

Modelo de distribución de *Leontochir ovallei* utilizando variables ambientales bioclimáticas, topográficas y de percepción remota.

Sergio Christian Payacan Castillo¹

¹ Recursos Naturales GHD S.A. Chile
Apoquindo 4775, Las Condes, Santiago - Chile
sergio.payacan@gmail.com

Abstract. Potential distribution of the species *Leontochir ovallei* is determined using spatial analysis tools of maximum entropy (Maxent) with a small number of existing records of its presence. Maxent is based on a machine learning response that is designed to make predictions based on incomplete data. This approach estimates the most uniform distribution (maximum entropy) of sampling points compared to background locations given the constraints derived from the data. Thirteen environmental variables were considered as potential predictors of the *L. ovallei* habitat distribution. The environmental variables used were bioclimatic, topographic and from remote sensing sources such as MODIS and Landsat. The quality of model predictions was evaluated with area under the receiver operating characteristic curve (AUC) and the jackknife test was used to assess the importance of predictor variables. The logistic habitat suitability map was generally consistent with the registers of areas sampled, local expert knowledge and experience. Cloud cover, annual precipitation and altitude were the most influential environmental variables on model generation, demonstrating the use of variables from remote sensing for predicting habitat suitability for rare and geographically restricted species. This work contributes to a more reliable estimate of the species distribution intended to support planning and management programs of conservation.

Key words: Maxent, remote sensing, Landsat, MODIS, biodiversity conservation, small sample size.

1. Introducción

Leontochir ovallei es una especie endémica de Chile, de género monotípico, de distribución geográfica reducida pero frecuente, que crece sólo en la región de Atacama, la cual es una hierba perenne que posee tallos gruesos, tendidos, que alcanzan hasta 2 m de longitud, que en sus extremos emerge una inflorescencia de tipo umbela compuesta de una gran número de flores rojas muy apreciada por su valor ornamental (Muñoz, 1973). El Decreto Supremo N° 50 de Chile (MINSEGPRES, 2008) la califica en Peligro debido a su distribución muy localizada y Rara porque se conoce su presencia en 2 o 3 localidades y solo florece en algunos momentos de abundancia relativa asociado al fenómeno del desierto florido. Los estudios sobre la biogeografía de la especie han sido principalmente narrativos y estimados a partir de la vegetación asociada (MINSEGPRES, 2008).

La predicción y cartografía de aptitud de hábitat para especies de plantas raras y en peligro de extinción es una herramienta útil para el monitoreo y restauración de poblaciones nativas en declinación, planificación y acciones destinadas a programas de conservación. La investigación de especies de plantas raras es parte de la necesidad de comprender y manejar la biodiversidad en un lugar determinado (Adam, 2002). En este sentido la predicción de los modelos de distribución de aptitud de hábitat es una poderosa herramienta para resolver preguntas en biología de la conservación y ecología aplicada (Guisan y Thuiller, 2005).

Los modelos de distribución de especies generados por el algoritmo de máxima entropía con el software Maxent (Phillips et al., 2004) han presentado un alto rendimiento con pequeños tamaños de muestras de presencia de especies (Hernandez, et al., 2006; Wisz, et al., 2008; Kumar y Stohlgren, 2009; Baldwin, 2009). Maxent es un algoritmo de aprendizaje de máquinas que usa distribuciones multivariadas de aptitud de hábitat inferida de registros de presencia para generar una superficie de probabilidad de presencia de la especie tan cercana a una distribución uniforme como sea posible, dada las limitaciones impuestas por los valores esperados de las características ambientales de cada ubicación donde la especie ha sido registrada.

Las variables ambientales comúnmente utilizadas en los modelos de aptitud de hábitat son, de acuerdo a Mateo et al. (2011) provenientes de datos climáticos, como la base de datos Worldclim desarrollada por Hijmans et al. (2005), información sobre edafología, litología y geología, elevación y variables derivadas como la exposición, pendiente y red de drenaje y variables obtenidas por datos de teledetección, las cuales han sido relativamente poco utilizadas pero de alto potencial. Al respecto, autores como Zimerman et al. (2007), González et al. (2013), Feilhauer et al. (2012), Kumar y Stohlgren (2009) y Kafley (2008) destacan la utilidad de los datos provenientes de percepción remota como variables predictoras de modelos de hábitat usando Maxent. Austin (2007) planteó que las respuestas de las especies dependerá de la naturaleza de las variables predictoras ambientales y los procesos ecológicos asociados, ante lo cual hay avances en el conocimiento acerca de los procesos ecofisiológicos y biofísicos que rigen las relaciones entre las especies y su entorno, saber que puede ser utilizado para seleccionar potenciales variables que describan la distribución espacial de las especies, es por ello que el presente trabajo integra variables climáticas (algunas raramente utilizadas en los modelos de distribución de especies, como la nubosidad), físicas y biológicas, que podrían tener vínculos ecofisiológicos con el potencial de distribución de *Leontochir ovallei*.

2. Metodología de Trabajo

Los rangos geográficos de distribución de la especie *Leontochir ovallei* fueron estimados usando máxima entropía (Maxent). La preparación de los archivos de registro de presencia de la especie, raster y cartografía final fueron procesados con el software QGIS versión 2.6 (QGIS, 2014). Se trabajó en un sistema de coordenadas geográficas datum WGS-84.

Fuentes de información. Se construyó una base de datos para los registros de *Leontochir ovallei*, mediante revisión bibliográfica, consultas en museos y referencias personales. La base de datos fue filtrada de acuerdo a la precisión de la fuente consultada con un total de 10 registros seleccionados para ingresar al modelo. Las coordenadas geográficas de los registros fueron convertidos a un archivo CVS. Las imágenes Landsat utilizadas se obtuvieron de la página Web del Servicio Geológico de los Estados Unidos (<http://earthexplorer.usgs.gov/>). El modelo digital de elevación se obtuvo de la página del Sistema Espacial Japonés (<http://gdem.ersdac.jpacesystems.or.jp/>). El producto cloud fraction de MODIS se obtuvo de la página de observaciones de la Tierra de la NASA (<http://neo.sci.gsfc.nasa.gov/>).

Selección de variables predictoras. Las variables predictoras fueron seleccionada a priori basado en el conocimiento de *Leontochir ovallei* en relación a aspectos climáticos, topográficos, vegetacionales e hidrografía. Inicialmente fueron usadas las 19 variables bioclimáticas provistas por la base de datos de WorldClim (Hijmans et al., 2005). Para remover la alta correlación entre las variables bioclimáticas se procedió a descartar aquellas con una correlación mayor a 0.7 (coeficiente de Pearson) y aquellas que no presentaban variabilidad en el área de estudio, seleccionando solo 5 de ellas: Temperatura media anual (Bio1), rango medio de temperatura diurna (Bio2), estacionalidad de la temperatura (Bio4), precipitación anual (Bio 12) y estacionalidad de la precipitación (Bio15). Se calcularon los promedios mensuales y anuales de fracción de nubosidad del área de estudio entre los años 2002 a 2013 provenientes de datos del producto Cloud Fraction derivado del producto MYD06_L2 de la plataforma Aqua del satélite MODIS, de los cuales se seleccionó el promedio de fracción de nubosidad del mes de septiembre por presentar una mayor contribución estadística al modelo en pruebas preliminares con Maxent. Las variables topográficas seleccionadas fueron la altitud, exposición y pendiente, derivadas del modelo digital de elevación GDEM (del sensor ASTER a bordo del satélite Terra). La distancia a la red hidrográfica fue calculada a partir de una capa vectorial de la hidrografía del área y transformada a una raster de distancia mediante el algoritmo de distancia euclidiana. Las

variables físicas relativas al suelo fueron derivadas de las escenas 001/079 y 001/080 del satélite Landsat 8 con fecha de adquisición el 9 de junio de 2014. Del mosaico de las escenas se calcularon las razones de bandas b4/b5, b4/b2 y b6/b7, de las cuales se seleccionaron las dos últimas por presentar menor correlación entre sí (coeficiente de Pearson < 0.7). Como variable biótica se utilizó el índice de vegetación de diferencia normalizada (NDVI) calculado de la relación entre las bandas correspondientes al rojo e infrarrojo próximo, proveniente de las escenas 001/079 y 001/080 del satélite landsat 5 TM con fecha de adquisición el 30 de septiembre de 1991, esta fecha fue seleccionada por la ocurrencia del fenómeno de desierto florido en el área de estudio, fenómeno asociado al evento meteorológico El Niño.

Procesamiento de las variables predictoras. Todas las variables fueron convertidas a formato GRID (raster), cortadas a un rombo de superficie de 20.465 Km² y remuestreadas a un tamaño de pixel de 90 m usando un interpolador bilinear. Posteriormente fueron convertidas a formato ASCII el cual es uno de los formatos que requiere el software Maxent.

La Tabla 1 presenta las variables utilizadas en el modelo de distribución para *Leontochir ovallei*.

Tabla 1 Variables ambientales

Variable predictora	Fuente	Tipo de dato	Unidad
Temperatura media anual (Bio1)	Worldclim	Contínuo	°C
Rango medio de temperatura diurna (Bio2)	Worldclim	Contínuo	°C
Estacionalidad de la temperatura (Bio4)	Worldclim	Contínuo	°C
Precipitación anual (Bio12)	Worldclim	Contínuo	mm
Estacionalidad de la precipitación (Bio15)	Worldclim	Contínuo	mm
Nubosidad promedio de septiembre (nubosidad_sep)	MODIS (MYD06)	Contínuo	0 a 1
Modelo digital de elevación (altitud)	ASTER GDEM	Contínuo	m
Exposición	Derivado de GDEM	Catagórico	4 categorías
Pendiente	Derivado de GDEM	Contínuo	%
Distancia al cauce	SIIT-Chile	Contínuo	m
b4/b2	Landsat 8	Catagórico	5 categorías
b6/b7	Landsat 8	Catagórico	5 categorías
NDVI	Landsat TM 5	Contínuo	-1 a +1

Modelo estadístico, validación y contribución de variables. Los rangos de distribución geográfica para la especie en estudio fueron estimados usando el software Maxent, mediante los registros de presencia de la especie *Leontochir ovallei* y las variables ambientales seleccionadas. La medida de ajuste implementada por Maxent es el área bajo la curva (AUC), tal ajuste compara la sensibilidad y la especificidad del modelo, tal que los modelos con un puntaje mayor a 0.5 son considerados mejores que una respuesta debida al azar. Maxent también estima la contribución relativa de todas las variables ambientales en el modelo, en tanto una prueba complementaria de jackknife se utilizó para analizar la importancia de cada variable en los resultados del modelo.

3. Resultados y discusión

El modelo presentado tuvo un alto desempeño con un valor de AUC de 0.985. La Figura 1 muestra el resultado del modelo de distribución de la especie *Leontochir ovallei* según la probabilidad de aptitud de hábitat, el cual presenta una capa temática de los pixeles con probabilidad mayores a 0.5 (probabilidades menores a 0.5 se aplicó una transparencia de 100%). Las áreas con mayor probabilidad de aptitud de hábitat corresponde a sectores costeros en un transecto longitudinal de 170 km aproximadamente (entre los valles de Vallenar y Copiapó), con áreas más fragmentados y de menor probabilidad hacia el interior de la costa.

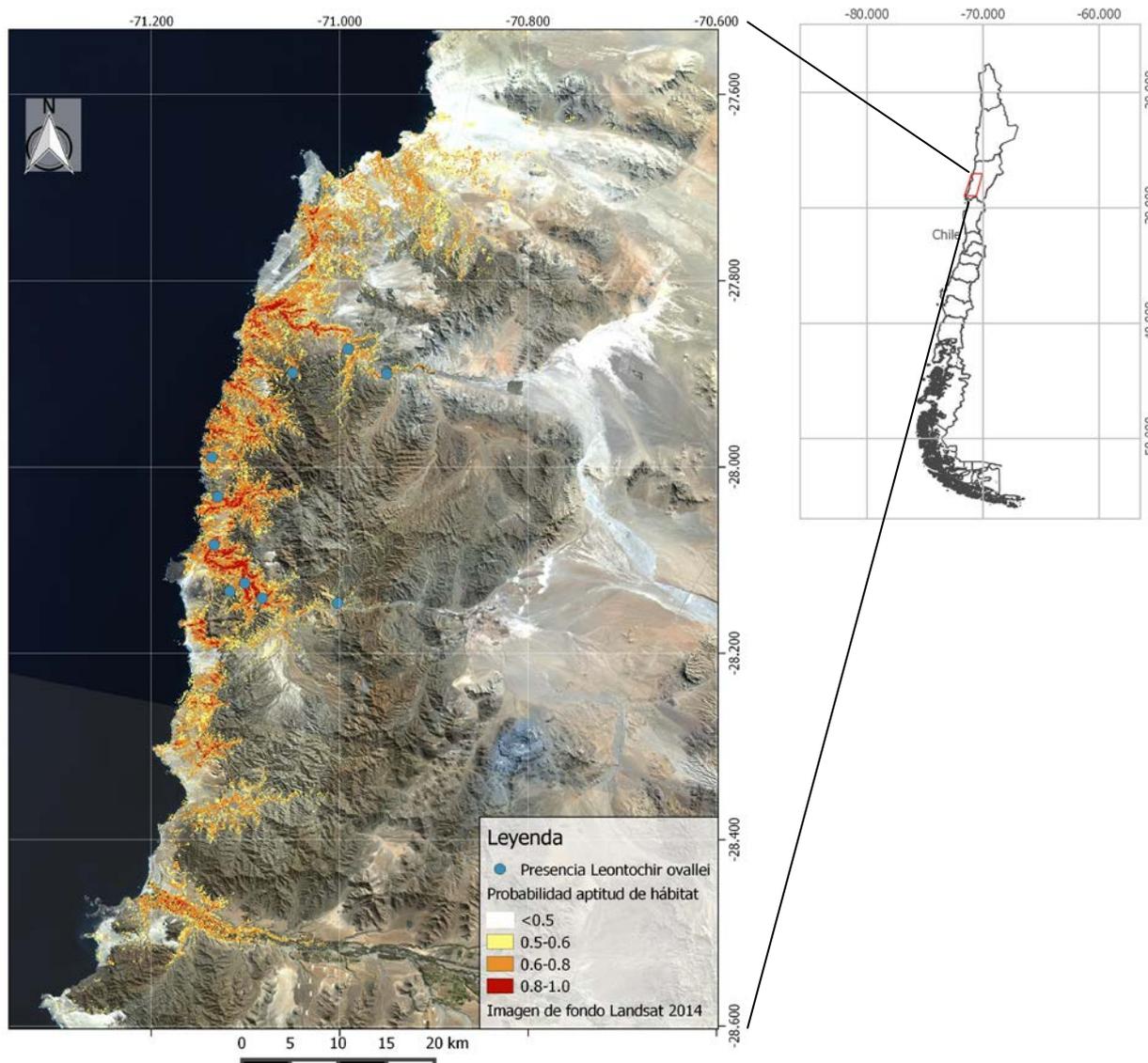


Figura 1 Cartografía de probabilidad de aptitud de hábitat de distribución para *Leontochir ovallei*

Entre las variables que presentaron una mayor contribución al desempeño del modelo están el promedio de nubosidad del mes de septiembre (49.3%), la precipitación anual Bio12 (15.6%), la altitud (12.8 %) y la distancia al cauce hidrográfico (12.4%). Finalmente las otras variables presentaron contribuciones al modelo menores al 10%.

Tabla 2 Estimación de la contribución relativa de las variables ambientales en el modelo generado por Maxent

Variable	Porcentaje de contribución	Importancia de permutación
nubosidad_sep	49.3	8.8
bio12	15.6	4.7
altitud	12.8	55.6
dis_cauce	12.4	24.4
Exposición	4.6	0.3
pendiente	1.9	1.4

Variable	Porcentaje de contribución	Importancia de permutación
ndvi	1.4	0.7
bio4	1.1	0
bio15	0.5	0.6
bio2	0.4	2.2
bio1	0	0
Landsat(b6_b7)	0	1.1
Landsat(b4_b2)	0	0.2

Nubosidad_sep: nubosidad promedio de septiembre; bio12:precipitación anual; dis_cauce: distancia al cauce hidrográfico mas cercano; bio4:estacionalidad de la temperatura; bio15:estacionalidad de la precipitación; bio2:rango medio de la temperatura diurna; bio1:temperatura promedio anual; Landsat(b6_b7): Razón de bandas b6/b7 (Landsat 8); Landsat(b4_b2): Razón de bandas b4/b2 (Landsat 8)

El test de Jackknife entregado por Maxent muestra la contribución de las variables expresado en términos de ganancia estadística, en la cual se determinó que la variable con mayor significancia estadística corresponde a la altitud, la nubosidad del mes de septiembre y la temperatura media anual, que contiene mayor información útil para el modelo cuando son consideradas individualmente (Figura 2)

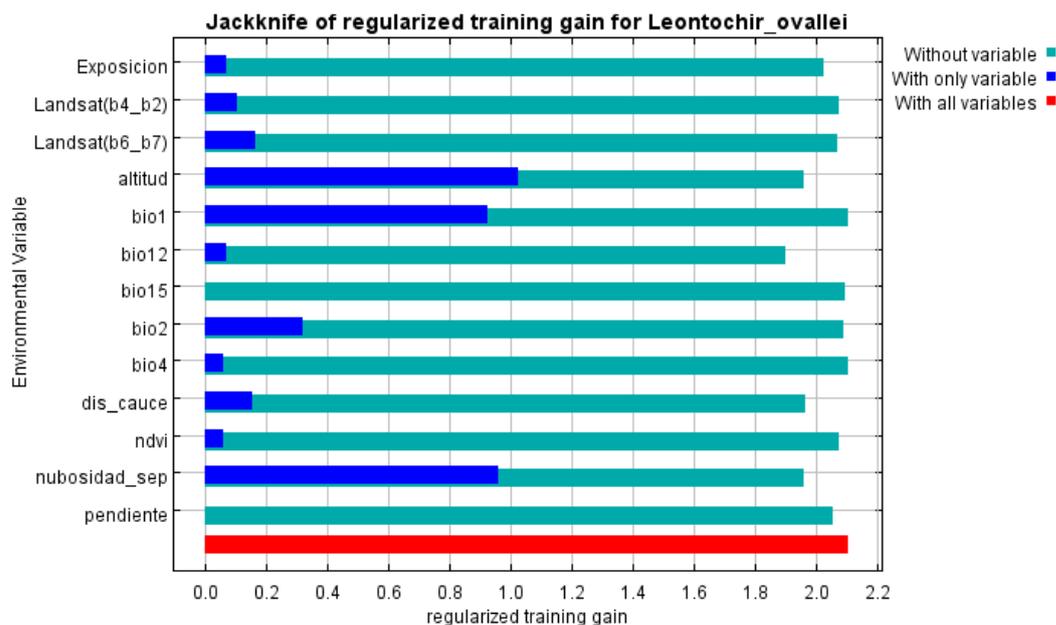


Figura 2 Prueba de jackknife para medir la importancia de las variables

La Figura 3 muestra las curvas de variación de las variables más importantes en la configuración del modelo de distribución, en la cual se puede apreciar la respuesta positiva de la presencia de *Leontochir ovallei* al incremento de variables como la nubosidad, la pendiente, el NDVI y el rango medio de temperatura diurna (Bio2); en contraste con la respuesta negativa a los incrementos de las variables precipitación anual (Bio12), altitud y distancia al cauce hidrográfico. La variable categórica exposición muestra la tendencia de la especie de estar ausente en laderas de exposición norte.

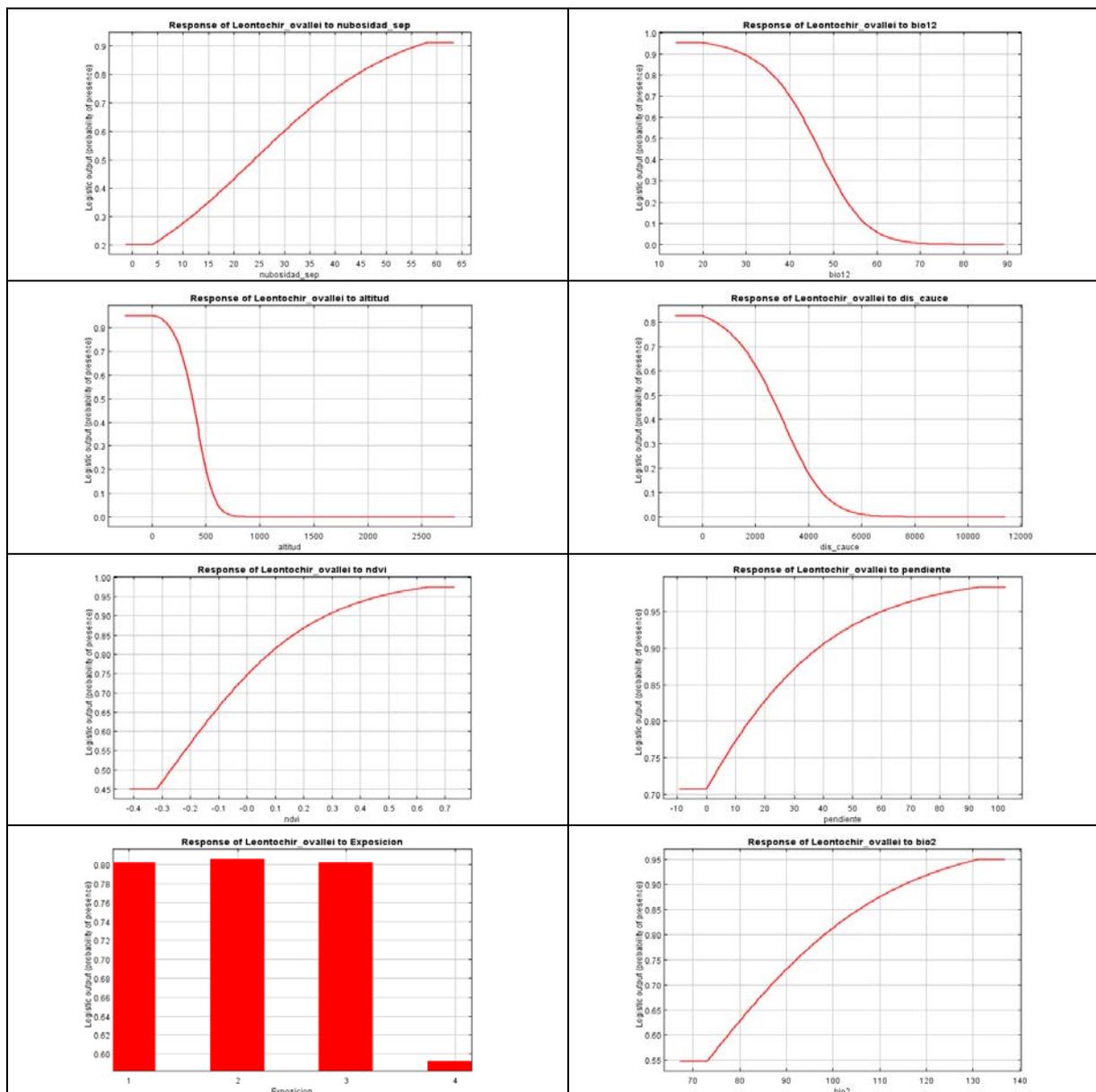


Figura 3 Curvas de respuesta de variables ambientales a la distribución de *Leontochir ovallei*

4. Conclusiones

El presente estudio muestra que los patrones de distribución de hábitat para *Leontochir ovallei* puede ser modelado eficientemente utilizando un pequeño número de registros de presencia y mediante una adecuada selección de variables ambientales por medio del software Maxent. Este trabajo proporciona predicciones de probabilidad de aptitud de hábitat para esta especie, mostrando y confirmando la reducida superficie en que se distribuye actualmente. Entre las variables ambientales incorporadas al modelo destacan positivamente las fuentes de datos de percepción remota de alto potencial de utilidad, como las derivadas de datos proporcionados por los satélites MODIS y Landsat que permitieron integrar variables de carácter abiótico (nubosidad, altitud y razón de bandas como índices de tipo litológico) y biótico (como NDVI en época de manifestación del fenómeno desierto florido). Particularmente influyente es la participación de la variable nubosidad del mes de septiembre en la construcción del modelo, la correlación positiva entre la presencia de la especie y la manifestación de nubosidad podría sugerir hipótesis de una dependencia fisiológica de la planta a condiciones de mayor humedad ambiental durante la etapa de crecimiento vegetativo,

que proporcionarían la nubosidad costera o fenómeno de la camanchaca característico del borde costero del norte de Chile. La participación de variables proporcionadas por la razón de bandas b_4/b_2 y b_6/b_7 utilizadas usualmente para discriminar tipos de rocas, muestra la contribución de la composición del suelo en la distribución de la especie (cuando son analizadas individualmente), lo cual concuerda con observaciones de campo que asocian la presencia de la especie a sectores rocosos. La influencia de las variables bioclimáticas como la precipitación anual y la estacionalidad de la precipitación explicarían en parte el desarrollo de la planta en un ambiente xérico y de características templadas, la cual no responde a ciclos anuales, sino, a ciclos interanuales influidos por el fenómeno climático El Niño al cual se asocia la expresión del desierto florido.

La aproximación del modelo al nicho fundamental de la especie utilizando variables ambientales puede generar una sobrepredicción en algunas áreas, sin embargo, la información producida por este estudio puede ser de gran utilidad como apoyo a programas de planificación y manejo para esta especie en peligro de extinción, descubrimiento de nuevas poblaciones en futuros muestreos, identificación de sitios prioritarios para la conservación y priorizar áreas para generar repoblamiento de la especie o restarurar hábitat naturales que posibilite la conservación efectiva de *Leontochir ovallei*.

Agradecimientos

Agradezco al Ministerio de Medio Ambiente por proporcionar los datos de presencia de la especie *Leontochir ovallei* y al revisor Sebastian Rosa por mejorar sustancialmente el escrito y proporcionar ideas fundamentales en el desarrollo de este trabajo. Al equipo de Recursos Naturales de GHD por su apoyo técnico y a Ximena Toro por su colaboración en la traducción.

Referencias Bibliográficas

- Adam, P. Rarity, rare plant species and the New South Wales Threatened Species Conservation Act – conservation opportunities and challenges. *Cunninghamia* 7, 651–669, 2000.
- Baldwin, R.A. "Use of Maximum Entropy Modeling in Wildlife Research." *Entropy* 11, no. 4: 854-866, 2009.
- González B.A., Samaniego H, Marín J.C., Estades C.F. Unveiling Current Guanaco Distribution in Chile Based upon Niche Structure of Phylogeographic Lineages: Andean Puna to Subpolar Forests. *PLoS ONE* 8(11): e78894. doi:10.1371/journal.pone.0078894, 2013
- Guisan, A. and Thuiller, W. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8: 993–1009. doi: 10.1111/j.1461-0248.2005.00792.x, 2005.
- Hernandez, P.A., Graham, C.H., Master, L.L. and Albert, D.L. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, 29: 773–785. doi: 10.1111/j.0906-7590.2006.04700.x, 2006.
- Hijmans R.J., Cameron S.E., Parra J.L., Jones, P.G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978, 2005.
- Kafley, H. Habitat Evaluation and Suitability Modeling of *Rhinoceros unicornis* in Chitwan National Park, Nepal: A Geospatial Approach. Institute of International Education World Wildlife Fund ITTO. p. 53, 2008.

Kumar, S. and Stohlgrenw, T.J. Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. *Journal of Ecology and The Natural Environment*. 1(4): 94-98, 2009.

Mateo, R.G., Felicísimo, A.M., & Muñoz, J. Modelos de distribución de especies: Una revisión sintética. *Revista chilena de historia natural*, 84(2), 217-240, 2011.

MINISEGPRES (Ministerio Secretaría General de la Presidencia, CL). Decreto Supremo N° 50/2008. Chile. Aprueba y oficializa nómina para el segundo proceso de clasificación de especies según su estado de conservación. Santiago, Chile, 2008.

Muñoz, C. Chile: Plantas en extinción. Edit. Univ., Santiago, Chile. 248 p., 1973

QGIS Development Team. QGIS Geographic Information System, Version 2.6 Brighton. Open Source Geospatial Foundation Project. <http://www.qgis.org>, 2014.

Phillips, S.J., Dudík, M., Schapire, R.E. A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-First International Conference on Machine Learning*, pages 655-662, 2004.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A. and NCEAS Predicting Species Distributions Working Group. Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, 14: 763–773. doi: 10.1111/j.1472-4642.2008.00482.x, 2008.

Zimmermann N.E., Moisen G.G., Frescino, T.S. & Blackard, J.A. Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology* 44: 1057-1067, 2007.