

COMISIÓN INTERAMERICANA DEL ATÚN TROPICAL

**2ª REVISIÓN DE LA EVALUACIÓN DE LA POBLACIÓN DE ATÚN ALETA
AMARILLA EN EL OCÉANO PACÍFICO ORIENTAL**

La Jolla, California (EE. UU.)
2-6 de diciembre de 2019

INFORME DE LA REUNIÓN

Shannon Cass-Calay (Presidente), Alistair Dunn, Adam Langley, Steve Teo, Laura Tremblay-Boyer

ÍNDICE

Resumen ejecutivo.....	1
1. Antecedentes	2
2. Comentarios generales por el Panel de Revisión.....	3
3. Revisión de las entradas de datos.....	3
4. Desplazamiento y estructura de la población.....	10
5. Biología del atún aleta amarilla del OPO	13
6. Estructura del modelo.....	15
7. Sensibilidad del modelo y diagnósticos	17
8. Resumen de los comentarios del público	18
9. Resumen de las conclusiones y trabajo futuro	18
Anexos:	
A: Areas de enfoque identificadas por el personal de la CIAT	23
B: Términos de referencia	24
C: Lista de participantes	26
D: Documentos disponibles al Panel de Revisión	26
E: Solicitudes y respuestas	29

RESUMEN EJECUTIVO

La [segunda revisión de la evaluación de la población de atún aleta amarilla en el Océano Pacífico oriental](#) (YFT del OPO) se llevó a cabo del 2 al 6 de diciembre de 2019 en el Hotel Marriott en La Jolla, California. De conformidad con los Términos de Referencia (TdR), el objetivo de la revisión no fue realizar una revisión de una evaluación específica en relación con la provisión de asesoramiento de ordenación, sino más bien identificar las investigaciones necesarias para mejorar la evaluación del aleta amarilla del OPO y para proveer asesoramiento que ayudaría al enfoque que debería adoptarse para desarrollar una evaluación para la 11ª Reunión del Comité Científico Asesor en mayo de 2020. Cuando se reunió el Panel de Revisión (PR), el personal de la CIAT todavía estaba en proceso de desarrollar un modelo de caso base para 2020. Por lo tanto, el PR no pudo proceder con una revisión formal de todos los aspectos de la evaluación, pero procuró en cambio proporcionar asesoramiento y recomendaciones sobre investigaciones, análisis y posibles opciones para desarrollar la evaluación.

El personal de la CIAT proporcionó documentos de referencia, numerosas presentaciones, y respuestas a las solicitudes del PR que abordaron temas relacionados con entradas de datos, ciclo vital y aspectos del modelado. La revisión se enfocó en las investigaciones necesarias para mejorar la evaluación actual. Los temas principales de la reunión incluyeron:

- ¿Qué causa el desfase en los índices de abundancia relativa basados en la CPUE de palangre y de cerco?
- ¿Cuál es la estructura de la población más adecuada para la evaluación del atún aleta amarilla?
- ¿Cuál es la estructura de la pesquería más adecuada para la evaluación del atún aleta amarilla?
- ¿Qué método debería usarse para abordar la incertidumbre en la talla de los individuos más viejos y el impacto que tiene en los resultados de la evaluación?
- ¿Cuál es la relación población-reclutamiento más adecuada?
- ¿Cómo se deberían usar los índices de abundancia de CPUE en la evaluación?
- ¿Se debería usar selectividad logística y para cuál pesquería/estudio?
- ¿Cómo se debería tratar el reciente aumento de la talla de los peces capturados con palangre?
- Mortalidad natural por edad y por sexo.

El Panel señala que los siguientes temas no fueron discutidos ni evaluados en la revisión:

- Monitoreo y estimación de descartes
- Parametrización de la capturabilidad

1. ANTECEDENTES

La revisión de la evaluación de la población de atún aleta amarilla en el Océano Pacífico oriental (YFT del OPO) se llevó a cabo del 2 al 6 de diciembre de 2019 en el Hotel Marriott en La Jolla, California. La reunión fue presidida por Shannon Cass-Calay (NOAA SEFSC) y los otros miembros del PR fueron Alistair Dunn (Consultor), Adam Langley (Consultor), Steve Teo (NOAA SWFSC), Laura Tremblay-Boyer (Consultora). La revisión se enfocó en las investigaciones necesarias para desarrollar una evaluación para el Comité Científico Asesor de la CIAT en 2020. El personal de la CIAT ya había reconocido que la evaluación previa del YFT del OPO no era suficientemente robusta para proveer asesoramiento fiable. Al momento de la revisión, el equipo de la CIAT estaba en proceso de revisar los insumos y la estructura del modelo para la evaluación de 2020.

Por consiguiente, el PR señaló que el propósito de la revisión de la evaluación del YFT del OPO del personal de la CIAT no era determinar si la evaluación actual o propuesta era adecuada para proporcionar asesoramiento de ordenación; la intención era brindar información al equipo de evaluación que pudiera mejorar la nueva evaluación. Las metas y objetivos de la revisión son:

- identificar la mejor ciencia disponible para uso en la evaluación;
- proveer una revisión independiente de la evaluación; y
- brindar asesoramiento sobre investigaciones futuras que mejoren la evaluación y la provisión de asesoramiento de ordenación.

Los TdR señalan que la principal responsabilidad del PR era realizar una revisión técnica adecuada de la evaluación.

El personal de la CIAT proporcionó documentos de referencia antes de la reunión, documentos preparados específicamente para la revisión, y presentaciones sobre entradas de datos, estructura de la población, biología y modelado de la evaluación. El PR hizo varias solicitudes al personal de la CIAT para mejorar su comprensión y consideración de los temas. En el **Anexo E** se detallan estas solicitudes y las respuestas del personal de la CIAT.

Como se señala anteriormente, la actual evaluación del YFT del OPO había sido rechazada por el personal de la CIAT y el modelo revisado no estaba desarrollado por completo al momento de la revisión. Por lo tanto, el PR enfocó su revisión en el desarrollo de entradas de datos, estructura de la población, biología y modelos de evaluación potenciales para el YFT del OPO. En las **secciones 3-6** se describen estas deliberaciones.

El PR destacó el profesionalismo y la experiencia del personal de la CIAT, incluyendo a Carolina Minte-Vera, Haikun Xu, Cleridy Lennert-Cody, Dan Fuller, Kurt Schaefer, Mark Maunder y Alexandre Aires Da Silva, y expresó su gratitud por el arduo trabajo y la voluntad para responder a las solicitudes del PR, y también por su apoyo, provisiones y hospitalidad durante la revisión.

2. COMENTARIOS GENERALES POR EL PANEL DE REVISIÓN

El PR agradeció al personal de la CIAT los documentos y presentaciones disponibles, pero señaló que ayudaría a los revisores externos si la CIAT elaborara documentación estandarizada de los datos de entrada y los análisis subsiguientes realizados sobre los datos de entrada para desarrollar información sobre el ciclo vital, parámetros del modelo, frecuencias de talla, e índices de abundancia. Esta documentación debería incluir una descripción de las fuentes de datos, la selección de datos, supuestos, métodos, diagnósticos del modelo pertinentes y una discusión.

El PR señaló que habían transcurrido más de seis años desde la revisión previa del YFT del OPO (Martell *et al.* 2013), y que las iteraciones y análisis subsiguientes de los datos de entrada y los métodos de evaluación estaban documentados en varias publicaciones principales y actualizaciones de informes. Además, varios análisis de entrada de datos y gráficas de diagnóstico no estaban disponibles para revisión científica o pública fuera del personal de la CIAT. El PR recomendó que el personal de la CIAT proporcionara documentos de referencia de los análisis usados para desarrollar los datos de entrada y el modelo de evaluación de poblaciones para la reunión del Comité en mayo de 2020 como parte de su desarrollo de la evaluación para su consideración para el asesoramiento de ordenación.

El PR también señaló que surgieron preguntas y preocupaciones importantes respecto a los índices de abundancia usados para informar el modelo de evaluación, particularmente el índice de CPUE palangrera (LL) de Japón y el índice de CPUE cerquera (PS) asociada a delfines. Se recomienda encarecidamente una documentación exhaustiva para facilitar la evaluación e interpretación de estos índices.

Finalmente, el PR señaló que se habían seguido varias líneas de investigación relativas a cuestiones de datos del atún patudo (cuya última evaluación también está siendo investigada por el personal de la CIAT) pero que no se habían repetido o priorizado para el atún aleta amarilla en el momento de la revisión. El PR aconsejaría que estos análisis también se extendieran al aleta amarilla.

3. REVISIÓN DE LAS ENTRADAS DE DATOS

Las siguientes entradas a la evaluación fueron evaluadas por el PR: captura, composición por especie, composición por talla y captura por unidad de esfuerzo (CPUE). El PR no revisó ni evaluó la estimación de los descartes.

3.1. Capturas

3.1.1. Captura cerquera

Las capturas de la pesquería de cerco (PS) sobre objetos flotantes (OBJ) están dominadas por el atún barrilete (SKJ), seguido por el patudo (BET); el atún aleta amarilla (YFT) representa una proporción relativamente pequeña de la captura total. La composición por especie de la captura se estima a partir de datos de muestreo en puerto debido a la necesidad de una especiación exacta de los peces más pequeños (es decir, de menos de 70 cm), principalmente debido a las similitudes entre el BET y el YFT pequeños. El personal de la CIAT informó al PR que los muestreos en puerto basados en especies han proporcionado estimaciones fiables de la captura de especies de la pesquería PS-OBJ desde 2000.

Las capturas más elevadas de BET de esta pesquería desde mediados de la década de 1990 han coincidido con la estimación de reclutamiento alto en el modelo de evaluación del BET. La revisión del modelo de evaluación del BET del OPO identificó la posibilidad de que la incertidumbre en las estimaciones de la

captura esté influyendo en la tendencia del reclutamiento de BET. Los modelos de evaluación del YFT del OPO han estimado niveles generales más bajos de reclutamiento a partir de 2000 en adelante. Existe la posibilidad de que la subestimación de las capturas de aleta amarilla de la pesquería PS-OBJ esté influyendo en las estimaciones del reclutamiento de aleta amarilla, aunque no existe un vínculo fuerte con una sobrestimación correspondiente de las capturas de patudo.

No obstante, dado que existen posibles problemas relacionados con la fiabilidad del muestreo de capturas de atunes más pequeños de capturas cerqueras, el PR recomendó que:

- Se evalúe la sensibilidad a la proporción actual de especies a través de niveles alternativos de captura de aleta amarilla de la pesquería OBJ.

Por ejemplo, el análisis podría determinar la magnitud del cambio en la captura de YFT por la pesquería PS OBJ que sería necesaria para eliminar el cambio drástico en el reclutamiento de aleta amarilla que sucedió alrededor de 2000.

Alternativamente, el cambio en el nivel estimado de reclutamiento de aleta amarilla desde 2000 podría estar asociado a un cambio ambiental (es decir, cambio de régimen). Por lo tanto, el PR recomendó que:

- Se debería realizar una evaluación de los mecanismos ambientales que pueden influir en el reclutamiento de aleta amarilla una vez concluida la nueva evaluación.

3.1.2. Captura palangrera

Los datos de captura y esfuerzo de palangre (LL) son reportados a la Secretaría de la CIAT en una escala agrupada de 5 x 5 grados, aunque algunos países proveen datos en resoluciones más finas. El PR no planteó ninguna preocupación específica sobre la fiabilidad de las capturas de palangre agrupadas, dado que estas capturas suelen estar bien determinadas en las OROP atuneras, y los individuos son capturados en una talla en la que la identificación de especies suele ser acertada.

Los datos operacionales de captura y esfuerzo, incluyendo el identificador del buque, para las flotas clave no son siempre accesibles para el personal de la CIAT, o solo están disponibles por tiempo limitado o bajo limitaciones específicas (por ejemplo, colaborar con personal visitante). Dada la influencia de los índices de CPUE LL estandarizados en la evaluación del YFT del OPO, el PR recomendó que:

- El personal de la CIAT tenga acceso continuo a un conjunto actualizado de datos operacionales de captura y esfuerzo para todas las flotas que operan en el área de la convención para asegurar que se puedan desarrollar análisis robustos e insesgados de este conjunto de datos. Este conjunto de datos debería incluir el identificador del buque y otras variables operacionales, como anzuelos entre flotadores, para informar la estandarización, idealmente para la serie de tiempo completa.

3.2. Composición por especie de la captura

El PR señaló que la composición por especie de la pesquería de cerco era más incierta antes de 2000. El PR reconoció que esto podría haber introducido sesgos no cuantificados a la evaluación, pero no hizo ninguna recomendación específica.

3.3. Composición por talla

Los datos de composición por talla de palangre provinieron principalmente de la flota japonesa. Los datos de los buques palangreros de entrenamiento de Japón no se incluyeron en el conjunto de datos de la evaluación ya que no se consideraron representativos de las operaciones de pesca comercial ya que los buques de entrenamiento operan en una zona limitada, principalmente al norte de 10N y en la región de Hawái, mientras que las operaciones de pesca comercial están más extendidas.

Los datos de composición por talla de la flota de palangre comercial se han recolectado de varias fuentes,

principalmente de pescadores comerciales y, en los últimos años, de observadores. Se desconoce la fuente de datos de talla recolectados antes de principios de la década de 1980. La mayor parte de los datos de talla del periodo inicial son datos de peso (kg). Estos datos han sido rechazados previamente del modelo de evaluación ya que no era posible conciliar las diferencias entre el peso y talla de peces individuales.

La mayor parte de los datos de composición por talla de palangre desde principios de la década de 1980 hasta la década de 2010 fue recolectada por pescadores comerciales. Una comparación de la talla de los peces a partir de múltiples fuentes de datos (es decir, peso promedio de los peces convertidos en talla de los pescadores comerciales, peso promedio de las bitácoras y del muestreo de peso) sugiere que puede haber habido un posible sesgo positivo en los datos de composición por talla recolectados por pescadores comerciales.

El personal de la CIAT señaló un aumento reciente en la talla promedio de los peces capturados con palangre. Ese aumento puede brindar una señal influyente de una disminución del esfuerzo (impacto) de pesca al modelo de evaluación. Por lo tanto, el PR recomendó:

- Verificar que no se deba a un problema en la recolección o tratamiento de los datos debería ser una prioridad.

El PR también consideró si el aumento de la talla promedio podría corresponder, al menos en parte, a una transición en 2010 de la recopilación de medidas de talla principalmente por las tripulaciones de los buques a los observadores. Sin embargo, esta hipótesis fue rechazada porque un análisis sugirió que las tripulaciones de los buques muestreaban peces más grandes que los observadores, al menos en los años 2011-2013, cuando se disponía de datos tanto de los observadores como de las tripulaciones de los buques. Otras hipótesis incluían cambios en el área de operación o cambios en la configuración de las artes de pesca, pero no se presentaron pruebas concluyentes.

Se disponía de los datos de composición por talla LL por sexo desde aproximadamente 1995, con una proporción de sexos sesgada hacia los machos de tallas grandes. Se desconoce la fiabilidad de la determinación de sexo ya que se señaló que las gónadas no suelen ser extraídas de la cavidad intestinal durante el procesamiento a bordo.

Si bien el modelo de evaluación actual está estructurado por sexo, no se ajustó a los datos de composición por talla por sexo. No hubo más discusión sobre el tema de la proporción de sexos, pero el PR recomendó que:

- Se investigara más a fondo el sesgo en la proporción de sexos si el personal de la CIAT quisiera avanzar hacia un desglose adicional por sexo para componentes específicos del modelo de evaluación.

3.3.1. Método de VAST para estandarizar la composición

Para estimar la composición por talla anual a escala de población para la pesquería PS asociada a delfines (DEL), el personal de la CIAT aplicó el paquete de *software* de VAST. La intención era incorporar las composiciones por talla a escala de población resultantes en el nuevo modelo de evaluación. El PR señaló que parecía haber datos limitados disponibles antes de principios de la década de 1980, con la proporción de lances sobre delfines inferior al umbral de 75% usado para identificar los buques durante todo el periodo de tiempo; durante este periodo la pesquería estaba compuesta principalmente por peces pequeños. Por lo tanto, el PR recomendó que:

- Se requieren análisis adicionales de los datos de composición por talla usando el método de VAST

antes de que estos índices puedan considerarse robustos, y el equipo de evaluación debería considerar la selección de los buques a incluir en el análisis, incluyendo la proporción de lances DEL por cada buque en un periodo de tiempo apropiado (por ejemplo, mensual o trimestral) como covariable potencial, y evaluar el potencial de confusión espacial en los análisis resultantes.

En general, el método de VAST estimó una población vulnerable (DEL PS) con una proporción mayor de peces grandes con respecto a la composición por talla nominal (a escala de captura). El PR recomendó:

- Comprender la fuente de este efecto antes de que estos índices puedan considerarse robustos para su uso en un modelo de evaluación.

Por ejemplo, podría derivarse de la interpolación entre áreas en las que la CPUE de los peces más grandes era mayor (y el nivel de captura era bajo), pero también de los posibles efectos de borde o de los supuestos de autocorrelación espacial uniforme a lo largo del dominio.

El PR señaló que también existe una estructura bimodal persistente en las composiciones por talla estandarizadas derivadas de los datos de la pesquería PS. Esto parece estar relacionado con la combinación de datos de pesquerías que operan por separado con peces más pequeños y más grandes, y capturas totales más bajas de peces en el rango de talla intermedia. Esto también podría estar relacionado con la estructura espacial de las pesquerías compuestas. De ser así, el PR sugirió que:

- Podría ser necesario ajustar a la composición por talla de la población usando una función de selectividad bimodal.

Surgieron problemas más pronunciados de una aplicación de VAST similar para los datos de composición por talla LL. El análisis tenía por objeto generar estimaciones anuales-trimestrales de la composición por talla de la población en toda la superficie espacial del OPO. La distribución espacial de estos datos fue limitada y variable entre intervalos de tiempo. Las composiciones por talla nominal y de VAST fueron muy similares, lo que sugiere que VAST interpoló una composición por talla relativamente homogénea a lo largo de la distribución espacial del análisis (en cada intervalo temporal). Esto fue bastante sorprendente en vista de los indicios de variación espacial en la talla del YFT capturado con palangre en el OPO entero con peces más grandes hacia el sureste de la pesquería. El PR esperaba que el modelo de VAST interpolara estos patrones para tomar en cuenta las diferencias en la distribución espacial del muestreo de tallas en cada intervalo temporal y, por lo tanto, generara composiciones por talla diferentes de la distribución nominal. No estaba claro que la implementación actual del modelo de VAST tuviera suficientemente en cuenta las diferencias espaciales en la composición por talla, al grado de que el análisis había eliminado las diferencias en la selectividad de la pesquería que podrían estar asociadas a los cambios en la distribución espacial de la pesquería (y/o de la recolección de muestras).

- En general, el PR recomendó el desarrollo de resúmenes adicionales para los datos de entrada, así como la inclusión de gráficas de diagnóstico estándar, desagrupadas por espacio y tiempo, para informar sobre el desempeño del modelo.

3.4. Índices de abundancia (CPUE)

La evaluación anterior del YFT del OPO incorporó dos conjuntos principales de índices de CPUE: índices de CPUE LL del sur derivados mediante el uso de un MLG, y la CPUE nominal (captura por día) de la pesquería cerquera DEL y NOA.

3.4.1. Palangre

En preparación para la evaluación de 2020, también se usó VAST para estandarizar la CPUE de la pesquería LL. El modelo de VAST de la pesquería LL incluye directamente el efecto de los anzuelos entre flotadores (AEF), sin embargo, la serie de tiempo está dividida en dos bloques de tiempo, divididos a principios de la

década de 1990 para tomar en cuenta los cambios de lances LL someros a profundos (como indican los AEF). El efecto de AEF sigue influyendo en la primera parte de la serie de tiempo.

En vista del importante cambio en la metodología de estandarización, el PR recomendó que:

- Sería informativo contar con una comparación directa entre los índices de VAST y los índices de MLG más tradicionales para determinar si el nuevo método de estandarización está influyendo, o no, en el conjunto final de índices de abundancia. Éste debería ser uno de los análisis de sensibilidad reportados para la evaluación de 2020.

La CPUE de la pesquería LL parecía estar relacionada con ENOS, y el PR señaló que los índices podrían estar reflejando cambios en la capturabilidad o en la abundancia. Si bien el PR señaló que se debería preferir un enfoque geoestadístico como VAST en lugar de un MLG con un efecto de celdas fijas para tomar en cuenta los cambios en el hábitat, los modelos geoestadísticos no resuelven la cuestión de diferenciar entre los efectos de abundancia y capturabilidad cuando no existen covariables representativas de uno u otro. Por ejemplo, es probable que la capturabilidad del aleta amarilla varíe con los cambios en la profundidad de la termoclina u otras condiciones oceanográficas relacionadas con la profundidad (aunque no en la misma medida que el patudo), pero es probable que el reclutamiento y el desplazamiento se vean afectados por otras variables oceanográficas (por ejemplo, TSM) que cambian durante los eventos de ENOS.

No se presentó ningún resumen de los datos de entrada para revisar adecuadamente la distribución del esfuerzo LL sobre la extensión espacial del análisis de VAST. Es evidente que la pesquería se ha contraído considerablemente y esto se ha acelerado en la última década. Esto se refleja parcialmente en los CV más elevados de los índices de años recientes, aunque el PR señaló que los CV eran todavía bastante bajos (< 10-15%) durante la mayor parte del periodo, excepto en los años terminales.

El PR cuestionó el supuesto de que el uso de VAST es una forma de control para el muestreo preferencial, y señaló que podría no ser válido. VAST, como la mayoría de los modelos geoestadísticos, supone que la variable de respuesta es independiente de la intensidad de muestreo. Este es un problema clave en el uso de la CPUE dependiente de la pesca como sustituto de la abundancia, dado que el esfuerzo de pesca suele concentrarse en áreas con tasas de captura más elevadas de las especies de interés. En el OPO, este problema, junto con la contracción del esfuerzo LL, aumenta el riesgo de que el método de VAST aplicado en su forma actual resulte en estimadores sesgados de las cantidades de interés que se modelan. Investigaciones recientes sobre este tema en el contexto de la pesca muestran que la densidad en las áreas con muestreo deficiente suele ser sobrestimada, lo que resulta en índices de CPUE hiperestables. El muestreo preferencial es una plaga para los análisis de datos dependientes de la pesca en la mayoría de las pesquerías y tomar en cuenta el muestreo preferencial sigue siendo un área activa de investigación en geoestadística. Según sabe el PR, ningún enfoque analítico aplicado a los sistemas de pesca ha logrado tomarlo en cuenta. Por lo tanto:

- El PR instó a la precaución en el uso e interpretación de índices derivados de VAST como estimadores poblacionales imparciales de abundancia y composición por talla. Cuando sea posible, el personal debería explorar si el modelo es sensible a los patrones de muestreo preferencial que ocurren en el OPO.
- Una opción sería desarrollar un modelo de simulación que combine la dinámica del esfuerzo LL japonés con escenarios de distribución espacial de la abundancia del aleta amarilla a lo largo del tiempo.

El PR también señaló que la incertidumbre en los índices de CPUE LL en el año terminal del modelo fue una de las razones para rechazar la evaluación anterior ya que cambiaba la condición estimada de la población. Sin embargo, el PR consideró que es poco probable que las deficiencias en el conjunto de datos

reciente se mitiguen adecuadamente con la aplicación del método analítico de VAST. Los datos de captura y esfuerzo LL tuvieron una cobertura espacial limitada durante la última década, particularmente en el área oriental de la pesquería. Esto significa que es poco probable que las predicciones para áreas grandes del dominio espacial sean fiables debido a la falta de observaciones en la última década. A fin de aclarar mejor la influencia del enfoque analítico en los índices finales, el PR recomendó que el personal de la CIAT:

- Considere desarrollar índices separados para áreas de la pesquería que han sido cubiertas por la pesquería a lo largo del periodo de datos (norte, suroeste, sureste). Estas áreas podrían usarse para generar tres conjuntos de índices de CPUE usando métodos tradicionales de estandarización. Los índices resultantes podrían entonces compararse más directamente con los resultados del modelo de VAST.

3.4.2. Red de cerco

Para la evaluación de 2020, el personal de la CIAT propuso aplicar un índice de CPUE derivado de la pesquería DEL PS para monitorear la abundancia del YFT en la región norte de la pesquería. El índice de CPUE propuesto se derivó usando una implementación simple en VAST. El PR tuvo varias preocupaciones sobre este enfoque, como se señaló durante la revisión.

El PR tuvo preocupaciones generales sobre la conveniencia de usar los datos de captura y esfuerzo de la pesquería DEL PS, especialmente considerando que el componente de búsqueda de la operación de pesca no estaba cuantificado adecuadamente y que la eficacia relativa de la operación de pesca había aumentado considerablemente con la introducción de nueva tecnología. El PR también discutió cuestiones relacionadas con la evolución de la pesquería de cerco asociada al desarrollo de la pesquería sobre objetos flotantes, y señaló que éstas podrían influir en la operación de la pesquería DEL PS. Se presentó información insuficiente para evaluar la continuidad de la operación de la pesquería (por ejemplo, con respecto a la longevidad de los buques individuales en la pesquería, la introducción de nueva tecnología y los cambios en la distribución del esfuerzo por tipo de lance [DEL vs OBJ]). Por lo tanto, el PR recomendó que:

- Se realice una caracterización exhaustiva del conjunto de datos de CPUE para permitir una evaluación de la utilidad de los índices de CPUE.

El PR también señaló que la “pesquería” DEL se había definido como aquellos buques que habían reportado al menos 75% de sus lances dirigidos a manadas de delfines cuando se agruparon los registros en la serie de tiempo completa. El PR expresó su preocupación ante la posibilidad que las prácticas reales de selección de objetivo de pesca por buques individuales cambiaran con el tiempo, y que la definición basada en el promedio a largo plazo pudiera no detectar cambios paulatinos en los objetivos de pesca y pudiera sesgar las tendencias de la CPUE estandarizada si no se toman en cuenta. El PR solicitó una gráfica diagnóstica durante la reunión que ilustrara la cantidad de esfuerzo por cada buque y la proporción de ese esfuerzo identificada como lances DEL. Esta gráfica sugirió una fuerte correlación entre el periodo pico de CPUE de aleta amarilla y una alta proporción de lances DEL por la mayoría de los buques. Por lo tanto:

- El PR recomendó que la clasificación de la pesquería DEL sea reconsiderada por el equipo de evaluación, y que se exploren definiciones alternativas para la pesquería DEL a partir de proporciones trimestrales o anuales de lances DEL. Además, algunos miembros del PR sugirieron que los analistas investiguen el uso de la proporción de lances DEL como variable explicativa en el análisis.

El modelo de CPUE de VAST para los datos DEL PS que se presentó al PR también mostró algunos fuertes patrones espaciales de los residuales, con una mayor prevalencia aparente de residuales negativos en los nudos que cubren las áreas más grandes en los bordes del dominio espacial. El PR recomendó que el personal de la CIAT explorara la razón de los fuertes patrones de residuales del análisis de VAST. Si el patrón de residuales es causado por el supuesto que la función de covarianza espacial no cambia a lo largo

del tiempo y el espacio, el PR recomendó que:

- Los analistas investiguen la división del dominio espacial en dos o tres componentes, ajustando el modelo por separado (como ya lo ha hecho la CIAT durante el análisis exploratorio), y verificando que el patrón de residuales se redujo.
- Incluir covariables oceanográficas pertinentes a la biología del atún aleta amarilla, como una covariable de densidad, también podría disminuir la dependencia del modelo de VAST de los efectos espaciotemporales estimados para los nudos.

El PR consideró que las covariables oceanográficas podrían incluirse como efectos de densidad o capturabilidad, pero concluyó que, debido a la falta de covariables de capturabilidad supuestas en el modelo actual, las predicciones del modelo mejorarían con la adición de covariables de densidad a menos que hubiera pruebas sólidas de que una covariable oceanográfica afectara la capturabilidad. Por lo tanto, a fin de evitar la confusión que se produce si una covariable afecta tanto la densidad como la capturabilidad, el PR acordó que las variables oceanográficas deberían incluirse generalmente como covariables de densidad.

El PR señaló que los patrones espaciales de los residuales eran generalmente menos severos en las áreas de la superficie donde se concentraban los nudos. Por lo tanto, el PR recomendó:

- Podría ser útil redefinir la malla para distribuir los nudos uniformemente en el dominio espacial, junto con un aumento en la resolución de la malla tanto para el área del núcleo como para los bordes exteriores del dominio espacial.
- Se debería examinar una gráfica del valor predicho y la varianza de cada nudo para verificar que no existan efectos de borde en la superficie estimada.

El PR señaló que el dominio espacial de la pesquería DEL PS era limitado antes de 1980 y que la cobertura espacial era bastante deficiente durante 1980-1985. La pesquería DEL PS también se extendió hacia el sur y hacia el oeste a partir de 2000. El PR consideró si la expansión espacial estaba relacionada con la interacción con la pesquería OBJ en expansión, y si la expansión de esa pesquería podría estar influyendo en las operaciones de la pesquería DEL PS. El PR señaló que:

- Una mejor comprensión de la interacción entre los componentes DEL y OBJ de la pesquería PS ayudaría en la implementación e interpretación del modelo.

Según entiende el PR, todas las celdas espaciales con al menos un registro de esfuerzo de pesca fueron incluidas en el dominio espacial para el análisis. A fin de asegurar que las celdas con observación deficiente no influyan excesivamente las predicciones finales, el PR recomendó:

- Es importante volver a ejecutar el modelo con un umbral diferente para la inclusión de celdas en el dominio espacial (por ejemplo, esfuerzo mínimo en lances durante un número determinado de trimestres).
- Idealmente, el dominio espacial debería ser refinado para excluir las celdas de borde que solo eran parte de la pesquería de manera anecdótica, pero se necesitaría determinar el umbral de inclusión exacto mediante la exploración de los insumos y de la sensibilidad de las predicciones del modelo a la definición del dominio espacial.

Se señaló que los índices de los primeros años (antes de 1980) tenían un CV asociado más alto y que esos índices influirían menos en la verosimilitud. El PR señaló que, en el caso del modelo del norte, no se disponía de otros índices de CPUE para ese periodo de tiempo, por lo que el modelo se ajustará a esos índices de todos modos. Sin embargo, el PR recomendó las siguientes exploraciones para facilitar la evaluación e interpretación de los índices cerqueros:

- Excluir los primeros índices del modelo para evaluar la influencia de estos datos.
- Si existen diferencias significativas en la operación de la pesquería en varios intervalos (por ejemplo, antes y después de la expansión de la pesquería OBJ), sería útil dividir los índices DEL PS en periodos de tiempo para tomar en cuenta los posibles cambios en la capturabilidad.

El PR señaló que un análisis de reducción (*depletion*) indicó que la capturabilidad de la pesquería PS podría ser variable; por ejemplo, el pico elevado en la CPUE PS podría haber sido causado por el aumento de la capturabilidad durante un periodo de mayor abundancia. Esto podría haber estado relacionado con la disponibilidad de YFT en la superficie y/o la proximidad de la biomasa al área principal de operación de la flota. De manera similar, el PR señaló que durante el periodo anterior a 1985 la proporción de lances no DEL fue considerablemente mayor. Esto refuerza la necesidad de excluir esos datos tempranos del análisis. Hubo también un periodo a mediados y finales de la década de 2000 en el que la proporción de lances DEL fue menor para la mayoría de los buques. Por consiguiente, el PR concluyó que:

- Sería preferible definir los criterios de selección de buque con base en la proporción de lances DEL para cada buque en escalas de tiempo más finas (por ejemplo, año o trimestre), como se comentó anteriormente.

El PR también hizo sugerencias sobre procedimientos generales y gráficas de diagnóstico pertinentes para la estandarización de la CPUE LL y DEL PS.

- Las gráficas escalonadas (Bentley *et al.* 2011) que muestran el efecto de capas consecutivas de filtrado y de los umbrales de filtrado candidatos sobre el índice estandarizado serían herramientas útiles para identificar si el índice de abundancia es sensible a las decisiones sobre la preparación de los datos.
- Teniendo en cuenta que los residuales en todo el dominio del modelo suelen comportarse bien, deberían considerarse gráficas que muestren los residuales desagrupados por espacio y tiempo siempre que sea posible. Por ejemplo, se deberían examinar las gráficas de QQ anuales de los residuales para verificar si existe una tendencia temporal en el ajuste del modelo.
- A fin de verificar si los efectos estacionales se están tomando en cuenta adecuadamente, se podría ajustar el modelo por separado para cada trimestre. Si se mejoran los diagnósticos para los modelos trimestrales, se requeriría algún tipo de flexibilidad adicional en el modelo de todos los trimestres. Una opción podría ser incluir covariables oceanográficas que varían estacionalmente, pero de manera desigual, en el dominio espacial (por ejemplo, TSM).
- Los residuales desagrupados en espacio y tiempo deberían compararse entre el método de VAST y el de MLG tradicional como apoyo a la metodología propuesta.
- Por último, los índices de CPUE PS y LL están compuestos por los dos componentes del modelo (probabilidad y magnitud de la captura). Sería importante comprender cuál componente del modelo es el que más influye en la determinación de los índices anuales/trimestrales compuestos. Esto se presentaría mejor como las tendencias anuales de cada componente (normalizadas) y el índice compuesto.

4. DESPLAZAMIENTO Y ESTRUCTURA DE LA POBLACIÓN

4.1. Desplazamiento

La CIAT presentó algunas pruebas de residencia regional de individuos marcados en las regiones del norte y del sur. Más específicamente, la mayoría de los desplazamientos lineales de las marcas convencionales parecieron ser menores de 1000 nm. Este desplazamiento limitado fue apoyado por datos de marcas archivadoras que también destacaron el potencial de residencia regional, especialmente alrededor de lugares específicos como las Islas Revillagigedo. Sin embargo, la cantidad exacta de intercambio norte-sur de las marcas convencionales fue difícil de evaluar con base en las visualizaciones presentadas al PR. También

parece que hubo liberaciones limitadas desde las áreas de alta mar.

También hubo algunas pruebas de un intercambio limitado entre las "poblaciones" del sur y del norte, a partir de los patrones en los datos de CPUE y de composición por talla. Sin embargo, no estaba claro que los patrones en los datos de CPUE y de composición por talla se debieran a un proceso biológico real, como el desplazamiento.

A fin de mejorar la comprensión del desplazamiento del YFT y la estructura de la población, el PR consideró que:

- Se debería diseñar y realizar, de manera prioritaria, un programa de marcado de YFT, especialmente en las áreas de alta mar y en conjunto con marcado de otolitos con oxitetraciclina. Considerando las limitaciones logísticas del marcado de aletas amarillas en las áreas de alta mar, el PR señaló que colaborar con la Comisión de Pesca del Pacífico Occidental y Central (WCPFC) podría ser una vía productiva para compartir las cargas logísticas y financieras de la organización de crueros de marcado en alta mar.

4.2. Hipótesis de la estructura de la población

Las evaluaciones previas han supuesto que el YFT en el OPO consistía en una sola población bien mezclada y que cualquier heterogeneidad espacial se tomaba en cuenta usando el enfoque de "flotas como áreas". Sin embargo, este enfoque supone que los individuos extraídos en cualquier área afectan a la población en todas las áreas al mismo tiempo. El trabajo exploratorio sobre el desfase temporal entre los índices DEL y LL (*cf. Xu - Hypotheses for the difference between PS and LL indices*) sugirió que las fuertes cohortes indicadas por los índices DEL y LL en 2001 y 2000, respectivamente, se debían a diferentes eventos de reclutamiento en 1999 y 1998, respectivamente. Esta diferencia en los eventos de reclutamiento entre los índices DEL y LL fue sugerida además como evidencia de la estructura espacial del YFT en el OPO, estructuración que no se toma en cuenta en la evaluación actual. Además, un análisis de los datos de composición por talla de las diferentes pesquerías y su relación con los índices DEL y LL (*cf. Minte-Vera et al. - Development of a new benchmark model for yellowfin tuna in the EPO*) sugirió que los datos de composición por talla de algunas pesquerías (algunas pesquerías DEL, algunas OBJ y algunas NOA) eran más consistentes con el índice DEL mientras que otras pesquerías (LL, algunas pesquerías OBJ y algunas NOA) eran más consistentes con el índice LL.

El personal de la CIAT también presentó pruebas de la estructura de la población del YFT para el OPO. Hubo un solo estudio genético con múltiples sitios en el OPO que no mostró diferenciación de las poblaciones en un gradiente norte-sur (Ward *et al.* 1997). Sin embargo, el PR señala que este estudio usó marcadores aloenzimáticos y mitocondriales, que son conocidos por ser muy sensibles a pequeñas cantidades de intercambio entre poblaciones¹. Por lo tanto, el PR recomendó:

- Un análisis genético actualizado que utilice marcadores moleculares modernos como SNP (polimorfismo de nucleótido único) de múltiples ubicaciones a lo largo de un gradiente norte-sur en el OPO podría brindar más información sobre la estructura de la población.

El personal de la CIAT presentó información adicional sobre las características conjuntas de las variables ambientales en todo el OPO (*cf. Minte-Vera et al. - Stock and spatial structure of yellowfin tuna in the Pacific Ocean*; y una respuesta a un PR re: *SST patterns in different quarters of the year*). El personal de la CIAT informó que el YFT desova continuamente en lugares con TSM de $\geq 24^{\circ}\text{C}$, y que el desove es óptimo

¹ Waples, R. S., & Gaggiotti, O. (2006). INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular ecology*, 15(6), 1419-1439.

cuando la TSM es de 28°C.

Dado que el índice DEL se basó principalmente en datos del norte y el índice LL se basó principalmente en datos del sur, se propuso que el YFT en el OPO se modelara como dos poblaciones separadas: una población "norte" y una población "sur". Por lo tanto, se desarrollaron y presentaron dos modelos de referencia. El modelo "norte" se basó en el índice DEL, e incluyó las pesquerías que se identificaron como consistentes con el índice DEL. El modelo "sur" se basó en el índice LL y las pesquerías que se identificaron como consistentes con el índice LL. Ambos modelos eran relativamente independientes entre sí y no compartían datos. Los parámetros biológicos (por ejemplo, crecimiento, mortalidad natural, relación población-reclutamiento) eran idénticos para ambos modelos. Aunque los modelos eran nominalmente espaciales, las definiciones de área variaban por tipo de pesquería y estación. Por ejemplo, las pesquerías DEL se dividieron aproximadamente a lo largo del paralelo 5°N en los modelos "norte" y "sur". En cambio, todas las pesquerías LL en el OPO entero, independientemente de la latitud, estaban en el modelo "sur", y solo una pesquería OBJ en los trimestres 2 y 3 se colocó en el modelo "norte".

Con base en la información presentada, el PR consideró que el YFT en el OPO podría consistir en dos poblaciones y que se podría explorar más a fondo una hipótesis de dos poblaciones. Sin embargo, el PR concluyó que:

- También se debería seguir desarrollando un modelo de una sola población como hipótesis alternativa porque se consideró que las pruebas que apoyan la hipótesis de dos poblaciones son sugerentes, más que concluyentes.

Además, la hipótesis de una sola población bien mezclada y la hipótesis de dos poblaciones independientes pueden considerarse casos extremos de un continuo, mientras que la estructura real de la población es probablemente un punto intermedio entre esos extremos. En resumen, no pareció haber pruebas suficientes para apoyar la hipótesis de dos poblaciones como la única hipótesis plausible.

Existieron más pruebas que sugerían que el YFT en el OPO estaba en algún punto entre una sola población bien mezclada y dos poblaciones independientes. Una comparación de los reclutamientos estimados para los modelos "norte" y "sur" indicó que la escala a largo plazo (multianual) y la tendencia del reclutamiento para ambos modelos independientes eran muy similares. Sin embargo, el reclutamiento para un trimestre específico podía diferir considerablemente en magnitud entre los modelos. Esto sugirió que la fuente a gran escala y/o las influencias ambientales en el reclutamiento para el "norte" y el "sur" eran similares, pero la repartición entre el "norte" y el "sur" podría ser altamente variable. Esto no descarta la hipótesis de dos poblaciones independientes porque las tendencias de reclutamiento similares a largo plazo podrían deberse a condiciones ambientales a gran escala que influyen en las dos poblaciones simultáneamente. No obstante, existe un riesgo no despreciable de que exista un grupo común de reclutas que se reparte entre las dos áreas. Aunque las consecuencias de ordenación no entran dentro del ámbito del PR, éste consideró que un análisis de riesgos podría ser útil para informar a los gestores, específicamente:

- El riesgo de cambiar a una hipótesis de dos poblaciones si la hipótesis de una población fuera realmente cierta y, al contrario, quedarse con la hipótesis de una población si la hipótesis de dos poblaciones fuera cierta, debería explorarse mejor antes de pasar unilateralmente a una hipótesis de dos poblaciones.

Además de las hipótesis de una y dos poblaciones presentadas por el personal de la CIAT, el PR también sugirió que se exploraran varias vías para tomar mejor en cuenta la estructura de la población en el modelo de evaluación. Éstas incluyen, entre otras:

- Una hipótesis de una población, pero suponiendo que solamente los índices LL son consistentes;
- Una hipótesis de una población, pero suponiendo que solamente los índices DEL son consistentes;

- Una hipótesis de dos poblaciones basada en dominios espaciales del norte y del sur, que pueden variar según la estación;
- Una hipótesis de dos poblaciones con una población basada en YFT asociado a delfines y otra población de YFT no asociado a delfines; y
- Una hipótesis de un grupo común de reclutas que se reparte entre las dos áreas.

Aunque los modelos de referencia “norte” y “sur” eran modelos nominalmente espaciales, los modelos fueron más consistentes con un modelo para el YFT asociado a delfines (modelo “norte”) y otro modelo para el YFT no asociado a delfines (modelo “sur”). Esto se debió a que el modelo “norte” estaba basado en pesquerías con datos más consistentes con el índice DEL, mientras que el modelo “sur” estaba basado en pesquerías con datos más consistentes con el índice LL. Las pesquerías DEL estaban predominantemente en el hemisferio norte y todas las pesquerías del modelo “norte” estaban en el norte. Sin embargo, las pesquerías y el índice LL en el modelo “sur” incluían datos del OPO entero. No hubo pruebas biológicas para apoyar la hipótesis de que los YFT asociados a delfines eran diferentes a los YFT no asociados a delfines. Por lo tanto, el PR recomendó que:

- Los datos y los modelos sean consistentes con hipótesis explícitas y plausibles o modelos conceptuales. Por ejemplo, si un modelo es nominalmente un modelo espacial, el dominio espacial debería definirse, y la preparación de datos y los modelos deberían ser consistentes para ese dominio espacial.

5. BIOLOGÍA DEL ATÚN ALETA AMARILLA DEL OPO

5.1. Crecimiento

La edad y el crecimiento del YFT en el OPO se basaron en incrementos diarios en otolitos validados con OTC hasta la edad 4. El PR señaló que la forma de la curva de crecimiento era esencialmente lineal hasta la edad 4, con pocos datos de edad para informar la relación de crecimiento después de la edad 4, ya sea de incrementos anuales o diarios. Es probable que esto resulte en un modelo de crecimiento sesgado y altamente incierto, especialmente el parámetro L_{inf} . A diferencia del modelo de crecimiento del BET, actualmente no se cuentan con suficientes datos de marcado de YFT grandes para reducir el sesgo y la incertidumbre del parámetro L_{inf} .

El personal de la CIAT ha concluido que el crecimiento del YFT en el OPO se estima mejor usando incrementos diarios en otolitos hasta la edad 4, y ha rechazado las conclusiones de otros estudios realizados sobre el patudo y el aleta amarilla en el Pacífico y otros océanos que sugieren que los incrementos anuales en otolitos podrían ser una medida de crecimiento alternativa plausible para los atunes tropicales, particularmente en tallas más grandes. Varios estudios han concluido que, en el caso de algunas especies de vida más larga, los incrementos diarios se vuelven muy estrechos una vez que la talla del pez se acerca a L_{inf} , lo que dificulta distinguirlos y cuantificarlos, y que puede haber áreas en un otolito que carecen de incrementos diarios o donde son difíciles de interpretar (por ejemplo, Sardenne *et al.* 2015). Esto podría causar una subestimación de la edad verdadera de los individuos más viejos.

No se dispone de una validación concluyente de deposición de incrementos anuales para el YFT. No obstante, un estudio preliminar reciente indicó que el YFT en el Atlántico deposita una sola zona opaca y translúcida cada año (Ailloud *et al.*, 2020). Si bien estos resultados fueron de alcance limitado, los autores concluyeron que el uso de incrementos diarios subestimó la edad del YFT del Atlántico de más de 55 cm TF, ya sea porque los incrementos no se depositaban sistemáticamente a diario y/o eran difíciles de interpretar. Un estudio realizado en el Océano Pacífico occidental y central (OPOC) por J. Farley (IATTC, 2019) también encontró que se observaron incrementos diarios claros en otolitos cerca del primordio, pero describió una interrupción a los ~150-180 incrementos que sugirió que los conteos de incrementos diarios no eran útiles para determinar la edad de aletas amarillas grandes/viejos en el OPOC. En vista de lo anterior,

el PR recomendó encarecidamente:

- Continuar los esfuerzos de investigación sobre la edad y el crecimiento del YFT, especialmente para YFT >150 cm.
- Dada la incertidumbre en el crecimiento, el PR apoya la estimación de los parámetros de crecimiento dentro del modelo integrado. El PR también sugiere que se consideren fuentes de datos adicionales para informar el ajuste de la curva de crecimiento (por ejemplo, datos de marcado);
- El PR sugirió una mayor colaboración con los científicos del OPOC que trabajan en el crecimiento de los atunes tropicales;
- El PR sugirió la exploración de métodos híbridos que combinen los incrementos diarios y anuales en las estimaciones finales de edad para equilibrar las fortalezas y debilidades de ambos métodos;
- A largo plazo, diseñar un programa de marcado de YFT en áreas más grandes más allá de las muestreadas actualmente, en conjunto con marcado con OTC; y
- El PR sugirió evaluar los posibles efectos de la densodependencia en el crecimiento.

La información disponible sugiere que las relaciones de crecimiento alternativas siguen siendo preliminares y que esto se debería identificar como una posible fuente de incertidumbre en la evaluación. Sin embargo, el PR señaló que, dada la incertidumbre en la estructura del modelo y en la interpretación espacio-temporal de la CPUE y de las observaciones de frecuencia de talla, los análisis de sensibilidad del modelo a relaciones alternativas de crecimiento deberían ser de prioridad baja en la evaluación de 2020.

5.2. Mortalidad natural (M)

El modelo actual de evaluación del atún aleta amarilla del OPO usó una función de mortalidad derivada de datos de marcado (Hampton, 2000) y supuso que M era específica por edad y sexo (Maunder y Aires-da-Silva, 2012). La M de las hembras aumenta después de la madurez, mientras que la M de los machos no. Se supuso que el cambio en la M de las hembras ocurría a los dos años (8 trimestres) en el aleta amarilla del OPO con base en un examen de la proporción de sexos.

Dada la escasez de información disponible para estimar la mortalidad natural, el PR no hizo recomendaciones específicas con respecto a las formas funcionales usadas en los modelos de población propuestos. Sin embargo, el PR sí señaló que la función de M supuesta será sensible a los cambios en la función de crecimiento supuesta y la proporción de sexos por edad, y que cualquier cambio en la parametrización de estas funciones requerirá que se reestime la M . El PR también recomendó diagnósticos y análisis de sensibilidad, como se describe a continuación.

- Explorar el potencial para estimar M dentro del modelo de evaluación mediante la pre-especificación de M a edad 0 y la estimación de la diferencia (compensación) en M entre hembras y machos maduros (por ejemplo, según lo informado por la proporción de sexos).

El PR también señaló que existe una gama de valores de M usados en evaluaciones del aleta amarilla de otras regiones. Se preferiría seleccionar valores de M que se han derivado de modelos que incorporan una cantidad considerable de datos de marcado que puedan informar al modelo sobre M , al menos para el rango de clases de edad representado en la porción marcada de la población. Dadas estas consideraciones, el PR recomendó que:

- El personal de la CIAT resuma la información disponible sobre las estimaciones de M y crecimiento del atún aleta amarilla, y documente las estimaciones usadas en la evaluación.
- La CIAT estudie relaciones alternativas plausibles, en particular tomando en consideración cualquier correlación entre M y crecimiento.
- Los análisis de sensibilidad del modelo deberían incluir (como mínimo) un nivel alternativo plausible de M que sea inferior al nivel general de M actualmente supuesto.

5.3. Reclutamiento y relación reproductor-recluta

El PR no discutió específicamente la elección de la relación reproductor-recluta, pero sí examinó la sensibilidad del modelo a σ_R . Actualmente, el modelo usa una relación Beverton-Holt con una inclinación fijada en 1.0 y σ_R fijado en 0.6 o 1.0. Debido a la naturaleza exploratoria del trabajo presentado, no fue posible determinar las parametrizaciones más adecuadas de la relación reproductor-recluta. Sin embargo, el PR recomendó que:

- Se consideren los procedimientos estándar de corrección de sesgo de *Stock Synthesis* (por ejemplo, Methot y Taylor, 2011) antes de seleccionar los modelos finales, incluyendo un intento de estimar σ_R .
- En los modelos exploratorios revisados por el PR, parecía haber cierto apoyo a los valores de σ_R superiores a 0.6. Si finalmente se fija σ_R , se recomienda un perfil de verosimilitud sobre σ_R para informar la selección del valor más adecuado para ese parámetro.

El PR también señaló que el valor de inclinación de 1.0 usado para la evaluación de atunes tropicales, incluyendo el YFT, es el más elevado de todas las OROP, pero no consideró más esta cuestión ya que ya había sido cubierta por la reciente revisión externa del atún patudo (CIAT, 2019).

El PR también señaló que los patrones de reclutamiento generales son comparables entre los dos conjuntos de índices de abundancia, aunque hay diferencias en ciertos periodos y estas diferencias son pronunciadas en algunos años (posiblemente/probablemente relacionadas con las condiciones de El Niño). Existe la posibilidad de desarrollar un modelo espacialmente estratificado para acomodarse a las diferencias en el reclutamiento entre las dos regiones de la pesquería, pero el PR también reconoció la mayor complejidad de ese modelo relativa a los supuestos de desplazamiento y el ajuste a escala de la biomasa relativa entre las dos regiones.

6. ESTRUCTURA DEL MODELO

6.1. Estructura de la pesquería

El personal de la CIAT propuso que el YFT en el OPO fuera considerado como dos poblaciones separadas para los fines de la evaluación (ver también la Sección 4 sobre desplazamiento y estructura de la población). Aunque existieron algunas pruebas de dos posibles grupos de reclutamiento (uno en el norte y otro en el sur), el PR solicitó:

- Que se realicen más análisis para evaluar los impactos de ordenación con dos poblaciones en vez de una, dado que seguía siendo muy probable que existiera cierto intercambio de individuos entre las poblaciones.
- Además de las hipótesis de una y dos poblaciones presentadas por el personal de la CIAT, el PR también sugirió que el personal de la CIAT explorara otros ejemplos de modelos e hipótesis plausibles con estructuras poblacionales alternativas (ver Sección 4.2).

El PR señaló que el método de usar bifurcaciones de árboles para desarrollar hipótesis para las pesquerías de cerco basadas en la homogeneidad de los datos de proporciones por edad por estación y celda espacial (celdas de 5x10 grados) era apropiado, pero que los análisis podrían no tomar en cuenta adecuadamente el crecimiento estacional, el reclutamiento variable o las áreas con datos de frecuencia de talla incompletos. Por lo tanto, el PR recomendó que:

- Se desarrollen más análisis de las pesquerías DEL, y sugirió que la CIAT considere convertir la frecuencia de talla a una frecuencia de edad aproximada mediante la curva de crecimiento promedio para probar las definiciones actuales de la pesquería.

El PR señaló la presencia de bimodalidad en algunas de las frecuencias de talla agrupadas en el árbol de

regresión, y sugirió que esta bimodalidad podría ser potencialmente el resultado de la agrupación a través de una escala espacial demasiado grande o podría representar pulsos de reclutamiento. Por lo tanto, el PR recomendó:

- Más análisis de los datos de frecuencia de talla para ayudar a determinar la causa de la bimodalidad.
- Además, el PR señaló que ya se dispone de datos adicionales a escala fina de la flota LL de Japón, y sugirió que se volviera a investigar la definición de las pesquerías LL usando estos nuevos datos.

6.2. Selectividad

El PR señaló que había pruebas sólidas de selectividades en forma de domo en algunas pesquerías basadas en la composición por talla de las muestras. Sin embargo, en el caso de las pesquerías que capturan peces más grandes, no estaba claro si la alta mortalidad natural o la especificación errónea de la curva de crecimiento podrían haber llevado a la selección, posiblemente inapropiada, de selectividades en forma de domo o logísticas. Por lo tanto, el PR señaló que:

- Podría ser necesario usar una forma de selectividad funcional más compleja o una forma no paramétrica (por ejemplo, *splines* cúbicos) para representar adecuadamente la selectividad de la población vulnerable a PS debido a la distribución bimodal de las composiciones por talla de la población.

Las pruebas del modelo con una selectividad con *splines* cúbicos estimaron una selectividad más compleja (no asintótica) y se ajustaron al hombro izquierdo de la distribución. La selectividad más compleja podría estar tomando en cuenta las deficiencias en la composición por talla de la población generada por VAST, que suaviza los datos de talla de pesquerías con peces más pequeños/grandes dependiendo del espacio.

- El PR también señaló que hay pruebas de un aumento reciente en la talla del atún aleta amarilla capturado por la pesquería LL de Japón (ver Sección 3). Esto sugiere que podría ser necesario explorar la selectividad variable con el tiempo en esa pesquería, considerando que también podría ser una característica de una población en respuesta a una reducción de la presión de pesca.

6.3. Capturabilidad

El PR no evaluó información relativa a un aumento de la capturabilidad, ni hizo recomendaciones específicas sobre la capturabilidad.

6.4. Año de inicio de la pesquería

Los modelos propuestos comienzan en 1975, y requieren la estimación de las condiciones iniciales usando únicamente los datos de composición por talla (ya que la sanción por captura de equilibrio inicial se desactiva en la verosimilitud). El PR señaló que las estimaciones iniciales de biomasa y reducción podrían ser sensibles a la especificación de las condiciones iniciales. En algunos casos, los modelos preliminares también estimaron una captura de equilibrio inicial mucho mayor que las capturas históricas observadas de la pesquería antes de 1975. Este comportamiento puede ser indicativo de una especificación errónea del modelo y no debe ser pasado por alto en los modelos finales.

La justificación original para que el modelo comenzara en 1975 fue la disponibilidad de anzuelos por canasta a partir de 1975 para la pesquería LL de Japón y el muestreo en puerto de composición por talla de cerco que comenzó en 1975. Los datos de captura LL y PS de Japón están disponibles desde la década de 1950 y, por lo tanto, podrían usarse para justificar que el modelo comience aun antes. Permitir que el modelo condicione las series de tiempo más largas de extracciones podría mejorar las estimaciones iniciales de reducción, particularmente si las series de tiempo pueden extenderse hasta un periodo en el pasado con extracciones insignificantes. En vista de estas consideraciones:

- El PR recomendó que se explore la sensibilidad del modelo a las condiciones iniciales mediante la estimación de las extracciones anteriores a 1975 (por ejemplo, en la década de 1950) y/o la implementación de una sanción por captura de equilibrio.

7. SENSIBILIDAD DEL MODELO Y DIAGNÓSTICOS

7.1. Ponderación de datos

El PR señaló que la ponderación de datos puede tener una gran influencia sobre los resultados del modelo, particularmente si los componentes de los datos están en conflicto entre sí o son inconsistentes con los supuestos del modelo. La ponderación de los datos en la evaluación del YFT se determinó mediante la desviación estándar lognormal de los índices de CPUE y el tamaño de la muestra multinomial para los datos de composición por talla. Se supusieron las estimaciones de CV producidas por VAST más una constante (para promediar 0.2) para el índice de CPUE LL. Los tamaños iniciales de las muestras de los datos de composición por talla PS se basaron en el número de bodegas muestreadas, mientras que los tamaños iniciales de las muestras de los datos de composición por talla LL se fijaron en valores que representan pesos comparables con la pesquería cerquera principal, pero reteniendo la ponderación relativa de muestras individuales basada en el número de peces muestreados. Se exploró la reponderación de datos usando el procedimiento integrado de reponderación de Francis (2014) en *Stock Synthesis*. El PR no discutió ni recomendó específicamente ningún esquema particular de ponderación de datos, pero se señaló que:

- El CV constante de 0.2 aplicado a los índices de CV parecía injustificado. Las estimaciones anuales de CV están disponibles y se deberían usar en los modelos finales. Si se desea un CV observado más alto en los índices, el PR recomienda que se ajuste la escala de los CV de tal manera que se conserve la variación interanual.

7.2. Diagnósticos

Dado que solo se disponía de modelos preliminares/exploratorios, se presentaron pocos diagnósticos para la revisión. El PR no solicitó un conjunto más completo de diagnósticos durante la reunión porque los modelos presentados sugirieron que aún se necesitaban cambios básicos en los datos y en la parametrización del modelo. Por ejemplo, la falta de ajuste a la composición por talla de VAST indicó un conflicto entre estos datos y otras fuentes de información en el modelo (por ejemplo, índice, datos de composición de la flota). Además, la falta de ajuste a los datos de composición por talla PS sugirió la necesidad de considerar una función de selectividad más flexible (es decir, *splines* cúbicos) para esas flotas.

A fin de examinar la calidad y estabilidad de los modelos finales, el PR recomienda una evaluación exhaustiva de un conjunto de diagnósticos que incluya:

- Diagnósticos mejorados de ajuste del modelo a los datos de composición por talla, incluyendo diagrama de caja de residuales de Pearson por clase de talla, año de observación, y potencialmente cohorte (pero menos probable que sea posible).
- Un examen de la matriz de covarianza para identificar pruebas de formulación inadecuada del modelo. En particular, estas tablas deberían examinarse a detalle para buscar pruebas de coeficientes de variación altos y correlaciones fuertes entre los patrones de selectividad y los parámetros de crecimiento.
- Gráficas de trazos para explorar las estimaciones de los parámetros con respecto a la fase de estimación. En un modelo que funciona debidamente, los parámetros no deberían cambiar mucho después de la fase final de estimación. Cambios grandes después de la fase final de estimación pueden mitigarse reconfigurando las fases de estimación.
- Un análisis de “*jitter*” para examinar el efecto de la variación de los valores de los parámetros

iniciales en los resultados del modelo. Un modelo que funciona debidamente debería converger en una solución global a través de una gama razonable de valores iniciales.

- Perfil de verosimilitud para evaluar el desempeño del modelo a través de una gama de valores para parámetros de entrada clave (por ejemplo, inclinación, R_0 , σ_R). Idealmente, el perfil debería ser una forma funcional suave. Un comportamiento anormal del modelo puede ser resultado de inestabilidad del mismo, y se indica mediante numerosos picos y perfiles de sierra.
- *Bootstrap*: El desempeño de *Stock Synthesis* también puede evaluarse graficando la distribución de las estimaciones de los parámetros y las cantidades derivadas a través de réplicas de *bootstrap*. En un modelo que funciona debidamente, la estimación de verosimilitud máxima debería ser similar al promedio de los *bootstraps*.
- Análisis retrospectivo para identificar los cambios en las estimaciones del tamaño de la población, la mortalidad por pesca, el reclutamiento, etc., que se producen a medida que se añaden o eliminan años de datos en el modelo (por ejemplo, Hurtado-Ferro *et al.*, 2015). La gravedad del patrón retrospectivo debería cuantificarse usando la p de Mohn.
- El *hindcasting* (Kell *et al.*, 2016) también puede ser un diagnóstico útil del modelo. En un *hindcast*, el modelo se ajusta a la primera parte de una serie de tiempo y luego se proyecta en el periodo omitido en el ajuste original. La habilidad de predicción puede entonces evaluarse comparando las predicciones de la proyección con las observaciones.

7.3. Otras recomendaciones

El PR señaló que al documentar la evaluación se debería incluir un análisis de puente (o "escalonado") para mostrar el impacto de cambios sucesivos en las decisiones del modelo y el tratamiento de los insumos entre el modelo de evaluación anterior y el actual.

8. RESUMEN DE LOS COMENTARIOS DEL PÚBLICO

El PR recibió comentarios del público después de la conclusión de la revisión. Éstos se resumen brevemente a continuación:

- El representante de Ecuador agradeció al personal científico de la CIAT y al PR. Señalando la falta de conclusiones firmes sobre la división espacial para el modelo de evaluación, Ecuador expresó preocupación sobre el asesoramiento de ordenación para la evaluación del próximo año.
- Un miembro del público quiso reconocer la presión para el personal de la CIAT de tener que proporcionar recomendaciones de ordenación al CCA de la CIAT en 2020, señalando que se espera que el personal presente resultados que informen la próxima medida de conservación. Se sugirió una opción alternativa por la cual, si no se resolvieran los problemas con el nuevo modelo del YFT en un plazo apropiado, el CCA podría recomendar prolongar la resolución actual por un año adicional para evitar presionar al personal de la CIAT. El miembro del público señaló que era mejor tener confianza en los resultados que no apoyar plenamente las conclusiones presentadas al CCA. También se agradeció al PR por su considerable esfuerzo en proporcionar un nuevo punto de vista.
- El representante de Japón agradeció al personal científico de la CIAT y al PR por su arduo trabajo, señalando también las colaboraciones en curso con la CIAT en materia de datos.
- La delegación de Corea agradeció al PR y mencionó su motivación para contribuir al análisis de la CPUE para el área del OPO y otros trabajos, así como su colaboración en curso con el personal de la CIAT.

9. RESUMEN DE LAS CONCLUSIONES Y TRABAJO FUTURO

Al concluir la reunión, el PR no consideró que ningún modelo que surgió fuese un candidato fuerte para la evaluación de 2020. En cambio, el PR concluyó que todavía era necesario desarrollar considerablemente

el modelo antes de que surgieran modelos candidatos.

Uno de los aspectos clave de este desarrollo del modelo requerido sería la producción de los índices de CPUE. Estos índices probablemente incluirían un índice de CPUE DEL PS (o índices) en el área al norte de la línea ecuatorial y un índice de CPUE LL al sur de la línea ecuatorial. La estructura del modelo candidato dependerá en cierta medida de la fiabilidad y representatividad de cada índice y de la coherencia en las tendencias de ambos índices dada la selectividad relativa de las pesquerías correspondientes.

No obstante, cabe señalar que el personal de la CIAT ya ha identificado conflicto entre estos índices. Por lo tanto, si el conflicto identificado no se resuelve, probablemente será necesario seleccionar un índice primario para la población entera (probablemente el índice LL como en el modelo de evaluación actual) y/o acomodar las diferencias mediante una división espacial del modelo. Si se adopta este enfoque de división espacial, el PR recomienda que el personal de la CIAT evalúe una serie de opciones de estructura espacial que no fueron presentadas al PR.

9.1. Estructura espacial

El PR observó algunas pruebas de que existe un intercambio limitado entre las 'poblaciones' del sur y del norte, tanto en los patrones en los datos de CPUE y de composición por talla, como en los patrones de desplazamiento derivados del marcado. No obstante, no está claro si los patrones en los datos de CPUE y de composición por talla se deban a un proceso biológico real. En esta etapa, el PR no cree que haya suficientes pruebas para justificar la hipótesis de dos poblaciones como la única hipótesis plausible.

Los datos y los modelos deberían ser consistentes con hipótesis explícitas y plausibles o con modelos conceptuales. Por ejemplo, si un modelo es nominalmente un modelo espacial, el dominio espacial debería definirse, y la preparación de datos y los modelos deberían ser consistentes para el dominio espacial. Dado que la información que apoya la hipótesis de dos poblaciones era sugerente más que concluyente, se deberían seguir desarrollando modelos alternativos que sean consistentes con las hipótesis de una y dos poblaciones, así como modelos en algún punto entre esos extremos. Si se desarrollan modelos con estructuras poblacionales alternativas, deberían evaluarse las consecuencias para la ordenación de usar una hipótesis de dos poblaciones si la hipótesis de una población fuera realmente cierta, y viceversa (por ejemplo, análisis de riesgos).

Para el desarrollo futuro de modelos, se debería investigar la posibilidad de incorporar una estructura espacial más compleja en un solo modelo del OPO entero. Además de las hipótesis de una y dos poblaciones presentadas por el personal de la CIAT, el PR también sugirió varios ejemplos de modelos e hipótesis plausibles con estructuras alternativas. Estos incluyen, entre otros: 1) una hipótesis de una población pero suponiendo que solamente los índices LL son consistentes; 2) una hipótesis de una población pero suponiendo que solamente los índices DEL son consistentes; 3) una hipótesis de dos poblaciones basada en los dominios espaciales del norte y del sur que pueden variar según la estación; 4) una hipótesis de dos poblaciones con una población basada en YFT asociado a delfines y otra población de YFT no asociado a delfines; y 5) una hipótesis de un grupo común de reclutas que se reparte entre las dos áreas.

9.2. Método de VAST

VAST en su formulación actual, como la mayoría de los modelos espaciales, supone que la variable de respuesta es independiente de la intensidad de muestreo. Parece que el personal de la CIAT supone que el uso de VAST es una forma de control para el muestreo preferencial. Esta es una cuestión clave para todas las pesquerías que dependen del uso de CPUE dependiente de la pesca como sustituto de la abundancia.

En el OPO, esta cuestión, junto con la contracción del esfuerzo palangrero, aumenta el riesgo de que el método de VAST aplicado en su forma actual resulte en estimadores sesgados de las cantidades de interés

modeladas, aun cuando los diagnósticos sean buenos. Si bien VAST puede ayudar en cuestiones de cobertura desigual del esfuerzo espaciotemporal, todavía no resuelve la cuestión del muestreo preferencial.

El PR recomienda desarrollar análisis de sensibilidad (CPUE y COMP) usando métodos tradicionales y de VAST, y que se evalúe la calidad de los ajustes de modelo resultantes usando diagnósticos adecuados. Los resultados del modelo de VAST deberían seleccionarse sobre los métodos tradicionales solo si se puede demostrar una mejora en el desempeño del diagnóstico. El personal también debería explorar si el modelo es sensible a los patrones de muestreo preferencial que se dan en el OPO. Por último, en caso posible, el personal debería explorar los impactos del muestreo preferencial en la fiabilidad de la CPUE estandarizada como índice de abundancia, así como investigar soluciones potenciales o colaborar con científicos pesqueros que ya estén realizando investigaciones sobre este tema.

9.2.1. Composición por talla de la población a partir de VAST

El PR tiene dudas con respecto a la fiabilidad de las composiciones por talla basadas en la población derivadas del método de modelado espacial de VAST. Sin embargo, no se ha proporcionado información suficiente para comprender el mecanismo de generación de conjuntos de datos de talla compuestos (por ejemplo, correlaciones espaciales). Los diagnósticos útiles y fáciles de interpretar, desagrupados en espacio y tiempo, podrían ser útiles en este sentido.

9.2.2. Índice de abundancia de la población

Es necesario evaluar la influencia de los dos componentes del modelo de CPUE (probabilidad de captura y abundancia), incluyendo un examen detallado (en espacio y tiempo) de los residuales del modelo, especialmente cuando se integran en un *software* complejo de modelado espacial como VAST.

9.2.3. Covariables ambientales de VAST

El PR reconoce que existe un problema con la inclusión de covariables oceanográficas como efectos de densidad o capturabilidad. Dada la falta de covariables de capturabilidad supuestas en el modelo actual (más allá de anzuelos entre flotadores), las predicciones del modelo mejorarían con la adición de covariables de densidad como la oceanografía. La inclusión de covariables oceanográficas podría:

- Brindar información adicional al modelo para imputar la CPUE en áreas grandes sin observaciones;
- Mejorar la capacidad del modelo para tomar en cuenta los efectos estacionales en la CPUE (más allá del uso de trimestres solamente); y
- Resolver parcialmente el posible problema de variación de la relación de covarianza espacial sobre el dominio espacial si parte de esta estructura de covarianza es causada por la distribución espacial de los datos oceanográficos.

9.3. Índices

9.3.1. Índices de cerco

Es probable que la CPUE PS de la pesquería sobre delfines se esté volviendo un insumo más importante para la evaluación dada la disminución de la extensión espacial de la pesquería palangrera. Existen limitaciones reconocidas en el uso de índices de CPUE PS en las evaluaciones de poblaciones de atún. Se requiere un examen más detallado para refinar el conjunto de datos de captura y esfuerzo PS DEL para asegurar que el conjunto de datos sea representativo con respecto a la operación de los buques y la estructura de la flota, y que pueda estandarizarse para obtener un índice de abundancia fiable. Se reconoce que es probable que haya tendencias a largo plazo en el aumento de la eficacia pesquera que no se han tomado en cuenta en la estandarización actual de los datos de captura y esfuerzo PS (relacionados con mejoras en la tecnología pesquera), especialmente dado el historial de actividad de los buques seleccionados como parte de la flota principal. El personal de la CIAT debería dar prioridad a investigar métodos

para tomar en cuenta la progresión del esfuerzo en los índices de cerco asociados a delfines.

Más específicamente, el PR recomienda que el personal de la CIAT considere:

- Trabajar con las partes interesadas (por ejemplo, países miembros, compañías pesqueras) para definir un cronograma de cambios tecnológicos que podrían informar sobre la Q variable con el tiempo.
- Refinar las covariables usadas para tomar en cuenta la progresión del esfuerzo más allá del identificador del buque.
- Mejorar la cuantificación del esfuerzo (por ejemplo, eliminar el tiempo de procesamiento del lance y otras actividades no relacionadas con el tiempo de búsqueda).
- Examinar la sensibilidad del modelo a incrementos (arbitrarios) en la potencia de pesca a lo largo del tiempo.

9.3.2. Índices de palangre

Es probable que los índices de CPUE de palangre se hayan vuelto menos fiables en la última década a medida que la pesquería se ha contraído. No obstante, para fines comparativos y mientras se desarrollan índices de CPUE PS fiables, es necesario mantener una opción de modelo alternativo que incorpore un índice de CPUE LL (derivado con métodos de estandarización más establecidos) como índice de abundancia principal.

9.4. Información del ciclo vital

9.4.1. Crecimiento

El personal de la CIAT concluyó que la mejor manera de estimar el crecimiento del YFT en el OPO es usar incrementos diarios en otolitos. El PR señaló que la forma de la curva de crecimiento era esencialmente lineal hasta la edad 4, con pocos datos de edad para informar la relación de crecimiento de peces más viejos/grandes. Dada la incertidumbre en el crecimiento, el PR apoya la estimación de los parámetros de crecimiento dentro del modelo integrado.

- Se deberían explorar métodos híbridos que combinen los incrementos diarios y anuales en las estimaciones finales de edad para equilibrar las fortalezas y debilidades de ambos métodos.
- El PR también sugiere que se consideren fuentes de datos adicionales en ese ajuste (por ejemplo, datos de marcado).

9.4.2. Mortalidad natural

Dada la escasez de información disponible para estimar la mortalidad natural, el PR no hizo recomendaciones específicas con respecto a las formas funcionales usadas en los modelos de población propuestos. Sin embargo, el PR señala que:

- M por edad es sensible a los supuestos sobre el crecimiento y la proporción de sexos por edad. Por lo tanto, cualquier cambio en la parametrización de estas funciones requerirá que se reestime la M .
- Los análisis de sensibilidad del modelo deberían incluir un nivel alternativo plausible de mortalidad natural que sea inferior al nivel general de M actualmente supuesto.

9.4.3. Reclutamiento y relación reproductor-recluta

Debido a la naturaleza exploratoria del trabajo presentado, no fue posible determinar las parametrizaciones más adecuadas de los parámetros R-R. Sin embargo, el PR recomienda que:

- Se consideren los procedimientos estándar de corrección de sesgo de *Stock Synthesis* (por ejemplo, Methot y Taylor, 2011) antes de seleccionar los modelos finales, incluyendo un intento de

estimar σ_R .

- Realizar un perfil de verosimilitud en σ_R para informar la selección del valor más adecuado para ese parámetro si finalmente se fija σ_R , dado que parecía haber cierto apoyo a los valores de σ_R superiores a 0.6.
- Considerar otros valores plausibles para la relación R-R y tomar en cuenta la sensibilidad del modelo a esas opciones al discutir los puntos de referencia, ya que éstos son altamente sensibles a estas selecciones, y el personal de la CIAT usa una parametrización predeterminada para la relación R-R significativamente diferente de la usada en otras evaluaciones del aleta amarilla (inclinación de 1.0).

9.5. Diagnósticos del modelo

A fin de examinar la calidad y estabilidad de los modelos finales, el PR recomienda una evaluación exhaustiva de un conjunto de diagnósticos que incluya:

- Diagnósticos mejorados de ajuste del modelo a los datos de composición por talla (por ejemplo, diagrama de caja de residuales de Pearson por clase de talla, año de observación, y potencialmente cohorte).
- Un examen de la matriz de covarianza para identificar pruebas de formulación inadecuada del modelo.
- Un análisis de “*jitter*” para examinar el efecto de la variación de los valores de los parámetros iniciales en los resultados del modelo.
- Perfil de verosimilitud para evaluar el desempeño del modelo a través de una gama de valores de un parámetro del modelo (por ejemplo, inclinación R_0 , σ_R).
- *Bootstrap*: En un modelo que funciona debidamente, la estimación de verosimilitud máxima debería ser similar al promedio de los *bootstraps*.
- Análisis retrospectivo: La gravedad de los patrones retrospectivos debería cuantificarse usando la ρ de Mohn.
- *Hindcasting* (Kell *et al.*, 2016).

9.6. Otras recomendaciones

El PR señaló que la metodología usada para preparar, procesar y analizar los insumos usados en la evaluación y los diagnósticos resultantes rara vez se incluía en los informes presentados al CCA. Describir la metodología exacta usada en la producción de los insumos para las evaluaciones presentadas para un año determinado, así como los diagnósticos clave que las acompañan, es fundamental para garantizar la transparencia, reproducibilidad y calidad de la ciencia usada para informar la ordenación pesquera.

El PR considera que esta es la mejor práctica para las evaluaciones de poblaciones, incluso cuando las partes interesadas no revisan los documentos técnicos. Esto es especialmente importante cuando la metodología evoluciona entre los años, como ocurre actualmente con la estandarización de la CPUE y el procesamiento de los datos de composición por talla. Otras OROP (por ejemplo, la CICAA, la CTOI y la WCPFC) incluyen informes de entrada de datos complementarios en sus evaluaciones.

ANEXO A: Áreas de enfoque identificadas por el personal de la CIAT

Main areas

1. What causes the mismatch in the longline and purse seine CPUE based indices of relative abundance?
2. What is the most appropriate stock structure for the yellowfin tuna stock assessment?
3. What is the most appropriate fishery structure for the yellowfin tuna stock assessment?
4. What approach should be used to deal with the uncertainty in the length of old individuals and the impact it has on the stock assessment results?
5. Can you estimate absolute abundance?
6. What is the appropriate stock-recruitment relationship?
7. How should the CPUE indices of abundance be used in the stock assessment?
8. Should logistic selectivity be used and for which fishery/survey?
9. How should the recent increase in the size of longline caught fish be dealt with?
10. Age- and sex-specific natural mortality

Other topics of interest

1. Time varying selectivity
2. Reference points
3. Using tagging data
4. Methods used to create the length-composition data
5. Inclusion of age conditioned on length data
6. Definition of spawning biomass
7. Diagnostics
8. Research recommendations
9. Data collection
10. Data analysis

ANEXO B: Términos de referencia

1. Metas y objetivos

La revisión de la evaluación por el personal de la CIAT de la población de aleta amarilla no tiene como propósito determinar el carácter adecuado o no de la evaluación actual para proporcionar asesoramiento de ordenación; la intención es brindar información al equipo de evaluación para mejorar la evaluación. Las metas y objetivos de la revisión son:

- a. identificar la mejor ciencia disponible para uso en la evaluación;
- b. brindar una revisión independiente de la evaluación; y
- c. proporcionar asesoramiento sobre investigaciones y recolección de datos a futuro que mejorarán la evaluación y la provisión de asesoramiento de ordenación.

2. Responsabilidades del Panel de Revisión

La responsabilidad principal del Panel de Revisión es realizar una revisión técnica adecuada de la evaluación. Los miembros del Panel deberían divulgar cualquier conflicto de intereses que podría perjudicar su objetividad de forma importante. Los conflictos de intereses incluyen, pero no están limitados a, intereses e inversiones financieras personales, afiliaciones del empleador, y arreglos, becas, o contratos consultativos.

Las responsabilidades específicas del Panel son:

- a. estar familiarizado con los términos de referencia, revisar los documentos informativos, insumos de datos, modelos analíticos, junto con otra información pertinente (por ejemplo, evaluaciones previas e informes previos del Panel de Revisión);
- b. discutir los méritos y deficiencias técnicas de los datos de entrada y métodos analíticos, trabajar con el personal de la CIAT para corregir las deficiencias, y, en caso posible, sugerir nuevas herramientas, análisis, y métodos de recolección de datos para mejorar las evaluaciones futuras; y
- c. redactar un informe de la reunión, para documentar las discusiones y recomendaciones.

Es responsabilidad del presidente del Panel coordinar las discusiones para que se lleve a cabo la revisión en el tiempo disponible.

3. Comentarios del público

Se asignan a tiempo durante la reunión para comentarios del público. El Panel tomará estos comentarios en consideración al elaborar su informe, según proceda.

4. Solicitudes de análisis adicionales

El propósito de la reunión es una revisión técnica de la metodología de evaluación, y podrían ser de beneficio análisis adicionales durante la reunión. Durante el transcurso de la reunión, el Panel podrá solicitar una cantidad razonable de análisis de sensibilidad, detalles adicionales para los modelos presentados, o análisis adicionales de ejecuciones alternativas. No obstante, cualquier solicitud de este tipo debe ser clara, explícita, presentada por escrito, y práctica en términos del tiempo disponible. Deberán ser detalladas individualmente en el informe del Panel, junto con su motivación y la respuesta. Al grado posible, los análisis solicitados por el Panel deberían ser llevados a cabo durante la reunión por el equipo de evaluación.

5. Informe del panel

El informe del Panel debería ser redactado y aprobado en un plazo corto después de la reunión. El proceso de elaborar el informe seguirá estos pasos:

- a. el Panel presenta una reseña del informe en la reunión;
- b. el Panel redacta y acuerda el borrador del informe;
- c. el Panel remite el borrador del informe al personal de la CIAT para comentarios sobre su exactitud técnica; y
- d. el Panel revisa los comentarios del personal, y modifica el informe según sea necesario.

El informe incluirá:

- a. nombres y afiliaciones de los miembros del Panel;
- b. una breve reseña de la reunión (lugar, agenda, principales recomendaciones por el Panel, *etc.*);
- c. breve resumen del modelo de evaluación actual, los datos usados, análisis presentados, y modelo de evaluación propuesto;
- d. lista de análisis solicitados por el Panel, motivación de cada solicitud, y un breve resumen de la respuesta;
- e. comentarios sobre los méritos y/o deficiencias en la evaluación, y recomendaciones para remediarlos;
- f. problemas no resueltos y principales incertidumbres, por ejemplo, cualquier problema especial que complique la evaluación y/o la interpretación de los resultados;
- g. problemas relacionados con los datos, la pesquería, o la evaluación planteados por el público; y
- h. recomendaciones de investigación y recolección de datos priorizadas para la evaluación subsiguiente.

El Panel y el personal de la CIAT procurarán resolver cualquier divergencia de opinión que pueda surgir con respecto al contenido del informe. Cualquier divergencia de opinión que no pueda ser resuelta deberá ser documentada y reflejada en el informe, el cual será publicado como informe especial de la CIAT.

Referencias

Nokome Bentley, Terese H. Kendrick, Paul J. Starr, Paul A. Breen, Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations, *ICES Journal of Marine Science*, 69 (1): 84–88.

Martell, S.J.D., de Bruyn, P., Davies, Nick M., Ernst, B. (2013). Recommendation of the review panel on the IATTC assessment of yellowfin tuna. Special report 20 submitted to the IATTC, La Jolla, California, USA.

ANEXO C: Lista de participantes

Nombre	Afiliación	Correo electrónico
Cass-Calay, Shannon	EE. UU.	shannon.calay@noaa.gov
Dreyfus, Michel	México	dreyfus@cicese.mx
Dunn, Alistair	Nueva Zelanda	alistair.dunn@oceanenvironmental.co.nz
Fleischer, Luis	México	lfleischer21@hotmail.com
Langley, Adam	Nueva Zelanda	adam_langley@xtra.co.nz
Kim, Doo Nam	Corea	doonam@korea.kr
Lee, Sung Il	Corea	k.sungillee@gmail.com
Morán, Guillermo	Ecuador	gamv6731@gmail.com
Piner Kevin	EE. UU.	kevin.piner@noaa.gov
Quiroz, Juan	Ecuador	jc.quirozespinoza@gmail.com
Satoh, Keisuke	Japón	kstu21@fra.affrc.go.jp
Schiff, Tom	EE. UU.	tschiffsd@aol.com
Teo, Steven	EE. UU.	steve.teo@noaa.gov
Tremblay, Laura	Nueva Zelanda	laura@dragonfly.co.nz
Urtizbera, Agurtzane	UE	aurtizbera@azti.es
Personal de la CIAT		
Compeán, Guillermo		gcompean@iattc.org
Aires-da-Silva, Alexandre		alexdasilva@iattc.org
Fuller, Dan		dfuller@iattc.org
Griffiths, Shane		sgriffiths@iattc.org
Hall, Martin		mhall@iattc.org
Lennert, Cleridy		clennert@iattc.org
López, Jon		jlopez@iattc.org
Maunder, Mark		mmaunder@iattc.org
Minte-Vera, Carolina		cminte@iattc.org
Morgan, Jeff		jmorgan@iattc.org
Sarazen, Robert		rsarazen@iattc.org
Schaefer, Kurt		kschaefer@iattc.org
Valero, Juan		jvalero@iattc.org
Xu, Haikun		hxu@iattc.org

ANEXO D: Documentos disponibles al Panel de Revisión

Informes técnicos:

- Ailloud L., Beare D., Farley J.H., and Krusic-Golub K. (2020) Preliminary results on AOTTP validation of otolith Increment deposition rates in yellowfin tuna in the Atlantic. COLLECTIVE VOLUME OF SCIENTIFIC PAPER (76): 156-163.
- Hall, M., & Román, M. H. (2019). THE FISHERY ON FISH-AGGREGATING DEVICES (FADs) IN THE EASTERN PACIFIC OCEAN. IATTC - Scientific Advisory Committee (May), 13–17.
- IATTC (2012). EXTERNAL REVIEW OF IATTC YELLOWFIN TUNA DOCUMENT YFT-01-01 INTRODUCTION TO THE REVIEW OF THE ASSESSMENT OF YELLOWFIN TUNA IN THE EASTERN PACIFIC OCEAN , 2012. (October), 15–19.
- IATTC (2016). STATUS OF THE TUNA AND BILLFISH STOCKS IN 2015. Stock Assessment Report 17. <https://doi.org/10.7551/mitpress/3298.003.0011>
- IATTC (2018). REPORT ON THE TUNA FISHERY , STOCKS , AND ECOSYSTEM IN THE EASTERN PACIFIC OCEAN IN 2018 Republic of Korea United States of America.
- IATTC (2019). External review of IATTC staff's stock assessment of bigeye tuna in the eastern Pacific Ocean - 11-15 March, 2019 - La Jolla, California, USA. https://www.iatcc.org/Meetings/Meetings2019/BET-02/Docs/_English/BET-02-AGN_External%20review%20of%20IATTC%20staff%E2%80%99s%20stock%20assessment%20of%20bigeye%20tuna%20in%20the%20eastern%20Pacific%20Ocean.pdf
- IATTC (2019). WORKSHOP ON AGE AND GROWTH OF BIGEYE AND YELLOWFIN TUNAS IN THE PACIFIC OCEAN. La Jolla, California (USA) 23-25 January 2019. https://www.iatcc.org/Meetings/Meetings2019/OTM-28/PDFs/_English/OTM-28-RPT_%20BET%20YFT%20Age%20and%20growth.pdf
- Martell, S. J. D., de Bruyn, P., Davies, N. M., & Ernst, B. (2013). RECOMMENDATIONS OF THE REVIEW PANEL ON THE IATTC ASSESSMENT OF YELLOWFIN TUNA.
- Maunder, M.N. and Aires-da-Silva, A., (2012). A review and evaluation of natural mortality for the assessment and management of yellowfin tuna in the eastern Pacific Ocean. Inter-Amer. Trop. Tuna Comm. Document YFT-01-07.
- Minte-Vera, C. V., Maunder, M. N., & Aires-da-Silva, A. (2018). Status of yellowfin tuna in the Eastern Pacific Ocean in 2017 and outlook for the future. IATTC - Scientific Advisory Committee, SAC-07-05b).
- Minte-Vera, C., Maunder, M. N., Aires-da-Silva, A., & Xu, H. (2019). EVALUATING INCONSISTENCIES IN THE YELLOWFIN ABUNDANCE INDICES. IATTC - Scientific Advisory Committee, (SAC-10 INF-F), 13–17.

Literatura revisada por pares:

- Aires-da-Silva, A. M., Maunder, M. N., Schaefer, K. M., & Fuller, D. W. (2015). Improved growth estimates from integrated analysis of direct aging and tag – recapture data : An illustration with bigeye tuna (*Thunnus obesus*) of the eastern Pacific Ocean with implications for management. Fisheries Research, 163, 119–126. <https://doi.org/10.1016/j.fishres.2014.04.001>
- Francis, R. I. C. C., Aires-da-Silva, A. M., Maunder, M. N., Schaefer, K. M., & Fuller, D. W. (2016). Estimating fish growth for stock assessments using both age – length and tagging-increment data. Fisheries Research, 180, 113–118. <https://doi.org/10.1016/j.fishres.2015.06.011>

- Hampton, J., 2000. Natural mortality rates in tropical tunas: size really does matter. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(5), pp.1002-1010. <https://doi.org/10.1139/f99-287>
- Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L. and Ono, K., 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models. *ICES Journal of Marine Science*, 72(1), pp.99-110. <https://doi.org/10.1093/icesjms/fsu198>
- Kell, L.T., Kimoto, A. and Kitakado, T., 2016. Evaluation of the prediction skill of stock assessment using hindcasting. *Fisheries Research*, 183, pp.119-127. <https://doi.org/10.1016/j.fishres.2016.05.017>
- Lennert-Cody, C. E., Minami, M., Tomlinson, P. K., & Maunder, M. N. (2010). Exploratory analysis of spatial-temporal patterns in length-frequency data: An example of distributional regression trees. *Fisheries Research*, 102(3), 323–326. <https://doi.org/10.1016/J.FISHRES.2009.11.014>
- Lennert-cody, C. E., Moreno, G., Restrepo, V., Roma, M. H., & Maunder, M. N. (2018). Recent purse-seine FAD fishing strategies in the eastern Pacific Ocean : what is the appropriate number of FADs at sea ? *ICES Journal of Marine Science*, 75, 1748–1757. <https://doi.org/10.1093/icesjms/fsy046>
- Maunder, M. N., & Deriso, R. B. (2013). A stock – recruitment model for highly fecund species based on temporal and spatial extent of spawning. *Fisheries Research*, 146, 96–101. <https://doi.org/10.1016/j.fishres.2013.03.021>
- Maunder, M. N., Deriso, R. B., Schaefer, K. M., Fuller, D. W., Aires, A. M., Minte, C. V., & Steven, V. (2018). The growth cessation model : a growth model for species showing a near cessation in growth with application to bigeye tuna (*Thunnus obesus*). *Marine Biology*, 165(4), 1–9. <https://doi.org/10.1007/s00227-018-3336-9>
- Methot Jr, R.D. and Taylor, I.G., (2011). Adjusting for bias due to variability of estimated recruitments in fishery assessment models. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(10), pp.1744-1760. <https://doi.org/10.1139/f2011-092>
- Minte-Vera, C. V, Maunder, M. N., Aires-da-Silva, A. M., Satoh, K., & Uosaki, K. (2017). Get the biology right , or use size-composition data at your own risk. *Fisheries Research*, 192, 114–125. <https://doi.org/10.1016/j.fishres.2017.01.014>
- Minte-Vera, C. V, Maunder, M. N., Schaefer, K. M., & Aires-da-Silva, A. M. (2019). The influence of metrics for spawning output on stock assessment results and evaluation of reference points : An illustration with yellowfin tuna in the eastern Pacific Ocean. *Fisheries Research*, 217, 35–45. <https://doi.org/10.1016/j.fishres.2018.09.022>
- Sardenne, F., Dortel, E., Le Croizier, G., Million, J., Labonne, M., Leroy, B., Bodin, N. and Chassot, E., 2015. Determining the age of tropical tunas in the Indian Ocean from otolith microstructures. *Fisheries Research*, 163, pp.44-57. <https://doi.org/10.1016/j.fishres.2014.03.008>
- Wang, S., Maunder, M. N., Piner, K. R., Aires-da-Silva, A., & Lee, H. (2014). Evaluation of virgin recruitment profiling as a diagnostic for selectivity curve structure in integrated stock assessment models. *Fisheries Research*, 158, 158–164. <https://doi.org/10.1016/j.fishres.2013.12.009>
- Xu, H., Lennert-cody, C. E., Maunder, M. N., & Minte-Vera, C. V. (2019). Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. *Fisheries Research*, 213 (September 2018), 121–131. <https://doi.org/10.1016/j.fishres.2019.01.013>
- Zhu, J., Maunder, M. N., Aires-da-Silva, A. M., & Chen, Y. (2016). Estimation of growth within Stock Synthesis models : Management implications when using length-composition data. *Fisheries Research*, 180, 87–91. <https://doi.org/10.1016/j.fishres.2015.09.019>

ANEXO E: Solicitudes y respuestas

The following requests were made by the RP to the IATTC staff (see responses [1](#) and [2](#))

1. Time-series of heatmaps for Japanese longline effort, longline yellowfin CPUE and bigeye CPUE for additional context and also in order to see whether there is a spatial pattern in the effort contraction in the Japanese longline fleet that might match one of the target species' CPUE. [LTB]
The heatmaps were provided by the IATTC staff and were used as a reference during the review.
2. Can we please see a plot of the Longline CPUE over time, and compare this with an equivalent index from the aggregated length frequency index from the VAST models [AD]
3. Can we please plot the %dolphin (by effort and by catch) per vessel per month as a time series box-plot for vessels that meet the threshold of at least 75% dolphin sets [AD]
4. A time-series showing the years of activity for each vessel in the purse seine DEL fleet from 1975 onwards, with points colored or scaled according to the proportion of dolphin sets. This plot was requested for additional context on the vessel history within the fleet and to check the relevance of using a vessel factor in the CPUE standardization as a proxy for operational changes.
The plot was provided by the IATTC staff and showed that many vessels had been active for a time period spanning many operational changes, and that there was a gradual change for some in the proportion of dolphin sets.
5. Regarding my request for the spatial comparisons of the distribution of archival tag data and purse seine fishing effort to evaluate the extent of fish movement within the domain of the area of operation of the purse seine fishery. [AL]

I suggest using the contour maps of archival release data sets – slide 18 of Kurt's presentation.

These could then be used to overlay the distribution of the PS DEL fishing effort for individual time period blocks (aggregated for 5? year intervals). Not necessary to do the entire time period of the PS fishery – maybe just the last 20 years.
6. Attached is a draft of the table requested on the **model structure**. We still need to fill in the information for how the purse seine catch and length comp data are calculated and the indices of abundance, but I thought you might want to look at it for the other information before tomorrow morning. You can see a map of the fisheries on the google drive. [SCC]
7. Please can you ask the assessment team to conduct an additional set of model runs (both north and south) with the all of the following changes to the current north and south models made simultaneously
 - Tighten the cv at length for the age-length relationship about 0.05
 - Relax the sigmaR value to something large, like 1.0
 - Fix the survey selectivity to be flat topped and asymptote at 1.0
 - Remove, or substantially reduce the F penalty from both models (currently set at 0.2 for year 176 – either remove it or set it to be something out of the way) [AD?]