IISE Transactions

Taylor & Francis
Taylor & Francis Group

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uiie21

A calibration-free method for biosensing in cell manufacturing

Jialei Chen , Zhaonan Liu , Kan Wang , Chen Jiang , Chuck Zhang & Ben Wang

To cite this article: Jialei Chen , Zhaonan Liu , Kan Wang , Chen Jiang , Chuck Zhang & Ben Wang (2020): A calibration-free method for biosensing in cell manufacturing, IISE Transactions, DOI: 10.1080/24725854.2020.1856982

To link to this article: https://doi.org/10.1080/24725854.2020.1856982

IISE TRANSACTIONS
https://doi.org/10.1080/24725854.2020.1856982

A calibration-free method for biosensing in cell manufacturing

Jialei Chen[a,b], Zhaonan Liu[b,c], Kan Wang[b], Chen Jiang[b,c], Chuck Zhang[a,b], and Ben Wang[a,b,c]

[a]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA; [b]Georgia Tech Manufacturing Institute, Georgia Institute of Technology, Atlanta, GA, USA; [c]School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

Chimeric antigen receptor T-cell therapy has demonstrated innovative therapeutic effectiveness in fighting cancers; however, it is extremely expensive, due to the intrinsic patient-to-patient variability in cell manufacturing. We propose in this work a novel calibration-free statistical framework to effectively deduce critical quality attributes under the patient-to-patient variability. Specifically, we model this variability via a patient-specific calibration parameter, and use readings from multiple biosensors to construct a patient-invariance statistic, thereby alleviating the effect of the calibration parameter. A carefully formulated optimization problem and an algorithmic framework are presented to find the best patient-invariance statistic and the model parameters. Using the patient-invariance statistic, we can deduce the critical quality attribute of interest, free from the calibration parameter. We demonstrate improvements of the proposed calibration-free method in different simulation experiments. In the cell manufacturing case study, our method not only effectively deduces viable cell concentration for monitoring, but also reveals insights for the cell manufacturing process.

ARTICLE HISTORY
Received 27 July 2020
Accepted 10 November 2020

KEYWORDS
CAR T-cell therapy; cell culture; invariance; patient-to-patient variability

1. Introduction

Cell therapy is one of the most promising treatment approaches to have emerged over the last decades, demonstrating great potential in treating cancers, including leukemia and lymphoma (Kim and de Vellis, 2009; Yin, 2017). Among those therapies, Chimeric Antigen Receptor (CAR) T-cell therapy (Bonifant *et al.*, 2016; June *et al.*, 2018), involving the reprogramming of a patient's T cells to effectively target and attack tumor cells, has shown innovative therapeutic effects in clinical trials, leading to a recent approval (i.e., the treatment of CD19+ hematological malignancies, see Prasad (2018)) by the FDA as a new cancer treatment modality. As illustrated in Figure 1, a typical CAR T-cell therapy involves four steps: deriving cells from a patient, genetically modifying the cells, culturing the cells, and re-administering back to the patient. With increasingly mature gene modification technology, more and more researchers focus on the culturing step (i.e., the red box in Figure 1), where the goal is to substantially increase the cell amount from a small batch to one dose for delivery to the patient. However, a key challenge is the intrinsic patient-to-patient variability in the starting material, i.e., cells derived from different patients vary in their viabilities, acceptance rates of genetic modification, and reactions to culture media (Hinrichs and Restifo, 2013). These variabilities introduce difficulties in cell culturing scale-up (i.e., cell manufacturing), and therefore, the current CAR T-cell therapy is hindered by low scalability, labor-intensive processes, and extremely high cost (Harrison *et al.*, 2019). To achieve high quality and acceptable vein-to-vein cost, we present in this work a statistical framework for online monitoring in cell manufacturing, which can alleviate the negative effect of the intrinsic patient-to-patient variability.

There are two reasons why a new statistical method is needed for monitoring critical quality attributes (exampled by the cell concentration) in cell manufacturing. First, a *direct* measurement method for cell concentrations is not suitable in cell manufacturing (Slouka *et al.*, 2016). Such a method typically requires experienced technicians to collect culture media, take microscopic images, and perform computation via an image-based software (e.g., ImageJ, see Collins (2007)). Therefore, it is labor-intensive, time-consuming, and may introduce contamination to the culture media. Furthermore, the direct measurement method is oftentimes *destructive* – the collected cells are killed in order to take microscopic images. Second, although there are *non-destructive* sensors available, these sensors need to be calibrated, due to the unknown parameters in the sensing relationship (Pan *et al.*, 2019). For example, impedance sensors (adopted in this work, see Figure 2), which measure the dielectric relaxation of a cell suspension, can be used to effectively estimate cell concentrations *after* the calibration of unknown electrical attributes, e.g., permittivity (Gheorghiu and Asami, 1998) and resistivity (Goh and Ram, 2010). However, due to the patient-to-patient variability, those electrical attributes not only are *unknown* but also *vary* among different patients, leading to difficulties in deducing cell concentrations from sensor readings.

We introduce in this work a calibration-free statistical method for online monitoring in cell manufacturing. Specifically, the intrinsic patient-to-patient variability is modeled by a patient-specific calibration parameter. We

CONTACT Jialei Chen jialei.chen@gatech.edu

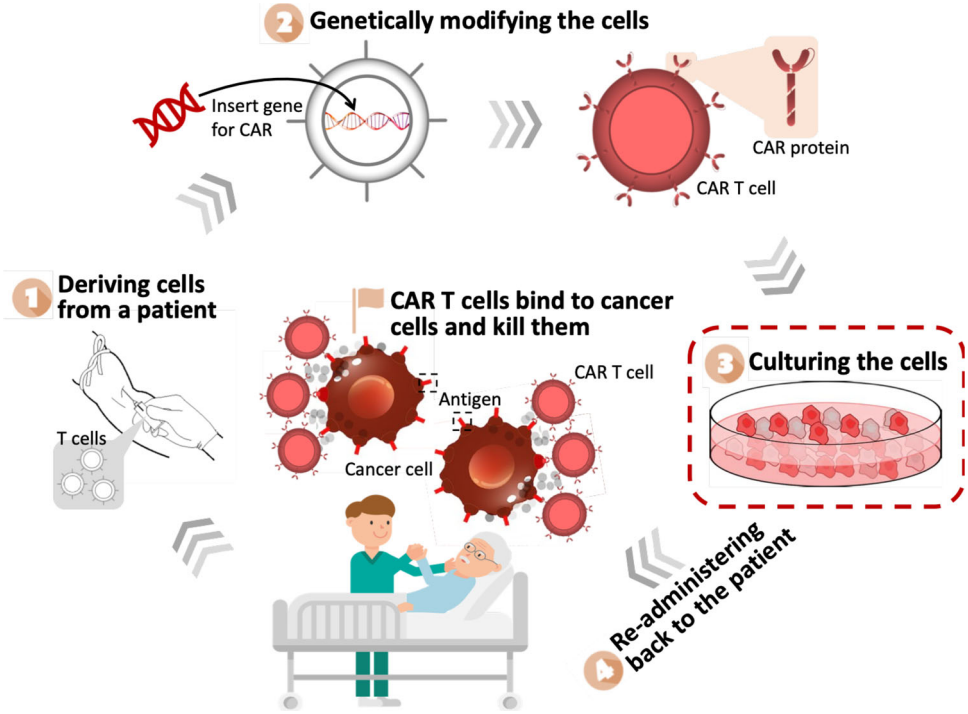

Figure 1. An illustration of the four steps in a typical CAR T-cell therapy. This work focuses on the cell culturing (or cell manufacturing) step, i.e., step 3.

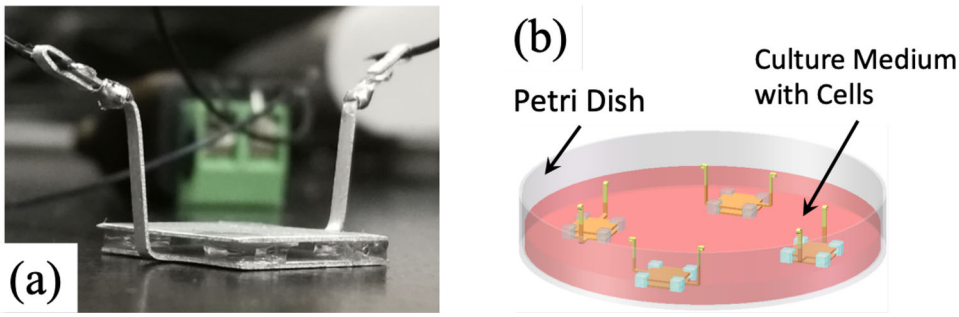

Figure 2. An illustration of the adopted impedance-based biosensors in the cell manufacturing application: (a) shows a photo of the biosensor design and (b) shows the biosensing setup.

propose to use multiple sensor readings to construct a patient-*invariance* statistic, where a transformation is adopted to isolate and alleviate the effect of the calibration parameter. The constructed invariance statistic is then used to model the critical quality attribute of interest. In the *training* stage, we use the historical data to estimate a transformation and model parameters via a carefully formulated optimization problem, rather than estimate the calibration parameter as in the standard calibration problem (Kennedy and O'Hagan, 2001; Tuo and Wu, 2015). In the *monitoring* stage, we use the online sensor readings to deduce the underlying critical quality attribute through the patient-invariance statistic, *free* from the calibration parameter. We demonstrate improvements of the proposed calibration-free method in both simulation experiments and a real-world case study of monitoring viable cell concentrations in cell manufacturing. The proposed approach provides an effective way to monitor cell manufacturing, and therefore, reduces the cost for the promising CAR T-cell therapy in treating cancers.

The remaining part of this article is organized as follows. In Section 2, we formulate the biosensing problem in cell manufacturing, with an emphasis on its challenging aspects. In Section 3, we present the proposed calibration-free method. A detailed simulation study and a real-world cell manufacturing case study are conducted in Sections 4 and 5, respectively. We conclude this work with future directions in Section 6.

2. Biosensing in cell manufacturing

We first describe the biosensing problem of deducing the Viable Cell Concentration (VCC) in cell manufacturing. We then discuss the key challenge – the patient-to-patient variability, and related works.

2.1. Impedance-based biosensing

As discussed in Section 1, the goal is to monitor VCC in cell manufacturing, thereby reducing the cost of the CAR

T-cell therapy. One state-of-the-art approach is to use biosensors to measure impedance signals, as indicators for the VCC of interest (Gheorghiu and Asami, 1998; Pan *et al.*, 2019). As illustrated in Figure 2, we adopt impedance-based biosensors with a Facing-Electrode (FE) design (Miura and Uno, 2019): Our FE biosensor consists of a pair of parallel-plate electrodes and silicone at four corners to maintain a gap between them; it is soaked in media to monitor floating cells in between the electrodes.

With the adopted biosensors, we need a biosensing method to deduce VCCs from impedance readings. From physical knowledge, we know that the impedance reading between the two electrodes reflects the cell amount, due to the capacitive property of viable cell membranes (Schwan, 1957). The sensing relationship is denoted by

$$y = f(x, \theta, \eta), \tag{1}$$

where y is the impedance reading, x is the underlying VCC of interest, θ denotes the sensor geometry, e.g., the gap width, and η models the underlying electrical attributes, e.g., permittivity and resistivity. Here, the relationship $f(\cdot)$ is *unknown,* due to the dynamic interaction between cells and biosensors, and extremely *challenging* to simulate via computer codes considering the micro size of cells. In the proposed calibration-free method, we first learn the sensing relationship $f(\cdot)$ from the historical data of different patients (i.e., the "training" stage), and then conduct online inference of the VCC using the impedance readings for new patients (i.e., the "monitoring" stage).

2.2. Patient-to-patient variability

The key challenge in biosensing is that the electrical attributes η are *unknown* in both training and monitoring and *different* for each patient (also see an illustrating gravity application in Section 4.1). Note that it is impractical to compute η for a patient from first principle since it represents the intrinsic properties from genetic material of the patient. One popular way is to model η as a *calibration* parameter, and estimate it from the training dataset $\{y^j, x^j, \theta\}_{j=1}^J$. Existing approaches include Bayesian implementation (Kennedy and O'Hagan, 2001), maximal likelihood estimation (Joseph and Melkote, 2009), and an interpretative l_2 optimization (Tuo and Wu, 2015). However, such methods are proposed specifically for data fusion of computer experiments and physical experiments, where η is *available* in the former and a *constant* in the latter. Whereas in biosensing, there is *no* effective computer simulator, and electrical attributes η *vary* among physical experimental runs for different patients. In the literature, this challenge is also related to the functional calibration problem (Plumlee *et al.*, 2016; Brown and Atamturktur, 2018; Ezzat *et al.*, 2018), where the calibration parameter $\eta = \eta(x)$ is modeled as a function of the input variables. In the biosensing application, however, the calibration parameter η varies among different patients, yet is a constant over different input VCCs for each patient.

Table 1. A comparison of the application scenarios of the proposed calibration-free method and other standard methods in the literature.

Methods	$f(\cdot)$	*Online* η^*	*Historical* η^j
Inverse problem (Aster *et al.*, 2018)	Known	Unknown	N.A.
Supervised learning (Bishop, 2006)	Unknown	Same	Same
Calibration (Kennedy and O'Hagan, 2001)	Unknown	Unknown	Known
Calibration-free (proposed)	Unknown	Unknown	Unknown

The biosensing problem is also related to the inverse problem in the literature (Aster *et al.*, 2018), where one would estimate both x and η via an optimization problem. However, such a method typically assumes that the sensing relationship $f(\cdot)$ is known or can be easily learned with a *complete* data $\{y^j, x^j, \theta, \eta^j\}_{j=1}^n$, whereas, the calibration parameter η is *unknown* even in the training stage in biosensing. Furthermore, one may regard η as an additional model parameter in the unknown relationship $f(\cdot)$, and adopt a standard supervised learning scheme (Bishop, 2006) for finding both $f(\cdot)$ and η; this implicitly assumes that η is a constant for different patients, which is *not* true in biosensing. Table 1 summarizes the related methods discussed; with extensive efforts in literature search, we have not found a standard method, which can be directly adopted for the biosensing problem in cell manufacturing.

3. Calibration-free biosensing method

We present the proposed calibration-free method in four parts. First, we discuss the sensing relationship for multiple sensors. We then introduce an invariance statistic to alleviate patient-to-patient variability. In the online monitoring stage, we use the invariance statistic to deduce VCCs. In the training stage, we propose a carefully constructed optimization problem and an algorithmic framework to estimate the underlying sensing model.

3.1. Sensing relationship with multiple sensors

The key idea of the calibration-free method is to use *multiple* sensors to address the unknown and patient-specific calibration parameter. For the ith sensor, we let θ_i be its geometry parameter. Consider first the sensing relationship for a given patient (or experiment). Denote $y_i[t]$ as the scalar impedance reading (more details in Section 5) from the ith sensor at experimental time t. Following (1), we model $y_i[t] = f(x[t], \theta_i, \eta)$ via the sensing relationship $f(\cdot) : \mathbb{R}^3 \mapsto \mathbb{R}$. Here, $x[t]$ is the VCC at experimental time t, and η is the calibration parameter. It is important to note that measurements from different sensors $\{y_1[t], y_2[t], \ldots\}$ can be modeled by the *same* calibration parameter η, featuring the underlying values for electrical attributes of the specific patient's cells. This patient-specific property is the key to "canceling out" the calibration parameter using multiple sensor readings (see Section 3.2).

We then introduce an additional superscript j for different patients:

$$y_i^j[t] = f(x^j[t], \theta_i, \eta^j), \quad i = 1, 2, ..., I, \quad t = 1, ...T, \quad \text{and,} \quad j = 1, ..., J. \tag{2}$$

Equation (2) further lays out our biosensing settings with multiple sensors:

1. We assume the *homogeneity* of VCC, i.e., at given time t and a given patient j, the VCC $x^j[t]$ is the same for different senors θ_i at different locations (see Figure 2(b)). This is because, in suspension cell manufacturing, the culture media is constantly stirred to ensure the homogeneity of nutrition, and thereby VCC (Haycock, 2011).
2. We use the same set of sensors with *known* parameters $\{\theta_i\}_{i=1}^I$ for all J patients and at different experimental time t. Those parameters are known from the fabrication process or can be easily measured from the sensors. Note that the proposed method is also effective for different sets of sensors, as long as those sensor parameters are known – the same sensor assumption is only for fabrication convenience and notation simplicity.
3. Also for different sensors θ_i, the calibration parameter η^j is the *same* among different values of the measurement time t. This is because η^j models the intrinsic property of the jth patient's cells, which typically does not change during cell manufacturing. After we clearly layout the above settings, we can then construct the patient-invariance statistic and deduce the underlying VCC.

3.2. Invariance statistic

For notation simplicity, we drop the experimental time $[t]$ and write θ_i in the subscript in this subsection. Furthermore, we rewrite (2) by decomposing $f(\cdot)$ into two parts:

$$y_i^j = f_{\theta_i}(x^j, \eta^j) = \mu_i(x^j) + \delta_i(x^j, \eta^j). \tag{3}$$

Here, for a given sensor θ_i, $\mu_i(\cdot) : \mathbb{R} \mapsto \mathbb{R}$ models the part of effect of VCC x on impedance reading y, *without* being hampered by the patient-specific calibration parameter η; and $\delta_i(\cdot) : \mathbb{R}^2 \mapsto \mathbb{R}$ is the remaining effect of *both* x and η on y. Intuitively speaking, $\mu_i(\cdot)$ can be viewed as the *mean* process of $f(\cdot)$ by plugging in some population average of $\{\eta^j\}_{j=1}^J$, ignoring the patient-to-patient variability; it can also be interpreted as the physical understanding of the sensing relationship. In practice, the mean relationship $\mu_i(\cdot)$ is oftentimes known, at least to a certain degree, prior to experimentation according to the domain-specific knowledge (e.g., the known set of basis functions, see Section 5). On the other hand, $\delta_i(\cdot)$ is the *variability* term, i.e., how patient-to-patient variability affects the impedance reading. Such a term leads to different readings, even when the VCC x is the same. Note that in one of the considered baseline methods (see Sections 4 and 5), we ignore $\delta_i(\cdot)$, i.e., assuming the calibration parameter is a constant; this will lead to noticeable errors when estimating x. This variability term $\delta_i(\cdot)$ is typically unknown. Such a decomposition of a mean trend and a variability term is widely assumed in different modeling methods (see, e.g., Guillas *et al.* (2020) and Chen, Mak, Joseph and Zhang (2020)).

Assume for now $\delta_i(x, \eta) = \delta_i(\eta)$, which suggests that the mean relationship $\mu_i(\cdot)$ extracts all the dependency of x on y (further discussion in Section 3.4). In other words, $f_i(x, \eta)$ is assumed to be *separable* for each sensor θ_i:

$$y_i^j = \mu_i(x^j) + \delta_i(\eta^j). \tag{4}$$

Now, we construct a statistic F that is invariant to the calibration parameter η. To gain intuition, consider the following illustrating example with *known* $\delta_i(\eta) = \theta_i \eta$ for $i = 1$, 2 (see the illustration application in Section 4.1). If we take a log-transformation to the variability term $\log \delta_i(\eta^j) = \log \theta_i + \log \eta^j$, the effect of the calibration parameter η^j is further separated from θ_i. Therefore, by subtracting the (log-transformed) variability at different sensors, one can obtain an invariance statistic $F = \log \delta_1(\eta^j) - \log \delta_2(\eta^j) = \log \theta_1 + \log \eta^j - (\log \theta_2 + \log \eta^j) = \log(\theta_1) - \log(\theta_2)$. Note that we incorporate the patient-*specific* property of the calibration parameter when constructing the invariance statistic.

Following the above intuition yet with the *unknown* variability term $\delta_i(\cdot)$, we construct the following statistic, via a linear combination of the transformed $\delta_i(\cdot)$:

$$F(\eta^j) = \sum_{i=1}^I c_i \mathcal{F}\left[\delta_i(\eta^j)\right] = \sum_{i=1}^I c_i \mathcal{F}\left[y_i^j - \mu_i(x^j)\right], \tag{5}$$

for patient j with η^j. Here, $c_1, ..., c_I$ are pre-defined combination coefficients, and $\sum_i c_i = 0$. With a properly selected transformation $\mathcal{F}[\cdot] : \mathbb{R} \mapsto \mathbb{R}$, (5) gives the target *invariance* statistic $F = F(\eta^j)$. Note that here we adapt a general transformation $\mathcal{F}[\cdot]$, instead of the specific log-transformation in the above example. The transformation $\mathcal{F}[\cdot]$ would be selected so that the dependency of the invariance statistic F to η^j is minimal; a detail estimation method for $\mathcal{F}[\cdot]$ will be discussed in Section 3.4.

It is important to note that we reconstruct the sensing model from (2) to (5) via the proposed invariance statistic. This is again due to the key challenge of patient-to-patient variability. Consider first using (2) for VCC deduction (also see the discussion in Section 2.2). Due to the unknown and patient-specific calibration parameter η^j, it is challenging to either learn a sensing model from training data or deduce VCCs for a new patient. However, the new model (5) contains only the invariance statistic, and is *free* from the calibration parameter. Thanks to the properly selected transformation and the combination (see Section 3.4), the invariance statistic would be approximately a constant for different patients. Therefore, our new model (5) allows an effective estimation of the sensing relationship (only the mean part needed) similar to the standard calibration problem with a *constant* calibration parameter (Tuo and Wu, 2015), and then a calibration-free deduction of the VCC of interest.

3.3. Online calibration-free deduction

We present next the method for deducing the VCC of interest x^*, in the *online monitoring* stage for a new patient denoted by $*$. At any time t, the sensor reading is denoted as $\mathcal{D}_{\text{monitor}} = \{y_i^*, \theta_i\}_{i=1}^I$ along with the unknown calibration

parameter η^*. Assume for now that the mean sensing relationship $\mu_i(\cdot)$ and the transformation $\mathcal{F}[\cdot]$ are known (see Section 3.4 for the estimation). We adopt the new sensing model (5) with the invariance statistic

$$F(\eta^*) = \sum_{i=1}^{I} c_i \mathcal{F}\left[y_i^* - \mu_i(x^*)\right], \tag{6}$$

where x^* is the target VCC. Note that the computed $F(\eta^*)$ in online monitoring is also *invariant* to the calibration parameter η^*. Therefore, the VCC of interest x^* can be deduced by minimizing the squared difference between the computed value and the underlying value $\bar{F}$ (see Section 3.4 for the estimation):

$$\hat{x}^* = \underset{x^*}{\operatorname{argmin}} \left(\sum_{i=1}^{I} c_i \mathcal{F}\left[y_i^* - \mu_i(x^*)\right] - \bar{F} \right)^2. \tag{7}$$

In the cell manufacturing application, we are interested in deducing a VCC curve $\hat{x}^*[t]$ over the whole manufacturing period $t = 1, \ldots, T$. To this end, we perform optimization (7) for T times corresponding to each experimental time t. Note that here we have not incorporated the time-dependency (or smoothness) of the deduced function $\hat{x}^*[t]$ in online deduction; one can use postprocessing methods or directly model $x^*[t]$ via a parametric form in the optimization (7). Readers are referred to functional data analysis literature (Ramsay, 2004; Wang *et al.*, 2016) for more discussion.

3.4. Parameter estimation

We estimate the unknown transformation $\mathcal{F}[\cdot]$, and the physical relationship $\mu_i(\cdot)$ using the training data $\mathcal{D}_{\text{train}} = \{y_i^j[t], x^j[t], \theta_i\}_{t=1}^{T}{}_{j=1}^{J}$ at hand. In our implementation, the transformation $\mathcal{F}[\cdot]$ is paramterized by the Box–Cox transformation (Box and Cox, 1964; Sakia, 1992)

$$\mathcal{F}_\lambda[z] = \begin{cases} \log(z) & \text{if} \quad \lambda = 0 \\ \dfrac{(z^\lambda - 1)}{\lambda} & \text{if} \quad \lambda \neq 0 \end{cases}. \tag{8}$$

Note that the log-transformation in the above example is a special case in the Box–Cox transformation. Here, the Box–Cox transformation contains an unknown parameter λ. A two-parameter Box–Cox or Yeo–Johnson transformation (Yeo and Johnson, 2000) can also be used if the data are not restricted to be non-negative. Furthermore, in this article, we will focus on the parametric transformation (8), but our proposed methods are general and can be applied to non-parametric cases.

As for the physical relationship $\mu_i(\cdot)$, we adopt the following basis decomposition:

$$\mu_{i,\boldsymbol{\beta}}(x) = \mu_{\boldsymbol{\beta}}(x, \theta_i) = \sum_{k=1}^{K} \beta_k \phi_k(x, \theta_i). \tag{9}$$

Here, $\phi_k(\cdot); k = 1, 2, \ldots, K$ are the pre-defined basis functions and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_K]^T$ denotes the vector of corresponding coefficients. Such a set of basis $\{\phi_k(\cdot)\}_{k=1}^{K}$ is selected by the physical knowledge of the cell manufacturing system or the observation from data.

Meanwhile, to account for the separable assumption $\delta_i(x, \eta) \approx \delta_i(\eta)$ in Section 3.2, we introduce slack variables $\boldsymbol{\Delta} = \{\Delta_i^j\}_{i=1}^{I}{}_{j=1}^{J}$ to account for the "goodness" of the assumption (more discussion below). Furthermore, the underlying value for the best invariance statistic $\bar{F}$ is also unknown and need to be estimated from data (see Section 3.3).

We propose to estimate the unknown parameters $\{\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}\}$ via the following optimization problem with two penalization terms:

$$\begin{aligned} \min_{\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}} l_\alpha(\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}) = \min_{\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}} & \sum_{t,j} \\ & \times \left[\sum_{i=1}^{I} c_i \mathcal{F}_\lambda \left[y_i^j[t] - \mu_{\boldsymbol{\beta}}(x^j[t], \theta_i) \right] - \bar{F} \right]^2 \\ & + \alpha_1 \sum_{i,j,t} \left[|y_i^j[t] - \mu_{\boldsymbol{\beta}}(x^j[t], \theta_i)|_1 - \Delta_i^j \right]^2 + \alpha_2 |\boldsymbol{\beta}|_1. \end{aligned} \tag{10}$$

Here, α_1 and α_2 are two penalization coefficients, and $|\cdot|_1$ denotes the vector l_1 norm.

The main objective term (i.e., the first term) in (10) is for achieving the best patient-invariance property of the constructed statistic F. Specifically, we minimize the Mean-Squared Error (MSE) of its underlying truth $\bar{F}$ and the computed value from data. This is equivalent to modeling the patient-invariance statistic F^j for each patient as independent and identically distributed (i.i.d.) random draws from a normal distribution $\mathcal{N}(\bar{F}, \sigma^2)$ with mean $\bar{F}$ and variance σ^2. Moreover, the first penalization term in (10) is for the separable assumption $\delta_i(x, \eta) \approx \delta_i(\eta)$. Here, we minimize the corresponding MSE of the set $\{\delta_i(x^j[t], \eta^j)\}_{t=1}^{T}$ to the underlying truth Δ_i^j, for each sensor i and patient j. Similarly, this can also be viewed as modeling $\delta_i(\cdot, \eta^j)$ via i.i.d. normal random variables; the corresponding penalization α_1 can then be interpreted as the ratio between the variances of the two normal distributions. Finally, the second penalization term $\alpha_2|\beta|_1$ is for basis selection, similar to the widely used LASSO method in the literature (Tibshirani, 1996; Donoho, 2006). This is because, in the cell manufacturing application, one would only have an intuitive understanding of the sensing relationship; we will collect a set of basis functions from experience and select the suitable ones via this penalization.

From the duality of the optimization problem, (10) can also be viewed as unpenalized log-likelihood combining both normal random variables with a sparsity constraint $||\boldsymbol{\beta}||_1 \leq \gamma$ (Boyd and Vandenberghe, 2004). The parameter α_1 sets the variance ratio between the two random variables, and α_2 controls the desired sparsity level as parameter γ. Since the objective is to obtain a high deduction accuracy of the VCC of interest, α_1 and α_2 would be tuned using cross-validation techniques (Friedman *et al.*, 2001).

Algorithm 1. The BCD algorithm for parameter estimation (10)

1: • Set initial values $\lambda \leftarrow 0$ and $\boldsymbol{\beta} \leftarrow 1_K$
2: • Set $I \times J \times T \times K$ tensor $\boldsymbol{\Phi}$ with each element $\Phi_{ijtk} \leftarrow \phi_k(x^j[t], \theta_i)$
3: **repeat**
4: Optimizing F and $\boldsymbol{\Delta}$:
5: • Set $I \times J \times T$ tensor $\mathbf{D}$ with each element $D_{ijt} \leftarrow y_i^j[t] - \sum_k \Phi_{ijtk}\beta_k$
6: • Update $\bar{F} \leftarrow \sum_{ijt} c_i \mathcal{F}_\lambda[D_{ijt}]/J/T$
7: • Update $\Delta_i^j \leftarrow \sum_t |c_i \mathcal{F}_\lambda[D_{ijt}]|_1/T$ for $i = 1, \ldots I$ and $j = 1, \ldots J$
8: Optimizing λ:
9: • Set $l_0(\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}) \leftarrow l_\alpha(\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta})$ with $\alpha_1 = \alpha_2 = 0$
10: • Update $\lambda \leftarrow \text{argmin}_\lambda l_\alpha(\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta})$
11: Optimizing $\boldsymbol{\beta}$:
12: • Update $\boldsymbol{\beta} \leftarrow \text{argmin}_{\boldsymbol{\beta}} l_\alpha(\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta})$ with the L-BFGS method
13: **until** λ, $\bar{F}, \boldsymbol{\Delta}$ and $\boldsymbol{\beta}$ converge
14: • **return** λ, $\bar{F}, \boldsymbol{\Delta}$ and $\boldsymbol{\beta}$

Consider now estimating the parameters $\{\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}\}$ via optimization (10) for fixed $\alpha_1 > 0$ and $\alpha_2 > 0$. We propose to use the following Blockwise Coordinate Descent (BCD) optimization algorithm, described below. First, assign initial values for $\{\lambda, \boldsymbol{\beta}, \bar{F}, \boldsymbol{\Delta}\}$. Next, iterate the following three steps until the convergence is achieved: (i) for fixed λ and $\boldsymbol{\beta}$, update $\{\bar{F}, \boldsymbol{\Delta}\}$ via the following estimates from first-order conditions

$$\hat{\bar{F}} = \frac{1}{JT}\sum_{i,j,t} c_i \mathcal{F}_\lambda\left[y_i^j[t] - \mu_\beta(x^j[t], \theta_i)\right], \qquad \hat{\Delta}_i^j = \frac{1}{T}\sum_{t=1}^{T} |y_i^j[t] - \mu_\beta(x^j[t], \theta_i)|_1; \tag{11}$$

(ii) for fixed $\boldsymbol{\beta}$ and $\bar{F}$, update the transformation parameter λ ignoring the two penalization terms; and (iii) for fixed λ, $\bar{F}$ and $\boldsymbol{\Delta}$, optimize for $\boldsymbol{\beta}$ via numerical line search methods, e.g., L-BFGS algorithm (Liu and Nocedal, 1989). The full optimization procedure is provided in Algorithm 1. Since (10) is a non-convex optimization problem, the proposed BCD algorithm only converges to a stationary solution (Tseng, 2001). Because of this, we suggest performing multiple runs of Algorithm 1 with random initializations for each run, then taking the converged estimates for the run with smallest objective function. These runs should be performed in parallel if possible, to take advantage of the parallel computing capabilities in many computing systems.

It is important to note the difference between the *training* stage in our calibration-free method and the *calibration* stage in the standard calibration problem (Kennedy and O'Hagan, 2001). In calibration methods, the calibration parameter η, assumed to be a constant, is directly estimated from the training set. This can be viewed as estimating a population average of the historical $\{\eta^j\}_{j=1}^J$, which would *not* be helpful in our cell manufacturing application. Due to the patient-to-patient variability, the calibration parameter η^* corresponding to the new patient can be completely different from the historical average value. In contrast, our calibration-free method adopts a patient-invariance statistic F, constructed from multiple sensor readings, to alleviate this patient-to-patient variability. In our training setup, we learn the unknown mean relationship $\mu_i(\cdot)$ and the transformation $\mathcal{F}[\cdot]$, which provide the best patient-invariance statistic F. We can then use the invariance statistic F to effectively deduce VCCs via (7), free from the patient-specific calibration parameter η^*.

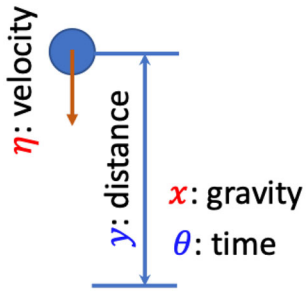

Figure 3. An illustration and notations of the toy application of deducing the gravitational acceleration coefficient.

4. Simulation study

A detailed simulation study is conducted in this section. We first look into a toy application of deducing gravitational acceleration coefficients, to show the applicability of the proposed method. We then discuss more simulation experiments with different sensing relationships.

4.1. A gravity application

Consider the following toy application, where the goal is to deduce the gravitational acceleration coefficient x, for a different planet. As illustrated in Figure 3, we drop a ball and measure the traveling distance y of the ball after a certain period of time θ. From physical knowledge, we have the relationship

$$y = f(x, \theta, \eta) = \frac{1}{2}x\theta^2 + \eta\theta. \tag{12}$$

Here, y is the traveling distance measured by, e.g., taking a photo, η is the initial velocity of the ball, and θ is the time period between dropping the ball and taking the photo. Suppose the ball is dropped by an engineer, meaning that the initial velocity η is non-zero and changes among different drops. With the collected data $\{y, \theta\}$, typically, one cannot deduce the gravitational acceleration x even with the known relationship (12). This is because the initial velocity η is also *unknown*. The key idea of the proposed calibration-free method is to take *multiple* photos at different times $\theta_i; i = 1, \ldots, I$. Therefore, more data $\{y_i, \theta_i\}_{i=1}^I$ is collected with the *same* initial velocity η. We can then use the proposed invariance statistic and Algorithm 1 to "cancel out" η and conduct inference on the gravitational acceleration x of interest.

The setup for deducing the gravity coefficient is as follows. We set the number of photos $I = 3$, with parameters (i.e., the time of taking photos) $\{\theta_i\}_{i=1}^3 = \{0.5, 1, 1.5\}$.

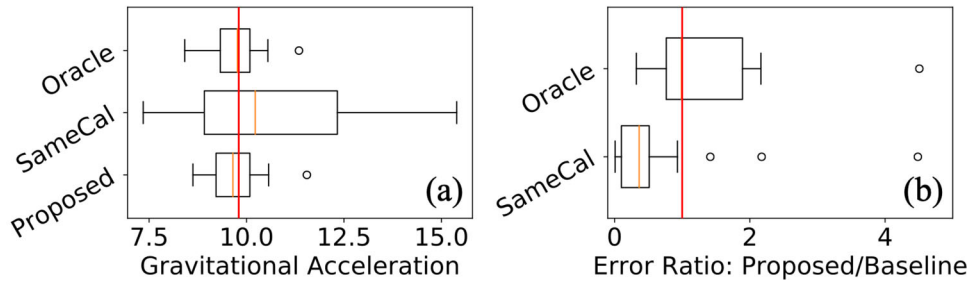

Figure 4. Results of the gravity application: (a) shows the deduced gravitational acceleration by the three considered methods. The red line marks the underlying truth $x^* = 9.8$. (b) shows the boxplots of absolute error ratios between the proposed method and baseline baseline methods. The red line marks the ratio of 1.0.

A historical dataset $\{y_i^j, x^j; \eta^j\}_{j=1}^J$ of size $J = 100$ is generated with calibration parameters (i.e., initial speed, *unknown*) $\eta^j \sim Unif(1,3)$, gravity coefficients $x^j \sim \mathcal{N}(9.8, 1^2)$, and each sensor reading (i.e., traveling distance) y_i^j simulated by (12) with an additional i.i.d. measurement error following $\mathcal{N}(0, 0.4^2)$. To test the deduction accuracy, we let the underlying truth $x^* = 9.8, \eta^* \sim Unif(1,3)$ randomly generated, and y^* obtained by (12) with the same measurement error.

The proposed calibration-free method (via Algorithm 1) is applied to find the best transformation $\mathcal{F}_{\hat{\lambda}}[\cdot]$ and then deduce the gravitational acceleration $\hat{x}^*$. The linear combination coefficients are $\{c_i\}_{i=1}^3 = \{1, 1, -2\}$, and the set of candidate basis functions is $\Phi = \{x, \theta, \theta x, \theta x^2, \theta^2 x\}$. The proposed calibration-free method is repeated, with newly generated test data $\{y^*, x^* = 9.8; \eta^*\}$, for 20 times.

We consider the following two baseline methods (also see Table 1). First, we implement the supervised learning setting (Bishop, 2006), i.e., assuming the calibration parameter $\eta = \eta^j = \eta^*$ is the *same* in both training stage and monitoring stage. Such an assumption is *not* true in the considered cell manufacturing application. Specifically, we use the historical dataset $\{y_i^j, x^j\}_{j=1}^{100}$ to estimate an $\bar{\eta}$ and a relationship $g(x, \theta) := f(x, \theta; \bar{\eta})$, which would then be used to deduce $\hat{x}^*$. Here, we use the same set of basis functions Φ for $g(\cdot)$, and a similar optimization scheme with LASSO-type penalization (Tibshirani, 1996) for estimating the coefficients $\boldsymbol{\beta}$. This method is referred to as "SameCal". The other baseline method is the standard l_2 calibration method suggested by Tuo and Wu (2015). In order to adopt this method, we need to assume that the calibration parameters $\{\eta^j\}_{j=1}^{100}$ in the historical data are *known*, which is *not* true in reality. Therefore, we refer to this method as "Oracle". To estimate the sensing relationship $f(\cdot)$, we adopt the set of basis functions $\Phi_o = \{x, \theta, \theta x, \theta x^2, \theta^2 x, \eta x, \eta\theta, \eta x\theta\}$ and a similar optimization scheme with LASSO-type penalization. Both baseline methods are implemented to deduce the x^* of interest via minimizing the squared difference similar to (7), using the same 20 simulated test data.

Figure 4(a) shows the boxplots of the estimated $\hat{x}^*$ using the proposed calibration-free method and the two baseline methods. The red line indicates the ground-truth value $x^* = 9.8$. Among the three methods, the Oracle baseline preforms the best, as it queries *additional* information of the calibration parameter in the training stage, which is again *not* feasible in reality. We notice that the proposed calibration-free method performs almost as good as Oracle. It can accurately deduce the true value, with the mean over the 20 estimates 9.7 and a relatively small variance, whereas for the SameCal baseline, the mean for 20 estimates is 10.8 and a noticeable bigger variance is observed.

We also conduct a pairwise comparison over the 20 test repetitions. Figure 4(b) shows the boxplots of absolute error ratios of the proposed method over the two baseline methods, with the red line indicating the ratio of 1.0. We notice that the proposed method is only slightly worse than Oracle; this is impressive since our method do *not* query the underlying calibration parameter in the training stage. Furthermore, the proposed calibration-free method is noticeably better in deducing the true x^* compared with SameCal. More specifically, our method outperforms SameCal with smaller errors in 17 estimates over 20 test runs in total. This is not surprising, as the calibration-free method, utilizing the patient-invariance statistic, can address the patient-specific calibration parameters η^*.

4.2. More experiments

Here, we conduct more experiments on the proposed calibration-free method. Specifically, we consider the following four underlying sensing relationships $f_k(x, \theta, \eta)$:

1. $f_1(x, \theta, \eta) = x\theta + \eta\theta^2$;
2. $f_2(x, \theta, \eta) = 3x + 2\theta x + x\theta\eta$;
3. $f_3(x, \theta, \eta) = \theta x + x^2 + \theta\eta^2 + \sqrt{\theta}\eta^2 + \sqrt{x}\eta/4$;
4. $f_4(x, \theta, \eta) = \sin(x) + (x + \eta)^{\theta} + \frac{x}{\theta + \eta}$.

Note that function $f_1(\cdot)$ is quite similar to the sensing relationship (12) in the gravity application in Section 4.1. For functions $f_2(\cdot)$ and $f_3(\cdot)$, we notice the existence of interaction terms between x and η, which means that the separable assumption $\delta_i(x, \eta) = \delta_i(\eta)$ in Section 3.2 does not hold. However, $f_2(\cdot)$ and $f_3(\cdot)$ can still be approximately represented by the adopted set of basis functions Φ. For function $f_4(\cdot)$, it is quite complex, and cannot be represented by Φ. We test all four functions, using the proposed calibration-free method and the same two baseline methods – SameCal and Oracle – introduced in Section 4.1. The detailed test procedure is the same as that in Section 4.1.

Figure 5 shows the boxplots of the absolute error ratios of the proposed method over the baseline SameCal method (a) and the baseline Oracle method (b), under all four

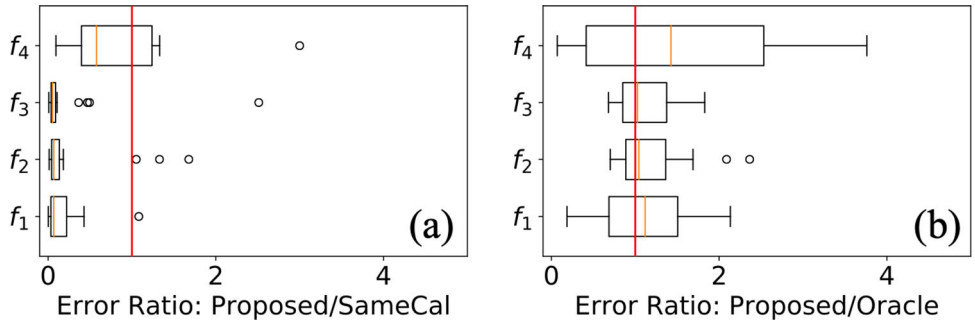

Figure 5. Boxplots of error ratios between the proposed method and the considered baselines, under different sensing relationships. The red line marks the ratio of 1.0, indicating the baseline method achieves the same accuracy as the proposed method.

underlying sensing functions $f_k(x, \theta, \eta); k = 1, ..., 4$. We notice the error ratios in Figure 5(a) are mostly smaller than 1.0, indicting that the proposed method can achieve smaller errors compared with SameCal. This is because the assumption of a constant calibration parameter in SameCal does *not* hold in cell manufacturing application (and thereby in this simulation study), whereas, our calibration-free method can address the patient-specific calibration parameter via a proper combination of multiple sensor readings. Moreover, compared with Oracle, the proposed method is only slightly worse. This shows the good performance of our calibration-free method: Although we do *not* know the values of the calibration parameter, we *can* deduce the underlying parameter of interest similar to the Oracle baseline, where its values *are* assumed accessible. Finally, we notice that for the sensing relationship $f_4(\cdot)$, although the proposed method adopts an inappropriate basis decomposition Φ, it still outperforms SameCal. This is again because the proposed calibration-free method introduces the invariance statistic to alleviate the effect of patient-to-patient variability, and therefore, shows improved performances in deducing the quantity of interest.

5. Cell manufacturing case study

In this section, we apply the proposed calibration-free method to the motivating case study of cell manufacturing. As discussed in Sections 1 and 2, we are interested in deducing and monitoring the VCC $x[t]$ at different values of the experimental time t. This is because the goal of cell manufacturing is to culture a small batch of cells to a significant amount, for an effective re-administering in the CAR T-cell therapy (see Figure 1).

5.1. Experimental setup

In our experiment, human leukemic T cells (Jurkat E6-1; American Type Culture Collection, ATCC®) are cultured in ATCC-formulated culture medium (RPMI-1640; GE Healthcare) with 10% fetal bovine serum, 2 mM L-glutamine, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 1 mM sodium pyruvate, 4500 mg/L glucose, and 1500 mg/L sodium bicarbonate in a 75 cm^2 Petri dish (Nunc™ EasYFlask™; ThermoFisher Scientific™). The cells are cultured in a humidified incubator controlled at 37°C and 5% CO_2, and the culture media is pre-heated to avoid the temperature effect on the impedance measurement (Dlugolecki *et al.*, 2010).

The impedance measurements are obtained by our electric cell-substrate impedance sensing setup. Figure 6 illustrates the experimental setup for the impedance measurement. Here, we use $I = 4$ sandwich shape 3D impedance sensors, consisting of a pair of parallel-plate electrodes and Polydimethylsiloxane (PDMS, Sylgard 184, Tow Corning) to maintain a gap between two electrodes (see Figure 2(b)). In our experimental setup, the geometry parameter θ of the sensors is the edge length of the electrode pads, and $\{\theta_i\}_{i=1}^4 = \{8\text{ mm}, 10\text{ mm}, 12\text{ mm}, 14\text{ mm}\}$. Impedance measurements are conducted by an LCR meter (E4980AL; Keysight Technologies) with a sinusoidal signal of 22 mVrms under multiple frequencies ranging from 500 Hz to 100 kHz. We let the impedance measurement y be the relaxation strength computed from the raw impedance readings over frequencies, i.e., the difference between high-frequency impedance and low-frequency impedance of the dielectric relaxation process, for its high dependency on the VCC of interest (Schwan, 1957). The measurement is taken every 15 minutes for around 30–35 hours. This is because typically after 35 hours, we have to change the culture media and move the cells to a bigger cell culture flask, which would inevitably interrupt the online monitoring. This results in an online monitoring dataset $\mathcal{D}_{\text{monitor}} = \{y_i^j[t], \theta\}_{t=1}^{T}{}_{j=1}^{J}$ with the underlying VCCs to be deduced.

The ground-truth VCCs are obtained by an automated cell counter (TC20™; Bio-Rad Laboratories, Inc.), and the concentration is maintained between 1×10^5 and 1×10^6 cells/mL. Multiple repetitions are performed, with the averaged value reported as the underlying VCC x. Note that the measurement procedure is not only labor-intensive, but also may introduce contamination to the culture media (see Section 1). We will only measure VCC around six times per cell culture experiment, which leads to a much smaller ($\tau \ll T$) yet full training dataset $\mathcal{D}_{\text{train}} = \{y_i^j[t], x^j[t], \theta\}_{t=1}^{\tau}{}_{j=1}^{J}$.

We conduct the cell culturing for $J = 5$ experiments, each with different starting materials, i.e., different calibration parameters η^j. We use the same $I = 4$ sensors $\{\theta_i\}_{i=1}^4 = \{8\text{ mm}, 10\text{ mm}, 12\text{ mm}, 14\text{ mm}\}$ for the $J = 5$ experiments. In experiment j, we measure and compute the relaxation strength of impedance $y_i^j[t]$ for each sensor i, at different time t (every 15 minutes, $T \approx 130$ in total). Meanwhile, we

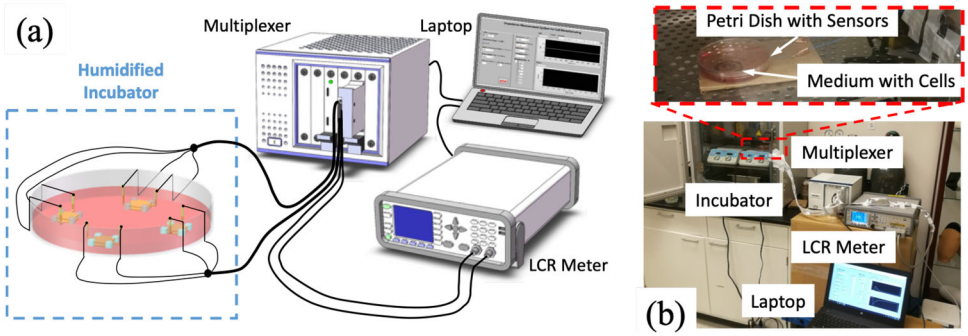

Figure 6. The cell manufacturing application: (a) an illustration and (b) the actual experimental setup with an emphasis on the impedance measurement part.

Table 2. Cross-validation errors of the deduced VCCs for the cell manufacturing case study, using the proposed calibration-free method and the baseline SameCal method.

	Exp 1	*Exp 2*	*Exp 3*	*Exp 4*	*Exp 5*	*Mean*
Proposed	**0.092**	0.270	**0.080**	**0.379**	**0.590**	**0.282**
SameCal	0.500	**0.174**	0.328	0.742	0.760	0.501

measure the ground-truth VCC $x^j[t]$ with a much lower resolution ($\tau \approx 6$ in total).

5.2. Cross validation of viable cell concentration

For the collected training dataset $\mathcal{D}_{\text{train}}$, we first preform a cross-validation test (Friedman *et al.*, 2001) on the deduced VCCs. More specifically, we apply the proposed calibration-free method (via Algorithm 1) to learn the sensing relationship using four out of five experiments, and then deduce VCC $\hat{x}^j[t]$ for the remaining experiment via (7). We let the linear combination coefficients $\{c_i\}_{i=1}^4 = \{1, -1, 1, -1\}$ and select the same set of basis functions Φ as in Section 4.1. Furthermore, a log-transformation on VCCs is performed prior to the analysis. We consider SameCal (see Section 4) as the baseline method. Such a method introduces an additional assumption that the calibration parameter is a constant. Note that the Oracle baseline cannot be adopted here since the actual values of the calibration parameter are always *unknown*, which is the key motivation of the proposed calibration-free method (also see Section 2.2 and Table 1).

Table 2 shows the absolute errors of the deduced log VCCs when the ground-truth VCCs are measured in each experiment. We observe that the proposed method outperforms the baseline SameCal method four experiments out of five. Furthermore, the mean error of the five experiments by the proposed method is 0.282, which is almost two times smaller than that of 0.501 by SameCal. This is due to the fact that the calibration parameter, which models the patient-to-patient variability, is *not* a constant in cell manufacturing (Hinrichs and Restifo, 2013); the proposed calibration-free method properly addresses this variability via the construction of the patient-invariance statistic.

5.3. Online deduction of viable cell concentration

We then perform VCC deduction on the online monitoring set $\mathcal{D}_{\text{monitor}}$. Here, the sensing relationship is estimated using all five experiments in the training set $\mathcal{D}_{\text{train}}$. Figure 7 shows the two deduced log VCC curves over the whole culture time $\log \hat{x}^j[t]$, via the proposed calibration-free method (in red line) and the baseline SameCal method (in green dash line). The ground-truth (log) VCC measurements in $\mathcal{D}_{\text{train}}$ are also plotted in black dots. We see that the proposed method deduces a meaningful estimation of VCC. The deduced $\log \hat{x}^j[t]$ increases approximately linearly over the culture time t, indicating $\hat{x}^j[t]$ grows exponentially in time; this matches the preliminary understanding in the cell culture literature (Haycock, 2011). Furthermore, the deduced $\hat{x}^j[t]$ approximately passes through the ground-truth measurements. However, due to the huge patient-to-patient variability, the baseline SameCal method struggles in either passing through the ground-truth experiments or providing reasonable estimates of VCC curves. Our calibration-free method, adopting the patient-invariance statistic, appears to alleviate such variability well.

The proposed calibration-free method can also provide important biological insights for cell growth in cell manufacturing. From Figure 7(b), we notice a decrease in the VCC curve at around hour 32. This may be due to the lack of nutrition in the media since the culture media typically needs to be changed after 30 hours. Furthermore, we observe from Figure 7 that the VCC curves decrease slightly in the first 2 hours in cell manufacturing. This may be because of the lack of viability of the cells at the beginning of the culture process – though we have already thawed cells and held them at a constant temperature for several minutes, it seems that a certain portion of cells still do not gain full viability and quickly die. As a result, we suggest standing the cell still longer for future experiments. Last but not least, we notice a small VCC decrease when conducting the ground-truth VCC measurements. One reason for this behavior is that the measurement itself is not in-line and needs to contact the culture media; it may introduce contamination, and therefore, kill a small portion of cells (Haycock, 2011). In contrast, the proposed calibration-free biosensing method, together with impedance-based biosensors, provides an in-line, non-destructive, and non-contact way for VCC monitoring in cell manufacturing.

6. Conclusion

In this work, we propose a new calibration-free method for monitoring VCC in cell manufacturing, which is a critical component in the promising CAR T-cell therapy. The key

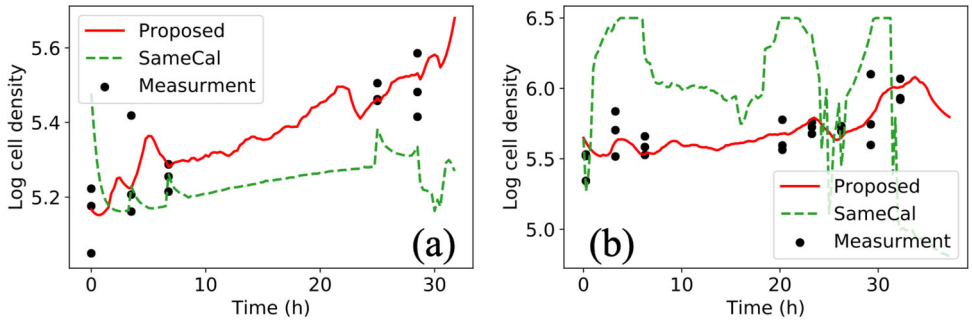

Figure 7. The deduced VCC over time of two cell manufacturing experiments under the two considered methods. The ground-truth VCC measurements are shown in dots.

challenge here is the patient-to-patient variability in the initial culturing material, leading to poor performances in deducing VCCs via existing methods. We propose to use multiple impedance-based biosensors with different geometries and an associated calibration-free statistical framework for online deduction of VCCs. Specifically, we model the patient-to-patient variability via a patient-specific calibration parameter. We then construct a patient-invariance statistic, which uses a transformation and a linear combination of sensor readings to alleviate the effect of the calibration parameter. In the training stage, we learn the best transformation and the sensing relationship via a carefully formulated optimization problem. In the online monitoring stage, VCCs can be deduced via the invariance statistic, free from the patient-specific calibration parameter. We then apply the proposed calibration-free method in different simulation experiments and a real-world case study of cell manufacturing, where the proposed method demonstrates substantial improvements against the existing methods. Therefore, we believe the proposed calibration-free method can play an essential role in cell manufacturing and reduce the cost of the promising CAR T-cell therapy.

Looking ahead, there are several interesting directions for future exploration. To begin with, a more thorough analysis of impedance-based sensors can be conducted, with a detailed comparison of sensitivity using different experimental settings such as sensor geometries and electrode materials. Moreover, we adopt in this work a parametric sensing relationship and a heuristic approach for parameter estimation. This is mainly due to the already improved performance compared with the baseline methods. A more flexible, and non-parametric Gaussian process regression method (Santner *et al.*, 2013; Lin and Joseph, 2020) with a rigorous likelihood-based parameter estimation scheme may lead to further improvements in deducing VCCs, as well as other critical quality attributes. Finally, micro cameras can also be used in cell manufacturing. Therefore, we are also interested in monitoring cell manufacturing based on cell morphology. In this case, physics-informed deep learning frameworks in the literature (Raissi *et al.*, 2017; Chen,Xie, Wang, Zhang, Vannan, Wang and Qian, 2020) appear to be suitable for deducing critical quality attributes in cell manufacturing.

Funding

This work is supported by National Science Foundation CMaT ERC (Award # NSF EEC-1648035), and FDA (Award # 1R01FD006598-01).

Notes on contributors

Jialei Chen is a PhD candidate with the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA. He received the BS degree in theoretical and applied mechanics from the University of Science and Technology of China, Hefei, China, in 2016. Mr. Chen's current research interests include physics-informed learning approaches and data-driven modeling methods for complex engineering system.

Zhaonan Liu is a PhD student with the School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA. He received the BS degree in physics from Peking University, Beijing, China, in 2016. Mr. Liu's current research interests include impedance-based biosensing and printed electronics.

Kan Wang is a research engineer with Georgia Tech Manufacturing Institute, Georgia Institute of Technology, Atlanta, GA, USA. He received the BS degree in theoretical and applied mechanics from Peking University, Beijing, China, in 2005, the M.S. degree in aircraft design from Beihang University, Beijing, in 2007, and the PhD degree in industrial and manufacturing engineering from Florida State University, Tallahassee, FL, USA, in 2013. Dr. Wang's research interests include cell manufacturing, additive manufacturing, and printed electronics technologies and their applications in smart materials and biomedical devices.

Chen Jiang is a PhD candidate with the School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA. She received the BS degree in Mechanical Engineering from Shanghai Institute of Technology, Shanghai, China, in 2014, the MS degree in Mechanical Engineering from Shanghai University, Shanghai, China, in 2017. Miss Jiang's current research interests include tissue engineering and regenerative medicines, biomanufacturing, electrospinning, and textile-based tissue engineering.

Chuck Zhang is Harold E. Smalley Professor with the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology, Atlanta, GA, USA. He received the BS and MS degrees in mechanical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1984 and 1987, respectively, an MS degree in Industrial Engineering from the State University of New York at Buffalo in 1990, and the PhD degree in industrial engineering from the University of Iowa, Iowa City, IA, USA, in 1993. Dr. Zhang has authored over 200 refereed journal articles and 220 conference papers. He also holds 25 U.S. patents. Dr. Zhang is a fellow of Institute of Industrial & Systems Engineers (IISE). His current research interests include additive manufacturing (3-D printing and printed electronics), advanced composite structures manufacturing and maintenance, manufacturing cybersecurity, and biomanufacturing.

Ben Wang is Executive Director of the Georgia Tech Manufacturing Institute, Atlanta. He is Gwaltney Chair in Manufacturing Systems in the School of Industrial and Systems Engineering, Georgia Institute of Technology, where he is also Professor in the School of Materials Science and Engineering. He has authored more than 260 refereed

journal papers journals and is a Co-Inventor of 35 patents or patent applications. With a primary research interest in applying emerging technologies to improve manufacturing competitiveness, he specializes in process development for affordable composite materials and is widely acknowledged as a pioneer in the growing field of nanomaterials. His current research interests include high performance and affordable composites, which is already changing product innovations worldwide. Dr. Wang is a fellow of the Institute of Industrial Engineers, Society of Manufacturing Engineers, and Society for the Advancement of Materials and Process Engineering.

ORCID

Jialei Chen http://orcid.org/0000-0002-6053-9302

References

Aster, R.C., Borchers, B. and Thurber, C.H. (2018) *Parameter Estimation and Inverse Problems*. Elsevier.

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. Springer.

Bonifant, C.L., Jackson, H.J., Brentjens, R.J. and Curran, K.J. (2016) Toxicity and management in CAR T-cell therapy. *Molecular Therapy-Oncolytics*, **3**, 16011.

Box, G.E.P. and Cox, D.R. (1964) An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, **26**, 211–252.

Boyd, S. and Vandenberghe, L. (2004) *Convex Optimization*. Cambridge University Press.

Brown, D.A. and Atamturktur, S. (2018) Nonparametric functional calibration of computer models. *Statistica Sinica*, **28**, pp. 721–742.

Chen, J., Mak, S., Joseph, V.R. and Zhang, C. (2020) Function-on-function kriging, with applications to three-dimensional printing of aortic tissues. *Technometrics*, pp. 1–12.

Chen, J., Xie, Y., Wang, K., Zhang, C., Vannan, M.A., Wang, B. and Qian, Z. (2020) Active image synthesis for efficient labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, to appear.

Collins, T.J. (2007) ImageJ for microscopy. *Biotechniques*, **43**(S1), S25–S30.

Dlugolecki, P., Ogonowski, P., Metz, S.J., Saakes, M., Nijmeijer, K. and Wessling, M. (2010) On the resistances of membrane, diffusion boundary layer and double layer in ion exchange membrane transport. *Journal of Membrane Science*, **349**(1-2), 369–379.

Donoho, D.L. (2006) For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, **59**(6), 797–829.

Ezzat, A.A., Pourhabib, A. and Ding, Y. (2018) Sequential design for functional calibration of computer models. *Technometrics*, **60**(3), 286–296.

Friedman, J., Hastie, T. and Tibshirani, R. (2001) *The Elements of Statistical Learning*, volume 1, Springer, New York, NY.

Gheorghiu, E. and Asami, K. (1998) Monitoring cell cycle by impedance spectroscopy: Experimental and theoretical aspects. *Bioelectrochemistry and Bioenergetics*, **45**(2), 139–143.

Goh, S. and Ram, R. (2010) Impedance spectroscopy for in situ biomass measurements in microbioreactor, in *Proceedings of the 14th International Conference on Miniaturized Systems for Chemistry and Life Science*, pp. 3–7.

Guillas, S., Sarri, A., Day, S.J., Liu, X. and Dias, F. (2018) Functional emulation of high resolution tsunami modelling over cascadia. *The Annals of Applied Statistics*, **12**(4), 2023–2053.

Harrison, R.P., Zylberberg, E., Ellison, S. and Levine, B.L. (2019) Chimeric antigen receptor–T cell therapy manufacturing: modelling the effect of offshore production on aggregate cost of goods. *Cytotherapy*, **21**(2), 224–233.

Haycock, J.W. (2011) 3D cell culture: A review of current approaches and techniques, in *3D Cell Culture. Methods in Molecular Biology (Methods and Protocols)*, vol. 695, Humana Press, pp. 1–15.

Hinrichs, C.S. and Restifo, N.P. (2013) Reassessing target antigens for adoptive T-cell therapy. *Nature Biotechnology*, **31**(11), 999–1008.

Joseph, V.R. and Melkote, S.N. (2009) Statistical adjustments to engineering models. *Journal of Quality Technology*, **41**(4), 362–375.

June, C.H., O'Connor, R.S., Kawalekar, O.U., Ghassemi, S. and Milone, M.C. (2018) CAR T cell immunotherapy for human cancer. *Science*, **359**(6382), 1361–1365.

Kennedy, M.C. and O'Hagan, A. (2001) Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B*, **63**(3), 425–464.

Kim, S.U. and de Vellis, J. (2009) Stem cell-based cell therapy in neurological diseases: A review. *Journal of Neuroscience Research*, **87**(10), 2183–2200.

Lin, L.-H. and Joseph, V.R. (2020) Transformation and additivity in Gaussian processes. *Technometrics*, **62**(4), 525–535.

Liu, D.C. and Nocedal, J. (1989) On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, **45**(1-3), 503–528.

Miura, T. and Uno, S. (2019) Computer simulation for electrochemical impedance of a living cell adhered on the inter-digitated electrode sensors. *Japanese Journal of Applied Physics*, **58**(SB), SBBG15.

Pan, Y., Hu, N., Wei, X., Gong, L., Zhang, B., Wan, H. and Wang, P. (2019) 3D cell-based biosensor for cell viability and drug assessment by 3D electric cell/matrigel-substrate impedance sensing. *Biosensors and Bioelectronics*, **130**, 344–351.

Plumlee, M., Joseph, V.R. and Yang, H. (2016) Calibrating functional parameters in the ion channel models of cardiac cells. *Journal of the American Statistical Association*, **111**(514), 500–509.

Prasad, V. (2018) Tisagenlecleucel – the first approved CAR-T-cell therapy: Implications for payers and policy makers. *Nature Reviews Clinical Oncology*, **15**(1), 11–12.

Raissi, M., Perdikaris, P. and Karniadakis, G.E. (2017) Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations. *arXiv preprint arXiv:1711.10561*.

Ramsay, J.O. (2004) Functional data analysis. *Encyclopedia of Statistical Sciences*, **4**.

Sakia, R.M. (1992) The Box–Cox transformation technique: A review. *Journal of the Royal Statistical Society: Series D (The Statistician)*, **41**(2), 169–178.

Santner, T.J., Williams, B.J. and Notz, W.I. (2013) *The Design and Analysis of Computer Experiments*. Springer.

Schwan, H.P. (1957) Electrical properties of tissue and cell suspensions, in *Advances in Biological and Medical Physics*, volume 5, Elsevier, pp. 147–209.

Slouka, C., Wurm, D.J., Brunauer, G., Welzl-Wachter, A., Spadiut, O., Fleig, J. and Herwig, C. (2016) A novel application for low frequency electrochemical impedance spectroscopy as an online process monitoring tool for viable cell concentrations. *Sensors*, **16**(11), 1900.

Tibshirani, R. (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**(1), 267–288.

Tseng, P. (2001) Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of Optimization Theory and Applications*, **109**(3), 475–494.

Tuo, R. and Wu, C.F.J. (2015) Efficient calibration for imperfect computer models. *The Annals of Statistics*, **43**(6), 2331–2352.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016) Functional data analysis. *Annual Review of Statistics and its Application*, **3**, 257–295.

Yeo, I.-K. and Johnson, R.A. (2000) A new family of power transformations to improve normality or symmetry. *Biometrika*, **87**(4), 954–959.

Yin, P. (2017) Keys to scale-up CAR T-cell therapy manufacturing. *BioProcess Online*.