

Determinación del Tamaño de Intervalo Óptimo para Análisis de Distribuciones de Frecuencias de Longitudes de *Epinephelus morio*

MIGUEL A. CABRERA-VÁZQUEZ y FRANCISCO ARREGUÍN-SÁNCHEZ
Centro de Investigación y de Estudios Avanzados del IPN Unidad Mérida
Apdo Postal 73 Cordemex
Mérida Yucatán, México

RESUMEN

Se determinó el intervalo de tallas óptimo para agrupar datos de longitudes de *Epinephelus morio* en distribuciones de frecuencias. Mediante simulación de Monte Carlo se construyeron distribuciones de frecuencias de longitudes con seis tamaños de intervalo (5, 10, 15, 20, 25 y 30 mm) y se compararon estadísticamente con la distribución real a través de pruebas de bondad de ajuste de Kolmogorov-Smirnov. Se encontró que estadísticamente no hubo diferencias significativas entre las distribuciones simuladas y la distribución real ($\alpha = 0.05$) cuando el agrupamiento fue de 5, 10 y 20 mm. Se observó una mayor similitud entre la distribución agrupada en este último intervalo con la real por lo que se recomienda utilizar intervalos de 20 mm para agrupar los datos de *E. morio*.

PALABRAS CLAVE: *Epinephelus morio*, tamaño de intervalo, distribuciones de frecuencias de longitudes, simulación.

ABSTRACT

An optimum interval size to group length frequency data for *Epinephelus morio* was determined. Length frequency distributions were constructed by a Monte Carlo simulation using size intervals of 5, 10, 15, 20, 25 and 30 mm and then compared with real distribution through a Kolmogorov-Smirnov test. When grouping was 5, 10 and 20 mm, no significant difference was observed between simulated and real distributions. However, an interval of 20 mm is recommended, because characteristics are similar in both real and simulated distributions.

KEY WORDS: *Epinephelus morio*, size class interval, length frequency data, simulation.

INTRODUCCION

En la evaluación de recursos pesqueros tropicales el uso de metodologías que se basan en el análisis de Distribuciones de Frecuencias de Longitudes (DFL) para estimar los parámetros de crecimiento de la ecuación de von Bertalanffy (principalmente K y L, ha venido cobrando mucha importancia debido a la dificultad de hacerlo a través de métodos directos. Sin embargo, su uso generalizado está limitado por las características de las DFL que sirven

como fuentes de datos, ya que se ha observado que existen algunos factores tales como la variabilidad en la distribución de las longitudes dentro de los grupos de edad y el agrupamiento en diferentes intervalos de longitud que pueden afectar la estimación de dichos parámetros (Hoenig *et al.*, 1987; Basson *et al.*, 1988; Wolff, 1989).

Las DFL de *Epinephelus morio* obtenidas de la flota comercial mexicana que opera en el Banco de Campeche se caracterizan por carecer de componentes modales evidentes, lo cual es un reflejo de algunos aspectos relacionados al ciclo de vida de la especie tales como una gran longevidad (25 años; Moe, 1969), una tasa de crecimiento lenta, y un período de desove prolongado (ocho meses; Mexicano-Cántora, 1990). Esto provoca un intenso traslapamiento de los grupos de edad en la DFL que se hace más crítico cuando son agrupadas en intervalos de tallas grandes. El uso y análisis de esta información para estimar los parámetros de crecimiento de esta especie a través de metodologías que se basan en datos de longitudes puede llevar a cometer sesgos en las estimaciones cuando no se consideran estas fuentes potenciales de error.

En este trabajo se presenta un análisis del efecto del agrupamiento en intervalos de talla de diferente tamaño, sobre el comportamiento de las DFL simuladas de *E. morio*. El objetivo es determinar un tamaño de intervalo óptimo para agrupar las DFL de esta especie, con miras a utilizarlo durante la aplicación de metodologías que se basan en este tipo de información.

MÉTODOS

Se construyó una distribución de frecuencias de longitudes simulada (DFLS) a partir de la simulación individual de los grupos de edad de acuerdo al procedimiento seguido por Cabrera (1991) y el cual se encuentra esquematizado de manera general en la Figura 1. Las longitudes medias para cada grupo de edad utilizadas para las simulaciones se obtuvieron de la lectura de marcas de crecimiento realizada en una muestra de 296 pares de otolitos de *E. morio* (Rodríguez, 1986). Esta muestra fue tomada de los desembarcos comerciales efectuados por la flota mayor y menor entre julio y noviembre de 1984 en los puertos de Progreso y Yucalpetén, Yucatán, México.

La distribución de frecuencias de longitudes observada (DFLO) representa mediciones de la longitud furcal de 4403 organismos que fueron tomadas durante el mismo período que la anterior. Esta distribución fue agrupada en seis intervalos de talla (5, 10, 15, 20, 25 y 30 mm), y comparada con la DFLS agrupada también en intervalos de la misma magnitud. Para elegir el intervalo de tallas óptimo, cada par de distribuciones (DFLS y DFLO) con el mismo tamaño de intervalo se compararon mediante pruebas de bondad de ajuste de Kolmogorov-Smirnov (K-S) para dos muestras con ($\alpha = 0.05$) (Payne, 1982). Se hizo una comparación entre estos resultados con los estimados mediante la aproximación dada por Wolff (1989).

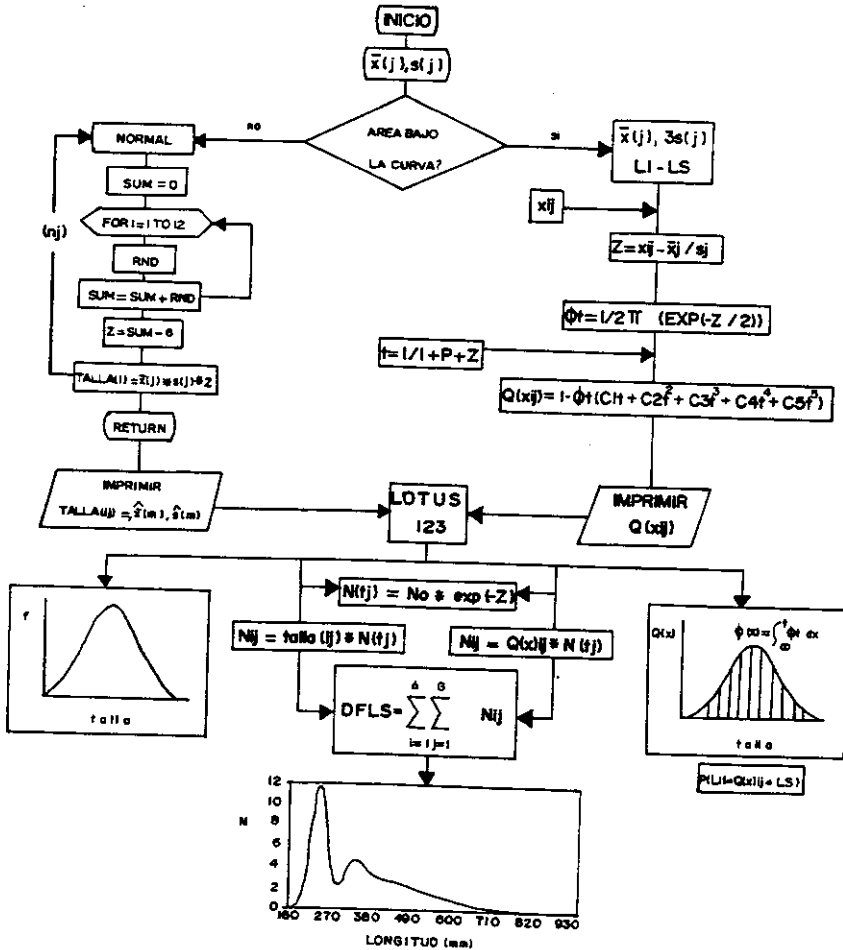


Figura 1. Representación esquemática del procedimiento empleado en la generación de las Distribuciones de Frecuencias de Longitudes Simuladas (DFLS).

RESULTADOS

En la muestra se encuentran representados 12 grupos de edad que abarcan un intervalo de tallas entre 215 mm y 800 mm de longitud; de estos, los últimos tres no están bien representados. La reconstrucción mediante simulación de los primeros nueve grupos de edad a partir de las longitudes medias de la muestra (Figura 2), revela un enorme traslapamiento en la distribución de las longitudes que existe sobre todo en los grupos 6, 7, 8 y 9.

Los diferentes tamaños de intervalo causaron cambios en el comportamiento de los componentes individuales, así como en la DFL total (Figura 3). En general, los cambios observados se relacionaron a cambios en la forma, en la posición de cada uno de ellos en la distribución y en el aumento en el traslapamiento de los grupos de edad más grandes.

El contorno de dichos componentes y de la distribución entera sufrió ligeros cambios. Al aumentar el tamaño del intervalo el contorno sufrió alteraciones más bruscas que hizo que perdieran su forma curva y se tornaran más agudos sobre todo en los grupos iniciales.

Se observaron modificaciones en la altura de los componentes (una mayor tendencia a la leptocurtosis), debido al agregamiento de las tallas de los organismos en los intervalos que están al inicio de la distribución.

El movimiento lateral de los componentes sobre el eje de las abscisas fue otro efecto observado. Este cambio de posición ocasionó un aumento en el traslapamiento, que se hizo más intenso en los grupos finales de la distribución. Asimismo, ocasionó un cambio de posición del valor modal asociado a cada componente con respecto al valor promedio verdadero y al intervalo de talla asociado.

Con respecto a la distribución total se observó que el aumento en el tamaño del intervalo hizo más nítida la presencia de algunos de los componentes que la caracterizan. Cuando el intervalo fue pequeño, la forma de la distribución fue casi continua y sin cambios marcados dificultando la identificación del número de grupos de edad que la componen. A cambios progresivos en el tamaño del intervalo se observaron cambios más marcados en el contorno de la distribución al comienzo de la misma y ligeros al final, y los grupos de edad se empezaron a definir sobre todo los iniciales. Este comportamiento aunque no fué muy notorio en los grupos finales permitió tener una idea más clara acerca de los grupos subyacentes en la distribución en especial los que están al comienzo de ella. Cuando el intervalo fue más amplio (mayor a 20 mm; Figuras 3e y 3f) la distribución perdió las características logradas con el agrupamiento quedando el contorno casi liso y enmascarando la presencia de los grupos que ya se habían definido en la distribución, dificultando su detección.

En relación a la determinación del intervalo de talla óptimo, la Figura 4 muestra una comparación gráfica del comportamiento de la DFLO y de la DFLS al haber sido agrupadas en los tamaños de intervalo considerados. La respuesta

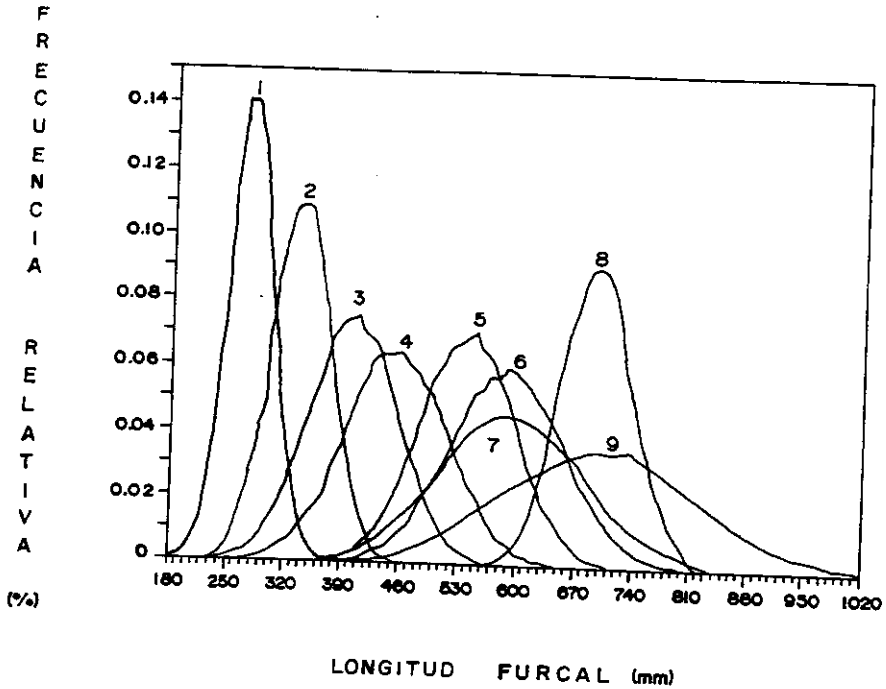


Figura 2. Simulación y reconstrucción de los primeros nueve grupos de edad basada en la longitud media y desviación estándar estimados para cada uno de ellos en la muestra (Rodríguez, 1986).

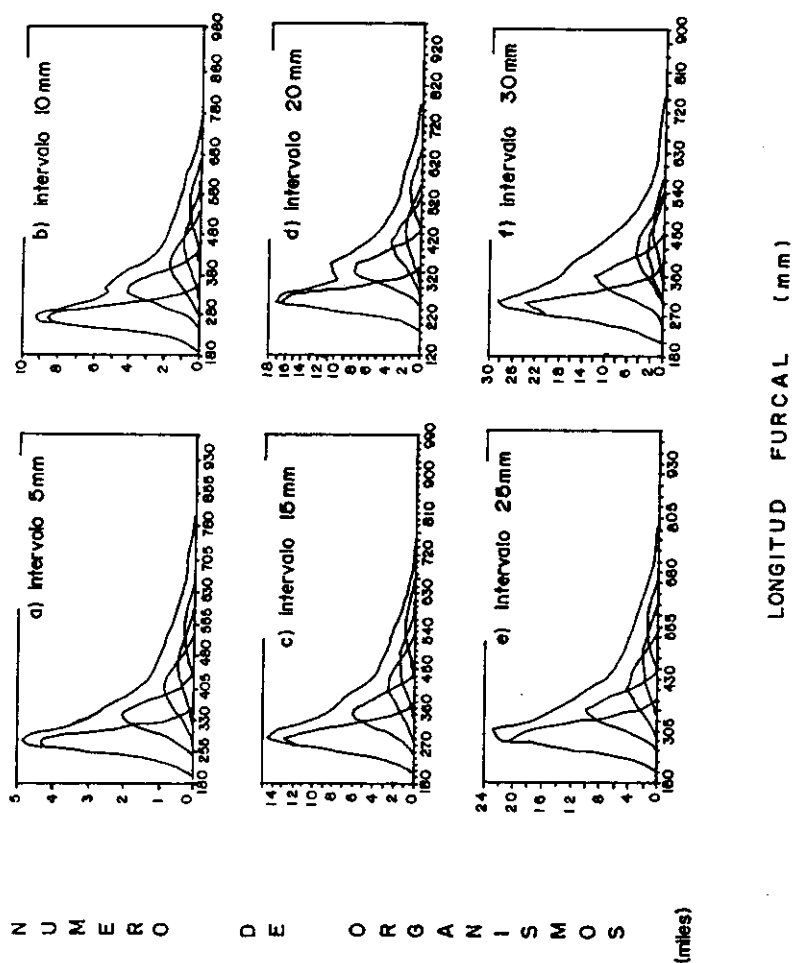


Figura 3. Distribuciones de Frecuencias de Longitudes Simuladas (DFLS) mostrando los componentes modales subyacentes y agrupadas en seis intervalos de talla.

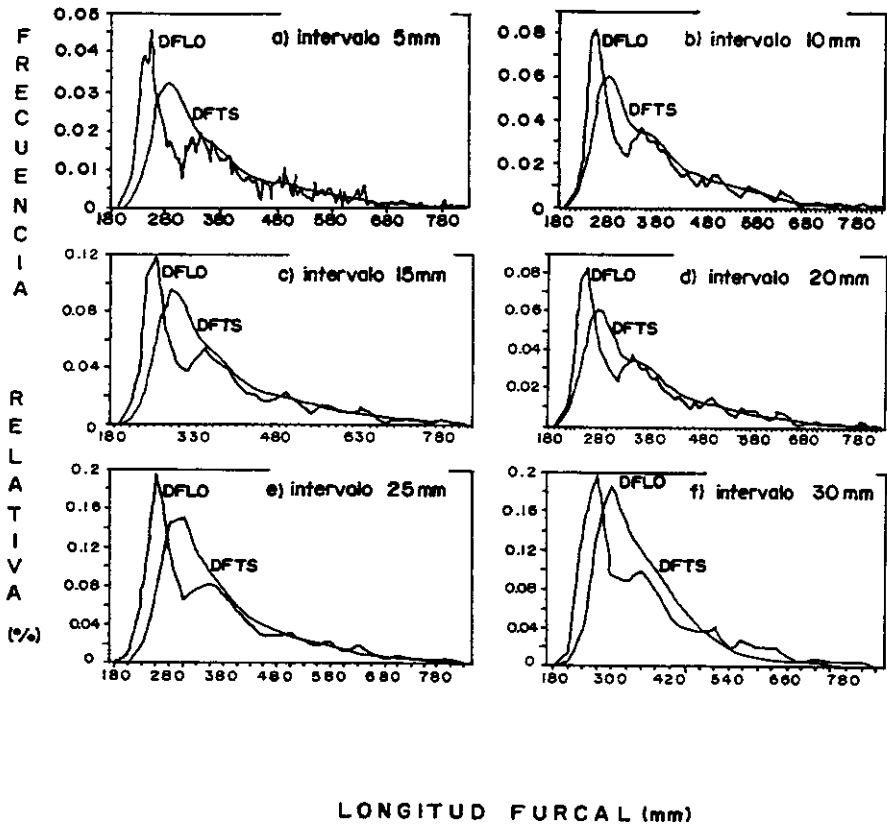


Figura 4. Comparación gráfica de las DFLS agrupadas en seis intervalos de talla con respecto a la distribución de frecuencias de longitudes observada (DFLO).

de la distribución observada al aumento del intervalo se manifestó en una mejor definición en las modas de la distribución, conforme la amplitud del intervalo aumentó. Este aumento modificó sensiblemente la forma de la distribución, hecho corroborado al comparar los resultados de la prueba K-S a través de los valores de la Distancia Máxima Absoluta (DMA) calculados (Figura 5). El comportamiento observado en los valores mostró que aumentos en la amplitud del intervalo causan una mayor diferencia entre las distribuciones comparadas.

En cuanto a la elección del intervalo de talla más adecuado para agrupar los datos, se encontró que la distribución agrupada en intervalos de 5 mm presentó el valor más bajo de DMA (0.085; Nivel de Significancia (NS) = 0.99; Figura 5), observándose muy poca diferencia con respecto al valor obtenido para las agrupadas en intervalos de 10 y 20 mm (0.092 y 0.091, respectivamente; NS = 1.0 en ambas). Esto sugiere que las distribuciones de *E. morio* pueden ser agrupadas para su análisis en cualquiera de los tres intervalos de talla, aunque de acuerdo a la forma de la distribución, la agrupada en intervalos de 20 mm fue más parecida a la observada (Figuras 3d y 4d).

Uno de estos resultados coincidió con los derivados de la aproximación de Wolff (1989), que consiste en dividir el valor de la longitud máxima observada (L_{max}) en la muestra entre el número de grupos de edad que se supone están representados en ella. En este caso, el intervalo óptimo estimado para todas las DFL estuvo entre 10 y 11 mm y coincidió con uno de los intervalos determinados por simulación.

DISCUSION

La reconstrucción de los grupos de edad demostró claramente el traslapamiento intenso que hay en los grupos más grandes. Los últimos cuatro grupos de edad tuvieron valores muy similares en sus longitudes medias, lo cual podría hacer pensar que los grupos 6 y 7 sean uno solo puesto que la variabilidad en ambos es muy semejante. Este caso no es muy claro para los grupos 8 y 9 puesto que aunque sus longitudes medias son muy aproximadas la variabilidad en las longitudes define la diferencia entre ambos grupos y esto probablemente sea debido a sesgos en el muestreo dirigido hacia esas edades que traen como consecuencia la poca representatividad de esos dos grupos en la muestra.

Si bien en la lectura de marcas de crecimiento en otolitos es posible cometer dos tipos de error que pueden llevar a subestimar o sobreestimar las longitudes medias de las edades de la muestra de *E. morio* (i. e., pérdida de claridad de las marcas de crecimiento, identificación errónea del patrón de formación), se partió de la suposición de que las longitudes medias estimadas para cada grupo de edad no presentaron sesgos en las estimaciones y que los grupos de edad encontrados en la muestra son independientes y representativos de la población. En consecuencia se asumió en este trabajo que no existen errores en la

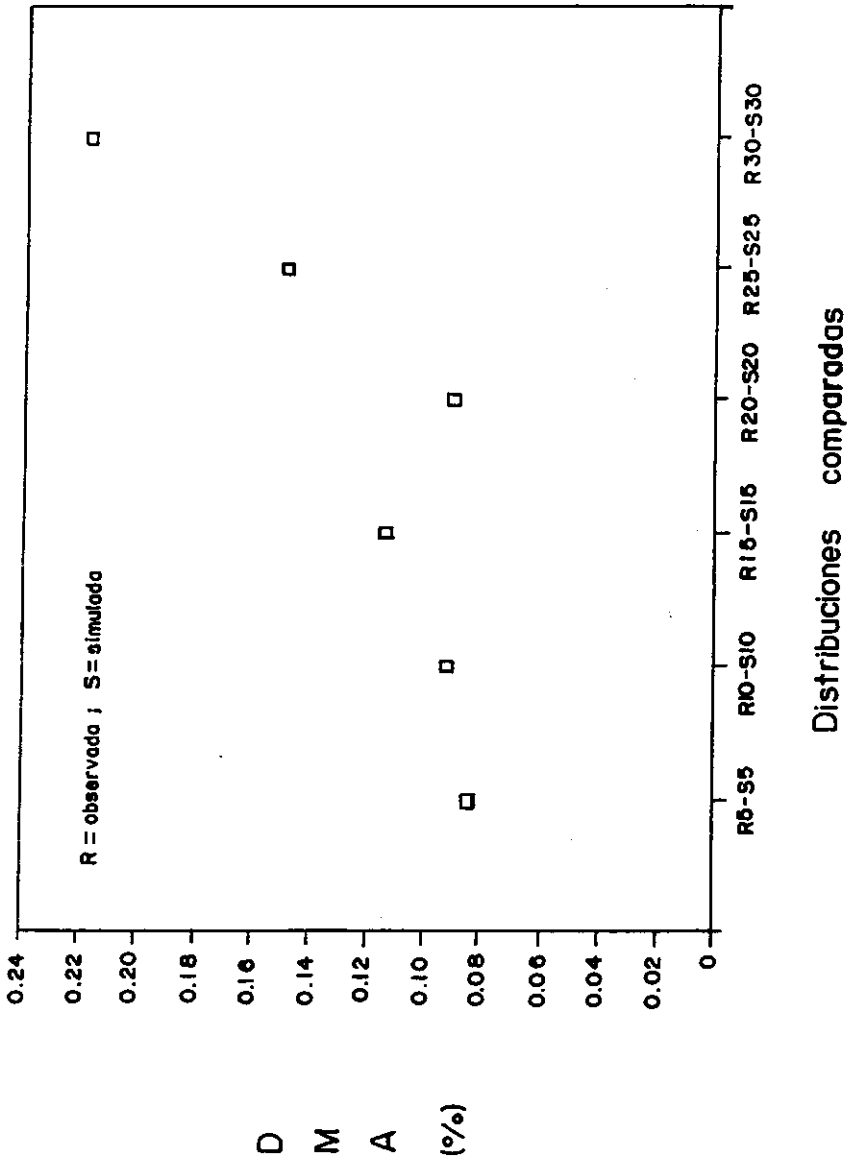


Figura 5. Valores de la Distancia Máxima Absoluta (DMA) obtenidos de la prueba de Kolmogorov-Smirnov (K-S) en la comparación de las DFLS y la DFLO para determinar el tamaño del intervalo óptimo.

determinación de edades y las respectivas distribuciones por edad determinadas en base a la aproximación de Rodríguez (1986).

La experiencia ha demostrado que agrupar una DFL en intervalos de talla de diferente amplitud puede causar cambios en la forma de la distribución y también en la estimación de K y L obtenidas mediante un procedimiento en particular, hecho ya mencionado para el programa ELEFAN I (Hoenig *et al.*, 1987; Majkowsky *et al.*, 1987). Los resultados obtenidos en este trabajo demostraron que se presentan efectos del agrupamiento sobre los componentes subyacentes y sobre la distribución total. De entre éstos, podemos considerar como más importantes a los cambios en la altura, en la amplitud y en la posición de los grupos de edad sobre el eje de las longitudes. La altura de los componentes en la distribución se incrementó con aumentos en el tamaño del intervalo como efecto del agregamiento de nuevos individuos, sobre todo a los grupos de edad iniciales. Los límites de los grupos de edad se modificaron sensiblemente y su efecto principal se notó sobre la distribución, lográndose longitudes máximas más pequeñas. A modo de ejemplo, de 830 mm para un intervalo de 5 mm disminuyó a aproximadamente 780 mm para un intervalo de 30 mm. Todo esto condujo por un lado a pérdida de información básica en lo que respecta a las longitudes máximas observadas, información que puede ser necesaria para hacer estimaciones previas de L por otros procedimientos que se basan en datos de DFL.

Si bien al aumentar la amplitud del intervalo hasta cierto punto se origina una mejor definición de los grupos de edad intermedios subyacentes en la distribución, el intenso traslapamiento que se provoca en las edades mayores cuando se continúa ampliando el intervalo de tallas, puede provocar un enmascaramiento completo de los grupos en la DFL haciéndola parecer como si fuera una sola moda e incluso algunos grupos llegar a sobreponerse completamente. La intensa sobreposición de los grupos de edad puede ser un factor crítico para aquellos procedimientos que se basan en la identificación de las modas en una distribución (e. g., ELEFAN: Pauly y David, 1981; BHATTACHARYA: modificado por Pauly y Caddy, 1985). La ampliación del intervalo de tallas implicó un cambio en la posición de las longitudes con respecto a un grupo de edad y con ellas el intervalo de tiempo transcurrido entre longitudes sucesivas, lo cual ocasionó cambios importantes en los valores modales. Asimismo las proporciones de las clases de longitud en los grupos de edad que componen a las distribuciones de frecuencias analizadas, también se ven afectadas por dicho cambio. Así, métodos que utilizan esta información como base para la estimación de L y K pueden también resultar afectados en sus estimaciones (SLCA: Shepherd, 1987a; PROJMAT: Shepherd, 1987b).

La determinación de un intervalo de tallas óptimo para agrupar las longitudes individuales de *E. morio* mostró que los datos pueden ser agrupados en distribuciones de 5, 10 y 20 mm debido a que estadísticamente no hay una

diferencia significativa entre los valores de DMA estimados para los pares de distribuciones comparadas. La elección de 20 mm como el intervalo de talla óptimo fue por las siguientes razones: 1) La DFL agrupada en intervalos de 20 mm reflejó más fielmente las características de la distribución simulada en cuanto a forma y definición de grupos. 2) Puesto que *E. morio* es una especie de crecimiento lento, los incrementos en longitud con respecto al tiempo son muy pequeños y por lo tanto estos cambios no pueden quedar bien representados en la DFL al agrupar los datos en intervalos pequeños (Caddy, 1986).

Estos resultados son comparables a los obtenidos por Erzini (1990) quien simulando distribuciones de frecuencias de longitudes para una especie hipotética de crecimiento lento y gran longevidad y cuyas características son semejantes a las de *E. morio* encontró que una amplitud de 20 mm puede considerarse como el intervalo de talla óptimo para agrupar datos de longitud. Asimismo encontró que la estimación del intervalo de talla óptimo está en función del tamaño de muestra y de las características biológicas de la especie (variabilidad de la longitud en cada edad, patrón de reclutamiento, tasa de crecimiento y talla máxima). De acuerdo a este autor estos factores pueden afectar la claridad y definición de las modas en la distribución.

Una ventaja asociada a las estimaciones realizadas en este trabajo y por Erzini (1990) es que la determinación del intervalo de talla óptimo está relacionada con las características de la especie reflejadas en la forma de la distribución usada, en tanto que la aproximación de Wolff (1989) depende claramente de la suposición del número de grupos de edad en la distribución, aunque puede ser bastante útil si se dispone de determinaciones previas de edad. Es claro que en *E. morio* este criterio no es del todo aplicable debido a las características de las distribuciones que anteriormente ya se han discutido.

LITERATURA CITADA

- Basson, M., A. Rosenberg, y J.R. Beddington. 1988. The accuracy and reliability of two new methods for estimating growth parameters from length-frequency data. *J. Cons. int. Explor. Mer*, 44: 277-285.
- Cabrera, M. A. 1991. Efecto de cambios en la varianza y tamaño del intervalo en la estimación de los parámetros de crecimiento del mero *Epinephelus morio*. Tesis de Maestría. CINVESTAV-IPN Unidad Mérida. 97 p.
- Caddy, J. F. 1986. Size frequency analysis in stock assessment, some perspectives, approaches and problems. *Proc. Gulf Carib. Fish. Inst.*, 37: 212-238.
- Erzini, K. 1990. Sample size and grouping of data for length-frequency analysis. *Fishery Research*. 9: 355-366.

- Hoenig, J. M., J. Csirke, M. J. Sanders, A. Abella, M. G. Andreoli, D. Levi, S. Ragonese, M. Al-Shoushani y M.M. El-Musa. 1987. Data acquisition for length-based stock assessment: Report of working group I. pp. 343-352. En D. Pauly y G.R. Morgan (eds.) *Length-based methods in fisheries research. ICLARM Conference Proceedings* 13.
- Majkowski, J., J. Hampton, R. Jones, A. Laurec y A. A. Rosenberg. 1987. Sensitivity of length-based methods for stock assessment: Report of working group III. pp. 363-372. En D. Pauly y G.R. Morgan (eds.) *Length-based methods in fisheries research. ICLARM Conference Proceedings* 13.
- Mexicano-Cántora G. 1990. Análisis preliminar de algunos aspectos reproductivos del mero (*Epinephelus morio*) de las costas de Yucatán. *Reporte de Especialización Académica*. CINVESTAV-IPN. Unidad Mérida. 42 p.
- Moe, M. A. 1969. Biology of the red grouper *Epinephelus morio* (Valenciennes) from the eastern Gulf of Mexico. *Florida Department of Natural Resources Professional Papers Ser.ies*. No.10:95.
- Pauly, D. y N. David. 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. *Sonderdruck aus Bd.*, 28: 205-211.
- Pauly, D. y J. F. Caddy. 1985. A modification of Bhattacharya's method for the analysis of mixture of normal distributions. *FAO Fish. Circ.* (781): 16p.
- Payne, J.A. 1982. *Introduction to simulation. Programming techniques and methods of analysis*. McGraw-Hill USA 324p.
- Rodríguez, S. H. 1986. Determinación de la edad y crecimiento del mero (*Epinephelus morio* Valenciennes) del Banco de Campeche utilizando dos estructuras óseas diferentes (Otolito y Hueso Mesopterigoides). Tesis licenciatura. UANL. Monterrey, México.
- Shepherd, J. G. 1987a. A weakly parametric method for estimating growth parameters from length composition data, pp. 113-119. En D. Pauly y G.R. Morgan (eds.) *Length-based methods in fisheries research. ICLARM Conference Proceedings* 13.
- Shepherd, J. G. 1987b. Towards a method for short-term forecasting of catch rates based on length compositions, pp. 167-176. En D. Pauly y G.R. Morgan (eds.) *Length-based methods in fisheries research. ICLARM Conference Proceedings* 13.
- Wolff, M. 1989. A proposed method for standardization of the selection of class intervals for length-frequency analysis. *Fishbyte*, 7: 5.