

UNIVERSIDAD SAN FRANCISCO DE QUITO

Colegio de Posgrados

Modelamiento con el apoyo de Sistemas de Información Geográfico de los efectos de cambio climático sobre parientes silvestres de cultivos prioritarios para la alimentación y la agricultura: Estudio de caso frijol común

Fredy Alexander Monserrate Rojas

Richard Resl, Ph.Dc., Director de Tesis

Tesis de grado presentada como requisito para la obtención del título de
Magister en Sistemas de Información Geográfica

Quito, Diciembre de 2012

**Universidad San Francisco de Quito
Colegio de Postgrados**

HOJA DE APROBACIÓN DE TESIS

Modelamiento con el apoyo de Sistemas de Información Geográfico de los efectos de cambio climático sobre parientes silvestres de cultivos prioritarios para la alimentación y la agricultura: Estudio de caso frijol común

Fredy Alexander Monserrate Rojas

Richard Resl. Ph.Dc.,
Director de Tesis

Pablo Cabrera Msc.,
Miembro del Comité de Tesis

Richard Resl. Ph.Dc.,
Director del Programa de Maestría en
Sistemas de Información Geográfica

Stella de la Torre, Ph.D.,
Decana del Colegio de
Ciencias Biológicas y Ambientales

Victor Viteri Breedy, Ph.D.,
Decano del Colegio de Postgrados

Quito, Diciembre de 2012

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído la Política de Propiedad Intelectual de la Universidad San Francisco de Quito y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo de investigación quedan sujetos a lo dispuesto en la Política.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo de investigación en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Firma:

Nombre: Fredy Alexander Monserrate Rojas

Pasaporte: 80217436

Fecha: Diciembre 2012

Dedicatoria

A las dos mujeres que amo Ana Isabel y Juana Marcela por su paciencia y sacrificio. Y al hermanit@ de Ana, que seguramente vendrá pronto...

A mis padres y hermanos

A Margarita y Jorge Enrique

Agradecimientos

Este trabajo fue desarrollado dentro del proyecto “Adapting agriculture to climate change: collecting, protecting and preparing crop wild relatives” liderado por Nora Patricia Castañeda y Andy Jarvis en el Centro Internacional de Agricultura Tropical (CIAT).

A Nora Castañeda por las ideas que dieron origen a este proyecto y su amistad

A Eduard Guevara por sus aportes desinteresados

A Juana Córdoba por su colaboración con la redacción

A Marcela Quintero y Gustavo Ligarreto por su comprensión y amistad

A Leonardo Cruz por su colaboración en la parte estadística

Muy especialmente a los participantes de los foros virtuales de Maxent, Python y R, por que sin ellos este trabajo no hubiera sido posible.... Mis colaboradores en la Nube!!!

Resumen

La evaluación del cambio climático en parientes silvestres de cultivos permite tomar decisiones para su conservación *in situ*. Las variables bioclimáticas sirven como predictores para modelar nichos y evaluar los efectos del cambio climático. En este estudio se utilizaron datos de ocurrencia/pseudo-ausencias de 34 taxones del género *Phaseolus* parientes del frijol común. Mediante un análisis de componentes principales (ACP) y la evaluación de la significancia biológica de las variables se removieron las que presentaban información redundante. Se modelaron nichos en Maxent en condiciones presentes y futuras (2020-2059 SRESA1B), usando todas y una selección de variables y teniendo en cuenta migración y no migración. Para 27 taxones se seleccionó el modelamiento con 10 variables seleccionadas. Ninguno de los taxones presentó peligro de extinción, pero de no migrar sus nichos se fragmentarán y perderían área, pero si hay migración habrá nichos ganadores y perdedores. La zona de mayor pérdida de biodiversidad ocurriría en la parte norte de la Sierra Madre Occidental en México y la mayor ganancia, cuando hay migración, estaría al sur de esta. No se encontró una tendencia clara entre el tamaño de los nichos actuales y el número de ocurrencias con el nicho futuro y su fragmentación. Los resultados permitieron la validación de la metodología utilizada para la remoción de variables.

Palabras clave: cambio climático, parientes silvestres de cultivos, sistemas de información geográfica (SIG), frijol común, modelamiento de distribución de especies (MDE), Maxent.

Abstract

Climate change assessment on crop wild relatives is useful for taking decisions about their *in situ* conservation. The bioclimatic variables can be used as environmental predictors for niche modeling and assessment of the climate change effect. In this study, data of occurrence/pseudo-absence from 34 taxa of the *Phaseolus* genera relatives of common bean were used. In order to remove the variables that present redundant information, Principal Component Analysis (PCA) and biological significance of the variables were analyzed. Niches were modeled using Maxent under present and future conditions (2020-2059 SRESA1B), using all and a selection of variables and under migration and no-migration. In 27 taxa only 10 variables were selected for modeling. No extinction danger were found for any taxon, however if the taxon were unable to migrate its niche would fragment and loss area, but under migration some niches would be winners and other losers. The zone that could lose more biodiversity is the north of the Sierra Madre Occidental in Mexico and under migration the south of this region would gain. No evident trend was seen about the actual size of the niches or the number of occurrences and the future niche and its fragmentation. The results allowed the validation of the method used for removing variables.

Key words: climate change, crop wild relatives, geographic information systems (GIS), common bean, spatial distribution model (SDM), Maxent.

TABLA DE CONTENIDO

Resumen	6
Abstract.....	7
LISTA DE TABLAS	11
LISTA DE FIGURAS	13
1. INTRODUCCIÓN.....	14
2. MARCO DE REFERENCIA	16
2.1 Sistemas de información geográfica (SIG) y la biodiversidad	16
2.2 Parientes silvestres de los cultivos de importancia para la alimentación y la agricultura	17
2.3 Conceptos de modelos de distribución de especies y nicho ecológico.....	18
2.4 Modelamiento de distribución de especies usando el principio de máxima entropía	20
2.5 El cambio climático y la evaluación de sus efectos.....	24
2.6 Variables bioclimáticas	25
3. MATERIALES.....	27
3.1 Bases de datos de germoplasma	27
3.2 Variables bioclimáticas actuales.....	27
3.3 Variables bioclimáticas bajo condiciones de cambio climático	28
3.4 Variables del terreno del sitio de colecta	28
3.5 Software utilizado	29
3.6 Hardware.....	30
4. PROCESO METODOLÓGICO.....	31
4.1 Organización de bases de datos	32
4.2 Análisis exploratorio de datos	33
4.3 Análisis de componentes principales.....	34
4.4 Selección de variables para el modelamiento	35

4.5 Modelamiento de nichos ecológicos en Maxent	35
4.6 Selección de modelos para el análisis espacial	38
4.6.1 Evaluación estadística de la precisión de los modelos generados en Maxent	38
4.6.2 Análisis de la estabilidad en la predicción logística realizada por Maxent	40
4.6.3 Criterios tenidos en cuenta para la selección del modelamiento a ser seleccionado	42
4.7 Proceso de análisis espacial	43
4.7.1 Pasos de análisis espacial seguido para la presentación y comparación de nichos ecológicos	44
4.7.2 Pasos de análisis espacial seguido para el análisis de fragmentación de nichos	47
4.7.3 Pasos de análisis espacial seguido para el análisis de riqueza	48
5. RESULTADOS	49
5.1 Análisis exploratorio de datos	49
5.1.1 Análisis de asociación de variables ambientales	49
5.1.2 Análisis exploratorio de datos a nivel espacial	51
5.2 Análisis de componentes principales.....	57
5.3 Selección de variables para el modelamiento	61
5.4 Modelamiento en Maxent.....	65
5.5 Evaluación estadística de la precisión de los modelos generados en Maxent	65
5.6 Análisis de la estabilidad en la predicción logística realizada por Maxent.....	68
5.7 Criterios tenidos en cuenta para la selección del modelamiento a ser seleccionado	71
5.8 Proceso de análisis espacial.....	73
5.8.1 Presentación y comparación de nichos ecológicos.....	74
5.8.2 Análisis de fragmentación de nichos ecológicos para las condiciones actuales y de cambio climático	84
5.8.3 Análisis de riqueza de especies silvestres del género Phaseolus	88
6. DISCUSIÓN.....	92
6.1 Selección de variables y de modelos de nicho ecológico usando el principio de máxima entropía	92
6.1.1 Determinación de criterios para la selección de variables	92
6.1.2 Selección de variables para el modelamiento de nichos ecológicos.....	93
6.1.3 Evaluación de parámetros para la selección de modelos	95

6.1.4 Selección de modelos de nicho ecológico obtenidos con Maxent.....	97
6.2 Evaluación de los efectos del cambio climático en parientes silvestres del frijol común	99
7. CONCLUSIONES Y RECOMENDACIONES	102
8. BIBLIOGRAFIA	104
8. ANEXOS	111
Anexo 1– Código en R para la realización del análisis de componentes principales	111
Anexo 2 – Código en R usado para la extracción de las predicciones logísticas de los modelamientos en Maxent del background y los datos de prueba	112
Anexo 3 – Scripts en Python 2.6 utilizados para la evaluación de la estabilidad en la predicción logística de los modelamientos.	113
Anexo 4 – Script usados en el proceso de análisis espacial	117

LISTA DE TABLAS

Tabla 1 - Resumen de estadísticos dependientes del umbral usados para evaluar la precisión estadística en modelos de distribución de especies.....	23
Tabla 2– Descripción de las variables incluidas en el análisis de componentes principales (ACP).	29
Tabla 3 – Número de datos de ocurrencia incluidos en el estudio por cada uno de los 34 taxones seleccionados con 10 o más datos.....	33
Tabla 4 – Contenido de los archivos <i>Phaseolus_all.bat</i> y <i>Phaseolus_sel.bat</i> corridos en Java para la generación de los modelos de nicho ecológico en condiciones presentes y futuras en Maxent.....	36
Tabla 5 – Matriz de coeficientes de determinación de Pearson (Parte Inferior) y significancia estadística de las correlaciones lineales (Parte Superior) para las variables incluidas en el estudio.	51
Tabla 6 – Valores propios, porcentaje de variabilidad explicada y acumulada de acuerdo al ACP realizado de las variables bioclimáticas, altura sobre el nivel del mar y pendiente del sitio de colecta para 34 taxones del genero <i>Phaseolus</i>	57
Tabla 7– Composición de los 4 primeros vectores propios y correlaciones de cada variable con el vector.	59
Tabla 8 – Variables de mayor participación en cada componente principal organizadas de acuerdo a su importancia y sentido de la correlación.....	62
Tabla 9 - Promedio de los índices AUC, Kappa y TSS calculados para los 10 modelos ajustados mediante validación cruzada en 34 taxones del genero <i>Phaseolus</i> usando 21 y 10 variables.	67
Tabla 10 – Área total de la cobertura de distribución potencial y porcentaje de la misma con desviaciones estándar menores a 0.1 de acuerdo a la totalidad (21 variables) o la selección (10 variables) de variables para 34 taxones del genero <i>Phaseolus</i>	70
Tabla 11 – Puntajes asignados a cada modelamiento de acuerdo a la evaluación de la precisión estadística y la estabilidad en la predicción logística para los 34 taxones del genero <i>Phaseolus</i>	73

Tabla 12 – Comparación de los tamaños de los nichos ecológicos en condiciones presentes y de cambio climático bajo los escenarios de no migración y migración para 34 taxones del genero <i>Phaseolus</i>	77
Tabla 13– Número (NP) y tamaño promedio de los parches (TPP) que conforman los nichos ecológicos de 34 taxones del genero <i>Phaseolus</i> bajo condiciones actuales y de cambio climático sin y con capacidad de migración.....	88

LISTA DE FIGURAS

Figura 1-Diagrama de flujo seguido en el desarrollo del estudio.....	31
Figura 2– Gráficos de dispersión con histogramas de las variables incluidas en el estudio	51
Figura 3– Distribución espacial de los datos de ocurrencia de 34 especies del género <i>Phaseolus</i> , su área nativa y el área del rectángulo más pequeño que contiene todos los datos de ocurrencia.....	54
Figura 4– Representación de las variables originales sobre los cuatro primeros vectores propios.....	61
Figura 5– Representación del área predicha mediante el modelamiento con la selección de variables (Eje X) y la totalidad de las mismas (Eje Y) para los 34 taxones del genero <i>Phaseolus</i>	71
Figura 6– Representación del tamaño de los nichos ecológicos de 34 taxones del genero <i>Phaseolus</i> bajo condiciones actuales y de cambio climático y dos escenarios de capacidad migratoria.	76
Figura 7– Comparación de los nichos ecológicos de 34 taxones del genero <i>Phaseolus</i> predichos con Maxent para las condiciones bioclimáticas actuales y de cambio climático.	81
Figura 8– Comparación de promedios del número de parches (A) y tamaño promedio de los parches (B) para los taxones agrupados en terciles de acuerdo al tamaño del nicho en condiciones actuales.....	85
Figura 9– Comparación de promedios del número de parches (A) y tamaño promedio de los parches (B) para los taxones agrupados de acuerdo al porcentaje de nuevas áreas que conformarían sus nichos en condiciones de cambio climático y migración.	87
Figura 10- Riqueza de taxones del genero <i>Phaseolus</i> modelada en Maxent para las condiciones actuales (A), de cambio climático y no migración (B) y su cambio (C).....	89
Figura 11- Riqueza de taxones del genero <i>Phaseolus</i> modelada en Maxent para las condiciones actuales (A), de cambio climático y migración (B) y su cambio (C).....	91

1. INTRODUCCIÓN

Los parientes silvestres de los cultivos de importancia para la alimentación y la agricultura son fuentes importantes de genes de resistencia a factores bióticos y abióticos de la producción, así como del mejoramiento de su calidad nutricional, rendimiento y otros caracteres de interés para el desarrollo del cultivo (Hajjar & Hodgkin, 2007). En el caso del fríjol común (*Phaseolus vulgaris* L.), la leguminosa de grano de mayor consumo en el mundo y fuente importante de proteínas y minerales (Broughton et al., 2003), sus parientes silvestres han sido de gran utilidad para su mejoramiento genético, por ejemplo han servido como fuente de resistencia a plagas como el gorgojo del fríjol (*zabrotes subfasciatus* (Boh.)) o a enfermedades como la bacteriosis común del fríjol (*Xanthomonas campestris* pv. *phaseoli* (Smith) Dye) (Acosta-Gallegos et al., 1998; Acosta-Gallegos & Gepts, 2007; Singh & Munoz, 1999). Parte de la diversidad de los parientes silvestres de cultivos han sido colectados y mantenidos a nivel mundial por herbarios y bancos de germoplasma, lo cual ha facilitado su uso para la agricultura. Sin embargo, la conservación *in situ* de estos parientes silvestres se hace necesaria debido a los constantes cambios en su diversidad genética estimulados por su entorno natural, pero en la actualidad los ecosistemas en los cuales estos recursos fitogenéticos se desarrollan están siendo amenazados entre otras razones por prácticas ambientalmente insostenibles y el cambio climático (Heywood, Casas, Ford-Lloyd, Kell, & Maxted, 2007; Maxted & Kell, 2009).

Una de las necesidades urgentes para la conservación de los parientes silvestres de los cultivos es la de poder estimar los efectos que el cambio climático pueda tener sobre el rango ambiental o nicho ecológico en el que estos se desarrollen (Jarvis, Lane, & Hijmans, 2008; Maxted et al., 2010). Para esto, la aproximación de mayor uso es el modelamiento espacial de los cambios en los nichos ecológicos (Morin & Lechowicz, 2008), los cuales se basan en la estimación en un espacio geográfico de la distribución potencial de las especies mediante la combinación de datos de ocurrencias y/o ausencias, variables ambientales o ecológicas en condiciones del pasado, presentes y futuras de acuerdo a diversos modelos climáticos y

diversos algoritmos disponibles para su cálculo (De Souza Muñoz et al., 2011; IPCC, 2007). Razones por las que el análisis de información, el modelamiento y el uso de herramientas de análisis espacial, en un ambiente de sistemas de información geográfico (SIG), se hace necesario para generar información útil para la toma de decisiones en torno a la conservación de estos recursos.

De esta forma, uno de los insumos fundamentales para realizar estos modelamientos son la información espacial a una escala espacio-temporal adecuada, por lo que se han utilizado variables bioclimáticas de las condiciones actuales y futuras, bajo varios escenarios de cambio climático (Hijmans et al., 2005; Ramirez & Jarvis, 2010), las cuales son índices con sentido biológico calculados a partir de modelos mensuales de precipitación y temperatura. Sin embargo, algunas de las variables bioclimáticas son calculadas a partir de otras o representan la misma medición en diferentes periodos de tiempo, y estas relaciones se hacen más fuertes de acuerdo a su localización geográfica. Razón por la cual el ajuste de modelos de nicho ecológico basado en la totalidad de estos predictores ambientales podría generar modelos con altos niveles de sobreajuste, especialmente en especies con pocos datos de ocurrencia (Guisan & Thuiller, 2005; Phillips & Elith, 2010).

Por tal motivo, en el presente estudio se busca evaluar los efectos del cambio climático entre los años 2020 a 2059, sobre 34 taxones del género *Phaseolus*, parientes silvestres del frijol común, para lo cual antes de realizar el modelamiento de nichos ecológicos, realizado en el aplicativo Maxent (Phillips, Dudik, & Schapire, 2011), se plantea realizar la remoción de algunas variables bioclimáticas basado en un análisis de componentes principales (ACP) a los datos de estas en los puntos de ocurrencia de los taxones.

2. MARCO DE REFERENCIA

2.1 Sistemas de información geográfica (SIG) y la biodiversidad. Los sistemas de información geográfica (SIG) tienen tres características de importancia las cuales son trabajar con información referenciada espacialmente, involucrar el uso de computadoras y permitir a las personas u organizaciones que los utilizan generar información para la toma de decisiones. Así, los SIG son sistemas asistidos por computador que permiten la captura, almacenamiento, recuperación, búsqueda, análisis, modelación y representación de información espacial, con el fin de ayudar en la toma de decisiones en torno a un problema (Chang, 2004; Cowen, 1988; Heikkila, 1998; Jankowski, 1995; Savigny & Wijeyaratne, 1995).

La información de un SIG se compone de tres tipos de datos que interactúan y forman un complejo espacio tridimensional y multivariado. Dicho espacio está compuesto por las coordenadas geográficas (dimensión espacial), coordenadas temporales (dimensión temporal) y múltiples atributos relacionados a los objetos geográficos representados (dimensión temática), los cuales al analizarlos en conjunto permiten la toma de decisiones (Openshaw, 1999).

Desde sus inicios en los años 1960, los SIG han estado ligados al medio ambiente, estos surgieron en Canada donde fueron utilizados para el manejo de los bosques y sus recursos naturales, al igual que en los Estados Unidos en donde además se utilizaron para el manejo urbano. Adicionalmente, aplicaciones dedicadas a lo ambiental fueron desarrolladas en la República Popular China, Hungría, República Federativa Checa y Eslovaca, junto a Japón (Foresman, 1998).

En términos de la biodiversidad, la generación de aplicaciones SIG, debido a la diversidad y complejidad de la información manejada (Salem, 2003), ha resultado ser una herramienta indispensable para su exploración, evaluación, monitoreo y conservación (Hijmans, Guarino, Cruz, & Rojas, 2001; Jones, Beebe, Tohme, & Galwey, 1997; Salem, 2003). De esta forma, diversos tipos y múltiples aplicaciones han sido desarrolladas en las últimas décadas,

considerándose incluso que los SIG han sido una de las herramientas más poderosas en el campo de la biogeografía y por ende en el de los análisis de biodiversidad (Foody, 2008).

2.2 Parientes silvestres de los cultivos de importancia para la alimentación y la agricultura.

Los parientes silvestres de los cultivos (CWR, por sus siglas en inglés¹) son especies relativamente cercanas a los cultivos, los cuales no han sufrido ningún proceso de selección por parte del hombre, incluyendo en este grupo a los progenitores de los cultivos actuales, los cuales tienen la capacidad potencial de ampliar la base genética de los mismos. Estos parientes silvestres son fuentes de genes para el mejoramiento genético de los cultivos, aportando genes responsables de la resistencia a múltiples plagas y enfermedades, incremento del rendimiento, estabilidad, adaptación a condiciones ambientales adversas y en general a muchos caracteres de interés para la agricultura y la alimentación (Heywood et al., 2007; Maxted & Kell, 2009; Maxted et al., 2010).

Aunque desde el inicio del desarrollo de la agricultura, los agricultores los han utilizado de manera consciente e inconsciente, por ejemplo mediante la selección de algunos de ellos como progenitores de los cultivos actuales, recién en el siglo XX, se ha reconocido su importancia de forma consciente y se han comenzado a utilizar en el mejoramiento moderno (Heywood et al., 2007). El primer investigador científico que reconoció su importancia fue Vavilov, quien en los años 1920 y 1930, los incluyó en sus programas de colección de germoplasma. Los investigadores agrícolas comenzaron su uso en los 1940 y 1950 para mejorar los principales cultivos. En los años 1960 y 1970, su uso exitoso en los programas de mejoramiento genético se aceleró, especialmente usando especies que hacen parte del acervo genético primario de los cultivos, es decir, aquellos parientes silvestres que generan fácilmente prole luego del cruzamiento con los cultivos de interés. Recientemente, a partir de los años 1980, con el desarrollo de la ingeniería genética el uso de parientes silvestres de otros acervos genéticos ha sido posible (Meilleur & Hodgkin, 2004).

La conservación tanto *ex situ* en herbarios y bancos de germoplasma como *in situ* en su ambiente natural son necesarias para su conocimiento, conservación y uso (Hunter &

¹ CWR: Crop Wild Relatives

Heywood, 2011; Maxted et al., 2010; Meilleur & Hodgkin, 2004). Aunque la conservación de estos recursos se ha enfocado en las condiciones *ex situ*, lo cual ha sido muy valioso y ha permitido su uso y conocimiento, la conservación *in situ* se hace necesaria debido entre otras a que al estar en constante contacto con su ambiente natural estos recursos están en un proceso constante de selección, competencia y evolución, permitiendo de esta forma cumplir con su función de donantes de genes. Por esta razón a nivel mundial en los años recientes se vienen liderando diferentes esfuerzos para su conservación (Croptrust, 2010; Heywood et al., 2007; Hunter & Heywood, 2011; Meilleur & Hodgkin, 2004). Así existe el tratado internacional sobre recursos fitogenéticos para la alimentación y la agricultura en el cual conscientes de la importancia y el riesgo en el que se encuentran estos recursos, ha acordado entre otras, las obligaciones de los actores mundiales en el tema, acciones y mecanismos para su conservación, prospección, recolección, caracterización, evaluación y documentación de estos recursos y prioriza los de importancia para ser conservados (Croptrust, 2010; FAO, 2004).

2.3 Conceptos de modelos de distribución de especies y nicho ecológico. Los modelos de distribución de especies (SDM- por sus siglas en inglés²), son modelos empíricos que relacionan las observaciones de campo (presencias, presencias/ausencias ó abundancias) que han colectado diversas personas a través del tiempo con variables o predictores ambientales y buscan determinar la distribución potencial de las especies en una determinada área geográfica (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). Desde los orígenes de los SDM, estos se han basado en la cuantificación de la relación especie-ambiente, pasando por tres etapas las cuales son la cuantificación estadística de dicha relación (cuantificación no espacial), basado en datos empíricos. En un segundo momento, fueron realizados basados en información de expertos (cuantificación no estadística, no espacial). En un tercer momento se han realizado mediante estadística explícitamente espacial y modelos empíricos de la distribución de las especies (Guisan & Thuiller, 2005).

En general el proceso seguido para construir un SDM comienza por una etapa de conceptualización y preparación de los datos, en la cual entre otros pasos, se recolecta información necesaria y relevante, se decide que variables usar y como lidiar con variables

² SDM: Spatial Distribution Model

relacionadas, se selecciona el algoritmo adecuado y los parámetros a usar en el mismo de acuerdo a la aplicación requerida. Luego, se realiza el ajuste y evaluación del modelo, para lo que se utiliza el algoritmo y parámetros seleccionados, se utilizan estadísticos, comparaciones y otros medios para la evaluación de los modelos. En tercer lugar, es realizada la predicción espacial, evaluación de la aplicabilidad del modelo, siendo un proceso iterativo con el fin de mejorar las predicciones basados en los aprendizajes obtenidos (Elith & Leathwick, 2009; Guisan & Thuiller, 2005).

El modelamiento de la distribución potencial de las especies es entonces el proceso de construir una representación espacial de los requerimientos ecológicos fundamentales para una especie y extrapolarlos dentro de una región geográfica, constituyendo así el nicho ecológico fundamental de la especie (De Souza Muñoz et al., 2011). Así aunque, los requerimientos ecológicos fundamentales, son aquellos recursos que le permiten a una especie establecerse y reproducirse en un sitio sin la necesidad de que otros individuos tengan que inmigrar para que la población se mantenga, las especies no ocupan la totalidad del espacio geográfico que potencialmente pueden ocupar, debido entre otras razones a limitaciones en la habilidad de dispersión, competencia y aislamiento geográfico de los espacios, por lo que solamente ocupan el nicho ecológico realizado (Peterson, 2003).

De esta forma, varios autores han discutido la relación entre los conceptos de modelos de distribución de especies (SDM) y nicho ecológico, aunque no se ha llegado a un concepto único (Elith & Leathwick, 2009; Guisan & Thuiller, 2005). Para varios de ellos, al ser los SDM calculados en base a datos de ocurrencia en donde la especie se encontraba en constante proceso de selección y competencia, los SDM permiten modelar los nichos ecológicos realizados, para otros permite representar los nichos ecológicos fundamentales, debido a la falta de información sobre competencia y otros factores limitantes (Guisan & Thuiller, 2005; Morin & Lechowicz, 2008; Wiens, Stralberg, Jongsomjit, Howell, & Snyder, 2009). Sin embargo, en la mayoría de trabajos, esta distinción no se ha realizado y se usa de forma neutral. Esto puede obedecer como lo sugieren Elith & Leathwick (2009), a que al evaluar los SDM y advertir sobre los cuidados en la interpretación de los modelos, promover el ímpetu por coleccionar mejores datos de ocurrencia, realizar experimentos ecológicos, modelar mecanismos de dispersión, efectos de competidores, entre otros. De esta forma, que los SDM puedan cumplir con los objetivos para los cuales fueron planteados como son: búsqueda de

especies raras y amenazadas, planificación de áreas de conservación, evaluación de los impactos de actividades humanas sobre la biodiversidad, predecir los impactos del cambio climático sobre la distribución de las especies, prevenir la dispersión de especies invasoras, identificar vectores de enfermedades, entre otras (De Souza Muñoz et al., 2011).

2.4 Modelamiento de distribución de especies usando el principio de máxima entropía. Existen múltiples y diversos tipos de algoritmos y aplicativos para modelar la distribución de especies, los cuales utilizan datos de abundancia, presencia-ausencia o solamente presencia (De Souza Muñoz et al., 2011). Sin embargo, el caso más común es la disponibilidad de datos únicamente de presencia y en ese caso el algoritmo de mayor utilización es el de máxima entropía implementada en el aplicativo Maxent (Phillips, Anderson, & Schapire, 2006; Phillips et al., 2011) el cual ha sido considerado como el de mejor desempeño para este tipo de aplicaciones con datos únicamente de presencia (Elith et al., 2011; Phillips & Dudík, 2008; Ramirez-Villegas et al., 2010).

Maxent es un aplicativo que utiliza como método el lenguaje de aprendizaje de maquina para realizar el modelamiento de nichos ecológicos (Phillips & Dudík, 2008). El fundamento del algoritmo que utiliza es estimar una distribución de probabilidad de máxima entropía, es decir, la que más expandida o más cercana a la uniformidad (Phillips, Anderson, & Schapire, 2006). Aunque el aplicativo es fácil de usar y de amplia aceptación para este tipo de análisis, su algoritmo es de difícil entendimiento, sin embargo Elith et al. (2011) plantean la explicación de su algoritmo en términos estadísticos.

De esta forma, el funcionamiento del aplicativo se basa en la introducción de datos de ocurrencia de una especie, variables ambientales cuantitativas o cualitativas de toda la zona de estudio en formato raster con igual resolución espacial y marco de referencia en formato ascii grid y generar puntos de pseudo-ausencias, los cuales pueden ser generados por Maxent aleatoriamente en el área de estudio o por el usuario. A partir de estos datos, Maxent genera modelos de probabilidad y ajusta modelos de acuerdo al comportamiento de cada variable y del conjunto de ellas, proyectando luego en un mapa en formato raster a la misma resolución de las variables de entrada, la probabilidad que maximiza la entropía de acuerdo a los valores de todas las variables en el pixel y el modelo ajustado. Esta proyección puede realizarse sobre

diferentes zonas y tiempos, siempre y cuando se cuente con mapas en formato raster a la misma resolución de las mismas variables. El valor de probabilidad representado en cada pixel de la proyección, representa la probabilidad condicional de tener al menos un porcentaje (por defecto el 50%) de las condiciones bioclimáticas favorables de acuerdo a los datos de presencia de la especie (Phillips et al., 2006). Sin embargo, la proyección en diferentes tiempos puede tener valores por fuera del rango de variación de los datos de entrada, caso en el cual Maxent asume el comportamiento del máximo valor de la condición actual, en lo que se denomina *Clamping*, esto lo hace debido a la incertidumbre sobre el comportamiento de las especies en ese escenario (Heikkinen et al., 2006; Wiens et al., 2009).

Otros de los aspectos que el usuario del aplicativo puede personalizar es el nivel de ajuste que desea de la curva de probabilidad, el cual se denomina la constante de regularización, la cual al incrementarse de 1 indica un menor ajuste a los datos de entrada y al disminuir uno mayor. Otro de gran importancia es el tipo y número de replicas que se desea correr del modelo, el cual define el tipo de validación a realizar, dentro de estas se encuentra la validación cruzada, en la que se divide el conjunto de ocurrencias en tantos grupos como validaciones cruzadas o modelos a correr, otro es el Bootstrap en el que se saca aleatoriamente una muestra cada vez que se corre el modelo y el tercero es un submuestreo en el que se selecciona una cantidad de ocurrencias para que sirvan como datos de prueba (Phillips & Dudík, 2008).

Adicionalmente a los mapas generados por Maxent, al igual que varios algoritmos de modelamiento que generan distribuciones de probabilidad de los modelos de distribución de especies, se generan varios parámetros de importancia para realizar la evaluación de los modelamientos realizados, siendo el de mayor utilización la curva ROC (*The receiver operating characteristic*) y el área debajo de esta curva (AUC- por sus siglas en inglés³) (Fielding & Bell, 1997).

Para entender que significa la curva ROC y la forma en que se realizan las evaluaciones de la precisión estadística en un modelamiento de distribución de especies, es necesario nombrar que el objetivo final es determinar una superficie binaria de presencias y ausencias de la especie (Liu, Berry, Dawson, & Pearson, 2005), esto a partir de la selección de un umbral. Las clasificaciones realizadas a partir de dicho umbral generan los valores de la denominada

³ AUC: Area Under Curve receiver operating characteristic

matriz de confusión: a) verdaderos positivos, b) falsos positivos, c) falsos negativos, d) verdaderos negativos. A partir de esta matriz de confusión se generan la mayoría de estadísticos para la evaluación estadísticas de los modelos, siendo las dos más sencillas la sensibilidad y la especificidad. La sensibilidad del modelo esta dada por el porcentaje de presencias predichas en el modelamiento que son observadas como son $(a / a+c)$ y cuantifica los errores de omisión. La especificidad es el porcentaje de ausencias correctamente predichas $(d / b+d)$ y cuantifica el error de comisión (Allouche, Tsoar, & Kadmon, 2006; Fielding & Bell, 1997; Liu et al., 2005).

De esta forma, la curva ROC es el resultado de dibujar en un plano XY todos los posibles valores de especificidad $(1 - \text{especificidad})$ en el eje X y los de sensibilidad en el eje Y, de acuerdo a todos los posibles umbrales o valores de probabilidad (valores entre 0 y 1). Por su parte, mediante la cuantificación del área debajo de la curva ROC (AUC), es posible tener una evaluación independiente del umbral seleccionado. Un valor de 0.5 de AUC, corresponde a una distribución completamente aleatoria y un área de 1 a una clasificación perfecta y sin errores (Fielding & Bell, 1997), sin embargo esta medida ha sido cuestionada debido a que puede ser dependiente de los datos de ocurrencia y la prevalencia de predicción, entre otras razones, por lo cual se recomienda hacer una evaluación conjunta con otros estadísticos basados en la predicción binaria (Allouche et al., 2006; Liu et al., 2005; Lobo, Jiménez-Valverde, & Real, 2008).

Por su parte, los estadísticos usados para evaluar la precisión estadística de los modelos y que dependen del umbral seleccionado son múltiples, algunos de ellos se nombran en la Tabla 1. Su uso e interpretación es igualmente diversa, sin embargo al depender del umbral seleccionado, la forma en que este se determiné es importante para la evaluación del modelo. Existen diversos criterios para definir el umbral en un SDM, los cuales pueden clasificarse en dos categorías: los subjetivos y los objetivos (Liu et al., 2005). En la primera categoría los de mayor uso son 0.5, 0.3 y 0.05, sin embargo estos son arbitrarios y carecen de bases ecológicas. Por su parte, los umbrales objetivos pueden clasificarse en varios grupos, entre los que se destacan los que maximizan alguno de los estadísticos, por ejemplo kappa y los que se basan en la combinación de la sensibilidad y especificidad, por ejemplo, los que maximizan la suma entre la sensibilidad y especificidad, los que buscan igualarla o aquellos que buscan maximizar

alguna de ellas o ambas, este último corresponde al punto de la curva ROC más cercano al punto con coordenadas (0,1) (Liu et al., 2005).

Tabla 1 - Resumen de estadísticos dependientes del umbral usados para evaluar la precisión estadística en modelos de distribución de especies

Medida	Cálculo*
Prevalencia	$(a + c)/N$
Poder diagnóstico general	$(b + d)/N$
Tasa de clasificaciones correctas o precisión general	$(a + d)/N$
Sensibilidad	$a/(a + c)$
Especificidad	$d/(b + d)$
Tasa de falsos positivos	$b/(b + d)$
Tasa de Falsos negativos	$c/(a + c)$
Precisión	$a/(a+b)$
Poder predictivo positivo (PPP)	$a/(a + b)$
Poder predictivo negativo (NPP)	$d/(c + d)$
Tasas de clasificaciones equivocadas	$(b + c)/N$
Odds-ratio	$(ad)/(cb)$
Kappa	$[(a + d) - (((a + c)(a + b) + (b + d)(c + d))/N)] / [N - (((a+c)(a+b) + (b+d)(c+d))/N)]$
Normalized mutual information statistisc (NMI)	$[-a*\ln(a)-b*\ln(b)-c*\ln(c)-d*\ln(d) + (a+b)*\ln(a+b)+(c+d)*\ln(c+d)] / [N*\ln N - ((a+c)*\ln(a+c) + (b+d)*\ln(b+d))]$
TSS (True Skill Statistic)	Sensitividad + Especificidad - 1

* $N = a + b + c + d$

Fuente: (Allouche et al., 2006; Fielding & Bell, 1997; Liu et al., 2005)

Luego del cálculo de los parámetros necesarios para evaluar la precisión estadística de los modelos de distribución de especies, con la finalidad de determinar el nivel de la misma, algunos autores han establecido escalas de acuerdo a los valores que toman los estadísticos. Por ejemplo, para el AUC los valores menores a 0.5 se consideran que son modelos con clasificaciones de menor precisión que un modelo aleatorio, para valores entre 0.5 y 0.7, modelos con baja precisión, modelos con valores entre 0.7 y 0.9 son modelos útiles y con precisiones mayores a 0.9 indican altas precisiones (Manel, Williams, & Ormerod, 2001). Según el estadístico kappa, valores menores a 0.05 no tienen ninguna precisión, entre 0.05 y 0.20 precisión muy pobre, entre 0.20 y 0.40 precisión pobre, entre 0.40 y 0.55 aceptable, entre 0.55 y 0.7 buena, entre 0.7 y 0.85 muy buena, entre 0.85 y 0.99 excelente y entre 0.99 y 1 perfecta (Monserud & Leemans, 1992).

2.5 El cambio climático y la evaluación de sus efectos. El panel intergubernamental sobre el cambio climático (IPPC) define el cambio climático como cualquier cambio del clima en el tiempo debido a la variabilidad natural o como resultado de la actividad humana (Parry, Rosenzweig, & Livermore, 2005). De la misma manera, advierte que la concentración global en la atmósfera de dióxido de carbono, metano y óxido nítrico en el año 2005 se había incrementado de forma importante desde 1750 entre 280 ppm y 379 ppm, como producto de la actividad humana, especialmente debido al uso de energías fósiles y el cambio en el uso del suelo. Esta situación se hizo evidente en las observaciones realizadas en los últimos años de la temperatura del aire y de los océanos, el constante y rápido derretimiento de las nieves y el hielo y el incremento del nivel del océano (IPPC, 2007).

A pesar de la evidencia en los cambios de las concentraciones de gases efecto invernadero que pueden generar el calentamiento o cambio climático global, estas concentraciones y sus posibles efectos sobre el clima, no son evidentes en el futuro, por tal motivo han sido planteados varios escenarios de emisiones de estos gases y del calentamiento global, claro esta bajo ciertos niveles de incertidumbre de acuerdo a la información disponible sobre cada uno de ellos (Nakicenovic et al., 2000). Así, fueron generados 4 familias de escenarios de emisiones y un total de 40 escenarios, los cuales difieren en muchos aspectos como el social, económico, tecnológico, ambiental y político. Las familias de escenarios fueron denominadas como A1, A2, B1 y B2. La familia A1 describe el futuro con un crecimiento económico rápido, bajo crecimiento poblacional y rápida introducción de nuevas y mejores tecnologías, con una reducción en las diferencias de ingreso regionales. El A2 describe un mundo heterogéneo, alto crecimiento poblacional y crecimiento económico enfocado en lo local, en este modelo el desarrollo tecnológico, la introducción de nuevas tecnologías es lenta y el crecimiento económico es muy fragmentado. El B1 describe un mundo con el mismo nivel de población que el A1, pero con rápidos cambios en las estructuras económicas, con reducción en el uso de materiales y la introducción de tecnologías limpias, son un desarrollo enfocado en lo global, pero sin iniciativas adicionales en cuanto al clima. El B2 describe un mundo en el cual el énfasis es hacia el desarrollo local, es un mundo con una población moderada, niveles intermedios de desarrollo económico y cambios tecnológicos menos rápidos que en el A1.

Basados en estos planteamientos, el IPPC delegó a varios institutos de investigación el cálculo de los escenarios de emisiones y los consecuentes modelos de circulación global (GCM⁴- por sus siglas en inglés), los cuales son los modelos numéricos que representan modelos físicos en la atmósfera, océano, en la criosfera y en la superficie terrestre. Estos modelos dividen la tierra en un grid tridimensional, que horizontalmente tiene una resolución entre 250 y 600 km y 10 a 20 capas en la atmósfera y en algunas ocasiones 30 en el océano (Nakicenovic et al., 2000).

Estas estimaciones, al tener una baja resolución espacial no permiten realizar la evaluación de los efectos del cambio climático al nivel de un paisaje. Por tal razón, basados en metodologías de análisis espacial y estadístico el Centro Internacional de Agricultura Tropical (CIAT), desarrolló metodologías para aumentar la resolución de las superficies generadas en cada (GCM). Este trabajo fue realizado para 24 GCMs y diferentes periodos de tiempo, para lo cual fueron calculadas las anomalías o deltas en los centriodos de cada GCM y luego estas son interpoladas espacialmente mediante el método *thin plate spline*, para finalmente ser adicionadas a una línea base, la cuales Worldclim (Hijmans et al., 2005) a varias resoluciones (30 segundos de arco, 2.5 segundos de arco, 5 segundos de arco y 10 segundos de arco) (Ramirez & Jarvis, 2010).

2.6 Variables bioclimáticas. Las variables bioclimáticas son índices con significado biológico que cuantifican los balances de energía y humedad de un sitio específico (Nix, 1986). Existen 36 variable bioclimáticas que pueden ser calculadas a partir de los promedios mensuales o semanales de máxima y mínima temperatura, precipitación, radiación solar y evapotranspiración, de las cuales las primeras 11(bio1 a 11) estan relacionadas con la temperatura, las siguientes 8 con la precipitación (bio12 a bio19), las siguientes 8 con la radiación solar (bio20 a 27) y las restantes 8 con la humedad (bio28 a 35)(Xu & Hutchinson, 2011). Cada una de estas variables puede resumir una condición promedio anual la cual es una aproximación gruesa a las entradas de energía y agua en el sitio, una condición promedio estacional la cual provee una medida de las condiciones prevalentes durante estaciones potencialmente activas y dormantes, un valor extremo en la estación que podría afectar o no la supervivencia de un organismo en el sitio o una condición de estacionalidad intra-anual la cual

⁴ GCM: Global Circulation Model

permite caracterizar los cambios en las condiciones durante el año (Busby & Austin, 1990). Para las 19 primeras variables bioclimáticas se encuentran estimativos con una resolución espacial de 1 km² en la página (<http://www.worldclim.org/bioclim>), estas variables fueron derivadas de las superficies con información de precipitación y temperatura Worldclim, las cuales son interpolaciones de datos de estaciones climáticas localizadas alrededor de todo el mundo y para el periodo entre 1950 a 2000 (Hijmans et al., 2005). Adicionalmente, es posible obtener información de las variables bioclimáticas proyectadas a condiciones de cambio climático bajo diferentes escenarios y tiempos, basados en parámetros del IPCC (2007) (Ramirez & Jarvis, 2010).

3. MATERIALES

3.1 Bases de datos de germoplasma. Fueron usadas las bases de datos de colecciones de germoplasma disponibles en la página web <http://gisweb.ciat.cgiar.org/GapAnalysis/> correspondientes al análisis de huecos en las colecciones de germoplasma (Gap Analysis) realizado para el acervo genético del género *Phaseolus* (Bioversity International, et al. , 2009). La base de datos original consta de 5146 registros correspondientes a 84 taxones del género *Phaseolus*, los cuales son las clasificaciones taxonómicas de mayor detalle encontradas en la base de datos y que será la clasificación tenida en cuenta para la realización de los análisis en este estudio. Así, en la base de datos, cada registro corresponde a un genotipo colectado en el campo, para el cual como mínimo se tiene la información de las coordenadas geográficas del sitio de colecta, datos sobre la o las instituciones que realizaron la colecta o mantienen la colección de germoplasma e información sobre la clasificación taxonómica, entre otras.

De la totalidad de los 5146 registros la mayoría de los taxones corresponden a las clasificaciones taxonómicas en género y especie (4343 registros), los restantes 803 registros corresponden a clasificaciones taxonómicas en género, especie y subespecie y/o forma varietal, formándose así los 84 taxones. En términos de la cantidad de registros por cada taxón 49 taxones tienen entre uno a diez registros, 22 entre 10 a 100 registros y solamente 13 taxones tienen más de 100 registros. Sin embargo, en el estudio se tendrán en cuenta registros con coordenadas únicas que representan en cada taxón una población, es decir, se asumirá que una población de un taxón específico será la que se encuentra en una misma coordenada geográfica o en su defecto en un territorio que no supera un kilómetro cuadrado de área (Ramírez-Villegas et al., 2010). En la sección de organización de bases de datos para el modelamiento se detalla el procedimiento seguido para este cálculo.

3.2 Variables bioclimáticas actuales. Fueron usadas 19 capas de las variables bioclimáticas las cuales se describen en la Tabla 2. Estos mapas contienen información a nivel mundial y una resolución espacial de 2.5" (5 Km aprox. en el Ecuador), estas variables bioclimáticas fueron obtenidas en formato ESRI-GRID del sitio web de worldclim (<http://www.worldclim.org/download>). Las variables bioclimáticas fueron desarrolladas a

partir de las superficies climáticas de worldclim (Hijmans et al., 2005), las cuales son superficies que contienen promedios mensuales de precipitación, temperatura máxima, promedio y mínima a nivel mundial. Adicionalmente, las superficies climáticas de worldclim son superficies climáticas desarrolladas mediante técnicas de interpolación espacial y con base en información entre los años 1950 y 2000 de estaciones climáticas a nivel mundial (Hijmans et al., 2005).

3.3 Variables bioclimáticas bajo condiciones de cambio climático. Fueron utilizadas las 19 capas de las variables bioclimáticas a la resolución de 2.5” generadas por el consorcio CCAFS para las condiciones futuras y el rango de años 2020 a 2059 (CCAFS, CIAT, ILRI, & IFPRI, 2012). Las capas usadas en este estudio son el producto del ‘downscaling’ estadístico (Ramirez & Jarvis, 2010) y posterior ensamblaje de las estimaciones realizadas para el rango de años entre el 2020 a 2059 realizado para 24 modelos de circulación global (GCM, por sus siglas en inglés) y el escenario de cambio climático SRES-A1B, el cual supone una utilización máxima de energía y un balance entre fuentes fósiles y alternativas de la misma, siendo estos modelos calculados en base a las recomendaciones del cuarto panel intergubernamental sobre cambio climático (IPPC, 2007).

3.4 Variables del terreno del sitio de colecta. Se incluyeron dos variables relacionadas con las condiciones del terreno del sitio de colecta del germoplasma, la altitud y la pendiente (Tabla 2). En este caso fue utilizado el modelo digital de elevación del terreno a la misma resolución espacial de las variables bioclimáticas descritas anteriormente, es decir 2.5”. La altitud se obtuvo del modelo digital de elevación del terreno (DEM) descargado del sitio web <http://srtm.csi.cgiar.org/>. A partir de este DEM y en haciendo uso de la herramienta slope (Spatial Analyst Tools>Surface>Slope) fue calculado el porcentaje de pendiente a la resolución de 2.5” en ArcGIS10(ESRI, 2011).

Tabla 2– Descripción de las variables incluidas en el análisis de componentes principales (ACP).

Label	Variable	Unidades	Grupo*
bio1	Temperatura media anual	°C	1
bio2	Rango medio de temperatura diurna	°C	4
bio3	Isotermalidad (bio2/bio7)*100	%	4
bio4	Estacionalidad de la temperatura (Desviación estandar*100)		4
bio5	Temperatura máxima del mes más caliente	°C	3
bio6	Temperatura mínima del mes más frío	°C	3
bio7	Rango anual de temperatura (bio5 - bio6)	°C	4
bio8	Temperatura media del trimestre más humedo	°C	2
bio9	Temperatura media del trimestre más seco	°C	2
bio10	Temperatura media del trimestre más caliente	°C	2
bio11	Temperatura media del trimestre más frío	°C	2
bio12	Precipitación anual	mm	1
bio13	Precipitación del mes más humedo	mm	3
bio14	Precipitación del mes más seco	mm	3
bio15	Estacionalidad de la precipitación (coeficiente de variación)	%	4
bio16	Precipitación del trimestre más humedo	mm	2
bio17	Precipitación del trimestre más seco	mm	2
bio18	Precipitación del trimestre más caliente	mm	2
bio19	Precipitación del trimestre más frío	mm	2
alt	Altitud	m	
pendiente	Pendiente	%	

* Hace referencia a la condición bioclimática que describe la variable: 1(Condiciones promedio anuales), 2 (Condiciones promedio estacionales), 3 (valores extremos en las condiciones bioclimáticas), 4 (Estacionalidad intra-anual)

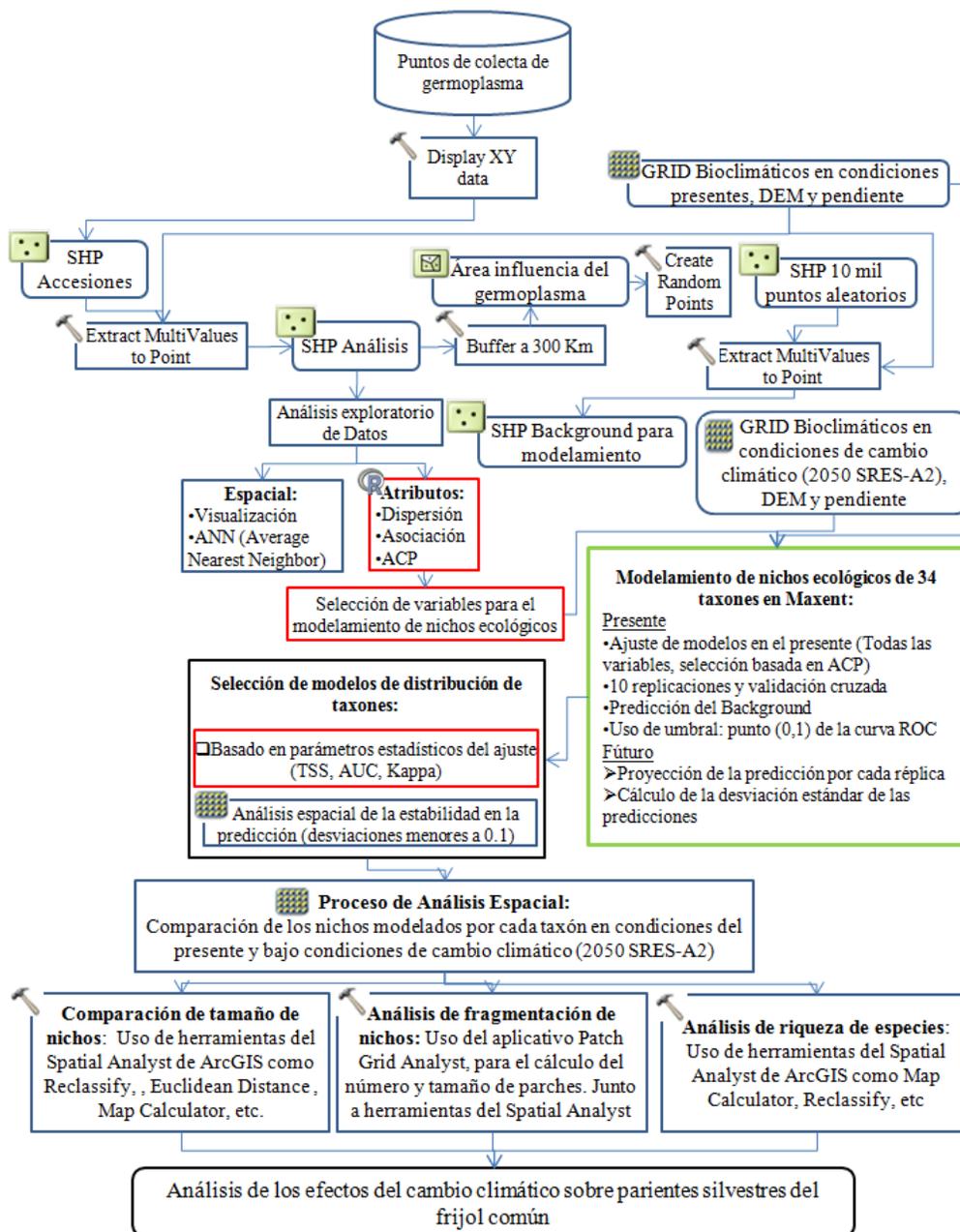
3.5 Software utilizado. El proceso de análisis espacial fue realizado en el Software ArcGis Desktop 10.0®, en donde fueron utilizadas las licencias ArcInfo, ArcEditor, ArcView y la extensión de análisis espacial (ESRI, 2011). Adicionalmente, debido a la necesidad de automatizar los procesos fueron desarrollados Script en Python 2.6 (Python Software Foundation, 2008), utilizando los paquetes propios del lenguaje de programación y el paquete arcpy (ESRI, 2011). Adicionalmente para el análisis de parches en la sección de análisis espacial fue utilizado el aplicativo Patch Analyst (Rempel, Kaukinen, & Carr, 2012). Por su parte, el análisis de nichos ecológicos fue realizado mediante el uso del software Maxent v3.3.1k (Steven J. Phillips et al., 2011). Finalmente, los análisis estadísticos fueron realizados en el lenguaje de programación R v. 2.15.0 (R Development Core Team, 2012), en donde fueron utilizados los paquetes básicos y el paquete FactoClass (Pardo & Del Campo, 2007).

3.6 Hardware. Este estudio fue desarrollado en su totalidad en un computador con sistema operativo Windows de 64 bits, procesador Intel Pentium Dual Core de 2.2 GHz, memoria RAM de 4 Gb y 250 Gb en disco Duro.

4. PROCESO METODOLÓGICO

En la Figura 1 se muestra un diagrama de flujo del proceso metodológico seguido, adicionalmente en las siguientes secciones se detalla el procedimiento seguido. En este esquema en recuadros azules se muestran los procesos desarrollados en ArcGis, en rojo los deguidos en R y en verde los desarrollados en Maxent.

Figura 1-Diagrama de flujo seguido en el desarrollo del estudio



4.1 Organización de bases de datos. A partir de la base de datos de germoplasma del género *Phaseolus* nombrada anteriormente, fue obtenida la base de datos definitiva para realizar el estudio mediante un proceso que consistió de tres pasos. En primer lugar, mediante el uso de las coordenadas geográficas de los puntos de colecta y dentro del software ArcGis 10.0 (ESRI, 2011), se generó un *shapefile* de puntos, donde fueron representados espacialmente todos los datos de ocurrencia, para lo cual se utilizó el elipsoide de referencia WGS84 en el cual se encontraban los mapas de las variables ambientales a usar en este estudio. En segundo lugar, para cada una de los taxones de la base de datos, se simplificó el número de puntos de ocurrencia a una resolución de 1 km², considerando que los puntos que se encontraban dentro del mismo pixel de 1 km² pertenecían a la misma población de plantas de acuerdo a lo planteado por (Ramirez-Villegas et al., 2010). En tercer lugar, luego de simplificadas los datos de ocurrencia por cada taxón, fueron removidos de la base de datos aquellos taxones con un número de datos de ocurrencia menores a 10, los cuales son los mínimos recomendables para realizar el modelamiento de nichos ecológicos en parientes silvestres de cultivos con importancia económica (Jarvis et al., 2008). De esta forma, en la Tabla 3 se muestra un resumen de los taxones incluidos en el estudio y el número de datos de ocurrencia que serán tenidos en cuenta para los análisis.

A partir de la base de datos 3560 datos de ocurrencia provenientes de 34 taxones, en el software ArcGis 10.0 (ESRI, 2011) y haciendo uso de la herramienta *Extract Multi Values to Points* de la extensión de análisis espacial (“*Spatial Analyst Tools > Extraction > Extract Multi Values to Points*”), fueron extraídos los atributos de los 21 mapas en formato raster con las variables a tener en cuenta en el presente estudio. Finalmente, la base de datos del *shapefile* con los 3560 registros y 25 variables (identificador único, nombre del taxón, latitud, longitud, 19 variables bioclimáticas, altitud y pendiente), fueron exportadas a formato *csv* el cual fue utilizado tanto para el modelamiento de nichos ecológicos como para el proceso de selección de variables.

Tabla 3 – Número de datos de ocurrencia incluidos en el estudio por cada uno de los 34 taxones seleccionados con 10 o más datos.

Taxón de Phaseolus	Ocurrencias
<i>P. acutifolius</i>	29
<i>P. acutifolius_acutifolius</i>	120
<i>P. acutifolius_tenuifolius</i>	154
<i>P. amblyosepalus</i>	10
<i>P. angustissimus</i>	181
<i>P. augusti</i>	40
<i>P. coccineus</i>	342
<i>P. coccineus_coccineus</i>	32
<i>P. costaricensis</i>	35
<i>P. dumosus</i>	124
<i>P. filiformis</i>	290
<i>P. glabellus</i>	15
<i>P. grayanus</i>	48
<i>P. leptostachyus</i>	192
<i>P. leptostachyus_leptostachyus</i>	43
<i>P. lunatus</i>	359
<i>P. maculatus</i>	10
<i>P. maculatus_maculatus</i>	130
<i>P. maculatus_ritensis</i>	105
<i>P. neglectus</i>	10
<i>P. nelsonii</i>	26
<i>P. oligospermus</i>	27
<i>P. pachyrrhizoides</i>	23
<i>P. parvifolius</i>	58
<i>P. parvulus</i>	70
<i>P. pauciflorus</i>	101
<i>P. pedicellatus</i>	59
<i>P. pluriflorus</i>	43
<i>P. polystachios</i>	322
<i>P. talamancensis</i>	11
<i>P. tuerckheimii</i>	18
<i>P. vulgaris</i>	490
<i>P. xanthotrichus</i>	31
<i>P. zimapanensis</i>	12
Total (34 Taxones)	3560

4.2 Análisis exploratorio de datos. Con los datos obtenidos luego de la organización de la información fue realizado un análisis exploratorio de datos. Fue realizado un análisis de dispersión y asociación de las variables utilizadas en el estudio y el análisis de la distribución espacial de la localización de los puntos de colecta de germoplasma. En el caso del análisis de distribución y asociación entre las variables, fueron realizados en R (R Development Core Team, 2012) histogramas de cada variable usada en el modelamiento y gráficos de dispersión entre las mismas, además de ser calculados coeficientes de correlación lineal de Pearson y sus

correspondientes valores P. En términos de la distribución espacial de los puntos de colecta de germoplasma utilizados para el análisis, en ArcGis 10 (ESRI, 2011) fue realizado un análisis exploratorio de datos a nivel espacial que corresponde a la visualización de la localización de los puntos de colecta de los 34 taxones y con la finalidad de determinar que tan concentradas se encontraban las observaciones de cada taxón con relación a la totalidad del área de estudio, fue utilizado el calculo del área del rectángulo más pequeño que incluye a todas las accesiones de cada taxón. Esta área fue calculada con la herramienta la herramienta de la distancia promedio al vecino más próximo (*Average Nearest Neighbor*) del Spatial Statistics (*Spatial Statistics Tools>Analyzing Patterns>Average Nearest Neighbor*) en ArcGIS 10. En este caso, únicamente fue utilizado como índice de concentración o de distribución de los puntos esta área, debido a que el cálculo completo de la concentración mediante esta herramienta es dependiente del número de puntos y el área mínima que los incluye y en este caso este número era diferente para cada taxón (ESRI, 2011; Scott & Janikas, 2010). En la sección de análisis de componentes principales se complementa el análisis espacial en términos de la exploración del comportamiento de cada variable y con relación a cada taxón, por lo cual se excluye de esta sección.

4.3 Análisis de componentes principales. Fue realizado un análisis de componentes principales (ACP) de la base de datos obtenida en la primera sección de este capítulo, la cual contiene información de 19 variables bioclimáticas, la altitud y la pendiente del sitio de colecta de un total de 3560 registros pertenecientes a 34 taxa del genero *Phaseolus* (ver detalles en la sección de generación bases de datos para el modelamiento). El ACP fue realizado haciendo uso del paquete FactoClass (Pardo & Del Campo, 2007) dentro del programa de análisis estadístico R (R Development Core Team, 2012), en donde se obtuvieron las estadísticas de descomposición de la variabilidad y fueron graficadas la participación de cada variable sobre los componentes principales de mayor importancia, lo cual sirvió como insumo para la selección de las variables para el modelamiento. En el Anexo 1– Código en R para la realización del análisis de componentes principales, se puede observar el código de R utilizado para realizar este análisis y graficar los resultados.

4.4 Selección de variables para el modelamiento. Con base en el análisis de componentes principales fueron seleccionadas un grupo de variables a tener en cuenta para el modelamiento en Maxent, para posteriormente comparar los modelos ajustados teniendo en cuenta la totalidad de variables y esta selección. Fueron tenidos en cuenta tres criterios para realizar la selección de las variables, el primero fue la importancia relativa de la variable o participación de la misma en los principales vectores propios, el segundo fue la inclusión de variables relacionadas con todas las condiciones bioclimáticas que buscan representar las variables bioclimáticas de acuerdo a lo planteado por (Nix, 1986) cuando planteó la necesidad de calcular las variables bioclimáticas a partir de la precipitación y temperatura (Busby & Austin, 1990) (Ver marco teórico), y el tercer criterio, fue tratar de eliminar variables redundantes, es decir aquellas que siendo de importancia, estuvieran muy correlacionadas y representarían el mismo tipo de información.

4.5 Modelamiento de nichos ecológicos en Maxent. Fueron modelados los nichos ecológicos en el Aplicativo Maxent versión 3.3.3.k (Phillips et al., 2006), siguiendo los parámetros descritos por Ramirez-Villegas et al. (2010), en donde realizó el modelamiento de nichos ecológicos para los parientes silvestres del frijol con las condiciones bioclimáticas actuales. En este estudio fueron realizados los modelamientos de nicho ecológico tanto para las condiciones actuales y futuras, además fue realizado el modelamiento utilizando la totalidad de las variables bioclimáticas, la altitud y la pendiente del terreno, además de las variables seleccionadas de acuerdo al análisis de componentes principales descrito en el paso anterior.

El modelamiento fue usado en modo “*Batch*” haciendo uso de un archivo construido para tal fin y ejecutado en java. En la Tabla 4 se puede observar el contenido del archivo ejecutado para realizar el modelamiento de 34 taxones de *Phaseolus* con la totalidad de variables del estudio.

Tabla 4 – Contenido de los archivos *Phaseolus_all.bat* y *Phaseolus_sel.bat* corridos en Java para la generación de los modelos de nicho ecológico en condiciones presentes y futuras en Maxent

Archivo	Código
Archivo <i>Phaseolus_all.bat</i> : código en java para generar los modelos de nicho basados en todas las variables del estudio	java -mx1300m -jar C:\maxent.jar -s C:\InputData\SWD_all.csv -e C:\InputData\background.csv -o C:\Output\all -j C:\InputData\pre,C:\InputData\fut -P -J -x nowriteclampgrid nowritemess appendtoresultsfile replicates=10 replicatetype=crossvalidate writebackgroundpredictions nowarnings -a @if errorlevel 1 pause
Archivo <i>Phaseolus_sel.bat</i> : código en java para generar los modelos de nicho basados en 10 variables seleccionadas	java -mx1300m -jar C:\maxent.jar -s C:\InputData\SWD_sel.csv -e C:\InputData\background.csv -o C:\Output\sel -j C:\InputData\pre,C:\InputData\fut -P -J -x nowriteclampgrid nowritemess appendtoresultsfile replicates=10 replicatetype=crossvalidate writebackgroundpredictions nowarnings -a @if errorlevel 1 pause

Es posible describir los parámetros tenidos en cuenta para el modelamiento realizado de acuerdo a la descripción del código presentado en la Tabla 4. Así, el primer parámetro del código “-mx1300m” implica el tamaño en Megabytes usado para que el computador realice el modelamiento, 1300 es capacidad es la definida como límite en computadores bajo el sistema Windows y con varias Gigabytes de memoria para realizar modelamientos en Maxent (Felipe, 2009). La siguiente sección del código “-jar C:\maxent.jar” simplemente invoca a la aplicación de Maxent que se encuentra en este caso en el disco “C:\”. Luego el parámetro “-s C:\InputData\SWD_all.csv”, invoca la localización del archivo de tipo *csv* que contiene la base de datos de las accesiones junto a los datos ambientales o archivo SWD (Samples With Data), construido en ArcGis10 (ESRI, 2011) de acuerdo a la descripción realizada en la sección que describe la organización de bases de datos para el modelamiento, se debe destacar que cuando fue realizado el modelamiento basado en la selección de variables realizada, en el archivo SWD usado se eliminaron los valores de las variables removidas.

Por su parte, la siguiente sección del código “-e C:\InputData\background.csv” invoca al archivo con extensión *csv* que contiene la información de 10000 puntos de pseudo-ausencia, los cuales fueron generados de forma aleatoria para el entrenamiento de los modelos, y sobre el área del rango distribucional del acervo genético del genero *Phaseolus*, esto fue realizado como estrategia para evitar el sobreajuste de los modelos debido la introducción de valores poco probables en las condiciones bioclimáticas de las especies a ser modeladas (Barbet-Massin, Jiguet, Albert, & Thuiller, 2012; Elith et al., 2011). Así, el procedimiento seguido en

ArcGis10(ESRI, 2011) para la generación del background, consistió en tres pasos, en el primero se generó un Buffer alrededor de todos los puntos de ocurrencia para el género *Phaseolus* de 300 Km, dentro del cual en el segundo paso y haciendo uso de la herramienta del ArcToolbox ‘*Create Random Points*’ (ArcToolbox>Data Management Tools>Feature Class>Create Random Points) fueron generados 10000 puntos aleatorios y en tercer lugar con el uso del shapefile generado y los ESRI Grids de las variables bioclimáticas, altitud y pendiente, se utilizó la herramienta *Extract Multi Values to Points* (ArcToolbox>Spatial Analyst Tools>Extraction>Extract Multi Values to Points) para extraer los valores ambientales para los puntos generados y poder generar el archivo background.cvs, este último exportando la base de datos del shapefile generado con el procedimiento anteriormente citado. En la siguiente sección del código ‘*-o C:\Output\all*’ simplemente se da la referencia a la localización de salida de los resultados del modelamiento. La siguiente sección del código ‘*-j C:\InputData\pre,C:\InputData\fut*’ define la localización de las carpetas con la información espacial de las variables bioclimáticas, altitud y pendiente sobre las cuales se va a proyectar los modelos ajustados para determinar la superficie de probabilidad que generan los modelamientos en Maxent, en este caso debido a que los modelos ajustados se van a proyectar a condiciones presentes y del futuro bajo condiciones de cambio climático, se definen dos carpetas que contienen la información de las condiciones actuales y futuras, localizaciones que están separadas por una coma y sin espacios. Se debe destacar que la información espacial contenida en estas carpetas debe estar en formato ASCII Grid, los cuales fueron generados a partir de archivos ESRI Grid de acuerdo al procedimiento presentado en la sección de organización de bases de datos para el modelamiento.

Dentro del código aparecen sentencias que permiten definir algunos parámetros de importancia en el modelamiento, por ejemplo, la letra ‘*P*’, permite graficar como la probabilidad de ocurrencia de la predicción depende de cada variable ambiental. Por su parte la letra ‘*J*’ permite hacer una prueba de Jackknife que establece como es afectada la curva ROC del modelo debido a la presencia o ausencia de cada variable. La letra ‘*x*’ implica que Maxent no va a generar ASCII Grids de las proyecciones por cada repetición del modelo, sino solamente del promedio, desviación estándar, máximo y mínimo valor del total de los modelos por cada taxón. Además las palabras *nowriteclampgrid* y *nowritemess*, hace que el modelo no genere este análisis para cada especie, el cual compara los valores de las variables originales

en la base de datos de análisis y los valores contenidos en las grillas bioclimáticas de las proyecciones.

Finalmente, una sección importante del código es '*replicates=10 replicatetype=crossvalidate writebackgroundpredictions*', las dos primeras sentencias definen el número de replicas que se van a correr del modelo por cada taxón y el tipo de validación a realizar, en este caso fueron definidas 10, a diferencia a lo realizado por (Ramirez-Villegas et al., 2010) que realizó 25 replicas, sin embargo debido a limitaciones de espacio en disco y de procesamiento del computador usado en este estudio fue necesario hacerlo únicamente con 10 replicas. Por su parte, la tercera sentencia hace referencia a la generación de un archivo con las predicciones para los 10000 puntos del background o puntos de pseudo-ausencias, los cuales se utilizarán para hacer las comparaciones de los modelos en la siguiente sección.

4.6 Selección de modelos para el análisis espacial. Para cada taxón fue seleccionado el modelamiento con un juego de variables, es decir entre aquellos que fueron ajustados teniendo en cuenta la totalidad de las variables y los que fueron ajustados con la selección de las mismas. Esta selección, fue realizada teniendo en cuenta dos análisis comparativos, el primero la evaluación estadística de la precisión de los modelos basados en la comparación de los índices AUC, Kappa y TSS(Allouche et al., 2006; Liu, White, & Newell, 2009; Lu, Jia, Lloyd, & Sun, 2012; Thuiller, Lafourcade, Engler, & Araújo, 2009).El segundo análisis, fue realizado a nivel espacial para evaluar la estabilidad en la predicción logística de los modelos proyectados en Maxent, para lo cual fue comparado el porcentaje de área con valores de desviación estándar menores de 0.1(PADM10) en la denominada cobertura de distribución potencial para cada especie (Ramirez-Villegas et al., 2010). A continuación, se describe el procedimiento desarrollado en cada comparación y finalmente el procedimiento mediante el cual se combinaron dichos criterios para la selección de los modelamientos a ser utilizados para el proceso de análisis espacial.

4.6.1 Evaluación estadística de la precisión de los modelos generados en Maxent: Tal y como se mencionó en el marco teórico, existen múltiples índices para evaluar el desempeño de los modelos de distribución de especies (Allouche et al., 2006; Liu et al., 2009; Liu et al.,

2005), en este caso se compararon los promedios de las 10 validaciones cruzadas realizadas en Maxent para cada combinación 'Taxón-'*juego de variables*', para esto fueron utilizados los índices AUC, Kappa y TSS, permitiendo evaluar la precisión de los modelos y seleccionar el mejor modelamiento para cada situación. Este procedimiento fue similar a lo planteado por la plataforma BIOMOD para este tipo de labor (Thuiller et al., 2009) y a los parámetros utilizados entre otros por (Lu et al., 2012). El primer índice usado fue el AUC -Area Under Receive Operation Curve- (Phillips et al., 2006; Ramirez-Villegas et al., 2010), el cual es un índice que es independiente del umbral seleccionado para la generación de los mapas binarios de predicción, pero que interpretado de forma independiente puede llevar a conclusiones erradas (Lobo et al., 2008) y los índices kappa (Guisan & Thuiller, 2005; Manel et al., 2001; Monserud & Leemans, 1992) y TSS-True Skill Statistic (Allouche et al., 2006; Loyola, Lemes, Faleiro, Trindade-Filho, & Machado, 2012; Lu et al., 2012), los cuales tienen en cuenta la sensibilidad y especificidad de los modelamientos y el último adicionalmente es independiente de la prevalencia de predicción, permitiendo evaluar la precisión de los modelos ajustados mediante la utilización de todas o la selección de variables utilizada.

Para la realización de las comparaciones, se utilizaron los resultados generados por Maxent en los modelamientos ajustados con el juego completo y la selección de variables realizada. Así, en cada modelamiento fue extraído del archivo *maxentResults.csv* el AUC de los 10 modelos ajustados para cada taxón. En el caso de los estadísticos Kappa y TSS, los cuales evalúan la precisión de los modelos binarios (presencia-ausencia) de nicho ecológico, del mismo archivo fue extraído el valor umbral para los 10 modelos ajustados en cada uno de los 34 taxones. El valor umbral seleccionado fue aquel que corresponde al punto en la curva ROC más cercano a la esquina superior izquierda (Ramirez-Villegas et al., 2010), el cual es el punto que maximiza la sensibilidad y especificidad del modelo y ha demostrado ser el de mayor utilidad en modelamientos de nicho ecológico con puntos de presencia y pseudo-ausencias (Hu & Jiang, 2011; Liu et al., 2005; Lu et al., 2012). Adicionalmente, fue necesario extraer los valores de las predicciones logísticas para los 10000 puntos del background o puntos de pseudo-ausencias y para las muestras o datos de prueba usados en cada modelamiento. Estos valores fueron extraídos de los 340 archivos en formato *csv* generados por Maxent para los puntos del background y los 340 para las muestras de prueba, los cuales son nombrados por el aplicativo como *Taxon_rep_backgroundPredictions.csv* y *Taxon_rep_samplePredictions.csv*,

respectivamente, donde en el *Taxon* se indica el nombre de cada taxón y en *rep* un número del 0 al 9 indicando el número del modelamiento. Estos valores fueron extraídos haciendo uso de un script en R que se puede observar en el Anexo 2.

Luego de haber extraído las predicciones logísticas y los umbrales de acuerdo al proceso previamente descrito, en una hoja de trabajo del Excel® fue calculada la matriz de confusión, la cual es utilizada para realizar el cálculo de los índices kappa y TSS (Allouche et al., 2006; Liu, Berry, Dawson, & Pearson, 2005), dicha matriz de confusión se compone de cuatro números que van identificados de la letra *a* hasta la *d*. La letra *a* cuantifica el número de verdaderos positivos que estima el modelo, que en este caso corresponde a la cantidad de muestras de prueba que fueron predichas con probabilidades mayores al valor umbral. Con la letra *b* se cuantifica el número de falsos positivos, que en este caso están representados por las muestras del background o pseudo-ausencias predichas como ciertas o con probabilidades logísticas mayores al umbral seleccionado. La letra *c*, corresponde a los falsos negativos o número de muestras de la muestra de prueba que fueron predichos con probabilidades menores al valor umbral y finalmente la letra *d* corresponde al número de verdaderos negativos o la cantidad de pseudo-ausencias predichas como negativas o por debajo del umbral. Adicionalmente los índices usados hacen uso de un valor *n*, que corresponde a la sumatoria de los valores anteriormente citados (Allouche et al., 2006), a partir de estos valores y haciendo uso de las formulas que se ilustran en el marco teórico y resumidas por (Allouche et al., 2006), fueron calculados los valores de Kappa y TSS para cada modelo. Finalmente y luego de haber extraído los valores de AUC y calculado los estadísticos Kappa y TSS a partir de la matriz de confusión de cada modelo, fueron comparados estadísticamente los promedios de los índices en cada modelamiento mediante la realización de una prueba de comparación de medias de t-student en R (R Development Core Team, 2012).

4.6.2 Análisis de la estabilidad en la predicción logística realizada por Maxent: En este caso se tuvo en cuenta el criterio de estabilidad en la predicción de un nicho ecológico manejado por (Ramirez-Villegas et al., 2010), que utilizó la proporción de área con altas desviaciones estándar en la predicción logística de modelos ajustados mediante validación cruzada para taxones del genero *Phaseolus*. De esta forma, en el presente estudio se tuvo en cuenta como criterio, la proporción del área de cobertura de distribución potencial con desviaciones

estándar menores a 0.10 (PADM10), considerando un modelo estable aquel con proporciones mayores al 90%. Por su parte, la cobertura de distribución potencial del taxón, hace referencia al área con predicción logística mayor al valor umbral seleccionado y que se encuentra dentro del área nativa de cada taxón (Ramirez-Villegas et al., 2010).

El procedimiento de cálculo seguido se describe a continuación:

- Conversión de las superficies de probabilidad y desviación estándar a formato ESRI Grid: Fue necesaria la conversión de los mapas generados por Maxent en formato ASCII Grid a formato ESRI Grid para realizar los análisis espaciales en ArcGIS10. Este procedimiento fue realizado haciendo uso de la herramienta del ArcToolBox *ASCII to Raster (Conversion Tools>To Raster>ASCII to Raster)* y para automatizar el proceso se desarrollo un Script Phyton 2.6 que se puede observar en el Anexo 3-A.
- Corte de las superficies de probabilidad logística en el área nativa de cada Taxón: Se realizó el corte de la superficie de probabilidad logística generada por Maxent en el área nativa de cada taxón. Para esto se multiplicaron las superficies promediadas de las predicciones logísticas y la superficie que representa el área nativa de cada taxón. Esta área fue tomada de (Bioversity International et al., 2009) y consiste en mapas en formato grid con un valor de 1 en las zonas reportadas en la pagina en mención, las cuales a su vez fueron producto de revisión bibliográfica y consultas a expertos. En el Anexo 3-B se muestra un Script de Phyton 2.6 usado para automatizar el proceso.
- Generación de los mapas binarios de predicción en el área nativa de cada Taxón: Fueron generados mapas binarios de predicción (1-presencia, NoData-ausencia), para cada uno de las 68 combinaciones Taxón-juego de variables seleccionadas (34*2) en el área nativa delimitada anteriormente. Esto fue realizado a partir de la reclasificación del mapa de predicciones logísticas promedio, para lo que se asignó un valor “NoData” (Ausencia) para las celdas con valores menores al umbral seleccionado y de 1(Presencia) para los valores promedio mayores al mismo valor. Para la realización de estas reclasificaciones fue generado un Script en Phyton 2.6, el cual utiliza la herramienta *Reclassify* del *Spatial Analyst* para reclasificar las predicciones logísticas con base en los valores umbrales seleccionados. Para automatizar el proceso fue desarrollado un Script en Phyton 2.6 que se muestra en el Anexo 3-C.

- Generación de las superficies con la desviación estándar en el área de la cobertura de distribución potencial: Luego, fue realizada una multiplicación entre el mapa binario generado en el paso anterior y el mapa de la desviación estándar en cada una de las 64 combinaciones *Taxón-juego de variables*. De forma similar al paso seguido anteriormente, fue desarrollado para esta operación un Script en Phyton 2.6 para realizar esta operación el cual se puede observar en el Anexo 3-D.
- Calculo de la PADM10 por cada combinación '*taxón-juego de variables*' utilizadas en el modelamiento: Finalmente, fue calculado la proporción del área de la cobertura de distribución potencial con desviación estándar menor a 0.1 (PADM10). Para esto se realizó una reclasificación de la superficie con la desviación estándar descrita en el paso anterior, en donde valores menores a 0.1 se reclasificaron con un 1 y los demás con un 0. Luego se calculo la proporción de pixeles con valores de 1. El proceso se automatizó mediante un Sript que s emuestra en el Anexo 3-E.

4.6.3 Criterios tenidos en cuenta para la selección del modelamiento a ser seleccionado: De acuerdo a la evaluación conjunta de los criterios de precisión estadística y estabilidad espacial, fue seleccionado el modelo de mejor desempeño ra acada taxón. Para tal fin, fue calculado el siguiente puntaje producto de la ponderación de los criterios estadísticos y espaciales descritos anteriormente:

$$PEM_i: 0.5*(A_i+B_i+C_i) + 0.5 (D_i)$$

Donde PEM es el puntaje de evaluación del modelamiento por cada combinación *Taxón-juego de variables*, A, B y C son los puntajes para la evaluación de la precisión estadística de los índices AUC, Kappa y TSS, respectivamente. Así, cada una de las letras puede tomar los valores 2, 1 y 0, donde 2 representa un promedio estadísticamente superior en la comparación realizada entre los promedios de los índices en cada combinación *Taxón-juego de variables*, 1 representa un promedio estadísticamente similar y 0 un promedio estadísticamente inferior. Sin embargo, si el promedio del índice presenta valores inferiores a los indicados para tener algún grado de precisión estos tomarán un valor de 0, sin importar los resultados de la prueba de comparación de promedios. Es decir, para los valores de AUC, de presentarse por debajo de 0.7 (Manel et al., 2001) y para los de Kappa y TSS por debajo de 0.4 (Monserud & Leemans,

1992).

En el caso de la letra D, esta indica una puntuación en función de la proporción de área de la cobertura de distribución potencial con desviaciones estándar menores a 0.1 (PADM10). Así, D toma valores entre 0 a 6, siendo 0 para los modelos con PADM10 menor al 90%, es decir para los modelos inestables, 3 para los modelos con PADM10 entre 90 y 92.5%, 4 para aquellos con PADM10 entre 92.5 a 95%, 5 para los que tienen valores de PADM10 entre 95 a 97.5% y 6 para los que tienen valores de PADM10 entre 97.5 a 100%, asegurándose de esta forma que los criterios estadísticos y espacial tuvieran el mismo peso.

4.7 Proceso de análisis espacial. El proceso de análisis espacial fue realizado en ArcGis10 (ESRI, 2011) y mediante la utilización de Script diseñados en Python 2.6 (Python Software Foundation, 2008) para automatizar algunos de los análisis. Este análisis fue realizado con las predicciones seleccionadas para cada taxón, de acuerdo a los procedimientos desarrollados en la sección anterior en donde se evaluó la precisión estadística de los modelamientos y la estabilidad en la predicción logística promedio de los mismos.

De esta forma, el proceso de análisis espacial seguido consta de tres grandes pasos, en el primero se presentan las comparaciones de los nichos ecológicos en las condiciones presentes y de cambio climático (2020-2050), bajo dos escenarios de capacidad migratoria de los taxones. Estos escenarios fueron sin capacidad de migración y con capacidad migratoria, para lo cual fue utilizado un máximo de 300 Km, distancia que ha sido considerada como la capacidad de migración para la mayoría de los parientes silvestres de cultivos de interés para un periodo de 50 años (Jarvis et al., 2008). En el segundo paso fue realizado un análisis comparativo de la fragmentación de los nichos ecológicos bajo los dos escenarios de migración de las especies, para lo cual fue utilizada extensión *Patch Grid Analyst* del paquete *Patch Grid* (Rempel et al., 2012) desarrollado para ArcGis10, haciendo un análisis comparativo del número de parches (NP) y el tamaño promedio de los parches (TPP). El tercer paso del proceso de análisis espacial fue la determinación de la riqueza de especies en condiciones actuales y la comparación con la riqueza en condiciones futuras bajo cada escenario de capacidad migratoria. En este caso, se identificaron las zonas en donde hay pérdida, conservación o ganancia en la biodiversidad en lo que se refiere a taxones del género

Phaseolus. A continuación se describen las operaciones de análisis espacial y procedimientos seguidos en cada paso.

4.7.1 Pasos de análisis espacial seguido para la presentación y comparación de nichos ecológicos: De acuerdo al procedimiento planteado en la sección de análisis de la estabilidad en la predicción logística realizada por Maxent, fueron desarrollados Script en Python 2.6 para realizar la conversión de los mapas generados por Maxent en formato ASCII Grid a formato ESRI Grid (Anexo 3-A). Igualmente, en la misma sección, se describió la forma en la que se generó los mapas binarios de predicción en las condiciones actuales, desarrollándose para eso un Script en Python 2.6 (Anexo 3-C). En esta sección, son utilizados los mapas binarios generados para las condiciones actuales y futuras y haciendo uso de los modelos de mejor ajuste en cada uno de los taxones. A continuación se describe el procedimiento seguido.

- *Generación de mapas binarios de predicción en las condiciones de cambio climático:* En este caso y a partir de las predicciones con las probabilidades logísticas en condiciones de cambio climático generadas por Maxent y los umbrales seleccionados para cada taxón, fueron generados los mapas binarios de predicción (1-presencia, NoData-ausencia). Dichos mapas contienen un 1 en los pixeles con probabilidades logísticas mayores al valor umbral en toda el área del estudio. Para la realización de esta operación fue realizada una reclasificación mediante la función *Reclassify* del *Spatial Analyst* y fue reutilizado el Script de Python 2.6 que se muestra en el Anexo 3-C, el cual fue utilizado anteriormente para hacer los mapas binarios en las condiciones actuales.
- *Generación de los mapas binarios de predicción en condiciones de cambio climático y el escenario de no migración:* En este caso, fue realizada la superposición espacial de los mapas binarios de las condiciones actuales y en las condiciones de cambio climático a nivel general desarrollados en el paso anterior. Para la realización de la superposición espacial en formato *raster* se realizó una multiplicación de los dos *raster* mencionados haciendo uso del *Map Calculator* de la extensión de análisis espacial de ArcGis10, en este caso, las áreas que conforman el nicho ecológico para cada taxón en condiciones de cambio climático y bajo el escenario de no migración tienen un valor de 1 y las restantes

un valor de NoData. Para automatizar esta operación fue desarrollado un Script en Python 2.6 que se muestra en el Anexo 4-A.

- *Generación de los mapas binarios de predicción en condiciones de cambio climático y el escenario de migración:* Para este cálculo fueron realizadas tres operaciones, en la primera a partir del mapa binario de las condiciones actuales en formato *raster* para cada taxón fue calculado un *raster* con la distancia euclídea hasta una distancia de 300 Km (2.69 DD, aproximadamente a la altura del Ecuador), de esta forma se calcula la distancia hasta la cual se estima pueden migrar los individuos de cada taxón con relación a la predicción de las condiciones presentes. En el segundo paso, se realiza la superposición espacial entre los *raster* de distancia euclídea calculados anteriormente y el mapa de predicción logística de las condiciones futuras a nivel general obtenido en el primer análisis de esta sección para cada taxón. Finalmente, se hace una reclasificación de los resultados a un valor de 1, para obtener el mapa binario en las condiciones de cambio climático y el escenario de movilidad tenido en cuenta en el presente estudio y para cada taxón. Para realizar este proceso fue desarrollado un Script en Python 2.6 que se muestra en el Anexo 4-B.
- *Proceso de superposición espacial y comparación de áreas de predicción bajo cada escenario:* Para realizar el análisis de comparación espacial entre las condiciones actuales y las de cambio climático fue realizado un proceso de análisis espacial que consistió en aplicar la siguiente formula en el *map calculator* de la extensión de análisis espacial de arcGis10:

$$\text{Superposición} = (\text{raster binario actual}_i) + 2 * (\text{raster binario futuro bajo el escenario de migración}_i)$$

Así, los pixeles con valores de 1 representan las áreas donde se predijo el nicho ecológico únicamente para las condiciones actuales en cada taxón o áreas de pérdida de nicho bajo el escenario de cambio climático, Los pixeles señalados con 2, son las áreas donde se predijo el nicho ecológico únicamente para las condiciones futuras o áreas donde el nicho ecológico tiene la posibilidad de incrementar su área siempre y cuando pueda migrar y los pixeles con un valor de 3, son las áreas donde se predijo los nichos ecológicos para ambas condiciones, estas áreas son las que corresponden al escenario de no migración bajo el

escenario de cambio climático. Este proceso de análisis espacial fue automatizado mediante la implementación de un Script en Python 2.6 que se muestra en el Anexo 4-C. En este estudio, se asume que las comunidades de individuos de cada taxón bajo condiciones de cambio climático pueden tener dos respuestas o se pueden presentar dos escenarios en términos de su capacidad migratoria, en la primera bajo el escenario en que no puedan migrar y establecerse en nuevos nichos con condiciones favorables, mantendrán o perderán área con relación a su nicho actual, por lo cual el área del nicho bajo condiciones de cambio climático y sin capacidad de migración serán menores o iguales al área actual. En el segundo escenario de migración bajo condiciones de cambio climático, los taxones tendrán la capacidad de establecerse en nuevos nichos con condiciones favorables, por lo cual habría nuevas áreas que tienen la probabilidad de ser ocupadas por los taxones, aunque hayan perdido parte o todo su nicho actual. A continuación se presenta el proceso de cálculo para determinar cada área:

- A. Área del nicho ecológico en condiciones actuales (Áreas con número 1 y 2 en cada raster)
 - B. Área del nicho ecológico bajo condiciones de cambio climático y sin capacidad de migración (Áreas con número 3 en cada raster)
 - C. Porcentaje de área del nicho que se pierde bajo condiciones de cambio climático ($[(B/A)*100]$)
 - D. Área del nicho ecológico en condiciones de cambio climático y el escenario de migración (Áreas con número 2 y 3 en cada raster)
 - E. Razón entre el área del nicho en condiciones de cambio climático y migración y el nicho en el presente ((D/A))
 - F. Porcentaje del nicho en condiciones de cambio climático que corresponde a nuevas áreas con posibilidad de ocupar ($[(\text{Áreas de cada raster con numero } 2/D)*100]$).
- Adicionalmente, para visualizar las probabilidades logísticas en cada caso, cada mapa binario fue multiplicado por el mapa de probabilidad logística con lo que se consiguió tener los mapas de probabilidad en la denominada cobertura de probabilidad potencial para cada nicho ecológico.

4.7.2 Pasos de análisis espacial seguido para el análisis de fragmentación de nichos: Este análisis fue realizado mediante la utilización del paquete *Patch Analyst* Versión 5.1 desarrollado para ArcGis10 (Rempel et al., 2012). Luego de realizar la instalación del programa descargado del sitio web <http://www.cnfer.on.ca/SEP/patchanalyst/>, fue realizado el análisis de parches, proceso que consta de tres partes. En la primera se realizó la proyección a coordenadas planas de los mapas binarios generados en el paso de análisis espacial anterior, esta proyección fue realizada mediante la herramienta *Project Raster* (*ArcToolBox>Data Management Tools>Projections and Transformations>Raster>Project Raster*) a la proyección *WGS84_World_Mercator*, con un tamaño de pixel de 5000 metros y técnica de remuestreo *NEAREST*, Para la realización de este proceso fue desarrollado un Script en Python 2.6 que se muestra en el Anexo 4-D.

En segundo lugar, fue utilizada la herramienta *Patch Grid Analyst* del paquete instalado para el análisis de parches. Para realizar el análisis de parches, se adicionaron a ArcGis 10 los *raster* con los mapas binarios de cada taxón y predicción a ser analizados, luego en el menú *Patch Grid* se abre la herramienta *Spatial Statistics (FragStats Interface)*, se configuran las siguientes opciones:

- ✓ **Layers:** Se selecciona un *raster* de los adicionados para analizar.
- ✓ **Class:** el atributo a analizar (Value)
- ✓ **Analyst by:** si se va a realizar para todo el área de trabajo (Landscape) y/o para cada valor del Class por aparte (Class), en este caso por tener solamente mapas binarios solamente es necesario hacerlo para Landscape.
- ✓ **Output Table:** Se introduce el nombre de la tabla en formato *dbf* y la ruta donde se almacenarán los resultados del análisis. La primera vez que se corre el programa, la opción *Append* o *Overwrite*, están deshabilitadas, para los siguientes análisis se selecciona la opción *Append* para que los resultados se almacenen en la misma tabla.
- ✓ **Options:** Finalmente, son seleccionados los estadísticos que se deseen obtener del análisis.

En tercer lugar, fueron resumidos y presentados los resultados del análisis en cuanto a los estadísticos solicitados. Finalmente, se debe destacar que los estadísticos generados son calculados por el paquete *Patch Analyst 5.1* a partir del formato *raster* mediante el uso de los algoritmos planteados por (McGarigal & Marks, 1995), para lo cual el programa genera

archivos temporales que almacenan un análisis del valor de los píxeles vecinos que tienen un mismo valor y hacen un análisis similar al de una topología en el modelo vectorial.

4.7.3 Pasos de análisis espacial seguido para el análisis de riqueza: En este caso y haciendo uso de los *raster* binarios generados en el primer paso del presente análisis, fueron calculados tres *raster* con los datos de riqueza de especies del género *Phaseolus* de acuerdo a las predicciones realizadas con los mapas de predicción de las condiciones presentes, y futuras bajo los escenarios de no migración y migración tenidos en cuenta en este estudio. Estos mapas de riqueza fueron generados haciendo uso del *Map Calculator* de la extensión de análisis espacial en ArcGis10. Igualmente, a partir de operaciones de álgebra de mapas, fueron cuantificadas las zonas con pérdidas y ganancias en la biodiversidad de acuerdo a la comparación realizada de las condiciones presentes y futuras bajo los dos escenarios de movilidad.

5. RESULTADOS

5.1 Análisis exploratorio de datos. A continuación se presenta el análisis de asociación lineal de variables ambientales y el análisis exploratorio de datos a nivel espacial de los puntos de ocurrencia de los 34 taxones incluidos en el estudio:

5.1.1 Análisis de asociación de variables ambientales: En la Figura 2, muestra una matriz de gráficos de dispersión de cada par de variables analizadas en el presente estudio, junto a los histogramas de cada variable en la diagonal principal. Así, por ejemplo, en la primera casilla de la matriz se muestra el histograma de la variable bio1 (Rango de temperatura anual en °C) y en la primera fila los gráficos de dispersión teniendo en los ejes Y a esta variable y en los ejes X a las demás comenzando por la bio2 hasta la altitud. De forma similar ocurre con las columnas de la matriz, además, en la parte superior e inferior se muestra una aproximación al rango de variación inter cuartil de cada variable, dando una idea de la variación de cada una de ellas. Finalmente, y para dar facilidad a la visualización del gráfico, sombreadas con un recuadro de color rojo se encuentran las variables asociadas con la temperatura y con uno de color azul las asociadas a la precipitación. En la Tabla 5, se muestra una matriz que resume abajo de la diagonal principal los coeficientes de determinación de Pearson de la correlación lineal de cada par de variables, mientras en la parte superior muestra la correspondiente significancia estadística de dicha correlación.

Al observar el Figura 2 y la Tabla 5, se hace evidente como algunas de las variables bioclimáticas incluidas en el estudio tienen una fuerte correlación lineal, especialmente entre variables asociadas a la temperatura o entre las asociadas a la precipitación, esto como consecuencia lógica en varios casos de su forma de calculo ó por ejemplo las correlaciones entre variables que miden el mismo parámetro en periodos de tiempo diferentes. Por ejemplo, la bio3, tiene una fuerte correlación negativa con las bio2 y especialmente con la bio7 ($r^2 = -0.36$ y -0.90 , para la bio3 y 7 respectivamente, ambas con $p < 0.001$) ó en el caso de las correlaciones positivas entre las bio11 y bio6 ($r^2 = 0.98$, $p < 0.01$) o entre las bio13 y 16 (0.99 , $p < 0.01$).

Comparando las correlaciones y el patrón de puntos generado entre variables que representan diferentes condiciones bioclimáticas (Tabla 5), se hace evidente que existen asociaciones de variables que aunque estadísticamente son significativas, la dispersión de los datos permite entender que su relación es más compleja que una simple correlación lineal. Por ejemplo entre las bio1 que representa una condición promedio anual con las bio 2, 3 y 4 que representan una condición bioclimática de estacionalidad intra-anual y la dispersión de los datos aparenta tener un comportamiento de mayor complejidad.

Por otra parte, observando los histogramas de distribución de las variables (Figura 2, diagonal principal), las variables asociadas a la temperatura que tienen una distribución aproximadamente normal son las bio1, bio5, bio6, bio8 y bio10, mientras que las que tienen una distribución bimodal son las bio3, bio4 y bio7 y finalmente las que tienen una distribución con algún sesgo son las bio2, bio8, bio9 y bio11. Por su parte, las variables asociadas a la precipitación tuvieron una distribución unimodal pero con sesgo hacia la derecha, a excepción de la bio15 que representa la variación en la precipitación. Esto indica que las variables asociadas a la precipitación de esta distribución tienen valores atípicos altos o algunos sitios con alta precipitación, mientras en términos de la temperatura las variables tienen una alta variabilidad. Estos resultados, como se discutirá más adelante implican por un lado entender la necesidad de usar aproximaciones de análisis multivariadas para entender la variabilidad y comportamiento de las condiciones bioclimáticas y por el otro analizar los mismos resultados del análisis multivariado.

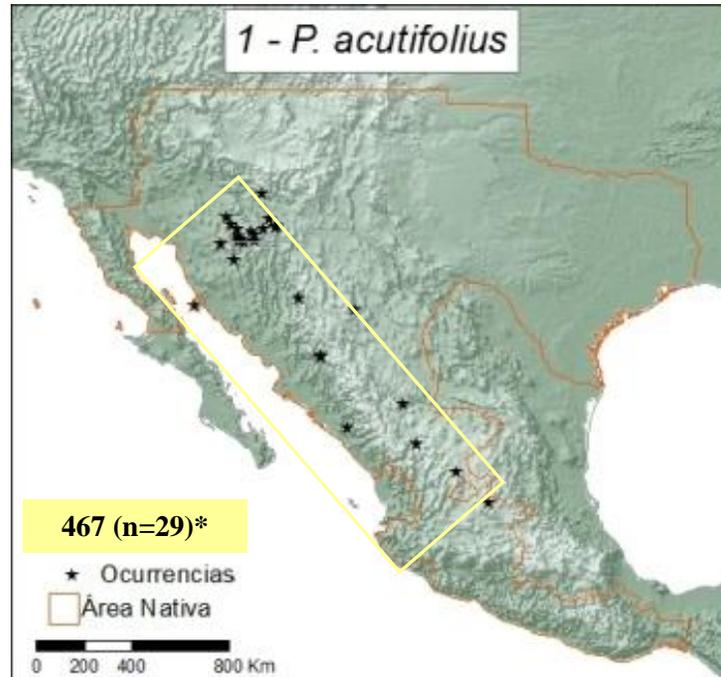
el número de ocurrencias (n). De esta forma, 28 taxones tienen una distribución espacial principalmente a lo largo de la sierra madre, desde el Sur de los Estados Unidos en el Estado de Arizona, hasta la extensión de esta cadena montañosa en Centro América. Únicamente los taxones *P. augusti* y *P. pachyrrhizoides* (Figura 3-6 y Figura 3-3, respectivamente) tienen una distribución espacial sobre los Andes, el primero desde Ecuador hasta el Norte de Argentina y el segundo en Perú. Finalmente, los taxones *P. coccineus*, *P. dumosus*, *P. lunatus* y *P. vulgaris* (Figura 3-7, -10, -16 y -32, respectivamente), es decir, los taxones con formas cultivadas tienen una distribución generalizada desde el Sur de México, hasta la Argentina, reflejándose esta distribución en que las áreas de los rectángulos de menor tamaño que contienen todos los puntos son los mayores con 25040, 9289, 29902 y 40534 miles de km² para *P. coccineus*, *P. dumosus*, *P. lunatus* y *P. vulgaris*, aunque no necesariamente estos sean los taxones con mayor número de puntos de ocurrencia, como se mostró en la Tabla 3.

Por su parte, los demás taxones tienen distribuciones más localizadas, lo cual se hace evidente por las áreas de los rectángulos mencionados anteriormente, las cuales no superan los 5000 miles de km², a excepción de *P. polistachios* para el que el rectángulo en mención mide 3674 miles de km², distribuidos exclusivamente en la parte sur occidental de los Estados Unidos (Figura 1R29). Es así como se presentan distribuciones espaciales desde el Sur de los Estados Unidos y pasando por varios estados mexicanos hasta Centro América en casos como los de *P. acutifolius acutifolius*, *P. acutifolius tenuifolius*, *P. coccineus coccineus*, *P. leptostachyus*, *P. leptostachyus leptostachyus*, *P. maculatus maculatus*, *P. parvifolius* y *P. pedicelatus* (Figura 3-2, -3, -8, -14, -15, -18, -24 y -27, respectivamente), los cuales a excepción del primero tienen áreas del rectángulo que incluye sus ocurrencias con valores mayores a los 1000 miles de km². Otras distribuciones como es el caso de *P. filiformis* y *P. grayanus* tienen distribuciones con rectángulos mayores a 1000 miles de km², sin embargo su rango de distribución incluye ocurrencias en la Sierra Madre Occidental, especialmente al norte de ella donde el ambiente es más bajo y seco y zonas bajas en el mismo rango latitudinal, en Nuevo México y la altillanura mexicana en el caso de *P. filiformis* (Figura 3-11) y en la altillanura en el caso de *P. grayanus* (Figura 3-13).

Finalmente, los demás taxones tienen distribuciones más localizadas que no superan los 1000 miles de km² en términos del rectángulo del que se trata en este apartado, lo que muestra el nivel de concentración de las ocurrencias. En este caso, los taxones de mayor concentración

concentrados en algunas partes de México son *P. amblyosepalus*, *P. glabellus* y *P. neglectus* (Figura 3-4, -12 y -20, respectivamente) y por fuera de México en Costa Rica son *P. costaricensis* y *P. talamancensis* (Figura 3-9 y -30), todos los menos de 100 mil km² en términos de las áreas de los mínimos rectángulos que incluyen todas las ocurrencias. Estos niveles de concentración en combinación con la cantidad de puntos de cada taxón pueden ser un indicativo de la cantidad de inercia que cada taxón puede recoger en términos de las diferencias en las condiciones bioclimáticas, lo cual se discutirá más adelante.

Figura 3– Distribución espacial de los datos de ocurrencia de 34 especies del género *Phaseolus*, su área nativa y el área del rectángulo más pequeño que contiene todos los datos de ocurrencia.



* Área en miles de km² del rectángulo más pequeño que contiene todos los puntos de ocurrencia del Táxon y número de ocurrencias (n) (ejemplo: Área del rectángulo amarillo para *P.acutifolius*)

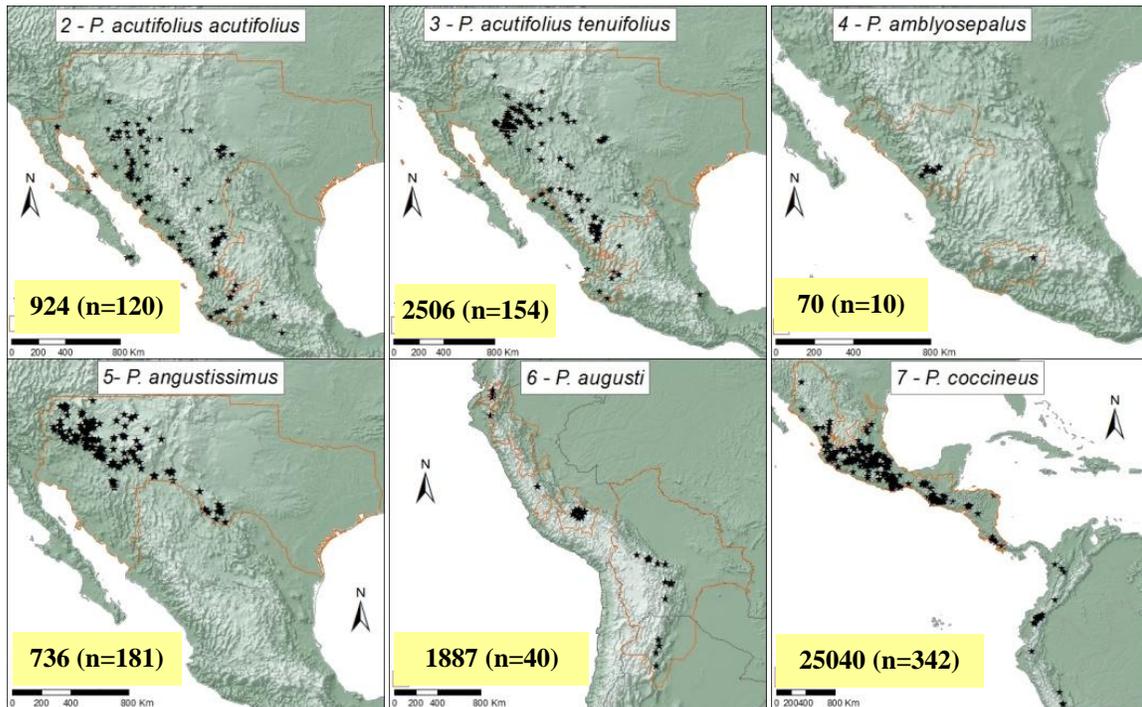


Figura 3 (Continuación)

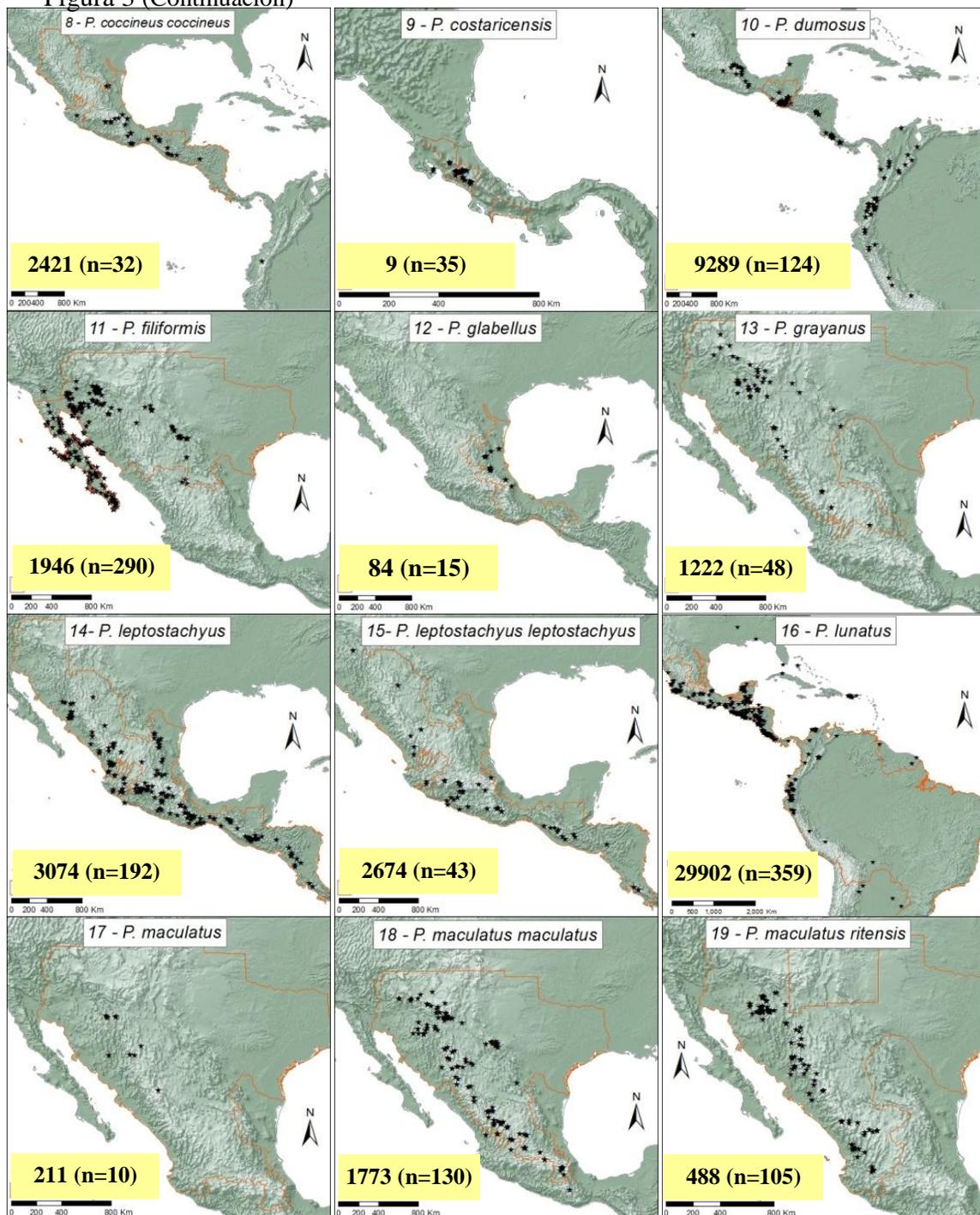
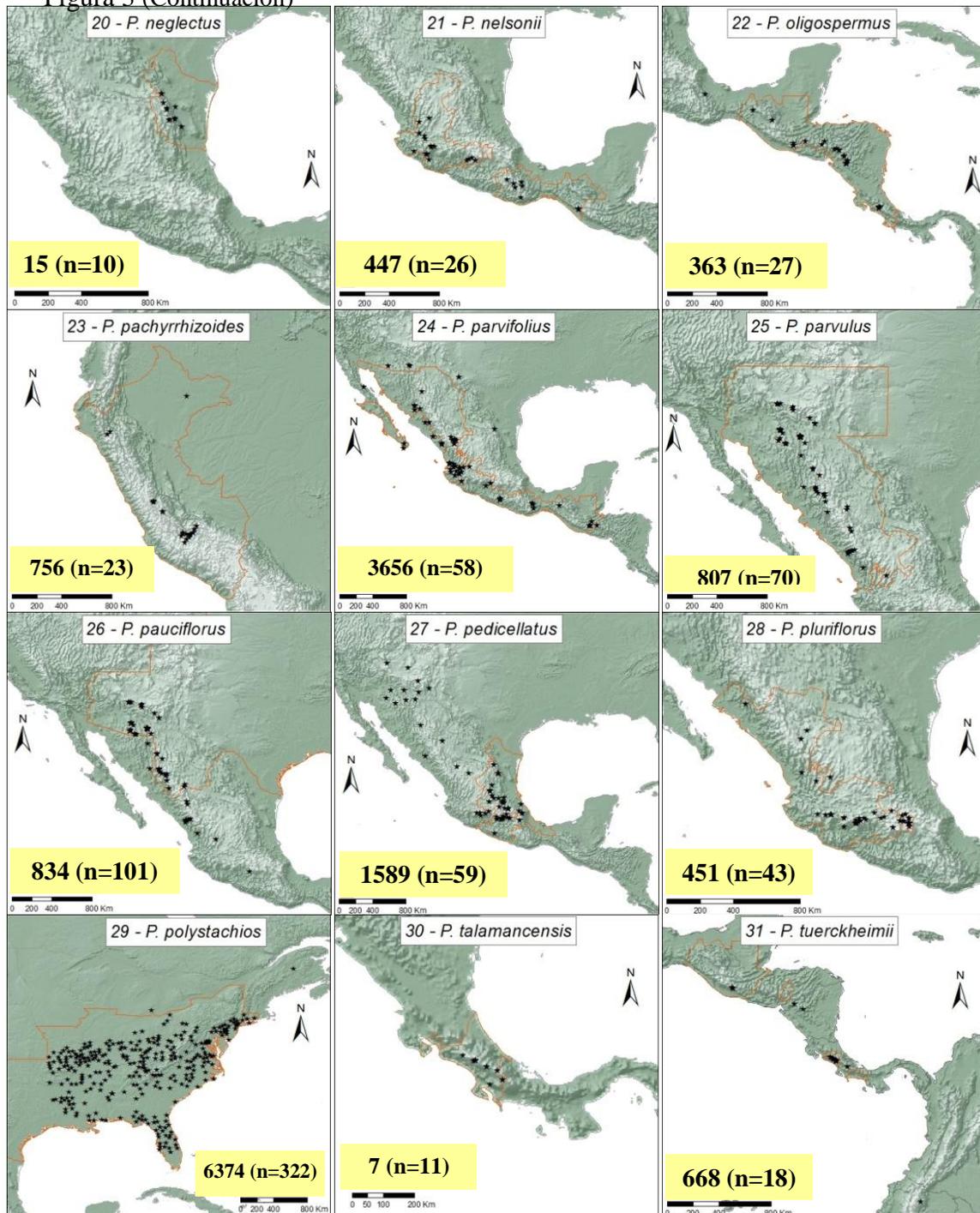
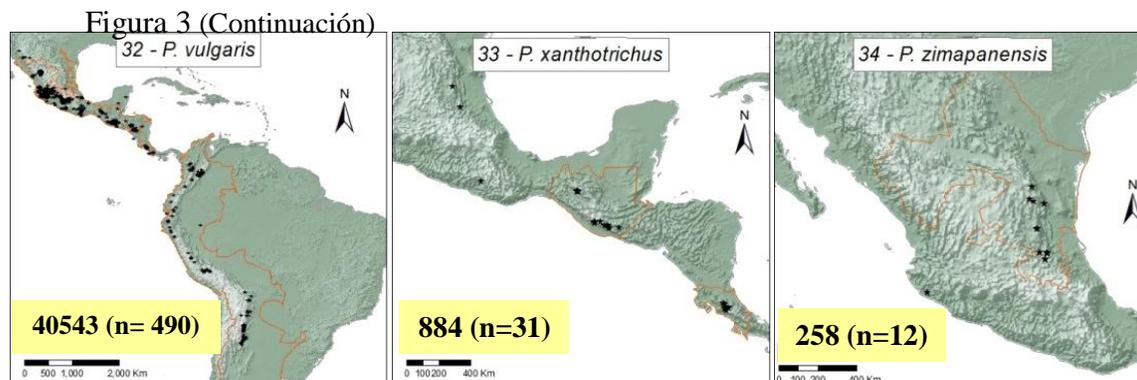


Figura 3 (Continuación)





5.2 Análisis de componentes principales. La Tabla 6 muestra los valores propios o magnitud de cada vector del ACP, junto al porcentaje de variabilidad explicada y acumulada por cada uno de ellos. En dicha tabla, se puede observar que es posible representar de forma adecuada la variabilidad de los datos con los primeros 4 componentes o vectores propios, con los cuales se representa un 86.5% de la variabilidad total. Este número de vectores fue seleccionado debido a que su magnitud es mayor o igual a la unidad y permiten explicar una gran parte de la variabilidad presente en los datos (Jolliffe, 2002). Así, los valores propios fueron de 7.8 para el primer vector, 5.2 para el segundo, 4.1 para el tercero y 1.0 para el cuarto, siendo sin duda los tres primeros vectores los de mayor importancia, llegando a explicar más de 80% de la variabilidad total.

Tabla 6 – Valores propios, porcentaje de variabilidad explicada y acumulada de acuerdo al ACP realizado de las variables bioclimáticas, altura sobre el nivel del mar y pendiente del sitio de colecta para 34 taxones del genero *Phaseolus*.

Componente	Valor propio	Porcentaje	Porcentaje acumulado
1	7.8	37.0	37.0
2	5.2	25.0	62.0
3	4.1	19.5	81.5
4	1.0	5.0	86.5
5	0.8	3.7	90.2
6	0.6	2.8	93.0
7	0.4	2.0	95.0
8	0.3	1.6	96.6
9	0.3	1.2	97.8
10	0.2	0.8	98.6
11	0.2	0.7	99.3
12	0.1	0.3	99.7
13	0.0	0.1	99.8
14	0.0	0.1	99.9
15	0.0	0.0	99.9
16	0.0	0.0	99.9
17	0.0	0.0	100.0
18	0.0	0.0	100.0
19	0.0	0.0	100.0
20	0.0	0.0	100.0

En la Tabla 7, se muestra la composición de los 4 primeros vectores propios y las correlaciones entre cada vector y la variable original. En el primer vector las variables de mayor correlación y participación ($r^2 > 0.8$) fueron en dirección positiva las bio7 y 4 (Rango anual de temperatura y estacionalidad de la temperatura), las cuales están relacionadas con la variación de temperatura anual, mientras las de mayor correlación negativa fueron las bio13, 16, 3, 6, 12 y 11 (Precipitación del mes más húmedo, precipitación del trimestre más húmedo, precipitación anual, isothermalidad, temperatura mínima del mes más frío, precipitación anual y temperatura media del trimestre más frío, respectivamente), estando las variables del sentido positivo del eje relacionadas con la variación intra-estacional de la temperatura y en el sentido negativo del eje se encuentran variables relacionadas con la precipitación de los periodos más húmedos (bio 13 y 16) y la total (bio 12), además de variables de temperatura de los periodos fríos del año (bio 6 y 11), a excepción de la bio3, la cual es inversamente proporcional a la bio7 de acuerdo a su cálculo. En el segundo vector, las variables con mayor participación tienen una correlación positiva con el vector ($r^2 > 0.8$) y están relacionadas con la temperatura, estas fueron las bio 1, 9, 8, 10 y 5 (Temperatura media anual, temperatura media del trimestre más seco, más húmedo, del más caliente y la máxima temperatura del mes más caliente, respectivamente). Por su parte en el tercer vector, las mayores correlaciones positivas ($r^2 > 0.7$) fueron para variables relacionadas con la precipitación de los periodos más secos y fueron las bio17 y 14 (precipitación del trimestre y del mes más seco, respectivamente), mientras en sentido negativo se encontró a la altitud ($R^2 = 0.81$), como las variables de mayor influencia en el vector. Finalmente, para el cuarto vector las correlaciones de las variables con este fueron muy bajas ($r^2 < 0.5$).

Tabla 7– Composición de los 4 primeros vectores propios y correlaciones de cada variable con el vector.

Variables	Eigenvectors*			
	EV1	EV2	EV3	EV4
bio_1	-0.17 (0.46)	<u>0.36</u> (0.82)	0.15 (0.31)	-0.06 (0.06)
bio_2	0.24 (0.68)	0.11 (0.25)	-0.16 (0.33)	0.35 (0.35)
bio_3	<u>-0.30</u> (0.84)	0.04 (0.08)	-0.19 (0.38)	-0.18 (0.19)
bio_4	<u>0.32</u> (0.88)	-0.05 (0.12)	0.19 (0.39)	0.15 (0.16)
bio_5	0.16 (0.44)	<u>0.32</u> (0.73)	0.23 (0.46)	0.16 (0.16)
bio_6	<u>-0.30</u> (0.82)	0.22 (0.51)	0.05 (0.09)	-0.20 (0.21)
bio_7	<u>0.34</u> (0.94)	-0.01 (0.02)	0.09 (0.18)	0.26 (0.26)
bio_8	0.01 (0.02)	<u>0.32</u> (0.74)	0.16 (0.32)	0.20 (0.2)
bio_9	-0.12 (0.35)	<u>0.33</u> (0.75)	0.11 (0.23)	-0.08 (0.08)
bio_10	0.09 (0.25)	<u>0.32</u> (0.73)	0.31 (0.62)	0.05 (0.05)
bio_11	<u>-0.27</u> (0.76)	0.27 (0.62)	0.00 (0)	-0.12 (0.13)
bio_12	<u>-0.29</u> (0.81)	-0.16 (0.36)	0.18 (0.36)	0.20 (0.2)
bio_13	<u>-0.32</u> (0.89)	-0.06 (0.13)	0.04 (0.09)	0.36 (0.37)
bio_14	-0.06 (0.16)	-0.26 (0.6)	<u>0.35</u> (0.72)	-0.05 (0.05)
bio_15	-0.09 (0.26)	0.25 (0.56)	-0.31 (0.62)	0.33 (0.34)
bio_16	<u>-0.32</u> (0.89)	-0.07 (0.16)	0.05 (0.1)	0.35 (0.36)
bio_17	<u>-0.06</u> (0.18)	-0.27 (0.61)	<u>0.35</u> (0.72)	-0.05 (0.05)
bio_18	-0.25 (0.69)	-0.14 (0.31)	0.10 (0.2)	0.44 (0.45)
bio_19	-0.14 (0.38)	-0.19 (0.43)	0.31 (0.62)	-0.08 (0.08)
Alt	-0.04 (0.11)	-0.19 (0.44)	<u>-0.40</u> (0.81)	0.06 (0.06)
slope	-0.13 (0.36)	-0.02 (0.05)	<u>-0.20</u> (0.4)	0.06 (0.07)

* Eigenvector (Correlations Variable - axes)

Por su parte, en la Figura 4 se muestra la representación de las variables originales sobre los tres primeros vectores propios. Las variables con vectores más largos y ángulos más pequeños con relación a cada eje o vector propio, tienen una mayor importancia relativa para explicar la variabilidad existente en los datos. A su vez, las variables de mayor asociación tienen un ángulo menor entre ellas. Así, en la Figura 4-A, en donde se muestra la representación de las variables originales sobre los dos primeros vectores propios, las variables bio7, 4 y 2 (Localizadas a la derecha) tienen una relación positiva con relación al primer componente, y están relacionadas entre ellas, aunque las bio7 y 4 tienen una mayor relación entre ellas y mayor importancia relativa con relación a la bio2, presentando los vectores más largos. En contraposición, las bio13, 16 y 3 (localizadas a la izquierda del gráfico), tienen una relación negativa con el primer componente y están altamente relacionadas, especialmente las dos primeras. De esta forma, el primer componente se presenta como un balance entre variables asociadas con la temperatura de forma positiva (Posición a la derecha en las Figura 4-A y -B) y 3 variables de precipitación, junto a tres de temperatura de forma negativa (Posición a la izquierda en las Figura 4-A y -B). En el caso del segundo componente, se debe destacar la relación de signo positivo de la mayoría de las variables asociadas a la temperatura a

excepción de las bio4 y 7. Siendo las de mayor importancia relativa las bio1 y 9 (Figura 4-A arriba y -C a la derecha). Para el tercer componente principal, los vectores más largos se pueden ver para las bio17 y 14 (localizados arriba en las Figura 4-B y -C), y en sentido negativo la altitud.

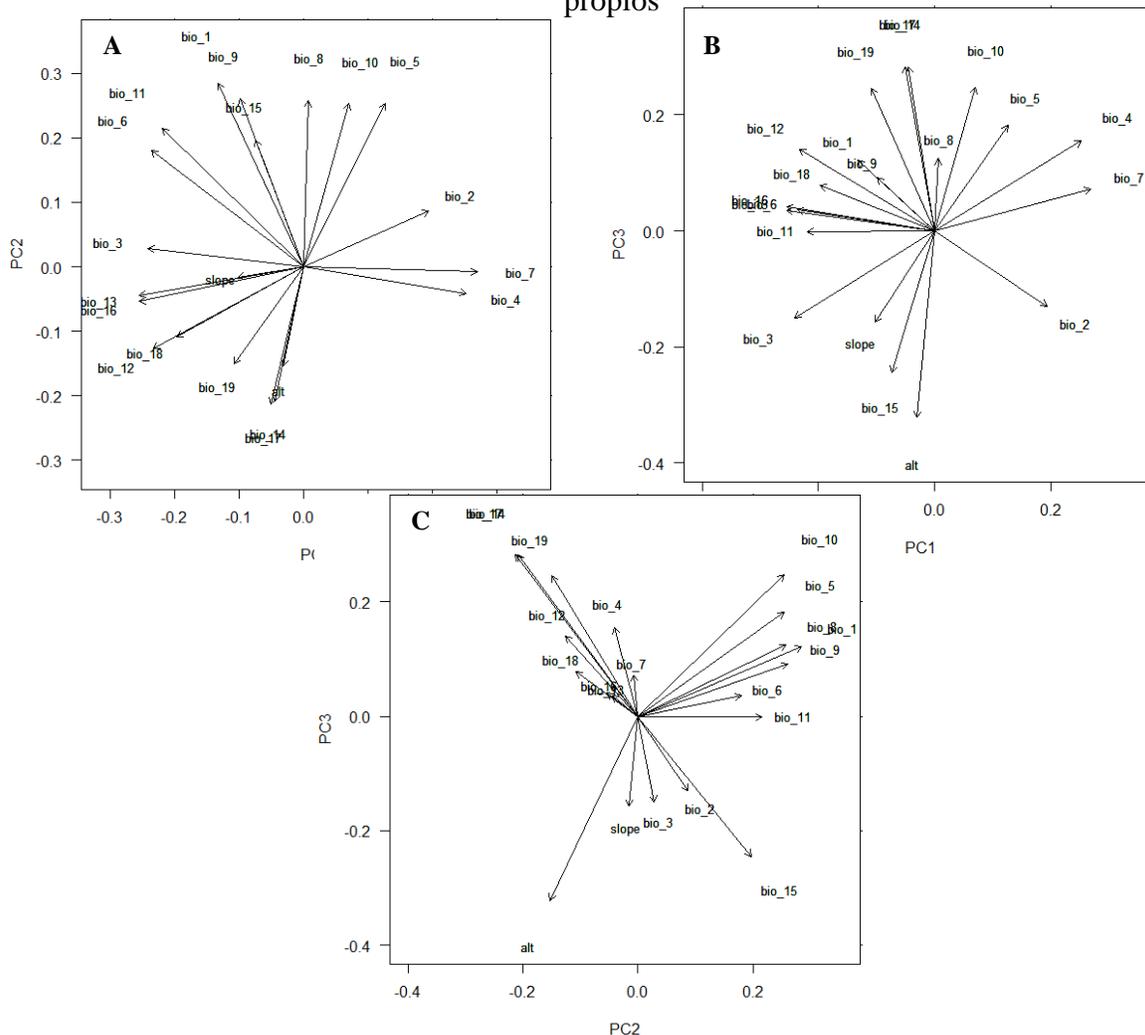
Así, en términos de la condición bioclimática que representan las variables con mayor importancia sobre cada eje, se podría decir que el primer vector propio se representa de forma importante la variación estacional de la temperatura, con dos variables que reflejan esta variación en sentido positivo (bio7 y 4) y una en sentido negativo (bio3), además en sentido negativo las variables de importancia relacionadas con la precipitación tienen que ver con condiciones promedio estacionales (bio 16), valores extremos en las condiciones bioclimáticas (bio 6) y condiciones promedio anuales (bio12). Por su parte las variables relacionadas con la temperatura, adicionalmente a la ya mencionada bio3 están relacionadas con las condiciones promedio estacionales (bio 11) y valores extremos en las condiciones (bio6). Esta variedad de condiciones representadas sobre el primer eje principal muestra la importancia de este vector en términos de sintetizar diferentes condiciones bioclimáticas.

Por su parte, en el segundo vector propio las variables de mayor importancia están asociadas con la temperatura y en el tercero con la precipitación, aunque al interior de cada eje se representan diferentes condiciones bioclimáticas. En el segundo vector, están representadas las condiciones promedio mensuales con una variable (bio1), las condiciones promedio estacionales con tres variables (bio9, 8 y 10) y valores extremos en las condiciones bioclimáticas con una variable (bio5). Mientras en el tercer vector propio se encuentran representadas las condiciones promedio estacionales con una variable (bio17), los valores extremos en las condiciones con una variable (bio14), ambas en sentido positivo y la altitud en sentido negativo que representa una condición propia del terreno.

Esta distribución de tipos variables y condiciones bioclimáticas en las variables de mayor importancia nos muestra que en términos de las variables estas están relacionadas con la temperatura en el primer y segundo vector propio y con la precipitación en el primer y tercer vector. Mientras en términos de las condiciones bioclimáticas, nos muestra una gran importancia de la estacionalidad intra-anual por su participación en el primer vector exclusivamente, mientras las condiciones promedio anuales participan en el primero y segundo vector. Por su parte, la importancia de las condiciones promedio estacionales y los

valores extremos se da por su participación en todos los ejes. Esta composición de los vectores propios se tiene en cuenta en la siguiente sección para realizar la selección de variables para el modelamiento con una preselección de variables.

Figura 4– Representación de las variables originales sobre los cuatro primeros vectores propios



5.3 Selección de variables para el modelamiento. En la Tabla 8 se muestra un resumen de las variables con mayor participación en los primeros tres vectores propios, las cuales están organizadas de acuerdo a la importancia relativa dentro del vector y al sentido de la relación con el mismo, adicionalmente, a la derecha de cada variable y entre paréntesis se colocó la condición bioclimática que cada una de ellas representa. De acuerdo al primer criterio de selección de variables, 15 de las 19 variables bioclimáticas y la altitud, resultaron con una alta correlación y participación con los tres primeros vectores propios, descartando

por este criterio únicamente a las variables bioclimáticas bio2, 15, 18 y 19 y la pendiente del sitio de colecta, esta última como variable propia del terreno del sitio.

Tabla 8 – Variables de mayor participación en cada componente principal organizadas de acuerdo a su importancia y sentido de la correlación

Tipo de Correlación	Componente Principal					
	CP1*		CP2		CP3	
Positiva	bio 7	(4)	bio 1	(1)	bio 17	(2)
	bio 4	(4)	bio 9	(2)	bio 14	(3)
			bio 8	(2)		
			bio 10	(2)		
			bio 5	(3)		
Negativa	bio 13	(3)			Altitud	
	bio 16	(2)				
	bio 3	(4)				
	bio 6	(3)				
	bio 12	(1)				
	bio 11	(2)				

* Los números entre paréntesis son la condición bioclimática que representa cada variable: 1. Condiciones promedio anuales, 2. Condiciones promedio estacionales, 3. Valores extremos en las condiciones bioclimáticas, 4. Estacionalidad intra-anual.

Para realizar una selección dentro de las quince variables bioclimáticas y la altitud se tuvo en cuenta un análisis realizado a las variables que participaban en cada vector teniendo en cuenta los criterios nombrados en la sección metodológica. De esta forma, como se muestra en la Tabla 8, en el primer vector propio participan tres variables que representan condiciones bioclimáticas de estacionalidad intra-anual de la temperatura, dentro de estas, la de mayor peso relativo para explicar la variabilidad existente en el conjunto de datos bioclimáticos o con mayor participación en el primer vector propio fue el rango de temperatura anual (bio7), estando muy relacionada a esta la estacionalidad en la temperatura (bio4) de forma positiva y a la Isotermalidad (bio3) de forma negativa, incluso en el cálculo de esta última está involucrado el rango de temperatura anual ($bio3 = bio2/bio7$), por lo que será seleccionada la bio7 para realizar el modelamiento y se descartarán la bio4 y bio3. Esto debido a que por una parte, la bio7 como ya se mencionó tiene una mayor importancia relativa, está involucrada en el cálculo de la bio3 presentando una relación inversamente proporcional con esta, lo que podría explicar la posición en el primer vector de cada variable, adicionalmente, la bio4 tiene un comportamiento muy similar a la bio7, especialmente en los dos primeros vectores, siendo

esta escogencia una forma de remover dos variables redundantes en términos de la condición bioclimática que representan para el conjunto de datos analizados.

Por otra parte, al observar las demás variables que conforman el primer vector propio, en el sentido negativo del eje aun se encuentran cinco variables, tres de las cuales están relacionadas con la precipitación y son en orden de importancia en términos de la explicación de la variabilidad de estos datos, la precipitación del mes y del trimestre mas húmedos (bio13 y 16, respectivamente), aunque estas dos variables representan dos condiciones bioclimáticas distintas, siendo la bio13 representante de las condiciones extremas y la bio16 de las condiciones promedio estacionales, tienen un comportamiento muy similar en los datos analizados, por lo que se puede asumir que el mes más húmedo hace parte del trimestre más húmedo, por tal motivo fue excluida la bio16, esto debido a que la bio13 tiene una mayor importancia relativa en la explicación de la variabilidad. La tercera variable en este eje relacionada con la precipitación fue la precipitación anual (bio12), la cual es la única variable que representa las condiciones promedio anuales en términos de la precipitación y al participar de forma importante en el primer vector propio fue incluida para el modelamiento. Así en términos de la precipitación, en este vector fue incluida la precipitación del mes más húmedo (bio13) y la precipitación anual (bio12), las cuales representan valores extremos y promedios anuales de precipitación.

Además, en el sentido negativo del primer vector se encuentran dos variables adicionales asociadas a la temperatura, las cuales son la temperatura mínima del mes más frío (bio6) y la temperatura mínima del trimestre más frío (bio11), estando estas dos relacionadas de manera similar con lo anteriormente descrito para las bio 13 y 16. Por tal razón, se decidió descartar la bio11, es decir que la selección en términos de las menores temperaturas representadas en este vector y debido a la importancia relativa de la variable, favoreció a la bio6 que representa un valor extremo en las condiciones bioclimáticas.

Por su parte, para el segundo vector propio se tienen cinco variables asociadas a la temperatura y correlacionadas positivamente con el vector, siendo la primera en orden de importancia la temperatura media anual (bio1), la cual representa condiciones promedio anuales en términos de temperatura, por lo cual se incluyó dentro de las variables para realizar el modelamiento. Las tres siguientes variables representan condiciones promedio estacionales para la temperatura y son en orden de importancia la temperatura media del trimestre más seco,

temperatura media del trimestre más húmedo y la temperatura media del trimestre más caliente (bio9, 8 y 10, respectivamente). En este caso, fueron seleccionadas las bio9 y 8 y descartada la bio10, esto debido por una parte a la mayor importancia de las dos primeras variables, las cuales representan temperaturas en dos periodos estacionales con diferente condición de humedad, y por otro lado, la relación con la variable bio5, la cual es la siguiente en orden de importancia en este vector, siendo esta última la temperatura máxima del mes más caliente y representando una condición extrema en la temperatura. En resumen, en este vector se seleccionaron las bio1, 9, 8 y 5, siendo la primera representante de las condiciones promedio anuales de la temperatura, las dos siguientes representantes de las condiciones promedio estacionales y la última de valores extremos en las condiciones bioclimáticas de temperatura, en la sección de discusión se profundizará sobre la selección de la variable 5 en cambio de la 10.

Finalmente, en el tercer vector propio y en sentido positivo se encontraron con importante participación las variables de precipitación del trimestre y mes más seco (bio17 y 14, respectivamente). En este caso, para evitar incluir información redundante, se seleccionó la bio17 que representa condiciones promedio estacionales, debido a su importancia relativa en este vector. Adicionalmente, fue incluida la altitud como variable representante de las condiciones del terreno y debido a su presencia en el sentido negativo del vector.

De esta forma fueron seleccionadas un total de 10 variables para realizar el modelamiento en Maxent y compararlo con el modelamiento realizado con la totalidad de las variables bioclimáticas, la altitud y la pendiente del sitio de colecta. Estas variables fueron la temperatura media y precipitación anual (bio1 y 12, respectivamente), representantes de las condiciones promedio anuales. La temperatura media del trimestre más seco, del trimestre más húmedo y la precipitación del trimestre más seco (bio8, 9 y 17, respectivamente), representantes de las condiciones promedio estacionales. La temperatura máxima del mes más caliente y la mínima del mes más frío, junto a la precipitación del mes más húmedo (bio 5, 6 y 13, respectivamente), representantes de valores extremos en las condiciones bioclimáticas. Finalmente, el rango de temperatura anual (bio7) que representa la estacionalidad intra-anual en la temperatura y la altitud del sitio de colecta del germoplasma.

5.4 Modelamiento en Maxent. Fueron ajustados modelamientos de nicho ecológico para 34 taxones del genero *Phaseolus* en dos juegos de variables, es decir, para la totalidad de las variables tenidas en cuenta y para la selección realizada a partir del ACP. Ambos modelamientos fueron realizados en un computador con sistema operativo Windows de 64 bits, procesador Intel Pentium Dual Core de 2.2 GHz y una memoria RAM de 4 Gb. El tiempo de ejecución y espacio de disco usado por el modelamiento con la totalidad de variables fue de 53.04 horas y 62.5 Gb, mientras que para la selección de variables fue de 20.5 horas y 60.5 Gb, haciéndose considerable la disminución en tiempo de ejecución.

5.5 Evaluación estadística de la precisión de los modelos generados en Maxent. En la Tabla 9, se presenta un resumen de los índices utilizados para evaluar los modelamientos realizados en los 34 taxones del genero *Phaseolus* incluidos en este estudio. Dicha tabla presenta el número de presencias utilizadas por el modelo, junto a los promedios de los índices AUC, Kappa y TSS por cada táxon y discriminados de acuerdo al modelamiento empleado, es decir, el modelamiento realizado con 21 variables (19 bioclimáticas, pendiente y altitud) y con 10 variables (9 bioclimáticas y la altitud). Adicionalmente, fueron señalados con un asterisco los promedios que resultaron estadísticamente superiores de acuerdo al grupo de variables utilizado.

Como se puede apreciar en la Tabla 9, en el caso del AUC, los promedios de este índice fueron estadísticamente similares en 32 de los 34 taxones (p -valor > 0.05) y presentaron promedios estadísticamente superiores para los modelos ajustados con todas las variables (p -valor < 0.05), únicamente en *P. acutifolius acutifolius* y en *P. neglectus*. En el caso de *P. acutifolius acutifolius* con promedios de AUC de 0.958 para el modelamiento con todas las variables y de 0.943 con la selección realizada. Mientras en el caso de *P. neglectus* los promedios fueron de 0.975 y 0.768, para los modelos ajustados con la totalidad y la selección de variables, respectivamente. Para los 32 casos restantes, los promedios en el AUC fueron mayores para 27 taxones cuando se realizó el modelamiento con todas las variables y en 5 ocasiones cuando fue realizado con la selección de las mismas, sin embargo estas diferencias no fueron estadísticamente significativas. Adicionalmente, los promedios de AUC para los 32 taxones donde no hubo diferencias estadísticas variaron entre 0.922 en *P. vulgaris* con la

selección de variables y 0.999 para el modelamiento realizado en *P. talamancensis* con todas las variables.

En el caso del índice Kappa, como se puede observar en la Tabla 9, 66 de los 68 índices Kappa estuvieron por debajo de 0.4, es decir, presentaron promedios inferiores al valor mínimo para considerar un modelo preciso mediante el índice, solamente *P. costaricensis* y *P. talamancensis* para los modelos ajustados con todas las variables superaron el umbral de 0.4 con 0.465 y 0.527, respectivamente, aunque sin que la diferencia fuera estadísticamente significativa con los modelamientos para dichos taxones realizados con la selección de variables. Así 31 de los 34 taxones tuvieron promedios similares estadísticamente y solamente *P. dumosus*, *P. pachyrrhizoides* y *P. tuerckheimii* presentaron promedios mayores de Kappa para los modelos realizados con todas las variables (0.060, 0.196 y 0.344, para *P. dumosus*, *P. pachyrrhizoides* y *P. tuerckheimii*, respectivamente) en comparación con los valores obtenidos con la selección de ellas (0.041, 0.035 y 0.106, para *P. dumosus*, *P. pachyrrhizoides* y *P. tuerckheimii*, respectivamente). Además, el índice fue mayor en 25 casos para los modelos ajustados con todas las variables.

Finalmente, en el caso del índice TSS, la Tabla 9 muestra que los valores promedio del índice son similares estadísticamente en 29 de los 34 taxones, mientras que para *P. oligospermus*, *P. pachyrrhizoides* y *P. talamancensis*, los promedios fueron superiores para el modelamiento realizado con todas las variables (0.856, 0.835 y 0.899, para *P. oligospermus*, *P. pachyrrhizoides* y *P. talamancensis*, respectivamente) en comparación con el ajuste realizado con la selección de las mismas (0.675, 0.562 y 0.499 para *P. oligospermus*, *P. pachyrrhizoides* y *P. talamancensis*, respectivamente). Adicionalmente, para *P. maculatus* y *P. pluriflorus*, los promedios fueron estadísticamente superiores para los modelos ajustados con la selección de variables (0.835 y 0.889 para *P. maculatus* y *P. pluriflorus*, respectivamente) en comparación con el ajuste realizado con todas las variables (0.443 y 0.782 para *P. maculatus* y *P. pluriflorus*, respectivamente). En el caso del índice TSS, los promedios tuvieron valores nominales mayores en comparación con los de Kappa variando entre 0.338 y 0.926. Es decir, asumiendo la misma clasificación de índices usada para Kappa estuvieron en las categorías de precisión pobre solamente 3 promedios de TSS (2 cuando fueron utilizadas todas las variables y 1 cuando fue utilizada la selección), 4 promedios en precisión aceptable (2 y 2, con todas y la selección de variables, respectivamente), buena (1 y 5, con todas y la selección de variables,

respectivamente), muy buena (18 y 17, con todas y la selección de variables, respectivamente) y excelente (11 y 9, con todas y la selección de variables, respectivamente).

Tabla 9 - Promedio de los índices AUC, Kappa y TSS calculados para los 10 modelos ajustados mediante validación cruzada en 34 taxones del género *Phaseolus* usando 21 y 10 variables.

Taxón de <i>Phaseolus</i> ^{1,2}	n	AUC		Kappa		TSS	
		(I)	(II)	(I)	(II)	(I)	(II)
<i>P. acutifolius</i>	29	0.963	0.967	0.042	0.054	0.755	0.826
<i>P. acutifolius acutifolius</i>	120	0.958*	0.943	0.021	0.014	0.844	0.815
<i>P. acutifolius tenuifolius</i>	154	0.965	0.959	0.053	0.036	0.833	0.836
<i>P. amblyosepalus</i>	10	0.967	0.938	0.218	0.002	0.468	0.338
<i>P. angustissimus</i>	181	0.980	0.977	0.075	0.065	0.918	0.912
<i>P. augusti</i>	40	0.979	0.976	0.042	0.018	0.829	0.823
<i>P. coccineus</i>	342	0.949	0.942	0.071	0.060	0.862	0.844
<i>P. coccineus coccineus</i>	32	0.974	0.968	0.029	0.022	0.773	0.791
<i>P. costaricensis</i>	35	0.996	0.995	0.465	0.323	0.850	0.883
<i>P. dumosus</i>	124	0.959	0.963	0.060*	0.041	0.848	0.825
<i>P. filiformis</i>	290	0.970	0.962	0.074	0.072	0.882	0.862
<i>P. glabellus</i>	15	0.986	0.985	0.250	0.238	0.682	0.680
<i>P. grayanus</i>	48	0.977	0.974	0.025	0.025	0.840	0.895
<i>P. leptostachyus</i>	192	0.959	0.954	0.039	0.037	0.848	0.837
<i>P. leptostachyus leptostachyus</i>	43	0.940	0.944	0.017	0.024	0.725	0.798
<i>P. lunatus</i>	359	0.929	0.923	0.047	0.045	0.784	0.776
<i>P. maculatus</i>	10	0.942	0.934	0.002	0.003	0.443	0.835*
<i>P. maculatus maculatus</i>	130	0.970	0.965	0.030	0.029	0.832	0.852
<i>P. maculatus ritensis</i>	105	0.984	0.981	0.065	0.046	0.913	0.892
<i>P. neglectus</i>	10	0.975**	0.768	0.005	0.000	0.376	0.469
<i>P. nelsonii</i>	26	0.989	0.990	0.039	0.051	0.866	0.819
<i>P. oligospermus</i>	27	0.986	0.978	0.079	0.043	0.856*	0.675
<i>P. pachyrrhizoides</i>	23	0.978	0.951	0.196*	0.035	0.835*	0.562
<i>P. parvifolius</i>	58	0.955	0.954	0.013	0.012	0.808	0.775
<i>P. parvulus</i>	70	0.988	0.988	0.069	0.055	0.860	0.875
<i>P. pauciflorus</i>	101	0.983	0.981	0.064	0.056	0.926	0.896
<i>P. pedicellatus</i>	59	0.940	0.931	0.012	0.014	0.731	0.753
<i>P. pluriflorus</i>	43	0.988	0.985	0.060	0.038	0.782	0.889*
<i>P. polystachios</i>	322	0.903	0.903	0.026	0.024	0.770	0.770
<i>P. talamancensis</i>	11	0.999	0.999	0.527	0.337	0.899*	0.499
<i>P. tuerckheimii</i>	18	0.996	0.994	0.344*	0.106	0.744	0.641
<i>P. vulgaris</i>	490	0.931	0.922	0.067	0.061	0.800	0.780
<i>P. xanthotrichus</i>	31	0.993	0.992	0.249	0.129	0.898	0.796
<i>P. zimapanensis</i>	12	0.950	0.940	0.069	0.025	0.347	0.587

1. (I) Modelamiento realizado con 21 variables, (II) Modelamiento realizado con 10 variables. 2. Significancia de la prueba t-Student: * p-valor < 0.05, ** p-valor < 0.001.

5.6 Análisis de la estabilidad en la predicción logística realizada por Maxent.

En la Tabla 10 se puede observar un resumen del área total y el porcentaje del área de la cobertura de distribución potencial con desviaciones estándar menores a 0.1 (PADM10), tanto para el modelamiento realizado con la totalidad de las variables (21 Variables), como con la selección de ellas realizada mediante un ACP (10 Variables) y para los 34 taxones del presente estudio. En términos del área de los nichos ecológicos, es posible observar que cuando fueron utilizadas la totalidad de las variables estas estuvieron entre los 2.4, 5.2 y 9.1 miles de km² para *P. talamancensis*, *P. costaricensis* y *P. tuerckheimii*, respectivamente, y los 2923.2, 2525.4 y 1558.4 miles de km² para *P. lunatus*, *P. polystachios* y *P. vulgaris*, respectivamente. Adicionalmente, cuando fueron utilizadas para el modelamiento la selección de variables, el orden relativo se mantuvo para los taxones nombrados con 2.3, 5.2 y 11.7 miles de km² para *P. talamancensis*, *P. costaricensis* y *P. tuerckheimii*, respectivamente, mientras que los mayores valores fueron 6352.2, 2680.5 y 1306.5 miles de km² para *P. lunatus*, *P. polystachios* y *P. vulgaris*, respectivamente, siendo especialmente notoria la diferencia para *P. Lunatus*. De la misma forma, en 15 de los 34 taxones el área de la superficie de distribución potencial fue mayor en el caso de los modelamientos ajustados mediante la utilización de la selección de variables, mientras en los restantes 19 taxones fue mayor el área para la predicción realizada cuando fue utilizada la totalidad de las mismas.

Por su parte en la Figura 5, se muestra un gráfico que compara las áreas de los nichos ecológicos para cada taxón haciendo uso de la selección de variables en el eje X y con la totalidad de las mismas en el eje Y. Además, cada punto fue representado con diferentes tamaños y colores en 4 categorías, las cuales representan la razón entre el área predicha para los modelos ajustados con la selección de variables y con la totalidad de las mismas. De esta forma, en la primera categoría se encuentran 4 taxones (*P. neglectus*, *P. lunatus*, *P. pachyrrhizoides* y *P. amblyosepalus*) representados con puntos de mayor tamaño y color azul claro, los cuales tienen una razón entre 4.4 y 2, es decir que la predicción hecha a partir de la selección de variables por lo menos duplica la realizada con la totalidad de las mismas, destacándose la diferencia en valor absoluto de *P. lunatus*. En el segundo grupo, se encuentran 5 taxones que tienen una razón entre 1.5 y 1.2 (*P. augusti*, *P. pluriflorus*, *P. tuerckheimii*, *P. maculatus*, *P. oligospermus*), en los cuales la predicción a partir de la selección de variables

supera a la realizada con la totalidad por entre el 20 y 50%. En el tercer grupo, se encuentran 20 taxones y tienen una razón entre 1.2 y 0.8, en los cuales el tamaño de los nichos es muy similar. Finalmente, en el cuarto grupo se encuentran 5 taxones que tienen un índice entre 0.8 y 0.6 (*P. leptostachyus leptostachyus*, *P. zimapanensis*, *P. acutifolius*, *P. acutifolius tenuifolius*, *P. pauciflorus*), en los cuales el tamaño del nicho ecológico predicho con la totalidad de las variables es mayor al encontrado para la selección de las mismas.

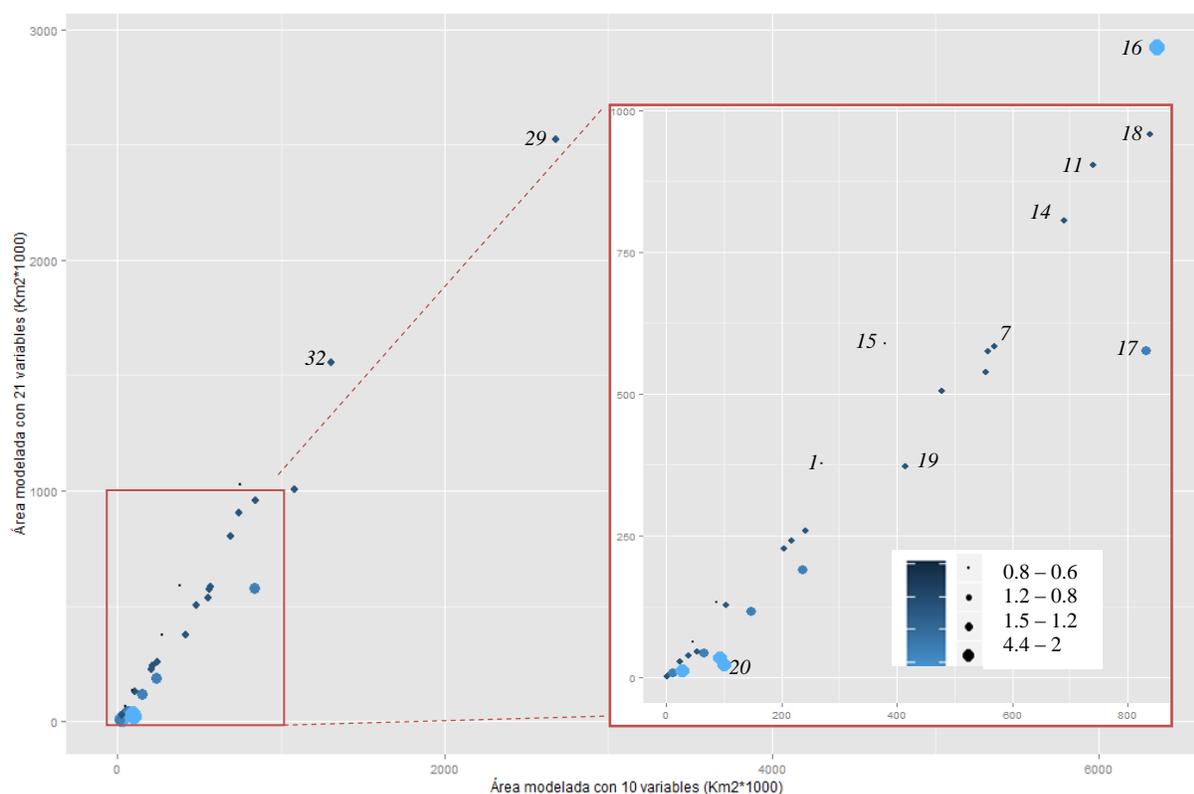
Por su parte, el valor de PADM10 para los modelos en donde se utilizaron la totalidad de las variables estuvo entre 83.2, 86.7 y 91.4% para *P. talamancensis*, *P. pauciflorus* y *P. glabellus*, respectivamente, y el 100% que se presentó en 4 taxones (*P. xanthotrichus*, *P. grayanus*, *P. parvulus*, *P. costaricensis*), siendo los modelos de *P. talamancensis* y *P. pauciflorus* menores al 90%, es decir, modelos con baja estabilidad en su predicción espacial. En el caso de los modelos ajustados mediante la selección de variables, el PADM10 varió entre el 92.1, 95.1 y 96.2 para *P. glabellus*, *P. pauciflorus* y *P. acutifolius acutifolius*, respectivamente, mientras que 12 taxones tuvieron una PADM10 del 100% (*P. neglectus*, *P. lunatus*, *P. amblyosepalus*, *P. maculatus*, *P. pluriflorus*, *P. talamancensis*, *P. parvifolius*, *P. pedicellatus*, *P. grayanus*, *P. parvulus*, *P. nelsonii*, *P. costaricensis*). , es decir, que mediante la utilización de la selección de variables realizada mediante el ACP, los modelos de la totalidad de los taxones tuvieron estabilidades adecuadas.

Tabla 10 – Área total de la cobertura de distribución potencial y porcentaje de la misma con desviaciones estándar menores a 0.1 de acuerdo a la totalidad (21 variables) o la selección (10 variables) de variables para 34 taxones del género *Phaseolus*.

Taxón de <i>Phaseolus</i>	21 Variables		10 Variables		Razón (10V/21V) ²
	Area (Km2*1000)	PADM10 (%) ¹	Area (Km2*1000)	PADM10 (%) ¹	
1 - <i>P. acutifolius</i>	377.7	97.2	269.1	99.5	0.7
2 - <i>P. acutifolius_acutifolius</i>	1008.2	96.4	1083	96.2	1.1
3 - <i>P. acutifolius_tenuifolius</i>	1024.6	99.7	747.8	99.3	0.7
4 - <i>P. amblyosepalus</i>	11.9	99.4	27.8	100	2.3
5 - <i>P. angustissimus</i>	538.4	99.2	554.2	97.4	1.0
6 - <i>P. augusti</i>	190.9	95.8	235.4	99.6	1.2
7 - <i>P. coccineus</i>	583.3	99.6	568.5	99.6	1.0
8 - <i>P. coccineus_coccineus</i>	259.3	95.1	241.1	99.9	0.9
9 - <i>P. costaricensis</i>	5.2	100	5.2	100	1.0
10 - <i>P. dumosus</i>	46.6	94.9	53.5	98	1.1
11 - <i>P. filiformis</i>	903.5	99.7	740.4	99.5	0.8
12 - <i>P. glabellus</i>	39.4	91.4	39.6	92.1	1.0
13 - <i>P. grayanus</i>	505.3	100	478.1	100	0.9
14 - <i>P. leptostachyus</i>	804.9	95.8	690.7	98.1	0.9
15 - <i>P. leptostachyus leptostachyus</i>	588.6	96	379.5	97	0.6
16 - <i>P. lunatus</i>	2923.2	99.6	6352.2	100	2.2
17 - <i>P. maculatus</i>	576.2	99.9	832.3	100	1.4
18 - <i>P. maculatus_maculatus</i>	957.2	99	839.7	98.3	0.9
19 - <i>P. maculatus_ritensis</i>	373.2	98.6	414	97.5	1.1
20 - <i>P. neglectus</i>	23	99.9	101.2	100	4.4
21 - <i>P. nelsonii</i>	127.5	99.5	104.1	100	0.8
22 - <i>P. oligospermus</i>	43.3	93.4	64.2	98.9	1.5
23 - <i>P. pachyrrhizoides</i>	34.5	99.8	93.1	97	2.7
24 - <i>P. parvifolius</i>	574.6	99.4	558.2	100	1.0
25 - <i>P. parvulus</i>	241.6	100	216.8	100	0.9
26 - <i>P. pauciflorus</i>	63.9	86.7	46.7	95.1	0.7
27 - <i>P. pedicellatus</i>	226.8	99.7	203.8	100	0.9
28 - <i>P. pluriflorus</i>	116.6	99.9	146.9	100	1.3
29 - <i>P. polystachios</i>	2525.4	98.1	2680.5	99.7	1.1
30 - <i>P. talamancensis</i>	2.4	83.2	2.3	100	1.0
31 - <i>P. tuerckheimii</i>	9.1	99.7	11.7	97.4	1.3
32 - <i>P. vulgaris</i>	1558.4	99.2	1306.5	99.5	0.8
33 - <i>P. xanthotrichus</i>	28	100	24.2	99.7	0.9
34 - <i>P. zimapanensis</i>	133.6	95.1	87.7	99.4	0.7

1- PADM10: proporción de área de la cobertura de distribución potencial con desviaciones estándar menores a 0.1, 2- Razón entre las áreas de los nichos ecológicos predichos con 10 variables y 21 variables

Figura 5– Representación del área predicha mediante el modelamiento con la selección de variables (Eje X) y la totalidad de las mismas (Eje Y) para los 34 taxones del género *Phaseolus*.



5.7 Criterios tenidos en cuenta para la selección del modelamiento a ser seleccionado. En la Tabla 11, se muestra un resumen de los puntajes asignados a cada modelamiento de acuerdo a la evaluación de la precisión estadística y su estabilidad espacial. Así, debajo de cada letra y el número I o II, según sea el caso, se encuentra el puntaje asignado para los modelamientos realizados con todas o la selección de variables. Así, en el caso del índice AUC (Tabla 11, debajo de la letra A), en 32 de los 34 taxones el puntaje fue de 1 para ambos modelamientos, a excepción de *P. acutifolius acutifolius* y *P. neglectus*, en donde la diferencia estadística entre los promedios favoreció a los modelos realizados con todas las variables, por lo que fue asignado un valor de 2 para estos modelos y de 0 para los ajustados con la selección realizada. Por su parte, en el caso del índice Kappa (Tabla 11, debajo de la letra B), fue asignado un valor de 0 en todos los taxones a excepción de los taxones *P. costaricensis* y *P. talamancensis*, en donde fue asignado un valor de 1 cuando los modelos

fueron ajustados con todas las variables, esto debido a que los valores de este índice a excepción de los dos citados tuvieron valores menores a 0.4, lo que corresponde a modelos impresos o de pobre precisión. Así, aunque en el caso de *P. dumosus*, *P. pachyrrhizoides* y *P. tuerckheimii*, los promedios resultaron estadísticamente superiores para los modelos ajustados con todas las variables, estos fueron inferiores al valor mínimo para ser considerados como indicadores de exactitud en los modelamientos.

Por su parte, en el caso del índice TSS (Tabla 11, debajo de la letra C), en 26 casos fue asignado un valor de 1 a ambos modelamientos debido a la igualdad estadística de los promedios para este índice, mientras que para *P. oligospermus*, *P. pachyrrhizoides* y *P. talamancensis*, fue asignado un valor de 2 para los modelos ajustados con todas las variables y para *P. maculatus* y *P. pluriflorus* fue asignado el mismo valor para los ajustados con la selección de variables, esto debido a la superioridad estadística mostrada por estos promedios en comparación con aquellos ajustados con el juego contrario de variables. Así, en el caso de *P. amblyosepalus*, *P. neglectus* y *P. zimapanensis* aunque tienen promedios similares, para *P. amblyosepalus* el valor de TSS fue menor de 0.4, mientras en los dos restantes los valores de TSS menores a 0.4 fueron para los modelos ajustados con la totalidad de las variables, por lo que para estos casos fue asignado el valor de 0.

De esta forma, la selección de los modelos fue realizada mediante la comparación de los puntajes obtenidos de acuerdo al número de variables utilizado en Maxent (Tabla 11, columna PEM) y bajo la columna denominada MS se define el modelo seleccionado. Para 27 taxones fue seleccionado el modelamiento realizado con la selección de las variables, es decir con 10 variables, de estos taxones en 10 casos el puntaje obtenido por el modelamiento realizado con la selección de las variables tuvo un mayor puntaje y en 17 casos el puntaje fue igual. Por su parte, para 7 taxones el puntaje obtenido por el modelamiento realizado con la totalidad de las variables (21 variables) fue mayor, por lo cual para estos taxones serán los únicos en los cuales el análisis espacial será realizado con este juego de variables, siendo los taxones *P. acutifolius acutifolius*, *P. amblyosepalus*, *P. angustissimus*, *P. costaricensis*, *P. neglectus*, *P. pachyrrhizoides* y *P. tuerckheimii*.

Tabla 11 – Puntajes asignados a cada modelamiento de acuerdo a la evaluación de la precisión estadística y la estabilidad en la predicción logística para los 34 taxones del genero *Phaseolus*.

Taxón de Phaseolus	A ¹		B ²		C ³		Precisión estadística ⁴		Estabilidad D ⁵		PEM ⁶		MS ⁷
	(I)	(II)	(I)	(II)	(I)	(II)	(I)	(II)	(I)	(II)	(I)	(II)	
	<i>P. acutifolius</i>	1	1	0	0	1	1	2	2	5	6	3.5	
<i>P. acutifolius_acutifolius</i>	2	0	0	0	1	1	3	1	5	5	4	3	21
<i>P. acutifolius_tenuifolius</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. amblyosepalus</i>	1	1	0	0	1	0	2	1	6	6	4	3.5	21
<i>P. angustissimus</i>	1	1	0	0	1	1	2	2	6	5	4	3.5	21
<i>P. augusti</i>	1	1	0	0	1	1	2	2	5	6	3.5	4	10
<i>P. coccineus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. coccineus_coccineus</i>	1	1	0	0	1	1	2	2	5	6	3.5	4	10
<i>P. costaricensis</i>	1	1	1	0	1	1	3	2	6	6	4.5	4	21
<i>P. dumosus</i>	1	1	0	0	1	1	2	2	4	6	3	4	10
<i>P. filiformis</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. glabellus</i>	1	1	0	0	1	1	2	2	3	3	2.5	2.5	10
<i>P. grayanus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. leptostachyus</i>	1	1	0	0	1	1	2	2	5	6	3.5	4	10
<i>P. leptostachyus_leptostachyus</i>	1	1	0	0	1	1	2	2	5	5	3.5	3.5	10
<i>P. lunatus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. maculatus</i>	1	1	0	0	0	2	1	3	6	6	3.5	4.5	10
<i>P. maculatus_maculatus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. maculatus_ritensis</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. neglectus</i>	2	0	0	0	0	1	2	1	6	6	4	3.5	21
<i>P. nelsonii</i>	1	1	0	0	1	1	1	1	6	6	3.5	3.5	10
<i>P. oligospermus</i>	1	1	0	0	2	0	3	1	4	6	3.5	3.5	10
<i>P. pachyrrhizoides</i>	1	1	0	0	2	0	3	1	6	5	4.5	3	21
<i>P. parvifolius</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. parvulus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. pauciflorus</i>	1	1	0	0	1	1	2	2	0	5	1	3.5	10
<i>P. pedicellatus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. pluriflorus</i>	1	1	0	0	0	2	1	3	6	6	3.5	4.5	10
<i>P. polystachios</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. talamancensis</i>	1	1	1	0	2	0	4	1	0	6	2	3.5	10
<i>P. tuerckheimii</i>	1	1	0	0	1	1	2	2	6	5	4	3.5	21
<i>P. vulgaris</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. xanthotrichus</i>	1	1	0	0	1	1	2	2	6	6	4	4	10
<i>P. zimapanensis</i>	1	1	0	0	0	1	1	2	5	6	3	4	10

(I) Puntajes para los modelos ajustados con todas las variables, (II) Puntajes para los modelos ajustados con la selección de variables. 1,2,3 – puntuaciones dadas a los modelos de acuerdo a la comparación entre los índices AUC, Kappa y TSS para los modelos ajustados con todas las variables y con la selección de ellas. 4- Sumatoria de A+B+C para cada modelamiento. 5-Puntaje dado de acuerdo al PADM10. 6- Puntaje de Evaluación del Modelamiento (PEM)= 0.5(A+B+C) + 0.5 (D). 7- modelamiento seleccionado: Todas las variables (21 Variables), Selección de variables (10 Variables).

5.8 Proceso de análisis espacial. En esta sección se presentan y comparan los nichos ecológicos de los 34 taxones incluidos en el estudio, luego se presenta el análisis de fragmentación de nichos y finalmente se presenta el análisis de riqueza de los taxones, todos

estos análisis en las condiciones bioclimáticas actuales y de cambio climático con o sin capacidad de migración.

5.8.1 Presentación y comparación de nichos ecológicos: De acuerdo a la metodología planteada para el modelamiento en Maxent, luego de ajustados los modelos con las condiciones bioclimáticas presentes en los puntos de ocurrencia, estos modelamientos fueron proyectados espacialmente a las condiciones presentes y de cambio climático para el periodo 2020 a 2059, denominado 2050s. Luego fue comparado el área de los nichos ecológicos en condiciones presentes y de cambio climático, en la Tabla 12, se presenta el resumen de las áreas obtenidas en cada caso, en ellas se puede comparar el área del nicho de cada taxón en miles de kilómetros cuadrados, en condiciones actuales (Tabla 12, Letra A), bajo condiciones de cambio climático y sin capacidad de migración (Tabla 12, Letra B) y con capacidad de migración (Tabla 12, Letra D). Así mismo, se puede ver el porcentaje de área del nicho perdida en condiciones de cambio climático (Tabla 12, Letra C) y el porcentaje de área del nicho en condiciones de cambio climático que corresponde a nuevas áreas con relación a las condiciones actuales (Tabla 12, Letra E).

De esta forma, para todos los taxones del genero *Phaseolus* incluidos en el presente estudio resultaron modeladas áreas potenciales con condiciones bioclimáticas favorables en condiciones de cambio climático, tanto para el escenario de no migración como el de migración. Es decir, ninguno de los tamaños de los nichos ecológicos tuvo un valor de 0 para las condiciones de cambio climático (Tabla 12, Letras B y D). De esta forma, los tamaños de los nichos variaron entre los 2.3, 0.8 y 0.8 miles de Km² para *P. talamancensis* en las condiciones actuales, de cambio climático sin migración y con migración, respectivamente, hasta los 6352.2, 4533.5 y 7183.0 miles de km² para *P. lunatus* en iguales condiciones (Tabla 12, Letras A, B, D).

Así, al comparar los tamaños de los nichos en condiciones presentes y de cambio climático bajo el escenario de no migración, estos tamaños disminuyeron para todos los taxones a excepción del nicho de *P. leptostachyus leptostachyus*, *P. maculatus*, *P. acutifolius acutifolius* y *P. neglectus*, en donde el área del nicho se mantuvo o disminuyó en menos del 2% del área con relación al nicho en condiciones actuales (Tabla 12, Columna C). Para los restantes taxones, la perdida de área del nicho estuvo entre el 10 y 30% para 7 taxones (*P.*

angustissimus, *P. vulgaris*, *P. acutifolius tenuifolius*, *P. parvifolius*, *P. pachyrrhizoides*, *P. pedicellatus*, *P. lunatus*), entre el 30 y 50% para 11 taxones (*P. filiformis*, *P. coccineus*, *P. coccineus coccineus*, *P. pluriflorus*, *P. augusti*, *P. glabellus*, *P. maculatus ritensis*, *P. zimapanensis*, *P. pauciflorus*, *P. leptostachyus*, *P. tuerckheimii*) y con un porcentaje de pérdida de área mayor al 50% del nicho ecológico actual se encontraron 12 taxones (*P. amblyosepalus*, *P. dumosus*, *P. nelsonii*, *P. grayanus*, *P. talamancensis*, *P. xanthotrichus*, *P. acutifolius*, *P. costaricensis*, *P. polystachios*, *P. parvulus*, *P. oligospermus*, *P. maculatus maculatus*).

Por su parte, cuando fue comparada el área del nicho ecológico predicho con condiciones bioclimáticas presentes con las de cambio climático y un escenario de movilidad, 15 taxones incrementaron el área total de su nicho en condiciones de cambio climático con relación a las condiciones actuales y en la Tabla 12 (Columna E), tienen una razón igual o superior a 1. Incluso, en el caso de los taxones *P. leptostachyus leptostachyus*, *P. maculatus*, *P. acutifolius acutifolius* y *P. neglectus*, duplicaron el área modelada para su nicho ecológico, sin perder su nicho actual como fue nombrado anteriormente. Para los restantes 19 taxones, la razón entre el área del nicho ecológico en condiciones de cambio climático en el caso de presentarse movilidad y el nicho en el presente, fue menor a uno, lo que indica que el nicho ecológico para estas especies se prevé va a ser menor en condiciones de cambio climático en cualquiera de los escenarios de movilidad.

El porcentaje de nuevas áreas con potencial de constituirse en parte del nicho ecológico de cada taxón en condiciones de cambio climático, está representado en la columna F de la Tabla 12 y determina en cada taxón el porcentaje de nuevas áreas a ser incluidas dentro del nicho ecológico de la especie en el futuro. Por ejemplo, en el caso de *P. zimapanensis*, aunque el nicho ecológico incrementa aproximadamente un 40% con relación al nicho ecológico actual (Razón Columna E de 1.4), el 60% del área del nicho bajo condiciones de cambio climático corresponde a nuevas áreas que según las condiciones bioclimáticas actuales no son adecuadas para la especie. Por lo cual, la posibilidad de ocupar efectivamente esas áreas dependerá de la capacidad de movilidad de las especies.

Para visualizar los resultados en términos de la comparación de tamaños de los nichos, en la Figura 6, se representa el tamaño de los nichos en condiciones actuales (Eje X) y en condiciones de cambio climático (Eje Y), donde en el eje Y se adicionaron dos series,

correspondientes al tamaño de los nichos cuando se asumió que las comunidades de los taxones pueden migrar (asteriscos de color rojo) y cuando no lo pueden hacer (cruces de color azul), junto a la diagonal principal, la cual indica que los puntos que se encuentren arriba de la misma tienen un nicho ecológico mayor en condiciones de cambio climático que en las condiciones actuales, adicionalmente se coloca el número del taxón (Ver Figura 6), junto al asterisco rojo. De esta forma, por una parte se puede observar que los valores de las áreas en miles de km^2 , siendo mayores para el caso de los nichos predichos con condiciones de migración (rojo), las diferencias tienen diferente magnitud de acuerdo a cada taxón y por otra parte, se pueden visualizar con facilidad los taxones para los cuales el tamaño del nicho incrementa con relación al actual, los cuales se nombraron anteriormente, teniendo una razón entre el nicho en condiciones de cambio climático y el actual mayor a 1.

Figura 6— Representación del tamaño de los nichos ecológicos de 34 taxones del genero *Phaseolus* bajo condiciones actuales y de cambio climático y dos escenarios de capacidad migratoria.

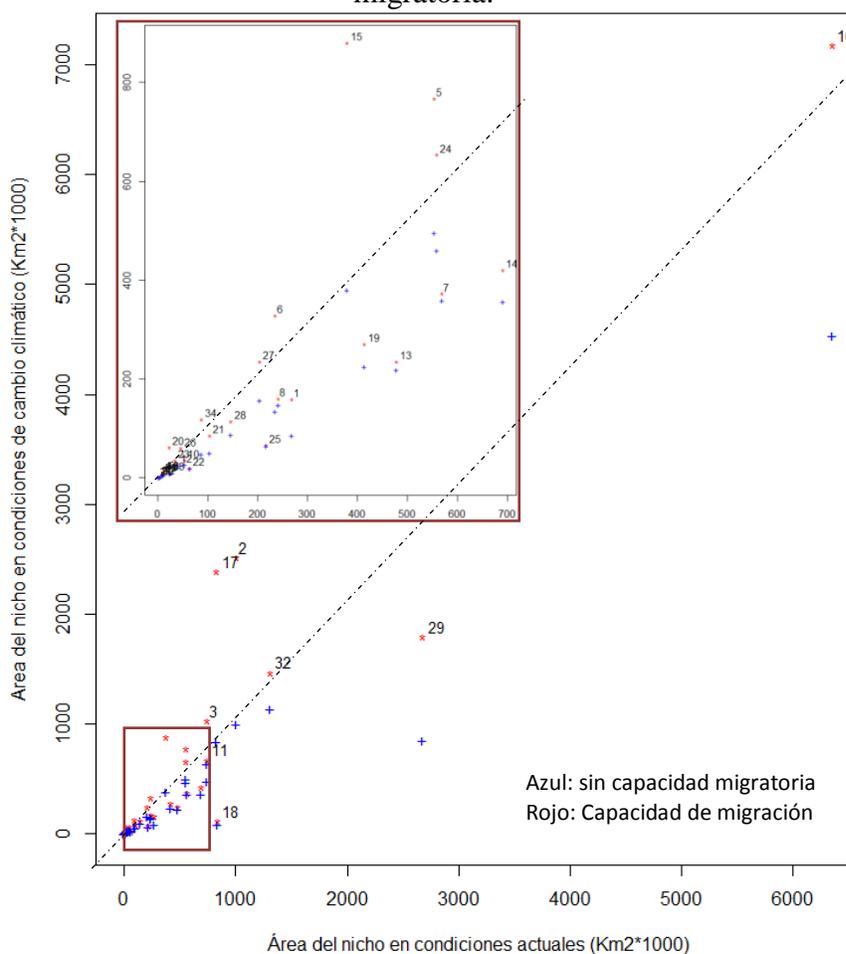


Tabla 12 – Comparación de los tamaños de los nichos ecológicos en condiciones presentes y de cambio climático bajo los escenarios de no migración y migración para 34 taxones del genero *Phaseolus*.

Taxon	A	B	C	D	E	F
	Km2*1000	Km2*1000	%	Km2*1000	(D/A)	%
1 - <i>P. acutifolius</i>	269.1	86.3	67.9	159.2	0.59	45.8
2 - <i>P. acutifolius_acutifolius</i>	1008.2	1000.4	0.8	2515.3	2.49	60.2
3 - <i>P. acutifolius_tenuifolius</i>	747.8	635.2	15.1	1027.9	1.37	38.2
4 - <i>P. amblyosepalus</i>	11.9	6	50	12.2	1.03	51.2
5 - <i>P. angustissimus</i>	554.2	495.1	10.7	768	1.39	35.5
6 - <i>P. augusti</i>	235.4	134.9	42.7	328.9	1.40	59
7 - <i>P. coccineus</i>	568.5	358.9	36.9	373	0.66	3.8
8 - <i>P. coccineus_coccineus</i>	241.1	147	39	161.1	0.67	8.8
9 - <i>P. costaricensis</i>	5.2	1.7	68.4	3.1	0.60	46.8
10 - <i>P. dumosus</i>	53.5	26.2	51.1	36.2	0.68	27.7
11 - <i>P. filiformis</i>	740.4	479	35.3	668.8	0.90	28.4
12 - <i>P. glabellus</i>	39.6	21.8	44.9	27.5	0.69	20.7
13 - <i>P. grayanus</i>	478.1	219	54.2	236	0.49	7.2
14 - <i>P. leptostachyus</i>	690.7	355.9	48.5	420.4	0.61	15.3
15 - <i>P. leptostachyus_leptostachyus</i>	379.5	379.5	0	880.5	2.32	56.9
16 - <i>P. lunatus</i>	6352.2	4533.5	28.6	7183	1.13	36.9
17 - <i>P. maculatus</i>	832.3	832.3	0	2387.6	2.87	65.1
18 - <i>P. maculatus_maculatus</i>	839.7	82.9	90.1	113	0.13	26.6
19 - <i>P. maculatus_ritensis</i>	414	225.4	45.6	271.1	0.65	16.8
20 - <i>P. neglectus</i>	23	22.6	1.7	62.1	2.70	63.7
21 - <i>P. nelsonii</i>	104.1	50.7	51.3	85.6	0.82	40.8
22 - <i>P. oligospermus</i>	64.2	18.6	71	20.1	0.31	7.3
23 - <i>P. pachyrrhizoides</i>	34.5	27.4	20.4	36.3	1.05	24.3
24 - <i>P. parvifolius</i>	558.2	460.4	17.5	654	1.17	29.6
25 - <i>P. parvulus</i>	216.8	64.5	70.3	65.9	0.30	2.2
26 - <i>P. pauciflorus</i>	46.7	25	46.4	60.3	1.29	58.5
27 - <i>P. pedicellatus</i>	203.8	156.9	23	235.5	1.16	33.4
28 - <i>P. pluriflorus</i>	146.9	87.2	40.7	114.9	0.78	24.2
29 - <i>P. polystachios</i>	2680.5	842.3	68.6	1789.7	0.67	52.9
30 - <i>P. talamancensis</i>	2.3	0.8	65.6	0.8	0.35	0
31 - <i>P. tuerckheimii</i>	9.1	4.7	48.6	8.2	0.90	43.5
32 - <i>P. vulgaris</i>	1306.5	1135.2	13.1	1466.8	1.12	22.6
33 - <i>P. xanthotrichus</i>	24.2	8.1	66.6	11.5	0.48	29.5
34 - <i>P. zimapanensis</i>	87.7	47.2	46.1	119.2	1.36	60.4

- A. Área del nicho ecológico en condiciones actuales
 B. Área del nicho ecológico bajo condiciones de cambio climático y sin capacidad de migración
 C. Porcentaje de área del nicho que se pierde bajo el escenario de no migración
 D. Área del nicho ecológico en condiciones de cambio climático y el escenario de migración
 E. Razón entre el área del nicho en condiciones de cambio climático y migración y el nicho en el presente (D/A)
 F. Porcentaje del nicho en condiciones de cambio climático que corresponde a nuevas áreas para el taxón

Adicionalmente, con la finalidad de describir la distribución espacial de los nichos con las condiciones bioclimáticas actuales y de cambio climático bajo los escenarios de no migración y de migración, en la **Figura 7** se presentan las comparaciones de los nichos ecológicos de los

34 taxones incluidos en el estudio. En cada mapa y con la finalidad de visualizar los nichos ecológicos de cada taxón, con color morado y naranja se representa el nicho ecológico en las condiciones bioclimáticas actuales. Así en color Morado se representa las áreas que fueron predichas por Maxent únicamente para las condiciones bioclimáticas actuales. Por su parte el color naranja, representa las áreas del nicho ecológico que fue modelado por Maxent tanto para las condiciones actuales como para las condiciones de cambio climático, estas áreas además coinciden con las áreas del nicho en condiciones de cambio climático cuando se asume que las comunidades de cada taxón no tendrán ninguna capacidad de migración con relación al nicho actual. Finalmente, con color azul y naranja, se representa el área del nicho ecológico de los taxones en condiciones de cambio climático y el escenario de migración de los taxones, aclarando que las áreas azules son las nuevas posibles áreas a donde potencialmente podrían migrar los taxones.

De esta forma, teniendo en cuenta la distribución espacial de los datos de ocurrencia de los taxones utilizados para el modelamiento en Maxent (Figura 7, estrellas rojas), los taxones *P. acutifolius*, *P. acutifolius acutifolius* y *P. acutifolius tenuifolius* (Figura 7 -1, -2 y -3, respectivamente), tienen una distribución que va desde los estados de Arizona y Nuevo México, hasta los estados de Michoacán y Guerrero, principalmente a lo largo de la Sierra Madre Occidental. Para estos, el nicho de mayor tamaño en todas las condiciones fue el de *P. acutifolius acutifolius*, para el que el área fue más del doble con relación al actual y con muy bajo porcentaje de pérdidas de nicho con relación al presente (Tabla 12), teniendo la particularidad de presentar una distribución de ocurrencias amplia que incluye ocurrencias a lo largo de la sierra madre occidental, el altiplano Mexicano y algunos puntos sobre la Sierra Madre Oriental, especialmente en algunas zonas de altitud media de los estados de Durango, Chihuahua y Texas, los cuales son zonas de menor humedad y con algunos ambientes áridos. Por su parte, el taxón de mayor pérdida de nicho con relación al actual (Color morado, Figura 7-1), fue *P. acutifolius*, el cual tiene una distribución más concentrada, especialmente sobre el norte de la Sierra Madre Occidental y algunas ocurrencias en zonas de menores latitudes.

Por su parte, para otros taxones con distribución principalmente sobre la parte norte de la Sierra Madre Occidental, la distribución del nicho ecológico es de forma similar a lo ocurrido con *P. acutifolius*, es decir para los taxones *P. grayanus*, *P. maculatus maculatus*, *P. maculatus ritensis*, *P. parvulus* (Figura 7-13, -18, -19 y -25, respectivamente) en donde se presenta una

perdida mayor al 50% del nicho en condiciones de cambio climático con relación al nicho actual, a excepción del nicho de *P. maculatus ritensis*, en donde la pérdida es del 45% y las nuevas áreas cuando se tuvo en cuenta la capacidad migratoria de los taxones no superó al 26% del nicho del taxón en esas condiciones. Se debe destacar otro taxón con distribución sobre el mismo sector que es *P. maculatus*, aunque en partes bajas de la Sierra Madre Occidental (Figura 7-17), el cual en términos de la proporción de áreas tiene un comportamiento similar a la de *P. acutifolius acutifolius*, sin área perdida del nicho actual y más del doble del área de nicho en condiciones de cambio climático con relación al área actual. Sin embargo, llama la atención la comparación entre los nichos, la cantidad de datos de ocurrencia y la configuración espacial de los nichos de *P. maculatus*, *P. maculatus ritensis* y *P. maculatus maculatus*, los cuales tienen 10, 105 y 130 datos de ocurrencia respectivamente, situados sobre la Sierra Madre Occidental y con menor concentración de las observaciones a medida que incrementa el número de muestras, para los cuales a medida que incrementaron el número de ocurrencias incrementó el área perdida o aquella que solo se predice en condiciones actuales.

Otros 12 taxones tuvieron puntos de ocurrencia principalmente entre el sur de los Estados Unidos y México los cuales son *P. amblyosepalus*, *P. angustissimus*, *P. coccineus*, *P. coccineus coccineus*, *P. filiformis*, *P. glabellus*, *P. neglectus*, *P. nelsonii*, *P. parvifolius*, *P. pauciflorus*, *P. pedicellatus*, *P. pluriflorus*, *P. zimapanensis* (Figura 7 -4, -5, -7, -8, -11, -12, -20, -21, -24, -26, -27, -28 y -34, respectivamente) de estos los que se distribuyen principalmente a lo largo de la Sierra Madre Oriental son *P. coccineus coccineus*, *P. glabellus* y *P. zimapanensis*, los cuales pierden área de nicho ecológico en ambos escenarios de migración, a excepción de *P. zimapanensis* en el caso de que pueda migrar de su nicho actual, adicionalmente estos tienen nuevas áreas potenciales en la Sierra Madre Occidental. Por su parte, *P. amblyosepalus*, *P. neglectus* y *P. pauciflorus*, tienen ocurrencias concentradas en alguna de las sierras y nichos relativamente pequeños en las condiciones actuales (11.9, 23 y 46.7 miles de km², respectivamente) y las nuevas áreas son más del 50% en el caso de que la especie tenga la capacidad de migrar, zonas que se concentran hacia el sur de su nicho actual. Dentro de este mismo grupo, se presentaron nichos que se perciben van a ser fragmentados y las nuevas áreas se encuentran en los extremos opuestos de su distribución actual, como por ejemplo los nichos de *P. filiformis*, *P. nelsonii*, *P. pedicellatus* y *P. pluriflorus*, mientras otros

de acuerdo a los resultados del modelamiento conservan su mayor núcleo como es el caso de *P. angustissimus*, *P. coccineus*, *P. parvifolius*.

Por su parte, seis de los taxones estudiados tienen una distribución localizada principalmente en Centro América los cuales son *P. costaricensis*, *P. dumosus*, *P. oligospermus*, *P. talamancensis*, *P. tuerckheimii*, *P. xanthotrichus* (Figura 7-9, -10, -22, -30, -31 y -33, respectivamente), para los cuales el nicho ecológico tiene un comportamiento muy característico, los cuales son nichos pequeños tanto en el presente como en el futuro en cualquiera de los escenarios de migración, los cuales nunca superan los 70 mil Km² y en todos los casos el nicho es menor en las condiciones de cambio climático con relación al actual. Adicionalmente, espacialmente sus nichos lucen con mayor fragmentación. Por otra parte un taxón con gran cantidad de ocurrencias localizadas, fue *P. polystachios* (Figura 7-29), el cual se distribuye en los estados del suroccidente de los Estados Unidos y sus nichos tienen la particularidad de perder una gran parte de su área actual, especialmente al sur de su distribución actual.

Adicionalmente, los taxones de *P. augusti* y *P. pachyrrhizoides* son los dos únicos que se distribuyen a lo largo de los Andes en Sur América exclusivamente (Figura 7-6 y 23). De los cuales, el primero es el de mayor tamaño del nicho en ambas condiciones, extendiéndose a lo largo de los Andes desde el sur de Ecuador, hasta el norte de Argentina, teniendo nuevas áreas de nicho en cambio climático que aparentemente permiten conectar zonas del nicho actual. Finalmente, los taxones de distribución más amplia a lo largo de la zona de estudio son *P. vulgaris*, *P. lunatus*, *P. leptostachyus* y *P. leptostachyus leptostachyus* (Figura 7-32, -16, -14 y -15, respectivamente). Los cuales a excepción de *P. leptostachyus* incrementan el tamaño de su nicho en condiciones de cambio climático y migración, llamando la atención que en este grupo de taxones, el área perdida con relación al área actual nunca supera el 50% del nicho original, incluso fue inferior al 29% para todos los taxones a excepción del caso de *P. leptostachyus*.

Figura 7– Comparación de los nichos ecológicos de 34 taxones del genero *Phaseolus* predichos con Maxent para las condiciones bioclimáticas actuales y de cambio climático.

Nota: Debido a la superposición realizada de los nichos la siguiente es la configuración de áreas de los nichos obtenidos:

Nicho Actual= Sumatoria de áreas de color naranja y morado (Color morado zonas en donde el nicho perderá área)

Nicho en condiciones de cambio climático y un escenario de no migración= áreas de color Naranja (Zonas en las que coincide el nicho actual y futuro sin migración)

Nicho en condiciones de cambio climático y capacidad de migración de 300 km= Sumatoria de áreas de color naranja y azul (Las áreas de color azul representan las nuevas áreas en un escenario de cambio climático con aptitud para la presencia del taxón).

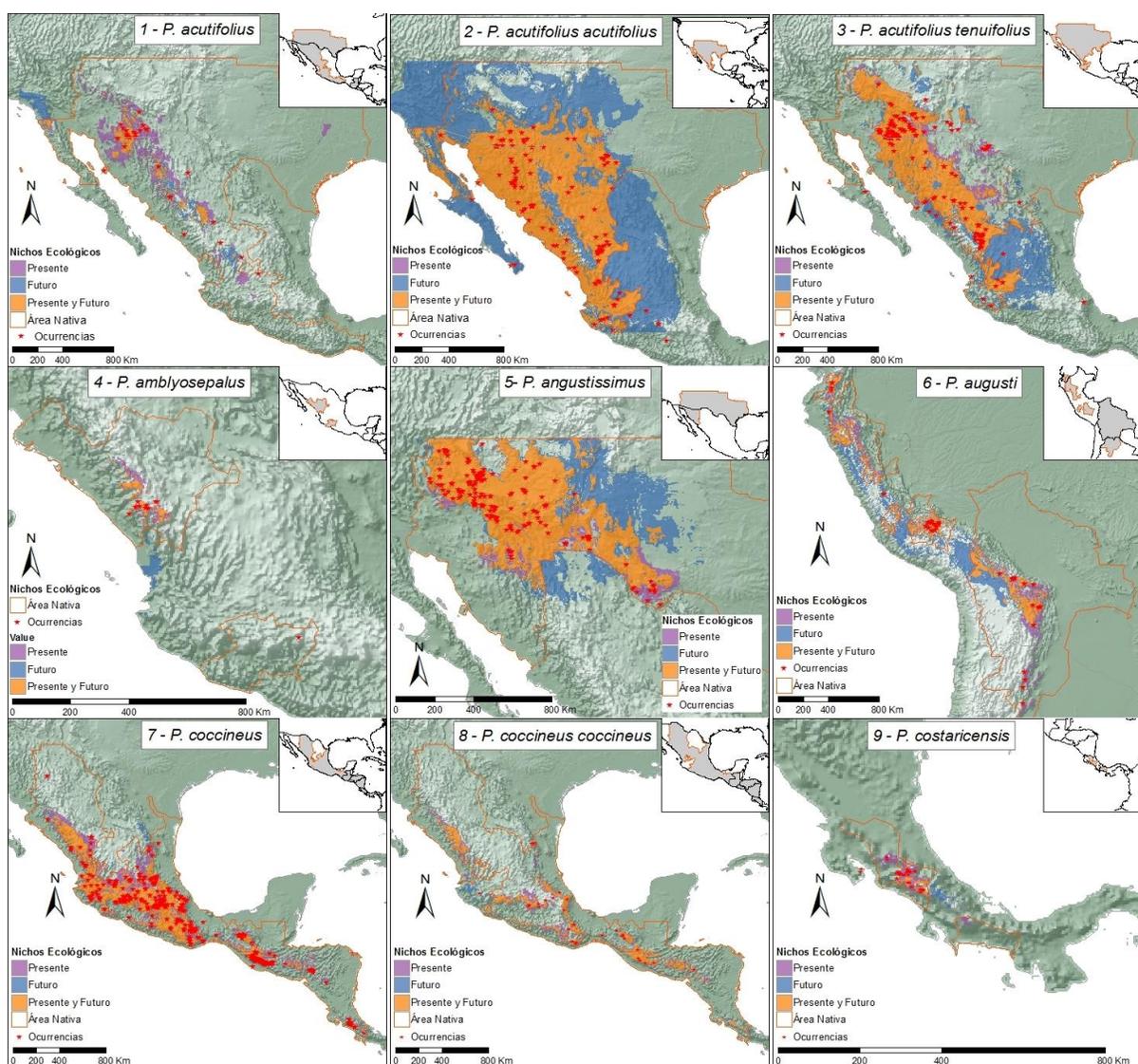


Figura 7 (Continuación)

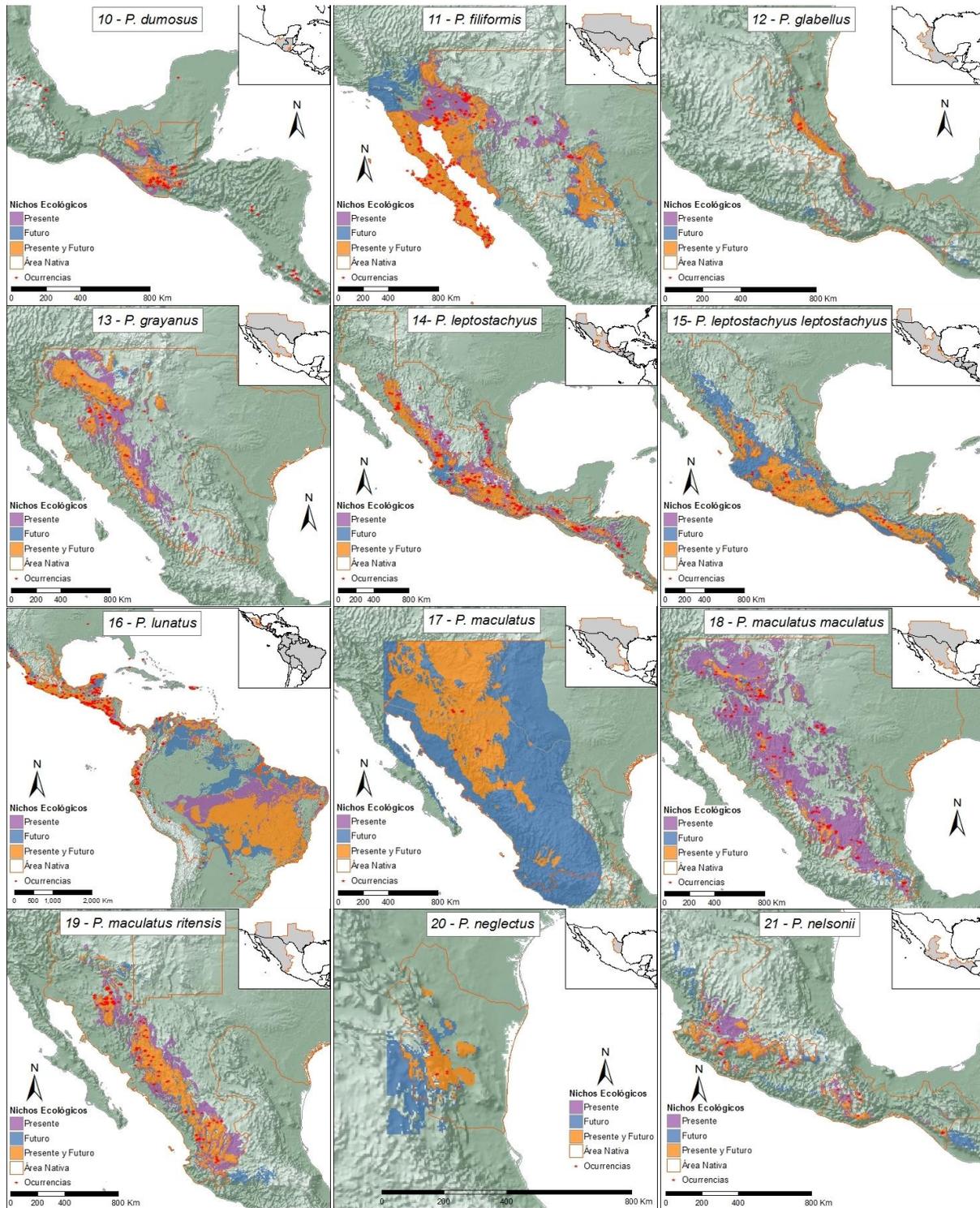


Figura 7 (Continuación)

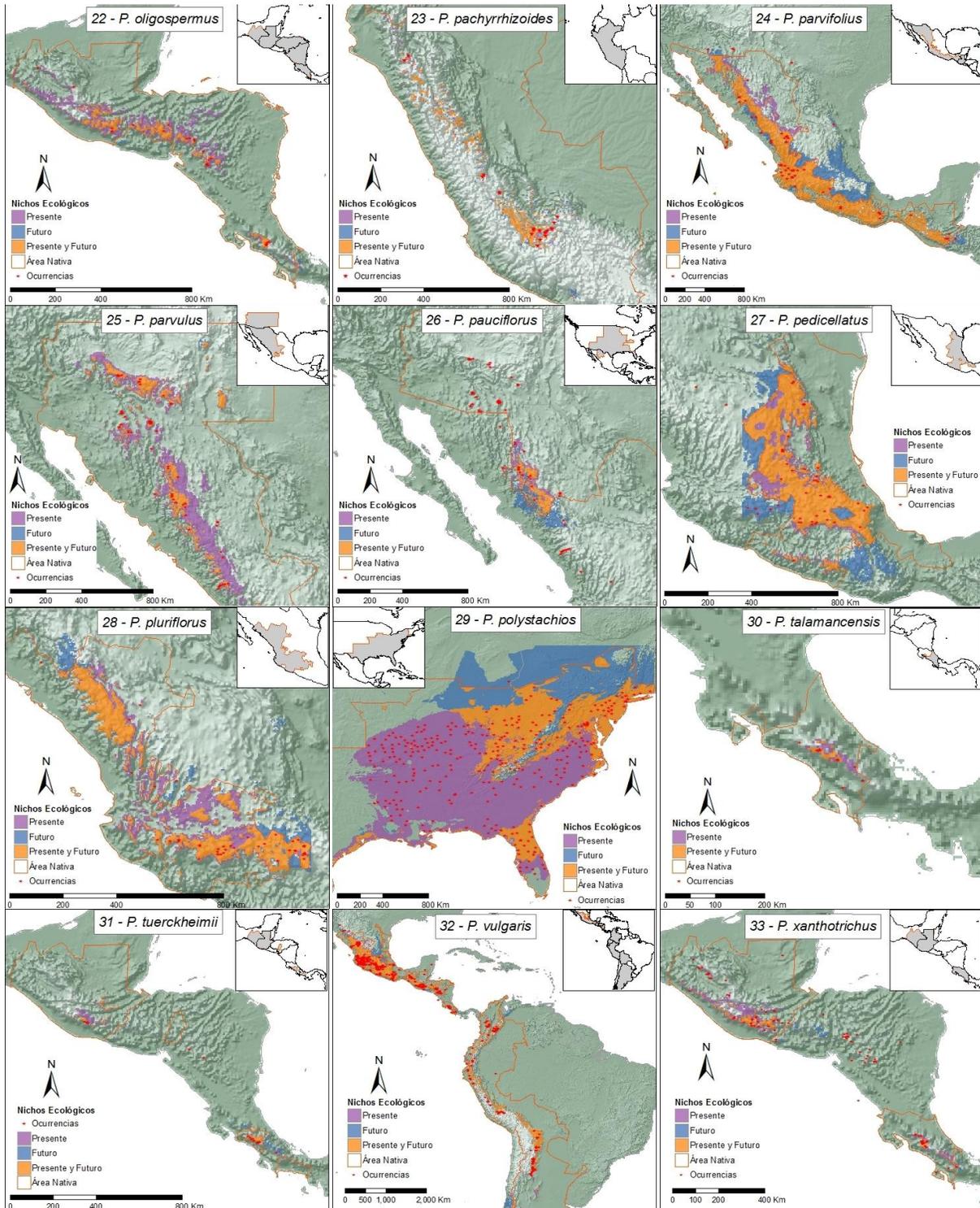
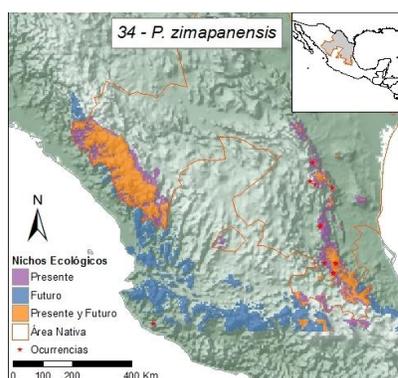


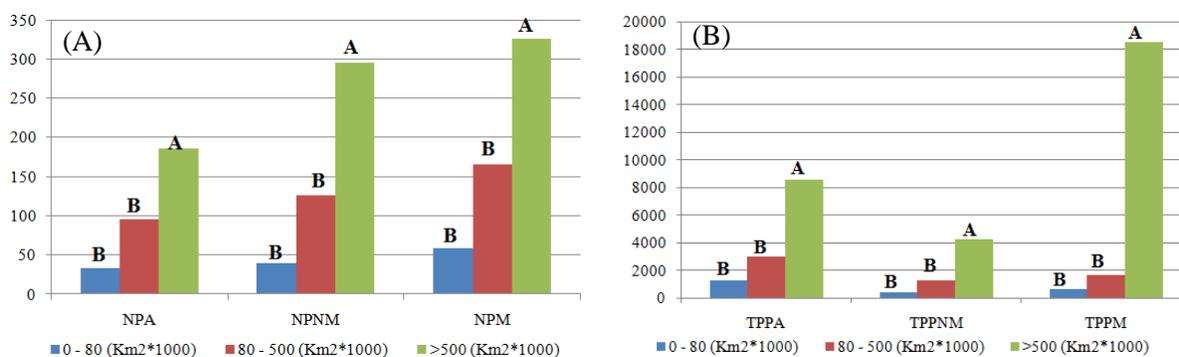
Figura 7 (Continuación)



5.8.2 Análisis de fragmentación de nichos ecológicos para las condiciones actuales y de cambio climático: Fue analizada la configuración de los nichos de cada taxón en condiciones actuales y de cambio climático en términos del número de parches que constituye cada taxón y su tamaño promedio en km² (Tabla 13). De esta forma, cuando fueron agrupados los taxones en terciles de acuerdo al tamaño del nicho ecológico en las condiciones actuales y adicionalmente fue calculado el número de parches y su tamaño promedio, se pudo observar para los tres escenarios generados (presente, cambio climático sin movilidad y con movilidad) que el número de parches y su tamaño promedio (Figura 8-A) se incrementaron con el tamaño promedio de los taxones, siendo en los tres casos el número de parches de los taxones con nicho mayor a 500 mil km² mayor estadísticamente ($p < 0.05$). Adicionalmente, al comparar la configuración del nicho ecológico para cada agrupación de taxones a través de los escenarios, se puede observar que para los taxones con nichos ecológicos menores a 80 mil km² (*P. amblyosepalus*, *P. costaricensis*, *P. dumosus*, *P. glabellus*, *P. neglectus*, *P. oligospermus*, *P. pachyrrhizoides*, *P. pauciflorus*, *P. talamancensis*, *P. tuerckheimii*, *P. xanthotrichus*) y para los que tienen un nicho entre 80 y 500 mil km² (*P. acutifolius*, *P. augusti*, *P. coccineus coccineus*, *P. grayanus*, *P. leptostachyus leptostachyus*, *P. maculatus ritensis*, *P. nelsonii*, *P. parvulus*, *P. pedicellatus*, *P. pluriflorus*, *P. zimapanensis*), a medida que el número de parches va en aumento a través de los tres escenarios con 33, 40 y 59 parches para el primer grupo y con 96, 127 y 166 para el segundo (Figura 8A), su tamaño promedio disminuye con relación al tamaño en las condiciones del presente con 1257 km² en condiciones presentes para el primer grupo en comparación con 427 y 664 en condiciones de cambio climático sin

movilidad y con movilidad, y de igual manera se presenta para el segundo grupo con 3024 miles de km^2 en comparación con los 1285 y 1685 miles de km^2 para las condiciones de cambio climático (Figura 8-B). Lo que implica, especialmente para el grupo con nichos más pequeños que estos son cada vez más segmentados, incluso, el área que tienen de ganancia o nuevas áreas con condiciones bioclimáticas favorables están más desconectadas, lo que puede dificultar la migración y su establecimiento. Mientras que para el caso de los taxones con nichos más grandes en las condiciones actuales (*P. acutifolius acutifolius*, *P. acutifolius tenuifolius*, *P. angustissimus*, *P. coccineus*, *P. filiformis*, *P. leptostachyus*, *P. lunatus*, *P. maculatus*, *P. maculatus maculatus*, *P. parvifolius*, *P. polystachios*, *P. vulgaris*) aunque tienen un comportamiento similar en términos del aumento en el número de parches con 186, 295 y 326, el tamaño promedio de los parches, aunque disminuye en condiciones de cambio climático cuando se asume la no migración (1285 miles de km^2 en comparación con los 8600 miles de km^2 del tamaño promedio en condiciones presentes), luego aumenta sustancialmente, cuando se asume migración con 18533 miles de km^2 , lo que indica que la configuración de su nicho se prevé más estable y el establecimiento en nuevas áreas dependerá de su propia capacidad de migración y no de barreras físicas debidas a la discontinuidad de su nicho.

Figura 8– Comparación de promedios del número de parches (A) y tamaño promedio de los parches (B) para los taxones agrupados en terciles de acuerdo al tamaño del nicho en condiciones actuales.



Un análisis similar fue realizado agrupando los taxones de acuerdo al porcentaje de nueva área que conforma el nicho ecológico de cada taxón en condiciones de cambio climático y asumiendo un escenario de movilidad, siendo los taxones que conforman el primer grupo aquellos que en condiciones de cambio climático tienen poca área con aptitud bioclimática

para establecer su nicho, es decir los que tienen nuevas áreas menores a 20% del nicho (*P. coccineus*, *P. coccineus coccineus*, *P. grayanus*, *P. leptostachyus*, *P. maculatus ritensis*, *P. oligospermus*, *P. parvulus*, *P. talamancensis*), el segundo grupo los que tienen nuevas áreas entre el 20 y 50% (*P. acutifolius*, *P. acutifolius tenuifolius*, *P. angustissimus*, *P. costaricensis*, *P. dumosus*, *P. filiformis*, *P. glabellus*, *P. lunatus*, *P. maculatus maculatus*, *P. nelsonii*, *P. pachyrrhizoides*, *P. parvifolius*, *P. pedicellatus*, *P. pluriflorus*, *P. tuerckheimii*, *P. vulgaris*, *P. xanthotrichus*) y finalmente en el tercer grupo los que tienen nuevas áreas mayores al 50% (*P. acutifolius acutifolius*, *P. amblyosepalus*, *P. augusti*, *P. leptostachyus leptostachyus*, *P. maculatus*, *P. neglectus*, *P. pauciflorus*, *P. polystachios*, *P. zimapanensis*), es decir, para los que el nicho lo conforman en su mayoría nuevas áreas. En este caso, no hubo ninguna diferencia estadística en términos del número promedio de parches cuando fue realizado el análisis de separación de medias dentro de cada escenario ($p > 0.05$, Letras A Figura 9-A), sin embargo, la tendencia en el número de parches fue a ser mayor para los taxones con nuevas áreas entre un 20 y 50% (128, 197 y 245 parches para las condiciones actuales, y de cambio climático sin y con movilidad, respectivamente), seguidos por los taxones con nuevas áreas menores al 20% (111, 155 y 174 parches) y los de menor número de parches fueron los taxones con nuevas áreas mayores al 50% (65, 87 y 91 parches), este comportamiento fue similar tanto para los modelos en las condiciones actuales y de cambio climático con y sin movilidad, presentándose además una tendencia hacia el incremento del número de parches cuando se paso de las condiciones actuales a las de cambio climático.

Por su parte, cuando se comparó el tamaño promedio de los nichos de acuerdo a esta agrupación de taxones (Figura 9-B), en las condiciones actuales no hubo diferencia estadísticas significativas ($p > 0.05$), aunque el tamaño promedio de los parches se incrementó cuando se paso de las condiciones actuales con 2749 km², a las de cambio climático sin movilidad con 3967 km² y con movilidad de los taxones con 6762 km². Por su parte, en condiciones de cambio climático y asumiendo la no movilidad de los taxones con relación a su nicho actual, el tamaño promedio de los parches fue inferior estadísticamente para los taxones con menor porcentaje de áreas nuevas en comparación con los promedios de los taxones con áreas nuevas entre el 20 y 50% y mayor al 50% (932 km² en comparación con 1889 y 3360 km²). Mientras en el caso del tamaño promedio de los parches cuando se asumió movilidad, el promedio del tamaño fue mayor estadísticamente ($p < 0.01$) para los taxones con áreas nuevas

superiores al 50% en comparación con las agrupaciones con nuevas áreas menores al 20% y entre el 20 y 50% (23360 km² en comparación con 920 y 1803 km²). Esto implica que para los taxones con nuevas áreas mayores al 50% sus nichos van a estar más consolidados con relación a los demás en condiciones de cambio climático asumiendo que puedan migrar a esas nuevas áreas.

Figura 9– Comparación de promedios del número de parches (A) y tamaño promedio de los parches (B) para los taxones agrupados de acuerdo al porcentaje de nuevas áreas que conformarían sus nichos en condiciones de cambio climático y migración.

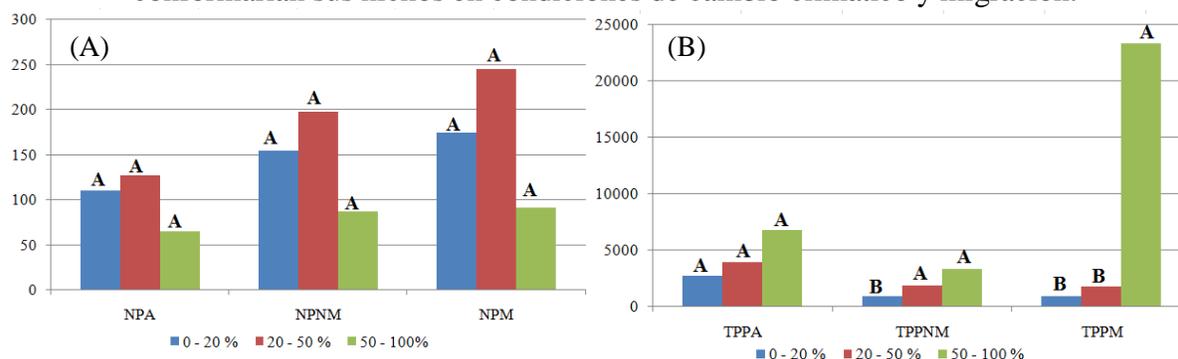


Tabla 13– Número (NP) y tamaño promedio de los parches (TPP) que conforman los nichos ecológicos de 34 taxones del genero *Phaseolus* bajo condiciones actuales y de cambio climático sin y con capacidad de migración.

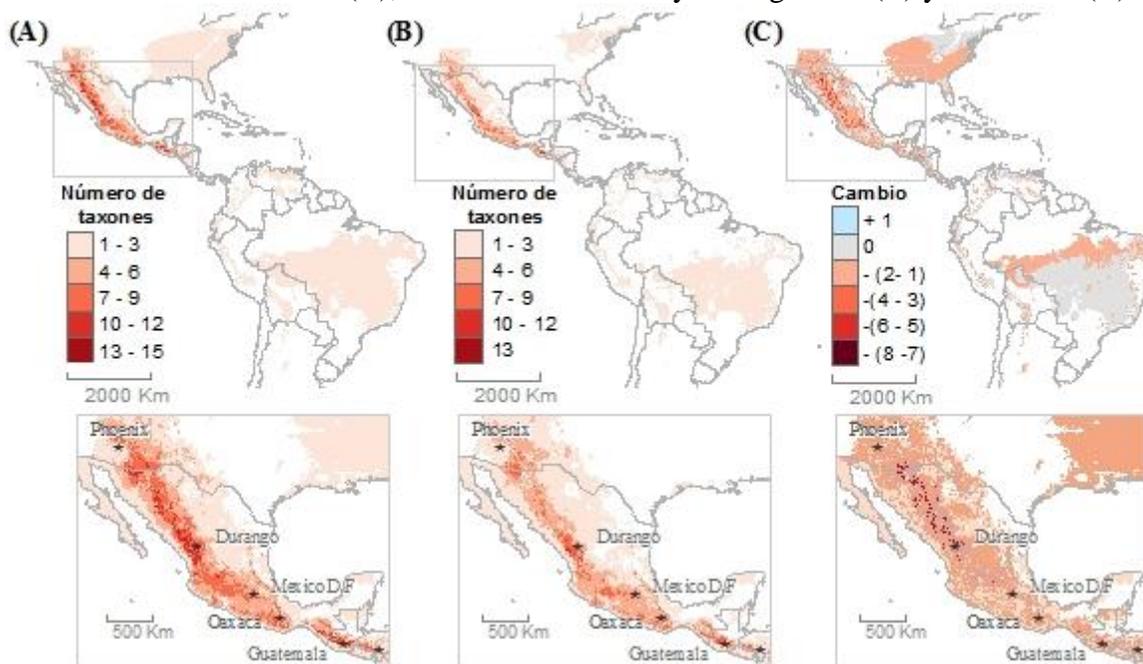
Taxón	Condiciones actuales		Cambio climático sin capacidad de migración		Cambio climático con capacidad de migración	
	NP	TPP (Km2)	NP	TPP (Km2)	NP	TPP (Km2)
<i>P. acutifolius</i>	159	1705.3	144	599.3	162	982.7
<i>P. acutifolius acutifolius</i>	181	5564.6	198	5052.3	97	25930.9
<i>P. acutifolius tenuifolius</i>	211	3540.4	239	2657.5	403	2550.6
<i>P. amblyosepalus</i>	19	647.4	26	228.8	44	277.3
<i>P. angustissimus</i>	62	8932.3	78	6347.4	125	6143.8
<i>P. augusti</i>	52	4544.2	111	1215.5	144	2284.0
<i>P. coccineus</i>	136	4179.2	206	1742.1	234	1594.0
<i>P. coccineus coccineus</i>	171	1403.5	223	659.1	274	588.0
<i>P. costaricensis</i>	3	1658.3	6	275.0	15	206.7
<i>P. dumosus</i>	21	2576.2	42	622.6	66	547.7
<i>P. filiformis</i>	72	10246.5	104	4605.3	197	3395.1
<i>P. glabellus</i>	26	1546.2	43	507.0	73	376.7
<i>P. grayanus</i>	143	3356.6	154	1422.2	171	1380.0
<i>P. leptostachyus</i>	126	5486.7	256	1390.0	263	1598.4
<i>P. leptostachyus leptostachyus</i>	104	3657.7	106	3580.2	126	6987.7
<i>P. lunatus</i>	629	10111.2	1287	3522.5	1411	5090.7
<i>P. maculatus</i>	68	12193.8	68	12239.3	15	159171.7
<i>P. maculatus maculatus</i>	224	3753.0	463	300.8	298	379.1
<i>P. maculatus ritensis</i>	109	3791.3	165	1366.1	202	1341.8
<i>P. neglectus</i>	19	1207.9	18	1252.8	16	3881.3
<i>P. nelsonii</i>	47	2217.6	111	456.3	191	448.2
<i>P. oligospermus</i>	66	978.0	63	295.2	72	278.8
<i>P. pachyrrhizoides</i>	139	252.7	139	197.3	193	187.8
<i>P. parvifolius</i>	71	7853.5	90	5115.8	171	3824.6
<i>P. parvulus</i>	130	1672.5	165	390.8	174	378.9
<i>P. pauciflorus</i>	19	2431.6	34	736.0	68	886.4
<i>P. pedicellatus</i>	35	5850.0	58	2704.7	96	2452.6
<i>P. pluriflorus</i>	71	2067.6	102	854.4	152	755.9
<i>P. polystachios</i>	97	27621.1	167	5043.7	181	9888.0
<i>P. talamancensis</i>	2	1125.0	4	193.8	4	193.8
<i>P. tuerckheimii</i>	22	419.3	23	202.2	28	293.8
<i>P. vulgaris</i>	351	3727.9	384	2956.2	517	2837.0
<i>P. xanthotrichus</i>	24	985.4	42	192.3	65	176.2
<i>P. zimapanensis</i>	28	2993.8	53	891.0	127	938.8

5.8.3 Análisis de riqueza de especies silvestres del género *Phaseolus*: En la Figura 10 se muestra el análisis de riqueza de especies del genero *Phaseolus*, de acuerdo al modelamiento de nichos ecológico realizado en Maxent. En la Figura, se muestra la comparación de la riqueza en las condiciones actuales (Figura 10-A), en condiciones de cambio climático asumiendo un escenario de no migración (Figura 10-B) y los cambios en la riqueza (Figura 10-C), es decir, muestra la cantidad de taxones que se pierden o se ganan en condiciones de

cambio climático y no migración. Adicionalmente, debajo de cada figura se muestra la zona ampliada de México, Guatemala y Honduras, en donde resultó una mayor riqueza de taxones en el presente estudio.

En la Figura 10A, se puede observar como la mayor riqueza de taxones en las condiciones actuales con entre 4 y 15 taxones (4 categorías de colores rojos oscuros), se encontraron a lo largo de la Sierra Madre Occidental desde los estados de Arizona en los Estados Unidos y pasando por los estados de Sonora, Chihuahua, Durango, Nayarit, Jalisco y Michoacán, además sobre la sierra madre sur en los estados de Guerrero y Oaxaca y luego en centro América sobre la misma cadena montañosa en Guatemala, Honduras y el norte de Nicaragua y la parte central en Costa Rica. De estas zonas, las de mayor importancia en términos de la riqueza predicha de especies resultaron ser en la parte norte y sobre la Sierra Madre Occidental, la parte centro norte del estado de Sonora, junto con el occidente de los estados de Chihuahua y Durango, en donde se concentran zonas con la presencia de nichos de más de 10 taxones (Figura 10-A Ampliación). La segunda zona con presencia de nichos de más de 10 taxones en las condiciones bioclimáticas actuales se encuentra localizada en la sierra madre sur entre el sur oriente del estado de Chiapas y siguiendo por la misma cadena montañosa atravesando todo Guatemala de occidente a oriente.

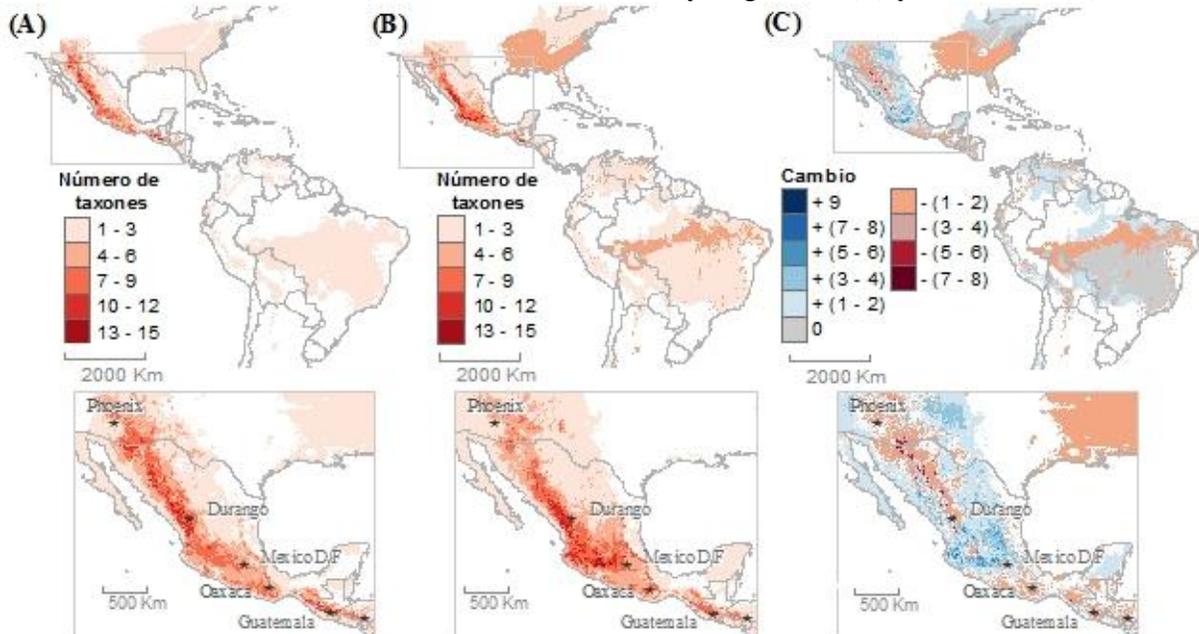
Figura 10- Riqueza de taxones del genero *Phaseolus* modelada en Maxent para las condiciones actuales (A), de cambio climático y no migración (B) y su cambio (C).



Por su parte, cuando fueron modelados los nichos ecológicos y se asumió la no migración de las especies desde su nicho actual, en la Figura 10-B, se muestra la distribución espacial de la riqueza de especies, en donde se hace evidente la disminución de las áreas con mayor número potencial de especies, especialmente disminuyen las áreas de las zonas con mayor número de especies, localizadas entre Chihuahua y Durango y entre Chiapas y Guatemala (Figura 10-B Ampliación). Los cambios entre el número de especies presenten en ambos escenarios bioclimáticos y asumiendo la no migración de los taxones, se hace evidente en la Figura 10-C, en donde con colores rojos de mayor intensidad se muestra el mayor número de taxones perdidos, con una alta concentración de pérdidas (Entre 5 y 8, especies perdidas) especialmente a lo largo de la Sierra Madre Occidental desde Sonora hasta el sur de Durango, pero en la parte oriental de la zona de mayor riqueza de especies en condiciones actuales (Figura 10-C Ampliación), para las demás zonas de México la pérdida de riqueza o taxones estuvo entre los 4 y 1 especie.

Así mismo en la Figura 11, se muestra la comparación de riqueza cuando se estimaron los nichos ecológicos en condiciones de cambio climático y se asumió la capacidad de migración de las especies. En este caso la Figura 11-A muestra el mismo mapa de la Figura 10-A, que al ser comparado con el mapa de la Figura 11-B, se puede observar como la riqueza de especies se incrementa a lo largo de la Sierra Madre Occidental, pero esta vez especialmente desde el estado de Durango alrededor de la zona de mayor riqueza actual y hacia el sur en los estados de Nayarit, Jalisco y Michoacán. Así, las zonas en donde se presentó un mayor cambio hacia la pérdida de biodiversidad (Figura 11-C), se mantuvo sobre la Sierra Madre Occidental hacia el norte y las zonas con mayor ganancia hacia el sur de la mayor zona de riqueza actual, siempre y cuando los taxones tengan la posibilidad de migrar hacia zonas más altas de la sierra en la parte sur de la misma.

Figura 11- Riqueza de taxones del genero *Phaseolus* modelada en Maxent para las condiciones actuales (A), de cambio climático y migración (B) y su cambio (C).



6. DISCUSIÓN

Los resultados presentados en este estudio se enfocaron en dos aspectos primordiales: Primero, la selección de variables bioclimáticas y del terreno para la estimación y evaluación estadística de modelos de nicho ecológico usando el aplicativo Maxent (Phillips et al., 2011) y la selección de modelos de acuerdo a la selección de variables realizada (Ver Tabla 8). Segundo, en la utilización de los modelos con mejor ajuste y su proyección al 2050, para predecir los efectos del cambio climático en la distribución de 34 taxones del género *Phaseolus*, parientes silvestres del frijol común (*Phaseolus vulgaris* L.) (Ver Tablas 9 y 10).

6.1 Selección de variables y de modelos de nicho ecológico usando el principio de máxima entropía. En esta sección se discuten los resultados generados desde la generación de criterios para seleccionar variables para el modelamiento de nichos ecológicos en Maxent hasta la propia escogencia de los modelamientos a usar para evaluar los efectos del cambio climático sobre los parientes silvestres del frijol común.

6.1.1 Determinación de criterios para la selección de variables: El primer criterio a tener en cuenta en la selección de variables era determinar si estas eran indispensables para el modelamiento o representaban información redundante. Para ello, se realizó el análisis exploratorio de datos a las variables bioclimáticas en los puntos de ocurrencia de género *Phaseolus*, confirmándose la multicolinealidad entre varias de ellas, principalmente entre variables asociadas a la temperatura y/o a la precipitación y que por ejemplo miden la misma característica en periodos de tiempo distintos, colocando en evidencia que existen variables que aportan información similar para el modelamiento (Ver Figura 2). Así, en un modelo ajustado con la totalidad de estos predictores se podría estar incluyendo información redundante (Elith et al., 2011; Guisan & Thuiller, 2005). Sin embargo, la asociación entre variables o índices bioclimáticos no fue, en todos los casos, completamente lineal o con coeficientes de determinación cercanos a 1. Esta situación fue evidente, en especial, cuando se compararon variables que representaban condiciones bioclimáticas distintas (p.e. condiciones promedio anuales vs estacionalidad intra-anual en la temperatura), aunque entre algunas de

ellas existió la correlación (Ver Figura 2). Estas características de los datos tuvieron varias implicaciones en la selección de variables. En primer lugar, se tuvo en cuenta el enfoque de (Guisan & Thuiller, 2005) los cuales plantean la necesidad de remover algunas de las variables o predictores ambientales con alta correlación antes de entrar a realizar el modelamiento de la distribución de especies, con el fin de evitar el sobreajuste de los modelos, especialmente cuando se tienen menos datos que predictores ambientales como se presenta en algunos taxones del presente estudio (Elith et al., 2011; Guisan & Thuiller, 2005; Wiens et al., 2009). Igualmente, varios estudios señalan que una de las formas de remover o incluir variables es el conocimiento de la(s) especies y/o las variables ambientales (Elith & Leathwick, 2009; Elith et al., 2011; Guisan & Thuiller, 2005; Wiens et al., 2009), por lo cual se hizo necesaria la inclusión, como criterio de selección, la significancia biológica de las variables bioclimáticas, definida conceptualmente por Nix (1986).

6.1.2 Selección de variables para el modelamiento de nichos ecológicos: Los resultados del análisis de componentes principales, en términos de la composición de los tres primeros vectores propios, permitieron describir de forma lógica las condiciones bioclimáticas de los sitios de ocurrencia de los taxones y realizar la posterior selección de variables (Ver Tabla 7). Así, la composición del primer vector propio permitió, por una parte, determinar que las condiciones de estacionalidad intra-anual en la temperatura, representadas por la participación significativa de las tres variables del grupo en el vector, son una condición de relevancia para explicar la variabilidad bioclimática en los datos. La inclusión de la variable bio7 (Rango anual de temperatura) que es la de mayor importancia relativa, permite representar la variabilidad e inercia que genera la estacionalidad intra-anual de la temperatura en los datos analizados. Por otra parte, el análisis de la participación de las variables en el sentido negativo del eje, permitió priorizar de forma consistente la selección de variables que representan condiciones bioclimáticas extremas tanto en temperatura como en precipitación sobre las variables que representaban condiciones promedio estacionales. Así se seleccionó la bio 6 (Temperatura mínima del mes más frío) sobre la bio11 (Temperatura mínima del trimestre más frío) y la bio13 (Precipitación del mes más húmedo) sobre la bio16 (Precipitación del trimestre más húmedo). Finalmente, al incluir la variable bio12 (Precipitación anual) que representa las condiciones promedio anuales en términos de precipitación, se configura este

eje como un balance entre variables de temperatura y precipitación que están asociadas con todas las condiciones bioclimáticas a excepción de las condiciones promedio estacionales (Ver Tabla 8).

En los ejes 2 y 3, se concentran principalmente las variables asociadas a las condiciones promedio estacionales, en el eje 2 con la participación de variables asociadas a la temperatura y en el 3 a la precipitación. Así en el eje 2, la selección fue realizada a favor de las bio9 (Temperatura media del trimestre más seco) y bio8 (Temperatura media del trimestre más húmedo) que representan condiciones promedio estacionales, junto a la bio1 (Temperatura media anual) y la bio5 (temperatura máxima del mes más caliente), que representan condiciones promedio anuales y de extremos, respectivamente. Esta última variable, fue incluida por encima de la bio10 (temperatura promedio del trimestre más caliente), a pesar que tuvo menor importancia relativa, debido a que la correlación lineal entre ambas fue altamente significativa ($r^2= 0.94$, $p<0.001$) y a que la condición extrema en términos de máxima temperatura no había sido incluida y puede definir, entre otras, la capacidad de un genotipo a estar adaptado a estas condiciones, las cuales pueden ser relevantes bajo un escenario de cambio climático. Finalmente, en el tercer eje fue seleccionada la bio17 (Precipitación del trimestre más seco), sobre la bio14 (precipitación del mes más seco), priorizando en este caso la condición promedio mensual sobre la extrema (Ver Tabla 8).

De esta forma, fue posible incluir la mayor cantidad de condiciones bioclimáticas con la menor cantidad de variables. Generando un grupo de predictores que en todos los casos es menor o al menos igual al número de observaciones de cada táxon, es decir 10, evitando así el sobreajuste de los modelos al menos por este factor (Elith & Leathwick, 2009; Guisan & Thuiller, 2005).

La aproximación utilizada se basó en criterios previamente establecidos, sin embargo puede ser subjetiva debido a que el análisis de componentes principales (ACP) es una técnica descriptiva los datos (Jolliffe, 2002) y la combinación de variables seleccionadas puede variar de acuerdo al criterio del investigador. Por ejemplo, al ser más estricto en seleccionar únicamente un número fijo de variables que haya participado en cada vector o en la explicación total de la variabilidad. Se hubieran podido utilizar otras aproximaciones de análisis multivariado, sin embargo, el ACP busca representar la máxima variabilidad posible

que genera la máxima inercia en los diferentes ejes y por ende, podría mejorar la predicción de la máxima entropía en los datos.

6.1.3 Evaluación de parámetros para la selección de modelos: La evaluación estadística y de estabilidad espacial de los modelamientos obtenidos en Maxent permitieron seleccionar los modelos con mejor desempeño de acuerdo a estos criterios (Ver Tabla 9). Adicionalmente, el análisis de los resultados de cada uno de ellos brindó la posibilidad de mejorar el entendimiento del modelamiento de nichos en estos taxones. Los tres estadísticos utilizados correspondientes a AUC, kappa y TSS pusieron en evidencia tres situaciones contrastantes. El AUC evidenció un posible sobreajuste de los modelos, el kappa mostró un bajo desempeño de todos los modelos asociado a una alta prevalencia de la predicción y el TSS, siendo un estadístico independiente de esta prevalencia, presentó un desempeño adecuado de los modelos, permitiendo discriminar, en algunos casos, entre los modelos ajustados con la totalidad de variables y los ajustados con la selección de las mismas.

En primer lugar, en el caso del AUC, todos los promedios, a excepción del obtenido para *P. neglectus* cuando se utilizó la selección de variables (AUC de 0.76), fueron superiores a 0.9 por lo que pueden considerarse como de alta precisión y en el caso mencionado de *P. neglectus* se puede considerarse aceptable (Manel et al., 2001) (Ver Tabla 9). No obstante, estos valores también podrían estar reflejando un sobreajuste de los modelos causado por la selección de las pseudo-ausencias y/o por la extensión en la que se están generando los modelos (Anderson & Gonzalez, 2011; Barbet-Massin et al., 2012; Lobo et al., 2008). La aproximación utilizada para la generación de los puntos de pseudo-ausencias se basó en la selección de 10000 puntos aleatorios de acuerdo a lo recomendado para Maxent (Elith et al., 2011; Phillips et al., 2011). En un primer intento, éstos fueron generados en el aplicativo y los valores de AUC estuvieron entre 0.997 y 0.999 superando siempre al valor de AUC calculado para los puntos de prueba en cada validación cruzada (datos no presentados). Posteriormente, en un segundo ensayo los puntos fueron generados en la totalidad del área nativa de los taxones obteniéndose resultados similares (resultados no presentados). Finalmente, los mejores resultados en términos de la comparación con los demás estadísticos y el AUC para los puntos de prueba se obtuvieron cuando las 10000 pseudo-ausencias fueron generadas de forma aleatoria 300 Km alrededor de las zonas de ocurrencia de todos los taxones y dentro de

su área nativa. Buscando constituir un *background* o conjunto de pseudo-ausencias con similares características al conjunto total de ocurrencias del género *Phaseolus*, en términos de los esfuerzos de colecta y características ambientales (Barbet-Massin et al., 2012; Phillips & Elith, 2010; Phillips et al., 2009). Los resultados obtenidos con esta metodología basada en la aleatoriedad y estratificación ambiental así como el número de pseudo-ausencias seleccionado tuvieron el mejor desempeño estadístico como lo reportaron Barbet-Massin et al. (2012) para modelos basados en el aprendizaje de máquina como Maxent. Sin embargo, es importante evaluar otras alternativas de generación de pseudo-ausencias como tener puntos de específicos para cada taxón y/o extensión del área a evaluar como lo sugieren Anderson & Gonzalez (2011) y Phillips et al. (2009).

En segundo lugar, los resultados de kappa fueron completamente contrastantes con los obtenidos con AUC (Ver Tabla 9). De acuerdo a (Monserud & Leemans, 1992), los promedios fueron bajos y propios de modelos que van desde los no precisos o con valores de Kappa < 0.05 (15 y 21 promedios, en el caso de los modelos realizados con todas y la selección de variables, respectivamente), con precisión muy pobre con valores de Kappa entre 0.05 y 0.2 (13 y 10 promedios, con los modelos con todas y la selección de variables), de precisión pobre con valores de Kappa entre 0.2 y 0.4 (4 y 3 promedios, para los modelos de todas y la selección de variables) y de precisión aceptable con valores de Kappa entre 0.4 y 0.55 (2 promedios de Kappa para taxones modelados con todas las variables). Esta situación puede deberse a que el valor de kappa está influenciado por la prevalencia del modelo o a la proporción de presencias predichas (Allouche et al., 2006; C. Liu et al., 2005), siendo una consecuencia lógica del área en la que se realizó el modelamiento. Situación que podría mejorarse limitando las áreas para las cuales se realiza el modelamiento para cada taxón. Sin embargo, este fraccionamiento podría dificultar la interpretación de los modelos y limitar la identificación de zonas bioclimáticamente aptas para los taxones en estudio. Así mismo, puede enmascarar zonas no muestreadas.

En tercer lugar, el estadístico TSS, siendo independiente de la prevalencia, pondera de igual forma la sensibilidad y especificidad de los modelamientos y permitió discriminar entre modelos ajustados con todas y la selección de variables (Ver Tabla 9). La comparación entre los estadísticos AUC y TSS descartó que la selección del umbral hubiera sido errónea y confirma que este es adecuado para este tipo de modelamientos de nicho ecológico en Maxent

(Allouche et al., 2006; Barbet-Massin et al., 2012; Thuiller et al., 2009). Sin embargo, los resultados deben manejarse con cuidado debido a que el umbral seleccionado y los cálculos posteriores dependen en gran medida de la curva ROC y como se mencionó en el caso del estadístico AUC, esta podría estar influenciada por un sobreajuste en los modelos.

Por su parte, al evaluar la estabilidad espacial de la predicción de nichos ecológicos, se pudo determinar que los modelos tienen una alta estabilidad con PADM10 mayores al 90% en todos los casos a excepción de los nichos de *P. pauciflorus* y *P. talamancensis* cuando los modelos fueron ajustados con todas las variables (Ver Tabla 10). Esto significa que las predicciones se mantienen independientemente del conjunto de datos que este siendo utilizado en la validación cruzada, tal y como lo plantea Ramirez-Villegas et al. (2010).

6.1.4 Selección de modelos de nicho ecológico obtenidos con Maxent: De acuerdo con la evaluación y ponderación de los criterios estadísticos y de estabilidad de los modelamientos, para 7 taxones (*P. acutifolius acutifolius*, *P. amblyosepalus*, *P. angustissimus*, *P. costaricensis*, *P. neglectus*, *P. pachyrrhizoides* y *P. tuerckheimii*) fue seleccionado el modelamiento realizado con la totalidad de las variables y para los restantes 27 el modelamiento con la selección de ellas (Ver Tabla 11).

En el caso de los 7 taxones, en 5 de ellos (*P. acutifolius acutifolius*, *P. amblyosepalus*, *P. costaricensis*, *P. neglectus* y *P. pachyrrhizoides*) los estadísticos AUC, kappa y TSS fueron el factor determinante para la selección y en dos (*P. angustissimus*, *P. tuerckheimii*), el factor determinante fue la estabilidad espacial. Cabe destacar que el tamaño del nicho ecológico predicho con todas las variables fue menor o al menos igual con relación al predicho con la selección de ellas. Aunque como se menciona más adelante, esta no fue la tendencia, debido a que para el grupo de taxones en donde los estadísticos favorecieron la selección de variables, existieron nichos de mayor y menor tamaño en relación a los predichos con la totalidad de variables. Adicionalmente, los taxones *P.amblyosepalus*, *P. neglectus* y *P. tuerckheimii* resultaron con mayor número de predictores ambientales que puntos de ocurrencia (21 predictores en comparación con 10, 10 y 18 puntos de ocurrencia, respectivamente), siendo esto contradictorio con lo planteado por Guisan & Thuiller (2005) en términos de buscar tener menor número de predictores que puntos de ocurrencia mediante la remoción de variables con información redundante. Así, aunque los estadísticos apoyaron la utilización de la totalidad de

las variables en estos casos, revisando los valores en la estabilidad de la predicción de los taxones con menos de 21 ocurrencias (*P. amblyosepalus*, *P. glabellus*, *P. maculatus*, *P. neglectus*, *P. talamancensis*, *P. tuerckheimii* y *P. zimapanensis*), la estabilidad espacial en la predicción fue mayor o al menos igual para los modelos ajustados con la selección de 10 variables, lo que apoyaría la idea de remover variables en los casos en que se tenga un menor número de ocurrencias que predictores ambientales, aunque para algunos casos los estadísticos no favorezcan esta iniciativa. Por lo tanto, se podría concluir que en los casos en que se tenga menor cantidad de ocurrencias que predictores ambientales, se debe dar mayor importancia a la estabilidad de los modelos que a los estadísticos.

En el caso de los restantes 27 taxones, donde fueron seleccionados los modelos con 10 variables, en 17 de ellos la puntuación obtenida tanto para los estadísticos AUC, Kappa y TSS, como para la estabilidad en la predicción fue igual. En estos casos, la utilización de los modelos con 10 variables se justificó debido a las ventajas de usar modelos ajustados sin variables redundantes y similar desempeño a los ajustados con la totalidad. Para estos taxones, al comparar la razón entre las áreas de sus nichos predichas con la selección y la totalidad de las variables (Tabla 10), en 13 casos (*P. coccineus*, *P. filiformis*, *P. glabellus*, *P. grayanus*, *P. maculatus maculatus*, *P. maculatus ritensis*, *P. nelsonii*, *P. parvifolius*, *P. parvulus*, *P. pedicellatus*, *P. polystachios*, *P. vulgaris* y *P. xanthotrichus*) variaron entre 0.81 y 1.1, lo que significa que el área predicha es muy similar en ambos casos. En el caso de *P. lunatus*, *P. acutifolius tenuifolius* y *P. leptostachyus leptostachyus*, la razón fue de 2.2, 0.73 y 0.64, respectivamente. Es decir, en el primer caso el tamaño del nicho se duplica con la selección de variables, mientras en los demás casos se reduce en más de un 30%. Estos resultados plantean al menos dos situaciones. La primera es que no hay un patrón en términos del tamaño entre los nichos obtenidos con la selección de 10 variables y la totalidad con relación al desempeño de los estadísticos. La segunda es que el desempeño de los modelamientos realizados con la selección de variables no se ve afectado por la distribución de los taxones, esto se hace evidente al comparar las amplias distribuciones de *P. vulgaris* y *P. lunatus* con las localizadas de *P. filiformis* y *P. glabellus*. Esto apoya la teoría que recomienda este tipo de análisis (Barbet-Massin et al., 2012; Guisan & Thuiller, 2005; Phillips & Elith, 2010; Thuiller et al., 2009).

Para los restantes 10 taxones del grupo anterior (*P. acutifolius*, *P. augusti*, *P. coccineus coccineus*, *P. dumosus*, *P. leptostachyus*, *P. maculatus*, *P. pauciflorus*, *P. pluriflorus*, *P. talamancensis* y *P. zimapanensis*), el desempeño en términos de los estadísticos y la estabilidad en la predicción fue mayor para los modelos obtenidos con la selección de 10 variables. En éstos se destaca que la estabilidad espacial de la predicción fue mayor en todos los casos y adicionalmente, en el caso de *P. maculatus*, *P. plurifolius* y *P. zimapanensis* los estadísticos, en especial el TSS, apoyaron esta selección a favor de los modelos con 10 variables. En este caso, se confirma la conveniencia de la utilización del estadístico TSS para identificar modelos que tienen una alta estabilidad espacial en la predicción.

6.2 Evaluación de los efectos del cambio climático en parientes silvestres del

frijol común. La comparación de los nichos ecológicos para los taxones evaluados en condiciones de cambio climático y distintos escenarios de migración, muestra que ninguno de los nichos perdió el 100% de su área (Ver Tabla 12 y Figura 6). Sin embargo, de no tener capacidad de migración, 32 nichos disminuirán su área, mientras de tener capacidad de migración y poder establecerse en otros sitios, el tamaño total del nicho igualmente disminuirá para 19 taxones y en 15 taxones se mantendrá o aumentará. Incluso en el caso de 5 taxones (*P. leptostachyus leptostachyus*, *P. maculatus*, *P. acutifolius acutifolius* y *P. neglectus*) duplicarán las áreas de su nicho con relación a las actuales. Estos resultados en principio plantean que existen taxones ganadores y perdedores en términos del área potencial de su nicho en condiciones de cambio climático, lo que concuerda con lo encontrado por Jarvis et al. (2008) para el modelamiento de los efectos del cambio climático en parientes silvestres de otras especies. Sin embargo, debido a que estos modelamientos se hacen bajo múltiples incertidumbres, por ejemplo en términos de la respuesta de adaptación de los taxones a las nuevas condiciones de clima, concentración de CO₂, presiones bióticas y urbanización entre otras (Heikkinen et al., 2006; Jarvis et al., 2008; Wiens et al., 2009), incluso la incertidumbre relativa a la verdadera capacidad de migración y mecanismos de dispersión de las especies (Pearson, 2006), estos resultados se deben manejar con cuidado, especialmente para aquellos taxones con distribución de sus ocurrencias muy concentrado, debido a que su distribución podría ser un indicador de su capacidad de migración.

La comparación de las áreas del nicho ecológico en condiciones actuales y las de cambio climático en cualquiera de los escenarios de migración pone en evidencia que estas van a cambiar. Sin embargo, al explorar diferentes tipos de asociación entre estos cambios de área y, por ejemplo, el tamaño del nicho actual, el número de ocurrencias, la concentración de las mismas, no fue posible encontrar un patrón claro. El único análisis en donde se encontró alguna concordancia es términos de la localización espacial de los puntos de ocurrencia. Así, fue posible identificar algunos de los comportamientos característicos de acuerdo a la distribución de sus ocurrencias (Ver Figura 7). Por ejemplo, para varios taxones con distribución de ocurrencias sobre la parte alta de la Sierra Madre Occidental, las pérdidas de área estarán por encima del 45% y las nuevas áreas no superarán el 25%. Para taxones con puntos de ocurrencia en la Sierra Madre Occidental, especialmente en la parte baja y algunos otros puntos sobre el altiplano mexicano, la Sierra Madre Oriental y en la península de Baja California, el tamaño de los nichos en general aumenta y se mantiene el nicho actual con pocas pérdidas, aunque fragmentado como se discutirá más adelante. Para aquellos que tienen puntos de ocurrencia sobre la Sierra Madre Oriental, aunque pierden área en ambos escenarios, ganan algunas áreas en la Sierra Madre Occidental. Por su parte los taxones distribuidos en Centro América tienen nichos pequeños y de menor tamaño en condiciones de cambio climático. Para los taxones con distribución más amplia, es decir, para *P.vulgaris* y *P.lunatus* las pérdidas de nicho fueron del 13% y 28%, mientras las ganancias estuvieron en 22% y 36%, manteniendo su nicho relativamente estable (Comparación Figura 7 y Tabla 12).

En términos del análisis de fragmentación del nicho ecológico, cuando fueron agrupados los taxones de acuerdo al tamaño de su nicho actual y se evaluó el número de parches y tamaño promedio de los mismos, el número de parches se incrementó para todos los tamaños de nichos, especialmente para los de tamaño mayor a 500 mil km² (Ver Figura 8). En términos del tamaño promedio de los parches que conforman cada nicho, éstos disminuyeron al compararlos con las condiciones de cambio climático y no migración, independientemente del tamaño del nicho. En el caso de las condiciones de cambio climático y migración, el tamaño promedio de los parches fue menor para los nichos con tamaños inferiores a 500 mil km² mientras que para los nichos con tamaño actual mayor a 500 mil km², el tamaño promedio se incrementaría drásticamente, confirmando que los taxones con nichos grandes en las condiciones actuales tienden a estar más estables y continuos en el futuro, mientras los de

menor tamaño se fragmentan en mayor medida. Esta relación no fue completamente lineal, lo que concuerda con lo encontrado por Jarvis et al. (2008) para taxones del género *Solanum* en Latinoamérica. Adicionalmente se encontró que los taxones que ganan mayor cantidad de área con relación a su nicho actual tienden a tener parches más grandes en condiciones de cambio climático.

Finalmente, cuando se realizó el análisis de riqueza de taxones en las condiciones actuales y se comparó con la riqueza de especies en condiciones de cambio climático y no movilidad (Ver Figura 10), se encontró que la zona con mayores pérdidas de biodiversidad en términos de estos taxones se encuentra concentrada especialmente sobre la Sierra Madre Occidental entre los estados de Chihuahua y Durango, en donde se concentran áreas con pérdidas previstas de entre 6 y 8 taxones. En contraste, cuando se compararon las zonas de mayores pérdidas y ganancias de biodiversidad en un escenario de cambio climático y migración (Ver Figura 11), las zonas de mayores pérdidas se mantuvieron, pero con pérdidas de entre 3 a 8 especies y las zonas con mayores ganancias se generaron hacia el sur de la zona actual de mayor riqueza, especialmente entre el sur de Durango y el norte de México D.F. En resumen, varios de los taxones ganarían zonas potencialmente aptas para el establecimiento de sus nichos en zonas de mayor altitud sobre la Sierra Madre las cuales se encuentran al sur de esta cadena montañosa.

7. CONCLUSIONES Y RECOMENDACIONES

- ✓ El análisis exploratorio de datos permitió determinar que existen variables redundantes
- ✓ El análisis de componentes principales (ACP) permitió explicar de forma coherente la relación entre variables y hacer una selección de ellas.
- ✓ La inclusión de la significancia biológica de las variables bioclimáticas como criterio de selección, junto al ACP, permitió seleccionar y descartar variables de forma coherente con relación al significado mismo de las variables, por lo que se propone como criterio para la selección de estas en ejercicios similares.
- ✓ El estadístico TSS fue el que presentó mayor poder de discriminación por lo que se propone como un estadístico a ser usado en estudios similares.
- ✓ Se propone explorar la configuración del *background* o conjunto de pseudo-ausencias en trabajos posteriores donde se pretenda comparar nichos de especies con diferente extensión geográfica.
- ✓ La selección de variables mostró ser una alternativa eficiente para mejorar el desempeño de los modelos de nicho ecológico.
- ✓ Bajo efectos de cambio climático y el modelo aplicado, ninguno de los 34 taxones parientes del frijol común presenta riesgo de extinción y se recomienda usar este modelo para otras especies/taxones, lo cual permitiría hacer una comparación geográfica de riesgo entre diferentes especies.
- ✓ Bajo efectos de cambio climático la configuración de los nichos cambiará, fragmentándose y variando su área y la de los parches que lo conforman en función de su capacidad de migración.
- ✓ No hay tendencia clara entre los cambios de área de los nichos y el tamaño del nicho actual, el número de ocurrencias y la concentración de las mismas.
- ✓ La zona de mayor riqueza actual y de mayor pérdida de biodiversidad en un escenario de cambio climático es la Sierra Madre Occidental entre los estados de Chihuahua y Durango.
- ✓ En un escenario de cambio climático y migración la zona hacia donde migrarían los taxones se localiza entre los estados de Durango y México D.F, en zonas de mayor altitud y humedad a las actuales.

- ✓ El análisis seguido en este trabajo permitirá hacer una cuantificación de los efectos del cambio climático sobre parientes silvestres de cultivos de importancia para la alimentación y la agricultura utilizando variables bioclimáticas no redundantes.
- ✓ En futuros trabajos se propone realizar un análisis de mayor especificidad con variables ambientales y de uso de la tierra específicas de las zonas de mayor riqueza de especies modelada, esto con el fin de poder discriminar al máximo las zonas prioritarias de colecta y conservación de biodiversidad.
- ✓ Este esquema de análisis, aunque plantea un esfuerzo adicional en términos estadísticos y de análisis espacial, puede ser utilizado para comparar la sensibilidad de las predicciones realizadas con la totalidad de variables bioclimáticas, representando un beneficio para la disminución de las incertidumbres asociadas a las predicciones.

8. BIBLIOGRAFIA

- Acosta-Gallegos, J. A., Quintero, C., Vargas, J., Toro, O., Tohme, J., & Cardona, C. (1998). A new variant of arcelin in wild common bean, *Phaseolus vulgaris* L., from southern Mexico. *Genetic Resources and Crop Evolution*, 45(3), 235–242.
- Acosta-Gallegos, J., & Gepts, J. D. (2007). Prebreeding in common bean and use of genetic diversity from wild germplasm. *Crop Science*, 47(S3), S44–S59.
doi:10.2135/cropsci2007.04.0008IPBS
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. doi:10.1111/j.1365-2664.2006.01214.x
- Anderson, R. P., & Gonzalez, I. (2011). Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, 222(15), 2796–2811.
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*. Recuperado a partir de <http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2011.00172.x/full>
- Bioversity International, IRRI, & CIAT. (2009). Gap Analysis. Recuperado junio 18, 2011, a partir de <http://gisweb.ciat.cgiar.org/GapAnalysis/>
- Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., & Vanderleyden, J. (2003). Beans (*Phaseolus* spp.) – model food legumes. *Plant and Soil*, 252, 55-128.
doi:10.1023/A:1024146710611
- Busby, C., & Austin, M. (1990). Chapter10 - BIOCLIM - A Bioclimate Analysis and Prediction System. En *Nature Conservation: Cost Effective Biological Surveys and Data Analysis*. Csiro Publishing.
- CCAFS, CIAT, ILRI, & IFPRI. (2012, enero). GCM- Downscaled GCM Data Portal. Recuperado enero 8, 2012, a partir de <http://ccafs-climate.org/>
- Chang, K.-T. (2004). *Introduction to geographic information systems*. McGraw-Hill Higher Education.
- Cowen, D. J. (1988). GIS versus CAD versus DBMS: What are the differences?. *PHOTOGRAMM. ENG. REMOTE SENS.*, 54(11), 1551–1555.
- Croptrust. (2010). The Global Crop Diversity Trust. *Global Crop Diversity Trust: A Foundation for Food Security*. Recuperado junio 18, 2011, a partir de <http://www.croptrust.org/main/ldiversity.php>

- De Souza Muñoz, M. E., De Giovanni, R., De Siqueira, M. F., Sutton, T., Brewer, P., Pereira, R. S., ... Canhos, V. P. (2011). OpenModeller: A generic approach to species' potential distribution modelling. *GeoInformatica*, 15(1), 111–135.
- Elith, J., & Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 677-697. doi:10.1146/annurev.ecolsys.110308.120159
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43–57. doi:10.1111/j.1472-4642.2010.00725.x
- ESRI. (2011). *ArcGIS Desktop*. Redlands, CA: Environmental Systems Research Institute: ESRI. Recuperado a partir de www.esri.com
- FAO. (2004). *Tratado internacional sobre los recursos fitogenéticos para la alimentación y la agricultura: un tratado mundial para la seguridad alimentaria y la agricultura sostenible*. FAO. Recuperado a partir de <http://www.fao.org/AG/cgrfa/Spanish/itpgr.htm>
- Felipe, B. (2009). Ayuda maxent. *Red Iberoamericana para el estudio de ecosistemas forestales ante el cambio climático global*. Recuperado septiembre 29, 2012, a partir de http://161.67.130.145/~red/index.php?option=com_content&view=article&id=25&Itemid=53
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation*, 24(01), 38–49.
- Foody, G. M. (2008). GIS: biodiversity applications. *Progress in Physical Geography*, 32(2), 223–235.
- Foresman, T. W. (1998). *The history of geographic information systems: perspectives from the pioneers*. Prentice Hall PTR.
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology letters*, 8(9), 993–1009.
- Hajjar, R., & Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica*, 156(1), 1–13.
- Heikkila, E. J. (1998). GIS is dead; long live GIS! *Journal of the American Planning Association*, 64(3), 350–360.

- Heikkinen, R. K., Luoto, M., Araújo, M. B., Virkkala, R., Thuiller, W., & Sykes, M. T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, 30(6), 751–777.
- Heywood, V., Casas, A., Ford-Lloyd, B., Kell, S., & Maxted, N. (2007). Conservation and sustainable use of crop wild relatives. *Agriculture, ecosystems & environment*, 121(3), 245–255.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., Jarvis, A., & others. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978.
- Hijmans, R. J., Guarino, L., Cruz, M., & Rojas, E. (2001). Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, 15–19.
- Hu, J., & Jiang, Z. (2011). Climate Change Hastens the Conservation Urgency of an Endangered Ungulate. *PLoS ONE*, 6(8). doi:10.1371/journal.pone.0022873
- Hunter, D., & Heywood, V. (2011). *Crop wild relatives: a manual of in situ conservation*. Earthscan/James & James. Recuperado a partir de http://books.google.com.co/books?hl=en&lr=&id=MPfZNXVZp-UC&oi=fnd&pg=PP2&dq=Crop+Wild+Relatives+A+Manual+of+in+situ+Conservation&ots=tKHellTUpl&sig=BbSjlfh9eFi20QX8WLMgI4O_a00
- IPPC, I. P. O. C. (2007). Climate change 2007: the physical science basis. *Agenda*, 6, 07.
- Jankowski, P. (1995). Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Science*, 9(3), 251–273.
- Jarvis, A., Lane, A., & Hijmans, R. J. (2008). The effect of climate change on crop wild relatives. *Agriculture, Ecosystems & Environment*, 126(1-2), 13–23.
- Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed., Vol. 2). Springer-Verlag New York, Inc.: Wiley Online Library. Recuperado a partir de http://onlinelibrary.wiley.com/mrw_content/esbs/articles/bsa501/image_n/bsa501.pdf
- Jones, P. G., Beebe, S. E., Tohme, J., & Galwey, N. W. (1997). The use of geographical information systems in biodiversity exploration and conservation. *Biodiversity and Conservation*, 6(7), 947–958.
- Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3), 385–393.

- Liu, C., White, M., & Newell, G. (2009). Measuring the accuracy of species distribution models: a review (pp. 4241 - 4447). Presentado en 18th World IMACS / MODSIM Congress, Cairns, Australia. Recuperado a partir de <http://mssanz.org.au/modsim09>
- Liu, Canran, Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3), 385–393. doi:10.1111/j.0906-7590.2005.03957.x
- Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2), 145–151. doi:10.1111/j.1466-8238.2007.00358.x
- Loyola, R. D., Lemes, P., Faleiro, F. V., Trindade-Filho, J., & Machado, R. B. (2012). Severe Loss of Suitable Climatic Conditions for Marsupial Species in Brazil: Challenges and Opportunities for Conservation. *PLOS ONE*, 7(9), e46257.
- Lu, N., Jia, C. X., Lloyd, H., & Sun, Y. H. (2012). Species-specific habitat fragmentation assessment, considering the ecological niche requirements and dispersal capability. *Biological Conservation*, 152, 102–109.
- Manel, S., Williams, H. C., & Ormerod, S. j. (2001). Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38(5), 921–931. doi:10.1046/j.1365-2664.2001.00647.x
- Maxted, N., & Kell, S. P. (2009). Establishment of a Global Network for the. *Situ Conservation of Crop Wild Relatives: Status and Needs. FAO Commission on Genetic Resources for Food and Agriculture, Rome, Italy*. Recuperado a partir de http://www.pgrforum.org/documents/Global_in_situ_CWR_conservation_network.pdf
- Maxted, N., Kell, S., Toledo, Á., Dulloo, E., Heywood, V., Hodgkin, T., Ford-Lloyd, B. (2010). A global approach to crop wild relative conservation: securing the gene pool for food and agriculture. *Kew Bulletin*, 65(4), 561–576.
- McGarigal, K., & Marks, B. J. (1995). Spatial pattern analysis program for quantifying landscape structure. *Gen. Tech. Rep. PNW-GTR-351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station*. Recuperado a partir de <http://www.umass.edu/landeco/pubs/mcgarigal.marks.1995.pdf>
- Meilleur, B. A., & Hodgkin, T. (2004). In situ conservation of crop wild relatives: status and trends. *Biodiversity and Conservation*, 13(4), 663–684.
- Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62(4), 275–293.
- Morin, X., & Lechowicz, M. J. (2008). Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters*, 4(5), 573–576.

- Nakicenovic, N., Alcamo, J., Davis, G., De Vries, B., Fenhann, J., Gaffin, S., ... Kram, T. (2000). *Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change*. Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US). Recuperado a partir de http://www.osti.gov/energycitations/product.biblio.jsp?osti_id=15009867
- Nix, H. A. (1986). A biogeographic analysis of Australian elapid snakes. *Atlas of Elapid Snakes of Australia*.(Ed.) R. Longmore, 4–15.
- Openshaw, S. (1999). Geographical data mining: key design issues (Vol. 99). Presentado en Proceedings of the 4th International Conference on GeoComputation, Virginia, USA. Recuperado a partir de http://www.geocomputation.org/1999/051/gc_051.htm
- Pardo, C. E., & Del Campo, P. C. (2007). Combination of Factorial Methods and Cluster Analysis in R: The Package FactoClass. *Revista Colombiana de Estadística*, 30(2), 231-245.
- Parry, M., Rosenzweig, C., & Livermore, M. (2005). Climate change, global food supply and risk of hunger. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1463), 2125.
- Pearson, R. G. (2006). Climate change and the migration capacity of species. *Trends in Ecology & Evolution*, 21(3), 111–113.
- Peterson, A. T. (2003). Predicting the geography of species' invasions via ecological niche modeling. *The quarterly review of biology*, 78(4), 419–433.
- Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19(1), 181–197.
- Phillips, S. J., & Elith, J. (2010). POC plots: calibrating species distribution models with presence-only data. *Ecology*, 91(8), 2476–2484.
- Phillips, S.J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190(3-4), 231–259.
- Phillips, Steven J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), 231-259. doi:10.1016/j.ecolmodel.2005.03.026
- Phillips, Steven J., & Dudík, M. (2008). Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175. doi:10.1111/j.0906-7590.2008.5203.x

- Phillips, Steven J., Dudik, M., & Schapire, R. E. (2011). *Maxent software for species habitat modeling*. AT&T Labs-Research, Princeton University, and the Center for Biodiversity and Conservation, American Museum of Natural History. Recuperado a partir de <http://www.cs.princeton.edu/~schapire/maxent/>
- Python Software Foundation. (2008). *Python*. Python Software Foundation. Recuperado a partir de <http://python.org/>
- R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Recuperado a partir de <http://www.r-project.org/>
- Ramirez, J., & Jarvis, A. (2010). Downscaling Global Circulation Model Outputs: The Delta Method Decision and Policy Analysis Working Paper No. 1. *International Center for Tropical Agriculture*. Available on: http://ccafs-climate.org/download_sres.html.
- Ramirez-Villegas, J., Khoury, C., Jarvis, A., Debouck, D. G., Guarino, L., & Fuller, D. Q. (2010). A Gap Analysis Methodology for Collecting Crop Genepools: A Case Study with. *PLoS ONE*, 5(10), 1 - 18. doi:10.1371/journal.pone.0013497
- Rempel, R., Kaukinen, D., & Carr, A. (2012). *Patch Analyst and Patch Grid*. Thunder Bay, Ontario: Centre for Northern Forest Ecosystem Research. Recuperado a partir de <http://www.cnfer.on.ca/SEP/patchanalyst/>
- Salem, B. B. (2003). Application of GIS to biodiversity monitoring. *Journal of Arid Environments*, 54(1), 91–114.
- Savigny, D. D., & Wijeyaratne, P. (1995). *Gis for Health and the Environment: Proceedings of an International Workshop Held in Colombo, Sri Lanka, 5-10 September 1994*. IDRC.
- Scott, L. M., & Janikas, M. V. (2010). Spatial Statistics in ArcGIS. En M. M. Fischer & A. Getis (eds.), *Handbook of Applied Spatial Analysis* (pp. 27-41). Springer Berlin Heidelberg. Recuperado a partir de <http://www.springerlink.com/content/x408660xh82541u6/abstract/>
- Singh, S. P., & Munoz, C. G. (1999). Resistance to common bacterial blight among Phaseolus species and common bean improvement. *Crop Science*, 39(1), 80–89.
- Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*, 32(3), 369–373.
- Wiens, J. A., Stralberg, D., Jongsomjit, D., Howell, C. A., & Snyder, M. A. (2009). Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, 106(Supplement_2), 19729-19736. doi:10.1073/pnas.0901639106

Xu, T., & Hutchinson, M. F. (2011). ANUCLIM Version 6.1 User Guide. *The Australian National University, Fenner School of Environment and Society, Canberra.*

Recuperado a partir de

http://fennerschool.anu.edu.au/files/panel/447/anuclim61_user_guide_pdf_19159.pdf

8. ANEXOS

Anexo 1– Código en R para la realización del análisis de componentes principales

```
# importando los datos desde el archivo phaseolus.csv
wd <- "C:/path/Datos/" # directorio de trabajo
setwd(wd)
frijol.pca <-
read.table("phaseolus.csv",header=T,row.names="ID_REG",sep=",")

# haciendo el ACP y la clasificación en FactoClass
#numero de factores 21, clases 6
FC.phaseolus <-FactoClass(frijol.pca, dudi.pca, nf=4, nfcl=21, k.clust=6,
scanFC="False")

# resultados del pca:
# variabilidad explicada por cada CP:

inertia.dudi(FC.phaseolus$dudi)

# Coordenadas normadas:
FC.phaseolus$dudi$c1

# Correlaciones con los ejes:
FC.phaseolus$dudi$co

# grafica de los Componentes con las variables
# PC1 vs PC2:
biplot(FC.phaseolus$dudi$c1[,1:2], FC.phaseolus$dudi$c1[,1:2], col=
c("black"), xlab= "PC1", ylab= "PC2", cex=0.8)
# PC1 vs PC3:
biplot(FC.phaseolus$dudi$c1[,1:3][-2], FC.phaseolus$dudi$c1[,1:3][-2], col=
c("black"), xlab= "PC1", ylab= "PC3", cex=0.8)
# PC2 vs PC3:
biplot(FC.phaseolus$dudi$c1[,2:3], FC.phaseolus$dudi$c1[,2:3], col=
c("black"), xlab= "PC2", ylab= "PC3", cex=0.8)
```

Anexo 2 – Código en R usado para la extracción de las predicciones logísticas de los modelamientos en Maxent del background y los datos de prueba.

```
# definicion y cambio del directorio de trabajo
wd <- "C:/Datos/Resultados_Modelamiento"
setwd(wd)

### Extracción de las predicciones logisticas del background
#'generación de un listado de los archivos con
# las predicciones del background'
back <- list.files(pattern="*backgroundPredictions.csv")
# 'generacion de un data.frame vacio de R con los campos
# que tienen todos los archivos y un campo para indicar el
# archivo del que provienen los datos'
all_back <- data.frame(back=NULL, x=NULL, y=NULL , raw=NULL, cumulative=NULL,
logistic=NULL)
  # lee todos los archivos de la lista back
  x <- read.table(back[i], header=TRUE, sep=",")
  # adiciona una columna con el nombre el archivo del que proviene cada dato
  x <- data.frame(file=rep(back[i], length(x[,1])), x)
  # agrega todos los datos al data.frame 'all_back'
  all_back <- rbind(all_back, x)
  def <- c("file", "logistic")
  # selecciona los campos file y logistic y los almacena en DF_DEF
  DF_def <- all_back[def]
  # crea una matriz de 340 columnas con el titulo de cada archivo del que
  # provienen los datos y 10000 filas de las predicciones logisticas
  uns_back <- unstack(DF_def, logistic~file)
  # exporta el data.frame 'uns_back a un archivo csv
  write.table(uns_back, file="logistic.csv", quote=FALSE, sep=",", col.names =
  NA)
}

###Extracción de las predicciones logisticas de los datos de prueba
# generacion de una lista de archivos con las predicciones de los muestreos
sample <- list.files(pattern="*samplePredictions.csv")
# crea un data.frame vacio con los campos de los archivos de muestras
all_sam <- data.frame(sample=NULL, x=NULL, y=NULL ,Test.or.train=NULL,
Raw.prediction=NULL, Cumulative.prediction=NULL, Logistic.prediction=NULL)
for(i in seq(along=sample)) {
  x <- read.table(sample[i], header=TRUE, sep=",")
  x <- data.frame(file=rep(sample[i], length(x[,1])), x)
  all_sam <- rbind(all_sam, x)
  defs <- c("file", "Test.or.train", "Logistic.prediction")
  DF_defs <- all_sam[defs]
  #selecciona unicamente los registros de prueba o test
  mat_test= DF_defs[DF_defs[, "Test.or.train"] == 'test',]
  write.table(mat_test, file="sample.csv", quote=FALSE, sep=",", col.names =
  NA)
}
```

Anexo 3 – Scripts en Phyton 2.6 utilizados para la evaluación de la estabilidad en la predicción logística de los modelamientos.

A – Script usado para la conversión de los mapas ASCII Grid generados por Maxent 3.3.3.k en formato ESRI Grid.

```
import arcpy
from arcpy import env

#Cambio del workspace al folder con los resultados del modelamiento
env.workspace = r"C:\Datos\Resultados\Modelamiento"
#generacion de un listado de los archivos que terminan en stddev.asc y avg.asc

desvListPre = arcpy.ListFiles("*pre_stddev.asc")
desvListFut = arcpy.ListFiles("*fut_stddev.asc")
promListPre = arcpy.ListFiles("*pre_avg.asc")
promListFut = arcpy.ListFiles("*fut_avg.asc")

# ciclo para convertir las desviaciones estandar del presente a ESRI Grid y
guardarlos en la carpeta desviaciones_pre con la terminacion dst

for ascp in ascListPre:
    nomp = ascp.split('_') # separamos la cadena por _
    nomlp= nomp[0][:3] + '_' +nomp[1][:3]
    if ascp[:8]== 'parvulus':
        a='parvu'
    elif'pre' in nomlp:
        a= ascp[:3]
    else:
        a= nomlp
    print a #para verificar el estado de avance
    arcpy.ASCIIToRaster_conversion(env.workspace + "/" + ascp ,
    "desviaciones_pre/" + a + "_dst","FLOAT")

# ciclo para convertir las desviaciones estandar del futuro a ESRI Grid y guardarlos
en la carpeta desviaciones_fut con la terminacion dst

for ascf in ascListFut:
    nomf = ascf.split('_') # separamos la cadena por _
    nomlf= nomf[0][:3] + '_' +nomf[1][:3]
    if ascf[:8]== 'parvulus':
        b='parvu'
    elif'fut' in nomlf:
        b= ascf[:3]
    else:
        b= nomlf
    print b #para verificar el estado de avance
    arcpy.ASCIIToRaster_conversion(env.workspace + "/" + ascf ,
    "desviaciones_fut/" + b + "_dst","FLOAT")

## Los ciclos para hacer la conversión de los promedios usan la misma estructura,
cambiando solamente la referencia al listado de los promedios, el folder en el cual
se guardan los resultados y su terminación.
```

B – Script usado para hacer el corte de las superficies de probabilidad logística en el área nativa de cada Taxón

```
import os
import arcpy
from arcpy.sa import *

arcpy.CheckOutExtension("spatial")

#Ruta al folder con las areas nativas de cada taxon
ws1=r"C:\Datos\Model\Mask_Folder"
# Ruta al folder con las proyecciones logisticas promedio
ws2=r"C:\Datos\Model\promedios
# Ruta del folder donde se van a almacenar los resultados
outputws=r"C:\Datos\Model\Promedios_AreaNativa

#Cambio de workspace y listado de raster
arcpy.env.workspace = ws1
rasters1 = arcpy.ListRasters("*mask")

#Cambio de workspace y listado de raster
arcpy.env.workspace = ws2
rasters2 = arcpy.ListRasters()

for raster in rasters1:
    if raster[:-4]+'ave' in rasters2:
        #Genera objetos Raster si la condición se cumple
        r1=Raster(os.path.join(ws1,raster))
        r2=Raster(os.path.join(ws2,raster[:-4]+'ave'))
        print "multiplicando en: " + raster
        #Multiplicacion
        output_ras=r1 * r2
        #salvar los resultados
        output_ras.save(os.path.join(outputws,raster[:-4]+'PAN'))

arcpy.CheckInExtension("spatial")
```

C – Script usado para generar los mapas binarios de predicción en el área nativa de cada Taxón

```
import os
import arcpy
from arcpy.sa import *

arcpy.CheckOutExtension("spatial")

# Leer archivo de umbrales que es un archivo de texto separado por tabulaciones y
que contiene el nombre de cada taxon y en frente el umbral.
inputUmbrales = r"C:\Datos\Umbrales.txt"
File = open(inputUmbrales, "r")
File.readline()
Umbrales_List=[]
# ciclo para importar los umbrales de cada especie y colocarlos en un listado
for line in File.readlines():
    Umbrales_List.append(float(line.split("\t")[1]))

ws_prom = r"C:\Datos\Model\promedios_AreaNativa"
ws_bin= r"C:\Datos\Model\Mapas_binarios"

#Cambio de workspace y listado de datos
```

```

arcpy.env.workspace = ws_prom
promedio = arcpy.ListRasters()

#ciclo para reclasificar los 34 raster de acuerdo a los 34 umbrales.
for k in range(34):
    a= Umbrales_List[k]
    b= promedio[k]
    print b + " Thresold: " + str(a)
    Reclass = Reclassify(b, "Value",RemapRange([[0,a,"NODATA"],[a,1,1]]))
    Reclass.save(os.path.join(ws_bin, b[:-3] +"bin"))

arcpy.CheckInExtension("spatial")

```

D – Script usado para la generación de las superficies con la desviación estándar en el área de la cobertura de distribución potencial

```

import os
import arcpy
from arcpy.sa import *

arcpy.CheckOutExtension("spatial")

ws_bin= r"C:\Datos\Model\Mapas_binarios"
ws_std= r"C:\Datos\Model\desviaciones_pre"
ws_std_CP= r"C:\Datos\Model\desviaciones_areaCoberturaDP"

#Cambio de workspace y listado de rasters
arcpy.env.workspace = ws_bin
bin = arcpy.ListRasters()

#Cambio de workspace y listado de rasters
arcpy.env.workspace = ws_std
std = arcpy.ListRasters()

#ciclo para hacer la multiplicacion entre los mapas binarios y el de desviaciones
estandar:

for raster in bin:
    if raster[:-3]+'dst' in std: #:
        #make Raster objects from filenames
        r1=Raster(os.path.join(ws_bin,raster))
        r2=Raster(os.path.join(ws_std,raster[:-3]+'dst'))
        print "multiplicando :" + str(r1)+ "*" + str(r2)
        #Multiplicacion
        output_ras=r1 * r2
        #save the output to the filesystem.
        output_ras.save(os.path.join(ws_std_CP,raster[:-3]+'cp'))

arcpy.CheckInExtension("spatial")

```

E – Script usado para el cálculo de la PADM10 por cada combinación ‘taxón-juego de variables’ utilizadas en el modelamiento

```

import os
import arcpy
from arcpy.sa import *

```

```
arcpy.CheckOutExtension("spatial")

ws_std_CP= r"C:\Datos\Model\desviaciones_areaCoberturaDP"
PADM10_sel= r"C:\Datos\Model\PADM10"

#Cambio de workspace y listado de rasters
arcpy.env.workspace = ws_ds_cp
dst = arcpy.ListRasters()

# ciclo para la reclasificacion de cada raster de desviacion estandar
for raster in dst:
    Reclass = Reclassify(raster, "Value",RemapRange([[0,0.1,1],[0.1,1,0]]))
    Reclass.save(os.path.join(PADM10_sel,raster[:-3]+"P10"))
    print "reclasificando: " + raster

arcpy.CheckInExtension("spatial")
```

Anexo 4 – Script usados en el proceso de análisis espacial

A – mapas binarios de predicción en condiciones de cambio climático y el escenario de no migración:

```

import os
import arcpy
from arcpy import env
from arcpy.sa import *

arcpy.CheckOutExtension("spatial")

#carpeta con los mapas binarios a nivel general:
bin_ftot = r"C:\Datos\Modeling\mapas_binarios_futuro"
#carpeta con los mapas binarios de las condiciones actuales:
bin_pre= r"C:\Datos\Modeling\mapas_binarios_presente"
#carpeta de resultados:
nomov= r"C:\Datos\Modeling\mapas_nomigracion"
# grid de referenciapara definir el extent y el snap.raster
phaseolus = r"C:\Datos\Referencia"

#definicion del espacio de trabajo y generacion de listados de raster
arcpy.env.workspace = bin_ftot
ftot = arcpy.ListRasters()

#definicion del espacio de trabajo y generacion de listados de raster
arcpy.env.workspace = bin_pre
pre = arcpy.ListRasters() #lista los raster mascararas de migracion de cada especie

for ras in ftot:
    if ras in pre:
        #make Raster objects from filenames
        r1=Raster(os.path.join(bin_ftot,ras))
        r2=Raster(os.path.join(bin_pre,ras))
        print "multiplicando: " + str(r1) + "*" + str(r2)
        #Multiply them
        output_ras=r1 * r2 * Raster(phaseolus)
        #configuración de los raster de salida
        desc = arcpy.Describe(phaseolus)
        ext = desc.extent
        arcpy.env.extent = ext
        arcpy.env.snapRaster = phaseolus
        print "reclasificando"
        #reclasificacion de los raster
        ReNM=Reclassify(output_ras,"Value",RemapValue([[1,1],["NODATA",0]]))
        ReNM.save(os.path.join(nomov,ras[:-3]+'nmov'))

arcpy.CheckInExtension("spatial")

```

B – Script usado para generar mapas binarios de predicción en condiciones de cambio climático y el escenario de migración:

```

import os
import arcpy
from arcpy import env
from arcpy.sa import *

```

```

arcpy.CheckOutExtension("spatial")

#carpeta con los mapas binarios a nivel general:
bin_ftot = r"C:\Datos\Modeling\mapas_binarios_futuro"
#carpeta con los mapas binarios de las condiciones actuales:
bin_pre= r"C:\Datos\Modeling\mapas_binarios_presente"
#carpeta de resultados:
mov= r" C:\Datos\Modeling\mapas_binarios_futuro_movilidad"
# grid de referencia para definir el extent y el snap.raster
phaseolus = r"C:\Datos\Referencia"
# carpeta para almacenar los grid con la distancia euclidea
eucl= r"C:\Datos\distancia_euclidea"

#definicion del espacio de trabajo y generacion de listados de raster
arcpy.env.workspace = bin_pre
pre = arcpy.ListRasters()

# ciclo para calcular la distancia euclidea
# desde el mapa binario de condiciones actuales
for bin in pre:
    EuD=Raster(os.path.join(bin_pre,bin))
    Dist= EucDistance (EuD, 2.697966)
    print "calculando la distancia euclidea de: " + bin
    Dist.save(os.path.join(eucl,bin[:-3]+"eu"))

# definicion del espacio de trabajo y generacion de listados de raster
arcpy.env.workspace = bin_ftot
ftot = arcpy.ListRasters()

# definicion del espacio de trabajo y generacion de listados de raster
arcpy.env.workspace = eucl
eucl = arcpy.ListRasters()

# ciclo para calcular los mapas de prediccion bajo el escenario de movilidad

for ras in ftot:
    if ras[:-3]+"eu" in eucl:
        r1=Raster(os.path.join(bin_ftot,ras))
        r2=Raster(os.path.join(eucl,ras[:-3]+"eu"))
        print "multiplicando: " + str(r1) + "*" + str(r2)
        #Multiply them
        output_ras=r1 * r2 * Raster(phaseolus)
        desc = arcpy.Describe(phaseolus) # configurar los grids de salida.
        ext = desc.extent
        arcpy.env.extent = ext
        arcpy.env.snapRaster = phaseolus
        print "reclasificando"
        ReM=Reclassify(output_ras,"Value",RemapValue([[1,1],["NODATA",0]]))
        ReM.save(os.path.join(mov,ras[:-3]+'mov1'))

arcpy.CheckInExtension("spatial")

```

C – Script usado para realizar el proceso de superposición espacial y comparación de áreas de predicción bajo cada escenario:

```

import os
import arcpy

```

```

from arcpy import env
from arcpy.sa import *

arcpy.CheckOutExtension("spatial")
#carpeta con los mapas binarios de condiciones actuales:
bin_pre= r"C:\Datos\Modeling\bin_pre"
#carpeta con los mapas binarios de condiciones futuras y escenario de movilidad:
mov= r" C:\Datos\Modeling\Movilidad"
#carpeta para guardar los resultados de la superposicion con los valores cero
sup0= r" C:\Datos\Modeling\superposicion\sup0"
#carpeta para guardar los resultados sin los valores cero
sup1= r" C:\Datos\Modeling\superposicion\sup1"
# grid para definir el extent y el snap.raster
phaseolus= r"C:\Datos\Referencia"

#cambio de workspace y listado de rasters
arcpy.env.workspace = bin_pre
pre = arcpy.ListRasters()

#cambio de workspace y listado de rasters
arcpy.env.workspace = mov
fut_m = arcpy.ListRasters()

# ciclo para hacer la superposición espacial
for ras in pre:
    if ras[:-3] + "mov1" in fut_m:
        r1=Raster(os.path.join(bin_pre,ras))
        r2=Raster(os.path.join(mov,ras[:-3]+"mov1"))
        print "superponiendo: " + str(r1) + "*" + str(r2)
        desc = arcpy.Describe(phaseolus)
        ext = desc.extent
        arcpy.env.extent = ext
        arcpy.env.snapRaster = phaseolus
        output_ras= r1 + (2*r2)
        print "reclasificando"
        output_ras.save(os.path.join(sup0,ras[:-3]+'sup'))
        RecS=Reclassify(output_ras,
            "Value",RemapValue([[1,1],[2,2],[3,3],[0,"NODATA"]]))
        RecS.save(os.path.join(sup1,ras[:-3]+'sup'))

arcpy.CheckInExtension("spatial")

```

D – Script usado para la proyección de los mapas binarios en coordenadas geográficas (WGS84) a coordenadas planas (WGS84_World_Mercator).

```

import os
import arcpy
from arcpy import env

bin= r"C:\Datos\Modeling\MapasBinarios"
bin_pro= r"C:\Datos\Modeling\MapasBinarios_proyectados"

# raster con sistema de coordenadas proyectado en WGS84_World_Mercator
phaseolus= "C:\Datos\Referencia\proyectado"

#cambio de workspace y listado de rasters
arcpy.env.workspace= bin
binList= arcpy.ListRasters()

```

```
#ciclo para proyectar los raster
for rs in binList:
    desc= arcpy.Describe(phaseolus)
    ext= desc.extent
    arcpy.env.extent= ext
    arcpy.env.snapRaster= phaseolus
    arcpy.ProjectRaster_management(rs,os.path.join(bin_pro,rs),phaseolus,
    "NEAREST","5000")
    print "proyectando: " + rs
```

