PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Validation-based model selection for ^{13}C metabolic flux analysis with uncertain measurement errors

Nicolas Sundqvist[1], **Nina Grankvist**[2,3,4], **Jeramie Watrous**[5], **Jain Mohit**[5], **Roland Nilsson**[2,3,4], **Gunnar Cedersund**[1,6]*

1 Linköping's University, Department of Biomedical engineering, Linköping, Sweden, **2** Cardiovascular Medicine Unit, Department of Medicine (Solna), Karolinska Institutet, Stockholm, Sweden, **3** Division of Cardiovascular Medicine, Karolinska University Hospital, Stockholm, Sweden, **4** Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden, **5** Department of Medicine & Pharmacology, University of California, San Diego, California, United States of America, **6** Center for Medical Image Science and Visualization, Linköping University, Linköping, Sweden

* gunnar.cedersund@liu.se

Abstract

Accurate measurements of metabolic fluxes in living cells are central to metabolism research and metabolic engineering. The gold standard method is model-based metabolic flux analysis (MFA), where fluxes are estimated indirectly from mass isotopomer data with the use of a mathematical model of the metabolic network. A critical step in MFA is model selection: choosing what compartments, metabolites, and reactions to include in the metabolic network model. Model selection is often done informally during the modelling process, based on the same data that is used for model fitting (estimation data). This can lead to either overly complex models (overfitting) or too simple ones (underfitting), in both cases resulting in poor flux estimates. Here, we propose a method for model selection based on independent validation data. We demonstrate in simulation studies that this method consistently chooses the correct model in a way that is independent on errors in measurement uncertainty. This independence is beneficial, since estimating the true magnitude of these errors can be difficult. In contrast, commonly used model selection methods based on the χ^2-test choose different model structures depending on the believed measurement uncertainty; this can lead to errors in flux estimates, especially when the magnitude of the error is substantially off. We present a new approach for quantification of prediction uncertainty of mass isotopomer distributions in other labelling experiments, to check for problems with too much or too little novelty in the validation data. Finally, in an isotope tracing study on human mammary epithelial cells, the validation-based model selection method identified pyruvate carboxylase as a key model component. Our results argue that validation-based model selection should be an integral part of MFA model development.

OPEN ACCESS

Citation: Sundqvist N, Grankvist N, Watrous J, Mohit J, Nilsson R, Cedersund G (2022) Validation-based model selection for ^{13}C metabolic flux analysis with uncertain measurement errors. PLoS Comput Biol 18(4): e1009999. https://doi.org/10.1371/journal.pcbi.1009999

Editor: Vassily Hatzimanikatis, Ecole Polytechnique Fédérale de Lausanne, SWITZERLAND

Received: June 24, 2021

Accepted: March 7, 2022

Published: April 11, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pcbi.1009999

Copyright: © 2022 Sundqvist et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data and supporting information are avilable at: https://gitlab.liu.se/nicsu70/val-based-model-selection_13cmfa.

Funding: This work was supported by the Swedish Foundation for Strategic Research, grant no. FFL12-0220 (RN, NG) and IMT17-0245 (RN, NG, NS, GC), the Swedish Research Council grant no. 2018-05418 (GC) and 2018-03319 (GC, NS), Karolinska Institutet (RN), CENIIT grant no. 15.09 (GC), SciLifeLab National COVID-19 Research Program, financed by the Knut and Alice Wallenberg Foundation grant no. 2020.0182 (GC), the H2020 project PRECISE4Q grant no. 777107 (GC), and from VINNOVA grants VisualSweden and 2020-04711 (GC). Additional support for GC came from the Swedish Fund for Research without Animal Experiments (F2019-0010), ELLIIT (2020-A12). Initials indicate that part of the author's salary has been paid for by the preceding funds. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Measuring metabolic reaction fluxes in living cells is difficult, yet important. The gold standard is to label extracellular metabolites with ^{13}C, to use mass spectrometry to find out where the ^{13}C-atoms ends up, and finally use mathematical modelling to calculate how quickly each reaction must have flowed, for the ^{13}C-atoms to end up like that. This measurement thus relies on usage of the right mathematical model, which must be *selected* among various candidate models. In this manuscript, we present a new way to do this model selection step, utilizing validation data. Using an adopted approach to calculate the uncertainty of model predictions, we identify new validation experiments, which are neither too similar, nor too dissimilar, compared to the previous training data. The model candidate that is best at predicting this new validation data is the one chosen. Tests on simulated data where the true model is known, shows that the validation-based method is robust when the magnitude of the error in the measurement uncertainty is unknown, something that conventional methods are not. This improvement is important since true uncertainties can be difficult to estimate for these data. Finally, we demonstrate how the new method can be used on real data, to identify fluxes and important reactions.

1. Introduction

Cellular metabolism is fundamental for all living organisms, involving thousands of metabolites and metabolic reactions that together form large interconnected metabolic networks [1,2]. While a substantial part of the human metabolic network has been reconstructed [2], measuring fluxes through individual reactions and metabolic pathways in living cells and tissues remains a challenge. This problem is central to a variety of medically relevant processes, including T-cell differentiation [3], caloric restriction and aging [4], cancer [5,6], the metabolic syndrome [7], and neurodegenerative diseases such as Parkinson's disease [8].

The gold standard method for measuring metabolic fluxes in a given system is model-based metabolic flux analysis (MFA) [9]. In this technique, cells or tissues are fed "labelled" substrates containing stable isotopes such as ^{13}C (Fig 1A). These substrates are metabolized to products containing various isotopic isomers (isotopomers) (Fig 1B). By measuring the abundance of these isotopomers, mass isotopomer distributions (MIDs, Fig 1C) are obtained for each metabolite [10]. Fluxes are then inferred by fitting a mathematical model $\mathcal{M}$ to the observed MID data D (Fig 1D).

While the above methodology is well established for assessing the fit of a given MFA model, several problems arise when it is used for model selection. In practise, MFA models are usually developed iteratively (Fig 1D), by repeatedly attempting to fit the same data to a sequence of models $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_k$ with successive modifications (adding or removing reactions, metabolites, and so on), until a model $\mathcal{M}_k$ is found acceptable, i.e. not statistically rejected. In practice, this means that the model $\mathcal{M}_k$ passes the χ^2-test for goodness-of-fit [11]. Given the iterative nature of modifying the model structures, model development thus turns into a model selection problem. Depending on the approach used to solve this model selection problem different model structures might be selected, given the same data set (Fig 1E). For instance, if the traditionally iterative modelling cycle is used, the first model that passes the χ^2-test might be selected and used for flux estimation. On the other hand, there might be multiple model structures that passes the χ^2-test. In this case, the model structure that passes the χ^2-test with the biggest margin may be a better option.

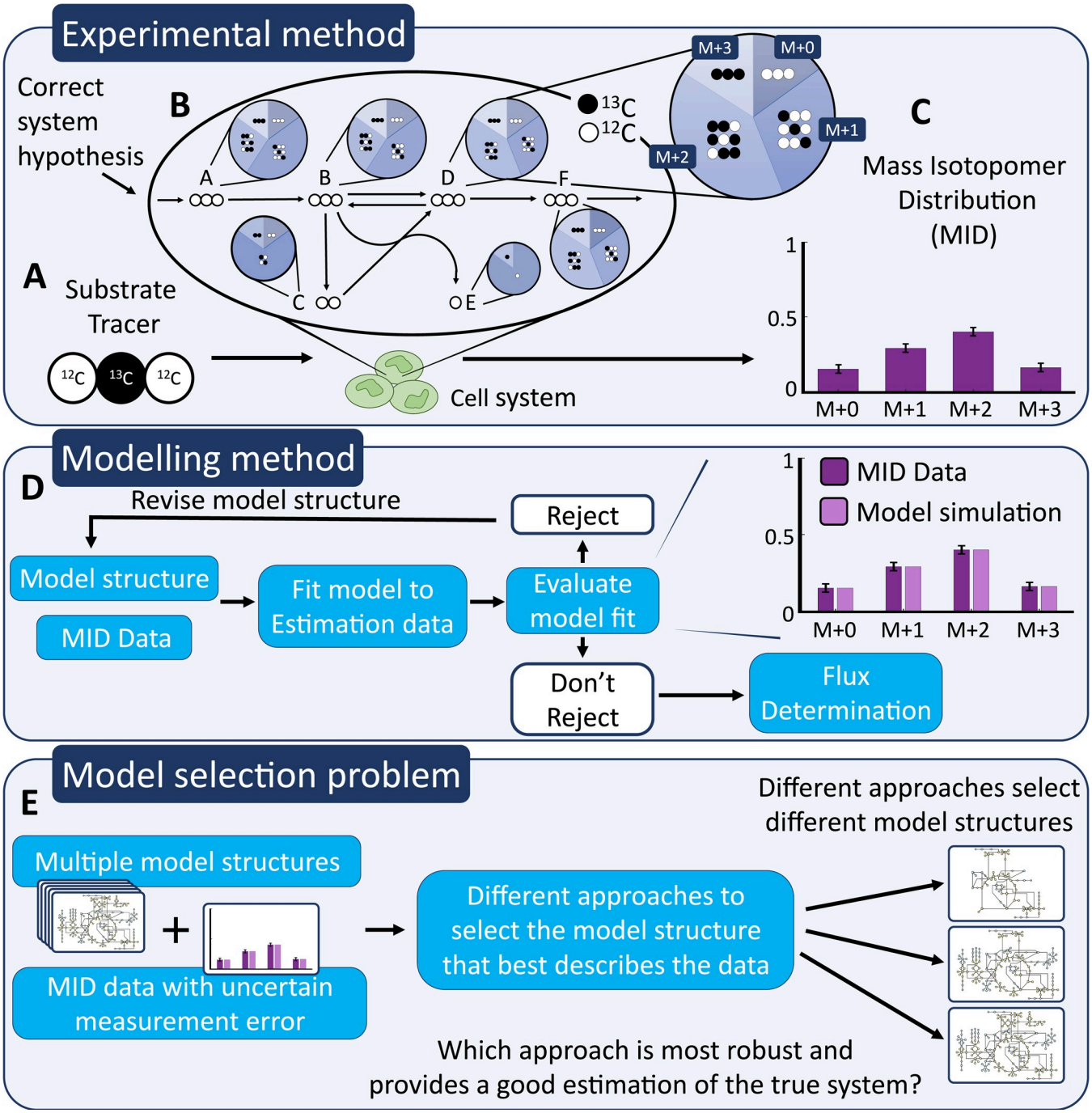

Fig 1. The basic steps in ^{13}C MFA and the model selection problem. (A) New substrates, containing ^{13}C (dark circles) are fed to the cells. (B) These substrates are consumed and converted to end products in the cells, according to its biochemical reactions. (C) The labelled ^{13}C molecules appear to various proportions in each of the mass isotopomers, and these proportions are summed up in these distribution bar charts for each detected metabolite. (D) The iterative modelling cycle in which a hypothesized model structure is fitted to MID data. The model fit is evaluated, usually with a χ^2-test, and either rejected or not. If the model structure is rejected it is revised and evaluated again. If the model structure is not rejected it is used for flux determination. (E) The iterative model development in (D) results in a model selection problem. Different approaches for solving this model selection problem might result in different model structures being selected. This paper evaluates how the uncertainty in measurement data affects uncertainty in model selection.

https://doi.org/10.1371/journal.pcbi.1009999.g001

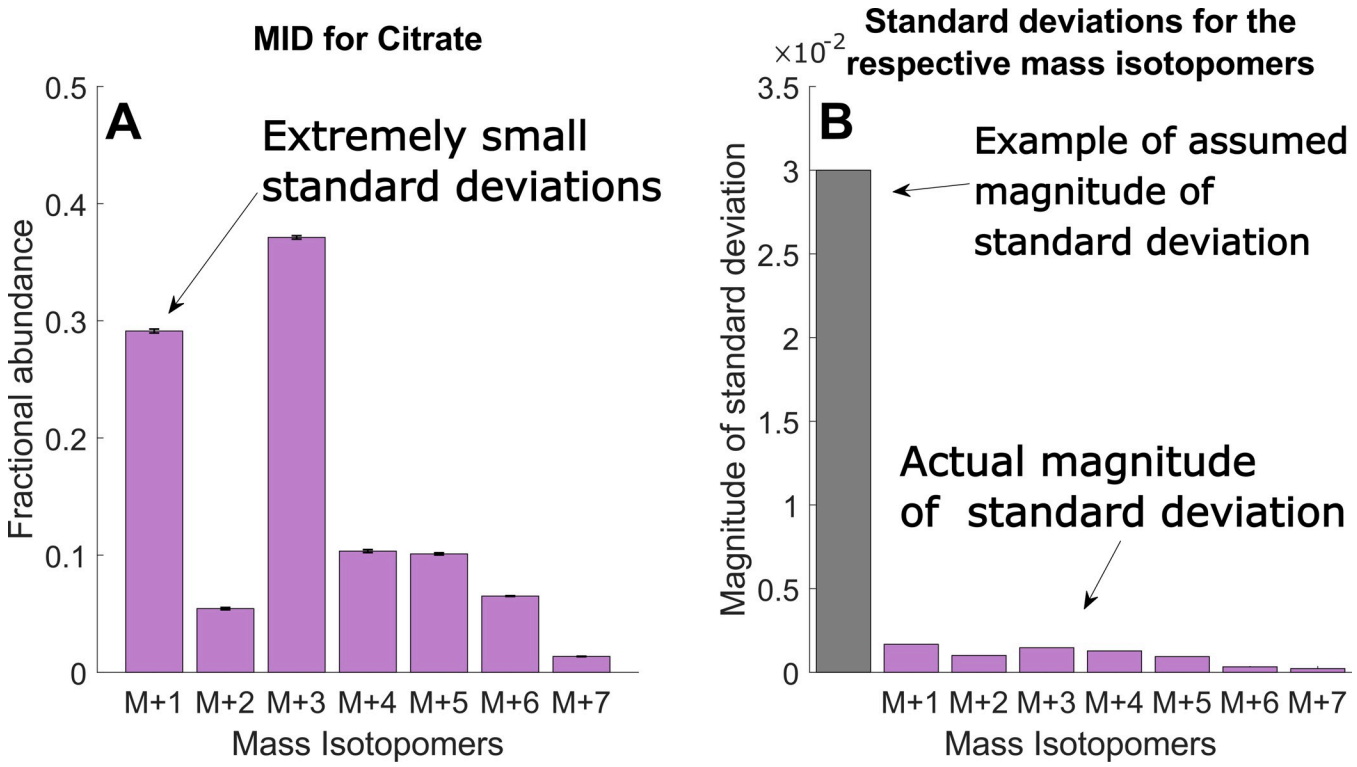

Fig 2. Example of MID sample standard deviation (A) Example of estimated mass isotopomer distribution (MID) of citrate from epithelial cells, as described in section 2.5. M+i indicate the fractional abundance of the i:th mass isotopomer. (B) Difference between the assumed magnitude of the standard deviations and the measured magnitudes.

https://doi.org/10.1371/journal.pcbi.1009999.g002

Generally, model selection approaches that rely solely on the χ^2-test to select a model can be problematic. First, correctness of the χ^2-test depends on knowing the number of identifiable parameters, which is needed to properly account for overfitting by adjusting the degrees of freedom of the χ^2 distribution [12], but can be difficult to determine for nonlinear models [13]. Second, the χ^2-test can be unreliable in practise since the underlying error model is often not accurate. Typically, the MID errors σ are estimated by sample standard deviations s from biological replicates, which for mass spectrometry data often is below 0.01, and even can be as low as 0.001 (Fig 2A). However, such low estimates may not reflect all error sources. For example, MI fractions obtained from orbitrap instruments can be biased so that minor isotopomers are underestimated [14,15]. Also, s does not account for experimental bias, such as deviations from metabolic steady-state that always occur in batch cultures. Some such problems can be detected by repeating experiments, but some others cannot. The normal distribution assumption itself is also questionable for MIDs, which are constrained to the n-simplex [16]. For these reasons, s can severely underestimate the actual errors, making it exceedingly difficult to find a model that passes a χ^2-test. In this situation, one is left with two bad choices: either arbitrarily increase s to some "reasonable" value to pass the χ^2-test (Fig 2B), or introduce more or less well-motivated extra fluxes into the model. The former alternative, increasing s, may lead to high uncertainty in the estimated fluxes and does not necessarily reflect the experimental bias one tries to account for. The latter approach, introducing additional fluxes, increases model complexity and can lead to overfitting.

While these issues with model selection are well known, they have to our knowledge not been treated systematically in the ^{13}C MFA field. Indeed, MFA model selection is typically

done in an informal fashion by trial-and-error, and the underlying procedure is rarely reported [17]. However, in other contexts where model fitting is central, such as systems biology, the problem of model selection has been treated extensively [13,18–27]. In these areas, a widely accepted solution is to perform model selection on a separate "validation" data set, which is not used for model fitting. Intuitively, this protects against overfitting by choosing the model that can best predict new, independent data. In this paper, we propose a formalized version of such a validation-based model selection approach for MFA. In a series of simulated examples, we demonstrate that this method consistently selects the correct metabolic network model, despite uncertainty in measurement errors, whereas "traditional" χ^2-testing on the estimation data does not. By quantifying prediction uncertainty using prediction profile likelihood, we can avoid cases where the validation data is too similar, or too dissimilar, to the estimation data. Finally, in an application to flux analysis on our own new data in human epithelial cells, we find that the same robustness to measurement uncertainty variations still holds, and that the validation-based model selection method can identify reactions that are known to be active in this cell type.

2. Results

To systematically examine the effects of the model selection procedure on MFA, we adopted a scheme where a sequence of models $\mathcal{M}_1, \mathcal{M}_2, \ldots$ with increasing complexity (increasing number of parameters) is tested by each model selection method, simulating typical iterative model development. We considered five possible model selection methods that use all available data for both parameter estimation and model evaluation (Table 1). Method "SSR" selects the model with the smallest weighted summed squared residuals (*SSR*) based on the data, included as a baseline. Method "First χ^2" selects the model with fewest parameters (the "simplest" model) that passes a χ^2-test, while accounting for overfitting by subtracting the number of free parameters p from the degrees of freedom in the χ^2-distribution (see Section 4.3). Method "Best χ^2" selects the model that passes the χ^2-threshold with the greatest margin. Methods "AIC" and "BIC" select the model that minimizes the Akaike Information Criterion or the Bayesian Information Criterion, respectively [28,29]. The five methods mentioned above all depend on the noise model Eq (5), and all except "SSR" also requires knowing the number of free parameters p. Considering common practices in the field, it is probable that some combination of the "First χ^2" and "Best χ^2" methods is the prevailing approach in MFA modelling [17,30], although this is not entirely clear since the model selection process is often not described.

In addition to these methods, we propose a validation-based model selection method ("Validation") that divides the data D into estimation data D^{est} and validation data D^{val}. For each model, parameter estimation (model fitting) is then done using D^{est}, and the model achieving the smallest SSR with respect to D^{val} is selected. The division into estimation and validation

Table 1. A summary of the different model selection approaches considered in this paper.

Method of model selection	Model selection criteria
Estimation SSR	Selects the model with the lowest SSR given D^{est}
First χ^2	Selects the first $\mathcal{M}_k$ that passes the χ^2 -test
Best χ^2	Selects the $\mathcal{M}_k$ that passes the χ^2-test with the greatest margin
AIC	Selects the $\mathcal{M}_k$ that minimizes the Akaike Information Criterion
BIC	Selects the $\mathcal{M}_k$ that minimizes the Bayesian Information Criterion
Validation	Selects the $\mathcal{M}_k$ with the smallest SSR with respect to D^{val}

https://doi.org/10.1371/journal.pcbi.1009999.t001

data must be done so that qualitatively new information is present in the validation data. This can be done by reserving data from distinct model inputs or new model outputs for validation. For all examples herein, data from distinct model inputs is used for validation. For the ^{13}C MFA examples, this means that data used for validation comes from a different tracer. Note that this proposed method allows for the selection of the most suitable model from a given set, but that it does not guarantee that the selected model is acceptable according to e.g. a χ^2-test. In other words, the model selected with our new Validation method would still need be subjected to some form of final model testing. A detailed description of the "SSR", "First χ^2", "Best χ^2", and "Validation" methods can be found in S1 Algorithm: A-D respectively (S1 Algorithm).

2.1 A motivating example

Before examining the behavior of the different model selection methods on metabolic network models, it may be helpful to illustrate their properties on a simple univariate example. For this purpose, we considered a model with a single input x and a single output $\hat{y}$, where model $\mathcal{M}_n$ is the n-th order polynomial

$$\hat{y} = h_n(x, u) = \sum_{k=0}^{n} u_k x^k \tag{1}$$

with parameter vector u. We assume that $\mathcal{M}_7$ is the correct model, with true parameters u_0, and sampled 20 measurements $y = h_7(x, u_0)+\epsilon$ for different values of x, where ϵ was drawn from $N(0, \sigma_r)$ with standard deviation $\sigma_r = 0.2$. To simulate uncertainty about the error model, we considered σ to be unknown, and let the various model selection methods choose among $\mathcal{M}_1, \ldots, \mathcal{M}_{14}$ with a "believed" standard deviation, denoted σ_b, in the range $[0.1\ \sigma_r, 10\ \sigma_r]$. For the "Validation" method, we reserved 4 of the 20 measurements for D^{val} (Fig 3, red error bars).

An illustration of the dependency on σ for a model selection method that does not use validation is shown in Fig 3. When D^{val} is not considered, we would expect larger values of σ_b to result in a simpler model, since almost all of the variation in the data is interpreted as noise (Fig 3A). Further, at very small values of σ_b an overly complex model will be required to obtain an acceptable fit to D^{est} (Fig 3C).

Applying the five model selection methods to data from this polynomial model gave different results (Figs 4 and S1). Since the model selection process is somewhat stochastic, we resampled the data 10,000 times, each time with a new error ϵ drawn from $N(0, \sigma_r)$, and report results as the fraction of times a particular model was chosen. As expected, "SSR" mostly selected the most complex polynomial $\mathcal{M}_{14}$ regardless of σ_b, as the most complex model always gives the lowest SSR (Fig 4A). In contrast, "First χ^2" or "Best χ^2" gave different results depending on σ_b. "First χ^2" selected the correct model $\mathcal{M}_7$ only when $\sigma_b \approx \sigma_r$. At $\sigma_b \approx 10\sigma_r$, only the low-degree polynomials ($\mathcal{M}_1, \mathcal{M}_2$ *and* $\mathcal{M}_3$) was chosen by the "First χ^2" method, while at $\sigma_b \approx 0.1\sigma$, an overly complex polynomial was chosen (Fig 4B). The "Best χ^2" method selected the correct model $\mathcal{M}_7$ for $\sigma_b \geq \sigma_r$, but selected overly complex models for smaller σ_b (Fig 4C). If σ_b were to increase further, "Best χ^2" would choose a lower degree polynomial. This is because, for these χ^2-based methods, the tradeoff between model complexity and goodness-of-fit is based on σ_b, and such a tradeoff is thus correct only if we happen to have $\sigma_b \approx \sigma_r$. Similar results are seen with the "AIC" and "BIC" methods, which also depend on σ_b (S1 Fig).

In contrast, the "Validation" method predominantly selected the correct model, $\mathcal{M}_7$, regardless of σ_b (Fig 4D). This happens because, even though a polynomial of the wrong degree may fit well on D^{est}, it fails to predict independent validation data, resulting in large SSR on D^{val}. Since the correct model structure $\mathcal{M}_7$ will best predict new data, agreement with validation data helps identify the right model, also in cases where the error model is inaccurate.

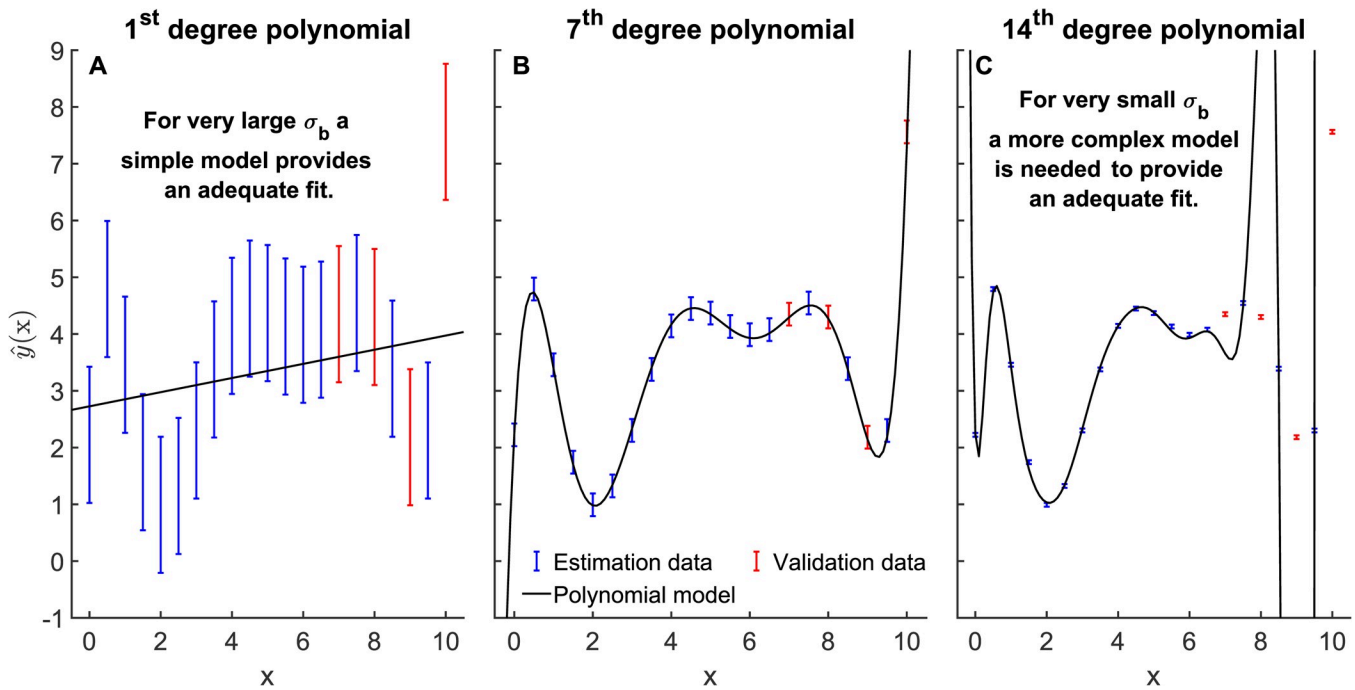

Fig 3. Example of how model selection is affected by σ_b, for the polynomial model. Error bars indicate data sampled from a 7th order polynomial $y = h_7(x, u_0) + \epsilon$ where ϵ is $N(0, \sigma_r)$, $\sigma_r = 0.2$. Colours indicate estimation data D^{est} (blue) and validation data D^{val} (red) used by the "Validation" method. Solid curves in (A–B) indicate polynomials chosen by an estimation-based method with different "believed" standard deviation σ_b. (A) $\sigma_b = 2$, chosen model h_1. (B) $\sigma_b = 0.2$ (the true value), chosen model h_7 (the correct model). (C) $\sigma_b = 0.02$, chosen model h_{14}.

https://doi.org/10.1371/journal.pcbi.1009999.g003

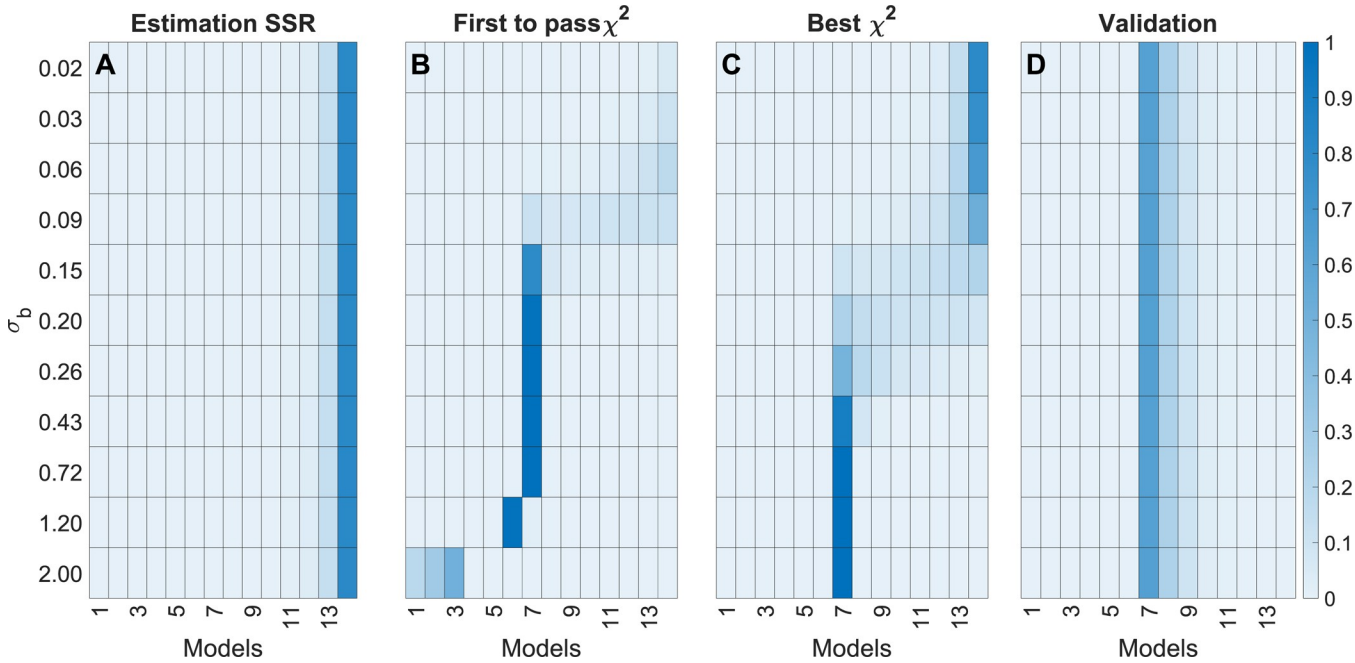

Fig 4. Model selection results for the polynomial model example. (A–D) Heatmaps represent results from the indicated selection methods, where rows represent different values of σ_b and columns represent the polynomial models $h_1, \ldots, h_{14}$. For each row, color indicates the fraction of times a model is selected for the given σ_b, out of 10,000 samples, as indicated by the color scale (right).

https://doi.org/10.1371/journal.pcbi.1009999.g004

Again, it should be recalled that the "Validation" method only is applicable to the task of selecting the best model, and that this selection approach should be followed by a final step that tests the quality of the model, e.g. using a χ^2-test.

2.2 Model selection on multivariate linear models

To investigate model selection in a setting more relevant to metabolic networks, we next considered a multivariate linear model, where the output vector $\hat{y}$ is a linear combination of the inputs x weighted by the model parameters. In this case, each model structure $\mathcal{M}_k$ is fully specified by a matrix A_k such that

$$\hat{y} = h_k(x, u) = A_k x \tag{2}$$

and where the free parameters are elements of A_k (Fig 5). This type of model is roughly analogous to a simple metabolic network, where x corresponds to labelled substrates and y corresponds to metabolic products. We constructed six such models (A_1–A_6) of increasing complexity, nested so that the parameter space of each A_k contains the parameter space of all models A_l for $l<k$. Model A_3 was used to simulate data from 6 distinct input vectors x, again with normal noise $N(0, \sigma_r)$ where $\sigma_r = 5$, and with a "believed" σ_b in the range [0.1 σ_r, 10 σ_r]. The believed σ_b is not scaled homogeneously across all data points rather the scaling is approximately 0.1σ_r and 10 σ_r, to reflect the more realistic scenario of σ_b being wrong to different degrees for different data points (Materials and methods). Again, to account for variance in the model selection process, the results are based on 1,000 different resamplings.

We then tested each of the six model selection methods on the generated data. As before, the "SSR" method chose the most complex models (A_5 and A_6, Fig 6A). For the other four methods that only use estimation data, the selected model again depended on σ_b. Method "First χ^2" selected the one of the simpler models A_2 at $\sigma_b \approx 10\ \sigma_r$, the correct model A_3 only when $\sigma_b \approx \sigma$, while at $\sigma_b \approx 0.1\ \sigma_r$, no model passed the χ^2-test (Fig 6B). The "Best χ^2" method again selected the correct model, A_3, for $\sigma_b \approx 10\ \sigma_r$ and $\sigma_b \approx \sigma_r$, and model A_6 for $\sigma_b \approx 0.1\ \sigma_r$

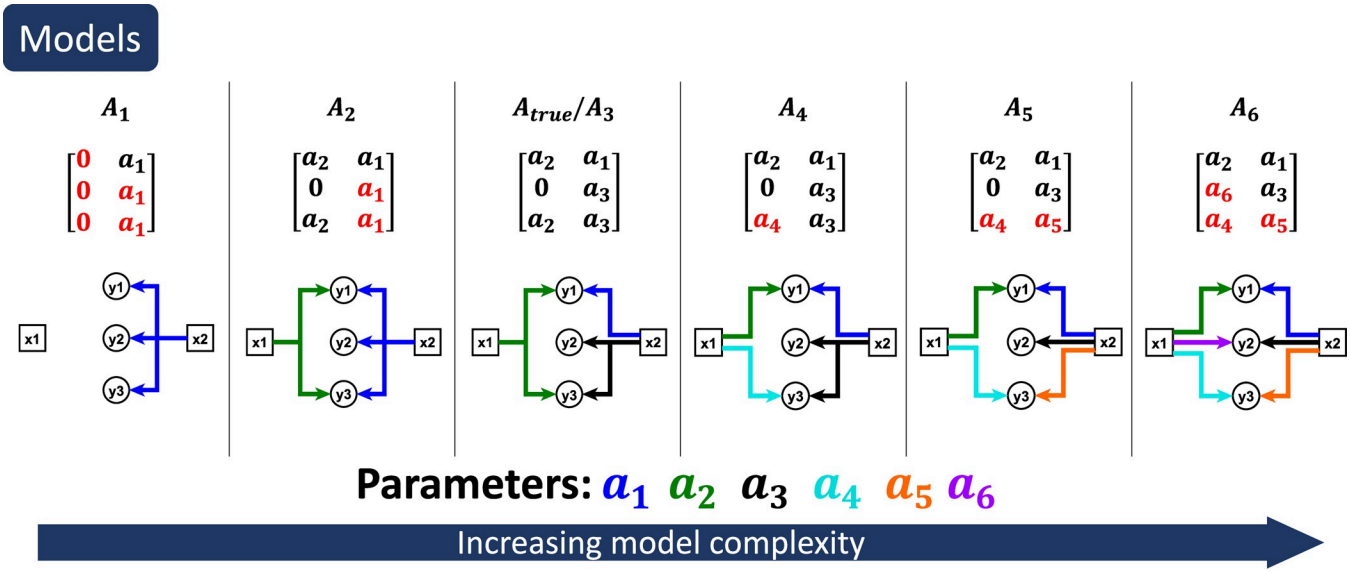

Fig 5. Six different model structures for the linear model. This example is chosen as a simple representation of a mass flow model. The top row shows the model names $A_1, \ldots, A_6$. The second row shows the matrices that constitute the model structures. The third row constitute visual illustrations of how the corresponding matrices connect the inputs x_i and the outputs y_i via the parameters $a_1, \ldots, a_6$.

https://doi.org/10.1371/journal.pcbi.1009999.g005

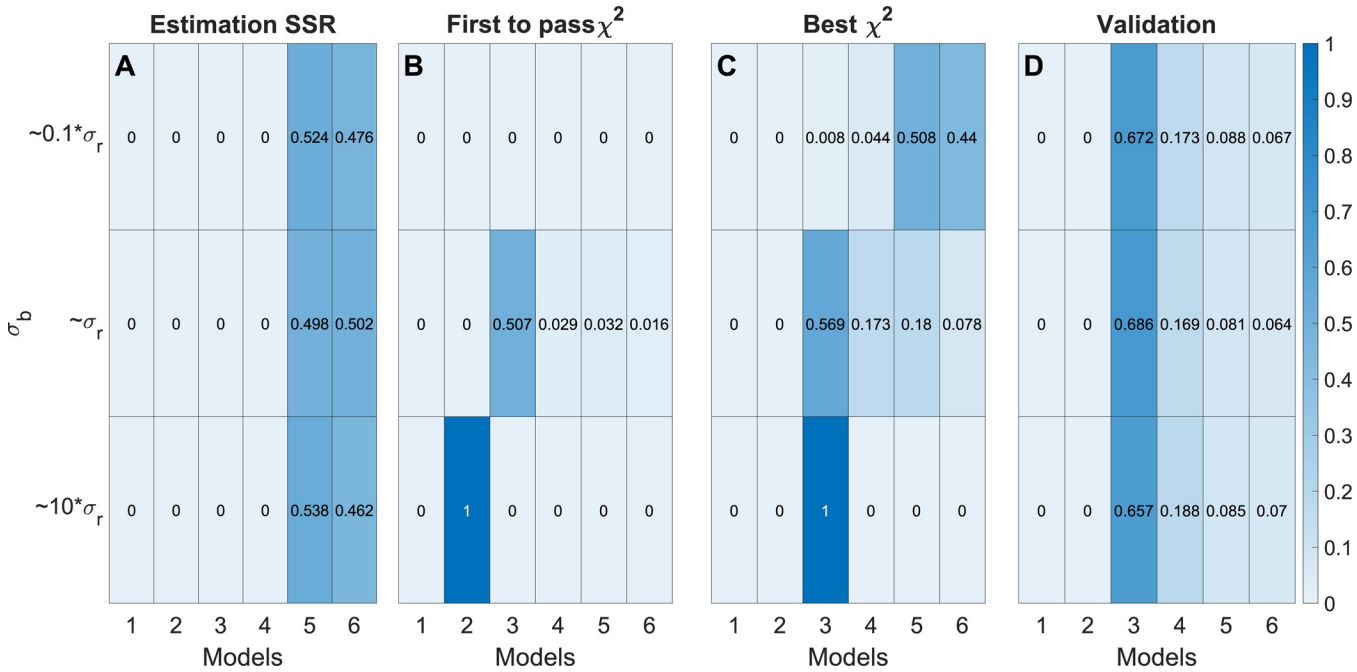

Fig 6. Model selection results for the linear model example. (A–D) Heatmaps represent results from the indicated selection methods, where rows represent different values of σ_b and columns represent the linear models $A_1,\ldots,A_6$. For each row, color indicates the fraction of times a model is selected for the given σ_b, out of 1000 samples, as indicated by the color scale (right).

https://doi.org/10.1371/journal.pcbi.1009999.g006

(Fig 6C). Methods "AIC" and "BIC" behaved similarly but chose somewhat different models (S2 Fig). Thus, which model is considered "best" depends on assumptions about the measurement noise, and established model selection methods give different results depending on what assumption is made.

For the "Validation" method, simulated data from 2 of the 6 distinct inputs x were reserved for validation data D^{val} (S3 Fig). Again, this method predominantly selected the true model structure A_3, regardless of σ_b (Fig 6D). Moreover, the model selection results for "Validation" method are consistent across all σ_b.

2.3 Model selection for simulated ^{13}C MFA models

Let us now turn to model selection for the multivariate, nonlinear MFA models. To simulate the process of MFA model development, we designed seven stoichiometric models $\mathcal{M}_1,\ldots,\mathcal{M}_7$ of the tricarboxylic acid (TCA) cycle and related reactions, with increasing model complexity (S1 Table and Fig 7). The full, atom-level models were generated using the EMU decomposition method (see methods). For all seven models, 51 MI fractions across nine metabolites (present in all models) were considered as measurement data. The data was simulated (Section 4.5) using model $\mathcal{M}_4$, with four different tracers separately used as inputs x in order to generate four separate sets of MID data. For this example, we resampled the data 100 times. Note that, unlike the previous examples, this model is nonlinear in the parameters.

As in the previous examples, the six methods of model selection were evaluated. As before, the "SSR" method always selected the most complex model $\mathcal{M}_7$ (Fig 8A). Method "First χ^2" selected different models depending on the value of σ_b (Fig 8B): it selects $\mathcal{M}_1$ for $\sigma_b\approx 10\sigma_r$, $\mathcal{M}_2$ for $\sigma_b\approx 3\sigma_r$,$\mathcal{M}_4$ (the true model) and $\mathcal{M}_2$ about 50% of the time respectively for $\sigma_b\approx\sigma_r$, and $\mathcal{M}_4$ for $\sigma_b\approx 0.1\sigma_r$ and $0.3\sigma_r$. In this example, "Best χ^2" method selected the correct model $\mathcal{M}_4$

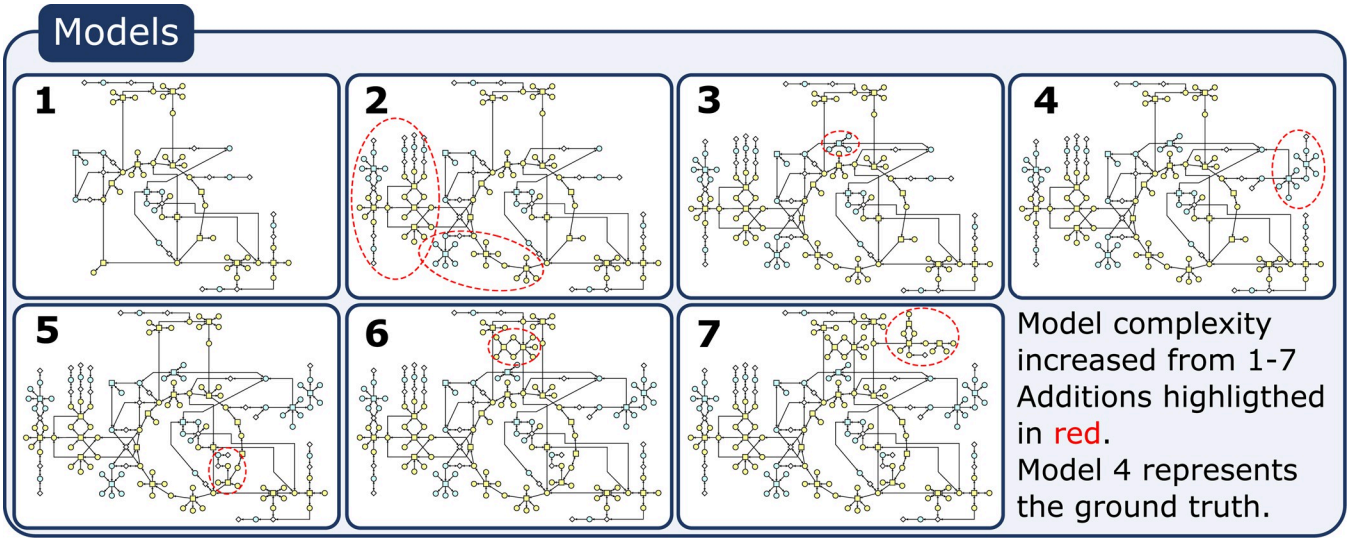

Fig 7. Seven different model structures included in the simulated EMU ^{13}C MFA example with simulated data. The added component to each model structure, compared to the previous model, with slightly smaller complexity, is found inside the red circle. The true model used to simulate the data is model nr 4. Detailed descriptions for each model can be found in the supplementary material (S1 Table).

https://doi.org/10.1371/journal.pcbi.1009999.g007

for $\sigma_b \approx 0.3\sigma_r$ and σ_r (Fig 8C). For $\sigma_b \approx \sigma_r$ "Best χ^2" show a fraction of the samples selecting $\mathcal{M}_2$ rather than $\mathcal{M}_4$ (Fig 8C) and for $\sigma_b \approx 10\sigma_r$, "Best χ^2" shift towards selecting the simpler model structures $\mathcal{M}_2$. Compared to previous examples, the AIC and BIC methods of model selection appear to be a bit more robust towards an unknown σ_b, selecting $\mathcal{M}_4$ for $\sigma_b \approx 0.1\sigma_r$, $0.3\sigma_r$, σ_r, and $3\sigma_r$. Nevertheless, for $\sigma_b \approx 0.1\sigma_r$, both the AIC and BIC show tendencies to prefer more

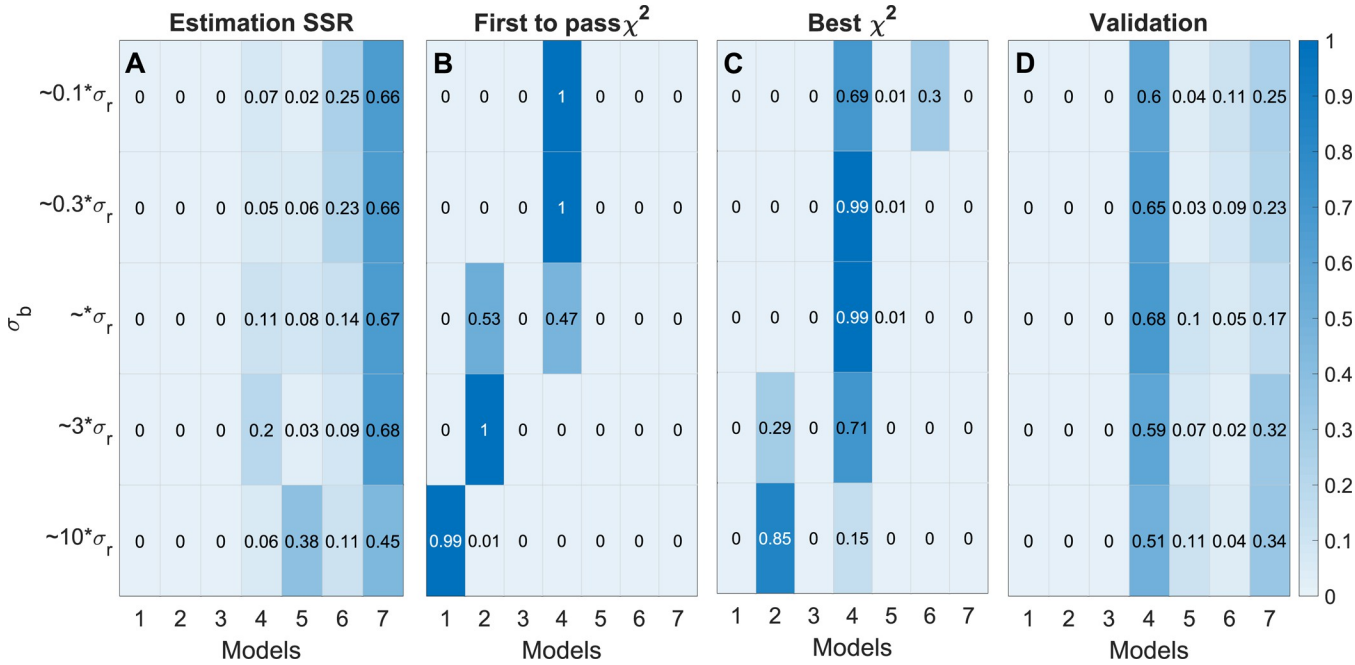

Fig 8. Model selection results for the simulated ^{13}C MFA model example. (A–D) Heatmaps represent results from the indicated selection methods, where rows represent different values of σ_b and columns represent the MFA models $\mathcal{M}_1, \ldots, \mathcal{M}_7$. For each row, color indicates the fraction of times a model is selected for the given σ_b, out of 100 samples, as indicated by the color scale (right).

https://doi.org/10.1371/journal.pcbi.1009999.g008

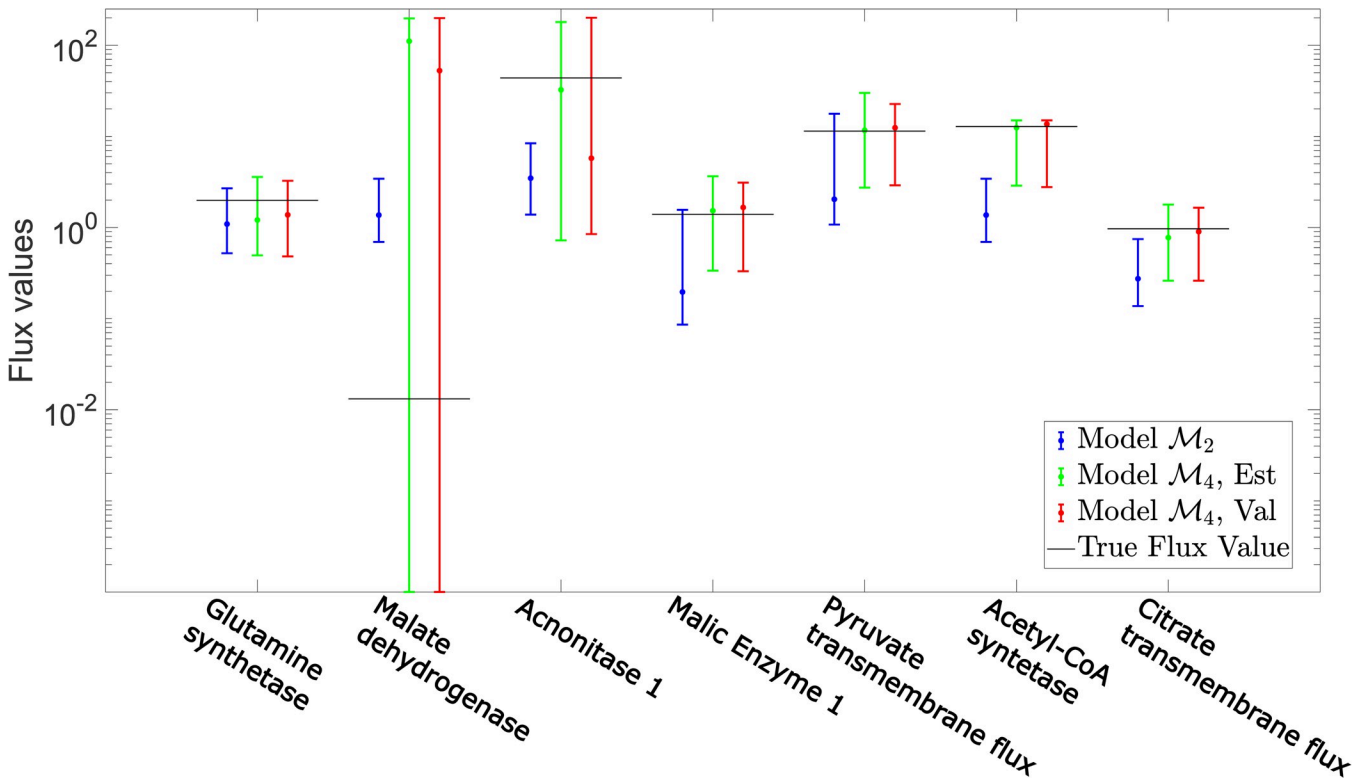

Fig 9. Comparison of estimated flux solutions for the simulated ^{13}C MFA example. The resulting flux values with 95% confidence intervals for seven of the fluxes that are overlapping between all model structures in the simulated ^{13}C MFA example. The confidence intervals correspond to the estimated fluxes for model $\mathcal{M}_2$ (Blue), model $\mathcal{M}_4$ with all data available (Green) and model $\mathcal{M}_4$ with the data split into D^{est} and D^{val} (Red). The figure illustrates the selecting the wrong model structure may result in incorrect flux estimations.

https://doi.org/10.1371/journal.pcbi.1009999.g009

complex models and for $\sigma_b \approx 10\sigma_r$ both the AIC and BIC selects too simple model structures, namely $\mathcal{M}_2$ and $\mathcal{M}_1$ respectively (S4 Fig).

For the "Validation" method, parameters were estimated using MID data from 3 out of 4 tracers, while the fourth set of MID data was used as validation data D^{val}; the exact division is described in Section 4.5. The "Validation" method selected the correct model $\mathcal{M}_4$ in 60–70% of cases, for all tested values of σ_b (Fig 8D). The key observation here is that the validation-based method obtains the same results independently of σ_b. However, it should be noted that the "Best χ^2" method does appear more robust in identifying the correct model when σ is correct.

By selecting the wrong model structure, methods that depend on σ_b can lead to poor estimates of metabolic fluxes. For instance, when investigating the estimated flux values for model $\mathcal{M}_2$ (model selected by "*First* χ^2" at $\sigma_b \approx 3\sigma_r$) it becomes clear that the "*First* χ^2" approach does not always capture the correct flux value with a 95% confidence interval (Fig 9). For instance, for the fluxes for mitochondrial Aconitase1 (ACONT1m) and Acetyl-CoA Synthetase (ASCm), the 95% confidence interval lies several standard deviations away from the estimated value, indicating that the confidence intervals are not reliable. In contrast, the flux solution, with a 95% confidence interval, for model $\mathcal{M}_4$ (selected by the "Validation" method) does contain the true flux value (Fig 9). These results show that selecting the wrong model structure leads to errors in flux estimation, and that the "Validation" method therefore is more advantageous for both of these tasks.

2.4 Assessing the novelty of a validation experiment using prediction uncertainty

As showed in the previous examples, validation data can be used for the purpose of model selection. However, one important aspect to consider for this new method in ^{13}C MFA, is the degree of novelty of the validation data used. There are essentially two pitfalls that one wants to avoid: 1) the validation data may be too similar to the estimation data (i.e. then the new data does not provide any new information), 2) the validation data may be too dissimilar from the estimation data (i.e. then no model is able to predict the validation data). Both of these pitfalls can be avoided by looking at the uncertainty of the model predictions for the chosen validation experiment. The model's prediction uncertainty is essentially a confidence interval for the predictions. In our case of ^{13}C MFA models, a confidence interval is an interval for the predicted MIDs. This prediction uncertainty will depend on the uncertainty of the estimated fluxes, the model structure, and on the connection between the validation data and the estimation data. If the validation data is not novel enough (pitfall 1 above), models will produce identical predictions which do not differ from the estimation data (Fig 10A). On the other hand, if the validation data is too novel (pitfall 2), the estimation data does not contain any new information regarding the predicted MIDs, and the uncertainty will be very large (Fig 10B). Together, this means that the degree of uncertainty in the model predictions, compared to the difference in predictions between estimation and validation data, can be used to assess the novelty of the validation data. The desired scenario would be to have validation data such that the predictions are well-determined and are different between estimation and validation data (Fig 10C). A general approach for determining prediction uncertainty has been outlined in previous work [31] and has been implemented here for ^{13}C MFA models in the EMU framework. A detailed description of this implementation is provided in Materials and methods, Section 4.4.

Another aspect that is important to consider for the case of ^{13}C MFA, if the validation data consists of MIDs from a new tracer experiment, is that the new tracer is suitable to be used for validation. One approach to ensure that the new tracer generates data that is truly independent of the estimation data is to perform an EMU basis vector analysis [32]. This approach ensures that the tracers for the estimation and validation data produce linearly independent EMU basis vectors, which guarantees that the experiments give complementary information. This also ensures that one avoids the pitfall of having the validation data containing the same information as the estimation data, i.e. that the validation data is too similar to the estimation data.

To demonstrate that the validation data used in the simulated ^{13}C MFA example above does not fall into these pitfalls, the prediction uncertainties for the chosen model structure $\mathcal{M}_4$ has been determined (Fig 11). As can be seen, the prediction uncertainties (light blue bars' error bars) are well determined for all MIDs. We have thus avoided pitfall 2 above, i.e. the validation data is not too novel. We have also avoided pitfall 1, since i) the predicted MIDs (light blue) are non-overlapping with the estimation data MIDs (red bars), for many of the MIDs, ii) the EMU basis vectors are linearly independent.

2.5 Model selection on cultured epithelial cells

Finally, we applied validation-based model selection on data from batch cultures of human cells. We performed two isotope labelling experiments with immortalized human mammary epithelium cells (HMECs), cultured with either U-^{13}C-glucose or U-^{13}C-glutamine for 6 cell doublings to achieve isotopic and metabolic steady-state (Materials and Methods, Section 4.6). MID data for nine metabolites were used as measurement data for this example, and the model structures used were the same as those presented previously in Section 2.3. The sample standard deviations from biological replicates were very small, around $s = 0.005$. In contrast to

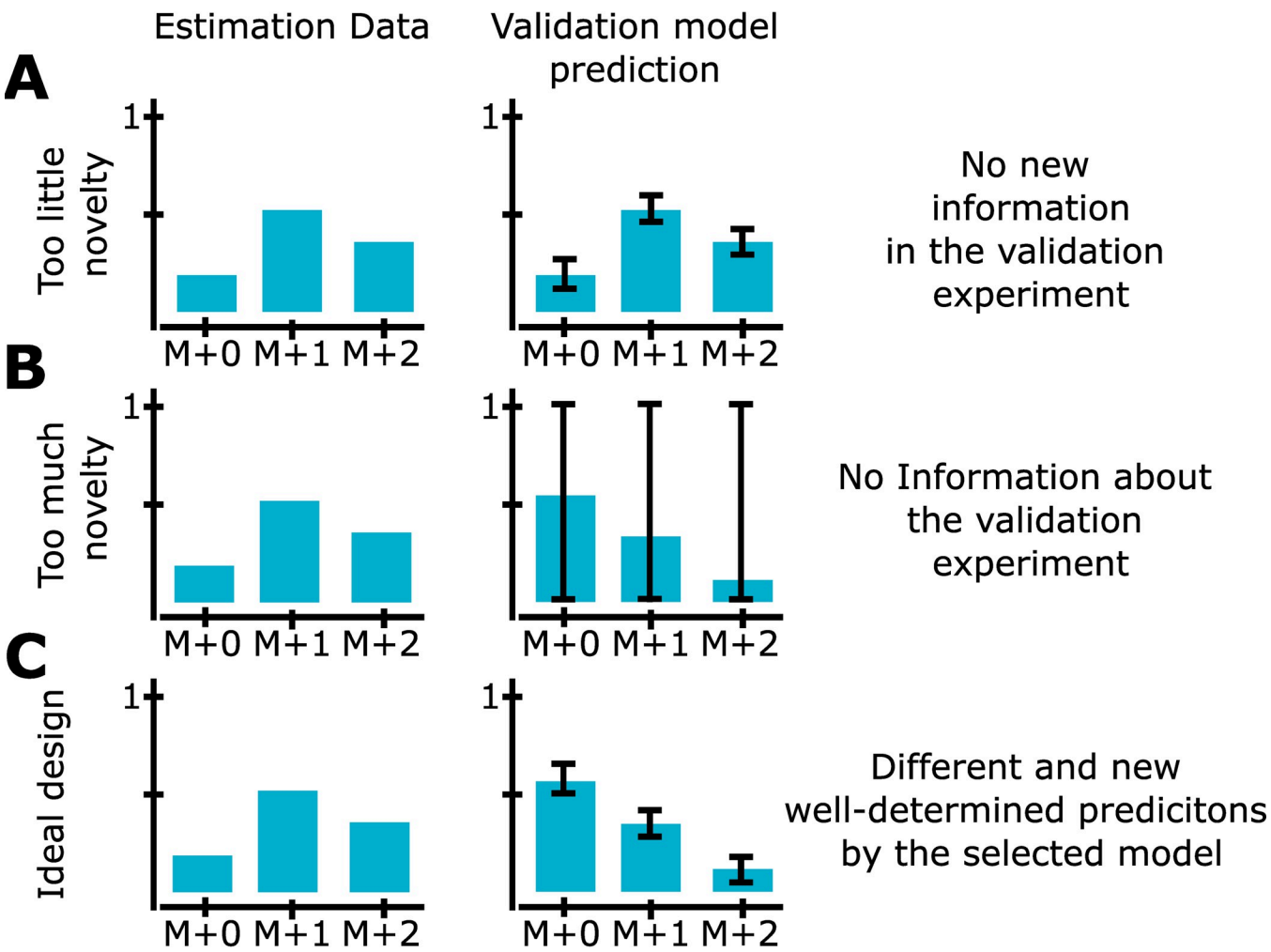

Fig 10. How prediction uncertainty can be used to assess the novelty in the validation data. (A) If there is too little novelty in the validation data, differences between estimation data and validation data will typically be smaller than the prediction and measurement uncertainty. (B) If there is too much novelty in the validation data, there is no information about the corresponding MIDs, and the prediction uncertainty will be large, approaching [0,1]. (C) An ideal design of validation data is thus to have well-determined predictions that are different compared to the estimation data. To be sure that there really is new information, one should also check that the new fluxes generate linearly independent EMU basis vectors (Section 2.4).

https://doi.org/10.1371/journal.pcbi.1009999.g010

the previous theoretical examples for this system, the true σ_r and the true model structure is now unknown. However, by evaluating the six model selection approaches for a range of different believed σ_b, it is clear that the results are consistent even for these data (Fig 12). The "SSR" method always chose the most complex model ($\mathcal{M}_7$, Fig 12A). The "First χ^2" method selected model $\mathcal{M}_4$ for $\sigma_b \approx 0.03$ (Fig 12B), while for the smaller σ_b, no model passed the χ^2-test. The "Best χ^2" method selected model $\mathcal{M}_6$ for $\sigma_b \approx 0.3$, 0.015, 0.003 (Fig 12C). Similarly, the BIC approach selects model $\mathcal{M}_6$ for all values of σ_b,while the AIC selects model $\mathcal{M}_6$ for $\sigma_b \approx 0.03$ and 0.015, and model $\mathcal{M}_7$ for $\sigma_b = 0.003$ (S5 Fig).

Similarly, the validation-based approach selected $\mathcal{M}_6$, regardless of σ_b (Fig 12D). Model $\mathcal{M}_6$ excludes reactions for unlabeled acetyl group entry into acetyl-CoA (included in $\mathcal{M}_7$), which represents catabolism of pre-existing fatty acids or acetate. Hence, the choice of $\mathcal{M}_6$ suggests that such entry does not occur in these cultures. This seems reasonable since the

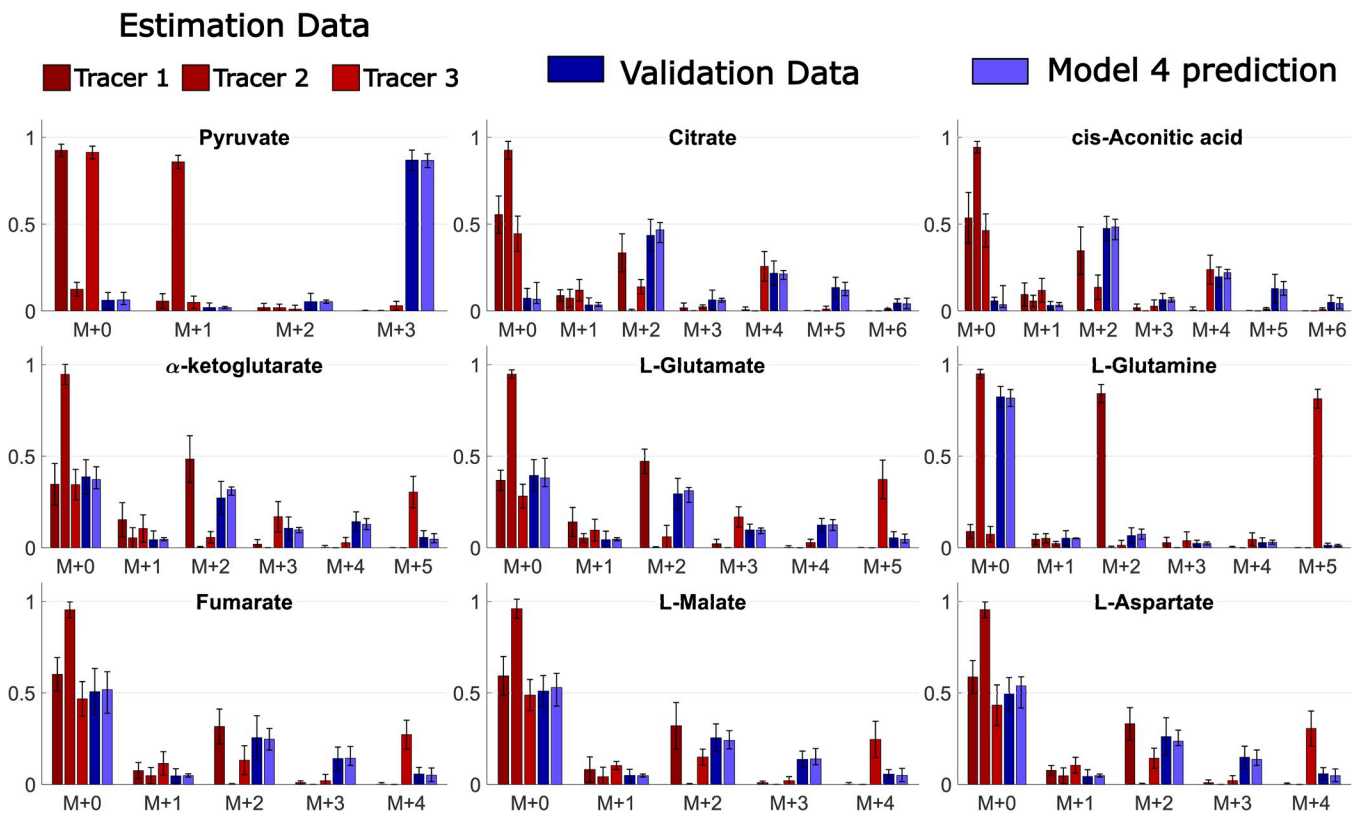

Fig 11. Usage of prediction uncertainty to demonstrate that the validation data has neither too little, nor too much, novelty, compared to the estimation data. This analysis shows the result from the simulated ^{13}C MFA example (Fig 7–9). The model was trained on estimation data corresponding to three tracers: Tracer 1 = 1,2-^{13}C-glutamine (dark red), Tracer 2 = 3-^{13}C-pyruvate (red), and Tracer 3 = U-^{13}C-glutamine (light red). The validation data (dark blue) came from usage of tracer U-^{13}C-pyruvate. For the experimental data, the error bars represent standard deviation, and for the model predictions (light blue), the error bars represent model uncertainty (Section 4.4).

https://doi.org/10.1371/journal.pcbi.1009999.g011

culture medium did not contain acetate, and was also free from serum and therefore contained very little fat. On the other hand, $\mathcal{M}_6$ includes the pyruvate carboxylase reaction while $\mathcal{M}_5$ does not, suggesting that this reaction was necessary to explain the data. Also, the pyruvate carboxylase flux was nonzero (95% confidence interval [0.08 0.98]). Interestingly, pyruvate carboxylase has been shown to be present in mammary epithelium *in vivo*, where it is important for *de novo* fatty acid synthesis [33] by replenishing the TCA cycle carbon that is consumed by citrate export ("anaplerosis"). To investigate if fatty acid synthesis also occurred in the cultured HMEC cells, we measured the MID of cellular lysophosphatidylcholine (LPC) 16:0 as a proxy for palmitate, which was not detectable with the methods used, after 7 days of ^{13}C labeling (Fig 13). The observed MID indicated that LPC 16:0 was a mixture of ^{13}C-labeled and unlabeled species, with higher mass isotopomers indicating that fatty acid synthesis indeed occurred. To further test the selected model structure $\mathcal{M}_6$, we used the estimated MID for cytosolic acetate (Fig 13B) from the fitted model to predict the MID of palmitate and LPC 16:0, assuming a linear mixture of pre-existing (unlabeled) and newly synthesized (labeled) species (Fig 13A). We found a reasonably good fit to the observed MID data at 82% newly synthesized LPC 16:0, indicating that the selected model $\mathcal{M}_6$ reflects actual lipid metabolism in this model system (Fig 13D).

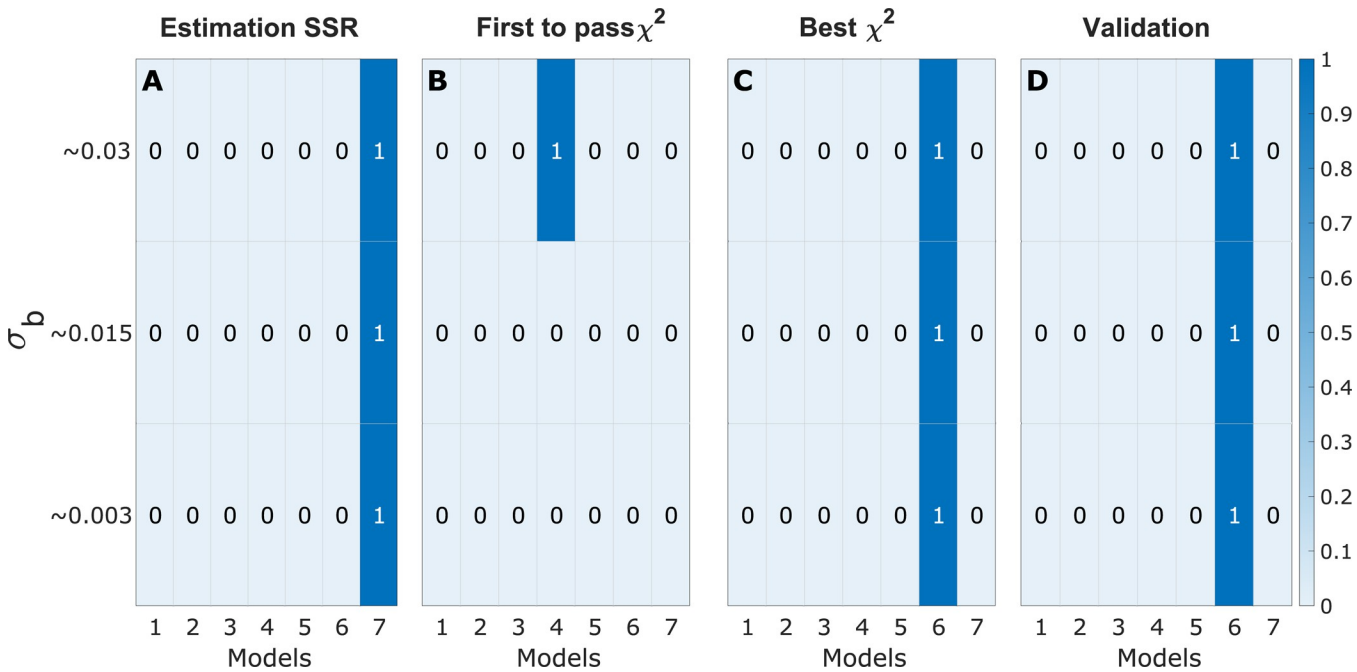

Fig 12. Model selection results for the cultures epithelial cell example. (A–D) Heatmaps represent results from the indicated selection methods, where rows represent different values of σ_b and columns represent the MFA models $\mathcal{M}_1, \ldots, \mathcal{M}_7$. For each row, color indicates the fraction of times a model is selected for the given σ_b, out of 1000 samples, as indicated by the color scale (right).

https://doi.org/10.1371/journal.pcbi.1009999.g012

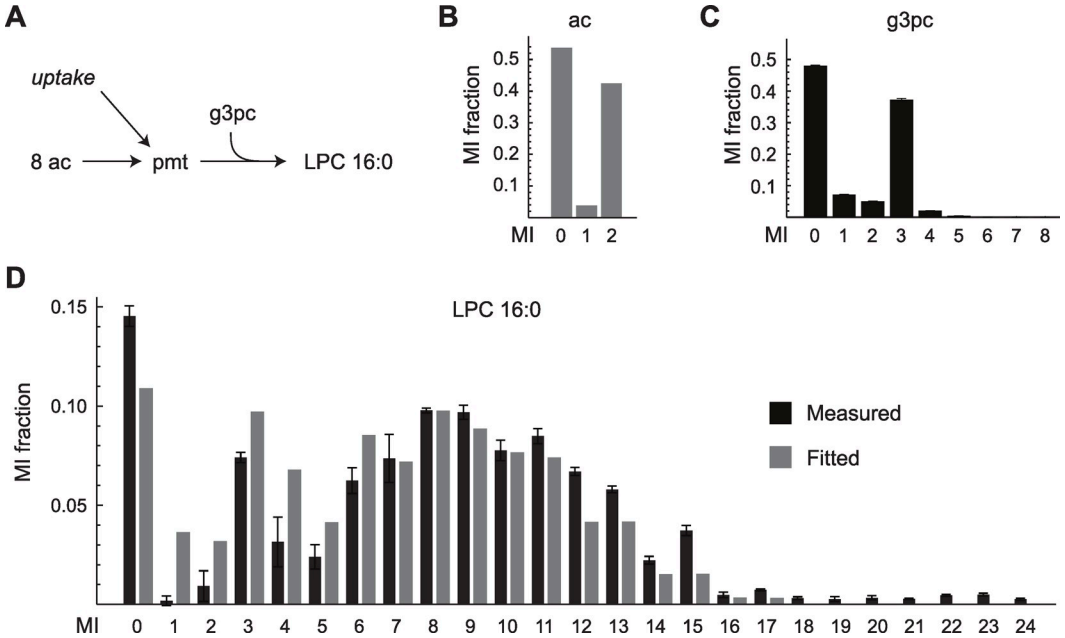

Fig 13. Validation of lipid synthesis in HMEC cultures. (A) Schematic of the model for lysophosphatidylcholine (LPC) 16:0 synthesis from acetate (ac). (B) Predicted MID of ac from the model selected by the "Validation" method. (C) Measured MID of glycerol-3-phosphocholine (g3pc). (D) Fitted (gray) and measured (black) MID of LPC 16:0. Mean values of biological triplicates are shown in (C, D). Error bars indicate standard deviation.

https://doi.org/10.1371/journal.pcbi.1009999.g013

3. Discussion

Since estimation of metabolic fluxes using ^{13}C MFA critically depends on the metabolic network model used, a systematic approach to model selection is of great importance. As we have demonstrated, commonly used model selection criteria such as the χ^2-test can give unpredictable results if the measurement error model is not accurate. Generally, we find that standard model selection methods that rely on a compensation for model complexity will choose different models depending on the "believed" standard deviation σ_b both for polynomial (Fig 4), linear model (Fig 6) and non-linear MFA models (Fig 8). Hence, when σ_b is inaccurate, these methods will over- or underfit the data, which naturally leads to errors in the estimated fluxes (Fig 9). Herein, we suggest remedying this problem by performing model selection on independent validation data, which is not used for estimating model parameters ("Validation" method). From our simulation studies, it is clear that this validation-based selection method indeed is more robust and selects the correct model in a way that is independent of errors in the size of σ_b (Figs 4, 6 and 8). Further, we demonstrate the importance of analysing the model's prediction uncertainty in order to generate confidence that the selected model accurately approximates the true metabolic system (Figs 10–11) Finally, to illustrate the potential with validation-based model development in MFA, we also applied it to new experimental data. For this data, the validation-based method consistently identifies a single model structure, whereas "traditional" methods that exclusively rely on estimation data again select different models depending on σ_b (Fig 12). Furthermore, we also support the choice of model by predicting the MID of LPS 16:0 with reasonable accuracy (Fig 13), which is synthesized as a result of the combination of factors that differentiates the selected model structure and the other alternatives. This illustrates yet another usage of validation data for model testing and model selection. In summary, validation-based model selection offers a more reliable approach to MFA model development when measurement errors are uncertain.

There are several reasons why the model (Eq (5)) of normal-distributed, independent errors may not be accurate for MID data. First, the since mass isotopomer fractions are constrained within [0,1] and sum to 1, strictly speaking they cannot be normal-distributed, nor independent. The normal assumption is particularly inaccurate for values close to 0 or 1, where the variance becomes very small. A better noise model for MIDs might be log-normal or other distributions on the n-simplex [16]. Moreover, MI fractions obtained from mass spectrometry can be biased for technical reasons: peak integration methods can affect MID accuracy [34], and minor isotopomers may underestimated due to limited sensitivity [15]. Finally, there are biological sources of error that are difficult to avoid. For example, in batch cultures, cells can never attain perfect metabolic steady-state, and there may be unforeseen kinds of compartmentalization, such as cell subpopulations, organelles, or reaction channeling [35]. Taken together, the result of these "hidden" factors is that observed standard deviations s will be artificially small compared to the residuals $y_i - \hat{y}_i$. While it could be argued that such biases constitute model error, and that the χ^2-test is correct in rejecting such models, it may be unrealistic to expect a perfect model fit in every scenario. Indeed, in many cases the estimated s is so small that it is exceedingly difficult to find a model that passes a χ^2-test, even for minor deviations from the error model. An important topic for future research is to address these issues by developing more suitable error models for MFA. However, in the meantime, validation-based model selection could offer a pragmatic way forward.

The rational for why the new validation-based method is robust with respect to errors in the magnitude of σ_b comes from general theory from the field System identification. This theory says that if the data has been generated by a "true" model structure $\mathcal{M}_0$ for some "true" parameters, θ_0, the estimated parameters will converge to θ_0 as the number of data points goes

to infinity. This theory assumes that the data used for model training is informative, i.e. that one would not gain any information regarding model distinction by exciting the system further, and that there is no redundancy of parameters, as is the case e.g. in structurally unidentifiable models. The convergence of the parameters to θ_0 holds for a large class of nonlinear model structures, which include all examples considered herein [36]. Furthermore, the overall magnitude of σ_b can be broken out from the cost function, and will thus not impact the location of any minima. These two facts together means that if the magnitude of sigma is the only thing that is wrong with sigma, the true model structure will still converge to the true parameters, while a too simple or too complex model structure will converge to the wrong parameters. This simple observation is the underlying motivation behind the validation-based method. Finally, note that the underlying theory from System identification assumes that in practice, the magnitude or scaling-error of σ_b is not necessarily the exact same for all datapoints, and the results herein indicate that the new validation-based method is a good choice also in such situations. In other words, the validation-based method is useful also in cases when the error in the believed σ is not homogeneously scaled for all data points. Furthermore, note that if one knows that the magnitude of the uncertainty for one metabolite is different form another metabolite, σ_b can still reflect this difference, and the presented results will still hold as long as the same difference in magnitude between metabolite measurement uncertainties also is the case for the true measurement uncertainty, σ_r. Finally, we believe that a validation-based approach is beneficial also in situations where σ is completely unknown, since a model that successfully predicts independent validation data probably is a decent description of reality. Note however that if the believed value of sigma, σ_b, is scaled wrong for some datapoint but not for others, the parameter estimation will be biased towards those datapoints and will converge to the wrong parameters. In this case, the predictions will be wrong, and the validation method may select the wrong model structure.

A key issue with the new validation method concerns how one divides data into an estimation and a validation data set. Clearly, the validation data must contain truly "novel" data: it is not sufficient to merely divide up replicate measurements y from the same experiment, which only differ by random noise. Herein, we have always used independent experiments with different inputs (tracers) x for validation data. In this form, validation-based model selection for MFA requires parallel experiments with distinct tracers, which naturally increases the experimental effort. An alternative might be to reserve certain measurement components y_i for the validation data set. In principle, the same methods for calculations of prediction uncertainty and validation-based model selection (Section 2.4) should be applicable also then, and preliminary analysis shows that this is indeed the case (S6 Fig).

The issue of which data points to reserve for the validation set is more difficult in our setting than in traditional cross-validation over statistically independent samples from a fixed data distribution. On one hand, highly dissimilar data points will be more difficult to predict, and should therefore provide a more stringent test for model selection. On the other hand, too dissimilar validation data ("extrapolation") may not be predictable by any model. To judge this tradeoff, the prediction uncertainty method in Section 2.4 is useful. This topic is also relevant for experimental design, which could be adapted to generate data suitable for informative validation data. Finally, an interesting aspect of these results is that also too simple models have a too large prediction uncertainty (S7 Fig), which is contrary to the traditional principle of bias-variance tradeoff, which says that only too complex models have a too high variance. This further emphasizes the fundamental differences between statistical methods based on sampling from the same distribution, compared to methods for mechanistic modelling, where data from different distributions can be used for the validation analysis. These results argue for a revision of such previously established truths, coming from statistics [13,31,37].

It is important to distinguish between model selection and model testing. As mentioned earlier, while our method allows selecting the best model from a given set, it does not guarantee that the best model is indeed acceptable. While a goodness-of-fit test could be performed on the validation data for the selected best model, such a test will in general be optimistic due to multiple testing over models. For proper model testing, a third "test" data set should be used, which is not used for either parameter estimation of model selection.

The analysis in Fig 8 is only meant to illustrate that errors in model structure can lead to errors in flux estimation, and is not an exhaustive analysis of what these errors might look like. The fluxes depicted in Fig 8 is considered a representative selection of the fluxes that overlaps between all model structures. The overall fact that errors in model structure may lead to more or less large errors in flux estimations should hold true.

Based on our results, we suggest that validation-based model selection should always be considered when developing MFA models. Nevertheless, for small errors in σ_b, the less computationally expensive methods, such as AIC and BIC may give the same results. The problem, however, is that one does not know when the dependency on errors in σ_b make those methods unreliable in cases were the magnitude of the experimental error is uncertain, and in such cases it is therefore safer to use a validation-based approach. Validation-based approaches also have important advantages related to interpretation, and are therefore common-place in other field.

We believe that the field of MFA modelling should take inspiration from such other fields of computational biology, where the ability to correctly predict independent data *not* used for parameter estimation is a standard criterion for model quality, and where such validation tests often are a requirement for publication [20,21,23,24,26,27]. Given that models are always simplifications of reality, such independent validation is important both for the modelling process and for communicating results to non-experts users. In other words, while it is almost always wrong to assume steady-state metabolism occurring in a single average cell, such a model may still be a good enough approximation of reality to produce realistic fluxes. Importantly, one way to demonstrate the realism and general predictive power of the chosen model is to show that it can predict new independent validation data. Notably, in guidelines issued by the US Food and Drug Administration (FDA), testing on independent validation data is necessary condition for a model to be considered trustworthy [38]. All in all, we believe that validation-based model selection provides sound and reliable checking of metabolic models, which we hope will be of value also to the ^{13}C MFA field.

4. Materials and methods

4.1 ^{13}C Metabolic flux analysis

As stated previously, the gold standard method for measuring metabolic fluxes in a given system is model-based metabolic flux analysis with isotopically labelled tracers. The model $\mathcal{M}$ includes the stoichiometry and atom mappings for each reaction, and is parameterized by the metabolic fluxes v, or more precisely, by the independent fluxes u. At steady state, the model-predicted MIDs $\hat{y}$ are uniquely determined by u together with the known isotope distributions x of the network substrates [9,39],

$$\hat{y} = h(x, u) \tag{3}$$

where the function h is determined by the model $\mathcal{M}$. Model fitting is done by seeking the vector u that minimizes the sum of the squared weighted residuals (SSR) between the model-

predicted and measured MIDs [40],

$$f(u) = SSR = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\hat{y}_j^i - y_j^i}{\sigma_{ij}}\right)^2 \tag{4}$$

$$y_j^i = h_j(x^i, u_0) + \epsilon_{ij}\ \epsilon_{ij}\ \epsilon\ N(0, \sigma_{ij}) \tag{5}$$

Here each measurement y_j^i is assumed to derive from the model prediction $\hat{y}_j^i$ at the true flux vector u_0, plus a normal-distributed noise ϵ_{ij} with standard deviation σ_{ij}. If there are several experiments with different tracers x^i, the sum is taken over all resulting measurement vectors y^i [41] Under these assumptions, $f(u)$ follows a χ^2-distribution, and so the χ^2-test can be used to assess model fit [42]. If this test does not reject, the model $\mathcal{M}$ and the inferred fluxes u are considered valid.

4.1 Construction of mathematical models: Predictors and the EMU framework

The mathematical models presented herein are formulated such that a mathematical structure describes one or more predictors $\hat{y}_i(\theta)$, given a set of model parameters θ. These mathematical structures or models can exist in different forms, such as polynomial models or ordinary differential equation (ODE) models. For MFA, a common approach for model formulation is the Elementary metabolite units (EMUs) framework [34]. In short, the EMU framework allows for a decomposition of the model such that only the information necessary to calculate a desired set of MIDs remains. The metabolic network is broken down into EMU subnetworks that are used to formulate equations, of the form in Eq (6) below [43].

$$A_{n,k}(v) * X_{n,k} = B_{n,k}(v) * Y_{n,k}(x, X_{n-1}, \ldots, X_1) \tag{6}$$

where index n indicates the size of the EMU network and index k is used to index several networks of the same size; where matrices A and B contains the model structure for the fluxes v, which can be parameterized according using a smaller set of independent fluxes u; where matrices $X_{n,k}$ and $Y_{n,k}$ contains the unknown and known EMU variables, respectively; and where x are the EMU variables that correspond to the system tracer [43]. In other words, for EMU models $\theta = u$.

4.2 Optimization of model parameters to fit the data

The objective of the parameter estimation step in any modelling problem is to minimize an objective function $f(\theta)$ which determines the agreement with data for a given the set of parameters θ. The general optimisation problem, which determines the optimal parameters θ^*, is formulated as

$$\theta^* = \arg\min_{\theta} f(\theta)$$

$$s.t. g_j(\theta) \geq 0\ \forall j \tag{7}$$

where $g_j(\theta)$ are functions describing constraints applied to the optimization problem. Again, for ^{13}C MFA modelling, the parameters θ are the independent fluxes, herein denoted u. Also,

for ^{13}C MFA modelling, two constraints are usually placed on these independent fluxes:

$$g_1(u): \ null(S) * u \geq 0$$

$$g_2(u): \ ub \geq u \geq lb$$

where $null(s)$ is a null space matrix of the network's stoichiometry matrix; and where ub and lb are the upper and lower bounds of u, respectively. The first condition ensures that all fluxes, which are given by the product of $null(S)$ and u, are positive. The second condition ensures that the independent fluxes are constrained within a predetermined interval. As for the detailed form of the objective function, f, it can vary depending on the specific analysis conducted, but will generally be some variant of the weighted SSR function, since this objective function has sound theoretical properties [36]. The SSR used herein is given by Eq (4). For the EMU model, the relationships between y, $\hat{y}$, σ, and the state variables X are given by:

$$\hat{y}_i(u) = X_{n,k}(l, m) \tag{8}$$

where y_i is the measured value for the mth mass fraction of the lth EMU in $X_{n,k}$, i.e. a specific bar in Fig 1C. If $\hat{y}_i$ is a mean value of multiple original data points, then the residuals $(y_i - \hat{y}_i(u))$, should be weighted with the standard error of the mean (SEM_i) rather than σ_i should be used in Eq (4). The relation between the two is given by:

$$SEM_i = \frac{\sigma_i}{\sqrt{N}} \tag{9}$$

where N is the number of sample points. However, all of these are theoretical truths; which denominator to use in Eq (4) is an unresolved issue for ^{13}C MFA models (see Section 3.2).

4.3 χ^2-test

In the ^{13}C MFA modelling field, the χ^2-test is the statistical hypothesis test that most commonly is employed to evaluate whether the SSR is small enough, i.e. if the model can be considered an accurate representation of the target system. In practice, the SSR is compared with the inverse cumulative χ^2-distribution, where the degrees of freedom is given by the number of datapoints adjusted for the fact that some independence between the datapoints and the model is lost by estimating parameters to the same data that is used for testing. This compensation can be done in different ways; the naive way is to do no compensation at all, and the most conservative way is to compensate for all parameters (free fluxes) in the model. The most accurate version is to instead use the number of practically identifiable parameters. In reality, this adjustment is done in different ways, often without justification, and these differences may be the reason why a model is, or is not, rejected. This ambiguity is one of the reasons arguing against the usage of this test. Its dependency on the value of σ_i is another such argument. With the most conservative choice, the algorithm becomes:

Input: model structure $\mathcal{M}$, parameters θ, data D with N datapoints.

1. Calculate the combined SSR for all data points N, using (Eq (4)).
2. *If* $SSR < \chi^{2,cuminv}(p = 0.95, N - \theta)$, i.e. the cumulative inverse of a χ^2 distribution, then model structure $\mathcal{M}$ is accepted

   else model structure $\mathcal{M}$ is rejected.

Output: FAIL if model structure $\mathcal{M}$ is rejected OR PASS if model structure $\mathcal{M}$ acceptable with respect to D.

4.4 Determining model prediction uncertainty

In this work two main approaches were used to determine the prediction uncertainties. The first and primary approach was a prediction profile likelihood (PPL) analysis. A prediction profile likelihood analysis is used to determine the uncertainty of a predicted model value or property [44–46]. The PPL-analysis was implemented by modifying the function for the SSR, seen in Eq (4), such that it contained an additional term such as:

$$f(\theta) = \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\hat{y}_j^i - y_j^i}{\sigma_{i,j}}\right)^2 + W * (p_{target}^{\omega} - p_{sim}^{\omega}(\theta))^2 \tag{10}$$

where p_{sim}^{ω} is the simulated relative abundance of mass fraction ω and is determined by the parameters θ. p_{target}^{ω} is a set target value for mass fraction ω and W is an integer with an arbitrary large value. By assigning a very large value to W, any difference between p_{sim}^{ω} and p_{target}^{ω} will be magnified. Thus, the optimization process will select parameters that minimizes this difference. Then, p_{target}^{ω} was gradually stepped away from the optimal simulated value of the mass fraction ω, until a cutoff value is reached. This stepping process is repeated for all mass fractions that are included in the prediction.

The second approach used for determining the model prediction uncertainty was estimation through Markov chain Monte Carlo (MCMC) sampling. For this analysis a posterior distribution of parameter values is generated, and all parameter sets that are acceptable with respect to the estimation data are collected. The model prediction uncertainty is then determined by the interval:

$$CI_{\alpha,Dof} = [f(\theta) \leq f(\theta^*) \pm \Delta_\alpha(\chi^2_{DoF})] \tag{11}$$

where $f(\theta)$ is the generalise form of the SSR objective function described in Eq (4); $f(\theta^*)$ is the SSR function value for the optimal parameters; α is the confidence level; $\Delta_\alpha(\chi^2)$ is the quantile of the χ^2-statistic; DoF is the degrees of freedom; and θ^* are the optimal parameters. In this work, the DoF is equal to the number of model parameters, i.e. in the ^{13}C MFA examples the number of free fluxes, and 10^5 samples were used for the sampling.

4.5 Simulated data generation

For the examples presented in this paper, simulated data is utilized to create scenarios for model estimation in which the ground truth is known. In each of these examples, a given model $\mathcal{M}_0$ with parameters θ_0 has been used to generate values for selected variables of interest, given a predetermined set of model parameters and inputs. To these values, a normally distributed noise was added, with a given true sigma, σ_r.

For the polynomial exampled a seventh order polynomial were used as true model (see S2 Table, for true parameter values), and a normally distributed noise was added with $\sigma_r = 0.2$. The model was then fitted to data corresponding to one realisation of this noise i.e. $i = 1$, using Eq (4).

For the linear model example, the model A_3 (Fig 5) was used as the true model (see S2 Table for true parameter values). A normally distributed noise was added with a true sigma of $\sigma_r = 5$ The model was then fitted to data corresponding to five realisations of this noise i.e. $i = 5$, using Eq (4).

For the EMU-model example, model 4 (Fig 7) was used as the true models (see S2 Table for true parameter values). A normally distributed noise was added with a true sigma of $\sigma_r = 0.03$ The model was then fitted to data corresponding to 3 realisations of this noise i.e. $i = 3$, using

Eq (4). The data was generated from four different tracers. These tracers were U-^{13}C-pyruvate, U-^{13}C-glutamine, 3-^{13}C-pyruvate, and 1,2-^{13}C-glutamine. For the "Validation" method the data from tracers U-^{13}C-glutamine, 3-^{13}C-pyruvate, and 1,2-^{13}C-glutamine was used as D^{est} while data from U-^{13}C-pyruvate was used as D^{val}

To generate the different values of σ_b, the sample standard deviation for each observation was scaled by a number drawn from a uniform random distribution. For $\sigma_b \approx 10^* \sigma_r$ the distribution range was between [8 12], for $\sigma_b \approx \sigma_r$ the distribution range was between [0.8 1.2], and for $\sigma_b \approx 0.1^* \sigma_r$ the distribution range was between [0.08 0.012].

4.6 Cell culture and isotope tracing

Human Mammary Epithelial Cells (HMECs) were obtained from the laboratory of William C. Hahn (Dana-Farber Cancer Institute, Boston, USA) and have been previously described [47]. HMECs were grown in custom-synthesized Mammary Epithelial Basal Medium (MCDB) 170 [48] supplemented with 1% Mammary Epithelial Growth Supplement (MEGS) (S0155, Gibco), 100 units/ml penicillin and 100 μg/ml streptomycin (15140122, Thermo Fisher Scientific). Cells were kept in a humidified atmosphere of 5% CO_2/95% air at 37˚C and washed and detached using ReagentPack Subculture Reagents (CC-5034, Lonza).

For isotope tracing experiments, 400,000 cells were seeded at day 0 in a T25 flask in 5mL medium and incubated overnight. On day 1, medium was changed to an MCDB 170 medium of the same molar composition, but with glucose or glutamine exchanged for U-^{13}C-glucose or U-^{13}C-glutamine (Cambridge Isotope Laboratories), respectively. On day 2, each T25 flask culture was detached and seeded into in two T25 flasks, using the same medium. On day 4, cells were detached and seeded into 6-well plates at 250,000 cells/well in 2mL of medium, in triplicate for each tracer. On day 7 (after roughly 6 cell divisions in the presence of each ^{13}C tracer), each multi-well plate was placed on ice, medium was aspirated and cells were washed twice with 1 mL of cold PBS. Then, 1 mL cold (–80˚C) methanol (JT Baker, BAKR8402.2500, VWR) was added to each well, cells were scraped using a 17mm cell scraper (83.1830, Sarstedt), and the extracts were carefully transferred to a new tube, vortexed for 30 seconds to break up aggregated cell material, and stored in -80˚C until analysis. LCMS analysis of cell extracts was performed using a pHILIC LC column coupled to a Thermo QExactive orbitrap mass spectrometer, as previously described [49]. All metabolite peaks reported were confirmed against pure standards. Peak areas were integrated directly from instrument data using the mzAccess data access framework [50] and Mathematica v.11.1 (Wolfram Research). Mass isotopomer distributions were calculated as the areas of each mass isotopomer peak, divided by the total peak area for all mass isotopomers.

4.7 Model for the LPC 16:0 MID

The MID x_{ac} for acetate (ac) was obtained from the fitted model as described above in section 2.5. The MID x_{pmt} of total cellular pool was modeled as a linear mixture

$$x_{pmt} = \alpha\, x_{pmt}^{new} + (1-\alpha) x_{pmt}^{old} \tag{12}$$

where x_{pmt}^{new} is the MID of newly synthesized palmitate (pmt), computed by convolution of x_{ac} eight times, modeling the condensation of eight ac molecules by fatty acid synthase; x_{pmt}^{old} is the natural MID, representing pre-existing palmitate; and α is the unknown mixture coefficient. The lysophosphatidylcholine MID x_{lpc} was modeled as a convolution of x_{pmt} and the measured glycerol-3-phosphocholine (g3pc) MID x_{g3pc}. This can be written as:

$$x_{lpc} = C(x_{g3pc}) x_{pmt} \tag{13}$$

where $C(x_{g3pc})$ is a matrix whose elements depend on x_{g3pc}. This yields an equation system linear in the unknown α, which was solved using the least-squared method.

4.8 Software

All analysis presented here were performed in MATLAB by The MathWorks Inc., release 2020b. For the models presented in this paper the open source MATLAB toolbox OpenFLUX2 [51] was employed to transform the network structure to the EMU equation systems.

Supporting information

S1 Fig. Model selection results for the polynomial model example. Heatmaps represent results from the AIC (left) and BIC (right) methods, where rows represent different values of σ_b and columns represent the polynomial models $h_1,\ldots,h_{14}$. For each row, colour indicates the fraction of times a model is selected for the given σ_b, out of 10,000 samples, as indicated by the colour scale (right).
(EPS)

S2 Fig. Model selection results for the linear model example. Heatmaps represent results from the AIC (left) and BIC (right) methods, where rows represent different values of σ_b and columns represent the polynomial models $A_1,\ldots,A_6$. For each row, colour indicates the fraction of times a model is selected for the given σ_b, out of 1 000 samples, as indicated by the colour scale (right).
(EPS)

S3 Fig. Illustration of the data and simulation for the Linear example. The simulated data for the linear example plotted with the different input vectors x along the x-axis and the model output y on the y-axis. The three different output variables are indicated by the different colours, y_1–purple, y_2–Green, and y_3–orange. The division into estimation data (left) and validation data (right) is indicated by the vertical line.
(EPS)

S4 Fig. Model selection results for the simulated 13C MFA model example. Heatmaps represent results from the AIC (left) and BIC (right) methods, where rows represent different values of σ_b and columns represent the polynomial models $\mathcal{M}_1,\ldots,\mathcal{M}_7$. For each row, colour indicates the fraction of times a model is selected for the given σ_b, out of 100 samples, as indicated by the colour scale (right).
(EPS)

S5 Fig. Model selection results for the epithelial cell example. Heatmaps represent results from the AIC (left) and BIC (right) methods, where rows represent different values of σ_b and columns represent the polynomial models $\mathcal{M}_1,\ldots,\mathcal{M}_7$.
(EPS)

S6 Fig. Usage of Validation data with prediction uncertainty where a sub part of a single data set has been reserved for validation. This preliminary analysis shows the result of using validation with prediction uncertainty in a scenario where a portion of a complete data has been reversed as validation data. In this example the model was trained on estimation data (dark red) consisting of MIDs from 8 metabolites, from two tracers, with the model fit to the estimation data (light red) is illustrated for each MID. The validation data consisted of the MID for α-ketoglutarate (dark blue) and the model prediction is illustrated (light blue) and shows good agreement with the validation data. For the experimental data, the error bars represent standard deviation, and for the model predictions, the error bars represent model

uncertainty (Section 4.4). The tracers are Tracer 1 = U-^{13}C-glutamine, Tracer 2 = U-^{13}C-pyruvate.
(EPS)

S7 Fig. A comparison of model predictions with uncertainties compared to the validation data for the simulated ^{13}C MFA example. The validation data for the simulated ^{13}CMFA example consisted of MIDs for nine metabolites from a [U-13C] pyruvate tracer. The metabolites are from top-left to bottom-right, pyruvate, citrate, cis-aconitic acid, alpha-ketoglutarate, L-glutamate, L-glutamine, fumarate, L-malate, and L-aspartate. The purple bars indicate the simulated MIDs, and the corresponding error bars indicate the experimental uncertainty of σ_r. The red bars indicate the predicted MIDs of model structure $\mathcal{M}_4$ and the corresponding error bars indicate the prediction uncertainty. The green bars indicate the predicted MIDs of model structure $\mathcal{M}_2$ and the corresponding error bars indicate the prediction uncertainty.
(EPS)

S1 Algorithm. Algorithm used for the model selection problem. The model selection algorithm takes a set of model structures and a set of data as inputs and selects the most appropriate model structure based on the sub type (A-D) and the data. Subtype A selects the model structure that yields the smallest summed squared residuals (SSR) given the entire data set. Subtype B selects the first/simplest model structure that can pass a χ^2-test. Subtype C selects the model structure that passes a χ^2-test with the largest margin. Finally, subtype D selects the model structure that yields the lowest SSR with respect to a validation subset of the data.
(DOCX)

S1 Table. Breakdown of reactions for the TCA-cycle models. The table contains a detailed breakdown of the reactions that are included in the TCA-models. In general, the model structures are derived from the most complex model structure $\mathcal{M}_7$ by successively removing and combining reactions and thus make each successive model structure simpler.
(DOCX)

S2 Table. A summary of the parameter values that were used to generate the simulated data for the three different examples that are used in the manuscript. The full parameter vector for the polynomial, linear and metabolic flux analysis model examples are given by the respective columns.
(DOCX)

Author Contributions

Conceptualization: Nicolas Sundqvist, Roland Nilsson, Gunnar Cedersund.

Data curation: Nicolas Sundqvist, Nina Grankvist, Jeramie Watrous, Jain Mohit.

Formal analysis: Nicolas Sundqvist, Jeramie Watrous, Jain Mohit, Roland Nilsson.

Funding acquisition: Roland Nilsson, Gunnar Cedersund.

Investigation: Nicolas Sundqvist, Roland Nilsson.

Methodology: Nicolas Sundqvist, Roland Nilsson, Gunnar Cedersund.

Project administration: Gunnar Cedersund.

Resources: Nina Grankvist.

Software: Nicolas Sundqvist.

Supervision: Roland Nilsson, Gunnar Cedersund.

Validation: Nicolas Sundqvist, Roland Nilsson, Gunnar Cedersund.

Visualization: Nicolas Sundqvist.

Writing – original draft: Nicolas Sundqvist, Roland Nilsson, Gunnar Cedersund.

Writing – review & editing: Nicolas Sundqvist, Roland Nilsson, Gunnar Cedersund.

References

1. Berg JM, Tymoczko JL, Stryer L. Biochemistry. [Internet]. 7th ed. W.H. Freeman & Company 2010; 2010. Available from: https://login.e.bibl.liu.se/login?url = https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=cat00115a&AN=lkp.573107&lang=sv&site=eds-live&scope=site
2. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. Nature Biotechnology. 2013 Mar 3; 31(5):419–25. https://doi.org/10.1038/nbt.2488 PMID: 23455439
3. Sinclair L V, Rolf J, Emslie E, Shi Y-B, Taylor PM, Cantrell DA. Control of amino-acid transport by antigen receptors coordinates the metabolic reprogramming essential for T cell differentiation. Nature Immunology. 2013 May 24; 14(5):500–8. https://doi.org/10.1038/ni.2556 PMID: 23525088
4. Anderson RM, Weindruch R. Metabolic reprogramming, caloric restriction and aging. Trends in Endocrinology & Metabolism. 2010 Mar 1; 21(3):134–41. https://doi.org/10.1016/j.tem.2009.11.005 PMID: 20004110
5. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell. 2011 Mar 4; 144(5):646–74. https://doi.org/10.1016/j.cell.2011.02.013 PMID: 21376230
6. Hiller K, Metallo CM. Profiling metabolic networks to study cancer metabolism. Current Opinion in Biotechnology. 2013 Feb 1; 24(1):60–8. https://doi.org/10.1016/j.copbio.2012.11.001 PMID: 23206561
7. Martin SD, McGee SL. Metabolic reprogramming in type 2 diabetes and the development of breast cancer. The Journal of endocrinology. 2018 May 1; 237(2):R35–46. https://doi.org/10.1530/JOE-18-0037 PMID: 29487204
8. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. Cell Communication and Signaling. 2013 May 17; 11(1):34. https://doi.org/10.1186/1478-811X-11-34 PMID: 23683503
9. Wiechert W. 13C Metabolic Flux Analysis. Metabolic engineering. 2001; 3(3):195–206. https://doi.org/10.1006/mben.2001.0187 PMID: 11461141
10. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, et al. A roadmap for interpreting 13 C metabolite labeling patterns from cells. Current Opinion in Biotechnology. 2015 Aug; 34:189–201. https://doi.org/10.1016/j.copbio.2015.02.003 PMID: 25731751
11. Long CP, Antoniewicz MR. High-resolution 13C metabolic flux analysis. Nat Protoc. 2019 Oct; 14(10):2856–77. https://doi.org/10.1038/s41596-019-0204-0 PMID: 31471597
12. Crown SB, Antoniewicz MR. Publishing 13C metabolic flux analysis studies: a review and future perspectives. Metabolic engineering. 2013 Nov; 20:42–8. https://doi.org/10.1016/j.ymben.2013.08.005 PMID: 24025367
13. Cedersund G, Roll J. Systems biology: model based evaluation and comparison of potential explanations for given biological data. FEBS Journal. 2009 Feb 1; 276(4):903–22. https://doi.org/10.1111/j.1742-4658.2008.06845.x PMID: 19215297
14. Nilsson R, Jain M. Simultaneous tracing of carbon and nitrogen isotopes in human cells. Molecular BioSystems. 2016 May; 12(6):1929–37. https://doi.org/10.1039/c6mb00009f PMID: 27098229
15. Su X, Lu W, Rabinowitz JD. Metabolite Spectral Accuracy on Orbitraps. Analytical chemistry. 2017 Jun; 89(11):5940–8. https://doi.org/10.1021/acs.analchem.7b00396 PMID: 28471646
16. Aitchison J. The Statistical Analysis of Compositional Data. Journal of the Royal Statistical Society Series B (Methodological). 1982 Jun; 44(2):139–77.
17. Dalman T, Wiechert W, Nöh K. A scientific workflow framework for (13)C metabolic flux analysis. Journal of biotechnology. 2016 Aug; 232:12–24. https://doi.org/10.1016/j.jbiotec.2015.12.032 PMID: 26721184
18. Johansson R, Strålfors P, Cedersund G. Combining test statistics and models in bootstrapped model rejection: it is a balancing act. BMC systems biology. 2014 Apr 17; 8:46. https://doi.org/10.1186/1752-0509-8-46 PMID: 24742065

19. Brännmark C, Palmér R, Glad ST, Cedersund G, Strålfors P. Mass and information feedbacks through receptor endocytosis govern insulin signaling as revealed using a parameter-free modeling framework. The Journal of biological chemistry. 2010/04/26 ed. 2010 Jun 25; 285(26):20171–9. https://doi.org/10.1074/jbc.M110.106849 PMID: 20421297

20. Nyman E, Brännmark C, Palmér R, Brugård J, Nyström FH, Strålfors P, et al. A hierarchical whole-body modeling approach elucidates the link between in Vitro insulin signaling and in Vivo glucose homeostasis. The Journal of biological chemistry. 2011 Jul; 286(29):26028–41. https://doi.org/10.1074/jbc.M110.188987 PMID: 21572040

21. Nyman E, Fagerholm S, Jullesson D, Strålfors P, Cedersund G. Mechanistic explanations for counter-intuitive phosphorylation dynamics of the insulin receptor and insulin receptor substrate-1 in response to insulin in murine adipocytes. The FEBS journal. 2012 Mar; 279(6):987–99. https://doi.org/10.1111/j.1742-4658.2012.08488.x PMID: 22248283

22. Schmidl D, Hug S, Li WB, Greiter MB, Theis FJ. Bayesian model selection validates a biokinetic model for zirconium processing in humans. BMC systems biology. 2012 Aug 5; 6:95. https://doi.org/10.1186/1752-0509-6-95 PMID: 22863152

23. Timmer J, Müller TG, Swameye I, Sandra O, Klingmüller U. Modeling the nonlinear dynamics of cellular signal transduction. International Journal of Bifurcation and Chaos in Applied Sciences and Engineering. 2004; 14(6):2069–79.

24. Müller TG, Faller D, Timmer J, Swameye I, Sandra O, Klingmüller U. Tests for cycling in a signalling pathway. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2004 Nov 1; 53(4):557–68.

25. Wagenmakers E-J, Ratcliff R, Gomez P, Iverson GJ. Assessing model mimicry using the parametric bootstrap. Journal of Mathematical Psychology. 2004; 48:28–50.

26. Mélykúti B, August E, Papachristodoulou A, El-Samad H. Discriminating between rival biochemical network models: three approaches to optimal experiment design. BMC systems biology. 2010 Apr 1; 4:38. https://doi.org/10.1186/1752-0509-4-38 PMID: 20356406

27. Roberts MAJ, August E, Hamadeh A, Maini PK, McSharry PE, Armitage JP, et al. A model invalidation-based approach for elucidating biological signalling pathways, applied to the chemotaxis pathway in R. sphaeroides. BMC systems biology. 2009 Oct 31; 3:105. https://doi.org/10.1186/1752-0509-3-105 PMID: 19878602

28. Schwarz G. Estimating the Dimension of a Model. The Annals of Statistics. 1978; 6(2):461–4.

29. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected Papers of Hirotugu Akaike [Internet]. New York, NY: Springer; 1998 [cited 2021 Apr 28]. p. 199–213. (Springer Series in Statistics). Available from: https://doi.org/10.1007/978-1-4612-1694-0_15

30. Antoniewicz MR. A guide to 13C metabolic flux analysis for the cancer biologist. Experimental & Molecular Medicine. 2018 Apr; 50(4):19. https://doi.org/10.1038/s12276-018-0060-y PMID: 29657327

31. Cedersund G. Conclusions via unique predictions obtained despite unidentifiability–new definitions and a general method. The FEBS journal. 2012 Sep; 279(18):3513–27. https://doi.org/10.1111/j.1742-4658.2012.08725.x PMID: 22846178

32. Crown SB, Antoniewicz MR. Selection of tracers for 13C-Metabolic Flux Analysis using Elementary Metabolite Units (EMU) basis vector methodology. Metabolic Engineering. 2012 Mar; 14(2):150–61. https://doi.org/10.1016/j.ymben.2011.12.005 PMID: 22209989

33. Gul B, Dils R. Pyruvate carboxylase in lactating rat and rabbit mammary gland. Biochem J. 1969 Feb; 111(3):263–71. https://doi.org/10.1042/bj1110263 PMID: 5812939

34. Antoniewicz MR, Kelleher JK, Stephanopoulos G. Accurate assessment of amino acid mass isotopomer distributions for metabolic flux analysis. Anal Chem. 2007 Oct 1; 79(19):7554–9. https://doi.org/10.1021/ac0708893 PMID: 17822305

35. Wahrheit J, Nicolae A, Heinzle E. Eukaryotic metabolism: Measuring compartment fluxes. Biotechnology Journal. 2011 Sep; 6(9):1071–85. https://doi.org/10.1002/biot.201100032 PMID: 21910257

36. Ljung L. System Identification (2nd Ed.): Theory for the User. Upper Saddle River, New Jersey 07458: Prentice Hall PTR; 1999.

37. Neal B, Mittal S, Baratin A, Tantia V, Scicluna M, Lacoste-Julien S, et al. A Modern Take on the Bias-Variance Tradeoff in Neural Networks. arXiv:181008591 [cs, stat] [Internet]. 2019 Dec 18 [cited 2021 Nov 9]; Available from: http://arxiv.org/abs/1810.08591

38. Viceconti M, Juarez MA, Curreli C, Pennisi M, Russo G, Pappalardo F. Credibility of in Silico Trial Technologies-A Theoretical Framing. IEEE Journal of Biomedical and Health Informatics. 2020 Jan 1; 24(1):4–13. https://doi.org/10.1109/JBHI.2019.2949888 PMID: 31670687

39. Antoniewicz MR. Methods and advances in metabolic flux analysis: a mini-review. Journal of industrial microbiology & biotechnology. 2015 Mar 23; 42(3):317–25. https://doi.org/10.1007/s10295-015-1585-x PMID: 25613286

40. Dai Z, Locasale JW. Understanding metabolism with flux analysis: From theory to application. Metabolic Engineering. 2016 Sep 22; 43:94–102. https://doi.org/10.1016/j.ymben.2016.09.005 PMID: 27667771

41. Antoniewicz MR. Parallel labeling experiments for pathway elucidation and 13C metabolic flux analysis. Current Opinion in Biotechnology. 2015 Dec 1; 36:91–7. https://doi.org/10.1016/j.copbio.2015.08.014 PMID: 26322734

42. Antoniewicz MR, Kelleher JK, Stephanopoulos G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metabolic engineering. 2006 Jul; 8(4):324–37. https://doi.org/10.1016/j.ymben.2006.01.004 PMID: 16631402

43. Antoniewicz MR, Kelleher JK, Stephanopoulos G. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. Metabolic engineering. 2007 Jan; 9(1):68–86. https://doi.org/10.1016/j.ymben.2006.09.001 PMID: 17088092

44. Kreutz C, Raue A, Timmer J. Likelihood based observability analysis and confidence intervals for predictions of dynamic models. BMC Systems Biology. 2012 Sep; 6(1):120. https://doi.org/10.1186/1752-0509-6-120 PMID: 22947028

45. Kreutz C, Raue A, Kaschek D, Timmer J. Profile likelihood in systems biology. The FEBS Journal. 2013; 280(11):2564–71. https://doi.org/10.1111/febs.12276 PMID: 23581573

46. Maiwald T, Hass H, Steiert B, Vanlier J, Engesser R, Raue A, et al. Driving the Model to Its Limit: Profile Likelihood Based Model Reduction. PLOS ONE. 2016 Sep 2; 11(9):e0162366. https://doi.org/10.1371/journal.pone.0162366 PMID: 27588423

47. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes & development. 2001 Jan; 15(1):50–65. https://doi.org/10.1101/gad.828901 PMID: 11156605

48. Hammond SL, Ham RG, Stampfer MR. Serum-free growth of human mammary epithelial cells: rapid clonal growth in defined medium and extended serial passage with pituitary extract. Proceedings of the National Academy of Sciences. 1984 Sep 1; 81(17):5435 LP– 5439. https://doi.org/10.1073/pnas.81.17.5435 PMID: 6591199

49. Grankvist N, Watrous JD, Lagerborg KA, Lyutvinskiy Y, Jain M, Nilsson R. Profiling the Metabolism of Human Cells by Deep (13)C Labeling. Cell chemical biology. 2018/09/27 ed. 2018 Nov 15; 25(11):1419–1427.e4.

50. Lyutvinskiy Y, Watrous JD, Jain M, Nilsson R. A Web Service Framework for Interactive Analysis of Metabolomics Data. Analytical Chemistry. 2017 Jun 6; 89(11):5713–8. https://doi.org/10.1021/acs.analchem.7b00890 PMID: 28514166

51. Shupletsov MS, Golubeva LI, Rubina SS, Podvyaznikov DA, Iwatani S, Mashko S V. OpenFLUX2: 13C-MFA modeling software package adjusted for the comprehensive analysis of single and parallel labeling experiments. Microbial Cell Factories. 2014 Dec 19; 13(1):152.