Les Cahiers du GERAD ISSN: 0711–2440

Informativeness-driven blackbox optimization of startup sequence selection for hydro turbine digital twins

M. Fahed, M. Gagnon, S. Alarie, J.-P. Gauthier, A. Favrel, M. Kokkolaras

G–2026–78

October 2026

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

Citation suggérée : M. Fahed, M. Gagnon, S. Alarie, J.-P. Gauthier, A. Favrel, M. Kokkolaras (Octobre 2026). Informativeness-driven blackbox optimization of startup sequence selection for hydro turbine digital twins, Rapport technique, Les Cahiers du GERAD G–2026–78, GERAD, HEC Montréal, Canada.

Avant de citer ce rapport technique, veuillez visiter notre site Web (https://www.gerad.ca/fr/papers/G-2026-78) afin de mettre à jour vos données de référence, s'il a été publié dans une revue scientifique.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

Suggested citation: M. Fahed, M. Gagnon, S. Alarie, J.-P. Gauthier, A. Favrel, M. Kokkolaras (October 2026). Informativeness-driven blackbox optimization of startup sequence selection for hydro turbine digital twins, Technical report, Les Cahiers du GERAD G–2026–78, GERAD, HEC Montréal, Canada.

Before citing this technical report, please visit our website (https://www.gerad.ca/en/papers/G-2026-78) to update your reference data, if it has been published in a scientific journal.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2026 – Bibliothèque et Archives Canada, 2026

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2026 – Library and Archives Canada, 2026

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Informativeness-driven blackbox optimization of startup sequence selection for hydro turbine digital twins

Melad Fahed [a, d]

Martin Gagnon [b]

Stéphane Alarie [b, d]

Jean-Philippe Gauthier [b]

Arthur Favrel [b]

Michael Kokkolaras [c, d]

[a] *Department of Mechanical Engineering, McGill University, Montréal (Qc), Canada, H3A 0C3*

[b] *Hydro-Québec Research Institute (IREQ), Varenne (Qc), Canada, J3X 1S1*

[c] *School of Mechanical and Automotive Engineering, Clemson University, South Carolina 29607, United States*

[d] *GERAD, Montréal (Qc), Canada, H3T 1J4*

melad.fahed@mail.mcgill.ca
gagnon.martin11@hydroquebec.com
alarie.stephane@hydroquebec.com
gauthier.jean-philippe4@hydroquebec.com
favrel.arthur@hydroquebec.com
mkokkol@clemson.edu

October 2026
Les Cahiers du GERAD
G–2026–78

Copyright © 2026 Fahed, Gagnon, Alarie, Gauthier, Favrel, Kokkolaras

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract : Recent efforts in hydropower innovation have focused on achieving a fully connected digital twin that continuously synchronizes data between the physical and virtual worlds. In an ideal scenario, this twin would enable real-time bidirectional data flow to support monitoring, prediction, and control. However, challenges remain, particularly in maintaining reliable data connectivity and ensuring that the digital twin remains representative of diverse operating conditions. This paper addresses the need for initializing and calibrating digital twins in the absence of representative, diverse, and real-time site measurements. We propose an offline trajectory-guided blackbox optimization framework for selecting startup sequences prior to measurement campaigns. It serves as both an initialization and backup step for measurement campaigns, preserving digital twin functionality when real-time measurement data is lost. Time histories of guide vane opening, blade angle, and rotational speed are generated for candidate sequences, and an objective function quantifying their informativeness based on diversity is proposed. Sequences are selected using a greedy max-min criterion and the Mesh Adaptive Direct Search algorithm for derivative-free optimization. The method is applied to a double-regulated Kaplan turbine to identify startup sequences that maximize coverage of the turbine operating range and enhance the generalizability of models within the digital twin paradigm.

Keywords : Kaplan turbines; startup transients; blackbox optimization; elastic distances; dynamic time warping

Acknowledgements: We thank Hydro-Québec, Mitacs and Natural Sciences and Engineering Research Council of Canada (NSERC) [funding reference number 589463] for funding this research, for their contribution, and we thank Hydro-Québec for providing the data used in this study. We also thank Christophe Tribes (GERAD/Polytechnique Montréal) for their support in configuring the Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD) optimization software.

1 Introduction

Hydroelectric utilities are increasingly pursuing fully connected digital twins (DTs) to bridge the physical and virtual worlds through continuous data exchange. In an ideal setup, a digital twin (DT) enables real-time, bidirectional data flow to support monitoring, prediction, and control of hydro turbine behavior. Hydropower units are renewable electricity generators and startup behavior directly affects unit availability and operational flexibility, so improving startup campaign design supports sustainable energy generation. Achieving this level of integration, however, requires robust and representative data across all operating conditions, which can be difficult to ensure consistently. One critical aspect of this effort is the accurate modeling of hydro turbine startup behavior, which is highly dynamic and impactful on turbine lifetime. Indeed, hydroelectric generating units may shut down and restart multiple times per day, and the startup transient induces large and fast variations in guide vane opening and flow that translate into cyclic loads on the runner, which accelerate fatigue and drive maintenance costs. Prior work has documented this reliability burden and its operational consequences. For example, Gagnon et al. quantify how startup procedures alter unit performance metrics and mechanical loading [1]; Trivedi et al. analyze the hydraulic transients that arise during startups and how they propagate through the waterway [2]; and Gagnon et al. investigate the practical optimization of startup settings under operational constraints, linking controllable settings to reduced transient severity [3]. More broadly, Liu et al. review fatigue mechanisms and assessment methods in hydro machinery [4], Soltani-Dehkharqani et al. discuss Kaplan-specific startup dynamics and controllability challenges [5], Gagnon et al. report data-informed evaluation of startup strategies grounded in field measurements [6], and Muser et al. introduce machine-learning approaches aimed at fast prediction and decision support for startup planning [7]. In practice, the tuning of DTs is realized using data collected during on-site measurement campaigns. These campaigns interrupt renewable electricity production and are limited in number and duration, so each selected startup test must provide as much information as possible for model calibration.

When real-time site measurement data are available, Mai et al. show that hydro turbine startup sequences can be selected adaptively during the campaign using active learning [8], where measured fatigue indicators guide which sequence to run next. In that online setting, *representativeness* means choosing a candidate that matches operating regimes that are already well observed [9, 10]. For hydro turbine startups, this corresponds to selecting sequences that resemble the stable procedures operators already use and that are known to induce low transient strain [8]. This paper addresses the converse situation where real-time measurements are unavailable or unreliable during the campaign. In that case, representativeness cannot be evaluated or used for selection, yet the planned tests must still produce informative data that improve predictive models and support future operational decisions.

1.1 Experiment design using trajectory similarity

Hydroelectric power startup transients are difficult to sample efficiently because field measurement campaigns are short and logistically constrained. Early design of experiments (DOE) approaches therefore sought to reduce uncertainty before testing by selecting representative sets of inputs. Kennard–Stone propose a greedy farthest-point selection strategy that iteratively spreads samples to improve surrogate generalization [11]. McKay et al. popularize Latin hypercube sampling (LHS) as a practical, broadly applicable space-filling baseline for high-dimensional design spaces [12]. In contrast, Morris designs target screening: they identify influential inputs with relatively few model evaluations, trading exhaustive coverage for diagnostic insight [13]. These ideas mainly target epistemic uncertainty that stems from incomplete knowledge and can be reduced by collecting informative data, versus aleatoric uncertainty due to irreducible noise in measurements or the environment [14].

Trajectory similarity measures have matured for comparing dynamic responses. Dynamic time warping (DTW) introduced elastic time alignment to handle local timing differences in trajectories [15]; the discrete Fréchet distance compares curves through a point-to-point leash-length notion [16]; and

time warp edit distance (TWED) adds edit operations with time and stiffness penalties to control warping and noise sensitivity [17]. Sarin et al. provide comparative guidance on how these elastic distances differ in practice and which failure modes to watch for [18]. More recently, Cuturi and Blondel propose soft-DTW as a differentiable relaxation suitable for learning with trajectory losses [19], and Blondel et al. discuss divergence variants that mitigate soft-DTW's bias [20]. Janati et al. combine soft-DTW with optimal transport (OT) to jointly align time and amplitude in spatiotemporal signals [21]. Comprehensive empirical work has also clarified when popular time-series distances behave and do not behave as commonly assumed, cautioning against uncritical metric choices in practice [22]. Beyond univariate series, Vlachos et al. study scalable multidimensional trajectory similarity and indexing strategies that generalize these distances to vector-valued paths [23].

1.2 Blackbox optimization

Blackbox optimization (BBO) considers the case where objective and constraint evaluations are possible only through experiments or simulations, without access to underlying mathematical models or knowledge of their properties. The Mesh Adaptive Direct Search (MADS) algorithm provides a convergence-backed direct-search framework [24], and the NOMAD software offers a practical implementation with constraint handling, mixed variables, and restarts [25]. In engineering practice, trajectory distances are sometimes embedded in optimization loops, but mostly for inverse problems and calibration against a single target trace. For example, Dubois et al. propose robust comparisons for partially observed curves [26], and Witowski et al. discuss multidimensional DTW-style comparisons [27]; these uses demonstrate curve-distance objectives in optimization, mainly for target matching rather than output-based diversity maximization.

When experiments are expensive and real-time measurement data are unavailable, allocating scarce experiments involves a design choice between representativeness and informativeness. Representative experiments concentrate on areas where performance is already good or is frequently encountered, while informative experiments emphasize coverage of poorly understood regions to reduce epistemic uncertainty when live feedback from real-time measurements is unavailable. Active learning formalized this trade-off by combining representativeness and informativeness in a single query logic so that selected points reflect both the data distribution and the model's uncertainty. Shen et al. emphasize uncertainty-aware querying to target informative samples [10], while He et al. propose an active-learning strategy that explicitly balances exploitation and exploration through alternating query types [28]. Related exploration–exploitation trade-offs appear in Bayesian optimization for water infrastructure and closely allied cyber-physical systems, where acquisition functions balance predicted reward and uncertainty to select the next operating condition under blackbox simulators such as EPANET/WNTR. Candelieri et al. describe Bayesian-optimization tooling for water-network decision support under expensive simulation [29], and the follow-up Better Call Bayes study demonstrates practical acquisition-driven operation under limited budgets [30]. Alnajim et al. apply Bayesian optimization to contamination-response settings where evaluations are costly and constraints matter [31], and they further extend the same paradigm to operational-response planning under uncertainty [32]. Recent work in hydroelectric power combines active learning with BBO for turbine startups by proposing virtual sensors and adaptive sampling rules, offering a solution specifically when real-time measurement data are available [8]. A contingency when such data are not available is to use physics-based simulation and DTs to generate synthetic data for rare events, thus increasing historical records and improving the robustness of predictive models. Guo et al. exemplify reliability-oriented DT data augmentation by simulating rare hydroturbine failure scenarios to improve generalization when field events are scarce [33].

Within water and energy systems, sequential blackbox and learning-based optimizers are iterative schemes that treat the simulator as a computationally expensive blackbox model and choose each new iteration based on results of previous ones. For water-network decision support, Better Call Bayes demonstrates acquisition-driven planning under limited evaluation budgets [30], while Alnajim

et al. focus on contamination-response decisions in which feasibility constraints interact with expensive hydraulic simulation [31]. For operational control and scheduling, Salomons et al. study real-time decision making under expensive simulations and delays [34], and Hu et al. address pump scheduling as a constrained, simulator-based optimization problem [35]. These studies operate on scalar or vector performance metrics and do not consider distances in trajectory space as the selection objective. In other words, they explore the input domain using statistical uncertainty as a surrogate of a scalar target, not an explicit diversity criterion on full-time output histories.

A separate line of practice addresses periods when real-time measurement data are unavailable or unreliable. The control and estimation literature provides fault-tolerant and intermittent measurement approaches that maintain state estimates when data streams are incomplete. Gu et al. study Kalman filtering under random delays [36], Li et al. develop distributed Kalman filtering for networked sensing [37], Gao et al. analyze stability conditions for Kalman filtering with imperfect communications [38], and Shan et al. propose robust filtering variants designed to tolerate intermittent measurements [39]. In hydroelectric power condition monitoring, when real-time measurement data are unavailable, operators rely on locally logged measurements, on-site inspections, and vibration monitoring systems that can operate autonomously until communications are restored. More recently, high-fidelity DTs have been used to generate synthetic data for rare events when real samples are unavailable. A recent DT study for the prediction of hydroturbine failure shows how simulated rare scenarios augment historical data to improve robustness and generalization when field events are scarce [33]. Practitioners inject information from physics, surrogates, or simulators to reduce the epistemic uncertainty before the next measurement window.

Three elements are well established in literature yet typically used in isolation: (i) input space coverage for selecting candidate experiments, (ii) trajectory distance for comparing transient output responses, and (iii) constraint aware blackbox planning for feasibility constrained campaigns. For (i), the Kennard Stone algorithm selects maximin points in the input space [11], while Latin hypercube sampling stratifies each dimension to improve space filling [12]. For (ii), we use DTW, which aligns two transient trajectories by allowing time warping and yields a similarity score for non synchronized dynamics [15]. For (iii), mesh adaptive direct search provides a practical framework for feasibility constrained blackbox optimization [24], implemented in NOMAD V4 [25]. Existing uses of trajectory distances in optimization mainly minimize misfit to a fixed target trace for calibration [26], while active learning and Bayesian optimization generally assume streaming feedback during the campaign [29, 30], including distance aware variants that update selection online [27].

Despite these advances, an offline planning capability is still lacking. What appears to be missing is an offline experimental design method that, before any measurements are taken, chooses and schedules a feasible set of hydro turbine startup sequences by maximizing how different the predicted transient trajectories are, while respecting timing and operational constraints. We therefore formulate startup experiments planning as an offline, trajectory-guided optimization problem whose objective is *informativeness*. In this context, informativeness quantifies the potential of a candidate sequence to expand model knowledge by introducing diverse and previously unseen dynamic responses [9, 10]. Each candidate sequence generates time histories of guide vane opening (GVO), blade angle, and rotational speed through a simulation model. After resampling and normalization of the sequences, we quantify dissimilarity between trajectories using DTW. A constraint-aware derivative-free optimizer then selects a feasible set of startup sequences that is maximally spread in trajectory space, yielding a limited campaign that spans a broad range of turbine transient behaviors even when no real-time feedback is available to adapt the plan.

This research makes three contributions.

1. It frames hydro turbine startup sequence selection as diversity maximization over multidimensional transient trajectories, explicitly targeting the reduction of epistemic uncertainty.

2. It integrates elastic trajectory distances with constraint-aware derivative-free optimization to propose feasible, executable sequences without gradients and without real-time site measurement data.
3. It demonstrates on a Kaplan turbine that the planned measurement campaigns span a wider range of transient behaviors than input space sampling baselines, improving the generalizability of models within the DT paradigm once campaign data are acquired.

The rest of the paper is organized as follows. Section 2 states the problem, lists the decision variables, describes feasibility rules, and presents the proposed method. Section 3 reports and analyzes results, and Section 4 presents the conclusion.

2 Problem formulation

A turbine startup can be represented as a finite, time-dependent trajectory generated by a low-dimensional control program executed by the speed governor. This research focuses on a Kaplan turbine, which is a double-regulated unit where both the guide vane opening and runner-blade pitch can be controlled. Let $t \in [0, t_s]$ denote time until synchronous speed is reached, where synchronous speed ω_s is the rotational speed at which the generator's frequency matches the grid frequency and the unit can safely connect to the grid. Let $\omega(t)$, $\alpha(t)$, and $\beta(t)$ denote the instantaneous rotational speed, GVO, and runner-blade pitch, respectively. Each startup sequence is defined by the vector:

$$\psi = [r_1,\ \alpha_1,\ \omega_{\text{thr}},\ r_2,\ \alpha_2,\ \beta_{\text{ini}}]^{\mathsf{T}}, \tag{1}$$

which encodes the piecewise ramps and plateaus drawn in Figure 1 for the GVO, namely $\alpha(t)$, the blade angle $\beta(t)$, and the resulting speed $\omega(t)$ throughout the startup sequence until synchronization speed ω_s is reached, where

r_1: *initial ramp rate* of $\alpha(t)$;
α_1: *initial plateau level* reached by $\alpha(t)$ after the initial ramp;
ω_{thr}: *speed threshold* at which the second ramp on $\alpha(t)$ is triggered;
r_2: *post-threshold ramp rate* of $\alpha(t)$;
α_2: *second plateau level* reached by $\alpha(t)$ after the second ramp;
β_{ini}: *initial blade angle* used at the beginning of the startup sequence.

The first five vector components induce the five GVO movement phases labeled (1)–(5) in Figure 1:

1. *Initial ramp:* with the unit initially at rest, guide vanes open at rate r_1 until $\alpha(t)$ reaches the initial plateau α_1; $\omega(t)$ begins to rise gradually.
2. *Initial plateau:* $\alpha(t)$ holds at α_1 while the unit accelerates; once $\omega(t)$ reaches the threshold ω_{thr}, the second ramp is triggered.
3. *Second ramp:* guide vanes open at rate r_2 until $\alpha(t)$ reaches the second plateau α_2, producing a steep speed increase.
4. *Second plateau:* $\alpha(t)$ holds at α_2 while $\omega(t)$ continues to increase rapidly.
5. *Feedback control loop:* as the speed nears its synchronization value ω_s, closed-loop control becomes effective and brings $\omega(t)$ smoothly to ω_s by t_s.

The movement of the blades is controlled exclusively by the sixth vector component, β_{ini}. The blade angle $\beta(t)$ is set at β_{ini} before the startup sequence is launched and is held at this value until $\omega(t)$ reaches a low speed threshold. The blade angle will then decrease at the maximum rate allowable by the mechanism until it reaches zero.

The speed threshold that affects the movement of the blades is independent from ω_{thr} and is one of many tunable speed governor settings that were kept constant for this study. This was done to keep the problem size manageable by focusing only on the most important parameters.

The governor maps ψ to GVO and blade angle instantaneous setpoints through its embedded algorithm. These setpoints are used in feedback loops to ultimately control $\alpha(t)$ and $\beta(t)$. The turbine–generator dynamics produce the rotational speed response $\omega(t)$. The resulting kinematic trajectory is

$$\tau(\psi) \;=\; \big\{\big(\alpha(t),\, \beta(t),\, \omega(t)\big) \;:\; t \in [0, t_{\mathrm{s}}]\big\}\,. \tag{2}$$

Figure 1 summarizes this six-components, five-phase structure. The components mirror how startup programs are implemented in the governor and therefore define the decision space for optimization. It should be noted that an example where $\alpha_2 > \alpha_1$ is shown, but the opposite case, $\alpha_2 < \alpha_1$, is also possible.

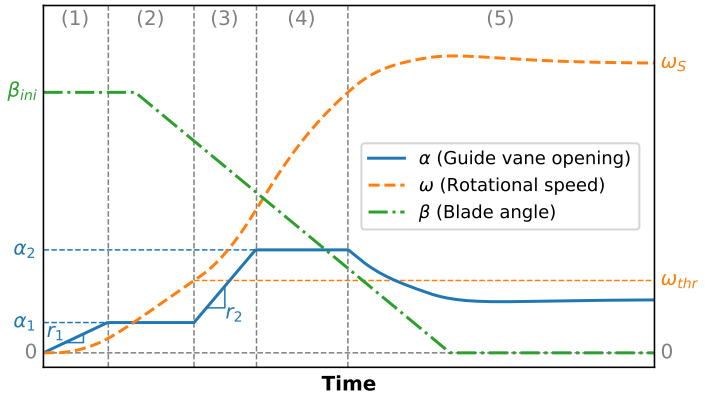

Figure 1: Startup sequence parameterization and phases

2.1 Representativeness-driven selection

When real-time measurement data are available during a measurement campaign, startup tuning can be formulated as an *active learning* problem [8]. During each startup, the blade strain $s(t)$ is measured. A scalar *strain severity score* $\mathcal{L}(\tau, s)$ is then computed to quantify how severe the transient is, for example as the peak strain amplitude or transient energy. Let $\mathcal{D}_t$ represent all measurements collected up to time t. The next startup configuration is chosen as the one predicted, based on the model's current state of knowledge, to minimize the expected strain severity:

$$\begin{aligned} \psi^{\star}_{\mathrm{rep}} \in \arg\min_{\psi \in \mathcal{F}} \mathbb{E}\left[\mathcal{L}\big(\tau(\psi),\, s\big) \,\middle|\, \mathcal{D}_t\right] \\ \text{s.t.} \quad t_{\mathrm{s}}(\psi) \le T_{\mathrm{s}}, \end{aligned} \tag{3}$$

where $\mathbb{E}[\cdot|\mathcal{D}_t]$ denotes the conditional expectation of the model based on observed data, s denotes the strain value and $\mathcal{F}$ is the feasible set that contains all physically and operationally admissible startups. This design is primarily representativeness-driven. It includes only a weak notion of informativeness, so it tends to remain near regions already seen in the training data and is not strongly rewarded for large exploratory moves.

2.2 Informativeness-driven selection

When real-time measurement data are unavailable or unreliable, the objective shifts from minimizing transient load to *maximizing trajectory diversity*. The goal is to pre-select a set $\mathcal{S} \subset \mathcal{F}$ of K feasible startups whose trajectories $\tau(\psi)$ are as distinct as possible in the kinematic space (α, β, ω). Formally:

$$\max_{\substack{\mathcal{S}\subset\mathcal{F} \\ |\mathcal{S}|=K}} \; \min_{\substack{\psi_i,\psi_j\in\mathcal{S} \\ i<j}} d\Big(\mathcal{N}\big(\tau(\psi_i)\big), \mathcal{N}\big(\tau(\psi_j)\big)\Big), \tag{4}$$

where $d(\cdot,\cdot)$ measures trajectory dissimilarity (e.g., using DTW, Soft-DTW, discrete Fréchet distance or TWED, among others) and $\mathcal{N}(\cdot)$ resamples and normalizes trajectories for consistent comparison. This *max–min* structure enforces a farthest-point design in trajectory space, ensuring that even the most similar pair of selected startups remain well separated, thereby targeting reduction of epistemic uncertainty. In this work, we adopt the well-established DTW for d because it robustly aligns trajectories with local time distortions, which is equivalent to the differentiable Soft-DTW formulation when the smoothing parameter $\gamma = 0$ [15, 19].

Because solving (4) exactly would require evaluating all combinations of feasible startups (a combinatorial problem) we use a greedy, sequential approximation. At iteration m, we already have a set of selected, feasible startups $\mathcal{S}_{m-1} = \{\psi_1, \ldots, \psi_{m-1}\}$. We then search for the next startup ψ_m that maximizes its minimum trajectory distance to all previously chosen ones:

$$\psi_m \in \arg\max_{\psi \in \mathcal{F}} \min_{\substack{\psi_i \in \\ \mathcal{S}_{m-1}}} d\Big(\mathcal{N}\big(\tau(\psi)\big), \mathcal{N}\big(\tau(\psi_i)\big)\Big). \tag{5}$$

Once ψ_m is found, it is appended to the set:

$$\mathcal{S}_m = \mathcal{S}_{m-1} \cup \{\psi_m\},$$

where $\mathcal{S}_{m-1}$ is the set of the first $m-1$ selected startup configurations, and $\mathcal{S}_m$ is the updated set after adding the m-th candidate. This "select–add–repeat" process continues until either the target number of planned startups K is reached ($|\mathcal{S}_m| = K$) or the overall campaign duration exceeds the total time budget T_{budget}. The result is a set of K feasible, information-rich startup sequences that uniformly span the space of achievable turbine responses.

2.3 Feasibility constraints

Using MADS algorithm, we optimize the six components of (1) with lower and upper bounds listed in the same order, denoted as $l \le \psi \le u$, where

$$\begin{aligned} l &= \{0.25,\ 0.10,\ 0,\ 0.25,\ 1.00,\ 50\} \\ u &= \{5.90,\ 2.00,\ 25,\ 5.90,\ 2.00,\ 100\}. \end{aligned}$$

Moreover, since bound constraints alone do not guarantee that the rotational speed reaches synchronization speed (ω_{s}), the simulation model therefore imposes the following hard constraints, causing any infeasible evaluation to be heavily penalized. These constraints are defined as:

$$\begin{aligned} &\omega\,(t_{\text{s}}) = \omega_{\text{s}}, \\ &t_{\text{s}}(\psi) \le T_{\text{s}}, \\ &K \text{ starts must fit within } T_{\text{budget}}, \end{aligned}$$

where all constraints are verified within the same hydroelectric power unit model that evaluates $\tau(\psi)$, ensuring that proposed startups are dynamically valid and executable within the campaign window.

2.4 Proposed method

The proposed framework implements the informativeness-driven optimization of Section 2 through a sequential, constraint-aware blackbox procedure. Each candidate startup configuration, defined by the vector ψ, is evaluated by a numerical simulation model of the hydroelectric power unit dynamics. This model consists in a system of nonlinear ordinary differential equations that represents the dynamic interactions between the turbine and the speed governor; its parameters are identified from historical/commissioning records, and validated to reproduce $\alpha(t)$, $\beta(t)$, $\omega(t)$ and t_{s} within acceptable tolerances [8]. For any ψ, the simulation provides both the trajectory $\tau(\psi)$ and the feasibility checks at

a cost suitable for iterative optimization. Because the diversity objective in (5) depends on simulated trajectories, uses non-smooth elastic distances, and involves discrete feasibility (e.g., failure to reach synchronous speed), the problem is nonconvex and lacks analytical gradients. We therefore use MADS, a derivative-free optimization algorithm. In each iteration, MADS conducts a *poll step*, which explores neighboring points on an adaptive mesh to guarantee local coverage, and a *search step*, which samples more globally promising regions to accelerate progress. We refer to this setup as a BBO framework, in which the optimizer interacts with the model solely through input–output evaluations of candidate configurations.

2.5 Workflow

The optimization loop implements the sequential informativeness strategy defined in (5). Each iteration consists of four coordinated stages that connect the simulation model model, constraint evaluation, and optimizer:

1. The MADS algorithm proposes a new startup vector ψ within the feasible domain $\mathcal{F}$.
2. The simulation model is used to predict the hydroelectric power unit dynamics for ψ, yielding $\tau(\psi) = \{(\alpha(t), \beta(t), \omega(t))\}$, and verifies all operational constraints.
3. The candidate trajectory is resampled and normalized through $\mathcal{N}(\cdot)$, and because NOMAD is strictly a minimization optimizer, we express its diversity score objective function as the negative of the minimum distance to all previously accepted trajectories:

$$J(\psi) = - \min_{\substack{\psi_i \\ \in \mathcal{S}_{m-1}}} d\left(\mathcal{N}\left(\tau(\psi)\right), \mathcal{N}\left(\tau(\psi_i)\right)\right). \tag{6}$$

4. If the candidate satisfies all feasibility checks and minimizes the negative of the current minimum pairwise distance (in other words, maximizes the current minimum pairwise distance), it is accepted into the set: $\mathcal{S}_m = \mathcal{S}_{m-1} \cup \{\psi\}$.

The entire workflow is depicted in Figure 2. This "propose–evaluate–accept" loop repeats until either (i) the target number of feasible startups is reached ($|\mathcal{S}_m| = K$) or (ii) the optimization or campaign time budget T_{budget} is exhausted. The final set $\mathcal{S}_K$ thus contains a diverse, feasible set of startup sequences ready for execution during the field campaign.

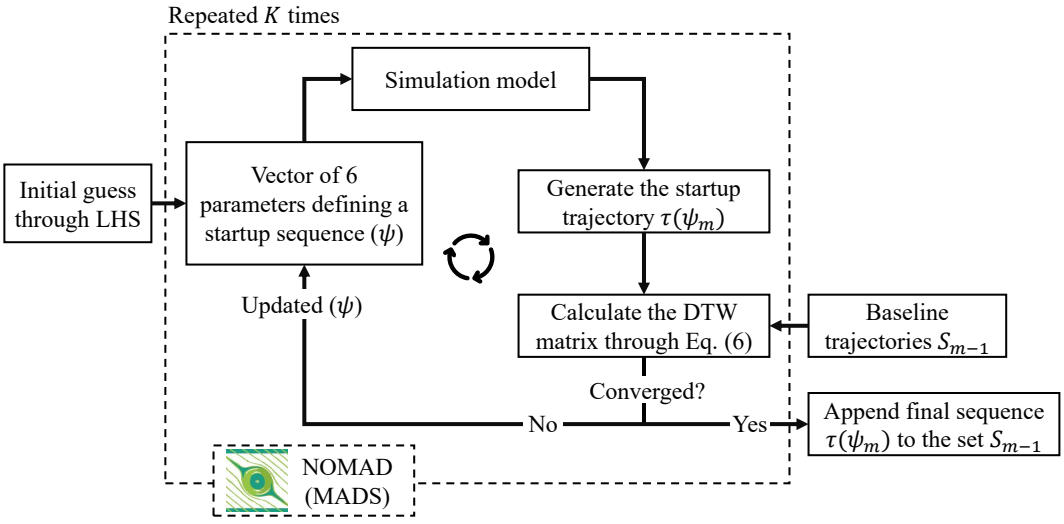

Figure 2: Optimization Workflow of the startup sequences. The process is repeated K times equal to the total number of startup sequences planned in a measurement campaign, or until the time budget T_{budget} is reached

2.6 Blackbox structure and constraint handling

The blackbox evaluation is the interface between the optimizer and the simulation model. Each MADS evaluation produces the three signals $\alpha(t)$, $\beta(t)$, and $\omega(t)$. Candidates that violate any constraint are penalized and therefore rejected. This strict feasibility enforcement ensures that the optimizer never considers physically non-executable or non-compliant trajectories.

The solver operates over the six-dimensional decision space defined in (1). An initial set of candidate points is generated through LHS to ensure broad coverage of the feasible region. At each subsequent iteration, MADS adaptively refines its search mesh around promising regions using the poll and search framework.

New candidates are accepted only if they improve the diversity score or feasibility margins. Convergence is declared when mesh refinements no longer yield feasible candidates that increase the minimum pairwise trajectory distance. This provides a principled stopping condition, rather than relying on heuristic iteration caps. Termination typically occurs when the solver reaches the target number of feasible startups K or the time limit $T_{\max}$. The final output is the set $\mathcal{S}_K$ of K feasible startups, each with its trajectory-distance metrics. This set represents the most diverse collection of startup behaviors attainable within the campaign budget T_{budget}, and is expected to maximize information gain for model calibration and uncertainty reduction.

3 Numerical experiments

This section reports the offline selection of informative startup sequences using the greedy rule in (5) and the objective in (6). Distances are computed with DTW after light preprocessing $\mathcal{N}$ that normalizes channels and resamples each trajectory to 100 points on a 250 s window. Raw traces are recorded at $\Delta t = 0.1$ s. For all reported experiments, we configured the optimization process to generate ten additional startup sequences. This number was chosen arbitrarily to avoid excessive visual clutter in trajectory plots. All ten optimization runs had a total time budget T_{budget} of 10,000 s wall time, and all accepted candidates satisfy feasibility constraints. The final set $\mathcal{S}_K$ thus contains $K = 15$ diverse and feasible startups, combining five baseline sequences with ten optimized additions. Results are reported for both the informativeness-driven selection using MADS and a classical space-filling benchmark using LHS, to compare diversity outcomes.

3.1 Baseline set

Five baseline sequences (*Seq 1*–*Seq 5*) supplied by the operator define the initial startup sequences set. Their kinematic trajectories $(\alpha(t), \beta(t), \omega(t))$ are shown in Figure 3. The panel also includes the provisional sixth candidate labeled *Target* (red $\times$), which is not part of the baseline but provisionally introduced by the optimizer to evaluate its distance from the current set. The 3D view omits the monotonically increasing time axis for clarity while preserving the sequence shapes.

3.2 Sequential selection with MADS

We use MADS algorithm to drive the simulation model at each evaluation. Each optimization run experiments hundreds of feasible sequences and returns the single candidate that maximizes the minimum elastic distance to the current set, per (5). The accepted candidate is appended to the set and becomes an input for the next run.

Figure 4 summarizes distances produced by all ten runs. The figure shows the lower triangular DTW matrix D, where each element D_{ij} is the distance between normalized trajectories τ_i and τ_j. By construction $D_{ii} = 0$, and symmetry $D_{ij} = D_{ji}$ makes the lower triangle sufficient. The last row corresponds to the current *Target*. At iteration m, the optimizer has already selected $\mathcal{S}_{m-1} =$

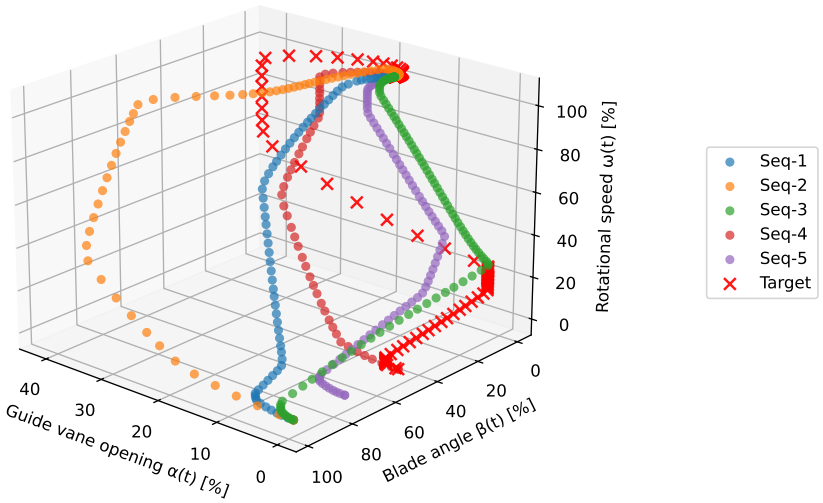

Figure 3: 3D plot of the initial five baseline sequences in $\alpha(t)$, $\beta(t)$ and $\omega(t)$ with the sixth provisional sequence Target

$\{\psi_1, \ldots, \psi_{m-1}\}$. A new candidate ψ_m is proposed, its distances to $\mathcal{S}_{m-1}$ fill the last row of a provisional matrix, and the optimizer focuses on the smallest off–diagonal entry in that row (the nearest neighbor). Maximizing this minimum is equivalent to minimizing $J(\psi_m)$ in (6). This rule expands coverage in trajectory space at each step.

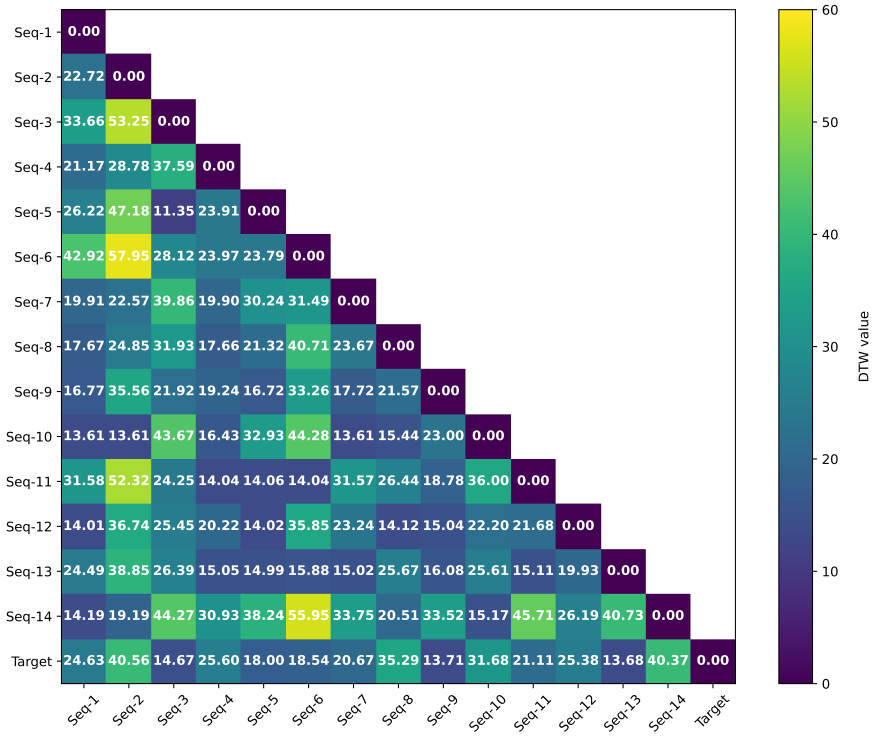

Figure 4: Pairwise DTW matrix for the optimization runs with MADS

After adding ten optimized sequences, the set contains *Seq 1–Seq 15*. The resulting $(\alpha(t), \beta(t), \omega(t))$ overlays are shown in Figure 5. The speed projections in Figure 6 show that all sequences reach 100% of synchronous speed in distinct trajectories.

For the ten additions chosen by MADS, the per-addition nearest-neighbor DTW values are: {23.79, 19.90, 17.71, 16.72, 13.61, 14.04, 14.01, 14.99, 14.19, 13.68}, which represent the minimum value of each row starting from Seq 6 to Seq 15 in Figure 4. Additional calculated statistics include a mean of 16.26, median of 14.59, range from 13.61 to 23.79 and a sum of 162.64.

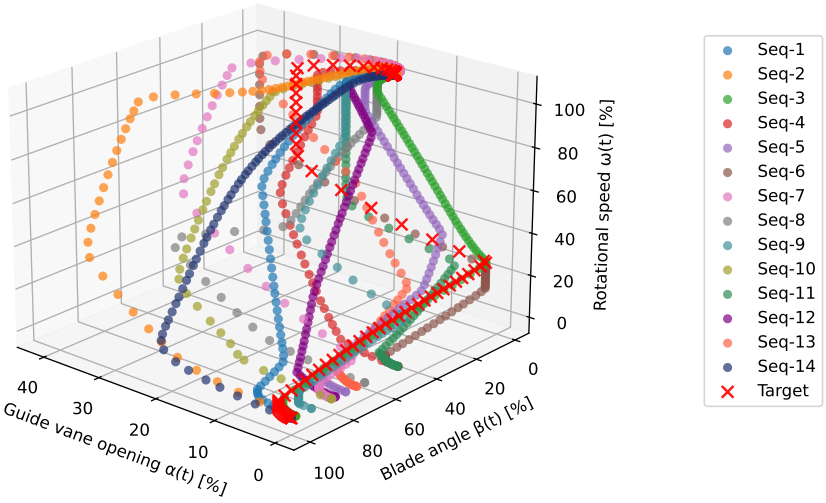

Figure 5: 3D plot of the five baseline sequences with the ten sequences selected by MADS (Seq 6–Seq 15)

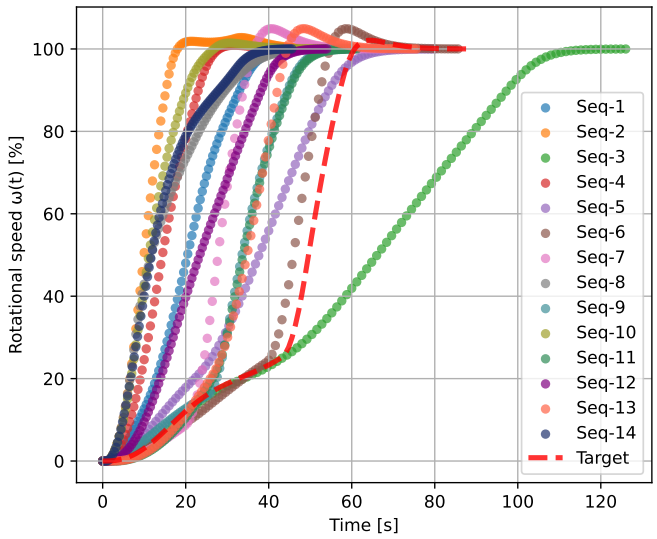

Figure 6: Rotational speed overlays for all MADS–generated sequences. All sequences reach 100% of synchronous speed in distinct trajectories

3.3 Classical space-filling baseline (LHS)

For benchmarking, we generated ten additions using a LHS within the same design space bounds and feasibility rules. The DTW matrix for this LHS startup sequences set is shown in Figure 7, and the corresponding trajectory overlay is in Figure 8.

For the ten LHS additions, the per-addition nearest-neighbor DTW values are {10.58, 10.27, 6.62, 11.31, 6.05, 10.72, 2.83, 6.86, 7.08, 5.96}, which represent the minimum value of each row starting from Seq 6 to Seq 15 in Figure 7. Additional calculated statistics include a mean of 7.83, median of 6.97, range from 2.83 to 11.31 and a sum of 78.28.

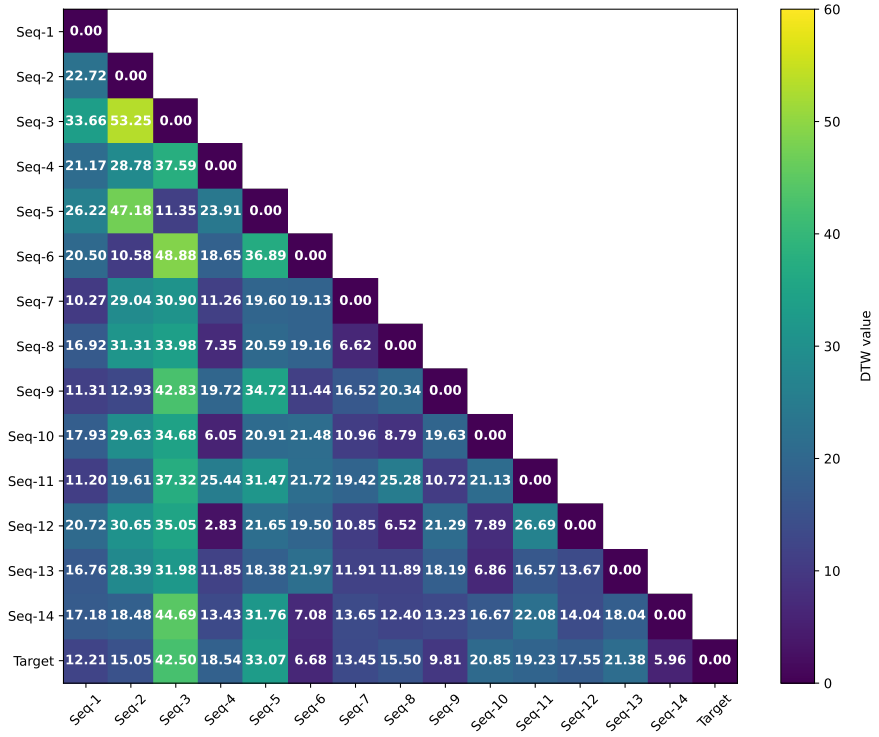

Figure 7: Pairwise DTW matrix for the LHS-generated sequences

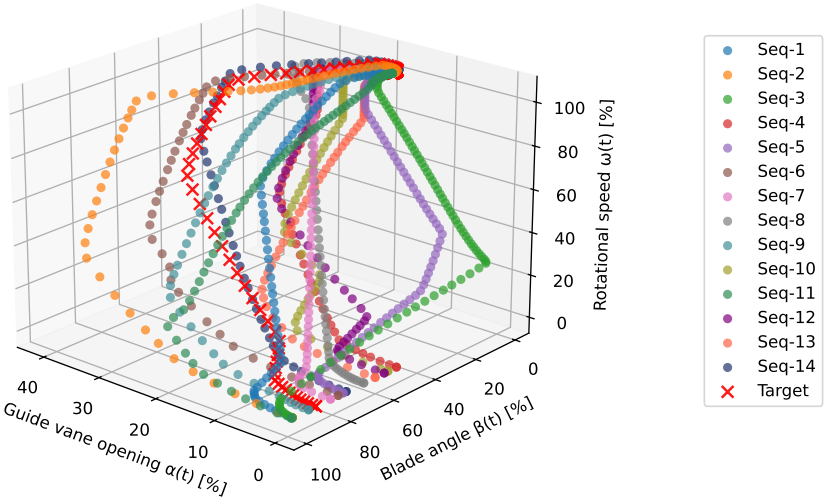

Figure 8: 3D plot of the five baseline sequences with the ten LHS-generated sequences

3.4 Discussion

The results indicate that selecting sequences by maximizing nearest neighbor distance in output trajectory space produces broader and more even coverage than sampling in input space. This is clear both visually and numerically. Figure 8 shows the ten added LHS sequences as more congested in trajectory space, while Figure 5 shows the ten MADS-selected sequences as more evenly spread, indicating broader coverage. The DTW matrices show higher separation for the optimized set than for the LHS (Figures 4 and 7). Numerically, across the ten additions the mean nearest neighbor DTW

is 16.26 for MADS versus 7.83 for LHS; the medians are 14.59 versus 6.97, and the totals are 162.64 versus 78.28. The optimized set therefore more than doubles spacing on average.

Two aggregate views help explain why. First, the per addition nearest neighbor DTW declines as coverage grows, which is consistent with a farthest point design on trajectories. For MADS, the average of the first three additions is 20.47 and the average of the last three is 14.29; for LHS the corresponding averages are 9.16 and 6.63 (Figure 9). Early additions close large gaps in the optimization-driven approach, and later additions fill smaller ones. Second, when we look at the per addition sum of distances to the already selected set, the optimization run adds more total coverage at every step (Figure 10). The coverage here is quantified as the total sum of each row in the DTW matrix. The per addition sums average 243.88 for MADS versus 186.14 for LHS (medians 229.58 versus 186.85), and the ten step totals are 2438.83 versus 1861.38.

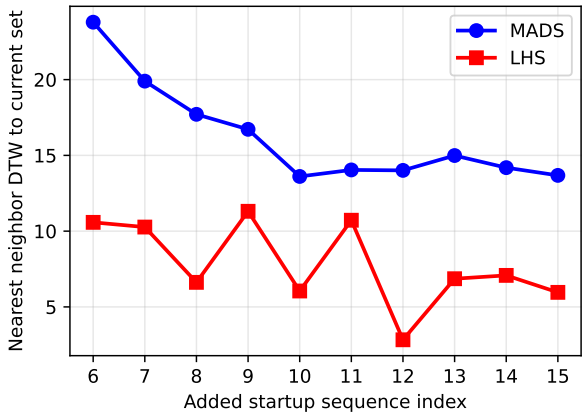

Figure 9: Per addition nearest neighbor DTW for the two sets. Early gains are large in MADS then taper as coverage improves

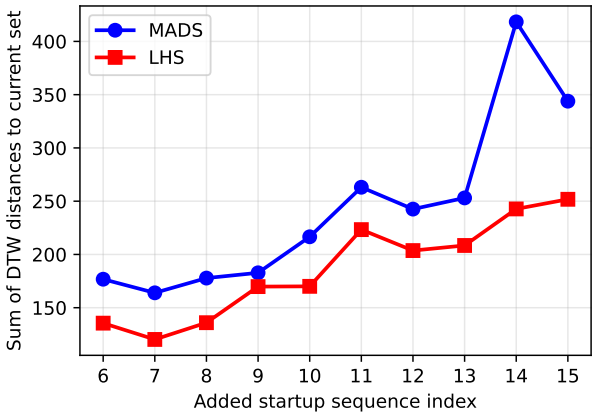

Figure 10: Per addition sum of distances to the already selected set. The set optimized with MADS adds more total coverage at most steps

These findings support a simple conclusion. If a credible simulation model is available, it is better to optimize diversity, namely in the shapes of $\alpha(t)$, $\beta(t)$, and $\omega(t)$ in the scope of this work. This targets a reduction of epistemic uncertainty directly, rather than relying on input spread as a proxy. If confidence in the simulation model is low, a classical LHS or a full factorial remain a more reasonable fallback.

The sequential design offers two practical benefits. It is interruptible, so the process can stop at a natural breakpoint and still retain the best additions found so far. It also provides a clear priority order.

For example, if ten startup sequences are planned but only eight fit within the experimental campaign window, execute the first eight returned by the optimizer since the startup sequences returned by the optimizer are ranked by informativeness. The choice of ten sequences is illustrative rather than intrinsic. The framework can be used for any number of startup sequences by running the optimizer sequentially. If more sequence are needed, just resume the sequence selection to extend the set.

The sequential design also offers a way to assess whether the number of startup sequences is sufficient to capture the diversity of machine behavior simulated to improve the generalizability of predictive models from DTs. By observing how the nearest neighbor DTW values evolve during selection, practitioners can gauge the marginal value of additional sequences. A consistent drop in DTW indicates improving coverage of the trajectory space, which is beneficial for reducing epistemic uncertainty in predictive models. However, when the decline becomes marginal, it suggests that further additions may offer limited new information. In contrast, if DTW values remain high and the drop is slow, it may be worth extending the measurement campaign to better capture the range of possible behaviors. This insight comes at no extra cost, as it leverages the same optimization data used for selection.

4 Conclusion

The developed work presented an offline procedure to plan informative Kaplan turbine startup tests when real-time measurements are unavailable. The method selects feasible governor-executable sequences by maximizing DTW-based diversity of the predicted transient trajectories under operational constraints. This serves both as a necessary initialization step to the DT, and as a contingency during the measurement campaign, allowing DT functionality to proceed when real-time measurement data connectivity is temporarily unavailable or disrupted. Instead of relying on real-time strain feedback and representativeness used in active learning-driven campaigns, the proposed method provides informativeness by maximizing diversity in multidimensional startup trajectories $(\alpha(t), \beta(t), \omega(t))$. Candidate startups are parameterized in terms of governor-executable ramps and plateaus, simulated through a site-calibrated simulation model of the hydroelectric power unit dynamics, and evaluated under strict feasibility constraints. We quantified informativeness as coverage in trajectory space using elastic distances, and we formulate sequence selection as a greedy max–min problem, in which each new startup is chosen to be as far as possible in output trajectory space from those already selected. The resulting BBO problem is solved using the derivative-free MADS algorithm.

The broader implication is that startup transient tuning can be reframed as an offline experimental design problem in trajectory space, rather than an online reaction to real-time measurements. The result is a measurement campaign that is more informative and more efficient for model calibration and reliability assessment. Future work will integrate physics-informed strain estimators and online updates, which will enable a unified acquisition policy that balances representativeness and informativeness to support a fully connected and robust DT capable of continuous learning and operation.

Statements and declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

M. Fahed: Conceptualization, Methodology, Software, Data Curation, Investigation, Visualization, Writing - Original Draft

M. Gagnon: Conceptualization, Methodology, Investigation, Resources, Writing – review & editing

S. Alarie: Conceptualization, Investigation, Writing – review & editing

J.-P. Gauthier: Investigation, Writing – review & editing

A. Favrel: Methodology, Software, Investigation, Writing – review & editing

M. Kokkolaras: Conceptualization, Methodology, Supervision, Writing – review & editing

References

[1] M. Gagnon, A. Tahan, P. Bocher, and D. Thibault. Impact of startup scheme on francis runner life expectancy. In IOP Conf. Ser.: Earth Environ. Sci., volume 12, page 012107, 2010. doi: https://doi.org/10.1088/1755-1315/12/1/012107.

[2] C. Trivedi, B. Gandhi, and M. Cervantes. Effect of transients on francis turbine runner life: a review. J. Hydraul. Res., 51(2):121–132, 2013. doi: https://doi.org/10.1080/00221686.2012.732971. URL

[3] M. Gagnon, N. Jobidon, M. Lawrence, and D. Larouche. Optimization of turbine startup: Some experimental results from a propeller runner. In IOP Conf. Ser.: Earth Environ. Sci., volume 22, page 032022, 2014. doi: https://doi.org/10.1088/1755-1315/22/3/032022. URL

[4] X. Liu, Y. Luo, and Z. Wang. A review on fatigue damage mechanism in hydro turbines. Renew. Sustain. Energy Rev., 54:1–14, 2016. doi: https://doi.org/10.1016/j.rser.2015.09.013. URL

[5] A. S. Dehkharqani, F. Engström, J. Aidanpää, and M. J. Cervantes. Experimental investigation of a 10 mw prototype kaplan turbine during start-up operation. Energies, 12(23):4582, 2019. doi: https://doi.org/10.3390/en12234582.

[6] M. Gagnon, D. Thibault, and M. Blain. On the expected monetary value of hydroelectric turbine start-up protocol optimisation. In J. Liyanage, J. Amadi-Echendu, and J. Mathew, editors, Engineering Assets and Public Infrastructures in the Age of Digitalization, Lect. Notes Mech. Eng., pages 209–216. Springer, Cham, 2020. doi: https://doi.org/10.1007/978-3-030-48021-9_24.

[7] T. Muser, E. Krymova, A. Morabito, M. Seydoux, and E. Vagnoni. Fatigue damage reduction in hydropower startups with machine learning. Nat. Commun., 16(1):2961, 2025. doi: https://doi.org/10.1038/s41467-025-58229-z.

[8] V. Mai, Q. H. Pham, A. Favrel, J.-P. Gauthier, and M. Gagnon. Active learning-based optimization of hydroelectric turbine startup to minimize fatigue damage. Renew. Energy, 256:124088, 2025. doi: https://doi.org/10.1016/j.renene.2025.124088.

[9] B. Settles. Active learning literature survey. Technical Report 1648, University of Wisconsin–Madison, 2009. URL https://minds.wisconsin.edu/handle/1793/60660.

[10] J. Shen and D. Tao. Exploring representativeness and informativeness for active learning. IEEE Trans. Cybern., 47(1):14–26, 2017. doi: https://doi.org/10.1109/TCYB.2016.2537180.

[11] R. W. Kennard and L. A. Stone. Computer aided design of experiments. Technometrics, 11(1):137–148, 1969. doi: https://doi.org/10.1080/00401706.1969.10490666.

[12] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 42(1):55–61, 2000. doi: https://doi.org/10.1080/00401706.2000.10485979.

[13] Max D. Morris and Toby J. Mitchell. Exploratory designs for computational experiments. J. Stat. Plann. Inference, 43(3):381–402, 1995. doi: https://doi.org/10.1016/0378-3758(94)00035-T.

[14] Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. Machine Learning, 110:457–506, 2021. doi: https://doi.org/10.1007/s10994-021-05946-3.

[15] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust., Speech, Signal Process., 26(1):43–49, 1978. doi: https://doi.org/10.1109/TASSP.1978.1163055.

[16] T. Eiter and H. Mannila. Computing discrete Fréchet distance. Technical Report CD-TR 94/64, TU Wien, 1994.

[17] P. F. Marteau. Time warp edit distance with stiffness adjustment for time series matching. IEEE Trans. Pattern Anal. Mach. Intell., 31(2):306–318, 2009. doi: https://doi.org/10.1109/TPAMI.2008.76.

[18] H. Sarin, M. Kokkolaras, G. D. Hulbert, P. Y. Papalambros, S. Barbat, and R.-J. Yang. Comparing time histories for validation of simulation models: Error measures and metrics. ASME J. Dyn. Syst., Meas., Control, 132(6):061401, 2010. doi: https://doi.org/10.1115/1.4002478.

[19] M. Cuturi and M. Blondel. Soft-DTW: A differentiable loss function for time series. In Proc. ICML, pages 894–903, 2017.

[20] M. Blondel, A. Mensch, and J.-P. Vert. Differentiable divergences between time series. In Proc. AISTATS, volume 130, pages 3857–3865, 2021.

[21] H. Janati, M. Cuturi, and A. Gramfort. Spatio-temporal alignments: Optimal transport through space and time. In Proc. AISTATS, pages 1695–1704, 2020.

[22] I. Paparrizos, L. Gravano, and T. Jayaraman. Debunking four long-standing misconceptions of time-series distance measures. In Proc. ACM SIGMOD, pages 1887–1906, 2020. doi: https://doi.org/10.1145/3318464.3389760.

[23] M. Vlachos, G. Kollios, and D. Gunopulos. Discovering similar multidimensional trajectories. In Proc. IEEE ICDE, pages 673–684, 2002. doi: https://doi.org/10.1109/ICDE.2002.994784.

[24] C. Audet and Jr. Dennis, J. E. Mesh adaptive direct search algorithms for constrained optimization. SIAM J. Optim., 17(1):188–217, 2006. doi: https://doi.org/10.1137/040603371.

[25] C. Audet, S. Le Digabel, V. Rochon Montplaisir, and C. Tribes. Algorithm 1027: NOMAD version 4: Nonlinear optimization with the MADS algorithm. ACM Trans. Math. Softw., 48(3):35:1–35:22, 2022. doi: https://doi.org/10.1145/3544489.

[26] S. Du Bois, A. Basudhar, K. Witowski, I. Gandikota, and N. Stander. An LS-OPT methodology for utilizing partial curve data for the calibration of material models. In Proc. 12th European LS-DYNA Conf., 2019.

[27] K. Witowski and N. Stander. Modified dynamic time warping for utilizing partial curve data to calibrate material models. In Proc. 16th Int. LS-DYNA Users Conf. (Optimization Track), June 2020. 16th International LS-DYNA Users Conference, June 10–11, 2020.

[28] G. He, Y. Duan, Y. Li, T. Qian, J. He, and X. Jia. Active learning for multivariate time series classification with positive unlabeled data. In Proc. 27th Int. Conf. Tools with Artif. Intell. (ICTAI), pages 178–185, Vietri sul Mare, Italy, 2015. doi: https://doi.org/10.1109/ICTAI.2015.38.

[29] A. Candelieri, R. Perego, and F. Archetti. Bayesian optimization of pump operations in water distribution systems. J. Global Optim., 71(1):213–235, 2018. doi: https://doi.org/10.1007/s10898-018-0641-2.

[30] A. Candelieri, A. Ponti, I. Giordani, and F. Archetti. Lost in optimization of water distribution systems: Better call Bayes. Water, 14(5):800, 2022. doi: https://doi.org/10.3390/w14050800.

[31] K. Alnajim and A. A. Abokifa. Bayesian optimization for contamination source identification in water distribution networks. Water, 16(1):168, 2024. doi: https://doi.org/10.3390/w16010168.

[32] K. Alnajim and A. A. Abokifa. Operational response to contamination in water distribution systems: a multi-objective Bayesian optimization approach. Front. Water, 7:1547112, 2025. doi: https://doi.org/10.3389/frwa.2025.1547112.

[33] J. Guo, Y. Zhang, and W. Wang. Predicting hydro turbine failures through digital twin simulations of rare real-world data. J. Comput. Methods Sci. Eng., 0(0), 2025. doi: https://doi.org/10.1177/14727978241304260.

[34] E. Salomons and M. Housh. Practical real-time optimization for energy efficient water distribution systems operation. J. Clean. Prod., 275:124148, 2020. doi: https://doi.org/10.1016/j.jclepro.2020.124148.

[35] S. Hu, J. Gao, D. Zhong, R. Wu, and L. Liu. Real-time scheduling of pumps in water distribution systems based on exploration-enhanced deep reinforcement learning. Systems, 11(2):56, 2023. doi: https://doi.org/10.3390/systems11020056.

[36] G. Gu, Y. Tang, and F. Qian. Kalman filtering over the random delay and packet drop channel. SIAM J. Control Optim., 59(4):2454–2476, 2021. doi: https://doi.org/10.1137/20M1379903.

[37] W. Li, Y. Jia, and J. Du. Distributed kalman consensus filter with intermittent observations. J. Franklin Inst., 352(9):3764–3781, 2015. ISSN 0016-0032. doi: https://doi.org/10.1016/j.jfranklin.2015.01.002.

[38] S. Gao, P. Chen, D. Huang, and Q. Niu. Stability analysis of multi-sensor kalman filtering over lossy networks. Sensors, 16(4):566, 2016. doi: https://doi.org/10.3390/s16040566.

[39] C. Shan, R. Li, and J. Cheng. A new robust kalman filter with measurement loss based on mixing distribution. Trans. Inst. Meas. Control, 44(10):2018–2030, 2022. doi: https://doi.org/10.1177/01423312211054942.