

## **INFORME DE LA REUNIÓN DE 2014 DEL GRUPO DE TRABAJO SOBRE MÉTODOS DE EVALUACIÓN DE STOCKS (WGSAM)**

*(Dublín, Irlanda, 7 - 11 de abril de 2014)*

### **1 Apertura de la reunión, adopción del orden del día y disposiciones para la reunión**

La reunión se celebró en las oficinas del Irish Sea Fisheries Board (BIM) en Dublín, Irlanda, del 7 al 11 de abril de 2014. Las disposiciones locales fueron llevadas a cabo por el Dr. Michael Keatinge. La Dra. Pilar Pallarés, en nombre del Secretario Ejecutivo de ICCAT, expresó su agradecimiento al BIM por acoger la reunión y facilitar las disposiciones logísticas.

El Dr. Michael Schirripa, Relator del Grupo de trabajo sobre métodos de evaluación de stocks, presidió la reunión. El Dr. Schirripa dio la bienvenida a los participantes (el Grupo) y procedió a revisar el orden del día que fue adoptado sin cambios (**Apéndice 1**).

La lista de participantes se adjunta como **Apéndice 2**.

La lista de documentos presentados a la reunión se adjunta como **Apéndice 3**.

Los siguientes participantes actuaron como relatores de las diversas secciones del informe:

<i>Sección</i>	<i>Relatores</i>
1,10	P. Pallarés
2	L. Kell
3	C. Porch
4	H. Arrizabalaga y G. Melvin
5	C. Brown
6	G. Díaz
7	M. Schirripa y V. Ortiz de Zárate
8-9	N. Abid

### **2 Examen de los métodos actuales de ICCAT para estimar la distribución de esfuerzo (EFFDIS)**

El SCRS/2014/026 presentaba una breve revisión de los métodos anteriores de estimación del esfuerzo total de palangre, así como sugerencias para mejorar esta estimación, tal y como discutió el WGSAM en 2013. El cálculo más reciente de EffDIS se realizó en 2009 y utilizó las nueve especies principales de túnidos y especies afines de ICCAT para obtener capturas nominales globales de Tarea I (en peso) y las CPUE a partir de estadísticas parciales de captura y esfuerzo (Tarea II). El supuesto base del modelo actual considera que las tasas de captura son equivalentes a nivel parcial y global. Al comparar los resultados con estimaciones anteriores (obtenidas durante reuniones intersesiones del Subcomité de ecosistemas en 2007 y 2008), los resultados globales no mostraban diferencias importantes. Sin embargo, a niveles más disgregados las diferencias eran mayores para algunos pabellones. En la mayoría de los casos, las importantes variaciones relativas se suelen asociar con diferentes mejoras y correcciones de algunos conjuntos de datos específicos. La distribución geográfica global muestra un pequeño incremento cerca de las aguas de Venezuela debido a las diferentes correcciones realizadas en las distribuciones espaciales de estadísticas de captura y esfuerzo de Venezuela durante varios años. Se presentan campos recomendados para mejorar la estimación de EffDIS con el fin de impulsar los debates y finalizar los métodos que deberían usarse en el futuro.

Se reconoce la importancia de tener en cuenta los cambios en las operaciones pesqueras y en las características de las flotas principales de cada CPC que opera en la zona del Convenio de ICCAT ya que estas afectan a la eficacia de las flotas a la hora de capturar especies objetivo y especies de captura fortuita. La documentación de estos cambios tecnológicos y de estrategia es especialmente importante para entender los informes nacionales de captura y esfuerzo que se presentan anualmente (Tarea II-CE). El Subcomité de Ecosistemas cuenta también con varias recomendaciones relacionadas, por ejemplo con i) restricciones y/o limitaciones potenciales de los datos y la información, ii) descripción de la incertidumbre, por ejemplo, sustituciones, ratios de extrapolación y

proporción de flotas sin clasificar ("otras"), iii) deberían explorarse métodos adicionales de extrapolar los datos y iv) deberían explorarse métodos como la validación cruzada.

El WGSAM discutió el uso de EffDIS y los temas relacionados con su estimación. En particular, ¿es el actual nivel temporal y espacial de agregación (es decir, cuadrículas de 5 grados y mes) adecuado para su uso previsto? esto podría evaluarse llevando a cabo estudios con diferentes niveles de agregación y comparando las conclusiones obtenidas con el conjunto de datos actual. Asimismo, era importante la clasificación de las flotas y también se pensó que sería necesario apoyo externo para ayudar a realizar el EffDIS en el futuro.

EffDIS es un recurso importante para el Subcomité de Ecosistemas y en 2013, el Subcomité de Ecosistemas recomendó que se hicieran esfuerzos para desarrollar estimaciones EffDIS similares para los artes de cebo vivo y cerco. Esto resultaría especialmente útil para las evaluaciones de medidas de ordenación de vedas espacio-temporales. Será importante identificar qué medida de esfuerzo global utilizar para estos dos artes.

A pesar de la importancia de EffDIS para el trabajo del SCRS, la última evaluación se llevó a cabo en 2007. Esto se debe principalmente a varias cuestiones metodológicas que aún deben ser resueltas (véase SCRS/2013/036 para un resumen completo) y a la falta de recursos para llevar a cabo el análisis. En concreto, es necesario validar el supuesto utilizado para construir la base de datos que está agregada a nivel de cuadrículas de 5 grados) y evaluar si conviene. Por lo tanto, el Grupo recomendó que:

- 1) Se exploren métodos para clasificar las flotas y
- 2) Se necesita apoyo externo para producir EffDIS en el futuro.

### **3 Cuantificación de la incertidumbre en las evaluaciones de ICCAT**

La Comisión espera un asesoramiento sobre medidas de ordenación basado en el riesgo, tal y como se establece en la matriz de estrategia de Kobe II y en su marco de toma de decisiones [Rec. 11-13]. Un aspecto importante a la hora de proporcionar dicho asesoramiento científico es cuantificar adecuadamente la incertidumbre sobre el estado del stock y las perspectivas futuras en escenarios de futuras opciones de ordenación. Con la llegada de modelos de evaluación de stock altamente parametrizados y más comúnmente aplicados, la inversión en informática para cuantificar la incertidumbre sobre el estado del stock y las perspectivas futuras es bastante elevada. Otras OROP de túnidos han tenido experiencias similares y se están aplicando una serie de aproximaciones para cuantificar ambos procesos y la incertidumbre por observación con el fin de desarrollar asesoramiento en materia de ordenación basado en el riesgo. Se pidió por tanto al WGSAM que proporcionara orientación sobre la evolución y la posibilidad de armonizar métodos para caracterizar la incertidumbre en los diferentes grupos de especies. En este sentido, el WGSAM indicó que todos los modelos de evaluación representan simplificaciones de un sistema pesquero más complejo y, por tanto, están intrínsecamente limitados en su capacidad de reflejar la variabilidad inherente de dicho sistema. En otras palabras, los gestores no deben esperar que la incertidumbre en el asesoramiento científico pesquero sea nunca completamente cuantificada. No obstante, el Grupo convino en que pueden hacerse mejoras en la forma en que se representa la incertidumbre en los diversos grupos de especies.

El Grupo identificó tres enfoques básicos para representar la incertidumbre en los resultados de los modelos de evaluación de stock: (1) un enfoque "basado en el modelo" que tenga en cuenta explícitamente las principales fuentes de incertidumbre percibidas utilizando una plataforma de modelación única y versátil, (2) promediando modelos, es decir, combinando ensayos alternativos de diferentes plataformas de modelación y (3) un enfoque "empírico" que utilice información histórica existente sobre la coherencia de los resultados de los modelos de evaluación de stock a lo largo del tiempo. El primer enfoque ha sido el más común en la práctica. Habitualmente, se desarrolla un solo modelo de base que incluye los parámetros estimables que representan lo que se percibe como las fuentes clave de la incertidumbre (con o sin distribuciones previas informativas). Posteriormente, se calculan las medidas de la incertidumbre en los parámetros clave de ordenación mediante métodos estándar (hessiano inverso, bootstrap, integración bayesiana) y se incorporan en las matrices de Kobe. Pueden ejecutarse los ensayos de sensibilidad o las plataformas alternativas del modelo para obtener más información sobre posibles incertidumbres, pero a menudo se usan principalmente para cualificar el asesoramiento del caso base del modelo y no se incorporan directamente en las matrices de Kobe. La eficacia de este enfoque depende de la capacidad para modificar el modelo de evaluación de una forma que aproxime razonablemente la fuente de incertidumbre.

Una segunda estrategia para incorporar la información facilitada por modelos alternativos, que se ha usado a veces, es la promediación de modelos, en la que se asignan ponderaciones por frecuencia a cada posible modelo, basándose tal vez en alguna medida del ajuste de los datos (por ejemplo, AIC, ponderación por la inversa de la varianza) o en la opinión de los expertos (véase también la sección 5). Una desventaja es que este enfoque tendrá probablemente como resultado una distribución multimodal de las medidas del estado del stock que se están considerando (por ejemplo, punto de referencia límite) o requerirán supuestos acerca de su forma de distribución (por ejemplo, normal o lognormal). El Grupo consideró que el éxito de cualquier enfoque basado en un modelo depende enormemente de la capacidad del grupo de evaluación para desarrollar distribuciones previas informativas o postular hipótesis alternativas y razonables antes de descubrir las implicaciones en cuanto a ordenación de estas hipótesis. Si no, existe el peligro de producir estimaciones sesgadas tanto de la tendencia central como de la incertidumbre descartando modelos considerados no informativos o introduciendo intencionadamente sesgos añadiendo modelos o alterando las distribuciones previas. Se mencionó que el Centro nacional de huracanes de Estados Unidos utiliza la promediación de modelos de forma rutinaria y deriva las ponderaciones para cada modelo candidato de su rendimiento en pasadas predicciones (es decir, comparando las predicciones del modelo con las trayectorias reales del huracán). Lamentablemente, a diferencia de las trayectorias de los huracanes, rara vez se conoce el verdadero estado de un stock de peces.

Los enfoques empíricos examinan el rendimiento pasado del modelo para deducir los rangos de la incertidumbre total. De forma ideal, se compararía el rendimiento del modelo con la verdadera disposición del stock, pero una vez más, en la práctica esto no es posible. Por otro lado, se podría comparar el rendimiento de dos evaluaciones de stocks históricas de referencia durante cada año que tienen en común. La incertidumbre total resultante incluiría implícitamente errores de estimación "dentro del modelo", así como errores sistémicos que podrían producirse con cambios en los modelos o en las filosofías del equipo de evaluación. Este enfoque ha sido ya adoptado por el Consejo de ordenación pesquera del Pacífico en Estados Unidos (véase Ralston et al. 2011) y se está considerando en otros consejos. Durante las discusiones del Grupo de trabajo se plantearon algunos posibles problemas con el método empírico. Se expresó la inquietud de que dicho enfoque no reflejaría la tendencia de las evaluaciones de stock de mejorar en precisión y exactitud con el tiempo mediante el añadido de nuevos datos, mejores modelos y una mejor recopilación de datos. Sin embargo, otros en el grupo de trabajo señalaron que las mejoras en los datos solo pueden disminuir parte de la incertidumbre y que la composición de los científicos del equipo de evaluación de stock, la composición del panel de revisión de la evaluación y los cambios en los métodos de evaluación de stock podrían tener la misma influencia en la coherencia de los resultados de la evaluación de stock. Se sugirió que, al menos, los análisis empíricos podrían utilizarse para ayudar a proporcionar una base real para las estimaciones de varianza derivadas de los actuales procedimientos basados en modelos.

Se realizó una presentación titulada "Cuantificando la incertidumbre debida al procesamiento de los datos en las evaluaciones de stocks estructuradas por edad". Se presentó el trabajo preliminar detallando un método para asignar los datos de talla faltantes a partir de observaciones basadas en su proximidad espacio-temporal (asumiendo que las observaciones eran aproximadamente multivariantes normales con covarianzas estimadas). Posteriormente, se utilizaron las distribuciones estimadas para generar muchos conjuntos de datos de talla completos (muestreando los parámetros asignados al modelo y posteriormente muestreando a partir de las observaciones de talla ponderadas), que fueron convertidos a distribuciones de edad utilizando diversos métodos de determinación de la edad (por ejemplo, separación de cohortes y claves de edad-talla). Cada conjunto de datos de edad se usó posteriormente en un VPA con diferentes tasas de mortalidad natural e índices de abundancia relativa. Los resultados preliminares sugieren que la atribución de tallas con un tamaño de muestra efectivo de 100 es demasiado baja para generar cualquier variabilidad en los conjuntos de datos de captura por talla derivados. La tasa de mortalidad natural (80%, 100% y 120% de los niveles de la evaluación más reciente) determina enormemente la  $F_{RMS}$  estimada a partir del análisis XSA.

El Grupo se mostró de acuerdo en que las variaciones en este enfoque serían una forma útil de intentar cuantificar cómo la incertidumbre en los datos se convertiría en incertidumbre en la evaluación. Dicho enfoque ayudaría también a identificar la importancia relativa de una fuente de incertidumbre sospechada y la correspondiente necesidad de incorporarla explícitamente al desarrollar las matrices de Kobe utilizadas para formular el asesoramiento científico a los gestores. Durante la discusión de evaluaciones de estrategias de ordenación, se señaló también que los modelos operativos deberían construirse teniendo en mente las mayores fuentes de incertidumbre percibidas y que las mismas podrían usarse para ayudar a cuantificar el alcance de propagación de estas incertidumbres en la incertidumbre del asesoramiento científico (que, de hecho, es una variación del enfoque discutido más arriba). Se indicó también que los análisis de riesgo como los llevados a cabo para el atún rojo (Leach et al. 2014) son útiles para identificar las principales fuentes de incertidumbre

percibidas por los científicos y otras partes interesadas y que podrían tener, además, el beneficio añadido de conseguir una mayor aceptación cuando se presentan los resultados.

#### 4 Descripción de la calidad de los datos pesqueros y la información biológica

Se presentó el documento SCRS/2014/035 con el objetivo de revisar la información de contexto y fomentar las discusiones sobre este punto del orden del día. La Resolución [13-15] para la estandarización de la información científica en el informe anual del SCRS requiere "...calificar la calidad de los datos de las pesquerías y de los datos relacionados con los conocimientos sobre las especies (por ejemplo, parámetros biológicos, datos históricos de patrones de distribución de la pesquería, selectividad) utilizados como entradas para las evaluaciones de stock. Las calificaciones cualitativas de los supuestos y datos de entrada podrían ser detalladas y deberían resumir el estado de los conocimientos de las diferentes entradas...". El Informe de 2013 del Subcomité de estadísticas proponía un método para calificar los diferentes elementos de datos. El método consistía en una tabla con una lista de categorías y puntos (dentro de una categoría) que eran evaluados y ponderados (por ejemplo, de acuerdo con la importancia relativa de los puntos dentro de cada categoría, en la evaluación de stock). La calificación de cada punto era el producto de su valor y su ponderación, y la calificación de cada categoría era la suma de las diferentes puntuaciones de cada punto dentro de una categoría. Se sugirió calcular la calidad de la información para una determinada evaluación de stock como la clasificación media de las categorías. A la propuesta le faltaban, no obstante, criterios predefinidos para basar la clasificación, como los utilizados en los Catálogos de datos de especies de la IOTC, en los que los límites de las puntuaciones están determinados por afirmaciones cualitativas acerca de la disponibilidad/calidad de los datos (por ejemplo, "disponible de acuerdo con los estándares", "no disponible de acuerdo con los estándares" o "no disponible en absoluto").

Por tanto, el SCRS/2014/035 proponía un marco genérico, algo similar a la propuesta del SC de estadísticas, pero con directrices de puntuación predefinidas y cualitativas. La puntuación de calidad propuesta (Q) de los diferentes conjuntos de información era:

$$Q=D*I$$

donde:

D=disponibilidad y calidad de los datos

Con

- 1= no hay datos o son de calidad mala/desconocida
- 2=si existen datos, pero la calidad no es muy buena
- 3=si existen datos de buena calidad

I=impacto de los datos en los resultados de la evaluación (cuán importante es disponer de este conjunto de información).

Con:

- 1=escaso impacto en los resultados de la evaluación
- 2=impacto moderado en los resultados de la evaluación
- 3=impacto elevado/desconocido en los resultados de la evaluación

El documento incluía varios ejemplos en los que se aplicaba provisionalmente la puntuación a diferentes stocks de atún blanco y listado. Se puntuaron los diferentes puntos relacionados con las pesquerías y la biología y la puntuación total de cada stock reflejaba la suma de los valores de Q en los diversos puntos.

El Grupo consideró que este marco genérico era útil como punto de partida y debatió varios puntos sobre ello. Se llegó al acuerdo de que las puntuaciones deberían basarse en una metodología objetiva, cuantificable y científicamente defendible. Aunque esto puede ser más fácil de lograr cuando se trata de puntuar la disponibilidad de datos, se reconoció que medir objetivamente la calidad de los datos podría ser algo más difícil. Sin embargo, el grupo sugirió mecanismos para avanzar en la puntuación en la medida de lo posible.

En esta etapa, una consideración importante era el propósito o utilidad de este ejercicio. Por una parte, es necesario responder a la Resolución 13-15 con el fin de que la Comisión cuente, además de la matriz de Kobe, con información acerca de la calidad de los datos utilizados para derivar la matriz de Kobe. Con este fin, se consideró alguna representación simple (similar a los ejemplos aportados en el SCRS/2014/035). No obstante, el

grupo indicó que las puntuaciones de calidad podrían usarse también para otros fines. Por ejemplo, podrían ser útiles para que los grupos de trabajo científicos describan las principales fuentes de incertidumbre y para mejorar, en la medida de lo posible, la formulación del asesoramiento científico (por ejemplo, seleccionando modelos adecuados y/o ponderando diferentes escenarios).

El Grupo constató que la importancia de los diferentes conjuntos de datos (puntuaciones "I") depende del modelo usado, así como del ciclo vital del stock. Por tanto, las puntuaciones totales "Q" podrían no ser comparables entre stocks. Para solucionar esto, el grupo recomendó cuantificar la ratio entre la puntuación total y la máxima puntuación posible teniendo en cuenta el vector I específico para cada stock, lo que permitiría alguna comparación basada en colores entre los stocks.

Por último, el Grupo recomendó desarrollar una base de metadatos que contenga información sobre la calidad y cantidad de los datos pesqueros y biológicos disponibles para las evaluaciones. Esto, junto con el mecanismo existente para evaluar los índices de abundancia independientes de la pesquería y de las CPUE, ayudaría a puntuar de forma más objetiva la calidad y cantidad de la información relacionada con la pesquería que se utiliza en la evaluación de las especies de ICCAT con el fin de aportar información a los científicos que hacen la evaluación de stock y proporcionar a la Comisión una base para que tenga una perspectiva global de la calidad de los datos, es decir, el valor de las puntuaciones "D". El Grupo discutió la posible estructura de una base de metadatos que puede posteriormente ser alimentada y actualizada a medida que se disponga de nueva información.

La base de metadatos constaría de tres componentes principales, uno sobre datos pesqueros, otro sobre datos biológicos y un tercero con estudios de marcado recaptura (**Apéndices 4 y 5**, respectivamente). La estructura sugerida es flexible y podría incorporar nuevos campos a medida que se prevea su necesidad. En el caso de los metadatos biológicos, se consideró que serían ya útiles los campos incluidos en "Información general", además de alguna información acerca de qué parámetros y cuán bien estimados están. Los campos incluidos en "estimaciones de parámetros" permitirían llevar a cabo metaanálisis que podrían ayudar de diversas formas, por ejemplo, describiendo la incertidumbre sobre determinados parámetros. No obstante, llenar estos campos sería, por supuesto, una intensa labor. El grupo consideró también que, en casos particulares, podría ser útil tener acceso a los datos en bruto de algunos estudios. La base de metadatos permitiría identificar fácilmente las fuentes de estos datos en bruto.

## 5 Conciliar los resultados cuando se utilizan múltiples métodos de modelación

Las evaluaciones del SCRS a menudo, y cada vez más, han incluido el uso de múltiples métodos de modelación (es decir, tipos diferentes de modelos, hipótesis alternativas) para estimar el estado del stock en relación con los elementos de referencia de conservación de ICCAT. El Grupo discutió las diversas razones de por qué se hace esto. Por ejemplo, los tipos de modelos pueden diferir en los supuestos subyacentes, y los participantes en la evaluación pueden no ser capaces de determinar cuál es más adecuado. En dichos casos, se prevé que incorporar los resultados de múltiples modelos en el asesoramiento sobre ordenación refleja mejor la incertidumbre en los resultados. En otros casos, las configuraciones alternativas de la misma plataforma o tipo de modelación pueden ser consideradas lo suficientemente plausibles para incluirlas en los resultados. El Grupo se mostró de acuerdo en que este enfoque tiene sus puntos buenos, pero reconoció que podría haber dificultades para determinar cuál es la forma más adecuada de combinar dichos resultados, lo que incluye asignar una ponderación adecuada, especialmente en aquellos casos en los que los diversos resultados del modelo pueden entrar en conflicto.

No hay ningún documento SCRS disponible sobre este tema. Sin embargo, el Grupo examinó el trabajo reciente de Deroba et al. (2014) empleando la simulación para probar la robustez de los modelos de evaluación de stock ante el error. Las conclusiones de este documento que están relacionadas con este tema incluyen:

- Las mayores diferencias se produjeron al comparar los resultados de diferentes tipos de modelos (modelo de producción excedente, basado en la edad, etc.); los modelos tendían a funcionar de forma similar a otros modelos del mismo tipo.
- La autopruueba es útil y recomendable, y la autopruueba y la prueba cruzada frecuentemente destacan la divergencia en los años más recientes de la serie temporal.
- Entre los modelos, la variabilidad puede ser considerada como un tipo de incertidumbre en la evaluación; esto tiene implicaciones al considerar si basar el asesoramiento sobre ordenación solo en un modelo configurado para ser el "mejor ajuste" para la evaluación o si incorporar los resultados de diferentes modelos (por ejemplo, la aplicación de la promediación de modelos).

Las conclusiones de Deroba et al. (2104) eran coherentes con las experiencias de las evaluaciones de stock del SCRS y proporcionan una confirmación analítica de la importancia de la selección del modelo (especialmente respecto al tipo de modelo), el examen de diagnósticos y la validación de los resultados, y las implicaciones de integrar o no los modelos en el asesoramiento final.

El Grupo consideró que un componente inicial esencial a la hora de reconciliar diferentes resultados del modelo es el proceso de selección de los modelos más adecuados a partir de los que desarrollar el asesoramiento en materia de ordenación. Este proceso debería iniciarse durante la reunión de preparación de datos, donde debe reservarse un tiempo para determinar qué modelos se usarán en la evaluación. Tal y como se ha descrito anteriormente, deberían considerarse para los diversos modelos los requisitos de datos, los supuestos y la capacidad de reflejar cambios importantes en la pesquería que se sabe que han ocurrido, teniendo en cuenta la calidad y cantidad de los datos disponibles. La implicación de los científicos de las CPC es clave en este proceso, ya que son los más familiarizados con los cambios en las diversas pesquerías que podrían merecer consideración en el proceso de evaluación, y están preparados para identificar cualquier inquietud con la adecuación de conjuntos de datos particulares para su uso en los diversos modelos. También debería considerarse el rendimiento de los diversos modelos utilizados en la evaluación anterior.

Durante la reunión de evaluación, debería asignarse tiempo suficiente para realizar un cuidadoso examen de los diagnósticos y el rendimiento del modelo. Las mediciones del rendimiento podrían incluir una evaluación de si los resultados parecen coherentes con lo que se conoce de la biología de la especie, así como de las tendencias pesqueras (un ejemplo de un enfoque así, "medición del rendimiento de la predicción retrospectiva", se detalla a continuación). Deberían considerarse la probabilidad y las consecuencias de la infracción de supuestos para cada modelo teniendo en cuenta la información disponible sobre la población y la pesquería. Una posible consecuencia de este proceso es que el modelo identificado durante la reunión de preparación de datos para su uso en la evaluación podría ser rechazado (adecuadamente) para su inclusión en el asesoramiento final en materia de ordenación.

El Grupo discutió si sería adecuado o no considerar los resultados de modelos más simples junto con los resultados de modelos más complejos (por ejemplo, usando los resultados de un modelo de producción excedente y los de un modelo más plenamente integrado, como stock synthesis). En otras palabras, si los datos fueran suficientes para ejecutar los modelos más complejos (que, en teoría, podrían tener en cuenta más aspectos y cambios en los parámetros del stock y la pesquería), ¿debería, en su lugar, basarse el asesoramiento solo en los resultados de los modelos más complejos? El Grupo consideró que podría haber circunstancias en las que es adecuado considerar los resultados de dichos tipos de modelos divergentes.

El Grupo debatió las opciones para procedimientos o criterios adicionales para la selección de los modelos. Se sugirió uno de estos procedimientos, basado en la medición del "desempeño de la predicción retrospectiva". La idea de este método es similar a la del análisis retrospectivo tradicionalmente utilizado en los VPA, pero ampliándola a la comparación de modelos múltiples. Por ejemplo, partiendo del supuesto de que se utilizan dos modelos en la evaluación que arrojan dos resultados diferentes. Las fases del proceso de medición de la predicción retrospectiva serían:

- truncar los datos disponibles, eliminando los xx años más recientes (en función de lo que sea apropiado para cada stock);
- volver a ejecutar cada modelo para obtener una estimación de la dinámica de la población durante xx años del pasado;
- proyectar el modelo hacia adelante utilizando las capturas reales hasta el año actual;
- calcular una tendencia de la CPUE (o quizá solo la tendencia de la biomasa estimada, lo que proceda en función de la complejidad del modelo) que se predice para los años proyectados hasta el año actual y
- comparar dicha tendencia de la CPUE con la(s) tendencia(s) real(es) de la CPUE (estandarizada) observadas.

Con este enfoque, pueden dejar de considerarse algunos modelos si los resultados de los análisis demuestran que existen incoherencias con la(s) tendencia(s) de la CPUE observada(s). El método puede ampliarse más para desarrollar ponderaciones para los diferentes resultados del modelo (que puede utilizarse cuando se combinan resultados entre modelos) mediante la definición de una medida de discrepancia entre las series de CPUE predichas y observadas (por ejemplo, suma de los residuos cuadrados dividida por SE de la CPUE observada). Se constató que este enfoque no requiere que los modelos utilicen los mismos conjuntos de datos. La aplicabilidad de este enfoque depende de que la(s) tendencia(s) de la CPUE real observada (estandarizada) refleje(n) adecuadamente la población y la ausencia de índices de CPUE en conflicto. El Grupo consideró que

este enfoque era prometedor, y recomendó que continúen los trabajos de investigación relacionados con la idoneidad de su utilización en la selección del modelo y/o en la ponderación.

El Grupo constató que también es importante la selección de los modelos que se van a utilizar para las proyecciones. En algunas circunstancias, para realizar las proyecciones puede utilizarse un modelo diferente al utilizado para estimar el estado actual del stock. Esto puede suceder, por ejemplo, en los casos en los que el modelo utilizado para estimar el estado del stock no puede utilizarse para las proyecciones o requiere un tiempo de procesamiento adicional (más del disponible durante la reunión de evaluación) para producir algunos elementos de salida que se requerirían para la proyección. El Grupo indicó que había surgido cierta inquietud en relación con la idoneidad de este enfoque. El Grupo identificó la necesidad de desarrollar enfoques para investigar esta cuestión. Se sugirió un enfoque potencial que implicaría volver a ejecutar el modelo utilizado para estimar el estado actual del stock, eliminando algunos de los años recientes de datos (como en el análisis retrospectivo). Entonces podría ejecutarse el modelo considerado para las proyecciones, proyectándose a partir del último año del ensayo del modelo retrospectivo, utilizando el historial de captura real hasta el año en curso. Los resultados del modelo de proyección podrían evaluarse mediante una comparación de la condición/estado proyectado en el año actual con la condición/estado estimado mediante el modelo de evaluación.

Después de que los científicos que realizan la evaluación se pongan de acuerdo sobre los modelos que se tienen que incluir para desarrollar el asesoramiento en materia de ordenación (por ejemplo, estimaciones del estado del stock con respecto a los puntos de referencia biológicos e incertidumbre asociada, matriz de estrategia de Kobe), queda la cuestión de la forma en que se tienen que presentar los resultados. Una opción que generalmente es la preferida sería combinar en cierto modo estos resultados para proporcionar niveles de referencia y representaciones adecuadas de la incertidumbre. Esto resulta particularmente difícil cuando los diversos resultados de los modelos entran en conflicto. El Grupo manifestó que resulta generalmente inapropiado combinar resultados que son totalmente incoherentes entre cada modelo, por ejemplo, cuando los diferentes modelos están estructurados basándose en dos hipótesis alternativas que reflejan estados asumidos muy diferentes, con poco o ningún solapamiento en las distribuciones de probabilidades. En dichos casos, combinar dichos resultados del mismo modo produciría estimaciones del estado del stock y/o tendencias promediadas entre las tendencias, lo que no es coherente con ninguna de las dos hipótesis alternativas. En dichos casos, se insta a los científicos a examinar con rigor los supuestos subyacentes de cada hipótesis y a considerar detenidamente si cada uno de ellos está respaldado por los datos existentes y la ciencia actual. Si este proceso resulta insuficiente para descartar una hipótesis, los científicos deberían considerar si se pueden asignar las probabilidades relativas de cada hipótesis. Sería aconsejable realizar más trabajos de investigación para recopilar datos que contribuyan a la aceptación/descarte de las hipótesis o a asignar probabilidades.

La práctica general del SCRS cuando se combinan resultados de múltiples modelos es ponderar igualmente los resultados de cada modelo. Tal y como se mencionó en la sección 3 de este informe, las ponderaciones podrían asignarse a los resultados de cada modelo basándose en alguna medición del ajuste a los datos (por ejemplo, AIC; ponderación de variación inversa). Además de las preocupaciones descritas en la sección 3, este enfoque es limitado en el sentido de que sólo es apropiado cuando los modelos potenciales utilizan los mismos conjuntos de datos.

El Grupo también debatió el potencial de utilizar la evaluación de estrategias de ordenación para aportar información a la ponderación/selección del modelo. El Grupo identificó un problema potencial en este enfoque en el sentido de que el modelo operativo de la MSE podría influir en la percepción del buen funcionamiento del modelo. Se sugirió que la utilización de modelos de producción excedente podría resultar problemática si las funciones de productividad presentan una fuerte asimetría hacia la izquierda y la productividad asciende a medida que los niveles del stock merman.

El Grupo reconoció que en un esquema de ponderación es inevitable que exista un elemento de subjetividad. Incluso la asignación por defecto de ponderación igual es implícitamente una decisión que podría dar más peso que el que se merece a algunos resultados, si la totalidad de los factores, como los supuestos del modelo, capacidad, calidad de los datos, tendencias de la pesquería y biología, pueden tenerse en cuenta de un modo objetivo. Por tanto, el Grupo considera que la opinión experta podría ser apropiada para la asignación de ponderaciones, si está respaldada por el conocimiento de estos factores y la postulación de mecanismos razonables que podrían respaldar una conclusión de que es más probable que ciertos modelos reflejen mejor la situación real que otros.

Otra práctica común en las evaluaciones del SCRS es descartar en última instancia para su inclusión en el asesoramiento en materia de ordenación varios modelos que producen resultados plausibles pero que podrían ser

considerados en cierto modo menos probables. El Grupo constató que se había manifestado la preocupación con respecto a que la práctica existente de seleccionar únicamente un subconjunto de modelos considerados más probables en los resultados finales podría dar lugar a una subestimación de la gama de incertidumbre. Se requieren más trabajos de investigación.

## **6 Puntos de referencia límite, normas de control de la captura y evaluaciones de estrategias de ordenación**

La evaluación de puntos de referencia límite (LRP) y normas de control de la captura (HCR) mediante el uso de evaluación de estrategias de ordenación (MSE) cuenta cada vez más con el reconocimiento de las OROP mundiales de túnidos como un medio eficaz de avanzar en el proceso de ordenación de pesquerías. Las evaluaciones de 2013 de atún blanco y pez espada se utilizaron como ejemplos de cómo un proceso MSE podría incluirse formalmente en la ordenación de estos stocks. El WGSAM prevé continuar con este esfuerzo mediante: (1) la continuación de la mejora de los métodos en el marco del proceso MSE; (2) la introducción de MSE en más evaluaciones cuando y donde proceda; y 3) el fomento de líneas de comunicación que mantengan informados a los gestores de sus beneficios y de sus puntos débiles. En lo que concierne al diálogo y la comunicación, hace poco ICCAT ha adoptado la *Recomendación de ICCAT para mejorar el diálogo entre los gestores y científicos pesqueros* [Rec. 13-18] que tiene como objetivo mejorar la comunicación y favorecer el entendimiento mutuo entre los científicos y gestores pesqueros con el fin de facilitar una toma de decisiones basada en la ciencia y simplificada. Asimismo, la Recomendación describe tareas específicas que se tienen que realizar durante la primera reunión del Grupo de trabajo permanente para mejorar el diálogo entre los gestores y científicos pesqueros (SWGSM) en 2014.

Se ha adoptado una norma de control de la captura (HCR) empírica para el atún rojo del sur (SBT) para establecer el total admisible de capturas (TAC). La HCR se basa en cambios de un año a otro en los índices a partir de un índice de CPUE de abundancia de adultos dependiente de la pesquería y de una prospección aérea de juveniles independiente de las pesquerías. Antes de que pueda implementarse la HCR, se deben seleccionar niveles de referencia apropiados para la captura y los índices; y los parámetros HCR deben adaptarse para cumplir los objetivos de ordenación mediante la evaluación de estrategias de ordenación (MSE).

En el documento SCRS/2014/036 se llevó a cabo una MSE preliminar para el atún rojo del Mediterráneo utilizando la HCR del SBT como parte de un procedimiento de ordenación (MP, a saber, la combinación de datos predefinidos, con un algoritmo al que se asignan los datos para proporcionar un valor para un TAC o una medida de control del esfuerzo). Los próximos pasos serán:

- 1) Identificar objetivos de ordenación y confrontarlos con las mediciones del desempeño con el fin de cuantificar en qué medida se han alcanzado.
- 2) Seleccionar hipótesis sobre la dinámica del sistema.
- 3) Modelos operativos (OM) en función de los datos y conocimientos, así como para el posible descarte y ponderación de las diferentes hipótesis.
- 4) Identificar estrategias de ordenación potenciales y codificarlas como procedimientos de ordenación.
- 5) Proyectar los modelos operativos hacia el futuro utilizando los MP como procedimientos de control de respuesta, y
- 6) Llegar a un acuerdo sobre los MP que cumplen mejor los objetivos de ordenación.

El Grupo resaltó la necesidad de seleccionar detenidamente los supuestos de los modelos operativos dado que los supuestos erróneos pueden afectar a los resultados de las evaluaciones de stock y de las evaluaciones de estrategias de ordenación (MS). Se debatió la importancia de la selección de la gama de escenarios utilizados en el modelo operativo (Fromentin et al., 2014; Leach et al., 2014). El Grupo acordó que en un marco MSE, el objetivo no es escoger el procedimiento de ordenación o las normas de control de la captura (HCR) que proporcionen los mejores resultados, sino los que sean más robustos en los escenarios escogidos. Por ejemplo, si el modelo operativo se ejecuta con cuatro niveles de inclinación, que se asume que son igualmente plausibles, la HCR elegida debería ser la que alcance los objetivos de ordenación con estos cuatro niveles de inclinación. El Grupo también debatió el hecho de que muchos modelos operativos asumen que la regulación de la población solo se produce en el reclutamiento, sin embargo, el modelo operativo debería captar también otras posibilidades de regulación (por ejemplo, mortalidad natural).

El Grupo acordó que el modelo operativo debería ser más complejo que el modelo de estimación. Por ejemplo, un modelo de producción excedente puede utilizarse como el modelo de estimación mientras que un modelo más



complejo, como SS3, puede utilizarse como modelo operativo. El documento presentaba un ejemplo en el que se utilizaron dos prospecciones como índices para reclutas y adultos. Se debatió la posibilidad de no tener que realizar estos tipos de prospecciones cada año para obtener índices de biomasa que aporten información a la HCR. Se indicó al Grupo que utilizar una HCR podría reducir los requisitos de datos en comparación con los datos requeridos actualmente en algunas evaluaciones de stock. El Grupo también debatió si la HCR debería incluir puntos de referencia límite más precautorios para aquellos casos en los que el índice de biomasa estimado a partir de las prospecciones presenta una mayor incertidumbre.

Se revisó la presentación que el Presidente del SCRS realizó a la Comisión en 2013. El Grupo resaltó la importancia de la próxima primera reunión del Grupo de trabajo permanente dedicado al diálogo entre los gestores y científicos pesqueros (SWGSM), durante la cual se debatirán cuestiones relacionadas con los puntos de referencia límite, probabilidades asociadas con la matriz de Kobe, periodos de recuperación para los diferentes stocks, etc. El Grupo resaltó que la HCR no podría desarrollarse sin este diálogo entre científicos y gestores.

En el documento SCRS/2014/025 se describe cómo los autores buscan evidencias de una relación stock-reclutamiento para el atún rojo, rabil y atún blanco. Los autores llegaron a la conclusión de que la evidencia de la existencia de una SRR para cualquiera de los stocks era débil y que los datos podrían respaldar también otras hipótesis (Vert-pre, 2014), por ejemplo, la de Gilbert (1998) de que el reclutamiento fluctúa en torno a un nivel medio durante un periodo de tiempo y luego se produce un cambio de régimen. Esto tiene obvias implicaciones para la evaluación de stock y el asesoramiento en materia de ordenación.

El Grupo debatió que la hipótesis de “cambio de régimen” presentada en el documento podría ser un producto de los modelos de estimación. Dicho de otro modo, aunque el “cambio de régimen” no es incoherente con la ausencia de una SRR observada en los stocks mencionados, esto, en sí mismo, no es una prueba de que se haya producido realmente un “cambio de régimen”.

En el documento se sugería que la SSB es una función de reclutamiento (en vez de que el reclutamiento sea una función de la SSB). El Grupo preguntó por qué parecía haber autocorrelación en los patrones de reclutamiento. Se debatió el hecho de que la autocorrelación observada podría deberse a la utilización de separación de cohortes para asignar edades a los stocks estudiados o a otros supuestos del modelo.

El documento sugería que a efectos de las proyecciones del estado del stock, debería darse más peso a los reclutamientos recientes observados en vez de ponderar todos los patrones de reclutamiento de la serie temporal del mismo modo (dicho de otro modo, es más probable que el reclutamiento de mañana sea más similar al de ayer que al reclutamiento de hace 25 años).

En el documento SCRS/2014/037 se utilizaron las series de SSB y de R procedentes de los escenarios del caso base para stock de atún rojo del Atlántico este y Mediterráneo, utilizando capturas “infladas” y “comunicadas” para ajustar tres modelos diferentes S-R (Beverton & Holt, Ricker and smooth Hockey-stick). Los resultados muestran que el stock mantuvo toda su capacidad reproductiva a lo largo de la serie temporal, desde 1950 y que la relación S-R de Ricker no puede descartarse para este stock. En el documento se afirmaba que este último resultado añade más incertidumbre a la estimación de  $B_{01}$ , lo que hace que no pueda utilizarse como punto de referencia de biomasa para este stock. Como alternativa el documento propone la utilización de  $B_{loss}$  como  $B_{lim}$  y después estimar  $B_{pa}$  a partir de dicho valor. Esto permitió a los autores simular el comportamiento de algunas HCR y seleccionar la más adecuada para este stock.

El Grupo debatió el hecho de que ICES está dejando de lado el marco de punto de referencia límite, que es lo contrario a lo que se propone en el documento. Al margen de esto, la utilidad del enfoque propuesto en el documento puede probarse mediante simulación. También se debatió la validez de establecer supuestos en los modelos operativos para que sean más precautorios (por ejemplo utilizando una relación de reclutamiento en forma de palo de hockey). En líneas generales, el Grupo manifestó que no estaba de acuerdo con este enfoque. El Grupo también debatió los factores que podrían contribuir a relaciones S-R de tipo Ricker en los atunes, pero no llegó a una conclusión final.

## **7 Incorporación de información sobre ecosistemas, clima y hábitat (ECH) en las evaluaciones de stock**

Los debates sobre la incorporación de información sobre ecosistema, clima y hábitat (ECH) en las evaluaciones de stock se centraron en los dos enfoques que pueden utilizarse para mejorar el proceso de evaluación de stock

de las especies de ICCAT. Se identificaron dos enfoques, uno cualitativo y otro cuantitativo, y este último tenía como objetivo reducir la incertidumbre en el proceso de evaluación de stock.

Bajo este punto del orden del día se expuso una única presentación bajo el título “Hipótesis de un reciente desplazamiento hacia los polos en la distribución del pez espada del Atlántico norte”, basándose en el documento SCRS/2013/161. Se describían los cambios en los índices de abundancia de pez espada con respecto a los indicadores de procesos medioambientales a gran escala (por ejemplo, AMO, NAO, AWP). Basándose en observaciones de tendencias opuestas en la abundancia de pez espada septentrional, el documento sugería la posibilidad de un desplazamiento de la abundancia de latitudes más cálidas y meridionales hacia latitudes más frías y septentrionales. Varios de los índices de abundancia observados cambiaban abruptamente de dirección de negativo a positivo, mientras que otros mostraban un cambio opuesto. Los cambios observados en la dirección de los índices de abundancia se corresponden con cambios en las tendencias en el tamaño de la Piscina Cálida del Atlántico (AWP), el cambio de signo de la Oscilación Multidecadal del Atlántico (AMO) y la Oscilación del Atlántico Norte (NAO). Para cuantificar una posible relación entre los cambios en la abundancia y los diversos índices medioambientales posibles, se ejecutó el modelo de evaluación sin la influencia de los datos medioambientales y se realizó una regresión de los residuos del ajuste de las CPUE a los diversos índices medioambientales. Teniendo en cuenta los presuntos límites de tolerancia del pez espada a la temperatura, es posible que o bien su hábitat preferido se haya desplazado hacia el norte o bien se haya desplazado su presa preferida o ambos.

Basándose en estos resultados, se inició un debate sobre las ventajas de la agregación de las diferentes CPUE por área, en vez de la agregación actual por flota disponible para el Grupo de especies sobre pez espada. Esto requerirá un acceso a los datos de la flota para cada operación de pesca, lo que permitirá la incorporación del efecto zona y flota, así como de la estructura espacial en los modelos de evaluación (a saber las CPUE del Atlántico este frente a las CPUE del Atlántico oeste). También se resaltó que hay otros enfoques para incorporar la zona como un factor de interacción en la modelación de las CPUE mediante la utilización de modelos mixtos. Los enfoques de ponderación podrían calcularse tras estandarizar las CPUE por zona.

El Grupo también presentó la utilización de la modelación del hábitat para incorporar la variabilidad regida por el medio ambiente, determinando la distribución espacial y temporal en la modelación de las CPUE.

Hubo una amplia discusión sobre cómo abordar la agregación de los datos disgregados por operaciones de pesca para las pesquerías de palangre dirigidas al pez espada, como continuación del estudio presentado. Un campo principal de discusión fue el modo de establecer un marco de colaboración apropiado entre los científicos nacionales para obtener datos de cada operación de pesca y analizar estos datos conjuntamente. Este intercambio podría realizarse en el marco del acuerdo de confidencialidad de ICCAT ya existente. Se reconoció que sería un proceso gradual, que posiblemente requerirá el apoyo de la Secretaría, así como trabajar en la “nube” con un acceso restringido a los científicos que participan en el análisis en colaboración de estos datos. Estos datos son necesarios para abordar una serie de preocupaciones/interacciones asociadas con la pesquería y los índices con el fin de mejorar la evaluación e incorporar en ella factores medioambientales.

## **8 Otros asuntos**

### **8.1 Colaboración con SISAM-ICES**

Se expuso al Grupo una presentación realizada por uno de los copresidentes de ICES\_SISAM (iniciativa estratégica sobre métodos de evaluación de stock). De este modo se resaltaba el trabajo realizado en el marco de esta iniciativa hasta la fecha, y se mantuvo un debate sobre la orientación futura de SISAM y su relevancia para la comunidad científica de ICCAT. Los científicos de ICCAT participaron en los trabajos de SISAM en 2013, a través del Congreso mundial sobre métodos de evaluación (WCSAM). Las actividades de SISAM en 2014 incluyen completar documentos para una edición especial del *ICES Journal of Marine Science* (basada en el WCSAM), y jornadas de trabajo de organizaciones asociadas como CAPAM, así como una sesión abierta que celebrará ICES en su ASC en septiembre de 2014. En los debates de la reunión se resaltó que ICES e ICCAT se enfrentan a cuestiones similares al menos en términos de desarrollo de métodos de evaluación de stocks y de provisión de asesoramiento sobre pesquerías, y que estas cuestiones podrían ser comunes a muchas OROP a nivel mundial. Estas cuestiones estaban relacionadas con los avances en las metodologías y el reto de mantener en la red de la organización a científicos con experiencia técnica suficiente para utilizar estos métodos de un modo adecuado. Junto con esto, se debatieron los esfuerzos compartidos para desarrollar marcos de evaluación y asesoramiento que abordan situaciones en las que los datos son limitados. Se acordó que sería beneficioso para

ICCAT mantener la colaboración con la iniciativa SISAM y, a través de esta plataforma, trabajar colaborando con científicos de otras OROP para contribuir a la resolución de estas cuestiones.

## 9 Recomendaciones

- El Grupo recomendó que se inste a las CPC a comunicar sus datos de captura y esfuerzo de Tarea II en una estratificación geográfica más fina (por ejemplo 1°x1°) en vez de comunicar estos datos en estratos de 5°x5°, ya que en algunos lugares dicha escala podría resultar demasiado tosca. El WGSAM solicita además que el Subcomité de estadísticas considere si debería requerirse también la comunicación en esta resolución de escala más fina.
- El Grupo acordó también que debería fomentarse la implementación del enfoque de evaluación de estrategias de ordenación (MSE) y el diálogo entre científicos y gestores pesqueros sobre normas de control de la captura y MSE para mejorar el asesoramiento científico proporcionado a la Comisión. Estos esfuerzos deberían incluir una revisión de los esfuerzos MSE realizados hasta la fecha considerando los logros, fracasos y los recursos que limitan los progresos de MSE futuras.
- El Grupo manifestó que pensaba que los diferentes grupos de trabajo podrían utilizar criterios simples para empezar a asignar una puntuación a la calidad de la información utilizada en diferentes evaluaciones de stock. Entre tanto, el Grupo recomienda que se sigan desarrollando modos objetivos más detallados y basados en la ciencia de proporcionar dichos sistemas de puntuación. Esto incluye desarrollar criterios para evaluar la importancia de los diferentes elementos de datos dependiendo del ciclo vital y/o del modelo de evaluación utilizado. En esta misma línea, el Grupo recomendó que prosiga el desarrollo de la base de metadatos con información sobre la cantidad y calidad de la información biológica y pesquera disponible.
- El Grupo instó a que la medición de la realización de predicción retrospectiva, descrita en la sección 5 de este informe, se evalúe mediante estudios de simulación, potencialmente con modelos condicionados por evaluaciones pasadas de ICCAT. A ser posible, debería desarrollarse un caso de estudio para un stock para el que no se hayan observado series de CPUE contradictorias.
- El Grupo instó a las CPC a proporcionar un acceso limitado a los datos de CPUE para cada operación de acuerdo con las necesidades y prioridades identificadas por los diferentes grupos de especies y subcomités. Esto permitiría al SCRS producir una amplia variedad de índices en una escala espacial más informativa. En un primer momento podría utilizarse la información de una sola especie y/o flota específica para ilustrar los beneficios de este enfoque. Se sugirió la utilización de las oportunidades que brinda la “nube” mantenida por la Secretaría para fines de archivo y acceso, con el fin de facilitar la colaboración multilateral. Este intercambio podría realizarse en el marco del acuerdo de confidencialidad de ICCAT ya existente.
- El Grupo consideró también que el SCRS debería continuar participando en la iniciativa ICES\_SISAM para seguir fomentando la colaboración para el desarrollo de metodologías de evaluación, para compartir y desarrollar los conocimientos sobre el modo de comunicar la incertidumbre a los gestores, para propiciar una colaboración más estrecha en las evaluaciones conjuntas (por ejemplo, marrajo sardinero), así como iniciativas prácticas, como compartir el orden del día de los grupos de trabajo sobre métodos de ICES y de ICCAT. Esto proporcionará un acceso abierto publicitado y estandarizado a las datos de entrada y salida de las evaluaciones, así como la posibilidad de compartir datos más detallados. Estos elementos podrían desarrollarse mediante la participación en el Grupo de trabajo global sobre métodos de evaluación (GAME).

## 10 Adopción del informe y clausura

El informe fue adoptado durante la reunión. El presidente del WGSAM expresó su agradecimiento a los organizadores locales por el excelente trabajo de organización de la reunión y a los participantes por su eficacia y el gran trabajo realizado. La Secretaría reiteró su agradecimiento a la *Irish Sea Fisheries Board* (BIM) por la excepcional organización de la reunión y por el cálido apoyo prestado a los participantes. La reunión fue clausurada.

## Referencias

- Leach, Adrian W., *et al.* "Identification and prioritization of uncertainties for management of Eastern Atlantic bluefin tuna (*Thunnus thynnus*)."*Marine Policy* 48 (2014): 84-92.
- Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., and R. J. Conser. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. *Fish. Bull.* 109:217–231 (2011)
- Fromentin J.-M., S. Bonhommeau, H. Arrizabalaga, and L. L. Kell. (2014). The spectre of uncertainty in management of exploited fish stocks: the illustrative case of Atlantic bluefin tuna. *Marine Policy*, 47: 8–14, 2014.
- Gilbert D. Towards a new recruitment paradigm for fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(4):969–977, 1997.
- Leach A., P. Levontin, J. Holt, L. Kell, and J. Mumford. (2014) Identification and prioritization of uncertainties for management of eastern Atlantic bluefin tuna (*Thunnus thynnus*). *Marine Policy*, 48: 84–92.
- Ralston, S., Punt, A. E., Hamel, O. S., DeVore, J. D., and R. J. Conser. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. *Fish. Bull.* 109:217–231 (2011)
- Vert-pre K. A., R. O. Amoroso, O. P. Jensen, and R. Hilborn. Frequency and intensity of productivity regime shifts in marine fish stocks. *Proceedings of the National Academy of Sciences*, 110(5):1779–1784, 2013.

## APÉNDICES

**Apéndice 1.** Orden del día

**Apéndice 2.** Lista de participantes.

**Apéndice 3.** Lista de documentos.

**Apéndice 4.** Campos sugeridos para la base de metadatos pesquerías y recuperación de marcas.

**Apéndice 5.** Bases de metadatos biológicos para estudios individuales.