

INFLUENCIA DE LAS INTERACCIONES CON EL FACTOR AÑO EN LOS ÍNDICES ANUALES DE ABUNDANCIA OBTENIDOS POR MODELOS LINEALES GENERALES UTILIZANDO DATOS DE CAPTURA POR UNIDAD DE ESFUERZO*

por

DANIEL R. HERNÁNDEZ¹ y RICARDO G. PERROTTA^{1,2}

¹Instituto Nacional de Investigación y Desarrollo Pesquero (INIDEP),
Paseo Victoria Ocampo N° 1 Escollera Norte, B7602HSA - Mar del Plata, Argentina
e-mail: danielh@inidep.edu.ar

²Departamento de Ciencias Marinas, Facultad de Ciencias Exactas y Naturales, Universidad
Nacional de Mar del Plata (UNMDP), Funes 3350, B7602AYL - Mar del Plata, Argentina

SUMMARY

The influence of interactions with the factor Year in the annual abundance index estimations obtained from general linear models using the commercial fleet CPUE data. The problem in the definition of a relative annual abundance index using a General Linear Model is posed. The possible factors generating changes in the catch per unit effort (CPUE) are discussed and the parameterization of the model is analyzed, in particular the Sigma-restrictive parameterization. This report considers the problem caused by the presence of interactions of first order that include the factor Year in the model. It has been mathematically demonstrated that in the presence of these interactions, the index series is not proportional to the abundance annual mean of the resource. An approach of practical order to evaluate the possible effect of the interactions that include the factor Year in the index is considered using the percentage of explained variance of the model with and without these interactions. Some methodological rules are also defined to attenuate several problems generated by the mentioned interactions: (1) evaluating if type of fleet, included as categories of the factor Ship, has experienced modifications throughout the period of study that could have generated changes in the efficiency of fishing units; (2) allowing for the possibility of working with indexes pondered by Area; (3) using the most convenient categorization for the factor Month (month, bimonthly, quarterly, fishery season, etc.), considering the seasonal distribution of the resource and the activity of the fleet, as well as the limitations generated by the presence of empty cells in the design. Two application examples are presented.

RESUMEN

Se plantea un problema central de la biología pesquera como es la definición de un índice de abundancia anual relativa, obtenido en este trabajo a partir de un Modelo Lineal General. Se discuten los factores que pueden generar cambios en la Captura por Unidad de Esfuerzo y se analiza la reparametrización del modelo, en particular la parametrización Sigma-restrictiva. Se considera el problema generado por la presencia en el modelo de interacciones de primer orden que incluyan el factor Año, y se demuestra matemáticamente que ante la presencia de estas interacciones, la serie de los índices no es proporcional a la abundancia anual media del recurso. Se plantea un criterio de orden práctico, orientado a evaluar el posible efecto de las interacciones que incluyan al factor Año sobre el índice, considerando el porcentaje de varianza explicada teniendo en cuenta el modelo con y sin estas interacciones. Se definen algunas reglas metodoló-

gicas, con el objeto de contribuir a atenuar el problema generado por las interacciones mencionadas: (1) evaluar si los estratos o tipos de flota, a incluir como categorías del factor Barco, han experimentado modificaciones a lo largo del período de estudio que puedan haber generado cambios en la eficiencia de las unidades de pesca; (2) considerar la posibilidad de trabajar con índices ponderados por las áreas; (3) utilizar la categorización más conveniente para el factor Mes (mes, bimestre, trimestre, temporada de pesca, etc.), teniendo en cuenta cuestiones relacionadas con la distribución estacional del recurso y con la actividad de la flota, así como las limitaciones generadas por la presencia de celdas vacías en el diseño. Se presentan dos ejemplos de aplicación.

Key words: Abundance annual index, multiplicative model, general linear model, catch per unit effort.

Palabras clave: Índices de abundancia anual, modelo multiplicativo, modelo lineal general, captura por unidad de esfuerzo.

INTRODUCCIÓN

Una cuestión muy importante en los estudios de las pesquerías es la determinación de índices de abundancia relativa que puedan ser utilizados, en primera instancia, como indicadores de la abundancia real de un recurso, permitiendo obtener una imagen de su evolución en el tiempo, o bien, que puedan ser utilizados para calibrar modelos de evaluación a los efectos de efectuar diagnósticos del estado de un recurso y estimar parámetros de manejo, necesarios para la explotación racional del recurso.

Los Modelos Lineales Generales (MLG) se utilizan cada vez más en aplicaciones orientadas a la estimación de esos índices (Large, 1992; Goñi *et al.*, 1999; Punt *et al.*, 2000; Horn, 2003; Perrotta y Hernández, 2005; Carozza *et al.*, 2004), dado que permiten integrar en forma simple los diferentes factores e interacciones que influyen sobre la variación de la captura por unidad de esfuerzo (CPUE). Esta característica –sumada a las facilidades computacionales que ofrecen los programas comerciales de análisis estadístico– (por ejemplo: SYSTAT, STATISTICA y S-PLUS), los ha popularizado como herramienta de análisis y estimación.

El propósito de este trabajo es mostrar, a partir de consideraciones de orden matemático, los problemas asociados con la presencia de interacciones que incluyan el factor Año cuando se construyen índices de abundancia anual y sugerir criterios prácticos para tener en cuenta al utilizar un MLG para ese fin.

MATERIALES Y MÉTODOS

Modelo multiplicativo

La ecuación básica que relaciona la CPUE de un buque, con su poder de pesca y la densidad sobre la cual está operando, está dada por Gulland (1971):

$$CPUE = p D \quad (1)$$

donde:

CPUE: captura por unidad de esfuerzo (por ejemplo kg/h o t/h).

p : poder de pesca del buque considerado.

D : densidad local sobre la cual opera el buque al efectuar el lance de pesca.

Los factores básicos que pueden generar cambios en los dos parámetros que figuran en el término derecho de (1) son:

- (1) La diferencia entre buques, debida a diversos factores (tamaño, potencia, etc.), que pueden generar cambios en el poder de pesca.
- (2) El área sobre la cual opera el buque, que tiene asociados cambios en la densidad local del recurso. El área indirectamente puede modificar el poder de pesca, ya que al modificar la densidad podría ocurrir, por ejemplo, que un mismo buque fuera más eficiente en un área de alta que en un área de baja densidad.
- (3) El mes (bimestre, trimestre, temporada de pesca, etc.) puede modificar la densidad global y, por supuesto, la densidad local, debido a las

variaciones estacionales en el patrón de distribución del recurso.

(4) El año también puede modificar la densidad global y local, debido a los cambios en la biomasa anual generados por el reclutamiento, el crecimiento, la mortalidad natural y la pesca. Existe también la posibilidad de que se produzcan cambios anuales en el patrón de distribución del recurso.

A estos factores los denominaremos: “factor Barco”, “factor Área”, “factor Mes”¹ y “factor Año”, respectivamente.

De esta forma, si queremos contemplar todas estas fuentes de variabilidad, debemos escribir el modelo (1) en la forma:

$$CPUE_{bamt} = p_{bamt} D_{amt} \quad (2)$$

donde b, a, m, t indican un cierto buque (o estrato de flota), operando en cierta área, cierto mes (bimestre, trimestre, etc.) y un año determinado, respectivamente.

Al modelo (2) se le agrega, además, un término de error multiplicativo, ya que si bien los factores: barco, área, mes y año son explicativos de una parte de la varianza de la CPUE, se sabe también que existe en general una porción de la varianza no explicada por el modelo (generalmente importante), debido a una infinidad de factores no contemplados en la ecuación (2), la cual es resumida en el término de error. De esta forma, el modelo (2) se puede escribir como:

$$CPUE_{bamt} = p_{bamt} D_{amt} e^{\varepsilon_{bamt}} \quad (3)$$

El modelo (3), como está expresado, es sumamente general pero tiene parámetros de más, lo cual no permite efectuar la estimación de los parámetros involucrados. Esto nos lleva a tener que plantear ciertos supuestos (por ejemplo, que el término

p_{bamt} se puede representar como el producto de otros términos), factorizando $p_{bamt} D_{amt}$ e introduciendo nuevos parámetros.

El término p_{bamt} se puede factorizar en la forma:

$$p_{bamt} = p_b p_{ba} p_{bm} p_{bt} \quad (4)$$

Por su parte, D_{amt} se puede factorizar y expresar en la forma:

$$D_{amt} = D_a D_m D_{am} D_{at} D_{mt} \bar{D}_t \quad (5)$$

donde \bar{D}_t la densidad media global correspondiente al año t .

Teniendo en cuenta (4) y (5), el modelo (3), incluyendo solo interacciones de primer orden, se puede escribir en la forma:

$$CPUE_{bamt} = p_b p_{ba} p_{bm} p_{bt} D_a D_m D_{am} D_{at} D_{mt} \bar{D}_t e^{\varepsilon_{bamt}} \quad (6)$$

El modelo (6) es denominado modelo multiplicativo, por razones obvias. En un caso particular bajo análisis, el modelo no tiene por qué incluir todas las interacciones, depende de las características del problema en estudio.

Modelo de ANOVA multifactorial

Si en el modelo (6) tomamos logaritmos en ambos miembros y ordenamos los términos, obtenemos el siguiente modelo:

$$\ln(CPUE_{bamt}) = \ln(p_b) + \ln(D_a) + \ln(D_m) + \ln(\bar{D}_t) + \ln(p_{ba}) + \ln(p_{bm}) + \ln(p_{bt}) + \ln(D_{am}) + \ln(D_{at}) + \ln(D_{mt}) + \varepsilon_{bamt} \quad (7)$$

Si llamamos:

$$y_{bamt} = \ln(CPUE_{bamt}) \quad (8)$$

$$\alpha_b = \ln(p_b), \beta_a = \ln(D_a), \gamma_m = \ln(D_m), \delta_t = \ln(\bar{D}_t)$$

¹Si bien a lo largo del texto hablaremos del “factor Mes”, está claro que en la práctica podemos usar el bimestre, el trimestre, temporada de pesca, etc. A lo largo del texto usaremos, por lo tanto, el término Mes por una cuestión de comodidad, pero la denominación debiera adaptarse a cada caso de aplicación en particular.

$$(\alpha\beta)_{ba} = \ln(p_{ba}), (\alpha\gamma)_{bm} = \ln(p_{bm}), (\alpha\delta)_{bt} = \ln(p_{bt}) \\ (\beta\gamma)_{am} = \ln(D_{am}), (\beta\delta)_{at} = \ln(D_{at}), (\gamma\delta)_{mt} = \ln(D_{mt})$$

entonces el modelo (8) se escribe como:

$$y_{bamt} = \alpha_b + \beta_a + \gamma_m + \delta_t + (\alpha\beta)_{ba} + (\alpha\gamma)_{bm} + \\ (\alpha\delta)_{bt} + (\beta\gamma)_{am} + (\beta\delta)_{at} + (\gamma\delta)_{mt} + \varepsilon_{bamt} \quad (9)$$

El modelo (9) es denominado modelo de ANOVA multifactorial con interacciones de primer orden.

La ventaja del modelo (9) con respecto al modelo (6) es que el modelo (9) es lineal en los parámetros. En particular, esto hace que la estimación de los parámetros del modelo (9), a partir del método de mínimos cuadrados o de máxima verosimilitud (en caso de suponer distribución normal en el término de error, determinan las mismas soluciones), se reduzca a la solución (generalizada) de un sistema de ecuaciones lineales. Además, en el modelo (9) la suma de cuadrados total de la variable dependiente (logaritmo de la CPUE) se puede descomponer en sumas de cuadrados, asociadas cada una de ellas con los efectos principales correspondientes a cada factor y a los efectos de interacción correspondientes a la combinación de los factores tomados de a dos (interacciones de primer orden), o más. Estas sumas de cuadrados permiten efectuar un análisis de la varianza (ANOVA), para analizar la significación estadística de los efectos de cada factor y de las interacciones, sobre la variabilidad total asociada con la variable dependiente. Descartando los factores e interacciones no significativas se simplifica la descripción del modelo correspondiente al caso en estudio.

Si se quisieran estimar los parámetros del modelo (9) por mínimos cuadrados o máxima verosimilitud, nos encontraríamos con que el modelo está “sobreparametrizado”. Es decir, existen parámetros de más, y cuando se plantean las ecuaciones normales se encuentra que estas ecuaciones no tienen una única solución (existen más incógnitas (= parámetros) que ecuaciones independientes). Esto lleva a la necesidad de tener que “reparametrizar” el modelo (9), lo que significa definir nuevos parámetros en función de los parámetros originales del

modelo. En los modelos de ANOVA, la “reparametrización” se obtiene incluyendo en (9) un término constante, μ , e imponiendo una serie de restricciones sobre los parámetros del modelo. Si n^b , n^a , n^m y n^t corresponden al número de niveles de los factores: Barco, Área, Mes y Año, respectivamente, el modelo y las restricciones sobre los parámetros que más comúnmente se imponen, son:

$$y_{bamt} = \mu + \alpha_b + \beta_a + \gamma_m + \delta_t + (\alpha\beta)_{ba} + (\alpha\gamma)_{bm} \\ + (\alpha\delta)_{bt} + (\beta\gamma)_{am} + (\beta\delta)_{at} + (\gamma\delta)_{mt} + \varepsilon_{bamt} \quad (10)$$

$$\sum_b \alpha_b = \sum_a \beta_a = \sum_m \gamma_m = \sum_t \delta_t = 0$$

$$\sum_b (\alpha\beta)_{ba} = \sum_a (\alpha\beta)_{ba} = 0 \text{ para } 1 \leq a \leq n_a, 1 \leq b \leq n_b$$

$$\sum_b (\alpha\gamma)_{bm} = \sum_m (\alpha\gamma)_{bm} = 0 \text{ para } 1 \leq m \leq n_m, 1 \leq b \leq n_b$$

(11)

$$\sum_b (\alpha\delta)_{bt} = \sum_t (\alpha\delta)_{bt} = 0 \text{ para } 1 \leq t \leq n_t, 1 \leq b \leq n_b$$

$$\sum_a (\beta\gamma)_{am} = \sum_m (\beta\gamma)_{am} = 0 \text{ para } 1 \leq m \leq n_m, 1 \leq a \leq n_a$$

$$\sum_a (\beta\delta)_{at} = \sum_t (\beta\delta)_{at} = 0 \text{ para } 1 \leq t \leq n_t, 1 \leq a \leq n_a$$

$$\sum_m (\gamma\delta)_{mt} = \sum_t (\gamma\delta)_{mt} = 0 \text{ para } 1 \leq t \leq n_t, 1 \leq m \leq n_m$$

Las ecuaciones (11) permiten completar el sistema de ecuaciones normales (número de incógnitas = número de parámetros = número de ecuaciones) y, de esta forma, se obtiene una única solución.

El modelo (10), con las restricciones (11), se considera que está “parametrizado” con la parametrización Sigma-restrictiva. Bajo esta parametrización μ representa la media global de los logaritmos de las CPUE.

Modelo Lineal General

El modelo (10) es un caso particular del denominado Modelo Lineal General, que se

escribe como:

$$Y = X\theta + \varepsilon \quad (12)$$

donde Y es el vector de valores observados de la variable dependiente (logaritmo de la CPUE), $\theta = (\mu, \alpha_1, \alpha_2, \dots, \beta_1, \beta_2, \dots, (\alpha\beta)_{11}, (\alpha\beta)_{12}, \dots)^t$ es el vector de parámetros (el supraíndice t indica el vector traspuesto), ε es el vector de términos de error y X es una matriz con 0's y 1's, denominada matriz de diseño.

Las ecuaciones normales se expresan:

$$(X^t X)\theta = X^t Y \quad (13)$$

donde X^t es la matriz traspuesta de la matriz X (las filas de X^t son las columnas de X).

Si no se imponen restricciones sobre los parámetros que conforman el vector θ , el sistema (13) tiene infinitas soluciones (la matriz $X^t X$ no tiene inversa). Si por el contrario, se imponen las restricciones (11), completando el sistema de ecuaciones normales, entonces existe una única solución.

Parametrizaciones

Según se expresó, la parametrización Sigma-restrictiva es frecuentemente usada en el contexto del análisis de la varianza. Por su parte, en el ámbito de la biología pesquera, al considerar el problema de la estimación de poderes de pesca relativos, se suele usar, teniendo en cuenta su simpleza intuitiva, la denominada parametrización con niveles estándar. En este tipo de parametrización, los parámetros del modelo son referidos a los parámetros correspondientes a: una categoría de barco (categoría de eslora, HP, etc.), un área, un mes y un año, considerados como estándar. Para esta parametrización, por ejemplo, el parámetro μ representa la media de los logaritmos de las CPUE, para la categoría de barcos, el área, el mes y el año estándares. Esta parametrización puede expresarse, como la Sigma-restrictiva,

como un conjunto de restricciones sobre los parámetros del modelo original (8).

Es necesario remarcar, que el análisis y los resultados obtenidos en este trabajo son independientes del tipo de parametrización considerada.

Estimación de índices de abundancia relativa

Es una práctica muy generalizada, normalizar el esfuerzo (previa estimación de los poderes de pesca relativos) y así construir series de CPUE normalizadas, para utilizarlas en la construcción de índices de abundancia (anuales o mensuales, bimestrales, etc.). Estos índices se consideran proporcionales a las abundancias reales.

Esta práctica adolece de por lo menos dos problemas:

- Es muy trabajosa y requiere efectuar cálculos complicados *a posteriori* de haber efectuado la estimación de los parámetros del modelo (9).
- No es posible obtener varianzas e intervalos de confianza para las CPUE normalizadas y se carece por lo tanto, de una medida de la precisión de nuestras estimaciones.

Si observamos el modelo (10), el parámetro asociado con las fluctuaciones anuales puras es δ_t ; los demás parámetros, o no, dependen de t o consisten en términos de interacción con los factores Barco, Área o Mes. Por lo tanto, consideraremos, en principio, en la escala logarítmica, el índice:

$$\ln(I_t) = \mu + \delta_t \quad (14)$$

y en la escala original:

$$I_t = e^{\mu + \delta_t} \quad (15)$$

Si el término de error del modelo (10) se distribuye normalmente y con la misma varianza σ_ε^2 , dentro de cada celda definida a partir de la combinación de los niveles de los factores, se puede observar (considerando las relaciones entre los

valores medios en la escala logarítmica y la escala original en una distribución lognormal) que:

$$E \left[\prod_b \prod_a \prod_m \overline{CPUE}_{bamt}^G \right]^{1/n_b n_a n_m} = e^{\mu + \delta_t + \frac{1}{2} \frac{\sigma_\varepsilon^2}{n_b n_a n_m \bar{n}_t^A}} \quad (16)$$

donde:

$$\bar{n}_t^A = \left[\frac{\sum_b \sum_a \sum_m \frac{1}{n_{bamt}}}{n_b n_a n_m} \right]^{-1}$$

la media armónica del número de observaciones por celda (Barco, Área, Mes) para el año t

y

$$\overline{CPUE}_{bamt}^G = \left[\prod_k CPUE_{bamt k} \right]^{1/n_{bamt}}$$

la media geométrica de los valores de CPUE observados (k indica la observación k -ésima) en la celda (Barco, Área, Mes) para el año t.

Desde el punto de vista práctico, el término

$$\frac{1}{2} \frac{\sigma_\varepsilon^2}{n_b n_a n_m \bar{n}_t^A}$$

puede desprejarse y entonces podemos considerar que:

$$E \left[\prod_b \prod_a \prod_m \overline{CPUE}_{bamt}^G \right]^{1/n_b n_a n_m} = e^{\mu + \delta_t} \quad (17)$$

De esta forma, el índice $I_t = e^{\mu + \delta_t}$ debe ser interpretado en función de la CPUE, como el valor esperado de la media geométrica de las medias geométricas de los valores de CPUE observados en las celdas (Barco, Área, Mes) para el año t. La interpretación del índice puede parecer compleja, pero no es sino el valor esperado de una media geométrica, más adecuada como medida de tendencia central que la media aritmética, cuando se consideran variables con distribuciones asimétri-

cas (como es el caso de las CPUE).

Con el supuesto de la distribución normal de los errores, por la propiedad de invarianza de los estimadores de máxima verosimilitud (Mood y Graybill, 1963), tenemos que, el estimador de máxima verosimilitud del índice I_t está dado por:

$$\hat{I}_t = e^{\hat{\mu} + \hat{\delta}_t} \quad (18)$$

donde $\hat{\mu}$ y $\hat{\delta}_t$ los estimadores de máxima verosimilitud (o de mínimos cuadrados) de μ y δ_t , respectivamente.

Nuevamente con el supuesto de normalidad de la distribución de los errores, los estimadores $\hat{\mu}$ y $\hat{\delta}_t$ tendrán una distribución normal (pues los estimadores de mínimos cuadrados, bajo el supuesto de normalidad del término de error del modelo lineal, tienen una distribución normal). De esta forma $\hat{\mu} + \hat{\delta}_t$ tendrá también una distribución normal y, por lo tanto, el índice $\hat{I}_t = e^{\hat{\mu} + \hat{\delta}_t}$ será una variable lognormal, con media y varianza según Kendall y Stuart (1967):

$$E(\hat{I}_t) = e^{\mu + \delta_t + \frac{1}{2} \sigma_{\hat{\mu} + \hat{\delta}_t}^2} \quad (19)$$

$$\sigma^2(\hat{I}_t) = e^{2(\mu + \delta_t) + \sigma_{\hat{\mu} + \hat{\delta}_t}^2} (e^{\sigma_{\hat{\mu} + \hat{\delta}_t}^2} - 1) \quad (20)$$

La ecuación (19) muestra que el estimador \hat{I}_t es sesgado. En caso de conocer $\sigma_{\hat{\mu} + \hat{\delta}_t}^2$ exactamente, bastaría con definir:

$$\tilde{I}_t = \hat{I}_t e^{-\frac{1}{2} \sigma_{\hat{\mu} + \hat{\delta}_t}^2} \quad (21)$$

para obtener un estimador insesgado. El hecho es que, como sólo tenemos una estimación de $\sigma_{\hat{\mu} + \hat{\delta}_t}^2$, entonces, el estimador que obtendremos considerando (21) y reemplazando $\sigma_{\hat{\mu} + \hat{\delta}_t}^2$ por su estimación $\hat{\sigma}_{\hat{\mu} + \hat{\delta}_t}^2$, aún será sesgado. La cuestión es que, para muestras pequeñas, no es claro cuál es la ganancia de usar (21), reemplazando $\sigma_{\hat{\mu} + \hat{\delta}_t}^2$ por $\hat{\sigma}_{\hat{\mu} + \hat{\delta}_t}^2$, teniendo en cuenta la incertidumbre asociada con $\hat{\sigma}_{\hat{\mu} + \hat{\delta}_t}^2$. Por su parte, para muestras grandes, teniendo en cuenta que los estimadores de máxi-

ma verosimilitud son consistentes en error medio cuadrático (Mood y Graybill, 1963), tenemos que $\sigma_{\hat{\mu}+\hat{\delta}_t}^2 \rightarrow 0$ cuando $n \rightarrow \infty$ y, por lo tanto, para muestras abundantes la corrección planteada en (21) será despreciable. De esta forma, resulta preferible seguir considerando el índice $\hat{I}_t = e^{\hat{\mu}+\hat{\delta}_t}$, que es asintóticamente insesgado.

El correspondiente estimador de la varianza de \hat{I}_t , está dado, teniendo en cuenta (20), por:

$$\hat{\sigma}^2(\hat{I}_t) = e^{2(\hat{\mu}+\hat{\delta}_t)} + \sigma_{\hat{\mu}+\hat{\delta}_t}^2 (e^{\hat{\mu}+\hat{\delta}_t} - 1) \quad (22)$$

donde: $\hat{\sigma}_{\hat{\mu}+\hat{\delta}_t}^2 = \hat{\sigma}_{\hat{\mu}}^2 + \hat{\sigma}_{\hat{\delta}_t}^2 + 2 \text{cov}(\hat{\mu}, \hat{\delta}_t)$, donde $\hat{\sigma}_{\hat{\mu}}^2$, $\hat{\sigma}_{\hat{\delta}_t}^2$ y $\text{cov}(\hat{\mu}, \hat{\delta}_t)$ son las estimaciones de las varianzas de $\hat{\mu}$ y $\hat{\delta}_t$ y la covarianza entre $\hat{\mu}$ y $\hat{\delta}_t$, respectivamente.

El estimador (22) es válido para muestras de cualquier tamaño, si bien su validez depende de asumir la normalidad estricta de la distribución del término de error del modelo lineal. Otra aproximación, válida para muestras grandes, pero que no depende de las características distribucionales del término de error, está dada por:

$$\hat{\sigma}_{\hat{I}_t}^2 = (\hat{I}_t)^2 \left[\hat{\sigma}_{\hat{\mu}}^2 + \hat{\sigma}_{\hat{\delta}_t}^2 + 2 \text{cov}(\hat{\mu}, \hat{\delta}_t) \right] \quad (23)$$

Obtenida a partir del método delta no paramétrico (Efron, 1983).

El intervalo de confianza para el índice I_t , con coeficiente de confianza $1-\alpha$, está dado por:

$$\begin{aligned} \text{Límite inferior} &= e^{\hat{\mu}+\hat{\delta}_t - t(v, \alpha/2) \hat{\sigma}_{\hat{\mu}+\hat{\delta}_t}} \\ \text{y Límite superior} &= e^{\hat{\mu}+\hat{\delta}_t + t(v, \alpha/2) \hat{\sigma}_{\hat{\mu}+\hat{\delta}_t}} \end{aligned} \quad (24)$$

donde v el número de grados de libertad correspondiente al cuadrado medio de error (igual al número total de observaciones, menos el número total de parámetros estimados en el modelo (10) y es, además:

$$\hat{\sigma}_{\hat{\mu}+\hat{\delta}_t} = \sqrt{\hat{\sigma}_{\hat{\mu}}^2 + \hat{\sigma}_{\hat{\delta}_t}^2 + 2 \text{cov}(\hat{\mu}, \hat{\delta}_t)} \quad (25)$$

Bajo el supuesto de normalidad en la distribución del término de error, el intervalo (24) es

exacto. En caso contrario, es una aproximación para muestras grandes.

Efecto de las interacciones que incluyen el factor Año sobre los índices de abundancia anuales

A continuación, mostraremos que la existencia de interacciones que incluyan el factor Año, en los modelos (9) y (10), determina que el índice (15) deje de ser proporcional a la abundancia media.

Para establecer el resultado necesitamos modificar en parte la notación considerada hasta el momento. Si bien en los modelos (9) y (10) hemos usado la misma notación para representar los parámetros de cada modelo, en realidad hay que observar que al reparametrizar el modelo (9), imponiendo las restricciones (11), los parámetros se modifican y dejan de tener la misma interpretación. De esta forma, para distinguir los parámetros de cada modelo, dejaremos el modelo (9) con la notación utilizada hasta ahora y escribiremos el modelo (10) y las restricciones (11), en la forma:

$$y_{bamt} = \mu' + \alpha'_b + \beta'_a + \gamma'_m + \delta'_t + (\alpha\beta)'_{ba} + (\alpha\gamma)'_{bm} + (\alpha\delta)'_{bt} + (\beta\gamma)'_{am} + (\beta\delta)'_{at} + (\gamma\delta)'_{mt} + \varepsilon_{bamt} \quad (26)$$

donde ahora, los nuevos parámetros satisfacen las restricciones:

$$\sum_b \alpha'_b = \sum_a \beta'_a = \sum_m \gamma'_m = \sum_t \delta'_t = 0 \quad (27)$$

$$\sum_b (\alpha\beta)'_{ba} = \sum_a (\alpha\beta)'_{ba} = 0 \quad \text{para } 1 \leq a \leq n_a, 1 \leq b \leq n_b$$

$$\sum_b (\alpha\gamma)'_{bm} = \sum_m (\alpha\gamma)'_{bm} = 0 \quad \text{para } 1 \leq m \leq n_m, 1 \leq b \leq n_b$$

$$\sum_b (\alpha\delta)'_{bt} = \sum_t (\alpha\delta)'_{bt} = 0 \quad \text{para } 1 \leq t \leq n_t, 1 \leq b \leq n_b$$

$$\sum_a (\beta\gamma)'_{am} = \sum_m (\beta\gamma)'_{am} = 0 \quad \text{para } 1 \leq m \leq n_m, 1 \leq a \leq n_a$$

$$\sum_a (\beta\delta)'_{at} = \sum_t (\beta\delta)'_{at} = 0 \quad \text{para } 1 \leq t \leq n_t, 1 \leq a \leq n_a$$

$$\sum_m (\gamma\delta)'_{mt} = \sum_t (\gamma\delta)'_{mt} = 0 \quad \text{para } 1 \leq t \leq n_t, 1 \leq m \leq n_m$$

En (26) y (27) se empleó la tilde para diferenciar los parámetros del modelo parametrizado, de los parámetros del modelo sin “reparametrización” presentado en (9).

Si la esperanza del término de error es igual a cero, esto es, si $E(\varepsilon_{bamt})=0$, entonces, la parte paramétrica de (9) y (26) deben coincidir. De esta forma, se debe satisfacer que:

$$\begin{aligned} \mu' + \alpha'_b + \beta'_a + \gamma'_m + \delta'_t + (\alpha\beta)'_{ba} + (\alpha\gamma)'_{bm} + (\alpha\delta)'_{bt} + & (28) \\ (\beta\gamma)'_{am} + (\beta\delta)'_{at} + (\gamma\delta)'_{mt} \\ = \alpha_b + \beta_a + \gamma_m + \delta_t + (\alpha\beta)_{ba} + (\alpha\gamma)_{bm} + (\alpha\delta)_{bt} + & \\ (\beta\gamma)_{am} + (\beta\delta)_{at} + (\gamma\delta)_{mt} & \end{aligned}$$

Si sumamos y promediamos consecutivamente, sobre los diferentes niveles de los factores: Barco, Área y Mes ambos miembros de (28) y, al efectuar las sumas tenemos en cuenta las restricciones (27) que pesan sobre los parámetros del primer miembro, obtenemos (ordenando los términos del segundo miembro):

$$\begin{aligned} \mu' + \delta'_t = \delta_t + \bar{\alpha} + \bar{\beta} + \bar{\gamma} + (\bar{\alpha\beta})_{..} + (\bar{\alpha\gamma})_{..} + (\bar{\beta\gamma})_{..} + & (29) \\ (\bar{\alpha\delta})_{.t} + (\bar{\beta\delta})_{.t} + (\bar{\gamma\delta})_{.t} & \end{aligned}$$

donde, por ejemplo:

$$\bar{\alpha} = \frac{\sum_{b=1}^{n_b} \alpha_b}{n_b}, \quad (\bar{\alpha\beta})_{..} = \frac{\sum_{b=1}^{n_b} \sum_{a=1}^{n_a} (\alpha\beta)_{ba}}{n_b n_a} \quad \text{y} \quad (30)$$

$$(\bar{\alpha\delta})_{.t} = \frac{\sum_{b=1}^{n_b} (\alpha\delta)_{bt}}{n_b}$$

Los demás términos se definen en forma análoga.

Los términos, desde el segundo hasta el séptimo, del segundo miembro de (29), son constantes y pueden resumirse en un único término al que denominaremos c_t . Los términos octavo, noveno y décimo son funciones de las interacciones Barco x Año, Área x Año y Mes x Año y podemos resumirlos en un término al que denominaremos ϕ_t ,

dependiente del año. De esta forma, la ecuación (29), la podemos escribir como:

$$\mu' + \delta'_t = \delta_t + c_t + \phi_t \quad (31)$$

El índice I_t , cuya estimación –en el texto principal– hemos denominado \hat{I}_t , está dado (con el cambio de notación que hemos aclarado) por:

$$I_t = e^{\mu' + \delta'_t} \quad (32)$$

Teniendo en cuenta (31) y (32), tenemos que:

$$I_t = e^{\delta_t + c_t + \phi_t} = c \psi_t \bar{D}_t \quad (33)$$

donde $c = e^{c_t}$ y $\psi_t = e^{\phi_t}$

De acuerdo con (33), vemos que, ante la presencia de cualquiera de las interacciones del tipo Barco x Año, Área x Año o Mes x Año, representadas por ψ_t , el índice deja de ser proporcional y, que cuando las interacciones no existen, $\psi_t = 1$ (pues $\phi_t = 0$), el índice es proporcional a la abundancia anual media \bar{D}_t , a través de la constante c .

Efecto de no incluir interacciones en el modelo sobre los índices de abundancia anuales cuando estas interacciones existen en la realidad

Si se supone que en la realidad existen interacciones de primer orden, pero que en el modelo no las incluimos, de forma tal que el modelo de trabajo se haya especificado a partir de:

$$y_{bamt} = \mu'' + \alpha''_b + \beta''_a + \gamma''_m + \delta''_t + \xi_{bamt} \quad (34)$$

Suponiendo además que se considera una “parametrización” Sigma-restrictiva, de forma que los parámetros en (34) satisfacen las siguientes restricciones:

$$\sum_b \alpha''_b = \sum_a \beta''_a = \sum_m \gamma''_m = \sum_t \delta''_t = 0 \quad (35)$$

En general, a no ser que las interacciones no incluidas en (34) se compensen de alguna forma, tendremos que:

$$E(y_{bamt}) - (\mu'' + \alpha_b'' + \beta_a'' + \gamma_m'' + \delta_t'') = \rho_{bamt} \quad (36)$$

donde ρ_{bamt} es función de las interacciones no incluidas en el modelo y equivalente al sesgo del modelo al aproximar el valor esperado de la variable (logaritmo de la CPUE) en la celda (b, a, m, t) .

Si suponemos $E(\varepsilon_{bamt}) = 0$, entonces según (9), (34) y (36), tenemos que:

$$\begin{aligned} \mu'' + \alpha_b'' + \beta_a'' + \gamma_m'' + \delta_t'' + \rho_{bamt} & \quad (37) \\ = \alpha_b + \beta_a + \gamma_m + \delta_t + (\alpha\beta)_{ba} + (\alpha\gamma)_{bm} + (\alpha\delta)_{bt} + & \\ (\beta\gamma)_{am} + (\beta\delta)_{at} + (\gamma\delta)_{mt} & \end{aligned}$$

De esta forma, si sumamos y promediamos consecutivamente, sobre los diferentes niveles de los factores: Barco, Área y Mes ambos miembros de (37) y al efectuar las sumas tenemos en cuenta las restricciones (35) impuestas sobre los parámetros del primer miembro, obtenemos (ordenando los términos del segundo miembro):

$$\begin{aligned} \mu'' + \delta_t'' + \bar{\rho}_{...t} = \delta_t + \bar{\alpha} + \bar{\beta} + \bar{\gamma} + (\bar{\alpha}\bar{\beta})_{..} + (\bar{\alpha}\bar{\gamma})_{..} + & \quad (38) \\ (\bar{\beta}\bar{\gamma})_{..} + (\bar{\alpha}\bar{\delta})_{.t} + (\bar{\beta}\bar{\delta})_{.t} + (\bar{\gamma}\bar{\delta})_{.t} & \end{aligned}$$

donde, en particular:

$$\bar{\rho}_{...t} = \frac{\sum_{b=1}^{n_b} \sum_{a=1}^{n_a} \sum_{m=1}^{n_m} \rho_{bamt}}{n_b n_a n_m} \quad (39)$$

los otros términos se han definido en (30).

Los términos, desde el segundo hasta el séptimo, del segundo miembro de (38), son constantes y pueden resumirse en un único término al que denominaremos c_7 . Los términos octavo, noveno y décimo son funciones de las interacciones Barco x Año, Área x Año y Mes x Año y podemos resumirlos en un término al que denominaremos ϕ_t , dependiente del año. De esta forma, la ecuación (38), podemos escribirla de esta forma:

$$\mu'' + \delta_t'' = \delta_t + c_1 + \phi_t - \bar{\rho}_{...t} \quad (40)$$

Tenemos entonces que:

$$I_t = e^{\mu'' + \delta_t''} = c \psi_t \bar{D}_t \quad (41)$$

donde ahora $\psi_t = e^{\phi_t - \bar{\rho}_{...t}}$

Vemos entonces que el procedimiento de no incluir las interacciones con el año en el modelo de trabajo, no solo no resuelve el problema de las interacciones con el factor Año, sino que agrava la situación, ya que, aunque $\phi_t = 0$, todavía ψ_t será función del tiempo, debido a la presencia de $\rho_{...t}$. Por otra parte, aun en el caso en que $E(\xi_{bamt}) = 0$, esto es, incluso cuando el modelo (34) permita obtener estimaciones insesgadas, el problema de las interacciones con el factor Año (si estas interacciones existen), no se resuelve por no incluir las interacciones en el modelo de trabajo.

Índices de abundancia ponderados por las áreas

Teniendo en cuenta la expresión (17), debe observarse que el tipo de promedio que involucra la definición del índice considerado, es no ponderado. En principio no debería generar dificultades con respecto a los factores Barco o Mes (a no ser que, en el último caso, en lugar del mes, bimestre, trimestre, etc., se consideren diferentes temporadas de pesca de desigual extensión en el tiempo). Pero, si las diferentes áreas (superficies en km² o mn²) involucradas en la definición del factor Área fueran muy distintas, puede construirse un índice ponderado.

Sea entonces A_a , el área correspondiente al nivel a (la región rotulada como a) del factor área y sea:

$$A_T = \sum_{a=1}^{n_a} A_a \quad (42)$$

el área total.

Definimos entonces:

$$w_a = \frac{A_a}{A_T} \quad (43)$$

En caso de existir interacción área x año, llamemos:

$$\hat{I}_{at} = e^{i\mu + \beta_a + \delta_t + (\overline{\beta\delta})_{at}} \quad (44)$$

donde con $(\overline{\beta\delta})_{at}$ indicamos el estimador de $(\beta\delta)_{at}$.

En caso de no existir interacción área x año:

$$\hat{I}_{at} = e^{i\mu + \beta_a + \delta_t} \quad (45)$$

El estimador de la abundancia anual, correspondiente al año t, ponderado por las áreas, lo definimos como:

$$\tilde{I}_t^w = \sum_{a=1}^{n_a} w_a \hat{I}_{at} \quad (46)$$

La varianza de \tilde{I}_t^w está dada por:

$$\sigma^2(\tilde{I}_t^w) = \sum_{a=1}^{n_a} w_a^2 \sigma^2(\hat{I}_{at}) + \sum_{a=1}^{n_a} \sum_{a'=1}^{n_a} w_a w_{a'} \text{cov}(\hat{I}_{at}, \hat{I}_{a't}) \quad (47)$$

donde $\text{cov}(\hat{I}_{at}, \hat{I}_{a't})$, la covarianza entre \hat{I}_{at} y $\hat{I}_{a't}$ (a y a' son niveles distintos dentro de los n_a niveles del factor Área).

El correspondiente estimador de (47) está dado por:

$$\hat{\sigma}^2(\tilde{I}_t^w) = \sum_{a=1}^{n_a} w_a^2 \hat{\sigma}^2(\hat{I}_{at}) + \sum_{a=1}^{n_a} \sum_{a'=1}^{n_a} w_a w_{a'} \hat{\text{cov}}(\hat{I}_{at}, \hat{I}_{a't}) \quad (48)$$

Para calcular $\hat{\sigma}_{\hat{I}_{at}}^2$, consideraremos el estimador delta no paramétrico, dado por:

$$\hat{\sigma}^2(\hat{I}_{at}) = (\hat{I}_{at})^2 \hat{\sigma}_{\mu + \beta_a + \delta_t + (\overline{\beta\delta})_{at}}^2 \quad (49)$$

donde:

$$\begin{aligned} \hat{\sigma}_{\mu + \beta_a + \delta_t + (\overline{\beta\delta})_{at}}^2 &= \hat{\sigma}_{\mu}^2 + \hat{\sigma}_{\beta_a}^2 + \hat{\sigma}_{\delta_t}^2 + \hat{\sigma}_{(\overline{\beta\delta})_{at}}^2 \quad (50) \\ &+ 2(\text{côv}(\mu, \beta_a) + \text{côv}(\mu, \delta_t) \\ &+ \text{côv}(\mu, (\overline{\beta\delta})_{at}) + \text{côv}(\beta_a, \delta_t) \\ &+ \text{côv}(\beta_a, (\overline{\beta\delta})_{at}) + \text{côv}(\delta_t, (\overline{\beta\delta})_{at})) \end{aligned}$$

Por otra parte, el estimador delta no paramétrico de los términos de covarianza está dado por:

$$\begin{aligned} \text{côv}(\hat{I}_{at}, \hat{I}_{a't}) &= (\hat{I}_{at} \hat{I}_{a't}) \left[\begin{aligned} &\hat{\sigma}_{\mu}^2 + \hat{\sigma}_{\delta_t}^2 + 2 \text{côv}(\mu, \delta_t) \\ &+ \text{côv}(\mu, \beta_a) + \text{côv}(\mu, (\overline{\beta\delta})_{at}) \\ &+ \text{côv}(\mu, \beta_a) + \text{côv}(\mu, (\overline{\beta\delta})_{at}) \\ &+ \text{côv}(\beta_a, \beta_a) + \text{côv}(\beta_a, \delta_t) \\ &+ \text{côv}(\beta_a, (\overline{\beta\delta})_{at}) + \text{côv}(\delta_t, \beta_a) \\ &+ \text{côv}(\delta_t, (\overline{\beta\delta})_{at}) + \text{côv}(\delta_t, (\overline{\beta\delta})_{a't}) \\ &+ \text{côv}((\overline{\beta\delta})_{at}, \beta_a) \\ &+ \text{côv}((\overline{\beta\delta})_{at}, (\overline{\beta\delta})_{a't}) \end{aligned} \right] \quad (51) \end{aligned}$$

En el caso de que no se considere el término de interacción Área x Año (por no ser significativo o por aportar muy poco a la explicación de varianza), en las fórmulas (49), (50) y (51), los términos que incluyan las interacciones $(\beta\delta)_{at}$ y $(\overline{\beta\delta})_{at}$ deben ser suprimidos.

Efecto de las interacciones sobre el índice de abundancia ponderado por las áreas

Teniendo en cuenta (26) con las restricciones (27) y considerando $E(\varepsilon_{bamt}) = 0$, podemos nuevamente establecer (28). Promediando entonces los dos miembros de (28) sobre los diferentes niveles de los factores Barco y Mes y teniendo en cuenta las restricciones (27), que pesan sobre los parámetros del primer miembro al efectuar las sumas, al ordenar los términos, obtenemos:

$$\begin{aligned} \mu' + \beta_a' + \delta_t' + (\beta\delta)_{at}' &= \bar{\alpha} + \bar{\gamma} + (\overline{\alpha\gamma})_{..} + \beta_a + \quad (52) \\ \delta_t + (\beta\delta)_{at} + (\overline{\alpha\beta})_{.a} &+ (\overline{\beta\gamma})_{.a} + (\overline{\alpha\delta})_{.t} + (\overline{\gamma\delta})_{.t} \end{aligned}$$

Si llamamos:

$$\begin{aligned} c &= \bar{\alpha} + \bar{\gamma} + (\overline{\alpha\gamma})_{..} \quad , \quad \phi_a = (\overline{\alpha\beta})_{.a} + (\overline{\beta\gamma})_{.a} \quad , \\ \phi_t &= (\overline{\alpha\delta})_{.t} + (\overline{\gamma\delta})_{.t} \end{aligned}$$

con definiciones de los términos como en (30)

entonces (52) se escribe como:

$$\mu' + \beta'_a + \delta'_t + (\beta\delta)'_{at} = c + \beta_a + \delta_t + (\beta\delta)_{at} + \phi_a + \phi_t \quad (53)$$

Teniendo en cuenta la definición de los parámetros del segundo miembro de (53) dada en (8), tenemos que:

$$I_{at} = e^{\mu + \beta'_a + \delta'_t + (\beta\delta)'_{at}} = c \psi_a \psi_{1t} D_a \bar{D}_t D_{at} \quad (54)$$

con definiciones obvias para c , ψ_a y ψ_{1t} .

Considerando (5) y promediando sobre los diferentes niveles del factor Mes, obtenemos la abundancia media en el área a , correspondiente al año t :

$$\bar{D}_{at} = \frac{\sum_{m=1}^{n_m} D_{amt}}{n_m} = D_a D_{at} \bar{D}_t \frac{\sum_{m=1}^{n_m} D_m D_{am} D_{mt}}{n_m} \quad (55)$$

De esta forma, a partir de (55), obtenemos que:

$$D_a D_{at} \bar{D}_t = \varphi_{at} \bar{D}_{at} \quad (56)$$

$$\text{donde: } \varphi_{at} = \left(\frac{\sum_{m=1}^{n_m} D_m D_{am} D_{mt}}{n_m} \right)^{-1}$$

Entonces, a partir de (54), tenemos:

$$I_{at} = c \psi_a \psi_{1t} \varphi_{at} \bar{D}_{at} \quad (57)$$

Por otra parte, el índice ponderado por las áreas está dado por:

$$I_t^w = \sum_{a=1}^{n_a} w_a I_{at} = c \psi_{1t} \sum_{a=1}^{n_a} w_a \psi_a \varphi_{at} \bar{D}_{at} \quad (58)$$

Definiendo entonces $\psi_{2t} = \sum_{a=1}^{n_a} w_a \psi_a \varphi_{at}$, tenemos que:

$$I_t^w = c \psi_{1t} \psi_{2t} \sum_{a=1}^{n_a} \pi_{at} \bar{D}_{at} \quad (59)$$

donde:

$$\pi_{at} = \frac{w_a \psi_a \varphi_{at}}{\sum_{a'=1}^{n_a} w_{a'} \psi_{a'} \varphi_{a't}}$$

Por último, si definimos:

$$\bar{D}_t^\pi = \sum_{a=1}^{n_a} \pi_{at} \bar{D}_{at} \quad (60)$$

que representa una abundancia media anual ponderada, ya que:

$$\sum_{a=1}^{n_a} \pi_{at} = 1$$

tenemos que:

$$I_t^w = c \psi_{1t} \psi_{2t} \bar{D}_t^\pi \quad (61)$$

El problema con \bar{D}_t^π es que, en ausencia de supuestos, los factores de ponderación π_{at} involucrados en su definición, dependen del año, que determina que los valores de \bar{D}_t^π , correspondientes a años diferentes, no sean comparables entre sí.

Se puede observar, teniendo en cuenta simplemente las definiciones de los términos involucrados, que en caso de que no existan las interacciones Barco x Año y Mes x Año, entonces: $\psi_{1t}=1$ y $\psi_{2t}=c_0$, donde c_0 es una constante. Además, se puede observar también que, en ausencia de estas interacciones, $\pi_{at}=\pi_a$, es decir, los factores de ponderación involucrados en la definición de \bar{D}_t^π ya no dependen del año, lo que confiere consistencia a \bar{D}_t^π como medida de abundancia media anual.

Por su parte, si además no existen las interacciones Barco x Área y Mes x Área, entonces, directamente es $\pi_{at}=w_a$ y de esta forma:

$$\bar{D}_t^\pi = \sum_{a=1}^{n_a} w_a \bar{D}_{at} \quad (62)$$

Entonces, la densidad media anual ponderada, \bar{D}_t^π es equivalente a la densidad media anual estratificada y ponderada por las áreas, caso que es

intuitivamente más simple y natural que la abundancia media anual definida en (60), ponderada por los pesos π_{at} que son función de los términos de interacción.

Ejemplo de aplicación

A los efectos de mostrar en la práctica el desarrollo teórico, se emplearon dos series de datos de captura en kilogramos y esfuerzo medidos en tiempo de arrastre expresado en horas, correspondientes a las especies corvina (*Micropogonias furnieri*) y pescadilla de red (*Cynoscion guatucupa*), en los períodos 1989-2003 y 1992-2003, respectivamente. Los datos básicos fueron obtenidos de los registros de la Dirección Nacional de Pesca y Acuicultura y provistos por el Centro de Cómputos del INIDEP. La base de datos consistió de 8.423 registros para corvina y 5.655 para pescadilla de red, con los cuales se calcularon los valores de CPUE por viaje, expresados en kg por hora de arrastre (Carozza *et al.*, 2004; Ruarte y Perrotta, en prensa).

Los factores empleados en el MLG fueron Barco, Área, Mes y Año y se definieron para cada especie (Carozza *et al.*, 2004; Ruarte y Perrotta, en prensa) como:

Corvina

Barco (3 niveles): pesca a la pareja, sin diferenciar esloras; pesca de arrastre, 6 a 17 m de eslora y pesca de arrastre >17 a 25 m de eslora.

Área (3 niveles): El Salado, Uruguay y Mar del Plata.

Mes (4 niveles): trimestre.

Año (14 niveles): 1989-2002.

Pescadilla

Barco (4 niveles): 8-20; >20-33; >33-38,30 y > 38,30.

Área (3 niveles): tres subáreas.

Mes (4 niveles): trimestre.

Año (14 niveles): 1989-2002.

RESULTADOS

Evaluación del efecto de las interacciones con el factor Año. Algunas reglas metodológicas a tener en cuenta en el análisis de los datos

A partir del estudio efectuado en Materiales y métodos, se pueden extraer una serie de procedimientos y pautas para considerar cuando se efectúa el análisis orientado a la construcción de los índices anuales de abundancia, utilizando un MLG.

Previo a su aplicación es necesario discutir críticamente los siguientes puntos:

1. Qué estratos, categorías o tipos de flota deberíamos incluir en el análisis. Analizar la conveniencia de trabajar con un estrato de flota patrón.
2. Necesidad y conveniencia de incluir, o no, una estratificación por áreas.
3. Evaluar el tipo de categorización más adecuada para el factor Mes: meses, bimestres, trimestres, etcétera.

Al analizar los puntos 1, 2 y 3, debe tenerse en cuenta la cuestión referida a la posibilidad de ocurrencia de celdas vacías en el diseño. En algunos casos puede ser necesario, al definir los niveles de algunos de los factores que se puedan incluir en el modelo, considerar categorías más amplias, orientadas a eliminar la profusión de celdas vacías, que puedan comprometer la estimación, en particular, de los efectos de algunas de las posibles interacciones.

Se pueden definir algunas reglas metodológicas que permitan evaluar la importancia de los términos de interacción que incluyan el factor Año y que puedan contribuir a disminuir —en caso de que existan— la influencia negativa de esas interacciones sobre el índice de abundancia:

1. Es importante tener en cuenta la posibilidad de que los términos de interacción, aunque

sean estadísticamente significativos (los no significativos deben eliminarse del modelo), puedan tener un efecto insignificante desde el punto de vista práctico. En general, cuando se analizan las CPUE de una flota, a lo largo de varios años, el número de observaciones puede alcanzar varios miles. Ante un tamaño muestral tan grande, cualquier efecto, aunque su aporte a la explicación de la varianza sea insignificante, puede resultar estadísticamente significativo. Esta cuestión debería ser analizada. Para hacerlo, como regla metodológica, se pueden calcular los coeficientes de determinación (porcentaje de varianza explicada por el modelo), considerando el modelo con y sin las interacciones con el factor Año y determinar la diferencia; esa diferencia corresponderá al porcentaje de varianza explicada por los términos de interacción. De esta forma, en caso de que el aporte de estas interacciones a la explicación de la varianza sea insignificante, podremos considerar los índices como representativos esencialmente de los cambios anuales de abundancia del recurso.

2. Por su parte, sería también importante no incluir en el análisis los estratos, las categorías o los tipos de flota que sepamos de antemano que han registrado ganancia o pérdida de eficiencia a lo largo de los años. En el caso de contar con un estrato de flota de la que podamos asumir que su eficiencia no ha cambiado con el paso de los años, y que haya tenido actividad en todo el período de estudio, podríamos tomarlo como “estrato de flota patrón” y así limitar el análisis, considerando solo los datos provenientes de este grupo de barcos. Este procedimiento eliminará directamente la existencia del factor Barco y, obviamente, la correspondiente interacción con el factor Año, y las dificultades asociadas con esta interacción.

3. Una cuestión que debe tenerse en cuenta es que cuando consideramos el estimador ponderado por las áreas, los problemas generados por la interacción Área x Año, desaparecen.

Este procedimiento podría ser una forma de eliminar la influencia de esta interacción sobre los índices de abundancia.

4. Puede ser útil, también, explorar las consecuencias de considerar diferentes tipos de categorizaciones del factor Mes (mes, bimestre, trimestre, etc.) y evaluar la importancia de la interacción Mes x Año puede ayudar a entender y ponderar su efecto sobre la imagen de las fluctuaciones reales de la abundancia del recurso, proporcionada por el índice.

Ejemplos de aplicación

En la Tabla 1 se muestran los coeficientes de determinación considerando el modelo con y sin las interacciones que incluyen el factor Año.

En el caso de la corvina las interacciones Barco x Año, Área x Año y Mes x Año, solo explican el 5,31% de la varianza de los logaritmos de la CPUE, lo cual representa un 9,37% de la varianza total explicada por el modelo. Por lo tanto, el efecto de estas interacciones sobre los índices de abundancia podría considerarse poco importante. De este modo el índice reflejaría los cambios de abundancia del recurso.

En el caso de la pescadilla, la situación es diferente. Las interacciones que incluyen el factor Año explican el 8,18% de la varianza, que representa un 22,25% de la varianza total explicada por el modelo. De esta forma, comparado con la corvina, podríamos esperar que los índices de abundancia para pescadilla tuvieran un grado mayor de “contaminación”, debido a la presencia de las interacciones Barco x Año, Área x Año y Mes x Año.

Para los dos ejemplos considerados, en el período de estudio, no hay razones para suponer que puede haber habido un cambio importante en las eficiencias de las unidades de pesca (modalidad de pesca, cambios tecnológicos, envejecimiento, etc.) (Ruarte *et al.*, 2000; Carozza *et al.*, 2001; Lasta *et al.*, 2001; Ruarte *et al.*, 2001). De hecho, en la Tabla 1 se puede observar que a la

interacción Barco x Año le corresponde un pequeño porcentaje de la varianza explicada, para ambas especies.

En referencia a la interacción Área x Año, como se ha mostrado, la utilización del estimador ponderado por las áreas puede eliminar el efecto generado por esta interacción sobre los índices de abundancia. En el caso de la pescadilla, en donde los índices de abundancia podrían estar más comprometidos que en el caso de la corvina (Tabla 1), debido al efecto de las interacciones que incluyen el factor Año, es aconsejable el empleo de este procedimiento (índices ponderados).

La elección del trimestre para categorizar el factor Mes estuvo condicionado, para ambas especies, por cuestiones relacionadas con la distribución del recurso y la actividad de la flota comercial (Carozza *et al.*, 2001; Ruarte *et al.*, 2001). Por su parte, el problema de las celdas vacías, frecuente en este tipo de análisis, también influyó en la determinación de dividir el año trimestralmente.

En la Figura 1 se pueden ver las series de los

índices de abundancia construidos a partir del MLG. En el caso de la pescadilla se presenta también el índice ponderado por las áreas.

DISCUSIÓN Y CONCLUSIONES

Con el propósito de construir índices anuales de abundancia a partir de la aplicación de un MLG, es común utilizar el efecto año (Hilborn y Walters, 1992; Large, 1992; Goñi *et al.*, 1999; García-Rodríguez, 2003), que en primera instancia se considera que resume las variaciones anuales *puras* de la abundancia del recurso considerado. Pero, como hemos visto, cuando están presentes interacciones que incluyan el factor Año, las series de abundancia anuales obtenidas a partir de los índices calculados en función de parámetros de estos modelos, pueden mostrar una imagen donde se mezclen cambios correspondientes a la abundancia real del recurso con efectos debidos a estas interacciones.

Tabla 1. Porcentaje de varianza explicada con y sin las interacciones que incluyen el factor Año y contribución de las interacciones para corvina y pescadilla.

Table 1. Percentage of explained variance considering and not considering the interaction that includes the factor Year and contribution interaction for Whitemouth croaker and Stripped weakfish.

	Corvina	Pescadilla
Porcentaje de varianza explicada por el modelo, con todas las interacciones de primer orden (R_0^2)	56,64	36,77
Porcentaje de varianza explicada por el modelo, sin las interacciones que incluyan el factor Año (R_1^2)	51,33	28,59
$R_0^2 - R_1^2$	5,31	8,18
Contribución porcentual de la interacción Barco x Año	2,19	1,02
Contribución porcentual de la interacción Área x Año	0,12	0,87
Contribución porcentual de la interacción Mes x Año	0,64	2,73

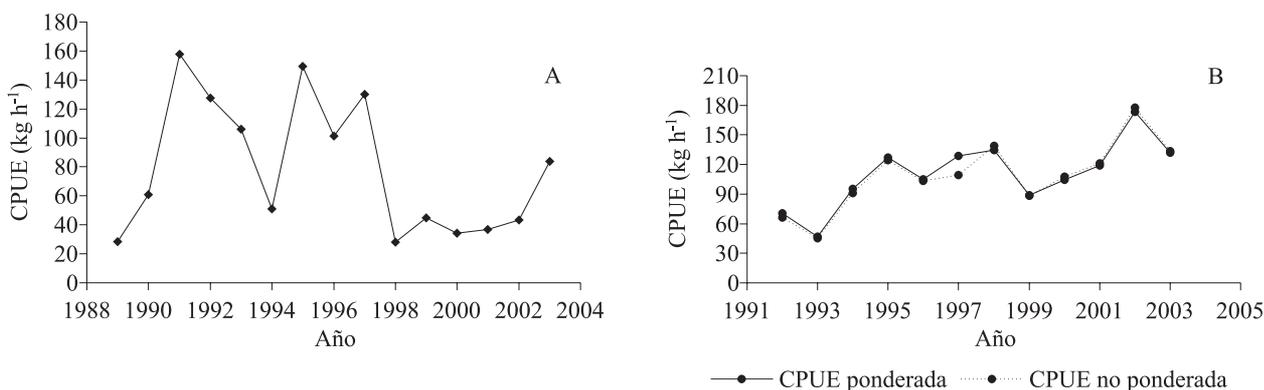


Figura 1. Series de índices de abundancia anual (CPUE). A) Corvina (tomado de Carozza *et al.*, 2004); B) Pescadilla (tomado de Ruarte y Perrotta, en prensa).

Figure 1. Abundance annual index (CPUE) series. A) Whitemouth croaker (from Carozza *et al.*, 2004); B) Stripped weakfish (from Ruarte and Perrotta, in press).

Es decir, supongamos que existiera una interacción entre el factor Barco y el factor Año, lo que sería indicativo de que los poderes de pesca, de al menos uno de los estratos de barcos considerados, habrían variado con el paso de los años. Por ejemplo, si existiera un efecto de deterioro, entonces la eficiencia de los barcos se vería disminuida con el paso del tiempo. Por el contrario, si se hubiesen introducido cambios tecnológicos, los poderes de pesca podrían haberse incrementado. En este último caso se podrían generar incrementos en la CPUE, más allá de los aumentos debidos a cambios reales en la abundancia del recurso. De hecho, en el caso particular en que la abundancia del recurso se mantuviera constante, el incremento en la eficiencia de la flota generará por sí solo una imagen en donde la CPUE se verá incrementada con el paso de los años. En realidad, a no ser que *a priori* se conozcan las eficiencias o poderes de pesca de la flota, no es posible separar estos efectos y entonces la imagen que se obtendrá dependerá siempre de una combinación de incrementos o disminuciones reales en la abundancia del recurso y de incrementos o disminuciones en la eficiencia de la flota. Matemáticamente se puede ver que cuando estas interacciones están presentes, la serie de índices de abundancia deja de ser pro-

porcional a la abundancia anual media y pierde, por lo tanto, utilidad como índice de abundancia relativa.

Autores como Robson (1966) y Gavaris (1980) plantean directamente el modelo sin términos de interacción, mientras que Large (1992) utiliza el procedimiento de eliminar del modelo de trabajo las interacciones que incluyen el factor Año, por no explicar un porcentaje importante de la varianza. Horn (2002, 2003) no considera en el modelo las interacciones que incluyan al factor Año (si bien considera la posibilidad de incluir otro tipo de interacciones), argumentando que el interés primario del análisis recae sobre el efecto año. Como se ha mostrado, esto último no soluciona necesariamente el problema y, en realidad, puede llegar a agravar la situación (“contaminación” del índice de abundancia anual).

Por otra parte, Punt *et al.* (2000) consideran en el modelo solamente el término de interacción Área x Año, pero utilizando a su vez el procedimiento de definir un estimador de la abundancia anual, ponderado por áreas.

El simple hecho de que las interacciones que incluyan el factor Año sean estadísticamente significativas no es razón suficiente para pensar que los índices de abundancia anuales obtenidos a partir del MLG, deban ser desechados. El problema,

por lo tanto, debe ser abordado desde un punto de vista práctico y para hacerlo se pueden seguir las reglas sugeridas.

Por último, consideramos que si bien los MLG son una herramienta elegante y poderosa para ser utilizada en la construcción de índices de abundancia anuales, deben ser empleados con cuidado cuando incluyen interacciones con el factor Año.

AGRADECIMIENTOS

Expresamos nuestro agradecimiento a los miembros del INIDEP, Claudia Carozza y Claudio Ruarte por habernos facilitado los datos de corvina y pescadilla empleados en este trabajo, y a Gabriela Irusta por acercarnos parte de la bibliografía y sus valiosos comentarios.

BIBLIOGRAFÍA

- CAROZZA, C.R., HERNÁNDEZ, D.R. & PERROTTA, R.G. 2004. Evaluación de corvina rubia (*Micropogonias furnieri*) en el área del Río de la Plata y Zona Común de Pesca Argentino-Uruguay por medio de un modelo de dinámica de biomasa. *Rev. Invest. Desarr. Pesq.*, 16: 77-90.
- CAROZZA, C., PERROTTA, R., COTRINA, C., BREMEC, C. & AUBONE, A. 2001. Análisis de la flota dedicada a la pesca de corvina rubia y distribución de tallas del desembarque. Período 1992-1995. *INIDEP Inf. Téc.*, 41, 17 pp.
- EFRON, B. 1983. The Jackknife, the bootstrap and other resampling plans. *Soc. Ind. Appl. Mathem.*, J.W. Arrowsmith Ltd., Bristol, 92 pp.
- GARCÍA-RODRÍGUEZ, M. 2003. Characterisation and standardisation of red shrimp, *Aristeus antennatus* (Risso, 1816), fishery off the Alicante gulf (SE Spain). *Sci. Mar.*, 67 (1):63-74.
- GAVARIS, S. 1980. Use of a multiplicative model to estimate catch rate and efforts from commercial data. *Can. J. Fish. Aquat. Sci.*, 37: 2272-2275.
- GOÑI, R., ÁLVAREZ, F. & ADLERSTEIN, S. 1999. Application of generalized linear modeling to catch rate analysis of Western Mediterranean fisheries: the Castellón trawl fleet as case study. *Fish. Res.*, 42: 291-302.
- GULLAND, J.A. 1971. Manual de métodos para la evaluación de las poblaciones de peces. Editorial Acribia, Zaragoza, 164 pp.
- HILBORN, R. & WALTERS, C.J. 1992. Quantitative fisheries stock assessment. Choice, dynamics & uncertainty. Chapman and Hall, Nueva York, 570 pp.
- HORN, P.L. 2002. CPUE from commercial line fisheries for ling (*Genypterus blacodes*) around the South Island (Fishstock LIN 3, 4, 5, 6, and 7). *New Zealand Fish. Assess. Report*, 17, 32 pp.
- HORN, P.L. 2003. CPUE from commercial line fisheries for ling (*Genypterus blacodes*) around the North Island, New Zealand: an evaluation of series for LIN 1, LIN 2, and Cook Strait. *New Zealand Fish. Assess. Report*, 13, 49 pp.
- KENDAL, M.G. & STUART, A. 1967. The advanced theory of statistics. Charles Griffin Company Ltd., Londres, Vol. 2, 690 pp.
- LARGE, P.A. 1992. Use of a multiplicative model to estimate relative abundance from commercial CPUE data. *ICES J. Mar. Sci.*, 49: 253-261.
- LASTA, C., RUARTE, C. & CAROZZA, C. 2001. Flota costera argentina. Situación actual y antecedentes. En: BERLOTTI, M., VERAZAY, G. & AKSELMAN, R. (Eds.). *El Mar Argentino y sus recursos pesqueros*. Tomo 3. Evolución de la flota pesquera argentina, artes de pesca y dispositivos selectivos. Publicaciones Especiales INIDEP, Mar del Plata: 89-106.
- MOOD, A.M. & GRAYBILL, F.A. 1963. Introduction to the theory of statistics. McGraw-Hill, New York, 443 pp.
- PERROTTA, R.G. & HERNÁNDEZ, D.R. 2005. Stock assessment of the bluemouth (*Helicolenus*

- dactylopterus*) in Azorean waters during the 1990-2002 period, applying a biomass dynamic model. *Rev. Invest. Desarr. Pesq.*, 17:
- PUNT, A.E., WALKER, T.I., TAYLOR, B.L. & PRIBA, C.F. 2000. Standardization of catch and effort data in a spatially-structured shark fishery. *Fish. Res.*, 45: 129-145.
- ROBSON, D.S. 1966. Estimation of the relative fishing power of individual ships. *Int. Comm. North. Atl. Fish.*, 3: 5-14.
- RUARTE, C., CAROZZA, C. & LASTA, C. 2001. La pesquería de pescadilla de red (*Cynoscion guatucupa*). Período 1975-1999. *Inf. Téc. Int. DNI-INIDEP N° 58/2001*, 17 pp.
- RUARTE, C. & PERROTTA, R.G. (en prensa). Estimación de un índice de abundancia anual estandarizado para pescadilla de red (*Cynoscion guatucupa*), mediante la aplicación de un modelo lineal general. Período 1992-2003. *INIDEP Inf. Téc.*, 64.
- Recibido: 04-08-04*
Aceptado: 20-07-05