



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

UN CASO DE ANÁLISIS DE SEÑALES EN PECES ELÉCTRICOS

Universidad de la República

Facultad de Ciencias Económicas y Administración

Licenciatura en Estadística

Informe de Pasantía

Gonzalo De Armas

Tutor: Prof. Ing. Enrique Cabaña

Tribunal:
Prof. Ing. Enrique Cabaña
Dra. Rossana Perrone
Dr. Marco Scavino

Montevideo, Noviembre de 2017

AGRADECIMIENTOS:

A Enrique Cabaña por su tutoría en esta pasantía.

A Ramon Alvarez, Laura Nalbarte, Marco Scavino y Florencia Santiñaque por su ayuda en la corrección del informe.

A Rossana Perrone y al Instituto de investigación biológica Clemente Estable por proporcionarme los datos y el asesoramiento técnico para realizar esta pasantía.

Dedicado a mis abuelos y a Darko (uno de mis gatos) que ya no están conmigo.

Índice

1. RESUMEN EJECUTIVO	6
2. OBJETIVOS	7
3. INTRODUCCIÓN	8
4. ANTECEDENTES	11
5. METODOLOGÍA UTILIZADA	12
5.1. Detección de pulsos y chirps	12
5.1.1. Detección de pulsos	12
5.1.2. Detección de chirps	12
5.2. Separación de señales	14
5.3. Análisis de chirps	15
5.3.1. Forma del chirp	15
5.3.2. Optimización - Algoritmo Nelder Mead	15
5.4. Análisis de componentes principales	16
5.5. Análisis de conglomerados	17
5.6. Prueba de modelos	19
5.6.1. Prueba de hipótesis	19
5.6.2. Prueba de ajuste de modelos de interacción en el presente estudio	19
5.7. Simulación	20
5.7.1. Método de la función inversa	20
6. RESULTADOS	21
6.1. Detección de pulsos y chirps	21
6.2. Separación de señales	22
6.3. Ajuste de curva	23
6.3.1. Parábola	23
6.3.2. Polinomio de grado 4	25
6.4. Análisis de componentes principales	26
6.5. Análisis de conglomerados	29
6.6. Ajuste de modelos	32
7. CONCLUSIONES	33
8. PERSPECTIVAS A FUTURO	34
9. APÉNDICE I : Ajustes de chirps	35
10. APÉNDICE II : PROGRAMAS	46
10.1. Detección pulsos y chirps	46
10.2. Identificación de pulsos según cada pez	47
10.3. Ajuste de función cuadrática	48
10.4. Ajuste por polinomio grado 4	49
10.5. Análisis de conglomerados	50
10.6. Evaluación de modelos	51
11. Bibliografía	54

Índice de figuras

1.	Voltaje relativo en función del tiempo	8
2.	Ejemplo de un Chirp en el contexto de la señal de pulsos	8
3.	Ejemplos de chirps.	9
4.	Distintos tipos de Chirps.	11
5.	Detección de pulsos	21
6.	Detección de pulsos y chirps mediante función de potencia.	21
7.	Clasificación de señales según pez emisor	22
8.	Ajuste de los chirps mediante una parábola	23
9.	Ajuste de los chirps mediante una parábola	24
10.	Ajuste de los chirps mediante una parábola	24
11.	Ajuste de los chirps mediante un polinomio de 4to grado	25
12.	Valores del R^2 según la cantidad de conglomerados elegidos	29
13.	Valores del Pseudo t^2 según la cantidad de conglomerados elegidos.	30
14.	Valores del Pseudo F según la cantidad de conglomerados elegidos	30
15.	Resultado del análisis de conglomerados	31
16.	Ajuste por polinomio de grado 4	35
17.	Ajuste por polinomio de grado 4	36
18.	Ajuste por polinomio de grado 4	37
19.	Ajuste por polinomio de grado 4	38
20.	Ajuste por polinomio de grado 4	39
21.	Ajuste por polinomio de grado 4	40
22.	Ajuste por polinomio de grado 4	41
23.	Ajuste por polinomio de grado 4	42
24.	Ajuste por polinomio de grado 4	43
25.	Ajuste por polinomio de grado 4	44
26.	Ajuste por polinomio de grado 4	45

Índice de cuadros

1.	Cosenos y cosenos al cuadrado	26
2.	Valores propios según cantidad de componentes	26
3.	Proporción de las varianzas según cantidad de componentes	26
4.	Coordenadas de las variables originales en los nuevos componentes	26
5.	Coordenadas de las variables originales en los nuevos componentes	28
6.	Valores de R^2 , <i>Pseudo F</i> y <i>Pseudo t^2</i> obtenidos por el algoritmo	29
7.	Percentiles del error de los datos reales	32

1. RESUMEN EJECUTIVO

En el contexto de trabajos de investigación hechos en el Instituto de Investigaciones Biológicas Clemente Estable, se hicieron análisis referentes a las señales eléctricas emitidas por una pareja de peces de la especie *Brachyhyppomus gauderio*, tanto para la automatización en la detección y clasificación de dichas señales, como también para demostrar la teoría de que dichas señales interactúan entre sí. En primera instancia se trabajó con el valor del voltaje de las señales y se acumuló su valor en un cierto intervalo para determinar en qué momento ocurre un pulso o un “chirp”. En cuanto a los pulsos, tomando como base que entre un pulso y el siguiente, el período no varía demasiado, se determinó a qué pez corresponde cada pulso, estimando donde estaría el siguiente pulso de cada pez, y clasificándolo como perteneciente a uno u otro pez según cual tenga menor diferencia respecto a lo proyectado. Luego se clasificaron a los chirps utilizando una función polinómica de cuarto grado que describa los chirps al actuar de forma similar a una envolvente, realizando con este polinomio un análisis de conglomerado jerárquico. Para ello se tomaron como datos de los chirps los valores del polinomio en cinco momentos de la duración del mismo: al comienzo, al momento del 25 %, 50 %, 75 % de su duración y finalmente su valor al momento de terminar. Finalmente para probar la interacción de las señales se sometió a prueba un modelo lineal sin interacciones contra uno no lineal que contiene interacciones en la secuencia de pulsos de cada pez y para ello se siguió el siguiente procedimiento:

- Se calculó el error con los datos reales en el modelo sin interacciones.
- Se simularon secuencias con el modelo sin interacción.
- Se aplicaron estos datos simulados al modelo con interacción calculando el error absoluto.
- Se ajustó el modelo con interacción a los datos reales calculando el error absoluto.
- Se ordenaron los valores de los errores obtenidos en las simulaciones y se observó que el valor del error obtenido en el modelo con interacciones usando los datos reales queda ubicado en el percentil 5 % inferior, de esta forma se rechaza la hipótesis del modelo sin interacciones.

Palabras claves: chirp, conglomerados, pulso, prueba de hipótesis, simulación.

2. OBJETIVOS

El objetivo general del presente trabajo es analizar las señales de la especie *Brachyhyppopomus gauderio*. Se analizarán tanto los pulsos que son señales de electrolocalización como los chirps que son señales denominadas sociales. Este trabajo tendrá como una primera consigna el poder desarrollar algoritmos que permitan la detección automática de ambos tipos de señales para luego modelar las mismas. En el caso de los pulsos se busca determinar si se puede contrastar un modelo donde la señal de un pez sea independiente del otro contra un modelo que evidencie un cambio en la frecuencia en caso de proximidad de los pulsos de uno y otro pez. En el caso de los chirps interesa analizar su forma para poder determinar cuántos tipos de chirps existen y sus características. Para realizar lo dicho en el párrafo anterior se establecen los siguientes objetivos específicos:

- Detectar los pulsos y chirps automáticamente.
- Identificar las secuencias de pulsos de cada uno de los peces.
- Determinar si hay distintos tipos de chirps considerando la forma general de los mismos.
 - Encontrar una función que describa la forma de cada chirp.
 - Realizar un análisis de conglomerado para investigar la existencia de chirps con características similares en cuanto a su forma.
- Determinar si, como se presume por parte de los investigadores del instituto Clemente Estable, al ocurrir dos pulsos en intervalos de tiempo muy pequeños, los peces varían la frecuencia de emisión de dichos pulsos.

3. INTRODUCCIÓN

El presente trabajo se presenta en el contexto de una investigación del Instituto Clemente Estable, sobre la especie *Brachyhyppomus gauderio* (pez de agua dulce) la cual utiliza la electricidad, en forma de impulsos eléctricos para localizar los objetos a su alrededor.

Esta especie genera una señal eléctrica sobre la cual, a intervalos aproximadamente regulares, ocurren pulsos eléctricos, utilizados para la electrolocalización¹ del pez, y que además ocasionalmente se transforma en una señal diferente denominada “chirp”. Se define como chirp a una desorganización de la señal entendida como un aumento en la frecuencia y una alteración en la amplitud de la señal. Los investigadores suponen que hay una relación entre esos chirps generados por el macho y un apagado de la señal por parte de la hembra durante el cortejo (Macadar & Silva, 2007). A continuación se presentan dos gráficos, por un lado en la Figura 1 se muestran los pulsos de ambos peces a lo largo del tiempo, lo que se puede apreciar es la variación del voltaje (voltaje relativo) al ocurrir un pulso, mientras que en las Figuras 2 y 3 se muestran ejemplos de chirps donde se aprecia la forma irregular de la señal a lo largo del tiempo.

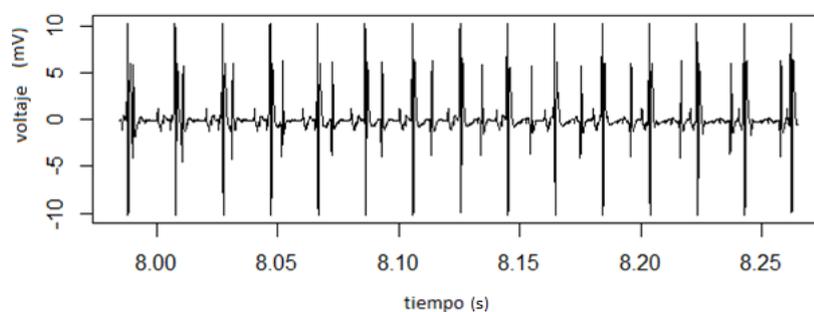


Figura 1: Voltaje relativo en función del tiempo
Fuente: Propia

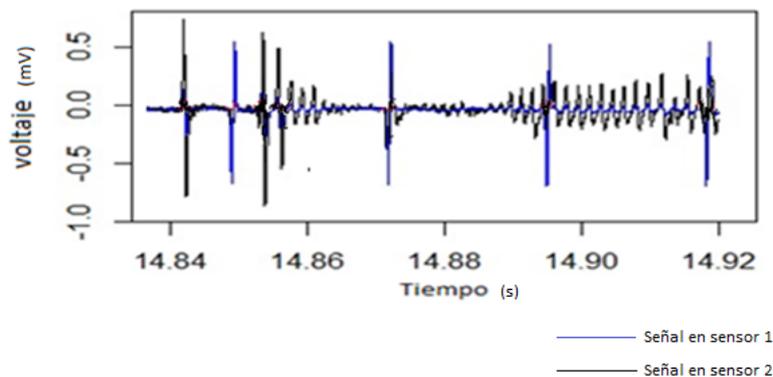


Figura 2: Ejemplo de un Chirp en el contexto de la señal de pulsos
Fuente: Propia

¹Se entiende por electrolocalización, el mecanismo utilizado por esta especie para recibir y hacer uso de los impulsos eléctricos, en este caso para ubicarse en el medio en que habita

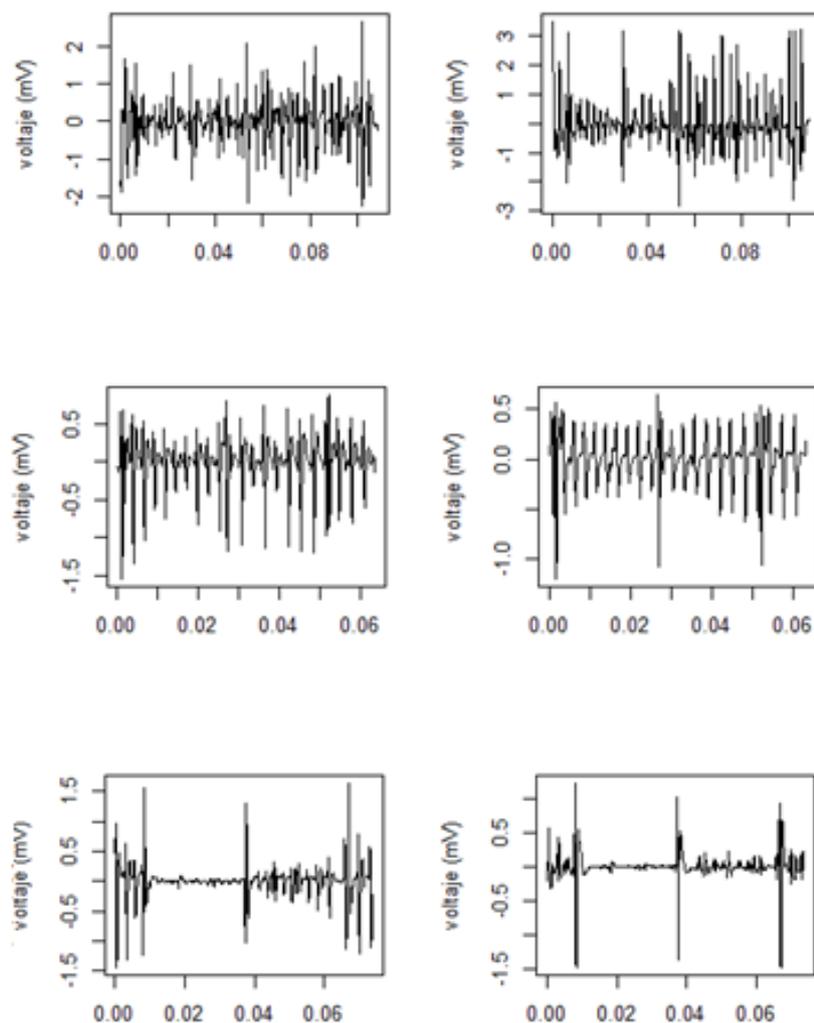


Figura 3: Ejemplos de chirps.
Fuente: Propia

Los peces son separados en parejas macho-hembra en peceras donde se simulan las condiciones de hábitat naturales y se registran los impulsos eléctricos mediante dos sensores colocados en diferentes lugares de la pecera. La necesidad de tener dos sensores se debe a que no se pierda la señal de uno de los peces debido a posiciones particulares del pez respecto a uno de los sensores. Estos sensores registran entonces las amplitudes de la señal instante a instante y a través de un software se obtienen los valores relativos del voltaje, exportándose en archivos de texto conteniendo tres columnas, la primera registra el tiempo que dependerá de la digitalización indicada en el software, la segunda registra la amplitud de la señal respecto al sensor 1, y la tercera registra la amplitud de la señal respecto al sensor 2. Cabe destacar que ambos sensores registran la señal de ambos peces simultáneamente como se ve en la Figura 1 donde se presume que los pulsos de amplitud mayor pertenecen a uno de los peces, y los de amplitud menor al otro pez.

Se asume que cada pez percibe el ambiente en un entorno del momento de emisión del pulso y que para evitar interferirse el uno al otro, si ambos pulsos ocurren en momentos muy próximos, se realiza un corrimiento en la emisión para evitar dicha interferencia. Dicho corrimiento será un parámetro a estimar en el desarrollo de esta pasantía. (Perrone, Macadar, & Silva, 2009)

Por otro lado una dificultad que se presentó en la investigación es que la señal obtenida por los sensores no registra solo los pulsos o chirps sino toda la emisión eléctrica del pez. Por tanto se plantea como un objetivo idear un método que permita detectar en qué momento ocurren tanto los pulsos como los chirps y poder separar en forma automática la señal perteneciente a cada pez.

Este trabajo plantea por lo tanto un estudio de caso de señales de peces eléctricos verificando que en caso de proximidad de los pulsos, los peces realizan un corrimiento del siguiente pulso para evitar la interferencia, así como desarrollar un método que permita detectar en qué momento ocurren los pulsos como los chirps y separar que señal pertenece a cada pez en forma automática, para luego clasificar los chirps.

La presente investigación se estructura de la siguiente manera: a continuación se introducirán los antecedentes relacionados, luego en la sección metodología utilizada se detallarán los procedimientos y algoritmos llevados a cabo para cumplir con los objetivos planteados. A continuación se presentarán los resultados y finalmente las conclusiones obtenidas.

4. ANTECEDENTES

En cuanto a la forma de los chirps en (Perrone, Macadar, & Silva, 2009) se detectaron cuatro tipos de chirps. Se puede observar en la Figura 4 que el chirp A está caracterizado por una larga duración, una fase inicial con máxima frecuencia y mínima amplitud y una segunda fase con picos conformados por oscilaciones en la amplitud después de que la frecuencia inicial aumente y la amplitud decrezca; el chirp denominado B se caracteriza por ser de corta duración y de amplitud decreciente, finalmente el chirp C se caracteriza por tener una duración intermedia a los chirps A y B, un comienzo como el chirp B seguido por una imagen espejo de sí mismo. También se menciona un chirp M de aparición exclusiva en las parejas macho-macho. Los parámetros usados para caracterizar estos chirps fueron: la duración del chirp, la amplitud intrachirp, el tiempo hasta registrar la mínima amplitud intrachirp y un índice de simetría.

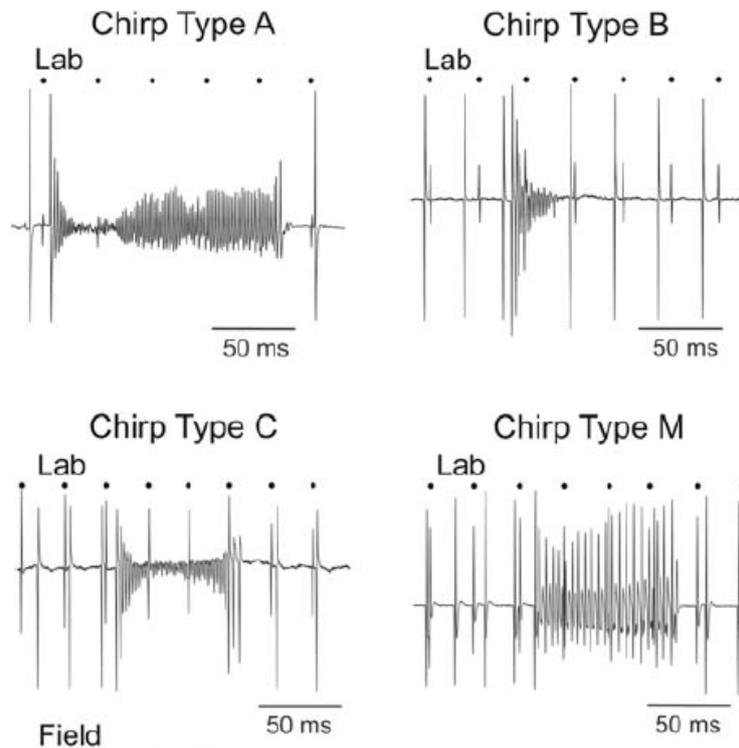


Figura 4: Distintos tipos de Chirps.

Fuente: “Social electric signals in freely moving dyads of *Brachyhypopomus pinnicaudatus*”

En cuanto a la detección automática de chirps, al momento de comenzar esta tesis no había antecedentes exitosos de lograr detectar con exactitud los tiempos en donde ocurre un chirp.

5. METODOLOGÍA UTILIZADA

5.1. Detección de pulsos y chirps

Se mencionó que los datos son arrojados en tres columnas, una con los tiempos y las otras dos con la amplitud de la señal en cada momento t . Debido a la forma en que el programa informático procesa y representa los datos, el tiempo se considera en un conjunto discreto de valores.

Sobre esos datos se tomó la señal de ambos sensores y se sumaron los mismos al cuadrado para que su signo sea positivo y no se compensen al ser sumados. La suma de ambos sensores se hace en la necesidad de evitar pérdida de señal por posiciones particulares respecto a los sensores. Las amplitudes de la señal se elevan al cuadrado porque como se vio en la Figura 1, tenemos valores de la variación del voltaje (valor relativo) al generarse los pulsos, por lo cual hay valores positivos y negativos, la suma al cuadrado evita que se compensen entre sí, luego aplicamos raíz cuadrada para volver a tener amplitudes más manejables.

Sean:

$x(t)$ el valor relativo del voltaje registrado en el primer sensor en el tiempo t

$y(t)$ el valor relativo del voltaje registrado en el segundo sensor en el tiempo t

Llamaremos $p(t)$ al resultado de sumar ambas señales al cuadrado y aplicarles la raíz cuadrada,

$$p(t) = \sqrt{x(t)^2 + y(t)^2}. \quad (1)$$

Se plantea identificar los pulsos realizados por los peces y los chirps, para ello definimos una función “potencia” que se encargue de acumular las amplitudes de la suma al cuadrado de las señales de los peces en ambos sensores, en un intervalo definido por el usuario.

Definimos:

$$potencia(t; n) = \sum_{i=t}^{t+n} p(i). \quad (2)$$

5.1.1. Detección de pulsos

Por las características ya descriptas que tienen los pulsos y en base a esa función de potencia, tomamos que ante un aumento pronunciado pero breve de la misma nos encontramos en la presencia de un pulso. Por tanto, para que un momento t pueda ser candidato a ser pulso debe cumplirse que:

$$potencia(t; n) > u. \quad (3)$$

$$potencia(t; n) \gg potencia(t - \delta; n) + potencia(t + \delta; n). \quad (4)$$

Dada la dificultad de la tarea de detección, se variaron empíricamente los parámetros que determinaron el umbral u y el intervalo δ a considerar. La estrategia utilizada fue tomar un δ arbitrario pero mayor que en el caso del procedimiento de detección de pulsos y variar el umbral u a modo de análisis de sensibilidad de los resultados.

5.1.2. Detección de chirps

Se aplica la función de potencia también para la detección de chirps puesto que dentro de un chirp existe un aumento de la frecuencia de los pulsos, entonces es presumible que si se toma una acumulación de la función potencia en un cierto intervalo (mayor que en el caso de la detección de los pulsos) se considere que existe un chirp utilizando el mismo criterio que los pulsos pero tomando un intervalo de comparación mayor. Al aplicar esta relación se encontraron muchos crecimientos de la función potencia

en lugares donde solo había pulsos, ya sea por aumentos bruscos de la frecuencia o amplitud debido al movimiento de los peces, entonces se buscaron alternativas a dicha estrategia en forma heurística siendo el siguiente procedimiento el que obtuvo mejores resultados.

Sea \mathbf{p} , un vector que contiene la potencia definida por la ecuación 2 para cada momento de tiempo t . El largo de este vector es equivalente a la cantidad de observaciones donde queremos analizar la existencia de chirps.

Sea:

$$m = \text{máx}(\mathbf{p}). \quad (5)$$

Sea \mathbf{p}' un vector de la misma dimensión que \mathbf{p} , que en cada posición contiene el mínimo entre \mathbf{p} y $\frac{m}{a}$,

donde a es un valor a determinar en cada caso en forma heurística, debido a que no pudo determinarse un valor único de pareja en pareja.

$$\mathbf{p}'[k] = \text{mín}\left(\frac{m}{a}, \mathbf{p}[k]\right), \quad k = 1, \dots, n.^2 \quad (6)$$

Esto tiene el efecto de acotar el valor de la función potencia en cada momento del tiempo.

Sea \mathbf{p}'' un vector que acumula la potencia obtenida en cada momento del tiempo en un cierto intervalo de tiempo i (con i determinado en forma heurística),

$$\mathbf{p}''[k] = \sum_{j=k}^{k+i} \mathbf{p}'[j]. \quad (7)$$

Entonces se considera que existe un chirp en aquellos momentos de tiempo t_i donde se cumpla que:

$$\mathbf{p}''[t_i] > \gamma \cdot \text{mediana}(\mathbf{p}''). \quad (8)$$

Donde γ es un valor a determinar en cada caso en forma heurística, debido a que no pudo determinarse un valor único de pareja en pareja.

² $\mathbf{p}'[k]$ refiere a la posición k -ésima del vector \mathbf{p}'

5.2. Separación de señales

Una vez obtenidos los momentos donde ocurren los pulsos debemos determinar a qué pez pertenece cada uno. Para ello tomamos como dato, a que pez pertenecen los primeros 5 pulsos, a tal efecto nos fijamos en las amplitudes y frecuencia de los pulsos. Se buscó, por supuesto, un punto de partida donde las amplitudes no varíen bruscamente y poder determinar de esa manera un comienzo para la secuencia sin confusiones.

Sean:

Z_k momento del k-ésimo pulso a clasificar, como perteneciente a uno de los peces.

$X_{k'}$ momento del k'-ésimo pulso del pez 1 (más próximo, previo, a Z_k).

$Y_{k''}$ momento del k''-ésimo pulso del pez 2 (más próximo, previo, a Z_k).

α : el período de los pulsos del pez 1.

β : el período de los pulsos del pez 2.

Entonces:

$$X_{k+1'} = X_{k'} + \alpha. \quad (9)$$

$$Y_{k+1''} = Y_{k''} + \beta. \quad (10)$$

Se toma el mínimo entre $(Z_k - X_{k+1'})$ y $(Z_k - Y_{k+1''})$ para determinar a qué pez corresponde el pulso, entonces si se encuentra que el mínimo está en $(Z_k - X_{k+1'})$ el pulso pertenece al pez 1, en caso contrario pertenece al pez 2.

No obstante se considera la posibilidad de que existan datos faltantes, o sea, que un pulso no haya sido detectado en el procedimiento de detección de pulsos, a la hora de calcular los mínimos, también planteamos la siguiente posibilidad:

Sean $X'_{k+1'}$ e $Y'_{k+1''}$ tales que:

$$X'_{k+1'} = X_{k'} + 2\alpha. \quad (11)$$

$$Y'_{k+1''} = Y_{k''} + 2\beta. \quad (12)$$

Si el mínimo se da en $(Z_k - X'_{k+1'})$ o en $(Z_k - Y'_{k+1''})$ significaría que tenemos un pulso “perdido” y lo imputamos como dato faltante perteneciente al pez cuya diferencia sea menor, para determinar su posición utilizamos la frecuencia α o β según corresponda a partir del pulso ocurrido en $X_{k'}$ o $Y_{k''}$ según corresponda.

5.3. Análisis de chirps

5.3.1. Forma del chirp

Se busca una función $f(t)$ que describa la forma de los chirps, esta función debe ser tal que $f(t)$ ajuste los máximos locales del chirp. Para ellos se aplica la optimización con el algoritmo de Nelder Mead implementado en R (R core team, 2015), obteniéndose un ajuste para cada chirp.

En este caso “relajado” el criterio de definición de esta función permite que algunos pulsos queden por encima de ella, de tal manera que la mayor parte del chirp quede bien definido por la función $f(t)$ y no se vea afectado por la aparición de pulsos del otro pez o de amplitudes demasiado grandes pero ocasionales del pez que está realizando el chirp. Este margen de tolerancia se determinó en tiempo de ejecución en forma heurística. Una vez encontrada esta función el chirp se describirá por el valor de $f(t)$ en cinco valores de t que son: en $t=0$, $t=0.25$, $t=0.5$, $t=0.75$ y $t=1$, siendo $t=0$ el comienzo del chirp y $t=1$ el final del mismo.

5.3.2. Optimización - Algoritmo Nelder Mead

“El algoritmo Nelder Mead es parte de un método simplex no degenerado en cuyos vértices tenemos aproximaciones al óptimo global. Dicho simplex se va modificando, bien al cambiar algunos de sus vértices por otro en el que la función objetivo haya disminuido, o bien al contraer todo el poliedro si no se ha conseguido mejorar la función objetivo con el simplex inicial. Es por tanto un algoritmo de carácter local que carece además de componente aleatoria” (Maria Jose Zapatero, 2011).

Para dos variables, un simplex es un triángulo y el método es un patrón de búsqueda que compara los valores de la función con los tres vértices del triángulo. El peor vértice, donde la función es más grande es rechazado y sustituido por un nuevo vértice, conformando un nuevo triángulo y continuando el proceso. Esto genera una sucesión de triángulos cada vez más pequeños que obtienen las coordenadas del mínimo. El algoritmo se puede generalizar a N variables tomándose un triángulo en N dimensiones.

Triángulo inicial BGW

$f(x, y)$ será la función a minimizar, que será evaluada en tres puntos: $B = (x_1, y_1)$, $G = (x_2, y_2)$, $W = (x_3, y_3)$, tal que $f(B) < f(G) < f(W)$.

$$M = \frac{B + G}{2} = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right). \quad (13)$$

La función decrece si nos movemos a lo largo de la arista WB y también decrece si nos movemos en la arista WG , por lo que es posible que $f(x, y)$ tome valores más pequeños a medida que nos alejamos de W . Se elige un punto R que se obtiene reflejando el punto W a través de la arista BG , entonces:

$$R = M + (M - W) = 2M - W. \quad (14)$$

Si $f(R) < f(W)$ entonces nos hemos movido en la dirección correcta hacia el mínimo, pero quizás, el mínimo este un poco más lejos del punto R por lo que extenderemos el segmento MR hasta el punto E de tal manera que MR sea congruente con RE , formado el triángulo BGE

$$E = R + (R - M) = 2R - M. \quad (15)$$

Si $f(E) < f(R)$ entonces encontramos un mejor vértice que R .

Si $f(R) = f(W)$ otro punto debe ser probado, quizás la función sea menor en M , pero no podemos reemplazar W con M porque debe mantenerse un triángulo, tomaremos dos puntos medios C_1 y C_2 tales que C_1 sea punto medio del segmento WM y C_2 punto medio del segmento MR , entonces:

$$C = \min(f(C_1), f(C_2)). \quad (16)$$

Si $f(C) \geq f(W)$ entonces los puntos G y W deben reducirse a través de B , el punto G reemplazará a M y W es reemplazado por S , siendo M punto medio de BG y S punto medio de WB . Un algoritmo eficiente solo hará comparaciones si son necesarias, en cada paso un nuevo vértice es encontrado que reemplazara a W . Una vez que se encuentra el mínimo no es necesario agregar más vértices y la iteración se termina. (Mathews & Fink, 1999)

5.4. Análisis de componentes principales

Debido a que la caracterización de los chirps se hace a través de una función que describa su forma y de ella se seleccionan 5 valores, es que la misma se hace en el espacio R^5 , aplicaremos el análisis de componentes principales (ACP) para reducir las dimensiones de los datos obtenidos a un espacio que permita trabajar más sencillamente con ellos, además de poder visualizarlos en un posterior análisis de conglomerado.

El ACP es una técnica factorial que se aplica a una tabla de datos cruzados de I individuos y J variables cuantitativas, formando así una matriz X_{IJ} . Cada observación está caracterizada por una fila en un espacio R^J ; cada variable será un vector de R^I . En este caso se trabaja con datos estandarizados para ello se les sustrae la media y se divide entre su desviación estándar. El objetivo del ACP es reducir el número de variables con las que se trabaja. Desde el punto de vista de las filas, busca encontrar individuos “parecidos” entre sí, o sea, busca tipologías de individuos. Desde el punto de vista de las columnas (variables), busca determinar la correlación existente entre ellas.

Para su aplicación es necesario definir distancias, métricas y pesos, ya sea para los individuos como para las variables. En el caso de los individuos, la métrica M utilizada es la matriz identidad y los pesos se representan en una matriz diagonal D , donde los elementos de la diagonal principal son $\frac{1}{I}$.

En el caso de las variables, cuando la matriz X_{IJ} está estandarizada, su métrica M coincide con la matriz de pesos D de los individuos y su matriz de pesos D coincide con M , la matriz identidad.

Cálculo de los componentes.

El primer componente principal será la combinación lineal de las variables originales que tenga varianza máxima. Los valores de este primer componente en los I individuos se representan con un vector \mathbf{z}_1 dado por:

$$\mathbf{z}_1 = X\mathbf{a}_1. \quad (17)$$

siendo \mathbf{a}_1 una dirección definida por un vector de norma unitaria.

Entonces:

$$Var(\mathbf{z}_1) = \mathbf{a}'_1 S \mathbf{a}_1. \quad (18)$$

donde S es la matriz de varianzas y covarianzas de las observaciones.

Para que la maximización de la ecuación (17) tenga solución se aplica la restricción de que $\mathbf{a}'_1 S \mathbf{a}_1$ y se obtiene la solución $S\mathbf{a}_1 = \lambda \mathbf{a}_1$ (Peña & Hill, 2002, pág. 141), que implica que \mathbf{a}_1 es un vector propio de la matriz S y λ su correspondiente valor propio, entonces la $Var(\mathbf{z}_1)$ es λ . Como se desea maximizar la varianza de \mathbf{z}_1 entonces λ es el mayor valor propio de la matriz S . Su vector propio asociado \mathbf{a}_1 define los coeficientes de cada variable en el primer componente. Análogamente se demuestra (Peña & Hill, 2002, págs. 144-145) que el espacio de dimensión r que mejor representa a los puntos está definido por los vectores propios asociados a los mayores valores propios de S . Una posible regla para determinar la cantidad de componentes a conservar, es quedarse con aquellas componentes cuyo valor propio sea mayores a 1 (Peña & Hill, 2002, pág. 159). Para determinar la calidad de representación de cada individuo X_i en las nuevas componentes, se mide si X_i es próximo a su proyección sobre el eje factorial a través del ángulo que forman la semirrectas $(0, X_i)$ y $(0, X'_i)$ siendo X'_i la proyección de X_i en el eje factorial. Cuando el valor del coseno cuadrado de dicho ángulo está cerca de 1 el ángulo es 0 o π y el individuo está bien representado en dicho eje factorial. Para determinar la calidad de representación de la variable X_j sobre el k -ésimo eje se calcula el $\cos^2(jk)$. Cuando este resultado está cerca de 1, entonces la variable j está bien representada en el k -ésimo componente.

5.5. Análisis de conglomerados

Con el objetivo de poder clasificar los chirps en grupos que tengan características en común es que se aplica un análisis de conglomerados (AC), dicha técnica tiene por objeto agrupar elementos en función de un conjunto de características. Normalmente se agrupan las observaciones (individuos), pero el AC puede también aplicarse para agrupar variables. La clasificación obtenida dependerá de las variables que se consideren. Esta técnica es de carácter exploratorio, es decir, describe una estructura de grupos en los que hay diferencias entre ellos, pero no es un análisis inferencial.

Para su aplicación es necesario definir criterios para unir individuos entre sí, individuos con grupos y grupos entre sí, lo que lleva implícito una definición de distancia. De esta forma los que están más “cerca” de acuerdo a algún criterio de distancia serán los que se unan.

Una vez definida la distancia entre individuos a utilizar, se requiere de un método que clasifique a los mismos. Los métodos de clasificación pueden ser jerárquicos o no jerárquicos.

Los métodos jerárquicos son métodos de particiones encajadas. Los grupos que se forman a un nivel de distancia comprenden grupos obtenidos a un nivel de distancia inferior. Los algoritmos utilizados en estos métodos pueden ser de agregación o de división. En los primeros, se parte de la situación en que cada individuo es un grupo. Se mide la distancia entre ellos (la que fue definida previamente), y se seleccionan las dos observaciones más próximas, las que pasarán a formar parte de un grupo. A partir de ahí, se debe definir cómo medir la distancia entre individuos y grupos y grupos entre sí. Una vez determinadas, se mide ahora la distancia de los demás elementos al grupo formado, y se unen aquellos que estén más “cerca” (puede ser que sean otros dos individuos para formar un nuevo grupo, o que un tercer individuo se una al grupo ya formado).

Para determinar la cantidad de conglomerados adecuada se utilizan tres indicadores:

1 – El R^2 establece la relación entre la variación explicada y la variación total, donde la variación explicada la representa la estructura de grupos hallada en cada nivel.

$$R^2 = 1 - \frac{\sum_{k=1}^K \sum_{i=1}^{n(k)} \sum_{j=1}^J (x_{ij} - \bar{x}_{kj})^2}{\sum_{i=1}^I \sum_{j=1}^J (x_{ij} - \bar{x}_{ij})^2}. \quad (19)$$

A mayor cantidad de conglomerados aumenta este indicador, por lo tanto se busca un valor de R^2 donde al aumentar la cantidad de conglomerados el aumento del R^2 no se considere “significativo”.

2 – El Pseudo³ F

$$\text{Pseudo } F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}}. \quad (20)$$

siendo k el número de conglomerados y n el número de observaciones. Un máximo local del pseudo F puede indicar una buena selección de conglomerados.

2 – El Pseudo⁴ t^2

Sean los conglomerados G y L que se unirán formando un nuevo conglomerado, este indicador es calculado como:

$$\text{Pseudo } t^2 = \frac{\text{tr}W_{GL} - (\text{tr}W_G + \text{tr}W_L)}{\frac{(\text{tr}W_G + \text{tr}W_L)}{n_G + n_L - 2}}. \quad (21)$$

³Se dice pseudo F porque no es una variable aleatoria con distribución F.

⁴Se dice pseudo t^2 porque no es una variable aleatoria con distribución t-student

siendo n_G y n_L el número de observaciones en los conglomerados G y L , trW_G y trW_L la traza de la matriz de variaciones intragrupos G y L , y trW_{GL} la traza de la matriz de variaciones que surge al unir los conglomerados G y L .

En el siguiente paso del algoritmo se evalúa el resultado de unir dos conglomerados, si el valor del indicador considerado es pequeño los conglomerados pueden ser combinados apropiadamente. Una regla para determinar la elección de la cantidad de conglomerados puede obtenerse al ordenar los valores de Pseudo t^2 según la cantidad de componentes, al obtenerse un valor marcadamente mayor al previo entonces se decide por conservar la cantidad de conglomerados del valor previo. (Peña & Hill, Análisis de datos multivariantes, 2002) (Blanco, 2006)

5.6. Prueba de modelos

Se plantea probar si la secuencia de pulsos de ecolocalización de los peces responde a un modelo donde solo depende del pulso anterior emitido por el propio pez, o si por el contrario responde a un modelo con interacciones, donde cada pez genera un corrimiento en sus pulsos dependiendo de la proximidad con el pulso del otro pez.

5.6.1. Prueba de hipótesis

“Consideramos dos modelos, H_0 y H_a que lo contiene, es decir, H_0 es el modelo obtenido para algunos valores particulares de los parámetros de H_a . Queremos probar la hipótesis nula de que el modelo H_0 es el adecuado para representar los datos X , contra la alternativa de que el modelo adecuado es H_a . Introducimos un estadístico $T(X)$ que cuando vale $H_a \setminus H_0$ es de esperar que sea menor que bajo H_0 . Si F_T es la distribución de T bajo H_0 , rechazaremos H_0 con nivel α si $T(X) < F^{-1}(\alpha)$. Si el estadístico está bien elegido, la potencia será buena, pero en cualquier caso, el nivel es α . Esto funciona de manera exacta si H_0 está completamente especificada, y como en general no se conoce la distribución de T bajo H_0 , se recurre a simular $T(X_j)$ para muestras X_j generadas con la distribución que especifica H_0 . Cuando H_0 no está completamente especificada, porque quedan parámetros por estimar, entonces estimamos los parámetros y generamos las X_j con la distribución del modelo H_0 con los parámetros estimados. Esto da lugar a una prueba aproximada. Si la estimación se hiciera de manera de minimizar el estadístico T , la prueba resultante será conservativa. En general no será conservativa pero será al menos aproximada.” (Cabaña)

5.6.2. Prueba de ajuste de modelos de interacción en el presente estudio

Se plantea probar si la secuencia de pulsos responde únicamente al pez emisor, sin tomar en cuenta el momento en que ha emitido un pulso el otro pez o si por el contrario, la proximidad en el tiempo de la emisión de pulsos de ambos peces genera una modificación en la frecuencia de emisión en cada uno de ellos.

$$H_0 : x_t = x_{t-1} + \beta_1 + \delta\epsilon_t. \quad (22)$$

$$H_a : x_t = x_{t-1} + \beta_1 + \beta_2 I(x_{t-1} - y_{t-1} < \gamma) + \beta_3 I(y_{t-1} - x_{t-1} < \gamma) + \beta_4 I(x_t - y_{t-1} < \gamma) + \delta\epsilon_t \quad (23)$$

Entendemos como:

β_2 el corrimiento dado por que las señales de cada pez en el momento anterior distan menos de γ siendo x_{t-1} la última señal.

β_3 el corrimiento dado por que las señales de cada pez en el momento anterior distan menos de γ siendo y_{t-1} la última señal.

β_4 el corrimiento dado por que la señal a pulsar por un pez, es demasiado próxima (menor que γ) a la última pulsada del otro.

Se verificará si el error de aplicar los datos reales de los instantes de los pulsos en el modelo con interacciones, está en el percentil 5% menor respecto a un conjunto de pulsos simulados con el modelo sin interacciones. Si esto ocurre se rechazaría que los pulsos “reales” puedan provenir de un modelo sin interacciones.

Se aplican los datos reales para estimar β_1 en cada pez, mediante mínimos cuadrados en el modelo H_0 , y luego se estima δ a través del desvío s de los datos de la muestra,

$$s = \sqrt{\frac{SCR}{n-1}} = \sqrt{\frac{\sum(\hat{x}_t - x_t)^2}{n-1}}. \quad (24)$$

Estimados los parámetros en la ecuación (22) para cada uno de los dos peces, y tomando como punto de partida la primera observación de cada pez, se generan secuencias simuladas de x e y con la misma cantidad de datos que la secuencia de datos reales. Se toma un punto de partida para x_1 arbitrario, se

simula que se asumió que se distribuye $N(0,1)$, y se obtiene x_2 , se repite el procedimiento hasta obtener todos los valores deseados de x_t , análogamente se obtiene la secuencia para y_t .

Se estima entonces el error empírico, al aplicar cada secuencia simulada bajo H_0 en el modelo H_a , puesto que este será el estadístico $T(X)$.

$$\hat{x}_t = x_{t-1} + \hat{\beta}_1 + \hat{\beta}_2 I(x_{t-1} - y_{t-1} < \gamma) + \hat{\beta}_3 I(y_{t-1} - x_{t-1} < \gamma) + \hat{\beta}_4 I(\hat{x}_t - y_{t-1} < \gamma). \quad (25)$$

$$Error = \sum (\hat{x}_t - x_t)^2. \quad (26)$$

Luego se ajusta el modelo bajo H_a a los datos reales (que no se sabe a qué modelo responden), si el error es mucho menor que los calculados a partir de la simulación generada con H_0 entonces estaremos en condiciones de rechazar H_0 .

Se repite el procedimiento de simulación de la secuencia de señales 100, 200 y 1000 veces para obtener el error y determinar que percentil ocupa el valor del error real si se lo compara con los errores de simulación. Finalmente se verifica si el resultado es consistente al variar la cantidad de veces que se realiza la simulación, de esta manera podemos garantizar la estabilidad de la metodología empleada.

5.7. Simulación

“La simulación es una técnica numérica que permite conducir experimentos en computadora para lo cual se necesitan modelos lógicos matemáticos que describan la conducta del negocio, fenómeno o realidad que se intenta reproducir, a través de periodos de tiempo real” (Rubistein, 1961).

5.7.1. Método de la función inversa

Suponga $U \sim U(0,1)$ y se desea simular una variable aleatoria continua X con función de distribución $F_X(X)$, sea $Y = F_X^{-1}(U)$ entonces:

$$F_Y(y) = P(Y \leq y) = P(F_X^{-1}(U) \leq y) = P(U \leq F_X(y)) = F_X(y)$$

Entonces Y tiene la misma distribución que X .

La utilidad de este proceso es que a partir de generación de variables aleatorias uniforme (0,1) (obtenibles con cualquier sistema informático, R en particular utiliza el generador congruencial Mersenne-Twister) y si existe la función inversa de $F_X(x)$ se pueden generar variables aleatorias con cualquier distribución utilizando la transformación en probabilidad. En caso de no existir la función inversa se pueden aplicar métodos numéricos para obtener datos simulados de dicha variable aleatoria. (Jones, Maillardet, & Robinson, 2014)

Este es el método usado por defecto en el R (R Documentation, s.f.) En el contexto de este trabajo, se utilizan técnicas de simulación para la generación de variables aleatorias con distribución normal utilizadas para modelar ϵ_t en la ecuación (22).

6. RESULTADOS

6.1. Detección de pulsos y chirps

En la Figura 5 puede apreciarse el resultado del algoritmo de detección de pulsos, se obtiene una sucesión de tiempos t_i candidatos a ser pulsos. En esta figura para una mejor visualización se superponen puntos rojos en el eje del tiempo donde mediante las desigualdades 3 y 4⁵ se detectan pulsos, con gráfico original para observar el resultado obtenido.

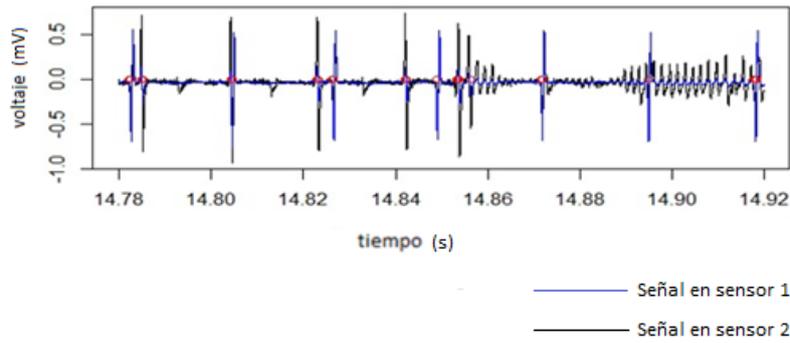


Figura 5: Detección de pulsos
Fuente: Propia

En la Figura 6 se grafican los valores obtenidos en la acumulación de potencia (p'') para cada momento del tiempo. Estos valores se grafican en color rojo y se superponen con la función $p(t)$ en color negro. Los valores donde la acumulación de potencia no supera el umbral⁶ determinado por la fórmula (8), se transforman en 0 de modo que cuando esta toma valores superiores a 0 es porque se ha cruzado el umbral y se ha detectado que en esos momentos del tiempo hay un chirp. Podemos apreciar que si bien puede detectarse el chirp con el algoritmo, el mismo no es preciso en indicar el comienzo y el final del mismo.

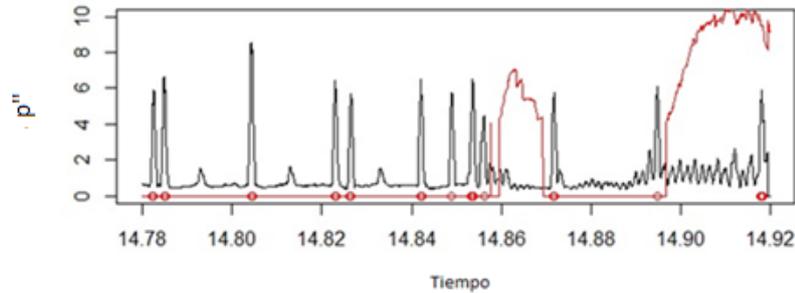


Figura 6: Detección de pulsos y chirps mediante función de potencia.
Fuente: Propia

⁵Para detectar pulsos, se tomó un entorno de 10 unidades de tiempo y un $u = 3$

⁶Heurísticamente se eligió como mejor opción $\gamma = 4$, $a = 5$ y se calculó la acumulación de potencia dado un momento t cualquiera.

6.2. Separación de señales

Aplicando el algoritmo detallado en las ecuaciones (9), (10), (11) y (12) para separar las señales de cada uno de los dos peces se obtienen los resultados graficados en la Figura 7. Los pulsos de un pez quedan marcados con el círculo de color blanco y los del otro con un punto de color verde. Nuevamente el programa entrega como resultado el tiempo y el valor del voltaje en cada uno de los sensores, permitiendo calcular la distancia (gap) entre pulsos consecutivos del mismo pez.

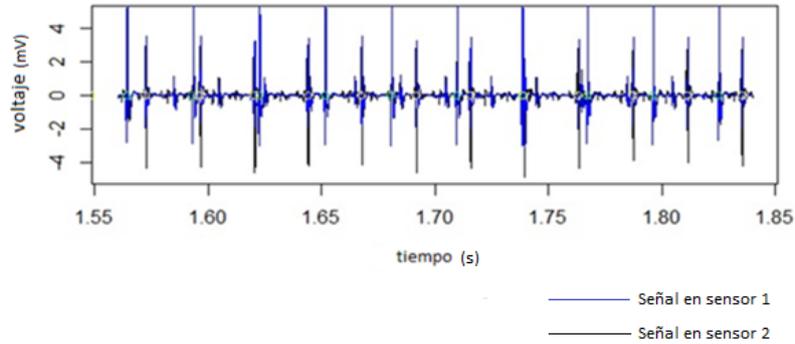


Figura 7: Clasificación de señales según pez emisor
Fuente: Propia

6.3. Ajuste de curva

Se aplica el ajuste planteado mediante optimización con el algoritmo de Nelder Mead implementado en R (R core team, 2015), a los 93 chirps que fueron aislados de los datos brindados por los investigadores del Clemente Estable. De cada uno de esos 93 chirps se obtuvieron dos, cada uno de ellos obtenido a partir de la señal del sensor 1 y el sensor 2. Se obtuvo así un ajuste para cada uno de esos 186 chirp. Las Figuras 8, 9 y 10 muestran tres ejemplos de chirps en sus valores positivos de voltaje relativo y el ajuste que se hizo en cada uno de ellos mediante un polinomio $f(t)$ de 2do grado. En las Figuras 8 y 9 se aprecia que el ajuste podría ser adecuado, pero la observación de la Figura 10 lleva a desechar este tipo de ajuste. Se debe aclarar que en todos casos la forma de determinar si el ajuste es apropiado ha sido visual.

6.3.1. Parábola

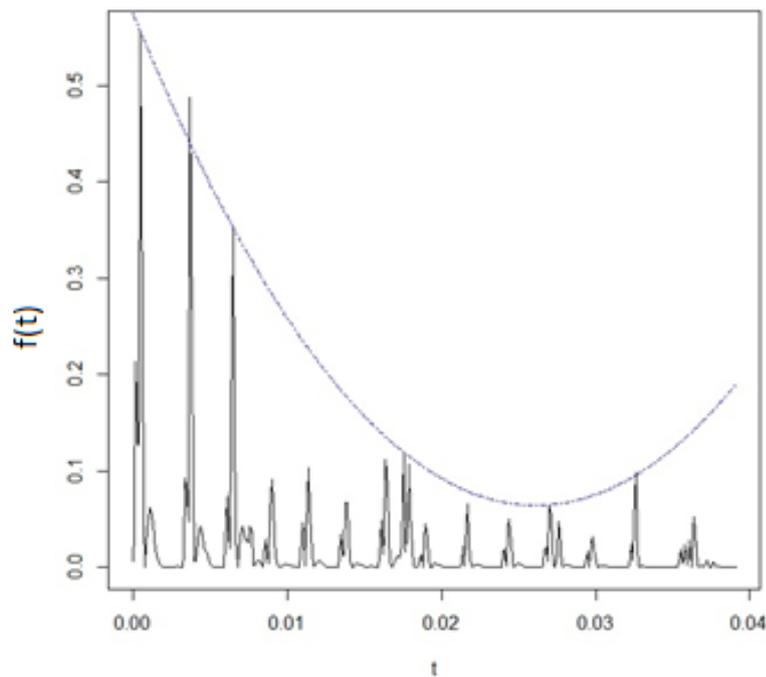


Figura 8: Ajuste de los chirps mediante una parábola
Fuente: Propia

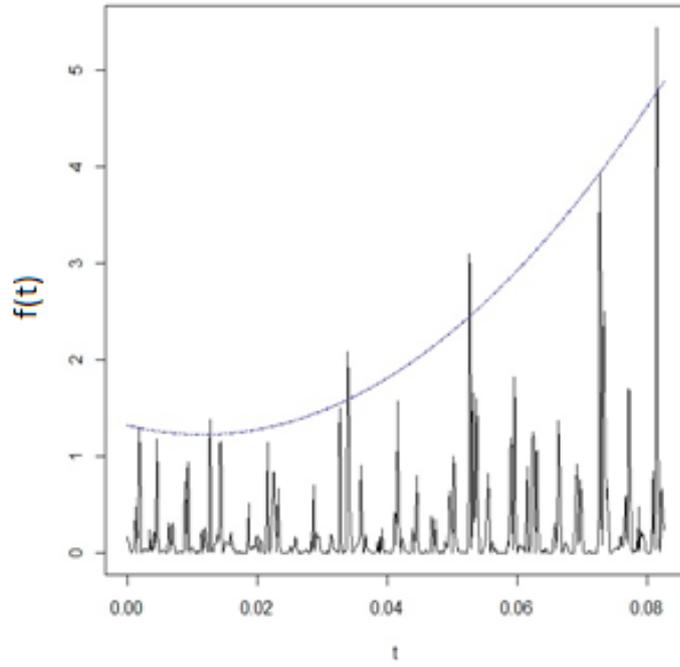


Figura 9: Ajuste de los chirps mediante una parábola
Fuente: Propia

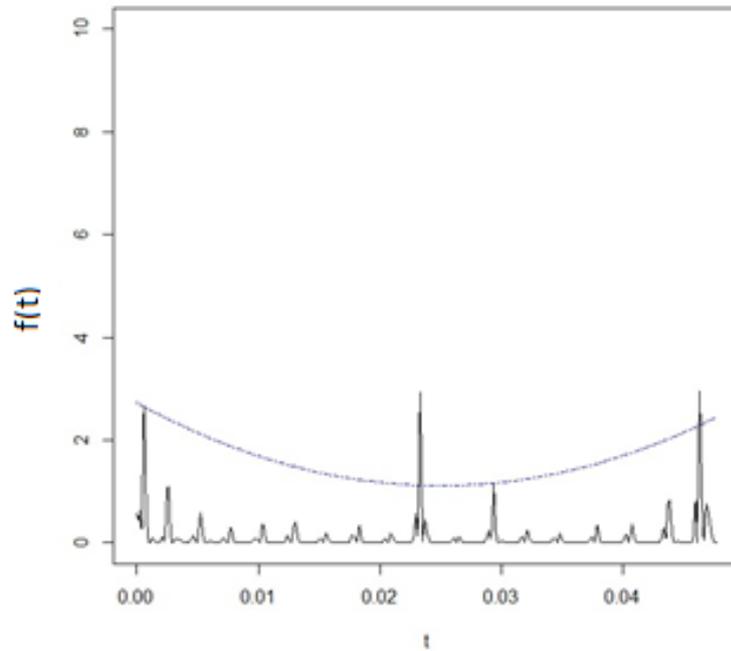


Figura 10: Ajuste de los chirps mediante una parábola
Fuente: Propia

6.3.2. Polinomio de grado 4

Como siguiente opción se pasa a ajustar los chirps por un polinomio $f(t)$ de grado 4, puede apreciarse un ejemplo de ajuste en la Figura 11. En el Apéndice I pueden verse los ajustes de todos los chirps, al no encontrar visualmente una cantidad significativa de resultados que indiquen desechar este ajuste se opta por modelar los chirps a través de polinomios de grado 4.

Decidido este modelo como adecuado, se procede a tomar cinco variables para posteriormente realizar un análisis de componentes principales, para ello se seleccionaron los valores del polinomio en cinco momentos de la duración del mismo: al comienzo, al momento del 25%, 50%, 75% de su duración y finalmente su valor al momento de terminar.

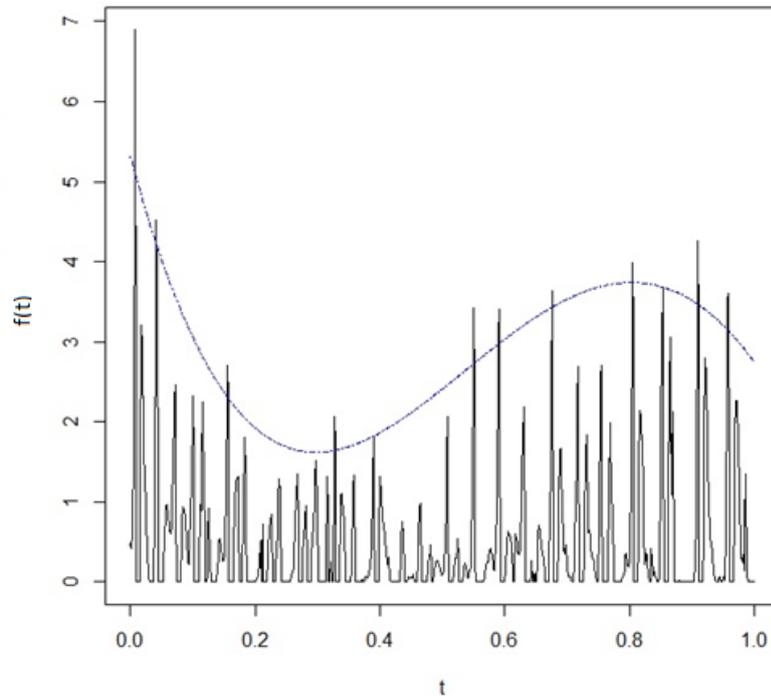


Figura 11: Ajuste de los chirps mediante un polinomio de 4to grado

Fuente: Propia

6.4. Análisis de componentes principales

Antes de aplicar el método de componentes principales, se toman como variables las indicadas en el marco teórico, con estas variables se obtienen los siguientes nuevos componentes y su vinculación con las variables. Sean:

Primero: el valor de la intensidad al comienzo del chirp.

Segundo: el valor de la intensidad en el 25 % de la duración del mismo.

Tercero: el valor de la intensidad en el 50 % de la duración del mismo.

Cuarto: el valor de la intensidad en el 75 % de la duración del mismo.

Quinto: el valor de la intensidad al final del chirp.

El cuadro 1 muestra cómo se relacionan las variables originales en las nuevas componentes (por cuestiones de practicidad solo se muestran las primeras dos componentes que son las que se conservarán para el análisis posterior), se aprecia allí el coseno entre la variable original y la componente nueva. Se toma como criterio que para que una variable esté bien representada en el nuevo sistema de ejes que el coseno al cuadrado no sea menor a 0,5. Optando por dos componentes todas las variables originales quedan bien representadas.

Cuadro 1: Cosenos y cosenos al cuadrado

Variable	cos		cos ²	
	Componente 1	Componente 2	Componente 1	Componente 2
Primero	0,797	-0,576	0,635	0,332
Segundo	-0,446	0,894	0,199	0,799
Tercero	-0,678	-0,707	0,460	0,500
Cuarto	0,946	0,284	0,895	0,082
Quinto	-0,904	-0,115	0,817	0,013

El cuadro 2 muestra los valores propios de cada componente, por regla general se considera apropiado decidir conservar las componentes con valores propios superiores a 1.

Cuadro 2: Valores propios según cantidad de componentes

Componente 1	Componente 2	Componente 3	Componente 4	Componente 5
2,335	1,229	0,732	0,679	0,034

Cuadro 3: Proporción de las varianzas según cantidad de componentes

	Componente 1	Componente 2	Componente 3	Componente 4	Componente 5
Desviación Estándar	1.528	1.109	0.855	0.818	0.185
Proporción de varianza	0.467	0.246	0.146	0.134	0.007
Varianza acumulada	0.467	0.713	0.859	0.993	1.000

Cuadro 4: Coordenadas de las variables originales en los nuevos componentes

Variable original	Componentes	
	1ra	2da
Primero	0,203	0,696
Segundo	-0,514	0,133
Tercero	-0,638	0,148
Cuarto	-0,487	0,252
Quinto	-0,225	-0,642

El cuadro 4 nos permite determinar las coordenadas de los individuos en las nuevas componentes de la siguiente manera:

Coordenada en componente 1 = $0,203 V_1 - 0,514 V_2 - 0,638 V_3 - 0,487 V_4 - 0,225 V_5$

Coordenada en componente 2 = $0,696 V_1 + 0,133 V_2 + 0,148 V_3 + 0,252 V_4 - 0,642 V_5$

Siendo V_i el valor de la i -ésima variable original.

Analizando el cuadro 4 se puede concluir que en la primera componente toman valores más grandes aquellos chirps que son grandes en su amplitud al principio y disminuyen su tamaño ya en el segundo quintil de la duración del chirp y toman valores más chicos aquellos que toman valores altos a partir del segundo quintil. Puede deducirse entonces que esta componente separa a aquellos chirps que son grandes en amplitud en el primer quintil de duración de aquellos que son grandes a partir del segundo quintil de duración. En la segunda componente toman valores grandes aquellos chirps que son grandes en amplitud en los primeros 4 quintiles de duración, en particular en el primer quintil, y toman valores pequeños aquellos chirps con valores grandes en el último quintil de duración. Puede considerarse que la segunda componente diferencia aquellos chirps que son grandes al principio y chicos al final de aquellos que son grandes al final y chicos al principio.

Cuadro 5: Coordenadas de las variables originales en los nuevos componentes

Chirp	1ra	2da	Chirp	1ra	2da	Chirp	1ra	2da
1	-0,807	1,628	61	-1,537	0,473	127	-0,680	1,264
3	0,762	0,632	62	1,693	0,733	128	1,080	0,182
4	0,340	-1,673	63	3,963	0,376	129	1,206	1,214
5	1,525	0,218	64	3,457	0,613	130	1,803	-2,238
6	-2,582	0,071	72	1,949	-2,038	133	0,439	0,468
7	0,931	0,600	75	2,316	-0,849	135	-0,030	1,534
8	0,475	-1,943	76	0,268	-1,332	136	0,302	0,975
11	-0,883	0,890	77	-1,298	0,712	137	-0,107	1,619
13	-0,526	-2,238	78	-0,695	0,355	138	0,203	1,193
14	0,913	0,147	79	-1,188	1,821	139	-1,739	-0,726
15	-2,152	0,195	80	0,276	0,623	142	-1,470	0,428
16	0,066	-2,296	81	-1,334	0,288	144	1,161	0,692
17	-0,892	-1,058	82	-1,621	-1,295	145	0,671	1,505
18	-1,663	-1,045	83	-1,267	0,457	146	0,335	1,736
19	-2,851	0,101	85	-2,387	-0,465	147	-1,566	1,397
21	-0,764	-1,750	86	-1,924	-0,093	148	2,682	0,736
22	-0,111	-1,523	87	1,147	0,762	150	-1,177	-0,439
23	-1,333	1,116	88	0,886	0,958	151	2,838	0,720
25	1,581	1,288	90	0,612	-1,628	152	2,583	0,744
26	-1,398	0,591	91	2,467	0,996	153	0,487	0,617
27	0,739	1,376	92	1,745	0,620	155	2,551	-1,040
28	0,224	0,726	93	2,292	0,570	156	2,408	-0,739
29	1,587	0,023	94	2,507	0,896	157	-1,713	0,989
31	-1,471	-0,140	95	-1,776	1,508	159	1,727	-0,434
32	-0,856	0,568	96	-2,508	0,885	162	2,754	-0,558
34	-2,711	0,476	97	-2,321	-0,092	164	0,708	-3,203
35	2,561	-1,277	99	-1,973	0,176	166	1,320	-2,998
38	2,197	0,460	100	-1,685	0,426	168	1,632	0,917
39	4,135	0,393	101	-1,770	-2,930	170	1,771	0,308
40	3,920	0,441	102	-0,307	-3,222	171	-1,993	0,627
42	-1,323	-0,009	104	-2,277	0,047	172	2,682	0,383
44	-1,894	1,158	105	2,465	-0,578	173	-0,843	1,655
45	0,913	1,177	106	-2,183	0,108	174	3,545	0,286
46	0,985	0,780	108	-1,381	0,374	175	-0,288	-1,848
47	-1,088	-0,375	109	-2,637	0,008	177	-1,147	0,063
48	-1,755	0,454	111	-2,198	-0,293	178	-2,189	1,030
50	2,015	-0,192	112	0,831	0,221	179	1,092	1,362
51	-2,735	-0,770	113	0,047	1,851	181	2,901	-0,057
52	-0,904	1,207	114	1,843	0,823	182	1,465	-0,113
53	0,104	-2,311	116	0,230	-1,510	183	0,121	-3,067
56	-0,079	0,633	121	-1,313	-0,574	186	2,355	0,099
57	0,701	-2,285	123	1,076	-0,193			
60	-0,313	-3,247	126	0,565	-3,730			

El cuadro 5 muestra las coordenadas de cada chirp en el nuevo eje de coordenadas. Se han eliminado aquellos chirps que han quedado mal representados, empleando también el criterio del coseno.

6.5. Análisis de conglomerados

Cuadro 6: Valores de R^2 , $Pseudo F$ y $Pseudo t^2$ obtenidos por el algoritmo

Cantidad de Conglomerados	R^2	$Pseudo F$	$Pseudo t^2$
13	0.941	152.63	16.04
12	0.935	150.67	24.96
11	0.928	149.78	23.84
10	0.921	151.45	35.9
9	0.909	147.65	23.89
8	0.895	145.39	24.04
7	0.875	140.38	22.68
6	0.847	134.4	34.09
5	0.819	137.69	41.71
4	0.776	142.63	28.33
3	0.661	121	84.87
2	0.386	78.851	118.9
1	0	NA	78.851

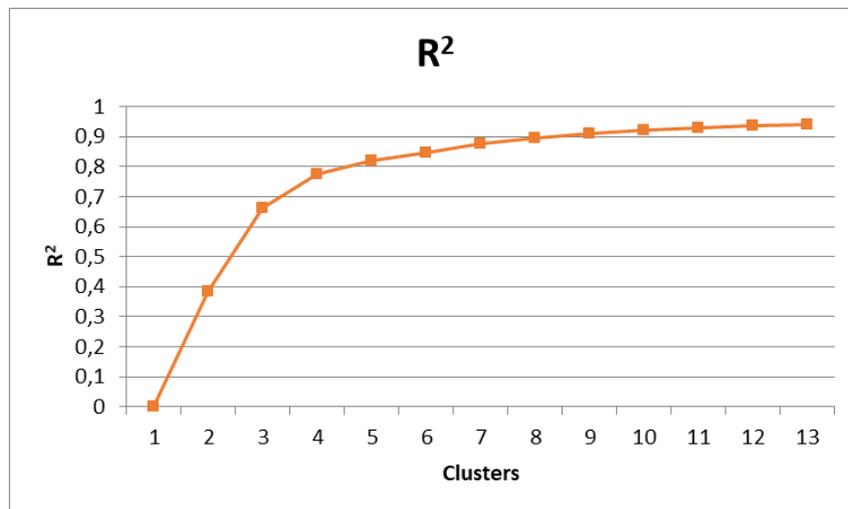


Figura 12: Valores del R^2 según la cantidad de conglomerados elegidos
Fuente: Propia

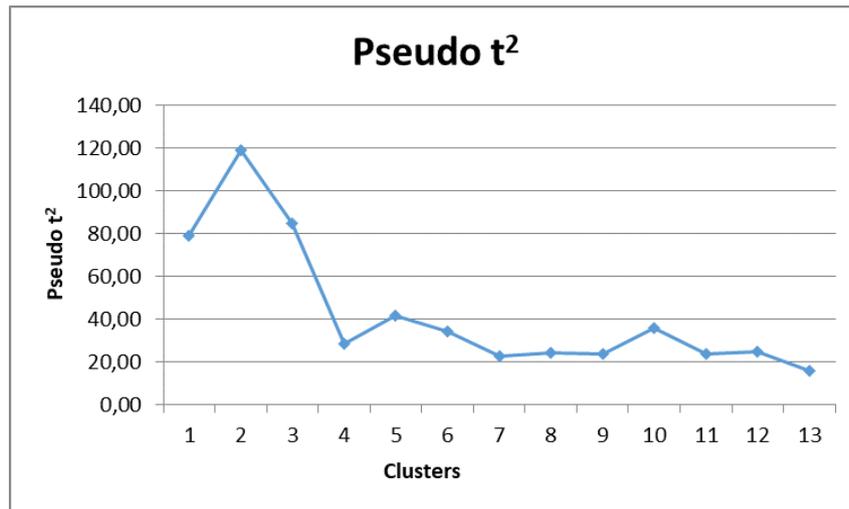


Figura 13: Valores del Pseudo t^2 según la cantidad de conglomerados elegidos.
Fuente: Propia

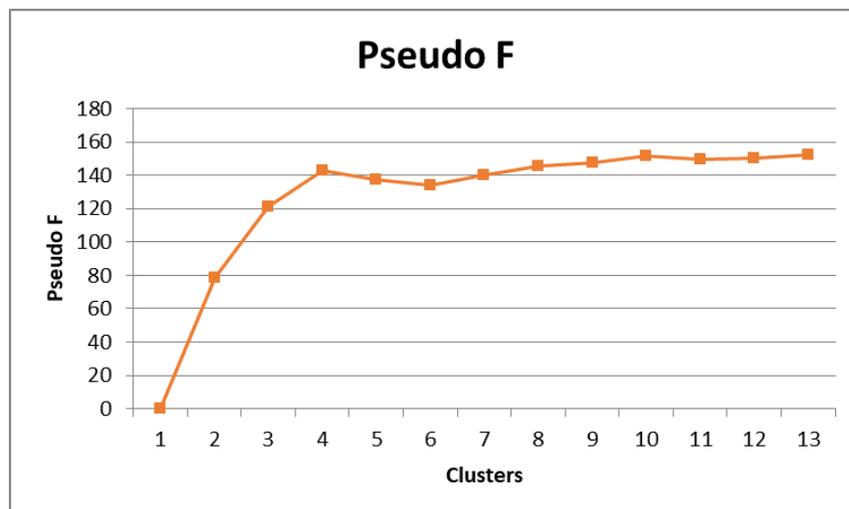


Figura 14: Valores del Pseudo F según la cantidad de conglomerados elegidos
Fuente: Propia

Dado que se aprecia en el cuadro 6 y la Figura 13 que se cumple que hay una caída en el pseudo t^2 y en la Figura 14 que hay un máximo local del pseudo F , la decisión es de determinar la existencia de 4 conglomerados de chirps, los cuales graficados según las dos componentes utilizadas se distribuyen de la siguiente manera:

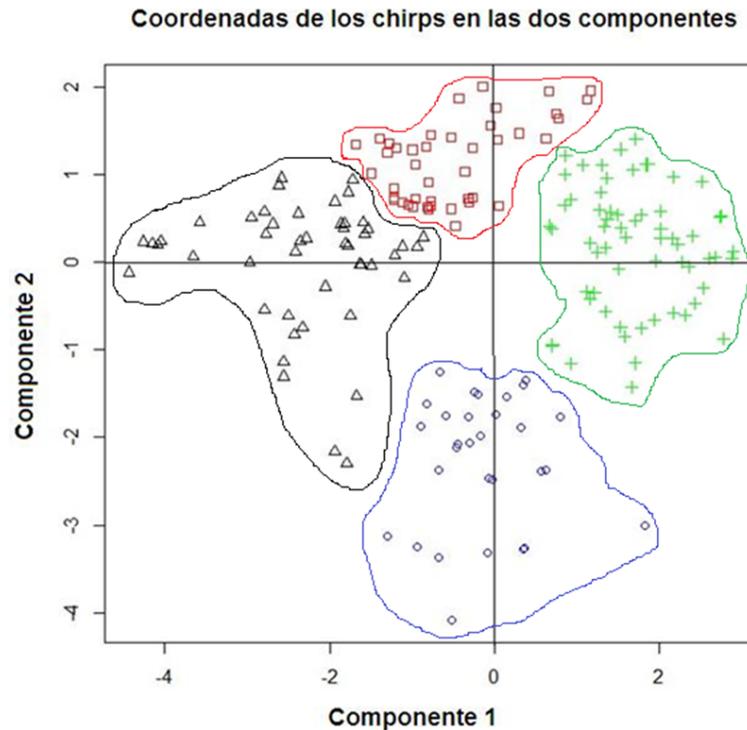


Figura 15: Resultado del análisis de conglomerados
Fuente: Propia

Utilizando la interpretación previamente hecha sobre el significado de cada componente podemos concluir que se ha obtenido una estructura de grupos con estas características:

1) La existencia de dos grupos de chirps con valores negativos y positivos respectivamente según la primera componente, aquellos con valores positivos presentan valores grandes en la primera variable (comienzo del chirp) y bajos en las siguientes, lo cual los catalogaría como chirps que son grandes en su amplitud al comienzo y más pequeños en el resto de su duración, por oposición. Los chirps que toman valores negativos en esta componente, son chirps que son pequeños en su amplitud al comienzo y más grandes a partir del 25 % de su duración.

2) La existencia de otros dos grupos de chirps, los cuales toman valores negativos y positivos respectivamente en la segunda componente. Los que toman valores positivos son chirps que son grandes en su amplitud al comienzo y pequeños al final, por el contrario, aquellos que toman valores negativos son chirps pequeños en su amplitud al comienzo y grandes al final.

En la sección anteriores se presentaron los resultados obtenidos en (Macadar & Silva, 2007), utilizando el análisis de conglomerados con distintas variables a las utilizadas en esta pasantía. Una diferencia sustancial es que en esa investigación se tomó en cuenta como elementos de clasificación la duración del chirp y el aumento de la frecuencia, mientras que este trabajo se concentró exclusivamente en la forma del chirp. No obstante estas diferencias se destaca que el chirp “B” en (Macadar & Silva, 2007), ilustrado en la Figura 4, coincide con uno de los grupos de chirps detectados en este trabajo, puesto que presenta una estructura donde la amplitud es grande al comienzo y decrece rápidamente, siendo pequeño ya a partir del 25 % de su duración.

6.6. Ajuste de modelos

Aplicamos los datos reales para estimar β_1 , mediante mínimos cuadrados en el modelo H_0 , y luego estimamos δ asumiendo que ϵ se distribuye $N(0,1)$.

Estimaciones usando los datos reales:

$\hat{\beta}_1$ para x_t es 0.029

$\hat{\beta}_1$ para y_t es 0.024

$\hat{\delta}_x=0,000197$

$\hat{\delta}_y=0,000034$

Se obtiene el error calculado con los datos reales en el modelo H_a y se lo ordena con los valores de los errores obtenidos con los datos simulados (considerando cierto H_0) aplicados a H_a . Se realizaron 100, 200 y 1000 simulaciones para verificar si la conclusión de rechazar o no rechazar H_0 variaba en estos tres casos, el cuadro 7 muestra los resultados obtenidos al aplicar el procedimiento de simulación y la ubicación del error expresado como percentil. En todos los casos el criterio de tomar un 5 % como criterio de rechazo se verifica por lo que aceptamos el modelo con interacciones como correcto.

Cuadro 7: Percentiles del error de los datos reales

Error Real en H_a respecto al error de los datos simulados	Cantidad de simulaciones		
	100	200	1000
menor al percentil 5 %	100 %	100 %	100 %
menor al percentil 4 %	100 %	100 %	100 %
menor al percentil 3 %	100 %	99,5 %	99,8 %
menor al percentil 2 %	99 %	92 %	97,3 %
menor al percentil 1 %	38 %	31,5 %	36,4 %

7. CONCLUSIONES

La presente pasantía tenía como primer objetivo el detectar automáticamente los pulsos para la ecolocalización como las señales sociales denominadas chirps. En ese aspecto y mediante la definición de funciones basadas en el voltaje acumulado durante un pequeño intervalo de tiempo, los pulsos lograron ser detectados e identificados correctamente. En el caso de los chirps la función detecta momentos del tiempo donde existen chirps, pero tiene como debilidad que deben ajustarse empíricamente en cada conjunto de datos los umbrales en los cuales detecta la presencia de un chirp. Se obtuvieron resultados de chirps detectados como tales que no lo eran, tampoco fue posible determinar con exactitud el comienzo y fin de cada chirp. Respecto a los chirps detectados se ajustó un polinomio de grado 4 para ser utilizado como una función que se ajusta por los puntos máximos en la intensidad durante la duración del chirp. De este polinomio se tomaron los valores del mismo en el comienzo, en el 25 %, en el 50 %, en el 75 % de la duración del chirp y al final del mismo. Al realizar un análisis de componentes principales para poder ver los datos en una dimensión menor se obtuvieron dos componentes, la primera de ellas toma valores mayores cuando la primera variable (intensidad al comienzo del chirp) toma valores mayores, y toma valores menores con un aumento de cualquiera de las otras variables. La segunda componente toma valores mayores cuando la primera variable es grande y toma valores menores cuando la última componente es grande. Al hacer un análisis de conglomerado con las coordenadas de los peces en el plano determinado por las dos componentes principales se obtuvieron cuatro grupos, los cuales coincidieron con estar ubicados en valores positivos o negativos sobre cada una de las dos componentes obteniéndose la siguiente clasificación:

- Un grupo de chirps con amplitudes grandes al comienzo y pequeñas en el resto de la duración del chirp.
- Un grupo de chirps con amplitudes pequeñas al comienzo y grandes en el resto de la duración del chirp.
- Un grupo de chirps con amplitudes grandes al comienzo y pequeñas al final de la duración del chirp.
- Un grupo de chirps con amplitudes pequeñas al comienzo y grandes al final de la duración del chirp.

En cuanto al análisis de las señales de ecolocalización, mediante técnicas de simulación se pudo probar, con un margen del 5 % de error, que los peces hacen una pequeña variación en la frecuencia de emisión de los pulsos en caso de que los pulsos de uno y otro pez se encuentren muy próximos uno del otro en el tiempo.

8. PERSPECTIVAS A FUTURO

En la presente pasantía se realizaron procedimientos que tuvieron como objetivo la detección automática de los pulsos y los chirps, además de la discriminación de los pulsos según cuál fuera el pez emisor. Estos procedimientos incluyeron una cierta cantidad de parámetros que debieron determinarse por ensayo y error, además de que no fueron únicos en las distintas secuencias, ni a lo largo del tiempo para una misma secuencia. Si bien esto se puede considerar como una dificultad a la hora del procesamiento de la información, es un avance respecto a la búsqueda manual de los eventos de interés, en particular de la ocurrencia de chirps. Quedan abiertas entonces líneas de investigación que permitan la detección automática de chirps, perfeccionando los procedimientos aplicados en esta pasantía mediante algún algoritmo de aprendizaje y determinación de los parámetros en forma automática.

9. APÉNDICE I : Ajustes de chirps

Todos los chirps detectados ajustados por un polinomio de grado 4.

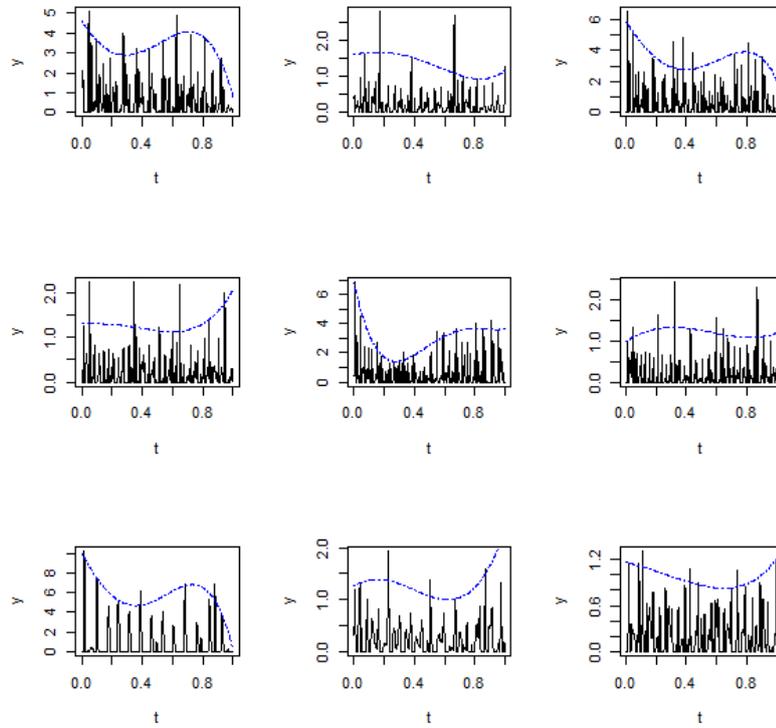


Figura 16: Ajuste por polinomio de grado 4
Fuente: Propia

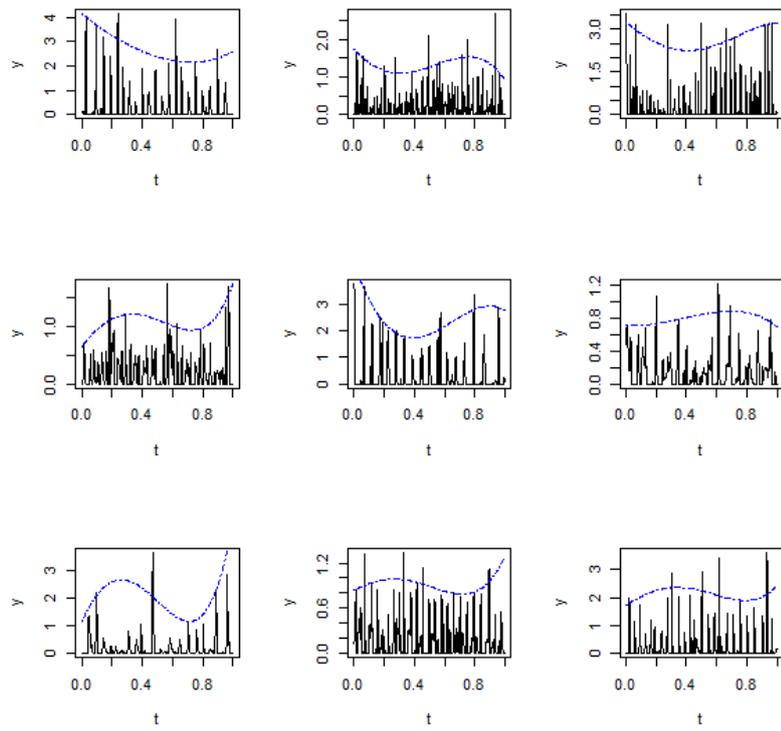


Figura 17: Ajuste por polinomio de grado 4
Fuente: Propia

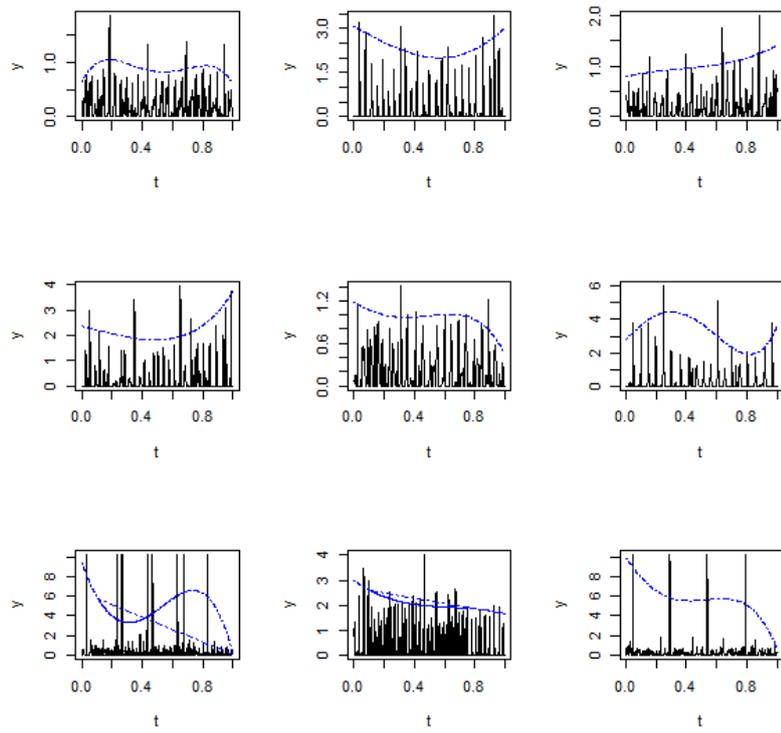


Figura 18: Ajuste por polinomio de grado 4
Fuente: Propia

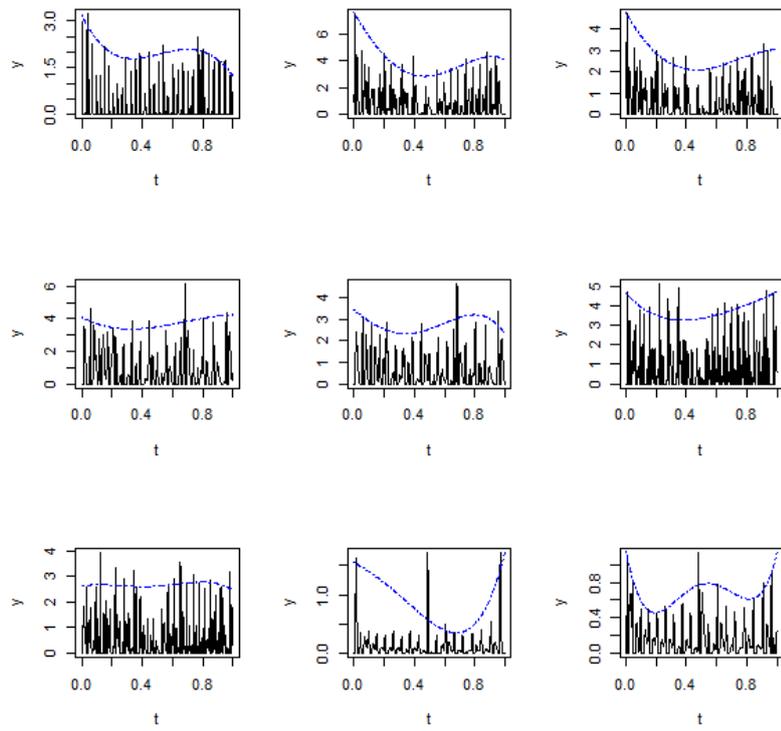


Figura 19: Ajuste por polinomio de grado 4
Fuente: Propia

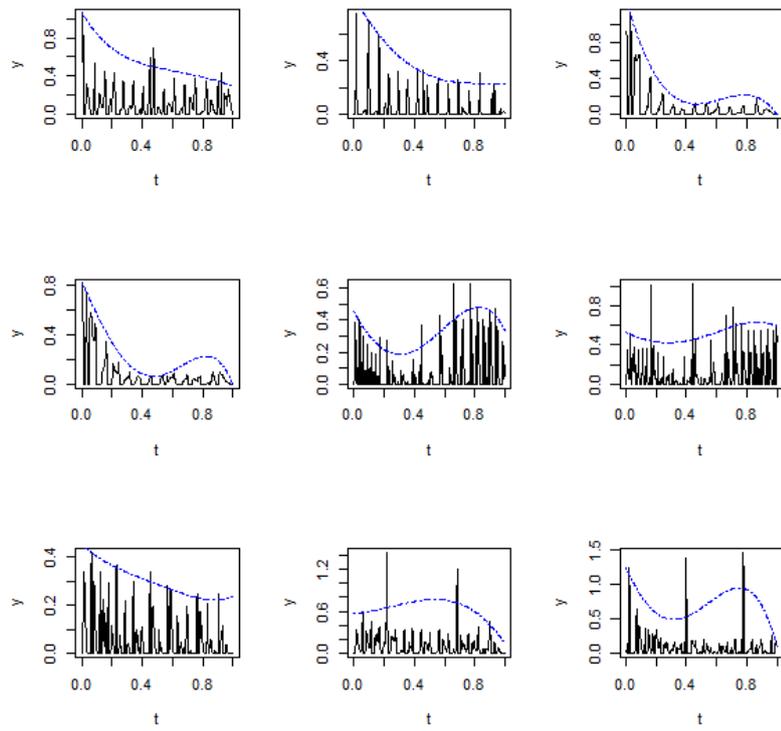


Figura 20: Ajuste por polinomio de grado 4
Fuente: Propia

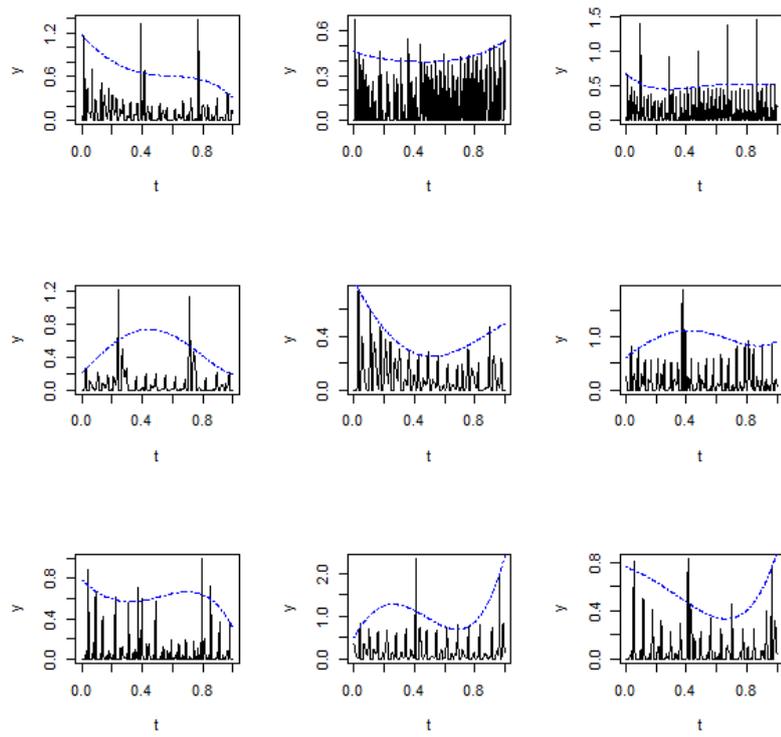


Figura 21: Ajuste por polinomio de grado 4
Fuente: Propia

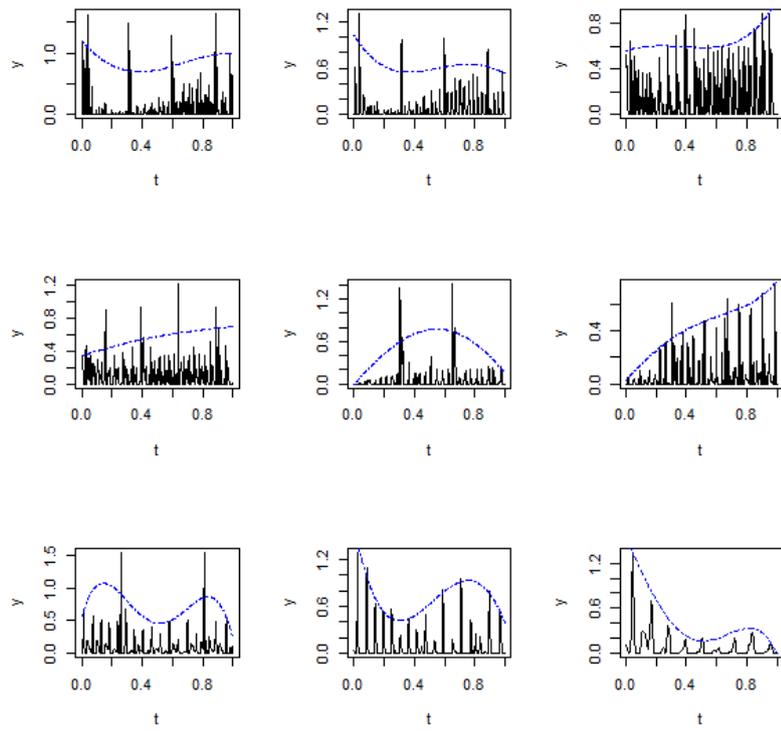


Figura 22: Ajuste por polinomio de grado 4
Fuente: Propia

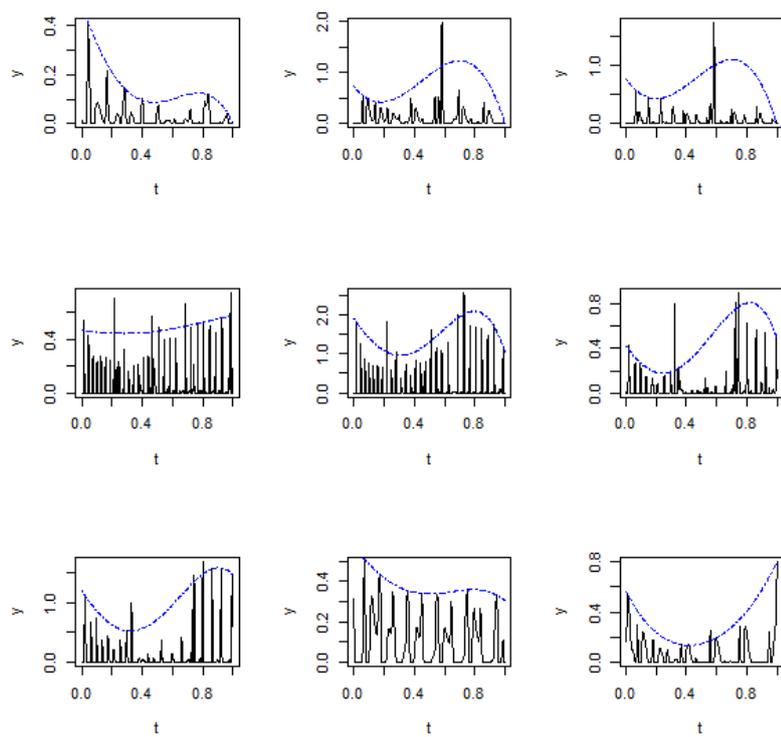


Figura 23: Ajuste por polinomio de grado 4
Fuente: Propia

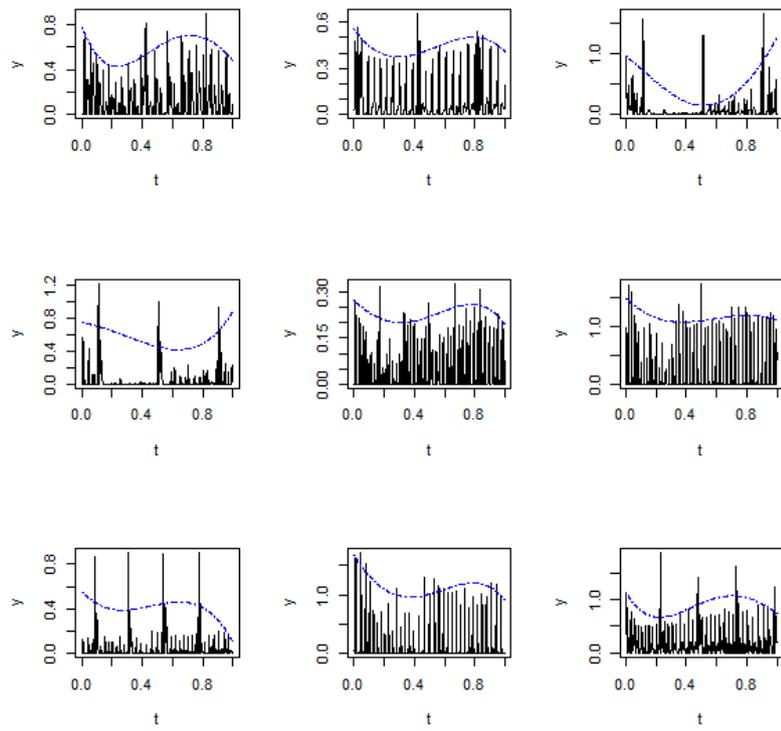


Figura 24: Ajuste por polinomio de grado 4
Fuente: Propia

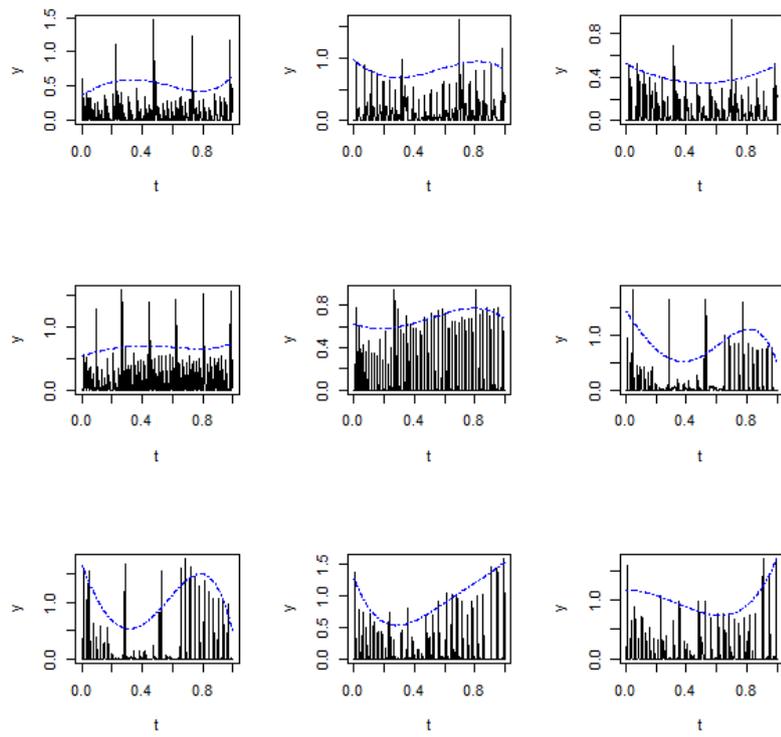


Figura 25: Ajuste por polinomio de grado 4
Fuente: Propia

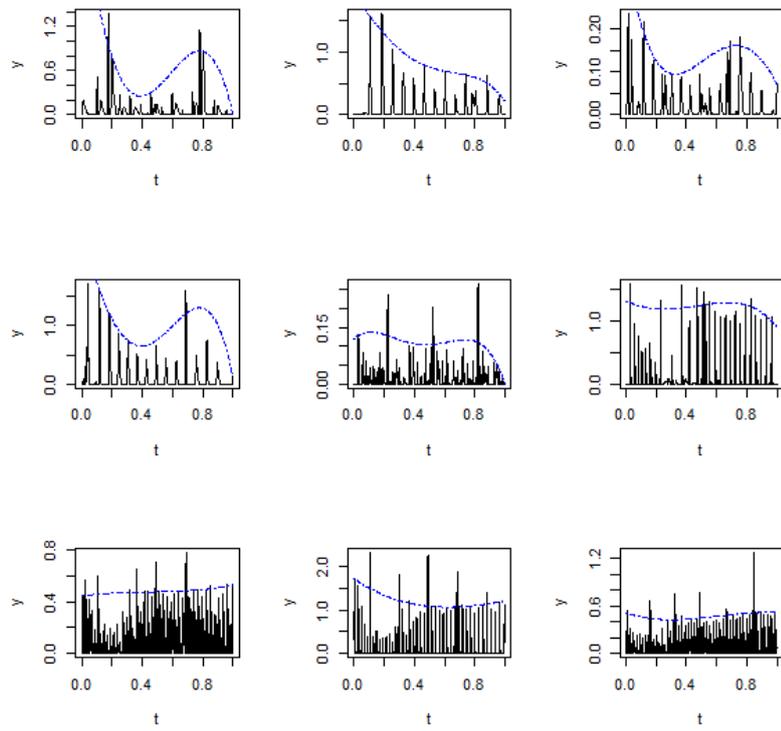


Figura 26: Ajuste por polinomio de grado 4
Fuente: Propia

10. APÉNDICE II : PROGRAMAS

10.1. Detección pulsos y chirps

```
lectura=function(arch,salt,laps){
long=20000*laps
pec=scan(file=arch,skip=salt*20000,dec='.',nlines=long)
time=pec[3*(0:(long-1))+1]
s1=pec[3*(0:(long-1))+2]
s2=pec[3*(0:(long-1))+3]
return(list(time,s1,s2))
}

graf=function(lec,yli=2,col1='blue',col2='red',typ='l'){
plot(lec[[1]],lec[[2]],type=typ,ylim=c(-yli,yli),col=col1)
lines(lec[[1]],lec[[3]],type=typ,ylim=c(-yli,yli),col=col2)
}

potencia=function(x,i){
poten<-array(dim=c(length(x),1),0)
b<-1
while(b<(length(x)-i)){
cant<-0
cant<-sum(x[b:(b+i)])
poten[b]<-cant
b<-b+1
}
return(poten)
}

absol=function(x,i){
a=max(x)
x=pmin(x,rep(a/5,length(x)))
xx=abs(x-c(0,x)[1:length(x)])
y=cumsum(xx)
abs=y-c(rep(0,i),y)[1:length(y)]
absco=abs-median(abs)
absco=absco*(absco>4)
return(absco)
}

graf12pot=function(salteo=0.5,intervalo=.1,margen=.02,ancho=10,ancho1=300,hpeaks=1.5,
pausa=1){
while(salteo<30){
lec=lectura("secuenciapulsos.txt",max(0,salteo-margen),intervalo+2*margen)
salteo=salteo+intervalo
t0=lec[[1]]
p0=sqrt((lec[[2]])^2+(lec[[3]])^2)
p=potencia(p0,ancho)
q=absol(p,ancho1)
pmas=c(p,0)[1+1:length(p)]
pmenos=c(0,p)[1:length(p)]
t=t0[(2*p-pmas-pmenos>hpeaks)&(p>3)&((p-pmas)&(p-pmenos)>=0)]
par(mfrow=c(2,1))
plot(t0,lec[[2]],type='l')
lines(t0,lec[[3]],type='l',col='blue')
lines(t,rep(0,length(t)),type='p',col='red')
plot(t0,p,type='l',ylim=c(0,10))
lines(t0,q,type='l',col='red')
lines(t,rep(0,length(t)),type='p',col='red')
if(pausa==1)a=readline()
if(a==-1)salteo=salteo-2*intervalo
if(a==0)pausa=0
}
}

graf12pot()
```

10.2. Identificación de pulsos según cada pez

```
pulsos<-read.table("pulsosgap.txt")

#se determina un comienzo manual para que el programa continúe la secuencia
datos<-c(1,2,2,1,2,1)
largo<-dim(pulsos)[1]
pulsos<-pulsos[2:largo,]
largo<-largo-1
ceros<-array(0,largo-length(datos))
peces<-c(datos,ceros)

pecespulsos<-cbind(pulsos,peces)
#extraigo los determinados manualmente
pez1<-pecespulsos[pecespulsos[,5]==1,]
pez2<-pecespulsos[pecespulsos[,5]==2,]

for (proximo in 7:largo){
#tomo el siguiente de los sin determinar, que está en la posición 19
l1<-dim(pez1)[1]
l2<-dim(pez2)[1]
#obtengo cuantos tengo de cada pez

periodo1<-pez1[l1,1]-pez1[l1-1,1]
periodo2<-pez2[l2,1]-pez2[l2-1,1]
#calculo el periodo según los últimos dos pulsos

estimado1<-pez1[l1,1]
estimado2<-pez2[l2,1]
#obtengo el ultimo pez de cada tipo

real<-pecespulsos[proximo,1]
#obtengo donde está el próximo sin determinar cuál es

#determino la diferencia entre donde debería estar el próximo pulso de cada pez y el real
gap1<-abs(real-(estimado1+periodo1))
gap2<-abs(real-(estimado2+periodo2))

#lo mismo por las dudas de que se haya saltado un pulso
dosgap1<-abs(real-(estimado1+2*periodo1))
dosgap2<-abs(real-(estimado2+2*periodo2))
#determino cual es el mas cercano
maschico<-min(gap1,gap2,dosgap1,dosgap2)

if (maschico<(min(periodo1,periodo2)*0.3)){
if (maschico==gap2) {
#agrego un pez2
pez2<-rbind(pez2,pecespulsos[proximo,])
}
if (maschico==gap1) {
#agrego un pez1
pez1<-rbind(pez1,pecespulsos[proximo,])
}
if (maschico==dosgap2) {
#agrego el actual y el faltante
donde<-estimado2+periodo2
pez2<-rbind(pez2,c(donde,0,0,donde-pez2[l2,1],2))
pez2<-rbind(pez2,pecespulsos[proximo,])
pecespulsos[proximo,5]==2
}
if (maschico==dosgap1) {
#agrego el actual y el faltante
donde<-estimado1+periodo1
pez1<-rbind(pez1,c(donde,0,0,donde-pez1[l1,1],1))
pez1<-rbind(pez1,pecespulsos[proximo,])
pecespulsos[proximo,5]==2
}
}
}
```

10.3. Ajuste de función cuadrática

```
#funcion a ajustar
parabola<-function(par , puntos){
  t=puntos [,1]
  h=par [1] *t ^2+par [2] *t+par [3]
  z=puntos [,2]
  f=sum((z>h) *100*(z-h)+h)
  return(f)
}

nombre1<-” chirp”

for (nombre2 in 1:93) {
  archivo<-paste(nombre1 , nombre2 , sep=”)
  archivo2<-paste(archivo , ” .txt” , sep=”)
  #leo los datos
  chirp<-read.table(archivo2)
  cant<-dim(chirp) [1]

  #maximo t
  maxtiempo<-chirp [cant ,1]

  for (sensor in 2:3){
    #elijo la señal
    pulso<-sensor
    t<-chirp [,1]
    y<-chirp [, pulso] ^2
    x<-cbind(t , y)

    #optimización de la parábola
    c0<-runif(3)
    op1<-optim(c0 , parabola , puntos=x)
    c1<-op1$par
    op2<-optim(c1 , parabola , puntos=x)
    c2<-op2$par
    op3<-optim(c2 , parabola , puntos=x)
    c3<-op3$par
    rbind(c1 , c2 , c3)
    plot(t , y , type=” l”)
    lines(t , c3 [1] *t ^2+c3 [2] *t+c3 [3]

  }
}
```

10.4. Ajuste por polinomio grado 4

```
#funcion a ajustar
polinomio4<-function(par, puntos){
  t=puntos[,1]
  h=par[1]*t^4+par[2]*t^3+par[3]*t^2+par[4]*t+par[5]
  z=puntos[,2]
  f=sum((z>h)*70*(z-h)+h)
  return(f)
}

aa<-4
nombre1<-"chirp"

for(nombre2 in 1:93){
  archivo<-paste(nombre1,nombre2,sep="")
  archivo2<-paste(archivo, ".txt", sep="")
  #leo los datos
  chirp<-read.table(archivo2)
  cant<-dim(chirp)[1]

  #maximo t
  maxtiempo<-chirp[cant,1]

  #en caso de querer graficar el chirp original con los dos sensores
  #par(mfrow=c(1,2))
  #plot(chirp[,1], chirp[,2], type="l")
  #plot(chirp[,1], chirp[,3], type="l")

  for(sensor in 2:3){
    #elijo la señal
    pulso<-sensor
    t<-chirp[,1]
    maxt<-max(t)
    t<-t/maxt
    y<-chirp[,pulso]
    y<-(y>0)*y
    x<-cbind(t,y)

    #optimización del polinomio
    c0<-runif(5)
    op1<-optim(c0, polinomio4, puntos=x)
    c1<-op1$par
    op2<-optim(c1, polinomio4, puntos=x)

    c2<-op2$par
    op3<-optim(c2, polinomio4, puntos=x)
    c3<-op3$par
    rbind(c1, c2, c3)
    if(aa==4){
      par(mfrow=c(2,2))
      aa<-0}
      aa<-aa+1

    plot(t, y, type="l")
    lines(t, c1[1]*t^4+c1[2]*t^3+c1[3]*t^2+c1[4]*t+c1[5], col='red')
    lines(t, c2[1]*t^4+c2[2]*t^3+c2[3]*t^2+c2[4]*t+c2[5], col='green')
    lines(t, c3[1]*t^4+c3[2]*t^3+c3[3]*t^2+c3[4]*t+c3[5], col='blue', lty=4)

    pri<-c3[1]*0^4+c3[2]*0^3+c3[3]*0^2+c3[4]*0+c3[5]
    sec<-c3[1]*0.25^4+c3[2]*0.25^3+c3[3]*0.25^2+c3[4]*0.25+c3[5]
    ter<-c3[1]*0.5^4+c3[2]*0.5^3+c3[3]*0.5^2+c3[4]*0.5+c3[5]
    qua<-c3[1]*0.75^4+c3[2]*0.75^3+c3[3]*0.75^2+c3[4]*0.75+c3[5]
    quin<-c3[1]*1^4+c3[2]*1^3+c3[3]*1^2+c3[4]*1+c3[5]
    #almaceno los resultados para hacer luego el analisis de clusters
    a<-as.vector(c(c3, pri, sec, ter, qua, quin, archivo2, pulso-1))
    write.table(t(a), file="ajuste_curva_para_cluster_\npol4.txt", sep="\t", append=T, col.
      names=F)
  }
}
```

10.5. Análisis de conglomerados

```
library(cluster)

 analisis<-read.table("ajustecurva_para_cluster.txt", header=F)
 clusters<- analisis[,c(5,6)]
 clusters2<- standard(clusters)
 agrupo<-agnes(clusters2, metric = "euclidean", stand = FALSE, method = "ward")
 ind<- indicadores(agrupo[4], clusters2)
 k<-3
 cl<- as.factor(cutree(agrupo[4],k) )
 grupos<- cbind(analisis ,cl)
```

Modificado a partir del programa realizado por los Lic Andres Castrillejo y Lic Oscar Gutierrez para el curso de análisis multivariado, licenciado bajo licencia GNU-GPL

10.6. Evaluación de modelos

```
totalSimulacion<-1000
resultado=array(0,totalSimulacion)

for (s in 1:totalSimulacion){

simulacion<-76

modeloInteraccionesX<-function(par,datos){
z<-0
n<-76 #cantidad de pulsos del pez1
pez1<-datos[datos[,5]==1,]
pez2<-datos[datos[,5]==2,]

for (i in 2:n){

xt<-pez1[i-1,1]
yt<-pez2[,1]

#ya = yt-1 anterior al xt-1
ya<-max(yt[yt<xt])
if (ya==(-Inf)) {ind1<-0} else {ind1<-(xt-ya<par[5])*1}

#yp = yt-1 posterior al xt-1
yp<-min(yt[yt>xt])
ind2<-(yp-xt<par[5])*1

#yaa = yt anterior a xt
yaa<-max(yt[yt<pez1[i,1]])
ind3<-(xt-yaa<par[5])*1

#calculo el error
z<-z+(pez1[i,1]-par[1]-pez1[i-1,1]-par[2]*ind1-par[3]*ind2-par[4]*ind3)
}
#retorno la suma de errores dividido n
return((z^2)/n-1)
}

modeloSimpleX<-function(par,pez1){
z<-0
n<-78 #cantidad de pulsos del pez1
for (i in 2:n){
xta<-pez1[i-1,1]
z<-z+(pez1[i,1]-xta-par)
}
return((z^2)/n)
}

modeloSimpleY<-function(par,pez2){
z<-0
n<-76 #cantidad de pulsos del pez2
for (i in 2:n){
yta<-pez2[i-1,1]
z<-z+pez2[i,1]-yta-par
}
return((z^2)/n-1)
}

pez1<-read.table("peces1.txt")
pez2<-read.table("peces2.txt")
pez1[,5]<-1
pez2[,5]<-2
peces<-rbind(pez1,pez2)

largoX<-dim(pez1)[1]
largoY<-dim(pez2)[1]

#estimo pX y pY por mínimos cuadrados
pX<-optim(0.2,modeloSimpleX,pez1=pez1)
pY<-optim(0.2,modeloSimpleY,pez2=pez2)
```

```

pX<-pX$par
pY<-pY$par

#estimo sigma para x modelo simple
sumaX<-0
sumaY<-0
largoX<-dim(pez1)[1]
largoY<-dim(pez2)[1]
for (i in 2:largoX){
sumaX<-sumaX+(pez1[i,1]-pez1[i-1,1]-pX)*(pez1[i,1]-pez1[i-1,1]-pX)
}

for (i in 2:largoY){
sumaY<-sumaY+(pez2[i,1]-pez2[i-1,1]-pY)*(pez2[i,1]-pez2[i-1,1]-pY)
}

sigx<-sumaX/largoX
sigy<-sumaY/largoY

parametros<-array(0,dim=c(100,5))
parametros2<-array(0,dim=c(100,5))
parametros3<-array(0,dim=c(100,5))
parametros4<-array(0,dim=c(100,5))

error1<-array(0,100)
error2<-array(0,100)
error3<-array(0,100)
error4<-array(0,100)

#aplico series de x en modelo con interaccion
xs<-array(pez1[1,1],c(100,largoX))
ys<-array(pez2[1,1],c(100,largoY))

for (i in 1:100){

for (t in 2:largoX) {
epsilon1<-rnorm(1,0,1)
xs[i,t]<-xs[i,t-1]+pX+sigx*epsilon1
}
for (t in 2:largoY) {
epsilon2<-rnorm(1,0,1)
ys[i,t]<-ys[i,t-1]+pY+sigy*epsilon2
}

vecX<-as.matrix(xs[i,])
vecXX<-as.matrix(cbind(vecX,0,0,0,1))
vecY<-as.matrix(ys[i,])
vecYY<-as.matrix(cbind(vecY,0,0,0,2))
simulados<-rbind(vecXX,vecYY)
parametros[i]<-optim(c(pX,0.005,0.005,0.005,0.005), modeloInteracionesX, datos=simulados)
par<-c(parametros[[1]],parametros[[2]],parametros[[3]],parametros[[4]],parametros[[5]])
error1[i]<-modeloInteracionesX(par, datos=simulados)

parametros2[i]<-optim(c(par[1],par[2],par[3],par[4],par[5]), modeloInteracionesX, datos=simulados)
parS2<-c(parametros2[[1]],parametros2[[2]],parametros2[[3]],parametros2[[4]],parametros2[[5]])
error2[i]<-modeloInteracionesX(parS2, datos=simulados)

parametros3[i]<-optim(c(parS2[1],parS2[2],parS2[3],parS2[4],parS2[5]), modeloInteracionesX, datos=simulados)
parS3<-c(parametros3[[1]],parametros3[[2]],parametros3[[3]],parametros3[[4]],parametros3[[5]])
error3[i]<-modeloInteracionesX(parS3, datos=simulados)

parametros4[i]<-optim(c(parS3[1],parS3[2],parS3[3],parS3[4],parS3[5]), modeloInteracionesX, datos=simulados)
parS2<-c(parametros4[[1]],parametros4[[2]],parametros4[[3]],parametros4[[4]],parametros4[[5]])
error4[i]<-modeloInteracionesX(parS3, datos=simulados)

```

```

}
#aplico datos reales en modelo con interacción

xs<-pez1
ys<-pez2
parReal<-optim(c(pX,0.005,0.005,0.005,0.005), modeloInteraccionesX, datos=peces)
parR<-c(parReal[[1]], parReal[[2]], parReal[[3]], parReal[[4]], parReal[[5]])
errorReal1<-modeloInteraccionesX(parR, datos=peces)

aa<-as.vector(parReal[1])
parReal2<-optim(c(aa$par[1], aa$par[2], aa$par[3], aa$par[4], aa$par[5]), modeloInteraccionesX
, datos=peces)
parR2<-c(parReal2[[1]], parReal2[[2]], parReal2[[3]], parReal2[[4]], parReal2[[5]])
errorReal2<-modeloInteraccionesX(parR2, datos=peces)

aa<-as.vector(parReal2[1])
parReal3<-optim(c(aa$par[1], aa$par[2], aa$par[3], aa$par[4], aa$par[5]), modeloInteraccionesX
, datos=peces)
parR3<-c(parReal3[[1]], parReal3[[2]], parReal3[[3]], parReal3[[4]], parReal3[[5]])
errorReal3<-modeloInteraccionesX(parR3, datos=peces)

aa<-as.vector(parReal3[1])
parReal4<-optim(c(aa$par[1], aa$par[2], aa$par[3], aa$par[4], aa$par[5]), modeloInteraccionesX
, datos=peces)
parR4<-c(parReal4[[1]], parReal4[[2]], parReal4[[3]], parReal4[[4]], parReal4[[5]])
errorReal4<-modeloInteraccionesX(parR4, datos=peces)
}

errores<-sort(error4)
posicion<-sum(errorReal4>error4)
resultado[s]<-posicion/simulacion
}

```

11. Bibliografía

1. Blanco, J. (2006). Introducción al análisis multivariado - Notas de clase.
2. Cabaña, E. (s.f.). Prueba de hipótesis para modelos no lineales usando simulación. Comunicación privada.
3. Jones, O., Maillardet, R., & Robinson, A. (2014). Introduction to Scientific Programming and Simulation Using R.
4. Macadar, O., & Silva, A. (2007). Comunicación eléctrica en peces sudamericanos del orden Gymnotiformes. Revista latinoamericana de psicología(39), 31-45.
5. Maria Jose Zapatero, J. A. (2011). Análisis de algunas metaheurísticas creadas a partir de "Optimización Gravitatoria".
6. Mathews, J., & Fink, C. (1999). Numerical Methods using Matlab.
7. Peña, D., & Hill, M. G. (2002). Análisis de datos multivariantes.
8. Perrone, R., Macadar, O., & Silva, A. (2009). Social electric signals in freely moving dyads of *Brachyhypopomus*.
9. R core team. (2015). A language and environment for statistical computing.
10. R Documentation. (s.f.). Obtenido de <http://127.0.0.1:24705/library/base/html/Random.html>
11. Rubinstein, R. (1961). Simulation and Monte Carlo method.