

ANALYSIS OF ELECTROCARDIOGRAM SIGNALS VIA WAVELET BASED BAYESIAN MODELS

RAQUEL PRADO¹ AND IRENE GARCÍA²

¹*Departamento de Cómputo Científico y Estadística and CESMa, Universidad Simón Bolívar,
Apartado 89000, Caracas, Venezuela*

²*Departamento de Ciencias Básicas and CESMa, Universidad Simón Bolívar, Núcleo Litoral,
Venezuela*

(Received November 15, 2000)

Abstract. The study of high resolution electrocardiogram signals (HRECGs) has proven useful in assessing the risk of arrhythmic events in patients suffering chagasic myocarditis. We present a wavelet based Bayesian modelling approach to the analysis and classification of multiple single-channel HRECG signals recorded on patients classified into different groups according to their level of cardiac damage. We compute a wavelet transform of each signal and propose a probability model that decomposes the expected value of each wavelet coefficient in two terms: a mean component common to all signals and a dispersion term capturing features that characterise each group. We use hierarchical mixture priors on the components associated to each wavelet coefficient. This prior structure has a level dependent positive probability mass at zero for vanishing terms and normal priors for non-zero terms. Model parameters are estimated using Markov chain Monte Carlo methods.

Key words and phrases: Wavelet based models; Bayesian hierarchical models; ECG; Chagasic myocarditis.

1. Introduction

The problem of classifying multiple signals has been studied by several authors from different perspectives. Motivated by the study of the high resolution electrocardiogram (HRECG) and its capabilities in predicting the risk of abnormal cardiac events in patients with chagasic myocarditis, we propose a wavelet based Bayesian modelling approach to the analysis and classification of multiple signals recorded in patients under various clinical conditions.

Chagasic myocarditis is characterised by a progressive inflammation of the cardiac muscle produced by lymphocytic infiltration of a parasite, the *Trypanosoma cruzi*. The inflammation process gradually leads to bundle branch block (BBB), hypertrophy and local disruption of ventricular conduction. A common endpoint is sudden cardiac death due to electric instability in myocardial contraction, which generally occurs after ventricular tachycardia (VT). Understanding the natural progression of chagasic myocarditis and exploring techniques for diagnosis of myocardial and electrical conduction damage in patients at the early stage of the disease is a key issue. Abnormal intra-QRS potentials (AIQP) and late potential (LP) measurements in the HRECG have been used as non-invasive markers of ventricular arrhythmia in patients with chagasic myocarditis (Gomis et al., 1996). Late potentials indicate slow or delayed conduction in the myocardium and are usually defined as abnormal signals that outlast the normal QRS period during sinus

rhythm. Time domain criteria have been widely used to quantify LPs, however, such criteria are not suitable in patients with prolonged QRS periods due to conduction abnormalities as BBB. Frequency domain and spectro-temporal analysis have also been used to identify abnormal potentials within the QRS of chagasic patients (Zareba et al., 1991) however, these techniques have implicit limitations due to spectral resolution and induced artifacts.

Time-frequency measurements have also been used to evaluate abnormal potentials in HRECG signals. Morlet et al. (1993) describes a method for discriminating patients with ventricular tachycardia based on detecting local maxima of the wavelet transform magnitude vectors of HRECG signals. In the context of chagasic myocarditis, Suárez et al. (1998) identify different stages of the disease using time-frequency indexes computed from short-time Fourier transforms, Wigner-Ville and wavelet representations of HRECG signals. Differences among some of the clinically different groups of chagasic patients are assessed by these indexes, however their predictive performance is not fully satisfactory (Suárez, 1997).

We aim to extend the previous developments using a wavelet based Bayesian modelling approach to the analysis and classification of unfiltered HRECG signals recorded in 44 subjects of which 11 were healthy subjects and 33 were seropositive patients with no evidence of cardiac damage, 25 chagasic patients with evidence of cardiac involvement including BBB and 11 patients in conditions similar to the previous group but with documented episodes of ventricular tachycardia. We compute a wavelet transform of each HRECG signal and propose a probability model that decomposes the expected value of each wavelet coefficient in two terms: a term common to all the traces and a component that characterises each group. In addition, following Müller and Vidakovic (1998), we use a hierarchical mixture of priors on the components associated to each wavelet coefficient. Such prior structure includes a level-dependent positive prior probability mass at zero for vanishing coefficients, implementing wavelet coefficient thresholding as a formal Bayes rule (Vidakovic, 1998). For non-zero coefficients shrinkage is introduced through Gaussian priors with different variances at each level of detail.

At this point it is worth emphasizing that our objective is not compression or noise reduction in ECG curves. We aim to detect abnormalities in the QRS complex in terms of a reduced number of wavelet coefficients that are useful for diagnosis purposes. In this sense, the work presented here is similar to the work of Morlet et al. (1993) in the ECG framework and to the work of Brown et al. (2001) in the context of spectroscopic calibration problems, to mention just a couple of references. Once a model with the structure summarised above is fitted to a particular dataset and posterior distributions for the model parameters are obtained, the focus relies on prediction. We evaluate the predictive capabilities of the proposed models through a Bayesian classification procedure. The method is based on computing the posterior odds or probability that a “new” ECG signal belongs to one particular group.

The paper is organised as follows. In Section 2, we give some preliminaries on wavelet transformations and describe the probability model and the simulation scheme to obtain posterior inference. Section 3 describes a Bayesian approach for classification of the HRECG signals based on the models developed in Section 2. Section 4 describes the analysis of the HRECG signals and evaluates the predictive performance of the models in a cross-validation study. Finally, conclusions and future developments are presented in Section 5.

2. The Model

2.1 Modelling multiple HRECG signals via wavelet transformations

We begin with some preliminaries on wavelet transforms. References on general wavelet theory include Daubechies (1993) and Meyer (1992) among others. For a comprehensive introduction to wavelets in statistics see Vidakovic (1999).

Let $\{\psi_{j,k}(x)\}$, with $j, k \in \mathbb{Z}$, be an orthonormal basis in $L_2(\mathbb{R})$ obtained as translations and dilations of a *mother wavelet* function ψ as $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$. Then, a function f can be represented as

$$f(x) = \sum_{j,k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(x),$$

with wavelet coefficients $d_{j,k} = \int f(x)\psi_{j,k}(x)dx$ describing the function f at location $2^{-j}k$ and frequency proportional to 2^j . In a discrete framework the data can be mapped from the time domain to the wavelet domain via a fast algorithm known as the discrete wavelet transformation (DWT). Let $\mathbf{y} = (y_1, \dots, y_n)'$, with $n = 2^J$ and $J \in \mathbb{Z}$, be the vector of observations usually corresponding to a sample of n equally spaced points in time. The data vector \mathbf{y} can be decomposed into an n -dimensional vector of wavelet coefficients \mathbf{d} via $\mathbf{d} = \mathbf{W}\mathbf{y}$, with \mathbf{W} an orthogonal matrix uniquely determined by the choice of the wavelet basis. The DWT algorithm is based on the application of recursive linear filters, so the matrix \mathbf{W} is not explicitly computed in practice. An algorithm to reconstruct \mathbf{y} from \mathbf{d} , known as the inverse discrete wavelet transformation (IDWT), is also available. Wavelets are localised in time and frequency, usually providing a good description of a variety of functions in terms of very few coefficients.

In our HRECG scenario we have M signals corresponding to high resolution signals obtained from patients originally classified into p different groups, each group formed by m_i patients with $i = 1, \dots, p$. Let $y_{il}(t)$ be the HRECG trace for patient l in group i at time t . Then, we assume the following model

$$y_{il}(t) = f_i(t) + \epsilon_{il}(t) = f(t) + \alpha_i(t) + \epsilon_{il}(t),$$

with $\epsilon_{il}(t) \sim N(0, \sigma^2)$ for all t and σ^2 unknown. In words, $f(t)$ represents the form of an average signal at time t , while $\alpha_i(t)$ accounts for group effect. Applying a DWT to the data vector $\mathbf{y}_{il} = (y_{il}(1), \dots, y_{il}(n))'$, we obtain

$$\begin{aligned} \mathbf{d}_{il} &\equiv \mathbf{W}\mathbf{y}_{il} = \mathbf{W}\mathbf{f} + \mathbf{W}\boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_{il} \\ (2.1) \qquad &= \boldsymbol{\theta} + \boldsymbol{\tau}_i + \boldsymbol{\epsilon}_{il}^*, \end{aligned}$$

where \mathbf{d}_{il} , $\boldsymbol{\theta} \equiv \mathbf{W}\mathbf{f}$ and $\boldsymbol{\tau}_i \equiv \mathbf{W}\boldsymbol{\alpha}_i$ are the vectors of wavelet coefficients of \mathbf{y}_{il} , \mathbf{f} and $\boldsymbol{\alpha}_i$ respectively, and $\boldsymbol{\epsilon}_{il}^*$ is an n -dimensional vector of i.i.d. innovations normally distributed with zero mean and variance σ^2 , $N(\boldsymbol{\epsilon}_{il}^* | \mathbf{0}, \sigma^2 \mathbf{I})$. The IDWT algorithm allows for the reconstruction of the signals from the coefficients, therefore, inferences on $\boldsymbol{\theta}$ and $\boldsymbol{\tau}_i$ translate into inferences on \mathbf{f} and $\boldsymbol{\alpha}_i$. Model (2.1) is completed by specifying a prior probability model on $\boldsymbol{\theta}$, $\boldsymbol{\tau}_i$ and σ^2 .

2.2 The prior structure

In order to allow thresholding of the wavelet coefficients we use a mixture prior model on the components $\boldsymbol{\theta}$ and $\boldsymbol{\tau}_i$. Following Vidakovic (1998) and Müller and Vidakovic

(1998), we take a prior structure that includes a positive probability mass at zero for vanishing coefficients and a Gaussian distribution for non-vanishing coefficients.

Let $\theta(j, k)$ be the wavelet coefficient of \mathbf{f} at level j , for $j = 1, \dots, J$ and shift k within the level with $k = 2^{(j-1)} + 1, \dots, 2^j$. We assume the following prior model,

$$\theta(j, k) \sim (1 - \pi_j^\theta)I_0(\theta(j, k)) + \pi_j^\theta N(\theta(j, k)|0, ur_j),$$

with $\pi_j^\theta = \alpha^j$ the prior probability of $\theta(j, k) \neq 0$ and $\alpha \in (0, 1]$. $I_0(\cdot)$ denotes the indicator function, that is, $I_0(\theta(j, k)) = 1$ if $\theta(j, k) = 0$ and $I_0(\theta(j, k)) = 0$ otherwise. If a coefficient is set to zero the corresponding column of the wavelet transformation matrix \mathbf{W} is not used in the description of \mathbf{f} . Note that the choice of $\pi_j^\theta = \alpha^j$ implies that the model sets smaller prior probabilities at non-zero wavelet coefficients for higher levels of details. Given that $\theta(j, k)$ is not vanishing, a normal zero mean prior distribution with level-dependent variance ur_j , with $r_j = 1/2^j$, is assumed. The scaling factors r_j are used to compensate for the factor $2^{j/2}$ in the definition of $\psi_{j,k}(\cdot)$ (see Müller and Vidakovic (1998) for a detailed discussion on this). Other choices for π_j^θ and r_j are possible (see for instance Chipman et al. 1997).

Similarly, if $\tau_i(j, k)$ denotes the wavelet coefficient of α_i at level j and scale k , we assume

$$\tau_i(j, k) \sim (1 - \pi_j^{\tau_i})I_0(\tau_i(j, k)) + \pi_j^{\tau_i} N(\tau_i(j, k)|0, v_i r_j),$$

with $\pi_j^{\tau_i} = \beta_i^j$ and $\beta_i \in (0, 1]$.

In addition, we set $\tau_{i_0}(j, k) = 0$ for all j, k and some group i_0 to guarantee parameter identifiability. We complete the model with hyperpriors for σ^2 , α , β_i , u and v_i . We take beta prior distributions on α and β_i , that is $\alpha \sim \text{Beta}(a_\alpha, b_\alpha)$ and $\beta_i \sim \text{Beta}(a_\beta, b_\beta)$, and inverse-gamma priors on σ^2 , u and v_i , $1/\sigma^2 \sim \text{Gamma}(a_\sigma, b_\sigma)$, $1/u \sim \text{Gamma}(a_u, b_u)$ and $1/v_i \sim \text{Gamma}(a_v, b_v)$ for all $i \neq i_0$.

2.3 Estimation of the parameters via Markov chain Monte Carlo

We briefly describe a Markov chain Monte Carlo (MCMC) simulation scheme to implement posterior inference in the proposed model. Details on the algorithm appear in García (1999). We are interested in making inferences on the set of model parameters $\mathbf{\Lambda} = \{\boldsymbol{\theta}, \boldsymbol{\tau}, \sigma^2, \alpha, \boldsymbol{\beta}, \mathbf{v}\}$ with $\boldsymbol{\theta} = \{\theta(j, k); \forall j, k\}$, $\boldsymbol{\tau} = \{\tau_i(j, k); \forall i \neq i_0, j, k\}$, $\boldsymbol{\beta} = \{\beta_i; \forall i \neq i_0\}$ and $\mathbf{v} = \{v_i; \forall i \neq i_0\}$.

Starting at some initial parameter values $\mathbf{\Lambda}^0$, the MCMC scheme iterates through the steps described below, generating samples from the desired joint posterior distribution once convergence is achieved.

- (i) *Update* $(\boldsymbol{\theta}|\boldsymbol{\tau}, \alpha, u)$. This reduces to updating $(\theta(j, k)|\tau_1(j, k), \dots, \tau_p(j, k), \alpha, u)$ independently for each j, k which is done through Gibbs steps as follows.

For each j, k , define $z_{il}(j, k) = d_{il}(j, k) - \tau_i(j, k)$ and $\mathbf{z}(j, k) = \{z_{il}(j, k), \forall i, l\}$. Then, sample $\theta(j, k) \sim N(\mu(\bar{\mathbf{z}}(j, k)|0, \rho^{-1}))$ with probability $(1 - p_0)$ and $\theta(j, k) = 0$ with probability p_0 , where

$$\mu(\bar{\mathbf{z}}(j, k)) = \frac{ur_j}{Mur_j + \sigma^2} \sum_{i=1}^p \sum_{l=1}^{n_i} z_{il}(j, k) \quad \rho = \frac{M}{\sigma^2} + \frac{1}{ur_j}$$

and

$$p_0 = \frac{(1 - \alpha^j) \prod_{i=1}^p \prod_{l=1}^{n_i} N(z_{il}(j, k)|0, \sigma^2)}{(1 - \alpha^j) \prod_{i=1}^p \prod_{l=1}^{n_i} N(z_{il}(j, k)|0, \sigma^2) + \alpha^j \prod_{i=1}^p \prod_{l=1}^{n_i} N(z_{il}(j, k)|0, \sigma^2 + ur_j)}.$$

- (ii) *Update* $(\boldsymbol{\tau}|\boldsymbol{\theta}, \boldsymbol{\beta}, \mathbf{v})$. This reduces to update $(\tau_i(j, k)|\boldsymbol{\theta}(j, k), \beta_i, v_i)$ independently for each j, k and each $i \neq i_0$. For each i define $z_{il}^*(j, k) = d_{il}(j, k) - \theta(j, k)$ and $\mathbf{z}_i^*(j, k) = \{z_{il}(j, k), l = 1, \dots, n_i\}$. Then, for each group i and each j, k

$$\tau_i(j, k) \sim p_0^i I_0(\tau_i(j, k)) + (1 - p_0^i) p(\tau_i(j, k) | \mathbf{z}_i^*(j, k), \beta_i, v_i),$$

with $p(\tau_i(j, k) | \mathbf{z}_i^*(j, k), \beta_i, v_i)$ and p_0^i computed as follows

$$p(\tau_i(j, k) | \mathbf{z}_i^*(j, k), \beta_i, v_i) = N(\tau_i(j, k) | \mu(\bar{\mathbf{z}}_i^*(j, k)), \rho_i^{*-1})$$

with

$$\mu(\bar{\mathbf{z}}_i^*(j, k)) = \frac{r_j v_i}{n_i r_j v_i + \sigma^2} \sum_{l=1}^{n_i} z_{il}^*(j, k) \quad \rho_i^* = \frac{n_i r_j v_i + \sigma^2}{\sigma^2 r_j v_i}$$

and

$$p_0 = \frac{(1 - \beta_i^j) \prod_{l=1}^{n_i} N(z_{il}^*(j, k) | 0, \sigma^2)}{(1 - \beta_i^j) \prod_{l=1}^{n_i} N(z_{il}^*(j, k) | 0, \sigma^2) + \beta_i^j \prod_{l=1}^{n_i} N(z_{il}^*(j, k) | 0, \sigma^2 + v_i r_j)}.$$

Then, we sample $\theta_i(j, k) \sim N(\mu(\bar{\mathbf{z}}_i^*(j, k)), \rho_i^{*-1})$ with probability $(1 - p_0^i)$ and $\tau_i(j, k) = 0$ with probability p_0^i .

- (iii) *Update* $(\alpha|\boldsymbol{\theta})$. This is done through a Metropolis step. Generate a proposal $\tilde{\alpha} \sim g_\alpha(\tilde{\alpha}|\alpha)$ and compute $a(\alpha, \tilde{\alpha}) = \min\{1, h(\alpha, \tilde{\alpha})\}$ with

$$h(\alpha, \tilde{\alpha}) = (\tilde{\alpha}/\alpha)^{a\alpha-1} \times ((1 - \tilde{\alpha})/(1 - \alpha))^{b\alpha-1} \times \prod_{j,k} [\tilde{\alpha}^j / \alpha^j]^{(1 - I_0(\theta(j,k)))} [(1 - \tilde{\alpha}^j)/(1 - \alpha^j)]^{I_0(\theta(j,k))}.$$

Then, with probability $a(\alpha, \tilde{\alpha})$ replace α by $\tilde{\alpha}$, otherwise keep α unchanged. We take $g_\alpha(\tilde{\alpha}|\alpha) = N(\alpha, w)$ and $(0.01)^2 < w < (0.1)^2$.

- (iv) *Update* $(\boldsymbol{\beta}|\boldsymbol{\tau})$. This is done through a collection of Metropolis steps, one for each β_i . For each i , generate a proposal $\tilde{\beta}_i \sim g_{\beta_i}(\tilde{\beta}_i|\beta_i)$ and compute

$$a(\beta_i, \tilde{\beta}_i) = \min\{1, \left(\tilde{\beta}_i/\beta_i\right)^{a\beta-1} \times \left((1 - \tilde{\beta}_i)/(1 - \beta_i)\right)^{b\beta-1} \times \prod_{j,k} [\tilde{\beta}_i^j / \beta_i^j]^{(1 - I_0(\tau_i(j,k)))} [(1 - \tilde{\beta}_i^j)/(1 - \beta_i^j)]^{I_0(\tau_i(j,k))}\}.$$

With probability $a(\beta_i, \tilde{\beta}_i)$ replace β_i by $\tilde{\beta}_i$. Otherwise keep β_i unchanged. Again, we use $g_{\beta_i}(\tilde{\beta}_i|\beta_i) = N(\beta_i, w)$ with $(0.01)^2 < w < (0.1)^2$.

- (v) *Update* $(u|\alpha, \boldsymbol{\theta})$. Sample $1/u \sim \text{Gamma}(a_u^*, b_u^*)$ with

$$a_u^* = \frac{\sum_{j,k} [1 - I_0(\theta(j, k))]}{2} + a_u, \quad b_u^* = \sum_{j,k} \frac{\theta(j, k)^2 [1 - I_0(\theta(j, k))]}{2r_j} + b_u.$$

- (vi) *Update* $(\mathbf{v}|\boldsymbol{\beta}, \boldsymbol{\tau})$. For each i , sample $1/v_i \sim \text{Gamma}(a_v^*, b_v^*)$ with

$$a_v^* = \frac{\sum_{j,k} [1 - I_0(\tau_i(j, k))]}{2} + a_v, \quad b_v^* = \sum_{j,k} \frac{\tau_i(j, k)^2 [1 - I_0(\tau_i(j, k))]}{2r_j} + b_v.$$

- (vii) *Update* $(\sigma^2|\boldsymbol{\theta}, \boldsymbol{\tau})$. Sample $1/\sigma^2 \sim \text{Gamma}(a_\sigma^*, b_\sigma^*)$ with

$$a_\sigma^* = a_\sigma + \sum_{i=1}^p \frac{n_i}{2}, \quad b_\sigma^* = b_\sigma + \sum_{i=1}^p \sum_{l=1}^{n_i} \frac{(d_{il}(j, k) - \theta(j, k) - \tau_i(j, k))^2}{2}.$$

3. Classification of a set of HRECG signals

Let $\mathbf{D} = \{\mathbf{d}_{il}, i = 1, \dots, p; l = 1, \dots, n_i^*\}$ be a set of M^* HRECG signals represented in the wavelet domain and $\tilde{\mathbf{D}} = \{\tilde{\mathbf{d}}_a, a = 1, \dots, q\}$ a set of q “new” signals also represented in the wavelet domain. Suppose that we want to determine what is the probability that each signal in $\tilde{\mathbf{D}}$ belongs to one of the p groups.

Let $p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_i)$ be the predictive density of $\tilde{\mathbf{d}}_a \in \tilde{\mathbf{D}}$ given model \mathbf{M}_i and the available data \mathbf{D} , where model \mathbf{M}_i is defined by the parameters $\Lambda_i = \{\boldsymbol{\theta}, \boldsymbol{\tau}_i, \sigma^2\}$ for $i = 1, \dots, p$. Then, $p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_i)$ is computed as

$$(3.1) \quad p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_i) = \int p(\tilde{\mathbf{d}}_a|\mathbf{D}, \Lambda_i, \mathbf{M}_i)p(\Lambda_i|\mathbf{D})d\Lambda_i.$$

The posterior probability that $\tilde{\mathbf{d}}_a$ belongs to group j is then given by

$$(3.2) \quad P(\mathbf{M}_j|\tilde{\mathbf{d}}_a, \mathbf{D}) = \frac{q_j p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_j)}{\sum_{i=1}^p q_i p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_i)},$$

where q_i is the prior probability that $\tilde{\mathbf{d}}_a$ belongs to group i . At this point we have an assessment of the probability that $\tilde{\mathbf{d}}_a$ belongs to group i . We may choose to assign $\tilde{\mathbf{d}}_a$ to the group for which $p(\mathbf{M}_j|\tilde{\mathbf{d}}_a, \mathbf{D})$ is a maximum or we may choose to withhold the classification unless the maximum is greater than some specified value.

If $q_i = 1/p$ for all i , equation (3.2) can be written as

$$(3.3) \quad P(\mathbf{M}_j|\tilde{\mathbf{d}}_a, \mathbf{D}) = \left(\sum_{i=1}^p \frac{p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_i)}{p(\tilde{\mathbf{d}}_a|\mathbf{D}, \mathbf{M}_j)} \right)^{-1} = \left(\sum_{i=1}^p B_{ij}^*(\tilde{\mathbf{d}}_a) \right)^{-1},$$

with $B_{ij}^*(\tilde{\mathbf{d}}_a)$ the predictive Bayes factor for \mathbf{M}_i versus \mathbf{M}_j . A log Bayes factor of 1 (-1) indicates evidence in favour of model i (j), a value of 2 or more (-2 or less) indicates strong evidence in favour of model i (j), while the value 0 indicates no evidence either way.

In our case, the expression (3.2) is not available in closed form so we use the approximation,

$$(3.4) \quad P(\tilde{\mathbf{d}}_a|\mathbf{M}_j, \mathbf{D}) \approx \frac{1}{S} \sum_{s=1}^S p(\tilde{\mathbf{d}}_a|\mathbf{D}, \Lambda_i^{(s)}, \mathbf{M}_i),$$

where each $\Lambda_i^{(s)}$ is a sample from the joint posterior distribution of $\Lambda_i = \{\boldsymbol{\theta}, \boldsymbol{\tau}_i, \sigma^2\}$.

4. Analysis of the HRECG data

Three groups of signals were studied. Group I was formed by 44 subjects of which 11 were healthy subjects and 33 were and blood seropositive patients with no echocardiographic or standard ECG evidence of cardiac damage. Group II was formed by 25 chagasic patients with evidence of cardiac involvement including either abnormal Holter, left ventricular ejection fraction or BBB. Finally, group III consisted of 11 patients with evidence of cardiac damage - similar to Group II - and documented episodes of ventricular tachycardia. The HRECGs were obtained using orthogonal XYZ leads, sampled at

1,000 Hz with signal-averaged techniques implemented in the Predictor system (Corazonix Corp., Oklahoma City). A preliminary analysis of some of the HRECG traces that are part of this dataset is described in García et al. (2000). Here, we extend the analysis and consider classification issues. As in García et al. (2000) we only consider signals from the X leads.

We applied a wavelet transform to each HRECG signal, obtaining a set of 256 wavelet coefficients. The MATLAB library for wavelet analysis *WaveLab v.701* (Wavelab Reference Manual, 1999) was used for this. Various wavelet bases, including minimum phase Daubechies wavelets, Coiflet and Symmlet families were considered, leading to similar results in terms of posterior inference and classification. The results presented here correspond to Daubechies wavelets with 5 vanishing moments. Following the notation of Section 2, the indexes $i = 1, 2, 3$ identify the groups I, II and III respectively. We set $\tau_1(j, k) = 0$ for all j, k to guarantee parameter identifiability. This restriction implies that θ characterises the wavelet coefficients of group I while τ_2 and τ_3 model effects of the wavelet coefficients of groups II and III over group I. Different values of the hyperparameters were considered for the hyperpriors on α , β_i , u and v_i and no significant impact in terms of posterior estimation. Table 1 shows the estimated posterior means for the parameters σ , \sqrt{u} , $\sqrt{v_2}$, $\sqrt{v_3}$, α , β_2 and β_3 using the complete data. The results are based on 2,000 posterior samples taken from 15,000 MCMC iterations after a burn in of 3,000 iterations. The posterior means and probability intervals obtained for α , β_2 and β_3 indicates that the level of thresholding is higher for the τ_2 and τ_3 coefficients (groups II and III) than for the θ coefficients (group I).

Table 1. Estimated posterior means for σ , \sqrt{u} , $\sqrt{v_2}$, $\sqrt{v_3}$, α , β_2 and β_3 and corresponding 95% probability intervals.

	Mean	95% probability interval
σ	0.1381	(0.1371 , 0.1392)
\sqrt{u}	4.9021	(3.9493 , 6.4400)
$\sqrt{v_2}$	1.3454	(0.9945 , 1.8140)
$\sqrt{v_3}$	3.3796	(2.4402 , 4.8838)
α	0.9137	(0.9033 , 0.9240)
β_2	0.6037	(0.5491 , 0.6593)
β_3	0.5713	(0.4997 , 0.6346)

Figure 1 displays the posterior estimated mean signals of the X leads for the 3 groups, $\hat{\mathbf{f}} = E[\mathbf{f}|\mathbf{y}]$ (solid line), $\hat{\mathbf{f}}_2 = E[\mathbf{f} + \alpha_2|\mathbf{y}]$ (dashed line) and $\hat{\mathbf{f}}_3 = E[\mathbf{f} + \alpha_3|\mathbf{y}]$ (dotted line) and 95% posterior intervals. The posterior estimated mean signals for groups II and III show more extended S waves in comparison to the S waves of group I. Group differences are also evident in terms of the amplitudes of the R waves. This is consistent with the results obtained in Suárez et al. (1998). Of all the coefficients at the 8 levels from coarsest to finest scales, only (100%, 100%, 75%, 37.5%, 25%, 0%, 0%, 0%) were used, respectively for each level, in computing the posterior estimates of θ . The corresponding percentages of coefficients for τ_1 and τ_2 at the 8 levels were (100%, 100%, 100%, 62.5%, 25%, 0%, 0%, 0%) and (100%, 100%, 75%, 50%, 12.5%, 0%, 0%, 0%).

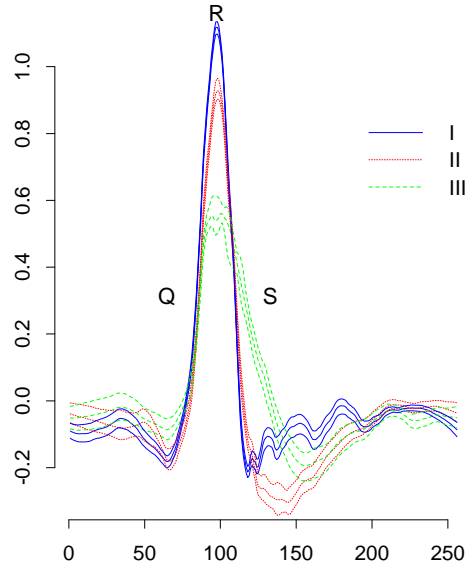


Fig. 1. Estimated posterior mean signals and 95% posterior intervals of the X leads for groups I (solid lines), II (dashed lines) and III (dotted lines).

4.1 Predictive performance

In order to assess the predictive performance of the models described in Section 2, we compare our results with those of late potential indexes based on time-frequency representations of the HRECG signals developed in Suárez (1997) and Suárez et al. (1998). For this purpose, we only discriminate among two groups of data at a time. We begin fitting model (2.1) to traces in groups I and II and evaluate the predictive performance of our model based methods in discriminating subjects in these two groups. Then, we fitted model (2.1) to traces in groups I and III and followed an analogous classification procedure.

We split the 69 signals of groups I and II randomly into 10 groups for cross-validation purposes, obtaining 9 groups with 7 signals each and a group with 6 signals. Each one of the first 9 groups contains 4 signals of group I and 3 signals of group II. We then used 9 of the 10 groups as training data and classified the traces in the 10th group. We repeated this procedure 10 times by using each group as a prediction dataset and the remaining 9 groups to fit the model. Similarly, for discrimination of signals of groups I and III, we split the 55 signals randomly into 10 groups obtaining 8 groups with 4 signals of group I and one signal of group III, one group with 6 signals of group I and one signal of group II and finally, one more group with 6 signals of group I and 2 signals of group III. Again, we repeatedly used 9 of the 10 groups to fit the model and one group for prediction.

Table 2 shows a comparison between the specificity, sensitivity and accuracy rates obtained using the wavelet based Bayesian models described in Section 2 and the rates obtained using late potential indexes based on short-time Fourier representations (LPST) described in Suárez (1997) and Suárez et al. (1998). Sensitivity and specificity are defined

Table 2. Specificity, sensitivity and accuracy rates.

	<i>Specificity (%)</i>	<i>Sensitivity (%)</i>	<i>Accuracy (%)</i>
Wavelet based Bayesian models (I, II)	81.82	60.00	73.91
LPST indexes (I, II)	24.24	65.38	42.37
Wavelet based Bayesian models (I, III)	90.91	90.91	90.91
LPST indexes (I, III)	45.45	15.15	22.72

as the percentage of abnormal and normal signals, respectively, correctly identified by the classification procedure. Accuracy is defined as the percentage of subjects correctly identified. Our classification criterium was to assign each particular signal to the group with the maximum posterior probability, as defined in equation (3.2). Classification in Suárez (1997) was performed via linear discriminant functions (Fisher, 1936). Wavelet based Bayesian methods provide considerably higher specificity, sensibility and accuracy rates when discriminating HRECG signals of groups I and III. LPST indexes show a higher sensitivity rate in the discrimination of groups I and II than the wavelet based Bayesian methods (65.30% vs 60%), however, the specificity and accuracy rates for the wavelet based Bayesian methods are considerably higher.

5. Closing remarks and future developments

We fitted wavelet based Bayesian models to three groups of clinically different HRECG signals. The models are able to capture group features in terms of a very reduced number of significant wavelet coefficients. Differences among groups are evident in the number, level and shift location and posterior distributions of non-vanishing wavelet coefficients. Such differences translate into different “average” group patterns, characterising common aspects of the morphology of the signals in each group. Our Bayesian model based classification approach shows, in general, a much better predictive performance than classification schemes based on short-time frequency late potential indexes. In addition, the proposed model based method provides a probabilistic assessment of the uncertainties involved in classifying the signals.

In order to improve the predictive performance, further developments should include extensions to multivariate models that account for the information provided by the XYZ channels simultaneously. Also, different prior structures should be considered.

Acknowledgements

The HECG data analysis arose from discussions with Pedro Gomis of Departamento de Tecnología Industrial, Universidad Simón Bolívar, with whom interactions have been most valuable.

REFERENCES

- Brown, P. J., Fearn, T. and Vanucci, M. (2001). Bayesian wavelet regression curves with application to a spectroscopic calibration problem, *Journal of the American Statistical Association*, (to appear).
- Chipman, H., Kolackzyk, E. and McCulloch, R. (1997). Adaptive Bayesian wavelet shrinkage, *Journal of the American Statistical Association*, **92**, 1413–1421.
- Daubechies, I. (1993). *Ten Lectures on Wavelets*. CBMS-NSF Conference Series, SIAM, **61**, Philadelphia.

- Fisher, R. (1936). The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, **7**, 179–188.
- García, I. (1999). Bayesian modelling of ECG signals. Master's Thesis, Universidad Simón Bolívar (in spanish).
- García, I., Prado R. and Gomis P. (2000). Wavelet based Bayesian methods for characterizing high-resolution ECG signals, *Computers in Cardiology* (to appear), IEEE Computer Society Press.
- Gomis, P., Suárez, K., Passariello, G., Mendoza, I., Caminal, P. and Lander, P. (1996). Abnormal intra-QRS signals and late potentials in the high resolution ECG associated with chagasic myocarditis, *Computers in Cardiology*, **23**, 633–636, IEEE Computer Society Press.
- Meyer, Y. (1992) *Wavelets and Operators*. Cambridge Studies in Advance Mathematics, **37**, Cambridge University Press, New York.
- Morlet, D., Peyrin, F., Desseigne, P., Touboul, P. and Rubel, P. (1993). Wavelet analysis of high-resolution signal-averaged ECGs in postinfarction patients. *Journal of Electrocardiology*, **26**, 311–320.
- Müller, P. and Vidakovic, B. (1998). Bayesian inference with wavelets: Density estimation, *J. Comput. Graph. Statist.*, **25**, 456–468.
- Suárez, K. (1997). *Evaluation of Chagas disease through HRECG time-frequency representations*, Master's Thesis (in spanish), Universidad Simón Bolívar.
- Suárez, K., Gomis, P. and Silva, J. (1998). Analysis of late potentials associated with chagasic myocarditis by means of time-frequency representations. *Computers in Cardiology*, **27**, 397–400, IEEE Computer Society Press.
- Vidakovic, B. (1998). Nonlinear wavelet shrinkage with Bayes rules and Bayes factors. *Journal of the American Statistical Association*, **93**, 173–179.
- Vidakovic B. (1999). *Statistical Modeling by Wavelets*, Wiley Series in Probability and Statistics, New York, USA.
- Wavelab Reference Manual (1999). URL <http://www-stat.stanford.edu/~wavelab>.
- Zareba, W., Guindo, J., Madoery, C., Vinolas, X., Bayes de Luna, A. (1991). Late potentials diagnosed by FFT frequency analysis in Chagas patients with and without bundle branch block: comparison of different area ratio. *Electrocardiology '91*, 339–342, World Scientific Pub, New York.