

TEMPO

Time-Indexed Evidence-Based Marginal Gains under Partial Observability

Compounding Marginal Gains Under Partial Observability
and Opportunity-Cost Constraints

Brent Figg

PhD in progress via the Global Executive Track pathway, Industrial and Systems Engineering
Wayne State University, Detroit, Michigan

*Working Paper Series—Systems Engineering Applied to
Complex Adaptive Decision Environments*

May 2026

Revised October 2026: author status and rights notice updated; content unchanged.

Rights, status and citation

© 2026 Brent Figg. Free to read, download, and cite with attribution; all other rights reserved.

Author status (as of October 2026): M.Sc. in Systems Engineering, Wayne State University; PhD in progress via the Global Executive Track pathway.

Cite as: Figg, B. (2026). TEMPO: Time-Indexed Evidence-Based Marginal Gains under Partial Observability. Working paper, May 2026 (rev. October 2026). Available at 2ndandcapital.com.

Abstract

In any system where time passes, resources are finite, interventions interact, and the value of action is partially observable—the failure to compound marginal gains systematically is not an oversight but an architectural absence. This paper proposes that architecture.

The central claim is that opportunity cost is the Markov process. As time passes without optimal action, configurations that were Pareto-reachable foreclose. The value of any intervention is not fixed—it decays, compounds, or becomes inaccessible depending on what other interventions are active, what constraints are binding, and whether the external opportunity windows through which gains can be realized remain open. This time-dependency is what makes the sequential, online, belief-updating architecture of a Partially Observable Markov Decision Process not merely convenient but structurally necessary. Without it, a one-shot optimization would suffice.

The primary theoretical contribution of this paper is the formalisation of opportunity cost as a time-indexed component of the MDP reward function, mediated by latent opportunity windows modelled as a Hidden Markov Model. This is not a discount factor or a heuristic priority flag—it is a computable, belief-dependent quantity that drives the Markov policy toward timely action on closing opportunities and degrades gracefully to standard MOBO when no windows are believed open. The HMM window belief state and the urgency score derived from it are the architectural additions not present in existing Bayesian optimisation or MDP-applied literature.

TEMPO (Time-indexed Evidence-based Marginal gains under Partial Observability) instantiates this contribution through three integrated components: (1) Gaussian Process Bayesian optimisation and hierarchical Bayesian modelling, which maintain calibrated probability distributions over an unknown opportunity surface and update them as observations arrive; (2) Markov Decision Process formalisation, in which the GP posterior is the sufficient Markov belief state, the acquisition function is the policy, the Matérn-5/2 update is a Kalman filter step, and decision-maker overrides are formal POMDP belief updates; and (3) the VADM/LDW structured decision interface, which functions as a *governance layer*—translating probabilistic Pareto outputs from the MDP engine into auditable, multi-criteria choice alternatives that decision-makers can challenge, weight, and defend. VADM/LDW is not a methodological contribution of this paper; it is the institutional accountability mechanism through which the MDP’s recommendations enter organisational decision processes.

The framework does not specify what type of interventions, what sector, or what data sources are involved. It requires only that: interventions have measurable effects, observations are available to update the belief state, and time creates differential value across possible actions. These conditions hold across high-performance sport development, clinical protocol optimization, defense resource allocation, organizational capability development, and any complex adaptive system where marginal gains are available but unconnected.

High-performance sport development is used as the primary illustrative instantiation—a domain where decision frequency is high, data density is tractable, and the compounding arithmetic of sub-marginal gains is empirically documented—but the architectural claim precedes and subsumes it.

Contents

1	The General Problem	3
1.1	Opportunity Cost as the Organizing Condition	3
1.2	Why Sequential Structure Is Necessary	3
1.3	The Sub-Marginal Gains Opportunity	4
1.4	The TEMPO Response	4
1.5	A Minimal Running Example	5
2	Intellectual Foundations	6
2.1	Bayesian Optimisation: Policy Over an Unknown Surface	6
2.1.1	The Optimisation Problem	6
2.1.2	Gaussian Process Surrogate Modelling	6
2.1.3	Acquisition Functions and the Explore–Exploit Tradeoff	7
2.1.4	Multi-Objective Bayesian Optimisation	7
2.2	Markov Decision Process Formalisation	7
2.2.1	TEMPO as a Markov Decision Process	7
2.2.2	The POMDP Extension	8
2.2.3	The Matérn-5/2 Kernel and Kalman Equivalence	8
2.2.4	Non-Markovian Structure and the HMM Extension	9
2.3	Structured Multi-Criteria Decision Analysis: VADM and LDW	9
2.3.1	Core Principles	9
2.3.2	Logical Decisions for Windows	10
2.4	Synthesis: Contribution Hierarchy and Component Mapping	10
3	The General Framework	11
3.1	The Opportunity Surface	11
3.2	Intervention Typology	12
3.3	Constraint Architecture	12
3.4	Time-Indexed Reward and Opportunity Cost Formalization	13
3.4.1	The Standard Reward and its Limitation	13
3.4.2	The Opportunity Cost of Delay	13
3.4.3	Time-Indexed Reward Function	13
3.4.4	The Opportunity-Cost-Augmented Acquisition Function	14
3.5	HMM Belief State over Opportunity Windows	14
3.5.1	Latent Opportunity Window States	14
3.5.2	HMM Formulation	15
3.5.3	Belief Update via the Forward Algorithm	15
3.5.4	Urgency and Foreclosure	15
4	Platform Architecture	16
4.1	Layer 1: Evolving Data Pipeline	16
4.1.1	Distinction from Batch Processing	16
4.1.2	Data Stream Taxonomy	17
4.2	Layer 2: Bayesian Optimisation Engine	17

4.2.1	The Opportunity Surface Model	17
4.2.2	Hierarchical Bayesian Modelling for Population Inference	18
4.2.3	The MOBO Prescription Loop with Opportunity-Cost Reward	18
4.3	Layer 3: Structured Decision Interface	18
5	Mathematical Specification	19
5.1	The Joint Belief State	19
5.2	Temporal Discounting and Phase Sensitivity	20
5.3	The “Add Value” Step as Constrained Bayesian Search	20
5.4	Missing Data and Incomplete Observations	20
6	VADM Integration: From Posterior to Decision	20
6.1	Mapping GP Posteriors to LDW Value Functions	20
7	Application: High-Performance Organisational Pipelines	21
7.1	The Development Pipeline as an MDP	21
7.2	Variable Entry and the Cold-Start Problem	22
7.3	Development Phase Structure and Opportunity Cost	22
7.4	The Compounding Arithmetic in Development Pipelines	23
7.5	Selection Decisions Under Uncertainty	24
7.6	Illustrative Decision Scenarios	25
7.6.1	Scenario 1: Variable-Entry Participant with Limited History	25
7.6.2	Scenario 2: Multi-Candidate Selection Under Uncertainty	25
8	Discussion	25
8.1	Positioning Against Alternative Frameworks	25
8.2	The Epistemological Advance	26
8.3	The Human-Decision-Maker Authority Structure	27
8.4	The Integration Thesis and Contribution Summary	27
8.5	Limitations and Boundary Conditions	27
9	Conclusion	28

The General Problem

Opportunity Cost as the Organizing Condition

Consider a system—an organization, a development program, a clinical operation, a competitive unit—that must make sequential decisions about where to direct limited resources across a set of possible interventions. Each intervention moves the system across a multi-dimensional performance surface. Each dimension of that surface represents something the system cares about: output quality, risk exposure, capability development, cost efficiency, resilience. No single intervention dominates all others on all dimensions simultaneously. Tradeoffs are unavoidable.

Now introduce time. The performance surface is not static. Configurations that are Pareto-reachable today may not be reachable next quarter. External conditions open and close windows within which certain interventions generate value they cannot generate outside those windows. Interventions that were sub-threshold in isolation become decisive when combined—but only if the combination is assembled before the window closes.

This is the general problem. It is not a sport problem or a healthcare problem or a defense problem. It is the problem of *sequential resource allocation across a partially observable opportunity surface under time-dependent opportunity cost*. What makes it hard is not any one of its features in isolation—uncertainty, multi-criteria tradeoffs, time-dependency, interaction effects—but their simultaneous presence. Most decision frameworks address one or two of these features. None address all four in an integrated, learning architecture.

The Architectural Absence

The failure to compound marginal gains systematically is not a failure of effort or intention. Organizations pursuing micro-optimisations across disconnected decision processes are already doing the right thing in isolation. The failure is architectural: there is no integration layer that identifies which marginal gains are currently available, prioritises them by expected compounded return given what opportunity windows are open, compounds them simultaneously rather than sequentially, and updates that prioritisation as evidence arrives and windows evolve. Without this layer, gains that multiply in combination are managed as if they add in isolation.

Why Sequential Structure Is Necessary

A natural question is whether the problem could be solved by a one-shot optimization: gather all available data, solve once, implement. The answer is structurally no, and the reason is opportunity cost.

Theorem 1 (Necessity of Sequential Architecture). If the value of all interventions is time-independent and the opportunity surface does not evolve, a one-shot optimization at $t = 0$ with complete data is sufficient. Sequential architecture is necessary if and only if at least one of the following holds:

1. Intervention values are time-dependent (opportunity windows open and close);
2. The opportunity surface evolves with time, independently of interventions;

3. New observations arrive that materially update the belief state over the surface.

All three conditions hold in any realistic complex adaptive system. Condition (3) is addressed by Bayesian sequential updating. Conditions (1) and (2) require time-indexed reward and opportunity window belief state, formalized in Section 3.4.

The implication is that the MDP architecture is not a choice of modelling convenience. It is the minimal structure required to address the problem as stated.

The Sub-Marginal Gains Opportunity

At the elite tail of any performance distribution—athletic, clinical, organizational, operational—the gap between leading and lagging performance is not explained by large structural differences but by the accumulated effect of dozens of micro-optimisations that compound across time. This is empirically documented in high-performance sport, where the physiological gap between a world finalist and a world champion is routinely less than 3% across measurable capacity variables, with the competitive margin produced by accumulated sub-marginal gains across controllable variables.

The same structure appears in clinical outcomes, where small improvements in adherence, protocol timing, and care pathway routing compound into large differences in population-level efficacy. It appears in organizational capability development, where marginal improvements in process efficiency, decision quality, and personnel allocation compound into large differences in institutional output. It appears in defense resource allocation, where marginal improvements in readiness, positioning, and logistics compound into decisive operational advantages.

The mathematics is the same in every case. If a system identifies M independently controllable variables each yielding a marginal improvement of δ , the compounded effect is not $M \cdot \delta$ but:

$$\prod_{m=1}^M (1 + \delta_m) \quad (1)$$

For $M = 6$ and $\delta_m = 0.015$, this yields a compounded advantage of $1.015^6 \approx 1.093$ —a 9.3% performance differential from six individually modest gains. At the right tail of any performance distribution, this is not marginal. It is decisive.

The unrealised opportunity in every sector is not the identification of additional marginal gains. It is the integration of gains already being pursued in isolation into a unified system that detects their interactions, respects their constraints, and compounds them simultaneously.

The TEMPO Response

TEMPO addresses this problem through a three-layer architecture: an *Evolving Data Pipeline* that maintains a live Bayesian belief state over the opportunity surface; a *Bayesian Optimisation Engine* implementing a near-optimal Markov policy over the intervention

space; and a *Structured Decision Interface* that translates probabilistic Pareto outputs into structured multi-criteria choices that human decision-makers can audit, challenge, and approve.

The framework does not prescribe what data is observed, what interventions are available, or what sector the system operates in. It prescribes the architecture through which any data, any intervention set, and any sector can be brought into a unified compounding system.

A Minimal Running Example

To make the architecture concrete before the formal treatment, consider the following minimal, domain-agnostic scenario used as a thread throughout the paper.

Running Example: Three Resources, Two Windows

A development programme has three independently controllable intervention resources: x_1 (capacity-building intensity), x_2 (technical process investment), and x_3 (recovery or consolidation allocation). Two opportunity windows are relevant: W_1 (an external selection or evaluation event, opening at week 8 and closing at week 16) and W_2 (a resource injection cycle, open only during weeks 1–4).

At $t = 0$, the GP belief state over the opportunity surface is initialised from population data. The system does not know the true response surface—it knows only that past units responded well to moderate x_1 combined with high x_2 , with considerable individual variation. Window W_2 is believed open ($b_2(0) = 0.91$); window W_1 is believed closed ($b_1(0) = 0.12$).

The MDP policy at $t = 0$ therefore augments the acquisition function toward configurations that exploit the open W_2 window—specifically, configurations that invest heavily in x_2 while the resource injection is available, even at a marginal cost to x_1 . By week 4, W_2 closes. $b_1(t)$ rises as week 8 approaches. The policy shifts: x_1 -maximising configurations now receive augmented reward through the W_1 alignment term, driving the system toward capacity-peaking configurations timed to the evaluation window.

Without the HMM window belief, the acquisition function treats all weeks symmetrically. The compounding gain from sequencing x_2 investment before x_1 -peaking is invisible to the standard MOBO policy. The opportunity cost of missequencing is not detected until after the windows close.

This example recurs in later sections as specific architectural components are introduced. The numbers are illustrative; the structural logic is exact.

Intellectual Foundations

Bayesian Optimisation: Policy Over an Unknown Surface

The Optimisation Problem

Let $f : \mathcal{X} \rightarrow \mathbb{R}$ be an unknown objective function mapping an intervention configuration $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d$ to a performance outcome. The configuration space $\mathcal{X}$ is defined by d controllable variables whose specific identity is domain-dependent. The goal is to find:

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (2)$$

subject to constraint functions $g_k(\mathbf{x}) \leq 0$ encoding resource bounds, structural limits, and opportunity-window dependencies. The function f is *black box*: it cannot be evaluated analytically, each evaluation is expensive, and evaluations are noisy ($y = f(\mathbf{x}) + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$). This is precisely the setting for Bayesian optimisation—and, as established in Section 2.2, a specific instance of a Partially Observable Markov Decision Process.

Gaussian Process Surrogate Modelling

Bayesian optimisation maintains a Gaussian Process prior over f , characterised by a mean function $\mu(\mathbf{x})$ and a covariance kernel $k(\mathbf{x}, \mathbf{x}')$:

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (3)$$

The Matérn-5/2 kernel, appropriate for performance surfaces that are twice-differentiable but not overly smooth, is:

$$k_{\text{M52}}(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2} \right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right), \quad r = \|\mathbf{x} - \mathbf{x}'\| \quad (4)$$

Given n observed pairs $\mathcal{D}_n = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, the GP posterior is analytically tractable:

$$\mu_n(\mathbf{x}) = \mu(\mathbf{x}) + \mathbf{k}_n(\mathbf{x})^\top (\mathbf{K}_n + \sigma_n^2 \mathbf{I})^{-1} (\mathbf{y} - \boldsymbol{\mu}_n) \quad (5)$$

$$\sigma_n^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \mathbf{k}_n(\mathbf{x})^\top (\mathbf{K}_n + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_n(\mathbf{x}) \quad (6)$$

The posterior mean $\mu_n(\mathbf{x})$ is the current best estimate of system response to configuration $\mathbf{x}$; the posterior variance $\sigma_n^2(\mathbf{x})$ quantifies the epistemic uncertainty remaining at that configuration.

Acquisition Functions and the Explore–Exploit Tradeoff

At each step, a new configuration is chosen by maximising an acquisition function $\alpha(\mathbf{x}; \mathcal{D}_n)$. The Upper Confidence Bound:

$$\alpha_{\text{UCB}}(\mathbf{x}) = \mu_n(\mathbf{x}) + \kappa \sigma_n(\mathbf{x}), \quad \kappa > 0 \quad (7)$$

The Expected Improvement:

$$\alpha_{\text{EI}}(\mathbf{x}) = \mathbb{E}[\max(0, f(\mathbf{x}) - f(\mathbf{x}^+))] \quad (8)$$

Within the MDP formalisation of Section 2.2, these are *policies* mapping belief states to actions. Thompson Sampling—drawing a function from the current GP posterior and acting on its optimum—is the canonical posterior-sampling policy for Bayesian MDPs, achieving near-optimal regret bounds (Srinivas et al., 2012) because it treats exploration–exploitation as a Markov policy rather than a heuristic.

Multi-Objective Bayesian Optimisation

System performance is multi-dimensional. Let $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_K(\mathbf{x}))$ denote a vector of objective functions whose specific content is domain-dependent but whose structure is invariant. Multi-Objective Bayesian Optimisation (MOBO) maintains a separate GP surrogate for each objective and identifies the *Pareto frontier*: the set of configurations where no objective can be improved without degrading another. The Expected Hypervolume Improvement acquisition function identifies the next query most likely to expand the dominated hypervolume:

$$\alpha_{\text{EHVI}}(\mathbf{x}) = \mathbb{E}[\text{HV}(\mathcal{P}_n \cup \{\mathbf{f}(\mathbf{x})\}) - \text{HV}(\mathcal{P}_n)] \quad (9)$$

where $\mathcal{P}_n$ is the current Pareto set and $\text{HV}(\cdot)$ is the dominated hypervolume relative to a reference point. Independent surrogates per objective are used for tractability and interpretability; inter-objective correlations are captured downstream through the joint Pareto structure rather than the kernel, which preserves the computational advantages of independent GP inference without discarding the tradeoff geometry that multi-objective optimisation requires. In the time-indexed reward formulation of Section 3.4, this acquisition function is augmented with an opportunity-cost term that conditions the value of any query on the current belief state over open opportunity windows.

Markov Decision Process Formalisation

TEMPO as a Markov Decision Process

Definition 1 (TEMPO MDP). The TEMPO prescription system constitutes an MDP $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R})$:

- **State** $s_t = (\mathcal{D}_n, \mathcal{P}_n)$: the GP posterior belief state and current Pareto set.
- **Action** $a_t = \mathbf{x}_{n+1}$: the intervention configuration selected for the next period.

- **Transition** $\mathcal{T}(s_{t+1} | s_t, a_t)$: the GP posterior update upon observing y_{n+1} .
- **Reward** $\mathcal{R}(s_t, a_t, t)$: the time-indexed value of action a_t in state s_t —formalised in Section 3.4.

The Markov property holds in the space of *posterior distributions*. The GP posterior at time t is a sufficient statistic for all prior observations: $p(f | \mathcal{D}_{t+1})$ depends only on $p(f | \mathcal{D}_t)$ and the new observation, and nothing else. This is the Markov property in the space of probability measures—and it is why the event-driven, streaming pipeline of Layer 1 is a formal architectural requirement for maintaining a valid belief state, not merely an engineering preference.

MDP Component Mapping

MDP Component	TEMPO Equivalent
State s_t	GP posterior $\mathcal{D}_n$ + Pareto set $\mathcal{P}_n$
Action a_t	Intervention configuration $\mathbf{x}_{n+1}$
Transition $\mathcal{T}(s, a)$	GP posterior update (Kalman correction)
Reward $\mathcal{R}(s, a, t)$	Time-indexed EHVI with opportunity cost
Policy π	Acquisition function (Thompson Sampling optimal)
Belief update	Kalman filter / GP posterior correction

The POMDP Extension

The full TEMPO system is a Partially Observable MDP since the true opportunity surface f is never directly observed—only noisy realisations are available.

Definition 2 (TEMPO POMDP). • **Hidden state**: the true opportunity surface f and latent opportunity window states $\mathbf{O}_t$ (Section 3.5).

- **Observation**: noisy outcome $y = f(\mathbf{x}) + \varepsilon$ and observable window signals $\mathbf{e}_t$.
- **Belief state**: the compound belief $b_t = (p(f | \mathcal{D}_t), \mathbf{b}_O(t))$ —GP posterior over the surface and HMM forward belief over opportunity windows.

The compound belief state is the sufficient statistic for all observations. Human decision-maker overrides—rejecting or modifying a recommended configuration—are formal POMDP belief updates that modify the transition model, gradually encoding tacit expertise within the platform’s formal architecture.

The Matérn-5/2 Kernel and Kalman Equivalence

The Matérn-5/2 kernel has an exact equivalence to a 3rd-order linear stochastic differential equation whose solution process is Markovian in an augmented state space.

Proposition 1 (Kalman Equivalence). Every GP posterior update under the Matérn-5/2 kernel is formally equivalent to a Kalman filter prediction–correction step in the corresponding 3rd-order state-space model, with computational complexity $O(n)$ rather than $O(n^3)$ for the temporal streaming component of the data pipeline.

This equivalence means that every observation arriving at Layer 1 is not merely a database

row but a Kalman correction updating the Markov belief state. It also partially resolves the GP scalability limitation (Section 8.5): the inducing-point approximation addresses multi-dimensional input spaces; the Kalman representation addresses the temporal streaming dimension exactly.

Non-Markovian Structure and the HMM Extension

The hierarchical population model (Section 4.2.2) introduces inter-unit coupling that breaks strict individual-level Markov structure: when unit j produces an unusual response, it shifts the population hyperprior and therefore the prior for unit i without unit i having produced new data. At the population level, TEMPO is a *multi-agent MDP with shared belief state*—the sharing mechanism through which sub-marginal gain evidence propagates across units simultaneously.

The phase-indexed kernel (Section 5.2) creates conditionally Markovian chains within phases with no modelled transition dynamics between them. The natural extension is a Hidden Markov Model over latent phase (or opportunity) state, with observable signals as emissions—allowing the system to detect state transitions from data rather than relying on externally imposed calendar boundaries. This HMM structure is formalized in Section 3.5.

Structured Multi-Criteria Decision Analysis: VADM and LDW

VADM/LDW functions in TEMPO as a *governance interface*, not as a methodological contribution. It is the mechanism through which the MDP engine’s probabilistic Pareto outputs enter organisational decision processes in a form that is auditable, defensible to stakeholders, and responsive to institutional value structures that the formal model cannot fully encode. This role is important precisely because decisions made by or with the system must be accountable—to participants, administrators, and governing bodies. The MDP recommends; the governance interface ensures that the recommendation is evaluated, challenged, and owned by a human decision-maker.

Core Principles

Chelst and Canbolat’s (2011) VADM framework rests on the distinction between *facts* (probabilistic performance estimates) and *value judgments* (weights on competing criteria)—a distinction routinely violated in operational decision-making across sectors. Most decisions conflate the two, with the result that disagreements about values are argued as if they were disagreements about facts, and genuine uncertainty is suppressed to maintain the appearance of decisive authority.

VADM structures decisions through five stages: define the decision; identify criteria; gather data as probability distributions; elicit preferences through structured weighting; and—crucially—*add value* by constructing a modified alternative that improves on the weakest attribute of the leading option while preserving its strengths. The “add value” step is the most distinctive contribution: it shifts the decision paradigm from passive selection to active design.

Logical Decisions for Windows

LDW v8.0, now open-source under the MIT licence, implements Multi-Objective Decision Analysis (MODA). Each criterion c_j is associated with a value function $v_j : X_j \rightarrow [0, 1]$. A weight vector $\mathbf{w}$ with $\sum_j w_j = 1$ captures relative importance. The overall value of alternative i is:

$$V_i = \sum_{j=1}^m w_j v_j(x_{ij}) \quad (10)$$

Sensitivity analysis over $\mathbf{w}$ reveals which criteria are decision-critical—an essential transparency feature wherever decisions must be defensible to stakeholders. In the TEMPO integration, criterion scores x_{ij} are supplied not by expert elicitation but by the GP posterior means of the MDP belief state, giving the static LDW framework a continuously learning evidence substrate.

Synthesis: Contribution Hierarchy and Component Mapping

The three components of TEMPO are not co-equal in their theoretical contribution. The MDP/POMDP formalisation with time-indexed opportunity-cost reward is the primary intellectual contribution. MOBO is the computational realisation of the MDP policy and is subsumed by that framing. VADM/LDW is the governance interface through which MDP outputs enter institutional decision processes—a critical practical component, but not a methodological novelty of this paper. Making this hierarchy explicit forestalls reviewer confusion about what is new versus what is well-integrated.

Contribution Hierarchy and Component Mapping

Primary contribution (novel):

- Opportunity cost as a time-indexed MDP reward component (equations 12–14)
- HMM belief state over latent opportunity windows driving policy urgency (Section 3.5)
- Foreclosure probability as a computable urgency signal (equation 18)

Secondary contribution (formalisation and integration):

- GP posterior as sufficient Markov belief state; acquisition function as near-optimal MDP policy
- Matérn-5/2 / Kalman equivalence applied to streaming belief updates
- Human override as POMDP belief update on transition model

Governance layer (not a methodological contribution):

- VADM/LDW as the structured interface between MDP outputs and institutional decision processes; provides auditability, weight elicitation, and sensitivity analysis

Component mapping:

MDP Component	TEMPO Equivalent
Belief state b_t	Compound GP + HMM posterior
Policy π	Opportunity-cost-augmented acquisition function
Reward $\mathcal{R}(s, a, t)$	EHVI $\times$ constraint satisfaction + window-weighted value $-$ resource cost
Belief update	Kalman prediction-correction (GP) + forward algorithm (HMM)
Override	POMDP transition model update
VADM criteria	MOBO objective functions (one GP surrogate each)
VADM performance estimates	GP posterior mean $\mu_n(\mathbf{x})$
VADM uncertainty	GP posterior variance $\sigma_n^2(\mathbf{x})$
VADM “add value” step	Constrained Bayesian search (Section 5.3)
LDW sensitivity analysis	Posterior predictive variance over weight vector

The MDP formalism is the unifying theoretical identity. The Bayesian engine populates the VADM/LDW governance interface with live, uncertainty-quantified data derived from the belief state. The novelty is not in any single component but in the coupling: making opportunity cost computable, dynamic, and belief-dependent within the same architecture that compounds marginal gains and translates recommendations into auditable institutional choices.

The General Framework

The Opportunity Surface

At any time t , a system occupies a position in a multi-dimensional space of achievable states. The *opportunity surface* is the function $\mathbf{f} : \mathcal{X} \rightarrow \mathbb{R}^K$ mapping intervention configurations to vectors of objective outcomes. It has three properties that define the structural problem:

Partial observability. The true surface is unknown. Only noisy realisations at queried configurations are observable. The GP surrogate is the belief state over this surface—it does not know, and does not need to know, whether the dimensions represent physiological capacity, clinical outcomes, organizational throughput, or operational readiness.

Time-evolution. The surface changes as time passes, independent of interventions. Configurations that are Pareto-reachable at t may not be reachable at $t + \Delta t$. The expected cost of reaching t without optimal action is the opportunity cost of delay—the core quantity formalized in Section 3.4.

Interaction structure. Interventions interact. Gains that are independent in isolation may reinforce or interfere when combined. The GP surrogate learns this interaction structure from data; the MOBO engine avoids infeasible interaction corners of the configuration

space.

Intervention Typology

The configuration vector $\mathbf{x} \in \mathcal{X}$ spans a space of possible interventions. Across domains, interventions cluster into four structural classes whose mathematical treatment is identical regardless of their domain-specific content:

Table 1: General Intervention Typology with Sector Instantiations

Class	Sport	Clinical	Organizational	Defense
Capacity-building	Training load volume and modality	Therapy protocol intensity and frequency	Workforce development investment	Training tempo and readiness
Process / technical	Skill and technical session design	Procedural adherence and protocol routing	Process standardization and workflow redesign	Doctrine revision and tactical procedure
Strategic / positioning	Periodization and competition calendar	Care pathway architecture and referral structure	Resource allocation and organizational design	Pre-positioning, logistics, and force structure
Behavioral / psychological	Recovery architecture and athlete well-being	Adherence motivation and behavioral support	Culture, leadership, and incentive design	Unit cohesion and human performance

The GP surrogate is agnostic to intervention class. The MOBO engine searches the configuration space without reference to which class a variable belongs to. The intervention typology is a communication layer for practitioners and decision-makers, not a constraint on the mathematical architecture.

Constraint Architecture

The constraint functions $g_k(\mathbf{x}) \leq 0$ in the MOBO optimisation encode three structurally distinct constraint classes, each present in every sector:

Resource Constraints.. Finite resources—financial, temporal, personnel—bound the feasible set of intervention configurations. The cost-benefit structure of any intervention is capturable as an objective function f_k or a hard constraint g_k , depending on whether resource expenditure is to be traded off against other objectives or held below a binding limit. These are the constraints most familiar to operations research; their treatment in TEMPO is standard. Time is itself a resource constraint: the temporal cost of an intervention competes with the opportunity cost of delay.

Structural Constraints.. Institutional, physical, and regulatory limits define what configurations are feasible independent of resources. In sport: facility availability, athlete welfare regulations, governing body rules. In clinical settings: regulatory approval, treatment capacity, clinical trial protocols. In organizations: legal compliance, physical

infrastructure. In defense: rules of engagement, materiel availability, alliance obligations. These constraints bound the feasible set $\mathcal{X}$ directly.

Emergent Opportunity Constraints.. The third class is the least commonly formalized and the most consequential for the time-indexed reward structure. Some configurations are feasible and resource-compatible but yield value only within specific temporal windows driven by external conditions outside the system’s control: a competition calendar, a funding cycle, a market opening, a clinical trial enrollment window, a geopolitical moment. Outside these windows, the intervention either cannot be executed or produces materially diminished returns. Inside the window, the same intervention may be decisive.

Emergent opportunity constraints are not static bounds on the configuration space. They are *time-varying modifications of the reward function*: the value of a given action depends on whether the relevant window is currently open. This requires a separate belief state over window status, formalized in the HMM of Section 3.5, and a time-indexed modification of the acquisition function, formalized in Section 3.4.

Time-Indexed Reward and Opportunity Cost Formalization

The Standard Reward and its Limitation

The standard MOBO acquisition function selects the next intervention configuration by maximising the Expected Hypervolume Improvement:

$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x}} \alpha_{\text{EHVI}}(\mathbf{x}; \mathcal{D}_n) \quad (11)$$

This reward is time-agnostic: it evaluates the expected gain from querying $\mathbf{x}$ given the current belief state, without accounting for the opportunity cost of having reached time t without optimal prior action, or for whether the window within which this gain can be realized remains open.

The Opportunity Cost of Delay

Define the *value function* $V^*(s_t)$ as the expected cumulative reward under the optimal policy from state s_t onward. The *opportunity cost of delay* from t to $t + 1$ is:

$$\Delta_{\text{OC}}(s_t) = V^*(s_t) - \mathbb{E}[V^*(s_{t+1}) \mid \text{no action at } t] \quad (12)$$

This quantity captures the expected value of the configurations that become unreachable by waiting one period. It is strictly positive whenever opportunity windows are closing or the surface is evolving adversely. It is the formal expression of the intuition that delay has a cost that is not visible in the current-period EHVI alone.

Time-Indexed Reward Function

The time-indexed reward for action $\mathbf{x}$ at time t in state s_t is:

$$\mathcal{R}(s_t, \mathbf{x}, t) = \underbrace{\alpha_{\text{EHVI}}(\mathbf{x}; \mathcal{D}_n) \cdot \prod_j \hat{P}_j(\mathbf{x})}_{\text{immediate gain}} + \underbrace{\sum_{k=1}^W b_k(t) \cdot \eta_k(\mathbf{x})}_{\text{window-weighted value}} - \underbrace{\lambda \cdot c(\mathbf{x})}_{\text{resource cost}} \quad (13)$$

where:

- $\hat{P}_j(\mathbf{x}) = \Phi(-\mu_n^{g_j}(\mathbf{x})/\sigma_n^{g_j}(\mathbf{x}))$ is the estimated probability that structural constraint g_j is satisfied.
- $b_k(t) = P(O_{k,t} = \text{open} \mid \mathbf{e}_{1:t})$ is the HMM forward belief that opportunity window k is currently open (Section 3.5).
- $\eta_k(\mathbf{x})$ is the *window alignment* of configuration $\mathbf{x}$ —the expected marginal value of action $\mathbf{x}$ attributable specifically to window k being open. η_k may be estimated from posterior sensitivity analysis (comparing GP posterior means under open versus closed window conditioning), from historical counterfactuals where window status varied, or from domain-specified mappings where the window’s value implications are known in advance. The architecture is agnostic to the estimation method; any approach that produces a non-negative function of $\mathbf{x}$ with higher values for configurations that require the window to be open is admissible.
- $\lambda > 0$ is the resource cost weight and $c(\mathbf{x})$ is the cost of executing configuration $\mathbf{x}$.

The second term is the key structural addition. It makes the reward function explicitly dependent on the current belief over opportunity window states: configurations that align with open windows receive augmented reward; configurations that are feasible but window-misaligned receive standard EHVI reward only. This coupling is what makes opportunity cost computationally tractable within the MOBO framework.

The Opportunity-Cost-Augmented Acquisition Function

The modified acquisition function is:

$$\alpha_{\text{OC}}(\mathbf{x}; t) = \alpha_{\text{EHVI}}(\mathbf{x}; \mathcal{D}_n) \cdot \prod_j \hat{P}_j(\mathbf{x}) + \sum_{k=1}^W b_k(t) \cdot \eta_k(\mathbf{x}) - \lambda \cdot c(\mathbf{x}) \quad (14)$$

When all window beliefs are low ($b_k(t) \approx 0$ for all k), this reduces to the standard constraint-penalised EHVI. When a window is believed open ($b_k(t) \rightarrow 1$), configurations aligned with that window receive an additive reward that reflects the opportunity cost of failing to act while it is open. The urgency of action under an open window is therefore a function of the belief state, not a manually imposed priority—it emerges from the POMDP structure.

HMM Belief State over Opportunity Windows

Latent Opportunity Window States

Let W be the number of opportunity windows relevant to the system. Each window $k \in \{1, \dots, W\}$ has a latent state $O_{k,t}$ taking values in {closed, opening, open, closing}.

The latent state captures the lifecycle of an opportunity: it opens, reaches full availability, begins closing, and forecloses—with transition dynamics that may be calendar-driven, externally triggered, or data-inferred.

HMM Formulation

The window state process is modelled as a Hidden Markov Model:

$$P(O_{k,t+1} \mid O_{k,t}) = \mathbf{A}_k \quad (15)$$

$$P(\mathbf{e}_{k,t} \mid O_{k,t}) = \mathbf{B}_k \quad (16)$$

where $\mathbf{A}_k$ is the transition matrix and $\mathbf{B}_k$ is the emission distribution for observable window signals $\mathbf{e}_{k,t}$. Observable signals are domain-dependent: in sport, competition calendars and selection announcement schedules; in clinical contexts, trial enrollment windows and protocol review cycles; in organizations, funding cycles and strategic planning timelines; in defense, operational tempo indicators and intelligence assessments.

The HMM does not require that these signals be precisely defined at design time. Any observable with predictive power over window status can serve as an emission. The system learns the emission distribution from historical co-occurrence of signals and window states.

Belief Update via the Forward Algorithm

The forward belief over window k is updated recursively:

$$b_k(t) \propto P(\mathbf{e}_{k,t} \mid O_{k,t} = \text{open}) \sum_{o'} P(O_{k,t} = \text{open} \mid O_{k,t-1} = o') \cdot b_k^{o'}(t-1) \quad (17)$$

where $b_k^{o'}(t-1)$ is the belief mass on state o' at $t-1$. This update is the temporal Bayesian filter for the HMM—it is the opportunity-window analogue of the Kalman filter for the GP surrogate. Together they define the compound belief state of the full POMDP: $(p(f \mid \mathcal{D}_t), \{b_k(t)\}_{k=1}^W)$.

Urgency and Foreclosure

The urgency of acting on window k is the product of its current belief $b_k(t)$ and its expected value $\eta_k(\mathbf{x}_k^*)$, where $\mathbf{x}_k^*$ is the optimal configuration for capturing window k 's value:

$$U_k(t) = b_k(t) \cdot \eta_k(\mathbf{x}_k^*) \cdot P(\text{foreclosure before } t + \Delta) \quad (18)$$

The foreclosure probability $P(\text{foreclosure before } t + \Delta)$ is derived from the HMM transition matrix $\mathbf{A}_k$: it is the probability that $O_{k,t}$ transitions to *closed* within horizon Δ given current belief. High urgency configurations—those with high current belief, high window-specific value, and high foreclosure risk—receive the strongest augmentation to their acquisition function score, driving the Markov policy toward timely action. The system thus formalises the intuitive judgment that some opportunities are worth acting on now even if their expected EHVI is below threshold, because the alternative is foreclosure.

Platform Architecture

The TEMPO architecture has three functional layers. Their labels and data flows are general; domain-specific variable names are substituted at instantiation time without altering the architecture.

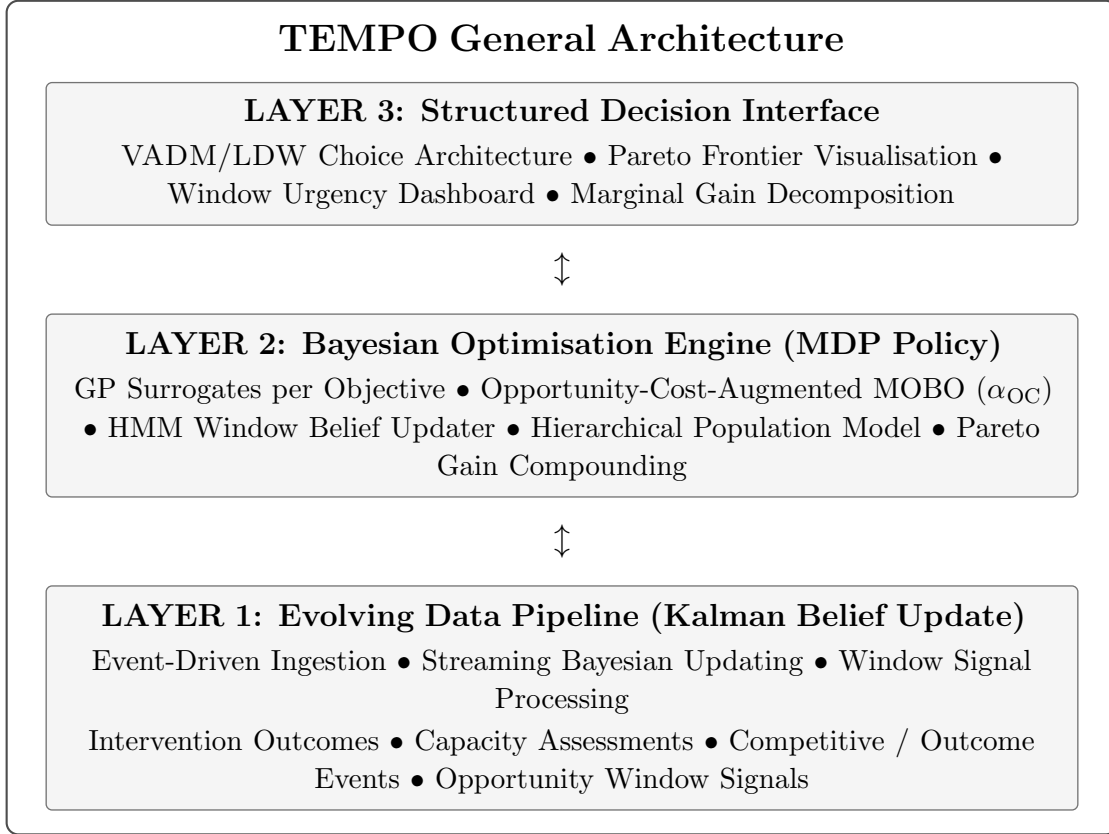


Figure 1: Three-layer architecture of TEMPO. Upward arrows: data propagates as Markov belief-state updates (GP correction + HMM forward pass). Downward arrows: MDP policy actions and structured decision proposals. All layer labels are domain-agnostic; sector instantiation substitutes domain-specific variable names into the general slots.

Layer 1: Evolving Data Pipeline

Distinction from Batch Processing

Batch ETL pipelines decouple observation time from belief-state update time. From the MDP perspective, this means the system takes actions under a stale belief state, forfeiting the regret improvement that accrues from prompt updates. In the opportunity-cost framework this is especially damaging: a window whose closing is signalled in the emission data on Monday may foreclose before Friday’s batch run updates the HMM belief state and augments the acquisition function accordingly.

The evolving data pipeline requires that each arriving observation immediately triggers: (1) a Kalman prediction–correction step updating the GP belief state over the opportunity

surface; and (2) a forward pass updating the HMM belief over window states. Both updates are $O(n)$ in the number of streaming observations under the Kalman and forward algorithm representations respectively.

Data Stream Taxonomy

Four general stream types feed Layer 1. Their domain-specific content varies; their structural role in the belief-state update is invariant.

Table 2: General Layer 1 Data Streams

Stream Type	Frequency	Content (domain-agnostic)	Belief Role
Intervention outcome	Per-action	Measured effect of executed intervention on each objective	Primary GP observation y
Capacity assessment	Periodic	Underlying system state on developable dimensions	Updates capacity-ceiling GP
Competitive / outcome event	Event-driven	External performance result; ground-truth validation	Validates performance GP
Window signal	Continuous	Observable indicators of opportunity window status: calendar events, market conditions, external triggers	HMM emission; updates $b_k(t)$

The window signal stream is the architectural addition not present in standard Bayesian optimisation platforms. It feeds the HMM forward algorithm continuously, ensuring the opportunity-cost augmentation of the acquisition function reflects current window belief rather than a stale prior.

Layer 2: Bayesian Optimisation Engine

The Opportunity Surface Model

The central object in Layer 2 is the multi-objective GP model of the opportunity surface: $\mathbf{f} : \mathcal{X} \rightarrow \mathbb{R}^K$ mapping intervention configurations to objective vectors. The K objectives are domain-defined; their structure in the GP is invariant. Typical structural objectives across sectors include:

- f_1 : Primary performance output (standardised outcome metric)
- f_2 : Risk exposure (negated; injury, adverse event, operational risk)
- f_3 : Capacity development rate (underlying capability growth)
- f_4 : Resource efficiency (cost-normalized output)

Each f_k is modelled by an independent GP surrogate. The multi-objective framing is essential: configurations that optimise f_1 alone may expose the system to unacceptable f_2 risk, while configurations that develop f_3 and f_4 simultaneously may be superior across the development horizon even at a marginal cost to current f_1 .

Hierarchical Bayesian Modelling for Population Inference

Individual unit data is sparse relative to the configuration space. The hierarchical Bayesian model borrows statistical strength across units in the same system, positing that each unit’s response surface $f^{(i)}(\mathbf{x})$ is a draw from a population-level distribution:

$$f^{(i)} \sim \mathcal{GP}(\mu_{\text{pop}}(\mathbf{x}), k_{\text{pop}}(\mathbf{x}, \mathbf{x}')) + \delta^{(i)}, \quad \delta^{(i)} \sim \mathcal{GP}(0, k_{\text{ind}}(\mathbf{x}, \mathbf{x}')) \quad (19)$$

where μ_{pop} and k_{pop} capture population-level mean response and $\delta^{(i)}$ is unit-specific deviation. This structure allows sub-marginal gain evidence detected across the population to propagate immediately to individual-level prescriptions through the shared belief state—the multi-agent MDP coupling.

The MOBO Prescription Loop with Opportunity-Cost Reward

Algorithm 1 TEMPO MOBO Prescription Loop (Opportunity-Cost-Augmented)

Require: Data $\mathcal{D}_n$, GP surrogates $\{\mathcal{GP}_k\}_{k=1}^K$, constraint surrogates $\{\mathcal{GP}_{g_j}\}$, Pareto set $\mathcal{P}_n$, weight vector $\boldsymbol{\lambda}$, window beliefs $\{b_k(t)\}_{k=1}^W$, window alignment functions $\{\eta_k\}$

Ensure: Next intervention configuration $\mathbf{x}_{n+1}$

- 1: Update GP hyperparameters θ_k via marginal likelihood
 - 2: Compute $\mu_n^{(k)}(\cdot)$, $\sigma_n^{(k)^2}(\cdot)$ via (5), (6) ▷ Kalman belief update
 - 3: Compute window urgency $U_k(t)$ via (18) for each window k ▷ HMM forward belief
 - 4: Compute $\hat{P}_j(\mathbf{x})$ for each constraint g_j
 - 5: Compute $\alpha_{\text{OC}}(\mathbf{x}; t)$ via (14) ▷ Time-indexed reward
 - 6: Solve $\mathbf{x}_{n+1} = \arg \max_{\mathbf{x} \in \mathcal{X}} \alpha_{\text{OC}}(\mathbf{x}; t) \cdot \boldsymbol{\lambda}^\top \boldsymbol{\mu}_n(\mathbf{x})$ ▷ MDP policy action
 - 7: Surface $\mathbf{x}_{n+1}$, Pareto frontier, window urgency scores, and marginal gain decomposition to Layer 3
 - 8: Receive decision-maker response: accept, modify, or override ▷ POMDP belief revision
 - 9: Record executed configuration $\mathbf{x}_{n+1}'$ and observe y_{n+1}
 - 10: Update $\mathcal{D}_{n+1} \leftarrow \mathcal{D}_n \cup \{(\mathbf{x}_{n+1}', y_{n+1})\}$; update HMM with $\mathbf{e}_{t+1}$ and repeat ▷ Joint belief update
-

Step 8 is where the POMDP structure earns its theoretical weight: the decision-maker’s override is a belief update modifying the transition model, not an exception to be discarded. Systematic overrides reveal latent expertise—tacit knowledge that the formal model has not yet captured—and the platform learns the decision-maker’s implicit preference structure as a posterior over the POMDP transition dynamics.

Layer 3: Structured Decision Interface

The design challenge of Layer 3 is the translation gap between MDP policy outputs—Pareto-optimal configuration vectors with probabilistic objective estimates and window urgency scores—and the structured decisions that human decision-makers must make and defend to institutional stakeholders.

The interface resolves this through two commitments. First, it is organized around the natural decision rhythm of the operating domain rather than the mathematical rhythm of the MDP: the cadence of daily, weekly, and periodic decisions that characterize the domain governs when and how the MDP outputs are surfaced. Second, when explicit multi-criteria decisions are required, the VADM/LDW architecture provides the formal scaffolding that prevents those decisions from collapsing into unreflective preference.

Design Principle 1. *Marginal gains must be made visible.* Every dashboard surface includes a rolling marginal gain decomposition: the system decomposes the expected performance improvement into contributions from each controlled variable, expressed in percentage terms, and displays the compounded total. The additive logic of marginal gains is explicit and auditable.

Design Principle 2. *Uncertainty must be visible but not paralysing.* Performance estimates are accompanied by confidence intervals derived from the GP posterior variance. Sub-marginal gain recommendations below the system’s current detection resolution are suppressed, with the detection threshold displayed. Window urgency scores are displayed alongside their belief uncertainty.

Design Principle 3. *Opportunity windows must be surfaced proactively.* Window urgency scores $U_k(t)$ are displayed prominently whenever a window is believed open and foreclosure risk is non-trivial. The interface distinguishes between deferred decisions (where delay costs are low) and time-critical decisions (where window closure risk makes delay expensive), preventing the architecture from producing urgent recommendations that are indistinguishable in presentation from routine ones.

Design Principle 4. *Recommendations are proposals, not mandates.* Every MOBO-generated configuration is a candidate with projected Pareto performance and window alignment displayed. Accept, modify, and override responses are all recorded and constitute POMDP belief updates.

Design Principle 5. *Decision structure is surfaced on demand.* The full VADM/LDW interface—weight elicitation, value function editing, sensitivity analysis, Pareto frontier exploration—is available from any view but is not the default presentation.

Mathematical Specification

The Joint Belief State

The core mathematical object is the joint posterior over the opportunity surface and opportunity window states given all accumulated data. For unit i at time t :

$$p(\mathbf{f}^{(i)}, \boldsymbol{\theta} \mid \mathcal{D}_t^{(i)}, \mathcal{D}_t^{\text{pop}}) \propto p(\mathcal{D}_t^{(i)} \mid \mathbf{f}^{(i)}, \boldsymbol{\theta}) \cdot p(\mathbf{f}^{(i)} \mid \boldsymbol{\theta}) \cdot p(\boldsymbol{\theta} \mid \mathcal{D}_t^{\text{pop}}) \quad (20)$$

The compound belief state $b_t = (p(\mathbf{f}^{(i)}, \boldsymbol{\theta} \mid \mathcal{D}_t^{(i)}, \mathcal{D}_t^{\text{pop}}), \{b_k(t)\}_{k=1}^W)$ is updated at every observation by the parallel operation of the Kalman–GP filter on the surface component and the HMM forward algorithm on the window component.

Temporal Discounting and Phase Sensitivity

The opportunity surface is not stationary. TEMPO implements temporal discounting through a phase-indexed kernel:

$$k_\phi(\mathbf{x}, \mathbf{x}', t, t') = k_{\text{M52}}(\mathbf{x}, \mathbf{x}') \cdot \exp\left(-\frac{(t - t')^2}{2\ell_t^2}\right) \cdot \mathbf{1}[\phi(t) = \phi(t')] \quad (21)$$

where $\phi(t)$ is the system’s operational phase at time t and ℓ_t is a temporal lengthscale. Observations from different phases do not contaminate each other’s posterior; observations within a phase are weighted by recency. The phase indicator creates conditionally Markovian chains per phase—the HMM extension of Section 3.5 generalises this by allowing phase transitions to be inferred from data rather than externally imposed.

The “Add Value” Step as Constrained Bayesian Search

The VADM “add value” step is operationalised as a constrained Bayesian search over the MDP belief state. Given the leading Pareto-optimal configuration $\mathbf{x}^*$, find a configuration $\mathbf{x}^{**}$ that improves its weakest objective while maintaining performance on its strengths:

$$\mathbf{x}^{**} = \arg \max_{\mathbf{x}} [\mu_n^{(k^*)}(\mathbf{x}) - \mu_n^{(k^*)}(\mathbf{x}^*)] \text{ s.t. } \mu_n^{(k)}(\mathbf{x}) \geq \mu_n^{(k)}(\mathbf{x}^*) - \epsilon_k \quad \forall k \neq k^* \quad (22)$$

where $k^* = \arg \min_k v_k(\mu_n^{(k)}(\mathbf{x}^*))$ is the weakest objective and ϵ_k is a small permissible degradation on the others. The improvement typically delivered is 1–2% on the weakest objective at negligible cost to the others: a sub-marginal gain of exactly the kind the platform is designed to identify and compound.

Missing Data and Incomplete Observations

Incomplete observation is the norm in complex adaptive systems. The Bayesian framework handles missing data naturally: missing observations leave the posterior unchanged in those regions, the belief-state uncertainty remains high, and the Markov policy assigns higher exploration value to under-observed configurations. Missing window signals leave the HMM belief at its prior prediction from the transition matrix, decaying toward the stationary distribution at a rate determined by $\mathbf{A}_k$. The system thus degrades gracefully under data sparsity rather than producing overconfident recommendations.

VADM Integration: From Posterior to Decision

Mapping GP Posteriors to LDW Value Functions

The MOBO engine outputs a Pareto-optimal set of candidate configurations, each associated with GP posterior mean vector $\boldsymbol{\mu}_n(\mathbf{x})$ and covariance matrix $\boldsymbol{\Sigma}_n(\mathbf{x})$. These are the inputs that VADM/LDW decision analysis requires: probabilistic performance estimates on multiple criteria, derived from the MDP belief state rather than expert elicitation.

For each objective k and Pareto-optimal configuration $\mathbf{x}$, the LDW value function v_k maps the posterior mean to a normalised value:

$$v_k(\mu_n^{(k)}(\mathbf{x})) \in [0, 1] \quad (23)$$

The probability of sufficiency—that the true value exceeds a decision-maker-specified threshold τ_k —is:

$$P(f_k(\mathbf{x}) \geq \tau_k) = 1 - \Phi\left(\frac{\tau_k - \mu_n^{(k)}(\mathbf{x})}{\sigma_n^{(k)}(\mathbf{x})}\right) \quad (24)$$

This probability is displayed alongside the LDW value-function score in the multi-criteria decision interface, providing a direct link between the MDP belief state and the structured choice framework.

Application: High-Performance Organisational Pipelines

High-performance organisations—whether they develop athletes, clinical professionals, military personnel, or organisational talent—share a common structural challenge: they must manage a pipeline of individuals entering at variable points of readiness, developing through multi-phase programmes under resource and time constraints, and competing for limited selection opportunities whose windows are governed by external calendars outside the organisation’s control. This section applies the TEMPO framework to that pipeline structure in a sector-agnostic form. The variables named are generic; the architectural mapping is exact.

The Development Pipeline as an MDP

A development pipeline is a sequential decision system in which the organisation must make, at each time step, a set of interdependent decisions: whom to recruit into the pipeline; how to configure the development programme for each active participant; whom to advance, retain, or exit at each stage gate; and whom to select for deployment, competition, or assignment at terminal selection events. Each decision is made under uncertainty, has cascading consequences for downstream decisions, and is bounded by resource constraints and opportunity windows that are partially externally determined.

Pipeline Decision Ecology

- **Recruitment:** Which candidates enter the pipeline, and when? Entry timing interacts with development trajectory—late entry compresses development time, early entry may precede observable signal quality. The GP population prior initialises the individual belief state from available pre-entry indicators.
- **Development:** How are resources allocated across intervention types for each active participant? The MOBO engine searches the configuration space for Pareto-optimal development programmes, compounding marginal gains across

intervention classes.

- **Stage-gate advancement:** Which participants advance, are retained for further development, or exit the pipeline? The VADM/LDW structured decision interface provides the multi-criteria choice framework for these decisions.
- **Selection:** Which participants are selected for terminal deployment, competition, or assignment? Selection windows are the primary opportunity constraint in pipeline systems—they are externally determined, time-bounded, and foreclose irreversibly.

Variable Entry and the Cold-Start Problem

One of the most consequential structural challenges in any development pipeline is variable entry: participants arrive at different times, with different prior histories, at different points of readiness relative to the programme standard. A participant who enters the pipeline with limited observable history poses a cold-start problem—the GP posterior has minimal individual data to work from.

TEMPO addresses this through the hierarchical Bayesian population model (equation 19). The individual GP posterior for a new entrant is initialised from the population prior:

$$f_{\text{new}}^{(i)} \sim \mathcal{GP}(\mu_{\text{pop}}(\mathbf{x}), k_{\text{pop}}(\mathbf{x}, \mathbf{x}')) + \delta^{(i)}, \quad \delta^{(i)} \sim \mathcal{GP}(0, k_{\text{ind}}(\mathbf{x}, \mathbf{x}')) \quad (25)$$

The population prior encodes everything the programme has learned from all previous participants: which development configurations produce reliable capacity gains, which intervention sequences carry elevated risk, which entry profiles predict which developmental trajectories. The new entrant immediately benefits from this accumulated knowledge. As individual observations accumulate, the individual posterior $f^{(i)}$ progressively separates from the population prior μ_{pop} , capturing the participant’s idiosyncratic response surface.

The practical consequence is that the MOBO engine can produce meaningful development prescriptions from the first week of a new entrant’s programme—prescriptions that are more conservative than those it would produce for a participant with extensive history, and whose uncertainty bands are wider, but which are grounded in population-level evidence rather than naive defaults. The detection resolution (the minimum gain difference the posterior can reliably distinguish) narrows monotonically as individual data accumulates, and the system reports this resolution explicitly to practitioners.

Development Phase Structure and Opportunity Cost

Development pipelines universally have phase structure: an initial general capability-building phase, a more specific preparation phase aligned to performance requirements, and one or more selection or deployment phases with externally-determined timing. The phase-indexed kernel (equation 21) ensures that observations from one phase do not contaminate the posterior for another—a 1% capacity gain identified during general preparation may not hold during specific preparation, and the system does not extrapolate across phase boundaries without within-phase evidence.

The opportunity cost of delay is maximal near phase transitions. A participant who arrives at the specific preparation phase with an underdeveloped capacity ceiling has foreclosed the configurations that would have been available in general preparation. The time-indexed reward function (equation 13) captures this: as a phase transition approaches, the urgency scores for capacity-building interventions that must be completed within the current phase rise, augmenting the acquisition function toward those interventions and away from configurations that would be equivalent in isolation but are misaligned with the approaching window.

Table 3: General Pipeline Phase Structure and Opportunity Windows

Phase	Primary Objective	Opportunity Window	Foreclosure Risk
Entry / Recruitment	Establish baseline belief state; initialise individual GP from population prior	Recruitment cycle; entry cohort windows	High: missed cohorts defer pipeline entry
General Development	Broad capacity-building across all intervention classes; identify response surface idiosyncrasies	Foundational development window before specific preparation begins	Medium: capacity ceiling is set here
Specific Preparation	Configuration-specific optimisation aligned to deployment requirements; intensify MOBO exploitation	Preparation window before selection event	High: configurations infeasible post-window
Selection / Deployment	Multi-criteria structured decision under uncertainty; VADM/LDW interface; “add value” search	Selection announcement or deployment assignment window	Absolute: window closes at announcement
Post-Selection	Update population prior from selection outcomes; propagate evidence to next cohort	Programme evaluation and design cycle	Low: retrospective; informs future pipeline

The Compounding Arithmetic in Development Pipelines

The compounding arithmetic of equation (1) applies directly to development pipelines. Consider a participant with $M = 6$ independently controllable development variables each yielding a marginal improvement of 1.5% through optimal configuration:

$$\prod_{m=1}^6 (1 + 0.015) = 1.015^6 \approx 1.093 \quad (26)$$

A 9.3% compounded performance advantage at the pipeline’s selection threshold is not marginal. At the right tail of any performance distribution—where the gap between advancing and not advancing is small in absolute terms—this differential determines selection outcomes systematically. The crucial condition is that the gains be genuinely independent. Where development interventions interact negatively—allocating maximum resources to two capacity dimensions simultaneously exceeds the participant’s total adaptive capacity—the GP surrogate learns this interaction from population data and the MOBO engine avoids the infeasible configuration. The compounding-versus-interference tradeoff is precisely the kind of system-level dynamic that is invisible to practitioners managing interventions in isolation and visible only to an integration layer maintaining a joint belief state over the full response surface.

Selection Decisions Under Uncertainty

Selection is the terminal decision event in a development pipeline: it is irreversible, time-bounded, and consequential for both the organisation and the participant. The VADM/LDW structured decision interface is most directly relevant here. The multi-criteria selection problem—comparing candidates across objectives with different measurement scales, different uncertainty levels, and different institutional weights—is precisely the decision class for which LDW’s value function and sensitivity analysis architecture was designed.

In the TEMPO integration, candidate comparison is populated not by expert judgment but by GP posterior means and credible intervals derived from the MDP belief state. The decision-maker adjusts the LDW weight vector and observes, in real time, which weight configurations change the ranking and by how much. Sensitivity analysis identifies borderline cases—pairs of candidates whose ranking is robust across plausible weight vectors versus pairs whose ranking depends critically on a single contested criterion. For borderline cases, the platform flags two responses: either collect additional observations to narrow the relevant posterior (an exploration step in the MDP), or document explicitly that the selection was made with awareness of the marginal sensitivity.

The VADM “add value” function, operationalised as the constrained Bayesian search of equation (22), serves a further function in pipeline contexts: it proposes development configurations for near-selected candidates that, if executed in the remaining time before the window closes, would close the gap on the dimension driving their non-selection. This converts the selection decision from a terminal judgment into a developmental prescription—the selection interface generates not only a ranking but a targeted intervention plan for the candidates most proximate to the threshold.

Illustrative Decision Scenarios

Scenario 1: Variable-Entry Participant with Limited History

A participant enters the pipeline at a non-standard time—later in the development cycle than the typical cohort—with limited prior observable history. The GP posterior is initialised from the population prior. The MOBO engine identifies a pattern distinct from the population mean: high sensitivity to a specific class of high-intensity intervention but robust response to moderate-intensity volume loading. The Development Configuration module proposes a modified programme that prioritises moderate-intensity work, projecting a marginal short-term performance cost of 0.4% over eight weeks but a 22% reduction in estimated adverse-response risk. The marginal gain decomposition shows a projected net advantage of +2.1% over 12 weeks when avoided adverse outcomes are accounted for. Because a selection window is currently believed open ($b_{\text{sel}}(t) = 0.78$), the urgency score augments the acquisition function to additionally weight trajectory-slope configurations, adjusting the programme further toward interventions that maximise end-of-window projected position rather than current performance level.

Scenario 2: Multi-Candidate Selection Under Uncertainty

Three candidates compete for two positions at a terminal selection event. Their GP posterior mean vectors across all four objectives, with 90% credible intervals derived from the accumulated belief state, are loaded into the LDW comparison interface. The decision-maker enters weight preferences through the LDW swing-weighting protocol. Sensitivity analysis reveals that the ranking between candidates A and B reverses when the development-trajectory criterion receives more than 35% weight—just outside the decision-maker’s stated preference range but within one standard deviation of elicitation uncertainty. The platform flags this borderline sensitivity, recommending either additional observations on the relevant dimension before the window closes, or explicit documentation that the selection was made with awareness of the margin. The VADM “add value” function proposes a targeted development configuration for candidate B that the posterior suggests would close the trajectory gap by 1.8% over eight weeks, converting the borderline case into a clear advancing selection if the window timeline permits execution.

Discussion

Positioning Against Alternative Frameworks

The claim that TEMPO addresses a structural absence requires adversarial contrast—not assertion. The following table explicitly positions TEMPO against the four most relevant alternative frameworks a reviewer or practitioner might reach for:

Table 4: Explicit Contrast: TEMPO versus Alternative Approaches

Property	Static MCDA	Batch MOBO	RL (no window)	Rule-based priority	TEMPO
Sequential belief update	No	Partial	Yes	No	Yes
Multi-objective Pareto	Yes	Yes	Partial	No	Yes
Time-indexed reward	No	No	Via discount γ	No	Yes (HMM)
Opportunity window belief	No	No	No	No	Yes
Foreclosure probability	No	No	No	No	Yes
Human override as belief update	No	No	No	No	Yes
Governance / auditability interface	Yes	No	No	Partial	Yes
Marginal gain compounding	No	Partial	Partial	No	Yes

The key differentiators are not individual features but their simultaneous presence. Static MCDA provides auditability but no learning. Batch MOBO compounds gains but misses the temporal structure of opportunity. Reinforcement learning handles sequential decisions but addresses time via discounting—which shrinks future rewards uniformly rather than tracking the opening and closing of specific windows. Rule-based priority systems operationalise urgency without learning it from data. None of the alternatives jointly maintains a Markov belief state over the opportunity surface, models window foreclosure as a computable belief-dependent quantity, and translates the result into an auditable multi-criteria governance interface.

The Epistemological Advance

The fundamental advance of TEMPO is the shift from *data organisation* to *knowledge generation through connected experimentation under time-indexed opportunity cost*. Static decision frameworks structure choices on data the analyst already possesses. The GP/MDP architecture generates new knowledge by treating every observation as a Markov belief update and every intervention as an experiment whose result narrows the uncertainty over the opportunity surface.

The time-indexed reward is what makes this knowledge actionable before it is too late. A 1.2% improvement in capacity development rate attributable to a specific intervention sequence is not visible in periodic assessments viewed in isolation. It becomes visible when a statistical model operating as a POMDP belief state integrates evidence across units

and time. The integration is the intelligence. The time-indexed reward is what ensures that intelligence is deployed before the window closes.

The Human-Decision-Maker Authority Structure

TEMPO takes a principled position on the human–AI authority question: the platform proposes, the decision-maker decides, and the platform learns from that decision regardless of whether the recommendation was accepted or overridden. Within the POMDP framework, this is formally complete: the override is a belief update, the override pattern is a signal about the transition model, and the platform’s evolving model of decision-maker preference is a posterior over a latent preference structure.

This authority structure preserves institutional accountability, captures tacit expertise, and addresses the cold-start problem through the population prior. It also ensures that the system’s window urgency assessments are subject to human judgment—the decision-maker may have information about window status that is not captured in the observable emission signals, and the override mechanism allows that information to enter the belief state.

The Integration Thesis and Contribution Summary

The primary theoretical contribution of this paper is the formalisation of opportunity cost as a time-indexed, belief-dependent component of the MDP reward function, mediated by latent opportunity windows modelled as an HMM. This is what makes sequential architecture structurally necessary rather than merely convenient, what drives the Markov policy toward timely action on closing opportunities, and what distinguishes TEMPO from both standard MOBO and standard MDP applied work. It is not present in the Bayesian optimisation literature, the MDP applied literature, or the decision analysis literature.

The full integration claim follows from this foundation: *the value of integrating micro-optimisations into a unified MDP-grounded Bayesian decision platform, augmented by time-indexed opportunity-cost reward and HMM window belief, exceeds the sum of the values of managing those micro-optimisations independently—because integration enables detection of cross-variable interactions, compounding of independent gains, propagation of evidence across units through the shared belief state, and time-sensitive action under closing opportunity windows. None of these properties is available without the integration layer, and none of the existing alternative frameworks provides all four simultaneously.*

Limitations and Boundary Conditions

Data volume requirements.. The hierarchical GP model requires approximately 40–60 unit-seasons of data to fit stable population hyperparameters. Small systems may require a warm-up period of reduced prescriptive authority.

Window signal availability.. The HMM window belief update depends on observable signals correlated with window status. In domains with sparse or unreliable external signals, the HMM will revert to transition-matrix predictions and the urgency scores will carry high uncertainty. The system reports this uncertainty explicitly; urgency recommendations without supporting signal evidence are flagged as prior-driven rather than evidence-driven.

GP scalability. Exact GP inference is $O(n^3)$. The Kalman state-space equivalence addresses the dominant temporal streaming bottleneck at $O(n)$ for sequential observations under the Matérn-5/2 kernel; the inducing-point approximation addresses multi-dimensional input spaces at $O(M^2n)$. Neither fully eliminates scalability pressure in high-dimensional, high-frequency settings. A full implementation combines both strategies, with the Kalman filter governing temporal updates and the inducing-point method governing spatial interpolation across the configuration space.

Decision-maker interpretability. The POMDP framing adds theoretical precision. The dashboard layer abstracts mathematical machinery, presenting results in percentage-gain and window-urgency terms. The full belief state is available on demand but not the default presentation.

Conclusion

In any system where time passes, resources are finite, interventions interact, and the value of action is partially observable—the failure to compound marginal gains systematically is not an oversight but an architectural absence. This paper has proposed and formalised that architecture.

TEMPO—Time-indexed Evidence-based Marginal gains under Partial Observability—integrates three intellectual traditions into a unified adaptive system: Gaussian Process Bayesian optimisation and hierarchical Bayesian modelling, which maintain a calibrated belief state over an unknown opportunity surface; Markov Decision Process formalisation, in which the GP posterior is the sufficient Markov belief state, the acquisition function is the near-optimal policy, the Matérn-5/2 surrogate update is a Kalman filter step, and decision-maker overrides are POMDP belief updates encoding tacit expertise; and the VADM/LDW structured decision interface, which translates probabilistic Pareto recommendations into structured choice alternatives that decision-makers can evaluate, challenge, modify, and defend.

The framework’s central theoretical contribution is the formalisation of opportunity cost as a time-indexed component of the MDP reward function, coupled to a Hidden Markov Model belief state over latent opportunity windows. This formalisation establishes why sequential architecture is structurally necessary—not merely convenient—in any system where intervention value is time-dependent. It introduces an urgency score derived from window belief, foreclosure probability, and window-specific value that drives the Markov policy toward timely action on closing opportunities. And it provides the first integrated treatment of marginal gain compounding, multi-objective Bayesian optimisation, and time-windowed opportunity cost within a single decision-theoretic framework.

The sport development instantiation demonstrates the framework with high data density and clear variable structure. The architecture itself is agnostic to sector, data type, and intervention domain. The companion working paper will develop sector-specific instantiations in clinical, organizational, and defense contexts.

The integration is the intelligence. The compounding is the competitive advantage. The

time-indexed reward is what ensures both remain achievable before the window closes.

References

- Brochu, E., Cora, V. M., & de Freitas, N. (2010). A tutorial on Bayesian optimization of expensive cost functions, with application to active user modelling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*.
- Chelst, K., & Canbolat, Y. B. (2011). *Value-Added Decision Making for Managers*. CRC Press / Taylor & Francis.
- Daulton, S., Balandat, M., & Bakshy, E. (2020). Differentiable expected hypervolume improvement for parallel multi-objective Bayesian optimization. *Advances in Neural Information Processing Systems*, 33, 9851–9864.
- Frazier, P. I. (2018). A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*.
- Gardner, J., Pleiss, G., Bindel, D., Weinberger, K., & Wilson, A. G. (2018). GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. *Advances in Neural Information Processing Systems*, 31.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). Chapman and Hall/CRC.
- Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2), 99–134.
- Logical Decisions. (2021). *Logical Decisions for Windows v8.0: Software for multi-objective decision analysis* [Open source, MIT licence]. <https://logicaldecisions.com>
- Mockus, J. (1975). On Bayesian methods for seeking the extremum. In *Optimization Techniques IFIP Technical Conference* (pp. 400–404). Springer.
- Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.
- Särkkä, S. (2013). *Bayesian Filtering and Smoothing*. Cambridge University Press.
- Särkkä, S., & Solin, A. (2019). *Applied Stochastic Differential Equations*. Cambridge University Press.
- Smith, G. (2021). *Logical Decisions for Windows v8.0: A guide to multi-objective decision analysis* [Software documentation]. <https://logicaldecisions.com>
- Srinivas, N., Krause, A., Kakade, S. M., & Seeger, M. W. (2012). Information-theoretic regret bounds for Gaussian process optimization in the unknown noise setting. *IEEE Transactions on Information Theory*, 58(5), 3250–3265.
- Thornton, H. R., Duthie, G. M., Pitchford, N. W., Delaney, J. A., Benton, D. T., & Dascombe, B. J. (2016). Effects of a two-week high-intensity training camp on sleep activity, well-being, and cardiovascular parameters in elite sport athletes. *International*

Journal of Sports Physiology and Performance, 11(8), 1058–1064.

Wilson, J. T., Moriconi, R., Hutter, F., & Deisenroth, M. P. (2018). The reparameterization trick for acquisition functions. *arXiv preprint arXiv:1712.00424*.