Journal of Computers Vol. 30 No. 2, 2019, pp. 41-54
doi:10.3966/199115992019043002004

# Examination of Time-Domain Features of EHG Data for Preterm-Term Birth Classification

Chomkansak Hemthanon[1], Suparerk Janjarasjitt[1*]

[1] Department of Electrical and Electronic Engineering, Ubon Ratchathani University, Warin Chamrap, Ubon Ratchathani 34190, Thailand
{chomkansak.he.56, suparerk.j}@ubu.ac.th

Received 15 December 2018; Revised 15 February 2019; Accepted 15 February 2019

Abstract. Prematurity is one of major health concerns that can cause short-term and long-term complications. There are evidences that uterine electromyography (EMG), also called electrohysterography (EHG), measuring the electrical activity of uterine muscles is a valuable diagnosis for term and preterm labor assessments. In this study, seventeen time-domain features commonly used in EMG signal processing and analysis of EHG data are examined and applied for preterm-term birth classifications. Two feature selection methods including Pearson's correlation coefficient and *p*-value of *t*-test are also applied to reduce the dimension of features. The classifications are performed using support vector machine (SVM) classifiers with polynomial kernel function. From the computational results, the best performance on preterm-term birth classifications of selected time-domain features of EHG data determined from the product of sensitivity and specificity achieved is the accuracy of 0.6667, the sensitivity of 0.7895 and the specificity of 0.6503. In addition, the computational results suggest that the remarkable time-domain features of EHG data for the preterm-term classifications are the difference absolute standard deviation value, the waveform length, the average amplitude change and the ν-order.

Keywords: classification, electrohysterogram, preterm birth, pregnancy, support vector machine

## 1 Introduction

Electromyography (EMG) measures an electrical signal corresponding to activity of muscles. EMG signals have been one of the most crucial biomedical signals that are used in various applications ranging from clinical diagnosis to human-machine interface. Typically, EMG data are recorded by attaching electrodes to skin where the muscles assessed are located. Recently, EMG data recorded from electrodes placed on abdominal around uterine, called uterine EMG or electrohysterography (EHG), for measuring the activity of uterine muscles [1] are applied for term and preterm labor assessments [1-3] Premature birth can cause several health problems to babies including short-term complications (e.g., breathing problems, heart problems, brain problems, blood problems and immune system problems) and also long-term complications (e.g., cerebral palsy, impaired cognitive skills, vision problems, hearing problems, chronic health issues and behavioral and psychological problems) [4]. An anticipation of premature birth is an exceptionally difficult task [1].

There are a number of digital signal processing and computational techniques have been applied to EHG data for various clinical applications including preterm birth classification. In [1], quantitative measures of EHG data obtained using linear digital signal processing techniques (i.e., the root-mean-square (RMS), the peak frequency, the median frequency, and the autocorrelation zero crossing) and also nonlinear digital signal processing techniques (i.e., the maximual Lyapunov exponent, the correlation dimension, and the sample entropy). Those quantitative measures were applied for differentiating EHG data associated with preterm birth from EHG data associated with term birth. The median frequency and

* Corresponding Author

the sample entropy of EHG recording associated with the preterm birth and those with the term birth are shown to be statistically, significantly different [1].

In [5], wavelet-based features, defined as the difference between logarithms of variances of detail coefficients corresponding to two consecutive levels of wavelet decomposition, of EHG data were examined applied for preterm birth classification. The performance on preterm birth classification using single wavelet-based features of EHG signals was evaluated using the leave-one-out cross validations. The wavelet-based features of EHG signals were shown to provide a reasonable performance on preterm birth classification [5] that is superior to the quantitative measures provided by and examined in [5]. The performance on preterm birth classification using single wavelet-based features of EHG data was furthermore reevaluated using the receiver operating characteristic (ROC) analysis in [6]. The best area under ROC curve obtained for the preterm birth classification using the wavelet-based features of EHG data is 0.7386.

The empirical mode decomposition (EMD), another fundamental computational technique applied to EMG data, was applied to EHG data for preterm birth discrimination in [7]. Quantitative measures used for discriminating EHG data associated with preterm birth are a ratio of Shannon entropies of intrinsic mode functions (IMFs) of EHG data obtained from the EMD [7]. A variety of classifiers consisting of support vector machine (SVM), random forests, multilayer perception, AdaBoost, Bayesian network, and simple logistic regression, are used and the classifications were performed using the 10-fold cross validation. The computational results showed that the classifications using the EMD are better than the classification without the EMD [7] where the area under ROC curve obtained from the classifications using the EMD ranges from 0.764 to 0.986 and the area under ROC curve obtained from the classifications without the EMD ranges from 0.503 to 0.914.

In this study, it is aimed to apply time-domain signal processing and analysis techniques commonly applied to EMG data to EHG data for evaluating their performances on preterm birth classification. The foremost advantage of time-domain signal processing and analysis techniques is the simplicity. Quantitative measures obtained from time-domain signal processing and analysis techniques are straightly determined from amplitudes of time series. Quantitative measures of EHG data obtained from common time-domain signal processing and analysis techniques including integrated EMG, mean absolute value, modified mean absolute value (type 1), modified mean absolute value (type 2), simple square integral, variance, absolute value of the 3rd, 4th and 5th temporal moments, root mean square (RMS), ν-order, log detector, waveform length, average amplitude change, difference absolute standard deviation value, zero crossing, and slope sign change are classified using SVM classifiers with polynomial kernel function. Furthermore, in this study, simple feature selection methods, i.e., Pearson's correlation coefficient and $p$-value of $t$-test, are also examined. The performance on preterm birth classifications is evaluated using the leave-one-out cross validation.

## 2 Methods

### 2.1 Subjects and EHG Data

Electrohysterogram (EHG) or uterine EMG data examined in this study are obtained from the Term-Preterm EHG Database (TPEHGDB) on PhysioNet (available online at http://www.physionet.org/ physiobank/database/ tpehgdb/. The EHG data were collected from 1997 until 2006 at the Department of Obstetrics and Gynecology, Medical Centre Ljubljana, Ljubljana [1, 8]. The EHG data were recorded from a general population of pregnant women during regular check-ups either around the 22nd week of gestation or around the 32nd week of gestation [1, 8]. The TPEHGDB contains a total of 300 recordings of EHG data. Those recordings can be divided into 262 recordings associated with term birth and 38 recordings associated with preterm birth. In addition, 143 of 262 recordings associated with term birth were recorded before the 26th week of gestation while the rest were recorded during or after the 26th week of gestation. Also, 19 of 38 recordings associated with preterm birth were recorded before the 26th week of gestation while the rest were recorded during or after the 26th week of gestation. The recordings of EHG data can therefore be classified into 4 classes, i.e., PE, PL, TE, and TL, according to their corresponding time of delivery (either term and preterm birth) and time of recordings (either early or late period of pregnancy). Table 1 summarizes the number of EHG data and their corresponding classes.

**Table 1.** Description of EHG data

| Case | Number of recording | Time of recording | Time of delivery |
|---|---|---|---|
| PE | 19 | Before the 26th week of gestation | Before or on the 37th week of gestation |
| PL | 19 | During or after the 26th week of gestation | After the 37th week of gestation |
| TE | 143 | Before the 26th week of gestation | Before or on the 37th week of gestation |
| TL | 119 | During or after the 26th week of gestation | After the 37th week of gestation |

The EHG data were recorded using the sampling frequency of 20 Hz. Each recording is composed of 3 channels, referred to as $s_1$, $s_2$, and $s_3$, recorded from 4 electrodes placed around the navel [1, 8]. In addition, the original EHG data were filtered with 3 different bandpass frequencies including (a) from 0.08 Hz and 4 Hz; (b) from 0.3 Hz to 3 Hz; and c) 0.3 Hz to 4 Hz. The first, second, and third subbands of EHG channels $s_1$, $s_2$, and $s_3$ are, respectively, referred to as $s_{1a}$, $s_{1b}$ and $s_{1c}$, $s_{2a}$, $s_{2b}$ and $s_{2c}$, and $s_{3a}$, $s_{3b}$ and $s_{3c}$. Similar to our previous study [5], only a segment with length of 8192 samples obtained from the middle section of each EHG channel is examined. Examples of EHG segments corresponding to channels $s_1$, $s_2$, and $s_3$ associated with preterm birth are illustrated, respectively, in Fig. 1 while Fig. 2 shows, respectively, exemplary EHG segments corresponding to channels $s_1$, $s_2$, and $s_3$ associated with term birth.

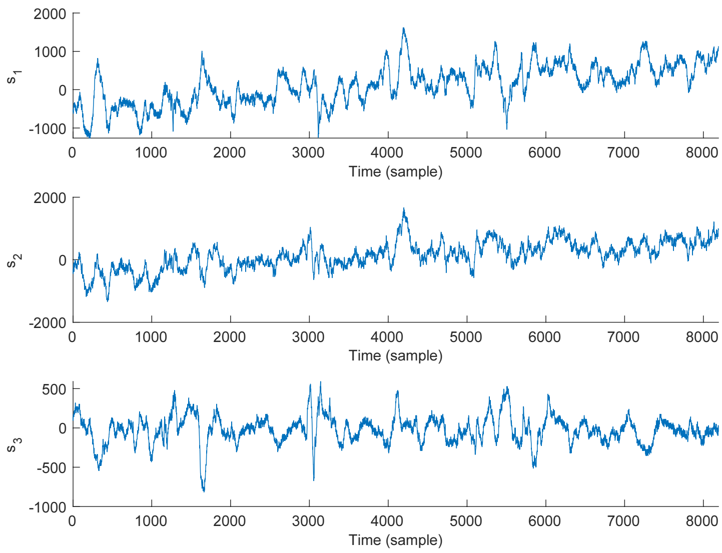



**Fig. 1.** Exemplary EHG segments associated with preterm birth

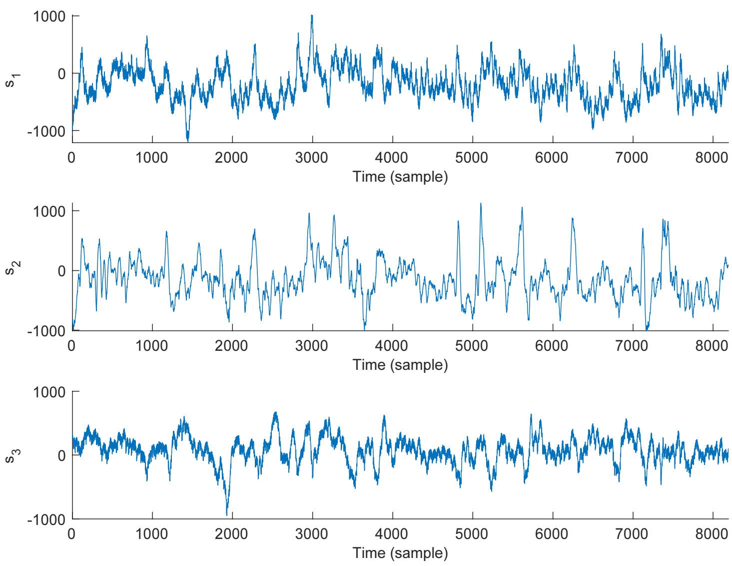



**Fig. 2.** Exemplary EHG segments associated with term birth

### 2.2 Time-Domain Features of EMG Data

There are 17 common time-domain features [9] of EMG data applied to EHG data in this study. Denote $\{x[n]\}$ where $n = 0, 1, 2, ..., N$-1 an EHG signal. Those time-domain features of EMG data are as follows.

(1) Integrated EMG (IEMG): The integrated EMG is typically used as an onset detection index [9]. It can be given by

$$F_1 = \sum_{n=0}^{N-1} \left| x[n] \right|. \quad \textbf{(1)}$$

(2) Mean Absolute Value (MAV): The mean absolute value, one of the most popular EMG data analysis techniques, is similar to the integrated EMG which is used for onset detection [9]. It can be given by

$$F_2 = \frac{1}{N} \sum_{n=0}^{N-1} \left| x[n] \right|. \quad \textbf{(2)}$$

(3) Modified Mean Absolute Value (Type 1): The mean absolute value (type 1) is modified from the mean absolute value where a weighted window function $w$ is applied [9]. It can be given by

$$F_3 = \frac{1}{N} \sum_{n=0}^{N-1} w[n] \left| x[n] \right|. \quad \textbf{(3)}$$

where

$$w[n] = \begin{cases} 1 & N/4 \le n \le 3N/4 \\ 0.5 & \textit{otherwise.} \end{cases}$$

(4) Modified Mean Absolute Value (Type 2): The mean absolute value (type 2) is another method modified from the mean absolute value. Similar to the modified mean absolute value (type 1), a smoother weighted window function $w$ is applied [9]. It can be given by

$$F_4 = \frac{1}{N} \sum_{n=0}^{N-1} w[n] \left| x[n] \right| \quad \textbf{(4)}$$

where

$$w[n] = \begin{cases} 1 & N/4 \le n \le 3N/4 \\ 4n/N & n < N/4 \\ 4(N-n-1)/N & \textit{otherwise.} \end{cases}$$

(5) Simple Square Integral (SSI): The simple square integral is simply an energy of EMG signal [9-10]. It can be given by

$$F_5 = \sum_{n=0}^{N-1} x^2[n]. \quad \textbf{(5)}$$

(6) Variance: The variance of EMG signal is another quantitative measure corresponding to an energy of EMG signal [9]. It can be given by

$$F_6 = \frac{1}{N-1} \sum_{n=0}^{N-1} \left| x[n] - \bar{x} \right|^2 \quad \textbf{(6)}$$

where $\bar{x}$ denotes the mean of $\{x[n]\}$.

(7) Absolute Value of the 3rd, 4th and 5th Temporal Moments: The temporal moments are statistical values originally proposed by [11]. Typically, the absolute value is taken to greatly reduce the within class separation for the odd moments [9]. The absolute value of the 3rd, 4th, and 5th temporal moments are, respectively, given by

$$F_7 = \left| \frac{1}{N} \sum_{n=0}^{N-1} x^3 \left[ n \right] \right|, \quad \textbf{(7)}$$

$$F_8 = \left| \frac{1}{N} \sum_{n=0}^{N-1} x^4 \left[ n \right] \right|, \quad \textbf{(8)}$$

and

$$F_9 = \left| \frac{1}{N} \sum_{n=0}^{N-1} x^5 \left[ n \right] \right|. \quad \textbf{(9)}$$

(8) Root Mean Square (RMS): The root mean square is another common quantitative measure applied for EMG analysis [9]. It can be given by

$$F_{10} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} x^2 \left[ n \right]}. \quad \textbf{(10)}$$

(9) ν-Order: The ν-order is a nonlinear detector that determines muscle contraction force [9]. It can be given by

$$F_{11} = \left( \frac{1}{N} \sum_{n=0}^{N-1} x^{\nu} \left[ n \right] \right)^{1/\nu} \quad \textbf{(11)}$$

where ν used in this study is 2 which is reported to be an optimal value [9, 12-13].

(10) Log Detector: The log detector is another quantitative measure used for determining muscle contraction force [9, 12-13]. It can be given by

$$F_{12} = e^{\frac{1}{N} \sum_{n=0}^{N-1} \log\left( \left| x[n] \right| \right)}. \quad \textbf{(12)}$$

(11) Waveform Length (WL): The waveform length is a quantitative measure of complexity of EMG signal [9, 14-15]. It is defined as a cumulative length of EMG signal [9]. It can be given by

$$F_{13} = \sum_{n=0}^{N-1} \left| x\left[ n+1 \right] - x\left[ n \right] \right|. \quad \textbf{(13)}$$

(12) Average Amplitude Change (AAC): The average amplitude change is generally equivalent to the waveform length [9]. It can be given by

$$F_{14} = \frac{1}{N} \sum_{n=0}^{N-1} \left| x\left[ n+1 \right] - x\left[ n \right] \right|. \quad \textbf{(14)}$$

(13) Difference Absolute Standard Deviation Value (DASDV): The difference absolute standard deviation value is a standard deviation value of the waveform length [9, 16]. It can be given by

$$F_{15} = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} \left( x\left[ n+1 \right] - x\left[ n \right] \right)^2}. \quad \textbf{(15)}$$

(14) Zero Crossing (ZC): The zero crossing is a quantitative measure quantifying a frequency information of EMG signal [9, 14] It quantifies a number of times that amplitude values of EMG signal cross zero level [9]. It can be given by

$$F_{16} = \sum_{n=1}^{N-1} \left[ \operatorname{sgn}\left( x\left[ n \right] \cdot x\left[ n-1 \right] \right) \cup \left| x\left[ n \right] - x\left[ n-1 \right] \right| \geq 0 \right] \quad \textbf{(16)}$$

where

$$\operatorname{sgn}\left( \alpha \right) = \begin{cases} 1 & \alpha \geq 0 \\ 0 & \textit{otherwise.} \end{cases}$$

(15) Slope Sign Change (SSC): The slope sign change is related to the zero crossing [9, 14]. It is another quantitative measure expressing a frequency information of EMG signal [9]. It can be given by

$$F_{17} = \sum_{n=1}^{N-2} f\left(\left(x[n]-x[n-1]\right)\cdot\left(x[n+1]-x[n]\right)\right) \quad \textbf{(17)}$$

where

$$f(\alpha) = \begin{cases} 1 & \alpha \geq 0 \\ 0 & otherwise. \end{cases}$$

### 2.3 Data Analysis and Classification

All 17 time-domain features denoted by $F_1$, $F_2$, $F_3$, ..., $F_{17}$ are extracted from the EHG segments. A Pearson's correlation coefficient, one of the simplest feature selection methods, [17-19] is applied to rank the time-domain features of EHG segments. The Pearson's correlation coefficient is defined as [17]

$$R(i) = \frac{\operatorname{cov}(x_i, Y)}{\sqrt{\operatorname{var}(x_i)\cdot \operatorname{var}(Y)}}$$

where $x_i$ is the $i$th variable, $Y$ is the corresponding class, and cov(·) and var(·) denote, respectively, the covariance and the variance. In addition, the two-sample $t$-test is also applied to rank the time-domain features of EHG segments by using the corresponding $p$-value.

A number of selected time-domain features of EHG segments with highest ranks, i.e., Pearson's correlation coefficients or $p$-values, are used and applied to feature vectors. The number of selected time-domain features of EHG segments examined in this study ranges from 2 to 9 (around a half of 17 time-domain features). The vectors of selected time-domain features of EHG segments are classified using support vector machine (SVM) classifiers with a kernel function of third-order polynomial. Five values of expected proportion of outliers, that are 0, 0.001, 0.005, 0.01, and 0.05, are applied in training SVM classifiers.

There are 2 preterm-term classifications performed corresponding to the same period of time of recordings. The first preterm-term classification is a classification between the vectors of selected time-domain features of EHG segments associated with preterm or term births recorded in the early period of pregnancy, i.e., PE vs TE. Another preterm-term classification is a classification between the vectors of selected time-domain features of EHG segments associated with preterm or term births recorded in the late period of pregnancy, i.e., PL vs TL. Due to the small number of EHG segments associated with the PE and PL classes, the leave-one-out (LOO) cross validation is applied to the preterm-term classifications.

The performance of the preterm-term birth classifications is evaluated using 3 conventional classification metrics: accuracy $Ac$, sensitivity $Se$, and specificity $Sp$ given by, respectively,

$$Ac = \frac{TP+TN}{TP+TN+FP+FN}, \quad \textbf{(18)}$$

$$Se = \frac{TP}{TP+FN}, \text{ and} \quad \textbf{(19)}$$

$$Sp = \frac{TN}{TN+FP} \quad \textbf{(20)}$$

where $TP$, $TN$, $FP$, and $FN$ denote a number of true positives, a number of true negatives, a number of false positives, and a number of false negatives. Furthermore, a product of sensitivity and specificity, i.e., $Se\times Sp$, is also determined to evaluate the performance on preterm-term classifications. The best performance on preterm-term birth classifications is determined with respect to the product of sensitivity and specificity that justifies both the true positive rate and the true negative rate.

# 3 Results

## 3.1 Ranks of Time-Domain Features of EHG Data

The ranks of time-domain features of EHG segments recorded in the early and late periods of pregnancy obtained using the Pearson's correlation coefficients are summarized in Table 2. Likewise, the ranks of time-domain features of EHG segments recorded in the early and late periods of pregnancy obtained using the *t*-tests are summarized in Table 3. It is shown that the ranks of time-domain features of EHG segments vary corresponding to their channels, the time of recordings, and the feature selection methods. There are however some time-domain features of EHG segments manifesting distinctive characteristics in the feature selections. Table 4 shows the time-domain features of EHG segments that are selected for most channels in each feature selection.

**Table 2.** Ranks of time-domain features of EHG segments using the Pearson's correlation coefficients

| Period of pregnancy | Rank | Channel | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $s_1$ | $s_{1a}$ | $s_{1b}$ | $s_{1c}$ | $s_2$ | $s_{2a}$ | $s_{2b}$ | $s_{2c}$ | $s_3$ | $s_{3a}$ | $s_{3b}$ | $s_{3c}$ |
| Early | 1 | $F_{12}$ | $F_{16}$ | $F_8$ | $F_8$ | $F_7$ | $F_7$ | $F_{17}$ | $F_{17}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ |
| | 2 | $F_7$ | $F_9$ | $F_9$ | $F_9$ | $F_{16}$ | $F_{17}$ | $F_7$ | $F_7$ | $F_{13}$ | $F_{13}$ | $F_{16}$ | $F_{13}$ |
| | 3 | $F_9$ | $F_{17}$ | $F_7$ | $F_{17}$ | $F_5$ | $F_9$ | $F_8$ | $F_8$ | $F_{14}$ | $F_{14}$ | $F_{10}$ | $F_{14}$ |
| | 4 | $F_{13}$ | $F_7$ | $F_{16}$ | $F_{13}$ | $F_8$ | $F_{16}$ | $F_{16}$ | $F_{11}$ | $F_6$ | $F_{11}$ | $F_{11}$ | $F_{10}$ |
| | 5 | $F_{14}$ | $F_8$ | $F_6$ | $F_{14}$ | $F_9$ | $F_8$ | $F_{10}$ | $F_{10}$ | $F_8$ | $F_{10}$ | $F_{13}$ | $F_{11}$ |
| | 6 | $F_8$ | $F_{13}$ | $F_5$ | $F_6$ | $F_1$ | $F_{11}$ | $F_{11}$ | $F_9$ | $F_5$ | $F_5$ | $F_{14}$ | $F_6$ |
| | 7 | $F_{15}$ | $F_{14}$ | $F_{10}$ | $F_5$ | $F_2$ | $F_{10}$ | $F_9$ | $F_1$ | $F_7$ | $F_6$ | $F_6$ | $F_5$ |
| | 8 | $F_5$ | $F_6$ | $F_{11}$ | $F_{16}$ | $F_{11}$ | $F_3$ | $F_1$ | $F_2$ | $F_9$ | $F_4$ | $F_5$ | $F_3$ |
| | 9 | $F_{16}$ | $F_5$ | $F_3$ | $F_{10}$ | $F_{10}$ | $F_1$ | $F_2$ | $F_5$ | $F_{10}$ | $F_3$ | $F_4$ | $F_4$ |
| | 10 | $F_1$ | $F_{15}$ | $F_1$ | $F_{11}$ | $F_6$ | $F_2$ | $F_5$ | $F_6$ | $F_{11}$ | $F_1$ | $F_3$ | $F_1$ |
| | 11 | $F_2$ | $F_{11}$ | $F_2$ | $F_7$ | $F_4$ | $F_4$ | $F_6$ | $F_{16}$ | $F_1$ | $F_2$ | $F_1$ | $F_2$ |
| | 12 | $F_6$ | $F_{10}$ | $F_4$ | $F_{15}$ | $F_3$ | $F_{15}$ | $F_3$ | $F_3$ | $F_2$ | $F_7$ | $F_2$ | $F_{16}$ |
| | 13 | $F_4$ | $F_1$ | $F_{17}$ | $F_3$ | $F_{15}$ | $F_6$ | $F_4$ | $F_4$ | $F_{16}$ | $F_8$ | $F_9$ | $F_9$ |
| | 14 | $F_3$ | $F_2$ | $F_{13}$ | $F_1$ | $F_{13}$ | $F_5$ | $F_{15}$ | $F_{15}$ | $F_3$ | $F_9$ | $F_{17}$ | $F_7$ |
| | 15 | $F_{11}$ | $F_3$ | $F_{14}$ | $F_2$ | $F_{14}$ | $F_{13}$ | $F_{13}$ | $F_{13}$ | $F_4$ | $F_{17}$ | $F_7$ | $F_8$ |
| | 16 | $F_{10}$ | $F_4$ | $F_{15}$ | $F_4$ | $F_{17}$ | $F_{14}$ | $F_{14}$ | $F_{14}$ | $F_{17}$ | $F_{16}$ | $F_8$ | $F_{17}$ |
| | 17 | $F_{17}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ |
| Late | 1 | $F_{16}$ | $F_{17}$ | $F_{17}$ | $F_{16}$ | $F_{15}$ | $F_{17}$ | $F_{16}$ | $F_{16}$ | $F_{13}$ | $F_{17}$ | $F_4$ | $F_{17}$ |
| | 2 | $F_{12}$ | $F_{16}$ | $F_4$ | $F_{17}$ | $F_{13}$ | $F_4$ | $F_{17}$ | $F_{17}$ | $F_{14}$ | $F_3$ | $F_{11}$ | $F_{16}$ |
| | 3 | $F_9$ | $F_4$ | $F_3$ | $F_4$ | $F_{14}$ | $F_{13}$ | $F_4$ | $F_{13}$ | $F_{15}$ | $F_1$ | $F_{10}$ | $F_{10}$ |
| | 4 | $F_8$ | $F_3$ | $F_{16}$ | $F_3$ | $F_7$ | $F_{14}$ | $F_{13}$ | $F_{14}$ | $F_{10}$ | $F_2$ | $F_3$ | $F_{11}$ |
| | 5 | $F_{17}$ | $F_1$ | $F_1$ | $F_1$ | $F_{17}$ | $F_{15}$ | $F_{14}$ | $F_{15}$ | $F_{11}$ | $F_{11}$ | $F_1$ | $F_4$ |
| | 6 | $F_7$ | $F_2$ | $F_2$ | $F_2$ | $F_{16}$ | $F_3$ | $F_3$ | $F_4$ | $F_1$ | $F_{10}$ | $F_2$ | $F_3$ |
| | 7 | $F_{13}$ | $F_{11}$ | $F_{11}$ | $F_{11}$ | $F_5$ | $F_{16}$ | $F_{15}$ | $F_3$ | $F_2$ | $F_4$ | $F_{16}$ | $F_1$ |
| | 8 | $F_{14}$ | $F_{10}$ | $F_{10}$ | $F_{10}$ | $F_9$ | $F_1$ | $F_1$ | $F_1$ | $F_3$ | $F_{16}$ | $F_5$ | $F_2$ |
| | 9 | $F_6$ | $F_8$ | $F_{13}$ | $F_{13}$ | $F_3$ | $F_2$ | $F_2$ | $F_2$ | $F_4$ | $F_5$ | $F_6$ | $F_5$ |
| | 10 | $F_5$ | $F_6$ | $F_{14}$ | $F_{14}$ | $F_4$ | $F_{11}$ | $F_{11}$ | $F_{11}$ | $F_6$ | $F_6$ | $F_{15}$ | $F_6$ |
| | 11 | $F_{10}$ | $F_5$ | $F_{15}$ | $F_6$ | $F_1$ | $F_{10}$ | $F_{10}$ | $F_{10}$ | $F_5$ | $F_{15}$ | $F_{13}$ | $F_{15}$ |
| | 12 | $F_{11}$ | $F_{13}$ | $F_6$ | $F_5$ | $F_2$ | $F_5$ | $F_6$ | $F_6$ | $F_8$ | $F_{13}$ | $F_{14}$ | $F_7$ |
| | 13 | $F_4$ | $F_{14}$ | $F_5$ | $F_{15}$ | $F_{10}$ | $F_6$ | $F_5$ | $F_5$ | $F_{16}$ | $F_{14}$ | $F_8$ | $F_8$ |
| | 14 | $F_{15}$ | $F_{15}$ | $F_8$ | $F_8$ | $F_{11}$ | $F_8$ | $F_8$ | $F_8$ | $F_{12}$ | $F_8$ | $F_7$ | $F_{13}$ |
| | 15 | $F_3$ | $F_7$ | $F_7$ | $F_7$ | $F_8$ | $F_7$ | $F_7$ | $F_7$ | $F_9$ | $F_9$ | $F_9$ | $F_{14}$ |
| | 16 | $F_1$ | $F_9$ | $F_9$ | $F_9$ | $F_6$ | $F_9$ | $F_9$ | $F_9$ | $F_7$ | $F_7$ | $F_{17}$ | $F_9$ |
| | 17 | $F_2$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{17}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ |

**Table 3.** Ranks of time-domain features of EHG segments using the *t*-test

| Period of pregnancy | Rank | Channel | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $s_1$ | $s_{1a}$ | $s_{1b}$ | $s_{1c}$ | $s_2$ | $s_{2a}$ | $s_{2b}$ | $s_{2c}$ | $s_3$ | $s_{3a}$ | $s_{3b}$ | $s_{3c}$ |
| | 1 | $F_{15}$ | $F_{16}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{16}$ | $F_{16}$ | $F_{16}$ |
| | 2 | $F_{17}$ | $F_{15}$ | $F_{11}$ | $F_{13}$ | $F_{17}$ | $F_{13}$ | $F_{13}$ | $F_{13}$ | $F_{17}$ | $F_{17}$ | $F_{15}$ | $F_{15}$ |
| | 3 | $F_{16}$ | $F_{13}$ | $F_{10}$ | $F_{14}$ | $F_{11}$ | $F_{14}$ | $F_{14}$ | $F_{14}$ | $F_{16}$ | $F_{15}$ | $F_7$ | $F_{13}$ |
| | 4 | $F_{10}$ | $F_{14}$ | $F_1$ | $F_{11}$ | $F_{10}$ | $F_{11}$ | $F_{10}$ | $F_{10}$ | $F_1$ | $F_{13}$ | $F_8$ | $F_{14}$ |
| | 5 | $F_{11}$ | $F_{11}$ | $F_2$ | $F_{10}$ | $F_1$ | $F_{10}$ | $F_{11}$ | $F_{11}$ | $F_2$ | $F_{14}$ | $F_9$ | $F_{17}$ |
| | 6 | $F_1$ | $F_{10}$ | $F_3$ | $F_1$ | $F_2$ | $F_3$ | $F_1$ | $F_1$ | $F_3$ | $F_8$ | $F_5$ | $F_7$ |
| | 7 | $F_2$ | $F_1$ | $F_4$ | $F_2$ | $F_{12}$ | $F_1$ | $F_2$ | $F_2$ | $F_4$ | $F_9$ | $F_6$ | $F_8$ |
| | 8 | $F_3$ | $F_2$ | $F_7$ | $F_3$ | $F_3$ | $F_2$ | $F_3$ | $F_3$ | $F_{11}$ | $F_7$ | $F_1$ | $F_9$ |
| Early | 9 | $F_4$ | $F_3$ | $F_8$ | $F_4$ | $F_6$ | $F_4$ | $F_4$ | $F_4$ | $F_{10}$ | $F_5$ | $F_2$ | $F_5$ |
| | 10 | $F_{13}$ | $F_4$ | $F_9$ | $F_6$ | $F_4$ | $F_6$ | $F_6$ | $F_6$ | $F_6$ | $F_6$ | $F_3$ | $F_6$ |
| | 11 | $F_{14}$ | $F_6$ | $F_5$ | $F_5$ | $F_5$ | $F_5$ | $F_5$ | $F_5$ | $F_5$ | $F_4$ | $F_4$ | $F_1$ |
| | 12 | $F_{12}$ | $F_5$ | $F_6$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_3$ | $F_{11}$ | $F_2$ |
| | 13 | $F_6$ | $F_8$ | $F_{13}$ | $F_8$ | $F_8$ | $F_8$ | $F_8$ | $F_8$ | $F_8$ | $F_1$ | $F_{10}$ | $F_3$ |
| | 14 | $F_5$ | $F_9$ | $F_{14}$ | $F_9$ | $F_9$ | $F_9$ | $F_9$ | $F_9$ | $F_9$ | $F_2$ | $F_{13}$ | $F_4$ |
| | 15 | $F_7$ | $F_7$ | $F_{16}$ | $F_{17}$ | $F_{13}$ | $F_{17}$ | $F_{17}$ | $F_{17}$ | $F_{12}$ | $F_{10}$ | $F_{14}$ | $F_{11}$ |
| | 16 | $F_8$ | $F_{17}$ | $F_{17}$ | $F_{16}$ | $F_{14}$ | $F_{16}$ | $F_{16}$ | $F_{16}$ | $F_{13}$ | $F_{11}$ | $F_{17}$ | $F_{10}$ |
| | 17 | $F_9$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{16}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{14}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ |
| | 1 | $F_3$ | $F_8$ | $F_7$ | $F_{17}$ | $F_5$ | $F_8$ | $F_{17}$ | $F_7$ | $F_6$ | $F_9$ | $F_8$ | $F_9$ |
| | 2 | $F_1$ | $F_9$ | $F_6$ | $F_7$ | $F_7$ | $F_9$ | $F_7$ | $F_4$ | $F_7$ | $F_7$ | $F_9$ | $F_7$ |
| | 3 | $F_2$ | $F_{17}$ | $F_5$ | $F_6$ | $F_8$ | $F_7$ | $F_{15}$ | $F_{10}$ | $F_{16}$ | $F_8$ | $F_{13}$ | $F_8$ |
| | 4 | $F_4$ | $F_7$ | $F_1$ | $F_5$ | $F_6$ | $F_6$ | $F_{13}$ | $F_{11}$ | $F_8$ | $F_5$ | $F_{14}$ | $F_5$ |
| | 5 | $F_{10}$ | $F_6$ | $F_2$ | $F_1$ | $F_9$ | $F_5$ | $F_{14}$ | $F_1$ | $F_5$ | $F_6$ | $F_7$ | $F_6$ |
| | 6 | $F_{11}$ | $F_5$ | $F_{10}$ | $F_2$ | $F_{10}$ | $F_{15}$ | $F_{10}$ | $F_2$ | $F_9$ | $F_{16}$ | $F_{16}$ | $F_{13}$ |
| | 7 | $F_5$ | $F_4$ | $F_{11}$ | $F_4$ | $F_{11}$ | $F_{13}$ | $F_{11}$ | $F_{15}$ | $F_{10}$ | $F_1$ | $F_{17}$ | $F_{14}$ |
| | 8 | $F_{12}$ | $F_1$ | $F_4$ | $F_3$ | $F_1$ | $F_{14}$ | $F_4$ | $F_3$ | $F_{11}$ | $F_2$ | $F_5$ | $F_{16}$ |
| Late | 9 | $F_8$ | $F_2$ | $F_3$ | $F_{10}$ | $F_2$ | $F_{17}$ | $F_1$ | $F_{13}$ | $F_{12}$ | $F_3$ | $F_6$ | $F_4$ |
| | 10 | $F_{15}$ | $F_3$ | $F_{16}$ | $F_{11}$ | $F_4$ | $F_{10}$ | $F_2$ | $F_{14}$ | $F_1$ | $F_{10}$ | $F_4$ | $F_{15}$ |
| | 11 | $F_{13}$ | $F_{10}$ | $F_9$ | $F_9$ | $F_3$ | $F_{11}$ | $F_3$ | $F_9$ | $F_2$ | $F_{11}$ | $F_{15}$ | $F_{17}$ |
| | 12 | $F_{14}$ | $F_{11}$ | $F_{17}$ | $F_{16}$ | $F_{16}$ | $F_3$ | $F_9$ | $F_{17}$ | $F_{13}$ | $F_4$ | $F_1$ | $F_1$ |
| | 13 | $F_9$ | $F_{15}$ | $F_8$ | $F_8$ | $F_{13}$ | $F_4$ | $F_8$ | $F_8$ | $F_{14}$ | $F_{13}$ | $F_2$ | $F_2$ |
| | 14 | $F_7$ | $F_{16}$ | $F_{13}$ | $F_{15}$ | $F_{14}$ | $F_1$ | $F_6$ | $F_6$ | $F_3$ | $F_{14}$ | $F_3$ | $F_3$ |
| | 15 | $F_6$ | $F_{13}$ | $F_{14}$ | $F_{13}$ | $F_{15}$ | $F_2$ | $F_5$ | $F_5$ | $F_4$ | $F_{17}$ | $F_{10}$ | $F_{11}$ |
| | 16 | $F_{16}$ | $F_{14}$ | $F_{15}$ | $F_{14}$ | $F_{12}$ | $F_{16}$ | $F_{16}$ | $F_{16}$ | $F_{17}$ | $F_{15}$ | $F_{11}$ | $F_{10}$ |
| | 17 | $F_{17}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{17}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ | $F_{15}$ | $F_{12}$ | $F_{12}$ | $F_{12}$ |

**Table 4.** The highly selected time-domain features of EHG segments

| Feature selection method | Period of pregnancy | Number of selected features | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Pearson's correlation coefficient | Early | $F_7$ | $F_7$ | $F_7$ | $F_8$ | $F_8$ | $F_8$ | $F_8$, $F_9$ | $F_5$, $F_{10}$ |
| | Late | $F_{17}$ | $F_{17}$ | $F_{17}$ | $F_{17}$ | $F_{17}$ | $F_{16}$, $F_{17}$ | $F_{16}$ | $F_3$, $F_{16}$ |
| *t*-test | Early | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ | $F_{15}$ |
| | Late | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ | $F_7$ |

The absolute values of the 3rd, 4th and 5th temporal moments ($F_7$, $F_8$ and $F_9$), the simple square integral ($F_5$), and the root mean square ($F_{10}$) are the time-domain features of EHG segments recorded in the early period of pregnancy that are highly selected using the Pearson's correlation coefficients while the slope sign change ($F_{17}$), the zero crossing ($F_{16}$), and the modified mean absolute value (type 1) ($F_3$) are the time-domain features of EHG segments recorded in the late period of pregnancy using the Pearson's correlation coefficients. On the other hand, the time-domain features of EHG segments recorded in the early and late period of pregnancy that are highly selected using the *t*-tests are, respectively, the difference absolute standard deviation value ($F_{15}$) and the absolute value of the 3rd temporal moment ($F_7$).

Furthermore, the modified mean absolute value (type 1) ($F_3$), the modified mean absolute value (type 2) ($F_4$), and the variance ($F_6$) are the time-domain features of EHG segments that are least selected using the

Pearson's correlation coefficients. The modified mean absolute value (type 2) ($F_4$), the variance ($F_6$), and the log detector ($F_{12}$) are the time-domain features of EHG segments that are least selected using the *t*-tests.

### 3.2 Performance on Preterm-Term Birth Classifications

The selected time-domain features of EHG segments recorded in the early and late periods of pregnancy are applied for preterm-term birth classifications. The best performances on preterm-term birth classifications for various numbers of selected time-domain features of EHG segments recorded in the early period of pregnancy are summarized in Table 5. The accuracy, sensitivity, and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period of pregnancy that are selected using the Pearson's correlation coefficients range between 0.5124 and 0.9259, 0.4211 and 0.8421, and 0.4825 and 0.9930, respectively. The corresponding products of sensitivity and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period range between 0.3522 and 0.4593.

**Table 5.** The best performance on preterm-term birth classifications of selected time-domain features of EHG segments recorded in the early period of pregnancy

| Feature selection method | Number of selected features | Channel | Outlier | Accuracy *Ac* | Sensitivity *Se* | Specificity *Sp* | $Se \times Sp$ |
|---|---|---|---|---|---|---|---|
| Pearson's correlation coefficient | 2 | $s_1$ | 0 | 0.9259 | 0.4211 | 0.9930 | 0.4181 |
| | 3 | $s_3$ | 0.001 | 0.6111 | 0.7368 | 0.5944 | 0.4380 |
| | 4 | $s_{3b}$ | 0 | 0.5802 | 0.8421 | 0.5455 | 0.4593 |
| | 5 | $s_{2a}$ | 0.001 | 0.6049 | 0.5789 | 0.6084 | 0.3522 |
| | 6 | $s_{2b}$ | 0.01 | 0.6049 | 0.5789 | 0.6084 | 0.3522 |
| | 7 | $s_3$ | 0 | 0.6296 | 0.5789 | 0.6364 | 0.3684 |
| | 8 | $s_{3a}$ | 0.01 | 0.7037 | 0.5263 | 0.7273 | 0.3828 |
| | 9 | $s_{3c}$ | 0.05 | 0.5123 | 0.7368 | 0.4825 | 0.3555 |
| *t*-test | 2 | $s_2$ | 0.01 | 0.6852 | 0.6316 | 0.6923 | 0.4372 |
| | 3 | $s_{3a}$ | 0.05 | 0.6481 | 0.5789 | 0.6573 | 0.3806 |
| | 4 | $s_{1a}$ | 0.001 | 0.6173 | 0.6316 | 0.6154 | 0.3887 |
| | 5 | $s_{1b}$ | 0 | 0.5432 | 0.6842 | 0.5245 | 0.3589 |
| | 6 | $s_{3a}$ | 0.01 | 0.7901 | 0.4737 | 0.8322 | 0.3942 |
| | 7 | $s_{3b}$ | 0.005 | 0.6111 | 0.6316 | 0.6084 | 0.3842 |
| | 8 | $s_{1a}$ | 0.01 | 0.5988 | 0.4737 | 0.6154 | 0.2915 |
| | 8 | $s_{2a}$ | 0.01 | 0.5988 | 0.4737 | 0.6154 | 0.2915 |
| | 9 | $s_{2c}$ | 0.01 | 0.6667 | 0.7895 | 0.6503 | 0.5134 |

Similarly, for the time-domain features of EHG segments selected using the *t*-tests, from Table 5 the accuracy, sensitivity, and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period of pregnancy range, respectively, between 0.5432 and 0.7901, 0.4737 and 0.7895, and 0.5245 and 0.8322 while the corresponding products of sensitivity and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period of pregnancy range between 0.2915 and 0.5134.

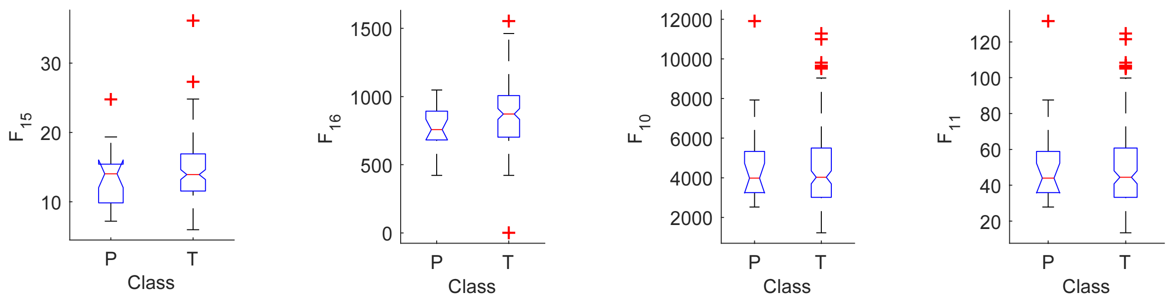
Table 6 summarizes the best performances on preterm-term birth classifications for various numbers of selected time-domain features of EHG segments recorded in the late period of pregnancy. The accuracy, sensitivity, and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the late period of pregnancy that are selected using the Pearson's correlation coefficients range, respectively, between 0.4815 and 0.6790, 0.3158 and 0.6316, and 0.4685 and 0.7273. The range of the corresponding products of sensitivity and specificity of preterm-term birth classifications is from 0.2208 to 0.3489.

The accuracy, sensitivity, and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the late period of pregnancy that are selected using the *t*-tests shown in Table 6 range between 0.5000 and 0.6543, 0.4211 and 0.7895, and 0.4615 and 0.6783, respectively. The corresponding products of sensitivity and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the late period of pregnancy range between 0.2856 and 0.3725.

**Table 6.** The best performance on preterm-term birth classifications of selected time-domain features of EHG segments recorded in the late period of pregnancy

| Feature selection method | Number of selected features | Channel | Outlier | Accuracy *Ac* | Sensitivity *Se* | Specificity *Sp* | *Se*×*Sp* |
|---|---|---|---|---|---|---|---|
| Pearson's correlation coefficient | 2 | $s_{3c}$ | 0.001 | 0.5617 | 0.6316 | 0.5524 | 0.3489 |
| | 3 | $s_1$ | 0.01 | 0.4815 | 0.5789 | 0.4685 | 0.2713 |
| | 4 | $s_2$ | 0 | 0.5000 | 0.6316 | 0.4825 | 0.3047 |
| | 5 | $s_{2c}$ | 0.05 | 0.6790 | 0.3684 | 0.7203 | 0.2654 |
| | 6 | $s_{3a}$ | 0.005 | 0.6111 | 0.3684 | 0.6434 | 0.2370 |
| | 7 | $s_3$ | 0.01 | 0.5123 | 0.4211 | 0.5245 | 0.2208 |
| | 8 | $s_{3b}$ | 0.05 | 0.5802 | 0.3684 | 0.6084 | 0.2241 |
| | 9 | $s_{1b}$ | 0 | 0.6790 | 0.3158 | 0.7273 | 0.2297 |
| *t*-test | 2 | $s_{2b}$ | 0.05 | 0.5000 | 0.7895 | 0.4615 | 0.3644 |
| | 3 | $s_{3c}$ | 0.05 | 0.5062 | 0.6842 | 0.4825 | 0.3301 |
| | 4 | $s_{2a}$ | 0 | 0.6358 | 0.5789 | 0.6434 | 0.3725 |
| | 5 | $s_{1a}$ | 0.001 | 0.5000 | 0.6316 | 0.4825 | 0.3047 |
| | 6 | $s_{1b}$ | 0 | 0.5432 | 0.5789 | 0.5385 | 0.3117 |
| | 7 | $s_{3a}$ | 0.05 | 0.6543 | 0.4737 | 0.6783 | 0.3213 |
| | 8 | $s_{3a}$ | 0.005 | 0.6481 | 0.4211 | 0.6783 | 0.2856 |
| | 9 | $s_1$ | 0.05 | 0.5309 | 0.6842 | 0.5105 | 0.3493 |

The best product of sensitivity and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period of pregnancy that are selected using the Pearson's correlation coefficient is 0.4593. This is obtained from the preterm-term classification using the SVM classifier with 4 time-domain features of EHG segments, i.e., the difference absolute standard deviation value, the zero crossing, the root mean square, and the ν-order, corresponding to the channel $s_{3b}$ and without an outlier. The distributions of the difference absolute standard deviation value, the zero crossing, the root mean square, and the ν-order of EHG segments recorded in the early period of pregnancy associated with preterm and term births are compared in Fig. 3(a) to Fig. 3(d), respectively.




(a) DASDV (b) ZC (c) RMS (d) ν-order

**Fig. 3.** Box plots of time-domain features of EHG segments recorded in the early period of pregnancy associated with preterm and term births selected using the Pearson's correlation coefficient

On the other hand, the best product of sensitivity and specificity of preterm-term birth classifications of time-domain features of EHG segments recorded in the early period of pregnancy that are selected using the *t*-test is 0.5134. This is obtained from the preterm-term classification using the SVM classifier with 9 time-domain features of EHG segments, i.e., the difference absolute standard deviation value, the waveform length, the average amplitude change, the root mean square, the ν-order, the integrated EMG, the mean absolute value, the modified mean absolute value (type 1), and the modified mean absolute value (type 2), corresponding to channel $s_{2c}$ and an outlier of 0.01. Fig. 4(a) to Fig. 4(k) compare, respectively, the difference absolute standard deviation value, the waveform length, the average amplitude change, the root mean square, the ν-order, the integrated EMG, the mean absolute value, the modified mean absolute value (type 1), and the modified mean absolute value (type 2) of EHG segments recorded in the early period of pregnancy associated with preterm and term births.

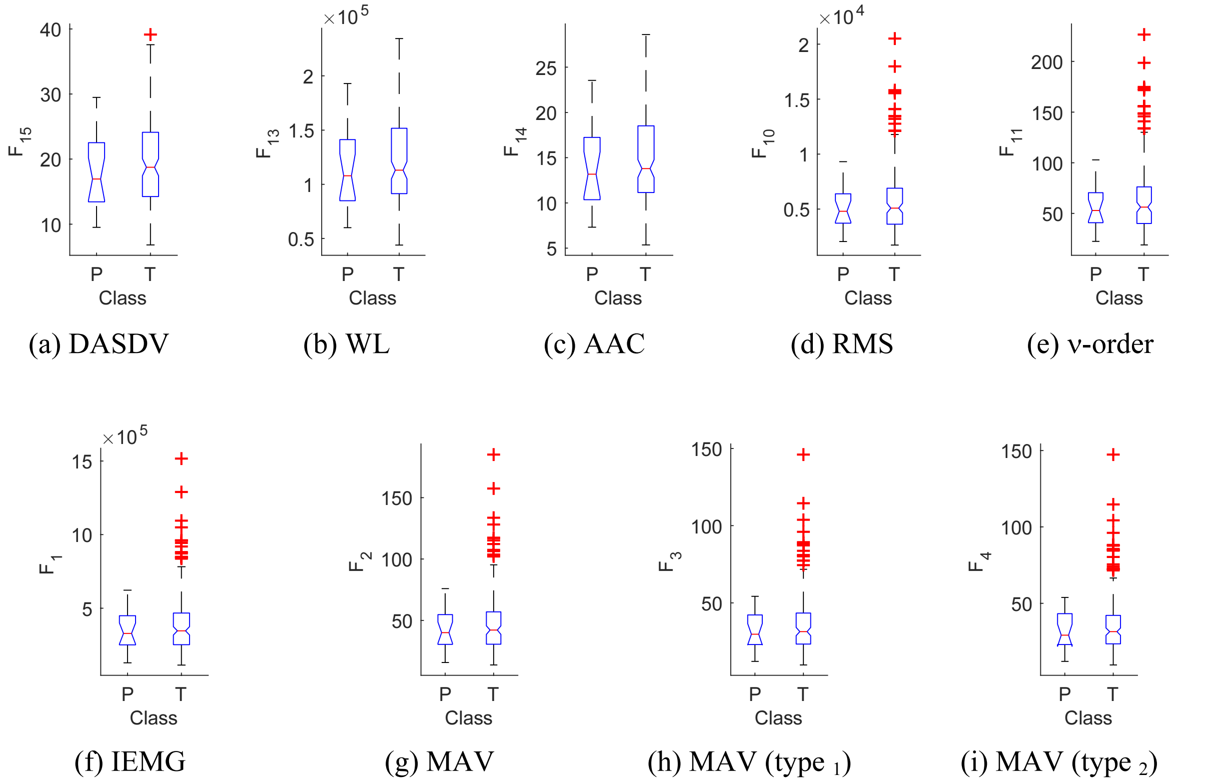



**Fig. 4.** Box plots of time-domain features of EHG segments recorded in the early period of pregnancy associated with preterm and term births selected using the *t*-test

For the EHG segments recorded in the late period of pregnancy, the best product of sensitivity and specificity of preterm-term birth classification using the time-domain features selected using the Pearson's correlation coefficient is 0.3489. This is obtained using the SVM classifier with 2 time-domain features of EHG segments, i.e., the difference absolute standard deviation value and the waveform length, corresponding to channel $s_{3c}$ and an outlier of 0.001. The difference absolute standard deviation value and the waveform length of EHG segments recorded in the late period of pregnancy associated with preterm and term births are compared in Fig. 5(a) to Fig. 5(b), respectively.

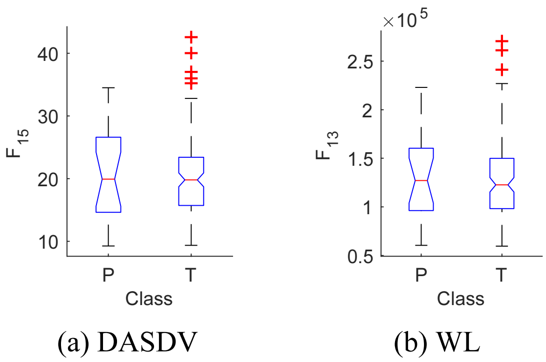
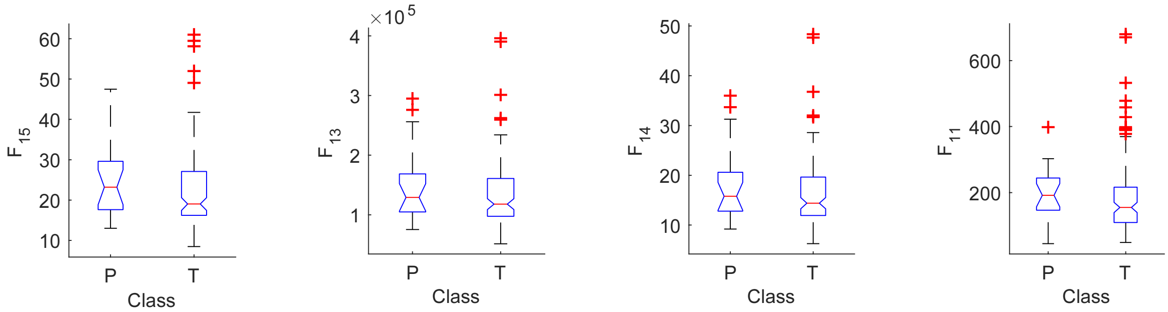



**Fig. 5.** Box plots of time-domain features of EHG segments recorded in the late period of pregnancy associated with preterm and term births selected using the Pearson's correlation coefficient

However, the best product of sensitivity and specificity of preterm-term birth classification using the time-domain features selected using the *t*-test is 0.3725. This is obtained using the SVM classifier with 4 time-domain features of EHG segments, i.e., the difference absolute standard deviation value, the waveform length, the average amplitude change and the ν-order, corresponding to channel $s_{3c}$ and without an outlier. The comparison of the difference absolute standard deviation value, the waveform length, the average amplitude change and the ν-order of EHG segments recorded in the late period of pregnancy associated with preterm and term births are, respectively, shown in Fig. 6(a) to Fig. 6(d).




(a) DASDV (b) WL (c) AAC (d) ν-order

**Fig. 6.** Box plots of time-domain features of EHG segments recorded in the late period of pregnancy associated with preterm and term births selected using the *t*-test

## 4 Discussion

From the computational results, it is shown that there are no obvious tendencies on a number of selected time-domain features of EHG segments, a channel of EHG data, and also an outlier applied to SVM classifiers for obtaining a better performance on preterm-term birth classification. For the EHG segments recorded in the early period of pregnancy, the time-domain features of EHG segments selected using the $t$-test provide the better performance on preterm-term birth classifications compare those selected using the Pearson's correlation coefficient when 2, 5, 6, 7, and 9 time-domain features selected. The time-domain features of EHG segments recorded in the late period of pregnancy selected using the *t*-test provide the better performance compare to those selected using the Pearson's correlation coefficient for all preterm-term birth classifications.

The best performance obtained from the preterm-term birth classifications of selected time-domain features of EHG segments recorded in the early period of pregnancy is better than the best performance obtained from the preterm-term birth classifications of selected time-domain features of EHG segments recorded in the late period of pregnancy. The best performance on the preterm-term birth classifications of selected time-domain features of EHG segments recorded in the early period of pregnancy is with the product of sensitivity and specificity of 0.5134 (the accuracy of 0.6667, the sensitivity of 0.7895, and the specificity of 0.6503). The EHG segments providing such best performance on the preterm-term birth classifications correspond to the 0.3-4.0-Hz subbands of channel $s_2$. The best performance on the preterm-term birth classifications of selected time-domain features of EHG segments recorded in the late period of pregnancy is with the product of 0.3725 (the accuracy of 0.6358, the sensitivity of 0.5789, and the specificity of 0.6434). The EHG segments providing such best performance on the preterm-term birth classifications correspond to the 0.08-4.0-Hz subbands of channel $s_2$.

These best performances on preterm-term birth classifications are furthermore better than the previous results obtained from the preterm-term birth classifications using the wavelet-based features of EHG segments recorded in the early and late periods of pregnancy [5] where the corresponding products of sensitivity and specificity are, respectively, 0.4880 and 0.3702.

## 5 Conclusions

In this study, seventeen common time-domain features of EMG data are applied to EHG or uterine EMG data recorded from pregnant women who delivered on term or prematurely. Two feature selection techniques, i.e., Pearson's correlation coefficient and *t*-test, are used to reduce the dimension of time-domain features of EHG data and then applied for preterm-term birth classifications using SVM classifiers. Furthermore, the leave-one-out cross validations are used to evaluate the performance of preterm-term birth classifications of selected time-domain features of EHG data recorded in the early and late period of pregnancy. From the computational results, it is shown that the best performance on preterm-term birth classifications of selected time-domain features of EHG data recorded in the early period of pregnancy with respect to the product of sensitivity and specificity is with the accuracy of 0.6667, the sensitivity of 0.7895 and the specificity of 0.6503 while the best performance on preterm-term birth classifications of selected time-domain features of EHG data recorded in the late period of pregnancy with respect to the product of sensitivity and specificity is with the accuracy of 0.6358, the sensitivity of 0.5789 and the specificity of 0.6434. The computational results also suggest that the difference absolute standard deviation value, the waveform length, the average amplitude change and the ν-order are the key time-domain features of EHG data for the preterm-term classifications.

## Acknowledgments

This work is supported by a TRF Research Career Development Grant, jointly funded by the Thailand Research Fund (TRF) and the Ubon Ratchathani University, under the Contract No. RSA6180041.

## References

[1] G. FeleŽorž, G. Kavšek, Ž. Novak-Antolic, F. Jager, A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups, Med. Biol. Eng. Comput. 46(2008) 911-922.

[2] W.L. Maner, R.E. Garfield, Identification of human term and preterm labor using artificial neural networks on uterine electromyography data, Ann. Biomed. Eng. 35(2007) 465-473.

[3] H. Leman, C. Marque, J. Gondry, Use of the electrohysterogram signal for characterization of contractions during pregnancy, IEEE Trans. Biomed. Eng. (46)(1999) 1222-1229.

[4] Mayo Clinic, Premature birth. <https://www.mayoclinic.org/diseases-conditions/premature-birth/symptoms-causes/syc-20376730>, 2017 (accessed 31.08.18).

[5] S. Janjarasjitt, Examination of single wavelet-based features of EHG signals for preterm birth classification, IAENG International Journal of Computer Science 44(2017) 212-218.

[6] S. Janjarasjitt, Evaluation of performance on preterm birth classification using single wavelet-based features of EHG signals, in: Proc. the 10th Biomedical Engineering International Conference, 2017.

[7] P. Ren, S. Yao, J. Li, P.A. Valdes-Sosa, K.M. Kendrick, Improved prediction of preterm delivery using empirical mode decomposition analysis of uterine electromyography signals, PLoS ONE 10(2015) e0132116.

[8] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101(2000) e215-e220.

[9] A. Phinyomark, P. Phukpattaranont, C. Limsakul, Feature reduction and selection for EMG signal classification, Expert Syst. Appl. 39(2012) 7420-7431.

[10] S. Du, M. Vuskovic, Temporal vs. spectral approach to feature extraction from prehensile EMG signals, in: Proc. IEEE International Conference on Information Reuse and Integration, 2004.

[11] G.N. Saridis, T.P. Gootee, EMG pattern analysis and classification for a prosthetic arm, IEEE Trans. Biomed. Eng. 29(1982) 403-412.

[12] D. Tkach, H. Huang, T.A. Kuiken, Study of stability of time-domain features for electromyographic pattern recognition, Journal of Neuro Engineering and Rehabilitation 7(2010). doi:10.1186/1743-0003-7-21.

[13] M. Zardoshti-Kermani, B.C. Wheeler, K. Badie, R.M. Hashemi, EMG feature evaluation for movement control of upper extremity prostheses, IEEE Trans. Rehabil. Eng. 3(1995) 324-333.

[14] B. Hudgins, P. Parker, R. Scott, A new strategy for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 40(1993) 82-94.

[15] M.A. Oskoei, H. Hu, Support vector machine based classification scheme for myoelectric control applied to upper limb, IEEE Trans. Biomed. Eng. 55(2008) 1956-1965.

[16] K.S. Kim, H.H. Choi, C.S. Moon, C.W. Mun, Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions, Current Applied Physics 11(2011) 740-745.

[17] G. Chandrashekar, F. Sahin, A survey on feature selection methods, Comput. Electr. Eng. 40(2014) 16-28.

[18] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3(2003) 1157-1182.

[19] R. Battiti, Using mutual information for selecting features in supervised neural net learning, IEEE Trans. Neural. Networks 5(1994) 537-550.