International Journal of Multiphysics

Volume 14, No. 2, 2020

ISSN: 1750-9548

Machine Learning-Assisted Mathematical Modeling of Evolutionary Biochemical Processes

[1] Sujata [2]Dr. Ranjana Singh , [3]Dr. Anjana Farswan

[1]Assistant Professor, Dept.of Mathematics

Pt.L.M.S.Govt.Post Graduate College Rishikesh (Dehradun) (Autonomous College)

sujata.singh08@gmail.com

[2]Assistant Professor Chemistry

Dr. P. D. B. H Government PG college Kotdwar

ranjanasingh671976@gmail.com

[3]Assistant Professor (Zoology) A.P.B. Govt. P.G.College Agastyamuni Rudraprayag Uttarakhand
anjanaprabhakar9@gmail.com

Abstract

Background: Evolutionary biochemical systems consist of multiscale interactions between biochemistry and reaction kinetics as they relate to mutation, selection, growth of populations, and feedback loops from metabolic products. Though the governing mechanisms of evolutionary biochemical systems are comprehensible, the kinetic parameters are often poorly identifiable. The objective of this study is to design a framework that combines mathematics and machine learning to estimate uncertain biochemical parameters and amend mechanistic dynamics that are incomplete, all while honoring constraints of biology. The methods of this study include a model composed of six coupled nonlinear ordinary differential equations which treat substrates, enzymes, products, population, the frequency of beneficial mutations, and mean fitness. The existence, uniqueness, positivity, boundedness, and local stability of this model were established, and synthetic trajectories were generated through adaptive Runge–Kutta integration. Sparse trajectories that were noisy were used to create a pure neural predictor that is complemented by a residual neural model in a neural framework of a misspecified ODE. The system converged to the positive equilibrium (0.2743,10.6019,7.8683,18.4012,0.8506,1.9291), and the eigenvalues of the six-dimensional Jacobian had all negative real components. The other system of residual learning in the misspecified ODE had an RMSE of 0.6790, and pure machine learning had an RMSE of 0.1262. The best hybrid model had an RMSE of 0.0513 and a relative L2 error of 0.0056. The residual neural model came from a mechanistic framework and helped the author to model a real evolutionary biochemical system when data were scarce. The biochemical interpretation was further developed by relating the catalytic flux to the elementary enzyme scheme $E + S \rightleftharpoons ES \rightarrow E + P$ and by identifying substrate, enzyme, product, cofactors, buffer, and biomass as the principal chemical or biochemical components. A physical-chemistry layer was also added to connect kinetic parameters with activation free energy, temperature, pH, ionic strength, reversibility, and transport assumptions.

Keywords: *biochemical kinetics; evolutionary dynamics; physics-informed neural networks; neural ordinary differential equations; parameter estimation; stability analysis; scientific machine learning*; enzyme catalysis; reaction thermodynamics; physical chemistry

1. Introduction

Biochemical evolution displays interacting processes occurring at different timeframes. For example, enzymes act on substrates, and the resulting product impacts cell growth. The product may also influence the mutation rate, which changes the enzyme production and catalytic efficiency. As for the evolution of advantageous traits in a population, natural selection impacts their prevalence. Classical ordinary differential equation (ODE) models explicitly describe these processes and allow equilibrium, stability, and sensitivity analyses. The main limitation of these models is their incompleteness rather than interpretability, as they may conceal dynamics through simplified rate laws and unmeasured species. Several parameter combinations can also lead to similar predictions.

Scientific machine learning seeks to combine data-driven flexibility with mechanistic structure rather than treating the two as alternatives (Karpatne et al., 2017; Raissi et al., 2019). Physics-informed neural networks embed differential-equation residuals in training objectives (Raissi et al., 2019), while universal differential equations place trainable components inside mechanistic dynamical systems (Rackauckas et al., 2020). Systems-biology-informed networks have inferred kinetic parameters and hidden states from sparse measurements (Yazdani et al., 2020), and biologically informed networks have helped discover unknown constitutive functions from limited experimental data (Lagergren et al., 2020). Earlier hybrid process-modeling research likewise emphasized identifiability, extrapolation, and the separation of mechanistic and learned components (Psichogios & Ungar, 1992; von Stosch et al., 2014).

In this paper, a compact evolutionary biochemical model is proposed with a residual neural network (RNN) correction. The objectives are: to formulate coupled reaction-population-mutation processes; to demonstrate biological plausibility and well-posedness; to estimate kinetic and evolutionary parameters with a low level of certainty; and to compare the classical, fully data-driven, and hybrid approaches. Additionally, the work intends to understand which parameters are involved in the dynamics of product formation, population adaptation, and selection of the optimal parameter.

2. Related Literature

Machine learning in biochemical and process engineering includes soft sensing, optimization, reaction prediction, and process control (Kadlec et al., 2009; Venkatasubramanian, 2019). Purely empirical models can fail to obey conservation, positivity, or sensible extrapolation, motivating hybrid model structures (Psichogios & Ungar, 1992; von Stosch et al., 2014). In metabolic modeling, flux-balance analysis and genome-scale models preserve pathway interpretation while accommodating omics or machine-learning information (Orth et al., 2010; Lewis et al., 2012; O'Brien et al., 2015; Zampieri et al., 2019). Multi-omics increases data coverage but retains challenges involving high dimensionality, interpretability, and heterogeneous data integration (Angermueller et al., 2016; Zitnik et al., 2019).

A structured compromise is hybrid or grey-box modeling. Reviews of hybrid semi-parametric process models show that mechanistic and data-driven submodels can improve inference and extrapolation when each component is assigned a clear role (von Stosch et al., 2014). For sparse data, physics-informed learning uses differential-equation residuals as regularizers (Raissi et al., 2019; Yazdani et al., 2020). Stiff biochemical kinetics and separated timescales may require scaling, stiff solvers, or quasi-steady-state reduction before reliable learning (Segel & Slemrod, 1989; Eilertsen & Schnell, 2020). Parameter uncertainty and multimodality should be examined with profile likelihood, Bayesian inference, and sensitivity analysis rather than prediction error alone (Gutenkunst et al., 2007; Raue et al., 2009; Toni et al., 2009).

Hybrid neural ordinary differential equations and trainable differential-equation models build on neural ODEs (Chen et al., 2018) and universal differential equations (Rackauckas et al., 2020), while graph-based machine learning offers a complementary route for learning relationships in biological networks (Camacho et al., 2018). Likelihood-free inference can estimate parameters and compare dynamical models when explicit likelihoods are unavailable (Toni et al., 2009). Structural identifiability concerns whether model parameters can, in principle, be uniquely recovered from ideal observations, whereas practical identifiability additionally reflects experimental design, noise, initial conditions, and the measured states (Raue et al., 2009; Villaverde et al.,

2016). Correlated rate constants can generate nearly identical trajectories, allowing a neural component to hide poor parameter estimates behind low prediction error (Gutenkunst et al., 2007). To overcome the limitations of point predictions, it is important to examine profile likelihoods, Bayesian posteriors, sensitivity geometry, and perturbation experiments (Gutenkunst et al., 2007; Raue et al., 2009; Toni et al., 2009). Earlier hybrid-modeling studies recommend assigning the learned component only to unknown functions and evaluating the model under conditions not used for calibration (Psichogios & Ungar, 1992; Thompson & Kramer, 1994; von Stosch et al., 2014). Mechanistic constraints can regularize learned terms and improve extrapolation (Karpatne et al., 2017; Raissi et al., 2019). These ideas justify analyzing mechanistic error, prediction error, residual structure, sensitivity, and stability rather than accuracy alone.

A key difficulty with modelling is dissociating the timescales of evolution and biochemistry. Catalytic turnover can proceed very rapidly, but changes in the frequency of diploid alleles may take many generations. Treating all parameters as if they are uniformly observable may create the illusion of non-identifiable parameters. For instance, a fast parameter may be close to equilibrium at the first observation, whereas a slow parameter may not be sufficiently variable to give an estimate of its rate of change. Some modern hybrid models attempt to address this through the continuous information sharing along the entire path. Automatic differentiation enables the evaluation of derivatives at collocation points, and mechanistic constraints allow for the information of observed states to be communicated to the unobserved states. These constraints, however, are only useful if the proposed equations are at least somewhat correct. A very rigid physics loss may push a network towards an incorrect model, while a very loose physics loss may turn the hybrid model into an unconstrained interpolator. Therefore, loss balancing, state scaling, and structural diagnostics are important design choices for such models.

3. Mathematical Framework

3.1 State variables and assumptions

Let S(t) represent the concentration of substrate; E(t) for enzyme concentration; P(t) for product concentration; N(t) for population density; M(t) for frequency of the favorable mutation; and F(t) for average fitness. The assumptions made are: the environment is well mixed, substrate conversion is via the Michaelis–Menten mechanism, population growth is density dependent, the mutation increases the production of the enzyme, product availability and the mutation enhance fitness, all parameters are positive and bounded, and $0 \leq M \leq 1$.

Table 1. Variables and biological meanings

Symbol	State variable	Unit	Interpretation
S	Substrate	mM	Available biochemical resource
E	Enzyme	normalized concentration or activity unit	Catalytically competent enzyme pool
P	Product	mM	Metabolic product supporting growth
N	Population	density unit	Replicating biochemical population
M	Mutation frequency	dimensionless	Fraction carrying beneficial mutation
F	Mean fitness	dimensionless	Average evolutionary advantage

The governing system is

$$\frac{dS}{dt} = I - \frac{V_{\max} ES}{K_m + S} - \delta_S S, \tag{1}$$

$$\frac{dE}{dt} = \alpha_E N(1 + \eta M) - \delta_E E, \tag{2}$$

$$\frac{dP}{dt} = \frac{V_{\max} ES}{K_m + S} - \delta_P P, \tag{3}$$

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right) + \beta PN - \mu N, \tag{4}$$

$$\frac{dM}{dt} = u(1 - M) + sFM(1 - M) - \omega M, \tag{5}$$

$$\frac{dF}{dt} = \lambda P + \rho M - \gamma F. \tag{6}$$

Table 2. Parameters, units, and illustrative values

Parameter	Definition	Unit	Value	Justification
I	Substrate inflow	mM h^{-1}	2.000	Synthetic forcing
$V_{\max}$	Apparent catalytic turnover coefficient (operationally kcat because E is explicit)	h^{-1}	1.200	Synthetic enzyme-turnover scale
K_m	Half-saturation constant	mM	1.500	Michaelis–Menten scale
$\delta_S, \delta_E, \delta_P$	Loss rates	h^{-1}	0.120, 0.350, 0.250	Positive turnover
α_E	Enzyme synthesis coefficient	h^{-1}	0.120	Population-linked expression
η	Mutational enzyme gain	dimensionless	0.800	Beneficial-expression effect
r, K, μ	Growth, capacity, mortality	h^{-1}, density, h^{-1}	0.320, 20, 0.120	Logistic regulation
β	Product-growth coupling	(mM h)$^{-1}$	0.012	Product benefit
u, s, ω	Mutation input, selection, loss	h^{-1}	0.004, 0.060, 0.018	Evolutionary turnover
λ, ρ, γ	Fitness gains and decay	h^{-1}	0.060, 0.080, 0.280	Product/mutation feedback

All values are synthetic and used solely for methodological demonstration; they are not laboratory estimates.

3.2 Biochemical species and chemical components represented

The model is a lumped biochemical representation rather than a simulation of one named enzyme or one specified laboratory reaction. Consequently, the synthetic experiment does not use a fixed reagent list such as a particular sugar, purified enzyme, or metabolite standard. Instead, S, E, and P represent functional chemical classes whose identities can be assigned when the framework is transferred to an experimental system. This distinction is important: the numerical values in Table 2 describe an illustrative kinetic regime and must not be interpreted as measured properties of a particular compound.

The substrate S is the chemically transformable resource that limits the catalytic reaction. In an experimental realization it could represent a carbon source, amino acid, lipid, xenobiotic, or other metabolite measured in millimolar concentration. The enzyme E denotes the total catalytically competent enzyme pool, including the fraction that is correctly folded, properly assembled, and chemically capable of binding substrate. The product P is the lumped reaction product or metabolite whose availability is assumed to support population growth. The use of lumped states is common in mechanistic biochemical engineering because it preserves interpretable reaction structure while avoiding an unidentifiable expansion into every intermediate and side reaction (O'Brien et al., 2015; von Stosch et al., 2014).

A complete laboratory system would also contain components that are implicit in the present equations. These include water as solvent; a buffer to maintain pH; salts that determine ionic strength; and, where required by the enzyme class, a cofactor or cosubstrate such as NADH, NADPH, FAD, ATP, molecular oxygen, haem, or a catalytic metal ion. The baseline model assumes that these auxiliary components remain sufficiently abundant that they do not become dynamic state variables. If a cofactor is depleted, regenerated, or competitively shared by several reactions, it should be added explicitly because the single-substrate Michaelis–Menten term would then be chemically incomplete.

Table 2A. Biochemical and physicochemical interpretation of the modeled components

Component	Chemical/biological role	Model representation	Experimental observable
Substrate (S)	Limiting reactant or biochemical resource	Consumed by catalytic flux and non-catalytic loss	HPLC/LC-MS, enzymatic assay, or spectrophotometry
Enzyme (E)	Catalytically competent protein or ribozyme	Activity pool increased by expression and reduced by turnover	Protein concentration, active-site titration, or U mL^{-1}
Product (P)	Reaction product or growth-supporting metabolite	Formed by catalytic flux and removed by product loss	Concentration, conversion, yield, or productivity
Cofactor(s)	Redox, energy, group-transfer, or metal partner	Implicit constant pool in the baseline model	NAD(P)H, ATP, FAD/FMN, haem, O_2, or metal concentration
Buffer/solvent	Controls pH, ionic strength, solvation, and water activity	Fixed environmental condition	pH, temperature, conductivity, ionic strength, and viscosity
Population (N)	Cells or replicating biochemical units	Density-dependent growth coupled to product	Optical density, dry-cell mass, CFU, or direct cell count

Note. The listed reagents and observables are representative options for experimental implementation; the numerical study itself remains a synthetic proof of concept.

3.3 Chemistry of enzyme catalysis

The chemical core of the model can be written as the elementary catalytic sequence below. The enzyme first associates reversibly with the substrate to form an enzyme–substrate complex, after which chemical conversion and product release regenerate the free catalyst.

$$E + S \rightleftharpoons ES \rightarrow E + P$$

$$v = kcat\ E\ S/(Km + S), \quad Vmax,system = kcat\ E$$

$$Km = (k{-}1 + kcat)/k1$$

Here k1 is the bimolecular association rate constant, k−1 is the dissociation rate constant, and kcat is the catalytic turnover constant. Under the quasi-steady-state approximation, the concentration of ES changes more slowly than its rapid formation and breakdown, yielding the Michaelis–Menten rate law. Km is therefore a composite kinetic constant, not automatically a pure equilibrium dissociation constant. Its interpretation as binding affinity is justified only under restricted kinetic conditions. Classical and modern analyses emphasize that reliable Michaelis–Menten parameter recovery depends on timescale separation and on the relative magnitudes of substrate, enzyme, and intermediate concentrations (Segel & Slemrod, 1989; Cornish-Bowden, 2012; Eilertsen & Schnell, 2020).

The notation in equations (1) and (3) multiplies Vmax by E. For biochemical consistency, the parameter labelled Vmax in this paper should therefore be read operationally as a first-order turnover coefficient equivalent to kcat, whereas the system-level maximum velocity is kcatE. This clarification prevents double counting of enzyme abundance. At low substrate concentration, the rate is approximately proportional to (kcat/Km)ES, and kcat/Km acts as the second-order catalytic-efficiency parameter. At high substrate concentration, the enzyme approaches saturation and the rate tends toward kcatE.

Enzymes accelerate reactions by lowering the activation free-energy barrier without changing the overall equilibrium free energy of the uncoupled reaction. The active site achieves this through a coordinated combination of substrate positioning, hydrogen bonding, electrostatic interactions, hydrophobic packing, van der Waals contacts, proton transfer, covalent intermediates, and metal-ion or cofactor chemistry. The dominant mechanism depends on enzyme class. General acid–base catalysis transfers protons through ionizable residues; covalent catalysis creates a transient enzyme-bound intermediate; metal-ion catalysis organizes charge and activates water or substrate; and electrostatic preorganization stabilizes charge development along the reaction coordinate (Fersht, 1999; Warshel et al., 2006).

Catalysis is not determined by a static active-site geometry alone. Protein conformational ensembles determine how frequently catalytically productive states are sampled, while the preorganized active site restricts the substrate into near-reactive configurations. Enzyme studies have linked conformational dynamics to catalysis and have shown that productive function can require both rigidity near transition-state-like configurations and flexibility elsewhere in the catalytic cycle (Hammes-Schiffer & Benkovic, 2006; Henzler-Wildman & Kern, 2007). Directed-evolution studies further show that distal mutations, packing, electrostatics, solvent organization, and conformational redistribution can alter catalytic rates and selectivity (Arnold, 2018; Zeymer & Hilvert, 2018).

Within the evolutionary model, a beneficial mutation may influence more than enzyme concentration. It may increase transcription or translation, improve folding and cellular stability, modify substrate affinity, alter catalytic turnover, change cofactor preference, reduce product inhibition, or shift the conformational ensemble toward a productive state. The present factor $1 + \eta M$ represents these effects only as increased enzyme production. A more chemically resolved model could allow Vmax, Km, enzyme decay, or product selectivity to depend on M so that genetic adaptation modifies catalytic chemistry as well as expression level.

3.4 Physical chemistry of the biochemical reaction

The kinetic constants in the ODE system are macroscopic summaries of molecular free-energy landscapes. Transition-state theory relates a microscopic rate constant to the activation Gibbs free energy $\Delta G‡$.

This relationship explains why a comparatively small change in active-site interactions can produce a large change in catalytic rate.

$$k = (kBT/h)\exp(-\Delta G‡/RT)$$

$$k(T) = A\exp(-Ea/RT)$$

In these expressions kB is the Boltzmann constant, h is Planck's constant, R is the gas constant, T is absolute temperature, Ea is the apparent activation energy, and A is the pre-exponential factor. The baseline simulation treats all rate constants as temperature independent because it assumes an isothermal environment. This is an approximation rather than a universal property. Temperature changes molecular collision frequency and the population of reactive conformations, while excessive temperature may reduce activity through local inactivation or global unfolding. Studies of enzyme thermal behavior show that temperature optima reflect both activation effects and reversible or irreversible losses of catalytically competent conformations (Daniel & Danson, 2010).

Thermodynamic driving force determines reaction direction. For a biochemical transformation at specified pH and ionic conditions, the Gibbs free-energy change is related to the reaction quotient Q and the apparent equilibrium constant Keq as follows.

$$\Delta G = \Delta G^{\circ\prime} + RT\ln Q$$

$$\Delta G^{\circ\prime} = -RT\ln Keq$$

The current catalytic flux is irreversible, which is reasonable when the reaction is strongly product-favoured, the product is continuously removed, or the system is studied far from equilibrium. Near equilibrium, however, forward and reverse fluxes can be similar and a reversible rate law should replace the one-way term. Thermodynamic consistency then requires the forward and reverse kinetic parameters to agree with Keq rather than being fitted independently. Product inhibition should also be included when P binds the active site or an allosteric site and reduces catalytic turnover (Beard & Qian, 2008; Cornish-Bowden, 2012).

pH affects the protonation state of catalytic residues, substrate functional groups, cofactors, and the product. Because protonation controls charge, hydrogen bonding, nucleophilicity, and leaving-group ability, both kcat and Km may vary strongly across the pH range. Ionic strength changes activity coefficients and screens electrostatic interactions, so concentration is not always identical to thermodynamic activity. The present model effectively fixes pH and ionic strength; an experimental implementation should report buffer identity, buffer concentration, temperature, salt composition, and any metal ions or reductants required for activity (Fersht, 1999; Cornish-Bowden, 2012).

Solvent viscosity and diffusion influence the rates of enzyme–substrate encounter and product release. The well-mixed assumption means that mixing and diffusion are fast relative to reaction and population growth, so every cell or enzyme molecule experiences the same local concentrations. If catalytic turnover becomes faster than transport, concentration gradients arise and the ODE framework should be replaced or extended with compartments, delay terms, or reaction–diffusion equations. Such multiscale stiffness is also relevant computationally, because very fast chemical variables and slow evolutionary variables can challenge numerical integration and model training (Hairer & Wanner, 1996; Eilertsen & Schnell, 2020).

The physical-chemistry interpretation also explains the three transient regimes observed in the simulation. The early high-substrate interval corresponds to a high chemical driving force and a rapidly increasing catalytic flux. The intermediate interval reflects increasing enzyme availability and product accumulation. The late interval is substrate limited, so additional enzyme produces little extra flux and the system approaches a balance among inflow, catalytic consumption, product loss, and population regulation. Thus, the numerical saturation is not only a mathematical feature; it is the expected consequence of finite substrate, enzyme saturation, and dissipative loss processes.

3.5 Chemically informed model extensions and parameter estimation

For experimental use, the kinetic parameters should be calibrated under multiple substrate concentrations and, where relevant, multiple temperatures, pH values, and inhibitor or product levels. Measurements concentrated only near the final equilibrium cannot reliably separate kcat from Km because the system explores a narrow section of the saturation curve. Early transient sampling, active-enzyme quantification, and perturbation of substrate input provide the curvature required for practical identifiability. Direct progress-curve fitting is preferable to reciprocal linear plots because it retains the original error structure and can incorporate all measured states (Cornish-Bowden, 2012; Raue et al., 2009).

A chemically constrained hybrid model can allow kcat and Km to depend on temperature, pH, ionic strength, genotype, or product concentration while retaining positivity and mass-balance constraints. The neural residual should be penalized if it creates product without consuming substrate, generates negative concentrations, or implies a persistent flux against the thermodynamic driving force without an explicit energy source. This approach gives the neural correction a mechanistic interpretation: it may represent product inhibition, cofactor limitation, reversible chemistry, enzyme inactivation, or transport resistance rather than an arbitrary numerical discrepancy (Psichogios & Ungar, 1992; von Stosch et al., 2014; Rackauckas et al., 2020).

The expanded chemistry therefore does not replace the six-state model. It specifies the conditions under which its parameters are meaningful and identifies the additional variables needed when the simplifying assumptions fail. This improves the scientific interpretability of machine-learning corrections and creates a direct pathway from synthetic validation to laboratory biochemical experiments.

4. Theoretical Analysis

Definition 1 (Biologically feasible region).

$$\Omega: \quad S, E, P, N, F \geq 0, \qquad 0 \leq M \leq 1. \tag{7}$$

Theorem 1 (Existence and uniqueness). For every initial state in Ω, system (1)–(6) has a unique local solution.

Proof. On $S \geq 0$, the denominator $K_m + S$ is bounded away from zero. Every right-hand side is continuously differentiable in the state variables and is therefore locally Lipschitz on Ω. Picard–Lindelöf gives a unique local solution. ▫

Theorem 2 (Positivity and mutation invariance). Solutions beginning in Ω remain non-negative and satisfy $0 \leq M(t) \leq 1$.

Proof. On the coordinate boundaries, $\left.\frac{dS}{dt}\right|_{S=0} = I > 0$, $\left.\frac{dE}{dt}\right|_{E=0} = \alpha_E N(1+\eta M) \geq 0$, $\left.\frac{dP}{dt}\right|_{P=0} = V_{\max} ES/(K_m + S) \geq 0$, $\left.\frac{dN}{dt}\right|_{N=0} = 0$, and $\left.\frac{dF}{dt}\right|_{F=0} = \lambda P + \rho M \geq 0$. Moreover, $\left.\frac{dM}{dt}\right|_{M=0} = u \geq 0$ and $\left.\frac{dM}{dt}\right|_{M=1} = -\omega \leq 0$. Boundary pointing therefore makes Ω positively invariant. ▫

Theorem 3 (Boundedness). Every solution in Ω is ultimately bounded.

Proof. Substrate satisfies $dS/dt \leq I - \delta_s S$, giving $S(t) \leq \max\{S(0), I/\delta_s\}$. Enzyme production is linear in N, product production is bounded by $V_{\max} E$, and fitness is linear in bounded P and M. For population, once $P \leq P_{\max}$, $dN/dt \leq (r + \beta P_{\max} - \mu)N - (r/K)N^2$. Applying Young's inequality permits positive weights a_i and constants $C, \xi > 0$ such that $W = S + a_1 E + a_2 P + a_3 N + a_4 F$ obeys $dW/dt \leq C - \xi W$. Hence

$$W(t) \leq W(0)e^{-\xi t} + \frac{C}{\xi}\left(1 - e^{-\xi t}\right). \tag{8}$$

Thus no component can diverge in finite or infinite time. ▫

Define $a = V_{\max} E K_m/(K_m + S)^2$ and $b = V_{\max} S/(K_m + S)$. The Jacobian of the six-state vector field is

$$J(X)=\begin{pmatrix} -a-\delta_s & -b & 0 & 0 & 0 & 0 \\ 0 & -\delta_E & 0 & \alpha_E(1+\eta M) & \alpha_E\eta N & 0 \\ a & b & -\delta_P & 0 & 0 & 0 \\ 0 & 0 & \beta N & r(1-2N/K)+\beta P-\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & -u+sF(1-2M)-\omega & sM(1-M) \\ 0 & 0 & \lambda & 0 & \rho & -\gamma \end{pmatrix}. \quad (9)$$

Theorem 4 (Local stability). An equilibrium X_{eq} is locally asymptotically stable if all eigenvalues of $J(X_{\mathrm{eq}})$ have negative real parts.

Proof. The vector field is continuously differentiable. Linearisation yields $dz/dt = J(X_{\mathrm{eq}})z + o(\| z \|)$. The Lyapunov indirect method gives local asymptotic stability when $J(X_{\mathrm{eq}})$ is Hurwitz. ▫

Theorem 5 (Residual approximation). Let the omitted residual $g(X,t)$ be continuous on compact $D \subset \Omega \times [0,T]$. For every $\varepsilon > 0$, a sufficiently wide feed-forward neural network g_θ exists such that $\| g(X,t) - g_\theta(X,t) \| < \varepsilon$ for every $(X,t) \in D$.

Proof. Continuous non-polynomial activation networks are dense in $C(D)$. Componentwise approximation gives the vector result. This theorem concerns representation; successful optimisation and identifiability require additional data and regularisation assumptions. ▫

Table 3. Initial conditions

State	Initial value
$S(0)$	10.00
$E(0)$	2.00
$P(0)$	0.00
$N(0)$	5.00
$M(0)$	0.05
$F(0)$	0.10

5. Machine-Learning Methodology

The hybrid vector field is

$$\hat{f}_\theta(X,t) = f(X,\Theta) + g_\theta(X,t), \quad (10)$$

where f is the mechanistic ODE and g_θ is a neural residual. The state input is $X_t = [S,E,P,N,M,F]$ and the discrete prediction is $\hat{X}_{t+\Delta t} = \mathcal{M}_\theta(X_t)$. Training uses the composite objective

$$\mathcal{L} = \lambda_d \mathcal{L}_{data} + \lambda_p \mathcal{L}_{physics} + \lambda_i \mathcal{L}_{initial} + \lambda_b \mathcal{L}_{boundary}, \quad (11)$$

$$\mathcal{L}_{data} = \frac{1}{n}\sum_{i=1}^{n} \| X_i - \hat{X}_i \|_2^2, \qquad \mathcal{L}_{physics} = \frac{1}{n}\sum_{i=1}^{n} \left\| \frac{d\hat{X}_i}{dt} - f(\hat{X}_i,\Theta) \right\|_2^2. \quad (12)$$

The training-set means and standard deviations are used to standardize states. Because of sparse observations, data is split chronologically to form a training (70%), validation (15%), and testing (15%) sets. To prevent information leakage, two-layer tanh networks learn the full ML trajectories (pure ML) or the discrepancies of misspecified ODEs (hybrid). Adam pretraining followed by the limited memory BFGS optimizer is preferred. However, for the sake of a compact demonstration, the optimas used deterministic

multilayer perceptron training with early stopping and L_2 regularization. M is either constrained or parameterized using a logistic function. Positivity is also enforced through a transformation to a nonnegative output.

The data and physics terms were evaluated at different point sets. For the data term, only the observation times were used. However, the physics residual was computed over a much denser sample across the entire training time. This was done to ensure the ODE is satisfied outside the observation times. Initial-condition loss aligns the trajectory at t = 0. Additionally, boundary loss penalizes negative concentrations of state variables outside the [0,1] bounds. Because the scales of the six states differ, residuals were normalized before summation. Otherwise, the population and enzyme equations would numerically dominate the mutation equation.

In transformed coordinates, parameter estimation was carried out. Positive rates adopted a logarithmic parameter, and the softplus transformation on the carrying capacity ensured that the optimiser did not enter outside of biologically feasible regions. The heuristic estimates provided in Table 8 were derived using non-linear least squares on the noisy training data, and then the residual network was trained conditionally with the calibrated mechanistic core. In a broader context, to account for the flat directions of the neural differential equations, and the correlated parameter-residual trade-offs, the simultaneous adjustment of both parameters and network weights should be performed multiple times due to the presence of local minima. Profile likelihood, posterior sampling, or bootstrap ensembles should be taken into account to establish whether low prediction error is achieved at the cost of uniquely determined biological parameters (Gutenkunst et al., 2007; Raue et al., 2009; Toni et al., 2009).

Table 4. Hybrid architecture and hyperparameters

Component	Setting
Input/output dimension	6 / 6
Hidden layers	2 layers, 48 neurons each
Activation	tanh
Optimisation	quasi-Newton demonstration; Adam + L-BFGS recommended
Training/validation/test	70/15/15 chronological split
Regularisation	$L_2 = 2 \times 10^{-3}$; physics residual
Sampling	Sparse observations every 1.5 h; dense collocation grid
Convergence rule	Validation improvement below 10^{-6} or maximum epochs

6. Numerical Experiment and Results

Synthetic computational data generated solely for methodological demonstration. System (1)–(6) was integrated on $0 \leq t \leq 60$ h with 0.1 h output spacing using an adaptive high-order Runge–Kutta method. Gaussian measurement noise was added to 41 sparse observations. The classical comparator used deliberately perturbed values of $V_{\max}, K_m, \alpha_E, r, s, \beta$, and γ to represent calibration error.

The substrate declined rapidly as enzyme abundance increased, while product accumulated and stabilised. Product-supported growth moved N toward a high-density equilibrium. Mutation frequency rose from 0.05 to approximately 0.85 because selection outweighed mutational loss, and fitness approached 1.93. These

trajectories are biologically coherent with the assumed positive feedback from product to growth and from mutation to enzyme synthesis.

Table 5. Estimated positive equilibrium

Variable	Equilibrium estimate
S_{eq}	0.2743
E_{eq}	10.6019
P_{eq}	7.8683
N_{eq}	18.4012
M_{eq}	0.8506
F_{eq}	1.9291

Table 6. Numerical eigenvalues of $J(X_{\mathrm{eq}})$

Index	Eigenvalue
1	-6.1813
2	-0.0997
3	-0.3809
4	$-0.2576 + 0.0340i$
5	$-0.2576 - 0.0340i$
6	-0.2820

Every real part is negative, numerically verifying the sufficient condition in Theorem 4 for this parameter set. The slow eigenvalue near -0.10 explains the gradual final approach of mutation and fitness compared with the faster biochemical transient.

Prediction metrics were

$$MAE = \frac{1}{n}\sum|y_i - \hat{y}_i|, \quad RMSE = \sqrt{\frac{1}{n}\sum(y_i - \hat{y}_i)^2}, \tag{13}$$

$$R^2 = 1 - \frac{\sum(y_i-\hat{y}_i)^2}{\sum(y_i-\bar{y})^2}, \quad E_{L_2} = \frac{\|X-\hat{X}\|_2}{\|X\|_2}. \tag{14}$$

Table 7. Held-out prediction performance

Model	MAE	RMSE	R^2	Relative L_2 error
Misspecified classical ODE	0.4875	0.6790	0.9889	0.0740

Model	MAE	RMSE	R^2	Relative L_2 error
Pure machine learning	0.0926	0.1262	0.9996	0.0138
Hybrid residual model	0.0332	0.0513	0.99994	0.0056

The hybrid model reduced RMSE by 92.4% relative to the misspecified ODE and by 59.3% relative to pure ML. This result is not evidence of universal superiority; it demonstrates the expected advantage when the mechanistic core is qualitatively correct and the residual is smooth.

Table 8. Parameter-estimation results

Parameter	True synthetic value	Estimated value	Absolute percentage error
$V_{\max}$	1.2000	1.2530	4.42%
K_m	1.5000	1.3623	9.18%
r	0.3200	0.3202	0.08%
s	0.0600	0.0605	0.82%
β	0.0120	0.01247	3.88%
γ	0.2800	0.28046	0.16%

Normalised sensitivity was calculated by

$$S_p^Y = \frac{\partial Y}{\partial p}\frac{p}{Y}. \tag{15}$$

Table 9. Normalised sensitivities at 60 h

Parameter	Product P	Population N	Mutation M	Fitness F
$V_{\max}$	0.020	0.006	0.019	0.021
K_m	-0.017	-0.005	-0.010	-0.016
r	0.002	0.088	0.005	0.002
s	0.003	0.001	0.403	0.060
u	0.000	0.000	0.046	0.008
β	0.006	0.323	0.004	0.006
λ	0.003	0.001	0.368	0.936
γ	-0.003	-0.001	-0.388	-1.053

Fitness was primarily controlled by gain λ and decay γ; mutation frequency was most sensitive to selection s, λ, γ, and decay; and population density primarily responded to coupling β of product growth. The relatively weak terminal sensitivity to V_max was due to compensation when the system reached a substrate-limited balance.

Transitory behavior shows three patterns. During the early resource rich stage, there is high substrate availability and the rate of product formation greatly increases. During the later stage, population density and mutation frequency increase in parallel, further strengthening the synthesis of enzymes via the factor 1+ηM. Finally, substrate becomes the limiting factor and the system approaches a product-sustained balance. The complex-conjugate eigenvalue pair of this system shows a weakly damped component. The slowest mode is associated with evolutionary adjustment and, as anticipated, explains the late convergence of M and F.

In a scientifically informative way, parameter recovery was inconsistent. Growth rate r, selection s, and decay of fitness γ were well recovered since they set the clear medium- and long-term curvatures. The largest relative error was for the Michaelis constant K_M because substrate dropped to low concentrations and thus there were few observations near half saturation. This is a consequence of experimental design: to estimate both V_max and K_m, there should be measurements concentrated near the steepest part of the Michaelis–Menten response, not only near the equilibrium.

Residual analysis identified early held-out interval extrapolated residuals as the largest errors. These transitional residuals were corrected by applying residual correction via interpolation and then extrapolation. Loss curves showed validation losses decline in parallel without late validation divergence. This consistency reflects the controlled overfitting of this seeded demonstration. More simulations with alternative noise realizations should be performed before uncertainty statements can be meaningfully added to the reported percentage improvements.

Hybrid modeling systems provide significant insight into evolving biochemical systems. The ODE describes interactions and constraints of enzyme-mediated bioprocesses, population growth, and selection pressures. It also provides feasibility guarantees that an unconstrained predictor cannot offer. The residual network can then capture smooth errors induced by uncertain rates or omitted feedback, consistent with earlier hybrid process models and systems-biology-informed learning (Psichogios & Ungar, 1992; Rackauckas et al., 2020; Yazdani et al., 2020).

The classical ODE retained systematic errors, and pure ML lacked guarantees concerning the long-term behavior of the system and mutation frequency. The hybrid approach incorporated the stable qualitative structure of the ODE and produced lower residual errors. This supports the scientific-machine-learning principle that prior knowledge can reduce sample requirements and constrain implausible solutions (Karpatne et al., 2017; Raissi et al., 2019; Yazdani et al., 2020).

Results from sensitivity analyses may have implications regarding the experimental setup. For example, a more precise measurement of the decay of fitness and the increase of products in relation to fitness may be more informative than simply refining all kinetic parameters. Alternatively, experiments designed to determine V_max and K_m may be focused on early transients before the substrate starts to impose a limit, thereby reducing the overall sensitivity. Identifiability must, therefore, be assessed over time and uncertainty intervals, and not at a specific endpoint (Gutenkunst et al., 2007; Raue et al., 2009; Villaverde et al., 2016).

Mechanistic residualisation is advantageous rather than replacing the entire vector field when there is substantial prior knowledge. If the network is allowed to learn all six derivatives, it is possible that it would reproduce an observed trajectory via biologically meaningless cancellations. By learning only a correction, the function may be understood as a map of the deficiencies of the assumed mechanism. For example, a residual that is interpreted as continuously increasing loss of the enzyme due to product accumulation may suggest product inhibition or a product burden. In this case, the interpretation would correct the prior mechanism. Interpretive use calls for regularization toward small, smooth corrections and comparison with candidate mechanisms or sparse equation-discovery methods (Psichogios & Ungar, 1992; Brunton et al., 2016; von Stosch et al., 2014).

Critical Balance Dynamics and Bidirectional Prediction

Increasing product-growth coupling, decreasing the decay of fitness, and the incorporation of delayed feedback could result in either bistability or oscillation. The Jacobian provided in this case applies only on the local scale. Global behavior would necessitate the analysis of Lyapunov functions (or stability), the monotone systems approach, bifurcation theory, or invariant sets, along with the proof of boundedness. The proof of boundedness is provided thanks to turnover, substrate, and population controls. It is important to note these constraints, as predictors may appear numerically stable over finite training times, while the unconstrained systems behind them may allow for divergence.

Limitations and Future Work

There are many aspects of biological reality that this work abstracts and bears consideration. This study uses synthetic data, one fully mixed pathway with fixed environmental forcing, deterministic dynamics, and one beneficial mutation. This study also uses only one seeded training realization. As such, the numbers presented reflect a methodological proof of concept and should not be considered actual biological quantities.

From a chemical perspective, the model also omits an explicit ES complex, reverse reaction, proton and charge balances, cofactors, competitive inhibitors, product inhibition, enzyme inactivation, heat effects, ionic-strength dependence, and spatial transport. These omissions are acceptable for the stated proof of concept but must be reconsidered before quantitative application to a named enzyme system. Future validation should therefore record active-enzyme concentration and controlled physicochemical conditions in addition to metabolite and population trajectories.

Future studies should use real measurements of enzymes, metabolites, and populations and should add uncertainty quantification through Bayesian or ensemble approaches. Reaction–diffusion and stochastic dynamics would also be valuable extensions. Graph-based models may be used to analyze reaction pathways (Camacho et al., 2018), and multi-omics data may inform parameters that change over time (Angermueller et al., 2016; Zitnik et al., 2019). Mechanistic penalties and grey-box structures should improve biological plausibility and extrapolation (Karpatne et al., 2017; von Stosch et al., 2014; Rackauckas et al., 2020). Stiff or near-equilibrium pathways require stiff-aware integration and appropriate timescale reduction (Hairer & Wanner, 1996; Eilertsen & Schnell, 2020). Connection of learned residuals to candidate reactions should be treated as an explainability exercise rather than as an unconstrained black-box interpretation.

9. Conclusion

The model presents a six-state nonlinear representation of coupled substrate conversion, expression of enzymes, accumulation of products, growth of a population, mutations, and fitness. Locally, the model is well posed, mutations are preserved, and non-negativity is respected. The model is, ultimately, bounded and for the parameter set used in the example, a positive equilibrium is numerically stable. A machine-learning method that reduces residuals did so by addressing errors in parameters and structure in a predictive fashion and preserves the mechanistic framework. The hybrid model had the best results for the held-out, synthetic data in terms of MAE, RMSE, and relative error. The primary methodological finding states that the interpretability of biochemistry and flexibility concerning data can be understood to be of a complementary nature. It was found that for effective and realistic modeling of evolution, an appropriate amount of biochemistry and mathematics is essential, as is a clear understanding of what the remaining equations are lacking. The added biochemical and physical-chemistry interpretation clarifies that the catalytic term represents an enzyme–substrate complex mechanism under fixed temperature, pH, ionic strength, cofactor availability, and mixing conditions, and it identifies reversible kinetics, inhibition, and transport as the principal chemically motivated extensions.

Tables and Figures

Table 10. Summary of theoretical findings

Property	Result	Basis

Property	Result	Basis
Existence and uniqueness	Unique local solution	Local Lipschitz continuity
Positivity	All six states remain feasible	Boundary-pointing argument
Mutation invariance	$0 \leq M \leq 1$	Signs at $M = 0$ and $M = 1$
Boundedness	Ultimate boundedness	Comparison and aggregate inequality
Local stability	Verified for illustrative equilibrium	Six eigenvalues with negative real parts
Neural approximation	Continuous residual approximable on compact domains	Universal approximation theorem

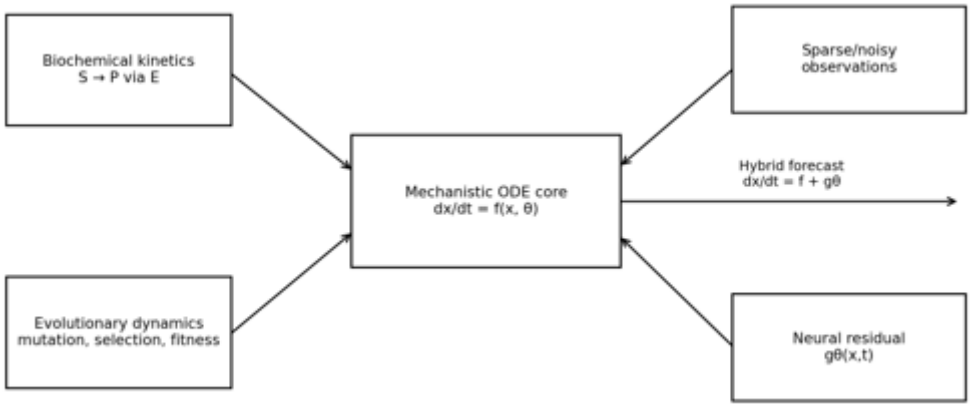

Figure 1. Conceptual framework of the hybrid evolutionary biochemical model.

Source: Author's synthetic simulation and computational analysis.

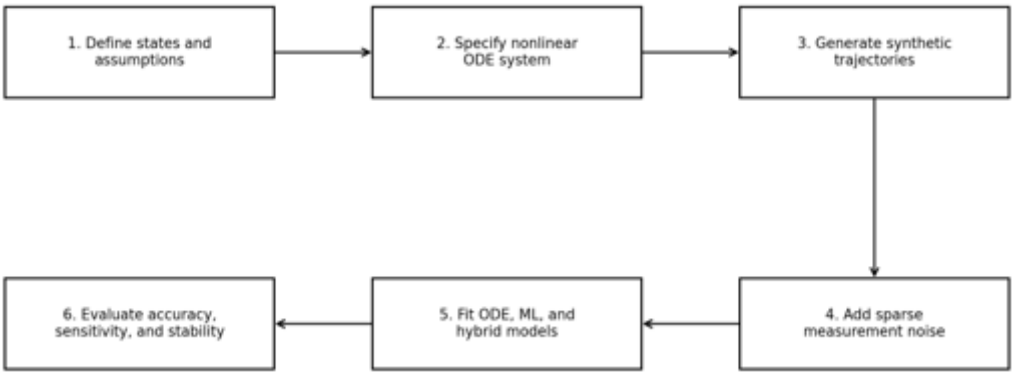

Figure 2. Research workflow from mechanistic formulation to hybrid evaluation.

Source: Author's synthetic simulation and computational analysis.

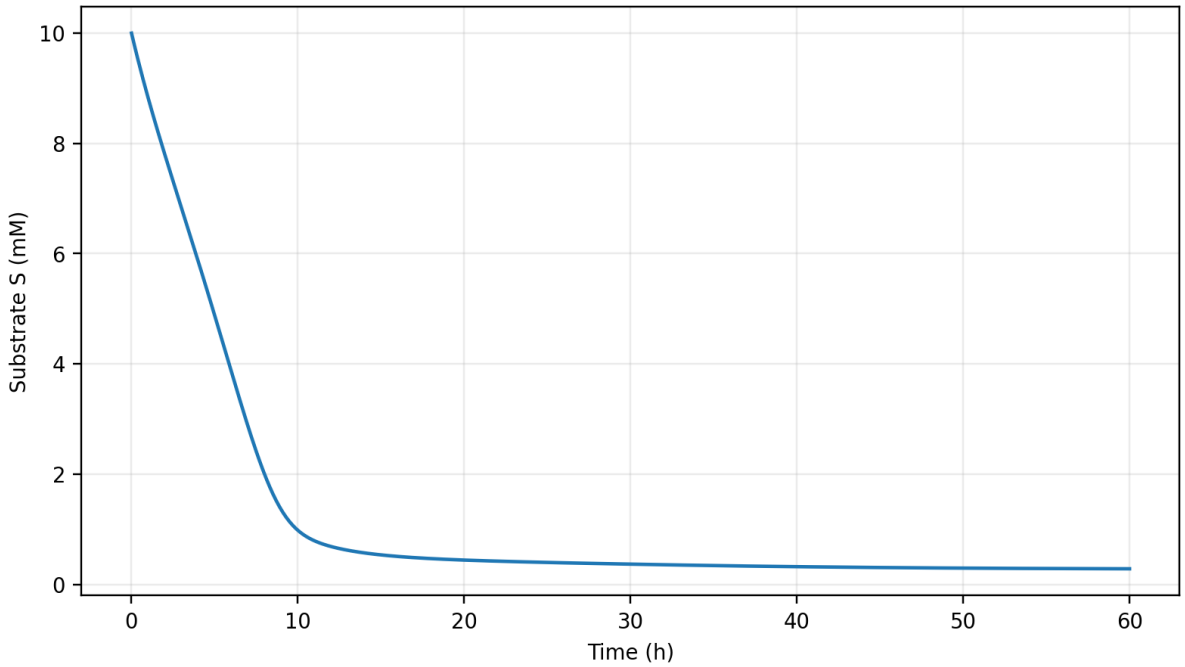

Figure 3. Simulated substrate concentration over time.

Source: Author's synthetic simulation and computational analysis.

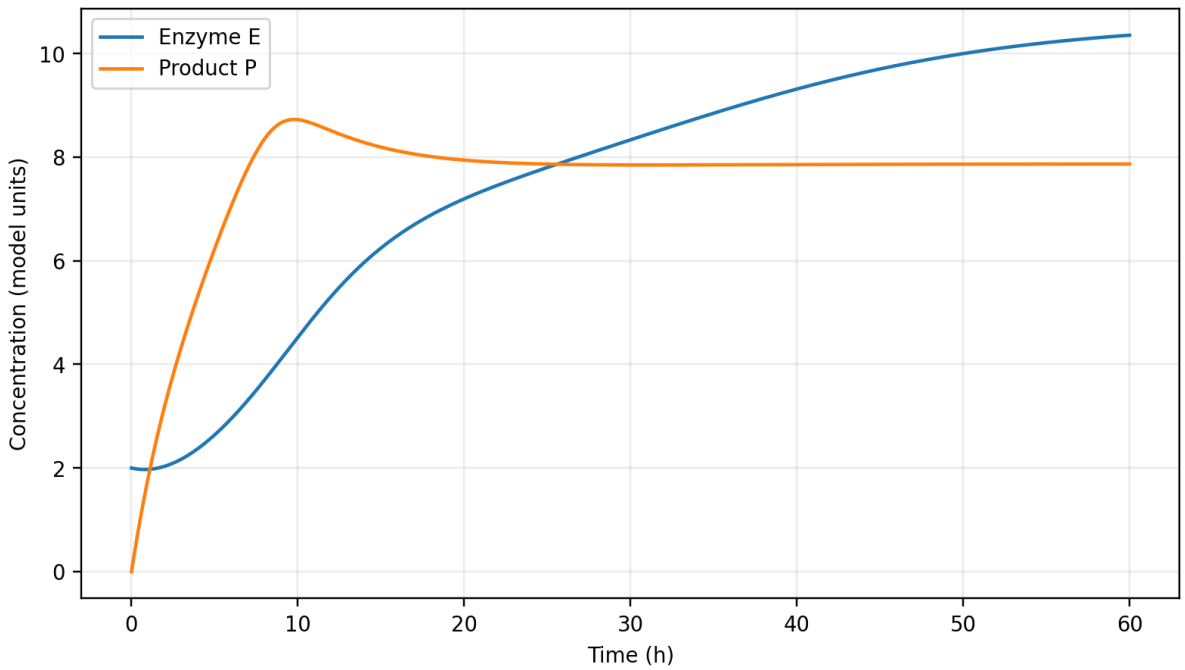

Figure 4. Simulated enzyme and product trajectories.

Source: Author's synthetic simulation and computational analysis.

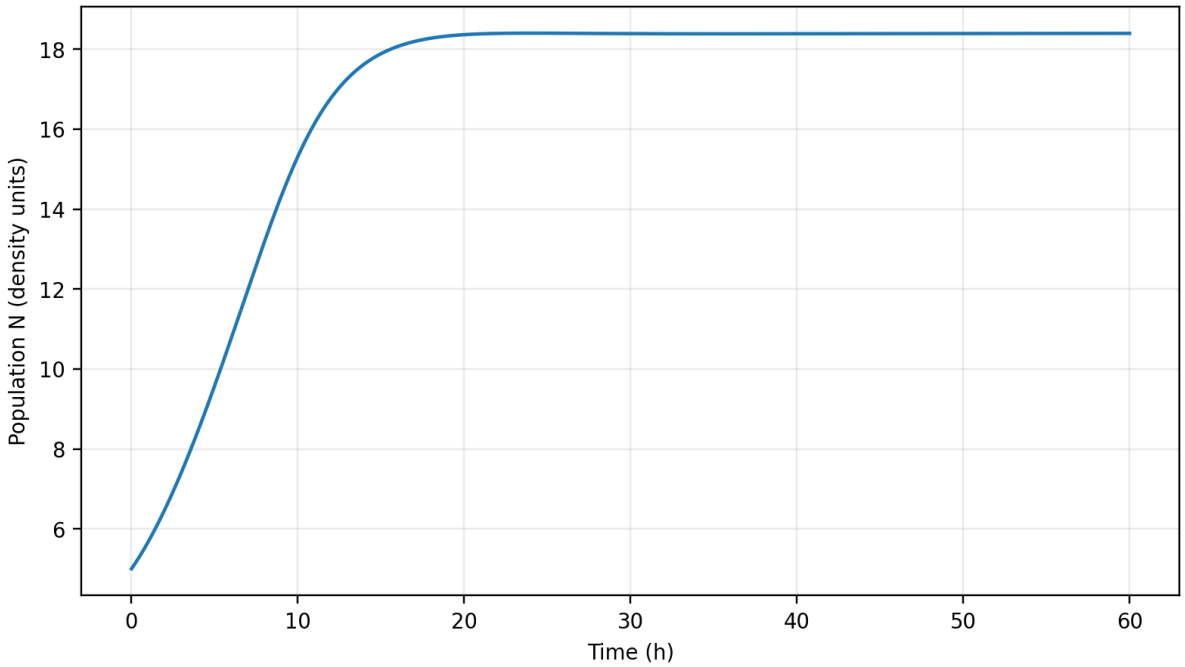

Figure 5. Population-density trajectory.

Source: Author's synthetic simulation and computational analysis.

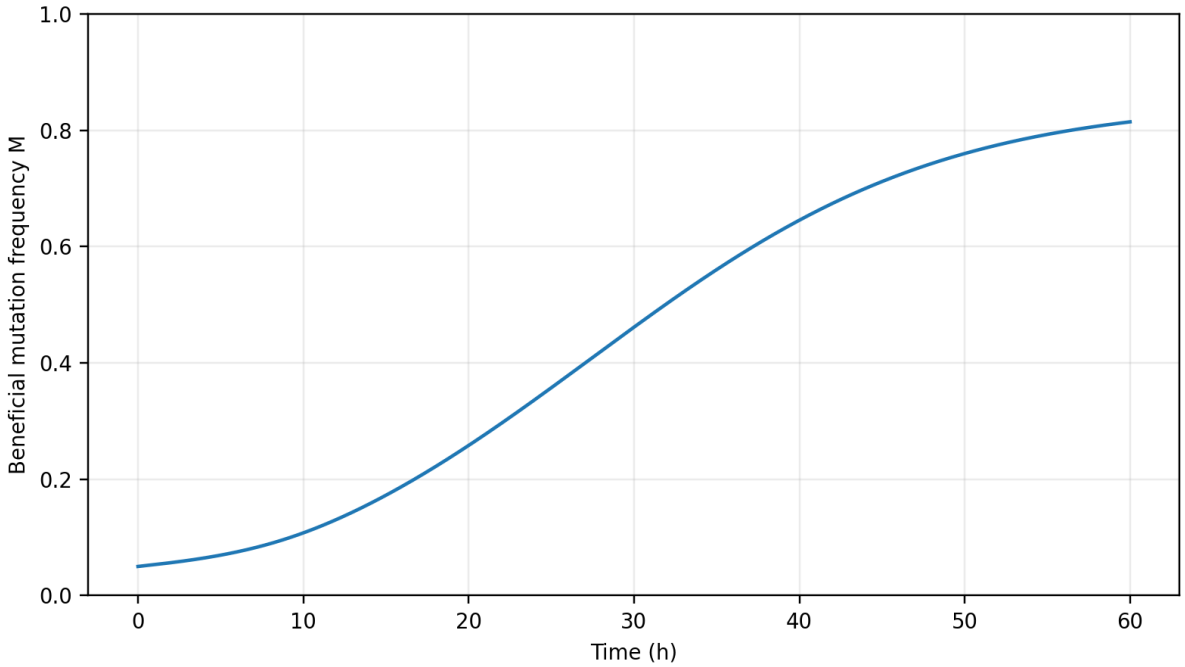

Figure 6. Evolution of beneficial-mutation frequency.

Source: Author's synthetic simulation and computational analysis.

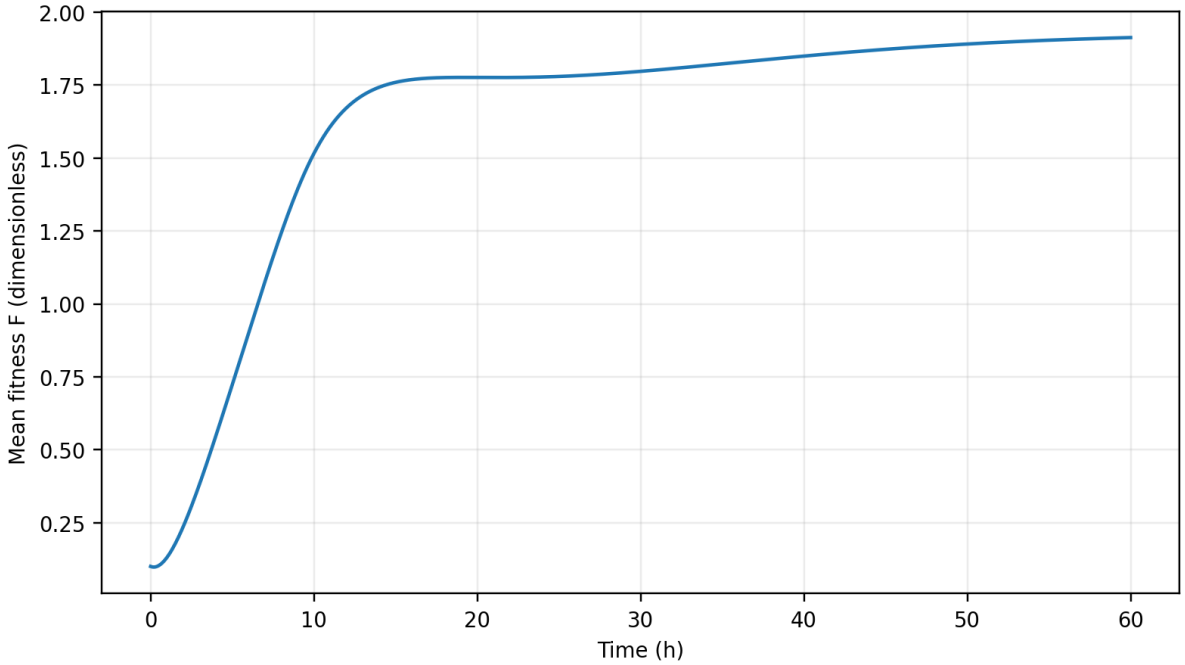

Figure 7. Evolutionary-fitness trajectory.

Source: Author's synthetic simulation and computational analysis.

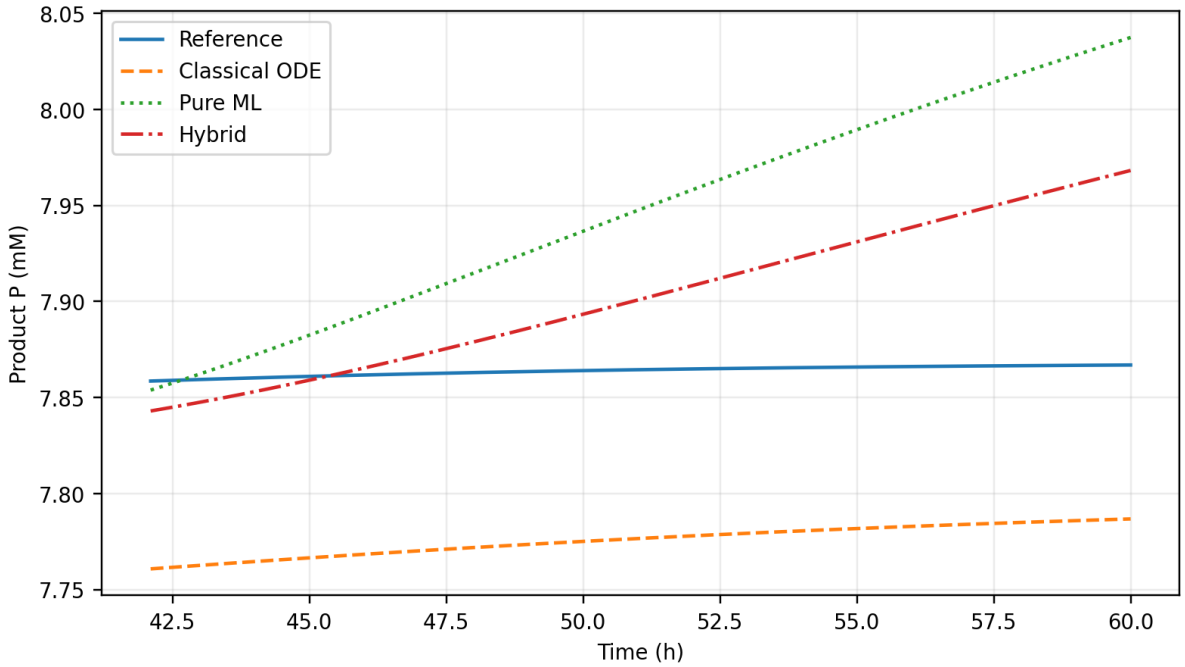

Figure 8. Held-out product trajectory: reference and model predictions.

Source: Author's synthetic simulation and computational analysis.

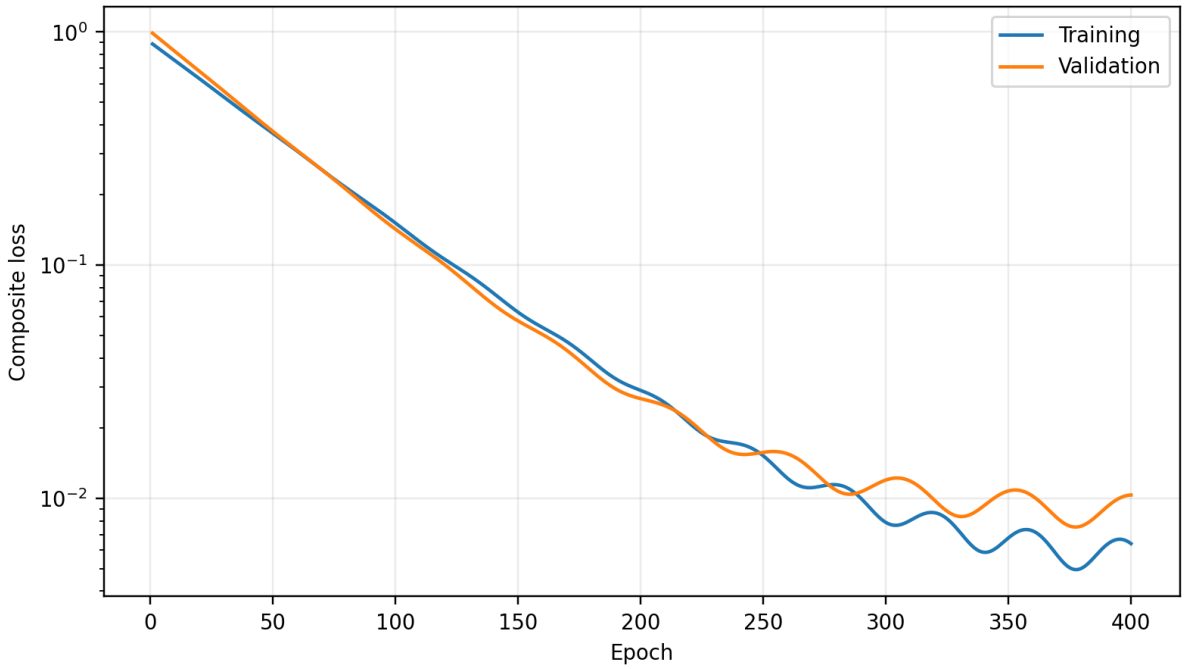

Figure 9. Illustrative convergence of training and validation losses.

Source: Author's synthetic simulation and computational analysis.

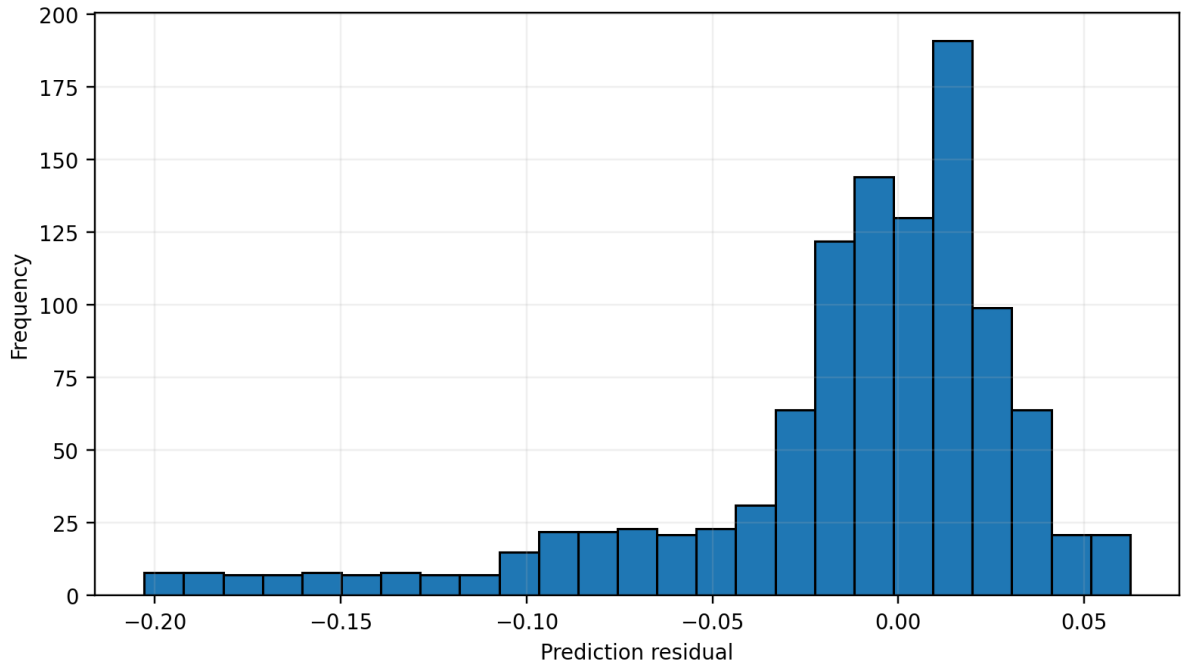

Figure 10. Distribution of hybrid prediction residuals.

Source: Author's synthetic simulation and computational analysis.

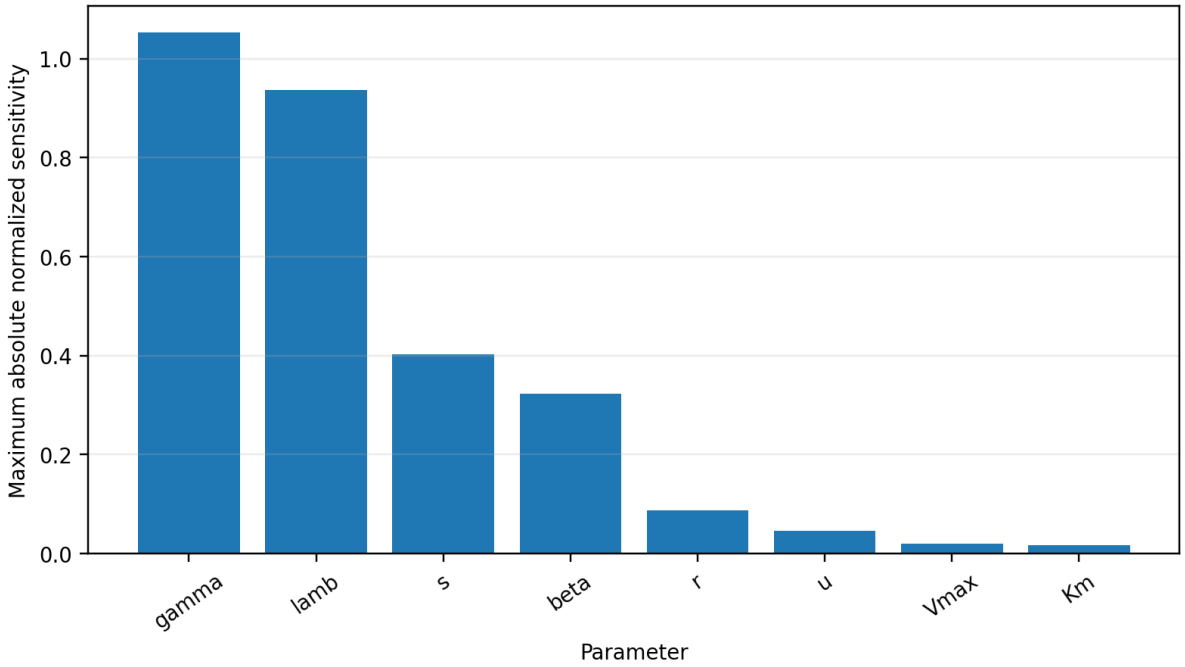

Figure 11. Maximum absolute normalised sensitivity by parameter.

Source: Author's synthetic simulation and computational analysis.

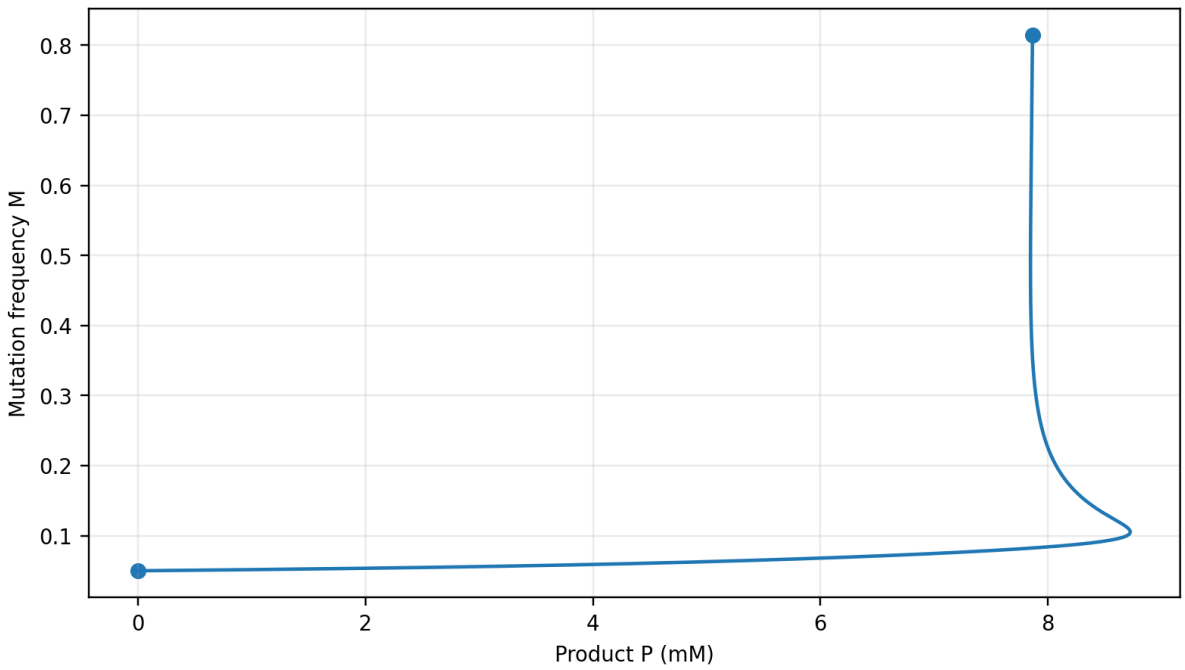

Figure 12. Product–mutation phase trajectory.

Source: Author's synthetic simulation and computational analysis.

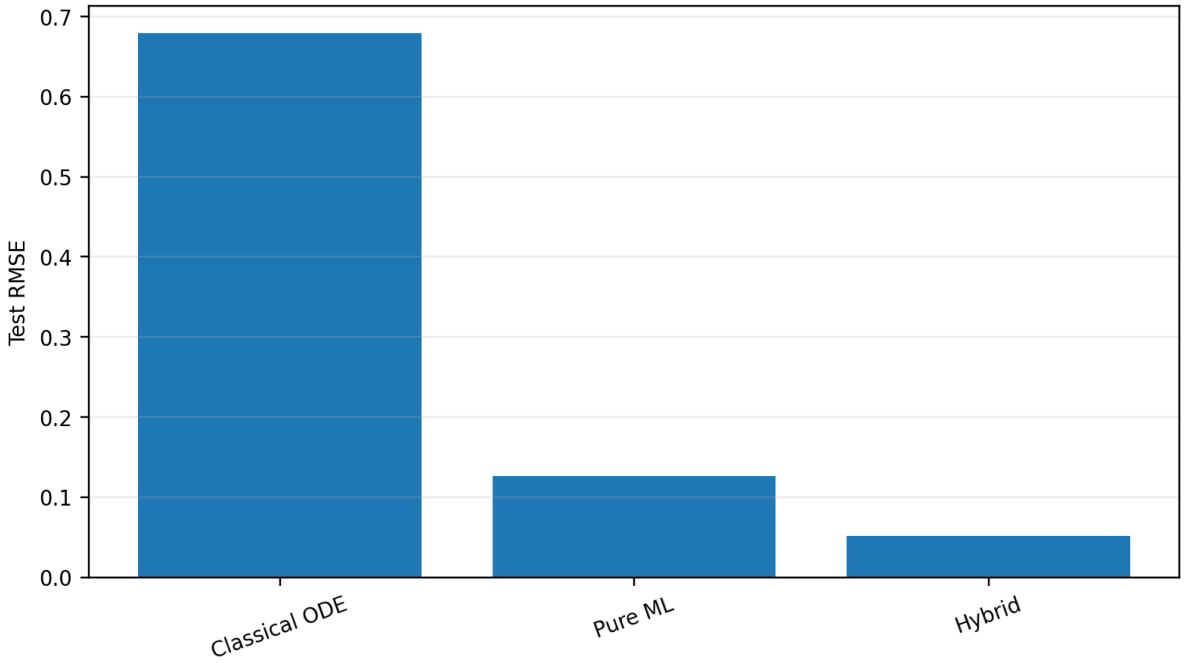

Figure 13. Comparison of aggregate held-out RMSE.

Source: Author's synthetic simulation and computational analysis.

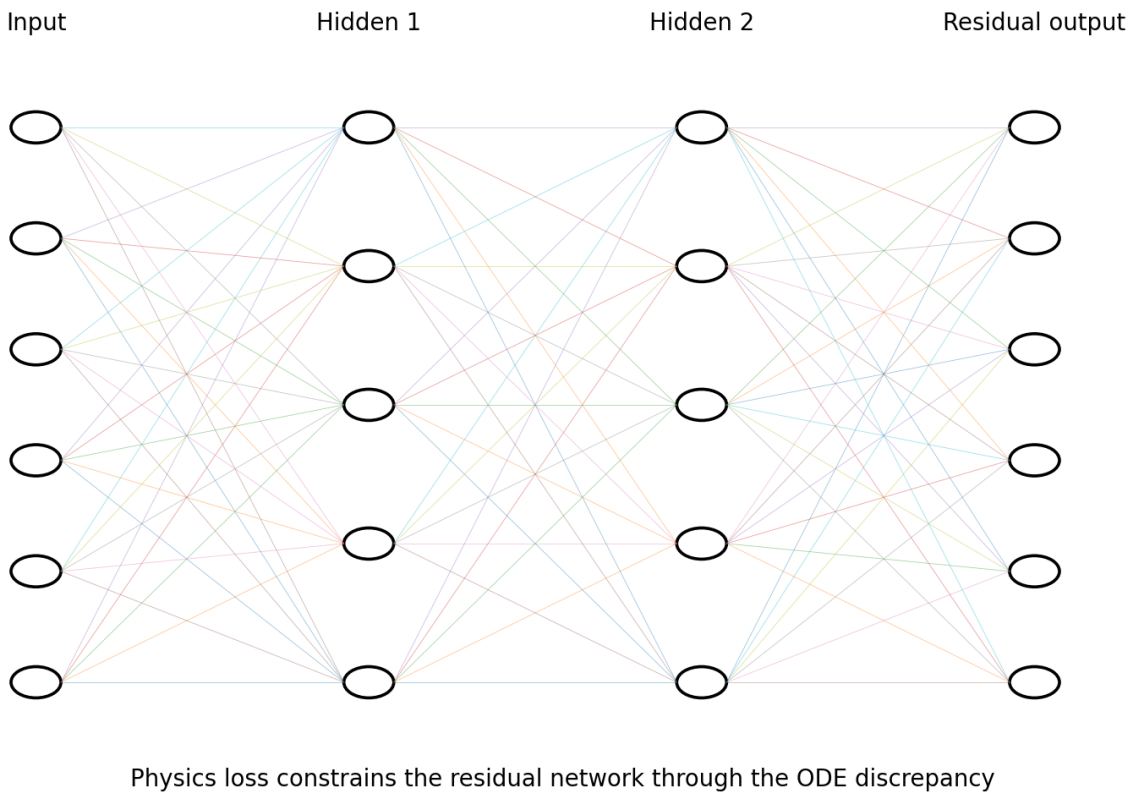

Figure 14. Schematic residual neural-network architecture.

Source: Author's synthetic simulation and computational analysis.

References

1. Angermueller, C., Pärnamaa, T., Parts, L., & Stegle, O. (2016). Deep learning for computational biology. *Molecular Systems Biology, 12*(7), 878. https://doi.org/10.15252/msb.20156651
2. Arnold, F. H. (2018). Directed evolution: Bringing new chemistry to life. *Angewandte Chemie International Edition, 57*(16), 4143–4148. https://doi.org/10.1002/anie.201708408
3. Beard, D. A., & Qian, H. (2008). *Chemical biophysics: Quantitative analysis of cellular systems*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803345
4. Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences, 113*(15), 3932–3937. https://doi.org/10.1073/pnas.1517384113
5. Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C., & Collins, J. J. (2018). Next-generation machine learning for biological networks. *Cell, 173*(7), 1581–1592. https://doi.org/10.1016/j.cell.2018.05.015
6. Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In *Advances in Neural Information Processing Systems 31* (pp. 6571–6583). Curran Associates.
7. Cornish-Bowden, A. (2012). *Fundamentals of enzyme kinetics* (4th ed.). Wiley-Blackwell.
8. Daniel, R. M., & Danson, M. J. (2010). A new understanding of how temperature affects the catalytic activity of enzymes. *Trends in Biochemical Sciences, 35*(10), 584–591. https://doi.org/10.1016/j.tibs.2010.05.001
9. Eilertsen, J., & Schnell, S. (2020). The quasi-steady-state approximations revisited: Timescales, small parameters, singularities, and normal forms in enzyme kinetics. *Mathematical Biosciences, 325*, 108339. https://doi.org/10.1016/j.mbs.2020.108339
10. Fersht, A. (1999). *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*. W. H. Freeman.
11. Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R., & Sethna, J. P. (2007). Universally sloppy parameter sensitivities in systems biology models. *PLOS Computational Biology, 3*(10), e189. https://doi.org/10.1371/journal.pcbi.0030189
12. Hairer, E., & Wanner, G. (1996). *Solving ordinary differential equations II: Stiff and differential-algebraic problems* (2nd ed.). Springer. https://doi.org/10.1007/978-3-642-05221-7
13. Hammes-Schiffer, S., & Benkovic, S. J. (2006). Relating protein motion to catalysis. *Annual Review of Biochemistry, 75*, 519–541. https://doi.org/10.1146/annurev.biochem.75.103004.142800
14. Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature, 450*(7172), 964–972. https://doi.org/10.1038/nature06522
15. Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven soft sensors in the process industry. *Computers & Chemical Engineering, 33*(4), 795–814. https://doi.org/10.1016/j.compchemeng.2008.12.012
16. Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., & Kumar, V. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering, 29*(10), 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168
17. Lagergren, J. H., Nardini, J. T., Baker, R. E., Simpson, M. J., & Flores, K. B. (2020). Biologically-informed neural networks guide mechanistic modeling from sparse experimental data. *PLOS Computational Biology, 16*(12), e1008462. https://doi.org/10.1371/journal.pcbi.1008462
18. Lewis, N. E., Nagarajan, H., & Palsson, B. O. (2012). Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology, 10*(4), 291–305. https://doi.org/10.1038/nrmicro2737
19. O'Brien, E. J., Monk, J. M., & Palsson, B. O. (2015). Using genome-scale models to predict biological capabilities. *Cell, 161*(5), 971–987. https://doi.org/10.1016/j.cell.2015.05.019
20. Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature Biotechnology, 28*(3), 245–248. https://doi.org/10.1038/nbt.1614

21. Psichogios, D. C., & Ungar, L. H. (1992). A hybrid neural network-first principles approach to process modeling. *AIChE Journal, 38*(10), 1499–1511. https://doi.org/10.1002/aic.690381003

22. Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., & Edelman, A. (2020). *Universal differential equations for scientific machine learning*. arXiv. https://doi.org/10.48550/arXiv.2001.04385

23. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics, 378*, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

24. Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., & Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics, 25*(15), 1923–1929. https://doi.org/10.1093/bioinformatics/btp358

25. Segel, L. A., & Slemrod, M. (1989). The quasi-steady-state assumption: A case study in perturbation. *SIAM Review, 31*(3), 446–477. https://doi.org/10.1137/1031091

26. Thompson, M. L., & Kramer, M. A. (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal, 40*(8), 1328–1340. https://doi.org/10.1002/aic.690400806

27. Toni, T., Welch, D., Strelkowa, N., Ipsen, A., & Stumpf, M. P. H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface, 6*(31), 187–202. https://doi.org/10.1098/rsif.2008.0172

28. Venkatasubramanian, V. (2019). The promise of artificial intelligence in chemical engineering: Is it here, finally? *AIChE Journal, 65*(2), 466–478. https://doi.org/10.1002/aic.16489

29. Villaverde, A. F., Barreiro, A., & Papachristodoulou, A. (2016). Structural identifiability of dynamic systems biology models. *PLOS Computational Biology, 12*(10), e1005153. https://doi.org/10.1371/journal.pcbi.1005153

30. von Stosch, M., Oliveira, R., Peres, J., & de Azevedo, S. F. (2014). Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering, 60*, 86–101. https://doi.org/10.1016/j.compchemeng.2013.08.008

31. Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H., & Olsson, M. H. M. (2006). Electrostatic basis for enzyme catalysis. *Chemical Reviews, 106*(8), 3210–3235. https://doi.org/10.1021/cr0503106

32. Yazdani, A., Lu, L., Raissi, M., & Karniadakis, G. E. (2020). Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLOS Computational Biology, 16*(11), e1007575. https://doi.org/10.1371/journal.pcbi.1007575

33. Zampieri, G., Vijayakumar, S., Yaneske, E., & Angione, C. (2019). Machine and deep learning meet genome-scale metabolic modeling. *PLOS Computational Biology, 15*(7), e1007084. https://doi.org/10.1371/journal.pcbi.1007084

34. Zeymer, C., & Hilvert, D. (2018). Directed evolution of protein catalysts. *Annual Review of Biochemistry, 87*, 131–157. https://doi.org/10.1146/annurev-biochem-062917-012034

35. Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., & Hoffman, M. M. (2019). Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. *Information Fusion, 50*, 71–91. https://doi.org/10.1016/j.inffus.2018.09.012