Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(6): 10570–10589.
DOI: 10.3934/mbe.2023467
Received: 28 December 2022
Revised: 29 March 2023
Accepted: 30 March 2023
Published: 11 April 2023

Research article

Realistic simulation of time-course measurements in systems biology

Janine Egert[1,2] **and Clemens Kreutz**[1,2,3,*]

[1] Institute of Medical Biometry and Statistics (IMBI), Faculty of Medicine and Medical Center – University of Freiburg, Stefan-Meier-Str. 26, 79104 Freiburg, Germany

[2] Centre for Integrative Biological Signalling Studies CIBSS, University of Freiburg, 79104 Freiburg, Germany

[3] Center for Data Analysis and Modelling (FDM), University of Freiburg, 79104 Freiburg, Germany

* **Correspondence:** Email: clemens.kreutz@uniklinik-freiburg.de; Tel: +4976127083741.

Abstract: In systems biology, the analysis of complex nonlinear systems faces many methodological challenges. For the evaluation and comparison of the performances of novel and competing computational methods, one major bottleneck is the availability of realistic test problems. We present an approach for performing realistic simulation studies for analyses of time course data as they are typically measured in systems biology. Since the design of experiments in practice depends on the process of interest, our approach considers the size and the dynamics of the mathematical model which is intended to be used for the simulation study. To this end, we used 19 published systems biology models with experimental data and evaluated the relationship between model features (e.g., the size and the dynamics) and features of the measurements such as the number and type of observed quantities, the number and the selection of measurement times, and the magnitude of measurement errors. Based on these typical relationships, our novel approach enables suggestions of realistic simulation study designs in the systems biology context and the realistic generation of simulated data for any dynamic model. The approach is demonstrated on three models in detail and its performance is validated on nine models by comparing ODE integration, parameter optimization, and parameter identifiability. The presented approach enables more realistic and less biased benchmark studies and thereby constitutes an important tool for the development of novel methods for dynamic modeling.
Availability: The approach is implemented in the MATLAB-based modelling toolbox Data2Dynamics and available at https://github.com/Data2Dynamics/d2d.

Keywords: systems biology; simulation; method evaluation; benchmark; model dynamics

1. Introduction

Systems biology deals with the mathematical analysis of complex biological processes and networks, often by means of large nonlinear systems. Therefore, many computational issues arise and large differences in performance between different application settings have been found [1–7]. Therefore, an important task is to evaluate the performance of methods and algorithms in both experimental applications and simulation studies.

Databases such as the BioModels database [8] offer a wide range of published biological models and provide equations for the dynamic variables of a desired biological system typically in the SBML format [9]. However, the BioModels database does not provide experimental data and how the dynamic equations are linked to the data. To fill this gap, there are attempts to assemble systems biology models together with the underlying experimental data in a benchmark repository [10, 11]. In addition, a standardized format for systems biology models including experimental information is developed [12]. This PEtab format aids in the validation and benchmarking of computational methods.

So far there are only a few published system biology models with experimental data. In order to enable benchmark studies on a larger set of test problems, we present a simulation approach for dynamic models. Within the presented approach, a realistic experimental design is generated and measurements are drawn to resemble real applications. The resulting data can be used to benchmark computational methods and in particular to investigate how performance varies over different applications and experimental designs.

A reasonable reflection of the experimental design including observables, observation time points, and a realistic data set with measurement errors is suggested. Our results are based on the experimental design of 19 biological models of a benchmark repository [10] and can be applied to a variety of dynamic models e.g., from the BioModels database [8]. While typical properties of the experimental design are provided by the 19 published models, the dynamics are defined by the desired model.

Although simulation studies can not replace experimental data, they are a valuable tool for evaluating different methods and algorithms used for systems analysis. In contrast to experimental data, they are advantageous in that an unlimited number of measurement and analysis repetitions can be simulated. Since the underlying truth of the system is known, additional statistical quantities, such as sensitivity and specificity, can be assessed from simulated data.

2. Materials and method

This work focuses on biological processes that can be modeled by ordinary differential equations (ODE)

$$\dot{x}(t) = f(x(t), u(t), p_x) \tag{2.1}$$

where x denotes the states of the biological system, p_x its parameters, and u the input function representing a stimulus of the system. Function f translates the interaction map of the biological processes into an ODE system. In a biological context, the states x correspond to molecular concentrations and represent the compartments of interest of the biological system. Because not all states x of a system can be measured in an experiment due to biological limitations or cost, often only data from a subset of states are collected. These measured states are defined by the experimenter prior to conducting the

experiment and are referred to as observables

$$y(t_i) = g(x(t_i), p_y) + \epsilon(t_i). \tag{2.2}$$

The function g links the dynamic states x with the measured values of the observable y and its parameters p_y, and may include a logarithmic transformation. The additive term ϵ represents the measurement noise.

2.1. Observables

Typical observation functions g are

$$\begin{aligned} g^{(1)}(x) &= x \\ g^{(2)}(x) &= a + s \cdot x \\ g^{(3)}(x) &= x_1 + x_2 \end{aligned} \tag{2.3}$$

Here, $g^{(1)}$ represents a direct measurement of the observable, $g^{(2)}$ a relative measurement with an offset a and a scaling factor s, and $g^{(3)}$ a combined measurement of two states x_1 and x_2, which in the following is also called a compound measurement. Common measurement techniques in biological applications for measurement type $g^{(2)}$ are e.g., Western-blot, PCR, sequencing and proteomics experiments, and for $g^{(3)}$ e.g., ligand-protein binding without discrimination of phosphorylation status. The measurement technique is chosen by the experimenter depending on measurability, accuracy, and cost of the experiment.

2.2. System dynamics

An essential part of setting up a realistic design is the choice of observation time points. A too short time duration does not capture the long-term dynamics, and larger time duration or larger intervals between the measurement points do not properly capture the fast dynamics. To define a reasonable set of observation time points which fit the model dynamics properly, the time scale of the desired model has to be estimated. Here, we characterized the time scale of the model dynamics by applying the transient function

$$\begin{aligned} f_{\mathrm{fr}}(t) = A_{\mathrm{off}} + A_{\mathrm{sig}} \cdot \Bigg[& A_{\mathrm{sus}} \cdot \left(1 - \mathrm{e}^{-\frac{t(\tau_{\mathrm{offset}})}{\tau_{\mathrm{sus}}}}\right) \\ & + A_{\mathrm{tr}} \cdot \left(1 - \mathrm{e}^{-\frac{t(\tau_{\mathrm{offset}})}{\tau_{\mathrm{sus}}}}\right) \cdot \mathrm{e}^{-\frac{(t_{\mathrm{real}}, \tau_{\mathrm{offset}})}{\tau_{\mathrm{tr}}}} \Bigg] \end{aligned} \tag{2.4}$$

where A_{off} is an additional offset, $A_{\mathrm{sig}} \in 0, 1$ indicates the direction of the response, A_{sus} and A_{tr} are the amplitudes of the sustained and the transient part of the response, and τ_{sus} and τ_{tr} are the time scales indicating the velocity of both parts of the response [13]. Because there are cases where the dynamics start with a delay τ_{offset}, the time predictor is shifted by a nonlinear transformation [13]

$$t(\tau_{\mathrm{offset}}) = \log_{10}(10^{\frac{10 \cdot t_{\mathrm{real}}}{\mathrm{range}(t_{\mathrm{real}})}} + 10^{\tau_{\mathrm{offset}}}) - \log_{10}(1 + 10^{\tau_{\mathrm{offset}}}).$$

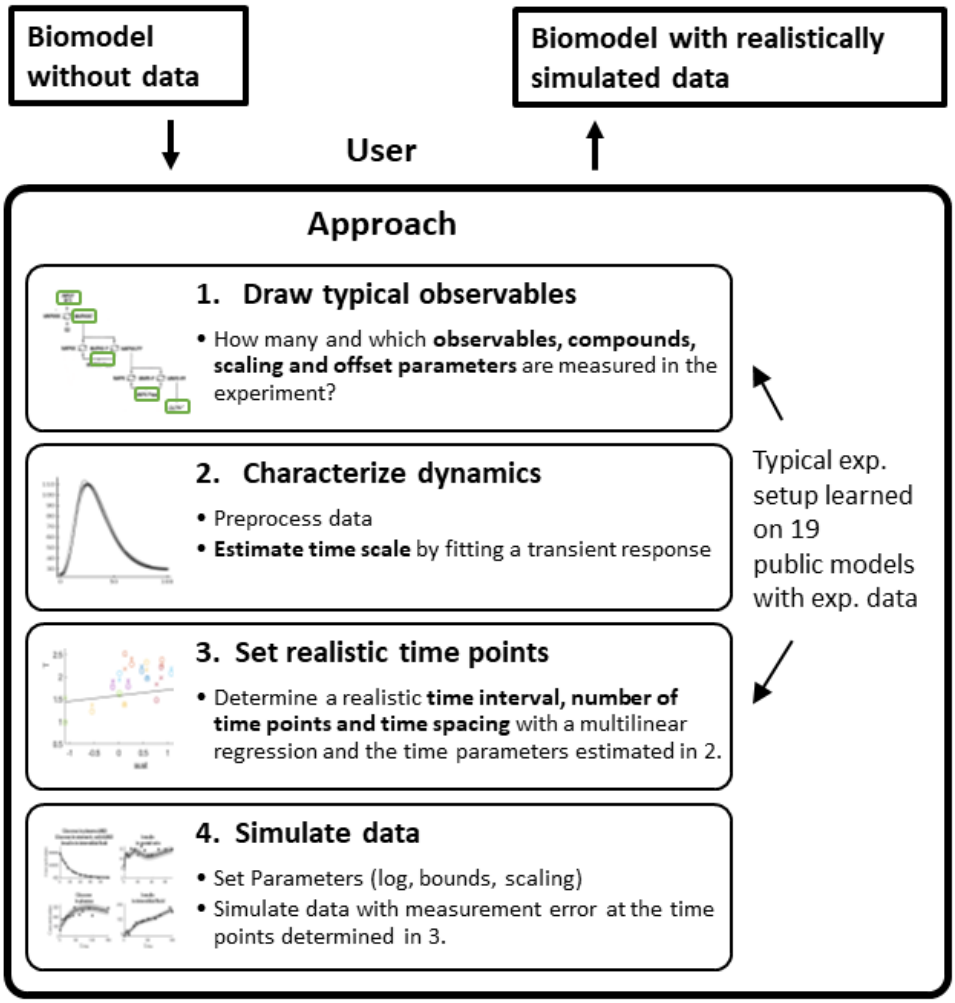


Figure 1. Main steps for setting up a realistic simulation for a desired model.

2.3. Realistic simulation

Biochemical reaction networks which can be modeled using transient dynamics are selected from the Data2Dynamics [14] examples folder, e.g., no oscillations are included. To this end, 19 biological pathway models [15–33] are analyzed for their experimental design. The analysis of the 19 benchmark experiments as well as the implementation of the realistic simulation are described in the following section. The main steps of the implemented algorithm are shown in Figure 1 and explained in more detail below.

2.3.1. Realistic observables and observation functions

The number of observables, relative observation functions, and compound measurements of the 19 published models are depicted in Table 1. In these 19 experiments, on average 50 % of the dynamic states are measured, of which 40 % are measured directly ($g^{(1)}$), 40 % with a scaling factor ($g^{(2)}$) and 20% contribute to a compound measurement ($g^{(3)}$). In order to guarantee a realistic simulation setup,

the number of observables and their observation function are drawn by a binomial draw from Table 1 Columns 4–8 and are then randomly assigned to the dynamic state variables.

Table 1. Frequency of observables and their observation type for the 19 published models.

Model	# Species	# Obs	$\frac{\#\text{Obs}}{\#\text{Species}}$ [%]	$\frac{\#\text{Compound}}{\#\text{Obs}}$ [%]	# Species in Compound	$\frac{\#\text{Scaling}}{\#\text{Obs}}$ [%]	$\frac{\#\text{Offset}}{\#\text{Scaling}}$ [%]	Ref.
Alkan	36	12	33	0	0	83	0	[15]
Bachmann	25	20	80	5	2	50	50	[16]
Becker	6	4	67	0	0	25	100	[17]
Boehm	8	3	38	100	2	0	0	[18]
Brannmark	9	3	33	33	2	0	0	[19]
Bruno	7	6	86	0	0	0	0	[20]
Crauste	5	4	80	0	0	0	0	[21]
Fiedler	6	2	33	0	0	0	0	[22]
Fujita	9	3	33	0	0	100	0	[23]
Hass	61	19	31	47	10.3	79	93	[24]
Isensee	25	3	12	0	0	67	0	[25]
Lucarelli	65	33	51	61	5.3	0	0	[26]
Merkle	23	22	96	5	2	50	82	[27]
Raia	14	8	57	13	3	63	0	[28]
Reelin	9	6	67	100	5	100	100	[29]
Schwen	11	4	36	100	2.8	50	100	[30]
Sobotta	35	18	51	0	0	67	100	[31]
Swameye	14	3	21	0	0	67	0	[32]
Zheng	15	15	100	0	0	0	0	[33]
$\mu\pm$ sd	20±20	10±9	50±30	20±40	2±3	40±40	20±30	

2.3.2. Observation time points

The choice of observation time points, i.e., the time duration and time spacing, play an important role in creating a realistic design. We observed that experimentally chosen time spacing in the published models commonly fall between linearly and exponentially distributed time intervals. Thus, we estimate the spacing of time points as:

$$t_{obs} = [\lambda \cdot x + (1 - \lambda) \cdot x \cdot (e^{\ln(2) \cdot x} - 1)^p] \cdot T \tag{2.5}$$

with N equidistantly distributed points in $x \in [0, 1]$, the total time length T, the linearity λ of the time spacing and the exponent p of the exponential time spacing. The influence of the linearity λ and the exponent p on the time spacing is visualized in Supplementary Figure S2. A multivariate linear

regression

$$\begin{pmatrix} \log(T) \\ \log(N) \\ \lambda \\ \log(p) \end{pmatrix} = \begin{pmatrix} \theta_{1T} & \theta_{2T} & ... & \theta_{9T} \\ \theta_{1N} & \theta_{2N} & ... & \theta_{9N} \\ \theta_{1\lambda} & \theta_{2\lambda} & ... & \theta_{9\lambda} \\ \theta_{1p} & \theta_{2p} & ... & \theta_{9p} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \log(\tau_{\text{sus}}) \\ \log(\tau_{\text{tr}}) \\ \log(\tau_{\text{sus}}) \cdot \log(\tau_{\text{tr}}) \\ \log(\tau_{\text{offset}}) \\ \log(A_{\text{sus}}) \\ \log(A_{\text{tr}}) \\ n_{\text{states}} \\ n_{\text{obs}} \end{pmatrix} \tag{2.6}$$

is performed between the characteristics N, T, λ, and p of the measurement times chosen in the experiment and the dynamics of the underlying system characterized by the parameters of the transient function f_{tr}. The offset A_{off} and the signum A_{sig} of the transient function are not used as predictors for the estimation of the time characteristics. Because we assume that an offset in the data and the direction of the response has no effect on the choice of the experimental measurement times. Because the number of observed quantities n_{obs} and the size of the model n_{states} might have an impact on chosen measurement times in practice, they are included as predictors in the multivariate linear regression. The time characteristics and the time predictors are analyzed on a logarithmic scale except for the linearity λ which lies between [0,1]. The multivariate linear regression is applied with intercepts $\theta_{1T}, \theta_{1N}, \theta_{1\lambda}$, and θ_{1p}. The estimated regression coefficients $\hat{\theta}$ of Eq 2.6 are shown in Supplementary Table S1. For the 4 time characteristics and 9 predictor variables, there are 36 regression parameters in total. Figure 2 shows the result of the regression analysis for the two most significant predictors for each time characteristic. All 4 time characteristics predicted by the multivariate linear regression show a strong positive correlation with the experimentally chosen time characteristics, with a Pearson correlation coefficient ρ between 0.67 and 0.93 (Figure 2).

Note that our analysis has been performed to be independent of physical units by considering that the time scale of the model dynamics defines the time constants τ and thus, the time characteristics, independently of their units.

2.3.3. Data simulation and measurement error

Because experimental data in molecular biology is typically log-normally distributed [34], the simulated data points

$$y(t_{obs}) \propto \text{Lognormal}(\mu, \text{sd}^2) \tag{2.7}$$

at the suggested observation times t_{obs} are drawn from a log-normal distribution, where the mean μ is given by the expectation value of the model dynamics. Due to different experimental techniques and measurement methods, the inter-model variation of the measurement error is usually larger than the intra-model variation. Therefore, we assume a linear mixed-effects model on a logarithmic scale to estimate the error parameters:

$$\text{sd} = \eta_1 + (\eta_2 \,|\, \text{Model}) + \epsilon \tag{2.8}$$

with $i = 1, 2, ..., n_{\text{obs}}$, a fixed effect η_1, a model-dependent random effect $\eta_2 \sim N(0, \sigma_m^2)$ and a Gaussian distributed term $\epsilon \sim \mathcal{N}(0, \sigma_g^2)$ that accounts for other effects on the magnitude of measurement errors. The estimated parameters for the 19 published models are $\eta_1 = -1.12$, $\sigma_m = 0.25$ and $\sigma_g = 0.015$.

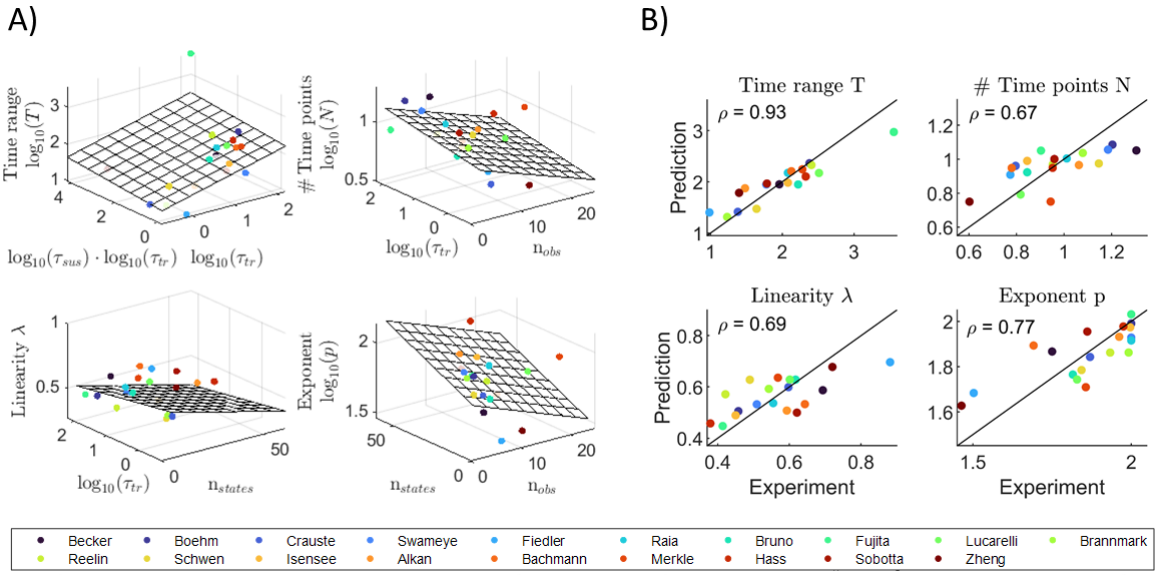


Figure 2. The experimentally chosen measurement time interval T, the number of time points N, the linearity λ, and the exponent p are compared with the predictions. Each dot represents one of the 19 published models listed in the legend below. **A)** The results of the multivariate linear regression is shown for the two most significant predictors for each time characteristic. **B)** The predicted time characteristics are plotted against the experimental time characteristics. A perfect agreement would be on the diagonal black line. The Pearson correlation coefficient ρ is greater than 0.6 for all four time characteristics.

This means that the errors are on average 7.8 %. An example of a realistic simulation of measurements is shown in Figure 3C and in Supplementary Figures S2 and S3.

2.4. Validation criteria

To validate how realistic the described simulation of measurements actually is, the realistic simulation is compared with a naive simulation, an unrealistic simulation, and the published model itself with the measured experimental data. The measurement characteristics of the naive and unrealistic simulation are displayed in Table 2.

The simulations are compared to the published model using the root mean square deviation

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{n_{\mathrm{simu}}} (\mathrm{criterium}_{\mathrm{simu,i}} - \mathrm{criterium}_{\mathrm{benchmark}})^2}{n_{\mathrm{simu}}}}$$

of 5 different validation criteria. The validation criteria are chosen to evaluate the parameter identifiability, integration of the ODEs, and parameter optimization performance of the model. For the least squares optimization

$$\hat{p}_{LS} = \arg\min \chi^2(p|y) \tag{2.9}$$

$$\chi^2(p|y) = \sum_i \frac{(y_i - g_i(p))^2}{\sigma_i^2(p)}, \tag{2.10}$$

Table 2. Measurement characteristics for the realistic, naive, and unrealistic simulation. To validate how realistic the described implementation is, it is compared to a naive and unrealistic simulation with fixed numbers of directly measured observables, data points, and measurement error, with linear time spacing and naively chosen time duration.

Measurement characteristics	Realistic design	Naive design	Unrealistic design
# Observables	Binomial draw [12–100] %	50 %	100 %
Observation type	g^1, g^2, g^3	g^1	g^1
# Data points	Multi linear regression	10 per obs.	100 per obs.
Measurement error	Mixed effects model	5 %	No error
Time spacing	Linear & exponential	Linear	Linear
Time duration	Multi linear regression	rounded to $10^n, n \in \mathcal{N}$	Longest T of all obs.

the number of fits n_{optfits} in the global optimum is evaluated by:

$$n_{\text{optfits}} = \sum_{i=1}^{n_{fits}} 1 \quad \text{if } (\chi_i^2 - \min(\chi^2)) < 0.01 \tag{2.11}$$

with n_{fits} being the number of fits performed for multi-start optimization using different random initial guesses. To evaluate the identifiability, the number of non-identifiable parameters is assessed with the identifiability test by radial penalization [35]. Further, the number of function evaluations of the converged fits, the computation time for fitting, and the number of steps for the ODE integration are compared.

2.5. Implementation

The algorithm for simulating realistic designs and experimental data is implemented in the MATLAB based modeling framework Data2Dynamics [14]. To import the model and export the realistic simulation in the standardized SBML [9] and PEtab [12] formats, the MATLAB code including the Data2Dynamics toolbox is as follows:

```
SBML2Model
arRealisticDesign
arExportPEtab
```

By default, the dynamic parameters of the ODE system are set to a logarithmic scale to improve the convergence of the optimization [2, 4] and the parameter bounds are set to four orders of magnitude.

3. Results

3.1. Application example

In the following, we demonstrate the general aspects of the algorithm on the 'Meal Simulation Model of the Glucose-Insulin System' by Dalla Man, Rizza and Cobelli, 2007 [36] which describes the physiology after meal uptake. It describes the concentration of insulin and glucose in different

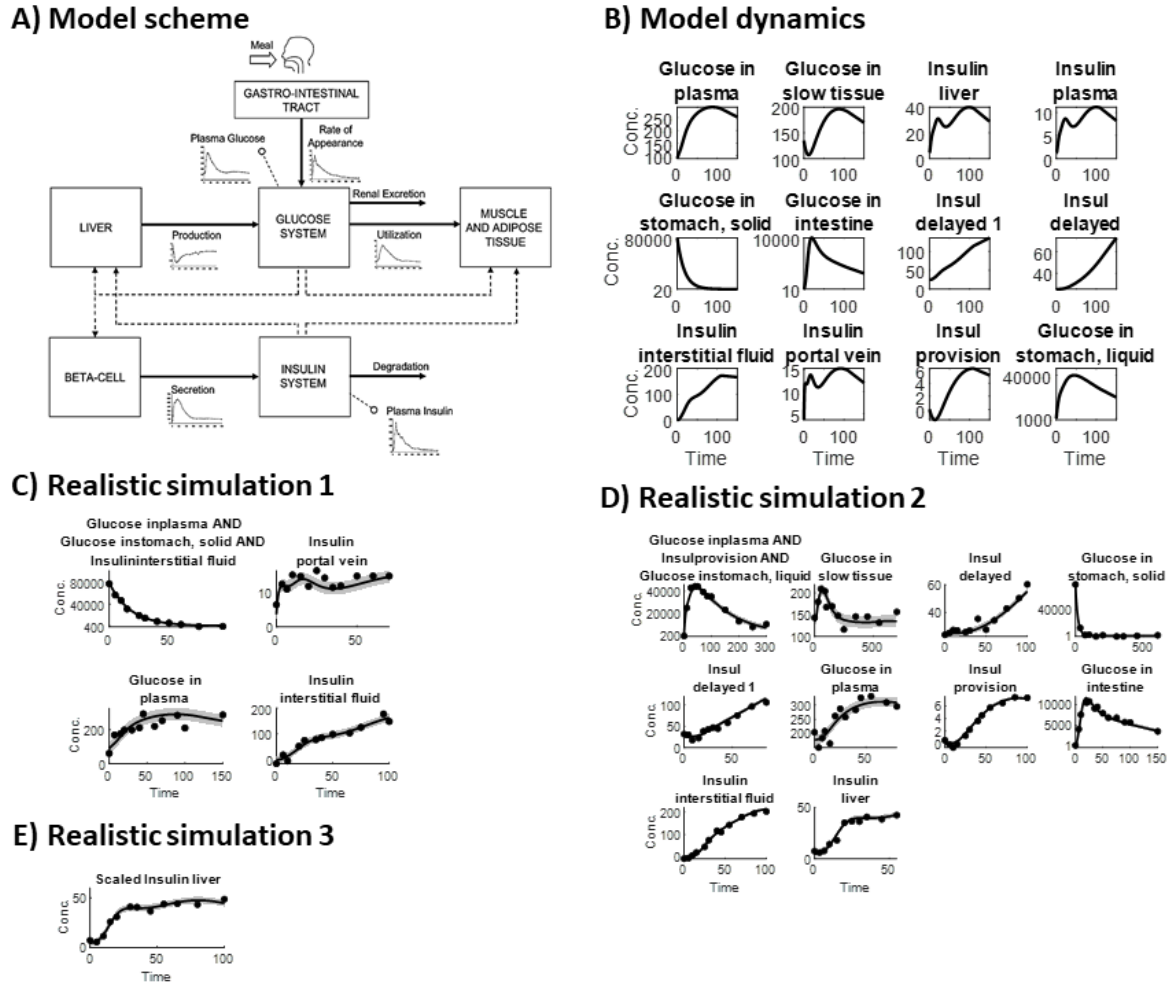


Figure 3. Example of creating a realistic design for a desired model. A) Model scheme of the glucose-insulin application example [36]. B) Model dynamics of the 12 glucose and insulin states. C) Simulation of a realistic measurement of the glucose-insulin model [36]. The model dynamics and simulated time and data points for the drawn observables are depicted.

compartments. Here, glucose and insulin fluxes are taken from a human database. The model scheme and the 12 state dynamics of the glucose-insulin system are shown in Figure 3A and B.

According to the main steps of the implemented algorithm displayed in Figure 1, first, the number of observables and their observation type are set by a binomial draw according to Table 1 and as a subset of the given states. In the examples shown in Figure 3C and D, there are 4 and 10 observables each and one compound consisting of three states. Our approach for generating realistic designs can also draw a compound observation function consisting of more states as illustrated in S3E. In the third example of Figure 3E, only a single observable has been drawn which is measured on a relative scale and, thus, has a scaling factor. After drawing the observables, the model is characterized for the observed dynamic states using the transient function (Eq 2.4). With Eqs 2.4 and 2.6, the total time range T, the number of data points N, the linearity λ of the time spacing, and the exponent p of the exponential time spacing is examined. The simulated measurement time points consist of N logarithmically spaced time values

between $t_0 = 0$ and T. The data is simulated at the suggested time points with a log-normal distribution whose mean is given by the model dynamics and the standard deviation is given by the error model (Eq 2.8). Three examples of simulated data are shown in Figure 3C, D and E. Further application examples of a mitogen-activated protein kinase (MAPK) cascade by Huang and Ferrell, 1996 [37] and of an Epidermal Growth Factor Receptor (EGFR) signaling by Kholodenko et al, 1999 [38] are shown in the Supplementary Figures S2 and S3.

3.2. Validation

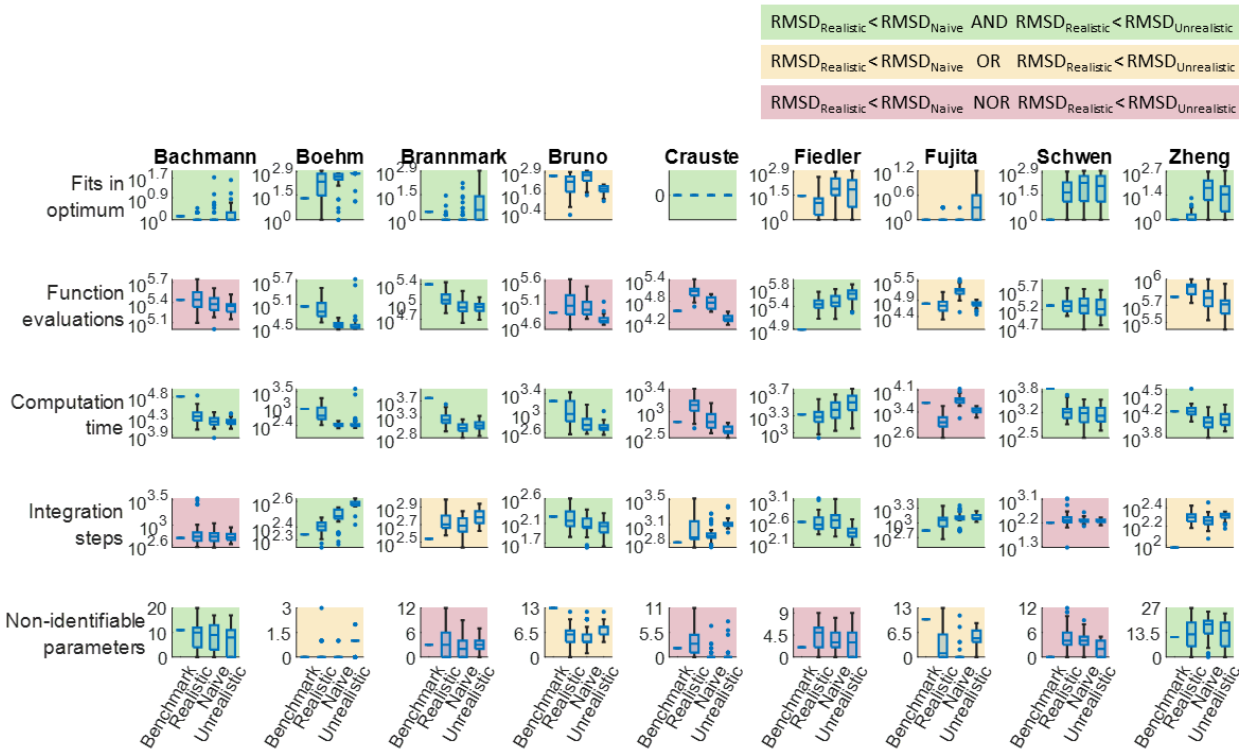


Figure 4. The optimization and integration performance for 9 published models and the realistic, naive, and unrealistic simulations for the 9 published models are shown as boxplots. Each boxplot comprises 50 simulations. To determine how realistic the simulations are, the optimization and integration performance should be close to the published model. This is illustrated by comparing the RMSD of the realistic simulation with the benchmark performance. If the realistic designs are closer to the published model in terms of the RMSD than the naive and unrealistic designs, the figure is highlighted in green. If the realistic designs are closer to the published model than the naive or unrealistic designs, the figure is highlighted in yellow. If the realistic designs deviate more from the published model than the naive and unrealistic designs, the figure is highlighted in red, which is the case for 11 of 45 comparisons.

Simulation studies for benchmarking should provide results that are as similar as possible to analyses based on real data. To validate how realistic the presented approach for the simulation of systems

biology measurements is, we evaluate the realistic simulations on 9 published models. To ensure implementation and data consistency, the 9 models are selected such that they support the PEtab import in the Data2Dynamics toolbox. To assess the optimization and integration performance, we compared 50 realistic, 50 naive and 50 unrealistic designs against the published model by evaluating the 5 different validation criteria introduced in section 2.4 using 1000 multi start optimization runs for each model and simulation.

The performance of the 5 validation criteria for the realistic simulation, naive simulation, unrealistic simulation, and the published model is shown in Figure 4. In most cases, the realistic simulation performs better than the naive and unrealistic designs, i.e., it is closer to the benchmark performance. For 8 of the 9 models, the median number of fits in the optimum of the realistic simulation is closest to the published model (all except Bruno [20]), and for 6 models also the RMSD is closest to the published model (green). In the case of integration steps, the RMSD of the realistic simulation is the smallest only in 4 out of 9 models. However, in the remaining 5 cases, there are no major deviations, and the performance of the realistic simulation is comparable to that of the naive and unrealistic design. For the optimization process, the realistic simulation generally requires more function evaluations and more computation time than the naive and unrealistic design (downward trend, except for Fiedler [22] and Fujita [23]). Moreover, the number of function evaluations and the computation time depend strongly on the number of parameters. For instance, in the case of the model by Crauste et al. [21], the number of parameters is highest for the unrealistic design, resulting in fewer converged fits and, thus, fewer function evaluations and a lower computation time. For the number of non-identifiable parameters (Figure 4) the realistic simulation deviates the most from the published model. The impact of the choice of measurement time points and observation type on the validation criteria is discussed in the supplement (Supplementary Figure S4 and S5).

4. Conclusions

In this work, we present an approach for creating realistic experimental designs and for simulating realistic data for any biological model. Many dynamic models of biological systems are publicly available but lack the experimental data, which is crucial for evaluating and benchmarking methods. Our approach fills this gap by providing a realistic setup for a given dynamic model, including observed quantities, observation time, number of data points, and measurement errors. This realistic simulation setup can be used as a basis for simulation studies to assess and improve methodological and computational procedures.

The analyzed models show large deviations in the number of observables, their observation type, and also in estimating the observation time points with the standard error having the same order of magnitude as the mean. These variations are the key to the realistic diversity of biological experiments. By using a binomial draw for these measurement characteristics, the realistic simulations reflect this diversity in real applications. Still, the observation types of the observables depend strongly on the studied system and the possible measurement technique. The time duration, time spacing, and number of data points for the suggested observables are set by the experimenter and thus vary between projects. Yet, the multivariate regression provides a good representation of these quantities.

The validation of the realistic simulation is challenging because the observables and measurement characteristics are drawn randomly and therefore vary widely, making the simulated experiments not

directly comparable to the single realization of the published model.

The variety of designs in the published models described above actually calls for more than 19 examples to define a typical experimental design. Also, more models would allow a better evaluation. In this work, we selected just one experimental condition for each model, i.e., the condition with the most species. Further studies with a more detailed representation could potentially improve our results. Possible refinements of a realistic model design are e.g., including repetition measurements, including multiple conditions such as genetic perturbations or stimulating with different doses, or a more representative selection of compound observables (which are more likely to be measured as a compound). Here, only a limited number of published models with data and comprehensive documentation of how the data is linked to the model, have been available for the analysis of a realistic design. The realistic reflection of an experimental design obtained in this work is particularly promising for signaling pathway models. However, the results can also be applied to pharmacokinetic, curated, non-curated, metabolic, disease-related, or other dynamic models.

The results and algorithm presented in this study represent a realistic reflection of dynamic biological models. This realistic representation is essential for simulation studies to depict real-world applications and to evaluate methodological and computational issues in a realistic biological setting. The realistic simulation approach is also beneficial for benchmark studies because it takes away the freedom and dependence on the choice of simulation setup. resulting in less bias.

Acknowledgments

This work was supported by the Federal Ministry of Education and Research of Germany [EA:Sys,FKZ031L0080 to J.E. and C.K.]; and the Excellence Initiative of the German Federal and State Governments [CIBSS-EXC-2189-2100249960-390939984 to C.K.].

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. A. Degasperi, D. Fey, B. N. Kholodenko, Performance of objective functions and optimisation procedures for parameter estimation in system biology models, *npj Syst. Biol. Appl.*, **3** (2017). https://doi.org/10.1038/s41540-017-0023-2
2. C. Kreutz, New Concepts for Evaluating the Performance of Computational Methods, *IFAC-PapersOnLine*, **49** (2016), 63–70. https://doi.org/10.1016/j.ifacol.2016.12.104
3. R. J. Prill, D. Marbach, J. Saez-Rodriguez, P. K. Sorger, L. G. Alexopoulos, X. Xue, et al., Towards a Rigorous Assessment of Systems Biology Models: The DREAM3 Challenges, *PLoS ONE*, **5** (2010), e9202. https://doi.org/10.1371/journal.pone.0009202
4. A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, et al., Lessons Learned from Quantitative Dynamical Modeling in Systems Biology, *PLoS ONE*, **8** (2013), e74335. https://doi.org/10.1371/journal.pone.0074335

5. P. Städter, Y. Schälte, L. Schmiester, J. Hasenauer, P. L. Stapor, Benchmarking of numerical integration methods for ODE models of biological systems, *Sci. Rep.*, **11** (2021), 2969. https://doi.org/10.1038/s41598-021-82196-2
6. P. Stapor, F. Fröhlich, J. Hasenauer, Optimization and profile calculation of ODE models using second order adjoint sensitivity analysis, *Bioinformatics*, **34** (2018), i151–i159. https://doi.org/10.1093/bioinformatics/bty230
7. A. F. Villaverde, F. Fröhlich, D. Weindl, J. Hasenauer, J. R. Banga, Benchmarking optimization methods for parameter estimation in large kinetic models, *Bioinformatics*, **35** (2018), 830–838. https://doi.org/10.1093/bioinformatics/bty736
8. N. Le Novere, B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, et al., BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems, *Nucleic Acids Res.*, **34** (2006), D689–D691. https://doi.org/10.1093/nar/gkj092
9. M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, et al., The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models, *Bioinformatics*, **19** (2003), 524–531. https://doi.org/10.1093/bioinformatics/btg015
10. H. Hass, C. Loos, E. Raimundez-Alvarez, J. Timmer, J. Hasenauer, C. Kreutz, Benchmark problems for dynamic modeling of intracellular processes, *Bioinformatics*, **35** (2019), 3073–3082. https://doi.org/10.1093/bioinformatics/btz020
11. A. F. Villaverde, D. Henriques, K. Smallbone, S. Bongard, J. Schmid, D. Cicin-Sain, et al., BioPreDyn-bench: a suite of benchmark problems for dynamic modelling in systems biology, *BMC Syst. Biol.*, **9** (2015). https://doi.org/10.1186/s12918-015-0144-4
12. L. Schmiester, Y. Schälte, F. T. Bergmann, T. Camba, E. Dudkin, J. Egert, et al., PEtab–Interoperable specification of parameter estimation problems in systems biology, *PLoS Comput. Biol.*, **17** (2021), e1008646. https://doi.org/10.1371/journal.pcbi.1008646
13. C. Kreutz, A New Approximation Approach for Transient Differential Equation Models, *Front. Phys.*, **8** (2020), 1–14. https://doi.org/10.3389/fphy.2020.00070
14. A. Raue, B. Steiert, M. Schelker, C. Kreutz, T. Maiwald, H. Hass, et al., Data2Dynamics: a modeling environment tailored to parameter estimation in dynamical systems, *Bioinformatics*, **31** (2015), 3558–3560. https://doi.org/10.1093/bioinformatics/btv405
15. O. Alkan, B. Schoeberl, M. Shah, A. Koshkaryev, T. Heinemann, D. C. Drummond, et al., Modeling chemotherapy-induced stress to identify rational combination therapies in the DNA damage response pathway, *Sci. Signal.*, **11** (2018), eaat0229. https://doi.org/10.1126/scisignal.aat0229
16. J. Bachman, A. Raue, M. Schilling, M. E. Böhm, C. Kreutz, D. Kaschek, et al., Division of labor by dual feedback regulators controls JAK2/STAT5 signaling over broad ligand range, *Mol. Syst. Biol.*, **7** (2011), 516. https://doi.org/10.1038/msb.2011.50
17. V. Becker, M. Schilling, J. Bachmann, U. Baumann, A. Raue, T. Maiwald, et al., Covering a broad dynamic range: information processing at the erythropoietin receptor, *Science*, **328** (2010), 1404–1408, https://doi.org/10.1126/science.1184913

18. M. E. B. Boehm, L. Adlung, M. Schilling, S. Roth, U. Klingmüller, W. D. Lehmann, Identification of Isoform-Specific Dynamics in Phosphorylation-Dependent STAT5 Dimerization by Quantitative Mass Spectrometry and Mathematical Modeling, *J. Proteome Res.*, **13** (2014), 5685–5694. https://doi.org/10.1021/pr5006923

19. C. Brännmark, R. Palmer, S. T. Glad, G. Cedersund, P. Strålfors, Mass and Information Feedbacks through Receptor Endocytosis Govern Insulin Signaling as Revealed Using a Parameter-free Modeling Framework*, *J. Biol. Chem.*, **285** (2010), 20171–20179. https://doi.org/10.1074/jbc.M110.106849

20. M. Bruno, J. Koschmieder, F. Wuest, P. Schaub, M. Fehling-Kaschek, J. Timmer, et al., Enzymatic study on AtCCD4 and AtCCD7 and their potential to form acyclic regulatory metabolites, *J. Exp. Bot.*, **67** (2016), 5993–6005. https://doi.org/10.1093/jxb/erw356

21. F. Crauste, J. Mafille, L. Boucinha, S. Djebali, O. Gandrillon, J. Marvel, et al., Identification of Nascent Memory CD8 T Cells and Modeling of Their Ontogeny, *Cell Syst.*, **4** (2017), 306–317. https://doi.org/10.1016/j.cels.2017.01.014

22. A. Fiedler, S. Raeth, F. Theis, A. Hausser, J. Hasenauer, Tailored parameter optimization methods for ordinary differential equation models with steady-state constraints, *BMC Syst. Biol.*, **10** (2016). https://doi.org/10.1186/s12918-016-0319-7

23. K. A. Fujita, Y. Toyoshima, S. Uda, Y. ichi Ozaki, H. Kubota, S. Kuroda, Decoupling of Receptor and Downstream Signals in the Akt Pathway by Its Low-Pass Filter Characteristics, *Sci. Signal.*, **3** (2010), ra56. https://www.science.org/doi/abs/10.1126/scisignal.2000810

24. H. Hass, K. Masson, S. Wohlgemuth, V. Paragas, J. E. Allen, M. Sevecka, et al., Predicting ligand-dependent tumors from multi-dimensional signaling features, *npj Syst. Biol. Appl.*, **3** (2017). https://doi.org/10.1038/s41540-017-0030-3

25. J. Isensee, M. Kaufholz, M. J. Knape, J. Hasenauer, H. Hammerich, H. Gonczarowska-Jorge, et al., PKA-RII subunit phosphorylation precedes activation by cAMP and regulates activity termination, *J. Cell Biol.*, **217** (2018), 2167–2184. https://doi.org/10.1083/jcb.201708053

26. P. Lucarelli, M. Schilling, C. Kreutz, A. Vlasov, M. E. Boehm, N. Iwamoto, et al., Resolving the Combinatorial Complexity of Smad Protein Complex Formation and Its Link to Gene Expression, *Cell Syst.*, **6** (2018), 75–89.e11. https://doi.org/10.1016/j.cels.2017.11.010

27. R. Merkle, B. Steiert, F. Salopiata, S. Depner, A. Raue, N. Iwamoto, et al., Identification of Cell Type-Specific Differences in Erythropoietin Receptor Signaling in Primary Erythroid and Lung Cancer Cells, *PLoS Comput. Biol.*, **12** (2016), e1005049. https://doi.org/10.1371/journal.pcbi.1005049

28. V. Raia, M. Schilling, M. Böhm, B. Hahn, A. Kowarsch, A. Raue, et al., Dynamic Mathematical Modeling of IL13-Induced Signaling in Hodgkin and Primary Mediastinal B-Cell Lymphoma Allows Prediction of Therapeutic Targets, *Cancer Res.*, **71** (2011), 693–704. https://doi.org/10.1158/0008-5472.CAN-10-2987

29. H. Hass, F. Kipkeew, A. Gauhar, E. Bouche, P. May, J. Timmer, et al., Mathematical model of early Reelin-induced Src family kinase-mediated signaling, *PLoS ONE*, **12** (2017), e0186927. https://doi.org/10.1371/journal.pone.0186927

30. L. O. Schwen, A. Schenk, C. Kreutz, J. Timmer, M. M. Rodriguez, L. Kuepfer, et al., Representative Sinusoids for Hepatic Four-Scale Pharmacokinetics Simulations, *PLoS ONE*, **10** (2015), e0133653. https://doi.org/10.1371/journal.pone.0133653

31. S. Sobotta, A. Raue, X. Huang, J. Vanlier, A. Jünger, S. Bohl, et al., Model Based Targeting of IL-6-Induced Inflammatory Responses in Cultured Primary Hepatocytes to Improve Application of the JAK Inhibitor Ruxolitinib, *Front. Physiol.*, **8** (2017), 775. https://doi.org/10.3389/fphys.2017.00775

32. I. Swameye, T. G. Müller, J. Timmer, O. Sandra, U. Klingmüller, Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling, *Proc. Natl. Acad. Sci. U.S.A.*, **100** (2003), 1028–1033. https://doi.org/10.1073/pnas.0237333100

33. Y. Zheng, S. M. M. Sweet, R. Popovic, E. Martinez-Garcia, J. D. Tipton, P. M. Thomas, et al., Total kinetic analysis reveals how combinatorial methylation patterns are established on lysines 27 and 36 of histone H3, *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012), 13549–13554. https://doi.org/doi/abs/10.1073/pnas.1205707109

34. E. Limpert, W. A. Stahel, M. Abbt, Log-normal Distributions across the Sciences: Keys and Clues: On the charms of statistics, and how mechanical models resembling gambling machines offer a link to a handy way to characterize log-normal distributions, which can provide deeper insight into variability and probability–normal or log-normal: That is the question, *BioScience*, **51** (2001), 341–352. https://doi.org/10.1641/0006-3568(2001)051[0341:LNDATS]2.0.CO;2

35. C. Kreutz, An easy and efficient approach for testing identifiability, *Bioinformatics*, **34** (2018), 1913–1921. https://doi.org/10.1093/bioinformatics/bty035

36. C. Dalla Man, R. Rizza, C. Cobelli, Meal Simulation Model of the Glucose-Insulin System, *IEEE Trans. Biomed. Eng.*, **54** (2007), 1740–1749. https://doi.org/10.1109/TBME.2007.893506

37. C.-Y. F. Huang, J. E. Ferrell, Ultrasensitivity in the mitogen-activated protein kinase cascade, *Proc. Natl. Acad. Sci.*, **93** (1996), 10078–10083. https://doi.org/10.1073/pnas.93.19.1007

38. B. N. Kholodenko, O. V. Demin, G. Moehren, J. B. Hoek, Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor, *J. Biol. Chem.*, **274** (1999), 30169–30181. https://doi.org/10.1074/jbc.274.42.30169

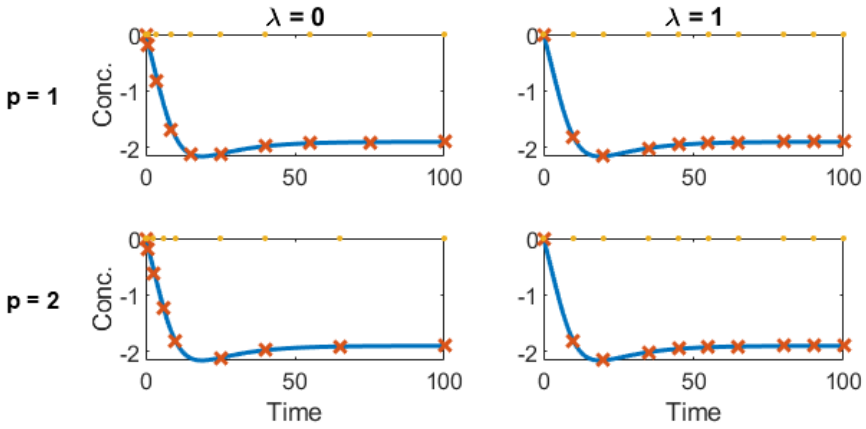
Supplementary

Multivariate linear regression for the measurement characteristics

For a realistic choice of observation time points, four time characteristics, the time duration T, the number of data points N, the linearity λ of the time spacing, and the exponent p of the exponential time spacing are predicted using Eq 2.6. The influence of the linearity λ and the exponent p is visualized in Supplementary Figure S1. With $\lambda = 1$, Eq 2.6 is linear and the time points are linearly distributed except for rounding. In this case, the exponent p has no effect on the sampling times (Figure S1 right). For $\lambda \neq 1$, the time points are exponentially spaced. with smaller distances between sampling times in the beginning and larger distances at the end. The estimated coefficient of the multivariate regression model after backward selection are shown in Supplementary Table S1.

Supplementary Table S1. Estimated parameters $\hat{\theta}$ of the multivariate linear regression in Eq 2.6 of the manuscript, for the overall time range T, the number of data points N, the linearity λ of the time spacing and the exponent p of the exponential time spacing and 9 predictor variables. Backward selection was performed using the likelihood ratio test, the parameters marked with (-) are not signifant.

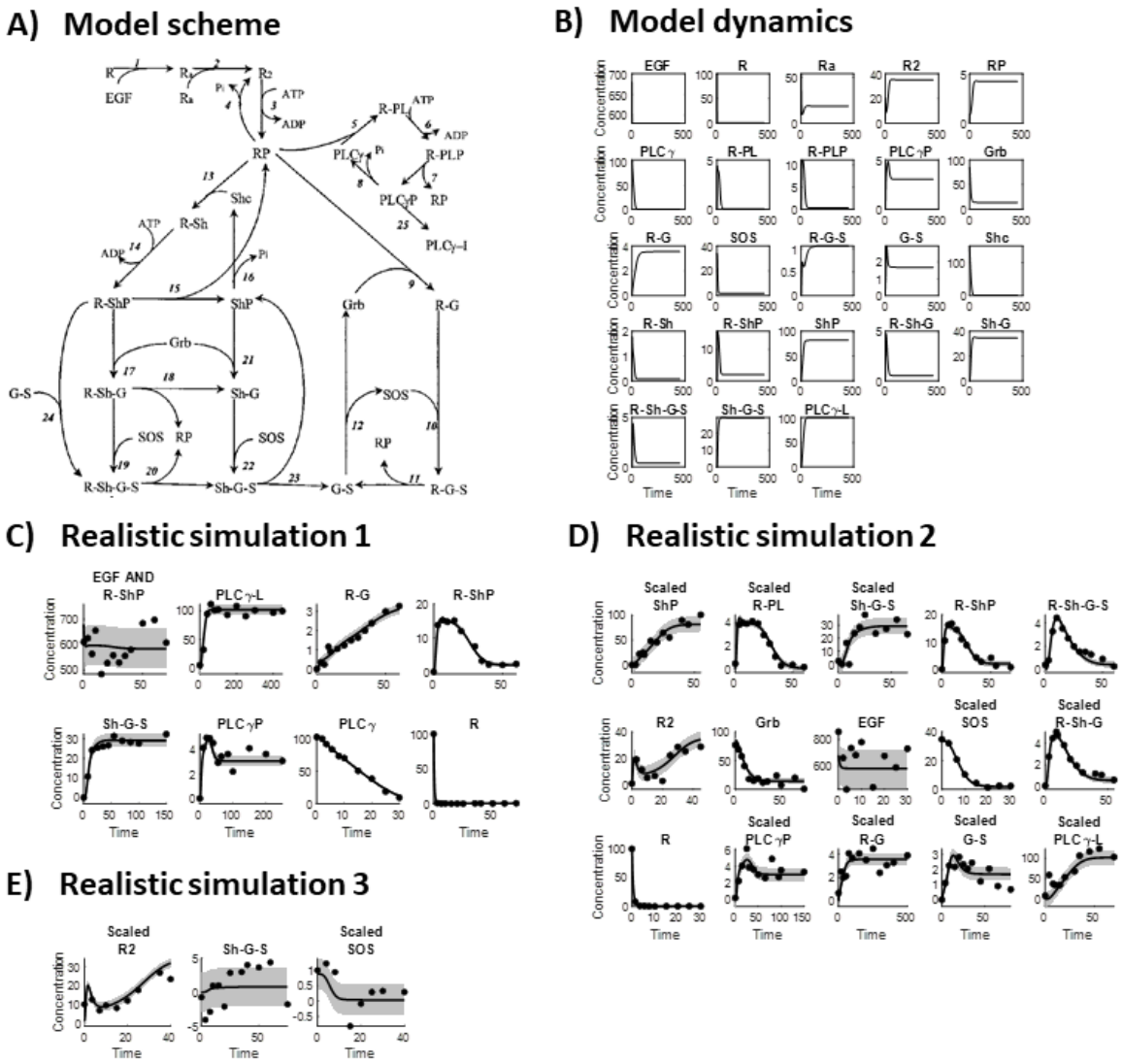
	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$	$\hat{\theta}_7$	$\hat{\theta}_8$	$\hat{\theta}_9$
Predictor	offset	τ_{sus}	τ_{tr}	$\tau_{\text{sus}} \cdot \tau_{\text{tr}}$	τ_{offset}	A_{sus}	A_{tr}	n_{states}	n_{obs}
T	1.58	0.13	0.28	0.042	0.00151	-	-0.019	-0.0046	0.016
N	1.02	0.037	0.059	-0.03	-0.00029	-	0.015	0.0016	-0.014
λ	0.616	-0.014	-0.044	-	-0.00015	-0.006	-0.006	-0.0039	0.0063
p	1.73	0.08	0.098	-0.035	-	-	0.031	0.0073	-0.013



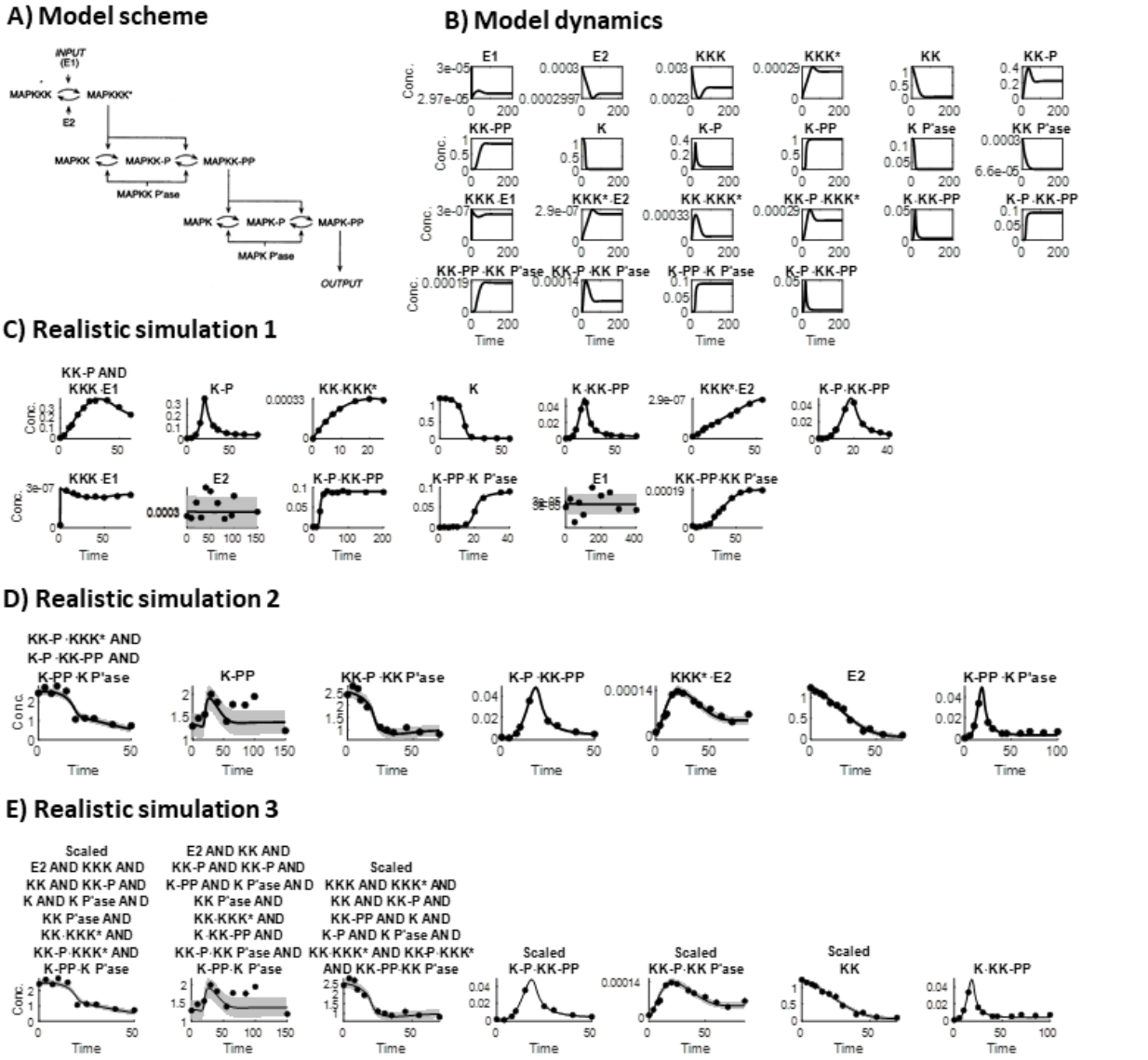
Supplementary Figure S1. The influence of the linearity λ and the exponent p on the measurement time points is demonstrated. On an exemplary model dynamics (blue), the measurement time points (red crosses and yellow dots) are calculated with Eq 2.6, $T = 100$, and $N = 10$. For $\lambda = 1$ the time points are linearly spaced and the exponent p has no influence (right column). For $\lambda < 1$, the time points are exponentially spaced. For larger p, the time intervals are shorter at the beginning and longer at the end.

Application examples

Supplementary Figure S2 shows additional examples for realistic designs of the Epidermal Growth Factor Receptor (EGFR) short term signaling by Kholodenko et al, 1999 [38] and Supplementary Figure S3 shows examples for realistic designs of the mitogen-activated protein kinase (MAPK) cascade by Huang and Ferrell, 1996 [37].



Supplementary Figure S2. Application example of an EGFR pathway model [38] using the realistic simulation approach presented in the manuscript. A) The model scheme of the EGFR short term signaling [37]. B) Model dynamics with the equations and parameters from the BioModels database. C) and D) and E) show three examples of data generated with different realistic simulation setups.



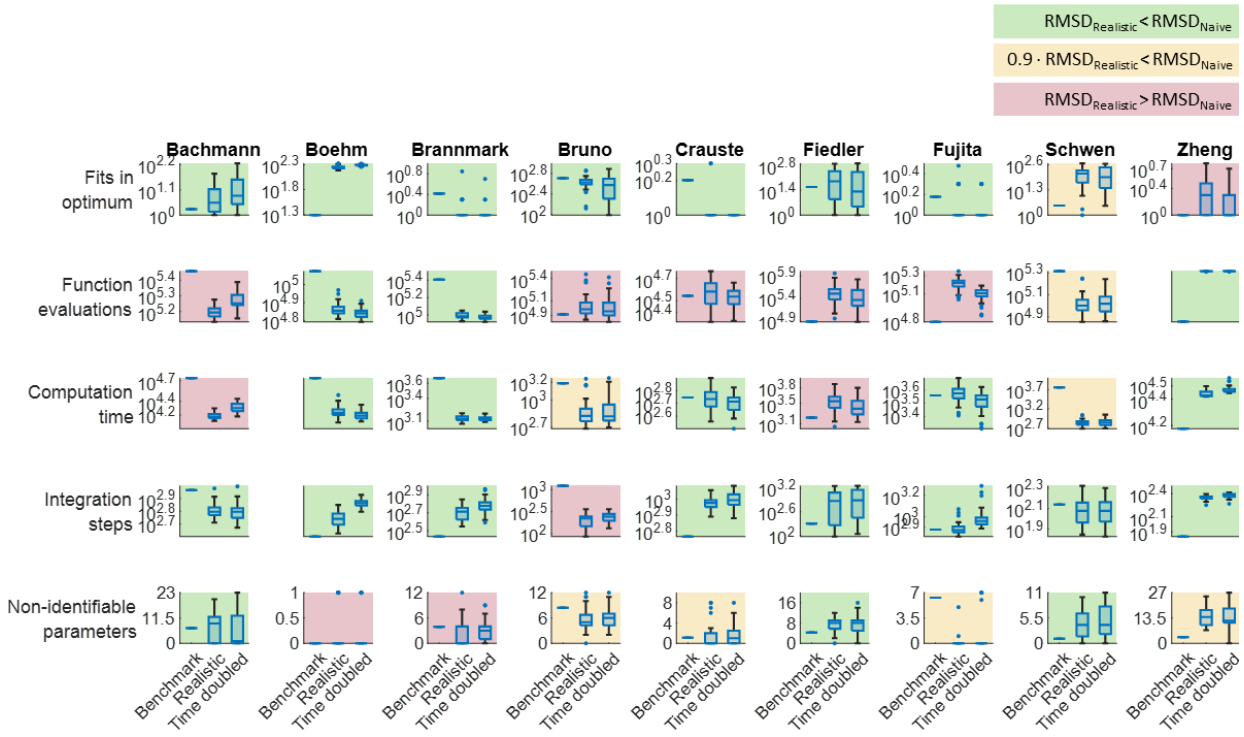
Supplementary Figure S3. Application example of a mitogen-activated protein-kinase (MAPK) cascade by Huang and Ferrell, 1996 [37] using the realistic simulation approach presented in the manuscript. A) The model scheme of the MAPK cascade [37]. B) Model dynamics with the equations and parameters from the BioModels database. C) and D) and E) Realistic simulations of measurements.

Validation

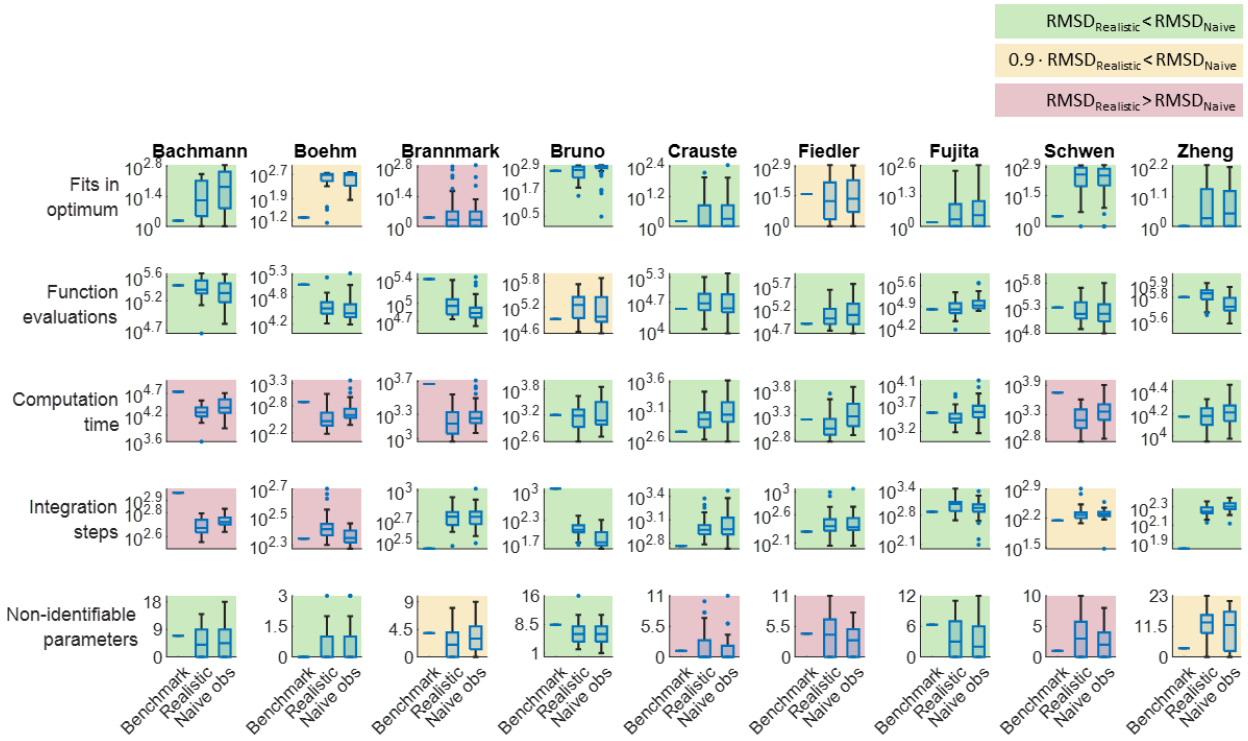
The impact of the presented simulation approach on different performance measures is further explored. To evaluate the impact of setting the measurement time points, the observables and observation functions stay the same as in the published model. Just the measurement time points are simulated as explained in the manuscript. The realistic simulation of time points is compared to a naive approach where each time point of the simulation is doubled. In general, the performance of the realistic sim-

ulation is closer to the published model than the naive approach (Figure S4). The number of function evaluations for optimization convergence is, in general, fewer for the naive approach. This is because, for the naive approach, the time points are not set according to the model dynamics, thus the model dynamics are not captured properly, and an easier fit is found.

To evaluate the impact of setting the observation type, the realistic simulation is compared to a naive approach with the same observables as in the realistic simulation, but only with direct measurements on an absolute scale. In explicit, for the naive approach, there are no compound measures or scaling factors. In general, the performance of the realistic simulation is closer to the published model than the naive approach (Figure S5).



Supplementary Figure S4. Performance measures for the presented approach of simulating realistic measurement time points compared to the published model and a naive approach. For the naive approach, the estimated time points are doubled. In general, the realistic approach of simulating measurement time points performs better (green). In 22% of the cases, the RMSD of the realistic simulation compared to the published model deviates more than 10% from the naive approach (red). In 18% of the cases, the RMSD of the realistic simulation deviates less than 10% from the naive approach (yellow).



Supplementary Figure S5. Performance measures for the presented approach of simulating the observation types compared to the published model and a naive approach. For the naive approach, the observables are measured directly without scaling or compound measurements. In general, the realistic approach of simulating the observation type performs better (green). In 22% of the cases, the RMSD of the realistic simulation compared to the published model deviates more than 10% from the naive approach (red). In 13% of the cases, the RMSD of the realistic simulation deviates less than 10% from the naive approach (yellow).