

RESEARCH ARTICLE

Open Access



Identifying effective evolutionary strategies-based protocol for uncovering reaction kinetic parameters under the effect of measurement noises

Hock Chuan Yeo¹, Varsheni Vijay³ and Kumar Selvarajoo^{1,2,3,4*} 

Abstract

Background The transition from explanative modeling of fitted data to the predictive modeling of unseen data for systems biology endeavors necessitates the effective recovery of reaction parameters. Yet, the relative efficacy of optimization algorithms in doing so remains under-studied, as to the specific reaction kinetics and the effect of measurement noises. To this end, we simulate the reactions of an artificial pathway using 4 kinetic formulations: generalized mass action (GMA), Michaelis–Menten, linear-logarithmic, and convenience kinetics. We then compare the effectiveness of 5 evolutionary algorithms (CMAES, DE, SRES, ISRES, G3PCX) for objective function optimization in kinetic parameter hyperspace to determine the corresponding estimated parameters.

Results We quickly dropped the DE algorithm due to its poor performance. Baring measurement noise, we find the CMAES algorithm to only require a fraction of the computational cost incurred by other EAs for both GMA and linear-logarithmic kinetics yet performing as well by other criteria. However, with increasing noise, SRES and ISRES perform more reliably for GMA kinetics, but at considerably higher computational cost. Conversely, G3PCX is among the most efficacious for estimating Michaelis–Menten parameters regardless of noise, while achieving numerous folds saving in computational cost. Cost aside, we find SRES to be versatily applicable across GMA, Michaelis–Menten, and linear-logarithmic kinetics, with good resilience to noise. Nonetheless, we could not identify the parameters of convenience kinetics using any algorithm.

Conclusions Altogether, we identify a protocol for predicting reaction parameters under marked measurement noise, as a step towards predictive modeling for systems biology endeavors.

Keywords Rate laws, Cell-free, Evolutionary algorithms, Parameter estimation, Predictive modeling, Systems biology, Michaelis–Menten, Mass action, Machine learning, Mechanistic modeling

*Correspondence:

Kumar Selvarajoo

kumar_selvarajoo@bii.a-star.edu.sg

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

In the book *A Brief History of Intelligence* [1], the author, Max S. Bennett, makes the important point of having a model that explicitly reproduces the fundamental features and interactions of a phenomenon (i.e., first-principled), to effectively learn the best course of action regarding it. For instance, the newly acquired perception of time and physical space by vertebrates has endowed a more advanced form of reinforcement learning, based on a model of the physical world. The new capability enables the schooling of sophisticated maneuvers, related to hunting and predator avoidance, as well as navigational skills, to enhance survival fitness in a harsh prehistoric world. Early mammals further acquire the capacity to flexibly simulate and hence discern alternative actions in their world model, providing them with a critical advantage over other animal classes, as to the efficiency and safety of learning. To note, the ancient competitive advantage of using a first-principle model for learning remains relevant in the current era of the fourth industrial revolution.

In an earlier commentary, we have set out the critical role of explicitly modeling the network and reaction kinetics for effective learning about biological systems [2]. For metabolic engineering [3], synthetic biology [4], or even precision drug endeavors [5], the holistic modeling of physico-chemical interactions among molecular components (systems biology [6]) is necessary for the reconstruction and elucidation of emergent behaviors and the identification of system-wide mechanisms and effects. Furthermore, the mathematical model of the change in component levels with time based on such interactions (i.e., rate laws or reaction kinetics), can be used to flexibly simulate, and predict component dynamics under broad conditions. From an application perspective, the mechanistic model also provides a structure grounded in first principles to help reduce sample size and datapoint requirement for training, the effect of biological and technical noises, data leakage, and batch effects [2]. The model also grounds predictions in molecular mechanism, thereby allowing for experimental verification, and enabling explainability. Furthermore, in recent years, there has been increasing interest in achieving the aspiration of digital twin [7] in various endeavors [3–5]. Regardless, however, its underlying principle compels the mirroring of a physical twin in its dynamic behavior, which is only truly compatible with an explicit model describing the underlying pertinent mechanisms. In all, the mechanistic model remains central to any machine learning/artificial intelligence approach for effective learning about biological networks.

While systems biology has begun its nascent transition from the explanatory modeling of fitted data to the

predictive modeling of unseen data [2, 8], there remain two broad objectives in the field:

1. The learning of a mechanistic model of the system for predicting unobserved dynamics, a shift from the earlier, more restricted objective of calibrating the parameters of a given model to best reproduce the fitted data [9];
2. The finding of an optimal set of system input variables, given the design objective(s) for the biological system. This is typically done by repeatedly simulating the system, according to values sampled from the input variable hyperspace. The design landscape is then separately modeled and searched to find the set of input variables corresponding to the optimal design objective.

In this regard, researchers are cognizant of the challenges in learning such a mechanistic model, both in determining the underlying rate laws and their parameter values, which are sensitive to biochemical contexts, such as pH, temperature, ion concentrations, and the intracellular microenvironment [10–15]. A suggested approach to bypass the difficulty is to model the individual reaction rates as a black box ML model [8] but this would result in a lack of clarity as to their biochemical basis. A more common heuristic approach is to first work on the selection or discernment of rate laws, followed by estimation of the associated parameters. Here, the point to note is that there remains room for improvement regarding the latter task in attaining the aspiration of predictive modeling for systems biology endeavors. The former task of determining the underlying form of the rate laws, and how to better integrate both tasks, presents further challenges that are not within the scope of the current study.

Currently, the various techniques available for parameter estimation are essentially of two main categories. The first consists of statistical (and sometimes tedious) procedures that have been prescribed for specific forms of reaction kinetics, including generalized mass action (GMA) [16–18], linear-logarithmic (Linlog) [19, 20], and convenience kinetics (CK) [21]. However, it is unclear what technique would then be easily applicable for the myriad combinations of such formulations and new ones that could be prescribed for the individual reactions. Certainly, costly ad hoc experiments [22] or deep learning methods [23–25] could be specified in specific circumstances but a standardized computational procedure would make the learning process more efficient and allow for automation. Thankfully, the predicament can be bypassed by rephrasing the problem generically as one of optimizing a nonlinear objective function in bounded parameter search space (nonlinear programming) [26].

What is only needed in this regard is an optimization algorithm that can traverse the landscape efficiently for sampling the objective function value. The rephrased problem thus trades the ill-defined task of prescribing parameter estimation procedure, by rate law formulation, with a more explicit one of finding an effective optimization algorithm.

The new task is, however, no walk in the park. Ideally, the algorithm must be capable of both exploring and exploiting diverse landscapes to find the global optimal among numerous local solutions. The biological landscape is characterized by terrains more than just numerous and steep optima. It may also be inconsistent because of technical and biological variabilities, and stochasticity, or be distorted by a non-representative or inadequate number of datapoints. The terrain may also interchange depending on the specific parameter dimension and thus exhibit sensitivity to small changes in the search space location. The terrain may further be sensitive to the input data (termed ill-conditioning), due to the prevalence of biochemical feedback/forward loops, for instance, and further aggravated by variability in experimental measurements. For these reasons, the optimization algorithm must minimally be robust and self-adapt quickly to the local terrain in carrying out its task.

Historically, researchers have recognized the advantage of using stochastic algorithms [9] over deterministic ones (e.g., gradient-based [27] and direct search methods [28, 29]) for the search and optimization of a multi-modal landscape. Among the more effective stochastic methods, evolutionary strategies (ESs) have been consistently reported to be more robust and efficient [30–34] than either simulated annealing (SA) [35] or genetic algorithms (GA) [36]. This may particularly be due to their capacity for self-adapting their strategy parameters, while having all the properties necessary for global optimization [32]. Their outperformance in problems having continuous search space has also been attributed to their specific design for them [26]. On the other hand, SA and GA are originally intended to solve combinatorial problems based on discrete variables [37, 38] with a simplified search space. Indeed, the efficacy of ESs has caught the attention of systems biologists. For example, they have been used for the explanative modeling of the circadian clock [39], development and pattern formation [40–43], and iron metabolism [44]. More broadly, they have been employed to help discern network mechanisms [45], reverse engineer pathways [46, 47], and infer signaling dynamics [48]. In all, the specific application of ESs has been largely based on earlier referenced works and/or through trials and errors. While such qualitative objectives can be adequately met by the ad hoc selection of

optimization algorithms, the aspiration of predictive modeling will require a more thorough characterization of promising algorithms in their ability to recover the underlying reaction parameters. For many real-world applications, it is as critical that the predictions are made on time [7], thereby placing a premium on the efficiency of the algorithm.

Using both criteria, we ask in this study if there is any evolutionary algorithm (EA) that can efficiently traverse the landscape of major reaction kinetics to find the globally optimal solution. Given the hard nature of the question, we also investigate if there is any value-add in prescribing specific algorithm for the individual formulations of reaction kinetics. In doing so, we recognize there could be no algorithm that outperforms others under all contexts, based on our metrics (no free lunch theorem [49]). Instead, we are searching for acceptable tradeoffs [50, 51] among our key criteria. To this end, we screen the effectiveness of 5 widely used and available EAs in estimating the kinetic parameters of an artificial *in silico* pathway, according to 4 canonical formulations. To mimic an actual application, we otherwise replicate the structure of the mevalonate pathway for limonene production. As a result, we find specific algorithms to be effective under empirical-like conditions, after applying key techniques and learning points to mitigate the effects of measurement errors. Our resulting protocol thus allows for the more accurate modeling of biological pathway dynamics, as a step towards realizing the aspiration of predictive modeling for real world endeavors.

Results

Screening evolutionary algorithms for parameter estimations of canonical reaction kinetics

We screen the effectiveness of 5 widely used and available EAs (overview provided in the “Methods” section):

- i. Differential evolution (DE),
- ii. Stochastic Ranking Evolutionary Strategy (SRES),
- iii. “Improved SRES” (ISRES),
- iv. “Covariance Matrix Adaptation Evolution Strategy” (CMAES), and
- v. “Generalized generation gap model with parent-centric combination” (G3PCX)

for 4 kinetic formulations (Additional file 1) of an artificial pathway (Fig. 1A and B):

- i. GMA [52],
- ii. Michaelis–Menten (MM) [53],
- iii. Linlog [19], and
- iv. CK [21]

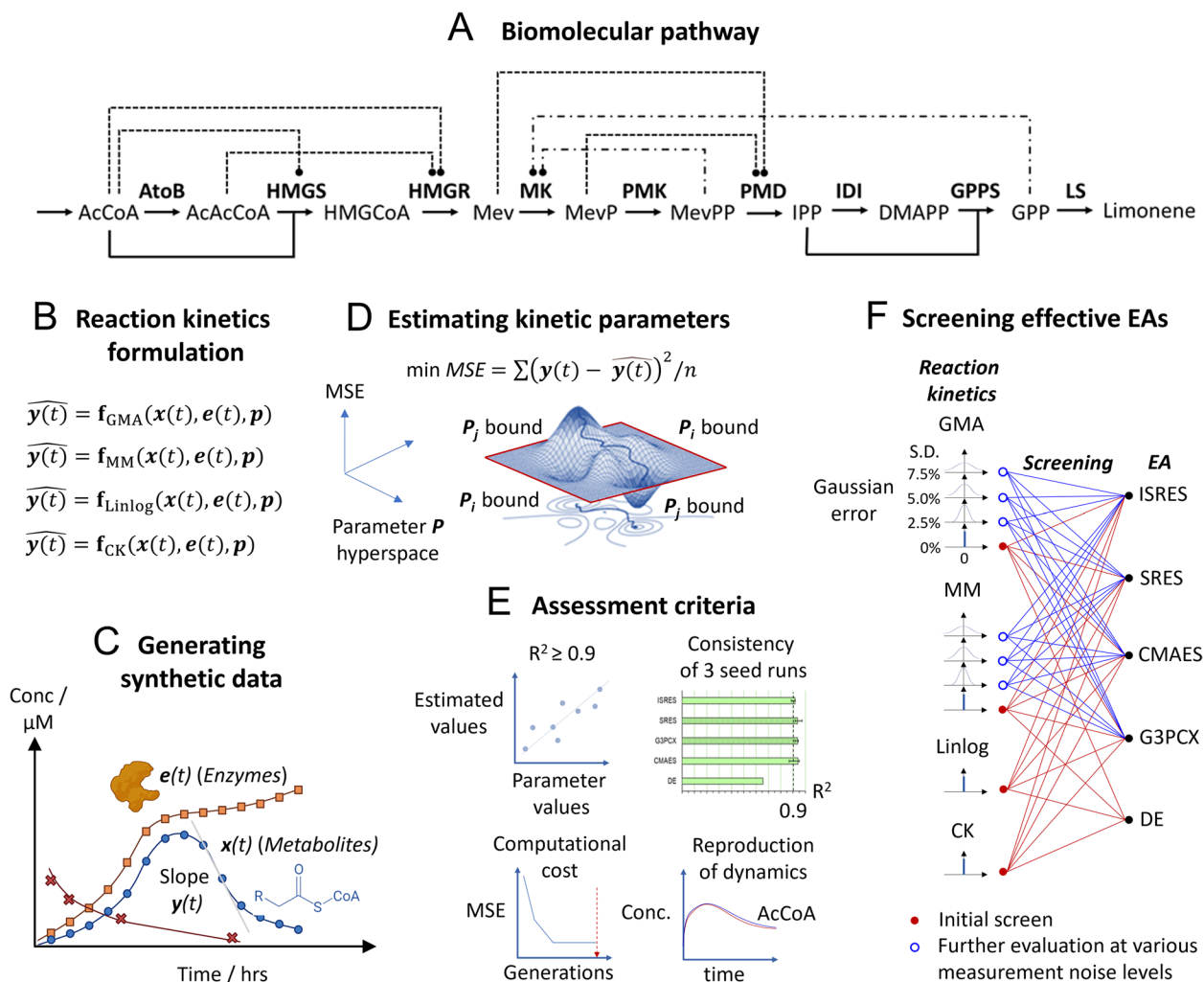


Fig. 1 Overview of the approach used in the study. **A** Artificial pathway used in the study by replicating the topology of the mevalonate pathway for limonene synthesis. Arrows refer to reaction. Dash lines represent feed-forward inhibition, whereas mixed dash and dot lines indicate feedback inhibition. **B** Formulation of separate reaction kinetics for the pathway [generalized mass action (GMA), Michaelis–Menten (MM), linear-logarithmic (Linlog), and convenience kinetics (CK)]. **C** Time-series data for metabolite concentrations $x(t)$ is generated by using the formulated reaction kinetics to simulate reaction progressions through time. The empirical net reaction rate for each metabolite at a given timepoint is then interpreted as the slope of its concentration in time. Separately, the dynamic concentration of enzyme participants $e(t)$ is each determined by a specific hyperbolic Hill function. **D** The mean square error (MSE) between empirical net reaction rates and the predicted values based on the kinetic formulation is then minimized as an objective function in kinetic parameter hyperspace. The corresponding parameter values are then reported as an estimation. **E** The quality of the parameter estimations is then assessed using 4 criteria/considerations: if the coefficient-of-determinant is greater than or equal to 0.9 and is consistently so for three different seed runs, the computational cost (by using the number of generations required for optimization as a proxy), and the degree of reproducibility of the underlying reaction dynamics with the estimated parameters. **F** Five widely available evolutionary algorithms (EAs) are then screened for their capacity to estimate the parameters of the different kinetic formulations. The initial screening is done using datapoints at 15 min intervals with no measurement noise. Further evaluations are conducted for selected EAs in estimating GMA and MM parameters at increasing measurement noise and datapoint spacing. The effect of taking parameter averages based on different seed solutions as well as data augmentation is also evaluated. *Metabolites* AcCoA, acetyl-coenzyme A; AcAcCoA, acetoacetyl-coenzyme A; HMGCōA, 3-hydroxy-3-methylglutaryl-coenzyme A; Mev, mevalonate; MevP, phosphomevalonate; MevPP, diphosphomevalonate; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate *Enzymatic reactions* AtoB, acetoacetyl-CoA thiolase; HMGS, hydroxymethylglutaryl-CoA Synthase, HMGR, 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; PMD, phosphomevalonate decarboxylase; IDI, isopentenyl-pyrophosphate delta-isomerase; GPPS, geranyl pyrophosphate synthase; LS, limonene synthase

Synthetic data of metabolite dynamics are generated based on each of the reaction kinetics (Fig. 1C), pre-processed, and then fitted back to the same formulation (Fig. 1D) by minimizing the mean square error (MSE) between predicted and empirical net reaction rates using all candidate EAs.

A total of 4 requirements/considerations are used for evaluating the corresponding parameter solution (Fig. 1E):

- i. A high quality of parameter estimations, as reflected by the coefficient of determination (R^2) being greater than 0.9,
- ii. Consistency in meeting the above requirement based on three different seed solutions,
- iii. The number of generations required for the EA to faithfully reach a plateau-state in terms of minimized MSE (figure conservatively rounded up to one significant figure based on the largest of three seeds), and
- iv. The ability to reproduce the underlying metabolite dynamics.

For criterion iii, we use the number of generations as a convenient proxy for the number of cost function evaluations (which is computationally expensive), as the same population size is used for all algorithms. To discern the effect of kinetic formulation and data quality on the performance of EAs, we first use “perfect” data (Fig. 1F), that is, with no measurement noise and a datapoint spacing that is 15 min apart. Subsequently, we introduce noise and increase the spacing to “strain-test” the algorithms. This strategy also allows us to efficiently filter out poor performers early on, before progressing to their evaluation under more challenging contexts.

Evolutionary algorithms differ widely in their effectiveness for parameter estimations

With perfect data, SRES stands out in consistently surpassing our R^2 threshold for all seed-instances of three out of the four reaction kinetics-of-interest. Its median R^2 values for GMA, Linlog, and MM kinetics are 0.993, 0.991, and 0.936, respectively (Fig. 2, left panel). We note that while SRES was previously reported to produce accurate parameter estimations for a MM model [26], its performance with regard to other reaction kinetics has not been studied.

In the same way, three other EAs exceed the R^2 threshold for just two classes of reaction kinetics. Specifically, CMAES and ISRES perform similarly well for the same formulations, achieving a respective median R^2 value of 1.00 and 0.996 for GMA kinetics, and 0.994 and 0.993 likewise for Linlog kinetics. We also find G3PCX to be

effective for Linlog formulation (median R^2 value of 0.99), but it is also as good as SRES in estimating MM parameters (0.936).

Conversely, DE is ineffective for all classes of investigated reaction kinetics, as its R^2 values mostly fall below 0.9, except for one seed-instance of GMA. Apparently, this is because of the inability of the algorithm to reach a plateau stage in terms of the minimized cost function (MSE-of-fit) (Additional file 1: Figs. S1–S4), within the budgeted number of generations (Table 1). In contrast, all other EAs can do so for the reaction kinetics that they did well for (as described above). As expected, its MSEs of fit are noticeably higher than those of other EAs for GMA, MM, and Linlog kinetics (Fig. 2, right panel). Furthermore, in contrast, the other EAs can achieve MSEs of fit within tighter ranges (implying consistency among seed-instances), which are also largely matching with each other (consistency among EAs). For instance, this is true for all other EAs in the case of MM and Linlog kinetics, while (I)SRES and CMAES algorithms are similarly so for GMA kinetics. One learning point from the juxtapositions is that we confirm that the MSE-of-fit, as to its magnitude and consistency among seed-instances, can be helpful in filtering out poor performers. The metric is especially valuable in practical settings, since the R^2 value cannot be assessed without knowledge of the true solution.

Effective evolutionary algorithms in the absence of measurement noise

Given the outright poor performance of DE, we dropped it for further evaluation; from this point onward, we use “All EAs” to refer to all other EAs except for DE. We then evaluate the efficacy of the estimated parameters in regenerating metabolite dynamics through simulation. For GMA kinetics, we find CMAES, SRES, and ISRES to closely reproduce all the metabolite dynamics (Additional file 1: Fig. S5A), which is consistent with their high R^2 value of parameter estimations (Fig. 2, left panel); those of CMAES are in fact virtually indistinguishable from the actual one. Also unsurprisingly given the poor R^2 performance, G3PCX could not qualitatively reproduce the dynamics of all metabolites for GMA kinetics. For example, we observe a time lag of 4–6 h for the concentration profile of the 5 most downstream metabolites. However, unexpectedly for MM kinetics, all shortlisted EAs can regenerate the dynamics almost identically (Additional file 1: Fig. S5B), despite the R^2 values being appreciably and broadly lower than in the case of GMA kinetics. This suggests the relative insensitivity of MM dynamics to the underlying parameters. We also note the same excellent recovery of Linlog-based dynamics for all EAs (Additional file 1: Fig. S6), which dovetails

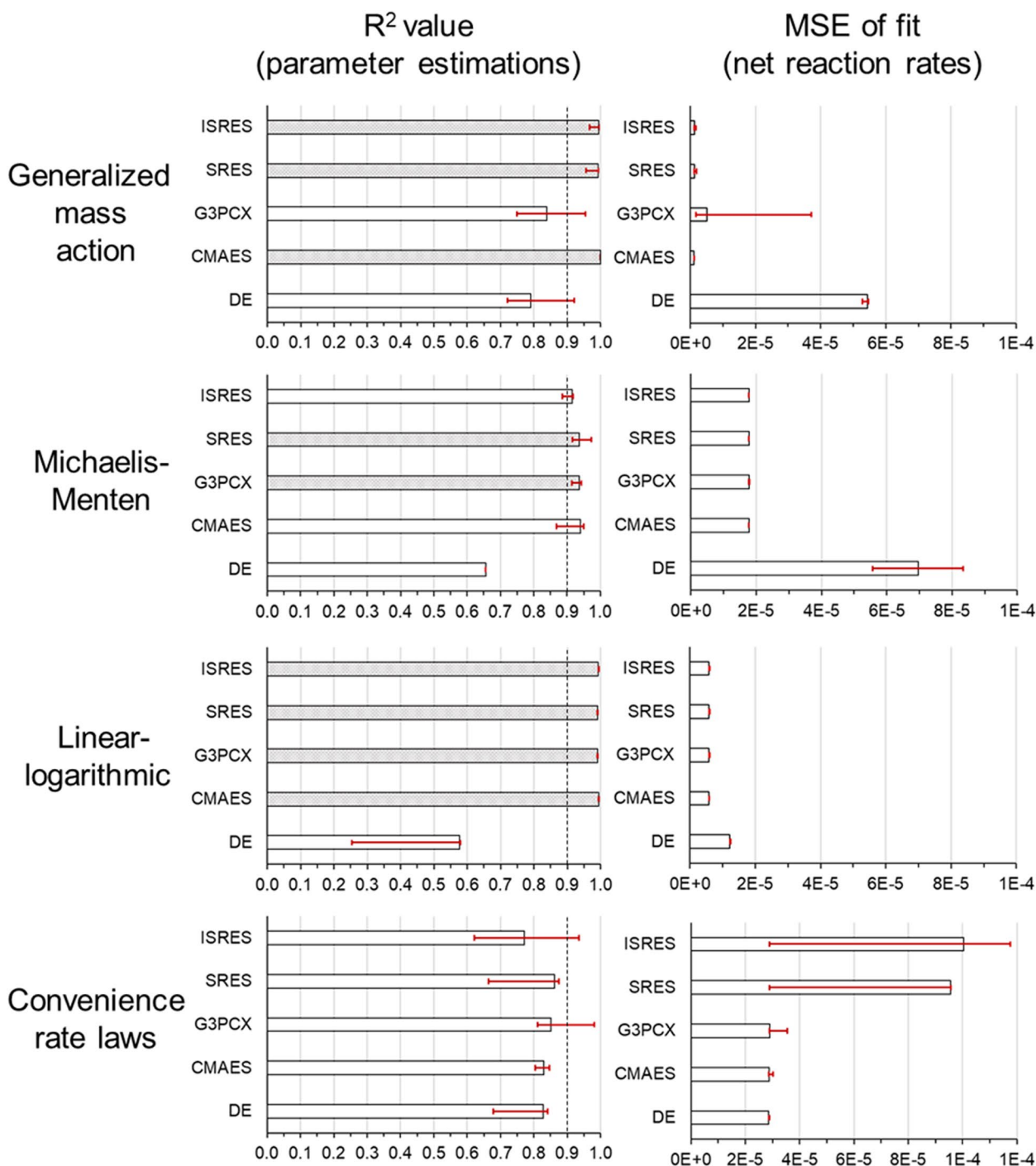


Fig. 2 Screening of evolutionary algorithms for parameter estimations of canonical rate law models. Data with no measurement error is used for screening. Triplicate estimates are conducted using different initial seed solutions. The bars represent the median R^2 or MSE-of-fit value, while the lower and upper bounds (in red) mark the smallest and largest values of the three initial seed solutions, respectively (Additional file 5: Data S4). Note that the computation of the standard deviation as a measure of the spread both above and below the median is inappropriate, as the distribution may not be symmetrical. Vertical dash lines mark the threshold value of 0.9 for R^2 -value metric (R -value of 0.95). A gray bar indicates that the values from all three seeds have exceeded the threshold. Note that only one R^2 -value has been reported in fitting the Michaelis–Menten model using the differential evolution (DE) algorithm; two others with negative R -value have not been shown

Table 1 Budgeted number of generations for parameter estimations of various kinetic models

Kinetic model	No. of generations
Generalized mass action	1×10^5
Michaelis–Menten	1×10^4
Linlog	1×10^5
Convenience	1×10^5

with their good metric performance ($R^2 > 0.99$). The corresponding MSE-of-fit is also found to be effectively minimized within 2000–50,000 generations (Additional file 1: Fig. S3, Additional file 2: Data S1).

When we further consider computational cost, CMAES stands out in requiring no more than 5000 and 2000 generations for GMA and Linlog kinetics, respectively (Additional file 2: Data S1). This corresponds to 1/6 and 1/20 of the generation number required for the next most cost-efficient EA (ISRES: 30,000 and 40,000, respectively) that also meets our other considerations. For MM kinetics, G3PCX is likewise most efficient in requiring only 1000 generations, which is $\frac{1}{2}$ of the other EA (SRES) that fulfills our considerations.

For all that, not a single EA assessed could consistently hit our R^2 threshold seed-wise for the CK formulation, despite all of them attaining the plateau stage during cost function minimization (Additional file 1: Fig. S4). The plateaued values further vary noticeably among EAs, suggesting a challenging cost function landscape with numerous local minima. More studies are needed to find an effective optimization method for CK-based models.

SRES and ISRES are most effective for GMA kinetics in the presence of measurement noise

For the assessment of EAs under experiment-like conditions, we add simulated errors to our metabolite and enzyme “measurements” before repeating our analysis. Again, DE is not considered due to its poor performance in earlier screening. We also exclude Linlog parameter estimations, as the kinetic formulation is a heuristic modeling approach, with no empirical profiling of the rate law to speak of. Given that random errors are expected to increase both the ruggedness and variability of the optimization landscape among replicates, we generate for each initial seed solution a different triplicate dataset to include the effects of measurement noise on parameter predictions. We also take the seed-average of the parameter predictions and compute its R^2 value to derive a measure of the *expected* quality of parameter estimation.

Contrary to our earlier finding in the absence of noise, CMAES becomes the worst performer for GMA kinetics upon the introduction of noise. Although its expected R^2

values (reported as filled circles in Fig. 3A) and correlation plots at various noise levels (Fig. 3B) are not very different from those of other EAs, it clearly exhibits greater variability in the metric value among seed-instances (open circles in Fig. 3A), compared to SRES/ISRES algorithms. CMAES likewise exhibits larger variation in the MSE-of-fit compared to other EAs, in the presence of noise (Additional file 1: Fig. S7, left column). When we test the efficacy of the resulting estimated parameters to recover the mass action dynamics under measurement noise, CMAES is unable to recapture the qualitative profile of the 4 most downstream metabolites, even at a noise level of just 2.5% (Additional file 1: Fig. S8A) and 5% standard deviation (Additional file 1: Fig. S8B). At the 7.5% level, the dynamics are off track for 7/10 metabolites (Fig. 3C), with an early reaction substrate (AcAcCoA) falsely predicted to be exhausted.

While G3PCX similarly has larger MSE-of-fit and R^2 variability compared to SRES/ISRES algorithms, this is also the case in the absence of noise. Thus, unsurprisingly, G3PCX cannot recover the qualitative profile of multiple downstream metabolites even with no noise, underscoring its unsuitability for estimating GMA parameters. Instead, SRES/ISRES perform the best for GMA kinetics, generating similar qualitative dynamics at 5% standard deviation or less, albeit at a cost of 100,000 generations.

With measurement noise, G3PCX remains most effective for recovering MM parameters

We similarly investigate the efficacy of EAs for estimating MM parameters when measurement errors are present. In this regard, the performance of all algorithms based on expected R^2 value remains relatively stable at increasing measurement noise (Fig. 4A), compared to the estimation of GMA parameters. To illustrate, the metric stays between 0.9 and 0.96 (filled ovals) for all EAs even at 7.5% noise level, whereas the same figure ranges between 0.78 and 1 for GMA parameter estimations. To view in another way, there is no discernable drop in the expected R^2 value for both ISRES and G3PCX algorithms with increasing noise, while those of CMAES and SRES decrease minorly (≤ 0.06). In contrast, we observe a decrease of 0.21 in the metric value for three EAs (CMAES, SRES, ISRES) with increasing noise, for the case of GMA kinetics. The similar correlation plots between the mean estimated parameters and actual values at increasing measurement noise further illuminate the more insensitive nature of MM parameter estimations to measurement errors (Fig. 4B).

The EAs also recover MM dynamics more reliably than that of GMA dynamics, as demonstrated by the largely unchanging metabolite profiles that are generated for

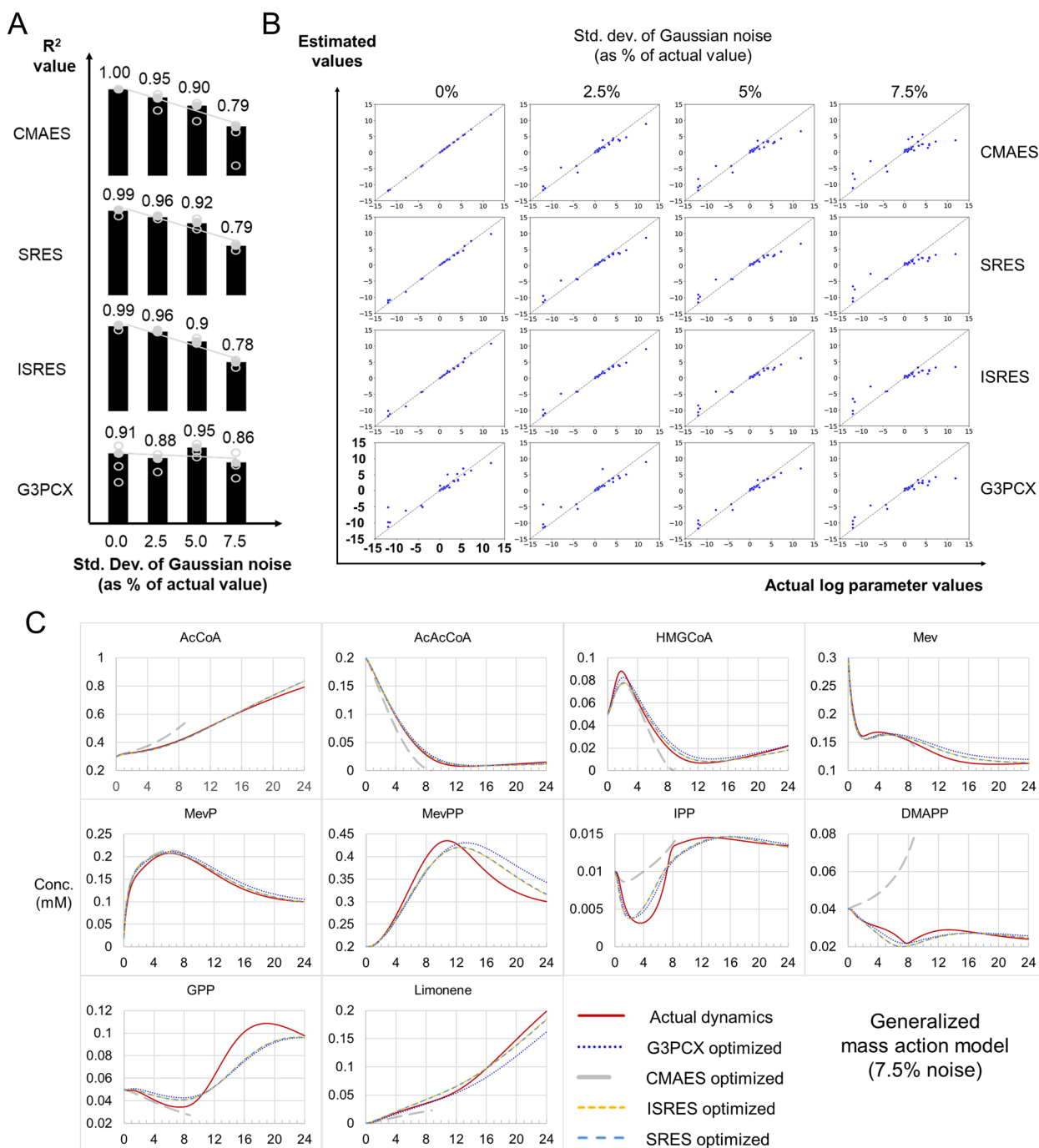


Fig. 3 Effectiveness of selected evolutionary algorithms for estimating parameters of generalized mass action model under simulated conditions of increasing measurement errors. **A** Quality of estimations at increasing measurement errors as quantified by R^2 -value. Each of the three open ovals represents a value derived from a distinct triplicate dataset for both metabolites and enzymes and a random seed solution, whereas a filled oval denotes the R^2 -value of their mean parameter values. The trend of the filled ovals is shown as a straight line. **B** Correlation plots of estimated parameter values (mean from the three initial seed solutions) against actual ones at increasing measurement errors for each algorithm. **C** Simulated metabolite dynamics based on the estimated parameters. In this case, the errors are sampled from a normal distribution centered on zero and a standard deviation equivalent to 7.5% of the underlying metabolite or enzyme concentration. Refer to Additional file 6: Data S5 for datapoints presented in the figure

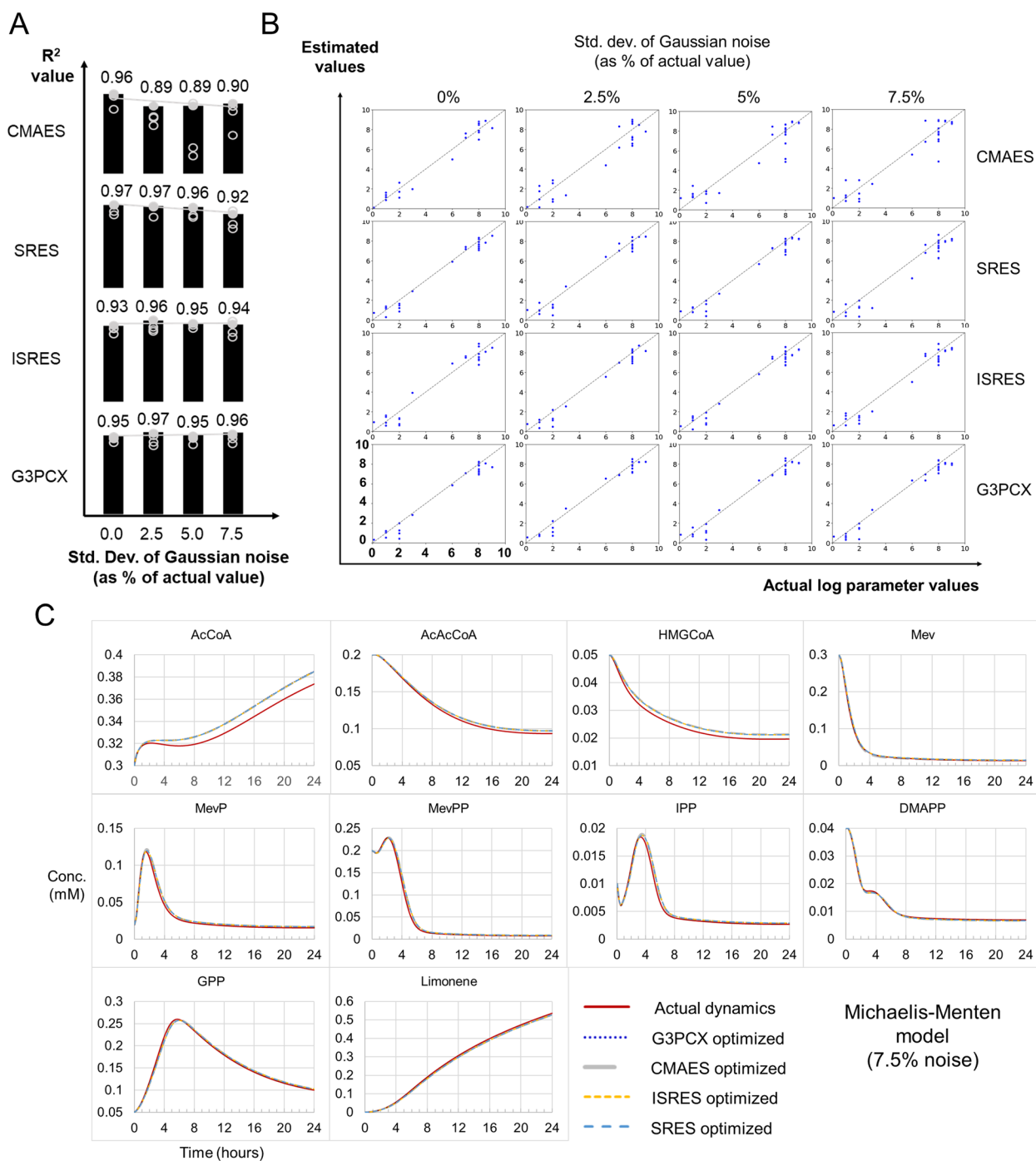


Fig. 4 Effectiveness of selected evolutionary algorithms for estimating parameters of Michaelis–Menten model under simulated conditions of increasing measurement noise. Refer to Fig. 3 for the legend of corresponding sub-figures **A–C** and Additional file 7: Data S6 for datapoints presented in the figure

MM kinetics with increasing noise for all EAs. Specifically, the profiles of all EAs are virtually identical to the actual dynamics at 0% (Additional file 1: Fig. S5B) and 2.5% noise levels (Additional file 1: Fig. S9A), while showing marginal differences for the three most upstream

metabolites at 5% (Additional file 1: Fig. S9B) and 7.5% noise levels (Fig. 4C). In this regard, all EAs perform similarly well in recovering MM parameters and dynamics when measurement noise is present. However, after taking computational cost into account, G3PCX remains

the most effective, by necessitating only 2000 generations to do so (Additional file 1: Fig. S10). The figure is 66% that of the next most efficient EA (SRES: 3000 in Additional file 2: Data S1) that satisfactorily meets all our other considerations. It would be interesting to verify for a variety of network topologies, if the measurement noise has, indeed, less effect on recovering the parameters and dynamics of a MM formulation, compared to its GMA equivalence.

CMAES is most sensitive to measurement noise

In addition to the case for GMA kinetics, we observe CMAES to show increasing inconsistency in MM parameter estimations with more measurement noise, as reflected by the larger variability in R^2 value (open ovals in Fig. 4A). Note that this is to a greater extent than other EAs. Taken together, the findings for both GMA and MM kinetics implicate the unreliability of CMAES under noisy experimental conditions.

We further note the inconsistency in MM parameter estimation by CMAES, as discussed above, has not been obvious from the distribution of the MSE-of-fit (Additional file 1: Fig. S7, right column). Instead, the MSE value for individual seed-instances is largely identical to that of other EAs, resulting in an akin overall distribution for all noise levels. Clearly, similar MSE values do not necessarily imply a similar quality of parameter estimations. (Yet conversely, we reason that MSE values and distributions that are distinctly larger should be seen as strong indicators of the relative ineffectiveness of parameter estimations.) Notably, the sensitivity of CMAES to noisy data would not be obvious, if not for the usage of the R^2 metric. In turn, the usage of the latter metric would only be possible if the underlying parameter values and reaction kinetics are known, such as in the context of an *in silico* pathway.

Recovery of parameters and dynamics can benefit from datapoints as close as 15 min apart

Given the superior applicability of SRES and G3PCX algorithms to the respective case of GMA and MM kinetics under noisy empirical conditions, we further investigate their sensitivity to another common factor affecting performance: the number of hourly empirical datapoints. To do so, we reduce the number of datapoints from four per hour to (i) two, and then (ii) one hourly. Then, in each case, we reevaluate the capability of the two EAs regarding parameter estimations and the recovery of metabolite dynamics.

For systems having reversals in dynamic trend (i.e., switching from increasing to decreasing concentrations, and vice versa) between two to four hours like our modeled system, our findings suggest the benefits of having

measurements more than once or even twice hourly. For a start, there is a tendency towards more accurate parameter estimations as exemplified by the R^2 value for the SRES case study, regardless of measurement noise (Fig. 5, top and bottom of left column). (Note that there is no clear deterioration in the recovered dynamics due to less datapoints [Additional file 1: Fig. S12 versus Additional file 1: Fig. S5A and Additional file 1: S8B]). Although the G3PCX estimation of MM parameters shows similarly high and stable metric value for all considered cases of hourly datapoints (one, two, and four) in the absence of measurement noise (Fig. 5, top and bottom of right column), the algorithm however performs poorly according to the metric for the case of one hourly datapoint with measurement noise. More illuminatingly, when we simulate the MM dynamics with the estimated parameters, we find the profiles based on one or two hourly datapoints do not faithfully reproduce all the underlying profiles (Additional file 1: Fig. S13), unlike the case when four hourly datapoints are used (Additional file 1: Fig. S5B). On the same note, in the presence of measurement noise, the regenerated dynamics of two hourly datapoints (Additional file 1: Fig. S13) are not as close to the real profile, compared to 4 datapoints (Fig. 4B). Taken together, our data suggests that having more than two hourly empirical datapoints (e.g., four) can be still helpful for improving parameter estimations and the recovery of dynamics, regardless of noise; certainly, this is of paramount importance in high-value applications, and when safety is critical [7].

No evidence that data augmentation improves parameter estimations

Our finding in the previous section begs the question if the augmentation of dynamic data could improve the recovery of parameters and metabolite dynamics, although it has been clear to us that it cannot be a substitute for empirical data. Again, we use the case studies of SRES and G3PCX algorithms for GMA and MM kinetics, respectively, both in the presence and absence of measurement noise. Initial datasets with one or two hourly datapoints are then augmented to four datapoints hourly before parameter estimations. We find at best, a small and negligible increase in the R^2 value of parameter estimations in many cases (Fig. 6) but can also potentially result in a sharp and significant decrease in the accuracy of estimations, as revealed by the data augmentation of one hourly datapoint for the G3PCX estimation of MM parameters (Fig. 6, bottom right). As such, we refrain from recommending data augmentation for the purpose of parameter estimation, as there is no evidence of its benefit from our analysis. In fact, our finding of a clearly adverse effect in one

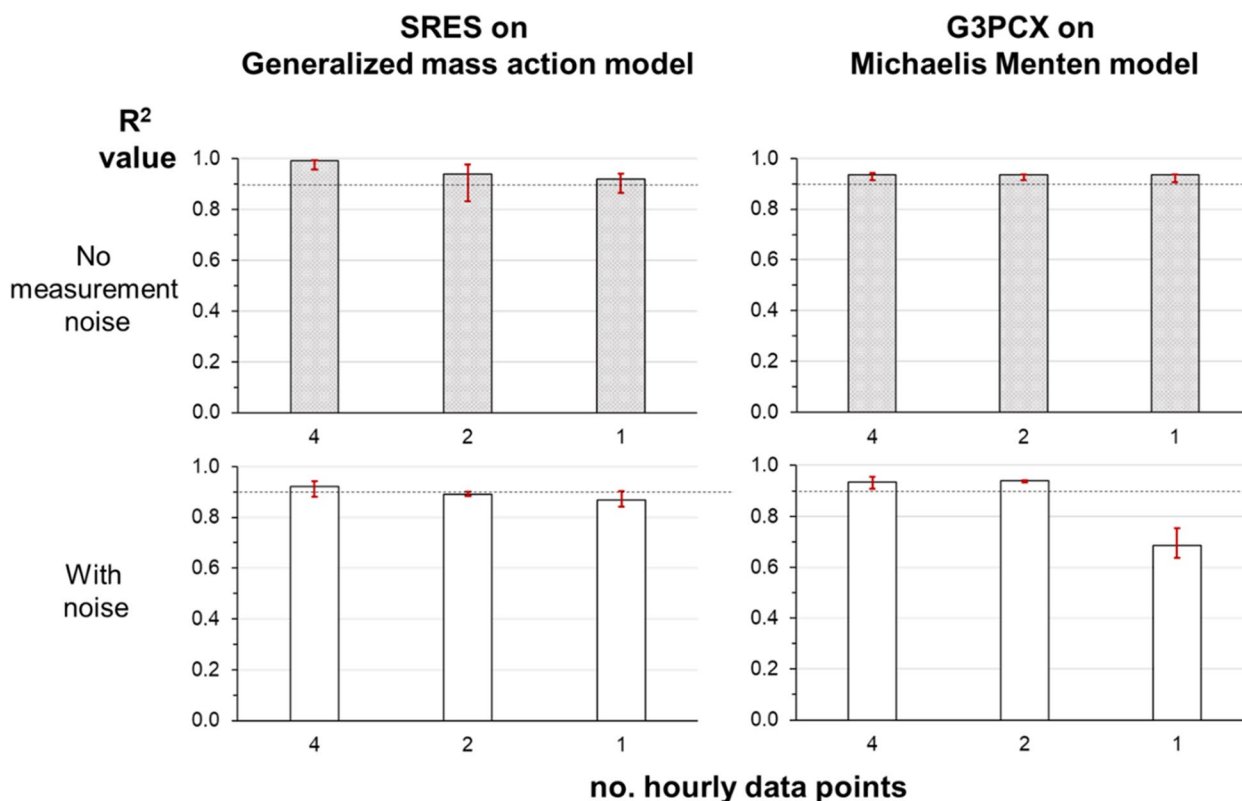


Fig. 5 The effect of the number of hourly data points on the quality of parameter predictions as reflected by the R^2 metric. Two case studies are shown: estimation of generalized mass action (GMA) parameters using the SRES algorithm (left column) and the Michaelis–Menten (MM) parameter estimations utilizing the G3PCX algorithm (right column). The top row and bottom row are for the cases without and with measurement noise, respectively. For GMA parameter predictions, the amount of measurement noise is sampled from a normal distribution with mean zero and a standard deviation equivalent to 5% of the underlying metabolite/enzyme measurement. The corresponding value is 7.5% for MM parameter predictions. The highest noise level has been chosen in each case such that the algorithm still performs well with 4 hourly datapoints. Triplicate values are derived using different initial seed solutions (Additional file 8: Data S7). The bars represent the median values while the lower and upper bounds (in red) mark the smallest and largest values based on the three initial seed solutions

case suggests that the repeated usage of data augmentation could result in a poorer quality of parameter estimations on average. Furthermore, any dataset sparser than those investigated here is likely to introduce more systematic bias in the augmented datapoints, as there is no information locally in time to reliably inform their estimate.

Protocol and learning points for uncovering reaction kinetic parameters in practice

We recapitulate the two key modules and component steps in our protocol contributing to effective parameter estimations (Fig. 7). Module one (first 5 steps in red) focuses on deriving high-quality estimates of the net reaction rate for all metabolites at each timepoint. In doing so, it forms the foundation for accurate parameter estimations in module two (the last three steps in blue). Their rationale or learning points are briefly as follows:

Module one

- In step one, the generation of timepoint data that are as close as possible reflects the value of having more timepoints in deriving accurate empirical values of net reaction rates. We exemplify this point by demonstrating the noticeable end improvement in parameter estimation accuracy by simply increasing from one to two and then four hourly datapoints (Fig. 5). The effect is even more pronounced in the typical case where measurement noise is present. This is despite the implementation of all the other steps for maximizing accuracy. Note that data augmentation is not an effective solution for insufficient timepoints (Fig. 6).
- Steps two and three include means for diminishing the effect of noise: taking an average of triplicate measurements and employing the Savitzky-Golay fil-

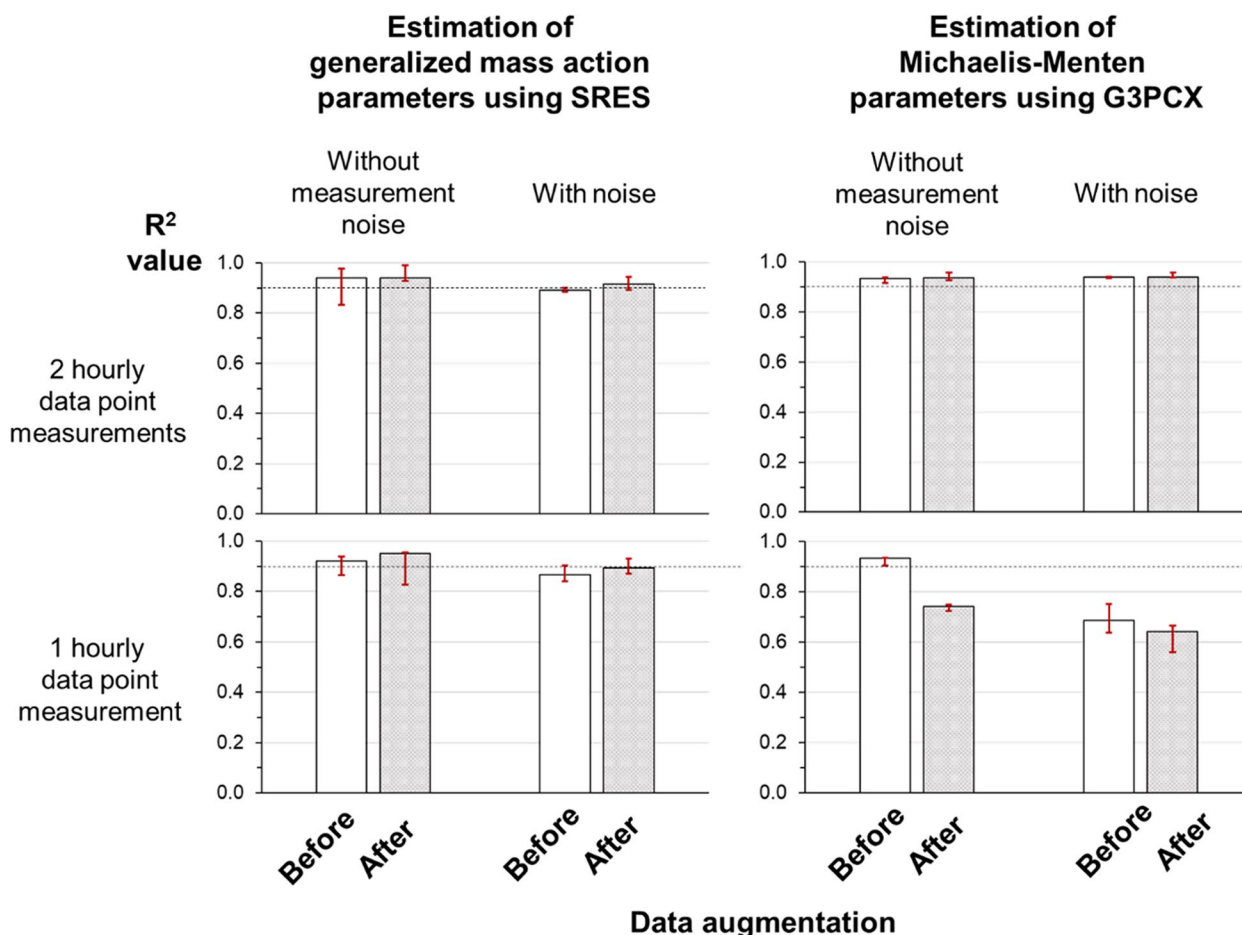


Fig. 6 The effect of data augmentation on the quality of parameter estimations as reflected by the R^2 metric. Two case studies are shown: estimation of generalized mass action (GMA) parameters using the SRES algorithm (left column) and the Michaelis–Menten (MM) parameter estimations utilizing the G3PCX algorithm (right column). Data augmentations (Materials and methods) are done in two cases to increase the hourly data points from 2 (top row) and 1 (bottom row) to 4 per hour. For GMA parameter predictions, the amount of measurement noise is sampled from a normal distribution with mean zero and a standard deviation equivalent to 5% of the underlying metabolite/enzyme measurement. The corresponding value is 7.5% for MM parameter predictions. Triplicate values are derived using different initial seed solutions (Additional file 9: Data S8). The bars represent the median values while the lower and upper bounds (in red) mark the smallest and largest values based on the three initial seed solutions

ter, which concurrently filters out noise, while computing the net reaction rate (materials and methods).

- Steps 3–5 further provide the mechanism for generating realistic time profiles of metabolites based on net reaction rates, by adjusting the Savitzky-Golay parameters (window size and polynomial degree for local fitting). By observing the similarity with experimental profiles, we obtain direct evidence of the reliability of derived net reaction rates. Among other possibilities, the MSE between regenerated and empirical concentration values (the average of triplicates) across all timepoints and metabolites can be used as a metric for assessing their similarity in an automated search for good Savitzky-Golay parameters.

Module two

- Step 6 estimates kinetic parameters (materials and methods) by using effective EA implementation according to the underlying rate law (Table 2), and their corresponding hyperparameters and termination criteria (materials and methods).

Importantly, the average of parameter values from multiple seed solutions (e.g., ≥ 10) is taken to improve their accuracy and decrease variability, compared to individual seed runs (Additional file 1: Fig. S11).

Learning points / rationale

Component steps

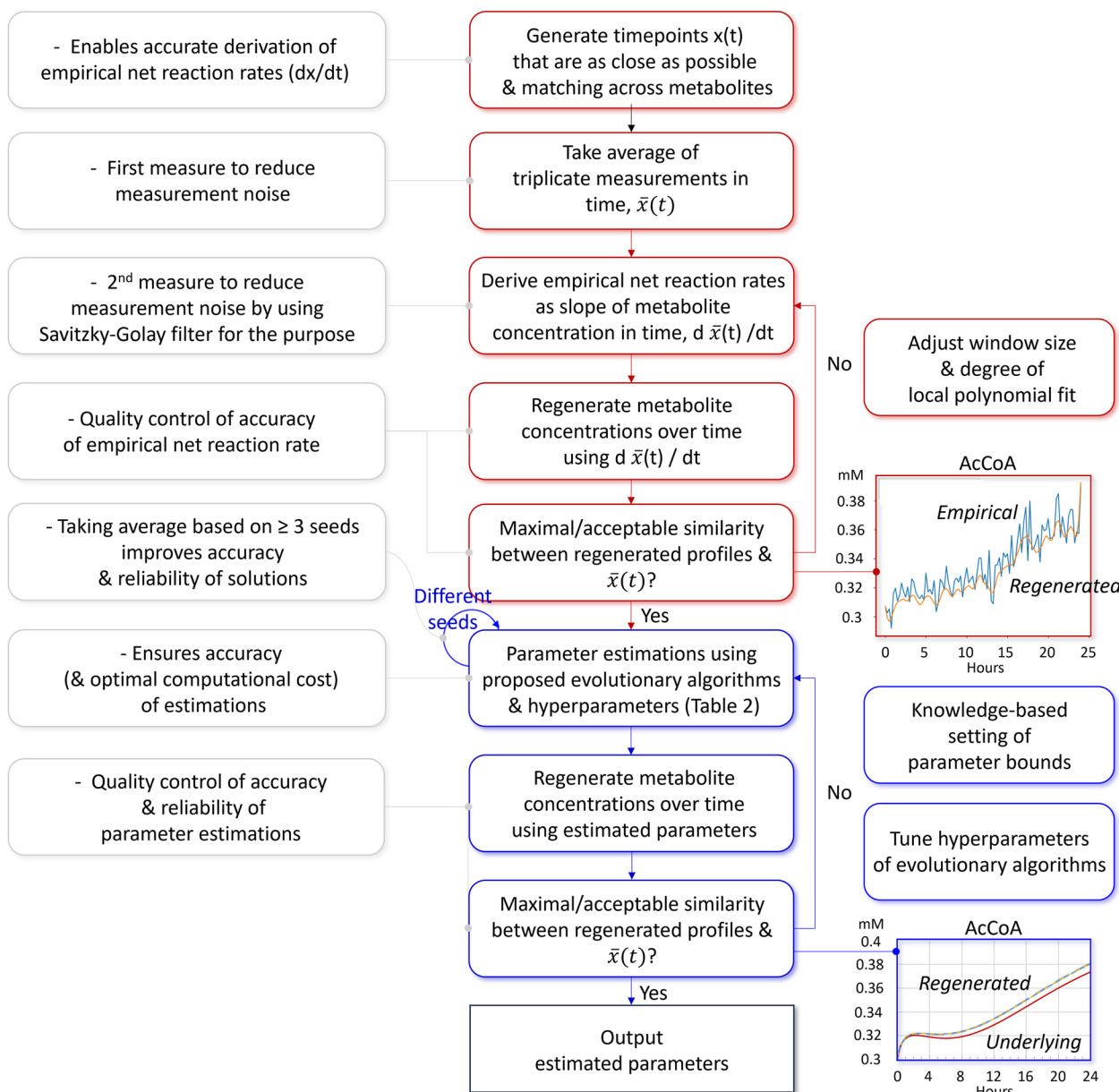


Fig. 7 Protocol and learning points for uncovering reaction kinetic parameters in practice. The gray boxes outline the rationale or learning point of the associated steps. The red boxes depict the 5 preprocessing steps to derive high-quality values for empirical net reaction rates, while the blue boxes set out the three subsequent steps for parameter estimations

- Steps 6–8 are collective mechanisms for enhancing the similarity between regenerated and experimental time profiles of metabolites, like steps 3–5. Without knowing the actual kinetic parameter values, the similarity serves as the end-all-be-all criteria for assessing the overall reliability of their estimations. The bounds of the parameter search space should be informed by domain knowledge as much

as possible. The hyperparameters of the employed EA can also be tuned.

By deconstructing the approach into two modules, it will be also easier to sieve out issues as to data quality and preprocessing, versus the actual parameter estimations, by examining the quality of regenerated profiles in modules one and two, respectively.

Table 2 Suggested evolutionary algorithms (EAs) for parameter estimations of known underlying reaction kinetics

System/model	Suggested EAs	Comment	Supporting figure(s)
<i>Generalized mass action system (without measurement error)</i>	CMAES	Most consistent and highest quality of parameter estimations, lowest computational cost	Fig. 2, Additional file 1: Fig. S5A, and Additional file 2: Data S1
<i>Generalized mass action system (with measurement errors)</i>	(I)SRES	Most consistent and highest quality of parameter estimations	Fig. 3 and Additional file 1: Fig. S8
<i>Michaelis–Menten system (with/without measurement error)</i>	G3PCX	Lowest computational cost, and among the best in consistency and quality of parameter estimations	Additional file 2: Data S1, Figs. 2 and 3, and Additional file 1: Figs. S5B and S9
<i>Linear logarithmic model^a</i>	CMAES	Lowest computational cost, among the best in consistency and quality of parameter estimations	Fig. 2 and Additional file 2: Data S1

^a The kinetic formulation is a heuristic modeling approach, with no empirical profiling of the rate law to speak of

Discussion

No free lunch: effectiveness and limitations of evolutionary algorithms

In pushing the performance of EAs to their limit, we realize the feasibility of deriving remarkably good parameter estimates in the presence of measurement errors, by using the appropriate EA (Table 2) with an entire repertoire of mitigation techniques in our protocol. Given that these EAs were not designed for systems biology applications, it is perhaps surprising to find their clear and distinct competitive traits, in terms of the accuracy and efficiency of parameters estimation as to individual rate laws, further differentiated by their versatility, robustness to measurement noise, and sensitivity to the initial seed solutions.

Notably, there appears to be trade-offs among their competitive advantages. For example, SRES may be compensating for their accuracy, versatility, as well as robustness to measurement noise and initial seed solutions, by expending large amounts of computation in exploring and exploiting the search landscape. To give another example, while CMAES could very well exploit the intrinsic relationships among the advantageous search directions of various GMA parameters, such relationships seem to be easily disrupted by measurement noise, which can create local kinks in the search terrain to disrupt the said relationships. These trade-offs attest to the veracity of the ‘no free lunch’ theorem, which implies that there is no algorithm that can outperform all others in all metrics under all contexts. In this light, perhaps, some limitations may be perceived as the necessary cost for outperformance in other aspects, and thus cannot be avoided. Even so, their specific implementations were largely ad-hoc, intuitive, and improvised stepwise, honed by the datasets-of-interest. This modus operandi is obvious from the originating publications, which mostly focus on spelling out the implementation and the comparative performance, without articulating their strategic

advantages over competing algorithms. The resulting lack of understanding of the reasons behind their effectiveness and limitations is thus carried forward and may become even more obscure in the context of new applications. In this regard, our curiosity provides the driving force behind the current work, which provides an elucidative but preliminary glimpse of the importance of customizing the optimization algorithm to the rate law formulation and experimental condition (e.g., the type of noise and noise level). Lastly, we quickly note that the poor performance of DE might very well be rooted in its origin in GA, which is designed for search in simplified discrete-variable space, rather than the continuous search space that is characteristic of parameters estimation.

Global (or near global) optimization requirement for parameters estimation

We take this opportunity to further elaborate on the difference in applicability between global and local optimization methods in the context of the Levenberg–Marquardt (LM) algorithm, which is commonly used to fit kinetic models. We have excluded the local search method from our study, as the aspiration of predictive modeling will require a more stringent recovery of the true parameter values, necessitating a global search. To recap, the LM algorithm is partly a gradient descent-based approach, whereby the model fitting error (sum-of-squares) is systematically being reduced by transiting in the direction of the steepest descent in the parameter search space. When the search location is determined to be near a minimum, another (Gauss–Newton) method is used to more efficiently find the minimum point, by approximating the error as a quadratic function, with respect to distance in the search direction. The algorithm thus works by leveraging on local terrain information to find a minimum point, akin to a ball rolling down a slope to the minimum point. In other words, it does not have a strategy for global

search. The solution that the LM algorithm finds, is therefore highly sensitive to the starting search location for complex terrain [54], which is a sufficient basis for us to exclude potential candidates in this study. Along the same line, we further evaluate the sensitivity of shortlisted EAs to different seed solutions in finding the actual parameter values. Notably, the only condition whereby the LM algorithm is guaranteed to find a global solution is when the error function is convex with respect to the parameter values within the search space. However, the required condition is not a given for most biochemical systems. Even for the relatively simple network in our study, the search terrain can be so complex that the true solution is not always found by our global optimization candidates.

Contribution to better modeling of biological systems

To spell out how our cumulative findings contribute to the better modeling of biological systems, we first recap that the aspiration of predictive modeling requires improvement in methodologies with respect to two key objectives in systems biology:

1. To uncover an accurate representation of the underlying rate laws
2. To accurately estimate the parameter values of the uncovered rate laws

In this light, the current study focuses on achieving objective two, by first demonstrating its achievability and then further contributing the following to enable its practice:

- Technology: an effective protocol for estimating the true parameters, based on:
 - i. Supporting preprocessing methods,
 - ii. Effective optimization algorithms, as to major formulations of reaction kinetics,
 - iii. While mitigating the effect of measurement noise.

The software has been provided with this work, allowing for immediate translational application and verification.

- Shift in understanding:
 - i. The imperativeness of having closely spaced time-series datapoints for most (if not all) molecular participants, which cannot be replaced by data augmentation.

- ii. It is possible to largely mitigate the effect of measurement errors on parameter estimation, even if the errors are heteroscedastic as commonly found in biological systems.
- iii. In contrast, explanative modeling does not require:
 - Comprehensiveness of data with respect to molecular participants
 - Close temporal spacing of datapoints
 - Global optimization to uncover the true parameters
 - Suitability of the optimization algorithm for the formulated rate law
 - That the chosen rate law formulation truly reflects the kinetics

This is because for explanative modeling, the accuracy of dynamic prediction (and hence the required accuracy of estimated parameters) is not foremost; its purpose is to merely provide a plausible explanation of the observed data in terms of pathway dynamics.

Translational and practical impacts

Indeed, it is both imperative and a delight to be able to use experimental data measured from biological systems to validate the proposed protocol. However, the closely spaced temporal data for all molecular participants, which is required for the emerging area of predictive modeling, is simply not yet available, e.g., time-series concentration data spaced 15 min apart or better, for dynamics with time scale of reversal (i.e., switching from increasing to decreasing concentrations, and vice versa) between 2 to 4 h, like our modeled system. Most if not all available datasets to date are sparse, with molecular participants mostly unprofiled. Intuitively, closely spaced datapoints for all participants are, however, required for accurately computing the instantaneous slope of their dynamic profiles to obtain the net reaction rates, i.e., the Y in the kinetic model: $Y=f(X)$. Without reliable data on Y , it will be impossible to find the parameters predicting the true Y .

Nonetheless, while the applicability to currently available data is low, we believe our work will impact how modeling and experimental data collection are being done going forward. Particularly, we foresee (or hope to see) the following practical and translational impacts:

- *A trend towards practices enabling more accurate modeling* For example, given the criticality of having closely spaced temporal datapoints for predictive modeling that is unreplaceable by data augmentation, we foresee an increase in experimental measurements that better fit our identified requirements for the purpose.

- *Accelerated technological and fundamental science developments enabling predictive modeling* By demonstrating the feasibility of predictive modeling, outlining the required conditions, and articulating the translational relevance, we aim to sow the seeds to drive its application in the relevant industries.
- *Direct and immediate applicability to cell-free biomanufacturing* Synthetic pathways for cell-free production are typically known (e.g., nuclei acids, monoclonal antibodies, high-value small molecules, etc.) and simpler than cellular systems and are thus likely to be the first bioproduction systems to achieve and benefit from predictive modeling. Together with the advent of technologies enabling automated online measurements, data sparsity both temporally and with respect to molecular participants may be more easily overcome, unlike the tedious and laborious measurements required for conventional biological domains, which are cell-based.

Caveats in practice

Lastly, there remain caveats in using our protocol. For example, the bounds of the parameter search space have an enormous effect on whether the true solution can ever be found. While we recommend constraining as much as possible the search space guided by domain knowledge, it may not be practically feasible for most parameters. It is also currently unclear if an EA is intrinsically suitable for the reaction kinetics that it performs well for in our study, and to what extent, it will remain so with an increasing number of reactions. Besides the need to address these questions, future works should explore if the hyperparameters tuning of EAs can make a material difference to their performance, such as in the required computational cost. To appreciate the latter's importance, imagine an early fish requiring minutes just to scan possible escape routes in the face of an ambush predator. Clearly, there is a lot more to learn from the history of life if predictive modeling is to become truly useful for real-world endeavors.

Conclusions

While systems biology has begun its nascent transition from the explanatory modeling of fitted data to the predictive modeling of unseen data, it still necessitates the effective recovery of all underlying reaction parameters. In this regard, we found algorithms that are effective under marked measurement noise for specific reaction kinetics, as a step towards predictive modeling for systems biology endeavors. We further provide the key learning points and the protocol to do so.

Methods

Candidate evolutionary algorithms

The following candidate EAs are evaluated in our work. Note that clustering techniques [55, 56] and adaptive stochastic methods [57, 58] are not considered, as they are deemed less effective [9]. Ant Colony optimization, Taboo Search, and particle swarm methods [59] are also not assessed, as they are not used commonly.

- DE [60] is a direct global search method, which is designed for optimizing nonlinear, non-differentiable, continuous objective functions. It iteratively improves a population of candidate vector solutions as follows:

- *Initialization*: A population of candidate solutions is first initialized.
- *Mutation*: For each candidate solution, three random individuals (a, b, c) are selected from the population and combined to form a trial vector by using the operation: $\text{trial_vector} = a + F * (b - c)$, where F is a user-defined amplification factor of (b-c).
- *Crossover*: operation then results in a segment of the trial vector replacing its counterpart in a pre-determined target. The starting location of the segment is randomly chosen, while a user-defined parameter, "crossover probability," determines the probability of extending the segment by an additional vector component.
- *Selection*: The resulting trial vector is then compared to the target based on objective function value and replaces the target if it gives a better value.
- *Termination*: The mutation, crossover, and selection steps are repeated until a termination condition is met, such as a predefined number of generations, convergence of the objective function to some plateau value, etc.

- ESs [61] are population-based like DE, but are designed for search in continuous space. They are highly varied in the specifics of implementation but generally consist of the following stages:

- *Initialization*: μ number of parental vector solutions are first initialized, based on a user-defined sampling scheme.
- *Recombination*: the μ parental solutions are duplicated, and then followed by the probabilistic exchange of solution components among them, via predefined rules, to generate a new population of λ "offspring" solutions. Note that the process

is carried out for the CMAES algorithm, but not SRES and ISRES.

- *Mutation*: the change in each component value of the ‘offspring’ solutions is sampled from the corresponding normal distribution with zero mean and some standard deviation. The latter also called the mutation step size, is a search strategy parameter, as it has the effect of balancing the exploration and exploitation of solutions. For example, the mutation step size can be increased to ensure a wider search when progress is poor and to decrease the step size when progress is promising. ES variants may implement a specific approach to automatically adapt the strategy parameter to the search landscape. Offspring solutions generated during the process may only be accepted (“non-lethal”) if they are within user-defined bounds.
- *Selection*: the fitness of each of λ offspring is then evaluated based on the objective function, with the best μ solutions chosen to make up the next generation of parents. The μ parents of the current generation may form part of the selection pool (not the case for our tested variants).
- *Termination*: the recombination, mutation, and selection steps are repeated until a termination condition is met, such as a predefined number of generations, convergence of the objective function to some plateau value, etc.

In our study, three ES variants were tested:

- SRES [33] probabilistically uses only objective function for comparing adjacently ranked solutions during bubble sort, without penalizing infeasible solution(s). It thus gives unacceptable solutions a chance to be selected, which could give rise to better and more feasible solutions in the next generation. In this way, the algorithm encourages exploration and helps prevent premature convergence.
- ISRES attempts to improvise over SRES by incorporating an alternative mutation strategy of using the differential between selected solutions to generate new ones [62], similarly to DE. This is carried out for variables in the search space that are dependent on others.
- CMAES [34] adapts the multi-variate normal distribution that is used for sampling mutation sizes, by overweighting those selected previously. As such, some components of the solution may have preferred mutation sizes that are correlated with each other. In this light, the covariance matrix of the distribution may become non-diagonal over generations. Besides adapting the covariance matrix, the algorithm also

adopts specific principles used in natural gradient descent to ensure that the distribution remains undistorted by one of its key steps (reparameterization).

3. G3PCX [63] presumes offspring solutions generated near parental ones in continuous search space to be more likely to be good candidates, given that the parents themselves have been previously selected based on their fitness. A particular recombination operation is used to ensure the offspring solutions are normally distributed and centered on individual parents (“parent-centric”). The algorithm also modifies its core model, the minimal generation gap (MGG) [64], to select the best two solutions separately from each batch of similar fitness in each generation, instead of using the roulette-wheel selection procedure, which is computationally expensive, to choose just one solution.

Generating synthetic enzyme datasets

An overview of the artificial pathway is provided in Fig. 1A. For each of the 4 *in silico* experiments, the final level for each of the 9 enzymes is chosen from three levels: low, medium, and high, via Latin Hypercube sampling [65]. The Hill function is then used to generate the concentration data $e_i(t)$ of enzyme i at time t of the run as follows:

$$e_i(t) = \frac{\alpha * k_{f,i} * t}{k_{m,i} + t} + e_{0,i}$$

Here, $k_{f,i}$ is the maximum increase in enzyme concentration, while the relative amplification factor α is 0.432, 1, and 2.486 for the low, medium, and high final level [66], respectively. Also, $k_{m,i}$ is the time required to increase the enzyme concentration by 50% of $\alpha * k_{f,i}$, while $e_{0,i}$ is the basal enzyme concentration prior to tuning. Other than α , the parameter values of each enzyme (Data S1) are uniformly sampled from predefined ranges.

Generating synthetic metabolite datasets by reaction kinetics

To do so, the initial concentration of metabolites (Additional file 3: Data S2) is set so that they are not exhausted during the run at the medium tunable level for all enzymes; enzyme concentrations over time are determined by using Hill’s function as above. Based on these conditions, metabolite concentration data are then generated at 15 min intervals over 24 h by simulating the underlying reactions using separate forms

of reaction kinetics for the system model (Additional file 1). The kinetic parameter values are obtained as follows: MM parameters are first arbitrarily chosen (Additional file 4: Data S3) and then used to simulate metabolite time-series data. The latter data is then used to fit the other kinetic models (using the CMAES algorithm) to efficiently engender reasonable parameter solutions. Parameters are also liberally altered to keep them within reasonable ranges.

The measurement error of metabolite and enzyme concentrations is sampled from a normal distribution with zero mean and a standard deviation equivalent to 2.5%, 5%, or 7.5% of the underlying value (otherwise referred to as 2.5%, 5%, 7.5% “noise levels”).

Data augmentation

Monotonic cubic splines are used piecewise (“Piecewise Cubic Hermite Interpolating Polynomial” [67]) to interpolate the metabolite concentration at new time points in between the regular temporal datapoints.

Computing metabolite net reaction rates from triplicate time-series data

For both metabolite and enzyme datasets, the average of triplicate measurements at regularly spaced matching timepoints is first taken. The net metabolite reaction rate y_t at each temporal data point is given by the instantaneous slope, i.e., the first derivative of the concentration curve in time. We use the Savitzky-Golay filter [68] for evaluation, as it intrinsically smoothens noisy measurements. For datasets with measurement noise, we apply the filter with a temporal window size of 5 datapoints and polynomial order 2. In the absence of noise, the values are 7 and 3, respectively.

To assess the accuracy of the computed y values, the metabolite concentration at each timepoint $t=t_k$ is then recovered by summing up its changes over time from its initial concentration:

$$x(t_k) = x(0) + \sum_{t=0}^{t=t_k-1} (y_t * \Delta t)$$

where Δt is the temporal spacing between neighboring datapoints, e.g., $\Delta t = 0.25$ h for 4 temporal datapoints per hour. The regenerated datapoints in time are then compared visually with the corresponding measurements by overlaying them.

Implementing evolutionary algorithms, hyperparameters, and termination criteria

We use evolutionary algorithms implemented in the python package, pymoo (version 0.6.0) [14]. Other than setting a population size of 64, the default hyperparameter settings

in pymoo are used for all algorithms. We set an upper limit of 1×10^5 generations for the termination criterion (Table 1), which allows most algorithms to reach a plateau state (in terms of the cost function) or to self-terminate for the CMAES algorithm (Additional file 1: Figs. S1–S4).

Estimating kinetic parameters for artificial pathway

The nonlinear programming problem is formulated as follows:

Find the vector of rate law parameter values (decision variable values), \mathbf{p} , to minimize the cost function, i.e., the overall mean square error MSE in our case:

$$MSE = \frac{1}{S * M * (T + 1)} \sum_{s=1}^S \sum_{m=1}^M \sum_{t=0}^T (y_{s,m,t} - \widehat{y}_{s,m,t})^2$$

where S is the total number of experimental runs, M is the total number of modeled metabolites, $T+1$ is the total number of time points, $y_{s,m,t}$ is the net reaction rate for metabolite m at time t in experiment s , and \widehat{y} is the corresponding prediction by the system model. The model is in the form of a set of ordinary differential equations as described in section one of Additional file 1.

MSE is further subjected to:

$$\begin{aligned} \widehat{y}_{s,m,t} &= \mathbf{f}_m(\mathbf{x}(s, t), \mathbf{e}(s, t), \mathbf{p}, \mathbf{v}) \\ \mathbf{x}(s, t) &= \mathbf{x}_{s,t} \\ \mathbf{e}(s, t) &= \mathbf{e}_{s,t} \\ \mathbf{v} &= \mathbf{v} \\ \mathbf{Plb} &\leq \mathbf{p} \leq \mathbf{Pub} \end{aligned}$$

where \mathbf{x} is the vector of metabolite concentrations as state variables, \mathbf{e} is the vector of enzyme concentrations as control variables, \mathbf{v} is the vector of other parameter values that are not being estimated in the current problem (usually invariant in the system with time). The constant inflow of AcCoA precursor into our modeled system is represented by one such parameter V_{in} .

Performance metric for parameter estimations: coefficient of determination

Recall that Pearson’s correlation coefficient indicates the degree to which two variables, y and x , are statistically related by a linear relationship between them, i.e., $y=mx+c$. To measure the quality of parameter predictions by candidate evolutionary algorithms, we use a similar metric to gauge the extent to which, the predicted values \widehat{k}_r are equivalent to the actual values \widehat{k}_r , i.e.,

$$k_r = \widehat{k}_r$$

This provides a direct and more stringent measure of estimation quality than otherwise if we merely assess the degree to which k_r and \widehat{k}_r are linearly related. Note that in this case, m and C are 1 and 0, respectively. Analogously, we calculate the coefficient of determinant R^2 :

$$R^2 = 1 - \frac{SS_r}{SS_t}$$

whereby the residual sum of square SS_r and the total sum of square SS_t are also analogously computed as follows:

$$SS_r = \sum_{i=1}^N (k_{r,i} - \widehat{k}_{r,i})^2$$

$$SS_t = \sum_{i=1}^N (k_{r,i} - \overline{k}_r)^2$$

$k_{r,i}$ is the actual value of the i th rate law parameter, while $\widehat{k}_{r,i}$ is the corresponding predicted value; \overline{k}_r is the mean of the actual value of all N rate law parameters.

Evaluating improvement in R^2 by taking average solution from multiple seed runs

The Wilcoxon-Mann-Whitney test is carried out using the software EDISON-WMW [69].

Abbreviations

CK	Convenience kinetics
CMAES	Covariance matrix adaptation evolution strategy
DE	Differential evolution
EAs	Evolutionary algorithms
ESs	Evolutionary strategies
G3PCX	Generalized generation gap model with parent-centric combination
GMA	Generalized mass action
GA	Genetic algorithms
ISRES	Improved SRES
LM	Levenberg-Marquardt
Linlog	Linear-logarithmic
MSE	Mean square error
MM	Michaelis-Menten
R^2	Coefficient of determination
SA	Simulated annealing
SRES	Stochastic ranking evolutionary strategy

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12915-024-02019-4>.

Additional file 1: Figure S1. Convergence profile in fitting the generalized mass action model by candidate evolutionary algorithms and seed-instances. Figure S2. Convergence profile in fitting the Michaelis-Menten model by candidate evolutionary algorithms and seed-instances. Figure S3. Convergence profile in fitting the Linlog model by candidate evolutionary algorithms and seed-instances. Figure S4. Convergence profile in fitting the convenience rate law model by candidate evolutionary algorithms and seed instances. Figure S5. Simulated generalized mass action (A) and Michaelis-Menten dynamics (B) based on parameters optimized using selected evolutionary algorithms for the case of zero measurement error. Figure S6. Metabolite dynamics of Linlog model based on parameters optimized using selected evolutionary algorithms for the case of zero measurement error. Figure S7. Box plot distribution of the final cost function (MSE-of-fit) values for candidate evolutionary

algorithms upon reaching the termination criterion. Figure S8. Simulated mass action dynamics based on parameters estimated using selected evolutionary algorithms. Figure S9. Simulated Michaelis-Menten dynamics based on parameter values that are optimized using selected evolutionary algorithms. Figure S10. Convergence profile in fitting the Michaelis-Menten model at various levels of Gaussian measurement noise using the G3PCX algorithm. Figure S11. Cumulative effect of individual optimizations on accuracy and reliability of parameter estimations. Figure S12. Simulated generalized mass action dynamics based on parameter estimations using (metabolite and enzyme concentration) data with reduced hourly datapoints, i.e., two and one hourly datapoints. Figure S13. Simulated Michaelis-Menten dynamics based on parameter estimations using (metabolite and enzyme concentration) data with reduced hourly datapoints, i.e., two and one hourly datapoints.

Additional file 2: Data S1. Number of generations required for cost function (MSE of fit) evaluation to reach a plateau (or self-terminate for CMAES) for various rate law formulations.

Additional file 3: Data S2. Hill parameters and initial metabolite concentrations used for generating synthetic data.

Additional file 4: Data S3. Upper and lower bounds of the parameter search space for the individual various rate laws.

Additional file 5: Data S4. R^2 and MSE-of-fit datapoints in Fig. 2.

Additional file 6: Data S5. Datapoints in Fig. 3.

Additional file 7: Data S6. Datapoints in Fig. 4.

Additional file 8: Data S7. R^2 datapoints in Fig. 5.

Additional file 9: Data S8. Datapoints in Fig. 6.

Additional file 10: Data S9. Datapoints in Figure S5.

Additional file 11: Data S10. Datapoints in Figure S6.

Additional file 12: Data S11. Datapoints in Figure S7.

Additional file 13: Data S12. Datapoints in Figure S8.

Additional file 14: Data S13. Datapoints in Figure S9.

Additional file 15: Data S14. Datapoints in Figure S11.

Additional file 16: Data S15. Datapoints in Figure S12.

Additional file 17: Data S16. Datapoints in Figure S13.

Acknowledgements

Publicly available codes were modified or reused for our work [8].

Authors' contributions

HCY conceptualized, planned, and conducted the research. HCY supervised the student and wrote the article. VV explored various parameter combinations for the Savitzky-Golay filter and methods for pooling solutions based on multiple seeds. KS supervised the work and reviewed the article. All authors read and approved the final manuscript.

Funding

This work was supported by the Agency for Science, Technology and Research (A*STAR).

Data availability

All data generated or analyzed during this study are included in this published article, its supplementary information files, and publicly available repositories (<https://zenodo.org/records/13831239>) [70]. The analyzed datasets are generated by simulation using the codes provided in this study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, Matrix #07-01, Singapore 138761, Republic of Singapore. ²Synthetic Biology Translational Research Program, Yong Loo Lin School of Medicine, National University of Singapore (NUS), 10 Medical Drive, Singapore 117597, Republic of Singapore. ³School of Biological Sciences, Nanyang Technological University (NTU), 60 Nanyang Drive, SBS-01s-45, Singapore 637551, Republic of Singapore. ⁴Synthetic Biology for Clinical and Technological Innovation (SynCTI), National University of Singapore (NUS), 28 Medical Drive, Centre for Life Sciences #02-07, Singapore 117456, Republic of Singapore.

Received: 12 March 2024 Accepted: 24 September 2024

Published online: 14 October 2024

References

- Bennett MS. A brief history of intelligence: humans, AI, and the five breakthroughs that made our brains. New York: HarperCollins; 2023. p. 320.
- Yeo HC, Selvarajoo K. Machine learning alternative to systems biology should not solely depend on data. *Brief Bioinform*. 2022;23(6):bbac436.
- Volk MJ, Tran VG, Tan S-I, Mishra S, Fatma Z, Boob A, et al. Meta-biochemical engineering: methodologies and applications. *Chem Rev*. 2023;123(9):5521–70.
- Hirschi S, Ward TR, Meier WP, Müller DJ, Fotiadis D. Synthetic biology: bottom-up assembly of molecular systems. *Chem Rev*. 2022;122(21):16294–328.
- Akhon N. Precision medicine: a new paradigm in therapeutics. *Int J Prev Med*. 2021;12:12.
- Kitano H. Systems biology: a brief overview. *Science*. 2002;295(5560):1662–4.
- Wright L, Davidson S. How to tell the difference between a model and a digital twin. *Adv Model Simul Eng Sci*. 2020;7(1):13.
- Costello Z, Martin HG. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *NPJ Syst Biol Appl*. 2018;4(1):19.
- Mendes P, Kell D. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*. 1998;14(10):869–83.
- Abernathy MH, He L, Tang YJ. Channeling in native microbial pathways: implications and challenges for metabolic engineering. *Biotechnol Adv*. 2017;35(6):805–14.
- Noor E, Bar-Even A, Flamholz A, Reznik E, Liebermeister W, Milo R. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput Biol*. 2014;10(2):e1003483.
- Gerosa L, Haverkorn van Rijsewijk Bart RB, Christodoulou D, Kochanowski K, Schmidt Thomas SB, Noor E, et al. Pseudo-transition analysis identifies the key regulators of dynamic metabolic adaptations from steady-state data. *Cell Syst*. 2015;1(4):270–82.
- Hackett SR, Zanotelli VRT, Xu W, Goya J, Park JO, Perlman DH, et al. Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science*. 2016;354(6311):aaf2786.
- Daran-Lapujade P, Rossell S, van Gulik WM, Luttki MAH, de Groot MJL, Slijper M, et al. The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *Proc Natl Acad Sci*. 2007;104(40):15753–8.
- Digel M, Ehehalt R, Stremmel W, Füllekrug J. Acyl-CoA synthetases: fatty acid uptake and metabolic channeling. *Mol Cell Biochem*. 2009;326(1):23–8.
- Hernández-Bermejo B, Fairén V, Sorribas A. Power-law modeling based on least-squares minimization criteria. *Math Biosci*. 1999;161(1–2):83–94.
- Savageau MA. Biochemical systems analysis: II. The steady-state solutions for an n-pool system using a power-law approximation. *J Theor Biol*. 1969;25(3):370–9.
- Savageau MA. Biochemical systems analysis: III. Dynamic solutions using a power-law approximation. *J Theor Biol*. 1970;26(2):215–26.
- Visser D, Heijnen JJ. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng*. 2003;5(3):164–76.
- Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ. Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng*. 2004;6(4):378–90.
- Liebermeister W, Klipp E. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model*. 2006;3:41.
- Hu CY, Varner JD, Lucks JB. Generating effective models and parameters for RNA genetic circuits. *ACS Synth Biol*. 2015;4(8):914–26.
- Sharma N, Liu YA. A hybrid science-guided machine learning approach for modeling chemical processes: a review. 2022;68(5):e17609.
- Yazdani A, Lu L, Raissi M, Karniadakis GE. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS Comp Biol*. 2020;16(11):e1007575.
- Lee D, Jayaraman A, Kwon JS. Development of a hybrid model for a partially known intracellular signaling pathway through correction term estimation and neural network modeling. *PLoS Comp Biol*. 2020;16(12):e1008472.
- Moles CG, Mendes P, Banga JR. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res*. 2003;13(11):2467–74.
- Coleman T, Branch MA, Grace A. Optimization toolbox. In: For use with MATLAB User's guide for MATLAB, vol. 5. 1999.
- Jones DR, Martins JRR. The DIRECT algorithm: 25 years later. *J Global Optim*. 2021;79(3):521–66.
- Nelder JA, Mead R. A simplex method for function minimization. *Comput J*. 1965;7(4):308–13.
- Hoffmeister F, Bäck T, editors. Genetic algorithms and evolution strategies: similarities and differences. Parallel problem solving from nature 1991. Berlin: Springer Berlin Heidelberg; 1991.
- Saravanan N, Fogel DB, Nelson KM. A comparison of methods for self-adaptation in evolutionary algorithms. *Biosystems*. 1995;36(2):157–66.
- Bäck T. Evolution strategies: an alternative evolutionary algorithm. In: Artificial evolution. Berlin: Springer Berlin Heidelberg; 1996.
- Runarsson TP, Xin Y. Stochastic ranking for constrained evolutionary optimization. *IEEE Trans Evol Comput*. 2000;4(3):284–94.
- Hansen N, Ostermeier A. Completely derandomized self-adaptation in evolution strategies. *Evol Comput*. 2001;9(2):159–95.
- Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. *Science*. 1983;220(4598):671–80.
- Holland JH. Genetic algorithms. *Sci Am*. 1992;267(1):66–73.
- Pincus M. Letter to the Editor—A Monte Carlo method for the approximate solution of certain types of constrained optimization problems. *Oper Res*. 1970;18(6):1225–8.
- Fogel DB. Evolutionary computation: the fossil record. Piscataway: Wiley-IEEE Press; 1998. p. 656.
- Lieberman AR, Kwon SB, Vu HT, Filipowicz A, Ay A, Ingram KK. Circadian clock model supports molecular link between PER3 and Human Anxiety. *Sci Rep*. 2017;7(1):9893.
- Hohm T, Zitzler E. Multicellular pattern formation. *IEEE Eng Med Biol Mag*. 2009;28(4):52–7.
- Fakhouri WD, Ay A, Sayal R, Dresch J, Dayringer E, Arnosti DN. Deciphering a transcriptional regulatory code: modeling short-range repression in the *Drosophila* embryo. *Mol Syst Biol*. 2010;6:341.
- Fomekong-Nanfack Y, Kaandorp JA, Blom J. Efficient parameter estimation for spatio-temporal models of pattern formation: case study of *Drosophila melanogaster*. *Bioinformatics*. 2007;23(24):3356–63.
- Bandodkar P, Shaikh R, Reeves GT. ISRES+: an improved evolutionary strategy for function minimization to estimate the free parameters of systems biology models. *Bioinformatics*. 2023;39(7):btad403.
- Parmar JH, Mendes P. A computational model to understand mouse iron physiology and disease. *PLoS Comput Biol*. 2019;15(1):e1006680.
- Maeda K, Westerhoff HV, Kurata H, Boogerd FC. Ranking network mechanisms by how they fit diverse experiments and deciding on *E. coli*'s ammonium transport and assimilation network. *npj Syst Biol Appl*. 2019;5(1):14.
- Ashyraliyev M, Siggins K, Janssens H, Blom J, Akam M, Jaeger J. Gene circuit analysis of the terminal gap gene *huckebein*. *PLoS Comput Biol*. 2009;5(10):e1000548.
- Jostins L, Jaeger J. Reverse engineering a gene network using an asynchronous parallel evolution strategy. *BMC Syst Biol*. 2010;4(1):17.

48. O'Connell MD, Reeves GT. The presence of nuclear cactus in the early drosophila embryo may extend the dynamic range of the dorsal gradient. *PLoS Comput Biol*. 2015;11(4):e1004159.
49. Wolpert DH, Macready WG. No free lunch theorems for optimization. *IEEE Trans Evol Comput*. 1997;1(1):67–82.
50. English TM. Optimization is easy and learning is hard in the typical function. In: Proceedings of the 2000 Congress on Evolutionary Computation. New York City: La Jolla, CA, USA. IEEE; 2000 16–19 July 2000.
51. Igel C, Toussaint M. A No-Free-Lunch theorem for non-uniform distributions of target functions. *J Math Model Algorithms*. 2005;3(4):313–22.
52. Voit EO. Biochemical systems theory: a review. *ISRN Biomathematics*. 2013;2013:897658.
53. Cornish-Bowden A. Chapter 2 - Introduction to enzyme kinetics. In: Cornish-Bowden A, editor. *Fundamentals of enzyme kinetics*. Oxford: Butterworth-Heinemann; 1979. p. 16–38.
54. Gavin HP, editor *The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems* c ©. 2013.
55. Törn AA, Goldberg W. (editor). *Global optimization as a combination of global and local search. Proceedings of computer simulation versus analytical solutions for business and economic models*. Gothenburg; 1973. <https://web.abo.fi/~atorn/ProbAlg/Page51J.html>.
56. Kan AHGR, Timmer GT. Stochastic global optimization methods. part 1: clustering methods. *Math Program*. 1987;39(1):27–56.
57. Banga JR, Casares JJ, editors. *ICRS: application to a wastewater treatment plant mode*. IChemE Symposium Series No 100. Oxford: Pergamon Press; p. 1987.
58. Goulcher R, Casares Long JJ. The solution of steady-state chemical engineering optimisation problems using a random-search algorithm. *Comput Chem Eng*. 1978;2(1):33–6.
59. David C, Marco D, Fred G, Dipankar D, Pablo M, Riccardo P, et al. New ideas in optimization. In: David C, Marco D, Fred G, Dipankar D, Pablo M, Riccardo P, et al., editors. UK: McGraw-Hill Ltd.; 1999.
60. Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *J Global Optim*. 1997;11(4):341–59.
61. Schwefel HP. Contemporary evolution strategies. In: European conference on artificial life. Heidelberg: Springer Link; 1995. p. 891–907.
62. Runarsson T, Yao X. Search biases in constrained evolutionary optimization. *IEEE Trans Syst Man Cybern C*. 2005;35:233–43.
63. Deb K, Anand A, Joshi D. A computationally efficient evolutionary algorithm for real-parameter optimization. *Evol Comput*. 2002;10(4):371–95.
64. Satoh H, Yamamura M, Kobayashi S. Minimal generation gap model for GAs considering both exploration and exploitation. In: Proceedings of 4th International Conference on Soft Computing, Iizuka. Scientific Research Publishing, Wuhan, PRC. 30 September-5 October 1996. p. 494–7.
65. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*. 1979;21(2):239–45.
66. Shukal S, Chen X, Zhang C. Systematic engineering for high-yield production of viridiflorol and amorphadiene in auxotrophic *Escherichia coli*. *Metab Eng*. 2019;55:170–8.
67. Fritsch FN, Butland J. A method for constructing local monotone piecewise cubic interpolants. *SIAM J Sci Stat Comput*. 1984;5(2):300–4.
68. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36(8):1627–39.
69. Marx A, Backes C, Meese E, Lenhof HP, Keller A. EDISON-WMW: exact dynamic programming solution of the Wilcoxon-Mann-Whitney Test. *Genom Proteom Bioinform*. 2016;14(1):55–61.
70. Yeo HC, Selvarajoo K, Varsheni V. Identifying effective evolutionary strategies-based protocol for uncovering reaction kinetic parameters under the effect of measurement noises. 2024. Zenodo. <https://doi.org/10.5281/zenodo.13788914>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.