

Caracterización de las cargas parasitarias de *C. rogercresseyi*
en la región de Aysén a través de Kriging Ordinario y distancia
euclíadiana y camino más corto



Francisco Vargas

Tutor: Jonathan Acosta

2024-03-13

Tabla de contenidos

Agradecimientos	4
RESUMEN	5
SUMMARY	6
I INTRODUCCIÓN Y METODOLOGÍA	7
Objetivos	8
Objetivo General	8
Objetivos Específicos	8
1 MARCO TEÓRICO	9
1.1 REGIÓN DE AYSÉN	9
1.2 SALMONICULTURA	11
1.3 <i>Caligus rogercresseyi</i>	11
1.3.1 GENERALIDADES	11
1.3.2 CICLO DE VIDA DEL PARÁSITO	12
1.3.3 TRATAMIENTOS	14
1.4 METODOLOGÍAS	15
1.4.1 MÉTODOS PREVIOS	15
1.4.2 MÉTODO GEOESTADÍSTICO PROPUESTO	18
1.4.3 DISTANCIAS	20
2 MATERIALES Y MÉTODOS	22
2.1 Datos	22
2.2 Distancias	23
2.3 Generalidades del Kriging Ordinario	23
2.4 Semivariogramas	25
II RESULTADOS	27
3 Análisis exploratorio	28

4 Kriging Ordinario	34
4.1 Máximos	34
4.2 Variogramas	35
4.3 Interpolación via Kriging	37
5 Conclusiones	42
6 Trabajos futuros	44
7 Apéndice	45
Referencias	53

Agradecimientos

En primer lugar, agradezco a mi pareja Pamela García y su familia: Ana Muñoz, Luz Muñoz y Marina Véliz por brindarme apoyo incondicional para cumplir mis objetivos tanto personales como académicos. Ellos me han acogido con brazos abiertos en este complejo período y me han apoyado en perseguir y continuar con mis estudios. Me han brindado soporte económico y emocional para poder seguir adelante. Además, quiero agradecer a mi hija Isabel Vargas, que en los pocos días desde que llegó al mundo me ha traído mucho amor y alegría constante.

En segundo lugar, quiero agradecer a mi tutor Jonathan Acosta por el apoyo en el desarrollo del presente trabajo durante el magíster, desde el envío de sus clases de estadística espacial hasta las reuniones frecuentes y la explicación de varios conceptos, los cuales me permitieron familiarizarme y desarrollarme en una de las disciplinas que desde el pre-grado de Medicina Veterinaria he considerado muy interesante.

Además, agradezco a todos los docentes de la Pontificia Universidad Católica que me enseñaron durante el transcurso del Magíster: Pedro Ramos, Manuel Galea, Alejandro Jara, Ana María Araneda, Ricardo Aravena, Ricardo Olea, Joshua Kunst, y María José García por su tiempo y dedicación a la docencia.

Agradezco también a los ayudantes: Josefa Silva, Jesús Achire, Julio Ávila, Nixon Jerez, Bladimir Morales y Alonso Campos por las ayudantías realizadas y por su disposición a resolver todo tipo de dudas y consultas.

Agradezco a todos mis compañeros por el buen ambiente que se generó durante el programa, tanto a nivel de colaboración como en la comunicación entre los distintos grupos.

Finalmente, agradezco a la Pontificia Universidad Católica por la oportunidad de haber realizado el magíster y por mantener un alto estándar de calidad y exigencia a nivel de docentes y estudiantes.

RESUMEN

El *Caligus rogercresseyi* es el ectoparásito que afecta más frecuentemente a la salmonicultura chilena. Se caracteriza por presentarse en el ambiente marino, en general en ambientes con salinidad sobre 26 ppm; presenta un ciclo con dos fases: una fase planktonica y una fase parásita, en donde el copepodito es el estadio infestante. Este parásito causa tanto pérdidas directas como indirectas, dado a los costos de tratamiento, pérdida de la conversión de alimento y pérdida de calidad en la planta de proceso, entre otras cosas.

El Servicio Nacional de Pesca y Acuicultura monitorea semanalmente los centros de cultivo activos a través del Programa Sanitario Específico de Vigilancia y Control de Caligus (PSEVC-Caligus), el cual establece que los centros deben realizar muestreos de caligus semanales (o mensuales, según período/especie/zona) y enviar estos muestreos al Servicio. SERNAPESCA además establece que los centros con promedio de hembras ovígeras mayores a 3 posterior a una ventana de tratamiento caen en categoría de Centro de Alta Diseminación (CAD), donde la normativa permite hacer una cosecha parcial de hasta el 50% de la biomasa si el centro es considerado CAD tres veces seguidas o del 25% si el centro es considerado CAD 3 veces en 6 evaluaciones.

La data utilizada corresponde al monitoreo de los centros de cultivo a través del programa de vigilancia de Caligidosis entre los años 2015 y 2022, y las distancias utilizadas fue una comparación entre el camino más corto (shortest path) y la distancia euclíadiana, aplicando distintos modelos de Semivariograma: Matérn, Gaussiano, Esférico y Cauchy al promedio de los máximos por ciclos para 158 centros de cultivo. Las predicciones fueron realizadas a través de la interpolación considerando Kriging Lognormal, dada la distribución asimétrica de los máximos.

Las diferencias observadas se encuentran principalmente en la varianza entre las distancias, donde los modelos utilizando la distancia euclíadiana presentó consistentemente mayor varianza que los modelos estimados utilizando la distancia shortest path.

SUMMARY

Caligus rogercresseyi is the ectoparasite that most often affects the chilean salmon farming. It's characterized by its presence in marine environments at over 26 ppm of salinity; it has a cycle with two main phases: a planktonic phase and a parasitic phase, where the infesting stage is the copepodid. This parasite causes both direct and indirect losses, due to increased treatment costs, loss of feed conversion and loss of quality at the processing plant, among other factors.

The National Aquaculture and Fisheries Service (SERNAPESCA) monitors *C. rogercresseyi* loads weekly through its Specific Vigilance and Control Sanitary Programme for Caligus (PSEVC - Caligus), which establishes that sites must do weekly parasite samplings (or monthly, in some specific time periods/species/regions) and they must send those samplings to the National Service. SERNAPESCA also establishes that any site with over three average ovigerous females on the site after a treatment window fall under the category of High Dissemination Site (CAD), where the law allows the government to demand up to 50% of biomass harvesting if the site was CAD 3 times in a row, or 25% if the site fell into the CAD category 3 times out of 6 evaluations.

The data used was gathered from the overall monitoring of sites on the Specific Vigilance and Control Sanitary Programme for Caligus during the years 2015 and 2022. The distances correspond to a comparison between the shortest-path distance, and euclidian distance, using different Semivariogram models: Matérn, Gaussian, Spherical and Cauchy on the average of the maximums per site, on 158 sites in total. Predictions were done using interpolation via Lognormal Kriging, due to the asymmetric distribution of the maxima.

The main differences were between the variances of each distance, where models using euclidian distance consistently showed a higher variance than the models estimated using the shortest-path.

Parte I

INTRODUCCIÓN Y METODOLOGÍA

Objetivos

Objetivo General

Evaluar el desempeño y validez de las distancias del camino más corto (shortest-path) con respecto a la distancia Euclidiana a través del Kriging Ordinario aplicado en datos reales de la carga total de adultos de *Caligus rogercresseyi* en la Región de Aysén durante el período 2015 - 2022.

Objetivos Específicos

1. Evaluar los modelos de covarianza generadas a partir de los distintos variogramas estimados utilizando la distancia obtenida a través del camino más corto y la distancia euclidiana.
2. Comparar las predicciones generadas por el Kriging Ordinario y los distintos modelos entre sus respectivas métricas de distancia.
3. Generar mapas en distintas utilizando distintos modelos de variograma que permitan caracterizar el riesgo a través de la covarianza a nivel global y local entre centros.
4. Caracterizar la carga promedio por ciclo de adultos totales de *C. rogercresseyi* a través de mapas generados por la interpolación del Kriging Ordinario, para evaluar el grado de dependencia espacial del piojo del salmón.

1 MARCO TEÓRICO

1.1 REGIÓN DE AYSÉN

A nivel oceanográfico, la Región de Aysén se caracteriza por ser un ecosistema sub-antártico con un complejo sistema de fiordos, canales, golfos, estuarios y bahías [1].

En estos sistemas de fiordos se producen complejos sistemas de gradientes de salinidad horizontales y verticales debido a la afluencia de agua dulce de múltiples sistemas [1].

En general, se observa una distribución de temperatura y salinidad distinta dependiendo de la zona en la que se encuentre. Las zonas norte y centrales de la región de aysén presentan salinidades menores a 34 ppt y temperaturas mayores a 10°C hasta llegar a las aguas sub-antárticas (aproximadamente desde los 55°S). En la zona sub-antártica se presentan masas de agua formadas remotamente, como las masas de aguas del frente polar que se caracterizan por temperaturas menores a 6°C y altas salinidades (33.7-34.5 ppt) [1].

Por otro lado, la relación entre el fitoplancton y la disponibilidad de macronutrientes, debido a la disponibilidad de nitratos y ácido silicílico en las aguas superficiales de algunos fiordos, junto con el ingreso de aguas continentales con bajo ratio de nitratos explica la dominancia de dinoflagelados en el fitoplancton de la región [2].

A continuación, se incluye la geografía de la Región de Aysén y lo que motiva el uso del camino más corto, la cantidad de fiordos y canales pequeños que tiene la región.

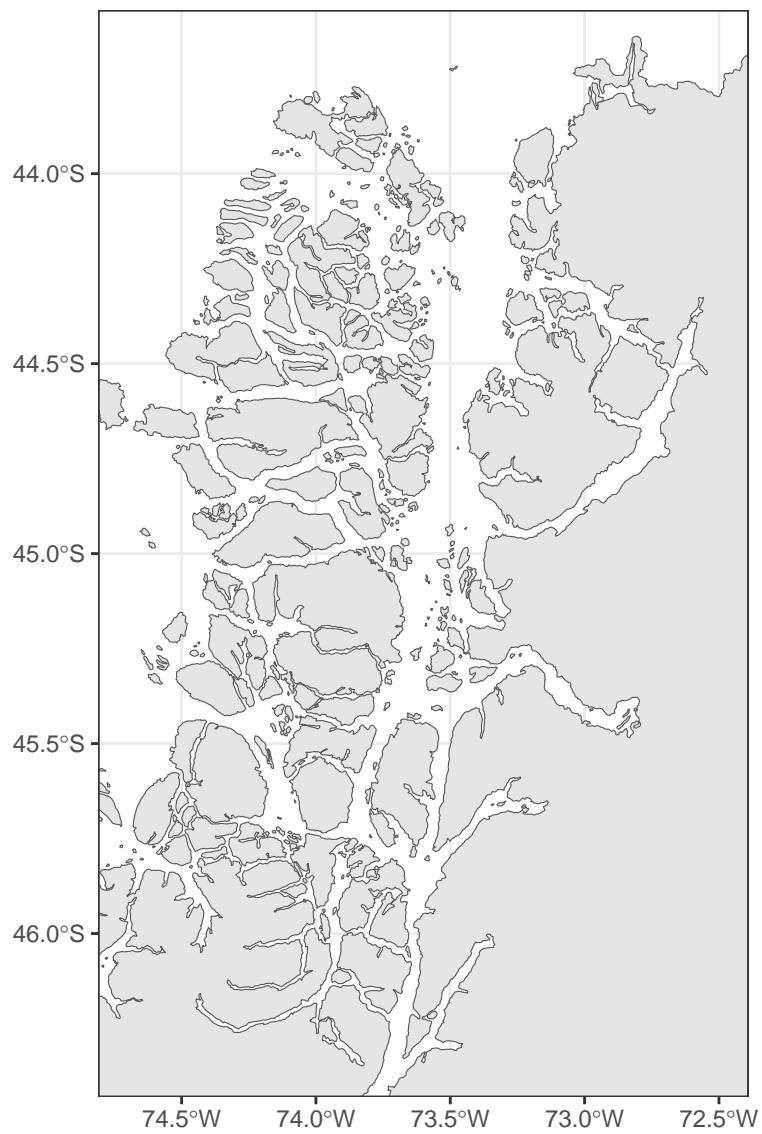


Figura 1.1: Región de Aysén

1.2 SALMONICULTURA

En la salmonicultura chilena se cultivan principalmente tres especies salmonídeas: Salmón del atlántico (*Salmo salar*), Trucha Arcoiris (*Oncorhynchus mykiss*) y el Salmón Coho o Salmón del pacífico (*Oncorhynchus kisutch*). De estas tres especies, las principales son el Salmón del atlántico y la Trucha Arcoiris.

Las especies en el ambiente marino son cultivadas en balsas-jaulas de 30 x 30 o 40 x 40 mts con profundidad variable según las redes instaladas.

Durante el año 2023 en la X región de Los Lagos, la XI región de Aysén y la XII región de Magallanes se produjeron 445.414, 513.084 y 128.876 toneladas de salmón respectivamente [3]. De estas 1.087.374 toneladas de salmón producidas en estas tres regiones, 74.6 % corresponde a Salmón del Atlántico y Trucha Arcoiris. En particular en la región de Aysén, el 86 % corresponde a producción de Salmón del Atlántico y Trucha Arcoiris [3], las cuales son las especies más susceptibles a la infestación por *C. rogercresseyi* [4].

El ciclo general de los salmonídeos se caracterizan principalmente por tener tres grandes fases: una fase de desarrollo inicial en agua dulce, una segunda fase de esmolteación donde se adaptan en aguas salobres, y una fase de crecimiento en agua de mar. En esta última fase, es donde el crecimiento a la talla comercial se realiza.

Salmón del atlántico: Esta especie se caracteriza por tener una temperatura ideal de crecimiento cercana a los 10 °C [5], similar a lo observado en las regiones de Los Lagos, Aysén y Magallanes. A nivel de producción, los ciclos en agua de mar que se encuentran entre 12 a 18 meses en agua de mar. El peso promedio de cosecha es cercano a los 5 kg promedio, dependiendo de los requerimientos de la empresa y mercado de destino.

Trucha arcoíris: La temperatura ideal de esta especie se encuentra entre 13 - 18°C [7], pero capaz de mantenerse en un rango amplio de temperaturas. A nivel productivo, se caracteriza por tener ciclos de producción cortos de 8 a 10 meses en agua de mar, llegando a un peso promedio de cosecha de aproximadamente 3.2 kg.

1.3 *Caligus rogercresseyi*

1.3.1 GENERALIDADES

La caligidosis es una enfermedad parasitaria propia de los peces, en Chile producida principalmente por el ectoparásito *C. rogercresseyi*, también conocido como piojo del salmón. Esta patología se considera una de los problemas sanitarios más relevantes durante la engorda en agua de mar debido a los costos de vigilancia y de control que implica.

Este parásito tiene una baja especificidad de hospedero, donde se ha identificado una gran de hospederos naturales, como *E. maclovinus*, *Odonthestes regia* y *Paralichthys microps*, las

cuales se encuentran frecuentemente alrededor de centros de producción de salmon atraídos por las excedentes de alimento.

Normativa: La normativa aplicable al control de caligus corresponde al PSEVC - Caligus o “Programa de Salud Específico de Vigilancia y Control de Caligidosis”, el cual establece las bases para el monitoreo y seguimiento de los centros de engorda en agua de mar, donde cada centro debe muestrear en 4 jaulas indicadas por SERNAPESCA 10 peces, y contar los distintos estadíos de *C. rogercresseyi* [8]. En particular, se destaca la clasificación de SERNAPESCA:

- Juveniles: Chalimus I – Chalimus IV.
- Adultos Moviles (AM): Machos + Hembras sin sacos ovígeros.
- Hembras Ovígeras (HO): Hembras con sacos ovígeros.

Esto principalmente porque la identificación de los distintos estadíos del parásito es compleja de realizar muestreos a simple vista.

SERNAPESCA establece distintas ventanas de tratamiento por Macrozona, en donde los centros tienen permitido realizar baños con productos farmacológicos durante ese período, con el fin de disminuir la chance de que el parásito genere resistencia.

El segundo método de control corresponde al establecimiento de Centros de Alta Diseminación, los cuales corresponden a centros que hayan terminado una ventana de baño con cargas de HO mayores a 3; a un CAD se le puede solicitar:

- Solicitud de cosecha anticipada del 50% de la biomasa si fue categorizado como CAD 3 veces consecutivas, la cual deberá efectuarse en 21 días consecutivos.
- Solicitud de cosecha anticipada del 25% de la biomasa para aquellos centros que presenten 3 categorizaciones de CAD en 6 evaluaciones consecutivas, la cual deberá efectuarse en 21 días consecutivos.

1.3.2 CICLO DE VIDA DEL PARÁSITO

C. rogercresseyi se caracteriza por tener un ciclo de vida directo en peces, donde se observan estadíos parásitos y estadíos de vida libre, sin observarse un hospedero intermediario.

En los estadíos de vida libre se encuentra el Nauplio I, Nauplio II y Copepodito. Los dos primeros estadíos corresponden a vida libre planctónico, mientras que el copepodito es el estadío infestante. El proceso desde eclosión hasta llegar a copepodito a 12°C demora aproximadamente 10 días en condiciones experimentales controladas [9].

El estadío parásito se caracteriza por presentar 5 estadíos: Chalimus I, Chalimus II, Chalimus III, Chalimus IV y el estadío adulto. Los estadíos de Chalimus son estadíos fijos, mientras que los adultos (inmaduros y maduros) son estadíos móviles. Este último presencia dimorfismo

sexual entre hembras y machos. Desde el momento de infestación hasta adulto maduro toma aproximadamente 18 días a 12°C [9].

El número de aproximado huevos por cada saco ovígero es de 45, el cual puede aumentar durante el invierno debido a la disminución de temperatura o en condiciones que sean más o menos apta para el parásito [9].

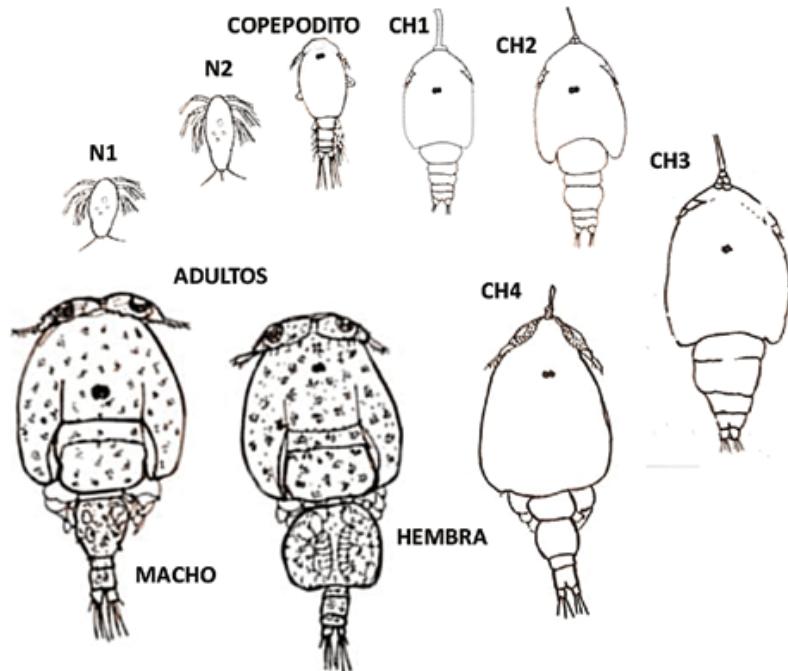


Figura 1.2: Ciclo de Caligus [10]

En el cultivo de salmones solo se observan los estadios parásitos, en donde SERNAPESCA clasifica a los Chalimus I - IV como “Juveniles” y los adultos son divididos entre “Adultos Móviles” (AM) y “Hembras Ovígeras” (HO). Esta última diferenciación se realiza dado que si bien ambos estadios son adultos, la diferenciación a simple vista entre estos estadios es compleja.

El ciclo del parásito se caracteriza por tener una duración promedio de 7 - 9 días como estadio libre antes de llegar a volverse copepodito, y entre 12 a 15 días para llegar a su respectivo estadio adulto una vez infesta a un salmón [9]. Al ser un animal ectotérmico, su ciclo y metabolismo se ven afectados directamente por la temperatura, por lo que temperaturas más altas reducen el tiempo que requiere el parásito para completar su ciclo [9].

1.3.3 TRATAMIENTOS

Actualmente se encuentran distintos tipos de tratamiento disponibles para este parásito y con distintas metodologías.

- Inmersión en Jaula: Azametifos, Peróxido de Hidrógeno, Deltametrina y Lyptus
- Inmersión en barco con estanco cerrado: Hexaflumurón, agua dulce.
- Tratamientos físicos en barcos en estancos cerrados: Hydrolicer, Thermolicer.
- En alimento en agua dulce: Benzoato de Emamectina, Lufenurón.

Cada uno de estos tratamientos tiene distintos costos, siendo los tratamientos de inmersión en jaula los más baratos (a excepción del peróxido de hidrógeno), y donde los tratamientos en barcos a través de estancos o métodos especializados permiten capturar los caligus que se desprenden del pez.

Además, los tratamientos como el azametifos, deltametrina, peróxido de hidrógeno y lyptus afectan principalmente el estadio adulto del parásito, aunque recientes investigaciones sugieren que también podría haber un efecto en Chalimus III y Chalimus IV [11]. En particular, el peróxido de hidrógeno no mata al parásito sino que logra un desprendimiento del estadio adulto, en donde se observa una reinfestación si el parásito no es removido del medio [12].

Los otros medios físicos como Hydrolicer y Thermolicer, en general se utilizan en menor medida dado al daño que le infringen a los peces en las escamas y piel, ya que este daño puede facilitar el ingreso de otros patógenos al debilitar la primera barrera de defensa.

1.4 METODOLOGÍAS

1.4.1 MÉTODOS PREVIOS

M. espacio-temporal estocástico [13]: Modelo propuesto por [13], que utiliza la distribución binomial negativa cero-inflada, en donde se modela la media a través de la siguiente especificación:

$$\mu_{it} = \mathbf{S}_{it} \cdot \kappa_{it}^{susc} \cdot (\lambda_{it}^w + \lambda_{it}^d + \lambda_{it}^o)$$

Donde los dos efectos multiplicativos:

- \mathbf{S}_{it} es una variable indicadora que toma el valor 1 cuando el sitio i está activo, o 0 en otro caso.
- κ_{it}^{susc} es un factor proporcional a la susceptibilidad del sitio i y depende de variables explicativas que caracterizan las condiciones de los peces de la granja i en el mes t . Este término toma la forma: $\kappa_{it}^{susc} = \exp(\sum_k \beta_k^{susc} \mathbf{x}_{ikt}^{susc})$, donde \mathbf{x}_{ikt}^{susc} son las variables explicativas del sitio i en el mes t y β_k son los coeficientes asociados a las covariables.

Los términos aditivos especificados corresponden a

1. λ_{it}^w representa la infección dentro de la granja actual de interés. Este parámetro es modelado a través de una suma ponderada de las últimas cargas de caligus del sitio i .
2. λ_{it}^d corresponde a la infección de los sitios vecinos, la cual depende de la distancia de camino más corto (seawater distance) a estos sitios y en sus abundancias de caligus. Este término corresponde a una suma de las contribuciones de todas las granjas $i \neq j$.
3. λ_{it}^o representa infección de otras fuentes no especificadas, como reservorios de infección o peces hospederos de vida libre. Se modela actualmente como una constante γ , y actúa como un término de intercepto.

El parámetros λ_{it}^w es modelado de la siguiente forma:

$$\lambda_{it}^w = \left(y_{i(t-1)} / \mathbf{n} + \sum_{l=2}^{L} \rho_l y'_{i(t-l)} / \mathbf{n} \right)^\alpha \cdot \mathbf{S}_{i(t-1)}$$

- $y'_{i(t-1)}$ es igual a $y_{i(t-1)}$ cuando el centro de cultivo ha estado activo des de los meses $t - 1$ a $t - l$, pero es cero si el centro ha estado inactivo en cualquiera de estos meses.
- α es un parámetro positivo que permite modelar dependencias no lineales en los conteos de caligus previos.

- $\rho_{l=2}, \dots, \rho_{l=L}$ son parámetros que consideran el efecto de cargas de los meses anteriores. Corresponde a una suma ponderada de las cargas de los meses previos.

El parámetro λ_{it}^d corresponde a

$$\lambda_{it}^d = \sum_{j \neq i} \exp(\phi_0 + \phi_1(d_{ij}^{\phi_2} - 1)/\phi_2) \kappa_{j(t-1)}^{inf} \cdot \left(y_{j(t-1)}/n + \sum_{l=2}^{l=L} \rho_l y'_{i(t-1)}/n \right)^{\alpha} \cdot S_{j(t-1)}$$

Donde,

- ϕ_0 cuantifica la importancia de infestaciones cercanas comparada con otras fuentes.
- d_{ij} es la distancia marítima (o camino más corto) entre los centros i y j
- ϕ_1 y ϕ_2 son parámetros que reflejan el efecto de la distancia marítima entre los centros vecinos. La transformación $(d_{ij}^{\phi_2} - 1)/\phi_2$ corresponde a la transformación de Box-Cox. cuando ϕ_2 se acerca a 0, la transformación corresponde a $\log(d_{ij})$.
- κ_{jt}^{inf} corresponde a un factor proporcional a la infecciosidad del centro j , dependiendo de las variables explicativas que caracterizan al centro j .

Este último término de infecciosidad tiene la forma:

$$\kappa_{jt}^{inf} = \exp \left(\sum_k \beta_k^{inf} x_{jkt}^{inf} \right)$$

Los parámetros $\alpha, \gamma, \phi_0, \phi_1, \phi_2, \rho_2, \dots, \rho_p$ y todos los β unidos en un vector de parámetros θ , el cual se encontró maximizando la log-verosimilitud.

M. AR de Estado Espacio [14]: Son modelos de estado-espacio multivariados (MARSS) propuestos por [14], se han utilizado para estimar la carga de pre-adultos y adultos móviles (PAAM) y hembras ovígeras (AF) de forma conjunta. El modelo tiene la siguiente especificación:

$$\tilde{\theta}_t = B\tilde{\theta}_{t-1} + u + \tilde{\beta}X + w_t$$

El cual corresponde a:

$$\begin{bmatrix} \theta_{p,t} \\ \theta_{a,t} \end{bmatrix} = \begin{bmatrix} b_p & 0 \\ b_{ap} & b_a \end{bmatrix} \begin{bmatrix} \theta_{p,t-1} \\ \theta_{a,t-1} \end{bmatrix} + \begin{bmatrix} u_p \\ u_a \end{bmatrix} \\ + \begin{bmatrix} \beta_{bath,p} & \beta_{feed,p} & \beta_{IIP,p} & \beta_{EIP,p} & \beta_{temp,p} \\ \beta_{bath,a} & \beta_{feed,a} & 0 & 0 & \beta_{temp,a} \end{bmatrix} \begin{bmatrix} x_{bath,t} \\ x_{feed,t} \\ x_{IIP,t} \\ x_{EIP,t} \\ x_{temp,t} \end{bmatrix} + \begin{bmatrix} W_{p,t} \\ W_{a,t} \end{bmatrix}$$

En donde:

- $\tilde{\theta}_t$ es el vector de abundancia de PAAM y AF en escala logarítmica en el tiempo t .
- \tilde{B} tiene los términos AR-1 para la abundancia de caligus en la diagonal, y en la off-diagonal se agrega un término para agregar la contribución de pre-adultos a la carga de hembras ovígeras (AF)
- El vector \tilde{u} es un vector de escala.
- \tilde{X}_t son las covariables y $\tilde{\beta}$ son sus respectivos coeficientes.
- \tilde{w}_t es un vector de errores en el tiempo t , los cuales se asumen que tienen una distribución normal multivariada con media cero, y matriz de varianza-covarianza $\tilde{Q} = \begin{bmatrix} q_p & q_{ap} \\ q_{ap} & q_a \end{bmatrix}$.

Estos modelos se enfocan en predecir las hembras ovígeras de manera separada al resto de la carga parasitaria, en particular porque estas son las que se consideran más relevantes para cuantificar la presión de infestación que aporta el centro en ese momento.

La estimación del modelo MARSS propuesto en este estudio fue realizada a través del algoritmo EM implementado en el paquete MARSS de R.

M. de dos partes con efectos aleatorios [15]: Corresponde a la estimación en dos partes propuesta por [15], se utiliza un modelo multivariable de efectos aleatorios para determinar los factores más relevantes en la abundancia promedio del estadio fijo del piojo del salmón en salmonidos de vida libre en British Columbia (Candadá) en una zona i y un mes t .

- Primera parte: Modelamiento de la chance de que la media de infestación fuese mayor que 0 utilizando regresión logística.
- Segunda parte: Modelamiento del conteo de piojos del salmón dado que los conteos eran mayor a 0 utilizando regresión Gamma.

- Efecto aleatorio: Los efectos aleatorios fueron incluídos considerando muestreos repetidos en cada zona i , y la potencial autocorrelación entre mediciones mensuales fue agregada considerando un outcome desfasado en $t - 1$ en las covariables.

Se destaca que este es un método que se utilizó en salmones de vida libre. Es decir, en ambiente no productivo. Dado que en Chile no tenemos salmones de forma endémica y solo en sistemas productivos, utilizar este modelo requeriría modificaciones.

Sin embargo, esta metodología podría ser considerada al evaluar especies nativas que sean hospederas de *C. rogercresseyi*.

1.4.2 MÉTODO GEOESTADÍSTICO PROPUESTO

El método propuesto consiste en la utilización de Kriging Ordinario, el cual asume que el proceso es *intrínsecamente estacionario*, donde nuestra media es constante y la matriz de varianza-covarianza depende de la distancia.

El Kriging es un método que nació inicialmente en Sudáfrica y en la minería del oro, donde se utilizó para interpolar en sitios desconocidos a través muestreos o localizaciones conocidas, para encontrar los lugares con mayor presencia de oro [16].

En este caso, se propone la utilización de Kriging Ordinario para evaluar las zonas que consistentemente presentan mayores cargas parasitarias asociadas.

Consideraremos una variable $\{\mathbf{Y}(\mathbf{s}) : \mathbf{s} \in D \subset \mathbb{R}^d\}$ o $\{\mathbf{Y}(\mathbf{s})\}$ es una colección de variables aleatorias indexadas por D , donde cada $\mathbf{s} \in D$ es una localización geográfica [16].

Consideraremos, además que $\mathbf{Y}(\mathbf{s})$ es un proceso *intrínsecamente estacionario*, donde la Varianza es libre de la localización, y depende solamente de la distancia: $Var[\mathbf{Y}(\mathbf{s} + \mathbf{h}) - \mathbf{Y}(\mathbf{s})]$ depende solo del vector \mathbf{h} [17].

Además de la estacionariedad, dentro de los supuestos de este modelo se encuentra la *isotropía*, donde consideramos que la función del semivariograma $\gamma(\mathbf{h})$ depende solo de la separación del vector solo a través de su largo $\|\mathbf{h}\|$ [18]

Variograma: El variograma es una función definida negativa que caracteriza la covarianza entre las localizaciones observadas a través de las diferencias en una distancia \mathbf{h} , en donde se espera que a medida que $\|\mathbf{h}\|$ aumenta, se espera que la diferencia sea mayor el variograma se puede representar como $2\gamma(h)$ y el semivariograma se considera $\gamma(h)$. El cálculo para el variograma empírico se ha definido de distintas maneras, pero las principales son:

- Semivariograma clásico: $\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} [Y(x_{i+h}) - Y(x_i)]^2$
- Semivariograma robusto de Cressie: $\gamma(h) = \frac{\left[\frac{1}{N_h} \sum_{i=1}^{N_h} |Y(x_{i+h}) - Y(x_i)|^{\frac{1}{2}} \right]^4}{0.914 + \frac{0.988}{N_h}}$

El variograma además, se compone de principalmente tres partes:

- Nugget: Corresponde a la semivarianza del primer desfase de distancia, en la discontinuidad en el origen
- Sill: El límite del variograma cuando $h \rightarrow \infty$
- Range: La distancia en la que la diferencia del variograma con respecto al Sill es mínima, y en el caso de los variogramas con Sill asintótica, corresponde al valor en el que la semivarianza alcanza el 95% del Sill.

Además, para una función de variograma válida, existe una función de covarianza definida positiva basada en la distancia que describe la relación entre los puntos observados y su distancia tal que:

$$C(\mathbf{h}) = C(\mathbf{0}) - \gamma(\mathbf{h})$$

En donde la función de covarianza entre dos puntos desaparece a medida que la distancia aumenta [18], siempre cuando el proceso sea ergódico, es decir, $\lim_{\|\mathbf{h}\| \rightarrow \infty} C(\mathbf{h}) = 0$ [18].

Estimación del variograma:

Hay distintos métodos para la estimación del variograma. El más frecuente es la inspección visual del patrón observado, pero [19] propuso un método numérico robusto a través de la utilización de mínimos cuadrados ponderados, para estimar el vector de parámetros $\hat{\boldsymbol{\lambda}}$:

$$\sum_{j=1}^k N_{h(j)} \left[\frac{\bar{\gamma}(h(j))}{\gamma(h(j)|\hat{\boldsymbol{\lambda}})} - 1 \right]^2$$

Kriging Ordinario: Este método nace originalmente de Matheron (1963) en honor a D.G. Krige, quien dio los primeros pasos para los métodos geoestadísticos actuales [18]. En este caso, consideramos a $\mathbf{Y} = (Y(s_1), \dots, Y(s_n))^T$, y definimos $Y(s_0)$ como la localización a interpolar.

Bajo este criterio, nuestro mejor predictor lineal (BLUP) basado en nuestros datos observados y considerando una función de pérdida cuadrática correspondería a: $\boldsymbol{\lambda}^T \mathbf{Y}$, donde para conseguir nuestros pesos óptimos λ resolvemos un problema de optimización con restricción $\sum_{i=1}^n \lambda_i = 1$

$$\hat{\boldsymbol{\lambda}} = \Gamma^{-1} \left(\boldsymbol{\gamma}_0 + \mathbf{1}^T (\mathbf{1}^T \Gamma^{-1} \mathbf{1})^{-1} (1 - \mathbf{1}^T \Gamma^{-1} \boldsymbol{\gamma}_0) \right)$$

O al expresarlo utilizando la covarianza:

$$\hat{\lambda} = \Sigma^{-1} \left[\mathbf{c}_0 + \mathbf{1}(\mathbf{1}^\top \Sigma^{-1} \mathbf{1})^{-1} (1 - \mathbf{1}^\top \Sigma^{-1} \mathbf{c}_0) \right]$$

Y la varianza de la predicción:

$$\sigma_{ok}^2 = \sigma_0^2 - \mathbf{c}_0^\top \Sigma^{-1} \mathbf{c}_0 + (1 - \mathbf{1}^\top \Sigma^{-1} \mathbf{c}_0)^\top (\mathbf{1}^\top \Sigma^{-1} \mathbf{1})^{-1} (1 - \mathbf{1}^\top \Sigma^{-1} \mathbf{c}_0)$$

Entonces, nuestro predictor lineal para una localización no conocida (o punto a interpolar), corresponde a:

$$p_{ok}(\mathbf{Z}; \mathbf{s}_0) = (\mathbf{1}^\top \Sigma^{-1} \mathbf{1})^{-1} \mathbf{1}^\top \Sigma^{-1} \mathbf{Z}(s) + \boldsymbol{\sigma}^\top \Sigma^{-1} (\mathbf{Z}(s) - \mathbf{1}(\mathbf{1}^\top \Sigma^{-1} \mathbf{1})^{-1} \mathbf{1}^\top \Sigma^{-1} \mathbf{Z}(s))$$

En donde:

- γ_0 es la semivarianza del lugar a interpolar, basada en la distancia de los sitios conocidos al nuevo punto.
- \mathbf{c}_0 es la covarianza del lugar a interpolar, basada en la distancia de los sitios conocidos al nuevo punto.

Kriging Ordinario log-normal:

El Kriging Ordinario log-normal considera que $Y(s) = \log\{Z(s)\}$ sigue una distribución normal multivariada [20]. En este caso, asumimos que $Y(\cdot)$ es intrínsecamente estacionario con media μ_Y , en donde el predictor lineal y la varianza corresponden a:

- $p(\mathbf{Z}; \mathbf{s}_0) = \exp \left(p_{ok}(\mathbf{Y}; s_0) + \frac{\sigma_Y^2}{2} - \frac{\text{Var}(p_{ok}(\mathbf{Y}; s_0))}{2} \right).$
- $E [(p_{olk}(\mathbf{Z}; \mathbf{s}_0) - Z(\mathbf{s}_0))^2] = \exp(2\mu_Y + \sigma_Y^2) \exp(\sigma_Y^2) \times \{(1 + (\exp\{-\sigma_{ok}^2(\mathbf{Y}; \mathbf{s}_0) + m_Y\}) \times (\exp\{m_Y\} - 2)\}$

Por lo que para el caso del kriging log-normal, requerimos estimar μ_Y y $\sigma_Y^2(\cdot)$ [20].

1.4.3 DISTANCIAS

Los métodos de Kriging Ordinario al ser utilizados con la distancia euclídea garantizan que la matriz de varianza-covarianza obtenida a través de las distancias entre los puntos conocidos está garantizada de ser definida positiva [21]. Sin embargo, situaciones como las que ocurren en la Región de Aysén, donde se observan múltiples barreras y canales. Dado esto, se propone la utilización de la distancia de camino más corto (shortest-path), la cual ha sido utilizada en los métodos anteriores.

Esta distancia fue calculada utilizando el paquete gdistance [22], el cual se basa en la utilización de rasters y matrices de transición. Esta matriz de transición utiliza la métrica de conducción para el cálculo de la distancia entre punto A y punto B.

Este cálculo de distancia se realiza utilizando el algoritmo de Dijkstra a través de la matriz de transición, considerando el mar como el terreno navegable a través de la inversión del raster basado en la geometría de la región de aysén.

Para garantizar que la matriz de varianza-covarianza sea definida positiva al utilizar la distancia del camino más corto, consideraremos una corrección llamada ‘ClosePD’ propuesta por [23], en donde la matriz de varianza-covarianza obtenida a través de la distancia pasa por un proceso de descomposición en los valores propios: $C([d_{ij}]) = V\Lambda V^\top$.

Se considerará un umbral ϵ , tal que $\epsilon = \frac{\lambda_1}{\tau}$, $\lambda_1 < \lambda_2 < \dots < \lambda_k$ donde τ es un valor de tolerancia predeterminado, en caso de ser necesaria la corrección.

El umbral se evalúa en cada λ_i :

$$\tilde{\Lambda}_i = \begin{cases} \lambda_i, & \lambda_i \geq \epsilon \\ \epsilon, & \lambda_i < \epsilon \end{cases}$$

A través de esto, se consigue una nueva matriz de covarianza basada en los valores propios corregidos, que se utiliza para reescalar la matriz de varianza-covarianza: $\tilde{C}([d_{ij}]) = V\tilde{\Lambda}V^\top$, donde $\tilde{\Lambda} = \text{diag}(\tilde{\Lambda}_1, \dots, \tilde{\Lambda}_n)$ corresponde a los valores propios de la matriz corregida co el método ‘ClosePD’.

2 MATERIALES Y MÉTODOS

2.1 Datos

El software utilizado fue R 4.3.3, la IDE RStudio 2023.12.1 Build 402 y el documento fue generado utilizando Quarto.

Los datos analizados corresponden a las declaraciones semanales de cargas parasitarias de juveniles (Juv), adultos móviles (AM) y hembras ovígeras (HO) de 476 sitios de producción de Salmón del atlántico y Trucha Arcoiris que fueron declaradas a SERNAPESCA durante el período 2015 - 2020. Se consideró como variable de interés el total de adultos (Hembras Ovígeras + Adultos Móviles), dado que son la fracción que se espera que sea capaz de desprendérse del pez y afectar otro hospedero. Además, esta fracción incluye tanto Hembras Ovígeras como Hembras sin sacos ovígeros.

La abundancia de Caligus es altamente asimétrica, con una gran proporción de valores cercanos a 0, en especial en períodos cercanos al inicio de ciclo dado que los peces ingresan con 0 carga parasitaria al ambiente marino.

Dada la naturaleza espacio-temporal de la fuente de datos, se optó por trabajar la metodología espacial separada de la temporal, en donde se consideraron tres métodos de resumen de datos: 1) El promedio de los máximos por ciclo productivo de cada centro de cultivo, 2) El promedio mensual, después de calcular los máximos por año/mes de cada centro y 3) el promedio de los máximos por estación, por año, por centro.

Dado que la normativa indica que las notificaciones son semanales para *S. salar* y *O. mykiss* además de un período de descanso de 3 meses entre ciclos productivos por ACS, se consideró un nuevo ciclo cuando había pasado más de 8 semanas entre la medición t y la medición $t - 1$, en cada centro de cultivo, para efecto del cálculo de promedio de máximos.

El análisis de la carga de adultos totales fue analizada de forma conjunta (sin diferenciar entre especie) para *S. salar* y *O. mykiss*, dado que ambas especies son susceptibles, y también a que si se utilizara Kriging Universal se requeriría de la covariante **especie** en los puntos a interpolar, algo que no es posible considerar debido a la naturaleza del problema.

2.2 Distancias

Las distancias corresponden a las distancias d_{ij} entre cada uno de los sitios en kilómetros. Las dos métricas de distancia utilizadas son la distancia euclíadiana y la distancia de camino más corto. La distancia euclíadiana fue calculada utilizando el paquete `sf` [24], mientras que la distancia del camino más corto fue calculada a través del paquete `gdistance` [22] utilizando la función `shortestPath`, la cual utiliza el algoritmo de Dijkstra y requiere una matriz de transición geo-corregida para calcular los valores de distancia correctamente.

El proceso para obtener la matriz de transición es a base de un raster, en este caso:

1. Creación de un raster de 5000 filas y 5000 columnas.
2. Especificación de `extent` del raster, en este caso (-75.5, -72) y (-46.5, -43.5).
3. Rasterización de los polígonos de la región utilizando este raster base.
4. Asignar a los valores NA del raster (parte terrestre), el valor -999.
5. Asignar un valor NA a las celdas del raster mayores a -999.
6. Asignar el valor 1 a las celdas con valor -999.

Este proceso permite invertir el raster, logrando que el área marítima y los canales sean considerados terreno navegable, y volviendo la parte terrestre una barrera impasable. Este cálculo también corresponde a la distancia marítima o seawater distance entre dos puntos.

La geo-corrección se le realiza a la matriz de transición a través de la función `geoCorrection`, considerando el argumento `type='c'` para la distancia del camino más corto.

Se destaca que en distancias cortas y sin barreras de por medio, esta distancia sería equivalente a la distancia euclíadiana dado que si no hay barreras que impidan el movimiento, el camino más corto es una línea recta entre el punto A y punto B.

Las distancias a los puntos de interpolación fueron calculadas de la misma manera que para las distancias entre los centros de cultivo.

2.3 Generalidades del Kriging Ordinario

En el presente trabajo se utilizó el Kriging Ordinario como método de interpolación y predicción espacial, en donde asumimos que el parámetro μ es desconocido pero constante, donde las variaciones que observamos dependen de la función de varianza-covarianza entre los puntos observados.

En este caso, el mejor predictor lineal insesgado (BLUP) bajo la función de pérdida cuadrática corresponde a la $E[\mathbf{Z}(\mathbf{s})] \equiv \mu\mathbf{1}$. El predictor lineal es aquel que minimiza $E[(p(\mathbf{Z}|\mathbf{s}_0) - Z(\mathbf{s}_0))^2]$ donde se asume que el modelo sigue:

$$\mathbf{Z}(s) = \mu \mathbf{1} + \mathbf{e}(s), \quad \mathbf{e}(s) \sim (0, \Sigma)$$

Donde en este caso, μ es desconocido, y Σ es conocido.

Entonces, nuestro predictor lineal toma la forma de $p(\mathbf{Z}|s_0) = \lambda_0 + \boldsymbol{\lambda}^\top \mathbf{Z}(s)$

2.4 Semivariogramas

Los semivariogramas empíricos fueron evaluados con el método clásico y con el estimador de módulo de Cressie. El ajuste de los semivariogramas fue realizado a través del método de mínimos cuadrados ponderados [19].

Para el semivariograma Matérn y los siguientes semivariogramas, los siguientes parámetros representan:

- ϕ : Rango o parámetro de correlación, asociado a la distancia.
- κ : Parámetro de suavidad, solo requerido en el semivariograma Matérn y Cauchy en esta instancia.
- σ^2 : Varianza del proceso.

En los siguientes modelos de semivariograma, la expresión se incluye en términos de la correlación.

Semivariograma Matérn:

El semivariograma Matérn se define como:

$$\gamma(\mathbf{h}) = \sigma^2 \left(1 - \frac{2^{1-\kappa}}{\Gamma(\kappa)} \left(\frac{\|\mathbf{h}\|}{\phi} \right)^\kappa K_\kappa \left(\frac{\|\mathbf{h}\|}{\phi} \right) \right)$$

Donde $\sigma^2 > 0$ es la varianza, $\phi > 0$ es la escala (o rango/range), $\kappa > 0$ es el parámetro de forma, $\Gamma(\cdot)$ es la función Gamma y $K_\kappa(\cdot)$ es la función de Bessel modificada de segundo tipo y orden κ y $\|\mathbf{h}\|$ es la norma del vector \mathbf{h} [25].

El modelo Matérn es un modelo que llega al *Sill* de manera asintótica [25], por lo que es práctica frecuente considerar el valor del *Sill* que contiene el 90% de los valores del semivariograma.

Y al expresarlo utilizando la covarianza:

$$\rho(\|\mathbf{h}\|) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\frac{\|\mathbf{h}\|}{\phi} \right)^\kappa K_\kappa \left(\frac{\|\mathbf{h}\|}{\phi} \right)$$

A continuación, los otros semivariogramas utilizados:

Semivariograma Gaussiano:

$$\rho(\|\mathbf{h}\|) = \exp \left[- \left(\frac{\|\mathbf{h}\|}{\phi} \right)^2 \right]$$

Semivariograma Esférico:

$$\rho(\|h\|) = \begin{cases} 1 - 1.5 \frac{\|h\|}{\phi} + 0.5 \left(\frac{\|h\|}{\phi} \right)^3, & h < \phi \\ 0, & \text{eoc.} \end{cases}$$

Semivariograma Cauchy:

$$\rho(\|h\|) = \left[1 + \left(\frac{\|h\|}{\phi} \right)^2 \right]^{-\kappa}$$

Parte II

RESULTADOS

3 Análisis exploratorio

En la actualidad tenemos 847 concesiones otorgadas para Salmón del Atlántico y Trucha Arcoíris entre la región de Los Lagos (X), Aysén (XI) y Magallanes (XII). Las principales regiones en las que la industria se ha establecido son la X y XI.

La XI región presenta un total de 37 ACS, las cuales fueron creadas por el Servicio Nacional de Pesca y Acuicultura en el año 2010 posterior a la crisis del virus ISA y, lo cual permite generar estrategias de baños antiparasitarios, estrategias de cosecha y períodos de descanso coordinado para cada una de las ACS.

Cada ACS y centro de cultivo tiene características específicas. A continuación, se presenta lo que ocurre a nivel de las 10 concesiones con más observaciones.

Se observa alta variabilidad independiente de cada ACS, principalmente debido a la naturaleza temporal y cero inflada de los datos recolectados por SERNAPESCA.

Tabla 3.1: Estadística descriptiva de adultos totales en las 10 ACS con mayor cantidad de observaciones

ACS	n	Media	Desv. Estándar	C. Variación	Mediana	Máximo
ACS 32	4729	2.17	4.15	191%	0.43	53.08
ACS 28 B	3619	0.71	1.84	260%	0.00	30.10
ACS 18 C	3332	6.33	15.26	241%	2.79	287.00
ACS 21 C	3280	4.33	5.20	120%	3.03	57.40
ACS 33	3104	3.98	7.08	178%	1.90	127.35
ACS 21 B	2702	5.96	10.49	176%	3.50	225.90
ACS 34	2478	5.77	15.30	265%	1.96	182.73
ACS 19 A	2307	4.41	6.17	140%	2.90	119.00
ACS 22 D	2057	4.58	6.41	140%	3.05	84.00
ACS 18 D	1897	4.36	6.57	151%	2.67	75.75

Tabla 3.2: Est. Descriptiva de los 10 principales centros de la ACS 32

Centro	n	Media	Desv. Estándar	C. Variación	Mediana	Máximo
110726	294	1.09	1.98	182%	0.03	10.35
110268	292	1.22	2.05	168%	0.05	11.77
110106	291	2.53	4.40	174%	0.55	25.25
110703	274	1.55	2.69	174%	0.38	18.28
[H]	110071	267	2.53	4.47	177%	0.42
	110519	238	2.31	3.63	157%	0.79
	110060	216	1.74	2.91	167%	0.49
	110058	211	2.29	3.95	172%	0.77
	110131	200	3.64	5.28	145%	1.47
	110178	186	2.57	4.57	178%	0.42
						28.65

Dentro de la ACS 32, los 10 principales centros también presentan un alto coeficiente de variación. Se destaca la diferencia entre las medias de estos centros. Si bien todos presentan variabilidad similar con respecto a su media, muchos de ellos se encuentran en promedio más cercano a 1, mientras que otros cercanos a un promedio de 2 o 3, lo que podría indicar ciertas características del manejo de la concesión, de los peces, o características inherentes a su localización geográfica.

Descripción temporal:

La base de datos original contiene un total de 106843 observaciones de 988 centro de cultivo de salmones en un período del 2015 a 2021, correspondientes a 88 Asociaciones de Concesiones (en adelante, ACS o barrios). Para el análisis, se consideraron solo los centros de cultivos de Salmón del atlántico (*Salmo salar*) y Trucha Arcoíris (*Oncorhynchus mykiss*) pertenecientes a la XI Región de Chile. La exclusión de las especies como el Salmón Coho es principalmente debido a su resistencia frente al piojo del salmón.

En el siguiente gráfico obseramos la carga total de parásitos de cada ACS en la XI Región de Aysén. Se excluyeron las concesiones con carga muy cercana a 0 durante todo el período, además de las que fueron otorgadas por un período específico de tiempo.

Se observan patrones estacionales marcados en cada uno de los barrios. Esto correspondería a las estaciones del año, junto con los descansos sanitarios que se realizan al cierre de las ACS.

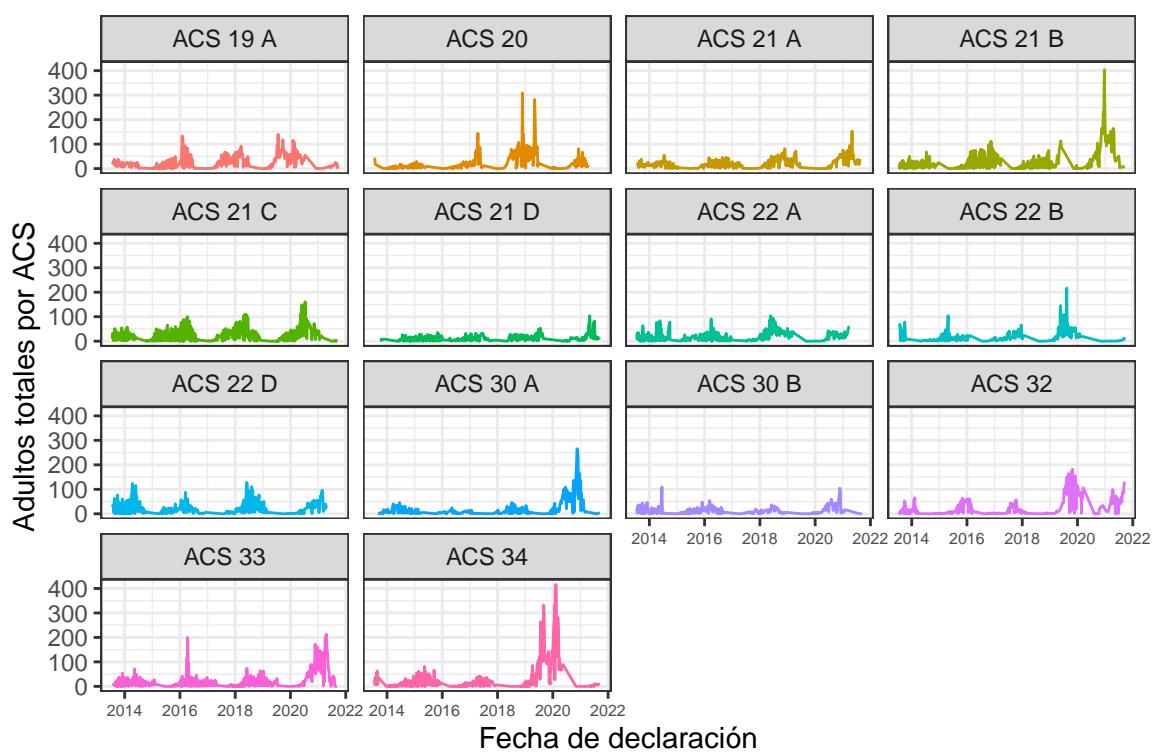


Figura 3.1: Adultos totales por ACS en la XI Región

Además, la serie temporal fue evaluada en la ACS 32, donde se observa que los centros activos dentro de un barrio varían en su fecha de ingreso y término, dado que algunos se encuentran presentes en distintos períodos o inactivos por largos períodos de tiempo.

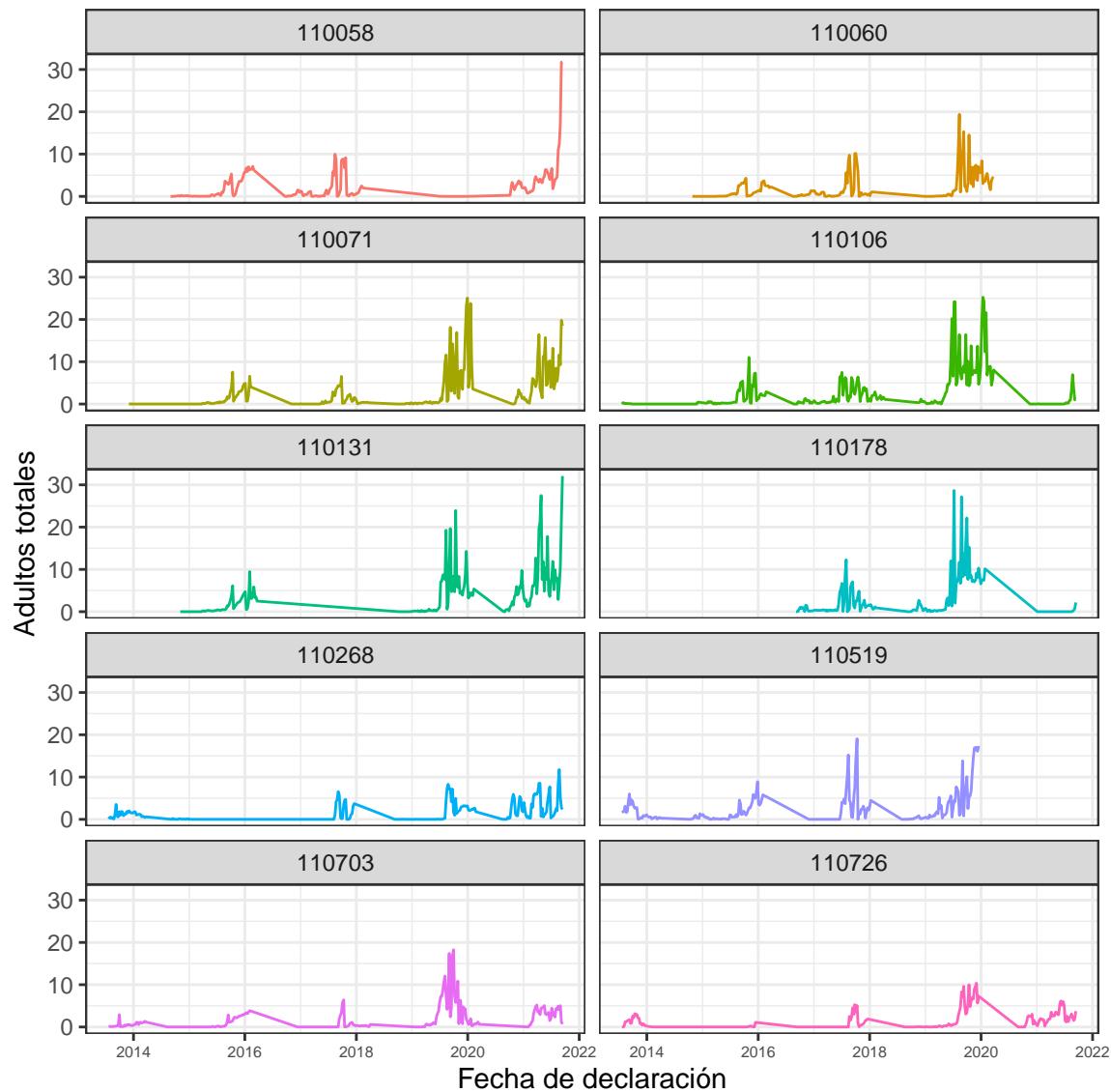


Figura 3.2: Carga de promedio de Adultos totales en centros de la ACS 32

Descripción geográfica:

A continuación se evaluó el máximo de adultos totales por centro de cultivo como medida resumen, debido a la naturaleza cero-inflada del conjunto de datos.

Se observa que en general, la mayoría de los centros presenta máximos menores a 100 adultos totales, y que los centros con máximos entre 100 y 150 adultos totales se encuentran en zonas cercanas.

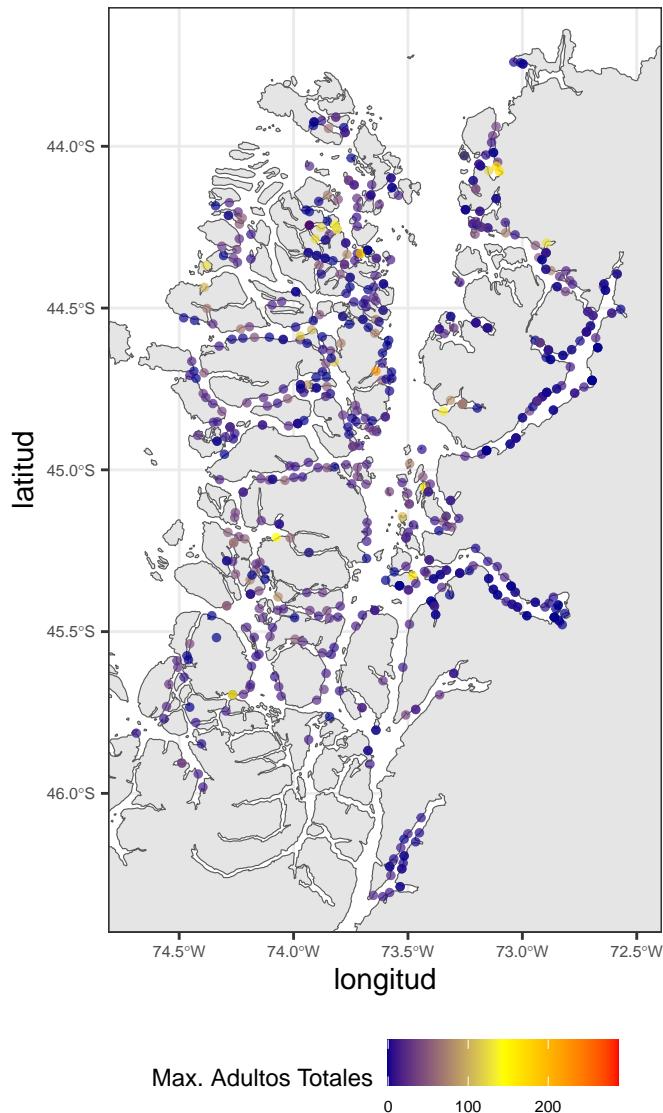


Figura 3.3: Máximo de adultos totales por centro de cultivo

Al evaluar la carga de adultos totales (Prom HO + Prom AM), se observa que corresponde a una variable cero-inflada que puede tomar valores extremos de hasta más de 200 adultos totales promedio. Esto ocurre dado que los centros al ingresar al ambiente marino no presentan carga de caligus, y una vez comienza a avanzar el ciclo se comienzan a infestar por los estadíos juveniles del parásito.

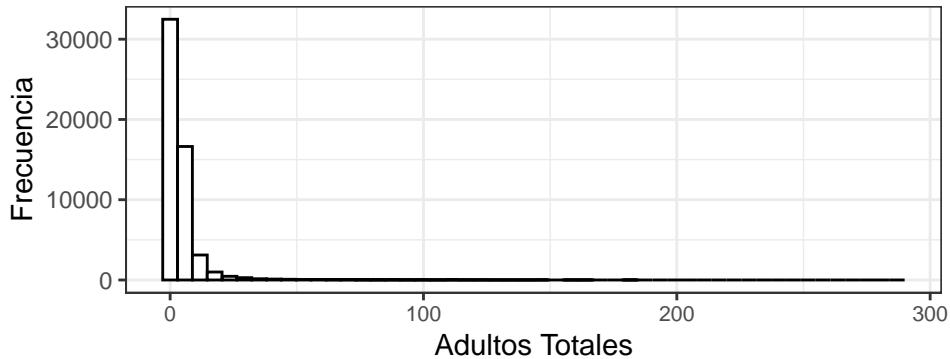


Figura 3.4: Histograma de la variable respuesta (Adultos totales) en la XI región de Aysén

Al evaluar ambas matrices de distancia podemos ver que en distancias menores a 25 km la distancia del camino más corto tiene valores relativamente cercanos a la distancia Euclíadiana, mientras que a medida que aumentamos la distancia, consistentemente se observan valores menores de distancia euclíadiana.

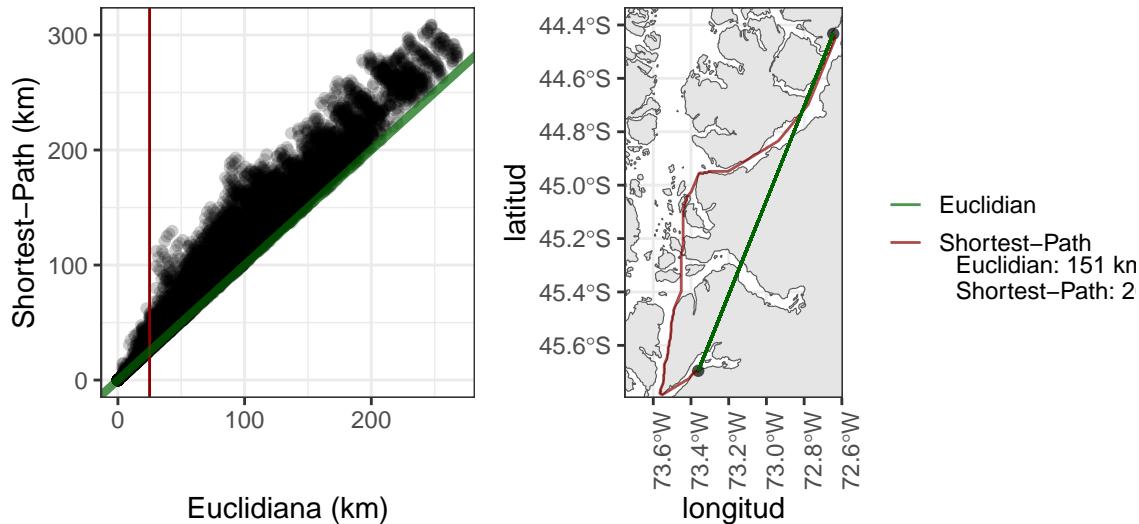


Figura 3.5: Comparación entre las distancias producidas por shortest-path, y la distancia euclidiana

4 Kriging Ordinario

4.1 Máximos

Se generó un promedio de los máximos por ciclo, por centro de salmones con el fin de realizar un resumen que no contuviera autocorrelación temporal. Los ciclos en cada centro pueden ser considerados como independientes, debido a que después de cada ciclo se cierra la ACS y con eso, todos los centros circundantes.

Para este caso, nuestra nueva variable distribuye de la siguiente manera:

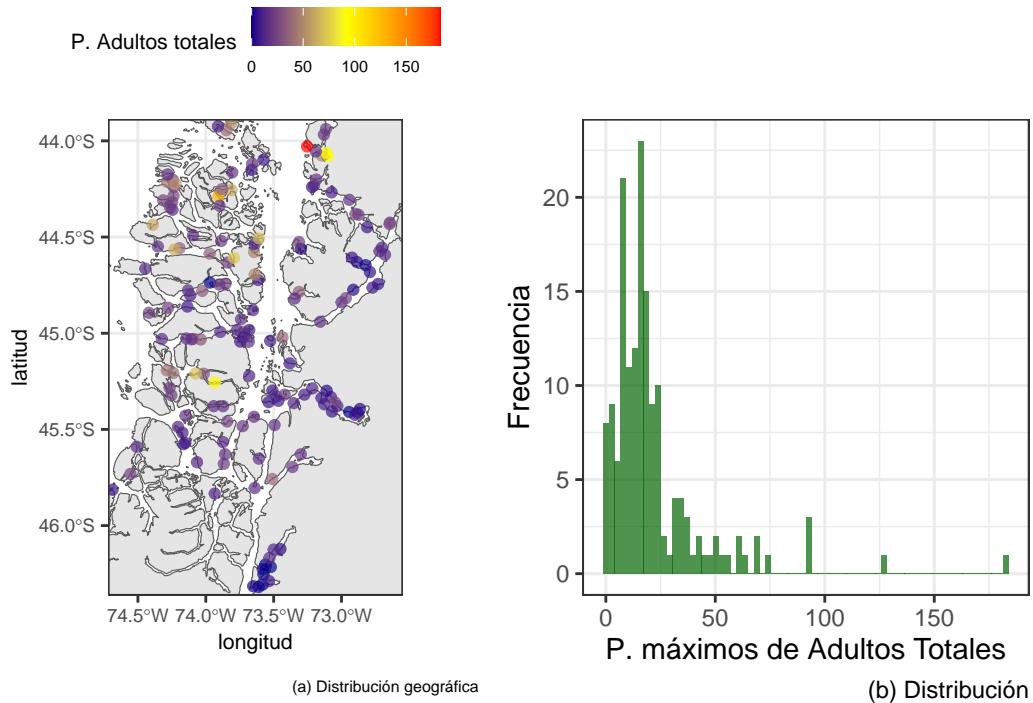


Figura 4.1: Promedio de máximos por ciclo

4.2 Variogramas

Dada la distribución asimétrica, con valores sobre 150 adultos promedio, se utilizó el semivariograma robusto de cressie del logaritmo de los adultos totales. Se consideró la distancia máxima como 60 km para ambas distancias.

A pesar de utilizar el estimador robusto del semivariograma, se observa que la distancia euclíadiana tiende a tener mayor varianza en distancias más cortas (<10 km).

El modelo Matérn en utilizando distancia euclidiana presenta baja suavidad, asociado a un parámetro de suavidad bajo, por lo que se observa un quiebre cuando este se acerca al 0.

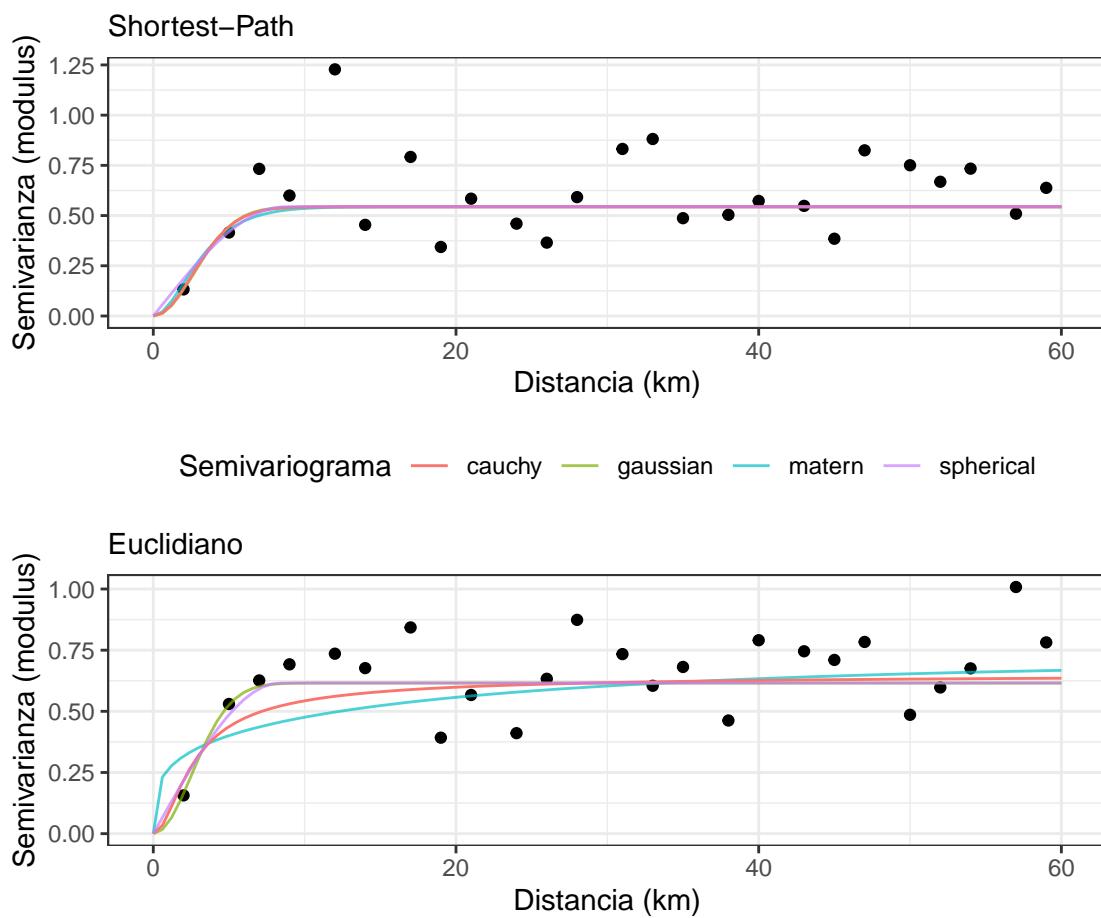


Figura 4.2: Semivariogramas robustos para distancia shortest-path y euclidiana

Todos los modelos de semivariograma llegan a valores similares del parámetro σ^2 y comparten un nugget de 0. Se observan distintos valores de phi en cada modelo. Al utilizar la distancia euclíadiana, los modelos comparten un nugget estimado de 0, y en general valores similares para σ^2 , exceptuando el modelo Matérn, que presenta un comportamiento distinto a todos los otros semivariogramas, principalmente asociado a las estimaciones de ϕ y κ .

Tabla 4.1: Parámetros estimados de los distintos semivariogramas para la distancia Shortest-path y Euclíadiana

Modelo	Nugget	σ^2	κ	ϕ
Shortest-path				
Matern	0	0.544	2.000	1.468
Gaussian	0	0.544	-	3.820
Spherical	0	0.543	-	8.646
Cauchy	0	0.544	14.723	14.249
Euclíadiana				
Matern	0	0.705	0.131	38.358
Gaussian	0	0.616	-	3.595
Spherical	0	0.616	-	8.375
Cauchy	0	0.652	0.522	1.837

Se evaluó el rango práctico para cada uno de los modelos de variograma, para la distancia Euclíadiana todos los modelos presentaron un rango práctico mayor que la distancia shortest-path. Además, el semivariograma Matérn Euclíadiano fue el que presentó el mayor rango práctico de todos los semivariogramas estimados.

Todos los modelos de variogramas ajustados en la distancia shortest-path presentaron rangos prácticos similares para el $\mathbf{Y}(\mathbf{s}) = \log \mathbf{Z}(\mathbf{s})$, mientras que en los semivariogramas estimados por la distancia euclíadiana, tanto el semivariograma Cauchy como el Matérn presentaron diferencias en comparación con el Gaussiano y Esférico.

Tabla 4.2: Comparación de rangos prácticos para los distintos modelos de semivariograma, y para cada métrica de distancia

Variograma	Shortest-Path	Euclidian
Matern	0.5440074	0.7049388
Gaussian	0.5436271	0.6156571
Spherical	0.5433373	0.6158074
Cauchy	0.5436143	0.6524011

Los valores propios de cada una de las matrices Σ generadas por cada métrica de distancia fueron evaluados para corroborar que las matrices fuesen definidas positivas.

Tabla 4.3: Mínimo de los valores propios de Σ para cada modelo y tipo de distancia estimadas

Modelo	Shortest-path	Euclidian
Matern	0.0445322	0.2574403
Gaussian	0.0295108	0.0371616
Spherical	0.0847541	0.0993695
Cauchy	0.0311105	0.0697865

Dado que tanto para la distancia euclíadiana como para el camino más corto y para todos los modelos de variograma los valores propios $\lambda_1, \dots, \lambda_n > 0$, no fue necesario realizar la corrección ClosePD indicada en el Marco teórico 1.

4.3 Interpolación via Kriging

A continuación se presentarán las diferencias entre la predicción interpolada via Kriging Log-normal, y la razón de varianzas de la interpolación. Los mapas de predicción se encuentran en Capítulo 7.

La razón de varianza o ratio de varianza considerado fue:

$$\frac{\text{var}(\text{Shortest path})}{\text{var}(\text{Euclidian})}$$

De esta forma, un ratio de 1 corresponde a que ambas varianzas son iguales, mientras que ratios menores a uno indican que la varianza de la distancia euclíadiana es mayor a la de la distancia shortest-path.

Además, para la comparación a nivel de predicción se consideró $|\text{Krig}_{sp} - \text{Krig}_{euc}|$ como métrica de comparación, dado que lo que es de interés es evaluar la diferencia entre predicciones, sin importar si es negativa o positiva.

Modelo Matérn:

El semivariograma Matérn presentó sus mayores diferencias a nivel de predicción a nivel de distancias intermedias, principalmente debido a la diferencia en el valor de ϕ estimado para la distancia euclíadiana, que indica una dependencia a mayor distancia.

A nivel de varianza, considerando se observa que está centrada en 0.85, donde las mayores diferencias se encuentran en las áreas cercanas a los centros. Se observó que en la interpolación alrededor de los centros cercanos la varianza de la distancia shortest-path es 0.6 veces la varianza de la distancia euclíadiana.

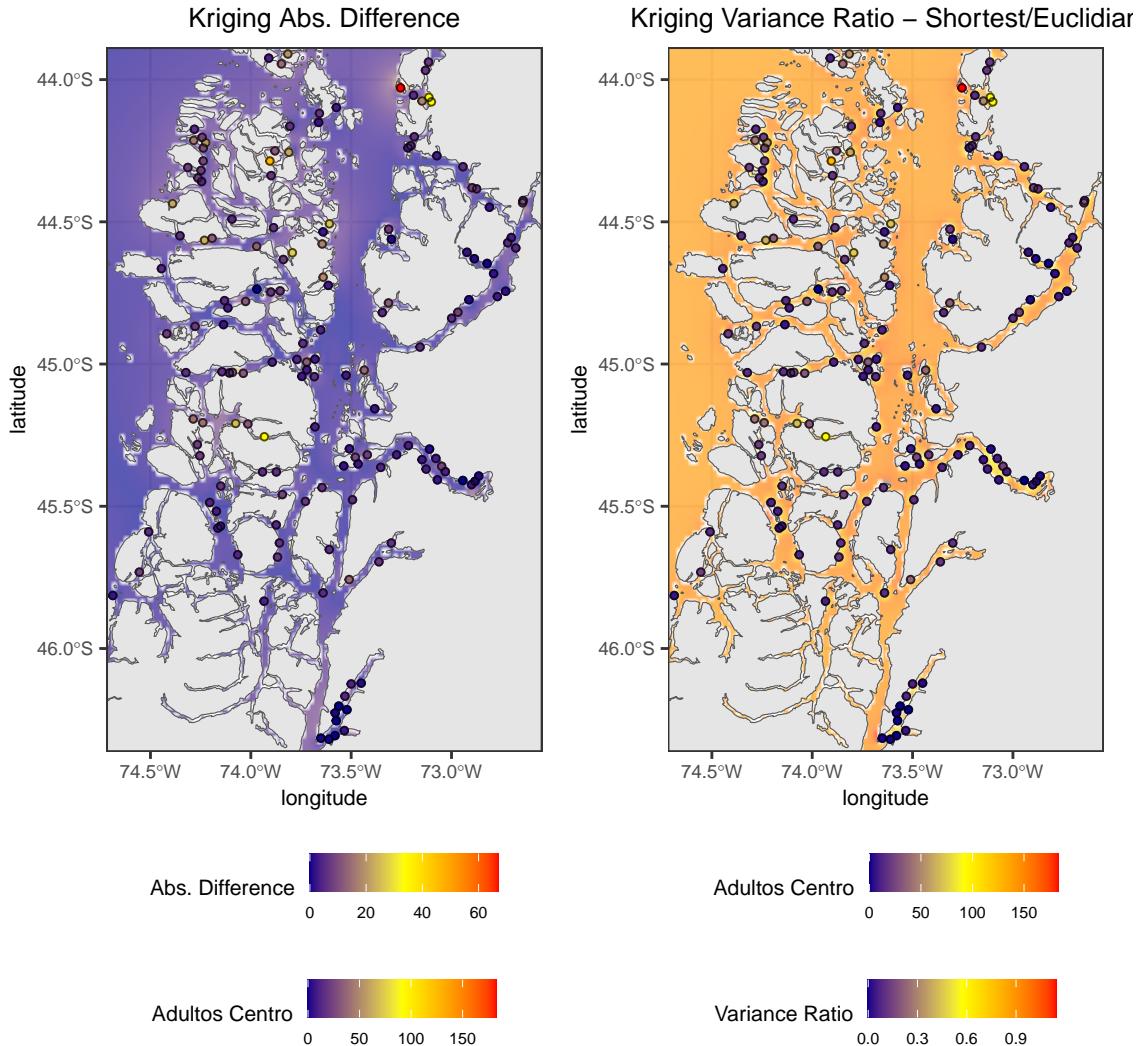


Figura 4.3: Diferencias en la interpolación del modelo Matérn shortest-path y euclidiano, a nivel de Kriging y Varianza de predicción.

Modelo Gaussiano:

El modelo Gaussiano presentó diferencias, principalmente asociadas a distancias cortas en los centros con mayor carga. También se observaron diferencias en las zonas bajo latitud 46 y longitud 74 - 73.5, debido a que la distancia euclídea no considera las barreras en la interpolación.

La varianza estimada se encuentra centrada en 1 en los puntos donde no hay centros observados, mientras que en los puntos que se tiene información, se observa que la varianza de la distancia euclídea es mayor que la distancia shortest-path.

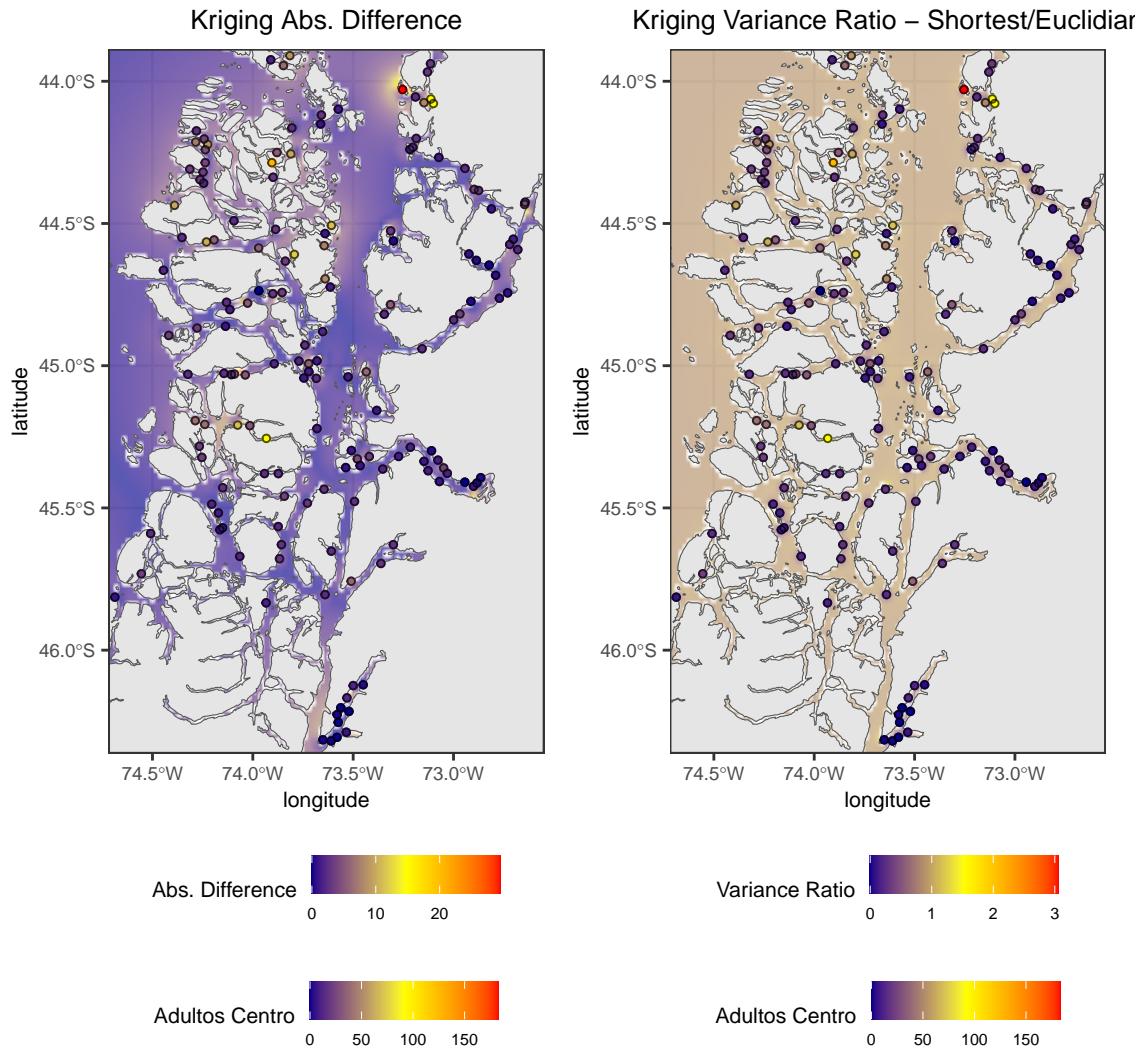


Figura 4.4: Diferencias en la interpolación del modelo Gaussiano shortest-path y euclidiano, a nivel de Kriging y Varianza de predicción.

Modelo Esférico:

En el modelo esférico, se observan mayores diferencias en los centros con mayor carga, y donde se encuentran más barreras o canales, observando diferencias de aproximadamente 10 adultos totales entre las distintas métricas de distancia.

A nivel de varianza, la varianza de la distancia shortest path es 0.8 veces la varianza de la distancia euclíadiana en la interpolación general, mientras que en las localizaciones con más información el ratio de varianzas tiende valores menores a 0.4, indicando que la distancia euclíadiana presenta menor certeza en la predicción.

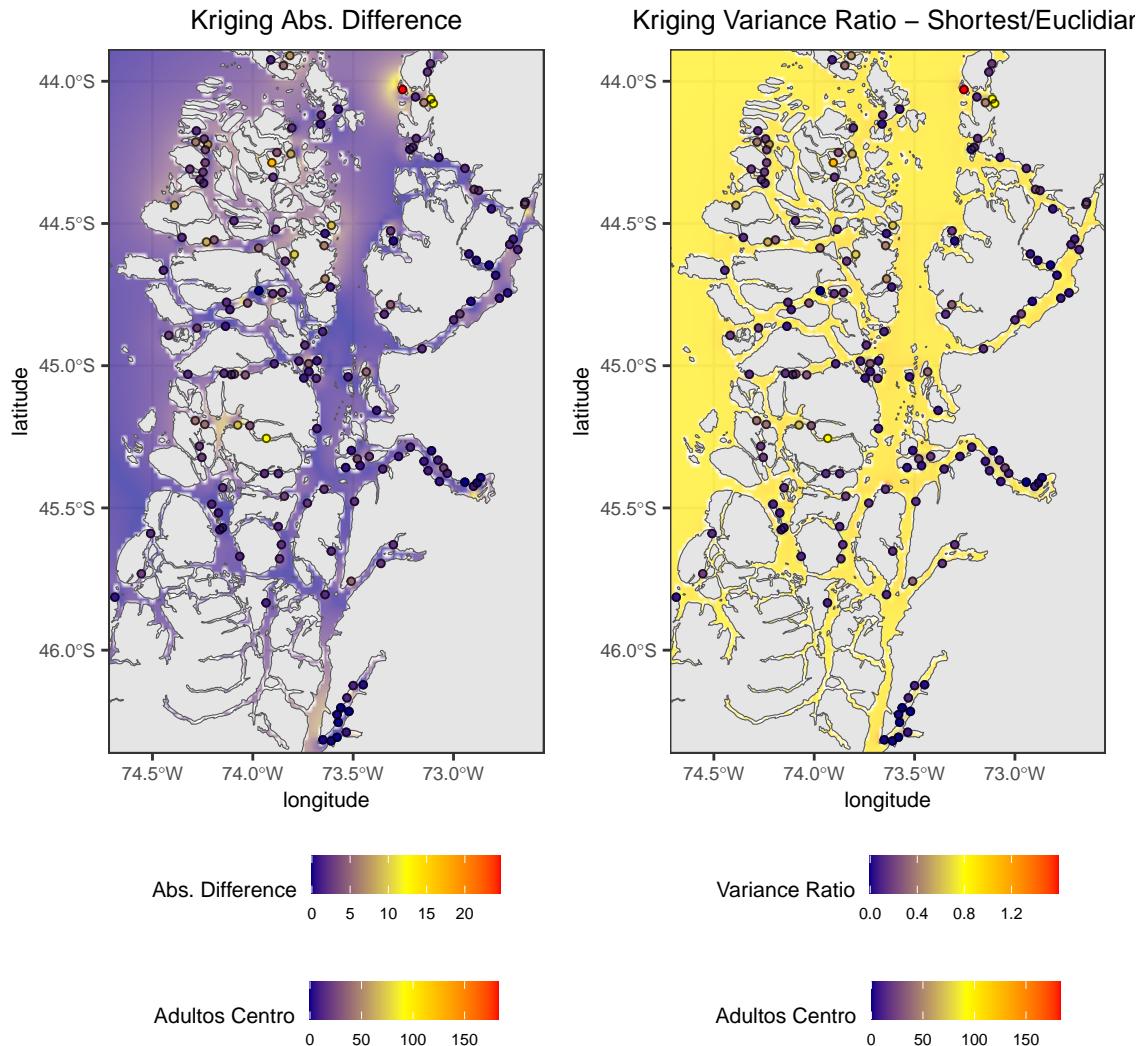


Figura 4.5: Diferencias en la interpolación del modelo Esférico shortest-path y euclidiano, a nivel de Kriging y Varianza de predicción.

Modelo Cauchy:

El modelo Cauchy presenta diferencias similares a los modelos anteriores, en general centrados a nivel de 10 adultos totales exceptuando en los centros con mayor carga. Las zonas en las que hay más diferencias a nivel de predicción son concordantes con lo observado en los modelos Gaussiano y Esférico.

A nivel de Varianza, también se observa un ratio en general centrado en 0.8, principalmente dado por el Sill estimado por cada uno de los modelos. En distancias cortas, la varianza del modelo euclíadiano continúa siendo mayor que la distancia del camino más corto.

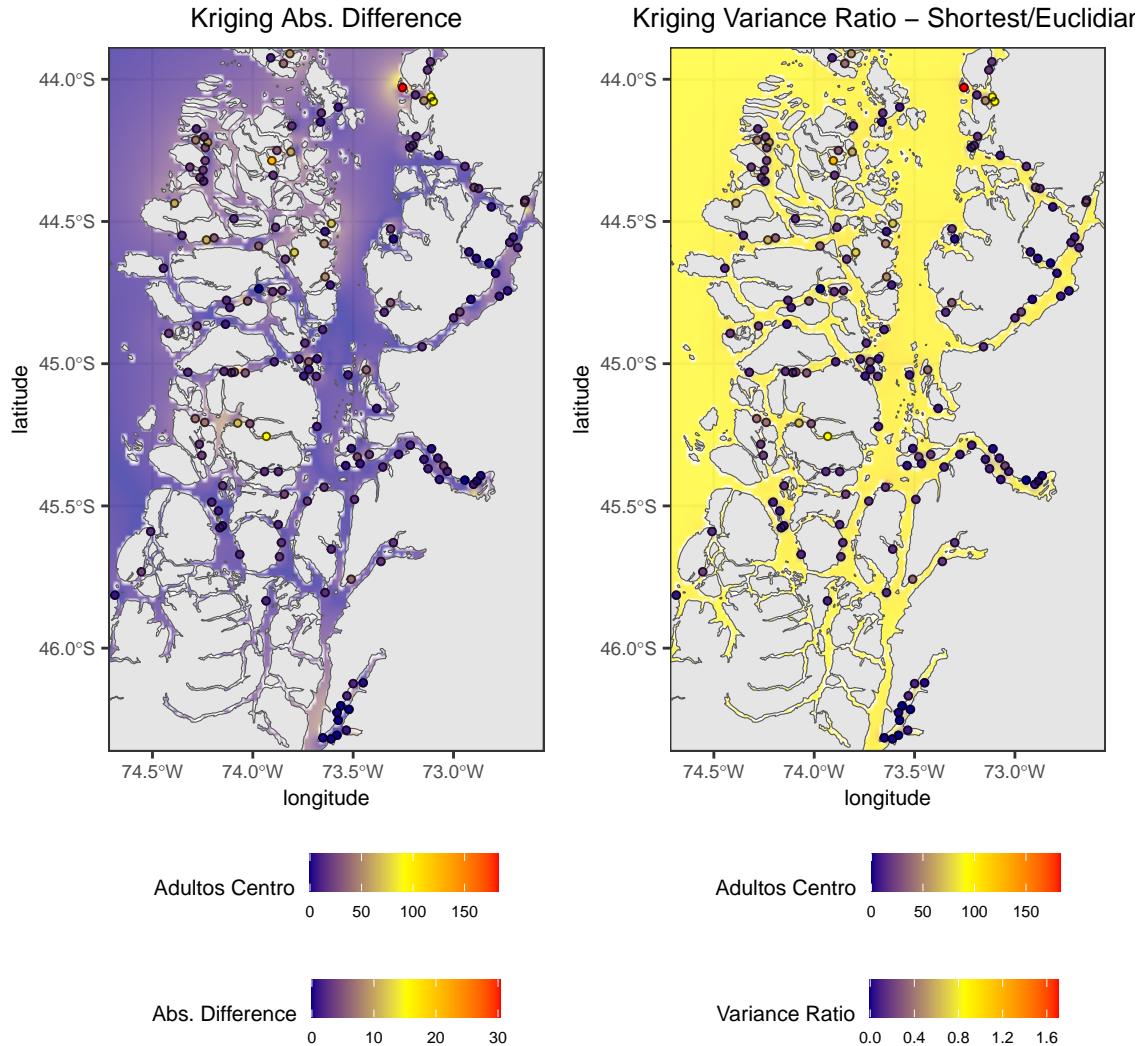


Figura 4.6: Diferencias en la interpolación del modelo Cauchy shortest-path y euclidiano, a nivel de Kriging y Varianza de predicción.

5 Conclusiones

La distancia del camino más corto o shortest-path representó de mejor manera la dependencia espacial que presentaron los centros de cultivo, en particular debido a la reducción de la varianza a nivel del semivariograma empírico y a los Sill teóricos estimados, en comparación con la distancia euclíadiana.

Con respecto a las predicciones a nivel de Kriging entre las distintas métricas de distancia, se observó de forma consistente que independientemente del modelo de semivariograma, la distancia shortest-path presentó menor varianza en los sitios no observados, lo que indica una menor incertidumbre en la predicción.

A nivel de riesgos, ambas distancias permitieron caracterizar las áreas que presentaron una mayor carga parasitaria en general, la diferencia se encuentra principalmente en la magnitud de la predicción.

Se observó en todos los modelos de variograma diferencias a nivel de varianza y de predicción para las distintas métricas de distancia. En particular, la distancia euclíadiana presentó un Sill mayor en todos los modelos ajustados con respecto a la distancia shortest-path. A pesar de esto, ambas métricas indican en general, una relación en distancias más cortas que a nivel regional, sugiriendo un modelo ajustado a nivel de vecinos cercanos.

Se destaca que la utilización de un modelo de vecinos cercanos en vez de un modelo regional podría entregar una mejor caracterización de las zonas con varios centros de cultivo y de su carga parasitaria, y dado que en distancias cortas la distancia euclíadiana es homologable a la shortest-path cuando no hay barreras de por medio, esta puede ser una opción a futuro para evaluar la dependencia espacio-temporal considerando solo los vecinos cercanos.

Por otro lado, a nivel regional los centros se distribuyen en barrios o asociación de concesiones (ACS) en donde estos sitios deben generar un plan de manejo común de caligidosis. Estas ACS fueron definidas por SERNAPESCA como división y ordenamiento general, pero no necesariamente con la consideración de la cercanía o similitud de cargas parasitarias de centros cercanos, o en qué nivel se encuentra esta dependencia espacial.

En el presente trabajo, se observó que la dependencia espacial del promedio de los máximos por ciclo fue principalmente de distancias cortas tanto para los modelos utilizando distancia euclíadiana como la el camino más corto, lo que es un indicio de que el manejo ideal de la carga parasitaria de los centros de cultivo podría no ser a nivel de ACS, sino teniendo en consideración centros con menor separación espacial.

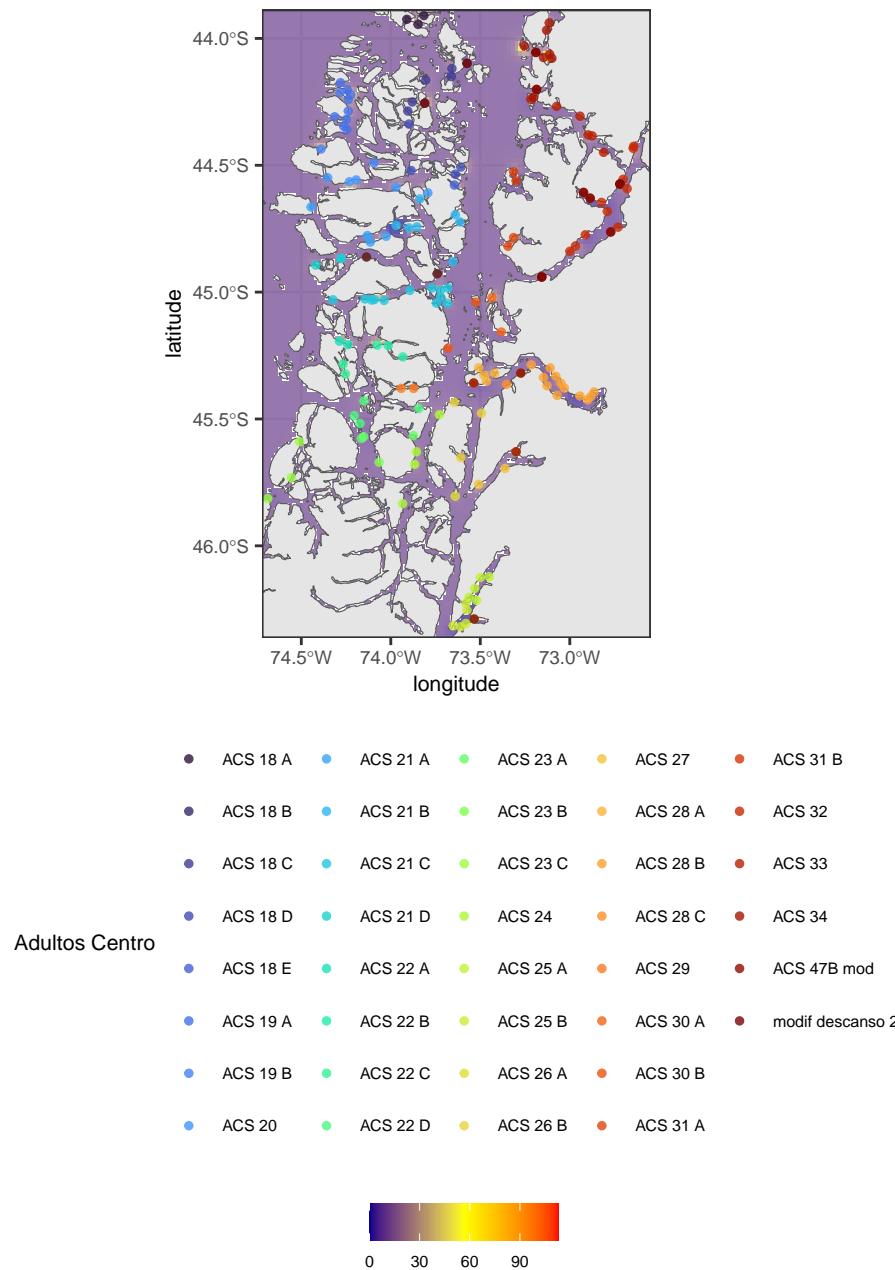


Figura 5.1: Interpolación utilizando Kriging Lognormal Matérn con distancia shortest-path, con los centros y sus distintas ACS.

Dada la dependencia a corta distancia, y la gran cantidad de ACS distintas en centros cercanos, una potencial alternativa al manejo de los centros sería la coordinación de tratamiento entre centros vecinos, a diferencia de la coordinación entre ACS completas.

6 Trabajos futuros

En este trabajo se consideró solamente el promedio de los máximos por ciclo en cada centro, pero la recopilación de los datos corresponde a una serie temporal cero-inflada con puntos en los que se van tratando los centros con distintos productos, por lo tanto las principales extensiones a considerar serían:

- Extensión a Kriging espacio-temporal, para tener en consideración la dependencia espacial y temporal de las cargas parasitarias.
- Cokriging, con el fin de capturar la dependencia entre estadíos juveniles - adultos móviles - hembras ovígeras por separado, aprovechando la asociación entre ellos sin incluirlos como covariables para evitar endogeneidad.
- Kriging Universal, para poder incluir covariables como temperatura, salinidad, tratamientos antiparasitarios y fase lunar, que podrían apotar información relevante para la estimación de las distintas cargas parasitarias.
- Consideración del número de peces y la carga parasitaria, dado que en este trabajo se consideró solo las cargas promedio de adultos totales y no el impacto que tiene cada centro según la cantidad de peces que tiene.

7 Apéndice

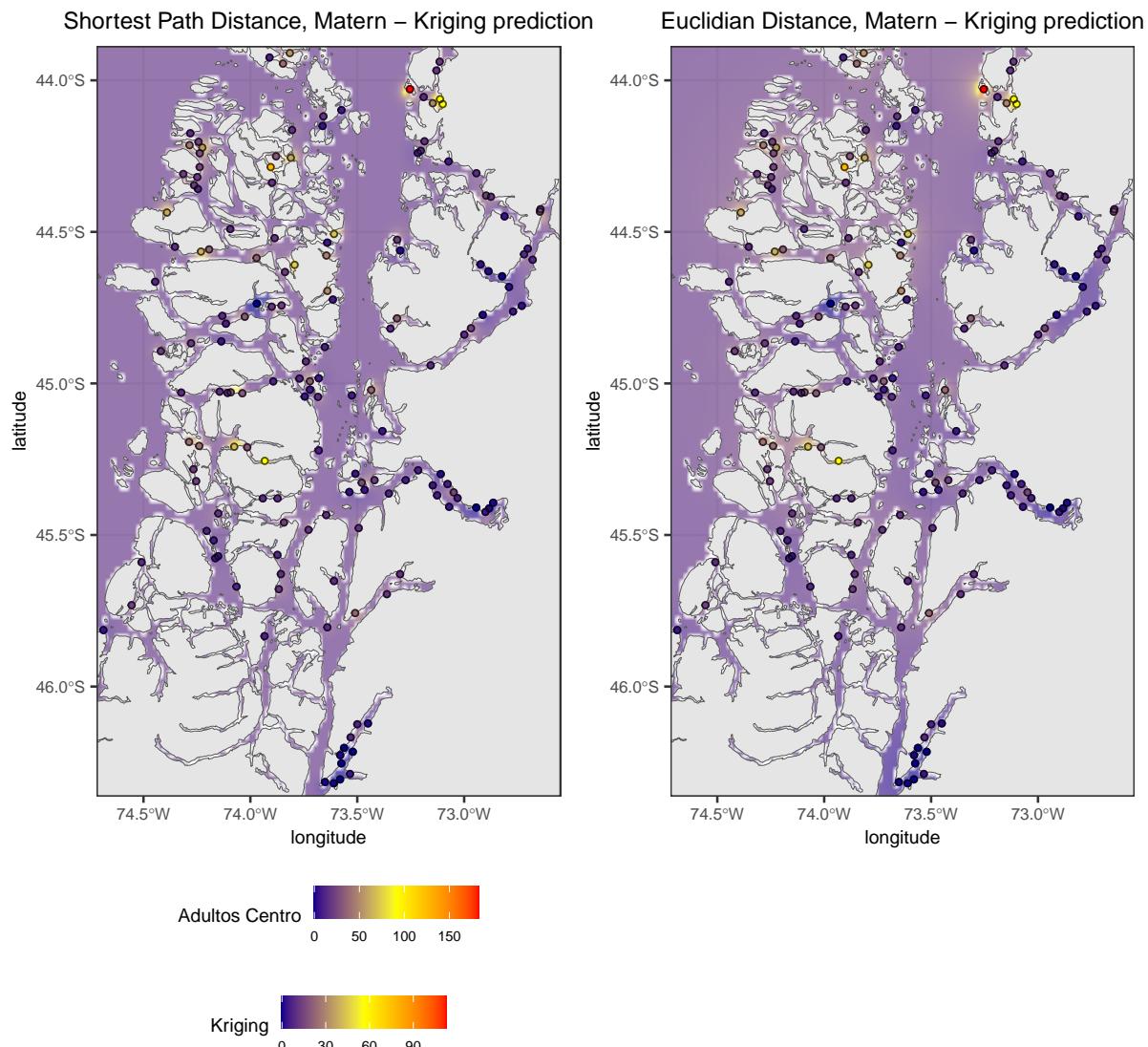


Figura 7.1: Interpolación utilizando Kriging Lognormal Matérn con distancia shortest-path y euclidiana

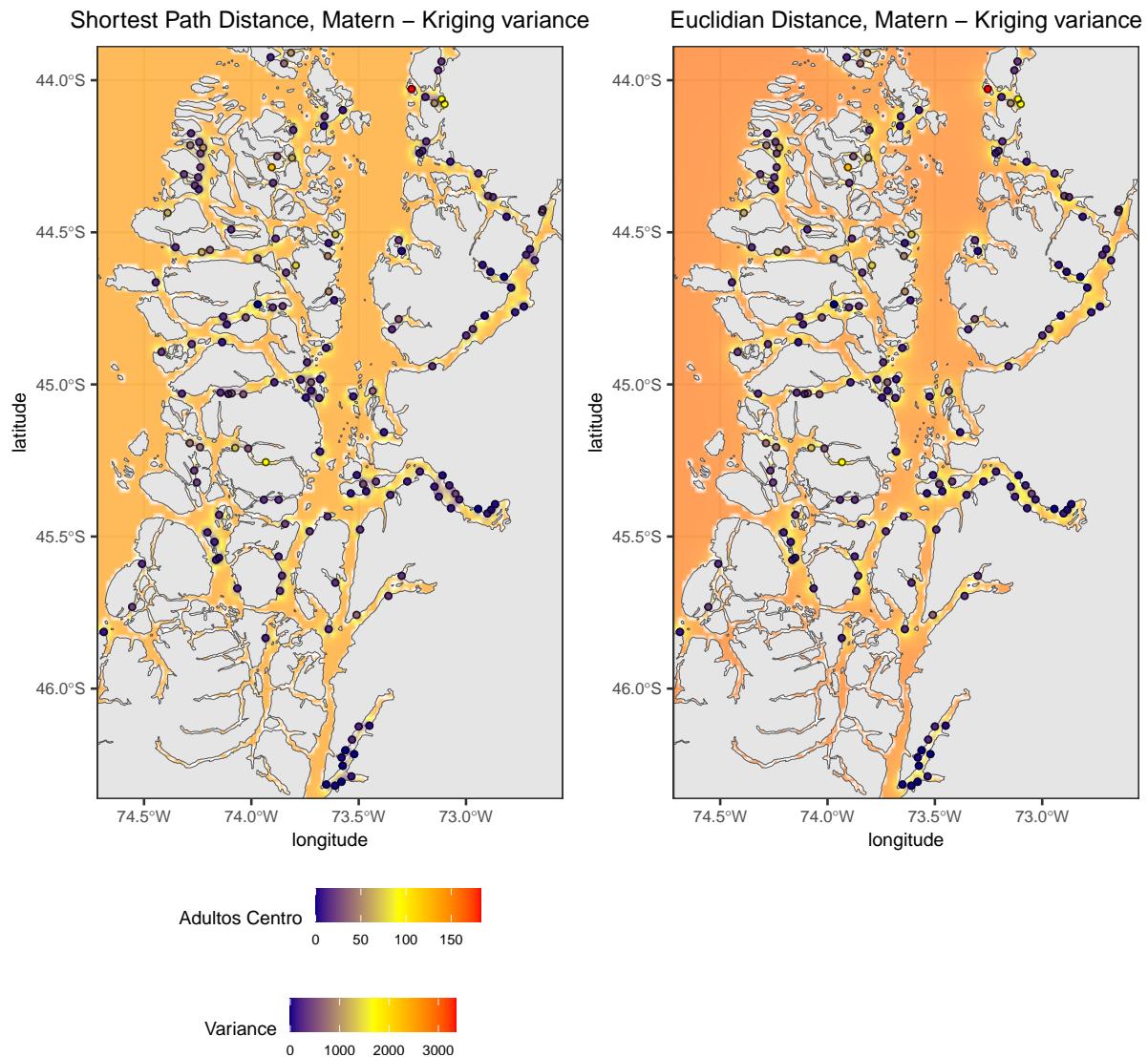


Figura 7.2: Interpolación de varianza utilizando Kriging Lognormal Matérn con distancia shortest-path y euclidiana

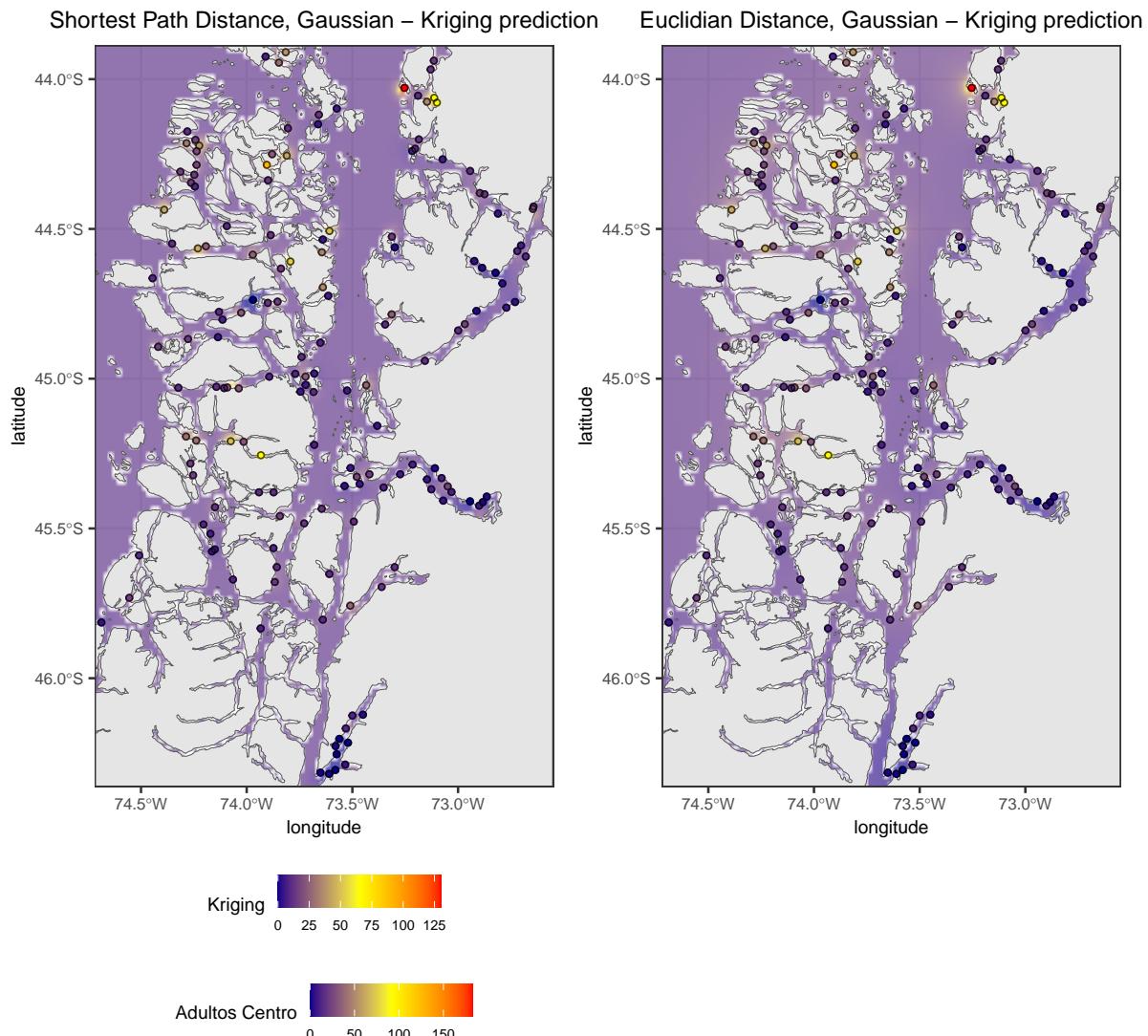


Figura 7.3: Interpolación utilizando Kriging Lognormal Gaussiano con distancia shortest-path y euclíadiana

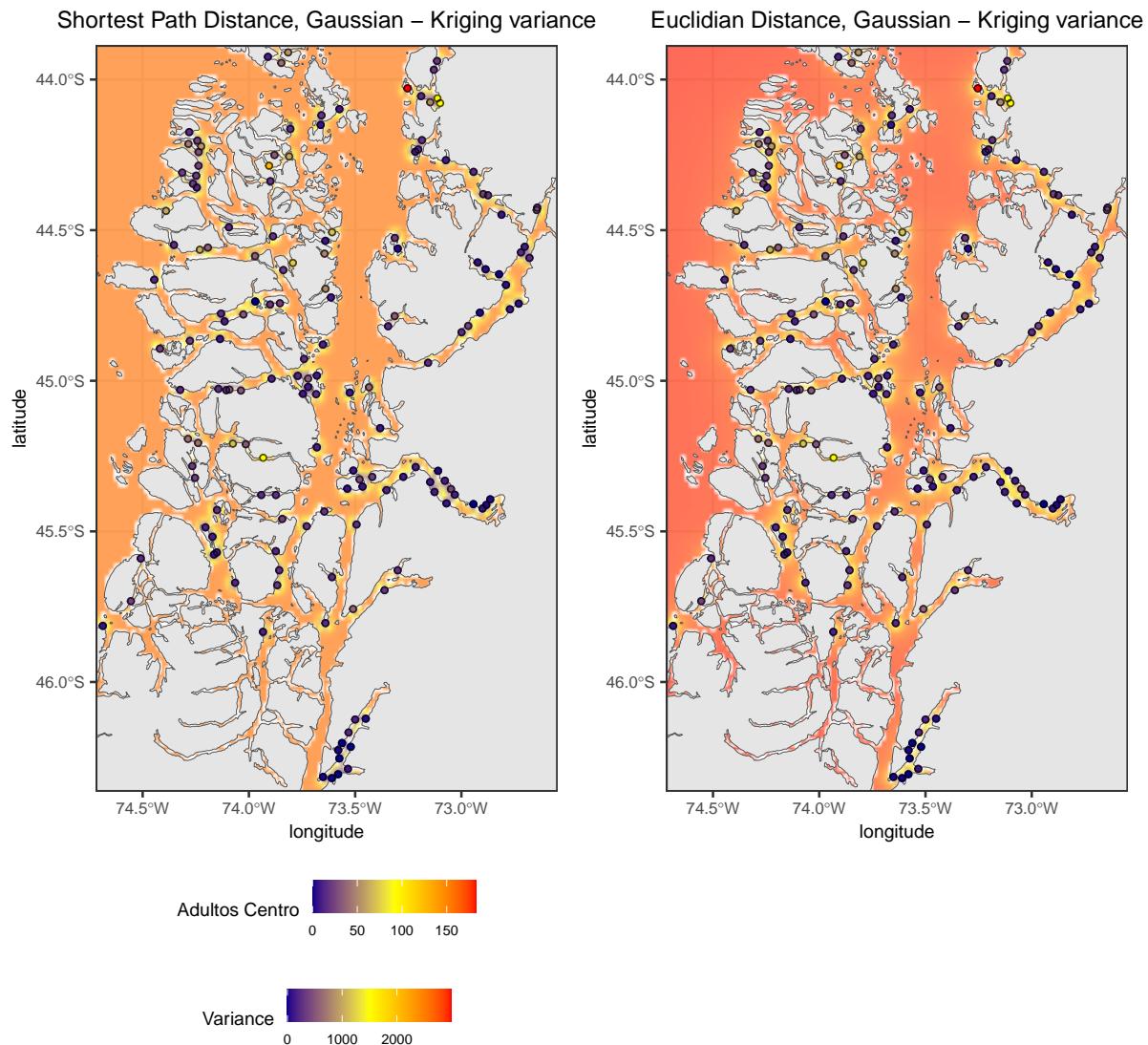


Figura 7.4: Interpolación utilizando Kriging Lognormal Gaussiano con distancia shortest-path y euclidiana

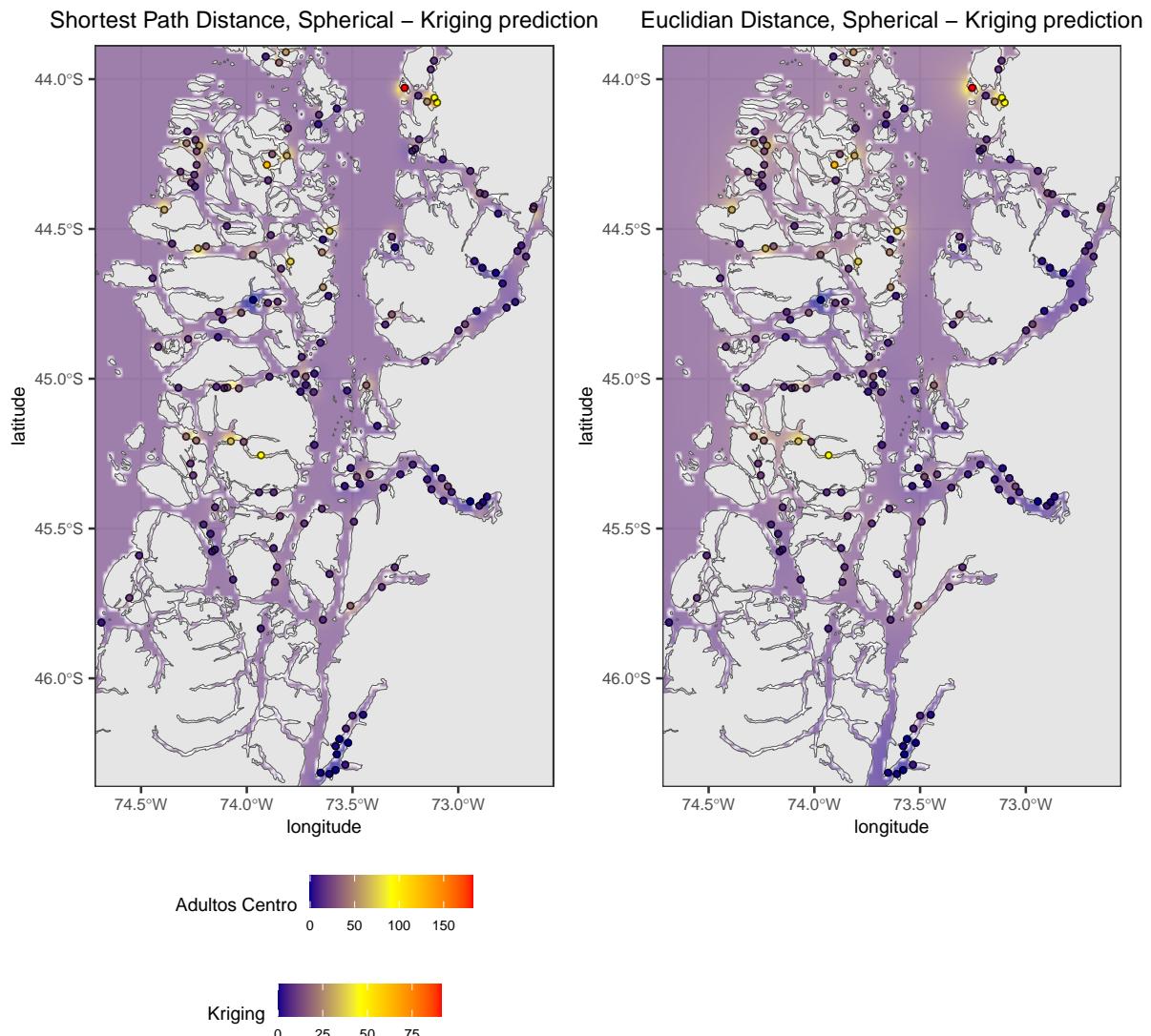


Figura 7.5: Interpolación utilizando Kriging Lognormal Esférico con distancia shortest-path y euclíadiana

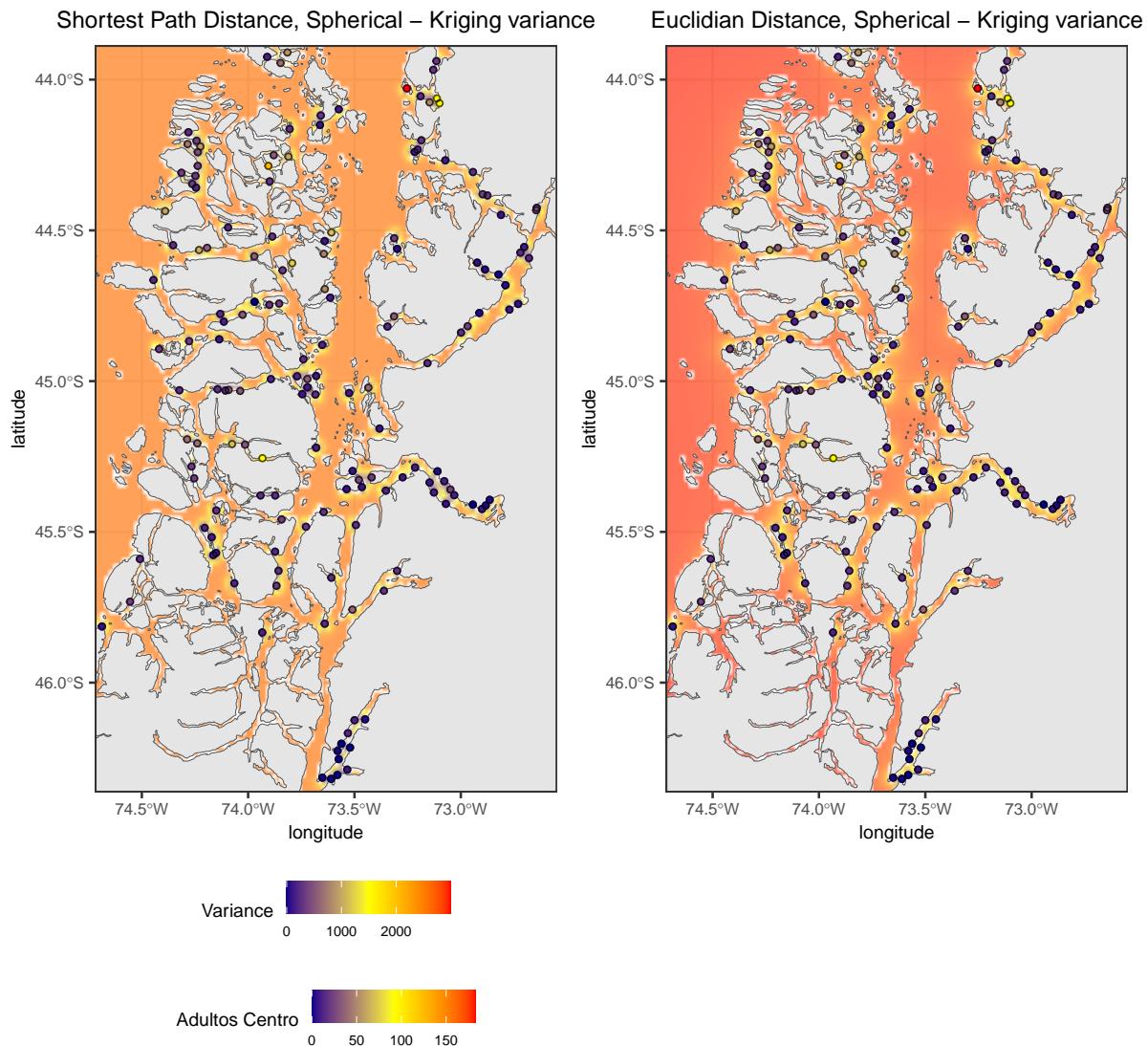


Figura 7.6: Interpolación utilizando Kriging Lognormal Esférico con distancia shortest-path y euclidianas

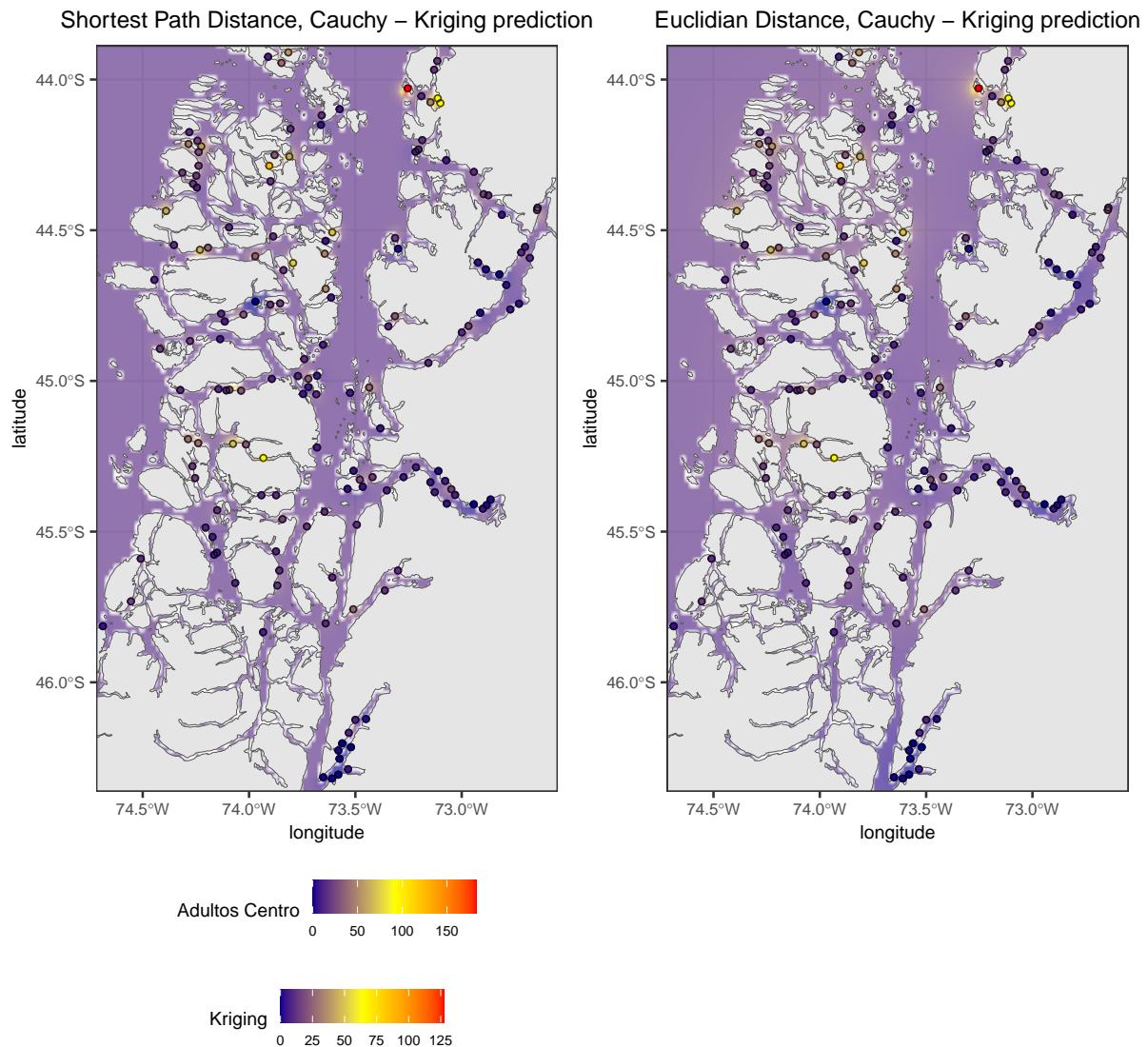


Figura 7.7: Interpolación utilizando Kriging Lognormal Cauchy con distancia shortest-path y euclíadiana

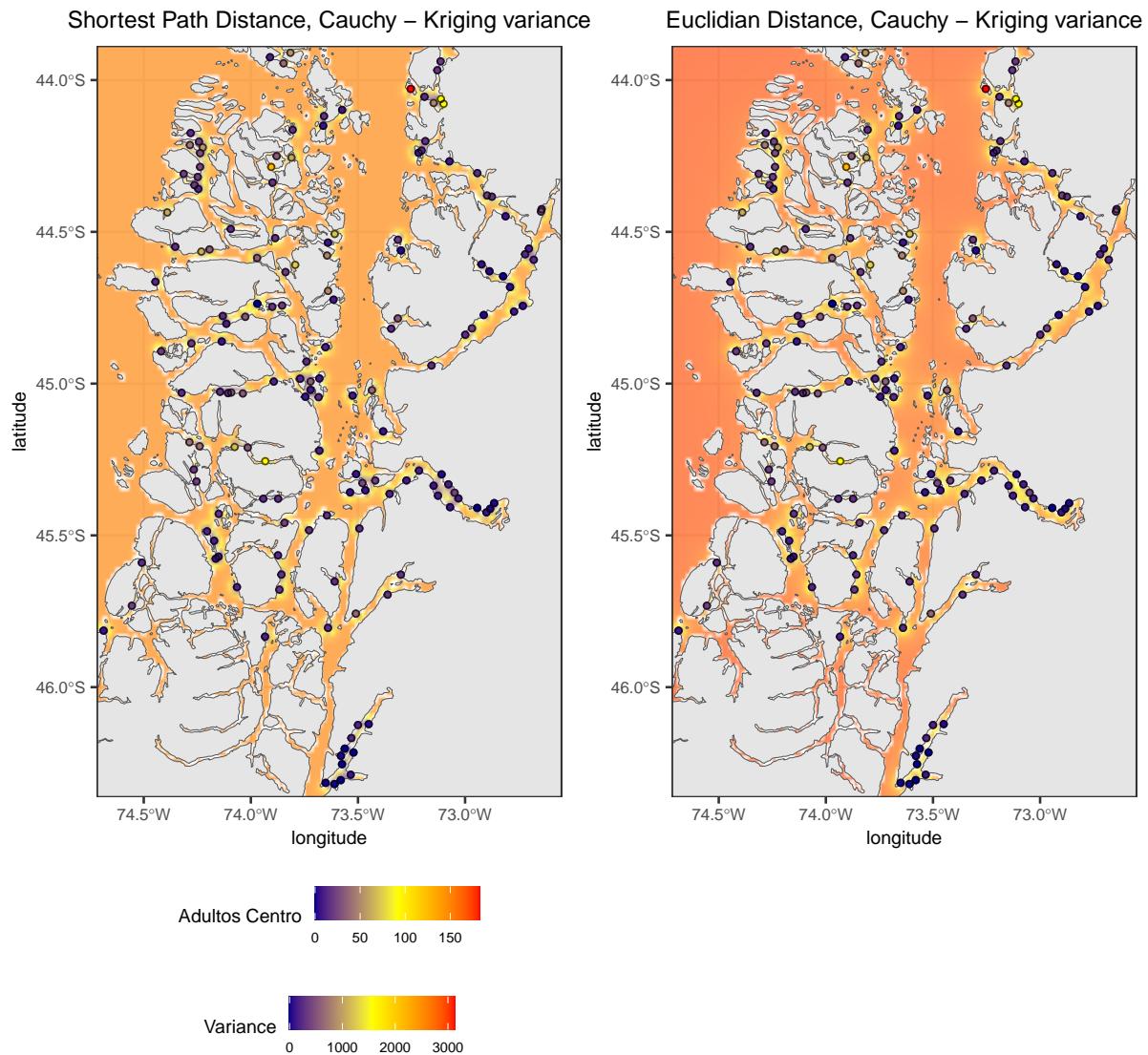


Figura 7.8: Interpolación utilizando Kriging Lognormal Cauchy con distancia shortest-path y euclíadiana

Referencias

1. Iriarte JL, Pantoja S, Daneri G. Oceanographic Processes in Chilean Fjords of Patagonia: From small to large-scale studies. *Progress in Oceanography*. 2014;129:1-7. doi:[10.1016/j.pocean.2014.10.004](https://doi.org/10.1016/j.pocean.2014.10.004)
2. Pantoja S, Luis Iriarte J, Daneri G. Oceanography of the Chilean Patagonia. *Continental Shelf Research*. 2011;31(3):149-153. doi:[10.1016/j.csr.2010.10.013](https://doi.org/10.1016/j.csr.2010.10.013)
3. SERNAPESCA. Anuario Estadístico de Pesca y Acuicultura 2023 - Chile, cosechas de centros de acuicultura por especie y región, 2023. Published online junio de 2024. <https://www.sernapesca.cl/informacion-utilidad/anuarios-estadisticos-de-pesca-y-acuicultura/>
4. Bravo S, Silva MT, Agusti C, Ponce N. Adult population structure of Caligus rogercresseyi after pharmacological treatments. *Aquaculture*. 2024;586:740788. doi:[10.1016/j.aquaculture.2024.740788](https://doi.org/10.1016/j.aquaculture.2024.740788)
5. Handeland SO, Imsland AK, Stefansson SO. The effect of temperature and fish size on growth, feed intake, food conversion efficiency and stomach evacuation rate of Atlantic salmon post-smolts. *Aquaculture*. 2008;283(1):36-42. doi:[10.1016/j.aquaculture.2008.06.042](https://doi.org/10.1016/j.aquaculture.2008.06.042)
6. Idenyi JN, Abdallah H, Adeyemi AD, et al. Optimizing growth and mitochondrial function in rainbow trout, *Oncorhynchus mykiss* through eco-friendly dietary and changes in water temperature regimen strategies. *Aquaculture*. 2025;595:741591. doi:[10.1016/j.aquaculture.2024.741591](https://doi.org/10.1016/j.aquaculture.2024.741591)
7. Jiang X, Dong S, Liu R, et al. Effects of temperature, dissolved oxygen, and their interaction on the growth performance and condition of rainbow trout (*Oncorhynchus mykiss*). *Journal of Thermal Biology*. 2021;98:102928. doi:[10.1016/j.jtherbio.2021.102928](https://doi.org/10.1016/j.jtherbio.2021.102928)
8. SERNAPESCA. Programa Sanitario Específico de Vigilancia y Control de Caligosis. Published online febrero de 2022. https://www.sernapesca.cl/app/uploads/2023/11/res.ex_.60-2022.pdf

9. Bravo S. The reproductive output of sea lice *Caligus rogercresseyi* under controlled conditions. *Experimental Parasitology*. 2010;125(1):51-54. doi:[10.1016/j.exppara.2009.12.001](https://doi.org/10.1016/j.exppara.2009.12.001)
10. Carvajal J, González Poblete L, George-Nascimento M. Native sea lice (Copepoda: Caligidae) infestation of salmonids reared in netpen systems in southern Chile. *Aquaculture*. 1998;166:241-246. doi:[10.1016/S0044-8486\(98\)00301-9](https://doi.org/10.1016/S0044-8486(98)00301-9)
11. Agusti C, Bravo S, Contreras G, et al. Sensitivity assessment of *Caligus rogercresseyi* to anti-louse chemicals in relation to treatment efficacy in Chilean salmonid farms. *Aquaculture*. 2016;458:195-205. doi:[10.1016/j.aquaculture.2016.03.006](https://doi.org/10.1016/j.aquaculture.2016.03.006)
12. Bravo S, Treasurer J, Sepulveda M, Lagos C. Effectiveness of hydrogen peroxide in the control of *Caligus rogercresseyi* in Chile and implications for sea louse management. *Aquaculture*. 2010;303(1):22-27. doi:[10.1016/j.aquaculture.2010.03.007](https://doi.org/10.1016/j.aquaculture.2010.03.007)
13. Aldrin M, Storvik B, Kristoffersen AB, Jansen PA. Space-Time Modelling of the Spread of Salmon Lice between and within Norwegian Marine Salmon Farms. *PLoS ONE*. 2013;8(5). doi:[10.1371/journal.pone.0064039](https://doi.org/10.1371/journal.pone.0064039)
14. Elghafghuf A, Vanderstichel R, Hammell L, Stryhn H. Estimating sea lice infestation pressure on salmon farms: Comparing different methods using multivariate state-space models. *Epidemics*. 2020;31:100394. doi:[10.1016/j.epidem.2020.100394](https://doi.org/10.1016/j.epidem.2020.100394)
15. Rees EE, St-Hilaire S, Jones SRM, et al. Spatial patterns of sea lice infection among wild and captive salmon in western Canada. *Landscape Ecology*. 2015;30(6):989-1004. doi:[10.1007/s10980-015-0188-2](https://doi.org/10.1007/s10980-015-0188-2)
16. Cressie NAC. Applications of Geostatistics. En: *Statistics for Spatial Data*. John Wiley & Sons, Ltd; 1993:211-275. doi:[10.1002/9781119115151.ch4](https://doi.org/10.1002/9781119115151.ch4)
17. Sahu S. Bayesian Modeling of Spatio-Temporal Data with R. *Routledge & CRC Press*. Published online 2022. Accedido noviembre 1, 2024. <https://www.routledge.com/Bayesian-Modeling-of-Spatio-Temporal-Data-with-R/Sahu/p/book/9781032209579>
18. Banerjee S, Carlin BP, Gelfand AE. *Hierarchical Modeling and Analysis for Spatial Data*. 2.^a ed. Chapman; Hall/CRC; 2014. doi:[10.1201/b17115](https://doi.org/10.1201/b17115)
19. Cressie N. Fitting variogram models by weighted least squares. *Journal of the International Association for Mathematical Geology*. 1985;17(5):563-586. doi:[10.1007/BF01032109](https://doi.org/10.1007/BF01032109)

20. Schabenberger O, Gotway CA. *Statistical Methods for Spatial Data Analysis*. Chapman; Hall/CRC; 2017. doi:[10.1201/9781315275086](https://doi.org/10.1201/9781315275086)
21. Ver Hoef JM. Kriging models for linear networks and non-Euclidean distances: Cautions and solutions. *Methods in Ecology and Evolution*. 2018;9(6):1600-1613. doi:[10.1111/2041-210X.12979](https://doi.org/10.1111/2041-210X.12979)
22. Van Etten J. Package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software*. 2017;76(13). doi:[10.18637/jss.v076.i13](https://doi.org/10.18637/jss.v076.i13)
23. Davis BJK, Curriero FC. Development and Evaluation of Geostatistical Methods for Non-Euclidean-Based Spatial Covariance Matrices. *Mathematical geosciences*. 2019;51(6):767. doi:[10.1007/s11004-019-09791-y](https://doi.org/10.1007/s11004-019-09791-y)
24. Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*. 2018;10(1):439. doi:[10.32614/RJ-2018-009](https://doi.org/10.32614/RJ-2018-009)
25. Pardo-Iguzquiza E, Chica-Olmo M. Geostatistics with the Matern semivariogram model: A library of computer programs for inference, kriging and simulation. *Computers & Geosciences*. 2008;34(9):1073-1079. doi:[10.1016/j.cageo.2007.09.020](https://doi.org/10.1016/j.cageo.2007.09.020)