Spring 5-12-2016

A correction method for systematic error in metabolomic time-course data

Stanislav Sokolenko
University of Waterloo, ssokolen@uwaterloo.ca

Marc Aucoin
University of Waterloo

A CORRECTION METHOD FOR SYSTEMATIC ERROR IN METABOLOMIC TIME-COURSE DATA

Stanislav Sokolenko, University of Waterloo
ssokolen@uwaterloo.ca
Marc G Aucoin, University of Waterloo

Key Words: Metabolomics, time-course, error, quantification, NMR

The growing ubiquity of metabolomic techniques has facilitated high frequency time-course data collection for many cell culture applications. Although the increasing resolution of metabolic profiles has potential to reveal important details about cell culture metabolism, more detailed results are subject to greater influence from measurement and data processing error. A number of common errors, stemming from metabolite extraction and internal standard addition, take the form of a dilution effect, where all observed concentrations feature a constant deviation relative to the true values. We have developed a simple technique to deal with such errors. A nonparametric smoothing fit was applied to all metabolite concentrations, with percent deviations from the fit calculated for each observation. Taking the median of these percent deviations for each sample (across multiple compounds) allowed the estimation of a systematic bias in the relative concentration of all compounds – typical of a dilution error. To validate this method, we developed a general framework for simulating metabolomic experiments. The correction was applied to simulated data sets composed of 20-60 metabolites and 10-20 timepoints. Deviations as small as 2.5% were successfully identified, although greater accuracy was achieved when more data was available. Given the pronounced influence of a small concentration bias on metabolic flux calculation, we were also interested in the effect of similar measurement errors on Metabolic Flux Analysis (MFA). To this end, a Chinese Hamster Ovary (CHO) cell model was used to simulate a set of realistic flux profiles, which were then perturbed with measurement error. Despite the considerable impact of measurement error on flux estimation, the standard χ^2-test was not able to identify erroneous data beyond the significance level. Our findings reinforce the need for validation at earlier stages of analysis in the development of rational strategies for metabolic engineering and media supplementation.

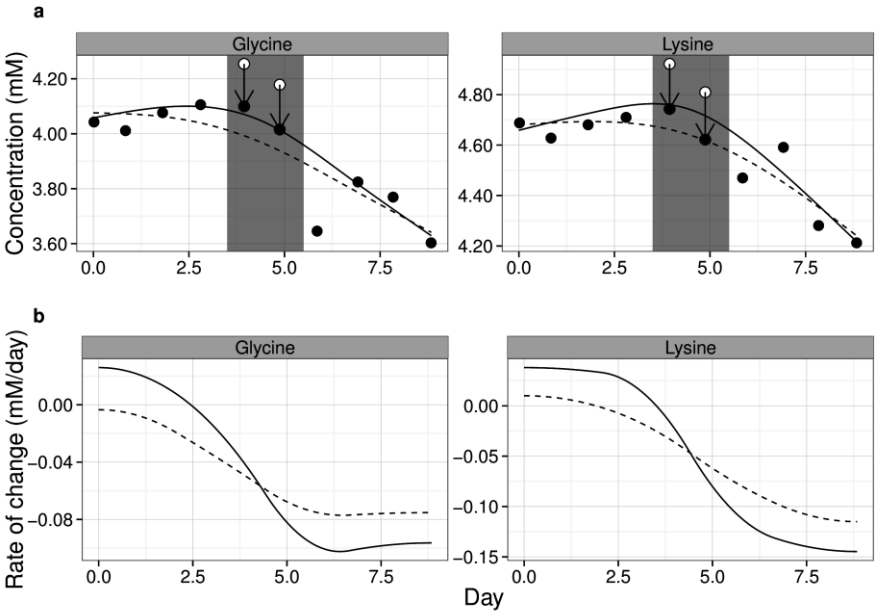

Figure 1 – Correction algorithm applied to example cell culture data. a) White points represent initially observed concentrations that were marked for correction by the algorithm, while black points represent final compound concentrations (with arrows signifying correction). Smoothing lines were generated with a cubic regression spline model using uncorrected (solid line) and corrected (dashed line) data. b) Derivatives calculated from the uncorrected (solid line) and corrected (dashed line) smoothed fits