



Análisis estadístico de capturas de camarón y fauna con la red prototipo RS-INP-MEX en el Alto Golfo de California durante la temporada de camarón 2010 / 2011

En el marco de las actividades y objetivos expresados en el protocolo de trabajo

Pesca Experimental de Camarón con la Red de Arrastre Prototipo RS-INP-MEX en el Alto Golfo de California: Capacitación al Sector Productivo en la Construcción, Operación y Mantenimiento de la Red y Colecta de Información en Campo. Campaña 2010/2011+

Luis Vicente González Ania
Daniel Aguilar Ramírez
Adauto Abel Flores Santillán
Araceli Ramos Montiel
Raymundo Torres Jiménez
Alejandro Liedo Galindo

Instituto Nacional de Pesca

*Dirección General de Investigación
Pesquera en el Pacífico Norte*

Material y método para análisis estadístico

Datos analizados

Se analizó la eficiencia de captura (kg/hora de arrastre) de camarón y de fauna de acompañamiento (FAC) con valor comercial, así como la proporción de camarón respecto a la captura total (kg de camarón/kg totales), en 1,084 lances de pesca experimental en pangas, realizados con cuatro tipos de redes de arrastre en el Alto Golfo de California, durante la temporada de pesca de camarón 2010 - 2011 (Tabla 1).

Tabla 1. Esfuerzo pesquero, eficiencia media de captura (CPUE) y proporción de camarón por tipo de arte.

Tipo de arte	Lances ¹	Esfuerzo promedio (horas/lance)	Captura por unidad de esfuerzo (kg/hora de arrastre)				Proporción de camarón por lance ²	
			Camarón	FAC		Total	Media	Coef. de variación ³
				Comercial	Descarte			
Prot. RSINP	134	0.95	4.73	8.45	9.32	22.49	0.19	1.05
Fantasma	600	0.96	1.73	5.95	5.73	13.41	0.12	1.39
Mixto	265	0.98	0.85	1.55	4.58	6.98	0.09	1.53
Super-mixto	85	0.98	0.20	2.65	3.90	6.75	0.03	1.77
TOTAL	1,084	0.97	1.77	4.92	5.75	12.44	0.11	1.42

¹ Lances con información completa y validada para los análisis.

² Estimada en una submuestra de 970 lances con captura total > 0.

³ s / \bar{x}

Modelos utilizados

Las tasas nominales de captura por unidad de esfuerzo (CPUE) de camarón y FAC comercial, al igual que las proporciones de camarón en la captura total, fueron analizadas en los 4 tipos de redes utilizando Modelos Lineales Generalizados, conocidos como GLM por sus siglas en inglés (Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989). Se plantearon 3 modelos: para CPUE de camarón, CPUE de FAC comercial y para proporción de camarón, los cuales se examinaron realizando pruebas de hipótesis. Los parámetros del GLM se estiman por el principio de máxima verosimilitud, mediante algoritmos iterativos, por lo que las pruebas de hipótesis se basan en comparaciones de verosimilitudes entre modelos anidados, evaluando una medida de bondad de ajuste que es la discrepancia (del inglés *deviance*) entre las observaciones y el modelo ajustado.

El modelo lineal generalizado extiende el concepto de los modelos lineales clásicos, dando cabida tanto a las respuestas con distribuciones no normales, como a las transformaciones para linearizar las relaciones entre las variables. El GLM también relaja los supuestos de igualdad o constancia en las varianzas, requeridos para las pruebas de hipótesis en los modelos lineales tradicionales.

En los modelos GLM se define una expresión lineal (predictor lineal η), que se relaciona con la variable de respuesta (media esperada: μ), mediante una función de enlace invertible [$g(\mu)$], de manera que en el presente análisis la expresión general de los modelos es la siguiente:

$$\eta_i = g(\mu_i) = \sum_{j=1}^k x_j \beta_j + \varepsilon$$

Donde μ representa la estimación media de la CPUE o de la proporción de camarón, los coeficientes β_j son los parámetros a estimar con los datos, x_j son los 4 niveles del predictor o variable explicativa categórica (factor tipo de red) y ε corresponde a un término de error residual aleatorio.

Para los análisis de CPUE, la estructura de error se eligió comparando modelos con 3 distribuciones diferentes, mediante el Criterio de Información de Akaike (1974), que permite comparar la verosimilitud de modelos no anidados, de manera que el modelo con mayor bondad de ajuste es el que presente una menor discrepancia entre las observaciones y el modelo ajustado, de acuerdo con el número de parámetros a estimar. En los modelos de CPUE se usó la función de enlace logarítmica [$g(\mu) = \log(\mu)$], que además de linearizar la relación, evita la predicción de tasas de captura negativas (que no son posibles), ya que cada valor estimado (μ_i), es el antilogaritmo del predictor lineal:

$$\mu_i = \exp(\eta_i) > 0$$

En el caso de la proporción de captura de camarón (variable de respuesta acotada por 0 y 1), se utilizó un modelo con error binomial y función de enlace probit.

Exploración del parámetro de dispersión

El parámetro de dispersión (ϕ) en los GLM permite expresar la varianza de la respuesta [$\text{var}(Y)$], como una función de la respuesta media [$V(\mu)$]:

$$\text{var}(Y) = \phi V(\mu)$$

En un modelo con estructura de error normal, el parámetro de dispersión equivale directamente a la varianza estimada de los datos [ya que la función de varianza es constante: $V(\mu) = 1$]. En cambio, en los modelos de un solo parámetro (binomial, exponencial y de Poisson), se asume que ϕ es constante ($\phi = 1$), porque la varianza es una función conocida de la media [$\text{var}(Y) = V(\mu)$]. En el modelo binomial negativo ϕ también se asume constante, pero la función de varianza contiene un parámetro adicional [$\text{var}(Y) = V(\mu, \phi)$].

Sin embargo, en el análisis de muchos problemas este supuesto no es válido (Crawley, 2004), por lo que el valor de ϕ debe estimarse y en su caso ser modelado adecuadamente, pues de lo contrario se puede incurrir en sesgos de las estimaciones, como los resumidos en la Tabla 2.

Tabla 2. Desviaciones del parámetro de dispersión ($\phi \neq 1$) y algunos posibles sesgos resultantes.

Valor de ϕ	Denominación	Efecto en pruebas de hipótesis	Error estándar	Intervalo de confianza
$\phi > 1$	Sobredispersión (varianza sobreestimada)	Mayor posibilidad de cometer error Tipo I, por <i>subestimación</i> de la probabilidad de la hipótesis nula.	Subestimado	Subestimado (más estrecho)
$\phi < 1$	Sub-dispersión (varianza subestimada)	Mayor posibilidad de cometer error Tipo II (menor poder de prueba), por <i>sobreestimación</i> de la probabilidad de la hipótesis nula.	Sobreestimado	Sobreestimado (más amplio)

En un modelo binomial negativo, las desviaciones ($\phi \neq 1$) pueden corregirse modificando el parámetro ϕ . En los modelos de un solo parámetro, el valor de ϕ se puede estimar y modelar sustituyendo la estructura de error original por una de quasi-verosimilitud, la cual requiere solamente las funciones de

enlace y varianza del modelo inicial. Para el modelo binomial, $g(\mu) = \text{probit}(\mu)$ y $\text{var}(Y) = \phi\mu(1-\mu)/n$, respectivamente. En un modelo de quasi-verosimilitud no es necesario conocer exactamente la distribución de error de la variable de respuesta y se asume menos durante la estimación e inferencia.

Pruebas de hipótesis

Inicialmente se realizó para cada modelo una prueba de contrastes *a priori*, análoga al análisis de varianza de una vía, bajo la hipótesis nula de iguales tasas de captura o proporciones medias de camarón en los 4 tipos de redes. La hipótesis alterna no es particularmente informativa, pues solamente implica que al menos una de las medias es significativamente distinta de las demás, por lo que en caso de encontrar un efecto significativo del factor red, es posible comparar redes específicas mediante un proceso *a posteriori* de simplificación del modelo, por el procedimiento de colapso o fusión de niveles del factor (Crawley 2004).

Esta técnica es de uso común en la modelación estadística y en el presente análisis consiste en reunir en un mismo nivel las observaciones de 2 tipos de redes a comparar. Con este factor colapsado se crea un modelo alternativo reducido en un grado de libertad, para compararlo contra el modelo original con una prueba de Anova. Si existe una diferencia significativa en la forma en que ambos modelos describen los datos (su poder explicativo), se puede admitir la hipótesis alterna, la cual asume que existe diferencia entre ambos tipos de redes.

Cuando un modelo no presenta desviaciones del parámetro de dispersión, las pruebas de hipótesis se realizan con el estadístico *Chi*. En caso de emplear modelos de quasi-verosimilitud, se aplica la *F* de Fisher.

Resultados del análisis estadístico

La Tabla 3 resume el análisis comparativo de los modelos de CPUE alternos por medio del Criterio de Información de Akaike (AIC), para los 3 tipos de distribución de error propuestos. Estos resultados muestran que la estructura de error binomial negativa es la que produce mejor bondad de ajuste tanto en los modelos para analizar la eficiencia de captura de camarón (AIC = 3,062.5), como de FAC comercial (AIC = 3,001.8).

Tabla 3. Comparación de modelos con diferentes estructuras de error mediante el criterio de Información de Akaike (AIC).

Estructura de error (ϵ)	Función de enlace [$g(\mu)$]	Grados de libertad	Respuesta (CPUE)	AIC
Binomial negativa	Logarítmica	5	Camarón	3,062.5
			FAC comercial	3,001.8
Gamma	Logarítmica	5	Camarón	4,174.4
			FAC comercial	3,939.2
Poisson	Logarítmica	5	Camarón	5,154.4
			FAC comercial	6,287.4

Eficiencia de captura de camarón

La Figura 1 muestra la CPUE de camarón para cada arte de pesca. Se puede observar que los mayores valores de media y mediana corresponden a la red prototipo RSINP y los menores a la red tipo Super-mixto.

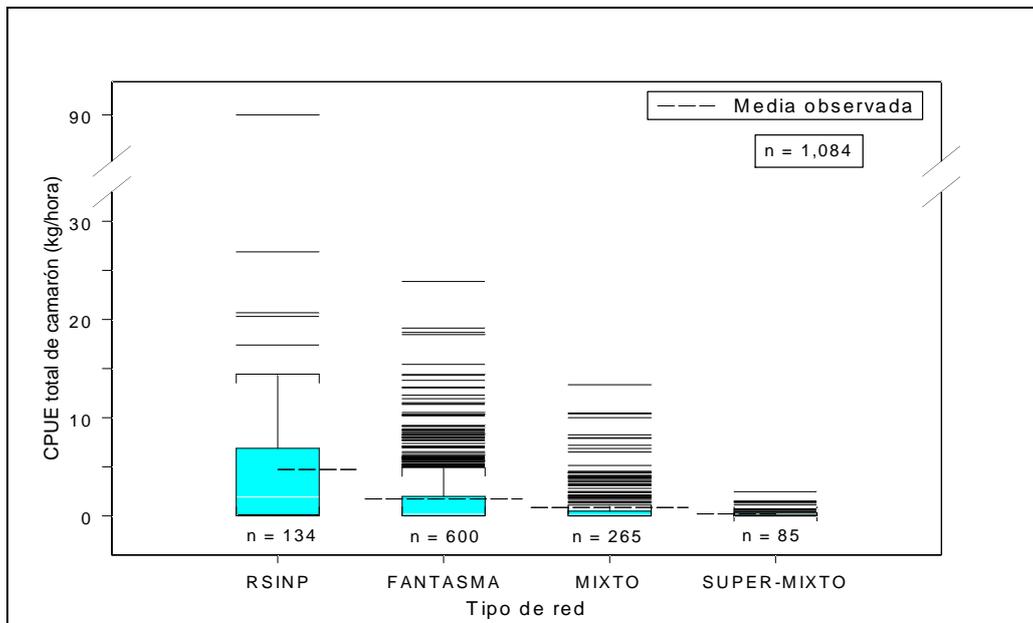


Figura 1. Tasas observadas de captura de camarón por tipo de arte. Los rectángulos centrales comprenden el rango intercuartílico y la línea blanca corresponde a la mediana. La línea discontinua representa la media de las observaciones.

La prueba inicial de contrastes *a priori* (Tabla 4), demostró que el tipo de red ejerce un efecto altamente significativo en la CPUE de camarón ($p < 1.00e-015$).

Tabla 4. Efecto del tipo de red en la CPUE de camarón.

Model	Terms	Df	Deviance	Resid. Df	Resid. Dev.	Pr(Chi)
NULL	----	----	----	1,083	1,151.772	----
I	RED	3	134.871	1,080	1,016.901	< 1.00e-015

En la Tabla 5 y Figura 2 se muestran las predicciones medias del modelo ajustado para la captura por unidad de esfuerzo de camarón, con intervalos de 95% de confianza.

Tabla 5. Predicciones del modelo ajustado para CPUE de camarón por tipo de red.

Arte de pesca	CPUE (kg/h)	Intervalo de 95% de confianza	
		Límite inferior	Límite superior
Prototipo RSINP	4.729	3.492	6.405
Fantasma	1.730	1.488	2.013
Mixto	0.852	0.667	1.090
Super-mixto	0.198	0.108	0.362

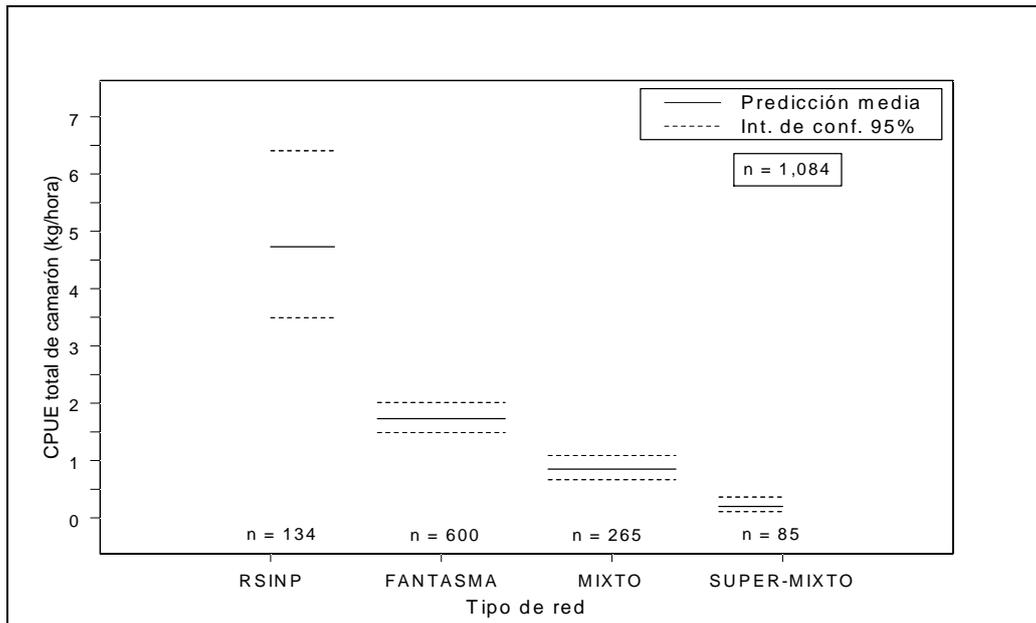


Figura 2. Predicciones del modelo ajustado para CPUE de camarón por tipo de red.

La Tabla 6 resume los resultados de las comparaciones entre redes. La eficiencia de todas es significativamente distinta y la red prototipo RSINP es la más eficiente.

Tabla 6. Comparación de la CPUE de camarón entre redes mediante colapso de niveles del factor red.

Model ¹	Test	Df	Resid. Df	Resid. Dev.	Df change	Dev. change	Pr(Chi)
Original	----	3	1,080	1,016.901	----	----	----
I	RSINP vs. Fantasma ²	2	1,081	1,056.521	-1	-39.620	3.09e-010
II	Fantasma vs. Mixto ²	2	1,081	1,038.860	-1	-21.959	2.79e-006
III	Mixto vs. Super-mixto ²	2	1,081	1,036.999	-1	-20.098	7.36e-006

¹ El factor red tiene 4 niveles en el modelo original; 3 niveles en los modelos simplificados (I, II y III).

² Contrastes contra el modelo original.

Eficiencia de captura de FAC comercial

La Figura 3 muestra la CPUE de FAC con valor comercial para cada arte de pesca. Las redes RSINP y Fantasma presentan las medias más altas y distribuciones semejantes entre sí.

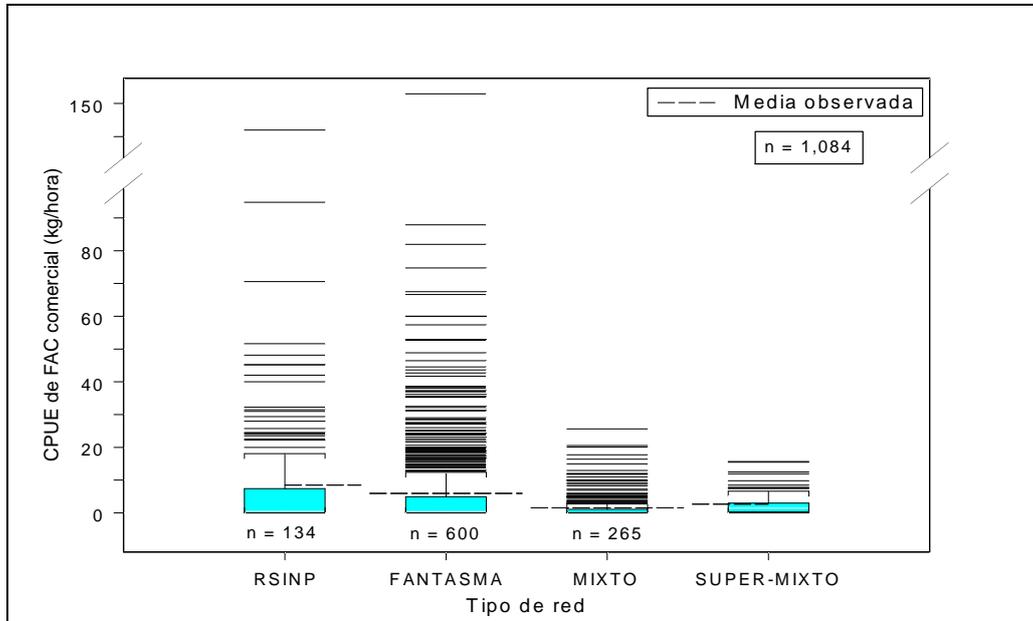


Figura 3. Tasas observadas de captura de FAC comercial por tipo de arte. Los rectángulos centrales comprenden el rango intercuartílico y la línea blanca corresponde a la mediana. La línea discontinua representa la media de las observaciones.

La prueba inicial de contrastes *a priori* (Tabla 7), mostró que el tipo de red produce un efecto altamente significativo en la CPUE de FAC comercial ($p = 2.78e-015$).

Tabla 7. Efecto del tipo de red en la CPUE de FAC comercial.

Model	Terms	Df	Deviance	Resid. Df	Resid. Dev.	Pr(Chi)
NULL	----	----	----	1,083	1,068.412	----
I	RED	3	70.869	1,080	997.543	2.78e-015

La Tabla 8 y Figura 4 contienen las predicciones medias del modelo para la captura por unidad de esfuerzo de FAC comercial, con los respectivos intervalos de 95% de confianza.

Tabla 8. Predicciones del modelo GLM para CPUE de FAC comercial por tipo de red.

Arte de pesca	CPUE (kg/h)	Intervalo de 95% de confianza	
		Límite inferior	Límite superior
Prototipo RSINP	8.449	5.845	12.213
Fantasma	5.948	4.993	7.086
Mixto	1.548	1.174	2.041
Super-mixto	2.647	1.646	4.257

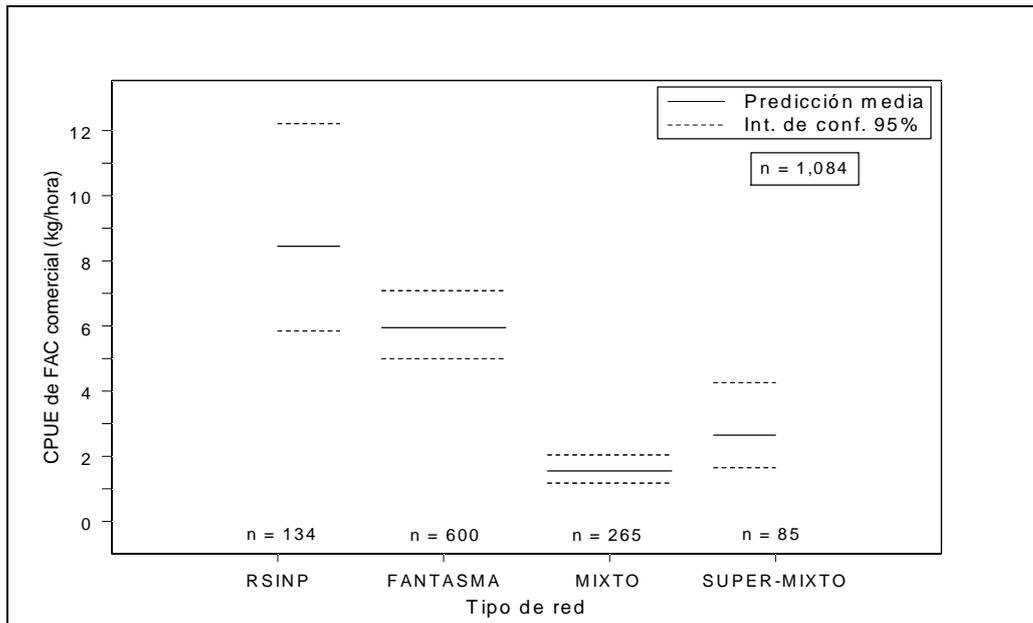


Figura 4. Predicciones del modelo GLM para CPUE de FAC comercial por tipo de red.

La Tabla 9 presenta los resultados de las comparaciones de CPUE entre redes. Si bien la red prototipo RSINP tiene la media más alta, no difiere estadísticamente de la red Fantasma. Tampoco hay diferencia significativa entre la eficiencia de los tipos Mixto y Super-mixto.

Tabla 9. Comparación de la CPUE de FAC comercial entre redes mediante colapso de niveles del factor red.

Model ¹	Test	Df	Resid. Df	Resid. Dev.	Df change	Dev. change	Pr(Chi)
Original	----	3	1,080	997.543	----	----	----
I	RSINP vs. Fantasma ²	2	1,081	1,000.593	-1	-3.050	0.08
II	Mixto vs. Super-mixto ³	1	1,082	1,004.524	-1	-3.931	0.05

¹ El factor red tiene 4 niveles en el modelo original; 3 niveles en el modelo I; 2 niveles en el modelo II.

² Contraste contra el modelo original.

³ Contraste contra el modelo I.

Estos resultados justifican la simplificación progresiva del modelo original (con 4 niveles del factor red), a uno en el que se reúnen las redes RSINP y Fantasma en un solo nivel y los tipos Mixto y Super-mixto en otro.

El análisis de este nuevo modelo simplificado para comparar ambos grupos de redes (RSINP y Fantasma vs. Mixto y Super-mixto), comprueba que difieren muy significativamente (Tabla 10).

Tabla 10. Efecto del grupo de redes en la CPUE de FAC comercial.

Model	Terms	Df	Deviance	Resid. Df	Resid. Dev.	Pr(Chi)
NULL	----	----	----	1,083	1,068.412	----
I	GRUPO.RED	1	63.888	1,082	1,004.524	1.33e-015

Proporción de camarón en la captura total

La proporción o fracción de camarón en la captura total representa el balance neto de eficiencia para cada arte de pesca utilizado, al medir simultáneamente la eficiencia de captura de camarón y de exclusión de fauna de acompañamiento total.

La Figura 5 muestra la proporción de captura de camarón para las cuatro redes de arrastre. Los valores más altos de media y mediana corresponden a la red prototipo RSINP y los menores a la red Super-mixto. Las redes de tipo Fantasma y Mixto presentan resultados semejantes entre sí.

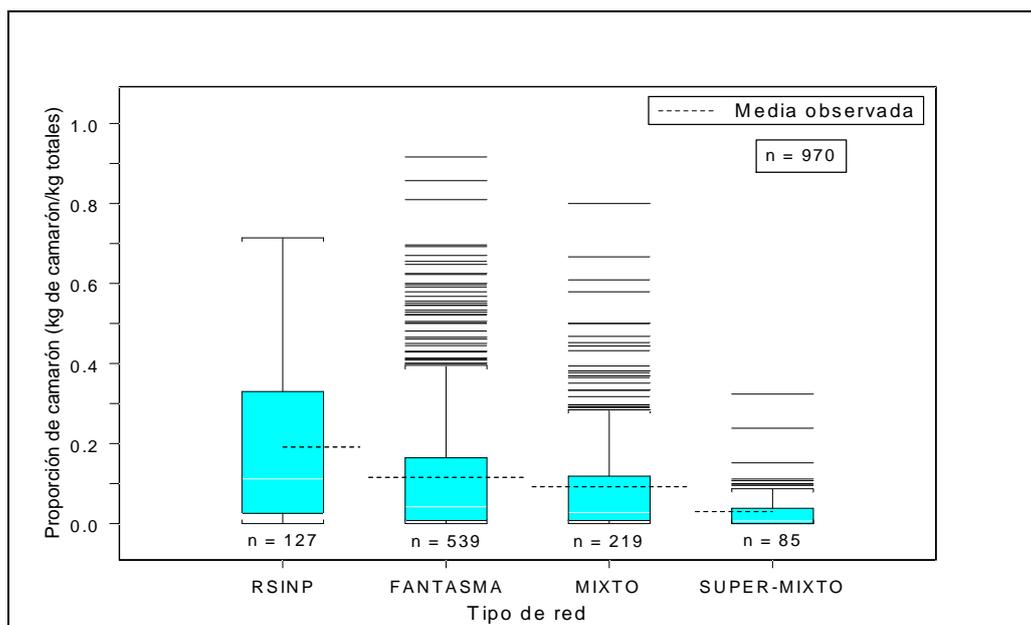


Figura 5. Proporción observada de camarón por tipo de arte. Los rectángulos centrales comprenden el rango intercuartílico y la línea blanca corresponde a la mediana. La línea discontinua representa la media de las observaciones.

El análisis estadístico de la pesca experimental realizada, revela un efecto muy significativo del tipo de red en la proporción de camarón (Tabla 11).

Tabla 11. Efecto del tipo de red en la proporción de captura de camarón.

Model	Terms	Df	Deviance	Resid. Df	Resid. Dev.	F Value	Pr(F)
NULL	----	----	----	969	211.118	----	----
I	RED	3	15.753	966	195.365	22.284	5.88e-014

En la Tabla 12 y Figura 6 se muestran las predicciones medias del modelo ajustado para la proporción de camarón en la captura total.

Tabla 12. Predicciones del modelo ajustado para proporción de camarón en la captura total, por tipo de red.

Arte de pesca	Proporción media de camarón	Intervalo de 95% de confianza	
		Límite inferior	Límite superior
Prototipo RSINP	0.191	0.160	0.226
Fantasma	0.116	0.103	0.129
Mixto	0.092	0.075	0.112
Super-mixto	0.029	0.016	0.052

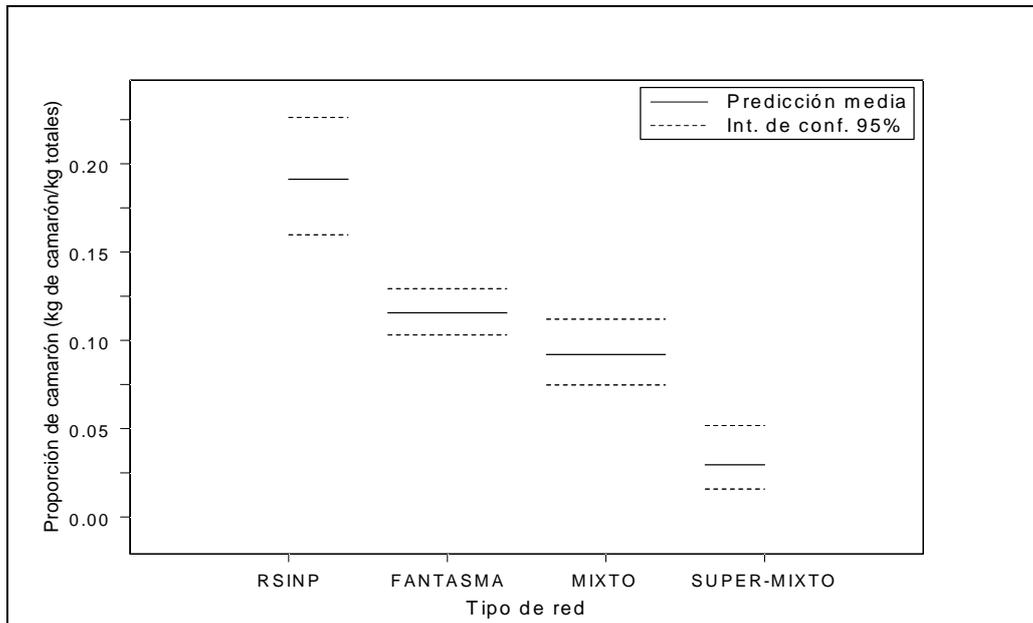


Figura 6. Predicciones del modelo para proporción de camarón en la captura total, por tipo de red.

Las pruebas de contrastes *a posteriori* (Tabla 13), señalan que la proporción de camarón en la captura total es básicamente la misma en las redes tipo Fantasma y Mixto, pero ambas en conjunto difieren en forma significativa tanto de la red prototipo RSINP, como de la tipo Super-mixto.

Tabla 13. Comparación de la proporción de camarón entre redes mediante colapso de niveles del factor red.

Model ¹	Test	Df	Resid. Df	Resid. Dev.	Df change	Dev. change	F Value	Pr(F)
Original	----	3	966	195.365	----	----	----	----
I	Fantasma vs. Mixto ²	2	967	196.288	-1	-0.922	3.914	0.05
II	Fantasma y Mixto vs. RSINP ³	1	968	202.512	-1	-6.224	26.308	3.52e-007
III	Fantasma y Mixto vs. Super-mixto ³	1	968	203.234	-1	-6.946	29.359	7.59e-008

¹ El factor red tiene 4 niveles en el modelo original; 3 niveles en el modelo I; 2 niveles en los modelos II y III.

² Contraste contra el modelo original.

³ Contrastes contra el modelo I.

Discusión y conclusiones del análisis estadístico

Los 1,084 lances de pesca analizados se caracterizaron en promedio por un esfuerzo cercano a 1 hora de arrastre (0.97 h), CPUE de camarón de 1.77 kg/hora, CPUE de FAC comercial de 4.92 kg/hora y 11% de camarón respecto a la captura total. Los resultados obtenidos corresponden prácticamente en su totalidad al camarón café, cuya captura representó el 94.77% (5.23% de camarón azul).

El tipo de red de arrastre empleado produjo efectos altamente significativos en la CPUE media de camarón ($p < 1.00e-015$) y de FAC con valor comercial ($p = 2.78e-015$), así como en la proporción media de camarón en la captura total ($p < 5.88e-014$).

Se encontraron diferencias sustanciales en la CPUE de camarón de todas las redes experimentadas ($3.09e-010$ $m p$ $m 7.36e-006$), siendo en este aspecto la red prototipo RSINP la más eficiente, con una CPUE media de 4.73 kg/hora (IC 95%: 3.49 . 6.40 kg/hora), mucho mayor que el promedio general.

En cuanto a las tasas medias de captura de FAC comercial, se identificaron 2 grupos de redes con niveles de desempeño estadísticamente muy distintos ($p = 1.33e-015$): uno de alta eficiencia integrado por las redes prototipo RSINP y Fantasma (8.45 y 5.95 kg/hora, respectivamente), y otro con los tipos Mixto y Super-mixto (1.55 y 2.65 kg/hora, respectivamente).

El resultado de eficiencia comparativa estimada como la proporción media que representó el camarón en la captura total de cada clase de red, confirmó que el prototipo RSINP tiene un desempeño significativamente superior a las demás (19%; $p = 3.52e-007$). Por otra parte, mientras las otras redes presentan un comportamiento muy disperso en este parámetro (1.39 m CV m 1.77), la red RSINP muestra mayor consistencia (CV = 1.05), lo que la hace más confiable y predecible en el resultado de proporción de captura de camarón que se puede obtener.

Referencias bibliográficas

- Akaike, H. 1974. A new look at statistical model identification. IEEE Transactions on Automatic Control AU-19, 716-722.
- Crawley, M. J. 2004. Statistical Computing. An Introduction to Data Analysis using S-Plus. John Wiley & Sons Ltd, West Sussex. 761 p.
- McCullagh, P. and J. A. Nelder. 1989. Generalized Linear Models, 2nd edition. Chapman and Hall, London. 511 p.
- Nelder, J. A. and R. W. M Wedderburn. 1972. Generalized linear models. Journal of the Royal Statistical Society, Series A 135: 370-384.

§ § §