



Universidad de Valladolid

Modelo para el análisis automático de latidos cardíacos. Implementación del algoritmo de estimación en aplicación de Shiny

Trabajo de Fin de Grado
Grado de Estadística

Autor: Christian Canedo Ortega
Tutor: Alejandro Rodríguez Collado y Cristina Rueda Sabater

29 de junio de 2021

Modelo para el análisis automático de latidos cardíacos. Implementación del algoritmo de estimación en aplicación de Shiny

Model for the automatic analysis of heartbeats. Implementation of the estimation algorithm in Shiny application

Christian Canedo Ortega

Resumen

La interpretación de los electrocardiogramas (ECG) juega un papel muy importante de cara a la identificación de enfermedades o trastornos cardiovasculares. Un análisis automático correcto es de gran ayuda en el diagnóstico y prevención de estas enfermedades.

El modelo Frequency Modulated Möbius (FMM), desarrollado recientemente por el grupo de investigación de Inferencia con Restricciones de la Universidad de Valladolid, es adecuado para el análisis de señales oscilatorias, entre las que se encuentra el ECG. En este trabajo se describe este modelo y la extensión desarrollada específicamente para el análisis de las señales ECG.

El objetivo principal del trabajo es la implementación de una aplicación en R Shiny que permite a un usuario general la estimación del modelo. Además, la aplicación proporciona dos marcadores nuevos, diseñados también por el grupo de investigación, que son las bases de reglas de diagnóstico para la detección de bloqueos de rama cardíacos.

Palabras clave: ECG, señal oscilatoria, FMM, R, aplicación Shiny, bloqueo de rama.

Abstract

The interpretation of electrocardiograms (ECG) plays a very important role in identifying cardiovascular diseases or disorders. A correct automatic analysis is very useful in the diagnosis and prevention of these diseases.

The Frequency Modulated Möbius model (FMM), recently developed by the Inferencia con Restricciones research group at the University of Valladolid, is suitable for the analysis of oscillatory signals, including the ECG. This work describes this model and the extension developed specifically for the ECG signals analysis.

The main goal of this work is the implementation of an application in R Shiny that allows a general user to estimate the model. In addition, the application provides two new markers, also designed by the research group, that define the diagnostic rules for the detection of cardiac bundle branch blocks.

Key words: ECG, oscillatory signal, FMM, R, Shiny app, bundle branch block.

Índice general

1. Introducción	1
1.1. Objetivos del trabajo.	1
2. El electrocardiograma (ECG).	3
3. Señales oscilatorias y modelos paramétricos. El modelo FMM.	9
3.1. Señales oscilatorias.	9
3.2. Modelos para señales oscilatorias	12
3.2.1. El modelo cosinor	12
3.2.2. El modelo de Fourier.	13
3.2.3. El modelo FMM.	14
4. Aplicación para el análisis de señales ECG.	25
4.1. Herramienta para la creación de la aplicación: R Shiny.	25
4.2. Descripción de la aplicación.	25
4.2.1. Input y preprocesado de datos.	26
4.2.2. Ajuste del modelo FMM_{ECG}	27
4.2.3. Salidas de la aplicación.	28
4.3. Ejemplos de uso de la aplicación.	29
4.3.1. Selección de datos.	29
4.3.2. ECG normal.	29
4.3.3. ECG con CBBB.	31
4.3.4. ECG con CLBBB.	32
5. Conclusión.	35
Bibliografía	37

1 Introducción

La interpretación de los electrocardiogramas (ECG) juega un papel muy importante de cara a la identificación de enfermedades o trastornos en el ámbito cardiovascular. Los ECG reflejan la actividad eléctrica del corazón, teniendo esta señal una particularidad que la convierte en un potencial objeto de estudio para el ámbito estadístico/matemático: se trata de una *señal oscilatoria*, aunque para nada sencilla.

Son muchas las propuestas de analizar los ECG y que utilizan una amplia variedad de técnicas, como las series temporales, modelos paramétricos y no paramétricos y sobretodo, las herramientas de análisis de grandes volúmenes de datos y *machine learning*.

Dentro de los modelos paramétricos, que no son nada usuales en este ámbito, se puede encontrar un modelo basado en la combinación de funciones gaussianas para modelar un único latido. Sin embargo, actualmente está en desuso dadas sus diversas limitaciones como la falta de flexibilidad o sencillez para describir ECGs que a menudo contienen ruido y presentan formas o componentes atípicas; la falta de interpretabilidad de los parámetros de los que depende este método o una eficacia computacional muy ineficiente.

Muy recientemente, en el grupo de investigación de Inferencia con Restricciones de la Universidad de Valladolid se ha desarrollado un nuevo procedimiento para el análisis de señales del ECG. El modelo FMM (Frequency Modulated Möbius) [1] es adecuado para modelizar sistemas de señales con patrones rítmicos o periódicos (señales circulares) y, en particular, el modelo permite descomponer las señales de los ECGs en cinco ondas fundamentales. Más adelante, en el capítulo donde se describe el FMM, se ahondará en las propiedades matemáticas de este modelo y sus ventajas.

El modelo FMM tiene un amplio rango de aplicaciones más allá de los ECGs. Algunos ejemplos son la neurociencia, caracterizando los registros de la actividad eléctrica de las neuronas; la cronobiología: ciclos circadianos y expresiones genéticas; o en la astrofísica, donde el patrón rítmico que describen las señales lumínicas de estrellas lejanas sirve para su clasificación.

1.1 Objetivos del trabajo.

Se han llevado a cabo los siguientes objetivos:

- Introducción de la señal del ECG, sus características y morfología, diferenciando casos típicos y casos patológicos.
- Estudio de señales oscilatorias. Funciones y modelos estadísticos para su modelización.
- Descripción y formulación del modelo FMM unicomponente y multicomponente. Propiedades matemáticas y procedimientos de inferencia. Descripción del modelo FMM específico para el análisis de señales ECG: FMM_{ECG} .
- Desarrollo de un aplicación en *R Shiny* que sirve de interfaz para analizar ECGs con el modelo y acercar el uso del FMM_{ECG} a un usuario general.

2 El electrocardiograma (ECG).

En este capítulo describiremos el ECG, el que es hasta la fecha el único método no invasivo que recoge de la actividad eléctrica del corazón. Un ECG permite la detección de alteraciones anatómicas del corazón, deficiencias rítmicas en el mismo y multitud de patologías cardíacas. A día de hoy, no caben dudas acerca de la importancia de los ECGs en el ámbito del diagnóstico médico.

La actividad eléctrica del corazón se origina en un conjunto de células localizado en la rama derecha del corazón denominado nódulo sinoauricular, lo que podríamos llamar el “marcapasos” del corazón. Este transmite su actividad por los distintos circuitos del corazón (figura 2.1) provocando su despolarización, lo que induce la contracción de las fibras que abrazan las cavidades del órgano: las aurículas y los ventrículos. Una vez se ha llevado a cabo la contracción completa del corazón, las células de miocardio (el tejido muscular del corazón) se relajan volviendo de nuevo a un estado polarizado.

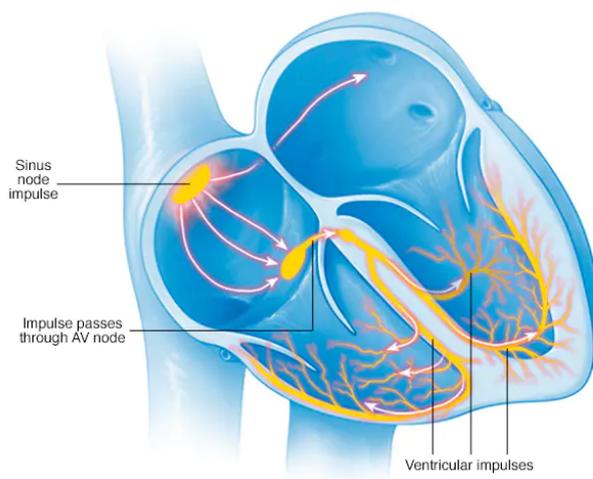


Figura 2.1: Recorrido de la actividad eléctrica del corazón.

El ECG es el registro que se toma de ese campo eléctrico producido por las células cardíacas que se despolarizan y repolarizan de manera rítmica o cíclica. Se lleva a cabo la anotación (analógica o digitalmente) de los distintos potenciales eléctricos mediante un dispositivo formado por varios electrodos durante un cierto período de tiempo, lo que genera las doce señales típicas que se ven en un ECG (figura 2.2).

La incidencia de la propagación de esta corriente en un electrodo es distinta dependiendo de dónde se sitúe el electrodo en base al origen de la actividad eléctrica (el propio corazón) y la dirección del desplazamiento de la corriente, por eso las doce señales tienen patrones distintos. La derivación más utilizada y analizada es la derivación II, porque se considera que registra de forma más completa la señal eléctrica. Un ejemplo de ECG normal se representa en la figura 2.2.

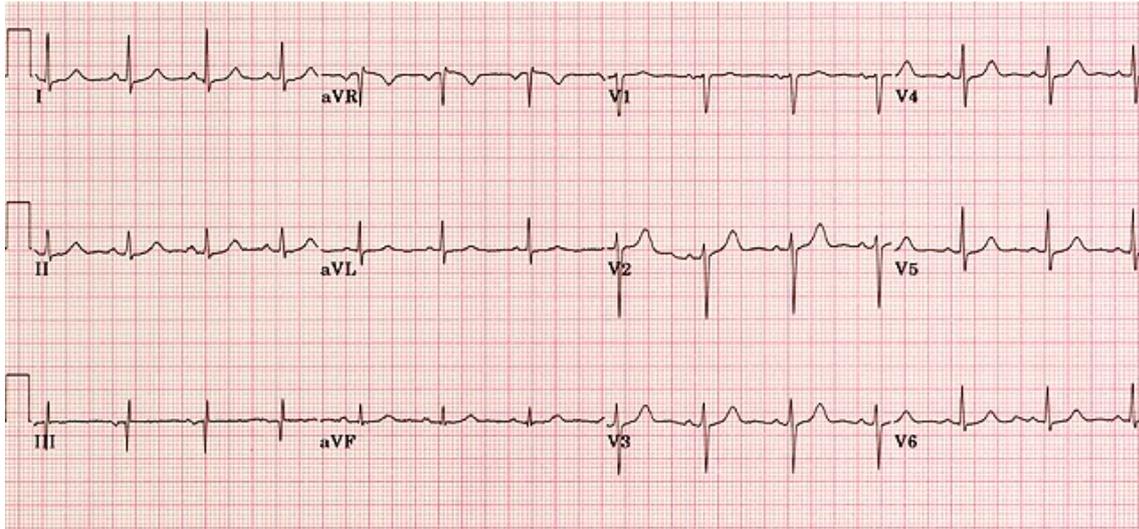


Figura 2.2: Ejemplo de ECG normal.

Cada una de las doce señales representa el funcionamiento eléctrico de diferentes partes del corazón. Son útiles para detectar alteraciones de la actividad normal que pueden ser signo de múltiples patologías. El registro completo consta de varios latidos consecutivos recogidos en un período variable. Además, es usual un primer instante de calibración que no guarda relación con la actividad cardiovascular como se ve en el inicio de las señales en la figura 2.2.

Un análisis adecuado de estas señales supone por lo tanto un gran paso en el diagnóstico de dichas patologías, en particular la descripción de la morfología específica de cada derivación, que ha sido ampliamente estudiada por cardiólogos e investigadores en la materia. Cada elemento estructural del ECG corresponde a la detección de la corriente eléctrica cuando esta se propaga por cada una de las partes del sistema conductor cardiaco.

En la figura 2.3 se representa el perfil típico de un latido de la derivación II donde se identifican las ondas que lo componen y algunos otros elementos. El latido comienza con la onda P, que se corresponde con la despolarización de las aurículas; le sigue el complejo QRS, que es el reflejo de la despolarización ventricular; terminando en la onda T, originada por la repolarización de los ventrículos. El complejo QRS tiene como componente principal la onda R, que está precedida por la onda Q y que es anterior que la onda S -ambas negativas-. Estas últimas no se detectan tan fácilmente como el resto y muchas veces no se aprecian debido a problemas en la propia medición, bien sea por falta de resolución o de exceso de ruido en la señal.

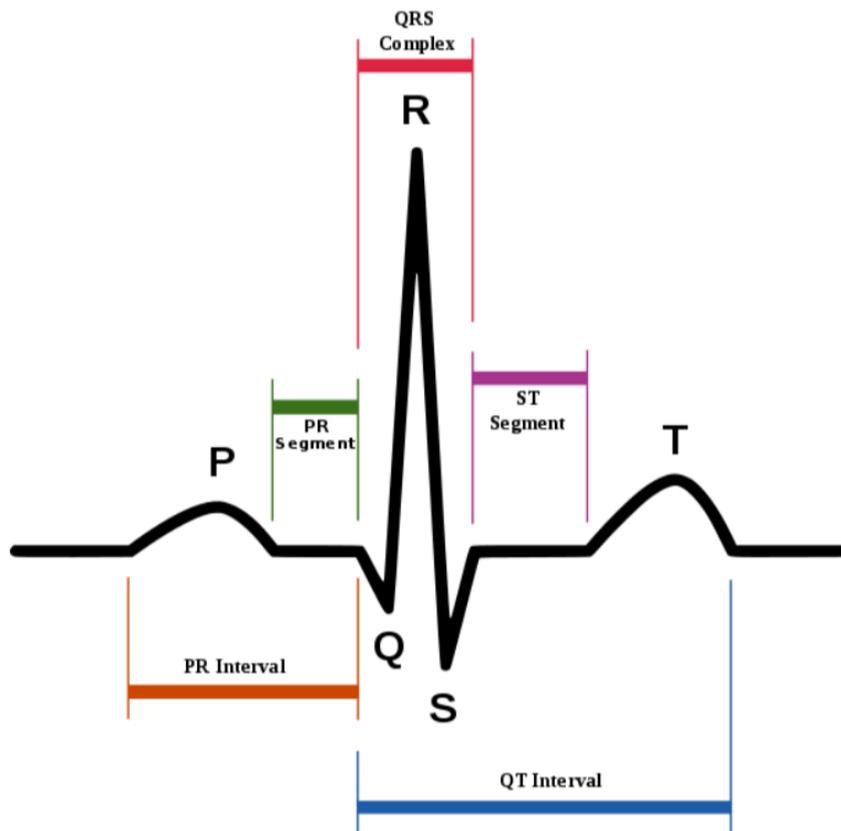


Figura 2.3: Esquema de latido típico de la derivación II

Existen otras medidas importantes y útiles en el diagnóstico al respecto de las ondas mencionadas. En 2.3 se ven los intervalos como es el periodo de estimulación auricular (intervalo PR), y ventricular (intervalo QT). En el análisis de un segmento formado por varios latidos consecutivos, la distancia entre los picos de las ondas R en latidos contiguos es una medida muy importante que describe el ritmo, ya que la onda R es la que mejor se identifica en un ECG.

El latido de la figura 2.3 corresponde a un individuo sano. La morfología de las ondas cambian de forma importante si hay una patología presente en el paciente.

A continuación se va a ilustrar con un ejemplo las alteraciones en la morfología del ECG. La figura 2.5 corresponde al ECG tomado a un paciente varón de 55 años, que manifestó tener opresión torácica, sensación de mareos y palpitaciones cardiacas (Caso: <https://urgenciasbidaso.wordpress.com/2020/04/27/caso-300-taquicardia-supraventricular-paroxistica-recidivante/>). Una vez analizado el ECG, al paciente se le diagnosticó una taquicardia supraventricular paroxística, esto es, una taquicardia rítmica (se aprecia en la derivación II la regularidad de los latidos) y rápida. En este paciente en concreto se han registrado 200 latidos por minuto en reposo, una tasa muy por encima de lo normal.

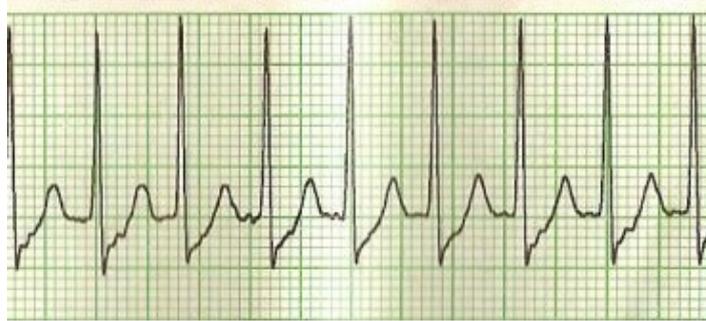


Figura 2.4: Latidos de la derivación II. Taquicardia supraventricular.

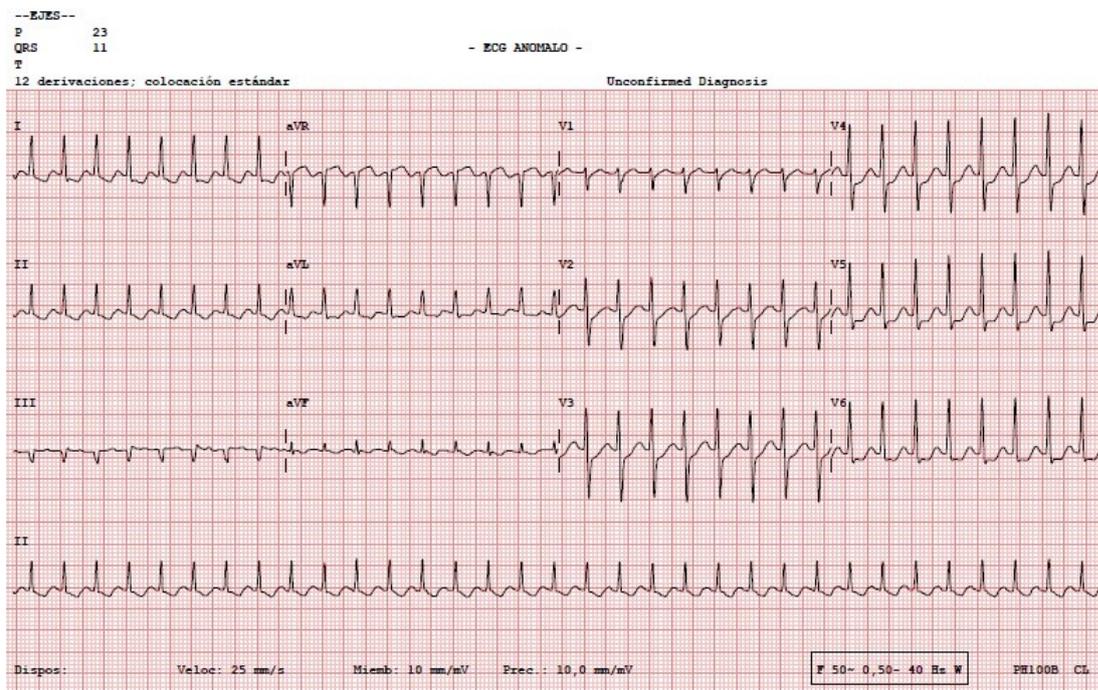


Figura 2.5: ECG de paciente con taquicardia supraventricular paroxística.

El origen de esta patología es el bloqueo de uno de los "motores" principales de la actividad cardíaca: el nodo AV, ver 2.1. Si comparamos detenidamente las figuras 2.3 y 2.4, concretamente las ondas S y T, se aprecia que debido al bloqueo del nodo AV se origina un descenso del segmento ST.

Otro caso importante, ya se refiere a una de las enfermedades más usuales y conocidas en este ámbito, es el infarto agudo de miocardio (IAM) [2]. En algunos pacientes se produce una oclusión de las vías cardiovasculares, provocada por varias posibles causas, como la acumulación de placas de colesterol o de células inflamatorias. Esto puede producir daños irreversibles en los tejidos y estructuras del corazón debido a la falta de riego sanguíneo, provocando la muerte de las células del miocardio de la zona afectada, lo que puede llevar al paciente a sufrir IAM.

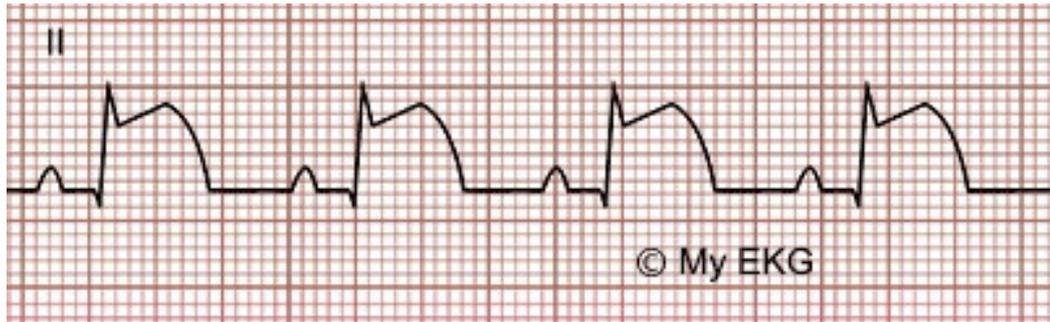


Figura 2.6: Latidos de la derivación II. Infarto agudo de miocardio (IAM).

En el caso de un ECG asociado al IAM se suele producir una elevación del segmento ST (al contrario que en la taquicardia), derivando en un latido anómalo que se repite a lo largo de todo el registro. Un caso de IAM reportado en: <https://urgenciasbidasoa.wordpress.com/2015/11/25/caso-202-elevacion-de-st-en-ecg-a-proposito-de-dos-casos/>, muestra un ECG con el patrón mencionado (fig. 2.7, derivación II).

En ECGs como el 2.5, la alteración de la morfología del latido se puede apreciar prácticamente en todas las derivaciones a simple vista. Sin embargo, en ECGs como el del IAM, mientras que en la derivación II se ve claramente una diferencia con el ECG normal (la mencionada elevación del ST), en otras derivaciones, como la I, sería muy complicado hacer un diagnóstico preciso del infarto. Esto no significa que la derivación no refleje el IAM, si no que los indicios, en caso de haberlos, son muy sutiles.

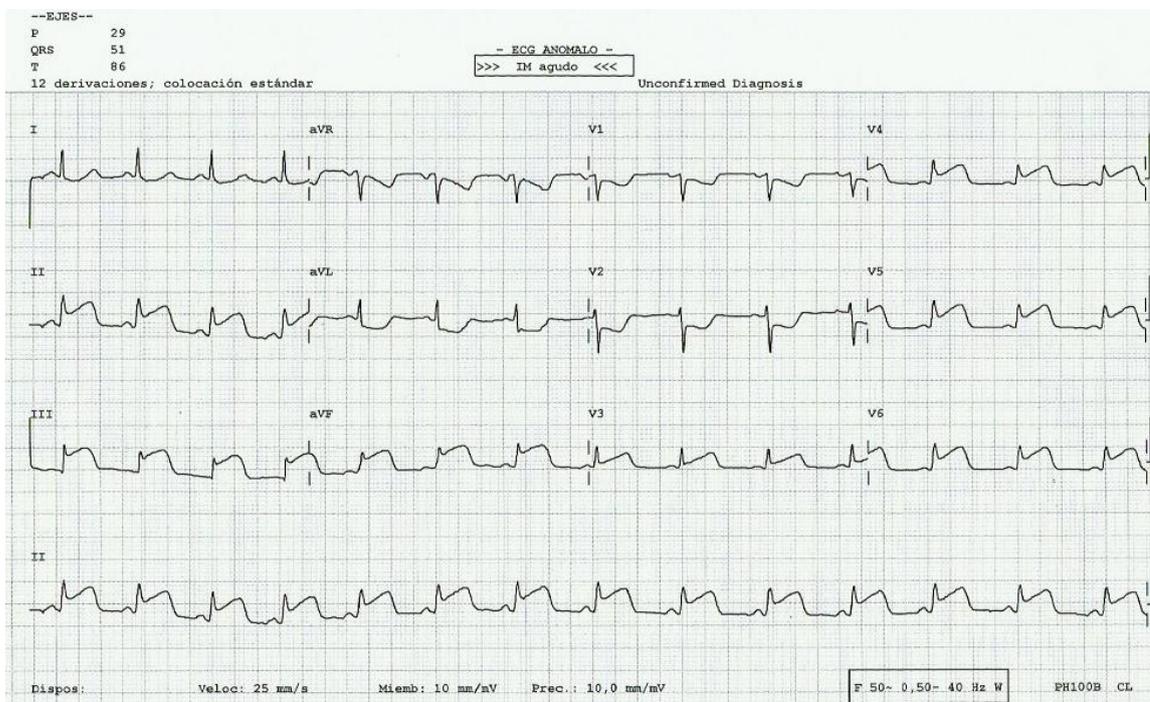


Figura 2.7: ECG de paciente con infarto agudo de miocardio (IAM).

El infarto agudo de miocardio y la taquicardia supraventricular son dos ejemplos de la gran variedad de patologías cardiovasculares que existen. El ECG es una herramienta imprescindible en la detección e identificación de estas enfermedades. El reconocimiento de una enfermedad concreta mediante su ECG asociado pasa, principalmente, por el análisis morfológico del mismo. También son importantes otras características como la tasa de latidos por minuto de un paciente.

3 Señales oscilatorias y modelos paramétricos. El modelo FMM.

Habiendo entendido la complejidad estructural del ECG, en este capítulo se describen las señales oscilatorias y las herramientas matemáticas y estadísticas usadas para su modelización. Dentro de estas últimas se incluye el modelo FMM y sus extensiones, en particular el modelo FMM_{ECG}

3.1 Señales oscilatorias.

Se ha hecho mención de una característica fundamental de los electrocardiogramas: son señales oscilatorias que fluctúan una o más veces en torno a un nivel basal, que en el caso del ECG suelen ser los $-90mV$ y que se repiten en el tiempo.

Las señales se pueden registrar en cualquier intervalo temporal, que en concreto en los ECGs suelen ser $80ms$ por latido aproximadamente, pero matemáticamente se trabaja con un intervalos común, $t \in [0, 2\pi]$ mediante la transformación lineal correspondiente. En este período de tiempo, la señal puede oscilar una o varias veces, llegando a describir desde señales simples, como la de la Figura 3.1, a señales oscilatorias más complejas, como el ECG. Una señal que oscila una única vez es lo que se denomina señal circular (3.1).

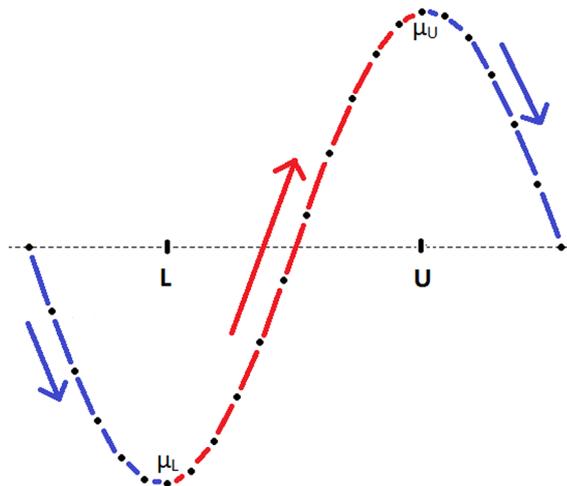


Figura 3.1: Esquema de una señal circular.

Estas señales pueden ser descritas, por ejemplo, de una manera no paramétrica mediante restricciones up-down-up.

Definición 3.1 Una señal real $\mu(t) \in \mathbb{R}$, $t \in [0, 2\pi]$, es circular si:

$$\begin{aligned} \exists t_U, t_L / \quad & a) \mu(t) \geq \mu(t'), t_L \leq t \leq t' \leq t_U, \\ & b) \mu(t) \geq \mu(t'), 0 \leq t \leq t' \leq t_L; t_U \leq t \leq t' \leq 2\pi \end{aligned}$$

bajo el supuesto de que $t_U \geq t_L$. Existe una definición análoga para el caso en el que $t_U \leq t_L$.

Esta definición corresponde a una señal en el espacio euclídeo, existiendo su equivalente en el círculo unidad, como sigue:

$$\phi(t) \leq \phi(t'), 0 \leq t \leq t' \leq 2\pi$$

Pueden definirse diferentes señales paramétricas que cumplen las propiedades anteriores, la señal más sencilla y probablemente más usada es la señal sinusoidal $\mu(t) = \cos(t + \varphi)$, donde la correspondiente señal circular es una función lineal $\phi(t) = t + \varphi$ que es lo que se ha representado en la señal de la Figura 3.1.

Una disciplina en la que encontramos muchos casos de señales oscilatorias es la cronobiología. Esta disciplina estudia los patrones rítmicos de los eventos fisiológicos y biológicos y cómo se ven afectados por condiciones externas. Concretamente, las expresiones de los genes involucrados en los ritmos circadianos describen una señal oscilatoria, tal y como se puede observar en la Figura 3.2 del gen IQGAP2 (NCBI GEO: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11923>). La cronobiología ha tenido un gran impacto en el tratamiento de enfermedades como la hipertensión, en la detección de genes asociados a enfermedades neurodegenerativas o a la depresión, y en la mejora de la efectividad de tratamientos contra el cancer [3].

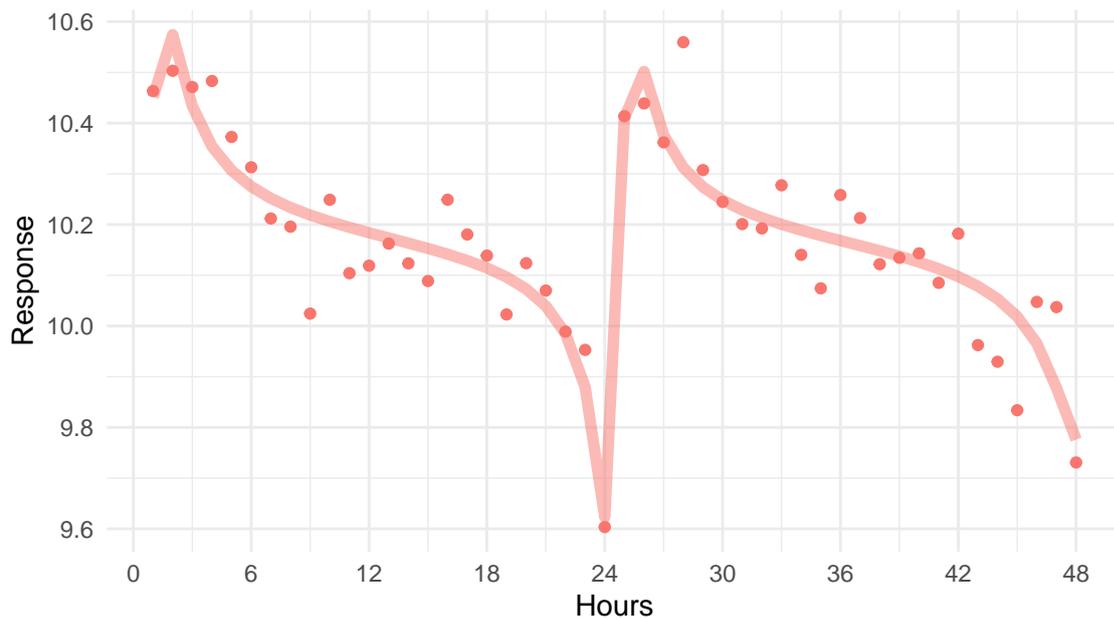


Figura 3.2: Expresión del gen IQGAP2 en ratones.

Volviendo a las matemáticas, el tipo de señales a las que se ha hecho referencia hasta el momento son señales reales, sin embargo, es habitual trabajar con señales que toman valores complejos. Es el caso de las señales analíticas:

Definición 3.2 Señal analítica.

$$S(t) = \mu(t) + i\nu(t) = \rho(t)e^{i\varphi(t)}, t \in [0, 2\pi]$$

$$\rho(t) = \sqrt{\mu(t)^2 + \nu(t)^2}$$

$$\varphi(t) = \arctan\left(\frac{\nu(t)}{\mu(t)}\right)$$

donde $\rho(t)$ y $\varphi(t)$ son habitualmente denominados amplitud instantánea y fase instantánea, respectivamente.

En particular, se define la señal real asociada a $S(t)$ como:

$$\text{Re}(S(t)) = \rho(t)\cos(\phi(t)), \quad t \in [0, 2\pi] \quad (3.1)$$

Una simplificación interesante usada en la práctica, que en particular verifican los modelos que veremos posteriormente, es que $\rho(t)$ sea constante. En estos casos $\phi(t)$, que es la fase, juega un papel fundamental en la configuración de la señal.

Un caso particular de señal compleja y analítica que merece la pena mencionar por su influencia en el desarrollo del modelo FMM del que hablaremos después es el átomo de Fourier, que se define como:

Definición 3.3 *Átomo de Fourier:*

$$S(t) = e^{i\phi_a(t)} = \tau_a(e^{it}), \quad t \in (-\infty, \infty), \quad a \in \mathbb{C}, |a| < 1,$$

donde $\tau_a(z)$ se conoce como la transformación de Möbius:

$$\tau_a(z) = \frac{z - a}{1 - \bar{a}z}, \quad |a| < 1$$

bajo la condición de que $\tau_a(a) = 0$.

Se muestran en los gráficos de la Figura 3.3 varios ejemplos de la parte real de átomos de Fourier para distintos valores del complejo a .

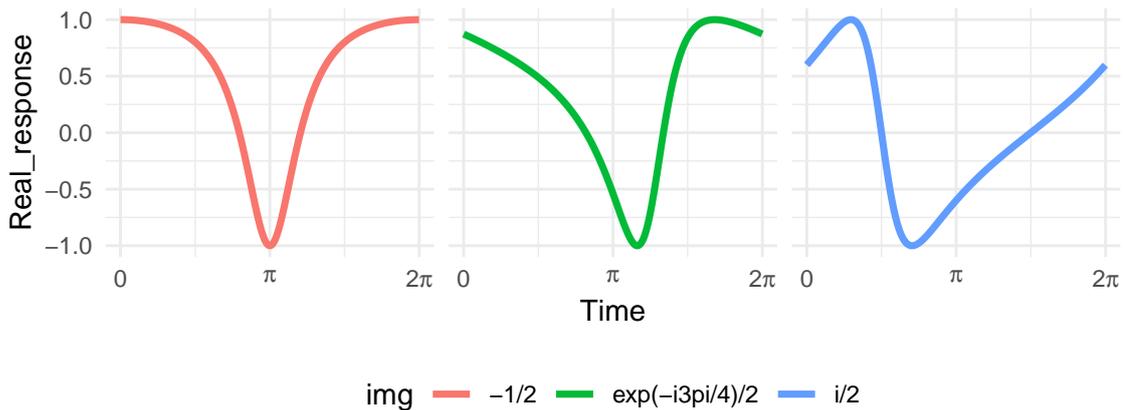


Figura 3.3: Ejemplos de átomos de Fourier, parte real.
(Izda.: $a = -1/2$, Med.: $a = \frac{1}{2}e^{-i\frac{3\pi}{4}}$, Dcha.: $a = i/2$)

Llama la atención la capacidad para describir señales con una única oscilación asimétricas, la cual es una característica notable que proporciona la transformación de Möbius y que es germen del modelo FMM. En concreto, alguno de los patrones de 3.3 se asimila a los patrones que describe la señal de la Figura 3.2.

3.2 Modelos para señales oscilatorias

En esta sección se introducen los modelos estadísticos más extendidos en su uso para el análisis de señales oscilatorias: Cosinor, Fourier y el FMM.

Se considera $X(t_i), t_1 < \dots < t_n$, el vector de valores reales donde se definen la señal original, siendo $X_i = X(t_i); i = 1, \dots, n$. Estos modelos se componen de una parte correspondiente a la señal ($\mu_i(t_i)$) y otra al ruido ($\epsilon(t_i)$), $i = 1, \dots, n$. El ruido se considera aleatorio e independiente de la señal, como es habitual.

3.2.1 El modelo cosinor

El modelo más sencillo que ajusta una función oscilatoria a una señal como las mencionadas con anterioridad es el Cosinor. En este modelo, la señal fluctúa en torno a un parámetro M . Su formulación, incluyendo el término de señal y de error:

Definición 3.4 Modelo Cosinor:

$$X(t_i) = \mu(t_i) + \epsilon(t_i) = M + A\cos(t_i + \varphi) + \epsilon(t_i); \quad i = 1, \dots, n$$

$$t_i, \varphi \in [0, 2\pi]$$

$$(\epsilon(t_1), \dots, \epsilon(t_n))' \sim N_n(0, \sigma^2 I_n)$$

Este es uno de los modelos más usados, en particular, en cronobiología[3]. En esta disciplina es usual denominar al parámetro M como *MESOR (Midline Estimating Statistic of Rhythm)*.

Una limitación importante del mismo es que describe únicamente señales simétricas (sinusoidales). Esto en la práctica, por lo general, no se verifica. Esta simplicidad se explica porque, en realidad, todos sus parámetros se podrían considerar parámetros de localización (M y ϕ) y de escala (A) sobre la función coseno en el plano de coordenadas cartesianas.

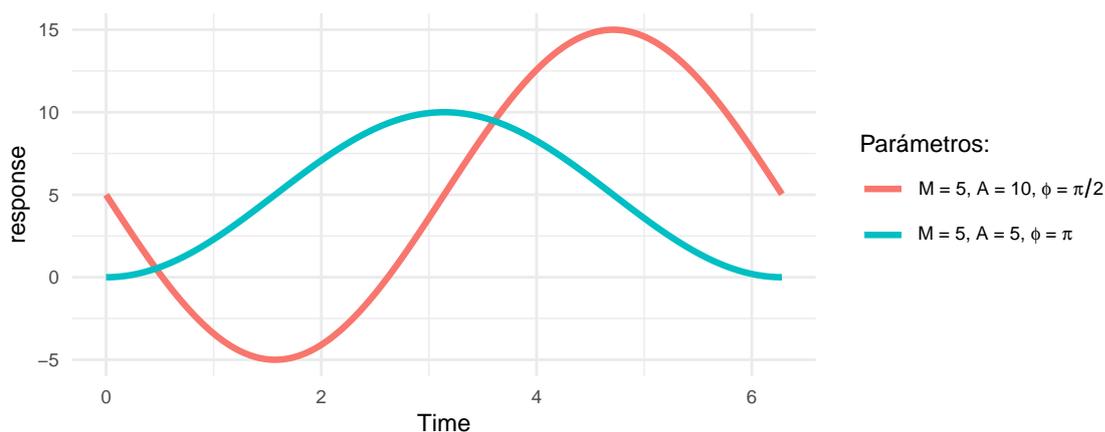


Figura 3.4: Ejemplos de modelo cosinor.

Se puede observar en la Figura 3.4 el término $\mu(t)$ de dos ejemplos de modelo cosinor con el mismo parámetro M y con distintos parámetros de amplitud (A) y fase (φ).

3.2.2 El modelo de Fourier.

El modelo FD (Fourier Decomposition) es capaz de describir formas algo más complejas que el modelo cosinor. Este es un modelo aditivo basado en la descomposición de Fourier, donde sus componentes individuales, también llamados armónicos, son funciones sinusoidales sencillas. Además de la señal, también se define un término de ruido o error. Un ejemplo es el FD^2 , el cual está formado únicamente por dos armónicos:

Definición 3.5 Modelo FD^2 :

$$X(t_i) = \mu(t_i) + \epsilon(t_i) = A_0 + A_1 \cos(t_i) + B_1 \sin(t_i) + A_2 \cos(2t_i) + B_2 \sin(2t_i) + \epsilon(t_i);$$

$$i = 1, \dots, n$$

En la Figura 3.5 se muestran algunos ejemplos de modelos FD^2 con $A_0 = 0$ para todos ellos.

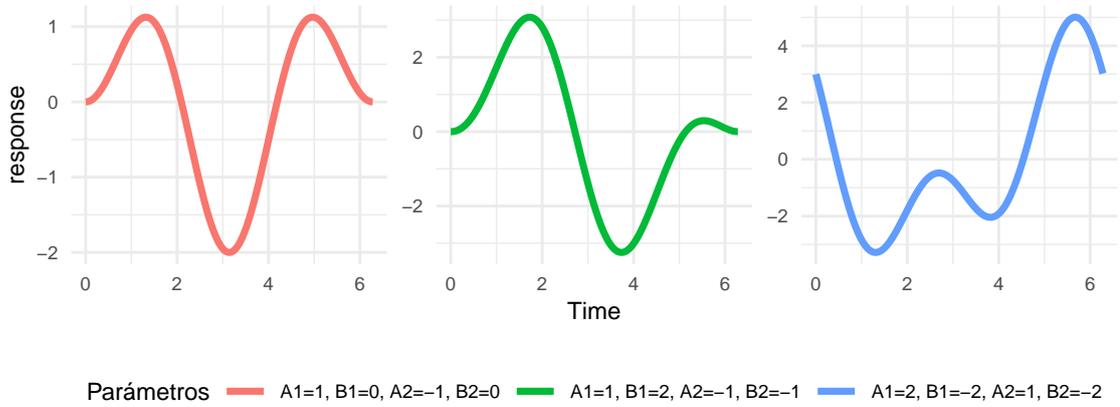


Figura 3.5: Ejemplos de modelo FD^2 .

Sin duda, este modelo es más flexible que el modelo cosinor y puede adaptarse mejor a un mayor número de configuraciones en los datos. Es inmediato ver que el modelo cosinor es un caso particular del FD. Sin embargo, mientras que el modelo cosinor tenía únicamente dos parámetros (amplitud y fase inicial), el modelo FD^2 tiene cinco parámetros: a_0 y dos coeficientes para el seno y el coseno de cada armónico. El aumento del número de armónicos nos lleva a modelos propensos a un efecto de sobreajuste, además de un coste computacional elevado en su ajuste.

Adecuación de los modelos.

Como medida de la bondad del ajuste se da la proporción de varianza explicada por el modelo sobre la variabilidad original de la señal:

$$R^2 = 1 - \frac{\sum_{i=1}^n (X(t_i) - \hat{\mu}(t_i, \theta))^2}{\sum_{i=1}^n (X(t_i) - \bar{X})^2} \quad (3.2)$$

donde $\hat{\mu}(t_i, \theta)$, $i = 1, \dots, n$ hace referencia al valor predicho o ajustado en el instante t_i y θ a los parámetros del modelo en su espacio paramétrico correspondiente.

3.2.3 El modelo FMM.

Este modelo se ha desarrollado muy recientemente en los artículos [1], [4], [5], [6], motivado por señales como la de la Figura 3.2 la cual no es correcto analizar con los modelos habituales. El modelo más sencillo se define para analizar señales con una única oscilación, y el modelo FMM_m se desarrolla posteriormente para analizar señales más complejas donde hay más oscilaciones.

FMM monocomponente.

El modelo FMM₁, cuya señal se compone de una única onda $W(t, v) = A \cos(\phi(t; \alpha, \beta, \omega))$, es capaz de describir patrones no sinusoidales, a diferencia de modelos como el cosinor. Esto es debido a la inclusión de la función de enlace de Möbius^{[7], [8]} como fase $\phi(t; \alpha, \beta, \omega)$, frente a la fase del cosinor que es simplemente una función lineal.

La formulación del modelo FMM de una única componente, que también es del tipo señal más ruido, es el siguiente:

Definición 3.6 Modelo FMM₁

$$X(t_i) = \mu(t_i) + \epsilon(t_i) = M + A \cos(\phi(t_i; \alpha, \beta, \omega)) + \epsilon(t_i), \quad i = 1, \dots, n,$$

$$(\epsilon(t_1), \dots, \epsilon(t_n))' \sim N_n(0, \sigma^2 I_n)$$

donde $M \in \mathbb{R}$, $A \in \mathbb{R}^+$ y $\phi(t; \alpha, \beta, \omega) = \beta + 2 \arctan(\omega \tan(\frac{t-\alpha}{2}))$; $\alpha, \beta \in [0, 2\pi)$, $\omega \in [0, 1]$.

Descripción de los parámetros

Respecto a la interpretación de los parámetros. M y A son parámetros de intercept y escala, respectivamente. Los parámetros α , β y ω definen la fase del modelo y que se explican detenidamente a continuación.

En primer lugar, α es un parámetro de localización de la fase. En 3.6 se muestran cuatro señales FMM estándar ($M = 0$ y amplitud unidad) con distintos valores para α manteniendo fijos el resto de parámetros.

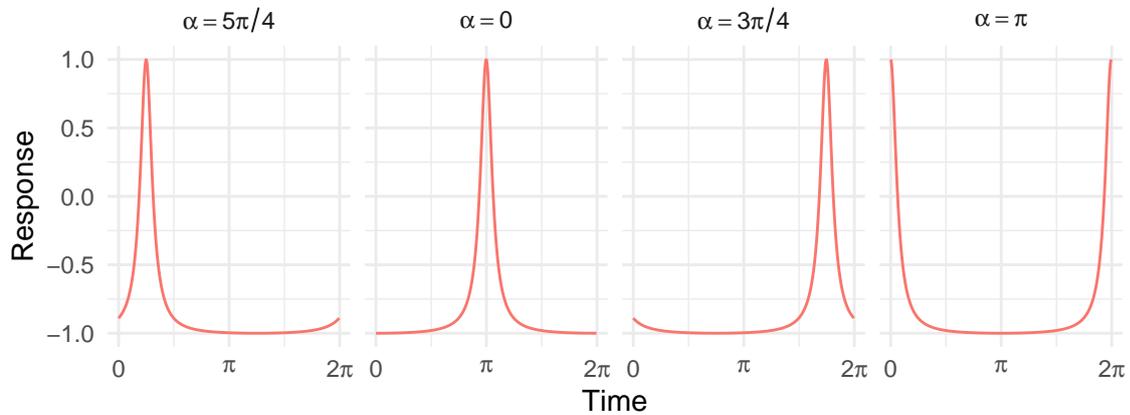


Figura 3.6: Modelos FMM. $M = 0$, $A = 1$, $\omega = 0.1$, $\beta = \pi$.

Por otro lado, ω y β son parámetros de forma. El parámetro β describe la asimetría de la onda y ω el apuntamiento. En la Figura 3.7 se ilustran algunos patrones asociados a valores específicos de los parámetros.

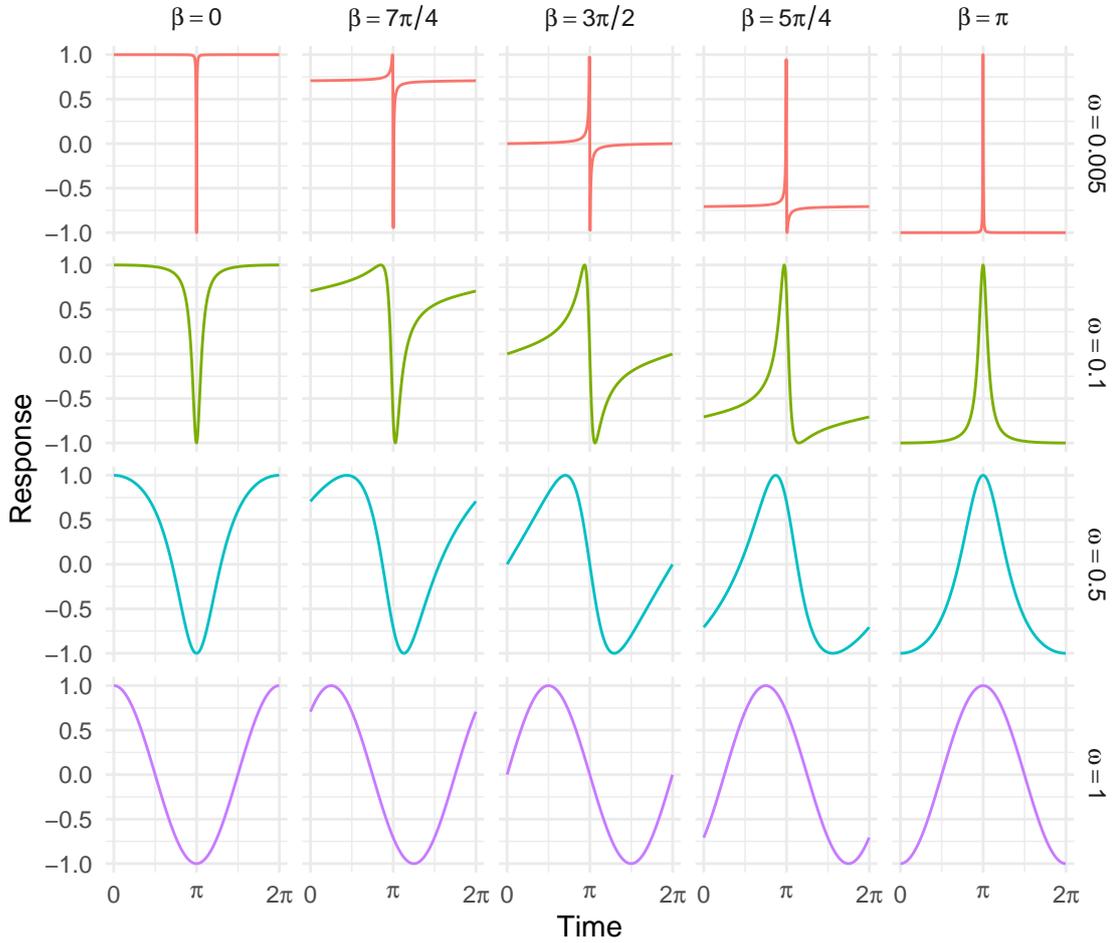


Figura 3.7: Modelos FMM. $M = 0$, $A = 1$, $\alpha = 0$.

La onda es totalmente simétrica cuando $\beta = 0$ y $\beta = \pi$, mientras que en sus valores intermedios se dan formas asimétricas. Otro aspecto remarcable de β es que dos ondas con parámetros β_1 y β_2 describen patrones inversos si verifican que $\beta_2 = \beta_1 + \pi$. Es inmediato comprobar esto dada la relación $\cos(x + \pi) = -\cos(x)$ (permaneciendo fijos el resto de parámetros). Es por ello que en la Figura 3.7 se omiten valores de β en el intervalo $[0, \pi]$. Aún así, en la figura se muestran los casos con $\beta = 0$ y $\beta = \pi$ en los que se puede apreciar que se describen ondas inversas.

ω mide el apuntamiento en la onda FMM. Aquellas ondas con un valor de ω cercano a 0 tienen un apuntamiento más pronunciado, mientras que valores de ω próximos a 1 describen formas más suaves, similares a ondas sinusoidales. En concreto, cuando $\omega = 1$, el FMM coincide con el modelo Cosinor (Ecuación 3.4) donde $\varphi = \beta - \alpha$, lo que se ve de manera directa:

$$\phi(t; u) = \beta + 2 \arctan(\omega \tan(\frac{t - \alpha}{2})) = \beta + 2 \arctan(\tan(\frac{t - \alpha}{2})) = t + \beta - \alpha$$

Además de los parámetros que podemos denominar como canónicos o básicos, se definen dos parámetros importantes por su interés en la práctica, que son los tiempos donde la onda alcanza su mínimo (t_L) y su máximo (t_U). Los valores que toma la onda en esos instantes, Z_U y Z_L para los valores máximo y mínimo. Estos parámetros tienen la siguiente expresión:

$$t_U = \alpha + 2\arctan\left(\frac{1}{\omega}\tan\left(\frac{-\beta}{2}\right)\right) \quad (3.3)$$

$$t_L = \alpha + 2\arctan\left(\frac{1}{\omega}\tan\left(\frac{\pi - \beta}{2}\right)\right)$$

$$Z_U = X(t_U) = M + A \quad (3.4)$$

$$Z_L = X(t_L) = M - A$$

Estos parámetros juegan un papel importante en la interpretación de las ondas FMM. Por ejemplo, en el caso de los genes, t_U corresponde con el momento del día en el que el gen se activa.

Estimación de los parámetros del modelo.

El problema de calcular el estimador máximo verosímil, en este caso, se reduce a un problema de mínimos cuadrados:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{i=1}^n (X(t_i) - \mu(t_i, \theta))^2 \quad (3.5)$$

donde Θ es el espacio paramétrico para $\theta = (M, A, \alpha, \beta, \omega)$, $\Theta = \mathbb{R} \times \mathbb{R}^+ \times [0, 2\pi) \times [0, 2\pi) \times [0, 1]$.

Para la estimación es conveniente una reformulación del modelo, como sigue:

$$\mu(t_i, \theta) = M + A\cos(t_i^* + \varphi), i = 1, \dots, n \quad (3.6)$$

donde $t_i^* = \alpha + \arctan(\omega \tan(\frac{t_i - \alpha}{2}))$ y $\varphi = \beta - \alpha$.

Mediante la identidad trigonométrica de la suma de ángulos se llega a una nueva expresión del modelo en forma lineal:

$$\mu(t_i, \theta) = M + \delta z_i + \gamma \omega_i, i = 1, \dots, n \quad (3.7)$$

donde $\delta = A\cos(\varphi)$, $\gamma = -A\sin(\varphi)$, $z_i = \cos(t_i^*)$, $\omega = \sin(t_i^*)$, para $i = 1, \dots, n$.

Para valores fijos y conocidos de α y ω , lo que supone valores conocidos de t_i^* , la estimación de los parámetros se reduce a un problema de mínimos cuadrados ordinarios. De esta forma, la estimación del resto de parámetros (M , A y β) queda como sigue:

$$\hat{M} = \bar{X} - \hat{\delta} \sum_{i=1}^n z_i - \hat{\gamma} \sum_{i=1}^n \omega_i \quad (3.8)$$

$$\hat{A} = \sqrt{\hat{\delta}^2 + \hat{\gamma}^2} \quad (3.9)$$

$$\hat{\beta} = \alpha + \varphi \quad (3.10)$$

Teniendo en cuenta el desarrollo anterior, el algoritmo de estimación de los parámetros del modelo se describe a continuación y se lleva a cabo en dos pasos:

1. Se define un grid de (α, ω) y para cada punto se obtienen los estimadores definidos en 3.8, 3.9 y 3.10. De las distintas soluciones se escoge la combinación de parámetros del grid que minimiza la suma de cuadrados residual.
2. En el segundo paso se usa método de optimización de Nelder-Mead para afinar la solución encontrada en el paso 1. Nelder-mead es un procedimiento para la búsqueda de un óptimo local partiendo de una solución inicial. El algoritmo consiste en evolucionar un sistema de puntos en n dimensiones (un simplex o “triángulo multi-dimensional”) mediante transformaciones en los mismos hasta llegar a una solución que el procedimiento no es capaz de mejorar después de un cierto número de transformaciones sin cambiar la solución óptima.

El problema 3.5 es computacionalmente intensivo. Para una rejilla de 48 valores de α y 24 valores de ω se resuelven un total de 1152 sistemas lineales y el posterior procedimiento de Nelder-Mead.

Para ilustrar un ejemplo de la estimación de los parámetros, en la Figura 3.8 se muestra el ajuste del modelo FMM a las primeras 24 horas de los datos de la expresión génica del gen IQGAP2, que se había mostrado con anterioridad en la Figura 3.2. La combinación de parámetros que minimiza la suma de cuadrados residual se detalla en la Tabla 3.1.

	\hat{M}	\hat{A}	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\omega}$	R^2
IQGAP2 gene	10.102	0.560	3.084	1.434	0.066	0.865

Tabla 3.1: Parámetros FMM₁ estimados en datos de expresión del gen IQGAP2 (24 horas).

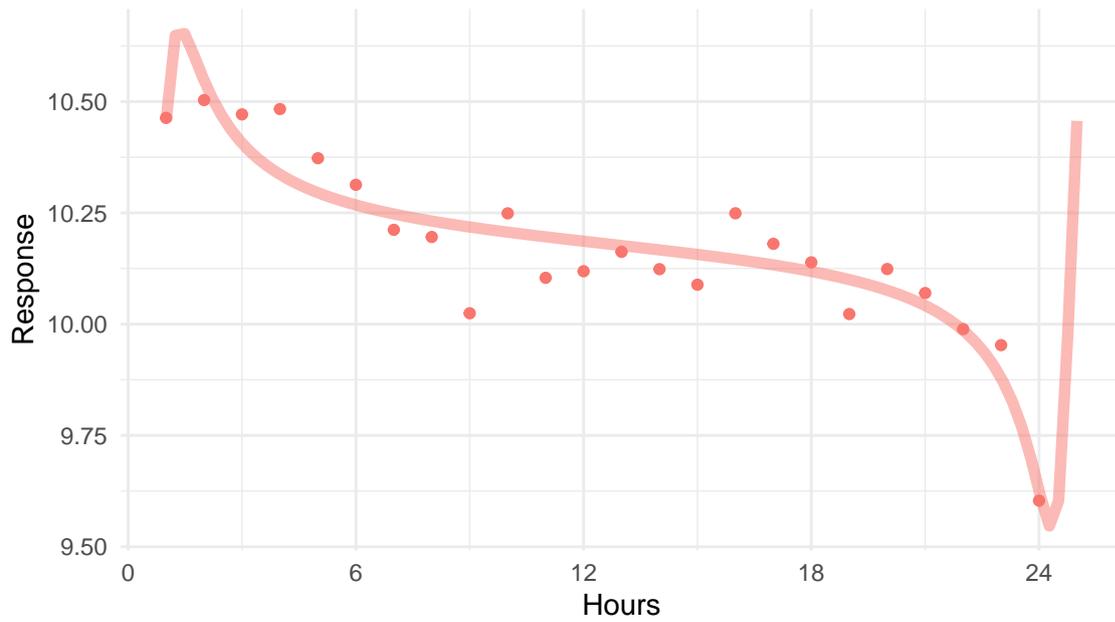


Figura 3.8: Ajuste del FMM_1 a expresión del gen IQGAP2 en ratones (24 horas).

La Figura 3.9 muestra la búsqueda de la solución óptima en un grid de (α, ω) de tamaño 200×200 para este mismo caso. Es una señal con una única oscilación y poco ruido, por lo que el algoritmo de estimación encuentra sin problema los parámetros óptimos. Sin embargo, un mayor ruido en la señal se traduce en un aumento de mínimos locales que dificultan la búsqueda de la solución óptima.

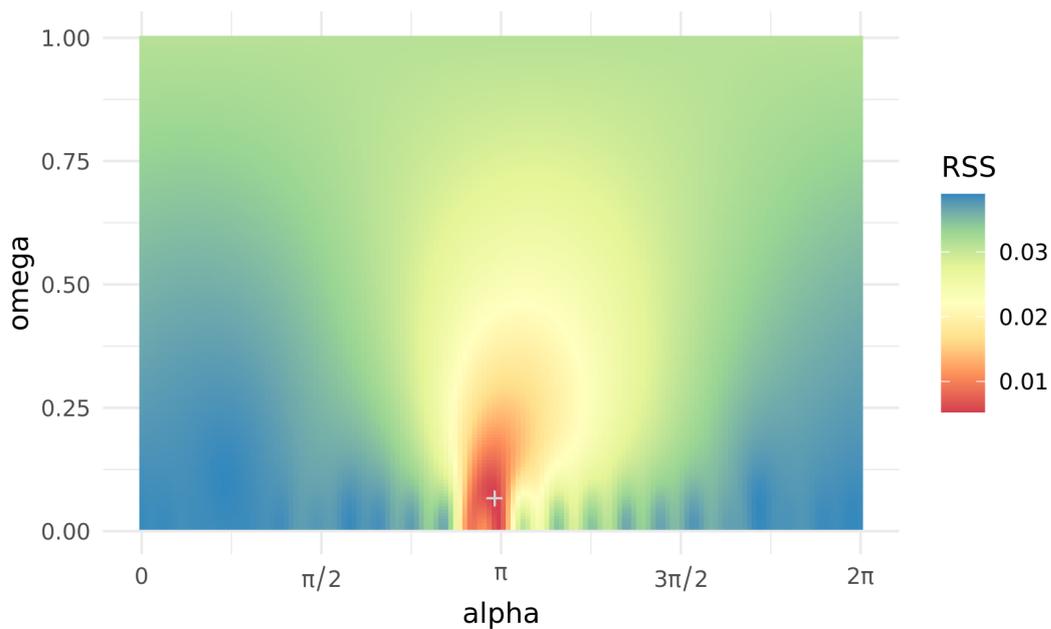


Figura 3.9: RSS para el grid (α, ω) en el análisis de datos de Expresión del gen IQGAP2 en ratones.

FMM multicomponente

El modelo FMM multicomponente se define como un modelo paramétrico formulado como la suma de m ondas más el ruido. Específicamente:

Definición 3.7 Modelo FMM_m

$$X(t_i) = \mu(t_i, \theta) + \epsilon(t_i), \quad i = 1, \dots, n,$$

donde:

1. $\mu(t_i, \theta) = M + \sum_{j=1}^m W(t; v_j)$ y $\theta = (M, v_1, \dots, v_m)$ que verifican que:

- $M \in \mathbb{R}$
- $v_j \in \Theta_j = \mathbb{R}^+ \times [0, 2\pi) \times [0, 2\pi) \times [0, 1]; j = 1, \dots, m$
- $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_m \leq \alpha_1$
- $A_1 = \max_j(A_j), j \in \{1, \dots, m\}$

2. $(\epsilon(t_1), \dots, \epsilon(t_n))' \sim N_n(0, \sigma^2 I_n)$

La ventaja de definir el modelo como una suma de ondas es que conocemos la interpretación individual los parámetros y sirve para describir cada una de las ondas que componen un modelo.

En la Figura 3.10 se muestra un ejemplo de ajuste de un FMM₅ a los datos de un latido de la derivación II de un ECG normal (sin patologías). En este caso, la señal oscila más de una vez. Un FMM₁ no sería adecuado para el ECG completo, puesto que se recogería únicamente la variabilidad de la oscilación más pronunciada. Pueden verse además las cinco ondas que componen la señal. La Tabla 3.2 contiene los distintos parámetros de las ondas.

\hat{M}	Componente	\hat{A}	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\omega}$
5.271	Wave 1 - Rojo	0.645	5.515	3.293	0.032
-	Wave 2 - Azul	0.099	4.420	3.770	0.136
-	Wave 3 - Verde	0.244	5.351	0.664	0.032
-	Wave 4 - Morado	0.316	5.592	4.865	0.013
-	Wave 5 - Naranja	0.067	1.799	2.128	0.163

Tabla 3.2: Parámetros de FMM₅ estimados en datos de ECG normal.

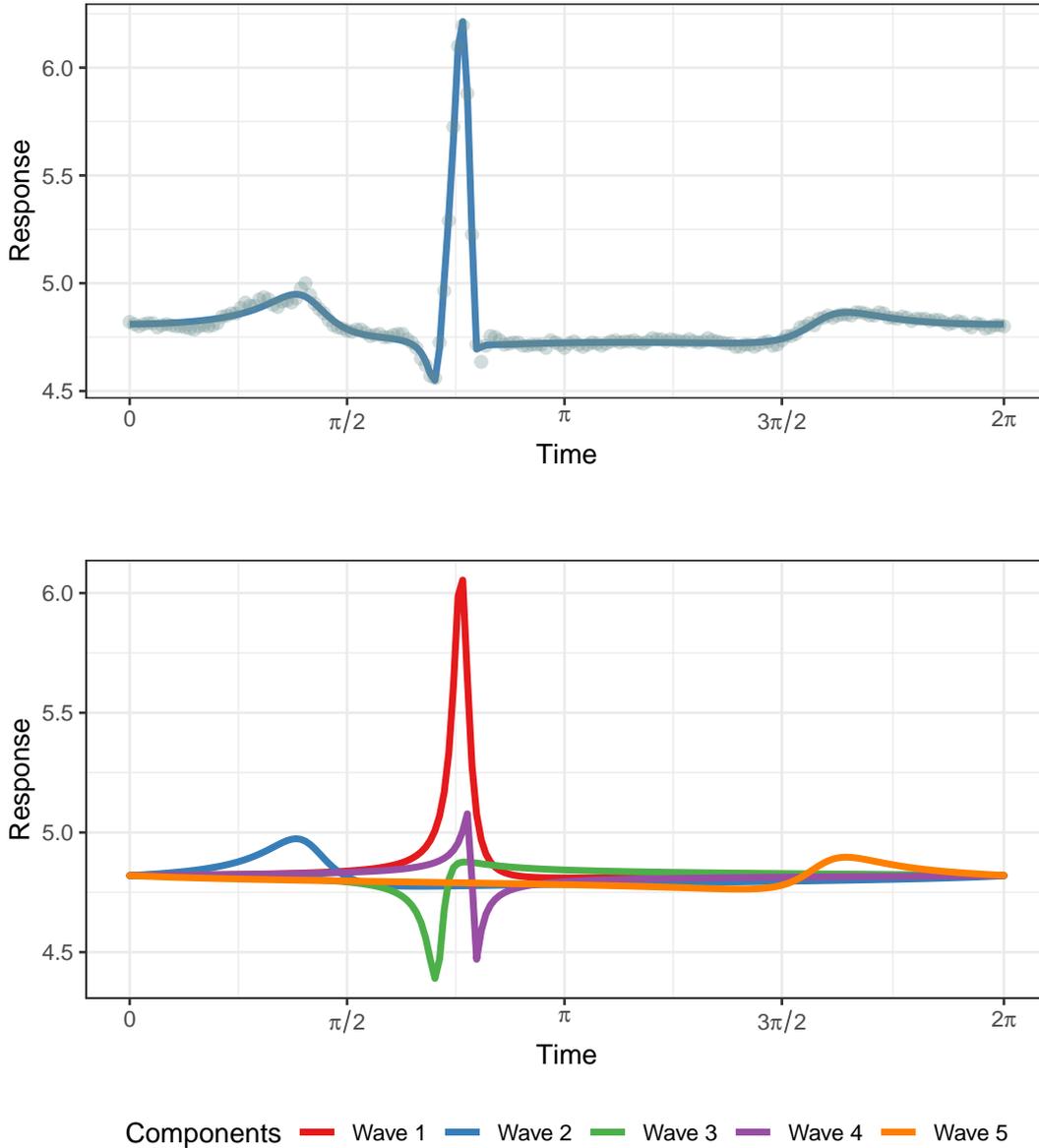


Figura 3.10: Ajuste de un FMM_5 a los datos de un ECG normal.

Estimación de los parámetros del modelo FMM_m

De igual forma que en la estimación de los parámetros en un FMM_1 , los parámetros del modelo FMM_m son la solución del problema de optimización:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{i=1}^n (X(t_i) - \mu(t_i, \theta))^2 \quad (3.11)$$

donde Θ hace referencia al espacio paramétrico de $\theta = (M, v_1, \dots, v_m)$ que se especifica en la definición del modelo (3.7).

La estimación conjunta y simultánea de los parámetros de un FMM_m en el problema 3.11 no es fácilmente resoluble. Para un modelo FMM_3 , la búsqueda se lleva a cabo en un grid multidimensional $(\alpha_1, \omega_1, \alpha_2, \omega_2, \alpha_3, \omega_3)$. Esto supone que para una longitud de cada

α_j de 48 y de ω_j de 24, $j = 1, 2, 3$, se tendrían que calcular más de mil millones de sistemas lineales, lo que en la práctica es inabordable. Por ello se recurre a un algoritmo de backfitting que se describe en [5]. El algoritmo obtiene de forma iterativa los estimadores de las ondas ajustando un FMM_1 a cada residuo.

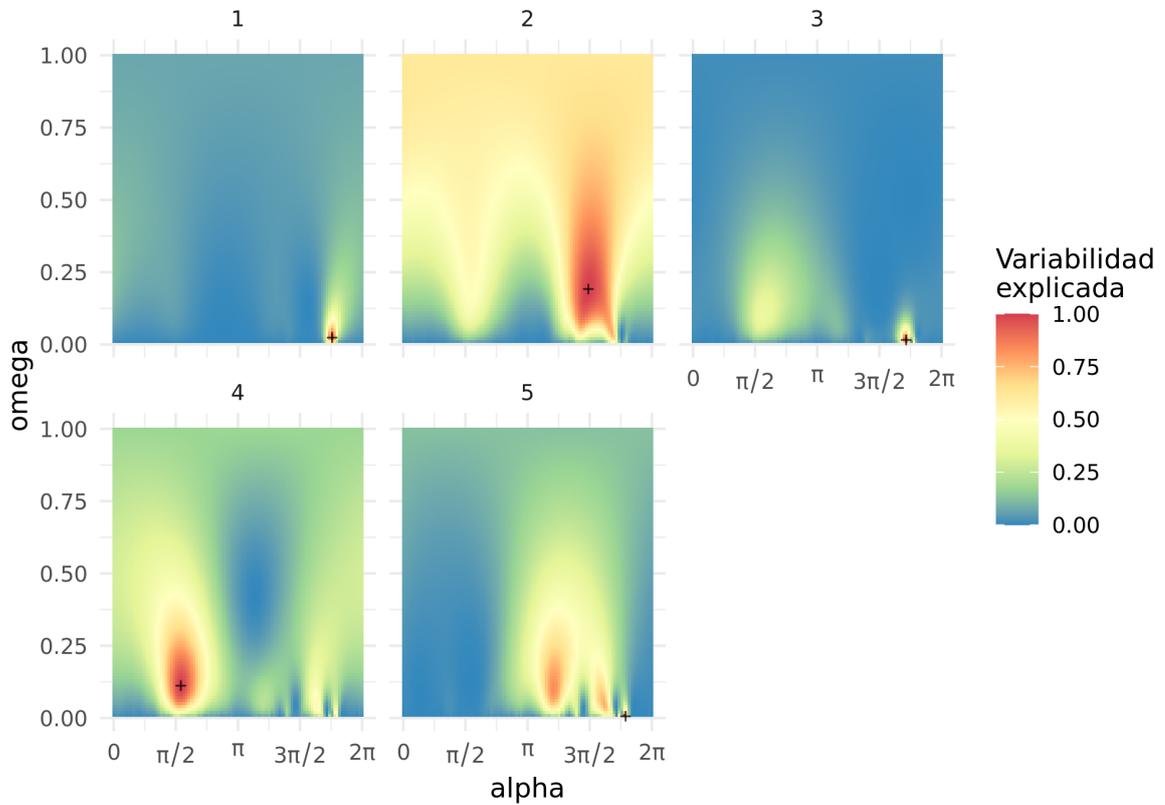


Figura 3.11: Búsqueda de parámetros óptimos en el algoritmo de backfitting (Datos de la Figura 3.10, FMM_5)

En la Figura 3.11 se ilustra la búsqueda de los parámetros de las sucesivas ondas FMM durante el proceso de backfitting, las cuales se muestran en 3.10. La escala de la proporción de varianza explicada en la Figura recorre desde el azul (0.00) que corresponde a la onda que peor ajusta a los datos, hasta la mejor onda en rojo (1.00). Mientras que en el caso monocomponente (Figura 3.9) no parecía haber un problema de existencia de múltiples mínimos locales, este caso evidencia la complejidad del problema de optimización.

Software: FMM en R.

Recientemente, el paquete $FMM^{[9]}$ ha sido incorporado al repositorio CRAN de R (Comprehensive R Archive Network). El paquete implementa los algoritmos de ajuste tanto de FMM_1 , como de FMM_m , incluyendo una versión de cómputo paralelizado de los mismos. El paquete también pone a disposición del usuario varios conjuntos de datos reales y diversas funcionalidades, como la simulación de señales a partir de parámetros específicos o la representación gráfica de modelos FMM.

FMM_{EKG}

Un caso particular del FMM_m es el FMM_{EKG}. Se trata de un modelo que concreta la formulación del FMM_m mediante la asignación de las ondas del FMM con las ondas fundamentales de un electrocardiograma, por lo que cada posible onda es asignada como $J \in \{P, Q, R, S, T\}$. Se describe a continuación:

Definición 3.8 Modelo FMM_{EKG}

$$X(t_i) = \mu(t_i, \theta) + \epsilon(t_i), \quad i = 1, \dots, n,$$

donde

$$\mu(t, \theta) = M + \sum_{J \in \{P, Q, R, S, T\}} W_J(t; v_J)$$

1. $\theta = (M, v_P, v_Q, v_R, v_S, v_T)$ que verifican que:

- $M \in \mathbb{R}$
- $v_J \in \Theta_J = \mathbb{R}^+ \times [0, 2\pi) \times [0, 2\pi) \times (0, 1]; J \in \{P, Q, R, S, T\}$
- $\underline{\alpha_P \leq \alpha_Q \leq \alpha_R \leq \alpha_S \leq \alpha_T}$

2. $(\epsilon(t_1), \dots, \epsilon(t_n))' \sim N_n(0, \sigma^2 I_n)$

La diferencia fundamental de este modelo es la restricción en los α de las ondas identificadas.

En [5] se diseña un algoritmo, denominado MI, de estimación y asignación de las ondas mediante dos pasos diferenciados: M , que realiza la estimación de K ondas, e I , que asigna las ondas del ECG. La Figura 3.12 muestra el diagrama de flujo del algoritmo, el cual recibe como input un segmento del ECG preprocesado que incluye la localización de los complejos QRS en cada latido.

El paso I (asignación) es el que merece más la pena detallar al ser específico del modelo FMM_{EKG}. Se asume que la localización del complejo QRS es conocida (t^{QRS}). La primera onda que se asigna es R . Es la onda de mayor variabilidad en el entorno de t^{QRS} . Además, para derivaciones como la II, donde R es claramente una cresta pronunciada, se busca que la onda tenga parámetros de forma tales como $\beta \in [\pi/2, 5\pi/3]$ o que tenga un apuntamiento pronunciado ($\omega < 0.12$).

Tras haber asignado la onda R , se hace una asignación previa del resto de ondas teniendo en cuenta que se debe verificar: $\alpha_P \leq \alpha_Q \leq \alpha_R \leq \alpha_S \leq \alpha_T$. Cuando se da un latido sin alteraciones morfológicas ni ruido en la señal, la asignación previa es la definitiva. Sin embargo, como se ven en los ejemplos de patologías cardiovasculares introducidas con anterioridad (Figuras 2.5, 2.4, 2.7 y 2.6), la presencia de enfermedades conlleva una alteración del ECG. Es por ello que para llevar a cabo una asignación completa de las cinco ondas son necesarias más reglas que las mencionadas, en las que se hace uso del resto de parámetros de las ondas.

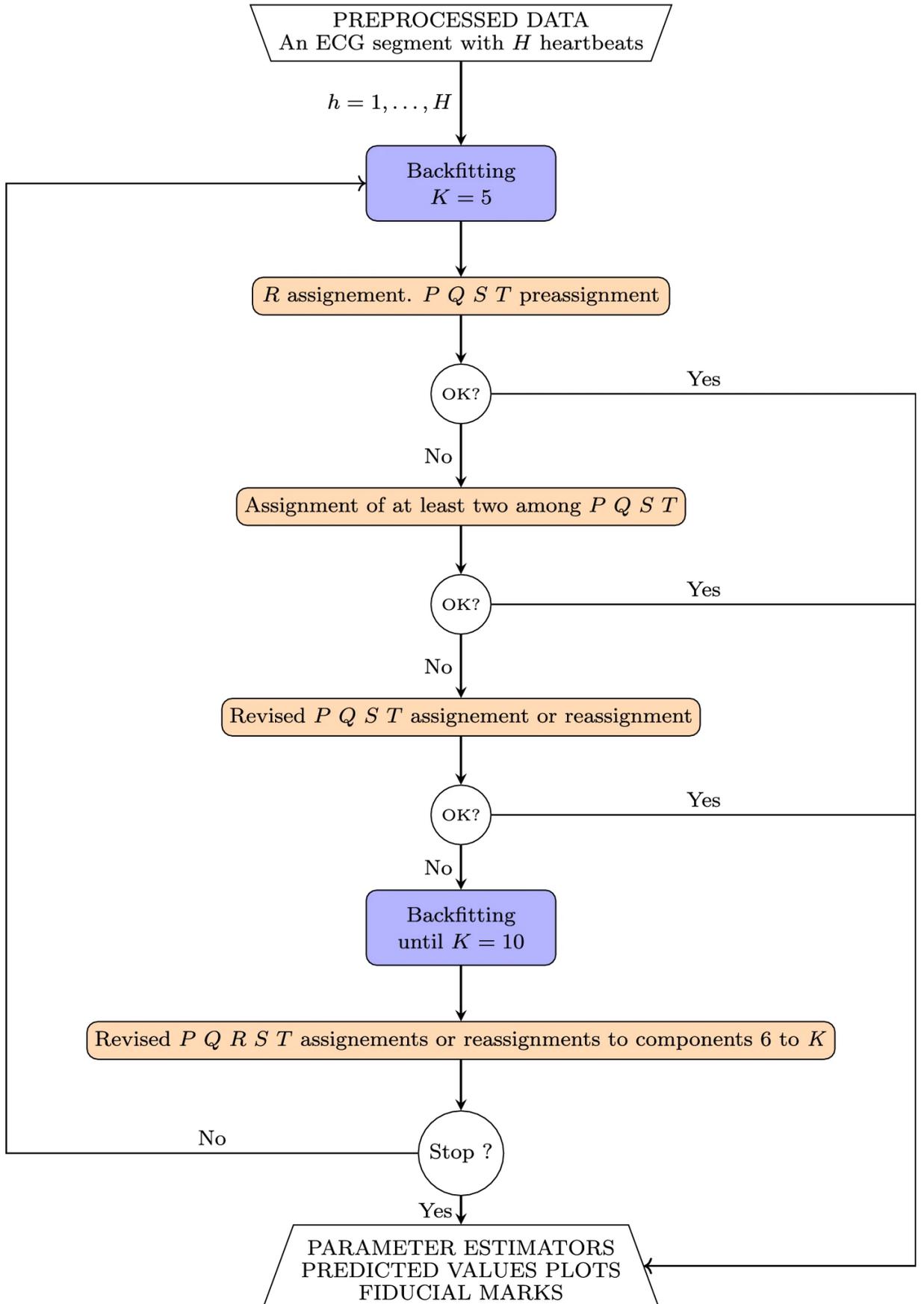


Figura 3.12: Algoritmo de asignación de ondas del ECG. Imagen de [5].

En la Figura 3.13 se muestran varios resultados de aplicar el modelo FMM_{ECG} en distintos pacientes con diferentes patologías. Se puede apreciar cómo las ondas tienen formas diferentes en función de cada caso. Son ejemplos analizados en [5], pertenecientes a la base de datos QT (<https://physionet.org/content/qtdb/1.0.0/>).

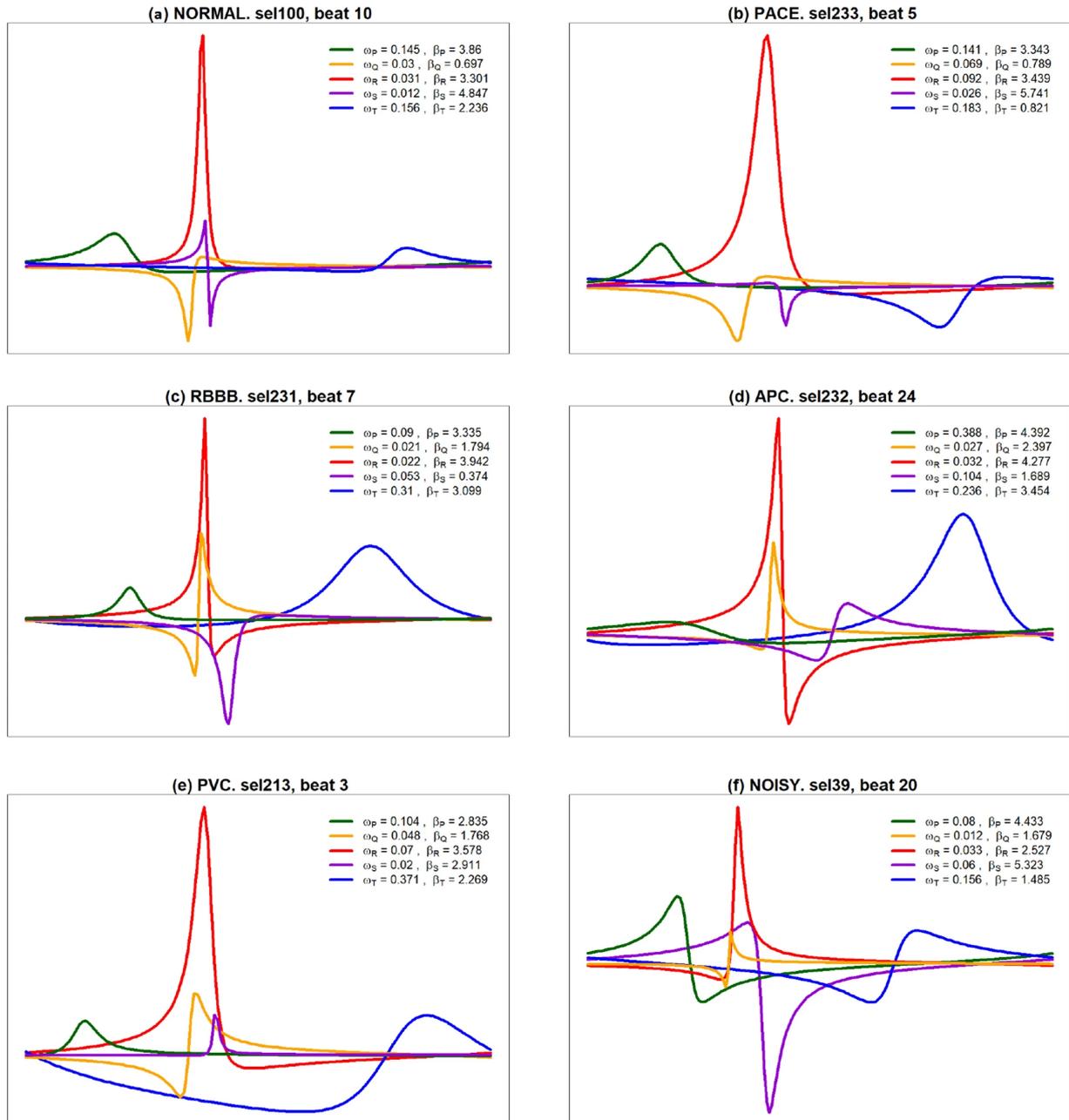


Figura 3.13: FMM_{ECG} para ejemplos de ECG a) Normal, b) PACE, c) RBBB, d) APC, e) PVC y f) con ruido. Imagen de [5].

4 Aplicación para el análisis de señales ECG.

En este capítulo se describe la aplicación creada para el análisis de latidos cardíacos. Por una parte se describen todas las partes que conforman la aplicación, y por otra parte, se muestran algunos ejemplos de su uso en los que se analizan casos reales de ECG.

4.1 Herramienta para la creación de la aplicación: R Shiny.

Uno de los lenguajes de programación actualmente más extendidos en la estadística y la ciencia de datos es R. De hecho, las implementaciones del FMM, disponible en CRAN y del FMM_{ECG} están escritos en este lenguaje. La aplicación interactiva integra estos procedimientos de R necesarios.

El paquete Shiny de R permite construir una aplicación interactiva directamente desde una sesión de R. La implementación de la aplicación puede dividirse en dos partes: la interfaz de usuario (UI) y el *server*. El diseño de la interfaz de usuario combina elementos de R Shiny, estilos de CSS, *widgets* escritos en HTML y acciones de JavaScript. Existen diversos paquetes, como *hinydashboards*, que facilitan la creación de interfaces interactivas, intuitivas y amigables al usuario. Por otra parte, el *server* se encarga de ejecutar, “por detrás” de la interfaz de usuario, todas las órdenes en lenguaje R que procedan.

La programación de aplicaciones Shiny pertenece al paradigma llamado “programación reactiva”. Este paradigma se ocupa de los flujos de datos en un sistema informático y su reacción a los cambios que ocurren en él. La propagación de los cambios se origina, o bien con la llegada de datos (asíncronos, no sabemos cuándo llegan al sistema o cuándo se producen), o bien con eventos “observables”, como un clic en un botón de la aplicación. Un sistema pueden estar formado por flujos de datos concurrentes, y la sincronización de estos también forma parte de la programación reactiva.

Un ejemplo claro de programación reactiva se da en las hojas de Excel. En ella, un usuario puede introducir datos en celdas y usar funciones que reciban la información de las celdas como argumento. La programación reactiva asegura que cualquier cambio en esas celdas desencadene una reacción en las celdas que dependan de ellas.

4.2 Descripción de la aplicación.

El flujo de datos de la aplicación se puede dividir en los bloques siguientes:

1. Selección de datos, opciones de análisis y preprocesado de los datos. El preprocesado incluye el escalado de los datos y el filtrado de la tendencia lineal (si existe).
2. Localización de los complejos QRS, división del segmento en latidos.
3. Ajuste del modelo FMM_{ECG} a las derivaciones seleccionadas.
4. Salidas: gráficos de la señal, tablas con los parámetros FMM y asignación de ondas, e indicadores patológicos.

En la Figura 4.1 se muestra un esquema general de la aplicación.

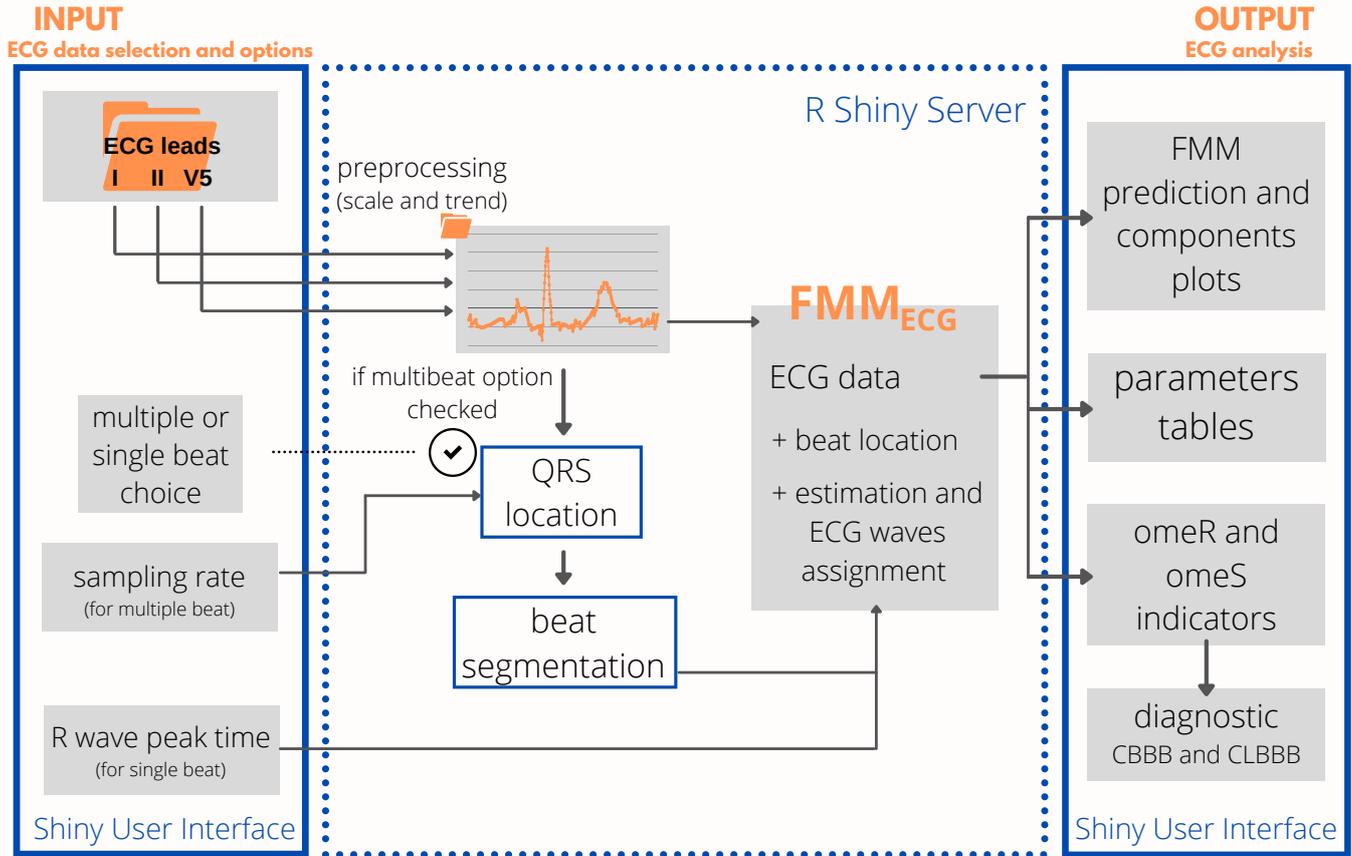


Figura 4.1: Esquema del funcionamiento de la aplicación.

4.2.1 Input y preprocesado de datos.

La aplicación permite seleccionar datos de entrada por parte del usuario. Estos datos deben ser un archivo de Excel (.xls, .xlsx) o .csv con cuatro columnas correspondientes al tiempo y al valor de la señal en las derivaciones I, II, y V5. Las cuatro columnas, por tanto, deben ser numéricas.

Una vez se han cargado los datos, previo al ajuste del modelo FMM_{ECG} , es necesario un preprocesado de los datos. Este preprocesado se basa, primeramente en escalar la señal, acotándola en el rango $[-1, 1]$. Así, para la señal X^D de la derivación D :

$$Y_i^D = 2 \frac{X_i^D - \min(X^D)}{\max(X^D) - \min(X^D)} - 1, \quad (4.1)$$

$$i = 1, \dots, n; D \in \{I, II, V5\}$$

Tras haber escalado la señal, se aplica un filtro para eliminar la tendencia.

En el caso de que la señal tenga múltiples latidos, para ajustar el modelo es necesario

segmentar la señal en latidos individuales, lo que requiere localizar los complejos QRS del segmento.

Esto se lleva a cabo mediante el algoritmo de Pan-Tompkins [10] (se debe especificar la frecuencia de muestreo de la señal). La implementación del algoritmo en R se encuentra disponible en el paquete de `rsleep`. Cuando la frecuencia de muestreo de la señal y el ritmo del latido (distancia RR) es constante, la localización del QRS (t^{QRS}) nos permite dividir el segmento en los latidos que lo componen de forma directa: un latido se corresponde con las observaciones $X(t_i)$, donde $t_i \in [t^{QRS} - 40\%RR, t^{QRS} + 60\%RR]$, como se indica en [5].

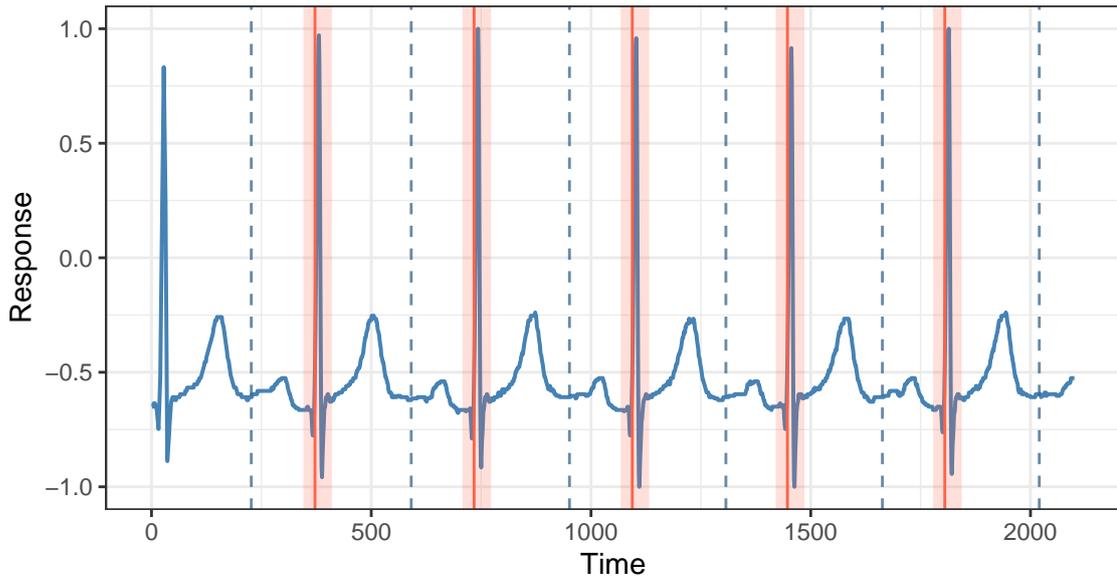


Figura 4.2: Ejemplo de división de segmento ECG. Las líneas en rojo muestran la localización de los complejos QRS

Un ejemplo de la aplicación de todos los pasos anteriores puede verse en la Figura 4.2. La señal ha sido tomada con una frecuencia de muestreo de 500GHz. Se detectan, en este caso, un total de cinco latidos completos. En el inicio de la señal se identifica un complejo QRS que no es seleccionado por el algoritmo de Pan-Tompkins al corresponder a un latido que no ha sido enteramente registrado.

4.2.2 Ajuste del modelo FMM_{ECG} .

El modelo para el análisis de los latidos ha sido descrito en el Capítulo 3, incluyendo la estimación y asignación de ondas (algoritmo MI). Se ajusta un modelo FMM_{ECG} por latido. En el caso de señales multilatido, la aplicación muestra el modelos FMM_{ECG} con los valores medianos de los parámetros.

El ajuste del modelo es computacionalmente pesado, por lo que la interfaz de usuario permite seleccionar qué derivaciones se analizan, así como el número de backfittings que realizará el algoritmo de estimación. El código que se utiliza para esta labor es una versión del que se usa en [5], desarrollada por los mismos autores. A su vez, el código depende del paquete `FMM` [9] para la estimación de los parámetros en el algoritmo MI.

4.2.3 Salidas de la aplicación.

Una vez se han estimado y asignado las ondas del modelo FMM_{ECG} , para cada una de las derivaciones seleccionadas por el usuario, la aplicación proporciona los resultados siguientes:

- Gráfico de la señal $\mu(t_i, \hat{\theta})$, $t_i \in [0, 2\pi]$; $i = 1, \dots, n$, correspondiente al modelo FMM_{ECG} del latido mediano. Cuando se analiza sólo un latido, el gráfico incluye la señal original escalada $Y(t_i)$.
- Gráfico de componentes $W_J(t_i, \hat{v}_J)$, $J \in \{P, Q, R, S, T\}$, $t_i \in [0, 2\pi]$; $i = 1, \dots, n$ del correspondiente modelo.
- Tabla con los valores $\hat{v}_J = (\hat{A}, \hat{\alpha}, \hat{\beta}, \hat{\omega})$ de cada onda asignada $J \in \{P, Q, R, S, T\}$, la proporción de variabilidad explicada por cada onda y el valor \hat{M} del modelo.

Además de los resultados anteriores, la aplicación proporciona dos marcadores útiles para el diagnóstico de las patologías CBBB y CLBBB. Estas patologías se dan por el bloqueo en una rama de uno de los circuitos principales del corazón: el haz de His. La actividad eléctrica en el haz de His se refleja en el complejo QRS de la señal del ECG. Con estas patologías presentes, la morfología del ECG se ve alterada, en particular en el ensanchamiento de las ondas que componen el complejo QRS. Los marcadores $omeR$ y $omeS$ miden la kurtosis de las ondas R y S, y se definen como sigue:

$$omeR = \max_{D=\{I, II, V5\}} \{\omega_R^D\} \quad (4.2)$$

$$omeS = \max_{D=\{I, II, V5\}} \left\{ \omega_S^D \mid \beta_S^D \in \text{arc} \left(\frac{5\pi}{3}, \frac{\pi}{3} \right) \right\} \quad (4.3)$$

A partir de los marcadores anteriores se diseñan las reglas para la detección de las patologías:

- CLBBB: $omeR > 0.06$.
- CBBB: $omeR > 0.025$ y $omeS > 0.05$.

La explicación completa de todo esto merecería otro capítulo entero que cae fuera de la extensión de este TFG. Es un trabajo que se está desarrollando actualmente en el grupo de investigación, en el cual se va a incluir la aplicación que se ha desarrollado. Se ha considerado interesante una breve explicación, ya que gracias a la implementación de estos marcadores la aplicación permite a un médico identificar patrones de una de estas dos patologías.

4.3 Ejemplos de uso de la aplicación.

En esta sección se muestran ejemplos del uso de la aplicación. En todos los casos se han analizado las señales ECG en las tres derivaciones (I, II y V5), además del número máximo de backfittings que se permiten (cinco). Los diferentes ejemplos forman parte de la base de datos PTBXL[REFERENCIA]. En concreto, latidos de los pacientes número 2305 (paciente sano), 8316 (paciente con CLBBB) y 10131 (paciente con CBBB).

Se asume que los inputs de “*Number of beats - Single/Multiple*”, “*Annotation*” y “*Sampling rate*” introducidos por el usuario, son correctos.

4.3.1 Selección de datos.

El primer paso para realizar un análisis es la selección de un conjunto de datos. Como se muestra en la Figura 4.3, la selección se realiza fácilmente mediante un cuadro de diálogo del sistema operativo. Como ya se ha mencionado, los formatos válidos son .xls, .xlsx o .csv, lo que supone que el cuadro de diálogo únicamente nos muestra los archivos con estas extensiones.

El usuario puede comprobar si los datos se han cargado correctamente, tanto por el mensaje “*Upload completed*”, como por el texto de “*Number of observations*” (Por defecto, marcará “-”).

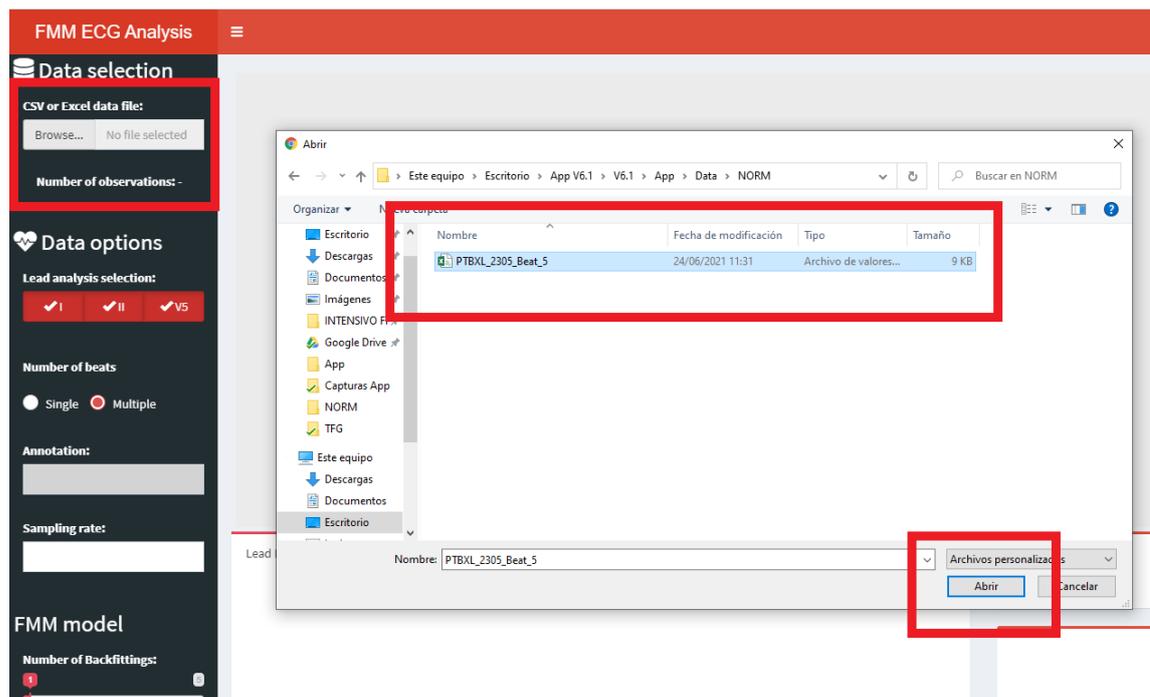


Figura 4.3: Selección de datos para el análisis.

Una vez se seleccionan los datos y se han rellenado o marcado las opciones correspondientes, el usuario deberá pulsar el botón de “*Fit Model*” para iniciar el algoritmo de estimación y asignación de ondas.

4.3.2 ECG normal.

Las Figuras 4.4 y 4.5 corresponden al análisis del ECG de un individuo sin patologías cardíacas. Se trata de un único latido, por lo que el usuario debe introducir en “*Annotation*” la localización del complejo QRS (en este caso, 224), además de marcar la opción de “*Number of observations - Single*”.

En la Figura de los componentes se aprecia la asignación de las cinco ondas del ECG a las componentes estimadas. Como se puede observar, corresponden con la descripción de un ECG normal que se daba en el Capítulo 2 sobre el ECG.

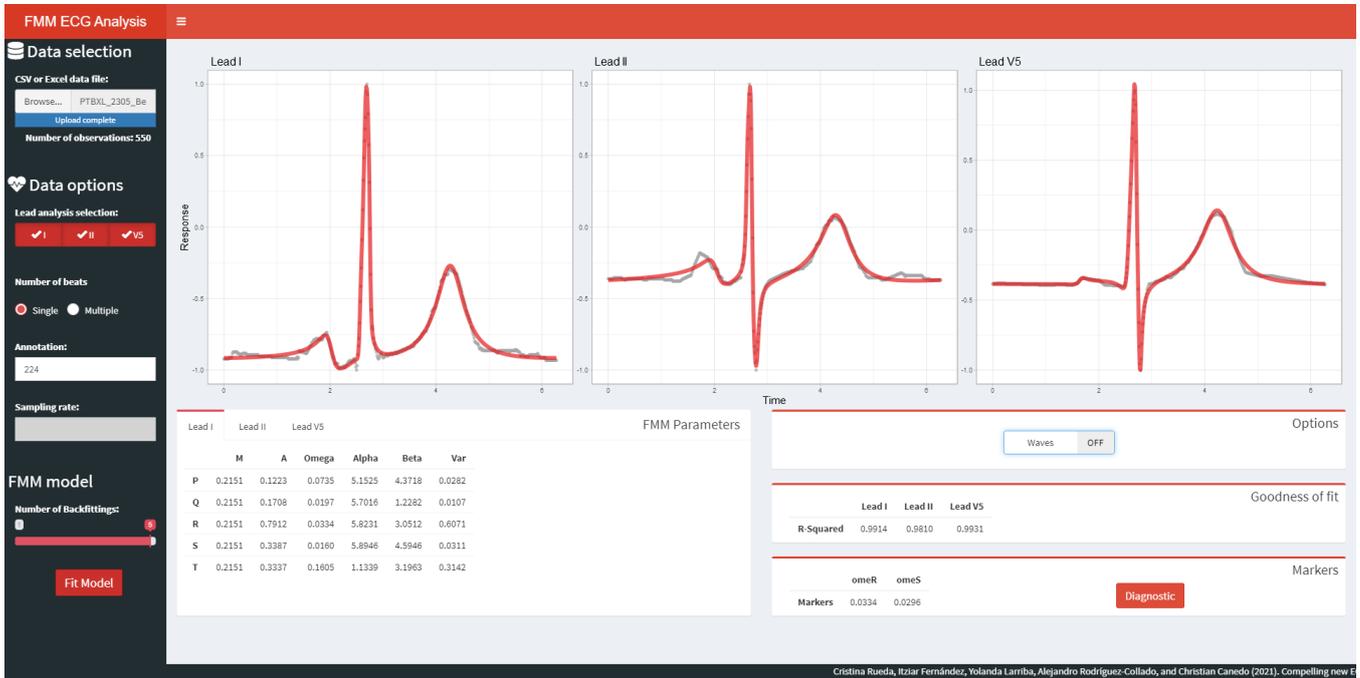


Figura 4.4: Análisis de ECG normal - señal FMM.



Figura 4.5: Análisis de ECG normal - componentes FMM.

Después de la asignación de ondas, mediante el botón “*Diagnostic*”, el usuario puede comprobar si el ECG del paciente muestra riesgo o no de alguna de las patologías para las que se ofrece diagnóstico automático. En la Figura 4.6 aparece el diagnóstico para el ECG de las Figuras 4.4 y 4.5 (ECG con marcadores sanos).

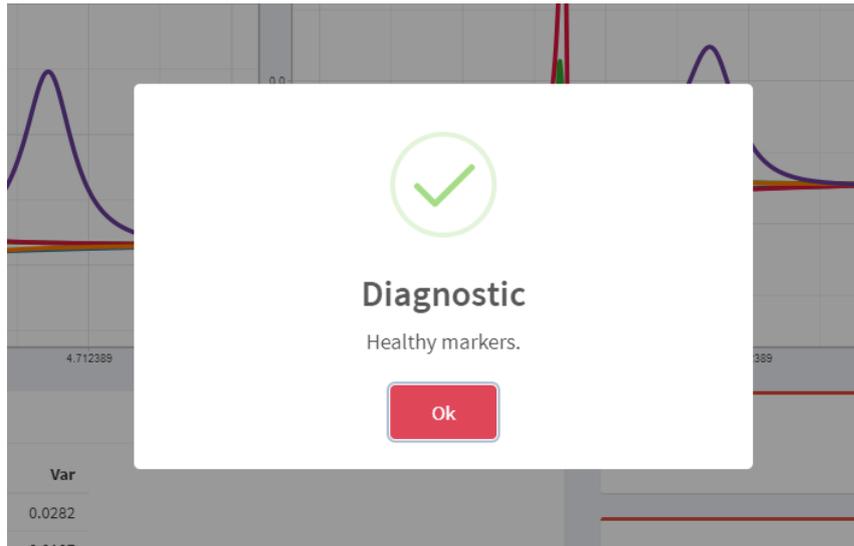


Figura 4.6: Análisis de ECG normal - diagnóstico.

4.3.3 ECG con CBBB.

El siguiente ejemplo de análisis corresponde con un individuo con la patología CBBB. Para el análisis de estas patologías es conveniente analizar segmentos del ECG con varios latidos, de forma que el diagnóstico obtenido a partir del latido mediano es más robusto que el de un latido.

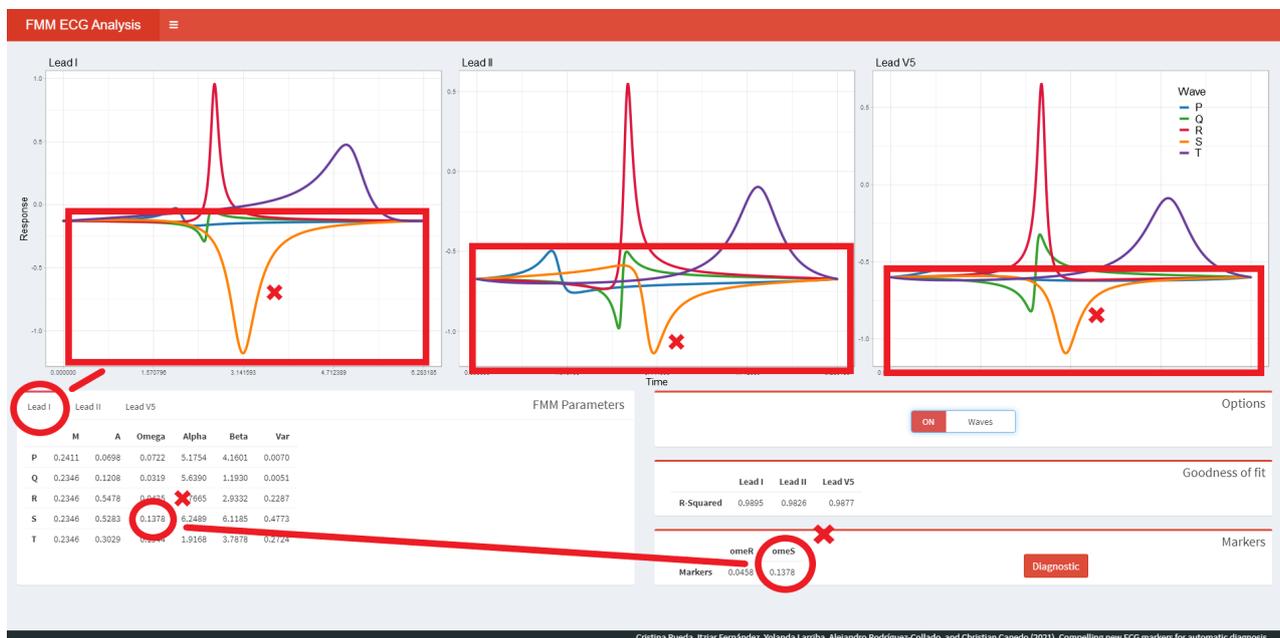


Figura 4.7: Análisis de ECG con CBBB - componentes FMM.

Como se puede comprobar en la Figura 4.7, las ondas S son anchas en comparación con las de un ECG normal. El parámetro ω_S , que mide el apuntamiento de la onda S, así como el marcador asociado omeS destacan esta anomalía. La aplicación indica el alto riesgo de CBBB en el paciente (Figura 4.8).

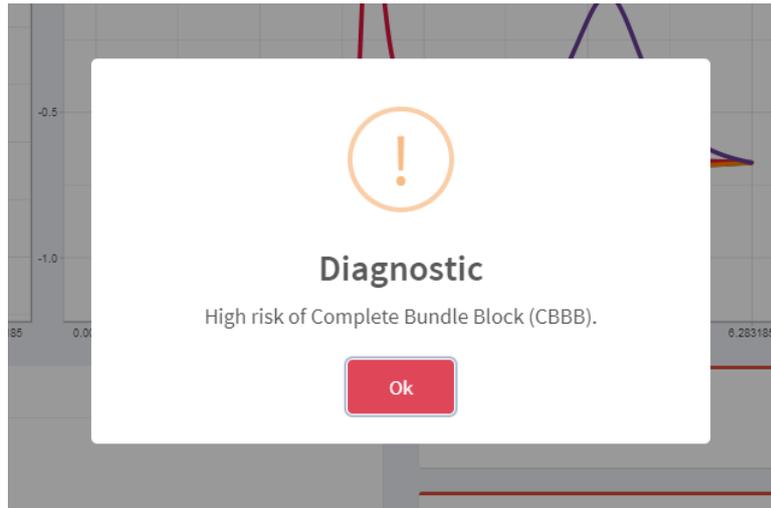


Figura 4.8: Análisis de ECG con CBBB - diagnóstico.

4.3.4 ECG con CLBBB.

Por último, se analiza el ECG de un paciente con CLBBB. En la Figura 4.9 donde se muestra la señal FMM, es notable un ensanchamiento del complejo QRS, lo que queda reflejado en los parámetros de forma de la onda R. El marcador omeR toma el valor máximo de ω_R en las tres derivaciones, que en este caso corresponde con la primera derivación, donde la onda R tiene un $\omega = 0.2308$.

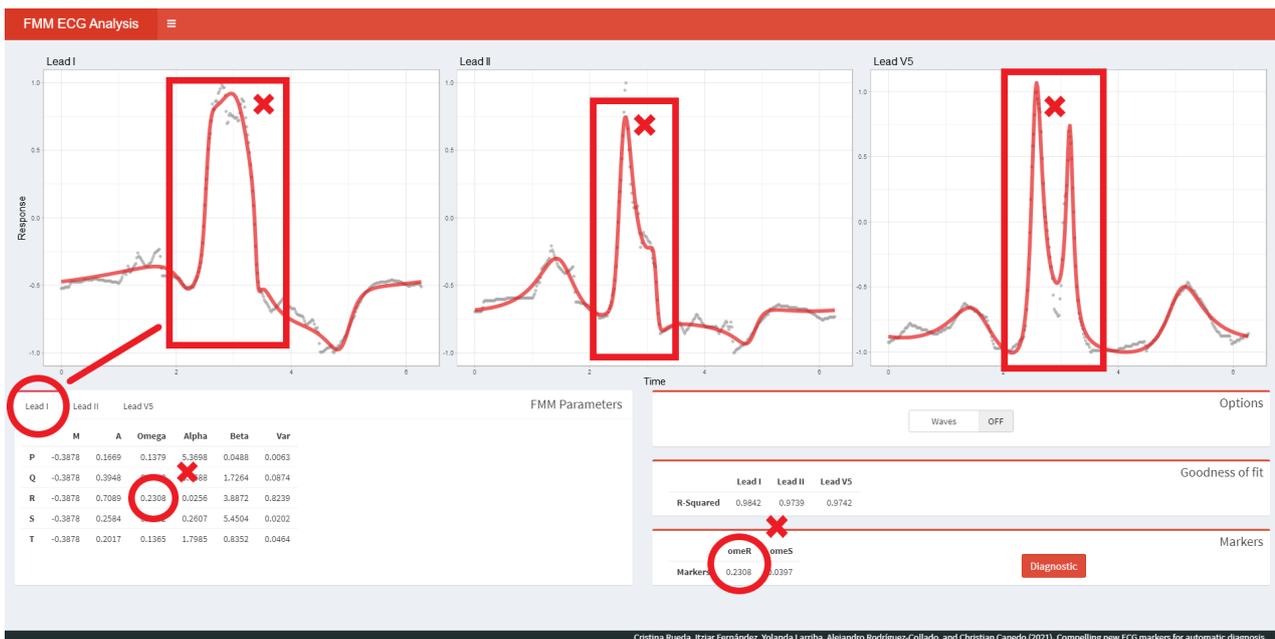


Figura 4.9: Análisis de ECG con CLBBB - señal FMM.

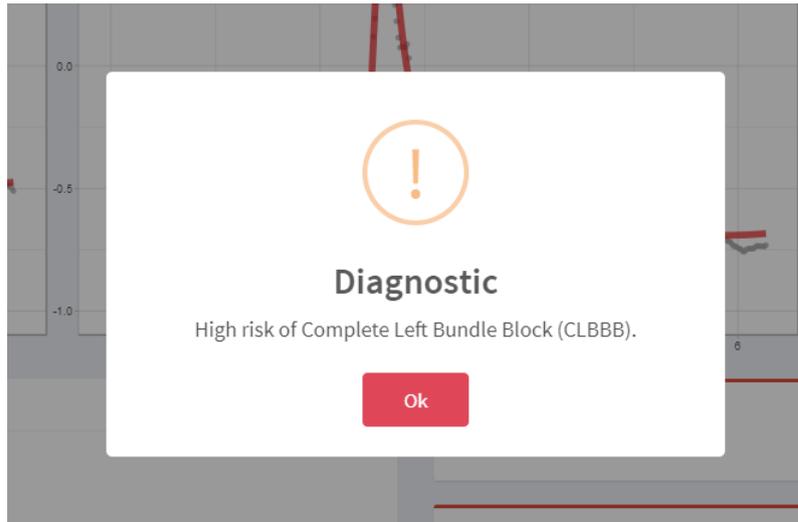


Figura 4.10: Análisis de ECG con CLBBB - diagnóstico.

Como resultado de este ensanchamiento (signo de CLBBB), el marcador omeR sobrepasa el umbral correspondiente y la aplicación diagnostica un alto riesgo de CLBBB.

5 Conclusión.

En este trabajo se han estudiado e introducido diferentes señales oscilatorias y su tratamiento mediante modelos paramétricos, destacando el modelo FMM.

El modelo FMM presenta ventajas importantes frente a alternativas, especialmente por ser paramétrico. Por una parte, las señales FMM pueden delinear una gran variedad de formas como las que se han mostrado en la memoria. Por otra parte, las ondas del modelo FMM, así como sus parámetros, son interpretables biológicamente, lo que es de gran utilidad en la práctica. En particular, en el análisis de señales ECG nos ayuda a describir las ondas de la contracción y relajación de ventrículos y aurículas, pudiendo extraer características de latidos sanos y patológicos.

La aplicación creada en este trabajo pretende facilitar el análisis con el FMM de las señales mediante una interfaz que incluye los gráficos de los valores predichos del modelo, la asignación de ondas del ECG y tablas con los valores estimados de los parámetros. La aplicación puede ser mejorada aumentando las funcionalidades, mejorando el diseño de la interfaz y, sobre todo, reduciendo el tiempo de cómputo de la estimación de los parámetros.

Existe mucho trabajo aún por delante, tanto en el ámbito teórico, como en el computacional y en el práctico. El modelo FMM, como tantos otros modelos estadísticos se compone de un término determinista y otro de error. Se asume que este error proviene de ruido en la medida totalmente independiente de la señal, sin embargo, es interesante tener en cuenta un posible error intrínseco a la actividad biológica o física asociado a la fase de la señal FMM.

En cuanto al aspecto computacional, la estimación de los parámetros del modelo en señales con varias oscilaciones conlleva un gran trabajo computacional. La dimensión del espacio paramétrico crece exponencialmente con el número de componentes y son necesarios métodos de optimización específicos para ello. El lenguaje en el que están implementados todos los procedimientos (R) es un lenguaje actualmente muy popular en Big Data y ciencia de datos, sin embargo, es habitual que la implementación de la estimación se realice en otros lenguajes de mayor eficiencia computacional. Entre otras cosas, ayudaría a mejorar los tiempos de reacción en la aplicación.

En cuanto a las aplicaciones, el modelo FMM está logrando buenos resultados en todos los campos donde las señales oscilatorias juegan un papel principal. En muchos de estos campos, existen modelos específicos que suelen ser complejos y difíciles de entender o interpretar, y es interesante contrastar su papel con la perspectiva que puede aportar el FMM.

Bibliografía

- [1] Cristina Rueda, Yolanda Larriba y Shyamal D. Peddada. «Frequency Modulated Möbius Model Accurately Predicts Rhythmic Signals in Biological and Physical Sciences». En: *Scientific Reports* 9 (dic. de 2019). doi: 10.1038/s41598-019-54569-1.
- [2] Antoni Bayes de Luna. *Manual de electrocardiografía básica*. 2014. Cap. 9.
- [3] Yolanda Larriba, Cristina Rueda, Miguel A. Fernández y Shyamal D. Peddada. «Order restricted inference in chronobiology». En: *Statistics in Medicine* 39.3 (2020), págs. 265-278. doi: <https://doi.org/10.1002/sim.8397>.
- [4] Cristina Rueda, Alejandro Rodríguez-Collado y Yolanda Larriba. «A Novel Wave Decomposition for Oscillatory Signals». En: *IEEE Transactions on Signal Processing* 69 (2021), págs. 960-972. doi: 10.1109/TSP.2021.3051428.
- [5] Cristina Rueda, Yolanda Larriba y Adrian Lamela. «The hidden waves in the ECG uncovered revealing a sound automated interpretation method». En: *Scientific Reports* 11 (feb. de 2021). doi: 10.1038/s41598-021-82520-w.
- [6] Alejandro Rodríguez-Collado y Cristina Rueda. «A Simple Parametric Representation of the Hodgkin-Huxley Model». En: *PLOS One* (2021). [Accepted, unpublished]. doi: 10.1101/2021.01.11.426189. url: <https://www.biorxiv.org/content/10.1101/2021.01.11.426189v1>.
- [7] T. D. Downs y K. V. Mardia. «Circular regression». En: *Biometrika* 89.3 (ago. de 2002), págs. 683-698. doi: 10.1093/biomet/89.3.683.
- [8] Shogo Kato, Kunio Shimizu y Grace S. Shieh. «A CIRCULAR–CIRCULAR REGRESSION MODEL». En: *Statistica Sinica* 18.2 (2008), págs. 633-645.
- [9] I. Fernández, A. Rodríguez-Collado, Y. Larriba, A. Lamela, C. Canedo y C. Rueda. «FMM: An R Package for Modeling Rhythmic Patterns in Oscillatory Systems». En: (2021). url: <https://arxiv.org/abs/2105.10168>.
- [10] Jiapu Pan y Willis J. Tompkins. «A Real-Time QRS Detection Algorithm». En: *IEEE Transactions on Biomedical Engineering* BME-32.3 (1985), págs. 230-236. doi: 10.1109/TBME.1985.325532.

