mathematics

MDPI

Article

A Bounded Adaptive Random Local Mutation Algorithm for Sparse-Data Parameter Identification in Nonlinear Bioprocess Models

Velislava Lyubenova [1,*,†], Olympia Roeva [2,3,†], Elena Chorukova [4,†], Maya Ignatova [1,†], Anastasiya Zlatkova [1], Denitsa Kristeva [1] and Gergana Roeva [1]

[1] Department of Mechatronic Bio/Technological Systems, Institute of Robotics, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 2, 1113 Sofia, Bulgaria; m_ignatova@ir.bas.bg (M.I.); a_zlatkova@ir.bas.bg (A.Z.); d_kristeva@ir.bas.bg (D.K.); g_roeva@ir.bas.bg (G.R.)
[2] Department of Bioinformatics and Mathematical Modelling, Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 105, 1113 Sofia, Bulgaria; olympia@biomed.bas.bg
[3] Centre of Excellence in Informatics and Information and Communication Technologies, 1113 Sofia, Bulgaria
[4] Department Biotechnology–Bioremediation and Biofuels, The Stephan Angeloff Institute of Microbiology, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 26, 1113 Sofia, Bulgaria; elena@microbio.bas.bg
* Correspondence: v_lyubenova@ir.bas.bg
† These authors contributed equally to this work.

Abstract

Sparse and unevenly informative observations can substantially limit parameter estimation in nonlinear bioprocess models. This proof-of-concept study investigates Bounded Adaptive Random Local Mutation (B-ARLM), a derivative-free algorithm for bounded parameter calibration, using a product-inhibited Andrews-type model of biological hydrogen production calibrated to six interval-average observations over 0–8 days. Andrews–P achieved the lowest in-sample calibration error among the three examined kinetic formulations, with a best-run NMSE of 1.29×10^{-3}. Under a common budget of 2250 objective-function evaluations and 30 repeated runs, B-ARLM significantly outperformed the Fixed-EA ablation and the study-specific sCA-ES reference, while its terminal performance was not statistically distinguishable from canonical CMA-ES or jDE after Holm correction. Hyperparameter, broad-start, profile-NMSE, and observation-error sensitivity analyses were additionally performed. Broad-start optimization and two-dimensional profiling revealed widely separated near-optimal parameter vectors and pronounced parameter compensation, showing that accurate trajectory reconstruction does not imply unique kinetic-parameter recovery. The results support B-ARLM as a competitive bounded calibration strategy for the examined sparse-data problem while emphasizing the distinction between optimization performance, practical identifiability, and biological interpretation.

Keywords: bounded adaptive random local mutation; derivative-free optimization; evolutionary optimization; sparse-data calibration; nonlinear bioprocess modelling; parameter identification; practical identifiability; parameter compensation; uncertainty analysis

MSC: 65K10; 62J02; 92C45

Academic Editor: Junyuan Yang

Received: 6 August 2026
Revised: 25 August 2026
Accepted: 26 August 2026
Published: 12 September 2026

1. Introduction

The problem of parameter estimation in biochemical pathways, formulated as a nonlinear programming problem subject to the pathway model acting as constraints, has received great attention [1–3]. Reliable parameter identification in nonlinear biochemical and bioprocess models is often challenging and may depend on the model structure, parameter domain, and calibration procedure [4–6]. Biochemical process engineering faces substantial challenges in dynamic modeling, model identification, and state estimation. Key areas of focus include recursive parameter estimation, as well as advanced control strategies, such as adaptive linearizing control and extremum-seeking control [7–9]. Related studies of nonlinear biological dynamics further illustrate the importance of accurately representing temporal and memory-dependent behaviour when modelling complex biological systems [10]. The difficulties become more pronounced when observations are sparse, noisy, unevenly distributed in time, or limited to a single measured output [11–14].

A low calibration objective indicates close agreement between simulated and observed trajectories, but it does not guarantee unique or precise recovery of the underlying kinetic parameters [11,12]. Because of limited observational support and parameter compensation, substantially different parameter combinations may produce similar trajectories. Variability introduced by the optimizer must therefore be distinguished from limitations in parameter recovery [13–15]. Calibration results should consequently be interpreted together with practical-identifiability and trajectory-uncertainty diagnostics [16–20].

Derivative-free optimization methods are widely used for nonlinear dynamic-model calibration because they do not require analytical derivatives and can address multimodal and ill-conditioned objective landscapes [21–23]. By relying solely on direct objective function evaluations, these algorithms bypass the need to compute complex symbolic or numerical gradients, which are often unavailable or unreliable in highly nonlinear system models. This structural independence makes them particularly well-suited for navigating intricate, non-convex parameter spaces that feature multiple local minima, heavy noise, or severe ill-conditioning [24,25]. Consequently, they provide a robust and adaptable framework for parameter identification in complex dynamic systems where classical gradient-based approaches frequently fail or stall [26,27].

Evolutionary methods differ in how they represent and adapt the mutation distribution. Fixed-mutation strategies retain prescribed amplitudes, whereas success-rule and self-adaptive evolution strategies modify one or more search-scale parameters during the run [28–32]. Coordinate-dependent methods introduce separate directional scales, while CMA-ES additionally learns cross-parameter correlations and rotations of the search distribution through covariance adaptation [28,33,34]. Differential Evolution uses population differences to generate trial vectors [24–26], and adaptive variants such as jDE modify their control parameters during the search [35]. Bayesian optimization follows a different surrogate-based paradigm and is particularly relevant when forward-model evaluations are expensive [23].

This work examines Bounded Adaptive Random Local Mutation (B-ARLM), a bounded derivative-free search mechanism in which each active parameter is assigned a normalized mutation amplitude. A generation-level success signal provides common expansion or contraction, while successful candidate displacements provide coordinate-wise scale information. The amplitudes remain within prescribed bounds, mutation is centred on the best-so-far solution, and elitism retains the current best candidate. Unlike CMA-ES, B-ARLM does not estimate a covariance matrix or evolution paths. The individual ingredients are established concepts; the contribution lies in their specific integration within a lightweight bounded best-centred mechanism.

B-ARLM is evaluated through repeated runs, a Fixed-EA ablation, the study-specific sCA-ES reference, and matched-budget comparisons with canonical CMA-ES and jDE. Hyperparameter sensitivity is examined separately. Calibration quality is assessed together with practical parameter support using profile NMSE, local Fisher-information diagnostics, broad-start optimization, two-dimensional profiling, leave-one-interval-out influence analysis, and conditional trajectory-uncertainty analyses under alternative prescribed observation-error models.

Accordingly, the contributions are: (i) the bounded adaptive B-ARLM mechanism; (ii) theoretical and numerical characterization of its adaptation behaviour and matched-budget optimizer performance; and (iii) an integrated sparse-data calibration analysis separating optimization performance from practical identifiability and trajectory uncertainty. The study is intended as a problem-specific proof of concept rather than evidence of universal optimizer superiority or unique kinetic-parameter recovery.

2. Related Work

2.1. Derivative-Free and Adaptive Optimization

Nonlinear dynamic-model calibration can be performed using local gradient-based methods, deterministic derivative-free procedures, or stochastic population-based algorithms [2,22]. Local methods may be efficient when a suitable initial estimate is available. Multimodality, simulation failures, ill-conditioning, and weak parameter support, however, often motivate the use of global search methods [36–38].

Evolution strategies differ in the amount of search-distribution information that is adapted during optimization. Classical success-rule methods primarily modify a global step size according to the frequency or history of successful trials, whereas self-adaptive strategies incorporate search parameters into the evolutionary process itself [29–32]. Coordinate-dependent and diagonal approaches, including diagonal/sep-CMA-ES variants, extend scalar adaptation by maintaining separate directional scales or diagonal covariance information, thereby accommodating anisotropy without estimating a full covariance matrix [34].

CMA-ES represents a more expressive adaptation mechanism because it jointly updates a global scale and a covariance matrix, allowing the search distribution to learn cross-parameter correlations and rotations [28,33]. Differential Evolution employs a different population-difference mutation mechanism [23–26]; adaptive variants such as jDE additionally modify control parameters during the run [35]. These distinctions are relevant to the present work because B-ARLM learns parameter-wise mutation scales but deliberately does not estimate cross-parameter covariance.

B-ARLM therefore occupies an intermediate position between scalar or coordinate-wise step-size adaptation and full covariance learning. It maintains bounded parameter-wise mutation amplitudes and combines a generation-level success signal with coordinate-wise scale information extracted from successful displacements. Unlike CMA-ES, no covariance matrix, evolution path, or rotational transformation of the search distribution is estimated. Unlike Differential Evolution, candidate displacements are not generated from differences between population members.

Table 1 summarizes broad method families rather than individual implementations. B-ARLM combines bounded coordinate-wise amplitudes with generation-level success control and successful-displacement information, but it does not estimate cross-parameter covariance. The numerical benchmark therefore focuses on Fixed-EA, the study-specific sCA-ES reference, canonical CMA-ES, and jDE under matched objective-evaluation budgets.

Meaningful stochastic-optimizer comparisons require matched evaluation budgets, repeated runs, and clearly reported initialization and evaluation procedures [39–41], which were applied here. Gaussian-process Bayesian optimization represents a different surrogate-

based derivative-free paradigm [23] and is discussed for context rather than included as a numerical comparator.

Table 1. Conceptual positioning of B-ARLM relative to representative mutation-scale adaptation mechanisms.

Method Family	Scale Representation	Primary Adaptation Signal	Cross-Parameter Covariance	Relation to B-ARLM
Fixed mutation	Fixed scalar or parameter-wise amplitudes	None	No	Control variant in which the adaptive amplitude updates are disabled
Scalar success-based ES	One adaptive global step size	Success rate or success history	No	Related to the generation-level multiplicative update
Coordinate-dependent/diagonal adaptation (e.g., diagonal/ sep-CMA-ES)	Separate coordinate-wise scales or diagonal covariance representation	Diagonal statistics or strategy-parameter updates	Usually absent or restricted	Related to parameter-wise scale learning
Full covariance adaptation	Global scale and covariance matrix	Ranked samples, evolution paths, and covariance updates	Yes	Represents correlations that B-ARLM does not estimate
B-ARLM	Bounded parameter-wise amplitudes	Generation-level success rate and median absolute coordinate-wise displacements of successful candidates	No	Combines bounded coordinate-wise amplitudes with success-rate and successful-displacement updates
Adaptive Differential Evolution	Mutation factor and crossover parameters	Population differences and adaptive control parameters	No explicit covariance matrix	Different population-difference mechanism; represented experimentally by jDE

2.2. Practical Identifiability and Uncertainty

Structural identifiability concerns whether model parameters can, in principle, be uniquely recovered from ideal observations for a specified model and observation structure, whereas practical identifiability concerns the support provided by finite and uncertain experimental data [42,43]. A structurally identifiable model may therefore remain practically difficult to calibrate when observations are sparse or parameter effects are strongly correlated.

Local sensitivity and Fisher-information analyses characterize parameter support in the neighbourhood of a selected calibration point [15,18,19]. Profile-based methods examine a broader parameter range by fixing one parameter and re-optimizing the remaining components [12,42]. These approaches are complementary, but local and one-dimensional diagnostics may fail to reveal separated low-error regions or nonlinear compensation between parameters. The present study therefore supplements the local FIM and one-dimensional profile NMSE with broad-start optimization over the admissible parameter domain and two-dimensional profile analysis.

Observation influence is examined separately through leave-one-interval-out recalibration. Because omission of one interval leaves only five observations for six Andrews–P parameters, this procedure is interpreted as an influence diagnostic rather than conventional predictive cross-validation.

Parameter uncertainty must also be distinguished from trajectory uncertainty. Substantially different parameter combinations may produce similar observable trajectories because of parameter compensation [11,12,15]. Conditional parametric bootstrap is therefore used to examine propagation of prescribed observation perturbations through recalibration. Its interpretation depends on the assumed error structure, perturbation magnitude, reference calibration, and optimization protocol [14,20,44]. Because replicate-derived error statistics are unavailable, the considered perturbation models are treated as sensitivity scenarios rather than empirically identified measurement-error distributions.

A formal structural-identifiability analysis is outside the scope of this study. Weak parameter support, broad profiles, and separated near-optimal solutions are therefore interpreted as practical limitations of parameter recovery under the examined observation design and are not attributed uniquely to structural non-identifiability.

3. Mathematical Model and Calibration Setting

The process is represented by a reduced-order product-inhibited model with biomass X, substrate S, and an accumulated inhibitory-product state P. The state P is interpreted as a lumped effective product-related inhibition pool rather than as a directly measured individual metabolite and is distinct from the measured hydrogen-production output $q_{\mathrm{H_2}}$. A cumulative hydrogen state H is introduced solely to implement the interval observation operator. The classical non-inhibited baseline follows Monod kinetics [45], whereas the inhibitory formulation is motivated by the Andrews inhibition principle [46] and by the broader use of combined substrate-product inhibition terms in biokinetic modelling [47]. In batch fermentation, accumulation of metabolic products can progressively reduce biological activity; the P-dependent term used here is intended to represent this effect at a phenomenological reduced-order level. The resulting formulation is denoted Andrews–P for the purposes of the present study and is not assumed to constitute a uniquely established mechanistic description of product inhibition.

The product-inhibited specific growth rate is

$$\mu(S, P; \boldsymbol{\theta}) = \frac{\mu_{\max} S}{K_S + S + P^2/k_{\mathrm{i}}}. \tag{1}$$

The quadratic product term is used as a study-specific phenomenological representation of progressively increasing inhibition. Its inclusion provides a flexible reduced-order description of product-dependent suppression but should not be interpreted as evidence that the underlying biological inhibition necessarily follows this particular quadratic mechanism. The state equations are

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \mu(S, P; \boldsymbol{\theta})X, \ \frac{\mathrm{d}S}{\mathrm{d}t} = -\gamma_1 \mu(S, P; \boldsymbol{\theta})X, \ \frac{\mathrm{d}P}{\mathrm{d}t} = \gamma_2 \mu(S, P; \boldsymbol{\theta})X. \tag{2}$$

The instantaneous volumetric hydrogen-production rate is defined as

$$q_{H_2}(t; \boldsymbol{\theta}) = \gamma_{H_2} \mu(S(t), P(t); \boldsymbol{\theta})\, X(t). \tag{3}$$

Because the experimental observations represent rates averaged over finite sampling intervals rather than instantaneous measurements, an auxiliary cumulative-output state was introduced:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = q_{\mathrm{H_2}}(t; \boldsymbol{\theta}), \quad H(0) = 0. \tag{4}$$

Accordingly, $q_{\mathrm{H_2}}$ has units of $\mathrm{dm}^3_{\mathrm{H_2}} \cdot \mathrm{L}^{-1} \mathrm{day}^{-1}$, whereas H has units of $\mathrm{dm}^3_{\mathrm{H_2}} \cdot \mathrm{L}^{-1}$.

The calibrated parameter vector is

$$\boldsymbol{\theta} = \left(\mu_{\max}, K_s, \gamma_1, \gamma_2, \gamma_{H_2}, k_i\right) \in \Theta. \tag{5}$$

The initial conditions were fixed throughout all calibration and diagnostic analyses as

$$X(0) = 0.1\ \mathrm{g{\cdot}L^{-1}},\ S(0) = 10\ \mathrm{g{\cdot}L^{-1}},\ P(0) = 0.1\ \mathrm{g{\cdot}L^{-1}},\ H(0) = 0\ \mathrm{dm^3_{H_2}{\cdot}L^{-1}}.$$

These initial-state values were prescribed model inputs and were not included among the calibrated quantities.

Calibration was based on $n = 6$ interval-average volumetric hydrogen-production rates. The normalized least-squares objective was

$$J(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\overline{q}_i(\boldsymbol{\theta}) - \overline{q}_i^{\mathrm{obs}}}{q_{\max}}\right]^2, q_{\max} = \max_{1\le i\le n} \mid \overline{q}_i^{\mathrm{obs}} \mid > 0. \tag{6}$$

For the $i-th$ observation interval $[a_i, b_i]$, the corresponding model prediction was defined by the cumulative-output difference

$$\overline{q}_i(\boldsymbol{\theta}) = \frac{H(b_i;\boldsymbol{\theta}) - H(a_i;\boldsymbol{\theta})}{b_i - a_i}, i = 1, \ldots, 6.$$

Here, $\overline{q}_i^{\mathrm{obs}}$ denotes the observed interval-average volumetric hydrogen-production rate over the same interval. Because all residuals were divided by the same positive quantity $q_{\max}$, Equation (6) rescales the objective function but does not apply observation-specific statistical weighting.

The admissible parameter domain is

$$\begin{aligned}\Theta = {} & [0.024,\ 22.8] \times [0.001,\ 10] \times [0.01,\ 10] \times [0.01,\ 10] \times \\ & [0.0001,\ 0.7] \times [0.1,\ 100].\end{aligned} \tag{7}$$

The intervals correspond, respectively, to $\mu_{\max}$, K_s, γ_1, γ_2, γ_{H_2}, and k_i. The same bounds were applied to all active parameters throughout the main calibration and all optimization-based diagnostic analyses. In the Monod formulation, k_i does not enter the kinetic expression and was therefore fixed at its nominal value and excluded from mutation and adaptation.

All objective-function evaluations used a fixed-step classical fourth-order Runge–Kutta scheme with

$$\Delta t = 0.02\ \mathrm{h} = \frac{0.02}{24}\ \mathrm{day}.$$

The system was integrated over the complete experimental duration, while only the cumulative output H at the required interval boundaries was retained for objective-function evaluation. Interval-average predictions were obtained from cumulative-output differences, avoiding full-trajectory storage and interpolation during optimization. The production discretization was verified against a reference step of 0.01 h. Nonnegative-state safeguards and the monotonicity constraint imposed on H are described in Supplementary Section S1.

For the primary calibration and kinetic-model comparison, the nominal values in Table 2 were used only to centre the initial bounded stochastic populations. They were not treated as prior estimates, independently measured kinetic constants, or expected calibrated values. The profile-NMSE and bootstrap analyses instead used their specified reference-centred initialization procedures. Substantial differences between the nominal and calibrated vectors are therefore admissible, particularly under parameter compensation and weak practical identifiability.

Table 2. Model parameters, definitions, units, nominal values, and calibration bounds.

Symbol	Definition	Unit	Nominal	Lower	Upper
$\mu_{\max}$	Maximum specific growth rate	day^{-1}	4.8	0.024	22.8
K_{s}	Substrate half-saturation constant	$\mathrm{g{\cdot}L}^{-1}$	0.05	0.001	10
γ_1	Substrate-consumption coefficient	$\mathrm{g_S{\cdot}g_X}^{-1}$	5.0	0.01	10
γ_2	Product-formation coefficient	$\mathrm{g_P{\cdot}g_X}^{-1}$	5.0	0.01	10
$\gamma_{\mathrm{H_2}}$	Hydrogen-production yield coefficient	$\mathrm{dm^3_{H_2}{\cdot}g_X}^{-1}$	0.0694	0.0001	0.7
k_{i}	Product-inhibition scale in P^2/k_{i}	$\mathrm{g{\cdot}L}^{-1}$	50.0	0.1	100

Here, $\mathrm{g_S}$, $\mathrm{g_P}$, and $\mathrm{g_X}$ denote grams of substrate, effective inhibitory product, and biomass, respectively, whereas L denotes litre of reactor working volume.

Experimental System and Data

The experimental dataset was generated specifically for the present study and has not been published previously. It was obtained from laboratory-scale batch anaerobic-fermentation experiments conducted for biological hydrogen production.

The experiments were performed in a bioreactor with a working volume of 3 L under mesophilic conditions at 35 ± 0.5 °C. The fermentation substrate consisted of a 1:1 mixture of corn steep liquor and cattle manure. A mixed microbial consortium obtained from the liquid phase of an operating biogas-production reactor was used as inoculum. The liquid fraction was sieved, centrifuged at 4500 rpm, washed twice with 0.9% NaCl solution, heat-treated at 75 °C for 30 min, cooled to room temperature, and introduced into the reactor at 10% (v/v) of the working volume.

The pH was maintained at 5.5 by automated addition of NaOH and HCl solutions. The hydrogen fraction in the produced biogas was measured using a Gasboard 3100P gas analyzer (Hubei Cubic-Ruiyi Instrument Co., Ltd., Wuhan, China). For each sampling interval, the hydrogen volume was obtained from the measured biogas volume and hydrogen fraction, and the interval-average volumetric production rate was calculated as

$$\bar{q}_i^{obs} = \frac{V_{H_2,i}}{V_R(b_i - a_i)},$$

where $V_{H_2,i}$ is the hydrogen volume produced during interval i, $V_R = 3$ L is the reactor working volume, and $(b_i - a_i)$ is the duration of the corresponding interval in days.

The calibration dataset comprised six interval-average rates over the intervals [0, 1], [1, 2], [2, 3], [3, 4], [4, 7], and [7, 8] days. The corresponding values were 0.95539, 0.27560, 0.11069, 0.15234, 0.00240, and 0 in $\mathrm{dm^3_{H_2}{\cdot}L_R^{-1}{\cdot}day^{-1}}$.

Replicate-level measurements and observation-specific variance estimates were not available to the modelling analysis. The available aggregated values were therefore used directly for calibration without observation-specific statistical weighting. The available dataset therefore contains six aggregated interval-average observations for six calibrated parameters in the Andrews–P formulation. No independent validation dataset was available, and a conventional train-test split was not performed because withholding observations would leave fewer calibration observations than unknown kinetic parameters. Accordingly, model comparisons in this study assess in-sample calibration performance rather than predictive generalization. The leave-one-interval-out analysis described below is used only as an observation-influence diagnostic and not as predictive cross-validation.

4. Forward-Model Well-Posedness, and Invariant Relations

The following properties establish nonnegativity, boundedness, and global existence for the continuous forward model. They also provide conditions under which the corresponding continuous calibration objective is well defined and attains a minimum on the compact admissible parameter domain. The numerical implementation, including

discretization, state safeguards, parameter projection, and the failure penalty, is considered separately below.

Proposition 1. *Positivity, boundedness, and invariant relations. Let the initial states be nonnegative and let all model parameters be strictly positive. Then, for every admissible parameter vector $\boldsymbol{\theta} \subset \Theta$, the unique solution of Equations (1)–(4) exists for all $t \geq 0$, remains nonnegative, and satisfies*

$$0 \leq S(t) \leq S(0),\ X(t) = X(0) + \tfrac{S(0)-S(t)}{\gamma_1},\ P(t) = P(0) + \tfrac{\gamma_2}{\gamma_1}[S(0) - S(t)],$$
$$\text{and } H(t) = H(0) + \tfrac{\gamma_{H_2}}{\gamma_1}[S(0) - S(t)].$$

Consequently,

$$X(t) \leq X(0) + \frac{1}{\gamma_1}S(0), \quad P(t) \leq P(0) + \frac{\gamma_2}{\gamma_1}S(0), \quad H(t) \leq H(0) + \frac{\gamma_{H_2}}{\gamma_1}S(0),$$

and the instantaneous hydrogen-production rate is nonnegative and bounded:

$$0 \leq q_{H_2}(t;\boldsymbol{\theta}) \leq \gamma_{H_2}\mu_{\max}\left[X(0) + \frac{S(0)}{\gamma_1}\right].$$

Proof. The vector field is locally Lipschitz continuous on the nonnegative state domain because the denominator in Equation (1) is bounded below by $K_s > 0$. Therefore, a unique local solution exists for every admissible parameter vector. For non-negative states and positive parameters, the denominator in Equation (1) satisfies

$$K_s + S + \frac{P^2}{k_i} \geq K_s > 0,$$

and therefore

$$0 \leq \mu(S,\ P,\ \boldsymbol{\theta}) \leq \mu_{\max}.$$

The behaviour of the vector field on each boundary of the nonnegative state domain can be verified explicitly. At $X = 0$ every reaction term contains the factor X, so $\mathrm{d}X/\mathrm{d}t = \mathrm{d}S/\mathrm{d}t = \mathrm{d}P/\mathrm{d}t = \mathrm{d}H/\mathrm{d}t = 0$, and the vector field is tangent to this boundary. At $S = 0$, the numerator of the growth-rate expression vanishes and therefore, $\mu = 0$; consequently, all four state derivatives vanish, and this boundary is also invariant. At $P = 0$, $\mathrm{d}P/\mathrm{d}t = \gamma_2$, $\mu X \geq 0$ and hence the vector field cannot point toward negative P. At $H = 0$, $\mathrm{d}H/\mathrm{d}t = \gamma_{H_2}\mu X \geq 0$, so the vector field likewise cannot point toward negative H. Thus, on every boundary face, the vector field is either tangent to or directed into the nonnegative state domain. The nonnegative state domain is positively invariant, and nonnegative initial states remain nonnegative for all $t \geq 0$. Moreover,

$$\frac{dS}{dt} = -\gamma_1\mu(S,P,\boldsymbol{\theta})X \leq 0,$$

so $S(t)$ is nonincreasing. From Equations (2) and (4),

$$\frac{dX}{dt} = -\frac{1}{\gamma_1}\frac{dS}{dt}, \qquad \frac{dP}{dt} = -\frac{\gamma_2}{\gamma_1}\frac{dS}{dt}, \quad \frac{dH}{dt} = -\frac{\gamma_{H_2}}{\gamma_1}\frac{dS}{dt}.$$

Integration yields the stated identities and bounds. Since all state variables remain bounded, finite-time blow-up is excluded, and the solution exists globally for all $t \geq 0$. □

For the continuous model, the vector field is continuous in the parameter vector $\boldsymbol{\theta}$, and, locally, Lipschitz is continuous in the state variables and applied uniformly over the compact admissible parameter set Θ. The corresponding forward solution therefore depends continuously on $\boldsymbol{\theta}$ over the finite experimental horizon. Each interval-average prediction, and hence the continuous objective defined in Equation (6), is consequently continuous on Θ. Because Θ is compact, the Weierstrass theorem guarantees that this continuous objective attains at least one global minimum on Θ. This existence statement does not imply uniqueness of the minimizer or practical identifiability of the calibrated parameters.

This argument applies to the ideal continuous forward model and objective and does not automatically extend to the complete implementation-level evaluation pipeline. In numerical optimization, candidate parameter vectors are projected onto the admissible bounds, the forward model is approximated by fixed-step RK4, nonnegative-state safeguards are applied, as described in the Supplementary Materials, and failed or nonfinite simulations are assigned the penalty value. Parameter projection and clipping-type safeguards may introduce non-smoothness, while implementation-level branching and hard failure penalization may introduce discontinuities. Accordingly, no continuity-based minimum-existence claim is made for the failure-penalized numerical mapping itself; the penalty is treated solely as an optimization safeguard.

5. B-ARLM Algorithm and Evaluation Protocol

B-ARLM combines generation-level success-rate control with coordinate-wise learning from successful candidate displacements within a bounded, elitist, best-centred mutation mechanism. Each active parameter direction has an individual normalized mutation amplitude, whereas no full covariance matrix is estimated.

5.1. Parameter-Wise Mutation Rule

Let

$$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_6) \in \Theta \subset \mathbb{R}^6$$

denote the six-component parameter vector, and let A denote the set of active parameter indices. For Andrews–P and Andrews–G, $d = |A| = 6$, whereas for Monod, $d = 5$, because k_i is inactive.

For generation $g \geq 2$, the successful candidate set and success rate are

$$S_g = \left\{ i : J\left(\theta_i^{(g)}\right) < J_{best}^{(g-1)} \right\}, r_g = \frac{|S_g|}{N} \tag{8}$$

where $i = 1, \ldots, N$ indexes the population candidates, N is the population size, $\theta_i^{(g)}$ is the i-th candidate parameter vector at generation g, and $J_{\text{best}}^{(g-1)}$ is the global best objective value available before generation g. Thus, only candidates that strictly improve upon the previous global best are classified as successful.

For every active direction $j \in A$, let $\sigma_j^{(g)}$ denote the normalized mutation amplitude for active parameter coordinate j at generation g. The intermediate globally adapted amplitude is denoted by $\overset{\sim}{\sigma}_j^{(g+1)}$, and the generation-level update is

$$\overset{\sim}{\sigma}_j^{(g+1)} = \begin{cases} \sigma_j^{(g)} / \alpha, & r_g > p^*, \\ \alpha \sigma_j^{(g)}, & r_g \leq p^*, \end{cases} \tag{9}$$

where p^* is the target success rate and $0 < \alpha < 1$ the multiplicative adaptation factor.

When $\mathcal{S}_g \neq \emptyset$, the coordinate-wise successful-displacement statistic is

$$m_{j,g} = 2\text{median}_{i \in \mathcal{S}_g}((|\ \theta_{ij}^{(g)} - \theta_{best,j}^{(g-1)}\ |)/(u_j - l_j)). \tag{10}$$

Here, $\theta_{ij}^{(g)}$ is the j-th component of candidate vector $\theta_i^{(g)}$, $\theta_{\text{best},j}^{(g-1)}$ is the j-th component of the global best parameter vector available before generation g, and l_g and u_g are the lower and upper bounds of parameter coordinate j respectively.

The coefficient 2 in Equation (10) is motivated by the centred uniform mutation model, whose median absolute unprojected displacement equals one half of the mutation amplitude. Because successful candidates form a selected subset and boundary projection can alter realized displacements, $m_{j,g}$ is treated as a robust scale-matching statistic.

The final update is

$$\sigma_j^{(g+1)} = \begin{cases} \Pi_{[\sigma_{\min}|\sigma_{\max}]}\left[(1-\eta)\tilde{\sigma}_j^{(g+1)} + \eta m_{j,g}\right], & \mathcal{S}_g \neq \emptyset, \\ \Pi_{[\sigma_{\min}|\sigma_{\max}]}\left[\tilde{\sigma}_j^{(g+1)}\right], & \mathcal{S}_g = \emptyset, \end{cases} \tag{11}$$

Here, $0 < \eta \leq 1$ is the coordinate-wise learning rate, $\sigma_{\min}$ and $\sigma_{\max}$ are the permanent lower and upper mutation-amplitude bounds, respectively, and

$$\Pi_{[a,b]}(x) = \min\{b, \max\{a, x\}\}.$$

The best-so-far solution is retained as the elitist member of the next population. The remaining candidates are generated by

$$\theta_{ij}^{(g+1)} = \Pi_{[l_j,u_j]}\left[\theta_{\text{best},j}^{(g)} + U_{ij}^{(g)}\sigma_j^{(g+1)}(u_j - l_j)\right], \tag{12}$$

In Equation (12), $\theta_{ij}^{(g+1)}$ denotes the j-th component of candidate i in the next population, and $\theta_{\text{best},j}^{(g)}$ is the j-th component of the current global best parameter vector with

$$U_{ij}^{(g)} \sim \mathcal{U}[-1,1], i = 2, \ldots, N, \quad j \in \mathbf{A}.$$

Inactive coordinates remain fixed at their nominal values.

The experiments used

$$p^* = 0.20,\ \eta = 0.20,\ \alpha = 0.95,\ \sigma_{\min} = 0.001,\ \sigma_{\max} = 0.20,\ \text{and}\ \sigma_0 \in \{0.05, 0.10, 0.20\} \tag{13}$$

Here, σ_0 denotes only the initial normalized mutation amplitude; subsequent amplitudes evolve according to Equations (9)–(11) while remaining within the permanent bounds [$\sigma_{\min}$ $\sigma_{\max}$]. The reference values $p^* = 0.20$, $\alpha = 0.95$, and $\eta = 0.20$ were selected as moderate predefined settings: p^* defines the success threshold, $\alpha = 0.95$ provides gradual multiplicative expansion/contraction, and $\eta = 0.20$ gives a smoothed coordinate-wise update. The reference settings from Equation (13) were fixed for the main comparative experiments. Their influence was subsequently examined by the one-factor-at-a-time sensitivity analysis described in Section 5.4, rather than being treated as evidence of an intrinsically optimal parameterization.

5.2. Algorithmic Workflow

Algorithm B-ARLM for bounded parameter calibration

1. Initialize active mutation amplitudes at σ_0; set inactive amplitudes to zero.

2. Construct a nominal-centred bounded population, retaining the nominal vector as its first member.
3. Let G denote the prescribed number of generations. For $g = 1,\ldots,G$:
   a. Evaluate all candidates and assign 10^{12} to failed or non-finite simulations;
   b. Update the global best-so-far solution;
   c. From $g = 2$ onward, compute $\mathcal{S}_g$ and r_g using Equation (8), and update the amplitudes using Equations (8)–(10);
   d. Record the best objective value and amplitude vector;
   e. When $g < G$ retain the best solution and generate the next population using Equation (11).
4. Return the best parameter vector, terminal objective value, convergence history, and amplitude histories.

Adaptation is omitted during the first generation because no previous finite best objective value is available.

5.3. Boundedness and Local Behaviour of the Adaptive Mutation Amplitudes

For every active parameter direction $j \in A$, the projection in Equation (11) guarantees

$$\sigma_{\min} \leq \sigma_j^{(g)} \leq \sigma_{\max} \tag{14}$$

for every generation g. Hence, the stored mutation amplitudes cannot collapse below $\sigma_{\min}$ or diverge above $\sigma_{\max}$. This boundedness result concerns the internal mutation scales; parameter-bound projection may still reduce the realized candidate displacement, particularly when the current best solution lies close to a parameter boundary.

To characterize the local behaviour of the adaptation rule, define $c_g = \alpha^{-1}$ when the generation-level success rate exceeds the target p^* and $c_g = \alpha$ otherwise.

Before projection, Equation (11) can then be written component-wise as

$$\sigma_{j,g+1}^{row} = (1-\eta)c_g\sigma_j^{(g)} + \eta m_{j,g}.$$

Here, $m_{j,g}$ denotes the coordinate-wise successful-displacement statistic from Equation (10). If the adaptation branch c_g and the displacement statistic $m_{j,g}$ are locally fixed, this expression defines an affine scalar mapping with slope $(1-\eta)cg$. The contraction branch therefore has slope $(1-\eta)\alpha$, whereas the expansion branch has slope $(1-\eta)/\alpha$. Both branches are locally contractive, provided

$$0 \leq (1-\eta)\alpha < 1 \text{ and } 0 \leq (1-\eta)/\alpha < 1$$

For $0 < \alpha < 1$, the first inequality is automatically satisfied for $0 < \eta \leq 1$, while the second requires $\eta > 1-\alpha$. With the reference values $\alpha = 0.95$ and $\eta = 0.20$, the corresponding slopes are 0.76 and approximately 0.842, respectively, and are therefore both below unity.

Projection $\Pi_{[\sigma_{\min},\sigma_{\max}]}$ onto $[\sigma_{\min}, \sigma_{\max}]$ is non-expansive and preserves boundedness. Under a fixed adaptation branch and a fixed successful-displacement statistic, satisfaction of the corresponding contraction condition therefore makes the projected affine scalar iteration contractive and implies convergence to a unique fixed point within the prescribed amplitude interval. However, the success branch cg can switch between generations, and $m_{j,g}$ is itself generated by the evolving successful candidate set. Consequently, the preceding affine contraction argument is local and conditional; it does not establish convergence of the complete stochastic adaptation process. Repeated changes of direction, temporary boundary occupancy, or oscillatory amplitude behaviour are possible in principle when suc-

cess signals and successful displacements vary across generations. The permanent bounds guarantee only that such behaviour remains confined to the prescribed amplitude interval.

Accordingly, no global stochastic-convergence theorem is claimed for B-ARLM. The empirical sensitivity and amplitude-occupancy diagnostics reported below are used to determine whether the reference configuration exhibits practically problematic dependence on the prescribed bounds or persistent boundary oscillation in the present calibration problem.

5.4. Hyperparameter-Sensitivity Protocol

The reference B-ARLM configuration used Equation (13). To assess whether the observed optimization behaviour depended strongly on these prescribed settings, a one-factor-at-a-time sensitivity analysis was performed around the reference configuration. Each factor was varied, while all remaining settings were held fixed. The examined alternatives were $p^* \in \{0.10, 0.30\}$, $\alpha \in \{0.90, 0.98\}$, $\eta \in \{0.10, 0.40\}$, $\sigma_{min} \in \{0.0005, 0.005\}$, and $\sigma_{max} \in \{0.10, 0.30\}$, giving 10 perturbed configurations in addition to the reference setting.

Each of the 11 configurations was evaluated in 30 reproducible Andrews–P calibration runs using the same population size, generation count, objective-evaluation budget, parameter bounds, and nominal-centred initialization protocol as the reference experiment. Run-specific random seeds were paired across configurations so that each perturbed setting could be compared with the corresponding reference run under matched stochastic initialization.

Terminal NMSE distributions were summarized by their median and interquartile range. Pairwise comparisons between each perturbed configuration and the reference configuration used the paired Wilcoxon signed-rank test. To control family-wise error across the 10 comparisons, the resulting p-values were adjusted using the Holm procedure. Paired rank-biserial correlation was reported as an effect-size measure. In addition to terminal NMSE, mutation-amplitude occupancy near the prescribed lower and upper bounds and generation-to-generation direction changes were examined to determine whether the permanent amplitude limits produced persistent boundary trapping or oscillatory behaviour.

This analysis is intended as a local robustness assessment over explicitly defined finite ranges, not as a global hyperparameter-optimization study. The reference configuration was retained unchanged for all primary comparisons regardless of the sensitivity-analysis outcomes, thereby avoiding post-hoc selection of the best-performing tested setting.

5.5. Comparator Implementations and Evaluation Protocol

Four comparators were used. Fixed-EA served as a direct ablation of B-ARLM and retained the same bounds, initialization, projection, failure handling, and elitist best-so-far tracking but used fixed parameter-wise mutation amplitudes. The study-specific sCA-ES was retained as a structured adaptive reference. Canonical CMA-ES [28,33] was added as an established full-covariance Evolution Strategy, and jDE [35] was included as a representative self-adaptive Differential Evolution method. All algorithms operated on the same Andrews–P calibration problem and used the same forward model, objective function, parameter bounds, and numerical-integration scheme.

Each method was evaluated in 30 independent runs under a common budget of 2250 objective-function evaluations per run. B-ARLM and Fixed-EA used paired seeds and matched initial populations because Fixed-EA constitutes a direct algorithmic ablation. The remaining methods retained their native stochastic sampling and adaptation mechanisms. Full implementation settings are provided in the Supplementary Materials and reproducibility package.

Terminal NMSE was summarized by the best, worst, mean, standard deviation, median, and interquartile range. B-ARLM and Fixed-EA were compared using the paired

Wilcoxon signed-rank test, whereas comparisons with sCA-ES, CMA-ES, and jDE used the Wilcoxon rank-sum test. The four pairwise *p*-values were adjusted by the Holm procedure. Effect sizes were reported as paired rank-biserial correlation for the Fixed-EA comparison and Cliff's delta for the remaining comparisons. Effect sizes were oriented so that positive values indicate lower terminal NMSE in favor of B-ARLM. The benchmark is interpreted as a problem-specific comparison under a matched evaluation budget rather than as evidence of universal optimizer superiority.

6. Identifiability and Uncertainty Diagnostics

Calibration results were interpreted using complementary diagnostics of local parameter support, observation influence, and conditional trajectory uncertainty.

6.1. Sensitivity and Fisher-Information Analysis

Let $\mathbf{S_{log}} = [S_{\log i,j}]$ denote the log-parameter sensitivity matrix of the six interval-average model outputs, where i = 1…6 indexes the observation intervals and j = 1…6 indexes the calibrated parameters:

$$S_{\mathbf{log},ij} = \frac{\partial \overline{\boldsymbol{q}}_i(\boldsymbol{\theta})}{\partial \mathbf{log}\boldsymbol{\theta}_j} = \boldsymbol{\theta}_j \frac{\partial \overline{\boldsymbol{q}}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j}, \quad \boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \ldots, \mathbf{6}. \tag{15}$$

Central multiplicative differences were used:

$$S_{\mathbf{log},ij} \approx \frac{\overline{\boldsymbol{q}}_i\left(\boldsymbol{\theta}^{(j,+)}\right) - \overline{\boldsymbol{q}}_i\left(\boldsymbol{\theta}^{(j,-)}\right)}{2\boldsymbol{\delta}}, \ \delta = \log(1 + 10^{-4}) \tag{16}$$

Here, δ is the multiplicative finite-difference perturbation, and $\theta^{(j,+)}$ and $\theta^{(j,-)}$ equal θ except for their *j*-th components,

$$\theta_j{}^{(j,+)} = \theta_j \mathrm{e}^{\delta}, \ \ \theta_j{}^{(j,-)} = \theta_j \mathrm{e}^{-\delta}.$$

The whitened sensitivity and scaled Fisher-information matrices were

$$\mathbf{S}_w = \frac{\mathbf{S_{log}}}{\boldsymbol{\sigma}_{\mathbf{obs}}}, \mathbf{F} = \mathbf{S}_w^{\mathrm{T}} \mathbf{S}_w, \quad \boldsymbol{\sigma}_{\mathbf{obs}} = \mathbf{0.05} \underset{i}{\mathbf{max}} | \ \overline{\boldsymbol{q}}_i(\boldsymbol{\theta}_{\mathbf{ref}}) \ | \tag{17}$$

Here, $\boldsymbol{\theta}_{\mathbf{ref}}$ denotes the selected reference Andrews–P parameter vector described in Section 7.2, and σ_{obs} is the prescribed sensitivity-whitening scale.

Effective rank was evaluated from the singular values of $\mathbf{S}_w$ and normalized by the largest singular value at thresholds 10^{-6}, 10^{-8}, and 10^{-10}. Because σ_{obs} is not estimated from replicate-level variance, the absolute FIM scaling is interpreted only as a local, reference-dependent diagnostic. Accordingly, the FIM results are interpreted as local practical-identifiability diagnostics and not as evidence of global parameter uniqueness.

6.2. Profile-NMSE Analysis

For each parameter θ_j, that parameter was fixed at α, while the remaining parameters were re-optimized:

$$\boldsymbol{J}_{(\mathbf{prof},j)}(\boldsymbol{a}) = \mathbf{min}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}: \boldsymbol{\theta}_j = \boldsymbol{a}} \ \ \boldsymbol{J}(\boldsymbol{\theta}), \boldsymbol{a} \in \left[\boldsymbol{l}_j, \boldsymbol{u}_j\right]. \tag{18}$$

Thus, $\boldsymbol{J}_{\mathbf{prof},j}(\boldsymbol{a})$ denotes the minimum NMSE attainable when parameter component $\boldsymbol{\theta}_j$ is fixed at α, while all remaining parameter components are re-optimized.

Each parameter grid comprised 17 logarithmically spaced values augmented with the reference value. At each grid point, the remaining five parameters were re-optimized in

three reference-centred B-ARLM runs using a population size of 20, 50 generations, and $\sigma_0 = 0.10$. For each profiled parameter, the same three deterministic seeds were reused across all grid values. Failed or non-finite values received a penalty of 10^{12}. The reported profile retained the lowest NMSE obtained from the three runs, while run-level summaries were preserved in the reproducibility outputs.

6.3. Broad-Start and Two-Dimensional Profile Analysis

To complement the reference-centred local diagnostics, global search behaviour was examined using 30 independent B-ARLM runs initialized from Latin-hypercube populations spanning the full admissible parameter box. The same Andrews–P objective, parameter bounds, population size, generation count, and objective-evaluation budget as in the primary calibration were retained. The resulting terminal solutions were compared in terms of NMSE and normalized parameter-space separation to determine whether similar objective values could be obtained from widely separated regions of the admissible domain.

Nonlinear parameter compensation was further examined by a two-dimensional profile analysis for γ_1 and γ_{H_2}, which showed the strongest association in the broad-start solutions. At each grid point, these two parameters were fixed, while the remaining four parameters were re-optimized in two B-ARLM runs using a population size of 20 and 40 generations. The lowest NMSE from the two runs was retained. This analysis was used to identify ridge-like low-error regions and parameter compensation beyond that visible in the one-dimensional profiles.

6.4. Repeated Leave-One-Interval-Out Observation-Influence Analysis

Each of the six interval-average observations was omitted in turn, and the remaining five observations were recalibrated in 30 B-ARLM runs using a population size of 30, 75 generations, and recorded deterministic run-specific seeds.

Let $\hat{\theta}_{-i}^{(r)}$ denote the recalibrated parameter vector obtained in run r after omission of observation interval i, where $r = 1,\ldots, R$, and $R = 30$ is the number of recalibration runs performed for each omitted interval. The omitted-interval absolute error was

$$e_{i,r}^{\mathbf{LOO}} = \mid \bar{q}_i\left(\hat{\theta}_{(-i)}^{(r)}\right) - \bar{q}_i^{\mathbf{obs}} \mid, \quad r = 1,\ldots,R. \tag{19}$$

Equation (19) defines the absolute omitted-interval prediction error for interval i in recalibration run r. The quantity $\tilde{e}_i^{\mathrm{LOO}}$ denotes the median of these absolute errors over the R recalibration runs and is defined by

$$\tilde{e}_i^{\mathrm{LOO}} = \mathrm{median}_{r=1,\ldots,R} e_{i,r}^{\mathrm{LOO}}. \tag{20}$$

The median and interquartile range characterize interval influence and recalibration variability. Because each omission leaves five observations for six calibrated parameters, the analysis is interpreted as an influence diagnostic rather than conventional predictive cross-validation.

6.5. Conditional Trajectory-Uncertainty and Error-Model Sensitivity

For bootstrap dataset b, synthetic interval-average observations were generated as

$$\bar{q}_i^{(b)} = \mathbf{max}\left\{0,\, \bar{q}_i(\theta_{\mathbf{ref}}) + \varepsilon_i^{(b)}\right\}, \varepsilon_i^{(b)} \sim \mathcal{N}\left(0, \sigma_\rho^2\right), \tag{21}$$

where $\varepsilon_i^{(b)}$ is the additive Gaussian perturbation applied to interval i in bootstrap dataset b, and $b = 1, \ldots, B$, and

$$\sigma_\rho = \rho \max_i | \bar{q}_i(\theta_{\mathrm{ref}}) |, \rho \in \{0.05, 0.10\}. \quad (22)$$

The quantity ϱ is the prescribed relative perturbation level, and σ_p is the corresponding Gaussian perturbation scale. The 5% and 10% perturbation levels were prescribed uncertainty scenarios and were not estimated from replicate-level measurement variance.

For each perturbation level, 200 datasets were recalibrated once using a reference-centred B-ARLM run with 50 generations, a population size of 20, and $\sigma_0 = 0.10$. The optimizer seed was fixed, whereas separate deterministic seed streams generated the measurement perturbations.

Let $\hat{\theta}^{(b)}$ denote the parameter vector obtained by recalibrating bootstrap dataset b, and let $B = 200$ denote the number of bootstrap datasets generated for each perturbation level. The resulting instantaneous-trajectory ensemble was

$$\mathcal{Q}_{\mathrm{boot}} = \left\{ q_{\mathrm{H_2}}\left(t; \hat{\theta}^{(b)}\right) \right\}_{b=1}^{B} \quad (23)$$

Pointwise empirical 2.5th and 97.5th percentiles define conditional pointwise 95% fitted-trajectory intervals. Because negative synthetic observations were clipped at zero and the optimizer seed was fixed, the resulting intervals primarily represent propagation of the imposed Gaussian observation perturbations and associated parameter compensation. They are conditional on the Andrews–P formulation, the selected reference calibration, the prescribed perturbation levels, the clipping procedure, and the fixed optimizer random-number stream. They should not be interpreted as replicate-derived confidence intervals or as predictive intervals for future noisy observations.

Because replicate-derived measurement-error statistics were unavailable, the dependence of the fitted-trajectory uncertainty on the assumed observation-error structure was additionally examined at the common 10% perturbation level. Three prescribed scenarios were considered: independent additive Gaussian perturbations, multiplicative log-normal perturbations, and additive Gaussian perturbations with first-order temporal correlation (AR(1), correlation coefficient fixed at 0.5). For each scenario, 200 synthetic datasets were recalibrated using the same reference-centred B-ARLM protocol. These scenarios are interpreted as error-model sensitivity analyses rather than as empirically identified stochastic descriptions of the experimental measurements.

7. Numerical Experiments and Results

All experiments used the same six interval-average observations and the normalized least-squares objective defined in Equation (6). Unless stated otherwise, the primary repeated-run experiments used a population size of 30, 30 independent runs, and 75 generations, corresponding to 2250 objective-function evaluations per run.

7.1. Comparative Kinetic-Model Calibration

Table 3 compares the in-sample calibration performance of the classical Monod formulation [45], the Andrews–G substrate-inhibited formulation [46], and the Andrews–P product-inhibited formulation defined in Section 3. Each formulation was calibrated in 30 independent B-ARLM runs using the same population size, 75 generations, interval observation operator, numerical-integration step, and nominal-centred initialization protocol. The same bounds were used for all active parameters; for the Monod formulation, k_i was fixed at its nominal value because it does not enter the kinetic law.

Table 3. Comparative in-sample calibration performance of the tested kinetic formulations over 30 independent B-ARLM runs.

Model	k	Best NMSE	Worst NMSE	Mean NMSE	SD NMSE
Monod	5	6.48×10^{-3}	2.03×10^{-2}	9.71×10^{-3}	5.96×10^{-3}
Andrews–G	6	6.48×10^{-3}	2.03×10^{-2}	8.79×10^{-3}	5.25×10^{-3}
Andrews–P	6	1.29×10^{-3}	6.47×10^{-3}	2.08×10^{-3}	1.75×10^{-3}

Note: k denotes the number of active calibrated parameters, and SD denotes the standard deviation of the terminal NMSE values. For the Monod model, the inhibition-scale parameter k_i does not enter the kinetic law and was therefore fixed at its nominal value and excluded from mutation and adaptation, giving an effective calibration dimension of five. The full-precision parameter vectors corresponding to the lowest-NMSE solutions are provided in Table 3 machine-readable outputs included in the reproducibility package.

Monod and Andrews–G reached essentially the same lowest recorded best-run NMSE in the present comparison. Andrews–P achieved substantially lower best, mean, worst, and standard-deviation NMSE values than either alternative formulation. These differences indicate a markedly better in-sample fit for Andrews–P under the examined dataset and calibration protocol, but they do not establish biological superiority of the corresponding inhibition mechanism.

Because Andrews–G and Andrews–P have the same effective calibration dimension, their performance difference cannot be attributed to parameter count alone. Within the examined dataset and calibration protocol, Andrews–P produced the lowest in-sample calibration errors. This result is interpreted as comparative calibration performance only and does not identify the Andrews–P inhibition term as a uniquely correct biological mechanism.

7.2. Reference Andrews–P Calibration

The reference Andrews–P solution was selected as the lowest-NMSE valid result among 30 independent B-ARLM runs using an initial mutation amplitude of $\sigma_0 = 0.10$, a population size of 30, and 75 generations.

The selected solution yielded an NMSE of 1.2876×10^{-3}, with the rounded reference parameter vector

$$\boldsymbol{\theta}_{\mathbf{ref}} \approx (18.2, 0.0106, 1.56, 5.31, 0.254, 2.32).$$

The components correspond, respectively, to μ_{max}, K_s, γ_1, γ_2, γ_{H_2}, and k_i. This vector represents one low-error calibration solution rather than a set of uniquely identified kinetic-parameter estimates.

Values in the main text are rounded to approximately three significant digits because the identifiability analysis does not support greater inferential precision. Full-precision computational values are retained in the accompanying CSV and MATLAB 2019a files for reproducibility.

Figure 1 shows the best-so-far convergence histories of the 30 B-ARLM runs used in the reference-calibration experiment. The median curve summarizes the typical convergence behaviour across the independent runs.

7.3. Initial-Amplitude and Hyperparameter Sensitivity

B-ARLM and Fixed-EA were compared using the three initial mutation amplitudes specified in Equation (13) (Figure 2). For B-ARLM, σ_0 specifies only the initial parameter-wise mutation amplitude, whereas all three configurations use the same permanent upper bound, $\sigma_{max} = 0.20$. In contrast, Fixed-EA retains the selected mutation amplitude throughout each run. Each configuration was evaluated in 30 independent runs using a population size of 30, 75 generations, and the same bounds for all active parameters. For each value

of σ_0, the corresponding B-ARLM and Fixed-EA runs used the same run-specific random seeds and identical initial populations.

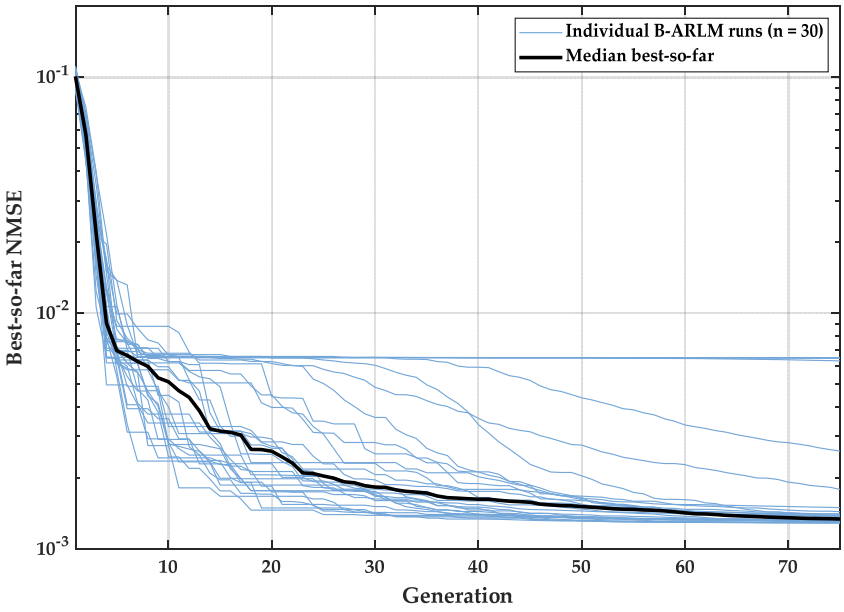


Figure 1. Best-so-far NMSE convergence trajectories of B-ARLM for the Andrews–P model over 30 independent runs. Thin blue curves represent individual optimization runs, while the thick black curve denotes the pointwise median best-so-far NMSE. Each run used a population of 30 individuals for 75 generations. The NMSE axis is logarithmic.

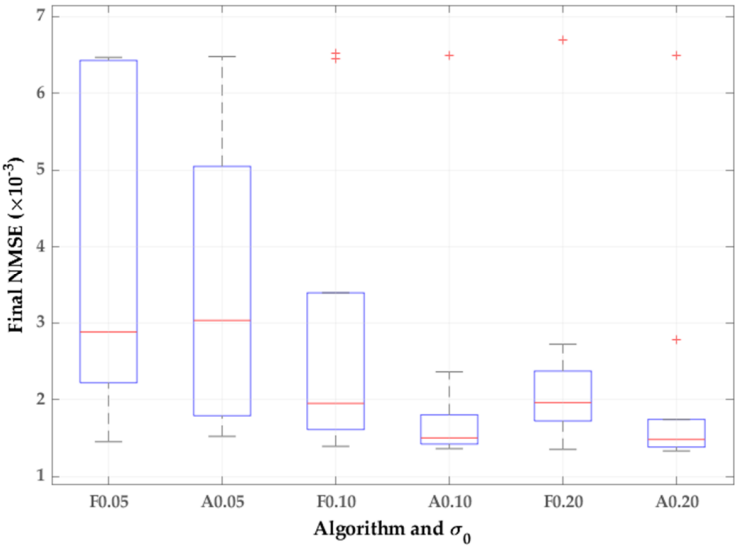


Figure 2. Terminal NMSE distributions over 30 runs for Fixed-EA (F) and B-ARLM (A) at the three tested mutation-amplitude settings σ_0 = 0.05, 0.10, and 0.20. For B-ARLM, σ_0 denotes the initial mutation amplitude, whereas Fixed-EA retains the selected amplitude throughout the run. Corresponding B-ARLM and Fixed-EA runs used matched random seeds and identical initial populations.

Across the three tested initial amplitudes, B-ARLM showed less variation in central terminal performance than Fixed-EA, consistent with subsequent adaptation of the mutation amplitudes within the prescribed permanent bounds.

The broader one-factor-at-a-time sensitivity analysis comprised the reference configuration and 10 single-factor perturbations, with 30 paired runs per configuration. Across the examined finite ranges, median terminal NMSE values varied from approximately 1.30×10^{-3} to 1.47×10^{-3}, compared with 1.36×10^{-3} for the reference setting. After

Holm correction, only increasing the permanent lower mutation bound from $\sigma_{min} = 0.001$ to 0.005 produced a statistically significant deterioration in terminal NMSE ($p_{Holm} \approx 0.047$).

The elevated lower bound also produced substantially greater lower-bound occupancy of the mutation amplitudes, consistent with insufficient late-stage contraction. None of the remaining tested hyperparameter perturbations remained statistically significant after multiplicity correction over the examined ranges. The reference setting is therefore retained as the predefined reference configuration, with limited sensitivity over the examined one-factor ranges, rather than interpreted as a globally optimized hyperparameter choice. Detailed sensitivity statistics and amplitude diagnostics are provided in the Supplementary Materials.

7.4. Practical Identifiability, Global Parameter Compensation, and Observation Influence

Practical parameter support was first examined through the profile-NMSE analysis shown in Figure 3. The profiles characterize the extent to which each parameter can vary while the remaining parameters are re-optimized.

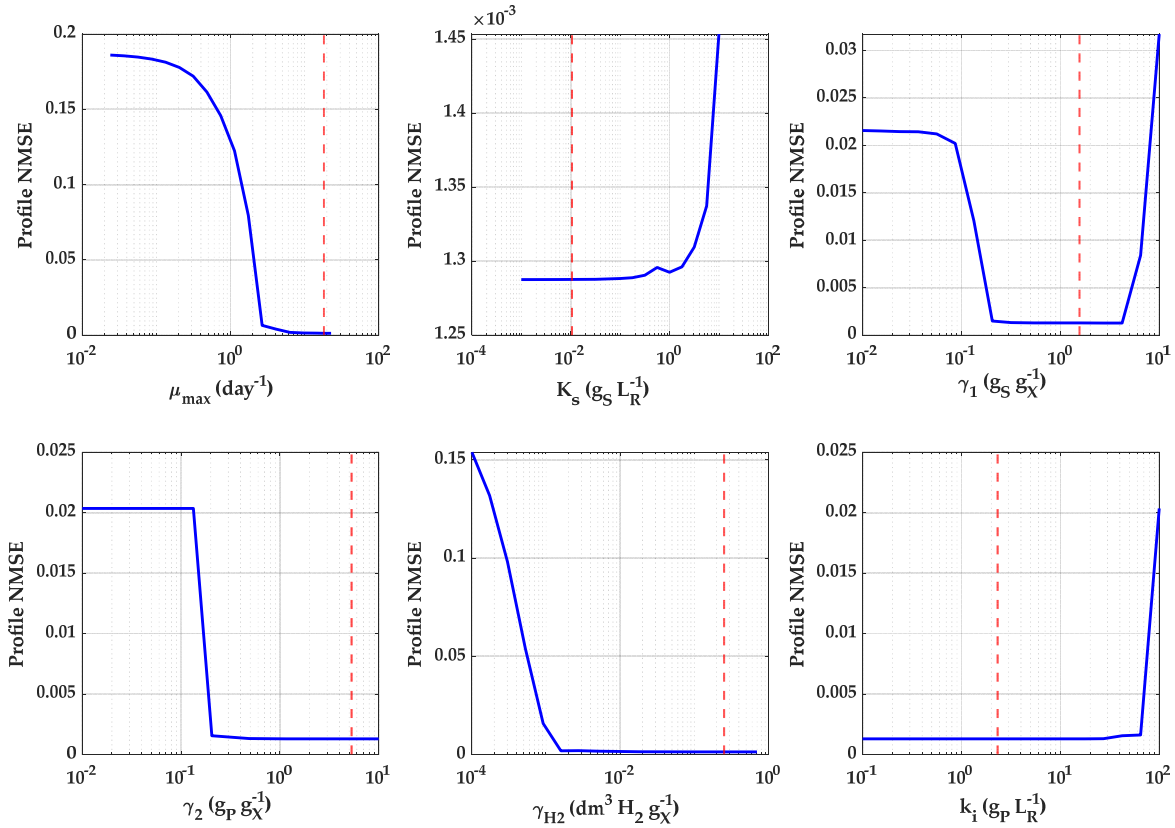


Figure 3. Profile-NMSE curves for the six calibrated Andrews–P parameters. Dashed vertical lines indicate the reference parameter values. For each parameter, 17 logarithmically spaced base-grid values were augmented with the reference value. At each grid point, the remaining parameters were re-optimized in three independent B-ARLM runs using 50 generations and a population size of 20, and the lowest obtained NMSE was retained as the profile value.

The profile-NMSE curves show markedly different levels of practical parameter support, including broad and asymmetric low-error regions consistent with parameter compensation. To test whether these features were specific to reference-centred initialization, 30 additional B-ARLM runs were initialized from Latin-hypercube populations spanning the full admissible domain. The best broad-start NMSE was 1.286×10^{-3}, essentially identical to the reference value of 1.288×10^{-3}; 16 of 30 runs remained within twice the

best value. Despite these similar errors, the near-optimal vectors were widely separated, with a median normalized pairwise distance of 0.842 and a maximum of 1.534.

The strongest broad-start compensation involved γ_1 and γ_{H_2}. Their two-dimensional profile produced a pronounced diagonal low-error valley; a distant combination near $\gamma_1 = 4.22$ and $\gamma_{H_2} = 0.70$ retained an NMSE of approximately 1.30×10^{-3}, only about 0.8% above the reference optimum (see Figure 4). Thus, the reference vector is not an isolated low-error parameter solution.

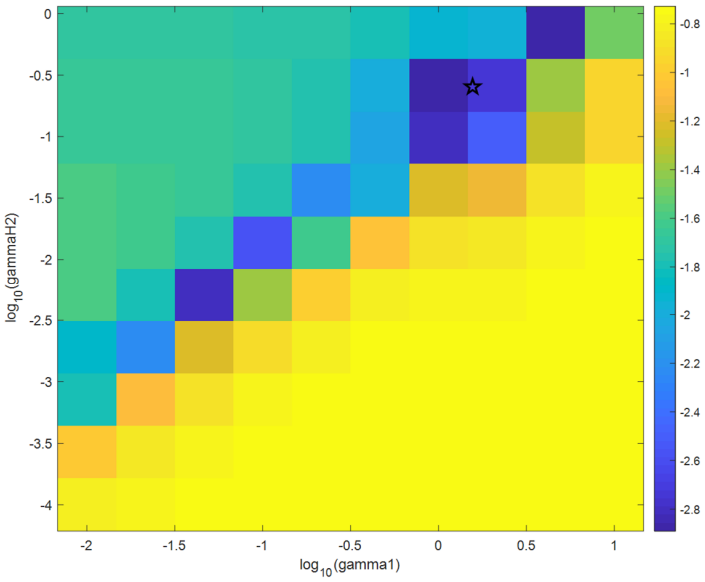
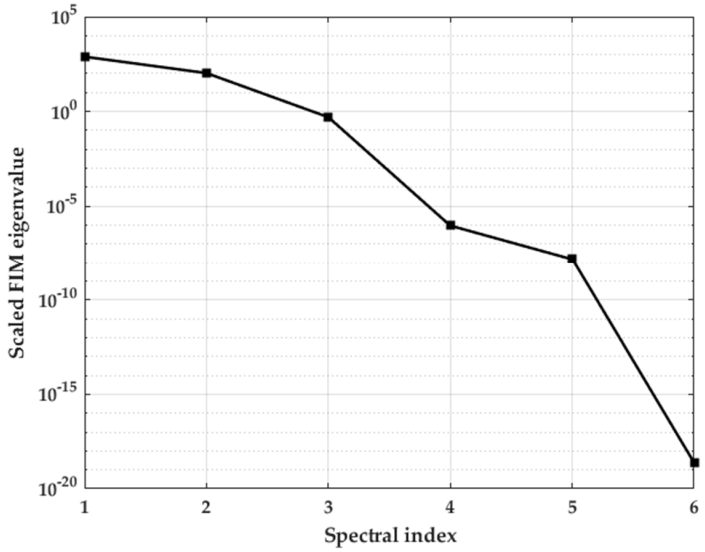


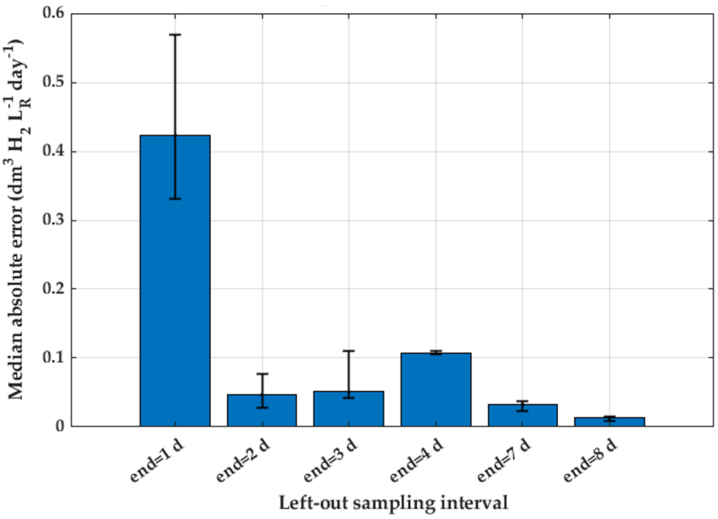
Figure 4. Two-dimensional profile-NMSE map for γ_1 and γ_{H_2}. At each grid point, the two displayed parameters were fixed, and the remaining four Andrews–P parameters were re-optimized in two B-ARLM runs using a population size of 20 and 40 generations. The star (★) denotes the reference parameter combination associated with the best-fit solution. The color scale represents $\log_{10}$ of the profile NMSE; the diagonal low-error region indicates strong parameter compensation.

Local parameter support and observation influence were further examined using the scaled Fisher-information spectrum and repeated leave-one-out recalibration results shown in Figure 5.



(**a**)

Figure 5. *Cont.*



(**b**)

Figure 5. Practical-identifiability and interval-influence diagnostics. (**a**) local-scaled Fisher-information eigenvalue spectrum based on whitened log-parameter sensitivities. (**b**) median absolute leave-one-interval-out prediction errors over 30 independent recalibrations per omitted interval, with error bars denoting the interquartile range.

For the displayed reference vector, the scaled FIM eigenvalues span approximately 21.5 orders of magnitude, with a condition number of approximately 3.23×10^{21}. The numerical rank was six, whereas the effective rank was five at all three examined relative singular-value thresholds; the smallest-scaled eigenvalue was approximately 2.40×10^{-19}. These results indicate severe local ill-conditioning and weak support for at least one parameter combination.

The leave-one-interval-out analysis comprised 180 recalibrations. The 0–1 day interval was the most influential, followed by 3–4 days, whereas the late 4–7 and 7–8 day intervals produced the smallest median omitted-interval errors. Taken together, the profile, broad-start, two-dimensional, FIM, and leave-one-interval-out analyses show that accurate trajectory reconstruction is compatible with widely separated parameter vectors and uneven observation influence, indicating substantial practical non-identifiability without establishing structural non-identifiability.

7.5. Conditional Trajectory Uncertainty and Error-Model Sensitivity

Using the protocol defined in Section 6.5, conditional fitted-trajectory uncertainty was evaluated at prescribed perturbation levels of 5% and 10%, with 200 synthetic datasets per level.

At both perturbation levels, the median trajectory retained the dominant early peak and subsequent decline. The conditional intervals were widest in the early dynamic region and narrowed during the late near-zero phase, indicating greater uncertainty in peak timing and magnitude than in the qualitative trajectory shape.

For the 5% perturbation level, the mean, median, and maximum pointwise 95% interval widths were 0.131, 0.0873, and 1.69 $\mathrm{dm}^3_{\mathrm{H}_2} \cdot \mathrm{L}_{\mathrm{R}}^{-1} \cdot \mathrm{day}^{-1}$, respectively (see Figure 6a). At the 10% perturbation level, the corresponding widths increased to 0.274, 0.154, and 2.74 $\mathrm{dm}^3_{\mathrm{H}_2} \cdot \mathrm{L}_{\mathrm{R}}^{-1} \cdot \mathrm{day}^{-1}$ (see Figure 6b). The increases in all three width summaries at the 10% perturbation level are consistent with greater fitted-trajectory uncertainty under the larger imposed observation-perturbation scale.

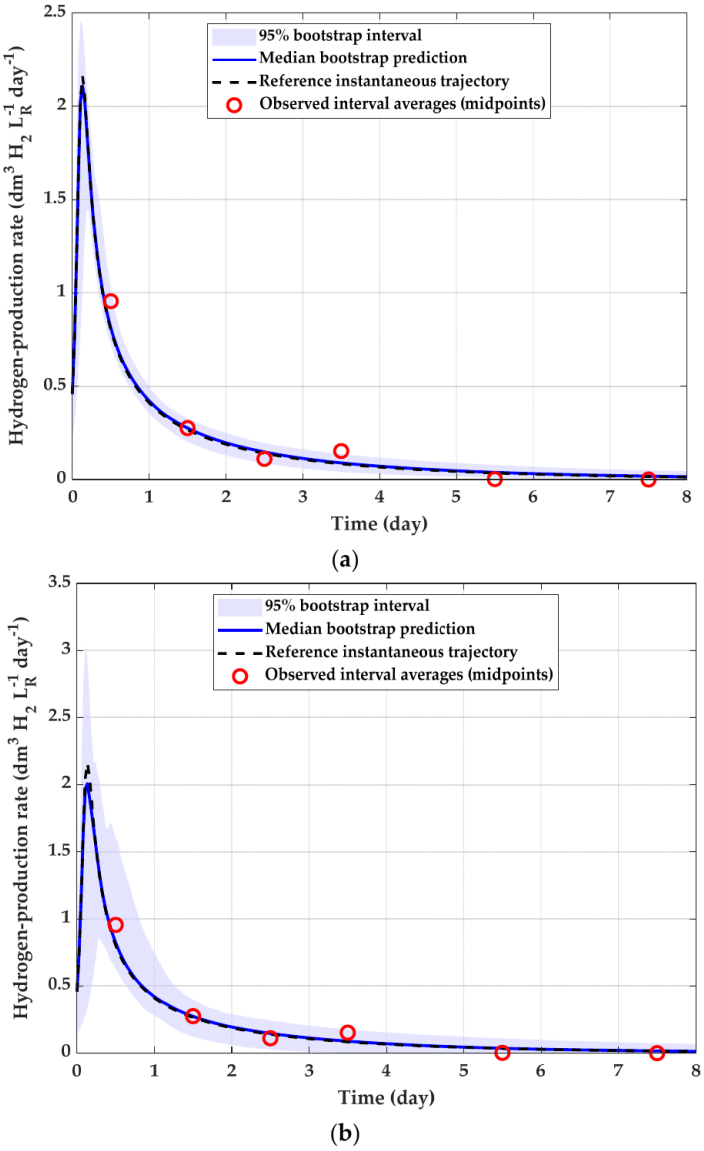


Figure 6. Conditional parametric-bootstrap fitted-trajectory uncertainty. Shaded regions denote pointwise 95% empirical trajectory intervals; solid and dashed curves represent the median recalibrated and reference Andrews–P instantaneous trajectories, respectively; circles denote the observed interval-average rates plotted at the interval midpoints. (**a**) 5% prescribed additive Gaussian perturbations with clipping at zero. (**b**) 10% prescribed additive Gaussian perturbations with clipping at zero.

The perturbation levels were not estimated from replicate-level measurement variance. Accordingly, the intervals characterize conditional fitted-trajectory uncertainty rather than replicate-derived confidence intervals or predictive intervals for future noisy observations. At the common 10% perturbation level, the mean pointwise 95% fitted-trajectory band width depended strongly on the assumed observation-error model: 0.280 for independent additive Gaussian noise, 0.101 for multiplicative log-normal noise, and 0.311 for additive Gaussian noise with first-order temporal correlation (AR(1), correlation coefficient = 0.5). Because replicate-level error statistics were unavailable, these results are interpreted as sensitivity to the prescribed error structure rather than evidence supporting any particular stochastic model.

7.6. Comparative Optimizer Performance

The matched-budget comparison of B-ARLM, Fixed-EA, sCA-ES, canonical CMA-ES, and jDE is summarized in Table 4 and Figure 7.

Table 4. Terminal optimizer performance over 30 runs per method under a common budget of 2250 objective-function evaluations per run. IQR denotes the interquartile range. Formal pairwise comparisons of B-ARLM with the four reference methods were Holm-adjusted as described in Section 5.5.

Optimizer	Median NMSE	IQR	Mean NMSE	SD	Worst NMSE
B-ARLM	1.339×10^{-3}	1.21×10^{-4}	2.076×10^{-3}	1.748×10^{-3}	6.473×10^{-3}
Fixed-EA	1.597×10^{-3}	4.14×10^{-4}	2.299×10^{-3}	1.687×10^{-3}	6.531×10^{-3}
sCA-ES	2.806×10^{-3}	4.90×10^{-3}	4.929×10^{-3}	4.513×10^{-3}	1.972×10^{-2}
CMA-ES	1.286×10^{-3}	8.41×10^{-3}	5.781×10^{-3}	6.815×10^{-3}	2.034×10^{-2}
jDE	1.401×10^{-3}	1.55×10^{-4}	1.439×10^{-3}	1.28×10^{-4}	1.806×10^{-3}

B-ARLM achieved a median terminal NMSE of 1.339×10^{-3} with an IQR of 1.21×10^{-4}. Fixed-EA produced a higher median of 1.597×10^{-3} and a wider IQR of 4.14×10^{-4}. The paired comparison confirmed a significant B-ARLM advantage over Fixed-EA after Holm correction ($p_{Holm} = 0.0011$; paired rank-biserial correlation = 0.746). B-ARLM also significantly outperformed sCA-ES ($p_{Holm} \approx 1.9 \times 10^{-5}$; Cliff's $\delta = 0.689$).

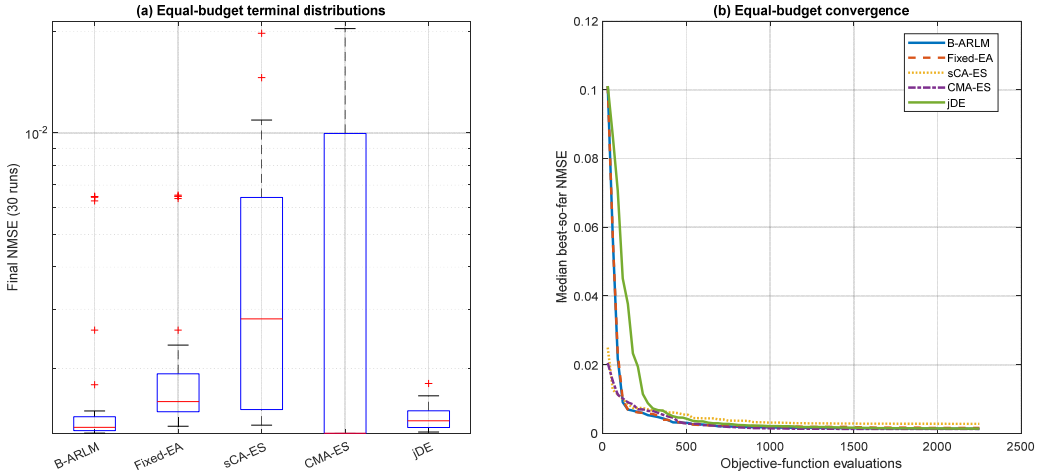


Figure 7. Comparative optimizer performance over 30 runs per method under a common budget of 2250 objective-function evaluations per run: (**a**) terminal NMSE distributions; (**b**) median best-so-far NMSE as a function of objective-function evaluations.

In contrast, terminal performance was not statistically distinguishable from canonical CMA-ES or jDE after Holm correction ($p_{Holm} \approx 0.316$ for both comparisons). CMA-ES achieved the lowest median NMSE but exhibited a pronounced upper tail, whereas jDE achieved the lowest mean, standard deviation, and worst-run NMSE among the five methods. B-ARLM therefore showed statistically supported improvement over its fixed-mutation ablation and the study-specific sCA-ES reference, while no statistically significant difference in terminal NMSE was detected between B-ARLM and CMA-ES or jDE under the present matched evaluation budget.

7.7. Computational Complexity and Runtime

With the production RK4 step, one objective evaluation required 9600 integration steps or 38,400 right-hand-side evaluations, corresponding to approximately 8.64×10^7 right-hand-side evaluations per 2250-evaluation run. B-ARLM requires $O(Nd)$ working

memory. Across five repeated measurements, the reference implementation required a median of 136.9 s per run (IQR 1.83 s). At the optimization layer, with G generations, population size N, d active parameters, and C_{obj} denoting the cost of one forward-model objective evaluation, the computational work scales as $O(GN(C_{obj} + d))$; for the present problem, this cost is dominated by RK4 forward integration. These timings characterize the present MATLAB implementation and are not interpreted as a runtime advantage over the comparator methods; additional details are provided in the Supplementary Material.

8. Discussion

The present study considered sparse-data bioprocess calibration as a combined optimization and parameter-recovery problem. Among the three examined kinetic formulations, Andrews–P produced the lowest in-sample calibration error for the six interval-average observations. This result supports its use as the best-fitting formulation within the present dataset and calibration protocol, but it does not establish that the quadratic product-inhibition term is a uniquely correct biological mechanism or that Andrews–P has superior predictive performance outside the observed experiment.

The extended optimizer benchmark provides a more balanced assessment of B-ARLM than comparison with the study-specific references alone. Under the common budget of 2250 objective-function evaluations per run, B-ARLM significantly outperformed the Fixed-EA ablation and sCA-ES reference after Holm correction, whereas its terminal performance was not statistically distinguishable from canonical CMA-ES or jDE. CMA-ES achieved the lowest median terminal NMSE but exhibited a broad upper tail, while jDE produced the lowest mean, standard deviation, and worst-run NMSE. These results support B-ARLM as a competitive bounded calibration strategy for the present problem rather than as a universally superior optimizer.

This interpretation is consistent with the methodological position of B-ARLM relative to established evolutionary strategies. Success-based step-size adaptation, coordinate-dependent search scales, and covariance adaptation represent progressively richer mechanisms for modifying the search distribution [28–34], while jDE adapts Differential Evolution control parameters through a different population-difference mechanism [35]. B-ARLM does not reproduce the covariance learning of CMA-ES or the population-difference mutation of jDE; instead, it combines generation-level success control with bounded coordinate-wise scale learning. Its potential advantage is therefore not greater representational power, but a comparatively lightweight adaptation mechanism for bounded calibration problems.

The hyperparameter analysis further showed that performance was not dominated by most of the tested one-factor perturbations over the examined ranges. Only increasing the permanent lower mutation bound from 10^{-3} to 5×10^{-3} remained significantly detrimental after multiplicity correction, consistent with excessive lower-bound occupancy and reduced late-stage contraction. The reference settings should nevertheless be regarded as a predefined reference configuration showing limited sensitivity over the examined one-factor ranges, rather than as globally optimized hyperparameters. Similarly, the branch-wise contraction analysis establishes only local conditional behaviour: amplitude projection guarantees bounded internal mutation scales but does not provide a global stochastic-convergence theorem.

The identifiability results substantially qualify the interpretation of the low calibration errors. Local Fisher-information and one-dimensional profile analyses indicated severe ill-conditioning, while broad-start optimization and two-dimensional profiling showed that widely separated parameter vectors could retain nearly identical objective values. The pronounced (γ_1 and γ_{H_2}) compensation valley demonstrates that this ambiguity is not restricted to the immediate neighbourhood of the reference solution. Such behaviour is

consistent with the distinction between successful trajectory fitting and practical parameter identifiability emphasized in the identifiability literature [12,15,18,19,42]. The leave-one-interval-out analysis further showed highly uneven observational influence, with the 0–1 day interval carrying substantially more information than the late near-zero intervals.

Fitted-trajectory uncertainty was also sensitive to the assumed observation-error structure. Additive Gaussian, multiplicative log-normal, and temporally correlated Gaussian perturbations produced materially different uncertainty band widths at the same nominal perturbation level. Because replicate-specific variance and temporal error statistics were unavailable, these scenarios should be interpreted as sensitivity analyses rather than identified experimental error laws. The principal limitations therefore remain the single sparse dataset, one measured output, six observations for six Andrews–P parameters, absence of independent predictive validation, lack of replicate-derived error statistics, and absence of formal structural-identifiability analysis. Broader assessment should include independent datasets, denser early-time sampling, additional measured states, and further dynamic model classes.

9. Conclusions

This proof-of-concept study evaluated B-ARLM as a bounded derivative-free strategy for parameter calibration in a sparse nonlinear bioprocess model. Among the three examined kinetic formulations, Andrews–P provided the lowest in-sample calibration error for the six interval-average observations. This result supports its use as the best-fitting formulation within the present dataset and calibration protocol but does not establish a uniquely correct biological inhibition mechanism or predictive superiority outside the observed experiment.

Under a common budget of 2250 objective-function evaluations per run, B-ARLM showed statistically significant improvement over the direct Fixed-EA ablation and the study-specific sCA-ES reference, while its terminal performance was not statistically distinguishable from canonical CMA-ES or jDE after Holm correction. The hyperparameter analysis indicated limited dependence on most examined one-factor perturbations, although an increased permanent lower mutation bound significantly degraded performance. These findings support B-ARLM as a competitive bounded calibration method for the present problem rather than as a universally superior optimizer.

The identifiability analyses showed that a low calibration error did not imply unique kinetic-parameter recovery. Broad-start optimization and two-dimensional profiling identified widely separated near-optimal parameter vectors and strong parameter compensation, while the fitted-trajectory uncertainty depended materially on the assumed observation-error structure. The main limitations remain the single sparse dataset, one measured output, absence of independent predictive validation, lack of replicate-derived error statistics, and absence of formal structural-identifiability analysis. Future work should evaluate B-ARLM on independent datasets, additional dynamic-model classes, and established benchmark problems. Further analysis is also required to characterize its convergence behaviour and computational performance beyond the present proof-of-concept setting.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/math14183315/s1.

Author Contributions: Conceptualization, V.L. and O.R.; methodology, V.L., O.R., and E.C.; software, V.L., O.R., A.Z., D.K., and G.R.; validation, V.L., O.R., E.C., and M.I.; formal analysis, V.L., O.R., E.C., and M.I.; investigation, V.L., O.R., E.C., A.Z., D.K., and G.R.; writing—original draft preparation, V.L., O.R., and E.C.; writing—review and editing, V.L., O.R., E.C., A.Z., D.K., and G.R.; visualization, D.K. and G.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are contained within the article and Supplementary Materials. Additional underlying experimental data are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

B-ARLM	Bounded Adaptive Random Local Mutation
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
DE	Differential Evolution
FIM	Fisher Information Matrix
Fixed-EA	Fixed-Mutation Evolutionary Algorithm
IQR	Interquartile Range
jDE	Self-Adaptive Differential Evolution
LOO	Leave-One-Interval-Out
NMSE	Normalized Mean Squared Error
RK4	Fourth-Order Runge–Kutta Method
sCA-ES	Simplified Covariance-Adaptive Evolution Strategy
Symbol	Definition
$\boldsymbol{\theta}$	Calibrated parameter vector; the individual kinetic parameters and their units are defined in Table 2
$\boldsymbol{\Theta}$	Admissible parameter domain
A	Set of active parameter-coordinate indices
$\boldsymbol{d}$	Number of active calibrated parameters
$\boldsymbol{G}$	Prescribed number of generations
$\boldsymbol{N}$	Population size
$\boldsymbol{J}(\boldsymbol{\theta})$	Normalized least-squares calibration objective
$\boldsymbol{J}_{\mathbf{best}}^{(g)}$	Global best objective value available at generation $\boldsymbol{g}$
$\boldsymbol{\mathcal{S}_g}$	Set of candidates that improve upon the previous global best at generation $\boldsymbol{g}$
$\boldsymbol{r_g}$	Generation-level success rate
$\boldsymbol{\sigma_j^{(g)}}$	Normalized mutation amplitude for parameter coordinate $\boldsymbol{j}$ at generation $\boldsymbol{g}$
$\boldsymbol{\sigma_o}$	Initial mutation amplitude
$\boldsymbol{\sigma_{\min}, \sigma_{\max}}$	Permanent lower and upper mutation-amplitude bounds
$\boldsymbol{p^*}$	Target success rate
$\boldsymbol{\eta}$	Coordinate-wise learning rate
$\boldsymbol{m_{j,g}}$	Successful-displacement scale statistic for coordinate j
$\boldsymbol{\alpha}$	Multiplicative success-rate adaptation factor
$\boldsymbol{c_g}$	Expansion/contraction factor
$\boldsymbol{q_{H_2}(t;\theta)}$	Instantaneous volumetric hydrogen-production rate
$\mathbf{S_{log}}$	Log-parameter sensitivity matrix
$\mathbf{S}_w$	Whitened log-parameter sensitivity matrix
$\mathbf{F}$	Scaled Fisher-information matrix
$\boldsymbol{\theta_{ref}}$	Selected reference Andrews–P parameter vector
$\boldsymbol{J_{prof,j}(a)}$	Minimum objective value attainable when parameter component $\boldsymbol{\theta_j}$ is fixed at $\boldsymbol{a}$
$\boldsymbol{\rho}$	Prescribed relative perturbation level
$\boldsymbol{\sigma_{obs}}$	Prescribed sensitivity-whitening scale

References

1. Banga, J.R. Optimization in computational systems biology. *BMC Syst. Biol.* **2008**, *2*, 47. [CrossRef]
2. Moles, C.G.; Mendes, P.; Banga, J.R. Parameter estimation in biochemical pathways: A comparison of global optimization methods. *Genome Res.* **2003**, *13*, 2467–2474. [CrossRef]
3. Ashyraliyev, M.; Fomekong-Nanfack, Y.; Kaandorp, J.A.; Blom, J.G. Systems biology: Parameter estimation for biochemical models. *FEBS J.* **2009**, *276*, 886–902. [CrossRef]
4. Grijspeerdt, K.; Vanrolleghem, P. Estimating the parameters of the Baranyi model for bacterial growth. *Food Microbiol.* **1999**, *16*, 593–605. [CrossRef]
5. Akkermans, S.; Logist, F.; Van Impe, J.F. Parameter estimations in predictive microbiology: Statistically sound modelling of the microbial growth rate. *Food Res. Int.* **2018**, *106*, 1105–1113. [CrossRef]
6. Zwietering, M.H.; Jongenburger, I.; Rombouts, F.M.; Van 't Riet, K. Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.* **1990**, *56*, 1875–1881. [CrossRef]
7. Dochain, D. (Ed.) *Bioprocess Control*; Wiley-ISTE: London, UK, 2008. [CrossRef]
8. Iglesias, P.A.; Ingalls, B.P. (Eds.) *Control Theory and Systems Biology*; MIT Press: Cambridge, MA, USA, 2010.
9. Villadsen, J.; Nielsen, J.; Lidén, G. *Bioreaction Engineering Principles*, 3rd ed.; Springer: New York, NY, USA, 2011. [CrossRef]
10. Joshi, H.; Jha, B.K. Modeling the spatiotemporal intracellular calcium dynamics in nerve cell with strong memory effects. *Int. J. Nonlinear Sci. Numer. Simul.* **2023**, *24*, 2383–2403. [CrossRef]
11. Linden, N.J.; Kramer, B.; Rangamani, P. Bayesian parameter estimation for dynamical models in systems biology. *PLoS Comput. Biol.* **2022**, *18*, e1010651. [CrossRef]
12. Simpson, M.J.; Maclaren, O.J. Profile-Wise Analysis: A profile likelihood-based workflow for identifiability analysis, estimation, and prediction with mechanistic mathematical models. *PLoS Comput. Biol.* **2023**, *19*, e1011515. [CrossRef]
13. Murphy, R.J.; Maclaren, O.J.; Simpson, M.J. Implementing measurement error models with mechanistic mathematical models in a likelihood-based framework for estimation, identifiability analysis and prediction in the life sciences. *J. R. Soc. Interface* **2024**, *21*, 20230402. [CrossRef]
14. Plank, M.J.; Simpson, M.J. Structured methods for parameter inference and uncertainty quantification for mechanistic models in the life sciences. *R. Soc. Open Sci.* **2024**, *11*, 240733. [CrossRef]
15. Heinrich, M.; Rosenblatt, M.; Wieland, F.-G.; Stigter, H.; Timmer, J. On Structural and Practical Identifiability: Current Status and Update of Results. *Curr. Opin. Syst. Biol.* **2025**, *41*, 100546. [CrossRef]
16. Heinrich, M.; Arutjunjan, R.; Timmer, J. On the Different Flavours of Practical Identifiability. *Curr. Opin. Syst. Biol.* **2025**, *42*, 100556. [CrossRef]
17. Kreutz, C. An Easy and Efficient Approach for Testing Identifiability. *Bioinformatics* **2018**, *34*, 1913–1921. [CrossRef]
18. Preston, S.P.; Wilkinson, R.D.; Clayton, R.H.; Chappell, M.J.; Mirams, G.R. Think before You Fit: Parameter Identifiability, Sensitivity and Uncertainty in Systems Biology Models. *Curr. Opin. Syst. Biol.* **2025**, *42*, 100563. [CrossRef]
19. Wang, S.; Hao, W. A Systematic Computational Framework for Practical Identifiability Analysis in Mathematical Models Arising from Biology. *Adv. Sci.* **2025**, *12*, e04346. [CrossRef]
20. Balsa-Canto, E.; Campo-Manzanares, N.; Moimenta, A.R.; Roudaut, G.; Troitiño-Jordedo, D. Quantifying and Managing Uncertainty in Systems Biology: Mechanistic and Data-Driven Models. *Curr. Opin. Syst. Biol.* **2025**, *42*, 100557. [CrossRef]
21. Mendes, P.; Kell, D.B. Nonlinear optimization of biochemical pathways: Applications to metabolic engineering and parameter estimation. *Bioinformatics* **1998**, *14*, 869–883. [CrossRef]
22. Balsa-Canto, E.; Banga, J.R. AMIGO, a toolbox for advanced model identification in systems biology using global optimization. *Bioinformatics* **2011**, *27*, 2311–2313. [CrossRef]
23. Carlozo, M.N.; Wang, K.; Dowling, A.W. Bayesian optimization methods for nonlinear model calibration. *Ind. Eng. Chem. Res.* **2025**, *64*, 18277–18297. [CrossRef]
24. Storn, R.; Price, K. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **1997**, *11*, 341–359. [CrossRef]
25. Das, S.; Suganthan, P.N. Differential evolution: A survey of the state-of-the-art. *IEEE Trans. Evol. Comput.* **2011**, *15*, 4–31. [CrossRef]
26. Price, K.V.; Storn, R.M.; Lampinen, J.A. *Differential Evolution: A Practical Approach to Global Optimization*; Springer: Berlin/Heidelberg, Germany, 2005. [CrossRef]
27. Kennedy, J.; Eberhart, R.C. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks (ICNN'95), Perth, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
28. Hansen, N.; Ostermeier, A. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **2001**, *9*, 159–195. [CrossRef]
29. Bäck, T. *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*; Oxford University Press: New York, NY, USA, 1996.

30. Beyer, H.-G.; Schwefel, H.-P. Evolution strategies—A comprehensive introduction. *Nat. Comput.* **2002**, *1*, 3–52. [CrossRef]
31. Morinaga, D.; Fukuchi, K.; Sakuma, J.; Akimoto, Y. Convergence Rate of the (1+1)-Evolution Strategy with Success-Based Step-Size Adaptation on Convex Quadratic Functions. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '21), Lille, France, 10–14 July 2021; pp. 1169–1177. [CrossRef]
32. Akimoto, Y.; Auger, A.; Glasmachers, T.; Morinaga, D. Global Linear Convergence of Evolution Strategies on More Than Smooth Strongly Convex Functions. *SIAM J. Optim.* **2022**, *32*, 1402–1429. [CrossRef]
33. Hansen, N. The CMA Evolution Strategy: A Tutorial. *arXiv* **2016**, arXiv:1604.00772. [CrossRef]
34. Akimoto, Y.; Hansen, N. Diagonal Acceleration for Covariance Matrix Adaptation Evolution Strategies. *Evol. Comput.* **2020**, *28*, 405–435. [CrossRef]
35. Brest, J.; Greiner, S.; Bošković, B.; Mernik, M.; Žumer, V. Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. *IEEE Trans. Evol. Comput.* **2006**, *10*, 646–657. [CrossRef]
36. Sun, W.; Yuan, Y.-X. *Optimization Theory and Methods: Nonlinear Programming*; Springer: New York, NY, USA, 2006. [CrossRef]
37. Edgar, T.F.; Himmelblau, D.M.; Lasdon, L.S. *Optimization of Chemical Processes*, 2nd ed.; McGraw-Hill: New York, NY, USA, 2001.
38. Schittkowski, K. *Numerical Data Fitting in Dynamical Systems: A Practical Introduction with Applications and Software*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2002.
39. Bäck, T.H.W.; Kononova, A.V.; van Stein, B.; Wang, H.; Antonov, K.A.; Kalkreuth, R.T.; de Nobel, J.; Vermetten, D.; de Winter, R.; Ye, F. Evolutionary Algorithms for Parameter Optimization—Thirty Years Later. *Evol. Comput.* **2023**, *31*, 81–122. [CrossRef]
40. Eiben, A.E.; Jelasity, M. A Critical Note on Experimental Research Methodology in EC. In Proceedings of the 2002 Congress on Evolutionary Computation (CEC 2002), Honolulu, HI, USA, 12–17 May 2002; Volume 1, pp. 582–587. [CrossRef]
41. Birattari, M.; Dorigo, M. How to Assess and Report the Performance of a Stochastic Algorithm on a Benchmark Problem: Mean or Best Result on a Number of Runs? *Optim. Lett.* **2007**, *1*, 309–311. [CrossRef]
42. Raue, A.; Kreutz, C.; Maiwald, T.; Bachmann, J.; Schilling, M.; Klingmüller, U.; Timmer, J. Structural and Practical Identifiability Analysis of Partially Observed Dynamical Models by Exploiting the Profile Likelihood. *Bioinformatics* **2009**, *25*, 1923–1929. [CrossRef]
43. Chis, O.-T.; Banga, J.R.; Balsa-Canto, E. Structural Identifiability of Systems Biology Models: A Critical Comparison of Methods. *PLoS ONE* **2011**, *6*, e27755. [CrossRef]
44. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall: New York, NY, USA, 1993.
45. Monod, J. The Growth of Bacterial Cultures. *Annu. Rev. Microbiol.* **1949**, *3*, 371–394. [CrossRef]
46. Andrews, J.F. A Mathematical Model for the Continuous Culture of Microorganisms Utilizing Inhibitory Substrates. *Biotechnol. Bioeng.* **1968**, *10*, 707–723. [CrossRef]
47. Van den Heuvel, J.C.; Beeftink, H.H. Kinetic Effects of Simultaneous Inhibition by Substrate and Product. *Biotechnol. Bioeng.* **1988**, *31*, 718–724. [CrossRef]