

Research article

Open Access

^{13}C labeling experiments at metabolic nonstationary conditions: An exploratory study

Sebastian Aljoscha Wahl*¹, Katharina Nöh² and Wolfgang Wiechert³

Address: ¹Bioprocess Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany, ²Fermentation Technology, Institute of Biotechnology 2, Research Centre Jülich GmbH, Jülich, Germany and ³Department of Simulation, Institute of Systems Engineering, University of Siegen, Siegen, Germany

Email: Sebastian Aljoscha Wahl* - awahl@mpi-magdeburg.mpg.de; Katharina Nöh - k.noeh@fz-juelich.de; Wolfgang Wiechert - wolfgang.wiechert@uni-siegen.de

* Corresponding author

Published: 18 March 2008

Received: 17 October 2007

BMC Bioinformatics 2008, 9:152 doi:10.1186/1471-2105-9-152

Accepted: 18 March 2008

This article is available from: <http://www.biomedcentral.com/1471-2105/9/152>

© 2008 Wahl et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Stimulus Response Experiments to unravel the regulatory properties of metabolic networks are becoming more and more popular. However, their ability to determine enzyme kinetic parameters has proven to be limited with the presently available data. In metabolic flux analysis, the use of ^{13}C labeled substrates together with isotopomer modeling solved the problem of underdetermined networks and increased the accuracy of flux estimations significantly.

Results: In this contribution, the idea of increasing the information content of the dynamic experiment by adding ^{13}C labeling is analyzed. For this purpose a small example network is studied by simulation and statistical methods. Different scenarios regarding available measurements are analyzed and compared to a non-labeled reference experiment. Sensitivity analysis revealed a specific influence of the kinetic parameters on the labeling measurements. Statistical methods based on parameter sensitivities and different measurement models are applied to assess the information gain of the labeled stimulus response experiment.

Conclusion: It was found that the use of a (specifically) labeled substrate will significantly increase the parameter estimation accuracy. An overall information gain of about a factor of six is observed for the example network. The information gain is achieved from the specific influence of the kinetic parameters towards the labeling measurements. This also leads to a significant decrease in correlation of the kinetic parameters compared to an experiment without ^{13}C -labeled substrate.

Background

The recent developments in Metabolomics and Fluxomics open up new perspectives for detailed *in vivo* studies of the cellular metabolism. Various experimental approaches using GC- or LC-MS enable the measurement of intracellular concentrations and isotope enrichments with increasing accuracy. Nowadays, almost all metabolites of the central carbon metabolism can be detected [1-3];

although not all can be quantified exactly. Metabolic fluxes depend on the metabolite concentrations, enzyme concentrations (regulated by transcription [4]) as well as regulatory mechanisms (like enzyme modification [5] and allosteric inhibition or activation [6]). Different information about the *in vivo* mechanisms and fluxes can be collected from already established experimental approaches.

Figure 1 depicts the most prominent approaches in Fluxomics and Metabolomics and their information targets. In general, experimental approaches can be divided into metabolic stationary and metabolic nonstationary experiments. Whereas metabolic stationary approaches aim to quantify intracellular fluxes, nonstationary experiments are used to identify reaction kinetic properties like substrate affinities and allosteric inhibition. Fluxes at different metabolic states can only be calculated from a reaction kinetic model (with identified parameters). The following approaches are currently established:

1. Metabolic Flux Analysis (left state in Figure 1): At metabolic steady state it is assumed, that the intracellular concentrations and fluxes are not changing during the time of the analysis. A model based on mass balances is used to quantify the flux through the network [7].
2. ¹³C Metabolic Flux Analysis (second left state in Figure 1): It is well known that solely stoichiometric balancing will not lead to a resolution of bidirectional or parallel

fluxes [8]. Experiments with additional ¹³C labeled substrates have been shown to increase the measurable information and to enable the quantification of exchange fluxes and parallel fluxes [9,10].

3. ¹³C Metabolic Flux Analysis at isotopically nonstationary state (third state in Figure 1). Here the organism is kept at metabolic stationary state while a ¹³C labeled substrate pulse is applied. Only recently, Nöh et al. [11] have evaluated an experiment of this type. It was shown that the duration of the labeling experiment can be drastically reduced from several hours to the order of minutes.
4. Stimulus-Response Experiments (fourth state in Figure 1). At a metabolic nonstationary state the intracellular concentrations (and consequently also the fluxes) are not constant. These conditions target the identification of *in vivo* enzyme kinetics with one single experiment. The currently most prominent approach in this field is given by Stimulus-Response Experiments [12-15]. A culture is driven to a substrate limited state and then excited by a

Metabolic state	stationary			instationary	
Isotopic state	no labeling	stationary	instationary	no labeling	instationary
Concentrations (intracellular)					
Fluxes					
Isotopomers					
Results	Net-fluxes	Net- and exchange-fluxes	Net- and exchange-fluxes	Enzyme kinetics	Enzyme kinetics

Figure 1
Category of experiments. Overview of experiments used for the estimation of intracellular fluxes and enzyme kinetics. The new metabolic and isotopic nonstationary state is analyzed here within a simulation study.

strong external stimulus, typically a substrate pulse. The metabolic response of the cells is tracked by rapid sampling on a sub-second timescale [16-18]. Although lots of data is generated from a stimulus response experiment [16,19], the approach has shown to be limited: still only a part of the parameters can be identified. Even if the data from several experiments is gathered, it is not possible to determine all kinetic parameters [20].

These experiments are now common procedure. Yet missing is the experiment that results from a fusion of approach 3 and 4, i.e. labeling the substrate in a metabolically dynamic experiment. This type of experiment is theoretically analyzed to evaluate whether this type of labeling experiment can increase the parameter estimation accuracy of a reaction kinetic model. To this end, it is first shown how the experiments can be modeled based on the known concepts from ^{13}C Flux Analysis and reaction kinetic modeling. A series of papers [21-23] are concerned with the integration of kinetic and isotopic information. In contrast to the present integration, the authors assumed most enzyme kinetic parameters to be known. Moreover the system was analyzed in a metabolic steady state and therefore resembles approaches 2 and 4 under the prerequisite of mostly known kinetic parameters. The values for these parameters are taken from literature or additional experiments [23]. The approach presented here is based on a dynamic experiment rather than one or several metabolic steady states and all enzyme kinetic parameters are to be estimated from the experimental data.

The approach is based on a complete isotopomer model that is capable of exploiting all available isotopomer measurements from MS(/MS) or NMR devices. In order to demonstrate the general mathematical framework a small example reaction network is discussed the first time. Comparing the outcomes of an experiment with labeling to those of a reference without the addition of ^{13}C substrate the information surplus from the metabolic and isotopic transient state is elucidated. To cover different experimental settings, three scenarios of available measurements are introduced and compared.

Modeling metabolic and isotopic nonstationary systems

A general model structure for metabolic and isotopic nonstationary systems can be obtained by combining models from both worlds. However, this hybrid structure also inherits some standard assumptions from kinetic and isotopic models. Precisely, all upcoming equations are only valid if the following assumptions [24] are accepted:

1. Homogeneity: All intracellular metabolites are distributed uniformly.

2. No isotopic effects: The enzymatic reactions are not influenced by the labeling state of the metabolites.

3. Fast equilibrium of enzyme complexes: Intermediate enzyme complexes rapidly reach equilibrium such that the quasi-stationary assumption of Michaelis-Menten [6] is valid.

Reaction kinetic model

With these assumptions, well known mechanistic or approximate enzymatic rate laws like Michaelis-Menten [6] can be used to describe the fluxes \mathbf{v} in dependence of substrate, product and effector concentrations, kinetic parameters α (like v_{max} , K_M) and external (experimental) parameters β (like substrate feed):

$$\mathbf{v} = \mathbf{v}(\mathbf{c}, \alpha, \beta)$$

The concentrations of metabolites (\mathbf{c}) do change depending on the in- and effluxes \mathbf{v} . The mass balances of the pools can be formulated using the well known stoichiometric matrix \mathbf{N} [25]:

$$\dot{\mathbf{c}} = \mathbf{N} \cdot \mathbf{v}(\mathbf{c}, \alpha, \beta)$$

^{13}C Balance equations

Here, the labeling state is modeled using the isotopomer concept [24] to keep the contribution understandable on a basic level. Nevertheless, more evolved but mathematically equivalent concepts like cumomers [26], bondomers [27] or elementary metabolite units [28] can also be used. A metabolite with n carbon atoms has 2^n labeling states. Thus, 2^n isotopomer fractions are required for each metabolite to describe its labeling state. The overall isotopic state of the network is described by a vector \mathbf{x} containing the isotopomer fractions of all metabolites. Isotopomer dynamics of a metabolite pool depends on the pool concentration \mathbf{c} , the input labeling \mathbf{x}^{inp} and the flux functions \mathbf{v} . The overall system for the labeling state is given by [29]:

$$\mathbb{Q}(\mathbf{c}) \cdot \dot{\mathbf{x}} = \mathbf{f}(\mathbf{v}(\mathbf{c}, \alpha, \beta), \mathbf{x}^{\text{inp}}, \mathbf{x})$$

Here, the operator $\mathbb{Q}(\mathbf{c})$ generates a diagonal matrix composed of concentrations c_i (with n_i repetitions which correspond to all isotopomer fractions, n_i being the number of isotopomers of metabolite i). The nonlinear function \mathbf{f} includes the stoichiometry of the isotopomer transitions, the vector \mathbf{x} describes the isotopomer distribution of the balanced intracellular pools (for an example take a look at Eq. (14) in the example section) and \mathbf{x}^{inp} is defined by the isotopomer fractions of the input substrate.

Notice that unlike in stationary isotopomer balances [30], the vector \mathbf{v} is not constant but changes over time in dependence of the metabolite concentrations. Combining both model equations (2) and (3), the metabolic and isotopic transients can now be completely described. Hence, the resulting model is a hybrid of a classical ^{13}C labeling and a reaction kinetic model. The functions \mathbf{f} is constructed based on matrix calculus as described earlier by Wiechert et al [24] but replacing the vector \mathbf{v} with a vector of kinetic rate terms. A clear difference appears at the point of solution; Instead of a numeric equation solver, numeric integration is applied to calculate the labeling transients.

Measurement model

The measurement model is a function that couples the dynamic isotopomer model with the available measurements. Within this section the measurement model will be described only very briefly in order to avoid rather unnecessary technical explanations (these can be found in Appendix A). Generally, at time point t_k measurements of the labeling and the concentrations are expected. Both quantities are gathered in the measurement vector $\mathbf{y}(t_k)$ that depends on the state of the system $\mathbf{c}(t_k)$ and $\mathbf{x}(t_k)$. Hence, we introduce a measurement function \mathbf{g} :

$$\mathbf{y}(t_k) = \mathbf{g}(\mathbf{x}(t_k), \mathbf{c}(t_k)), k = 1, \dots, N$$

Collecting all sampling time points into a vector $\xi = (t_1, \dots, t_N)^T$, Eq. (4) compactly reads:

$$\underbrace{\begin{pmatrix} \mathbf{y}(t_1) \\ \mathbf{y}(t_2) \\ \vdots \\ \mathbf{y}(t_N) \end{pmatrix}}_{\mathbf{y}(\xi)} = \underbrace{\begin{pmatrix} \mathbf{g}(\mathbf{x}(t_1), \mathbf{c}(t_1)) \\ \mathbf{g}(\mathbf{x}(t_2), \mathbf{c}(t_2)) \\ \vdots \\ \mathbf{g}(\mathbf{x}(t_N), \mathbf{c}(t_N)) \end{pmatrix}}_{\mathbf{g}(\mathbf{x}(\xi), \mathbf{c}(\xi))}$$

A concrete function \mathbf{g} is presented within the example given in the appendix for massisotopomer measurements.

Parameter sensitivity

The final goal of the nonstationary experiment is the determination of the kinetic parameters α . All other quantities (fluxes \mathbf{v} , concentrations \mathbf{c}) follow from these parameters. Sensitivities quantify the influence of α on the model response, namely the concentration time course $\mathbf{c}(t)$ and the isotopomer distributions $\mathbf{x}(t)$. Implicit derivation of the differential equations (2) and (3) with respect to the parameters α leads to the sensitivity differential equations associated with the model (2)–(3):

$$\begin{aligned} \frac{d\mathbf{c}}{d\alpha} &= \mathbf{N} \cdot \left[\frac{d\mathbf{v}}{d\alpha} + \frac{d\mathbf{v}}{d\mathbf{c}} \cdot \frac{d\mathbf{c}}{d\alpha} \right] \\ \square(\mathbf{c}) \cdot \frac{d\mathbf{x}}{d\alpha} &= \frac{d\mathbf{f}}{d\mathbf{v}} \cdot \frac{d\mathbf{v}}{d\alpha} + \frac{d\mathbf{f}}{d\mathbf{x}} \cdot \frac{d\mathbf{x}}{d\alpha} - \frac{d\mathbf{N}}{d\alpha} \cdot \frac{d\mathbf{c}}{d\alpha} \cdot \dot{\mathbf{x}} \end{aligned}$$

Besides the influence of the parameters α on the system state, the influence on the expected measurements is now easily calculated using the chain rule:

$$\frac{d\mathbf{y}}{d\alpha} = \frac{d\mathbf{g}}{d\mathbf{x}} \cdot \frac{d\mathbf{x}}{d\alpha} + \frac{d\mathbf{g}}{d\mathbf{c}} \cdot \frac{d\mathbf{c}}{d\alpha}$$

As shown in the following sections, the sensitivities are essential for the statistical evaluation and comparison of different experimental configurations [10].

Nonlinear regression model

Regression analysis is the chosen method to estimate the kinetic parameters from the experimental data. Assuming that the model adequately describes reality and that experimental measurements \mathbf{y} have some unknown error ϵ , Eq. (5) extends to:

$$\eta(\xi) = \mathbf{g}(\mathbf{x}(\xi), \mathbf{c}(\xi)) + \epsilon(\xi)$$

Parameter estimation is performed by using weighted least square minimization. Expecting that the errors $\epsilon(\xi)$ are normally distributed with expectation $E[\epsilon(\xi)] = \mathbf{0}$ and the given measurement covariances are collected in the measurement covariance matrix Σ , the weighted least squares functional for the calculation of the optimal parameters $\hat{\alpha}$ reads [31]:

$$\hat{\alpha} = \arg \min_{\alpha} (\eta - \mathbf{g})^T \Sigma^{-1} (\eta - \mathbf{g})$$

Various algorithms can be applied to solve this minimization problem like derivation-free simplex methods [32], evolutionary strategies [33], or gradient-based methods [34].

Statistical evaluation

In order to obtain a statement about whether or not the information contained in the measurements is sufficient to identify the parameter values within a certain precision, it is essential to analyze the accuracy of the parameter estimation $\hat{\alpha}$ found by solving Eq. (8). Assuming that the estimate $\hat{\alpha}$ is close to the real solution α^* application of linearized regression analysis is feasible. In experimental design studies usually a parameter set $\hat{\alpha} = \alpha^*$ is chosen. It

is known, that linear statistics can produce imprecise estimates of parameter confidence regions [35]. Nevertheless, this tool still fulfils its purpose when it is used to compare different experiments. Here the relative improvement of the confidence regions is much more important than their precise knowledge. Moreover, the concept of correlation coefficients that characterizes the confidence regions is an inherently linear concept. It will turn out that parameter correlations are relevant because with ^{13}C labeled substrates parameter correlations strongly decrease that lead to smaller confidence regions.

The covariance of the parameters $\hat{\alpha}$ is calculated using the derivative and the covariance of the measurements Σ by [36]:

$$\text{cov}^{-1}(\alpha) = \left(\frac{dy}{d\alpha} \right)^T \Sigma^{-1} \left(\frac{dy}{d\alpha} \right)$$

In the appendix section it is explained how the different types of measurements and additional knowledge are integrated.

Statistical measures to compare experiments

The parameter covariance matrix $\hat{\alpha}$ contains all information needed to judge the information gain of an experiment, namely parameter standard deviations, correlations and confidence regions. A drawback of using the covariance matrix is its size. Even for small models it is rather difficult to evaluate and compare experiments on its basis. Therefore statistical measures are used that reduce the complexity and facilitate an automated comparison of experiments. Usually, a reference experiment is specified to which all other experiments are compared.

A commonly used measure is the D -criterion, which is calculated from the determinant of the covariance matrix. It comprehensively summarizes the available parameter information. To compare different experiments Möllney et al. [10] propose to use a root of the D -criterion. The scaled ratio I_D between an experiment and the reference experiment (index *ref*) is defined by:

$$I_D = \sqrt[2 \dim(\alpha)]{\frac{D_{\text{ref}}}{D}}, \quad D = \det(\text{cov}(\alpha))$$

Roughly speaking the I_D value reflects an overall information gain based on the volume of the covariance ellipsoid (for details see [10]). The standard deviations and correlations of the judged experiment are – on an average – I_D times smaller than those from the reference experiment.

However, this does not guarantee that every single standard deviation is improved.

Besides the D -criterion, the A -criterion is frequently applied. The latter criterion reflects the mean of the diagonal elements of the covariance matrix. Again, a scaled scalar is introduced to compare two experiments:

$$I_A = \sqrt{\frac{A_{\text{ref}}}{A}} \quad A(\hat{\alpha}) = \frac{1}{N} \text{trace}(\text{cov}(\hat{\alpha}))$$

In contrast to the D -criterion, correlations between parameters are not taken into account; the A -criterion is based on the parameter variances only. Therefore, the square root has to be taken.

An exploratory study

Within this section the described framework is applied to an example network. As an example, some equations are explicitly shown to facilitate the understanding of the presented concepts. Some different conceivable scenarios regarding the available measurements are also introduced to cover different measurement setups and analyze the influence of, for example, a missing concentration measurement.

Example network

The network chosen for this study is shown in Figure 2. Although it is simple, it nonetheless reflects all typical reaction mechanisms of the central carbon metabolism. A linear reaction sequence (v_{upt} , v_1 , v_2) with a feedback inhibited entry reaction (v_{upt}) as well as a cyclic reaction sequence (v_3 , v_4) with some fillup reaction (v_5) and an exit (v_6) is included. Some reversible reaction steps (v_1 , v_5) are incorporated as well as a bimolecular reaction step (v_2) with Hill-type kinetic mechanism and irreversible split reactions (v_3 , v_4) are present (see Table 1 for a complete listing of reaction mechanisms and the kinetic expressions).

The assumed metabolic regulation reflects some basic principles like product inhibition (reaction v_{upt} by A , v_2 by C) and allosteric inhibition (v_4 by C). Short chain molecules were chosen to keep the dimension of the isotopomer equations low. The substrate of the system contains only 2 carbon atoms (A_{feed}) and the molecule formed in the bimolecular reaction entering the cycle has 4 carbon atoms.

Isotopomer balance equations and simulation

Some assumptions about the experimental setup and the network will be needed to perform the simulation study of the nonstationary state: (1) the external influences β , (2) assumptions about the reaction mechanisms that will

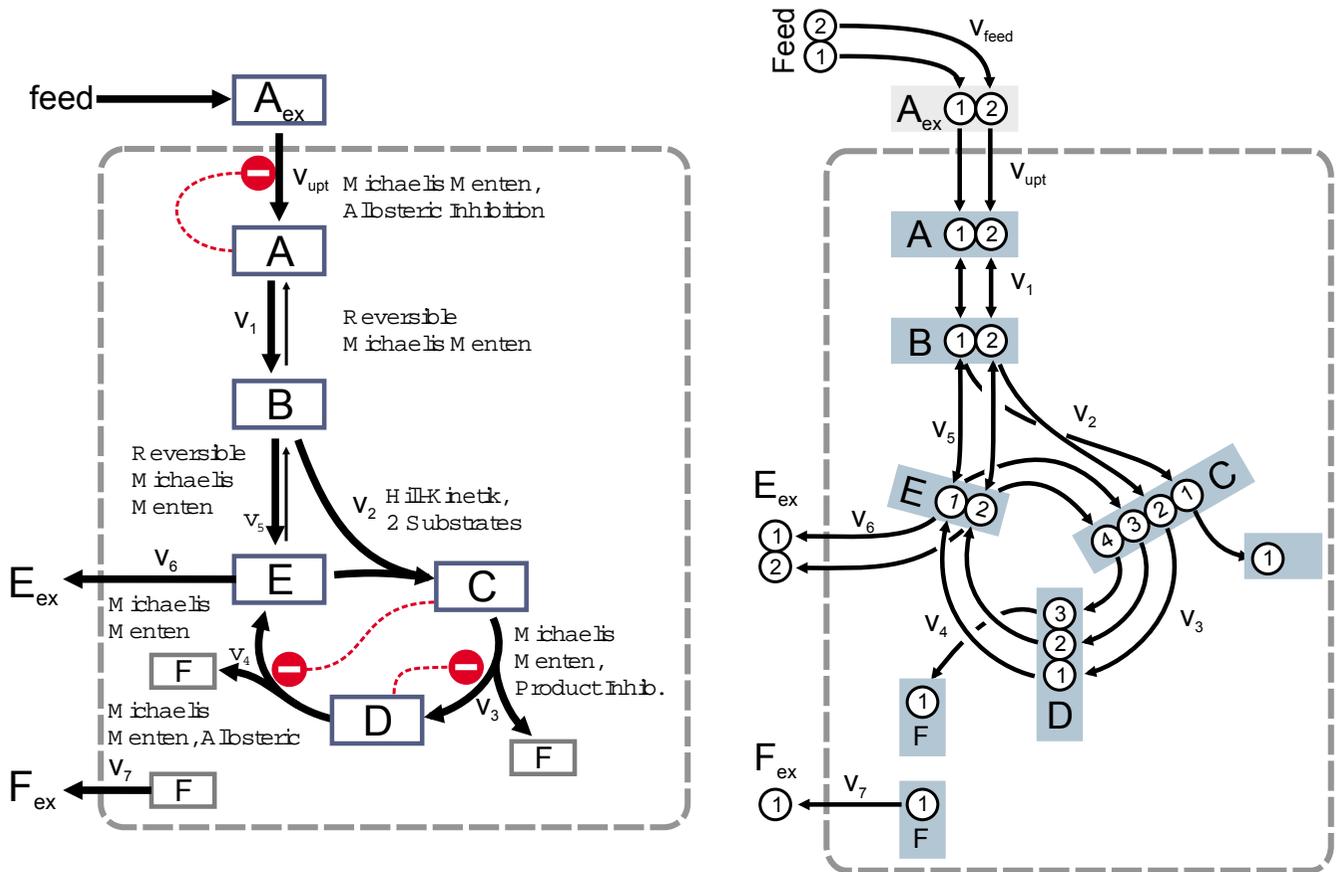


Figure 2
Example network structure including atom transitions. The reaction kinetics are given in Table 1.

form the functions v and (3) estimates for the unknown kinetic parameters α .

The balance equations for the metabolite pools are not shown here because they are well documented in literature [7]; five pools are balanced for the example network. In total, 24 kinetic parameters are needed for the reaction kinetics. In order to model the reactor feeding profile, one input parameter (\dot{v}_{feed}) is introduced. The values used within this study are listed in Table 2. Having the balances in hand, the time course of the concentrations $c(t)$ is calculated by integrating Eq. (2). Figure 3 shows the simulation result. Concentrations of the metabolite pools A and B show a short overshoot, while concentrations C, D and E show a monotonic increase after the pulse. Pools A and B rapidly reach a steady state after about 10 s, while the dynamics in C, D and E do not reach a new steady state in the given simulation time of 20 s.

The atom transitions have to be known to balance the isotopomer fractions (Eq. (3)). These are shown in Figure 2 and are also listed in Table 3. Forty isotopomer fractions are balanced in the example. Four input variables are needed to describe the labeling of A_{ex} (here set to $\mathbf{x}^{inp} = (0.01, 0.01, 0.01, 0.97)^T$). Exemplary, the balance equations for the isotopomer fractions $\mathbf{a} = (a_{00}, a_{01}, a_{10}, a_{11})^T$ of pool A are shown.

$$\frac{dA}{dt} = v_{upt}^{\rightarrow} - v_1^{\rightarrow} + v_1^{\leftarrow}$$

$$A(t) \frac{da_{00}}{dt} = a_{ex00} v_{upt}^{\rightarrow} - a_{00} v_1^{\rightarrow} + b_{00} v_1^{\leftarrow}$$

$$A(t) \frac{da_{01}}{dt} = a_{ex01} v_{upt}^{\rightarrow} - a_{01} v_1^{\rightarrow} + b_{01} v_1^{\leftarrow}$$

$$A(t) \frac{da_{10}}{dt} = a_{ex10} v_{upt}^{\rightarrow} - a_{10} v_1^{\rightarrow} + b_{10} v_1^{\leftarrow}$$

$$A(t) \frac{da_{11}}{dt} = a_{ex11} v_{upt}^{\rightarrow} - a_{11} v_1^{\rightarrow} + b_{11} v_1^{\leftarrow}$$

Table 1: Kinetic rate laws used in the small example network.

Reaction	Mechanism	Inhibitor	Kinetic
upt	Michaelis-Menten, I Inhibitor	A	$\frac{v_{\max}A}{K_{mA}\left(1+\frac{A}{K_I}\right)+A_{ex}\left(\frac{A}{K_I}\right)}$
v1	reversible Michaelis-Menten		$\frac{v_{\max}\left(A-\frac{B}{K_{eq}}\right)}{K_{mA}\left(1+\frac{B}{K_{mP}}\right)+A}$
v2	2 Substrate Hill-Kinetic		$\frac{v_{\max}B^{hA}E^{hB}}{\left(K_{mA}+B^{hA}\right)\left(K_{mB}+E^{hB}\right)}$
v3	Michaelis-Menten, I competitive Inhibitor	D	$\frac{v_{\max}C}{K_{mA}\left(1+\frac{D}{K_I}\right)+C}$
v4	Michaelis-Menten, I competitive Inhibitor	C	$\frac{v_{\max}D}{K_{mA}\left(1+\frac{C}{K_I}\right)+D}$
v5	reversible Michaelis-Menten		$\frac{v_{\max}\left(A-\frac{E}{K_{eq}}\right)}{K_{mA}\left(1+\frac{E}{K_{mP}}\right)+B}$
v6	Michaelis-Menten		$\frac{v_{\max}E}{K_{mA}+E}$

Again it must be noted that the fluxes are functions of kinetic parameters (see Table 1). For example, the flux v_1 ($A \rightarrow B$) is a reversible flux. Its rate is given by:

$$v_1 = \frac{v_{\max}\left(A-\frac{B}{K_{eq}}\right)}{K_{mA}\left(1+\frac{B}{K_{mP}}\right)+A}$$

With known initial isotopomer fractions x_0 at time point $t = 0$ (usually determined by the natural isotope enrichments), the integration of Eq. (3) yields the time course $x(t)$ of the isotopomer enrichments. In Figure 4 it can be seen that pools A and B are rapidly enriched with labeled carbon. Only at 2 s respectively at 5 s after the pulse they do show the labeling of the input substrate A_{ex} . Pools C, D and E exhibit a different behavior. Especially in C and D an interim labeling state can be observed, that results from the reaction of unlabeled E#00 with fully labeled B#11. This state is still seen in D while in E it is not as pronounced because E is fed from B and D. This already hints at additional information contents, because a more detailed behavior in comparison to the one shown Figure 3 is observed. In the following sections, it will be seen that

the kinetic parameters do influence the isotopomer distribution differently to the concentration time course.

Scenarios

Since there is still much development in the analytical methods, scenarios with different measurements are investigated. In order to estimate the influence of increased or decreased availability of measurement information the following cases will be compared (summarized in Figure 5):

- Scenario S_{all} : This is an optimistic scenario where all metabolite concentrations can be quantified and all mass isotopomer fractions are available. Here, the mass isotopomers of the entire carbon skeleton are measured without any fragmentation.
- Scenario S_{D-rel} : Here, less information is available compared to S_{all} ; this scenario is derived from findings that for some metabolites the concentration could not be determined [11]. Nevertheless, it was possible to measure the mass isotopomer ratios. As an example, the concentration of D cannot be determined, but its mass isotopomer ratios are measured.

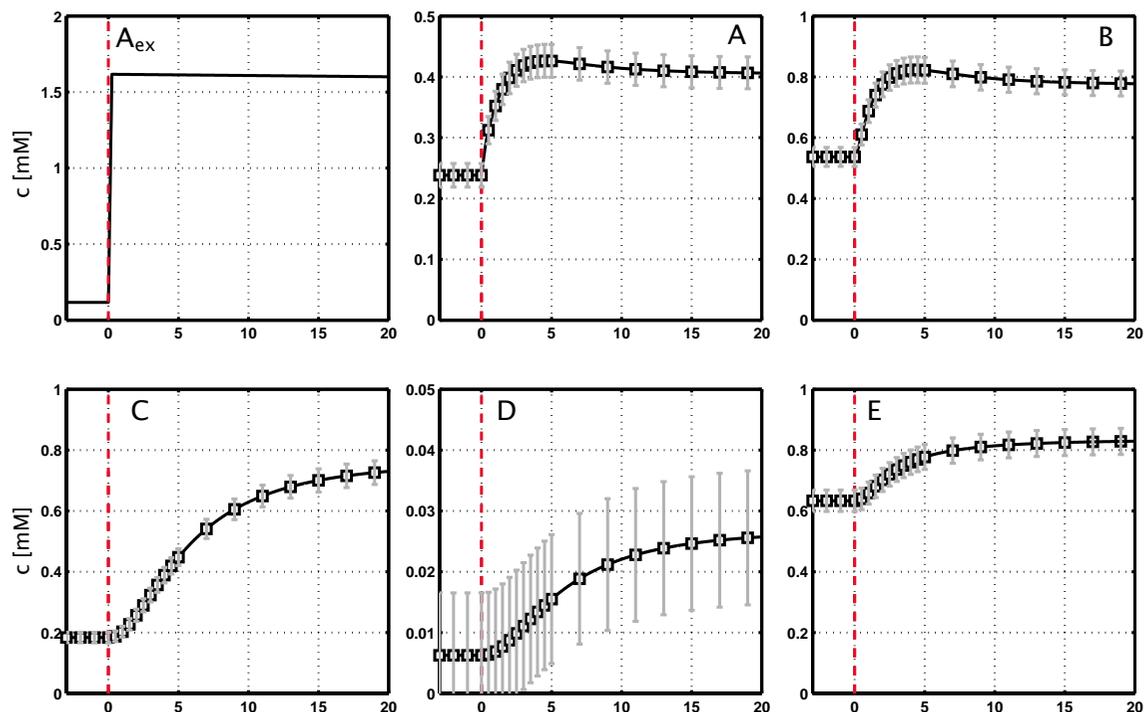


Figure 3

Simulation of the example network and the assumed measurement standard deviations. The extracellular concentration (A_{ex}) rises from 0.05 mM to 1.5 mM by the pulse. The intracellular concentrations A and B show a short overshoot, C, D and E increase continuous and reach the steady state after about 25 s. The squares show the noise free measurements generated from the simulation. The standard deviations include the errors from sample processing and the MS measurements.

- Scenario $S_{C,D-frag}$: Compared to scenario S_{all} additional mass measurements of fragments are available. As an example, the labeling distribution of the pools C and D are measured in more detail because of a specific fragmentation in MS/MS measurement mode (see Figure 5). Metabolite D fragments into a C_2 and a C_1 skeleton. Here, the mass isotopomers of the C_2 body can be measured additionally. Instead of four mass isotopomers, in total six tandem mass isotopomer measurements are generated [37]. Metabolite C fragments into a C_3 and a C_1 body. Consequently, instead of five mass isotopomer ratios, eight tandem mass isotopomers are measured.

To summarize, scenario S_{all} is optimistic, $S_{C,D-frag}$ is even more optimistic, whereas in S_{D-rel} some information is missing.

Sampling and quality of the measurements

Besides the measurements, the sampling frequency is also crucial for parameter identification [29]. For all scenarios, the following sampling setup is used:

- Three samples are taken every 1 s before the pulse (-3...0 s),
- In the following 5 s samples are taken every 0.5 s (0.5...5 s),
- After 7 s samples are taken every 2 s (7...19 s)

The quality of the measurements depends on the accuracy of the measurement device and possible errors during sample preparation. Figure 6 depicts the sampling steps usually needed to determine intracellular concentrations. Essentially, the sampling steps lead to a dilution ϕ . The signal intensity depends on the concentration in the sample and the response factor γ that is determined from

Table 2: Kinetic parameters for the simulation study.

Reaction	Parameter	Value
feed	c	0.01
upt	v_{max}	1
	K_I	2
	K_{mA}	0.001
v1	v_{max}	3
	K_{eq}	3
	K_{mA}	0.1
	K_{mP}	3
v2	v_{max}	2.5
	K_{mA}	0.25
	hA	2
	K_{mB}	2
	hB	3
v3	v_{max}	2
	K_{mA}	2
	K_I	0.05
v4	v_{max}	3
	K_{mA}	0.1
	K_I	1
v5	v_{max}	2
	K_{eq}	4
	K_{mA}	1
	K_{mP}	1
v6	v_{max}	2
	K_{mA}	3

standard additions [38]. More details about the measurement model can be found in the appendix section. All sampling steps and possible errors have to be taken into account for the statistical analysis. Based on the current knowledge for the standard deviations the following factors are considered:

Table 3: Atom transitions for the small example network

Flux	Substrates	Products
v_{feed}	Feed	A_{ex}
	#ab	#ab
v_{upt}	A_{ex}	A
	#ab	#ab
v_1	A	B
	#ab	#ab
v_2	B + E	C
	#ab + #cd	#abcd
v_3	C	D + F
	#abcd	#bcd + #a
v_4	D	E + F
	#ab	#ab + #c
v_5	B	E
	#abc	#ab
$v_{6,ex}$	E	E_{ex}
	#ab	#ab
$v_{7,ex}$	F	F_{ex}
	#a	#a

The notation of Wiechert and de Graaf [24] is used.

- The response factors γ (set to 1 if determined) have a standard deviation σ_γ of 1%.

- The dilution factors ϕ (ξ) (also set to 1) have a standard deviation of 3%.

- The MS(/MS) measurement (area) itself is composed of a relative error which is assumed in the order of 1% and an absolute error that was set to 10 nM. The latter one appears rather high, but as the dilution usually will be around a factor 20 it is within a realistic range:

$$\sigma_{MS}(t_k) = 10 \text{ nM} + 0.01 \gamma \gamma_{MS}$$

Results

Statistical analysis

The example system consists of 40 isotopomer and six concentration states with 24 parameters. Hence, the parameter sensitivity matrix (Eq. (6)) contains $46 \cdot 24 = 1104$ entries for each of the 19 sampling time points. Although interesting, a detailed analysis of these large matrices is extraordinarily time-consuming. In fact, the influence on the measurable signals is more relevant. As it can be seen from equations, and the measurement equations are linear. Thus, the calculation of the output sensitivity is straight forward (Eq. (7)).

Figure 7 shows the course of the sensitivities with respect to kinetic parameters of the uptake reaction v_{upt} (K_I) and reaction v_1 (K_{mA} , v_{max} and K_{eq}). The sensitivities have to be compared on a comparable scale. Recall that the mass isotopomers are given as fractions of the pool. Multiplying the fractions with the pool concentration, the mass isotopomer concentrations are immediately obtained by $c \cdot x$.

These quantities can be directly compared to metabolite concentrations. Their parameter sensitivities can be simply obtained from Eq. (7) by using the chain rule. The sensitivity matrix $dy/d\alpha$ includes the sensitivities of the concentrations and of the mass isotopomers with respect to kinetic parameters. These sensitivities are plotted in Figure 7. It is observed that:

- After the pulse there is a strong increase in sensitivity of mass isotopomer concentrations of fully labeled pools (A+2, B+2, E+2, with a less extend C+4, D+3)

- The sensitivity of unlabeled concentrations (A+0, B+0, C+0, D+0, E+0) reaches zero, with some intermediate peaks.

- Some parameters show maxima in sensitivity, e.g. $d(B \cdot b_{+0})/dK_{mA}$, $d(C \cdot c_{+2})/dK_{mA}$, $d(E \cdot e_{+2})/dK_{mA}$.

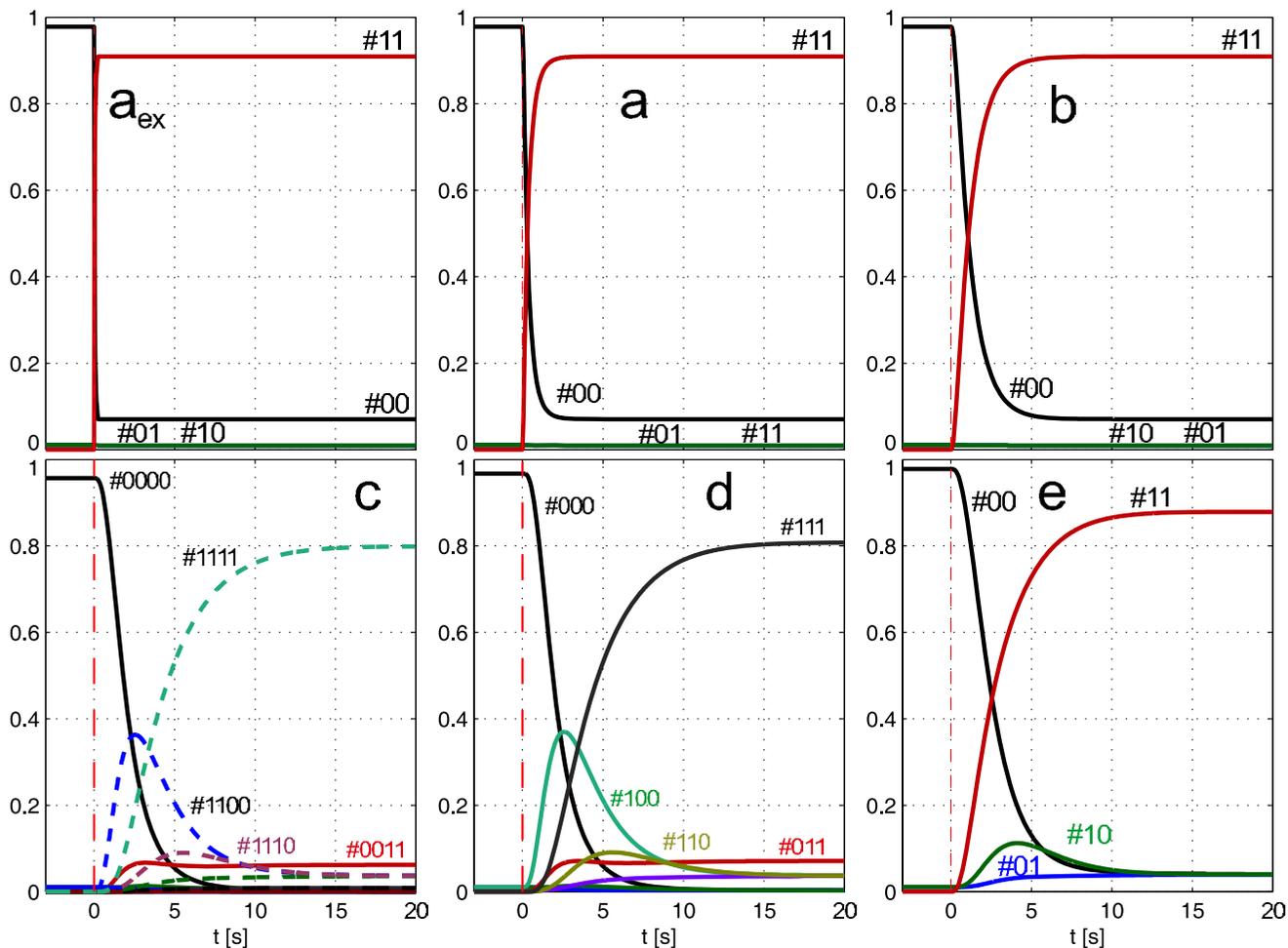


Figure 4
Simulation of the isotopomer enrichments. The extracellular substrate A_{ex} immediately switches from unlabeled to labeled. The enrichment of A and B is fast (steady state after <10 s), pool C shows intermediate labeling patterns before reaching the fully labeled steady state. These intermediate patterns can also be seen in D and E.

- Some parameters first show positive, then negative influence on the mass isotopomers (and vice versa), e.g. $d(B \cdot b_{+0})/dK_{m_A}$, $d(B \cdot b_{+0})/dK_{e_q}$, $d(B \cdot b_{+0})/dV_{max}$.
- Most parameter sensitivities reach a new steady state within the given simulation interval of 20 s, e.g. $d(A \cdot a_{+2})/dK_{m_A}$, $d(A \cdot a_{+2})/dK_{e_q}$, $d(A \cdot a_{+2})/dV_{max}$.

These findings clearly show that the kinetic parameters do not only influence the total metabolite concentrations, but also influence the labeling dynamics. Within the statistical analysis section it will be seen that this will help to achieve an identification of the enzyme parameters.

Parameter standard deviations

Figure 8 shows a cumulative plot of the estimated relative standard deviations of all parameters. Obviously, the majority of the parameters can hardly be determined from the (unlabeled) reference experiment. The best estimate is found to have about 46% relative error. The second step is seen at about 61%, which is the second best estimate. Three parameters are found within a standard deviation of about 137%. These high deviations are observed because the scenario includes 3% sample dilution error and additional noise from the MS device. Especially for low concentrated metabolites (like E, see Figure 3), high standard deviations occur due to the ground noise (10 nM) of the MS measurements. It has to be taken into account too,

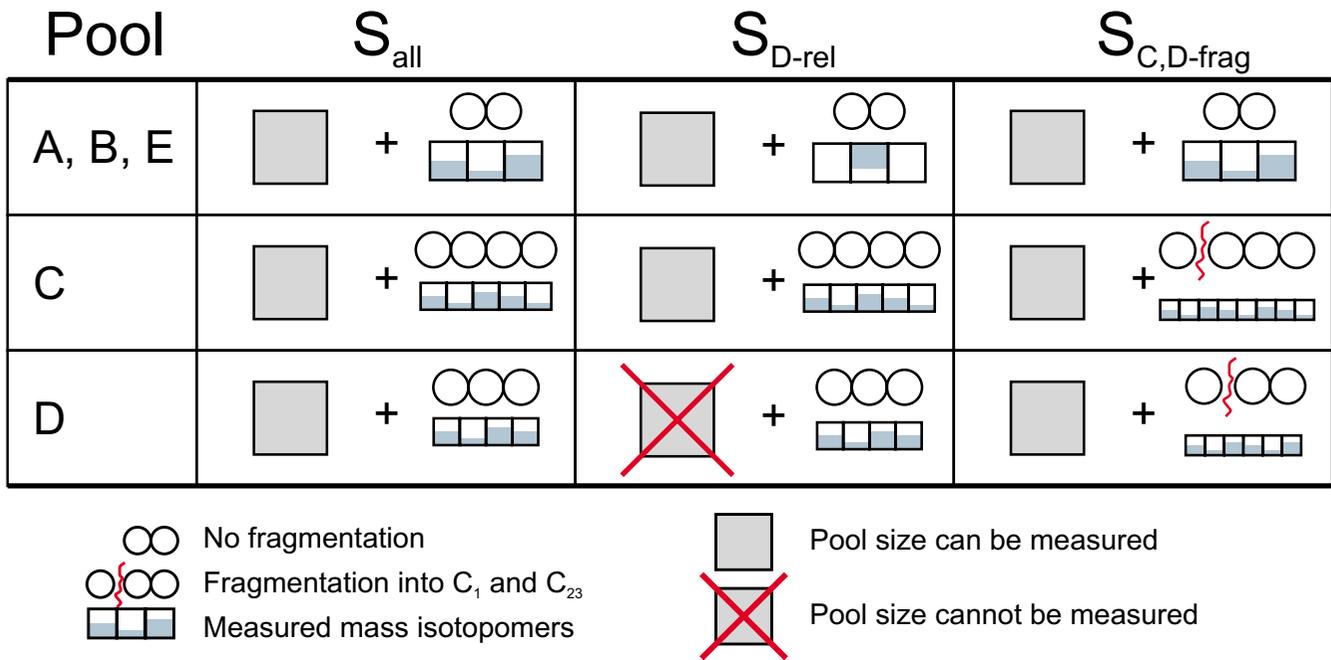


Figure 5
Summary of the compared measurement scenarios. In S_{all} it is assumed, that all concentrations and mass isotopomers can be measured. Scenario S_{D-rel} is less informative, the concentration of metabolite D cannot be quantified. $S_{C,D-frag}$ assumes MS/MS measurements with fragmentation for the metabolites C and D.

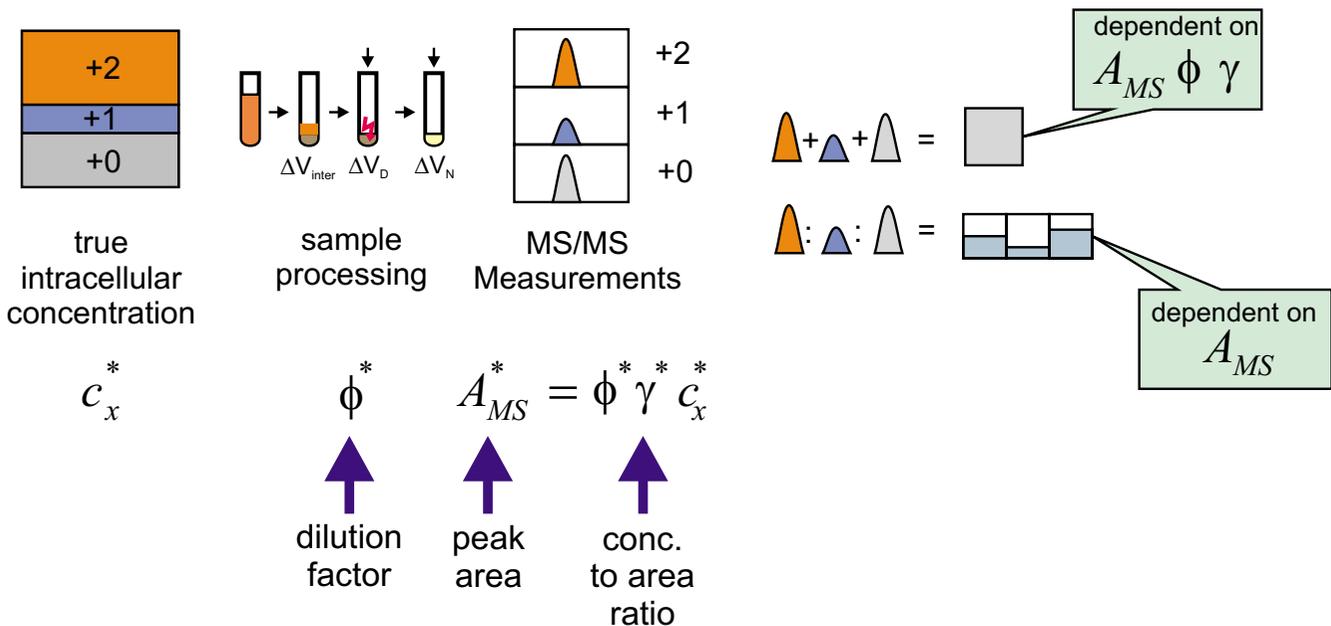


Figure 6
Scheme of the sample processing steps. The single sampling steps and their effects on concentration and labeling measurements.

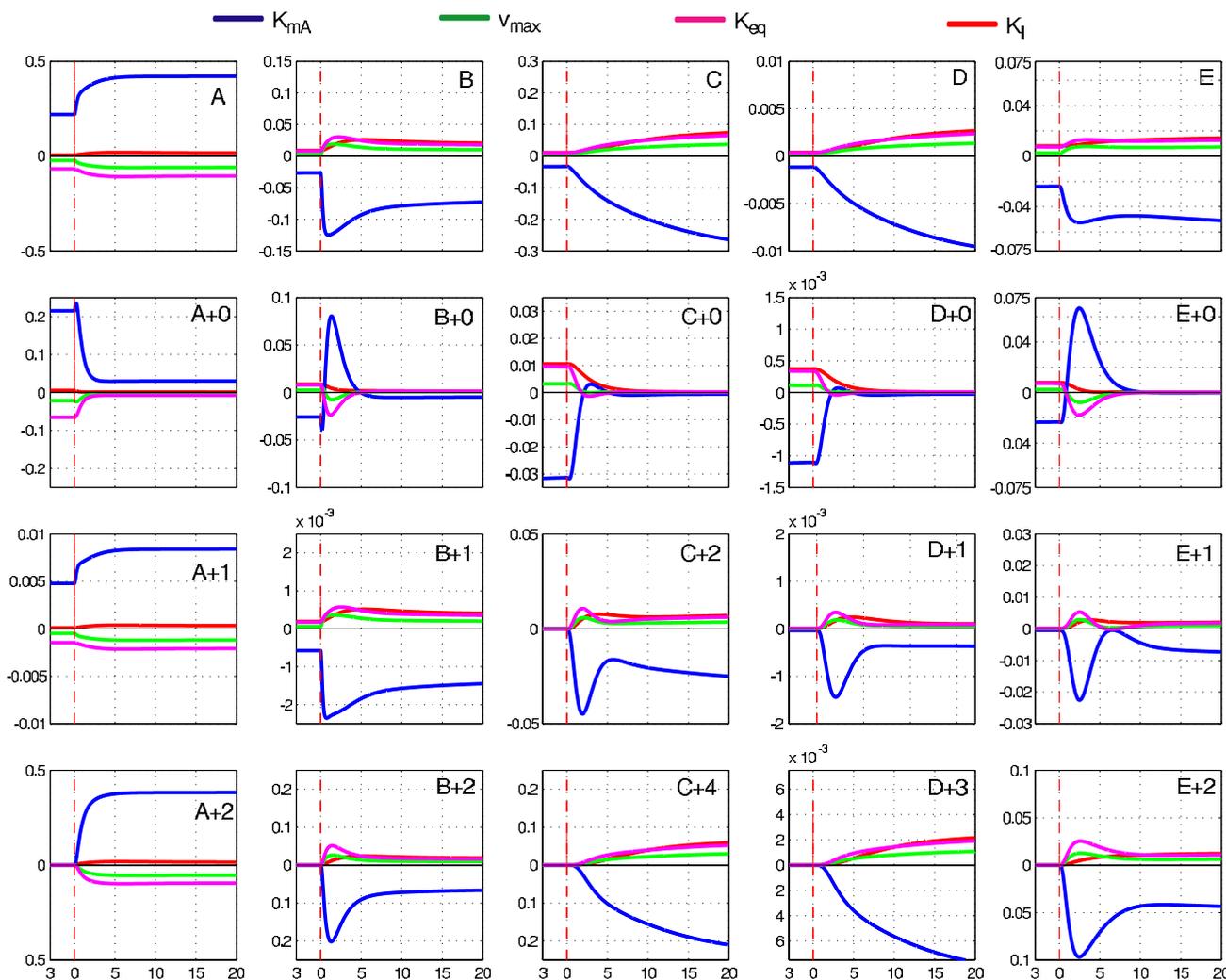


Figure 7
Parameter sensitivity plots. First row: Sensitivities of the concentrations to the parameters of reaction v_1 (K_{mA} , v_{max} and K_{eq}) and v_{upt} (K_I). Second to last row: Sensitivities of selected mass isotopomer concentrations to the same parameters.

that there are more than the 24 kinetic parameters. Another $19 \times 5 + 5 = 100$ parameters (five ϕ 's for each measured metabolite and each sampling time point

and five γ 's) are introduced for the measurement model. Although these parameters are measured directly from the sampling volume, added reagents and standard additions, these additional parameters do consume some information. It is seen from the statistical analysis that these parameters are also influenced and their estimation accuracy gets higher compared to the experimental errors.

A pronounced increase in estimation accuracy of the kinetic parameters is seen when comparing scenario S_{all} with the unlabeled reference (Figure 8). The line is shifted about a factor of 10 to the left. The increase is even more pronounced for the most accurate estimation (0.92% that is about a factor of 50). With up to 70% standard deviation, 10 parameters are found for the labeled experiment compared to only two for the reference.

When comparing the three labeled scenarios only little differences can be found. With some additional labeling information (scenario $S_{C,D-frag}$) the estimates become slightly more accurate (red, dash-dotted line). A missing concentration (scenario S_{D-rel}) can obviously be compensated by the labeling information. The dashed line is

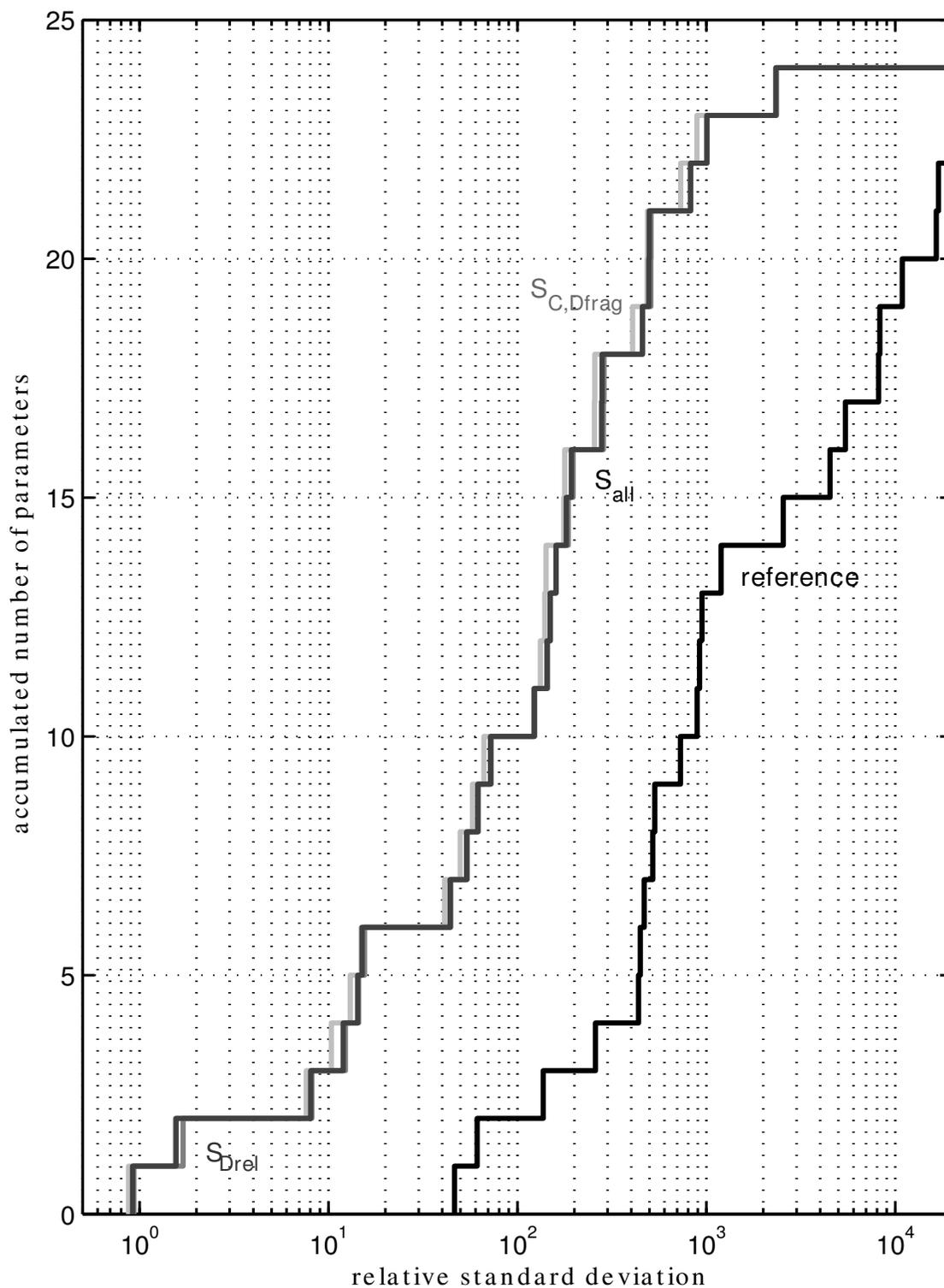


Figure 8
Cumulative distribution of the estimated standard deviations. In the unlabeled reference experiment the most accurate parameter is determined with a relative standard deviation of about 46% (first step of the black line). Only two parameters can be determined with less than 70% standard deviation. In contrast, using a labeled substrate, 10 parameters (about 40% of the parameters) can be determined with a standard deviation of less than 70%.

shifted to the right slightly. The most accurate estimate has a standard deviation of 0.93%.

These findings are also reflected in the *D*- and *A*-criterion (Table 4). The *A*-criterion is based on the variances and reflects an information increase that is comparable to observations from Figure 8. The increase in information is 11.01 (scenario S_{all}). The information gain from scenario $S_{C,D-frag}$ is about 1% higher; for S_{D-rel} about 1% lower.

The information gain calculated from the *D*-criterion is lower. Here a value of 6.91 is calculated for scenario S_{all} . With some more available labeling measurements (scenario $S_{C,D-frag}$) a factor of 7.31 is found. The information gain for scenario S_{D-rel} decreases only marginally compared to S_{all} (6.80).

Parameter correlations

So far, the improvement of the absolute variances of all parameters was investigated. In order to characterize the influence of the additional information from ¹³C labeling on the relationship among different parameters, the correlation matrix (Figure 9) will now be analyzed. The correlations for the reference experiment and the labeled scenario S_{all} are shown. The parameters are grouped with respect to the reactions and a block structure becomes visible.

At first glance a strong decrease in correlation can be observed on the areas outside of the main diagonal blocks. There are blocks with only little changes close to the main diagonal. These blocks are formed from groups of parameters that belong to one single reaction step. For example, the kinetic parameters v_{max} and K_{mA} of the reaction v_1 are strongly positive correlated (blue). I.e. if one is estimated too high, the other will also be estimated too high. Also in the labeled scenario, these parameters cannot be determined separately. It is well known from

enzyme kinetic studies that v_{max} and K_{mA} values can hardly be independently determined, if the experimental concentration range is not chosen correctly.

In the following it will be seen that the effect of improved parameter distinction between different reaction steps is based on indirect flux measurements. The parameters K_{mA} (of reaction v_1) and K_I (of v_{upt}) were chosen here as an example, because a strong decrease in correlation is observed. In Figure 7 the sensitivities of these parameters are plotted for the reference and the scenario S_{all} . Clearly, parameters with distinct influence on measurements will be less correlated. Within the reference the parameters exhibit a rather comparable influence on all measurements (though the magnitude is distinct). As a consequence, positive correlation is observed.

If labeling is introduced, the picture gets different. The influences on the measured mass isotopomers show some behavior that leads to a decoupling of parameters especially within the first 5 s after the pulse. The intermediate labeling patterns, particularly E+2, B+2 as well as E+0 and B+0 exhibit a distinct behavior for parameters K_{mA} and K_I . Whereas K_{mA} exhibits a strong influence on the time course of these mass isotopomers, the influence of K_I is relatively small. Additionally, the sensitivities of K_{mA} show a strong time dependency that increases the independent measurement of this parameter.

Conclusion

The exploratory study strongly suggests that performing a stimulus response experiment with ¹³C labeled substrate and measuring the intracellular labeling states can significantly increase the accuracy of kinetic parameter estimation. A decrease in the variance of more than a factor of 10 was observed for the analyzed example network (Figure 8). With labeling, three parameters are determined with less than 10% relative standard deviation. More than 40%

of the parameters (i.e. 10 of 24) can be determined with less than 70% standard deviation (Figure 8).

Additionally it was found that various parameter correlations are drastically reduced (Figure 9). When comparing the experiment with labeled substrate to its reference counterpart a seven fold information gain (based on the *D*-criterion) was observed. Interestingly, the information loss due to the unknown concentration of pool D (scenario S_{all}) can be compensated by the measured mass isotopomer distributions. This scenario shows only slightly less accurate determined parameters (Figure 8 and Table 4)

It can clearly be seen from the sensitivities that the parameters influence the labeling time course more specifically

Table 4: Comparison of the D- and A-criteria.

Substrate	Criterion	Scenario		
		S_{all}	S_{D-rel}	$S_{C,D-frag}$
¹² C, ξ	D	2.20 10 ⁸		
	I_D	1		
U- ¹³ C ₂ , ξ	D	5.05 10 ⁻²⁶	2.57 10 ⁻²⁵	3.58 10 ⁻²⁷
	I_D	5.02	4.85	5.31
¹² C, ξ	A	1.06 10 ⁶		
	I_A	1		
U- ¹³ C ₂ , ξ	A	1.02 10 ³	1.07 10 ³	0.91 10 ³
	I_A	10.17	9.95	10.77

List of D- and A-criteria for the different measurement scenarios of an experiment without the addition of labeling (¹²C) and an experiment with fully labeled substrate (U-¹³C₂). The calculation of D, A and the information index is given in Eq. (11) and (12).

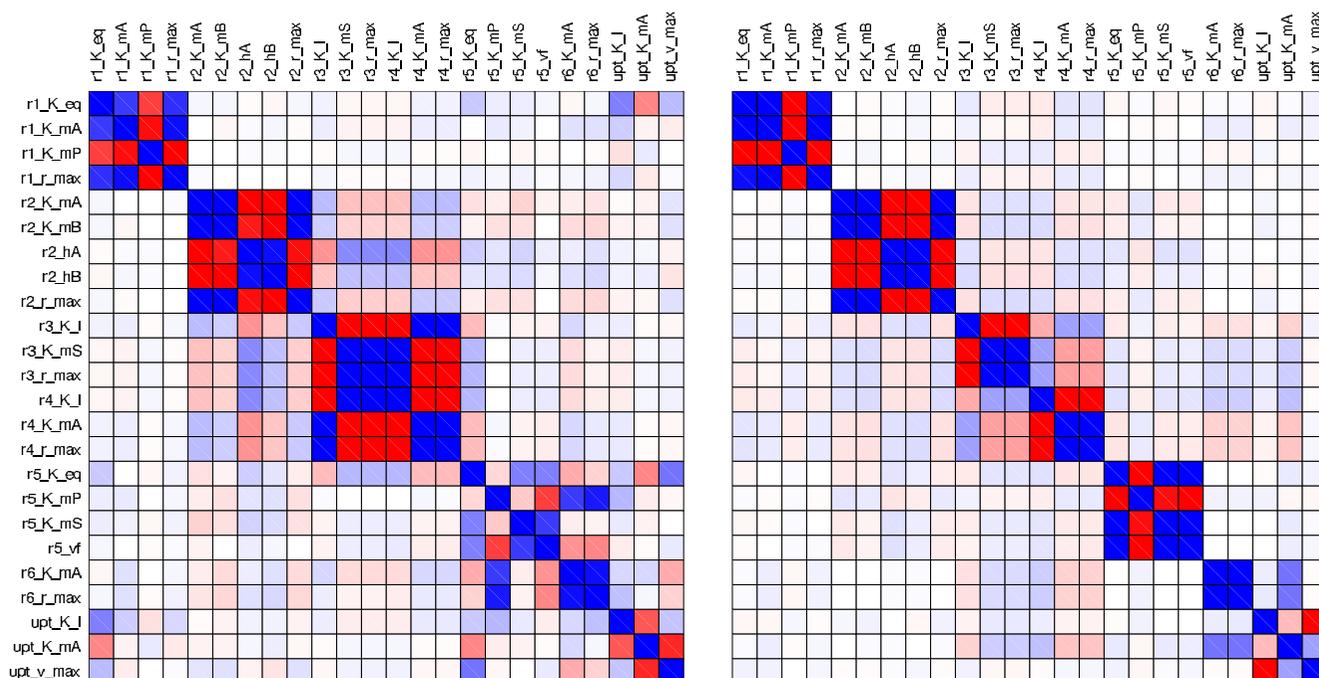


Figure 9
Color visualization of the correlation matrix. The parameter correlation matrix for the reference experiment (left) and the labeled experiment S_{all} (right). Deep blue represents a correlation of 1, deep red represents -1.

than the concentration time course. When only measuring the concentrations, a concentration increase can be explained by an increased influx or a decreased efflux. The addition of isotopic transients to the concentration courses serves as additional information about the pool exchange. A higher influx will lead to a faster labeling enrichment, while a decreased efflux will slow down the labeling enrichment. Thus, the labeling transients include information about the pool exchange that delivers valuable constraints for parameter identification.

In the case of isotopically nonstationary experiments under metabolic steady-state conditions, a similar example network was analyzed to estimate the information gain compared to a classical ^{13}C labeling experiment at isotopically steady-state [39]. The conclusions from the simulation study have shown to be in accordance with findings from a complex *E. coli* network [11]. Therefore, the results obtained from the performed exploratory study should be transferable to real sized metabolic networks.

Clearly, an automated simulation tool is required to set up the equations based on the stoichiometry and the atom transitions. To solve these equations algorithms comparable to the tool of Nöh et al. [29] could be applied. Nöh et al. [29] have shown that the solution of approxi-

mately 4000 equations can be handled without problems. Compared to this number, the additional kinetic equations (1) can be neglected. Also, the sensitivities can be rapidly calculated using a cluster computer [29]. As further refinement, Antoniewicz et al. [28] demonstrated that the dimension of isotopomer balances can be further reduced using the concept of Elementary Metabolite Units (EMU).

Authors' contributions

SAW: Did the simulations and statistical evaluation, writing of the manuscript, KN: Wrote the section about statistical measures, correction of the manuscript, WW: Supervision of the presented work, correction of the manuscript.

All authors read and approved the manuscript.

Appendix: Details on the measurement model

Within scenario S_{all} all pool concentrations can be measured and the labeling state is observed by mass isotopomers from unfragmented molecules over time (Section "Scenarios"). For a metabolite with n carbon atoms, a total of $n+1$ different mass isotopomers will be measured. For example, the unlabeled pool A#00 will be found at a mass $M(A)+0$ (henceforth denoted as $A+0$). The signal

A+1 is evoked by the singly labeled isotopomers A#01 and A#10. The fully labeled A#11 is found in the signal A+2. Usually some linear relationship between the signal intensity γ_A and the concentration can be supposed. A scalable measurement model with some scaling factor ω is introduced by Möllney et al. [10]. Following this approach, the mass isotopomer measurements for pool A can be expressed by:

$$\begin{aligned} \gamma_{+0}(t_k) &= \omega(t_k) \cdot a_{00}(t_k) \\ \gamma_{+1}(t_k) &= \omega(t_k) \cdot (a_{01}(t_k) + a_{10}(t_k)) \\ \gamma_{+2}(t_k) &= \omega(t_k) \cdot a_{11}(t_k) \end{aligned}$$

which can be shortly denoted by

$$\mathbf{y}(t_k) = \omega(t_k) \cdot \mathbf{M} \mathbf{x}(t_k)$$

The measurement equations with tandem mass isotopomers [37] are a little bit more evolved. The construction of the associated measurement matrix can be found in [37]. Its implementation is then done analogously to the example shown.

In contrast to an experiment without labeling, a metabolite pool is now distributed over various mass isotopomers depending on the labeling state. The total concentration has to be reconstructed from the single mass isotopomer measurements. Assuming that the labeling has no influence on the ionization, the calibration for the unlabeled metabolite can be used for all mass traces. Usually some linear relationship between the signal intensity γ and the sample concentration is supposed, given by the response factor γ . The sample concentration originates from various pre-processing steps (Figure 6). Thus, the measured signal is obtained from the intracellular concentration $c(t_k)$, the dilution factor $\phi(t_k)$ (sample preparation) and the response factor γ :

$$\gamma(t_k) = \gamma \phi(t_k) c(t_k) \quad \text{with} \quad c(t_k) = \sum_{i=0}^n c_{+i}(t_k)$$

Some specialities of the different factors have to be pointed out:

1. The response factor γ is a device dependent factor that will usually be constant for all measurement time points, but different for each metabolite:

$$\gamma = (\gamma_A, \dots, \gamma_E)^T$$

2. The dilution factor $\phi(t_k)$ reflects the sampling procedure and, thus, is time dependent (e.g. due to different amounts for neutralization). With regard to the metabo-

lites, different scenarios could be taken into account. Current observations show, that during quenching metabolites leak to different extents [40]. Therefore in this study, a dilution factor for each metabolite is introduced that is independent. The vector of dilution factors thus reads:

$$\phi(\xi) = (\phi_A(t_1), \dots, \phi_E(t_1), \dots, \phi_A(t_N), \dots, \phi_E(t_N))^T$$

Here a difference between the labeled and the non-labeled experiment has to be noticed. The dilution factor $\phi(t_k)$ is a pool factor, the isotopomers of this pool are diluted to the same extent.

3. The scaling factor $\omega(t_k)$ is used for scaling the simulated labeling enrichments to the absolute scale of the measurement device. Consequently it can be replaced by the previously discussed scaling factors:

$$\omega(\xi) = (\gamma_A \phi_A(t_1)/A, \gamma_B \phi_B(t_1)/B, \dots, \gamma_E \phi_E(t_N)/E)^T$$

The errors of the single factors used for the simulation study have been mentioned in the main text. How these errors can be included to the regression model has not been discussed. The response factor γ is determined from a standard addition and is assumed to be constant over the measurement sequence. This factor is determined for each metabolite.

In contrast, $\phi(t_k)$ is needed for each measurement time point. Thus, the regression Eq. (8) has to be extended by the scaling factors γ and $\phi(t_k)$:

$$\begin{pmatrix} \eta(\xi) \\ \vdots \\ \tilde{\phi}_A(t_1) \\ \tilde{\phi}_B(t_1) \\ \vdots \\ \tilde{\phi}_E(t_N) \\ \tilde{\gamma}_A \\ \vdots \\ \tilde{\gamma}_E \end{pmatrix} = \begin{pmatrix} \mathbf{g}(\mathbf{x}(\xi), \mathbf{c}(\xi)) \\ \vdots \\ \phi_A(t_1) \\ \phi_B(t_2) \\ \vdots \\ \phi_E(t_N) \\ \gamma_A \\ \vdots \\ \gamma_E \end{pmatrix} + \begin{pmatrix} \varepsilon(\xi) \\ \vdots \\ \varepsilon_{\phi_A}(t_1) \\ \varepsilon_{\phi_B}(t_2) \\ \vdots \\ \varepsilon_{\phi_E}(t_N) \\ \varepsilon_{\gamma_A} \\ \vdots \\ \varepsilon_{\gamma_E} \end{pmatrix}$$

Construction of the measurement covariance matrix

The covariance matrix Σ contains the measurement variances that can be obtained from multiple measurements or more common from measurement validation. Its construction depends on the regression model (Eq. (21)). In case that $\mathbf{g}(\mathbf{x}(\xi), \mathbf{c}(\xi))$ calculates the measurement vector ($A_{+0}, A_{+1}, A_{+2}, B_{+0}, B_{+1}, B_{+2}, C_{+0}, C_{+1}, \dots, E_{+2}$) at t_1 and in the following of t_2 until t_N . For sample preparation we have to take into account the dilution (Φ) of each metabolite at

36. Chatterjee S, Hadi AS: **Sensitivity Analysis in Linear Regression, Probability and Mathematical Statistics**. New York, John Wiley & Sons; 1988.
37. Rantanen A, Rousu J, Kokkonen JT, Tarkiainen V, Ketola RA: **Computing positional isotopomer distributions from tandem mass spectrometric data**. *Metab Eng* 2002, **4**:285-294.
38. Bader A: **A systematic approach to standard addition methods in instrumental analysis**. *J Chem Edu* 1980, **57**:703-706.
39. Nöh K, Wiechert W: **Experimental design principles for isotopically instationary ¹³C labeling experiments**. *Biotechnol Bioeng* 2006, **94**:234-251.
40. Bolten CJ, Kiefer P, Letisse F, Portais JC, Wittmann C: **Sampling for metabolome analysis of microorganisms**. *Anal Chem* 2007, **79**:3843-3849.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

