From Metabolite Concentration to Flux – A Systematic Assessment of Error in Cell Culture Metabolomics

by

Stanislav Sokolenko

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2016

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

For Chapter 2, the author designed the experiments, supervised and assisted with NMR spectra profiling, and performed most of the analysis. Ryan McKay and David Chang assisted with NMR interpretation. Ryan McKay performed the NMR scans. Michael J. Lewis and Ben George performed primary NMR profiling. Eric Blondeel assisted with profiling and manuscript preparation. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 3, the author designed the experiments, supervised and assisted with NMR data collection, and performed most of the analysis. David Chang assisted with NMR interpretation. Eric Blondeel and Ben George performed the NMR scans. Nada Azlah performed primary NMR profiling. Eric Blondeel assisted with manuscript preparation. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 4, the author designed and implemented the algorithms and carried out the data collection. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

For Chapter 5, the author designed and implemented the algorithms and carried out most of the data collection. Marco Quattrociocchi performed primary NMR profiling. Marc G. Aucoin was involved in the conceptualization of the work, oversaw the implementation, aided in the analysis, as well as reviewed and edited the manuscript.

Abstract

The growing availability of genomic, transcriptomic, and metabolomic data has opened the door to the synthesis of multiple levels of information in biological research. As a consequence, there has been a push to analyze biological systems in a comprehensive manner through the integration of their interactions into mathematical models, with the process frequently referred to as "systems biology". Despite the potential for this approach to greatly improve our knowledge of biological systems, the definition of mathematical relationships between different levels of information opens the door to diverse sources of error, requiring precise, unbiased quantification as well as robust validation methods. Failure to account for differences in uncertainty across multiple levels of data analysis may cause errors to drown out any useful outcomes of the synthesis. The application of a systems biology approach has been particularly important in metabolic modeling. There has been a concentrated effort to build models directly from genomic data and to incorporate as much of the metabolome as possible in the analysis. Metabolomic data collection has been expanded through the recent use of hydrogen Nuclear Magnetic Resonance (^{1}H-NMR) spectroscopy for cell culture monitoring. However, the combination of uncertainty from model construction and measurement error from NMR (or other means of metabolomic) analysis complicates data interpretation. This thesis establishes the precision and accuracy

of NMR spectroscopy in the context of cell cultivation while developing a methodology for assessing model error in Metabolic Flux Analysis (MFA).

The analysis of cell culture media via NMR has been made possible by the development of specialized software for the "deconvolution" of complex spectra, however, the process is semi-qualitative. A human "profiler" is required to manually fit idealized peaks from a compound library to an observed spectra, where the quality of fit is often subject to considerable interpretation. Work presented in this thesis establishes baseline accuracy as approximately 2%-10% of the theoretical mean, with a relative standard deviation of 1.5% to 3%. Higher variabilities were associated primarily with profiling error, while lower variabilities were due in part to tube insertion (and the steps leading up to spectra acquisition). Although a human profiler contributed to overall uncertainty, the net impact did not make the deconvolution process prohibitively imprecise. Analysis was then expanded to consider solutions that are more representative of cell culture supernatant. The combination of metabolites at different concentration levels was efficiently represented by a Plackett-Burman experiment. The orthogonality of this design ensured that every level of metabolite concentration was combined with an equal number of high and low concentrations of all other variable metabolites, providing a worst-case scenario for variance estimation. Analysis of media-like mixtures revealed a median error and standard deviation to be approximately 10%, although estimating low metabolite concentrations resulted in a considerable loss of accuracy and precision in the presence of resonance overlap. Furthermore, an iterative regression process identified a number of cases where an increase in the concentration of one metabolite resulted in increased quantification error of another. More importantly, the analysis established a general methodology for estimating

the quantification variability of media-specific metabolite concentrations.

Subsequent application of NMR analysis to time-course data from cell cultivation revealed correlated deviations from calculated trends. Similar deviations were observed for multiple (chemically) unrelated metabolites, amounting to approximately 1%-10% of the metabolite's concentration. The nature of these deviations suggested the cause to be inaccuracies in internal standard addition or quantification, resulting in a skew of all quantified metabolite concentrations within a sample by the same relative amount. Error magnitude was estimated by calculating the median relative deviation from a smoothing fit for all compounds at a give timepoint. A metabolite time-course simulation was developed to determine the frequency and magnitude of such deviations arising from typical measurement error (without added bias from incorrect internal standard addition). Multiple smoothing functions were tested on simulated time-courses and cubic spline regression was found to minimize the median relative deviation from measurement noise to approximately 2.5%. Based on these results, an iterative smoothing correction method was implemented to identify and correct median deviations greater than 2.5%, with both simulation and correction code released as the "metcourse" package for the R programming language.

Finally, a t-test validation method was developed to assess the impact of measurement and model error on MFA, with a Chinese hamster ovary (CHO) cell model chosen as a case study. The standard MFA formulation was recast as a generalized least squares (GLS) problem, with calculated fluxes subject to a t-significance test. NMR data was collected for a CHO cell bioreactor run, with another set of data simulated directly from the model and perturbed by observed measurement error. The frequency of rejected fluxes in the simulated data (free of model error) was attributed to measurement uncertainty alone. The

rejection of fluxes calculated from observed data as non-significant that were not rejected in the simulated data was attributed to a lack of model fit i.e. model error. Applying this method to the observed data revealed a considerable level of error that was not identified by traditional χ^2 validation. Further simulation was carried out to assess the impact of measurement error and model structure, both of which were found to have a dramatic impact on statistical significance and calculation error that has yet to be addressed in the context of MFA.

Acknowledgements

I would like to thank Marc Aucoin for going above and beyond the role of supervisor – it has been reassuring to know that you always have my best interests in mind.

I would also like to thank Raymond Legge, for initially getting me interested in research and the biological aspects of Chemical Engineering; Eric Jervis, for fueling my growing appreciation for numerical methods; Thomas Duever, for teaching me the value of statistical design; Brian Ingalls, for expanding my understanding of biological modeling; Ali Elkamel, for encouraging and supporting my teaching aspirations; and Christine Moresoli, for helping to guide my teaching practice. I am particularly grateful to Thomas Duever, Brian Ingalls, and Ali Elkamel for sitting on my thesis advisory committee as well as Ana Teixeira for serving as an external examiner.

A lot of the research in this thesis would not have happened if not for the collaboration and assistance of David Chang (and Chenomx Inc. more broadly) as well as Ryan McKay (and the National High Field Nuclear Magnetic Resonance Centre); thank you very much for your instruction and support.

While everyone in the Aucoin lab will hold a special place in my heart, I want to thank Eric Blondeel, Sandi Yen, Steffen Schulze, and Mark Bruder in particular for their engaging conversations – which I value whether I agreed with you or not.

Even though I will likely never meet him, I want to thank Hadley Wickham for his work on the R Programming Language. Apart from saving me hours of labour, you have served as a true model for the elegant transformation of data into information.

Finally, I want to thank National Science and Engineering Research Council for funding my research at both undergraduate and graduate levels with the Undergraduate Student Research Awards and the Canada Graduate Scholarships.

Dedication

To the reader.

Table of Contents

List of Figures

List of Tables

Introduction

With sales of $140 billion in 2013 (Walsh 2014), biopharmaceutical production is an important target for bioprocess development. Cell culture expression systems, which are used to manufacture practically all biopharmaceuticals currently available on the market, offer unparalleled potential for monitoring and control when compared to many other biotechnology processes. A high level of process control is particularly important for biopharmaceuticals produced in mammalian cell lines, which are more sensitive to process conditions than bacterial or yeast systems (Schmidt 2004). Beyond just keeping the cells alive, expanding the range of observed process parameters and applying multivariate data analysis has made it possible to drastically increase product yield (Schaub et al. 2012). Moreover, the typical requirements of high product quality and yield are further amplified by the use of biopharmaceuticals like monoclonal antibodies (mAbs) to treat life-threatening conditions such as cancer.

In 2004, the Food and Drug Administration (FDA) issued a guidance on the Process Analytical Technology (PAT) framework for pharmaceutical development (FDA 2004). The document highlighted the importance of process control through the identification of critical quality attributes (CQAs) that impact final product characteristics. The FDA argued against freezing manufacturing practices for the benefit of simplifying regulatory burdens.

Rather, the suggestion is to focus on continued process development and understanding. According to the guidance, "a process is generally considered well understood when (1) all critical sources of variability are identified and explained; (2) variability is managed by the process; and (3) product quality attributes can be accurately and reliability predicted. . . ". This guidance has been interpreted as a push for the greater application of design of experiment (DOE) methodology and multivariate analysis in biopharmaceutical process development (Read et al. 2010).

The focus on greater process monitoring and understanding has coincided with the development of new technologies for genomic, transcriptomic, proteomic, and metabolomic analysis – referred to together as multi-omic technology. These developments have spurred the development of a "systems biology" approach, characterized by the combination of comprehensive multi-omic data collection and its incorporation in descriptive mathematical models (Ideker et al. 2001). An important ambition of this strategy is to identify research areas that may have been overlooked by traditional hypothesis-driven investigation. Despite the potential benefits of adopting a systems biology perspective, important caveats have emerged from dealing with large-scale data collection and analysis. For example, in-depth investigation of microarray data has revealed considerable issues with reproducibility and sensitivity that may not always be considered (Fathallah-Shaykh 2005). The results of modeling may be similarly questionable. Gutenkunst et al. (2007) have found evidence of universal "sloppiness" in dynamic biological models, leading to high levels of uncertainty in parameter estimation.

The recent application of hydrogen nuclear magnetic resonance (^{1}H-NMR) spectroscopy to cell culture monitoring falls directly in line with the systems biology approach. Unlike

^{13}C-NMR, which has been used for targeted metabolic analysis through ^{13}C labeling studies, ^{1}H-NMR can be used to assess the concentrations of all hydrogen-containing small molecules in extracellular or intracellular environments. In contrast to (chromatographic) analytical methods such as high performance liquid chromatography (HPLC) or gas chromatography mass spectrometry (GCMS), NMR uses a single procedure for the uniform detection of all small molecules. The lack of compound-specific sample preparation reduces the possibility of (hypothesis) confirmation bias and falls directly in-line with the discovery-driven science championed by systems biology. However, the lack of chromatographic separation results in the generation of highly overlapping spectra when analyzing complex mixtures, complicating concentration quantification. An increasingly common approach to the peak fitting problem has made use of a human to manually match spectral peaks of interest to idealized ones generated from a library – commonly termed targeted profiling (Weljie et al. 2006). Following use in clinical biofluid analysis and other applications, the methodology has recently expanded to cell culture monitoring (Aranibar and Reily 2014; Bradley et al. 2010; Khoo and Al-Rubeai 2009; Read et al. 2013). However, despite the high possibility of bias or imprecision, there is a serious lack of quantification validation in this context. The validation that has been done has tended to focus on spectra integration (rather than human peak matching) or metabolite identification (rather than concentration quantification) and typically in the context of biofluid analysis. Although ^{1}H-NMR offers great potential for unbiased cell culture process monitoring in the context of biopharmaceutical production, potential sources of quantification variability need to be identified and explained to bring it in line with the recommendations of the PAT framework.

Genomic technology has also seen new applications in recent years. Although genetic modification has practically defined the field of bioengineering, ready access to genomic sequencing is relatively new in some bioprocess contexts. A draft genome of a Chinese hamster ovary (CHO) K1 cell line was only published in 2011 (Xu et al. 2011) despite the commercial importance of CHO cells in biopharmaceutical production. Similarly, the first genome assembly of a *Spodoptera frugiperda*, Sf21 cell line, used in the baculovirus expression vector system (BEVS), was published in 2014 (Kakumani et al.). High quality genomic data has a direct application in rational bioengineering strategies. Moreover, it allows for the generation of metabolic models based on the observed genes. Although metabolic models can take many forms, a common approach focuses on major material flows in the cell to give a broad overview of the metabolic network. The typical approach to developing such models has relied on generic Eukaryotic cell pathways available from bioengineering texts. The growing availability of genomic data has allowed for a more tailored approach of simplifying the metabolic model directly from the genome scale (Quek et al. 2010). However, the process of model simplification, whether from traditional bioengineering texts or genomic data allows for a considerable amount of uncertainty that has yet to be sufficiently explored.

The overall objective of this research project is to address the variability of a metabolic modeling toolchain from the systems biology perspective i.e. integrating error or uncertainty associated with metabolite concentration, flux estimation, and model choice. ^{1}H-NMR metabolomics through the targeted profiling method offers considerable possibilities to bioprocess monitoring and control; however, continued implementation must address potential sources of error. Similarly, the process of model development is not trivial, and

modeling must address the possibility of disagreement between the model and data being due to model error. Model simplification involved in Metabolic Flux Analysis (MFA) makes the analysis of particular interest for validation. The thesis is divided into 5 chapters. Chapter 1 consists of a literature review of cell culture monitoring and modeling via MFA. Chapter 2 establishes a base level of variability for targeted profiling quantification, while Chapter 3 presents a detailed investigation of variability in the context of cell culture media analysis. Chapter 4 addresses bias correction in ^{1}H-NMR data, required for accurate flux estimation, and Chapter 5 presents a generalized least-squares (GLS) framework for MFA validation and assesses the impact of both measurement and model error. Chapters 2–4 take the form of published manuscripts, while Chapter 5 has been formatted for forthcoming submission.

Chapter 1

Background and Literature Review

1.1 Metabolomics in cell culture

Although precise definitions vary, the metabolome is generally taken as the collection of small molecules that participate in cellular metabolism. What constitutes as "small" remains quite vague, with some authors excluding polymeric compounds such as peptides. On the other hand, Khoo and Al-Rubeai (2007) suggest a 1 kDa cut-off that includes polypeptides and compounds that are not direct substrates or products in cellular reactions. The term "metabolome" first appeared around 1998 (Oliver et al. 1998; Tweeddale et al. 1998), and was used to emphasize the holistic analysis of all observable metabolites. As it stands, only a fraction of over 1400 unique metabolites included in recent metabolic models of *Escherichia coli* (Weaver et al. 2014) can be detected. Nonetheless, metabolomics (like all 'omic analyses) is distinguished by the goal of establishing a collective profile of cellular function, rather than focusing on individual components (Lederberg and McCray 2001).

As the chemical environment of the cell, the metabolome represents the phenotype of higher level processes. Evaluation of genomic or proteomic perturbation requires a

measurement of outcome, and observation at the metabolic level has the potential to capture the full impact of small changes in enzyme activity (Dunn and Ellis 2005). This idea has made metabolomic analysis popular across a wide array of applications. Chief among these is the detection of disease biomarkers in biofluids such as urine or plasma (reviewed by Zhang et al. 2012b). Similarly, the metabolic response to drug exposure has also seen considerable study both *in vitro* and *in vivo* (reviewed by Schnackenberg and Beger 2008). Although the study of metabolic response to drug administration was initially coined as metabonomics (Nicholson et al. 1999), the distinction in terminology has largely faded away since then (Cuperlović-Culf et al. 2010). More broadly, metabolomic analysis has also been used to study the impact of environmental toxins (García-Sevillano et al. 2015). Although many developments are shared across the broader field of metabolomics, monitoring of recombinant protein production bioprocesses presents a number of unique challenges and opportunities.

1.1.1 Footprints and fingerprints

In the context of cell culture monitoring, the broader concept of metabolomics can be divided into the observation of extracellular or intracellular metabolites, referred to respectively as footprinting (or exo-metabolomics) and fingerprinting (or endo-metabolomics) (Oldiges et al. 2007). Extracellular metabolite analysis is relatively simple but remains limited to metabolic substrates and excreted products. Sample preparation consists of centrifugation or fast filtration to remove cellular material. Further processing may consist of derivatization or fractionation, depending on the analytical method. Intracellular analysis allows access to the greatest number of metabolites, which can include short-lived

intermediates. However, observation is severely limited by difficulties in sample preparation. To ensure an accurate snapshot of the cellular metabolism, all enzyme activity must be halted as the cells are sampled. This is referred to as the quench step and typically consists of a rapid reduction in cell sample temperature. Metabolites must then be extracted from the cellular debris, which may require membrane disruption along with the selection of an extraction buffer. In many cases, however, this division is not clear and the two steps overlap.

The most common quenching procedure consists of mixing cell culture broth with a larger volume of methanol-water mixture cooled to −40 °C, followed by either fast filtration or centrifugation to eliminate the remains of extracellular fluid. The protocol is generally credited to de Koning and van Dam (1992), who were studying short-lived glycolytic metabolites in yeast cells. Since then, it has been applied to everything from bacteria to mammalian tissue culture, with mixed success. Bolten et al. (2007) compared the use of a 60% methanol mixture at −58 °C with 10 mM HEPES buffer (and centrifugation) to fast vacuum filtration through a nominal pore size of 0.2 µm for a number of Gram-positive and Gram-negative bacteria. The methanol quench resulted in the leakage of more than 60% of many metabolites, quantified by measuring metabolite levels in the supernatant. In contrast, leakage from fast-filtration only impacted Gram-negative species and could be largely eliminated through the use of a 2.6% NaCl solution during filter wash. Methanol-induced leakage was considerably worse for amino acids and TCA cycle intermediates than phosphorylated glycolysis intermediates. Similar leakage problems were reported for different bacteria in the comparison of the cold methanol quench (buffered or unbuffered) to a number of other methods (Shin et al. 2010; Taymaz-Nikerel et al. 2009; Winder et al.

2008). Of the alternatives, recent application of glycerol saline quenching solutions appears quite promising (Link et al. 2008; Villas-Bôas and Bruheim 2007).

Reports of leakage as a result of methanol-quenching in Eukaryotic cells are more conflicted. Working with *Saccharomyces cerevisiae* cells, Villas-Bôas et al. (2005) have observed considerable membrane damage (as assessed by propidium iodide) resulting from a 60% or 75% methanol quench at $-40\,^{\circ}\mathrm{C}$, with a number of amino and organic acids being almost entirely excreted from the cell. Canelas et al. (2008) have found similar issues, but showed dramatic leakage reduction using pure methanol at high ratios of quench solution to sample broth. In contrast, quenching comparisons performed on *Pichia pastoris* showed minimal leakage across practically all methanol concentrations (Carnicer et al. 2012; Tredwell et al. 2011b). Similarly, no issues with leakage could be identified in the methanol-quench of *Aspergillus niger* (Lameiras et al. 2015) or *Penicillium chrysogenum* (de Jonge et al. 2012). While the conflicting reports may be attributed to the different species, there remain a number of non-systematic deviations in the various protocols. For example, Carnicer et al. (2012) used fast filtration in combination with the quench whereas Canelas et al. (2008) used centrifugation.

Systematic comparison of quenching methods for CHO cells have only appeared quite recently. Sellick et al. (2009) tested the use of 60% methanol at $-40\,^{\circ}\mathrm{C}$ with a number of different supplements and found the addition of 0.85% ammonium bicarbonate to give best results with minimal leakage. However, their analysis focused on larger molecular weight metabolites such as ATP and glucose-6-phosphate, which have been previously shown to suffer fewer leakage problems in other cell types (Bolten et al. 2007; Canelas et al. 2008). On the other hand, both Dietmair et al. (2010) and Kronthaler et al. (2012) have shown

considerable damage to the cell membrane from the use of methanol, in contrast to 0 °C water with 0.9 w/v NaCl and PBS respectively. The use of fast filtration for CHO cell quenching has also shown potential (Volmer et al. 2011).

Whereas quenching comparisons have generally focused on methanol-water mixtures and the question of filtration vs. centrifugation, metabolite extraction protocols are much more varied. A number of the more aggressive protocols have survived from the 40s and 50s, including the use of perchloric acid (Hancock 1958), hot water (Gale 1947; Work 1949), or hot ethanol (Fuerst and Wagner 1957). Recent method comparisons have also tended to include pure cold methanol or various combinations of cold methanol, chloroform, acetonitrile, and water (Canelas et al. 2009; Faijes et al. 2007; Maharjan and Ferenci 2003; Sellick et al. 2010; Villas-Bôas et al. 2005; Winder et al. 2008). In general, hot water or ethanol extractions ensure enzyme denaturation, but may cause side-reactions such as amino acid interconversion or protein hydrolysis (Canelas et al. 2009; Dietmair et al. 2010). While the use of cold organic solvents is meant to slow down reactions, some enzymes may continue to function at temperatures as low as −100 °C (Bragger et al. 2000), and Canelas et al. (2009) have suggested that freeze-thaw methods commonly employed with methanol extraction raise the temperature enough for considerable enzyme activity to occur.

As it stands, both quenching and extraction processes are based on compromise. An aggressive quench and rinse step may reduce the turnover of highly energetic compounds such as ATP or the concentration of extracellular compounds, while facilitating the leakage of smaller acids. Similarly, the use of boiling ethanol for extraction may prevent interconversion of extracellular glucose into glucose-6-phosphate but cause changes to amino acid pools. Finally, different extraction solutions will naturally favour some metabolites over

others, which may or may not be observable with various available analytical methods. At the same time, continued developments in rapid (< 1 second) sampling (Canelas et al. 2008; Canelas et al. 2009; Lameiras et al. 2015; Lange et al. 2001) as well as chemical-free quenching using microstructure heat exchangers (Wiendahl et al. 2007) may be able to reduce experimental variability and ease future method comparison.

1.1.2 Applications

The ultimate goal of metabolomic analysis is the integration of the greater metabolome into a systematic understanding of cellular metabolism (Chen et al. 2016; Wang et al. 2006). However, the lack of a truly comprehensive analytical method and the complexity of modeling interactions across multiple levels of metabolic regulation has limited the current scope of metabolomic applications. As it stands, most studies use metabolomics to identify a subset of chemical species that can be linked to suboptimal culture process performance. Cell culture media optimization represents a particularly common target of analysis. The systematic quantification of all media components over the course of a cultivation allows for simple identification of limiting substrates. For example, Read et al. (2013) identified the depletion of aspartate, cysteine, methionine, tryptophan, and tyrosine in the cultivation of murine hybridoma cells producing an IgG antibody. Supplementation of these compounds resulted in 20% increased yield. In a study on IgG producing CHO cells, Sellick et al. (2011) identified the depletion of aspartate, asparagine, glutamate, and pyruvate. Supplementation of these compounds maintained exponential growth for an extra day, increasing overall yield by over 100%. Although the identification of metabolite depletion is not novel in itself, metabolomic methods allow for a much broader scope of

quantification using a single analytical method (Aranibar et al. 2011; Bradley et al. 2010).

Even if no limiting substrates can be identified, metabolomic analysis can nonetheless reveal targets for bioprocess improvement. For example, Mohmad-Saberi et al. (2013) compared CHO cell growth on two different medias and connected consumption of asparagine with improved cell health and the production of ornithine with the onset of apoptosis. Indeed, multiple metabolites produced during the course of cell growth have been found to induce apoptosis in CHO cells (Chong et al. 2011). A number of studies have focused on the overflow of glucose into lactate in CHO cell metabolism. Ma et al. (2009) developed an optimized media that improved maximum cell density and supported a shift from lactate production to consumption. Metabolomic analysis revealed that the extracellular concentrations of metabolites "downstream" of pyruvate and "upstream" of fumarate imitated lactate's time-course profile. Subsequent studies have also connected sorbitol, fructose (Luo et al. 2012), and glycerol (Carinhas et al. 2013) to inefficient glucose processing associated with lactate production. Broader overflow metabolism has also been studied for different microorganisms (Paczia et al. 2012). Ideally, metabolomic analysis can reveal targets for feed optimization or genetic engineering. One example of this process can be seen in the work of Hasunuma et al. (2011), who investigated the negative impact of acetate on the growth of *S. cerevisiae* on lignocellulosic feedstock. Intracellular metabolomic analysis identified the accumulation of pentose phosphate pathway (PPP) intermediates such as sedoheptulose-7-phosphate, ribulose-5-phosphate, ribose-5-phosphate and erythrose-4-phosphate. To deal with the accumulation, new strains were constructed to overexpress either transaldolase (TAL) or transketolase (TKL) and shown to have improved acetate tolerance in the form of improved cell growth and productivity.

The amount of data generated by metabolomic analysis (and its complexity) frequently requires the use of multivariate statistics or multivariate data analysis (MVDA). Key challenges involve identifying related metabolites and connecting metabolic phenotype to cellular function. Among the available techniques (reviewed by Cuperlović-Culf et al. 2010; Rathore et al. 2015; Wang et al. 2006), principal components (PCA) and partial least squares (PLS) regression[1] are some of the most commonly used. PCA is used to map changes in concentrations of many metabolites onto a small set of "principal components" that best explain overall data variability. This process can be used to identify groupings among different samples and propose biological or mechanistic relationship between grouped metabolites. PLS is similar, but identifies a set of principal components that best explain the variability in a set of variables designated as a response. PCA and PLS (as well as related approaches) have been used to discriminate between different microbial strains (Pope et al. 2007), identify gene deletions (Mas et al. 2007), detect contamination (Sue et al. 2011), and study the effect of process parameters such as temperature (Wagstaff et al. 2013) or growth inhibiting additives (Badsha et al. 2016). In one particularly interesting study, Chong et al. (2012) used PLS to link high antibody production in different CHO clones to high levels of NADH and glutathione, suggesting that clone redox/oxidative state may have a particularly important role in protein productivity.

[1]PLS is used as an acronym for "partial least squares" as well as "projection to latent structures". In the context of metabolomics, the two terms are generally used interchangeably.

1.2 Analytical methods

The primary analytical methods of metabolomics are mass spectrometry (MS) and nuclear magnetic resonance (NMR) (Dunn and Ellis 2005; Zhang et al. 2012a). Mass spectrometry is based on sample ionization and observation of the resulting mass-to-charge ratios, which can be related to metabolite identity. Ionization causes many metabolites to fragment and the fragmentation pattern can be included as a component of the analysis. Nuclear magnetic resonance, on the other hand, relies on the fact that certain nuclei will resonate at radio frequencies when placed in a strong magnetic field, with the resonant frequency depending on the nucleus and strength of the magnetic field. Electron density around functional groups can alter the local magnetic field that a nucleus experiences – causing the resonance of nuclei found in different functional groups to appear at different points of a Fourier transformed spectrum. Although both methods have seen extensive use in broader metabolomics applications, cell culture and bioprocess monitoring has tended to favour mass spectrometry. Due to the chemical complexity of the samples involved, a mass spectrometer detector is typically coupled with sample separation using gas or liquid chromatography (GC and LC respectively). Despite the growing popularity of capillary electrophoresis (CE) separation, CE-MS has thus far been limited to more targeted approaches (Monton and Soga 2007). Together, coupled separation-mass spectrometry methods such as GC-MS and LC-MS are referred to as "hyphenated" methods. Despite their general similarity, GC-MS and LC-MS differ considerably in their processing steps as well as the results they are able to generate.

GC-MS offers high resolution and reproducible separation with the added benefit of

standardized electron impact ionization (Gao and Xu 2015; Lei et al. 2011). This allows for relatively simple database generation (Babushok et al. 2007; Kind et al. 2009; Koo et al. 2014). However, GC-MS is mostly limited to volatile and thermally stable metabolites such as short chain alcohols, acids, esters, and hydrocarbons (Lei et al. 2011). While the quantification of such compounds may be sufficient for some microbial applications (Gao and Xu 2015), analysis of other compounds requires derivatization (Chace 2001). LC-MS is able to analyze a much broader collection of metabolites with less sample preparation; however, commonly used electrospray ionization suffers from inconsistent ionization suppression that depends on sample composition (King et al. 2000; Lei et al. 2011). The total number of metabolites that can be identified by GC-MS and LC-MS (and the quality of concentration quantification) will depend on both the specific protocol of chromatographic separation and ionization (as reviewed by Gao and Xu 2015). In some cases, protocol optimization may result in a tradeoff between different metabolites. For example, Danielsson et al. (2010) found that increasing purge vent time and injector temperature in their GC-MS protocol resulted in an overall increase of peak area but considerable decreases of peak areas corresponding to phosphoenolpyruvate, asparagine, histidine, glutamine, and glutamate (in their derivatized form). As the work of van der Werf et al. (2007) highlights, truly comprehensive metabolomic analysis requires the combination of multiple chromatographic methods as well as multiple separation protocols.

In contrast to MS, typical NMR experiments require no sample preparation beyond the addition of an internal standard solution and no separation step (Barding et al. 2012). Furthermore, quantification is independent of many chemical properties such as polarity or hydrophobicity (assuming that a metabolite can stay in solution). These factors make

NMR a near-"universal" method (Leenders et al. 2015), considerably simplifying analysis. But universality comes at the cost of sensitivity. While MS can detect metabolites at nanomolar concentrations (Lei et al. 2011), NMR is constrained by limits of detection in the high micromolar range (Leenders et al. 2015). At the same time, NMR requires up to 500x greater sample volumes. It should be noted, however, that NMR analysis is non-destructive, allowing sample reuse. Whereas hyphenated MS techniques are primarily distinguished by their method of separation, NMR relies on magnet strength and the detection of multiple nuclei to discriminate between different metabolites. Four biologically relevant nuclei can be detected by NMR: ^{1}H, ^{13}C, ^{31}P, and ^{15}N, of which ^{1}H and ^{31}P are found at high natural abundance. Considering the prevalence of ^{1}H in practically all metabolites, ^{1}H-NMR has become the default NMR technique for a wide variety of metabolomic applications (Larive et al. 2015). The abundance of ^{1}H nuclei in metabolomic samples also results in a high degree of resonance (peak) overlap in ^{1}H-NMR spectra. Although characteristic peak shapes can aid the deconvolution process, the primary driver for peak separation is magnet strength. Typically measured as the frequency of ^{1}H resonance (which increases in stronger magnetic fields), 600 MHz instruments and above have become the norm for analyzing complex samples. Apart from increasing magnet strength (which can become prohibitively expensive beyond 1 GHz resonance), it's also possible to perform 2D experiments to identify coupled nuclei. Whether the coupled nuclei are the same (homonuclear experiments e.g. ^{1}H coupled to ^{1}H) or different (heteronuclear experiments e.g. ^{1}H coupled to ^{13}C), 2D experiments spread resonances into two dimensions, thereby reducing overlap and simplifying identification (Leenders et al. 2015).

1.2.1 ^{1}H-NMR targeted profiling

Despite the popularity of 1D-^{1}H-NMR in metabolomic applications, spectral overlap remains a considerable challenge for metabolite quantification, even at high magnetic strengths. Although 2D NMR can be a powerful method for spectral deconvolution, there is a considerable cost in the form of decreased sensitivity and increased experimental time. Recent developments of so-called "ultrafast" 2D methods (as reviewed by Akoka and Giraudeau 2015; Gal and Frydman 2015) and their application to metabolomic analysis (Guennec et al. 2014) are paving the way towards more common use, but thus far, 1D-^{1}H-NMR experiments have remained the norm. The extraction of quantitative data from 1D-^{1}H-NMR spectra can be accomplished through a number of different means. Fundamentally, NMR quantification relies on the principle that the integrals of resonance peaks are proportional to the number of nuclei that make up the resonances (Bharti and Roy 2012). Quantification schemes primarily differ in how the area is estimated. One common approach is to divide the spectra into a series of section or "bins" and calculate the area for each bin, which may be equidistant, specified manually, or set according to an adaptive algorithm (De Meyer et al. 2010; De Meyer et al. 2008). Multivariate analysis like PCA can then be used to identify spectral regions that correlate with a response of interest. Although some resonances can be linked to specific metabolites, other spectral regions will remain convoluted. Another approach is to decompose the spectra into a subset of simpler peaks. For example, "total-line-shape fitting" has been used to convert whole spectra into a series of Gaussian and Lorentzian peaks (Laatikainen et al. 1996; Soininen et al. 2005). Alternatively, spectral features can be extracted based on variations between different samples (Eads et al. 2004; Ochs et al. 1999). Features extracted in this way may happen to corre-

spond to specific metabolite resonances (Eads et al. 2004), but the overall process does not take advantage of metabolite-specific resonance information. The third major approach is to use a spectral library to fit collected data. Although automated methods are becoming increasingly sophisticated (Crockford et al. 2005; Gipson et al. 2006; Gómez et al. 2014; Hao et al. 2012; Mercier et al. 2011; Ravanbakhsh et al. 2015; Zheng et al. 2011), commonly used commercial software requires a human to perform metabolite resonance assignment and fitting.

First introduced by Weljie et al. (2006), "targeted profiling" refers to the process of manually matching computer generated resonance peaks to a spectra as a means of "deconvoluting" resonance overlap – with the functionality provided by NMR Suite software (Chenomx Inc., Edmonton, Canada). The generated curves are drawn from a library of pure compound spectra, collected over a range of pH values (since pH can have a considerable impact on the chemical shift, or position, of the resonances). The library can be used to identify metabolites based on the location of individual resonances, their characteristic shapes, and the correlation of multiple metabolite resonances across the spectra. A human "profiler" chooses metabolites from a shortlist generated by the software to arrive at the best overall fit. Ideally, subtracting the generated resonances from the original spectra should result in a "subtraction line" that is consistently near 0. Comparing the integrals of metabolite resonances identified in this way to the resonance integral of a known quantity of internal standard allows for the accurate quantification of all identified metabolites. However, as resonance integral values can depend on the specific parameters of the NMR experiment, quantification is contingent on experimental data and the database sharing the same NMR pulse sequence.

The relative novelty of the targeted profiling method and the possibility of human error in the quantification raises a number of questions around expected variability. Previous work with single compound solutions or well resolved peaks has suggested that NMR quantification can reach relative standard deviations of approximately 0.5% (Malz and Jancke 2005; Maniara et al. 1998; Wells et al. 2002), even when comparing analysis by different laboratories (Malz and Jancke 2005). However, the method of quantification was limited to simple integration. A number of studies have also examined the analytical reproducibility of highly convoluted biofluid samples in the context of spectral binning (Dumas et al. 2006; Keun et al. 2002b; Parsons et al. 2009; Viant et al. 2009; Ward et al. 2010). While it's possible to conclude that spectra acquisition by NMR is quite robust, less can be said about the method of deconvolution. Two studies have explicitly considered variability associated with human profiling (Slupsky et al. 2007; Tredwell et al. 2011a), however, analysis focused on between-profiler variability. Although coefficients of variation as high as 25% or more were routinely observed, the source of variability remains unclear and a number of questions remain. Is profiling alone to blame, or does underlying spectral variability contribute to human inaccuracy? Furthermore, is there a way to predict variability estimates from solution composition?

1.3 Metabolic flux analysis

Due to the complexity of cellular metabolism – consisting of overlapping systems that incorporate genetic information, transcription/translation levels, and enzyme activities (Lee et al. 2005; Tyo et al. 2007) – the conversion of metabolomic data to metabolic informa-

tion typically requires some form of mathematical modeling. A system-level approach to metabolic engineering that considers all levels of information has been a major goal in the field for some time (Stephanopoulos et al. 1999). Although some studies have begun to collect and analyze combinations of metabolomic, proteomic, and transcriptomic data (frequently referred to as multi-omic) (Al Zaid Siddiquee et al. 2004; Krömer et al. 2004; Wang et al. 2003; Yoon et al. 2003), the incorporation of this data into a single comprehensive model has proven to be challenging (Dauner 2010; Stephanopoulos et al. 2004). The relationship between (and regulation of) gene expression level, protein concentration, and metabolite production has yet to be fully elucidated (Blank and Kuepfer 2010; Carinhas et al. 2012; Kim et al. 2008; Maertens and Vanrolleghem 2010; Shimizu 2009). As such, most modeling efforts have focused on relatively simple stoichiometric balancing (Boghigian et al. 2010), with the calculated intracellular flux seen as the cumulative output of all the regulatory systems (Blank and Kuepfer 2010; Boghigian et al. 2010; Kim et al. 2008; Niklas et al. 2010).

At its core, the stoichiometric model is nothing more than a system of linear equations representing mass or energy balances around the chemical reactions that take place inside a cell. The scope of the model, however, has changed considerably over the years from simple element balances (Stephanopoulos et al. 1998) to full genome-scale reaction networks (Quek et al. 2010). Despite the recent advances, the process of translating genomic information to cellular reactions is still under development. Even the well-studied genomes of *Escherichia coli* and *Saccharomyces cerevisiae* still have approximately 20% of their open reading frames (ORFs) uncharacterized (Dauner 2010) and the development of reaction networks requires a significant amount of curation (Boghigian et al. 2010; Dauner 2010;

Maertens and Vanrolleghem 2010). Furthermore, the large number of reactions that can occur inside a cell make it difficult to calculate metabolite flux through the network as a function of extracellular metabolite concentrations. Thus, where intracellular flux calculation is required, the reaction network must be reduced by removing reactions that are presumed to be irrelevant under particular experimental conditions and combining smaller reactions into pooled pathways (Quek et al. 2010). The calculation of intracellular fluxes from such a reduced, overdetermined stoichiometric model is generally referred to as simple or traditional metabolic flux analysis (MFA). For the purpose of this document, MFA will always refer to traditional overdetermined MFA, unless noted otherwise.

The process of generating an MFA stoichiometric model can be argued as being somewhat *ad hoc*, with model curation guided by (generally well-argued) assumptions about cellular metabolism and culture conditions (Quek et al. 2010). Despite this freedom in model formulation, the effect of model structure on the calculated fluxes is not well studied (Maertens and Vanrolleghem 2010). Furthermore, frequently employed model and data validation techniques have remained largely unchanged since their first introduction (van der Heijden et al. 1994), in contrast to the continued growth of metabolic model size and scope. The accepted validation strategies have to be carefully evaluated given the advances in model formulation, and updated where necessary. Moreover, new techniques are needed to incorporate a model's inherent explanatory power in the interpretation of experimental data.

1.3.1 Formulation

Although the development of MFA has been a gradual process, its current mathematical formulation is generally attributed to Stephanopoulos et al. (1998). The basis of metabolic flux analysis takes the form of a mass balance around intracellular metabolite concentrations:

$$\frac{dC_{met}}{dt} = r_{met} - \mu C_{met} \tag{1.1}$$

where $\frac{dC_{met}}{dt}$ is the "change in concentration" vector for intracellular metabolites, r_{met} is the "net rates of formation" vector (generation and consumption), and μC_{met} is the "biomass dilution" term, which represents the amount of metabolite leaving the cell upon cell division. From Equation 1.1, two major simplifications are made: assumption of steady state and incorporation of biomass growth as a consumption reaction, resulting in:

$$0 = r_{met} = Sv \tag{1.2}$$

where S is the stoichiometric matrix and v is the vector of fluxes that correspond to reactions defined by S. The steady state assumption follows from the observation that metabolite pools are often much smaller than the metabolite production and consumption fluxes and can be ignored. A similar assumption is applied to pool dilution as a result of cell division. Even when a metabolic pool is split in half, the amount of metabolite needed to reestablish the pool is small in comparison to the flux passing through it.

The Sv matrix can be be separated into $S_c v_c + S_o v_o$, where c stands for calculated flux and o for observed flux. The manner of solution for the unknown fluxes depends on

the form of S_c. Taking S_c to be an n-by-m matrix, three cases are possible: i) n < m, ii) n > m, iii) n = m. In the first case, the matrix is undetermined and must be solved through optimization – the most common result when using the full scope of a genomic model. In the second, the matrix is overdetermined and must be solved using a pseudo-inverse, which is equivalent to least-squares regression[2]. The third case is quite rare in practice and can be accomplished with a simple matrix inverse.

1.3.2 Strategies for estimating unknown fluxes

The differences in the solution of underdetermined and overdetermined systems influences the types of questions that these models can answer. The most important distinction is that the analysis of underdetermined systems, or flux balance analysis (FBA) as it is frequently referred to, defines a feasible solution space for the unknown reaction fluxes based on the metabolic network, while MFA calculates a single flux estimate (Toya et al. 2011). The number of reactions that occur inside a cell is generally far greater than then number of observed metabolites (Quek et al. 2010). FBA takes advantage of detailed metabolic network data and known flux constraints to restrict the possible solution space of the unknown fluxes (Kauffman et al. 2003). The calculation of a single flux estimate requires the assumption of a metabolic objective, such as the maximization of growth rate or the maximization of glucose consumption efficiency (among others) (Schuetz et al. 2007). By specifying the objective as an optimization criteria, a single flux estimate can be calculated by linear (or quadratic) programming. This approach is typically used to predict the effects of specific gene insertions or deletions (Kim et al. 2008).

[2]Overdetermined MFA is covered in greater detail in Chapter 5.

While FBA models can offer a very detailed picture of a microorganism's metabolism, they also pose considerable challenges. FBA models require the specification of each enzymatic reaction (Wiechert et al. 2001) and are typically used on well-curated genome-scale metabolic models (Lewis et al. 2012). Detailed models will naturally result in more precise feasible solution spaces. Unfortunately, animal cell lines relevant to cell culture applications frequently lack the necessary genomic data (Carinhas et al. 2012), with mouse hybridoma cells a notable exception (Quek et al. 2010; Selvarasu et al. 2009). Even the well-studied *S. cerevisiae* has approximately 20% of its protein-coding genes uncharacterized and the formulation of the various available databases has required a considerable amount of manual confirmation and curation (Christie et al. 2009). Furthermore, the relation between the presence of a gene sequence and enzymatic activity is not always obvious (Maertens and Vanrolleghem 2010). A combined transcriptomic-metabolomic modeling study of *E. coli* has revealed the existence of redundant gene expression where no flux was observed (Shlomi et al. 2007). Meanwhile, a study of lysine-producing *Corynebacterium glutamicum* metabolism suggested that while the expression of some genes appears tightly coupled to metabolic fluxes, others can remain practically constant despite considerable changes in metabolic flux (Krömer et al. 2004). Even where genomic data is readily available, its translation into accurate metabolic models can be limited.

To avoid the analysis of underdetermined systems, the formulation of an overdetermined system can be achieved with the combination of model simplification (Quek et al. 2010) and/or the collection of extra information in the form of isotopomer analysis (^{13}C-MFA) (Wiechert et al. 2001). The latter approach takes into account the carbon distribution of site-specific enzyme reaction products by using ^{13}C-labeled substrate. Unlike FBA and

traditional MFA, this method is able to calculate unique intracellular fluxes of parallel and bi-directional reactions (Iwatani et al. 2008; Wiechert and Nöh 2005). On the other hand, the introduction of a labeled substrate adds a considerable amount of complexity in the form of labeling transients, as isotopic equilibrium is achieved in both intracellular pools and biomass components (Wiechert and Nöh 2005). To avoid these transients, experiments must be performed in continuous culture - in contrast to batch-mode operation frequently used in industrial production (Wiechert et al. 2001). Semi-continuous cultures have also been used for this purpose (Sheikholeslami et al. 2013), but the high cost of labeled substrate can make high volume experiments prohibitive (Wiechert et al. 2001). While more advanced, non-stationary ^{13}C-MFA approaches have the ability to overcome continuous-culture requirements by directly considering the transient profiles of isotopic labels, new complications arise in the form of intracellular pool quantification and the need for very rapid sampling (Nöh and Wiechert 2011). Accurate quantification of intracellular metabolite pools is particularly problematic for Eukaryotes, as some metabolites may have distinct pools between mitochondria and cytosol, subject to different enzymes (Nöh and Wiechert 2011).

Although both ^{13}C-MFA and FBA offer significant advantages over the traditional MFA approach, they also pose considerable challenges. The use of complex models requires information or techniques that may not be currently available for mammalian or insect cell cultures. On the other hand, simpler, mass balance approaches such as traditional MFA can still be used to gain a better understanding of cellular metabolism. Such models continue to see use for the metabolic analysis of mammalian cell lines (Niklas et al. 2012; Niklas et al. 2011; Nyberg et al. 1999; Priesnitz et al. 2012; Quek et al. 2010; Xing et al.

2011) and have recently been applied to insect cell cultures as well (Bernal et al. 2009; Drugmand 2007). A number of extensions to traditional MFA have also been developed. Goudar et al. (2006) implemented a quasi real-time strategy where MFA calculation is performed as data is collected during a perfusion reactor cultivation. More recently, a dynamic MFA (or DMFA) approach has been proposed to integrate flux estimation from metabolite concentration trends and stoichiometric flux modeling (Leighty and Antoniewicz 2011). Whereas the original formulation was limited to piecewise linear trends (Leighty and Antoniewicz 2011), Martínez et al. (2015) extended the approach to use B-spline curve fitting. Ongoing development and continued implementation have kept traditional MFA relevant despite the existence of more sophisticated methods.

1.3.3 Current validation methods

Model and data validation have been an important component of MFA for some time (Stephanopoulos et al. 1998). The basis of validation was established by van der Heijden et al. (1994) and revolves around the idea of "gross error detection". The basis for this analysis stems from the fact that measurement error will result in a difference between some hypothetically true value of measured fluxes and those that are observed. If such deviations stem from random measurement noise, they can be assumed to follow a normal distribution. Thus, a χ^2-test can be used to determine the presence of significant outliers signifying the presence of a gross error[3]. If no such errors are identified then all observed values must contain only random noise, which can be "balanced" by finding an estimate for the observed values that minimizes the sum of squared errors for the observations (Stephanopoulos et al. 1998). This idea of "balancing" a set of observations based on a stoichiometric model has

been used extensively as a form of data reconciliation around chemical reactors to account for observation error (Madron et al. 1977). A traditional application of this approach to cell culture has been to make use of known element balances to correct observed values before calculating culture parameters such as growth yields (Solomon et al. 1982).

The χ^2-test and the accompanying observed value correction have been established as the default validation method for MFA models. Recent work continues to use this test to determine model and observation accuracy (Bernal et al. 2009; Niklas et al. 2012; Niklas et al. 2011; Nyberg et al. 1999; Priesnitz et al. 2012; Quek et al. 2010; Xing et al. 2011). However, today's models are much larger and more complex than the ones used for test development (van der Heijden et al. 1994). Furthermore, the initial argument for observed value correction, to use a well established empirical model for the correction of relatively questionable data (Wang and Stephanopoulos 1983), has remained unchanged despite greater availability of high-throughput data and the growing complexity of model formulation.

Complementary forms of model validation also consider error sensitivity (Niklas et al. 2012; Niklas et al. 2011; Nyberg et al. 1999; Priesnitz et al. 2012; Xing et al. 2011). The calculation of stoichiometry matrix condition number[4] has long been established as one method to predict the effect of observation error on calculated fluxes (Vallino and Stephanopoulos 1990). Condition numbers smaller than 100 are generally seen as reasonable, whereas those greater than 1000 suggest serious issues in the model where even small experimental errors could be magnified to unacceptable levels. Minimizing condition numbers has also been suggested as an objective function for deciding which fluxes need to be

[3]Mathematical details are presented in Chapter 5.

observed (Savinell and Palsson 1992). More recently, Goudar et al. (2009) have proposed the use of a normalized sensitivity matrix as a more specific measure of error effect. As opposed to using a single estimate of error magnification in the form of condition number, this method calculates the effect of all observed variables for each calculated flux. The researchers found that while large calculated fluxes could tolerate observation error reasonably well, calculated fluxes of lower magnitude were very susceptible to even relatively small changes in observed fluxes of high volume.

[4]The condition number of a matrix A is defined as $\kappa(A) = ||A|| \cdot ||A^{-1}||$. Given the equation $Ax = b$, $\kappa(A)$ represents how sensitive the solution of x is to error in b when using A^{-1}. In loose terms, $\kappa(A)$ serves as an upper boundary of the scaling factor that relates a relative change in b to a relative change in x (Cheney and Kincaid 2007).

Chapter 2

Establishing the Base Level Variance of Targeted Profiling

As discussed in Chapter 1, cell culture monitoring represents a relatively new application for targeted profiling methods. While a number of studies have considered variability due to human error (Slupsky et al. 2007; Tredwell et al. 2011a), the context was primarily biofluid analysis. The differences between biofluid and cell culture samples have a conflicting impact on the quality of metabolite quantification. The relative simplicity of cultured organisms in comparison to the *Homo sapien* (or even *Mus musculus* used in toxicological studies) leads to much simpler metabolic phenotypes and fewer challenges in metabolite identification. The smaller number of metabolites leads to less spectral convolution and background "chemical noise". On the other hand, differences in cell culture phenotypes may be biologically relevant at very small concentration and are unlikely to be as drastic as in biofluid samples, where a disease or metabolic disturbance can have a pronounced impact on metabolic phenotype. To get a better idea of how useful targeted profiling quantification would be for cell culture analysis, it was necessary to establish a baseline of quantification variability.

The overall process of extracting a metabolite concentration value from an NMR scan using the targeted profiling method can be broken down into three steps – tube insertion and shimming, acquisition of the spectrum, and profiling. The position of the NMR tube in the sample holder influences the orientation of the sample within the instrument's magnetic field and the receiver coil used for spectra acquisition. Not every part of the tube may experience the exact same magnetic field, with potential inhomogeneity caused by the design of the magnet itself, certain materials used in internal instrumentation, small variations in the thickness of the sample tube, or the presence of paramagnetic particles in the sample. The correction of field inhomogeneity is referred to as "shimming" and consists of passing low levels of current through special coils around the sample in order to generate magnetic fields capable of balancing the deviations. Failure to correct field inhomogeneity typically results in a broadening of individual chemical resonances (as molecules will resonate at frequencies dictated by the local magnetic fields they experience).

A basic ^{1}H-NMR scan consists of exciting all protons in a sample with a single strong radio signal and recording the resulting radio wave emissions as excited protons drop to a lower energy state. This output signal contains a mixture of sine waves with varying frequencies and amplitudes that slowly decay over time (an example of this output can be seen in Figure 2.1a). A Fourier transform converts this information into a series of peaks (as shown in Figure 2.1b), with peak location indicative of resonance frequency, i.e., the electron environment of the proton, and peak area corresponding to the abundance of that proton. As most metabolites contain multiple protons in different electron environments, more than one peak "cluster" is generally observed per metabolite. Although the ideal peak shape should take the form of a Lorentzian function, interactions between neighbouring

protons can result in much more complex shapes. If a metabolite resonance is well resolved, its area can be calculated through simple integration. However, most complex mixtures feature a considerable degree of overlap, requiring peak separation. The targeted profiling method provides the user with a library of ideal spectra that can be overlaid through a "drag-and-drop" type interface to match observed results. An illustration of this process can be seen in Figure 2.2. Regardless of how peak area is calculated, comparing the area of a metabolite resonance to the area of an internal standard resonance allows absolute concentration quantification.

Given the qualitative nature of profiling and the potential for errors at all stages of the process, a number of questions had to be considered. First, what is the minimum expected variability for a relatively simple mixture? Second, is baseline variability the product of human error alone, or are other factors involved? Third, what is the impact of small differences between the excitation pulse sequence used during NMR spectra acquisition and library generation i.e. how robust is the quantification? While answering these questions was important for future cell culture work, the results were judged to be relevant for biofluid analysis as well. To broaden potential interest in the results, we partnered with Ryan McKay from the National High Field Nuclear Magnetic Resonance Centre (NANUC) as well as Chenomx Inc. and the results of the investigation were published in the journal Metabolomics (Sokolenko et al. 2013). The manuscript is presented below in its original form. Online resources available with the publication are provided in Appendix A.

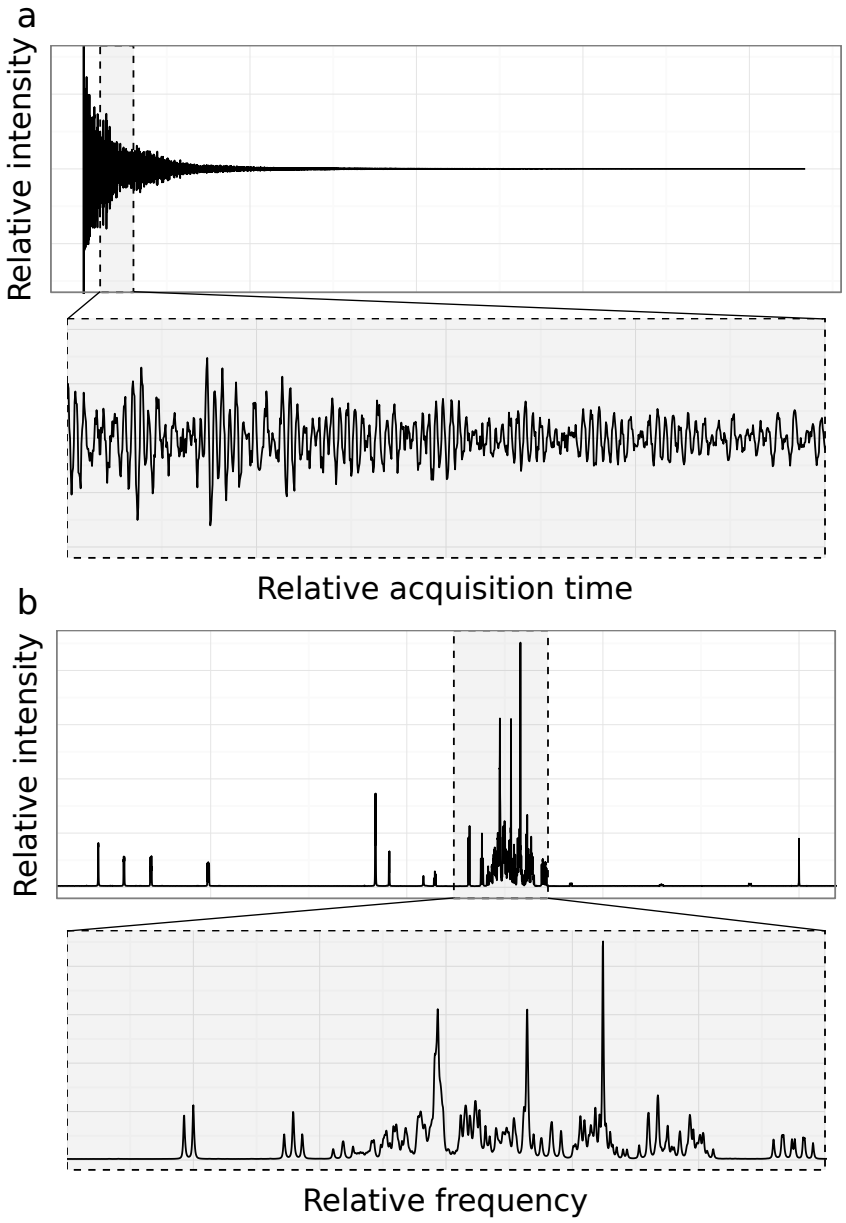


Figure 2.1: **a)** Example of resonance data observed using NMR and **b)** the corresponding Fourier transformed spectrum.

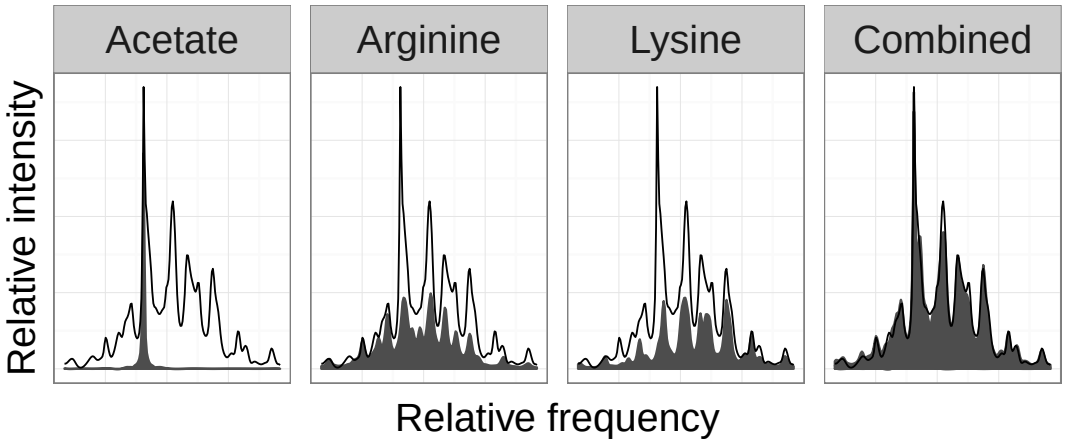


Figure 2.2: An illustration of targeted profiling with three metabolite peaks overlaid to match the observed spectrum.

Understanding the variability of compound quantification from targeted profiling metabolomics of 1D-^{1}H-NMR spectra in synthetic mixtures and urine with additional insights on choice of pulse sequences and robotic sampling

Stanislav Sokolenko[1], Ryan McKay[2], Eric J. M. Blondeel[1], Michael J. Lewis[3], David Chang[3], Ben George[1], Marc G. Aucoin[1]

[1] Department of Chemical Engineering, Waterloo Institute for Nanotechnology, University of Waterloo. [2] National High Field Nuclear Magnetic Resonance Centre (NANUC) Department of Chemistry, University of Alberta. [3] Chenomx Inc., Edomonton, Alberta.

2.1 Abstract

The growing use of 'targeted profiling' approaches for the deconvolution of 1D-^{1}H-NMR spectra by comparison to a pure compound library has created a need for an in-depth characterization of quantification variability that is beyond what is currently available in the literature. In this study, we explore the underlying source of quantification variability (tube insertion, spectra acquisition, and profiling) as well as a number of other factors, such as temporal consistency of repeated NMR scans, human consistency in repeated profiles, and human versus machine sampling. We also look at the effect of different pulse sequences on the differences between acquired spectra and the peak reference library. Two sample types were considered for this work – a synthetic five compound mixture as well as human urine. The result is a comprehensive examination of 1D-^{1}H-NMR quantification error. Our investigation into variability sources revealed that apart from profiling, sample insertion and/or shimming can play a significant role in final quantification, a finding that is equally applicable to all integration-based methods of quantification. Both sources of error were also found to have temporal relationships, with bias identified as a function of both scan and profiling order, reinforcing the need for randomization in scanning and profiling. As well as presenting a practical estimate of variability in human urine samples, we have uncovered a considerable amount of complexity in underlying NMR variability that will hopefully serve as impetus for future exploration in this area.

2.2 Introduction

The study of small molecules originating from endogenous and exogenous metabolism has developed rapidly over the past several years. Routine identification and quantitation of metabolites is performed on a number of human bodily fluids including, but not limited to, whole blood, serum, plasma, cerebrospinal fluid, and urine. Mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) are the predominant analytical techniques used to this end, although rarely used in concert (undoubtedly due to cost restraints). MS has the advantage of rapid individual component identification and inherent sensitivity. While NMR has no theoretical sensitivity limit due to the ability to add repeated scans, data collection much below the micromolar concentration level quickly becomes impractical for metabolic samples in terms of required acquisition time. Biofluid analysis is further limited by considerable levels of peak overlap. On the other hand, NMR has the advantage of being able to acquire not only qualitative (composition) but also quantitative data on intact biological samples without the chromatographic separation needed in most MS experiments. Combined together, the two techniques offer unparalleled identification capabilities of a range of concentrations, and cover a broad range of source compounds (i.e. types of solubility, molecular size, functional groups etc.).

Despite the fact that all NMR applications are based on the same physio-chemical quantum mechanical principles, experimentation and data processing vary considerably in the results they can generate. The most obvious division is between one or multi-dimensional experiments (1D-NMR and 2D-NMR respectively). As it stands now, 1D-NMR techniques have the advantage of considerably shorter acquisition time and better sensitivity,

although new developments for 2D-NMR analysis are quickly bridging this gap (McKenzie et al. 2011). 1D-NMR can be further subdivided into how the generated spectral data is used. It is possible to use spectral information directly for sample classification (as in disease biomarker identification), such as with the commonly used combination of spectral binning followed by principal components analysis (PCA). Another approach is to convert spectral information to metabolite concentrations before classification, frequently referred to as (targeted) profiling (Weljie et al. 2006). The application of spectral binning for the purpose of disease biomarker identification through discriminant analysis is extremely powerful, but only scratches the surface of the information held within a spectrum. Spectral deconvolution allows the profiling of individual compounds and provides researchers further information towards deciphering cellular function. A number of software packages have been developed to aid a human profiler in the task of spectra deconvolution (as reviewed in (Dieterle et al. 2011)), including Bio-Rad Laboratories Know-It-All (Bio-Rad Laboratories, Philadelphia, PA),Bruker Biospin Biofluid Spectra Base (Bruker Biospin GmbH, Rheinstetten), and Chenomx NMR Suite (Edmonton, Alberta), the software considered in this work. More recently, fully automated deconvolution algorithms have also begun to appear (Hao et al. 2012; Mercier et al. 2011; Zheng et al. 2011). The development of new techniques requires a new look at measurement reproducibility as previously reported results (such as those from binning spectral data) cannot be applied.

With the increasing number of NMR-based metabolomics studies of human diseases (Nagrath et al. 2011; Psychogios et al. 2011; Wei 2011; Zulyniak and Mutch 2011), there is a need for a comprehensive examination of NMR quantification variability, similar to the work done by Zelena et al. (2009) for UPLC-MS. A fundamental question that has

yet to be answered is how much variability exists, and where does the variability in an NMR quantified concentration come from, given that there are numerous potential sources of error e.g. tube positioning and shimming, spectra acquisition, metabolite assignment and deconvolution, to name a few. More generally, what are the error bounds that one can expect in a typical biofluid sample following quantitation with software such as the Chenomx NMR Suite – and how should these bounds be interpreted? From a commercial point of view, an in-depth error analysis is key for personalized medicine applications of biofluid analysis, especially because small spectral differences are being encouraged as patentable intellectual property. These differences can be viewed as biomarkers for disease identification, disease severity evaluation, and optimal treatment Nagrath et al. (2011). While there have been studies regarding sample preparation, sample handling, and how the subsequent data should be analyzed (Beckonert et al. 2007; Begley et al. 2009; Beltran et al. 2012; Foxall et al. 1993; Garde et al. 2004; Kohl et al. 2012; Lauridsen et al. 2007; Saude et al. 2007; Saude and Sykes 2007; Sukumaran et al. 2009; Warrack et al. 2009; White et al. 2010), few studies have addressed the very serious need for 'error bars' on the raw information, lacking from virtually all metabolic data published to date. Where spectral reproducibility has been addressed directly, the context has often been the integration of relatively unconvoluted peaks, more typical of analytical chemistry and limited to only select biofluid constituents (Malz and Jancke 2005; Maniara et al. 1998; Viant et al. 2009; Ward et al. 2010; Wells et al. 2002), or binning approaches for convoluted spectra (Dumas et al. 2006; Keun et al. 2002a; Parsons et al. 2009; Viant et al. 2009; Ward et al. 2010). The reproducibility of software assisted profiling of biofluid spectra has so far largely focused on 'between-person' comparisons (Slupsky et al. 2007; Tredwell et al. 2011b).

The issue of error quantification is further confounded by the effect of solvent suppression. As the most convenient biological NMR samples are aqueous in nature, successful NMR analysis requires a 10^5–10^6–fold reduction of the inherent ^{1}H signal from the solvent in order to have any chance of viewing the solute signals (i.e. metabolites). Fortunately there has been a wealth of excellent development in this regard (for reviews, see Liu and Mao 1999; McKay 2009; Price 1999 and references therein). While the number of possible pulse sequences makes the selection of a standard technique quite daunting, one has nonetheless emerged. The first increment of the 2D [^{1}H,^{1}H]-nuclear Overhauser spectroscopy (NOESY) sequence has a simple setup (with few parameters to optimize), robust solvent suppression, and simple hardware requirements. For example, the sequence does not require pulse field gradients, waveform generators, or even multi-channel/nuclei consoles and probes. Essentially, even the simplest instrumentation will likely be able to run this pulse sequence with little trouble and the experimental protocol comes standard with the operating software from most major instrument vendors. While the pulse sequence itself is capable of multi-dimensional data acquisition, only the first dimension is collected. The NOESY utilizes both presaturation, as well as inherent solvent suppression capabilities from the delays and phase cycling used in the sequence (Mckay 2011). The simplicity, water suppression, low hardware requirements, and the use by major research groups made the 1D-NOESY the sequence of choice for the Chenomx software package (Chenomx Inc, Edmonton, Alberta). While protocols have been established for the collection of highly reproducible data, for particular samples (such as highly dilute urine) the residual solvent signal can be unexpectedly difficult to suppress, requiring the exploration of pulse sequence alterations or entirely new sequences with improved water suppression.

To address this issue, the effect of using a non-standard version of the NOESY and a different pulse sequence on the Chenomx NMR Suite analysis was investigated. The comparison was performed on a simple five-compound mixture as well as a human urine sample. In the process, we present a comprehensive examination of the precision that can be expected from metabolite quantification, ranging from the ideal scenario of only somewhat convoluted spectra to the more realistic case of heavily convoluted samples i.e. a great deal of spectral overlap. A simple mixture was then used to investigate the source of variability inherent to NMR analysis, considering the impact of tube insertion and shimming, spectral acquisition, and lastly the profiling itself. A fully randomized experimental design also allowed the investigation of time/order related effects (such as measurement drift) at the sample insertion/shimming and profiling levels. In this manuscript, we have gone a long way towards elucidating the error contribution of each aspect of the experiment to suggest protocols for monitoring, and possible avenues for improving metabolomics data for all interested groups.

2.3 Materials and methods

2.3.1 Sample composition

Simple mixture

A five-compound, synthetic mixture consisting of glucose (Aldrich, ChEBI:17634), glycine (Aldrich, ChEBI:15428), maltose (Sigma, ChEBI:17306), sucrose (Sigma, ChEBI:17992) and nicotinic acid (Aldrich, ChEBI:15940) was used to probe the base-level accuracy and

variability of the 1D-^{1}H-NMR data. Each compound was dissolved to a randomly chosen concentration close to 10 mM. Compound concentration has been previously observed not to influence quantification precision, as long as the concentration remains above the limit of detection. A relatively high concentration was chosen to limit the impact of weighing and volumetric errors. Briefly, dry compound stocks were dissolved in 70 mL of 30 mM PO4 buffer. The pH of the solution was adjusted to 6.9, and 8 mL of an internal standard solution was added. The internal standard consisted of 99.9% D_2O with 4.6861 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, see Markley et al. 1998) serving as an internal chemical shift reference and 0.2% w/v sodium azide (NaN3) to inhibit bacterial growth (Chenomx Inc., Edmonton, Alberta). The mixture was brought up to a final volume of 80 mL with 30 mM PO_4 buffer, to produce a solution containing 9.92 mM glycine, 10.30 mM maltose, 10.58 mM sucrose, 8.61 mM glucose, 10.18 mM nicotinic acid, and 0.4686 mM DSS. The solution was stored frozen at $-80\,^{\circ}$C before being separated into 21 tubes for spectra acquisition and profiling.

Urine

A single human urine sample was selected randomly from a healthy donor. The sample was stored at $-80\,^{\circ}$C until scanned. The sample was filtered through a 0.22 µm filter (VWR Scientific, Mississauga, ON), and transferred to a Wilmad 535P-8 inch NMR tube (Wilmad Glass Inc, Vineland NJ) for spectra acquisition and profiling.

2.3.2 NMR

Pulse sequences

Three pulse sequences (hereafter referred to as 1D-NO-ESY_s1a4, 1D-NOESY_s2a3, and grd-NOESY_s1a4) were investigated to determine their relative contribution to measurable error and to test whether a database created with a single pulse sequence (Chenomx Inc.) would still yield reliable qualitative and quantitative information when other pulse sequences were utilized. The nomenclature indicates the initial presaturation and acquisition times, e.g. in 1D-NOESY_s1a4, s1 corresponds to $\sim$ 1 s of presaturation (10 ms of relaxation delay with 990 ms of active water suppression at an induced field strength/gammaB_1 of $\sim$ 80 Hz, see Mao and Chen 1996) while the term a4 corresponds to an acquisition period of 4 s. The recommendations when using the Chenomx Software database for biomarker identification require a 10 ms delay, 990 ms presaturation period, 100 ms mixing time (with saturation), and a 4 s acquisition period for a total recycle delay of 5 s (i.e. relaxation can occur freely during this entire period, but not during mixing as the evolution is perturbed by the final pulses). 1D-NOESY_s2a3 is the same general pulse sequence as 1D-NOESY_s1a4 with the presaturation period extended by 1 s and the acquisition time shortened by the same amount, yielding an identical total relaxation delay. Finally, grd-NOESY_s1a4 was a pulse sequence developed for improved solvent suppression (McKay et al. 2009). Where the sequence names appear in figures, a shortened form is used – grd_s1a4, met_s1a4, and met_s2a3, where the 'grd' stands for grd_NOESY and 'met' refers to the frequently used name of 'metnoesy' in reference to the 1D-NOESY sequence. Additional information on the pulse sequences used can be found in Appendix A.

Sample acquisition

800 MHz NMR The single urine sample (600 µL) was repeatedly scanned at 25 °C on an Inova 18.8T (800 MHz) NMR Spectrometer (Varian Inc. now Agilent, California) using VNMRJ 2.2D controller software on a RHEL 5.3 operating system. The instrument was equipped with a 5 mm HCN Z-pulsed field gradient cryogenically cooled probe. The magnetic field was manually optimized until a line shape for the DSS methyl peak (see above) of less than 1.0, 12, and 20 Hz at 50, 0.55 and 0.11% respectively, was achieved. Deuterium lock was maintained on the 10% D_2O added as part of the internal standard. The sweep width was 9599.2 Hz; presaturation ($\sim$ 80 Hz gammaB_1) time was 990 ms with a 10 ms relaxation delay; mixing time was 100 ms (also with saturation power); and all pulse lengths were manually calibrated to optimal 90° excitation. The acquisition period was 4 s (76794 real plus imaginary points) with 32 scans collected per experiment after 4 steady state scans to establish equilibrium. Digital oversampling of a factor of 20 was utilized during acquisition, and the extraneous points automatically removed prior to storage.

600 MHz NMR Simple mixture NMR samples were transferred to 5 mm diameter Wilmad 512P-4-inch tubes with robotic caps. Sample volumes were made up to precisely 700 µL for dependable automated optimization of the spectrometer magnetic field (i.e. shimming). Samples were run on a Varian (now Agilent Inc., California) VNMRS console equipped with a Varian 768AS automatic sampling handling system controlled by VNMRJ 2.2C software in a RHEL 4u3 environment. Spectrometer lock was established on the 10% D_2O and maintained throughout the experiments. Automated Z-axis pulsed field gradient

shimming was performed on each sample. Line shape of the methyl group resonance of DSS was evaluated and was only deemed acceptable when the width of the peak was <1.0, 12 and 20 Hz for the 50, 0.55 and 0.11% peak heights, respectively. The effective 90° pulse width was automatically calibrated using nutation theory (Wu and Otting 2005) along with the optimal carrier position for water saturation. A simple presaturation pulse sequence (Figure A.1a) was used for sample parameter optimization (i.e. spectrometer locking, shimming, pulse width, and carrier position). Optimized parameters were then transferred to the 1D-NOESY (Figure A.1b) and grd-NOESY (Figure A.1c) experiments, respectively.

Spectra were collected with 32 transients and 4 steady state scans. Manual handling of the sample included sample length measurement and position of the sample in the spinner for centering in the receiver NMR coil. Robotic sample handling was carried out using the Varian 768AS sample handling system. The 1D-NOESY pulse sequence is shown in Figure A.1b and the phase cycling has been described in detail previously (Mckay 2011).

Spectra processing and profiling

All spectra were processed and profiled using NMR Suite 7.1 (Chenomx Inc., Edmonton, Alberta) by a single profiler. Spectrum phasing and baseline correction were applied automatically. Reference deconvolution to remove line asymmetry (Morris et al. 1997) and spectrum profiling were done manually. The pre-profiling steps of spectrum phasing, baseline correction, and reference deconvolution (together, termed processing) were performed in 1 day as a single stage of analysis for all of the samples before a second stage of profiling carried over a 4 day period. Urine profiles were reviewed and any ambiguous compound

assignments were removed. For more information about the theoretical principles behind the software, consult http://www.chenomx.com/ and the work of Weljie et al. (2006).

2.3.3 Experimental design

Simple mixture

The basis of the simple mixture design was to (a) estimate underlying sources of 1D-^{1}H-NMR quantification variability, (b) compare the effect of human and machine sampling on quantification and (c) compare quantification accuracy and variability from the use of different pulse sequences. A single experimental run was carried out to this end on 21 prepared sample tubes, divided into two experimental sets – the first with human sample handling and the second with machine manipulation. In the first set, each tube was positioned in the spinner (centering the sample in the receiver coil space), inserted, and scanned with each of the three pulse sequences (1D-NOESY_s1a4 followed by 1D-NOESY_s2a3 and grd-NOESY_s1a4) before removal. In the second set, the above procedure was repeated on the same 21 samples with the modification that each of three spectra acquisitions were done using only the 1D-NOESY_s1a4 sequence. Spectra from the first set were processed and quantified once; spectra from the second set were duplicated and randomized before processing/profiling. The first set was used to compare the quantification of three different pulse sequences. The latter set allowed statistical separation between tube insertion/shimming, spectra acquisition, and processing/profiling for the estimation of variability from these different sources.

Urine

A single human urine sample tube was scanned eleven times using each of the three pulse sequences described in 2.3.1.

2.3.4 Statistical analysis

All statistical analysis was carried out using the R programming language (R Core Team 2012). In particular, plyr (Wickham 2011) and ggplot2 (Wickham 2009) packages were extensively used for data manipulation and presentation, respectively. All mean comparisons were performed using two-tailed t-tests and variance comparisons using F tests, both at the 95% confidence level. The term 'significant' is generally reserved for use in the statistical sense, unless it's clear that no numerical comparison was performed. In some figures, a smoothing trend line was added using locally weighted scatterplot smoothing (LOESS). This line is not meant to signify a statistically significant trend; it is used for visual aid only.

2.4 Results and discussion

2.4.1 Design justification

Pulse sequences

While the 1D-NOESY is one of, if not the most used pulse sequence for metabolomic biofluids analysis via NMR (Beckonert et al. 2007), we have experienced difficulty with receiver

overloading issues on highly dilute samples. This prompted part of the initial exploration to see if modifying the length of internal delays, without altering the total experiment time, would perturb analysis results. To this end we performed the standard 1D-NOESY_s1a4 pulse sequence, then increased the presaturation period from 990 to 1990 ms with a concomitant decrease of the acquisition time from 4 s down to 3 s. It is crucial to maintain the total experiment length otherwise T_1 relaxation issues can come into play and distort the relative peak intensities expected by the Chenomx software, thereby disrupting the carefully determined database of metabolite peak intensities for quantitation. We also included the grd-NOESY_s1a4 sequence, reported previously to have superior water suppression capabilities (McKay et al. 2009) but not certified for use on metabolomics samples. We wished to ascertain the identification and quantitation capabilities of the modified 1D-NOESY_s1a4 (i.e. 1D-NOESY_s2a3) and grd-NOESY against the standard 1D-NOESY_s1a4 sequence.

Compound selection

The selected compound mixture was restricted to five compounds to ensure that fundamental NMR quantification variability would not be masked by excess convolution. To underline the applicability of the simple mixture to deconvolution based quantification (of the type frequently carried out on more complex samples like urine), these compounds nevertheless featured a small degree of known convolution. While this peak overlap did not significantly impact quantification using Chenomx NMR Suite software, it did rule out simple integration-based quantification techniques for all compounds except nicotinate, the only unconvoluted compound. Metabolites were chosen to cover a broad range of chemical

functionalities (including a simple amino acid, a vitamin and a number of sugars) and molecular weights (from 75 to 342 g/mol), while remaining relevant to biofluids (and cell culture applications). Three sugars were used to provide a reasonable amount of convolution and served to compare the properties of compounds having similar functionality and structure. Although the use of only three compounds meant that variability could not be generalized, differences could be used to disprove a generalization.

Sample number

The choice of sample number was based on the desire to maximize statistical power for variance comparison while maintaining a reasonable total number of samples. The power of a test is defined based on a hypothetical "real" variance that one is trying to observe. Sets of 21 samples allowed variance estimates with 20 degrees of freedom, equating to a 2% probability of incorrectly determining that the variances were not different (rejecting the null hypothesis) when one variance is, in reality, two times the other. While a twofold difference may be considered as too lenient, we believe that it is an acceptable sensitivity threshold as smaller variance differences can still be correctly identified (albeit with lower probability). The end result of using sample sets much larger than the frequently encountered triplicates was a greater confidence in negative results i.e. failure to see a difference between two variances was much more likely to mean that they are actually very similar. A somewhat smaller sample size of 11 was chosen for the analysis of urine due to practical time limitations of profiling such a complex mixture.

2.4.2 Pulse sequence comparison

One of the primary goals of this study was to determine the effect that the choice of NMR parameters, or pulse sequence, would have on quantification variability. As inherently greater variability in more complicated (and convoluted) mixtures have the possibility of obscuring or distorting this effect, we opted to focus the analysis on a representative simple mixture to establish the points of comparison between the three pulse sequences. Real urine data was reserved as a case study of their application.

Mixture mean concentration

The use of each tested pulse sequences resulted in mean concentration values that were comparable in their difference between observed and theoretical values (Table 2.1). Means of compound concentrations generated using the grd-NOESY_s1a4 sequence were closest to theoretical concentrations for glucose, maltose, and nicotinate, while both 1D-NOESY sequences resulted in closer agreement to theoretical values for sucrose (glycine concentration was similarly close for all pulse sequences). Figure 2.3a presents a graphical interpretation of these results. While no one sequence can be said to be most accurate, a definite relationship was observed between the three sequences. The grd-NOESY_s1a4 sequence gave consistently lower concentrations than either of the 1D-NOESY ones. Between 1D-NOESY_s1a4 and 1D-NOESY_s2a3, the latter generally resulted in slightly higher concentrations with sucrose concentration as the only exception (highlighted in Figure 2.3b).

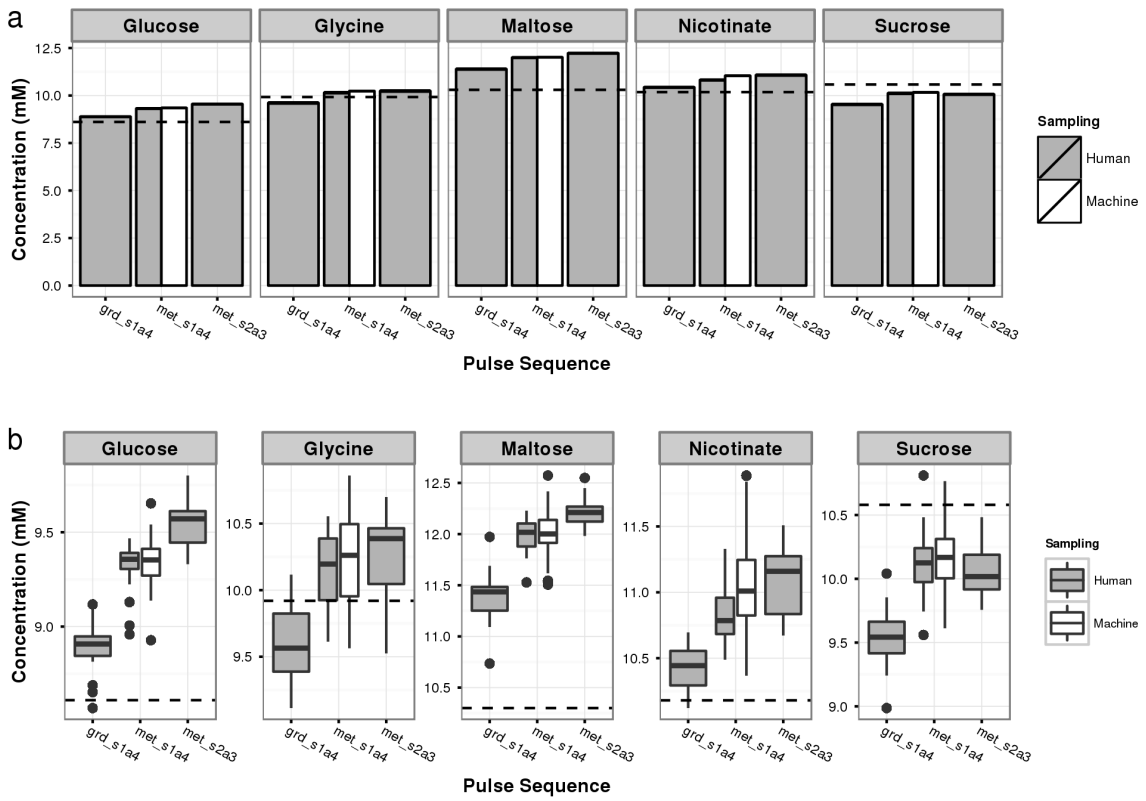


Figure 2.3: **a)** Comparison of mean observed concentration values and b observed concentration distributions between three pulse sequences. Theoretical compound concentrations are represented by dashed lines. In **b)**, thick lines represent data medians, box edges represent 25th and 75th percentiles, whiskers extend out to the furthest point up to a maximum of 1.5 times the inter-quartile range beyond the box edges, and dots represent outliers falling more than 1.5 times the interquartile range away from the box edges. 95% confidence intervals around the mean values are on the same magnitude as individual line widths and were excluded from the plots.

Table 2.1: Mean observed concentrations and their deviations from theoretical values for the 5-compound simple mixture sample.

Sequence	grd-NOESY_s1a4		1D-NOESY_s1a4				1D-NOESY_s2a3	
Sampling	Human		Machine		Human		Human	
	Mean concentration (mM)	Difference from theoretical value (%)	Mean concentration (mM)	Difference from theoretical value (%)	Mean concentration (mM)	Difference from theoretical value (%)	Mean concentration (mM)	Difference from theoretical value (%)
Glucose	8.89	−0.96	9.35	4.18	9.32	3.83	9.55	6.44
Glycine	9.61	−4.14	10.24	2.11	10.15	1.23	10.24	2.11
Maltose	11.39	4.00	12.01	9.66	12.00	9.54	12.22	11.61
Nicotinate	10.43	0.42	11.04	6.28	10.82	4.13	11.07	6.62
Sucrose	9.53	−10.34	10.16	−4.39	10.11	−4.86	10.06	−5.36

Mixture variability

Apart from influencing observation accuracy, the choice of pulse sequence was also hypothesized to have an effect on quantification variability. A graphical comparison of variability is presented in Figure 2.4 as distributions of individual observations from their mean values (which we have termed deviations in the figure). Overall, the distributions from the three sequences were quite similar, with only the concentration of nicotinate showing signs of bimodality when using the 1D-NOESY_s2a3 sequence. Pairwise variance comparisons using F-tests identified only the variance of maltose concentrations to be different for the three pulse sequences, with grd-NOESY_s1a4 being the odd one out. It should be noted that the observations for maltose using the grd-NOESY_s1a4 pulse sequence featured two prominent outliers (Figure 2.4) that skewed the variance estimation considerably.

The results of the experiment extend beyond just a pulse sequence comparison. More generally, we have established an approximate 1.5–3.5% base estimate for standard devia-

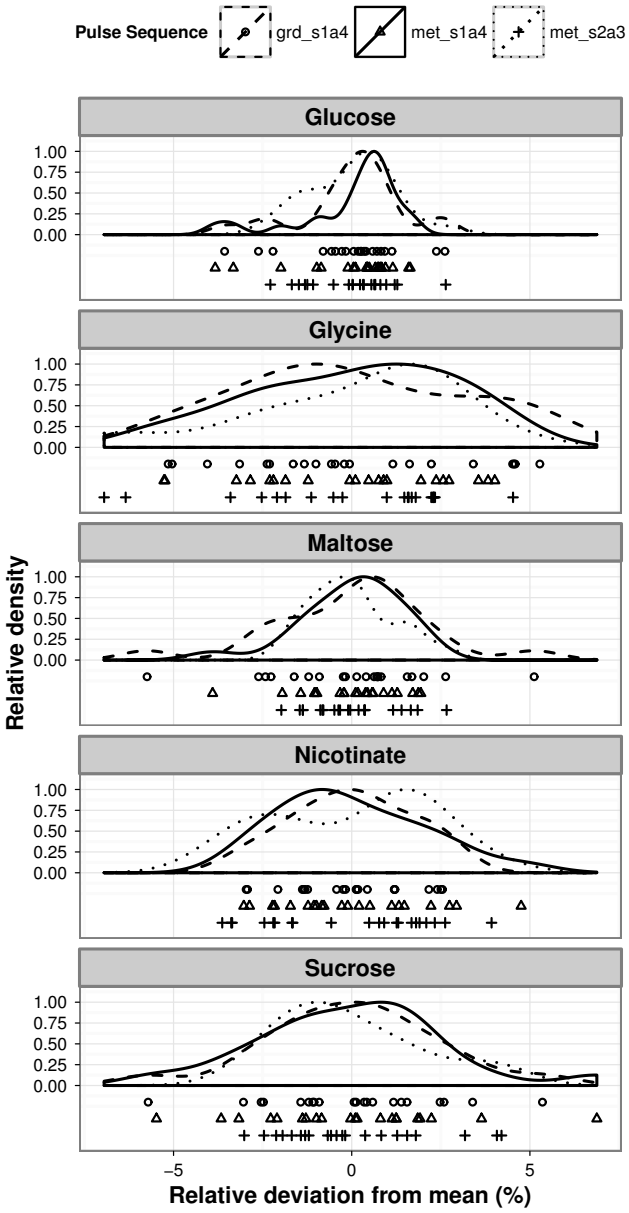


Figure 2.4: Comparison of observation distributions around mean values for each compound and pulse sequence tested. Distribution curves were generated via kernel density estimates. Individual observations are indicated by points below the distributions.

tion when profiling with Chenomx software (Table 2.2). While profiling resulted in higher standard deviation than previously reported integration methods (Malz and Jancke 2005), this level of variability is exceptional when considering that it is unaffected by a considerable amount of convolution – where integration techniques will falter. The concentration of nicotinate, the only unconvoluted compound had a higher standard deviation than both maltose and glucose (species with spectral overlap) when using the 1D-NOESY_s1a4 pulse sequence (Table 2.2). Therefore, convolution does not imply greater variability as long as NMR clusters are not entirely obscured by other compounds, which can drastically reduce measurement precision. While no trends were expected to be observed with only five compounds in the standard, the relative difference in the variability of sucrose and maltose, two relatively similar compounds, suggests the idea that functionality and molecular weight are unlikely to act as predictors for measurement variability. For applications where twofold differences in standard deviation can be problematic, we would therefore recommend that generic variability estimates not be used.

Table 2.2: Observed standard deviations for the 5-compound simple mixture sample, human-sampled.

Sequence	grd-NOESY_s1a4		1D-NOESY_s1a4		1D-NOESY_s2a3	
	Standard deviation (mM)	Relative standard deviation (%)	Standard deviation (mM)	Relative standard deviation (%)	Standard deviation (mM)	Relative standard deviation (%)
Glucose	0.13	1.45	0.14	1.46	0.11	1.18
Glycine	0.31	3.21	0.29	2.81	0.30	2.97
Maltose	0.26	2.25	0.17	1.40	0.15	1.21
Nicotinate	0.18	1.68	0.22	2.05	0.25	2.28
Sucrose	0.23	2.45	0.27	2.70	0.21	2.06

Urine sample data variability

The variability range presented in Section 2.4.2 is, naturally, an ideal case. NMR spectra of real urine samples have long been known to suffer from significant levels of 'chemical noise', i.e. unidentified compounds that mask the baseline and other known compounds (Nicholson and Wilson 1989). While the deconvolution process is still able to quantify the general concentration of such compounds, high quantification precision cannot be expected from all of them.

As this study did not consider the problem of compound mis-identification, a conservative approach was taken during profiling. Of the 304 biologically and clinically relevant compounds found in the Chenomx library, 56 compounds were unique in their ability to match the observed spectra and have all been previously observed in urine samples. It should be noted that if compound mis-identification did occur, it would not influence the goal of assessing quantification precision. Approximately 30 more compounds from the Chenomx compounds database could have been profiled, but their presence in the samples was deemed ambiguous and would have required confirmation by other analytical methods, as recommended by Tredwell et al. (2011a). While conservative, the compound set did encompass a greater collection of compounds than previously considered in urine variability studies employing Chenomx software – 9 by Slupsky et al. (2007), 24 by Saude et al. (2007), and 37 by Tredwell et al. (2011a). Of the 56 identified compounds, 18 were found to have relative standard deviations below the previously identified maximum of 3.5%. Indeed, 9 of the compounds actually had a relative standard deviation smaller than the 1.5% minimum of the simple mixture set, with the lowest equal to $\sim 0.5\%$ and comparable to integration of unconvoluted spectra (Malz and Jancke 2005). Creatinine was among the most precisely

quantified compounds, suggesting that the common practice of normalizing concentration to that of creatinine does not impose a significant loss of observation precision at the individual sample level. 16 more compounds had a relative standard deviation in the range of 3.5–10%, while the remaining 22 compounds had a relative standard deviation higher than 10% (with some reaching the 50–70% range).

In 5 replicated profiles of a single urine spectra, (Tredwell et al. 2011a) reported that 32 out of 37 compounds had a relative standard deviation (referred to as coefficient of variance in their work) of under 10% (as compared to 34 of 56) and a median relative standard deviation of 2.4%. While the number of compounds in the stated range of precision was comparable, we did not observe as low of a median. However, as will be described in Section 2.4.3, replicated profiling of the same acquired spectra does not capture the full extent of variability with this method. The reported inter-profiler variability (Tredwell et al. 2011a), however, is much closer to the range of variability observed in this work, suggesting that the two are of similar magnitude. A full list of concentration means and standard deviations for all quantified compounds (including the compounds left out of the analysis above due to identification ambiguity) can be found in Appendix A.

To add further context to the urine sample results, observed mean and standard deviation values of compounds clearly identified using the Chenomx Inc. software were compared to concentration ranges reported in literature (Wishart et al. 2007) and compiled in the Human Metabolome Database (HMDB) (Table 2.3). In most cases, the results indicated that observation variability is much less than the variability encountered in urine samples from healthy adults (5%). Only rarely did the fraction increase as high as 20% (Table 2.3). Thus, where profiled compounds can be clearly identified, the observed concentration error

of compounds with highly overlapping spectra is still likely to be only a fraction of overall variability in normal, healthy individuals. A similar comparison for compounds whose identity was less certain, but still profiled, is given in Table A.2 in Appendix A.

Table 2.3: Observed mean and standard deviations of compounds unambiguously identified in urine sample, normalized to creatinine concentration (µmol/mmol creatinine units), as compared to literature values taken from the Human Metabolome Database (HMDB), version 2.5. HMDB concentrations were limited to urine from normal, healthy adults (18+) reported in the database in µmol/(mmol creatinine). Information from the indicated number of sources was compiled and a single grand range presented as overall minimum and maximum. Where the observed mean did not deviate more than 25 % from the minimum–maximum range found in literature, the observed standard deviation was expressed as a percent of the range span to illustrate the proportional differences between observation variability and normal variation between healthy patients.

Compounds	Observed mean (µmol/mmol)	Observed standard deviation		HMDB Literature		
		(µmol/mmol)	(% of HMDB range)	Min (µmol/mmol)	Max (µmol/mmol)	Number of sources
Urea	4226.45	46.61	0.12	6.58	38 812.40	10
Hippurate	611.79	8.46	0.93	19.30	932.66	22
Citrate	222.62	1.85	0.33	46.87	599.90	21
Trigonelline	122.42	0.96	0.92	5.50	109.30	3
Trimethylamine N-oxide	84.51	0.46	0.09	0	509.60	11
Glycine	63.12	0.94	NA	NA	NA	0
Formate	55.75	0.65	0.33	0.65	195.63	14
Glycolate	49.32	3.63	2.97	0.04	122.10	14
3-Indoxylsulfate	47.67	2.42	3.77	0.66	64.80	10
Ethanolamine	38.83	2.52	4.92	4.91	56.20	12
Glucose	32.47	2.23	2.01	0	111.07	13
Pyroglutamate	28.57	7.06	27.37	6.80	32.60	6
Dimethylamine	24.36	0.63	1.57	19.20	59.20	14
Creatine	23.61	1.42	NA	NA	NA	0
cis-Aconitate	22.79	3.31	3.57	2.70	95.30	7

Continued on next page

Table 2.3 – continued from previous page

Compounds	Observed mean (µmol/ mmol)	Observed standard deviation (µmol/ mmol)	Observed standard deviation (% of HMDB range)	HMDB Literature Min (µmol/ mmol)	HMDB Literature Max (µmol/ mmol)	HMDB Literature Number of sources
1,3-Dimethylurate	18.82	0.28	NA	1.34	10.10	5
Taurine	18.28	3.24	1.31	4.00	251.60	12
Malonate	17.67	1.26	NA	0.60	9.40	3
3-Aminoisobutyrate	17.02	0.67	NA	NA	NA	0
Glutamine	16.51	5.41	7.85	9.00	77.90	15
Succinate	13.89	8.93	27.05	0.30	33.30	23
Alanine	12.34	0.46	0.94	7.10	56.15	23
Carnitine	12.08	1.60	5.96	0.70	27.50	11
Histidine	11.71	3.21	1.64	17.10	212.81	20
Betaine	11.70	0.22	0.25	2.70	92.70	11
Xanthine	11.66	1.08	NA	2.20	3.75	2
π-Methylhistidine	11.62	4.87	7.33	2.80	69.27	10
Acetate	11.50	0.26	0.25	0	106.00	7
τ-Methylhistidine	11.27	2.04	1.32	0	153.77	26
4-Hydroxyphenylacetate	10.63	0.33	NA	NA	NA	0
Phenylalanine	10.16	1.49	9.02	1.65	18.17	21
Methylamine	9.04	0.30	1.94	1.33	16.70	6
Asparagine	7.72	5.57	11.13	3.00	53.01	21
Dimethyl sulfone	7.46	0.14	0.28	1.30	48.90	4
Sucrose	7.30	0.41	2.11	0	19.50	4
3-Methylxanthine	7.28	0.41	2.51	1.40	17.80	5
trans-Aconitate	6.90	1.59	6.84	1.80	25.10	3
Lactate	5.94	0.54	0.12	3.50	444.29	25
2-Hydroxyisobutyrate	5.93	0.12	2.96	1.70	5.90	5
3-Hydroxyisovalerate	5.93	1.06	4.85	3.20	25.00	7
Tyramine	5.56	0.78	NA	0.20	0.78	5
1,6-Anhydro-β-D-glucose	5.44	0.79	2.90	2.40	29.60	2
Tyrosine	5.29	0.50	1.39	2.57	38.66	20
Ethanol	5.28	2.01	NA	NA	NA	0
Lysine	5.21	2.48	3.39	2.11	75.29	16
Methanol	5.11	0.09	0.09	9.60	116.60	4
Hypoxanthine	5.09	0.24	1.01	0	24.10	8
N,N-Dimethylglycine	4.69	1.72	16.37	0.70	11.20	6
Acetone	4.50	0.14	0.17	0.20	86.14	9
1-Methylnicotinamide	3.99	0.39	2.65	0.24	15.00	12

Continued on next page

Table 2.3 – continued from previous page

Compounds	Observed mean (µmol/ mmol)	Observed standard deviation (µmol/ mmol)	Observed standard deviation (% of HMDB range)	HMDB Literature Min (µmol/ mmol)	HMDB Literature Max (µmol/ mmol)	HMDB Literature Number of sources
N-Acetylaspartate	3.06	1.57	17.71	0	8.84	10
Trimethylamine	2.31	0.97	5.08	0.30	19.40	4
Leucine	2.14	0.76	4.27	1.20	19.07	16
Valine	1.91	0.14	0.58	1.10	25.62	24
Isobutyrate	0.56	0.09	0.07	1.20	122.00	5

In most cases, the general range of precision could have been identified a priori in a qualitative manner based on the profiling process alone. Compounds with unconvoluted NMR peaks (most frequently found in the aromatic range, ($\sim$ 6.5–8.5 ppm) or peaks that stood out from other resonances could be quantified with similar precision as the compounds in the simple mixture. While the convolution level was generally similar for these compounds, their exact standard deviation would still vary from compound to compound in the range of $\sim$ 0.5–3.5%. More severe cases of convolution resulted in a marked loss of precision, with the lowest precision (30–70% relative standard deviation) observed for compounds found almost entirely within regions of unprofilable 'chemical noise'. It has been our experience, however, that the levels of background noise and local convolution can vary considerably between different samples, making it impossible to pinpoint general precision levels as a function of the compounds observed. In lieu of full replication as a concrete estimate of variability, it would be advisable for researchers to apply a contextual filter in the presentation and interpretation of their results. Uncertainty during the profiling process will most likely result in uncertainty of the final quantification.

Pulse sequence comparison for urine sample data

Given the large number of compounds identified in urine, comparing the effect of pulse sequences on individual compound quantification, although possibly necessary, is tedious and challenging. Furthermore, it was highly likely that the effect of pulse sequences could be masked by the inherent variability of profiling (as recorded by the relative standard deviation). We, therefore, investigated a subset of compounds with small standard deviations ($< 5\%$).

After excluding the highly variable compounds, differences between the means resulting from the pulse sequences were plotted as distributions (Figure 2.5). Overall, the 1D-NOESY_s2a3 showed less deviation from the 1D-NOESY_s1a4 observations than did the grd-NOESY_s1a4. Moreover, the deviations in 1D-NOESY_s2a3 appeared more uniform. A closer look at the results from the grd-NOESY_s1a4 pulse sequence revealed that the five points in the -10–0% difference range corresponded to 2-hydroxyisobutyrate, acetate, alanine, formate, and trigonelline – easily profilable compounds with clusters on the edges of the spectra. All but one of the compounds in the -20 to 10% difference range had clusters closer to the water peak. The fact that compounds with clusters on the edges of the spectra experienced less quantification underestimation suggested that the primary factor for the difference was water suppression bleaching. This effect could also be seen by looking at the clusters of trigonelline, which happen to be generally free of background noise and found across the 9.8–4.4 ppm region. The 4.4 ppm cluster was indeed considerably lower than the ideal curve generated from the Chenomx library, as compared to the 9.8 and 8.8 ppm clusters. Final confirmation was achieved by binning the spectral data (normalizing to total spectral area) and comparing the differences between average sums as a function of chemical

shift (Figure 2.6). Once again, grd-NOESY_s1a4 showed a pronounced underestimation near the water peak when compared to 1D-NOESY_s1a4 results. While 1D-NOESY_s2a3 sequence also showed some difference from the 1D-NOESY_s1a4, the trend was much less prominent. Performing the binning normalized to the internal standard also revealed that grd-NOESY_s1a4 had a smaller total spectral area in proportion to the DSS. The overall phenomena of underestimation, when using the grd-NOESY_s1a4 sequence, could be the result of a relative overestimation of the reference standard and underestimation in the proximity of the water peak.

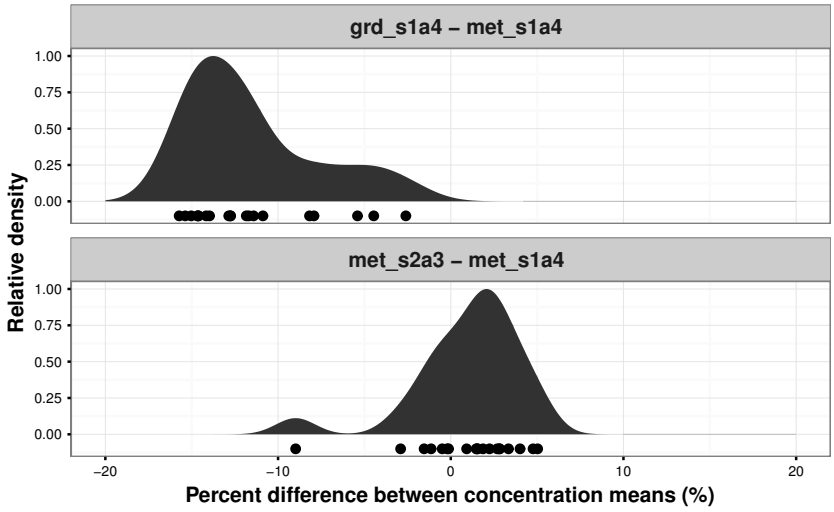


Figure 2.5: Distribution of differences of mean compound concentrations obtained with different pulse sequences. Only compounds with relative standard deviations less than 5% were considered in this analysis.

The results from the urine data expand on and reinforce the simple standard comparison. The five compounds used in the simple standard had at least some of their clusters

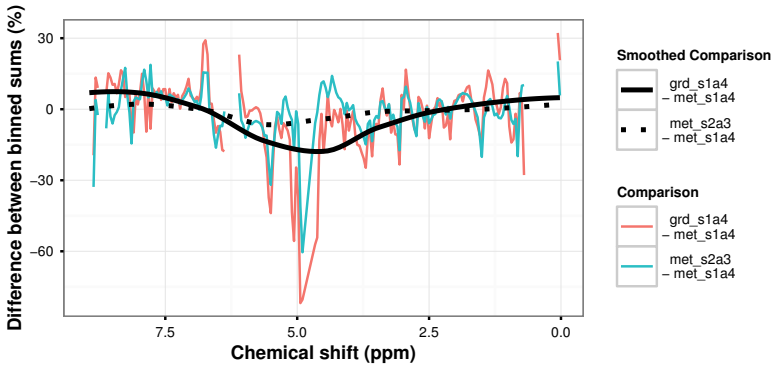


Figure 2.6: Averaged percent difference in binned spectral data between grd-NOESY_s1a4 and 1D-NOESY_s1a4 as well as 1D-NOESY_s2a3 and 1D-NOESY_s1a4 as a function of chemical shift (thin coloured lines) with LOESS trend lines added for visual aid (thick black lines). The area under the spectra was divided into bins 0.04 ppm wide (with the 4.68–4.88 ppm water region excluded) and averaged across the 11 replicates for each of the pulse sequences. Excessive noise was removed by filtering all bins that accounted for less than 0.0001 of the total area.

relatively far away from the water peak and were present at much higher concentrations than those of typical urine constituents. The suppression effect was therefore less noticeable and the human profiler was able to balance out some of the discrepancies by referencing other clusters that were generally available. The final underestimation ($\sim$ 4%) was in the same general range as that of 2-hydroxyisobutyrate, acetate, alanine, formate, and trigonelline in the urine data. Many of the compounds found in urine, however, had only one or two readily profilable clusters. Their underestimation would further depend on their relative proximity to the water peak (anywhere between 2 and 6 ppm). Although the grd-NOESY_s1a4 resulted in a greater mean deviation from the 1D-NOESY_s1a4 than did the 1D-NOESY_s2a3, this deviation was also more predictable.

A few practical implications can be drawn from these results. The use of pulse sequences other than 1D-NOESY_s1a4 with Chenomx software can result in statistically significant changes in observed mean compound concentration values. Where the method of water suppression is not altered, these changes are likely to be minimal ($\pm$5% as in the use of 1D-NOESY_s2a3). The introduction of a different water suppression technique, such as with the grd-NOESY_s1a4 sequence, however, can result in changes as a function of chemical shift. Hypothetically, it may be possible to incorporate water suppression effects into peak fitting software to correct for the underestimation; however, such developments lie outside the scope of this work. Results from simple standard analysis suggested that quantification variability remained relatively independent from the choice of pulse sequence. Comparison of urine samples gave a similar impression, although inherent measurement variability made the comparison less precise. This similarity in measurement variability bodes well for applications such as disease biomarker identification, where relative changes

in compound concentration may be more important than absolute accuracy. Our results suggest that using pulse sequences other than 1D-NOESY_s1a4 in this fashion would be no less effective.

2.4.3 Underlying NMR variability

Variance source breakdown

With the simple mixture set, we also used a hierarchical (or nested) statistical design with analysis of variance (ANOVA) to split the sources of quantification variability into three component parts—tube insertion/shim, spectra acquisition, spectra processing/profiling. This was done to determine the fraction of total variability that can be attributed to human profiling as compared to the inherent variability of NMR analysis, with the goal of contextualizing observed variability and improving quantification strategies.

A breakdown of observed variability (in the form of variance) can be seen in Figure 2.7. Similar to the rest of the results presented thus far, there are considerable differences among the five compounds. While profiling was found to be the only significant source of variability for glycine and sucrose concentrations, it made up less than half of the observed variability of glucose and nicotinate concentrations. Of the three sources considered, spectra acquisitions had the least impact on net variability, suggesting that it is the most reproducible aspect of the process. The same could not be said for tube insertion/shimming, in the cases where it was possible to separate it from profiling. The most direct application of these results pertains to replication strategies. It would appear that re-profiling acquired spectra to get a more consistent result is not guaranteed to have a significant effect and

may create a false impression of sample consistency when a considerable portion of the variability can originate from tube insertion/shimming. Due to the low variability associated with sample acquisition, the same can be said for performing two spectra acquisitions and profiling each of the resulting spectra. The only recommended replication strategy for improving quantification precision is complete, randomized sample reinsertion for replicated scanning. Although our findings could not differentiate between the effects of tube insertion and shimming, both are factors that are not frequently considered in the topic of quantification precision and may warrant closer investigation.

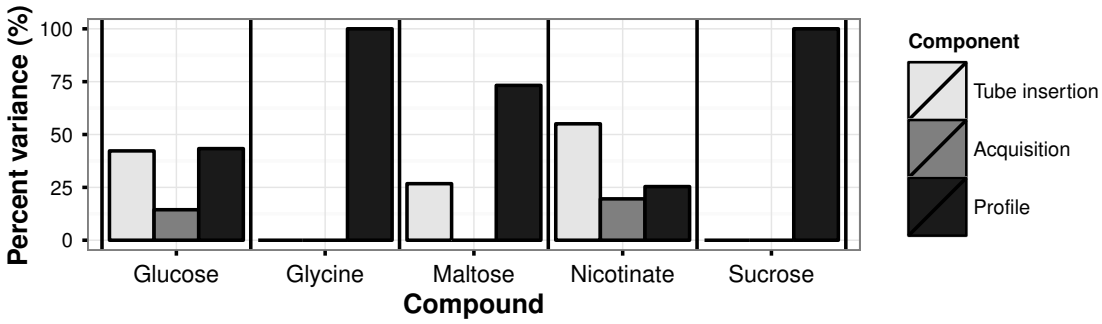


Figure 2.7: Comparison of variance sources as a percentage of total variance for each compound. Component variance values were estimated from an ANOVA of hierarchical design data.

Comparison to integration quantification methods

With the inclusion of a single compound with entirely unambiguous clusters (nicotinate), it was possible to confirm whether variability corresponding to tube insertion was related

to the area of the NMR peaks (standardized to DSS). While it was not surprising to find the actual integration calculation had zero associated variability, the results also revealed that tube insertion/shim was indeed the greatest source of variability with the acquisition step accounting for no more than 5% of peak cluster area variability. This suggests that the recommended replication strategy is equally applicable for quantification via integration. More interesting was the comparison of overall distributions from integration and human profiling (Figure 2.8). Not only did the distributions associated with the four integrated peak cluster vary considerably in their shapes, but all featured pronounced multimodality. Taking the average of all the areas did not correct the multimodality, with the resulting distribution appearing similar to the distribution of the 8.3 ppm peak. Although human profiling had an undoubtedly larger quantification range, the observation distribution was more robust, with a greater proportion of observations clustering around the overall mean.

Profiling variability

As the profiling was carried in a random order, independent of other experimental conditions, it was possible to assess if any profile-order-dependent trends were present in the data. In particular, it was sought to determine if a human profiler would go through a 'burn-in' period of initially high variability before converging on a long-term mean with lower variability. A plot of observation deviations from overall means as a function of profiling order can be seen in Figure 2.9a. 1D-NOESY_s1a4 data from the first experimental set (with human sampling) was added as a control to guarantee that observed trends are indeed the result of profiling order, independent of other conditions. The most prominent profiling-related trend was observed for glycine, resulting in up to 10% differences between-

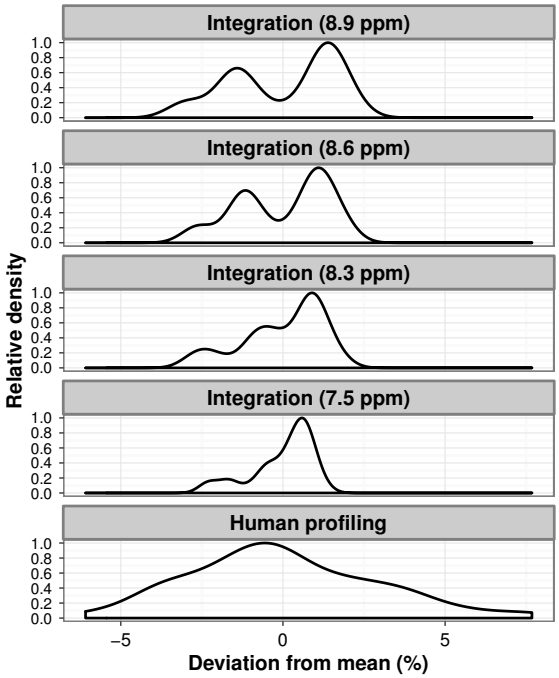


Figure 2.8: Comparison of observation distributions around mean values for human profiling and integration quantification of nicotinate. Integration was performed on each of the clusters associated with nicotinate and normalized by the area of the DSS peak. Distribution curves were generated via kernel density estimates.

quantification performed at different times. Three different profiling interpretations could be observed with relatively rapid shifts between them occurring in the middle of day 2 as well as between days 3 and 4. The most likely explanation is that glycine NMR spectra consists of only a single peak, making it impossible to validate the quantification if there is a slight difference in shape between the spectra and the ideal peak generated from the Chenomx database.

Daily statistics from Figure 2.9b reveal trends that are not immediately obvious from Figure 2.9a, such as the fact that all standard deviations except glycine vary more in the first 3 days than between day 3 and 4, which can be interpreted as a variation of the previously mentioned 'burn-in' effect, where the variability fluctuates until converging on a more or less fixed value. Perhaps more prominent is the fact that most of the trends for both means and standard deviations appear very similar. Thus, apart from the 'burn-in' period, there also exists evidence for global profiling patterns impacting the accuracy and variability of more than one compound at once. Together, these observations reinforce the need for randomization at the profile level, independent of both scan order and any natural groupings of the samples themselves.

Temporal variability

In a similar fashion to the profile order investigation, it was also possible to assess if trends could be identified as a function of scanning time. Concentration data from both the pulse sequence comparison and variance breakdown data is presented as a function of time in Figure 2.10. The analysis revealed correlated changes in all of the compounds beginning at hour 8. Glucose, maltose, and most prominently, nicotinate, all saw a change in their time

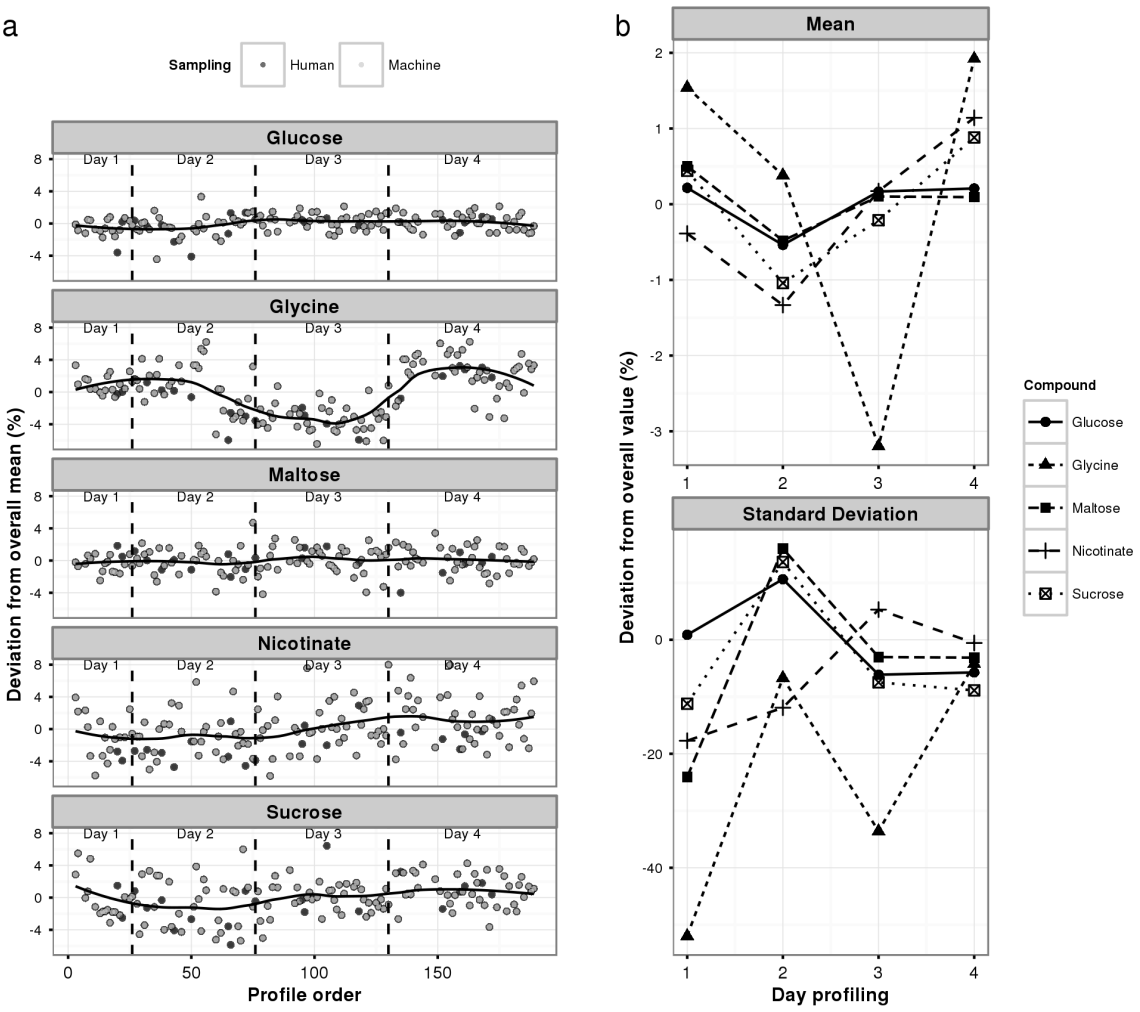


Figure 2.9: **a)** Observed concentrations represented as percent deviations from mean compound concentration as a function of the order in which the spectra were profiled (including duplicated spectra) over a 4-day period. Trend lines were added using LOESS for visual aid only. **b)** Daily mean and standard deviations of observed concentrations as a percentage of overall values for each of the compounds over the 4-day period.

trends at approximately the 8 h mark. Since glucose and maltose both see an increase, compound breakdown can be ruled out. Comparing the 6–8 h period to the 8–10 h period, mean glucose and maltose concentrations both increased by approximately 1% while that of nicotinate decreased by approximately 2% (statistically significant at 95% confidence). While not large in absolute terms, the changes still represent as much as half of total variability. It would be tempting to suggest that these time-dependent trends are singular phenomena that are not present in everyday NMR applications; however, it should be noted that the nature of the performed experiment was particularly well-suited to observe this type of deviation. Rarely would identical samples be scanned multiple times over a period of 2 h, much less with repeated spectral acquisitions.

2.4.4 Human vs. machine sampling

Human and machine sampling were compared based on the means and standard deviations of observations for each compound. Mean concentration values were found to be statistically equal for all compounds except nicotinate, where machine sampled concentration data was approximately 2% higher than human sampled data. The increase in nicotinate concentration can be observed to some extent in Figure 2.10, just as the sampling switches from human to machine. The standard deviation of nicotinate was found to be larger for machine sampling than human sampling by approximately 1% of the mean nicotinate concentration, which is not surprising given the trend observed in Figure 2.10 between 8 and 10 h. Given that these differences pertain to only one of the compounds, we are unable to conclude that any of the differences are due specifically to the nature of the sampling.

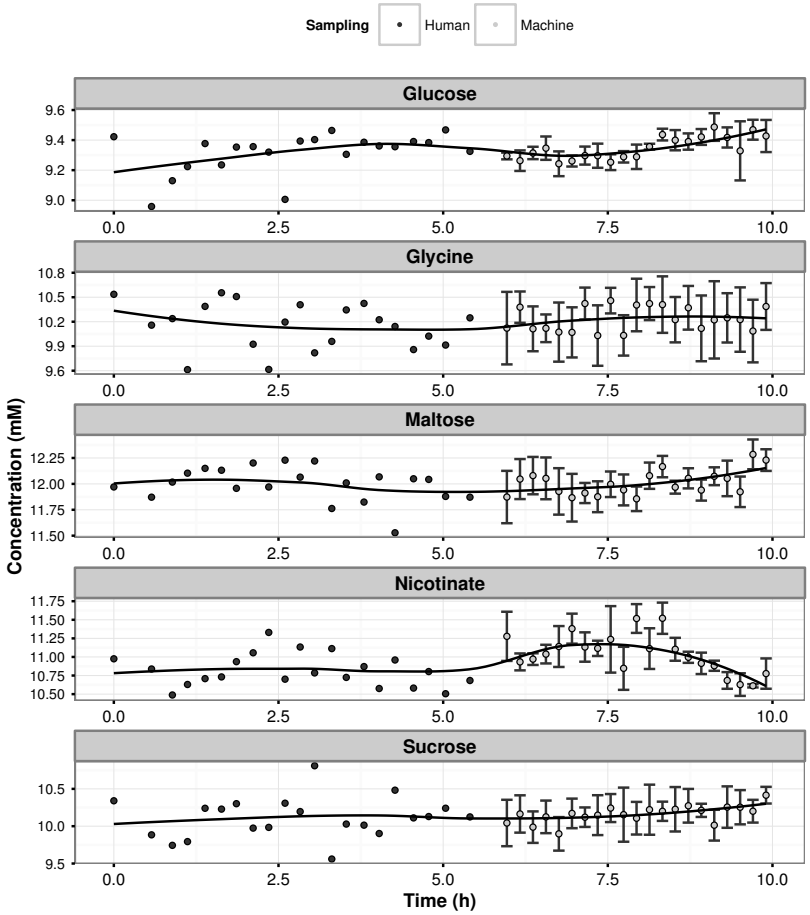


Figure 2.10: Combined human and machine sampled data as a function of experimental time. Human sampled data represents individual 1D-NOESY_s1a4 observations recorded as part of pulse sequence comparison. Machine sampled data represents sample means of three spectra acquisitions per sample, each profiled twice, with error bars signifying observation standard deviations. Both experimental sets were carried out on the same 21 samples, scanned from 1 to 21. Trend lines were added using LOESS for visual aid only.

2.4.5 Concluding remarks

The results presented in this work establish a baseline for quantification variability via targeted profiling approaches and highlight the complexity of NMR quantification that may have otherwise been tempting to overlook. First and foremost, we have established that the choice of pulse sequence has a tangible effect when compound concentration measurement is carried out by comparison to a reference library, even with pulse sequences very similar to the one used for library generation. Differences in solvent suppression were found to have two effects – general underestimation as a result of relative overestimation of the reference peak, coupled with a resonance frequency specific underestimation presumably as a result of bleaching resonances in proximity to the solvent. More generally, we have identified that a large proportion of total spectral variability is due to sample tube insertion and/or shimming – a source of variability that does not generally receive attention. Furthermore, we were able to identify quantification bias as a function of scan and profiling order, providing direct evidence for the importance of randomization at multiple stages of experimentation and analysis. While the magnitude of the observed effects differed considerably in each case, the overall impression is clear – all sources of quantification variability must be considered for the continuing development of NMR technology and practice.

Acknowledgments

The authors would like to acknowledge that the grd-NOESY_s1a4 used in this study was first presented as a poster by RM et al. (2009) at the Experimental Nuclear Magnetic Resonance Conference in Pacific Grove, California and co-developed with Leo Spyracopou-

los of the Department of Biochemistry, University of Alberta. The work presented in this manuscript was supported in part by Natural Science and Engineering Research Council of Canada (NSERC) Discovery and NSERC Strategic Network (MabNet) Grants to MGA, and NSERC Canada Graduate Scholarship to SS. We would like to thank the Canadian National High Field NMR Centre (NANUC) for their assistance and use of the facilities. Operation of NANUC was funded by NSERC and the University of Alberta.

Chapter 3

Targeted Profiling for Cell Culture Applications

Following the determination of basic profiling variability, analysis was expanded to consider the context of cell culture media. Given the dependence of measurement variability on the particular convolution pattern of a specific mixture, it was necessary to tailor analysis to the media of interest. A Plackett-Burman design was implemented as a convenient way to generate orthogonal combinations of high and low concentration levels of each metabolite. Traditionally, the Plackett-Burman design is used as a screening experiment to identify the main effects of a large number of experimental factors in a small number of experimental runs. For this application, factors corresponded to the concentration levels of each metabolite and each run represented a single mixture. An example of an 8 mixture design combining high and low concentration levels of 7 metabolites is presented in Figure 3.1. The orthogonal nature of the design can be seen by focusing on the mixtures where the

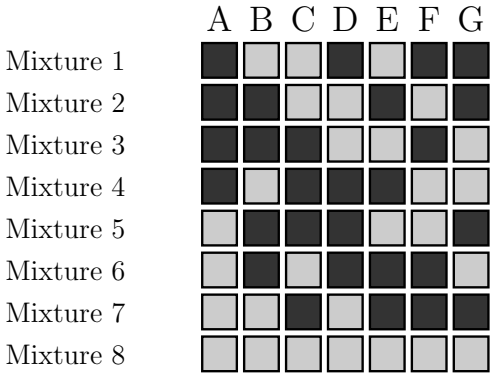


Figure 3.1: An example of a Plackett Burman design for 8 mixtures combining 7 metabolites at low (light) and high (dark) concentration levels.

concentration of metabolite A is high (mixtures 1–4) – each other metabolite is found at high and low levels in exactly half of these mixtures (independently of all other metabolite concentrations). Plackett Burman designs can be generated for any number of mixtures n, where n is a multiple of 4. Up to $n - 1$ metabolites can be combined in an n mixture design.

By combining an equal number of high and low concentration levels in this fashion, variances calculated for the measurement of each metabolite serve as "worst-case" estimates, featuring many combinations of different overlap patterns. Considering the deviation from mean concentration as a response variable, it also becomes possible to use the Plackett Burman design in traditional regression to estimate the impact of all metabolite concentration levels on the calculation error of each metabolite. Although estimated main effects would be confounded with two-way interaction terms, significant interaction effects are not expected for this application. As an illustration, consider the resonance of metabolite A to

be overlapped by the resonances of metabolites B and C. An increase in the concentration of B may hide portion x of A's resonance. Likewise, an increase in the concentration of C may hide portion y of A's resonance. An increase in both B and C would hide portions $x + y$, with the net effect equivalent to a sum of the two main effects (with little to no interaction effect). This principle forms the basis of Plackett Burman design justification presented in section 3.3.1. With the efficiency of this design and the large amount of information that could be extracted from a small number of runs, the results were published in the journal Analytical Chemistry (Sokolenko et al. 2014). The manuscript is presented below in its original form. Supporting information that was made available online is provided in Appendix B.

Profiling convoluted single-dimension proton NMR spectra: A Plackett-Burman approach for assessing quantification error of metabolites in complex mixtures with application to cell culture

Stanislav Sokolenko[1], Eric J. M. Blondeel[1], Nada Azlah[1], Ben George[1], Steffen Schulze[1,2], David Chang[3], and Marc G. Aucoin[1]

[1] Department of Chemical Engineering, Waterloo Institute for Nanotechnology, University of Waterloo. [2] Institute of Applied Biotechnology, University of Applied Sciences Biberach. [3] Chenomx Inc., Edomonton, Alberta.

3.1 Abstract

Single-dimension hydrogen, or proton, nuclear magnetic resonance spectroscopy (1D-^{1}H NMR) has become an attractive option for characterizing the full range of components in complex mixtures of small molecular weight compounds due to its relative simplicity, speed, spectral reproducibility, and noninvasive sample preparation protocols compared to alternative methods. One challenge associated with this method is the overlap of NMR resonances leading to "convoluted" spectra. While this can be mitigated through "targeted profiling", there is still the possibility of increased quantification error. This work presents the application of a Plackett-Burman experimental design for the robust estimation of precision and accuracy of 1D-^{1}H NMR compound quantification in synthetic mixtures, with application to mammalian cell culture supernatant. A single, 20 sample experiment was able to provide a sufficient estimate of bias and variability at different metabolite concentrations. Two major sources of bias were identified: incorrect interpretation of singlet resonances and the quantification of resonances from protons in close proximity to labile protons. Furthermore, decreases in measurement accuracy and precision could be observed with decreasing concentration for a small fraction of the components as a result of their particular convolution patterns. Finally, the importance of *a priori* concentration estimates is demonstrated through the example of interpreting acetate metabolite trends from a bioreactor cultivation of Chinese hamster ovary cells expressing a recombinant antibody.

3.2 Introduction

In recent years, metabolomics has become an attractive experimental tool for generating multivariate quantification data from systems featuring complex mixtures of small molecular weight compounds, with applications in cell culture bioprocessing (Read et al. 2013), biofluid analysis (Fukuhara et al. 2013), environmental toxicology (Davis et al. 2013), and pharmacology (Liu et al. 2014), as well as food and nutrition (Jung et al. 2013), among others. Analytical methods employed in this field are derived from classical analytical chemistry methodologies, such as mass spectrometry (MS), nuclear magnetic resonance spectroscopy (NMR), and high- and ultraperformance liquid chromatography (HPLC and UPLC).

While mass spectrometry is capable of metabolite quantifications in the picogram range, offering the greatest sensitivity of available metabolomics methods, this sensitivity varies as a function of sample characteristics such as acidity and hydrophobicity (Pan and Raftery 2007). Furthermore, sample preparation and processing can be relatively time-consuming and multiple MS methods may need to be employed for final determination of unknown species.

Single-dimension hydrogen, or proton, nuclear magnetic resonance spectroscopy (1D-^{1}H NMR) offers numerous competitive advantages over MS and HPLC. NMR spectra are highly reproducible and can be rapidly acquired; furthermore, sample-preparation is simple, noninvasive, and nondestructive (Bradley et al. 2010; Issaq et al. 2009; Khoo and Al-Rubeai 2007). Metabolite detection occurs uniformly, providing detailed structural information for compound identification (Bradley et al. 2010; Issaq et al. 2009). Challenges associated with

employing 1D-^{1}H NMR for global metabolite analysis include a lower sensitivity (typically in the micromolar range), the overlap or "convolution" of spectral resonance peaks of different metabolites, and differences between samples due to matrix effects such as altered pH and ionic strength (Issaq et al. 2009). The latter of these may cause shifts in the positions of metabolite peaks, hampering tools such as automated spectral binning.(Weljie et al. 2006)

To overcome these challenges, 1D-^{1}H NMR spectra can be analyzed through "targeted-profiling", a method whereby metabolite peaks of NMR spectra are directly matched against pure compound spectra from a database by a human "profiler" (Weljie et al. 2006). This method is tolerant to peak shifts from matrix effects and also allows for "deconvolution" of overlapping peaks (Weljie et al. 2006). With respect to our work, identification and quantification of metabolites through targeted profiling was achieved by using Chenomx NMR Suite 7.5 (Chenomx Inc., Edmonton, Alberta). While multidimensional NMR has been argued as a possible solution to the problem of convolution, 1D-^{1}H NMR remains an attractive option for quantification on the basis of sensitivity and scan time (McKenzie et al. 2011).

Despite the appearance of automated or semiautomated deconvolution methods over 10 years ago (Wishart 2008) and their growing application to cell culture (Behrends et al. 2013; Read et al. 2013), little in depth validation has been done to assess their applicability to particular complex mixtures. A lot of work has focused on quantifying the error associated with the integration of well-dispersed peaks (Malz and Jancke 2005; Maniara et al. 1998; Pinciroli et al. 2004; Wells et al. 2002), but these studies treated spectral overlap as an occasional problem rather than the norm, making their conclusions only loosely relevant

to analysis of cell culture supernatant and other complex spectra, where overlap is highly prevalent. NMR spectra reproducibility has also been examined in the context of a large-scale split-sample study (Dumas et al. 2006). However, the focus was on binning, meaning that metabolites found at low concentrations, whose resonances are almost entirely hidden, would not have had a significant impact on the analysis, despite their possible importance. In more recent years, at least two studies tackled the issue of metabolite quantification with significant convolution with the use of Chenomx software (Slupsky et al. 2007; Tredwell et al. 2011a). Unsurprisingly, the general level of precision was much lower than the $\sim 1.5\%$ standard deviation reported for the integration of well-dispersed peaks (Malz and Jancke 2005). Slupsky et al. (2007) also found that mean concentrations could systematically deviate from theoretical values by as much as 22% for the nine compounds tested. In our work, we have found that relative standard deviation can vary from 1% to 70% depending on a compound's concentration and the environment around its NMR resonance (Sokolenko et al. 2013). These studies make it clear that to determine the precision and accuracy of NMR quantification, convolution must be considered explicitly. Unfortunately, the nature of convolution is that it is entirely dependent on the metabolic makeup of a tested sample, meaning that its effect must be considered on an application basis.

In this work, we apply 1D-^{1}H NMR to a typical animal cell culture media to determine the metabolic profiles NMR analysis can reveal. More importantly, we make use of a common Plackett-Burman design to rigorously probe the observed data for quantification accuracy and precision through the analysis of synthetic metabolite mixtures. The chosen design assesses orthogonal (independent) combinations of compound concentrations, allowing the estimation of accuracy and precision of a given compound considering possible

variation in other (convoluting) compounds (Plackett and Burman 1946). Consistent with the idea of tailored media compositions and rational nutrient feed regimes, the metabolic profile results are specific to the cell line and media under consideration and cannot be generalized. However, we show how a simple, 20–30 sample synthetic mixture design is able to reveal a wealth of useful information about 1D-^{1}H NMR quantification, necessary for understanding application-specific results.

3.3 Experimental section

3.3.1 Experimental design

The purpose of this investigation was to determine the impact of spectral overlap on the precision and accuracy of metabolite quantification in cell culture supernatant by 1D-^{1}H NMR. The hypothesis was that the concentration of one compound may, by spectral overlap, hinder the accuracy and precision in the quantification of another. A more pronounced effect was expected for resonances whose peaks were found exclusively in overlapping regions corresponding to multiple metabolites, particularly in circumstances of significant concentration differences between the chemical species. In order to achieve an accurate measure of quantification error, it was necessary to simultaneously vary multiple compounds at known concentrations. Therefore, synthetic media formulations were constructed with defined two-level component concentrations and varied according to a Plackett-Burman design, allowing for a realistic amount of variability in an experiment of practical size.

One limitation of this design (as compared to the full or fractional factorial design) was the convolution of main effects with two-factor interaction terms. However, Plackett and Burman (1946) have previously argued that interaction terms can be validly neglected if they are of the same order of magnitude as main effects. As spectral convolution is summative in nature, the presence of multiple convoluting compounds was not expected to produce an effect that is substantially greater than the sum of their individual effects. Thus, there was no reason to suspect that factor interactions, if present, would exceed the magnitude of the main effects, allowing the use of the more efficient Plackett-Burman design rather than the full or fractional factorial approach. More information about the Plackett-Burman design as well as steps to reproduce it can be found in many textbooks on statistical design, such as "Statistical Design and Analysis of Experiments: With Applications to Engineering and Science" by Mason et al. (2003).

3.3.2 Synthetic media formulation

The compound concentration levels used in the synthetic media formulations were generated from commonly observed metabolic time course data collected in our lab using Chenomx NMR Suite, excluding those that were not biologically relevant or were identified inconsistently. The remaining compounds were divided into two groups – compounds whose concentration remains approximately constant during fermentation and those whose concentration changes with respect to time. Compounds that were found to have a significant correlation to time (using a Spearman rank correlation coefficient at a 95% confidence level) were added at two concentration levels equal to the minimum and maximum observed concentration values extended by 25% of the minimum to maximum range. The

lower level concentrations were not prevented from reaching zero. For these compounds, it was deemed more important to assess the possibility of false detection, to determine if a true zero concentration would actually be observed. Compounds without a significant correlation to time were added at a single constant concentration level equal to their median observed value. pH was not controlled other than through the use of a pH buffer. Thus, compound composition was allowed to dictate natural pH variation, as it would during a fermentation process, e.g., higher lactate concentration resulting in increased acidity.

Concentrated stocks were made from pure chemical stock (Table B.1) dissolved in PBS buffer. For powder stock, the minimum amount of chemical added was kept at a level sufficient to achieve 99.9% weighing accuracy. The amount of PBS added was determined by the solubility of each compound. The formulation of individual mixtures was performed by volumetric addition of the concentrated stocks with the further addition of PBS to ensure accurate total volume measurement. The use of PBS was a significant simplification when considering the more complex salt and buffer composition of most media. While this ignored the effect of NMR-detectable buffers such as HEPES, the impact on quantification had been previously assessed as minimal (data not shown), in that nonmetabolite resonances did not overlap with metabolite resonances to a significant degree. A full list of compound concentrations and measured pH values for all mixtures can be seen in Table B.2.

3.3.3 NMR analysis

An amount of 630 µL of each clarified medium sample was mixed with 70 µL of internal NMR standard composed of 99.9% D_2O with 5 mM 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) serving as a chemical shape indicator (CSI) and 0.2% w/v sodium azide (NaN_3) to

inhibit bacterial growth (Chenomx Inc., Edmonton, Canada). The solution was vortexed, and 700 µL was pipetted into a 5 mm NMR tube (NE-UL5-7, New Era Enterprises Inc., Vineland, NJ) for scanning.

NMR spectra were acquired using the first increment of a nuclear Overhauser effect spectroscopy (NOESY) pulse sequence with a 1 s presaturation pulse (10 ms relaxation delay, 990 ms water suppression), 100 ms mixing time, and a 4 s acquisition time on a Bruker Avance 600.13 MHz spectrometer with a triple resonance probe (TXI 600). Following acquisition, spectra were imported into Chenomx NMR Suite 7.5 (Chenomx Inc., Edmonton, Canada). Phase and baseline corrections were carried out manually. Line asymmetry correction (reference deconvolution, see Morris et al. 1997) was performed automatically on each sample by the software based on a manual comparison of ideal and observed DSS peaks. Compounds were quantified ("profiled") by a single person using the software's built-in 600 MHz compound library by comparison to a known amount of DSS as the internal standard (see http://www.chenomx.com/ or Weljie et al. 2006 for more information on targeted profiling). Briefly, all resonances were considered in the fitting of a compound's concentration. Isolated resonances were given preference, with concentrations confirmed based on the overall fit of overlapping regions.

3.3.4 General statistical analysis

All data manipulation and analysis were carried out using the R programming language (R Core Team 2012). All plots were generated with the ggplot2 package (Wickham 2009). Unless otherwise noted, the significance of a difference between an observed sample mean and a theoretical value was determined with the use of a two-tailed t test at a significance

level of 0.05. Likewise, the significance of a difference between the variability of two observed samples was determined using a two-tailed F-test at a significance level of 0.05.

3.3.5 Modeling the effect of convoluting compounds on the mean of individual compounds

The experimental design allowed the estimation of spectral overlap effect on the observed concentrations in the form of the following model:

$$\mathbf{y}_i = \mu_i + \sum_{j \in J, j \neq i} a_{ij} \lambda_j + \boldsymbol{\varepsilon}_i \tag{3.1}$$

where $\mathbf{y}_i$ is the column vector of concentrations of compound i from all samples where compound i is at a single concentration level, μ_i is the mean concentration of compound i across those samples, λ_j is the concentration level of compound j whose resonance overlaps with that of compound i (1 for high, 0 for low), a_{ij} is the regression coefficient, J is the set of all varied compounds, and $\boldsymbol{\varepsilon_i}$ is a column of residuals. In this way, it was possible to determine if the concentration level of one compound resulted in an increased or decreased observed concentration of another when the two have overlapping resonances. The effect was quantified using iterative linear regression. When no effect is observed, the model breaks down to the base case of mean estimation. For each compound, a list of compounds with overlapping resonances was compiled based on the spectra obtained in the experiment (Table B.3). Any compounds with overlapping resonances large enough to have a possible effect on the quantification of a target compound were added to the list (judged subjectively – corresponding to the presence of resonance sections with more than $\sim 5\%$ overlap), as

were any compounds whose resonances could be mistaken for those of the target compound, or vice versa.

Linear models of each compound's concentration mean were generated by iteratively considering the effect of overlapping compounds' concentration level. Overlapping compound levels were included in the model one at a time, and any whose level was found to have a significant effect on the measured concentration of a target compound (as judged by an F-test with a significance level of 0.05) was added to a model shortlist. From this list, the compound whose inclusion in the model resulted in the greatest increase in adjusted R^2 value was kept in the model. The process was then repeated to test if further compound terms should be added. Once no improvement could be made to the adjusted R^2 value, the iteration was terminated with the selection of current best model. A schematic of this procedure is presented in Figure B.1.

3.4 Results and discussion

Overall, 28 biologically relevant compounds were used in the synthetic formulation, with the bulk consisting of amino acids, making the formulation typical of a generic animal cell culture process. Of the 28, 16 compounds were included at two levels (corresponding to a significant change with respect to time), while the rest were kept at constant concentrations. With two levels per varying compound, this corresponded to a 20 run Plackett-Burman experiment. While synthetic complex mixtures have previously been used to probe different aspects of NMR-based analysis (Alves et al. 2012; Athersuch et al. 2013; Lewis et al. 2007); this application of synthetic formulations is the first to account for the significant amount

of variation in the metabolite composition of cell culture supernatant. The following is presented as a case study to illustrate the accuracy and precision that can be achieved with the application of 1D-^{1}H NMR to the observation of cell culture, explicitly considering realistic resonance overlap.

3.4.1 Measurement accuracy

A comparison of observed means to theoretical concentrations is presented in Table 3.1 (a box-plot representation can be found in Figure B.2). The level of accuracy was found to vary considerably between different compounds as well as for different concentrations of the same compound. Overall, the absolute percent differences between observed and theoretical concentrations ranged from a low of 1.8% to a singular high of 193.0%, with a median deviation of 9.7%.

Table 3.1: Measurement errors and standard deviations for all compounds included in the synthetic media. An error significance of *** represents 95% confidence that the mean observed deviation from theoretical concentration is significantly different from 0 as determined via a t-test. Compounds marked with * have significantly different percent standard deviations between their high and low concentration levels as judged by an F-test at a 95% confidence level.

Compound	Level	Theoretical Concentration (mM)	Measurement Error (%)	Measurement SD (%)	Error Significance
Glucose	High	19.944	−1.941	3.801	
	Low	9.650	−4.064	4.404	***

Continued on next page

Table 3.1 – continued from previous page

Compound	Level	Theoretical Concentration (mM)	Measurement Error (%)	Measurement SD (%)	Error Significance
*Lactate	High	6.222	7.576	3.809	***
	Low	0.834	16.029	11.638	***
*Glutamine	High	3.966	−7.289	5.345	***
	Low	0	NA	NA	***
Alanine	High	2.348	−4.945	3.066	***
	Low	0	NA	NA	***
Proline	High	2.047	−10.796	8.793	***
	Low	1.261	−11.796	10.468	***
*Isoleucine	High	1.403	−14.149	1.639	***
	Low	0.314	−13.849	9.236	***
Valine	High	1.238	−4.876	2.181	***
	Low	0.534	−5.28	2.996	***
*Leucine	High	1.154	−10.604	5.979	***
	Low	0.248	−3.843	2.823	***
*Glycine	High	1.038	−13.802	5.588	***
	Low	0.033	−16.152	12.121	***
Threonine	Constant	0.888	−8.979	3.941	***
Lysine	Constant	0.814	−10.452	4.791	***
Arginine	Constant	0.692	−4.425	10.838	
Glutamate	High	0.560	−5.819	16.071	
	Low	0.049	192.952	242.857	***
*Formate	High	0.570	6.021	2.632	***
	Low	0	NA	NA	***
Pyruvate	Constant	0.497	−4.736	2.616	***
Tyrosine	Constant	0.454	18.676	8.590	***
Phenylalanine	Constant	0.431	3.741	5.104	***
Serine	Constant	0.42	−14.107	4.286	***
Methionine	High	0.360	2.178	4.444	
	Low	0.211	2.587	4.739	
Pyroglutamate	Constant	0.269	22.204	24.907	***
Asparagine	Constant	0.144	14.362	9.028	***

Continued on next page

Table 3.1 – continued from previous page

Compound	Level	Theoretical Concentration (mM)	Measurement Error (%)	Measurement SD (%)	Error Significance
*Tryptophan	High	0.118	1.772	5.085	
	Low	0.063	18.042	12.698	***
myo-Inositol	Constant	0.111	4.611	10.811	
Arabinose	High	0.101	4.873	17.822	
	Low	0.035	48.647	54.286	***
Acetate	Constant	0.087	21.168	11.494	***
*Choline	High	0.059	−11.113	1.695	***
	Low	0.015	−14.595	13.333	***
Aspartate	Constant	0.057	17.912	17.544	***
Succinate	High	0.045	−6.223	4.444	***
	Low	0.002	NA	0	***

On the basis of the results presented in Table 3.1, the consistency of measurement accuracy across concentration levels depended in large part on the compound in question. Three of the compounds – glutamate, tryptophan, and arabinose – had considerable spikes in percent measurement error at lower concentrations. In all three cases, the magnitude of percent measurement error at the higher concentration level was below or at the median error, in the range of 1.8–10.0%. At lower concentration levels, the lowest percent error for measurement was tryptophan at 18.0%. Analysis of the profiled spectra revealed that the particularly high percent error for glutamate (193.0%) was due primarily to its significant spectral overlap as both of the other compounds had nonconvoluted resonances at 7.5 and 7.7 ppm for tryptophan and 4.5 ppm for arabinose, which could be used for quantification. However, the presence of isolated resonances did not guarantee perfect accuracy. With a lower concentration of approximately 0.05 mM, the resonances of both tryptophan and

arabinose were very close to the baseline and proved difficult to quantify accurately. For many of the other compounds, the general trend was of constant percent error. Proline, isoleucine, valine, glycine, methionine, and choline were all observed with similar percent error, whether at high or low concentrations. Glucose is singular in that it actually had a relatively constant absolute error (0.39 mM) with a corresponding change in percent error at different concentration levels. As the experimental conditions included the variation of all compounds whose concentrations were previously observed to vary with time, the measurement errors presented in Table 3.1 are good general estimates for the expected bias in the concentration measurement of the above compounds. Overall, a compound's concentration was not an effective predictor of its quantification error, as highlighted by the fact that the compound found at one of the lowest concentrations, succinate, was observed with an absolute percent error of only 6.2%. This compound-dependent nature of the observations reinforces the need for validation experiments to include as many relevant compounds as possible and at realistic concentrations.

Depending on the application, consistent biases in metabolite quantifications with 1D-^{1}H NMR may not be problematic. Often, it is not the absolute concentration values that are being analyzed, but correlations (or lack thereof) between time course trends (Aranibar et al. 2011; Bradley et al. 2010; Schaub et al. 2012). For these types of applications, a constant deviation of a reasonable magnitude ($\sim$ 10–20%) is unlikely to result in disputed conclusions. As we have been able to identify, however, the accuracy of some compounds is not consistent over all concentrations. As compound concentrations are generally correlated with time, this can result in biased time trends whether looking at the trends of compound concentrations (Aranibar et al. 2011; Bradley et al. 2010; Schaub et al. 2012) or their rates

of change (Schaub et al. 2012).

3.4.2 False positive identification

Three of the compounds included in the formulation – glutamine, alanine, and formate – had a theoretical concentration of 0 mM as their low concentration level. The use of a 0 mM lower level prevented their comparison to other compounds as they could only be overestimated at this level. It was judged more important to know the prevalence of false positives for these compounds rather than get an estimate of their precision at a low concentration. At low concentration levels, glutamine and formate were both "observed" despite a theoretical concentration of 0 mM; alanine was correctly identified as not being present. The measured concentrations were between 0.005 and 0.015 mM for formate and between 0.01 and 0.05 mM for glutamine, making these values a good estimate of their detection limits. Observations in these ranges can therefore be said to be likely the result of noise and indistinguishable from 0 mM. While false positives in this range are unlikely to influence general time trends, a good understanding of quantification limits can help distinguish between the cases when a metabolite has been entirely depleted versus when its uptake has been down-regulated.

3.4.3 Improving measurement accuracy

The use of synthetic compound mixtures with known compound concentrations allowed a direct assessment of profiling technique. Analysis of the deconvoluted spectra revealed peak-specific strategies that can be used to improve accuracy. Tredwell et al. (2011a)

have approached this subject already by recommending group overview of profiled spectra for the generation of a master profiling list as a way to avoid inconsistent metabolite identification. Our results go a step further and illustrate how a validation experiment can provide actionable information for the improvement of quantification accuracy.

In Chenomx NMR Suite 7.5 software, fit quality is generally determined with the use of a "subtraction line", the difference between the observed spectra and the fitted curve. While the ideal is to get the subtraction line all the way to zero, the reality is that the line may be high in some parts of the curve and low in others, resulting in a considerable amount of profiling ambiguity. How much the sum line tends to vary across the span of a curve is often resonance-dependent. This was found to be particularly problematic for compounds whose spectra consisted of only one or two narrow peaks, such as choline, formate, glycine, and lactate, resulting in inaccurate quantification (for a comparison of peak width, see Figure B.3). The knowledge of true concentrations from the synthetic mixture study made it possible to adjust the profiling pattern to achieve better accuracy by identifying the specific shape of the subtraction line that resulted in the most accurate fit. In particular, glycine and choline concentration accuracy was improved when the total area between the subtraction line and the baseline was set close to zero (treating sections of the subtraction line below the baseline as having a negative area), ignoring poor fitting at the peak itself. The concentrations of formate and lactate, on the other hand, were more accurately fit by using the peaks alone (1.3 ppm doublet for lactate). For lactate, quantification of the 1.3 ppm doublet is preferred to that of the quartet at 4.1 ppm due to the latter's proximity to the suppressed water peak. It should be noted, however, that these techniques may not be valid for different sample compositions.

It was also found that less reliable quantification was obtained for resonances resulting from labile protons, more specifically, hydrogen atoms bonded to nitrogen atoms, or protons vicinal to labile protons. While this configuration is technically found in all amino acids, such resonances were more characteristically problematic for proline, lysine, and isoleucine, where they caused considerable underestimation of the true concentration of the compounds. Even in cases where there was no significant underestimation, these problematic resonances were frequently wider than the pure-compound spectra of the Chenomx NMR Suite 7.5 library (an example can be seen in Figure 3.2). While this would not be a problem for scans of pure compounds, where all resonances are available for quantification and problematic ones can be ignored, complex mixtures such as cell culture supernatant do not always allow the luxury of resonance comparison due to significant spectral overlap from a myriad of metabolites. While the above strategies pertain only to Chenomx software, the synthetic mixture design proposed in this work can be used to identify compound- or peak-specific issues regardless of the software used for quantification.

3.4.4 Effect of spectral overlap on accuracy

From the results of this experiment, two effects of spectral overlap were identified. As mentioned in Section 3.4.1, overlap can have an indirect impact on the accuracy of compound quantification by masking quantifiable resonances. This can occur when the relative intensity of a compound resonance is large enough to completely absorb visible signs of smaller overlapping compound peaks. As expected for animal cell culture supernatant, this effect was observed most prominently for (but not exclusive to) glucose, due to its high concentration. If alternative quantifiable resonances of a compound are available, the net

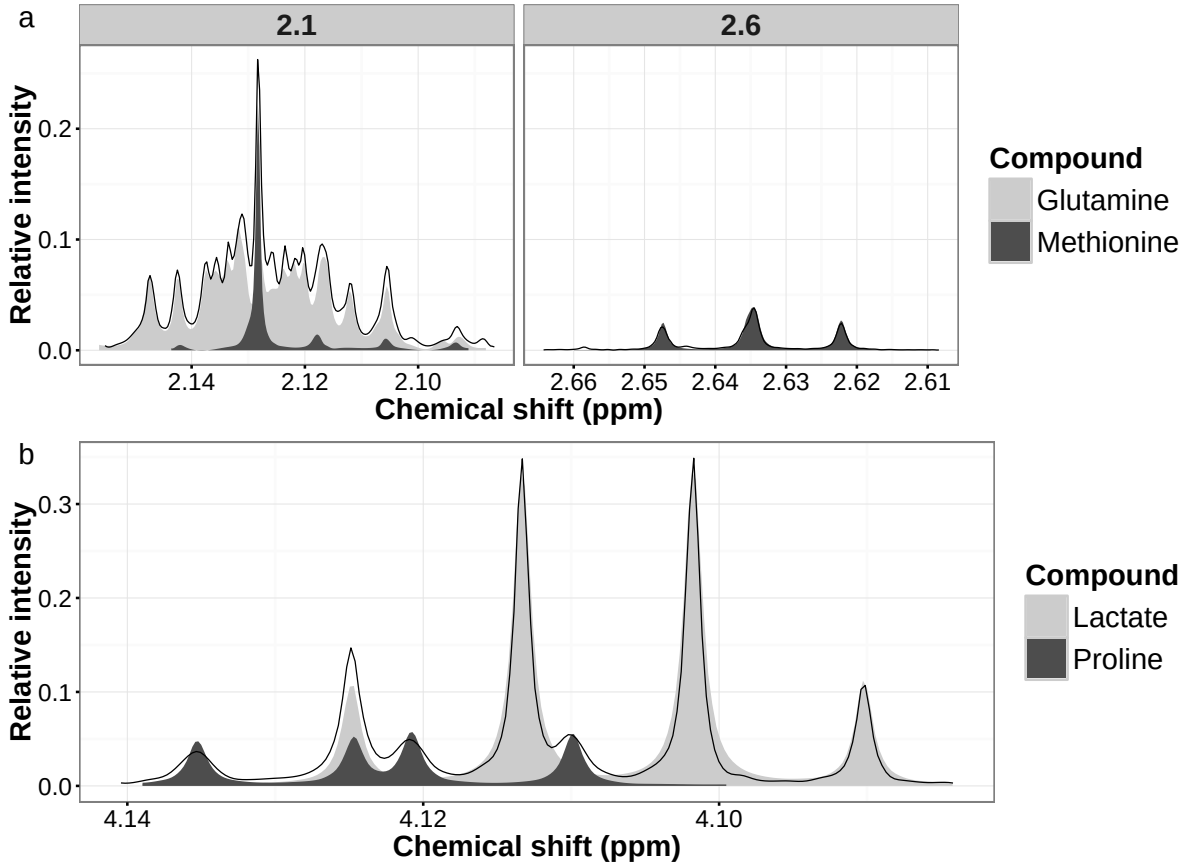


Figure 3.2: Examples of convolution between methionine and glutamine **a)** as well as proline and lactate **b)**. Shaded areas represent pure compound spectra generated from the Chenomx library, with the observed spectra superimposed (thin black line). The differences in peak width of observed spectra when compared to the generated spectra of proline **b)** are also an example of the shape distortion observed for some peaks associated with a labile hydrogen.

effect may be negligible. However, it can also force a profiler to rely on less dependable resonances, such as those corresponding to hydrogen atoms near an amine group, or simply more overlapping regions, leading to a loss in precision. As long as the masking compound resonance stays significantly larger than the masked compound, the overall effect will not be dependent on the exact concentrations of the compounds involved.

A more direct impact of spectral overlap was also identified. When overlapping compound resonances were of similar size, resonance peaks could be misattributed, resulting in a measurement bias that was dependent on the concentration of overlapping compounds. This was tested by modeling the effect of overlapping compound concentration level on the measured concentration of a given compound with linear regression (described in Section 3.3.5), corresponding to main effect estimation generally performed on Plackett-Burman data. Unlike in general practice, however, the effects had to be calculated separately for each level of a target compound as measured concentration distributions were not consistent between levels of the target compound (described in greater detail in Section 3.4.5, below).

Of all the compound relationships tested, three were found to have been statistically significant. Two of these involved methionine and proline quantification. At high levels of methionine, high glutamine levels corresponded to a statistically significant increase in the observed concentration of methionine. Of methionine's five resonances, three are of sufficiently large relative intensity to allow quantification – one at 2.6 ppm and two at 2.1 ppm. The resonances at 2.1 ppm generally overlapped with glutamine (Figure 3.2a). Following the results of the analysis, it was found that when the methionine concentration level was high, high levels of glutamine resulted in a slight (but significant) overestimation

of methionine. At low methionine concentrations, a similar effect was also observed, but it was not found to be statistically significant. The results suggest that some of glutamine's resonances at 2.1 ppm were being incorrectly attributed to methionine, which is in agreement with the general pattern of glutamine underestimation. A similar relationship was observed between proline and lactate, whose resonances overlap considerably at 4.1 ppm (Figure 3.2b). At low proline concentration levels, high lactate levels corresponded to significantly lower proline concentrations, despite the fact that proline has at least four other resonances suitable for verification. A similar but nonsignificant trend could also be observed for high proline concentrations. Similarly to the issues described in the context of direct profiling improvements (Section 3.4.3), these biases can be avoided by a human profiler once identified.

The third identified relationship was also the most directly relevant to cultures supplemented with glutamine. Pyroglutamate, present at only one level, was observed to be approximately 30% higher when the glutamine concentration level was high. An analysis of the spectra revealed that this was not an example of spectral overlap but, rather, evidence of glutamine breakdown, a phenomenon that has been previously observed in literature (Gawlitzek et al. 1998). An increase in glutamine concentration from 0 to approximately 4.0 mM resulted in an increase of pyroglutamate concentration of approximately 0.1 mM and is responsible for most of the overestimation seen in Table 3.1. Similar forms of degradation, if present, could also be detected with this form of analysis.

3.4.5 Measurement variance

Standard deviation values (Table 3.1), like the differences between observed mean and theoretical concentrations, did not have a single general trend. While there was a slight tendency for compounds found at lower concentrations to have a higher relative standard deviation, the results for choline, tryptophan, and methionine are clear counterexamples. When looking at the concentration levels of each compound individually, however, there did appear to be an indication that lower concentrations translate into higher relative variability. Of all the varied compounds, only leucine had a significantly lower standard deviation at lower concentrations for reasons that have not been identified. The rest of the compounds that had a nonzero lower concentration level either showed no change in standard deviation (glucose, proline, valine, glutamate, and methionine) or a significant increase of 2-fold or more (lactate, isoleucine, glycine, tryptophan, arabinose, and choline). While the increase in the standard deviation of glutamate was not technically significant, as it was already high to begin with, its magnitude suggests that this compound belongs more with the latter group than the former. At lower concentrations, the relative standard deviations of these compounds frequently increased to a magnitude of 10% or more. As a 95% confidence interval around an observed concentration with a standard deviation of 10% translates into a range of ±20%, required precision levels should be carefully considered when dealing with low concentrations of these compounds.

The applications of estimating standard deviation values for metabolite quantification may not be as easily apparent as those of bias estimation, but they are just as important. *A priori* information about the confidence of a metabolite observation can be used to confirm the applicability of 1D-^{1}H NMR observation for a given task. The particular overlap

patterns of specific media or cell type can render some metabolites (such as glutamate at low concentrations in the case of certain animal cell culture media) practically unquantifiable. While general glutamate levels can still convey a considerable amount of information, it should be clear that studies focusing on glutamate in particular should seek more tenable methods of quantification.

3.4.6 Application to cell culture results

One example of how these results can be applied relates to the monitoring of acetate concentration in Chinese hamster ovary (CHO) culture. A mammalian bioprocess is typically sampled only once per day, resulting in four or more sample points per compound, depending on the length of cultivation. Acetate concentrations from the first four days of culture can be seen in Figure 3.3a. Due to the sparseness of the data, it was difficult to conclude whether the concentration fluctuations corresponded to a parabolic trend or were the result of measurement noise around a nearly constant value. Adding the variability estimates from the Plackett-Burman results (Figure 3.3b) suggested that the acetate concentrations on days 1 and 2 may be different from those on days 0 and 3, a hypothesis that could be confirmed with the use of t tests. For this particular cultivation, sampling was performed three times a day, rather than only once. Adding the extra data points to the plot (Figure 3.3c) confirmed the calculated trend. While acetate is not typically seen as a particularly important metabolite in mammalian cell culture, its concentration profile can be important in monitoring pyruvate metabolism. This application serves to demonstrate how the generation of a priori variability allows better data interpretation while diminishing the need for high-frequency or repeated sampling.

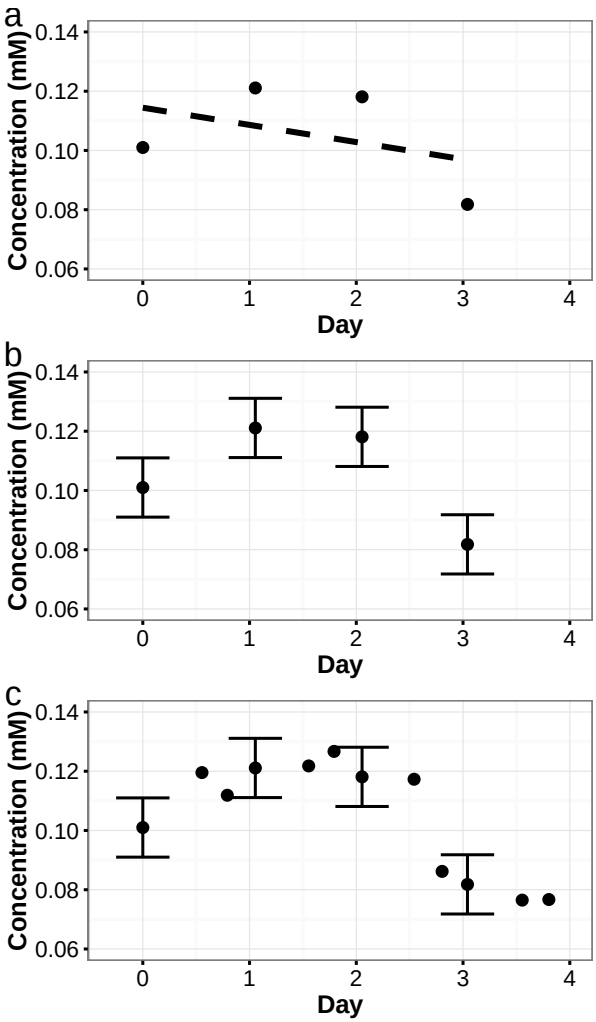


Figure 3.3: Application of variability estimates to acetate concentration data from a 3 L bioreactor (Applikon Biotechnology Inc., Foster City, U.S.A.). Cells were seeded at a concentration of 3×10^5 cells/mL with a working volume of 1.5 L. Supernatant samples were taken three times a day and stored at $-80\,^{\circ}\mathrm{C}$ until NMR analysis 1 week later. The process parameters were set to a pH of 7.4, 50% dissolved oxygen (DO), $37\,^{\circ}\mathrm{C}$, and an agitation speed of 120 rpm. **a)** Daily samples with the possibility of high variability around an approximately constant or decreasing concentration. **b)** Error bars in the form of standard deviations added to the observed data, allowing the use of t tests to compare acetate concentrations at days 2 and 3 to those on day 0 and day 3. **c)** High-frequency sampling from the same cultivation reinforcing the trends determined with the use of variability estimates.

Acknowledgments

This work was supported in part by NSERC ENGAGE and NSERC Strategic Network (MabNet) Grants to MGA, NSERC Canada Graduate Scholarship to SS, and a King Abdullah Scholarship (Saudi Arabia) to NA.

3.5 Further analysis

The manuscript submission was structured around a case study of a specific media composition; however, a number of more general trends were considered outside the submission. Relative standard deviations and absolute bias estimates were combined and plotted as kernel density estimates in Figures 3.4 and 3.5, respectively. The observations below the overall density estimates highlight the overall effect of concentration. Although some metabolites found at low concentrations had relatively low variability and error, decreasing the concentration of any given metabolite generally increased both. Plotting the absolute measurement error versus relative standard deviation (Figure 3.6) reinforced the idea that both bias and variability are effected by resonance overlap. Whereas averaging out multiple uncertain measurement can decrease variability, bias would remain unchanged.

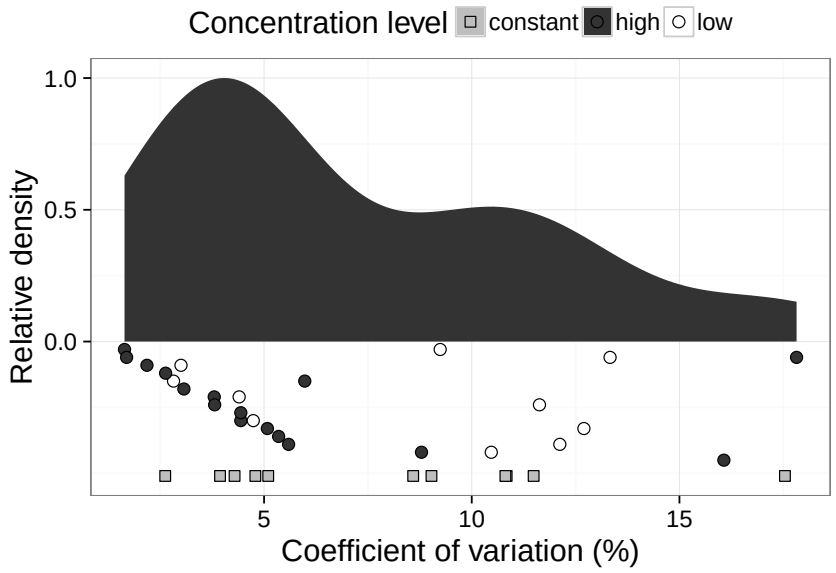


Figure 3.4: Kernel density estimate for the distribution of relative standard deviations of all metabolites except those with singularly high deviations like glutamate and arabinose. Relative standard deviations of metabolites added at low and high concentration levels are shown one metabolite per line below the overall density estimate. Relative standard deviations of metabolites added at a single constant concentration are shown together on a single line.

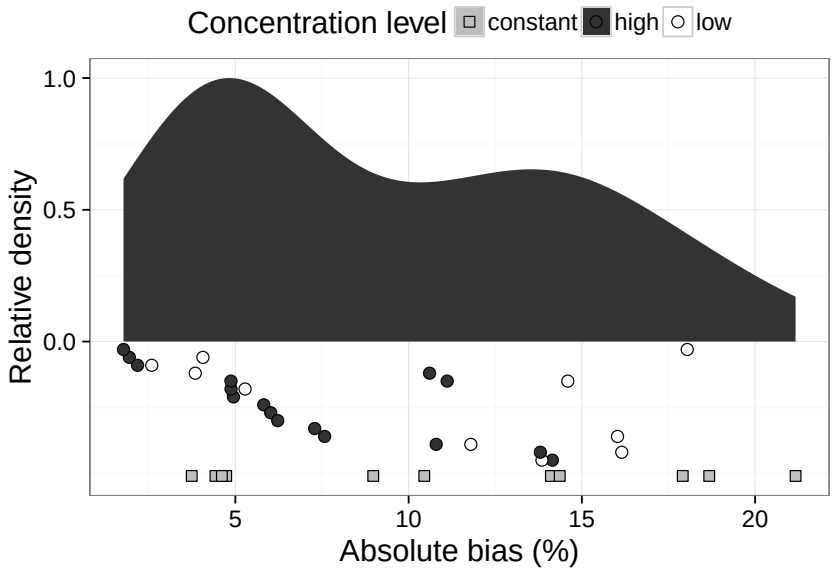


Figure 3.5: Kernel density estimate for the distribution of absolute measurement error of all metabolites except those with singularly high deviations like glutamate and arabinose. Absolute measurement errors of metabolites added at low and high concentration levels are shown one metabolite per line below the overall density estimate. Absolute measurement errors of metabolites added at a single constant concentration are shown together on a single line.

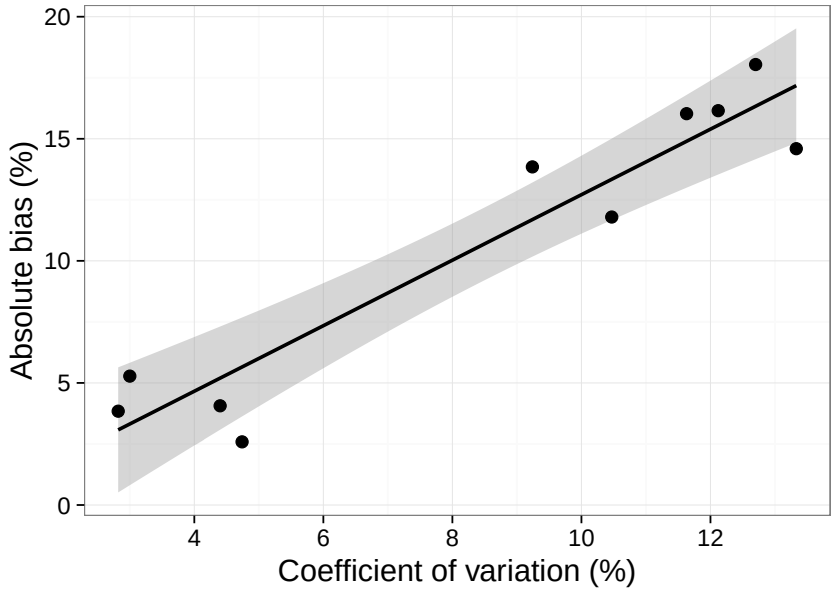


Figure 3.6: Absolute measurement error plotted versus relative standard deviation for all estimates made at low metabolite concentration levels. The line represents a linear regression of the data, with the standard error represented by the shaded region.

Chapter 4

Nonparametric Smoothing of Time-course Trends

Metabolic flux analysis (and other flux balancing methods) naturally rely on the accurate estimation of metabolic flux from extracellular concentrations. Although it's possible to estimate a flux from the slope of the line between two concentration levels (as in Bernal et al. 2009, for example), increasing the sampling resolution allows application of nonparametric curve fitting and derivative estimation (Niklas et al. 2012; Niklas et al. 2011; Priesnitz et al. 2012). However, applying nonparametric regression methods to metabolic time-courses collected in our lab revealed the presence of suspiciously correlated deviations across multiple metabolites at some of the sampled timepoints. Given our understanding of single compound variability, it was possible to simulate realistic time-course trends with the expected amount of uncorrelated error to determine the probability of observing correlated deviations. The results ruled out the possibility of such deviations resulting

from random chance, requiring the development of a correction strategy. The resulting algorithm and validation were published in the journal BMC Systems Biology (Sokolenko and Aucoin 2015). The manuscript is presented below in its original form. Although the publication included a number of additional files providing source code for the analysis, this code is not included here as it has been made available in the form of the metcourse R package available on github (`https://github.com/ssokolen/metcourse`).

A correction method for systematic error in ^{1}H-NMR time-course data validated through stochastic cell culture simulation

Stanislav Sokolenko and Marc G. Aucoin

Department of Chemical Engineering, Waterloo Institute for Nanotechnology, University of Waterloo.

4.1 Abstract

4.1.1 Background

The growing ubiquity of metabolomic techniques has facilitated high frequency time-course data collection for an increasing number of applications. While the concentration trends of individual metabolites can be modeled with common curve fitting techniques, a more accurate representation of the data needs to consider effects that act on more than one

metabolite in a given sample. To this end, we present a simple algorithm that uses nonparametric smoothing carried out on all observed metabolites at once to identify and correct systematic error from dilution effects. In addition, we develop a simulation of metabolite concentration time-course trends to supplement available data and explore algorithm performance. Although we focus on nuclear magnetic resonance (NMR) analysis in the context of cell culture, a number of possible extensions are discussed.

4.1.2 Results

Realistic metabolic data was successfully simulated using a 4-step process. Starting with a set of metabolite concentration time-courses from a metabolomic experiment, each time-course was classified as either increasing, decreasing, concave, or approximately constant. Trend shapes were simulated from generic functions corresponding to each classification. The resulting shapes were then scaled to simulated compound concentrations. Finally, the scaled trends were perturbed using a combination of random and systematic errors. To detect systematic errors, a nonparametric fit was applied to each trend and percent deviations calculated at every timepoint. Systematic errors could be identified at timepoints where the median percent deviation exceeded a threshold value, determined by the choice of smoothing model and the number of observed trends. Regardless of model, increasing the number of observations over a time-course resulted in more accurate error estimates, although the improvement was not particularly large between 10 and 20 samples per trend. The presented algorithm was able to identify systematic errors as small as 2.5% under a wide range of conditions.

4.1.3 Conclusion

Both the simulation framework and error correction method represent examples of time-course analysis that can be applied to further developments in ^{1}H-NMR methodology and the more general application of quantitative metabolomics.

4.2 Background

Hydrogen nuclear magnetic resonance (^{1}H-NMR) spectroscopy is an emerging tool for metabolomic analysis of cell culture. In contrast to the established use of ^{13}C-NMR for targeted elucidation of intracellular metabolic flux (reviewed in (Szyperski 1998)), the quantification of a broader cellular metabolome with ^{1}H-NMR in the context of recombinant protein production has been much more recent (Aranibar et al. 2011; Bradley et al. 2010; Khoo and Al-Rubeai 2009; Read et al. 2013; Yen et al. 2014). Unlike ^{13}C-NMR, which requires relatively expensive ^{13}C labelled compounds and often complex interpretation, ^{1}H-NMR benefits from simple sample preparation and non-selective data acquisition. The result is that a single scan can reveal the concentration of many small molecules in an unbiased manner, with concentration levels reaching as low as the micromolar range. Despite the maturity of ^{1}H-NMR technology, the context of cell culture metabolomics offers opportunities for further developments in both acquisition and post-processing of metabolomic time-course data.

Quantitative NMR relies on the principle that the integrals of resonance peaks are proportional to the number of nuclei that make up the resonances (Bharti and Roy 2012). The absolute area of the integrals is also dependent on spectrometer and sample properties

that include the relaxation time of various metabolites, pulse excitation, and broad-band decoupling. While the effect of relaxation time can be ignored with a sufficiently long acquisition time (or measured and factored in directly – see (Bharti and Roy 2012)), the effect of other factors is accounted for by comparison to a calibration standard. Typical calibration standards can be broadly categorized as internal (where a known quantity of a compound is added directly to the sample), external (where a known quantity of a compound is scanned in a co-axial tube), or electronic (where a synthetic signal generated inside the NMR is used as reference) (see Giraudeau et al. 2014 for an in-depth review). Regardless of how the reference signal is generated, metabolite quantification relies on the ratio of target resonance and reference peak integrals. Unlike typical measurement variability, error in the generation or measurement of the reference signal will have the same relative impact on all the quantified metabolites and represents one example of a systematic bias.

Error related to the reference standard can stem from sample preparation (in the form of pipetting) as well as spectra processing and analysis. Although external and electronic standards do not rely on the addition of a chemical standard, the lack of internal standard introduces extra variability from the the amount of sample analyzed. Proper technique can ensure good reproducibility, but occasional mistakes are nonetheless possible. More importantly, the reference peak is subject to the same variability as any other resonance. Phase and baseline correction, which are typically performed on all NMR spectra, are known to have a considerable impact on the accuracy of peak area integration (Griffiths and Irving 1998). Malz and Jancke (2005) have observed that while routine standard deviation can be reduced to 1.5% of mean concentration, the relative uncertainty can be as

high as 11% with just "slightly" wrong phase and baseline corrections. Other factors may also come into play depending on the quantification method. Some commercial packages such as Chenomx NMR Suite (Chenomx Inc., Edmonton, Canada), which has been used in recent cell culture applications (Aranibar et al. 2011; Bradley et al. 2010; Read et al. 2013; Yen et al. 2014), require the user to match the observed internal standard peak to an idealized representation. Apart from introducing user uncertainty, this method may be particularly sensitive to line shape variability. Discrepancies between the ideal and observed shapes of the internal resonance peak due to imperfect shimming are a likely source of quantification error.

While errors from standard quantification impact practically all NMR samples to some extent, biofluid and cell culture samples are also subject to dilution effects. Urine samples vary in their water content, which is corrected by normalization to either total spectrum area or a reference metabolite such as creatinine (reviewed in (Smolinska et al. 2012) and (Giraudeau et al. 2014)). The metabolomic analysis of cell lysates, common to many cell culture applications such as drug discovery (Powers 2014), suffers from similar problems due to the variability of extraction efficiency. The effect of variable solvent concentration results in the same systematic error as from reference quantification – a global underestimation or overestimation in the relative concentrations of all observed metabolites in a given sample.

The application of NMR spectroscopy and other metabolomic approaches to time-course samples presents both a unique challenge and opportunity in dealing with systematic errors. On the one hand, a single biased sample can skew the trends of multiple compounds and suggest false metabolic relationships. On the other hand, the time-course trends of metabolite concentrations have a significant degree of implicit replication that can be ex-

ploited through mathematical means. Recent work with cell culture (Niklas et al. 2011) and biofluid (Berk et al. 2011) data has used nonparametric curve fitting techniques to model metabolite concentration trends by leveraging the inherent smoothness of biological trends. This work extends the concept by identifying systematic deviations across a number of metabolites. In the same way that a dramatic deviation from an overall trend of a metabolite's concentration is identified as measurement error via smoothing spline regression, the deviations of many metabolites in one sample can be identified as the result of reference error or a dilution effect.

In the context of cell culture process monitoring, a subset of compound concentration trends from a batch culture shown in Figure 4.1 illustrates the confusion that can arise from possible systematic errors (details provided in the Cell culture section of Methods). The jumps in concentrations of glycine and lysine on days 4 and 5 correspond with the exhaustion of choline and the peak of o-phosphocholine concentration. The question is whether these deviations from the general trend of the compounds can be interpreted as a physiological shift in cellular metabolism or if they are more likely to be the result of systematic error that is associated with internal standard addition. This work presents a simple iterative smoothing algorithm as a means to address this issue. The method is tested by the stochastic generation of cell culture trends subject to simulated observation error to ensure that identified systematic errors are independent of measurement uncertainty.

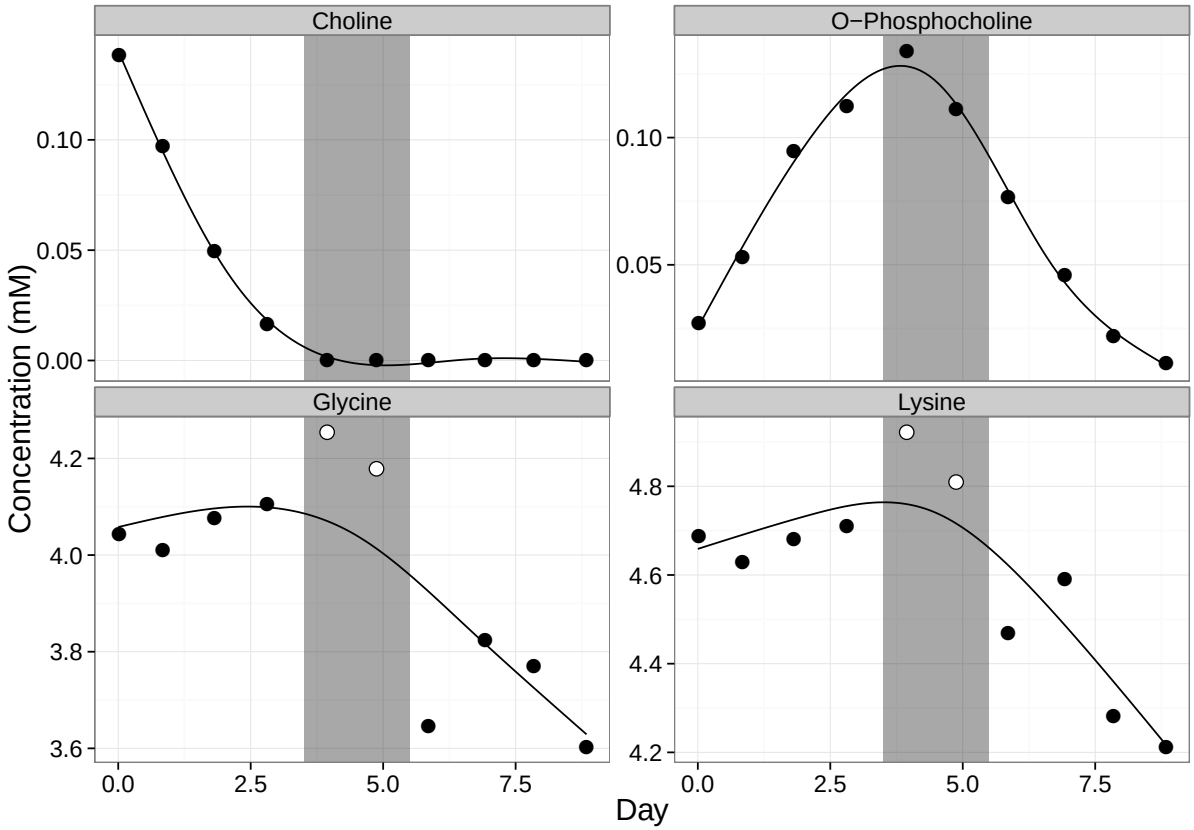


Figure 4.1: An example of 4 metabolite trends from a metabolic study. Jumps in glycine and lysine concentration trends (highlighted as white points) were hypothesized to be the result of choline exhaustion (region highlighted in grey). Time-course data was collected as described in the Cell Culture subsection of the Methods section.

4.3 Methods

4.3.1 Cell culture

Metabolic data presented in this work originated from an insect cell media supplementation experiment. *Spodoptera frugiperda* (Sf9) cells were grown in shake flasks at 27 °C and 130 RPM using in-house supplemented IPL-41 media (Weiss et al. 1981). The cells were routinely split to $0.5{\cdot}10^6$ cells/mL upon reaching a concentration of $2{\cdot}10^6$ cells/mL, with experiments carried out on cells that have undergone less than 30 passages. A 1 L mother flask was seeded at $0.5{\cdot}10^6$ cells/mL with a working volume of 250 mL and grown up to $2{\cdot}10^6$ cells/mL. This flask was used to seed 125 mL flasks at $0.5{\cdot}10^6$ cells/mL with a working volume of 30 mL. Cells were counted and sampled for NMR every 24 hours until reaching their maximum concentration (of approximately $7{\cdot}10^6$ cells/mL). 1 mL samples of cell culture media were collected and centrifuged for 8 minutes at 250 g, with the supernatant collected and stored at −80 °C until NMR analysis.

The experimental data used as a template for stochastic trend generation, hereafter referred to as reference data, consisted of 4 different carbohydrate supplemented flasks cultured over a period of 10 days. The cultures were identical and seeded from the same stock, but with varying concentrations of glucose and maltose. 43 Compounds were profiled for a total of 172 model trends across the 4 flasks. Although many of the compound concentration trends were similar across the flasks, the use of different conditions resulted in more general trends than would be available from replicates.

4.3.2 NMR

The collected supernatant samples were thawed at room temperature and NMR samples prepared by the addition of 70 µL internal standard to 630 µL supernatant. The standard consisted of 5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) and 0.2% w/v sodium azide preservative dissolved in 99.9% D_2O (Chenomx Inc., Edmonton, Canada). The NMR sample solutions were vortexed and pipetted into 5 mm NMR tubes (NE-UL5-7, New Era Enterprises Inc., Vineland, NJ). Samples were randomized and scanned over a two day period on a Bruker Avance 600 MHz spectrometer with a triple resonance probe (TXI 600). Scans were performed using the first increment of a 1D-NOESY pulse sequence with a 1 s presaturation pulse, 100 ms mixing time, and a 4 s acquisition. The acquired spectra were re-randomized (Sokolenko et al. 2013) and analyzed using Chenomx NMR Suite 7.7 (Chenomx Inc., Edmonton, Canada). Phasing and baseline correction were done automatically by the software and adjusted by a human profiler. Compound concentrations were calculated using the "targeted profiling" method (see Weljie et al. 2006 for more information). Briefly, the observed spectra were fit by the overlay of idealized NMR resonance peaks from the software library, with compound concentration quantified by comparison to an idealized fit of the DSS resonance peak.

4.3.3 Systematic error correction

Starting with all compound concentration time-courses from a single cell culture, a non-parametric (smoothing) model was fit to each time-course. Percent deviations from the fits were calculated at each timepoint and for each compound ($\varepsilon_{time=i,compound=j} = (y_{i,j,observed} -$

$y_{i,j,smoothed})/y_{i,j,smoothed}$). A median percent deviation was taken at each timepoint, corresponding to sorting all the deviations at a given timepoint from lowest ($\varepsilon_{time=i,1}$) to highest ($\varepsilon_{time=i,n}$), and focusing on the middle (or median) value ($\varepsilon_{time=i,n/2}$). If the largest median percent deviation exceeded a specified threshold, it was subtracted from the observed concentrations of all compounds at the corresponding timepoint. The process was repeated until the largest deviation failed to exceed the specified threshold. An overview of the algorithm is presented as a flowchart in Figure 4.2.

In principle, the algorithm takes advantage of the fact that an error in internal standard addition or quantification will result in a deviation for all quantified compounds relative to their concentration. As the percentage error from measurement uncertainty can be quite high for some media components (Sokolenko et al. 2014), the median of relative deviations was chosen as a conservative statistic that could still be capable of identifying systematic error. Mean values were also tested but found to be more susceptible to random noise. An iterative process was used to account for the effect an erroneous measurement can have on a smoothing trend. Once a systematic deviation is identified, the deviating timepoint is corrected and the trend re-smoothed to calculate new deviations. Although the elimination of a deviating timepoint would also be suitable, correction has been chosen in this work as it conserves more of the observed data in the form of a consensus between all compound trends.

The choice of smoothing model and median deviation threshold are two important parameters for error detection. A smoothing model should be chosen according to the expected smoothness of compound concentration trends i.e. how likely they are to exhibit rapid fluctuations. A high-density cell culture or one subject to perturbation may

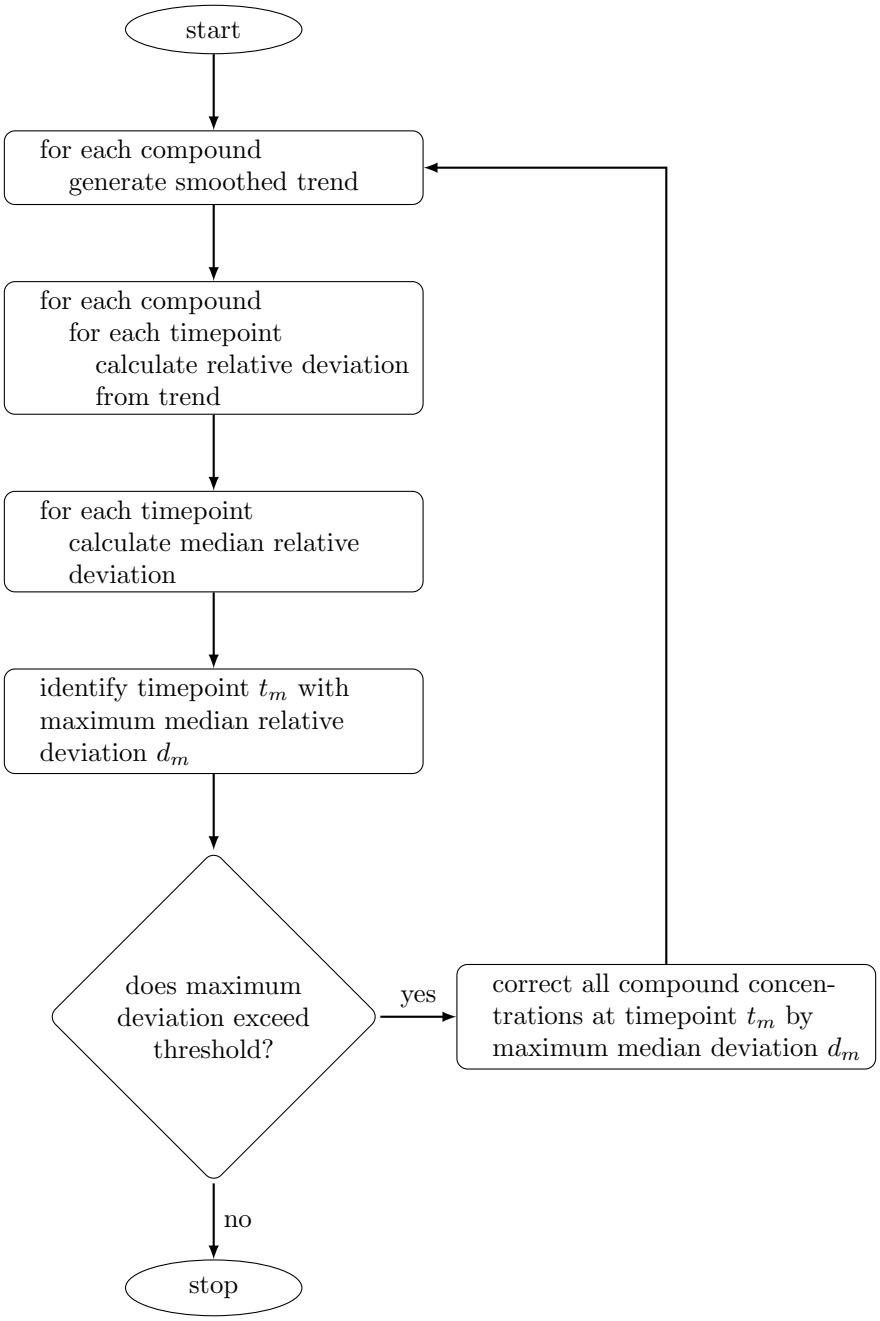


Figure 4.2: Algorithm flowchart. Step by step description of the internal standard error correction algorithm. Corrected values can be kept or flagged for further investigation/removal.

require less smoothing to ensure that rapid physiological changes are not mistaken for internal standard error. On the other hand, a slow-growing or continuous culture could use a much greater degree of smoothing. The median deviation threshold represents the minimum amount of deviation that can be attributed to come from systematic error rather than random measurement uncertainty. High measurement uncertainty is reflected in the variability of median deviation, requiring higher thresholds to prevent false bias detection. However, a number of other factors can also have an impact, including the number of observed compounds and the number of timepoints included in the trend. The effect of these factors on the threshold is explored in this work using stochastic trend generation.

4.3.4 Stochastic trend generation

The development of a framework for stochastic generation of extracellular compound concentration trends was based on the need to estimate the variability of median relative deviations from a smoothing fit. Trend simulation was reduced to four general parameters – overall trend shape, maximum compound concentration, percent change in compound concentration, and measurement variability. The framework was developed around a reference of collected data and consisted of four steps. First, the reference trends were classified as either increasing, decreasing, concave, or approximately constant. A parametric model was chosen for each classification, and representative curves generated with a domain and range of 0 to 1. The combination of simulated maximum compound concentrations and percent changes were used to generate maximum and minimum concentration values to scale the trend. Finally, measurement variability was simulated and applied to the data. The combination of multiple trends with varying parameters was taken to be a represen-

tative of the data one would collect from the time-course of a single culture and is termed "an experiment" throughout the text.

Trend classification

Initial classification of the reference data identified trends with a net change in concentration greater than 10%. Concentrations with changes of less than 10% were taken as having approximately constant concentrations, or "unclassified". Simple linear regression was used to classify trends as either increasing or decreasing if the slope was found to be statistically different from 0 at a 95% confidence level using a t-test. Compound concentrations that had a statistically significant increase followed by a statistically significant decrease were classified as concave (none of the trends could be statistically determined as convex). Trends were left unclassified if the classification of a compound differed across the different experimental conditions. This was done to ensure that classification was restricted to general patterns rather than singular observations. In this way, 15 compounds were classified as decreasing, 14 increasing, 2 concave, and 12 were left unclassified. To allow changes in the number of simulated compounds, these numbers were reformulated and rounded to 35%, 30%, 5% and 30% of the total compounds respectively.

Trend shape

Classified reference data trends were smoothed using cubic regression splines with an upper limit of 4 degrees of freedom (Figure 4.3a). When normalized to the same domain and range, most of the concentration trends appeared to take very similar shapes. Sigmoidal

equations (with 2 parameters) were used to model the increasing/decreasing trends while the concave curves were approximated by a truncated beta distribution density function:

$$\text{sigmoidal decrease:} \quad y = \frac{1}{1 + e^{\frac{x-a}{b}}} \qquad x \in [0, 1] \tag{4.1}$$

$$\text{sigmoidal increase:} \quad y = 1 - \frac{1}{1 + e^{\frac{x-a}{b}}} \qquad x \in [0, 1] \tag{4.2}$$

$$\text{concave:} \quad y = x^{a-1} \cdot (1 - x)^{b-1} \qquad x \in [c \geq 0, d \leq 1] \tag{4.3}$$

The sigmoid functions were defined over a domain of 0 to 1, while the beta function's domain was kept variable. The extra parameters offered greater flexibility in controlling the rate of concentration changes. The y values (and beta distribution x values) were scaled to a range of 0 to 1 after simulation for easier comparison. Unclassified compounds were assumed to follow a linear trend with equal probability of either increasing or decreasing. The linear trend was used to convey a lack of information rather than a strictly linear relationship in compound concentration i.e. the case where a true trend was dwarfed by relative measurement error.

Model parameter ranges were selected by trial and error to visually match the observed trends. As the increasing/decreasing trends showed evidence of two distinct patterns each, two sets of parameters were chosen for the sigmoidal curves along with a separate parameter that related the probability of sampling from one population or the other. The parameters in Table 4.1 were used to generate the trends in Figure 4.3b. Overall, the simulated trends were highly comparable to the observed ones. Although there was less agreement between the concave trends, parameter constraints were kept flexible to account for the low number of concave reference curves.

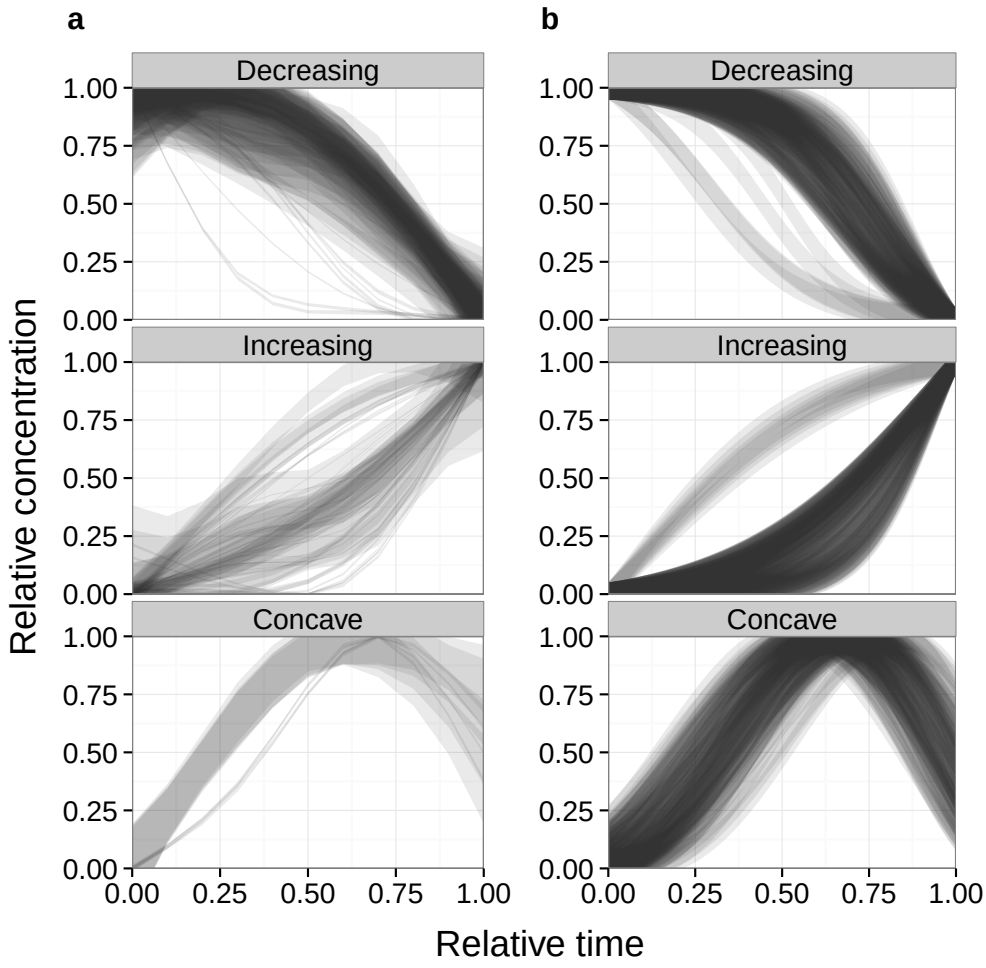


Figure 4.3: Trend shape simulation. A comparison of **a)** observed and **b)** simulated compound concentration trends mapped to a domain and range of 0 to 1. Observed data has been smoothed using cubic regression splines with a maxium of 4 degrees of freedom. The line widths represent estimated standard error ranges from the regression spline model. Simulated trend generation is described in the text and the line widths have been set to a constant 5% of maximum value. Lines were plotted using a high degree of transparency, with darker areas indicating a higher density of curves passing through them. In **b)**, 100 trends were generated for each of the classifications using Equations 4.1-4.3 and parameter values listed in Table 4.1.

Table 4.1: Parameter ranges used in the trend shape simulation of reference data (via Equations 4.1-4.3). Parameter values were drawn from a uniform distribution constrained to the given ranges. Where two ranges are given, the range was chosen randomly for each trend based on the given probability.

Trend	Parameter	Range 1	Range 2	P(Range 1)
Sigmoidal decrease	a	0.200-0.600	0.600-0.900	0.05
	b	0.100-0.180	NA	1.00
Sigmoidal increase	a	0.045-0.055	0.945-0.955	0.15
	b	0.200-0.400	0.100-0.300	0.15
Concave	a	3.500-4.500	NA	1.00
	b	2.500-3.500	NA	1.00
	c	0.000-0.200	NA	1.00
	d	0.800-0.900	NA	1.00

Trend range

The conversion of idealized trend shapes to realistic concentration time-courses required the generation of minimum and maximum values. The distribution of maximum compound concentrations from the reference data is shown in Figure 4.4b. Compounds increasing in concentration were observed to have lower maximum concentrations than decreasing ones, requiring the simulation to be based on trend classification (with concave compounds being treated as increasing). On a logarithmic scale, the spread of maximum concentrations was reasonably modelled by a mixture of two normal distributions with means of -0.4 and 0.8 (corresponding to approximately 0.4 mM and 6.3 mM respectively) and standard deviations of 0.35. The probability density functions of the resulting distributions can be seen in Figure 4.4a with the comparison to observed values in Figure 4.4b. The proportions between the lower and higher concentration clusters were chosen as 0.20, 0.70, and 0.35 for

the decreasing, increasing, and unclassified trends respectively. Although a greater degree of fine tuning was possible to achieve better agreement between observed and simulated distributions, the marginal improvement did not warrant deviating from more general consistency.

To avoid dealing with the correlation between maximum and minimum concentrations (for compounds with relatively small changes in concentration), minimum values were generated from the simulation of net concentration change as a fraction of maximum value. Relative concentration changes were assumed to be less dependent than minimum concentrations on maximum values. As compounds with increasing concentrations were generally observed to have an initial concentration of approximately 0, their percent change was taken as 100% for the purpose of the simulation. The distribution of fractional changes for decreasing compound concentrations is shown in Figure 4.4d. One compound was practically exhausted in all 4 of the tested conditions, with the remainder of the compounds being consumed to various degrees but clustering around 25% reduction. No change of less than 10% can be observed as this value had been chosen as a cutoff for separating compounds with a significant trend. The simulation distribution was modelled by a mixture of two beta distributions – one to represent the distribution of non-exhausted compounds ($\alpha = 2$, $\beta = 5$) and another to increase the probability of values close to 0 and 1 ($\alpha = 0.5$, $\beta = 0.5$), with the proportion between the two set to 0.7 (Figure 4.4c). The simulated distribution was truncated to the range of 0.1-1.0 to reflect the reference data. Figure 4.4d suggests that the simulation was in good agreement with the reference data.

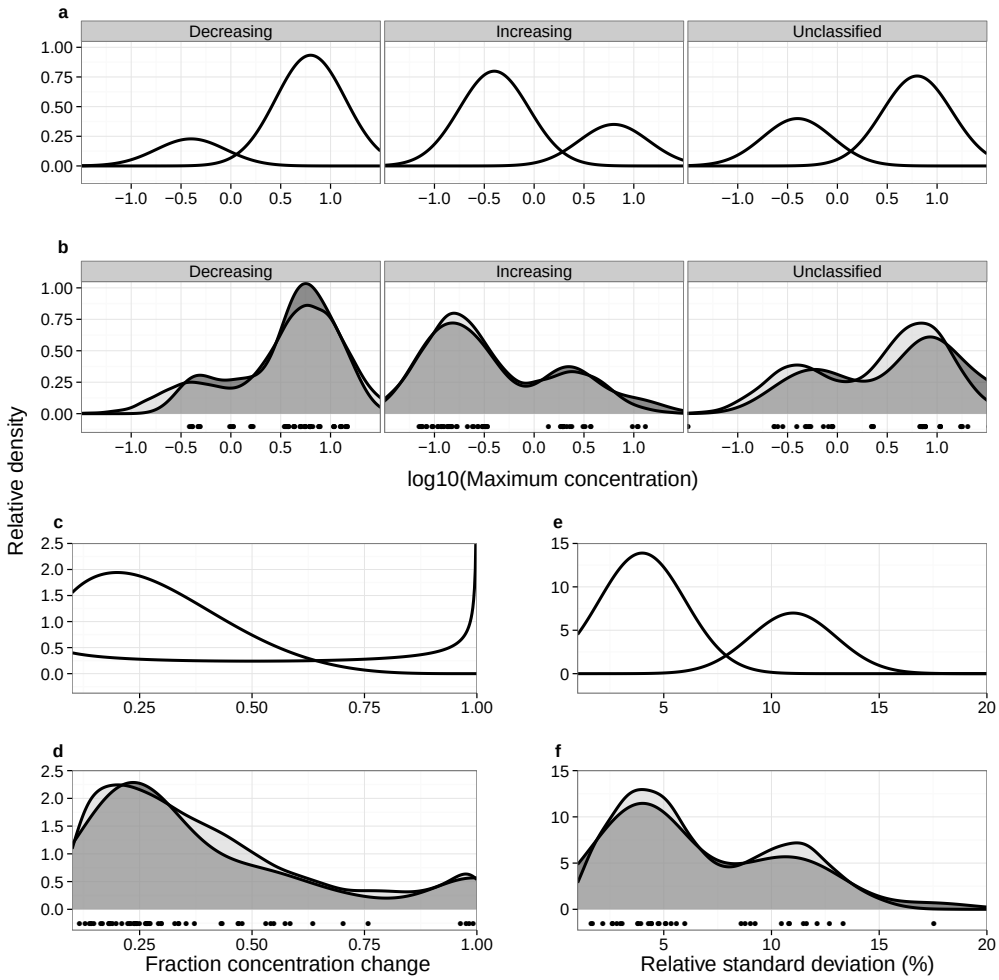


Figure 4.4: Trend scale and perturbation simulation.**a)** Probability density functions used to simulate maximum compound concentration distributions. **b)** A comparison of observed (darker grey) and simulated (lighter grey) maximum compound concentration distributions (with a semi-transparent overlap). **c)** Probability density functions used to simulate fraction concentration change distribution. **d)** A comparison of observed (darker grey) and simulated (lighter grey) fraction concentration change distributions (with a semi-transparent overlap). **e)** Probability density functions used to simulate fraction relative standard deviation distribution. **f)** A comparison of observed (darker grey) and simulated (lighter grey) relative standard deviation. The curves represent kernel density estimates, with the simulated data generated from 10000 samples per classification from mixtures of two normal distributions (see text for details). Observed data points are shown below the curves.

Measurement variability

A measurement variability distribution was developed from our previous work on estimating ^{1}H-NMR measurement uncertainty for cell culture applications (Sokolenko et al. 2014). Briefly, a Plackett-Burman design was used to generate a series of media-like formulations with an orthogonal combination of high and low compound concentrations. In this way, measurement standard deviations for each compound could be estimated independently of other compound concentrations. The result was a collection of relative standard deviations (otherwise referred to as the coefficients of variation) for all compounds in the media. Relative standard deviations for compounds with a statistically significant change in concentration during cell growth were estimated at both high and low concentrations; a single estimate was used for compounds without a significant change.

As the differences in relative standard deviation between compound concentrations were not typically large, all of the relative standard deviations were pooled together into a single distribution of measurement uncertainty (Figure 4.4f). Three of the compounds that were particularly challenging to quantify (Sokolenko et al. 2014) (and had correspondingly high uncertainties) were excluded as they were not representative of typical quantification – compounds identified to have low concentrations and considerable resonance overlap were not quantified in this work. The resulting distribution took the shape of a bimodal normal distribution (Figure 4.4f) with means of 4% and 11% and a common standard deviation of 2% (probability density function shown in Figure 4.4e).

4.3.5 Algorithm validation

The simulation framework was applied to answer two fundamental questions. What is the minimum level of bias that can be identified given normal measurement variability? How is bias identification impacted by the choice of smoothing model and experimental parameters? Two smoothing models were considered – local linear least squares regression and a cubic regression spline. The former was implemented by the loess function in base R and the latter as a general additive model (gam) provided by the mgcv package (Wood 2011). Both models made use of a smoothing parameter. The loess approach required a span that dictated what fraction of data points to use in local regression. This parameter was varied from 2.0 (less smooth) to 0.5 (more smooth). The gam approach required the choice of basis dimension number, which was varied from 3 (less smooth) to 6 (more smooth). In the text, models are referred to by their smoothing parameter i.e. loess-0.5 or gam-6. Combined with model type and smoothing parameter, the number of quantified compounds (20–60) and the number of observed data points (10–20) were also seen as important factors that could influence bias detection.

1000 experiments were simulated for each factor combination (with the number of trends making up a single experiment varied as a parameter). Half of the experiments were subject to normal measurement variability, while half were further perturbed with a systematic bias of 5% at a single randomly selected timepoint. Algorithm performance was assessed by smoothing the simulated data using a given model and calculating the median relative deviation of observations from the fit for each timepoint in each experiment. The result was a pool of median values for each timepoint corresponding to a certain factor combination.

4.3.6 Implementation

The algorithms and all analysis has been implemented in the R programming language (R Core Team 2012). Figures were generated using the ggplot2 package (Wickham 2009).

4.4 Results and Discussion

4.4.1 Application

The correction algorithm was applied to the example data from Figure 4.1 and the results can be seen in Figure 4.5. Although only glycine and lysine results are shown, all 43 observed compounds were used in the calculation (using a gam-5 smoothing model and a threshold of 2.5% median deviation). The algorithm provided strong evidence that the jumps in glycine and lysine concentration were not due to metabolic shifts but were the result of a systematic error. Figure 4.5a also demonstrates that random measurement error such as the pronounced deviation in glycine concentration on day 6 was not impacted by the correction, as it was not general to all metabolites. The influence of the correction was most pronounced in the rates of concentration changes calculated as the derivatives of the smoothing curves (Figure 4.5b). As a result of the changes in concentration, both compounds went from being produced then consumed to a steady pattern of increasing consumption. More importantly, the correction of only two points resulted in considerable changes to derivative estimates across all time-points. This can have an important impact on the use of spline smoothing for flux estimation in metabolic flux analysis (as in (Niklas et al. 2011), for example).

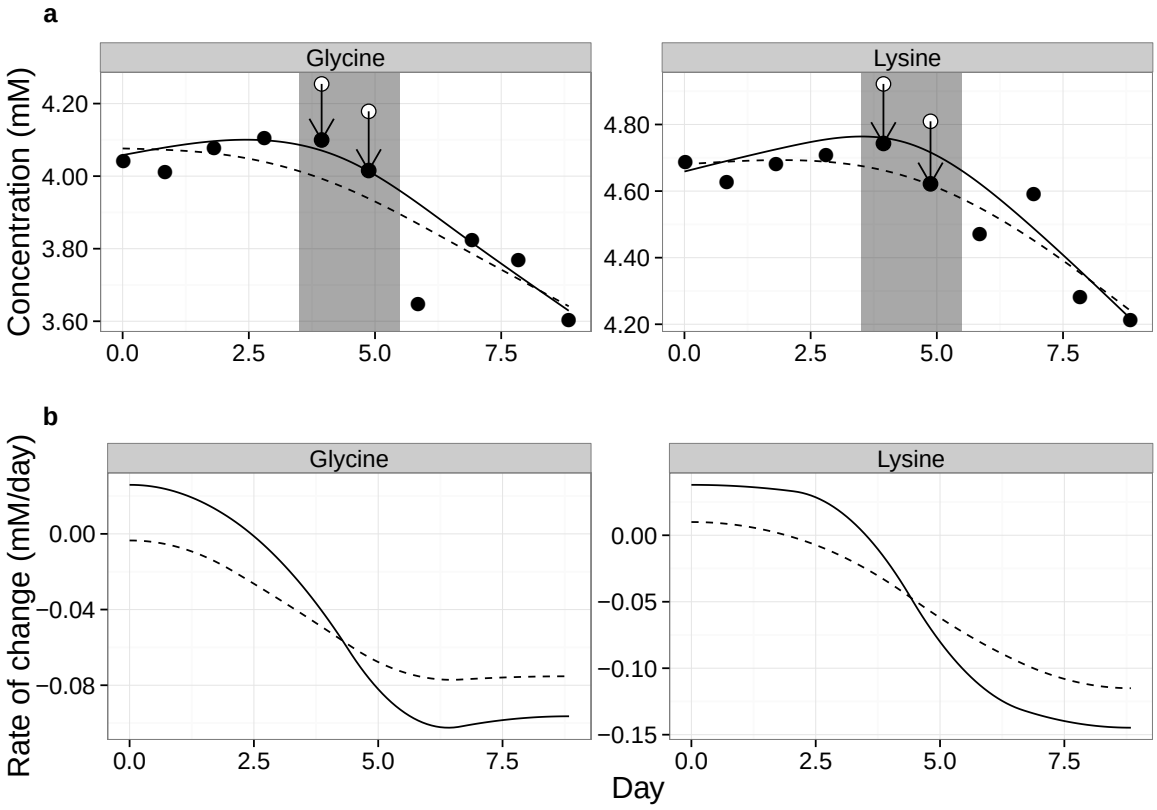


Figure 4.5: Correction applied to example data. **a)** White points represent initially observed concentrations that were marked for correction by the algorithm, while black points represent final compound concentrations (with arrows signifying correction). Smoothing lines were generated using the gam-5 model using uncorrected (solid line) and corrected (dashed line) data. Time-course data was collected as described in the Cell Culture subsection of the Methods section. **b)** Derivatives calculated from the uncorrected (solid line) and corrected (dashed line) smoothed fits.

4.4.2 Validation

Smoothing bias

The smoothing model used in the correction algorithm must strike a balance in having enough flexibility to follow metabolism related changes in compound concentrations while avoiding undue influence from deviating observations. A lack of flexibility can result in systematic deviations from a smoothing fit where no errors are present, while too much flexibility can underestimate deviations due to error. The simulated trends described in the Algorithm validation section were smoothed using loess and gam models (with varying smoothness parameters) and the median deviations from each experiment were averaged to identify overall trends (Figure 4.6). Unsurprisingly, a greater degree of smoothing resulted in less biased deviations i.e. loess-0.5 and gam-6/gam-5 models had practically constant deviations across all timepoints. On the other hand, using an inadequate amount of smoothing generally resulted in an underestimated fit early in the culture (positive deviations from the smoothing fit) and an overestimated fit later. Between the two smoothing functions, gam was found to have a better discrimination of artificially biased timepoints than loess at comparable smoothing levels (gam-5/6 and loess-0.5) – the deviations were more consistent across different timepoints and were not as sensitive to the number of observations. Although the jump from loess-0.5 to loess-1.0 in Figure 4.6 is quite considerable, further analysis using other span parameters reinforced the observation that gam smoothing is superior for bias discrimination. As gam-5 requires less information than gam-6, it can be seen as a good compromise between an unbiased fit and deviation identification. For best results, the smoothing model should be tailored to the data under

study.

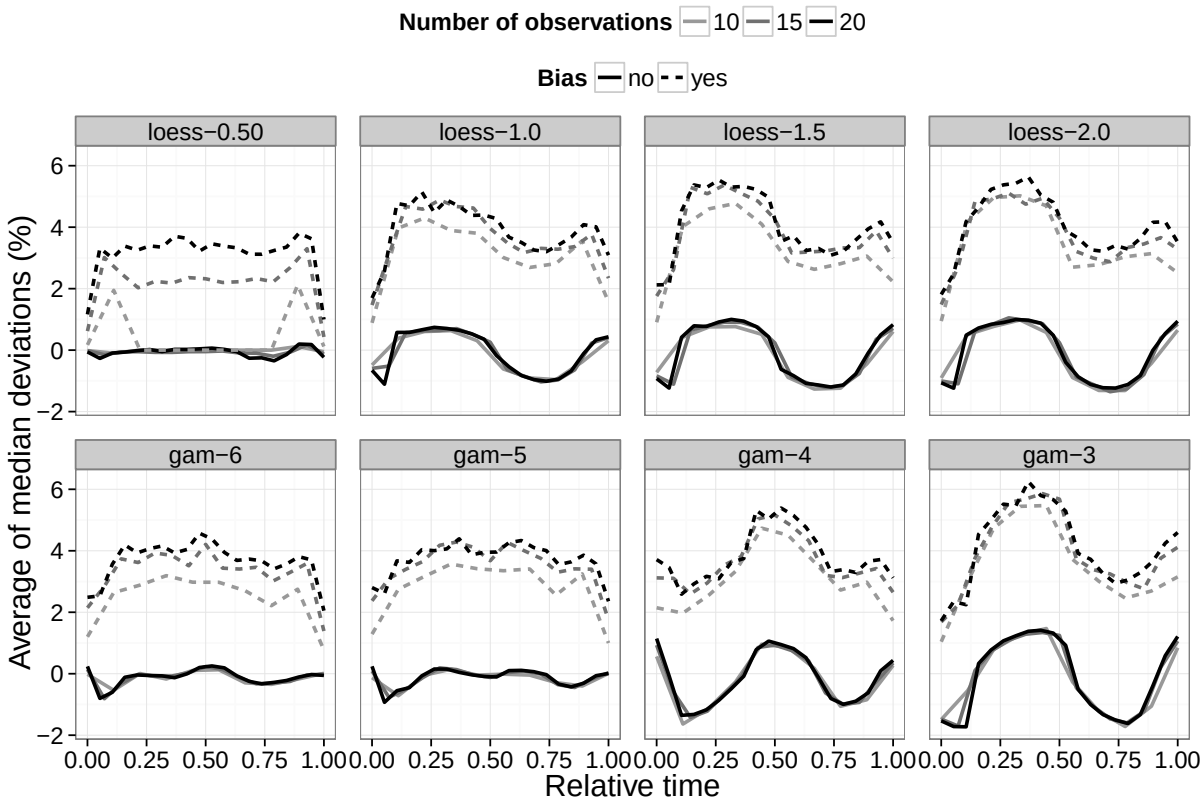


Figure 4.6: Average bias as a function of smoothing model. Lines represent averages of simulated median relative deviations from smoothing fits. Dashed lines are used to distinguish timepoints simulated with a 5% bias. The gam-6/loess-0.5 models correspond to a greater degree of smoothing in comparison to gam-3/loess-2.0. The number of observations corresponds to the number of timepoints in the generated metabolic trends. Although the number of compounds per time-course set was varied in the simulation, these were found to have no impact on average trends.

Apart from smoothing model, the number of observations over the course of a culture was also found to have an influence on deviation estimation (Figure 4.6). Increased sample

frequency yielded a more accurate deviation estimate for biased timepoints. However, the net impact of having a greater number of observations remained quite small. For gam-5, for example, a true bias of 5% was estimated as approximately 4% with 15 or 20 observations and closer to 3.5% with only 10 observations. Further simulations on lower observation numbers suggested that comparable performance could be attained down to 8 observations before degrading to a significant degree (data not shown). As batch processes may be operational for as few as 5 days, this translates to a required sampling frequency of two samples a day. Since 12 hour sampling may not always be practical, the effect of a staggered sampling on the correction algorithm was also investigated. With gam-5 smoothing, little to no difference was observed between even 12 hour sampling and a routine where 2 samples are taken 8 hours apart, followed by a break of 16 hours (data not shown).

Confidence intervals

The variability of median deviations is particularly important for the selection of a correction threshold. The threshold must be high enough to avoid correcting deviations due to random measurement noise while remaining sensitive to systematic sources of error. Empirical 90% confidence intervals were constructed from the simulated data by excluding the 5% highest and 5% lowest median deviations at each timepoint (Figure 4.7). Between the number of compounds and the number of observations, only the number of compounds was found to have an effect on confidence interval width. Naturally, the observation of more compounds reduced the impact of measurement noise and allowed for a more robust median estimate. However, the simulation of more compounds assumed equal quantification quality. If the number of observed compounds is increased by profiling highly convoluted

or otherwise poorly quantifiable compound resonances, the beneficial impact is likely to be limited.

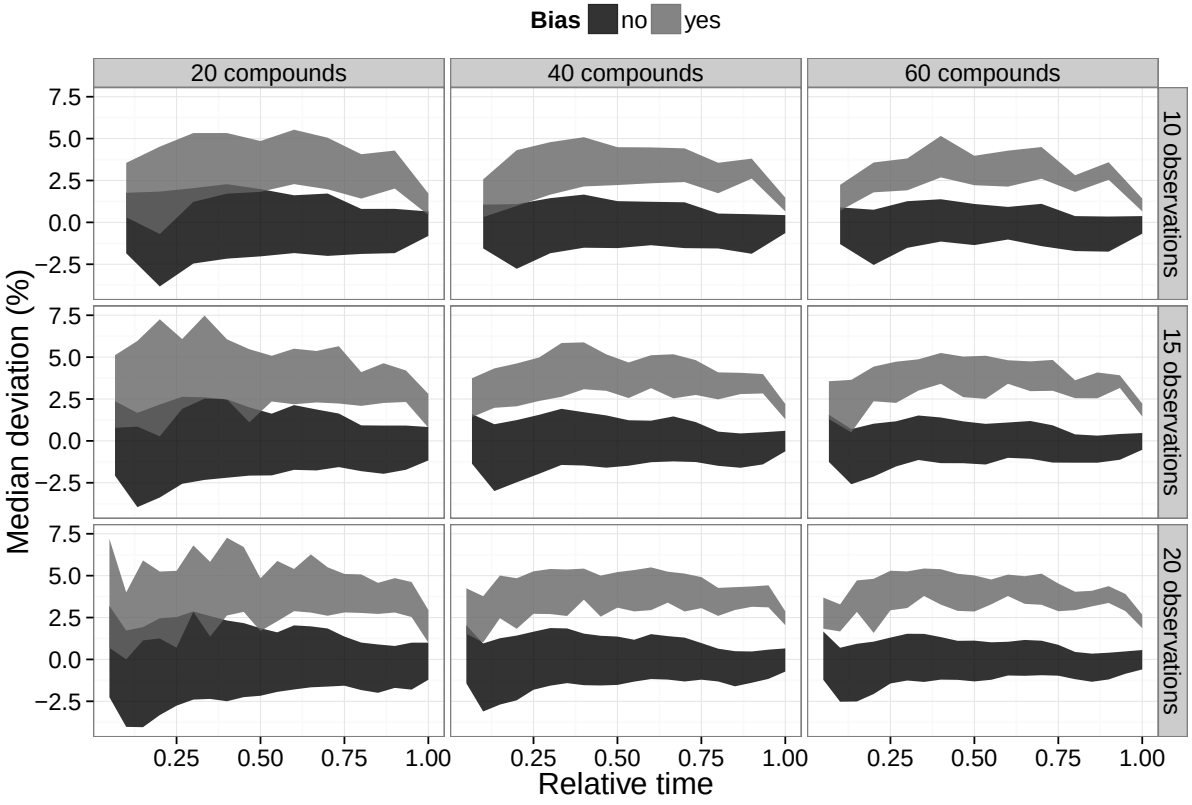


Figure 4.7: 90% confidence interval around median deviations. Empirical 90% confidence intervals were constructed from the simulated data by excluding the 5% highest and 5% lowest median deviations at each timepoint. Lighter grey colour is used to distinguish timepoints simulated with a 5% bias.

Based on the results, the observation of 40 compounds at 10 timepoints (typical of the data obtained in our lab) will exhibit a natural variation in median deviation of ap-

proximately 2–2.5%. Thus, deviations beyond this threshold have a high probability of occurring due to a source of bias such as internal standard addition or quantification. The results also show that a 5% bias is more likely to be identified as anywhere between 2.5–5% (with further reduction in performance at earlier timepoints), meaning that a subtraction of the estimated median deviation is more likely to dampen the bias, rather than remove it. Reduced performance at the end-points reflects the relative lack of trend data and can be ameliorated by adding replicates or extending the observation time beyond the span of direct interest.

Simulation extension

To determine how robust the correction method is to changes in the underlying data, four modifications to the simulated data were considered. The ratio of decreasing to increasing trends (intially taken as 35%:30% based on our cell culture data) was set to 60%:5% as well as 5%:60%. Despite these dramatic shifts, both average bias and confidence interval trends remained very similar to those presented in Figures 4.6 and 4.7. The only exception was at the end points, where lower concentration magnitudes resulted in much more variable relative deviations. Since increasing and decreasing trends reach their minimum concentrations at different endpoints, the overall effect on median relative deviations is not pronounced when the two trends are balanced in number. However, the extreme case of a 12:1 imbalance between increasing and decreasing trends resulted in larger variability ranges at time-course edges. With 60% of the trends increasing, the bias threshold at early timepoints increased from 2.5% to 5%. With 60% of the trends decreasing, the bias threshold increased at late timepoints but did not go beyond the overall average of 2.5% (as the

threshold at these timepoints was already low). The difference between the two conditions can be explained by the fact that all increasing trends start at or very close to 0, while only some of the decreasing trends reach such low concentrations. Since a more balanced proportion of increasing and decreasing trends is expected in real data, the overall effect would be minimal. Two other conditions – increasing the net concentration changes of decreasing trends (by doubling the proportion of compounds with large relative changes) and increasing the variability of observations (by doubling the standard deviation of highly variable compounds) did not appear to have any impact on the threshold calculation. For all conditions, gam-5 smoothing remained the best choice.

Taken together, these results suggest that a bias threshold of approximately 2.5% using gam-5 smoothing would be an adequate default choice for diverse data sets. Beyond cell culture applications, we predict the bias correction algorithm to be just as useful for other time-course metabolomic data. One such example is biofluid analysis in toxicology. The Consortium for Metabonomic Toxicology (COMET) has already established a large collection of time-course urine samples that meet the requirements for systematic error correction (Lindon et al. 2005). While the proposed correction is not designed to replace standard normalization techniques, it can build on the development of recent smoothing spline techniques (Berk et al. 2011) and serve a complementary role in the identification of spurious results. Further extension to mass spectrometry (MS) methods is also likely to be fruitful. Techniques such as multiple reaction monitoring (MRM) are commonly used for pharmaceutical and toxicological metabolomics (Kitteringham et al. 2009; Steven H et al. 2001) and suffer from similar dilution effects as NMR (exacerbated by the need for more sample manipulation such as liquid extraction steps). The correction of system-

atic biases may serve to reduce the relative standard deviations of quantified compound concentrations.

4.5 Conclusions

The growing popularity of quantitative metabolomics for time-course applications presents a new context for data processing and acquisition. While this work deals primarily with the correction of internal standard quantification in cell culture data, it's not difficult to imagine similar approaches applied to other analytical methods. Improvements in accuracy, precision, and analysis speed can be best achieved by leveraging the replication inherent to the parallel observation of multiple metabolite trends. The algorithm presented in this work took advantage of inherent autocorrelation to identify and correct systematic bias originating from internal standard addition and quantification. The gam-5 model was identified as the best smoothing function for the task, with the ability to detect a bias greater than 2.5% across most of a culture's time-course. The simulation framework followed the context-driven approach by capturing the key elements of a cell culture time-course. Although the presented validation has focused on trends typically observed in our lab, full code has been provided to allow rapid adaptation to user needs.

Availability and requirements

Project name: metcourse

Project home page: https://github.com/ssokolen/metcourse

Operating system: Platform independent (tested on Linux)

Programming language: R (version 3.2.1)

Other requirements: R packages – dplyr (version 0.4.2), mgcv (version 1.8-6)

License: Apache (version 2.0)

Any restrictions to use by non-academics no

Author's contributions

SS designed and implemented the algorithms and carried out the data collection. MA assisted in data interpretation and manuscript preparation. All authors read and approved the final manuscript.

Acknowledgements

This work was supported in part by an NSERC Canada Graduate Scholarship to SS as well as NSERC Discovery and Strategic Project grants to MA. The authors thank Marco Quattrociocchi for his assistance with cell culturing, Sandi Yen for her critical review of the manuscript, and Larry Campbell for discussion of MS techniques.

Chapter 5

Metabolic Flux Validation

The development of robust variance estimates and the correction of quantification bias opened the door to accurate estimation of intracellular flux from extracellular concentration data using metabolic flux analysis. Previously published MFA models were applied to cultures of CHO and Sf9 (*Spodoptera frugiperda* insect) cells. Whereas no errors were identified through the standard detection of gross measurement errors, some of the results featured unexpectedly large material imbalances. This prompted an investigation into alternative forms of validation that could be used to verify accuracy and precision. The result was a combined simulation and t-test strategy that could address both the impact of measurement uncertainty and lack of model fit. This work was submitted to the journal BMC Systems Biology, with the manuscript is presented below in its original form. Additional files meant to support the manuscript that would be available online are provided in Appendix C.

Identifying model error in metabolic flux analysis – A generalized least squares approach

Stanislav Sokolenko, Marco Quattrociocchi, and Marc G. Aucoin

Department of Chemical Engineering, Waterloo Institute for Nanotechnology, University of Waterloo.

5.1 Abstract

5.1.1 Background

The estimation of intracellular flux through traditional metabolic flux analysis (MFA) using an overdetermined system of equations is a well established practice in metabolic engineering. Despite the continued evolution of the methodology since its introduction, there has been little focus on validation and identification of poor model fit outside of identifying "gross measurement error". The growing complexity of metabolic models, which are increasingly generated from genome-level data, has necessitated robust validation that can directly assess model fit.

5.1.2 Results

In this work, MFA calculation is framed as a generalized least squares (GLS) problem, highlighting the applicability of the common t-test for model validation. To differentiate between measurement and model error, we simulate ideal flux profiles directly from the

model, perturb them with estimated measurement error, and compare their validation to real data. Application of this strategy to an established CHO cell model shows how fluxes validated by traditional means may be largely non-significant due to a lack of model fit. With further simulation, we explore how t-test significance relates to calculation error and show that fluxes found to be non-significant have 2-4 fold larger error (if measurement uncertainty is in the 5%-10% range).

5.1.3 Conclusions

The proposed validation method goes beyond traditional detection of "gross measurement error" to identify lack of fit between model and data. Although the focus of this work is on t-test validation and traditional MFA, the presented framework is readily applicable to other regression analysis methods and MFA formulations.

5.2 Background

As the metabolic phenotype of the cell, the flow of material through intracellular reactions (or metabolic flux) represents the sum total of all underlying cellular processes. The accurate determination of metabolic flux is becoming increasingly important for assessing the impact of metabolic engineering or feeding strategies on cellular metabolism (Chen et al. 2016). In lieu of *in vivo* observation, the inference of intracellular fluxes is commonly accomplished through metabolic flux analysis (MFA). At its most basic, MFA refers to the process of modeling intracellular flux via a stoichiometric balance of metabolic reaction and transport rates (assuming a "pseudo steady-state" in the form of negligible molecule

accumulation) (Stephanopoulos et al. 1998). The original applications of the technique centered on using simple element balances as a means to correct unreliable measurements (Wang and Stephanopoulos 1983). However, the increasing availability of data from multi-omic technologies has led to the development of metabolic flux models that extend far beyond these foundations.

The basis of MFA is the stoichiometry matrix. In the typical arrangement, rows represent balances on molecular species, with each column encoding the stoichiometry of a reaction (see Stephanopoulos et al. 1998 for details). As cellular reaction networks generally have more reactions than species, the resulting stoichiometry matrix is typically underdetermined. The estimation of a single flux profile requires that the number of unknown reaction rates be equal to or less than the number of molecular species, and this has traditionally been accomplished by observing as many extracellular transport rates as possible. However, the growing availability of genomic data has opened the door to developing models that may contain thousands of reactions, complicating the calculation of a unique flux profile.

A considerable amount of metabolic information can be gathered without calculating a unique flux profile through constraint-based reconstruction and analysis (COBRA) methods. The combination of mass balance constraints from stoichiometric relations as well as other factors such as enzyme capacity and reaction thermodynamics can be used to generate a feasible solution space for cellular metabolism. If a unique flux profile is required, one can be estimated by assuming an objective function such as cell growth maximization. However, it is also possible to study the solution space directly (for a detailed review, see Bordbar et al. 2014). The popularity of COBRA methods has resulted in the develop-

ment of a large number of software packages that have considerably simplified analysis (see Lewis et al. 2012). However, the complexity of genome-scale models remains an on-going challenge.

Despite the recent advances, the process of translating genomic information to cellular reactions is still under development. Even the well-studied genomes of *Escherichia coli* and *Saccharomyces cerevisiae* had approximately 20% of their open reading frames (ORFs) uncharacterized as recently as 2010 (Dauner 2010) and the development of reaction networks requires a significant amount of curation (Boghigian et al. 2010; Dauner 2010; Maertens and Vanrolleghem 2010). Furthermore, the relation between the presence of a gene sequence and enzymatic activity is not always obvious (Maertens and Vanrolleghem 2010). A combined transcriptomic-metabolomic modeling study of *E. coli* has revealed the existence of redundant gene expression where no flux was observed (Shlomi et al. 2007). Meanwhile, a study of lysine-producing *Corynebacterium glutamicum* metabolism suggested that while the expression of some genes appears tightly coupled to metabolic fluxes, others can remain practically constant despite considerable changes in metabolic flux (Krömer et al. 2004). The popular Chinese Hamster Ovary (CHO) cell line has an added problem of high genetic variability that may question the generality of a given model (Feichtinger et al. 2016; Kaas et al. 2015). Taken together, these issues add a considerable amount of uncertainty to modeling efforts, especially for less studied expression systems.

The addition of isotopically labelled substrate and the analysis of resulting metabolites through ^{13}C-MFA can be a powerful means to gain better understanding of a metabolic system. But despite the ready availability of algorithms and software packages to assist with everything from identifying optimal labelling strategies to final analysis (as reviewed

in Antoniewicz 2015; Young 2014), ^{13}C-MFA is not always practical. Isotopic labelling is expensive, especially for large volume bioreactor cultivation, and can not be used to monitor ongoing production processes. Moreover, studying transient labelling patterns requires accurate intracellular metabolite quantification, which is not always straightforward (Mashego et al. 2007), and increased computational resources (Antoniewicz 2015).

As such, one approach to dealing with genome-scale model uncertainty and complexity has been to simplify the models to a level where they can be solved directly from measured extracellular transport rates (Quek et al. 2014; Quek et al. 2010), continuing the use of traditional overdetermined MFA. The simplification can be aided by software such as CellNetAnalyzer that can deal with both underdetermined COBRA models and overdetermined MFA formulations (Klamt et al. 2007; Klamt and von Kamp 2011). Recent developments have also led to an automation of the model simplification process (Erdrich et al. 2015). Despite increasing model size, overdetermined MFA has continued to see use over the last 10 years (Bernal et al. 2009; Carinhas et al. 2010; Carinhas et al. 2013; Niklas et al. 2012; Niklas et al. 2011; Priesnitz et al. 2012; Quek et al. 2010; Xing et al. 2011), especially for less commonly used cell lines that lack well curated genomic and transcriptomic data. However, the reduction of genome-levels models in this fashion is an inversion of the original MFA foundations. In contrast to the use of a simple, reductive model for the reconciliation of questionable data, it is the accuracy of the model that is becoming increasingly variable – making it necessary to rigorously assess the validity of model simplification.

A number of strategies are currently available for model validation. The stoichiometric matrix can be probed directly by checking its condition number (Vallino and Stephanopou-

los 1990) or by determining the sensitivity of calculated fluxes to measurement error (Goudar et al. 2009). The incorporation of measurement flux uncertainty allows the use of gross measurement error detection (van der Heijden et al. 1994), which identifies whether deviations between observed and fit data are normally distributed through a χ^2-test. While useful for identifying singular errors of large magnitude, this statistic does not asses the overall quality of fit – errors may be unreasonably large while remaining normally distributed. Despite the increasing consideration of confidence intervals around calculated fluxes in recent studies (Antoniewicz et al. 2006; Leighty and Antoniewicz 2011), the question of whether a set of data fits a given metabolic model has thus far remained open.

In this work, we propose the use of a standard t-test as a natural extension of the least-squares calculation that underpins traditional MFA calculation. Applying MFA to a Chinese Hamster Ovary (CHO) cell culture, t-tests were used to determine whether each calculated flux could be deemed sufficiently distinct from zero. Once non-significant fluxes were identified, we explored whether the uncertainty in calculated fluxes could be explained by measurement uncertainty alone, or if a lack of model fit could be to blame. To do this, the solution space of the stoichiometric model was constrained by observed flux ranges and hypothetical flux profiles were generated directly from the model. The profiles were perturbed by measurement error and collected to establish a baseline of calculated flux significance given perfect model fit.

5.3 Methods

5.3.1 Theoretical principles[1]

The material balance on molecular species that forms the basis of MFA is typically expressed as

$$Sv = 0 \tag{5.1}$$

where S is the stoichiometric matrix and v is the vector of fluxes that correspond to reactions defined by columns of S. This formulation proceeds from a pseudo steady-state assumption that changes in metabolite pools (as a result of cell division or other processes) are much smaller than metabolite production and consumption fluxes and can therefore be ignored. The Sv matrix can be be separated into $S_c v_c + S_o v_o$, where c stands for calculated flux and o for observed flux.

$$S_c v_c + S_o v_o = 0 \tag{5.2}$$

$$-S_o v_o = S_c v_c \tag{5.3}$$

Since v_o is a vector of observed data, $S_o v_o$ can be calculated directly. The dimension of S_c depends on how many fluxes can be observed, i.e., the length of v_o. S_c must have no more columns than rows to calculate a unique flux profile, although the observation of more fluxes (and the accompanying reduction in the number of S_c columns) is useful for error estimation [2]. Pooling cyclic or parallel pathways may be required in the initial formulation

[1] A more detailed discussion of the theoretical principles, including a worked example and some proofs, is available in Appendix C.1

of S to ensure the required form of S_c is obtained.

Assuming that an overdetermined form of S_c can be formulated (with sufficient information to calculate v_c), Equation (5.3) is equivalent to linear regression and can be solved in a similar fashion.

Linear regression		MFA	
$y = X\beta + \varepsilon$	(5.4)	$-S_o v_o = S_c v_c + \varepsilon$	(5.5)
$\hat{\beta} = \left(X^T X\right)^{-1} X^T y$	(5.6)	$\hat{v}_c = -\left(S_c^T S_c\right)^{-1} S_c^T S_o v_o$	(5.7)

With this formulation, ε represents the deviation between observed and calculated fluxes that may be the result of either measurement error or lack of model fit. Equation (5.7) assumes ε is independently and identically distributed, which is unlikely to be the case. The variance-covariance matrix $\mathrm{Cov}(\varepsilon)$ can be expressed as a scalar σ^2 multiplied by a matrix of relative covariance terms V, i.e., $\mathrm{Cov}(\varepsilon) = \sigma^2 V$. If observed fluxes do not covary and have equal variance, then $V = I$, where I is the identity matrix. Otherwise, Equation (5.5) needs to be rescaled by the matrix square root of V. Taking $V = PP$, the scaled form of Equation (5.5) is:

$$-P^{-1} S_o v_o = P^{-1} S_c v_c + P^{-1}\varepsilon \tag{5.8}$$

[2]It is typically assumed that S_c is sufficient for the estimation of all v_c values. However, failure to observe a key metabolite may result in a case where not all values of v_c can be estimated despite S_c appearing determined or overdetermined. See van der Heijden et al. 1994 for details on stoichiometry matrix classification.

where $P^{-1}\varepsilon$ now satisfies the assumptions of linear regression. Formally, this is equivalent to generalized least squares (GLS) regression, however, incorporating P^{-1} directly into each term allows the use of all ordinary least squares techniques. Letting $P^{-1}S_o = S'_o$, $P^{-1}S_c = S'_c$, and $P^{-1}\varepsilon = \varepsilon'$:

$$\hat{v}_c = -\left(S_c'^T S'_c\right)^{-1} S_c'^T S'_o v_o \tag{5.9}$$

The calculation of P^{-1} requires the estimation $\mathrm{Cov}(\varepsilon)$ from the variance of observed fluxes. Calculating the covariance-variance matrix of both sides of Equation (5.5):

$$\mathrm{Cov}(-S_o v_o) = \mathrm{Cov}(S_c v_c + \varepsilon) \tag{5.10}$$

$$\mathrm{Cov}(\varepsilon) = S_o \mathrm{Cov}(v_o) S_o^T \tag{5.11}$$

Since $\mathrm{Cov}(\varepsilon) = \sigma^2 V$ for any value of σ, σ is set to 1 so that $V = \mathrm{Cov}(\varepsilon)$. In practice, $\mathrm{Cov}(v_o)$ need only capture the relative magnitudes of observed flux variances as $\hat{\sigma}$ is estimated during regression. Balances around molecular species that do not include an observed flux v_o will have a row of zeros in $\mathrm{Cov}(\varepsilon)$, which prevents the calculation of a matrix inverse (required to get P^{-1}). Although this mathematically equates to a variance of zero for those balances, a better interpretation is that there is an unknown variance around the "observation" of no net flux. The simplest solution is to add a small non-zero value to each diagonal entry of $\mathrm{Cov}(\varepsilon)$, representing the confidence of the calculated fluxes being fully balanced. If there is more uncertainty around some balances than others, this information could be encoded in the magnitude of the added variance. P can then be calculated via a

matrix square root of estimated Cov(ε). Since a variance (covariance) matrix is positive semi-definite, P is known to be unique.

Whereas calculated fluxes $\hat{v}_c$ are commonly estimated using a very similar "weighted" least squares approach, the use of validation methods that are part of the regression framework have yet to be explored. The common χ^2 test can still be used to detect gross measurement errors in estimated residuals ($\hat{\varepsilon}$); however, the validation of a regression model also requires the use of t-tests to ensure the significance of calculated fluxes. Confidence and prediction intervals are also highly relevant to MFA. Estimated fluxes require a confidence interval to report the uncertainty of calculation, while a prediction interval around a predicted balance can be used to judge the validity of that balance being closed. The calculation of a t-statistic follows from normal regression:

Linear regression		MFA	
$t_{\hat{\beta}_i} = \dfrac{\hat{\beta}_i}{\text{se}(\hat{\beta}_i)}$	(5.12)	$t_{\hat{v}_{c,i}} = \dfrac{\hat{v}_{c,i}}{\text{se}(\hat{v}_{c,i})}$	(5.13)

Thus:

$$t_{\hat{v}_{c,i}} = \frac{\left(-\left(S_c'^T S_c'\right)^{-1} S_c'^T S_o' v_o\right)_i}{\hat{\sigma}\sqrt{(S_c'^T S_c')^{-1}_{i,i}}} \tag{5.14}$$

The estimated standard deviation of ε (or $\hat{\sigma}$) is calculated as follows:

$$\hat{\sigma}^2 = \frac{\sum(\hat{\varepsilon}_i')^2}{n_b - n_c - 1} \tag{5.15}$$

where:

$$\hat{\varepsilon}' = -S'_o v_o + S'_c \left(S'^T_c S'_c\right)^{-1} S'^T_c S'_o v_o \tag{5.16}$$

and n_b is the number of balances (rows of S'_c) while n_c is the number of fluxes to be calculated (columns of S'_c). If the model is correct and $\mathrm{Cov}(\varepsilon)$ was correctly estimated, $\hat{\sigma}^2$ should be approximately equal to 1. Once the t-value is calculated, a flux can be judged statistically significant if $|t_{\hat{v}_{c,i}}| \geq t_{\alpha/2, n_b - n_c - 1}$ where α is the significance level.

The identification of non-significant flux may be interpreted in two ways. The measurement error around observed fluxes may be too high to allow robust flux calculation. In that case, non-significant fluxes should be treated as having a flux of zero and excluded from the model or further analysis. Alternatively, non-significance may be the result of excess variability from a lack of fit between the model and observed data, requiring model correction. To distinguish between these cases, it is necessary to separate model error from measurement uncertainty. One way to accomplish this is to reduce measurement uncertainty through added replication, however, the required effort can make this approach practically infeasible. Another solution is to simulate a set of feasible fluxes directly from the stoichiometric model (and therefore free of model error) for comparison to the observed data.

The simulation of feasible fluxes can be simplified by eliminating flux equality constraints expressed by the stoichiometry matrix. Essentially, only $n_c - n_b$ fluxes have to specified in order to generate all the other values. More formally, the relationships between the fluxes can be succinctly summarized through the nullspace (or kernel) of S, which describes all flux balance conservations in the model. This makes it possible to

calculate all fluxes from a smaller set of variables referred to as the basis. Unlike fluxes, which must satisfy constraints imposed by $Sv = 0$, the basis can take any arbitrary value to generate fluxes that satisfy all required constraints. Expressed mathematically,

$$\text{Null}(S) = K \tag{5.17}$$

$$Kb = v \tag{5.18}$$

where b is a basis vector of any value with the same number of rows as columns of K. While all values of b satisfy $Sv = 0$, it is still necessary to constrain fluxes to a set of realistic values representative of a cell cultivation. The space of all feasible fluxes v can be constrained by defining upper and lower bounds on each observed flux:

$$\begin{aligned} & v = Kb \\ \text{subject to} \quad & K_i b \leq v_i + a \cdot \text{sd}(v_i) \\ & K_i b \geq v_i - a \cdot \text{sd}(v_i) \end{aligned} \tag{5.19}$$

where v_i is an observed flux, K_i is the corresponding row of K, and a is a scaling constant that can be set to $t_{\alpha/2,df}$ to specify a confidence interval around v_i. As the basis solution space is only constrained by inequalities, it is readily amenable to stochastic sampling. All values of v that satisfy Equation (5.19) represent feasible fluxes that would perfectly satisfy the stoichiometric model while remaining within measurement uncertainty of real observations. If the resulting space is infeasible, then the observed data does not fit the specified model. Otherwise, a random sample of feasible fluxes can be taken for comparison to observed results. If the addition of measurement error to simulated fluxes results in less

uncertainty than from observed results, then model error is to blame.

5.3.2 Cell culture

CHO-BRI cells were grown in a 3 L bioreactor (Applikon Biotechnology Inc., Foster City, CA) in serum-free BioGro-CHO media (BioGro Technologies Inc., Winnipeg, Canada) with an in-house amino acid supplement (manuscript submitted). The culture was seeded at $0.3 \cdot 10^6$ cells/ml with a working volume of 2 L. Temperature, pH, dissolved oxygen, and agitation speed were held at 37°C, 7.4, 50%, and 120 RPM respectively. Samples were taken three times a day for offline analysis. Cell density was determined using a Coulter Counter Z2 (Beckman Coulter, Miami, FL) calibrated to results from trypan blue exclusion analysis. Aliquots were centrifuged, with the supernatant collected and stored at -80°C until NMR analysis. Dry cell mass was calculated by vacuum filtering 15 mL of cell culture through a type A/D glass filter (Pall Corporation, Port Washington, NY) and weighing the filter after drying it for 24 hours at 50°C.

5.3.3 Metabolite quantification

NMR spectra acquisition, metabolite quantification, and internal standard correction are described by Sokolenko and Aucoin (2015). In brief, samples were scanned on a Bruker Avance 600 MHz spectrometer using the first increment of a 1D-NOESY pulse sequence with metabolite quantification carried out using Chenomx NMR Suite 8.1 (Chenomx Inc., Edmonton, Canada). GlutaMAX™ was added manually to the software library using the Chenomx NMR Suite's 'compound builder' tool. All compounds were profiled in triplicate.

Ammonia measurements were taken using an Orion Star™Plus ISE Meter (Thermo Fisher Scientific, Waltham, MA).

5.3.4 MFA model

A CHO cell MFA model was taken from the work of Altamirano et al. (2001). New transport fluxes were added for acetate, formate, pyruvate, citrate, malate, pyroglutamate, and GlutaMAX™ (the fluxes of which could all be observed via NMR). The transport of GlutaMAX™ was grouped together with the conversion of the dipeptide into glutamine and alanine. The transport of cystine was grouped together with the reduction of cystine into cysteine. A new reaction was added for the conversion of glutamate into pyroglutamate (Kumar and Bachhawat 2012) (via a number of possible enzymatic and non-enzymatic reactions). New reactions were also added for acetyl-CoA hydrolase and formate-tetrahydrofolate ligase to explain acetate and formate production. A full list of reactions and an outline of metabolite flow can be found in Appendix C.2. As in the original formulation, a number of unbalanced species were removed from the model before analysis, including O_2, CO_2, ATP, NADH, NADPH, and FADH (NADH and NADPH were later reintroduced in a modified form of the model).

5.3.5 Flux estimation

Metabolite and cell concentration timecourse data was fit by a regression spline with 4 cubic basis functions, provided by the gam function (Wood 2011) in the R programming language (R Core Team 2016). Measurement error was estimated by calculating the vari-

ance of observation deviation from the fit. 1000 predicted concentration timecourses were simulated for each trend by adding normally distributed error corresponding to the sum of regression and measurement variance. A new regression split fit was calculated for each of the simulated timecourses. Metabolite transport fluxes were calculated by dividing the derivative of the metabolite concentration fit by cell concentration ($v_o = \frac{1}{X}\frac{\mathrm{d}C_o}{\mathrm{d}t}$). The mean and variance of the simulated fluxes at each time-point were used for all MFA analysis. Biomass fluxes were calculated as by Altamirano et al. (2001), with the exception that dry cell mass measured to be 0.24 mg/10^6 cells. A single mid-exponential time-point of 66 hours was chosen for MFA analysis to fulfill pseudo steady-state conditions.

5.3.6 Implementation

All MFA calculations, validation, and sampling were carried out using the omfapy Python package, developed in-house. The package as well as analysis code is available on github (`https://github.com/ssokolen/omfapy`). Basic functionality was based on theoretical principles presented by Stephanopoulos et al. (1998). Sampling of a feasible flux space was implemented using the random direction algorithm (Smith 1984) as well as the mirror algorithm presented by Van den Meersche et al. (2009). Although slower than the random direction algorithm, the mirror algorithm was able to generate more even coverage of the sampling space.

5.4 Results

5.4.1 Identification of model error

Observed uptake fluxes and their corresponding coefficients of variation 66 hours post inoculation are shown in Table 5.1, with overall metabolite concentration profiles and cell density in Figure 5.1. As usual for CHO cells, the metabolic profile was dominated by large fluxes of glucose and lactate. Considerable fluxes of alanine, GlutaMAX^{TM}, ammonia, and glutamine were also observed. The median coefficient of variation was found to be 9.3%. Although this was similar to previously reported estimates for concentration quantification via NMR (Sokolenko et al. 2014), incorporating the uncertainty of derivative calculation resulted in a somewhat larger probability of high variance values. As described by Sokolenko et al. (2014), the singularly high variability of glutamate flux was primarily due to its low concentration and heavy spectral convolution.

Table 5.1: Observed uptake fluxes and coefficients of variation (standard deviation of flux divided by flux) 66 hours post inoculation.

	Flux ($\frac{\text{nmol}}{10^6\text{cells}\cdot\text{h}}$)	CV (%)
Acetate	−1.03	5.08
Alanine	−33.95	3.32
Ammonia	−17.65	23.10
Arginine	2.52	16.33
Asparagine	2.21	7.64
Aspartate	2.14	7.09
Carbohydrates	−2.13	12.25

(Continued on next page)

Table 5.1 – continued from previous page

	Flux ($\frac{\text{nmol}}{10^6\text{cells}\cdot\text{h}}$)	CV (%)
Citrate	−1.56	7.14
Cystine	0.33	19.04
DNA	−0.31	13.15
Formate	−7.52	2.06
Glucose	161.87	2.89
Glutamate	−0.17	213.18
Glutamax	17.98	10.69
Glutamine	7.35	12.48
Glycine	−2.25	8.79
Histidine	1.02	14.92
Isoleucine	1.52	8.13
Lactate	−283.53	3.19
Leucine	2.66	9.47
Lipids	−1.36	14.86
Lysine	1.80	8.05
Malate	−0.40	13.78
Methionine	0.89	6.44
Phenylalanine	1.19	7.04
Proline	1.94	9.17
Protein	−32.69	13.11
Pyroglutamate	−3.86	3.86
Pyruvate	−2.62	5.74
RNA	−0.89	13.77
Serine	2.64	12.36
Succinate	−0.15	15.52
Threonine	1.70	11.45
Tryptophan	0.34	17.40
Tyrosine	1.11	6.57
Valine	2.24	5.26

The incorporation of the observed fluxes into the MFA model showed no issues using typical metrics. The condition number of the reduced stoichiometry matrix was consider-

Figure 5.1: Observed time-course trends. Panels depict **a)** metabolites that changed by more than 50% of their maximum concentration, **b)** those that changed by less than 50%, and **c)** cell density. All metabolite concentrations are expressed as fractions of their maximum value. Curves were calculated from cubic regression spline fits constrained to 4 basis functions. Grey area designates 99% prediction interval used for sampling.

ably below 1000 and the χ^2 p-value was 0.93, indicating little evidence of gross measurement error. However, t-test analysis on the calculated fluxes using the GLS framework revealed that only 15 of 47 fluxes were statistically significant (at the standard 5% significance level). The statistically significant fluxes were primarily those that related to glycolysis – offering only a shallow look at cellular metabolism. All of the TCA and many of the amino acid degradation fluxes were deemed non-significant. To determine whether measurement variability or model error was to blame, 100 flux profiles were sampled from the stoichiometric matrix bounded by 99% confidence intervals on the measured fluxes (fluxes generated directly from the model in this way will be referred to as "balanced"). 99% intervals were chosen to include practically all possible flux values. The sampled fluxes had good coverage of the constraint space, suggesting that the model was flexible enough to fit fluxes similar to those observed. Each balanced flux profile was then perturbed 100 times using normally distributed noise generated from observed flux standard deviations. The result was 10 000 sets of fluxes subject to observed measurement error but no model error.

Figure 5.2 compares the percentage of simulated (balanced) fluxes found to be non-significant to the results from observed data. The simulation revealed that approximately

half of the calculated fluxes (and all TCA fluxes) are entirely non-significant even when there is no model error (Figure 5.2B). Many of the other fluxes were only significant for 50% of the simulations or fewer. The lack of significance showed that the model was incapable of providing high confidence results for the collected data. Along with the overall low significance, evidence of model error could also be observed. Focusing on approximately 20 of the lowest magnitude fluxes, all were deemed to be non-significant based on the observed data. Comparing the simulated data, the same fluxes were rejected as non-significant 50%-95% of the time. Taken together, the probability of all the low magnitude fluxes being observed as non-significant is extremely low, giving strong indication of poor fit beyond the effect of measurement error alone, i.e., as a result of model error. Although model correction is outside the scope of this work, the proposed methodology was successful in identifying a considerable degree of uncertainty overlooked by commonly used validation methods.

5.4.2 Effect of measurement noise

An extended simulation was carried out to determine whether the lack of statistical significance was due to measurement variability. The flux constraints were extended beyond 66 hours post inoculation to consider the broader applicability of the model. 99% confidence intervals were generated for all fluxes 18-80 hours post inoculation with the minimum and maximum values for each flux used to bound the flux solution space. 100 balanced flux profiles were generated with 100 sets of measurement error drawn from a normal distribution using 5%, 10%, 15%, and 20% coefficients of variation for each flux. The 45 calculated fluxes spanned more than 3 logarithms of values from approximately 0.1 $\frac{\text{nmol}}{10^6\text{cells}\cdot\text{h}}$

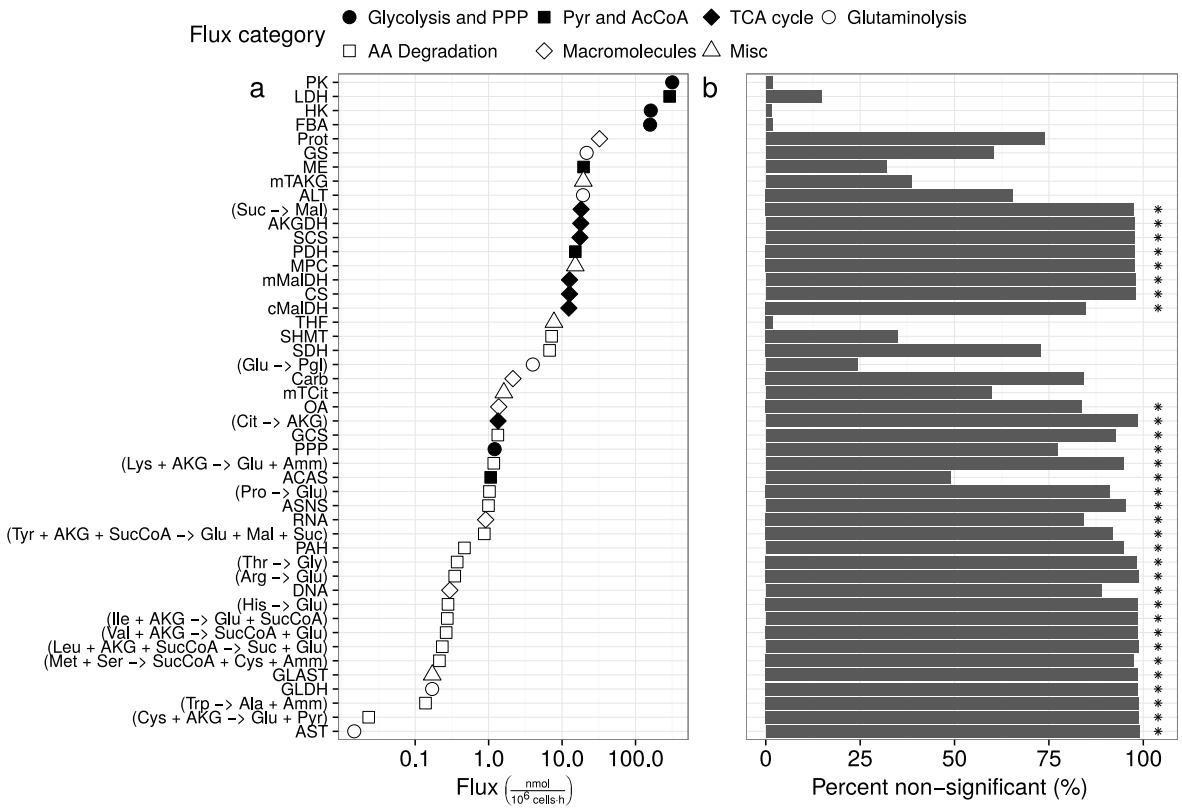


Figure 5.2: Comparison of flux rejection between observed and simulated data. Panels depict **a)** calculated flux magnitude and **b)** the percent of simulations in which the calculated fluxes were found to be non-significant (with asterisks indicating fluxes calculated to be non-significant using observed data). Simulated data was drawn from the stoichiometric model described in the Materials and Methods section, constrained by 99% confidence intervals on fluxes observed at 66 hours post inoculation. 100 balanced flux profiles were generated with 100 random generated sets of measurement error applied to each. See Appendix C.2 for reaction definitions.

to 400 $\frac{\text{nmol}}{10^6\text{cells}\cdot\text{h}}$ (Figure 5.3A). Fluxes had variable magnitudes across the simulations, so all analysis was performed as a function of flux rank, where a rank of 1 indicates the smallest magnitude flux in a given flux profile.

All the simulated flux profiles were subject to a χ^2 test, with only 5% of the simulations rejected (equal to the false positive rate). The remainder of the fluxes are show in Figure 5.3. As the simulated fluxes included both observed and calculated values, a percent error could be calculated for each calculated flux. Despite passing the χ^2 test, most fluxes were characterized by median errors of 10%-20% (Figure 5.3B), increasing with measurement variability. It should be noted that the median is a relatively conservative statistic. By definition, half of the calculated fluxes featured much greater errors than the reported values. The pronounced jump in error for flux ranks of 36 to 44 was traced to the TCA fluxes, which had high error despite large flux magnitudes. Similar to median error, the percentage of fluxes identified as non-significant increased with measurement variability (Figure 5.3C). However, even measurements with 5% coefficient of variation resulted in rejection rates of 50% or more across practically all fluxes. The TCA fluxes in particular (ranks 36 to 44) were rejected as non-significant 75% of the time or more (at all levels of measurement variability). The high level of flux rejection at low levels of measurement variability suggested the uncertainty in MFA calculation using observed data was primarily due to model structure rather than the uncertainty of observed data. Despite passing traditional validation tests, the simulation of stoichiometrically balanced fluxes revealed that the model is incapable of explaining observed metabolic profiles with an acceptable degree of confidence.

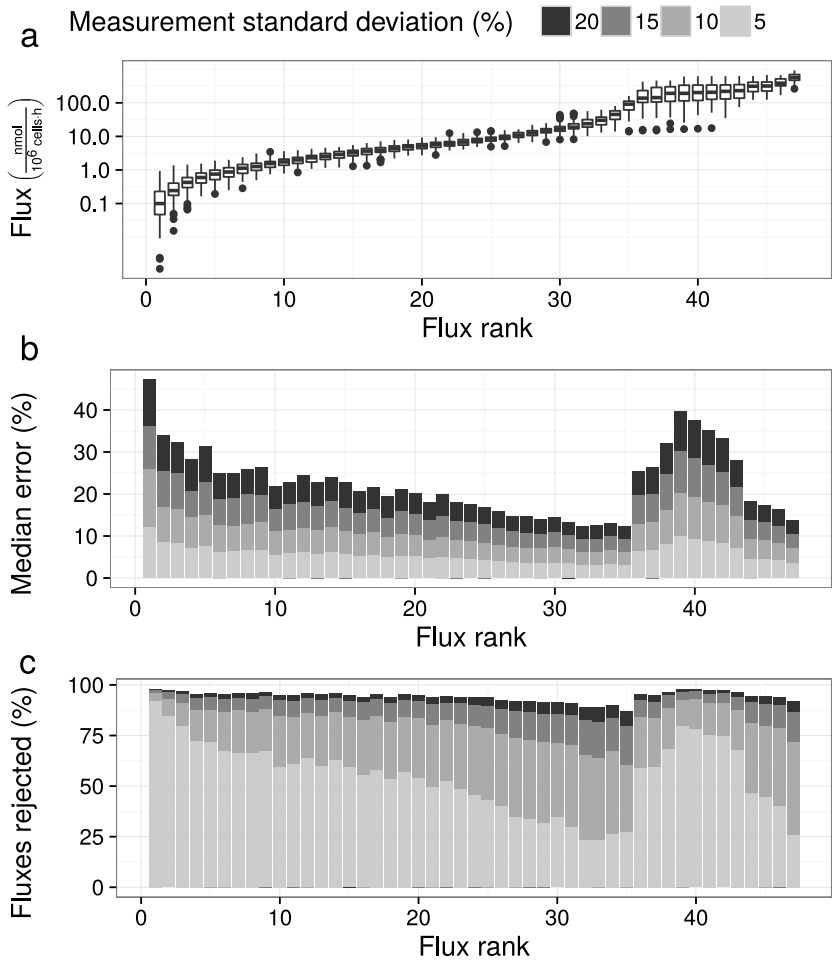


Figure 5.3: Comparison of fluxes simulated with different measurement errors. Panels depict **a)** flux magnitude, **b)** median error, and **c)** percent non-significance. Simulated data was drawn from the stoichiometric model described in the Materials and Methods section, constrained by 99% confidence intervals on fluxes observed between 18 and 80 hours post inoculation. 100 balanced flux profiles were generated with 100 random generated sets of measurement error applied to each. Each balanced flux profile was ordered according to increasing absolute flux magnitude to generate an associated rank from 1 to 45.

5.4.3 Effect of model structure

To test the influence of model structure on the significance of calculated fluxes, we simulated the effect of a broken electron transport chain – allowing a closed balanced on NADH and NADPH. Essentially, NADH and NADPH were reintroduced into the model and assumed to be balanced by the defined stoichiometric relations. Although arbitrary, this assumption is consistent with largely anaerobic metabolism of CHO cells (termed the "Warburg Effect") and allowed the addition of balances around intermediate compounds participating in many reactions. Incorporating the modified model into analysis of the observed fluxes at 66 hours post inoculation revealed no sign of gross measurement error (χ^2 p-value of 0.91) and decreased the number of non-significant fluxes from 32 (of 47) to 16. As before, 10 000 sets of fluxes were simulated from 99% confidence intervals around the observed measurement fluxes, subject to observed measurement error (Figure 5.4). In comparison to Figure 5.2B, Figure 5.4B reveals a considerable increase in significance across a large number of fluxes, consistent with the idea that model structure plays an important role in uncertainty around calculated fluxes. The impact was particularly drastic for TCA fluxes, most of which changed from entirely non-significant to significant. Despite the improvement in model fit, some model error could also be observed – too many of the low magnitude fluxes calculated from observed data were found to be non-significant when compared to the simulated results.

The modified model was also tested with an extended simulation (Figure 5.5). As with the original model, 99% confidence intervals were generated for all fluxes 18-80 hours post inoculation with the minimum and maximum values for each flux used to bound the flux solution space. The most pronounced impact of the modification was on the rate of flux

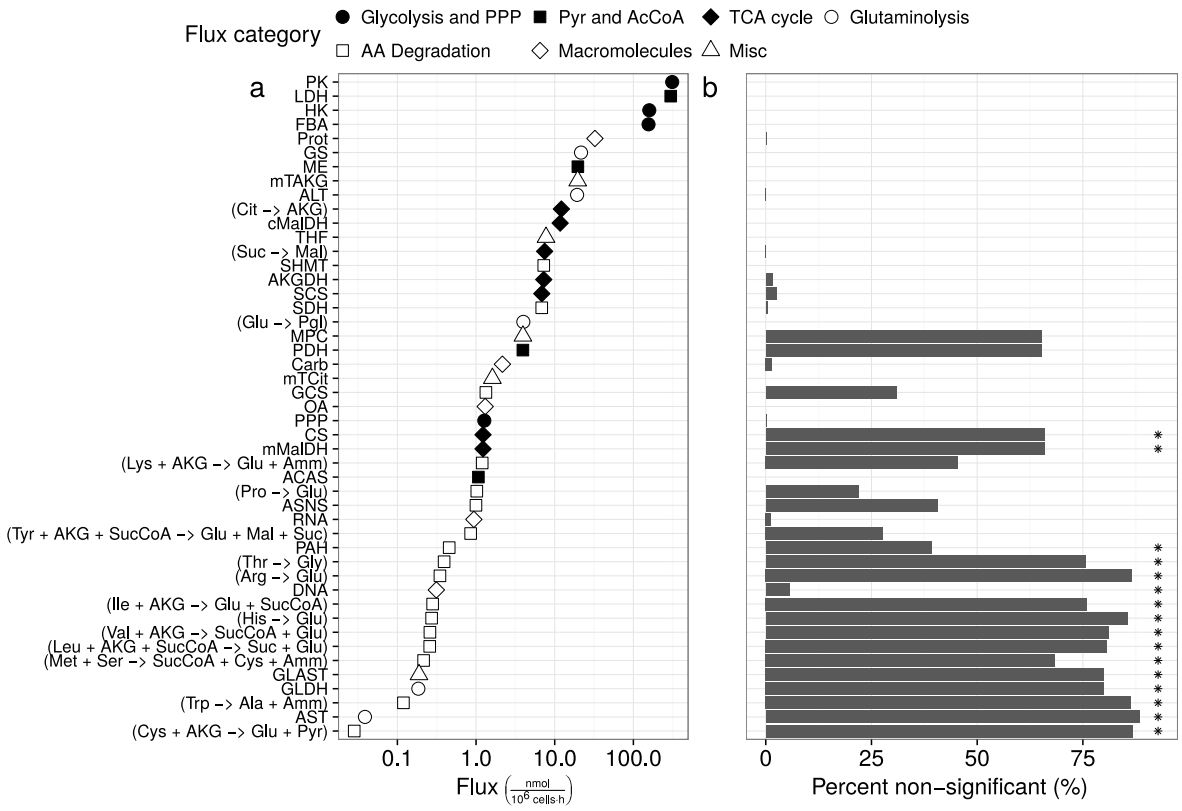


Figure 5.4: Comparison of flux rejection between observed and simulated data following model modification. Panels depict **a)** calculated flux magnitude and **b)** the percent of simulations in which the calculated fluxes were found to be non-significant (with asterisks indicating fluxes calculated to be non-significant using observed data). Simulated data was drawn from a modification of the stoichiometric model described in the Materials and Methods section (with balances on NADH and NADPH), constrained by 99% confidence intervals on fluxes observed at 66 hours post inoculation. 100 balanced flux profiles were generated with 100 random generated sets of measurement error applied to each. See supplementary materials for reaction definitions.

rejection (Figure 5.5C). At 5% measurement variability, approximately two thirds of the fluxes were always significant. The remaining third of the lowest magnitude fluxes were significant at least 50% of the time. In comparison, none of the fluxes calculated with the original model were significant for more than 75% of the simulations. To get a better idea of how the t-test metric related to flux inaccuracy, median errors were separated for significant and non-significant fluxes. At 5% coefficient of variation, fluxes deemed statistically significant had a constant median error of less than 5% (with relation to flux rank), while non-significant fluxes had considerably higher errors (Figure 5.6). Increasing coefficients of variation resulted in dramatic increases in overall rates of flux rejection (Figure 5.5C). However, the median error of statistically significant fluxes also increased, diminishing the ability of the t-test metric to identify inaccuracy in higher magnitude fluxes (Figure 5.6). In comparison, the typical χ^2 test retained a 5% rejection rate for all measurement errors (equal to the false positive rate).

5.5 Discussion

Taken together, the results of the simulations suggest that both measurement uncertainty and model structure have an impact on MFA results that are not assessed by typical validation methods. The structure of the model may lead to a considerable amount of uncertainty around calculated fluxes despite a high level of measurement precision. Mathematically, this impact can be seen in the $(S_c'^T S_c')^{-1}$ term that stems from the variance of estimated regression parameters, i.e., $\text{Cov}(\hat{\beta})$. Less formally, it may be intuitive that a model featuring a balance on important intermediate metabolites such as NADH and

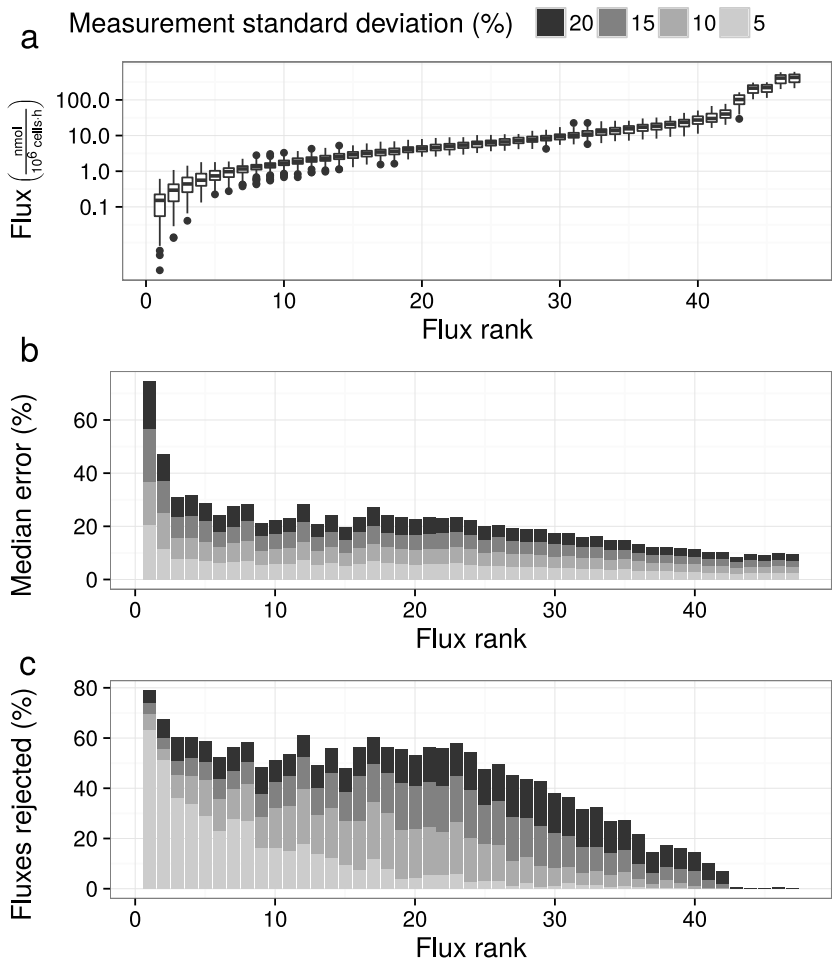


Figure 5.5: Comparison of fluxes simulated with different measurement errors following model modification. Panels depict **a)** flux magnitude, **b)** median error, and **c)** percent non-significance of fluxes simulated with different measurement errors. Simulated data was drawn from a modification of the stoichiometric model described in the Materials and Methods section (with balances on NADH and NADPH), constrained by 99% confidence intervals on fluxes observed between 18 and 80 hours post inoculation. 100 balanced flux profiles were generated with 100 random generated sets of measurement error applied to each. Each balanced flux profile was ordered according to increasing absolute flux magnitude to generate an associated rank from 1 to 45.

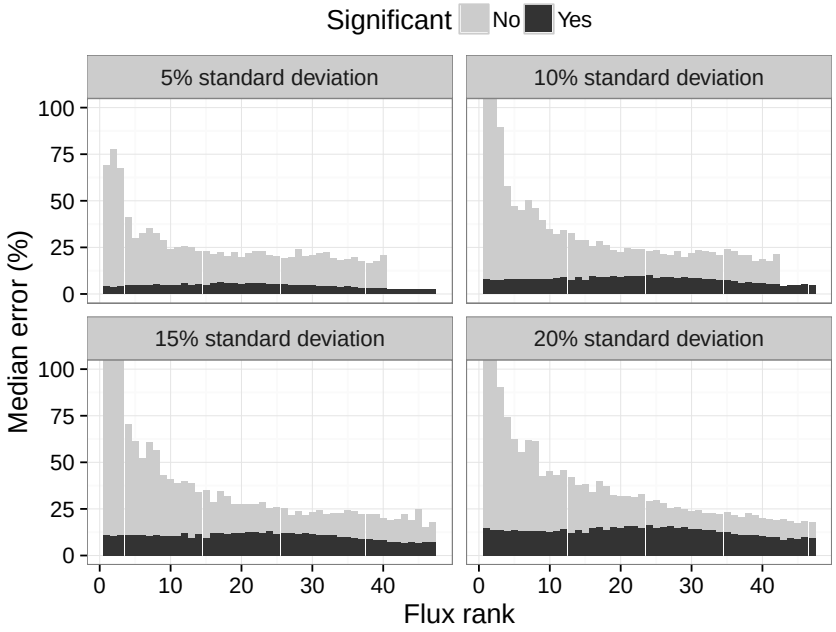


Figure 5.6: Comparison of median error of significant and non-significant fluxes (determined by t-test with $\alpha = 0.05$) simulated with different measurement errors. Simulated data was drawn from a modification of the stoichiometric model described in the Materials and Methods section (with balances on NADH and NADPH), constrained by 99% confidence intervals on fluxes observed between 18 and 80 hours post inoculation. 100 balanced flux profiles were generated with 100 random generated sets of measurement error applied to each. Each balanced flux profile was ordered according to increasing absolute flux magnitude to generate an associated rank from 1 to 45.

NADPH would be able to estimate intracellular fluxes with a greater degree of confidence than a model without the extra information afforded by the balance. Naturally, the addition of isotopically labelled substrates can add a much greater degree of certainty. Indeed, an important application of the proposed testing and simulation framework is to provide a rigorous assessment of when extra information from sources such as labelled substrate would be essential for accurate flux calculation.

Beyond offering an *a priori* determination of a model's predictive strength, the proposed methodology explicitly considers the lack of fit between model and measured data. Model validity is particularly important in the context of overdetermined MFA due to the large degree of simplification involved in model generation. However, lack of fit is rarely considered outside of "gross measurement error" detection. The combination of t-test validation and balanced flux simulation avoids the assumption of model validity in the determination of significance. Identified lack of fit is independent of gross measurement errors and is overlooked by standard χ^2 tests.

It is important to note that the GLS framework for validation is also more robust to estimated measurement error than the standard χ^2 test. GLS regression only requires an estimate of relative measurement variance and covariance in the form of V. Residual variance magnitude ($\hat{\sigma}^2$) is still estimated from the model. On the other hand, variance scaling in the χ^2 test allows for large measurement variance to reduce the χ^2 statistic. Effectively, high variability leads to a lower confidence that deviations are not normally distributed. Given that variance does not factor into any other aspect of validation, assuming a large variance can serve as a way to avoid dealing with lack of fit.

Following the case study presented in this work, we recommend the following valida-

tion procedure. Before any experiments are carried out (but after a model of interest has been identified), construct reasonable limits around each observable flux from literature or other available data. Simulate flux profiles from the constrained flux space and perturb them with a range of measurement errors. If the flux space is infeasible, there is considerable disagreement between fluxes and the model that needs to be resolved. Otherwise, generate confidence intervals around the calculated fluxes and calculate the proportion of simulated fluxes that are non-significant. If many high magnitude fluxes are found to be non-significant in the majority of simulations (regardless of measurement error), then the model may have structural issues that need to be resolved. Alternatively, extra flux information may be required. If the model is sound, then experiments can be carried out and collected data analyzed via MFA. Apply the model and generate confidence intervals around calculated fluxes. Construct limits in close vicinity of observed values, simulate flux profiles, and perturb them with estimated measurement error. If the confidence intervals of simulated fluxes are considerably smaller than those of observed fluxes, then the model may have errors resulting in a lack of fit.

5.6 Conclusion

The interpretation of MFA through the GLS framework opens the door to a suite of readily applied validation methods. Furthermore, the mathematical equivalence of MFA and regression suggests that the failure to follow good practices of regression analysis can lead to questionable results. This work establishes the use of t-tests for the detection of error due to measurement variability as well as presenting a means to directly assess

model error via flux profile simulation. At the same time, we highlight the impact of measurement variability on calculation error and validation, underlining the need for better reporting. Although this work has focused on the validation of a traditional MFA model via t-test analysis, the overall framework is likely to be just as applicable to other regression validation methods or alternative MFA formulations (such as dynamic MFA).

Availability of data and material

All experimental data used in this manuscript and all of the code used to generate the figures is available on github (https://github.com/ssokolen/omfapy). Simulation results are not made available due to their large size, but can be recreated with provided code.

Funding

The work was supported in part by an NSERC Canada Graduate Scholarship to SS as well as NSERC Discovery (RGPIN 355513-2012) and Mabnet Strategic Network (NETGP 380070-08) grants to MGA. The funding provider did not play a role in the design of the study; collection, analysis, and interpretation of data; or writing the manuscript.

Author's contributions

SS designed the validation method, performed the analysis, and assisted with data collection. MQ assisted with data collection and performed NMR quantification. MGA assisted

in data interpretation and manuscript preparation. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Steffen Schulze and Eric J Blondeel for their assistance with bioreactor cultivation.

Conclusions and Recommendations

The increased pace of metabolomic data collection associated with systems biology research requires an increased vigilance against bias and variability. At its simplest, a greater volume of data will naturally contain a greater number of erroneous observations. More importantly, however, increasing synthesis of metabolomic and genomic information (as well as the eventual integration of transcriptome regulation) has the potential to hide errors under layers of modeling. At the same time, a greater volume of data also allows for improved error detection and reconciliation. Although this thesis applies a relatively narrow focus on targeted profiling quantification of ^{1}H-NMR data and traditional metabolic flux analysis, the results of the investigations have broader implications.

Despite its human component, the targeted profiling method was found to be remarkably precise, with a relative standard deviations as low as 0.5–1.5% at the base level – comparable to spectral integration. However, the accuracy of quantification was found to be highly dependent on the water suppression technique implemented in the pulse sequence. Considering the difficulty of generating a spectral library from pure chemical stocks across a range of pH values, the development of new pulse sequences will need to show a considerable improvement over the accepted standards to warrant implementation.

This can be seen as a considerable limitation of library based approaches. Indeed, the issue of library robustness highlights the potential value of accurately simulating the impact of water suppression or even whole library generation through quantum mechanical modeling of NMR principles.

The implementation of a Plackett-Burman design for application-specific variance estimation revealed further useful information about quantification accuracy and precision. Although few general trends could be identified, the efficiency of the experiment provides a convenient framework for further application-specific studies. Estimating variance is essential for low sample number time-course trends, which may otherwise be confused for nothing more than noise around a straight line. Even though high throughput sampling reduces the need for *a priori* variance estimates (as variance can be estimated directly from a fitted trend), bias remains a source of concern. Even high sampling rates will not be able to detect, for example, when high concentrations of one metabolite inflate the concentration estimates of another, especially if both metabolites have correlated trends. Perhaps more interesting, however, is that both accuracy and precision stem from relatively predictable convolution patterns. Thus, further development of quantification methods may be able to provide estimates of error alongside the concentration values.

The in-depth understanding of quantification error developed as part of this thesis also led to better data interpretation. Without an understanding of expected variance, it would have been difficult to detect correlated deviations of multiple metabolites as a result of dilution error. Although the developed correction strategy was designed primarily to resolve issues of inconsistent internal standard addition, its implications are far broader. A similar approach may be used to account for variable extraction efficiency and cell

concentration in intracellular metabolomic analysis. Moreover, the correction strategy represents a move towards more integrated methods of analysis that incorporate as much of the data as possible, illustrated by the fact that observing 80 or more metabolites reduces the threshold of dilution error detection to 1% or lower. In contrast to commonly applied unsupervised methods of analysis, understanding and incorporating the structure of the data and sources of variation can result in more accurate models to come.

The identification of biases in sampling or metabolite concentration estimation has a considerable impact on the use of metabolic models. Whereas the traditional use of simple stoichiometric models was to reconcile possible error in metabolite detection, the growing complexity of model development requires a stronger emphasis on model validation. Despite its age, traditional metabolic flux analysis remains a convenient means of estimating overall flux flows. However, the accuracy of the metabolic model does not always receive enough attention. The ability to generate good variance estimates and correct biases before the modeling stage allows a more stringent examination of the model. The application of simple t-test validation revealed considerable deviations between data and model that were not detected by traditional methods focusing on gross measurement error. Pairing t-test validation with the simulation of theoretically valid fluxes, it became possible to differentiate between measurement uncertainty and lack of model of fit. Separating data and model error and developing improved validation methods will only become more important as modeling complexity increases.

References

Akoka S. and Giraudeau P. (2015).
"Fast hybrid multi-dimensional NMR methods based on ultrafast 2D NMR."
Magnetic Resonance in Chemistry 53.11, pp. 986–94. DOI: 10.1002/mrc.4237.

Al Zaid Siddiquee K., Arauzo-Bravo M. J., and Shimizu K. (2004).
"Metabolic flux analysis of pykF gene knockout Escherichia coli based on 13C-labeling experiments together with measurements of enzyme activities and intracellular metabolite concentrations". *Applied Microbiology and Biotechnology* 63.4, pp. 407–17.
DOI: 10.1007/s00253-003-1357-9.

Altamirano C., Illanes A., Casablancas A., Gámez X., Cairó J. J., and Gòdia C. (2001).
"Analysis of CHO cells metabolic redistribution in a glutamate-based defined medium in continuous culture." *Biotechnology Progress* 17.6, pp. 1032–41.
DOI: 10.1021/bp0100981.

Alves A. C., Li J. V., Garcia-Perez I., Sands C., Barbas C., Holmes E., and Ebbels T. M. D. (2012). "Characterization of data analysis methods for information recovery from metabolic 1H NMR spectra using artificial complex mixtures". *Metabolomics* 8.6, pp. 1170–1180. DOI: 10.1007/s11306-012-0422-8.

Antoniewicz M. R. (2015).
"Methods and advances in metabolic flux analysis: a mini-review."
Journal of industrial microbiology & biotechnology 42.3, pp. 317–25.
DOI: 10.1007/s10295-015-1585-x.

Antoniewicz M. R., Kelleher J. K., and Stephanopoulos G. (2006). "Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements."
Metabolic Engineering 8.4, pp. 324–37. DOI: 10.1016/j.ymben.2006.01.004.

Aranibar N., Borys M., Mackin N. A., Ly V., Abu-Absi N., Abu-Absi S., Niemitz M., Schilling B., Li Z. J., Brock B., Russell R. J., Tymiak A., and Reily M. D. (2011). "NMR-based metabolomics of mammalian cell and tissue cultures." *Journal of biomolecular NMR* 49.3-4, pp. 195–206. DOI: 10.1007/s10858-011-9490-8.

Aranibar N. and Reily M. D. (2014). "NMR Methods for Metabolomics of Mammalian Cell Culture Bioreactors". In: ed. by R. Pörtner. Vol. 1104. Methods in Molecular Biology. Humana Press, pp. 223–236. DOI: 10.1007/978-1-62703-733-4_15.

Athersuch T. J., Malik S., Weljie A., Newton J., and Keun H. C. (2013). "Evaluation of 1H NMR metabolic profiling using biofluid mixture design." *Analytical Chemistry* 85.14, pp. 6674–81. DOI: 10.1021/ac400449f.

Babushok V. I., Linstrom P. J., Reed J. J., Zenkevich I. G., Brown R. L., Mallard W. G., and Stein S. E. (2007). "Development of a database of gas chromatographic retention properties of organic compounds." *Journal of Chromatography A* 1157.1-2, pp. 414–21. DOI: 10.1016/j.chroma.2007.05.044.

Badsha M. B., Kurata H., Onitsuka M., Oga T., and Omasa T. (2016). "Metabolic analysis of antibody producing Chinese hamster ovary cell culture under different stresses conditions." *Journal of Bioscience and Bioengineering*. DOI: 10.1016/j.jbiosc.2015.12.013.

Barding G. A., Salditos R., and Larive C. K. (2012). "Quantitative NMR for bioanalysis and metabolomics." *Analytical and Bioanalytical Chemistry* 404.4, pp. 1165–79. DOI: 10.1007/s00216-012-6188-z.

Bax A. (1985). "A spatially selective composite 90° radiofrequency pulse". *Journal of Magnetic Resonance* 65, pp. 142–145.

Beckonert O., Keun H. C., Ebbels T. M. D., Bundy J., Holmes E., Lindon J. C., and Nicholson J. K. (2007). "Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts." *Nature Protocols* 2.11, pp. 2692–703. DOI: 10.1038/nprot.2007.376.

Begley P., Francis-McIntyre S., Dunn W. B., Broadhurst D. I., Halsall A., Tseng A., Knowles J., Goodacre R., and Kell D. B. (2009). "Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum." *Analytical Chemistry* 81.16, pp. 7038–46. DOI: 10.1021/ac9011599.

Behrends V., Bell T. J., Liebeke M., Cordes-Blauert A., Ashraf S. N., Nair C., Zlosnik J. E. A., Williams H. D., and Bundy J. G. (2013). "Metabolite profiling to characterize disease-related bacteria: gluconate excretion by Pseudomonas aeruginosa mutants and clinical isolates from cystic fibrosis patients." *The Journal of Biological Chemistry* 288.21, pp. 15098–109. DOI: 10.1074/jbc.M112.442814.

Beltran A., Suarez M., Rodríguez M. A., Vinaixa M., Samino S., Arola L., Correig X., and Yanes O. (2012). "Assessment of compatibility between extraction methods for NMR- and LC/MS-based metabolomics." *Analytical Chemistry* 84.14, pp. 5838–44. DOI: 10.1021/ac3005567.

Berk M., Ebbels T., and Montana G. (2011). "A statistical framework for biomarker discovery in metabolomic time course data." *Bioinformatics* 27.14, pp. 1979–85. DOI: 10.1093/bioinformatics/btr289.

Bernal V., Carinhas N., Yokomizo A. Y., Carrondo M. J. T., and Alves P. M. (2009). "Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism". *Biotechnology and Bioengineering* 104.1, pp. 162–80. DOI: 10.1002/bit.22364.

Bharti S. K. and Roy R. (2012). "Quantitative 1H NMR spectroscopy". *TrAC Trends in Analytical Chemistry* 35, pp. 5–26. DOI: 10.1016/j.trac.2012.02.007.

Blank L. M. and Kuepfer L. (2010). "Metabolic flux distributions: Genetic information, computational predictions, and experimental validation". *Applied Microbiology and Biotechnology* 86.5, pp. 1243–55. DOI: 10.1007/s00253-010-2506-6.

Boghigian B. A., Seth G., Kiss R., and Pfeifer B. A. (2010). "Metabolic flux analysis and pharmaceutical production". *Metabolic Engineering* 12.2, pp. 81–95. DOI: 10.1016/j.ymben.2009.10.004.

Bolten C. J., Kiefer P., Letisse F., Portais J.-C., and Wittmann C. (2007). "Sampling for metabolome analysis of microorganisms." *Analytical Chemistry* 79.10, pp. 3843–9. DOI: 10.1021/ac0623888.

Bordbar A., Monk J. M., King Z. A., and Palsson B. O. (2014). "Constraint-based models predict metabolic and associated cellular functions." *Nature Reviews Genetics* 15.2, pp. 107–20. DOI: 10.1038/nrg3643.

Bradley S. A., Ouyang A., Purdie J., Smitka T. A., Wang T., and Kaerner A. (2010). "Fermentanomics: monitoring mammalian cell cultures with NMR spectroscopy". *Journal of the American Chemical Society* 132.28, pp. 9531–3. DOI: 10.1021/ja101962c.

Bragger J. M., Dunn R. V., and Daniel R. M. (2000). "Enzyme activity down to -100°C". *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology* 1480.1–2, pp. 278–282. DOI: http://dx.doi.org/10.1016/S0167-4838(00)00081-9.

Canelas A. B., Ras C., ten Pierick A., van Dam J. C., Heijnen J. J., and van Gulik W. M. (2008). "Leakage-free rapid quenching technique for yeast metabolomics". *Metabolomics* 4.3, pp. 226–239. DOI: 10.1007/s11306-008-0116-4.

Canelas A. B., ten Pierick A., Ras C., Seifar R. M., van Dam J. C., van Gulik W. M., and Heijnen J. J. (2009). "Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics." *Analytical Chemistry* 81.17, pp. 7379–89. DOI: 10.1021/ac900999t.

Carinhas N., Bernal V., Monteiro F., Carrondo M. J. T., Oliveira R., and Alves P. M. (2010). "Improving baculovirus production at high cell density through manipulation of energy metabolism." *Metabolic Engineering* 12.1, pp. 39–52. DOI: 10.1016/j.ymben.2009.08.008.

Carinhas N., Duarte T. M., Barreiro L. C., Carrondo M. J. T., Alves P. M., and Teixeira A. P. (2013). "Metabolic signatures of GS-CHO cell clones associated with butyrate treatment and culture phase transition." *Biotechnology and Bioengineering* 110.12, pp. 3244–57. DOI: 10.1002/bit.24983.

Carinhas N., Oliveira R., Alves P. M., Carrondo M. J. T., and Teixeira A. P. (2012). "Systems biotechnology of animal cells: the road to prediction". *Trends in Biotechnology* 30.7, pp. 377–85. DOI: 10.1016/j.tibtech.2012.03.004.

Carnicer M., Canelas A. B., Ten Pierick A., Zeng Z., van Dam J., Albiol J., Ferrer P., Heijnen J. J., and van Gulik W. (2012). "Development of quantitative metabolomics for Pichia pastoris." *Metabolomics* 8.2, pp. 284–298. DOI: 10.1007/s11306-011-0308-1.

Chace D. H. (2001). "Mass spectrometry in the clinical laboratory." *Chemical Reviews* 101.2, pp. 445–77. DOI: 10.1021/cr990077+.

Chen C., Le H., and Goudar C. T. (2016). "Integration of systems biology in cell line and process development for biopharmaceutical manufacturing". *Biochemical Engineering Journal* 107, pp. 11–17. DOI: 10.1016/j.bej.2015.11.013.

Cheney W. and Kincaid D. (2007). *Numerical mathematics and computing*. Cengage Learning, p. 321.

Chong W. P. K., Yusufi F. N. K., Lee D.-Y., Reddy S. G., Wong N. S. C., Heng C. K., Yap M. G. S., and Ho Y. S. (2011). "Metabolomics-based identification of apoptosis-inducing metabolites in recombinant fed-batch CHO culture media." *Journal of Biotechnology* 151.2, pp. 218–24. DOI: 10.1016/j.jbiotec.2010.12.010.

Chong W. P. K., Thng S. H., Hiu A. P., Lee D.-Y., Chan E. C. Y., and Ho Y. S. (2012). "LC-MS-based metabolic characterization of high monoclonal antibody-producing Chinese hamster ovary cells." *Biotechnology and Bioengineering* 109.12, pp. 3103–11. DOI: 10.1002/bit.24580.

Christie K. R., Hong E. L., and Cherry J. M. (2009). "Functional annotations for the Saccharomyces cerevisiae genome: the knowns and the known unknowns". *Trends in Microbiology* 17.7, pp. 286–94. DOI: 10.1016/j.tim.2009.04.005.

Crockford D. J., Keun H. C., Smith L. M., Holmes E., and Nicholson J. K. (2005). "Curve-fitting method for direct quantitation of compounds in complex biological mixtures using 1H NMR: application in metabonomic toxicology studies." *Analytical Chemistry* 77.14, pp. 4556–62. DOI: 10.1021/ac0503456.

Cuperlović-Culf M., Barnett D. A., Culf A. S., and Chute I. (2010). "Cell culture metabolomics: applications and future directions." *Drug Discovery Today* 15.15-16, pp. 610–21. DOI: 10.1016/j.drudis.2010.06.012.

Danielsson A. P. H., Moritz T., Mulder H., and Spégel P. (2010). "Development and optimization of a metabolomic method for analysis of adherent cell cultures." *Analytical Biochemistry* 404.1, pp. 30–9. DOI: 10.1016/j.ab.2010.04.013.

Dauner M. (2010). "From fluxes and isotope labeling patterns towards in silico cells". *Current Opinion in Biotechnology* 21.1, pp. 55–62. DOI: 10.1016/j.copbio.2010.01.014.

Davis J. M., Collette T. W., Villeneuve D. L., Cavallin J. E., Teng Q., Jensen K. M., Kahl M. D., Mayasich J. M., Ankley G. T., and Ekman D. R. (2013). "Field-based approach for assessing the impact of treated pulp and paper mill effluent on endogenous metabolites of fathead minnows (Pimephales promelas)." *Environmental Science & Technology* 47.18, pp. 10628–36. DOI: 10.1021/es401961j.

De Meyer T., Sinnaeve D., Van Gasse B., Rietzschel E.-R., De Buyzere M. L., Langlois M. R., Bekaert S., Martins J. C., and Van Criekinge W. (2010). "Evaluation of standard and advanced preprocessing methods for the univariate analysis of blood serum 1H-NMR spectra." *Analytical and Bioanalytical Chemistry* 398.4, pp. 1781–90. DOI: 10.1007/s00216-010-4085-x.

De Meyer T., Sinnaeve D., Van Gasse B., Tsiporkova E., Rietzschel E. R., De Buyzere M. L., Gillebert T. C., Bekaert S., Martins J. C., and Van Criekinge W. (2008). "NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm." *Analytical Chemistry* 80.10, pp. 3783–90. DOI: 10.1021/ac7025964.

De Jonge L. P., Douma R. D., Heijnen J. J., van Gulik W. M., Faijes M., Mars A. E., and Smid E. J. (2012). "Optimization of cold methanol quenching for quantitative metabolomics of Penicillium chrysogenum." *Metabolomics* 8.4, pp. 727–735. DOI: 10.1007/s11306-011-0367-3.

De Koning W. and van Dam K. (1992). "A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH." *Analytical Biochemistry* 204.1, pp. 118–23.

Dieterle F., Riefke B., Schlotterbeck G., Ross A., Senn H., and Amberg A. (2011). "NMR and MS methods for metabonomics." In: vol. 691. Methods in Molecular Biology. Humana Press, pp. 385–415. DOI: 10.1007/978-1-60761-849-2_24.

Dietmair S., Timmins N. E., Gray P. P., Nielsen L. K., and Krömer J. O. (2010). "Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol." *Analytical Biochemistry* 404.2, pp. 155–64. DOI: 10.1016/j.ab.2010.04.031.

Drugmand J.-C. (2007). "Study of the metabolism and physiology of High-Five insect cells for the development of processes for the production of recombinant protein". PhD thesis.

Dumas M.-E., Maibaum E. C., Teague C., Ueshima H., Zhou B., Lindon J. C., Nicholson J. K., Stamler J., Elliott P., Chan Q., and Holmes E. (2006). "Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP Study." *Analytical Chemistry* 78.7, pp. 2199–208. DOI: 10.1021/ac0517085.

Dunn W. B. and Ellis D. I. (2005). "Metabolomics: Current analytical platforms and methodologies". *TrAC Trends in Analytical Chemistry* 24.4, pp. 285–294. DOI: http://dx.doi.org/10.1016/j.trac.2004.11.021.

Eads C. D., Furnish C. M., Noda I., Juhlin K. D., Cooper D. A., and Morrall S. W. (2004). "Molecular factor analysis applied to collections of NMR spectra." *Analytical Chemistry* 76.7, pp. 1982–90. DOI: 10.1021/ac035301g.

Erdrich P., Steuer R., and Klamt S. (2015). "An algorithm for the reduction of genome-scale metabolic network models to meaningful core models." *BMC Systems Biology* 9, p. 48. DOI: 10.1186/s12918-015-0191-x.

Faijes M., Mars A. E., and Smid E. J. (2007). "Comparison of quenching and extraction methodologies for metabolome analysis of Lactobacillus plantarum." *Microbial Cell Factories* 6, p. 27. DOI: 10.1186/1475-2859-6-27.

Fathallah-Shaykh H. M. (2005). "Microarrays: applications and pitfalls." *Archives of Neurology* 62.11, pp. 1669–72. DOI: 10.1001/archneur.62.11.1669.

FDA (2004). *PAT – A framework for innovative pharmaceutical development, manufacturing, and quality assurance*. Tech. rep. Rockville, MD: U.S. Department of Health et al.

Feichtinger J., Hernández I., Fischer C., Hanscho M., Auer N., Hackl M., Jadhav V., Baumann M., Krempl P. M., Schmidl C., Farlik M., Schuster M., Merkel A., Sommer A., Heath S., Rico D., Bock C., Thallinger G. G., and Borth N. (2016). "Comprehensive genome and epigenome characterization of CHO cells in response to evolutionary pressures and over time." *Biotechnology and Bioengineering*. DOI: 10.1002/bit.25990.

Foxall P. J., Spraul M., Farrant R. D., Lindon L. C., Neild G. H., and Nicholson J. K. (1993). "750 MHz 1H-NMR spectroscopy of human blood plasma." *Journal of Pharmaceutical and Biomedical Analysis* 11.4-5, pp. 267–76.

Fuerst R. and Wagner R. P. (1957). "An analysis of the free intracellular amino acids of certain strains of Neurospora." *Archives of Biochemistry and Biophysics* 70.2, pp. 311–26.

Fukuhara K., Ohno A., Ota Y., Senoo Y., Maekawa K., Okuda H., Kurihara M., Okuno A., Niida S., Saito Y., and Takikawa O. (2013). "NMR-based metabolomics of urine in a mouse model of Alzheimer's disease: identification of oxidative stress biomarkers." *Journal of Clinical Biochemistry and Nutrition* 52.2, pp. 133–8. DOI: 10.3164/jcbn.12-118.

Gal M. and Frydman L. (2015). "Multidimensional NMR spectroscopy in a single scan." *Magnetic Resonance in Chemistry* 53.11, pp. 971–85. DOI: 10.1002/mrc.4271.

Gale E. F. (1947). "The assimilation of amino-acids by bacteria; the passage of certain amino-acids across the cell wall and their concentration in the internal environment of Streptococcus faecalis." *Journal of General Microbiology* 1.1, pp. 53–76. DOI: 10.1099/00221287-1-1-53.

Gao P. and Xu G. (2015). "Mass-spectrometry-based microbial metabolomics: recent developments and applications." *Analytical and Bioanalytical Chemistry* 407.3, pp. 669–80. DOI: 10.1007/s00216-014-8127-7.

García-Sevillano M. Á., García-Barrera T., and Gómez-Ariza J. L. (2015). "Environmental metabolomics: Biological markers for metal toxicity." *Electrophoresis*. DOI: 10.1002/elps.201500052.

Garde A. H., Hansen A. M., Kristiansen J., and Knudsen L. E. (2004).
"Comparison of uncertainties related to standardization of urine samples with volume and creatinine concentration." *The Annals of Occupational Hygiene* 48.2, pp. 171–9.
Gawlitzek M., Valley U., and Wagner R. (1998).
"Ammonium ion and glucosamine dependent increases of oligosaccharide complexity in recombinant glycoproteins secreted from cultivated BHK-21 cells."
Biotechnology and Bioengineering 57.5, pp. 518–28.
Gipson G. T., Tatsuoka K. S., Sweatman B. C., and Connor S. C. (2006).
"Weighted least-squares deconvolution method for discovery of group differences between complex biofluid 1H NMR spectra."
Journal of Magnetic Resonance 183.2, pp. 269–77.
DOI: 10.1016/j.jmr.2006.09.003.
Giraudeau P., Tea I., Remaud G. S., and Akoka S. (2014). "Reference and normalization methods: essential tools for the intercomparison of NMR spectra."
Journal of pharmaceutical and biomedical analysis 93, pp. 3–16.
DOI: 10.1016/j.jpba.2013.07.020.
Gómez J., Brezmes J., Mallol R., Rodríguez M. A., Vinaixa M., Salek R. M., Correig X., and Cañellas N. (2014). "Dolphin: a tool for automatic targeted metabolite profiling using 1D and 2D (1)H-NMR data."
Analytical and Bioanalytical Chemistry 406.30, pp. 7967–76.
DOI: 10.1007/s00216-014-8225-6.
Goudar C. T., Biener R., Konstantinov K. B., and Piret J. M. (2009).
"Error propagation from prime variables into specific rates and metabolic fluxes for mammalian cells in perfusion culture." *Biotechnology Progress* 25.4, pp. 986–98.
DOI: 10.1002/btpr.155.
Goudar C., Biener R., Zhang C., Michaels J., Piret J., and Konstantinov K. (2006).
"Towards industrial application of quasi real-time metabolic flux analysis for mammalian cell culture."
Advances in Biochemical Engineering/Biotechnology 101, pp. 99–118.
DOI: 10.1007/10_020.
Griffiths L. and Irving A. (1998).
"Assay by nuclear magnetic resonance spectroscopy: quantification limits".
Analyst 123.5, pp. 1061–1068. DOI: 10.1039/A800625C.
Guennec A. L., Giraudeau P., and Caldarelli S. (2014).
"Evaluation of fast 2D NMR for metabolomics."
Analytical Chemistry 86.12, pp. 5946–54. DOI: 10.1021/ac500966e.

Gutenkunst R. N., Waterfall J. J., Casey F. P., Brown K. S., Myers C. R., and Sethna J. P. (2007).
"Universally sloppy parameter sensitivities in systems biology models."
PLoS Computational Biology 3.10, pp. 1871–78.
DOI: 10.1371/journal.pcbi.0030189.

Hancock R. (1958).
"The intracellular amino acids of Staphylococcus aureus: release and analysis."
Biochimica et Biophysica Acta 28.2, pp. 402–12.

Hao J., Astle W., De Iorio M., and Ebbels T. M. D. (2012).
"BATMAN–an R package for the automated quantification of metabolites from nuclear magnetic resonance spectra using a Bayesian model".
Bioinformatics 28.15, pp. 2088–2090. DOI: 10.1093/bioinformatics/bts308.

Hasunuma T., Sanda T., Yamada R., Yoshimura K., Ishii J., and Kondo A. (2011).
"Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of Saccharomyces cerevisiae."
Microbial Cell Factories 10.1, p. 2. DOI: 10.1186/1475-2859-10-2.

Ideker T., Galitski T., and Hood L. (2001).
"A new approach to decoding life: systems biology."
Annual Review of Genomics and Human Genetics 2, pp. 343–72.
DOI: 10.1146/annurev.genom.2.1.343.

Issaq H. J., Van Q. N., Waybright T. J., Muschik G. M., and Veenstra T. D. (2009).
"Analytical and statistical approaches to metabolomics research."
Journal of Separation Science 32.13, pp. 2183–99. DOI: 10.1002/jssc.200900152.

Iwatani S., Yamada Y., and Usuda Y. (2008).
"Metabolic flux analysis in biotechnology processes".
Biotechnology Letters 30.5, pp. 791–9. DOI: 10.1007/s10529-008-9633-5.

Jung J.-Y., Jung Y., Kim J.-S., Ryu D. H., and Hwang G.-S. (2013).
"Assessment of peeling of Astragalus roots using 1H NMR- and UPLC-MS-based metabolite profiling."
Journal of Agricultural and Food Chemistry 61.43, pp. 10398–407.
DOI: 10.1021/jf4026103.

Kaas C. S., Kristensen C., Betenbaugh M. J., and Andersen M. R. (2015).
"Sequencing the CHO DXB11 genome reveals regional variations in genomic stability and haploidy." *BMC Genomics* 16, p. 160. DOI: 10.1186/s12864-015-1391-x.

Kakumani P. K., Malhotra P., Mukherjee S. K., and Bhatnagar R. K. (2014).
"A draft genome assembly of the army worm, Spodoptera frugiperda."
Genomics 104.2, pp. 134–43. DOI: 10.1016/j.ygeno.2014.06.005.

Kauffman K. J., Prakash P., and Edwards J. S. (2003).
"Advances in flux balance analysis".
Current Opinion in Biotechnology 14.5, pp. 491–496.
DOI: 10.1016/j.copbio.2003.08.001.

Keeler J., Clowes R. T., Davis A. L., and Laue E. D. (1994).
"Pulsed-field gradients: theory and practice."
Methods in Enzymology 239, pp. 145–207.

Keun H. C., Beckonert O., Griffin J. L., Richter C., Moskau D., Lindon J. C., and Nicholson J. K. (2002a).
"Cryogenic probe 13C NMR spectroscopy of urine for metabonomic studies."
Analytical Chemistry 74.17, pp. 4588–93. DOI: 10.1021/ac025691r.

Keun H. C., Ebbels T. M. D., Antti H., Bollard M. E., Beckonert O., Schlotterbeck G., Senn H., Niederhauser U., Holmes E., Lindon J. C., and Nicholson J. K. (2002b).
"Analytical reproducibility in (1)H NMR-based metabonomic urinalysis."
Chemical Research in Toxicology 15.11, pp. 1380–6. DOI: 10.1021/tx0255774.

Khoo S. H. G. and Al-Rubeai M. (2007).
"Metabolomics as a complementary tool in cell culture."
Biotechnology and Applied Biochemistry 47.Pt 2, pp. 71–84.
DOI: 10.1042/BA20060221.

Khoo S. H. G. and Al-Rubeai M. (2009). "Metabolic characterization of a hyper-productive state in an antibody producing NS0 myeloma cell line."
Metabolic Engineering 11.3, pp. 199–211. DOI: 10.1016/j.ymben.2009.02.001.

Kim H. U., Kim T. Y., and Lee S. Y. (2008).
"Metabolic flux analysis and metabolic engineering of microorganisms".
Molecular BioSystems 4.2, pp. 113–20. DOI: 10.1039/b712395g.

Kind T., Wohlgemuth G., Lee D. Y., Lu Y., Palazoglu M., Shahbaz S., and Fiehn O. (2009). "FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry."
Analytical Chemistry 81.24, pp. 10038–48. DOI: 10.1021/ac9019522.

King R., Bonfiglio R., Fernandez-Metzler C., Miller-Stein C., and Olah T. (2000).
"Mechanistic investigation of ionization suppression in electrospray ionization."
Journal of the American Society for Mass Spectrometry 11.11, pp. 942–50.
DOI: 10.1016/S1044-0305(00)00163-X.

Kitteringham N. R., Jenkins R. E., Lane C. S., Elliott V. L., and Park B. K. (2009).
"Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics." *Journal of Chromatography B* 877.13, pp. 1229–39.
DOI: 10.1016/j.jchromb.2008.11.013.

Klamt S., Saez-Rodriguez J., and Gilles E. D. (2007).
"Structural and functional analysis of cellular networks with CellNetAnalyzer".
BMC Systems Biology 1.1, pp. 1–13. DOI: 10.1186/1752-0509-1-2.
Klamt S. and von Kamp A. (2011).
"An application programming interface for CellNetAnalyzer."
BioSystems 105.2, pp. 162–8. DOI: 10.1016/j.biosystems.2011.02.002.
Kohl S. M., Klein M. S., Hochrein J., Oefner P. J., Spang R., and Gronwald W. (2012).
"State-of-the art data normalization methods improve NMR-based metabolomic analysis." *Metabolomics* 8.Suppl 1, pp. 146–160. DOI: 10.1007/s11306-011-0350-z.
Koo I., Shi X., Kim S., and Zhang X. (2014). "iMatch2: compound identification using retention index for analysis of gas chromatography-mass spectrometry data."
Journal of Chromatography A 1337, pp. 202–10.
DOI: 10.1016/j.chroma.2014.02.049.
Krömer J. O., Sorgenfrei O., Klopprogge K., Heinzle E., and Wittmann C. (2004).
"In-depth profiling of lysine-producing Corynebacterium glutamicum by combined analysis of the transcriptome, metabolome, and fluxome".
Journal of Bacteriology 186.6, pp. 1769–1784. DOI: 10.1128/JB.186.6.1769.
Kronthaler J., Gstraunthaler G., and Heel C. (2012).
"Optimizing high-throughput metabolomic biomarker screening: a study of quenching solutions to freeze intracellular metabolism in CHO cells." *OMICS* 16.3, pp. 90–7.
DOI: 10.1089/omi.2011.0048.
Kumar A. and Bachhawat A. K. (2012).
"Pyroglutamic acid: throwing light on a lightly studied metabolite".
Current Science 102.2, pp. 288–97.
Laatikainen R., Niemitz M., Malaisse W. J., Biesemans M., and Willem R. (1996).
"A computational strategy for the deconvolution of NMR spectra with multiplet structures and constraints: analysis of overlapping 13C-2H multiplets of 13C enriched metabolites from cell suspensions incubated in deuterated media."
Magnetic Resonance in Medicine 36.3, pp. 359–65.
Lameiras F., Heijnen J. J., and van Gulik W. M. (2015). "Development of tools for quantitative intracellular metabolomics of Aspergillus niger chemostat cultures."
Metabolomics 11.5, pp. 1253–1264. DOI: 10.1007/s11306-015-0781-z.
Lange H. C., Eman M., van Zuijlen G., Visser D., van Dam J. C., Frank J., de Mattos M. J., and Heijnen J. J. (2001). "Improved rapid sampling for in vivo kinetics of intracellular metabolites in Saccharomyces cerevisiae."
Biotechnology and Bioengineering 75.4, pp. 406–15.

Larive C. K., Barding G. A., and Dinges M. M. (2015).
"NMR spectroscopy for metabolomics and metabolic profiling."
Analytical Chemistry 87.1, pp. 133–46. DOI: 10.1021/ac504075g.
Lauridsen M., Hansen S. H., Jaroszewski J. W., and Cornett C. (2007).
"Human urine as test material in 1H NMR-based metabonomics: recommendations for sample preparation and storage." *Analytical Chemistry* 79.3, pp. 1181–6.
DOI: 10.1021/ac061354x.
Lederberg J. and McCray A. T. (2001).
"'Ome Sweet 'Omics – a genealogical treasury of words". *The Scientist* 15.8.
Lee S. Y., Lee D.-Y., and Kim T. Y. (2005).
"Systems biotechnology for strain improvement".
Trends in Biotechnology 23.7, pp. 349–58. DOI: 10.1016/j.tibtech.2005.05.003.
Leenders J., Frédérich M., and de Tullio P. (2015). "Nuclear magnetic resonance: a key metabolomics platform in the drug discovery process."
Drug Discovery Today: Technologies 13, pp. 39–46.
DOI: 10.1016/j.ddtec.2015.06.005.
Lei Z., Huhman D. V., Sumner L. W., Lei Z., Huhman D. V., and Sumner L. W. (2011).
"Mass spectrometry strategies in metabolomics."
The Journal of Biological Chemistry 286.29, pp. 25435–42.
DOI: 10.1074/jbc.R111.238691.
Leighty R. W. and Antoniewicz M. R. (2011). "Dynamic metabolic flux analysis (DMFA): a framework for determining fluxes at metabolic non-steady state."
Metabolic Engineering 13.6, pp. 745–55. DOI: 10.1016/j.ymben.2011.09.010.
Levitt M. H. (1986). *Composite pulses*. Wiley Online Library.
Lewis I. A., Schommer S. C., Hodis B., Robb K. A., Tonelli M., Westler W. M., Sussman M. R., and Markley J. L. (2007).
"Method for determining molar concentrations of metabolites in complex solutions from two-dimensional 1H-13C NMR spectra."
Analytical Chemistry 79.24, pp. 9385–90. DOI: 10.1021/ac071583z.
Lewis N. E., Nagarajan H., and Palsson B. O. (2012). "Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods."
Nature Reviews Microbiology 10.4, pp. 291–305. DOI: 10.1038/nrmicro2737.
Lindon J. C., Keun H. C., Ebbels T. M., Pearce J. M., Holmes E., and Nicholson J. K. (2005). "The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements." *Pharmacogenomics* 6.7, pp. 691–9.
DOI: 10.2217/14622416.6.7.691.

Link H., Anselment B., and Weuster-Botz D. (2008).
"Leakage of adenylates during cold methanol/glycerol quenching of Escherichia coli".
Metabolomics 4.3, pp. 240–247. DOI: 10.1007/s11306-008-0114-6.
Liu M. and Mao X.-a. (1999). "Solvent Suppression Methods in NMR Spectroscopy". In:
Oxford: Elsevier, pp. 2145–2152. DOI: 10.1006/rwsp.2000.0385.
Liu S., Wang W., Zhou X., Gu R., and Ding Z. (2014).
"Dose responsive effects of cisplatin in L02 cells using NMR-based metabolomics."
Environmental Toxicology and Pharmacology 37.1, pp. 150–7.
DOI: 10.1016/j.etap.2013.11.016.
Luo J., Vijayasankaran N., Autsen J., Santuray R., Hudson T., Amanullah A., and Li F.
(2012). "Comparative metabolite analysis to understand lactate metabolism shift in
Chinese hamster ovary cell culture process."
Biotechnology and Bioengineering 109.1, pp. 146–56. DOI: 10.1002/bit.23291.
Ma N., Ellet J., Okediadi C., Hermes P., McCormick E., and Casnocha S. (2009).
"A single nutrient feed supports both chemically defined NS0 and CHO fed-batch
processes: Improved productivity and lactate metabolism."
Biotechnology Progress 25.5, pp. 1353–63. DOI: 10.1002/btpr.238.
Madron F., Veverka V., and Vaněček V. (1977).
"Statistical analysis of material balance of a chemical reactor".
AIChE Journal 23.4, pp. 482–486. DOI: 10.1002/aic.690230412.
Maertens J. and Vanrolleghem P. A. (2010). "Modeling with a view to target
identification in metabolic engineering: A critical evaluation of the available tools".
Biotechnology Progress 26.2, pp. 313–31. DOI: 10.1002/btpr.349.
Maharjan R. P. and Ferenci T. (2003). "Global metabolite analysis: the influence of
extraction methodology on metabolome profiles of Escherichia coli."
Analytical Biochemistry 313.1, pp. 145–54.
Malz F. and Jancke H. (2005). "Validation of quantitative NMR".
Journal of pharmaceutical and biomedical analysis 38.5, pp. 813–23.
DOI: 10.1016/j.jpba.2005.01.043.
Maniara G., Rajamoorthi K., Rajan S., and Stockton G. W. (1998).
"Method performance and validation for quantitative analysis by (1)H and (31)P
NMR spectroscopy. Applications to analytical standards and agricultural chemicals."
Analytical Chemistry 70.23, pp. 4921–8. DOI: 10.1021/ac980573i.
Mao X.-a. and Chen J.-h. (1996).
"Radiation damping effects in solvent preirradiation experiments in NMR".
Chemicaly Physics 202, pp. 357–366.

Markley J. L., Bax A., Arata Y., Hilbers C. W., Kaptein R., Sykes B. D., Wright P. E., and Wüthrich K. (1998). "Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy." *Journal of Biomolecular NMR* 12.1, pp. 1–23.

Martínez V. S., Buchsteiner M., Gray P., Nielsen L. K., and Quek L.-E. (2015). "Dynamic metabolic flux analysis using B-splines to study the effects of temperature shift on CHO cell metabolism". *Metabolic Engineering Communications* 2, pp. 46–57. DOI: 10.1016/j.meteno.2015.06.001.

Mas S., Villas-Bôas S. G., Hansen M. E., Akesson M., and Nielsen J. (2007). "A comparison of direct infusion MS and GC-MS for metabolic footprinting of yeast mutants." *Biotechnology and Bioengineering* 96.5, pp. 1014–22. DOI: 10.1002/bit.21194.

Mashego M. R., Rumbold K., De Mey M., Vandamme E., Soetaert W., and Heijnen J. J. (2007). "Microbial metabolomics: past, present and future methodologies." *Biotechnology Letters* 29.1, pp. 1–16. DOI: 10.1007/s10529-006-9218-0.

Mason R. L., Gunst R. F., and Hess J. L. (2003). *Fractional Factorial Experiments*. John Wiley & Sons, Inc., pp. 228–270. DOI: 10.1002/0471458503.ch7.

McKay R. T., Mercier P., and Sykes B. D. (2009). *Comparative analysis of solvent suppression techniques and improvements for high resolution 1H NMR metabonomics studies*.

McKay R. T. (2009). "Recent Advances in Solvent Suppression for Solution NMR: A Practical Reference". In: vol. 66. New York: Academic Press, pp. 33–76. DOI: 10.1016/S0066-4103(08)00402-X.

Mckay R. T. (2011). "How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution". *Concepts in Magnetic Resonance A* 38A.5, pp. 197–220. DOI: 10.1002/cmr.a.20223.

McKenzie J. S., Donarski J. A., Wilson J. C., and Charlton A. J. (2011). "Analysis of complex mixtures using high-resolution nuclear magnetic resonance spectroscopy and chemometrics." *Progress in Nuclear Magnetic Resonance Spectroscopy* 59.4, pp. 336–59. DOI: 10.1016/j.pnmrs.2011.04.003.

Mercier P., Lewis M. J., Chang D., Baker D., and Wishart D. S. (2011). "Towards automatic metabolomic profiling of high-resolution one-dimensional proton NMR spectra." *Journal of Biomolecular NMR* 49.3-4, pp. 307–23. DOI: 10.1007/s10858-011-9480-x.

Mohmad-Saberi S. E., Hashim Y. Z. H.-Y., Mel M., Amid A., Ahmad-Raus R., and Packeer-Mohamed V. (2013). "Metabolomics profiling of extracellular metabolites in CHO-K1 cells cultured in different types of growth media." *Cytotechnology* 65.4, pp. 577–86. DOI: 10.1007/s10616-012-9508-4.

Monton M. R. N. and Soga T. (2007). "Metabolome analysis by capillary electrophoresis-mass spectrometry." *Journal of Chromatography A* 1168.1-2, 237–46, discussion 236. DOI: 10.1016/j.chroma.2007.02.065.

Morris G. A., Barjat H., and Home T. J. (1997). "Reference deconvolution methods". *Progress in Nuclear Magnetic Resonance Spectroscopy* 31.2–3, pp. 197–257. DOI: http://dx.doi.org/10.1016/S0079-6565(97)00011-3.

Nagrath D., Caneba C., Karedath T., and Bellance N. (2011). "Metabolomics for mitochondrial and cancer studies." *Biochimica et Biophysica Acta* 1807.6, pp. 650–63. DOI: 10.1016/j.bbabio.2011.03.006.

Nicholson J. K., Lindon J. C., and Holmes E. (1999). "'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data." *Xenobiotica* 29.11, pp. 1181–9. DOI: 10.1080/004982599238047.

Nicholson J. K. and Wilson I. D. (1989). "High resolution proton magnetic resonance spectroscopy of biological fluids". *Progress in Nuclear Magnetic Resonance Spectroscopy* 21, pp. 449–501.

Niklas J., Priesnitz C., Rose T., Sandig V., and Heinzle E. (2012). "Primary metabolism in the new human cell line AGE1.HN at various substrate levels: increased metabolic efficiency and α1-antitrypsin production at reduced pyruvate load". *Applied Microbiology and Biotechnology* 93.4, pp. 1637–50. DOI: 10.1007/s00253-011-3526-6.

Niklas J., Schneider K., and Heinzle E. (2010). "Metabolic flux analysis in eukaryotes". *Current Opinion in Biotechnology* 21.1, pp. 63–9. DOI: 10.1016/j.copbio.2010.01.011.

Niklas J., Schräder E., Sandig V., Noll T., and Heinzle E. (2011). "Quantitative characterization of metabolism and metabolic shifts during growth of the new human cell line AGE1.HN using time resolved metabolic flux analysis." *Bioprocess and Biosystems Engineering* 34.5, pp. 533–45. DOI: 10.1007/s00449-010-0502-y.

Nöh K. and Wiechert W. (2011). "The benefits of being transient: isotope-based metabolic flux analysis at the short time scale". *Applied Microbiology and Biotechnology* 91.5, pp. 1247–65. DOI: 10.1007/s00253-011-3390-4.

Nyberg G. B., Balcarcel R. R., Follstad B. D., Stephanopoulos G., and Wang D. I. (1999). "Metabolism of peptide amino acids by Chinese hamster ovary cells grown in a complex medium." *Biotechnology and Bioengineering* 62.3, pp. 324–35.

Ochs M. F., Stoyanova R. S., Arias-Mendoza F., and Brown T. R. (1999). "A new method for spectral decomposition using a bilinear Bayesian approach." *Journal of Magnetic Resonance* 137.1, pp. 161–76. DOI: 10.1006/jmre.1998.1639.

Oldiges M., Lütz S., Pflug S., Schroer K., Stein N., and Wiendahl C. (2007). "Metabolomics: current state and evolving methodologies and tools." *Applied Microbiology and Biotechnology* 76.3, pp. 495–511. DOI: 10.1007/s00253-007-1029-2.

Oliver S. G., Winson M. K., Kell D. B., and Baganz F. (1998). "Systematic functional analysis of the yeast genome." *Trends in Biotechnology* 16.9, pp. 373–8.

Paczia N., Nilgen A., Lehmann T., Gätgens J., Wiechert W., and Noack S. (2012). "Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms." *Microbial Cell Factories* 11, p. 122. DOI: 10.1186/1475-2859-11-122.

Pan Z. and Raftery D. (2007). "Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics." *Analytical and Bioanalytical Chemistry* 387.2, pp. 525–7. DOI: 10.1007/s00216-006-0687-8.

Parsons H. M., Ekman D. R., Collette T. W., and Viant M. R. (2009). "Spectral relative standard deviation: a practical benchmark in metabolomics." *The Analyst* 134.3, pp. 478–85. DOI: 10.1039/b808986h.

Pinciroli V., Biancardi R., Visentin G., and Rizzo V. (2004). "The Well-Characterized Synthetic Molecule: A Role for Quantitative 1H NMR". *Organic Process Research & Development* 8.3, pp. 381–384. DOI: 10.1021/op0341925.

Plackett R. L. and Burman J. P. (1946). "The design of optimum multifactorial experiments". *Biometrika* 33.4, pp. 305–325. DOI: 10.1093/biomet/33.4.305.

Pope G. A., MacKenzie D. A., Defernez M., Aroso M. A. M. M., Fuller L. J., Mellon F. A., Dunn W. B., Brown M., Goodacre R., Kell D. B., Marvin M. E., Louis E. J., and Roberts I. N. (2007). "Metabolic footprinting as a tool for discriminating between brewing yeasts." *Yeast* 24.8, pp. 667–79. DOI: 10.1002/yea.1499.

Powers R. (2014). "The current state of drug discovery and a potential role for NMR metabolomics." *Journal of Medicinal Chemistry* 57.14, pp. 5860–70. DOI: 10.1021/jm401803b.

Price W. S. (1999). "Water signal suppression in NMR spectroscopy". In: New York: Academic Press, pp. 289–354.

Priesnitz C., Niklas J., Rose T., Sandig V., and Heinzle E. (2012). "Metabolic flux rearrangement in the amino acid metabolism reduces ammonia stress in the α1-antitrypsin producing human AGE1.HN cell line." *Metabolic Engineering* 14.2, pp. 128–37. DOI: 10.1016/j.ymben.2012.01.001.

Psychogios N., Hau D. D., Peng J., Guo A. C., Mandal R., Bouatra S., Sinelnikov I., Krishnamurthy R., Eisner R., Gautam B., Young N., Xia J., Knox C., Dong E., Huang P., Hollander Z., Pedersen T. L., Smith S. R., Bamforth F., Greiner R., McManus B., Newman J. W., Goodfriend T., and Wishart D. S. (2011). "The human serum metabolome." *PloS One* 6.2, e16957. DOI: 10.1371/journal.pone.0016957.

Quek L.-E., Dietmair S., Hanscho M., Martínez V. S., Borth N., and Nielsen L. K. (2014). "Reducing Recon 2 for steady-state flux analysis of HEK cell culture." *Journal of Biotechnology* 184, pp. 172–8. DOI: 10.1016/j.jbiotec.2014.05.021.

Quek L.-E., Dietmair S., Krömer J. O., and Nielsen L. K. (2010). "Metabolic flux analysis in mammalian cell culture." *Metabolic Engineering* 12.2, pp. 161–71. DOI: 10.1016/j.ymben.2009.09.002.

R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Rathore A. S., Kumar Singh S., Pathak M., Read E. K., Brorson K. A., Agarabi C. D., and Khan M. (2015). "Fermentanomics: Relating quality attributes of a monoclonal antibody to cell culture process variables and raw materials using multivariate data analysis." *Biotechnology Progress* 31.6, pp. 1586–99. DOI: 10.1002/btpr.2155.

Ravanbakhsh S., Liu P., Bjorndahl T. C., Bjordahl T. C., Mandal R., Grant J. R., Wilson M., Eisner R., Sinelnikov I., Hu X., Luchinat C., Greiner R., and Wishart D. S. (2015).
"Accurate, fully-automated NMR spectral profiling for metabolomics."
PloS One 10.5, e0124219. DOI: 10.1371/journal.pone.0124219.

Read E. K., Park J. T., Shah R. B., Riley B. S., Brorson K. A., and Rathore A. S. (2010).
"Process analytical technology (PAT) for biopharmaceutical products: Part I. Concepts and applications." *Biotechnology and Bioengineering* 105.2, pp. 276–84.
DOI: 10.1002/bit.22528.

Read E. K., Bradley S. A., Smitka T. A., Agarabi C. D., Lute S. C., and Brorson K. A. (2013). "Fermentanomics informed amino acid supplementation of an antibody producing mammalian cell culture." *Biotechnology Progress* 29.3, pp. 745–53.
DOI: 10.1002/btpr.1728.

Saude E. J., Adamko D., Rowe B. H., Marrie T., and Sykes B. D. (2007).
"Variation of metabolites in normal human urine". *Metabolomics* 3.4, pp. 439–451.
DOI: 10.1007/s11306-007-0091-1.

Saude E. J. and Sykes B. D. (2007).
"Urine stability for metabolomic studies: effects of preparation and storage".
Metabolomics 3.1, pp. 19–27. DOI: 10.1007/s11306-006-0042-2.

Savinell J. M. and Palsson B. O. (1992). "Optimal selection of metabolic fluxes for in vivo measurement. I. Development of mathematical methods".
Journal of Theoretical Biology 155.2, pp. 201–14.

Schaub J., Clemens C., Kaufmann H., and Schulz T. W. (2012).
"Advancing biopharmaceutical process development by system-level data analysis and integration of omics data."
Advances in Biochemical Engineering/Biotechnology 127, pp. 133–63.
DOI: 10.1007/10_2010_98.

Schmidt F. R. (2004). "Recombinant expression systems in the pharmaceutical industry."
Applied Microbiology and Biotechnology 65.4, pp. 363–72.
DOI: 10.1007/s00253-004-1656-9.

Schnackenberg L. K. and Beger R. D. (2008).
"The role of metabolic biomarkers in drug toxicity studies."
Toxicology Mechanisms and Methods 18.4, pp. 301–11.
DOI: 10.1080/15376510701623193.

Schuetz R., Kuepfer L., and Sauer U. (2007). "Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli".
Molecular Systems Biology 3.119, p. 119. DOI: 10.1038/msb4100162.

Sellick C. A., Croxford A. S., Maqsood A. R., Stephens G., Westerhoff H. V., Goodacre R., and Dickson A. J. (2011).
"Metabolite profiling of recombinant CHO cells: designing tailored feeding regimes that enhance recombinant antibody production."
Biotechnology and Bioengineering 108.12, pp. 3025–31. DOI: 10.1002/bit.23269.

Sellick C. A., Hansen R., Maqsood A. R., Dunn W. B., Stephens G. M., Goodacre R., and Dickson A. J. (2009). "Effective quenching processes for physiologically valid metabolite profiling of suspension cultured Mammalian cells."
Analytical Chemistry 81.1, pp. 174–83. DOI: 10.1021/ac8016899.

Sellick C. A., Knight D., Croxford A. S., Maqsood A. R., Stephens G. M., Goodacre R., and Dickson A. J. (2010).
"Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets".
Metabolomics 6.3, pp. 427–438. DOI: 10.1007/s11306-010-0216-9.

Selvarasu S., Wong V. V. T., Karimi I. A., and Lee D.-Y. (2009). "Elucidation of metabolism in hybridoma cells grown in fed-batch culture by genome-scale modeling".
Biotechnology and Bioengineering 102.5, pp. 1494–504. DOI: 10.1002/bit.22186.

Sheikholeslami Z., Jolicoeur M., and Henry O. (2013). "Probing the metabolism of an inducible mammalian expression system using extracellular isotopomer analysis".
Journal of Biotechnology 164.4, pp. 469–78. DOI: 10.1016/j.jbiotec.2013.01.025.

Shimizu K. (2009). "Toward systematic metabolic engineering based on the analysis of metabolic regulation by the integration of different levels of information".
Biochemical Engineering Journal 46.3, pp. 235–251.
DOI: 10.1016/j.bej.2009.06.006.

Shin M. H., Lee D. Y., Liu K.-H., Fiehn O., and Kim K. H. (2010).
"Evaluation of sampling and extraction methodologies for the global metabolic profiling of Saccharophagus degradans." *Analytical Chemistry* 82.15, pp. 6660–6.
DOI: 10.1021/ac1012656.

Shlomi T., Eisenberg Y., Sharan R., and Ruppin E. (2007).
"A genome-scale computational study of the interplay between transcriptional regulation and metabolism". *Molecular Systems Biology* 3.101, p. 101.
DOI: 10.1038/msb4100141.

Slupsky C. M., Rankin K. N., Wagner J., Fu H., Chang D., Weljie A. M., Saude E. J., Lix B., Adamko D. J., Shah S., Greiner R., Sykes B. D., and Marrie T. J. (2007).
"Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles." *Analytical Chemistry* 79.18, pp. 6995–7004.
DOI: 10.1021/ac0708588.

Smith R. (1984). "Efficient Monte Carlo Procedures for Generating Points Uniformly Distributed Over Bounded Regions". *Informs* 32.6, pp. 1296–308.

Smolinska A., Blanchet L., Buydens L. M. C., and Wijmenga S. S. (2012). "NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review." *Analytica Chimica Acta* 750, pp. 82–97. DOI: 10.1016/j.aca.2012.05.049.

Soininen P., Haarala J., Vepsäläinen J., Niemitz M., and Laatikainen R. (2005). "Strategies for organic impurity quantification by 1H NMR spectroscopy: Constrained total-line-shape fitting". *Analytica Chimica Acta* 542.2, pp. 178–185. DOI: 10.1016/j.aca.2005.03.060.

Sokolenko S. and Aucoin M. G. (2015). "A correction method for systematic error in (1)H-NMR time-course data validated through stochastic cell culture simulation." *BMC Systems Biology* 9, p. 51. DOI: 10.1186/s12918-015-0197-4.

Sokolenko S., Blondeel E. J. M., Azlah N., George B., Schulze S., Chang D., and Aucoin M. G. (2014). "Profiling convoluted single-dimension proton NMR spectra: a Plackett-Burman approach for assessing quantification error of metabolites in complex mixtures with application to cell culture." *Analytical Chemistry* 86.7, pp. 3330–7. DOI: 10.1021/ac4033966.

Sokolenko S., McKay R., Blondeel E. J. M., Lewis M. J., Chang D., George B., and Aucoin M. G. (2013). "Understanding the variability of compound quantification from targeted profiling metabolomics of 1D-1H-NMR spectra in synthetic mixtures and urine with additional insights on choice of pulse sequences and robotic sampling". *Metabolomics* 9.4, pp. 887–903. DOI: 10.1007/s11306-013-0503-3.

Solomon B. O., Erickson L. E., Hess J. E., and Yang S. S. (1982). "Maximum likelihood estimation of growth yields". *Biotechnology and Bioengineering* 24.3, pp. 633–49. DOI: 10.1002/bit.260240309.

Stephanopoulos G. N., Alper H. S., and Moxley J. (2004). "Exploiting biological complexity for strain improvement through systems biology". *Nature Biotechnology* 22.10, pp. 1261–7. DOI: 10.1038/nbt1016.

Stephanopoulos G. N., Aristidou A. A., and Nielsen J. (1999). "Metabolic fluxes and metabolic engineering". *Metabolic Engineering* 1.1, pp. 1–11. DOI: 10.1006/mben.1998.0101.

Stephanopoulos G., Aristidou A., and Nielsen J. (1998). *Metabolic engineering: principles and methodologies*. Academic Press.

Steven H H., Kenneth L M., Kenneth D G., Timothy R B., Kevin L H., and Roy L M D. (2001). "Transformations in pharmaceutical research and development, driven by innovations in multidimensional mass spectrometry-based technologies". *International Journal of Mass Spectrometry* 212.1–3, pp. 135–196. DOI: http://dx.doi.org/10.1016/S1387-3806(01)00499-7.

Sue T., Obolonkin V., Griffiths H., and Villas-Bôas S. G. (2011). "An exometabolomics approach to monitoring microbial contamination in microalgal fermentation processes by using metabolic footprint analysis." *Applied and Environmental Microbiology* 77.21, pp. 7605–10. DOI: 10.1128/AEM.00469-11.

Sukumaran D. K., Garcia E., Hua J., Tabaczynski W., Odunsi K., Andrews C., and Szyperski T. (2009). "Standard operating procedure for metabonomics studies of blood serum and plasma samples using a 1H-NMR micro-flow probe." *Magnetic Resonance in Chemistry* 47 Suppl 1, S81–5. DOI: 10.1002/mrc.2469.

Szyperski T. (1998). "13C-NMR, MS and metabolic flux balancing in biotechnology research." *Quarterly reviews of biophysics* 31.1, pp. 41–106.

Tate S.-I. and Inagaki F. (1992). "Reduction of the water hump using a composite refocusing pulse". *Journal of Magnetic Resonance* 96.3, pp. 635–643.

Taymaz-Nikerel H., de Mey M., Ras C., ten Pierick A., Seifar R. M., van Dam J. C., Heijnen J. J., and van Gulik W. M. (2009). "Development and application of a differential method for reliable metabolome analysis in Escherichia coli." *Analytical Biochemistry* 386.1, pp. 9–19. DOI: 10.1016/j.ab.2008.11.018.

Toya Y., Kono N., Arakawa K., and Tomita M. (2011). "Metabolic flux analysis and visualization". *Journal of Proteome Research* 10.8, pp. 3313–23. DOI: 10.1021/pr2002885.

Tredwell G. D., Behrends V., Geier F. M., Liebeke M., and Bundy J. G. (2011a). "Between-person comparison of metabolite fitting for NMR-based quantitative metabolomics." *Analytical Chemistry* 83.22, pp. 8683–7. DOI: 10.1021/ac202123k.

Tredwell G. D., Edwards-Jones B., Leak D. J., and Bundy J. G. (2011b). "The development of metabolomic sampling procedures for Pichia pastoris, and baseline metabolome data." *PloS One* 6.1, e16286. DOI: 10.1371/journal.pone.0016286.

Tweeddale H., Notley-McRobb L., and Ferenci T. (1998). "Effect of slow growth on metabolism of Escherichia coli, as revealed by global metabolite pool ("metabolome") analysis." *Journal of Bacteriology* 180.19, pp. 5109–16.

Tyo K. E., Alper H. S., and Stephanopoulos G. N. (2007). "Expanding the metabolic engineering toolbox: more options to engineer cells". *Trends in Biotechnology* 25.3, pp. 132–7. DOI: 10.1016/j.tibtech.2007.01.003.

Vallino J. J. and Stephanopoulos G. N. (1990). "Flux determination in cellular bioreaction networks: applications to lysine fermentations". In: ed. by S. K. Sikdar and M. Bier. Boulder, Colorado, pp. 205–19.

Van den Meersche K., Soetaert K., and Van Oevelen D. (2009). "xsample(): An R Function for Sampling Linear Inverse Problems". *Journal of Statistical Software* 30.1.

Van der Heijden R. T., Romein B., Heijnen J. J., Hellinga C., and Luyben K. C. (1994). "Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors." *Biotechnology and Bioengineering* 43.1, pp. 11–20. DOI: 10.1002/bit.260430104.

Van der Werf M. J., Overkamp K. M., Muilwijk B., Coulier L., and Hankemeier T. (2007). "Microbial metabolomics: toward a platform with full metabolome coverage." *Analytical Biochemistry* 370.1, pp. 17–25. DOI: 10.1016/j.ab.2007.07.022.

Viant M. R., Bearden D. W., Bundy J. G., Burton I. W., Collette T. W., Ekman D. R., Ezernieks V., Karakach T. K., Lin C. Y., Rochfort S., de Ropp J. S., Teng Q., Tjeerdema R. S., Walter J. A., and Wu H. (2009). "International NMR-based environmental metabolomics intercomparison exercise." *Environmental Science & Technology* 43.1, pp. 219–25. DOI: 10.1021/es802198z.

Villas-Bôas S. G. and Bruheim P. (2007). "Cold glycerol-saline: the promising quenching solution for accurate intracellular metabolite analysis of microbial cells." *Analytical Biochemistry* 370.1, pp. 87–97. DOI: 10.1016/j.ab.2007.06.028.

Villas-Bôas S. G., Højer-Pedersen J., Akesson M., Smedsgaard J., and Nielsen J. (2005). "Global metabolite analysis of yeast: evaluation of sample preparation methods." *Yeast* 22.14, pp. 1155–69. DOI: 10.1002/yea.1308.

Volmer M., Northoff S., Scholz S., Thüte T., Büntemeyer H., and Noll T. (2011). "Fast filtration for metabolome sampling of suspended animal cells." *Biotechnology Letters* 33.3, pp. 495–502. DOI: 10.1007/s10529-010-0466-7.

Wagstaff J. L., Masterton R. J., Povey J. F., Smales C. M., and Howard M. J. (2013). "^{1}H NMR spectroscopy profiling of metabolic reprogramming of Chinese hamster ovary cells upon a temperature shift during culture." *PloS One* 8.10, e77195. DOI: 10.1371/journal.pone.0077195.

Walsh G. (2014). "Biopharmaceutical benchmarks 2014." *Nature Biotechnology* 32.10, pp. 992–1000. DOI: 10.1038/nbt.3040.

Wang N. S. and Stephanopoulos G. (1983). "Application of macroscopic balances to the identification of gross measurement errors." *Biotechnology and Bioengineering* 25.9, pp. 2177–208. DOI: 10.1002/bit.260250906.

Wang Q.-Z., Wu C.-Y., Chen T., Chen X., and Zhao X.-M. (2006). "Integrating metabolomics into a systems biology framework to exploit metabolic complexity: strategies and applications in microorganisms." *Applied Microbiology and Biotechnology* 70.2, pp. 151–61. DOI: 10.1007/s00253-005-0277-2.

Wang W., Sun J., Hartlep M., Deckwer W.-D., and Zeng A.-P. (2003). "Combined use of proteomic analysis and enzyme activity assays for metabolic pathway analysis of glycerol fermentation by Klebsiella pneumoniae." *Biotechnology and Bioengineering* 83.5, pp. 525–36. DOI: 10.1002/bit.10701.

Ward J. L., Baker J. M., Miller S. J., Deborde C., Maucourt M., Biais B., Rolin D., Moing A., Moco S., Vervoort J., Lommen A., Schäfer H., Humpfer E., and Beale M. H. (2010). "An inter-laboratory comparison demonstrates that [H]-NMR metabolite fingerprinting is a robust technique for collaborative plant metabolomic data collection." *Metabolomics* 6.2, pp. 263–273. DOI: 10.1007/s11306-010-0200-4.

Warrack B. M., Hnatyshyn S., Ott K.-H., Reily M. D., Sanders M., Zhang H., and Drexler D. M. (2009). "Normalization strategies for metabonomic analysis of urine samples." *Journal of Chromatography B* 877.5-6, pp. 547–52. DOI: 10.1016/j.jchromb.2009.01.007.

Weaver D. S., Keseler I. M., Mackie A., Paulsen I. T., and Karp P. D. (2014). "A genome-scale metabolic flux model of Escherichia coli K-12 derived from the EcoCyc database." *BMC Systems Biology* 8, p. 79. DOI: 10.1186/1752-0509-8-79.

Wei R. (2011). "Metabolomics and its practical value in pharmaceutical industry." *Current Drug Metabolism* 12.4, pp. 345–58.

Weiss S. A., Smith G. C., Kalter S. S., and Vaughn J. L. (1981). "Improved method for the production of insect cell cultures in large volume". *In Vitro* 17.8, pp. 744–744. DOI: 10.1007/BF02628413.

Weljie A. M., Newton J., Mercier P., Carlson E., and Slupsky C. M. (2006). "Targeted profiling: quantitative analysis of 1H NMR metabolomics data." *Analytical chemistry* 78.13, pp. 4430–42. DOI: 10.1021/ac060209g.

Wells R. J., Hook J. M., Al-Deen T. S., and Hibbert D. B. (2002).
"Quantitative nuclear magnetic resonance (QNMR) spectroscopy for assessing the purity of technical grade agrochemicals: 2,4-dichlorophenoxyacetic acid (2,4-D) and sodium 2,2-dichloropropionate (Dalapon sodium)."
Journal of Agricultural and Food Chemistry 50.12, pp. 3366–74.
DOI: 10.1021/jf0114379.

White B. C., Jamison K. M., Grieb C., Lally D., Luckett C., Kramer K. S., and Phillips J. (2010). "Specific gravity and creatinine as corrections for variation in urine concentration in humans, gorillas, and woolly monkeys."
American Journal of Primatology 72.12, pp. 1082–91. DOI: 10.1002/ajp.20867.

Wickham H. (2009). *ggplot2: Elegant Graphics for Data Analysis*.
Springer-Verlag New York.

Wickham H. (2011). "The Split-Apply-Combine Strategy for Data Analysis".
Journal of Statistical Software 40.1, pp. 1–29.

Wiechert W., Möllney M., Petersen S., and de Graaf A. A. (2001).
"A universal framework for 13C metabolic flux analysis".
Metabolic Engineering 3.3, pp. 265–83. DOI: 10.1006/mben.2001.0188.

Wiechert W. and Nöh K. (2005).
"From stationary to instationary metabolic flux analysis".
Advances in Biochemical Engineering/Biotechnology 92, pp. 145–172.
DOI: 10.1007/b98921.

Wiendahl C., Brandner J. J., Küppers C., Luo B., Schygulla U., Noll T., and Oldiges M. (2007).
"A microstructure heat exchanger for quenching the metabolism of mammalian cells".
Chemical Engineering & Technology 30.3, pp. 322–328.
DOI: 10.1002/ceat.200600362.

Winder C. L., Dunn W. B., Schuler S., Broadhurst D., Jarvis R., Stephens G. M., and Goodacre R. (2008). "Global metabolic profiling of Escherichia coli cultures: an evaluation of methods for quenching and extraction of intracellular metabolites."
Analytical Chemistry 80.8, pp. 2939–48. DOI: 10.1021/ac7023409.

Wishart D. S. (2008). "Quantitative metabolomics using NMR".
TrAC Trends in Analytical Chemistry 27.3, pp. 228–237.
DOI: 10.1016/j.trac.2007.12.001.

Wishart D. S., Tzur D., Knox C., Eisner R., Guo A. C., Young N., Cheng D., Jewell K., Arndt D., Sawhney S., Fung C., Nikolai L., Lewis M., Coutouly M.-A., Forsythe I., Tang P., Shrivastava S., Jeroncic K., Stothard P., Amegbey G., Block D., Hau D. D., Wagner J., Miniaci J., Clements M., Gebremedhin M., Guo N., Zhang Y., Duggan G. E., Macinnis G. D., Weljie A. M., Dowlatabadi R., Bamforth F., Clive D., Greiner R., Li L., Marrie T., Sykes B. D., Vogel H. J., and Querengesser L. (2007). "HMDB: the Human Metabolome Database."
Nucleic Acids Research 35.Database issue, pp. D521–6. DOI: 10.1093/nar/gkl923.

Wood S. N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models".
Journal of the Royal Statistical Society: Series B 73.1, pp. 3–36.
DOI: 10.1111/j.1467-9868.2010.00749.x.

Work E. (1949). "Chromatographic investigations of amino acids from micro-organisms".
Biochimica et Biophysica Acta 3, pp. 400–411.
DOI: 10.1016/0006-3002(49)90111-0.

Wu P. S. C. and Otting G. (2005).
"Rapid pulse length determination in high-resolution NMR."
Journal of Magnetic Resonance 176.1, pp. 115–9. DOI: 10.1016/j.jmr.2005.05.018.

Xing Z., Kenty B., Koyrakh I., Borys M., Pan S.-H., and Li Z. (2011). "Optimizing amino acid composition of CHO cell culture media for a fusion protein production".
Process Biochemistry 46.7, pp. 1423–1429. DOI: 10.1016/j.procbio.2011.03.014.

Xu X., Nagarajan H., Lewis N. E., Pan S., Cai Z., Liu X., Chen W., Xie M., Wang W., Hammond S., Andersen M. R., Neff N., Passarelli B., Koh W., Fan H. C., Wang J., Gui Y., Lee K. H., Betenbaugh M. J., Quake S. R., Famili I., Palsson B. O., and Wang J. (2011).
"The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line."
Nature Biotechnology 29.8, pp. 735–41. DOI: 10.1038/nbt.1932.

Yen S., Sokolenko S., Manocha B., Blondeel E. J. M., Aucoin M. G., Patras A., Daynouri-Pancino F., and Sasges M. (2014). "Treating cell culture media with UV irradiation against adventitious agents: minimal impact on CHO performance."
Biotechnology Progress 30.5, pp. 1190–5. DOI: 10.1002/btpr.1942.

Yoon S. H., Han M.-J., Lee S. Y., Jeong K. J., and Yoo J.-S. (2003).
"Combined transcriptome and proteome analysis of Escherichia coli during high cell density culture". *Biotechnology and Bioengineering* 81.7, pp. 753–67.
DOI: 10.1002/bit.10626.

Young J. D. (2014). "(13)C metabolic flux analysis of recombinant expression hosts."
Current Opinion in Biotechnology 30, pp. 238–45.
DOI: 10.1016/j.copbio.2014.10.004.

Zelena E., Dunn W. B., Broadhurst D., Francis-McIntyre S., Carroll K. M., Begley P., O'Hagan S., Knowles J. D., Halsall A., Consortium H., Wilson I. D., and Kell D. B. (2009). "Development of a robust and repeatable UPLC-MS method for the long-term metabolomic study of human serum." *Analytical Chemistry* 81.4, pp. 1357–64. DOI: 10.1021/ac8019366.

Zhang A., Sun H., Wang P., Han Y., and Wang X. (2012a). "Modern analytical techniques in metabolomics analysis." *The Analyst* 137.2, pp. 293–300. DOI: 10.1039/c1an15605e.

Zhang A., Sun H., Wang P., Han Y., and Wang X. (2012b). "Recent and potential developments of biofluid analyses in metabolomics." *Journal of Proteomics* 75.4, pp. 1079–88. DOI: 10.1016/j.jprot.2011.10.027.

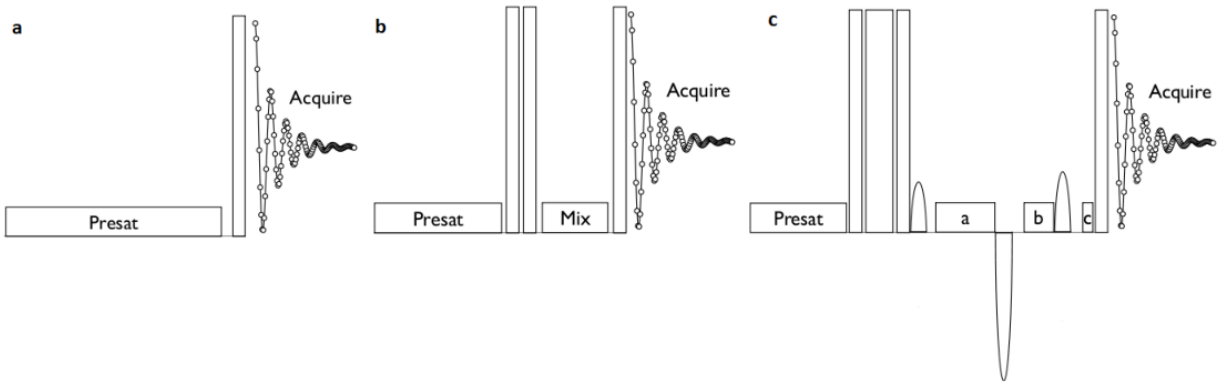
Zheng C., Zhang S., Ragg S., Raftery D., and Vitek O. (2011). "Identification and quantification of metabolites in (1)H NMR spectra by Bayesian model selection." *Bioinformatic* 27.12, pp. 1637–44. DOI: 10.1093/bioinformatics/btr118.

Zulyniak M. A. and Mutch D. M. (2011). "Harnessing metabolomics for nutrition research." *Current Pharmaceutical Biotechnology* 12.7, pp. 1005–15.

Appendix A

Supplementary Information for Chapter 2

A.1 Pulse sequence details



a) Presat 1D-1H A simple 1D-1H presaturation experiment was used for carrier frequency and pulse width optimization, and is universally available on major manufacturer's spec-

trometers. The presaturation period (wide low rectangle denoting relatively low power) was comprised of continuous wave irradiation at the carrier position tuned to the solvent peak via an empirical array of the frequency and then observing the residual signal minima. The power setting was adjusted to deliver $\sim$ 80 Hz of gamma-B_1 induced field strength for convenient periods of $\sim$ 2.5 seconds. The optimized pulse width for high power ($\sim$ 33 kHz gamma-B_1) excitation (shown as tall narrow rectangles) was either determined by arraying the pulse period and observing a 360° rotation, and then dividing by a factor of 4 for the 90° (Agilent Inova console), or by use of the newer nutation single pulse calibration available on digital architectures (e.g. Agilent VNMRS console). The pulse sequence typically performs little phase cycling except basic Cyclops (i.e. rotation of the excitation and observed phase by 90° per increment). **b) 1D-NOESY_s1a4** The 1D-NOESY, or metnoesy, contained an initial delay (10 ms), then a presaturation period of 990 ms using the same power settings as the 1D-1H presaturation detailed above. Following the presaturation period, two 90° high power pulses were followed by a "mixing" period (100 ms) also executed with saturation of the solvent peak (same power as presaturation period). Lastly a 90° high power pulse is executed before a 4 second acquisition period. The phase cycle and subsequent magnetization behaviour can be deceptively complex, and has been detailed elsewhere (Mckay 2011). **c) grd-NOESY_s1a4** Lastly, for comparison, a modified 1D-NOESY utilizing a simple composite inversion pulse ($90^\circ_x 180^\circ_y 90^\circ_x$ (Bax 1985; Levitt 1986; Tate and Inagaki 1992)) and pulsed field gradients during the mixing time was acquired for each sample. The grd-NOESY has the same period and power presaturation as the standard 1D-NOESY but is followed by the composite pulse that alternates between an effective 180° and 0° pulse on odd and even transients, respectively via the phase cycle. This was followed by a "mix"

delay totaling 100 ms in length, split into three solvent saturation of 59, 29, and 9 ms each. These sections were separated by pulse field gradients (Keeler et al. 1994) of 1 ms duration and amplitudes of 3.75, -18.75, and 5.625 Gauss/cm (i.e. 2:-10:3) respectively. Lastly a 90° pulse perturbed bulk magnetization into the transverse plane for the 4 second acquisition period. The sweep width used on all experiments (600 MHz instrument) was 7225.4 Hz, with 4 steady states before collection of 32 transients. Data were zero filled to twice the number of acquired points, and an apodization window of 0.5 Hz line broadening was applied before fast Fourier transformation and analysis with Chenomx Suite software.

Bax A. (1985). "A spatially selective composite 90° radiofrequency pulse". *Journal of Magnetic Resonance* 65, pp. 142–145.

Keeler J., Clowes R. T., Davis A. L., and Laue E. D. (1994). "Pulsed-field gradients: theory and practice." *Methods in Enzymology* 239, pp. 145–207.

Levitt M. H. (1986). *Composite pulses*. Wiley Online Library.

Mckay R. T. (2011). "How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution". *Concepts in Magnetic Resonance A* 38A.5, pp. 197–220. DOI: 10.1002/cmr.a.20223.

Tate S.-I. and Inagaki F. (1992). "Reduction of the water hump using a composite refocusing pulse". *Journal of Magnetic Resonance* 96.3, pp. 635–643.

A.2 Compounds quantified in urine

Table A.1: Full list of compounds profiled in the urine sample. Mean and standard deviation values calculated from profiling 11 fully replicated scans (including sample re-insertion) of a single NMR tube.

Compound	Mean (mM)	Relative standard deviation (%)
Urea	29.136	0.81
Creatinine	6.894	0.47
Hippurate	4.218	1.38
Citrate	1.535	0.77
Trigonelline	0.844	0.86
Trimethylamine N-oxide	0.583	0.46
Glycine	0.435	1.40
Formate	0.384	1.12
Glycolate	0.34	7.40
3-Indoxylsulfate	0.329	4.95
Ethanolamine	0.268	6.32
Glucose	0.224	6.88
Pyroglutamate	0.197	24.96
Dimethylamine	0.168	2.44
Creatine	0.163	5.99
cis-Aconitate	0.157	14.53
1,3-Dimethylurate	0.130	1.31
Taurine	0.126	17.66
Malonate	0.122	7.04
3-Aminoisobutyrate	0.117	3.92
Glutamine	0.114	33.01
Succinate	0.096	64.26
Alanine	0.085	3.86
Carnitine	0.083	13.24

Continued on next page

Table A.1 – continued from previous page

Compound	Mean (mM)	Relative standard deviation (%)
Histidine	0.081	27.48
Betaine	0.081	1.80
Xanthine	0.080	9.54
π-Methylhistidine	0.080	41.98
Acetate	0.079	2.15
τ-Methylhistidine	0.078	18.09
4-Hydroxyphenylacetate	0.073	3.25
Phenylalanine	0.070	14.98
Methylamine	0.062	3.40
Asparagine	0.053	72.22
Dimethyl sulfone	0.051	1.65
Sucrose	0.050	5.72
3-Methylxanthine	0.050	5.82
trans-Aconitate	0.048	22.98
Lactate	0.041	9.13
2-Hydroxyisobutyrate	0.041	2.23
3-Hydroxyisovalerate	0.041	17.83
Tyramine	0.038	14.10
1,6-Anhydro-β-D-glucose	0.037	14.35
Tyrosine	0.036	9.59
Ethanol	0.036	37.99
Lysine	0.036	47.82
Methanol	0.035	1.82
Hypoxanthine	0.035	4.82
N,N-Dimethylglycine	0.032	36.85
Acetone	0.031	3.32
1-Methylnicotinamide	0.028	9.93
N-Acetylaspartate	0.021	51.13
Trimethylamine	0.016	41.61
Leucine	0.015	35.75
Valine	0.013	7.68
Isobutyrate	0.004	16.16

Table A.2: List of compounds profiled in the urine sample that were excluded from the main analysis as their presence in the sample was deemed ambiguous without confirmation by other analytical methods. Mean and standard deviation values calculated from profiling 11 fully replicated scans (including sample re-insertion) of a single NMR tube.

Compound	Mean (mM)	Relative standard deviation (%)
O-Phosphoserine	0.665	17.70
Homoserine	0.344	29.19
Gluconate	0.313	45.63
Serine	0.301	20.53
Glycylproline	0.263	24.24
Arginine	0.213	23.92
Mannitol	0.157	52.94
Guanidoacetate	0.142	26.42
Carnosine	0.103	21.95
N-Methylhydantoin	0.089	36.77
Lactose	0.089	62.04
Glutamate	0.088	27.69
Methylsuccinate	0.085	16.73
Tryptophan	0.059	46.21
N-Acetylglycine	0.055	49.91
Acetoacetate	0.054	28.38
Benzoate	0.052	35.66
Phenylacetate	0.051	69.55
Succinylacetone	0.049	25.17
N-Nitrosodimethylamine	0.049	58.75
Tiglylglycine	0.048	21.88
2-Oxoglutarate	0.047	24.10
Imidazole	0.047	59.04
Anserine	0.045	23.27
Salicylurate	0.044	38.69
Xylose	0.040	28.56

Continued on next page

Table A.2 – continued from previous page

Compound	Mean (mM)	Relative standard deviation (%)
Fucose	0.040	42.54
Homocystine	0.039	26.65
O-Acetylcarnitine	0.034	36.47
N-Acetylglutamate	0.033	33.94
Methionine	0.021	64.89
Thymol	0.019	33.43
Tartrate	0.018	17.19
Butanone	0.017	34.06
Isoleucine	0.015	23.63
Isopropanol	0.013	22.46
Fumarate	0.008	19.78
N-Acetylglutamine	0.001	331.66

Appendix B

Supplementary Information for Chapter 3

B.1 Supplementary tables

Table B.1: List of compounds used to make concentrated compound stocks.

Compound	Source	Cas No.	CheBI
Acetate	Sigma Aldrich	64-19-7	CHEBI:15366
Alanine	Sigma Aldrich	56-41-7	CHEBI:16977
Arabinose	Sigma Aldrich	5328-37-0	CHEBI:30849
Arginine	Acros Organics	1119-34-2	CHEBI:16467
Asparagine	Sigma Aldrich	70-47-3	CHEBI:17196
Aspartate	Sigma Aldrich	56-84-8	CHEBI:17053

Continued on next page

Table B.1 – continued from previous page

Compound	Source	Cas No.	CheBI
Choline	Sigma Aldrich	67-48-1	CHEBI:133341
Formate	Sigma Aldrich	64-18-6	CHEBI:30751
Glucose	Thermo Scientific	50-99-7	CHEBI:17634
Glutamate	Alfa Aesar	56-86-0	CHEBI:16015
Glutamine	Thermo Scientific	56-85-9	CHEBI:18050
Glycine	Sigma Aldrich	56-40-6	CHEBI:15428
Isoleucine	Sigma Aldrich	73-32-5	CHEBI:17191
Lactate	Sigma Aldrich	79-33-4	CHEBI:422
Leucine	Nutritional Biochemical Corp.	61-90-5	CHEBI:15603
Lysine	Sigma Aldrich	56-87-1	CHEBI:18019
Methionine	Sigma Aldrich	63-68-3	CHEBI:16643
myo-Inositol	Sigma Aldrich	87-89-8	CHEBI:17268
Phenylalanine	Sigma Aldrich	63-91-2	CHEBI:17295
Proline	Acros Organics	147-85-3	CHEBI:17203
Pyroglutamate	Sigma Aldrich	98-79-3	CHEBI:18183
Pyruvate	Acros Organic	127-17-3	CHEBI:32816
Serine	Sigma Aldrich	56-45-1	CHEBI:17115
Succinate	Sigma Aldrich	110-15-6	CHEBI:15741
Threonine	Sigma Aldrich	72-19-5	CHEBI:16857
Tryptophan	Sigma Aldrich	73-22-3	CHEBI:16828
Tyrosine	Sigma Aldrich	60-18-4	CHEBI:17895
Valine	Sigma Aldrich	72-18-4	CHEBI:16414

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
pH	7.12	6.24	6.24	7.21	7.12	6.5	6.31	6.4	6.5	6.98	6.13	7.21	6.28	7	7.16	7.17	7.03	6.19	6.14	7.23
Glucose	20	20	9.7	9.7	20	20	20	20	9.7	20	9.7	20	9.7	9.7	9.7	9.7	20	20	9.7	9.7
Lactate	0.83	6.2	6.2	0.83	0.83	6.2	6.2	6.2	6.2	0.83	6.2	0.83	6.2	0.83	0.83	0.83	0.83	6.2	6.2	0.83
Proline	2.1	1.3	2.1	2.1	1.3	1.3	2.1	2.1	2.1	2.1	1.3	2.1	1.3	2.1	1.3	1.3	1.3	1.3	2.1	1.3
Alanine	2.4	2.4	0	2.4	2.4	0	0	2.4	2.4	2.4	2.4	0	2.4	0	2.4	0	0	0	0	0
Isoleucine	0.31	1.4	1.4	0.31	1.4	1.4	0.31	0.31	1.4	1.4	1.4	1.4	0.31	1.4	0.31	1.4	0.31	0.31	0.31	0.31
Valine	0.53	0.53	1.2	1.2	0.53	1.2	1.2	0.53	0.53	1.2	1.2	1.2	1.2	0.53	1.2	0.53	1.2	0.53	0.53	0.53
Leucine	0.25	0.25	0.25	1.2	1.2	0.25	1.2	1.2	0.25	0.25	1.2	1.2	1.2	1.2	0.25	1.2	0.25	1.2	0.25	0.25
Methionine	0.21	0.21	0.21	0.21	0.36	0.36	0.21	0.36	0.36	0.21	0.21	0.36	0.36	0.36	0.36	0.21	0.36	0.21	0.36	0.21
Glycine	1	0.033	0.033	0.033	0.033	1	1	0.033	1	1	0.033	0.033	1	1	1	1	0.033	1	0.033	0.033
Glutamate	0.049	0.56	0.049	0.049	0.049	0.049	0.56	0.56	0.049	0.56	0.56	0.049	0.049	0.56	0.56	0.56	0.56	0.049	0.56	0.049
Tryptophan	0.12	0.063	0.12	0.063	0.063	0.063	0.063	0.12	0.12	0.063	0.12	0.12	0.063	0.063	0.12	0.12	0.12	0.12	0.063	0.063
Arabinose	0.035	0.1	0.035	0.1	0.035	0.035	0.035	0.035	0.1	0.1	0.035	0.1	0.1	0.035	0.035	0.1	0.1	0.1	0.1	0.035
Formate	0.57	0	0.57	0	0.57	0	0	0	0	0.57	0.57	0	0.57	0.57	0	0	0.57	0.57	0.57	0
Glutamine	4	4	0	4	0	4	0	0	0	0	4	4	0	4	4	0	0	4	4	0
Choline	0.059	0.059	0.059	0.059	0.015	0.059	0.015	0.059	0.015	0.015	0.015	0.015	0.059	0.059	0.015	0.059	0.059	0.015	0.015	0.015
Succinate	0.002	0.045	0.045	0.045	0.045	0.002	0.045	0.002	0.045	0.002	0.002	0.002	0.002	0.045	0.045	0.002	0.045	0.045	0.002	0.002
Arginine	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
Lysine	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
Threonine	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89
Tyrosine	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
Phenylalanine	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43
Serine	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Pyroglutamate	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27
Pyruvate	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Asparagine	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Myo-Inositol	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Aspartate	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057	0.057
Acetate	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087

Table B.2: Theoretical compound concentrations (mM) and observed pH values of the synthetic mixtures

Table B.3: Convolution matrix of profiled compounds. Compounds were judged to convolute if any of their spectra clusters had more than ∼5% overlap or if there was a possibility of misidentification

Compound	Cluster location (ppm)	Convoluting compounds
Acetate	1.9	arginine, lysine
Alanine	3.8	glucose, glutamine, arginine
	1.5	isoleucine, lysine
Arabinose	4.1	asparagine, phenylalanine, serine, glucose
	3.9	serine
	3.6	isoleucine
Arginine	3.8	glucose, glutamine
	3.2	glucose
	1.9	lysine, acetate
	1.7	leucine, lysine
Asparagine	4.1	tryptophan, myo-inositol, choline
Aspartate	3.9	glucose
	2.6	methionine
Choline	4.1	myo-inositol, tryptophan
	3.2	tyrosine
Glucose	3.8	methionine, serine
	3.7	leucine
	3.4	proline
Glutamate	3.7	glucose, glutamine, lysine
	2.3	pyroglutamate, proline
	2.1	glutamine, methionine
	2.0	proline, pyroglutamate
Glutamine	3.7	glucose, arginine
	2.5	pyroglutamate
	2.4	pyroglutamate, succinate
	2.1	methionine, proline, glutamate

Continued on next page

Table B.3 – continued from previous page

Compound	Cluster location (ppm)	Convoluting compounds
Isoleucine	2.0	proline
	1.5	alanine, lysine
	0.9	leucine
Lactate	4.1	proline
	1.3	threonine
Leucine	3.7	glucose, lysine
	1.7	lysine, arginine
	0.9	isoleucine, valine
Lysine	3.7	glucose, glutamine, leucine
	3.0	tyrosine
	1.9	arginine, acetate
	1.7	leucine, arginine
	1.5	isoleucine, lysine, alanine
Methionine	3.8	glucose, serine
	2.6	aspartate
	2.1	glutamine, glutamate
Myo-inositol	4.1	choline, tryptophan, asparagine
	3.6	valine
	3.5	glucose
	3.3	glucose, phenylalanine, arginine, tryptophan
Phenylalanine	7.3	tryptophan
	4.0	serine
	3.3	glucose, myo-inositol, arginine
	3.1	
Proline	4.1	lactate
	3.4	glucose
	3.3	tryptophan
	2.3	pyruvate, glutamate
	2.1	glutamine, glutamate, methionine, pyroglutamate
	2.0	pyroglutamate, isoleucine
Pyroglutamate	2.5	glutamine

Continued on next page

Table B.3 – continued from previous page

Compound	Cluster location (ppm)	Convoluting compounds
	2.4	succinate
	2.3	glutamine
	2.1	proline, glutamate
Pyruvate	2.3	proline
Serine	4.0	phenylalanine
	3.9	arabinose, glucose, tyrosine
	3.8	glucose, methionine
Succinate	2.4	pyroglutamate, glutamine
Threonine	1.3	lactate
Tryptophan	7.3	phenylalanine
	7.2	tyrosine
	4.1	asparagine, myo-inositol, choline, arabinose
	3.5	glucose
	3.3	proline, myo-inositol
Tyrosine	7.2	tryptophan
	3.9	serine, glucose
	3.2	choline
	3.0	lysine
Valine	3.6	myo-inositol
	0.9	leucine

B.2 Supplementary figures

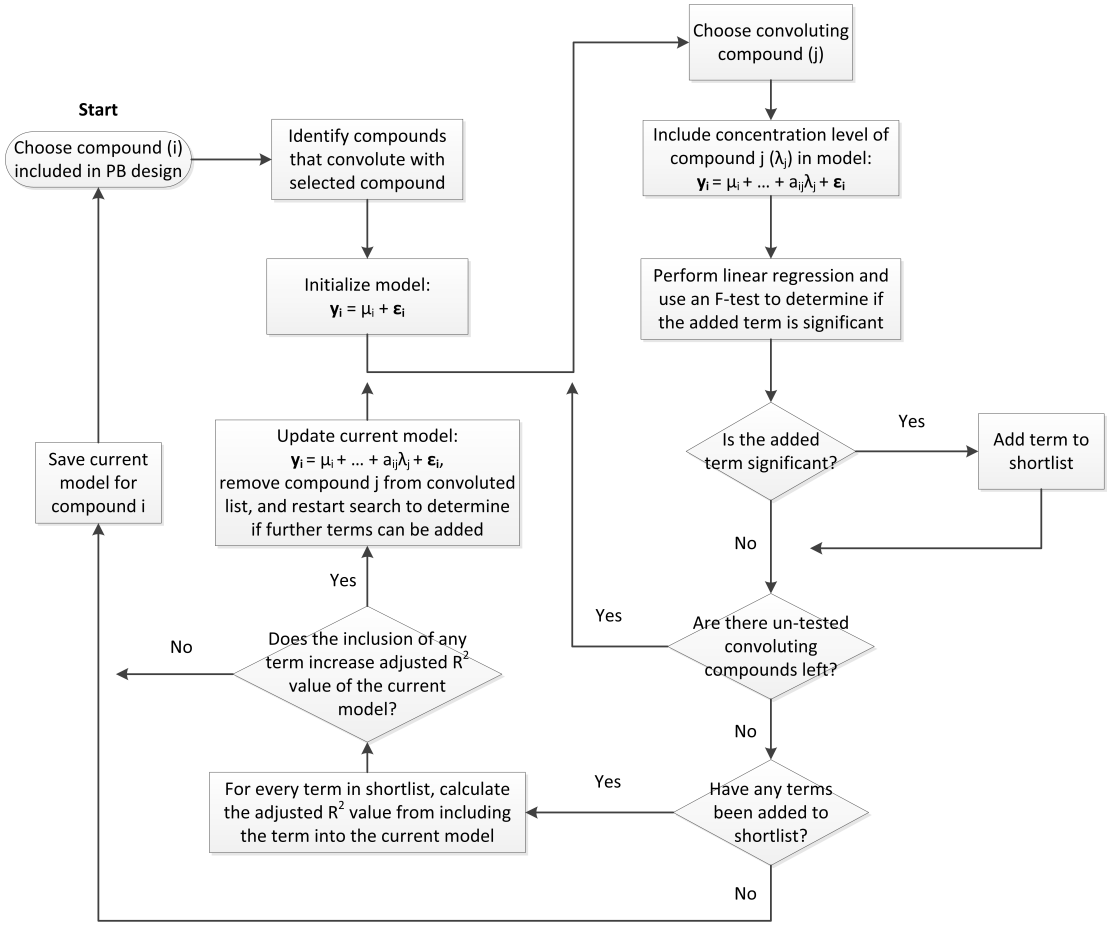


Figure B.1: A schematic of the iterative linear regression algorithm used to estimate the effect of one compound's concentration level on the measured concentration of another. In the equations, $\mathbf{y}_i$ is the column vector of concentrations of compound i from all samples where compound i is at a single concentration level, μ_i is the mean concentration of compound i across those samples, λ_j is the concentration level of compound j whose resonance overlaps with that of compound i (1 for high, 0 for low), a_{ij} is the regression coefficient, J is the set of all varied compounds, and $\boldsymbol{\varepsilon_i}$ is a column of residuals.

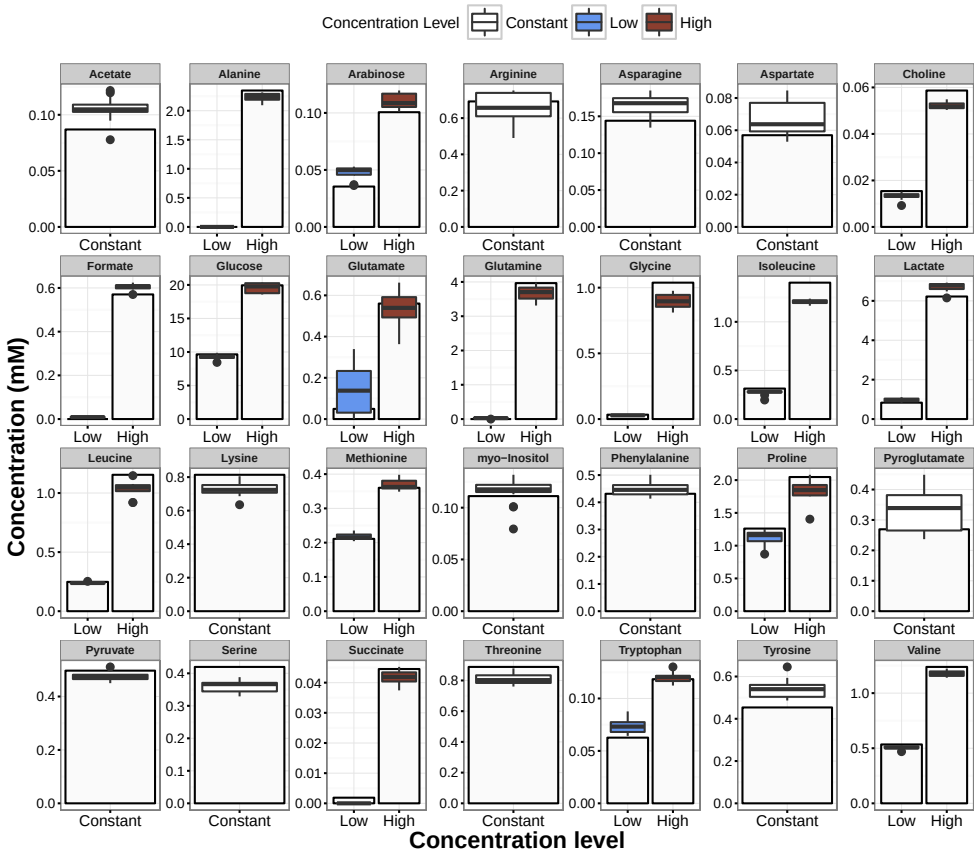


Figure B.2: Comparison of theoretical compound concentrations (grey bars) to overlaid box and whisker plots of observed values for all concentration levels from the Plackett-Burman experiment. Each box plot represents the distribution of all the observed concentrations for each compound. For compounds that were added at two levels, the data was split into a low and high level accordingly. Box plots were generated in standard fashion, with the whiskers extending out to the furthest observation that is still within a length of two interquartile ranges from the median. Outliers that extend beyond these ranges are identified by individual points. As the interpretation of a percentage value can be quite difficult, the tabular data presented in the main body of the manuscript is depicted here in a graphic form. From this plot, it's possible to get context about what the percentage values represent in terms of observed concentrations (such as the high error and variability of glutamate). It is also possible to see the prevalence of outliers (such as those of myo-inositol), which cannot be determined from a simple standard deviation value.

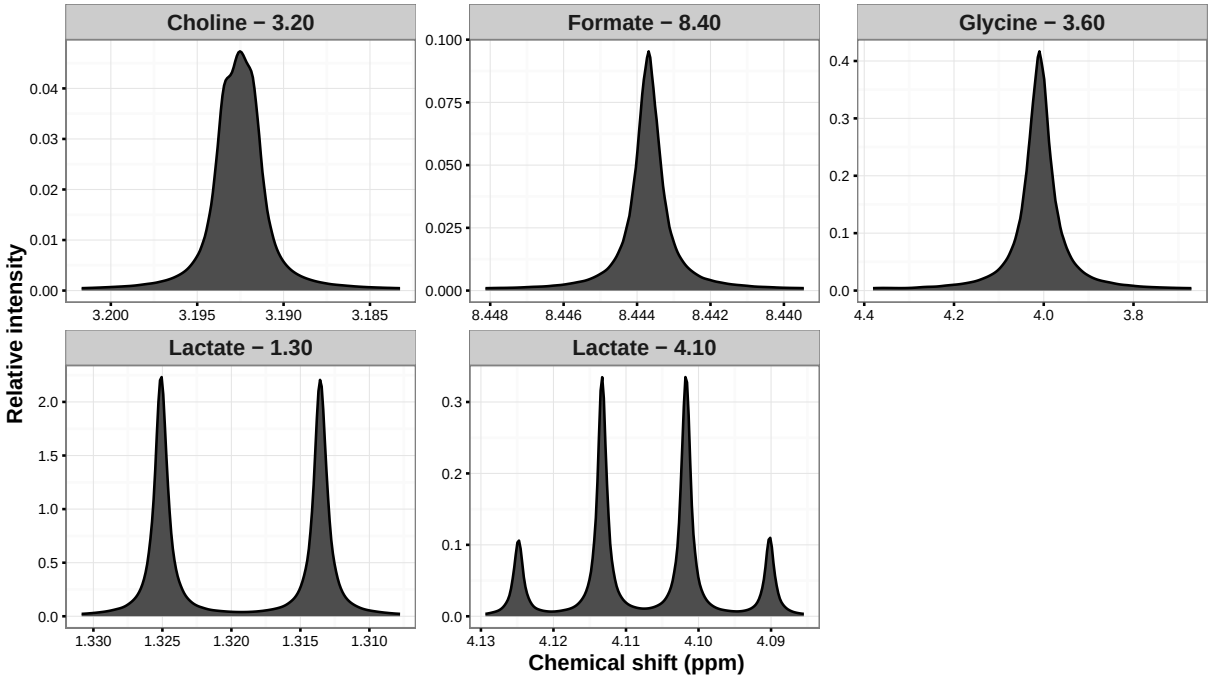


Figure B.3: An example of the type of sharp pure compound spectra peaks generated form the Chenomx library that resulted in ambiguous quantification (choline, formate, glycine, and lactate at their high concentration levels). For a relative comparison of peak sharpness, the chemical shift range of each cluster's subplot corresponds to relative intensity values greater than 1% of that cluster's peak intensity.

Appendix C

Supplementary Information for Chapter 5

C.1 Extended theoretical principles

To make the proposed protocol as accessible as possible, the theoretical section has been extended with a number extra details as well as a simplified example model.

C.1.1 Basic principles of MFA[1]

The basis of MFA stems from a mass balance on each intracellular compound (Equation C.1).

$$\text{rate of change} = \text{transport in} - \text{transport out}$$

[1]Mathematical description of MFA is largely based on Stephanopoulos et al. 1998.

$$+ \text{ generation } - \text{ consumption} \tag{C.1}$$

The right hand side of the equation can be separated into a net rate of formation and a dilution term that results from increasing cellular volume through cell division (Equation C.8),

$$\frac{\mathrm{d}C}{\mathrm{d}t} = r - \mu C \tag{C.2}$$

where $\frac{\mathrm{d}C}{\mathrm{d}t}$ is a column vector of metabolite rates of change, r is a column vector of net fluxes, and μC is the dilution term. r combines transport fluxes and consumption/production from reactions (which include biomass production). Due to high turnover rate of intracellular metabolite pools, both $\frac{\mathrm{d}C}{\mathrm{d}t}$ and μC are generally assumed to be negligible in relation to large flux magnitudes (termed "pseudo-steady state"). The result is that the sum of transport and reaction fluxes must be equal to zero.

$$0 = r \tag{C.3}$$

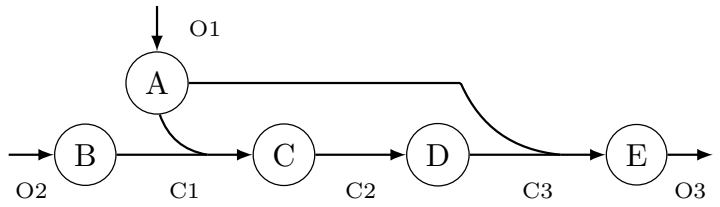
The sum of transport and reaction fluxes can be expressed as a product of a stoichiometry matrix S and a column vector of fluxes v (Equation C.4).

$$r = Sv \tag{C.4}$$

Each row of S corresponds to a metabolite, with the columns encoding the stoichiometry of reactions included in the model. As an example, a simple set of reactions is presented in Box 1, with the corresponding stoichiometry matrix defined in Equation C.5. Fluxes

O1-O3 can be observed, while fluxes C1-C3 need to be calculated.

O1	$\mathrm{A}_{\mathrm{external}} \longrightarrow \mathrm{A}$
O2	$\mathrm{B}_{\mathrm{external}} \longrightarrow \mathrm{B}$
C1	$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{C}$
C2	$\mathrm{C} \longrightarrow \mathrm{D}$
C3	$\mathrm{A} + \mathrm{D} \longrightarrow \mathrm{E}$
O3	$\mathrm{E} \longrightarrow \mathrm{E}_{\mathrm{external}}$



Box 1: Simple set of reactions defining a metabolic model.

$$S = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{cccccc} C1 & C2 & C3 & O1 & O2 & O3 \end{array} \\ \begin{bmatrix} -1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} \end{array} \tag{C.5}$$

Letters A, B, C, D, and E designate row names and correspond to balances on metabolites A, B, C, D, and E. C1, C2, C3, O1, O2, and O3 designate column names and refer to observed and calculated reactions. S can be partitioned into observed S_o and unknown (or calculated S_c) components,

$$Sv = S_c v_c + S_o v_o \tag{C.6}$$

where the observed fluxes tend to be metabolite uptake or secretion rates. For the reactions

in Box 1,

$$S_o = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \quad S_c = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{bmatrix} \end{array} \tag{C.7}$$

Combining Equations (C.3), (C.4), and (C.6),

$$0 = S_c v_c + S_o v_o \tag{C.8}$$

Since S_o, S_c and v_o are all known, v_c can then be calculated. Going back to the example, assume that the uptakes of A and B have been observed as 2.1 and 0.9, while the secretion of E was found to be 1.2:

$$0 = S_c v_c + S_o v_o \tag{C.9}$$

$$0 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} C1 \\ C2 \\ C3 \end{array} \begin{bmatrix} v_{C1} \\ v_{C2} \\ v_{C3} \end{bmatrix} + \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{array} \begin{array}{c} O1 \\ O2 \\ O3 \end{array} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} \tag{C.10}$$

The calculation and validation of v_c can be approached in multiple ways.

C.1.2 Traditional calculation

Assuming that sufficient fluxes can be observed, Equation (C.8) can be used to solve for the value of v_c through linear algebra:

$$0 = S_c v_c + S_o v_o \tag{C.11}$$

$$-S_c v_c = S_o v_o \tag{C.12}$$

To isolate for v_c, S_c must be invertible. If S_c has more rows than columns (and all rows are independent), then multiply both sides by S_c^T. Since $S_c^T S_c$ is square, it's possible to calculate v_c using a matrix inverse.

$$(S_c^T) - S_c v_c = (S_c^T) S_o v_o \tag{C.13}$$

$$v_c = -\left(S_c^T S_c\right)^{-1} S_c^T S_o v_o \tag{C.14}$$

For the example model:

$$\begin{matrix} C1 \\ C2 \\ C3 \end{matrix} \begin{bmatrix} 0.975 \\ 1.050 \\ 1.125 \end{bmatrix} \tag{C.15}$$

If v_o is known exactly, then Equation (C.14) can be substituted back into Equation (C.8) to define a redundency matrix R, which represents the flux balances around each metabolite

as a function of the observed fluxes alone.

$$0 = S_c \left[- \left(S_c^T S_c \right)^{-1} S_c^T S_o v_o \right] + S_o v_o \tag{C.16}$$

$$0 = \left[-S_c \left(S_c^T S_c \right)^{-1} S_c^T S_o \right] v_o + S_o v_o \tag{C.17}$$

$$0 = \left[S_o - S_c \left(S_c^T S_c \right)^{-1} S_c^T S_o \right] v_o \tag{C.18}$$

$$0 = R v_o \tag{C.19}$$

For the example model:

$$R = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.3333 & -0.3333 & -0.3333 \\ -0.3333 & 0.5833 & 0.0833 \\ 0 & 0.2500 & -0.2500 \\ 0 & 0.2500 & -0.2500 \\ 0.3333 & -0.0833 & -0.5833 \end{bmatrix} \end{array} \tag{C.20}$$

Given that $Rv_o = 0$, each row of R represents a relationship that must be satisfied by v_{O1}, v_{O2}, and v_{O3}. For example the balance on A will only hold if $v_{O1} = v_{O2} + v_{O3}$. Whereas v_o is a set of ideal observations that satisfy the balance perfectly, fluxes observed in practice ($\hat{v}_o$) are subject to experimental error and will result in some of the flux balances failing to close:

$$\varepsilon = R\hat{v}_o \tag{C.21}$$

where ε is the vector of residual fluxes associated with the flux balances around each of the metabolites (the length of ε is equal to the number of metabolites in the model). For

the example model:

$$\varepsilon = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.3333 & -0.3333 & -0.3333 \\ -0.3333 & 0.5833 & 0.0833 \\ 0 & 0.2500 & -0.2500 \\ 0 & 0.2500 & -0.2500 \\ 0.3333 & -0.0833 & -0.5833 \end{bmatrix} \end{array} \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} \tag{C.22}$$

$$= \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{bmatrix} 0 \\ -0.075 \\ -0.075 \\ -0.075 \\ -0.075 \end{bmatrix} \tag{C.23}$$

By considering $\hat{v}_o$ as the sum of hypothetically true values (v_o) and a vector of observation errors (δ),

$$\hat{v}_o = v_o + \delta \tag{C.24}$$

it's possible to describe ε as a function of observation error alone. Starting from Equation C.21:

$$\begin{aligned} \varepsilon &= R\hat{v}_o \\ &= R(v_o + \delta) \end{aligned} \tag{C.25}$$

Given that $Rv_o = 0$ by definition:

$$\varepsilon = R\delta \tag{C.26}$$

Relationships among the reactions in the metabolic model will cause linear dependencies in the redundancy matrix (R). In the simple reaction network of $A \rightarrow B \rightarrow C$, for example, a discrepancy in the observed transport fluxes of A and C will cause the same residual fluxes around A, B, and C, stemming from three linearly dependent rows in R. The degree of freedom that can be used to assess whether a significant observation error has been made is equal to the number of independent rows (rank) of R. For this reason, the redundancy matrix is reduced to only the linearly independent rows (R'), resulting in a corresponding change to ε (ε'). For the example model, R has a rank of two, so R' can be taken as two linearly independent rows of R:

$$R' = \begin{array}{c} \begin{matrix} O1 & O2 & O3 \end{matrix} \\ \begin{bmatrix} 0.3333 & -0.3333 & -0.3333 \\ 0 & 0.2500 & -0.2500 \end{bmatrix} \end{array} \tag{C.27}$$

The vector of residuals ε is similarly reduced to ε':

$$\varepsilon' = R'\hat{v}_o \tag{C.28}$$

$$\varepsilon' = R'\delta \tag{C.29}$$

For the example model:

$$\varepsilon' = \begin{array}{c} \begin{matrix} O1 & O2 & O3 \end{matrix} \\ \begin{bmatrix} 0.3333 & -0.3333 & -0.3333 \\ 0 & 0.2500 & -0.2500 \end{bmatrix} \end{array} \begin{array}{c} \begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \\ \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} \end{array} \tag{C.30}$$

$$= \begin{bmatrix} 0 \\ -0.075 \end{bmatrix} \tag{C.31}$$

The test for "gross measurement error" consists of determining whether the residual fluxes (ε') are normally distributed. If this is the case, then it can be concluded that the residuals are due to random noise in the observations. If the residuals are not normally distributed, at least one of the observations must contain a significant (gross) error. The sum of squares of standard normally distributed ($\mathcal{Z}$) random variables (x_i) is known to be chi-squared (χ^2) distributed. Expressed mathematically:

$$x_i \sim \mathcal{Z} \tag{C.32}$$

$$x = [x_1, x_2, \cdots, x_{n-1}, x_n]^T \tag{C.33}$$

$$\sum_{i=1}^{n} (x_i^2) \sim \chi^2 \tag{C.34}$$

The sum of squares can also be expressed in linear algebra as:

$$\sum_{i=1}^{n} (x_i^2) = x^T x \tag{C.35}$$

$$x^T x \sim \chi^2 \tag{C.36}$$

From this relationship, the "gross measurement error" test is a χ^2 test on the sum of squares of (row-reduced) residual fluxes, with the ε' terms standard normal variance of 1.

The χ^2 statistic, h, is estimated as follows:

$$h = \varepsilon'^T \mathrm{Cov}(\varepsilon')^{-1} \varepsilon' \tag{C.37}$$

$$\sim \chi^2 \tag{C.38}$$

$\mathrm{Cov}(\varepsilon')$ is the variance-covariance matrix of ε' can be calculated using (C.26):

$$\varepsilon' = R'\delta$$

$$\mathrm{Cov}(\varepsilon') = \mathrm{Cov}(R'\delta) \tag{C.39}$$

$$= R'\mathrm{Cov}(\delta)R'^T \tag{C.40}$$

The transition from Equation (C.39) to Equation (C.40) can be shown from the definition of a variance-covariance matrix:

$$\mathrm{Cov}(X) = \mathrm{E}([X - \mathrm{E}(X)][X - \mathrm{E}(X)]^T) \tag{C.41}$$

$$\mathrm{Cov}(RX) = \mathrm{E}([RX - \mathrm{E}(RX)][RX - \mathrm{E}(RX)]^T) \tag{C.42}$$

Since $\mathrm{E}(RX) = R\mathrm{E}(X)$,

$$\mathrm{Cov}(RX) = \mathrm{E}([RX - R\mathrm{E}(X)][RX - R\mathrm{E}(X)]^T) \tag{C.43}$$

$$= \mathrm{E}([(R)(X - \mathrm{E}(X))][(R)(X - \mathrm{E}(X))]^T) \tag{C.44}$$

Since $(RX)^T = X^T R^T$

$$\mathrm{Cov}(RX) = \mathrm{E}(R[X - \mathrm{E}(X)][X - \mathrm{E}(X)]^T R^T) \tag{C.45}$$

$$= R\mathrm{E}([X - \mathrm{E}(X)][X - \mathrm{E}(X)]^T)R^T \tag{C.46}$$

$$= R\mathrm{Cov}(X)R^T \tag{C.47}$$

$\mathrm{Cov}(\delta)$ is a covariance matrix of observed flux values that can be calculated directly through replication or estimated from prior information. For the example problem, v_o was taken to be $[2.1, 0.9, 1.2]^T$. Following the estimation of a covariance matrix around the observations, the χ^2 would be carried out as follows:

$$\mathrm{Cov}(\delta) = \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.04 \end{bmatrix} \end{array} \tag{C.48}$$

$$\mathrm{Cov}(\varepsilon') = \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.3333 & -0.3333 & -0.3333 \\ 0 & 0.2500 & -0.2500 \end{bmatrix} \end{array}$$

$$\cdot \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.04 \end{bmatrix} \end{array} \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \\ \begin{bmatrix} 0.3333 & 0 \\ -0.3333 & 0.2500 \\ -0.3333 & -0.2500 \end{bmatrix} \end{array} \tag{C.49}$$

$$= \begin{bmatrix} 0.006667 & 0.002500 \\ 0.002500 & 0.003125 \end{bmatrix} \tag{C.50}$$

$$h = \begin{bmatrix} 0 & -0.075 \end{bmatrix} \begin{bmatrix} 0.006667 & 0.002500 \\ 0.002500 & 0.003125 \end{bmatrix} \begin{bmatrix} 0 \\ -0.075 \end{bmatrix} \tag{C.51}$$

$$= 2.5714 \tag{C.52}$$

Since 2.5714 is smaller than the critical value of $\chi^2_{0.05} = 5.991$ for two degrees of freedom, no gross errors can be identified in the observations.

The suggested procedure is to use the χ^2 test to identify whether any significant errors are present among the observations. If no gross errors are identified then all observed values must contain only random noise, which can be "balanced" by finding an estimate for the observed values that minimizes the sum of squared errors for the observations ($\bar{v}_o$) given by:

$$\bar{v}_o = (I - \mathrm{Cov}(\delta) R'^T \mathrm{Cov}(\varepsilon')^{-1} R') \hat{v}_o \tag{C.53}$$

For the example problem, the corrected observations are $\bar{v}_o = [2.0571, 1.0286, 1.0286]^T$.

C.1.3 Generalized least squares calculation

Following the generalized least squares framework, the residual term ε is not assumed to be the result of measurement error alone and is introduced earlier in the formulation.

$$0 = S_c v_c + S_o v_o + \varepsilon \tag{C.54}$$

Since S_o and v_o are known, $S_o v_o$ is a column vector, and Equation (C.54) can be rearranged to look more like linear regression,

$$-S_o v_o = S_c v_c + \varepsilon \tag{C.55}$$

$$y = X\beta + \varepsilon \tag{C.56}$$

If the elements of ε are independently and identically distributed, then $\hat{v}_c$ can be estimated through ordinary regression (note that residuals don't have to be normally distributed for the estimate to be a best linear unbiased estimator).

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T y \tag{C.57}$$

$$\hat{v}_c = -\left(S_c^T S_c\right)^{-1} S_c^T S_o v_o \tag{C.58}$$

Going back to the example model in Box 1,

$$S_o v_o = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{array} \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} \tag{C.59}$$

$$= \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{bmatrix} 2.1 \\ 0.9 \\ 0 \\ 0 \\ -1.2 \end{bmatrix} \tag{C.60}$$

$$\begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix}\begin{bmatrix} 2.1 \\ 0.9 \\ 0 \\ 0 \\ -1.2 \end{bmatrix} = \begin{matrix} \\ A \\ B \\ C \\ D \\ E \end{matrix}\begin{matrix} C1 & C2 & C3 \\ \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \end{matrix} \begin{matrix} C1 \\ C2 \\ C3 \end{matrix}\begin{bmatrix} v_{C1} \\ v_{C2} \\ v_{C3} \end{bmatrix} \tag{C.61}$$

$$\hat{v}_c = \begin{matrix} C1 \\ C2 \\ C3 \end{matrix}\begin{bmatrix} 0.975 \\ 1.050 \\ 1.125 \end{bmatrix} \tag{C.62}$$

With this formulation, ε represents the deviation between observed and calculated fluxes that may be the result of either measurement error or lack of model fit. However, the assumption of ε being independently and identically distributed is unlikely to be true. Balances around large magnitude fluxes are likely to have higher associated variabilities. Mathematically, it is necessary to have $\text{Cov}(\varepsilon) = \sigma^2$ with a variance-covariance matrix $\text{Cov}(\varepsilon) = I\sigma^2$, whereas real balances may have unequal variances and non-zero covariance terms, equivalent to $\text{Cov}(\varepsilon) = V\sigma^2$.

If Equation (C.55) is scaled by matrix A, then the variance-covariance of the scaled error term becomes

$$\text{Cov}(A\varepsilon) = A\text{Cov}(\varepsilon)A^T \tag{C.63}$$

$$= AV\sigma^2A^T \tag{C.64}$$

$$= \sigma^2AVA^T \tag{C.65}$$

where A must be chosen such that $AVA^T = I$ to meet the required conditions of linear

regression. This is true if A is taken to be the inverse of the matrix square root of V, i.e., $A = P^{-1}$ and $PP = V$. Finding the square root of a matrix P, such that $PP = V$ can be performed by matrix diagonalization of V. To prove this, start from the matrix P and assume that that there exists a diagonal matrix D_P and matrix Γ, for which:

$$P = \Gamma D_P \Gamma^{-1} \tag{C.66}$$

$$PP = (\Gamma D_P \Gamma^{-1})(\Gamma D_P \Gamma^{-1}) \tag{C.67}$$

$$= \Gamma D_P D_P \Gamma^{-1} \tag{C.68}$$

Since $PP = V$ and the square of a diagonal matrix is that same matrix with all diagonal entries squared, then $D_P D_P$ can be redefined as diagonal matrix D_V:

$$V = \Gamma D_P D_P \Gamma^{-1} \tag{C.69}$$

$$V = \Gamma D_V \Gamma^{-1} \tag{C.70}$$

P can thus be calculated by diagonalizing V and taking the square root of its diagonal matrix, which is equal to the square root of the individual diagonal values. Since V is symmetric ($\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(X_2, X_1)$),

$$V = \Gamma D_V \Gamma^T \tag{C.71}$$

Therefore, P must also be symmetric:

$$P^T = (\Gamma D_P \Gamma^T)^T \tag{C.72}$$

$$= (\Gamma^T)^T (D_P)^T (\Gamma)^T \tag{C.73}$$

$$= \Gamma D_P \Gamma^T \tag{C.74}$$

$$= P \tag{C.75}$$

Combining $PP = V$ and $P = P^T$:

$$\text{Cov}(P^{-1}\varepsilon) = P^{-1}\text{Cov}(\varepsilon)(P^{-1})^T \tag{C.76}$$

$$= P^{-1} V \sigma^2 (P^T) - 1 \tag{C.77}$$

$$= \sigma^2 P^{-1} P P P^{-1} \tag{C.78}$$

$$= \sigma^2 I \tag{C.79}$$

as required.

Scaling Equation (C.55) by P^{-1}:

$$-P^{-1} S_o v_o = P^{-1} S_c v_c + P^{-1}\varepsilon \tag{C.80}$$

where $P^{-1}\varepsilon$ now satisfies the assumptions of linear regression. Formally, this is equivalent to generalized least squares (GLS) regression, however, incorporating P^{-1} directly into each term allows the use of all ordinary least squares techniques:

$$-S'_o v_o = S'_c v_c + \varepsilon' \tag{C.81}$$

$$\hat{v}_c = -\left(S'^T_c S'_c\right)^{-1} S'^T_c S'_o v_o \tag{C.82}$$

The calculation of P^{-1} requires the estimation $\text{Cov}(\varepsilon)$ from the variance of observed fluxes. Calculating the covariance-variance matrix of both sides of Equation (C.55):

$$\text{Cov}(-S_o v_o) = \text{Cov}(S_c v_c + \varepsilon) \tag{C.83}$$

$$\text{Cov}(\varepsilon) = S_o \text{Cov}(v_o) S_o^T \tag{C.84}$$

Since $\text{Cov}(\varepsilon) = \sigma^2 V$ for any value of σ, σ is set to 1 so that $V = \text{Cov}(\varepsilon)$. In practice, $\text{Cov}(v_o)$ need only capture the relative magnitudes of observed flux variances as $\hat{\sigma}$ is estimated during regression. Balances around molecular species that do not include an observed flux v_o will have a row of zeros in $\text{Cov}(\varepsilon)$, which prevents the calculation of a matrix inverse (required to get P^{-1}). Although this mathematically equates to a variance of zero for those balance, a better interpretation is that there is an unknown variance around the "observation" of no net flux. The simplest solution is to add a small non-zero value to each diagonal entry of $\text{Cov}(\varepsilon)$, representing the confidence of the calculated fluxes being fully balanced. If there is more uncertainty around some balances than others, this information could be encoded in the magnitude of the added variance. P can then be calculated via a matrix square root of $\text{Cov}(\varepsilon)$. Since a variance (covariance) matrix is positive semi-definite, P is known to be unique.

Coming back to the example problem, assume that $\text{Cov}(v_o)$ has been defined as follows:

$$\text{Cov}(v_o) = \sigma^2 \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.04 \end{bmatrix} \end{array} \tag{C.85}$$

Then,

$$\mathrm{Cov}(\varepsilon) = \sigma^2 \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{array}$$

$$\cdot \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \\ 0 & 0 & 0.04 \end{bmatrix} \end{array} \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \end{array} \tag{C.86}$$

$$= \sigma^2 \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.04 \end{bmatrix} \end{array} \tag{C.87}$$

Rows corresponding to balances on C and D do not include any observed fluxes, and are therefore all zero. The variance on the balance can be set to 0.0001 arbitrarily to reflect the relative confidence of the balance being closed:

$$\mathrm{Cov}(\varepsilon) = \sigma^2 V \tag{C.88}$$

$$\mathrm{Cov}(\varepsilon) = \sigma^2 \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 & 0 \\ 0 & 0 & 0.0001 & 0 & 0 \\ 0 & 0 & 0 & 0.0001 & 0 \\ 0 & 0 & 0 & 0 & 0.04 \end{bmatrix} \end{array} \tag{C.89}$$

Since V is a diagonal matrix, P^{-1} can be calculated as the reciprocal square root of each element.

$$P^{-1} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \end{array} \tag{C.90}$$

$$S_o' = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \end{array} \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{array} \tag{C.91}$$

$$= \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -5 \end{bmatrix} \end{array} \tag{C.92}$$

$$S'_c = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \end{array} \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} -1 & 0 & -1 \\ -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \tag{C.93}$$

$$= \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} -10 & 0 & -10 \\ -10 & 0 & 0 \\ 100 & -100 & 0 \\ 0 & 100 & -100 \\ 0 & 0 & 5 \end{bmatrix} \end{array} \tag{C.94}$$

Performing the least squares calculation as before:

$$S'_o v_o = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccc} O1 & O2 & O3 \end{array} \\ \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -5 \end{bmatrix} \end{array} \begin{array}{c} \\ O1 \\ O2 \\ O3 \end{array} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} \tag{C.95}$$

$$= \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{bmatrix} 21 \\ 9 \\ 0 \\ 0 \\ -6 \end{bmatrix} \tag{C.96}$$

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{bmatrix} 21 \\ 9 \\ 0 \\ 0 \\ -6 \end{bmatrix} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} -10 & 0 & -10 \\ -10 & 0 & 0 \\ 100 & -100 & 0 \\ 0 & 100 & -100 \\ 0 & 0 & 5 \end{bmatrix} \end{array} \begin{array}{c} C1 \\ C2 \\ C3 \end{array}\begin{bmatrix} v_{C1} \\ v_{C2} \\ v_{C3} \end{bmatrix} \tag{C.97}$$

$$\hat{v}_c = \begin{array}{c} C1 \\ C2 \\ C3 \end{array}\begin{bmatrix} 1.0278 \\ 1.0287 \\ 1.0295 \end{bmatrix} \tag{C.98}$$

Whereas calculated fluxes $\hat{v}_c$ are commonly estimated using a very similar "weighted" least squares approach, the use of validation methods that are part of the regression framework have yet to be explored. The common χ^2 test can still be used to detect gross measurement errors in estimated residuals ($\hat{\varepsilon}$), however, the validation of a regression model also requires the use of t-tests to ensure the significance of calculated fluxes. Confidence and prediction intervals are also highly relevant to MFA. The calculation of a t-statistic follows from normal regression:

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_i}{\text{se}(\hat{\beta}_i)} \tag{C.99}$$

$$t_{\hat{v}_{c,i}} = \frac{\hat{v}_{c,i}}{\text{se}(\hat{v}_{c,i})} \tag{C.100}$$

$$t_{\hat{v}_{c,i}} = \frac{\left(-\left(S_c'^T S_c'\right)^{-1} S_c'^T S_o' v_o\right)_i}{\hat{\sigma}\sqrt{(S_c'^T S_c')^{-1}_{i,i}}} \tag{C.101}$$

The estimated standard deviation of ε (or $\hat{\sigma}$) is calculated as follows:

$$\hat{\sigma}^2 = \frac{\sum(\hat{\varepsilon}'_i)^2}{n_b - n_c - 1} \tag{C.102}$$

where:

$$\hat{\varepsilon}' = -S'_o v_o + S'_c \left(S'^T_c S'_c\right)^{-1} S'^T_c S'_o v_o \tag{C.103}$$

and n_b is the number of balances (rows of S'_c) while n_c is the number of fluxes to be calculated (columns of S'_c). If the model is correct and $\mathrm{Cov}(\varepsilon)$ was correctly estimated, $\hat{\sigma}^2$ should be approximately equal to 1. Once the t-value is calculated, a flux can be judged statistically significant if $|t_{\hat{v}_{c,i}}| \geq t_{\alpha/2, n_b - n_c - 1}$ where α is the significance level.

As $\hat{v}_c$ has already been calculated for the example model, only $\hat{\sigma}$ and $\sqrt{(S'^T_c S'_c)^{-1}_{i,i}}$ are required to perform a t-test.

$$\hat{\varepsilon}' = \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} \begin{bmatrix} -0.4261 \\ 1.2784 \\ 0.0852 \\ 0.0852 \\ 0.8523 \end{bmatrix} \tag{C.104}$$

$$\hat{\sigma} = 1.60 \tag{C.105}$$

$$\sqrt{(S'^T_c S'_c)^{-1}_{i,i}} = \begin{matrix} C1 \\ C1 \\ C3 \end{matrix} \begin{bmatrix} 0.0441 \\ 0.0442 \\ 0.0444 \end{bmatrix} \tag{C.106}$$

$$t_{0,\hat{v}_{c,i}} = \begin{matrix} C1 \\ C1 \\ C3 \end{matrix} \begin{bmatrix} 14.59 \\ 14.55 \\ 14.51 \end{bmatrix} \tag{C.107}$$

The critical t-value for a two-tailed test with 1 degree of freedom and 5% significance level is 12.71, making all three calculated fluxes statistically significant.

The identification of non-significant flux may be interpreted in two ways. The measurement error around observed fluxes may be too high to allow robust flux calculation. In that case, non-significant fluxes should be treated as having a flux of zero and excluded from the model or further analysis. Alternatively, non-significance may be the result of excess variability from a lack of fit between the model and observed data, requiring model correction. To distinguish between these cases, it is necessary to separate model error from measurement uncertainty. One way to accomplish this is to reduce measurement uncertainty through added replication, however, the required effort can make this approach practically infeasible. Another solution is to simulate a set of feasible fluxes directly from the stoichiometric model (and therefore free of model error) for comparison to the observed data.

The simulation of feasible fluxes can be simplified by eliminating flux equality constraints expressed by the stoichiometry matrix. Essentially, only $n_c - n_b$ fluxes have to specified in order to generate all the other values. More formally, the relationships between the fluxes can be succinctly summarized through the nullspace (or kernel) of S, which describes all flux balance conservations in the model. This makes it possible to calculate all fluxes from a smaller set of variables referred to as the basis. Unlike fluxes,

which must satisfy constraints imposed by $Sv = 0$, the basis can take any arbitrary value to generate fluxes that satisfy all required constraints. Expressed mathematically,

$$\text{Null}(S) = K \tag{C.108}$$

$$Kb = v \tag{C.109}$$

where b is a basis vector of any value with the same number of rows as columns of K. While all values of b satisfy $Sv = 0$, it is still necessary to constrain fluxes to a set of realistic values representative of a cell cultivation. The space of all feasible fluxes v can be constrained by defining upper and lower bounds on each observed flux:

$$\begin{aligned} & v = Kb \\ \text{subject to} \quad & K_i b \leq v_i + a \cdot \text{sd}(v_i) \\ & K_i b \geq v_i - a \cdot \text{sd}(v_i) \end{aligned} \tag{C.110}$$

where v_i is an observed flux, K_i is the corresponding row of K, and a is a scaling constant that can be set to $t_{\alpha/2,df}$ to specify a confidence interval around v_i. As the basis solution space is only constrained by inequalities, it is readily amenable to stochastic sampling. All values of v that satisfy Equation (C.110) represent feasible fluxes that would perfectly satisfy the stoichiometric model while remaining within measurement uncertainty of real observations. If the resulting space is infeasible, then the observed data does not fit the specified model. Otherwise, a random sample of feasible fluxes can be taken for comparison to observed results. If the addition of measurement error to simulated fluxes results in less uncertainty than from observed results, then model error is to blame.

The example model can be solved by specifying only 1 flux value, meaning that K is a column vector.

$$K = \begin{matrix} C1 \\ C1 \\ C3 \\ O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} \tag{C.111}$$

Based on K, a flux profile is consistent with the model as long as the flux of compound A (v_{O1}) is two times greater than all other fluxes. From the estimated variance-covariance matrix, the standard deviations of the observed fluxes are $[0.1, 0.1, 0.2]^T$. Assuming that these estimates came from a large number of samples, approximate 95% confidence intervals can be constructed by multiplying the standard deviation values by 2.

$$v = \begin{matrix} C1 \\ C1 \\ C3 \\ O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} b \tag{C.112}$$

$$\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} b \leq \begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} + 2 \left(\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 0.1 \\ 0.1 \\ 0.2 \end{bmatrix} \right) \tag{C.113}$$

$$\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} b \leq \begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2.3 \\ 1.1 \\ 1.6 \end{bmatrix} \tag{C.114}$$

$$\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} b \geq \begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2.1 \\ 0.9 \\ 1.2 \end{bmatrix} - 2 \left(\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 0.1 \\ 0.1 \\ 0.2 \end{bmatrix} \right) \tag{C.115}$$

$$\begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} b \geq \begin{matrix} O1 \\ O2 \\ O3 \end{matrix} \begin{bmatrix} 1.9 \\ 0.7 \\ 0.8 \end{bmatrix} \tag{C.116}$$

With only a single basis value b, the constraints can be combined to limit b between 0.95 and 1.1. Random values can then be generated from a uniform distribution. For basis vectors with more than one value, independent limits on values of b can not be isolated and the solution space must be sampled directly. Once basis vectors are generated, flux profiles can be calculated from $Kb = v$ and subject to measurement error.

References

Stephanopoulos G., Aristidou A., and Nielsen J. (1998).
Metabolic engineering: principles and methodologies. Academic Press.

C.2 Model definition

The model used in this work was largely taken from Altamirano et al. 2001 with only minimal changes (listed in the Materials and Methods section of the manuscript). A full list of reactions is presented below with a rough outline of matetrial flow in figure C.1.

C.2.1 Glycolysis and PPP

HK	$\mathrm{Glc} + \mathrm{ATP} \longrightarrow \mathrm{ADP} + \mathrm{G6P}$
FBA	$\mathrm{G6P} + \mathrm{ATP} \longrightarrow \mathrm{ADP} + 2\,\mathrm{DHAP}$
PK	$\mathrm{DHAP} + 2\,\mathrm{ADP} + \mathrm{NAD} + \mathrm{P} \longrightarrow 2\,\mathrm{ATP} + \mathrm{NADH} + \mathrm{Pyr} + \mathrm{H_2O}$
PPP	$\mathrm{G6P} + 2\,\mathrm{NADP} + \mathrm{H_2O} \longrightarrow 2\,\mathrm{NADPH} + \mathrm{R5P} + \mathrm{CO_2}$

HK hexokinase

FBA fructose-biphosphate aldolase

PK pyruvate kinase

PPP pentose phosphate pathway

C.2.2 Pyr and AcCoA

LDH	$\mathrm{Pyr} + \mathrm{NADH} \longrightarrow \mathrm{Lac} + \mathrm{NAD}$
ACAS	$\mathrm{mAcCoA} + \longrightarrow \mathrm{mCoA} + \mathrm{AcOH}$
ME	$\mathrm{Mal} + \mathrm{NADP} \longrightarrow \mathrm{Pyr} + \mathrm{CO_2} + \mathrm{NADPH}$
PDH	$\mathrm{mPyr} + \mathrm{mCoA} + \mathrm{NAD} \longrightarrow \mathrm{mAcCoA} + \mathrm{NADH} + \mathrm{CO_2}$

LDH lactate dehydrogenase

ACAS acetyl-coa synthetase

ME NADP-malic enzyme

PDH pyruvate dehydrogenase

C.2.3 TCA cycle

CS	$mAcCoA + mOxal + H_2O \longrightarrow mCit + mCoA$
	$mCit + NAD \longrightarrow mAKG + NADH + CO_2$
AKGDH	$mAKG + mCoA + NAD \longrightarrow mSucCoA + NADH + CO_2$
SCS	$ADP + mSucCoA + P \longrightarrow ATP + mCoA + mSuc$
	$mFAD + mSuc + H_2O \longrightarrow mFADH + mMal$
mMalDH	$mMal + NAD \longrightarrow mOxal + NADH$
cMalDH	$NADH + Oxal \longrightarrow Mal + NAD$

CS citrate synthase

AKGDH alpha-ketoglutarate dehydrogenase

SCS succinyl-coa synthetase

mMalDH mitochondrial malate dehydrogenase

cMalDH cytosolic malate dehydrogenase

C.2.4 Glutaminolysis

ALT	$Glu + Pyr \longrightarrow AKG + Ala$
GLDH	$mGlu + NAD + H_2O \longrightarrow Amm + mAKG + NADH$
GS	$Glu + ATP + Amm \longrightarrow Gln + ADP + P$
AST	$Glu + Oxal \longrightarrow Asp + AKG$
	$Glu \longrightarrow Pgl$

ALT alanine transaminase

GLDH glutamate dehydrogenase

GS glutamine synthetase

AST aspartate transaminase

C.2.5 Amino acid degradation

SDH	$\mathrm{Ser} \longrightarrow \mathrm{Pyr} + \mathrm{Amm}$
SHMT	$\mathrm{Ser} + \mathrm{FH_4} \longrightarrow \mathrm{Gly} + \mathrm{N_5N_{10}methyleneFH_4} + \mathrm{H_2O}$
GCS	$\mathrm{CO_2} + \mathrm{Amm} + \mathrm{N_5N_{10}methyleneFH_4} + \mathrm{NADH} \longrightarrow \mathrm{Gly} + \mathrm{FH_4} + \mathrm{NAD}$
	$\mathrm{Thr} + \mathrm{mCoA} + \mathrm{NAD} \longrightarrow \mathrm{Gly} + \mathrm{mAcCoA} + \mathrm{NADH}$
	$\mathrm{Cys} + \mathrm{AKG} \longrightarrow \mathrm{Glu} + \mathrm{Pyr}$
	$\mathrm{His} + \mathrm{FH_4} + \mathrm{NADPH} + 2\,\mathrm{H_2O} \longrightarrow$ $\mathrm{Glu} + \mathrm{N_5N_{10}methyleneFH_4} + \mathrm{NADP} + 2\,\mathrm{Amm}$
	$\mathrm{Arg} + \mathrm{AKG} + \mathrm{NAD} + 2\,\mathrm{H_2O} \longrightarrow 2\,\mathrm{Glu} + \mathrm{NADH}$
	$\mathrm{Pro} + 2\,\mathrm{NAD} + 2\,\mathrm{H_2O} \longrightarrow \mathrm{Glu} + 2\,\mathrm{NADH}$
ASNS	$\mathrm{Asn} + \mathrm{H_2O} \longrightarrow \mathrm{Amm} + \mathrm{Asp}$
	$\mathrm{Ile} + \mathrm{AKG} + \mathrm{ATP} + 2\,\mathrm{mCoA} + \mathrm{mFAD} + 2\,\mathrm{NAD} + \mathrm{H_2O} \longrightarrow$ $\mathrm{mSucCoA} + \mathrm{Glu} + \mathrm{mAcCoA} + \mathrm{ADP} + \mathrm{P} + \mathrm{mFADH} + 2\,\mathrm{NADH}$
	$\mathrm{Leu} + \mathrm{ATP} + \mathrm{AKG} + \mathrm{NAD} + \mathrm{mCoA} + \mathrm{mFAD} + \mathrm{H_2O} + \mathrm{mSucCoA} \longrightarrow$ $2\,\mathrm{mAcCoA} + \mathrm{mSuc} + \mathrm{Glu} + \mathrm{NADH} + \mathrm{mFADH} + \mathrm{ADP} + \mathrm{P}$
	$\mathrm{Lys} + \mathrm{AKG} + 2\,\mathrm{mCoA} + 0.5\,\mathrm{O_2} + 2\,\mathrm{H_2O} + 2\,\mathrm{mFAD} + 2\,\mathrm{NAD} \longrightarrow$ $\mathrm{Glu} + 2\,\mathrm{mAcCoA} + 2\,\mathrm{CO_2} + \mathrm{Amm} + 2\,\mathrm{mFADH} + 2\,\mathrm{NADH}$
	$\mathrm{Met} + \mathrm{Ser} + 3\,\mathrm{ATP} + \mathrm{FH_4} + 2\,\mathrm{NAD} + \mathrm{ATP} + \mathrm{mCoA} + 4\,\mathrm{H_2O} \longrightarrow$ $\mathrm{mSucCoA} + \mathrm{Cys} + \mathrm{Amm} + 3\,\mathrm{ADP} + 3\,\mathrm{P} + \mathrm{N_5N_{10}methyleneFH_4} + 2\,\mathrm{NADH} + \mathrm{ADP}$
PAH	$\mathrm{Phe} + \mathrm{NADPH} + \mathrm{O_2} \longrightarrow \mathrm{Tyr} + \mathrm{NADP} + \mathrm{H_2O}$
	$\mathrm{Trp} + 2\,\mathrm{mCoA} + 3\,\mathrm{O_2} + 4\,\mathrm{H_2O} + \mathrm{NADPH} + 3\,\mathrm{NAD} + \mathrm{mFAD} \longrightarrow$ $2\,\mathrm{mAcCoA} + 4\,\mathrm{CO_2} + \mathrm{Ala} + \mathrm{NADP} + 3\,\mathrm{NADH} + \mathrm{mFADH} + \mathrm{Amm}$
	$\mathrm{Tyr} + \mathrm{AKG} + \mathrm{mCoA} + \mathrm{mSucCoA} + 2\,\mathrm{O_2} + 2\,\mathrm{H_2O} \longrightarrow$ $\mathrm{Glu} + \mathrm{Mal} + 2\,\mathrm{mAcCoA} + \mathrm{mSuc} + \mathrm{CO_2}$
	$\mathrm{Val} + \mathrm{AKG} + 2\,\mathrm{ATP} + \mathrm{mCoA} + \mathrm{mFAD} + 3\,\mathrm{NAD} + 4\,\mathrm{H_2O} \longrightarrow$ $\mathrm{mSucCoA} + \mathrm{Glu} + 2\,\mathrm{ADP} + 2\,\mathrm{P} + \mathrm{mFADH} + 3\,\mathrm{NADH} + \mathrm{CO_2}$

SDH serine dehydratase

SHMT serine hydroxymethyltransferase

GCS glycine cleavage system

ASNS asparagine synthetase

PAH phenylalanine hydroxylase

C.2.6 Macromolecules

Carb	$\mathrm{G6P} + 3.5\,\mathrm{ATP} \longrightarrow \mathrm{Carb} + \mathrm{ADP}$
OA	$9\,\mathrm{mCit} + 9\,\mathrm{Mal} + 17\,\mathrm{ATP} + 17\,\mathrm{NADPH} + 9\,\mathrm{NADH} + \mathrm{O_2} \longrightarrow$ $\mathrm{OA} + 9\,\mathrm{mMal} + 9\,\mathrm{Oxal} + 17\,\mathrm{ADP} + 17\,\mathrm{P} + 7\,\mathrm{NADP} + 9\,\mathrm{NAD} + \mathrm{H_2O}$
DNA	$1.9\,\mathrm{Gln} + 1.3\,\mathrm{Asp} + 7.5\,\mathrm{ATP} + 0.5\,\mathrm{Gly} + 1.3\,\mathrm{N_5N_{10}methyleneFH_4} +$ $0.7\,\mathrm{NAD} + 0.3\,\mathrm{NADPH} + \mathrm{R5P} + 0.3\,\mathrm{NADH} + 0.5\,\mathrm{CO_2} + 3.1\,\mathrm{H_2O} \longrightarrow$ $1.9\,\mathrm{Glu} + 0.8\,\mathrm{Mal} + 7.5\,\mathrm{ADP} + 7.5\,\mathrm{P} + 1.3\,\mathrm{FH_4} +$ $0.7\,\mathrm{NADH} + 0.3\,\mathrm{NADP} + 0.3\,\mathrm{NAD} + \mathrm{DNA}$
RNA	$2.091\,\mathrm{Gln} + 1.194\,\mathrm{Asp} + 7.487\,\mathrm{ATP} + 0.489\,\mathrm{Gly} + 0.978\,\mathrm{N_5N_{10}methyleneFH_4} +$ $0.806\,\mathrm{NAD} + 0.978\,\mathrm{NADP} + \mathrm{R5P} + 0.194\,\mathrm{NADH} +$ $0.489\,\mathrm{CO_2} + 4.59\,\mathrm{H_2O} + 0.097\,\mathrm{O_2} \longrightarrow$ $2.091\,\mathrm{Glu} + 0.683\,\mathrm{Mal} + 7487\,\mathrm{ADP} + 7.487\,\mathrm{P} + 0.978\,\mathrm{FH_4} +$ $0.806\,\mathrm{NADH} + 0.978\,\mathrm{NADPH} + 0.194\,\mathrm{NAD} + \mathrm{RNA}$
Prot	$0.095\,\mathrm{Ala} + 0.048\,\mathrm{Asp} + 0.039\,\mathrm{Asn} + 0.063\,\mathrm{Arg} + 0.028\,\mathrm{Cys} +$ $0.052\,\mathrm{Gln} + 0.064\,\mathrm{Glu} + 0.078\,\mathrm{Gly} + 0.022\,\mathrm{His} + 0.052\,\mathrm{Ile} +$ $0.088\,\mathrm{Leu} + 0.089\,\mathrm{Lys} + 0.02\,\mathrm{Met} + 0.021\,\mathrm{Phe} + 0.028\,\mathrm{Pro} +$ $0.057\,\mathrm{Ser} + 0.061\,\mathrm{Thr} + 0.006\,\mathrm{Trp} + 0.02\,\mathrm{Tyr} + 0.059\,\mathrm{Val} + 4\,\mathrm{ATP} + 4\,\mathrm{H_2O} \longrightarrow$ $\mathrm{Prot} + 4\,\mathrm{ADP} + 4\,\mathrm{P}$

OA oleic acid (lipid) synthesis

C.2.7 Misc.

	$\mathrm{mMal} + \mathrm{AKG} \longrightarrow \mathrm{mAKG} + \mathrm{Mal}$
	$\mathrm{mMal} + \mathrm{Cit} \longrightarrow \mathrm{mCit} + \mathrm{Mal}$
GLAST	$\mathrm{Glu} \longrightarrow \mathrm{mGlu}$
MPC	$\mathrm{Pyr} \longrightarrow \mathrm{mPyr}$
	$\mathrm{FH_4} + \mathrm{FoOH} + \mathrm{ATP} + \mathrm{NADH} \longrightarrow \mathrm{ADP} + \mathrm{P} + \mathrm{NAD} + \mathrm{N_5N_{10}methyleneFH_4}$

GLAST glutamate aspartate transporter

MPC mitochondrial pyruvate carrier

C.2.8 Phosphorylation

$3\,\mathrm{ADP} + \mathrm{NADH} + 0.5\,\mathrm{O_2} + 3\,\mathrm{P} \longrightarrow 3\,\mathrm{ATP} + \mathrm{NAD} + 4\,\mathrm{H_2O}$
$2\,\mathrm{ADP} + \mathrm{mFADH} + 0.5\,\mathrm{O_2} + 2\,\mathrm{P} \longrightarrow 2\,\mathrm{ATP} + \mathrm{mFAD} + 2\,\mathrm{H_2O}$

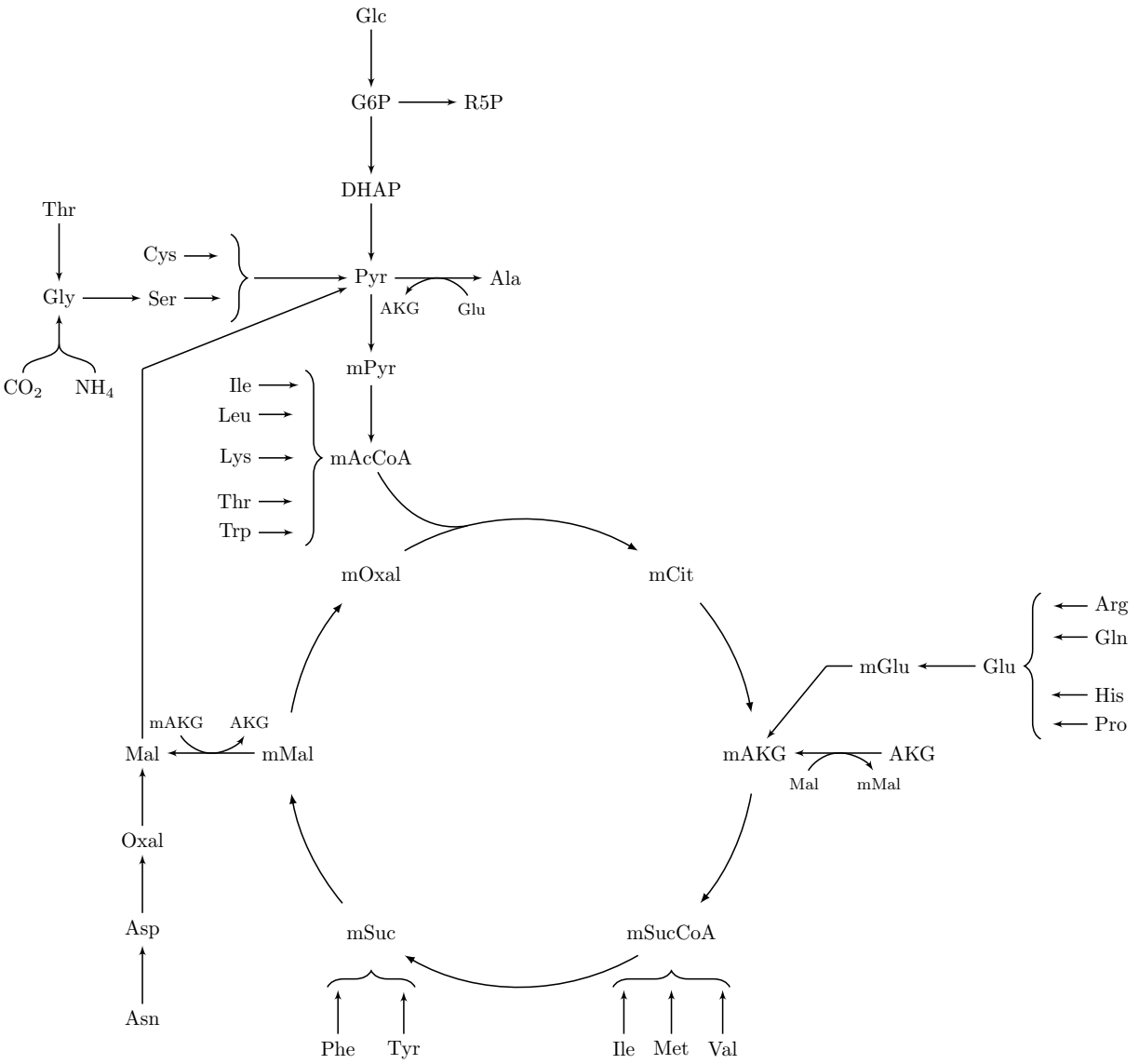


Figure C.1: Outline of metabolic network used in this work. A number of intermediate compounds have been omitted to clarify overall material flow.

References

Altamirano C., Illanes A., Casablancas A., Gámez X., Cairó J. J., and Gòdia C. (2001). "Analysis of CHO cells metabolic redistribution in a glutamate-based defined medium in continuous culture." *Biotechnology Progress* 17.6, pp. 1032–41. DOI: 10.1021/bp0100981.