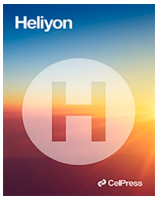
Heliyon 10 (2024) e40433

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

# Characteristics of phase synchronization in electrohysterography and tocodynamometry for preterm birth prediction

Jae-Hwan Kang [a,b], Young-Ju Jeon [a,b], In-Seon Lee [c], Junsuk Kim [d,*]

[a] Digital Health Research Division, Korea Institute of Oriental Medicine, Daejeon, South Korea
[b] Aging Convergence Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, South Korea
[c] College of Korean Medicine, Kyung Hee University, Seoul, South Korea
[d] School of Information Convergence, Kwangwoon University, Seoul, South Korea

ARTICLE INFO

*Keywords:*
Electrohysterogram
Preterm birth prediction
Phase synchronization
Mean phase coherence

ABSTRACT

Preterm birth prediction is important in prenatal care; however, it remains a significant challenge due to the complex physiological mechanisms involved. This study aimed to explore the feasibility of phase synchronization of multiple oscillatory components across electrohysterography (EHG) and tocodynamometry (TOCO) signals to identify preterm births using advanced machine-learning techniques. Using an open-access EHG dataset, we first assessed the degree of phase synchronization of five specified frequency ranges from 0.08 to 5.0 Hz in three individual EHG signals by constructing two distinct sets of mean phase coherence: the inclusion or exclusion of TOCO signals. We then employed two machine-learning models, XGBoost and TabNet, to classify preterm and term delivery conditions and analyze the predictive potential of these features. The models' performance was evaluated by considering varying lengths of time windows and the use of overlapping windows. Our results demonstrate the importance of lower-frequency EHG signals and synchronization patterns across the horizontal plane of the abdomen, particularly synchronization between the upper and lower regions of the uterus. Furthermore, we observed a distinctive pattern in the high-frequency band (1.0–2.2 Hz), emphasizing the important role of the lower horizontal regions with other sites in the synchronization process. Interestingly, our findings indicated that TOCO signals, while not substantially enhancing the overall prediction performance, contributed to slightly improved accuracy rates when combined with EHG signals. This study suggests the critical role of EHG signals and their intricate spatiotemporal patterns in predicting preterm birth, providing insights for the development of more accurate and efficient prediction models.

## 1. Introduction

Preterm birth (PTB), defined as premature delivery before 37 weeks of gestation [1], poses a serious risk to both neonates and pregnant women [2]. Early prediction of PTB is one of the most promising research topics, as it can reduce the rate of neonatal mortality and prevent serious pregnancy complications [3]. However, it is hard to predict the exact time of PTB [4]. Although there are various risk factors such as maternal behaviors (tobacco, alcohol, and drug abuse) and medical conditions (obesity, diabetes, and heart

* Corresponding author. School of Information Convergence, College of AI Convergence, Kwangwoon University, Nowon-gu, 01897, Seoul, South Korea.
*E-mail address:* junsuk.kim@kw.ac.kr (J. Kim).

https://doi.org/10.1016/j.heliyon.2024.e40433
Received 13 January 2024; Received in revised form 7 November 2024; Accepted 13 November 2024
Available online 15 November 2024

disease) related to PTB, as well as an important index such as the shortening cervical length assessed by transvaginal ultrasound at about 20 weeks of gestation, there are still no rigid standard parameters to identify PTB [4]. For the purpose of reliable PTB prediction, a variety of studies have paid attention to the characteristics of uterine contractility for the examination of the functional status of the uterus during pregnancy [5–7]. In this respect, there are two types of important time-series signals for uterus contractility [8,9]. One is the tocodynamometry (TOCO), and the other is the electrohysterography (EHG) signals. The external TOCO signals are time-varying mechanical pressures of the uterus recorded by a strain gauge belt against the pregnant woman's abdomen. It has been widely used for the detection of uterine contractions and labor progression in pregnancy [10]. However, there are still some issues about the feasibility of TOCO signals in PTB with low performance due to the high sensitivity of maternal body mass index (BMI) and movements [11,12]. The EHG signals, also known as uterine electromyography signals, are time-varying records that capture changes in the electrical activity of the uterine muscles [13,14]. These signals are gathered using several electrodes placed on the abdomens of pregnant women. Compared to the TOCO, the EHG is a more prominent long-term monitoring system for the prediction of PTB due to the advantage of not only noninvasive recording but also relatively lower-sensitivity to BMI [10,11]. In this study, we aimed to propose a new analytic approach to PTB by assessing the degree of synchronization in several spectral phases' information across multiple EHG signals and a TOCO signal. Over the decades, many preterm-related EHG signaling studies have proposed specific useful EHG features that not only reflect the condition of the uterus in terms of pregnancy and parturition [8,15] but also serve as vital inputs for machine- or deep-learning models to evaluate the health status of pregnant women and their babies or to predict the likelihood of PTB. To summarize the main types of EHG features proposed so far, we can broadly categorize them into the following four groups: (1) temporal, (2) spectral, (3) nonlinear, and (4) propagation velocity parameters, regardless of whether single or multiple EHG signals are considered.

First, temporal features are defined as quantified temporal characteristics of the raw or preprocessed single-channel EHG signals without considering predefined, specific frequency ranges. The most commonly used temporal parameters are the peak or median amplitude and the root mean square (RMS) value of the absolute EHG signal [16]. Several studies on PTB have employed the RMS values because they are highly related to the strength of the EHG contraction and burst [15,16]. However, these amplitude-based temporal parameters can easily be contaminated by artifacts originating from the individual conditions of a woman's tissue and impedance changes in the attached EHG electrodes.

Second, spectral power features within some specific frequency ranges have been traditionally employed to predict PTB. In particular, the narrow frequency range of 0.34–1.0 Hz of the EHG signal, which has been named 'uterine-specific' frequency, has provided a useful information on the status of uterine activity by avoiding some unwanted artifacts from maternal respiratory and heart activities [17,18]. Most studies on PTB have widely used the 0.1–3.0 Hz of the EHG signals [19–22]. This range of frequency is also divided into two oscillatory signals: 0.1–0.6 Hz (fast wave low, FWL) and 0.6–3 Hz (fast wave high, FWH) [14]. It has been reported that the FWL corresponds to the propagation of electrical activity and the FWH corresponds to the excitability of the uterus [14,23]. Moreover, many studies have extended to regions of interest with frequency ranges up to below 16 Hz [24].

Third, nonlinear or nonstationary analyses, free from any type of linear model to quantify time-series signals, have also been used to represent the temporal characteristics of EHG signals. For instance, several entropy-based techniques, such as sampling, permutation, wavelet, and multivariate multiscale entropy, estimate the corresponding entropy of single- or multichannel EHG signals. These techniques measure the stability, predictability, and complexity of time-series signals [25–27]. Sample entropy, particularly notable for its ability to measure the complexity of finite-length time-series signals, has proven to be effective in distinguishing between term and preterm deliveries [26,27]. Statistical analyses have revealed that sample entropy is typically higher in term deliveries than in preterm deliveries and progressively decreases as gestation progresses [26].

Finally, regarding propagation velocity parameters, a bundle of EHG studies have intensively investigated changes in the conduction velocity and direction of EHG bursts originating from action potentials in the uterine muscles [28–30]. For example, Rabotti et al. obtained both row and column continuous delays from 64 channel (arranged in an $8 \times 8$ matrix) EHG signals by adopting a fast Fourier transform and maximum likelihood estimation [28,30]. Using these continuous two-dimensional delays, the authors were able to estimate the conduction velocity of EHG bursts in the uterus, revealing that the conduction velocity in the labor group was faster than that in the non-labor group [28]. Another simulation study improved our understanding of the physiology of the pregnant uterus in relation to preterm labor detection by designing three types of models: electrical, force, and deformation. These models simulate the electrical activity at the cellular level, the mechanical force at the muscular level, and the deformation model between them in the entire uterus [31]. Another study emphasized the capability of EHG wave propagation conduction velocity to distinguish between preterm and term deliveries. This was achieved by adopting the velocities in both the vertical and horizontal directions, as well as the sample entropy of the EHG signals [32].

As mentioned above, various EHG features have been proposed to predict PTB and monitor uterine health in pregnant women. Recent studies have focused on exploring the relationships and correlations across multiple EHG signal channels beyond the physical properties of single-channel signals. Despite their potential benefits, these methods have inherited limitations, including strict mathematical assumptions and complex computational requirements. To overcome these challenges, we introduced a method commonly used in EEG research: phase synchrony. This approach allowed us to assess the time-varying functional connectivity among multiple EHG signals simultaneously by quantifying the degree of phase synchronization of certain frequency bands. Given the diverse spectral components of EHG signals, we hypothesized that the phase synchronization information of specific unique frequency bands could serve as a crucial factor in a preterm birth prediction model.

Therefore, this study focused on the application of our proposed method for classifying term and preterm deliveries and evaluated its performance under various conditions. By comparing the results and conducting thorough post-hoc statistical analyses, we aimed to propose an optimal solution for integrating EHG phase synchrony into preterm delivery prediction.

## 2. Materials and methods

### 2.1. Term-Preterm ElectroHysteroGram Dataset with Tocogram (TPEHGT DS)

This study utilized one of the most popular EHG databases, the TPEHGT DS, which is publicly accessible on the PhysioNet website [18,33]. Our main objectives were twofold. First, we aimed to explore the possibility of utilizing phase synchrony analysis to portray EHG characteristics under different pregnancy conditions, specifically preterm and term deliveries. For this purpose, we chose an EHG database that contains EHG signals for both preterm and term conditions. Second, we sought to understand the variation in the classification accuracy of the two pregnancy scenarios when TOCO signals were present or absent in conjunction with the EHG signals. Unlike other EHG databases, TPEHGT DS includes TOCO signals that record the external pressure on the abdomen of the participants.

Participants were categorized into three groups based on their pregnancy conditions: non-pregnant, preterm, and term delivery. In this database, monopolar signals were initially captured using four $AgCl_2$ electrodes placed on the abdominal surfaces of both non-pregnant and pregnant women. These electrodes were positioned in two horizontal rows, above and below the navel, with a gap of 7 cm between them. The captured monopolar signals were transformed into three bipolar electromyography signals: S1 (E2-E1) from the highest electrode, S2 (E2-E3) from the leftmost electrodes, and S3 (E4-E3) from the lower electrodes, where E1, E2, E3, and E4 represent the monopolar surface potentials of electrodes 1, 2, 3, and 4, respectively. Along with the three bipolar EHG signals, the fourth signal, the TOCO signal, represents the external mechanical pressure on the uterus recorded from the top of the fundus. Both the EHG and TOCO signals were prefiltered with a bandwidth of 0–5.0 Hz and digitized at a sampling rate of 20 Hz. All signals were recorded concurrently for a 30-min period. Fig. 1A illustrates the common position of electrodes and the description of the three EHG signals in the TPEHGT DS database. Fig. 1B shows an example of all raw time-varying signals consisted of the 3 EHG and TOCO signals with color-marked annotations (contraction and dummy intervals) in a record.

The TPEHGT DS consists of 31 individual records: five from the non-pregnant, 13 from the preterm, and 13 from the term delivery conditions. For the preterm and term conditions, each record had a corresponding annotation file indicating the periods of contraction and non-contraction (dummy). In terms of variety, there were an equal total of 53 contraction and dummy intervals in the term condition and 47 in the preterm condition.

### 2.2. EHG and TOCO signal analyses

#### 2.2.1. Band pass filtering

Before conducting the phase synchrony analysis, we applied a band-pass filter to all the EHG and TOCO signals. This was performed using a filter bank set to several predetermined frequency ranges, enabling us to identify the unique characteristics of the EHG phase synchronization of each band. As mentioned in the Introduction, it is well-documented that the primary frequency component of EHG signals ranges from 0 to 5 Hz [14,18,32,34,35]. Within this frequency spectrum, several distinctive bands can be recognized, each of which has different clinical implications. For instance, a study conducted by Jager et al. [32] on predicting preterm delivery segmented the valid EHG frequency range into five distinct bands: (1) 0.3–1.0 Hz, (2) 0.3–4.0 Hz, (3) 1.0–2.2 Hz, (4) 2.2–3.5 Hz, and (5) 3.5–5.0 Hz. Guided by past EHG research and our preliminary results, we selected the following five distinct frequency bands for this study as (1) B0: 0.08–0.34 Hz, (2) B1: 0.34–1.00 Hz, (3) B2: 1.00–2.20 Hz, (4) B3: 2.20–3.50 Hz, and (5) B4: 3.50–5.00 Hz. Bandpass filtering was performed using finite impulse response filters with 256 filter orders. To minimize the potential impact of phase distortion resulting from digital filtering, this process is usually implemented in a forward and backward manner.

#### 2.2.2. Phase synchrony

In the main phase synchrony analysis of the EHG and TOCO signals, it is essential to extract instantaneous phase information for all the signals within each individual band. To achieve this, we first applied the Hilbert transform to the bandpass-filtered signals, yielding the associated analytic signals. Because these filtered and analytical signals correspond to the real and imaginary parts of a complex

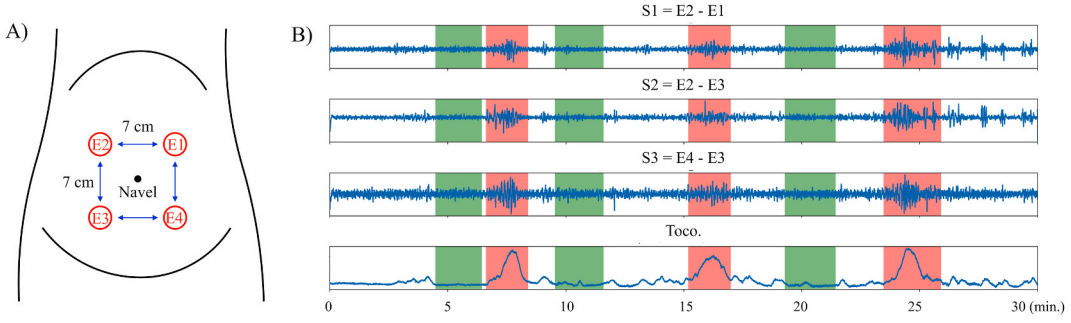



**Fig. 1.** Fig. 1A depicts the layout of the four $AgCl_2$ electrodes and the description of the three EHG signals in the TPEHGT DS. Fig. 1B shows all raw time-varying traces as the three EHG (S1, S2, S3) and TOCO signals with color-marked annotations (contraction intervals in red and dummy intervals in green) of one example record (tpeght_p008).

plane, they can be represented in a complex domain.

We calculated the instantaneous phase information as the angles of all vectors on this complex plane, which was simply determined as the radian of the inverse tangent between the real and imaginary parts. Within each frequency band, we also computed all instantaneous phase differences among all six pairs comprising the three EHG and one TOCO signals. Next, to gauge the degree of phase synchronization between the two signals, we utilized a well-known measure of phase synchrony, known as the Mean Phase Coherence (MPC), as proposed by Mormann et al. [36]. As an efficient measure of the consistency of phase differences between two signals over a specific time period, the MPC has the statistical advantage of normalization, in which all MPC values numerically range from 0 to 1.

As the values of the instantaneous phase differences are in radians, they were transformed into vectors on a unit circle in the complex plane using sine and cosine functions. We then determined the mean vector lengths by summing all complex vectors over a certain time duration. MPC is the mean of the vector lengths. Normalization of the MPC values from 0 to 1 is necessary because they are derived from the vectors on the unit circle in the complex plane. A constant instantaneous phase difference between the two signals indicates a strong phase coupling between them, and as a result, the MPC approximates 1. Conversely, if the phase differences exhibit substantial fluctuations, the MPC tends towards zero.

#### 2.2.3. Constructing MPC features

Determining a fixed time duration is crucial for calculating the MPC because it represents the mean of the vector sum within a given timeframe. Apart from the clinical meaning of uterine contraction during pregnancy, it is still controversial whether the EHG signals from the contraction intervals are more useful to classify preterm and term delivery compared with those from the dummy intervals [18,37]. In this study, we ignored these contraction and dummy intervals. Instead, we used the complete EHG and TOCO signals across the entire duration by sliding a predefined time window with a set overlap rate.

Specifically, the MPC features that represented phase synchrony within a specific timeframe among all pairs were formulated based on two factors: the duration of the time window and the overlap rates. For the non-overlapping condition, we established nine distinct time window conditions of 2, 5, 10, 20, 40, 60, 80, 100, and 120 s without any overlap in the time window. For the overlapping condition, the fixed time windows were divided by varying the overlap rates, resulting in 18 individual conditions. This was performed by combining six different time window durations (20, 40, 60, 80, 100, and 120 s) with three overlap rates between the time windows (25, 50, and 75 %). In summary, each time window for the construction of the MPC features can be individually arranged through nine non-overlapping conditions and 18 overlapping conditions. The MPC measure within a fixed time window served as the set of MPC features used in this study for both preterm and term classification tests, as well as for subsequent statistical analysis.

### 2.3. Classification analyses

#### 2.3.1. Preterm vs. term classification

Using this set of MPC features, we conducted classification tests for preterm and term deliveries to assess the suitability of these features, which are linked to the unique characteristics of EHG and TOCO signals. These classification tests had two variations in the MPC features, depending on whether they included TOCO signals or not. Because the MPC features are derived from a pair of signals, the number of feature dimensions is contingent on whether TOCO signals are included. If the set of MPC features are compiled from the EHG and TOCO signals, it would yield a 30-dimensional feature set. However, without the TOCO signal, this is a 15-dimensional feature set. In addition, we tested all MPC features by employing two different classifiers based on the TabNet and eXtreme Gradient Boosting (XGBoost) algorithms (refer to the following section for further details). Consequently, each individual test generated four different performance results (two different algorithms $\times$ two variations of the MPC features) for binary classification. We evaluated the performance results of the binary classification originating from all 27 conditions based on the time window and overlap rate (nine non-overlapping and 18 overlapping conditions).

#### 2.3.2. TabNet and XGBoost

TabNet [38], a deep-learning algorithm specializing in tabular data, combines the advantages of a tree-based ensemble model. It comprises two main components: an encoder network and a decision network. The encoder network processes the input features and performs a series of nonlinear transformations to produce a set of feature embeddings. These embeddings are then utilized by the decision network to generate predictions. The TabNet algorithm also features a unique feature-selection mechanism that learns to selectively focus on certain input features, thereby improving prediction accuracy. This mechanism operates using a sparse attention system that learns to identify the most pertinent features of each decision step.

XGBoost is a machine-learning algorithm widely used for regression and classification tasks on large, high-dimensional datasets [39]. XGBoost implements gradient-boosting trees, which is an ensemble learning method that amalgamates multiple decision trees to generate predictions. In XGBoost, each decision tree is constructed sequentially, with each subsequent tree striving to rectify errors in the preceding tree. The model learns by minimizing the loss function, which quantifies the disparity between predicted and actual values. Loss function optimization was performed using gradient descent, calculating the gradient of the loss function and updating the model parameters to reduce the loss.

For the assignment of hyperparameters in both the TabNet and XGBoost models, we used an efficient framework of hyperparameter tunings called *optuna*, which was proposed by Akiba et al. [40], with several advantages, such as ease of setup, flexibility, and visualization. We employed the TabNet algorithm with default hyperparameters provided by the PyTorch platform [41]. Some hyperparameters were selected by means of the *optuna* toolbox such as the: (1) number of independent (“n_independent”) and shared

("n_shared") gated linear units, respectively; (2) coefficient for feature reusage in the masks ("gamma"); and (3) the extra sparsity loss of coefficient ("lambda_sparse"). Similar to TabNet, most of the XGBoost hyperparameters were set as the default values and the four following hyperparameters were tuned by the *optuna* toolbox the: (1) maximum depth of a tree ("max_depth"), (2) minimum sum of instance weight ("min_child_weight"), (3) minimum loss reduction ("gamma"), and (4) learning rate ("eta").

In each preterm and term delivery classification test, a set of MPC features was initially constructed under specific conditions determined by three factors: the time window, the overlap rate, and whether the TOCO signal was used. All data comprising the MPC features in each segmented time window were collected and pooled into two types of delivery conditions across all the subjects. Subsequently, the set of MPC features was randomly split into a dataset for the training session (90 %) and another for the test session (10 %) using a 10-fold cross-validation method. In the training session, the data for model training were further divided into 90 % training data and 10 % validation data. The two classifier models, TabNet and XGBoost, were individually trained using the training data in accordance with the corresponding hyperparameters. During the testing session, the learned classifiers predicted the remaining test datasets. All classification test results were evaluated based on the accuracy rate, sensitivity, specificity, and F1-score.

### 2.4. MPC feature evaluation

Based on the results of the preterm and term classification tests, we conducted three further analyses to evaluate the use of individual MPC features to represent the unique EHG properties related to a woman's pregnancy. To conduct these analyses, it is first necessary to select a specified set of MPC features. We chose a set of MPC features obtained from the 100-s interval with a 75 % overlap rate, considering the balance with the highest classification accuracies in both the TabNet (80-s time interval) and XGBoost (120-s interval) models. First, we examined the significance of each feature to the output of the binary classifier model by utilizing a specific function, which is termed as *feature importance*, available in most decision-tree models, including the XGBoost algorithm. This function typically provides three measures of the feature importance: weight, cover, and gain. In decision-tree models, weight indicates how often a feature is chosen across all trees during data splitting. Cover is similar to weight, but is additionally weighted by the number of training data points. Gain represents the average reduction of training loss achieved at the data split. In this study, we adopted feature importance as the weight in both the TabNet and XGBoost models. Second, we applied the SHapley Additive exPlanations (SHAP) method to inspect individual feature attributes in the binary classifier XGBoost model. SHAP was developed to address some limitations of the feature importance function embedded in decision-tree-based models. For instance, measurements from feature-importance methods can lead to inconsistent and inaccurate interpretations because they appear to be sensitive to different classifier models [42]. The SHAP analysis offers the advantage of providing both local and global interpretability of input features, independent of the model used, by comparing the contribution of individual features to the final output relative to the overall average predicted by the model. The overall flowchart of analytic procedures in this study were illustrated in Fig. 2.

### 2.5. Statistical analysis on MPC features

To assess the statistical characteristics of MPC features, we performed further statistical analyses to verify significant differences across preterm, term, and non-pregnancy conditions in the TPEHGT DS. Prior to these analyses, we selected a representative set of MPC features calculated from a 100-s time window with a 75 % overlap rate. This set was identical to the one used in the SHAP analysis, as it showed one of the best performances in the main preterm and term delivery classification tasks. The dataset consisted of 115 samples from 5 non-pregnant individuals, 299 samples from 13 preterm deliveries, and 299 samples from 13 term deliveries, regardless of contraction, dummy, and other remaining intervals. For each individual feature, we first performed the Kruskal-Wallis H-test, a nonparametric statistical test for one-way ANOVA on ranks, to assess whether significant differences exist across the three pregnancy

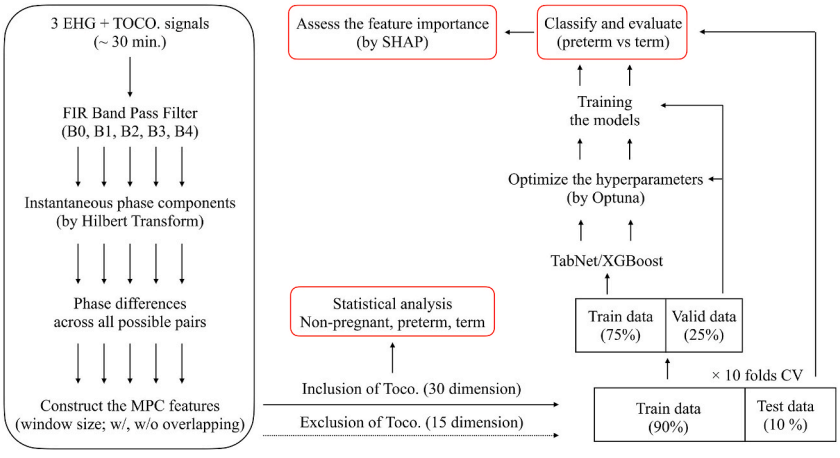



**Fig. 2.** Overall flowchart of the analytic procedures in this study starting from the preprocessing of raw EHG and TOCO signals to binary classification tests using XGBoost and TabNet algorithms.

conditions. In cases where the Bonferroni-adjusted p-values were below 0.05, we performed additional post-hoc analyses using the Mann-Whitney *U* test on all possible pairs among the three conditions to determine which conditions were different.

## 3. Results

### 3.1. Preterm vs. term classification

To assess the overall performance of term and preterm delivery classification tasks in terms of the two binary classifiers and the use of TOCO signals, we summarized all accuracy rates as a representative metric of performance in both the non-overlap and overlap conditions. Table 1 and Fig. 3 display all accuracy rates with respect to various fixed time windows in the non-overlapping condition, where the MPC features were obtained without any overlap between the segmented time windows. As depicted in Fig. 3, the two classifiers, TabNet and XGBoost, show different trends according to the length of the time window. The performance of XGBoost increased linearly until a window time of nearly 40 s and then remained constant despite longer window times. Conversely, the performance of TabNet increased at the 20-s window time, but dramatically decreased from the 40-s window time. This observation seems to originate from the differing procedures of hyperparameter tuning between TabNet and XGBoost and will be discussed in more detail below.

In comparison with two different sets of MPC features, where the TOCO signals were either included (30 feature dimensions) or excluded (15 feature dimensions), the accuracy was slightly higher when the TOCO signals were used. However, it is unlikely that the use of TOCO signals significantly improved the overall performance because these differences under all conditions were minimal. The highest accuracy rate for XGBoost was 0.900 in the 100-s time window with the TOCO signal, while that for TabNet was 0.820 in the 20-s time window with the TOCO signal. The corresponding performance metrics in the non-overlapping condition are listed in Table 2.

In the overlapping condition, where the MPC features were extracted from the overlapping time windows, Table 3 and Fig. 4 present the accuracy rates with respect to both the time window and overlap rates. Note that, in addition to the three different overlap rates of 25 %, 50 %, and 75 %, we also included a no-overlap rate of 0 % for a more direct comparison with Table 1 and Fig. 3. The results showed that larger overlap rates corresponded to higher accuracy rates, with all rates derived from a 75 % overlap surpassing those of the other overlap percentages. In general, XGBoost outperformed TabNet under most conditions, except in the shortest time interval (20 s) or at smaller time windows (20, 40, 60, and 80 s), with higher overlap rates (75 %). Similar to the non-overlapping condition, the inclusion of TOCO signals did not significantly boost accuracy rates. The peak accuracy rate for TabNet reached 0.983 at an 80-s window time with a 75 % overlap rate and the inclusion of the TOCO signal, whereas XGBoost peaked at 0.978 at a 100-s window time with the same overlap rate and the inclusion of the TOCO signal. The highest accuracy rates with the TOCO signals were only 0.035 and 0.009 higher than those without TOCO signals by the TabNet and XGBoost classifiers, respectively. For the comparison of these results with the classification performance reported in previous studies, we summarized the main metrics of classification performance on preterm and term deliveries corresponding to the TPEHGT DS database in Table 4.

### 3.2. MPC feature evaluation and statistical analysis related to different delivery conditions

For a more comprehensive analysis, considering not only the statistical difference of individual features but also the feature importance in differentiating preterm and term deliveries, we relied on MPC features with TOCO signals extracted from a 100-s time window with a 75 % overlap rate, as summarized in Table 5. The following analyses utilized the same MPC features extracted under these conditions.

Fig. 5A illustrates all the SHAP values that indicated the effect of each MPC on the preterm and term classification tasks, and the mean of the absolute SHAP values is depicted in Fig. 5B. In Fig. 5A, each data point, represented by either a red or blue dot, signifies a SHAP value associated with the prediction of preterm or term delivery. For example, if an individual feature has many red dots located on the positive x-axis, this indicates that the corresponding feature led to a term delivery prediction. The red dots on the negative x-axis represent preterm delivery predictions. Fig. 5C illustrates the importance of each MPC feature in the preterm and term delivery classification tasks. The importance of a feature was defined by the frequency of its appearance in both models. Note that the values of

**Table 1**
Accuracy rates of preterm and term delivery classification tasks without overlap rate.

| Window Time (sec) | TabNet | | XGBoost | |
|---|---|---|---|---|
| | without Toco | with Toco | without Toco | with Toco |
| 2 | 0.624 | 0.653 | 0.641 | 0.670 |
| 5 | 0.702 | 0.728 | 0.718 | 0.755 |
| 10 | 0.749 | 0.789 | 0.767 | 0.807 |
| 20 | 0.791 | 0.820 | 0.816 | 0.845 |
| 40 | 0.565 | 0.636 | 0.852 | 0.885 |
| 60 | 0.578 | 0.620 | 0.859 | 0.865 |
| 80 | 0.568 | 0.547 | 0.879 | 0.888 |
| 100 | 0.554 | 0.577 | 0.878 | 0.900 |
| 120 | 0.467 | 0.495 | 0.866 | 0.882 |

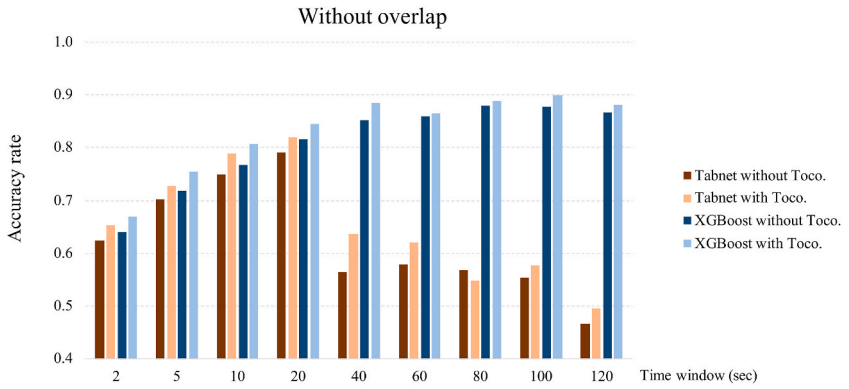



**Fig. 3.** The graph represents the accuracy rates of the binary classifiers, TabNet and XGBoost, comparing accuracy across various fixed time windows under non-overlapping conditions where the MPC features were obtained. The x-axis represents the length of the time window, and the y-axis represents the accuracy rate of each classifier.

**Table 2**
The best performance metrics of non-overlapping conditions at 20 s in TabNet and 100 s in XGboost.

| 0 % overlap | Toco | accuracy | sensitivity | specificity | F1 score |
|---|---|---|---|---|---|
| 20 s (TabNet) | without Toco | 0.791 | 0.784 | 0.798 | 0.792 |
| | with Toco | 0.820 | 0.849 | 0.791 | 0.806 |
| 100 s (XGBoost) | without Toco | 0.878 | 0.883 | 0.873 | 0.876 |
| | with Toco | 0.900 | 0.918 | 0.882 | 0.897 |

**Table 3**
Accuracies of preterm and term delivery classification tasks with overlap rates.

| Window Time (sec) | Overlap percentage (%) | TabNet | | XGBoost | |
|---|---|---|---|---|---|
| | | without Toco | with Toco | without Toco | with Toco |
| 20 | 0 | 0.791 | 0.820 | 0.816 | 0.845 |
| | 25 | 0.812 | 0.865 | 0.830 | 0.872 |
| | 50 | 0.844 | 0.888 | 0.841 | 0.884 |
| | 75 | 0.876 | 0.937 | 0.873 | 0.918 |
| 40 | 0 | 0.565 | 0.636 | 0.852 | 0.885 |
| | 25 | 0.743 | 0.770 | 0.870 | 0.895 |
| | 50 | 0.832 | 0.904 | 0.889 | 0.910 |
| | 75 | 0.926 | 0.969 | 0.920 | 0.950 |
| 60 | 0 | 0.578 | 0.620 | 0.859 | 0.865 |
| | 25 | 0.616 | 0.675 | 0.867 | 0.904 |
| | 50 | 0.742 | 0.758 | 0.914 | 0.923 |
| | 75 | 0.948 | 0.978 | 0.937 | 0.965 |
| 80 | 0 | 0.568 | 0.547 | 0.879 | 0.888 |
| | 25 | 0.561 | 0.590 | 0.900 | 0.906 |
| | 50 | 0.635 | 0.700 | 0.921 | 0.930 |
| | 75 | 0.886 | 0.983 | 0.955 | 0.964 |
| 100 | 0 | 0.554 | 0.577 | 0.878 | 0.900 |
| | 25 | 0.538 | 0.587 | 0.918 | 0.906 |
| | 50 | 0.559 | 0.631 | 0.922 | 0.928 |
| | 75 | 0.903 | 0.844 | 0.966 | 0.970 |
| 120 | 0 | 0.467 | 0.495 | 0.866 | 0.882 |
| | 25 | 0.572 | 0.544 | 0.895 | 0.899 |
| | 50 | 0.596 | 0.567 | 0.927 | 0.930 |
| | 75 | 0.671 | 0.797 | 0.969 | 0.978 |

feature importance in Fig. 5C were obtained from the MPC with TOCO signals at a 100-s time window with a 75 % overlap rate, in which the performance of the XGBoost classifier was almost at the highest level, but those of the TabNet classifier (0.844) were not better than those of the same case in the 80-s time window (0.983), as shown in Table 3.

Considering the evidence of feature importance shown in Fig. 5, which is more focused on the SHAP results (Fig. 5A and B), we can summarize three main findings. First, the use of MPC features for preterm prediction could be interpreted in terms of two distinct frequency ranges, B0, B1, B2, B3, and B4, either above or below 2.2 Hz. The MPC features obtained from the lower frequency ranges

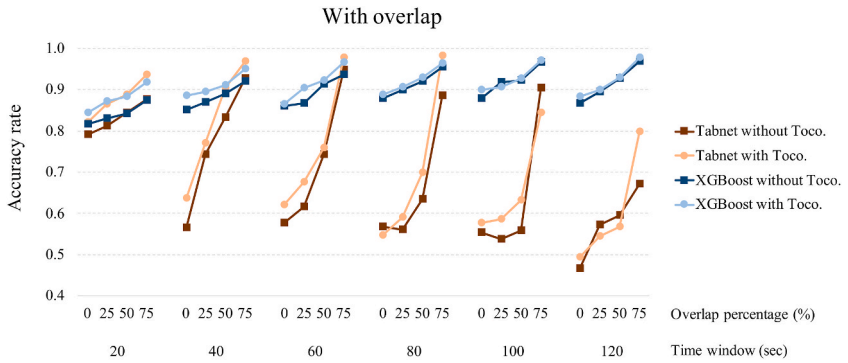



**Fig. 4.** This figure illustrates the accuracy rates of the binary classifiers, TabNet and XGBoost, under overlapping conditions. The accuracy rates are presented in relation to both the length of the time window and the degree of overlap (0 %, 25 %, 50 %, and 75 %). Each line corresponds to a different overlap rate, showing that larger overlap rates generally lead to higher accuracy rates.

**Table 4**
Summary of the resultant metrics of preterm and term delivery classification in TPEHGT DB.

| Works | Signals | Segments/Intervals | Features | Methods | SEN (%) | SPE (%) | ACC (%) |
|---|---|---|---|---|---|---|---|
| Jager et al., 2018 [18] | S2, S3 | Contraction interval<br>Dummy interval | Sample entropy, Median frequency, | SFS + | 89.0<br>87.0 | 89.0<br>91.0 | 88.7<br>88.8 |
| | S2, TOCO | Contraction interval<br>Dummy interval | Peak amplitude | QDA | 89.0<br>91.0 | 92.0<br>92.0 | 90.6<br>91.5 |
| Chen et al., 2020 [52] | S1, S2, S3, TOCO | 51.2 s segments in the contraction and dummy intervals manually | Sample entropy, Wavelet entropy | DNN<br>DBN<br>H-ELM | 98.0<br>91.6<br>93.2 | 97.7<br>94.2<br>87.5 | 97.9<br>93.0<br>90.4 |
| This study | S1, S2, S3, TOCO | 100 s segments without overlap in the entire signal<br>100 s segments with 75 % overlap in the entire signal | Mean phase coherence | XGBoost | 91.8<br>96.9 | 88.2<br>97.1 | 90.0<br>97.0 |

**Table 5**
The performance metrics of the TabNet and XGBoost classifiers using MPC features extracted from 100 s with overlap of 75 %.

| 100 s overlap 75 % | TabNet | | | | XGBoost | | | |
|---|---|---|---|---|---|---|---|---|
| | accuracy | sensitivity | specificity | F1 score | accuracy | sensitivity | specificity | F1 score |
| without Toco | 0.903 | 0.955 | 0.850 | 0.915 | 0.966 | 0.973 | 0.959 | 0.966 |
| with Toco | 0.844 | 0.949 | 0.739 | 0.770 | 0.970 | 0.969 | 0.971 | 0.970 |

(B0, B1, and B2) below 2.2 Hz were more informative for preterm prediction than those of higher frequencies (B3 and B4). Second, in the B2 band, the two highest SHAP values were for the S3-S2 and TO-S3 pairs, with high feature importance in the XGBoost model at S3-S2. Third, the S3-S1 pair played an important role in both the B0 and B1 bands, with the highest feature importance in the XGBoost model for the B1 band. These findings indicate that the EHG signals in the lower horizontal direction (S3) play a more important role in the phase network in the low-frequency ranges below 2.2 Hz across other signals, including TOCO signals.

These findings of feature importance were further verified by additional statistical analyses across three different pregnancy conditions (preterm, term, and non-pregnancy), which are summarized in both Table 6 and Fig. 6. Using two main nonparametric statistical analyses, the Kruskal-Wallis and Mann-Whitney rank tests, we identified significant differences in MPC features across various pregnancy conditions. As shown in Table 6, most MPC features across B0 to B4 bands exhibited statistically significant differences between the non-pregnancy and pregnancy conditions. Notably, the lower frequency ranges (B0, B1, B2) have more powerful features with significant differences compared to the higher frequency ranges (B3, B4). In particular, all six features in each of the B0 and B2 bands, along with four features in the B1 band, demonstrated strong statistical differences (p-value $<0.01$). In line with the objective of this study, we observed strong statistical differences between preterm and term deliveries in three of four MPC features (S3-S1 in B0 (Fig. 6A), S3-S1 in B1 (Fig. 6B), TO-S3 in B2 (Fig. 6D)) that showed the highest feature importance in the SHAP analysis. Specifically, horizontal synchronization (S3-S1 pair) in the B0 band was higher in the preterm condition, while vertical synchronization (S3-S2 pair) in the B2 band was higher in the term condition (Fig. 6C). In addition, the synchronization between the lower horizontal direction (S3) and TOCO signals in the B2 band was relatively higher in the preterm condition. Based on these statistical

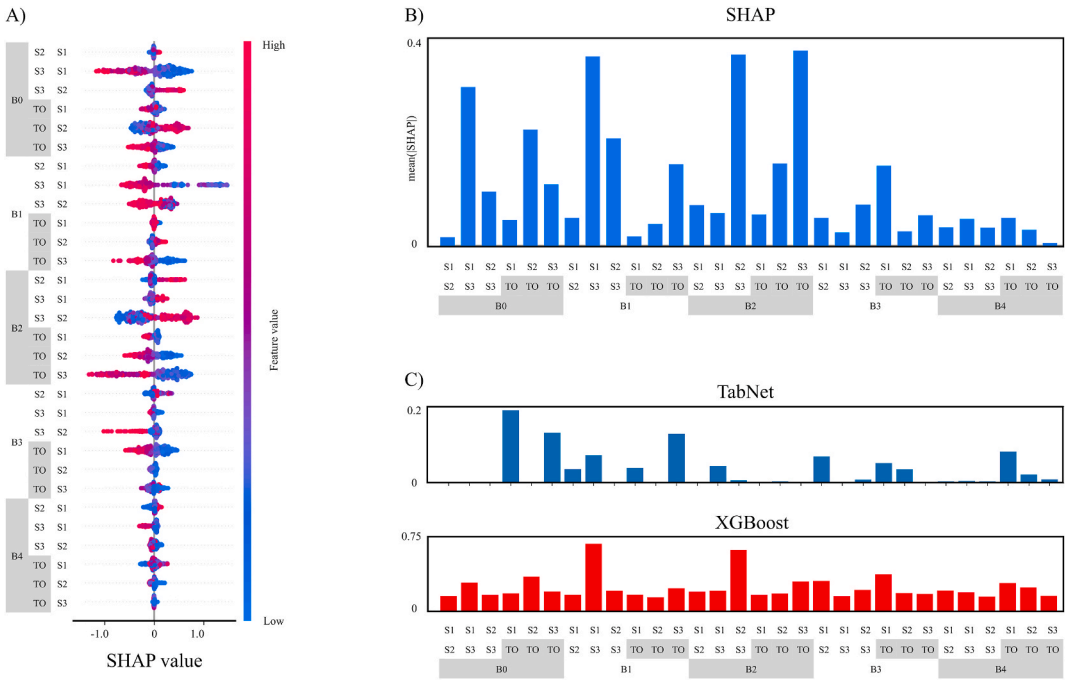



**Fig. 5.** Fig. 5A shows the individual SHAP values. Each data point, represented by either a red or blue dot, signifies a SHAP value associated with the prediction of preterm or term delivery. Fig. 5B shows the mean absolute SHAP values in Fig. 5A. Fig. 5C shows the feature importance (weight) of individual MPC features in the TabNet and XGBoost models, respectively.

results, we confirmed that the MPC features related to different pregnancy conditions, which were primarily used as input variables for the XGBoost classifier, significantly contributed to the improved classification performance between preterm and term deliveries.

## 4. Discussion

### *4.1. Spectral characteristics of phase synchronization in the EHG and TOCO signals related to preterm and term deliveries*

Our study aimed to determine the predictive potential of uterine EHG signals in distinguishing between preterm and term births using advanced machine-learning algorithms. We identified two EHG phase synchronization features as significant predictors. The first key feature included EHG signals from the S1-S3 pairs within the B0 and B1 frequency bands (0.08–1.0 Hz), and the second included EHG signals from the S2-S3 and S3-TO pairs in the B2 frequency band (1.0–2.2 Hz). The crucial influence of EHG signals in the low-frequency range (<2.2 Hz) on preterm birth prediction was apparent in our study.

The utility of lower-frequency EHG signals for preterm birth prediction is consistent with previous findings [26,30,43]. For instance Ref. [30] underlined that the spectral power within the 0.3–3.0 Hz range could successfully differentiate between term and preterm deliveries. Fele-Zorz et al. [26] further supported the importance of low-frequency signals in predicting labor, emphasizing the significance of spectral power in low-frequency bands (<3 Hz) in differentiating between true and false labor. Moreover, a previous study found significant changes in power spectral density and propagation velocity in the low-frequency range (0.3–0.8 Hz) of EHG signals during labor [34]. Our study further validated these earlier findings and refined the understanding of this frequency range (<2.2 Hz), which appeared to be the most critical in preterm labor prediction.

In the higher frequency band (0.34–2.2 Hz), our research underscored the importance of phase synchronization features from the S2-S3 and S3-TO pairs. Notably, S3 (lower horizontal) appeared to serve as a central point for both vertical (S2) and TOCO signal synchronization within this band. This finding is particularly relevant, as it points to the spatiotemporal propagation of uterine contractions, a critical aspect of labor that has been explored in previous studies [44,45]. Our study suggests that while the lower-frequency signals (<1.0 Hz) might be tied to the overall synchronization of contractions across the horizontal plane, the higher-frequency signals (1.0–2.2 Hz) could be representative of localized myometrial activity and the propagation of contractions involving both vertical and horizontal directions. This interpretation aligns with the findings of Leman et al. [46] who suggested that the different frequency components of EHG signals may reflect distinct aspects of the contractile process.

Our results revealed a distinctive pattern of phase synchronization across the vertical plane of the abdomen, as represented by the S2-S3 pairs in the 0.08–1.0 Hz band. Synchronization across the upper (S1) and lower (S3) uterine regions is indicative of an intricate network of electromechanical activities occurring during labor. Several studies have underscored the critical role of this synchronization in labor progression and effective expulsion of the fetus [28,47]. This synchronization pattern may reflect the coordinated

**Table 6**
Statistical analysis on non-pregnancy, preterm, and term delivery conditions by means of the set of MPC features extracted from 100 s with overlap of 75 %. For the statistical results, single asterisk markers indicate the marginally significant differences across the corresponding groups ($p < 0.05$) and double asterisk markers indicate the strongly significant differences ($p < 0.01$).

| Band | Pair | MPC (avg ± std) | | | Krustal-Wallis test | Post-hoc (Mann-Whitney rank test) | | |
|---|---|---|---|---|---|---|---|---|
| | | NP | P | T | | NP vs P | NP vs T | P vs T |
| B0 | S2_S1 | 0.51 ± 0.22 | 0.24 ± 0.14 | 0.26 ± 0.17 | ** | ** | ** | |
| | S3_S1* | 0.47 ± 0.24 | 0.30 ± 0.19 | 0.23 ± 0.16 | ** | ** | ** | ** |
| | S3_S2 | 0.77 ± 0.15 | 0.35 ± 0.16 | 0.41 ± 0.23 | ** | ** | ** | |
| | TO_S1 | 0.72 ± 0.19 | 0.42 ± 0.20 | 0.45 ± 0.24 | ** | ** | ** | |
| | TO_S2 | 0.90 ± 0.05 | 0.45 ± 0.26 | 0.54 ± 0.26 | ** | ** | ** | ** |
| | TO_S3 | 0.48 ± 0.19 | 0.49 ± 0.16 | 0.44 ± 0.16 | ** | ** | | |
| B1 | S2_S1 | 0.44 ± 0.19 | 0.52 ± 0.17 | 0.46 ± 0.14 | ** | ** | ** | |
| | S3_S1* | 0.73 ± 0.20 | 0.78 ± 0.14 | 0.72 ± 0.17 | * | | | * |
| | S3_S2 | 0.88 ± 0.10 | 0.80 ± 0.15 | 0.75 ± 0.17 | ** | ** | ** | ** |
| | TO_S1 | 0.94 ± 0.05 | 0.90 ± 0.07 | 0.87 ± 0.10 | ** | ** | ** | |
| | TO_S2 | 0.48 ± 0.29 | 0.34 ± 0.18 | 0.35 ± 0.19 | * | ** | | |
| | TO_S3 | 0.36 ± 0.24 | 0.34 ± 0.18 | 0.25 ± 0.17 | ** | ** | ** | |
| B2 | S2_S1 | 0.59 ± 0.22 | 0.32 ± 0.15 | 0.37 ± 0.23 | ** | ** | ** | |
| | S3_S1 | 0.64 ± 0.24 | 0.41 ± 0.18 | 0.46 ± 0.23 | ** | ** | ** | |
| | S3_S2* | 0.86 ± 0.08 | 0.43 ± 0.26 | 0.54 ± 0.26 | ** | ** | ** | ** |
| | TO_S1 | 0.55 ± 0.20 | 0.24 ± 0.15 | 0.17 ± 0.10 | ** | ** | ** | ** |
| | TO_S2 | 0.36 ± 0.17 | 0.33 ± 0.19 | 0.24 ± 0.15 | ** | ** | ** | |
| | TO_S3* | 0.23 ± 0.18 | 0.39 ± 0.20 | 0.32 ± 0.23 | ** | ** | ** | ** |
| B3 | S2_S1 | 0.17 ± 0.14 | 0.20 ± 0.14 | 0.19 ± 0.18 | | | | |
| | S3_S1 | 0.21 ± 0.22 | 0.09 ± 0.05 | 0.09 ± 0.06 | ** | ** | ** | |
| | S3_S2 | 0.64 ± 0.22 | 0.28 ± 0.18 | 0.20 ± 0.12 | ** | ** | ** | ** |
| | TO_S1 | 0.34 ± 0.20 | 0.39 ± 0.21 | 0.28 ± 0.17 | ** | ** | | |
| | TO_S2 | 0.22 ± 0.15 | 0.18 ± 0.10 | 0.24 ± 0.23 | | | | |
| | TO_S3 | 0.15 ± 0.12 | 0.13 ± 0.07 | 0.17 ± 0.17 | | | | |
| B4 | S2_S1 | 0.20 ± 0.19 | 0.08 ± 0.04 | 0.09 ± 0.05 | ** | ** | ** | |
| | S3_S1 | 0.40 ± 0.24 | 0.26 ± 0.17 | 0.19 ± 0.11 | ** | ** | ** | ** |
| | S3_S2 | 0.28 ± 0.19 | 0.33 ± 0.19 | 0.30 ± 0.21 | | | | |
| | TO_S1 | 0.21 ± 0.15 | 0.37 ± 0.21 | 0.31 ± 0.22 | ** | ** | ** | ** |
| | TO_S2 | 0.16 ± 0.11 | 0.19 ± 0.14 | 0.18 ± 0.17 | | | | |
| | TO_S3 | 0.20 ± 0.18 | 0.09 ± 0.05 | 0.09 ± 0.06 | ** | ** | ** | |

The four MPC features superscripted with an asterisk are the same as the four features with the highest SHAP values and shown in Fig. 6.

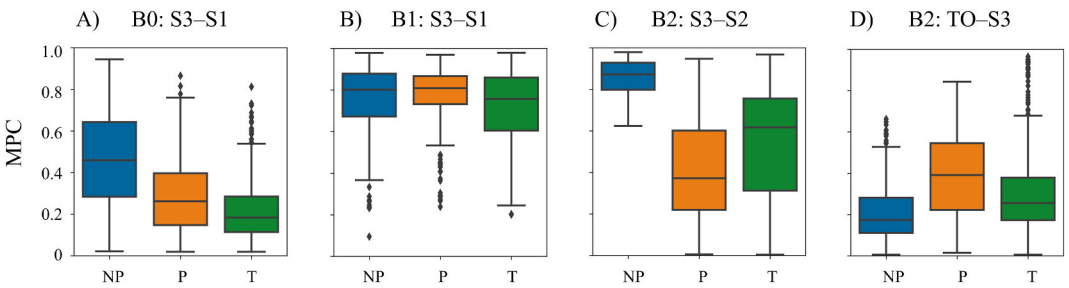



**Fig. 6.** The distribution of three different pregnancy conditions (non-pregnancy, preterm, and term) in the four most important MPC features by SHAP analysis at 100-s segment with 75 % overlap. The errorbar means the standard error in each group.

contractile behavior of the uterus during labor. Moreover, these signals within the 0.08–1.0 Hz band are recognized as uterine-specific frequencies, corresponding to the natural frequencies of the uterine muscle [10,48]. Euliano et al. [49] further supported this finding by demonstrating the sensitivity of low-frequency signals to oxytocin-induced uterine contractions. Therefore, our findings regarding the S1-S3 pairs in this band provided a valuable link between this uterine-specific frequency and the synchronization of uterine contractions during labor. Additionally, we examined the role of TOCO signals in predicting preterm delivery. Consistent with previous studies [8,50], our findings suggest that TOCO signals may not substantially boost the overall performance of preterm delivery predictions. However, the inclusion of the TOCO signals contributed to slightly higher accuracy rates in our classification tasks. This is attributed to the distinct types of information captured by the EHG and TOCO signals. EHG signals directly reflect the electrical activity of the myometrium, which induces physical contractions, whereas TOCO signals provide a mechanical measure of these uterine contractions. This distinction implies that EHG signals may provide a more precise depiction of uterine activity, particularly regarding the initiation and propagation of contractions. However, the potential synergistic effect of combining EHG and TOCO signals in prediction models should not be overlooked and warrants further investigation.

### 4.2. Performance of classification between preterm and term deliveries

As summarized in Table 4, two previous studies [18,52] have reported classification test results between preterm and term deliveries using the same TPEHGT DS as this study. In the study [18], that initially introduced the TPEHGT DS, the authors utilized sample entropy, median frequency, and peak amplitude derived from single-channel signals as the main EHG features. Using sequential forward selection (SFS) for feature selection and a quadratic discriminant analysis (QDA) classifier, they achieved the best accuracies of 90.6 % for contraction intervals and 91.5 % for dummy intervals. Although the occurrence of contractions is clinically important [43,51], as these results indicate, the differentiation between contraction and dummy intervals does not impact the classification of preterm and term deliveries using EHG and TOCO signals. In another study [52], the authors calculated sample entropy and wavelet entropy from manually segmented windows of 51.2 s (1024 samples) without distinguishing between contraction and dummy intervals. These constructed EHG features were then utilized as input features in several types of neural network-based classifiers, resulting in classification accuracies between preterm and term deliveries ranging from 90.4 % to 97.9 %. However, despite their outstanding classification performance, neural network-based classifiers in high-dimensional spaces require substantial computing resources for model training and have the limitation that results cannot be easily explained in terms of clinically interpretable EHG characteristics.

Unlike the aforementioned studies, our study focused on spectral phase synchronization between two signals rather than specific linear or nonlinear dynamics of a single signal. Additionally, we implemented two different types of classifiers: XGBoost and TabNet. One is a representative method based on machine learning techniques, and the other is based on deep learning techniques. Under various conditions defined by window size and overlap rate, we used spectral phase synchronization information obtained from 100-s segmented signals without considering contraction or dummy intervals. This approach resulted in accuracies of 90.0 % (without overlap) and up to 97.0 % (with a 75 % overlap rate). By exploring the effect of overlapping time windows on prediction accuracy, we observed that a higher overlap ratio led to increased accuracy. The use of overlapping windows in time-series analysis is a widely accepted approach in signal processing because it ensures the continuity of information and prevents the potential loss of crucial data points at the window boundaries [53,54]. Specifically, our results showed the highest accuracy with a 75 % overlap ratio. This trend can be attributed to the continuity of the signal and the nature of labor contractions, which are characterized by a wave-like pattern with a relatively long duration. Higher overlap ratios allowed for a more fine-grained capture of these contraction patterns, thereby providing more information to the classification models. This is corroborated by a previous study [55] that found that overlapping windows improved the performance of activity recognition systems. However, it is important to consider this trade-off, as a higher overlap could also lead to a high correlation between samples, potentially resulting in overfitting. The improvements in accuracy at a 75 % overlap ratio in our study suggest that the benefits of capturing more fine-grained temporal information and increasing the sample size may outweigh the potential overfitting risks at this level of overlap.

Interestingly, the XGBoost and TabNet classifiers exhibited apparent differences depending on the length of the time window, which is directly related to the number of samples fed into the classifiers. As the window size increases, the number of samples decreases. We believe these different performance trends are mainly due to inherent differences in the architectures of these two models in dealing with tabular data. XGBoost is a gradient-boosting model that operates by building an ensemble of weak prediction models. TabNet, a deep learning model, adopts a different approach. Its architecture relies on attention mechanisms, allowing it to select useful features and skip irrelevant ones [38].

Recently, numerous deep learning-based methodologies have been developed in the field of data science. In particular, studies dealing with image or language-type data have reported dramatically superior performance compared to traditional machine learning-based methodologies. However, when considering tabular data formats, some open questions remain. Recent studies have reported that tree-based models, such as XGBoost, outperform neural network-based models [56,57]. While further research is ongoing regarding the underlying causes, a recent study [56] reported that neural network-based models exhibited lower performance than the XGBoost algorithm in handling tabular datasets due to unwanted bias toward overly smooth solutions, as well as high sensitivity to uninformative features and the rotation of original data during training. One of our results, which showed higher performance of XGBoost compared to TabNet under conditions of longer window size and small overlap rate, also demonstrates the outperformance of tree-based classifiers in small datasets with more complex patterns in the decision hyperplane.

In conclusion, our study reaffirmed the predictive potential of EHG signals for preterm birth while providing specific insights into the frequency and spatial patterns of uterine contractions. By aligning with and expanding the existing body of research, our findings significantly contribute to the ongoing development of more accurate and effective preterm prediction models.

## CRediT authorship contribution statement

**Jae-Hwan Kang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Conceptualization. **Young-Ju Jeon:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization. **In-Seon Lee:** Writing – original draft, Investigation, Conceptualization. **Junsuk Kim:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Investigation, Funding acquisition, Formal analysis, Conceptualization.

## Data availability statement

The data used in this study, the Term-Preterm ElectroHysteroGram Dataset with Tocogram (TPEHGT DS), is publicly accessible on the PhysioNet website (https://physionet.org/).

## Funding

The present study was supported by the National Research Council of Science & Technology (NST) Aging Convergence Research Center (CRC22013-500) and the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) (2021R1F1A1055814).

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] WHO, WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths, Acta Obstet. Gynecol. Scand. 56 (1977) 247–253. Modifications recommended by FIGO as amended October 14, 1976.
[2] S.R. Walani, Global burden of preterm birth, Int. J. Gynecol. Obstet. 150 (2020) 31–33.
[3] J.P. Vogel, S. Chawanpaiboon, A.-B. Moller, K. Watananirun, M. Bonet, P. Lumbiganon, The global epidemiology of preterm birth, Best Pract. Res. Clin. Obstet. Gynaecol. 52 (2018) 3–12.
[4] F.W.G.o.G.C.P.i.M.F. Medicine, Good clinical practice advice: prediction of preterm labor and preterm premature rupture of membranes, Int. J. Gynecol. Obstet. 144 (2019) 340–346.
[5] D.L. Bentley, J.L. Bentley, D.L. Watson, R.A. Welch, R.W. Martin, K.S. Gookin, R.A. Knuppel, M.F. Lake, W.C. Hill, A.D. Fleming, et al., Relationship of uterine contractility to preterm labor, Obstet. Gynecol. 76 (1990) 36s–38s.
[6] I.L. Buxton, W. Crow, S.O. Mathew, Regulation of uterine contraction: mechanisms in preterm labor, AACN Clin Issues 11 (2000) 271–282.
[7] J.R. Challis, S.J. Lye, W. Gibb, W. Whittle, F. Patel, N. Alfaidy, Understanding preterm labor, Ann. N. Y. Acad. Sci. 943 (2001) 225–234.
[8] J. Garcia-Casado, Y. Ye-Lin, G. Prats-Boluda, J. Mas-Cabo, J. Alberola-Rubio, A. Perales, Electrohysterography in the diagnosis of preterm birth: a review, Physiol. Meas. 39 (2018) 02tr01.
[9] L. Mendis, M. Palaniswami, F. Brownfoot, E. Keenan, Computerised cardiotocography analysis for the automated detection of fetal compromise during labour: a review, Bioengineering 10 (2023) 1007.
[10] D. Schlembach, W.L. Maner, R.E. Garfield, H. Maul, Monitoring the progress of pregnancy and labor using electromyography, Eur. J. Obstet. Gynecol. Reprod. Biol. 144 (Suppl 1) (2009) S33–S39.
[11] J. Jezewski, K. Horoba, A. Matonia, J. Wrobel, Quantitative analysis of contraction patterns in electrical activity signal of pregnant uterus as an alternative to mechanical approach, Physiol. Meas. 26 (2005) 753.
[12] M. Almeida, H. Mouriño, A.G. Batista, S. Russo, F. Esgalhado, C.R.P. dos Reis, F. Serrano, M. Ortigueira, Electrohysterography extracted features dependency on anthropometric and pregnancy factors, Biomed. Signal Process Control 75 (2022) 103556.
[13] A. Shafik, Electrohysterogram: study of the electromechanical activity of the uterus in humans, Eur. J. Obstet. Gynecol. Reprod. Biol. 73 (1997) 85–89.
[14] D. Devedeux, C. Marque, S. Mansour, G. Germain, J. Duchene, Uterine electromyography: a critical review, Am. J. Obstet. Gynecol. 169 (1993) 1636–1653.
[15] J.S. Xu, Z.Q. Chen, H.X. Lou, G.J. Shen, A. Pumir, Review on EHG signal analysis and its application in preterm diagnosis, Biomed Signal Proces 71 (2022).
[16] P. Fergus, P. Cheung, A. Hussain, D. Al-Jumeily, C. Dobbins, S. Iram, Prediction of preterm deliveries from EHG signals using machine learning, PLoS One 8 (2013).
[17] W.L. Maner, L.B. MacKay, G.R. Saade, R.E. Garfield, Characterization of abdominally acquired uterine electrical signals in humans, using a non-linear analytic method, Med. Biol. Eng. Comput. 44 (2006) 117–123.
[18] F. Jager, S. Libensek, K. Gersak, Characterization and automatic classification of preterm and term uterine records, PLoS One 13 (2018) e0202125.
[19] 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.
[20] S. Vinothini, N. Punitha, P.A. Karthick, S. Ramakrishnan, Automated detection of preterm condition using uterine electromyography based topological features, Biocybern. Biomed. Eng. 41 (2021) 293–305.
[21] G. Fele-Zorz, G. Kavsek, Z. 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.
[22] X. Song, X. Qiao, D. Hao, L. Yang, X. Zhou, Y. Xu, D. Zheng, Automatic recognition of uterine contractions with electrohysterogram signals based on the zero-crossing rate, Sci. Rep. 11 (2021) 1956.
[23] J. Xu, Z. Chen, J. Zhang, Y. Lu, X. Yang, A. Pumir, Realistic preterm prediction based on optimized synthetic sampling of EHG signal, Comput. Biol. Med. 136 (2021) 104644.
[24] P. Fergus, P. Cheung, A. Hussain, D. Al-Jumeily, C. Dobbins, S. Iram, Prediction of preterm deliveries from EHG signals using machine learning, PLoS One 8 (2013).
[25] M.U. Ahmed, T. Chanwimalueang, S. Thayyil, D.P. Mandic, A multivariate multiscale fuzzy entropy algorithm with application to uterine EMG complexity analysis, Entropy-Switz 19 (2017).
[26] G. Fele-Zorz, G. Kavsek, Z. 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.
[27] L.L. Chen, H.Y. Xu, Deep neural network for semi-automatic classification of term and preterm uterine recordings, Artif. Intell. Med. 105 (2020).
[28] H. de Lau, C. Rabotti, R. Bijloo, M.J. Rooijakkers, M. Mischi, S.G. Oei, Automated conduction velocity analysis in the electrohysterogram for prediction of imminent delivery: a preliminary study, Comput. Math. Methods Med. 2013 (2013) 627976.
[29] C. Rabotti, M. Mischi, Propagation of electrical activity in uterine muscle during pregnancy: a review, Acta Physiol. 213 (2015) 406–416.
[30] C. Rabotti, M. Mischi, S.G. Oei, J.W. Bergmans, Noninvasive estimation of the electrohysterographic action-potential conduction velocity, IEEE Trans. Biomed. Eng. 57 (2010) 2178–2187.
[31] M. Yochum, J. Laforet, C. Marque, An electro-mechanical multiscale model of uterine pregnancy contraction, Comput. Biol. Med. 77 (2016) 182–194.
[32] F. Jager, K. Gersak, P. Vouk, Z. Pirnar, A. Trojner-Bregar, M. Lucovnik, A. Borovac, Assessing velocity and directionality of uterine electrical activity for preterm birth prediction using EHG surface records, Sensors 20 (2020).
[33] A.L. Goldberger, L.A.N. 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, Circulation 101 (2000) e215–e220.
[34] M. Lucovnik, R.J. Kuon, L.R. Chambliss, W.L. Maner, S.Q. Shi, L.L. Shi, J. Balducci, R.E. Garfield, Use of uterine electromyography to diagnose term and preterm labor, Acta Obstet. Gynecol. Scand. 90 (2011) 150–157.
[35] J. Peng, D. Hao, L. Yang, M. Du, X. Song, H. Jiang, Y. Zhang, D. Zheng, Evaluation of electrohysterogram measured from different gestational weeks for recognizing preterm delivery: a preliminary study using random Forest, Biocybern. Biomed. Eng. 40 (2020) 352–362.
[36] F. Mormann, K. Lehnertz, P. David, C.E. Elger, Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients, Physica D 144 (2000) 358–369.

[37] F. Jager, K. Geršak, P. Vouk, Ž. Pirnar, A. Trojner-Bregar, M. Lučovnik, A. Borovac, Assessing velocity and directionality of uterine electrical activity for preterm birth prediction using EHG surface records, Sensors 20 (2020) 7328.
[38] S.O. Arik, T. Pfister, TabNet: attentive interpretable tabular learning, Aaai Conf Artif Inte 35 (2021) 6679–6687.
[39] T.Q. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, Kdd'16, Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining (2016) 785–794.
[40] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: a next-generation hyperparameter optimization framework. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Association for Computing Machinery, Anchorage, AK, USA, 2019, pp. 2623–2631.
[41] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z.M. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J.J. Bai, S. Chintala, PyTorch: an imperative style, high-performance deep learning library, Adv Neur In (2019) 32.
[42] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, Adv Neur In 30 (2017).
[43] R.E. Garfield, W.L. Maner, Physiology and electrical activity of uterine contractions, Semin. Cell Dev. Biol. 18 (2007) 289–295.
[44] B. Karlsson, J. Terrien, V. Gudmundsson, T. Steingrimsdottir, C. Marque, Abdominal EHG on a 4 by 4 grid: mapping and presenting the propagation of uterine contractions, Ifmbe Proc 16 (2007) 139. +.
[45] L. Lange, A. Vaeggemose, P. Kidmose, E. Mikkelsen, N. Uldbjerg, P. Johansen, Velocity and directionality of the electrohysterographic signal propagation, PLoS One 9 (2014).
[46] 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.
[47] C.S. Buhimschi, I.A. Buhimschi, Advantages of vaginal delivery, Clinical obstetrics and gynecology 49 (2006) 167–183.
[48] 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.
[49] T. Euliano, M. Skowronski, D. Marossero, J. Shuster, R. Edwards, Prediction of intrauterine pressure waveform from transabdominal electrohysterography, the journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet 19 (2006) 811–816.
[50] V. Berghella, E. Hayes, J. Visintine, J.K. Baxter, Fetal fibronectin testing for reducing the risk of preterm birth, Cochrane Database Syst. Rev. 2008 (2008) CD006843.
[51] F. Esgalhado, A.G. Batista, H. Mouriño, S. Russo, C.R.P. dos Reis, F. Serrano, V. Vassilenko, M. Duarte Ortigueira, Automatic contraction detection using uterine electromyography, Appl. Sci. 10 (2020) 7014.
[52] L. Chen, H. Xu, Deep neural network for semi-automatic classification of term and preterm uterine recordings, Artif. Intell. Med. 105 (2020) 101861.
[53] A. Atyabi, S. Fitzgibbon, D.M.W. Powers, Biasing the overlapping and non-overlapping sub-windows of EEG recording. The 2012 International Joint Conference on Neural Networks (IJCNN), 2012, pp. 1–6.
[54] A. Dehghani, O. Sarbishei, T. Glatard, E. Shihab, A quantitative comparison of overlapping and non-overlapping sliding windows for human activity recognition using inertial sensors, Sensors 19 (2019) 5026.
[55] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, I. Rojas, Window size impact in human activity recognition, Sensors 14 (2014) 6474–6499.
[56] L. Grinsztajn, E. Oyallon, G. Varoquaux, Why do tree-based models still outperform deep learning on typical tabular data?. Proceedings of the 36th International Conference on Neural Information Processing Systems Curran Associates Inc., New Orleans, LA, USA, 2024. Article 37.
[57] R. Shwartz-Ziv, A. Armon, Tabular data: deep learning is not all you need, Inf. Fusion 81 (2022) 84–90.