Análisis exploratorio de datos de la componente biológica de *Donax trunculus* y *Chamelea gallina*

Proyecto RECLAM

García, A.; Mardones, M; Silva, L; Delgado, M.

09 April, 2025

Table of Contents

# 1 CONTEXTO

El proyecto RECLAM tiene entre sus principales objetivos establecer la estructura poblacional y la abundancia de las poblaciones de *Donax trunculus* (coquina) y *Chamelea gallina* (chirla) en el Golfo de Cádiz y Valencia. Estas especies representan un recurso importante para la pesca artesanal y semiindustrial en la costa sur peninsular, siendo esenciales para la sostenibilidad económica de comunidades costeras ([Rico, 2023](#ref-RECLAM2023); [Silva et al., 2019](#ref-Silva2019)).

Este objetivo se articula en base a la tarea 1 (“Parámetros poblacionales de la coquina y chirla”) y subtareas 1.1 (“Muestreos dirigidos a la estimación de la densidad y la estructura poblacional”), 1.2 (“Estructura poblacional”) y 1.3 (“Estimación del estado del stock”), de las que es responsable el Centro Oceanográfico de Cádiz (IEO-CSIC; COCAD). En el presente informe se reporta el desarrollo de actividades realizadas hasta la fecha, y relacionadas con las subtareas **1.1** y **1.2**.

# 2 MATERIALES Y METODOS

La dinámica poblacional de estos bivalvos puede verse afectada por múltiples factores, incluyendo la presión pesquera, la variabilidad ambiental y la estacionalidad reproductiva. En este contexto, la generación de información estandarizada a través de muestreos sistemáticos es clave para comprender la estructura de las poblaciones y orientar estrategias de gestión basadas en evidencia. Los principales objetivos de estas tareas en el marco del Proyecto RECLAM son:

* Realizar muestreos mensuales en los principales caladeros de coquina.
* Registrar la información biológica y pesquera de manera estandarizada.
* Generar una base de datos estructurada y reutilizable para análisis poblacionales.
* Identificar patrones espacio-temporales y estructura de tallas de ambas especies.
* Evaluar la composición poblacional y el reclutamiento estacional.

Los muestreos se realizan con una frecuencia mensual, cubriendo las principales áreas de pesca previamente identificadas.

## 2.1 Zona de Estudio

La Figura 2.1 muestra la zona de estudio en la costa sur de España, donde se lleva a cabo el monitoreo de las pesquerías de Chirla y Coquina. El área comprende una franja costera entre los 36.6º y 37.4º de latitud norte, con puntos de muestreo distribuidos desde el Golfo de Cádiz hasta las cercanías de Huelva. Se identifican múltiples polígonos de seguimiento, cada uno etiquetado con un código (por ejemplo, AND101 a AND112 y AND601 a AND108), que representan unidades espaciales de monitoreo donde se recopila información biológica y pesquera. Estos polígonos incluyen tanto zonas someras cercanas a la costa como áreas más abiertas, lo que permite cubrir el rango de hábitats utilizados por ambas especies. (la distribución de los puntos de monitoreo lo haré pronto)

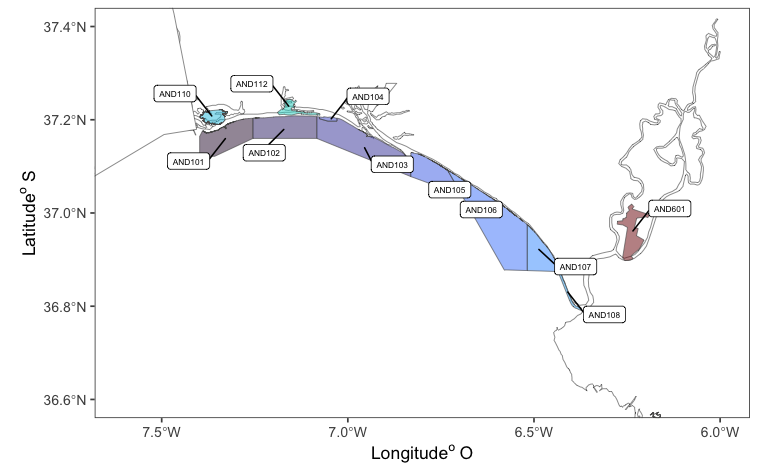


Figure 2.1: Distribución espacial de los polígonos de muestreo utilizados en el monitoreo de las pesquerías de Chirla y Coquina en la costa suroccidental de España. Cada polígono (etiquetado con su correspondiente código AND) representa una unidad de marisqueo definido por la Junta de Andalucía.

## 2.2 Estrategia de muestreo

### 2.2.1 Muestreo dirigido a la estimación de la estructura poblacional y densidad de *D. trunculus*

En el Golfo de Cádiz se han establecido tres estaciones de muestreo entre los ríos Guadalquivir y Guadiana, en concreto, en el Parque Nacional de Doñana. En el Golfo de Valencia, se han establecido cinco estaciones de muestreo en Oliva-Denia, Gandia, Tabernes, Cullera y Valencia. Estas estaciones se visitan con periodicidad mensual para la toma de muestras comerciales y poblacionales.

Para la recogida de muestras poblacionales, se han establecido tres transectos equidistantes paralelos a la costa (réplicas) en cada estación. En cada transecto, los observadores científicos realizaron arrastres de 25 m de largo paralelo a la costa georreferenciado mediante un GPS para calcular el área muestreada (). La recogida de muestras comerciales se realiza solo sobre un transecto. Los arrastres se realizan utilizando dos rastros a pie, uno similar a los que utilizan los mariscadores locales de bivalvos (rastro comercial) y otro modificado que permite la captura de individuos de pequeño tamaño (rastro poblacional). Estos rastros tienen una estructura de hierro con una apertura de 44,5 cm de ancho que excava profundamente en el sedimento (hasta los 15 cm), utilizando un tamaño de malla experimental (poblacional: 3 x 3 mm) para permitir la recolección de los individuos más pequeños y juveniles, y un tamaño de malla comercial (comercial: 7 x 7 mm). La captura (organismos) retenida en cada rastro se almacena en la playa y se transporta al laboratorio. Todos los organismos se identifican siguiendo la nomenclatura WORMS (www.marinespecies.com) y literatura especializada en taxones específicos, se cuentan y pesan. Se obtuienen datos de densidad (ind/) y biomasa (g/).

### 2.2.2 Muestreo dirigido a la estimación de la estructura poblacional y densidad de de *C. gallina*

En el caso de C. gallina, el muestreo se llevó a cabo utilizando una draga hidráulica comercial en el Golfo de Cádiz. Se definieron tres transectos equidistantes paralelos a la costa, con estaciones a 5 y 8 metros de profundidad durante la marea baja, totalizando 6 estaciones. Se realizaron tres réplicas en cada estación cada mes. En cada estación, la draga se desplegó y se arrastró paralelamente a la costa durante 10 minutos. Cada transecto también fue georreferenciado mediante un GPS para calcular el área muestreada (). En el mar Mediterráneo, los muestreadores de la Universitat Politècnica de València ya estaban realizando muestreos similares, ajustando la profundidad a la geomorfología del Golfo de Valencia, que es muy diferente a la del Golfo de Cádiz. En el laboratorio, se registró el número de individuos retenidos de las especies objetivo para estimar la densidad poblacional (ind/) y la biomasa (g/) a partir de las muestras experimentales, y el rendimiento (kg por tiempo de pesca) a partir de las muestras comerciales. La longitud de la concha (SL) se midió con un calibrador digital Vernier hasta el 0,1 mm para producir distribuciones de frecuencias de longitud de todas las muestras. Se realizó un análisis exploratorio de las distribuciones de frecuencia de longitud de la valva mediante histogramas y estadística descriptiva por muestra y tipo de arte de pesca. Se estimaron parámetros poblacionales mediante modelos lineales y se evaluaron patrones espacio-temporales de variables clave, tales como la longitud media, abundancia relativa y proporción de adultos y juveniles, integrando herramientas gráficas y analíticas para caracterizar la dinámica poblacional de la especie.

## 2.3 Estandarización de Datos

Todos los datos fueron registrados en formularios digitales y posteriormente almacenados en una base de datos relacional en formato .csv y .rds, diseñada para facilitar análisis poblacionales, cálculo de estructura de tallas y estimación de indicadores. Una base llamada “Base\_RECLAM\_Tallas.rds”, con datos de tallas y la otra, llamada “Base\_RECLAM\_lp.rds” con los datos de la relación Longitud-Peso. Cabe señalar que esta última base de datos no tiene la dimensión espacial, dado que es un análisis global.

El objetivo actual es coordinar que el muestreo de Valencia tenga el mismo formato para luego unificar las bases. Todos los análisis fueron realizados con las librerías “tidyverse” y “easystats” ([Lüdecke et al., 2022](#ref-Ludecke2022); [Wickham et al., 2019](#ref-Wickham2019)).

# 3 RESULTADOS

## 3.1 Datos obtenidos

A continuación, se presenta de forma resumida la tipología del dato obtenido durante los primeros meses de muestreo. En la Figura 3.1, se presentan el número de individuos por especie, estación de muestreo y réplica, para cada mes de muestreo. Estos datos también se muestran desglosados en la Tabla 1 agregados por estación.

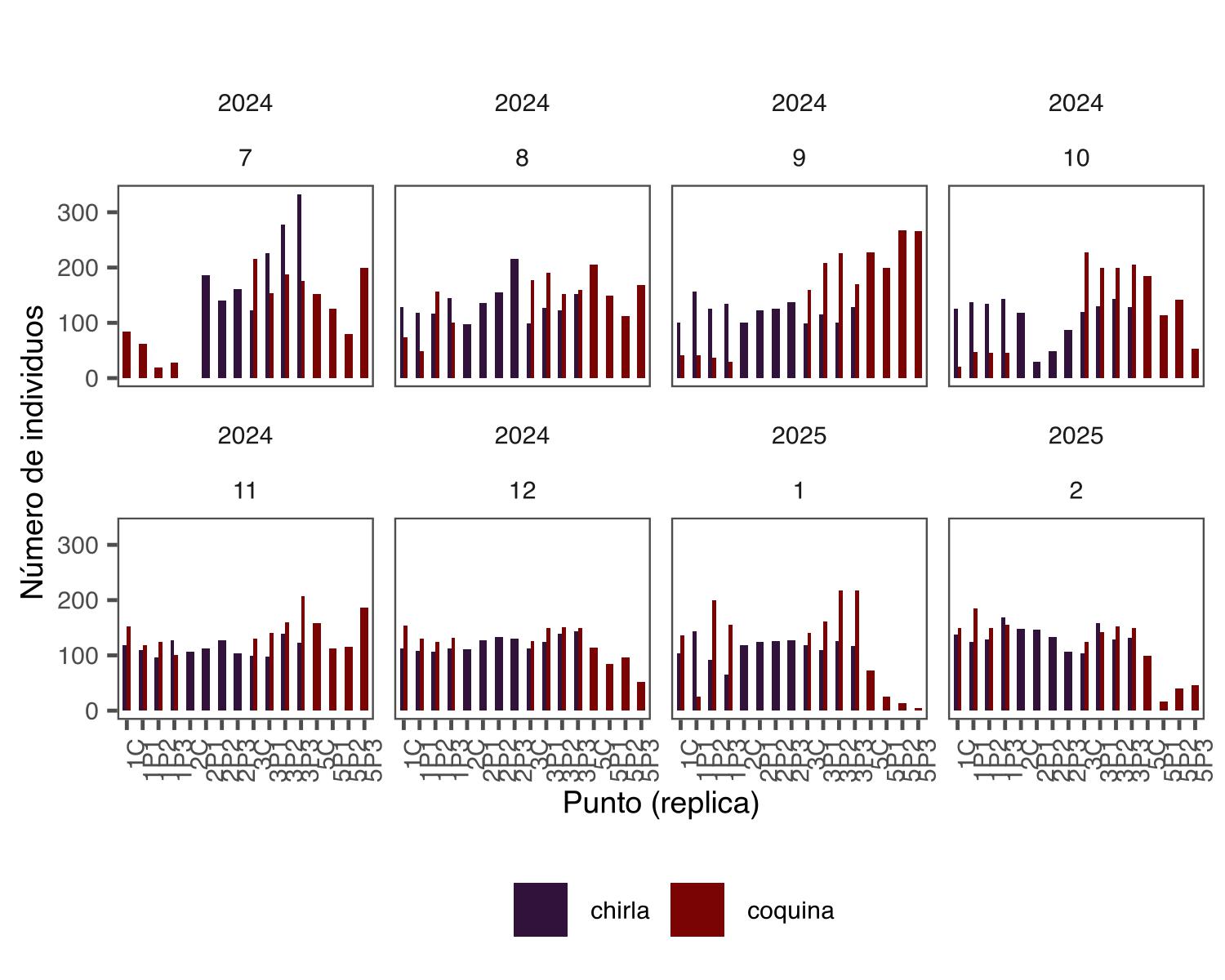


Figure 3.1: Registros totales por especie, estación de muestreo y réplica, para cada mes de muestreo

Tabla 1 y 2 muestra el resumen de los individuos muestreados por especie, identificando, periodo y punto de muestreo.

Table 3.1: Número de individuos muestreados de Chirla por mes y punto de muestreo

| **ANO** | **MES** | **especie** | **PUNTO** | **N\_individuos** |
| --- | --- | --- | --- | --- |
| 2,025 | 1 | chirla | 1 | 406 |
| 2,025 | 1 | chirla | 2 | 496 |
| 2,025 | 1 | chirla | 3 | 470 |
| 2,025 | 2 | chirla | 1 | 559 |
| 2,025 | 2 | chirla | 2 | 533 |
| 2,025 | 2 | chirla | 3 | 522 |
| 2,024 | 7 | chirla | 2 | 487 |
| 2,024 | 7 | chirla | 3 | 958 |
| 2,024 | 8 | chirla | 1 | 508 |
| 2,024 | 8 | chirla | 2 | 603 |
| 2,024 | 8 | chirla | 3 | 501 |
| 2,024 | 9 | chirla | 1 | 516 |
| 2,024 | 9 | chirla | 2 | 486 |
| 2,024 | 9 | chirla | 3 | 443 |
| 2,024 | 10 | chirla | 1 | 540 |
| 2,024 | 10 | chirla | 2 | 283 |
| 2,024 | 10 | chirla | 3 | 521 |
| 2,024 | 11 | chirla | 1 | 451 |
| 2,024 | 11 | chirla | 2 | 450 |
| 2,024 | 11 | chirla | 3 | 460 |
| 2,024 | 12 | chirla | 1 | 441 |
| 2,024 | 12 | chirla | 2 | 502 |
| 2,024 | 12 | chirla | 3 | 519 |

Table 3.2: Número de individuos muestreados de Ccoquina por mes y punto de muestreo

| **ANO** | **MES** | **especie** | **PUNTO** | **N\_individuos** |
| --- | --- | --- | --- | --- |
| 2,025 | 1 | coquina | 1 | 516 |
| 2,025 | 1 | coquina | 3 | 737 |
| 2,025 | 1 | coquina | 5 | 115 |
| 2,025 | 2 | coquina | 1 | 640 |
| 2,025 | 2 | coquina | 3 | 570 |
| 2,025 | 2 | coquina | 5 | 203 |
| 2,024 | 7 | coquina | 1 | 193 |
| 2,024 | 7 | coquina | 3 | 732 |
| 2,024 | 7 | coquina | 5 | 557 |
| 2,024 | 8 | coquina | 1 | 380 |
| 2,024 | 8 | coquina | 3 | 678 |
| 2,024 | 8 | coquina | 5 | 635 |
| 2,024 | 9 | coquina | 1 | 149 |
| 2,024 | 9 | coquina | 3 | 763 |
| 2,024 | 9 | coquina | 5 | 960 |
| 2,024 | 10 | coquina | 1 | 160 |
| 2,024 | 10 | coquina | 3 | 831 |
| 2,024 | 10 | coquina | 5 | 492 |
| 2,024 | 11 | coquina | 1 | 495 |
| 2,024 | 11 | coquina | 3 | 637 |
| 2,024 | 11 | coquina | 5 | 573 |
| 2,024 | 12 | coquina | 1 | 541 |
| 2,024 | 12 | coquina | 3 | 577 |
| 2,024 | 12 | coquina | 5 | 348 |

La Figura 3.2 muestra histogramas de frecuencia de tallas mensuales para Chirla (panel superior, en negro) y Coquina (panel inferior, en rojo) durante los años 2024 y 2025. En Chirla, las distribuciones del segundo semestre de 2024 presentan modas bien definidas entre 18 y 22 mm, mientras que en enero y febrero de 2025 aparece una nueva moda cercana a los 10 mm, lo que indica un posible evento de reclutamiento. En Coquina, se observan modas múltiples a lo largo del año, con una clara concentración de individuos pequeños (~10 mm) entre julio y septiembre de 2024, que se desplaza hacia tallas mayores en los meses siguientes, reflejando el crecimiento de esa cohorte. A comienzos de 2025 también se observa la aparición de una nueva moda de tallas pequeñas, compatible con un nuevo reclutamiento. En ambas especies, las variaciones en la forma y posición de las modas a lo largo del tiempo permiten identificar patrones estacionales de reclutamiento y seguimiento del desarrollo de las cohortes.

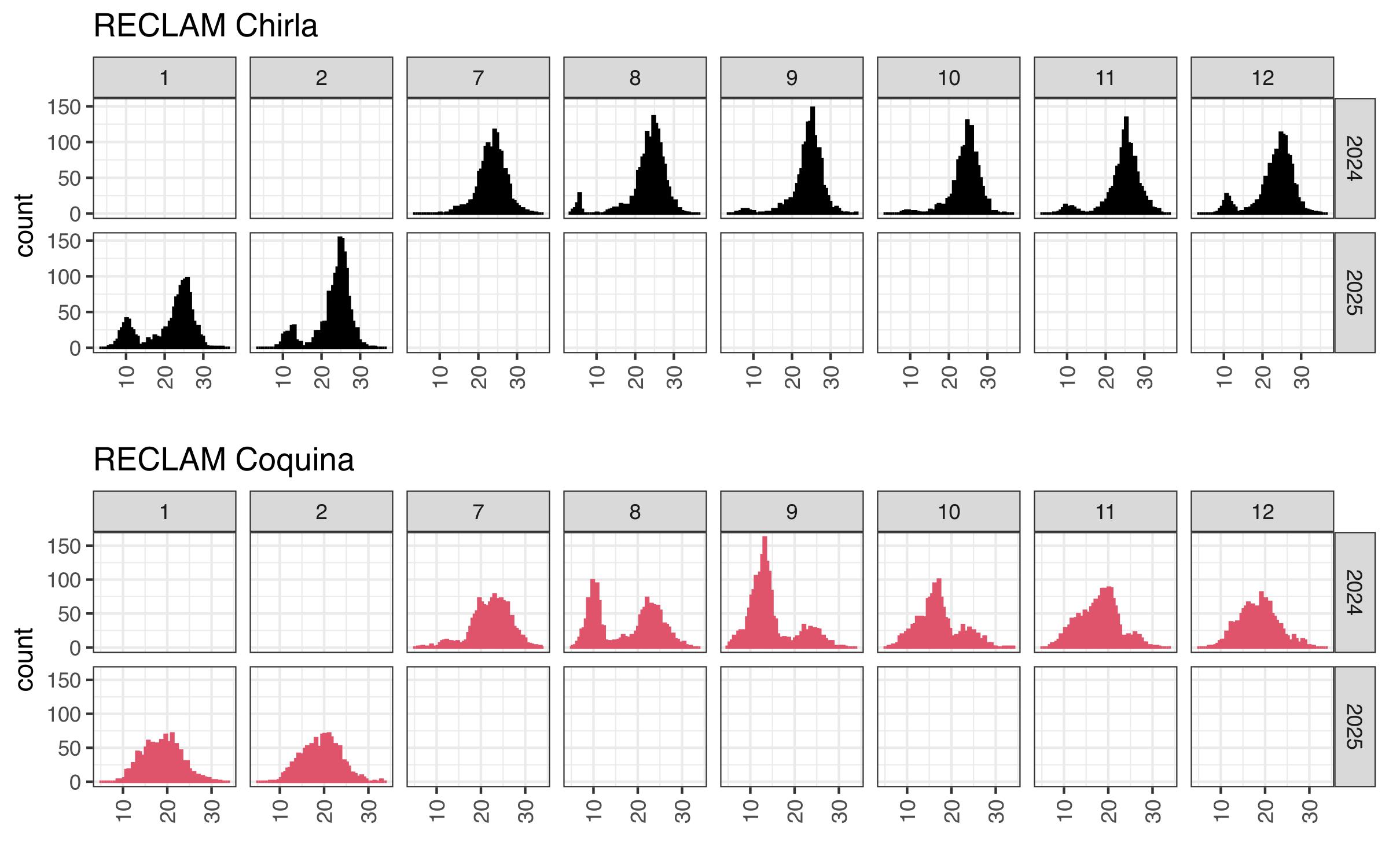


Figure 3.2: Estructura de tallas de Chamelea gallina (chirla) y Donax trunculus (coquina) en los muestreos mensuales.

## 3.2 Estructura poblacional

### 3.2.1 Relaciones Alométricas

Generar relaciones Alómetricas permitirá estimar el peso esperado de cada especie en función de su longitud, facilitando el cálculo de biomasa como paso previo al desarrollo de evaluaciones y otros análisis comparativos espacio-temporales y entre poblaciones. Los datos utilizados en estos análisis preliminares se corresponden con los obtenidos en el mes de Junio 2024. Durante el mes de Marzo del 2025 se ha realizado otro muestreo de relación talla - peso, pero sus datos aún no están sistematizados. La s relaciones de longitud peso para ambas especies realizada durante un primer muestreo se ilustran en la Figura 3.3

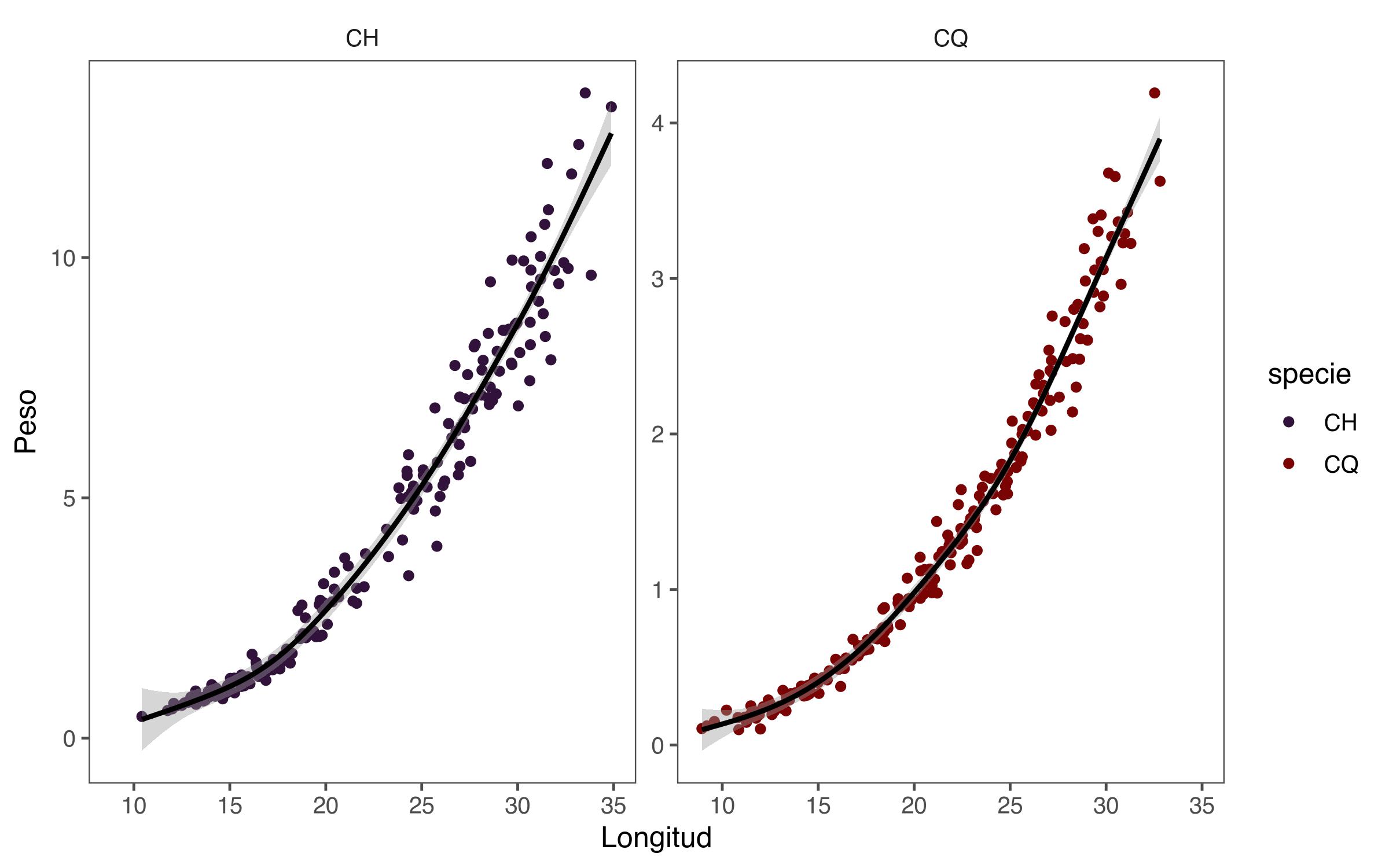
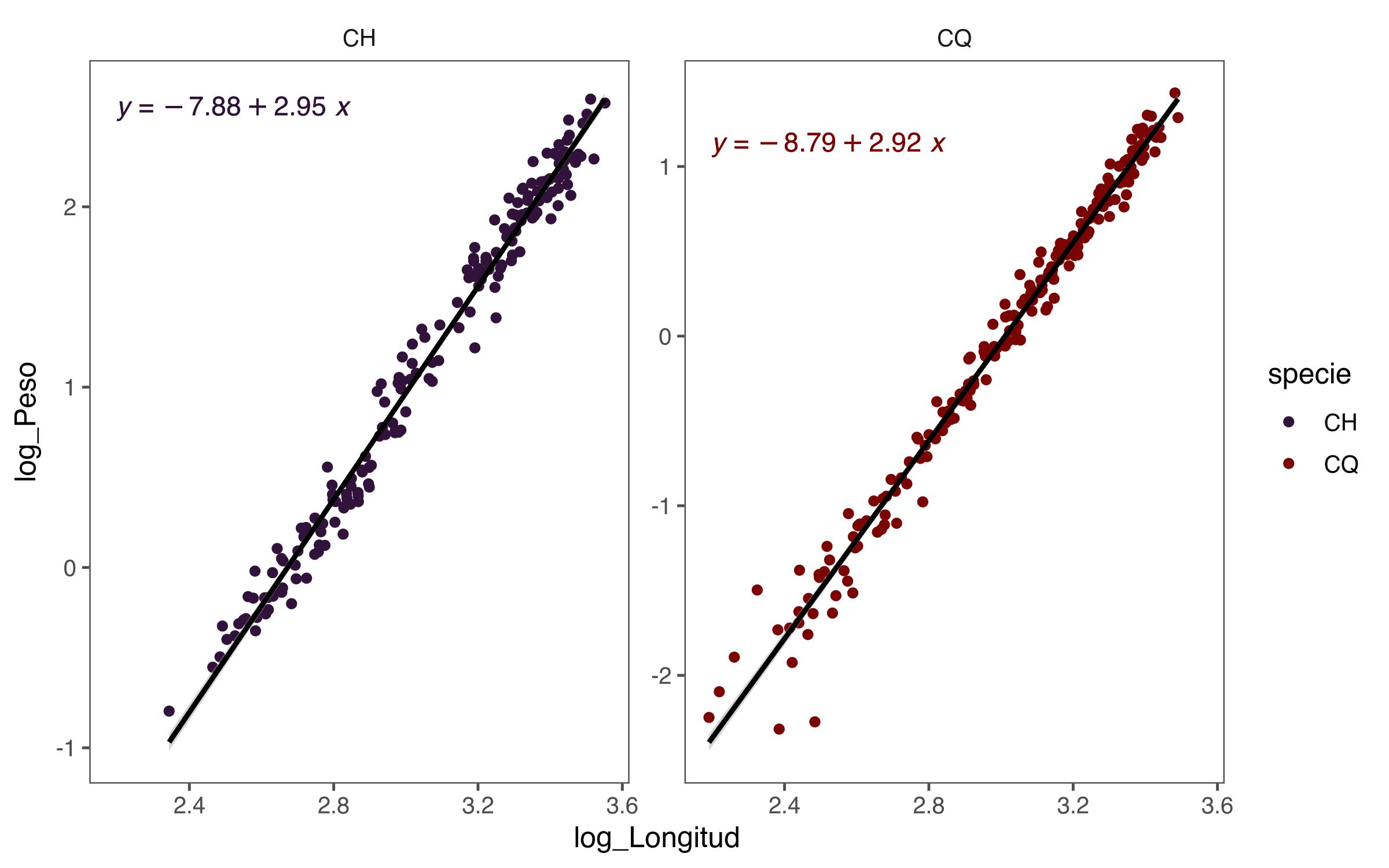


Figure 3.3: Relaciones talla-peso para *Chamelea gallina* y *Donax trunculus* (junio 2024)

Se presenta la relación alométrica entre el logaritmo del peso (log\_Peso) y el logaritmo de la longitud (log\_Longitud) para las especies **Chirla** y **Coquina**.



La transformación logarítmica se aplicó sobre la ecuación clásica del modelo alométrico , lo que permite linealizar la relación mediante la expresión:

En el caso de la Chirla, la ecuación ajustada fue:

Mientras que para la Coquina, la relación estimada fue:

Ambas ecuaciones muestran una pendiente cercana a 3, lo cual sugiere un patrón de crecimiento **isométrico**, es decir, que el peso corporal aumenta proporcionalmente al volumen, sin grandes cambios en la forma o densidad a lo largo del crecimiento.

La dispersión de los datos es baja en ambas especies, y la línea de tendencia se ajusta de forma precisa al patrón observado, indicando un ajuste robusto del modelo. Aunque los interceptos difieren ligeramente, con valores de -6.98 en Chirla y -6.78 en Coquina, esta diferencia sugiere que, a longitudes pequeñas, la Coquina podría tener un peso ligeramente superior al de la Chirla, aunque la diferencia es marginal.

Volviendo a la escala original, se pueden expresar las ecuaciones como:

* Para la Chirla:
* Para la Coquina:

Estas funciones permiten estimar el peso esperado de cada especie en función de su longitud, facilitando el cálculo de biomasa, conversiones talla-peso, y análisis comparativos entre poblaciones.

### 3.2.2 Índice de Reclutamiento para coquina y chirla

se calculará un índice de reclutamiento simple en las especies de coquina y chirla, utilizamos un proceso de estandarización relativa de los datos de tallas menores a 15 mm. La transformación se aplica a la variable longitud (mm), que representa la proporción de individuos bajo talla, transformada mediante el logaritmo de la proporción. El cálculo del índice de reclutamiento sigue los siguientes pasos:

1. Se definen los valores mínimos y máximos para el rango objetivo de normalización. En este caso, los límites de normalización son:
   * Límite inferior:
   * Límite superior:
2. Se calculan el valor mínimo y máximo de la longitud de la valva de los cuales se quiere obtener el índice de reclutamiento:
   * *min\_x*: el valor mínimo de la variable.
   * *max\_x*: el valor máximo de la variable.
3. Se aplica la siguiente fórmula de normalización para transformar los datos y obtener el índice de reclutamiento:

* En esta fórmula:
  + *Prop Length* es el valor de la proporción de individuos bajo la talla.
  + *min\_x* y *max\_x* son los valores mínimo y máximo de *Prop Length*, respectivamente.
  + *a* y *b* son los límites inferior y superior del rango de normalización, establecidos en -1 y 1, respectivamente.

Este índice permite identificar problemas en el proceso de reclutamiento, como condiciones ambientales adversas, sobrepesca o destrucción de hábitats, que podrían dificultar la regeneración de la población. En términos de gestión pesquera, el índice es esencial para ajustar las cuotas de captura y establecer medidas de protección, como tamaños mínimos de captura o áreas de veda, para garantizar la sostenibilidad de las poblaciones. También facilita la comprensión de la dinámica poblacional y la predicción de tendencias futuras.

### 3.2.3 Estimación de Parámetros:

Para el análisis de la distribución de frecuencias de tallas se empleará la versión modificada de la función de crecimiento de Von Bertalanffy (VBGF) que incorpora oscilación estacional en el crecimiento, y se ajustará utilizando el método ELEFAN para análisis de datos de frecuencia de longitud ([Pauly, 1980](#ref-Pauly1980a)) a travéss de la aplicación del paquete de R ([R Core Team, 2024](#ref-R2024)) llamado “TropfishR” ([Mildenberger et al., 2017](#ref-Mildenberger2017)). Los índices compuestos de growth performance, como phi prime, permiten realizar comparaciones inter e intraespecíficas (entre poblaciones). Este índice de phi prime se calculará utilizando las expresiones de Pauly & Morgan ([1987](#ref-Pauly1987)) y Munro y Pauly (1983). El método Bahttacharya se aplicará para el análisis de cohorte. Estos análisis se llevarán a cabo una vez se cuente con mínimo un año de seguimiento mensual de las poblaciones.

### 3.2.4 Densidad poblacional

La densidad poblacional se calcula como el número de individuos por unidad de área, en este caso, individuos por metro cuadrado (ind/). Para obtener esta densidad, se debe considerar el número de organismos capturados durante los arrastres, los cuales se realizan a lo largo de transectos georreferenciados para ambas especies. La densidad se calcula dividiendo el número total de individuos capturados en cada arrastre entre el área muestreada. El área muestreada depende de la longitud del arrastre y el ancho del rastro (ya sea manual o draga hidráulica) utilizado. Entonces, el área de cada arrastre es multiplicado por el ancho del arraste, y la densidad se obtiene dividiendo el número de individuos por el área de muestreo en metros cuadrados.

Para realizar un análisis comparativo temporal y espacial de la densidad poblacional de chirla y coquina, se pueden utilizar tanto técnicas paramétricas como no paramétricas, dependiendo de la distribución de los datos. En el análisis temporal, si los datos siguen una distribución normal, se puede aplicar un análisis de varianza (ANOVA) ([Kaufmann & Schering, 2014](#ref-ANOVA2014)) o una prueba t para comparar las medias de densidad en diferentes momentos. Si los datos no son normales, se puede utilizar la prueba de Kruskal-Wallis o la prueba de Wilcoxon para comparar medianas. En el análisis espacial, para comparar las densidades entre estaciones, se puede usar ANOVA o Kruskal-Wallis, dependiendo de la distribución de los datos.

### 3.2.5 Evaluación de stock

Por ahora se pondrá el método que utiliza Luis. Cuando Luis y Jorge tengan preparadas las bases de datos podrá Mauricio aplicar nuevas metodologías. Mientras no nos cojamos los dedos, y pongamos solo la que se ha aplicado por ahora.

Marina. Recuerda que tambien esta realizdaa la evaluacion de stock de Coquina con el modelo integrado. Podemos mostrar variables relativas de salida. Si bien no se hizo con los datos del RECLAM, pasa lo mismo con el modelo data poor que usa Luis, dado que también usaría datos históricos.

### 3.2.6 Producción anual secundaria de *Chamelea gallina*.

Está prevista la determinación de la producción anual total (junio de 2024-mayo de 2024) en la zona submareal. Se calculará mediante el método de la tasa de crecimiento específica de biomasa (Brey, 2001). Este método requerirá datos sobre la relación SL-AFDW, distribuciones de frecuencia de longitud de valva de muestras agrupadas y parámetros VBGF. Para ello se recogerán muestras adicionales de 50 individuos (tomados al azar) en el mes de mayo de 2025 para los análisis de biomasa, mediante la determinación del peso seco de tejidos blandos y valvas libres de ceniza (AFDW).

## 3.3 Análisis Estadístico de las Diferencias entre Poblaciones de Valencia y Cádiz

Una vez obtenidas las muestras de las poblaciones de Valencia, se procederá con un análisis estadístico exhaustivo para evaluar la existencia de diferencias significativas entre las poblaciones de ambas especies en función de diversas variables. Para ello, se aplicarán pruebas estadísticas de comparación de pendientes y *ANCOVA* (Análisis de Covarianza ([Cook & Campbell, 1979](#ref-cook1979)) con el objetivo de identificar posibles diferencias en tallas, parámetros y patrones de crecimiento.

* Prueba de Homogeneidad de Pendientes:

Para evaluar si las poblaciones presentan la misma respuesta a las variables explicativas, se realizará una prueba de homogeneidad de pendientes. Este análisis permite verificar si la pendiente de la relación entre la variable dependiente (por ejemplo, biomasa, densidad, etc.) y la variable independiente (por ejemplo, tamaño, profundidad, etc.) es homogénea a través de las distintas poblaciones. Si se detectan diferencias significativas en las pendientes, esto indicaría que las poblaciones responden de manera diferente a los factores explicativos y justificaría el uso de modelos separados para cada población.

* Análisis de Covarianza (ANCOVA):

Para investigar las diferencias entre las poblaciones, se aplicará un ANCOVA. Este modelo estadístico permite evaluar las diferencias en las medias de la variable dependiente entre las poblaciones, ajustando por posibles efectos de covariables que puedan influir en la variable dependiente (por ejemplo, la edad o el tamaño de los individuos). El ANCOVA no solo compara las medias, sino que también tiene en cuenta la variabilidad de las covariables, lo que ofrece una estimación más precisa de las diferencias poblacionales.

El modelo de ANCOVA se especificará como sigue:

Donde: - es el valor observado de la variable dependiente (por ejemplo, densidad o biomasa) para el individuo en la población , - es la media general, - es el efecto específico de la población , - es el coeficiente de la covariable (por ejemplo, tamaño o edad), - es la covariable (por ejemplo, tamaño o profundidad), - es el error aleatorio.

Los resultados de estas pruebas proporcionarán información clave sobre las diferencias en la dinámica poblacional entre las poblaciones de Valencia y Cádiz, lo que permitirá inferir si existen factores ecológicos, biológicos o ambientales que afectan de manera diferenciada a cada población. Dependiendo de los resultados, se podrán realizar recomendaciones sobre la gestión de las poblaciones y la conservación de los recursos marinos en la región.

# REFERENCIAS

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field setting*. Rand McNally.

Kaufmann, J., & Schering, A. (2014). Analysis of variance ANOVA. In *Wiley StatsRef: Statistics reference online*. John Wiley & Sons, Ltd. https://doi.org/<https://doi.org/10.1002/9781118445112.stat06938>

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Wiernik, B. M., Bacher, E., Thériault, R., & Makowski, D. (2022). Easystats: Framework for easy statistical modeling, visualization, and reporting. *CRAN*. <https://doi.org/10.32614/CRAN.package.easystats>

Mildenberger, T. K., Taylor, M. H., & Wolff, M. (2017). TropFishR: an R package for fisheries analysis with length-frequency data. *Methods in Ecology and Evolution*, *8*(11), 1520–1527. <https://doi.org/10.1111/2041-210X.12791>

Pauly, D. (1980). On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *Journal Cons. Int. Explor. Mer*, *39*(2), 175–192. <https://doi.org/10.1093/icesjms/39.2.175>

Pauly, D., & Morgan, G. R. (1987). *Length-based methods in fisheries research* (Vol. 13, p. 468).

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Rico, C. (2023). *MEMORIA CIENTÍFICO-TÉCNICA Plan Complementario de Ciencias Marinas y del Plan de Recuperación , Transformación y Resiliencia .* (pp. 1–20).

Silva, L., Delgado, M., Cojan, M., Rodriguez-Rua, A., Román, S., Blanco, R., García, E., Terrón, A., & Collado, J. (2019, September). Research and management as key drivers in the evolution of the fishery of striped venus (chamelea gallina) in the spanish waters of the gulf of cádiz. *XX Iberian Symposium on Marine Biology Studies (SIEBM)*.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>