Data-driven digital twin models for forecasting multi-step ahead profiles of mammalian cell culture performance

Dong-Yup Lee[1], Seo-Young Park[1], Sun-Jong Kim[1], Cheol-Hwan Park[1], and Jiyong Kim[1]

[1]Sungkyunkwan University School of Chemical Engineering

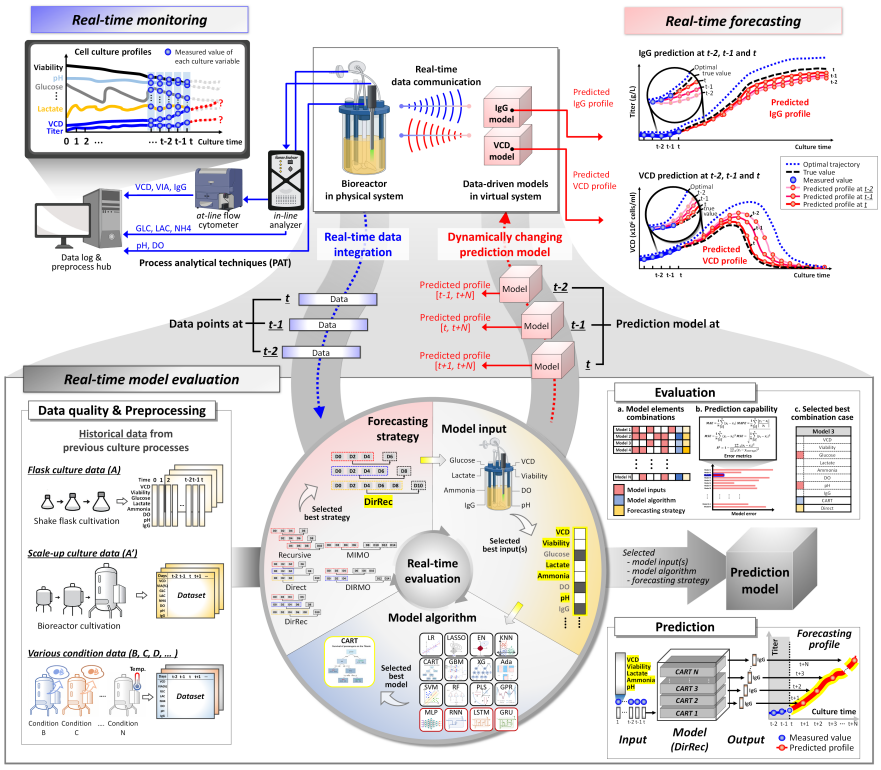
December 2, 2022

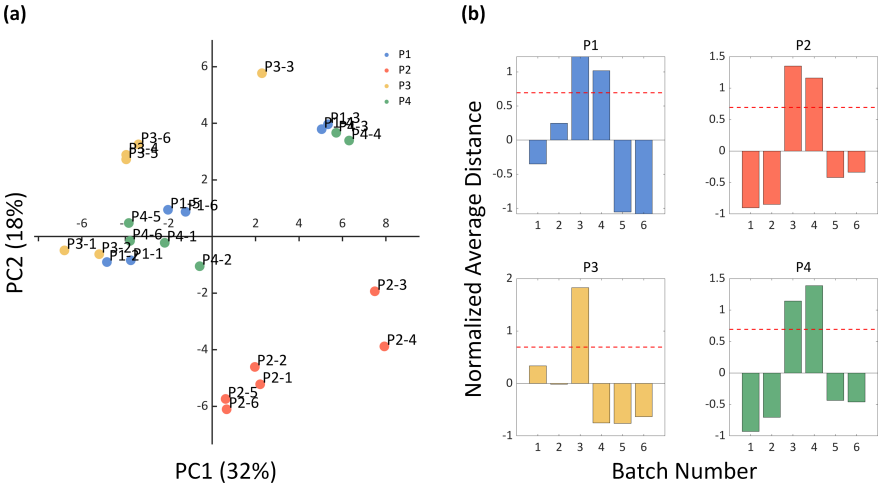
Abstract

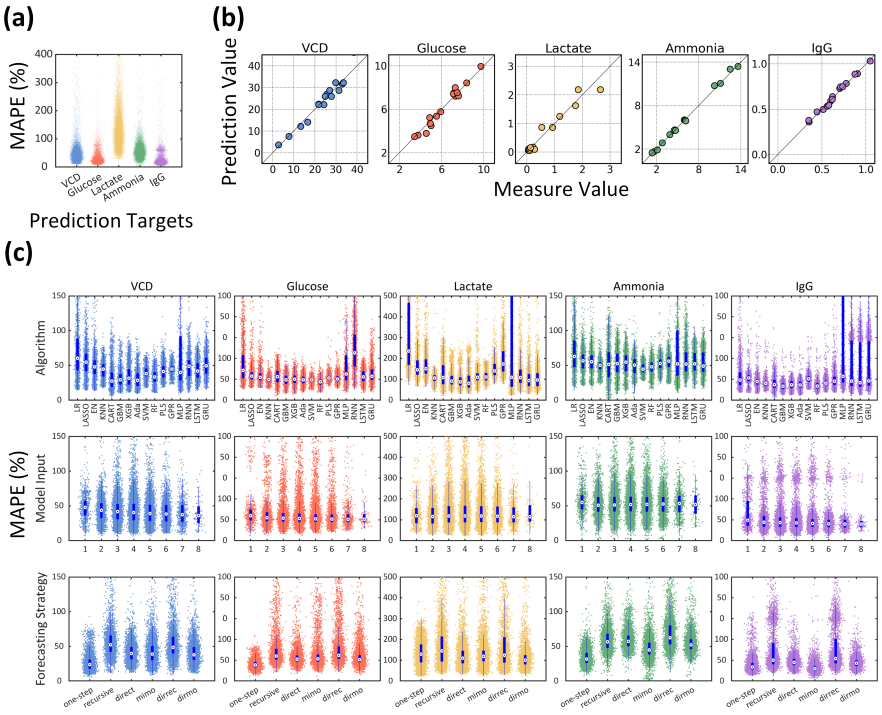
Recently, enormous culture profiles and datasets from biomanufacturing processes to produce recombinant therapeutic proteins (RTP) such as monoclonal antibodies (mAbs) could be generated by virtue of the advancement in process analytical techniques and artificial intelligence (AI). Thus, now it is highly necessary to develop AI-based data-driven models (DDMs) and exploit them accordingly in order to further enhance operational efficiency and accelerate reliable product supply. Since bioprocess is a complex and dynamic system, DDMs are practical and particularly useful to describe the intrinsic relationship among biological and process parameters and cell culture conditions by capturing inherent patterns and to produce high-quality RTP under consistent operations as well as to decrease cost and time by predicting incipient or abrupt faults during the cell cultures. In this work, we provide the practical guideline for choosing the best DDM on given mAb-producing Chinese hamster ovary (CHO) cell culture data sets, enabling us to forecast culture performance such as VCD, and mAb titer as well as glucose, lactate and ammonia concentrations in real time manner. Via the case study with 32 fed-batch data sets of CHO cell cultures, we suggested best combination of model elements including AI algorithms and multi-step ahead forecasting strategies, for good prediction in terms of the computational load as well as the model accuracy and reliability, which is applicable to implementation of interactive data-driven model within bioprocess digital twins. We believe this systematic study can help bioprocess engineers to start developing predictive DDMs with their own data and learn how their cell cultures behave in near future, thereby making proactive decision possible.

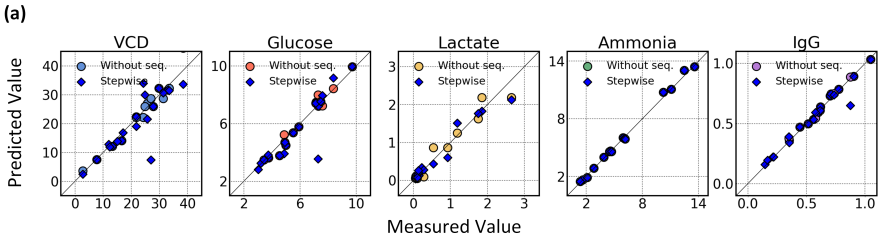
Hosted file

`Manuscript_BnB_Bioprocess AI.docx` available at https://authorea.com/users/399277/articles/609362-data-driven-digital-twin-models-for-forecasting-multi-step-ahead-profiles-of-mammalian-cell-culture-performance

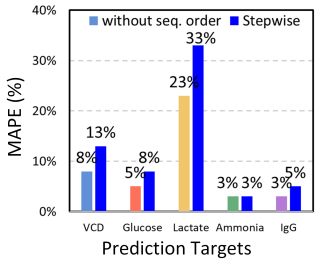








(b)

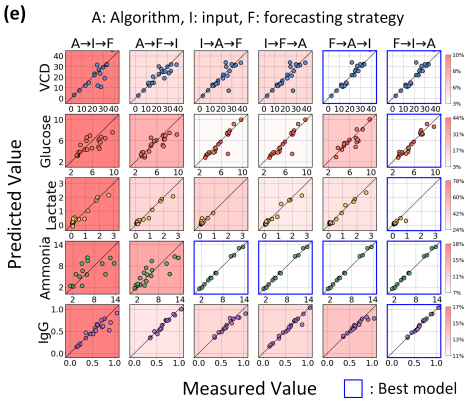


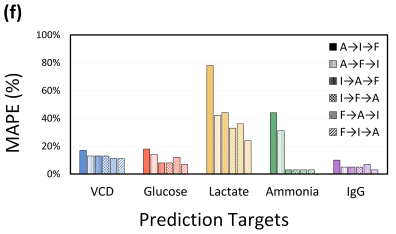
(c)

Elapsed time (min)	Prediction model				
	VCD	Glucose	Lactate	Ammonia	IgG
Without Seq.	751.6	832.5	760.0	754.6	749.4
Stepwise (I > F > A)	36.4	35.1	39.7	35.6	43.9

(d)

	VCD	Glucose	Lactate	Ammonia	IgG
Rank	43 / 20400	7 / 20400	21 / 20400	13 / 20400	45 / 20400





(g)

	AIF	AFI	IAF	IFA	FAI	FIA
VCD	333	32	43	43	10	10
Glucose	3651	563	7	7	195	5
Lactate	3724	153	231	21	43	2
Ammonia	4144	634	13	13	13	13
IgG	1609	18	49	45	119	6