

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

SESGO Y PRECISIÓN DE UNA MODIFICACIÓN DEL MODELO DE JOLLY-
SEBER PARA INCLUIR EL TAMAÑO INICIAL DE LA POBLACIÓN Y CONTEOS
PARCIALES DE MORTALIDAD

Tesis sometida a la consideración de la Comisión del Programa de Estudios de
Posgrado en Estadística para optar al grado y título de Maestría Académica en
Estadística

LUIS ALEJANDRO RODRÍGUEZ CAMPOS

Ciudad Universitaria Rodrigo Facio Brenes, Costa Rica

2022

Dedicatoria

A Javier Arce, cuya experiencia me motivó a escribir este trabajo.

A mis abuelos en el cielo Dora, Alicia y Juan Luis, con cariño eterno.

Agradecimientos

A Dios, que me da la vida y la oportunidad de escribir este trabajo.

A mis padres, Juan Antonio y Carmen, que me han apoyado e impulsado para dar este paso. No hay manera de agradecerles que haga justicia a todo su esfuerzo.

A don Fernando Ramírez, por su paciencia y apoyo como director de tesis.

A Gilbert Brenes y Juan Bautista Chavarría, por sus invaluable aportes al trabajo en su papel de asesores.

A Johnny Madrigal Pana, que junto con mi padre Juan Antonio y don Fernando, me motivaron a escribir sobre este tema.

A Shu Wei Chou, Mauricio Campos (Moe) y Russell Alpízar Jara, que me apoyaron con ideas muy valiosas en diferentes fases de la tesis.

A los contribuyentes que sostienen el sistema de educación superior pública y nos permiten, aún sin beca, obtener títulos a bajísimo costo. Muchos de ellos, lastimosamente, no tienen o no pudieron gozar de estas oportunidades.

Esta Tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Estadística

Dr. Shu Wei Chau Chen
**Representante de la Decana
Sistema de Estudios de Posgrado**

M.Sc. Fernando Ramírez Hernández
Director de Tesis

M.Sc. Juan Bautista Chavarría Chaves
Asesor

Dr. Gilbert Brenes Camacho
Asesor

Ph.D. Arodys Robles Soto
Representante del Director del Programa de Posgrado en Estadística

Luis Alejandro Rodríguez Campos
Estudiante

Tabla de Contenidos

DEDICATORIA	II
AGRADECIMIENTOS	III
HOJA DE APROBACIÓN	IV
TABLA DE CONTENIDOS	V
RESUMEN	VII
ABSTRACT	IX
ÍNDICE DE CUADROS	XI
ÍNDICE DE FIGURAS	XIV
1 INTRODUCCIÓN	1
2 OBJETIVOS	3
2.1 General	3
2.2 Específicos.....	3
3 MARCO TEÓRICO	4
3.1 ¿Qué es la acuicultura y cuáles son sus retos más importantes?.....	4
3.2 Breve explicación del sistema de producción de tilapia y de las implicaciones del muestreo de captura y recaptura en este ambiente	4
3.3 Estimación del tamaño poblacional mediante métodos de captura y recaptura	8
3.3.1 <i>Técnicas para el análisis de poblaciones abiertas</i>	9
3.3.2 <i>Precisión de las estimaciones de abundancia y sobrevivencia del modelo de Jolly-Seber</i>	23
3.3.3 <i>Sesgo en las estimaciones de abundancia y sobrevivencia a partir de JS y CJS</i>	26
3.3.4 <i>Técnicas para el manejo de heterogeneidad en las probabilidades de captura</i>	31
4 PROPUESTA TEÓRICA	38

5	ANÁLISIS DE LA MORTALIDAD OBSERVABLE.....	44
5.1	Introducción.....	44
5.2	Ajuste de curvas de supervivencia	45
5.3	Resultados de los modelos ajustados	50
5.4	Prueba de simulación	53
6	METODOLOGÍA DE SIMULACIÓN	56
6.1	Fuente de los datos.....	56
6.2	Método de simulación de la población.....	58
6.3	Diseño de los escenarios de estimación.....	64
6.4	Análisis de los resultados de estimación	64
7	RESULTADOS	67
7.1	Abundancia	67
7.2	Sobrevivencia.....	73
7.3	Recuperabilidad	79
8	CONCLUSIONES Y RECOMENDACIONES	87
9	LITERATURA CITADA	90
10	ANEXO 1: SALIDAS A PARTIR DE DATOS CON COLAS RECORTADAS.....	98
10.1	Abundancia	98
10.2	Sobrevivencia.....	102
10.3	Recuperabilidad	106

Resumen

La estimación de los inventarios de peces es crucial en la producción acuícola, pues permite un manejo apropiado de las tasas de alimentación y de la programación de las cosechas.

Los métodos de captura y recaptura han sido ampliamente utilizados para estimar parámetros demográficos en poblaciones animales silvestres. Específicamente, el modelo de Jolly-Seber (JS) modela la abundancia y sobrevivencia de la población; sin embargo, es restrictivo en cuanto asume probabilidades iguales de captura y sobrevivencia para todos los individuos. Adicionalmente, en la acuicultura comercial se tiene información del número de peces sembrados y conteos parciales de mortalidad que podrían ayudar a generar modelos más robustos y precisos.

Considerando lo mencionado anteriormente, el presente trabajo busca determinar el sesgo y precisión de estimadores de sobrevivencia y abundancia de JS a partir de diferentes tamaños de muestra, para la estimación de sobrevivencia y tamaño instantáneo, en poblaciones artificiales con capturabilidad desigual, tamaño inicial conocido y conteos parciales de mortalidad, usando simulación.

Para ello, se simuló el crecimiento y mortalidad de 36 ciclos de producción de tilapia nilótica (*Oreochromis niloticus*) que se muestrearon para probar en cada una, dos diferentes números de ocasiones de muestreo (4 y 6) y tres diferentes tamaños de muestra en cada captura ($n = 50, 100$ y 150 , equivalentes aproximadamente al 5, 10 y 15% de la población inicial), por lo cual se obtuvieron 909 estimaciones de abundancia y 680 de sobrevivencia y recuperabilidad. Como factores a nivel de ciclo productivo, se incluyó la forma de la relación entre sobrevivencia aparente y sobrevivencia observada, así como la época del año.

Los datos simulados se analizaron mediante un modelo bayesiano de captura y recaptura que permite probabilidades desiguales de captura y la inclusión de conteos de mortalidad. Las estimaciones obtenidas de abundancia, sobrevivencia y recuperabilidad se analizaron por su nivel de sesgo y error estándar relativo.

Los resultados obtenidos para las estimaciones de abundancia mostraron una concordancia adecuada con los tamaños poblacionales reales, con bajos niveles de sesgo (<7%) y error estándar relativo (<3%). Se evidenció un sesgo importante si el estanque presenta pérdida o robo, pero no se vieron severamente afectadas por la época de siembra. Las estimaciones de sobrevivencia mostraron incluso menor nivel de sesgo y error estándar relativo, pero la distribución de las estimaciones tendió a mostrar una importante asimetría, que podría indicar falta de convergencia. Las estimaciones de recuperabilidad presentaron elevados niveles de sesgo y error estándar relativo, pero ambos disminuyen al aumentar la edad del estanque.

Si bien el modelo bayesiano de captura y recaptura presentado es innovador y tiene potencial para ser implementado, debe ser mejorado para reducir el tiempo que dura en converger y aumentar la calidad de las estimaciones, a fin de que pueda ser útil en situaciones de acuicultura industrial.

Abstract

Fish inventory estimation is a vital task in aquaculture because it allows an appropriate management of feeding rates and harvesting programs.

Capture-recapture methods have been widely utilized to estimate demographic parameters in wildlife populations. Specifically, Jolly-Seber (JS) model estimates abundance and survival rates of a population, however, it has some restrictive assumptions like homogeneity in catchability and survival probabilities. In addition, in commercial aquaculture there is information about the fish stocking and partial mortality counts that may help to generate more robust and precise models.

For these reasons, the aim of this work is to determine the bias and precision of JS survival and abundance estimators with different sample sizes, for the estimation of survival and instantaneous population size, in artificial populations with unequal catchability, known stocking size and partial mortality counts, via simulation.

The growth and mortality of 36 production cycles of Nile tilapia (*Oreochromis niloticus*) were simulated and sampled to test two different number of sampling occasions (4 and 6 per cycle) and sample sizes ($n = 50, 100, \text{ and } 150$; or approximately 5, 10 and 15% of initial population) on each cycle, thus 909 abundance and 680 survival and recuperability estimations were obtained. The form of the apparent survival vs observed survival and season were evaluated as external factors.

A Bayesian capture-recapture model allowing unequal catchability and inclusion of partial mortality counts was generated. Relative bias and standard error for abundance, survival and recuperability estimates were analyzed.

In general, abundance estimates showed good concordance with real data, showing low bias (<7%) and relative standard error (<3%). Estimates showed significant bias if production cycle presented fish loss/theft, but there was no significant effect of stocking season. Survival estimates showed even lower relative bias and standard errors, but the distribution of estimates showed

important asymmetry, indication of lack of convergence. Recuperability estimates showed unacceptable levels of relative bias and standard error.

Although our Bayesian capture-recapture model is innovative and has some potential to be implemented, it must be enhanced to reduce convergence time and estimation quality, to be useful in industrial aquaculture.

Índice de cuadros

Cuadro 1. Matriz triangular de estadísticos suficientes para el modelo de Cormack-Jolly-Seber.....	14
Cuadro 2. Matriz triangular para el manejo de los datos en el análisis de recuperación de marcas.....	20
Cuadro 3. Matriz triangular para el manejo de los datos en el modelo CJS con recuperación de marcas.....	21
Cuadro 4. Valores esperados para los conteos de recuperaciones y recapturas en un modelo de CJS con información adicional de recuperación de marcas	22
Cuadro 5. Modelos posibles que usan la información completa de historias de captura en el diseño robusto	32
Cuadro 6. Diferentes modelos que consideran heterogeneidad en la sobrevivencia y/o probabilidad de sobrevivencia, con K muestras y C clases.	34
Cuadro 7. Coeficientes, errores estándar y valores p para el modelo no lineal de Richards para la mortalidad de los estanques sembrados en época térmica cálida.	51
Cuadro 8. Valores de desviación estándar de los efectos aleatorios para cada uno de los parámetros del modelo no lineal de Richards para la mortalidad en los estanques sembrados durante la época cálida y sus correlaciones.	51
Cuadro 9. Coeficientes, errores estándar, valores p y desviación estándar de los efectos aleatorios para los parámetros del modelo logístico doble para la mortalidad de los estanques sembrados durante la época térmica fría.	52
Cuadro 10. Coeficientes del modelo para la predicción de la densidad de biomasa sembrada por estanque (kg/m ²)	52
Cuadro 11. Coeficientes, error estándar y valores p para el modelo de predicción de la edad del estanque.....	53

Cuadro 12. Datos de siembra, cosecha y duración del ciclo en 186 ciclos de producción de tilapia en estanques semiintensivos, llevados a cabo en una empresa ubicada en Costa Rica, durante la campaña 2012-2013.	57
Cuadro 13. Parámetros de simulación de la ganancia de peso vivo (g/d) por mes de siembra del estanque.	59
Cuadro 14. Indicadores de sesgo y el error estándar relativo de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra	67
Cuadro 15. Indicadores de sesgo y varianza de los estimadores de abundancia según la forma asumida para la relación entre mortalidad observada y mortalidad real.....	71
Cuadro 16 Indicadores de sesgo y error estándar relativo de los estimadores de abundancia según la época climática de siembra, comparando dos métodos de estimación	72
Cuadro 17. Indicadores de sesgo y error estándar relativo de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra	73
Cuadro 18. Indicadores de sesgo y error estándar relativo de los estimadores de sobrevivencia según la forma asumida para la relación entre mortalidad observada y mortalidad real	77
Cuadro 19. Comparación del sesgo relativo (SR) y el Error Estándar Relativo (EER) entre el modelo propuesto en este trabajo y el modelo usado en la empresa con relación a las épocas climáticas.	78
Cuadro 20. Indicadores de sesgo y varianza de los estimadores de recuperabilidad según el número de ocasiones de muestreo y el tamaño de la muestra	80
Cuadro 21. Indicadores de sesgo y error estándar relativo de los estimadores de recuperabilidad según la forma asumida para la relación entre mortalidad observada y mortalidad real.	84

Cuadro 22. Sesgo relativo (SR) y Error Estándar Relativo (EER) para las estimaciones de recuperabilidad por época climática.	85
---	----

Índice de figuras

Figura 1. Curvas de mortalidad, estimadas por el método de Kaplan-Meier para 189 estanques de producción semiintensiva de tilapia.	46
Figura 2. Curvas de mortalidad para 189 estanques, clasificadas por épocas térmicas, de acuerdo con la clasificación de Arce-Vega (2014).....	47
Figura 3. Diagrama de flujo del proceso de simulación.....	54
Figura 4. Curvas de mortalidad simuladas para 200 estanques semiintensivos de producción de tilapia.	55
Figura 5. Comparación entre diferentes valores de sobrevivencia obtenidos en 186 estanques de producción comercial de tilapia, en Costa Rica, durante la campaña 2012-2013.	58
Figura 6. Curva de crecimiento de tilapias, según Zuniga-Jara y Goycolea-Homann (2014).	60
Figura 7. Asociación entre la abundancia estimada y la abundancia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.....	68
Figura 8. Distribución del error porcentual de la estimación de abundancia por tamaños de muestra y número de ocasiones de muestreo.....	69
Figura 9. Asociación entre la edad del estanque y el error porcentual en las estimaciones de abundancia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.	70
Figura 10. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.	71
Figura 11. Error porcentual en la estimación de la abundancia por edad del estanque y época de siembra.	72
Figura 12. Asociación entre la sobrevivencia estimada y la sobrevivencia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.	74

Figura 13. Distribución del error porcentual en la estimación de sobrevivencia por tamaños de muestra y número de ocasiones de muestreo.....	75
Figura 14. Asociación entre la edad del estanque y el error porcentual en las estimaciones de sobrevivencia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.....	76
Figura 15. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.	78
Figura 16. Error porcentual de la estimación de sobrevivencia por edad y época del lote de producción.	79
Figura 17. Asociación entre la recuperabilidad estimada y la recuperabilidad real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.	81
Figura 18. Distribución del error porcentual en la estimación de recuperabilidad por tamaños de muestra y número de ocasiones de muestreo.....	82
Figura 19. Asociación entre la edad del estanque y el error porcentual en las estimaciones de recuperabilidad, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.....	83
Figura 20. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.	84
Figura 21. Error porcentual de la estimación de recuperabilidad por edad del cultivo y época climática del año.....	85



UNIVERSIDAD DE
COSTA RICA

SEP Sistema de
Estudios de Posgrado

Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.

Yo, Luis Alejandro Rodríguez Campos, con cédula de identidad 1-1494-0284, en mi condición de autor del TFG titulado SESGO Y PRECISIÓN DE UNA MODIFICACIÓN DEL MODELO DE JOLLY-SEBER PARA INCLUIR EL TAMAÑO INICIAL DE LA POBLACIÓN Y CONTEOS PARCIALES DE MORTALIDAD

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI NO *

*En caso de la negativa favor indicar el tiempo de restricción: _____ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

FIRMA ESTUDIANTE

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

1 INTRODUCCIÓN

La estimación de los inventarios de peces reviste una importancia crucial en la producción acuícola, pues un control adecuado de estos permite un manejo apropiado de las tasas de alimentación y de la programación de las cosechas (Killian, Heikes, Wyk, Masser y Engle, 1998). Con el fin de determinar el tamaño de los inventarios en sistemas de producción acuícola se han propuesto metodologías como estudios históricos, estimaciones volumétricas, conteos directos, mediante el uso del sonar, estudios de remoción, ajustes con funciones de sobrevivencia o regresión curvilínea, entre otras (Killian et al., 1998). Sin embargo, aunque los estudios de remoción y el uso de sonares son de las alternativas más confiables para estas estimaciones, no son aplicables en todas las condiciones de producción acuícola.

Los métodos de captura y recaptura han sido ampliamente utilizados para estimar parámetros demográficos en poblaciones de animales silvestres. Específicamente, el modelo de Jolly-Seber (JS) es un método que se enfoca en las pérdidas de individuos de la población, para modelar la abundancia y sobrevivencia de la población (Williams, Nichols, y Conroy, 2002). Dicho modelo, sin embargo, es restrictivo en cuanto a los supuestos en que se basa. En especial, los supuestos de probabilidades iguales de captura y sobrevivencia son los más difíciles de cumplir en la naturaleza (Krebs & Futuyma, 1973). Para relajar estos supuestos se ha propuesto el uso de covariables relacionadas con el tiempo, grupos múltiples o cohortes que comparten probabilidades de captura o sobrevivencia, covariables individuales que afectan la probabilidad de sobrevivencia y modelos de respuesta a la captura (*trap-shy* o *trap-happy*) (Williams, Nichols, y Conroy, 2002).

Por otra parte, en la acuicultura comercial se tiene información de ciertos valores que difícilmente se poseen en las poblaciones silvestres: dado que es un sistema de producción artificial, se conoce con cierta precisión el número de peces sembrado, asimismo se mide una serie de covariables ambientales y se dispone

de conteos parciales de mortalidad que podrían ayudar a generar modelos más robustos y precisos, reduciendo el tamaño de muestra y el tiempo consumido en las labores de muestreo. No obstante, la metodología estadística actual no presenta, según se ha constatado mediante la revisión bibliográfica, una sistematización tal que permita utilizar estas fuentes de información en el modelo Jolly-Seber de forma que sea aplicable a la acuicultura comercial. Por tanto, este trabajo busca, en primer lugar, contribuir significativamente a la Estadística como disciplina, al evaluar el efecto en el sesgo y la precisión de estimadores de sobrevivencia obtenidos mediante el modelo de Jolly-Seber de la inclusión del conocimiento del tamaño inicial de la población y conteos parciales de individuos. En segundo lugar, se desea contribuir al área de la acuicultura, al tener a la mano métodos de estimación del tamaño poblacional que provean información con mayor exactitud permitiendo ajustar de manera óptima las raciones, pronosticar más adecuadamente la pesca y controlar las pérdidas por robos o depredación, entre muchas otras ventajas.

Este documento se estructura de la siguiente manera. En el capítulo 2 se explicitan el objetivo general y los objetivos específicos de este trabajo. Seguidamente, en el capítulo 3, se hace un repaso de los antecedentes teóricos requeridos para la comprensión del trabajo, tanto desde el área estadística como desde el punto de vista zootécnico. Después, en el capítulo 4, se presenta el desarrollo teórico de los modelos utilizados en la tesis. Luego, en el capítulo 5, se realiza un análisis de sobrevivencia para modelar la mortalidad observable en el sistema de producción analizado. A continuación, en el capítulo 6, se procede a describir la metodología de las simulaciones realizadas, cuyos resultados se presentan en el capítulo 7, en el cual también se incluye la discusión de los resultados. Finalmente, en el capítulo 8 se muestran las conclusiones y recomendaciones del trabajo.

2 OBJETIVOS

En el capítulo anterior se mencionó la importancia que tiene el tema de esta tesis en la estadística y la acuicultura. En el presente capítulo se delimitarán los objetivos de la tesis, de forma que quede claro a lo que se quiere llegar con ésta.

2.1 General

Determinar el sesgo y precisión de estimadores de sobrevivencia y abundancia de Jolly-Seber a partir de diferentes tamaños de muestra, para la estimación de sobrevivencia y tamaño instantáneo, en poblaciones artificiales con capturabilidad desigual, tamaño inicial conocido y conteos parciales de mortalidad, usando simulación.

2.2 Específicos

1. Plantear una modificación al método de Jolly-Seber, que permita incluir información adicional disponible como el tamaño de población inicial y conteos parciales de mortalidad en poblaciones con capturabilidades asociadas a una covariable.
2. Contrastar el sesgo y la precisión de estimadores de sobrevivencia y abundancia poblacional obtenidos mediante el método de Jolly-Seber, con diferentes tamaños de captura, e incluyendo información adicional como tamaño de muestra inicial, conteos parciales de mortalidad, por medio de simulaciones.

El cumplimiento de estos objetivos generará información técnica que permita evaluar la utilidad de las metodologías de captura y recaptura en la estimación del tamaño poblacional en la acuicultura.

3 MARCO TEÓRICO

Este capítulo está organizado en dos secciones: la primera detalla las particularidades del sistema de producción acuícola, con especial enfoque a la producción de tilapia y su comportamiento. La segunda sección considera la teoría de los métodos de captura y recaptura, y como estos podrían ser aplicados en la industria acuícola.

3.1 ¿Qué es la acuicultura y cuáles son sus retos más importantes?

La acuicultura es un sector de gran crecimiento en el mundo. Mientras que para el 2016 representaba origen del 40% de los productos pesqueros, para 2016 dicha proporción fue de 46% (FAO, 2018). En Costa Rica, se producen unas 21 000 t de productos acuícolas, de las cuales el 80% corresponde a tilapia (Peña y Chacón, 2019). Es en el cultivo de este pez que se centrará la presente investigación.

3.2 Breve explicación del sistema de producción de tilapia y de las implicaciones del muestreo de captura y recaptura en este ambiente

Por el nombre *tilapia* se denominan tres géneros de cíclidos de origen africano (*Oreochromis*, *Sarotherodon* y *Tilapia*); sin embargo, las principales especies comerciales pertenecen al género *Oreochromis* y son: *O. niloticus*, *O. aureus*, *O. urolepis hornorum* y *O. mossambicus* (Popma & Masser, 1999). De estas, la tilapia más común en sistemas productivos a nivel mundial es *O. niloticus*, y es, a su vez, la cuarta especie acuática más cultivada a nivel mundial, con 4.2 millones de toneladas métricas para 2016, que representa un 7% de la producción mundial de productos acuícolas (FAO, 2018).

FUNPROVER (2008) señala las siguientes etapas de producción de tilapia: crianza (0-5 g), desarrollo o preengorde (5-80 g), engorde (80 g hasta el peso de cosecha) y reproducción. Estos son lineamientos generales, pues, a partir de este conocimiento, las empresas elaboran las etapas productivas requeridas según su conveniencia. Por ejemplo, Barboza Coto (2008), a partir de su experiencia en una

empresa de producción de tilapia en el Pacífico Norte de Costa Rica, las siguientes etapas:

- Reproducción: padrotes cuyo ciclo productivo es de 30 días de producción y 30 días de descanso
- Reversión sexual: Comprende de 0,02 g a 1 g, en un periodo de 30 días.
- Nursery 1: Comprende de 1g a 15 g en un período de 60 días.
- Nursery 2: Comprende de 15 a 40 g en un período de 90 días.
- Nursery-Engorde: Comprende de 40g a 250 g en un período de 120 días.
- Engorde 2: Comprende de 250 g a 950 g, que es el peso de salida a mercado, con una duración de 120 días.

En cada una de estas etapas se siembra una cantidad de peces de acuerdo con la disponibilidad de oxígeno en el agua y la tasa de recambio de esta, tomando en consideración que los niveles de oxígeno disuelto recomendables son de aproximadamente 4,5 mg/L (FUNPROVER, 2008). Por tanto, aquellos estanques con una mayor tasa de recambio de agua y mayor concentración de oxígeno disuelto permitirán sostener una mayor biomasa de peces por metro cúbico.

A partir de Barboza Coto (2008) se puede esquematizar el siguiente proceso de producción:

1. Se siembra una cantidad N_0 de tilapias, la cual es conocida con exactitud o, al menos, con incertidumbre instrumental.
2. Los peces, pasado un período de tiempo t , se cosechan, ya sea para pesca interna (traslado entre etapas) o para salida al mercado. El número de peces cosechado N_c , es conocido también con exactitud o, en su defecto, con incertidumbre instrumental.
3. La sobrevivencia del ciclo productivo se calcula como $\phi = N_c/N_0$.

Durante todo el tiempo en que los peces están bajo el agua, se registra la mortalidad recuperada dicha información, en conjunto con modelos ajustados con base en datos previos de la explotación, se estima la cantidad de peces en el agua.

Sin embargo, no toda la mortalidad es considerada, pues pueden existir escapes accidentales, robos, depredación o errores en la estimación del número de peces sembrado que podrían repercutir en una sobreestimación de la cantidad de peces que existen en un momento determinado en el estanque. Esto repercute en costos de alimentación, pues se alimenta considerando una cantidad de peces superior a la que realmente hay en el estanque, y en faltantes a nivel de cosecha (diferencias entre la cantidad cosechada y la cantidad que se esperaba cosechar) (Barboza Coto, 2008; Múnera Miranda, 2013).

En las explotaciones de producción de tilapia es común que se realicen muestreos periódicos para estimar el peso promedio de los peces y, con base en los datos, hacer ajustes en la abundancia de peces. Estos se hacen mediante un método indirecto y se basa en experiencia: si los peces están más pesados y con mejor condición corporal de lo esperado, es probablemente debido a que hay menos peces en el estanque de lo estimado, por lo cual es necesario reducir los valores esperados de abundancia. De igual forma, si la condición corporal y el peso de los animales es inferior a la esperada, y no se reportan problemas de enfermedad o calidad del alimento, se puede achacar a que la cantidad de peces es superior a la usada para calcular las tablas de alimentación (Barboza Coto, 2008).

Estos muestreos suelen realizarse atrayendo al animal con alimento balanceado hacia una de las esquinas del estanque, lo cual podría provocar que, en caso de que el proceso de muestreo no se controle de manera oportuna, sean los peces más dominantes los que se acerquen primero, si se supone un efecto *trap-happy*, ya que se brinda alimento, o bien, se alejen de la zona de muestreo, asumiendo una respuesta *trap-shy*, probablemente debida al estrés de la manipulación durante el experimento. Trabajos de Oliveira y Almada (1996a, 1996b) han constatado la presencia de relaciones sociales agonísticas de dominancia entre especímenes de *O. mossambicus*, donde las relaciones generalmente vienen dominadas por machos de mayor tamaño. Dicho tipo de relaciones podrían producir sesgo en las estimaciones de biomasa y las correcciones en el tamaño poblacional. Dinámicas sociales similares también se han reportado para *O.*

niloticus (Barcellos, Nicolaiewsky, de Souza y Lulhier, 1999; Fernandes y Volpato, 1993).

Por otra parte, contrario a las estimaciones en poblaciones naturales, la dinámica del estanque de tilapias es más sencilla, ya que durante la etapa de reversión sexual se administra la hormona 17- α -metil-testosterona, para favorecer la expresión gonadal masculina en toda la población, lo que favorece al sistema productivo, principalmente porque la tilapia nilótica es un pez que incuba sus huevos en la boca, por lo cual, las hembras en período reproductivo no comen ni ganan peso (El-Sayed, 2006). Aunque existen algunas ineficiencias en este proceso, ya que no todos los animales logran una reversión exitosa y, por lo tanto, algunos llegan a reproducirse, la proporción de errores de reversión es pequeña y se puede ignorar. Por esta razón, los posibles nacimientos ocurrirían en la etapa de engorde, por lo que no serían de interés para el sistema productivo y son fácilmente separables del resto. Además, no es común que se añadan nuevos peces a un ciclo de producción ya iniciado, por ello, se asume fácilmente que no hay nacimientos ni inmigración. Sin embargo, sí puede haber salidas de individuos de la población, ya sea por procesos de mortalidad o de emigración.

Adicionalmente, mientras en poblaciones naturales es común realizar estimaciones anuales de abundancia o sobrevivencia (Fletcher et al., 2012; Pollock, 1981), tal periodicidad entre las capturas sería insuficiente para los ciclos de producción acuícola, que no suelen ser tan extensos. Sin embargo, los muestreos deben mantenerse al mínimo posible y recolectando la menor cantidad de peces, no solamente para fomentar una mayor eficiencia en el uso de la mano de obra, sino por el efecto negativo que podría tener la manipulación excesiva sobre el metabolismo del crecimiento y la respuesta inmunológica de los peces (Akinrotimi, Abu, Ansa, Edun, y George, 2009; Barcellos et al., 1999; Fernandes y Volpato, 1993; Foo y Lam, 1993; Vijayan, Pereira, Grau, y Iwama, 1997).

Con base en lo discutido en esta primera sección, se puede sintetizar que un método apropiado para estimar el tamaño poblacional o la sobrevivencia en condiciones de producción acuícola comercial debería:

1. Hacer uso del conocimiento adicional que se tiene, como el tamaño de la población inicial y los conteos parciales de peces muertos, para evitar estimaciones con poco sentido, como que se señale que en el estanque haya más peces que la resta de los sembrados menos los muertos observados (considerando que ambas cantidades se saben con exactitud).
2. Ser altamente preciso, para evitar gastos excesivos en el consumo de alimento.
3. Hay que considerar que existen relaciones de dominancia social entre los peces, que podrían conducir a sesgos en la estimación de los parámetros de interés.

En la siguiente sección se analizan los principales modelos que se han desarrollado para las estimaciones de captura y recaptura de poblaciones naturales, y se evalúa cuánto satisfacen las condiciones antes mencionadas para la utilidad en la acuicultura comercial.

3.3 Estimación del tamaño poblacional mediante métodos de captura y recaptura

El problema de la estimación del tamaño poblacional en poblaciones silvestres ha sido un reto para los investigadores. Dada la naturaleza móvil de los animales, las técnicas de muestreo tradicional no resultan apropiadas. En estos casos, es prácticamente imposible contar todos los animales para calcular, de manera convencional, el tamaño de muestra, por lo cual es necesario realizar ajustes por subconteo (Chao, Tsay, Lin, Shau, & Chao, 2001).

Según lo plantean Chao et al. (2001), Laplace fue el primero en utilizar las técnicas de captura y recaptura con el fin de estimar el tamaño de la población de Francia en 1786. Lincoln, Petersen y Dahl, por su parte, fueron los primeros en aplicarlas a problemas ecológicos, mientras que Darroch estableció los fundamentos matemáticos

El modelo más básico, conocido como Lincoln-Petersen, intenta estimar la cantidad de individuos en la población, N , a partir de una cierta cantidad de animales marcados (n_1) y devueltos al sistema en una primera ocasión de

muestreo. En una segunda ocasión, se obtiene una muestra de tamaño n_2 , de la cual una cantidad $m \leq n_2$ se encuentran marcados. Si se supone que $m/n_2 = n_1/N$ se puede obtener un estimador $\hat{N} = n_1 n_2 / m$, con varianza aproximada $(n_1 + 1)(n_2 + 1)(n_1 - m)(n_2 - m) / [(m + 1)^2(m + 2)]$ (Chao, Pan y Chiang, 2008).

Cabe agregar que para obtener resultados válidos de dicho modelo es necesario que se cumplan los siguientes supuestos: A) la población es cerrada, por tanto, ningún individuo sale o entra a la misma, ya sea por migraciones, muertes o nacimientos. B) todos los animales tienen la misma probabilidad de ser capturados en cada muestreo. C) no hay pérdida de marcas y estas son siempre visibles al observador (Pollock, 2000). Aunque las poblaciones naturales, bajo circunstancias específicas, pueden cumplir una o varias de estas condiciones, no puede obviarse que dichas restricciones pueden ser limitantes en el contexto de producción comercial productivo que se planteó anteriormente.

Las siguientes secciones del marco teórico se desarrollarán bajo el siguiente esquema: primero, se explorará una serie de extensiones al modelo básico de Petersen-Lincoln con el fin de relajar el supuesto de que la población sea cerrada. Luego, se hará hincapié en la precisión y sesgo de estos modelos. Seguidamente, se procederá a comentar sobre las técnicas para relajar otros supuestos, como el de la homogeneidad de probabilidades de captura y sobrevivencia.

De esta manera, la anterior organización prevé desarrollar los temas teóricos necesarios para proponer posteriormente un modelo de captura y recaptura adaptado a la producción acuícola industrial.

3.3.1 Técnicas para el análisis de poblaciones abiertas

En la presente sección se analizarán los modelos aplicables al análisis de poblaciones abiertas; sin embargo, no se busca hacer una revisión exhaustiva de todos los métodos disponibles, sino solamente de los más aplicables o relacionados al problema original de estudio. Dicha sección se estructura para facilitar la comprensión de las relaciones entre los diferentes modelos.

3.3.1.1 Modelo de Jolly-Seber (JS)

Este modelo permite estimar la abundancia (o tamaño poblacional), sobrevivencia, tasa de dilución (nacimientos + inmigración), tasa de mortalidad aparente (mortalidad + emigración) y tasas de captura. Adicionalmente, algunas extensiones multigrupo permiten estimar tasas de reclutamiento entre cohortes, tasas de crecimiento poblacional, entre otros parámetros de interés para el investigador (Schwarz, 2001).

Esta metodología se basa en el análisis de t ocasiones de muestreo, donde se capturan individuos de la población. Cabe decir que el término captura, en este caso, se refiere a encontrar al individuo mediante una captura propiamente dicha (trampa, por ejemplo), un avistamiento (cámaras trampa), que aparezca en una lista, entre otras. Con esta metodología es importante que se pueda identificar de manera única el individuo, ya sea mediante collares de telemetría, marcas corporales, bandas, etcétera. Si el individuo capturado no está marcado, se procederá a marcarlo con un identificador único y liberarlo.

Cada individuo asociado con el estudio tendrá una historia de captura ω , que es un vector de tamaño t , donde $\omega_i = 1$ si el animal fue capturado y $\omega_i = 0$ en caso contrario ($i \in \{1, 2, \dots, t\}$). Dado que hay 2^t historias de captura posibles, en vez de tener un archivo con n líneas, una para cada individuo capturado al menos una vez, se puede obtener las frecuencias para cada historia de captura particular (denotada X_ω) excepto la historia donde $\omega_i = 0$ para todo $i \in \{1, 2, \dots, t\}$ (Link y Barker, 2005). Es común que se agreguen líneas adicionales con frecuencias negativas para indicar la cantidad de individuos que se capturaron pero no se liberaron al entorno para una cierta historia de captura (por ejemplo, murieron durante el marcaje o fueron removidos por el investigador) (Williams et al., 2002). Para utilizar este método se requiere definir los siguientes estadísticos suficientes (Krebs, 1999; Schwarz, 2001):

- m_i : número de individuos capturados en el momento i que habían sido previamente marcados

- u_i : número de individuos capturados en el momento i que no habían sido previamente marcados
- n_i : número de individuos capturados en el momento i ($m_i + u_i$)
- R_i : número de individuos liberados en el momento i (n_i menos pérdidas del muestreo más inyecciones de individuos)
- $r_j = \sum_{j=i+1}^K m_{ij}$: total de individuos liberados en el período i que son recapturados en algún momento posterior.
- z_j : conteo de individuos capturados antes del período i , no capturados en i pero en algún momento posterior a i .
- d_i, d'_i : son el conteo de individuos (marcados y no marcados, respectivamente) que no se liberan después de la captura en el período i

Con lo cual, se pueden definir los siguientes parámetros (Schwarz, 2001; Williams et al., 2002):

- p_i : probabilidad de captura en el momento i
- φ_i : probabilidad de que un individuo permanezca en la población entre los muestreos i e $i + 1$, dado que estaba vivo en i .
- N_i : tamaño de la población o abundancia al tiempo i .
- B_i : número de animales que entran en la muestra en el tiempo i , sobreviviendo hasta $i + 1$. Esta cantidad es conocida como nacimientos netos, aunque incluye todas las fuentes de incremento de población.
- λ_i : tasa de adición de nuevos individuos (natalidad más inmigración) entre i e $i + 1$, también conocida como tasa de dilución.
- ψ_i : probabilidad de que un individuo que entró a la población esté todavía vivo y no sea visto después del i -ésimo momento de captura.
- U_i : número total de individuos sin marcas en tiempo i .
- η_i, η'_i las probabilidades de liberación de los individuos capturados (marcados y no marcados, respectivamente) después de la captura en el período i

Williams et al. (2002) indican que la función de distribución de las variables observadas $\{u_i\}, \{d_i, d'_i\}, \{m_{ij}\}$ puede describirse a partir de tres componentes:

1. Un componente que considera la probabilidad de captura de animales no marcados a partir de una distribución binomial:

$$P(\{u_i\}|\{U_i\}, \{p_i\}) = \prod_{i=1}^k \binom{U_i}{u_i} p_i^{u_i} (1 - p_i)^{U_i - u_i} \quad (1)$$

2. Un componente que calcula la probabilidad de que algún individuo, marcado o no, no sea devuelto a la población:

$$P(\{d_i, d'_i\}|\{m_i, u_i\}, \{\eta_i, \eta'_i\}) \\ = \prod_{i=1}^k \binom{u_i}{d'_i} \eta_i^{u_i - d'_i} (1 - \eta'_i)^{d'_i} \times \prod_{i=1}^k \binom{m_i}{d_i} \eta_i^{m_i - d_i} (1 - \eta_i)^{d_i} \quad (2)$$

3. Un tercer componente que estima la probabilidad de que un individuo sea capturado, dado que fue capturado previamente:

$$P(\{m_{ij}\}|\{R_i\}, \{\varphi_i\}, \{p_i\}) \\ = \prod_{i=1}^{K-1} \frac{R_i}{(m_{i,i+1})! (m_{i,i+2})! \dots (m_{i,K})! (R_i - r_i)!} (\varphi_i p_{i+1})^{m_{i,i+1}} \\ \times \{[\varphi_i (1 - p_{i+1}) \varphi_{i+1} p_{i+2}]^{m_{i,i+2}} \dots [\varphi_i (1 - p_{i+1}) \dots \\ \times \varphi_{K-1} p_K]^{m_{i,K}} \chi_i^{R_i - r_i}\} \quad (3)$$

Con lo cual, la función de distribución de probabilidad es:

$$P(\{u_i\}, \{d_i, d'_i\}, \{m_{ij}\}) \\ = P(\{u_i\}|\{U_i\}, \{p_i\}) \times P(\{d_i, d'_i\}|\{m_i, u_i\}, \{\eta_i, \eta'_i\}) \\ \times P(\{m_{ij}\}|\{R_i\}, \{\varphi_i\}, \{p_i\}) \quad (4)$$

Krebs (1999) indica que se pueden obtener los siguientes estimadores a partir del modelo anterior:

1. Abundancia o tamaño poblacional: $\hat{N}_i = \hat{M}_i / \hat{\alpha}_i$, donde \hat{M}_i es el tamaño de la población marcada para el momento i , y su fórmula es $\hat{M}_i = m_i + z_i(R_i + 1)/(r_i + 1)$. Por su parte, $\hat{\alpha}_i$ es la proporción de animales marcados para el i -ésimo momento de muestreo, que se obtiene mediante la fórmula $\hat{\alpha}_i = (m_i + 1)/(n_i + 1)$.

2. Probabilidad de supervivencia: en realidad se puede plantear como una tasa de pérdida de individuos en la población (es una supervivencia aparente). Se estima como $\hat{\varphi}_i = \hat{M}_{i+1} / [\hat{M}_i + (R_i - m_i)]$.
3. La tasa de dilución: $\hat{\lambda} = \hat{N}_{i+1} / \{\hat{\varphi}_i [\hat{N}_i - (n_i - R_i)]\}$.
4. El número de adiciones a la población (nacimientos e inmigración) $\hat{B}_i = \hat{N}_{i+1} - \hat{\varphi}_i [\hat{N}_i - (n_i - R_i)]$

Este modelo asume probabilidades iguales de captura para todos los individuos en cada momento de muestreo, así como supervivencia idéntica para todos los individuos entre el período i e $i + 1$. De igual manera, se asume que no hay pérdida de marcas ni errores de avistamiento de estasy que cada momento de muestreo es casi instantáneo, por lo que las probabilidades de mortalidad o nacimientos durante el período de muestreo i son despreciables (Krebs, 1999).

Una variante interesante de considerar del modelo de Jolly-Seber es cuando se puede asumir que no hay migración ni nacimientos en la población, solo muertes. En este caso, Jolly (1965) propone los siguientes estimadores:

$$\hat{N}_i = \frac{R_i + 1}{r_i + 1} z'_i + n_i \quad (5)$$

$$\hat{\varphi}_i = \frac{\hat{N}_{i+1}}{\hat{N}_i - n_i + R_i} \quad (6)$$

$$\hat{p}_i = \frac{n_i}{\hat{N}_i} \quad (7)$$

Aquí $z'_i = z_i + \sum_{j=i+1}^S u_j$ con $j = i+1, \dots, K$.

Este modelo presenta ventajas, por ejemplo, permite la estimación de abundancia y supervivencia en poblaciones abiertas. Adicionalmente, no es condicional a R_1 , la cantidad de individuos liberados en $i = 1$ (como en otros modelos que se verán más adelante) más bien incluye la probabilidad de capturar individuos no marcados previamente (Williams et al., 2002). Empero, no está adaptado para la inclusión de probabilidades de captura desiguales ni considera el número de peces

sembrado, lo cual podría dar lugar a estimaciones que sean incluso superiores que el número inicial de peces.

Con respecto a la posibilidad de usar la información de conteos parciales de mortalidad, es necesario evaluar si tales conteos se pueden acumular y considerar como individuos capturados pero no devueltos a la población.

A continuación, se presentan algunos métodos alternativos que intentan solucionar, en parte, dichas falencias, basándose en la idea misma del modelo de Jolly-Seber (JS).

3.3.1.2 Modelo de Cormack-Jolly-Seber (CJS)

Jolly y Seber, plantearon, de manera independiente en 1965, una extensión a la propuesta de Cormack (1964) para hacer estimaciones por muestreo de captura y recaptura en poblaciones abiertas (Link y Barker, 2005).

Tal como se muestra en el Cuadro 1, es común representar los datos de captura-recaptura como matrices triangulares, donde se indica el número de individuos liberados en una ocasión de captura (R_i) y las recapturas de estos en periodos subsecuentes (m_{ij}):

Cuadro 1. Matriz triangular de estadísticos suficientes para el modelo de Cormack-Jolly-Seber

Periodo	Liberaciones	Periodos de captura		
		2	3	4
1	R_1	m_{12}	m_{13}	m_{14}
2	R_2		m_{23}	m_{24}
3	R_3			m_{34}

Fuente: Williams et al. (2002)

Si un animal es marcado en algún período, pero no se libera al entorno, simplemente no se suma al valor R_i correspondiente.

En relación con lo anterior, Williams et al. (2002) señalan que si se definen los siguientes estadísticos:

- $m_j = \sum_{i=1}^{j-1} m_{ij}$ Número de individuos marcados que fueron capturados en el período de muestreo j .
- n_i Cantidad total de animales capturados en el período i (marcados o no)

- $r_j = \sum_{j=i+1}^K m_{ij}$ Total de individuos liberados en el período i que son recapturados en algún momento posterior.
- z_j Conteo de individuos capturados antes del período i , no capturados en i pero en algún momento posterior a i .
- M_i Variable aleatoria que indica la cantidad de animales marcados en la población al momento del muestreo i .

Se puede obtener la función de verosimilitud de los datos de recaptura, m_{ij} , condicional a las liberaciones (R_i) y los parámetros de sobrevivencia y capturabilidad previamente definidos, para un estudio de K períodos:

$$\begin{aligned}
 & P(\{m_{ij}\}|\{R_i\}, \{\varphi_i\}, \{p_i\}) \\
 &= \prod_{i=1}^{K-1} \frac{R_i}{(m_{i,i+1})! (m_{i,i+2})! \dots (m_{i,K})! (R_i - r_i)!} (\varphi_i p_{i+1})^{m_{i,i+1}} \\
 & \times \{[\varphi_i(1 - p_{i+1})\varphi_{i+1}p_{i+2}]^{m_{i,i+2}} \dots [\varphi_i(1 - p_{i+1}) \dots \\
 & \times \varphi_{K-1}p_K]^{m_{i,K}} \chi_i^{R_i - r_i}\}
 \end{aligned} \tag{8}$$

Donde χ_i es la probabilidad de que un individuo vivo y en la población de estudio en el período de muestreo i no sea capturado en ningún momento posterior al período i . Nótese que ((8) es equivalente a (3), pues la probabilidad de recaptura, condicionada por la previa liberación, es la única parte de interés en el modelo CJS.

Además, se pueden obtener los siguientes estimadores de forma cerrada:

- El estimador de la población marcada en el período i : $\widehat{M}_i = m_i + R_i Z_i / r_i$
- La proporción capturada o probabilidad de captura: $\widehat{p}_i = m_i / \widehat{M}_i$
- La sobrevivencia entre el período i y el período $i+1$: $\varphi_i = \frac{\widehat{M}_{i+1}}{\widehat{M}_i - m_i + R_i}$

Algunas características importantes en el modelo CJS son las siguientes:

1. Sus estimadores son condicionales a la captura, o sea, sus estimaciones de sobrevivencia están completamente basadas en los individuos previamente capturados (Schwarz, 2001; Seber y Schofield, 2019).

2. La tasa de sobrevivencia calculada es aparente, pues no se puede separar el efecto de las posibles emigraciones (Williams et al., 2002). De hecho, más que sobrevivencia, podría denominarse una tasa de permanencia en la población.
3. Williams et al. (2002) indican que el modelo CJS asume que:
 - a. Todos los individuos marcados en la población en el período i tienen la misma probabilidad p_i de ser recapturada o reobservada (homogeneidad de probabilidades de captura).
 - b. Todos los individuos presentes en la población en el momento inmediato posterior al período de muestreo i tienen la misma probabilidad φ_i de sobrevivir al período $i+1$ (homogeneidad de probabilidades de sobrevivencia).
 - c. Las marcas no se pierden ni se ignoran, además son reportadas correctamente.
 - d. Los períodos de muestreo son instantáneos u ocurren en períodos muy cortos por lo cual, la migración o mortalidad que ocurren durante los mismos son despreciables.
 - e. Toda emigración del área muestreada es permanente.
 - f. Las probabilidades de sobrevivencia y captura de cada animal son independientes de otros animales.

En especial, los dos primeros supuestos suelen ser muy restrictivos en términos de lo que ocurre en las poblaciones naturales. Diferentes metodologías han sido propuestas para relajar estos supuestos, tales como modelos multi-estado, múltiples grupos, inclusión de covariables individuales o dependientes del tiempo, modelos de respuesta a la captura, entre otros; cuyo detalle puede explorarse más a fondo en Williams et al. (2002).

En el caso particular del interés de esta tesis, es importante señalar que dicho modelo aventaja al JS en que, al hacer una estimación de la sobrevivencia únicamente, nunca podría arrojar estimaciones de abundancia superiores al número inicial sembrado, ya que si se quiere conocer la abundancia, simplemente se multiplica la sobrevivencia por el número de peces sembrado y se obtiene el

resultado. Además, los intervalos de confianza de dicho estimador están acotados entre 0 y 1, con lo cual los intervalos de confianza también tendrían sentido. Empero, el hecho de que las estimaciones sean condicionales a los individuos capturados permite preguntarse si es correcto suponer la misma sobrevivencia para todos los individuos del estanque, más considerando que podrían existir probabilidades de captura diferentes, debido a las relaciones de dominancia en el estanque. De este se extrae otro problema, pues como se mencionó, el modelo CJS supone capturabilidad homogénea.

3.3.1.3 Modelos de recuperación de marcas

Esta metodología se utiliza para la estimación de tasas de sobrevivencia en poblaciones explotadas (Williams et al., 2002). Consiste en hacer un marcaje al inicio de un período dado y, posteriormente, recuperar las marcas de animales que salen de la población, ya sea por mortalidad natural, depredación o cacería (Seber y Schofield, 2019).

Williams et al. (2002) consideran este modelo un caso particular del modelo CJS, ya que se enfoca en la estimación de tasas de sobrevivencia y probabilidades de muestreo, mas no otros parámetros como abundancia o ingreso de nuevos individuos a la población. Por eso, en ambos métodos es común el uso de modelos de probabilidad multinomial para las estimaciones de máxima verosimilitud, parámetros de importancia biológica que ayudan a simplificar el modelo y pruebas de bondad de ajuste para seleccionar el modelo que mejor se ajusta a los datos. Sin embargo, Williams et al. (2002) indican que sus principales diferencias son las siguientes:

- Mientras en el CJS hay posibilidad de recapturar un mismo individuo varias veces, en los modelos de recuperación de marcas únicamente se puede recuperar una sola vez.
- En el CJS el marcaje de nuevos animales y la recaptura de animales marcados ocurren al mismo tiempo en períodos claramente establecidos. En los modelos de recuperación de marcas los procesos de marcaje y

recuperación se dan en momentos separados y solo el marcaje ocurre en momentos claramente definidos.

- Lo anterior conlleva a que mientras en el CJS hay igual número de períodos de recaptura que de marcaje, en los modelos de recuperación de marcas hay más momentos de recuperación que de marcaje.
- En los modelos de recuperación de marcas se hace una estimación de S_{ij} , la sobrevivencia real en el período entre el marcaje i y la j -ésima ocasión de recuperación. Por su parte, en los modelos CJS solo se puede estimar la sobrevivencia aparente φ_{ij} , que está contaminada por los individuos que emigran e inmigran.
- Finalmente, los métodos de recuperación de marcas asumen que el número de individuos marcados es fijo (de hecho, es especificado *a priori*, como parte del diseño del estudio). Por su parte, en el CJS se considera que esta es una variable aleatoria.

El método, como lo describen Seber y Schofield (2019), consiste en marcar individuos al inicio de un período durante s períodos, con una marca única para cada uno. Se registra la recuperación de marcas por t períodos ($t \geq s$). Sea x_{ij} una variable indicadora que toma el valor de 1 si un individuo particular marcado en el período i se recupera en el período j ($j = i, i + 1, \dots, t$), y 0 en otro caso. Por tanto, un individuo particular generará un vector de captura¹ $\mathbf{x}_{ij} = (x_{ii}, x_{i,i+1}, \dots, x_{it})'$ que seguirá una distribución con una probabilidad multinomial:

$$P(\mathbf{x}_{ij}|\boldsymbol{\theta}) = \prod_{j=i}^{t+1} \pi_{ij}^{x_{ij}} \quad (9)$$

¹ Nótese el uso del término vector de captura y, más adelante, historias de captura. Dichos términos, tomados prestados del análisis de captura-recaptura, podrían no ser estrictamente correctos en el método de recuperación de marcas, pero tienen sentido considerando la relación intrínseca entre ambas metodologías.

Donde $\theta = (\pi_{ii}, \pi_{i,i+1}, \dots, \pi_{it})'$ es un vector de probabilidades de recuperación de las marcas (π_{ij}) .

Como lo indican Williams et al. (2002), si las historias de captura individuales se asumen independientes, la probabilidad de que un cierto número de marcas m_i colocadas en el período i sean recuperadas sería:

$$P(m_i|\pi_i) = \frac{R_i!}{\prod_j m_{(ij)}! (R_i - \sum_j m_{ij})!} \left(\prod_{j=1}^t \pi_{ij}^{m_{ij}} \right) \pi_{i,t+1}^{R_i - \sum_j m_{ij}} \quad (10)$$

Donde R_i es la cantidad de individuos marcados y liberados en el i -ésimo período. De lo anterior se desprende la función de verosimilitud indicada por Seber y Schofield (2019):

$$L(\theta) = P(\{m_{ij}\}|\theta) = \prod_{i=1}^s \left\{ \binom{R_i}{m_{ii}, \dots, m_{it}, m_{i,t+1}} \prod_{j=1}^{t+1} \pi_{ij}^{x_{ij}} \right\} \quad (11)$$

Williams et al. (2002) indican que si se supone que la sobrevivencia de un individuo no se ve afectada por el período en el que fue marcado, se puede expresar las probabilidades de recuperación π_{ij} de un individuo marcado en el período i y recuperado en el período j de la siguiente forma:

$$\pi_{ij} = \prod_{m=i}^{j-1} S_{ij} f_j \quad (12)$$

Donde f_j se define como el producto de la probabilidad de cosecha y la probabilidad de que la marca sea vista y reportada en el período j (probabilidad de recuperación). De esta forma, los datos se pueden visualizar como una matriz triangular, como se muestra en el Cuadro 2.

Cuadro 2. Matriz triangular para el manejo de los datos en el análisis de recuperación de marcas

Marcas en el período i	Recuperaciones en período j					No recuperadas
	1	2	3	...	t	
R_1	m_{11}	m_{12}	m_{13}	...	m_{t1}	$R_1 - \sum_j m_{1j}$
R_2		m_{22}	m_{23}	...	m_{t2}	$R_2 - \sum_j m_{2j}$
R_3			m_{33}	...	m_{t3}	$R_3 - \sum_j m_{3j}$
\vdots				...	\vdots	\vdots
R_s				...	m_{ts}	$R_s - \sum_j m_{sj}$

Fuente: Williams et al. (2002).

Donde los valores de las celdas m_{ij} tienen un valor esperado de $R_i \pi_{ij}$. Los parámetros \hat{S}_{ij} y \hat{f}_i se pueden obtener de esta matriz, previamente calculando los

estimadores $m_{i.} = \frac{\sum_{j=1}^t m_{ij}}{t}$, $m_{.j} = \frac{\sum_{i=1}^s m_{ij}}{s}$:

$$\hat{S}_{ij} = \frac{m_{i.}}{R_i} \left[1 - \frac{m_{.j}}{T_i} \right] \frac{R_{i+1}}{m_{i+1}}, \quad i = 1, \dots, t - 1 \quad (13)$$

3.3.1.4 Modelos de captura-recaptura con información auxiliar de recuperación de marcas

Williams et al. (2002) plantean, a partir del trabajo de Burnham (1993), que dada la relación de la información de los modelos de recuperación de marcas y el CJS es posible combinarlos para mejorar las predicciones en especies animales que están bajo explotación y, posiblemente, estimar la emigración permanente.

Siguiendo con la notación desarrollada anteriormente, deben definirse los estadísticos suficientes m_{ijl} y m_{ijd} que representan los animales liberados en la i -ésima ocasión de captura y encontrados vivos o muertos en la ocasión j , respectivamente. Así las cosas, la matriz triangular de datos se muestra como en el Cuadro 3.

Cuadro 3. Matriz triangular para el manejo de los datos en el modelo CJS con recuperación de marcas

Liberaciones	Encuentros en período j								
	1		2		3		t		
	C ¹	R. ²	C	R	C	R	C	R	
R_1	$-^3$	m_{11d}	m_{12l}	m_{12d}	m_{13l}	m_{13d}	...	m_{1tl}	m_{1td}
R_2				m_{22d}	m_{23l}	m_{23d}	...	m_{2tl}	m_{2td}
R_3						m_{33d}	...	m_{3tl}	m_{3td}
\vdots							...	\vdots	\vdots
R_s							...		m_{std}

Fuente: Williams et al. (2002).

1Recapturas de animales vivos

2Recuperaciones de las marcas animales muertos

3En el período 1 solo hay capturas y liberaciones, pero no recapturas

Defínanse, además, $r_{il} = \sum_{j=i+1}^k m_{ijl}$ y $r_{id} = \sum_{j=i}^k m_{ijd}$. Considerando lo anterior, Williams et al (2002) indican que se puede modelar el Cuadro 3 como el producto de las multinomiales de cada cohorte de liberación (R_i). Sean π_{ijl} y π_{ijd} las probabilidades de las celdas asociadas con cada m_{ijl} y m_{ijd} , respectivamente. Además, sea $\lambda_{il} = \sum_{j=i+1}^k \pi_{ijl}$ para $i = 1, \dots, k-1$ y $\lambda_{id} = \sum_{j=i}^k \pi_{ijd}$ para $i = 1, \dots, k$. Entonces:

$$\begin{aligned}
& P(\{m_{ijd}, m_{ijl}\} | \{R_i\}) \\
&= \prod_{i=1}^{k-1} \left(m_{i id}, m_{i, i+1, l}, \dots, m_{i kl}, m_{i kd}, R_i - r_{id} - r_{il} \right) \\
&\times \prod_{i=1}^{k-1} \left[\left(\prod_{j=i}^{k-1} \pi_{i jd} \pi_{i, j+1, l} \right) \left(\pi_{i kd} (1 - \lambda_{id} - \lambda_{il})^{R_i - r_{id} - r_{il}} \right) \right] \quad (14) \\
&\times \binom{R_k}{m_{k kd}} \left[\pi_{k kd}^{m_{k kd}} (1 - \pi_{k kd})^{R_k - m_{k kd}} \right]
\end{aligned}$$

A su vez, se pueden plantear los valores π_{ijl} y π_{ijd} como funciones de parámetros de importancia biológica, según fueron definidos previamente:

$$\pi_{i, i+1, l} = \varphi_i p_{i+1} \quad (15)$$

$$\pi_{ijl} = \varphi_i (1 - p_{i+1}) \times \dots \times \varphi_{j-2} (1 - p_{j-1}) \times (\varphi_{j-1} p_j), \quad \forall j > i + 1 \quad (16)$$

$$\pi_{iid} = f_i \quad (17)$$

$$\pi_{ijd} = S_i \dots S_{j-1} \gamma_{ij} f_i, \quad \forall j > i \quad (18)$$

Donde S_i es la probabilidad de sobrevivencia entre el período de captura i e $i + 1$, f_i es la probabilidad de recuperación de marcas en el periodo i (i.e. hallar el individuo muerto) y γ_{ij} es un parámetro no introducido anteriormente, que indica la probabilidad de que un animal liberado después de la i -ésima ocasión de captura y que permanece vivo en la ocasión j , no haya sido recapturado posteriormente. Este parámetro puede ser replanteado recursivamente en función de las probabilidades de captura p_i y la probabilidad de que el animal no emigre permanentemente entre el periodo i y el $i + 1$ (Fidelidad, F_i):

$$\gamma_{ij} = 1 - F_i + F_i(1 - p_{i+1})\gamma_{i+1,j} = 1 - F_i[1 - q_{i+1}\gamma_{i+1,j}] \quad (19)$$

Donde $1 - F_i$ es la probabilidad de que el animal emigre del área de estudio entre i e $i + 1$, y $F_i(1 - p_i)\gamma_{i+1,j}$ es la probabilidad de que aún quedándose en el área y sobreviviendo, no haya sido capturado. Además, $q_{i+1} = 1 - p_{i+1}$.

Con base en lo anterior, se pueden expresar los valores esperados para los m_{ijl} y m_{ijd} del Cuadro 3, como funciones de los parámetros $R_i, f_i, S_i, \gamma_{ij}$ y p_i (Cuadro 4).

Cuadro 4. Valores esperados para los conteos de recuperaciones y recapturas en un modelo de CJS con información adicional de recuperación de marcas

Número Liberado	Encuentros en período j					
	1		2		3	
	C ¹	R. ²	C	R	C	R
R_1	⁻³	$R_1 f_1$	$R_1 S_1 F_1 p_2$	$R_1 S_1 \gamma_{12} f_2$	$R_1 S_1 F_1 q_2 S_2 F_2 p_3$	$R_1 S_1 S_2 \gamma_{13} f_3$
R_2				$R_2 f_2$	$R_2 S_2 F_2 p_3$	$R_2 S_2 \gamma_{23} f_3$
R_3						$R_3 f_3$

Fuente: Williams et al. (2002).

¹Recapturas de animales vivos

²Recuperaciones de las marcas animales muertos

³En el período 1 solo hay capturas y liberaciones, pero no recapturas

De los modelos analizados hasta el momento, este es el que se aproxima más a lo deseado, ya que permite combinar la información de captura y recaptura con la recuperación de animales muertos, para estimar la sobrevivencia, con lo cual, las

estimaciones de abundancia para cualquier momento $t < 0$ siempre serían menores o iguales que el número de peces sembrado (abundancia en $t = 0$). Sin embargo, al igual que todos los modelos derivados del CJS, tiene la desventaja de ser condicional a las capturas y de asumir probabilidades de captura y sobrevivencia homogéneas.

3.3.2 Precisión de las estimaciones de abundancia y sobrevivencia del modelo de Jolly-Seber

Hasta el momento se ha procedido a presentar los estimadores puntuales en todos los modelos, mas no las varianzas de los estimadores. Lo anterior se debe a que se quería dedicar una sección particular del trabajo a analizar la complejidad de la obtención de la varianza de los estimadores de JS en situaciones reales.

Jolly (1965) plantea originalmente los siguientes estimadores, donde sustituye las variables aleatorias por sus esperanzas:

$$var(\hat{N}_i) = N_i(N_i - n_i) \left\{ \frac{M_i - m_i + s_i}{M_i} \left(\frac{1}{r_i} - \frac{1}{R_i} \right) + \frac{1 - \alpha_i}{m_i} \right\} + N_i - \sum_{h=0}^{i-1} N_i^2(h) \quad (20)$$

$$var(\hat{\phi}_i) = \phi_i^2 \left\{ \frac{(M_{i+1} - m_{i+1})(M_{i+1} - m_{i+1} + R_{i+1})}{M_{i+1}^2} \left(\frac{1}{r_{i+1}} - \frac{1}{R_{i+1}} \right) + \frac{M_i - m_i}{M_i - m_i - R_i} \left(\frac{1}{r_{i+1}} - \frac{1}{R_{i+1}} \right) + \frac{1 - \phi_i}{M_{i+1}} \right\} \quad (21)$$

$$var(\hat{N}_i) = N_i(N_i - n_i) \left\{ \frac{M_{i+1} - m_{i+1} + R_{i+1}}{M_{i+1}} \left(\frac{1}{r_{i+1}} - \frac{1}{R_{i+1}} \right) + \frac{1 - \alpha_i}{m_i} \right\} + N_i \sum_{h=0}^{i-1} \frac{N_i^2(h)}{B_h} \quad (22)$$

Estas varianzas obtenidas por Jolly (1965) asumen una aproximación normal, por lo que los intervalos de confianza de los estimadores siguen la forma convencional de $\theta \pm Z_{1-\frac{\alpha}{2}} \sqrt{var(\theta)}$, donde $Z_{1-\frac{\alpha}{2}}$ es un valor de la distribución normal estándar, hasta el cual se acumula $1 - \frac{\alpha}{2}$ del área bajo la curva normal; mientras que α es el complemento del nivel de confianza deseado.

Sin embargo, las fórmulas anteriores han sido criticadas, ya que estudios de simulación demuestran que si se subestima el parámetro de interés, también las varianzas tendrán subestimación, principalmente porque dichas fórmulas incluyen el parámetro que se está estimando (Manly, 1984). Adicionalmente, se señala que los intervalos de confianza podrían no ser simétricos en estos casos (por lo cual es irracional asumir una distribución normal del estimador) y que existe un sesgo importante en muestras pequeñas (Manly, 1984; Seber y Manly, 1985).

A raíz de lo anterior, Manly (1984) propone que la varianza del estimador de abundancia \hat{N}_i , anteriormente mencionado, se puede aproximar usando la metodología de series de Taylor a partir de la transformación $T_1(\hat{N}_i) = \ln(\hat{N}_i) + \ln\left[\frac{1}{2}(1 - p_i/2 + \sqrt{1 - p_i})\right]$, donde $p_i = n_i/\hat{N}_i$. La varianza de $T_1(\hat{N}_i)$ está dada por:

$$var[T_1(\hat{N}_i)] = \left(\frac{\hat{M}_i - m_i + R_i + 1}{\hat{M}_i + 1}\right) \left(\frac{1}{r_i + 1} + \frac{1}{R_i + 1}\right) + \frac{1}{m_i + 1} \quad (23)$$

En el caso del estimador de sobrevivencia, la transformación recomendada por Manly (1984) es: $T_2(\hat{\phi}_i) = \ln\left\{\left[1 - (1 - \hat{\gamma}_i\hat{\phi}_i)^{\frac{1}{2}}\right]\left[1 + (1 - \hat{\gamma}_i\hat{\phi}_i)^{\frac{1}{2}}\right]^{-1}\right\}$, donde $\hat{\gamma}_i = C_i/(B_i + C_i)$, y, a su vez:

$$B_i = \frac{(\hat{M}_{i+1} - m_{i+1} + 1)(\hat{M}_{i+1} - m_{i+1} + R_{i+1} + 1)}{(\hat{M}_{i+1} - 1)^2} \left(\frac{1}{r_{i+1} + 1} - \frac{1}{R_{i+1} + 1}\right) + \frac{(\hat{M}_i - m_i + 1)}{(\hat{M}_i - m_i + R_i + 1)} \left(\frac{1}{r_i + 1} - \frac{1}{R_i + 1}\right) \quad (24)$$

$$C_i = \frac{1}{\hat{M}_{i+1} + 1} \quad (25)$$

Entonces $var[T_2(\hat{\phi}_i)] = B_i + C_i$, que al igual que en el caso de $var[T_1(\hat{N}_i)]$ no dependen del valor estimado para ϕ o N , respectivamente.

El método anterior requiere consideraciones adicionales para obtener los intervalos de confianza. En el caso de $T_1(N_i)$, su intervalo de confianza sería:

$$T_1(\hat{N}_i) - k_1 \left(\text{var} \left(T_1(\hat{N}_i) \right) \right)^{\frac{1}{2}} < T_1(N_i) < T_1(\hat{N}_i) + k_2 \left(\text{var} \left(T_1(\hat{N}_i) \right) \right)^{\frac{1}{2}} \quad (26)$$

Con $k_1 = -2,4$ y $k_2 = 1,6$ para un 95% de confianza, mientras que para un 99% de confianza $k_1 = -2,2$ y $k_2 = 3$.

A su vez, el intervalo de confianza para $T_2(\varphi_i)$ está dado por:

$$T_2(\hat{\varphi}_i) - h_1 \left(\text{var}[T_2(\hat{\varphi}_i)] \right)^{\frac{1}{2}} < T_2(\varphi_i) < T_2(\hat{\varphi}_i) + h_2 \left(\text{var}[T_2(\hat{\varphi}_i)] \right)^{\frac{1}{2}} \quad (27)$$

Donde los valores de $h_1 = -1,9$ y $h_2 = 2,1$ se usan cuando se quiere obtener un intervalo de 95% de confianza, pero si se quiere un 99% de confianza se recomienda $h_1 = -2,5$ y $h_2 = 2,7$.

Un paso adicional para obtener los intervalos de confianza es retornar los valores a su escala original. Para esto se usa la fórmula $[4 \exp(L_i) + n_i]^2 [16 \exp(L_i)]^{-1}$, donde los L_i son los límites del intervalo de confianza de $T_1(\hat{N}_i)$. Para el caso de $T_2(\hat{\varphi}_i)$, el retorno a la escala original se obtiene con $\hat{\gamma}_i^{-1} \{1 - [1 - \exp(L_i)]^2 [1 + \exp(L_i)]^{-2}\}$.

Hay que decir que existe polémica en cuanto a cómo deberían estimarse las varianzas del modelo de JS. Pollock, Nichols, Brownie y Hines (1990), y también Seber y Schofield (2009), proponen la utilización de la metodología Jolly (1965), pues señalan que la propuesta de Manly (1984) es arbitraria. Sin embargo, Krebs (1999) apoya la metodología de Manly e indica que sus estimadores tienen un comportamiento satisfactorio en la práctica.

Otros autores como Buckland y Garthwaite (1991) y Williams et al. (2002) presentan ideas diferentes, como estimar la varianza e intervalos de confianza mediante *bootstrapping*. Adicionalmente, Williams et al. (2002) indican que de preferirse un estimador con forma cerrada, se recurra al siguiente:

$$\text{var}(\hat{N}_i) = \frac{n_i^2 \text{var}(\hat{p}_i)}{\hat{p}_i^4} + \frac{n_i(1 - \hat{p}_i)}{\hat{p}_i^2} \quad (28)$$

Donde n_i "es el número de animales capturados [...] tomado directamente del muestreo y las estimaciones de $\text{var}(\hat{p}_i)$ y \hat{p}_i que son calculadas por el software

usado para calcular un modelo particular (por ejemplo MARK)” (Williams et al., 2002, p. 503)². Para obtener los intervalos de confianza, recomiendan la metodología sugerida por Chao (1989), donde es necesario estimar $\hat{f}_{0i} = \hat{N}_i - n_i$, que corresponde al número de individuos no capturados en i . Luego $\ln(\hat{f}_{0i})$ es una variable aleatoria aproximadamente normal y se puede obtener el intervalo de confianza $\left[\hat{N}_i - \frac{\hat{f}_{0i}}{C}, \hat{N}_i + \hat{f}_{0i}C \right]$, donde:

$$C = \exp \left\{ Z_{1-\frac{\alpha}{2}} \left[\ln \left(1 + \frac{\text{var}(\hat{N}_i)}{\hat{f}_{0i}^2} \right)^{\frac{1}{2}} \right] \right\} \quad (29)$$

En resumen, es necesaria aún más investigación en esta área, para lograr un consenso acerca de los estimadores de varianza en el modelo de JS.

3.3.3 Sesgo en las estimaciones de abundancia y sobrevivencia a partir de JS y CJS

Jolly (1965) planteó su estimador de máxima verosimilitud para M_i como $\hat{M}'_i = m_i + z_i R_i / r_i$, señalando que el mismo presentaba un sesgo positivo e igual a $[E(M_i) - E(m_i)][E(r_i)^{-1} - E(R_i)^{-1}]$. Por tanto, los estimadores de abundancia y sobrevivencia originales también estarían sesgados. Trabajos posteriores (Jolly 1982 y Seber 1982, citados por Pollock et al., 1990) idearon los estimadores aproximadamente insesgados que se presentaron en apartados anteriores.

A pesar de lo anterior, dichos estimadores siguen siendo sesgados especialmente en condiciones de recapturas escasas y heterogeneidad en las probabilidades de captura. A continuación, se detalla más el sesgo y posibles soluciones en dichas condiciones.

² Traducción del autor. El texto, en su lenguaje original, se transcribe fielmente a continuación: “... the numbers of animals caught n_i , are the statistics that come directly from the sampling, and the estimates \hat{p}_i and $\text{var}(\hat{p}_i)$ are computed by the software used to fit the particular model (e.g. MARK)”

3.3.3.1 Sesgo por recapturas escasas

Seber y Schofield (2009) señalan que los estimadores aproximadamente insesgados tienen mejor control del sesgo cuando m_i y r_i son mayores a 10. Huggins, Stoklosa, Roach, y Yip (2018) se centran en este aspecto y ofrecen tres estimadores para N_i , basados en trabajos previos en modelos de población cerrada (Chao, 1989), para condiciones con recapturas escasas:

1. El estimador *local crudo de Chao*:

$$\hat{N}_1(t) = S(t, 1) + \frac{f_1(t, 1)^2 - \sum_{j=t-k_1}^{t+k_2} Z_j(t, 1)^2}{2\hat{f}_2(t, 1)}, \quad (30)$$

Donde:

- t es el momento de muestreo, con $t = 1, 2, \dots, \tau$
- $k_1 = \max(1, t - \delta)$, $k_2 = \min(1, t + \delta)$ donde δ es un número entero pequeño que sirve para definir un vecindario de t que debería tener un comportamiento similar de las variables estudiadas, lo cual facilita la estimación (generalmente $\delta = 1$).
- $f_n(t, 1)$ es la frecuencia absoluta de individuos capturados exactamente n veces en las ocasiones $t - 1, t, t + 1$. El estimador $\hat{f}_2(t, 1)$ se obtiene mediante mínimos cuadrados ordinarios.
- $Z_j(t, 1)$ es el número de individuos capturados únicamente en las ocasiones $t - 1, t, t + 1$.
- $S(t, 1)$ es el número de total de individuos capturados en las ocasiones $t - 1, t, t + 1$.

2. El estimador *local suavizado de Chao*:

$$\hat{N}_2(t) = S(t, 1) + \frac{f_1(t, 1)^2 - \sum_{j=t-k_1}^{t+k_2} Z_j(t, 1)^2}{2\tilde{f}_2(t, 1)}, \quad (31)$$

en el cual t , k_1 , k_2 , $Z_j(t, 1)$, $S(t, 1)$ y $f_n(t, 1)$ se definen igual que en (30) pero $\tilde{f}_2(t, 1)$ representa la estimación de $f_2(t, 1)$ obtenida por técnicas de suavizamiento (*thin plate regression splines*) a partir de:

$$\ln[E\{f_2(t, 1)\}] + (\nabla - 3) \ln(1 - p) = \ln\left\{\frac{6}{\nabla(\nabla - 1)}\right\} + \ln(f_2(t, \delta')). \quad (32)$$

En este caso, δ' es un entero entre 1 y Δ . Mayores valores de Δ favorecen el ajuste del modelo, pero introducen un sesgo positivo en la estimación de $f_2(t, 1)$. ∇ , por otro lado, equivale al número de ocasiones de muestreo en un vecindario dado ($2\delta' + 1$). Finalmente, el 3 corresponde al número de ocasiones de muestreo consideradas en el período $t - 1, t, t + 1$.

3. *El estimador doblemente suavizado de Chao* (Huggins et al., 2018), que aplica un suavizamiento adicional de la estimación de $\ln(\hat{N}_2(t))$ como una función de t .

Con estos estimadores Huggins et al. (2018) lograron importantes mejoras en resultados de simulación y datos reales, en comparación con el modelo JS.

3.3.3.2 *Sesgo por heterogeneidad de las probabilidades de captura.*

Tanto el modelo CJS como el JS suponen, como se indicó previamente, homogeneidad en las probabilidades de captura de los individuos. Sin embargo, es probable que esto no se cumpla en todos los casos y, aunque los estimadores de sobrevivencia son robustos a la violación de dicho supuesto, los estimadores de abundancia o tamaño poblacional no lo son (Pledger, Pollock y Norris, 2010; Williams et al., 2002).

Otis et al. (1978) señalan que las poblaciones naturales pueden desviarse del supuesto de capturabilidades iguales debido a:

- efectos del tiempo (t)
- efectos de comportamiento (b)
- heterogeneidad individual (h)

Dichos efectos pueden actuar de manera individual o en conjunto en una población, por lo que se debe determinar si se presentan en una población para aplicar las correcciones correspondientes. Así se pueden identificar diferentes modelos de capturabilidad: M_0 = Capturabilidades homogéneas, M_b = Capturabilidades desiguales por comportamiento en respuesta a la captura, M_t =

Capturabilidades desiguales por efectos temporales, M_h = Capturabilidades desiguales por heterogeneidad individual, y sus combinaciones (M_{bh} , M_{th} , M_{bt} , M_{tbh}). En cuanto a esto, es importante el criterio biológico para definir la presencia o ausencia de alguno de estos efectos (Otis et al., 1978).

Cuando se habla de diferencias en la probabilidad de captura debidas a comportamiento en respuesta a la captura, Otis et al. (1978) se refieren a diferencias en la probabilidad de captura debidas a la experiencia previa de haber sido capturados. Después de la primera captura y del manejo que se haya dado durante esta, se pueden presentar situaciones en las que el individuo previamente capturado, tenga más propensión a aparecer en capturas subsecuentes, en comparación con los que fueron capturados, efecto conocido como *trap happy*. Esto podría suceder si la recompensa por la captura es beneficiosa para el animal (por ejemplo, puede obtener algún alimento). Caso contrario, en aquellos en que la captura resulte en un evento traumático o poco gratificante para el individuo y que provoque que los individuos previamente capturados sean reacios a caer en muestreos posteriores, se le denomina *trap shy*.

Por otra parte, las capturabilidades desiguales por efectos temporales puede ser una respuesta a condiciones climáticas adversas que cambien la probabilidad de que algunos individuos aparezcan en ciertas muestras, pero aumenten su probabilidad en condiciones atmosféricas diferentes. Otra causa podría ser el uso de métodos de muestreo diferentes en cada ocasión (Otis et al., 1978).

Finalmente, la heterogeneidad individual en las probabilidades de captura surge de posibles diferencias de jerarquía social o comportamiento en distintos grupos de la población, debido a sus características intrínsecas, como sexo, edad o tamaño. Para mejorar la estimación en estos casos, es necesario utilizar información adicional que pueda ingresar al modelo como covariables (Otis et al., 1978).

Algunos autores indican que los estimadores de sobrevivencia de CJS y JS son relativamente robustos a violaciones de esos supuestos (Pledger, Pollock y Norris, 2003; Williams et al., 2002); sin embargo, estudios de Pollock y Ravelling (1982) y

Nichols et al. (1982) (citados por Pledger et al., 2003) han mostrado que el nivel de sesgo podría ser importante si: a) las probabilidades de sobrevivencia están correlacionadas entre ocasiones de muestreo o b) hay correlación entre la probabilidad de captura y la probabilidad de sobrevivencia.

Por otro lado, Devineau, Choquet y Lebreton (2006) muestran que a mayor diferencia en las probabilidades de captura, mayor sesgo en la estimación de la sobrevivencia, en especial si se consideran modelos más complejos (por ejemplo, sobrevivencias diferentes para cada ocasión de muestreo o uso de covariables, en comparación con sobrevivencia constante).

Fletcher et al. (2012) refuerzan lo anterior, mostrando variedad de casos:

- Si se asumen dos grupos con capturabilidades diferentes (p_1 y p_2 , respectivamente), el sesgo en la estimación de ϕ es negativo, e incrementa concomitantemente con el aumento en $|p_1 - p_2|$. En este caso, si $|p_1 - p_2|$ se mantiene fijo, la magnitud del sesgo en la estimación de ϕ empeora cuando p_1 y p_2 disminuyen.
- En modelos con respuesta a la captura, donde p_0 es la probabilidad de captura de los individuos no atrapados en la ocasión de captura anterior y p_1 es dicha magnitud para los individuos que cayeron en la trampa en la ocasión de captura anterior, se tiene que los aumentos en $|p_0 - p_1|$ empeoran el sesgo. Si $p_0 < p_1$, el sesgo es negativo y si $p_0 > p_1$, es positivo.
- Adicionalmente estimaron el sesgo en modelos con emigración temporal entre dos áreas, con una probabilidad de emigración del área 1 al área 2 de a , probabilidad de captura p en el área 1 y 0 en el área 2. En este caso, si $a < 0,5$, el sesgo es negativo y si $a > 0,5$, el sesgo es positivo. Además, el sesgo aumenta cuando $|a - 0,5|$ aumenta.

Otros estudios en esta área probaron el efecto de asumir probabilidades de captura iguales en condiciones de: a) probabilidades de captura simétricas alrededor de un valor medio, b) asimetría positiva, c) asimetría negativa, d) dos grupos con $p_1 = 0,2$ y $p_1 = 0,8$ (Abadi, Botha y Altwegg, 2013). En todas sus estimaciones, el sesgo fue negativo, ya que muestra valores menores a 2% en los

escenarios a y c, pero mayores a 6% en los otros escenarios. Estos autores señalan que, aunque en algunos casos el nivel de sesgo puede ser pequeño, ignorar la heterogeneidad podría dar lugar a malas decisiones de manejo. Analizando dos casos reales, determinaron que el efecto de incluir la heterogeneidad en el modelo es pequeño y tiene mayor importancia en las ocasiones de muestreo finales que en las iniciales.

3.3.4 Técnicas para el manejo de heterogeneidad en las probabilidades de captura

Los modelos de captura-recaptura en poblaciones cerradas fueron los primeros en desarrollar técnicas para el manejo de probabilidades de captura desiguales, lo que se puede constatar en el importante desarrollo que hacen Otis et al. (1978). Pollock (1982), como se cita en Williams et al. (2002), propone combinar tales avances con el modelo CJS para poder dar seguimiento a una población a través del tiempo, en lo que actualmente se conoce como el *modelo robusto*.

Dicho modelo se compone de una estructura anidada, con K ocasiones de muestreo primario, las cuales se dan en un intervalo de tiempo suficientemente amplio como para considerar la población como abierta. En cada ocasión primaria de muestreo, se realizan I ocasiones secundarias de muestreo, espaciadas de tal forma que la población pueda considerarse cerrada en medio de estas (Williams et al., 2002).

El planteamiento original de Pollock, por un lado, fue estimar la abundancia de la población basado en la información de cada ocasión primaria de muestreo, usando modelos de población cerrada. Por otro lado, combinando la información en cada ocasión primaria de muestreo, de forma que un individuo era marcado como capturado si aparecía en al menos una ocasión secundaria de muestreo y como no capturado en otro caso. Usando estas historias de captura combinadas, se podrían obtener estimadores de sobrevivencia de CJS. Adicionalmente, el modelo permitiría obtener la tasa de nacimientos o nuevos reclutamientos en las diferentes categorías poblacionales, mediante la comparación de la tasa de sobrevivencia con los estimadores de abundancia obtenidos. Esto lo convierte en un modelo *ad*

hoc, más que uno basado en una sola función de verosimilitud (Williams et al., 2002). En cuanto al modelo robusto, Kendall et al. (1995) plantean un acercamiento formal mediante una función de verosimilitud para dicho modelo. Adicionalmente, extienden los planteamientos de Otis et al. (1978) mediante una nomenclatura que permite considerar distintas clases de fuentes de heterogeneidad entre y dentro de ocasiones primarias de muestreo, donde los superíndices reflejan la fuente de variación entre ocasiones primarias de muestreo y los subíndices lo que sucede dentro de cada ocasión primaria (Cuadro 5). Algunas de las combinaciones posibles de modelos no se encuentran en el Cuadro 5, por ejemplo, ningún modelo usa heterogeneidad individual entre ocasiones primarias de muestreo, esto por cuanto los autores consideraron que dichas combinaciones no eran plausibles en la realidad.

Cuadro 5. Modelos posibles que usan la información completa de historias de captura en el diseño robusto

Modelo	Fuentes de variación en probabilidad de captura	
	Dentro	Entre
M_0^0	Ninguna	Ninguna
M_0^t	Ninguna	Tiempo
M_t^0	Tiempo	Ninguna
M_t^t	Tiempo	Tiempo
M_b^0	Comportamiento	Tiempo
M_b^b	Comportamiento	Comportamiento
M_b^{tb}	Comportamiento	Tiempo y comportamiento
M_h^0	Heterogeneidad individual	Ninguna
M_h^t	Heterogeneidad individual	Tiempo
M_{tb}^0	Tiempo y comportamiento	Ninguna
M_{tb}^t	Tiempo y comportamiento	Tiempo
M_{tb}^b	Tiempo y comportamiento	Comportamiento
M_{tb}^{tb}	Tiempo y comportamiento	Tiempo y comportamiento
M_{th}^0	Tiempo y heterogeneidad individual	Ninguna
M_{th}^t	Tiempo y heterogeneidad individual	Tiempo
M_{bh}^0	Comportamiento y heterogeneidad individual	Ninguna
M_{bh}^t	Comportamiento y heterogeneidad individual	Tiempo
M_{bh}^{tb}	Comportamiento y heterogeneidad individual	Tiempo y comportamiento
M_{tbh}^0	Todas	Ninguna
M_{tbh}^t	Todas	Tiempo
M_{tbh}^b	Todas	Comportamiento
M_{tbh}^{tb}	Todas	Comportamiento y tiempo

Fuente: Kendall et al. (1995).

Ahora bien, ¿qué pasa si no se puede o no se quiere usar este esquema de doble muestreo? La alternativa es usar el modelo CJS/JS, pero intentando modificarlo.

McDonald y Amstrup (2001) enfrentan el problema del sesgo en los estimadores de abundancia replanteando los mismos mediante un estimador de tipo Horvitz-Thompson $\hat{N}_j = \sum_{i=1}^n x_{ij} / \hat{p}_{ij}$, donde x_{ij} es una variable indicadora que toma valor de 1 si el individuo i es capturado en la j -ésima ocasión de captura, y 0 si no es capturado. Los \hat{p}_{ij} se estiman a partir de un modelo logístico basado en k covariables individuales:

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_1 x_{ij1} + \dots + \beta_k x_{ijk} \quad (33)$$

y se sigue una metodología similar para estimar la sobrevivencia ($\hat{\phi}_{ij}$). La varianza de \hat{N}_j está dada por:

$$\text{var}(\hat{N}_j) = \sum_{i=1}^n \left[\frac{x_{ij}(1-\hat{p}_{ij})}{\hat{p}_{ij}^2} + \frac{x_{ij}\sigma_{\hat{p}_{ij}}^2}{\hat{p}_{ij}^3} + \frac{x_{ij}(1-\hat{p}_{ij})\sigma_{\hat{p}_{ij}}^2}{\hat{p}_{ij}^4} \right] \quad (34)$$

Pledger et al. (2003) replantean el modelo CJS para incluir la heterogeneidad de las probabilidades de captura y sobrevivencia a partir de mezclas finitas. Para ello, indican que cada historia de captura (CH_i) tendrá una probabilidad de ocurrencia, condicionada a la primera captura (f_i) dada por:

$$P(CH_i|f_i) = \sum_{d=l_i}^K \left\{ \left(\prod_{j=f_i}^{d-1} \varphi_j \right) (1-\varphi_d) \left(\prod_{j=f_i+1}^d p_j^{x_{ij}} (1-p_j)^{1-x_{ij}} \right) \right\}, \quad (35)$$

donde d es el momento en que el animal sale de la población (por muerte o emigración), que es un valor desconocido mayor o igual a l_i , el último período en el que el individuo fue observado vivo después de la primera captura f_i (Pledger et al., 2003).

En este, se supone que la población se puede dividir en C clases, cuyos miembros no necesariamente se conocen *a priori*, sin embargo, cada individuo tiene una probabilidad π_c de pertenecer a la clase c . Cada clase está definida por una capturabilidad de p_{jc} y sobrevivencia de φ_{jc} , donde c toma valores entre 1 y C . Por tanto, la contribución de cada individuo a la función de verosimilitud global es la suma de los productos de la probabilidad de que ese animal pertenezca a una

clase c con la probabilidad de su historia de captura, dada esa clase y el momento de su primera captura. Por tanto, la verosimilitud global para los n animales se define como:

$$L = \prod_{i=1}^n \sum_{c=1}^C \sum_{d=l_i}^K \left\{ \pi_c \left(\prod_{j=f_i}^{d-1} \varphi_j \right) (1 - \varphi_d) \left(\prod_{j=f_i+1}^d p_j^{x_{ij}} (1 - p_j)^{1-x_{ij}} \right) \right\}. \quad (36)$$

Cuadro 6. Diferentes modelos que consideran heterogeneidad en la sobrevivencia y/o probabilidad de sobrevivencia, con K muestras y C clases.

Modelo	Número de parámetros independientes estimables			
	π	φ	p	Total
$\{\phi(\cdot), p(\cdot)\}$	0	1	1	2
$\{\phi(\cdot), p(t)\}$	0	1	$K - 1$	K
$\{\phi(\cdot), p(hc)\}$	$C - 1$	1	C	$2C$
$\{\phi(\cdot), p(t + hc)\}$	$C - 1$	1	$K + C - 2$	$K + 2C - 2$
$\{\phi(\cdot), p(t \times hc)\}$	$C - 1$	1	$(K - 1)C$	KC
$\{\phi(t), p(\cdot)\}$	0	$K - 1$	1	K
$\{\phi(t), p(t)\}$	0	$K - 1$	$K - 1$	$2K - 3^*$
$\{\phi(t), p(hc)\}$	$C - 1$	$K - 1$	C	$K + 2C - 2$
$\{\phi(t), p(t + hc)\}$	$C - 1$	$K - 1$	$K + C - 2$	$2K + 2C - 4$
$\{\phi(t), p(t \times hc)\}$	$C - 1$	$K - 1$	$(K - 1)C$	$KC + K - 3^*$
$\{\phi(hc), p(\cdot)\}$	$C - 1$	C	1	$2C$
$\{\phi(hc), p(t)\}$	$C - 1$	C	$K - 1$	$K + 2C - 2$
$\{[\phi(h), p(h)]_c\}$	$C - 1$	C	C	$3C - 1$
$\{[\varphi(h), p(t + h)]_c\}$	$C - 1$	C	$K + C - 2$	$K + 3C - 3$
$\{[\phi(h), p(t \times h)]_c\}$	$C - 1$	C	$(K - 1)C$	$KC + C - 1$
$\{\phi(t + hc), p(\cdot)\}$	$C - 1$	$K + C - 2$	1	$K + 2C - 2$
$\{\phi(t + hc), p(t)\}$	$C - 1$	$K + C - 2$	$K - 1$	$2K + 2C - 4$
$\{[\phi(t + h), p(h)]_c\}$	$C - 1$	$K + C - 2$	C	$K + 3C - 3$
$\{[\phi(t + h), p(t + h)]_c\}$	$C - 1$	$K + C - 2$	$K + C - 2$	$2K + 3C - 5$
$\{[\phi(t + h), p(t \times h)]_c\}$	$C - 1$	$K + C - 2$	$(K - 1)C$	$KC + K + C - 4$
$\{\phi(t \times hc), p(\cdot)\}$	$C - 1$	$(K - 1)C$	1	KC
$\{\phi(t \times hc), p(t)\}$	$C - 1$	$(K - 1)C$	$K - 1$	$KC + K - 3^*$
$\{[\phi(t \times h), p(h)]_c\}$	$C - 1$	$(K - 1)C$	C	$KC + C - 1$
$\{[\phi(t \times h), p(t + h)]_c\}$	$C - 1$	$(K - 1)C$	$K + C - 2$	$KC + C - 4^*$
$\{[\phi(t \times h), p(t \times h)]_c\}$	$C - 1$	$(K - 1)C$	$(K - 1)C$	$2KC - 2C - 1^{**}$

Fuente: *Pledger et al. (2003)*

La notación $[\phi p]_c$ indica mezcla de C clases para ambos parámetros

*Un parámetro no es identificable

** C parámetros no son identificables

El modelo expresado en (36) es una mezcla finita de C componentes que permite que tanto la sobrevivencia como la probabilidad de captura presenten efectos multiplicativos de heterogeneidad individual y temporal (Pledger et al., 2003). Adicionalmente, estos autores señalan que mediante algunas

reparametrizaciones y usando una transformación *logit*, se podría simplificar el modelo para obtener efectos aditivos (sin interacciones) así como una gran cantidad de combinaciones (

Cuadro 6).

En un trabajo posterior, estos autores incorporan este mismo concepto en el modelo JS, permitiendo la mejora de los estimadores de abundancia en condiciones de heterogeneidad de probabilidades de captura (Pledger et al., 2010). Adicionalmente, se reduce el sesgo en la estimación de la sobrevivencia ocasionado por condicionar los estimadores a una captura previa. Para esto reconocen a $\theta = [N, \beta, \pi, p, \phi]^t$ como el vector paramétrico de interés, donde β_j corresponde a la proporción de individuos que entra a la población en la j -ésima ocasión de muestreo, haciéndose disponible para captura a partir de ahí. La función de verosimilitud de ese vector paramétrico corresponde a (37).

$$L(\theta) = \frac{N!}{(N-D)! \prod_h n_h!} \times \prod_{i=1}^D L_i \times L_0^{N-D} \quad (37)$$

Donde h es el subíndice que representa cada una de las historias de captura, D es el número total de individuos distintos observados, y

$$L_i = \sum_{b=1}^{f_i} \sum_{c=1}^C \sum_{d=l_i}^K \left\{ \pi_c \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{jc} \right) (1 - \phi_{dc}) \left(\prod_{j=b}^d p_{jc}^{x_{ij}} (1 - p_{jc})^{1-x_{ij}} \right) \right\} \quad (38)$$

En este caso, el subíndice b representa los posibles momentos de nacimiento o ingreso a la población y d los posibles momentos de salida.

Adicionalmente de (37), L_0 , se define como la probabilidad de que un animal nunca sea capturado:

$$L_0 = \sum_{b=1}^{f_i} \sum_{c=1}^C \sum_{d=l_i}^K \left\{ \pi_c \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{jc} \right) (1 - \phi_{dc}) \left(\prod_{j=b}^d (1 - p_{jc}) \right) \right\} \quad (39)$$

Esta propuesta metodológica incluye una serie de simplificaciones del modelo por lo que permite diferentes combinaciones de efectos aditivos y multiplicativos.

Una alternativa al uso de modelos de mezclas finitas para incluir la heterogeneidad no medida es el uso de modelos mixtos (Gimenez, Cam, y Gaillard, 2018), donde se ajusta un modelo logístico $logit(\phi_i) = \beta_0 + \epsilon_i$, donde los ϵ_i 's siguen una distribución normal e distribuyen normales con media 0 y varianza σ^2 , la cual se debe estimar posteriormente. En el anterior modelo el subíndice i identifica a cada individuo. Estas estimaciones de efectos aleatorios, obtenidas mediante métodos bayesianos o frecuentistas, se introducen para modelar la probabilidad de cada historia de captura.

Se concluye este breve repaso sobre algunas técnicas de manejo de heterogeneidad en las probabilidades de captura y sobrevivencia con los hallazgos de Crespin, Choquet, Lima, Merritt y Pradel (2008), quienes sugieren que la heterogeneidad en probabilidades de captura se deben a factores intrínsecos de la población en estudio (por ejemplo, dominancia social) así como a factores extrínsecos, generados por el diseño de muestreo (efecto de borde, dependencia de trampas, respuestas de comportamiento a la trampa, transiencia³). Proponen el uso de modelos multi-evento, donde consideran dos subpoblaciones, aquellos altamente capturables (con probabilidad de captura p_1) y aquellos con baja capturabilidad (probabilidad de captura p_2 proporcional a p_1 , o sea $p_2 = \alpha p_1$ donde α es un coeficiente de proporcionalidad). Estos estados son permanentes a través del estudio, en comparación con los efectos *trap-shy* y *trap-happy* de la modelación tradicional que son dependientes de la ocasión de muestreo. La pertenencia a cada una de las clases (alta o baja capturabilidad) no es conocida para cada individuo, pero se puede estimar una probabilidad (π) de pertenencia a alguna de estas clases. Comparando esta nueva propuesta con modelos que incluyen efectos de transiencia y heterogeneidad de probabilidades de captura por comportamiento, se encuentran mejores ajustes usando el nuevo modelo propuesto.

³ El efecto de transiencia se refiere a una serie de individuos cuya probabilidad de captura es nula por definición (Crespin et al., 2008)

Una revisión menos detallada pero más extensiva del tema se puede hallar en el trabajo de Gimenez et al (2018).

Para finalizar este capítulo, se puede aseverar que, considerando el problema que origina esta investigación, se ha hecho una revisión de una gran cantidad de metodologías que pueden servir para combinar información de captura y recaptura, recuperación de mortalidad y tamaños de muestra iniciales en situaciones de muestreo donde se violan los supuestos de homogeneidad de capturas y sobrevivencia. Sin embargo, ninguno de los métodos analizados logra, al menos teóricamente, combinar todas las propiedades deseadas. Por tanto, es necesario plantear una modificación que logre unificar parte de estas propuestas para lograr la meta planteada en un inicio, usando los hallazgos y desarrollos recientes en el tema.

4 PROPUESTA TEÓRICA

Si bien en apartado anterior se hizo mención del planteamiento matemático básico del modelo JS y algunas versiones de este, también se mencionó que necesita adaptaciones para poder aplicarse en situaciones de acuicultura. Por lo que, en este capítulo, se desarrollará un planteamiento teórico que permitirá utilizar el modelo JS en situaciones de acuicultura, donde el tamaño inicial es conocido, se tiene conteos parciales de mortalidad y hay heterogeneidad en las probabilidades de captura.

Para ello, se puede hacer uso del replanteamiento del modelo JS como un modelo espacio-estado de cadenas de Markov ocultas, tal y como lo sugieren autores como Kéry y Schaub (2012) y Royle y Dorazio (2008). Este planteamiento sugiere que hay una matriz \mathbf{Z} de dimensiones $N \times t$, donde cada elemento $z_{i,j}$ corresponde al estado de un individuo en un tiempo determinado pudiendo tomar los valores de 1 si el i -ésimo animal está vivo en la j -ésima ocasión de muestreo, o 0 en caso contrario. Cuando $j = 1$ entonces $z_{i,1} = 1$, pues todos los animales de la población entran vivos a esta. Para todo $j > 1$:

$$z_{i,j} \sim \text{Bernoulli}(\phi_j z_{i,j-1}) \quad (40)$$

En otras palabras, el estado actual depende del estado anterior multiplicado por la probabilidad de sobrevivir (ϕ_j) entre las ocasiones de muestreo j y $j + 1$. Estos estados no son directamente observables y, de hecho, son desconocidos para los individuos analizados, pero se logran inferir a partir de la relación entre \mathbf{Z} y una matriz de observaciones \mathbf{Y} que contiene una línea para cada individuo observado al menos una vez. En esta matriz con $2t$ columnas, la mitad de estas contiene la información de la historia de captura, mientras que la otra mitad, señala la información de recuperación de marcas. En ambos casos, un 1 indica una observación, un 0 que el individuo no fue capturado ni recuperado en ese momento. Para el caso de este trabajo se usó un formato donde las primeras t columnas aportan la información de capturas y las columnas $t + 1, \dots, 2t$ contienen la información de recuperación de marcas (formato LL...DD..., donde las L indican

observaciones de animales vivos y las D observaciones de animales muertos). No obstante, es posible modificarlo para usarlo con codificación LDLD tal y como hacen otros programas más utilizados, por ejemplo Mark (Cooch y White, 2019), en los modelos que combinan información de capturas y recuperación de marcas.

Más allá de la codificación, las recuperaciones seguirán una distribución de Bernoulli, con parámetro dependiente de los estados actual y previo del individuo, así como de una tasa de recuperación de los cadáveres r_j , dependiente del tiempo.

En otras palabras:

$$r_{i,j} \sim \text{Bernoulli}(r_j z_{i-1,j} (1 - z_{ij})) \quad (41)$$

Las capturas seguirán una distribución Bernoulli, cuyo parámetro depende del estado del individuo en la ocasión de muestreo y la probabilidad de captura, la cual varía de acuerdo con el individuo y la ocasión de muestreo:

$$y_{i,j} \sim \text{Bernoulli}(p_{ij} z_{i-1,j}) \quad (42)$$

Anteriormente se mencionó que las probabilidades de captura podrían verse afectadas por el comportamiento de dominancia de las especies del género *Oreochromis*, asimismo, la probabilidad de captura podría variar de ocasión a ocasión, en especial si el tamaño de muestra es fijo. Por ello, es prudente considerar la heterogeneidad individual y temporal en la capturabilidad. Para esto se considera el siguiente modelo logístico de efectos mixtos:

$$\log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \mu_t + \epsilon_i \quad (43)$$

$$\epsilon_i \sim N(0, \sigma^2)$$

Donde μ_t corresponde al efecto de la ocasión de muestreo sobre la probabilidad de captura, mientras que ϵ_i es un efecto aleatorio individual, distribuido normalmente con media cero y varianza σ^2 .

Considerando lo anterior, podrían plantearse dos versiones del modelo. La primera versión sería un modelo JS completo que permita estimar de manera directa todas las cantidades de interés: N_j, ϕ_j, p_{ij}, r_j . Este tipo de modelo está

basado en la reformulación del modelo de JS a partir de modelos de ocupación restringida, tal y como lo plantean Kéry & Schaub (2012).

En el caso particular en estudio, se puede utilizar la estimación bayesiana auxiliada por la técnica de aumento de datos, para obtener una cantidad M de historias de captura, de las cuales $M - n$ son historias de captura en blanco de posibles individuos que puedan ser parte de la población.

Es necesario estimar un parámetro γ que es la probabilidad de entrada a la población en la primera ocasión de muestreo, con lo cual la estimación de N_1 se convierte en la realización de un experimento binomial con parámetros γ y M . Dado que se conoce el tamaño de la población inicial y un indicador de precisión de este, se puede ajustar M en conformidad.

El Recuadro 1 muestra el código para plantear este modelo en lenguaje BUGS. Se asumió una sola sobrevivencia y recuperabilidad por momento de captura t para todos los individuos en el estanque. A ambas se les asignó una distribución previa $Uniforme(0,1)$. Asimismo, para el parámetro γ , previamente explicado, se asume una distribución previa $Uniforme(0,1)$ indicando nula información acerca del mismo.

Para las probabilidades de captura de cada individuo, se usó un modelo logístico tal como se indicó en (43). Para el logaritmo natural de la probabilidad media de captura para un momento de captura t , denotado como μ_t en (43) o $logp[t]$ en el Recuadro 1 se utilizó una distribución previa t generalizada, con parámetro de localización 0, escala 2,5 y 7 grados de libertad. Dicha distribución es recomendada como una distribución mínimamente informativa y que permite regularizar los coeficientes de una regresión logística (Gelman, Jakulin, Pittau, y Su, 2008). Por su parte, para las contribuciones individuales (ϵ_i) se utilizó una distribución normal con desviación estándar distribuida según una exponencial con parámetro de tasa en 0,14, que se puede considerar mínimamente informativa y penalizadora de la complejidad del modelo (Simpson, Rue, Riebler, Martins, y Sørbye, 2017).


```

model {
  # Distribuciones previas y restricciones
  for (i in 1:M){
    for (t in 1:(n.occasions-1)){
      phi[i,t] <- mean.phi[t]
      r[i,t] <- mean.r[t]
    } #t
    for (t in 1:n.occasions){
      logit(p[i,t]) <- logp[t] + epsilon[i]
    } #t
    epsilon[i] ~ dnorm(0,tau)
  } #i
  for (t in 1:(n.occasions-1)){
    mean.phi[t] ~ dunif(0, 1)
    mean.r[t] ~ dunif(0, 1)
  }
  for (t in 1:n.occasions){
    logp[t] ~ dt(0, pow(2.5,-2),7)
    mean.p[t] <- exp(logp[t])/(1+exp(logp[t]))
  }
  gamma ~ dunif(0, 1)
  sigma ~ dexp(0.14)
  tau <- pow(sigma,-2)

  # Función de verosimilitud
  for (i in 1:M){
    # Primera ocasión de muestreo
    # Matriz de estado
    z[i,1] ~ dbern(gamma)
    # Observaciones
    mul[i] <- z[i,1] * p[i,1]
    y[i,1] ~ dbern(mul[i])
    # Ocasiones de muestreo subsiguientes
    for (t in 2:n.occasions){
      # Matriz de estado
      mu2[i,t] <- phi[i,t-1] * z[i,t-1]
      z[i,t] ~ dbern(mu2[i,t])
      # Matriz de capturas
      mu3[i,t] <- z[i,t] * p[i,t]
      y[i,t] ~ dbern(mu3[i,t])
      # Matriz de recuperaciones
      mu4[i,t] <- z[i,t-1] * (1-z[i,t]) * r[i,t-1]
      d[i,t] ~ dbern(mu4[i,t])
    } #t
  } #i
  # Cálculos derivados
  for (t in 1:n.occasions){
    N[t] <- sum(z[1:M,t]) # Tamaño poblacional
  } #t
  for (i in 1:M){
    Nind[i] <- sum(z[i,1:n.occasions])
    Nalive[i] <- 1-equals(Nind[i], 0)
  } #i
}

```

Recuadro 1. Código BUGS para el modelo JS completo.

Dicho planteamiento tiene la ventaja que permitiría ajustar el número de peces sembrado, en caso de reportes erróneos, intencionales o accidentales. Sin embargo, en las pruebas realizadas se encontraron problemas de convergencia, que requieren una gran cantidad de iteraciones y tiempo computacional para obtener resultados consistentes.

La segunda propuesta de modelo, no estima directamente la abundancia, sino que considera el tamaño inicial de población reportado como verdadero. En este caso, se puede estimar N_j usando la ecuación $N_j = \phi_{j-1}N_{j-1}$ para todo $j > 1$. Por su parte, ϕ_j se puede obtener a partir de una idea similar al modelo de Burnham, que combina información de capturas y recuperaciones, pero parametrizado de tal forma que se consideren las recuperaciones como una fracción de la sobrevivencia aparente. El Recuadro 2 muestra el planteamiento de este modelo que alcanza más fácilmente la convergencia, pero tiene el inconveniente de que no ajusta el tamaño de la población inicial, que se considera verdadero. Se hace uso de las mismas distribuciones previas indicadas anteriormente.

Este capítulo presentó dos versiones del modelo JS adaptado para aplicar en situaciones de acuicultura comercial, pues permite incluir información del número de peces sembrado y conteos parciales de mortalidad, de forma que se pueda estimar los parámetros deseados (abundancia, sobrevivencia) más un parámetro instrumental de recuperabilidad, que refleja la proporción de individuos muertos cuyos cadáveres han sido contabilizados. El segundo modelo será el usado en el trabajo para obtener las estimaciones.

```

model {
  # previas y restricciones
  for (i in 1:M){
    for (t in 1:(n.occasions-1)){
      phi[i,t] <- mean.phi[t]
      r[i,t] <- mean.r[t]
    } #t
    for (t in 1:n.occasions){
      logit(p[i,t]) <- logp[t] + epsilon[i]
    } #t
    epsilon[i] ~ dnorm(0,tau)
  } #i
  for (t in 1:(n.occasions-1)){
    logphi[t] ~ dt(0,pow(2.5,-2),1)
    logr[t] ~ dt(0,pow(2.5,-2),1)
    mean.phi[t] <- exp(logphi[t])/(1+exp(logphi[t]))
    mean.r[t] <- exp(logr[t])/(1+exp(logr[t]))
  }
  for (t in 1:n.occasions){
    logp[t] ~ dt(0,pow(2.5,-2),1)
    mean.p[t] <- exp(logp[t])/(1+exp(logp[t]))
  }
  sigma ~ dexp(0.14)
  tau <- pow(sigma,-2)
  # verosimilitud
  for (i in 1:M){
    # Primera ocasión
    # Proceso de estado
    z[i,1] <- 1
    # Proceso de observación
    mul[i] <- z[i,1] * p[i,1]
    y[i,1] ~ dbern(mul[i])
    # Ocasiones subsiguientes
    for (t in 2:n.occasions){
      # Proceso de estado
      mu2[i,t] <- phi[i,t-1] * z[i,t-1]
      z[i,t] ~ dbern(mu2[i,t])
      # Proceso de observación
      mu3[i,t] <- z[i,t] * p[i,t]
      y[i,t] ~ dbern(mu3[i,t])
    }
    # recuperación
    mu4[i,t] <- z[i,t-1] * (1-z[i,t]) * r[i,t-1]
    d[i,t] ~ dbern(mu4[i,t])
  } #t
  } #i
  # Cálculos derivados
  for (t in 3:n.occasions){
    N[t]<-N[t-1]*mean.phi[t-1]
  }#t
  N[2]<-N0*mean.phi[1]
}

```

Recuadro 2. Código BUGS para modelo JS condicional a la primera captura, con información adicional de capturas.

5 ANÁLISIS DE LA MORTALIDAD OBSERVABLE

Cuando se desea hacer simulaciones, es importante que la toma de decisiones acerca de la distribución de los datos produzca resultados generalizables a situaciones reales, para ello, una posible estrategia es partir de un juego de datos conocido, que sirva como ejemplo motivador (Burton, Altman, Royston y Holder, 2006). Este capítulo sigue esa premisa, pues se analiza un juego de datos de sobrevivencia de una empresa acuícola industrial de Costa Rica y se usa como machote para la simulación. El capítulo se divide en tres secciones: una introducción acerca del uso de métodos de análisis de sobrevivencia en la acuicultura, la presentación de la metodología y el análisis de los resultados.

5.1 Introducción

A pesar de que la sobrevivencia es una característica importante de considerar en la producción acuícola, generalmente sus mediciones y los modelos para su estimación corresponden con la mortalidad total del ciclo productivo. Por ejemplo, Lorenzen (1996) desarrolla un modelo para pronosticar la mortalidad en función del peso inicial del estanque. Por su parte, Zuniga-Jara y Goycolea-Homann (2014), desarrollan un modelo bioeconómico para la producción de tilapia roja en Ecuador, en el que las tasas de mortalidad son fijas para diferentes periodos de la producción. Específicamente, usan tasas de mortalidad diarias de 0,009673; 0,006957 y 0,001470 para los periodos de 0 a 76, 77 a 109 y 110 días de edad del estanque en adelante, respectivamente.

La utilización de modelos de sobrevivencia propiamente dichos es escasa en la literatura encontrada. Si bien el método de Kaplan-Meier ha sido utilizado por diversos autores (Wonmongkol et al., 2018; Foyosal et al., 2019; Sukhavachana, Poompuang, Onming y Luengnaruemitchai, 2019; Yilmaz, 2019 y Sukhavachana et al., 2020), son pocos los que usan modelos paramétricos o semiparamétricos de sobrevivencia (Wilson, Nibbelink y Peterson, 2009 y Gosselin y Anderson, 2020).

En el presente estudio, es de crucial importancia que la simulación de la sobrevivencia sea lo más apegada a la realidad. Burton et al. (2006) sugieren que para la simulación de tiempos a la ocurrencia de un evento, se debe partir de un juego de datos similar, para obtener tanto la distribución de los datos de sobrevivencia y de su relación con las covariables. En aras de cumplir este objetivo, y dado que se contaba con los datos, se procedió a analizar la situación de la mortalidad contada u observable en el sistema productivo en cuestión, con el fin de obtener parámetros de simulación adecuados y apegados a la realidad.

5.2 Ajuste de curvas de sobrevivencia

Para el ajuste de las curvas de sobrevivencia, se partió de los datos de 189 ciclos de producción de tilapia, llevados a cabo en dos fincas ubicadas en Cañas, Guanacaste, Costa Rica, durante el 2012 y el 2013, ambas pertenecientes a una compañía productora de tilapia para exportación. En estos estanques se registró el destino de aproximadamente 9 millones de tilapias, cuyas observaciones de tiempo de vida fueron censuradas en el momento de la cosecha, para aquellos casos en que no fueron anotadas como muertas en algún momento del ciclo productivo.

La Figura 1 muestra la gran diversidad de curvas de sobrevivencia para estos estanques. Se observa que hay diferentes tipologías de curvas de mortalidad; mientras unas muestran una alta mortalidad en los primeros días del ciclo, otros más bien muestran altas mortalidades después del día 60 del ciclo. Algunas otras ni siquiera muestran mortalidad apreciable hasta muy cerca de finalizar el ciclo (~120 días). En muchos casos se observan patrones similares a escalones, en contraste con las curvas suavizadas esperadas, lo cual coincide con Gosselin y Anderson (2020), quienes apuntan que esto es efecto de estresores ambientales y la presencia de heterogeneidad en la población.

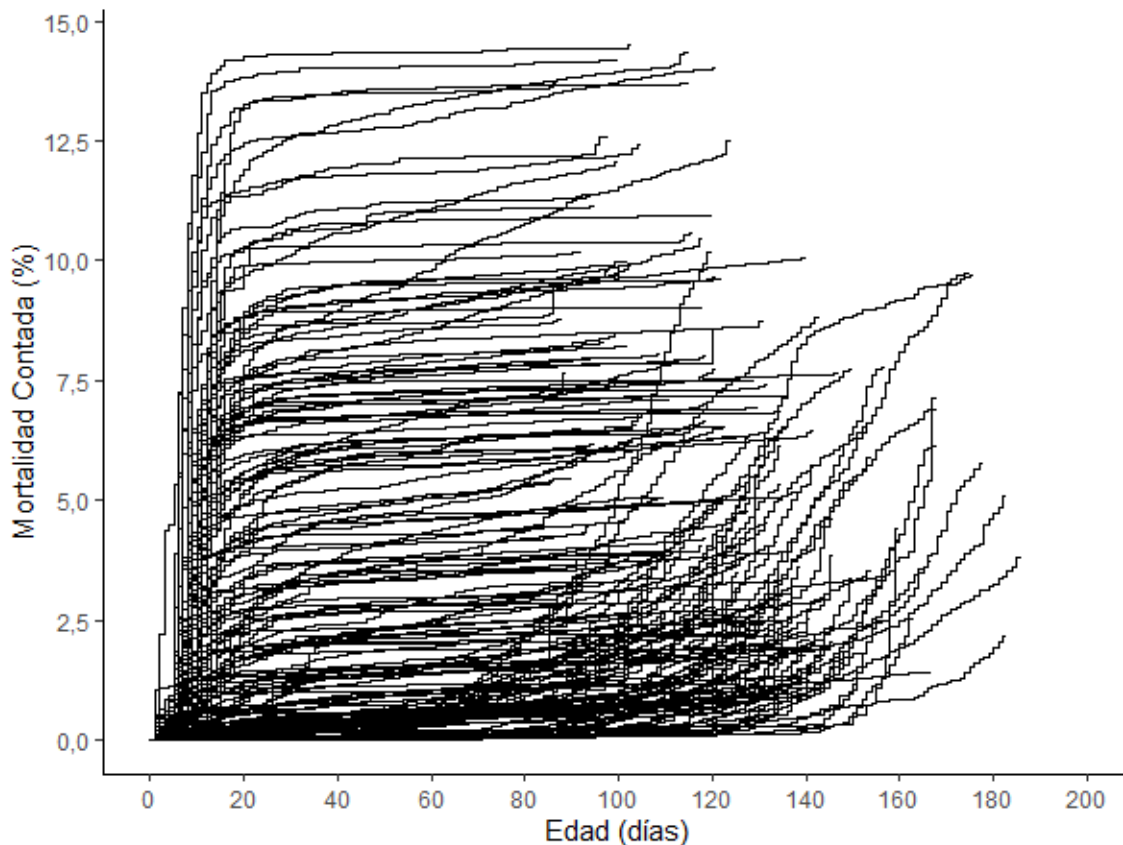


Figura 1. Curvas de mortalidad, estimadas por el método de Kaplan-Meier para 189 estanques de producción semiintensiva de tilapia.

Gran parte de la heterogeneidad observada en la forma de las curvas de mortalidad puede explicarse por la época térmica en la que se realizó la siembra del estanque. Arce-Vega (2014) determinó la influencia de la época térmica en el crecimiento de las tilapias, en el mismo sistema productivo analizado. Sus resultados mostraron que durante los meses de abril a octubre se presenta una época cálida, con temperaturas de 27°C en el agua y tasas de crecimiento altas. Por su parte, los meses restantes se clasifican en la época fría, con temperaturas de 25°C aproximadamente y menores tasas de crecimiento. La Figura 2 muestra que dicha clasificación permite desagregar la gran divergencia observada en el comportamiento de la mortalidad en los datos analizados. Adicional a las dos épocas térmicas, se decidió separar los meses de octubre y abril, que podrían clasificarse como meses de "transición" y en los cuales las curvas de mortalidad muestran características mixtas.

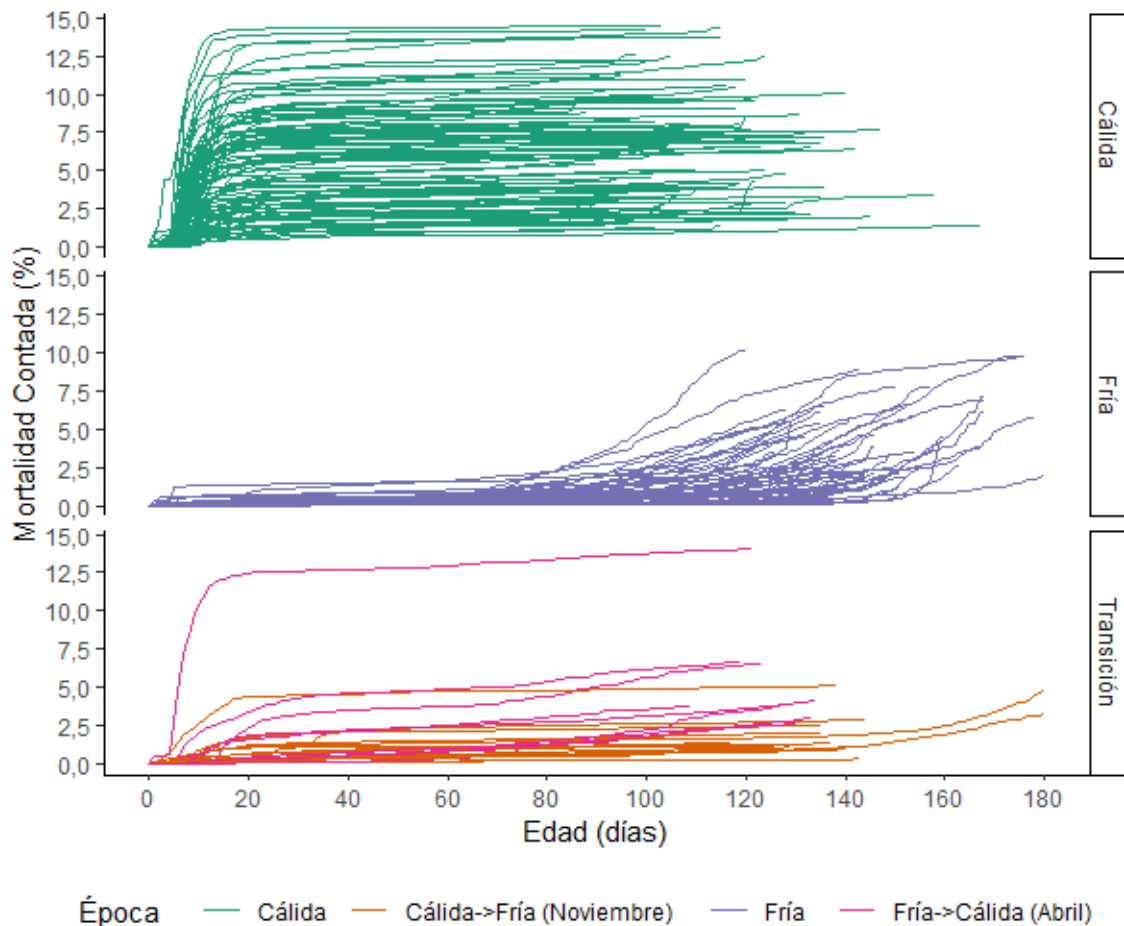


Figura 2. Curvas de mortalidad para 189 estanques, clasificadas por épocas térmicas, de acuerdo con la clasificación de Arce-Vega (2014)

Para poder realizar las simulaciones de forma que se conserve la enorme variabilidad observada, se ajustó un modelo de sobrevivencia no paramétrico usando el estimador de Kaplan-Meier para cada estanque. Los resultados de mortalidad ($1 - \text{sobrevivencia}$) estimados para cada día y estanque se utilizaron posteriormente para ajustar dos modelos no lineales mixtos.

El primer modelo, ajustado usando únicamente las curvas de mortalidad de los estanques sembrados durante la época cálida, se realizó usando el modelo no lineal de Richards:

$$y_{jk} = \frac{A_j}{\left(1 + m * \exp\left(-r_j(t_{jk} - i_j)\right)\right)^{\frac{1}{m}}} + \epsilon_{jk} \quad (44)$$

Donde

- y_{jk} corresponde a la observación k -ésima, realizada en el estanque j -ésimo
- t_{jk} corresponde al tiempo en el que se tomó la observación k -ésima, en el estanque j -ésimo
- A_j es el valor asintótico al que tiende la mortalidad cuando $t \rightarrow +\infty$ en el estanque j -ésimo. La estimación de este parámetro incluye efectos fijos de la densidad de biomasa (D) y la finca (F), así como el efecto aleatorio del estanque sobre el intercepto (b_{Aj}), según la ecuación $A_j = \mu_A + b_{Aj} + \beta_1 D_j + \beta_2 F_j$.
- m es el parámetro de forma de la función no lineal de Richards.
- r_j es el parámetro de tasa de la función de Richards para el estanque j -ésimo. La estimación de este parámetro incluye efectos fijos de la densidad de biomasa (D) y la finca (F), así como el efecto aleatorio del estanque sobre el intercepto (b_{rj}), según la ecuación $r_j = \mu_r + b_{rj} + \beta_1 D_j + \beta_2 F_j$.
- i_j es el punto de inflexión de la curva de mortalidad del estanque j -ésimo. La estimación de este parámetro incluye efectos fijos de la densidad de biomasa (D) y la finca (F), así como el efecto aleatorio del estanque sobre el intercepto (b_{ij}), según la ecuación $i_j = \mu_i + b_{ij} + \beta_1 D_j + \beta_2 F_j$.

Es importante indicar que el vector de efectos aleatorios $[b_{Aj}, b_{rj}, b_{ij}]' \sim MVN(0, \Sigma)$, donde Σ es una matriz triangular con descomposición log-Cholesky. Finalmente, los residuales para la observación ϵ_{jk} k en el tiempo j se asumen independientes e idénticamente distribuidos, con distribución normal y varianza constante.

Para las curvas de mortalidad de los estanques sembrados durante la época fría, se usó un modelo logístico doble:

$$y_{jk} = \frac{A}{1 + \exp(-r_j(t_{jk} - i_j))} + \frac{A'_j}{1 + \exp(-r'_j(t_{jk} - i'))} + \epsilon_{jk} \quad (45)$$

Donde

- y_{jk} corresponde a la observación k -ésima, realizada en el estanque j -ésimo
- t_{jk} corresponde al tiempo en el que se tomó la observación k -ésima, en el estanque j -ésimo
- A y A'_j corresponden a los valores asintóticos de las respectivas curvas logísticas. En el caso de la primera logística, es un valor fijo para la población, mientras que la segunda incluye un efecto aleatorio para cada estanque ($b_{A'j}$), según la ecuación $A'_j = \mu_{A'} + b_{A'j}$.
- r_j y r'_j son las tasas de aumento de la mortalidad para ambas ecuaciones logísticas en el estanque j -ésimo. La estimación de estos parámetros incluye el efecto aleatorio del estanque sobre el intercepto ($b_{rj}, b_{r'j}$), según las ecuaciones $r_j = \mu_r + b_{rj}$ y $r'_j = \mu_{r'} + b_{r'j}$.
- i_j e i' corresponden a los puntos de inflexión de las curvas logísticas del estanque j -ésimo. i' es un valor fijo para la población, pero i_j contiene el efecto aleatorio del estanque (b_{ij}), según la ecuación $i_j = \mu_i + b_{ij}$.

Es importante indicar que el vector de efectos aleatorios $[b_{A'j}, b_{rj}, b_{r'j}, b_{ij}]' \sim MVN(0, \Sigma)$, donde $\Sigma = [\sigma_{A'j}, \sigma_{rj}, \sigma_{r'j}, \sigma_{ij}]' I_{[4 \times 4]}$. En otras palabras, se asume que los efectos aleatorios para los diferentes parámetros no están correlacionados entre sí. Finalmente, los residuales (ϵ_{jk}) se asumen independientes, con distribución normal y una función de varianza que aumenta exponencialmente con el logaritmo natural del tiempo.

En el caso de los estanques de época de transición, no se realizó un modelo específico para estos, ya que presentan una mezcla de curvas que serían caracterizadas por los dos modelos anteriores. El ajuste de los modelos se realizó usando el paquete nlme (Pinheiro, Bates, DebRoy, Sarkar y R Core Team, 2020), complementario de R 4.0.4 (R Core Team, 2020).

Adicional al de la mortalidad, se realizó el modelaje de la biomasa sembrada y de la duración del estanque. La primera porque que influye directamente en los parámetros de las curvas de mortalidad, mientras que la segunda permite ver la longitud de la curva de mortalidad.

Para la biomasa sembrada se utilizó un modelo lineal con distribución normal, donde la media de la densidad se estimó a partir del mes de siembra y la finca. Adicionalmente, la varianza de los errores se asumió dependiente de la finca. Para el ajuste del modelo mencionado se usó el paquete glmmTMB (Brooks et al., 2017) complementario de R 4.0.4 (R Core Team, 2020). En el caso de la edad del estanque, se utilizó un modelo de regresión lineal donde los predictores eran la finca y un polinomio grado cuatro del número de mes (desde enero = 1 hasta diciembre = 12), ajustado con las funciones base de R.

5.3 Resultados de los modelos ajustados

El Cuadro 7 muestra los resultados del modelo para la mortalidad en la época seca. Cabe resaltar que las covariables Densidad de Biomasa y Finca afectaron de manera significativa ($p < 0,05$) los valores de cada uno de los parámetros del modelo no lineal de Richards, excepto el caso del punto de inflexión. Con respecto al parámetro A, que corresponde a la mortalidad potencial, se puede ver que la finca S tiene un valor 2% menor que la finca L, que es la situación de referencia. Asimismo, en otros parámetros como la tasa de mortalidad (r) y el parámetro de forma (m) se registran valores menores en la finca S, lo cual puede deberse a que las aguas que llegan a esta finca han sido menos reutilizadas que las que llegan a la finca L.

Todos los parámetros, por su parte, se ven aumentados ante incrementos en la densidad de masa sembrada, expresada como kg/m^2 , lo que es de esperarse, pues el mayor hacinamiento hace más fácil la transmisión de enfermedades y requiere mayores niveles de oxígeno y una mejor calidad de agua para poder ser sostenidas.

Cuadro 7. Coeficientes, errores estándar y valores p para el modelo no lineal de Richards para la mortalidad de los estanques sembrados en época térmica cálida.

Parámetro ¹	Predictor	Coefficiente	EE	Valor p
A	Intercepto	0,005	0,015	0,7479
	Densidad BM	0,065	0,016	<0,0001
	Finca S	-0,019	0,008	0,0141
r	Intercepto	0,069	0,046	0,1338
	Densidad BM	0,265	0,050	<0,0001
	Finca S	-0,204	0,023	<0,0001
i	Intercepto	9,373	1,325	<0,0001
	Densidad BM	-1,220	1,449	0,3996
	Finca S	0,018	0,667	0,9783
m	Intercepto	-0,448	0,006	<0,0001
	Densidad BM	0,417	0,007	<0,0001
	Finca S	-0,354	0,004	<0,0001

¹Parámetros de la ecuación $y = A / (1 + m * \exp(-r(t - i)))^{\frac{1}{m}}$

En el Cuadro 8 se puede observar que los efectos aleatorios de algunos de estos parámetros mostraron correlaciones importantes y que, en especial para el punto de inflexión (i) existe una importante heterogeneidad no modelada, lo cual se refleja en una alta desviación estándar de los efectos aleatorios.

Cuadro 8. Valores de desviación estándar de los efectos aleatorios para cada uno de los parámetros del modelo no lineal de Richards para la mortalidad en los estanques sembrados durante la época cálida y sus correlaciones.

Parámetro	Desviación Estándar Efectos Aleatorios	Correlación	
		A	r
A	0,031	-	-
r	0,094	0,509	-
i	2,584	-0,285	-0,657
residual	0,003	-	-

Por su parte, para el caso de los estanques sembrados durante la época fría, la inclusión de las covariables no se realizó, ya que afectaba la convergencia de los modelos, y cuando el modelo convergía, los resultados eran poco plausibles, por tanto, en este modelo toda la heterogeneidad observada en los parámetros se consideró latente. El Cuadro 9 muestra los coeficientes obtenidos, tanto para los efectos fijos como los aleatorios.

Cuadro 9. Coeficientes, errores estándar, valores p y desviación estándar de los efectos aleatorios para los parámetros del modelo logístico doble para la mortalidad de los estanques sembrados durante la época térmica fría.

Parámetro ¹	Coeficiente	EE	Valor p	DEEA ^{2,3}
A	0,096	0,0003	<0,0001	-
i	153,685	3,0218	<0,0001	23,519
r	0,073	0,0061	<0,0001	0,047
A'	0,002	0,0004	<0,0001	0,003
i'	8,641	0,1833	<0,0001	0,073
r'	0,111	0,0128	<0,0001	-

¹Parámetros de la ecuación $y = A/(1 + \exp(-r(t - i))) + A'/(1 + \exp(-r'(t - i')))$

²Desviación estándar de los efectos aleatorios

³Desviación estándar residual = 0,0005

En el Cuadro 10 se muestran los resultados del modelo para estimar la densidad de biomasa. Se observa una tendencia al aumento en la densidad de biomasa sembrada hacia los meses de setiembre y noviembre, mientras que los meses de enero y marzo son los que presentan promedios menores de densidad de biomasa sembrada.

Cuadro 10. Coeficientes del modelo para la predicción de la densidad de biomasa sembrada por estanque (kg/m²)

Parámetro	Predictor	Coeficiente	E.E.	Valor p
Media	Intercepto	0,51	0,040	<0,0001
	Mes febrero	0,13	0,054	0,0166
	Mes marzo	0,04	0,056	0,4512
	Mes abril	0,17	0,072	0,0204
	Mes mayo	0,40	0,068	<0,0001
	Mes junio	0,31	0,055	<0,0001
	Mes julio	0,27	0,051	<0,0001
	Mes agosto	0,35	0,059	<0,0001
	Mes setiembre	0,56	0,055	<0,0001
	Mes octubre	0,45	0,052	<0,0001
	Mes noviembre	0,30	0,059	<0,0001
	Mes diciembre	0,12	0,066	0,0594
	Finca S	0,23	0,025	<0,0001
Varianza	Intercepto	-3,91	0,160	<0,0001
	Finca S	0,51	0,216	0,0180

El Cuadro 10 también permite comparar la situación de la densidad de biomasa inicial entre la Finca S y la Finca L, que es, nuevamente, la situación de referencia para la interpretación de los coeficientes. En cuanto al parámetro de la media de la biomasa, la Finca S presentó una densidad 0,23 kg/m² mayor que en la Finca

L, lo cual se debe a que tiene mejor calidad de agua, por lo cual se puede sembrar una mayor cantidad de peces por unidad de área. En cuanto a la varianza del modelo, en la Finca S es 66% ($e^{0,51}$) superior con respecto a la finca L, que indica menor consistencia en la siembra.

Finalmente, en el Cuadro 11 se pueden observar los resultados del modelo para la edad o duración del estanque. En este modelo se aprecia que los estanques sembrados hacia mitad de año (época cálida) tienden a durar significativamente menos que los estanques de inicio y principio de año (época fría), lo cual se debe a que las tasas de crecimiento son mayores en la época cálida (Arce-Vega, 2014) y a que, ante la menor calidad de agua y mayor mortalidad, es necesario tomar decisiones de cosecha de manera más temprana.

Cuadro 11. Coeficientes, error estándar y valores p para el modelo de predicción de la edad del estanque.

Parámetro	Predictor	Coeficiente	EE	Valor p
Media	Intercepto	163,72	0,815	<0,0001
	Mes ¹	-5,86	0,806	<0,0001
	Mes ²	-3,70	0,238	<0,0001
	Mes ³	0,64	0,027	<0,0001
	Mes ⁴	-0,03	0,001	<0,0001
	Finca	6,48	0,201	<0,0001
Desviación Estándar Residual		14,81		

^{1,2,3,4} Coeficientes polinómicos para la variable Mes, que se ingresa al modelo como variable cuantitativa usando los valores usuales de 1 = enero, 2 = febrero, ..., 12 = diciembre

5.4 Prueba de simulación

Tras ajustar los modelos anteriores, es necesario verificar si, al incluirlos en un proceso de simulación, se consiguen resultados similares a los observados en los datos recolectados. Para lograr una simulación de los resultados se empleó el siguiente procedimiento, cuya esquematización se puede observar en la Figura 3:

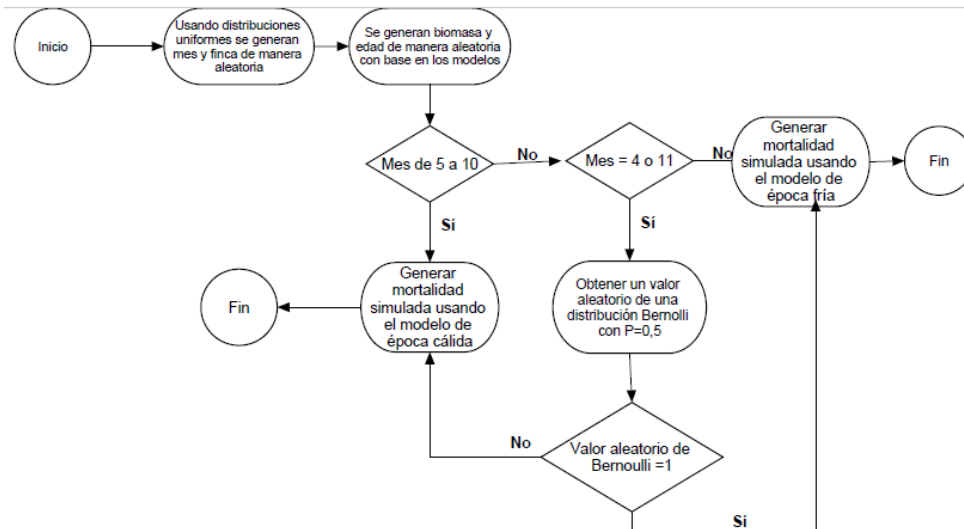


Figura 3. Diagrama de flujo del proceso de simulación.

1. Se toma un valor aleatorio de una distribución discreta uniforme entre 1 y 12, que corresponde al mes en que se inició el ciclo.
2. Se toma un valor aleatorio de una distribución uniforme discreta entre 0 y 1, si sale 1, se asigna a la finca "S", caso contrario a la finca "L".
3. Con los valores anteriores se obtiene el valor esperado de densidad de biomasa sembrada y edad, usando el modelo correspondiente. El valor predicho por ese modelo, en conjunto con la desviación estándar residual de cada modelo, se usan para obtener un valor aleatorio de una distribución normal que será usado como el valor observado de biomasa y edad.
4. Los valores de biomasa, finca y tiempo se incluyen en los modelos no lineales que correspondan, dependiendo del mes de siembra. Si el mes de siembra toma valores entre 5 y 10 se usa el modelo para época caliente. Si el mes de siembra es 12, 1, 2 o 3, se usa el modelo de época fría. Finalmente, si el mes es 4 o 11, se genera un número aleatorio de una distribución de Bernoulli, con $p = 0,5$. Si el valor es igual a 1, se usa la ecuación de época fría, en otro caso, la de época caliente.

Para verificar el resultado del proceso de simulación, se generaron 200 simulaciones de mortalidad. Puede notarse que la forma de las simulaciones coincide, en general, con los datos reales, se considera que las simulaciones aproximan mucho la realidad. La Figura 4 muestra las 200 curvas de mortalidad simuladas, muestra una importante similitud con las curvas reales observadas en la Figura 1, lo cual indica que el proceso de modelaje y simulación ha sido apropiado, por lo menos a nivel visual.

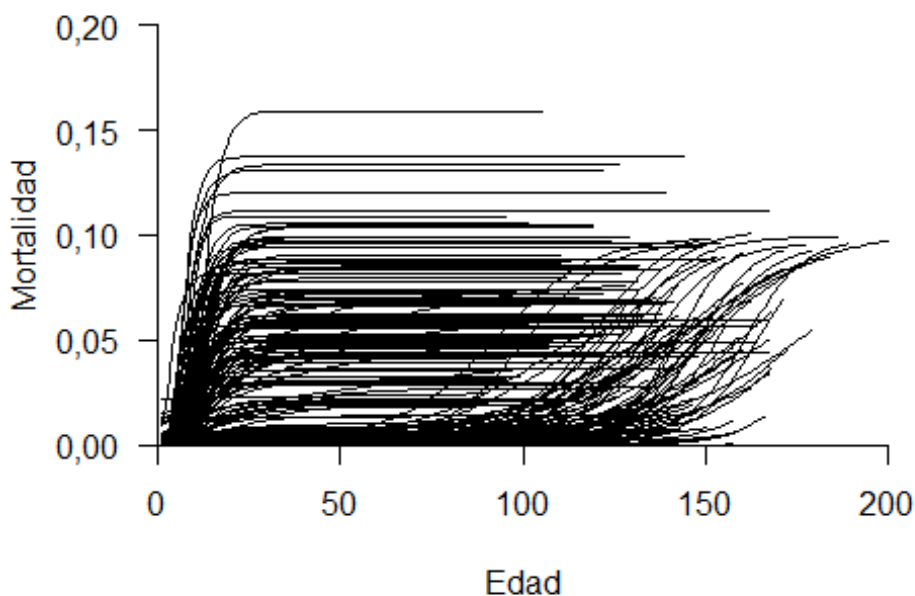


Figura 4. Curvas de mortalidad simuladas para 200 estanques semiintensivos de producción de tilapia.

A modo de cierre de este capítulo, se puede decir que se mostró un análisis de la mortalidad observada en una empresa acuícola costarricense en el que se descubrió un patrón de diferencias en la forma de las curvas de mortalidad que coincide con estudios anteriores que muestran diferencias por época del año en las tasas de crecimiento. Asimismo, se generan modelos estadísticos que van a servir, en los siguientes capítulos, para realizar las simulaciones de mortalidad.

6 METODOLOGÍA DE SIMULACIÓN

Este capítulo está dividido en cuatro secciones, cada una enfocada en aspectos necesarios para comprender cómo se realizaron las simulaciones y el análisis de sus resultados. Para garantizar que los parámetros asumidos son válidos se necesita que los parámetros de simulación sean similares a los que se encontrarían en explotaciones reales, por eso, la sección primera aborda el análisis descriptivo de los datos de una empresa productora de tilapia que sirven como fundamento para las simulaciones. Asimismo, la sección segunda, basada en datos reales y artículos de investigación, plantea el mecanismo para simular las poblaciones de peces. La tercera sección versa sobre los métodos que se usaron, para estimar, en cada caso, los parámetros de interés, mediante metodologías de captura-recaptura. Finalmente, la cuarta sección aborda el tema de cómo se comparan la precisión y el sesgo de los resultados.

6.1 Fuente de los datos

Los datos para realizar la simulación se obtuvieron a partir de los resultados de la campaña 2012-2013 de engorde de una empresa de producción de tilapia, radicada en Costa Rica. Durante esta campaña, se reportaron 186 ciclos de producción en 72 estanques de engorde semiintensivos de tilapia, ubicados en dos fincas propiedad de la compañía (Fincas denominadas para la investigación como S y L).

Algunos estadísticos descriptivos de dichos ciclos de producción se encuentran en el Cuadro 12. Adicionalmente, en la **¡Error! No se encuentra el origen de la referencia.** se muestra que la sobrevivencia estimada únicamente a partir de los conteos de mortalidad sobreestima entre 0 y 60% la sobrevivencia final del estanque y que, el modelo usado para estimar la sobrevivencia final en aquel momento lograba evitar dicho sesgo, pero su exactitud era baja.

Cuadro 12. Datos de siembra, cosecha y duración del ciclo en 186 ciclos de producción de tilapia en estanques semiintensivos, llevados a cabo en una empresa ubicada en Costa Rica, durante la campaña 2012-2013.

Variable	Mínimo	Cuartil 1	Mediana	Promedio	Cuartil 3	Máximo
Duración del ciclo (d)	84,00	103,25	123,00	123,82	139,75	186,00
N_0 - Peces sembrados	22499	42856	49530,5	48418,1	56965	76891
Área del estanque (Ha)	0,59	1,29	1,53	1,62	1,88	3,34
Densidad de biomasa sembrada (kg/m ²)	0,19	0,64	0,93	0,98	1,25	3,26
N_f - Peces cosechados	12284	33970	41326	40252	47085	63462
ϕ - Supervivencia aparente ¹	0,53	0,78	0,83	0,83	0,87	0,99
M_c - Mortalidad observada ²	137	1139	1981	2551	3708	9470
ϕ_c - Supervivencia observada ³	0,85	0,93	0,96	0,95	0,97	1,00
M_e - Mortalidad estimada ⁴	2374	5328	6908	7517	9374	17288
ϕ_e - Supervivencia estimada ⁵	0,72	0,81	0,86	0,85	0,88	0,95

$$^1\phi = N_f/N_0$$

²Conteo de peces muertos recuperados por estanque

$$^3\phi_o = (N_0 - M_c)/N_0$$

⁴A partir de la mortalidad observada, modelos matemáticos y ajustes por experiencia

$$^5\phi_e = (N_0 - M_e)/N_0$$

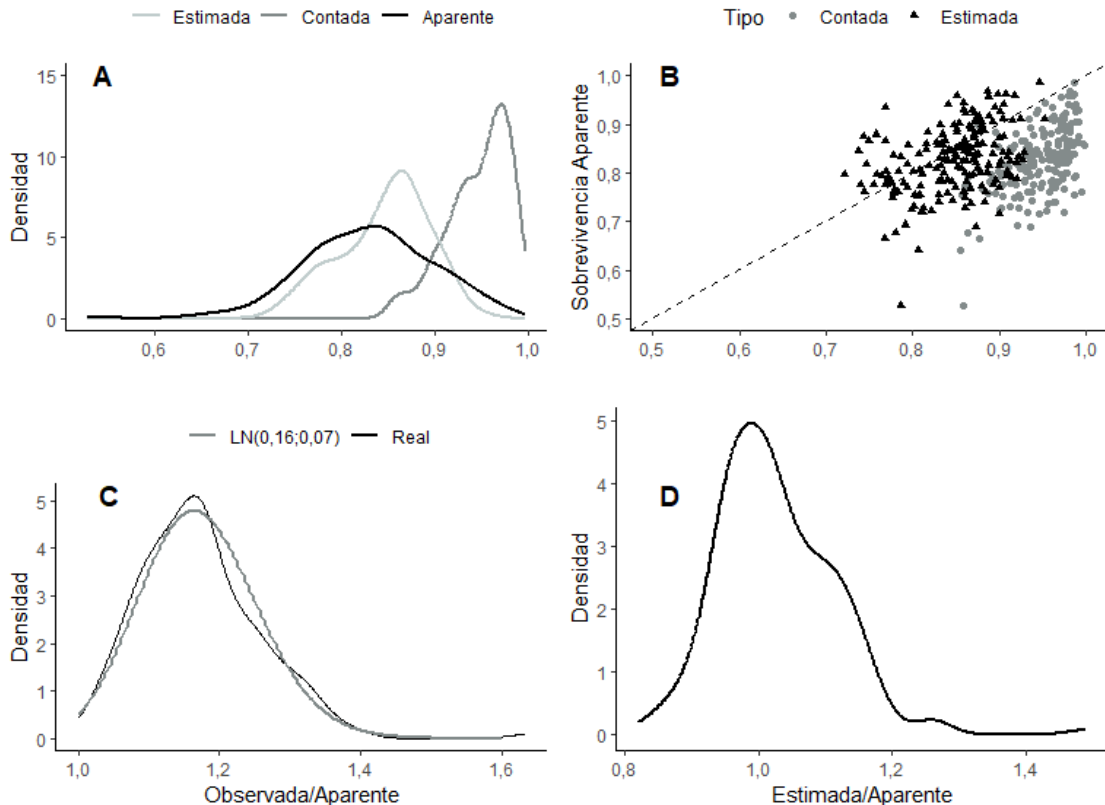


Figura 5. Comparación entre diferentes valores de sobrevivencia obtenidos en 186 estanques de producción comercial de tilapia, en Costa Rica, durante la campaña 2012-2013.

A) Gráficos de densidad de probabilidad para cada estimación de la sobrevivencia (para ver las definiciones de cada una, revisar el pie del Cuadro 1). B) Relación entre la sobrevivencia aparente con la sobrevivencia contada y la estimada. C) Gráfico de densidad de la razón entre sobrevivencia observada y aparente, en comparación con la densidad de probabilidad de una distribución lognormal con log-media = 0,16 y log-desviación = 0,07. D) Gráfico de densidad de probabilidad para la razón entre sobrevivencia estimada y aparente.

6.2 Método de simulación de la población

Se simularon 36 ciclos de producción con un número inicial de 1000 peces, lo cual, aunque es significativamente menor que el número inicial de los ciclos de producción en los datos reales, permitió que el modelo bayesiano propuesto arrojara las estimaciones en un tiempo razonable.

A cada ciclo de producción se le asignó un mes de inicio, muestreado de una distribución aleatoria entre 1 y 12, pues el mes de siembra es una variable importante dado que determina la época térmica en la que el ciclo productivo se

lleva a cabo, y tiene implicaciones en la tasa de crecimiento y la sanidad de las tilapias (Arce-Vega, 2014). De igual forma, la selección de la finca fue aleatoria.

Con fin de obtener el progreso en el peso de los peces cada día, se le asignó a cada pez un peso inicial aleatorio en gramos, tomado de una distribución normal con $\mu = 292$ y $\sigma = 65$, valores similares a los encontrados en el sistema de producción analizado. A cada pez se le estableció una ganancia diaria de peso aleatoria y fija a través de toda la simulación, muestreada de una distribución normal con parámetros dependientes del mes de siembra (Cuadro 13). Esto permitió simular un crecimiento lineal, que, aunque no corresponde con las curvas de crecimiento sigmoideas de los peces, es plausible en esta etapa de vida (**¡Error! No se encuentra el origen de la referencia.**).

Cuadro 13. Parámetros de simulación de la ganancia de peso vivo (g/d) por mes de siembra del estanque.

Mes Siembra	Promedio	Desviación Estándar
Enero	5,48	0,52
Febrero	5,84	0,44
Marzo	6,03	0,33
Abril	5,82	0,65
Mayo	6,12	0,51
Junio	6,13	0,42
Julio	5,74	0,85
Agosto	5,14	0,53
Setiembre	5,65	0,43
Octubre	6,02	0,37
Noviembre	6,05	0,73
Diciembre	5,43	0,48

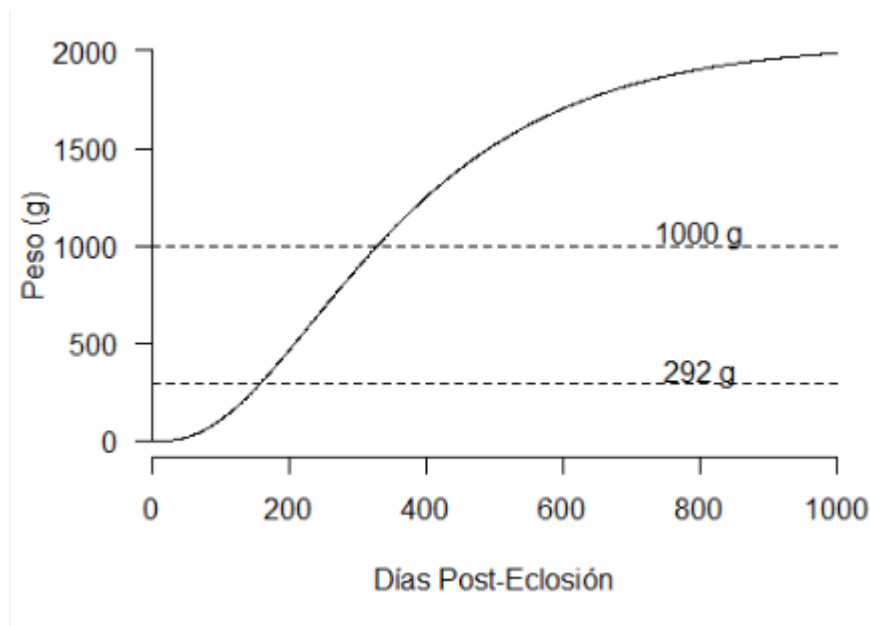


Figura 6. Curva de crecimiento de tilapias, según Zuniga-Jara y Goycolea-Homann (2014). El rango de pesos utilizado en el presente estudio se encuentra entre las líneas punteadas. Nótese que el crecimiento es aproximadamente lineal en esta etapa.

La mortalidad se determinó siguiendo los hallazgos descritos en el Capítulo 5. A continuación se describe el procedimiento:

1. Se toma un valor aleatorio de una distribución discreta uniforme entre 1 y 12, que corresponde al mes en que se inició el ciclo.
2. Se toma un valor aleatorio de una distribución de Bernoulli con $\pi = 0,5$. Si el valor resultante es un 1, el estanque se asigna a la finca "S", caso contrario a la finca "L"
3. Con los valores anteriores se obtiene el valor de densidad de biomasa (D_{BM}) sembrada, muestreado de una distribución normal con $\sigma = 0,39$ y μ dada por:

$$\begin{aligned} \mu_{BM} = & 0,40 + 0,16(\text{Mes}_2) + 0,22(\text{Mes}_3) + 0,30(\text{Mes}_4) + \\ & 0,58(\text{Mes}_5) + 0,32(\text{Mes}_6) + 0,30(\text{Mes}_7) + 0,59(\text{Mes}_8) + \\ & 0,76(\text{Mes}_9) + 0,94(\text{Mes}_{10}) + 0,56(\text{Mes}_{11}) + 0,24(\text{Mes}_{12}) + \\ & 0,28(\text{Finca}_S) \end{aligned}$$

Donde las variables Mes_i corresponden a variables indicadoras para cada mes de siembra, y la variable Finca_S es la indicada en el inciso 2. Este modelo se ajustó a partir de los datos de los 186 estanques, según se indica en el Capítulo 5.

4. En este punto se generan las curvas de supervivencia observable, dependiendo de si el mes se clasifica en la época cálida o fría:
- 4.1. Si el número de mes corresponde a la época cálida se obtienen tres valores aleatorios para cada población a partir de una distribución normal multivariada con promedio 0 y matriz de varianza-covarianza obtenida mediante la fórmula $\Sigma = \text{Diag}(\mathbf{S}) \cdot \mathbf{R} \cdot \text{Diag}(\mathbf{S})$ donde \mathbf{R} corresponde a la matriz de correlaciones entre los parámetros del modelo no lineal de Richards indicada en el Cuadro 8, mientras que \mathbf{S} corresponde a la columna de Desviaciones estándar de los efectos aleatorios, indicada en ese mismo cuadro, mientras que la función $\text{Diag}(\cdot)$ convierte el vector \mathbf{S} en una matriz cuadrada con los valores de \mathbf{S} en la diagonal y ceros en otro caso.
- 4.1.1. El primer valor aleatorio obtenido en 4.1 se suma al resultado de la multiplicación de los coeficientes del parámetro A , según lo indicado en Cuadro 7 con los valores de las variables $Fincas$ y D_{BM} obtenidos en los puntos 2 y 3, respectivamente, para lograr el valor de A (mortalidad potencial) correspondiente al estanque.
- 4.1.2. Se sigue el mismo procedimiento para los parámetros r (tasa de mortalidad) e i (localización del punto de inflexión), pero sumando el segundo y tercer valor aleatorio obtenidos en el punto 4.1, respectivamente.
- 4.1.3. Para el caso del parámetro de forma de la función no lineal de Richards (m), dado que no se encontró evidencia que apoyara la heterogeneidad individual del mismo, según se indicó en el Capítulo 5, se procedió a simplemente multiplicar los coeficientes indicados en el Cuadro 7 con los valores de $Fincas$ y D_{BM} .
- 4.1.4. Con los parámetros obtenidos se encuentra la curva de mortalidad observada diaria, mediante la primera derivada de (44), de la cual solamente falta sumar el error aleatorio $\epsilon_{jk} \sim N(0; 0,003)$, según lo indicado en el Cuadro 8.

- 4.2. Si el número del mes en que se sembró el estanque corresponde a la época fría, es necesario aplicar la primera derivada de la función doble logística indicada en (45). Los coeficientes para obtener los valores de los parámetros A, A', r, r', i e i' se indican en el Cuadro 9. Los interceptos aleatorios para cada población se obtienen como valores aleatorios de poblaciones normales independientes con media cero y las desviaciones estándar indicadas en la última columna del Cuadro 9. Una vez derivados los valores para los coeficientes, se aplica la fórmula doble logística y se le suma un error aleatorio $\epsilon_{jk} \sim N(0; 0,0005)$.
5. En los datos que se tienen de la empresa se conoce la sobrevivencia observable ϕ_c y la sobrevivencia aparente ϕ (Cuadro 12) y que su cociente Ω se distribuye aproximadamente como *LogNormal*(0,16; 0,07) (**¡Error! No se encuentra el origen de la referencia.**). No obstante, estos son resultados finales del ciclo, pero no se conoce la relación entre dichas variables para un día cualquiera del ciclo de producción, por lo cual es necesario suponer posibles formas para dicha relación. Para hacer la simulación se empieza tomando un valor aleatorio de la distribución *log-normal*, previamente mencionada, pero truncada en 1 para evitar situaciones en las que $\phi < \phi_c$. Para cualquier estanque con una edad de cosecha a se puede pensar que $\phi = \prod_{t=1}^a \vartheta_t$, donde ϑ_t es la tasa de sobrevivencia del día t . Sea, a su vez, $\phi_c = \prod_{t=1}^a \vartheta_{c_t}$. Recordando que $\frac{\phi}{\phi_c} = \Omega$, por tanto $\phi = \Omega \prod_{t=1}^a \vartheta_{c_t} = \prod_{t=1}^a \omega(t) \vartheta_{c_t}$, donde $\omega(t)$ es alguna función que define la relación entre ϑ_t y ϑ_{c_t} para cualquier día t del ciclo productivo tal que $\prod_{t=1}^a \omega(t) = \Omega$. Puede haber muchas posibles formas funcionales, pero se escogieron las siguientes:
- Forma 1: Se asume que $\omega(t) = \Omega^{\frac{1}{a}}$, lo que es equivalente a decir que hay un coeficiente de recuperabilidad fijo para todo el ciclo de producción.
 - Forma 2: En este caso, el valor de $\omega(t)$ aumenta con el tiempo, según la función $\omega(t) = \Omega^{\frac{2t-1}{a^2}}$, partiendo de la suposición de que la exactitud en el reporte de la mortalidad es mayor hacia el inicio del ciclo que al final de

este, por cuanto al inicio es donde generalmente se dan los episodios de mortalidad más importantes, en especial en la época cálida.

- Forma 3: En este caso se asume un coeficiente de recuperabilidad fijo para todo el ciclo de producción, tal como en la forma 1, pero se selecciona un día aleatorio d de la segunda mitad del ciclo productivo para introducir una pérdida puntual. La importancia relativa de esta pérdida puntual, respecto a toda la mortalidad no observada está dada por una constante κ , que para efectos de la simulación es muestreada de una distribución uniforme entre 0,2 y 0,7:

$$\omega(t) = \begin{cases} \Omega \frac{1-\kappa}{a}, & \text{si } t = d \\ \Omega \frac{\kappa}{a}, & \text{en otro caso} \end{cases} \quad (1)$$

Se le asignó una forma a cada población y se generó una curva de mortalidad total, a partir de la mortalidad observada.

6. Se generó una matriz con la cantidad de animales sembrados, todos vivos al día de la siembra (día 1). Para cualquier día $t > 1$ se obtuvo un valor aleatorio $z_{i,t}$ que indica el estado (1 = vivo, 0 = muerto) para cada animal i a partir de una distribución de Bernoulli con parámetro $\vartheta_t z_{i,t-1}$. El valor de ϑ_t se obtuvo a partir de la curva de mortalidad diaria total (observable y no observable) generada en el punto 5.

Considerando el mencionado efecto de dominancia por tamaño que se ha observado en ciertas especies de tilapia (Barcellos et al., 1999; Fernandes y& Volpato, 1993), se calcularon las probabilidades de captura a partir del cuantil de la distribución de pesos, ajustados para que las probabilidades sumen 1.

La siembra se usó como la primera ocasión de captura, donde se marcará una cantidad m_0 de peces. La última ocasión de muestreo se fijó 7 días antes de la fecha de cosecha. Los restantes muestreos se situaron en puntos equidistantes en el tiempo.

6.3 Diseño de los escenarios de estimación

Los escenarios de estimación se realizaron siguiendo un esquema anidado. Los 36 ciclos productivos se muestrearon para probar en cada uno dos diferentes números de ocasiones de muestreo (4 y 6) y tres diferentes tamaños de muestra en cada captura ($n = 50, 100$ y 150 , equivalentes aproximadamente al 5, 10 y 15% de la población inicial), por lo cual se obtuvieron 909 estimaciones de abundancia y 680 de sobrevivencia y recuperabilidad. Como factores a nivel de ciclo productivo se analizó la forma de la relación entre sobrevivencia aparente y sobrevivencia observada, así como la época del año.

6.4 Análisis de los resultados de estimación

El modelo bayesiano se escribió y se estimó mediante la plataforma JAGS (Plummer, 2017) que usa el algoritmo de cadenas de Markov y Montecarlo con muestreo de Gibbs. La verificación de convergencia se realizó en uno solo de los modelos ajustados, de forma que se obtuvieran los valores necesarios para todas las poblaciones y condiciones de estimación. El criterio de Gelman-Rubin, la inspección de los diagramas de traza y de los gráficos de autocorrelación se usaron para verificar la convergencia.

Para cada estimación se generaron tres cadenas, con puntos iniciales aleatorios diferentes, de forma que se pudiera verificar la convergencia, con $5,00 \times 10^5$ iteraciones cada una, de las cuales se descartaron $4,50 \times 10^5$ en la fase de quemado (*burn in*). Además, se hizo un adelgazamiento (*thinning*) de 1 de cada 100 resultados.

Para el ajuste del modelo se usaron dos computadoras, una con procesador AMD Ryzen 3 con 2,0 GHz y 12 GB de memoria RAM, la otra un servidor virtual con 16 GB de RAM y un procesador Intel Scalable 3,0 GHz, como interfaz.

Los resultados de sobrevivencia y abundancia estimados se compararon con los valores usados para generar cada simulación, con el fin de obtener las siguientes métricas:

1. Sesgo relativo (SR%): Se refiere a la diferencia entre el valor esperado del estimador $E(\theta)$, respecto al valor del parámetro (θ). En términos matemáticos se expresa como:

$$SR(\%) = \frac{E(\theta) - \theta}{\theta} \times 100 \quad (2)$$

2. Error estándar proporcional (EE%): Es el error estándar de la estimación de θ , expresado como proporción del valor de dicha estimación:

$$EE(\%) = \frac{EE(\hat{\theta})}{\hat{\theta}} \times 100 \quad (3)$$

Recordando que $EE(\hat{\theta}) = \sqrt{var(\hat{\theta})/n}$, donde n es la cantidad de ciclos productivos usadas en la estimación de θ y $var(\hat{\theta})$ es la varianza residual de un modelo lineal con población como factor aleatorio.

El cálculo se realiza de esta manera para evitar el efecto de pseudorreplicación que se tendría si se considera cada observación como independiente (Lazic, Clarke-Williams y Munafò, 2018).

3. Error porcentual: Es la diferencia entre el valor real de θ y el valor estimado, dividida por el valor real:

$$EP(\%) = \frac{\hat{\theta} - \theta}{\theta} \times 100 \quad (4)$$

Esta métrica se empleó en gráficos para poder mostrar, a la vez, la varianza y el sesgo en las estimaciones, en relación con otras variables independientes analizadas. En dichos gráficos se agregaron estimaciones suavizadas para los percentiles 5 y 95, así como para el promedio. Las estimaciones de regresión suavizada de los percentiles 5 y 95 se obtuvieron mediante la función `rqss` del paquete `quantreg` (Koenker, 2022) usando los valores por defecto. La estimación de regresión suavizada para el promedio se obtuvo mediante la función `gam` del paquete `mgcv` (Wood, 2017). Ambos son paquetes complementarios al lenguaje de programación estadística R, versión 4.0.2 (R Core Team, 2020). En este mismo sistema se realizaron las simulaciones de población, se ajustaron los modelos

bayesianos y se obtuvieron resultados y se realizaron los cuadros y gráficos mediante diferentes funciones disponibles.

Los cuadros y gráficos con los indicadores anteriores se obtuvieron en dos condiciones: usando todos los resultados y recortando 10% de valores extremos (5% en cada cola). No obstante, al analizar los resultados obtenidos con datos recortados, no se detectó una tendencia clara en los indicadores ni una mejoría general en la interpretabilidad, por lo que se determinó solamente presentar el análisis con los datos completos. Los cuadros y figuras obtenidos, eliminando las colas, se encuentran en el Anexo 1.

El capítulo que termina mostró la metodología empleada para la simulación de los datos y la obtención de las estimaciones. En el siguiente capítulo se presentará la evaluación de los resultados para los parámetros de abundancia, sobrevivencia y recuperabilidad, por medio del sesgo relativo y el error estándar relativo.

7 RESULTADOS

En el presente capítulo se analizan los resultados de las simulaciones por medio de los indicadores de Sesgo Relativo (SR) y de Error Estándar Relativo (EER). Se empieza con por el análisis de las estimaciones de abundancia, para posteriormente analizar las estimaciones de sobrevivencia y de recuperabilidad.

7.1 Abundancia

Respecto de la estimación de abundancia, en el Cuadro 1 Cuadro 14 se muestra que al aumentar el número de ocasiones de muestreo y el tamaño de muestra, hay reducciones en el EER, lo cual es una señal de consistencia en este estimador. Por otra parte, el sesgo, que en los tamaños de muestra bajos es negativo, se reduce y se hace positivo con tamaños de muestra más amplios y mayor número de ocasiones de muestreo. Los mejores resultados se dan usando 6 ocasiones de muestreo y un tamaño de muestra de 100 individuos por ocasión, aunque con valores de sesgo y error estándar relativo muy similares a los que se dan con 4 ocasiones de muestreo y 150 individuos por ocasión. En ambos casos el sesgo y el error estándar rondan el 1%.

Cuadro 14. Indicadores de sesgo y el error estándar relativo de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%) ¹	EER (%) ²
4	50	-7,1	1,4
	100	-2,3	2,1
	150	1,1	1,3
6	50	-3,3	1,2
	100	0,9	1,0
	150	2,5	0,8

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

En la **¡Error! No se encuentra el origen de la referencia.** se muestra que generalmente hay una buena correspondencia entre las estimaciones de abundancia obtenidas y los valores reales para la población analizada. Empero,

para valores bajos de abundancia que ocurren especialmente al final del ciclo en casos de alta mortalidad, podría haber un sesgo mayor.

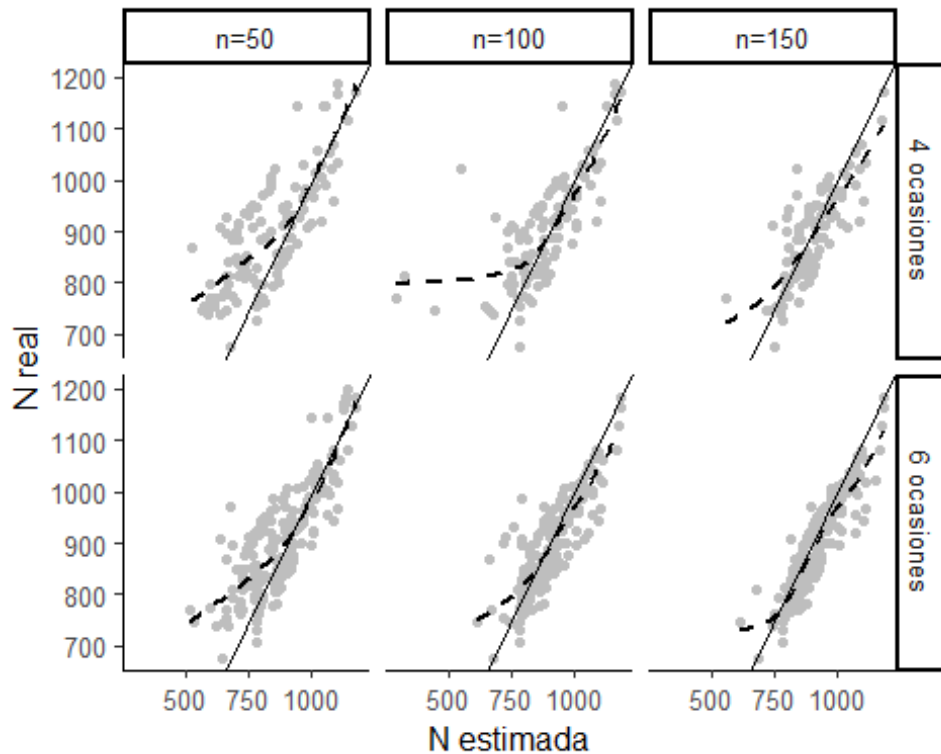


Figura 7. Asociación entre la abundancia estimada y la abundancia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de la relación entre los valores estimados y los reales

La **¡Error! No se encuentra el origen de la referencia.** muestra de manera gráfica lo anteriormente apuntado el sesgo y la dispersión de las estimaciones se reducen, tanto al aumentar el tamaño de muestra como el número de ocasiones de muestreo. No obstante, adicionalmente muestra que las distribuciones de los estimadores presentan una asimetría negativa, con una importante probabilidad de valores extremos a la izquierda. Asimismo, se destaca la similitud de las distribuciones de resultados al combinar 4 ocasiones de muestreo con 150 individuos muestreados o 6 ocasiones con $n = 100$.

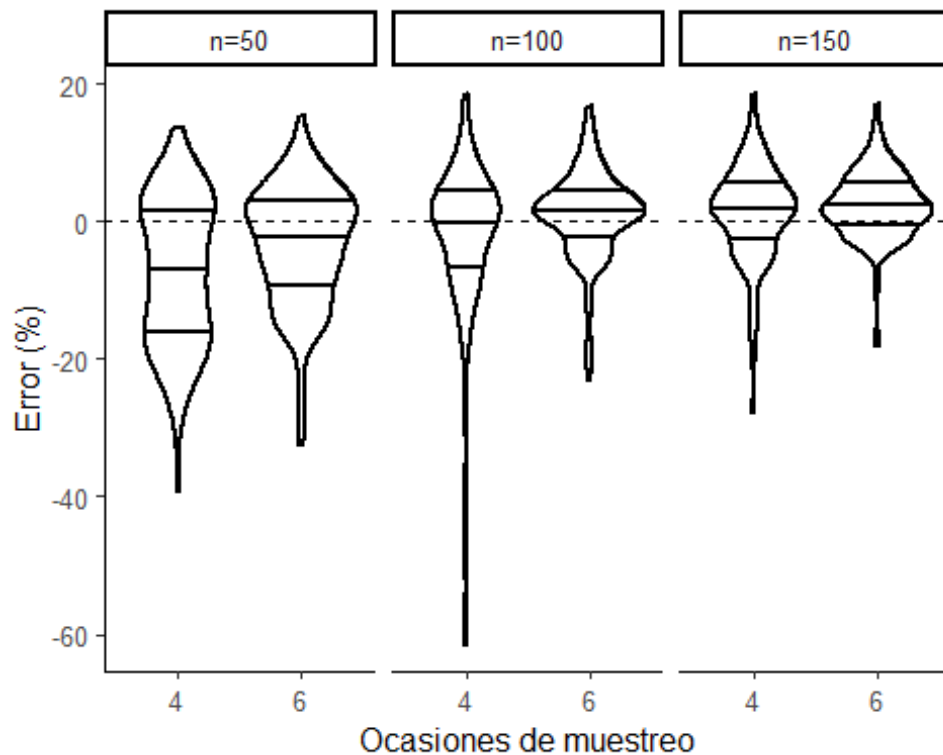


Figura 8. Distribución del error porcentual de la estimación de abundancia por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se marcan los percentiles 25, 50 y 75.

La **¡Error! No se encuentra el origen de la referencia.**, por su parte, indica que tanto el sesgo y la dispersión de las estimaciones aumenta con la edad, lo cual es resultado lógico del método de estimación, que parte del valor “verdadero” sembrado, pero, posteriormente, agrega una serie de estimaciones dependientes de las anteriores, por lo cual las desviaciones podrían verse amplificadas.

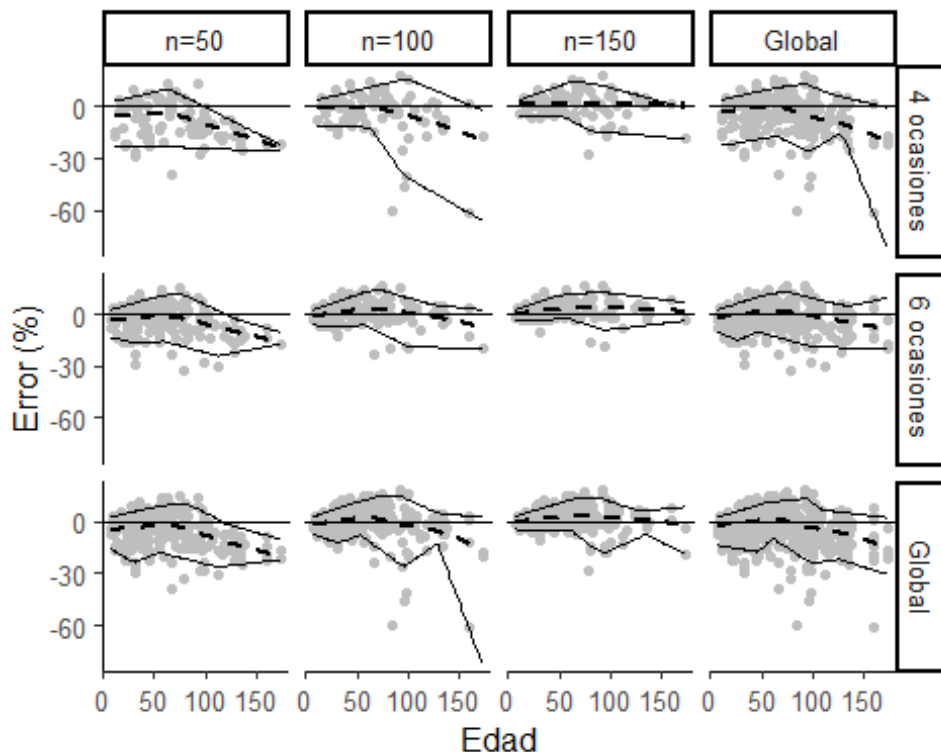


Figura 9. Asociación entre la edad del estanque y el error porcentual en las estimaciones de abundancia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos.

Se aprecia que los estimadores empleados son sensibles a la forma de la relación entre la mortalidad real y la observada, pero la presencia o ausencia de pérdidas puntuales afectan poco su comportamiento (Cuadro 15). La forma 3, aquella donde se presentan robos o pérdidas puntuales de peces muestra un sesgo negativo importante, en comparación con asumir una recuperabilidad constante o que aumenta con el tiempo (formas 1 y 2 respectivamente) (Cuadro 15). El error a través de los días de edad del estanque es, sin embargo, bastante similar al de la forma 2, aunque con valores más bajos en su extremo superior, lo cual puede dar lugar a ese rendimiento global (Figura 10).

En general las estimaciones que se obtuvieron de abundancia fueron cercanas a los valores reales, con niveles de sesgo y errores estándar relativos menores o

iguales a 7 y 3% respectivamente (Cuadro 14, Cuadro 15). Por su parte, el método usado por la empresa para proyectar las cosechas tiene sesgo relativo de 7,6% y error estándar relativo de 1,5%, lo cual señala que el método propuesto tiene menor sesgo, pero una varianza levemente superior.

Cuadro 15. Indicadores de sesgo y varianza de los estimadores de abundancia según la forma asumida para la relación entre mortalidad observada y mortalidad real.

Forma	SR (%) ¹	EER (%) ²
1	1,7	2,0
2	-1,4	2,7
3	-4,0	2,3

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

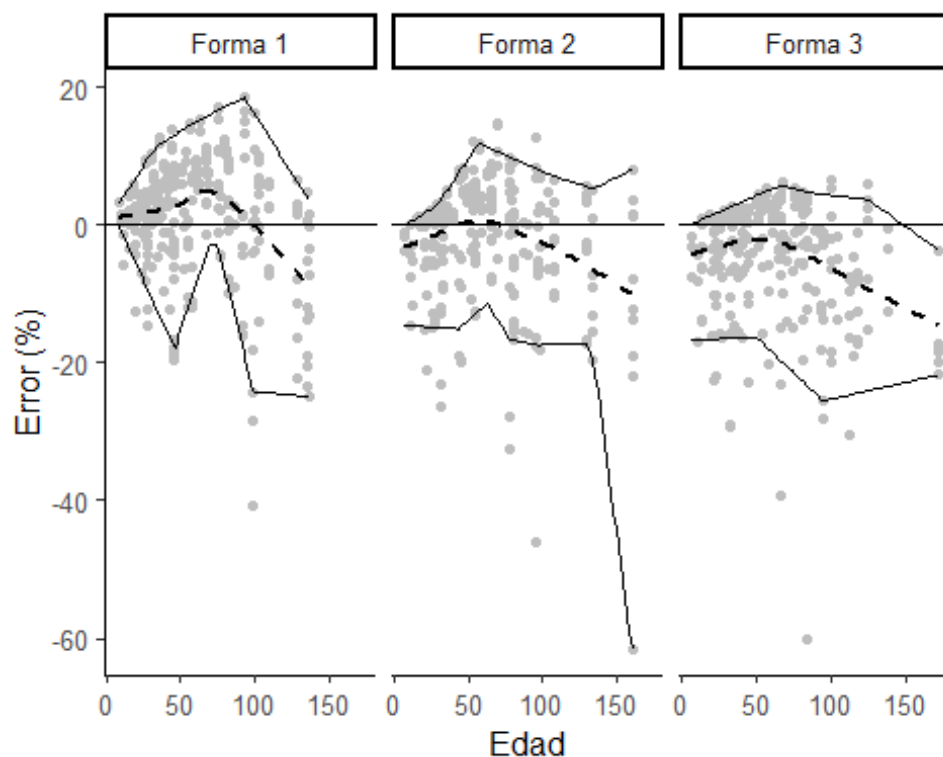


Figura 10. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos.

En cuanto a la época de siembra, los resultados de la época fría suelen tener un nivel considerable de subestimación (Cuadro 16). Sin embargo, la dispersión del error de las estimaciones en dicha época es la más constante a través del ciclo de

cultivo, mientras que la época caliente y de transición muestran resultados cuyas estimaciones tienen mayor dispersión al final del ciclo (**¡Error! No se encuentra el origen de la referencia.**).

Aunque es un resultado esperable, pues las estimaciones al final del ciclo combinan la incertidumbre de las estimaciones previas, esto podría generar estimaciones de cosecha imprecisas. No obstante, el método que se propone en este trabajo es similar en precisión al modelo de la empresa, aunque con un sesgo menor para la época cálida (Cuadro 16).

Cuadro 16 Indicadores de sesgo y error estándar relativo de los estimadores de abundancia según la época climática de siembra, comparando dos métodos de estimación

Época	Este trabajo		Modelo de la Empresa	
	SR (%) ¹	EER (%) ²	SR (%)	EER (%)
Cálida	2,5	1,9	11,9	2,1
Fría	-6,0	2,5	2,2	2,7
Transición	-2,9	2,8	3,7	3,3

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

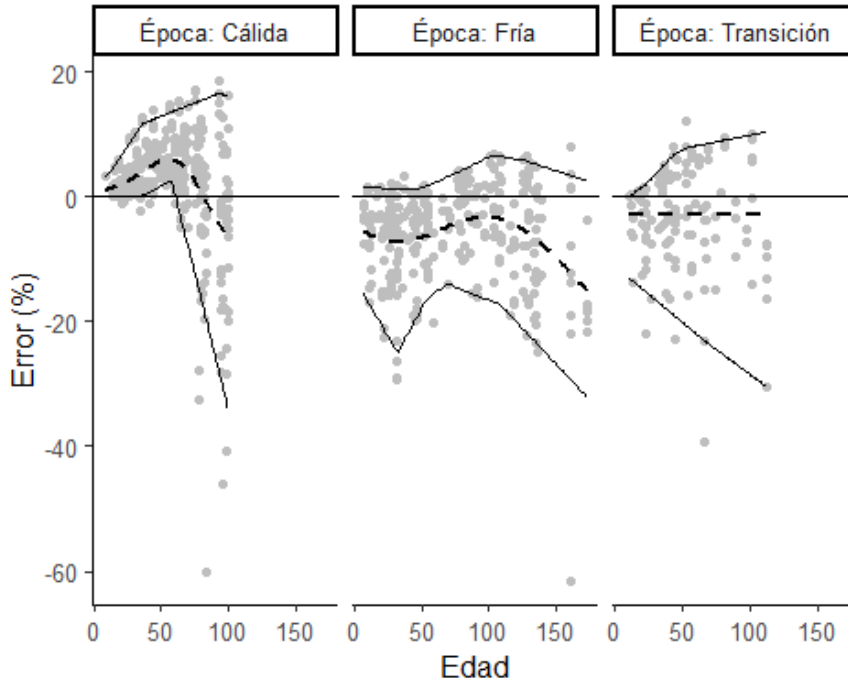


Figura 11. Error porcentual en la estimación de la abundancia por edad del estanque y época de siembra.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos

7.2 Sobrevivencia

El Cuadro 17 muestra que, en general el sesgo y el error estándar relativo del estimador de sobrevivencia fue incluso menor que para el caso de la abundancia. Cuando se descuenta el efecto del número de ocasiones de muestreo no es particularmente importante el tamaño de muestra, pues ambos indicadores se mantienen bastante similares cuando utilizan 6 ocasiones de muestreo. Sin embargo, en el caso de 4 ocasiones de muestreo, sí hay variabilidad importante en sesgo y errores estándar relativos, interesantemente con el tamaño de muestra intermedio se obtienen los peores resultados, lo cual es contradictorio a lo esperado (mayor tamaño de muestra y más ocasiones de muestreo deberían resultar en mejor precisión).

No obstante, lo anteriormente mencionado no ocurre cuando las ocasiones de muestreo son 6. Se observa en estas nuevamente que los resultados de 4 ocasiones de muestreo y $n = 150$ son similares a los de 6 ocasiones de muestreo y $n = 100$, pues coinciden en valores similares de sesgo, e indican que cerca de los 600 puntos de muestreo podría haber un óptimo.

Cuadro 17. Indicadores de sesgo y error estándar relativo de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%) ¹	EER (%) ²
4	50	-1,4	1,3
	100	-2,5	2,4
	150	0,2	1,3
6	50	-0,7	1,1
	100	0,1	0,9
	150	0,6	0,7

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

En general, se observa una buena correspondencia entre los valores estimados y los reales, no obstante, la existencia de valores estimados extremadamente bajos (0,4-0,9) mucho menores que la escala de valores reales afecta severamente los

resultados (**¡Error! No se encuentra el origen de la referencia.**). El análisis de la **¡Error! No se encuentra el origen de la referencia.**, sin embargo, muestra un panorama relevante para la investigación, ya que hay una tendencia a que entre el 50% y el 75% de los valores estimados tienen un error positivo, pero los valores extremos hacen que haya compensación en la media de los resultados, haciendo lucir el sesgo relativo como relativamente pequeño.

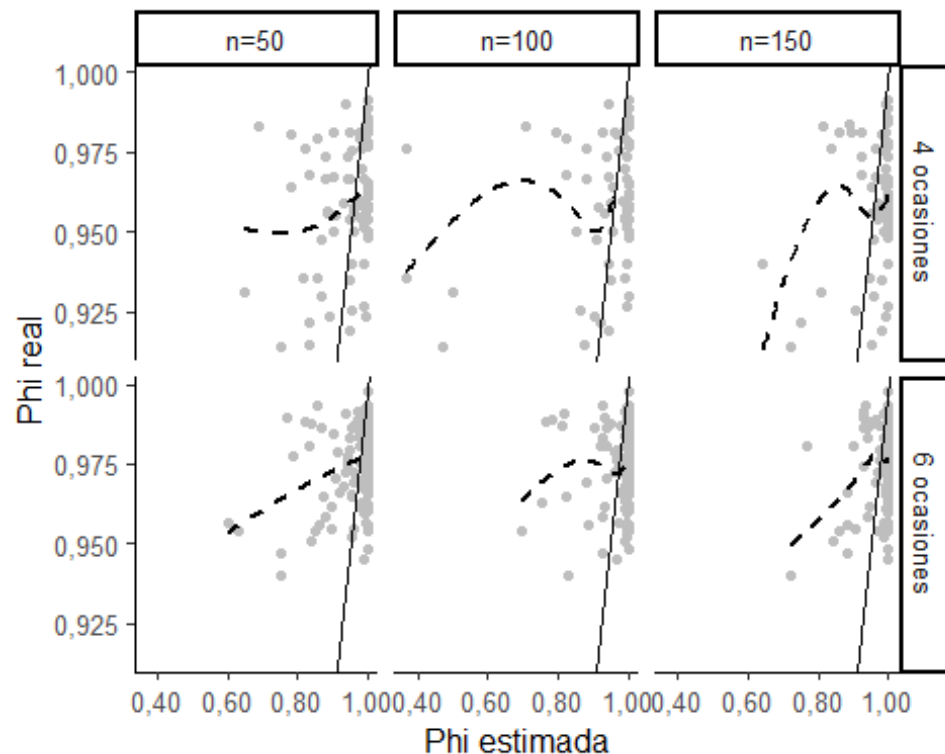


Figura 12. Asociación entre la sobrevivencia estimada y la sobrevivencia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de la relación entre los valores estimados y los reales

La **¡Error! No se encuentra el origen de la referencia.** muestra, además, un resultado interesante, pues el nivel de sobreestimación de la sobrevivencia aumenta en conjunto con el tamaño de las muestras obtenidas y el número de ocasiones de muestreo. Concretamente, cuando se tienen cuatro ocasiones de

muestreo con $n = 50$, aproximadamente se obtiene 50% de estimaciones con error positivo, pero este valor va aumentando con el tamaño de muestra, hasta llegar a 75% si se toma $n = 150$, manteniendo el número de ocasiones de muestreo en cuatro. Existe un comportamiento de crecimiento más sutil, pero de mucha menor magnitud cuando el número de ocasiones de muestreo es igual a 6.

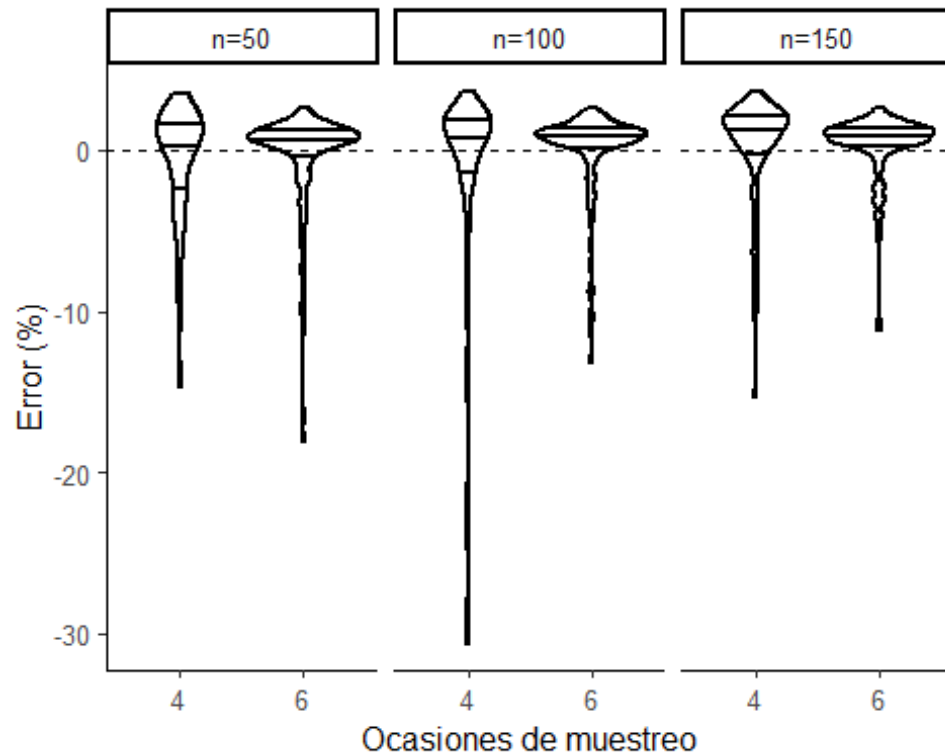


Figura 13. Distribución del error porcentual en la estimación de sobrevivencia por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se muestran los percentiles 25, 50 y 75.

La edad del estanque en la cual se obtiene la estimación de sobrevivencia no es relevante, en general, como para afectar el sesgo de las estimaciones obtenidas. En la **¡Error! No se encuentra el origen de la referencia.** se muestra que el promedio del error (línea punteada) tiende a mantenerse cercano a la línea de 0% de error, sin tendencias claras a la baja o al alza, salvo en ciertos casos puntuales, donde ciertos valores extremos ejercen una influencia importante.

No obstante, se logra notar a simple vista que hay una tendencia generalizada a un mayor error estándar relativo en las estimaciones (la cual se puede observar como la distancia entre las líneas delgadas continuas, correspondientes a los percentiles 5 y 95) cerca de 100 días. Este comportamiento puede deberse a la conjunción de datos de estanques que son cosechados precozmente, con estanques en la mitad de su vida de crecimiento, que ya se han alejado lo suficiente en edad de la siembra, donde la población es conocida, pero no se ha recolectado información suficiente para tener estimaciones más precisas.

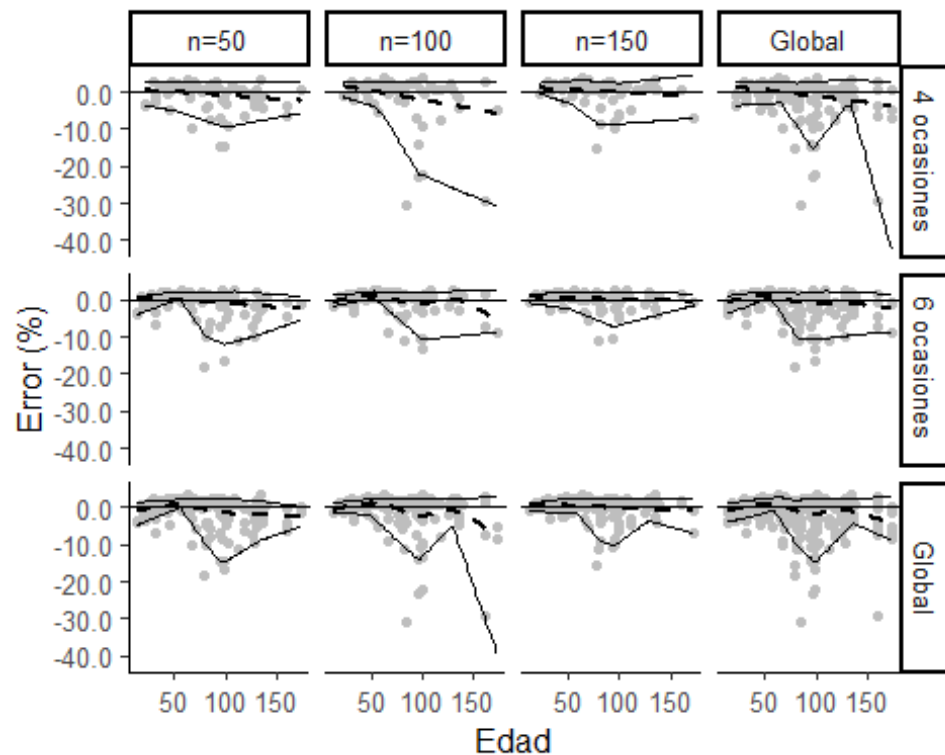


Figura 14. Asociación entre la edad del estanque y el error porcentual en las estimaciones de supervivencia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos

El Cuadro 18 y la **¡Error! No se encuentra el origen de la referencia.** muestran que las estimaciones de supervivencia se ven poco afectadas si la tasa de

recuperación es constante o disminuye con la edad del estanque se da una importante subestimación de la sobrevivencia aparente al haber posibilidad de pérdidas puntuales. Este es un resultado relevante para la investigación porque lo esperado era una sobreestimación, asumiendo que no se descubre el problema de la pérdida de los peces. No obstante, tal parece que el método logra ajustar las estimaciones, aunque con un poco de rigurosidad. **¡Error! La autoreferencia al marcador no es válida.** Resulta importante, pues, que este tipo de análisis de métodos de estimación se hagan siguiendo, no solamente las fórmulas convencionales, sino también analizando la distribución de los resultados obtenidos, lo cual permitiría descubrir comportamientos como estos.

Cuadro 18. Indicadores de sesgo y error estándar relativo de los estimadores de sobrevivencia según la forma asumida para la relación entre mortalidad observada y mortalidad real

Forma	SR (%) ¹	EER (%) ²
1	0,6	1,7
2	0,1	2,5
3	-2,1	2,2

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

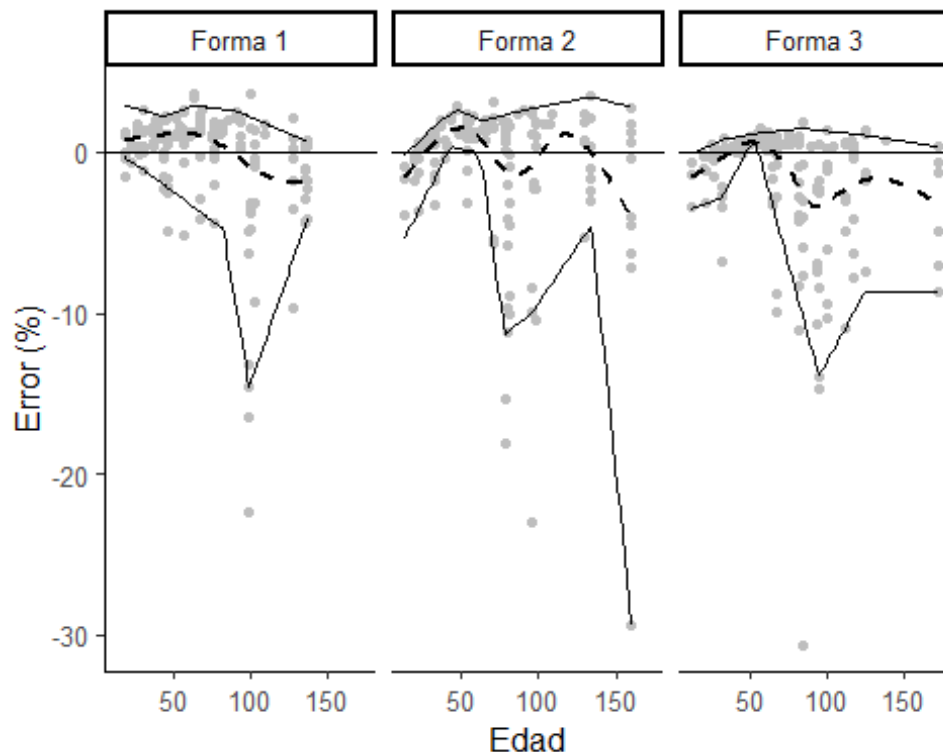


Figura 15. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95

A pesar de los problemas e inconsistencias detectadas, el modelo bayesiano de captura y recaptura usado superó el rendimiento del método usado en la empresa, que logra 8,12% de SR con 0,4% de EER. El desglose del error por época permite ver que el sesgo es despreciable en el modelo propuesto, mientras que el de la empresa muestra ventaja en precisión medida con el EER, ya que se obtienen estimaciones entre dos y cuatro veces más precisas (Cuadro 19). No obstante, es importante mejorar las estimaciones extremas y la subestimación en caso de pérdidas o robos.

Cuadro 19. Comparación del sesgo relativo (SR) y el Error Estándar Relativo (EER) entre el modelo propuesto en este trabajo y el modelo usado en la empresa con relación a las épocas climáticas.

Época	Este Trabajo		Modelo de la empresa	
	SR (%) ¹	EER (%) ²	SR (%)	EER (%)
Cálida	-0,9	2,1	12,9	0,5
Fría	-0,1	1,8	2,4	0,5

Época	Este Trabajo		Modelo de la empresa	
	SR (%) ¹	EER (%) ²	SR (%)	EER (%)
Transición	0,0	2,1	4,0	1,1

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

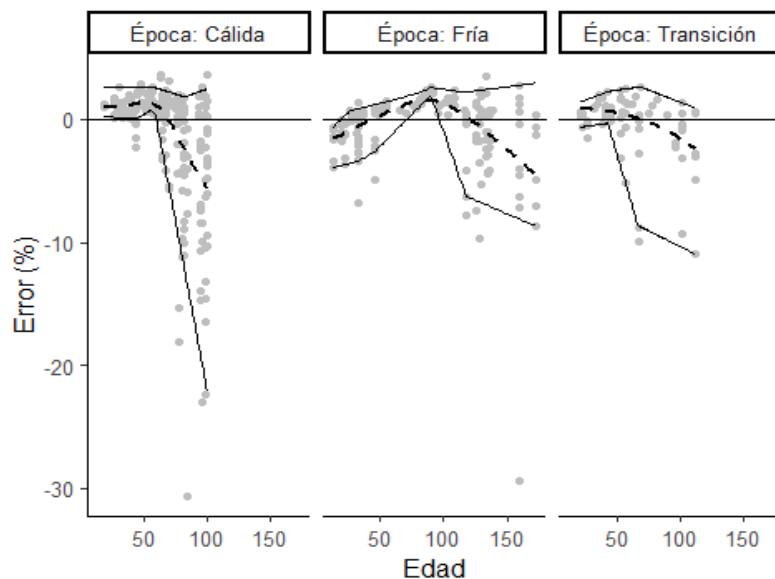


Figura 16. Error porcentual de la estimación de sobrevivencia por edad y época del lote de producción.

La línea punteada representa una estimación suavizada del promedio y las líneas continuas representan los percentiles 5 y 95.

7.3 Recuperabilidad

La recuperabilidad, o la tasa de cadáveres recuperados como proporción de la mortalidad total es un parámetro importante de estimar, pues está relacionado con la estimación de la mortalidad real. Sin embargo, los resultados, en términos de SR y EER (Cuadro 20) son bastante altos en comparación con los resultados obtenidos en las variables analizadas anteriormente. Los sesgos relativos superan por mucho el 100% mientras que error estándar relativo es superior al 13%, indicando resultados nada buenos del método de estimación.

Cuadro 20. Indicadores de sesgo y varianza de los estimadores de recuperabilidad según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%) ¹	EER (%) ²
4	50	868,2	18,5
	100	816,0	18,5
	150	763,9	19,0
6	50	693,9	13,1
	100	658,9	13,2
	150	622,0	13,2

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

La **¡Error! No se encuentra el origen de la referencia.** muestra que, mientras las estimaciones del modelo empleado se distribuyen en todo el rango entre 0 y 1, los valores reales son, mayoritariamente, inferiores a 0,1 y, los pocos que lo superan, llegan como máximo a 0,8. Por lo que se puede pensar que, dado que los valores reales suelen ser tan bajos y el modelo empleado los sobreestima, los errores se magnifican al presentarlos como porcentaje, dado que son coeficientes de variación, lo que permite y facilita su comparación.

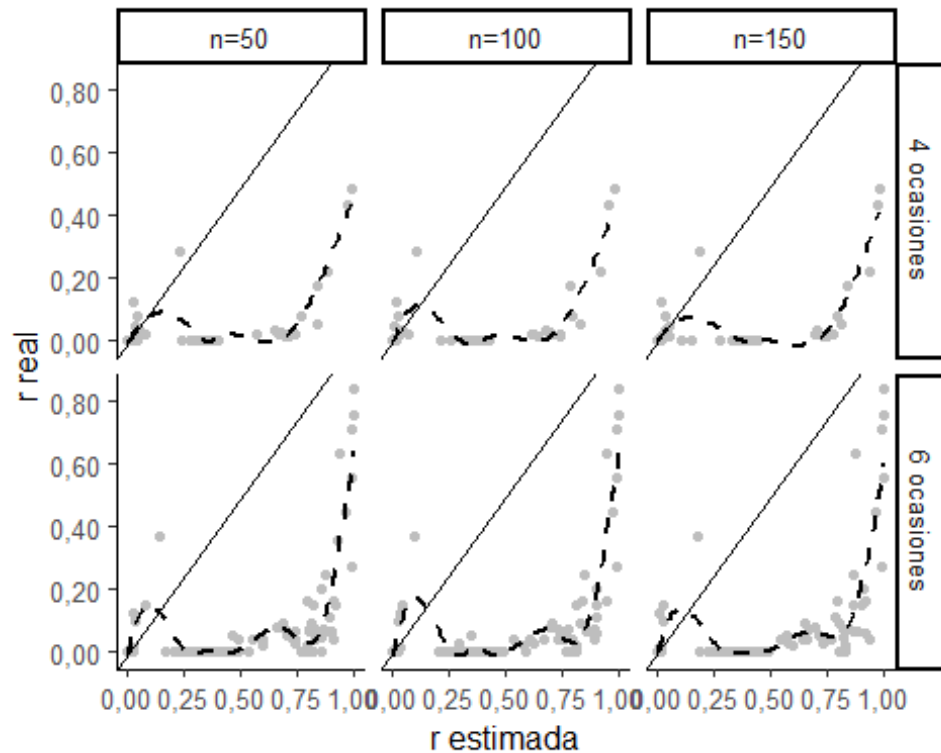


Figura 17. Asociación entre la recuperabilidad estimada y la recuperabilidad real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de la relación entre los valores estimados y los reales

La **¡Error! No se encuentra el origen de la referencia.** muestra otra perspectiva del problema. Específicamente, 75% o más de las estimaciones están por encima de los valores reales. Además, la distribución de los datos tiende a ser bimodal, con una moda entre 30% y 50% de error y la otra alrededor de 0% de error. La **¡Error! No se encuentra el origen de la referencia.**, sin embargo, muestra un comportamiento notable, que el nivel de error de las estimaciones tiende a disminuir con la edad del estanque, lo que señala que el problema podría ser de falta de información suficiente para identificar adecuadamente el parámetro, por lo cual, al aumentar la edad, y tener más ocasiones de muestreo acumuladas, el error se reduce. Sobre este aspecto, lo destacable para la investigación es que, si bien el tamaño del error se reduce cuando la cosecha del estanque se acerca y es

beneficioso para hacer predicciones de cosecha precisas los días antes de esta, es problemático para saber el tamaño poblacional en puntos intermedios.

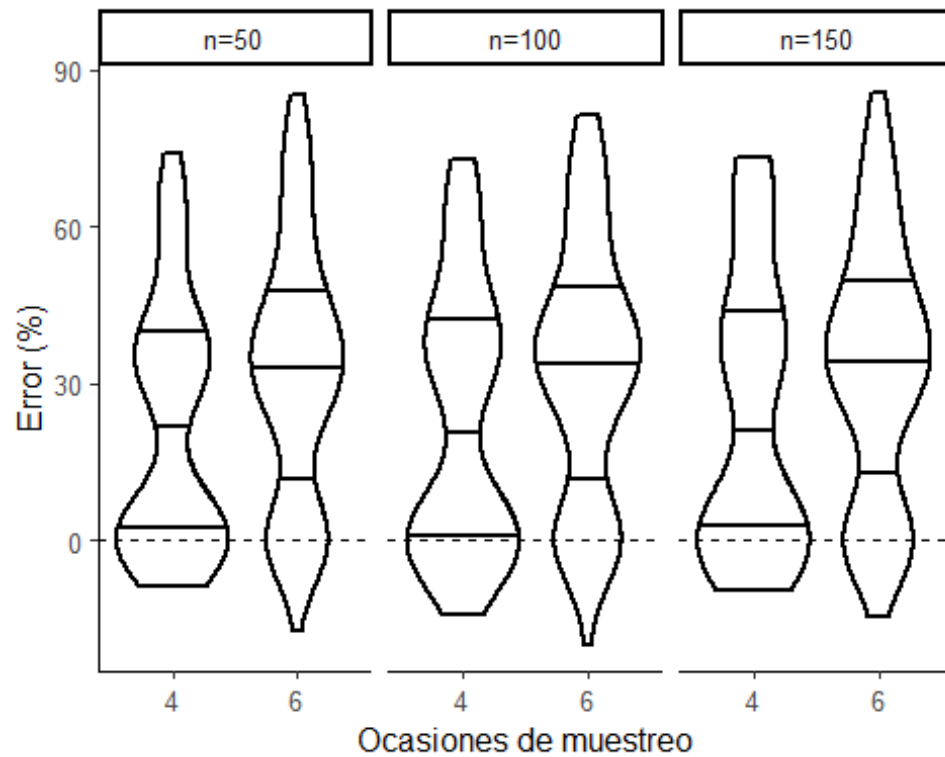


Figura 18. Distribución del error porcentual en la estimación de recuperabilidad por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se muestran los percentiles 25, 50 y 75.

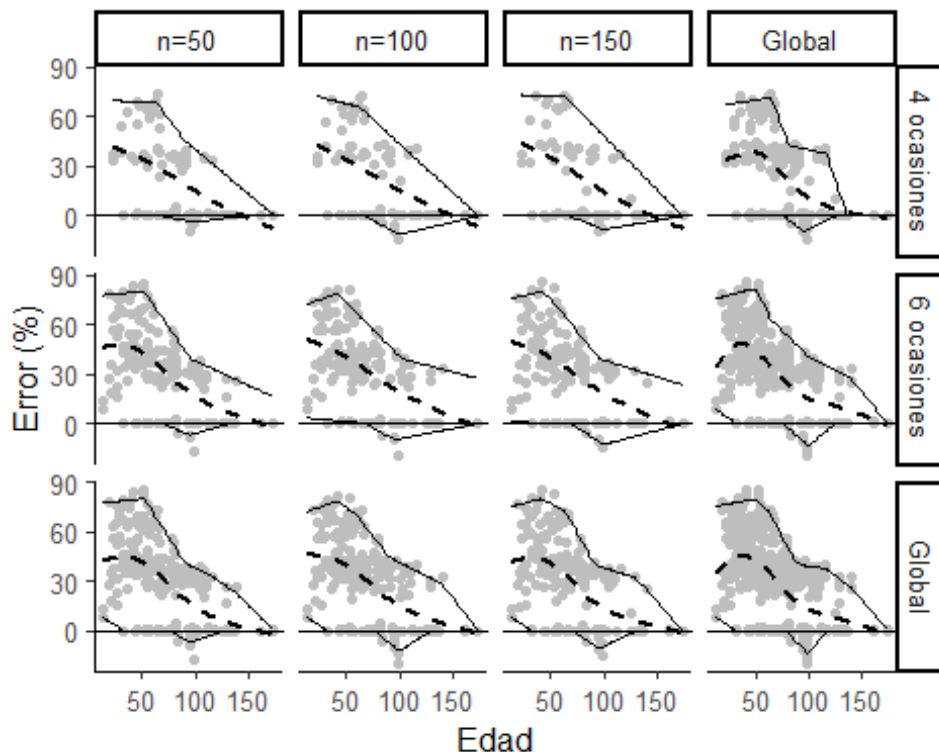


Figura 19. Asociación entre la edad del estanque y el error porcentual en las estimaciones de recuperabilidad, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos.

Según se puede ver en la información presentada en el **¡Error! La autoreferencia al marcador no es válida.**, se puede decir que la forma de la relación entre mortalidad real y mortalidad observada es irrelevante, pues, aunque pareciera ser que los estanques donde hay robo hay menor sesgo se siguen presentando valores de sesgo muy altos y, cuando se analiza la **¡Error! No se encuentra el origen de la referencia.**, se descubre que el comportamiento de las estimaciones para cada una de las formas asumidas es realmente muy similar como para ser discutido en este documento.

En cuanto a la recuperabilidad, este es un parámetro que no se mide directamente en la empresa, sino que se utiliza en el presente trabajo para facilitar el ajuste entre la mortalidad real y la observada. Al ser este parámetro una innovación del

presente estudio, no es posible compararlo con los datos previos obtenidos en la empresa, pero de todas formas, dado su rendimiento, no pareciera ser necesario.

Cuadro 21. Indicadores de sesgo y error estándar relativo de los estimadores de recuperabilidad según la forma asumida para la relación entre mortalidad observada y mortalidad real.

Forma	SR (%) ¹	EER (%) ²
1	952,6	24,5
2	837,1	24,0
3	464,5	25,8

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

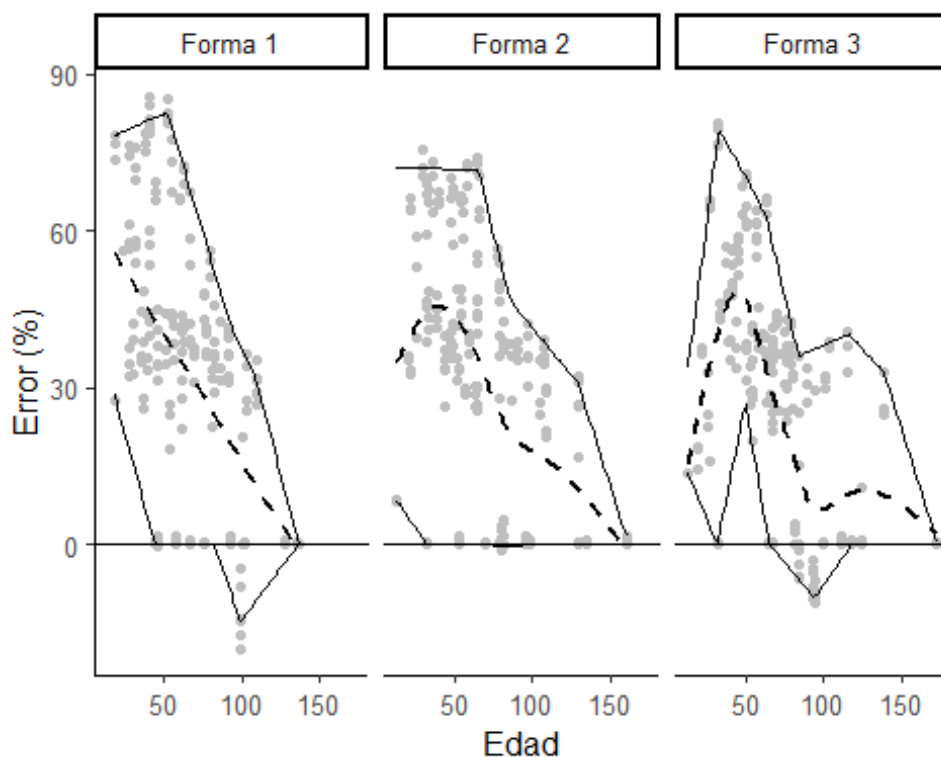


Figura 20. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95.

Si bien la recuperabilidad es difícil de estimar, fue en la época cálida donde se presentó mayor problema en su cálculo, con un sesgo de 1123% y un error

estándar relativo de 19% (Cuadro 22). Esto es importante porque es la época cuando es más común observar diferencias importantes entre la mortalidad observada y la real, por lo cual es una limitación importante del modelo, que no se logre estimar correctamente esta variable. La **¡Error! No se encuentra el origen de la referencia.** muestra cómo el nivel de error se reduce con la edad del ciclo para los estanques de época seca, lo que no sucede con los otros equipos.

Cuadro 22. Sesgo relativo (SR) y Error Estándar Relativo (EER) para las estimaciones de recuperabilidad por época climática.

Época	SR (%) ¹	EER (%) ²
Cálida	1 122,8	19,3
Fría	451,4	25,5
Transición	568,3	37,3

¹SR: Sesgo Relativo

²EER: Error Estándar Relativo

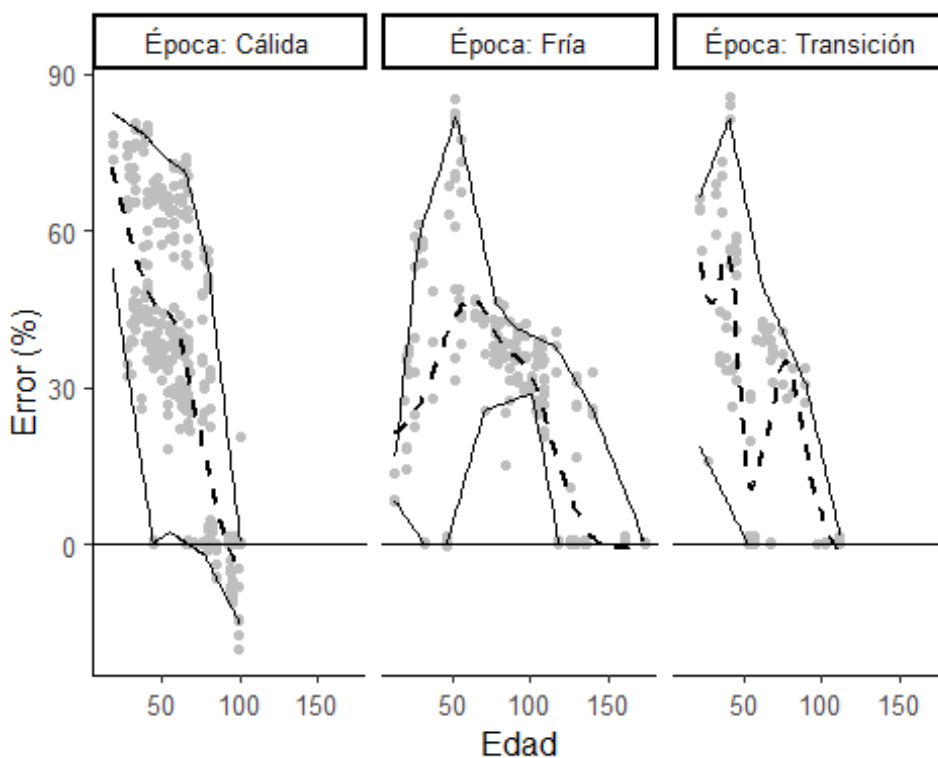


Figura 21. Error porcentual de la estimación de recuperabilidad por edad del cultivo y época climática del año.

La línea punteada es una estimación del promedio del error dada la edad y las líneas continuas representan los percentiles 5 y 95.

En el presente capítulo se han presentado los resultados de las estimaciones del modelo de captura y recaptura propuesto, obtenidas a partir de las simulaciones generadas. En el siguiente capítulo se presentarán las conclusiones y recomendaciones.

8 CONCLUSIONES Y RECOMENDACIONES

Después de analizar los principales resultados se procede en este capítulo a exponer las conclusiones propias para cada objetivo y se plantea algunas recomendaciones a la luz de estos datos.

En primer lugar, el presente trabajo logró simular con bastante realismo el comportamiento de estanques semiintensivos de producción de tilapia, tanto en términos de mortalidad como de crecimiento. Las metodologías usadas aquí podrían favorecer la experimentación *in silico* de este tipo de sistemas.

En segundo lugar, se generó un modelo bayesiano que parte de estimadores del modelo de Jolly-Seber para el análisis del tamaño poblacional y la sobrevivencia aplicado a poblaciones acuícolas, donde se conoce el tamaño inicial (con cierta precisión) y conteos parciales de mortalidad. Específicamente, se generaron dos propuestas de modelo, seleccionando la segunda -en la que se trabaja a partir de un tamaño de siembra constante- por ser más estable en sus estimaciones.

Por un lado, los estimadores de abundancia bayesianos obtenidos muestran propiedades aceptables en términos de sesgo y precisión, mejorando sustancialmente respecto a lo que se usa en la empresa. Por otro lado, los estimadores de sobrevivencia mostraron un sesgo más considerable, así como fallas para el ajuste cuando el valor de sobrevivencia era bajo. No obstante, las combinaciones de 4 ocasiones de muestreo con un tamaño de muestra de 150 peces por ocasión y 6 ocasiones de muestreo con 100 peces por ocasión muestran resultados muy buenos para estos parámetros de alto interés.

En tercer lugar, aunque no era parte de los objetivos del trabajo, se evaluó el comportamiento de un parámetro más que era necesario para el ajuste del modelo dada la información adicional proveniente de la mortalidad observada. Dicho parámetro, nombrado *recuperabilidad* ya que debería señalar la proporción de peces cuyos cadáveres se recuperan y se contabilizan, mostró valores de sesgo y precisión poco atractivos, de modo que se indica que no fue posible que

convergiere apropiadamente, lo cual lleva a pensar que la información adicional aportada por los conteos de mortalidad no fue suficiente para que su valor se actualizara adecuadamente en las cadenas de Markov-Monte Carlo. Empero, se nota que el nivel de error se disminuye al acercarse el momento de la cosecha, por lo que es posible obtener estimaciones de cosecha precisas los días antes de la cosecha.

Se notó, además, que el modelo puede ser “pesado” para correr en situaciones reales, pues incluso considerando que se redujo el tamaño poblacional (de unos 100.000 animales a solamente 1.000), tomó más de un mes de trabajo obtener los resultados a partir de las poblaciones simuladas. Esto es problemático pues su implementación puede requerir computadoras avanzadas para un sistema de acuicultura industrial.

Asimismo, se realizaron aportes adicionales en el estudio de las curvas de mortalidad observada para los estanques del sistema de producción que aportó los datos. Se muestra una importante relación con estudios elaborados en dicha empresa acerca del crecimiento. Es conocida la relación entre crecimiento y salubridad en la producción acuícola, así que estos resultados concuerdan con ese saber.

En cuanto a las recomendaciones, se sugiere trabajar en la eficiencia de este modelo, que, aunque arroja resultados prometedores, tiene un margen importante de mejora, para obtener estimaciones ágiles y precisas. Asimismo, podría incluirse otra información adicional provista por la empresa, a la cual no se tuvo acceso para este estudio. A saber:: valores de variables ambientales, de calidad del agua, alimentación, entre otras, de forma que se puedan obtener resultados con un mejor ajuste.

Este tesis aporta a la acuicultura industrial las bases para el establecimiento de modelos de captura-recaptura como procedimiento operativo para conocer el tamaño poblacional, que, combinados con otras técnicas, podrían ser beneficiosos

en el control de inventarios, medición de la mortalidad y sobrevivencia, entre otras variables.

Adicionalmente, si se logran modelos que brinden un mejor ajuste, pero es necesario que se evalúe la factibilidad técnica y económica de la implementación de metodologías de captura y recaptura, que podrían implicar una erogación económica importante, tanto por el hecho de necesitar marcas para identificación única de los peces, como de mayor mano de obra para el muestreo. Estos datos serían útiles de estimar, pero se salen del alcance de este trabajo.

9 LITERATURA CITADA

- Abadi, F., Botha, A., & Altwegg, R. (2013). Revisiting the Effect of Capture Heterogeneity on Survival Estimates in Capture-Mark-Recapture Studies: Does It Matter? *PLoS ONE*, 8(4), 20-22. <https://doi.org/10.1371/journal.pone.0062636>
- Akinrotimi, O. A., Abu, O. M. G., Ansa, E. J., Edun, O. M., & George, O. S. (2009). Hematological responses of *Tilapia guineensis* to acute stress. *International Journal of Natural and Applied Sciences*, 5(4), 338-343. Recuperado de <https://www.ajol.info/index.php/ijonas/article/view/76681>
- Arce-Vega, J. J. (2014). *Determinación de la tasa de alimentación para tilapia del Nilo (*Oreochromis niloticus*, Linnaeus, 1759) en condiciones comerciales corregida por el Coeficiente Térmico de Crecimiento* (Universidad de Costa Rica). Universidad de Costa Rica, San José, Costa Rica. Recuperado de <http://www.zootecnia.ucr.ac.cr/images/tesis/pdfs/arce-vega-jose-javier.pdf>
- Barboza Coto, J. (2008). *Práctica Profesional en el Departamento de Pesca Interna de AquaCorporación Internacional S.A.* Universidad de Costa Rica.
- Barcellos, L. J. G., Nicolaiewsky, S., de Souza, S. M. G., & Lulhier, F. (1999). The effects of stocking density and social interaction on acute stress response in Nile tilapia *Oreochromis niloticus* (L.) fingerlings. *Aquaculture Research*, 30(11-12), 887-892. <https://doi.org/10.1046/j.1365-2109.1999.00419.x>
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal*, 9(2), 378-400. <https://doi.org/10.32614/rj-2017-066>
- Buckland, S. T., & Garthwaite, P. H. (1991). Quantifying Precision of Mark-Recapture Estimates Using the Bootstrap and Related Methods. *Biometrics*, 47(1), 255-268. <https://doi.org/10.2307/2532510>
- Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of

- simulation studies in medical statistics. *Statistics in Medicine*, 25, 4267-4278. <https://doi.org/10.1002/sim.2673>
- Chao, A. (1989). Estimating Population Size for Sparse Data in Capture-Recapture Experiments. *Biometrics*, 45(2), 427-438.
- Chao, A., Pan, H. Y., & Chiang, S. C. (2008). The Petersen - Lincoln estimator and its extension to estimate the size of a shared population. *Biometrical Journal*, 50(6), 957-970. <https://doi.org/10.1002/bimj.200810482>
- Chao, A., Tsay, P. K., Lin, S.-H., Shau, W.-Y., & Chao, D.-Y. (2001). The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20, 3123-3157. <https://doi.org/10.1002/sim.996>
- Cooch, E. G., & White, G. C. (2019). *Program MARK: a gentle introduction 19th edition* (17th ed.; E. G. Cooch & G. C. White, Eds.). Fort Collins, CO: Colorado State University.
- Crespin, L., Choquet, R., Lima, M., Merritt, J., & Pradel, R. (2008). Is heterogeneity of catchability in capture-recapture studies a mere sampling artifact or a biologically relevant feature of the population? *Population Ecology*, 50(3), 247-256. <https://doi.org/10.1007/s10144-008-0090-8>
- Devineau, O., Choquet, R., & Lebreton, J. D. (2006). Planning Capture – Recapture Studies : Straightforward Precision , Bias , and Power Calculations. *Wildlife Society Bulletin*, 34(4), 1028-1035.
- El-Sayed, A. F. M. (2006). *Tilapia culture*. Wallingford, Reino Unido: CABI.
- FAO. (2018). *The State of World Fisheries and Aquaculture - Meeting the Sustainable Development Goals*. Roma: FAO. <https://doi.org/10.1093/japr/3.1.101>
- Fernandes, M. de O., & Volpato, G. L. (1993). Heterogeneous growth in the Nile tilapia: Social stress and carbohydrate metabolism. *Physiology and Behavior*, 54(2), 319-323. [https://doi.org/10.1016/0031-9384\(93\)90117-X](https://doi.org/10.1016/0031-9384(93)90117-X)

- Fletcher, D., Lebreton, J. D., Marescot, L., Schaub, M., Gimenez, O., Dawson, S., & Slooten, E. (2012). Bias in estimation of adult survival and asymptotic population growth rate caused by undetected capture heterogeneity. *Methods in Ecology and Evolution*, 3(1), 206-216. <https://doi.org/10.1111/j.2041-210X.2011.00137.x>
- Foo, J. T. W., & Lam, T. J. (1993). Serum cortisol response to handling stress and the effect of cortisol implantation on testosterone level in the tilapia, *Oreochromis mossambicus*. *Aquaculture*, 115(1-2), 145-158. [https://doi.org/10.1016/0044-8486\(93\)90365-6](https://doi.org/10.1016/0044-8486(93)90365-6)
- Foyosal, M. J., Alam, M., Momtaz, F., Chaklader, M. R., Siddik, M. A. B., Cole, A., ... Rahman, M. M. (2019). Dietary supplementation of garlic (*Allium sativum*) modulates gut microbiota and health status of tilapia (*Oreochromis niloticus*) against *Streptococcus iniae* infection. *Aquaculture Research*, 50(8), 2107-2116. <https://doi.org/10.1111/are.14088>
- FUNPROVER. (2008). *Manual de producción de tilapia con especificaciones de calidad e inocuidad*. Veracruz, México: Fundación Produce Veracruz.
- Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.*, 2(4), 1360-1383. <https://doi.org/10.1214/08-AOAS191>
- Gimenez, O., Cam, E., & Gaillard, J. M. (2018). Individual heterogeneity and capture–recapture models: what, why and how? *Oikos*, 127(5), 664-686. <https://doi.org/10.1111/oik.04532>
- Gosselin, J. L., & Anderson, J. J. (2020). Step-patterned survivorship curves: Mortality and loss of equilibrium responses to high temperature and food restriction in juvenile rainbow trout (*Oncorhynchus mykiss*). *PLoS ONE*, 15(5), e0233699. <https://doi.org/10.1371/journal.pone.0233699>
- Huggins, R., Stoklosa, J., Roach, C., & Yip, P. (2018). Estimating the size of an open population using sparse capture–recapture data. *Biometrics*, 74(1), 280-

288. <https://doi.org/10.1111/biom.12718>
- Jolly, G. M. (1965). Explicit estimates from capture-recapture data with death and immigration-stochastic model. *Biometrika*, 52(1), 225-247.
- Kendall, W. L., Pollock, K. H., & Brownie, C. (1995). A Likelihood-Based Approach to Capture-Recapture Estimation of Demographic Parameters under the Robust Design. *Biometrics*, 51(1), 293-308.
- Kéry, M., & Schaub, M. (2012). *Bayesian Population Analysis using WinBUGS*. Waltham, MA: Academic Press.
- Killian, H. S., Heikes, D., Wyk, P. Van, Masser, M., & Engle, C. R. (1998). Inventory Assessment Methods for Aquaculture Ponds. *Southern Regional Aquaculture Center Technical Publication Series*, (395), 1-4.
- Koenker, R. (2022). *quantreg: Quantile Regression*. Recuperado de <https://cran.r-project.org/package=quantreg>
- Krebs, C. J. (1999). *Ecological Methodology* (2nd ed.). Menlo Parck, CA: Addison Wesley Longman.
- Krebs, C. J., & Futuyma, D. J. (1973). Ecology. The Experimental Analysis of Distribution and Abundance. *The Quarterly Review of Biology*, 48(1), 47-48. <https://doi.org/10.1086/407551>
- Lazic, S. E., Clarke-Williams, C. J., & Munafò, M. R. (2018). What exactly is 'N' in cell culture and animal experiments? *PLOS Biology*, 16(4), e2005282. <https://doi.org/10.1371/JOURNAL.PBIO.2005282>
- Link, W. A., & Barker, R. J. (2005). Modeling Association among Demographic Parameters in Analysis of Open Population Capture – Recapture Data. *Biometrics*, 61, 46-54.
- Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: A comparison of natural ecosystems and aquaculture. *Journal of Fish Biology*, 49(4), 627-647.

<https://doi.org/10.1006/jfbi.1996.0192>

- Manly, B. F. J. (1984). Obtaining Confidence Limits on Parameters of the Jolly-Seber Model for Capture-Recapture Data. *Biometrics*, 40(3), 749-758.
- McDonald, T. L., & Amstrup, S. C. (2001). Estimation of Population Size Using Open Capture-Recapture. *Journal of Agricultural, Biological, and Environmental Statistics*, 6(2), 206-220.
<https://doi.org/10.1016/j.carres.2004.07.015>
- Múnera Miranda, L. G. (2013). *Revisión del nivel de cumplimiento de las buenas prácticas acuícolas en el departamento de producción de finca primaria Aquacorporación Internacional S.A., ubicada en Cañas, Guanacaste*. Universidad de Costa Rica.
- Oliveira, R. F., & Almada, V. C. (1996a). Dominance hierarchies and social structure in captive groups of the Mozambique tilapia *Oreochromis mossambicus* (Teleostei: Cichlidae). *Ethology Ecology and Evolution*, 8(1), 39-55. <https://doi.org/10.1080/08927014.1996.9522934>
- Oliveira, R. F., & Almada, V. C. (1996b). On the (In)stability of dominance hierarchies in the cichlid fish *Oreochromis mossambicus*. *Aggressive Behavior*, 22(1), 37-45. [https://doi.org/10.1002/\(SICI\)1098-2337\(1996\)22:1<37::AID-AB4>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1098-2337(1996)22:1<37::AID-AB4>3.0.CO;2-R)
- Otis, D. L., Burnham, K. P., White, G. C., & Anderson, D. R. (1978). Statistical Inference from Capture Data on Closed Animal Populations. *Wildlife Monographs*, p. 135. Logan, UT: The Wildlife Society.
<https://doi.org/10.2307/2287873>
- Peña, N., & Chacón, J. (2019). Acuicultura en Costa Rica. *World Aquaculture*, (July), 23-28. Recuperado de https://www.researchgate.net/publication/334726301_Acuicultura_en_Costa_Rica
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). *nlme*:

- Linear and Nonlinear Mixed Effects Models*. Vienna, Austria: R Foundation for Statistical Computing. Recuperado de <https://cran.r-project.org/package=nlme>
- Pledger, S., Pollock, K. H., & Norris, J. L. (2003). Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics*, *59*, 786-794. <https://doi.org/10.1111/j.1541-0420.2009.01361.x>
- Pledger, S., Pollock, K. H., & Norris, J. L. (2010). Open capture-recapture models with heterogeneity: II. Jolly-Seber model. *Biometrics*, *66*(3), 883-890. <https://doi.org/10.1111/j.1541-0420.2009.01361.x>
- Plummer, M. (2017). *JAGS Version 4.3.0 user manual*. Recuperado de https://people.stat.sc.edu/hansont/stat740/jags_user_manual.pdf
- Pollock, K. H. (1981). *A Capture-Recapture sampling design robust to unequal catchability*. Raleigh, NC: North Carolina State University.
- Pollock, K. H. (2000). Capture-Recapture Models. *Journal of the American Statistical Association*, *95*(449), 293-296.
- Pollock, K. H., Nichols, J. D., Brownie, C., & Hines, J. E. (1990). Statistical Inference for Capture-Recapture Experiments. *Wildlife Monographs*, p. 97. Raleigh, NC: The Wildlife Society. <https://doi.org/10.1002/jwmg.113>
- Popma, T., & Masser, M. (1999). *Tilapia: Life history and biology* (p. 4). p. 4. Stoneville, MS: Southern Regional Aquaculture Center.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Recuperado de <https://www.r-project.org/>
- Royle, A. J., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology*. Oxford, UK: Academic Press.
- Schwarz, C. J. (2001). The Jolly – Seber Model: More Than Just Abundance. *Journal of Agricultural, Biological, and Environmental Statistics*, *6*(2), 195-205.

- Seber, G. A. F., & Manly, B. F. J. (1985). Approximately Unbiased Variance Estimation for the Jolly-Seber Mark-Recapture Model: Population Size. En B. J. T. Morgan (Ed.), *Statistics in Ornithology* (pp. 363-371). Berlin, Alemania: Springer.
- Seber, G. A. F., & Schofield, M. R. (2019). *Capture-recapture: parameter estimation for open animal populations*. Cham, Switzerland: Springer. <https://doi.org/10.1007/978-3-030-18187-1>
- Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors. <https://doi.org/10.1214/16-STS576>, 32(1), 1-28. <https://doi.org/10.1214/16-STS576>
- Sukhavachana, S., Poopuang, S., Onming, S., & Luengnaruemitchai, A. (2019). Heritability estimates and selection response for resistance to *Streptococcus agalactiae* in red tilapia *Oreochromis* spp. *Aquaculture*, 502, 384-390. <https://doi.org/10.1016/j.aquaculture.2018.12.075>
- Sukhavachana, S., Tongyoo, P., Massault, C., McMillan, N., Luengnaruemitchai, A., & Poopuang, S. (2020). Genome-wide association study and genomic prediction for resistance against *Streptococcus agalactiae* in hybrid red tilapia (*Oreochromis* spp.). *Aquaculture*, 525, 735297. <https://doi.org/10.1016/j.aquaculture.2020.735297>
- Vijayan, M. M., Pereira, C., Grau, E. G., & Iwama, G. K. (1997). Metabolic responses associated with confinement stress in tilapia: The role of cortisol. *Comparative Biochemistry and Physiology - C Pharmacology Toxicology and Endocrinology*, 116(1), 89-95. [https://doi.org/10.1016/S0742-8413\(96\)00124-7](https://doi.org/10.1016/S0742-8413(96)00124-7)
- Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and Management of Animal Populations*. San Diego CA: Academic Press.
- Wilson, J. C., Nibbelink, N. P., & Peterson, D. L. (2009). Thermal tolerance

experiments help establish survival probabilities for tilapia, a group of potentially invasive aquatic species. *Freshwater Biology*, 54(8), 1642-1650. <https://doi.org/10.1111/j.1365-2427.2009.02214.x>

Wonmongkol, P., Sukhavachana, S., Ampolsak, K., Srisapoome, P., Suwanasopee, T., & Poompuang, S. (2018). Genetic parameters for resistance against *Flavobacterium columnare* in Nile tilapia *Oreochromis niloticus* (Linnaeus, 1758). *Journal of Fish Diseases*, 41(2), 321-328. <https://doi.org/10.1111/jfd.12728>

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R* (2.^a ed.). Chapman and Hall/CRC.

Yilmaz, E. (2019). Effects of dietary anthocyanin on innate immune parameters, gene expression responses, and ammonia resistance of Nile tilapia (*Oreochromis niloticus*). *Fish and Shellfish Immunology*, 93, 694-701. <https://doi.org/10.1016/j.fsi.2019.08.033>

Zuniga-Jara, S., & Goycolea-Homann, M. (2014). A bioeconomic model for red tilapia culture on the coast of Ecuador. *Aquaculture International*, 22, 339-359. <https://doi.org/10.1007/s10499-013-9643-8>

10 Anexo 1: Salidas a partir de datos con colas recortadas

Los siguientes resultados se obtuvieron después de eliminar 5% de los datos en cada cola. Los resultados se muestran como referencia.

10.1 Abundancia

Cuadro A1: Indicadores de sesgo y varianza de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%)	EER (%)
4	50	-2,1	1,0
	100	-0,9	0,8
	150	0,4	0,8
6	50	-2,3	0,8
	100	0,4	0,7
	150	1,5	0,7

SR: Sesgo Relativo

EER: Error Estándar Relativo

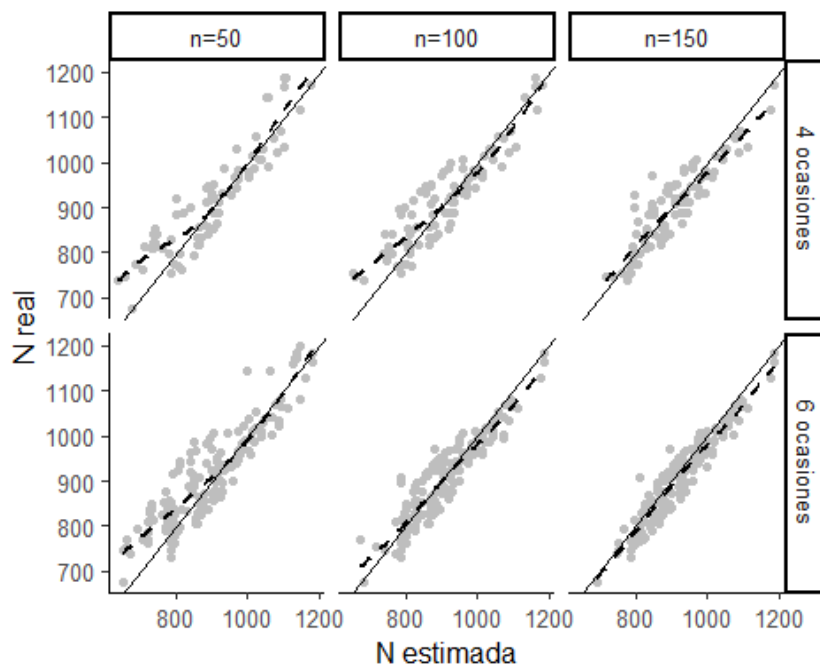


Figura A 1. Asociación entre la abundancia estimada y la abundancia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de las relación entre los valores estimados y los reales

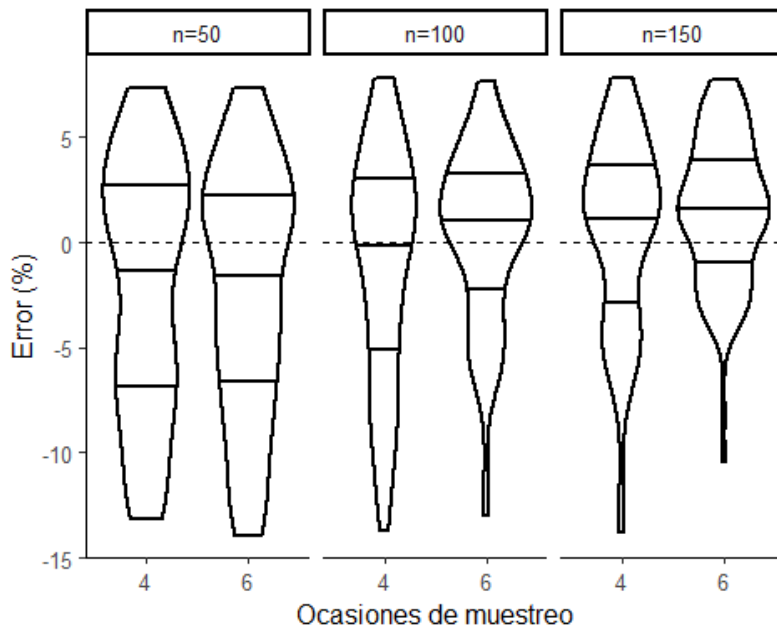


Figura A 2. Distribución del error porcentual en la estimación de abundancia por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se muestran los percentiles 25, 50 y 75

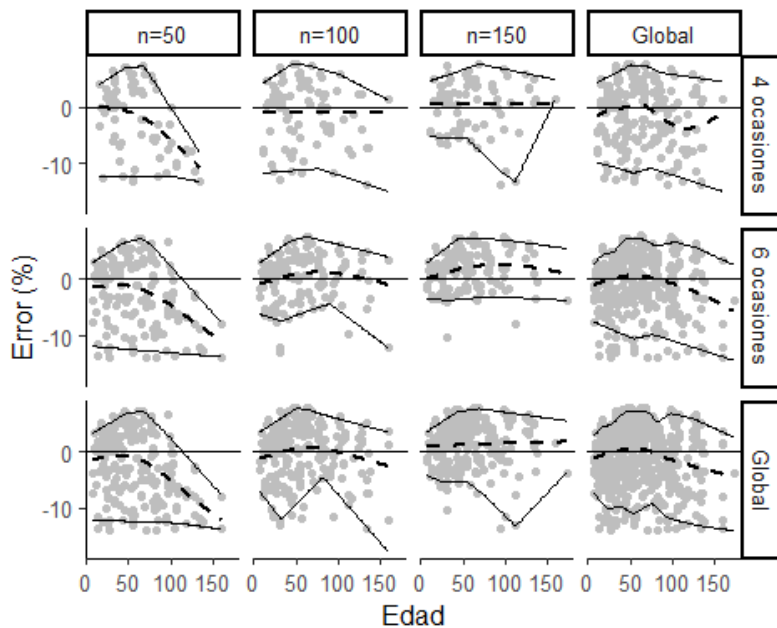


Figura A 3. Asociación entre la edad del estanque y el error porcentual en las estimaciones de abundancia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos

Cuadro A2: Indicadores de sesgo y varianza de los estimadores de abundancia según el la forma asumida para la relación entre mortalidad observada y mortalidad real

Forma	SR (%)	EER (%)
1	0,7	1,5
2	-0,2	1,5
3	-1,6	1,3

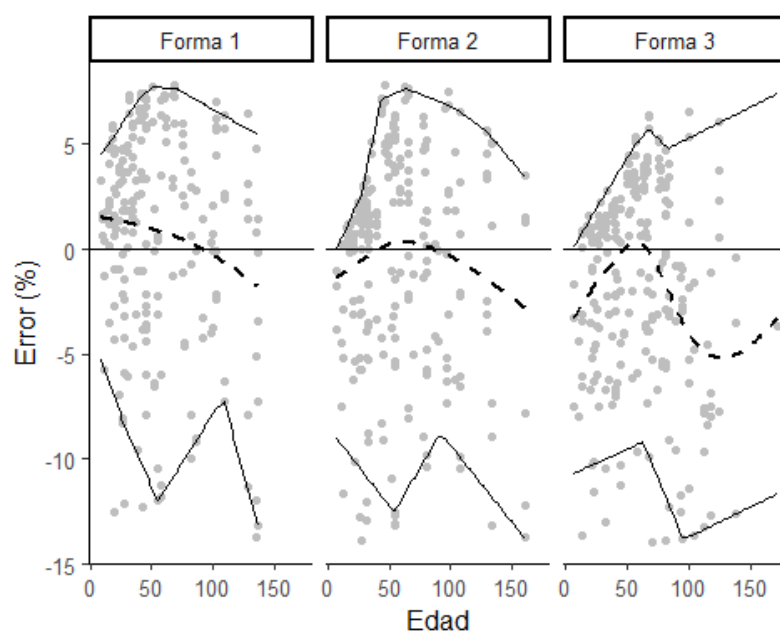


Figura A 4. Error porcentual por edad de la estimación y forma de la relación entre mortalidad real y mortalidad observada.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95.

Cuadro A3: Indicadores de sesgo y varianza de los estimadores de abundancia según la época climática de siembra

Epoca	SR (%)	EER (%)
Cálida	2,2	1,1
Fría	-3,5	1,5
Transición	-1,6	1,9

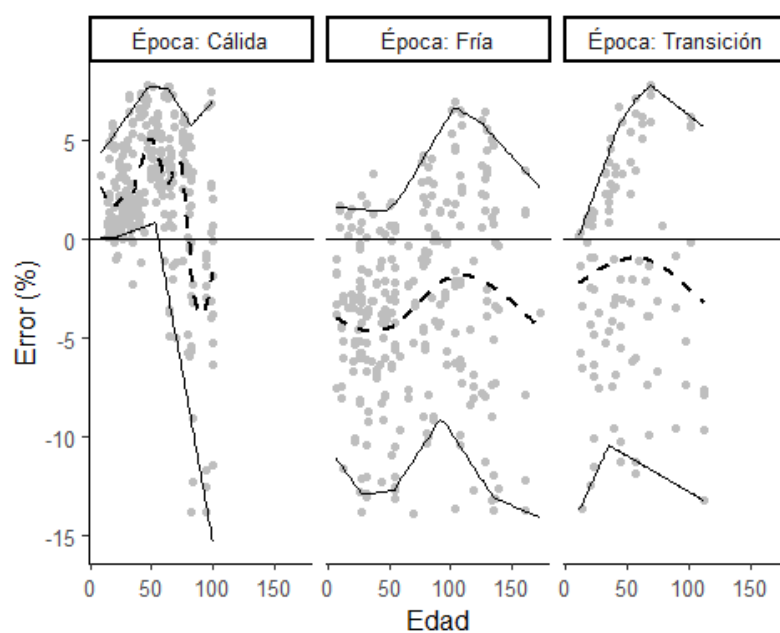


Figura A 5. Error porcentual por edad de la estimación y época de siembra.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95

10.2 Sobrevivencia

Cuadro A4: Indicadores de sesgo y varianza de los estimadores de abundancia según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%)	EER (%)
4	50	0,4	0,6
	100	0,8	0,6
	150	1,4	0,5
6	50	1,0	0,4
	100	1,2	0,4
	150	1,1	0,4

SR: Sesgo Relativo

EER: Error Estándar Relativo

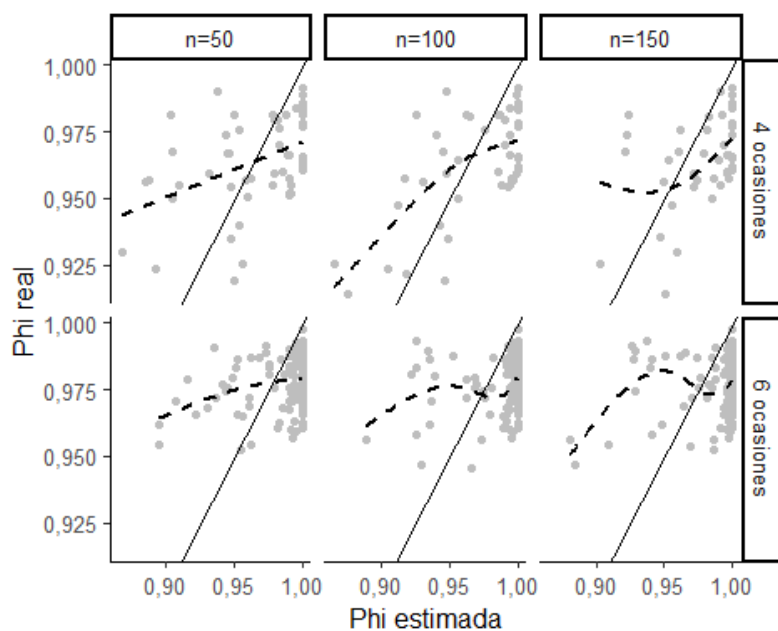


Figura A 6. Asociación entre la sobrevivencia estimada y la sobrevivencia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de la relación entre los valores estimados y los reales

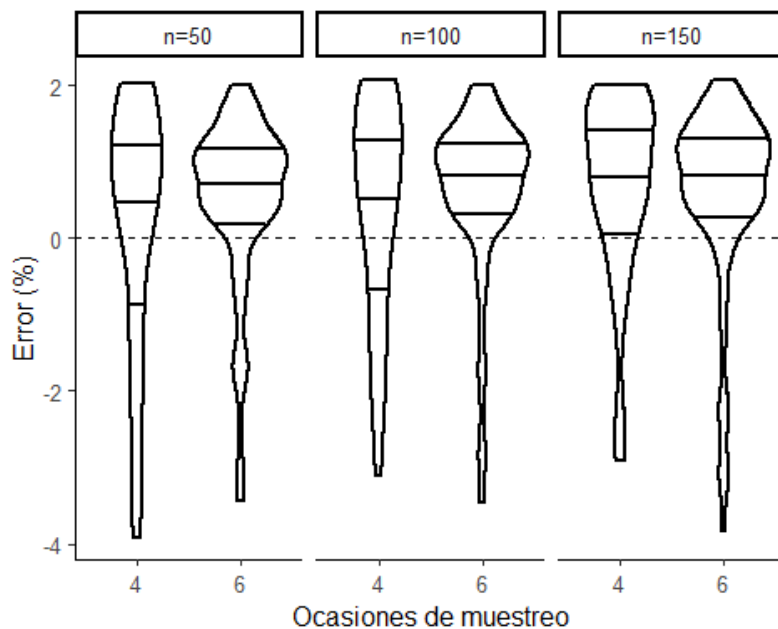


Figura A 7. Distribución del error porcentual en la estimación de sobrevivencia por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se muestran los percentiles 25, 50 y 75

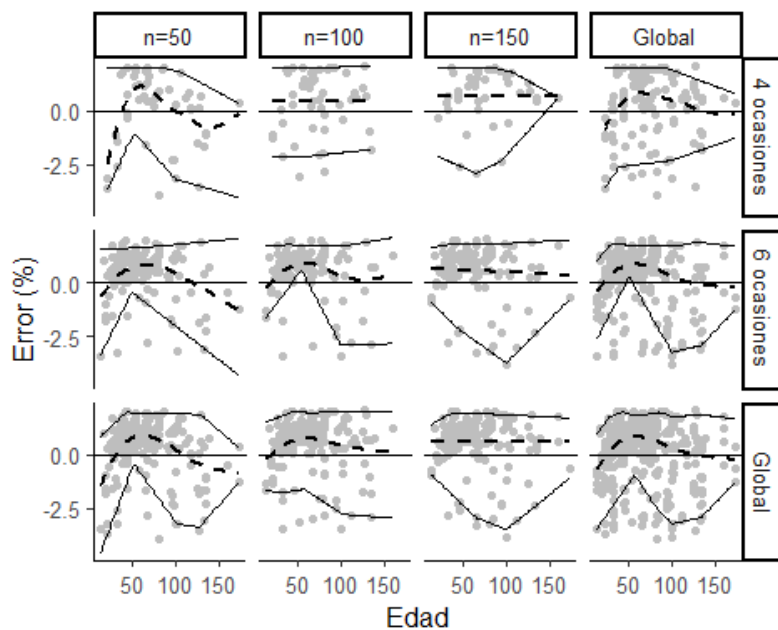


Figura A 8. Asociación entre la edad del estanque y el error porcentual en las estimaciones de sobrevivencia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos

Cuadro A5: Indicadores de sesgo y varianza de los estimadores de supervivencia según el la forma asumida para la relación entre mortalidad observada y mortalidad real

Forma	SR (%)	EER (%)
1	1,4	0,8
2	1,5	0,9
3	0,2	0,6

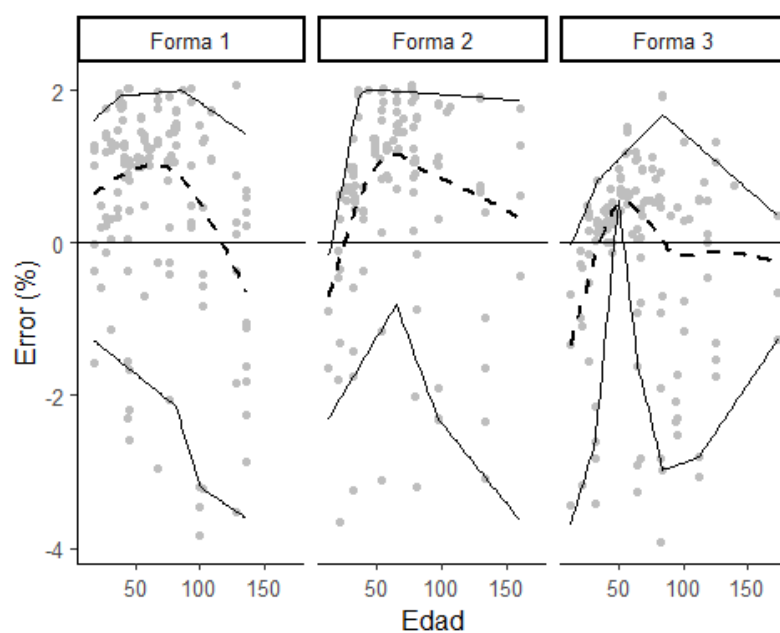


Figura A 9. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95

Cuadro A6: Indicadores de sesgo y varianza de los estimadores de sobrevivencia según la época climática de siembra

Epoca	SR (%)	EER (%)
Cálida	1,4	0,6
Fría	0,7	0,8
Transición	0,7	1,1

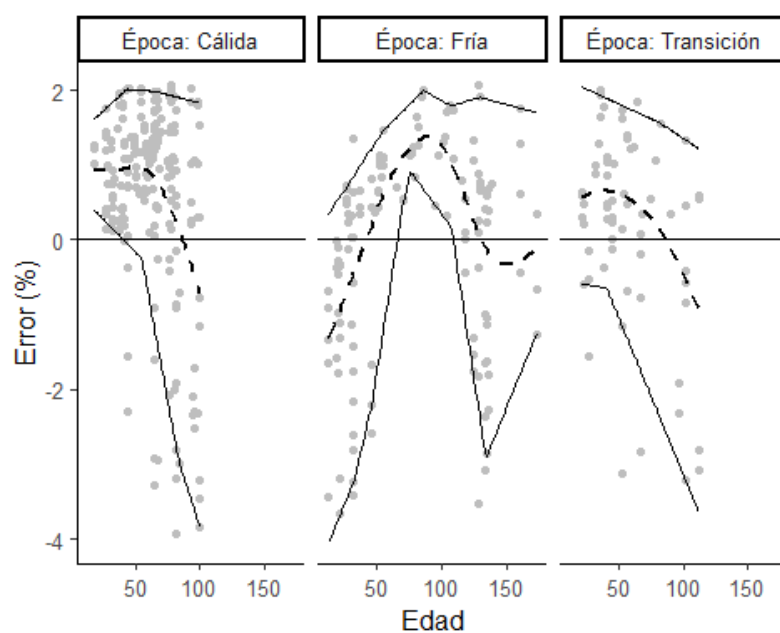


Figura A 10. Error porcentual por edad de la estimación de sobrevivencia y época de siembra.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95

10.3 Recuperabilidad

Cuadro A7: Indicadores de sesgo y varianza de los estimadores de recuperabilidad según el número de ocasiones de muestreo y el tamaño de la muestra

Ocasiones de muestreo	Tamaño de muestra	SR (%)	EER (%)
4	50	1 020,5	19,2
	100	993,6	18,0
	150	783,6	19,7
6	50	644,4	13,0
	100	637,0	12,1
	150	597,0	11,1

SR: Sesgo Relativo

EER: Error Estándar Relativo

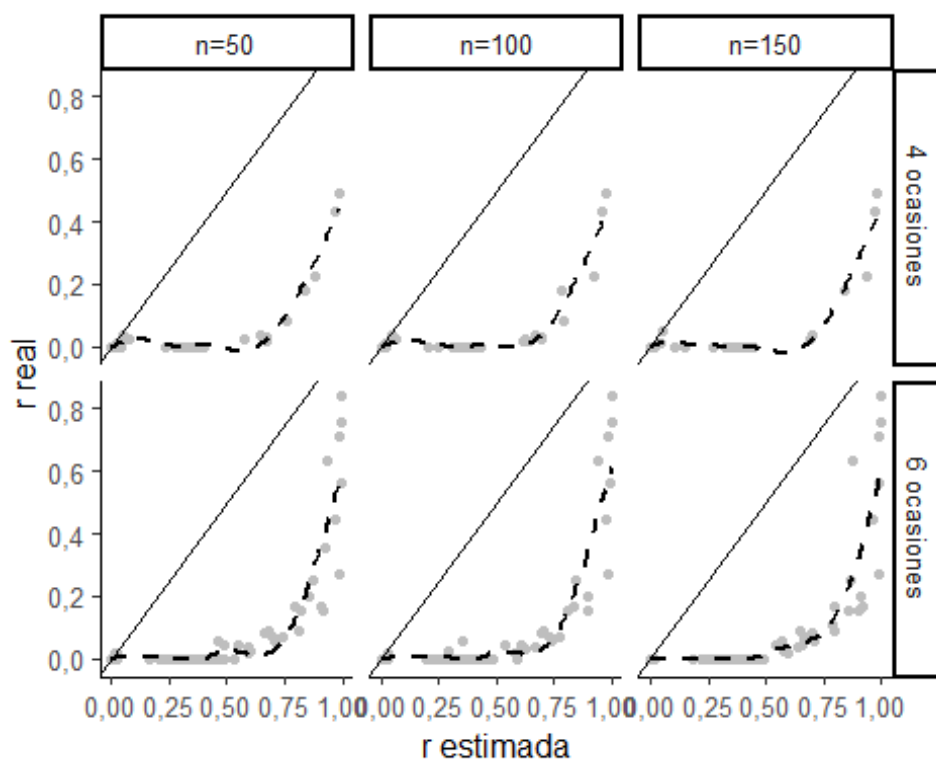


Figura A 11. Asociación entre la sobrevivencia estimada y la sobrevivencia real por número de ocasiones de muestreo y el tamaño de la muestra en cada ocasión.

La línea negra continua muestra la ecuación de $Y = X$ y la línea discontinua una estimación suavizada del promedio de la relación entre los valores estimados y los reales

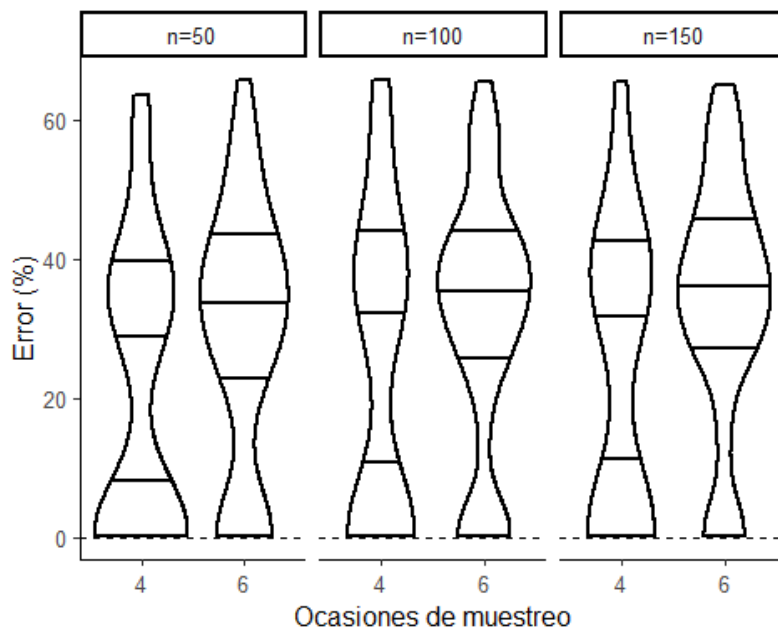


Figura A 12. Distribución del error porcentual en la estimación de sobrevivencia por tamaños de muestra y número de ocasiones de muestreo.

En cada violín se muestran los percentiles 25, 50 y 75

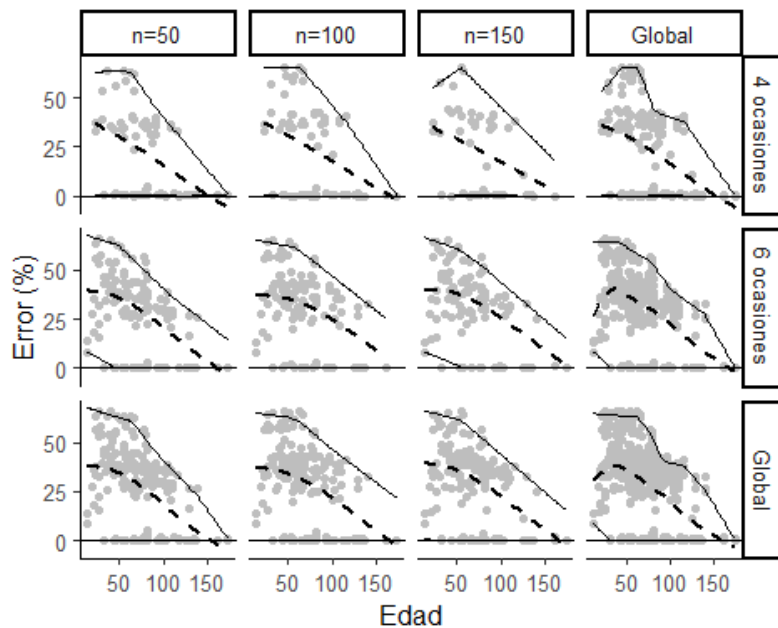


Figura A 13. Asociación entre la edad del estanque y el error porcentual en las estimaciones de abundancia, por número de ocasiones de muestreo y tamaño de muestra en cada ocasión.

La línea punteada muestra un modelo suavizado para la media y las líneas continuas indican los percentiles 5 y 95, para indicar la variabilidad de los datos

Cuadro A8: Indicadores de sesgo y varianza de los estimadores de sobrevivencia según el la forma asumida para la relación entre mortalidad observada y mortalidad real y la presencia de pérdida puntual

forma	SR (%)	EER (%)
1	1 084,7	23,1
2	745,9	23,7
3	504,8	20,4

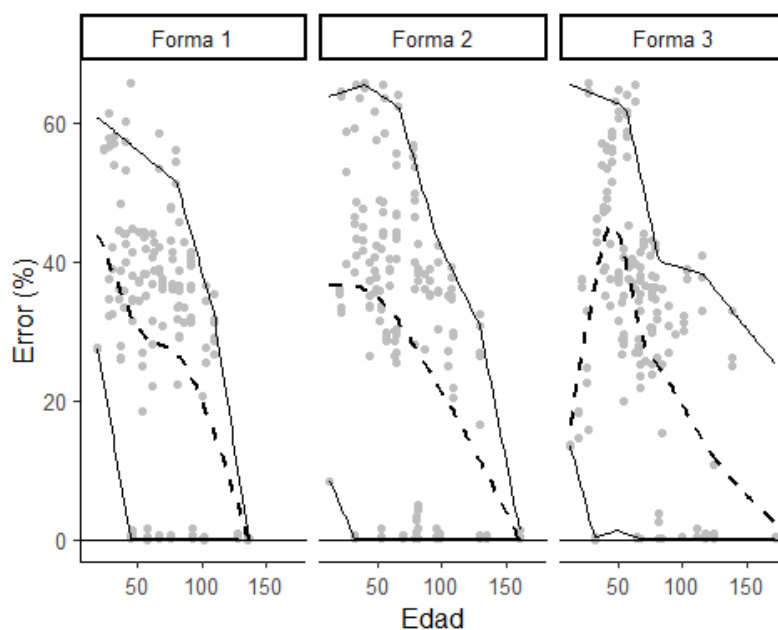


Figura A 14. Error porcentual por edad de la estimación, forma de la relación entre mortalidad real y mortalidad observada y presencia o ausencia de pérdida.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95

Cuadro A9: Indicadores de sesgo y varianza de los estimadores de recuperabilidad según la época climática de siembra

Epoca	SR (%)	EER (%)
Cálida	1 974,0	16,2
Fría	410,3	23,3
Transición	491,4	35,9

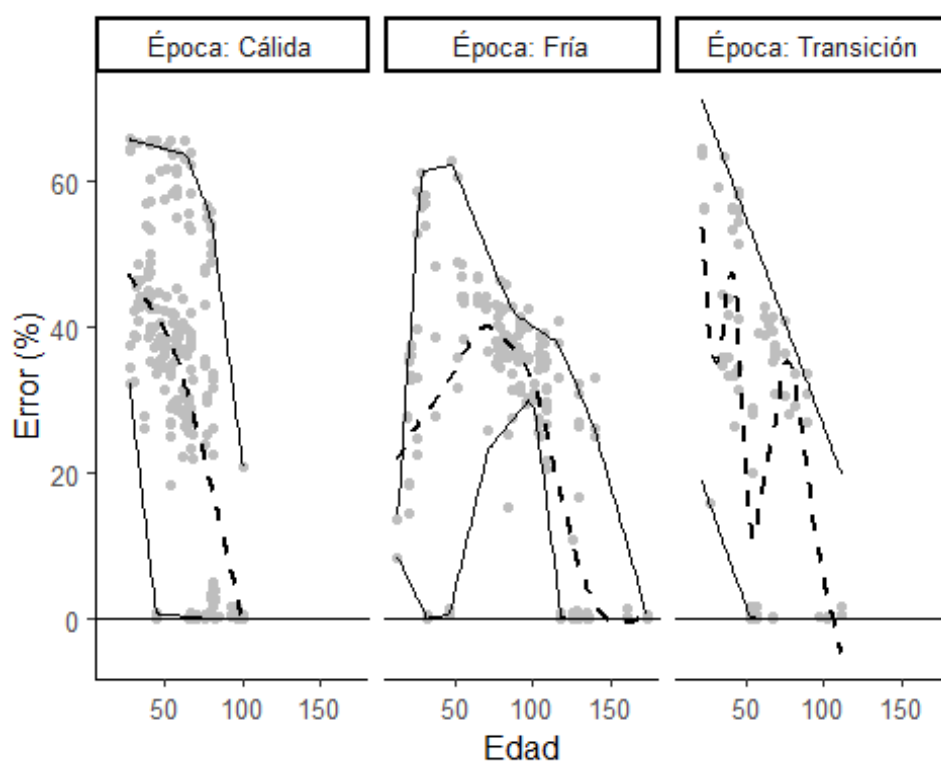


Figura A 15. Error porcentual por edad de la estimación de recuperabilidad y época de siembra.

La línea punteada muestra una estimación suavizada del promedio, mientras que las líneas continuas los percentiles 5 y 95