



# A Strategy for Functional Interpretation of Metabolomic Time Series Data in Context of Metabolic Network Information

Thomas Nägele<sup>1,2\*†</sup>, Lisa Fürtauer<sup>1†</sup>, Matthias Nagler<sup>1†</sup>, Jakob Weiszmann<sup>1†</sup> and Wolfram Weckwerth<sup>1,2</sup>

<sup>1</sup> Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria, <sup>2</sup> Vienna Metabolomics Center, University of Vienna, Vienna, Austria

## OPEN ACCESS

### Edited by:

Guowang Xu,  
Chinese Academy of Sciences, China

### Reviewed by:

Michał Jan Markuszewski,  
Medical University of Gdansk, Poland  
Huub C. J. Hoefsloot,  
University of Amsterdam, Netherlands

### \*Correspondence:

Thomas Nägele  
thomas.naegle@univie.ac.at

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Metabolomics,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 11 December 2015

**Accepted:** 19 February 2016

**Published:** 07 March 2016

### Citation:

Nägele T, Fürtauer L, Nagler M,  
Weiszmann J and Weckwerth W  
(2016) A Strategy for Functional  
Interpretation of Metabolomic Time  
Series Data in Context of Metabolic  
Network Information.  
*Front. Mol. Biosci.* 3:6.  
doi: 10.3389/fmolb.2016.00006

The functional connection of experimental metabolic time series data with biochemical network information is an important, yet complex, issue in systems biology. Frequently, experimental analysis of diurnal, circadian, or developmental dynamics of metabolism results in a comprehensive and multidimensional data matrix comprising information about metabolite concentrations, protein levels, and/or enzyme activities. While, irrespective of the type of organism, the experimental high-throughput analysis of the transcriptome, proteome, and metabolome has become a common part of many systems biological studies, functional data integration in a biochemical and physiological context is still challenging. Here, an approach is presented which addresses the functional connection of experimental time series data with biochemical network information which can be inferred, for example, from a metabolic network reconstruction. Based on a time-continuous and variance-weighted regression analysis of experimental data, metabolic functions, i.e., first-order derivatives of metabolite concentrations, were related to time-dependent changes in other biochemically relevant metabolic functions, i.e., second-order derivatives of metabolite concentrations. This finally revealed time points of perturbed dependencies in metabolic functions indicating a modified biochemical interaction. The approach was validated using previously published experimental data on a diurnal time course of metabolite levels, enzyme activities, and metabolic flux simulations. To support and ease the presented approach of functional time series analysis, a graphical user interface including a test data set and a manual is provided which can be run within the numerical software environment Matlab<sup>®</sup>.

**Keywords:** metabolic network, data integration, metabolomics, time series analysis, systems biology, network dynamics

## INTRODUCTION

The functional interpretation of experimental data in context of biochemical network information represents one of the central challenges in current biological research. While genome sequencing projects have enabled the reconstruction of genome-scale metabolic networks, their high dimensionality precludes a direct and intuitive application to interpret experimental data. Hence, although genome sequence information and metabolic networks have become available for

numerous organisms, tissues, or cell types (Herrgard et al., 2008; Chang et al., 2011; De Oliveira Dal'Molin and Nielsen, 2013; Thiele et al., 2013), functional metabolic data interpretation still represents a major obstacle in systems biology. Various mathematical and computational strategies from the fields of multivariate statistics, ordinary, and partial differential equations (ODEs/PDEs), optimization or statistical time series analysis have been developed and applied to reveal a biologically meaningful interpretation of comprehensive and multidimensional experimental data sets. For example, a computational model of starch metabolism in plants enabled the analysis of starch kinetics through diurnal metabolic and circadian sensors (Pokhilko et al., 2014). The authors developed a model of 28 ODEs which were numerically simulated in order to analyze diurnal kinetics of carbon metabolism *in silico*. In another study, the response of *Escherichia coli* to varying oxygen concentrations was analyzed applying a mathematical model of the central metabolism (Ederer et al., 2014). Here, the authors derived a prediction about the impact of product formation on biomass concentration using steady state simulations at varying environmental conditions.

Both examples for mathematical modeling differ in organism and application. Besides, the dynamic approach can be distinguished from the steady state approach. However, in both approaches, dynamics of metabolic systems can be described by sets of ODEs. If sufficient kinetic information is available, such ODEs can be numerically integrated revealing simulated metabolic concentrations depending on time, enzyme parameters, thermodynamic constraints, etc. Yet, statistically robust experimental enzyme kinetic information often limits the applicability of such modeling approaches. Particularly, the resolution of enzyme activities, substrate affinities, or inhibitory constants is very laborious and only possible if well-established experimental assays and sufficient biochemical knowledge are available. Additionally, uncertainties about model structures and reaction kinetics complicate the interpretation of a numerically simulated output (Schaber et al., 2009). Such limitations have been addressed by different theoretical approaches, for example by structural kinetic modeling, SKM (Steuer et al., 2006). In the SKM approach, local linear models are applied to explore and statistically analyze a given parameter space without the need for explicit information about functional forms of enzyme kinetics and rate equations. Finally, a Jacobian matrix is derived which characterizes the dynamic capabilities of a metabolic system at a certain steady state. In previous publications, we have developed a procedure to determine Jacobian matrices directly from experimental metabolomics data (Nägele, 2014; Nägele et al., 2014). Based on experimental metabolic (co)variance information a metabolic regulator was identified indicating a strategy how plant metabolism is reprogrammed during exposure to energy limiting conditions. In a different context, other studies have also shown that it is possible to infer regulatory information about metabolic steady states from experimental data with such approaches (see e.g., Steuer et al., 2003; Sun and Weckwerth, 2012; Kügler and Yang, 2014).

Beyond these approaches of dynamic and steady state modeling, time series analysis and related regression models

offer another mathematical strategy to reveal information about molecular system dynamics (Schelter, 2006). For example, Dutta and co-workers developed an algorithm for identification of differentially expressed genes in a time series experiment (Dutta et al., 2007), which they also applied to integrate transcriptome and metabolome data (Dutta et al., 2009). In another study, statistical modeling and regression analysis revealed a nitrogen-dependent modulation of root system architecture in the genetic model plant *Arabidopsis thaliana* (Araya et al., 2015). While these exemplarily mentioned studies present only a very small fraction of possible statistical applications, it already becomes evident that these are promising and necessary mathematical approaches to reveal biologically meaningful information from comprehensive experimental data sets being preliminary for hypothesis generation and experimental validation. However, a common problem of regression and correlation approaches in a biochemical context is a missing functional linkage of the results to causal biochemical interrelations, i.e., enzymatically driven reactions. To overcome this limitation and to facilitate the biochemical interpretation of the statistical results, the present study derives a theoretical connection between mathematical approaches of ODE-based dynamic modeling and statistical time series analysis. Based on the stoichiometric matrix information of a metabolic network, ratios of time-dependent derivative functions were built providing an estimate for the strength and probability of a metabolic interaction during the time course. The suggested strategy was tested using previously published experimental data sets on diurnal and stress-induced dynamics of metabolite concentrations and related enzyme kinetic information. Finally, a graphical user interface for Matlab is provided which intends to facilitate the application of the presented strategy.

## RESULTS

### Deriving Metabolic Functions by Inverse Variance-Weighted Regression Analysis

Time-dependent dynamics of metabolite concentrations in a biochemical network can be described by a set of ODEs:

$$\frac{d}{dt}\mathbf{M}(t) = \mathbf{N}\mathbf{v}(\mathbf{M},\mathbf{p},t) = \mathbf{f}(\mathbf{M},\mathbf{p},t) \quad (1)$$

Here,  $\mathbf{M}$  represents an  $n$ -dimensional vector of mean metabolite concentrations ( $c_n$ ),  $\mathbf{N}$  is the  $n \times k$  stoichiometric matrix and  $\mathbf{v}$  describes the  $k$ -dimensional vector of reaction rates which depend on metabolite concentrations  $\mathbf{M}$ , enzyme kinetic parameters  $\mathbf{p}$  and time  $t$ . The right side of the ODE system can be summarized by metabolic functions,  $\mathbf{f}(\mathbf{M},\mathbf{p},t)$ . Hence, these metabolic functions define the time-dependent changes in metabolite concentration as a sum of all biochemical reactions either consuming or producing a metabolite. A metabolic steady state is described by ODEs which equal zero:

$$\frac{d}{dt}\mathbf{M}(t) = 0 \quad (2)$$

Linearization enables the investigation of the system behavior close to a steady state. The linearization process results in the so-called Jacobian matrix  $J$  which characterizes the dynamic properties of the system at a steady state:

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial c_1} & \dots & \frac{\partial f_1}{\partial c_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial c_1} & \dots & \frac{\partial f_n}{\partial c_n} \end{pmatrix} \quad (3)$$

Hence, in a biochemical context, the Jacobian matrix  $J$  describes the behavior of metabolic functions  $f_i$  (for  $i = 1, \dots, n$ ) of a metabolic network with regard to small changes of variables  $c_i$  (for  $i = 1, \dots, n$ ), i.e., metabolite concentrations at the considered steady state. The information if a metabolic function  $f_i$  biochemically depends on the concentration of a metabolite  $c_i$  is provided by the stoichiometric matrix  $N$  of a metabolic network (see Equation 1).

To derive, i.e., predict, time continuous information from time discrete experimental observations, interpolation methods can be applied. To prevent unrealistic oscillations of high-degree polynomial interpolation, intervals of approximation can be partitioned in subintervals which can be approximated, for example, by cubic polynomials which form a cubic spline  $S_{c_i}(t)$  (see e.g., Bronstein et al., 2008):

$$S_{c_i}(t) = \alpha_{ij} + \beta_{ij}(t - t_j) + \gamma_{ij}(t - t_j)^2 + \delta_{ij}(t - t_j)^3 \quad (4)$$

Here, it is  $t \in [t_j, t_{j+1}]$  with ( $j = 1, 2, \dots, z-1$ ), where  $z$  represents the number of interpolation nodes ( $t_j, g_{ij}$ ), and it is  $S_{c_i}(t_j) = g_{ij}$ . Interpolation coefficients are represented by  $\alpha, \beta, \gamma$ , and  $\delta$ . Due to the occurrence of experimental errors, the requirement of  $S_{c_i}(t_j) = g_{ij}$  is not fulfilled which prevents the suitability of such a type of interpolation. Instead, a smoothing element can be introduced accounting for those experimental errors:

$$\min \left( \sum_{j=1}^z w_{ij} [g_{ij} - S_{c_i}(t_j)]^2 + \lambda \int_{t=t_1}^{t_z} [S_{c_i}''(t)]^2 dt \right) \quad (5)$$

Here,  $w_{ij}$  represents a weighting factor,  $S_{c_i}''(t)$  is the second derivative of  $S_{c_i}(t)$ , and  $\lambda$  (with  $\lambda \geq 0$ ) represents a smoothing factor. For  $\lambda = 0$ , one obtains the cubic spline interpolation, while the degree of smoothing increases with the value of  $\lambda$ . To connect the smoothing spline generation to experimentally observed errors we defined the weighting factor  $w_{ij}$  to equal the inverse variance information, i.e., the inverse squared standard deviation:

$$w_{ij} = \sigma_{ij}^{-2} = \left( \frac{1}{r-1} \sum_{k=1}^r (c_{ij,k} - \bar{c}_{ij})^2 \right)^{-1} \quad (6)$$

Here,  $r$  represents the number of experimental and independent replicates.

Merging Equations (1), (5) and (6) and replacing  $g_{ij}$  by the mean concentration of metabolite  $i$  at time point  $j$ ,  $\bar{c}_{ij}$ , reveals a

description of metabolic functions by inverse variance-weighted regression analysis:

$$\begin{aligned} \frac{d}{dt} M_i(t) &= f_i(\mathbf{M}, \mathbf{p}, t) \\ &= \frac{d}{dt} \left[ \min \left( \sum_{j=1}^z \left[ \left( \frac{1}{r-1} \sum_{k=1}^r (c_{ij,k} - \bar{c}_{ij})^2 \right)^{-1} \right. \right. \right. \\ &\quad \left. \left. \left. [\bar{c}_{ij} - S_{c_i}(t_j)]^2 \right] + \lambda \int_{t=t_1}^{t_z} [S_{c_i}''(t)]^2 dt \right) \right] \quad (7) \end{aligned}$$

Hence, building the first derivative of the smoothed interpolation of experimental time-course data reveals information about the connected metabolic function. In the present study, this approach was applied to evaluate a diurnal time course of previously published metabolite concentrations (Nägele et al., 2012) belonging to the central carbohydrate metabolism in leaves of the genetic model plant *A. thaliana*. Diurnal dynamics of metabolic functions are shown exemplarily (**Figure 1**) for the metabolite pools of sucrose (Suc) and sugar phosphates (SP) in a control experiment (non-cold acclimated, na) and after exposure to low temperature (acc).

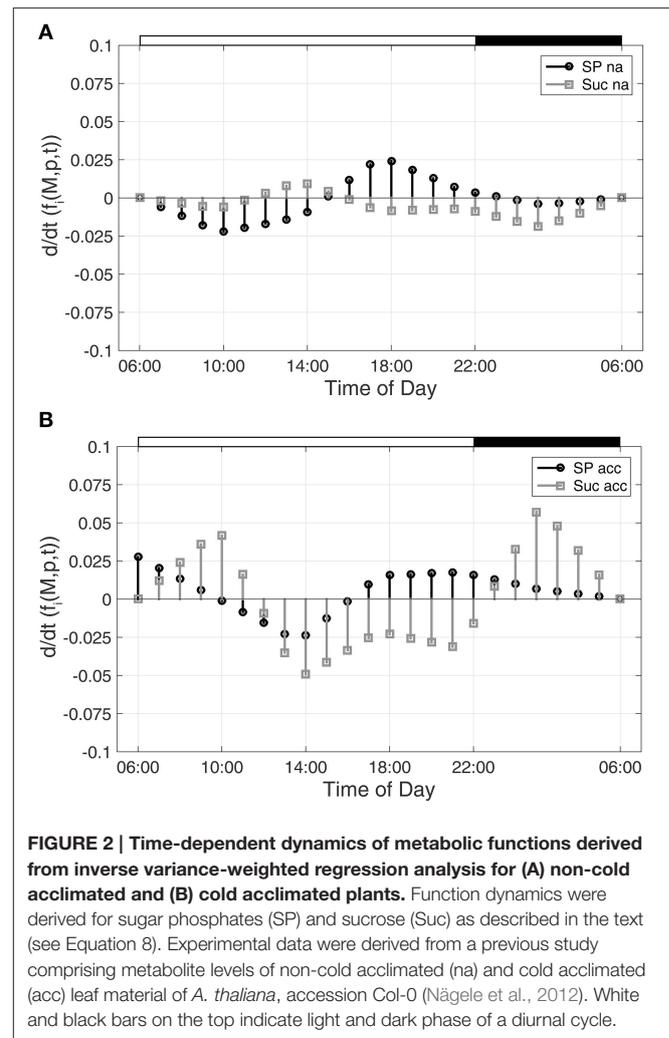
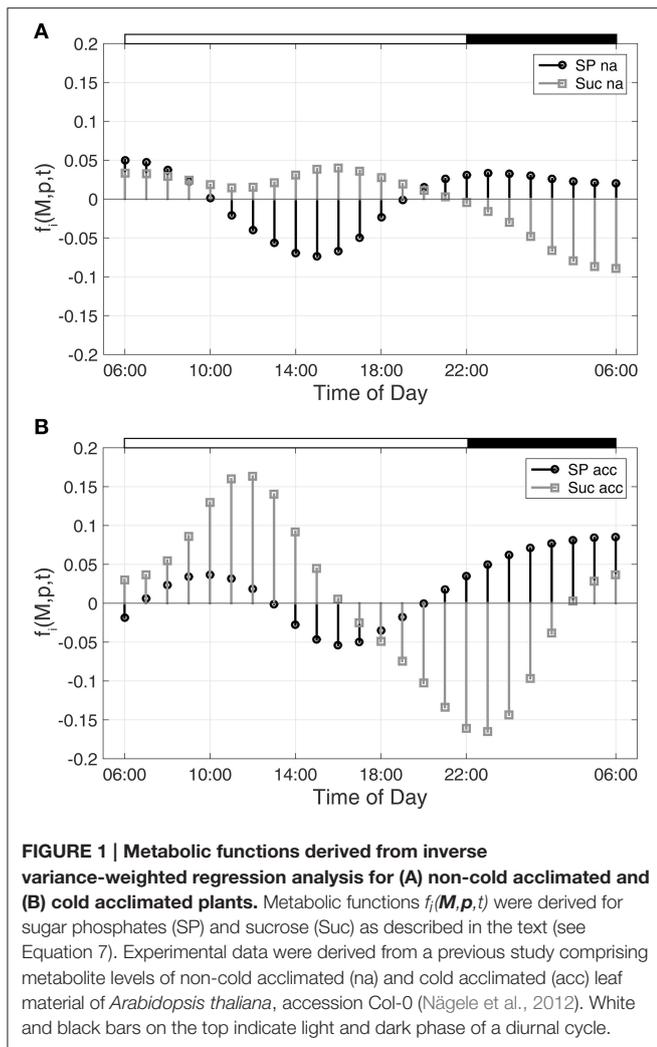
To characterize time-dependent changes in metabolic functions, the second time-dependent derivative was built from the approximated diurnal time course of metabolite concentrations:

$$\begin{aligned} \frac{d^2}{dt^2} M_i(t) &= \frac{d}{dt} f_i(\mathbf{M}, \mathbf{p}, t) \\ &= \frac{d}{dt} \left[ \min \left( \sum_{j=1}^z \left[ \left( \frac{1}{r-1} \sum_{k=1}^r (c_{ij,k} - \bar{c}_{ij})^2 \right)^{-1} \right. \right. \right. \\ &\quad \left. \left. \left. [\bar{c}_{ij} - S_{c_i}(t_j)]^2 \right] + \lambda \int_{t=t_1}^{t_z} [S_{c_i}''(t)]^2 dt \right) \right] \quad (8) \end{aligned}$$

As described for **Figure 1**, diurnal dynamics of those time-dependent changes of metabolic functions are also shown exemplarily for Suc and SP (**Figure 2**).

## Connecting Metabolic Functions Based on Biochemical Network Information

While the metabolic time-course information derived before characterizes time-dependent rates of changes in each considered metabolite concentration separately (see Equations 7 and 8), information of biochemical interdependencies, i.e., the information about a substrate - product relationship between two or more metabolites, is only contained implicitly. To explicitly analyze and visualize these biochemical interdependencies with regard to the time-dependent rates of concentrations changes, a metabolic  $n \times n$  interaction matrix,  $Y$ , was derived where  $n$  represents the number of metabolites comprised by the model. In  $Y$ , each entry indicates whether two metabolites are biochemically connected (entry: 1) or not (entry: 0). The interaction is characterized analogous to entries of the Jacobian



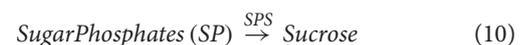
matrix (Equation 3): if the metabolic function of metabolite A is biochemically connected to changes in concentrations of metabolite B, the corresponding entry in  $\mathbf{Y}$  is 1. Information about metabolic functions is given row-wise, while biochemically connected metabolites are indicated column-wise for each function. In a simple example, containing three reactions ( $r_1$ – $r_3$ ) and four metabolites (A–D), the construction and content of  $\mathbf{Y}$  is exemplified (Figure 3). The diagonal entries indicate the biochemical dependencies of metabolic functions on substrate concentrations. For example,  $Y_{11} = 1$  indicates that metabolic function  $f(A, t)$  depends on the concentration of A( $t$ ). The non-diagonal entries describe interdependencies between different metabolite pools. For example,  $Y_{21} = 1$  indicates that metabolic function  $f(B, t)$  depends on the concentration of A( $t$ ).

Based on a previously published metabolic model (Nägele et al., 2012), an interaction matrix  $\mathbf{Y}$  was derived for the central carbohydrate metabolism in leaves of *A. thaliana*. The metabolic functions (Equation 7) and their time-dependent derivatives (Equation 8) were related to each other according to the entries of  $\mathbf{Y}$ . This finally resulted in functions  $\omega(a \rightarrow b, t)$  indicating changes in metabolic functions of  $b$  in context of concentration

changes of  $a$  which might represent substrates, inhibitors or activators:

$$\omega(a \rightarrow b, t) = \frac{\frac{d}{dt} f_b(\mathbf{M}, \mathbf{p}, t)}{f_a(\mathbf{M}, \mathbf{p}, t)}, \quad D = \{\mathbb{R} \setminus f_a(\mathbf{M}, \mathbf{p}, t) = 0\} \quad (9)$$

With regard to the analyzed time-course of sugar phosphate (SP) and sucrose (Suc) concentrations (see Figures 1, 2),  $\omega(\text{SP} \rightarrow \text{Suc}, t)$  revealed information about the reaction of sucrose biosynthesis, catalyzed by the enzyme sucrose phosphate synthase (SPS):



In detail,  $\omega(\text{SP} \rightarrow \text{Suc}, t)$  described changes in the metabolic function of sucrose in context of concentration changes of its biochemical substrate sugar phosphates:

$$\omega(\text{SP} \rightarrow \text{Suc}, t) = \frac{\frac{d}{dt} f_{\text{Suc}}(\mathbf{M}, \mathbf{p}, t)}{f_{\text{SP}}(\mathbf{M}, \mathbf{p}, t)} \quad (11)$$

Comparing  $\omega(\text{SP} \rightarrow \text{Suc}, t)$  for na and acc plants revealed a noticeable difference between both conditions within the first

4 h of the day (Figure 4). Interestingly, in the same time period, simulations of sucrose biosynthesis, based on a system of ordinary differential equations (ODEs), revealed a similar picture in which rates of sucrose biosynthesis were decreased only in acc plants (Nägele et al., 2012).

### Characterization of $\omega(t)$

$\omega(t)$  represents a ratio of metabolic functions and related derivatives. Hence, the unit of  $\omega(t)$  is derived from the flux unit of a metabolic function,  $[mM s^{-1}]$ :

$$|\omega(a \rightarrow b, t)| = \frac{\left[ \frac{d}{dt} f_b(\mathbf{M}, \mathbf{p}, t) \right]}{[f_a(\mathbf{M}, \mathbf{p}, t)]} = \frac{[mM s^{-2}]}{[mM s^{-1}]} = \left[ \frac{1}{s} \right] \quad (12)$$

Here, concentrations are given in mM ( $mmol l^{-1}$ ) and the time unit is seconds (s).

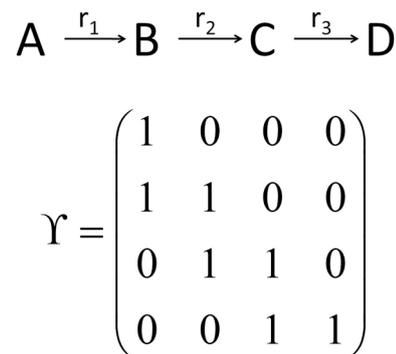
This results in the unit of a rate or frequency. Hence,  $|\omega(a \rightarrow b, t)|$  was interpreted as oscillations of a metabolic function per time-period with reference to a biochemical effector.

In the case of  $|\omega(a \rightarrow b, t)| \rightarrow \infty$  for  $t \rightarrow \tau$ , the influence of the biochemical effector on a metabolic function was defined to be strong, while  $|\omega(a \rightarrow b, t)| \rightarrow 0$  for  $t \rightarrow \tau$  indicated a weak effect. In detail,  $|\omega(a \rightarrow b, t)| \rightarrow \infty$  for  $t \rightarrow \tau$  indicates that it is  $|d/dt(f_b(\mathbf{M}, \mathbf{p}, t))| \gg |f_a(\mathbf{M}, \mathbf{p}, t)|$ . Vice versa,  $|\omega(a \rightarrow b, t)| \rightarrow 0$  for  $t \rightarrow \tau$  indicates that  $|d/dt(f_b(\mathbf{M}, \mathbf{p}, t))| \ll |f_a(\mathbf{M}, \mathbf{p}, t)|$ .

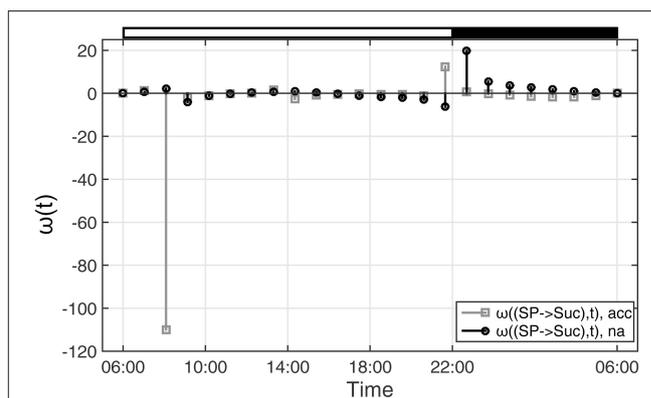
### Application Example: Stress-Induced Metabolic Reprogramming in *Arabidopsis thaliana*

While in the above mentioned example, the calculation and interpretation of  $\omega(t)$  was demonstrated in context of a previously published kinetic ODE model, another published data set was analyzed by this strategy comprising metabolite levels of the primary and secondary metabolism in *A. thaliana* (Doerfler et al., 2013). In the experiment performed by Doerfler and co-workers, a combined strategy of gas chromatography and liquid chromatography coupled to mass spectrometry was applied in order to reveal a comprehensive picture of metabolic reprogramming during exposure to low temperature and high light intensity. The time period of stress exposure comprised more than 2 weeks which allowed for the analysis of a short- and long-term acclimation response in the metabolome of *A. thaliana*, accession Col-0. A central output of the study was the characterization of metabolomic and regulatory dynamics at the interface of primary and secondary metabolism. The authors observed a fast increase of stress-responsive compounds, e.g., sucrose, which became significant already after 2 days of stress exposure, while the interaction with the secondary metabolism, resulting in biosynthesis of flavonoids, became most significant after 8 days of stress exposure.

To prove the suitability of deriving the absolute value function  $|\omega(a \rightarrow b, t)|$  in order to reveal steps of metabolic regulation within a considered time interval, regression analysis and metabolic functions were calculated for the dataset of Doerfler et al. (2013) and compared to their observations. The metabolic interaction matrix  $\mathbf{Y}$  was derived from the metabolic



**FIGURE 3 | Schematic reaction chain and the derived interaction matrix  $\mathbf{Y}$ .** Rows (metabolic functions) and columns (metabolites) of  $\mathbf{Y}$  describe biochemical interactions of metabolites A (first row/column), B (second row/column), C (third row/column), and D (fourth row/column). Entries, i.e., 0 and 1, indicate if two metabolites interact (entry 1) or not (entry 0).



**FIGURE 4 | Functions  $\omega(t)$  indicating changes in metabolic functions in context of concentration changes of biochemical interaction partners.**  $\omega(t)$  was calculated as described in the main text (see Equation 9). Results of  $\omega(t)$  are shown for the reaction of sucrose biosynthesis for non-cold acclimated (na) and cold acclimated (acc) leaf material of *A. thaliana*, accession Col-0 (Nägele et al., 2012). White and black bars on the top indicate light and dark phase of a diurnal cycle.

network model which was previously suggested and applied for inverse approximations of the Jacobian matrix (Doerfler et al., 2013). For regression analysis and for integration of metabolic network information we developed and applied a graphical user interface (FEMTO, Functional Evaluation of Metabolic Time series Observations) which is based on the numerical software environment Matlab<sup>®</sup> (<http://www.mathworks.com>), and which is provided in the supplements together with a user manual (Supplementary Files S1, S2).

To characterize sucrose metabolism, changes of the metabolic function of sucrose were related to changes in sucrose concentrations:

$$|\omega(\text{Suc} \rightarrow \text{Suc}, t)| = \left| \frac{\frac{d}{dt} f_{\text{Suc}}(\mathbf{M}, \mathbf{p}, t)}{f_{\text{Suc}}(\mathbf{M}, \mathbf{p}, t)} \right| \quad (13)$$

For time-dependent characterization of flavonoid dynamics, changes in the flavonoid (Flav) metabolic function were related to substrate concentration changes, i.e., phenylalanine (Phe) dynamics:

$$|\omega(\text{Phe} \rightarrow \text{Flav}, t)| = \left| \frac{\frac{d}{dt} f_{\text{Flav}}(\mathbf{M}, \mathbf{p}, t)}{f_{\text{Phe}}(\mathbf{M}, \mathbf{p}, t)} \right| \quad (14)$$

Results of metabolic function analysis and the resulting time course of  $|\omega(t)|$  revealed an early de-regulation of sucrose metabolism during the first 2 days of stress exposure (**Figure 5A**), while the peak value of  $|\omega(t)|$  for flavonoid biosynthesis occurred delayed after 8 days (**Figure 5B**).

These findings coincide with the previous findings described by Doerfler and colleagues who applied the method of Granger causality time-series correlation and a covariance-based inverse approximation of Jacobian matrices to reveal strategies of metabolic regulation (Doerfler et al., 2013). Conclusions which have been drawn from the  $|\omega(t)|$  calculation were found to be highly similar to the output of other statistical methods, finally substantiating the validity of the suggested workflow and the derived method to unravel time points of regulatory perturbation in a biochemical system.

## DISCUSSION

Mathematical analysis of biochemical system dynamics represents a central focus of current biomathematical, biochemical and biotechnological research due to the need for methods and algorithms enabling a functional interpretation of experimental data in context of a biochemical network. Particularly, system dynamics which arise due to circadian regulation (Harmer, 2009; Kumar Jha et al., 2015), diurnal metabolic adjustment (Geiger and Servaites, 1994; Pokhilko et al., 2014) or stress-induced metabolic reprogramming (Jozefczuk et al., 2010; Kanshin et al., 2015) are hardly traceable by intuition. Hence, this indicates a strong need for suitable theoretical approaches being capable of resolving and functionally connecting molecular moieties with underlying biochemical regulation.

Various theoretical strategies have addressed this complex issue, providing a comprehensive methodological platform for time-series analysis, dynamic flux balance analysis, kinetic and Boolean modeling (see e.g., Mahadevan et al., 2002; Schelker, 2006; Rohwer, 2012; Steinway et al., 2015). In a recent approach, Willemsen and colleagues have modified the approach of dynamic flux balance analysis by incorporating time-resolved metabolomics measurements (Willemsen et al., 2015). With their extended method, the authors derived an estimate of dynamic flux profiles which allowed them to generate and test hypotheses related to environmentally induced molecular dynamics. In another recent study, a computational approach was suggested to translate metabolomics data into flux information (Cortassa et al., 2015). One main methodological difference between the studies of Willemsen et al. and Cortassa et al. was the extent of kinetic information which was needed to estimate cellular behavior and metabolic fluxes. While Willemsen et al. focused

on minimalistic kinetic information, the study of Cortassa and co-workers used a detailed kinetic model of glucose catabolic pathways to derive flux information.

In our presented approach, flux information, which was implicitly derived from spline interpolation, was interpreted only indirectly by comparing time-dependent changes in metabolic functions to concentration changes of biochemical reaction partners. This procedure revealed information about a rate which was interpreted in terms of metabolic functions related to concentration changes in a substrate or co-substrate. Comparing derived results to other methods, it was shown that changes in ratios of second- to first-order derivatives between functionally connected variables potentially reveal time points of regulatory perturbation within a biochemical interaction. Hence, these observed perturbations might indicate a change in enzymatic activity, protein abundance, or allosteric regulation ultimately leading to a change in the metabolic functions.

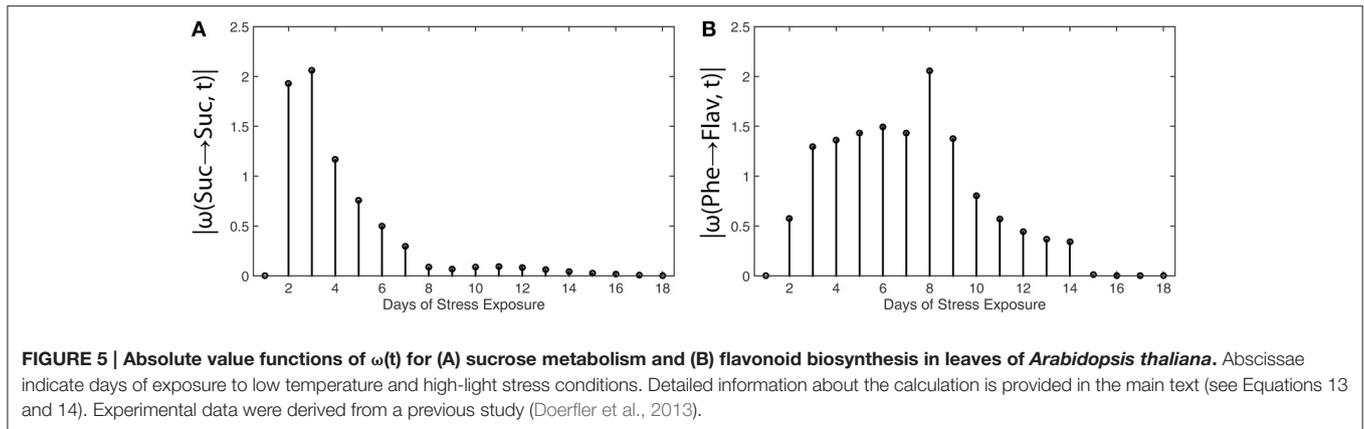
The information content of the introduced time-dependent functions  $\omega(t)$  is related to entries of the Jacobian matrix  $J$  (see Equation 3) indicating the dynamics of metabolic functions with respect to (small) concentration changes at a certain steady state. This theoretical connection of  $J$  and  $\omega(t)$  at a considered time point  $t_0$  might be illustrated in a simple first-order reaction scheme.



Here, substance  $A$  is interconverted into substance  $B$ , and the reaction velocity is characterized by the rate constant  $k$ . The time-dependent change in concentration of  $A$  equals  $dA/dt = -k \cdot A$ . Hence, a general solution of this ODE is given by  $A(t) = A_0 e^{-kt}$  which finally yields  $J_{11}(t_0) = \omega(A \rightarrow A, t_0) = -k$ .

With this, the information of  $\omega(t)$  becomes comparable to entries of the Jacobian matrix  $J$ . Yet, in contrast to entries of  $J$ , characterizing dynamic properties of a metabolic steady state ( $d/dt \mathbf{M}(t) = 0$ ), functions  $\omega(t)$  were derived from a time series of experimental data and might rather be valid for a non-infinitesimal than for an infinitesimal time frame. While for  $\lim_{t \rightarrow t_0} |\omega(t)|$ ,  $|\omega(t)|$  might be assumed to approach entries of  $J$ , this was not tested in the present study and would need experimental validation. In addition, while a connection, and probable correlation, to other molecular levels, such as the proteome or transcriptome, was not experimentally analyzed, this might be a promising target for analysis in future studies. However, the incorporation of an interaction matrix, which, in the present study, was derived from a previously published reaction network, and which might be derived from genome-scale metabolic reconstruction works in future studies (Weckwerth, 2011; King et al., 2015), provides direct evidence for the biochemical and physiological relevance of the performed theoretical analysis.

While our results indicate a realistic and biochemically interpretable output of the presented method, limitations of application might occur due to several reasons. First, the presented method significantly depends on the knowledge about the biochemical network structure and involved regulatory interactions, e.g., feedback inhibition or feedforward activation. Although regression analysis of time series data might be



performed for all network components independently, deriving a reliable biochemical interaction matrix  $Y$  is essential to reveal realistic information about time-dependent changes in metabolic interactions. A second central prerequisite for a meaningful regression analysis is the design of an adequate experimental setup. This comprises the number of biological (independent) replicates as well as the number and interval of sampling points. It is hardly possible to generalize a number of replicates or sampling points due to heterogeneous technical or environmental fluctuations which are introduced by different analytical techniques, growth conditions or sample types. Yet, spanning various experimental scenarios, it might be generalized that the interval of sampling points is crucial to be able to discriminate between metabolic fast or short-term responses and slow or long-term responses. Particularly to resolve fast metabolic regulation, a narrow sampling interval is needed in order to prevent any over-interpretation of regression analysis and related derivatives. Comparing the presented approach to methods of metabolic modeling, a third major limitation is the missing predictive output by model simulations. For example, enzyme kinetic models of metabolism aim at going beyond the time interval of measured rate constants or metabolite concentrations to predict changes in system dynamics under changing environmental conditions or due to a mutated gene. However, although our presented method cannot afford this simulation output, time-dependent changes within the considered time interval might indicate regulatory bottlenecks and kinetic information supporting the numerical solution and simulation of metabolic ODE models.

In summary, the suggested approach intends to promote the functional interpretability of metabolic time series data in context of metabolic network information. Particularly with regard to multidimensional metabolomics data sets, this might unravel strategies of complex biochemical regulation and might overcome some limitations in the generation of testable hypotheses as we have discussed previously (Nägele and Weckwerth, 2012). Finally, the direct integration of biochemical network information with experimental data promises to enable the functional interpretation and the causal connection of various levels of molecular organization.

## MATERIALS AND METHODS

The described procedure of data analysis, spline interpolation and graphical representation was performed within the numerical software environment Matlab<sup>®</sup>. A Matlab-based graphical user interface (FEMTO, Functional Evaluation of Metabolic Time series Observations) was developed and is provided, together with a user manual, in the supplements (Supplementary Files S1, S2).

## AUTHOR CONTRIBUTIONS

TN, LF, MN, and JW performed data analysis, statistics, wrote the source code of the graphical user interface and wrote the paper. TN and WW conceived the study and wrote the paper. All authors read and approved the final version of the manuscript.

## FUNDING

This work was supported by the Austrian Science Fund, FWF, Project P 26342 and Project I 2071.

## ACKNOWLEDGMENTS

We would like to thank the members of the MoSys Department for constructive advice and fruitful discussions. Further, we thank Arnd G. Heyer from the Department of Plant Biotechnology at the University of Stuttgart, Germany, for providing experimental data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmolb.2016.00006>

**Supplementary File S1 | Matlab code-file and test data file.**

**Supplementary File S2 | Manual for the graphical user interface FEMTO 1.0.**

## REFERENCES

- Araya, T., Kubo, T., von Wirén, N., and Takahashi, H. (2015). Statistical modeling of nitrogen-dependent modulation of root system architecture in *Arabidopsis thaliana*. *J. Integr. Plant Biol.* doi: 10.1111/jipb.12433. [Epub ahead of print].
- Bronstein, I. J. N., Semendjajew, K. A., Musiol, G., and Mühlig, H. (2008). *Taschenbuch der Mathematik*. Frankfurt am Main: Wissenschaftlicher Verlag Harri Deutsch GmbH.
- Chang, R. L., Ghamsari, L., Manichaikul, A., Hom, E. F., Balaji, S., Fu, W., et al. (2011). Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Mol. Syst. Biol.* 7, 518. doi: 10.1038/msb.2011.52
- Cortassa, S., Caceres, V., Bell, L. N., O'Rourke, B., Paolocci, N., and Aon, M. A. (2015). From metabolomics to fluxomics: a computational procedure to translate metabolite profiles into metabolic fluxes. *Biophys. J.* 108, 163–172. doi: 10.1016/j.bpj.2014.11.1857
- De Oliveira Dal'Molin, C. G., and Nielsen, L. K. (2013). Plant genome-scale metabolic reconstruction and modelling. *Curr. Opin. Biotechnol.* 24, 271–277. doi: 10.1016/j.copbio.2012.08.007
- Doerfler, H., Lyon, D., Nägele, T., Sun, X., Fragner, L., Hadacek, F., et al. (2013). Granger causality in integrated GC-MS and LC-MS metabolomics data reveals the interface of primary and secondary metabolism. *Metabolomics* 9, 564–574. doi: 10.1007/s11306-012-0470-0
- Dutta, B., Kanani, H., Quackenbush, J., and Klapa, M. I. (2009). Time-series integrated “omic” analyses to elucidate short-term stress-induced responses in plant liquid cultures. *Biotechnol. Bioeng.* 102, 264–279. doi: 10.1002/bit.22036
- Dutta, B., Snyder, R., and Klapa, M. I. (2007). Significance analysis of time-series transcriptomic data: a methodology that enables the identification and further exploration of the differentially expressed genes at each time-point. *Biotechnol. Bioeng.* 98, 668–678. doi: 10.1002/bit.21432
- Ederer, M., Steinsiek, S., Stagge, S., Rolfe, M. D., Ter Beek, A., Knies, D., et al. (2014). A mathematical model of metabolism and regulation provides a systems-level view of how *Escherichia coli* responds to oxygen. *Front. Microbiol.* 5:124. doi: 10.3389/fmicb.2014.00124
- Geiger, D. R., and Servaites, J. C. (1994). Diurnal regulation of photosynthetic carbon metabolism in C3 plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 45, 235–256. doi: 10.1146/annurev.pp.45.060194.001315
- Harmer, S. L. (2009). The circadian system in higher plants. *Annu. Rev. Plant Biol.* 60, 357–377. doi: 10.1146/annurev.arplant.043008.092054
- Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* 26, 1155–1160. doi: 10.1038/nbt1492
- Jozefczuk, S., Klie, S., Catchpole, G., Szymanski, J., Cuadros-Inostroza, A., Steinhauser, D., et al. (2010). Metabolomic and transcriptomic stress response of *Escherichia coli*. *Mol. Syst. Biol.* 6, 364. doi: 10.1038/msb.2010.18
- Kanshin, E., Kubiniok, P., Thattikota, Y., D'Amours, D., and Thibault, P. (2015). Phosphoproteome dynamics of *Saccharomyces cerevisiae* under heat shock and cold stress. *Mol. Syst. Biol.* 11, 813. doi: 10.15252/msb.20156170
- King, Z. A., Lloyd, C. J., Feist, A. M., and Palsson, B. O. (2015). Next-generation genome-scale models for metabolic engineering. *Curr. Opin. Biotechnol.* 35, 23–29. doi: 10.1016/j.copbio.2014.12.016
- Kügler, P., and Yang, W. (2014). Identification of alterations in the Jacobian of biochemical reaction networks from steady state covariance data at two conditions. *J. Math. Biol.* 68, 1757–1783. doi: 10.1007/s00285-013-0685-3
- Kumar Jha, P., Challet, E., and Kalsbeek, A. (2015). Circadian rhythms in glucose and lipid metabolism in nocturnal and diurnal mammals. *Mol. Cell Endocrinol.* 418(Pt 1), 74–88. doi: 10.1016/j.mce.2015.01.024
- Mahadevan, R., Edwards, J. S., and Doyle, F. J. III (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi: 10.1016/S0006-3495(02)73903-9
- Nägele, T. (2014). Linking metabolomics data to underlying metabolic regulation. *Front. Mol. Biosci.* 1:22. doi: 10.3389/fmolb.2014.00022
- Nägele, T., Mair, A., Sun, X., Fragner, L., Teige, M., and Weckwerth, W. (2014). Solving the differential biochemical Jacobian from metabolomics covariance data. *PLoS ONE* 9:e92299. doi: 10.1371/journal.pone.0092299
- Nägele, T., Stutz, S., Hörmiller, I., and Heyer, A. G. (2012). Identification of a metabolic bottleneck for cold acclimation in *Arabidopsis thaliana*. *Plant J.* 72, 102–114. doi: 10.1111/j.1365-313X.2012.05064.x
- Nägele, T., and Weckwerth, W. (2012). Mathematical modeling of plant metabolism—from reconstruction to prediction. *Metabolites* 2, 553–566. doi: 10.3390/metabo2030553
- Pokhilko, A., Flis, A., Sulpice, R., Stitt, M., and Ebenhöf, O. (2014). Adjustment of carbon fluxes to light conditions regulates the daily turnover of starch in plants: a computational model. *Mol. Biosyst.* 10, 613–627. doi: 10.1039/c3mb70459a
- Rohwer, J. M. (2012). Kinetic modelling of plant metabolic pathways. *J. Exp. Bot.* 63, 2275–2292. doi: 10.1093/jxb/ers080
- Schaber, J., Liebermeister, W., and Klipp, E. (2009). Nested uncertainties in biochemical models. *IET Syst. Biol.* 3, 1–9. doi: 10.1049/iet-syb:20070042
- Schelter, B. (2006). *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*. Weinheim: WILEY-VCH.
- Steinway, S. N., Biggs, M. B., Loughran, T. P. Jr., Papin, J. A., and Albert, R. (2015). Inference of network dynamics and metabolic interactions in the gut microbiome. *PLoS Comput. Biol.* 11:e1004338. doi: 10.1371/journal.pcbi.1004338
- Steuer, R., Gross, T., Selbig, J., and Blasius, B. (2006). Structural kinetic modeling of metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11868–11873. doi: 10.1073/pnas.0600013103
- Steuer, R., Kurths, J., Fiehn, O., and Weckwerth, W. (2003). Observing and interpreting correlations in metabolomic networks. *Bioinformatics* 19, 1019–1026. doi: 10.1093/bioinformatics/btg120
- Sun, X. L., and Weckwerth, W. (2012). COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8, S81–S93. doi: 10.1007/s11306-012-0399-3
- Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31, 419–425. doi: 10.1038/nbt.2488
- Weckwerth, W. (2011). Unpredictability of metabolism—the key role of metabolomics science in combination with next-generation genome sequencing. *Anal. Bioanal. Chem.* 400, 1967–1978. doi: 10.1007/s00216-011-4948-9
- Willemsen, A. M., Hendrickx, D. M., Hoefsloot, H. C., Hendriks, M. M., Wahl, S. A., Teusink, B., et al. (2015). MetDFBA: incorporating time-resolved metabolomics measurements into dynamic flux balance analysis. *Mol. Biosyst.* 11, 137–145. doi: 10.1039/C4MB00510D

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Nägele, Fürtauer, Nagler, Weiszmann and Weckwerth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.