



Facultad de Economía y Negocios  
Escuela de Ingeniería Informática Empresarial

**PREDICIENDO LAS CONSECUENCIAS DE LOS INCENDIOS FORESTALES  
SOBRE LOS AGENTES POLINIZADORES: UN ENFOQUE DESDE LA  
MINERÍA DE DATOS**

Autores: Camilo Andrés Rodríguez Villagra

Marcelo Andrés Hurtado Morales

Prof. guía: Eduardo Álvarez

Proyecto de Memoria para optar al título de INGENIERO INFORMÁTICO  
EMPRESARIAL

TALCA – CHILE

2022

## CONSTANCIA

La Dirección del Sistema de Bibliotecas a través de su unidad de procesos técnicos certifica que el autor del siguiente trabajo de titulación ha firmado su autorización para la reproducción en forma total o parcial e ilimitada del mismo.



Talca, 2023

## **Agradecimientos**

### **De Marcelo Andrés Hurtado Morales**

En primera instancia me gustaría agradecer a mi familia, quienes me han brindado todo el amor, comprensión y apoyo posible, tanto hoy como antes de entrar a esta carrera, tanto en las buenas como en las malas. Agradecer también a mi pareja, Francisca, un pilar fundamental para mí en estos últimos años, que me ayudó a no rendirme cuando pensé en hacerlo.

A Eduardo Álvarez, por ser además de nuestro profesor guía, un mentor para mí y mi compañero. No sólo nos ayudó cumpliendo con sus responsabilidades como profesor, también nos entregó enseñanzas que nos servirán para el desarrollo de nuestras vidas como profesionales y apoyo incluso cuando no le correspondía.

A Yony Ormazábal, por su disposición y atención en relación con nuestras dudas, consultas y estar presente cada vez que le necesitábamos.

Finalmente, a nuestros colaboradores: Camilo Matus, Ian acuña, Fulgencio Lisón y Jaime Carrasco, quienes desde un principio estuvieron allí, apoyándonos y dándonos la información necesaria para desenvolvemos en un proyecto de tal magnitud. Sin ellos, nada de esto hubiera sido posible.

## **De Camilo Rodríguez**

El camino para llegar a este punto ha sido largo y arduo es por ello que en primer lugar quiero agradecer a mi madre, a mis tíos y primos, los cuales me han apoyado a lo largo de toda mi trayectoria académica y vida. Quiero agradecer también a Valentina Muñoz mi pareja que me impulsó a hacer las cosas de la mejor manera siempre.

Agradezco también la colaboración de mi compañero Marcelo Hurtado, con quien ha sido un agrado trabajar tanto durante este proceso de tesis, así como a mis amigos y compañeros de generación, quienes siempre con la mejor de las actitudes me animaron en los momentos más estresantes de este proceso.

Asimismo no es posible dejar fuera de estos agradecimientos a quien cumplió el rol de profesor guía, el Dr. Eduardo Álvarez, quién no solo nos guio y apoyo de la mejor forma posible a pesar de su apretado horario y escaso tiempo, si no que nos apoyó a través de la fundación Mi Casa, dirigida por el mismo.

Sumado a lo anterior, es innegable el aporte a esta tesis del profesor Yony Ormazábal, quien fue más allá de los deberes correspondidos como profesor informante y nos aconsejó, incluso cuando no se lo exigía su puesto.

Para finalizar hay muchas personas que nos han brindado su apoyo, conocimiento y recomendaciones en esta tesis, desde sus áreas de expertiz, siendo estas nuevas para nosotros y sin este apoyo el desarrollo de esta memoria hubiese sido mucho más largo y difícil, quizá incluso no hubiese podido ser terminado, estas personas son: Fulgencio Lisón, Ian Acuña, Jaime Carrasco y Camilo Matus.

## Índice de Contenidos

<b>Agradecimientos</b> .....	2
<b>Resumen</b> .....	8
<b>Abstract</b> .....	9
<b>I. Introducción</b> .....	10
<b>II. Marco Teórico</b> .....	12
2.1. Servicios ecosistémicos.....	12
2.1.2. Polinización como servicio.....	12
2.1.3. Incendios forestales en Chile.....	13
2.1.4. Distribución espacial de incendios en Chile.....	13
2.1.5. Disrupción del servicio por incendios.....	13
2.2. Sistemas de Información Geográfica (SIG).....	14
2.2.1. ¿Qué son los Sistemas de Información Geográfica?.....	14
2.2.2. Categorías de SIG.....	14
2.3. Modelos predictivos.....	15
2.3.1. ¿Qué son los algoritmos de minería de datos?.....	15
2.3.2. ¿Qué son los métodos predictivos?.....	15
2.3.3. Clasificaciones de modelos predictivos.....	16
2.3.4. Aprendizaje supervisado y no supervisado.....	16
<b>III. Metodología</b> .....	22
3.1. Obtención de datos de polinizadores.....	22
3.1.1. Obteniendo la lista de polinizadores en Chile.....	22
3.1.2. Obteniendo los registros de avistamiento.....	22
3.2. Tratamiento de datos de polinizadores.....	23
3.2.1. Unificando los datos.....	23
3.2.2. Eliminando duplicados.....	23
3.2.3. Eliminando coordenadas de no interés.....	24
3.2.4. Creando el formato espacial.....	24
3.3. Obtención y tratamiento de datos predictores ambientales.....	25
3.3.1. Descargando datos de predictores.....	25

3.3.2. Tratando datos de predictores .....	25
3.4. Desarrollo de modelos .....	26
3.5. Generación de mapas .....	27
3.5.1. Obteniendo el mapa ideal.....	27
3.5.2. Cruzando el mapa ideal con el mapa de incendios.....	28
3.6. Herramienta de visualización.....	28
3.6.1. Definiendo las librerías y lenguajes a utilizar.....	28
3.6.2. Mostrando los archivos.....	28
<b>IV. Presentación y análisis de resultados .....</b>	<b>30</b>
4.1. Primeros mapas de observaciones .....	30
4.1.1. Mapas de puntos de observaciones .....	30
4.1.2. Mapas de presencias potenciales (simulaciones).....	31
4.1.3. Mapas Árboles de Decisión.....	35
4.1.4. Mapas MaxEnt .....	36
4.2. Mapa potencial, observado y resultados con incendios (nueva capa de incendios).....	39
4.3. Herramienta de visualización para la toma de decisiones .....	50
<b>V. Conclusiones .....</b>	<b>54</b>
<b>VI. Bibliografía.....</b>	<b>57</b>
<b>Anexos .....</b>	<b>61</b>

## Índice de Tablas

Tabla 1. Ejemplo de datos.....	23
Tabla 2. Ejemplo de datos inválidos.....	24
Tabla 3. Ejemplo con datos climáticos.....	25
Tabla 4. Resultados AUC.....	31
Tabla 5. AUC finales.....	37
Tabla 6. Hábitat perdido por especie.....	48

## Índice de Figuras

Figura 1. Aprendizaje Supervisado.....	17
Figura 2. Árboles de decisión. ....	18
Figura 3. Ejemplo KNN. ....	19
Figura 4. Aprendizaje No Supervisado. ....	20
Figura 5. Ejemplo de Clustering. ....	21
Figura 6. Primeras funciones de herramienta de visualización. ....	29
Figura 7. Mapa de observaciones. ....	30
Figura 8. Mapa de distribución potencial <i>Sephanoides sephanoides</i> con Random forest.....	35
Figura 9. Mapa de distribución potencial <i>Sephanoides sephanoides</i> con MaxEnt. ....	36
Figura 10. Gráficas AUC MaxEnt. ....	37
Figura 11. Mapa de distribución observada. ....	40
Figura 12. Mapa de riqueza de especies.....	41
Figura 13. Mapa de efectos de los incendios sobre el hábitat de los polinizadores. ....	42
Figura 14. Mapa de área afectada por los incendios.....	43
Figura 15. Mapa de efectos graves de los incendios sobre el hábitat de polinizadores. ....	44
Figura 16. Especies más afectadas por los incendios según hábitat total.....	45
Figura 17. Mapa hábitat potencial de los polinizadores sin efectos directos de incendios. ....	46
Figura 18. Mapa con zonas de recurrencia de incendios.....	47
Figura 19. Hábitat perdido por los polinizadores anualmente desde 2000 hasta 2018.....	50
Figura 20. Inicio de herramienta de visualización.....	51
Figura 21. Ejemplo 1 de selección en la herramienta.....	51
Figura 22. Ejemplo 2 de selección en la herramienta.....	52
Figura 23. Ejemplo 3 de selección en la herramienta.....	53
Figura 24. Ejemplo 4 de selección en la herramienta.....	53

## **Resumen**

Durante los últimos 20 años se han incrementado los incendios forestales en nuestro país. De estos eventos se han identificado y contabilizado las pérdidas relacionadas en terrenos agrícolas, recursos forestales, zonas residenciales, entre otros, pero hasta ahora no se ha consolidado una estimación de las pérdidas de hábitat y servicios ecosistémicos presentes en ellos. En este contexto, el objetivo del presente trabajo es generar e implementar una herramienta predictiva que permita estimar la presencia de organismos y agentes polinizadores, y la degradación en la población como consecuencia de incendios forestales. Para esto, se analizó y utilizó la información disponible en las bases de datos de plataformas y organizaciones que tienen como finalidad divulgar información acerca de entidades polinizadoras, junto con algoritmos predictivos de Machine Learning evaluar la distribución y riqueza potencial de zonas centrales de nuestro país.

Como producto de este estudio se obtuvieron una serie de mapas y datos, mediante los cuales se pudo estimar el área promedio perdida por incendios forestales para los polinizadores. Este resultado se produjo gracias a utilización de un algoritmo de minería de datos que identifica las zonas que tienen las condiciones adecuadas para que una especie los habite, de acuerdo a las condiciones ambientales de dicha zona. De esta forma, el área total habitable para las especies polinizadoras se estimó en 186450 km<sup>2</sup>, de los cuales 8000 km<sup>2</sup> aproximadamente han sido afectados por incendios, lo que equivalente a un 4,2% del total.

## **Abstract**

During the last 20 years forest fires have increased in our country. For these events related losses agricultural land, forest resources, residential areas, among others, have been identified and accounted, but until now an estimate of the losses of habitat and ecosystem services present in them has not been consolidated. In this context, the objective of this work is to generate and implement a predictive tool that allows estimating the presence of pollinating organisms and agents, and degradation in the population as a result of forest fires. For this, the information available in the databases of platforms and organizations that aim to disseminate information about pollinating entities was analyzed and used, together with predictive algorithms of Machine Learning to evaluate the distribution and potential wealth of central areas of our country.

As a product of this study, a series of maps and data were obtained, through which it was possible to estimate the average area lost by forest fires for pollinators. This result was produced thanks to the use of a data mining algorithm that identifies the areas that have the appropriate conditions for a species to inhabit them, according to the environmental conditions of said area. In this way, the total habitable area for pollinating species was estimated at 186450 km<sup>2</sup>, of which approximately 8000 km<sup>2</sup> have been affected by fires, equivalent to 4.2% of the total.

## **I. Introducción**

En Chile cada vez es más común la ocurrencia de incendios forestales, en especial en época de verano, propiciados por las alzas de temperatura y el poco cuidado de la gente. Estos siniestros no solo afectan las masas forestales, sino que también a los organismos y agentes polinizadores del ecosistema (Brown et al., 2017), lo que a su vez puede afectar cultivos cercanos que dependan de esto u ecosistemas enteros, dando como resultado la pérdida de cultivos o de biodiversidad.

En la actualidad no existen estudios en Chile respecto de los efectos de los incendios forestales en la población de agentes polinizadores, por tanto, este crucial aspecto no es incluido en los actuales planes preventivos para campos de cultivo agrícola y de mitigación de incendios forestales. En consecuencia, no se consideran de forma adecuada o se desconocen las eventuales pérdidas de biodiversidad y de productividad de las zonas agrícolas cercanas a dichos eventos y que podrían estar relacionadas con la pérdida en los agentes polinizadores.

Si bien existe información respecto a los agentes polinizadores y se encuentra actualmente disponible, está distribuida en distintas partes. Por esto, una etapa inicial del presente proyecto de Memoria fue buscar asesoría en distintas entidades universitarias, de manera de obtener apoyo científico y técnico del tema, además de información disponible y generada en otros estudios. A partir de esto, se realiza el presente proyecto que apunta al diseño e implementación de una herramienta predictiva que permita la estimación de la presencia y degradación de la población de agentes polinizadores como consecuencia de los incendios forestales en Chile.

Para el diseño de dicha herramienta, se consideraron registros de incendios forestales ocurridos en diferentes zonas en Chile, incluyendo bases de datos de información histórica respecto a su distribución espacial y de las características ecológicas de población y presencia de organismos y agentes polinizadores. Dichos datos fueron proporcionados por investigadores de la Universidad de Concepción, y se utilizan en el diseño e implementación de un algoritmo predictivo de tipo supervisado que permite estimar los efectos que tendría la ocurrencia de un incendio forestal, de ciertas características y en una

zona determinada, en la biodiversidad de agentes polinizadores del área de ocurrencia del evento.

Por tanto, el objetivo general del presente trabajo es generar e implementar una herramienta predictiva que permita estimar la presencia de organismos y agentes polinizadores, y la degradación en la población como consecuencia de incendios forestales. Como objetivos específicos, se plantean:

- Analizar la información disponible en las bases de datos de los investigadores de la Universidad de Concepción.
- Seleccionar un algoritmo de minería de datos, basándose en su desempeño prediciendo cómo son afectados los entes polinizadores.
- Diseñar una herramienta de visualización de los resultados y predicciones del modelo para un determinado incendio.

Los resultados que genera esta herramienta, podrían ser utilizados para mejorar el diseño de planes preventivos y de mitigación de incendios forestales. Adicionalmente, con el desarrollo de esta solución se espera motivar a más profesionales e investigadores a interesarse en el área de los polinizadores y los incendios forestales, así como fomentar y aportar a la generación de planes de contingencia y medidas preventivas en caso de estos siniestros.

## **II. Marco Teórico**

### **2.1. Servicios ecosistémicos**

Para definir qué son los servicios ecosistémicos, nos basaremos en la definición de la Millenium Ecosystem Assessment (2005) que explica:

*“Los servicios ecosistémicos son los beneficios que la gente puede obtener de los ecosistemas. Estos incluyen servicios de aprovisionamiento de alimentos, agua, madera y fibra; servicios reguladores que afectan el clima, las inundaciones, las enfermedades, los desechos y la calidad del agua”.*

Con esto podemos clasificarlos como todos los servicios provistos por los ecosistemas que nos permiten abastecernos de elementos vitales, pudiendo incluso a veces llegar a ser esenciales para la sociedad y sus actividades. Como se detalla en la investigación de Potts et al. (2016), en Estados Unidos, en 2006, los servicios ecológicos entregados por insectos (polinización, control de plagas, descomposición, entre otros) contemplaban un valor anual de al menos \$57 mil millones de dólares.

#### **2.1.2. Polinización como servicio**

La polinización es básicamente el transporte de polen de la parte masculina de una flor a la parte femenina de otra, permitiendo así su reproducción (García et al., 2016). Los agentes polinizadores cumplen un rol muy importante, debido a que el servicio que brindan está vinculado directamente con el bienestar humano, ayudando a la reproducción de las plantas y por tanto al funcionamiento de los ecosistemas y la producción de cultivos. Esto último tiene un efecto directo en la seguridad alimentaria, ya que la polinización incide en la calidad y/o el rendimiento de aproximadamente un 75% de los cultivos de importancia mundial; sin los servicios de polinización se estima que entre un 5% y un 8% de la producción agrícola mundial se vería gravemente afectada (Potts et al., 2016). Cabe mencionar que durante marzo de 2020 los valores globales del servicio de polinización variaban aproximadamente entre los \$195 y los \$387 (promedio de \$267-657) mil millones dólares anuales (Porto et al., 2020).

Los servicios de polinización biótica (basada en organismos vivos) es realizada por diversos tipos de animales, pudiendo ser aves, insectos e incluso mamíferos. Entre estos destacan principalmente los insectos (Nates-Parra, 2016), que en el caso de Chile corresponden, en mayo del 2022, a alrededor del 98,6% de las especies de polinizadores según la red chilena de polinización (Inaturalist, 2022).

### **2.1.3. Incendios forestales en Chile**

Los incendios forestales son incendios que se producen y desarrollan en zonas donde no hay desarrollo urbano, consumiendo y utilizando como combustible las zonas vegetales (CONAF, 2022).

En Chile estos siniestros tienen lugar en varias épocas del año, siendo la principal de ellas el verano. Algunos ejemplos de paisajes donde pueden ocurrir son los cultivos forestales y agrícolas. Así mismo, suelen producirse entre las regiones IV y XII, siendo algunas de sus causas fogatas descuidadas o cigarro mal apagados (CONAF, 2006).

### **2.1.4. Distribución espacial de incendios en Chile**

En Chile, la Corporación Nacional Forestal (CONAF) estima una superficie afectada por los incendios forestales de 2.665.432,51 hectáreas, entre 1985 y 2021. Las zonas más afectadas fueron las regiones Centro-Sur y Sur del país, siendo de alto impacto mediático los casos de mega incendios ocurridos entre 2016 y 2017 donde desde la V hasta la IX región se registraron más de 546.000 mil hectáreas forestales afectadas.

Cabe destacar que los incendios además de afectar a zonas forestales también pudieron llegar a afectar zonas agrícolas dado que las regiones VII a la IX concentran la mayor superficie cultivada del país (ODEPA, 2019). Según datos de CONAF, en la temporada 2016-2017 se afectaron directamente cerca de 21.000 hectáreas de uso agrícola por los incendios forestales.

### **2.1.5. Disrupción del servicio por incendios**

No se conoce en su totalidad cómo los incendios afectan a las comunidades de polinizadores, aunque sabemos ciertamente que las plantas dependen de estos animales para el transporte de polen. Por tanto, aunque no es seguro asociar al fuego con las bajas de actividad polinizadora en una zona afectada por incendios, entender la relación

incendio/polinizador podría ser de ayuda para reducir este efecto, y por ende no poner en riesgo la vida vegetal silvestre y los cultivos que permiten la seguridad alimentaria (Brown et al., 2016).

## **2.2. Sistemas de Información Geográfica (SIG)**

### **2.2.1. ¿Qué son los Sistemas de Información Geográfica?**

Para este término se tiene la definición dada por Chang (2019), que indica que:

*“Un sistema de información geográfica (GIS) es un sistema informático para capturar, almacenar, consultar, analizar y mostrar datos geoespaciales. Los datos geoespaciales describen tanto la ubicación como los atributos de las características espaciales”.*

Se entiende que el uso de las funciones descritas por Chang para los SIG permite observar, manipular y visualizar de mejor manera la información geográfica, por ejemplo, a un nivel nacional de eventos como los incendios forestales, así como de las características geoespaciales de los terrenos donde ocurrieron los siniestros. En este sentido, los SIG son capaces de mostrar información de variadas características geográficas y para lograrlo usan un sistema de coordenadas, que son una forma de distribución del planeta para poder ubicar los objetos que lo conforman. Los SIG no solo se centran en la ubicación y representación gráfica de los objetos que conforman un terreno, sino que también son capaces de mostrar sus características, por ejemplo, si tuviéramos los datos y mediciones de una determinada red de caminos, podríamos visualizar junto con el trazado las características registradas para los mismos (Bolstad, 2006).

### **2.2.2. Categorías de SIG**

Warf (2010) clasifica a los Sistemas de Información Geográfica en: *SIG de escritorio, Sistemas de gestión de bases de datos espaciales (SDBMS), Servidor de mapas web, SIG de servidor, Clientes SIG Web, SIG Móvil, y Bibliotecas y Extensiones*. Siendo la categoría de SIG de escritorio la más popular, y se entiende como las aplicaciones que se instalan en un computador para realizar las acciones básicas como consultar, editar y crear, entre otras, en un ambiente local.

## **2.3. Modelos predictivos**

### **2.3.1. ¿Qué son los algoritmos de minería de datos?**

Data mining o minería de datos es un término que tiende a confundir mucho con knowledge discovery from data (KDD). Este último es el proceso completo de extracción de información desde los datos, mientras que la minería de datos es solamente un paso de KDD, concretamente, la tarea donde aplicamos métodos o algoritmos para buscar patrones dentro de las bases de datos, lo que más tarde se traduce mediante otros procesos en conocimiento (Han et al., 2012). Por otro lado, Witten (2005) dice que la minería de datos es el proceso de descubrir patrones que sean significativos dentro de los datos, estos patrones al ser descubiertos pueden ser usados para obtener una ventaja.

Ambas fuentes coinciden en que es durante la minería de datos que encontramos patrones, sin embargo, Han et al. (2012) dicen que el proceso de obtener conocimiento de estos patrones identificados viene de otras tareas y Witten (2005) considera la minería de datos como el proceso completo, incluyendo la transformación de los patrones descubiertos a conocimiento o ventajas. De cualquier manera, teniendo en cuenta que el fin de esta investigación es encontrar la relación (los patrones) de cómo son afectadas las poblaciones de polinizadores por incendios forestales, ambas definiciones son igualmente útiles para este estudio.

Según Grus (2019) los modelos son una forma de expresar relaciones matemáticas o probabilísticas entre diferentes variables. Teniendo en cuenta esto, los algoritmos de minería de datos son métodos que permiten extraer conocimientos importantes de grandes cantidades de datos, tales como en sistemas con información alimentados de base de datos masivas, donde en principio una persona no podría analizar tal cantidad de datos. Algunos ejemplos de datos masivos podrían ser las bases de datos municipales de una comuna o región, las bases de datos de una empresa multinacional con sede en muchos países, entre otros (Fadl Elsid y Eltahir, 2014).

### **2.3.2. ¿Qué son los métodos predictivos?**

Para describir los métodos predictivos nos basaremos en la definición del análisis predictivo que según Nyce (2007) corresponde a:

*“Un término amplio que describe una variedad de técnicas estadísticas y analíticas utilizadas para desarrollar modelos que predicen eventos o comportamientos futuros”.*

A partir de esto entendemos que los métodos predictivos son herramientas matemáticas diseñadas para realizar predicciones precisas sobre el comportamiento (numérico) de un fenómeno que se planea predecir. Estos modelos son usados en minería de datos para explorar las bases de datos y obtener los anteriormente mencionados patrones o conocimiento.

### **2.3.3. Clasificaciones de modelos predictivos**

Para esta investigación se utilizan modelos predictivos basados en Machine Learning (ML), los cuales se clasifican en supervisados y no supervisados. Existen muchas definiciones de ML algunas de estas son:

- *“Una forma de Inteligencia Artificial en la que permite a un sistema aprender de los datos en lugar de ser por una programación explícita”* (Hurwitz y Kirsch, 2018).
- *“La creación y el uso de modelos que aprenden a partir de datos, para predecir varios resultados nuevos para estos datos”* (Grus, 2019).

Considerando esto, definiremos ML como una forma de inteligencia artificial que aprende por datos, de los cuales puede obtener información valiosa; esto cuando es usada en minería de datos para obtener patrones dentro de una base de datos o para explicar las relaciones probabilística o matemáticas entre variables (Grus, 2019).

### **2.3.4. Aprendizaje supervisado y no supervisado**

De acuerdo con la definición de Müller y Guido (2016), el aprendizaje supervisado se usa para predecir una salida en base a una entrada determinada. Este tipo de aprendizajes es usado en modelos de ML como algunos de los que se mencionarán más adelante.

Considerando el esquema de la Figura 1, podemos identificar dos etapas. La primera es donde se entrena a la Inteligencia Artificial (IA), es decir, es aquí donde ocurre el componente supervisado, el cual consiste, en este caso, de datos que vienen etiquetados para que la IA pueda procesarlos e identificar patrones dentro de estos. Dichos patrones

servirán para la segunda etapa donde la IA debe identificar nuestro objeto (para este ejemplo fotos de hombres), en base a los patrones que aprendió.

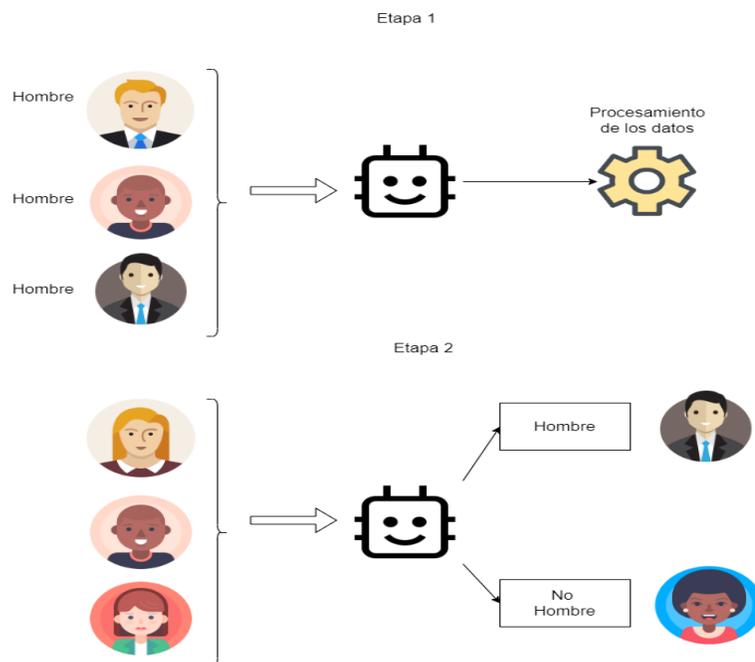


Figura 1. Aprendizaje Supervisado.

(Elaboración propia, basado en Müller y Guido, 2016).

Algunos ejemplos de aprendizajes supervisados son los siguientes:

- Árboles de decisión: Según la definición de James et al. (2013):  
*“Los árboles de decisión incorporan un enfoque de clasificación supervisada. La idea surgió de la estructura de árbol ordinaria que se compone de una raíz y nodos (las posiciones donde se dividen las ramas), ramas y hojas. De manera similar, un árbol de decisión se construye a partir de nodos que representan círculos y las ramas están representadas por los segmentos que conectan los nodos”.*

Para ejemplificar esta definición explicaremos la ilustración de la Figura 2 en la cual nuestra raíz (señalada con el color rojo) corresponde a la pregunta de si el animal en el ejemplo **¿Tiene plumas?**, para continuar con los caminos indicados por las ramas que en este caso sería **SI** o **NO**, luego dependiendo de la rama que corresponda a la observación, nos derivará a uno de los dos siguientes nodos:

¿Puede volar? o ¿Es un mamífero?, reiteraremos el proceso de responder la pregunta en base a nuestra observación, en caso del nodo señalado con el color azul (¿Es un mamífero?) continuaremos recorriendo las ramas y nodos hasta llegar a algún nodo terminal como el nodo verde (**Mono**).

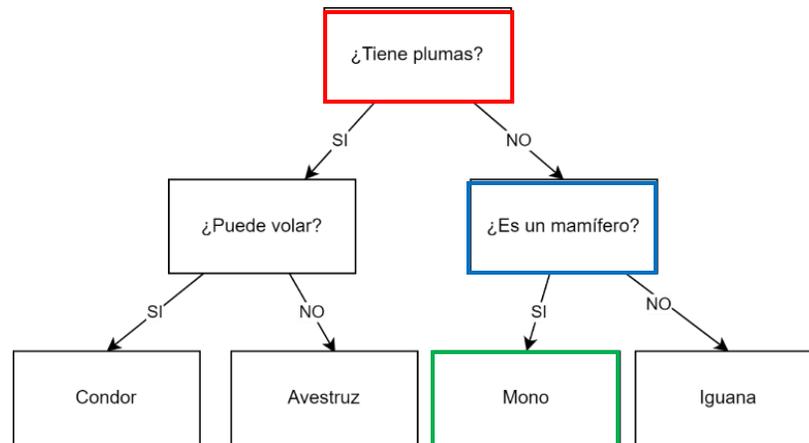
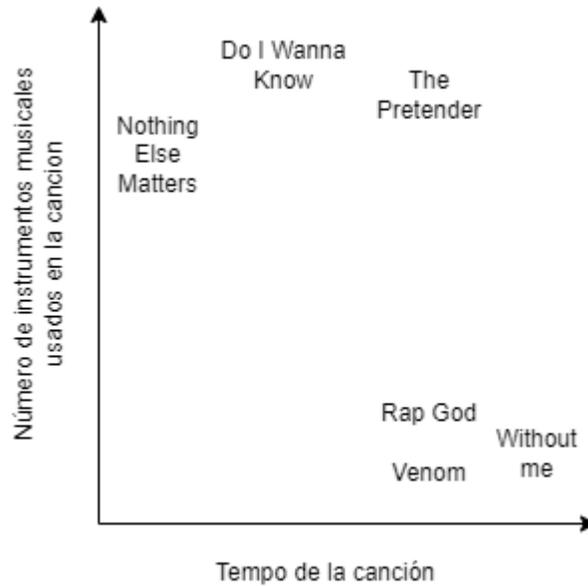


Figura 2. Árboles de decisión.

(Adaptado de Müller y Guido, 2016).

- K-Nearest Neighbors: El KNN es un algoritmo supervisado, principalmente usado para clasificar. El algoritmo toma observaciones cuyo grupo desconocemos y las asigna a un grupo con el que tenga más similitud. Esto se decide comparando las K observaciones más cercanas con la observación sin clasificar (Albon, 2018). En el ejemplo de la Figura 3, podemos apreciar el comportamiento a priori del algoritmo, en la cual el algoritmo identifica en base a parámetros de las canciones a qué grupo deberían de pertenecer, agrupándolas en este caso en base al número de instrumentos musicales que incorporan y a su tempo.



*Figura 3. Ejemplo KNN.*

*(Adaptado de Ejemplo de Han et al., 2012).*

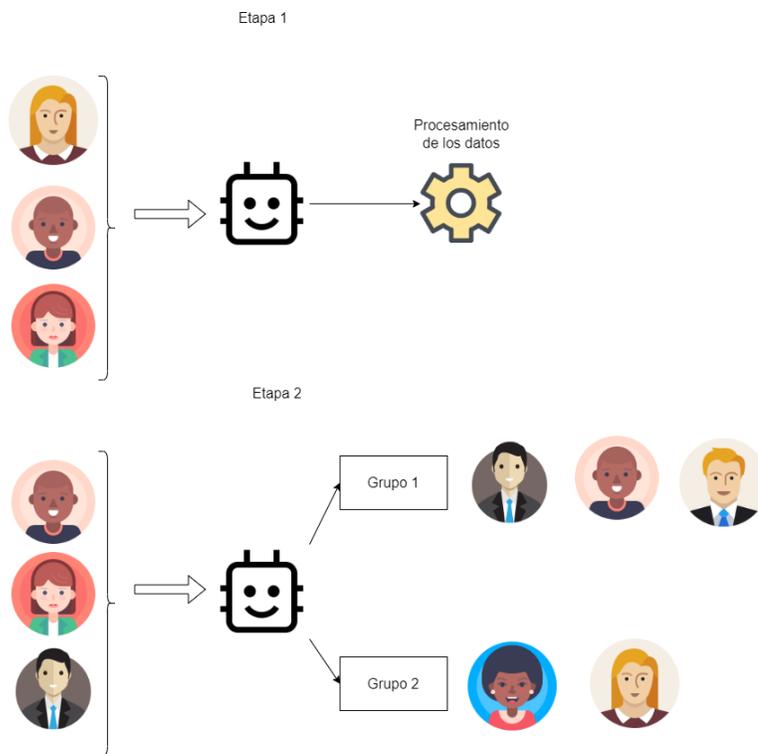
- MaxEnt: Es un algoritmo utilizado para modelar la distribución de especies. Contando solo con las presencias, este modelo nos permite modelar distribuciones de especies, donde el núcleo de salida de este modelo es brindar información sobre las características ideales para la posible presencia relativa de un punto o lugar frente a otro (Elith et al., 2010). Gracias a este modelo se pueden construir mapas de posible presencia de especies, que muestren la posible presencia de flora o fauna en un territorio, basado en las observaciones de presencia con las que se debe entrenar al modelo. De esta manera, podemos predecir en base a características, para luego mediante SIG hacer que esto sea traspasado a un plano visual.

Por otra parte, y de acuerdo con la definición de Russel y Norvig (2020), en el aprendizaje no supervisado:

*“El agente aprende patrones en la entrada, aunque no se proporcione retroalimentación explícita. La tarea de aprendizaje no supervisado más común es la clusterización:*

*detectar grupos de ejemplos de entrada potencialmente útiles. Por ejemplo, un agente de taxis podría desarrollar gradualmente un concepto de "días de buen tráfico" y "días de mal tráfico" sin que un maestro le dé ejemplos etiquetados de cada uno".*

En el ejemplo de la Figura 4, nuevamente podemos identificar dos etapas: la primera en la cual, en base a los datos de entrada, el algoritmo aprende e identifica patrones entre las observaciones; y la segunda, donde usa los patrones aprendidos para dividir las observaciones según los patrones que encuentra.



*Figura 4. Aprendizaje No Supervisado.*

*(Basado en Patel, 2019).*

Un ejemplo de aprendizaje no supervisado corresponde a:

- Clustering: Donde, según Müller y Guido (2016) “Los algoritmos de agrupamiento asignan o predicen un número en cada punto de datos indicando a que grupo pertenece este dato en particular”. Así, los algoritmos de clusterización agrupan observaciones que sean similares en base a sus características, para lo que predicen de forma no supervisada.

Un ejemplo de estos métodos de clusterización es el de k-means (Figura 5), este intenta encontrar puntos de conglomerado de los datos que sean representativos para cada zona de estos. Como se observan en la imagen, el algoritmo asigna primero cada punto del dato al conglomerado con el que tenga más similaridad, luego establece un centro para cada clúster basado en las observaciones que lo conforman, y luego nuevamente asigna cada punto al clúster con el que tenga más similaridad. Este proceso es iterativo, y su fin se ve cuando ninguna observación cambia de clúster (Müller y Guido, 2016).

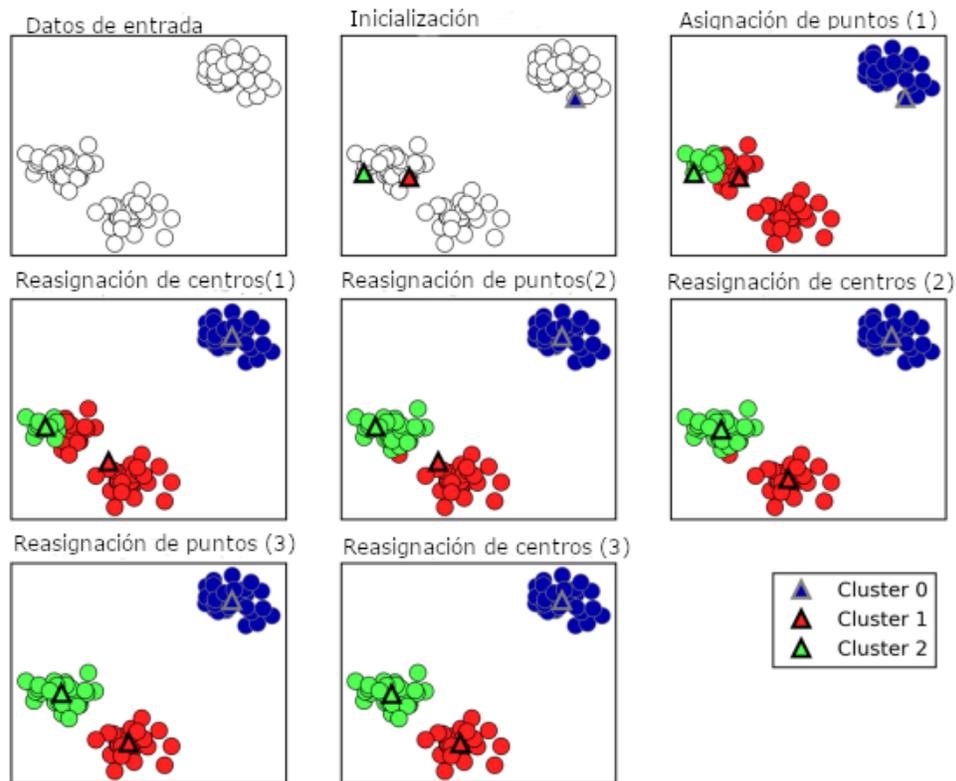


Figura 5. Ejemplo de Clustering.  
(Adaptado de Müller y Guido, 2016).

### **III. Metodología**

#### **3.1. Obtención de datos de polinizadores**

##### **3.1.1. Obteniendo la lista de polinizadores en Chile**

Para obtener la lista de los polinizadores en Chile, se utilizó la información disponible en la red chilena de polinización de Inaturalis (mayo de 2022), para la extracción de estos datos se descargó el elemento “div” que contiene toda la información de las especies de polinizadores mostrados en su página web. Después de esto mediante la herramienta de programación Visual Estudio Code que permite eliminar coincidencias con su editor de texto, se eliminaron todas las líneas de código que no eran necesaria en los registros dejando solo el nombre de las especies, obteniendo finalmente un listado de 466 polinizadores. Luego se procedió a eliminar algunas especies que se encontraban registradas por su género o familia, puesto que para poder obtener los registros de presencia de dicha especie más adelante es necesario tener el nombre científico de la misma.

##### **3.1.2. Obteniendo los registros de avistamiento**

Una vez limpia la lista de polinizadores presentes en Chile, el siguiente paso para tener un set de datos que permita entrenar modelos de Machine Learning fue descargar los avistamientos de las especies listadas. En este caso se usaron dos fuentes de datos: Inaturalist (2022) y Global Biodiversity Facility (2022), de ahora en adelante, GBIF. La descarga de datos se hizo mediante el ambiente de programación Rstudio y el lenguaje R (ver código en Anexo1), además de las librerías (rinat, rgbif, y scrubr) que nos permiten descargar los datos en el lenguaje, esto dado que pueden conectarse a las bases de datos de las fuentes antes mencionadas cuando se les da como parámetro una especie por su nombre científico.

De esta forma usando el código que está en el anexo, es posible automatizar la consulta a la base de datos de estas fuentes. Este proceso genera un archivo por cada fuente de datos, por lo que se obtuvieron dos archivos (uno proveniente de GBIF y otro de Inaturalist), que en cada hoja tiene los registros de avistamientos de una especie distinta, junto con algunas variables y las coordenadas de avistamiento en UTM.

## 3.2. Tratamiento de datos de polinizadores

### 3.2.1. Unificando los datos

Como se mencionó, el producto del código usado para obtener los registros fueron dos archivos Excel, cada página de estos archivos contenía los registros de una especie y variables asociadas a esta, el siguiente paso para darles uso a estos registros y crear modelos de Machine Learning es el dejarlos en un solo archivo unificado, en el cual los registros provenientes de ambas fuentes deben tener mismas variables (columnas), y adicionalmente todas las especies y sus registros deben estar en la misma hoja. Este segundo proceso se realizó con un código Python, que implementa la librería Pandas que facilita el trabajo con archivos de datos y permite su manipulación de forma más eficiente (ver códigos en Anexos 2 y 3).

La eliminación de variables y unificación se realizó borrando las columnas no deseadas, algunas de estas, por ejemplo, eran referentes a la ubicación donde se registró el avistamiento, pero poseían demasiados registros en blanco, por lo que se descartó hacer análisis con ellas. Una vez eliminadas las variables de no interés se copiaron los registros de un archivo a otro usando la herramienta Excel, unificando así el archivo.

*Tabla 1. Ejemplo de datos.*

Espece	Latitud	Longitud	Fecha Identified	Provincia	Fuente
<b>Phulia nymphula</b>	- 18,05400	- 69,20000		Arica y Parinacota	GBIF
<b>Pierphulia rosea</b>	- 18,05485	- 69,20023	2021-08-01 13:00:00 -0400	Caquena, Putre, Arica y Parinacota	Inaturalist
<b>Vanessa carye</b>	- 18,06523	- 69,06888	2019-05-16 14:30:00 +0200	Laguna Sora Pata	Inaturalist
<b>Oreotrochilus estella</b>	- 18,11080	- 69,29927		Arica y Parinacota	GBIF

### 3.2.2. Eliminando duplicados

Una vez obtenido el documento unificado en Excel, se procedió a eliminar los duplicados que había, es decir, observaciones que aparecían tanto en Innaturalis como en GBIF y que podían alterar la veracidad de nuestros datos, esto se dado que la segunda incorpora cada cierto tiempo algunos registros de la primera. En este caso usamos Python 3.7 y la librería

Pandas para las estructuras de datos. Este algoritmo elimina aquellos registros en que una especie, en este caso identificada por su nombre científico, se encuentre en las mismas coordenadas exactas que una de su misma especie. Se inicio con un total de 46.000 observaciones aproximadamente y luego de estos filtros se llegó a 23.518.

### 3.2.3. Eliminando coordenadas de no interés

Se eliminaron 168 coordenadas en el documento Excel que tenían como valor 0 ó correspondían a coordenadas de Chile no continental (mayoritariamente datos de Juan Fernández e Isla de Pascua) (ver Tabla 2). Esta operación se realizó aplicando un filtro en Excel que consistió en excluir aquellos registros que fueran igual a 0 ó estuvieran fuera de Chile continental, esto basándonos en los límites del país.

*Tabla 2. Ejemplo de datos inválidos.*

Especie	Latitud	Longitud	dateIdentified	Provincia	Fuente
<b>Acamptopoeum submetallicum</b>	0	0		Metropolitana de Santiago	GBIF
<b>Alloscirtetica gayi</b>	0	0		Coquimbo - Región IV	GBIF
<b>Anthidium chilense</b>	0	0		Atacama - Region III	GBIF
<b>Anthophora paranensis</b>	0	0		Metropolitana de Santiago	GBIF
<b>Sephanoides fernandensis</b>	-33,64223	-78,8273		Valparaíso	GBIF
<b>Sephanoides fernandensis</b>	-33,64082	-78,81269		Valparaíso	GBIF

### 3.2.4. Creando el formato espacial

Hasta el momento las coordenadas de cada registro en el archivo Excel leído por los códigos, no estaban en un formato espacial, uno que permitiera un análisis con coordenadas, por ello se transformaron usando la librería SP de R “a spatial points”, esto genera puntos que representan un cierto lugar en la tierra y los datos se asocian a este lugar mediante coordenadas. Si bien existen diferentes sistemas de coordenadas, en este caso se mantuvieron en el formato original, el cual corresponde a UTM (Universal Transversa de Mercator) que da a la tierra una proyección cilíndrica y divide el planeta en husos. Con esto fue posible realizar los primeros gráficos de cómo se distribuyen las presencias de las

especies de polinizadores a lo largo de Chile. Más adelante este proceso facilitará el pasar los datos a un formato ráster en el cual realizar los modelos que ayudarán a identificar patrones dentro de las presencias de las especies.

### 3.3. Obtención y tratamiento de datos predictores ambientales

#### 3.3.1. Descargando datos de predictores

Para poder añadir los datos ambientales que servirán como predictores de las especies, fue necesario acceder (vía librerías de R, en este caso Ráster) a las bases de datos presentes en la web de Fick et al. (2017). Mediante esta librería se pudo descargar la información climática de todo el mundo, el método para realizar esto fue crear la lista con las coordenadas de presencias en formato spatial data, como el mencionado anteriormente, lo que nos permitió procesarlo como datos espaciales (es decir, puntos geográficos con coordenadas asociadas), para luego extraer estas y posteriormente asociarlas a los registros de presencia.

A continuación, para ilustrar los datos, en Tabla 3 se observan 6 de las 19 variables bioclimáticas descargadas, así como 4 registros de especies. Finalmente, aquí quedamos con 23.352 observaciones aproximadamente.

*Tabla 3. Ejemplo con datos climáticos.*

Especie	Latitud	Longitud	dateIdentified	Provincia	Fuente	bio1	bio2	bio3	bio4	bio5	bio6
<b>Manduca sexta</b>	- 35,43016	- 71,67192	2020-12-18 13:24:51 - 0300	5 1/2 Sur A, Talca, Maule, CL	Inaturalist	147	141	52	4547	305	37
<b>Autoplusia gammoides</b>	- 35,43213	- 71,63952	2020-11-17 10:53:24 - 0300	18 y Medio Oriente, Talca, Maule, CL	Inaturalist	147	142	52	4599	307	37
<b>Vanessa carye</b>	- 35,43322	- 71,61628	2021-04-11 12:52:20 - 0300	3 Norte, Talca, Maule, CL	Inaturalist	145	143	52	4619	307	35
<b>Vanessa tersichore</b>	- 35,43334	- 71,61615	2021-09-04 09:36:54 - 0300	3 Norte, Talca, Maule, CL	Inaturalist	145	143	52	4619	307	35

#### 3.3.2. Tratando datos de predictores

Al contar ya con toda la información climática de Chile se procedió a enlazar la presencia de nuestras especies de polinizadores con esta información, logrando así un matriz de información en la cual las especies fueron agregadas como otra variable. Así, para los

píxeles con ausencia de alguna especie se creó una nueva columna con el nombre de la especie, rellena con un 0 y con un 1 cuando sí estaba, logrando así finalmente identificar las condiciones climáticas presentes en los píxeles donde existía presencia de una especie.

### **3.4. Desarrollo de modelos**

Se usó Ensamble de Pequeños Modelos (ESM, por sus siglas en inglés) para modelar la distribución potencial de las especies que presentaron menos de 150 registros en el área de estudio, la cual abarca desde la región de Valparaíso hasta el extremo Sur. Se eligieron estos modelos porque han demostrado una capacidad predictiva sobresaliente en comparación con modelos de distribución especies estándares, cuando se quiere estimar la distribución potencial de especies raras o con pocos registros de presencias, como es el caso de algunas especies contempladas en este trabajo (Breiner et al., 2015). Además, los ESM permiten usar una gran cantidad de predictores sin riesgo de sobreajuste (Breiner et al., 2015), ya que los ESM modelan la distribución de una especie mediante el ensamble de modelos bivariados, es decir, modelos que tienen solo dos predictores, eliminando de esta forma la posibilidad de sobreajuste. Así, las especies con menos de 150 registros se modelaron con los ESM que se usaron un total de 15 predictores no correlacionados (Coeficiente de Pearson  $< 0.6$  y  $> -0.6$ ), lo que resultó en un total de 105 modelos bivariados por especie. Todos los ESM fueron ajustados mediante la función `ecospat.ESM.Modeling` del paquete `ecospat` del software y lenguaje de programación R (Di Cola et al., 2017). Se usó MaxEnt para ajustar los modelos bivariados, ya que ha demostrado una gran capacidad predictiva sobre otros algoritmos cuando se estima la distribución de especies mediante el uso de ESM (Breiner et al., 2018), además de que ha demostrado superar a otros modelos como se mencionó con anterioridad. El ensamble de modelos bivariados para crear el ESM por especie, se basó en el promedio de las probabilidades de presencias estimadas, ponderadas por la capacidad predictiva de los modelos bivariados. Se usó el AUC (Area Under the Curve) para la ponderación de los pequeños modelos, quedaron fuera del ensamble los modelos bivariados con un AUC menor a 0.5, siguiendo los ejemplos de Breiner et al. (2015) y Breiner et al. (2018). Las especies con 150 o más registros fueron modeladas de la manera estándar, es decir, usando los 15 predictores a la vez. También se usó MaxEnt para estimar la distribución potencial

de estas especies, debido a su alta capacidad predictiva (Elith et al. 2006; Elith and Graham 2009; Heikkinen et al. 2012), siendo lo más adecuado para modelar especies con registros de “solo presencia” (Phillips et al., 2006), es decir, especies que se estima que están en muchos más lugares fuera de lo observado por los investigadores. Esto se ajusta al caso de estudio, ya que como se mencionó, los registros que usamos vienen de observación de ciudadana u de otras entidades, por lo que no es posible determinar cada lugar donde puede habitar o no una especie. Se usó la función BIOMOD Modelling del paquete biomod2 para ajustar los modelos (Thuiller et al., 2019). A todos los modelos se les buscó la mejor configuración de parámetros (es decir, opciones de modelado) del algoritmo MaxEnt para obtener la mejor predicción posible por especie. Específicamente, se probaron todas las combinaciones posibles de restricciones o “característica”, las cuales fueron: lineal, cuadrática, producto, umbral y bisagra. Se evaluó cada configuración de parámetros de cada modelo mediante validación cruzada de 10 veces, usando aleatoriamente en cada iteración un 80% de los datos para calibración (es decir, ajuste o entrenamiento) y 20% para validación. Se usó AUC para evaluar los modelos. La configuración de los parámetros que resultó en el valor esperado de AUC más alto para el modelo de una especie, se mantuvo para predecir la distribución potencial de la especie en cuestión sobre toda el área de estudio. Se aclara que, en el caso de los ESM, cada modelo bivariado para una especie fue evaluado buscando la mejor configuración de parámetros descrita anteriormente, pero el ESM final fue vuelto a evaluar mediante validación cruzada de 10 veces, para así calcular la capacidad predictiva final del ESM.

### **3.5. Generación de mapas**

#### **3.5.1. Obteniendo el mapa ideal**

Para la creación del mapa ideal se usaron los píxeles mencionados anteriormente, los cuales tenían asociadas variables ambientales que fueron utilizadas como variables independientes y entrada de los modelos, lo cual nos permitió para cada especie predecir una potencial presencia/ausencia sobre un píxel, todo esto basándonos en los registros de presencia obtenidos al principio.

### **3.5.2. Cruzando el mapa ideal con el mapa de incendios**

Se cruzaron ambos mapas para determinar qué área fue la afectada por los incendios de la temporada pasada y poder determinar cuánta área de alta biodiversidad en Chile se perdió. El procedimiento consistió en usar una herramienta de SIG de escritorio, en este caso QGIS, la cual permitió enlazar las coordenadas disponibles y que fueron obtenidas en pasos anteriores, con datos de incendios nacionales, proporcionados por colaboradores de otras universidades. Una vez obtenidos dichos datos, se pudo calcular la superficie quemada para cada píxel de los archivos ráster.

## **3.6. Herramienta de visualización**

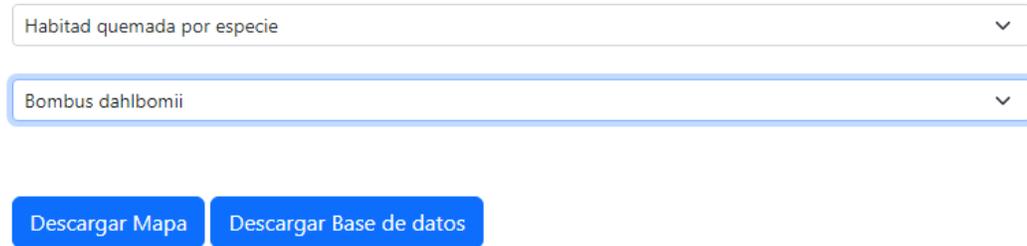
### **3.6.1. Definiendo las librerías y lenguajes a utilizar**

Esta herramienta fue desarrollada utilizando el lenguaje de etiquetas HTML y JavaScript con la librería jQuery, para dar funcionalidad a estas etiquetas. Este método fue elegido dado que permite usar hostings gratuitos, lo que facilita el alojar la aplicación web una vez finalizada su creación. Asimismo, la creación y personalización de los estilos se usó la librería Bootstrap 5, que permite cargar plantillas de páginas web ya creadas, lo cual facilita el desarrollo de una nueva. Así mismo permite dar responsividad a la página, de forma que sea usable en múltiples dispositivos sin importar el tamaño de su pantalla.

### **3.6.2. Mostrando los archivos**

Los archivos como imágenes o datos para las tablas están guardados en carpetas, cada uno con un nombre que les permite identificarse. Estos archivos son los mapas previamente generados durante este trabajo. Para poder mostrarlos se definió usar selectores, es decir, botones que al hacer click en ellos desplieguen una serie de opciones de las cuales se debe seleccionar una (Figura 6). Dependiendo de la opción, podría aparecer un segundo selector, destinado a seleccionar una especie. Por ejemplo, si se elige para visualizar en el primer selector “Mapa de presencia estimada en %”, aparecería bajo este un segundo selector destinado a elegir una especie para el mapa.

# Selector mapas de incendios y otras graficas



The image shows a web interface with two dropdown menus and two buttons. The first dropdown menu is labeled "Habitad quemada por especie" and has a downward arrow. The second dropdown menu is labeled "Bombus dahlbomii" and also has a downward arrow. Below these are two blue buttons: "Descargar Mapa" and "Descargar Base de datos".

*Figura 6. Primeras funciones de herramienta de visualización.*

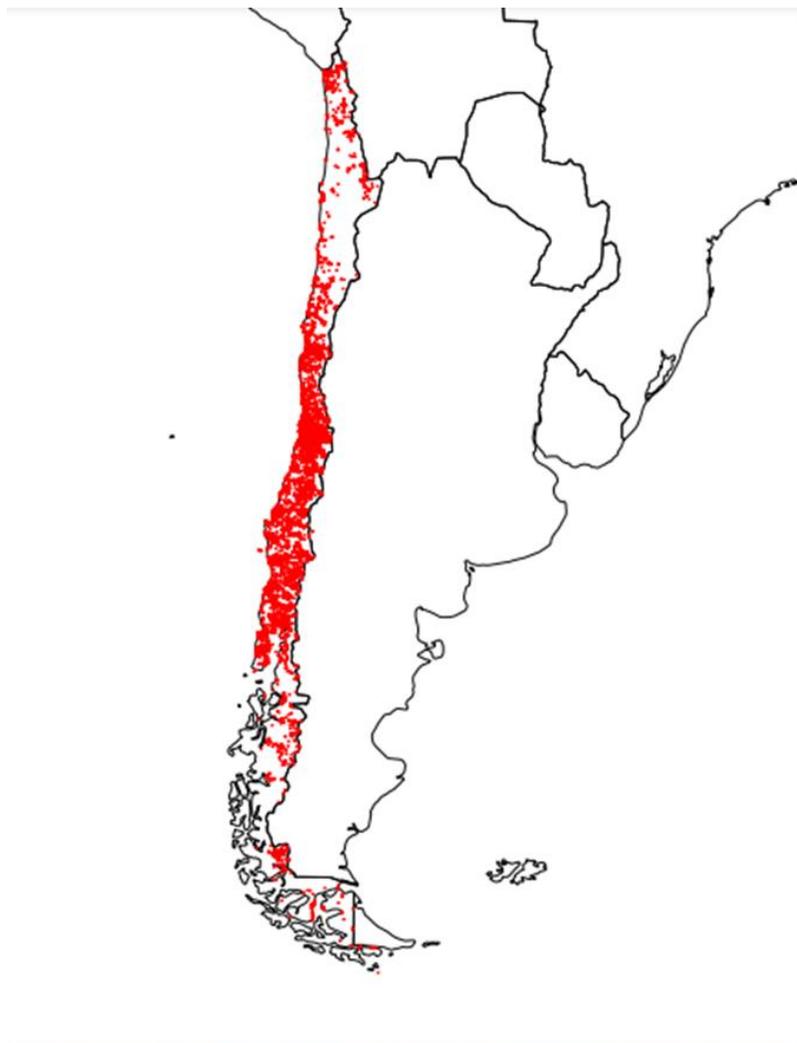
El archivo organizador.js es el encargado de cambiar el atributo a la imagen principal de la página con la dirección correspondiente de esta, según las opciones indicadas por el selector, es decir, según la opción que se elija. Supongamos, para hábitat quemada por especie de “Bombus dahlbomii”, se buscará el mapa asociado a este nombre y se mostrará. De este mismo modo, se puede descargar el mapa generado, usando la dirección del archivo que el usuario requiera con el botón “Descargar Mapa”. Por último, el botón “Descargar base de datos” se encarga de descargar un archivo CSV con los datos que se recabaron durante la investigación y con los que están desarrollados los distintos mapas expuestos en la herramienta.

## IV. Presentación y análisis de resultados

### 4.1. Primeros mapas de observaciones

#### 4.1.1. Mapas de puntos de observaciones

Como primer resultado se generó el mapa de observaciones donde se muestra la distribución de las observaciones a lo largo de Chile, para todas las especies (Figura 7). Se utilizó principalmente para identificar puntos que no correspondían a Chile continental, y por ende, lugares de los que no se disponía de información de incendios, además de dar una idea de en qué lugar se aglomeran principalmente las observaciones.



*Figura 7. Mapa de observaciones.*

Cabe mencionar que muchos de los puntos fueron descartados al encontrarse dentro del mismo píxel, de forma que a la hora de pasar este mapa de avistamientos totales a presencia binaria (es decir, si se encuentra o no una especie en un píxel) se superpusieron las observaciones, dejándolas solo como una.

#### 4.1.2. Mapas de presencias potenciales (simulaciones)

En Tabla 4 se muestran los valores de AUC registrados para cada especie en los modelos de Árboles de Decisión y MaxEnt Phillips. En términos generales, Árboles de Decisión presenta mejores resultados en las primeras especies de la lista, las cuales tienen la mayor cantidad de registros, sin embargo, la diferencia entre el desempeño de ambos algoritmos se reduce a medida que se avanza por la lista, lo que se debe a que MaxEnt obtiene mejores predicciones en casos donde la especie presenta menor cantidad de registros.

*Tabla 4. Resultados AUC.*

Especie	AUC Árboles de decisión	AUC MaxEnt Phillips
<b>Sephanoides sephaniodes</b>	89,19%	87,06%
<b>Patagona gigas</b>	97,10%	95,65%
<b>Rhodopis vesper</b>	97,29%	97,21%
<b>Bombus dahlbomii</b>	80,73%	89,23%
<b>Bombus terrestris</b>	92,74%	88,85%
<b>Apis mellifera</b>	96,79%	97,92%
<b>Vanessa carye</b>	95,99%	97,25%
<b>Astylus trifasciatus</b>	97,59%	95,24%
<b>Oreotrochilus leucopleurus</b>	95,40%	98,58%
<b>Megachile saulcyi</b>	88,46%	91,99%
<b>Cadeguala occidentalis</b>	89,34%	93,98%
<b>Eristalis tenax</b>	91,68%	95,74%
<b>Tatochila mercedis</b>	93,78%	96,20%
<b>Megachile pollinosa</b>	89,42%	92,51%
<b>Anthidium chilense</b>	95,51%	93,22%
<b>Acamptopoeum submetallicum</b>	84,73%	88,43%
<b>Colias vauthierii</b>	87,89%	88,68%
<b>Hylephila fasciolata</b>	94,04%	93,69%
<b>Oreotrochilus estella</b>	94,58%	99,77%
<b>Ormiscodes cinnamomea</b>	93,40%	96,98%
<b>Svastrides melanura</b>	87,66%	96,19%

<b>Ruizantheda proxima</b>	94,28%	91,99%
<b>Diphaglossa gayi</b>	96,72%	98,66%
<b>Thaumastura cora</b>	99,23%	99,45%
<b>Pieris brassicae</b>	95,86%	98,57%
<b>Centris nigerrima</b>	90,43%	96,59%
<b>Xylocopa augusti</b>	99,91%	99,95%
<b>Notanthidium steloides</b>	89,92%	93,47%
<b>Toxomerus vertebratus</b>	88,28%	95,69%
<b>Ormiscodes amphinome</b>	83,47%	91,43%
<b>Battus polydamas</b>	99,34%	98,26%
<b>Diadasia chilensis</b>	90,47%	96,77%
<b>Leptotes trigemmatius</b>	99,54%	98,60%
<b>Auca coctei</b>	95,37%	96,93%
<b>Corynura chloris</b>	92,93%	97,68%
<b>Vanessa terpsichore</b>	95,98%	92,18%
<b>Bilyaxia concinna</b>	98,21%	97,99%
<b>Manuelia gayatina</b>	92,10%	95,17%
<b>Hyles annei</b>	98,38%	97,09%
<b>Tatochila autodice</b>	94,55%	94,23%
<b>Calliopsis trifasciata</b>	92,37%	96,60%
<b>Manuelia gayi</b>	88,78%	93,92%
<b>Centris chilensis</b>	88,62%	96,61%
<b>Isepeolus luctuosus</b>	85,42%	91,62%
<b>Corynura chilensis</b>	93,86%	94,91%
<b>Alloscirtetica gayi</b>	93,37%	96,29%
<b>Yramea cytheris</b>	89,60%	94,27%
<b>Cadeguala albopilosa</b>	89,93%	93,10%
<b>Ectinogonia buquetii</b>	96,83%	99,30%
<b>Allograpta hortensis</b>	89,19%	92,63%
<b>Corynura corinogaster</b>	85,05%	89,44%
<b>Anthidium gayi</b>	80,50%	90,39%
<b>Syrphus octomaculatus</b>	92,69%	92,23%
<b>Caupolicana gayi</b>	87,90%	94,96%
<b>Chilesia rudis</b>	93,78%	94,22%
<b>Bombus ruderatus</b>	93,08%	89,83%
<b>Palpada elegans</b>	90,40%	89,89%
<b>Polythysana cinerascens</b>	95,01%	92,10%
<b>Megachile semirufa</b>	89,22%	92,35%
<b>Colletes cyanescens</b>	94,03%	97,79%
<b>Adetomeris erythropros</b>	95,92%	93,48%
<b>Arthrobrachus nigromaculatus</b>	99,61%	95,93%
<b>Caupolicana fulvicollis</b>	98,82%	98,25%

<b>Coelioxys pergandei</b>	89,65%	93,13%
<b>Pseudolucia chilensis</b>	98,69%	96,06%
<b>Cylindrophora maulica</b>	91,18%	96,23%
<b>Heliopetes americanus</b>	92,48%	96,97%
<b>Centris cineraria</b>	94,23%	97,03%
<b>Liphanthus sabulosus</b>	97,88%	95,45%
<b>Melectoides triseriatus</b>	98,24%	93,35%
<b>Ctenucha vittigera</b>	98,59%	99,26%
<b>Syncirsodes primata</b>	89,53%	94,30%
<b>Plagiodera erythroptera</b>	96,47%	99,32%
<b>Doeringiella gayi</b>	76,98%	93,17%
<b>Hylephila signata</b>	92,07%	96,67%
<b>Austroscaeva melanostoma</b>	90,79%	98,09%
<b>Allograpta pulchra</b>	90,03%	95,13%
<b>Manuelia postica</b>	91,56%	92,22%
<b>Epiclopus gayi</b>	92,97%	94,01%
<b>Megachile rotundata</b>	98,40%	97,37%
<b>Colletes nigritulus</b>	86,76%	96,37%
<b>Trichothurgus dubius</b>	93,32%	97,96%
<b>Megachile euzona</b>	95,99%	97,50%
<b>Strymon eurytulus</b>	94,26%	95,31%
<b>Phoebis sennae</b>	96,64%	96,52%
<b>Ennada flavaria</b>	93,16%	93,02%
<b>Megachile cinerea</b>	94,71%	96,23%
<b>Dolichogyna chilensis</b>	86,03%	94,66%
<b>Notanthidium rudolphi</b>	88,24%	98,37%
<b>Caupolicana hirsuta</b>	95,44%	98,64%
<b>Ruizanthedella mutabilis</b>	91,02%	99,15%
<b>Xylocopa splendidula</b>	98,76%	98,64%
<b>Lygaeus alboornatus</b>	95,82%	95,18%
<b>Pseudagapostemon citricornis</b>	82,66%	95,62%
<b>Malamblia melas</b>	91,87%	93,90%
<b>Calliopsis hirsutula</b>	85,76%	95,58%
<b>Caupolicana quadrifasciata</b>	96,73%	98,69%
<b>Phaedon semimarginatus</b>	92,16%	90,74%
<b>Platycheirus chalconota</b>	87,25%	92,63%
<b>Camponotus morosus</b>	95,10%	96,46%
<b>Camponotus chilensis</b>	88,07%	96,88%
<b>Chalepogenus rozeni</b>	95,85%	97,49%
<b>Lonchopria zonalis</b>	94,64%	94,29%
<b>Rhodometra sacraria</b>	93,08%	97,45%
<b>Mathania leucothea</b>	83,04%	95,36%

<b>Colletes bicolor</b>	93,60%	96,40%
<b>Lasionota rouletii</b>	98,53%	97,29%
<b>Polythysana rubescens</b>	89,89%	91,97%
<b>Ormiscodes joiceyi</b>	90,99%	95,96%
<b>Centris buchholzi</b>	93,75%	97,96%
<b>Corynura herbsti</b>	96,14%	98,09%
<b>Chalepogenus caeruleus</b>	86,67%	92,86%
<b>Epistomentis pictus</b>	87,45%	92,08%
<b>Ruizantheda mutabilis</b>	88,43%	94,53%
<b>Tatochila theodice</b>	90,98%	98,53%
<b>Caenurgia runica</b>	85,10%	93,89%
<b>Xanthogaleruca luteola</b>	98,24%	97,98%
<b>Autoplusia gammoides</b>	95,29%	90,35%
<b>Chirgus fides</b>	93,33%	99,12%
<b>Orgyia antiqua</b>	92,94%	93,03%
<b>Svastra flavitarsis</b>	90,55%	97,07%
<b>Castnia eudsmia</b>	98,32%	97,86%
<b>Rheumaptera exacta</b>	84,66%	91,45%
<b>Auca barrosi</b>	99,16%	99,21%
<b>Ormiscodes schmidtnielseni</b>	94,54%	97,87%
<b>Cosmosatyrus leptoneuroides</b>	95,02%	97,76%
<b>Colibri coruscans</b>	99,32%	100,00%
<b>Argyrophorus chiliensis</b>	96,15%	92,65%
<b>Strymon bicolor</b>	97,06%	95,96%
<b>Toxomerus calceolatus</b>	97,06%	97,92%
<b>Lerodea eufala</b>	93,44%	96,63%
<b>Copestylum scutellatum</b>	94,57%	97,45%
<b>Isepeolus septemnotatus</b>	92,16%	96,42%
<b>Endrosis sarcitrella</b>	97,06%	95,39%
<b>Eulidia yarrellii</b>	97,55%	100,00%
<b>Eristalinus taeniops</b>	94,85%	99,94%
<b>Austroscaeva occidentalis</b>	89,95%	93,69%
<b>Hagnagora vittata</b>	88,73%	95,58%
<b>Plodia interpunctella</b>	98,04%	98,88%
<b>Romanophora verecunda</b>	99,51%	98,55%
<b>Chalepogenus herbsti</b>	89,84%	99,08%
<b>Eroessa chiliensis</b>	95,45%	93,61%
<b>Dioprosopa clavata</b>	99,47%	99,93%
<b>Psilaspilates signistriata</b>	83,42%	93,53%
<b>Hylaeus punctatus</b>	95,19%	97,47%
<b>Alloscirtetica rufitarsis</b>	93,05%	94,37%
<b>Conognatha viridiventris</b>	92,78%	95,53%

<b>Platycheirus fenestrata</b>	91,47%	95,45%
<b>Callistochlora chloris</b>	92,35%	97,37%
<b>Anthidium manicatum</b>	99,41%	100,00%
<b>Syrphus reedi</b>	92,65%	97,76%
<b>Strichosa eburata</b>	93,82%	97,05%
<b>Palpada meigenii</b>	96,47%	96,74%
<b>Catophoenissa dibapha</b>	92,06%	96,14%

#### 4.1.3. Mapas Árboles de Decisión

Como primer mapa potencial se desarrolló este ejemplo del polinizador *Sephanoides sephanoides* el cual muestra su distribución estimada porcentual (Figura 8). Igual finalmente fue descartado debido a su menor precisión comparado con los mapas desarrollados posteriormente usando el algoritmo con más precisión.

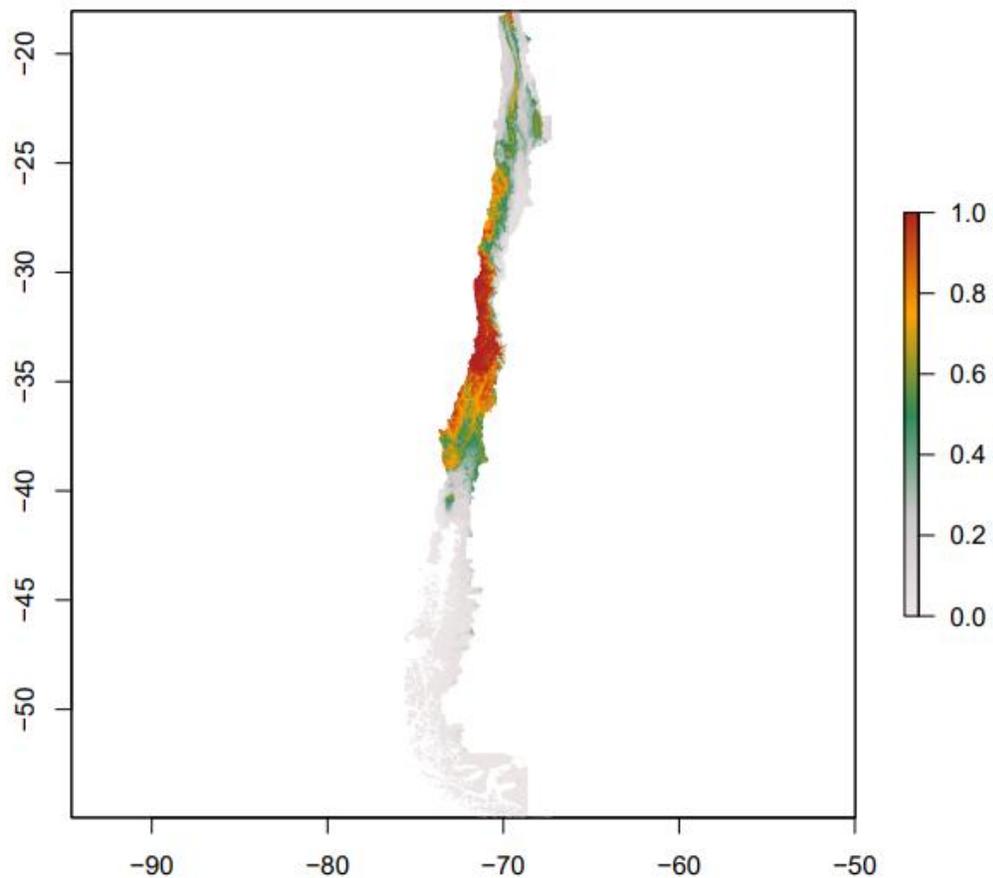


Figura 8. Mapa de distribución potencial *Sephanoides sephanoides* con Random forest.

#### 4.1.4. Mapas MaxEnt

El mapa MaxEnt (Figura 9) es la primera versión del algoritmo utilizado finalmente para modelar. Este no cuenta con el método de feature selection, pero sirvió para identificar las diferencias con su similar desarrollado por Árboles de Decisión, lo cual se puede notar contrastando ambas figuras (8 y 9) apreciando que la zona potencial para la especie cambia drásticamente.

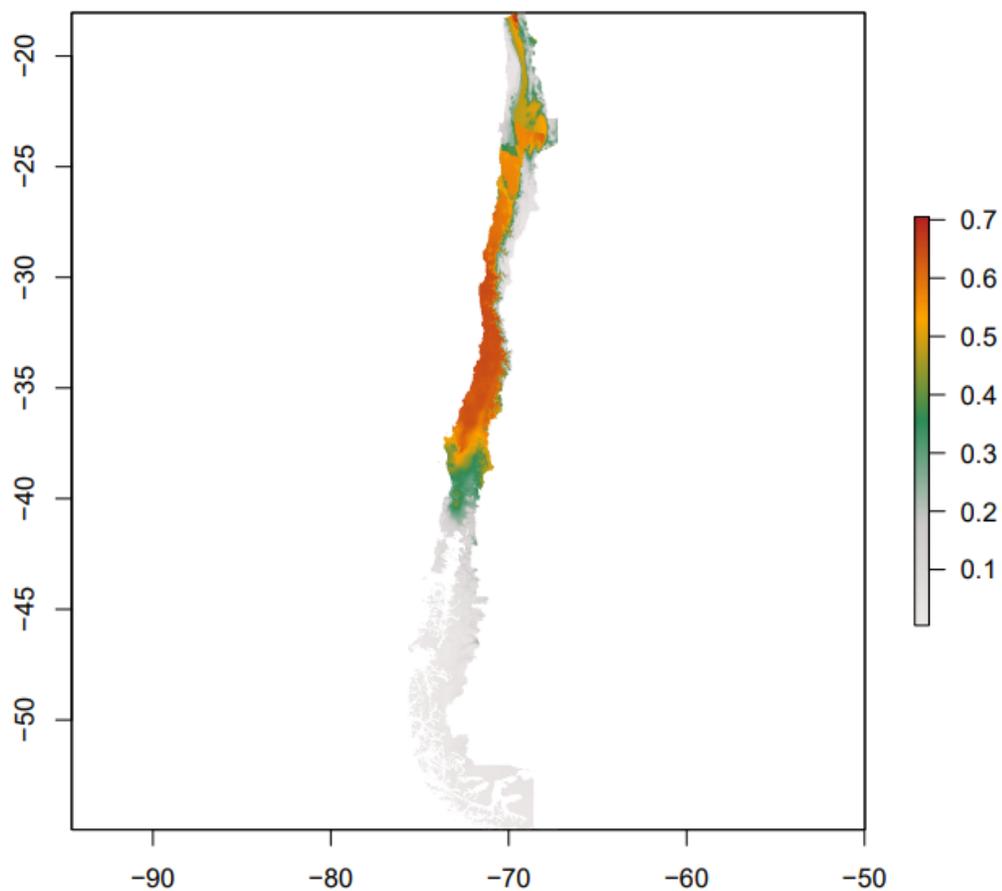


Figura 9. Mapa de distribución potencial *Sephanoides sephanioides* con MaxEnt.

Dentro de los 9.000 píxeles que se consideraron para este estudio, los cuales estaban en Chile continental, solo 7.458 tenían las características adecuadas para que habitara un

polinizador y de estos 2.715 fueron afectados por incendios, equivalente al 36% de su hábitat.

En la Figura 10 se muestra la gráfica AUC para el método MaxEnt, mientras que en la Tabla 5 se incluyen los promedios AUC finales del método de modelado diferenciado para especies normales y raras siendo las especies modeladas con el método antiguo solo *Sephanoides.sephaniodes*, *Bombus.dahlbomii*, *Bombus.terrestris*, *Apis.mellifera*, *Vanessa.carye* y *Astylus.trifasciatus*. Igualmente se agrega otro indicador que es el continuous boyce index el cual también es utilizado para medir desempeño de los modelos ausencia/presencia.

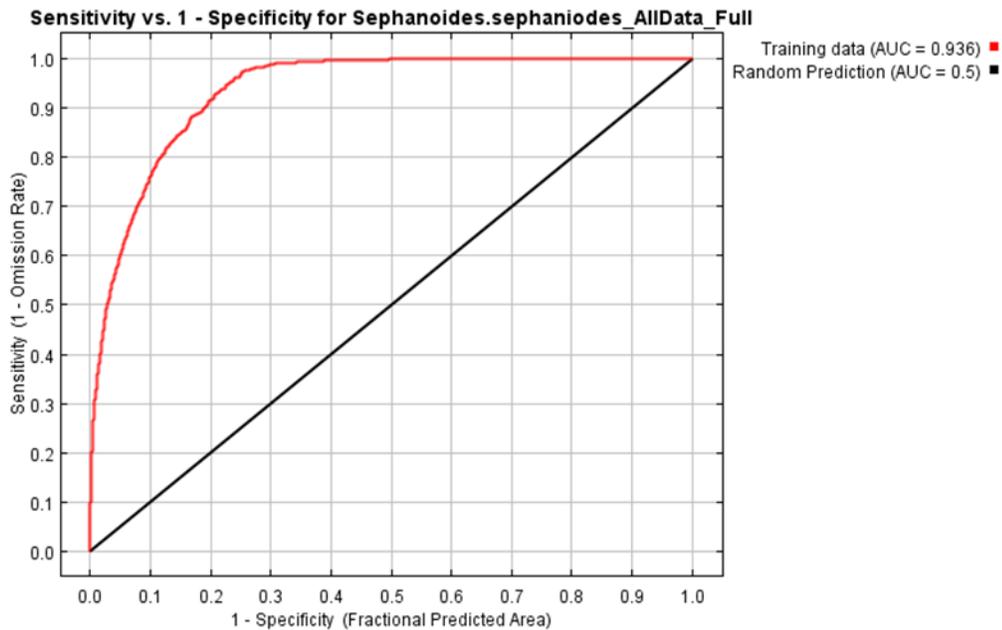


Figura 10. Gráficas AUC MaxEnt.

Tabla 5. AUC finales.

Especies	Promedio continuo boyce index	Promedio AUC
<b>Sephanoides.sephaniodes</b>	92,0%	86,0%
<b>Bombus.dahlbomii</b>	95,0%	78,0%
<b>Bombus.terrestris</b>	89,0%	89,0%
<b>Apis.mellifera</b>	92,0%	90,0%
<b>Vanessa.carye</b>	88,0%	90,0%

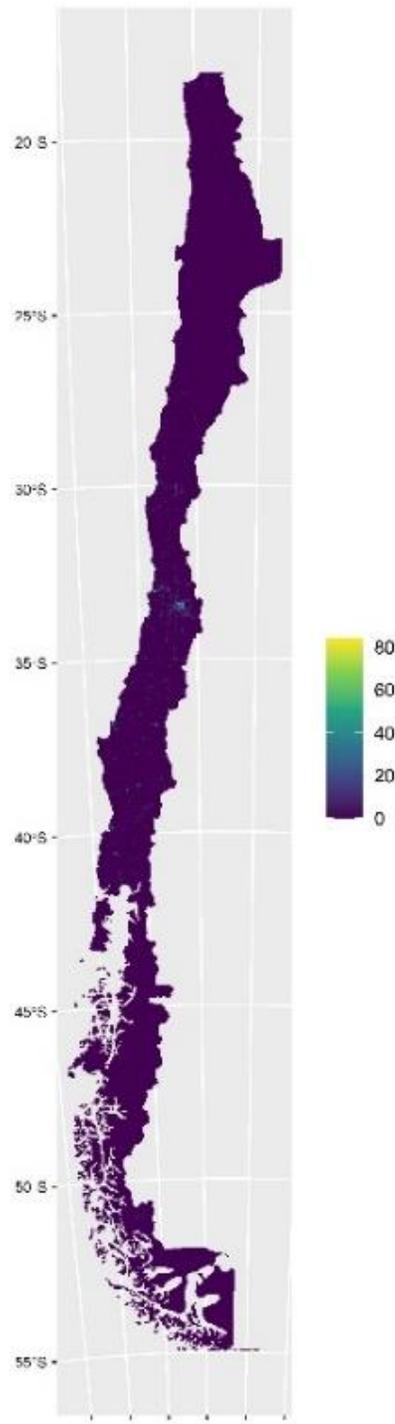
<b>Astylus.trifasciatus</b>	90,0%	90,0%
<b>Adetomeris.erythropros</b>	78,0%	87,0%
<b>Allograpta.hortensis</b>	67,8%	74,6%
<b>Allograpta.pulchra</b>	69,0%	77,5%
<b>Argyrophorus.chiliensis</b>	68,8%	73,4%
<b>Arthrobrachus.nigromaculatus</b>	66,2%	75,3%
<b>Auca.barrosi</b>	62,8%	79,0%
<b>Auca.coctei</b>	73,0%	78,8%
<b>Austroscaeva.melanostoma</b>	78,9%	79,0%
<b>Autoplusia.gammoides</b>	77,4%	79,9%
<b>Battus.polydamas</b>	78,9%	82,2%
<b>Bilyaxia.concinna</b>	76,6%	79,2%
<b>Bombus.ruderatus</b>	75,6%	78,1%
<b>Cadeguala.occidentalis</b>	77,8%	78,4%
<b>Caenurgia.runica</b>	82,7%	81,3%
<b>Camponotus.chilensis</b>	73,9%	79,9%
<b>Camponotus.morosus</b>	76,8%	82,2%
<b>Castnia.eudesmia</b>	73,7%	82,4%
<b>Centris.nigerrima</b>	79,6%	84,2%
<b>Chilesia.rudis</b>	85,3%	84,2%
<b>Chirgus.fides</b>	82,6%	85,6%
<b>Colias.vauthierii</b>	79,8%	84,3%
<b>Cylindrophora.maulica</b>	78,9%	83,6%
<b>Diphaglossa.gayi</b>	85,6%	85,2%
<b>Ectinogonia.buquetii</b>	81,2%	86,4%
<b>Ennada.flavaria</b>	83,2%	86,2%
<b>Epistomentis.pictus</b>	77,5%	88,2%
<b>Eristalis.tenax</b>	79,0%	87,0%
<b>Heliopetes.americanus</b>	81,8%	87,3%
<b>Hylephila.fasciolata</b>	75,4%	87,3%
<b>Hylephila.signata</b>	88,9%	88,2%
<b>Hyles.annei</b>	86,9%	89,1%
<b>Lasionota.rouletii</b>	89,6%	89,8%
<b>Leptotes.trigemmatas</b>	84,9%	89,8%
<b>Lygaeus.alboornatus</b>	78,3%	90,1%
<b>Malamblia.melas</b>	85,9%	88,4%
<b>Mathania.leucothea</b>	77,2%	83,3%
<b>Megachile.pollinosa</b>	78,2%	85,0%
<b>Megachile.saulcyi</b>	78,3%	90,0%
<b>Orgyia.antiqua</b>	80,4%	93,0%
<b>Ormiscodes.amphinome</b>	84,4%	90,8%

<b>Ormiscodes.cinnamomea</b>	84,7%	91,1%
<b>Patagona.gigas</b>	78,5%	90,9%
<b>Phaedon.semimarginatus</b>	66,7%	82,7%
<b>Phoebis.sennae</b>	75,7%	90,6%
<b>Pieris.brassicae</b>	71,1%	94,1%
<b>Polythysana.cinerascens</b>	78,7%	93,5%
<b>Polythysana.rubescens</b>	88,7%	92,2%
<b>Pseudolucia.chilensis</b>	80,3%	95,2%
<b>Rhodometra.sacraria</b>	84,3%	94,6%
<b>Ruizantheda.proxima</b>	80,2%	94,1%
<b>Ruizanthedella.mutabilis</b>	82,9%	96,7%
<b>Strymon.bicolor</b>	83,4%	96,8%
<b>Strymon.eurytulus</b>	85,9%	96,4%
<b>Syncirsodes.primata</b>	81,1%	95,3%
<b>Syrphus.octomaculatus</b>	78,3%	95,8%
<b>Tatochila.autodice</b>	82,8%	96,6%
<b>Tatochila.mercedis</b>	80,3%	96,1%
<b>Toxomerus.vertebratus</b>	80,7%	94,2%
<b>Vanessa.terpsichore</b>	85,6%	91,7%
<b>Xylocopa.augusti</b>	83,7%	99,0%
<b>Yramea.cytheris</b>	84,0%	95,7%

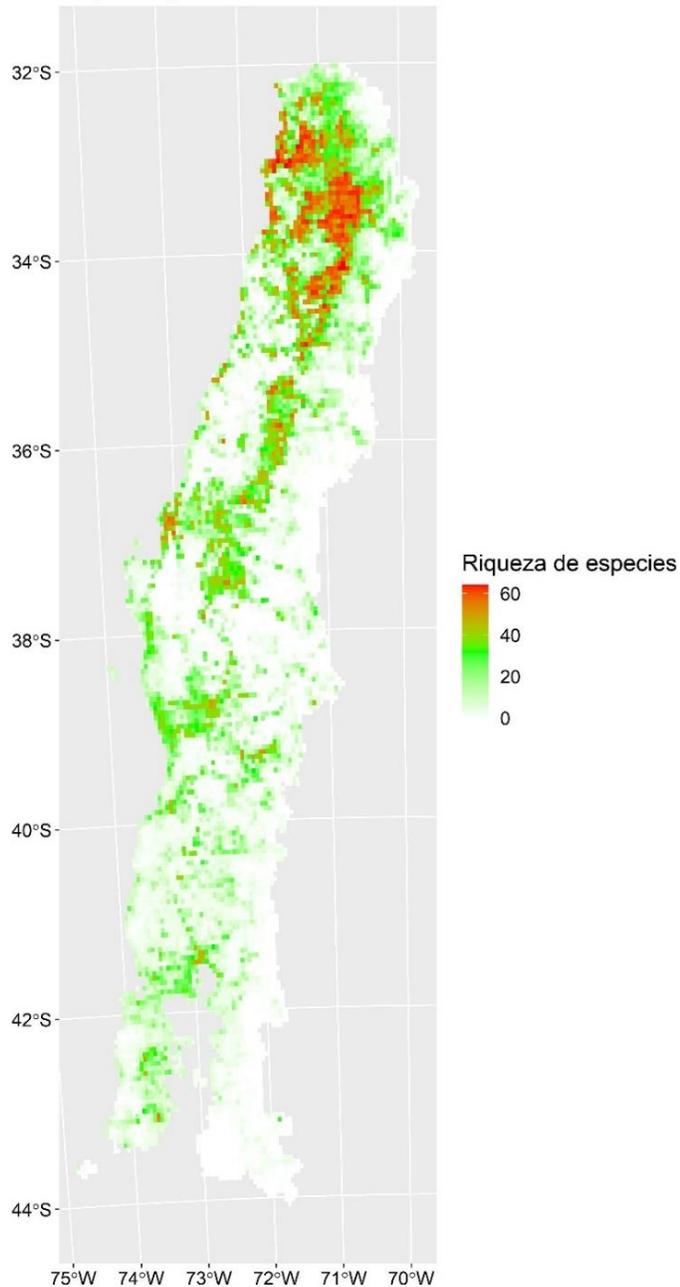
#### **4.2. Mapa potencial, observado y resultados con incendios (nueva capa de incendios)**

Inicialmente se diseñó el mapa de distribución observada (Figura 11) basado en todas las observaciones con la que se desarrolló el trabajo. Este mapa corresponde al primer paso por lo que es el más impreciso ya que solo se basa en las presencias observadas.

Luego el mapa de riqueza de especies de polinizadores potenciales (Figura 12) es el resultado de las últimas simulaciones de hábitat potencial de los polinizadores, destacando en un color rojo los píxeles de más importancia ya que presentan una mayor riqueza de especies. Allí podrían habitar un conjunto máximo de 60 especies de polinizadores, las cuales corresponden a la lista final de especies que se estudió. Los píxeles con mayor importancia son los que están más cerca del límite superior del mapa anterior, esta zona corresponde a la zona central de Chile, en este caso, cercana a la región metropolitana.

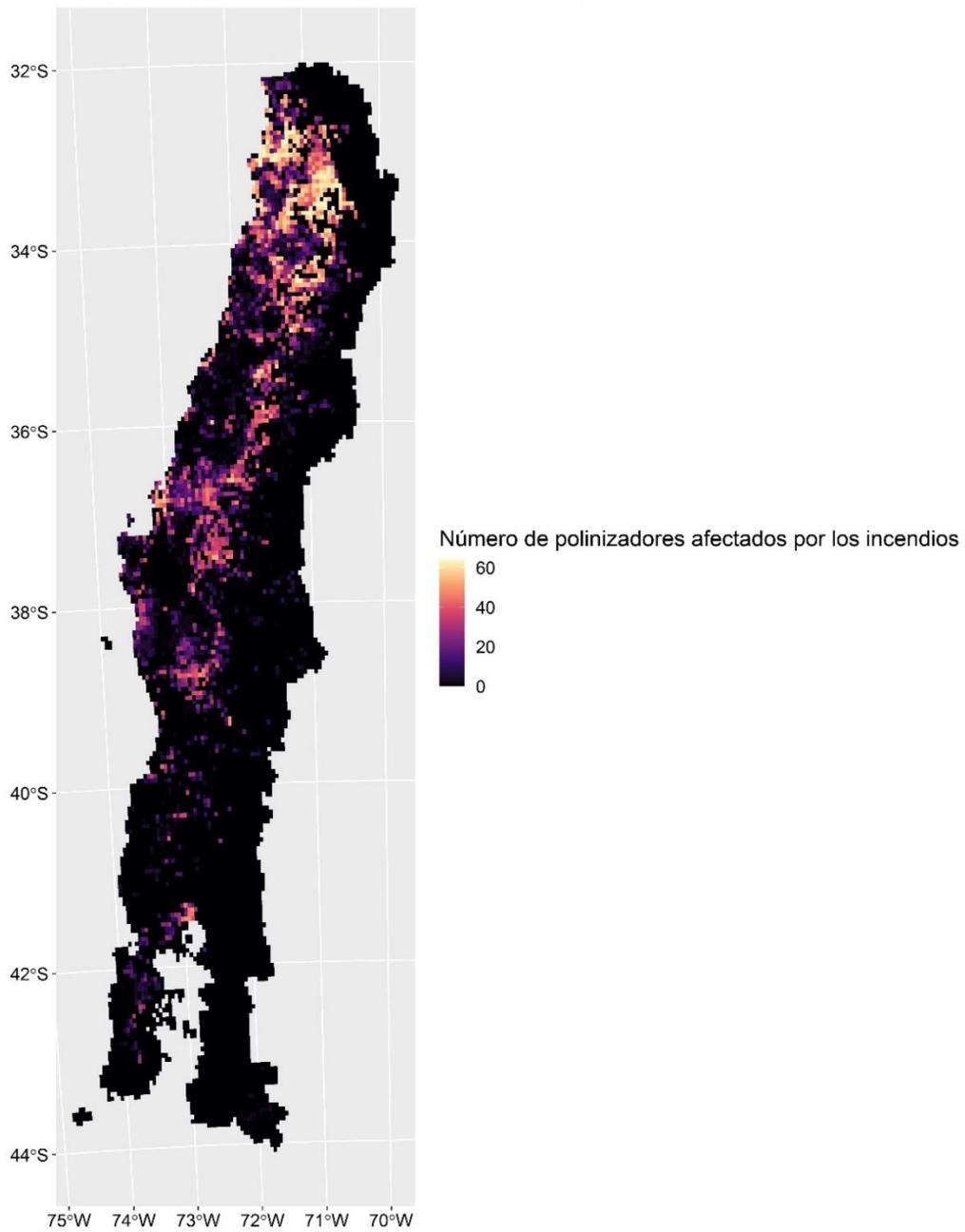


*Figura 11. Mapa de distribución observada.*



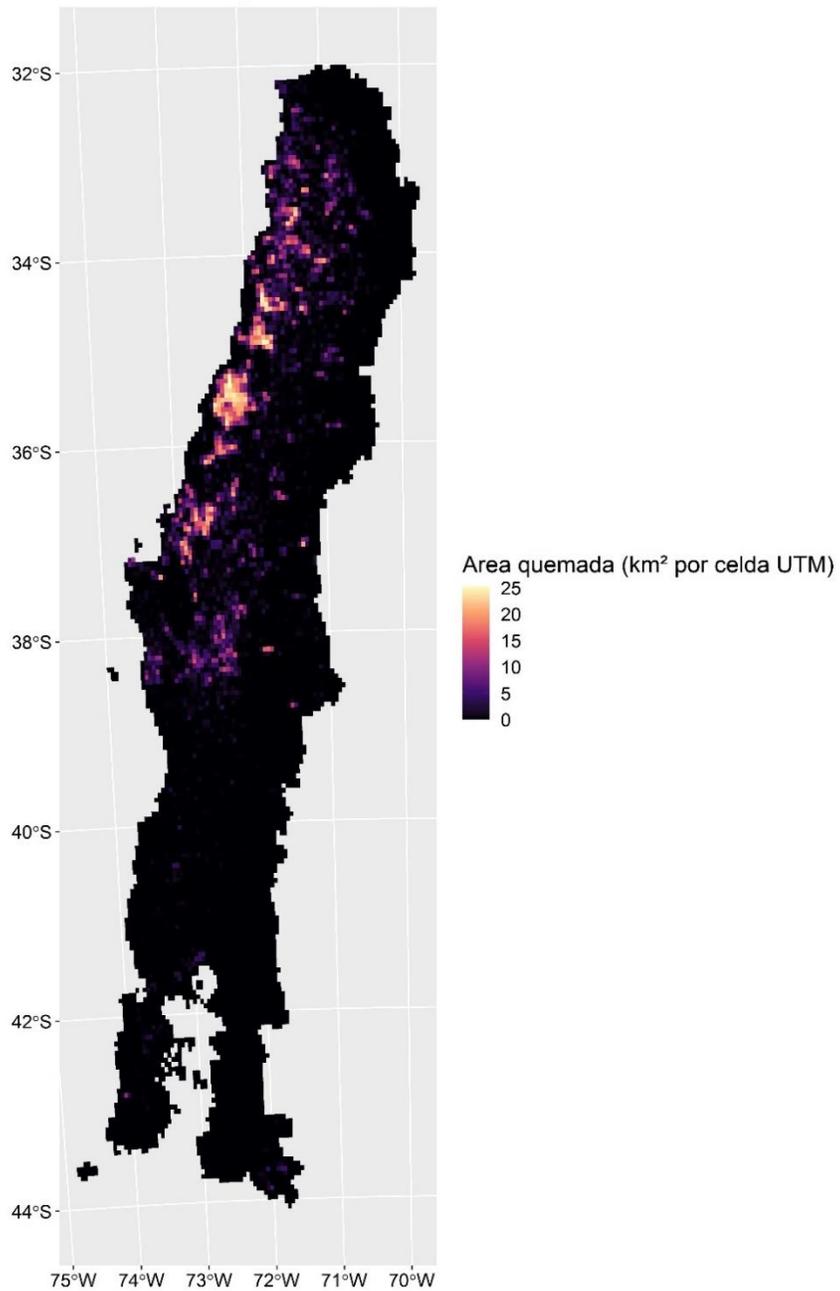
*Figura 12. Mapa de riqueza de especies.*

Una vez que se unió el mapa de riqueza con la información de incendios forestales, se logró identificar los píxeles afectados por los incendios forestales desde 2000 hasta 2018, creando así el mapa de Figura 13 que muestra de forma más precisa las zonas donde podrían habitar más polinizadores, aunque estas igualmente fueron afectadas, eliminando la flora con la que los polinizadores podrían alimentarse.



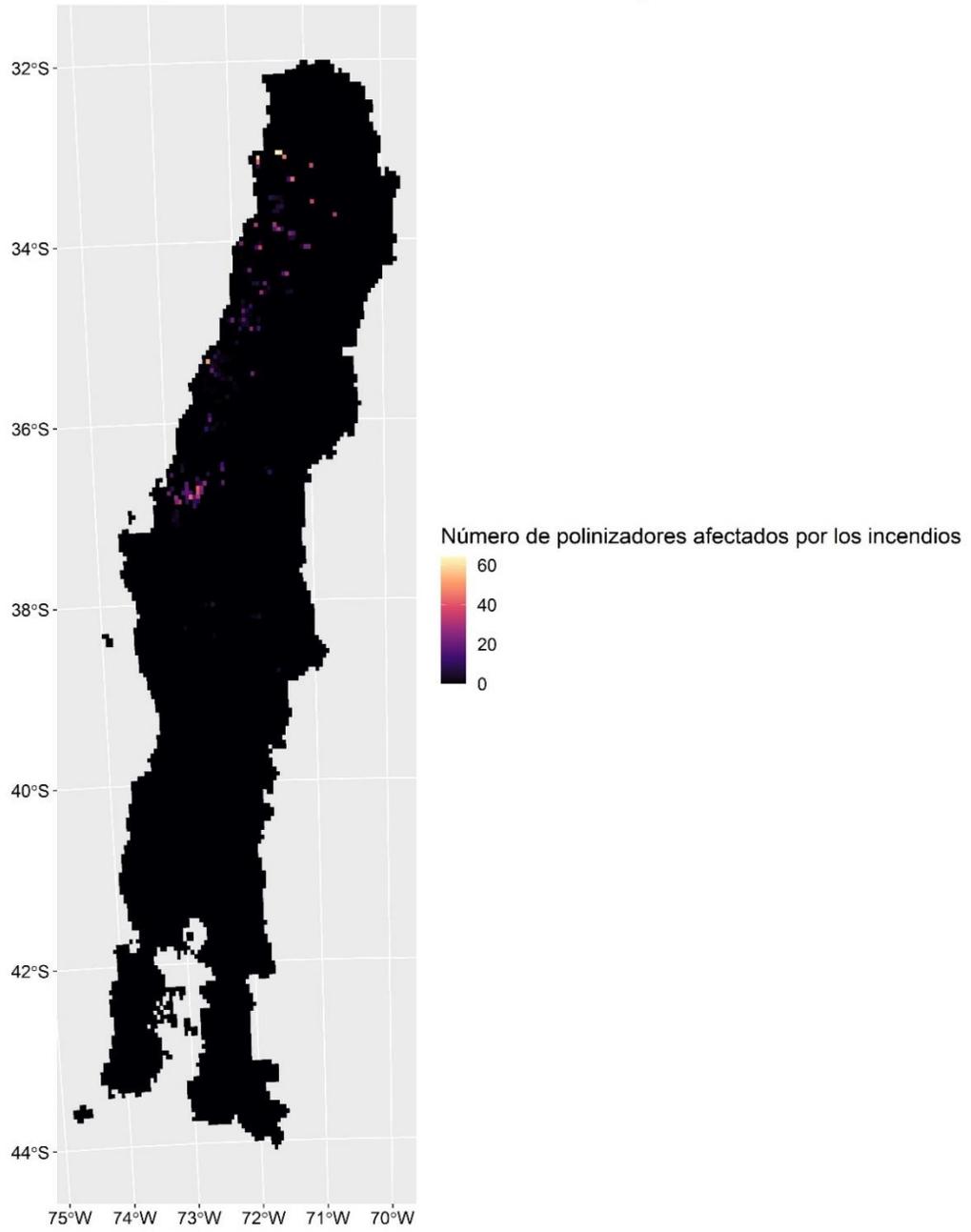
*Figura 13. Mapa de efectos de los incendios sobre el hábitat de los polinizadores.*

Como se describió anteriormente, la mayoría de los incendios ocurren en la zona central del país, lo cual se ve representado en el mapa de Figura 14, siendo las áreas más claras aquellas con más incendios desde el año 2000 en adelante.



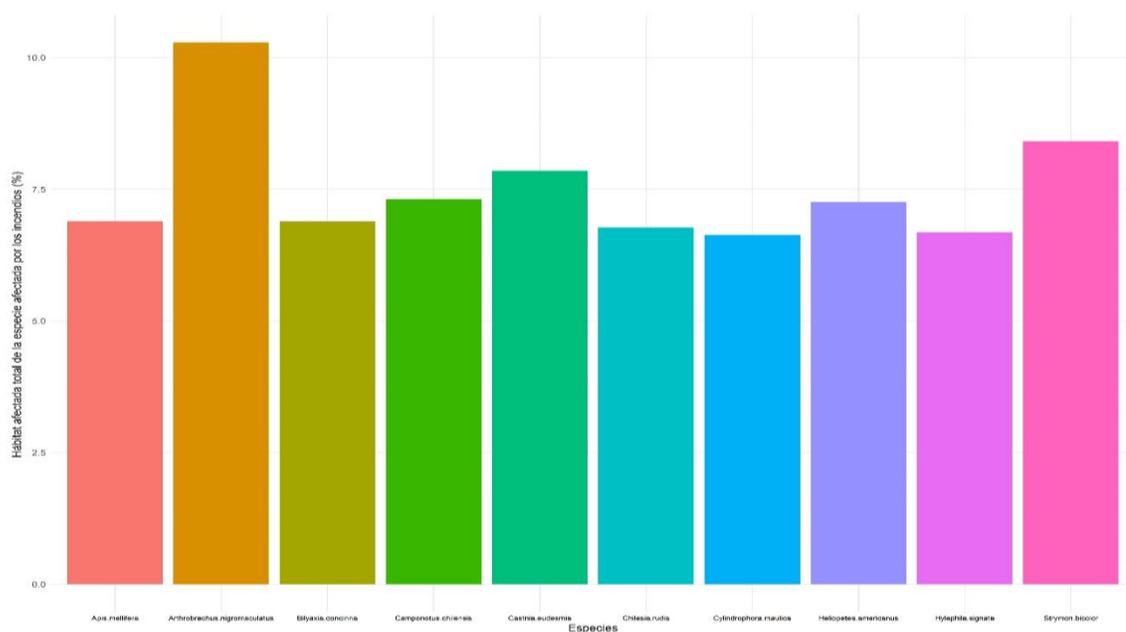
*Figura 14. Mapa de área afectada por los incendios.*

El mapa de Figura 15 muestra las áreas con efectos graves por incendios sobre el hábitat de polinizadores, donde se muestra el área destacada según el número de especies de polinizadores que fueron afectadas, con píxeles que tienen un área quemada mínima de más de un kilómetro cuadrado.



*Figura 15. Mapa de efectos graves de los incendios sobre el hábitat de polinizadores.*

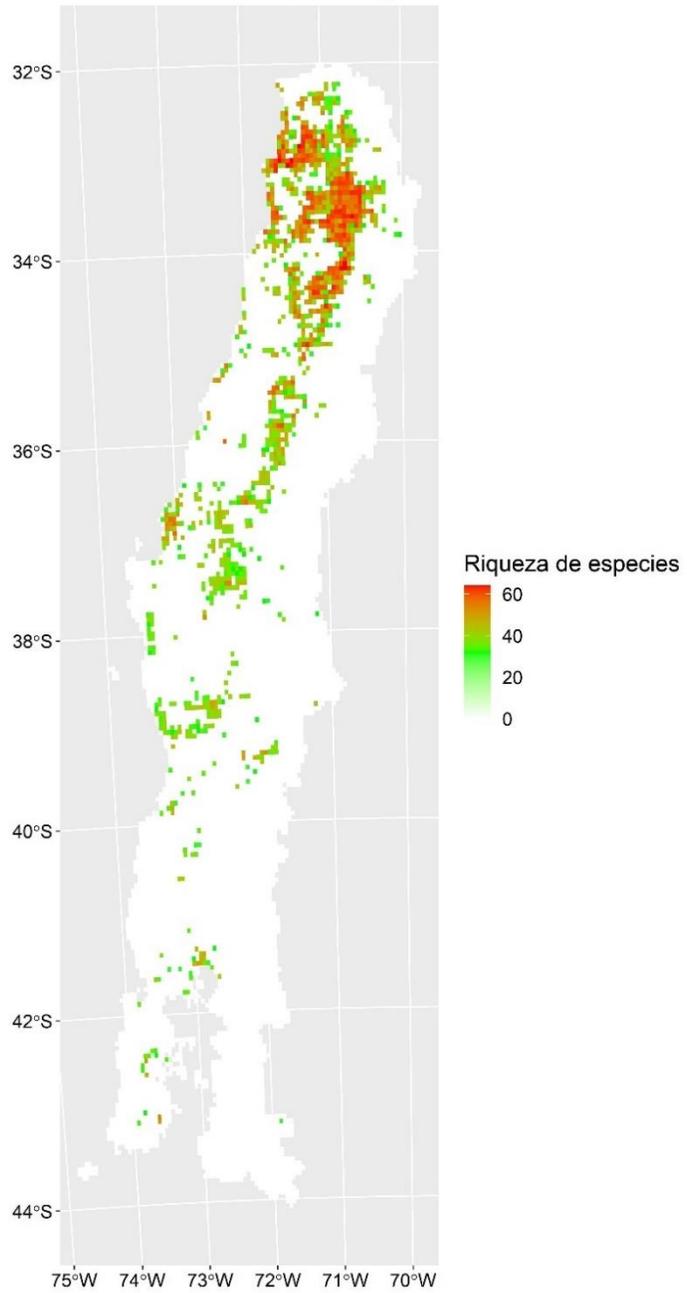
En base al mapa de efectos graves, se pudo generar el gráfico de Figura 16, que muestra las especies que perdieron más del 49% del área total de la celda UTM en cuestión (sobre 12,5 km cuadrados), afectando gravemente el hábitat potencial de los polinizadores.



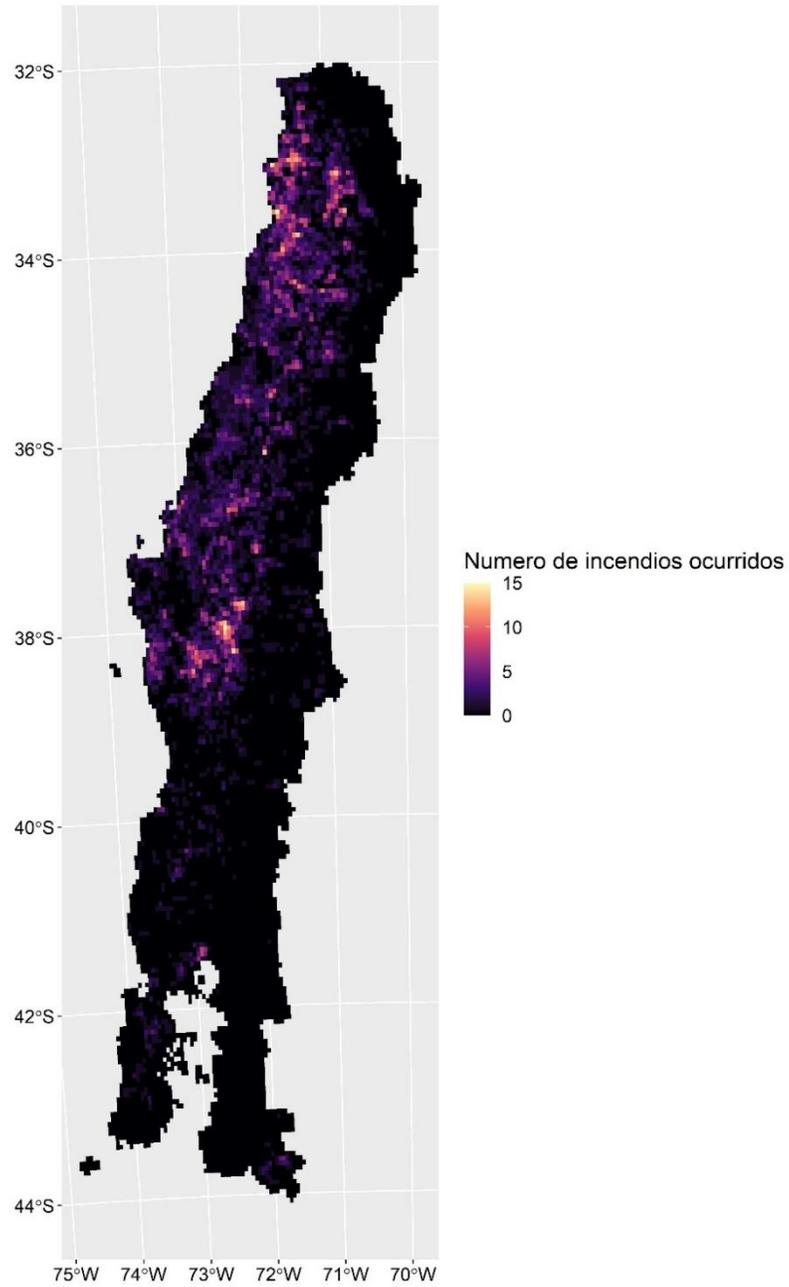
*Figura 16. Especies más afectadas por los incendios según hábitat total.*

El gráfico muestra que la especie que más hábitat potencial perdió por los incendios es *Arthrobrachus nigromaculatus*, lo que representa una de las situaciones más graves ya que un 10% de todo su hábitat potencial fue afectado por los incendios.

En contrapartida, la Figura 17 muestra un mapa del hábitat potencial de los polinizadores que no tuvo efectos directos por incendios, lo cual está en directa relación con el mapa de la Figura 18 que muestra las zonas con mayor recurrencia de incendios en los últimos 18 años.



*Figura 17. Mapa hábitat potencial de los polinizadores sin efectos directos de incendios.*



*Figura 18. Mapa con zonas de recurrencia de incendios.*

En Tabla 6 se muestra un recuento de hábitat por especie, indicando el área potencial total y el área de esta que fue quemada, por lo que se determinó lo que perdió cada especie en hectáreas y porcentaje.

Tabla 6. Hábitat perdido por especie.

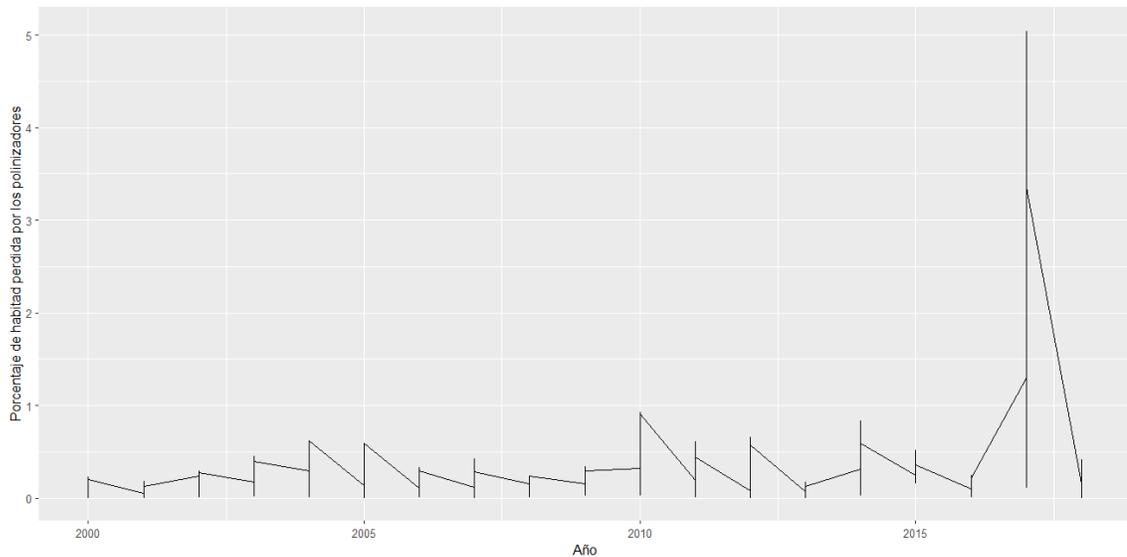
Especie	Clase	Estado de conservación	Hábitat total de la especie (km2)	Hábitat quemado (km2)	Hábitat perdido (km2)	Area total quemada (%)
<b>Arthrobrachus,nigromaculatus</b>	Insecto		44950.00	4622.64	10.28	10%
<b>Strymon,bicolor</b>	Insecto		31700.00	2665.80	8.41	8%
<b>Castnia,eudesmia</b>	Insecto		44725.00	3510.37	7.85	8%
<b>Camponotus,chilensis</b>	Insecto		45675.00	3339.47	7.31	7%
<b>Heliopetes,americanus</b>	Insecto		70450.00	5111.71	7.26	7%
<b>Apis,mellifera</b>	Insecto	DD	43450.00	2993.41	6.89	7%
<b>Bilyaxia,concinna</b>	Insecto		37350.00	2571.74	6.89	7%
<b>Chilesia,rudis</b>	Insecto		50800.00	3440.20	6.77	7%
<b>Hylephila,signata</b>	Insecto		28600.00	1909.27	6.68	7%
<b>Cylindrophora,maulica</b>	Insecto		38500.00	2553.34	6.63	7%
<b>Patagona,gigas</b>	Ave	LC	34450.00	2256.95	6.55	7%
<b>Auca,coctei</b>	Insecto		43025.00	2771.34	6.44	6%
<b>Battus,polydamas</b>	Insecto	LC	26875.00	1731.05	6.44	6%
<b>Phaedon,semimarginatus</b>	Insecto		38350.00	2413.79	6.29	6%
<b>Camponotus,morosus</b>	Insecto		42100.00	2599.60	6.17	6%
<b>Centris,nigerrima</b>	Insecto		35050.00	2148.26	6.13	6%
<b>Tatochila,mercedis</b>	Insecto		40600.00	2466.65	6.08	6%
<b>Vanessa,carye</b>	Insecto		45725.00	2775.62	6.07	6%
<b>Rhodometra,sacraria</b>	Insecto		21275.00	1285.24	6.04	6%
<b>Strymon,eurytulus</b>	Insecto		19975.00	1179.58	5.91	6%
<b>Tatochila,autodice</b>	Insecto		38300.00	2258.08	5.90	6%
<b>Allograpta,hortensis</b>	Insecto		28625.00	1683.42	5.88	6%
<b>Astylus,trifasciatus</b>	Insecto		39325.00	2295.67	5.84	6%
<b>Cadeguala,occidentalis</b>	Insecto		44100.00	2567.45	5.82	6%
<b>Adetomeris,erythrospis</b>	Insecto		48700.00	2833.13	5.82	6%
<b>Allograpta,pulchra</b>	Insecto		34025.00	1955.75	5.75	6%
<b>Mathania,leucothea</b>	Insecto		11575.00	651.60	5.63	6%
<b>Autoplusia,gammoides</b>	Insecto		19225.00	1075.64	5.59	6%
<b>Syncirsodes,primata</b>	Insecto		50625.00	2778.80	5.49	5%
<b>Argyrophorus,chiliensis</b>	Insecto		49925.00	2737.81	5.48	5%
<b>Ormiscodes,amphinome</b>	Insecto		60350.00	3241.13	5.37	5%
<b>Ectinogonia,buquetii</b>	Insecto		30625.00	1642.73	5.36	5%
<b>Ruizanthedella,mutabilis</b>	Insecto		18900.00	997.94	5.28	5%

<i>Pieris,brassicae</i>	Insecto	LC	27300.00	1426.13	5.22	5%
<i>Ormiscodes,cinnamomea</i>	Insecto		62000.00	3172.19	5.12	5%
<i>Pseudolucia,chilensis</i>	Insecto		23300.00	1192.03	5.12	5%
<i>Orgyia,antiqua</i>	Insecto		24675.00	1248.51	5.06	5%
<i>Eristalis,tenax</i>	Insecto	LC	52700.00	2638.74	5.01	5%
<i>Phoebis,sennae</i>	Insecto	LC	26200.00	1304.76	4.98	5%
<i>Megachile,pollinosa</i>	Insecto		36575.00	1820.78	4.98	5%
<i>Syrphus,octomaculatus</i>	Insecto		47950.00	2355.93	4.91	5%
<i>Auca,barrosi</i>	Insecto		28625.00	1393.87	4.87	5%
<i>Polythysana,cinerascens</i>	Insecto		36325.00	1763.22	4.85	5%
<i>Lygaeus,alboornatus</i>	Insecto		19900.00	956.44	4.81	5%
<i>Megachile,saucyi</i>	Insecto		34225.00	1574.19	4.60	5%
<i>Bombus,terrestris</i>	Insecto	LC	64000.00	2857.60	4.47	4%
<i>Xylocopa,augusti</i>	Insecto		8875.00	395.97	4.46	4%
<i>Hylephila,fasciolata</i>	Insecto		36425.00	1613.58	4.43	4%
<i>Lasionota,rouletii</i>	Insecto		34300.00	1498.66	4.37	4%
<i>Chirgus,fides</i>	Insecto		15775.00	673.99	4.27	4%
<i>Vanessa,terpsichore</i>	Insecto		59650.00	2513.06	4.21	4%
<i>Leptotes,trigemmatas</i>	Insecto		16700.00	690.53	4.13	4%
<i>Malamblia,melas</i>	Insecto		74950.00	3088.75	4.12	4%
<i>Hyles,annei</i>	Insecto		19475.00	787.91	4.05	4%
<i>Ennada,flavaria</i>	Insecto		26125.00	1055.45	4.04	4%
<i>Yramea,cytheris</i>	Insecto		50450.00	2029.55	4.02	4%
<i>Caenurgia,runica</i>	Insecto		27050.00	1074.02	3.97	4%
<i>Toxomerus,vertebratus</i>	Insecto		33475.00	1284.40	3.84	4%
<i>Austrocaeva,melanostoma</i>	Insecto		33175.00	1247.39	3.76	4%
<i>Sephanoides,sephanoides</i>	Ave	LC	44125.00	1641.51	3.72	4%
<i>Bombus,ruderatus</i>	Insecto		63500.00	2224.92	3.50	4%
<i>Ruizantheda,proxima</i>	Insecto		34700.00	1103.96	3.18	3%
<i>Colias,vauthierii</i>	Insecto		53350.00	1632.03	3.06	3%
<i>Diphaglossa,gayi</i>	Insecto		88250.00	2632.45	2.98	3%
<i>Bombus,dahlbomii</i>	Insecto	EN	74300.00	1169.17	1.57	2%
<i>Polythysana,rubescens</i>	Insecto		53850.00	799.78	1.49	1%
<i>Epistomentis,pictus</i>	Insecto		29825.00	201.16	0.67	1%

De acuerdo a lo anterior, se estima que en promedio los polinizadores considerados en este estudio (67 especies) perdieron un 5% de su habitad debido a los incendios. La gráfica

de la Figura 19 muestra el comportamiento del hábitat de los polinizadores perdido por incendios forestales desde 2000 hasta 2018.

Cabe mencionar que el área habitable potencial de las especies compone una zona de 186450 km<sup>2</sup> aproximadamente, de los cuales se quemaron alrededor de 8000 km<sup>2</sup> debido a los incendios forestales, lo que representa un 4,2% aproximadamente del área total habitable.



*Figura 19. Hábitat perdido por los polinizadores anualmente desde 2000 hasta 2018.*

### **4.3. Herramienta de visualización para la toma de decisiones**

La herramienta en su primera vista de página contempla 2 selectores que principalmente apuntan a mostrar la distribución potencial de los polinizadores juntos con mapas de su riqueza (Figura 20 y 21). Esto fue desarrollado para ayudar a tomadores de decisiones que desean ver la distribución de una determinada especie con motivos de conservación o estudio, pudiendo visualizar dónde se podrían encontrar estas especies. Por su parte, el mapa de riqueza les permite visualizar fácilmente en que zonas se encuentran las mayores concentraciones de especies de polinizadores distintas.

## Selector de Mapas Simulados

Seleccione el tipo de mapa

Descargar Mapa

Descargar Base de datos

Figura 20. Inicio de herramienta de visualización.

## Selector de Mapas Simulados

Mapas de distribución potencial estimada

Bombus dahlbomii

Descargar Mapa

Descargar Base de datos

Dist Pot. Bombu

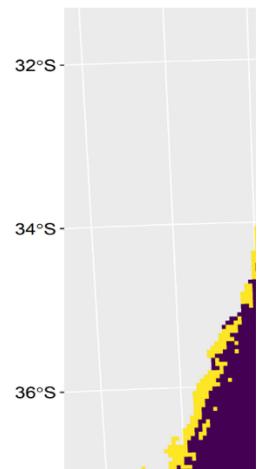
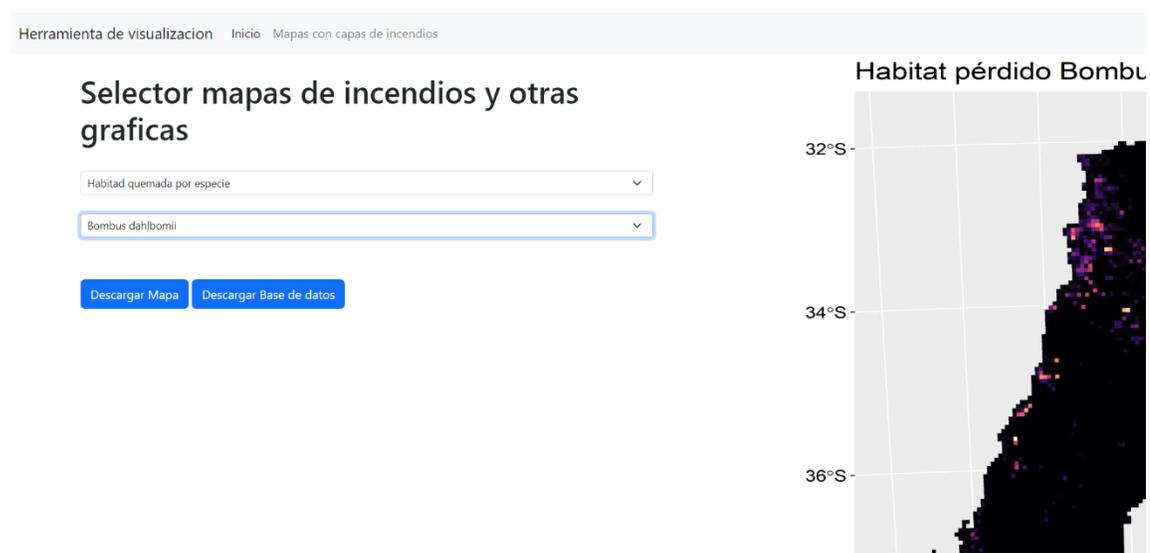


Figura 21. Ejemplo 1 de selección en la herramienta.

La segunda página permite ver los resultados contrastados con los incendios, proporcionando gráficas, mapas de la distribución de los incendios dentro del área de los polinizadores individualmente, mostrando además zonas donde hay incendios más recurrentes, áreas aun no afectadas directamente por los polinizadores, entre otros resultados (Figuras 22, 23 y 24).

Esto permitirá a los tomadores de decisiones e investigadores visualizar de una manera más rápida los daños de los incendios de una especie o del conjunto de 67 especies, lo cual puede ser útil para apoyar planes de conservación específicos, o para otras investigaciones que podrían verificar mediante el análisis en terrenos consecuencias indirectas o no medidas en este proyecto.



*Figura 22. Ejemplo 2 de selección en la herramienta.*

En términos generales, la finalidad principal de esta herramienta es servir como base para generar nueva información más específica a futuro, constituyendo por ahora una guía para decisiones futuras, como por ejemplo, cuantificar cifras de daños aproximados.

La herramienta actualmente se encuentra en línea en la página <http://Polinizadores Data CL.surge.sh> debido a que este es un servidor gratuito que sirve para visualizar y revisar esta primera versión.

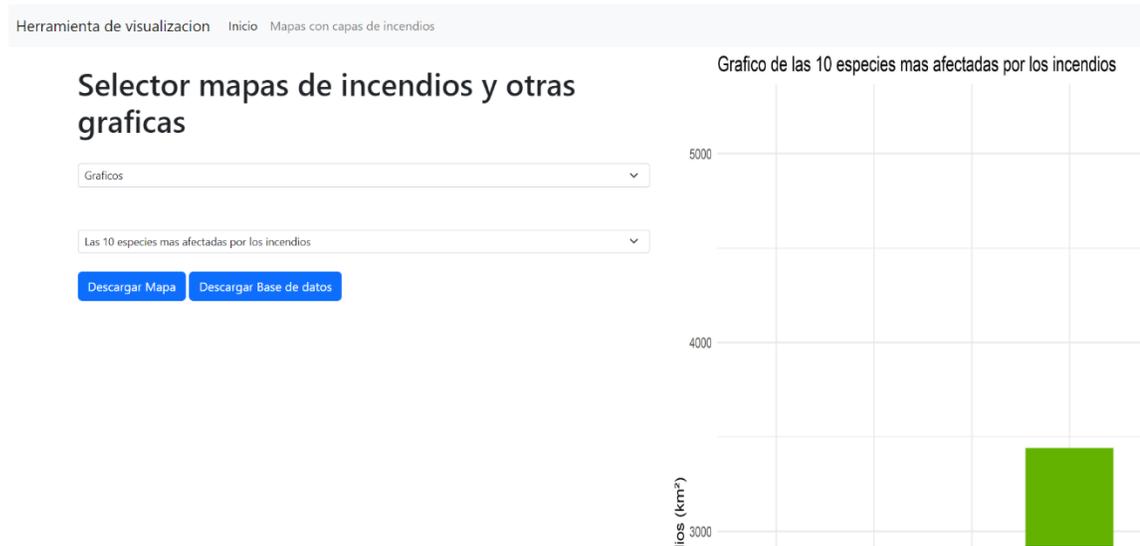


Figura 23. Ejemplo 3 de selección en la herramienta.

The screenshot shows a table of species affected by fires. The table has four columns: 'Id de especies afectadas', 'Clase', 'Estado de conservación', and 'Especie'. The table lists 30 species, including various birds, bees, and other insects. The species names are listed in the 'Especie' column, and their corresponding IDs are in the 'Id de especies afectadas' column.

Id de especies afectadas	Clase	Estado de conservación	Especie
1641.5387	Aves	LC	Sephanoides sepharoides
1169.1713391104	Insecta	EN	Bombus darwini
2887.801331548395	Insecta	LC	Bombus terrestris
2983.60883932008	Insecta	DD	Apis mellifera
2778.82547138287	Insecta	NA	Vanessa carye
2266.8714834864	Insecta	NA	Artibeus affinis
2266.89121811719	Aves	LC	Falco sparverius
1574.1623122209	Insecta	NA	Megachile laevis
2887.6547899483	Insecta	NA	Colletes cacaeciarum
2638.74148210438	Insecta	LC	Eristalis taraxacum
3466.6521933265	Insecta	NA	Trochilidactylus
1622.7709674124	Insecta	NA	Megachile patinoana
1622.62810380228	Insecta	NA	Colletes caudatum
1613.5883441786	Insecta	NA	Hyalictus leucostictus
3172.188854827	Insecta	NA	Osmia delongi
1102.897502403	Insecta	NA	Aculepeira proxima
2632.44834813338	Insecta	NA	Diprionidactylus
1424.12877649801	Insecta	LC	Pteris caudata
2148.26387320823	Insecta	NA	Centris nigripes
395.89647043518	Insecta	NA	Hylocichla ustulata
1284.39814483788	Insecta	NA	Toxomeris venestratus
3341.12821448502	Insecta	NA	Osmia delongi
1731.04819795621	Insecta	LC	Bombus polydorus
680.53219073079	Insecta	NA	Lepidoptera eremita
2771.32789720441	Insecta	NA	Auca doctus
2313.0588607378	Insecta	NA	Vanessa terpsichore
2391.73709897071	Insecta	NA	Blythia condita
787.80687121487	Insecta	NA	Hyalictus
2268.87708128914	Insecta	NA	Trochilidactylus
2029.89388482289	Insecta	NA	Tramea carye
1642.7284848487	Insecta	NA	Estrogonia busuetti
1683.41835882887	Insecta	NA	Allograpta horreana

Figura 24. Ejemplo 4 de selección en la herramienta.

## V. Conclusiones

En Chile, si bien existen muchos datos registrados sobre polinizadores y sobre incendios, no existen estudios acerca de cómo los primeros son afectados por los segundos. Esto, en un país que presenta gran cantidad de incendios principalmente en la zona central, ocasiona que los tomadores de decisiones u otras partes interesadas no tengan la información correcta para diseñar sus planes de contingencia, prevención o conservación.

La presente investigación delimitó y cuantificó la pérdida del hábitat potencial de los polinizadores debido a los incendios forestales ocurridos desde 2000 hasta 2018, generando resultados tanto gráficos como numéricos. De esta forma, se ha podido estimar que el área total que los polinizadores podrían llegar a utilizar es de 186.450 km<sup>2</sup>, donde el área quemada es de alrededor del 4,2%, lo que arroja un promedio de hábitat perdido por especie de 5%. Esto contrasta con el 36% de píxeles habitables por polinizadores determinados como afectados por incendios, dado que este último porcentaje considera como perdida cada píxel donde se produjo un incendio, sin importar el área de este que haya sido afectada.

Así mismo, se estimó que el porcentaje perdido por especies en peligro es un 2%, siendo esta *Bombus dahlbomii* la única en nuestro país de las estudiadas que tiene ese estatus. En el caso de las especies clasificadas como sin datos suficientes según la IUCN, han perdido un 7% de su hábitat equivalente a 3.000 km<sup>2</sup>, siendo la *Apis mellifera* la única en esta categoría. Por su parte, las especies que son clasificadas como preocupación menor han perdido un 5%.

Es de suma importancia recalcar que las observaciones usadas para poder entrenar los modelos de minería de datos proceden de observación de ciudadanos o investigadores y entidades como GBIF o iNaturalist, que reportan avistamiento de la presencia de un polinizador en determinadas coordenadas. Por ello no es posible concluir que, en caso de no tener registro de un polinizador en una zona con un incendio registrado, este haya sido afectado al presentar dicho lugar condiciones óptimas. Considerando esto, el tener bases

de datos que ratifiquen la ausencia de especies (lo cual es una tarea ardua, ya que implicaría buscar polinizadores en terreno a lo largo de todo el país), facilitaría mucho la realización de mejores análisis, los cuales podrían ser de provecho para elaborar planes preventivos de conservación de especies mucho más efectivos.

Se debe considerar también que los incendios aumentan en tamaño cada vez a lo largo de los años, y si bien no se dispuso de información sobre la intensidad de dichos incendios en el presente estudio, se puede entender que si los incendios tienen cada vez mayor extensión o nivel de daño, existirá mayor probabilidad de que las zonas que se ven afectadas por su aparición terminen con un daño mucho mayor, afectando con ello la disponibilidad y calidad de los hábitats.

Por otra parte, la herramienta presentada permite que los tomadores de decisiones que tengan relación con entes polinizadores puedan tener una idea del valor potencial de un cierto territorio para una especie, lo cual podría servir para poder diseñar planes de protección de especies más específicos. Esto gracias a que se puede visualizar en ella el área que aún no ha sido afectada por los incendios, además de la riqueza de polinizadores presente en ese territorio. Si bien una meta del proyecto era generar una herramienta que permitiera simular las pérdidas, la falta de datos de ausencia/presencia de polinizadores y del antes y después de un incendio no permitió llegar a ello. No obstante, la herramienta cumplen cabalmente con la finalidad de difundir y facilitar la obtención de información y datos en estas materias, pudiendo por ejemplo, apoyar el diseño de planes de contingencia o de prevención de incendios para aquellas zonas que presenten más incendios recurrentes a lo largo del tiempo.

En cuanto a los modelos de minería de datos, se pudo evidenciar que los modelos como Árboles de Decisión funcionan mejor en aquellas especies que tienen más observaciones, a diferencia del modelo MaxEnt Philips que si bien no deja de tener un buen desempeño en las instancias donde existen grandes cantidades de observaciones, obtiene mayor AUC en especies con menores observaciones, siendo este el caso para la mayoría de los agentes polinizadores que abarca este estudio.

Es esperable que otras organizaciones puedan facilitar el acceso a los datos relacionados con polinizadores, dado que, si bien fue posible acceder a estos al momento de trabajar en este proyecto, el tener no tener que realizar Web Scrapping o limpiar datos de otros países incluidos como si estuvieran en territorio chileno, podría facilitar más los estudios a futuro en el área.

Finalmente, es importante destacar que esta investigación no es más que la primera de la que esperamos sean muchos estudios sobre el daño de los incendios en los agentes polinizadores. Así mismo esperamos que tanto esta investigación como los datos facilitados vía herramienta Web sean de utilidad para los investigadores y faciliten los estudios posteriores en el área como, por ejemplo, usar los datos generados para comprobar dentro de las zonas incendiadas si existen registros de una especie de polinizador, de forma que se pueda saber si el incendio en cuestión alteró el interés de los polinizadores en el área.

## VI. Bibliografía

Albon, C. (2018). Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning. O'Reilly Media.

Bolstad, P. (2016). GIS Fundamentals: A First Text on Geographic Information Systems, Fifth Edition (5th edition). XanEdu Publishing Inc.

Breiner, F. T., Guisan, A., Bergamini, A. & Nobis, M. P. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, 6(10), 1210-1218. <https://doi.org/10.1111/2041-210x.12403>

Breiner, F. T., Nobis, M. P., Bergamini, A. & Guisan, A. (2018). Optimizing ensembles of small models for predicting the distribution of species with few occurrences. *Methods in Ecology and Evolution*, 9(4), 802-808. <https://doi.org/10.1111/2041-210x.12957>

Brown, J., York, A., Christie, F., & McCarthy, M. (2016). Effects of fire on pollinators and pollination. *Journal of Applied Ecology*, 54(1), 313-322. <https://doi.org/10.1111/1365-2664.12670>

Chang, K. (2019). Geographic Information System. *International Encyclopedia of Geography*, 1-10.

Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J. & Liu, H. (2018, 30 noviembre). Feature Selection. *ACM Computing Surveys*, 50(6), 1-45. <https://doi.org/10.1145/3136625>

CONAF. (s. f.). Recuperado 11 de octubre de 2022, de <https://www.conaf.cl/incendios-forestales/incendios-forestales-en-chile/>

CONAF (2006). Manual con Medidas para la Prevención de Incendios Forestales IX Región. Gobierno de Chile. [https://www.conaf.cl/wp-content/files\\_mf/1367247874manualaraucania.pdf](https://www.conaf.cl/wp-content/files_mf/1367247874manualaraucania.pdf)

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F. T., D'Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R. G., Hordijk, W., Salamin, N. & Guisan, A. (2017). ecospat: an R package to support spatial analyses and

modeling of species niches and distributions. *Ecography*, 40(6), 774-787.  
<https://doi.org/10.1111/ecog.02671>

Elith, J. & Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32(1), 66-77. <https://doi.org/10.1111/j.1600-0587.2008.05505.x>

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129-151.  
<https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., & Yates, C.J. (2010). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43-57.  
<https://doi.org/10.1111/j.1472-4642.2010.00725.x>

Fadl Elsid, T.O., & Eltahir M.A. (2014). *Data Mining: Classification Techniques of Students' Database A Case Study of the Nile Valley University*, North Sudan. *International Journal of Computer Trends and Technology (IJCTT)*, 16, p. 192–203.

Fawcett, T. (2006, junio). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>

Fick, S., R. Hijmans, 2017. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37 (12): 4302-4315.

for Species Distribution Modeling. R package version 3.3-7.1.

García García, M., Ríos Osorio, L. A., & Álvarez Del Castillo, J. (2016). La polinización en los sistemas de producción agrícola: revisión sistemática de la literatura. *Idesia (Arica)*, 34(3), 53–68. <https://doi.org/10.4067/s0718-34292016000300008>

Global Biodiversity Information Facility. (no publicado). GBIF API (Versión V1) [Conjunto de datos].

- Grus, J. (2019). *Data Science from Scratch: First Principles with Python* (2nd ed.). O'Reilly Media
- Han, J., Kamber, M., Pei, J. (2012). *Data Mining Concepts and Techniques*. Burlington, Massachusetts: Morgan Kaufmann Publishers.
- Harrell, F. E., Lee, K. L., Califf, R. M., Pryor, D. B. & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine*, 3(2), 143-152. <https://doi.org/10.1002/sim.4780030207>
- Heikkinen, R. K., Marmion, M. & Luoto, M. (2012). Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, 35(3), 276-288. <https://doi.org/10.1111/j.1600-0587.2011.06999.x>
- Hurwitz, J., & Kirsch, D. (2018). *Machine Learning for Dummies®*, IBM Limited Edition (IBM Limited Edition). John Wiley & Sons, Inc.
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2021, 30 julio). *An Introduction to Statistical Learning: With Applications in R* (2nd 2021 ed.). Springer.
- Millennium Ecosystem Assessment. (2005). *Ecosystems and Human Well-Being: Synthesis* (2nd ed.). Island Press.
- Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists* (1.a ed.). O'Reilly Media.
- Nates-Parra, G. (ed.). (2016). *Iniciativa Colombiana de Polinizadores ~ Abejas ~ ICPA*. Bogotá, D. C. Departamento de Biología, Universidad Nacional de Colombia. 364 pp.
- Nyce, C. (2007). *Predictive analytics white paper*. America Institute for CPCU. Insurance Institute of America, 9-10.
- Oficina de Estudios y Políticas Agrarias. (2019). *Panorama de la agricultura chilena*. Gobierno de Chile. <https://www.odepa.gob.cl/wp-content/uploads/2019/09/panorama2019Final.pdf>
- Patel, A.A. (2019). *Hands-On Unsupervised Learning Using Python: How to Build Applied Machine Learning Solutions from Unlabeled Data*. O'Reilly Media.

Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3-4), 231-259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>

Porto, R.G., de Almeida, R.F., Cruz-Neto, O., Tabarelli, M., Viana, B.F., Peres, & Lopes, A.V. (2020). Pollination ecosystem services: A comprehensive review of economic values, research funding and policy actions. *Food Sec.* 12, 1425–1442. <https://doi.org/10.1007/s12571-020-01043-w>

Potts, S.G., Imperatriz-Fonseca, V., Ngo, H.T., Aizen, M.A., Biesmeijer, J.C., Breeze, T.D., Dicks, L.V., Garibaldi, L.A., Hill, R., Settele, J., & Vanbergen, A.J. (2016). Safeguarding pollinators and their values to human well-being. *Nature*, 540(7632), 220–229. <https://doi.org/10.1038/nature20588>

Red Chilena de Polinizaci3n. (s. f.). iNaturalist. Recuperado 11 de octubre de 2022, de <https://www.inaturalist.org/projects/red-chilena-de-polinizacion?tab=stats>

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Shcheglovitova, M. & Anderson, R. P. (2013). Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecological Modelling*, 269, 9-17. <https://doi.org/10.1016/j.ecolmodel.2013.08.011>

Thuiller W, Georges D, Engler R, Breiner F. 2019. biomod2: Ensemble Platform

Warf, B. (2010). *Encyclopedia of Geography* (1.<sup>a</sup> ed.). Sage Publications, Inc.

Warren, D. L. & Seifert, S. N. (2011). Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21(2), 335-342. <https://doi.org/10.1890/10-1171.1>

Witten, I. H. (2005). *Data Mining Practical Machine Learning Tools and Techniques* (2.a ed.). Morgan Kaufmann Pub.

## Anexos

### Anexo 1. Código 1:

```
library(rgbif)
library(scrubr)
library(maps)
lista <- read.csv("ListaPolinizadoresInaturalist.csv")
for (especie in lista) {
  esp <- especie
}
for (especie in lista) {
  Datos <- occ_data(scientificnam = especie ,hasCoordinate = TRUE, limit = 2000,
country = "CL")
  Datos [["data"]]
}
Datos2 = c()
for (v in esp [1:446]) {
  vet <- Datos[[v]][["data"]]
  Datos2[[length(Datos2) +1 ]] <- vet
}
write.csv(Datos2,"C:/Users/marce/Documents/R/DatosPolinizadoresGBIF.csv")
Datos = c()
for (name in esp[1:466]) {
  especies <- get_inat_obs(taxon_name = name , maxresults=2000 , place_id "chile")
  Datos[[length(Datos) +1 ]] <- especies
}
```

```
write.csv(Datos,"C:/Users/marce/Documents/R/DatosPolinizadoresInat.csv")
```

### **Anexo 2. Código 2:**

```
Import pandas as pd
```

```
df_temp=pd.read_excel(r"C:\Users\cami_\Documents\DatosPolinizadoresgbif.xlsx,  
sheet_name=None)
```

```
lista_especies=list(df_temp.keys())
```

```
lista=pd.DataFrame()
```

```
for i in range(356):
```

```
    lista=lista.append(df_temp[lista_especies[i]])
```

```
writer
```

```
=pd.ExcelWriter('demo.xlsx',engine='xlsxwriter',options={'strings_to_urls':False})
```

```
writer.save()
```

### **Anexo 3. Código 3:**

```
Import pandas as pd
```

```
df_temp=pd.read_excel(r"C:\Users\cami_\Documents\DatosPolinizadores.xlsx,  
sheet_name=None)
```

```
lista_especies=list(df_temp.keys())
```

```
lista=pd.DataFrame()
```

```
for i in range(462):
```

```
lista=lista.append(df_temp[lista_especies[i]])
```

```
writer  
=pd.ExcelWriter('consolidadoinna.xlsx',engine='xlsxwriter',options={'strings_to_urls':False})  
writer.save()
```

#### **Anexo 4. Código 4:**

```
#librerías  
library(raster)  
library(sp)  
library(xlsx)  
library(readxl)  
  
# r es los datos climaticos del mundo  
r <- getData("worldclim",var="bio", res=2.5,lon=-72,lat=-35)  
  
#cambio de directorio para leer el archivo  
setwd("C:/Users/Marcelo/Documents/R/")  
  
#se lee el archivo con las coordenadas  
sp.data <- read_excel("consolidado_sin_duplicados.xlsx")  
  
#se revisan los datos  
summary(sp.data)  
  
#se realizan 2 for para tener todas las coordenadas juntas  
for (i in sp.data["lat"]) {  
  lats <- c(i)  
}  
  
for (i in sp.data["lon"]) {
```

```

lons <- c(i)
}

# se les da un formato de data frame
coords <- data.frame(x=lons,y=lats)

#se pasan a puntos espaciales para coincidir con los de r que es el ráster con todos los
datos climáticos
points <- SpatialPoints(coords, proj4string = r@crs)
#se extraen los datos climáticos de los puntos y se guardan en una variable
values <- extract(r,points)

#se guarda esta como un excel
write.xlsx(values,"C:/Users/marce/Documents/Tesis/R/Datospuntos.xlsx")

```

### **Anexo 5. Código 5:**

```

library(sf)
require(ggplot2)
require(stringi)
require(stringr)
require(foreign)
library(dplyr)
require(Rmisc)

setwd("C:/Users/Marcelo/Documents/anaData-20221011T203508Z-
001/anaData/4graficosMaps")

```

```

d1 <- read.csv("dataBruta.csv")

spp <- unique(d1$species)

d2 <- c()
for (i in 1:length(spp)) {

  d2.h <- d1[d1$species==spp[i,]][1,c("species","CLASS","Conservation")]

  d2 <- rbind(d2,d2.h)}

names(d2) <- c("namesSpp","Grupo","estado")

d2$namesSpp <- str_replace_all(d2$namesSpp," ",".")

db <- read.csv("disPotSppAndBurnAreaCellUTM.csv")
db2 <- read.csv("disPotPorbsSppAndBurnAreaCellUTM.csv")
spps <- names(db)[23:ncol(db)]

estOut <- c()
gOut <- c()
sumOut <- c()
espOut <- c()
for (i in 1:length(spps)) {

  spp.h <- spps[i]

```

```

espOut <- c(espOut,spp.h)

est <- d2[d2$namesSpp==spp.h,"estado"]
g.h <- d2[d2$namesSpp==spp.h,"Grupo"]

sum.h <- sum(db[db[,spp.h]==1,"area"])

sumOut <- c(sumOut,sum.h)
gOut <- c(gOut,g.h)
estOut <- c(estOut,est)
}

dat1 <- data.frame(sumOut,gOut,estOut,espOut)

#####
head(dat1)#data habitat quemado por especie
#####
dat1max <- dat1[order(sumOut),]
dat1max <- dat1max[56:66,]

jpeg("Grafico-10-max.jpg", width = 14, height = 10, units = 'in', res = 500)
ggplot(data=dat1max, aes(x=espOut, y=sumOut , fill = espOut)) +
geom_bar(stat="identity", show.legend = FALSE)+
theme_minimal()+
theme(
  axis.text.x = element_text(size = 9,colour = "black")
)

```

```

)+
labs(title = "Grafico de las 10 especies mas afectadas por los incendios")+
ylab("Hábitat afectada por incendios (km²)")+
xlab("Especies")
dev.off()

anhos <- seq(2000,2018,1)

dat.out <- c()
for (u in 1:length(anhos)) {

  anho.h <- anhos[u]

  estOut <- c()
  gOut <- c()
  sumOut <- c()
  espOut <- c()

  for (i in 1:length(spps)) {

    spp.h <- spps[i]

    espOut <- c(espOut,spp.h)

    est <- d2[d2$namesSpp==spp.h,"estado"]
    g.h <- d2[d2$namesSpp==spp.h,"Grupo"]

    sum.h <- sum(db[db[,spp.h]==1,paste("area",anhos[u],sep = "_")])

```

```

sumOut <- c(sumOut,sum.h)
gOut <- c(gOut,g.h)
estOut <- c(estOut,est)

}

dat <- data.frame(sumOut,gOut,estOut,espOut)
dat$anho <- anho.h

dat.out <- rbind(dat.out,dat)

}

#####
head(dat.out) #data habitat quemado por especie y año
#####

```

**Anexo 6. Código 6:**

```

#####
##algunos gráficos
#####

# summarySE provides the standard deviation, standard error of the mean, and a (default
95%) confidence interval

```

```
tgc2 <- summarySE(dat.out, measurevar="sumOut", groupvars=c("estOut", "anho"))
tgc2
```

```
tgc2[tgc2$estOut=="", "estOut"] <- "NoData"
```

```
pd <- position_dodge(0.5) # move them .05 to the left and right
```

```
jpeg("habitatLostAnhoEC.jpg", width = 14, height = 10, units = 'in', res = 500)
ggplot(tgc2, aes(x=anho, y=sumOut, colour=estOut)) +
  geom_errorbar(aes(ymin=sumOut-ci, ymax=sumOut+ci), width=.8, position=pd) +
  geom_line(position=pd) +
  geom_point(position=pd)+
scale_color_viridis_d("E. conservación")+
  theme(
    axis.text.y = element_text(size = 16, colour = "black"),
    axis.title.y = element_text(size = 25, colour = "black"),
    axis.text.x = element_text(size = 16, colour = "black"),
    axis.title.x = element_text(size = 25, colour = "black"),
    plot.title = element_text(size = 11, face = "italic"),
    plot.subtitle = element_text(size = 10, face = "italic"),
    legend.title = element_text(size=20),
    legend.text = element_text(size=20),
    legend.key.height= unit(1, 'cm'),
    legend.key.width= unit(1, 'cm')
  )+
```

```

labs(title = "")+
ylab("Hábitat promedio de las especies incendiado medio (km²)")+
xlab("Año")
dev.off()

# summarySE provides the standard deviation, standard error of the mean, and a
# (default 95%) confidence interval
tgc <- summarySE(dat.out, measurevar="sumOut", groupvars=c("gOut","anho"))
tgc[tgc$gOut=="A","gOut"] <- "Aves"
tgc[tgc$gOut=="M","gOut"] <- "Mamíferos"
tgc

jpeg("habitatLostAnhoGrupo.jpg", width = 14, height = 10, units = 'in', res = 500)
pd <- position_dodge(0.5) # move them .05 to the left and right
ggplot(tgc, aes(x=anho, y=sumOut, colour=gOut)) +
  geom_errorbar(aes(ymin=sumOut-ci, ymax=sumOut+ci), width=.5, position=pd) +
  geom_line(position=pd) +
  geom_point(position=pd)+
  scale_color_manual("Grupo",values = c("tomato3","dodgerblue3"))+
  theme(
    axis.text.y = element_text(size = 16,colour = "black"),
    axis.title.y = element_text(size = 25,colour = "black"),
    axis.text.x = element_text(size = 16,colour = "black"),
    axis.title.x = element_text(size = 25,colour = "black"),
    plot.title = element_text(size = 11,face = "italic"),

```

```

plot.subtitle = element_text(size = 10,face = "italic"),
legend.title = element_text(size=20),
legend.text = element_text(size=20),
legend.key.height= unit(1, 'cm'),
legend.key.width= unit(1, 'cm')

)+
labs(title = "")+
ylab("Hábitat incendiado medio (km²)")+
xlab("Año")
dev.off()

dat1[dat1$gOut=="A","gOut"] <- "Aves"
dat1[dat1$gOut=="M","gOut"] <- "Mamíferos"

jpeg("boxplotHabLostGrupo.jpg", width = 10, height = 10, units = 'in', res = 500)
ggplot(dat1, aes(x=gOut, y=sumOut))+
  geom_boxplot()+
  theme(
    axis.text.y = element_text(size = 16,colour = "black"),
    axis.title.y = element_text(size = 20,colour = "black"),
    axis.text.x = element_text(size = 16,colour = "black"),
    axis.title.x = element_text(size = 20,colour = "black"),
    plot.title = element_text(size = 11,face = "italic"),
    plot.subtitle = element_text(size = 10,face = "italic"),
    legend.title = element_text(size=20),
    legend.text = element_text(size=20),

```

```

    legend.key.height= unit(1, 'cm'),
    legend.key.width= unit(1, 'cm')

)+
labs(title = "")+
ylab("Hábitat incendiado (km²)")+
xlab("Grupo")
dev.off()

jpeg("boxplotHabLostEC.jpg", width = 10, height = 10, units = 'in', res = 500)
ggplot(dat1, aes(x=estOut, y=sumOut))+
  geom_boxplot()+
  theme(
    axis.text.y = element_text(size = 16,colour = "black"),
    axis.title.y = element_text(size = 20,colour = "black"),
    axis.text.x = element_text(size = 16,colour = "black"),
    axis.title.x = element_text(size = 20,colour = "black"),
    plot.title = element_text(size = 11,face = "italic"),
    plot.subtitle = element_text(size = 10,face = "italic"),
    legend.title = element_text(size=20),
    legend.text = element_text(size=20),
    legend.key.height= unit(1, 'cm'),
    legend.key.width= unit(1, 'cm')

)+
labs(title = "")+
ylab("Hábitat incendiado (km²)")+

```

```
xlab("Estado de conservación")
dev.off()
```

### **Anexo 7. Código 7:**

```
#####
#Algunos mapas
#####

data <- st_read("grilla_final.shp")

divIn <- rowSums(
db[,names(db)[names(db)%in%d2[d2$Grupo=="Insecta","namesSpp"]]]##Diversidad
potencial de insectos

divAv <- rowSums(
  db[,names(db)[names(db)%in%d2[d2$Grupo=="Aves","namesSpp"]]]##Diversidad
potencial de aves

divPo <- rowSums(db[23:89])

dAvIn <- db[,c("arclm_d", "ID2")]
dAvIn2 <- db2[,c("arclm_d", "ID2")]

dAvIn$divAv <- divAv
```

```
dAvIn$divIn <- divIn
```

```
dAvIn$divPo <- divPo
```

```
dbF <- merge(data,dAvIn,by="arclm_d")
```

```
dbF <- dbF[order(dbF$ID2.x),]
```

```
nrow(dAvIn)
```

```
dbF$divHA <- 0
```

```
dbF$divHA <- ifelse(dbF$area != 0,dbF$divPo,0 )
```

```
dbF$divHP <- 0
```

```
dbF$divHP <- ifelse(dbF$area > 12500000 ,dbF$divPo,0)
```

```
jpeg("mapaInsec.jpg", width = 10, height = 12, units = 'in', res = 500)
```

```
ggplot(dbF)+
```

```
  geom_sf(aes(fill=(divIn)),color=NA)+
```

```
  #scale_fill_viridis_c("Riqueza de especies")+
```

```
  scale_fill_gradientn("Riqueza de especies",colours = c("white","green","red"))+
```

```
  # theme(legend.position="none")+
```

```
  theme(
```

```
    axis.text.y = element_text(size = 12,colour = "black"),
```

```
    axis.title.y = element_text(size = 7,colour = "black"),
```

```
    axis.text.x = element_text(size = 12,colour = "black"),
```

```

axis.title.x = element_text(size = 7,colour = "black"),
legend.title = element_text(size=15),
plot.title = element_text(size = 18),
plot.subtitle = element_text(size = 10),
legend.text = element_text(size=12))+
labs(title = "Diversidad potencial de Insectos")
dev.off()

```

```

jpeg("mapaAves.jpg", width = 10, height = 12, units = 'in', res = 500)
ggplot(dbF)+
  geom_sf(aes(fill=(divAv)),color=NA)+
  #scale_fill_viridis_c("Riqueza de especies")+
  scale_fill_gradientn("Riqueza de especies",colours = c("white","green","red"))+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),
    legend.title = element_text(size=15),
    plot.title = element_text(size = 18),
    plot.subtitle = element_text(size = 10),
    legend.text = element_text(size=12))+
  labs(title = "Diversidad potencial de Aves")
dev.off()

```

```

jpeg("mapapolinizadores.jpg", width = 10, height = 12, units = 'in', res = 500)
ggplot(dbF)+
  geom_sf(aes(fill=(divPo)),color=NA)+
  #scale_fill_viridis_c("Riqueza de especies")+
  scale_fill_gradientn("Riqueza de especies",colours = c("white","green","red"))+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),
    legend.title = element_text(size=15),
    plot.title = element_text(size = 18),
    plot.subtitle = element_text(size = 10),
    legend.text = element_text(size=12))+
  labs(title = "Diversidad potencial de Polinizadores")
dev.off()

jpeg("mapahabitadadeespeciesefectadasporincendios.jpg", width = 10, height = 12, units
= 'in', res = 500)
ggplot(dbF)+
  geom_sf(aes(fill=(divHA)),color=NA)+
  scale_fill_viridis_c("Numero de polinizadores afectados por los incendios",option =
"A")+
  # theme(legend.position="none")+
  theme(

```

```

axis.text.y = element_text(size = 12,colour = "black"),
axis.title.y = element_text(size = 7,colour = "black"),
axis.text.x = element_text(size = 12,colour = "black"),
axis.title.x = element_text(size = 7,colour = "black"),
legend.title = element_text(size=15),
plot.title = element_text(size = 18),
plot.subtitle = element_text(size = 10),
legend.text = element_text(size=12))+
labs(title = "Efectos de los incendios sobre habitad de polinizadores")
dev.off()

jpeg("mapahabitadadeespeciesefectadasgravementeporincendios.jpg", width = 10, height
= 12, units = 'in', res = 500)

ggplot(dbF)+
  geom_sf(aes(fill=(divHP)),color=NA)+
  scale_fill_viridis_c("Número de polinizadores afectados por los incendios",option =
"A")+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),
    legend.title = element_text(size=15),
    plot.title = element_text(size = 18),
    plot.subtitle = element_text(size = 10),
    legend.text = element_text(size=12))+
  labs(title = "Efectos graves de los incendios sobre hábitat de polinizadores")

```

```

dev.off()

jpeg("mapa de incendios.jpg", width = 10, height = 12, units = 'in', res = 500)
ggplot(dbF)+
  geom_sf(aes(fill=area/1000000),color=NA)+
  scale_fill_viridis_c("Area quemada (km² por celda UTM)",option = "A")+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),
    legend.title = element_text(size=15),
    plot.title = element_text(size = 18),
    plot.subtitle = element_text(size = 10),
    legend.text = element_text(size=12))+
  labs(title = "Area afectada por los incendios")
dev.off()

db.h <- merge(data[,c(1,2)],db,by="arclm_d")

db.h <- db.h[order(db.h$ID2.x),]

db2.h <- merge(data[,c(1,2)],db2,by="arclm_d")

db2.h <- db2.h[order(db2.h$ID2.x),]

```

```

jpeg("mapaBTerrestris.jpg", width = 10, height = 12, units = 'in', res = 500)
ggplot(db.h)+
  geom_sf(aes(fill=(as.factor(Bombus.terrestris))),color=NA)+
  scale_fill_viridis_d("Ausencia/Presencia")+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),
    legend.title = element_text(size=15),
    plot.title = element_text(size = 18),
    plot.subtitle = element_text(size = 10),
    legend.text = element_text(size=12))+
  labs(title = "Dist Pot. Bombus terrestris")
dev.off()

setwd("C:/Users/Marcelo/Documents/anaData-20221011T203508Z-
001/anaData/4graficosMaps/mapas/prob")
for (i in spp) {
  tryCatch({
    jpeg(paste(i,".jpg",sep = ""), width = 10, height = 12, units = 'in', res = 500)
    print(ggplot(db2.h)+
      geom_sf(aes(fill = get(i)),color=NA)+
      #scale_fill_gradient("Porcentaje",colours = c("white","gray","black"))+
      scale_fill_viridis_c("Porcentaje")+
      # theme(legend.position="none")+

```

```

theme(
  axis.text.y = element_text(size = 12,colour = "black"),
  axis.title.y = element_text(size = 7,colour = "black"),
  axis.text.x = element_text(size = 12,colour = "black"),
  axis.title.x = element_text(size = 7,colour = "black"),
  legend.title = element_text(size=15),
  plot.title = element_text(size = 18),
  plot.subtitle = element_text(size = 10),
  legend.text = element_text(size=12))+
labs(title = paste("Dist Pot. % ",i,sep = "")))
dev.off()
Sys.sleep(3)
}, error=function(e){cat("ERROR :",conditionMessage(e), "\n")})
}

setwd("C:/Users/Marcelo/Documents/anaData-20221011T203508Z-
001/anaData/4graficosMaps/mapas")
for (i in spp) {
  tryCatch({
    jpeg(paste(i,".jpg",sep = ""), width = 10, height = 12, units = 'in', res = 500)
    print(ggplot(db.h)+
      geom_sf(aes(fill=(as.factor(get(i))),color=NA))+
      scale_fill_viridis_d("Ausencia/Presencia")+
      # theme(legend.position="none")+
      theme(
        axis.text.y = element_text(size = 12,colour = "black"),
        axis.title.y = element_text(size = 7,colour = "black"),
        axis.text.x = element_text(size = 12,colour = "black"),

```

```

axis.title.x = element_text(size = 7,colour = "black"),
legend.title = element_text(size=15),
plot.title = element_text(size = 18),
plot.subtitle = element_text(size = 10),
legend.text = element_text(size=12))+
labs(title = paste("Dist Pot.",i,sep = "")))
dev.off()
Sys.sleep(3)
}, error=function(e){cat("ERROR :",conditionMessage(e), "\n")})
}

```

```
db.h$areaBombus.terrestris <- 0
```

```
db.h[db.h$Bombus.terrestris==1,"areaBombus.terrestris"] <-
db.h[db.h$Bombus.terrestris==1,"area"]
```

```

jpeg("mapaBTerrestris2.jpg", width = 10, height = 12, units = 'in', res = 500)
ggplot(db.h)+
  geom_sf(aes(fill=(areaBombus.terrestris),color=NA))+
  scale_fill_viridis_c("Habitat quemado (km2 por celda UTM)",option = "A")+
  # theme(legend.position="none")+
  theme(
    axis.text.y = element_text(size = 12,colour = "black"),
    axis.title.y = element_text(size = 7,colour = "black"),
    axis.text.x = element_text(size = 12,colour = "black"),
    axis.title.x = element_text(size = 7,colour = "black"),

```

```

legend.title = element_text(size=15),
plot.title = element_text(size = 18),
plot.subtitle = element_text(size = 10),
legend.text = element_text(size=12))+
labs(title = "Habitat perdido Bombus terrestris")
dev.off()

consta=24
for (i in spp) {
  tryCatch({

    cons <- paste("area",i,sep = "")

    db.h[paste(cons)] <- 0

    db.h[unlist(db.h[,consta])==1 ,paste("area",i,sep = "")] <-
db.h[unlist(db.h[,consta])==1 ,"area"]

    consta <- consta + 1

    print(consta)

  }, error=function(e){cat("ERROR :",conditionMessage(e), "\n")})
}

setwd("C:/Users/Marcelo/Documents/anaData-20221011T203508Z-
001/anaData/4graficosMaps/mapas/incendios")

for (i in spp) {
  tryCatch({

    cons <- paste("area",i,sep = "")

    jpeg(paste(i,".jpg",sep = ""), width = 10, height = 12, units = 'in', res = 500)

    print(ggplot(db.h)+

      geom_sf(aes(fill=(get(cons))),color=NA)+

      scale_fill_viridis_c("Habitat quemado (km² por celda UTM)",option = "A")+

```

```

# theme(legend.position="none")+
theme(
  axis.text.y = element_text(size = 12,colour = "black"),
  axis.title.y = element_text(size = 7,colour = "black"),
  axis.text.x = element_text(size = 12,colour = "black"),
  axis.title.x = element_text(size = 7,colour = "black"),
  legend.title = element_text(size=15),
  plot.title = element_text(size = 18),
  plot.subtitle = element_text(size = 10),
  legend.text = element_text(size=12))+
  labs(title = paste("Habitat perdido ",i,sep = "")))
dev.off()
Sys.sleep(3)
}, error=function(e){cat("ERROR :",conditionMessage(e), "\n")})
}

```