Robust Nonlinear State Space Model Identification for Hemorrhage Resuscitation

Elham Estiri, *Student Member, IEEE* and Hossein Mirinejad, *Senior Member, IEEE*

***Abstract*— This paper presents a novel method for predicting hemodynamic responses in hemorrhage resuscitation. The proposed approach, namely, robust nonlinear state space modeling (RNSSM), aims to overcome challenges of identifying reliable models using limited and noisy critical care data by innovatively integrating autoencoder learning and variational Gaussian inference in a unified framework. Simulation results demonstrate the initial feasibility and performance evidence of the RNSSM approach as a digital twin of an animal study in hemorrhage resuscitation scenarios.**

***Clinical Relevance*— Enabling reliable, personalized hemodynamic models amenable to the closed-loop control design can potentially lead to development of efficient model-informed precision dosing strategies, promoting patient safety and outcomes in critical care.**

I. Introduction

Fluid resuscitation is a medical intervention commonly used in hypovolemic scenarios to compensate for the lost blood volume and stabilize critically ill patients. Fluid management is currently ad-hoc and dependent on the physician's style and expertise [1-5]. These ad-hoc protocols lack the capability to accurately adjust fluid infusion dosages due to their empiric nature, especially in the presence of clinical disturbances [6], posing a significant risk of adverse effects such as under- and over-dosing. Consequently, treatment performance is compromised due to the absence of appropriate dosage adjustment tools.

A few studies have considered the modeling of hemodynamic responses in fluid resuscitation therapies [1-3]. In [1], a simplified lumped-parameter model replicating blood volume responses in different physiological states following fluid infusion was presented. In [2], a control-theoretic, physic-based model of hemodynamic variables, including blood volume, cardiac output, and blood pressure response, in hemorrhage resuscitation was presented. In [3], a cyber-physical fluid resuscitation test bed that included hemodynamic responses to blood volume perturbations was developed.

Automated modeling and control methodologies have recently received great attention in physiological modeling and dosage adjustment [7-12]. To design a model-informed dose adjustment tool, an appropriate dose-response model is needed, and the success of the control approach is highly dependent on the availability of a reliable model. Such a model must (i) be simple enough to be amenable to the design of the controller using limited clinical data and (ii) make use of limited, noise-distorted, patient-specific clinical data with unknown baseline conditions. Development of a reliable, personalized hemodynamic model amenable to the closed-loop control design remains challenging due to the limited, distorted clinical data, large inter- and intra- patient variability, and the complexity of modeling physiological variables, and this work intends to address it in the context of fluid resuscitation.

Machine learning algorithms have been recently used for dose-response modeling [13-15]. In [13], a machine learning algorithm was used to predict the dose-adjusted concentrations of lamotrigine based on noninvasive clinical parameters. In [14], a method that combines model predictive control and reinforcement learning (RL) was presented to address the challenge of drug administration variability in the treatment of renal anemia. A complete literature review exploring the use of data mining and machine learning techniques for disease prediction using complete blood count data was presented in [15]. Most of these models predominantly rely on population-based data, limiting their applicability to individual subjects. In our previous study [16], we designed an individual-based fluid dosing algorithm using a model-free RL. While this approach showed promising results in fluid management, it required a substantial amount of data for training and provided an inferior performance in the presence of clinical disturbances.

To address the aforementioned challenges, this paper presents a novel modeling framework namely, robust nonlinear state space modeling (RNSSM), for predicting hemodynamic responses in hemorrhage resuscitation. The proposed approach integrates autoencoder learning and variational Gaussian inference (VGI) into a unified framework to develop nonlinear state space models that are highly amenable to the closed-loop control design from limited, noisy critical care data. The goal is to develop subject-specific models that can reliably predict mean arterial pressure (MAP) responses to fluid infusion in hemorrhage scenarios. The RNSSM approach improves (1) model accuracy by considering subject-specific characteristics and drug attributes and (2) model reliability by accounting for uncertainties inherently present in clinical data.

The rest of the paper is organized as follows: Section II describes the proposed methodology leveraging autoencoder learning and Gaussian inference for developing RNSSM models. Section III presents the results and discussions, and Section IV draws the conclusions.

* This material is based upon work supported by the National Science Foundation under Grant No. 2138929.

E. Estiri and H. Mirinejad are with the College of Aeronautics and Engineering, Kent State University, Kent, OH 44242 USA (correspondence phone: 330-672-7403; e-mail: hmiri@kent.edu).

II. Material and methods

We derive a new methodology to predict robust, individualized MAP responses to fluid infusion in hemorrhage scenarios. The proposed methodology focuses on identifying reliable nonlinear state space models from limited, noisy clinical data using machine learning algorithms. Consider a multiple-input/multiple-output nonlinear state space model (NSSM) in a general form:

$$x_{k+1} = f(x_k, u_k, \theta) + v_k$$
$$y_k = g(x_k, u_k, \theta) + \omega_k \quad (1)$$

where x_k represents the hidden state variable in $\mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ denotes the observed input, and $y_k \in \mathbb{R}^{n_y}$ represents the measured output. The functions $f(.)$ and $g(.)$ capture the state transition and output measurement, respectively, while $\theta \in \mathbb{R}^{n_\theta}$ represents a vector of unknown parameters. The terms v_k and ω_k account for disturbance and measurement noise, respectively, both described by Gaussian probability density functions.

We aim to enable reliable, individualized prediction of hemodynamics by amalgamating autoencoder learning with VGI techniques.

A. Autoencoder Learning of Nonlinear State Space Models

Autoencoder is a type of artificial neural network (ANN) used for representation learning [17]. It learns how to efficiently compress and encode data then learns how to reconstruct the data back from the reduced representation to a representation that is as close to the original input as possible. This is accomplished through a bottleneck in the ANN forcing a compressed knowledge representation of the original input. The use of autoencoders for learning state space models has received great attention recently [18-20] due to their capability to impose regular geometry on the learned latent space— an abstract space that positions similar samples close to each other. Here, we aim to capture the intricate dynamics of MAP responses to fluid infusion changes using autoencoder learning. The structure was adopted from [20] where the autoencoder learns a nonlinear state space representation from a given subject-specific input/output dataset. Suppose we are given a dataset of input/output $\boldsymbol{I}_k = \{u_1, u_2, \dots, u_N, y_1, y_2, \dots, y_N\}$, where $\boldsymbol{u}_k \in \mathbb{R}^{n_u}$ is the vector of inputs (fluid dosages) and $\boldsymbol{y}_k \in \mathbb{R}^{n_y}$ is the vector of measured outputs (MAP responses). The objective is to find optimal values for functions $e: \mathbb{R}^{n_I} \to \mathbb{R}^{n_x}$, $f: \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$, and $g: \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ by minimizing the following fitting criterion:

$$\min_{e,f,g} \mathcal{L}(e, f, g, Z) = \min_{e,f,g} \sum_{k=k_0}^{N} L(\hat{\boldsymbol{y}}_k, \boldsymbol{y}_k) \quad (2)$$

where e, f, g, are the functions describing the encoder, bridge, and decoder models, Z describes the training dataset, $L: \mathbb{R}^{2n_y} \to \mathbb{R}$ is the loss function, and $\boldsymbol{y}$ and $\hat{\boldsymbol{y}}_k$ are the measured and predicted outputs, respectively.

To determine the appropriate nonlinear state space model that yields an acceptable mismatch between the predicted value, $\hat{\boldsymbol{y}}_k$, and the measured value, $\boldsymbol{y}_k$, we need to design a suitable ANN architecture for training these functions.

The autoencoder model, used in this work, consists of three main components: (1) A multilayer ANN encoder for predicting $\boldsymbol{x}_k$ from $\boldsymbol{I}_{k-1}$; (2) A multilayer ANN decoder for predicting $\boldsymbol{y}_k$ from $\boldsymbol{x}_k$; and (3) A bridge network, also a multilayer ANN model, for modeling the function f that maps $\boldsymbol{x}_k$ to $\boldsymbol{x}_{k+1}$. Since direct access to the internal dynamics of the system to obtain $\boldsymbol{x}_{k+1}$ is unavailable, a second autoencoder is defined that simultaneously maps $\boldsymbol{I}_k$ to $\boldsymbol{x}_{k+1}$ and $\boldsymbol{y}_{k+1}$ using the same weights as the first autoencoder. The following criterion was chosen to train the model:

$$\min_{e,f,d} \Sigma_{k=k_0}^{N-1} L_1(\hat{\boldsymbol{y}}_k, \boldsymbol{y}_k) + L_1(\hat{\boldsymbol{y}}_{k+1}, \boldsymbol{y}_{k+1}) + L_2(\boldsymbol{x}_k, \boldsymbol{x}_{k+1}) + L_3(\boldsymbol{y}_{k+1}, \tilde{\boldsymbol{y}}_{k+1}) \quad (3)$$

where $\boldsymbol{y}_k = [y_1, y_2, \dots, y_k]$ and $\boldsymbol{y}_{k+1} = [y_2, y_3, \dots, y_{k+1}]$ are the vectors of measured outputs, and $\hat{\boldsymbol{y}}_k = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_k]$, $\hat{\boldsymbol{y}}_{k+1} = [\hat{y}_2, \hat{y}_3, \dots, \hat{y}_{k+1}]$, and $\tilde{\boldsymbol{y}}_{k+1} = [\tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_{k+1}]$ are the predicted outputs. L_1 and L_2 are the loss functions for training the autoencoders and the bridge network, respectively. L_3 is introduced to prevent the error introduced by the bridge network f from being amplified by the decoder.

B. Variational Gaussian Inference (VGI)

State space models based purely on autoencoders, as identified in part A, fail to account for external uncertainties (e.g., measurement noise), as well as internal sources (e.g., unmodeled dynamics) [21]. This becomes particularly problematic in fields where data is inherently distorted. For instance, clinical data is often tainted by various factors, including measurement noises. To tackle this challenge, we are integrating VGI techniques into autoencoder learning, aiming to enhance the robustness of the identified models. This integration enables us to determine underlying model parameters, factoring in the uncertainties inherent in training data—referred to as aleatoric uncertainty [22]. Specifically, in this context, it relates to the uncertainties observed in fluid infusion doses and their corresponding hemodynamic responses. Such variability can arise from physiological factors, measurement inaccuracies, and other stochastic elements [9].

VGI enhances the performance of a standard autoencoder by introducing a probabilistic interpretation. Instead of encoding input data into a *fixed* latent space representation, the encoder maps the data onto a *probability distribution* over the latent space, typically assumed to be a multivariate Gaussian function. The decoder reconstructs the data by sampling from this latent distribution. This integrated system, often referred to as a variational autoencoder (VAE), leverages the benefits of a generative model. The enforced probability distribution in the VAE's latent space allows for efficient sampling and data point generation. Such a generative feature is especially valuable for dynamic state space modeling in time-series data. As the model produces results based on ever-changing inputs, this methodology becomes versatile for tasks demanding accurate reconstruction and generation within a dynamic data landscape.

In the VAE, the focus is on approximating the true posterior distribution $p_\theta(x|y)$, where x denotes latent variables and y is the observed data. Computing the true posterior $p_\theta(x|y)$ is analytically intractable, prompting the introduction of a variational inference that approximates the

posterior using a simpler variational distribution $q_\phi(x|y)$, parameterized by $\phi = (\mu, \sigma)$. Here, μ and σ denote the mean and standard deviation of the distribution, respectively. Their values are typically set to establish the prior distribution as a standard normal distribution, i.e., $\mu = 0$ and $\sigma = 1$. The training goal for a VAE is to determine model parameters making the variational distribution $q_\phi(x|y)$ closely match the true posterior distribution $p_\theta(x|y)$. This is achieved by minimizing the Kullback-Leibler (KL) divergence between the two distributions, defined as [23]:

$$KL_{Divergence} = D_{KL}(q_\phi(x|y)\big|p_\theta(x|y)\big) = -\frac{1}{2}\Sigma_{N=1}^{D}(1 + \log\sigma^2 - \mu^2 - \sigma^2). \quad (4)$$

The model's total loss function is derived as:

$$Total\ loss = L_1(\hat{y}_k, y_k) + L_1(\hat{y}_{k+1}, y_{k+1}) + L_1(\tilde{y}_{k+1}, y_{k+1}) + L_2(\tilde{x}_{k+1}, x_{k+1}), \quad (5)$$

with $L_2(x, \hat{x}) = ||x - \hat{x}||_2$, and L_1 defined as:

$$L_1 = \big|\big|x - \hat{x}\big|\big|_2 - \frac{1}{2}\Sigma_{N=1}^{D}(1 + \log\sigma^2 - \mu^2 - \sigma^2). \quad (6)$$

Here, L_1 comprises two parts: a reconstruction term and a KL divergence term. The former measures the quality of reconstructed data, while the latter ensures the learned latent space aligns with the assumed prior distribution.

The resulting algorithm, integrating VGI and autoencoder learning for hemodynamics prediction, is named the RNSSM framework, as detailed in Fig. 1.

III. Results & Discussion

The proposed RNSSM approach leverages autoencoder learning and VGI within a unified framework to develop robust nonlinear state space models that are highly amenable to closed-loop control design. The dataset was sourced from an animal study conducted at the Resuscitation Research Laboratory, University of Texas Medical Branch [24], where different sheep underwent high and medium hemorrhage procedures accompanied by fluid infusion. The data collection process followed the study protocol approved by the Institutional Animal Care and Use Committee [24].

In the animal study, a hemorrhage rate of 25 ml/kg was administered to the subjects within the first 15 minutes, after which it was halted. At times t = 52 and t = 72 minutes, two smaller hemorrhage rates of 5 ml/kg were applied to each subject for a duration of 2 minutes. Fluid resuscitation using Lactated Ringer's solution began 30 minutes after the start of study. MAP measurements were recorded every 5 minutes for a duration of 180 minutes. The input data for the model consisted of fluid infusion and hemorrhage rates, and the output data was the corresponding MAP values.

The RNSSM training dataset consisted of N=1441 samples. An early-stopping strategy was employed using 10% of the training dataset to verify the stopping criterion. The dataset used for cross-validation comprised 131 samples. All three modules of the model (encoder, decoder, and bridge) were designed with two hidden layers, each consisting of 30 neurons. The ANNs consisted of 3651 weights, evenly distributed among the modules. Simulations were conducted on data from six animal subjects.

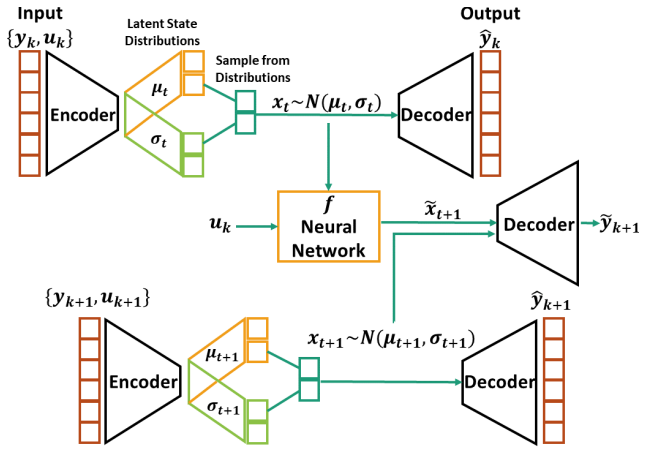

Figure 1. Robust nonlinear state space modeling (RNSSM) framework integrating autoencoder learning and variational Gaussian inference for identification of hemodynamics from limited, noise-distorted clinical data. The framework leverages two distinct VAEs: The first VAE is focused on estimating y_k, and the second VAE is dedicated to approximating x_{k+1} through the state space transition function f. The estimation of f is routed through the same decoder as the first VAE, ensuring that the inferred x_{k+1} results in reliable reconstructions. Furthermore, the reconstructed $\tilde{x}_{k+1}$ is also processed through the same decoder to mitigate the error amplification when data navigate through the f neural network.

Fig. 2a displays the fluid and hemorrhage rates (inputs) during the animal study, while Fig. 2b illustrates the predicted MAP responses from the RNSSM model alongside the measured MAP for a sample subject. Fig. 2 demonstrates the model's capability to track the real trend of time-series data and effectively capture MAP fluctuations caused by hemorrhage, highlighting the model's robustness against external disturbances.

Additionally, Table I presents the performance metrics, including root mean square error (RMSE), mean absolute error (MAE), and median absolute percentage error (MDAPE) for all subjects. These metrics provide quantitative insights into the model's accuracy. The outcomes highlight the model's capability to robustly and accurately capture MAP responses throughout the hemorrhage resuscitation, bolstering the case for continued exploration and enhancement of the RNSSM framework. The presented method addresses several key challenges in the field. It effectively tackles the issues caused by the noise and external disturbances, as well as the limited data availability. Also, the proposed method has border applicability in other fields such as robotics and manufacturing, where data availability is limited, and measurements are prone to noise.

It's essential, moving forward, to present a comparison study against other resuscitation models. Additionally, a deeper probe into the robustness of the RNSSM model, especially in the face of uncertainties, is needed in the near future. The formulation of a closed-loop controller for fluid management, leveraging the RNSSM models, also presents an intriguing avenue for future research.

TABLE I. Performance Metrics for All Subjects

	RMSE (%)	MAE (%)	MDAPE (%)
MEAN	0.37	0.28	0.29
STD	0.013	0.011	0.1

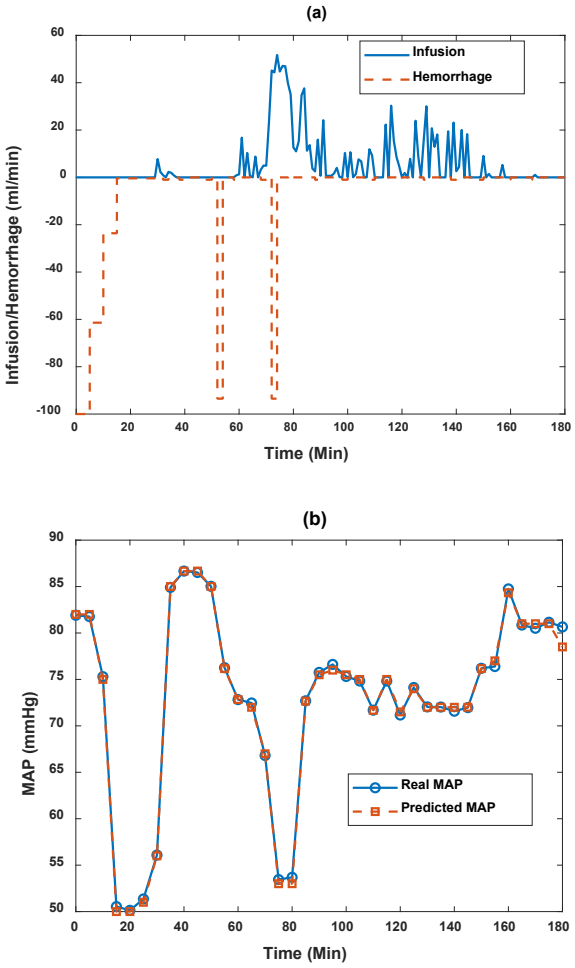

Figure 2. (a) Fluid infusion and hemorrhage rates during the animal study, and (b) measuered and predicted MAP responses from the RNSSM method.

IV. Conclusion

RNSSM, a novel modeling framework for predicting hemodynamic responses in hemorrhage resuscitation was presented. The approach combined autoencoder learning and VGI to overcome the challenges of identifying reliable models using limited and noisy critical care data. Simulation results were highly promising, encouraging further investigation of the RNSSM approach against state-of-the-art digital twin models in the near future.

References

[1] R. Bighamian, A. T. Reisner, and J. Hahn, "A lumped-parameter subject-specific model of blood volume response to fluid infusion," *Front. Physiology*, vol. 7, p. 390, 2016.

[2] R. Bighamian, B. Parvinian, C. G. Scully, G. C. Kramer, and J. O. Hahn, "Control-oriented physiological modeling of hemodynamic responses to blood volume perturbation," *Control Eng. Pract*., vol. 73, pp. 149-160, 2018.

[3] H. Mirinejad, B. Parvinian, M. Ricks, Y. Zhang, S. Weininger, J. O. Hahn, and C. G. Scully, "Evaluation of fluid resuscitation control algorithms via a hardware-in-the-loop test bed," *IEEE Trans. Biomed. Eng*., vol. 67, no. 2, pp. 471-481, Feb. 2020.

[4] E. Estiri and H. Mirinejad, "Autoencoder learning and variational Gaussian inference for predicting mean arterial pressure in fluid resuscitation", IEEE Conf. Systems, Man, Cybern. (SMC'23), Honolulu, HI, Oct. 2023.

[5] J. Rinehart, B. Alexander, Y. Le Manach, C. Hofer, B. Tavernier, Z. N. Kain, and M. Cannesson, "Evaluation of a novel closed-loop fluid-administration system based on dynamic predictors of fluid responsiveness: an in silico simulation study," *Crit. Care*, vol. 15, no. 6, p. R278, 2011.

[6] D. De Backer and M. R. Pinsky, "Can one predict fluid responsiveness in spontaneously breathing patients?" *Intensive Care Med*., vol. 33, no. 7, pp. 1111–1113, 2007.

[7] H. Mirinejad, A. E. Gaweda, M. E. Brier, J. M. Zurada, T. Inanc, "Indicidualized drug dosing using EBF_Galerkin method: Case of anemia management in chronic kidney disease," *Computer Methods and Programs in Biomedicine,* vol. 148, pp. 45-53, 2017.

[8] P. E. Marik, M. Baram, and B. Vahid, "Does central venous pressure predict fluid responsiveness? A systematic review of the literature and the tale of seven mares," *Chest*, vol. 134, no. 1, pp. 172–178, 2008.

[9] G. C. Kramer, M. P. Kinsky, D. S. Prough, J. Salinas, J. L. Sondeen, M. L. Hazel-Scerbo, and C. E. Mitchell, "Closed-loop control of fluid therapy for treatment of hypovolemia," *J. Trauma*, vol. 64, no. 4, pp. S333–S341, 2008.

[10] H. Mirinejad and T. Inanc, "Individualized anemia managemenet using a radial basis function method," in *Proc. IEEE Great Lakes Biomedical Conf. (GLBC)*, May 2015.

[11] J. M. Laínez-Aguirre and G. V. Reklaitis, "A stochastic optimization approach for the design of individualized dosage regimens," *AIChE J*., vol. 59, no. 9, pp. 3296-3307, 2013.

[12] H. Mirinejad and T Inanc, "RBF method for optimal control of drug administration in the anemia of hemodialysis patients," in *Proc. 41st Northeast Bioeng. Conf. (NEBEC)*, Apr. 2015.

[13] X. Zhu, W. Huang, H. Lu, Z. Wang, X. Ni, J. Hu, S. Deng, Y, Tan, L. Li, M. Zhang, and C. Qiu, "A machine learning approach to personalized dose adjustment of lamotrigine using noninvasive clinical parameters," *Scientific Reports*, vol. 11, p. 5568, 2021.

[14] A. E. Gaweda, M. K. Muezzinoglu, A. A. Jacobs, G. R. Aronoff, and M. E. Brier, "Model predictive control with reinforcement learning for drug delivery in renal anemia management," in *Proc. IEEE* Intern. Conf. of Eng. Med. Biol., pp. 5177-5180, 2006.

[15] M. Jaiswal, A. Srivastava, and T. J. Siddiqui, "Machine learning algorithms for Anemia Disease prediction," *Recent trends in Communication, Computing, and Electronics*, vol. 534, pp. 463-469, 2018.

[16] E. Estiri and H. Mirinejad, "Precision dosing in critical care: application of machine learning in fluid therapy," in *Proc. IEEE Intern. Conf. Digital Health (ICDH)*, Chicago, IL, Jul. 2023.

[17] C. C. Tan and C. Eswaran, *Autoencoder Neural networks,* Lambert, 1st ed., 2010.

[18] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Found. Trends Machine Learn.*, vol. 12, no. 4, 2019.

[19] S. E. Otto and C. Rowley, "Linearly recurrent autoencoder networks for learning dynamics," *SIAM J. Appl. Dynam. Sys*., vol. 18, no. 1, pp. 558-593, 2019.

[20] D. Masti and A. Bemporad, "Learning nonlinear state-space models using deep autoencoders," *Automatica*, vol. 129, 2021.

[21] N. Stahl, G. Falkman, A. Karlsson, and G. Mathiason, "Evaluation of uncertainty quantification in deep learning," *Information Proc. Management Uncertainty Knowledge-Based Sys*., pp. 556-568, May 2020.

[22] E. Daxberger and J. M. Hernandez-Lobato, "Bayesian variational autoencoders for unsupervised out-of-distribution detection," *arXiv preprint arXiv: 1912.05651,* 2019.

[23] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," in *Proc. IEEE Intern. Conf. Acoustics, Speech Signal Proc. (ICASSP'07)*, 2007.

[24] A. D. Rafie, P. A. Rath, M. W. Michell, R. A. Kirschner, D. J. Deyo, D. S. Prough, J. J. Grady, and G. C. Kramer, "Hypotensive resuscitation of multiple hemorrhages using crystalloid and colloids," *Shock*., vol. 22, pp. 262-269, Sep. 2004.