshold enter a shadow register used only in sensitivity analysis. The earthquake-storm hypothesis is then tested as a **point-process hypothesis** — a conditional-intensity test against a homogeneous Poisson null at the region's long-term seismicity rate — not as a narrative assessment.

◆ Stream M — Archaeometallurgical and provenance

The purpose is to reconstruct *flow*, not merely presence. Lead-isotope analysis is adopted with its standard cautions promoted to first-class model components rather than footnotes:

- ore-field isotopic fields overlap, so provenance is a **posterior distribution over sources**, never a point assignment;
- recycling and mixing shift composition along mixing lines, so mixing is modelled explicitly as a two- or three-endmember problem with unknown proportions;
- reference-database coverage is incomplete, so an “unmatched” result is evidence of a gap in the reference set as much as of an exotic source;
- trace-element and, where available, copper-isotope data are required as an independent check before any provenance claim is admitted.

ULUBURUN IS CALIBRATION, NOT UNIVERSE

The wreck carried roughly **10 tonnes of copper and 1 tonne of tin** — a 10:1 ratio matching the alloy proportion of standard tin bronze. It is a single cargo and cannot stand as a sample of the trade. It is used for exactly three purposes: to calibrate the copper:tin ratio moving as a bundled consignment; to establish the compositional diversity of one shipment; and as a **taphonomic anchor** — goods in transit, uncontaminated by the deposition biases of settlement contexts.

◆ Stream T — Textual

Three corpora supply the political-dependency layer and the only direct evidence of contemporaneous perception: the Amarna correspondence (EA 1–382, Moran 1992), the Hittite state archives from Boğazköy/Ḫattuša, and the Ugaritic and Akkadian correspondence from Ras Shamra.

HANDLING RULE

A letter reporting a grain shortage is evidence that a shortage was *asserted by an interested party in a specific rhetorical context*. It is coded as `{sender, recipient, date-range, commodity, direction of request, rhetorical register, corroboration status}` and never entered as a measured famine. Each cited tablet carries its siglum, publication reference, and a stated confidence in its date and context — several of the most-quoted texts have contested find-contexts. **No text is quoted from a secondary synthesis.**

Chronological normalization

The hardest problem in the design. The approach is to abandon absolute placement: **do not attempt to put events on a calendar; estimate the joint posterior over event orderings and intervals.**

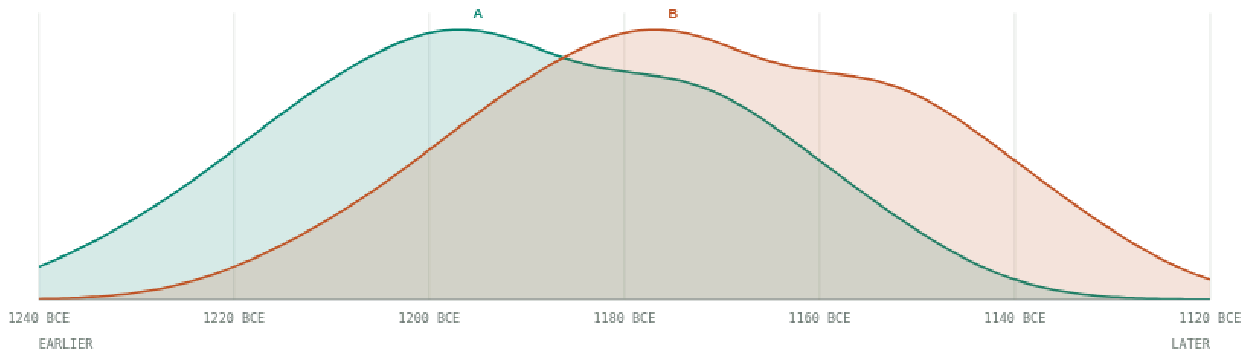
All determinations are modelled in OxCal against IntCal20 with `Sequence` / `Phase` / `Boundary` structures encoding stratigraphy; general and charcoal outlier models on every determination; a jointly estimated regional offset parameter `Δ_R` with a prior informed by the demonstrated Mediterranean offsets; ceramic synchronisms as informative priors with explicit uncertainty; and `KDE_Plot` so that apparent clustering can be distinguished from the sum of individually wide calibrated distributions.

The ordering test — the actual inferential workhorse

Causal claims here are claims about order and lag. For every pair of events the protocol computes, directly from the joint chronology, the posterior probability `P(t_A < t_B | D)` and the posterior distribution of the lag `Δ_AB`.

INTERACTIVE · PHASE 2 ADMISSIBILITY THRESHOLD P ≥ 0.90

Can we say which happened first?



Event A — destruction at an Aegean centre   Event B — destruction at a Levantine centre

P(A BEFORE B)	MEDIAN GAP	A · 95.4% SPAN	VERDICT
0.72	**20 yr**	**88 yr**	● UNRESOLVED

Drag the centres apart, or tighten precision, until the ordering becomes admissible. With realistic wiggles and ±25-year determinations, two events **20 years apart cannot be ordered at all** — which disqualifies a large share of the sequences the narrative literature relies on. Curves are illustrative of calibrated posterior shape, not calibrations of specific samples.

Coping with irreducible imprecision

1. **Coarsening.** Aggregate to 25-year bins and accept the loss of power rather than manufacture precision.
2. **Wiggle-matching.** Where sequences of dated samples exist, wiggle-matching can reach decadal resolution — so sampling programmes that would yield such sequences are prioritized (§5.4).
3. **Order-free tests.** Some predictions need no ordering at all. That failure severity correlates with centrality after controlling for exposure is a cross-sectional test, and survives total chronological failure. **The design deliberately front-loads such tests** so the study still discriminates H_1, H_3 and H_4 in the worst chronological case.

The destruction-horizon re-audit

Before any site-level failure enters the model, its destruction claim is re-derived from primary excavation documentation: is there a published section or plan; is burning architectural or localized to a kiln or hearth; is the assemblage consistent with sudden abandonment or planned departure; are trauma indicators published or merely asserted; is the deposit dated independently or by assumed correlation with the collapse horizon; and is there immediate reoccupation?

`destroyed (violent)` `destroyed (cause indeterminate)` `burnt (localized)`

`abandoned (gradual)` `abandoned (rapid, non-violent)` `continuity` `insufficient documentation`

Outcome is a **categorical response variable with measurement error**, not a binary. Circular dating — a layer dated to 1200 BCE *because* it is a destruction layer — is flagged, and the site is excluded from chronological tests while remaining in cross-sectional ones.

Survivorship, selection, and the causal firewall

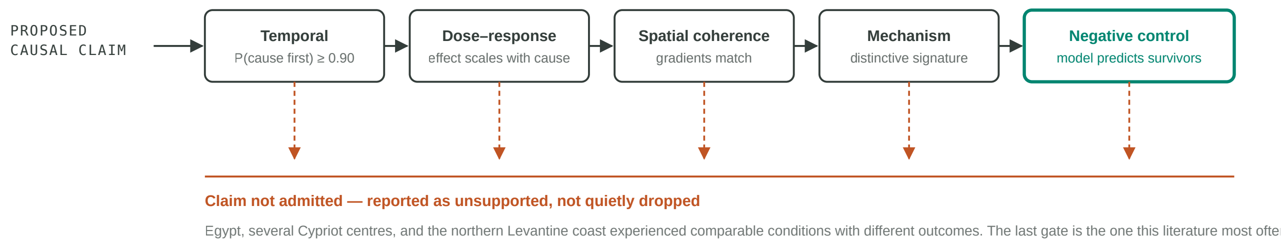
Four distinct biases, each with a distinct correction. Conflating them is the standard error in this literature.

BIAS	MECHANISM	WHICH HYPOTHESIS IT FLATTERS	CORRECTION
Excavation intensity	Large, rich, famous sites are excavated, published, and dated more	Any account centred on major palaces	Excavation-intensity covariate in every model; systematic survey data as a second, differently-biased sampling frame
Destruction preservation	Catastrophic burning preserves floor assemblages and bakes tablets; gradual abandonment leaves swept floors	**H_2 (invasion)** — the record over-represents violent endings	Model `P(detect | ending type)` explicitly; inverse-probability weighting; report weighted and unweighted
Archive survival	The Ugaritic crisis letters survive *because* Ugarit burned. The textual record of the collapse is conditioned on the collapse.	Any reading of the texts as a neutral account of causes	Texts may generate and characterize hypotheses but may **never** be the outcome variable in a causal test; comparable archive-less sites coded missing-not-at-random
Chronological resolution	Better-dated sites are better-funded, hence larger, hence more central	Naively, the network hypothesis itself	Dating quality as covariate; matched analysis pairing high- and low-centrality sites of comparable dating quality

COLLIDER STRATIFICATION · THE MOST SEDUCTIVE TRAP IN THE CORPUS

Selecting on the outcome and then reading the surviving texts as an unbiased account of causes is a textbook error. It is the reason the archive-survival row above is the one gate this protocol will not relax for any argument.

FIGURE 1 · THE CAUSAL FIREWALL



All five gates, in order. No causal claim is admitted unless it clears every one. A model that reproduces the failures but cannot reproduce the survivals is rejected at the final gate regardless of how well it fits everything upstream.

Data infrastructure

A PostgreSQL + PostGIS schema in which **no uncertain quantity is stored as a scalar**: dates are posterior distributions, provenance is a distribution over sources, edge weights are distributions. Every row references a source and carries a tier. Git-tracked ETL, every figure regenerable by one command, releases DOI-minted, CIDOC CRM alignment for interoperability, coordinates deliberately fuzzed for sites at looting risk, and code and derived data released under open licence at publication.

PHASE 3

Systems Modelling and Network Vulnerability Testing

Build a spatially and temporally explicit multilayer model, subject it to historically constrained perturbations, and identify the conditions under which localized failure becomes systemic. The model is built to be broken — and to be broken in ways that could have failed.

DESIGN COMMITMENT

Every result is reported against degree-preserving, spatial, and volume-matched null models. **A cascade that also occurs in the null models is not evidence for H_4.**

Formal specification

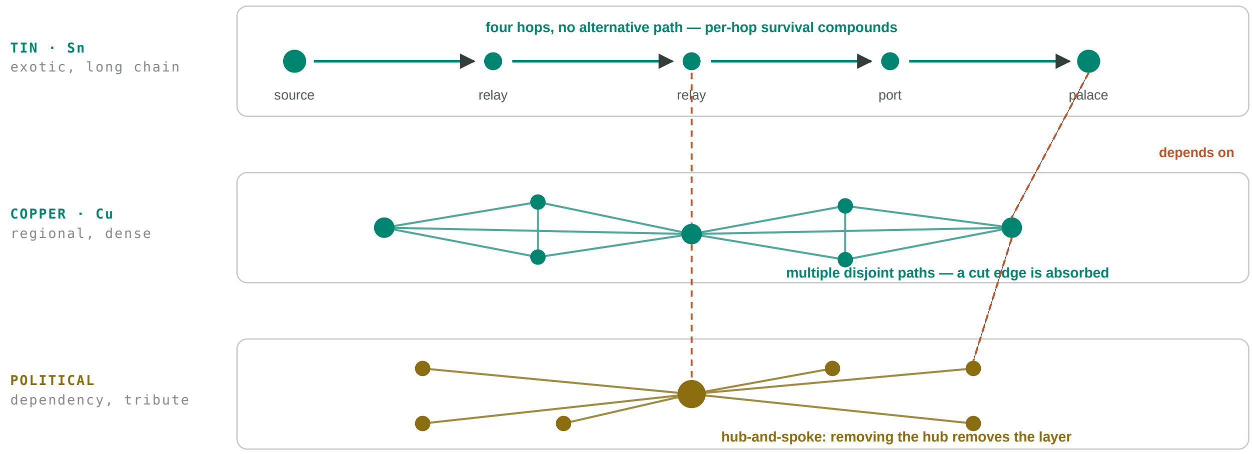
The system is a temporal, weighted, directed multiplex:

```
M(t) = ( V, {E^(α)(t)}, {W^(α)(t)}, D(t) )   for layers α ∈ L

  V        nodes: polities and exchange places, |V| ≈ 120–250
  L        layers: Cu · Sn · Au/Ag · grain · finished goods · political
  E^(α)    directed edges in layer α at time t
  W^(α)    edge weights: expected commodity flow, or dependency intensity
  D(t)     inter-layer dependency: node i in layer α given its state in β
```

Separating commodities into layers is not cosmetic. Copper and tin have different source geographies, different substitutability, and different route structures; collapsing them into one “trade” graph destroys precisely the asymmetry the study is testing.

FIGURE 2 · WHY THE MODEL IS A MULTIPLEX



The same node behaves differently in each layer, and the layers depend on each other. Tin arrives through a long chain with no alternative path; copper circulates in a redundant regional mesh; political dependency is hub-and-spoke. The dashed links are what convert a gradual decline into a discontinuous one: a palace that loses its metals function stops performing its grain-redistribution function too.

Nodes, edges, weights

Node attributes are carried as distributions, never scalars: location, settlement extent, administrative complexity c_i (archives, sealing practice, standardized weights, storage architecture, craft specialization), agricultural capacity π_i, storage buffer σ_i converted to person-months of supply, metallurgical dependency μ_i, and political dependency ρ_i.

> **NO NODE IS PRIVILEGED**
>
> Ḫattuša, Mycenae, Ugarit and Pi-Ramesses enter as ordinary nodes. Their prominence in the narrative literature is itself a bias to be controlled: prominence correlates with excavation intensity, and excavation intensity is already a covariate.

An edge requires positive evidence of directional flow, from three independent classes combined by evidence synthesis rather than union — **material provenance** (discounted by the provenance posterior and the number of plausible intermediaries), **textual attestation** (full strength on the political layer, weak on commodity layers, because gift-exchange is a poor proxy for bulk volume), and **route feasibility** (least-cost paths over sea accounting for winds, currents, seasonality and the coastal-tramping character of Bronze Age sailing, and over land across terrain cost surfaces).

```
P(e_ij^(α) = 1 | evidence)  ∝  P_prov(·) · P_text(·) · P_route(·)^θ

Volume weight      w_ij^(α)   estimated flow, order-of-magnitude posterior
Dependency weight  d_ij^(α) = w_ij^(α) / Σ_k w_kj^(α)
```

`d` is what matters for cascade dynamics. A node may carry small absolute volume and still be structurally critical if it is a sole supplier.

> **ANACHRONISM GUARD ON THE AMARNA LAYER**
>
> The Amarna archive dates predominantly to the **mid-14th century** — roughly a century before the transition window. It parameterizes *baseline* political topology and the persistence of dependency structures, with an explicit decay model for extrapolation forward. Using it directly as a 1200 BCE snapshot is forbidden. Coding is done in duplicate with inter-coder reliability reported (Krippendorff's $\alpha \geq 0.80$ for admission).

Metrics and cascade dynamics

Definitions are stated explicitly so results are reproducible.

```
Strength            s_i^(α) = Σ_j w_ij^(α)
Betweenness         B_i     = Σ_{s≠i≠t} σ_st(i) / σ_st
Supply criticality  SC_i^(α)= Σ_{j≠i} d_ij^(α) · ( 1 − r_j^(α) )

Percolation (Cohen et al. 2000)
    κ   = ⟨k²⟩ / ⟨k⟩            giant component exists iff κ > 2
    f_c = 1 − 1/(κ₀ − 1)        critical fraction under random removal

Global efficiency   E_glob  = 1/(N(N−1)) · Σ_{i≠j} 1/d_ij
Robustness          R       = (1/N) · Σ_{q=1..N} S(q)        R ∈ (0, 0.5]
Brittleness index   B       = 1 − 2R                          B ∈ [0, 1)

Cascade (Motter–Lai 2002)
    C_i = (1 + α) · L_i(0)      node fails at step τ if L_i(τ) > C_i
```

> **ROBUST YET FRAGILE — THE FORMAL CONTENT OF THE THESIS**
>
> For heavy-tailed degree distributions with $2 < \gamma \leq 3$, the second moment diverges, $\kappa_0 \to \infty$, and $f_c \to 1$. The network becomes extremely tolerant of *random* failure and acutely vulnerable to *targeted or correlated* failure. Measuring γ is therefore a primary Phase 3 result — estimated by the Clauset–Shalizi–Newman maximum-likelihood procedure and tested against log-normal and stretched-exponential alternatives. A heavy tail is a hypothesis here, not an assumption; the sample is small enough that naive log-log fitting would mislead.

A IS MEASURED, NOT TUNED

The tolerance parameter `α` is the fraction of spare capacity a polity holds above ordinary throughput — granary volume, standing surplus, redundant shipping, alternative suppliers. It is not free. A command economy optimizing for extraction and display consumption drives it down. **The over-optimization hypothesis just is the claim that the system ran at low α**, so α is estimated from excavated storage capacity against estimated consumption and reported with its uncertainty as a headline result.

Interdependent cascade. Where a node's function in one layer depends on its state in another — a palace cannot administer grain redistribution once its metallurgical and prestige economy has failed, and cannot sustain that economy without agricultural surplus — the iterative failure process between coupled layers produces a **first-order, discontinuous percolation transition**, in contrast to the continuous transition of a single network. Interdependence converts graceful degradation into abrupt collapse.

This is the sharpest signature available. Under H_1 decline should be roughly proportional to shock magnitude. Under H_4 with interdependence there is a threshold below which the system absorbs the shock and above which it disintegrates.

INTERACTIVE · PHASE 3 · THE CORE MECHANISM

Where localized failure becomes systemic

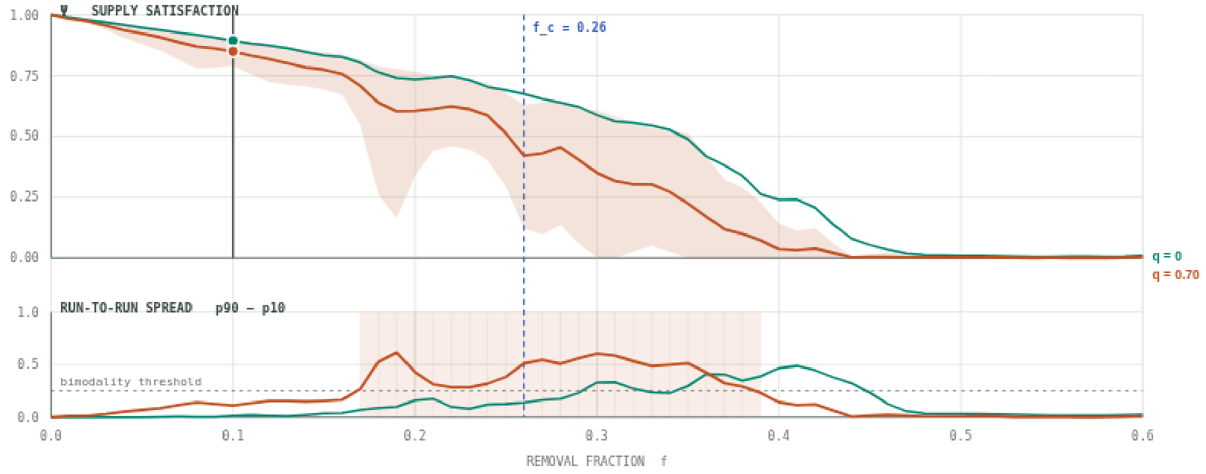
BIMODAL NEAR THRESHOLD — THE WARNING INTERVAL CLOSES

Supplied — within five hops of a tin source

Cut off — still standing, but the tin no longer arrives

Removed by the shock or the cascade

Ψ · SUPPLY SATISFIED	F_C · HALF UNSUPPLIED	BIMODAL ZONE WIDTH	TRANSITION ORDER
0.70	0.26	Δf 0.22	● DISCONTINUOUS



Independent layers (q = 0) Coupled at the current q — shaded band is the 10th–90th percentile across 16 runs

f c — removal fraction at which half the system is unsupplied

Set coupling to zero and the curve bends gently: a **second-order** transition, where decline is roughly proportional to damage. Raise coupling and the same shock schedule produces a cliff — a **first-order** transition, where the warning interval between “stressed” and “failed” shrinks toward nothing. Switching from random removal to hubs-first collapses the network at a fraction of the damage, which is the robust-yet-fragile asymmetry made visible. The lower panel is the run-to-run spread. In a finite system a discontinuous transition does not show up as a vertical drop in the average — averaging smooths it away — it shows up as **bimodality**: near the threshold some runs survive intact and others are wiped out, so the ensemble spread balloons. A wide bimodal zone is the finite-size signature §3.6 prescribes. Simulated on a synthetic two-layer network with the qualitative properties described in §3.1 — an exchange layer and a hub-and-spoke political layer, coupled at strength q, with load redistributed on a degree proxy. It illustrates the mechanism; it is not a reconstruction of the Bronze Age.

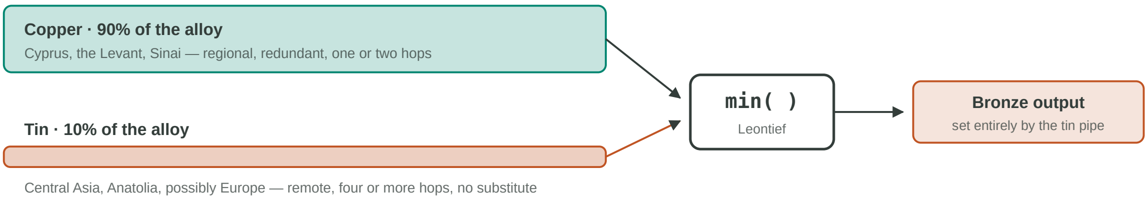
The tin bottleneck

The metals system is modelled explicitly because it is where the argument is most testable. Tin bronze at the standard ~10% Sn requires inputs in near-fixed proportion — a Leontief technology with negligible short-run substitution:

```
b_i = min( cu_i / (1 − θ) , sn_i / θ )        θ ≈ 0.10
```

Copper was available within the Eastern Mediterranean, Cypriot deposits above all. Tin had no abundant, securely exploited regional source at the scale of demand — small Anatolian deposits notwithstanding — and travelled great distances from Central Asian, Anatolian, and possibly European sources. The binding term is overwhelmingly `sn_i / θ`.

FIGURE 3 · WHY TIN, AND NOT COPPER, IS THE CONSTRAINT



Widening the copper pipe changes nothing. Elasticity of output w.r.t. tin ≈ 1; w.r.t. copper ≈ 0.

Criticality is a function of substitution elasticity and chain length — not of cost share, which is what procurement accounting measures.

The binding constraint is the small input, not the large one. Tin was a tenth of the alloy and a tenth of the cargo mass, and an absolute limit on bronze production. This asymmetry generates a specific archaeological prediction, testable before any network modelling: declining Sn content in bronzes, rising recycling signatures, and increasing use of low-tin or unalloyed copper should appear *before* settlement failure at affected nodes.

THE RESULT THE STUDY MOST NEEDS TO PRODUCE

The Powell et al. 2022 provenancing cuts both ways. Remote sources **lengthen** the chain, adding intermediaries and compounding per-hop failure — exposure rises as `Π p_survive` over hops. But multiple extraction sources feeding the system through small-scale exchange **reduce** source-concentration risk. The two effects act in opposite directions, and the net effect is an empirical question the model must answer rather than assume.

Substitution. The ferrous transition is a substitution response, not merely a technological succession. Substitution capacity `r_j` becomes time-varying, and the aftermath window tests whether the network reconfigured around a commodity with ubiquitous sources — structurally, a shift from a long-chain, high-criticality input to a short-chain, low-criticality one.

Perturbation operators and the experimental matrix

ID	OPERATOR	EMPIRICAL CALIBRATION	FREE PARAMETERS
O1	Hydroclimate stress	`π_i → (1−δ_i(t))·π_i` from the spatially varying anomaly field, carrying its between-record disagreement	Severity scaling; lag from anomaly to yield loss
O2	Seismic node damage	Capacity and administrative function degraded at sites in the admitted seismic register, on the estimated event chronology	Damage-to-function mapping; recovery rate
O3	Edge severance	Maritime and overland edges removed at random, by weight, by betweenness, or in geographic clusters	Fraction removed; duration
O4	Node deletion	Random, degree-targeted, betweenness-targeted, or on the attested destruction schedule from the re-audit	Fraction; schedule

ID	OPERATOR	EMPIRICAL CALIBRATION	FREE PARAMETERS
O5	Demand shock	Reduction in elite prestige-goods demand, modelling loss of legitimacy expenditure	Magnitude
O6	**Endogenous mobility**	`m_i(t) = f(unmet subsistence, neighbour failure)`; displaced groups then act on the network as raiders and migrants	Threshold; mobility rate; predation intensity

O6 IS THE WHOLE ARGUMENT, OPERATIONALIZED

Under H_2 the mobile actors are exogenous inputs. Under H_4 they are *outputs* that then feed back. The two are formally distinguishable: under H_4 the model must reproduce the observed pattern of intrusive material culture and raiding attestation **without any exogenous injection of raiders**, predicting their timing and geography from prior stress. **If an exogenous injection is required to fit the data, H_4's strongest form is falsified.**

FACTOR	LEVELS
Shock set	none · O1 · O2 · O1+O2 · O1+O2+O3 · full (O1–O6)
Network	Empirical reconstruction · degree-preserving rewire · spatially constrained random · volume-matched random · **redundancy-augmented counterfactual** · low-efficiency counterfactual
Tolerance `α`	Empirical posterior · 0.05 · 0.1 · 0.2 · 0.5 · 1.0
Tin configuration	Polycentric (Powell et al.) · single-source concentrated · intermediate, weighted by the provenance-debate prior
Coupling `q`	0 (independent layers) → 1 (full coupling), 11 levels
Substitution `r`	Fixed low · fixed high · time-varying with the ferrous transition

Ensemble ≥ 10^4 realizations per cell, or until Monte Carlo standard error falls below 1% of the outcome range. Latin hypercube sampling over continuous parameters; full seed and configuration logging; every figure regenerable.

THE DECISIVE COMPARISON

The **redundancy-augmented counterfactual** holds constant the number of nodes, the total trade volume, and the geography, while adding alternative paths. If the historical shock schedule collapses the empirical network but *not* this one, the brittleness claim is supported in the strongest available sense: the collapse is attributable to topology *given* the shock — neither to the shock alone nor to the topology alone.

Locating the tipping point

The central question is answered as a critical-phenomena measurement, not a narrative judgement. Rather than the bare largest-component fraction, the order parameter is a *functional* outcome — whether the system is doing its job:

```
Ψ(f) = ( Σ_i min( supply_i(f), demand_i ) ) / ( Σ_i demand_i )      on the bronze layer

χ(f) = Σ'_s s²·n_s / Σ'_s s·n_s      largest cluster excluded; χ peaks at f_c

Finite-size scaling:   f_c(N) = f_c(∞) + a · N^(−1/ν)
```

- **Order of the transition** is discriminated by a discontinuity in Ψ under infinite-size extrapolation, bimodality of Ψ across ensemble members near `f_c`, and hysteresis under a reverse sweep restoring capacity.
- **Finite-size analysis is mandatory.** With `N ≈ 120–250`, apparent sharpness may be an artefact. Any claim of a sharp transition unaccompanied by finite-size analysis is inadmissible.
- **Early-warning signals** — rising autocorrelation, variance and skewness before a bifurcation — are directly measurable in the model. The empirical counterpart (settlement-size variance, import-diversity variance, hoarding frequency, assemblage heterogeneity) is deliberately weaker and is presented as exploratory: archaeological resolution is close to the limit this test needs, and a negative result would be uninformative, which must be said when it is reported.
- **Attribution decomposition.** A Shapley-value decomposition over the factorial design splits the systemic outcome into shock magnitude, topology, and their interaction. H_4 predicts the interaction term carries the largest share — which converts “the shocks interacted with a brittle network” from a metaphor into a number with a confidence interval.

Validation — a model that fits the collapse is worthless unless it could have failed to

1. **Out-of-sample temporal validation.** Fit on 1400–1250 BCE baseline dynamics only; predict 1250–1100. No collapse-window data touches the fitting stage.
2. **Spatial hold-out.** Withhold entire regions — Cyprus, or the southern Levant — fit on the remainder, predict the withheld pattern of failure *and survival*.
3. **Independent-event benchmark.** Run the identical pipeline on the Theran eruption horizon, where the Knappett–Rivers–Evans results provide a published comparator. If it cannot reproduce known results there, it is not trusted here.
4. **Survival prediction** scored by balanced accuracy and Matthews correlation coefficient, because the classes are imbalanced and raw accuracy would mislead.
5. **Adversarial fitting.** An independent team fits the same observations with H_1 and H_2 maximally favoured, on an equal computational budget. Bayes factors are computed between the best H_4 model and the best adversarial model.
6. **Sensitivity.** Global Sobol indices over all parameters; results reported for the full posterior, never a maximum-likelihood point.

7. **Prior sensitivity.** Every substantive conclusion re-run under the advisory panel's most sceptical elicited prior. Conclusions that do not survive are reported as prior-dependent.

PHASE 4

Brittleness and the Limits of Modern Parallels

Convert the network results into a general account of how efficiency-seeking generates fragility — and establish, under strict conditions, what if anything transfers to modern systems.

STANDING WARNING

This phase carries the highest risk in the project. The pull toward a satisfying contemporary moral is strong, the incentives to indulge it are strong, and the resulting literature is largely worthless. A **formal homology test** must be passed before any parallel is asserted, and disanalogies are published with equal prominence. If the test fails, Phase 4 reports that it fails — a legitimate and publishable outcome.

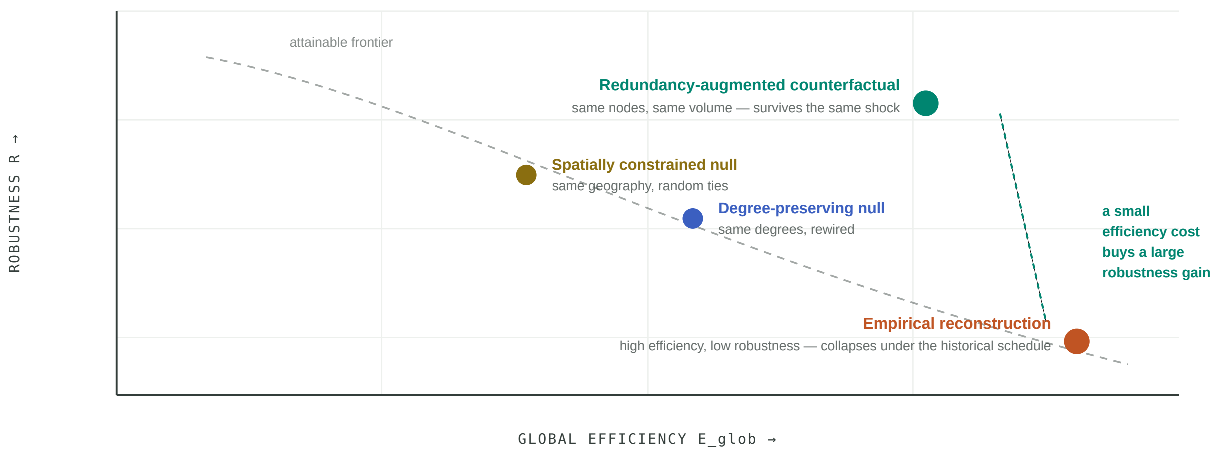
Operationalizing brittleness

Brittleness is not a mood. It is a measurable property of a response function, defined here as a five-component vector so that no single number can hide a trade-off.

COMPONENT	SYMBOL	DEFINITION	MEASURED FROM
Redundancy deficit	$1 - \bar{R}$	Mean edge-disjoint path count, normalized, dependency-weighted	Reconstructed topology
Buffer deficit	$1 - \bar{\alpha}$	Spare capacity relative to throughput	Excavated storage volume vs. estimated consumption
Input criticality	SC	Supply criticality on the binding commodity	Commodity-flow layer
Substitution rigidity	$1 - \bar{r}$	Inability to substitute inputs or suppliers within the shock's timescale	Leontief constraint; source geography
Coupling	q	Strength of inter-layer dependency	Multiplex dependency links

The composite index $B = 1 - 2R$ is the headline scalar; the *vector* is what gets interpreted, because two systems can share a B for entirely different reasons and will fail differently.

FIGURE 4 · THE EFFICIENCY–RESILIENCE FRONTIER



Schematic of the hypothesis, not a result. H_4 predicts the empirical network sits measurably above the nulls in efficiency and below them in robustness. **If it does not, the over-optimization thesis is falsified at its root** — and that is a live risk: a network reconstructed from surviving imports is biased *toward* apparent efficiency, because evidence for a route survives better when the route was heavily used. That bias must be quantified by simulation before this figure is ever filled with real values.

The mechanism, stated so it can be checked

1. **Centralized redistribution suppresses redundancy.** When a palace stops functioning, the redistribution it performed does not degrade — it stops. *Testable:* the sharpest discontinuities occur at sites with the highest administrative-complexity index, and dependent settlements fail at a short, measurable lag behind their palace.
2. **Elite prestige demand rewards long chains.** Value accrues to the exotic, so legitimacy expenditure commits the system to long, thin supply chains — the configuration with the lowest per-hop survival probability.
3. **Specialization raises efficiency and lowers substitutability.**
4. **Buffers are politically expensive.** Stored surplus that is neither consumed nor displayed returns no legitimacy, so under competitive display pressure α is driven down. *Testable, and genuinely surprising if true:* storage capacity relative to estimated catchment yield should **decline** through the thirteenth century at high-complexity sites.
5. **Coupling is a by-product of integration.** The same institutions manage metals, grain and diplomacy — and strong coupling makes the transition discontinuous.

Point 4 is the sharpest empirical hook in Phase 4 and should be prioritized: it is a claim about excavated storage volumes over time, it is measurable, and it could easily come out the other way.

The homology test for modern parallels

No parallel is asserted unless this gate is passed, and the modern comparator must be **named and characterized**

in advance — not left as “globalization.”

#	DIMENSION	LBA MEASUREMENT	MODERN COUNTERPART	PASS CONDITION
1	Degree-distribution class	`γ` by Clauset MLE with model comparison	Same estimator on the modern supplier graph	Same class, overlapping credible intervals
2	Critical-input concentration	`SC` for tin; source-geography HHI	`SC` and HHI for the named critical input	Same order of magnitude
3	Buffer-to-throughput	`α` from storage vs. consumption	Days of inventory; strategic reserve coverage	Same order of magnitude
4	Substitution elasticity	Leontief rigidity of the alloy constraint	Short-run elasticity for the critical input	Both below 0.2 in the relevant horizon
5	Route redundancy	Edge-disjoint paths on the critical layer	Independent logistics corridors	Comparable normalized `R`
6	Cross-layer coupling	`q`	Coupling of logistics, finance, energy, information	Comparable, or explicitly bounded

Scoring: a parallel is asserted only where at least four of six pass and none fails catastrophically. Where fewer pass, the finding is reported as *structural difference* — equally informative, and considerably rarer in the literature.

The disanalogy register — published, not buried

DIFFERENCE	WHY IT MATTERS
Information velocity	Bronze Age disruption was detectable only after it had propagated. Instant signalling enables faster correction *and* faster contagion; the sign of this difference is genuinely ambiguous and must not be asserted.
Price signals	Palatial redistribution and gift-exchange are not price-clearing markets. Modern chains reallocate through prices — a powerful adaptive mechanism entirely absent from the ancient case. **The single largest disanalogy.**
Capital and technological substitution	Modern systems build new capacity in months to years. A Bronze Age polity could not conjure a new tin source.
State capacity	Strategic reserves, industrial policy, and coordinated crisis response have no close ancient analogue.
Energy basis	Solar-agricultural versus fossil, nuclear and renewable energy produce different constraint structures entirely.
Scale and demographic buffer	Populations, urbanization, and mobility differ by orders of magnitude.
Data asymmetry	Modern supply-chain data is dense and directly measured; LBA data is sparse, indirect, and centuries-averaged. Confidence intervals must reflect that rather than being presented on a

DIFFERENCE	WHY IT MATTERS
	common footing.

What actually transfers, if the gate is passed

The transferable content is **structural and conditional**, never predictive.

1. **Efficiency and robustness trade off, and the trade-off is invisible in normal operation.** A network optimized for throughput emits no distress signal while conditions stay in-distribution. The absence of failure is not evidence of robustness.
2. **Robust-yet-fragile is a topological property, not a management failure.** Systems that have survived many random shocks may have learned precisely the wrong lesson about their own resilience.
3. **Interdependence converts gradual degradation into discontinuous collapse.** Coupling logistics, finance, energy and information moves the transition from second-order to first-order — the warning interval between “stressed” and “failed” shrinks toward zero.
4. **The binding constraint is the long-chain, low-substitutability input, whatever its share of value.** Criticality is a function of substitution elasticity and chain length, not cost share — exactly the quantity conventional procurement accounting does not measure.
5. **Correlated shocks defeat diversification.** When drought, seismicity and unrest are spatially correlated, apparently independent routes fail together and measured diversification overstates real protection.
6. **The visible agent of collapse may be an output of the system's own stress.** If the endogenous-mobility result holds, attributing collapse to whoever appears at the moment of failure is a category error. *This is the study's contribution to general theory and the claim most at risk of being over-stated; it must travel with the model's limitations attached.*

Conclusion architecture

The concluding argument is written to a fixed structure that makes the strength of each claim visible: **(1)** what the evidence shows, with tiers and uncertainties, independent of any model; **(2)** what the model shows, always paired with the null comparison; **(3)** posterior model probabilities for H_0–H_4 with prior sensitivity; **(4)** what remains unresolved and the specific evidence that would resolve it; **(5)** what generalizes, gated by the homology test; **(6)** what does not.

If the hypothesis is supported, the concluding claim is bounded in this form and no stronger:

> The observed pattern of failure and survival between 1250 and 1100 BCE is better explained by exogenous climatic and seismic shocks propagating through a highly efficient, low-redundancy, strongly coupled exchange network than by those shocks alone, by network structure alone, or by exogenous migration. The mobile groups recorded in Egyptian sources are, on this account, consistent with an output of the failing system rather than its initiating cause — though the textual record is too thin, and its survival too strongly conditioned on the collapse itself, to exclude an independent migratory contribution.

If the hypothesis is not supported, the corresponding statement is written in advance, so the framing cannot drift after the results are seen.

SECTION 05

Governance, Falsification, and Programme Management

Pre-registered before data compilation, with a timestamped hash. Seven conditions reject the study hypothesis; one of them means the study cannot answer the question at all, and the two must never be conflated in reporting.

Falsification criteria

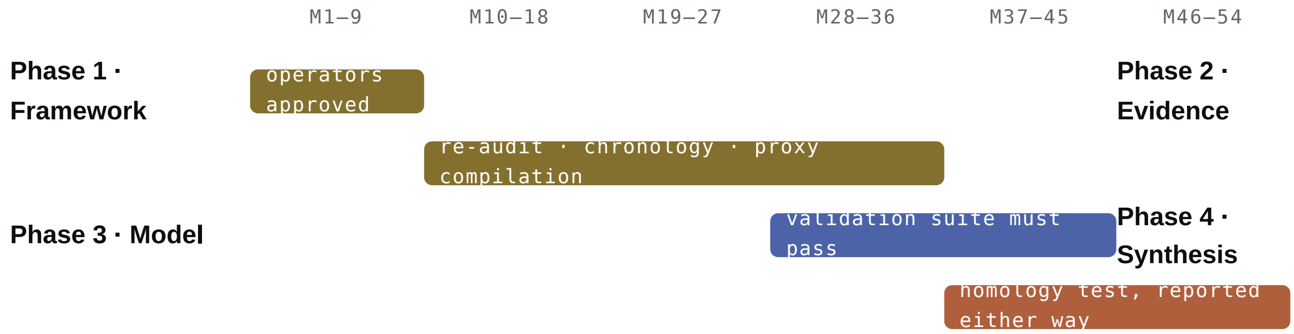
#	H_4 IS REJECTED IF…
F1	The reconstructed network's efficiency and robustness fall within the 90% interval of degree-preserving and spatially constrained nulls — it is not unusually brittle.
F2	Failure incidence and severity show no association with network position after controlling for hydroclimate exposure, seismic exposure, excavation intensity, and dating quality.
F3	The historical shock schedule collapses the redundancy-augmented counterfactual as readily as the empirical network.
F4	Failure times show no path-ordered structure: lag between connected failures is unrelated to network distance, at the resolution the chronology permits.
F5	The earliest robust attestations of Sea Peoples activity **precede** the earliest robust indicators of supply-chain and hydroclimate stress in the source regions.
F6	The re-audit reduces securely attested destructions below the threshold needed to detect the predicted spatial pattern at power ≥ 0.8.
F7	The Bayes factor for H_4 against the best adversarial H_1/H_2 model is below 3.

F6 IS NOT A NEGATIVE RESULT

F6 is a condition under which the study **cannot answer the question**, not one under which the answer is no. If it obtains, the deliverable becomes the audited evidence base plus a power analysis specifying what new excavation and dating would be required — which is a real contribution, and must not be written up as a refutation.

Programme, team, and risk

FIGURE 5 · PHASE STRUCTURE, WITH HARD GATES



Phases overlap deliberately. The Phase 2 → Phase 3 gate is hard: **modelling on unaudited data is prohibited.**

Team. A network scientist with percolation and multilayer expertise; a computational archaeologist; four regional specialists (Aegean, Anatolian, Levantine/Cypriot, Egyptian); a philologist covering Akkadian, Ugaritic and Hittite; a palaeoclimatologist; an archaeometallurgist; a Bayesian chronologist; a data engineer; and a heritage-liaison officer.

ANTI-SILOING REQUIREMENT

Regional specialists see every quantitative result in domain terms before publication and may veto claims that misrepresent their evidence. Modellers see every archaeological interpretation and may flag unfalsifiable claims. **Disagreements that survive are published as disagreements.**

Data acquisition priorities

Ranked by expected reduction in posterior uncertainty:

1. High-resolution, well-dated hydroclimate records for the Anatolian plateau and northern Levant — the largest spatial gaps in the drought field.
2. Dendrochronologically anchored or wiggle-matched sequences from destruction contexts — the only realistic route to decadal ordering.
3. Systematic lead- and tin-isotope programmes on stratified assemblages spanning 1300–1050 BCE, converting provenance snapshots into a time series.
4. Re-excavation or archival re-study of destruction contexts flagged as insufficiently documented.
5. Systematic quantification of storage capacity by phase at palatial sites — the weakest link in the α estimate, on which the sharpest Phase 4 prediction depends.

RISK	LIKELIHOOD	IMPACT	MITIGATION
Destruction corpus collapses under re-audit	Medium	High	Front-load the audit; design order-free tests that survive it
Chronology cannot resolve ordering	High	High	Cross-sectional tests as primary; wiggle-matching programme; report unresolved pairs honestly
Network reconstruction is circular — routes inferred from the same finds used to test the model	High	Critical	Strict separation of edge-construction evidence from outcome evidence; spatial hold-out; simulate the survival bias and quantify its effect on the frontier
Overfitting to a single realization	High	High	Out-of-sample and hold-out validation; Thera benchmark; Sobol sensitivity
Modern-parallel over-claiming	High	High (reputational)	Homology gate; disanalogy register; external review by an economist and a supply-chain specialist
Proxy disagreement dismissed as noise	Medium	High	Between-record disagreement carried as parameter uncertainty, never averaged away

Ethics

Coordinates for sites at looting risk are fuzzed in public releases. Sampling requires national permits and is minimized in favour of legacy-collection reanalysis. Regional scholars and institutions in the countries where the evidence originates are collaborators with authorship, not data providers. Published language avoids framing collapse as civilizational judgement, and avoids deploying ancient migration as commentary on modern migration — a use of this material that is both historically indefensible and politically instrumentalized.

Deliverables

ID	DELIVERABLE
D1	Historiographical synthesis and formal operator table
D2	Audited, versioned, DOI-minted multiproxy evidence base with full uncertainty representation
D3	Destruction-horizon re-audit catalogue — a standalone contribution regardless of the study's outcome
D4	Bayesian chronological model with published posterior orderings and a register of unresolved pairs
D5	Open-source multilayer network and cascade simulation suite, containerized and reproducible
D6	Primary results paper: model comparison and tipping-point analysis

ID	DELIVERABLE
D7	Methods paper: multiproxy normalization and the causal firewall, written for transfer to other collapse studies
D8	Phase 4 synthesis: brittleness, the frontier, and the homology assessment
D9	Public interactive visualization of the network and its failure dynamics

APPENDIX A

Notation

SYMBOL	MEANING	SYMBOL	MEANING
$M(t)$	Temporal multiplex network	E_{glob}	Global efficiency
V, N	Node set; number of nodes	R_{ij}, $\bar{R}$	Edge-disjoint path count; weighted mean
L, α	Layer set; layer index	R	Schneider robustness index
$w_{ij}^{(\alpha)}$	Volume weight, i → j, layer α	B	Brittleness index, 1 − 2R
$d_{ij}^{(\alpha)}$	Dependency weight	L_i, C_i	Node load; node capacity
$\langle k \rangle$, $\langle k^2 \rangle$	Degree moments	α (cascade)	Tolerance, $C_i = (1+\alpha)L_i(0)$
κ	$\langle k^2 \rangle / \langle k \rangle$, Molloy–Reed parameter	q	Inter-layer coupling strength
f, f_c	Removal fraction; critical fraction	$\Psi(f)$	Supply-satisfaction ratio
γ	Degree-distribution tail exponent	$\chi(f)$	Percolation susceptibility
B_i	Betweenness centrality	ν	Finite-size scaling exponent
$SC_i^{(\alpha)}$	Supply criticality	π_i, σ_i, c_i	Agricultural capacity; storage buffer; complexity
$r_j^{(\alpha)}$	Substitution capacity	θ	Tin fraction in bronze (≈ 0.10)
$\delta_i(t)$	Hydroclimate capacity reduction	Δ_R	Regional radiocarbon offset

APPENDIX B

Verified Source Register

Every item was checked against a publisher, journal, or institutional record during protocol preparation. Identifiers are reproduced as published.

ANCHOR MONOGRAPHS AND EDITIONS

- Cline, E.H. *1177 B.C.: The Year Civilization Collapsed.* Princeton UP, 2014; rev. ed. 2021. ISBN 9780691208015 (rev.) · 9780691140896 (1st) · JSTOR j.ctv15r58dw
- Cline, E.H. *After 1177 B.C.: The Survival of Civilizations.* Princeton UP, 2024. ISBN 9780691192130
- Tainter, J.A. *The Collapse of Complex Societies.* Cambridge UP, 1988. ISBN 9780521386739
- Renfrew, C. “Systems Collapse as Social Transformation,” in *Transformations: Mathematical Approaches to Culture Change*, 1979. Academic Press
- Millek, J.M. *Destruction and Its Impact on Ancient Societies at the End of the Bronze Age.* Lockwood Press, 2023. ISBN 9781948488839
- Killebrew, A.E. & Lehmann, G. (eds.) *The Philistines and Other “Sea Peoples” in Text and Archaeology.* SBL, 2013. ISBN 9781589831292
- Leidwanger, J. & Knappett, C. (eds.) *Maritime Networks in the Ancient Mediterranean World.* Cambridge UP, 2018. ISBN 9781108429948
- Knappett, C. (ed.) *Network Analysis in Archaeology.* Oxford UP, 2013.
- Moran, W.L. *The Amarna Letters.* Johns Hopkins UP, 1992. Standard edition, EA 1–382

CLIMATE AND ENVIRONMENT

- Langgut, D., Finkelstein, I. & Litt, T. 2013. “Climate and the Late Bronze Collapse: New Evidence from the Southern Levant.” *Tel Aviv* 40(2): 149–175. DOI 10.1179/033443513X13753505864205
- Kaniewski, D. et al. 2013. “Environmental Roots of the Late Bronze Age Crisis.” *PLOS ONE* 8(8): e71004. DOI 10.1371/journal.pone.0071004
- Drake, B.L. 2012. “The influence of climatic change on the Late Bronze Age Collapse and the Greek Dark Ages.” *Journal of Archaeological Science* 39(6): 1862–1870.
- Finné, M., Woodbridge, J., Labuhn, I. & Roberts, C.N. 2019. “Holocene hydro-climatic variability in the Mediterranean.” *The Holocene*. DOI 10.1177/0959683619826634

- Jacobson, M.J., Seguin, J. & Finné, M. 2024. “Holocene hydroclimate synthesis of the Aegean.” *The Holocene*. DOI 10.1177/09596836241275028

- Weiberg, E. & Finné, M. 2018. “Resilience and persistence of ancient societies in the face of climate change.” *World Archaeology*. DOI 10.1080/00438243.2018.1515035

SEISMICITY

- Nur, A. & Cline, E.H. 2000. “Poseidon's Horses: Plate Tectonics and Earthquake Storms in the Late Bronze Age Aegean and Eastern Mediterranean.” *Journal of Archaeological Science* 27(1): 43–63.

- Stiros, S. 1996. “Identification of earthquakes from archaeological data: methodology, criteria and limitations,” in Stiros & Jones (eds.), *Archaeoseismology*, Fitch Laboratory Occasional Paper 7.

METALS AND PROVENANCE

- Powell, W. et al. 2022. “Tin from Uluburun shipwreck shows small-scale commodity exchange fueled continental tin supply across Late Bronze Age Eurasia.” *Science Advances* 8(48): eabq3766. DOI 10.1126/sciadv.abq3766

- Stos-Gale, Z.A. et al. 1997. “Lead isotope characteristics of the Cyprus copper ore deposits applied to provenance studies of copper oxhide ingots.” *Archaeometry* 39(1). DOI 10.1111/j.1475-4754.1997.tb00792.x

- OXALID lead-isotope reference database, University of Oxford.

CHRONOLOGY

- Reimer, P.J. et al. 2020. “The IntCal20 Northern Hemisphere radiocarbon age calibration curve.” *Radiocarbon* 62(4).

- Manning, S.W. et al. 2020. “Mediterranean radiocarbon offsets and calendar dates for prehistory.” *Science Advances*. DOI 10.1126/sciadv.aaz1096

- Manning, S.W. et al. 2020. “Radiocarbon offsets and old world chronology as relevant to Mesopotamia, Egypt, Anatolia and Thera.” *Scientific Reports*. DOI 10.1038/s41598-020-69287-2

- Manning, S.W. et al. 2018. “Fluctuating radiocarbon offsets observed in the southern Levant.” *PNAS* 115(24): 6141.

NETWORK SCIENCE

- Knappett, C., Evans, T. & Rivers, R. 2008. “Modelling maritime interaction in the Aegean Bronze Age.” *Antiquity* 82(318): 1009–1024.

- Knappett, C., Rivers, R. & Evans, T. 2011. “The Theran eruption and Minoan palatial collapse.” *Antiquity* 85(329): 1008–1023. DOI 10.1017/S0003598X00068459

- Cline, D.H. & Cline, E.H. 2015. “Text Messages, Tablets, and Social Networks: The ‘Small World’ of the Amarna Letters,” in *There and Back Again — the Crossroads II*, 17–44. Charles University, Prague.

- Cohen, R., Erez, K., ben-Avraham, D. & Havlin, S. 2000. “Resilience of the Internet to random breakdowns.” *Phys. Rev. Lett.* 85(21): 4626. DOI 10.1103/PhysRevLett.85.4626
- Motter, A.E. & Lai, Y.-C. 2002. “Cascade-based attacks on complex networks.” *Phys. Rev. E* 66: 065102(R). DOI 10.1103/PhysRevE.66.065102
- Buldyrev, S.V. et al. 2010. “Catastrophic cascade of failures in interdependent networks.” *Nature* 464: 1025–1028. DOI 10.1038/nature08932
- Schneider, C.M. et al. 2011. “Mitigation of malicious attacks on networks.” *PNAS* 108(10): 3838–3841.
- Latora, V. & Marchiori, M. 2001. “Efficient behavior of small-world networks.” *Phys. Rev. Lett.* 87: 198701.
- Clauset, A., Shalizi, C.R. & Newman, M.E.J. 2009. “Power-law distributions in empirical data.” *SIAM Review* 51(4): 661–703.
- Scheffer, M. et al. 2009. “Early-warning signals for critical transitions.” *Nature* 461: 53–59.

[DATA REQUIREMENT] — EVIDENCE CLASSES SPECIFIED BUT NOT RESOLVED TO PARTICULAR DATASETS

- Comprehensive site-level destruction and abandonment catalogue for 1250–1100 BCE meeting the re-audit rubric.
- Anatolian and northern Levantine hydroclimate records meeting the Stream P acceptance criteria.
- Complete published corpus of oxhide-ingot and tin isotope determinations within the temporal universe, with analytical metadata.
- Storage-capacity-by-phase measurements at palatial sites, for estimation of α.
- Systematic excavation-intensity metadata for all sites in the universe.
- Regional field-survey datasets, systematically collected and published, as a second sampling frame.
- Palaeoseismic trench and offset-feature data for the relevant fault systems.

Protocol v1.0 — methodology only. No empirical results are claimed or implied. All quantities marked `[DATA REQUIREMENT]` must be resolved to specific, verifiable datasets before any analysis is executed. The interactive panels simulate the mechanisms described in the text on synthetic networks and illustrative distributions; they are teaching instruments, not reconstructions of the Late Bronze Age.