Soft Sensing of Intracellular States in Bioprocessing with Ensemble Kalman Filters

Luxi Yu [a] , Ehecatl Antonio del Río Chanona [a] , Cleo Kontoravdi [a]

a Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, United Kingdom.

Email: luxi.yu15@imperial.ac.uk; a.del-rio-chanona@imperial.ac.uk; cleo.kontoravdi98@imperial.ac.uk

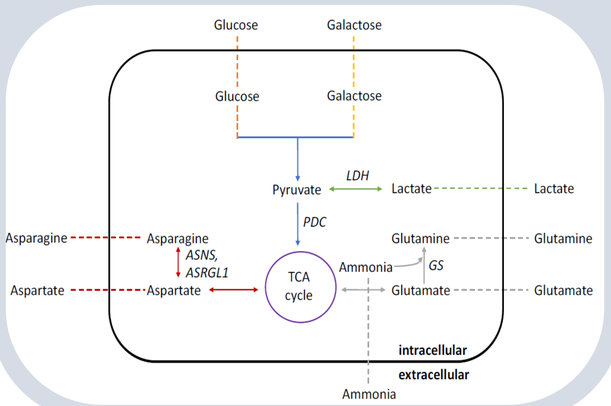
Motivation & Goal

- Glycosylation represents one of the most important quality attributes for a growing number of biotherapeutics such as monoclonal antibodies.
- Nucleotide sugar donors (NSD) are derived from monosaccharides (glucose, galactose, etc.) metabolism and are direct co-substrates for the glycosylation process.
- NSD dynamics provide **tremendous insights** on intracellular states for a variety of purposes, including **glycan profile prediction**[1]. It is highly desirable to predict NSD abundance.
- However, intracellular measurements are often challenging to obtain in the lab and are not routinely taken in industry. In comparison, extracellular metabolite measurements are usually easily accessible.
- The objective of this work is to employ an **Ensemble Kalman Filter (EnKF)**, utilizing a series of **measurements** over time and **the dynamic process model**, which connects the metabolites (measured states correction) with NSDs (unmeasured states estimation).

CHO Cell Metabolism Model

$$\frac{d[met]}{dt} = F_{in}[met]_{feed} - F_{out}[met]_{feed} + Q_{cons}$$

$$q_{mAb} = a * mu + b$$



specific growth rate

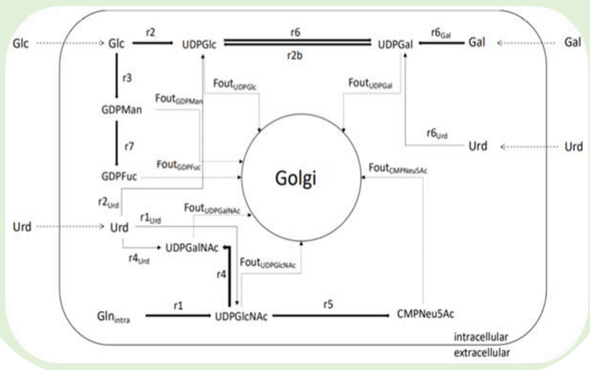
sp. production rate

NSD Synthesis Model

$$r_{prod/cons} = f_{M.M.}(substrate)$$

$$r_{NSD} = r_{prod} - r_{cons}$$

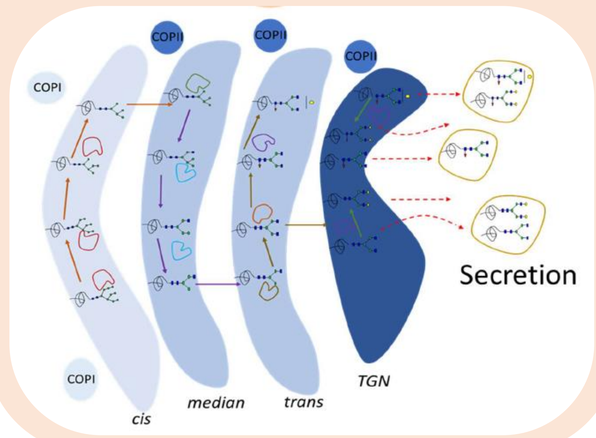
$$\frac{d[NSD]}{dt} = r_{NSD} - Fgolgi_{NSD}$$



NSD

Glycan Distribution (%)

Glycan prediction from NSD with Artificial Neural Network [1] or mechanistic model[2]



Existing Approach

- Whilst some of the existing mechanistic models linking the metabolites with NSDs are highly accurate[2, 3, 4], they are usually specifically designed for the system and require extensive reparameterization for a slight change in experimental condition.

Methodology

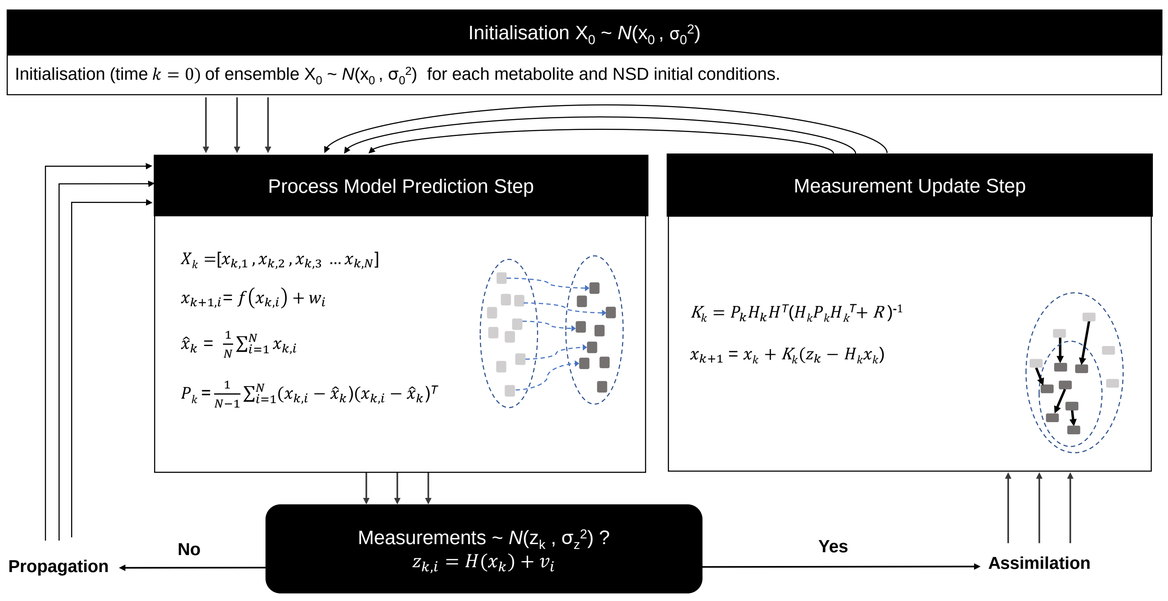
Experimental Set-Up

- Two feeding strategies were carried out with a CHO cell line producing an IgG antibody[2].
- The control experiment includes feeding of glucose and amnio acid nutrients on even days of the cell culture period, whereas the 10G5U experiment includes additional 10mM galactose and 5mM uridine feeding on both Day 4 and Day 8.

Simulation Set-Up

- True values are simulated by adding process variance to a previous mechanistic model[2], measurements are generated by adding the measurement variance obtained from experiments to the true values.
- In real-life scenarios, true value is unknown. The simulation of true value serves as a measure to check the effectiveness of EnKF.

Ensemble Kalman Filter Algorithm



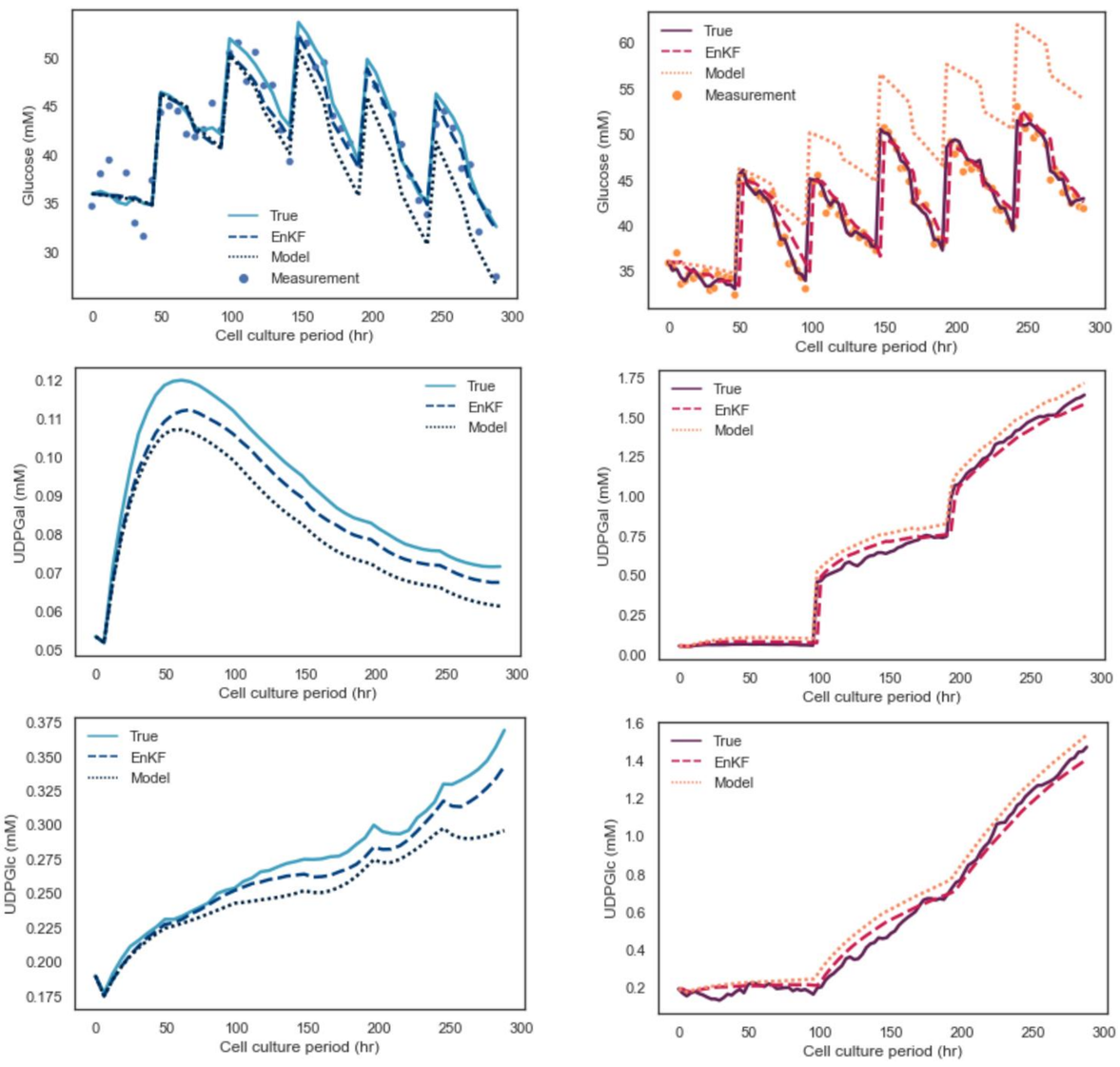
- Ensemble Kalman Filter is a Monte-Carlo implementation of Kalman filtering. EnKF stores, propagates, and updates the ensemble of vectors that approximates the state distribution.
- EnKF uses a collection of state vectors to represent the distribution of the system state and replace the covariance matrix by the sample covariance computed from the ensemble.

Notations: *X*, sampled ensemble; *f*, process model; *w*, Gaussian white noise of process; *x*, process states; *P*, state covariance; *K,* Kalman gain; *H*, measurement function; *R*, measurement noise covariance; z, measurement; *v*, Gaussian white noise of measurement

Results

- The metabolites are corrected by the experimental measurements and are closer to the true value than either the model or the measurement alone.
- NSDs are successfully estimated for both the control and feeding experiments, with filtered results converge to the true value as time evolves.
- The EnKF is able to capture the feeding event and reflected on UDP-Gal (NSD).

CONTROL, NO ADDITIONAL FEEDING | **10mM GAL & 5mM URD FED ON DAY 4 and 8**



Conclusion & Outlook

- EnKF unfolds the possibility of acquiring accurate intracellular NSD information in a more general approach, reducing extensive laboratory work or re-parametrization of the model.
- The versatility of EnKF could be further exploited with more accessible NSD information, bridging the gap between the extracellular metabolites and the antibody quality attributes, and setting the groundwork for in silico glycan optimization of recombinant proteins.

References

1. Jedrzejewski, P.M. *et al. Int J Mol Sci* **15**, 4492-4522 (2014).
2. Kotidis, P. *et al. Biotechnol Bioeng* **116**, 1612-1626 (2019).
3. Sou, S.N. *et al. Biotechnol Bioeng* **114**, 1570-1582 (2017).
4. Jimenez del Val, I., Nagy, J.M. & Kontoravdi, C. *Biotechnol Prog* **27**, 1730-1743 (2011).