Towards an affordable magnetomyography instrumentation and low model complexity approach for labour imminency prediction using a novel multiresolution analysis

Ejay Nsugbe [1*] and Ibrahim Sanusi[2]

[1] Independent Researcher
[2] Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, U.K.
* Correspondence: ennsugbe@yahoo.com

Abstract

The ability to predict the onset of labour is seen to be an important tool in a clinical setting. Magnetomyography has shown promise in the area of labour imminency prediction, but its clinical application remains limited due to high resource consumption associated with its broad number of channels. In this study, five electrode channels, which account for 3.3% of the total, are used alongside a novel signal decomposition algorithm and low complexity classifiers (logistic regression and linear-SVM) to classify between labour imminency due within 0–48hrs and >48hrs. The results suggest that the parsimonious representation comprising of five electrode channels and novel signal decomposition method alongside the candidate classifiers could allow for greater affordability and hence clinical viability of the magnetomyography-based prediction model, which carries a good degree of model interpretability.

Keywords: Labour Signal Processing, Obstetrics, Preterm, Signal Decomposition, Kernel Methods, Logistic Regression, Optimisation

Introduction

The process of expulsion of a foetus from the womb is termed as labour. Due to the complexity of the overall process, it is desirable and imperative to be able to have a good idea as to when this is going to occur – particularly in the case of a preterm delivery – in order for the clinicians to be able to commence prompt care and treatments.[1–3] From previous studies it has been shown that contraction signals acquired from the uterine wall convey information that can be used to forecast a labour imminency timeframe with greater accuracy than traditional methods.[1–3] Acquired contraction signals reflect the immediate physiological state with greater quality, and could be used to predict a labour due period, thus presenting an additional layer of information which could serve a decision support platform and allow for proactive care strategies for pregnant patients.[1–3] These uterine contraction signals also manifest themselves as electrophysiological signals, which are typically characterised by a flow of ionic current that allows for the recording of these bioelectrical manifestations of the contraction signals to be done by either an electric or magnetic recording instrumentation.[4]

A bulk of the work in the area of pregnancy monitoring and has been done with electrohysterogram (EHG), which acquire bio-electric signals associated with anatomical contractions. The limitation of this is tied to the biophysics of the propagation of electrical signals through tissue, which attenuate due to tissue conductivities as they travel from source to receiver.[5] Due to the principle of electrodynamics, the orthogonal offset of these bio-electric signals is accompanied by a related magnetic field signal that can be acquired using magnetomyography (MMG) instrumentation which, unlike EHG, is not dependent on tissue conductivities and does not require coupling between the sensor and the surface of the uterine walls.[6] A number of authors have investigated the use of MMG for labour imminency prediction since the early work done by Eswaran et al.[7], who applied the high resolution superconducting quantum interference device array for reproductive assessment (SARA).Where it can be noted that all 151channels were used for the

analysis process. The authors of this paper have also investigated the use of MMG across various case studies spanning labour imminency prediction using a 0–48hrs and >48hrs prediction timeline and an ethnic-specific labour prediction model due to the influence of ethnicity in gestation [8]. The results from these prior studies were seen to produce a high accuracy in the >90% region and produced technical contributions as 148 electrode channels were used.[8] However, these prior studies have not addressed ways in which the instrumentation can increase its appeal by dropping the overall cost, which has been seen to be a crucial factor in the clinical adoption of the instrumentation.[8]

Two of the primary sources of the cost associated with the MMG instrumentation include the need for electromagnetic shielding due to its susceptibility to magnetic interferences and its associated broad number of channels.[9] A workaround for the magnetic shielding could involve hardware modifications where materials which minimise magnetic field interferences from fields outside its primary source could act as an insulation, while an exercise on the application of a decomposition algorithm – which can aid in the maximisation of the signal quality produced by each electrode channel – could be looked at as a means of investigating how a low number of electrodes could lead to a parsimonious yet robust approach towards the acquisition of MMG signals from a pregnant patient.

Nsugbe et al.[10–16] have previously designed a decomposition algorithm which was initially used for the source separation of a mixture given an acquired non-stationary time domain signal. This decomposition is done using a series of linear thresholds of varying amplitudes as the basis for the decomposition, alongside heuristic reasoning which promotes the flexibility of the approach to be adapted to different signal processing problems.[10–16] The approach has recently been trialled on the decomposition of neuromuscular signals from electromyography for phantom motion decoding and brain waves from electroencephalography for motor imagery decoding, in the application area of upper-limb prosthesis control.[16] During that study, its performance was also seen to surpass the wavelet transform from both a decoding accuracy perspective and in computational efficiency.[16,17] The results from this showed the promise of the designed approach towards effective decomposition of physiological signals which, like the signals from the original application (mixture signals), have been seen to be stochastic and non-linear.[16,17]

In this paper, the focus is on the investigation of a decomposition method which can be used to enhance the signal quality and lead to a parsimonious electrode selection, driving down the overall instrumentation cost in the process. We evaluate the feasibility of the application of the proposed signal decomposition method towards enhancing the quality of signal acquired from an MMG electrode channel for labour prediction imminency, using MMG signals available from the Physionet database.[18] Specifically speaking, the following are the main contributions of this paper:

- Application of a designed signal decomposition algorithm on MMG signals to allow for a reduced channel configuration and therein contribute towards potentially reducing the cost of the instrumentation;

- Showcasing the feasibility of a reduced feature extraction and low complexity classifier towards labour imminency prediction, which would ultimately enhance model interpretability and overall clinical appeal.

Materials and Methods

MMG Dataset

The MMG dataset used was downloaded from the publicly available Physionet database[18], which hosts the MMG signals acquired using the 151 channel SQUID Array for Reproductive Assessment (SARA), and is located at UAMS, Little Rock, USA. The dataset comprised of 25 patients, all of whom were in their third trimester, delivered in a mixture of term and preterm, and whose ethnicity spanned Caucasian, Black and Hispanic.[18] The data was separated into two classes: i) data from patients who delivered within 48 hours, and ii) data from patients who delivered over 48 hours from when the data was collected.[8,18] The data was initially acquired at a rate of 250Hz, then down-sampled to 32Hz: further details on the acquisition instrumentation and method can be seen on the Physionet database.[8] An illustration of the acquisition of the MMG field signal can be seen in Figure 1. Due to errors encountered while downloading the dataset, 22 patients' data were used for final analysis.

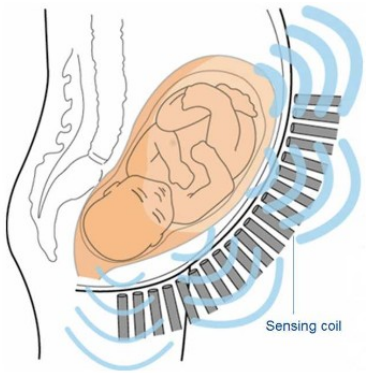

Figure 1: Acquisition of MMG field signals[4]

MMG Theoretical Principle

Ionic flow in biological tissue is seen to be accompanied by a perpendicular magnetic field, as deduced in the seminal work by Maxwell[4,6]. The governing equations underpinning the propagation of electromagnetic signals within the uterus during pregnancy can be approximated with the quasi-static versions of Maxwell's equations, with the assumption of a linear model as follows[4,6]:

$$\nabla x E = 0 \tag{1}$$

$$\nabla . E = \frac{\rho}{\varepsilon_0} \tag{2}$$

$$\nabla . B = \mu_0 J \tag{3}$$

$$\nabla . B = 0 \tag{4}$$

E and B are the electrical and magnetic fields, ε_0 and μ_0 are the permittivity and permeability of free space respectively, and ρ and J are the current and charge density. Given a situation with continuous current flow, the Biot-Savart's law which encapsulates the behaviour of a magnetic field from an electric source, can thus be used to estimate the magnetic field as seen in equation 5[4]:

$$B(r) = \frac{\mu_0}{4\pi} \int \frac{J(r') x l}{l^3} dv' \tag{5}$$

$l = r - r'$ is a vector from the source r' to a point of observation of magnitude ¿$\|l\|_2$, and $J(r')$ represents the current density from the source. Re-expressing the ratio of l/l^3 as $-\nabla\left(\frac{1}{l}\right) = \nabla'\left(\frac{1}{l}\right)$, equation 5 can thus be redefined as follows[4,6]:

$$B(r) = \frac{\mu_0}{4\pi} \int J(r') x \nabla' \frac{l}{l} dv'$$

In a scenario where there is rapid current density dissipation (6) can be reformulated as:

$$B(r) = \frac{\mu_0}{4\pi} \int \frac{\nabla' x J(r')}{l} dv' \quad (7)$$

With the superposition principle, a linear can be established between the magnetic field B and an accompanying source current density J^p, which thus allows for the computation of the potential V as below[4,6]:

$$B(r) = \frac{\mu_0}{4\pi} \int \left(J^p(r') + V(r')\nabla'\sigma(r')\right) x \frac{l}{l^3} dv' \quad (8)$$

A full list of the derivation used to obtain equation (8) can be seen in Zhang et al.[4]

Signal Processing

Signal Decomposition Algorithm

For a given non-stationary time-series comprising a single event of a unit impulse accompanied by a decay, $s(t) = e^{-at}u(t) + \zeta$ for $u(t) = \begin{cases} 0, \wedge a < 0 \\ 1, \wedge a \geq 0 \end{cases}$, the time-series can be characterised by the magnitude of its impulse alongside temporal characteristics, where, $u(t)$ is a step function, *t* is the impulse time, and a is an indication value which shows that the function is 0 until $a \geq 0$, and ζ is an additive white noise. When consecutive events occur within the time-series, an overlap can be seen which causes the associated impulse decay characteristics to be challenging to localise [10–16]. Thus in reality the amplitude characteristics represent a source of information which can be used to model and describe certain time series as described by Nsugbe et al.[10–16]. Thus, for a non-stationary stochastic time-series, an optimal region within a given time-series is one in which signal information is maximised, and associated uncertainties are minimal, as can be inferred from the amplitude characteristics of the signal peaks within the time-series.[10–17]

The proposed method works with a series of heuristically tuned linear threshold as a basis for the decomposition alongside a peak identification method, to decompose a time-series into a number of sub time-series.[10–17] The obtained optimal decomposition region within the time-series can be described by its parameters in time and space. The corresponding variables for these parameters are referred to as X_{opt} and represent the region within the signal which all subsequent time-series should be analysed, as this has been seen to generalise across all subsequent signals acquired from the source, providing the signal acquisition instrumentation remains unchanged. The decomposition method can be seen to represent a first stage dimensionality reduction stage which will pay dividends in the reduced computational load during the feature extraction stage.

Below are the steps taken to implement the proposed signal decomposition method and identify the optimal region within a given time-series: $S_n = x_1, x_2 \ldots . x_N$ transformed to its absolute form $|S_n|$:

Step 1: Set an initialisation threshold for $|S_n|$. In this work, this was chosen to be 50% of $max|S_n|$, although this is flexible and has been seen to be an a priori value in a previous study.[10–16] The application of the initialisation threshold to the time-series would yield a set of sub time-series $X=\{X_{11,}X_{12}\}$ where X_{11} is a sub time-series corresponding to the upper portion of the initialisation threshold, and X_{12} is a sub time-series corresponding to the lower portion of the initialisation threshold.

Step 2: For every sub time-series in X, form a new set $X_{filt}=\{X_{11_{filt}},X_{12_{filt}}\}$ by identifying the peaks within each sub time-series, where the definition of a 'peak' is a data sample with amplitude greater than or equal to its nearest neighbours, and is mathematically expressed as

$$S_{n.\,peaks}(x)=\begin{cases}x_{peak.n}, x_n \geq x_{n-1} \wedge ¿ \wedge ¿ \wedge x_{n+1}\\ 0, Otherwise\end{cases}$$

Once peaks are identified, a performance index is to be conducted to assess the 'goodness' of each obtained sub time-series. For this, two candidate time-series from separate classification classes are required (i.e. 0–48hrs and >48hrs), and the respective candidate time-series needs to have been processed through Steps 1 and 2 beforehand, to produce $X_{filt_{class\,1}}=\{X_{11_¿},X_{12_{filt_{class\,1}}}\}$ and $X_{filt_{class2}}=\{X_{11_¿},X_{12_{filt_{class\,2}}}\}$. Features, mean of the peaks (MP) and waveform length (WL) were extracted from each sub time-series to form a feature vector. A geometric based performance index J was chosen to be the Euclidean distance,[19] normalised by the standard deviation of each time-series in question, as expressed in equations 9–11 for an example exercise constituting of sub time-series $X_{11_¿}$ and $X_{11_¿}$

$$ED(p,q)=\sqrt{¿¿¿} \tag{9}$$

$$\sigma=\sqrt{\sum_{w=1}^{N_w} ¿¿¿¿} \tag{10}$$

$$J=(p,q)=\frac{ED(p,q)}{\sigma_m} \tag{11}$$

ED is the Euclidean distance, p and q are co-ordinates of the features in the feature vector projected in a Euclidean space, w is the wth feature, N_w, r_w is a feature within the feature vector, μ is the mean of the feature vector, and σ_m is the mean of the standard deviations from the two time-series. The computed value for J should be stored in an array and be used for a maxima seeking exercise at a later stage.

In order to promote consistency, it is important to note that J should only be calculated for equivalent threshold regions and iterations pairings, for a true reflection; i.e., $\{X_{xy_¿}$ and $X_{xy_¿}\}$.

Step 3: Subsequent decomposition of $X_{11_{filt}(upper\,threshold),} \wedge X_{12_{filt(lower\,threshold)}}$, this time using a separate threshold tuning heuristic and scale factor as defined below:

Upper threshold decomposition parameters alongside iteration numbers:

$$T_{l_¿}=50\%\,of\,max|S_n|, T_{l_¿}=\frac{T_{l_¿}}{2}+T_{l_¿}, T_{l_¿}=\frac{T_{l_¿}}{2}+T_{l_¿}, \ldots\ldots\ldots T_{l_¿}=\frac{T_{l_¿-1}}{2}+T_{l_¿-1} \tag{12}$$

Lower threshold decomposition parameters alongside iteration parameters:

$$T_{l_i} = 50\%\, of\ max\left|S_n\right| , T_{l_i} = \frac{T_{l_i}}{2}, T_{l_i} = \frac{T_{l_i}}{2} \ldots\ldots\ldots T_{l_i} = \frac{T_{l_i - 1}}{2} \quad (13)$$

For every subsequent decompositions of $X_{11_{filt}(upper\ threshold)}, \wedge X_{12_{filt(\ lower\ threshold)}}$, only the time-series peaks above the re-tuned upper thresholds are considered for subsequent analysis in the case of the upper threshold, while for the lower threshold only the time-series peaks below the re-tuned lower thresholds are considered for subsequent analysis. This structure has been determined as a means of best practice and should be repeated for any further threshold iterations.[10–16]

The stopping criteria for the iterations was chosen for when the number of peaks within the threshold iteration was less than the number of samples required for the computation of some of the features required to be subsequently extracted from the signal. Note that the algorithm stopping criteria is flexible and adds a degree of freedom to the decomposition algorithm, which allows for it to be adapted to suit the case study at hand.

Step 4: A maxima seeking exercise[20] from both threshold regions where $J_{max}(upper\ threshold)$ and $J_{max}(lower\ threshold)$ are identified and the ultimate maximum from the pair is selected as $J_{optimal}$. The associated threshold used to obtain this is stored as X_{opt}, thus referring to the optimal decomposition parameters to be used for subsequent signal decomposition from the same source.

An illustration showing the various steps used to find the optimal decomposition region within the signal can be seen in Figure 2.

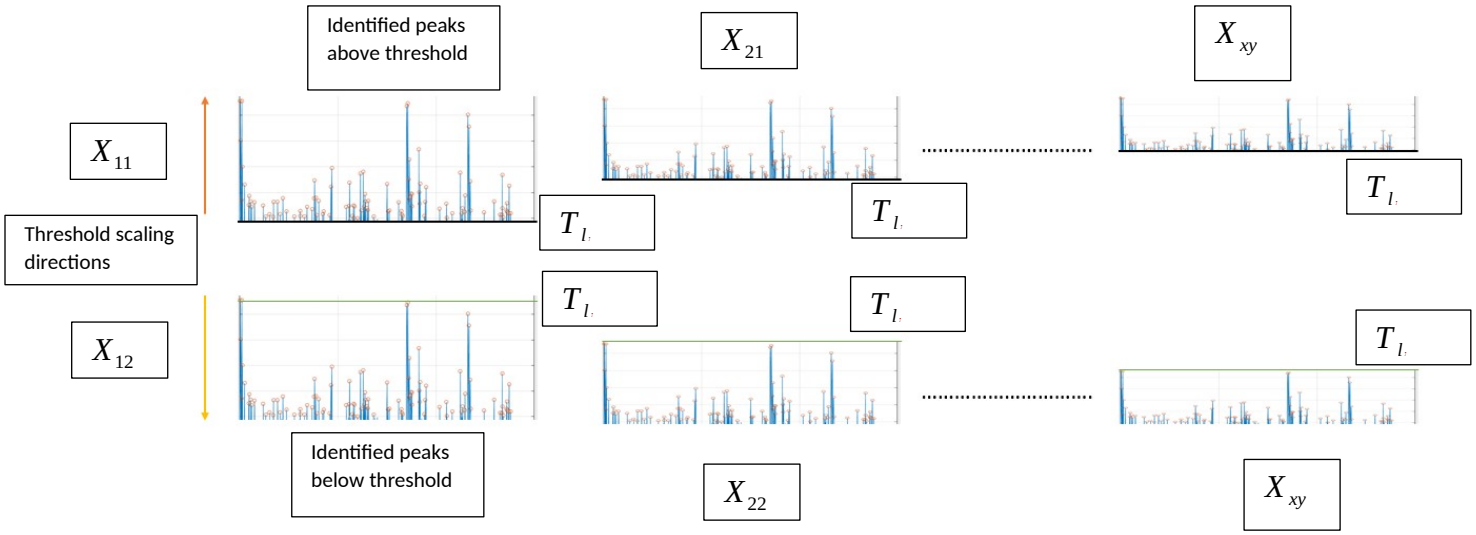

Figure 2: Visual illustration of the signal decomposition approach

Feature Extraction

Two groups of features were used for the classification exercises. Group 1 features comprised a reduced set of features whereas Group 2 features comprised the full list of features used in previous work.[21–24]

Group 1 Features: MP, WL, slope sign change (SSC), root mean squared (RMS), sample entropy (SampEN), Cepstrum (Ceps), maximum fractal length (MFL), median frequency of power spectrum (MF), simple squared integral (SSI) and variance (VAR).

Group 2 Features: MP, WL, SSC, RMS, SampEN, Ceps, MFL, MF, SSI, VAR, 4th order auto-regression co-efficient (AR), Higuchi fractal dimension (HFD), detrended fluctuation analysis (DFA), peak frequency (PF), sum of peaks (SP).

SMOTE

The SMOTE algorithm was applied to increase the amount of training examples and for class balancing. The nearest neighbour parameters were selected to be 5, while *k* was 2.7 for class 1, and 5 for class 2 for the sample generation process.[25]

Classifiers

Two classifiers were selected for the classification exercises, namely the logistic regression (LR), and linear support vector machine (L-SVM).[26–28] LR was chosen due to its low model complexity thus classification decision transparency, while the SVM is an iterative classifier with a bit more complexity but a linear kernel was chosen to help minimise overfit and reduce complexity.[26–28]

- LR: for a binary decision defined by the set $Y \in \{0,1\}$, a generalised function parametrised by θ can be used to approximate values for Y as follows[26]:

$$h_\theta(X) = \frac{1}{1+e^{-\theta^T X}} = Pr(Y=1|X_i;\theta) / Pr(Y=0|X_i;\theta) = 1 - h_\theta(X) \quad (14)$$

Where X_i is an input. For an independently Bernoulli-distributed observation, the log likelihood function can be expressed as:

$$L(\theta|y;x) = Pr(Y|X;\theta) \quad (15)$$

$$\prod_i Pr(y_i|x_i;\theta) \quad (16)$$

$$\prod_i h_\theta(x_i)^{y_i}\left(1-h_\theta(x_i)\right)^{(1-y_i)} \quad (17)$$

where x_i is a data sample and y_i is an approximated output.

- L-SVM: given a feature vector x, the SVM solves an optimisation problem as follows[27,28]:

$$min \frac{1}{2} w^t w + R \sum_{i=1}^{N} \zeta_i \quad (18) \qquad w, b, \zeta$$

$$s.t.\ y_i\left(w^T \Phi(x_i) + b\right) \geq 1 - \zeta_i, \quad \zeta_i \geq 0, i = 1, \ldots, N$$

Where ζ is a slack variable, *R* is a regulariser and y is an indication vector. The parameters used for the SVM implementation of the linear kernel are expressed as $k(x_i, x_j) = x_i^T x_j$ with a one-vs-one multiclass method.

Results

To demonstrate the effectiveness of the proposed approach, only the first five electrode channels which were arbitrarily chosen were used for signal processing work, which represents 3.3% of the total electrode channels. To assess the performance of the classifier, four select classifier metrics

were employed; namely the classification accuracy (CA), sensitivity (Sen), specificity (Spec) and area under the curve (AUC).[29] Below are the values obtained for J prior to the fulfilling of the algorithm stopping criteria, where the maximum was found to be upper threshold region at the third iteration.

Results from upper threshold decompositions can be seen below:

$$T_{l_i}=1.906,\ T_{l_i}=2.291,\ T_{l_i}=2.628$$

Results from lower threshold decompositions can be seen below:

$$T_{l_i}=0.237, T_{l_i}=0.489, T_{l_i}=0.776$$

The results from the LR and L-SVM classifiers can be seen in Figures 3 and 4 respectively. The result for the proposed method (bar 1) is around 90% (except for sensitivity), which is about 20% greater than the other two methods being contrasted.

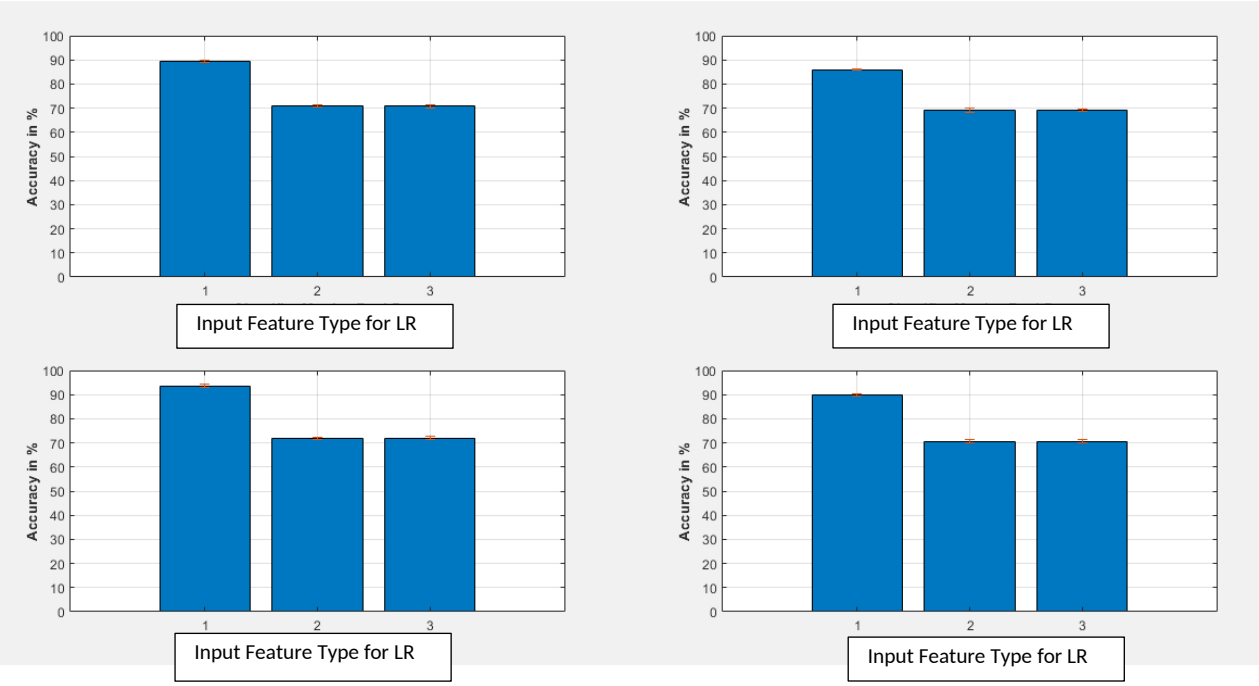

Figure 3: Classification results for LR classifier, Top Left-CA, Top Right-Sen, Bottom Left-Spec, and Bottom Right-AUC

1. Proposed signal decomposition method + Group 1 features
2. Group 1 features
3. Group 2 features

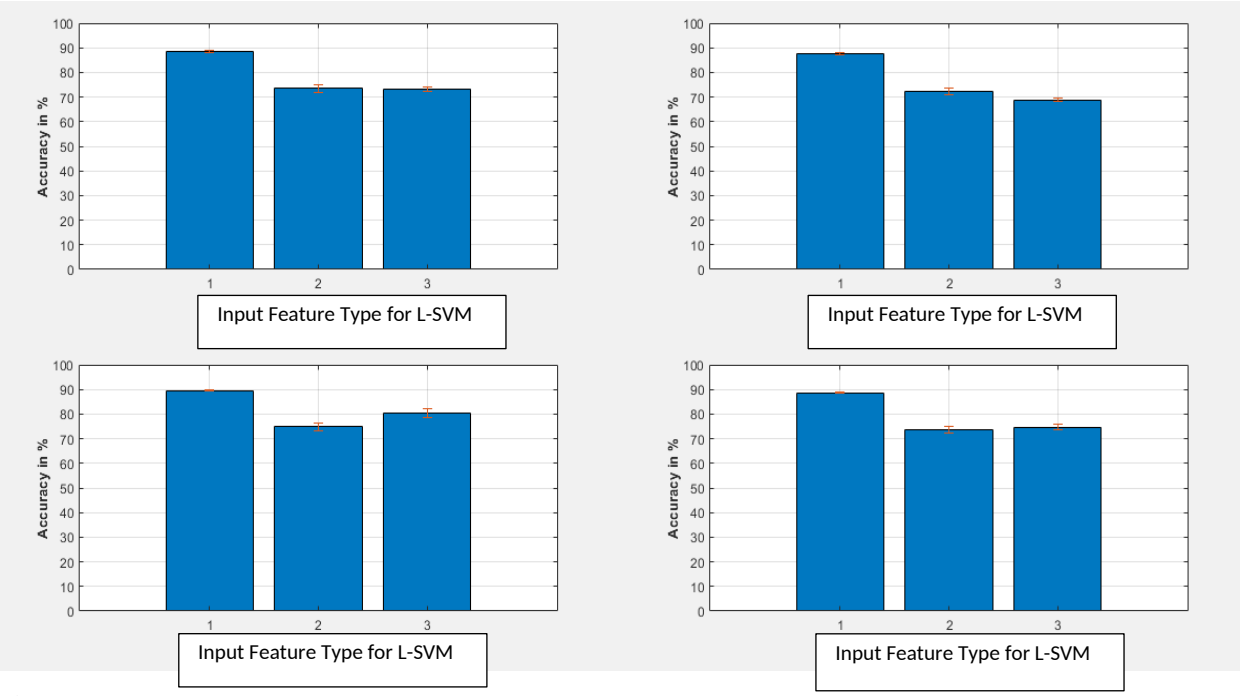

Figure 4: Classification results for L-SVM classifier, Top Left-CA, Top Right-Sen, Bottom Left-Spec, and Bottom Right-AUC

The same trend continues in the case of the L-SVM where the results from the proposed methods are superior to its counterparts, albeit with a slightly lower performance overall than the LR, implying the compatibility of the proposed method with regression-based classifiers. The effects of additional features (bar 3) are more pronounced in the case of the L-SVM where an increase in the specificity is observed, which contributed to yield a greater AUC value.

A visualisation of the respective clusters from the two classes can be seen from the principal component analysis (PCA) plot with the first two principal components shown in Figure 5.[30] A qualitative assessment of the PCA plot from the proposed method shows highly linear projections of the data clusters, while the subsequent two show more variability, scatter and overlap associated with their projections. The presence of additional features can be seen to enhance the class separability of the data points when plots 2 and 3 are contrasted, but the scatter associated with their projections remain. Thus, it can be said that the proposed signal decomposition method promotes compact linear projections with a reduced variability, and therein the effectiveness of linear decision boundaries, while the inclusion of additional features enhances the separability between clusters in feature/Euclidean space.

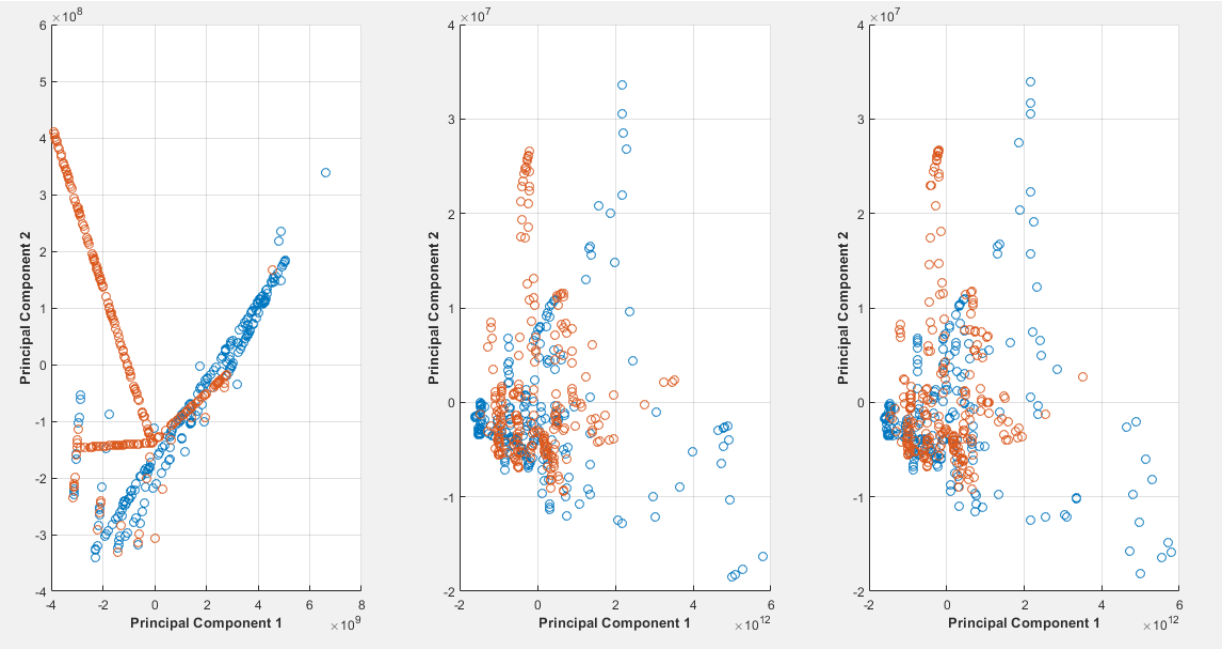

Figure 5: PCA results plots: Left) Results using novel signal decomposition method and Group 1 features; Middle) Results without signal decomposition and Group 1 features; Right) Results without signal decomposition and Group 2 features

Conclusions

In this paper, a highly reduced channel electrode channel selection comprising 3.3% (5 out of 151 channels) of the original electrode segments, alongside a novel decomposition algorithm and low order classifiers have been combined to form a parsimonious method of labour imminency prediction using MMG. This idea of parsimony promotes reduced cost, computational load and transparency of the decision-making process of the model, and will contribute towards the overall clinical appeal and viability of the full system architecture. Further work in this area will involve the identification of norms and statistical features which can contribute towards greater accuracies using the candidate classifiers.

Acknowledgments

The authors would like to thank Dr Michael Provost and Dr Ali Al-Timemy for providing thoughts and feedback on the manuscript and Brian Kerr from Kerr Editing for proofreading the manuscript.

References

1. Babu TA, Kumar PR. Features extraction and classification of uterine magnetomyography signals. *International Journal of Current Engineering and Scientific Research*. 2018;5(3).

2. Offiah I, ODonoghue K, Kenny L. Clinical Risk Factors for Preterm Birth. In: Morrison J, ed. *Preterm Birth - Mother and Child*. InTech; 2012. doi:10.5772/27439

3. Figueroa JP, Massmann A, Pimentel G, Nathanielsz PW. Characteristics of the electromyogram recorded from the mesometrium of the pregnant ewe from 106 days' gestation to delivery: Similarities with and differences from the Electromyogram obtained from the myometrium. *Am J Obstet Gynecol*. 1987;157(4):991-998. doi:10.1016/S0002-9378(87)80102-3

4. Zhang M, La Rosa PS, Eswaran H, Nehorai A. Estimating uterine source current during contractions using magnetomyography measurements. Young RC, ed. *PLoS ONE*. 2018;13(8):e0202184. doi:10.1371/journal.pone.0202184

5. Jager F, Libenšek S, Geršak K. Characterization and automatic classification of preterm and term uterine records. Uthman O, ed. *PLoS ONE*. 2018;13(8):e0202125. doi:10.1371/journal.pone.0202125

6. Maxwell J. VIII. A dynamical theory of the electromagnetic field. *Phil Trans R Soc*. 1865;155:459-512. doi:10.1098/rstl.1865.0008

7. Eswaran H, Preissl H, Wilson JD, Murphy P, Lowery CL. Prediction of labor in term and preterm pregnancies using non-invasive magnetomyographic recordings of uterine contractions. *Am J Obstet Gynecol*. 2004;190(6):1598-1602. doi:10.1016/j.ajog.2004.03.063

8. Nsugbe E, Obajemu O, Samuel OW, Sanusi I. Application of non-invasive magnetomyography in labor imminency prediction for term and preterm pregnancies and ethnicity specific labor imminency prediction. *In peer review*.

9. Understanding EMI/RFI Shielding to Manage Interference. Ceptech. Accessed February 1, 2021. https://ceptech.net/understanding-emi-rfi-shielding-to-manage-interference/

10. Nsugbe E, Starr A, Ruiz Carcel C. Monitoring the particle size distribution of a powder mixing process with acoustic emissions: a review. *Engineering & Technology Reference*. 2012;1(1). doi:10.1049/etr.2016.0139

11. Nsugbe E, Starr A, Foote P, Ruiz-Carcel C, Jennions I. Size differentiation of a continuous stream of particles using acoustic emissions. *IOP Conf Ser: Mater Sci Eng*. 2016;161:012090. doi:10.1088/1757-899X/161/1/012090

12. Nsugbe E, Starr A, Jennions I, Carcel CR. Particle size distribution estimation of a mixture of regular and irregular sized particles using acoustic emissions. *Procedia Manufacturing*. 2017;11:2252-2259. doi:10.1016/j.promfg.2017.07.373

13. Nsugbe E. Particle size distribution estimation of a powder agglomeration process using acoustic emissions. Published online 2017. https://dspace.lib.cranfield.ac.uk/handle/1826/14378

14. Nsugbe E, Starr A, Jennions I, Carcel CR. Online particle size distribution estimation of a mixture of similar sized particles with acoustic emissions. *J Phys: Conf Ser*. 2017;885:012009. doi:10.1088/1742-6596/885/1/012009

15. Nsugbe E, Ruiz-Carcel C, Starr A, Jennions I. Estimation of fine and oversize particle ratio in a heterogeneous compound with acoustic emissions. *Sensors*. 2018;18(3):851. doi:10.3390/s18030851

16. Nsugbe E, Starr A, Jennions I, Ruiz-Carcel C. Estimation of online particle size distribution of a particle mixture in free fall with acoustic emission. *Part Sci Technol*. 2019;37(8):953-963. doi:10.1080/02726351.2018.1473540

17. Nsugbe E, Samuel O, Asogbon M, Li G. A contrast of multi-resolution analysis methods for transhumeral phantom motion decoding. *In peer review*.

18. Escalona-Vargas D, Govindan RB, Furdea A, Murphy P, Lowery CL, Eswaran H. MMG Database. Published online 2016. doi:10.13026/C2ZW95

19. Phinyomark A, Limsakul C, Phukpattaranont P. Application of wavelet analysis in EMG feature extraction for pattern classification. *Measurement Science Review*. 2011;11(2). doi:10.2478/v10048-011-0009-y

20. Zhang C, Ordóñez R. *Extremum-Seeking Control and Applications: A Numerical Optimization-Based Approach*. Springer; 2012.

21. Nsugbe E, Phillips C, Fraser M, McIntosh J. Gesture recognition for transhumeral prosthesis control using EMG and NIR. *IET Cyber-Systems and Robotics*. 2020;2(3):122-131. doi:10.1049/iet-csr.2020.0008

22. Nsugbe E, Samuel OW, Asogbon MG, Li G. Phantom motion intent decoding for transhumeral prosthesis control with fused neuromuscular and brain wave signals. *In peer review*.

23. Phinyomark A, Khushaba RN, Scheme E. Feature extraction and selection for myoelectric control based on wearable EMG sensors. *Sensors*. 2018;18(5):1615. doi:10.3390/s18051615

24. Fergus P, Cheung P, Hussain A, Al-Jumeily D, Dobbins C, Iram S. Prediction of preterm deliveries from EHG signals using machine learning. Wei Z, ed. *PLoS ONE*. 2013;8(10):e77154. doi:10.1371/journal.pone.0077154

25. Blagus R, Lusa L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinf*. 2013;14(1):106. doi:10.1186/1471-2105-14-106

26. cs229-notes1.pdf. Accessed February 2, 2021. https://akademik.bahcesehir.edu.tr/~tevfik/courses/cmp5101/cs229-notes1.pdf

27. Scikit Learn. 1.4. Support Vector Machines — scikit-learn 0.23.2 documentation. Published n.d. Accessed October 12, 2019. https://scikit-learn.org/stable/modules/svm.html

28. Bishop CM. *Pattern Recognition and Machine Learning*. Springer; 2006.

29. Agarwal R. The 5 Classification Evaluation metrics every Data Scientist must know. Medium. Published September 11, 2020. Accessed December 6, 2020. https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226

30. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Phil Trans R Soc A*. 2016;374(2065):20150202. doi:10.1098/rsta.2015.0202