Time Progression of a Parametric Impulse Response Function Estimate from Intra-Partum Cardiotocography for Normal and Hypoxic Fetuses

PA Warrick[1,4], RE Kearney[1], D Precup[2], EF Hamilton[3,4]

[1]Biomedical Engineering Department, McGill University, Montreal, Quebec, Canada
[2]School of Computer Science, McGill University, Montreal, Quebec, Canada
[3]Department of Obstetrics and Gynecology, McGill University, Montreal, Quebec, Canada
[4]LMS Medical Systems, Montreal, Quebec, Canada

Abstract

We compared the time evolution of a parametric system identification model of cardiotocography (CTG) during the last three hours of labour and delivery. The models are the impulse response functions (IRFs) of the system relating the uterine pressure (UP) and fetal heart rate (FHR). We used surrogate data to assess the likelihood that the model successfully identified system dynamics. We then investigated the ability of the models to indicate fetal distress by comparing normal and hypoxic cases.

For pathological and normal fetuses, model output standard deviation $\sigma_{\hat{y}}$, first IRF minimum time T_{min} and variance accounted for (VAF) were discriminating in the first two hours. In the final hour however, only T_{min} remained discriminating. For intermediate and normal cases, $\sigma_{\hat{y}}$ was always discriminating while T_{min} discriminated in the final hour only. These results suggest that detecting some pathological and intermediate cases before injury occurs may be possible; this would be very useful clinically.

1. Introduction

The difficulties of visual cardiotocography (CTG) interpretation have been discussed in many previous clinical and technical studies: the sensitivity is clinically useful but the low specificity can increase cesarean section rates[1]. We would like to use automated methods to model the maternal-fetal interaction available via CTG and eventually use these models to improve the assessment of fetal tolerance to labour.

We examine the interactions between the CTG signal pair of uterine pressure (UP) and FHR, which can be viewed from the perspective of maternal stimulus and fetal response. In a previous study [2] we have done this system identification using a low-order parametric model.

In this study, we examined the use of surrogate data to discriminate spurious models from those identifying true system dynamics. We then assessed the ability of the models classified as 'significant' to discriminate hypoxic and normal fetuses over the last three hours of labour.

2. Methods

2.1. Data

The database consisted of 717 intrapartum CTG tracings for pregnancies having a birth gestational age greater than 36 weeks and having no known genetic malformations. The FHR was acquired at $f_S = 4$Hz while the UP was acquired at 1Hz and up-sampled to 4Hz. Only records with at least 3 hours of recording were considered. The examples were labelled by outcome according to their arterial umbilical-cord base deficit (BD) and neonatal indications of severe neurological impairment. Base deficit is considered an important marker for hypoxia leading to intrapartum asphyxia with metabolic acidosis [3]. A minority of the cases were severely pathological (58 'A': BD $\geq$ 12 mmol/L, compromised neurological function), while the other were either intermediate (425 'C': BD $\geq$ 8 mmol/L) or normal (234 'D': BD $<$ 8 mmol/L). The letter labels 'A', 'C', and 'D' are our own internal labelling scheme.

2.2. System identification

As described previously [2], we perform UP-FHR system identification by fitting a delayed second-order dynamic model to the singular value decomposition (SVD) of the the nonparametric impulse response function (IRF). We generate models over 20-minute epochs having 50% overlap.

2.3. Model significance using surrogates

Safeguards are employed during system identification to reduce overfitting over the short 20-min epoch. Maintaining the epoch-to-lag ratio N:M as high as possible (10:1) using the shortest possible input lag M promotes better signal-to-noise ratio (SNR) in the non-parametric IRF linear regression. The problem of narrow-bandwidth input UP is mitigated by the SVD step. Finally, the order search of the parametric modelling reduces the order below that overestimated by MDL [2].

Nevertheless, whether true system dynamics have been identified by the modelling remains questionable. Often the model output VAF is quite low (i.e. <20%, see figure 2), which can be interpreted in one of two ways. The model could include true dynamics but system non-linearities or large disturbances in the output from other sources (e.g. intrinsic heart-rate variability) remain unaccounted for. Alternatively, there is no fetal sensitivity to uterine activity and the model should be discarded.

To overcome this ambiguity, we compared the system identification at the linear regression step for the measured FHR with that from multiple FHR surrogates. The surrogates were generated using the amplitude adjusted Fourier transform (AAFT) algorithm [4], which preserves the power spectrum (or equivalently, autocorrelation) and the amplitude probability distribution of the original FHR. While we could have alternatively generated UP surrogates, we processed FHR due to its richer spectral content. Equivalent synthesis of such surrogates is linearly filtered Gaussian noise followed by a static nonlinearity.

Comparing the VAF metric with the original FHR to the ensemble of VAFs associated with the surrogates allows us to assess the significance of the model. For arbitrary metric distributions, [4, 5] suggest a rank-order test where $\alpha = \frac{K}{M+1}$ is the probability that by chance, the VAF of the original FHR has the K-th largest VAF when compared to M surrogate VAFs. Using this test, we assigned the level of significance $\gamma = 1-\alpha$ to our models. To reduce computational requirements we generated M=20 surrogates, sufficient to determine a significance level of 95%.

3. Results

Figure 1 plots the percentage of models retained using two different filtering criteria: γ (top) and VAF (bottom). At the 95% significance level, the γ filter retained 43.9% of the models. At an equivalent acceptance rate, the simpler VAF filter had a 23% threshold (dot-dashed lines).

Figure 2 confirms a trend towards greater γ with increasing VAF, but models with low VAF and high γ were not uncommon. Using the γ=95% threshold rejected 89.0% of models with low VAF (<20, green region), 30% of intermediate VAF (orange region), and 3.0% of models with high VAF (>40, red region). Example models from the most extreme regions are shown in figures 4 and 5. In figure 4 both sensor disturbance and response nonlinearity likely contributed to low VAF, yet the model with high γ apparently identified the FHR response. This model would be rejected using the VAF filtering scheme with equivalent model retention. In figure 5, the FHR response to UP is clear (note the IRF minimum at 60s, an estimate of time to deceleration nadir), yet γ was low. Inspection of the surrogates FHRs indicated that the very dominating contraction frequency permitted surrogates to be modelled well despite their phase randomization. As indicated above, rejection of these high VAF models was relatively infrequent.

The ability of select model parameters to distinguish model classes are assessed by the hypothesis tests shown in tables 1 and 2. We display only those parameters that demonstrated the most discrimination, namely, $\sigma_{\hat{y}}$ (model output standard deviation), T_{min} (time to first IRF minimum) and the VAF. Asterisks indicate that the null hypothesis can be rejected at the p <0.01 significance level. For A-D comparison, in hours -3 and -2, all parameters showed statistically significant differences from the null hypothesis for all three tests. In hour -1, however, only T_{min} remained discriminating. For C-D comparison, the $\sigma_{\hat{y}}$ was discriminating across all three hours while T_{min} was discriminating in the final hour only.

These model parameter differences are reflected in the time plots of figure 3 over the last three hours (18 epochs). The measured FHR standard deviation σ_y is also shown as a reference for $\sigma_{\hat{y}}$. The effect of the γ filter on these features is shown in 3(a) and 3(b). It is most noteworthy that the filtering enhanced the difference between the C and D T_{min} values in the final hour. C-D $\sigma_{\hat{y}}$ differences reduced in the final hour as a result of the filter, likely indicating that fewer Ds with low VAF were retained.

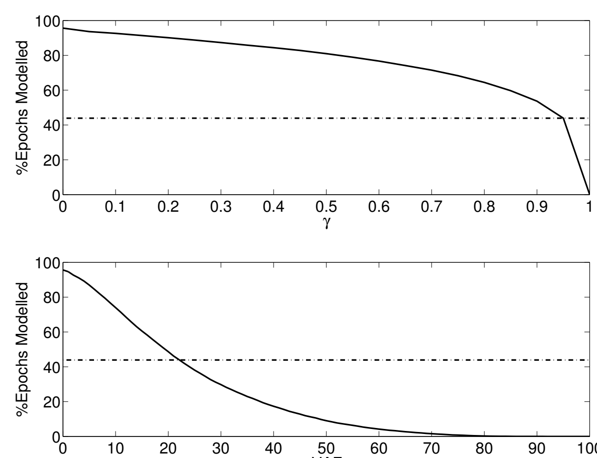


Figure 1. Comparison of percentage of models retained by thresholding γ (top) or VAF (bottom).

Hour		t-Test	RankSum	K-S
-1	T_{min}	9.42e-06*	9.38e-06*	6.34e-06*
-1	VAF	7.37e-02	1.43e-01	2.56e-02
-1	$\sigma_{\hat{y}}$	3.29e-02	2.20e-01	3.26e-01
-2	T_{min}	7.80e-04*	4.03e-04*	6.28e-06*
-2	VAF	5.12e-10*	5.13e-09*	1.58e-08*
-2	$\sigma_{\hat{y}}$	5.18e-13*	7.04e-05*	1.34e-06*
-3	T_{min}	1.58e-04*	4.29e-03*	3.30e-05*
-3	VAF	4.26e-06*	4.88e-05*	6.54e-05*
-3	$\sigma_{\hat{y}}$	3.26e-06*	5.72e-04*	2.89e-04*

Table 1. Hypothesis tests: A vs. D.

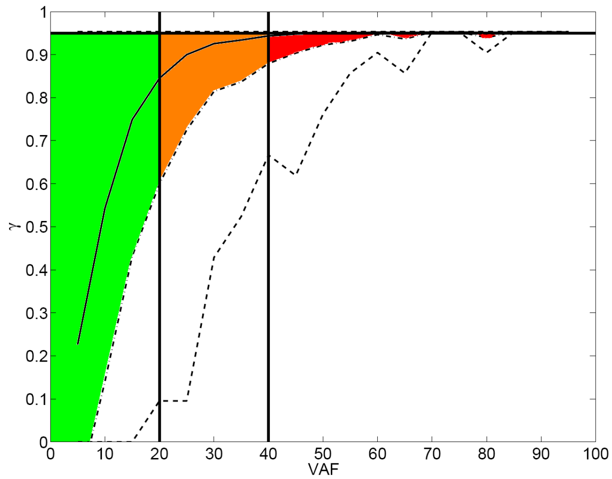


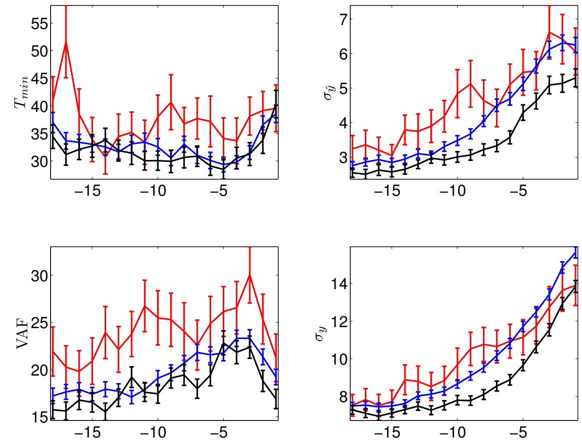
Figure 2. γ vs. VAF. The mean (solid), 2 standard deviations below the mean (dot-dashed) and minimum (dashed) values are shown. Relative rejection rates indicated by green (89.0%), orange (30.0%) and red (3.0%) areas.

4. Discussion and conclusions

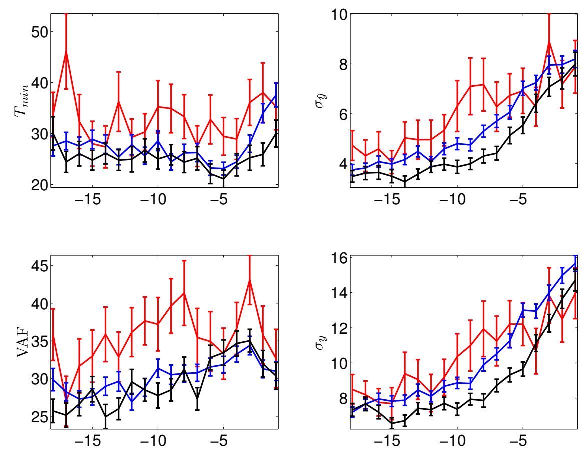
While it would be possible to address model significance with a fixed VAF threshold, our method is arguably preferable because in each epoch it adapts to the input-output interaction noise floor which is subject to varying disturbance energies from sensor artifact, system nonlinearities and intrinsic heart rate variability.

In [5] a method is proposed to reduce periodicity artifacts in surrogates (responsible for the false rejection of figure 5). We will investigate better surrogate algorithms.

The γ filter removed epochs that are likely to be uninformative for discrimination purposes. This accentuated T_{min} differences between intermediate 'C's and 'D's in the last hour. We suspect that the A-D discrimination reduction in last hour was due to reducing the already small numbers of accepted 'A' epochs. Other features remained discrimi-



(a)



(b)

Figure 3. Model parameters over last 3 hours before (a) and after (b) γ filter. Time measured in epochs before delivery. Pathological 'A' (red), intermediate 'C' (blue) and normal 'D' (black) cases plotted individually. Units are seconds (T_{min}) and beats per minute (σ_y and $\sigma_{\hat{y}}$).

Hour		t-Test	RankSum	K-S
-1	T_{min}	3.08e-03*	2.17e-03*	1.91e-04*
-1	VAF	2.39e-01	3.05e-01	5.78e-01
-1	$\sigma_{\hat{y}}$	5.24e-05*	4.43e-04*	8.07e-03*
-2	T_{min}	3.43e-01	9.65e-01	2.00e-01
-2	VAF	1.72e-01	2.07e-01	1.75e-02
-2	$\sigma_{\hat{y}}$	7.24e-08*	4.31e-05*	5.08e-06*
-3	T_{min}	1.61e-01	4.03e-01	1.60e-01
-3	VAF	3.92e-03*	5.60e-03*	4.58e-02
-3	$\sigma_{\hat{y}}$	1.63e-04*	5.03e-04*	5.05e-04*

Table 2. Hypothesis tests: C vs. D

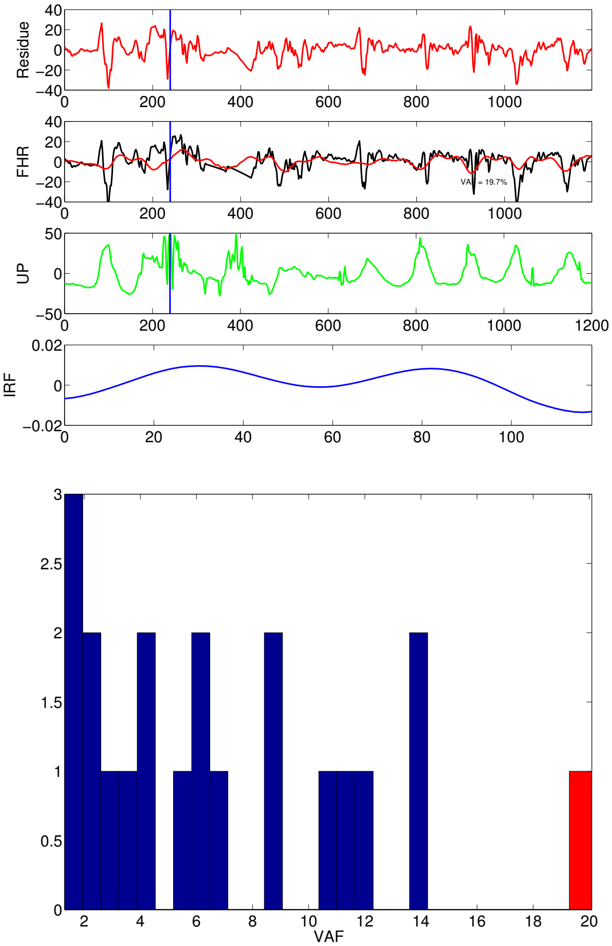


Figure 4. Histogram of surrogate (blue) and original (red) VAFs for an accepted CTG with low VAF (19.7%). Epoch residual, FHR, UP and IRF signals shown above.

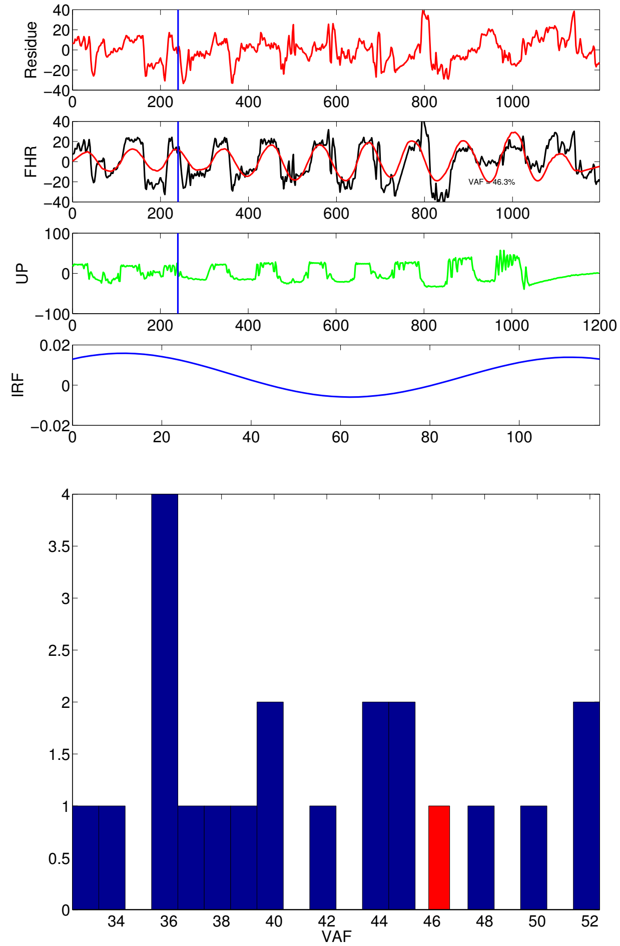


Figure 5. Histogram of surrogate (blue) and original (red) VAFs for a rejected CTG with high VAF (46.3%). Epoch residual, FHR, UP and IRF signals shown above.

nating for 'A's after γ filtering. We anticipate that the filter will improve overall classification; it may also highlight near-hypoxic 'D's with strong responses.

Acknowledgements

The authors acknowledge the financial support of this work by LMS Medical Systems, Inc.

References

[1] Low J, Victory R, Derrick E. Predictive value of electronic fetal monitoring for intrapartum fetal asphyxia with metabolic acidosis. Obstet Gynecol 1999;93:285–291.

[2] Warrick PA, Kearney RE, Precup D, Hamilton EF. Low-order parametric system identification for intrapartum uterine pressure-fetal heart rate interaction. In The 2007 IEEE Engineering in Medicine and Biology 29th Annual Conference. 2007; 5043–5046.

[3] Parer JT, King T, Flanders S, Fox M, Kilpatrick SJ. Fetal acidemia and electronic fetal heart rate patterns: Is there evidence of an association? Journal of Maternal Fetal Neonatal Medicine May 2006;19(5):289–294. ISSN 14767058.

[4] Theiler J, Eubank S, Longtin A, Galdrikian B, Doyne Farmer J. Testing for nonlinearity in time series: the method of surrogate data. Physica D Nonlinear Phenomena September 1992; 58(1-4):77–94.

[5] Schreiber T, Schmitz A. Surrogate time series. Physica D Nonlinear Phenomena August 2000;142(3-4):346–382.

Address for correspondence:

Philip A. Warrick
Department of Biomedical Engineering
McGill University
3775 rue University
Montreal, QC H3A 2B4 Canada
philip.warrick@mcgill.ca