UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

ESCUELA DE POSTGRADO

MAESTRÍA EN ESTADÍSTICA APLICADA



"RELACIONES ONTOGÉNICAS Y ESPACIO-TEMPORALES EN LA DIETA DEL CALAMAR GIGANTE (*Dosidicus gigas*) EN PERÚ, UTILIZANDO UN MODELO ADITIVO GENERALIZADO"

Presentada por:

ANA RENZA PAOLA ALEGRE NORZA SIOR

TESIS PARA OPTAR EL GRADO DE MAGISTER SCIENTIAE EN:

ESTADÍSTICA APLICADA

Lima - Perú

2011

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

ESCUELA DE POSTGRADO

MAESTRÍA EN ESTADÍSTICA APLICADA

"RELACIONES ONTOGÉNICAS Y ESPACIO-TEMPORALES EN LA DIETA DEL CALAMAR GIGANTE (*Dosidicus gigas*) EN PERÚ, UTILIZANDO UN MODELO ADITIVO GENERALIZADO"

TESIS PARA OPTAR EL GRADO DE MAGÍSTER SCIENTIAE

Presentada por:

ANA RENZA PAOLA ALEGRE NORZA SIOR

Sustentada y aprobada por el siguiente jurado:

M Sc. Rino Sotomayor Ruiz PRESIDENTE M Sc. Víctor Maehara Oyata PATROCINADOR

M Sc. Clodomiro Miranda Villagómez MIEMBRO M Sc. César Menacho Chiok MIEMBRO

UNIVERSIDAD NACIONAL AGRARIA LA MOLINA Escuela de Post Grado Telf. 614-7800 Anexo 209 Telefax 614-7142 Apartado 12-056 La Molina LIMA - PERU

ACTA DE SUSTENTACION

Los Miembros del Jurado que suscriben, reunidos para evaluar la sustentación de tesis presentada por la alumna ANA RENZA PAOLA ALEGRE NORZA, denominada: "RELACIONES ONTOGÉNICAS Y ESPACIO-TEMPORALES EN LA DIETA DEL CALAMAR GIGANTE (*Dosidicus gigas*) EN PERÚ, UTILIZANDO UN MODELO ADITIVO GENERALIZADO", para cumplir con uno de los requisitos para optar el grado académico de *Magister Scientiae* en la Especialidad de ESTADÍSTICA APLICADA.

Teniendo en consideración los méritos del referido trabajo así como los conocimientos demostrados por la sustentante, declaramos la tesis como:

APROBADA EXCLIENTE

con el calificativo (*) de

En consecuencia, queda en condición de ser considerada APTA por el Consejo Universitario y recibir el grado académico de *Magister Scientiae*, de conformidad con lo estipulado en el Artículo 41° del Reglamento de la Escuela de Post Grado.

La Molina, 27 de diciembre del 2011

Mg.Sc. Rino Sotomayor Ruiz PRESIDENTE

Mg.Sc. Clodomiro Miranda Villagómez MIEMBRO

M ara Ovata ictor Mae PATROCIMADOR

NACIO

Escuela da Post Grada

A MOLIN

Mg.Sc. César Menacho Chiok MIEMBRO

(*) De acuerdo con el Artículo 17º del Reglamento de Tesis, éstas deberán ser calificadas con términos de: EXCELENTE, MUY BUENO, BUENO o REGULAR. Dedicado a mis padres Mario⁺ y Angélica⁺ que son mi fuente de inspiración y a mis hijos Paula y Salvador quienes son el motor que me impulsa a cumplir mis sueños.

AGRADECIMIENTOS

En primer lugar quiero agradecer a mi profesor M Sc. Víctor Maehara, mi mentor, patrocinador y maestro, por ayudarme minuciosamente con el desarrollo de esta tesis, siendo siempre muy acucioso en sus críticas; pero también por darme la confianza de estar desarrollando un buen trabajo.

También quiero dar las gracias al Dr. Arnaud Bertrand (IRD - Francia) por haberme dado la oportunidad de estudiar esta maestría, por su incansable apoyo y sabios consejos, por darme el ánimo para seguir en el camino y no desfallecer, por ser amigo y compañero en esta lucha, sin su voto de confianza no hubiera podido subir este peldaño en mi carrera. Al Dr. Fréderic Ménard (IRD – Francia) quien me enseñó una parte de las técnicas estadísticas que he usado en este trabajo.

Asimismo, quiero agradecer a los profesores M Sc. Rino Sotomayor, M Sc. César Menacho y M Sc. Clodomiro Miranda, por su tiempo y disposición para revisar esta tesis. Agradezco también a los profesores que tuve en la maestría en estadística aplicada de la Universidad Nacional Agraria, cuyas enseñanzas me sirvieron como herramientas imprescindibles en el desarrollo de mis investigaciones; no quiero olvidar a mis compañeros de clase con los cuales compartí momentos de alegría, estrés y de apoyo mutuo cuando formábamos grupos de estudio, a Jesús, Joel, Alberto, Carlos, María, Sandra, Karina y Adrián; sin obviar a la señora Rosita secretaria del departamento académico quien me dio su apoyo desinteresado para hacer el seguimiento a mi documentación.

Quiero agradecer también al Instituto del Mar del Perú que me proporcionó la infraestructura necesaria para la obtención de los datos usados en esta tesis, a mis compañeros de trabajo quienes me han dado apoyo y ánimo para seguir adelante con esta ardua tarea, a Pepe, Iván, David, César, Juan y Anatolio; especialmente a Ricardo Tafur quien ha analizado junto conmigo durante 6 años, el contenido de más de 4000 de estómagos de calamar gigante.

Finalmente, quiero dar las gracias a mi familia, a mis hermanas Fanny y Angélica que me alentaron a seguir con mis estudios, a mi hermano Mario quien fue como un padre y el primero en sembrar en mi la curiosidad por responder preguntas científicas; a mi esposo Flavio quien me brindó su apoyo incondicional en esta aventura, acompañando a mis hijos en mi ausencia cuando llegaba tarde de mis clases, por extenderme la mano en los momentos más difíciles y darme el amor que me inspira a seguir adelante, a mis dos hermosos hijos Paula y Salvador que son mi motor, que tuvieron paciencia cuando tenía que pasar horas en la computadora, sin hacerme ruido cuando tenía que estudiar, esperándome con alegría los sábados que llegaba de clases, a Sonia por cuidarlos con tanto cariño. No quiero olvidar a dos seres que se fueron apenas comenzaba a estudiar el pregrado, mis padres Mario y Angélica, a él por enseñarme que el conocimiento es la mayor riqueza que puede portar un ser humano y a mi madre quien siempre confió en mi capacidad y cuyo último deseo fue que sea consecuente con mis elecciones y que termine la carrera que ya había comenzado.

A todos ellos, mi más sincero agradecimiento.

RESUMEN

Alegre A. 2011. Relaciones ontogénicas y espacio-temporales en la dieta del calamar gigante (*Dosidicus gigas*) en Perú, utilizando un Modelo Aditivo Generalizado. Tesis para obtener el grado de Magister en Ciencias. Escuela de Postgrado. Maestría en Estadística Aplicada. Universidad Nacional Agraria La Molina. 100pp

El calamar gigante Dosidicus gigas (d'Orbigny, 1835) es un depredador importante en el ecosistema del Perú. Se postula que el papel del calamar gigante varía teniendo en cuenta la talla, tiempo, hora, temperatura y distribución espacial. Para comprobar esta hipótesis se aplicó un modelo aditivo generalizado (GAM) en datos biológicos de alimentación de 4178 calamares gigantes capturados por la flota industrial pesquera a lo largo del litoral peruano (3°S a 18°S) desde 2 a 299 millas náuticas (mn) de distancia a la costa desde el año 2004 a 2009 realizados por el Laboratorio de Ecología Trófica del Instituto del Mar del Perú (IMARPE). La talla de los calamares estudiados fluctuó entre 14 y 112 cm de longitud de manto (LM). En total 43 item-presa fueron registrados, los grupos más importantes fueron los cefalópodos (Dosidicus gigas), Teleosteii (Photichthyidae, Myctophidae y Nomeidae) y Malacostraca crustáceos (Euphausiidae). Las presas principales fueron D. gigas (indicando canibalismo) en términos gravimétricos (% W=35.4), los otros cephalopodos en frecuencia de ocurrencia (FO=14.4), y los eufáusidos en términos de abundancia relativa (% N=62.2). Estos resultados reflejan una alta variabilidad de la dieta, y un espectro trófico similar en comparación con otras latitudes en ambos hemisferios (México y Chile). Los modelos GAM muestran que todas las variables predictoras fueron significativas en relación a la variable respuesta llenura estomacal (p <0.0001). La llenura estomacal fue mayor en los individuos juveniles, también durante la noche hubo mayor consumo, mientras no se reflejaron tendencias en la alimentación con relación a la temperatura superficial del mar (TSM), pero espacialmente se observan cambios en la dieta, aumentando el porcentaje de llenura a medida que esta especie se aleja de la costa. Por lo tanto se concluye que la dieta del calamar gigante depende de la talla y su distribución espacio-temporal.

Palabras clave: Dosidicus gigas, dieta, modelo aditivo generalizado, llenura.

ABSTRACT

Alegre A. 2011. Ontogenetic and spatio-temporal relationships in the diet of jumbo squid (*Dosidicus gigas*) off Peru, using a Generalized Additive Model. Thesis for obtain the Degree of Master in Science. Postgraduate School. Master of Applied Statistics. National Agrarian University La Molina.100pp

The giant squid *Dosidicus gigas* (d'Orbigny, 1835) is a important predator in the ecosystem of Peru. We hypothesize that the giant squid predator role varies considering the size, time, hour, temperature and spatial distribution. To test this hypothesis was proposed to apply a generalized additive model (GAM) on biological data of 4178 giant squid caught by industrial fishing, along the coast of Peru (3°S to 18°S) from 2 to 299 nautical miles (nm) away to the coast from 2004 to 2009 on by the Laboratory of Trophic Ecology of Marine Institute of Peru (IMARPE). The size of squid studied ranged between 14 and 112 cm mantle length (ML). In total 43 item-prey were recorded, the most important were the cephalopods (Dosidicus gigas), Teleosteii (Photichthyidae, Myctophidae and Nomeidae) and Malacostraca crustaceans (Euphausiidae). The main prey were D. gigas (indicating cannibalism) in terms of gravimetric (%W=35.4), the other Cephalopoda in frequency of occurrence (FO=14.4), and euphausiids in terms of relative abundance (%N=62.2). These results reflect a high variability in the diet, and a similar trophic spectrum compared to other latitudes in both hemispheres (Mexico and Chile). GAM models show that all predictor variables were significant variable in relation to the stomach fullness (p <0.0001). Stomach fullness was higher in the juvenile individuals, also during the evening there was increased consumption, while not reflected trends in food in relation to sea surface temperature (SST), but spatially observed changes in diet, increasing the percentage of fullness as this species is far from the coast. It is therefore concluded that the giant squid diet depends of the size and the spatiotemporal distribution.

Keywords: Dosidicus gigas, diet, fullness, Generalized Additive Model.

ÍNDICE GENERAL

I.	INT	RODUCCIÓN	1
II.	REV	VISIÓN DE LITERATURA	4
	2.1	ANTECEDENTES BIOLÓGICOS DEL CALAMAR GIGANTE	4
		2.1.1 TAXONOMÍA	4
		2.1.2 DISTRIBUCIÓN A NIVEL MUNDIAL	5
		2.1.3 CICLO DE VIDA	6
		2.1.4 MORFOLOGÍA EXTERNA	7
		2.1.5 MADUREZ	8
		2.1.6 ECOLOGÍA TRÓFICA	8
	2.2	ANTECEDENTES OCEANOGRÁFICOS	10
		2.2.1 EL SISTEMA DE LA CORRIENTE DE HUMBOLDT	10
		2.2.2 ZONA MÍNIMA DE OXÍGENO	11
	2.3	ÍNDICE DE DIVERSIDAD DE MARGALEF (d)	12
		2.3.1 CÁLCULO DEL ÍNDICE DE LLENURA ESTOMACAL	12
	2.4	ANTECEDENTES ESTADÍSTICOS	13
		2.4.1 GRÁFICO DE PROBABILIDADES NORMALES	13
		2.4.2 GRAFICO DE HOMOGENEIDAD DE VARIANCIA	13
		2.4.3 PRUEBA DE NORMALIDAD DE SHAPIRO & WILK	14
		2.4.4 MÉTODOS DE SELECCIÓN DE VARIABLES	15
		2.4.5 MODELOS LINEALES GENERALIZADOS	17
		2.4.6 MODELOS ADITIVOS GENERALIZADOS	18
III	MA	TERIALES Y MÉTODOS	24
	3.1	MATERIALES	24
		3.1.1 MATERIAL BIOLÓGICO	24
		3.1.2 MATERIAL DE LABORATORIO	24
		3.1.3 MATERIAL DE ESCRITORIO	24
	3.2	POBLACIÓN Y MUESTRA	24
	3.3	PROCESAMIENTO DE MUESTRAS	28
	3.4	PROCESAMIENTO DE DATOS	29
		3.4.1 FORMULACIÓN DE HIPÓTESIS	29
		3.4.2 MODELO GAM	29
		3.4.3 DISTRIBUCIONES TWEEDIE	38
		3.4.4 LIBRERÍAS DE "R" USADAS	38
IV	. RES	SULTADOS Y DISCUSIÓN	39
	4.1	IDENTIFICACIÓN DE VARIABLES	39
		4.1.1 VARIABLES DEPENDIENTES O RESPUESTA	39
		4.1.2 VARIABLES INDEPENDIENTES O PREDICTORAS	39
	4.2	VARIABILIDAD DE LA DIETA	39
		4.2.1 DIETA GLOBAL	39
		4.2.2 VARIABILIDAD DE LA DIETA RESPECTO A LA TALLA	41
		4.2.3 VARIABILIDAD TEMPORAL DE LA DIETA	43
		4.2.4 VARIACIÓN ESPACIAL DE LA DIETA	47
	4.3	ANÁLISIS EXPLORATORIO	49
		4.3.1 RELACIONES GENERALES ENTRE VARIABLES	49
		4.3.2 GRÁFICOS DE PROBABILIDAD NORMAL	51

4.4	MODELO ADITIVO GENERALIZADO (GAM)	54
	4.4.1 MÉTODOS DE SELECCIÓN DE VARIABLES	54
	4.4.2 TIPO DE DISTRIBUCIÓN	
	4.4.3 ÍNDICE DE LLENURA ESTOMACAL	56
	4.4.4 ÍNDICE DE DIVERSIDAD DE MARGALEF	60
V.	CONCLUSIONES	64
VI.	RECOMENDACIONES	66
VII.	REFERENCIAS BIBLIOGRÁFICAS	67
VIII.	ANEXOS	82

ÍNDICE DE FIGURAS

Figura 1. Desembarques anuales del calamar gigante Dosidicus gigas, en el litoral	
peruano de 1980 al 2009 (Fuente: IMARPE)	1
Figura 2. Distribución geográfica del calamar gigante, Dosidicus gigas (Fuente:	
Salinas et al. 2010)	6
Figura 3. Calamar gigante Dosidicus gigas. A, vista dorsal de un ejemplar grande	
(50-80 cm LM) (Roper et al. 1984); B, vista ventral de un ejemplar de 25.4 cm LM	
(Fuente: Wormuth 1976)	8
Figura 4. Principales predadores y presas del calamar gigante Dosidicus gigas	
(Fuente: IMARPE)	10
Figura 5. Esquema del efecto de la Zona Mínima de Oxigeno (ZMO) en la	
distribución de las especies según día y noche (Fuente: Ballón et al. 2011)	12
Figura 6. Estaciones de muestreo de estómagos de calamar gigante Dosidicus gigas	
a lo largo del litoral del 2004 al 2009	25
Figura 7. Gráfico de dispersión del contenido del estómago del calamar gigante	
(peso en gramos) sobre la base de <i>D. gigas</i> y otras especies (sin el canibalismo) de	
2004 a 2009	27
Figura 8. Disección del estómago de calamar gigante <i>Dosidicus gigas</i>	28
Figura 9. Metodología de muestreo y separación de estructuras duras de peces y	
cetalópodos utilizadas para la identificación indirecta de las especies presa del	•
calamar gigante <i>Dosidicus gigas</i>	29
Figura 10. Número de especies, indices de diversidad, porcentaje de llenura	
estomacal y composición de la presas en peso (g) en la dieta del calamar gigante	40
<i>Dosidicus gigas</i> respecto a grupos de talla (cm) del 2004 al 2009	43
Figura 11. Numero de especies, indices de diversidad, porcentaje de llenura	
estomacal y composicion de la presas en peso (g) en la dieta del calamar gigante	4.4
<i>Dostatcus gigas</i> respecto a las estaciones y el ano del 2004 al 2009	44
Figura 12. Numero de especies, indices de diversidad, porcentaje de lienura	
Desidieus ciago recepcto el periodo hororio del 2004 el 2000	17
<i>Dostateus gigas</i> respecto al periodo norano del 2004 al 2009	47
la presas en peso (g) en la dieta del calamar gigante Desidicus gigas respecto a la	
distancia a la costa (mn) del 2004 al 2009	18
Figura 14 Número de especies índices de diversidad porcentaie de llenura	40
estomacal y composición de la presas en peso (g) en la dieta del calamar gigante	
Dosidicus gigas respecto a la latitud (°S) del 2004 al 2009	49
Figura 15 Scaternlot entre las variables llenura estomacal (fullness) número de	77
especies (species) índice de diversidad de Margalef (Margalef) talla (size)	
estacionalidad (seasvear) periodicidad horaria (hour) distancia a la costa	
(<i>distcoast</i>) latitud (lat) y temperatura superficial del mar (SST) en la dieta del	
calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	50
Figura 16. Análisis de normalidad y homocedasticidad de varianzas del modelo	
lineal: porcentaie de llenura estomacal del calamar gigante, respecto a las variables	
predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud v	
temperatura.	52
Figura 17. Análisis de normalidad y homocedasticidad de varianzas de la variable	
respuesta del modelo: índice de diversidad de Margalef respecto a las variables	
predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y	
-	

temperatura	53
Figura 18. Histograma de frecuencia y densidad de la variable respuesta llenura estomacal en el calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	55
Figura 19. Gráfico <i>saddlepoint approximation</i> , se puede apreciar un "p" máximo de 2.2 de la variable respuesta llenura estomacal en el calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	56
Figura 20. Modelo Aditivo Generalizado (GAM) relacionando a la respuesta de la llenura estomacal con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	57
Figura 21. Pruebas de bondad de ajuste y homocedasticidad de varianzas de la variable respuesta del modelo: llenura estomacal respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y temperatura	59
Figura 22. Modelo Aditivo Generalizado (GAM) relacionando a la respuesta del índice de diversidad de Margalef con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	61
Figura 23. Pruebas de bondad de ajuste y homocedasticidad de varianzas de la variable respuesta del modelo: llenura estomacal respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y	67
ютротацита	02

ÍNDICE DE CUADROS

Cuadro 1. Familia de distribución y funciones de enlace correspondientes	21
Cuadro 2. Librerías de "R" usadas	38
Cuadro 3. Influencia de las variables predictoras en el comportamiento de la variable	
respuesta llenura estomacal, usando las pruebas de bondad de ajuste AIC y	
desvianza explicada	54
Cuadro 4. Selección de variables predictoras que más influyen en el modelo usando	
el método Forward	55
Cuadro 5. Coeficientes del modelo por cada tramo de <i>splines</i>	60
Cuadro 6. Coeficientes del modelo por cada tramo de <i>splines</i>	63

ÍNDICE DE ANEXOS

ANEXO 1a. Espectro trófico del calamar gigante Dosidicus gigas en el litoral
peruano del 2004 al 2009
ANEXO 1b. Espectro trófico del calamar gigante <i>Dosidicus gigas</i> en el litoral
peruano del 2004 al 2009
ANEXO Ic. Espectro trófico del calamar gigante <i>Dosidicus gigas</i> en el litoral
peruano del 2004 al 2009
ANEXO 1d. Espectro trófico del calamar gigante <i>Dosidicus gigas</i> en el litoral
peruano del 2004 al 2009
ANEXO If. Espectro trófico del calamar gigante <i>Dosidicus gigas</i> en el litoral
peruano del 2004 al 2009
ANEXO 2. Histogramas de frecuencias de las variables respuestas (a) porcentaje
de llenura estomacal (%), (b) número de especies e (c) índice de diversidad de
Margalef y las predictoras (covariables) (d) talla (cm), (e) estacionalidad, (f)
periodicidad horaria (24 horas), (g) distancia a la costa (mn), (h) latitud (°S) y (i)
temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i>
del 2004 al 2009
ANEXO 3. Diagrama de cajas (boxplot) de las variables respuestas (a) porcentaje
de llenura estomacal (%), (b) número de especies e (c) índice de diversidad de
Margalef y las predictoras (covariables) (d) talla (cm), (e) estacionalidad, (f)
periodicidad horaria (24 horas), (g) distancia a la costa (mn), (h) latitud (°S) y (i)
temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i>
del 2004 al 2009
ANEXO 4. Diagrama de cajas (boxplot) de las variables predictoras (covariables):
(a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a
la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación
a la variable respuesta porcentaje de llenura estomacal en la dieta del calamar
gigante <i>Dosidicus gigas</i> del 2004 al 2009
ANEXO 5. Diagrama de cajas (boxplot) de las variables predictoras (covariables):
(a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a
la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación
a la variable respuesta número de especies en la dieta del calamar gigante <i>Dosidicus</i>
<i>gigas</i> del 2004 al 2009
ANEXO 6. Diagrama de cajas (boxplot) de las variables predictoras (covariables):
(a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a
la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación
a la variable respuesta índice de diversidad de Margalef en la dieta del calamar
gigante <i>Dosidicus gigas</i> del 2004 al 2009
ANEXO 7. Script para medir la influencia de las variables predictoras en el
comportamiento de la variable respuesta llenura estomacal, usando las pruebas de
bondad de ajuste AIC y desvianza explicada
ANEXO 8. Script para medir la influencia de las variables predictoras en el
comportamiento de la variable respuesta llenura estomacal, usando las pruebas de
bondad de ajuste AIC y desvianza explicada
ANEXO 9. Script para medir la respuesta de la llenura estomacal con las variables

predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a		
la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del		
calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	97	
ANEXO 10. Script para medir la respuesta de índice de diversidad de Margalef con		
las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas),		
distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la		
dieta del calamar gigante Dosidicus gigas del 2004 al 2009		
la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009 ANEXO 10. Script para medir la respuesta de índice de diversidad de Margalef con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante <i>Dosidicus gigas</i> del 2004 al 2009	97 99	

I. INTRODUCCIÓN

La actividad pesquera en el Perú es la segunda fuente generadora de divisas después de la minería, como en el año 2008, que representó aproximadamente dos mil millones de dólares en exportaciones (Díaz *et al.* 2009); esta importancia está sustentada en los recursos pesqueros marinos tales como el calamar gigante (*Dosidicus gigas*), que actualmente es una de las principales pesquerías en el Perú, pues durante el año 2011 la exportación de esta especie sumó 113 millones de dólares hasta mediados de año destinando este recurso principalmente a España, China, Tailandia, República de Corea, Japón, Italia y Venezuela, cabe destacar que los desembarques anuales hasta el 2008 han presentado una tendencia creciente hasta alcanzar valores de 287 mil toneladas (Figura 1).



Figura 1. Desembarques anuales del calamar gigante *Dosidicus gigas*, en el litoral peruano de 1980 al 2009 (Fuente: IMARPE)

Es importante el estudio de alimentación de esta especie porque es a través de los análisis de los contenidos estomacales que se consigue información de utilidad que permite cuantificar las especies de mayor relevancia en la estructuración de la red trófica marina¹, la cual se basa en las relaciones predador-presa, sin embargo, estas relaciones no son lineales e incluso aplicando transformaciones se hace muy difícil aplicar metodología de estadística paramétrica para modelar este tipo de comportamiento. Los Modelos Lineales Generalizados (siglas en inglés GAM) ya han sido aplicados en este tipo de información

¹ Red trófica marina: describen los hábitos alimentarios y las interacciones que se dan entre los individuos de una comunidad marina.

(Espinoza y Bertrand 2008, Espinoza *et al.* 2009); sin embargo, no se ha profundizado en la aplicación de métodos de suavización que permitan un mejor ajuste de este tipo de datos.

Los estudios de alimentación del calamar gigante, son realizados con datos provenientes de muestras recolectadas *in situ*, los cuales presentan muchas limitaciones; por ejemplo, frecuentemente no presentan distribución normal que imposibilitan el uso de métodos paramétricos. El campo no paramétrico o semiparamétrico es uno de los más populares y se está empleando como una nueva herramienta de análisis estadístico. Esta herramienta ofrece una alternativa más adecuada en la exploración de datos univariados o multivariados sin presuponer ninguna distribución específica.

Las técnicas de regresión no paramétrica y semiparamétricas se presentan como una alternativa para la descripción de variables con distribución desconocida y para reducir la dispersión de las estimaciones de variables biológicas y fisicoquímicas tomadas *in situ*, estimando objetivamente la relación funcional del término predictor de un modelo aditivo (Hastie y Tibshirani 1990). La idea básica de estos modelos es reemplazar la función lineal usual de cada variable por una función de suavizado, permitiéndole a los datos mostrar la forma funcional más adecuada.

En este estudio, nuestro interés principal es aplicar técnicas de inferencia estadística utilizadas para datos con distribución diferente a la normal, como son los Modelos Aditivos Generalizados para conocer mejor el comportamiento trófico² del calamar gigante, relacionando el índice de llenura del estómago términos de porcentaje en relación al peso del cuerpo y la diversidad de presas (usando índices ecológicos como el índice de Margalef) con variables predictoras bióticas³ y abióticas⁴ (ontogénicas y espaciotemporales respectivamente). En este estudio se utilizó información de más de 4000 estómagos de calamar gigante de una muestra tomada entre el año 2004 y 2009. En tal contexto, se plantea la hipótesis que existen relaciones entre las variables respuestas: índice de llenura estomacal e índice de diversidad de las presas del calamar gigante (*Dosidicus gigas*) respecto a las variables predictoras: talla, latitud, distancia a la costa, tiempo (años, estación) y temperatura.

² Trófico: referidos a las relaciones de alimentación predador-presa.

³ Biótico: hace referencia a lo característico de los seres vivos o que está vinculado a ellos.

⁴ Abiótico: se refiere al medio en el que no es posible la vida.

Asimismo, este trabajo tiene como objetivo principal aplicar un Modelo Aditivo Generalizado (GAM) en datos biológicos de alimentación del calamar gigante *Dosidicus gigas*, para describir las relaciones ontogénicas⁵ y espacio-temporales en su dieta; determinando las relaciones de las variables respuestas cuantitativas continuas: índice de llenura estomacal e índice de diversidad de Margalef, respecto a las variables predictoras cuantitativas continuas: talla (cm), tiempo (años, estación), periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (TSM).

Además tiene como objetivos específicos probar que no existe relación o función lineal entre las variables respuesta: índice de llenura estomacal e índice de diversidad de Margalef y las variables predictoras: talla, tiempo, hora, distancia a la costa, latitud y temperatura. Probar que el modelo no pertenece a la familia de distribución Gausiana o normal. En caso se pruebe la no linealidad del modelo, determinar la familia de distribución a la que mejor se ajusta el modelo. Determinar entre las variables predictoras: talla, tiempo, hora, distancia a la costa, latitud y temperatura; cuáles son las variables que explican o influyen más en el comportamiento de las variables respuesta: índice de llenura estomacal e índice de diversidad de Margalef. Analizar las tendencias de las relaciones entre las variables respuesta: índice de llenura estomacal e índice de diversidad de Margalef y las variables predictoras: talla, tiempo, hora, distancia a la costa, latitud y temperatura.

Para cumplir estos objetivos se aplicó el modelo GAM (Hastie y Tibshirani 1990, Venables y Ripley 2002, Wood y Augustin 2002, Venables y Dichmont 2004, Wood 2006). El GAM es un método semiparamétrico que mediante funciones suavizadas llamadas *splines* permite determinar la forma de la relación entre variables ambientales y bióticas (Swartzman 1997, Piet 2002, Cianelli *et al.* 2008).

Estos resultados constituirán una base útil para la administración pesquera de estas poblaciones y para estudios bio-económicos de esta actividad pesquera.

⁵ Ontogénico: relacionado al desarrollo del individuo respecto a la talla.

II. REVISIÓN DE LITERATURA

2.1 ANTECEDENTES BIOLÓGICOS DEL CALAMAR GIGANTE

2.1.1 TAXONOMÍA

Reino: Animalia

Filo: Mollusca

Clase: Cephalopoda

Orden: Teuthida

Suborden: Oegopsida

Familia: Ommastrephidae

Subfamilia: Ommastrephinae

Género: Dosidicus

Especie: D. gigas (d'Orbigny 1835)

Nombres comunes:

Calamar gigante (México, Perú), jibia (Chile, Perú), pota (Perú), calamar rojo (Chile).

La primera referencia al calamar gigante la realiza el abate Molina (1788), quien la denomina *Sepia tunicata* sugiriendo el parecido con las sepias (Cephalopoda: Sepioidea) que él hubo de ver en Europa (Schneider 1930, Wilhelm 1954). d'Orbigny (1835) fue, sin embargo, el primero en describir formalmente a esta especie, nombrándola *Ommastrephes gigas*, en clara alusión a su tamaño. Posteriormente Steenstrup (1857) describió un nuevo género y especie, *Dosidicus eschrichti*. Steenstrup siempre distinguió su *Dosidicus* del *O. gigas* descrito por d'Orbigny (1835). Posteriormente, se ha reconocido la sinonimia de ambas especies (Berry 1912, Pfeffer 1912, Wormuth 1998), quedando establecido que *Dosidicus gigas* es el único representante vivo del género.

2.1.2 DISTRIBUCIÓN A NIVEL MUNDIAL

Dosidicus gigas es el calamar ommastrefido de mayor abundancia y tamaño en el Pacifico Este. Se distribuye desde California (37°N) hasta el sur de Chile (47°S). Las mayores concentraciones se encuentran en la parte central de su rango de distribución (Figura 2), particularmente en las aguas relacionadas a las zonas de alta producción primaria, alta biomasa de zooplancton y peces mesopelágicos⁶ y batipelágicos⁷ (Nesis 1970, 1983, Nigmatullin et al. 2001). La mayoría de los estudios indican que diferentes poblaciones de calamar gigante se distribuyen el pacifico este. Así, en base a observaciones morfológicas, Clarke y Paliza (2000) concluyen que el calamar gigante de la corriente de Perú, que habita aguas frías representa una población diferente, separada por la contracorriente ecuatorial de la población del norte, la cual vive en aguas cálidas y es de menor tamaño. De la misma manera, usando la misma metodología Ehrhardt et al. (1983a,b, 1986) concluyen que los calamares del Golfo de California constituyen un sólo stock compuesto por varias cohortes⁸. Más recientemente, métodos genéticos confirmaron estas conclusiones. Enríquez-Paredes (1999) discriminó, mediante la técnica de ampliación aleatoria de polimorfimos de ADN, hasta 6 cohortes⁸, y al menos una, la de individuos pequeños, de tallas menores a 35 cm LM, se diferencia bien de las cohortes con tallas de madurez mayores a 50 cm LM. Finalmente Sandoval et al. (2007), usando distintos métodos aplicados a datos obtenidos con dos marcadores moleculares (RAPDs⁹ y secuencias de DNA mitocondrial identificadas mediante SSCPs¹⁰), probaron que los calamares gigantes del Golfo de California (México) y del Mar de Perú presentan una divergencia genética.

⁶ Peces mesopelágicos: peces que viven a media agua aproximadamente a 200 m de profundidad.

⁷ Peces batipelágicos: peces que viven en el fondo desde los 1000 m de profundidad.

⁸ Cohorte: conjunto de individuos de una población que han nacido en una misma época del año.

⁹ RAPDs: polimorfismos de ADN amplificados al azar, marcadores moleculares que amplifican lateralmente los segmentos de ADN en una gran variedad de especies.

¹⁰ SSCPs: polimorfismos conformacionales de cadena simple, una técnica adicional que revela los cambios en las secuencias de ADN con base en las cadenas sencillas de ADN.



Figura 2. Distribución geográfica del calamar gigante, *Dosidicus gigas* (Fuente: Salinas et al. 2010).

2.1.3 CICLO DE VIDA

D. gigas es una especie monocíclica¹¹ igual que la mayoría de los cefalópodos, y muere después del primer ciclo de reproducción (Rocha *et al.* 2001, Nigmatullin *et al.* 2001). Esta especie es de rápido crecimiento y corta vida (Arkhipkin y Murzov 1986, Masuda *et al.* 1998, Arguelles *et al.* 2001, Markaida *et al.* 2004) siendo la edad de la mayoría de ejemplares no mayores a un año. Sin embargo, parece que la edad de los ejemplares de mayor tamaño puede estar subestimado pudiendo ser mayor a los dos años (Arkhipkin, citado por Nigmatullin *et al.* 2001). Asimismo, otros estudios encontraron variabilidad en crecimiento, longevidad y tamaño máximo entre diferentes cohortes (Keyl 2009, Keyl *et al.* 2010).

¹¹ Monocíclica: se reproducen sólo una vez en su vida.

2.1.4 MORFOLOGÍA EXTERNA

Aletas romboidales, musculosas y anchas, su ancho ocupa el 49-65% de la longitud del manto (LM), y su longitud 41-49% de la LM; ángulo de la aleta 50-65°. Foveola¹² con 5 a 9 pliegues junto con 2 a 8 bolsas laterales; dáctilo de la masa tentacular con 4 hileras de ventosas; entre 60 y más de 200 pares de ventosas en el brazo I; membrana protectora del brazo III igual al ancho del brazo, membranas de los otros brazos nunca más altas que las ventosas; brazo IV derecho o izquierdo hectocotilizado¹³, sin ventosas y pedúnculos en la punta y membranas protectoras perforadas y engrosadas; pequeños fotóforos subcutáneos en la superficie ventral del manto, cabeza y brazos III y IV (Nesis 1971, Wormuth, 1976, 1998, Roper *et al.* 1984, 1995). La característica más conspicua es tal vez la prolongación de los extremos distales de los brazos en filamentos largos y delgados, conteniendo entre 100 y 200 pares de ventosas diminutas muy agrupadas, 3 veces más numerosas que en los demás Ommastrephidae¹⁴ (Nesis 1971, Zuev *et al.* 1975, Roper *et al.* 1984, 1995, Nigmatullin *et al.* 2001). Descripciones más detalladas del calamar gigante se encuentran en Berry (1912) y Wormuth (1976). Wormuth (1976, 1998) añade la presencia de dos fotóforos¹⁵ intestinales en su diagnóstico (Figura 3).

¹² Foveola: En los cefalópodos un pliegue transversal membranoso en el conducto sifonal.

¹³ Hectocotilo: Tentáculo de los cefalópodos que generalmente está modificado en su extremo para transferir el esperma a la hembra.

¹⁴ Ommastrephidae: perteneciente a la familia de los calamares gigantes.

¹⁵ Fotóforos: órgano luminoso especializado que se halla en la superficie de muchos organismos que habitan en la profundidad.



Figura 3. Calamar gigante *Dosidicus gigas*. A, vista dorsal de un ejemplar grande (50-80 cm LM) (Roper *et al.* 1984); B, vista ventral de un ejemplar de 25.4 cm LM (Fuente: Wormuth 1976).

2.1.5 MADUREZ

Nesis (1983) y Nigmatullin *et al.* (1991, 1999) han descrito tres grupos de tallas de madurez desconociendo su relación genética, Yokawa (1995) basado en análisis de isozimas¹⁶, halló en el mar peruano hembras maduras que pertenecían a la misma población de 24.3-26.5 cm LM, 40.8-47.7 cm LM y 83-92.3 cm LM. Wormuth (1976, 1998) señaló diferencias morfométricas¹⁷ significativas entre calamares maduros desde el norte al sur del Perú, siendo la más aparente la diferencia de tamaño, que puede deberse a disponibilidad de alimento.

2.1.6 ECOLOGÍA TRÓFICA (ALIMENTACIÓN)

Debido a su rápido crecimiento esta especie tiene un alto requerimiento energético, siendo una de sus principales presas los peces mesopelágicos (Shchetinnikov 1989). Sus

¹⁶ Isozimas o Isoenzimas son proteínas con diferente estructura pero que catalizan la misma reacción.

¹⁷ Morfométricas: referido a la forma, talla, peso, etc.

principales predadores son los cachalotes, los delfines y peces grandes como el pez espada, el merlín, el atún (Zuev y Nesis 1971).

El calamar gigante tiene los hábitos de alimentación oportunista (Amaratunga 1983) y lleva a cabo amplias migraciones verticales¹⁸ de hasta 1200 m de profundidad para alimentarse (Nesis 1970, Nigmatullin *et al.* 2001, Klett 1981, Markaida *et al.* 2005, Gilly *et al.* 2006). Estas migraciones verticales proporcionan un flujo de energía del eficiente ecosistema desde la superficie hasta las aguas más profundas (Klett 1981, Markaida *et al.* 2005, Gilly *et al.* 2005, Gilly *et al.* 2006).

Los calamares gigantes son predadores de una gran variedad de especies incluyendo a ellos mismos (Nesis 1970, Markaida y Sosa 2003) y tienen una gran importancia en la estructura trófica de los ecosistemas que habitan (Rodhouse y Nigmatullin 1996) (Figura 4). En el Perú entre 1991 y 1995, encontraron eufáusidos¹⁹ en calamares gigantes pequeños, mientras que los medianos y mayores se alimentan de peces mesopelágicos⁶ como mictófidos, peces voladores y otros peces, junto con una alta incidencia de canibalismo. El plancton disminuyó con el aumento del canibalismo en calamares mayores de 30 cm de longitud del manto LM (Yamashiro *et al.* 1998, Benites 1985, Benites y Valdivieso 1986, Nesis 1970, 1983, Kuroiwa 1998). Schetinnikov (1986a, b, 1989), mostró que mientras que esta especie se desplaza más al sur del Perú y a aguas oceánicas aumenta la importancia de mictófidos y más al norte aumenta el canibalismo. Otros investigadores en el pacífico este se han encontrado los mismos grupos dietarios (Ehrhardt *et al.* 1983,1991, Nigmatullin *et al.* 2001, Markaida y Sosa-Nishizaki 2003, Chong *et al.* 2005).

¹⁸ Migraciones verticales: son migraciones que se llevan a cabo por algunas especies marinas para buscar alimento, durante el día se profundizan y en la noche suben a la superficie.

¹⁹ Eufáusidos: crustáceos miembros del zooplancton parecido al krill.



Figura 4. Principales predadores y presas del calamar gigante *Dosidicus gigas* (Fuente: IMARPE).

2.2 ANTECEDENTES OCEANOGRÁFICOS

2.2.1 EL SISTEMA DE LA CORRIENTE DE HUMBOLDT

El sistema de la Corriente de Humboldt (SCH) abarca las costas de Chile y Perú, es uno de los sistemas costeros de surgencia²⁰ más productivos del mundo. El norte del sistema la corriente de Humboldt (NSCH) frente a Perú representa menos del 0,1% de la superficie oceánica mundial, pero su producción es del 10% de las capturas mundiales de pescado (Chávez *et al.* 2008). Esta alta producción en peces no se dan en los otros tres sistemas de afloramientos²¹ más importantes del mundo (Benguela, California y Canarias) y la mayor productividad de la pesca de la NHCS no corresponde a una mayor productividad primaria (Chávez *et. al.* 2008). Tal comparación hace suponer que existe una transferencia de energía de un nivel a otro más eficiente en el NSCH que en los demás sistemas de afloramiento.

Esta producción biológica excepcional, se debe a la presencia de zonas de surgencia o afloramiento que se producen principalmente por la acción del viento, el que desplaza

²⁰ Surgencia: afloramiento costero conjunto de nutrientes que migran desde el fondo a la superficie por acción del viento.

²¹ Sistemas de afloramiento: véase el punto 18, referido a surgencia.

grandes cantidades de aguas superficiales, creando un espacio que es llenado por el ascenso de aquellas que se encuentran en profundidades que fluctúan entre los 150 y 200 metros. Estas aguas son frías y muy ricas en nitratos y fosfatos, sustancias que provienen de la descomposición de los organismos del mar y/o de los residuos que llegan a él y que son básicas para el desarrollo de la vida. Las aguas ricas en nutrientes, al llegar a la superficie y por acción de la energía solar, facilitan la proliferación de las algas microscópicas que forman el fitoplancton, iniciándose así una cadena alimentaria. Este sistema tiene otras particularidades, algunas de las cuales han sido sugeridas como responsables de la alta producción (Bakun y Weeks 2008, Chávez *et al.* 2008, Chaigneau *et al.* 2009, Ballón 2010): (i) es el sistema de surgencia más cercano al Ecuador, (ii) tiene la más superficial e intensa zona de mínimo oxigeno (ZMO)²², (iii) tiene la más estable y menos intensa variabilidad física a mesoescala²³ comparada con otros sistemas de surgencia, y (iv) está íntimamente ligado a las dinámicas del Pacifico ecuatorial y está sujeto a fluctuaciones largas interanuales y multi-décadas.

2.2.2. ZONA MÍNIMA DE OXÍGENO

Los océanos incluyen extensas áreas llamadas zonas de mínimo oxígeno (ZMOs) que son zonas de contenido reducido de oxígeno situadas a nivel de los márgenes orientales de los océanos tropicales, a una profundidad de 20 a 800 metros (Figura 5). Las ZMOs se sitúan debajo de una capa superficial surtida de oxígeno disuelto. La frontera entre estos dos dominios forma la oxiclina²⁴ (Bertrand *et al.* 2010). La casi ausencia de oxígeno en la ZMO obliga a muchas especies a concentrarse en la capa superficial oxigenada, en ocasiones, teniendo tan sólo una decena de metros de altura. Este fenómeno podría agravarse debido a que las ZMO se amplían bajo el efecto del cambio climático. Algunas especies de zooplancton, peces mesopelágicos y calamares han adaptado su metabolismo para habitar temporalmente o permanentemente en la ZMO e incluso algunas la utilizan como un refugio a la depredación. Incidentalmente, la escasez de oxígeno en el fondo preserva los organismos muertos que se sedimentan y cuyo estudio permite reconstruir las condiciones climáticas y de productividad pasadas (Chávez *et al.* 2008).

²² Zona de mínimo oxígeno: son masas de aguas que, se encuentran entre los 100 y los 400 metros de profundidad.

²³ Mesoescala: escala espacial correspondiente entre los 100 y 1000 km.

²⁴ Oxiclina: capa del océano donde el oxígeno disminuye bruscamente con la profundidad.



Figura 5. Esquema del efecto de la Zona Mínima de Oxigeno (ZMO) en la distribución de las especies según día y noche (Fuente: Ballón *et al.* 2011).

2.3 ÍNDICE DE DIVERSIDAD DE MARGALEF (d)

El índice de riqueza de Margalef (*d*) permite comparar la cantidad de especies independientemente del número de individuos (N) (Margalef 1974). Se estimó a partir de la ecuación (2.1):

$$d = \frac{S-1}{LnN} \tag{2.1}$$

2.3.1 CÁLCULO DEL ÍNDICE DE LLENURA ESTOMACAL

La llenura estomacal es un índice que indica el consumo en relación al peso total de cada individuo. Generalmente suele ser mayor en las primeras etapas de crecimiento de las especies. Esta información permitió estratificar el consumo para cada parámetro como son: tallas (cada 10 cm), latitud (cada dos grados), de distancia a la costa (cada 30 millas náuticas), periodos de tiempo (anual, estacional), periodicidad horaria (día, atardecer, noche y aurora). El cálculo se realiza según la ecuación (2.2) dividiendo el peso total del contenido estomacal, que es la sumatoria del peso individual de cada presa, dividido entre el peso del individuo eviscerado por cien obteniendo un porcentaje de "*Fullness Weight Index*", (*FWI*, en porcentaje) (Rasero *et al.* 1996).

$$FWI = \frac{SCW \times 100}{BW - SCW}$$
(2.2)

Dónde:

SCW = peso de contenido estomacal (g)

BW= peso total del individuo (g).

2.4 ANTECEDENTES ESTADÍSTICOS

2.4.1 GRÁFICO DE PROBABILIDADES NORMALES DE RESIDUALES

El gráfico de probabilidad normal (ecuación 2.3) consiste en un gráfico de los escores normales (estadísticos de orden normales) versus los residuales estudiantizado internamente ordenados en magnitud. Los escores normales representan los valores esperados de observaciones ordenadas que provienen de una distribución normal estándar. El i-ésimo score normal es aproximado por:

$$z_{(i)} = \Phi^{-1} \left(\frac{i - \frac{3}{8}}{n + \frac{1}{4}} \right)$$
(2.3)

Donde Φ representa la función de distribución acumulada de una normal estándar y n (n>5) es el número de observaciones en la muestra. Para que cumpla con el supuesto de normalidad los puntos deben estar alineados alrededor de una recta que pasa por el origen. Si la tendencia de los puntos es curveada entonces la distribución es asimétrica. El tipo de asimetría es determinada por el lado donde está la parte curveada. Un gráfico de normalidad que produce una curva en forma de S indica que la distribución tiene una cola pesada o liviana dependiendo de la forma de la *S*. Si la *S* es alargada entonces la cola es liviana.

2.4.2 GRAFICO PARA VERIFICAR HOMOGENEIDAD DE VARIANCIA

a. Residuales contra los valores ajustados.

Los residuales no deben tener relación con estas variables y la dispersión de los residuales debe ser el mismo en todo el rango. De otra manera (la dispersión en forma de embudo, formas no lineales, residuales alejados horizontal o verticalmente) podría sugerir violación de supuestos de homogeneidad de variancia, falta de inclusión de términos no

lineales en el modelo, o presencias de valores altos (outlier) o extremos (leverage) en el modelo.

b. Raíz cuadrada de valores absolutos de residuales estudentizados internamente contra valores ajustados.

Es una línea de regresión no paramétrica. *Lowess* de los residuales estudentizados internamente en función de los valores ajustados. Si el supuesto de homogeneidad se cumple el *lowess* debe ser aproximadamente una línea horizontal; es decir, con pendiente cero.

c. Contraste de Breusch-Pagan (ncv.test)

Calcula una calificación en las pruebas de la hipótesis de la varianza del error constante contra la alternativa de que hay cambios de la varianza de error con el nivel de la respuesta (valores ajustados), o con una combinación lineal de los predictores. Esta prueba requiere de la librería car del paquete R.2.12.1 (Fox y Weisberg 2011). Esta prueba se conoce como la prueba de Breusch-Pagan, se propuso de forma independiente por Cook y Weisberg (1983).

2.4.3 PRUEBA DE NORMALIDAD DE SHAPIRO & WILK

Existe una extensa lista de pruebas estadísticas diseñadas para verificar la distribución normal de un conjunto de datos. Se usó la prueba de Shapiro & Wilk (Shapiro y Wilk 1965) por su simplicidad, las hipótesis para comprobar la normalidad de los datos sería:

Ho: Los datos presentan distribución normal

Ha: Los datos no presentan distribución normal

La prueba W de Shapiro & Wilk contempló los siguientes pasos:

- Ordenar en forma ascendente las *j* observaciones para la variable de estudio
- Calcular $\sum_{j=1}^{n} (X_j \overline{X})^2$ a partir de los datos del inciso 1
- Si j es par, j = 2k, calcular en (2.4):

$$b = \sum_{j=1}^{n} a_{n-j+1} (X_{n-j+1} - x_j)$$
(2.4)

- Si *j* es impar, j = 2k + 1, entonces se omite la mediana de los datos X_{k+1} , y se calcula el valor de *b* de acuerdo a la fórmula
- Calcular el valor de W_{cal} , el estadístico de prueba propuesto en la ecuación (2.5):

$$W_{cal} = \frac{b^2}{\sum_{j=1}^{n} (X_j - \overline{X})^2}$$
(2.5)

- Comparar W_{cal} con el valor de W_{tab} (los valores de W_{tab}). Si W_{cal} es menor que W_{tab} se rechaza la hipótesis de normalidad de los datos. Si se compara con un valor p teórico digamos de 0.05, se dice que se rechaza la hipótesis de normalidad de los datos si *p*-value calculado es menor que el valor teórico.

2.4.4 MÉTODOS DE SELECCIÓN DE VARIABLES

a. Selección "Forward" (Selección hacia adelante).

Aquí se empieza con la regresión no paramétrica simple que considera como variable predictora a aquella que está más altamente correlacionada (sin tomar en cuenta el signo) con la variable de respuesta. Si esta primera variable no es significativa entonces se para el proceso y se considera el modelo $\hat{Y} = \overline{Y}$, de lo contrario en el siguiente paso se añade al modelo la variable que reúne cualquiera de estos requisitos equivalentes:

- Aquella variable que tiene el estadístico de F o de T (sin tomar en cuenta el signo)
 más grande entre las variables no incluidas aún en el modelo.
- Aquella variable que produce el mayor incremento en el R2 al ser añadida al modelo.
- Aquella variable que tiene la correlación parcial más alta (en valor absoluto) con la variable de respuesta, tomando en cuenta las variables ya incluidas en el modelo.
 Toda variable que es añadida al modelo ya no puede salir. El proceso termina cuando se cumple una de las siguientes condiciones:

- Se llega a un modelo con un número prefijado p* de variables predictoras.
- El valor de la prueba de F para cada una de las variables no incluidas aun en el modelo es menor que un número prefijado F-in (por lo general este valor es 4, o el F correspondiente a un nivel de significación prefijado, digamos 15%). O en forma equivalente se para cuando el valor absoluto del estadístico de t es menor que la raíz cuadrada de F-in (por lo general, |t|<2).

b. Criterio de Información de Akaike (AIC)

El criterio de información de Akaike (Akaike 1973), tiene su origen en conceptos de teoría de información y está basado en la minimización de la distancia Kullback-Leibler entre la distribución de la variable de respuesta *Y* bajo el modelo reducido y bajo el modelo completo. Se define en (2.6):

$$AIC = -2ln L + 2 * (p)$$

(2.6)

Dónde:

p = número de parámetros

L = máximo del log likelihod

Este índice evalúa tanto el ajuste del modelo a los datos como la complejidad del modelo. Mientras menor sea el valor de AIC el ajuste del modelo será mejor.

c. Desvianza explicada

Como en GAM no se tiene un coeficiente de determinación, un estadístico similar es la desvianza explicada, llamado también como pseudo R^2 (Dobson 2002), que nos da una idea de la variabilidad de los datos y se calcula como:

Desviación explicada =
$$100 \times \frac{\text{desviación nula - desviación residual}}{\text{desviación nula}}\%$$
 (2.7)

Donde la desviación nula se calcula igual que la desviación residual pero en vez del modelo ajustado utiliza el modelo nulo.

2.4.5 MODELOS LINEALES GENERALIZADOS

Los modelos lineales generalizados (Nelder y Wedderburn 1972) unifican los modelos tanto con variable respuesta numérica como categórica, por lo cual se considera como distribución para la variable respuesta a parte de la Normal, otras distribuciones tales como; la Binomial, Poisson, Logística, etc. Los modelos lineales generalizados, son una extensión de los modelos lineales clásicos que permiten modelar variables de respuestas tanto cuantitativas y cualitativas

Estos modelos consisten en un componente aleatorio, un componente sistemático y la función de enlace que une los dos componentes. La respuesta *Y* de la variable aleatoria asume que su función de densidad pertenece a la familia exponencial.

$$f_{Y}(y;\theta;\phi) = \exp\left\{\frac{y\theta - b(\theta)}{\alpha(\phi)} + c(y,\phi)\right\},$$
(2.8)

Donde θ es llamado el parámetro natural y φ es el parámetro de escala. Este es el componente aleatorio del modelo. También suponemos que el valor esperado de *Y*, denotado por μ , está relacionado con el conjunto de covariables X_I , X_2 , ..., X_p por $(g)\mu = \eta$, donde $\eta = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p$. η llamado predictor lineal es el componente sistemático y g(.) es la función de enlace. Nótese que la media μ está relacionada al parámetro natural θ por $\mu = b'(\theta)$; Además, la función de enlace más utilizada para un *f* dado se llama vínculo canónico, por lo que $\eta = \theta$. Para definir el modelo en términos de μ y $\eta = (g)\mu$ y por lo tanto θ no juega un rol importante. Por lo tanto, cuando sea conveniente vamos a escribir $f_Y(y, \theta, \varphi)$ como $f_Y(y, \mu, \varphi)$.

La estimación de la $\mu = E[Y_i]$ no implicará el parámetro de escala φ , por lo que para simplificar esto se asume conocida.

Teniendo en cuenta las especificaciones del componente aleatorio y sistemático, considerando una función de enlace y un conjunto de *n* observaciones, la estimación de máxima verosimilitud de $\hat{\beta} = (\hat{\beta}_o, \hat{\beta}_1, ..., \hat{\beta}_p)$ se puede encontrar por un procedimiento de puntuación de Fisher. Sea $\hat{\eta}$ (una estimación actual del predictor lineal), con un valor ajustado correspondiente a $\hat{\mu}$, que define la variable dependiente ajustada.

$$z = \hat{\eta} + (y - \hat{\mu}) \left(\frac{dn}{d\mu}\right).$$
(2.9)

Definiendo los pesos W por:

$$(W)^{-1} = \left(\frac{dn}{d\mu}\right)^2 V, \qquad (2.10)$$

Donde V es la varianza de Y_i en $\mu_i = \hat{\mu}$. El algoritmo procede mediante una regresión de z en 1, x_1 ,..., x_p con los pesos W para obtener una estimación de $\hat{\beta}$. Usando $\hat{\beta}$, y un nuevo $\hat{\mu}$ y \hat{n} se calculan; un nuevo z se calcula y se repite el proceso, hasta que el cambio en la desviación sea suficientemente pequeño.

$$dev(y_i, \hat{\mu}) = 2[l(y_i) - l(\hat{\mu})]$$
 (2.11)

De lo anterior, $l(\mu_i)$ es la función log verosimilitud de probabilidad $\sum \log f_Y(y_i,\mu_i,\phi)$. Nelder y Wedderburn (1972) muestran que el algoritmo ajustado de la variable dependiente es equivalente al sistema de puntuación de Fisher, que es lo mismo a la secuencia de las estimaciones. Una descripción completa de los modelos lineales generalizados se da por McCullagh y Nelder (1983).

2.4.6 MODELOS ADITIVOS GENERALIZADOS (EXTENSIÓN DE LA SUAVIZACIÓN DE LOS MODELOS LINEALES GENERALIZADOS).

Considerando que el modelo de regresión lineal normal y el modelo de regresión lineal generalizado asumen una relación lineal (o alguna otra paramétrica) para formar covariables X_1 , X_2 , ..., X_p desde una regresión lineal usando un modelo con solo una variable predictora, que está dada por:

$$Y_i = \beta_0 + \beta_j X_{ij} + \varepsilon_i, \text{ donde } \varepsilon_i \stackrel{a}{\sim} N(0, \sigma^2)$$
(2.12)

Donde Y_i es la variable respuesta proveniente de una distribución de la familia exponencial, α es el intercepto, β_j es la pendiente, X_{ij} es la variable predictora y ε_i es el

residuo. Se asume teóricamente en las regresiones lineales, que la distribución de la variable respuesta es normal y la relación entre la variable respuesta y predictiva es lineal.

En el caso de los modelos de clase aditivos generalizados (GAM), se sustituye a la forma lineal $\sum \beta_j X_{ij}$ por una suma de funciones suavizadas $\sum s_j(X_{ij})$, los $s_j(.)$'s son funciones no especificadas que se estiman utilizando un diagrama de dispersión más suave, en un proceso iterativo que llamamos el algoritmo de calificación local.

El modelo aditivo generalizado extiende el componente sistemático de un modelo lineal generalizado basado en la función de suavización de las covariables (Hastie y Tibshirani 1986).

Para comprender los modelos aditivos generalizados, la forma más fácil de iniciar es desde una regresión lineal usando solo una variable predictora, la primera y más notable diferencia es que se usa una función para encontrar la relación entre Y_i y X_i , donde:

$$Y_i = \beta_0 + f(X_{ij}) + \varepsilon_i, \text{ donde } \varepsilon_i \stackrel{a}{\sim} N(0, \sigma^2)$$
(2.13)

La única diferencia a simple vista entre ambos modelos es el reemplazo de βX_i por $f(X_i)$ en la ecuación (2.13). Sin embargo, este cambio es de fundamental importancia. La regresión lineal da como resultado una fórmula y la relación entre las variables respuesta y predictora es cuantificada por un solo parámetro de regresión más intervalos de confianza. De otro lado, en GAM se tiene una función que relaciona la variable respuesta con las variables predictoras mediante suavizadores. Donde $f(X_i)$ es calculada usando una funciones base:

$$f(X_i) = \sum_{j=1}^{p} \beta_j \times b_j(X_i)$$
(2.14)

Donde β son los parámetros y los $b_j(X_i)$ llamadas funciones básicas son los suavizadores. Existen muchas clases de suavizadores para implementar los GAM. Entre ellos los más utilizados son los *spline* regresivos penalizados debido a que el grado de suavizado es estimado como parte del ajuste del modelo (Wood y Augustin 2002).

Los spline regresivos penalizados (Wood, 2006) están expresados por:

$$||Y - X\beta||^2 = \lambda \int f''(x)^2 dx$$
 (2.15)

En la ecuación (2.15) al estimar los parámetros β también se estiman los suavizadores *X*. La primera parte de la Ecuación (2.15), $||Y - X\beta||^2$, se refiere al ajuste de la curva mediante el método de mínimos cuadrados. En la suma de cuadrados se obtiene los parámetros β_j , y es útil para medir cuan cercana es la estimación de los datos (Venables y Dichmont 2004). La suma de los cuadrados a minimizar está dada por:

$$S = \sum_{j=1}^{n} (Y_j - \mu)^2 = \|Y - \mu\|^2 = \|Y - X\beta\|^2$$
(2.16)

Donde: Y_j es el valor observado o valor estimado para cada observación, y *j* es el número de observaciones. En la ecuación (2.16), la notación || || es debido a la norma Euclideana; **Y** contiene todos los valores observados en formato de vector, $\boldsymbol{\beta}$ son los parámetros y **X** son las funciones base o suavizadores.

La segunda parte de la Ecuación (2.15) $(\lambda \int f''(x)^2 dx)$ es una penalización, por esta razón se llaman *spline* penalizados, y mide el grado de "rudeza" en la estimación de la función (Venables y Dichmont 2004). Este contiene un parámetro del suavizado (λ) , λ es una constante de ajuste que afecta el balance entre la precisión y el suavizado. En el caso de utilizar validación cruzada general (GCV) no tiene que ser estimado directamente. Este parámetro de suavizado (λ) es afectado por una integral de la derivada de segundo orden de la función de suavizado *f* que muestra cuán suavizada es la curva.

Un alto valor de f'' significa que el suavizador f es altamente no-lineal y la curva es muy flexible mientras que un valor de cero en la segunda derivada resulta en una línea recta. Por otro lado, si λ es muy grande, la penalización por no tener una curva no suavizada es grande y el suavizador resultante puede ser una línea y si λ es pequeño, luego hay una baja penalización para la no suavización y probablemente se obtendrá una curva menos suavizada.

Este control en los dos sentidos hace necesario que se busque un punto intermedio entre baja penalización y óptimo suavizado. Para obtener el valor óptimo de suavizado sin tener que calcular λ ; Wood (2006) implementó el cálculo del GCV para estimar los parámetros de suavizado. En lugar de proveer un resultado de la validación cruzada en términos de λ , las modificaciones implementadas por Wood (2006) utilizan un término llamado grados de libertad efectivos (*edf*). Este es un valor entre 0 e infinito, un alto número de *edf* significa que es menos lineal la función de suavizado utilizada. Para variables respuesta discretas el cálculo del GCV se modifica el estimador imparcial de riesgo.

Una penalización aun mayor son los *spline* regresivos penalizados con *shrinkage* en los cuales se puede definir la cantidad máxima de suavización antes del cálculo del GCV.

En GAM, además de los parámetros de suavizado, la asunción de una familia de distribución de la variable respuesta (Y_i) es vital ya que no se asume que las variables tienen distribución normal. La distribución puede ser cualquiera dentro de la familia exponencial.

La elección de la familia de distribución para la variable respuesta debe ser hecha *a priori* basada en el conocimiento de la variable respuesta (Hastie y Tibshirani 1990, Peltonen *et al.* 2007). Para variables binarias la elección es obvia (distribución de Bernoulli) pero hay muchas más opciones cuando la variable es discreta o continua (Zuur *et al.* 2009) (Cuadro 1). Otro aspecto importante es que la relación entre la variable respuesta y la variable predictora con función de enlace conocida también puede variar. Las funciones de enlace son las responsables de modelar la relación entre la variable respuesta y las variables predictoras. Hay varias opciones en la función de enlace para una misma familia de distribución, por lo tanto se recomienda explorar antes la relación entre las variables predictoras y la respuesta (Cuadro 1).

Distribución	Tipo de datos	Funciones de enlace
Normal	Continuos	identity, log, inverse
Poisson	Discretos (Conteo)	log, identity, sqrt
Quasipoisson	Discretos (Conteo)	logit, probit, cloglog, identity, inverse, log, sqrt
Negativa binomial	Discretos sobredispersos	log, sqrt, identity
Gamma	Continuos	inverse, identity, log
Binomial	Proporcionales	logit, probit, cauchy, log, cloglog
Bernoulli	Presencia y ausencia	logit, probit, cauchy, log, cloglog

Cuadro 1. Familia de distribución y funciones de enlace correspondientes.

Según el Cuadro 1 para una variable respuesta continua puede utilizarse una distribución gamma o una distribución normal. Una aproximación primaria para elegir entre ambas familias es realizar un análisis exploratorio de normalidad a la variable respuesta debido a que a según esta distribución se va a tratar de modelar las variables predictoras. Cuando se utiliza un GAM modelado con distribución normal para la variable respuesta transformada previamente mediante $log_{10}(\mu_i+1)$ y una función de enlace identidad, el modelo estaría definido por:

$$identity(\log(\mu_i + 1) = \alpha_i + f(X_i) + \varepsilon_i$$
(2.17)

Donde:

$$E(Y_i) = \mu_i \text{ y } Var(Y_i) = \sigma_i^2$$

 μ_i es la media y $var(Y_i)$ es la varianza. La transformación de la variable respuesta es utilizada para disminuir la variabilidad.

Si se utilizara la distribución Gamma, esta distribución puede ser asumida para variables *Y* de respuestas continuas que tengan valores positivos ($Y_i > 0$). Para una distribución Gamma con función de enlace "*log*" el GAM estaría definido por:

$$\log(\mu_i) = \alpha_i + f(X_i) + \varepsilon_i \tag{2.18}$$

Donde:

$$E(Y_i) = \mu_i \text{ y } Var(Y_i) = \frac{\mu^2}{v}$$

 μ_i es la media y $var(Y_i)$ es la varianza, v es un parámetro de dispersión. Un valor pequeño de v indica que la dispersión de los datos es grande y un valor grande de v da como resultado una curva simétrica y en forma de campana, en ese caso una distribución normal puede ser usada; sin embargo, los valores de Y_i deben ser mayores a 0 (cero).

Para corroborar que la distribución y la función de enlace son adecuadas para el modelo se debe validar el modelo mediante el análisis de los residuos.

Las formas graficas de verificación de un buen ajuste del modelo pueden ser graficas de residuos versus cada una de las variables predictoras. En un buen ajuste del modelo estos gráficos no deben mostrar patrones. Si existen patrones en la anchura de banda de la gráfica, significa que existe una sobredispersión de los datos (se deben
eliminar los valores atípicos) o el uso de la relación entre la media y la varianza es incorrecta produciendo heterogeneidad (se ha elegido una distribución incorrecta). Una relación con una curvatura determinada puede indicar que existe un patrón en los datos no incluido en el modelo). Los residuos comúnmente utilizados son los residuos de desvianza, debido a que son cercanos a la normalidad.

Los Modelos Aditivos Generalizados (GAM) tienen aplicación en muchos campos de regresión con respuesta perteneciente a la familia exponencial (Binomial, Gaussiana, Poisson, Binomial Negativa, etc.). Esto modelos pueden ser utilizados con fines exploratorios, predictivos, explicativos o bien para detectar posibles interacciones entre los efectos de covariables. Su aplicación es muy amplia y se extiende a todos los campos de la Biología.

III. MATERIALES Y MÉTODOS

3.1 MATERIALES

3.1.1 MATERIAL BIOLÓGICO

Se empleará la información del IMARPE sobre la dieta del calamar gigante en individuos muestreados entre verano de 2004 y primavera de 2009 procedentes del seguimiento de pesquería industrial a lo largo del litoral peruano.

3.1.2 MATERIAL DE LABORATORIO

- Tamiz de acero inoxidable de 500 μ
- Microscopio estereoscópico (60X y 120X)
- Balanza analítica de 0.001 g de precisión.
- Placas Petri de 20 cm de diámetro.

3.1.3 MATERIAL DE ESCRITORIO

- Claves y listas taxonómicas adecuadas para el reconocimiento de las taxa-presa.
- Software: R 2.12.1.
- Computadora Core Duo.

3.2 POBLACIÓN Y MUESTRA

La población objetivo del estudio estuvo constituida por los calamares gigantes que habitan entre 3° y 18°S en la zona económica exclusiva (ZEE) desde 14 hasta 200 mn y fuera de la ZEE, hasta las 299 millas náuticas (mn). La muestra está conformada por el contenido estomacal de 4178 calamares gigantes capturados por la flota industrial pesquera desde el año 2004 a 2009 (Figura 6).

La colecta de calamares se realizó a bordo de embarcaciones calamareras de bandera coreana y japonesa, utilizando artes de pesca dirigidas a atraer mayor cantidad del

recurso, en este caso máquinas poteras llamadas *jiggins*²⁵, la concentración del recurso alrededor de este arte de pesca se consigue con la ayuda de potentes luces (focos de 2 KW).

Una vez realizada la captura se procedió a hacer la colecta de calamares para posteriormente extraer los estómagos, por el método de muestreo aleatorio, agrupando 20 individuos por cada estación de muestreo, teniendo en cuenta colectar individuos de todos los rangos de tallas registrados en el momento de la captura. Para efectos de este estudio los ejemplares se agruparon por estratos según talla (cada 10 cm), latitud (cada 2° de Latitud Sur – LS), distancia a la costa (cada 30 millas náuticas – mn), intervalo de tiempo (anual, estacional), periodicidad horaria (día, atardecer, noche y aurora). Paralelamente se tomó información de temperatura, posición geográfica y hora. En base a estos grupos se utilizó la información obtenida del cálculo de llenura estomacal (Rasero *et al.* 1996) y los índices de diversidad (Margalef 1974).

En el 80% de los muestreos, los individuos completos fueron colectados, en el 20% restante el muestreo biológico fue realizado a bordo y sólo fueron colectados los estómagos. En cada caso, la muestra completa fue congelada a bordo a -2°C y posteriormente enviada al Laboratorio de Ecología Trófica de la sede central del IMARPE.

²⁵ Las *jiggins* constan de una línea de pesca que a lo largo tiene anzuelos con carnadas luminosas especiales para pesca de calamares.



Figura 6. Estaciones de muestreo de estómagos de calamar gigante *Dosidicus gigas* a lo largo del litoral del 2004 al 2009.

Hay que tener en cuenta que el tipo de arte de pesca descrita y utilizada por la flota industrial para capturar el calamar gigante, ocasiona que este recurso se concentre demasiado en la zona de pesca, incrementando artificialmente el canibalismo. En efecto, muchos autores (Bennett 1978, Merdsoy 1978, Ehrhardt *et al.* 1983 a, 1986, Bazanov, 1986 Cubillos *et al.* 2004) han enseñado que cuando la flota calamarera utiliza focos para atraer a los recursos, se genera una alta concentración de calamares que aumenta artificialmente el canibalismo.

Para poder estudiar la dieta real de *D. gigas* y eliminar el canibalismo inducido por la pesca se trabajó en dos etapas. Primero se eliminó de los estómagos los trozos frescos de *D. gigas* (trozos carnosos de color rosáceo de $0.5-1 \text{ cm}^3$, ver Baral 1967). En efecto, la presencia de estos trozos frescos indica que fueron ingeridos recientemente y que el canibalismo se dio durante las maniobras de pesca. A pesar de este proceso, el análisis de la base de datos indica una tasa de canibalismo muy alta (>75%), incompatible con los desarrollos de una población en términos energéticos. Estos resultados preliminares indicaron que no se eliminó totalmente el efecto del canibalismo inducido por la pesca. En efecto, el calamar gigante tiene una alta tasa de digestibilidad. Así, Bazanov (1986) estableció las etapas de digestión y tiempo de evacuación de 1 a 2 horas. Esta tasa hace que se pueda encontrar presas en un grado de digestión medio, debido al tiempo que dura la maniobra de pesca. Por tanto, podría existir una sobrestimación del canibalismo.

Para contar con una base de datos más fiable se hizo una segunda selección. Para eso se estudió la distribución de los contenidos estomacales del calamar gigante según el tipo de presa: *D. gigas* versus todas las otras especies. Los resultados (Figura 7) enseñaron una forma de 'exclusión' con 18.1% de estómagos conteniendo 100% de *D. gigas*. Estos estómagos, cuyo único ítem-presa fue *D. gigas*, fueron eliminados para disminuir el sesgo del canibalismo inducido por la pesca, quedándonos con 3421 estómagos.



Figura 7. Gráfico de dispersión del contenido del estómago del calamar gigante (peso en gramos) sobre la base de *D. gigas* y otras especies (sin el canibalismo) de 2004 a 2009.

3.3 PROCESAMIENTO DE MUESTRAS

El análisis biológico se realizó midiendo la longitud total del manto (en centímetros), el peso total del individuo (en gramos), el sexo y el estadio de madurez sexual. Luego se extrajo el estómago (Figura 8) y cada contenido estomacal se colocó en un tamiz de malla de 500 µm para ser lavado con agua corriente; una vez lavada la muestra se procedió a separar manualmente los diferentes elementos de la dieta tales como otolitos (*sagitta*), huesos, picos de cefalópodos y cefalotórax de crustáceos. Luego, los elementos dietarios fueron vertidos en placas *Petri* y observados con microscopio estereoscópico (60X y 120X) para su determinación al mínimo taxón posible con el empleo de bibliografía especializada para teleósteos (Fitch y Brownell 1968, Chirichigno 1998, García-Godos 2001), cefalópodos (Wolf 1984), crustáceos (Newel 1963, Méndez 1981), zooplancton (Trégouboff y Rosse 1957) y moluscos (Alamo y Valdivieso 1987) (Figura 9).



Figura 8. Disección del estómago de calamar gigante Dosidicus gigas

Las presas debidamente identificadas fueron cuantificadas en número y peso para cada estómago. Luego, la información fue procesada y expresada en cantidades relativas de acuerdo a los criterios de agrupamiento de tallas, trimestre y zona de procedencia según los métodos numérico (% N), gravimétrico (% P) y de frecuencia de ocurrencia (FO) (Berg 1979, Hyslop 1980).



Figura 9. Metodología de muestreo y separación de estructuras duras de peces y cefalópodos utilizadas para la identificación indirecta de las especies presa del calamar gigante *Dosidicus gigas*.

3.4 PROCESAMIENTO DE DATOS

3.4.1 FORMULACIÓN DE HIPÓTESIS

Hipótesis Planteada (H_{θ}):

La relaciones entre la llenura estomacal y los índices de diversidad de las presas del calamar gigante (*Dosidicus gigas*) respecto a la talla, latitud, distancia a la costa, tiempo (años, estación) y temperatura son aleatorias.

Hipótesis Alternativa (*H*₁):

Existen relaciones en la llenura estomacal y los índices de diversidad de las presas del calamar gigante (*Dosidicus gigas*) respecto a la talla, latitud, distancia a la costa, tiempo (años, estación) y temperatura.

3.4.2 MODELO GAM

Se aplicó el modelo GAM (Hastie y Tibshirani 1990, Venables y Ripley 2002, Wood y Augustin 2002, Venables y Dichmont 2004, Wood 2006), que es definido como un método semiparamétrico que mediante funciones suavizadas permite determinar la forma de la relación entre variables ambientales y bióticas (Swartzman 1997, Piet 2002, Cianelli *et al.* 2008), por ello en este estudio se busca encontrar posibles influencias entre la llenura de estómago y diversidad en la dieta, con respecto al tamaño del individuo, la hora del día, la latitud, la distancia a la costa y la temperatura superficial del mar (TSM) de acuerdo a la forma:

$$Y_i = \beta_0 + \sum_{j=1}^p f_j(x_i) + \varepsilon_i.$$
 (3.1)

 $log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + f_6(TSM) + \varepsilon_i$

 $log(E[diversidad]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + f_6(TSM) + \varepsilon_i$

Los datos de dieta del calamar gigante reúnen gran complejidad, no son balanceadas y se espera que no tengan relaciones lineales con las variables ambientales. El *"Generalyzed Additive Models*" GAM (Hastie y Tibshirani 1990) es un tipo especial de regresión semiparamétrica que soporta la típica falta de normalidad de los datos de dieta y son muy útiles para analizar grandes cantidades de información (Stefánsson y Pálsson 1997).

El modelo aditivo es una generalización de un modelo lineal, y por tanto, tiene una característica importante: el efecto de una variable predictora en una variable respuesta se suma a los efectos de cualquier otra variable. En otras palabras, podemos verificar la contribución individual de cada variable predictora en la explicación de la variabilidad de la variable respuesta.

En esta clase de modelos, hay tres tipos de enfoques: sólo paramétrico, estrictamente no paramétrico y una mezcla de los dos métodos anteriores, como se muestra a continuación.

De manera similar a lo establecido en el modelo de regresión considerando *n* pares de observaciones (x_i, y_i) , i = 1, 2, ..., n, y vamos a tener una función *f* que establece la relación entre las variables *x* e *y* de la forma:

$$y_i = f(x_i) + \mathcal{E}_i, \tag{3.2}$$

Siendo *f* una función que se estima y suponiendo que los errores aleatorios ε_i se tiene $E(\varepsilon_i) = 0$ y $Var(\varepsilon_i) = \sigma^2$. Si tenemos en cuenta un conjunto de *k* variables explicativas representadas en una matriz **X** de dimensión $n \times k$ y el rango completo con *i-ésima* fila dada por $X_i = x_{i1}, x_{i2}, ..., x_{ik}$, tenemos una función *f* como tal que $y_i = f(x_{i1}, x_{i2}, ..., x_{ik}) + \varepsilon_i$.

Para construir un modelo lineal de las variables regresivas, Buja *et al.* (1989) y Hastie y Tibshirani (1990) consideran la función f como una suma de funciones f_j , j = 1,...,k, para cada una de las columnas de X, mientras que para la *i-ésima* línea de X tenemos:

$$f(X_i) = f_1(x_{i1}) + f_2(x_{i2}) + \dots + f_k(x_{ik}), \qquad i = 1, 2, \dots, n,$$
(3.3)

Entonces el modelo (3.3) queda expresado por:

$$y_i = f_1(x_{i1}) + f_2(x_{i2}) + \dots + f_k(x_{ik}) + \varepsilon_i, \qquad i = 1, 2, \dots, n,$$
(3.4)

Donde $y^{T} = (y_1, ..., y_n)$ es un vector *n* x 1 de respuestas y una *i-ésima* fila de la matriz *X* es $X_i = (x_{i1}, ..., x_{ik})$ es un vector *n* de variables explicativas con $E(\varepsilon_i) = 0$ y $Var(\varepsilon_i) = \sigma^2$ (Bidurin y Andre 2001). Este modelo es llamado por Buja *et al.* (1989) de regresión aditivo o modelo aditivo.

El predictor obtenido en el modelo corresponde a un modelo totalmente no paramétrico y cada función f_j para j = 1,...,k es una función univariada arbitraria que va a ser estimado por medio de un procedimiento no paramétrico (mecanismo de suavización). Aquí, el problema es superado una vez que cada función f_j es estimada de modo univariado. Luego, el modelo aditivo generaliza el modelo de regresión lineal múltiple con la función lineal sustituida por una cantidad aditiva de funciones univariadas.

Los modelos cuyos predictores combinan formas paramétricas de algunas (g) variables predictoras en términos no-paramétricos de otras (k - g) variables son también parte de esta clase. En este caso, el predictor se puede escribir como:

$$y_{i} = \beta_{1}x_{i1} + \dots + \beta_{g}x_{ig} + f_{1}(x_{ig+1}) + \dots + f_{k-g}(x_{ik}) + \varepsilon_{i}$$
(3.5)

Estos modelos son denominados semiparamétricos, también referenciados en la literatura como modelos parcialmente lineales (Speckman 1988).

Se añade a ésto, que los resultados asintóticos de la eficiencia y la consistencia de los estimadores obtenidos por Schick (1986, 1993 y 1996) y Bhattacharya y Zao (1997) han ayudado a consolidar la aplicación de modelos semiparamétricos, que han sido ampliamente discutidos por varios estadísticos de renombre (Stone 1985, Hastie y Tibshirani 1990, Lee 1990).

a. Función de densidad

La función de densidad de probabilidad caracteriza completamente un espacio muestral compuesto por una determinada variable aleatoria y se define como una función real medible no negativa y satisfaciendo:

$$\int_{-\infty}^{\infty} f(y) dy = 1.$$
(3.6)

Sea Y_i una variable aleatoria con función de densidad de probabilidad f. La especificación de la función f ofrece una descripción natural de la distribución de la variable y permite que las probabilidades asociadas con Y_i sean determinadas por la relación:

$$P(a < Y_i < b) = \int_a^b f(y) dy, \text{ para todo } a < b.$$
(3.7)

Se entiende por el proceso de estimación de la densidad a la construcción de una función de probabilidad que representa un determinado conjunto de datos. Con su identificación proporciona información valiosa en el análisis exploratorio, en aspectos tales como la multimodalidad, asimetría de tipo de cola larga, curtosis, entre otras características, es también un análisis confirmatorio o como instrumento indicativo para el uso de diferentes métodos (análisis discriminante, el análisis de conglomerados, pruebas de la moda, etc.).

Silverman (1986) también destaca que la función de densidad estimada tiene un efecto visual impactante que facilita la comprensión del problema, incluso por personas no familiarizadas con la estadística. Varios procedimientos no paramétricos para estimar la función de densidad de probabilidad están disponibles en la literatura (Silverman 1986,

Pagan y Ulah 1999, Härdle 1990) y se refieren a menudo como los métodos de suavización.

b. Método de suavización por splines

Para Lima *et al.* (2001), el método de suavización por *splines* es una suavización (también denotado en la literatura por suavizado) es una herramienta que describe la variación de la media de una variable Y_i en función una o más variables no estocásticas X_1 ,..., X_k . Cuando la variación de la media de Y_i se describe en términos de una sola variable X_i , el suavizador es denominado unidimensional. Cuando k variables, $X_1,...,X_k$, son consideradas como un suavizador multidimensional.

En la mayoría de los casos, el suavizador se utiliza para ajustar el modelo:

$$y_i = f(X_i) + \mathcal{E}_i,$$

$$i = 1, \dots, n, \qquad (3.8)$$

Donde *f* es una función no especificada y ε_i son errores aleatorios distribuidos de forma independiente con media cero y varianza σ^2 .

En gran parte de las técnicas de suavización, el valor suavizado \hat{y}_i se obtiene sobre la base a un promedio de *t* observaciones alrededor de un determinado valor x_i . Diferentes formas de cálculo de este promedio alrededor de x_i definen diferentes métodos de suavización.

La curva suavizada y curva ajustada se construye sobre la base de puntos (x_i, \hat{y}_i) , donde \hat{y}_i es el valor esperado (por suavización) de *Y* para $X = x_i$. Estos valores se obtienen sin la adopción de un modelo paramétrico sobre *Y* y *X*.

El término *splines* se tomó prestado de un dispositivo mecánico muy utilizado (antes de computación gráfica) para dibujar las secciones del casco y las líneas férreas. Los *spline* mecánicos constituían tiras flexibles de madera que eran forzadas a pasar por determinados puntos fijos, a fin de lograr la posición deseada (Bowman y Azzalini 1997). La idea detrás de este dispositivo era permitir a la confluencia de dos o más curvas de manera suave, pasando por puntos pre-fijados.

La implementación de esta idea para el ajuste del modelo se lleva a cabo con el fin de tener una regresión polinomial por "pedazos". Para trabajar con el ajuste y la interpolación de funciones, la aproximación por polinomios es muy conveniente, ya que los polinomios tienen muchas propiedades interesantes que permiten calcular las derivadas de cualquier orden de los polinomios. Sin embargo, la necesidad de muchas derivadas de la función que se aproxima por polinomios puede ser muy restrictiva. Una forma de evitar este problema es el uso de polinomios por partes, de esta manera se puede escapar del análisis de todo el intervalo, permitiendo discontinuidades de las derivadas de orden superior en algunos puntos. Esa característica de "*pseudo*-análisis" ofrece funciones polinómicas por partes, denominadas *splines*, con propiedades de aproximación, de estabilidad y convergencia (Cunha 2000).

La suavización por *splines* (*smoothing spline*) fue creada por Whittaker (1923). Sin embargo, fue Schoenberg (1964) que, utilizando los métodos desarrollados por Whittaker, obtiene un estimador de suavización por *splines* (Dias 2001). Las funciones *splines* están asociadas a la partición de un intervalo [a, b] de dominio D de f en que se pretende trabajar. Una partición I que será definida de los puntos $x_0, ..., x_k$ tal que:

$$a = x_0 < x_1 < \dots < x_{k-1} < x_k = b.$$
(3.9)

En cada subintervalo $[x_i, x_{i+1}]$, i = 0,...,k, los *splines* son polinomios de un determinado grado *d*. Estos pedazos de polinomios están convenientemente ubicados para que algunas derivadas de orden dictada por el problema, existan en todo el intervalo [a, b]. Hay una relación entre el grado de dos pedazos de polinomio y el orden de las derivadas necesarias en los puntos de la partición. Así, algunas restricciones deben ser impuestas en la definición general de *spline* (Cunha 2000).

La función s(x) es llamada de *splines* de grado *d*, con nudos (*knots*) en $\{x_i\}_{i=1}^k$, asociada a una partición de [a, b] si:

s (x) es un polinomio de grado d en cada intervalo de subíndices consecutivos [x_i, x_i
 +1];

- s(x) tiene d - 1 derivadas continuas en cada x_i , y por lo tanto en [a, b].

El conjunto de funciones S_d ($x_0,...,x_k$) es un espacio lineal y recibe el nombre de espacio *splines* (en inglés, *splines space*), en la que sus elementos son funciones *splines*. Aunque varias configuraciones sean posibles, una opción bastante usada es el *splines* cúbico (d = 3), que consisten en un conjunto de polinomios cúbico con primera y segunda derivadas continuas en los nudos.

Así que si $b_j(x)$ es la base de la *j-ésima* función que define su espacio s(x) puede ser representado por la expresión:

$$s(x) = \sum_{j=1}^{q} b_j(x) \beta_j,$$
(3.10)

para algunos el valor del parámetro desconocido de β_j . Por ejemplo, si *s* es una función polinomial de grado 5, su base está dada por $b_1(x)=1$, $b_2(x)=x$; $b_3(x)=x^2$; $b_4(x) = x^3$, $b_5(x)=x^4$ y $b_6(x)=x^5$. Con esto, la ecuación anterior puede ser expresada por:

$$s(x) = \beta_1 + x\beta_2 + x^2\beta_3 + x^3\beta_4 + x^4\beta_5 + x^5\beta_6.$$
(3.11)

c. Penalización de suavidad

Un buen ajuste a los datos no es el único objetivo en la estimación de una curva, existe otro objetivo, que es obtener una estimación que no oscile rápidamente.

Considere un modelo de regresión propuesto y supongamos que $\hat{f}(x)$ estima la función f(x). El criterio de bondad de ajuste puede ser dado por la suma de los cuadrados de los residuos

$$\sum_{i=l}^{n} \left[y_i - \hat{f}(x_i) \right]^2.$$
(3.12)

Si f es asumida con una forma sin restricciones (concepto no paramétrico), entonces el valor de la suma anterior se reduce a 0 para alguna curva en particular, y para algún comportamiento específico de datos. Como señalan Rosa y Soler (2004), solo, la suma de los cuadrados de los residuos no es un buen criterio de ajuste, ya que se trata de modelos numéricos creados originalmente para la interpolación.

Se añadió entonces un criterio de penalización para la falta de suavidad de ajuste a partir de derivadas de función *f*, propuesto por Handscomb (1966), de la forma:

$$\int_{a}^{b} \left[f^{m}(x) \right]^{2} dx, \qquad (3.13)$$

Para *a* y *b* tal que $a \le x_0 \le ... \le x_k \le b$ y *m* es un parámetro del orden de la derivada de la función *f* que está relacionada con el grado *d* de los polinomios que se ajustan; es decir, f^m denota la *m*-ésima derivada. En el caso de *splines* cúbico (d = 3 y m = 2), para penalizar la segunda derivada, o la aceleración de la curva.

Componiendo los dos criterios en una sola ecuación, tenemos la forma tradicional del *spline* suavizado, que en realidad es la búsqueda de una función f(x) con m derivadas continuas que minimiza la suma de los cuadrados de los residuos penalizados.

$$\sum_{i=1}^{n} \left[y_i - f(x_i) \right]^2 + \lambda \int_{a}^{b} \left[f^m(x) \right]^2 dx,$$
(3.14)

Sabiendo que la primera parte penaliza la falta de ajuste de la función de regresión a los datos y la segunda parte sanciona la falta de suavidad de la función $\hat{f}(x)$ y λ es el parámetro de suavización que determina la suavidad de la estimación, el control de cuanto se dirigió a la interpolación de los datos o en la dirección de suavizado excesivo.

Cabe destacar que los *splines* gozan de unas buenas propiedades tales como la existencia, unicidad y flexibilidad, además de ser fáciles de calcular computacionalmente en comparación con otros métodos de suavizado. En particular los *splines* cúbico (Eubank 1994) han sido ampliamente utilizados en la estimación no paramétrica y son bastante empleados para la solución de criterios anteriormente descritos.

Estudios sobre el comportamiento asintótico del estimador obtenido utilizando el método *spline* de suavizado se puede encontrar en Silverman (1984) y Eubank (1988). El primer autor también muestra que, bajo ciertas condiciones, el *splines* suavizado

corresponde aproximadamente a la suavización por *kernel* con ventana *h* en función de la densidad local entre dos puntos de observación (Hardle 1990).

d. Estimación del parámetro de suavización λ

Según Souza (2008), existen dos enfoques diferentes en la elección el parámetro de suavizado. El primer enfoque es la libre elección de suavización de parámetros como una característica importante del procedimiento. Lo que se hace es utilizar diferentes valores y por lo tanto elegir el que de alguna manera produce estimaciones con mejor ajuste de la curva visual sin problemas a los datos. Esto hace que este método sea subjetivo, sin embargo es ampliamente utilizado en la práctica. Constituye una óptima opción cuando se desea ajustar una única curva

Otro enfoque con la necesidad de contar con un procedimiento automático para la elección de λ en base de los datos. Se podría decir que condicionando la elección de método automático a ser usado, esta es una forma objetiva para elegir λ .

Entre los procedimientos automáticos para la elección del parámetro de suavizado λ , el más conocido de todos es el criterio GCV, que tiene este aspecto cuando la estimación de *f* está dada por el método *splines*:

$$GCV(\lambda) = \frac{\frac{1}{n} \sum_{i=1}^{n} \left[y_i - \hat{f}(x_i) \right]^2}{\left[I = \overline{h_i}(\lambda) \right]^2} = \frac{\frac{1}{n} \left\| (I - H_\lambda) y \right\|^2}{\left[\frac{1}{n} tr(I - H_\lambda)^2 \right]},$$
(3.15)

Siendo $\|\cdot\|$ la norma euclidiana, $\overline{h}(\lambda) = (1/n)tr(H_{\lambda}) y tr(H_{\lambda})$ es la traza de la matriz $H_{\lambda} = (X_{\lambda}^{T}X_{\lambda})^{-1}X_{\lambda}^{T}$. La elección del parámetro de suavizado se hace encontrando el valor λ de que minimiza el criterio $GCV(\lambda)$. También, es importante tener en cuenta que los métodos de validación cruzada están bien definidos para un conjunto de puntos $\{x_i\}_{i=1}^n$ distintos; y, por lo tanto, hay que tener cuidado en su aplicación para eliminar grupos de puntos iguales antes de realizar el procedimiento de optimización (Ramsay y Silverman 2006).

3.4.3 DISTRIBUCIONES TWEEDIE (APROXIMACIÓN SADDLEPOINT)

La familia de distribuciones *Tweedie* pertenece a la clase de modelos de dispersión exponencial (MDE), famoso por su papel en los modelos lineales generalizados. Las distribuciones *Tweedie* son los MDE con una variación de la forma $var[Y] = \varphi \mu^p$ donde pes mayor o igual a uno, o menor o igual a cero. Esta función sólo se evalúa para p mayor o igual a uno. Casos especiales incluyen la normal (p = 0), poisson ($p = 1 \text{ con } \varphi = 1$), gamma (p = 2) e inversa gaussiana (p = 3) distribuciones. Cuando el valor de p se encuentra entre 1 y 2 (1), la distribución continua de <math>Y es mayor que cero, con una masa positiva en Y = 0. Para p > 2, la distribución es continua en Y mayor que cero. Esta función se aproxima a la densidad de uso de la aproximación *saddlepoint* definido por Nelder y Pregibon (1987). El valor *saddlepoint* aproximado de densidad para la distribución *Tweedie* dado con el poder de los parámetros de μ y φ .

3.4.4 LIBRERÍAS DE "R" USADAS

Cuadro	2.	Librerí	as de	"R"	usadas
Cuadro	2.	Libreri	as de	"R"	usadas

Prueba estadística	Librería	Referencia
Scatterplot.matrix	MASS	Ripley et al. 2011
Contraste de Breusch-Pagan (ncv.test)	car	Fox y Weisberg 2011
Prueba de Kruskal - Wallis	agricolae	Mendiburu 2010
aproximación saddlepoint	tweedie	Dunn 2011
Modelo GAM	mgcv	Wood 2007
Modelo GAM con aproximación	tweedie	Dunn 2011

IV. RESULTADOS Y DISCUSIÓN

4.1 IDENTIFICACIÓN DE VARIABLES

4.1.1 VARIABLES DEPENDIENTES O RESPUESTA:

Las variables respuesta o dependientes usadas en los modelos fueron de tipo cuantitativa continua.

 Y_i = *Indices* de diversidad de contenido en la dieta del calamar gigante *Dosidicus gigas* (variable cuantitativa continua).

 Y_j = Llenura estomacal en peso por cada grupo de presas contenido en la dieta del calamar gigante *Dosidicus gigas* (variable cuantitativa continua).

4.1.2 VARIABLES INDEPENDIENTES O PREDICTORAS (COVARIABLES):

$X_1 =$	Talla (cm) del	calamar gigante	(variable cua	ntitativa continua)
---------	----------------	-----------------	---------------	---------------------

- X_2 = Tiempo (año, estación) (variable cuantitativa continua)
- X_3 = Hora de captura (variable cuantitativa continua)
- X_4 = Distancia a la costa (mn) (variable cuantitativa continua)
- X_5 = Latitud (°S) (variable cuantitativa continua)

 X_6 = Temperatura (°C), (variable cuantitativa continua)

4.2 VARIABILIDAD DE LA DIETA

4.2.1 DIETA GLOBAL

En el anexo 1 se puede apreciar la composición anual de la dieta del calamar gigante en términos de porcentaje en número, porcentaje en peso y frecuencia de ocurrencia desde el 2004 al 2009. De un total de 3421 estómagos analizados el 80.2% se encontraron llenos en individuos de 14.3 a 112.5 cm de longitud del manto (LM); determinándose en total 43 taxones presas. El calamar gigante (canibalismo) fue la presa dominante en el peso (%W) (35.4%), los eufáusidos fueron los más abundantes (% N) (62.2%), mientras que en términos de frecuencia (FO), los cefalópodos indeterminados

(34.6%), el pez mesopelágico *Vinciguerria lucetia* (36.4%) y los mictófidos en general (45.3%) fueron las presas más importantes.

La importancia de estos taxones presa es similar a lo observado por otros autores en el Pacífico Oriental (Ehrhardt *et al.* 1983a, Schetinnikov 1986b, Ehrhardt 1991, Nigmatullin *et al.* 2001, Markaida y Sosa Nishizaki 2003, Chong *et al.* 2005).

La importancia de los peces mesopelágicos en la dieta del calamar gigante fue reportada por varios autores (Markaida y Sosa-Nishizaki 2003, Chong et al. 2005, Markaida 2006, Fitch 1968, Fitch 1969, Coleman y Nafpaktitis 1972); estos peces tienen fotóforos que podrían atraer a los depredadores. Además, según Nesis (1970, 1983), el calamar gigante se alimenta de cualquier presa que se mueva, mientras sea abundante y grande. La mayoría de los peces mesopelágicos (los 'activos' o 'nictoepipelágicos en Nesis 1983) realizan migraciones nictemerales y se distribuyen generalmente en la capa mesopelagica (entre 200 m y 1000 m) durante el día y migran hacia la superficie durante la noche, patrón de migración similar al de D. gigas (Sielfeld et al. 1995, Moku et al. 2001, 2002, Uchikawa et al. 2002, Yasuma et al. 2006). Es el caso de las principales presas mesopelágicas encontradas en los estómagos del calamar gigante: Vinciguerria lucetia y los mictófidos Myctophum nitidulum, Myctophum aurolaternatum y Diogenichthys laternatus. Estimaciones acústicas de la biomasa muestran que Vinciguerria lucetia es un organismo muy abundante e lo largo de las costas del Perú y que los movimientos migratorios de los calamares gigantes parece relacionados a su distribución (Rosas et al. 2011).

Rosas (2007), señala que la gran variabilidad de presas consumidas apoyaría la hipótesis de la no selectividad en la alimentación del calamar gigante propuesta por varios autores (Klett 1981, Clarke y Paliza 2000, Nigmatullin *et al.* 2001). Sin embargo, la presencia de una gran mayoría de presas de tipo micronectónicas mesopelágicas estaría sugiriendo una preferencia hacia este tipo de presas (Markaida y Sosa 2003, Markaida 2006). Ménard *et al.* (2006) postulan por lo general que, en los ecosistemas de mar abierto, los grandes depredadores de peces adoptan una conducta de alimentación oportunista dentro de su rango de alimentación.

Al igual que en Markaida (2001), en este estudio se halló un bajo porcentaje de estómagos vacíos (16%). Al contrario, otros estudios reportan un gran porcentaje de

estómagos vacíos (Wilhelm 1954, en Clarke y Paliza 2000, Baral 1967, Ehrhardt *et al.* 1983 a, b 1986, Benites 1985, Segura *et al.* 1996, Ganoza *et al.* 1997, Mariátegui *et al.* 1998b). Algunos autores lo atribuyen a una baja disponibilidad de presas (Benites 1985, Segura *et al.* 1996, Ganoza *et al.* 1997). Por el contrario, Baral (1967), Ehrhardt *et al.* (1983a, b, 1986), Marcial-Ramos (1996) y Mariátegui *et al.* (1998b) atribuyen el alto porcentaje de estómagos vacíos principalmente a una alta tasa de digestión más que a la falta de alimento.

4.2.2 VARIABILIDAD DE LA DIETA RESPECTO A LA TALLA

La dieta de calamar gigante en términos de peso varía según el tamaño (Figura 10). El patrón principal fue el aumento constante de canibalismo con el tamaño de *D. gigas* que representa el 8,3% de la dieta de calamares pequeños (LM <30 cm) hasta llegar a 61,9% en calamares con LM> 90 cm (Figura 10c). Por el contrario, la importancia de otros cefalópodos, teleosteii, mictófidos y pterópodos disminuyó con la LM del calamar gigante, mientras que la importancia de la *V. lucetia* aumentó hasta alcanzar el 31,0% en individuos de 51-60 cm de LM, para disminuir luego en las tallas grandes (3,1%) (Figura 5c). Los euphausiidae no presentan una tendencia clara; alcanzaron un alto valor (hasta 21,2%) en los calamares gigantes pequeños y grandes (Figura 10c). El índice de diversidad de Margalef también vario según el tamaño de los calamares gigantes. Este índice creció de pequeños calamares (0,68 por LM <30 cm) para alcanzar 1,08 para calamares de 41-50 cm y disminuir luego a 0,45 en los individuos más grandes (LM> 90 cm) (Figura 10a). Por último, la llenura del estómago presenta la tendencia clásica de disminución con el aumento de LM (Figura 10b).

Schetinnikov (1989), estudió el contenido estomacal de 280 calamares de entre 2 y 42 cm LM en aguas peruanas y demostró que el cambio de alimentación a nivel de tamaño es común; también, se ha inferido que el canibalismo del calamar gigante se incrementa con relación al tamaño corporal de la especie (Nigmatullin *et al.* 2001, Markaida 2001, Markaida y Sosa-Nishizaki 2003). En cuanto al grado de llenura estomacal, Nesis (1970, 1983) indica un aumento en los índices de llenado con la talla del calamar (desde 1.6-2.0 a los 20-39 cm hasta los 2.5-3.0 a los 45-59 cm LM). Sin embargo, en este estudio se reportan valores altos en calamares pequeños y una progresiva disminución con la talla lo que explica la alta tasa metabólica en las primeras etapas de la vida del calamar.

Rosas (2007) encontró en Baja California una división entre dos grandes grupos de tallas: los individuos con talla hasta los 50 cm de LM se alimentaron de un alto número de presas como el pez *V. lucetia*, la langostilla *Pleuroncodes planipes*, y los calamares *P. giardi* y *H. heteropsis*, mientras que los calamares de más de 50 cm LM se alimentaron casi solamente de *P. planipes* y *H. heteropsis*. En este estudio el cambio de dieta por tamaños fue progresivo, sin embargo, se presentaron dos grupos de tallas: los calamares menores de 60 cm consumieron en su mayoría peces mesopelágicos y cefalópodos indeterminados, y los calamares mayores de 60 cm consumieron mayormente cefalópodos.

El cambio ontogenético de una dieta basada en invertebrados meso y macroplanctónicos (crustáceos) a peces y calamares micronectónicos ha sido documentado para el calamar gigante (Nesis 1970 y 1983, Shchetinnikov 1989), así como para otros omastréfidos oceánicos (Amaratunga 1983, Rodhouse y Nigmatullin 1996) y Sthenoteuthis oualaniensis (Shchetinnikov 1992). Otros autores también reportan en otras latitudes que el canibalismo se acentúa en los calamares más grandes (Baral 1967, Fitch 1968, Nesis 1970, Fitch 1976, Nesis 1983, Ehrhardt et al. 1983b y 1986, Shchetinnikov 1989, Marcial-Ramos 1996, Markaida 2006). En el presente estudio se observo alto canibalismo en tallas grandes (mayores de 60 cm) pero también, alto consumo de eufáusidos. Ayón et al. (2008) postulan que Euphausia mucronata (o 'krill'; Escribano et al. 2000, Antezana 2006) es uno de los macrocrustáceos pelágicos característicos en el Sistema de la Corriente de Humboldt (HCS). Esta especie puede ser muy abundante en el Sistema de la Corriente de Humboldt (SCH) (Antezana 2002, Ballón et al. 2011). Esta abundancia podría estar relacionada con un uso más eficiente de la producción primaria, y/o una fuerte conexión entre las zonas costeras y los ecosistemas pelágicos en alta mar (Espinoza y Bertrand 2008, Ballón et al. 2011). Además, estos crustáceos son la principal fuente de alimento para adultos de anchoveta (Engraulis ringens) (Espinoza y Bertrand 2008) y de muchos otros depredadores, tales como merluza (Mejía et al.1971, 1973, 1980). Esto indicaría que por las características del comportamiento del calamar gigante y la abundancia de esta presa, que se distribuye principalmente afuera de la plataforma continental (Ayón et al. 2008, Ballón et al. 2011), fácilmente los calamares podrían predar sobre ella. Estos resultados confirman que, a nivel individual, el calamar gigante adulto podía alimentarse principalmente de eufáusidos disminuyendo su nivel trófico Lorrain et al. (2011).



Figura 10. Número de especies, índices de diversidad (a), porcentaje de llenura estomacal (b) y composición de la presas (c) en peso (g) en la dieta del calamar gigante *Dosidicus gigas* respecto a grupos de talla (cm) del 2004 al 2009.

4.2.3 VARIABILIDAD TEMPORAL DE LA DIETA

La variación temporal en la dieta de calamar gigante también ilustra la importancia del canibalismo en algunas estaciones del año pero sin seguir un patrón claro, siendo dominante en el 2004 y 2007. Otras presas importantes en algunos periodos fueron los eufáusidos (veranos 2004, 2006 y 2008; otoño 2005 y 2008 e invierno 2004), *Cubiceps* sp. (veranos del 2005, 2007 y 2008); y *V. lucetia* (primaveras del 2006 y 2008) (Figura 11c). La llenura del estómago fluctuó fuertemente entre 0,08% en el verano de 2009 a 0,53% en la primavera de 2007 (Figura 11b). El índice de diversidad de Margalef (Figura 11a) no presento una tendencia clara. Los valores más altos del índice de diversidad de Margalef se presentaron en verano de 2004 y el invierno de 2006 (1,63 y 2,4, respectivamente) (Figura 11a).

Rosas (2007) en la costa occidental de la Península de Baja California indicó que los peces V. lucetia y T. mexicanus dominaron en la dieta de D. gigas durante los meses de marzo-abril y octubre. En nuestro estudio se aprecia mayor contribución de V. lucetia en la dieta del calamar gigante durante las primaveras del 2006 y 2008.

El *Cubiceps* sp. es uno de los peces más abundantes de los niveles tróficos intermedios y es un componente importante de las redes tróficas marinas en los ecosistemas pelágicos en mar abierto (Potier et al. 2008). Se concentra grandes en zonas de alta productividad y su abundancia tiene una fuerte estacionalidad en la transferencia de energía a niveles tróficos más altos (Potier et al. 2008). En SCH se ha detectado la presencia de esta especie en la dieta a lo largo del tiempo; sin embargo, es durante el verano que tuvo mayor relevancia en la dieta. En el Océano Índico se ha observado entre diciembre y abril asociado al Monzón Nor-oriental (Potier et al. 2008).



Figura 11. Número de especies, índices de diversidad (a), porcentaje de llenura estomacal (b) y composición de la presas (c) en peso (g) en la dieta del calamar gigante *Dosidicus gigas* respecto a las estaciones y el año del 2004 al 2009.

Los resultados indican que el calamar gigante se alimenta durante todos los períodos del día (Figura 12). Sin embargo, la llenura fue más alta durante la noche y más baja en la madrugada (Figura 12b). La composición de la dieta también varío según el ciclo diario. Por la noche, D. gigas fue la presa principal seguido por los peces mesopelágicos y cefalópodos. Al amanecer, los peces mesopelágicos, en particular Vincigerria lucetia (51,2%) y mictófidos (17,5%) dominaron fuertemente, y D. gigas representó un 7.8% solamente. Durante el día la dieta estuvo más equilibrada entre los taxones, las presa más importantes fueron cefalópodos (22.1%), D. gigas (21.9%), V. lucetia (18.6%), euphausiidae (11.4%) y teleósteos (8.9%). Finalmente, al atardecer, los cefalópodos fueron las presas dominantes (41.6%), seguido de D. gigas (22.9%) y peces mesopelágicos (16.6%) (Figura 12c). El índice de diversidad Margalef (diversidad: 1.4) fue más alto en la madrugada e inferior en el día (0.5). El número de especies fue mayor durante los crepúsculos (Figura 7a). Después de haber comido durante la noche, la llenura estomacal del calamar gigante en la madrugada presentó mayor incidencia en peso (0.2 g/ind) (Figura 12b). El calamar gigante es capaz de cazar en la ZMO (Nigmatulin et al. 2001) por lo tanto, puede comer mesopelágicos a cualquier hora, de día cuando están en la ZMO, de noche cuando están en superficie y durante los crepúsculos cuando migran. Es durante la madrugada que se observó mayor incidencia de Vinciguerria lucetia y mictófidos, esto coincidiría con la accesibilidad de estas especies durante la noche en la superficie y durante su migración a la ZMO en la madrugada.

Al igual que en este estudio, otros investigadores han indicado máximos de alimentación durante la noche (Nesis 1983, Bazanov 1986, Clarke y Paliza 2000, Nigmatullin *et al.* 2001). También se ha indicado, que esta especie se alimenta a cualquier hora, entre la superficie y los 500 m o más (Fitch 1976). En otras especies, la mayor ingesta de alimento ha sido registrada entre las 14:00 y las 20:00 horas para *I. argentinus* (Ivanovic 2000) y durante la noche-amanecer para *Nototodarus gouldi* (O'Sullivan y Cullen 1983).

D. gigas migra dentro de la zona mínima de oxígeno (ZMO) durante el día, donde se buscan los alimentos que les dan más energía para satisfacer sus necesidades metabólicas altas por ser una especie de ciclo de vida corto (Nesis 1970 y 1983). Esta asociación ha sido reportada en el Golfo de California, donde el calamar se mantuvo en gran medida en la zona hipóxica (ZMO) (Salinas *et al.* 2010). Yatsu *et al.* (1999) colocaron

transmisores ultrasónicos en calamares gigantes en el Domo de Costa Rica y frente a Perú, los calamares nadaron sobre los 200 m de profundidad durante el día y bajaron al menos hasta los 1000 m al amanecer, no siendo afectados por la capa de oxígeno mínimo, entre los 200 y 800 m, durante el descenso.

Según Rosas *et al.* (2011) las relaciones entre *D. gigas* y *V. lucetia* registradas en señales acústicas en el Perú, son evidentes durante el día y la noche y *D. gigas* probablemente se alimenta de *V. lucetia* en aguas más profundas durante el día (Rosas *et al.* 2011). Alarcón *et al.* (2004) llevó a cabo pesca experimental con redes de arrastre semipelágicas para *V. lucetia* y registró grandes capturas de calamar gigante. Rosas *et al.* (2011) estimaron la biomasa en las señales acústicas de *D. gigas* y *V. lucetia* en la columna de agua, sugiriendo una estrecha relación entre presa y depredador cerca del límite de la plataforma continental. En este mismo estudio se observó que durante la noche, *D. gigas* estuvo cerca de la superficie del agua rodeado de *V. lucetia* hasta las primeras horas del día en que ambos *V. lucetia* y *D. gigas* se migraron verticalmente, compartiendo la misma capa de agua durante el día, regresando a la superficie del agua en las primeras horas de la noche. Esta relación trófica se reflejó en el análisis del contenido estomacal de los calamares capturados en las primeras horas de la mañana, en el que *V. lucetia* fue el principal componente de la dieta de calamar gigante.



Figura 12. Número de especies, índices de diversidad (a), porcentaje de llenura estomacal (b) y composición de la presas (c) en peso (g) en la dieta del calamar gigante *Dosidicus gigas* respecto al periodo horario del 2004 al 2009.

4.2.4 VARIACIÓN ESPACIAL DE LA DIETA

D. gigas fue la presa principal en todas partes excepto en los individuos recolectados lejos de la costa a más de 240 mn, donde *V. lucetia* fue la presa dominante. Los eufáusidos tuvieron mayor contribución a una distancia menor de 120 nm (Figura 13c). La llenura estomacal aumentó con la distancia a la costa (Figura 13b), así como el número de especies de presa, mientras que los índices de diversidad no mostraron una tendencia clara (Figura 13a). Schetinnikov (1986a) destaca el incremento del canibalismo conforme esta especie se acercaba a la costa y en aguas oceánicas se destacó la importancia del mictófido *M. nitidulum*.





Ballón *et al.* (2011) enseñaron que la biomasa en eufáusidos es baja en la plataforma continental y se incrementa drásticamente a partir del talud donde pueden aprovechar de la productividad generada por el afloramiento costero y empieza a disminuir a unos 150 km de la costa. En este estudio todas las muestras fueron colectadas en la zona de mayor concentración del recurso afuera de la plataforma continental, sin embargo la mayor importancia de esta presa se observó por debajo de las 120 mn.

Latitudinalmente no surgieron fuertes patrones (Figura 14), excepto la gran importancia de *D. gigas* en la zona norte. *V. lucetia* fue más importante en el centro y los teleosteii en el sur de Perú. El índice de Margalef no presentó variación latitudinal.

Según Schetinnikov (1986a) mientras esta especie se desplaza más hacia el sur entre Perú y Chile, la importancia de los mictófidos disminuye. Bazanov (1986) se refirió a que la actividad alimenticia decrece cuando la abundancia de peces mictófidos disminuye y es en este momento cuando los calamares inician el canibalismo. En este estudio el canibalismo fue más representativo en la zona norte y los mesopelágicos en la zona central del Perú.



Figura 14. Número de especies, índices de diversidad (a), porcentaje de llenura estomacal (b) y composición de la presas (c) en peso (g) en la dieta del calamar gigante *Dosidicus gigas* respecto a la latitud (°S) del 2004 al 2009.

4.3 ANÁLISIS EXPLORATORIO PARA DETERMINAR LA NO LINEALIDAD DEL MODELO

4.3.1 RELACIONES GENERALES ENTRE VARIABLES

En la Figura 15 se aprecia la no linealidad de la relación de las variables respuesta: llenura estomacal (*fullness*) y diversidad de Margalef con las predictoras: talla, hora, tiempo, distancia a la costa, latitud y temperatura superficial del mar. Estas distribuciones se asemejan a una distribución gamma. La variable número de especies no presenta linealidad, más bien se observa una tendencia errática. Las variables predictoras talla y temperatura presentan una distribución normal mientras el tiempo presenta una estacionalidad. La periodicidad presenta asimetría negativa al igual que la latitud y, finalmente, la distancia a la costa presenta asimetría positiva.



Figura 15. *Scaterplot* entre las variables llenura estomacal (*fullness*), número de especies (*species*), índice de diversidad de Margalef (*Margalef*), talla (*size*), estacionalidad (*seasyear*), periodicidad horaria (*hour*), distancia a la costa (*distcoast*), latitud (*lat*) y temperatura superficial del mar (*SST*) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.

4.3.2 GRÁFICOS DE PROBABILIDAD NORMAL DE RESIDUOS Y PRUEBAS DE NORMALIDAD

En el modelo general de regresión lineal con variable respuesta porcentaje de llenura estomacal (Figura 16) se observa no linealidad de la variable respuesta con relación a las variables predictoras. La forma de embudo en el gráfico de residuales respecto a los valores ajustados (Figura 16a) indica que no se cumple el supuesto de homocedasticidad de las varianzas.

El contraste de Breusch-Pagan es significativo con un nivel $\alpha = 0.05$ (Chi cuadrado = 957.42, *p_valor* < 0.0001) rechaza la hipótesis nula que la varianza del error del modelo es constante, aceptando la hipótesis alternativa que hay cambios de la varianza de error con el nivel de la respuesta (valores ajustados), o con una combinación lineal de los predictores.

En el gráfico de probabilidad normal de los residuos se observa asimetría positiva. Los puntos 1819, 2191 y 3377 pueden ser considerados como valores atípicos 'outlier' o 'leverage' pero no influencian en el modelo. Las Figuras 16c y 16d de localización y escala (*lowess* de los residuales estandarizados en función de los valores ajustado) no muestran una distribución homogénea de los residuos, no se aproximan a una línea horizontal con pendiente cero, lo que indica la no linealidad de la variable respuesta del modelo.

La prueba de normalidad de Shapiro – Wilk es significativa (W = 0.39, p-valor < 2.2e-16) por lo que se rechaza la hipótesis nula que postula que el modelo presenta distribución normal.

Según lo descrito en el punto 4.1, gráficamente el comportamiento de la variable respuesta en relación a las variables predictoras no presenta linealidad, este modelo corresponderían a otro tipo de relación.



Figura 16. Análisis de normalidad y homocedasticidad de varianzas del modelo lineal: porcentaje de llenura estomacal del calamar gigante, respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y temperatura.

El modelo general con variable respuesta índice de diversidad de Margalef no cumple el supuesto de homocedasticidad de las varianzas observándose una ligera forma de embudo de los residuales (Figura 17a). En el gráfico de probabilidad normal de los residuales se observa la no linealidad del modelo. Los puntos 1261, 69 y 830 son valores outlier y leverage, lo mismo ocurre en las gráficas de localización y escala que no muestran una distribución homogénea de los residuales (Figuras 17b, c y d).

La prueba de normalidad de Shapiro – Wilk es significativa (W = 0.13, p-valor < 2.2e-16) por lo que se rechaza la hipótesis nula que el modelo tiene distribución normal.

La variable respuesta índice de diversidad de Margalef que compara la cantidad de especies presa en cada contenido estomacal no presenta distribución normal. En efecto, los residuos no presentan una distribución normal. En este caso, se tiene una dieta conformada por un número reducido de presas en cada estómago. De hecho, muchas veces se encontró sólo una presa en el contenido estomacal y en otros casos los estómagos se encontraron vacíos, lo que podría explicar este tipo de distribución. El calamar es una especie altamente migratoria (Nesis 1970, 1983), esto lleva a suponer que en largos viajes no siempre tienen disponibilidad de alimento, ejerciendo muchas veces canibalismo.



Figura 17. Análisis de normalidad y homocedasticidad de varianzas de la variable respuesta del modelo: índice de diversidad de Margalef respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y temperatura.

4.4 MODELO ADITIVO GENERALIZADO (GAM)

4.4.1 MÉTODOS DE SELECCIÓN DE VARIABLES

Se usaron los métodos AIC y la desvianza explicada del modelo GAM, para ver que variables tienen mayor influencia en el comportamiento de la variable respuesta del modelo (llenura estomacal). Para seleccionar las variables predictoras que mejor explican el modelo se usó el método *Forward*. Según se aprecia en el cuadro 3, la covariable que explica la mayoría de la varianza es *talla* con una desvianza explicada de 8% y un valor AIC de 1673.37. Las demás variables tienen una contribución menor y homogénea (desvianza explicada entre 2.21% y 3.89%). El calamar gigante es una especie de ciclo de vida corto y crecimiento rápido (Nigmatullin *et al.* 2001, Arkhipkin y Murzov 1986, Masuda *et al.* 1998, Arguelles *et al.* 2001, Markaida *et al.* 2004), observándose una alta tasa metabólica durante los primeros estadíos de vida (Shchetinnikov 1989), esta característica podría influir en la diferencia de la llenura estomacal durante su ciclo de vida.

Cuadro 3. Influencia de las variables predictoras en el comportamiento de la variable respuesta llenura estomacal, usando las pruebas de bondad de ajuste AIC y desvianza explicada.

Modelo (contribución de cada variable)	Desvianza explicada (%)	AIC
$log(E[llenura]) = \beta_0 + f_1(Talla) + \varepsilon_i$	8.00	1673.37
$log(E[llenura]) = \beta_0 + f_1(Tiempo) + \varepsilon_i$	2.94	1966.84
$log(E[llenura]) = \beta_0 + f_1(Hora) + \varepsilon_i$	2.59	1564.91
$log(E[llenura]) = \beta_0 + f_1(Distancia) + \varepsilon_i$	3.89	1910.46
$log(E[llenura]) = \beta_0 + f_1(Latitud) + \varepsilon_i$	2.21	2007.41
$log(E[llenura]) = \beta_0 + f_1(TSM) + \varepsilon_i$	2.04	2015.35

El modelo que explica mejor la probabilidad de cambios en la llenura del estómago es el que está constituido por las 6 variables predictoras o covariables: talla, tiempo, hora, distancia a la costa, latitud y temperatura. Dicho modelo explica el 21.30% de la desvianza (Cuadro 4), siendo este porcentaje mayor que para los otros modelos y presenta el menor valor de AIC (741.58). Además, todas las relaciones fueron significativas (p_valor<0.00001). La desvianza explicada encontrada en este modelo es relativamente baja. Sin embargo, en el ambiente marino muchas variables influyen el comportamiento de una determinada variable respuesta, debido a la complejidad y al alto número de

interacciones que existen entre los distintos componentes del ecosistema. Por lo tanto, para poder medir las covariables necesarias para encontrar un ajuste estadísticamente aceptable, sería necesario mover una logística sumamente costosa, así que se opta por medir las variables que por experiencia sabemos nos brindaran una explicación ecológica adecuada para responder nuestras preguntas.

Cuadro 4. Selección de variables predictoras que más influyen en el modelo usando el método *Forward*.

Modelo (selección de variables por el método "Forward")	Desvianza explicada (%)	AIC
$log(E[llenura]) = \beta_0 + f_1(Talla) + \varepsilon_i$	8.00	1673.37
$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + \varepsilon_i$	10.90	1517.86
$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + \varepsilon_i$	14.70	1028.69
$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + \varepsilon_i$	16.70	941.87
$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + \varepsilon_i$	20.00	789.31
$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + f_6(TSM) + \varepsilon_5$	22.10	741.58

4.4.2 TIPO DE DISTRIBUCIÓN

Utilizando *saddlepoint approximation* en el paquete tweedie (Dunn 2011) del programa R 2.12.1, sugiere que $var[Y] = \varphi \mu^p$, donde *p* es mayor o igual a uno, o menor o igual a cero, los casos usados por esta función incluyen la distribución normal (*p* = 0), Poisson (*p* = 1), gamma (*p* = 2) e inversa gaussiana (*p* = 3). Usando el análisis gráfico histograma de frecuencias y densidad (Figuras 18a y b) se puede observar una distribución parecida a la Gamma; y utilizando el gráfico *saddlepoint approximation*, se puede apreciar un "*p*" máximo de 2.2 lo cual indica que la distribución de la variable respuesta llenura estomacal se aproxima a una distribución gamma (Figura 19).



Figura 18. Histograma de frecuencia y densidad de la variable respuesta llenura estomacal en el calamar gigante *Dosidicus gigas* del 2004 al 2009.



Figura 19. Gráfico *saddlepoint approximation*, se puede apreciar un "*p*" máximo de 2.2 de la variable respuesta llenura estomacal en el calamar gigante *Dosidicus gigas* del 2004 al 2009.

4.4.3 ÍNDICE DE LLENURA ESTOMACAL

La relación de la variable respuesta llenura estomacal y las variables predictoras talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante es no lineal (*edf* >1) y se ha determinado que existen relaciones significativas y positivas entre ellas. El modelo global sugerido sería:

$$log(E[llenura]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + f_6(TSM) + \varepsilon_{i}$$

La desvianza explicada del modelo propuesto fue de 21.30% y se puede concluir con un nivel de significación $\alpha = 0.05$, que los efectos de todas las variables predictoras fueron significativos al modelo (p-valor < 0.0001) para explicar el comportamiento de la variable respuesta llenura estomacal.



Figura 20. Modelo Aditivo Generalizado (GAM) relacionando la respuesta de la llenura estomacal con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.

La respuesta de la variable predictora talla fue la más importante en el modelo (desvianza explicada 8%). La talla presenta una relación negativa con respecto a la llenura del estómago lo que confirma (Figura 20a) que los calamares de tallas menores se alimentan más en relación a su peso que los de tallas grandes, debido a la alta tasa metabólica que presenta esta especie en las primeras etapas de vida (Shchetinnikov 1989) (*edf*=5.82, p_valor<0.001). Variación en el tiempo no presenta una tendencia clara

(*edf*=8.36, p_valor<0.001). Se observa menor probabilidad de llenura estomacal durante el verano del 2009 (Figura 20b).

La periodicidad horaria indica una disminución de la llenura durante la madrugada hasta las primeras horas de la mañana (edf=8.62, p_valor<0.001). La llenura aumenta a medida que los calamares se alejan de la costa (edf=2.19, p_valor<0.001). Latitudinalmente no se aprecia un patrón (edf=8.69, p_valor<0.001), la temperatura superficial del mar fue significativa para el modelo (edf=8.23, p_valor<0.001) pero no hubo una tendencia en el consumo de alimento de los calamares (Figura 20).

Según documentan algunos autores (Mangold 1983, Voss 1983, Boyle 1990, Hanlon y Messenger 1996, Rodhouse y Nigmatullin 1996, Nixon 1987) los cefalópodos son depredadores activos y rápidos que se alimentan de presas vivas: crustáceos, peces y otros cefalópodos, habiendo por lo general una sucesión en la alimentación (estratificación por tallas), en este estudio se destaca la influencia de la talla en la llenura estomacal distinguiéndose altos valores en individuos juveniles. Nesis (1970, 1983) reporta un aumento en los índices de llenura con la talla del calamar (desde 1.6-2.0 a los 20-39 cm hasta los 2.5-3.0 a los 45-59 cm LM). Sin embargo, la alta variabilidad del llenado y del peso del contenido estomacal registrado por Markaida (2001) no permitió percibir diferencias entre la llenura y la talla, madurez sexual y sexo.

La periodicidad horaria de alimentación presenta un débil aumento en la llenura durante la tarde y la noche decreciendo en la madrugada. Sin embargo, en la flota industrial, la mayor actividad de pesca se realiza durante la noche debido a la vulnerabilidad del recurso, puesto que el calamar gigante sube cerca a la superficie para alimentarse, debido a las migraciones nictimerales de sus presas, teniendo mayor probabilidad de encontrar abundante alimento. Yatsu *et al.* (1998) señala que las bajas o nulas capturas durante el día sugieren que el calamar se halla en profundidad, inaccesible al arte de pesca. Nigmatullin *et al.* (2001) señalan que el calamar permanece bajo la capa mínima de oxígeno durante el día y se alimenta de noche en superficie. Otros autores (Nesis, 1983, Clarke y Paliza 2000, Bazanov 1986, Shchetinnikov 1986 a y b, 1989). Fitch (1976), indicaron que el calamar gigante se alimenta a cualquier hora entre la superficie y los 500 m.
Fuera de la costa la llenura estomacal es mayor, esto sugeriría un mayor número de presas fuera de la costa. Nesis (1970, 1983) señala que las mayores concentraciones de calamar se hallan en la zona donde la productividad primaria es alta pero no máxima, la biomasa de zooplancton es relativamente alta y el número de peces meso y batipelágicos es máximo. Latitudinalmente no se observa un patrón en la llenura estomacal. Schetinnikov (1986b) analizó el contenido estomacal de 120 calamares colectados en Perú y Chile y determinó que entre más se desplaza hacia el sur, la importancia mictófido *Mycthopum nitidulum* disminuye incrementando la importancia del canibalismo.



Figura 21. Pruebas de bondad de ajuste y homocedasticidad de varianzas de la variable respuesta del modelo: llenura estomacal respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y temperatura.

La temperatura tuvo un efecto altamente significativo en la llenura pero sin presentar un patrón claro. Eso se puede relacionar al hecho que el calamar gigante es una especie euritérmica, hallándose en aguas con un amplio rango de temperatura superficial, entre 15-28°C Nesis (1983). Si además se considera su rango vertical de distribución este rango es mucho mayor.

La variable respuesta llenura estomacal no presenta un buen ajuste de los residuales (Figura 21). Sin embargo, presentó un mejor ajuste que al aplicar un modelo lineal, esto se aprecia principalmente en el histograma y la gráfica de residuales ajustados.

En el cuadro 5 se observan los coeficientes por cada variable predictora, al aplicar el método de *splines*, se obtuvieron 9 tramos o nodos y en consecuencia 9 coeficientes, por cada tendencia observada en cada covariable respecto a la variable respuesta. En todos los casos se observaron tendencias negativas y positivas por cada tramo.

-	61	62	63	64	65	66	67	68	69
	s(size).1	s(size).2	s(size).3	s(size).4	s(size).5	s(size).6	s(size).7	s(size).8	s(size).9
Talla	0.105	-0.164	-0.536	-0.744	-0.819	-0.929	-0.935	-1.238	-0.974
	s(seasyear).1	s(seasyear).2	s(seasyear).3	s(seasyear).4	s(seasyear).5	s(seasyear).6	s(seasyear).7	s(seasyear).8	s(seasyear).9
Tiempo	-0.260	0.080	-0.213	-0.124	0.463	0.326	-0.396	-1.319	-0.041
	s(hour2).1	s(hour2).2	s(hour2).3	s(hour2).4	s(hour2).5	s(hour2).6	s(hour2).7	s(hour2).8	s(hour2).9
Hora	0.264	0.284	0.149	0.255	0.290	0.631	-0.183	0.558	-0.997
	s(distcoast).1	s(distcoast).2	s(distcoast).3	s(distcoast).4	s(distcoast).5	s(distcoast).6	s(distcoast).7	s(distcoast).8	s(distcoast).9
Distancia	-0.165	-0.012	0.056	0.112	0.144	0.265	0.040	0.440	0.426
	s(lat).1	s(lat).2	s(lat).3	s(lat).4	s(lat).5	s(lat).6	s(lat).7	s(lat).8	s(lat).9
Latitud	0.314	0.128	0.537	1.024	0.543	0.473	1.368	0.601	0.372
	s(SST).1	s(SST).2	s(SST).3	s(SST).4	s(SST).5	s(SST).6	s(SST).7	s(SST).8	s(SST).9
Temperatura	-0.020	-0.295	-0.143	0.252	-0.014	-0.155	0.314	0.256	-0.117

Cuadro 5. Coeficientes del modelo por cada tramo de splines.

4.4.4 ÍNDICE DE DIVERSIDAD DE MARGALEF

La relación de la variable respuesta diversidad de Margalef y las variables predictoras talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante es no lineal (*edf* >1) y se ha determinado que existen relaciones significativas y positivas entre ellas menos con la variable predictora temperatura (ns). El modelo global sugerido sería:

$$log(E[diversidad]) = \beta_0 + f_1(Talla) + f_2(Tiempo) + f_3(Hora) + f_4(Distancia) + f_5(Latitud) + f_6(TSM) + \varepsilon_0(TSM) +$$

La desvianza explicada del modelo propuesto fue de 16.5% y se puede concluir con un nivel de significación $\alpha = 0.05$, que los efectos las variables predictoras talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) fueron significativos al modelo (p-valor < 0.0001) para explicar el comportamiento de la variable respuesta diversidad de Margalef.



Figura 22. Modelo Aditivo Generalizado (GAM) relacionando la respuesta de la índice de diversidad de Margalef con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.

La diversidad, al igual que la llenura fue mayor en individuos pequeños, tendiendo a decrecer a medida que los individuos crecieron, esta tendencia es opuesta a lo observado por Rosas (2007), quien observó un bajo número de presas en individuos más pequeños y no se observó el aumento del canibalismo en individuos grandes. En la Figura 21a, se aprecia el claro aumento del canibalismo en relación a la talla, viéndose una distribución más homogénea y mayor número de las especies en las primeras etapas de vida, además se aprecian varios grupos en la dieta de individuos juveniles, como crustáceos, peces mesopelágicos, pterópodos y otros cefalópodos lo que podrían estar aumentando la diversidad (*edf*=7.45, p_valor<0.001).



Figura 23. Pruebas de bondad de ajuste y homocedasticidad de varianzas de la variable respuesta del modelo: llenura estomacal respecto a las variables predictoras talla, estacionalidad, periodicidad horaria, distancia a la costa, latitud y temperatura.

El cambio ontogenético de una dieta con mayor diversidad en individuos pequeños ha sido argumentado para el calamar gigante por algunos autores (Fitch 1976, Nesis 1970, 1983, Shchetinnikov 1989). En el tiempo (Figura 21b), hay una disminución importante de la diversidad durante el año 2008, sin embargo, no se presenta una tendencia estacional clara (edf=8.54, p_valor<0.001). La mayor diversidad de la dieta se presentó en las primeras horas de la mañana, puede que el calamar se halla estado alimentando durante la madrugada (edf=8.51, p_valor<0.001) de varias especies peces mesopelágicos que suelen presentar migraciones verticales desde la Zona Mínima de Oxígeno (ZMO) (Nesis 1970, 1983). Fuera de las 100 mn la diversidad es acentuadamente mayor que en la zona costera (Figura 21d), tendiendo a aumentar fuera de la costa (edf=7.64, p_valor<0.001). Latitudinalmente, se observan cambios en la zona norte y sur, lo cual debe ser tomado con cuidado pues los límites de confianza se alejan por la falta de ajuste debido a la menor cantidad de datos (edf=8.90, p_valor<0.001). La temperatura no presentó relación alguna con la diversidad (Figura 22).

La variable respuesta índice de diversidad de Margalef, no muestra un buen ajuste de los residuales (Figura 23). Sin embargo, al igual que en el caso de la llenura estomacal, presentó un mejor ajuste que al aplicar un modelo lineal, esto se aprecia principalmente en el histograma.

En el cuadro 6 se observan los coeficientes por cada variable predictora, al aplicar el método de *splines*, se obtuvieron 9 tramos o nodos y en consecuencia 9 coeficientes, al igual que en el caso de la variable respuesta llenura estomacal. Asimismo, en todos los casos se observaron tendencias negativas y positivas por cada tramo.

	61	62	63	64	65	66	67	68	69
	s(size).1	s(size).2	s(size).3	s(size).4	s(size).5	s(size).6	s(size).7	s(size).8	s(size).9
Talla	0.483	0.181	0.059	-0.031	-0.108	-0.292	-0.460	-0.378	-1.062
	s(seasyear).1	s(seasyear).2	s(seasyear).3	s(seasyear).4	s(seasyear).5	s(seasyear).6	s(seasyear).7	s(seasyear).8	s(seasyear).9
Tiempo	-0.115	-0.091	-0.151	0.314	-1.520	-0.694	-0.275	-0.224	-0.116
	s(hour2).1	s(hour2).2	s(hour2).3	s(hour2).4	s(hour2).5	s(hour2).6	s(hour2).7	s(hour2).8	s(hour2).9
Hora	0.269	0.171	-0.352	-0.085	-0.211	-0.068	0.148	-0.034	0.197
	s(distcoast).1	s(distcoast).2	s(distcoast).3	s(distcoast).4	s(distcoast).5	s(distcoast).6	s(distcoast).7	s(distcoast).8	s(distcoast).9
Distancia	0.476	0.254	0.552	0.442	0.294	0.083	-0.320	0.516	0.222
	s(lat).1	s(lat).2	s(lat).3	s(lat).4	s(lat).5	s(lat).6	s(lat).7	s(lat).8	s(lat).9
Latitud	-0.899	-0.568	-0.724	-0.758	-0.518	-1.109	-0.588	-0.003	0.595
	s(SST).1	s(SST).2	s(SST).3	s(SST).4	s(SST).5	s(SST).6	s(SST).7	s(SST).8	s(SST).9
Temperatura	a -0.102	0.041	-0.198	-0.042	-0.408	0.231	-0.328	-0.172	0.433

Cuadro 6. Coeficientes del modelo por cada tramo de splines.

V. CONCLUSIONES

- Las variables respuestas llenura del estómago e índice de diversidad de Margalef en relación a las variables predictoras no pertenecen a la familia de distribución normal. El modelo se aproxima a una distribución gamma con un valor de *p*=2.2.
- La variable que más influye en la llenura del estómago del calamar gigante fue la talla (8.00%), las demás variables tuvieron una influencia débil y homogénea (entre 2.04% y 3.89%).
- La dieta del calamar gigante cambia con relación a la talla, se refleja mayor consumo en los individuos juveniles debido a su alta tasa metabólica, disminuyendo progresivamente en los adultos.
- Se ha observado mayor diversidad en los calamares más pequeños y una disminución a medida que aumentó la talla.
- Hay mayor influencia del canibalismo en la dieta en calamares mayores de 60 cm, siendo los peces mesopelágicos más importantes en individuos más pequeños.
- Se reflejan cambios significativos respecto al tiempo, pero se probó que no se puede definir patrón estacional en los modelos GAM.
- La periodicidad de alimentación presentó cambios significativos, reflejándose un aumento del consumo durante la noche y decreciendo en la madrugada.
- La llenura del estómago aumenta significativamente a medida que los individuos se alejan de la costa, esto también se ve reflejado en la diversidad y en el aumento en el consumo de mesopelágicos a partir de las 240 mn.
- Respecto a la latitud, se observó un patrón de llenura estomacal a lo largo de la costa aumentando al norte con respecto a la variable respuesta llenura estomacal y disminuyendo en la zona central respecto a la variable respuesta índice de diversidad de Margalef.
- Respecto a la temperatura no se observaron tendencias, debido a que el calamar gigante es una especie euritérmica, hallándose en aguas con un amplio rango de temperatura.

- Finalmente Se ha determinado que existen relaciones significativas y positivas entre la llenura estomacal y la diversidad de las presas y las variables explicativas talla (cm), tiempo (estacionalidad), periodicidad horaria, distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C).

VI. RECOMENDACIONES

- Se recomienda complementar la información obtenida de la pesca industrial con datos provenientes de la flota artesanal que opera cerca a la costa, dentro de la plataforma continental. Asimismo, datos de cruceros dirigidos especialmente al recurso (*Dosidicus gigas*) que obtienen información fuera de la Zona Económica Exclusiva (ZEE) y durante las 24 horas del día. De esta manera se podrá obtener estómagos en todo el espectro espacial, obtener un mejor ajuste estadístico y una visión más completa de la influencia de las variables predictoras sobre el modelo.

- Se recomienda analizar como variables respuesta, la llenura estomacal de individualmente en las principales especies de la dieta, pues los patrones encontrados en la conducta alimenticia del calamar gigante podrían ser el reflejo de la disponibilidad de estas especies en el medio.

- Por último se recomienda aumentar el número de variables predictoras (covariables) como salinidad, corrientes, nutrientes, etc. para obtener un mejor ajuste del modelo.

VII. REFERENCIAS BIBLIOGRÁFICAS

Akaike, Hirotogu, 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. Petrov and F. Csaki, *eds*. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, pp. 267–281.

Alamo, V; Valdivieso; V. 1987. Lista sistemática de moluscos marinos del Perú. *Bol. Inst.Mar del Perú Vol. Extraordinario*: 205 p.

Alarcón, J.; Salazar, C.; Ganoza, F.; Chacon, G.; Calderon, J.; Takhia, M.; Wiji, O.; Masatsugu, Y. 2004. Efectos del incremento del tamaño de malla del copo de la red semipelágica en la captura de *Vinciguerria lucetia*. Reunión Científica Internacional. Instituto del Mar del Peru-Japan Deep Sea Trawler Association. Perú.

Amaratunga, T. 1983. The role of cephalopods in the marine ecosystem. *In*:Caddy, J. (*Ed*.) Advances in assessment of world cephalopod resources. FAO *Fisheries Technical Paper* 231: 379-415.

Antezana, T., 2002. Adaptative behaviour of Euphausia mucronata in relation to the oxygen minimum layer of the Humboldt Current. In: Färber, Jaime. (Ed.), Oceanography of the Eastern Pacific II. Centro de Investigación Científica y de Educación Superior de Ensenada CICESE, México, pp. 29–40.

Antezana, T., 2006. Euphasia mucronata a keystone and prey of the Humboldt Current. In: Book of Extended Abstracts, International Conference the Humboldt Current System: Climate, Ocean Dynamics, Ecosystem Processes, and Fisheries, Lima, Peru, November 27–December 1, 2006, pp. 51–53.

Argüelles, J; Rodhouse, P; Villegas, P; Castillo, G. 2001. Age, growth and population structure of the jumbo flying squid *Dosidicus gigas* in Peruvian waters. *Fish. Res.* 54, 51-61.

Arkhipkin, A; Murzov, S. 1986. Age and growth patterns *of Dosidicus gigas* (Ommastrephidae). *In*: Ivanov, B. (*Ed.*), Present State of Fishery for Squids and Prospects of its Development. *VNIRO Press*, Moscow, p:107-123.

Ayón, P., Criales-Hernandez, M.I., Schwamborn, R., Hirche, H.J., 2008a. Zooplankton research off Peru: a review. Progress in Oceanography 79, 238–255.

Ballon, M. (2010). Étude acoustique du macrozooplancton au Pérou: estimation de biomasse, distribution spatiale, impact du forçage physique, et conséquences sur la distribution des poissons fourrage. These pour obtenir le grade de Docteur de L'Universite Montpellier II. Discipline: Biologie des Populations et Écologie. Ecole Doctorale : Systèmes intégrés en Biologie, Agronomie, Géosciences, Hydrosciences et Environnement (SIBAGHE). 205 pp.

Ballón M., Bertrand A., Lebourges-Dhaussy A., Gutiérrez M., Ayón P., Grados D., Gerlotto F. 2011. Is there enough zooplankton to feed forage fish population off Peru? An acoustic (positive) answer. Progress in Oceanography, 00:000-000.

Bakun, A., & Weeks, S. J. (2008). The marine ecosystem off Peru: What are the secrets of its fishery productivity and what might its future hold? Progress In Oceanography, 79(2-4), 290-299. Elsevier Ltd. doi: 10.1016/j.pocean.2008.10.027.

Bhattacharya, P; Zao, P. 1997. Semiparametric inference in a partial linear model. Annals of Statistics. 25:244-262.

Baral, A. 1967. Some data on the biology of Southeastern Pacific squid. Ryb. Khoz., 43(8): 15-17. [En ruso]

Bazanov, S. I. 1986. Feeding behavior of *Dosidicus gigas* and some quantitative aspects of its feeding. pp. 124-131 En: Resursy i perspektivy ispol'zovaniia kal'marov mirovogo okeana: sbornik nauchnykh trudov / [redaktsionnaia kollegiia: B. G. Ivanov (otv. red.) (i dr.)]. Moskva: [VNIRO], 1986. Series title: Sbornik nauchnykh trudov (Vsesoiuznyi nauchno-issledovatel'skii institut morskogo rybnogo khoziaistva i okeanografii (Soviet Union) ocean. pp. 124-131. Moscú, VNIRO Press [En ruso, resumen en inglés]

Berg, J. 1979. Discussion of methods of investigating the food of fishes, with reference to a preliminary study of the prey of *Gobiusculus flavescens* (Gobiidae). Mar. Biol. 50: 263-273.

Benites, C. 1985. Resultado de las investigaciones biológico-pesqueras de la jibia *Dosidicus gigas* (d'Orbigny, 1835) en el litoral peruano de julio de 1983 a marzo 1984. En:

A. E. Tresierra (ed.). Anales del Congreso Nacional de Biología Pesquera (28 jun.-01 jul., 1984). Trujillo – Perú. 10-15 p.

Bennet, B. A. 1978. Underwater observations of the squid Illex illecebrosus in the Newfoundland inshore waters. In N. Balch, T. Amaratunga y R. K. O'Dor (Eds.). Proceedings of the workshop on the squid *Illex illecebrosus*. Dalhousie Univ., Halifax, Nova Scotia, May 1978. Tech. Rep. Fish. Mar. Serv., Envir. Can., 833: 12.1-12.9.

Bertrand, A., Ballón, M., Chaigneau, A. 2010. Acoustic observation of living organisms reveals the oxygen minimum zone. PLoS ONE 5(4): e10330.

Berry, S. 1912. A review of the cephalopods of Western North America. *Bull. Bur. Fish. Wash.*, 30: 269-336.

Bidurin, C; André, L. 2001. Modelos semiparamétricos: uma revisão. Revista Brasileira de Estatística 62: 71-90.

Bowman, A; Azzalini, A. 1997. Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations. New York: Oxford University.

Boyle, P. R. 1990. Cephalopod biology in the fisheries context. Fish. Res. 8: 303-321.

Buja, A; Hastie, T; Tibshirani, R. 1989. Linear smoothers and additive models. Annals of Statistics 17: 453-510.

Chavez, F., Bertrand, A., Guevara-Carrasco, R., Soler, P., Csirke, J. (2008) The northern Humboldt Current System: Brief history, present status and a view towards the future. Progress in Oceanography 79, 95–1051.

Chirichigno, N. 1998. Clave para identificar los peces marinos del Perú, 2ª ed. Instituto del Mar de Perú. Publicación Especial. Callao, 496 p.

Chong, J; Oyarzún, C; Galleguillos, R; Tarifeño, E; Sepúlveda, R; Ibáñez, C. 2005. Parámetros biológico-pesqueros de la jibia, *Dosidicus gigas* (Orbigny, 1835) (Cephalopoda: Ommastrephidae). Gayana 69(2): 319-328.

Cianelli, L; Fauchald, P; Chan, K; Agostini, V; Dingsør, G. 2008. Spatial fisheries ecology: Recent progress and future prospects. *J. Mar. Sys.* 71:223-236.

Clarke, R; Paliza O. 2000. The Humboldt Current squid *Dosidicus gigas* (Orbigny, 1835). *Rev. Biol. Mar. y Ocean.* 35(1): 1-39.

Coleman L. R. y B. G. Nafpaktitis. 1972. Dorsadena yaquinae, a new genus and species of Myctophid fish from the eastern North Pacific Ocean. Contributions in Science. Los Angeles, California U. E. A. 225: 1-11.

Cook, R. D. and Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. Biometrika, 70, 1–10.

Cubillos, L; Ibáñez, C; González, C; Sepúlveda, A. 2004. Pesca de investigación: Pesca de jibia (*Dosidicus gigas*) con red de cerco entre la V y X regiones, año 2003. Informe Final. Instituto de Investigaciones Pesqueras, Talcahuano, Chile. 48 pp.

Cunha, M. 2000. Métodos Numéricos, 2ª ed. São Paulo: Unicamp.

d'Orbigny, A. 1835-1843. Voyage dans l'Amérique méridionale exécuté pendant les anneés 1826-1827, 1828, 1829, 1830, 1831, 1832, et 1833. Paris et Strassbourg, Tom V pt. 3: Mollusques, Cephalopoda: 1-64, pp. 1-4.

Dias, R. 2001. O uso de Spline em Regressão Não-Paramétrica. Disponible en www.ime.unicamp.br/~dias/me661.html.

Díaz, F; Berger, K; Barrenechea, M; Mendoza, V. 2009. Informe anual desenvolvimiento del comercio exterior pesquero en el Perú 2008. Comisión de Promoción del Perú para la Exportación y el Turismo: PromPerú. Lima. pp 4-111.

Dobson, A. 2002. An introduction to Generalized Linear Models. New York. Library of Congress Cataloging-in-Publication Data. Chapman & Hall/CRC. 221 pp.

Dunn, 2011. Package 'tweedie'. Disponible en http://www.r-project.org/package=tweedie

Ehrhardt, N; Jacquemin, P; García, F; González, G; López, J; Ortiz, J; Solís, A. 1983a. On the fishery and biology of the jumbo squid *Dosidicus gigas* in the Gulf of California, México. FAO Fisheries Technical Paper. 231: 306-340.

Ehrhardt, N; Jacquemin, P; García, F; González, G; López, J; Ortiz J; Solís, A. 1983b. On the fishery and biology of the giant squid *Dosidicus gigas* in the Gulf of California,

Mexico. *In* J. F. Caddy (Ed.). Advances in assessment of world cephalopod resources. FAO *Fisheries Technical Paper*. 231: 306-339.

Ehrhardt, N; Solís, A; Jacquemin, P; Ortiz, J; Ulloa, P; González, G; García, F. 1986. Análisis de la biología y condiciones del stock del calamar gigante *Dosidicus gigas* en el Golfo de California, México, durante 1980. Ciencia Pesquera, 5: 63-76.

Ehrhardt, N. 1991. Potencial impact of a seasonal migratory jumbo squid (*Dosidicus gigas*) stock on a Gulf of California sardine (*Sardinops sagax caerulea*) population. Bulletin of Marine Science 49 (1-2): 325-332.

Enríquez-Paredes, L. 1999. Estructura genética de las agregaciones de calamar gigante (*Dosidicus gigas*) en el Golfo de California, México. Tesis de Maestría, Fac. Ciencias Marinas, UABC, 85 p.

Escribano, R, Hidalgo, P., Krautz, C., in press. Zooplankton associated with the oxygen minimum zone system in the northern upwelling region of Chile during March 2000. Deep-Sea Research II.

Espinoza, P; Bertrand, A. 2008. Revisiting Peruvian anchovy (*Engraulis ringens*) trophodynamics provides a new vision of the Humboldt Current system. Progress in Oceanography 79:215-227.

Espinoza, P; Bertrand, A; Lingen, C van der; Garrido, S; Rojas de Mendiola, B. 2009. Diet of sardine (*Sardinops sagax*) in the northern Humboldt Current system and comparison with the diets of clupeoids in this and other eastern boundary upwelling systems. Progress in Oceanography 83: 242-250.

Eubank, R. 1988. Spline Smoothing and Nonparametric Regression. New York: Marcel Dekker.

Eubank, R. 1994. A simple smoothing spline. American Statistician 2, 103-106.

Fitch, J; Brownell, R. 1968. Fish Otoliths in Cetacean Stomachs and Their Importance in Interpreting Feeding Habits. Journal of Fisheries Research Board of Canada, 25(12):2561-2574.

Fitch J. 1969. Fossil lanterfish otholiths of California, whit notes on fossil Myctophidae of

North America. Contributions in Science. Los Angeles, California U. E. A. 173: 1-20.

Fox J. y Weisberg S. 2011. Package 'car'. Disponible en <u>http://www.r-</u> project.org/package=car

Fitch, J. 1976. Food habits of the jumbo squid *Dosidicus gigas*. CalCOFI Conference Abstracts for 1976, p. 21.

Ganoza, F., J. Rubio, O. Morón y M. Girón. 1997. Investigaciones sobre el calamar gigante o pota en primavera en el área de Puerto Pizarro a Paita. Crucero BIC SNP-1 9609-10. Inf. Prog. Inst. Mar Perú, 65: 23-53.

García-Godos, I. 2001. Patrones morfológicos del otolito sagitta de algunos peces óseos del mar peruano. Boletín Instituto del Mar del Perú 20 (1-2): 83 pp.

Gilly, W; Markaida, U; Baxter, C; Block, B; Boustany, A; Zeidberg, L; Reisenbichler, K; Robinson, B; Bazzino, G; Salinas, C. 2006. Vertical and horizontal migrations by squid *Dosidicus gigas* revealed by electronic tagging. *Mar. Ecol. Series*. 324:1-17.

Handscomb, D. 1966. Spline functions. In Methods of Numerical Approximation. Oxford: Pergamon Press.

Hanlon, R.; Messenger J. 1996. Cephalopod behaviour. *Cambridge Univ. Press*, Gran Bretaña. 232 p.

Härdle, W. 1990. Applied Nonparametric Regression. Cambridge University. Cambridge.

Hastie, T; Tibshirani, R. 1986. Generalized Aditive Models. *Statistical Science*. 1(3):297-318.

Hastie, T; Tibshirani, R. 1990. Generalized Additive Models. Chapman and Hall, London.

Hyslop, E. 1980. Stomach contents analysis-a review of methods and their applications. *Journal of Fish Biology*, 17 (4): 411-429.

Ivanovic, M. L. 2000. Alimentación y relaciones tróficas del calamar Illex argentinus en el ecosistema pesquero. Tesis para optar el título de Doctor en Ciencias Biológicas, Universidad Nacional de Mar del Plata, Argentina. 248 p.

72

Keyl, F. 2009. The Cephalopod Dosidicus gigas of the Humboldt Current System Under the Impact of Fishery and Environmental Variability. Universität Bremen

Keyl, F; Argüelles, J; Mariategui, L; Tafur, R; Wolff, M; Yamashiro, C. 2008. A hypothesis on range expansion and spatio-temporal shifts in size-at-maturity of jumbo squid (*Dosidicus gigas*) in the eastern pacific ocean. CalCOFI Rep., vol. 49, 119-128.

Keyl, F; Argüelles, J; Tafur, R. 2010. Interannual variability in size structure, age, and growth of jumbo squid (*Dosidicus gigas*) assessed by modal progression analysis. ICES Journal of Marine Science, doi:10.1093/icesjms/fsq167.

Klett, A. 1981. Estado actual de la pesquería del calamar gigante en el estado de Baja California Sur. Depto. Pesca México. Serie Cient., 21: 1-28.

Kuroiwa, M. 1998. Explorations of the Jumbo squid, *Dosidicus gigas*, Resources in the Southeastern Pacific Ocean with notes on the history of jigging surveys by the Japan Marine Fishery Resources Center. In Contributed papers to international symposium on large pelagic squid. Japan Marine Fishery Resources Research Center, Tokyo. P 89-105.

Lee, D. 1990. Cross-validation in semiparametric models: some Monte Carlo results. Journal of Statistical Computation and Simulation 37, 171-187.

Lima, L; André, C; Singer, J. 2001. Modelos aditivos generalizados: metodologia e prática. *Revista Brasileira de Estatística* 62, 37-69.

Lorrain, A.; Arguelles, J.; Alegre, A.; Bertrand, A.; Richard, P.; Cherel, Y. 2011. Sequential isotopic signature along gladius highlights contrasted individual foraging strategies of pelagic squids. PLoS ONE, 6(7) 1-6.

Mangold, K. 1983. Food, feeding and growth in cephalopods. Mem. Nat. Mus. Victoria, 44: 81-92.

Marcial-Ramos, R. W. 1996. Estructura poblacional, madurez sexual y alimentación de *Dosidicus gigas*, desembarcada por la pesca artesanal en Paita, septiembre de 1995 - agosto de 1996. Tesis de maestría. Univ. Nal. De Trujillo, Trujillo, Perú. 28 pp.

Margalef, R. 1974. Ecología. Ediciones Omega. Barcelona.

Mariátegui, L., G. Castillo, C. Ruiz, C. Pais, B. Díaz y O. Valladares. 1998b. Pesquería artesanal del calamar gigante (*Dosidicus gigas*) en el litoral norte, octubre a diciembre 1997. Inf. Prog. Inst. Mar Perú, 77: 27-38.

Markaida. U. 2001. Biología del calamar gigante Dosidicus gigas Orbigny, 1835 (cephalopoda: ommastrephidae) en el Golfo de California, México. Tesis para obtener el grado de Doctor en Ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada. División de Oceanología, Departamento de Ecología. 426 pp.

Markaida, U; Sosa-Nishizaki, O. 2003. Food and feeding habits of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, México. *J. Mar. Biol. Ass. U. K.* 83:4162/1-16.

Markaida, U; Quiñónez-Velázquez, C; Sosa-Nishizaki, O. 2004. Age, growth and maturation of jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) from the Gulf of California, Mexico. *Fish. Res.*, 66(1): 31-47.

Markaida, U; Rosenthal, J; Gilly, W. 2005. Tagging studies on the jumbo squid (*Dosidicus gigas*) in the Gulf of California, México. *Fish Bull*. 103: 219-226.

Markaida, U. 2006. Food and Feeding of jumbo squid Dosidicus gigas in the Gulf of California and adjacent waters alter the 1997-98 El Niño event. Fish. Res. 1-12.

Masuda, S; Yokawa, K; Yatsu, A; Kawahara, S. 1998. Growth and population structure of *Dosidicus gigas* in the Southeastern Pacific. In: Okutani, T. (*Ed.*). Large Pelagic Squid, 107-118 p.

McCullag, P; Nelder, J. 1983. Generalized Linear Models. Chapman and Hall. London.

Ménard F.; Labrune C; Shin Y.; Asine A.; Bard F. 2006. Opportunistic predation in tuna: a size-based approach. Mar Ecol Prog Ser. Vol. 323: 223–231

Méndez, M. 1981. Claves de identificación y distribución de los langostinos y camarones (Crustacea: Decapoda) del mar y ríos de la costa del Perú. Boletín Instituto del Mar del Perú 5:170 p.

Mendiburu, F. 2010. Package 'agricolae'. Disponible en <u>http://cran.r-</u> project.org/web/packages/agricolae/agricolae.pdf Merdsoy, B. 1978. In situ observations of squid. In N. Balch, T. Amaratunga y R. K. O'Dor (Eds.), Proceedings of the workshop on the squid *Illex illecebrosus*. Dalhousie Univ., Halifax, Nova Scotia, May 1978. *Tech. Rep. Fish. Mar. Serv., Envir. Can.*, 833: 11.1-11.2.

Moku M., K. Ishimaru & K. Kawaguchi. 2001. Growth of larval and juvenil Diaphus theta (Pises: Myctophidae) in the transitional water of the western Nort Pacific. Ichthyol. Res. 48: 385-390.

Moku M., A. Tsuda & K. Kawaguchi. 2002. Spawning season and migration of the myctophid fish Diaphus theta in the wester North Pacific. Ichthyol. Res. 50: 52-58.

Molina, J. 1788. Comp. Hist. Geog. Nat. y Civ. del Reino de Chile. Madrid. 217-218 p.

Nelder, J.; Wedderburn, R. 1972. Generalized linear models. J. Roy. Statist. Soc. Ser. A. 135: 370-384.

Nelder, J.; Pregibon, D. 1987. An extended quasi-likelihood function. Biometrika, 74(2), 221–232.

Nesis, K. 1970. The biology of the jumbo squid of Perú and Chile, *Dosidicus gigas*. Oceanology 10, 108-118.

Nesis, K. 1971. Genus *Dosidicus* Steenstrup, 1957. *In* Zuev, G. V. y K. N. Nesis (Eds.), Squids (Biology and Fisheries). Fishing Industry Press, Moscow: 209-215 p.

Nesis, K. 1983. *Dosidicus gigas. In:* Boyle, P. (*Ed.*), Cephalopod Life Cycles. Volume 1. Academic Press, London, 215-231 p.

Newell, G. 1963. Marine plankton a practical guide. Hutchison Biological Monographs. 207 p.

Nigmatullin, C; Nesis, K; Arkhipkin A. 2001. A review of the biology of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae). Fisheries Research 54: 9-19.

Nixon, M. 1987. Cephalopod diets. Pp. 201-219. En P. R. Boyle (Ed.), Cephalopod Life Cycles Vol. II, Comparative reviews. Academic Press, London: 441 pp.

O'Sullivan, D.; Cullen, J. 1983. Food of the squid *Nototodarus gouldi* in Bass Strait. Australian Journal of Marine and Freshwater Research 34: 261-285.

Pagan, A; Ulah, A. 1999. Nonparametric Econometrics. Cambridge: Cambridge University.

Peltonen, H; Luoto, M; Pääkkönen, J; Karjalainen, M; Tuomaala, A; Pönni, J; Viitasalo,
M. 2007. Pelagic fish abundance in relation to regional environmental variation in the Gulf of Finland, northern Baltic Sea. – *ICES J. Mar. Sci.* 64: 487–495.

Pfeffer, G. 1912. Die Cephalopoden der Plankton-Expedition. Zugleich eine monographische Übersicht der oegopsiden Cephalopoden. Ergeb. Plankton-Exped. Humboldt-Stiftung, Bd. II, F. a. Lipsius and Tischer. Kiel and Leipzig. 815 pp, 48.

Piet, G. 2002. Using external information and GAMs to improve catch-at-age indices for North Sea plaice and sole. – *ICES J. Mar. Sci.*, 59: 624–632.

Potier M.; Romanov E.; Cherel Y.; Sabatié R.; Zamorov V.; Ménard F. 2008. Spatial distribution of *Cubiceps pauciradiatus* (Perciformes: Nomeidae) in the tropical Indian Ocean and its importance in the diet of large pelagic fishes. Aquat. Living Resour. 21, 123–134

Ramsay, J; Silverman, B. 2006. Functional Data Analysis, 2nd Ed. New York: Springer.

Rasero, M; Gonzalez, A; Castro, B; Guerra, A. 1996. Predatory relationships of two sympatric squid, *Todaropsis eblanae* and *Illex coindetii* (Cephalopoda: Ommastrephidae) in Galician waters. Journal of the Marine Biological Association of the United Kingdom, 76, 73-87.

Ripley B.; Hornik K.; Gebhardt A. 2011. Package 'MASS'. Disponible en <u>http://cran.r-</u> project.org/web/packages/MASS/MASS.pdf

Rocha, F; Guerra, A; Gonzalez, A. 2001. A review of reproductive strategies in cephalopods *Biol. Rev.* (2001), 76, pp. 291±304

Rodhouse, P; Nigmatullin, C. 1996. Role as consumers. Philosophical Transactions of the Royal Society London, B, 351-1003-1022.

Rosas, R. 2007. Descripción de la alimentación del calamar gigante (*Dosidicus gigas*) D'Orbigny, 1835 en la costa occidental de la península de Baja California. Master thesis. CIBNOR, México. 100 pp.

Rosas, R., Tafur R., Alegre A., Castillo P., Cornejo R., Salinas C., Sánchez P. 2011. Trophic relationships between jumbo squid Dosidicus gigas D'Orbigny, 1835 and the mesopelagic fish *Vinciguerria lucetia* Garman, 1899 in the Humboldt Current System of Peru. Scientia Marina. 75(3) ISSN: 0214-8358.

Roper, C; Sweeney, M; Nauen, C. 1984. Cephalopods of the World. An annotated and illustrated catalogue of especies of interest to fisheries. *FAO Fish. Synop.*, 125(3): 277 p.

Roper, C; J. Sweeney M; Hochberg, F. 1995. Cefalópodos. *In* Fisher, W., F. Krupp, W. Schneider, C. Sommer, K. E. Carpenter y V. H. Niem (*Eds.*). Guía FAO para la identificación de especies para los fines de la pesca. Pacífico centro-oriental. Vol. I. Plantas e invertebrados. 305-353 p.

Rosa, F; Soler, J. 2004. Avaliando técnicas de normalização para Microarrays de cDNA. In: 16 Sinape, Caxambu - MG. Anais do 16 Sinape.

Salinas-Zavala C.A., G Bazzino-Ferreri, S., Camarillo-Coop, R., Rosas-Luis, A. Mejía-Rebollo, J.E. Ramos-Castillejos. 2010. El calamar gigante *Dosidicus gigas* (D'Orbigny 1835). Dinámica del ecosistema pelágico frente a Baja California, 1997-2007. pp: 469-486.

Sandoval, E., Uribe, M., Diaz, P., 2007. Population genetic structure of jumbo squid (*Dosidicus gigas*) evaluated by RAPD analysis. Fisheries Research 83, 113–118.

Schneider, C. 1930. Notas sobre la jibia chilena (*Ommastrephes gigas*). Bol. Soc. Biol. Concepción (Chile), III y IV, 1929-1930: 117-124.

Segura, M., R. Castillo, J. Rubio, A. Tello, M. Gutierrez y A. Aliaga. 1996. Crucero de investigaciones sobre el recurso calamar gigante o pota a bordo del BICSNP-1. CR.9510-11. Inf. Prog. Inst. Mar Perú, 34: 29-65.

Shapiro, S; Wilk, M. 1965. An analysis of variance test for normality (complete samples). *Biometrika* (1965), *Printed in &eat Britain*. Vol. 52, No. 3/4. 591-611 p. Schetinnikov, A. 1986a. Geographical variability of food spectrum of *Dosidicus gigas* (Ommastrephidae) pp. 143-153 In: Resursy i perspektivy ispol'zovaniia kal'marov mirovogo okeana : sbornik nauchnykh trudov / [redaktsionnaia kollegiia: B. G. Ivanov (otv. red.) (i dr.)]. Moskva : [VNIRO], 1986. Series title: Sbornik nauchnykh trudov (Vsesoiuznyi nauchno-issledovatel'skii institut morskogo rybnogo khoziaistva i okeanografii (Soviet Union)) Series title: Sbornik nauchnykh trudov (Vsesoiuznyi nauchno-issledovatel'skii institut morskogo rybnogo khoziaistva i okeanografii (Soviet Union))

Schetinnikov, A. 1986b. Food spectrum of *Dosidicus gigas* (Ommastrephidae) and its variation during night. *In:* Ivanov B. (*Ed.*). Resources and fishery perspectives of squid of the world ocean. VNIRO Press. Moscu. 132-142 p.

Schetinnikov, A. 1989. Food spectrum of *Dosidicus gigas* (Oegopsida) in the ontogenesis. Zoologicheskii Zhurnal. 68: 28-39 (In Russian with English abstract).

Shchetinnikov, A. S. 1992. Feeding spectrum of squid Sthenoteuthis oualaniensis (Oegopsida) in the eastern Pacific. J. mar. biol. Ass. U.K., 72: 849-860.

Schick, A. 1986. On asymptotically e-cient estimation in semiparametric models. Annals of Statistics 14: 1139-1151.

Schick, A. 1993. On e-cient estimation in regression models. Annals of Statistics 21: 1486-1521.

Schick, A. 1996. Root-n-consistent and e-cient estimation in semiparametric additive regression models. Statistics & Probability Letters 30: 45-51.

Schoenberg, I. 1964. Spline interpolation and best quadrature formulae. Bulletin of the American Mathematical Society 70: 143-148.

Silverman, B. 1984. Spline smoothing: the equivalent variable kernel method. Annals of Statistics 12: 898-916.

Silverman, B; Green, P. 1986. Density Estimation for Statistics and Data Analysis. London: Chapman and Hall. Sielfeld W., M. Vargas & R. Fuenzalida. 1995. Peces mesopelágicos frente a la costa norte de Chile (18°25´-21°47´S). Invest. Mar. 23: 83-97.

Souza, C. 2008. Testes de Hipóteses para Dados Funcionais Baseados em Distâncias: Um Estudo Usando Spline. Dissertação (Mestrado em Estatística) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica (UNICAMP/IMECC), Campinas-SP.

Speckman, P. 1988. Kernel smoothing in partial linear models. Journal of the Royal Statistical Society 50, 413-436.

Steenstrup, J. 1857. En ny Slägt af Bläcksprutter, *Dosidicus Eschrichtii* gen. et sp. n. Overs. Dansk. Vidensk. Selsk. Forh.: 11-14. [En danés] "Prof. Steenstrup gave the following information on a new species of cephalopods" Trad. Inglesa en p. 22 *In*: The Cephalopod Papers of Japetus Steenstrup, (Copenhagen, 1962).

Stefánsson, G; Pálsson, O. 1997. Statistical evaluation and modelling of the stomach contents of Icelandic cod (*Gadus morhua*). Canadian Journal of Fisheries and Aquatic Sciences 54: 169-181 p.

Stone, C. 1985. Additive regression and other nonparametric models. Annals of Statistics 13: 689-705.

Swartzman, G. 1997. Analysis of the summer distribution of fish schools in the Pacific Eastern Boundary Current. ICES J. Mar. Sci. 54,1: 106–116.

Trégouboff, G; Rosse, M. 1957. Manuel de plactonologie mediterraneénne. Tomo I y II. 268 p.

Uchikawa, K., Yamamura, O; Kitagawa, D; Sakurai; Y. 2002. Diet of the mesopelágico fish Notoscopelus japonicus (Family: Myctophidae) associated with the continental slope off the Pacific coast of Honshu, Japan. Fish. Scien. 68: 1034-1040 p.

Venables, W; Ripley, B. 2002. *Modern applied statistics with S*. 4th ed. Springer-Verlag, New York. 60 p.

Venables, W; Dichmont, C. 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. *Fish. Res.* 70:315-333.

Voss, G. L. 1983. A review of cephalopod fisheries biology. Mem. Nat. Mus. Victoria, 44: 229-241.

Whittaker, E. 1923. On new method of graduation. Proceedings of the Edinburgh Mathematical Society 41: 63-75.

Wilhelm, O. 1954. Algunas observaciones acerca de las mortandades de jibias (*Dosidicus gigas* d'Orbigny), en el litoral de Concepción. Rev. Biol. Mar., 4: 196-201.

Wolff, C. 1984. Identification and estimation of size from the beaks of 18 species of cephalopods from the Pacific Ocean. NOAA Technical Report NMFS, 17: 1-50 p.

Wood, S; Augustin, N. 2002. GAMs with integrated model selection using penalized regression spline and applications to environmental modelling. *Ecol. Model.* 157:157-177.

Wood, S. 2006. Generalized Additive Models: An Introduction with R. CRC/Chapman & Hall.

Wood, S. 2007. Package 'mgcv'. Disponible en <u>http://ryle-if-</u> a.mirrorservice.org/sites/lib.stat.cmu.edu/R/CRAN/doc/packages/mgcv.pdf

Wormuth, J. 1976. The biogeography and numerical taxonomy of the oegopsid squid family Ommastrephidae in the *Pacific Ocean. Bull. Scripps. Inst. Oceanogr.*, 23: 90 p.

Wormuth, J. 1998. Workshop deliberations on the Ommastrephidae: A brief history of their systematics and a review of the systematics, distribution, and biology of the genera *Martialia* Rochebrune and Mabille, 1889, *Todaropsis* Girard, 1890, *Dosidicus* Steenstrup, 1857, *Hyaloteuthis* Gray, 1849, and *Eucleoteuthis* Berry, 1916. *In* Voss, N., Vecchione, A., Toll R., Sweeney, M. (*Eds.*), *Systematics and biogeography of cephalopods. Vol. I.* Smith. Contrib. Zool., 586 (II): 373-383.

Yasuma H., Y. Takao, K. Sawada, K. Miyashita & I. Auki. 2006. Target strength of the lanternfish, Stenobrachius leucopsarus (family Myctophidae), a fish without an airbladder, measured in the Bering Sea. ICES Jour. of Mar. Scien. 63: 000-000.

Yatsu, A., Yamanaka, K., Yamashiro, C., 1999. Tracking experiments of the jumbo flying squid Dosidicus gigas with an ultrasonic telemetry system in the Eastern Pacific Ocean. Bul. Nat. Res. Inst. Far Seas Fisheries 36:55-60.

Yokawa, K. 1995. Isozyme comparison of large, medium and small size specimens of Dosidicus gigas. Proc. Res. Conf. Squid Resources & Fish. Cond., Hachinohe, f.y. 1993: 48-52.

Zar, J. 1999. Biostatistical Analysis. 4th Ed. Prentice Hall, New Jersey. 663 pp.

Zuev, G; Nesis, K. 1971. Role of the squid in the food chains of the ocean. Squid (biology and fishery) Vol. II. Moscow, Pishchevaya Promyshlennost' 360 p. Smithsonian Institution Washington, D.C. USA. 360 p.

Zuev, G; Nesis, K; Nigmatullin. C. 1975. System and evolution of the squid genera *Ommastrephes* and *Symplectoteuthis* (Cephalopoda, Ommastrephidae). *Zool. Zhurnal.*, 54(10): 1468-1479.

Zuur, A; Ieno, E; Walker, N; Saveliev, A; Smith, G. 2009. Mixed effects models and extensions in ecology with R. Springer.

VIII. ANEXOS

ANEXO 1a. Espectro trófico del calamar gigante *Dosidicus gigas* en el litoral peruano del 2004 al 2009.

	2004			2005 2		2006			2007		2008				2009		Total				
	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W
MOLLUSCA																					
PTEROPODA																					
Cavolinia uncinata	0.00	0.00	0.00	4.02	3.92	0.92	1.46	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.58	0.41	0.07
Diacria sp.	13.46	17.00	2.03	0.00	0.00	0.00	1.25	0.14	0.06	0.63	0.39	0.02	2.21	0.10	0.05	4.46	2.75	1.05	4.48	6.00	0.69
Pteropoda n/i	0.00	0.00	0.00	0.00	0.00	0.00	0.63	0.25	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.04	< 0.01
HETEROPODA																					
Atlanta sp	5.00	0.10	0.00	0.00	0.00	0.00	1.25	0.22	0.00	0.63	0.03	< 0.01	0.00	0.00	0.00	0.27	0.11	< 0.01	1.31	0.08	0.00
GASTROPODA																					
Natica sp.	2.88	0.09	0.01	1.34	0.03	0.00	2.71	0.26	0.00	0.63	0.03	< 0.01	0.92	0.02	0.00	1.49	0.07	< 0.01	1.79	0.09	0.00
Gastropoda n/i	0.19	< 0.01	< 0.01	0.00	0.00	0.00	0.21	0.02	< 0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	< 0.01	< 0.01
PELECYPODA																					
Larvas de Pelecypoda	0.38	< 0.01	< 0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.02	< 0.01	0.18	< 0.01	< 0.01
CEPHALOPODA																					
Abrialopsis affinis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.87	0.16	0.86	10.27	1.30	5.24	3.53	0.20	1.23
Argonauta sp.	2.69	0.05	0.23	7.59	0.38	4.27	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.02	0.06	3.92	0.36	1.93	2.33	0.11	0.78
Dosidicus gigas	15.96	0.29	47.49	14.29	0.24	20.56	11.48	0.41	22.87	23.75	1.42	69.18	11.07	0.19	19.72	11.89	0.74	27.72	14.36	0.40	35.38
Loliginidae	1.54	0.02	0.18	0.45	0.02	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.01	0.05
Cephalopoda n/i	35.58	0.92	9.02	29.46	0.85	28.91	35.49	1.32	20.40	22.50	1.53	7.48	29.34	0.56	10.66	40.27	2.07	28.81	34.62	1.06	16.54

	2004			2005			2006			2007			2008			2009			Total		
	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W
CRUSTACEA																					
Crustacea n/i	1.92	0.03	< 0.01	1.34	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.74	0.02	0.02	0.41	0.02	0.02	0.73	0.02	0.01
COPEPODA																					
Copepoda n/i	0.00	0.00	0.00	0.45	9.56	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.98	< 0.01
ZOEA																					
Zoea n/i	0.00	0.00	0.00	0.89	6.42	0.25	0.84	1.12	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.83	0.02
GAMMARIDA																					
Gammarida n/i	0.58	0.20	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.10	0.01	0.00	0.00	0.00	0.15	0.09	0.01
EUPHAUSIACEA																					
Euphausiidae	13.27	71.83	18.02	4.46	56.39	2.17	9.19	55.05	5.36	4.69	55.80	2.63	12.18	58.31	9.45	7.70	60.71	4.56	9.51	62.23	8.11
DECAPODA																					
Pleuroncodes monodon	0.38	0.01	< 0.01	4.02	2.25	6.45	0.21	< 0.01	< 0.01	0.00	0.00	0.00	4.98	1.84	2.69	1.76	1.25	0.91	1.90	0.83	1.18
Decapoda n/i	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.22	0.09	0.03	0.33	0.01	0.01
STOMATOPODA																					
Squilla panamensis	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.03	0.97	1.99	0.55	0.12	0.41
Stomatopoda n/i	0.38	0.02	0.01	0.45	0.01	< 0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.37	0.01	0.00	0.27	0.01	< 0.01	0.26	0.01	< 0.01

ANEXO 1b. Espectro trófico del calamar gigante Dosidicus gigas en el litoral peruano del 2004 al 2009.

ANEXO 1c. Espectro trófico del calamar gigante Dosidicus gigas en el litoral peruano del 2004 al 2009.

	2004		2005			2006			2007			2008			2009			Total			
-	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W
ACTINOPTERYGII																					
Teleosteii n/i	28.08	2.24	9.79	37.05	1.62	4.87	16.08	0.39	2.72	8.13	0.44	3.25	14.21	0.30	3.95	12.03	0.63	2.94	18.15	1.13	4.80
Huevos de Teleosteii n	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	24.72	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.04	1.28	0.01
PHOTICHTHYIDAE																					
Vinciguerria lucetia	35.58	5.43	7.00	46.88	9.78	11.93	44.05	33.74	36.50	23.44	9.03	3.66	41.88	31.29	27.18	26.49	22.04	13.61	36.41	18.66	16.86
BATHYLAGIDAE																					
Leuroglossus sp.	0.38	< 0.01	0.08	0.89	0.05	0.03	1.67	0.14	0.01	1.56	0.17	0.24	3.69	0.15	0.13	1.89	0.17	0.11	1.86	0.09	0.10
MYCTOPHIDAE																					
Diogenichthys laternat	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.31	4.70	2.46	9.86	2.99	1.88	5.69	1.50	0.84
Lampanyctus sp.	11.92	0.30	1.30	25.00	1.13	3.62	20.88	1.50	3.78	13.75	1.63	2.79	13.10	0.60	2.34	12.16	0.97	2.15	15.42	0.79	2.47
Myctophidae	20.77	0.84	0.76	25.45	5.78	3.63	21.29	1.38	1.18	17.19	2.81	2.59	7.38	0.89	0.44	12.30	2.03	1.07	16.51	1.69	1.32
Myctophum aurolatern	2.12	0.03	0.21	1.34	0.05	0.15	0.21	< 0.01	0.01	0.00	0.00	0.00	0.74	0.02	< 0.01	1.76	0.18	0.62	1.17	0.04	0.18
Myctophum nitidulum	0.96	0.01	0.04	4.91	0.46	0.22	4.38	0.16	0.52	4.38	1.12	0.93	2.58	0.07	0.49	1.08	0.05	0.11	2.66	0.16	0.36
Myctophum sp.	13.27	0.43	0.65	7.59	0.36	0.70	3.34	3.61	1.30	0.31	0.05	< 0.01	0.00	0.00	0.00	0.54	0.03	0.10	3.90	0.74	0.43
SPHYRAENIDAE																					
Sphyraena sp.	1.92	0.04	0.07	3.57	0.21	1.84	1.67	0.04	0.25	1.25	0.06	0.24	0.92	0.08	0.59	0.41	0.02	0.23	1.38	0.07	0.39

	2004		2005			2006		2007		2008			2009			Total					
	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W
NOMEIDAE																					
Cubiceps sp.	0.00	0.00	0.00	2.23	0.05	2.83	2.51	0.06	2.74	3.13	0.24	3.64	2.77	0.18	10.18	1.62	0.07	1.30	1.97	0.08	3.35
Psenes sio	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.16	0.10	0.29	0.58	0.01	0.06
MORIDAE																					
Physiculus sp.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.01	0.32	0.04	0.00	0.07
ENGRAULIDAE																					
Engraulidae	1.92	0.05	0.73	5.36	0.17	2.05	1.25	0.05	0.27	0.94	0.09	0.15	1.11	0.07	0.64	1.08	0.10	0.32	1.64	0.07	0.57
Engraulis ringens	1.54	0.04	0.82	6.70	0.26	4.43	0.42	0.03	0.22	5.00	0.45	3.13	6.46	0.30	4.85	2.03	0.14	1.92	3.32	0.16	2.31
CARANGIDAE																					
Trachurus murphyi	0.00	0.00	0.00	0.00	0.00	0.00	1.04	0.03	1.60	0.00	0.00	0.00	0.92	0.01	2.67	0.00	0.00	0.00	0.36	0.01	0.76
SCOMBRIDAE																					
Scombridae	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.01	0.77	0.04	0.00	0.16
EXOCOETIDAE																					
Exocoetus sp.	0.19	0.02	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.17
MERLUCCIDAE																					
Merluccius gayi peruanus	0.38	< 0.01	0.72	0.00	0.00	0.00	0.21	< 0.01	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	< 0.01	0.18

ANEXO 1d. Espectro trófico del calamar gigante Dosidicus gigas en el litoral peruano del 2004 al 2009.

	2004			2005		2006			2007			2008				2009		Total			
	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W	FO	%N	%W
THALIACEA																					
SALPIDA																					
Salpidae	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.01	0.00	0.04	< 0.01	< 0.01
РНАЕОРНУСЕАЕ																					
Laminariaceae	0.00	0.00	0.00	0.00	0.00	0.00	0.21	< 0.01	0.03	0.00	0.00	0.00	0.18	< 0.01	0.55	0.00	0.00	0.00	0.07	< 0.01	0.11
N° Presas	27		24			26				18			25			31			43		
Estómagos analizados:		650		283			589			320			657			922			3421		
Estómagos llenos:		520			224		479			239			542			740			2744		
% Estómagos llenos:		80.00			79.15		81.32			74.69			82.50			80.26			80.21		
Rango de tallas (cm):	2	1.0-104	.5	2	28.7-91.	0		27.4-98.0			28.3-109.5			4.3-112.	5	23.6-111.5			14.3-112.5		
Latitud (°S):		5°-15°			4°-15°			5°-15°			4°-10°			5°-17°			3°-15°		3°-17°		
Longitud (°W):	75°-82°				76°-82°			76°-81°			79°-83°			74°-82°			76°-84°			74°-84°	•
Dist. A la costa (mn)	21.9-115.5			25.8-137.3			2	4.7-105.	.5	18	8.5-271.4	45	2	0.3-185.	2	14	4.1-298.	6	14.1-298.6		
TSM (°C)	19.70			19.15			20.02			17.92		19.98			20.77			19.88			
Llenura (%)	15.91		12.62			13.19			21.59			13.16			10.35			17.10			

ANEXO 1f. Espectro trófico del calamar gigante *Dosidicus gigas* en el litoral peruano del 2004 al 2009.

ANEXO 2. Histogramas de frecuencias de las variables respuestas (a) porcentaje de llenura estomacal (%), (b) número de especies e (c) índice de diversidad de Margalef y las predictoras (covariables) (d) talla (cm), (e) estacionalidad, (f) periodicidad horaria (24 horas), (g) distancia a la costa (mn), (h) latitud (°S) y (i) temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.



ANEXO 3. Diagrama de cajas (boxplot) de las variables respuestas (a) porcentaje de llenura estomacal (%), (b) número de especies e (c) índice de diversidad de Margalef y las predictoras (covariables) (d) talla (cm), (e) estacionalidad, (f) periodicidad horaria (24 horas), (g) distancia a la costa (mn), (h) latitud (°S) y (i) temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.



ANEXO 4. Diagrama de cajas (boxplot) de las variables predictoras (covariables): (a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación a la variable respuesta porcentaje de llenura estomacal en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.



ANEXO 5. Diagrama de cajas (boxplot) de las variables predictoras (covariables): (a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación a la variable respuesta número de especies en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.



ANEXO 6. Diagrama de cajas (boxplot) de las variables predictoras (covariables): (a) talla (cm), (b) estacionalidad, (c) periodicidad horaria (24 horas), (d) distancia a la costa (mn), (e) latitud (°S) y (f) temperatura superficial del mar (°C), en relación a la variable respuesta índice de diversidad de Margalef en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.



ANEXO 7. Script para medir la influencia de las variables predictoras en el comportamiento de la variable respuesta llenura estomacal, usando las pruebas de bondad de ajuste AIC y desvianza explicada.

```
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
          edf Ref.df
                      F p-value
s(size) 5.271 6.443 45.37 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.036 Deviance explained =
                                             8%
GCV score = 1.5132 Scale est. = 1.5104
                                         n = 3421
> AIC(explo.gampresencepota)
[1] 1673.369
>
> explo.gampresencepota<-gam(fullness~s(seasyear,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(seasyear, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.47663
                      0.02848 -51.85 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
            edf Ref.df
                          F p-value
s(seasyear) 8.37 8.876 11.04 <2e-16 ***
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0097 Deviance explained = 2.94%
GCV score = 1.5993 Scale est. = 1.5949
                                        n = 3421
> AIC(explo.gampresencepota)
[1] 1966.835
> explo.gampresencepota<-gam(fullness~s(hour2,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness \sim s(hour2, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
                      0.03203 -46.01 <2e-16 ***
(Intercept) -1.47398
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
          edf Ref.df F p-value
s(hour2) 8.612 8.927 8.187 5.16e-12 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.00534
                      Deviance explained = 2.59%
GCV score = 1.62 Scale est. = 1.6143 n = 2734
```

```
> AIC(explo.gampresencepota)
[1] 1564.904
>
> explo.gampresencepota<-gam(fullness~s(distcoast,bs="cs"),</pre>
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(distcoast, bs = "cs")
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
                       0.02843 -52.41 <2e-16 ***
(Intercept) -1.48994
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
              edf Ref.df F p-value
s(distcoast) 6.873 7.741 17.23 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0164 Deviance explained = 3.89%
GCV score = 1.5822 Scale est. = 1.5785 n = 3421
> AIC(explo.gampresencepota)
[1] 1910.462
>
> explo.gampresencepota<-gam(fullness~s(lat,bs="cs"),</pre>
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(lat, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
                     0.02851 -51.45 <2e-16 ***
(Intercept) -1.46692
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
         edf Ref.df F p-value
s(lat) 8.374 8.83 8.867 4.17e-13 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.00496
                      Deviance explained = 2.21%
                                        n = 3421
GCV score = 1.6113 Scale est. = 1.6069
> AIC(explo.gampresencepota)
[1] 2007.414
>
> explo.gampresencepota<-gam(fullness~s(SST,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
```

```
Link function: log
Formula:
fullness \sim s(SST, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.46114 0.02863 -51.04 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
       edf Ref.df
                     F p-value
s(SST) 7.98 8.696 7.075 6.18e-10 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.00712 Deviance explained = 2.04%
GCV score = 1.6187 Scale est. = 1.6144 n = 3399
> AIC(explo.gampresencepota)
[1] 2015.353
>
```

ANEXO 8. Script para medir la influencia de las variables predictoras en el comportamiento de la variable respuesta llenura estomacal, usando las pruebas de bondad de ajuste AIC y desvianza explicada.

```
> explo.gampresencepota<-gam(fullness~s(size,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.55045
                      0.02825 -54.88 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
         edf Ref.df F p-value
s(size) 5.271 6.443 45.37 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.036 Deviance explained =
                                            8%
GCV score = 1.5132 Scale est. = 1.5104 n = 3421
> AIC(explo.gampresencepota)
[1] 1673.369
>
> explo.gampresencepota<-gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs"),</pre>
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs") + s(seasyear, bs = "cs")
```
```
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.5904
                        0.0281 -56.61 <2e-16 ***
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
              edf Ref.df
                           F p-value
s(size)
            5.211 6.385 46.48 <2e-16 ***
s(seasyear) 8.313 8.852 12.57 <2e-16 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0486 Deviance explained = 10.9%
GCV score = 1.4719 Scale est. = 1.4656 n = 3421
> AIC(explo.gampresencepota)
[1] 1517.861
>
> explo.gampresencepota<-
gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness \sim s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
   bs = "cs")
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.65078 0.03143 -52.52 <2e-16 ***
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
             edf Ref.df
                             F p-value
s(size) 6.189 7.387 31.261 < 2e-16 ***
s(seasyear) 8.673 8.960 16.229 < 2e-16 ***
           8.493 8.883 7.669 4.38e-11 ***
s(hour2)
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0569 Deviance explained = 14.7%
GCV score = 1.4339 Scale est. = 1.4211 n = 2734
> AIC(explo.gampresencepota)
[1] 1028.687
>
> explo.gampresencepota <-
gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs")+s(distcoast,bs=
"cs"),
+
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
    bs = "cs") + s(distcoast, bs = "cs")
Parametric coefficients:
```

```
Estimate Std. Error t value Pr(>|t|)
                      0.03128 -53.63 <2e-16 ***
(Intercept) -1.67733
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
              edf Ref.df
                              F p-value
             6.180 7.379 25.919 < 2e-16 ***
s(size)
s(seasyear) 8.377
                   8.863 11.168 < 2e-16 ***
            8.510 8.888 6.364 6.87e-09 ***
s(hour2)
s(distcoast) 5.586 6.477 10.257 5.75e-12 ***
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0701
                     Deviance explained = 16.7%
GCV score = 1.4061 Scale est. = 1.3909 n = 2734
> AIC(explo.gampresencepota)
[1] 941.8683
>
> explo.gampresencepota<-
gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs")+s(distcoast,bs=
"cs")+s(lat,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
   bs = "cs") + s(distcoast, bs = "cs") + s(lat, bs = "cs")
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.72510
                       0.03105 -55.55 <2e-16 ***
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
              edf Ref.df
                           F p-value
             5.888 7.099 24.174 < 2e-16 ***
s(size)
s(seasyear) 8.381 8.869 15.955 < 2e-16 ***
           8.400 8.840 4.406 1.14e-05 ***
s(hour2)
s(distcoast) 4.399 5.185 10.101 7.17e-10 ***
         8.563 8.913 13.189 < 2e-16 ***
s(lat)
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
                                            20%
R-sq.(adj) = 0.0826 Deviance explained =
GCV score = 1.3573 Scale est. = 1.3391 n = 2734
> AIC(explo.gampresencepota)
[1] 789.3068
>
> explo.gampresencepota<-
gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs")+s(distcoast,bs=
"cs")+s(lat,bs="cs")+s(SST,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
```

```
bs = "cs") + s(distcoast, bs = "cs") + s(lat, bs = "cs") +
    s(SST, bs = "cs")
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.74451
                       0.03102 -56.25 <2e-16 ***
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
              edf Ref.df
                             F p-value
s(size)
             5.816 7.030 22.777 < 2e-16 ***
s(seasyear) 8.358 8.857 13.817 < 2e-16 ***
s(hour2) 8.622 8.930 5.457 2.05e-07 ***
s(distcoast) 2.189 2.687 13.513 4.53e-08 ***
s(lat) 8.698 8.956 14.851 < 2e-16 ***
            8.226 8.805 4.676 4.39e-06 ***
s(SST)
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0847
                      Deviance explained = 21.3%
GCV score = 1.3423 Scale est. = 1.3212 n = 2730
> AIC(explo.gampresencepota)
[1] 741.5762
```

ANEXO 9. Script para medir la respuesta de la llenura estomacal con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.

```
> rm(list=ls())
> require(mgcv)
> require(tweedie)
> data<-read.table("F:\\dgigassincanib.csv", sep=";",header=T)#data</pre>
> names(data)
[1] "year"
                    "date"
                                    "month"
                                                   "seasvear"
[5] "seas_year"
                   "season"
                                   "hourl"
                                                   "hour"
[9] "hour2"
                    "hour3"
                                   "SST"
                                                   "SST1"
[13] "depth"
                    "lat"
                                    "lat1"
                                                   "long"
[17] "distcoast"
                  "distcoast1"
                                   "distshelf"
                                                   "zone"
[21] "size"
                   "classize"
                                  "sizel"
                                                   "weight"
[25] "sex"
                    "mat"
                                   "fullness"
                                                   "D_gigas"
[29] "Cephalopoda" "V_lucetia"
                                   "Myctophidae"
                                                   "Cubiceps_sp"
[33] "E_ringens"
                    "M_gayi"
                                   "T_murphyi"
                                                   "Teleosteii"
[37] "Euphausiidae" "P_monodon"
                                                   "Pteropoda"
                                    "Crustacea"
[41] "Others"
                   "withoutcanib" "species"
                                                   "num"
                   "Pielou"
[45] "Margalef"
                                   "Shanon"
                                                   "Simpson"
[49] "WSC"
                    "D_gigas1"
                                    "Cephalopoda1" "V_lucetia1"
[53] "Myctophidae1" "Cubiceps_sp1"
                                   "E_ringens1" "M_gayi1"
                                    "Euphausiidae1" "P_monodon1"
[57] "T_murphyi1"
                    "Teleosteiil"
[61] "Crustaceal"
                    "Pteropodal"
                                    "Others1"
> explo.gampresencepota<-
gam(fullness~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs")+s(distcoast,bs=
"cs")+s(lat,bs="cs")+s(SST,bs="cs"),
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
>
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
fullness ~ s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
   bs = "cs") + s(distcoast, bs = "cs") + s(lat, bs = "cs") +
   s(SST, bs = "cs")
```

```
Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
                                           <2e-16 ***
(Intercept) -1.74451
                        0.03102 -56.25
_ _ _
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
               edf Ref.df
                               F
                                  p-value
             5.816
s(size)
                    7.030 22.777
                                  < 2e-16
                    8.857 13.817
             8.358
                                  < 2e-16
                                          * * *
s(seasyear)
s(hour2)
             8.622
                    8.930 5.457 2.05e-07
                                          * * *
s(distcoast) 2.189
                    2.687 13.513 4.53e-08 ***
                    8.956 14.851 < 2e-16 ***
s(lat)
             8.698
                           4.676 4.39e-06 ***
s(SST)
             8.226
                    8.805
_ _ _
                0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Signif. codes:
R-sq.(adj) = 0.0847
                       Deviance explained = 21.3%
GCV score = 1.3423 Scale est. = 1.3212
                                            n = 2730
> par(mfrow=c(2,3))
> plot(explo.gampresencepota,se=T,residuals=F,ylim=c(-2,2.5))
>
```





ANEXO 10. Script para medir la respuesta de índice de diversidad de Margalef con las variables predictoras: talla (cm), estacionalidad, periodicidad horaria (24 horas), distancia a la costa (mn), latitud (°S) y temperatura superficial del mar (°C) en la dieta del calamar gigante *Dosidicus gigas* del 2004 al 2009.

```
> explo.gampresencepota<-
gam(Margalef~s(size,bs="cs")+s(seasyear,bs="cs")+s(hour2,bs="cs")+s(distcoast,bs=
"cs")+s(lat,bs="cs")+s(SST,bs="cs"),
+
data=data,na.action=na.exclude,method="GCV.Cp",family=Tweedie(1.628571,link="log"
),control = gam.control(irls.reg=0.0,mgcv.tol=1e-6,mgcv.half=15))
>
> summary(explo.gampresencepota)
Family: Tweedie(1.628571)
Link function: log
Formula:
Margalef ~ s(size, bs = "cs") + s(seasyear, bs = "cs") + s(hour2,
   bs = "cs") + s(distcoast, bs = "cs") + s(lat, bs = "cs") +
   s(SST, bs = "cs")
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
                       0.04265 -14.10 <2e-16 ***
(Intercept) -0.60147
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
Approximate significance of smooth terms:
                edf Ref.df F p-value
           7.45173 8.41510 7.115 1.14e-09 ***
s(size)
s(seasyear) 8.54288 8.92706 10.697 3.38e-16 ***
            8.50595 8.91750 2.950 0.00182 **
s(hour2)
s(distcoast) 7.63722 8.40993 7.094 1.25e-09 ***
s(lat) 8.89975 8.99412 10.648 3.24e-16 ***
s(SST)
           0.01198 0.02361 0.012
                                        NA
Signif. codes: 0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1
R-sq.(adj) = 0.0291
                     Deviance explained = 16.5%
GCV score = 2.6506 Scale est. = 2.5896
                                        n = 1828
> par(mfrow=c(2,3))
> plot(explo.gampresencepota,se=T,residuals=F,ylim=c(-2,2.5))
```



distcoast

