Practical Use of Regularization in Individualizing a Mathematical Model of Cardiovascular Hemodynamics Using Scarce Data

Ali Tivay1, Xin Jin1, Alex Kai-Yuan Lo1, Christopher G. Scully2 and Jin-Oh Hahn1*

1 Department of Mechanical Engineering, University of Maryland, College Park, College Park, MD, United States, 2 Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, United States

Individualizing physiological models to a patient can enable patient-specific monitoring and treatment in critical care environments. However, this task often presents a unique “practical identifiability” challenge due to the conflict between model complexity and data scarcity. Regularization provides an established framework to cope with this conflict by compensating for data scarcity with prior knowledge. However, regularization has not been widely pursued in individualizing physiological models to facilitate patient-specific critical care. Thus, the goal of this work is to garner potentially generalizable insight into the practical use of regularization in individualizing a complex physiological model using scarce data by investigating its effect in a clinically significant critical care case study of blood volume kinetics and cardiovascular hemodynamics in hemorrhage and circulatory resuscitation. We construct a population-average model as prior knowledge and individualize the physiological model via regularization to illustrate that regularization can be effective in individualizing a physiological model to learn salient individual-specific characteristics (resulting in the goodness of fit to individual-specific data) while restricting unnecessary deviations from the population-average model (achieving practical identifiability). We also illustrate that regularization yields parsimonious individualization of only sensitive parameters as well as adequate physiological plausibility and relevance in predicting internal physiological states.

Keywords: individualization, physiological model, regularization, practical identifiability, volume kinetics, cardiovascular hemodynamics, hemorrhage, resuscitation

INTRODUCTION

Clinical care automation has been a domain of interest for a few decades by virtue of its potential for error-free and vigilant performance of routine and low-level patient monitoring and treatment tasks (Hemmerling et al., 2010; Salinas et al., 2011; Dussaussoy et al., 2014; Rinehart et al., 2015; Brogi et al., 2017; Hundeshagen et al., 2017; Pasin et al., 2017), yet realizing this potential is contingent upon establishing the safety and the performance characteristics of clinical care automation. Patient physiology models built upon physical principles (hereafter called physiological models) can facilitate the development (Jin-Oh et al., 2012; Bighamian et al., 2014;...
Jin et al., 2018) and the testing (Kovatchev et al., 2009; Ortiz et al., 2010; Brown et al., 2015) of clinical care automation capabilities. However, individualizing physiological models (which can enable the systematic development and the testing of patient-specific clinical care automation) presents formidable challenge due to the conflict between model complexity and data scarcity. That is, physiological models are complex and involve a large number of unknown patient-specific parameters that frequently exhibit interaction properties, whereas clinical data pertaining to an individual patient are scarce in both quantity and quality. For example, blood pressure (BP) can increase in response to an increase in cardiac output (CO) and an increase in total peripheral resistance (TPR), but routine clinical measurements often provide only BP which does not offer sufficient information to uniquely determine CO- and TPR-related model parameters. In fact, such a conflict appears to be a main obstacle in characterizing a wide range of mechanistic models in physiology as well as in broader fields of biology and physics, which has been called the sloppiness property (Transtrum et al., 2010, 2015; Machta et al., 2013; White et al., 2016) or lack of practical identifiability (Raue et al., 2009; Maiwald et al., 2011) in different contexts. If not properly addressed, such conflict may yield a physiological model suffering from non-unique and physically irrelevant parameter values as well as poor individual-specific internal (i.e., unmeasured) state predictions.

The existing body of work has addressed the practical identifiability challenge mainly in two ways: (i) by modifying the model structure or parameterization toward more simplified or lumped models and (ii) by utilizing some form of regularization (or more generally, prior knowledge) to additionally inform the modeling procedure and effectively reduce the complexity of the model search space. Analyzing and modifying the model structure has particularly strong precedents in physiological modeling, where effective tools such as variation-based sensitivity analysis (Sobol, 2001; Saltelli et al., 2008, 2010) and profile-likelihood (Raue et al., 2009, 2014; Maiwald et al., 2016) methods are used to assess the appropriateness of a model structure and the importance of model parameters with respect to given data. While useful in informing potential modifications to the model structure and parameterization (Machta et al., 2013; Mattingly et al., 2018), many of these tools are formalized on a pre-determined level of data availability, making them less powerful if data availability varies (which is quite common in clinical scenarios). Regularization is an established framework to cope with conflicts between model complexity and data scarcity, by which one can potentially compensate for the scarce data available to individualize a physiological model based on prior knowledge. To date, the notion of prior distribution (Toni et al., 2008; Brown et al., 2015) and its closely related counterpart in regularization (Benning and Burger, 2018) have been pursued in a wide range of deterministic and stochastic modeling problems to incorporate additional information into the modeling problem and control the complexity of the model search space (Davidian and Giltinan, 2003; Tivay et al., 2019; Zhang et al., 2019).

The goal of this work is to garner potentially generalizable insight into the role of regularization in individualizing a complex physiological model using scarce data by studying its effect in a clinically significant critical care case study of blood volume (BV) kinetics and cardiovascular hemodynamics in hemorrhage and circulatory resuscitation. We construct a population-average physiological model as prior knowledge and individualize the physiological model via regularization to elucidate how regularization enables the physiological model to capture salient individual-specific characteristics (leading to the goodness of fit to individual-specific data) while restricting unnecessary deviations from prior knowledge (achieving practical identifiability). We also analyze how regularization affects the plausibility and the relevance of the physiological model under varying levels of scarcity in individual-specific data in terms of its parameter values and internal state predictions.

This paper is organized as follows: section “Cardiovascular Hemodynamics in Hemorrhage and Circulatory Resuscitation” outlines the physiological model and the experimental data used in the case study. Section “Materials and Methods” describes a step-by-step procedure for applying regularization to individualize the physiological model. Section “Results” presents the results, which are discussed in section “Discussion”. Section “Conclusions” concludes the paper with suggested future work.

CARDIOVASCULAR HEMODYNAMICS IN HEMORRHAGE AND CIRCULATORY RESUSCITATION

Credible physiological models of BV kinetics and cardiovascular hemodynamics have the potential to facilitate the development and the testing of circulatory resuscitation devices and algorithms (Parvinian et al., 2018). Prior work has reported a wide spectrum of physiological models (in terms of scale and detail) to represent volume kinetics (Hahn, 2010; Bighamian et al., 2016) and cardiovascular hemodynamics (Moss et al., 2012; Bighamian et al., 2018) in response to stimuli. In some cases, these physiological models were also analyzed for practical identifiability (Moss et al., 2012; Marquis et al., 2018; Pathmanathan et al., 2019; Pironet et al., 2019). In this work, we employ an extended and enhanced version of a physiological model reported in our prior work (Bighamian et al., 2018), which concerns the effect of hemorrhage and circulatory resuscitation on the changes in BV kinetics and cardiovascular hemodynamics.

Model Description

The physiological model is composed of component mathematical models that represent BV kinetics, cardiovascular function, and BP regulation, all at the systems level (Figure 1).

The BV kinetics model is a lumped-parameter model that simulates the effect of fluid gain, blood loss, and inter-compartmental fluid exchange on changes in BV. In this model, the time rate of change in BV (∆V_B) can be described by the following differential equation:

\[ \Delta V_B(t) = u(t) - v(t) - u_{ex}(t) \] (1)

where \( u \), \( v \), and \( u_{ex} \), respectively, denote the rate of external fluid gain (e.g., fluid resuscitation), external fluid loss (e.g.,
hemorrhage and urinary output), and internal fluid exchange between BV and interstitial tissue compartments. The rate of fluid exchange is macroscopically modeled as a control input signal that regulates BV to a target value \( r_{BV} \) as follows:

\[
u_{ex}(t) = -K_p \left[ r_{BV}(t) - \Delta V_B(t) \right]
\]  

(2)

where \( K_p \) is the controller gain and \( r_{BV} \) represents the target change in BV, which is computed according to the history of external fluid gain and loss as follows:

\[
r_{BV}(t) = \frac{1}{1 + \alpha} \int_0^t u(\tau) \, d\tau - \frac{1}{1 + \alpha} \int_0^t v(\tau) \, d\tau
\]

(3)

where the ratio \( \frac{1}{1 + \alpha} \) determines the fraction of the total external fluid gain that eventually remains within the BV compartment (with the rest shifted to the interstitial tissue compartment), and the ratio \( \frac{1}{1 + \alpha} \) determines the fraction of the total fluid loss that is not compensated for by a refill from the interstitial tissue compartment. In our prior work, we demonstrated that the macroscopic expressions above can reproduce BV kinetic responses to volume perturbations with crystalloid, colloid, and blood (Bighamian et al., 2016, 2018).

The cardiovascular function model is built to compute CO from BV based on the CO–venous return equilibrium principle (Guyton et al., 1975) and the left ventricular (LV) pressure–volume relationship (Lankhaar et al., 2009). The details of such a derivation can be found in our previous work (Bighamian et al., 2018), which yields the following model relating BV to CO:

\[
CO(t) = \frac{HR(t)}{A + \frac{B}{r} \frac{HR(t)}{TPR(t)}} \log \left( \frac{\gamma k}{B} TPR(t) CO(t) \right)
\]

\[
+ \frac{\gamma}{BC_S} V_B(t) - \frac{\gamma}{BC_S} \eta V_{BV} + 1
\]

(4)

where \( HR(t) \) is the heart rate, \( TPR(t) \) is the TPR associated with the arterial circulation, \( C_S \) is the systemic capacitance, \( E_s \) is the left ventricular elastance, \( A \) and \( B \) define the shape of the LV pressure–volume loop, \( \gamma \) represents the approximate proportionality constant between central venous \( (P_{CV}) \) and LV end diastolic \( (p_{LVED}) \) pressures \( (\gamma \approx \frac{p_{LVED}}{P_{CV}}) \), \( k \) denotes the ratio between the resistance to venous return and TPR \( (k = \frac{R_{VR}}{TPR}) \), \( V_{BV} \) is the initial BV, and \( \eta \) represents a proportional relationship between \( V_{BV} \) and the unstressed BV \( V_{BU} \): \( V_{BU} = \eta V_{BV} \).

The BP regulation component of the model approximates the regulatory mechanism in the body that monitors the discrepancy between a pre-specified, intrinsic mean arterial BP (MAP) set point \( (MAP_{target}) \) versus the subject’s actual MAP \( (MAP(t)) \) and adjusts TPR (and accordingly, the resistance to venous return) to maintain MAP near the set point level. This component is modeled with the following equations:

\[
MAP(t) = CO(t) TPR(t) = CO(t) [TPR_0 + \Delta TPR(t)]
\]

(5a)

\[
\Delta TPR(t) + \rho_{TPR} \Delta TPR(t) = k_{TPR} [MAP_{target} - MAP(t)]
\]

(5b)

where \( TPR_0 \) is the initial value of TPR, which can be found from the initial value of CO and MAP \( (TPR_0 = MAP_0/CO_0) \), \( k_{TPR} \) is the gain of MAP regulation, and \( \rho_{TPR} \) represents the speed of MAP regulation. Due to the relatively short time period (180 min) associated with the dataset as well as the presumably maximal level of short-term BP regulation responses due to large initial hemorrhage, the full mechanistic details of BP regulation mechanisms (e.g., the long-term aspects associated with the responses of the renin–angiotensin–aldosterone and anti-diuretic hormone systems) are not considered, and the collective short-term effects of BP regulation mechanisms are macroscopically expressed by Eq. 5.

**Experimental Dataset**

The experimental measurements acquired from animals in an array of prior work (Rafie et al., 2004; Vaid et al., 2006; Marques et al., 2017) were used as data in this case study. A total of \( N = 23 \) sheep constituted the population considered in this study. The
data included the time series sequences of BV, CO, and MAP acquired from these sheep subjects. In each sheep, initial BV was measured using a dye concentration method. Then, subsequent changes in BV were measured with hemoglobin dilution after correction for the loss of red blood cells. CO was measured with a flow probe placed on the pulmonary artery. MAP was measured with arterial catheterization. These signals were measured at 5-min intervals during a measurement period of 180 min. Each sheep was subjected to a large initial hemorrhage (25 ml/kg; see "H" at 0–15 sec in Figure 4) followed by resuscitation (see "I" at 30–180 min in Figure 4) and subsequent minor hemorrhages (5 ml/kg; see "H" at 50–55 and 70–75 min in Figure 4). Fluid resuscitation was performed by previously developed closed-loop control algorithms (Rafie et al., 2004; Vaid et al., 2006; Marques et al., 2017) designed to maintain MAP. Hemorrhage and fluid resuscitation yielded diverse BV, CO, and MAP responses in the datasets, which is appropriate to analyze the role of regularization in individualizing the physiological model (Table 1).

![Image of the page content here](https://example.com/page-image.png)

**MATERIALS AND METHODS**

In this section, we present a procedure for applying regularization to the physiological model outlined in section “Cardiovascular Hemodynamics in Hemorrhage and Circulatory Resuscitation”. Special emphasis is given to the construction of prior knowledge, which involves the choice of a regularization function employed in individualizing the physiological model. Then, data analysis details are presented.

**Individualizing the Physiological Model**

Given the data associated with an individual subject $i$, the inputs given to the subject (i.e., hemorrhage and fluid resuscitation) are denoted by $U^d_i = \{u(t_1), v(t)\}$, while the outputs of the subject are denoted by a vector $Y^d_i$. For instance, in case BV, CO, and MAP data are available in the individual $i$, then $Y^d_i$ is defined as:

$$Y^d_i = \begin{bmatrix} BV^d_i (t_1, \ldots, t_{n_{BV}}) & CO^d_i (t_1, \ldots, t_{n_{CO}}) \\ MAP^d_i (t_1, \ldots, t_{n_{MAP}}) \end{bmatrix}$$ (6)

where $X^d_i (t_1, \ldots, t_{n_X})$ is a row vector containing the measurements of $X$ at times $(t_1, \ldots, t_{n_X})$. The nominal noise covariance matrix associated with $Y^d_i$ is $\Sigma = \text{diag} (\sigma^2_{BV}, \sigma^2_{CO}, \sigma^2_{MAP})$.

To obtain model predictions $Y_i(\theta, U^d_i)$ corresponding to the elements of the data vector $Y^d_i$, the model equations in section “Cardiovascular Hemodynamics in Hemorrhage and Circulatory Resuscitation” are numerically solved given the inputs $U^d_i$ and the vector of lumped model parameters $\theta$, which is defined as follows (see Appendix Table A1 for details and ranges):

$$\theta = \left[ a_u \ a_v \ K_p \ A \ \frac{y_k}{E_s} \ \frac{y}{B} \ \frac{\eta}{BC_r} \ k_{TPR} \ P_{TPR} \ MAP_{target} \right]$$ (7)

The problem of individualizing the physiological model can be formulated as solving the following maximum a posteriori estimation problem (Gelman et al., 2013):

$$\hat{\theta}_i = \arg\max_{\theta} P(\theta|U^d_i, Y^d_i) = \arg\max_{\theta} P(Y^d_i|U^d_i, \theta)P(\theta)$$

$$= \arg\min_{\theta} \left[ -\log P \left( Y^d_i | U^d_i, \theta \right) - \log P(\theta) \right]$$

$$= \arg\min_{\theta} \left[ J_1(\theta) + J_2(\theta) \right]$$ (8)

where $\hat{\theta}_i$ is the vector of model parameters associated with the individual $i$, $J_1(\theta) = -\log P \left( Y^d_i | U^d_i, \theta \right)$ corresponds to the likelihood of the individual-specific data with respect to the model output, and $J_2(\theta) = -\log P(\theta)$ is a regularization term that corresponds to prior knowledge $P(\theta)$ about the parameter values. Assuming that $P \left( Y^d_i | U^d_i, \theta \right)$ is a Gaussian distribution due to noise on the output measurements, we can set $J_1(\theta)$ as the covariance-weighted mean-squared output errors:

$$J_1(\theta) = \left[ Y_i \left( U^d_i, \theta \right) - Y^d_i \right] \Sigma^{-1} \left[ Y_i \left( U^d_i, \theta \right) - Y^d_i \right]^T$$ (9)

In case there is no informative prior knowledge about the probability distribution of the model parameters, $J_2(\theta)$ in Eq. 8 does not depend on $\theta$ and is effectively removed. However, in that case, the optimization problem in Eq. 8 is likely to become ill-posed and suffer from practical identifiability challenges when the physiological model is too complex relative to the available individual-specific data (which is frequently the case).

The main idea to address the practical identifiability problem in the absence of sufficient individual-specific data is to construct an appropriate prior distribution $P(\theta)$ to additionally inform the modeling problem. This prior distribution can be constructed based on the prior knowledge about plausible parameter ranges as well as heterogeneous data from a population of individuals. In this work, we formulate a prior in the shape of a Laplace distribution of the following form:

$$P(\theta) \propto \prod_{m=1}^{m_{\text{th}}} \exp \left( -\frac{|\theta_{m} - \bar{\theta}_m|}{b_m} \right) \rightarrow J_2(\theta) = \lambda \sum_{m=1}^{m_{\text{th}}} \frac{1}{b_m} |\theta_{m} - \bar{\theta}_m| $$ (10)

where $\theta_{m}$ is the $m$-th element in $\theta$, $\bar{\theta}_m$ is the $m$-th element of the mode of the distribution $\bar{\theta}$, $\lambda$ is the regularization weight, and $b_m$ is a normalization factor for $\theta_{m}$, which reflects prior knowledge about reasonable and/or physiological ranges for the model parameter (i.e., given physiological upper $\bar{\theta}_m$ and lower $\bar{\theta}_m$) bounds on $\theta_m$, we set $b_m = (\bar{\theta}_m - \bar{\theta}_m)$; see Appendix Table A1.

**TABLE 1 |** Range of blood volume (BV), cardiac output (CO), and mean arterial BP (MAP) responses at important time intervals during the experimental protocol [mean (SD)], $N = 23$.

| Time (min)          | BV (L) | CO (L/min) | MAP (mmHg)  |
|---------------------|--------|------------|-------------|
| Baseline pre-hemorrhage ($t = 0$) | 2.37 (0.44) | 4.40 (0.96) | 91.22 (10.60) |
| Immediate post-hemorrhage ($t = 15$) | 1.83 (0.44) | 2.23 (0.83) | 52.33 (16.73) |
| Start of resuscitation ($t = 30$)  | 1.86 (0.41) | 2.21 (0.79) | 53.17 (10.71) |
| End of resuscitation ($t = 180$)   | 2.22 (0.58) | 4.67 (1.15) | 82.80 (6.31)  |
for \( \hat{\theta}_m \) and \( \bar{\theta}_m \) associated with the physiological model in section “Cardiovascular Hemodynamics in Hemorrhage and Circulatory Resuscitation”]. Equation 10 is a form of \( L_1 \)-regularization (Bach et al., 2011; Jenatton et al., 2011), which represents the inter-individual variability observed in the population data by compressed parametric deviations from a mode denoted by \( \bar{\theta} \). This representation may be possible if the studied physiological modeling problem has the sloppiness property (Machta et al., 2013; Transtrum et al., 2015), i.e., a wide range of parameter sensitivities, implying compressibility (see section “Role of Regularization in Individualizing Physiological Models” for more details as far as our case study is concerned).

Having Eq. 10, the problem of constructing the prior reduces to determining the parameters \( \lambda \) and \( \bar{\theta} \) from heterogeneous population data. Techniques exist in the field of non-linear mixed-effects modeling (Davidian and Giltinan, 2003) and system identification (Pan et al., 2018) that can deal with such problems, often with comparable results (Kataria et al., 1994; Hahn et al., 2011). In this work, we define \( \bar{\theta} \) as the solution to the following maximum-likelihood optimization problem based on the population data from \( L = N - 1 \) individuals:

\[
\bar{\theta} = \arg\min_{\theta} \sum_{i=1}^{L} \left[ Y_i \left( U_i^d, \theta \right) - Y_i^d \right] \Sigma^{-1} \left[ Y_i \left( U_i^d, \theta \right) - Y_i^d \right]^T
\]  

(11)

Note that the population-based modeling problem in Eq. 11 is supposedly less susceptible to practical identifiability challenges than the individualization problem in Eq. 8 by virtue of the larger amount of heterogeneous data available from the population. The resulting model with \( \bar{\theta} \) will hereafter be called the population-average model, which is intended to represent an aggregated and sensible model of typical physiological behavior.

The regularization weight \( \lambda \) serves as the relative scaling factor between the spread of the prior in Eq. 10 and the spread of the likelihood in Eq. 9. To find an appropriate individual-specific weight \( \lambda_i^* \), we adopt an approach related to the L-curve method (Hansen, 1992; Hansen and O’Leary, 1993).

The weight \( \lambda_i^* \) may be estimated by solving the optimization problem in Eq. 8 for a range of \( \lambda \in [0, \lambda_{\text{max}}] \) and plotting the resulting values for the goodness of fit cost \( J_1(\theta) \) and the deviation distance \( D(\theta) = J_2(\theta)/\lambda \). (Figure 2). An initial increase of \( \lambda \) from zero tends to decrease \( D \) by compressing parametric deviations from \( \bar{\theta} \) in low-sensitivity directions. This increases \( J_1 \), but the extent is not large relative to the decrease in \( D \) since the effect of restricting low-sensitivity (i.e., sloppy) deviations on \( J_1 \) is, by definition, small. This trend continues until \( \lambda \) reaches a critical value. However, beyond this critical value, sensitive deviations also become restricted, leading to a substantial increase in \( J_1 \) relative to the decrease in \( D \). The appropriate \( \lambda_i^* \) can be chosen as the value of \( \lambda \) in between these regime changes (Figure 2).

In sum, individualizing a physiological model via regularization can be performed in two steps: (i) constructing a prior (regularization function) from heterogeneous data measured in a population by solving Eq. 11 and (ii) individualizing the physiological model with regularization using scarce individual-specific data by solving Eq. 8 (Figure 3).

**Data Analysis**

We applied the individualization procedure outlined in section “Individualizing the Physiological Model” to the physiological model outlined in section “Cardiovascular Hemodynamics in Hemorrhage and Circulatory Resuscitation”. In particular, we constructed an appropriate regularization function and individualized the physiological model under varying levels of scarcity in the individual-specific data. To enable the data analysis with relatively small sample size, we performed a leave-one-out type of data analysis (details follow).

We individualized the physiological model under three levels of increasing data scarcity: (i) case 1, in which the physiological model was individualized given BV, CO, and MAP data, (ii) case 2, in which the physiological model was individualized given CO and MAP data, and (iii) case 3, in which the model was individualized given only MAP data. We considered these data scarcity scenarios for two investigative reasons: (i) comparison of the physiological models individualized with and without the use of regularization in the three cases may offer insight into how regularization affects the parameter values in response to varying data scarcity and (ii) analysis of the prediction errors for unmeasured internal states (i.e., BV in case 2 and BV and CO in case 3) with and without the use of regularization may offer insight into how regularization affects the physiological plausibility and the relevance of the physiological model under data scarcity.

For each animal subject, we analyzed the data pertaining to the rest of the \( L = 22 \) animal subjects with the maximum-likelihood method in Eq. 11 to derive a population-average physiological model, characterized by \( \theta \), that predicts BV, CO, and MAP responses to hemorrhage and fluid resuscitation inputs. Then, we individualized the physiological model to the (excluded) animal via regularization by solving Eq. 8 with Eq. 10 as the regularization function. We selected the...
appropriate value of $\lambda$ for each individual subject as described in section "Individualizing the Physiological Model" (Figure 2). We repeated individualizing the physiological model in this way for the three dataset scarcity cases. For comparative analysis, we also individualized the physiological model to each animal without using regularization for all the three data scarcity scenarios.

With the goal of assessing the plausibility and the relevance of the physiological models, we compared the models individualized with and without regularization as well as the population-average model. We devised two quantitative error metrics for this purpose: (i) output prediction error $e_1$, defined as the normalized root-mean-square error (RMSE) between model-predicted and measured output signals (among $BV$, $CO$, and $MAP$) explicitly used for individualizing the physiological model and (ii) state prediction error $e_2$, defined as the normalized RMSE between model-predicted and measured internal state signals not used in individualizing the physiological model. For example, in case 3, $e_1$ is the normalized RMSE associated with $MAP$:

$$e_1 = \frac{1}{N \sigma_{MAP}} \sum_{i=1}^{N} \left[ \frac{1}{\sigma_{MAP}} \| MAP_i(t_1, \ldots, t_{n_{MAP}}) - MAP_{id}(t_1, \ldots, t_{n_{MAP}}) \|_2 \right]$$

(12)

while $e_2$ is the normalized RMSE associated with $BV$ and $CO$:

$$e_2 = \frac{1}{N(n_{BV} + n_{CO})^{1/2}} \sum_{i=1}^{N} \left[ \frac{1}{\sigma_{BV}} \| BV_i(t_1, \ldots, t_{n_{BV}}) - BV_{id}(t_1, \ldots, t_{n_{BV}}) \|_2^2 + \frac{1}{\sigma_{CO}} \| CO_i(t_1, \ldots, t_{n_{CO}}) - CO_{id}(t_1, \ldots, t_{n_{CO}}) \|_2^2 \right]^{1/2}$$

(13)

We employed the Wilcoxon signed-rank test in conjunction with the Bonferroni correction to assess the statistical significance of the difference in these metrics between the models.

A numerical simulation of the physiological model was performed using MATLAB®’s ODE solvers in the Simulink® environment. The numerical optimization was performed using MATLAB®’s Optimization Toolbox. Data analysis and visualization was performed using the seaborn and the matplotlib libraries in Python.

RESULTS

Table 2 shows the output ($e_1$) and the internal state ($e_2$) prediction errors associated with the population-average model.
TABLE 2 | Output and state prediction errors associated with population-average model as well as individualized models with and without the use of regularization (mean (SD), N = 23).

| Population average | Individualized/no regularization (λ = 0) | Individualized/regularization |
|-------------------|------------------------------------------|-------------------------------|
| **Case 1: individualizing the physiological model with BV, CO, and MAP** (individualized/regularization, λ = 0.16/0.13) | | |
| BV (L) | 0.26 (0.19) | 0.10 (0.03) | 0.10 (0.04) |
| CO (L/m) | 1.31 (0.80) | 0.34 (0.18) | 0.34 (0.17) |
| MAP (mmHg) | 17.1 (8.88) | 7.23 (2.01) | 7.40 (2.05) |
| | e_1 | 0.25 (0.14) | 0.08 (0.02) \( \dagger \) | 0.08 (0.02) \( \dagger \) |
| **Case 2: individualizing the physiological model with CO and MAP** (individualized/regularization, λ = 0.16/0.13) | | |
| BV (L) | 0.26 (0.19) | 0.29 (0.21) | 0.25 (0.18) |
| CO (L/m) | 1.31 (0.80) | 0.30 (0.18) | 0.32 (0.15) |
| MAP (mmHg) | 17.1 (8.88) | 6.94 (2.20) | 7.92 (2.42) |
| | e_1 | 0.30 (0.19) | 0.09 (0.03) \( \dagger \) | 0.10 (0.03) \( \dagger \) |
| | e_2 | 0.15 (0.15) | 0.16 (0.17) | 0.14 (0.14) |
| **Case 3: individualizing the physiological model with MAP** (individualized/regularization, λ = 0.22/0.17) | | |
| BV (L) | 0.26 (0.19) | 0.43 (0.35) | 0.27 (0.18) |
| CO (L/m) | 1.31 (0.80) | 13.0 (18.0) | 1.19 (0.94) |
| MAP (mmHg) | 17.1 (8.88) | 4.85 (1.43) | 7.48 (2.03) |
| | e_1 | 0.24 (0.13) | 0.07 (0.02) \( \dagger \) | 0.10 (0.04) \( \dagger \) |
| | e_2 | 0.26 (0.15) \( * \) | 2.08 (2.83) | 0.24 (0.20) \( * \) |

\( \dagger \) \( p < 0.016 \) compared to the individualized model without the use of regularization.

\( * \) \( p < 0.016 \) compared to the population-average model.

and the individualized models (both with and without the use of regularization), while Figure 4 presents the measured versus the model-predicted output and internal state signals associated with the population-average and the individualized physiological models. Figure 5 presents the representative behaviors of the output prediction error (\( e_1 \)), the internal state prediction error (\( e_2 \)), and the parametric deviation distance (\( D \)) with respect to the regularization weight \( \lambda \) in an individual subject. Figure 6 presents the regularized parametric deviations from the population-average model with respect to varying degrees of data scarcity.

**DISCUSSION**

With the long-term goal of advancing the future development of data scarcity, the testing of clinical care automation based on physiological models, we sought to investigate the role of regularization in individualizing a physiological model using scarce data. Our findings from the case study of blood volume kinetics and cardiovascular hemodynamics in hemorrhage and circulatory resuscitation suggest that regularization can be effective in individualizing the physiological model to capture salient individual-specific characteristics and restrict unnecessary deviations from the population-average model. Furthermore, regularization results in appropriately varying parametric deviations to cope with the varying levels of data scarcity, thereby securing the physiological plausibility and the relevance of the individualized physiological model (details follow).

**Role of Regularization in Individualizing Physiological Models**

First, the use of regularization effectively maintained the goodness of fit to individual-specific data while reducing parametric deviations in insensitive directions. In all scenarios associated with the three levels of data scarcity (see section “Data Analysis”), the difference in \( e_1 \) was small between the physiological models individualized with and without the use of regularization (Table 2 and Figure 4). In contrast, the parametric deviations [in terms of the distance \( D(\theta) = J_2(\theta)/J_1 \), averaged across all the animals] were smaller when regularization was employed (1.19 in case 1, 1.14 in case 2, and 0.30 in case 3) than when it was not (1.87 in case 1, 1.89 in case 2, and 2.82 in case 3). This suggests that regularization assists in individualizing a physiological model to achieve a moderate level of goodness of fit while restricting parametric deviations that may only result in overfitting.

Second, the physiological model individualized with the use of regularization exhibited improved physiological plausibility and relevance in comparison with the one individualized without the use of regularization in terms of the accuracy in predicting internal state signals. In general, the goodness of fit associated with BV and CO responses was deteriorated when the corresponding measurement was removed in individualizing the physiological model. Figure 4 shows that indeed (i) BV prediction is deteriorated when BV data become unavailable (case 1 → case 2 and case 3) and (ii) CO prediction is likewise deteriorated when CO data become unavailable (case 2 → case 3). In case 2, \( e_2 \) (i.e., BV prediction error) was smaller when regularization was employed than when it was not employed (\( p = 0.03 \) (Table 2). In case 3, \( e_2 \) (i.e., BV and CO prediction errors) was likewise smaller when regularization was employed than when it was not employed (\( p < 0.01 \) (Table 2). Interestingly, MAP prediction was notably improved as data became more scarce when regularization was not employed (case 1 (7.23 mmHg) to case 2 (6.94 mmHg) to case 3 (4.85 mmHg)), all in terms of mean errors; see Table 2. Thus, the large deterioration in \( e_2 \) may be attributed to overfitting to MAP data. In conjunction with the finding above on the role of regularization in restricting unnecessary parametric deviations, this suggests that regularization improves the ability of the individualized physiological model to predict the internal states accurately by preventing the undesired over-tuning of model parameters away from their population-average values. One additional minor note is that BV prediction achieved with regularization was comparable between case 2 and case 3. This can be interpreted as follows: (i) both case 2 and case 3 did not use BV data in individualizing the physiological model, and regularization tended to reduce the BV component in the physiological model to the population-average model (as can be seen by the comparable BV prediction errors associated with these models; see Table 2) and (ii) presumably CO data did not provide much implications on the behavior of BV on top of the
prior knowledge [as can be seen by the BV prediction errors in case 2 with and without the use of regularization (0.25 L and 0.29 L, both in terms of mean errors; see Table 2).]

Third, the individualized physiological model exhibited goodness of fit parametric deviation behavior with respect to the regularization weight \( \lambda \) as anticipated in Figure 2, especially in the case of very scarce data (case 3; see Figure 5). A sharp initial decrease in \( D(\theta) = J_2(\theta)/\lambda \) relative to a modest increase in \( e_1 \) was observed when \( \lambda \) was initially increased from zero. Further increase in \( \lambda \) after a critical point caused \( e_1 \) to largely increase relative to a steady decrease in the parametric deviation distance and eventually approach that of the population-average model. In sum, an appropriate \( \lambda^*_i \) could be selected as outlined in section "Individualizing the Physiological Model" (\( \lambda^* = 0.15 \)).

Interestingly, there appeared to be a notable link between the behaviors of \( D(\theta) \) and \( e_2 \): \( e_2 \) initially showed a notably decreasing trend similar to that of \( D(\theta) \), which suggests that regularization improves the ability of the individualized model to predict internal states (Figure 5). For larger values of \( \lambda \), \( e_2 \) increased, which can be attributed to the restriction of sensitive parametric deviations, degenerating the individualized model to the population-average model.

Fourth, the physiological model individualized with the use of regularization was associated with a significantly superior output prediction error and a comparable internal state prediction error in comparison with the population-average model (Table 2). This suggests that regularization is effective in judiciously individualizing sensitive model parameters to achieve desirable

---

**FIGURE 4** Comparison of measured data with model outputs in all \( N = 23 \) individualized models, with and without the use of regularization, in three cases of data scarcity. Solid lines correspond to the mean response across all individuals and the shaded bands show the standard deviation of the response across all individuals. The lines on the horizontal axis indicate the timings of hemorrhage and infusion in the experimental protocol.
goodness of fit to individual-specific data while at the same time capitalizing on prior knowledge (in the form of regularization) to preserve the ability to adequately predict unmeasured internal states.

**Parametric Deviations in Relation to Varying Data Scarcity**

In our case study, regularization-induced deviations from the population-average model showed two noteworthy trends. First, with regularization, deviations tended to decrease as the data became more scarce (Figure 6). In other words, (i) regularization allowed meaningful deviations if ample data were available so that the physiological model could be better individualized by absorbing the individual-specific characteristics represented by the data, whereas (ii) it restricted unwarranted deviations if scarce data were available so that the physiological model could fall back to the population-average model. This can be viewed as a desirable behavior in that increasingly leveraging the prior knowledge (i.e., the population-average model) as the data scarcity increases is the intended effect of regularization. In contrast, such a parametric deviation trend was not observed when regularization was not used. The opposite occurred instead: scarce data resulted in more aggressive deviations from the population-average model (see the deviation distance results reported in section “Role of Regularization in Individualizing Physiological Models”), which consequently
compromised the fidelity of the individualized physiological models (e.g., its internal state prediction was deteriorated; see Table 2). Second, regularization appeared to effectively individualize those parameters relevant to the presented data. For example, the most visibly deviated parameter in case 3 was the one associated with BP regulation ($k_{TPR}$), which is reasonable in that only MAP measurements were presented to individualize the physiological model in case 3 (Figure 6C). The parameters associated with both CV function ($\gamma/BC_i$ and $\eta$) and BP regulation ($k_{TPR}$ and MAPtarget) were likewise visibly deviated in case 2, in which CO and MAP measurements were presented (Figure 6B). Finally, the parameters associated with BV were likewise largely deviated ($\alpha$, in particular) when BV measurements were presented in addition to CO and MAP data (Figure 6A). One caveat is that, despite the presence of these trends, the studied parameters belong to a complex physiological system and it may not be possible to attribute each parametric deviation completely to the availability of a specific type of measurement.

### Study Limitations

This work has limitations. First, this work investigated the role of regularization in a specific case study. Hence, the insights obtained from the case study may not be universally true in all physiological modeling problems. In particular, the legitimacy of the key assumption used to address the adverse effect of scarce data on the quality of individualized models (that the inter-individual variability observed in the population data can be represented by compressed deviations from a population-average model according to the Laplace distribution in Eq. 10) may depend on the specificity of the physiological modeling problem at hand. On the one hand, such a compressed representation may indeed be valid in many real-world physiological modeling problems with the sloppiness property (Machta et al., 2013; Transtrum et al., 2015) (which implies parameter-space compressibility). On the other hand, regularization may not prove effective if the quality of the data and the model structure are such that the majority of the model parameters are associated with sensitive deviations (where individualizing the physiological model requires deviations in all the model parameters; in this case, regularization is not needed) or insensitive deviations (where a population-average model may suffice because of negligible inter-individual variability). In any case, the validity of the “parameter-space compressibility” assumption in a specific physiological modeling application can be assessed based on model prediction errors and parameter values in the individualized physiological models: the bias introduced by the inadequacy of the assumption will manifest itself as poor goodness of fit of the individualized physiological model to the data, while a lack of compressibility will manifest itself as estimated elements of $\hat{\theta}_i$ in Eq. 8 that all deviate from $\tilde{\theta}$ in Eq. 11. Second, regularization does not guarantee, and may even prevent, the convergence of the physiological model to the “true” individual-specific physiological model. Indeed its main purpose is to minimize the deviations from the prior knowledge (i.e., population-average model). Thus, the quality of a physiological model individualized with regularization, especially using highly scarce data, hinges upon the quality of prior knowledge. Hence, the physiological model individualized with regularization is a candidate approximation to the ideal individual-specific physiological model (which would be obtained with sufficiently diverse and high-quality data obtained from an individual). Taking these aspects into account, regularization may be viewed as an effective tool to individualize a physiological model in cases with a relatively complex model and scarce individual-specific data.

### CONCLUSION

We sought to garner potentially generalizable insight into the role of regularization as prior knowledge in individualizing physiological models using scarce data. The findings obtained from a clinically significant case study illustrated that regularized individualization creates physiological models equipped with several desirable properties: (i) an adequate trade-off between goodness of fit to scarce individual-specific data and deviation from a population-average model, (ii) improved physiological plausibility and relevance, and (iii) parametric deviations relevant to the scarcity level of the data. Noting that regularization is not prevalently used in the physiological modeling domain, future work must be devoted to exploring the use of regularization in individualizing a range of physiological models and developing appropriate theory and novel structures for the regularization functions based on appropriate physiological assumptions.

### DATA AVAILABILITY STATEMENT

The datasets analyzed in this study can be obtained from Dr. George C. Kramer at the University of Texas Medical Branch (gkramer@utmb.edu).

### AUTHOR CONTRIBUTIONS

AT, CS, and J-OH formulated the procedure to individualize the physiological model via regularization and wrote and revised the manuscript. XJ, AL, CS, and J-OH developed the physiological model. AT and J-OH analyzed the data. XJ and AL reviewed the manuscript.

### FUNDING

This work was supported by the United States National Science Foundation (NSF) CAREER Award (Grant No. CNS-1748762) and the United States Office of Naval Research (ONR) (Grant No. N00014-19-1-2402).

### ACKNOWLEDGMENTS

We thank Dr. George C. Kramer at the University of Texas Medical Branch for provision of the experimental dataset for analysis as well as his physiological consultation and guidance in data analysis.
Salinas, J., Chung, K. K., Mann, E. A., Cancio, L. C., Kramer, G. C., Serio-Melvin, M. L., et al. (2011). Computerized decision support system improves fluid resuscitation following severe burns: an original study. *Crit. Care Med.* 39, 2031–2038. doi: 10.1097/CCM.0b013e31821cb790

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Commun.* 181, 259–270. doi: 10.1016/j.cpc.2009.09.018

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., and Cariboni, J. (2008). *Global Sensitivity Analysis: The Primer*. Hoboken, NJ: John Wiley & Sons.

Sobol, I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* 55, 271–280. doi: 10.1016/S0378-4754(00)00270-6

Tivay, A., Arabi, G., Dor, D., Bighamian, R., Kramer, G. C., and Hahn, J.-O. (2019). “A regularized system identification approach to subject-specific physiological modeling with limited data,” in *Proceedings of 2019 American Control Conference*, (Philadelphia, PA: American Automatic Control Council), 3468–3473.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* 6, 187–202. doi: 10.1098/rsif.2008.0172

Transtrum, M. K., Machta, B. B., Brown, K. S., Daniels, B. C., Myers, C. R., and Sethna, J. P. (2015). Perspective sloppiness and emergent theories in physics, biology, and beyond. *J. Chem. Phys.* 143,010901. doi: 10.1063/1.4923066

Transtrum, M. K., Machta, B. B., and Sethna, J. P. (2010). Why are nonlinear fits to data so challenging? *Phys. Rev. Lett.* 104,060201. doi: 10.1103/PhysRevLett.104.060201

Vaid, S. U., Shah, A., Michell, M. W., Rafie, A. D., Deyo, D. J., Prough, D. S., et al. (2006). Normotensive and hypotensive closed-loop resuscitation using 3.0% NaCl to treat multiple hemorrhages in sheep. *Crit. Care Med.* 34, 1185–1192. doi: 10.1097/01.CCM.0000207341.78696.3A

White, A., Tolman, M., Thomas, H. D., Withers, H. R., Mason, K. A., and Transtrum, M. K. (2016). The limitations of model-based experimental design and parameter estimation in sloppy systems. *PLoS Comput. Biol.* 12:e1005227. doi: 10.1371/journal.pcbi.1005227

Zhang, L., Jiang, Z., Choi, J., Lim, C. Y., Maiti, T., and Baek, S. (2019). Patient-specific prediction of abdominal aortic aneurysm expansion using bayesian calibration. *IEEE J. Biomed. Heal. Informatics* 23, 2537–2550. doi: 10.1109/JBHI.2019.2896034

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
APPENDIX

TABLE A1 | List of lumped parameters individualized in the physiological model with nominal upper bounds, lower bounds, and population-average values.

| Model parameter | Unit       | Lower bound ($\tilde{\theta}_m$) | Upper bound ($\tilde{\theta}_m$) | Population-average value ($\bar{\theta}$) |
|-----------------|------------|----------------------------------|----------------------------------|------------------------------------------|
| $\alpha_u$      | (1)        | 0                                | 6                                | 4.29                                     |
| $\alpha_v$      | (1)        | 0                                | 6                                | 2.28                                     |
| $K_p$           | (1/min)    | $10^{-2}$                        | 0.5                              | 0.21                                     |
| $A$             | (1/L)      | 0                                | 0.3                              | 0.23                                     |
| $A/E_s$         | (1/mmHg)   | 0                                | $10^{-3}$                        | $2.54 \times 10^{-4}$                    |
| $\gamma_k/B$    | (1/mmHg)   | 0                                | $10^{-4}$                        | $3.62 \times 10^{-5}$                    |
| $\gamma/BC_s$   | (1/L)      | 0                                | 0.1                              | $2.43 \times 10^{-2}$                    |
| $\eta$          | (1)        | 0.2                              | 1                                | 0.43                                     |
| $kTPR$          | (1/L)      | 0                                | 1                                | $5.87 \times 10^{-2}$                    |
| $pTPR$          | (1/min)    | 0                                | 1                                | 0.66                                     |
| MAP$_{target}$  | (mmHg)     | 40                               | 100                              | 70.1                                     |