Biomedical Signal Processing and Control 62 (2020) 102126

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Unsupervised anomaly detection in peripheral venous pressure signals with hidden Markov models

Md Abul Hayat [a], Jingxian Wu [a,*], Patrick C. Bonasso [b], Kevin W. Sexton [b], Hanna K. Jensen [b,c], Melvin S. Dassinger [b], Morten O. Jensen [c]

[a] Department of Electrical Engineering, University of Arkansas, AR 72701, USA
[b] Department of Surgery, University of Arkansas for Medical Sciences, AR 72205, USA
[c] Department of Biomedical Engineering, University of Arkansas, AR 72701, USA

ARTICLE INFO

Keywords:
Peripheral venous pressure (PVP)
Anomaly detection
Motion and noise artifacts (MNA)
Hidden Markov model (HMM)
State Space Model (SSM)
Dynamic linear model (DLM)

ABSTRACT

This paper proposes an automatic anomaly detection and removal algorithm for peripheral venous pressure (PVP) signals, which can be used to predict intravascular volume loss in humans. PVP signal collection is a minimally invasive procedure that can be performed by using a standard peripheral intravenous (PIV) catheter and a commercial pressure-monitoring transducer. PVP signals are highly susceptible to motion and noise artifacts such as patient movements or unintended manipulation of PIV lines. Anomalies in PVP signals can corrupt useful information and seriously affect the integrity of PVP signal analysis. We propose to detect and remove such anomalies by exploiting the properties of PVP signals. Specifically, a dynamic linear model (DLM) with a Kalman filter is used to track and predict the time-domain evolution of PVP signals. The prediction residuals of the Kalman filter are then modeled with a hidden Markov model (HMM), with the normal and anomalous status of the signal modeled by using binary states of a hidden Markov chain. The HMM parameters along with the hidden states are iteratively estimated by using an unsupervised learning algorithm with a modified Baum–Welch method. The anomaly detection algorithm is applied to clinical data from a cohort of 24 pediatric patients with hypertrophic pyloric stenosis. Experimental results demonstrate that the proposed unsupervised anomaly detection algorithm can efficiently remove anomalies in PVP signals without the need of a training phase. The algorithm can also be applied to other time series signals, such as Electrocardiography (ECG) and Photoplethysmogram (PPG) signals.

1. Introduction

Dehydration or loss of intravascular blood volume is a common and potentially life-threating condition. Dehydration affects 30 million children annually and accounts for 400,000 pediatric emergency room visits in the United States [1,2]. Often severe dehydration (hypovolemia) or occult bleeding is not recognized by monitoring vital signs until the stage where end-organ damage occurs [3,4]. However, there is no standardized measurement for intravascular volume in adults or children. This necessitates the development of technologies that would accurately assess the volume status of a patient to guide resuscitation and treatment.

Analysis of peripheral venous pressure (PVP) waveforms is a novel method of monitoring intravascular volume and may provide valuable information in earlier detection of volume depletion [5,6]. Previous studies have demonstrated that PVP signals strongly correlate with central venous pressure under different conditions [7–11]. The shapes of PVP waveforms could be affected by a number of physiological phenomena, such as respiration, heart rate (HR), and/or systolic blood pressure (SBP). With porcine models, it has been shown in [5] that PVP signals can be a good indicator of intravascular volume depletion due to hemorrhage. Similarly, the HR frequency components of PVP signals have been used to quantify blood loss with porcine models in [12].

Pilot studies have shown evidence that PVP signal can be used for early hemorrhage detection in humans [13]. It has been demonstrated in previous studies that PVP signals can be used to predict dehydration in pediatric patients [6,14,15]. PVP signal collection is a minimally invasive technology consisting of a standard peripheral intravenous (PIV) catheter and a commercial pressure-monitoring transducer. However, PVP signals are highly susceptible to non-stationary noise, such as

* Corresponding author.
E-mail addresses: mahayat@uark.edu (M.A. Hayat), wuj@uark.edu (J. Wu), pcbonasso@uams.edu (P.C. Bonasso), kevin.sexton@uams.edu (K.W. Sexton), hkjensen@uark.edu (H.K. Jensen), dassingermelvins@uams.edu (M.S. Dassinger), mojensen@uark.edu (M.O. Jensen).

https://doi.org/10.1016/j.bspc.2020.102126
Received 6 November 2019; Received in revised form 23 July 2020; Accepted 1 August 2020
Available online 22 August 2020

random patient movements or unintended physical manipulation of the PIV catheter during the measurement process. The PVP signals will be corrupted under such events, which results in anomalies in the collected data. The presence of anomalies will seriously undermine the integrity of the analysis results. Thus it is necessary to detect and remove such anomalies before any further processing.

Anomaly detection has a wide range of applications, and it is critical for biomedical signal processing [16]. One of the main sources of anomalies in time series biomedical signals are motion and noise artifacts (MNA), which impose serious challenges in the accurate collection, modeling, and analysis of signals such as PVP [14], Electrocardiography (ECG) [17], and Photoplethysmography (PPG) [18]. Anomalies in PVP signals are detected and removed manually through visual inspections by clinical staffs in [6] and [15], where sections of PVP signals with large variations are removed before further processing. A simple automatic anomaly detection method is adopted in [14], where signals with excessively large variances are labeled as anomalies. An automatic MNA detection method for ECG signals is developed in [17] by using empirical mode decomposition, and the cleaned signals are used to detect diseases like atrial fibrillation (AF). An adaptive step-size least mean squares (AS-LMS) filter is used to reduce motion artifacts in PPG signals in [18]. Anomaly detections in time series can also be performed by using deep learning algorithms, such as recurrent neural networks with long short-term memory (LSTM) [19,20].

The objective of this paper is to develop an efficient and unsupervised anomaly detection algorithm by exploiting the features of PVP signals. To model the correlation of consecutive samples of the PVP signal, we propose to use a dynamic linear model (DLM) [21] with Kalman filter to track and predict the evolution of the PVP signal in the time domain. It is expected that the model residual, which is the difference between the actual signal and the one predicted by the Kalman filter, is a stationary random process with small variances for normal PVP signals. On the other hand, since anomalies or bad data violate the normal dynamics of PVP signals, the DLM residual will be large for anomalies. Motivated by this fact, we propose to detect and remove PVP anomalies by analyzing the differences in statistical behaviors of Kalman filter residuals.

Specifically, the Kalman filter prediction residuals are modeled by using a hidden Markov model (HMM). HMM is a powerful statistical tool that has been widely used in many different disciplines for various applications [22]. For example, it has been used for protein structure prediction [23], genome sequencing [24], speech processing [22,25], network intrusion detection [26], traffic sensing [27], etc. In this paper, the status of the PVP data is modeled by using a Markov chain with two hidden states: normal or anomalous states. The distributions of the prediction residuals depend on the two hidden states of the PVP signals. The unknown parameters of the statistical model are iteratively learned by using a modified Baum–Welch algorithm [28]. The original Baum–Welch algorithm [28] suffers from mathematical underflow due to limited computer precision. A variable scaling approach [22] is adopted in this paper to ensure the numerical stability of the iterative algorithm. In addition, we provide detailed explanations on physical meanings of the scaled probability variables which to our understanding are not available in the literature. The results of the modified Baum–Welch algorithm are used to detect and remove anomalies in PVP signals. The algorithm has been applied to clinical data from a cohort of 24 pediatric patients. Experiment results demonstrate that the algorithm can efficiently remove anomalies in PVP signals without a training phase. The proposed anomaly detection algorithm can also be applied to other time-series biomedical signals, such as ECG and PPG signals.

The remainder of this paper is organized as follows. The data acquisition process is described in Section 2. Section 3 presents the representation of PVP signals with the DLM and HMM. The anomaly detection is performed with a modified Baum–Welch algorithm, which is also described. Experimental results are presented in Section 4, and Section 5 concludes the paper.

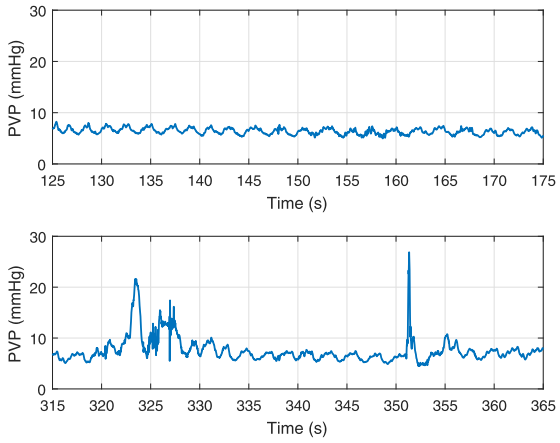

Fig. 1. Exemplary PVP signal (a) without anomaly and (b) with anomaly.

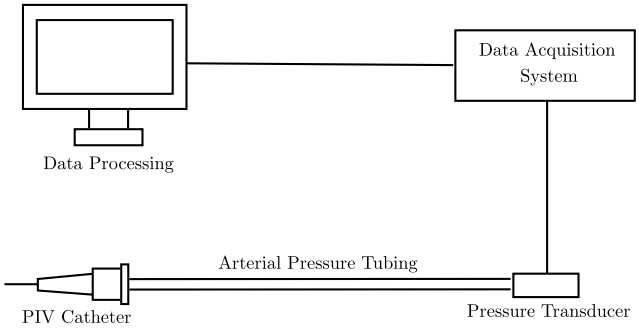

Fig. 2. Schematic diagram of the data acquisition system. Peripheral intravenous (PIV) catheter is inserted into peripheral vein of a patient. Arterial pressure tubing connects the PIV catheter to pressure transducer that converts pressure to equivalent electrical signal. This signal is acquired and processed afterwards.

2. Methods

PVP waveform data was collected in a closed system via a 22-gauge Insyte-N Autoguard PIV catheter (Becton Dickinson Infusion Therapy Systems, Sandy, Utah, USA) connected to a 48-inch arterial pressuring tubing (Smiths Medical, Dublin, OH, USA). The arterial pressure tubing was connected to a Deltran II pressure transducer (ADInstruments, Colorado Springs, CO, USA) interfaced with a PowerLab data acquisition system (ADInstruments) (see Fig. 2). A similar data collection procedure was used in [14] and [29]. Primary data manipulation and visualization have been done using LabChart v8 (ADInstruments) software of PowerLab. Acquired data was saved in .adicht and .mat formats.

Patients were defined as euvolemic or hydrated when the concentration of serum chloride was $\geq$ 100 mmol/L or bicarbonate $<$ 30 mmol/L and hypovolemic or dehydrated when serum chloride $<$ 100 mmol/L or bicarbonate $\geq$ 30 mmol/L. Data were collected on emergency department admission of patients suffering from Hypertropic pyloric stenosis.

The study was approved by the University of Arkansas for Medical Sciences (UAMS) Institutional Review Board (IRB) (Protocols 206193 & 207085). All procedures performed on the study participants were in accordance with the ethical standards of the IRB and with the 1964 Helsinki declaration and its later amendments. Informed consent was obtained from the legal guardians of the children participating in the study.

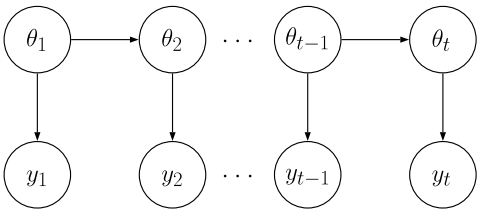

Fig. 3. Dependence structure of dynamic linear model. Here, θ_i's are forming a first-order Markov chain. Also, θ_{i+1} and y_i follow a Gaussian distribution depending on θ_i under a linear relationship. $\{\theta_i\}$ and $\{y_i\}$ are continuous random variables.

3. Theory

In this section, the PVP signals are represented and modeled by using dynamic linear models (DLM) [21]. The results will be used to detect and remove anomalies in PVP signals. The PVP signals are collected by using PIV catheters from a peripheral part of the body [14]. As a result, the signal is highly susceptible to non-stationary noise caused by events such as the random movement of the patients or an unintended touch of the catheter during the measurement process. The PVP signals will be corrupted under such events, which result in anomalies in the collected data. Fig. 1 shows a normal PVP signal and a corrupted PVP signal. It is apparent that the statistical behaviors of the PVP signals with anomalies are considerably different from its normal counterpart. The presence of anomalies will seriously undermine the integrity of the analysis results. Thus it is necessary to detect and remove such anomalies before any further processing.

We propose to capture the dynamic behaviors of PVP signals by using a DLM [21], which is a special case of the Gaussian state space model and commonly used for time series modeling and forecast. The DLM can be used in combination with the Kalman filter to track and predict the dynamic evolution of the PVP signal in the time domain. Under normal conditions, it is expected that Kalman filter with the DLM can accurately predict normal PVP signal values, and this will result in very small prediction residuals (prediction errors). On the other hand, abnormal signals will lead to large prediction residuals due to the unpredicted nature of abnormal events. Thus anomalies can be detected by analyzing the statistical properties of prediction residuals of Kalman filter.

3.1. Dynamic linear model with Kalman filter

The discrete-time PVP signal can be modeled as a time series y_t, for $t = 1, 2, \ldots, T$. Define $\mathbf{y}_{1:T} = [y_1, y_2, \ldots, y_T]$ to simplify notation. In this DLM, it is assumed that each sample of the PVP signal, y_t, is associated with a time-varying state, θ_t, as shown in Fig. 3. The DLM can be represented as [21]

$$y_t = F\theta_t + v_t; \tag{1a}$$

$$\theta_t = G\theta_{t-1} + w_t. \tag{1b}$$

where (1a) is the observation model describing the relationship between state variable θ_t and observation variable y_t, with v_t being the observation noise, and (1b) is the state transition model with w_t being model uncertainties. Here, the coefficients F and G are constants for a given patient, and it is assumed that both v_t and w_t are independent zero-mean Gaussian distributed with variances σ_v^2 and σ_w^2, respectively, that is $v_t \sim \mathcal{N}(0, \sigma_v^2)$, $w_t \sim \mathcal{N}(0, \sigma_w^2)$ and $v_t \perp w_t$.

With the DLM given in (1), we can apply the Kalman filter to iteratively predict the values of the PVP signal y_t and the state variable θ_t by using all previous observations $\mathbf{y}_{1:t-1}$. In the implementation of the Kalman filter, denote $\hat{\theta}_{t|\tau}$ as the estimation of θ_t by using $\mathbf{y}_{1:\tau}$ with $\tau = t-1$ or t, $R_{t|\tau} = \mathbb{E}[|\hat{\theta}_{t|\tau} - \theta_{t|\tau}|^2]$ as the corresponding error variance; denote $\hat{y}_{t|\tau}$ as the estimation of y_t by using $\mathbf{y}_{1:\tau}$ with $\tau = t-1$, with error variance $Q_{t|\tau} = \mathbb{E}[|\hat{y}_{t|\tau} - y|^2]$. The values of $\hat{\theta}_{t|\tau}$, $R_{t|\tau}$, $\hat{y}_{t|\tau}$, and $Q_{t|\tau}$ can be iteratively updated by using the Kalman filter, and details are given in Algorithm 1.

The DLM is completely represented by the initial parameter set, $\{F, G, \sigma_v^2, \sigma_w^2 \hat{\theta}_{0|0}, R_{0|0}\}$. The values of these parameters for a given patient can be estimated by applying maximum likelihood estimation on the data of that patient. In this paper, the parameters are estimated by using the maximum likelihood estimation algorithm L-BFGS-B [30]. Once the parameters are estimated, Kalman filter can then be applied to track and estimate the dynamic evolution of the data in the time domain.

Algorithm 1: Kalman Filter.

1: **Input:** Discrete time dataset $\mathbf{y}_{1:T}$;
2: Estimate $\{F, G, \sigma_v^2, \sigma_w^2 \hat{\theta}_{0|0}, R_{0|0}\}$ by applying the L-BFGS-B algorithm on $\mathbf{y}_{1:T}$; initialize $t = 0$.
3: **do**
4: $t \leftarrow t + 1$;
5: Prediction of θ_t:

$$\hat{\theta}_{t|t-1} = G\hat{\theta}_{t-1|t-1} \tag{2}$$

$$R_{t|t-1} = G^2 R_{t-1|t-1} + \sigma_w^2 \tag{3}$$

6: Prediction of y_t:

$$\hat{y}_{t|t-1} = F\hat{\theta}_{t|t-1} \tag{4}$$

$$Q_{t|t-1} = F^2 R_{t|t-1} + \sigma_v^2 \tag{5}$$

7: Calculate prediction residual: $x_t = y_t - \hat{y}_{t|t-1}$;
8: Update estimation of θ_t:

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + F R_{t|t-1} Q_{t|t-1}^{-1} x_t \tag{6}$$

$$R_{t|t} = R_{t|t-1} - F^2 R_{t|t-1}^2 Q_{t|t-1}^{-1} \tag{7}$$

9: **while** $t \leq T$
10: **Output:** Prediction residual $\mathbf{x}_{1:T}$.

The Kalman filter described in Algorithm 1 iteratively tracks and predicts the state and observation variables using the DLM. The output of the Kalman filter is the prediction residual, $x_t = y_t - \hat{y}_{t|t-1}$, which is the difference between the predicted and actual value of y_t. It is expected that the prediction residuals under normal and abnormal conditions will exhibit different statistical properties, thus they can be used for anomaly detection.

3.2. Modeling PVP prediction residuals with HMM

The prediction residuals at the output of the Kalman filter are modeled by using HMM. Each residual x_t is associated with a binary hidden state $s_t \in \{0, 1\}$, where $s_t = 0$ indicates normal data, and $s_t = 1$ indicates anomalies. The HMM is illustrated in Fig. 4. Under the HMM, the residuals are conditionally independent with each other conditioned on their respective hidden states. The dependence of the residuals are represented through the hidden states. To simplify notation, define $\mathbf{x}_{1:T} = [x_1, x_2, \ldots, x_T]$.

The hidden states form a first-order Markov chain with transition probability, $a_{ij} = \Pr(s_{t+1} = j | s_t = i)$, for $i, j \in \{0, 1\}$. Define the transition probability matrix as

$$\mathbf{A} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \tag{8}$$

The initial distribution of the two states in the Markov chain is assumed to be $\pi_k = \Pr(s_1 = k)$ for $k \in \{0, 1\}$. Define the initial probability vector as $\boldsymbol{\pi} = [\pi_0, \pi_1]^T$, where the operator $(\cdot)^T$ represents matrix transpose.

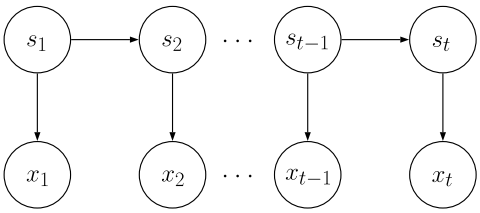

Fig. 4. Dependence structure of first-order hidden Markov model. Unlike Fig. 3, here $\{s_i\}$ are discrete random variables and x_i follows a Gaussian distribution depending on s_i.

The statistical distribution of each prediction residual depends on its hidden state. For state $s_t = k$, it is assumed that the prediction residuals follows a Gaussian distribution with mean μ_k and variance σ_k^2, that is,

$$x_t|(s_t = k) \sim \mathcal{N}(\mu_k, \sigma_k^2), \quad \text{for} \quad k \in \{0, 1\} \tag{9}$$

It is assumed that $\sigma_1^2 > \sigma_0^2$ due to the presence of noise or interference in the corrupted signals. Define the sample distribution parameter set as $\mathcal{B} = \{\mu_0, \mu_1, \sigma_0^2, \sigma_1^2\}$.

Based on the above definitions, the HMM can be represented by using the parameter set $\lambda = \{\mathbf{A}, \boldsymbol{\pi}, \mathcal{B}\}$. The HMM is assumed to be time-homogeneous, that is, the model parameters do not change over time.

The objective of anomaly detection is to identify the state $\{s_t\}$, by using the prediction residuals $\{x_t\}$. The identification of the state variables require the knowledge of the parameter set λ, which can be directly learned from $\{x_t\}$ by using the Baum–Welch algorithm.

Define the posterior probability of the hidden state variable as

$$\gamma_t(i) = \Pr(s_t = i|\mathbf{x}_{1:T}, \lambda). \tag{10}$$

If the HMM parameter set λ is known, then the anomaly detection algorithm can be formulated as

$$\hat{s}_t = \underset{i\in\{0,1\}}{\operatorname{argmax}} \quad \gamma_t(i). \tag{11}$$

The parameter set λ is unknown, and all parameters can be estimated from prediction residual $\{x_t\}$ by using the Baum–Welch algorithm.

3.3. Unsupervised learning with a modified Baum–Welch algorithm

In this section, unsupervised learning is performed to detect the anomalies in PVP signals with the help of a modified Baum–Welch algorithm [28]. The Baum–Welch algorithm is an expectation–maximization algorithm [31] that can iteratively learn the values of the unknown parameter set, λ, by maximizing the log-likelihood function

$$\ell = \log p(\mathbf{x}_{1:T}|\lambda) \tag{12}$$

The unsupervised learning process is carried out by using the observations, $\mathbf{x}_{1:T}$, without the need of a training phase.

The implementation of the Baum–Welch algorithm requires the iterative calculation of a forward probability, $\alpha_t(i)$, and a backward probability, $\beta_t(i)$, which are defined, respectively, as

$$\alpha_t(i) = \begin{cases} \pi_i p(x_1|s_1 = i, \lambda), & t = 1, \\ \Pr(\mathbf{x}_{1:t}, s_t = i|\lambda), & 2 \le t \le T \end{cases}$$

$$\beta_t(i) = \begin{cases} p(\mathbf{x}_{(t+1):T}|s_t = i, \lambda), & 1 \le t \le T-1, \\ 1, & t = T. \end{cases}$$

Based on the above definitions and the Markov property of the hidden states, the forward and backward probabilities can be iteratively updated as

$$\alpha_t(i) = \sum_{j\in\{0,1\}} \alpha_{t-1}(j) a_{ji} p(x_t|s_t = i, \lambda), \quad 2 \le t \le T$$

$$\beta_t(i) = \sum_{j\in\{0,1\}} a_{ij} p(x_{t+1}|s_{t+1} = j, \lambda)\beta_{t+1}(j),$$
$$\text{for} \quad 1 \le t \le T-1.$$

As the number of samples, T, increases, the values of $\alpha_t(i)$ and $\beta_t(i)$ get very small and technically become zero due to limited computer precision. To solve the problem of mathematical underflow, the 'log-sum-exponent' trick is used in many cases. However, the 'log-sum-exponent' trick is an approximate solution and it introduces additional computation time. To achieve mathematically accurate and robust estimation, we adopt the variable scaling approach as proposed in [22]. The notation is ambiguous while defining the scaled probabilities in [22]. This paper tries to give the physical meaning and better understanding of the scaled variables.

3.4. Modified Baum–Welch algorithm

Define a modified forward probability variable, $\tilde{\alpha}_t(i)$, as

$$\tilde{\alpha}_t(i) = \begin{cases} \Pr(x_1, s_1 = i|\lambda), & t = 1 \\ \Pr(x_t, s_t = i|\mathbf{x}_{1:t-1}, \lambda), & t > 1 \end{cases} \tag{13}$$

The variable $\tilde{\alpha}_t(i)$ can be interpreted as a scaled version of $\alpha_t(i)$ as

$$\tilde{\alpha}_t(i) = C_{t-1}\alpha_t(i) \tag{14}$$

where the scaling variable C_t is defined as follows

$$C_t^{-1} = \begin{cases} 1, & t = 0, \\ p(\mathbf{x}_{1:t}|\lambda) = \alpha_t(0) + \alpha_t(1), & t \ge 1 \end{cases}$$

Lemma 1. *The scaled forward probability variable, $\tilde{\alpha}_t(i)$, can be calculated in an iterative manner as*

$$\tilde{\alpha}_t(i) = c_{t-1} \sum_{j\in\{0,1\}} \tilde{\alpha}_{t-1}(j) \cdot a_{ji} \cdot p(x_t|s_t = i, \lambda), \tag{15}$$

for $2 \le t \le T$ and $\tilde{\alpha}_1(i) = \alpha_1(i)$, and c_t is defined as

$$c_t^{-1} = p(x_t|\mathbf{x}_{1:t-1}, \lambda) = \tilde{\alpha}_t(0) + \tilde{\alpha}_t(1) \tag{16}$$

Proof.

$$\begin{aligned} \tilde{\alpha}(i) &= \Pr(x_t, s_t = i|\mathbf{x}_{1:t-1}, \lambda) \\ &= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\lambda)}{p(\mathbf{x}_{1:t-1}|\lambda)} \\ &= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\lambda)}{p(x_{t-1}|\mathbf{x}_{1:t-2}, \lambda) \cdot p(\mathbf{x}_{1:t-2}|\lambda)}, \end{aligned} \tag{17}$$

The numerator in the above expression can be calculated as

$$\begin{aligned} &\Pr(\mathbf{x}_{1:t}, s_t = i|\lambda) \\ &= \sum_{j\in\{0,1\}} \Pr(\mathbf{x}_{1:t}, s_{t-1} = j, s_t = i|\lambda) \\ &= \sum_{j\in\{0,1\}} \Pr(\mathbf{x}_{1:t-1}, s_{t-1} = j|\lambda) \cdot a_{ji} \cdot p(x_t|s_t = i, \lambda) \\ &= p(\mathbf{x}_{1:t-2}|\lambda) \cdot p(x_t|s_t = i, \lambda) \\ &\times \sum_{j\in\{0,1\}} \left[\Pr(x_{t-1}, s_{t-1} = j|\mathbf{x}_{1:t-2}, \lambda) \cdot a_{ji}\right] \end{aligned} \tag{18}$$

Substituting (17) into (18) results in

$$\tilde{\alpha}(i) = c_{t-1} \sum_{j\in\{0,1\}} \tilde{\alpha}_{t-1}(j) \cdot a_{ji} \cdot p(x_t|s_t = i, \lambda).$$

This completes the proof. ■

Similarly, define a scaled backward probability variable, $\tilde{\beta}_t(i)$, as

$$\tilde{\beta}_t(i) = \begin{cases} \frac{\Pr(\mathbf{x}_{t+1:T}|s_t=i,\lambda)}{\Pr(\mathbf{x}_{t+1:T}|\mathbf{x}_{1:t}=i,\lambda)}, & 1 \le t \le T-1 \\ 1, & t = 1. \end{cases} \tag{19}$$

Also,

$$\tilde{\beta}_t(i) = D_{t+1}\beta_t(i), \tag{20}$$

where the scaling variable D_t is defined as

$$D_t^{-1} = \begin{cases} \Pr(\mathbf{x}_{t:T}|\mathbf{x}_{1:t-1}, \lambda), & 1 \le t \le T, \\ 1, & t = T+1. \end{cases} \tag{21}$$

Lemma 2. *The scaled backward probability variable, $\tilde{\beta}_t(i)$, can be calculated in an iterative manner as*

$$\tilde{\beta}_t(i) = c_{t+1} \sum_{j\in\{0,1\}} a_{ij}\, p(x_{t+1}|s_{t+1} = j, \lambda)\tilde{\beta}_{t+1}(j) \tag{22}$$

where $1 \le t \le T-1$ and $\tilde{\beta}_T(i) = \beta_T(i) = 1$.

Proof.

$$\begin{aligned} \tilde{\beta}_t(i) &= \frac{\Pr(\mathbf{x}_{t+1:T}|s_t = i, \lambda)}{p(\mathbf{x}_{t+1:T}|\mathbf{x}_{1:t}, \lambda)} \\ &= \frac{\Pr(\mathbf{x}_{t+1:T}|s_t = i, \lambda)}{p(x_{t+1}|\mathbf{x}_{1:t}, \lambda)\cdot p(\mathbf{x}_{t+2:T}|\mathbf{x}_{1:t+1}, \lambda)} \\ &= c_{t+1} \sum_{j\in\{0,1\}} a_{ij}\, p(x_{t+1}|s_{t+1} = j, \lambda)\tilde{\beta}_{t+1}(j). \end{aligned}$$

The last equality in the above equation is based on the fact that

$$\begin{aligned} &\Pr(\mathbf{x}_{t+1:T}|s_t = i, \lambda) \\ &= \sum_{j\in\{0,1\}} a_{ij}\, p(x_{t+1}|s_{t+1} = j, \lambda)\Pr(\mathbf{x}_{t+2:T}|s_{t+1} = j, \lambda). \end{aligned}$$

This completes the proof. ■

Based on the definitions of C_t, D_t, and c_t, we have the following relationship among them

$$C_t = c_1 c_2 \cdots c_t = \prod_{i=1}^{t} c_i = c_t C_{t-1}$$

$$D_t = c_t c_{t+1} \cdots c_T = \prod_{i=t}^{T} c_i = c_t D_{t+1}$$

The likelihood function can then be calculated as

$$p(\mathbf{x}_{1:T}|\lambda) = C_T^{-1} = \frac{c_t}{C_t D_t} = C_t^{-1} D_{t+1}^{-1} = c_t^{-1} C_{t-1}^{-1} D_{t+1}^{-1} \tag{23}$$

The posterior probability $\gamma_t(i)$ defined in (10) can be calculated as

$$\begin{aligned} \gamma_t(i) &= \frac{p(\mathbf{x}_{1:T}, s_t = i|\lambda)}{p(\mathbf{x}_{1:T}|\lambda)} \\ &= \frac{\Pr(\mathbf{x}_{1:t}, s_t = i|\lambda)\cdot p(\mathbf{x}_{(t+1):T}|s_t = i, \lambda)}{p(\mathbf{x}_{1:T}|\lambda)} \\ &= \alpha_t(i)\beta_t(i) C_{t-1} D_{t+1} c_t \\ &= c_t \tilde{\alpha}_t(i)\tilde{\beta}_t(i) \end{aligned} \tag{24}$$

During implementation, the relationship $\gamma_t(0) + \gamma_t(1) = 1$ can be used for sanity check. Due to the structure of HMM, $\mathbf{x}_{1:t}$ and $\mathbf{x}_{t+1:T}$ are conditionally independent given s_t.

With $\tilde{\alpha}_t(i)$ and $\tilde{\beta}_t(i)$, we can calculate the posterior probability of transitioning from state i at time t to state j at time $t+1$ as the implementation of the Baum–Welch algorithm requires the definition of the following probability

$$\begin{aligned} \xi_t(i,j) &= \Pr(s_t = i, s_{t+1} = j|\mathbf{x}_{1:T}, \lambda) \\ &= \frac{\alpha_t(i)\beta_{t+1}(j) a_{ij}\, p(x_{t+1}|s_{t+1} = j)}{p(\mathbf{x}_{1:T}|\lambda)} \\ &= c_t a_{ij} \tilde{\alpha}_t(i)\tilde{\beta}_t(j) p(x_{t+1}|s_{t+1} = j) \end{aligned} \tag{25}$$

The Baum–Welch algorithm updates the values of the parameter set λ in an iterative manner. Each iteration contains an expectation step (E-step) and a maximization step (M-step).

In the E-step, the probabilities, $\tilde{\alpha}_t(i)$, $\tilde{\beta}_t(i)$, $\gamma_t(i)$, and $\xi_t(i,j)$, are calculated by using (15), (22), (24) and (25) with the parameter set λ from the previous iteration.

In the M-step, the parameter set λ are updated by using the probabilities $\gamma_t(i)$ and $\xi_t(i,j)$ obtained in the E step. The parameter estimations [32] performed at the M-step in each iteration are

$$\pi_i = \gamma_1(i); \tag{26}$$

$$\mu_i = \frac{\sum_{t=1}^{T} \gamma_t(i) x_t}{\sum_{t=1}^{T} \gamma_t(i)} \tag{27}$$

$$\sigma_i^2 = \frac{\sum_{t=1}^{T} \gamma_t(i)(x_t - \mu_i)^2}{\sum_{t=1}^{T} \gamma_t(i)} \tag{28}$$

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{29}$$

for $i, j \in \{0, 1\}$. It is important to note that the estimated parameters are calculated using $\tilde{\alpha}_t(i)$ and $\tilde{\beta}_t(i)$ instead of using α_t and β_t.

In the kth iteration, we can update the log-likelihood function as

$$\ell(k) = -\log C_T. \tag{30}$$

The iterative procedure is terminated if $\ell(k+1) - \ell(k)$ is less than a predefined threshold ϵ.

At the beginning of the iteration process, the parameters are initialized as

$$\pi_0 = \pi_1 = 0.5 \tag{31a}$$

$$\mathbf{A} = \tfrac{1}{2}\mathbf{I}_2 \tag{31b}$$

$$\mu_0 = 0 \tag{31c}$$

$$\mu_1 = 0 \tag{31d}$$

$$\sigma_0 = s \tag{31e}$$

$$\sigma_1 = \frac{1}{2}\max(|\mathbf{x}_{1:T}|). \tag{31f}$$

Here, $\bar{x} = \sum_{i=1}^{T} x_i$ and $s^2 = \frac{1}{T-1}\sum_{i=1}^{T}(x_i - \bar{x})^2$ are sample mean and sample variance of $\mathbf{x}_{1:T}$, respectively, and $\mathbf{I}_2$ is a size 2 identity matrix. Also, $|.|$ is the absolute value operator, operating individually on each element of the vector. Proper care should be taken while selecting values of σ_0 and σ_1, if these values are too big or small then likelihood $\Pr(x_t|s_t = i)$ can be very small which will force subsequent forward and backward probabilities to zero.

The modified Baum–Welch algorithm is summarized in Algorithm 2.

Algorithm 2: Baum–Welch Algorithm.

1: **Input:** Discrete time dataset $\mathbf{x}_{1:T}$;
2: Initialize λ with Eq. (31), set $k = 0$ and $\ell(0) = -\infty$;
3: **do**
4: $k \leftarrow k + 1$;
5: Calculate $\tilde{\alpha}_t(i)$ and c_t using Eq. (15) and Eq. (16);
6: Calculate $\tilde{\beta}_t(i)$ using Eq. (22);
7: Calculate $\gamma_t(i)$ and $\xi_t(i,j)$ using Eq. (24) and Eq. (25);
8: Update $\pi_i, \mu_i, \sigma_i^2, \mathbf{A}$ using Eq. (26)-Eq. (29), here $\sigma_1^2 > \sigma_0^2$;
9: Calculate $\ell(k)$ using Eq. (30);
10: **while** $\ell(k) - \ell(k-1) > \epsilon$
11: **Output:** λ, $\gamma_t(i)$, for $t = 1, \cdots, T$ and $i \in \{0, 1\}$.

3.5. Anomaly removals

Based on the output of the Baum–Welch algorithm, the hidden state $\hat{s}_t$ of each sample can be detected by applying $\gamma_t(i)$ in (11). Samples with estimated hidden state $\hat{s}_t = 1$ are labeled as anomalies.

Since PVP signals are usually analyzed in the frequency domain, we need to ensure the continuity of the signal in the time domain after the removal of the anomalous samples. To achieve this goal, the PVP signals are first divided into non-overlapping windows with w samples

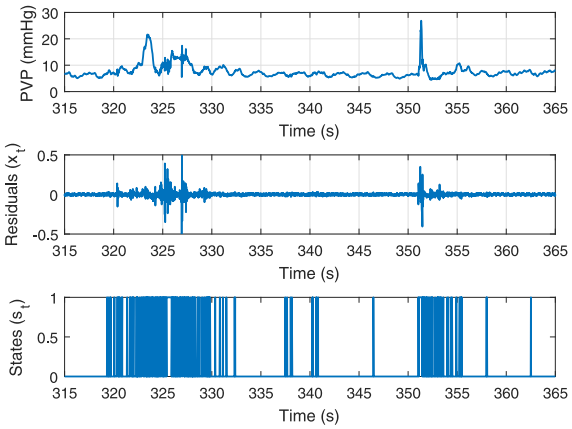

Fig. 5. (Top) Exemplary PVP signal from Fig. 1; (middle) Prediction residual of the Kalman filter; and (bottom) estimated hidden states ($s_t = 0$: normal sample; $s_t = 1$: anomalous sample).

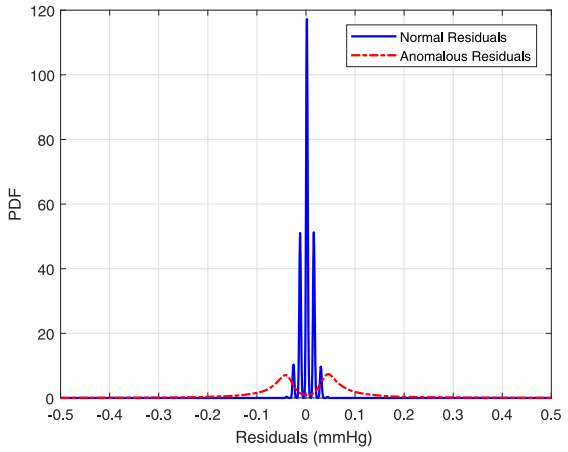

Fig. 6. Empirical probability density function of prediction residuals (x_t) of patient 10.

per window. If the percent of corresponding anomalous residuals within a window exceeds a certain threshold ζ (e.g. $\zeta = 15\%$), then all samples within this window are discarded. Such a window-based anomaly removal ensures the time continuity of samples within the same window. The remaining windows are considered as normal windows with valid data. Fast Fourier transform (FFT) can then be applied to each normal window to obtain the frequency domain representation of the PVP signal. With a window size of w samples with sampling period T_0, the frequency resolution of the frequency domain signal is $\frac{1}{wT_0}$.

The window-based anomaly removal approach will remove both normal and anomalous sample inside a window if the percentage of anomalous samples inside that window exceeds a certain threshold. Thus using a smaller window size can effectively reduce the amount of normal samples being removed. On the other hand, a smaller window will results in a lower resolution in the frequency domain. Thus the window size determines a tradeoff between the precision of anomaly removal and frequency domain resolution. In this paper, we pick a window size of 10,000 samples, which corresponds to a duration of 10 s with a sampling rate of 1 kHz, to ensure a frequency domain resolution of 0.1 Hz. Algorithm 3 summarizes the procedure of window-based anomaly removal.

Algorithm 3: Window-based Anomaly Removal.

1: **Input:** PVP signal $\mathbf{y}_{1:T}$, window size w, threshold ζ;
2: Infer $\hat{\mathbf{s}}_{1:T}$ using Algorithms 1 and 2.
3: **for** $n = 1$ to $\lfloor T/w \rfloor$ **do**
4: Calculate the percentage of anomaly samples in the n-th window

$$\tau_n = \frac{1}{w}\sum_{i=1}^{w} \hat{s}_{(n-1)w+i}$$

5: If $w \geq \zeta$, discard the n-th window;
6: **end for**
7: **Output:** Normal windows.

4. Results

The proposed anomaly detection and removal algorithm is applied to clinical data to verify its effectiveness. The data used in the experiments are PVP signals collected from a cohort of 24 pediatric patients suffering from Pyloric stenosis during emergency department admissions. The characteristics of the patient population are summarized in Table 2. Based on the concentration of serum chloride or bicarbonate of each patient, patients were classified as either euvolemic or hypovolemic as described in Section 2. Among the 24 patients, 14 were labeled as euvolemic and 10 were labeled as hypovolemic. The signals for each patient are collected over a continuous period of time

Table 1
Number of normal and anomalous windows for each patient ($\zeta = 15\%$).

Patient	Status	Normal	Anomaly	Anomaly (%)
5	Hypo.	55	9	14.06
6	Hypo.	65	2	02.99
7	Hypo.	14	34	70.83
9	Euvo.	55	1	01.79
10	Euvo.	42	24	36.36
12	Euvo.	35	23	39.66
18	Hypo.	27	17	38.64
20	Euvo.	36	5	12.20
22	Hypo.	42	5	10.64
23	Euvo.	26	5	16.13
24	Euvo.	26	7	21.21
25	Hypo.	49	9	15.52
26	Hypo.	45	4	08.16
27	Euvo.	22	11	33.33
28	Euvo.	45	2	04.26
29	Hypo.	43	3	06.52
30	Euvo.	37	9	19.57
31	Euvo.	52	9	14.75
32	Hypo.	38	9	19.15
33	Euvo.	30	14	31.82
34	Hypo.	30	15	23.33
35	Euvo.	50	7	12.28
37	Euvo.	40	13	24.53
39	Euvo.	40	9	18.37

Table 2
Patient characteristics.

	Euvo.	Hypo.
Patients	14	10
Average weight (kg)	4.17	3.89
Minimum weight (kg)	2.76	2.72
Maximum weight (kg)	5.82	4.72
Std. deviation (kg)	0.78	0.70
Mean age (days)	38.1	40.1

Table 3
Number of windows in dataset.

	Training set			Testing set		
	Euvo.	Hypo.	Total	Euvo.	Hypo.	Total
Raw data	466	354	820	209	161	370
Manual	333	273	606	155	124	279
Algorithm in [14]	372	306	678	172	139	311
Proposed algorithm	371	281	652	165	127	292

Table 4
Testing classification results.

Parameter	Raw data	Manual	Algorithm in [14]	Proposed algorithm
True positive rate	45.96%	69.35%	63.31%	71.65%
True negative rate	76.08%	77.42%	79.65%	81.21%
Precision	59.68%	73.83%	71.54%	74.60%
F1 score	52.00%	75.41%	67.18%	73.09%
Accuracy	62.97%	71.07%	72.35%	77.05%
Windows used	100%	70.20%	84.05%	78.92%

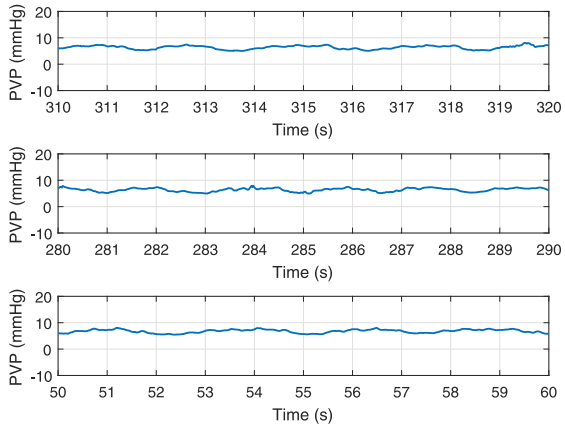

Fig. 7. Example of normal windows inferred by the proposed model. Windows have a periodic structure and the amplitude does not change abruptly.

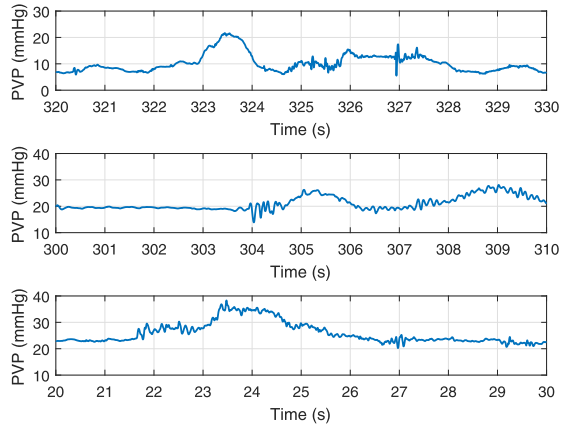

Fig. 8. Example of anomalous windows inferred by the proposed model. In these windows, amplitude changes abruptly (10–15 mmHg higher than average) indicating random MNAs.

with a sampling rate of $f_s = 1$ kHz. The PVP signal is also affected by heart rate in the range 1.9–2.5 Hz (114 – 150 bps) and respiratory rate at 0.25–0.45 Hz (15 – 27 bps) and their higher order harmonics. Amplitude of the signal varies from $5 - 40$ mmHg. The signal duration collected from each patient ranges from 310 to 670 s.

For each patient, the unsupervised anomaly detection algorithm is applied to the entire data sequence collected from that patient. As an example, Fig. 5 shows the PVP signal along with the estimated hidden states from patient 20 for a duration of 50 s. Hidden states 1 and 0 correspond to anomalous samples or normal samples, respectively. The probability density functions of anomalous and normal data residuals for patient 10 are shown in Fig. 6. As can be seen from the results, all large fluctuations in the signals are labeled as anomalous by the proposed algorithm. The large fluctuations are mainly caused by sudden movements of the pediatric patients during the data collection process, and they deviate significantly from normal PVP signals. The results in Fig. 6 demonstrate that the distributions of residuals from anomalous data are quite different from their normal counterpart, thus the algorithm can effectively distinguish between the two by analyzing their statistical distributions.

Among all PVP signal windows, 20.67% are detected as anomalous by the HMM-based anomaly detection algorithm. Here, $\epsilon = 10^{-5}$ has been used. A window is declared as anomalous if it contains more than $\zeta = 15\%$ of anomalous residual samples. Figs. 7 and 8 show a few examples of the PVP signals in normal and anomalous windows, respectively. Table 1 lists the number of normal and anomalous windows for each patient.

The direct quantitative evaluation of the anomaly detection algorithm requires labeled data with each sample classified as either normal or anomalous. However, due to the unexpected nature of anomaly events, no labeled data is available for testing of the HMM-based anomaly detection algorithm. Manually labeling the data might introduce subjective bias, which will affect the objectiveness of the evaluation results. The only objective labels that are available to the data are the euvolemic or hypovolemic status of the patients. It has been shown in [6] that there is a strong correlation between PVP signals and the volume status of pediatric patients. That is, the dehydration status of a pediatric patient can be accurately predicted by analyzing PVP signals. Motivated by this result, we propose to indirectly evaluate the effectiveness of the HMM-based anomaly detection algorithm through volume status prediction.

Based on the results in [6], our hypothesis is that removing anomaly in PVP signals can improve the prediction accuracy of dehydration in pediatric patients. To verify the hypothesis, the experiment is designed as follows. First, the original raw PVP signals without cleaning are used to train and detect the volume status of the patients. The accuracy of the prediction can be quantitatively evaluated given that all patients are objectively labeled as euvolemic or hypovolemic. Second, apply the HMM-based anomaly detection algorithm to the PVP signals, and remove all anomalous windows. Then perform volume status prediction with the cleaned PVP signal after anomaly removal. It is expected that the cleaned signal can achieve a higher accuracy of volume status prediction compared to its uncleaned counterpart.

The volume status prediction is performed by using regularized logistic regression as described in [6]. The regularized logistic regression is a supervised learning algorithm for classification, thus it requires a training phase to build the classification model. It should be noted the training is only used for volume status classification, and no training is needed for the proposed anomaly detection algorithm.

During the analysis, the data are divided into non-overlapping windows with $w = 10{,}000$ samples per window. Thus the duration of each window is $T_w = 10$ seconds. Signals in each window are converted to the frequency domain by using FFT. The frequency resolution of the frequency domain signals is $f_0 = \frac{1}{T_w} = 0.1$ Hz, with the highest frequency being $\frac{1}{2}f_s = 500$ Hz. Since PVP signals are mainly in the low frequency range, only signal components below 20 Hz are used in the analysis. This results in a frequency domain vector of length 200 for each window. For volume status prediction with regularized logistic regression, the training-testing data was split with a $70\% - 30\%$ ratio. During training, a 5-fold cross validation was used to tune the regularization parameter. The number of windows used in both cases are presented in Table 3.

The classification results obtained from raw data, manually cleaned data, data cleaned by a simple anomaly removal method in [14], and

data cleaned by the proposed algorithm are summarized in Table 4. For the algorithm in [14], the model parameters are set as $a = 4$, $b = -0.1$ and $w_s = 100$. Even though the denominator used by different algorithms are different in Tables 3 and 4, the difference is relatively small. For example, the numbers of testing samples for manual cleaning, algorithm in [14], and proposed algorithm are 279, 311, and 292, respectively, which are within 5.7% of their mean value. As can be seen from the results, after removing anomaly with the proposed anomaly detection algorithm, the accuracy of volume status prediction is considerably improved compared to the results obtained from the uncleaned raw data. The proposed algorithm also outperforms manual cleaning or the cleaning algorithm in [14]. These results demonstrate that the HMM-based anomaly detection algorithm can effectively remove the anomalous data from the collected PVP signals, thus more accurate volume status prediction is achieved by using the cleaned PVP signals.

5. Conclusion

An unsupervised anomaly detection algorithm has been proposed to remove MNA in PVP signals. The PVP signals were modeled by using a DLM to capture correlation of neighboring samples of the PVP signals. Kalman filter was applied to the DLM to track and predict the PVP signal in the time domain. The prediction residuals were then modeled by using an HMM, where the normal and anomalous states of the PVP signals were represented by using the binary states of the hidden Markov chain. The HMM parameters along with the hidden states were iteratively estimated with a modified Baum–Welch algorithm to ensure the numerical stability. The proposed algorithm was verified by comparing the hypovolemic prediction accuracy of PVP signals before and after the anomaly processing. Experimental results with clinical data of a cohort of 24 pediatric patients demonstrated that the algorithm can efficiently remove anomalies in PVP signals. The algorithm can also be used for anomaly detection in other time series signals, such as ECG and PPG signals.

CRediT authorship contribution statement

Md Abul Hayat: Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Jingxian Wu:** Methodology, Writing - original draft, Writing - review & editing, Funding acquisition. **Patrick C. Bonasso:** Conceptualization, Investigation, Resources, Data curation, Writing - review & editing. **Kevin W. Sexton:** Conceptualization, Investigation, Resources, Data curation, Supervision, Writing - review & editing, Funding acquisition. **Hanna K. Jensen:** Conceptualization, Investigation, Supervision, Writing - review & editing. **Melvin S. Dassinger:** Conceptualization, Investigation, Supervision, Writing - review & editing, Funding acquisition. **Morten O. Jensen:** Conceptualization, Investigation, Supervision, Writing - review & editing, Funding acquisition.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.bspc.2020.102126.

Acknowledgments

The authors want to thank Dr. Giovanni Petris (Professor of Mathematical Sciences, University of Arkansas) for his thoughtful remarks and ideas. All supplementary materials and codes of this paper can be found at this github repository: https://github.com/i2pt/UAD_PVPS_HMM

The work of M. A. Hayat and J. Wu was supported in part by the U.S. National Science Foundation (NSF) under Award Number ECCS-1711087. The work of P. C. Bonasso and K. W. Sexton was supported in part by the Arkansas Children's Hospital, USA and the University of Arkansas for Medical Sciences (UAMS), USA Clinician Scientist Program.

References

[1] M. Niescierenko, R. Bachur, Advances in pediatric dehydration therapy, Curr. Opin. Pediatr. 25 (3) (2013) 304–309.
[2] J.E. Wathen, T. MacKenzie, J.P. Bothner, Usefulness of the serum electrolyte panel in the management of pediatric dehydration treated with intravenously administered fluids, Pediatrics 114 (5) (2004) 1227–1234.
[3] L. Paladino, R. Sinert, D. Wallace, T. Anderson, K. Yadav, S. Zehtabchi, The utility of base deficit and arterial lactate in differentiating major from minor injury in trauma patients with normal vital signs, Resuscitation 77 (3) (2008) 363–368.
[4] V.A. Convertino, Blood pressure measurement for accurate assessment of patient status in emergency medical settings, Aviat. Space Environ. Med. 83 (6) (2012) 614–619.
[5] P.C. Bonasso, M.S. Dassinger, B. McLaughlin, J.M. Burford, K.W. Sexton, Fast fourier transformation of peripheral venous pressure changes more than vital signs with hemorrhage, Mil. Med. 184 (Suppl. 1) (2019) 318–321.
[6] P.C. Bonasso, K.W. Sexton, M.A. Hayat, J. Wu, H.K. Jensen, M.O. Jensen, J.M. Burford, M.S. Dassinger, Venous physiology predicts dehydration in the pediatric population, J. Surg. Res. 238 (2019) 232–239.
[7] D. Amar, J.A. Melendez, H. Zhang, C. Dobres, D.H. Leung, R.E. Padilla, Correlation of peripheral venous pressure and central venous pressure in surgical patients, J. Cardiothorac. Vasc. Anesth. 15 (1) (2001) 40–43.
[8] R. Desjardins, A.Y. Denault, S. Bélisle, M. Carrier, D. Babin, S. Lévesque, R. Martineau, Can peripheral venous pressure be interchangeable with central venous pressure in patients undergoing cardiac surgery? Intensive Care Med. 30 (4) (2004) 627–632.
[9] J.R. Munis, S. Bhatia, L.J. Lozada, Peripheral venous pressure as a hemodynamic variable in neurosurgical patients, Anesth. Analg. 92 (1) (2001) 172–179.
[10] N. Hoftman, M. Braunfeld, G. Hoftman, A. Mahajan, Peripheral venous pressure as a predictor of central venous pressure during orthotopic liver transplantation, J. Clin. Anesth. 18 (4) (2006) 251–255.
[11] N. Hadimioglu, Z. Ertug, A. Yegin, S. Sanli, A. Gurkan, A. Demirbas, Correlation of peripheral venous pressure and central venous pressure in kidney recipients, Transplant. Proc. 38 (2) (2006) 440–442.
[12] K.M. Hocking, B. Sileshi, F.J. Baudenbacher, R.B. Boyer, K.L. Kohorst, C.M. Brophy, S.S. Eagle, Peripheral venous waveform analysis for detecting hemorrhage and iatrogenic volume overload in a porcine model, Shock 46 (4) (2016) 447–452.
[13] B. Sileshi, K.M. Hocking, R.B. Boyer, F.J. Baudenbacher, K.L. Kohurst, C.M. Brophy, S. Eagle, Peripheral venous waveform analysis for detecting early hemorrhage: a pilot study, Intensive Care Med. 41 (6) (2015) 1147–1148.
[14] P.C. Bonasso, K.W. Sexton, S.C. Mehl, M.S. Golinko, M.A. Hayat, J. Wu, M.O. Jensen, S.D. Smith, J.M. Burford, M.S. Dassinger, Lessons learned measuring peripheral venous pressure waveforms in an anesthetized pediatric population, Biomed. Phys. Eng. Express 5 (3) (2019) 035020.
[15] P.C. Bonasso, K.W. Sexton, A. Hayat, A. Al-Alawi, J. Wu, H.K. Jensen, M.O. Jensen, S.D. Smith, J.M. Burford, M.S. Dassinger, Venous physiology predicts anesthetic induced hypotension in infants, J. Am. Coll. Surg. 227 (4) (2018) e116.
[16] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, ACM Comput. Surv. (CSUR) 41 (3) (2009) 15.
[17] J. Lee, D.D. McManus, S. Merchant, K.H. Chon, Automatic motion and noise artifact detection in Holter ECG data using empirical mode decomposition and statistical approaches, IEEE Trans. Biomed. Eng. 59 (6) (2011) 1499–1506.
[18] M.R. Ram, K.V. Madhav, E.H. Krishna, N.R. Komalla, K.A. Reddy, A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter, IEEE Trans. Instrum. Meas. 61 (5) (2011) 1445–1457.
[19] P. Malhotra, L. Vig, G. Shroff, P. Agarwal, Long short term memory networks for anomaly detection in time series, in: Proceedings, Presses universitaires de Louvain, 2015, p. 89.
[20] F. Karim, S. Majumdar, H. Darabi, S. Chen, LSTM fully convolutional networks for time series classification, IEEE Access 6 (2017) 1662–1669.
[21] G. Petris, S. Petrone, P. Campagnoli, Dynamic linear models, in: Dynamic Linear Models with R, Springer, 2009, pp. 31–84.
[22] L.R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE 77 (2) (1989) 257–286.
[23] A. Krogh, B. Larsson, G. Von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (3) (2001) 567–580.
[24] V. Narasimhan, P. Danecek, A. Scally, Y. Xue, C. Tyler-Smith, R. Durbin, BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data, Bioinformatics 32 (11) (2016) 1749–1751.
[25] M. Gales, S. Young, et al., The application of hidden Markov models in speech recognition, Found. Trends Signal Process. 1 (3) (2008) 195–304.
[26] C.-M. Chen, D.-J. Guan, Y.-Z. Huang, Y.-H. Ou, Anomaly network intrusion detection using hidden Markov model, Int. J. Innov. Comput. Inform. Control 12 (2016) 569–580.

[27] C.Y. Goh, J. Dauwels, N. Mitrovic, M.T. Asif, A. Oran, P. Jaillet, Online map-matching based on hidden markov model for real-time traffic sensing applications, in: 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE, 2012, pp. 776–781.

[28] L.E. Baum, T. Petrie, G. Soules, N. Weiss, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, Ann. Math. Stat. 41 (1) (1970) 164–171.

[29] P. Bonasso, M. Dassinger, M. Jensen, S. Smith, J. Burford, K. Sexton, Optimizing peripheral venous pressure waveforms in an awake pediatric patient by decreasing signal interference, J. Clin. Monitor. Comput. 32 (6) (2018) 1149.

[30] G. Petris, R. An, An R package for dynamic linear models, J. Stat. Softw. 36 (12) (2010) 1–16.

[31] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc. Ser. B Stat. Methodol. 39 (1) (1977) 1–22.

[32] W. Zucchini, I.L. MacDonald, R. Langrock, Hidden Markov Models for Time Series: An Introduction Using R, Chapman and Hall/CRC, 2016.