Universidad de Valparaíso
CHILE

Facultad de Ciencias
Instituto de Estadística
Ingeniería en Estadística

Modelos de series de tiempo para los desembarques de anchoveta en la zona centro sur de Chile, con y sin variable exógena.

Felipe Antonio Sánchez González
13 marzo 2023

Profesor Guía
Mauricio Tejo Arriagada Ph.D.
Instituto de Estadística, Universidad de Valparaíso

Proyecto de titulación para optar al grado académico de:
Magíster en Estadística

Índice general

Agradecimientos **3**

Resumen **4**

Objetivos **5**

1. Contexto ecológico pesquero **6**
1.1. Pesquería de la anchoveta en la zona centro-sur de Chile . . . 7
1.2. Área de estudio . . . 7
1.3. Variables ambientales y pesqueras . . . 8

2. Series de tiempo **10**
2.1. Introducción a las series de tiempo . . . 10
2.2. Enfoque clásico de series de temporales . . . 10
2.2.1. Tendencia . . . 11
2.2.2. Componente estacional . . . 11
2.2.3. Componente irregular o aleatorio . . . 11
2.2.4. Datos atípicos (*Outlier*) . . . 11
2.3. Descomposición de una serie de tiempo . . . 12
2.4. Propiedades de las distribuciones marginales . . . 12
2.4.1. Procesos estacionarios . . . 13
2.5. Serie estacionaria . . . 13

3. Modelo propuesto **15**
3.1. Introducción al modelo GARCH-X . . . 15
3.2. Modelo GARCH-X . . . 16
3.2.1. Modelos ARCH y GARCH . . . 16
3.3. Teoría asintótica . . . 17
3.3.1. Caso estacionario . . . 19
3.3.2. Proceso estocástico estacionario . . . 21

4. Análisis estadístico de los datos **22**
4.1. Selección de variables . . . 22
4.2. Correlación de las variables . . . 22
4.3. Análisis de Componente Principales (ACP) . . . 22
4.4. Variables del modelo GARCH-X y estacionariedad . . . 25
4.4.1. Serie de los desembarques de anchoveta . . . 26
4.4.2. Serie de la Temperatura superficial del mar . . . 26

4.5. Prueba de estacionariedad Dickey Fuller 26
4.6. Estimación del modelo GARCH-X para los desembarques de anchoveta 30

5. Conclusiones **34**

6. Discusión **35**

Resumen

Las fluctuaciones en abundancia, biomasa, estructura de edad y patrones de distribución de los recursos pesqueros responden, entre otros, a la variabilidad ambiental. Dentro de estos recursos se encuentra la anchoveta (*Engraulis ringens*), estos peces habitan en la región nerítica, es decir, aquella masa de agua situada sobre la plataforma continental (aproximadamente hasta los 200 m de profundidad) en esta zona desarrollan casi el 90 % de las pesquerías mundiales Camiñas y Baro (2004). La relación ambiente-recurso ha sido ampliamente abordada, en consideración a que el entendimiento de los procesos naturales de variabilidad en las diferentes escalas de tiempo y sus impactos sobre el ecosistema son esenciales para un efectivo manejo de los recursos hidrobiológicos, generando el conocimiento que permita administrar de manera eficiente y sustentable los recursos naturales. En consecuencia, este proyecto estudió desde el punto de vista de las series de tiempo, un modelo para los desembarques de anchoveta en la zona centro sur de Chile, seleccionando con técnicas estadísticas como variable exógena la temperatura superficial del mar captada por satélites NOAA en la zona donde se distribuye este recurso, con la cual se obtuvo un mejor desempeño que el modelo que no incluyó dicha variable. Para esto, fue necesario describir los datos, se identificó la variable exógena, se revisó la estacionalidad de las variables de entrada al modelo y se implementaron dos modelos, un modelo GARCH y un modelo GARCH-X (con variable exógena), para luego, estimar sus respectivos parámetros, validarlos, junto con discutir los resultados. La implementación del modelo GARCH-X del presente trabajo en la pesquería de anchoveta, en la zona cetro sur de Chile, brinda información sustancial para la toma de decisiones del manejo pesquero. Que en conjunto con otros enfoques metodológicos, y el conocimiento de expertos, permite a obtener una mayor comprensión de la dinámica de este recurso que se caracteriza por un alta variabilidad ante los cambios del ambiente.

Objetivos

Objetivo General

Implementar un modelo de serie temporal GARCH-X con variable exógena, con el objeto de predecir las fluctuaciones de la abundancia de anchoveta en la zona centro sur de Chile.

Objetivos Específicos

1. Describir los datos e identificar la variable exógena a considerar.
2. Implementar el modelo GARCH-X y estimar los parámetros.
3. Validar el modelo y discutir los resultados.

Capítulo 1

Contexto ecológico pesquero

Las poblaciones de pequeños peces pelágicos proporcionan aproximadamente 25 % del rendimiento total anual de la captura de peces en todo el mundo y de muchas comunidades costeras, en particular en los países en desarrollo (Alheit y Peck, 2019; Canales et al., 2020; Grisa y Sabourin, 2019). La Corriente de Humboldt es uno de los ecosistemas marinos más ricos en productividad biológica, lo que permite sustentar una de las mayores producciones pesqueras del planeta (Gonzalez et al., 1998; Iriarte et al., 2000). Esto se debido al transporte de nutrientes por la advección horizontal a gran escala y el afloramiento costero persistente (Suryanarayana et al., 2008; Carr, 2001). Respecto a Chile, se realiza la pesca de múltiples especies y dentro de ellas los pequeños pelágicos como la anchoveta (*Engraulis ringens*; Jenyns, 1842), uno de los recursos pesqueros de mayor importancia para la flota cerquera industrial y artesanal de la zona centro sur de Chile (33°00"S - 41°30"S) (S. Cubillos et al., 1998; Yáñez Rodríguez, 1998). Esta pesquería se inicia en la década del 40" con la explotación conjunta de sardina común y anchoveta, presentando un continuo crecimiento en sus desembarques hasta 1969. Sin embargo durante los años siguientes estos presentaron fuertes variaciones de carácter interanual e interdecadal, evidenciando el carácter fluctuante de esta pesquería.

Las fluctuaciones en abundancia, biomasa, estructura de edad y patrones de distribución de los recursos pesqueros responden, entre otros, a la variabilidad ambiental (Yáñez, 1998; Parada et al., 2013). Dentro de los recursos pesqueros, se ha demostrado que el ciclo de vida de los pequeños pelágicos está estrechamente relacionado con el ciclo estacional de temperatura superficial del mar (L. A. Cubillos et al., 2007; Yáñez et al., 2008), que es el caso de la anchoveta, pez pelágico costero de importancia comercial para los países limítrofes del Sistema de la Corriente de Humboldt. Las poblaciones de estos peces experimentan fluctuaciones extremas en la abundancia y tienen una amplia distribución geográfica (Alheit y Peck, 2019).

Los mecanismos que causan las fluctuaciones en de los pequeños pelágicos han recibido una atención considerable en todo el mundo (Alheit et al., 2009; MacCall, 2009; Alheit y Bakun, 2010). La opinión general es que las variaciones en la abundancia de las poblaciones están dirigidas principalmente por controles ascendentes, como la temperatura del océano, el afloramiento, la composición del plancton, el oxígneo, el canibalismo y la pesca (Chavez et al., 2003; Alheit y Niquen, 2004; Ayón et al., 2008; Bertrand et al., 2011) procesos top-down, como la depredación intragremial y los depredadores naturales (Irigoien y de Roos, 2011; Checkley Jr et al., 2017; KL y MC, s.f.; Canales et al., 2016, 2020); y los rasgos intrínsecos de la especie (Checkley Jr et al., 2017).

La relación ambiente-recurso ha sido ampliamente abordada (Crawford, 1991; Schwartzlose y Alheit, 1999; Cury et al., 2000; Yáñez et al., 2001; Yáñez et al., 2003; Yáñez et al., 2008), lo que es necesario debido a que el entendimiento de los procesos naturales de variabilidad en las diferentes escalas de tiempo y sus impactos sobre el ecosistema son esenciales para un efectivo manejo de los recursos vivos marinos (Alheit y Niquen, 2004). Los mecanismos que definen la variabilidad ambiental se evidencian a través de diferentes indicadores climáticos.

Respecto a los cambios océano-climáticos interanuales, estos son asociados principalmente al fenómeno El Niño, lo que influye directamente en el ciclo de vida de la anchoveta, dado que su ciclo reproductivo está acoplado al ciclo estacional de surgencia costera (L. Cubillos et al., 2001).

Por lo tanto, cualquier alteración en esta, afectará en la reproducción de las especies. Los cambios de régimenes interdecadales son reflejados por la Oscilación Decadal del Pacífico (S. R. Hare, 1996; Mantua et al., 1997). Los cambios océano-climáticos alteran los ecosistemas marinos en diversas escalas (Alheit, 1993; S. Hare et al., 2000; Chavez et al., 2003; Yáñez et al., 2003; Alheit y Niquen, 2004). Pesquerías como la de la anchoveta están sometidas a fluctuaciones intraanuales, de marcada estacionalidad Aranis et al. (2013), y fluctuaciones de carácter interanuales e interdecadales (Alheit, 1993; Chavez et al., 2003; Yáñez et al., 1992).

1.1. Pesquería de la anchoveta en la zona centro-sur de Chile

La anchoveta, es uno de los recursos de mayor importancia pesquera para la flota industrial y artesanal de cerco que opera en la zona centro-sur de Chile 1.1, cuyos principales puertos de descarga de capturas se sitúan en Talcahuano y Corral. Estos recursos se caracterizan por tener una distribución costera y ser muy influenciados por las condiciones ambientales, presentando además un corto ciclo de vida, con alta mortalidad natural, elevado crecimiento y con tendencia a formar cardúmenes altamente densos (L. Cubillos et al., 2001; L. A. Cubillos y Arcos, 2002). En relación a la macrozona de pesca de estas especies, es decir la zona costera del centro sur de Chile, la pesquería se desarrolla estacionalmente con capturas fuertemente dependientes del pulso de reclutamiento anual, proceso que tiende a ocurrir en pleno período estival (S. Cubillos et al., 1998).

La estacionalidad de las capturas se mantiene y se ha hecho más notoria en los últimos años, debido a que el reclutamiento de estos recursos es en el último trimestre y primer semestre. De este modo se mantiene la estacionalidad característica del sistema pretérito de libre acceso o sin cuota. Considerando esto, bajo el actual sistema administrativo, la estacionalidad es modulada por la asignación trimestral de las cuotas, que han mantenido las proporciones estacionales históricas.

1.2. Área de estudio

El área de estudio comprende la zona centro sur, la cual se extiende entre Valparaíso (33°00”S) y Los Lagos (41°30”S). Figura 1.1.

En la Figura 1.2 se presentan los datos georeferenciados por flota de pesca artesanal e industrial y por distintas categorías operacionales. En primer lugar, se muestra que actividad pesquera sobre sardina común y anchoveta en la zona centro sur entre los años 2001 y 2014, observándose que la

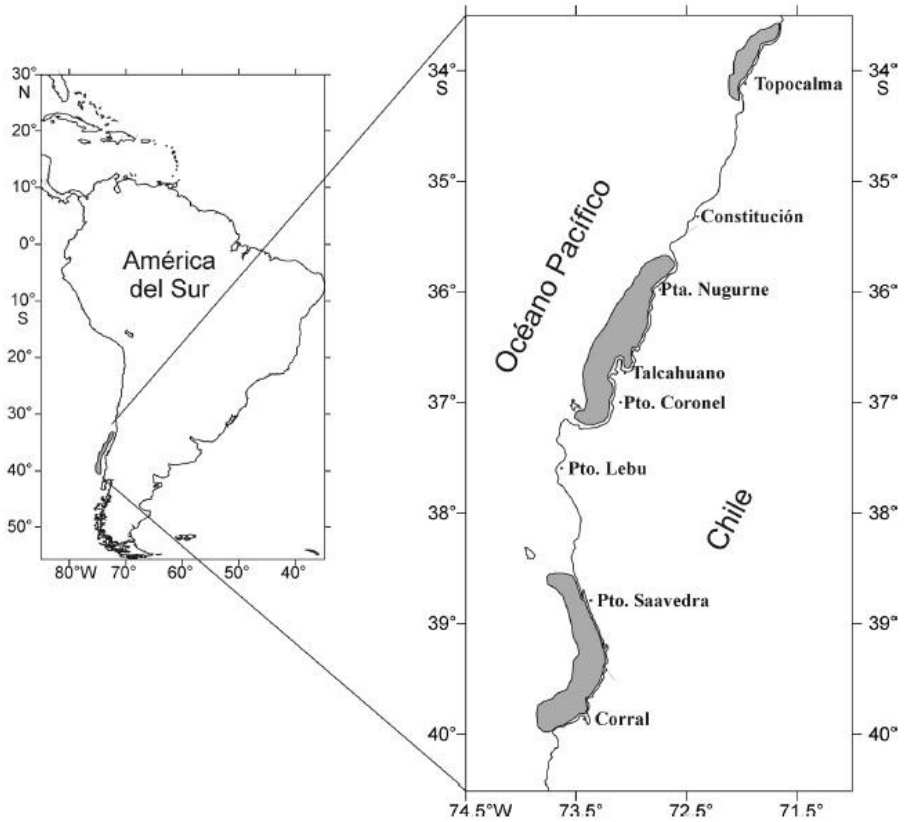


Figura 1.1: Área principal de captura de anchoveta en la zona centro sur de Chile. Gatica et al. (2007).

mayor captura ocurre en las Regiones del BíoBío y Los Ríos, siendo la bahía de Concepción y el golfo de Arauco las zonas que sostienen la mayor recurrencia de pesca, relegando al segundo lugar la zona entre el sur de puerto Saavedra y Corral.

1.3. Variables ambientales y pesqueras

Las variables ambientales (locales y globales) y pesqueras mensuales del período 1983-2012. Las variables locales temperatura superficial del mar (TSM) y nivel medio del mar (NMM) provienen del Servicio Hidrográfico y Oceanográfico de la Armada (SHOA); mientras que la temperatura del aire (TA), magnitud y dirección del viento provienen del Servicio Meteorológico de Chile (SMCH). Ambos conjuntos de datos corresponden a las estaciones oceanográficas y meteorológicas de Talcahuano. Con la magnitud y dirección del viento de Carriel Sur se estiman el Indice de Turbulencia (IT) y el Transporte de Ekman (TE). La TSM captada a través de los satélites NOAA se obtiene del Earth System Research Laboratory (ESRL). Los índices globales se recuperan de informes de libre disposición (www.cpc.ncep.noaa.gov/data/indices). Tabla 1.1.

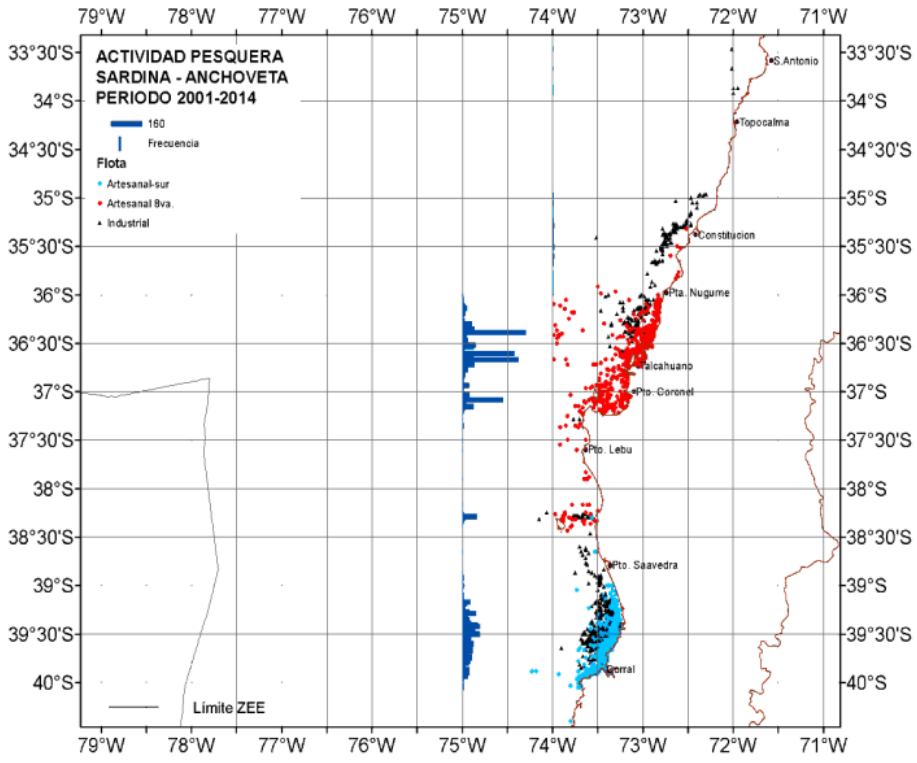


Figura 1.2: Distribución espacial de los muestreos de sardina y anchoveta realizados a la flota artesanal e industrial de la Región del BíoBío y de la Región de los Ríos, desde 2001 al 2014. IIP (2016)

Tabla 1.1: Variables ambientales y pesqueras mensuales utilizadas en el estudio.

Variables		Símbolo	Unidad
Locales	Indice de turbulencia	IT	m^3/s^3
	Transporte de Ekman	TE	m^3/s^3
	Temperatura superficial del mar	TSM	°C
	Temperatura superficial del mar	TSM-NOAA	°C
	Temperatura del aire	TA	°C
	Nivel medio del mar	NMM	cm
Globales	Oscilación decadal del pacífico sur	ODP	°C
	TSM en la región Niño 1+2	NIÑO 1+2	°C
	TSM en la región Niño 3.4	NIÑO 3.4	°C
	Índice de oscilación del sur	IOS	Mbar
	Índice de Humboldt	HCI	Mbar
	Oscilación Antártica	AAO	Mbar
	Índice de Lengua Fría	ILF	°C * 100
Pesquera	Desembarque	Des	t

Capítulo 2

Series de tiempo

2.1. Introducción a las series de tiempo

Una serie de tiempo corresponde a un proceso estocástico $\{Z_t\}_{t\in\mathcal{T}}$ definido sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, adaptado a una filtración $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathcal{T}}$, en donde $\mathcal{T}$ corresponde a un conjunto numerable, típicamente $\mathbb{Z}$ o $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. En nuestro caso, dado el contexto de nuestra aplicación, consideraremos $\mathcal{T} = \mathbb{Z}_+$, y asumiremos que la filtración $\mathbb{F}$ corresponderá a la filtración natural del proceso a modelar; esto es, para cada $t \in \mathbb{Z}_+$, $\mathcal{F}_t$ denotará los conjuntos medibles de la historia de nuestro proceso, desde el inicio de nuestras observaciones en el tiempo 0, hasta el tiempo actual t.

Una serie temporal es un conjunto de datos que se registran a lo largo del tiempo, y son comunes en muchas áreas de estudio. Según Hyndman y Athanasopoulos (2018), una serie temporal es una secuencia ordenada de observaciones tomadas a intervalos regulares en el tiempo. Estos datos pueden ser cualquier tipo de variable que cambie a lo largo del tiempo, como el precio de una acción, el índice de temperatura, el tráfico en una red de computadoras, etc. Las series temporales son útiles para predecir el comportamiento futuro de una variable a partir del comportamiento pasado, y también para entender cómo se relaciona una variable con otras. Según Chatfield et al. (2014), el análisis de series temporales es una parte fundamental de la estadística y tiene muchas aplicaciones prácticas, incluyendo la predicción de la demanda, la gestión de inventarios, la planificación de la producción y la evaluación de la estabilidad de los sistema.

Para analizar una serie temporal, se suele dividir en dos componentes: la tendencia y la estacionalidad. La tendencia es la dirección general en la que se está moviendo la serie temporal, mientras que la estacionalidad se refiere a patrones periódicos en la serie, como cambios estacionales en la temperatura o en las ventas de navidad. Según Box y Jenkins (1976), "La tendencia y la estacionalidad son dos componentes clave en la mayoría de las series temporales, y deben ser identificadas y modeladas adecuadamente para poder hacer predicciones precisas".

2.2. Enfoque clásico de series de temporales

El enfoque clásico de series temporales hace una referencia a que toda serie tiene cuatro componentes dominantes, la tendencia, componente cíclico, componente estacional y componente irregular. En esta metodología se considera como componentes deterministas a la tendencia, al componente

cíclico y la componente estacional, por lo tanto cada una de estas componentes pueden ser estimadas por diversos métodos, entre ellos se encuentra para el caso de la tendencia, el método de media móvil y regresión con respecto al tiempo, así también, para el caso de la componente estacional se suele utilizar el análisis espectral. Tanto la componente estacional como la ciclicidad corresponden a patrones que se repiten en la serie en determinado periodo de tiempo, donde la componente estacional hace referencia a patrones con ocurrencia en períodos inferiores a un a no y generalmente se atribuye a los cambios de temperatura, cambios de estaciones del a no o eventos provocados por el calendario anual. Por otra parte, la componente cíclica ocurre con periodicidad superior a un a no, como es el caso de la corriente del ni no o algunos fenómenos económicos con ciclicidad no perfecta. Para el caso de la componente irregular o aleatoria la estimación es un poco más compleja ya que corresponden a eventos fortuitos que se presentan en el transcurso del tiempo, por lo cual para estimar esta componente se utilizan métodos estocásticos. De la Fuente 2022.

2.2.1. Tendencia

Representa la evolución o cambio de la serie en el largo plazo. La tendencia está estrechamente relacionada con los cambios en la media, que al momento de ser bruscos provocan que la estructura de la serie cambie dramáticamente. Entre los tipos de tendencia determinista destaca la tendencia lineal, cuadrática, exponencial, logarítmica, entre otras. De la Fuente (2022).

2.2.2. Componente estacional

La componente estacional hace referencia a variaciones o patrones constantes que se repiten en la serie a lo largo del tiempo, donde generalmente estas variaciones suelen presentarse durante períodos mensuales, trimestrales o anuales, junto a la componente cíclica. Así como existen métodos para estimar la tendencia, también existen métodos que permiten estimar la estacionalidad. Principalmente, la estacionalidad hace referencia a aquellos fenómenos o factores que intervienen en el desenlace de la serie, ocasionando que esta presente un patrón sistemático en función de aquellos factores. Entre los factores que provocan estacionalidad en una serie, destacan, las estaciones del a no y meses del a no, las temperaturas y principalmente en series medioambientales las condiciones climáticas, por el lado de las series económicas la estacionalidad puede producirse debido a múltiples factores que intervengan en el proceso. Como una orientación, es importante mencionar que para analizar series de tiempo que presentan estacionalidad se sugiere estudiar aquellos factores que intervienen en el desenlace de esta. De la Fuente (2022).

2.2.3. Componente irregular o aleatorio

El componente irregular o aleatorio presente en algunas series de tiempo, corresponde a todo aquel comportamiento que parece ser impredecible, que no presenta patrón alguno y que exhibe fluctuaciones producidas por eventos fortuitos y no controlados. De la Fuente (2022).

2.2.4. Datos atípicos (*Outlier*)

Un *Outlier* hace referencia a todas aquellas observaciones o datos en nuestra variable de estudio que se escapan de lo normal, pueden ser atribuidos a un acontecimiento extraordinario o errores

de mediciones. Existen situaciones en la que los datos atípicos padecen de una explicación natural, en otros casos pueden estar relacionados a hechos puntuales desconocidos, por lo tanto bajo estas circunstancias se debe llevar a cabo un análisis más profundo de estos valores, dado que afectarán de manera negativa la estimación de los parámetros, en la identificación del modelo y en consecuencia el eventual pronóstico. De la Fuente (2022).

2.3. Descomposición de una serie de tiempo

En general toda serie de tiempo puede ser descompuesta matemáticamente en sus respectivos componentes, dando lugar a tres modelos, el modelo aditivo, el modelo multiplicativo y el modelo mixto. De la Fuente (2022).

Modelo aditivo

$$Z_t = T_t + C_t + S_t + I_t$$

Modelo Multiplicativo

$$Z_t = T_t \cdot C_t \cdot S_t \cdot I_t$$

Mixto

$$Z_t = (T_t \cdot C_t \cdot S_t) + I_t,$$

donde Z_t corresponde al valor que toma la serie de tiempo Z en el instante de tiempo t, T_t representa la tendencia presente en la serie, C_t es el componente cíclico presente en la serie, S_t representa la componente estacional que actúa en la serie y el componente irregular o aleatorio presente en la serie está dado por I_t.

2.4. Propiedades de las distribuciones marginales

Las medidas descriptivas asociadas a una serie de tiempo, como es el caso de la media, varianza y autocovarianza están dadas en función del tiempo. La función de medias del proceso corresponde al cálculo de la media para cada instante de tiempo t definido y su valor se denota por:

$$\mu_t = \mathbb{E}(Z_t), \text{para todo } t \in T$$

Un caso particular importante, es cuando la función de medias es constante. En este caso la serie no mostrará ninguna tendencia y se concluirá que el proceso es estable en media. Por otra parte, en el caso que la media cambie en el tiempo, se visualiza que las observaciones en distintos momentos exhibirán esos cambios.

La función de varianza del proceso se define como la varianza obtenida en cada instante temporal y es denotada como:

$$\sigma_t^2 = Var(Z_t) = \mathbb{E}[(Z_t - \mu_t)^2]$$

Se concluirá que el proceso es estable en la varianza si esta es constante en el tiempo. Es necesario mencionar que un proceso puede ser estable en la media y no en la varianza o viceversa.

Por otra parte, la estructura de dependencia lineal entre las variables aleatorias del proceso se representa por las funciones de autocovarianza y correlación. La función de autocovarianza corresponde a la covarianza obtenida entre dos variables del proceso en dos instantes cualquiera.

$$\gamma(t, t+j) = Cov(Z_t, Z_{t+j}) = \mathbb{E}[(Z_t - \mu_t)(Z_{t+j} - \mu_{t+j})]$$

En particular se tiene que,

$$\gamma(t, t) = Var(Z_t) = \sigma_t^2$$

Otra función importante en el estudio de series de tiempo es la función de autocorrelación. Se denomina coficiente de autocorrelación de orden $(t, t+j)$ al coeficiente de correlación entre las variables Z_t y Z_{t+j} y función de autocorrelación, a la función de dos argumentos que describe estos coeficientes para dos valores cualquiera de las variables. Esta función se define como:

$$\rho(t, t+j) = \frac{\gamma(t, t+j)}{\sigma_t \sigma_{t+j}}.$$

2.4.1. Procesos estacionarios

Toda serie de tiempo puede ser clasficada en dos tipos de series, en general, las series de tiempo de tipo estacionarias y las no estacionarias.

2.5. Serie estacionaria

Una serie de tipo estacionaria hace referencia al **proceso estocástico** cuyas propiedades estadísticas para la secuencia $Z_{t_1}; Z_{t_2}; ...; Z_{t_n} (n \geq 1)$ de componentes de Z_t es **semejante** a la de la

secuencia $Z_{t_1} + h$; $Z_{t_2} + h$;...; $Z_{t_n} + h$ para cualquier número entero $h = \pm 1; \pm 2; ...,$. De la Fuente (2022).

Una serie de tiempo $X = (X_t, t \in T)$ es:

- Estable en media si $\mu_t) = \mu$ (μ constante)
- Estable en varianza si $\sigma_t = \sigma_x$ (σ_x constante)
- Estable en covarianza si $\gamma(t_1, t_1 + j) = \gamma(t_2, t_2 + j) = \gamma(0, |j| :=\gamma(|j|)$, para todo $t_1, t_2 \in T$ y para todo j, tal que $t_1 + j, t_2 + j \in T$

Existen dos nociones de estacionariedad, la primera es en sentido débil y la segunda en sentido estricto. El primer concepto está relacionado con propiedades distribucionales y la segunda con la trayectoria (realización) de la serie.

- Una serie de tiempo es estacionaria en sentido débil si es estable en media y autocovarianza.
- Una serie de tiempo es estacionaria en sentido estricto si

$$F_{t_1,...,t_k}(x_1, ..., x_k) = F_{t_1+h,...,t_k+h}(x_1, ...x_k)$$

para todo $k, t_1, ...t_k \in T$ y para todo h tal que $t_1 + h, ..., t_k + h \in T$, donde F denota la función de distribución acumulada del vector $(Z_{t_1}, Z_{t_2}, ..., Z_{t_n})(n \leq 1)$.

Capítulo 3

Modelo propuesto

3.1. Introducción al modelo GARCH-X

El modelo GARCH-X son una extensión del modelo GARCH (Generalized Autoregressive Conditional Heteroscedasticity) que incluye un término adicional, conocido como el término exógeno o término X, que permite modelar la dependencia de la varianza de una serie temporal con una variable exógena. Los modelos GARCH-X se utilizan en diversas aplicaciones, como la predicción de precios de acciones y la evaluación del riesgo financiero, y son útiles para mejorar la precisión de las predicciones de "volatilidad". Según Fan et al. (2003), "Los modelos GARCH-X proporcionan una forma de modelar la dependencia de la varianza en una serie temporal con otras variables exógenas, lo que puede mejorar la precisión de las predicciones de la varianza".Los modelos GARCH-X se dividen en tres componentes: el componente autoregresivo (AR), el componente de media móvil (MA) y el término X. El componente AR se refiere a la relación entre la varianza actual y la varianza rezagada, el componente MA se refiere a la relación entre la varianza actual y el error residual de la serie, y el término X se refiere a la relación entre la varianza actual y las variables exógenas.

Desde el punto de vista fenomenológico y del modelado, la introducción de las variables exógenas dentro del coeficiente de ruido no es antojadiza. El ruido viene a representar las perturbaciones aleatorias que afectan a nuestros observables, debido a la interacción de nuestro sistema principal (esto es, todo lo que podemos controlar o medir directamente a través de la observación) y el ambiente, entorno o reservorio, con el cual nuestro sistema principal interactúa. Esta interacción detenta un intercambio de energía la cual asumimos "equilibrada", lo cual se refleja en la elección de un ruido centrado en cero, como lo es el ruido blanco. Por ende, esta relación no es observable (al menos totalmente), pero podemos tener alguna idea de la magnitud de tal interacción, cantidad que se ve descrita en el término σ_t. Dado que nuestras variables exógenas contienen precisamente las características abióticas del ambiente o entorno en la cual vamos a observar nuestra variable de respuesta, es que se justifica plenamente que estos observables externos sean parte de la magnitud de la interacción sistema principal/ambiente. Adicionalmente, la manera en que éstas entran en nuestra dinámica (es decir, su relación funcional con el coeficiente σ_t), extiende de manera natural la estructura GARCH original, la cual recoge la acumulación temporal de estas externalidades que han ido perturbando nuestro sistema principal en cada instante t (Ver: Von Bertalanffy (1950)).

3.2. Modelo GARCH-X

Se define $\{Z_t, n \in \mathbb{Z}\}$ como un proceso $GARCH(p,q)$-X si satisface el siguiente par de ecuaciones:

$$Z_t = \sigma_t \xi_t,\ \xi_t \overset{\text{iid}}{\sim} N(0,1)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-1}^2 + \sum_{j=1}^{q} \beta_j \xi_{t-j}^2 + \sum_{k=1}^{n} \pi_k x_{t-k}^2$$

donde $\xi_t \sim iidN(0,1)$, $w > 0, \alpha_i \geq 0, 1 \leq i \leq p$ y $\beta_j \geq 0, 1 \leq j \leq q, \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. Además la covariable del proceso exógeno usado para mejorar el comportamiento del modelo es positiva.

3.2.1. Modelos ARCH y GARCH

Previo al modelo GARCH está el modelo ARCH el cual tiene la desventaja de que a menudo requiere de altos valores en los desfaces para describir adecuadamente la variabilidad a lo largo del tiempo. Bollerslev (1986) introdujo una extensión del modelo ARCH llamado modelo ARCH generalizado, o GARCH, para superar este problema. El modelo GARCH (p, q) asume que $Z_t = \sigma_t \xi_t$, donde ξ_t nuevamente son variables iid aleatorias con media cero y varianza 1, y donde σ_t está dado por:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-1}^2 + \sum_{j=1}^{q} \beta_j \xi_{t-j}^2$$

con $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \ldots, p-1$, $\alpha_p > 0$, $\beta_j \geq 0$, $j = 1, \ldots, q-1$ y $\beta_q > 0$. Estas restricciones de parámetros son suficientes para la varianza condicional σ_t^2 sea positiva. Nelson y Cao (1992) demostraron que estas restricciones se pueden relajar ligeramente para permitir que algunos de los parámetros sean negativos mientras que la varianza condicional sigue siendo positiva. La restricción adicional $\sum_{i=1}^{m}(\alpha_i + \beta_i) < 1$, donde $m = \text{máx}(p, q)$ con $\alpha_i = 0$, para $i > p$, y $\beta_i = 0$, para $j > q$, asegura que la varianza incondicional σ^2 es finita.

El modelo más simple y más utilizado en esta clase es el modelo GARCH (1, 1) donde:

$$\sigma_t^2 = E[a_t^2 | F_{t-1}] = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Dado que las constantes α_1 y β_1 son positivas, vemos que un valor grande de a_{t-1}^2 o σ_{t-1}^2 da como resultado un valor grande de σ_t^2. En cuanto al proceso ARCH, este modelo, por lo tanto, da cuenta de la agrupación de variabilidad.

Asumiendo que $\alpha_1 + \beta_1 < 1$, la varianza incondicional de a_t es:

$$\sigma_a^2 = var[a_t]| = \alpha_0/[1-(\alpha_1+\beta_1)]$$

Además, asumiendo que las distribuciones condicionales son normales, el cuarto momento incondicional de a_t es finito siempre que $(\alpha_1+\beta_1)^2+2+\alpha_1^2 < 1$ (Bollerslev, 1986). Además, la curtosis de la distribución marginal de a_t es igual a:

$$k = \frac{E(a_{t^4})}{[E(a_{t^2})]^2} = \frac{3[1-(\alpha_1+\beta_1)^2]}{1-(\alpha_1+\beta_1)^2-2\alpha_1^2} > 3$$

Como en el caso ARCH, la distribución incondicional de a_t tiene colas más pesadas que la distribución normal y se espera que dé lugar a una mayor frecuencia de observaciones extremas o "valores atípicos"que en el caso de normalidad.

Ahora sea $v_t = a_t^2 - \sigma_t^2$ de modo que $\sigma_t^2 = a_t^2 - v_t$, donde v_t tiene media cero y no está correlacionado en serie. Luego vemos que el modelo GARCH $(1,1)$ se puede reorganizar como a_t^2 - v_t $= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1(a_{t-1}^2 - v_{t-1})$, o que tiene la forma de un proceso ARMA para a_t^2 con orden AR igual a $m = max(p,q)$. La estructura de autocorrelación de a_t^2 también imita la del proceso ARMA siempre que el cuarto momento incondicional de a_t sea finito (Bollerslev, 1988). La condición necesaria y suficiente para la estacionariedad de segundo orden del proceso GARCH (p,q) es:

$$\sum_{i=1}^{p}\alpha_i + \sum_{i=1}^{q}\beta_i = \sum_{i=1}^{m}(\alpha_i+\beta_i) < 1$$

Cuando se cumple esta condición, la varianza incondicional es:

$$\sigma_a^2 = var[a_t] = \alpha_0/\left[1-\sum_{i=1}^{m}(\alpha_i+\beta_i)\right]$$

3.3. Teoría asintótica

Los argumentos básicos utilizados para establecer la distribución asintótica del QMLE son idénticos para los regresores estacionarios o no estacionarios. La única diferencia es que en el caso no estacionario, se mantiene ω fijo en ω

promedio dado que este no se identifica en esta situación, se puede estimar ω junto con los demás parámetros. La estrategia de prueba es el vector de puntuación denotado por:

$$S_{n,\gamma}(\gamma) = \partial L_n(\gamma)/(\partial\gamma) = \sum_{t=1}^{n} S_{\gamma,t}(\gamma)$$

y la matriz Hessiana por:

$$H_{n,\gamma\gamma}(\gamma) = \partial^2 L_n(\gamma)/\partial_\gamma = \sum_{t=1}^{n} H_{\gamma\gamma,t}(\gamma)$$

respectivamente, donde $\gamma = \vartheta$ o $\gamma = \theta$. Dependiendo el caso estacionario o no estacionario. La distribución asintótica del QMLE,$\hat{\gamma}$, se obtiene a partir de la expansión de Taylor de primer orden del vector de puntuación, i.e.

$$S_{n,\gamma}(\gamma) = S_{n,\gamma}(\gamma_0) + H_{n,\gamma\gamma}(\bar{\gamma})(\hat{\gamma} - \gamma_0),$$

Donde $\hat{\gamma}$ se encuentra en el segmento de la linea que conecta $\hat{\gamma}$ y γ_0. Si $\hat{\gamma}$ es una solución interna, tenemos $S_n(\hat{\gamma}) = 0$. Por lo tanto, se puede escribir a partir de la ecuación anterior.

$$\sqrt{n}(\hat{\gamma} - \gamma_0) = -[H_{n,\gamma\gamma}(\bar{\gamma})/n]^{-}1[S_n, \gamma(\gamma_0)\sqrt{n}]$$

Luego se verifican las siguientes condiciones ML1-ML2; estos juntos producirán el resultado deseado:

ML1 $S_{n,\gamma}(\gamma_0)/\sqrt{n} \to_d N(0, V\gamma\gamma)$.

ML2 $-H_{n,\gamma\gamma}(\bar{\gamma})/n \to_p H_{\gamma\gamma}$. para todo $H_{\gamma\gamma} > 0$

Combinando la ecuación anterior con ML1-ML2, sigue directamente que:

$$\sqrt{n}(\hat{\gamma} - \gamma_0) = -[H_{n,\gamma\gamma}(\bar{\gamma})/n]^{-}1[S_n, \gamma(\gamma_0)\sqrt{n}] + o_p(1) \to_d N(0, H_{\gamma\gamma}^{-1} V_{\gamma\gamma} H_{\gamma\gamma}^{-1})$$

Queda

$$V_{\vartheta\vartheta}^{stat} = \begin{bmatrix} V_{\vartheta\vartheta}^{stat} & V_{\theta\omega}^{stat} \\ V_{\omega\theta}^{stat} & V_{\omega\omega}^{stat} \end{bmatrix} \in \mathbb{R}^{4x4}, H_{\vartheta\vartheta}^{stat} = \begin{bmatrix} H_{\theta\theta}^{stat} & H_{\theta\omega}^{stat} \\ H_{\omega\theta}^{stat} & H_{\omega\omega}^{stat} \end{bmatrix} \in \mathbb{R}^{4x4},$$

denotan las matrices derivadas en el caso estacionario, mientras que $V_{\omega\omega}^{nonstat} \in \mathbb{R}^{3x3}$ denote las matrices obtenidas en uno no estacionario.

Como veremos, dependiendo de si x_t es estacionario o no, los términos limitantes de la covarianza $(V_{\theta\theta}^{stat}, H_{\theta\theta}^{stat})$ y $(V_{\theta\theta}^{nonstat}, H_{\theta\theta}^{nonstat})$ estarán diferente. Sin embargo, al mismo tiempo, como parte de nuestras pruebas, mostramos que los siguientes dos estimadores convergen hacia las versiones relevantes de $V_{\theta\theta}$ y $H_{\theta\theta}$ independientemente de si estamos en régimen estacionario o no estacionario:

$$\hat{V}_{\theta\theta} = \frac{1}{n}\sum_{t=1}^{n} s_{\theta,t}(\hat{\theta}) s_{\theta,t}(\hat{\theta})', \hat{H}_{\theta\theta} = \frac{1}{n}\sum_{t=1}^{n} h_{\theta\theta,t}(\hat{\theta})$$

Esto implica que las versiones de muestra estándar de las estadísticas t y las estadísticas de la razón de (cuasi) verosimilitud siguen las mismas distribuciones asintóticamente ya sea que x_t sea estacionario o no.

3.3.1. Caso estacionario

Los estimadores son globalmente consistentes bajo las siguientes condiciones con F_t denotando la filtración natural:

Supuesto 1

(i) $\{(\varepsilon_t, x_t)\}$ estacionario y ergódico con $\mathbb{E}[\varepsilon_t|F_{t-1}] = 0$ y $\mathbb{E}[\varepsilon_t^2|F_{t-1}] = 1$

(ii) $\mathbb{E}[(log(\alpha_0\varepsilon_{t-1}^2 + \beta_0)] < 0$ y $\mathbb{E}[x_t^{2q}] < \infty$ para todo $0 < q < \infty$.

(iii) $\Theta = \left\{\vartheta : \underline{\omega} \leq \omega \leq \bar{\omega}, 0 \leq \alpha \leq \bar{\alpha}, 0 \leq \beta \leq \bar{\beta}, 0 \leq \pi \leq \bar{\pi}\right\}$

donde $0 < \underline{\omega} \leq \bar{\omega} < \infty, \bar{\alpha}\ \beta < 1$ y $\bar{pi} < \infty$. El valor verdadero, $\vartheta_0\ \epsilon\ \Theta$ con $(\alpha_0, \pi_0) \neq (0, 0)$.

(iv) Para todo $a, b\ \epsilon|\mathbb{R} : a\varepsilon_t^2 + bx_t^2|F_{t-1}$, tiene una distribución no dejenerada.

El supuesto $1(i)$ es una generalización de las condiciones encontradas en Escanciano (2009) quien deriva el propiedades asintóticas de QMLE para procesos GARCH puros (es decir, no hay covariables exógenas incluidos) con errores de diferencia de martingala. Es idéntico al supuesto impuesto en Kristensen y Rahbek (2005).

Las condiciones de momento en el Supuesto $1(ii)$ implican que una solución estacionaria de las ecs. (1)-(2) en el verdadero valor del parámetro existe y tiene un momento polinómico finito, c.f. Lema 1 siguiente. Aquí se permiten procesos GARCH integrados ($\alpha + \beta = 1$). Se trabajará bajo el supuesto implícito de que hemos observado la solución estacionaria.

La condición de compacidad en el Supuesto $1(iii)$ debe ser posible debilitar siguiendo el argumentos de Kristensen y Rahbek (2005). En este caso no se puede eliminar la restricción en el espacio de parámetros de $\beta < 1$, ya que esto conducirá a $\sigma_t^2(\vartheta)$, siendo no-estacionario. El requisito de que $(\alpha_0, \pi_0) \neq (0, 0)$ es necesario para asegurar la identificación de β_0 ya que en el caso donde $(\alpha_0, \pi_0) \neq (0, 0)$, $\sigma_t^2 = \sigma_t^2(\vartheta_0) \rightarrow_{a.s.} \omega_0/(1 - \beta_0)$ por lo que no se podrían identificar conjuntamente ω_0 y β_0.

La condición de no degeneración del Supuesto $1(iv)$ también es necesaria para la identificación. Se descarta la colinealidad (dinámica).

La condición de no degeneración en el Supuesto 1(iv) también es necesaria para la identificación. Se descarta colinealidad (dinámica) entre y_{t-1}^2 y x_t^2.

Para derivar las propiedades asintóticas de $(\hat{\vartheta})$, se establecen algunos resultados preliminares. El primer lema establece que existe una solución estacionaria para el modelo en los valores verdaderos de los parámetros:

Lema 1. Bajo el Supuesto 1: existe una solución estacionaria y ergódica para las ecuaciones (1)-(2) en ϑ_0, que satisface Que satisface $\mathbb{E}[\sigma_t^{2s}] < \infty$ y $\mathbb{E}[y_t^{2s}] < \infty$ para todo $0 < s < 1$.

A continuación, se demostrará que para cualquier valor en el espacio de parámetros, la definición de recursividad $\sigma_t^2(\vartheta)$ posee una solución estacionaria:

Lema 2. Bajo el Supuesto 1: Para todo $\vartheta \in \Theta$, ecuación (5) tiene una solución ergódica estacionaria dada por:

$$\sigma_{0,t}^2(\vartheta) = \sum_{i=0}^{\infty} \beta_{\omega_{t-i}}^i(\vartheta), \omega_t = \omega + \alpha y_{t-1}^2 + \pi x_{t-1}^2$$

Satisface $\mathbb{E}[sup_{\vartheta\in\Theta}\sigma_{0,t}^{2s}(\vartheta)] < \infty$ con $s > 0$ dado en el Lema 1.

Finalmente, se demuestra que el valor inicial elegido para $\sigma_t^2(\vartheta)$ es asintóticamente irrelevante tal que $\sigma_t^2(\vartheta)$ es asintóticamente equivalente de primer orden a la solución estacionaria correspondiente:

Lema 3. Bajo el Supuesto 1: Con $s > 0$ se tiene en el Lemma 1, existe algún $K_s < 1$ tal que:

$$\mathbb{E}\left[\sup_{\vartheta\in\Theta} |\sigma_t^2(\vartheta) - \sigma_{0,t}^2(\vartheta)|^s\right] \leq K_s\beta^{st}$$

Teorema 4 Bajo el Supuesto 1 el QMLE en la ecuación (6) es consistente. Habiendo demostrado que el QMLE es consistente, se verifica que ML1-ML2 bajo el siguiente adicional suposición:

Supuesto 2

(i) $k_4 = \mathbb{E}[(\varepsilon_t^2 - 1)^2] < \infty$

(ii) ϑ_0 esta contenido en Θ

El supuesto 2(i) se utiliza para demostrar que existe la varianza de la puntuación. Se necesita el supuesto 2(ii) para emplear los argumentos de expansión de Taylor descritos anteriormente al derivar la distribución asintótica.

Para verificar ML1-ML2 indicado anteriormente, el siguiente lema resulta útil. Básicamente muestra que las derivadas del proceso de volatilidad $\sigma_t^2(\vartheta)$ tienen propiedades similares a las de

$\sigma^2_{0,t}(\vartheta)$ Las recursiones que definen las derivadas tienen soluciones estacionarias con momentos adecuados:

Lema 5. Bajo los Supuestos 1-2: Los modelos que definen $\partial\sigma^2_t(\vartheta)/(\partial\vartheta)$ y $\partial^2\sigma^2_t(\vartheta)/(\partial\vartheta\partial\vartheta')$ tienen soluciones estacionarias y ergódicas que se denotan $\partial\sigma^2_{0,t}(\vartheta)/(\partial\vartheta)$ y $\partial^2\sigma^2_{0,t}(\vartheta)/(\partial\vartheta\partial\vartheta')$.

3.3.2. Proceso estocástico estacionario

Un proceso estocástico es un proceso matemático que describe una evolución aleatoria a lo largo del tiempo. En otras palabras, un proceso estocástico es una sucesión de eventos aleatorios que se describen mediante un conjunto de reglas matemáticas. Estos procesos son ampliamente utilizados en una variedad de campos, como la física, las ciencias económicas y financieras, la ingeniería, la medicina, entre otros.

Un proceso estocástico estacionario es un tipo especial de proceso estocástico en el que las propiedades estadísticas (como la media y la varianza) son constantes a lo largo del tiempo y no dependen de la posición en el tiempo. En otras palabras, si se toma una muestra del proceso en cualquier momento, se tendrá la misma distribución de probabilidad para esa muestra.

Capítulo 4

Análisis estadístico de los datos

4.1. Selección de variables

Según Pérez-Magariño et al. (2004), cuando existen muchas variables es posible que parte importante sea redundante, en cuyo caso es necesario eliminar el exceso y dejar sólo aquellas que tengan representatividad dentro del conjunto. Esto se consigue con la aplicación de las técnicas multivariantes de reducción de la dimensión.

En este caso se realiza un análisis de correlación entre las variables, junto con un análisis de componentes principales (ACP), con el objeto de excluir aquellas que se encuentren fuertemente correlacionadas, para así disminuir la dimensión de las variables de ingreso al modelo. Luego de seleccionar las variables se prepara la serie de datos a utilizar para entrenar los modelos.

4.2. Correlación de las variables

Algunos modelos (LM, GLM entre otros) se ven perjudicados si incorporan predictores altamente correlacionados. Por esta razón, es conveniente estudiar el grado de correlación entre las variables disponibles (4.1).

4.3. Análisis de Componente Principales (ACP)

El análisis de componentes principales (ACP) es una técnica estadística multivariante utilizada para reducir la dimensionalidad de un conjunto de datos mediante la identificación de patrones subyacentes en los datos. Se busca transformar las variables originales en un conjunto de variables no correlacionadas denominadas çomponentes principales", que se ordenan según la cantidad de varianza que explican en los datos.

Desarrollado originalmente por Pearson (1901), el Análisis de Componentes Principales (ACP), es un procedimiento matemático que transforma un número de variables posiblemente correlacionadas.

Se considera una serie de variables $x^T = (x_1, x_2, ..., x_m)$ sobre un grupo de objetos o individuos y se trata de calcular, a partir de ellas, un nuevo conjunto de variables $y^T = (y_1, y_2, ..., y_m)$ no

Tabla 4.1: Matriz de correlación variables ambientales zona centro sur

MATRIZ DE CORRELACIÓN

	TA	TSMT	TSM NOAA	NMM (m)	TE (CS)	IT (CS)	PDO	NINO12	NINO34	IOS	ILF	AAO	HCI
TA	1,000	0,682	0,932	-0,120	0,342	0,591	0,013	0,523	-0,225	0,023	-0,013	0,027	-0,075
TSMT	0,682	1,000	0,685	-0,012	0,130	0,340	0,176	0,585	0,060	-0,124	0,153	0,001	-0.037
TSM NOAA	0,932	0,685	1,000	-0,018	0,320	0,556	0,131	0,669	-0,140	-0,013	0,011	0,049	-0,167
NMM (m)	-0,120	-0,012	-0,018	1,000	0,151	-0,450	0,260	0,238	0,334	-0,291	0,220	-0,084	-0,130
TE (CS)	0,342	0,130	0,320	0,151	1,000	-0,115	0,093	0,018	-0,128	-0,092	0,022	-0,143	-0,298
IT (CS)	0,591	0,340	0,556	-0,450	-0,115	1,000	-0,030	0,183	-0,213	0,042	0,011	0,078	-0,056
PDO	0,013	0,176	0,131	0,260	0,093	-0,030	1,000	0,280	0,447	-0,406	0,442	-0,122	-0,313
NIÑO12	0,523	0,585	0,669	0,238	0,018	0,183	0,280	1,000	0,384	-0,178	0,292	-0,016	0,062
NIÑO34	-0,225	0,060	-0,140	0,334	-0,128	-0,213	0,447	0,384	1,000	-0,661	0,831	-0,184	-0,050
IOS	0,023	-0,124	-0,013	-0,291	-0,092	0,042	-0,406	-0,178	-0,661	1,000	-0,720	0,142	0,234
ILF	-0,013	0,153	0,011	0,220	0,022	0,011	0,442	0,292	0,831	-0,720	1.000	-0,191	-0,084
AAO	0,027	0,001	0,049	-0,084	-0,143	0,078	-0,122	-0,016	-0,184	0,142	-0,191	1,000	0,116
HCI	-0,075	-0,037	-0,167	-0,130	-0,298	-0,056	-0,313	0,062	-0,050	0,234	-0,084	0,116	1,000

correlacionadas entre sí, cuyas varianzas vayan decreciendo progresivamente.

Cada $y_j(i = 1, ..., m$ es una combinación lineal de las variables $(x_1, x_2, ..., x_p)$ originales, es decir:

$$y_i = a_{1i}x_1 + a_{2i}x_2 + ... + a_{mi}x_m, i = 1, 2, ..., m,$$

donde a es el peso correspondiente y x_j es la j-ésima variable de la serie original. El primer componente principal y_1 tiene la mayor varianza, $\lambda_1 = Var(y_1)$, el segundo componente principal contiene la proporción más alta de la varianza restante, $\lambda_2 = Var(y_2)$, y así, sucesivamente. De esta forma, un número p de componentes principales, pueden retener la mayor parte de la varianza presente en las q variables originales, donde $p < m$. Así, la reducción de dimensionalidad de una serie de datos multivariados con q variables, puede ser alcanzada mediante una proyección en un subespacio p-dimensional. Este subespacio puede ser construido mediante una matriz de valores propios a partir de la matriz de covarianza de $\sum X$. De acuerdo a la decomposición de valores singulares, la matriz de covarianza puede ser expresada de la siguiente forma:

$$\sum = U\Lambda U^T$$

donde U contiene los pesos de los componentes principales y la matriz Λ contiene las varianzas correspondientes, así, el primer vector columna de U es el vector de pesos del primer componente principal y su varianza es el primer elemento de la matriz Λ a lo largo de la diagonal. Para reducir la dimensionalidad de la serie de datos multivariados, los primeros p componentes principales son retenidos, los primeros p vectores propios son utilizados para construir el nuevo subespacio $A_mxp = U(1 : m; 1 : p)$. Los componentes pueden ser formados como se indica a continuación:

$$Y_{mxp} = X_{mxp}A_{mxp}$$

,

donde n es es el largo de la serie de datos multivariada, con $p < myp < n$. De esta forma, se puede transformar una serie de datos multivariada de tamaño nxm en una representación reducida, de tamaño nxp.

A partir del ACP se generan 13 factores que su conjunto contienen el 100 % de la varianza. Tablas 4.2 y 4.3.

Tabla 4.2: Valores propios calculados a partir del ACP de las variables ambientales.

Componente	Valor propio	% Varianza	% Acumulado
1	**3.571**	**27,472**	**27,472**
2	3,162	24,322	51,795
3	1,497	11,515	63,31
4	1,184	9,11	72,42
5	0,939	7,223	79,642
6	0,744	5,72	85,363
7	0,557	4,282	89.645
8	0,441	3,39	93,035
9	0,392	3,018	96,052
10	0,254	1,952	98,004
11	0,145	1,115	99,12
12	0,081	0,619	99,739
13	0,034	0,261	100
Promedio	1		

Tabla 4.3: Correlación de cada variable con cada componente.

	Componente			
	1	2	3	4
TA	**0,908**	-0,135	-0,212	-0,166
TSMT	**0,817**	0,131	-0,053	0,024
TSM-NOAA	**0,948**	-0,059	-0,120	-0,174
NMM (m)	0.061	0,250	0,822	-0,109
TE (CS)	0,265	-0,157	0,288	-0,712
IT (CS)	0.483	-0,014	-0,774	0,009
PDO	0,140	0,599	0,115	-0,295
NIÑO12	0,780	0,333	0,235	0,231
NIÑO34	-0.040	0.911	0.215	0,127
IOS	-0,024	-0,813	-0.044	0,208
CTI	0,068	0,908	0,003	-0.026
AAO	0.087	-0,237	-0,020	0,366
HCI	-0.010	-0,156	0,085	0,792

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

Considerando los resultados obtenidos de la matriz de correlación entre las variables ambientales y los resultados del ACP (Tablas 4.1, 4.2 y 4.3), del cual se consideran las variables cuyo valor de correlación con la componente es mayor a $0,7$ (HAIR et al., 1999) (Tabla 4.3), se selecciona la variable ambiental TSM-NOAA.

4.4. Variables del modelo GARCH-X y estacionariedad

La estacionariedad es una suposición clave en el modelo GARCH . Según Bollerslev (1986), dado que el modelo GARCH se basa en la hipótesis de que la varianza de la serie temporal es condicionalmente constante, lo que significa que la varianza es constante en promedio a lo largo del tiempo. Si los datos no cumplen con esta suposición, los resultados del modelo GARCH pueden no ser fiables.

Para verificar si una serie temporal es estacionaria, se pueden realizar pruebas estadísticas como la prueba de Dickey-Fuller. Si la serie temporal es estacionaria, entonces el modelo GARCH puede ser aplicado con confianza. Si no es estacionaria, se pueden aplicar técnicas de diferenciación para transformar la serie temporal en una serie estacionaria y luego aplicar el modelo GARCH.

A continuación se presentan las gráficas de las series incluidas en el modelo, las cuales de manera

visual, permiten inferir la estacionariedad de ambas.

4.4.1. Serie de los desembarques de anchoveta

La estacionalidad de los desembarques de anchoveta es una de las principales caracteristicas de esta pesquería, en efecto, L. Cubillos et al. (2001) y Castillo-Jordán et al. (2010) determinan una estacionalidad importante en el crecimiento de la anchoveta, aspecto que tendría un impacto importante en el inicio de la temporada de pesca sobre estos recursos, que comienza en la época estival. Esta especie es fuertemente dependiente de los ciclos estacionales asociados a los eventos de surgencia en la zona centro-sur de Chile (L. Cubillos et al., 2001).

A continuación se presentan la gráfica de los desembarques y su tendencia, junto con la gráfica de boxplot mensual y la estacionalidad interanual. Figuras 4.1, 4.2, 4.3.

4.4.2. Serie de la Temperatura superficial del mar

La temperatura superficial del mar en la zona centro-sur de Chile tiene una estacionalidad pronunciada, con valores máximos en verano y mínimos en invierno, de manera inter-anual esta es influenciada principalmente por el fenómeno del Niño, que se manifiesta en forma de anomalías térmicas durante los meses de verano y otoño, y por el fenómeno del Niña, que se manifiesta en forma de anomalías térmicas durante los meses de invierno y primavera (González y Cortés, 2005).

A continuación se presentan la gráfica de la temperatura superficial del mar y su tendencia, junto con la gráfica de boxplot mensual y la estacionalidad interanual. Figuras 4.4, 4.5, 4.6.

A continuación se presentan la gráfica de los desembarques y su tendencia, junto con la gráfica de boxplot mensual y la estacionalidad interanual de los desembarques de anchoveta. Figuras 4.4, 4.5, 4.6.

4.5. Prueba de estacionariedad Dickey Fuller

Esta prueba fue popularizada por Dickey y Fuller (1979), la hipótesis nula implica que hay una raíz unitaria en un modelo AR, lo cual implica que la serie es no estacionaria. La hipótesis alternativa es generalmente estacionariedad. De esta manera se puede definir como:

- H_0: Los datos son no estacionarios.
- H_1: Los datos son estacionarios.

El test estadístico es:

$$Z_t = \rho Z_{t-1} + \xi_t, t = 1, 2, ..., n \tag{4.1}$$

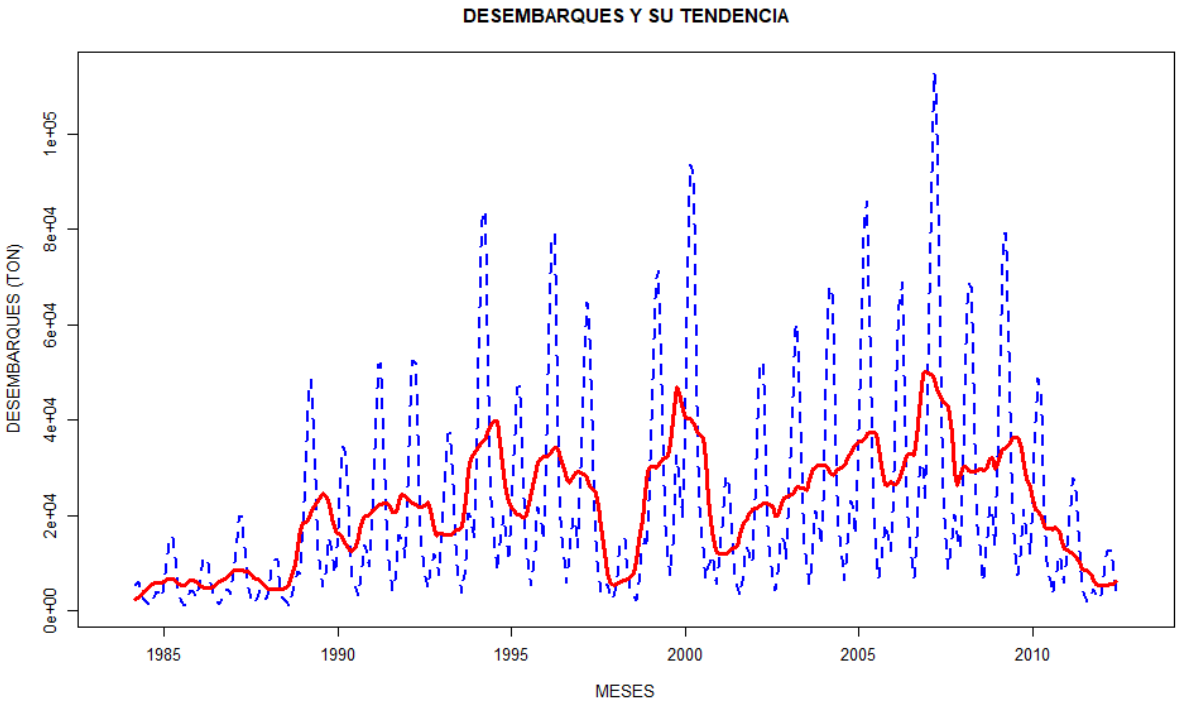


Figura 4.1: Desembarques de anchoveta y su tendencia.

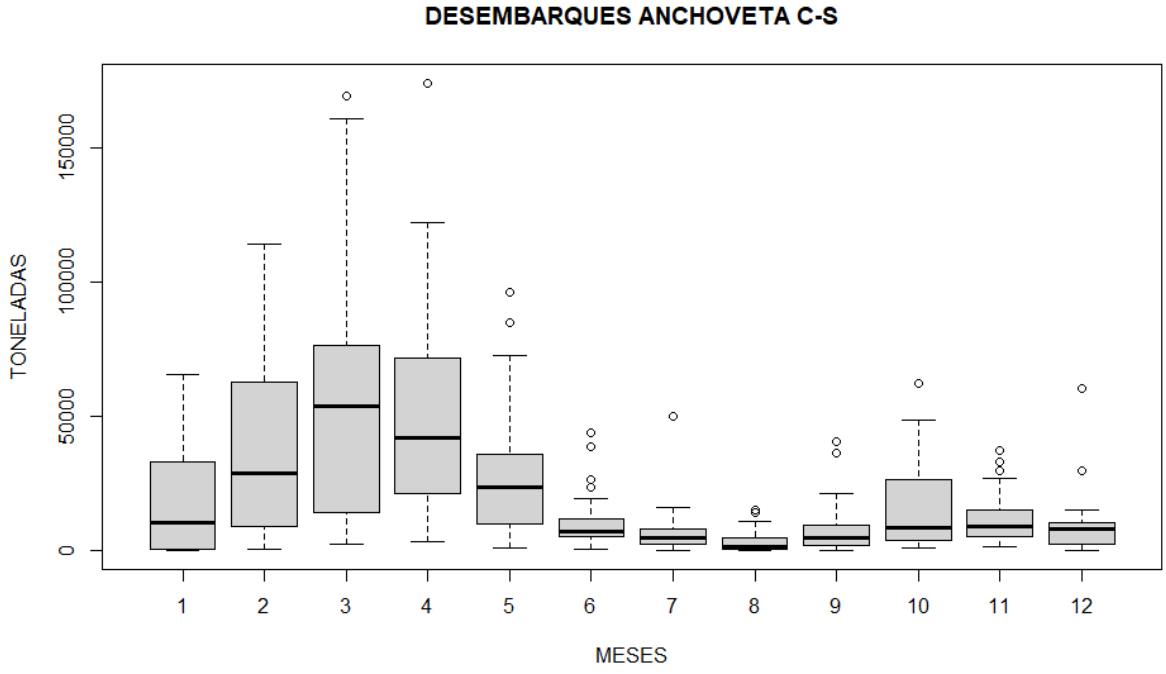


Figura 4.2: Boxplot de los desembarques mensuales de anchoveta.

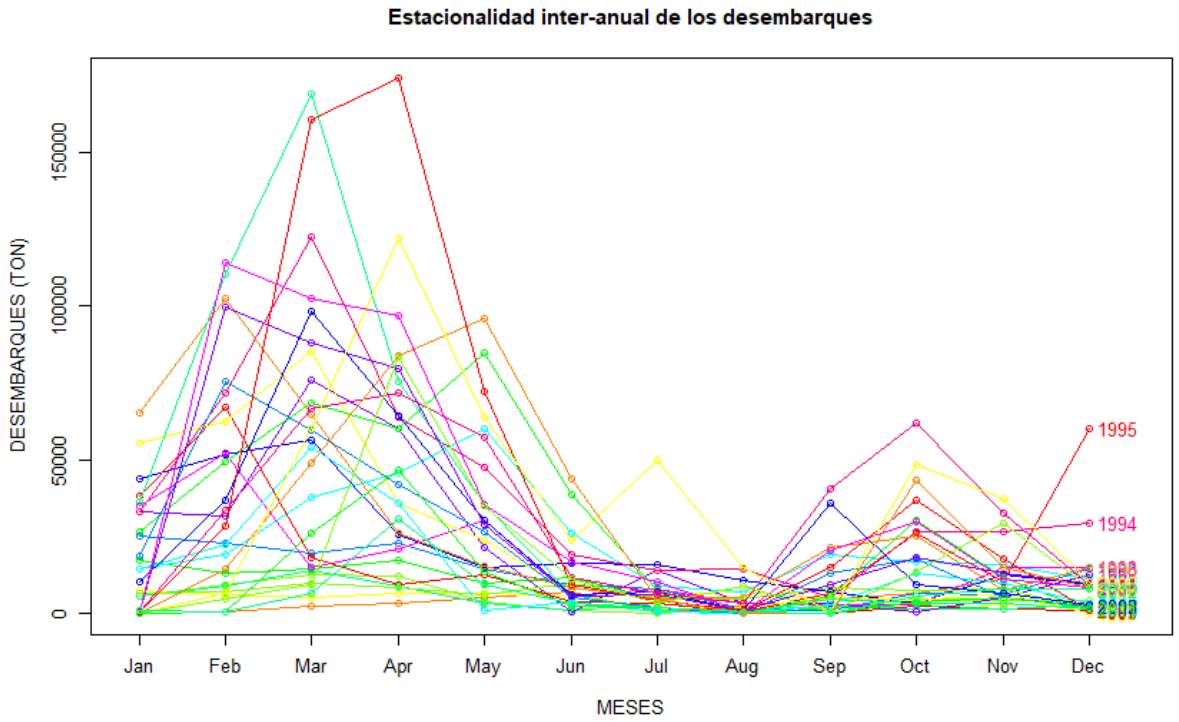


Figura 4.3: Estacionalidad inter-anual de los desembarques de anchoveta.

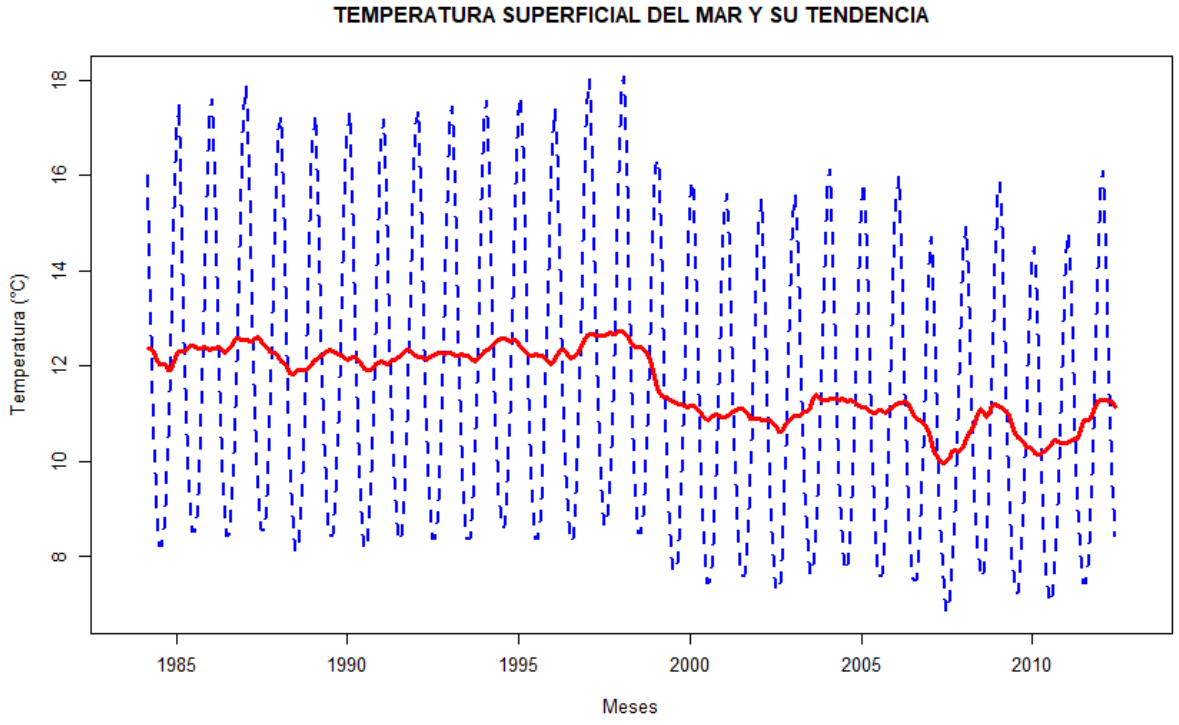


Figura 4.4: Temperatura superficial del mar y su tendencia.

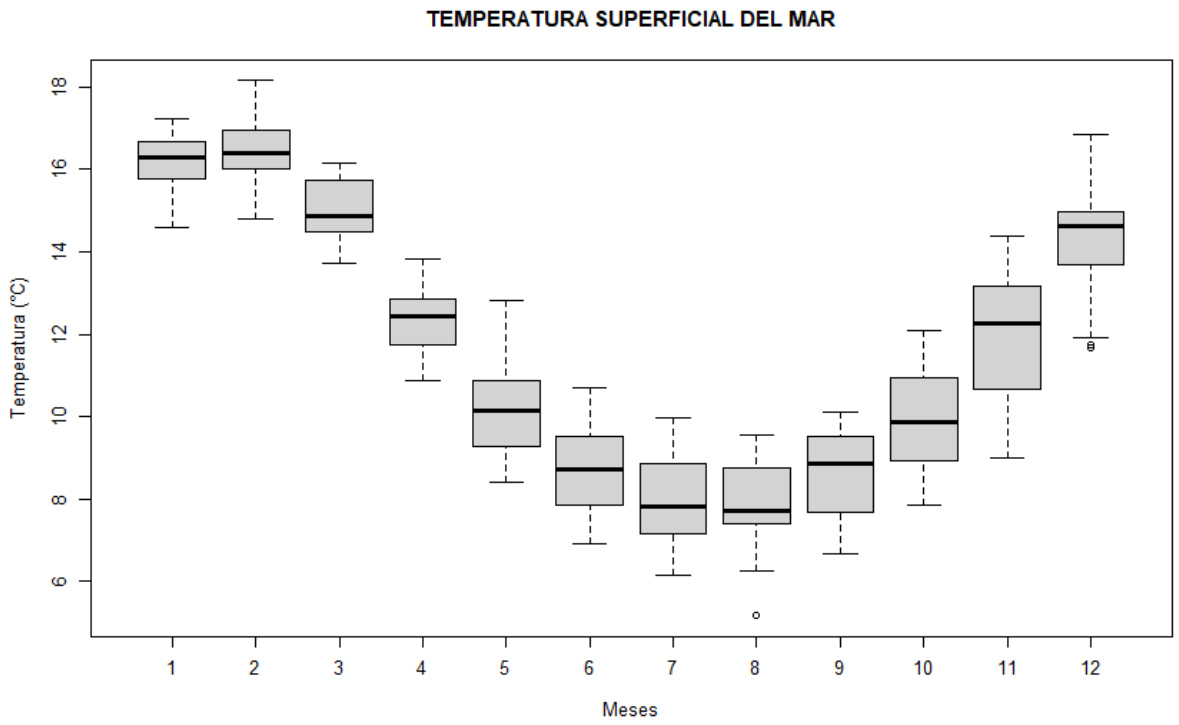


Figura 4.5: Boxplot de la de la temperatura superficial del mar.

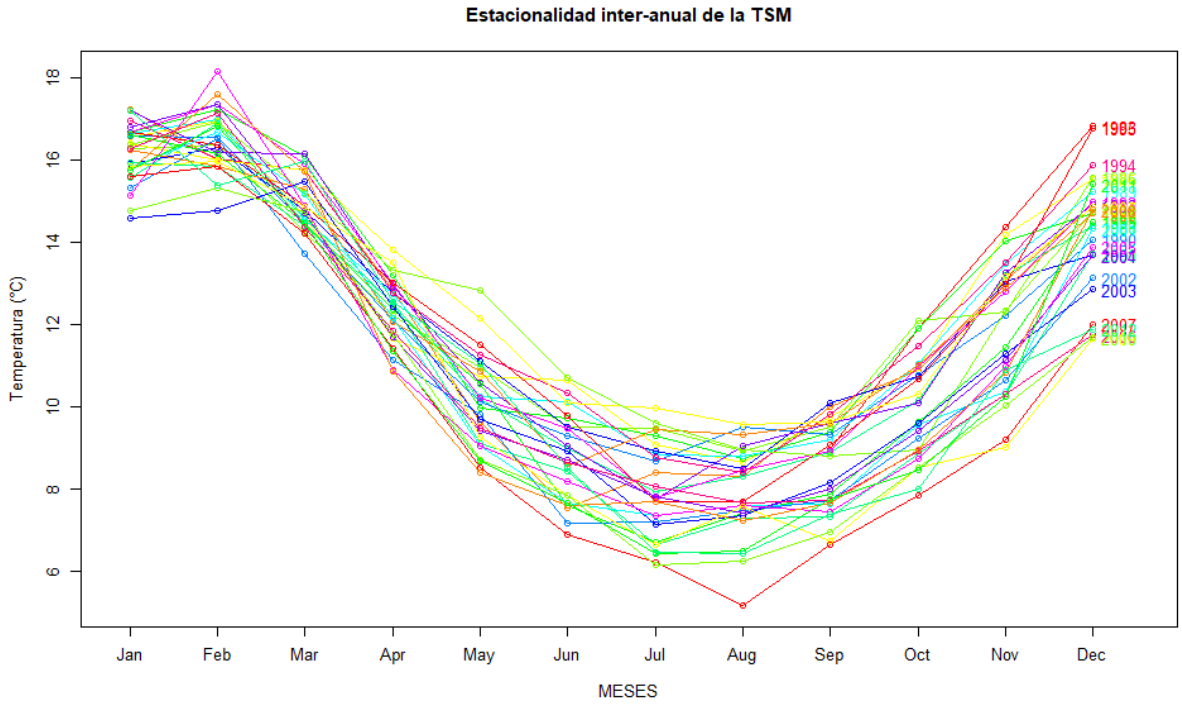


Figura 4.6: Estacionalidad inter-anual de la temperatura superficial del mar.

Donde $Z_0 = 0$, ρ es un número real, y ξ_t es una secuencia de variables aleatorias normales independientes con media cero y varianza σ^2 4.1.

Dada las n observaciones $Z_1, Z_2, ..., Z_n$ el estimador de máxima verosimilitud de ρ es:

$$\widehat{\rho} = (\sum_{t=1}^{n} Z_{t-1}^2)^{-1} \sum_{t=1}^{n} Z_t Z_{t-1} \tag{4.2}$$

La ecuación 4.2 muestra que $\widehat{\rho}$ es un estimador consistente para todos los valores de ρ.

La ecuación 4.1 representa un proceso AR de primer orden ($AR(1)$) implica que Z_t depende de su pasado solamente. ρ es el coeficiente de interés para la determinación de raíz unitaria o estacionariedad. Si el valor de $\rho < 1$, los choques en Z_t serán transitorios y desaparecerán con el pasar del tiempo. En este caso, la serie será de naturaleza estacionaria. Por el contrario, si el valor de $\rho = 1$ los choques en la serie de datos Z_t son permanentes en naturaleza y no desaparecen con el pasar del tiempo, en este escenario la serie es caracterizada como una raíz unitaria. Si $\rho > 1$ la serie será explosiva.

El problema de la ecuación que se plantea es que supone la presencia de estacionariedad, que es la razón de porque la ecuación se estima en niveles. Sin embargo, la teoría sugiere que se debe partir de la duda de estacionariedad, es decir, que la serie sea raíz unitaria. Si partimos de ese supuesto entonces debemos reescribir el modelo de la siguiente forma.

$$(Z_t - Z_{t-1}) = (\rho - 1)Z_{t-1} + u_t, t = 1, 2, ..., n$$
$$\triangle Z_t = \delta Z_{t_1}, t = 1, 2, ..., n$$

La suposición inicial es que la serie es no estacionaria y es por eso que el modelo se estima de forma diferente. Aquí (δ) es el coeficiente de interés para determinar raíz unitaria o estacionariedad. Si el valor de $\delta < 1$, los choques en $\triangle Z_t$ serían transitorios y desaparecerán con el pasar del tiempo. En este caso, la serie sería de naturaleza estacionaria. Por el contrario, si el valor de $\delta = 1$ los choques en la serie de datos $\triangle Z_t$ son permanentes en naturaleza y no desaparecen con el pasar del tiempo, en este escenario la serie es caracterizada como una raíz unitaria. Si $\delta > 1$ la serie sería explosiva.

Al aplicar el test de Dickey Fuller a las series de desembarque y temperatura superficial del mar, dado que el p-value, para ambos casos es $0,01$ se rechaza la hipótesis nula, H_0, que existe raíz unitaria, por lo tanto ambas series son estacionarias Tabla (4.4).

4.6. Estimación del modelo GARCH-X para los desembarques de anchoveta

Las variables de entrada al modelo son los desembarques mensuales de anchoveta en la zona centro sur de Chile en toneladas y la temperatura superficial del mar captada por satélites de la NOAA en °C.

Tabla 4.4: Test de Dickey Fuller aplicado a las series de desembarque y temperatura superficial del mar.

Test de Dickey Fuller		
Variable	Desembarques	TSM-NOAA
Dickey-Fuller	-8,7523	-5,6987
Lag order	0	0
p-value	0,01	0,01

Dado que las variables de estudio poseen distintas unidades, estas entran al modelo estandarizadas, al estandarizar las variables, estas se convierten en valores adimensionales (sin unidad) y se eliminan las diferencias en la escala y la unidad, lo que permite una comparación más justa entre ellas, esto ayuda a controlar los efectos de los cambios en las unidades de medida o los cambios en la escala. Al tener unidades de medida diferentes pueden tener un peso desproporcionado, lo que puede sesgar los resultados del análisis (F. Hair Jr et al., 2014; Stevens, 2009; Rupert et al., 2003). De esta manera, la estandarización de la variable implica que posea una media igual a cero, una desviación estándar igual a uno, una distribución normal, valores adimensionales y un coeficiente de correlación invariable, así las variables cumplen con las condiciones de entrada al modelo GARCH-X.

Los modelos GARCH con variables exógenas o GARCH-X están implementados en la librería garchx versión 1.1 (Sucarrat, 2020), que tiene como objetivo proporcionar un entorno de prueba y modelado GARCH univariante flexible.

En las Tablas 4.5 y 4.6 se pueden observar los resultados obtenidos al aplicar los modelos GARCH y GARCH-X en donde la variable endógena es los desembarques mensuales de anchoveta (t) y la exógena la TSM-NOAA (°C). Al comparar el ajuste de ambos modelos, se verifica que el modelo GARCH con variable exógena presenta valores ajustados de mayor significancia, lo que demuestra la incidencia de la la TSM-NOAA en los desembarques de anchoveta, los demás parámetros también son significativos y además $\omega, \alpha, \beta > 0$ y la suma de estos parámetros es menor que la unidad, lo que cumple la estacionariedad en covarianza, que es uno de los supuestos del modelo.

Tabla 4.5: Parámetros óptimos modelo GARCH.

Parámetros óptimos para el modelo GARCH univariado				
Parámetro	Estimación	Error Est.	t-calculado	P-Valor
$\alpha(arch1)$	0,51628	0,31433	1,64250	0,05024
$\beta(garch1)$	0,00000	0,54172	0,00000	0,50000
Fuente: Elaboración propia				

En las ecuaciones 4.3 y 4.4, se muestran las ecuaciones de los modelos ajustados GARCH y GARCH-X respectivamente, en donde se aprecia el error estándar de los parámetros estimados para ambos modelos, junto con los valores de Log-verosimilitud. Se verifica un mejor ajuste del modelo GARCH-X, a partir del valor de estos parámetros respectivamente.

Tabla 4.6: Parámetros óptimos modelo GARCH-X.

Parámetros óptimos para el modelo GARCH-X				
Parámetro	Estimación	Error Est.	t-calculado	P-Valor
$\alpha(arch1)$	0,47986	0,20056	2,39266	0,00836
$\beta(garch1)$	0,00000	0,19729	0,00000	0,50000
$\pi(xreg1)$	0,36255	0,10344	3,50493	0,00023
Fuente: Elavoración propia				

Modelo GARCH univariado:

$$\begin{aligned} \widehat{\sigma}_t^2 &= \underset{(0,3333)}{0,4778} + \underset{(0,3143)}{0,5163}\alpha + \underset{(0,5417)}{0,0000}\beta \\ &\text{Log-verosimiltud: } -446,9312 \qquad T = 351 \end{aligned} \tag{4.3}$$

Modelo GARCH con variable exógena:

$$\begin{aligned} \widehat{\sigma}_t^2 &= \underset{(0,0893)}{0,1375} + \underset{(0,2006)}{0,4799}\alpha + \underset{(0,1973)}{0,0000}\beta + \underset{(0,1034)}{0,3625}\pi \\ &\text{Log-verosimiltu: } -426,5353 \qquad T = 351 \end{aligned} \tag{4.4}$$

En su forma más general, el modelo GARCH-X se puede escribir como:

$$y_t = \mu_t + \epsilon_t$$

$$\epsilon_t = \sigma_t \cdot z_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{k=1}^{m} \gamma_k X_{kt}^2$$

$$f(P) = \frac{e(P)}{\sum P' \epsilon P^e(P')}; P \epsilon \rho$$

donde:

y_t es la variable dependiente en el período de tiempo t.

μ_t es la media condicional de la variable y_t.

ϵ_t es el error condicional en el período de tiempo t.

σ_t es la desviación estándar condicional de ϵ_t en el período de tiempo t.

z_t es una variable aleatoria con media cero y varianza uno.

ω es la varianza constante de los errores.

α_i son los coeficientes de la parte ARCH del modelo que miden el impacto de los errores pasados en la volatilidad actual.

β_j son los coeficientes de la parte GARCH del modelo que miden el impacto de la varianza pasada en la volatilidad actual.

γ_k son los coeficientes de las variables exógenas (X) que miden el impacto de las variables exógenas en la volatilidad actual.

X_{kt} es el valor de la variable exógena k en el período de tiempo t. p, q, y m son los órdenes del modelo ARCH, GARCH y de variables exógenas, respectivamente.

En resumen, la ecuación del modelo GARCH-X describe cómo la volatilidad de la serie de tiempo se relaciona con la varianza pasada, los errores pasados y las variables exógenas. Con esta ecuación, es posible estimar los coeficientes del modelo y, por lo tanto, predecir la volatilidad futura de la serie de tiempo.

Capítulo 5

Conclusiones

El presente trabajo plantea implementar un modelo GARCH con variable exógena para los desembarques mensuales de anchoveta en la zona centro sur de Chile, con el objeto de predecir las fluctuaciones de abundancia de esta pesquería, a partir de una base ambiental y pesquera multivariada. Para tal fin, se realizó un análisis exploratorio y de selección entre las variables, para determinar la variable exógena a incorporar.

Con el objeto de comparar el grado de ajuste de los parámetros de interés, se realizaron dos modelos, uno sin variable exógena y otro con variable exógena. La variable seleccionada, fue la temperatura superficial del mar captada por satélites NOAA. Es importante destacar desarrollar la incorporación de técnicas no lineales, para verificar si estas relaciones mejorarían la selección del a variable y por ende el ajuste del modelo.

La implementación del modelo GARCH-X del presente trabajo en la pesquería de anchoveta, en la zona cetro sur de Chile, brinda información sustancial para la toma de decisiones del manejo pesquero. Que en conjunto con otros enfoques metodológicos, y el conocimiento de expertos, permite a obtener una mayor comprensión de la dinámica de este recurso que se caracteriza por un alta variabilidad ante los cambios del ambiente.

Capítulo 6

Discusión

El calentamiento global provocará cambios significativos en la biología de los océanos (Boyd y Doney, 2002; **?**), con grandes impactos para las pesquerías y las comunidades humanas a diferentes escalas (Roessig et al., 2004; R. Perry y Barange, 2009; R. I. Perry et al., 2010). Estos efectos del calentamiento global en las condiciones del océano tendrán un impacto importante en la historia y ciclo de vida de los peces, afectando la reproducción, y la distribución espacial en sus diferentes etapas de desarrollo (J. A. Hare et al., 2010; A. L. Perry et al., 2005; R. Perry y Barange, 2009; Merino et al., 2012). Los modelos de excedente de producción convencionales empleados en evaluación de stock normalmente utilizan sólo una variable explicatoria, el esfuerzo de pesca (Fréon y Yáñez, 1995). Estos modelos no siempre son adecuados, porque las variaciones del esfuerzo de pesca explican sólo una parte de la variabilidad de los desembarques. A menudo la variación residual es originada por fenómenos ambientales que afectan la abundancia y/o la capturabilidad del stock de un año a otro (Fréon y Yáñez, 1995). Por lo tanto, las predicciones de la abundancia de especies marinas en escenarios futuros son herramientas importantes para explorar los impactos potenciales del cambio climático en los servicios de los ecosistemas, como en las pesquerías de los pequeños pelágicos. El estudio de la selección de variables es una de las decisiones más importantes en la investigación empírica según Hair Jr et al. (2017). Además, mediante el análisis de correlación, la selección de variables puede ser determinada por la teoría, la experiencia previa o técnicas de selección de variables basadas en modelos estadísticos.

En el norte de Chile, Yáñez et al. (2010a), estudiaron la predicción de abundancia de anchoveta y sardina española, utilzando como una de variables exógenas la TSM. Cabe señalar que Plaza et al. (2008), Plaza et al. (2018), Gutiérrez-Estrada et al. (2007), Gutiérrez-Estrada et al. (2009); Yáñez et al. (2010b), Yáñez et al. (2017), (Naranjo et al., 2015) realizaron analisis de regresion y componentes principales para reducir las variables de entrada en modelos RNAs, obteniendo importantes resultados, en donde también fue seleccionada la TSM como variable de entrada a los modelos. Por otra parte, Canales et al. (2020) encontraron que la TSM fue el principal impulsor de la dinámica poblacional de la anchoveta. En la primera etapa de selección del presente estudio se realizó el análisis de correlación entre variables ambientales, donde se pudo observar fuertes relaciones entre estas, por lo cual se descartó del análisis la TA y TSM, por estar altamente correlacionada con la TSM-NOAA, $0,93$ y $6,685$ respectivamente. Respecto al análisis de componentes principales, y en función del análisis de correlación, se seleccionó la variable TSM-NOAA, dado que esta es la que presenta el mayor valor de correlación con la componente 1 $(0,948)$, la cual explica un $27,5\,\%$ de la variabilidad total de los datos. Sin embargo, Gutiérrez-Estrada et al. (2007) sugieren la inclusión de relaciones no lineales en la selección de variables de entrada a los modelos

La anchoveta es fuertemente dependiente de los ciclos estacionales asociados a los eventos de surgencia en la zona centro-sur de Chile (L. Cubillos et al., 2001). Su estacionalidad de los desembarques es una de las principales caracteristicas de esta pesquería. Del mismo modo, la temperatura superficial del mar en la zona centro-sur de Chile tiene una estacionalidad pronunciada, con valores máximos en verano y mínimos en invierno, de manera inter-anual esta es influenciada principalmente por el fenómeno del Niño, que se manifiesta en forma de anomalías térmicas durante los meses de verano y otoño, y por el fenómeno del Niña, que se manifiesta en forma de anomalías térmicas durante los meses de invierno y primavera (González y Cortés, 2005). A través del test de Dickey Fuller (Dickey y Fuller, 1979), se verificó la estacionariedad de ambas variables, premisa importante para la implementación de los modelos GARCH.

Desde el punto de vista de la administración pesquera, sería de mucha utilidad poder predecir los desembarques para el próximo año utilizando enfoques de series de tiempo donde los datos se actualizan anualmente (Czerwinski et al., 2007). Preciado et al. (2006), implementaron en España los modelos ARIMA y función de transferencia para los desembarques de merluza (*Merluccius merluccius*, L., 1758). Las predicciones efectuadas por el modelo ARIMA ajustado a la serie 1983-1997 y las estimaciones del modelo de función de transferencia, considerando la serie 1983-2000, permitieron completar vacíos de la serie, así como detectar errores o cambios en el comportamiento de la flota. La comparación de ambos modelos confirma la estabilidad de la serie, validando las estimaciones efectuadas en los dos últimos años. La aplicación de modelos de heteroscedasticidad condicional (es decir, algún tipo de modelo ARCH/GARCH) en la pesca, se asocia principalmente con los mercados financieros de pescado y la volatilidad de su precio (Bose, 2004, Ko, 2009, Dahl y Oglend, 2014, Yoon y Yoon , 2017). Sin embargo, Lai et al. (2005), obtuvieron ajustes importantes para la predicción de la captura mensual de peces costeros de Australia Occidental, a través de modelos ARIMA y GARCH. Kim et al. (2015) analizaron la captura mensual por unidad de esfuerzo (CPUE) para 1987-2012 utilizando métodos de suavizado exponencial, promedio móvil integrado autorregresivo estacional (SARIMA), redes neuronales autorregresivas (ANN) y modelos de redes neuronales recurrentes autorregresivas (ARNN) para pronosticar la abundancia de anchoveta en función de las condiciones de pesca. Por otra parte, Sanz-Fernández y Gutiérrez-Estrada (2021), evaluaron desde una perspectiva univariante los patrones de variación de los desembarques de besugo (de 1983 a 2016) en el Estrecho de Gibraltar e investigaron los efectos de los factores ambientales y cambios en los desembarques. Utilizaron enfoques univariados estacionales y no estacionales como Holt-Winters, modelos ARIMA y modelos GARCH. Estos modelos fueron capaces de detectar una fuerte dependencia lineal entre los desembarques actuales y los desembarques anteriores, así como los efectos estacionales, mientras que los modelos GARCH indicaron la presencia de una intensa incertidumbre o volatilidad alrededor de dos períodos (1993-1998 y 2007-2011), asociados con posibles efectos de factores biológicos, ambientales-climáticos y regulatorios subyacentes.

En este trabajo, previo a la implementación del modelo GARCH-X para los desembarques de anchoveta en la zona centro-sur de Chile, se ajustó un modelo GARCH, sin variable exógena, para así poder comparar el desempeño de ambos modelos; de esta manera, se verificó que el modelo GARCH con variable exógena presenta valores ajustados de mayor significancia, lo que demuestra la incidencia de la la TSM-NOAA en los desembarques de anchoveta. La implementación del modelos GARCH-X en la pesquería de anchoveta, en la zona cetro sur de Chile, proporcionaría información valiosa para la toma de decisiones en el manejo pesquero. Sin embargo, también es importante tener en cuenta que estos modelos se basan en supuestos simplificados y que es necesario complementarlos con otros enfoques para obtener una comprensión completa de la dinámica de la

pesquería.

Referencias

Alheit, J. (1993). Effects of physical and biological changes on the biomass yield of the humboldt current ecosystem. *Stress, Mitigation and Sustainability*.

Alheit, J., y Bakun, A. (2010). Population synchronies within and between ocean basins: apparent teleconnections and implications as to physical–biological linkage mechanisms. *Journal of Marine Systems*, *79*(3-4), 267–285.

Alheit, J., y Niquen, M. (2004). Regime shifts in the humboldt current ecosystem. *Progress in Oceanography*, *60*(2-4), 201–222.

Alheit, J., y Peck, M. A. (2019). Drivers of dynamics of small pelagic fish resources: biology, management and human factors. *Marine Ecology Progress Series*, *617*, 1–6.

Alheit, J., Roy, C., y Kifani, S. (2009). Decadal-scale variability in populations. *Climate change and small pelagic fish*, 64–87.

Aranis, A., Ahumada, M., Arcos, D., Araya, M., Bravo, L., Candia, A., ... others (2013). *Seguimiento hidroacústico de la biomasa de anchoveta y sardina común en la zona centro-sur de chile durante el año 2012* (Inf. Téc.). Instituto de Fomento Pesquero (IFOP).

Ayón, P., Criales-Hernandez, M. I., Schwamborn, R., y Hirche, H.-J. (2008). Zooplankton research off peru: a review. *Progress in Oceanography*, *79*(2-4), 238–255.

Bertrand, A., Chaigneau, A., Peraltilla, S., Ledesma, J., Graco, M., Monetti, F., y Chavez, F. P. (2011). Oxygen: a fundamental property regulating pelagic ecosystem structure in the coastal southeastern tropical pacific. *PloS one*, *6*(12), e29558.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Bollerslev, T. (1988). On the correlation structure for the generalized autoregressive conditional heteroskedastic process. *Journal of time series analysis*, *9*(2), 121–131.

Box, G. E., y Jenkins, G. M. (1976). Time series analysis: Forecasting and control san francisco. *Calif: Holden-Day*.

Boyd, P. W., y Doney, S. C. (2002). Modelling regional responses by marine pelagic ecosystems to global climate change. *Geophysical Research Letters*, *29*(16), 53–1.

Camiñas, J., y Baro, J. (2004). *Osteictios de andalucía. in book: Proyecto andalucia: Naturaleza osteictios marinos*.

Canales, T. M., Law, R., y Blanchard, J. L. (2016). Shifts in plankton size spectra modulate growth and coexistence of anchovy and sardine in upwelling systems. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(4), 611–621.

Canales, T. M., Lima, M., Wiff, R., Contreras-Reyes, J. E., Cifuentes, U., y Montero, J. (2020). Endogenous, climate, and fishing influences on the population dynamics of small pelagic fish in the southern humboldt current ecosystem. *Frontiers in Marine Science*, *7*, 82.

Carr, M.-E. (2001). Estimation of potential productivity in eastern boundary currents using remote sensing. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(1-3), 59–80.

Castillo-Jordán, C., Cubillos, L. A., y Navarro, E. (2010). Inter-cohort growth rate changes of common sardine (strangomera bentincki) and their relationship with environmental conditions off central southern chile. *Fisheries Research*, *105*(3), 228–236.

Chatfield, K., Simonyan, K., Vedaldi, A., y Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*.

Chavez, F. P., Ryan, J., Lluch-Cota, S. E., y Ñiquen, M. (2003). From anchovies to sardines and back: multidecadal change in the pacific ocean. *science*, *299*(5604), 217-221.

Checkley Jr, D. M., Asch, R. G., y Rykaczewski, R. R. (2017). Climate, anchovy, and sardine. *Annual Review of Marine Science*, *9*, 469–493.

Crawford, R. (1991). An empirical investigation of trans-oceanic linkages between areas of high abundance of sardine. *Long-term variability of pelagic fish populations and their environment*.

Cubillos, L., Arcos, D., Bucarey, D., y Canales, M. (2001). Seasonal growth of small pelagic fish off talcahuano, chile (37 degrees s, 73 degrees w): a consequence of their reproductive strategy to seasonal upwelling? *Aquatic Living Resources (France)*.

Cubillos, L. A., y Arcos, D. F. (2002). Recruitment of common sardine (strangomera bentincki) and anchovy (engraulis ringens) off central-south chile in the 1990s and the impact of the 1997–1998 el niño. *Aquatic Living Resources*, *15*(2), 87–94.

Cubillos, L. A., Ruiz, P., Claramunt, G., Gacitúa, S., Núñez, S., Castro, L. R., y Sepúlveda, A. (2007). Spawning, daily egg production, and spawning stock biomass estimation for common sardine (strangomera bentincki) and anchovy (engraulis ringens) off central southern chile in 2002. *Fisheries Research*, *4*(86), 228-240.

Cubillos, S., y cols. (1998). Evaluación indirecta del stock de anchoveta y sardina común en la zona centro-sur: informe final.

Cury, P., Bakun, A., Crawford, R. J., Jarre, A., Quinones, R. A., Shannon, L. J., y Verheye, H. M. (2000). Small pelagics in upwelling systems: patterns of interaction and structural changes in "wasp-waist" ecosystems. *ICES Journal of Marine Science*, *57*(3), 603–618.

Czerwinski, I. A., Gutiérrez-Estrada, J. C., y Hernando-Casal, J. A. (2007). Short-term forecasting of halibut cpue: Linear and non-linear univariate approaches. *Fisheries Research*, *86*(2-3), 120–128.

Dickey, D. A., y Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427–431.

Escanciano, J. C. (2009). Quasi-maximum likelihood estimation of semi-strong garch models. *Econometric Theory*, *25*(2), 561–570.

Fan, J., Yao, Q., y Cai, Z. (2003). Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society: series B (statistical methodology)*, *65*(1), 57–80.

F. Hair Jr, J., Sarstedt, M., Hopkins, L., y G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (pls-sem) an emerging tool in business research. *European business review*, *26*(2), 106–121.

Fréon, P., y Yáñez, E. (1995). Influencia del medio ambiente en evaluación de stock: una aproximación con modelos globales de producción. *Investigaciones marinas*, *23*, 25–47.

Gatica, C., Arteaga, M., Giacaman, J., y Ruiz, P. (2007). Tendencias en la biomasa de sardina común (strangomera bentincki) y anchoveta (engraulis ringens) en la zona centro-sur de chile, entre 1991 y 2005. *Investigaciones marinas*, *35*(1), 13–24.

Gonzalez, H., Daneri, G., Figueroa, D., Iriarte, J., Lefevre, N., Pizarro, G., ... Troncoso, A. (1998). Producción primaria y su destino en la trama trófica pelágica y océano-atmósfera de co2 en la zona norte de la corriente de humboldt (23◦ s): Posibles efectos del evento el niño, 1997–1998 en chile. *Rev. Chil. Hist. Nat*, *71*, 429–458.

González, M., y Cortés, G. (2005). Variabilidad interanual de la temperatura superficial del mar en la costa norte de chile y su relación con el evento el niño-oscilación del sur. *Investigaciones Marinas*, *33*(2).

Grisa, C., y Sabourin, E. (2019). Agricultura familiar: de los conceptos a las políticas públicas en américa latina y el caribe. *2030–Alimentación, agricultura y desarrollo rural en América Latina y el Caribe*.

Gutiérrez-Estrada, J. C., Silva, C., Yáñez, E., Rodríguez, N., y Pulido-Calvo, I. (2007). Monthly catch forecasting of anchovy engraulis ringens in the north area of chile: non-linear univariate approach. *Fisheries Research*, *86*(2-3), 188-200.

Gutiérrez-Estrada, J. C., Yáñez, E., Pulido-Calvo, I., Silva, C., Plaza, F., y Bórquez, C. (2009). Pacific sardine (sardinops sagax, jenyns 1842) landings prediction. a neural network ecosystemic approach. *Fisheries research*, *100*(2), 116-125.

HAIR, J., JF, R. A., RL, T., y WC, B. (1999). Análisis de regresión múltiple. *Análisis multivariante. Pretince Hall Iberia, 5ª ed., Madrid*, *4*, 143–226.

Hair Jr, J. F., Matthews, L. M., Matthews, R. L., y Sarstedt, M. (2017). Pls-sem or cb-sem: updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, *1*(2), 107–123.

Hare, J. A., Alexander, M. A., Fogarty, M. J., Williams, E. H., y Scott, J. D. (2010). Forecasting the dynamics of a coastal fishery species using a coupled climate–population model. *Ecological Applications*, *20*(2), 452-464.

Hare, S., Minobe, S., Wooster, W., y McKinnell, S. (2000). An introduction to the PICES symposium on the nature and impacts of North Pacific climate regime shifts. *Progress in Oceanography*(47), 99-102.

Hare, S. R. (1996). *Low frequency climate variability and salmon production.* University of Washington.

Hyndman, R. J., y Athanasopoulos, G. (2018). Forecasting. *OTexts, Heathmont, Vic*.

IIP, F. (2016). *Antecedentes sobre áreas de pesca de la flota artesanal e industrial en sardina común y anchoveta en la zona centro sur 2016* (Inf. Téc.). Instituto de Investigación Pesquera.

Iriarte, J., Pizarro, G., Troncoso, V., y Sobarzo, M. (2000). Primary production and biomass of size-fractionated phytoplankton off antofagasta, chile (23–24 s) during pre-el nino and el nino 1997. *Journal of Marine Systems*, *26*(1), 37–51.

Irigoien, X., y de Roos, A. (2011). The role of intraguild predation in the population dynamics of small pelagic fish. *Marine Biology*, *158*(8), 1683–1690.

Kim, J. Y., Jeong, H. C., Kim, H., y Kang, S. (2015). Forecasting the monthly abundance of anchovies in the south sea of korea using a univariate approach. *Fisheries Research*, *161*, 293-302.

KL, S., y MC, S. (s.f.). Cc, 2015. fishing amplifies forage fish population collapses. *Proc. Natl. Acad. Sci*, *112*, 66486652.

Kristensen, D., y Rahbek, A. (2005). Asymptotics of the qmle for a class of arch (q) models. *Econometric Theory*, *21*(5), 946–961.

Lai, E., Cheng, Y., y McAleer, M. (2005). Predicting monthly catch for some western australia coastal finfish species with seasonal arima-garch models.

MacCall, A. D. (2009). Mechanisms of low-frequency fluctuations in sardine and anchovy populations. *Climate change and small pelagic fish*, *285*.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., y Francis, R. C. (1997). A pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the american Meteorological Society*, *78*(6), 1069–1080.

Merino, G., Barange, M., Blanchard, J. L., Harle, J., Holmes, R., Allen, I., ... others (2012). Can marine fisheries and aquaculture meet fish demand from a growing human population in a changing climate? *Global Environmental Change*, *22*(4), 795-806.

Naranjo, L., Plaza, F., Yáñez, E., Barbieri, M. Á., y Sánchez, F. (2015). Forecasting of jack mackerel landings (t rachurus murphyi) in central-southern c hile through neural networks. *Fisheries Oceanography*, *24*(3), 219-228.

Nelson, D. B., y Cao, C. Q. (1992). Inequality constraints in the univariate garch model. *Journal of Business & Economic Statistics*, *10*(2), 229–235.

Parada, C., Yannicelli, B., Hormazábal, S., Vásquez., A., Porobic., S., Ernst, J., ... Gretchina, A. (2013). Variabilidad ambiental y recursos pesqueros en el pacífico suroriental: estado de la investigación y desafíos para el manejo pesquero. *Latin American Journal of Aquatic Research*, *41*(1), 1-28.

Pearson, K. (1901). Principal components analysis. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *6*(2), 559.

Pérez-Magariño, S., Ortega-Heras, M., González-San José, M., y Boger, Z. (2004). Comparative study of artificial neural network and multivariate methods to classify spanish do rose wines. *Talanta*, *62*(5), 983–990.

Perry, A. L., Low, P. J., Ellis, J. R., y Reynolds, J. D. (2005). Climate change and distribution shifts in marine fishes. *science*, *308*(5730), 1912-1915.

Perry, R., y Barange, M. (2009). Global ocean ecosystem dynamics (globec).

Perry, R. I., Ommer, R. E., Barange, M., y Werner, F. (2010). The challenge of adapting marine social–ecological systems to the additional stress of climate change. *Current Opinion in Environmental Sustainability*, *2*(5-6), 356-363.

Plaza, F., Salas, R., y Yáñez, E. (2018). Identifying ecosystem patterns from time series of anchovy (engraulis ringens) and sardine (sardinops sagax) landings in northern chile. *Journal of Statistical Computation and Simulation*, *88*(10), 1863-1881.

Plaza, F., Yáñez, E., Rodríguez, N., Gutiérrez, J., y Pulido, I. (2008). Predicción de capturas de anchoveta (engraulis ringens) en el norte de chile mediante redes neuronales artificiales: un enfoque multivariado. *Nuevas aproximaciones metodológicas para el análisis de pesquerías. Universidad de Huelva, Huelva*, 39–57.

Preciado, I., Punzón, A., Gallego, J., Vila, Y., y cols. (2006). Using time series methods for completing fisheries historical series. *Bol. Inst. Esp. Oceanogr*, *22*(1-4), 83-90.

Roessig, J. M., Woodley, C. M., Cech, J. J., y Hansen, L. J. (2004). Effects of global climate change on marine and estuarine fishes and fisheries. *Reviews in fish biology and fisheries*, *14*(2), 251-275.

Rupert, T. J., Single, L. E., y Wright, A. M. (2003). The impact of floors and phase-outs on taxpayers' decisions and understanding of marginal tax rates. *Journal of the American Taxation Association*, *25*(1), 72–86.

Sanz-Fernández, V., y Gutiérrez-Estrada, J. C. (2021). Historical changes of blackspot seabream (pagellus bogaraveo) landing patterns in the strait of gibraltar from 1983 to 2016: Environmental and legislation effects. *Fisheries Oceanography*, *30*(2), 111-126.

Schwartzlose, R., y Alheit, J. (1999). Worldwide large-scale fluctuations of sardine and anchovy populations. *African Journal of Marine Science*, *21*.

Stevens, J. (2009). *Applied multivariate statistics for the social sciences (pp. 292–294)*. Hillsdale, NJ: Erlbaum.

Sucarrat, G. (2020). garchx: Flexible and robust garch-x modelling.

Suryanarayana, I., Braibanti, A., Rao, R. S., Ramam, V. A., Sudarsan, D., y Rao, G. N. (2008). Neural networks in fisheries research. *Fisheries Research*, *92*(2-3), 115–139.

Von Bertalanffy, L. (1950). An outline of general system theory. *British Journal for the Philosophy of science*.

Yáñez, E., Barbieri, M., y Santillán, L. (1992). Long-term environmental variability and pelagic fisheries in talcahuano, chile. *South African Journal of Marine Science*, *12*(1), 175-188.

Yáñez, E., Barbieri, M., y Silva, C. (2003). Fluctuaciones ambientales de baja frecuencia y principales pesquerías pelágicas chilenas. *Actividad pesquera y de acuicultura en Chile*, 109–121.

Yáñez, E., Barbieri, M., Silva, C., Nieto, K., y Espíndola, F. (2001). Climate variability and pelagic fisheries in northern chile. *Progress in Oceanography*, *49*, 581 - 596. Descargado de http://www.sciencedirect.com/science/article/pii/S0079661101000428 (Pacific climate variability and marine ecosystem impacts) doi: http://dx.doi.org/10.1016/S0079-6611(01)00042-8

Yáñez, E., Hormazábal, S., Silva, C., Montecinos, A., Barbieri, M. A., Valdenegro, A., ... Gómez, F. (2008, 00). Coupling between the environment and the pelagic resources exploited off northern Chile: ecosystem indicators and a conceptual model. *Latin american journal of aquatic research*, *36*, 159-181. Descargado de http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-560X2008000200003&nrm=iso

Yáñez, E., Plaza, F., Gutiérrez-Estrada, J., Rodríguez, N., Barbieri, M., Pulido-Calvo, I., y Bórquez, C. (2010a). Anchovy (engraulis ringens) and sardine (sardinops sagax) abundance forecast off northern chile: a multivariate ecosystemic neural network approach. *Progress in Oceanography*, *87*(1), 242–250.

Yáñez, E., Plaza, F., Gutiérrez-Estrada, J. C., Rodríguez, N., Barbieri, M., Pulido-Calvo, I., y Bórquez, C. (2010b). Anchovy (engraulis ringens) and sardine (sardinops sagax) abundance forecast off northern chile: a multivariate ecosystemic neural network approach. *Progress in Oceanography*, *87*(1-4), 242-250.

Yáñez, E., Plaza, F., Sánchez, F., Silva, C., Barbieri, M. Á., y Bohm, G. (2017). Modelling climate change impacts on anchovy and sardine landings in northern chile using anns. *Latin american journal of aquatic research*, *45*(4), 675-689.

Yáñez Rodríguez, E. (1998). *Fluctuations des principaux stocks de poissons pélagiques exploités au chili et leurs relations avec les variations de l'environnement* (Tesis Doctoral no publicada). Brest.

Yáñez, E. (1998). *Fluctuations des principaux stocks de poisons pélagiques exploités au chili et leurs relations avec les variations de l" environnement.*