Article

Constraint-aware prediction for novel ecosystems: Delimiting the predictable domain and the emergence boundary in the Anthropocene

WenJun Zhang
School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China
E-mail: wjzhang@iaees.org, zhwj@mail.sysu.edu.cn

Received 3 September 2026; *Accepted 26 September 2026*; *Published online 31 October 2026*; *Published 1 September 2027*
(cc) BY

Abstract
Forecasting methodologies for novel ecosystems share a structural dependence on prior information: historical baselines, structural priors about bifurcation neighborhoods, or assumptions of pattern transferability from training distributions. Ecological novelty does not invalidate these priors uniformly. It erodes them differentially across dynamical features, and existing frameworks lack a systematic language for expressing which features remain constrainable under a given novelty profile and which do not. This paper develops Constraint-Aware Prediction (CAP), a methodological framework that reframes novelty from a property of predicted outcomes into an informational condition of prediction itself. CAP partitions dynamical features into three domains. The constrained domain contains features for which predictive distributions contract relative to priors in a testable manner. The coupled domain contains features governed by cross-scale transmission whose conductance can be estimated but whose parameters are uncertain under novelty. The emergent domain contains features arising from interactions among novel species combinations whose specific configurations cannot be deduced from prior information. For each domain, the framework specifies an operational criterion. Constraint degree is estimated through Bayesian predictive contraction, cross-condition functional stability, and network-topological extrapolability. Social processes are modeled not as external scenario inputs but as structural generators of the constraint set, capable of expanding or compressing the constrained domain through governance responses, institutional lock-in, and knowledge production. CAP does not aim to eliminate uncertainty. It aims to make prediction honestly bounded and therefore usable under conditions without historical precedent. Three testable inferences are derived, and the framework is connected to hierarchical predictive theories of biodiversity change and to the literature on regime shifts, machine learning transferability, and social-ecological feedbacks.

Keywords novel ecosystems; constraint-aware prediction; predictable domain; emergence boundary; ecological novelty; regime shifts; social-ecological systems; adaptive governance; Bayesian predictive contraction; transferability.

Computational Ecology and Software
ISSN 2220-721X
URL: http://www.iaees.org/publications/journals/ces/online-version.asp
RSS: http://www.iaees.org/publications/journals/ces/rss.xml
E-mail: ces@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Novel ecosystems are defined by combinations of species and abiotic conditions that have not previously occurred within a given biome, and they arise through deliberate or inadvertent human action (Hobbs et al., 2006). The concept carries two defining features: novelty in species composition and functioning, and human agency as the generative force (Hobbs et al., 2006; Hobbs et al., 2013; Staples et al., 2025). A working definition holds that a novel ecosystem is a system of abiotic, biotic and social components that, by virtue of human influence, differs from those that prevailed historically, has a tendency to self-organize (Zhang, 2014, 2015, 2016b. 2026a-c, 2027a-b) and manifests novel qualities without intensive human management (Hobbs et al., 2013). Novel ecosystems are distinguished from hybrid systems by practical limitations, combining ecological, environmental and social thresholds, on the recovery of historical qualities (Hobbs et al., 2013). As human transformation of the Earth intensifies, these systems increase in importance while remaining relatively little studied (Hobbs et al., 2006; Zhang and Chen, 2011; Zhang and Liu, 2012).

The predictive challenge posed by novel ecosystems is not simply that extrapolation distances become longer. The information base on which prediction rests undergoes a structural change. Historical baselines lose their referential function because the system has moved beyond the threshold separating historical from novel configurations (Hobbs et al., 2013). Structural priors about the neighborhood of a bifurcation lose their warrant because the generative processes reorganize species pools and functional relationships (Kerr et al., 2026). Training distributions lose their coverage because the target conditions have no analogue in the observational record (Wu et al., 2026; Farchadi et al., 2025).

Existing methodological responses to this challenge cluster into three families, and each carries an implicit prior. Early warning signal approaches assume that the system resides in or near a bifurcation neighborhood where critical slowing down produces detectable statistical signatures (Radhakrishnan et al., 2025). Process-based frameworks assume that the mapping from generative processes to novel outcomes is identifiable, at least in principle, through measurement of composition, function, abiotic conditions and social drivers (Kerr et al., 2026). Data-driven approaches assume that structural relationships learned from training data remain transferable to target conditions, an assumption that machine learning research in ecology has repeatedly found to be fragile under extrapolation (Wu et al., 2026; Kearney et al., 2026).

The central argument of this paper is that these priors do not fail simultaneously or uniformly. Some dynamical features, such as response diversity and the dominance of negative feedback, retain constrainability longer than others, such as the specific configuration of interactions among newly assembled species. The differential erosion of prior information is the phenomenon that a methodology for novel ecosystem prediction must address. The first step is not the selection of a more powerful model but the delimitation, for a given novelty profile, of the boundary of the predictable domain. This delimitation can itself be made a rigorous methodological object.

The paper proceeds as follows. Section 2 examines how the three families of prediction methods depend on prior information and why their failure under novelty is layered rather than total. Section 3 develops the Constraint-Aware Prediction framework, defining constraint degree, the three-domain partition, and the criteria that operationalize each domain. Section 4 specifies estimation strategies and derives testable inferences. Section 5 discusses the epistemological stance and the implications for adaptive governance.

2 The Layered Failure of Prior-Dependent Prediction

2.1 Three disguises of the baseline

Equilibrium thinking persists in prediction methodology under three guises. The historical baseline treats a past state as the reference against which deviation is measured. The statistical baseline treats the training

distribution as the universe of conditions for which a model is calibrated. The structural baseline treats a bifurcation neighborhood, or more generally a recognizable dynamical skeleton, as the condition under which indicators carry information. Novelty compromises all three, but not to the same degree and not through the same mechanism.

The historical baseline is compromised by definition: novel ecosystems differ from historical ones precisely in the combinations of species and abiotic conditions they contain (Hobbs et al., 2006). The structural baseline is compromised when the system is generating its own structure rather than moving within a fixed structure (Kerr et al., 2026). The statistical baseline is compromised when the mechanisms linking predictors to responses are themselves altered by environmental change, a condition that violates a fundamental assumption of supervised learning (Zhang, 2007a-b; Zhang et al., 2008; Wu et al., 2026). A methodology that does not declare which baseline its predictions depend on, and to what degree, cannot be evaluated under novelty (Staples et al., 2025).

2.2 Early warning signals and the boundaries of their applicability

Early warning signals rest on critical slowing down, the phenomenon in which a system's response to perturbation relaxes more slowly as a bifurcation is approached (Radhakrishnan et al., 2025). Variance and autocorrelation are the canonical indicators, and they have been reported to increase before shifts in ecosystems, climate, food webs and other dynamical systems (Radhakrishnan et al., 2025; Zhang, 2012, 2018, 2026a-b, 2027a-b). The applicability of critical slowing down is limited to systems experiencing critical transitions that can be modeled using bifurcation theory (Radhakrishnan et al., 2025). Conventional signals also struggle to distinguish oscillatory behavior associated with Hopf bifurcations from shifts to a distant attractor associated with fold bifurcations, and they become less reliable in systems affected by density-dependent noise (Bury et al., 2021).

The empirical record is sobering. Analyses of multi-trophic planktonic data from lakes worldwide found few instances of critical transitions, with different trophic levels often expressing different forms of abrupt change and most indicators performing no better than chance (O'Brien et al., 2023). A machine learning model applied in the same study performed poorly (O'Brien et al., 2023). The authors concluded that predictive ecology should move away from the concept of critical transitions and develop methods suitable for predicting resilience loss not limited to the strict bounds of bifurcation theory (O'Brien et al., 2023). When parameters change rapidly, early warnings can arrive after the bifurcation point has already been passed, and the fast rate of parameter change can delay the tipping event, opening a borrowed window for intervention whose duration depends on the interaction between warning time, prevention action and rate of change (Radhakrishnan et al., 2025).

These findings do not invalidate early warning signals. They locate their applicable domain. That domain is bounded by the requirement that the system be close to a recognizable bifurcation type and that the parameter drift be slow relative to the system's relaxation dynamics. Novel ecosystems that assemble through species pool reorganization and functional reconfiguration may never satisfy the first requirement, because the structure within which a bifurcation would occur is itself being generated (Kerr et al., 2026).

2.3 Data-driven prediction and the conditional nature of transferability

Machine learning in ecology faces a spatial transferability bottleneck (Zhang, 2007a-b, 2010; Zhang and Barrion, 2006; Zhang and Zhang, 2008; Zhang et al., 2007, 2008). Models trained in local environments often degrade when extrapolated to regional or global scales because ecological relationships are shaped by site-specific factors such as microclimate and land-use history (Wu et al., 2026). Carbon flux models trained on local eddy covariance observations can capture local relationships yet lose accuracy across biomes. Soil property products show low correlations between predicted and observed values in heterogeneous regions

(Guevara et al., 2018; Tifafi et al., 2017).

In the temporal dimension, the breakdown of stationarity is the central obstacle. Ecological forecasting increasingly faces situations in which the governing processes are themselves evolving, as climate change, land-use change, biological invasions and shifting disturbance regimes reshape ecosystem structure and function and create conditions with no historical analogue (Wu et al., 2026). Relationships learned from historical observations may no longer remain valid when projected into future scenarios (Dietze et al., 2018; Reichstein et al., 2019). This challenge extends beyond distribution shift because environmental change can alter the ecological mechanisms linking predictors and responses, violating a fundamental assumption of supervised learning (Wu et al., 2026).

Empirical evaluations of cross-validation practice reinforce the point. Random k-fold cross-validation produced overly optimistic performance estimates compared with more extrapolative validation schemes, and simpler algorithms such as partial least squares regression were more consistent and outperformed complex algorithms for the most extrapolative tasks (Kearney et al., 2026). The authors concluded that random k-fold cross-validation likely produces unrealistically optimistic expectations and could be associated with major prediction misses when models are used in novel environmental conditions (Kearney et al., 2026). Transfer learning and meta-learning can draw on data-rich regions or tasks to support predictions in data-sparse contexts, but their effectiveness depends on the extent to which ecological patterns can be transferred, and they remain constrained by limited observations in underrepresented areas (**Nathaniel et al., 2023**).

These results establish that transferability is not a property that can be assumed. It is a conditional property that must be estimated for each combination of model, training distribution and target novelty profile. Data integration can improve forecast skill under novel ocean conditions, with model-based integration of diverse data types outperforming simpler approaches, but performance still declined as environmental novelty increased across all models tested (Farchadi et al., 2025). Novelty erodes the statistical baseline continuously rather than switching it off at a threshold (Staples et al., 2025).

2.4 Social processes as externalized drivers

The social dimension of novel ecosystems is acknowledged in the foundational literature. A novel ecosystem is a system of abiotic, biotic and social components (Hobbs et al., 2013), and human agency is a defining characteristic rather than a background condition (Hobbs et al., 2006). Governance research has examined how institutional arrangements could enable adaptive responses to transformative changes such as novel ecosystems, and has argued for a more inclusive paradigm of collaborative governance drawing on multiple forms of knowledge (McKelvey et al., 2021; Chaffin et al., 2014).

In prediction methodology, however, social processes are typically handled as external scenario inputs. Participatory scenario narratives generate storylines about alternative futures, and these storylines are then translated into driver values that parameterize biophysical models. This architecture leaves the formal structure of the ecological model unchanged across scenarios. The social process influences the values of driving variables but not the feedback structure or the feasible state space of the system.

The limitation of this treatment becomes visible when social responses alter the system's constraint structure. A governance intervention that maintains response diversity expands the set of dynamical features that remain constrainable. Institutional lock-in that forecloses adaptive options compresses that set. Social-ecological systems research has developed non-recursive structural equation models that estimate feedback loops between ecological health, governance and user behavior, and has shown that the same system can move from a regime in which endogenous management is effective to one in which external climatic forcing dominates and institutional responses lose their regulatory capacity (Bozzeda and Defeo, 2026). When such structural shifts occur, a prediction that treats governance as an external driver will misrepresent the

system at the level of its feedback architecture, not merely at the level of parameter values.

3 Constraint-Aware Prediction: Framework Construction (Fig. 1)

3.1 Constraint degree as a first-order predictive quantity

The central construct of the framework is constraint degree. Let $\mathcal{N}$ denote a novelty profile described across four dimensions: compositional, functional, abiotic and social. Let $\mathcal{D}$ be a set of dynamical features of interest, such as the persistence of a functional group, the sign of a dominant feedback, or the occurrence of a particular oscillatory mode. Constraint degree is defined as

$$C(\mathcal{D}|\mathcal{N}) = \mathbb{E}_{\mathcal{N}}[(D_{\text{prior}}(\mathcal{D}) - D_{\text{post}}(\mathcal{D}|\mathcal{N}))/D_{\text{prior}}(\mathcal{D})]_+$$

where D_{prior} is a dispersion functional for the predictive distribution of the feature under the prior information available before conditioning on system-specific data, D_{post} is the corresponding dispersion after conditioning, and the positive part ensures that only genuine contraction counts as constraint. The expectation is taken over the novelty profile. A feature with C near 1 is strongly constrained by current theory and data. A feature with C near 0 has a posterior distribution that has not contracted relative to its prior, meaning that the available information does not discriminate among plausible values.

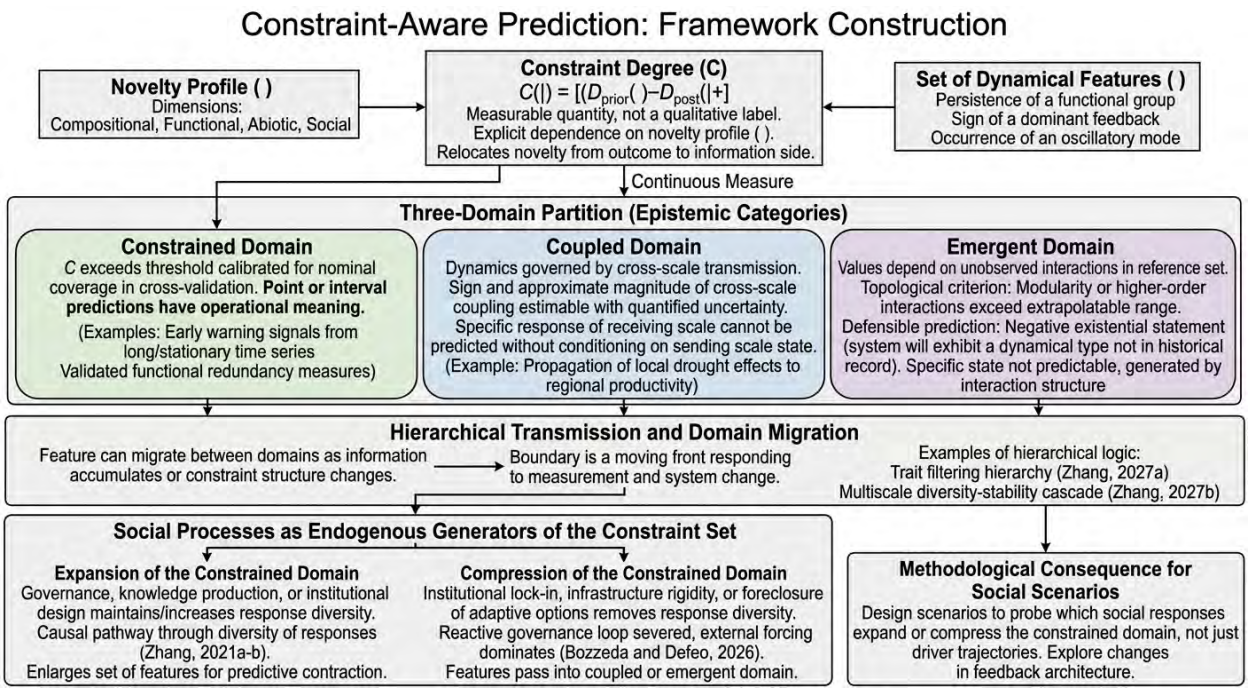

Fig. 1 Constraint-Aware Prediction: Framework Construction.

This definition makes constraint degree a measurable quantity rather than a qualitative label. It also makes explicit the dependence on novelty. C is not a property of a feature in isolation but of a feature under a specified novelty profile. The same feature, such as the response of a functional group to warming, may be well constrained in a system with historical analogues and poorly constrained in a system whose species pool has been reorganized. The definition thus relocates novelty from the outcome side of prediction to the information side.

3.2 Three domains and their criteria

Constraint degree provides a continuous measure. The framework converts this measure into a three-domain partition by specifying thresholds that carry distinct methodological consequences.

The constrained domain contains features for which C exceeds a threshold calibrated so that interval predictions achieve nominal coverage in cross-validation. Within this domain, point or interval predictions have operational meaning. The criterion is not a fixed numerical cutoff but a calibration requirement: the predictive interval must have demonstrated coverage under novelty levels comparable to the target. Early warning signals computed from sufficiently long and stationary time series can fall in this domain when the structural baseline is approximately satisfied, and functional redundancy measures can fall here when the mapping from redundancy to function has been validated across a novelty gradient.

The coupled domain contains features whose dynamics are governed by cross-scale transmission. Local responses propagate through spatial or organizational scales, and the conductance of this propagation can be estimated, but the parameters of the transmission are uncertain under novelty. The criterion for membership is that the sign and approximate magnitude of the cross-scale coupling can be estimated with quantified uncertainty, while the specific response of the receiving scale cannot be predicted without conditioning on the actual state of the sending scale. A concrete example is the propagation of local drought effects to regional productivity through dispersal and landscape connectivity. The direction of the effect may be estimable; its magnitude under a novel species composition may not.

The emergent domain contains features whose values depend on interactions that have not been observed in the reference set. The criterion for membership is topological: when the modularity of the interaction network, or the strength of higher-order interactions, exceeds the range that can be extrapolated from historical or reference networks, the specific configuration of the feature is treated as emergent. Within this domain, the defensible prediction is a negative existential statement: the system will exhibit a dynamical type that is not present in the historical record. The specific state cannot be predicted because it is generated by the interaction structure rather than determined by prior constraints.

The three domains are not ontological categories of nature. They are epistemic categories relative to the current state of theory, data and monitoring. A feature can migrate between domains as information accumulates or as the system's constraint structure changes.

3.3 Hierarchical transmission and domain migration

The domain partition is dynamic. Monitoring does not merely update parameter estimates within a fixed domain; it can reveal that a feature has moved from the constrained to the coupled domain or from the coupled to the emergent domain. The logic of hierarchical transmission is borrowed from predictive theories that connect organizational levels through explicit mappings. A trait filtering hierarchy connects individual fitness surfaces, population growth rates, community assembly and regional species pools through a measurable mapping from environmental states and trait combinations to persistence probabilities (Zhang, 2027a). A multiscale diversity-stability cascade connects local stability through spatial asynchrony to regional stability (Zhang, 2027b). In both cases, prediction is not a point estimate at a single scale but a transmitted functional relationship across levels.

Constraint-Aware Prediction adopts this hierarchical logic but applies it to the epistemic status of features rather than to their values. A feature that is constrained at one level may enter the coupled domain when its determinants at a lower level become uncertain, and may enter the emergent domain when the lower-level interactions are themselves novel. The boundary between domains is therefore a moving front that responds to both measurement and system change.

3.4 Social processes as structural generators of the constraint set

The framework treats social processes as endogenous generators of the constraint set rather than as external drivers. Two pathways are distinguished.

Expansion of the constrained domain occurs when governance, knowledge production or institutional design maintains or increases the response diversity available to the system. A governance regime that preserves functional redundancy across scales, or that supports monitoring capable of detecting domain migration, enlarges the set of features for which predictive contraction can be achieved. The causal pathway runs through the diversity of responses available to the system (Zhang, 2021a-b), not through the values of driving variables.

Compression of the constrained domain occurs when institutional lock-in, infrastructure rigidity or the foreclosure of adaptive options removes response diversity and narrows the range of dynamical behaviors the system can express. In the language of social-ecological feedback models, when the reactive governance loop is severed and external forcing dominates, endogenous management efforts are rendered ineffective and the system's adaptive capacity is reduced (Bozzeda and Defeo, 2026). In constraint terms, features that were previously constrainable through management-responsive feedbacks pass into the coupled or emergent domain.

This treatment has a methodological consequence. Social scenarios should not be designed solely to generate driver trajectories. They should be designed to probe which social responses expand or compress the constrained domain. A scenario that changes the values of temperature and precipitation while leaving governance structure fixed explores a different kind of uncertainty than a scenario that changes the feedback architecture itself. Constraint-Aware Prediction requires the second kind of exploration alongside the first.

4 Operationalization and Testable Inferences

4.1 Estimating constraint degree

Three estimation strategies correspond to the three domains and can be combined in practice.

The first estimation strategy is Bayesian predictive contraction. For a feature of interest, specify a prior predictive distribution based on available theory and data, then update with system-specific observations to obtain a posterior predictive distribution. Constraint degree is the relative reduction in a discrete functional such as predictive variance or interval width. This strategy has the advantage of being applicable to any feature for which a probabilistic model can be formulated, and it makes the dependence on prior choice explicit. Its limitation is that the prior itself may be misspecified under novelty, a problem that can be partially addressed by prior sensitivity analysis across a hierarchy of increasingly weak priors.

The second estimation strategy is cross-condition functional stability. For functional features such as the relationship between response diversity and the maintenance of a process rate, estimate the relationship in multiple reference systems spanning a novelty gradient and test whether its sign and approximate magnitude remain stable. If the relationship retains its direction across the gradient, the feature belongs to the constrained domain with respect to that gradient. If the relationship weakens or changes sign beyond a threshold of novelty, the feature has migrated to the coupled or emergent domain. This strategy operationalizes the intuition that functional redundancy carries information that is partly independent of species identity, while making the limits of that independence an empirical question.

The third one is network-topological extrapolability. For features determined by interaction structure, estimate the modularity and higher-order interaction strength of the interaction network (Zhang, 2012, 2016b, 2018) and compare these quantities to the range observed in the reference set. When modularity or higher-order strength exceeds the reference range, the feature is assigned to the emergent domain. The threshold is not arbitrary: it is the upper bound of the reference distribution, and it can be updated as the

reference set grows. This strategy connects the framework to network-based approaches in ecology that formalize the relationship between interaction structure and dynamical behavior (Zhang, 2015, 2016a-b, 2018).

4.2 Monitoring as domain-boundary detection

The operationalization changes the function of monitoring. In a conventional forecasting architecture, monitoring generates data for parameter estimation and forecast verification. In Constraint-Aware Prediction, monitoring also generates evidence about whether the constraint structure itself has shifted.

A monitoring program designed for domain-boundary detection tracks indicators of constraint degree alongside the features being predicted. When constraint degree for a feature falls below the calibrated threshold, or when a network-topological indicator crosses the reference upper bound (Zhang, 2016a), the appropriate response is not to update a parameter within the existing model but to switch the prediction mode. The feature moves from the constrained to the coupled domain, or from the coupled to the emergent domain, and the output of the predictive system changes accordingly: from an interval prediction to a conditional scenario spectrum, or from a scenario spectrum to a negative existential statement.

This architecture has an important consequence for the interpretation of forecast error. In the conventional architecture, a large forecast error is evidence that the model is wrong. In Constraint-Aware Prediction, a large forecast error may be evidence that the domain assignment was wrong, which is a different kind of information. The two diagnoses call for different responses: model revision in the first case, domain reassignment in the second.

4.3 Social-ecological scenarios as constraint-space exploration

Scenario methods can be redesigned to serve constraint-space exploration. Instead of asking stakeholders to describe plausible future driver values, the exercise asks which social responses would change the feedback structure of the system. A governance intervention that creates a new monitoring capacity, or that removes a legal barrier to adaptive management, expands the constrained domain. A decision that locks in a particular infrastructure pathway compresses it. The scenario output is not a set of driver trajectories but a map of which social choices move which features across domain boundaries.

This redesign connects the framework to adaptive governance research. Governance arrangements that maintain experimental capacity, reversibility and response diversity are not only normatively attractive; in constraint terms they preserve the predictable residual of the system (McKelvey et al., 2021; Chaffin et al., 2014). A system whose governance forecloses adaptive options loses constrainability along with adaptive capacity, and the loss is measurable through the constraint-degree indicators described above.

4.4 Testable inferences

The framework yields three inferences that can be evaluated against data:

(1) Under high novelty, functional predictions based on response diversity and redundancy should exhibit greater cross-condition stability than predictions based on species composition. The rationale is that response diversity encodes backup capacity that is partly independent of species identity, whereas composition-based predictions depend on the historical mapping from species to function. The inference can be tested by comparing the transferability of composition-based and redundancy-based models across a novelty gradient.

(2) Models that treat social processes as generators of the constraint set should outperform models that externalize social processes when predicting system trajectories after management interventions. The rationale is that interventions change feedback structure, and a model that holds feedback structure fixed across scenarios will misrepresent the post-intervention dynamics. The inference can be tested using structural equation models that estimate feedback loops and comparing their out-of-sample performance against models with exogenous social drivers.

(3) When interaction-network modularity exceeds the reference range, the calibration error of point predictions

should increase systematically (Zhang, 2016a), while the hit rate of negative existential predictions, statements that a dynamical type absent from the historical record will appear, should exceed the hit rate of specific state predictions. The rationale is that beyond the topological threshold, the specific configuration is generated rather than constrained. The inference can be tested in systems with sufficient time-series data to estimate both calibration error and the occurrence of previously unobserved dynamical types.

5 Discussion

5.1 Not an abandonment of prediction

A predictable objection holds that Constraint-Aware Prediction lowers the ambition of ecological forecasting by conceding large domains to unpredictability. The objection misreads the framework's epistemic target. The target is not precision but warranted credibility allocation. Under novelty, a prediction that declares a feature to be in the emergent domain, and therefore declines to specify its value while asserting that it will belong to a dynamical class absent from the historical record, can be more informative than a point prediction with unacknowledged and violated priors. The choice is not between prediction and no prediction. It is between prediction that misrepresents its own basis and prediction that states its basis and its limits.

This argument aligns with calls to move beyond the strict bounds of bifurcation theory in predictive ecology (O'Brien et al., 2023) while retaining the aspiration to anticipate change. It also aligns with the observation that forecasting performance of complex nonlinear models is context-dependent, with the best performance on chaotic and oscillating time series and weaker performance elsewhere (Buckner et al., 2026). A framework that specifies the conditions under which a method performs well is more useful than a method that performs well on average and fails unpredictably.

5.2 Relation to hierarchical predictive theories and constraint-first methodology

The framework has a structural affinity with a family of predictive theories that place constraints before mechanisms. The Niche-Energy-Time hypothesis treats thermodynamic, structural and historical constraints as irreducible determinants of species diversity patterns (Zhang, 2026b-c). The trait filtering-diversity dynamics hierarchy formalizes the mapping from environmental states and trait combinations to persistence probabilities through a filtering operator that can be estimated with hierarchical Bayesian models, structural causal models and interpretable machine learning (Zhang, 2027c). The multiscale diversity-stability cascade connects local, spatial and regional stability through explicit transmission relationships (Zhang, 2027b). The critical transition framework unifies radiations and mass extinctions as behaviors of the same complex adaptive system on either side of a threshold (Zhang, 2026a). The self-organized criticality framework places ecosystem phase transitions under the control of energy fluctuation variables (Zhang, 2026b).

Constraint-Aware Prediction inherits three features from this tradition. It treats constraint as first-order information rather than a residual category. It organizes prediction hierarchically, with transmission between levels. It insists on operationalizing qualitative concepts into measurable quantities. Its distinctive contribution is to apply this logic to the epistemic status of prediction itself rather than to the values of ecological quantities. Constraint degree, domain membership and domain migration are properties of the predictive situation, not of the ecosystem alone. This reflexive move is what allows the framework to address novelty directly: novelty changes the predictive situation, and the framework's first task is to describe that change.

5.3 Implications for adaptive governance

If the predictable residual of a system depends on the size and stability of its constrained domain, then managing the constraint set becomes a governance function. Maintaining response diversity, preserving experimental capacity and avoiding irreversible institutional commitments are strategies for keeping features in the constrained domain. Foreclosing adaptive options compresses the constrained domain and reduces the

system's predictable residual along with its adaptive capacity.

This implication connects constraint-aware prediction to the literature on governance of novel ecosystems. Governance research has argued that novel ecosystems require collaborative, adaptive arrangements that draw on multiple knowledge systems and can respond to constant change (McKelvey et al., 2021; Chaffin et al., 2014). Constraint-Aware Prediction supplies a methodological rationale for this argument. Adaptive governance is not only a response to normative commitments or to irreducible uncertainty. It is a strategy for preserving the informational conditions under which prediction remains possible.

5.4 Limitations

Several limitations bound the current formulation. The calibration of constraint-degree thresholds requires empirical work across diverse systems, and the thresholds may prove to be scale-dependent or system-specific. The domain boundaries are treated as sharp for operational purposes, but in reality they are likely fuzzy, and a fuzzy-set formulation may be more adequate. The network-topological criterion for the emergent domain relies on the availability of interaction-network estimates, which are difficult to obtain in species-rich systems (Zhang, 2016a). The formal modeling of social processes as constraint generators remains underdeveloped; the framework specifies the pathways through which social processes act but does not yet provide a general dynamic model of how governance and knowledge production alter the constraint set. These limitations define a research agenda rather than undermining the framework's core logic.

6 Conclusion

Novel ecosystems challenge prediction at the level of its information base, not merely at the level of model choice. The failure of prior-dependent methods under novelty is layered: some dynamical features leave the constrained domain earlier than others, and a methodology that treats novelty as a uniform degradation of predictability will either overstate or understate what can be said. Constraint-Aware Prediction responds to this condition by defining constraint degree as a first-order predictive quantity, partitioning dynamical features into constrained, coupled and emergent domains with explicit criteria, treating domain boundaries as dynamic and responsive to monitoring, and modeling social processes as structural generators of the constraint set. The framework does not promise to predict the specific configurations that novel ecosystems will assume. It promises something more defensible: a systematic account of which features remain constrainable under a given novelty profile, which have passed into the domain of cross-scale transmission with uncertain parameters, and which must be addressed through negative existential prediction rather than point forecasting. In the Anthropocene, the ambition of predictive ecology should be to state, with warranted confidence, what can still be said, and to mark honestly where speech must give way to monitored silence.

Appendix: Computational Details

A.1 Bayesian predictive contraction

Let θ denote the parameter vector of a model for feature $d \in \mathcal{D}$. Let $p(\theta|\mathcal{I}_0)$ be the prior distribution reflecting information available before system-specific data, and let $p(\theta|\mathcal{I}_0, \mathbf{y})$ be the posterior after observing data $\mathbf{y}$. The prior predictive distribution for d is $p(d|\mathcal{I}_0) = \int p(d|\theta)p(\theta|\mathcal{I}_0)d\theta$. The posterior predictive distribution is $p(d|\mathcal{I}_0, \mathbf{y}) = \int p(d|\theta)p(\theta|\mathcal{I}_0, \mathbf{y})d\theta$. The discrete functional is taken as the width of the central 90 percent predictive interval, $W_{0.9}(\cdot)$. Constraint degree for feature d is estimated as $\hat{C}(d) = \max\{0, [W_{0.9}(p(d|\mathcal{I}_0)) - W_{0.9}(p(d|\mathcal{I}_0, \mathbf{y}))]/W_{0.9}(p(d|\mathcal{I}_0))\}$. This quantity is computed for each feature and each novelty profile. Prior

sensitivity is assessed by repeating the calculation under a hierarchy of priors with increasing variance and comparing the resulting $\hat{C}$ values. A feature whose $\hat{C}$ remains above the calibrated threshold across the prior hierarchy is assigned to the constrained domain. The threshold is calibrated by simulation: synthetic systems with known domain membership are generated, $\hat{C}$ is computed, and the threshold that maximizes classification accuracy is selected.

A.2 Cross-condition functional stability

For a functional relationship F between an indicator r (such as response diversity) and a process rate g, estimate the relationship in K reference systems ordered by a novelty index. For system k, fit $g_k = \alpha_k + \beta_k r_k + \epsilon_k$. Compute the sign stability $S_{\text{sign}} = (1/K)\ \Sigma_k \mathbf{1}[\text{sign}(\hat{\beta}_k)=\text{sign}(\hat{\beta}_{k'})]$ for all pairs (k, k') and the magnitude consistency $S_{\text{mag}} = 1 - \text{CV}(|\widehat{\beta}_k|)$, where CV is the coefficient of variation. If S_{sign} exceeds a threshold (for example 0.8) and S_{mag} exceeds a threshold (for example 0.5), the relationship is classified as constrained with respect to the novelty gradient. If sign stability is high but magnitude consistency is low, the relationship is classified as coupled. If sign stability is low, it is classified as emergent. The thresholds are calibrated on simulated systems with known stability properties.

A.3 Network-topological extrapolability

For an interaction network with adjacency matrix **A** (Zhang, 2016b, 2018), compute modularity Q using a community detection algorithm such as Louvain or Leiden. Compute higher-order interaction strength H as the proportion of variance in node-level growth rates explained by terms involving three or more nodes beyond what is explained by pairwise terms, estimated through a regression of growth rates on pairwise and higher-order interaction terms. For the reference set of historical networks, compute the empirical distribution of Q and H. The upper reference bounds are $Q_{\max}$ and $H_{\max}$, taken as the 95th percentiles. A target network with $Q > Q_{\max}$ or $H > H_{\max}$ is classified as emergent for features whose dynamics depend on interaction structure. The choice of the 95th percentile is a calibration decision that trades false positives against false negatives and should be sensitivity-tested.

References

Bozzeda F, Defeo O. 2026. Understanding feedbacks in social–ecological systems using non-recursive structural equation models. Ecological Modelling, 519: 111628. https://www.sciencedirect.com/science/article/pii/S0304380026001560

Bury TM, Sujith RI, Pavithran I, Scheffer M, Lenton TM, Anand M, Bauch CT. 2021. Deep learning for early warning signals of tipping points. Proceedings of the National Academy of Sciences, 118(39): e2106140118. https://www.pnas.org/doi/10.1073/pnas.2106140118

Buckner JH, Meunier ZD, Arroyo-Esquivel J, et al. 2026. Recovering complex ecological dynamics from time series using state-space universal dynamic equations. Communications Earth & Environment, 7: 112. https://www.nature.com/articles/s43247-025-03130-2

Chaffin BC, Gosnell H, Cosens BA. 2014. A decade of adaptive governance scholarship: Synthesis and future directions. Ecology and Society, 19(3): ES-06824-190356. https://www.mendeley.com/catalogue/a488ffe5-d400-359e-96cd-7ee17cb86b95/

Dietze M C, Fox A, Beck-Johnson LM, et al. 2018. Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences of the United States of America, 115(7): 1424–1432. https://doi.org/10.1073/pnas.1710231115

Farchadi N, Braun CD, Arostegui MC, et al. 2025. Data integration improves species distribution forecasts under novel ocean conditions. Ecography, e07997.

https://nsojournals.onlinelibrary.wiley.com/doi/10.1002/ecog.07997

Guevara M, Olmedo GF, Stell E, et al. 2018. No silver bullet for digital soil mapping: Country-specific soil organic carbon estimates across Latin America. Soil, 4(3): 173-193. https://doaj.org/article/5891a424cc194173a6d3d0789ed398fa

Hobbs RJ, Arico S, Aronson J, et al. 2006. Novel ecosystems: Theoretical and management aspects of the new ecological world order. Global Ecology and Biogeography, 15(1): 1-7. https://onlinelibrary.wiley.com/doi/10.1111/j.1466-822X.2006.00212.x

Hobbs RJ, Higgs ES, Hall CM. 2013. Defining Novel Ecosystems. In: Novel Ecosystems (Hobbs RJ, Higgs ES, Hall CM, eds). Wiley-Blackwell, Oxford, UK. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118354186.ch6

Kearney SP, Augustine DJ, Porensky LM, et al. 2026. Bringing cross-validation into the real world to evaluate transferability of satellite-based vegetation models. Scientific Reports, 16: 9383. https://www.nature.com/articles/s41598-026-39866-w

Kerr MR, Svenning J-C, Riede F, et al. 2026. The making of novel ecosystems: A process-based framework for measurement, analysis and application. Methods in Ecology and Evolution, 17(3): 683-704. https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.70255

McKelvey KS, Block WM, Jain TB, et al. 2021. Adapting research, management, and governance to confront socioecological uncertainties in novel ecosystems. Frontiers in Forests and Global Change, 4: 644696. https://www.frontiersin.org/journals/forests-and-global-change/articles/10.3389/ffgc.2021.644696/full

Nathaniel J, Liu JG, Gentine P. 2023. Metaflux: meta-learning global carbon fluxes from sparse spatiotemporal observations. Scientific Data, 10: 440. https://www.nature.com/articles/s41597-023-02349-y

O'Brien DA, Deb S, Gal G, et al. 2023. Early warning signals have limited applicability to empirical lake data. Nature Communications, 14: 7942. https://www.nature.com/articles/s41467-023-43744-8

Radhakrishnan R, Pavithran I, Livina V, Kurths J, Sujith RI. 2025. Early warnings are too late when parameters change rapidly. Scientific Reports, 15: 20256. https://www.nature.com/articles/s41598-025-06525-5

Reichstein M, Camps-Valls G, Stevens B, et al. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743): 195-204. https://www.nature.com/articles/s41586-019-0912-1

Staples TL, Blois J, Cramer KL, et al. 2025. A conceptual framework for measuring ecological novelty. Global Ecology and Biogeography, 34(2): e70005. https://onlinelibrary.wiley.com/doi/10.1111/geb.70005

Tifafi M, Guenet B, Hatté C. 2017. Large differences in global and regional total soil carbon stock estimates based on SoilGrids, HWSD, and NCSCD: intercomparison and evaluation based on field data from USA, England, Wales, and France. Global Biogeochemical Cycles, 32(1): 42-56. https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017GB005678

Wu JP, Luo YQ, Liu FY, et al., 2026. Machine learning in ecology: Current status and future directions in the big data era. Geography and Sustainability, 7(5): 100533. https://www.sciencedirect.com/science/article/pii/S2666683926001288

Zhang WJ. 2007a. Pattern classification and recognition of invertebrate functional groups using self-organizing neural networks. Environmental Monitoring and Assessment, 130: 415-422. https://link.springer.com/article/10.1007/s10661-006-9432-1

Zhang WJ. 2007b Supervised neural network recognition of habitat zones of rice invertebrates. Stochastic Environmental Research and Risk Assessment, 21: 729-735. https://link.springer.com/article/10.1007/s00477-006-0085-y

Zhang WJ. 2010. Computational Ecology: Artificial Neural Networks and Their Applications. World Scientific, Singapore. https://www.worldscientific.com/worldscibooks/10.1142/7436#t=aboutBook

Zhang WJ. 2012. Computational Ecology: Graphs, Networks and Agent-based Modeling. World Scientific, Singapore. https://www.worldscientific.com/worldscibooks/10.1142/8118#t=aboutBook

Zhang WJ. 2014. Selforganizology: A more detailed description. Selforganizology, 1(1): 31-46. http://www.iaees.org/publications/journals/selforganizology/articles/2014-1(1)/5-Zhang-Abstract.asp

Zhang WJ. 2015. A generalized network evolution model and self-organization theory on community assembly. Selforganizology, 2(3): 55-64. http://www.iaees.org/publications/journals/selforganizology/articles/2015-2(3)/3-Zhang-Abstract.asp

Zhang WJ. 2016a. Network robustness: Implication, formulization and exploitation. Network Biology, 6(4): 75-85. http://www.iaees.org/publications/journals/nb/articles/2016-6(4)/1-Zhang-Abstract.asp

Zhang WJ. 2016b. Selforganizology: The Science of Self-Organization. World Scientific, Singapore. https://www.worldscientific.com/worldscibooks/10.1142/9685

Zhang WJ. 2018. Fundamentals of Network Biology. World Scientific Europe, London, UK. https://www.worldscientific.com/worldscibooks/10.1142/q0149

Zhang WJ. 2021a. A statistical simulation method for causality inference of Boolean variables. Network Biology, 11(4): 263-273. http://www.iaees.org/publications/journals/nb/articles/2021-11(4)/2-Zhang-Abstract.asp

Zhang WJ. 2021b. Causality inference of nominal variables: A statistical simulation method. Computational Ecology and Software, 11(4): 142-153. http://www.iaees.org/publications/journals/ces/articles/2021-11(4)/1-Zhang-Abstract.asp

Zhang WJ. 2026a. What triggers explosive radiations and mass extinctions? A unified critical transition framework. Proceedings of the International Academy of Ecology and Environmental Sciences, 16(3): 89-122. http://www.iaees.org/publications/journals/piaees/articles/2026-16(3)/2-Zhang-Abstract.asp

Zhang WJ. 2026b. Ecological thermodynamic fluctuation and phase transition: A self-organized criticality-based framework for ecosystem responses to global warming. Selforganizology, 13(3-4): 66–103. http://www.iaees.org/publications/journals/selforganizology/articles/2026-13(3-4)/2-Zhang-Abstract.asp

Zhang WJ. 2026c. The Niche-Energy-Time triadic synergy hypothesis (NET Hypothesis): A unified explanatory framework for determinants of species diversity. Computational Ecology and Software, 16(3): 243-272. http://www.iaees.org/publications/journals/ces/articles/2026-16(3)/3-Zhang-Abstract.asp

Zhang WJ. 2027a. General design principles for combining functional modules into complete cells: From compatibility to emergence. Network Biology, 17(2): 538-585. http://www.iaees.org/publications/journals/nb/articles/2027-17(2)/3-Zhang-Abstract.asp

Zhang WJ. 2027b. How biodiversity determines ecosystem stability at large scales? Multiscale diversity-stability cascade and the β diversity insurance hypothesis. Computational Ecology and Software, 17(2): 69-94. http://www.iaees.org/publications/journals/ces/articles/2027-17(2)/2-Zhang-Abstract.asp

Zhang WJ. 2027c. A hierarchical predictive theory of biodiversity change driven by environmental filtering of species traits: From environmental filtering to multi-scale dynamic integration. Proceedings of the International Academy of Ecology and Environmental Sciences, 17(1): 1-18. http://www.iaees.org/publications/journals/piaees/articles/2027-17(1)/1-Zhang-Abstract.asp

Zhang WJ, Bai CJ, Liu GD. 2007. Neural network modeling of ecosystems: A case study on cabbage growth system. Ecological Modelling, 201: 317-325. https://www.sciencedirect.com/science/article/pii/S0304380006004480

Zhang WJ, Barrion AT. 2006. Function approximation and documentation of sampling data using artificial neural networks. Environmental Monitoring and Assessment, 122: 185-201. https://link.springer.com/article/10.1007/s10661-005-9173-6

Zhang WJ, Chen B. 2011. Environment patterns and influential factors of biological invasions: a worldwide survey. Proceedings of the International Academy of Ecology and Environmental Sciences, 1(1): 1-14. http://www.iaees.org/publications/journals/piaees/articles/2011-1(1)/1-Zhang-Abstract.asp

Zhang WJ, Liu CH. 2012. Some thoughts on global climate change: will it get warmer and warmer? Environmental Skeptics and Critics, 1(1): 1-7. http://www.iaees.org/publications/journals/environsc/articles/2012-1(1)/1-Zhang-Abstract.asp

Zhang WJ, Zhong XQ, Liu GH. 2008. Recognizing spatial distribution patterns of grassland insects: neural network approaches. Stochastic Environmental Research and Risk Assessment, 22(2): 207-216, https://link.springer.com/article/10.1007/s00477-007-0108-3

Zhang WJ, Zhang XY. 2008. Neural network modeling of survival dynamics of holometabolous insects: a case study. Ecological Modelling, 211: 433-443. https://www.sciencedirect.com/science/article/pii/S030438000700498X