

EVALUATING FUTURE IMPACTS OF CLIMATE CHANGE
ON TRADITIONAL MEXICAN MAIZE SUITABILITY
AND INDIGENOUS COMMUNITIES IN MEXICO

by

Steven C. Gonzalez, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Geography
May 2018

Committee Members:

Russell Weaver, Chair

Jennifer Jensen

Thomas Ballinger

COPYRIGHT

by

Steven C. Gonzalez

2018

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgment. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Steven C. Gonzalez, refuse permission to copy in excess of the "Fair Use" exemption without my written permission.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my thesis advisor Dr. Russell Weaver for the continuous support throughout the development of this research. His immense knowledge greatly contributed to the direction of this thesis' results, particularly in relating its ecological aspects to sociological dilemmas.

Further, I express a species thanks to my thesis committee members, Dr. Jennifer Jensen and Dr. Thomas Ballinger. Dr. Jensen's acute attention to detail ensured that I be always on my toes. I am extremely gratefully for her expert and valuable guidance and inspiration. Dr. Ballinger's expertise in climatology greatly aided the direction of this thesis, particularly in variable selection and data processing.

Without my committee's support, critique, encouragement, and inspiration, this thesis would have never been completed.

I'd like to thank my family for their never-ending encouragement.

Lastly, I'd like to extend special acknowledgement to my wife, Alejandra, for unwavering support through this arduous process called graduate school.

Thank you all sincerely!

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	x
ABSTRACT	xiv
CHAPTER	
I. INTRODUCTION	1
II. IMPORTANCE OF RESEARCH.....	8
III. PURPOSE STATEMENT	10
IV. RESEARCH QUESTIONS	12
V. CONCEPTUAL FRAMEWORK.....	13
VI. LITERATURE REVIEW AND THEORETICAL FRAMEWORK	16
Maize Domestication and History	16
Maize as a Mirror of Humanity	19
Hutchinson’s Niche	25
Fundamental and Realized Niches	26
Ecological Niche Modeling (ENM)	27
MaxEnt	28
VII. DATA AND METHODS	32
Site and Situation.....	33
WorldClim Climatologies	33
WorldClim and ENVIREM Bioclimatic Variable Processing	36
CONABIO Maize Survey Records	40

	DEM-derived Topographic Indices.....	42
	Harmonized Soil Database v 1.2 Land Cover	43
	CONABIO Ethnolinguistic Data.....	43
VIII.	ANALYSIS AND TECHNIQUE	49
	MaxEnt Variable Selection	49
	MaxEnt in biomod2.....	52
	Technical Details.....	56
IX.	RESULTS	57
	Ensemble model performance	57
	Niche Overlap.....	64
	Multi-variate Environmental Similarity Surfaces (MESS)	68
	Maize Landrace Niche Size.....	72
	Geographic Distribution of Mexican Maize Diversity.....	73
	Influences of Indigeneity on Maize Landrace Diversity	81
	Geographic Clusters of Group 2 Diversity and Indigenous Presence	87
	Changes in Maize Landrace Diversity with Climate Change	92
	Clusters of Group 2 Maize Diversity Changes and Indigenous Communities	104
X.	CONCLUSIONS.....	108
	Maize Landrace Diversity	108
	Associations between Indigeneity and Maize Landrace Diversity	110
	Climate Change and Maize Landraces: Model Performance and Forecasts.....	111
XI.	CONTRIBUTIONS	114

XII. LIMITATIONS OF PROPOSED DATA AND METHODS	117
XIII. CONCLUDING REMARKS.....	119
APPENDIX SECTION.....	120
REFERENCES	221

LIST OF TABLES

Table	Page
1. Maize landraces with sample size in entire georeferenced CONABIO dataset (Total) with sample sizes after cleaning data of erroneous records (Cleaned). Then, count in presence/absence matrix (PAM) for spatially-filtered grid at 1 km ² . P/C Ratio refers to the count in the PAM divided by the ‘cleaned’ count, with ‘-’ indicating maize landraces with < 15 samples in PAM that were excluded from modeling, or failed in modeling.....	5
2. Indigenous language groups in Mexico by language family, including number of dialects per language group and estimated number of speakers above the age of 5	23
3.- Nineteen bioclimatic variables calculated with ‘dismo’ in R, corresponding to those found in the WorldClim database.....	38
4. Fifteen bioclimatic variables calculated using ‘ENVIREM’ in R, newly-proposed complementary extensions to the WorldClim’s bioclimatic variables.	39
5. Final 20 uncorrelated variables selected for modeling, with variable name, Variance Inflation Factor, source or R-package used for calculation, and a brief description.	51
6. Mean evaluation metric scores on 30% holdout data across 3 weighted-mean ensemble models. Each weighted-mean ensemble model performed similarly for each metric.	59
7. AUC values of weighted-mean ensemble by ROC (AUC) with the threshold cutoff with the minimum difference between sensitivity and specificity (SS) (SS Min Diff), the threshold cutoff with the maximum SS difference (SS Max), the minimum error rate (Min Error), and the threshold cutoff resulting in the minimum error (Min Error Cutoff).	61

8. Ensemble forecasted 1970 – 2000 realized niche size (km ²), ranked from most generalist to most specialist species. Landrace groups are exponentiated at the end of each landrace name.....	72
9. Games-Howell post-hoc analysis for Kruskall Wallis test indicating significant paired differences of median diversity per <i>municipio</i> (n = 2,511) by indigenous population presence group.....	81
10. Games-Howell post-hoc analysis for Kruskall-Wallis and Brown-Forsythe tests indicating differences of median diversity per <i>municipio</i> (n = 2,511) by indigenous marginality level.....	84
11. Forecasted species range percent change by maize landrace for 2041 – 2060. Landrace groups are exponentiated at end of landrace name.	102
12. Forecasted species range change by maize landrace for 2041 – 2060. Landrace groups assigned to each maize landraces are exponentiated at end of landrace name. Percent (%) Loss values greater than .90 are in bold.	103

LIST OF FIGURES

Figure	Page
1. Heuristic BAM diagram depicting the interaction between <i>biotically-suitable environments</i> (B), <i>abiotically-suitable environments</i> (A), and <i>mobility</i> (M) interplay to constitute species ecological niches. Ones (1) indicate observed species that exist in both <i>biotically-suitable & abiotically-suitable</i> environments. Twos (2) indicate observed species that exists outside of <i>abiotically-suitable</i> environments, but still persist. Threes (3) represent observed species that exists outside of <i>biotically-suitable</i> environments, but still persist. Fours (4) represent observed species that exists outside of both <i>biotically-suitable & abiotically-suitable</i> environments, but still persist.	14
2. Location of the Balsas River, overlaid the 30 arc-second WorldClim altitude raster to depict the profound drainage basin, with scalebar.	16
3. A visual representation of phenotypic diversity of maize landraces in Mexico.	20
4. Workflow overview, generally following the workflow of Ecological Niche Modeling as a whole, but with ensembling.	32
5. Workflow (Part A), bioclimatic variable processing and variable selection by VIF thresholding.	37
6. Workflow (Part B), cleaning of georeferenced maize landrace observation to create Presence/Absence Matrix (PAM) for variety observation and target background sampling.	40
7. Maize landrace observations, after cleaning, with each landrace assigned a different color, overlaid the Kernel Density Estimation of sampling density bias present in CONABIO maize observations.	41

8. Percent indigenous population by <i>municipio</i> (municipality). Data via <i>el Comision Nacional para el Conocimiento y Uso de la Biodiversidad</i> (CONABIO). 2010.....	44
9. Spatial distribution of the majority indigenous language by <i>municipio</i> . Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por <i>municipio</i> , 2010'	45
10. Spatial distribution of the second major indigenous language by <i>municipio</i> . Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por <i>municipio</i> , 2010'	46
11. Spatial distribution of the third major indigenous language by <i>municipio</i> . Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por <i>municipio</i> , 2010'	47
12. Spatial distribution of the fourth major indigenous language by <i>municipio</i> . Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por <i>municipio</i> , 2010'	48
13. Correlation matrix of bioclimatic variables. WorldClim bioclimatic variables are prefixed with 'BIO', whereas ENVIREM bioclimatic variables are prefixed with 'EVM'	50
14. Workflow (Part C), 'biomod2' Data Formatting and MaxEnt Modeling.....	52
15. Workflow (Part D), Ensemble Forecasting and Total Consensus Ensembling	54
16. ROC curves for the average weighted-mean ensemble predictions according to each maize landrace used to calculate AUC values.....	63
17. Cluster dendrogram using the correlation matrix of distances of variable importance between each pair of maize landraces using Ward's hierarchical clustering method with correlative distances. Approximately unbiased (AU) values are highlighted in red, with values > 95 indicating strong support by	

the data. Five groups identified from the cluster are enclosed in dashed boxes.	66
18. Combined correlation matrix of <i>I</i> similarity statistic (upper triangle) and variable importance distances (lower triangle) with 5 different groups identified than when using Ward’s hierarchical clustering method on variable importance alone.....	67
19. Multivariate Environmental Similarity Surface (MESS) maps using all predictor variables for 1970 – 2000 (top) and 2041 – 2060 (bottom) time periods.	69
20. Multivariate Environmental Similarity Surface (MESS) maps using only bioclimatic predictor variables for 1970 – 2000 (top) and 2041 – 2060 (bottom) time periods.....	70
21. Ensemble forecast of alpha diversity of all maize landraces for 1970 - 2000	75
22. Ensemble forecast of alpha diversity of Group 1 maize landraces for 1970 – 2000.....	77
23. Ensemble forecast of alpha diversity of Group 2 maize landraces for 1970 – 2000.....	77
24. Ensemble forecast of alpha diversity of Group 3 maize landraces for 1970 – 2000.....	78
25. Ensemble forecast of alpha diversity of Group 4 maize landraces for 1970 – 2000.....	79
26. Ensemble forecast of alpha diversity of Group 5 maize landraces for 1970 – 2000.....	80
27. Significant clusters of Group 2 mean diversity by <i>municipio</i> using Getis-Ord with 99,999 permutations using False Discovery Rate at $p = 0.001$	88

28. Detail of figure 27, focusing on the five geographic clusters of significantly high Group 2 mean diversity by <i>municipio</i> using Getis-Ord with 99,999 permutations using False Discovery Rate at $p = 0.001$	89
29. Conditional plot of Group 2 quantiles of mean diversity by levels of marginality, with <i>municipios</i> colored by indigenous population percent with a 1.5 hinge (POBPCNT).....	91
30. Overall maize diversity change, calculated by subtracting forecasted 2041 – 2060 alpha diversity from 1970 – 2000 alpha diversity.	92
31. Overall maize diversity change, calculated by subtracting forecasted 2061 – 2080 alpha diversity from 1970 – 2000 alpha diversity.	93
32. Group 1 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.....	95
33. Group 2 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.....	96
34. Group 3 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.....	97
35. Group 4 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.....	98
36. Group 5 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.....	99
37. Bivariate local indicator of spatial association (BiLiSA) map of Group 2 mean diversity change and indigenous population percentage by <i>municipio</i>	105

ABSTRACT

Anthropogenic climate change is expected to have dramatic impacts on natural and human systems worldwide in the near future. Ensuring food security for an exponentially-increasing world population requires extensive research from a variety of disciplines. Maize, a globally-important food crop, evolved from a single location in the Balsas River Basin in southwestern Mexico thousands of years ago via artificial selection from indigenous Mexican communities. Ample research has evaluated the effects of bioclimatic and societal influences on maize diversity and distribution in Mexico in contemporary eras. This research utilizes ensemble maximum-entropy ecological niche modeling (ENM) using biomod2 and MaxEnt in R to evaluate current and future patterns of traditional Mexican maize landrace ($n = 46$) distributions and diversity using high-resolution data at the national-level in Mexico. This research uncovered an unambiguous statistical association between indigeneity and the distribution of maize diversity across Mexico from groupings of maize landraces derived from hierarchical clustering of ensemble variable importance loadings. Forecasts of maize landrace distributions to the 2041 – 2060 time-period reveal drastic changes in traditional Mexican maize suitability nationally, with a median decrease in current suitable maize land area of 51% across landraces modeled by the 2041 – 2060 time-period. This research further finds that particular indigenous communities in southern Mexico are at higher risk of being impacted by maize diversity loss via climate change relative to non-indigenous populations for specific maize groupings identified. This thesis identifies the locations of

these communities to prioritize future sustainable maize agricultural practices, as well as community outreach, in order to maintain current levels of maize productivity and diversity in Mexico, but also linkages to the associated ancestral food traditions.

I. INTRODUCTION

Anthropogenic climate change is expected to have devastating impacts on natural and human systems worldwide in the near future (Walther, G.-R. et al., 2002; IPCC 2007, 2013; Kang, Y, S. Khan 2009; Hoegh-Guldberg, O. 2010). Analysis of proxy and observational data suggest that global average annual temperatures have increased at unprecedented rates during the past 100 years (Houghton JT et al. 2001). For the past three decades surface temperatures have been successively warmer than any decade previous to 1850 (IPCC 2013). Regardless of the source of heightened atmospheric greenhouse gas (GHG) concentrations, GHGs undoubtedly trap longwave radiation within Earth's atmosphere, altering temperature and precipitation patterns, and leading to profound, though nuanced, impacts at both global and regional scales (Sala et al., 1999; IPCC 2007, 2013; Mendelsohn 2008; Weyant et al. 2009).

While some research suggests that certain crops might respond positively to elevated carbon dioxide emissions associated with climate change (Jablonski, Wang, and Curtis 2002; Tubiello and Fischer 2007), all of the climate projections from the Intergovernmental Panel on Climate Change (IPCC) suggest that the effects of climate change on domesticated crop yields and productivity will be negative, regardless of the emission scenario implemented (Brown and Funk 2008; Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012; IPCC 2013). Further, external effects of anthropogenic climate change, including global disruption of terrestrial and marine ecologies (Walther et al. 2002; Parmesan 2006), rising average global sea level (IPCC 2013), destabilization of global and regional food networks (Brown and Funk 2008; Kang, Y, S. Khan 2009; Chung et al. 2014), prolonging of severe droughts (Dai 2011), and societal impacts on

vulnerable communities (UNU-IAS 2008; Esperón-Rodríguez, Bonifacio-Bautista, and Barradas 2016), are expected to far outweigh any positive effect that excessive CO₂ levels might have on crop productivity.

Maize (*Zea mays* subsp. *mays* L.) — together with wheat, rice, potatoes, barley, sweet potatoes, and cassava — supplies half of human nutrition globally (Kehoe 2003; Sealing 2003). Maize is one of the most widely-consumed cereal grain worldwide, supplying a staple food crop for more than 200 million people worldwide (Nuss and Tanumihardjo 2010; Ureta et al. 2013). Maize provides an estimated 15% of global human protein intake and 20% of global human calories (Nuss and Tanumihardjo 2010). This is especially true in Mexico, maize's Center of Domestication (COD), where maize constitutes a major source of protein and carbohydrates for all ethnic groups (Ruiz Corral et al. 2008).

Mexico's traditional maize landraces account for 60% of maize genetic diversity globally (Ureta et al. 2013). As a domesticated crop, maize is intimately contingent upon societal influences. The high genetic diversity of maize and maize landraces today is inextricably related to both the environmental conditions and the cultural traditions within which they evolved. Each unique landrace population makes tangible the dynamic evolution of maize through varied environmental conditions, management techniques, breeding agendas and cultural assemblages experienced over thousands of years. As such, maize landrace morphology and genetics today reflect the lengthy and dynamic genealogy of maize within societies through history. Many Mexican maize farmers still implement traditional agrotechnology, regulating maize productivity as highly dependent upon environmental factors; however, the genetic makeup of domesticated crops is

inherently shaped by artificial selective forces. Historically, seed management practices by indigenous communities, together with environmental forces and random gene flow, altered the diversity and distribution of maize in Mesoamerica. Today, social factors, including community ethno-linguistic diversity, racial composition, and seed management practices, still influence maize landrace diversity and distribution in Mexico.

At present, more than one-half of Mexico's cropland is used to cultivate traditional maize landraces (Perales and Golicher 2014), harboring at least 59 reputed traditional maize landraces and reportedly thousands of varieties of landraces (Table 1). Mexican maize farmers primarily grow traditional maize landraces for human consumption, with commercial and hybrid maize seed restricted to only a small percent of farms (Perales, Benz, and Brush 2005). This trend lies in stark contrast with maize's lifecycle in the U.S., which grows primarily hybridized and genetically-modified varieties, 90% of which is directed to non-human consumption, namely for ethanol, feedstock, corn syrup, or international export, increasingly to Mexico.

This complex network of interactions amongst environment conditions, social factors, and maize diversity has intrigued scholars for decades, leading to voluminous research from a variety of disciplines. Because of the increased risks that climate change is expected to induce on agriculture, an increasing amount of research has evaluated the influences of climate change on maize production to assess future potential risks with one of the world's most valuable crops. Situated on that body of work, this research hopes to assess the impacts that global climate change might have on the distribution and diversity

of traditional maize landraces in Mexico to further identify the communities that might be most heavily affected.

Table 1 – Maize landraces with sample size in entire georeferenced CONABIO dataset (Total) with sample sizes after cleaning data of erroneous records (Cleaned). Then, count in presence/absence matrix (PAM) for spatially-filtered grid at 1 km². P/C Ratio refers to the count in the PAM divided by the ‘cleaned’ count, with ‘-’ indicating maize landraces with < 15 samples in PAM that were excluded from modeling, or failed in modeling. Tau refers to the estimated *a-priori* prevalence value utilized in modeling, where values marked with ‘ ’ indicate default prevalence in lieu of data (discussed in more detail in ‘Analysis and Technique’). ND = Prime maize landrace not determined, and thus not utilized. Data via CONABIO (2015).

	Landrace	Total	Cleaned	PAM	P/C	Tau
1	Ancho	324	314	177	0.563	0.034
2	Apachito	65	62	53	0.854	0.020
3	Arrocillo Amarillo	316	306	112	0.366	0.010
4	Azul	99	99	84	0.848	0.015
5	Blando	44	41	26	0.634	0.045
6	Bofo	37	37	23	0.621	0.109
7	Bolita	533	524	210	0.400	0.043
8	Cónico	2,000	1,915	811	0.423	0.016
9	Cónico Norteño	1,188	1,154	732	0.634	0.195
10	Cacahuacintle	68	62	57	0.919	0.033
11	Celaya	939	910	566	0.621	0.011
12	Chalqueño	631	602	366	0.607	0.015
13	Chapalote	31	29	15	0.517	0.011
14	Chiquito	43	43	18	0.418	0.006
15	Choapaneco	5	5	1	-	-
16	Comiteco	1,280	1,278	174	0.136	0.073
17	Complejo Serrano de Jalisco	18	18	9	-	-
18	Conejo	119	115	87	0.756	0.106
19	Coscomatepec	119	117	55	0.470	0.003
20	Cristalino de Chihuahua	313	306	148	0.483	0.019

Table 1 (cont.)– Maize landraces with sample sizes after cleaning data of erroneous records. Maize landraces with <15 sample were excluded from modeling (n = 46). ND = landrace not determined.

	Landrace	Total	Cleaned	PAM	P/C	Tau
21	Cubano Amarillo	65	65	40	0.615	0.010'
22	Dulce	42	42	34	0.809	0.085
23	Dulcillo del Noroeste	37	34	21	0.617	0.043
24	Dzit Bacal	116	100	72	0.72	0.046
25	Elotero de Sinaloa	77	76	69	0.907	0.035
26	Elotes Cónicos	572	557	400	0.718	0.059
27	Elotes Occidentales	395	382	275	0.719	0.109
28	Gordo	69	68	58	0.852	0.039
29	Harinoso de Ocho	2	2	2	-	-
30	Jala	43	39	19	0.48	0.004
31	Mixeño	12	9	6	-	-
32	Mixteco	4	4	3	-	-
33	Motozinteco	5	5	2	-	-
34	Mountain Yellow	8	8	6	-	-
35	Mushito	382	373	172	0.461	0.039
36	Mushito de Michoacán	54	54	36	0.666	-
37	Nal-tel	148	122	95	0.778	-
38	Nal-tel de Altura	19	18	16	0.888	-
39	ND	4119	-	-	-	-
40	Negrito	8	8	7	-	-

Table 1 (cont.)– Maize landraces with sample sizes after cleaning data of erroneous records. Maize landraces with <15 sample were excluded from modeling (n = 46). ND = landrace not determined.

	Landrace	Total	Cleaned	PAM	P/C	Tau
41	Negro de Chimaltenango	2	1	1	-	-
42	Olotón	583	582	89	0.152	0.085
43	Olotillo	1111	1086	494	0.454	0.005
44	Onaveño	96	84	62	0.738	0.097
45	Palomero de Chihuahua	6	6	5	-	-
46	Palomero de Jalisco	3	3	1	-	-
47	Palomero Toluqueño	53	43	36	0.837	0.017
48	Pepitilla	313	299	193	0.645	0.041
49	Quicheño	1	1	1	-	-
50	Ratón	467	431	362	0.839	0.229
51	Reventador	90	86	55	0.639	0.063
52	Serrano	8	8	5	-	-
53	Serrano Mixe	31	31	13	-	0.002
54	Tablilla de Ocho	44	43	35	0.813	0.110
55	Tabloncillo	602	581	371	0.638	0.105
56	Tabloncillo Perla	196	181	97	0.535	0.084
57	Tehua	45	43	22	0.511	0.004
58	Tepecintle	528	510	275	0.539	0.079
59	Tuxpeño	3,454	3,232	1298	0.4011	0.200
60	Tuxpeño Norteño	327	300	191	0.636	0.225
61	Uruapeño	1	1	1	-	-
62	Vandeno	316	307	232	0.755	0.135
63	Zamorano Amarillo	81	80	64	0.8	0.041
64	Zapalote Chico	140	137	67	0.489	0.005
65	Zapalote Grande	84	81	63	0.777	0.056

II. IMPORTANCE OF RESEARCH

The preservation of agrobiodiversity at Centers of crop Domestication (COD) and diversification is critical for global food security, particularly in the current era of climate change (Thrupp 2000; Esquinas-Alcázar 2005; Ureta et al. 2013). As such, evaluating the impacts of climate change on traditional Mexican maize landrace distribution is a timely investigation. This research aims to visualize the geographic distributional change of maize suitability to aid in identifying priority areas needed for maize conservation efforts and sustainable agricultural practices. In addition, the study will attempt to identify and visualize the location of indigenous communities that might be most heavily impacted by negative maize distributional change. The motivation for this latter goal lies in recognition that indigenous communities are likely to suffer disproportionate impacts of anthropogenic climate change (IPCC 2007; Liu et al. 2013), despite their relatively negligible contributions to it (UNU-IAS 2008; IPCC 2013). In addition, indigenous peoples can offer valuable contributions to a variety of discussions, including climate change debates, low-impact agricultural techniques, sustainable seed management practices, and expanded local climatological knowledge, particularly relating to sub-grid scale relationships of climate and agriculture (J Luis Hernandez-Stefanoni, Pineda, and Valdes-Valadez 2006; Alexander, Chief Clarence, Nora Bynum, Liz Johnson, Ursula King, Tero Mustonen, Peter Neofotis, Noel Oetlé, Cynthia Rosenzweig, Chie Sakakibara, Chief Vyacheslav Shadrin, Marta Vicarelli, Jon Waterhouse 2011; Sánchez-Cortés and Chavero 2011). For that reason, failing to consider the locations in which indigenous populations coincide with projected negative maize distribution change may result not only in a loss of rich genetic information, but also irrecoverable deep

knowledge and wisdom surrounding the cultivation and management of a globally important crop.

III. PURPOSE STATEMENT

Distributions of maize landraces are incontrovertibly influenced by environmental forces (Ruiz C., Sánchez G., and Goodman 1998; Benz, Perales, and Brush 2007; Brush and Perales 2007; Ruiz Corral et al. 2008; Ureta et al. 2013; Perales and Golicher 2014; García-Martínez et al. 2016). They are further influenced, as research has shown, by ethnicity and other sociocultural forces (Brush and Perales 2007; Bird 2010; Ureta et al. 2013; Perales and Golicher 2014). While there has been ample research on the effects of climate change on maize distribution, few studies have sought to better understand how climate change-induced effects on maize might be linked to indigeneity at a national level (Kehoe 2003; Monterroso Rivas et al. 2011; Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012).

This research aims to model the distribution of maize landraces using their associated social and bioclimatic environments. Through a modeling approach, this thesis strives to (1) identify which maize landraces are most susceptible to extinction with climate change, and, in turn, (2) identify which indigenous communities are most vulnerable to negative maize distribution change; (3) verify projections of future maize distributions from previous research (Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012; Perales and Golicher 2014; Ureta et al. 2015); and (4) identify how the current geographic associations of indigenous populations and maize landraces might change within the next century. Assessing the current relationships between distinct indigenous populations and maize landraces will provide both the groundwork for evaluating the temporal variability of these relationships and a provide launching point to commence

more qualitative discourse with indigenous populations regarding maize productivity and habitat at localized scales.

IV. RESEARCH QUESTIONS

This thesis draws on a wealth of literature that examines the social, biogeographical, and local climatological conditions that have historically influenced maize distribution and productivity in Mexico.

To examine the intersectionality of maize diversity, climate change, and indigeneity in Mexico, this research addresses the following questions:

1. Where do maize landrace exist and what are the spatial patterns of Mexico's maize landrace diversity? How does indigeneity covary with their distributions and diversity?
2. How is climate change expected to alter current maize landrace distributions and diversity?
3. Where might climate change, through changing maize's distribution and diversity, have the greatest impacts on indigenous communities?

V. CONCEPTUAL FRAMEWORK

This research's conceptual framework (Fig. 1) draws upon a heuristic diagram commonly used to formulate species distribution modeling methodologies (Peterson and Soberón 2012; Zimmermann, Niklaus E. Thuiller, Wilfried Guisan 2017). The Biotic-Abiotic-Mobility (BAM) framework utilizes a Venn diagram composed of three ecological spaces – (1) Biotically-suitable (B), (2) Abiotically-suitable (A), (3) and Mobile (M) – within Geographic (G) space to understand the intricacies of species distribution modeling (SDM) and/or ecological niche modeling (ENM). Areas within the intersection of the B, A, and M spaces represent locations in ecological space that is most suitable for a species and constitute its 'realized niche' (1's in Figure 1).

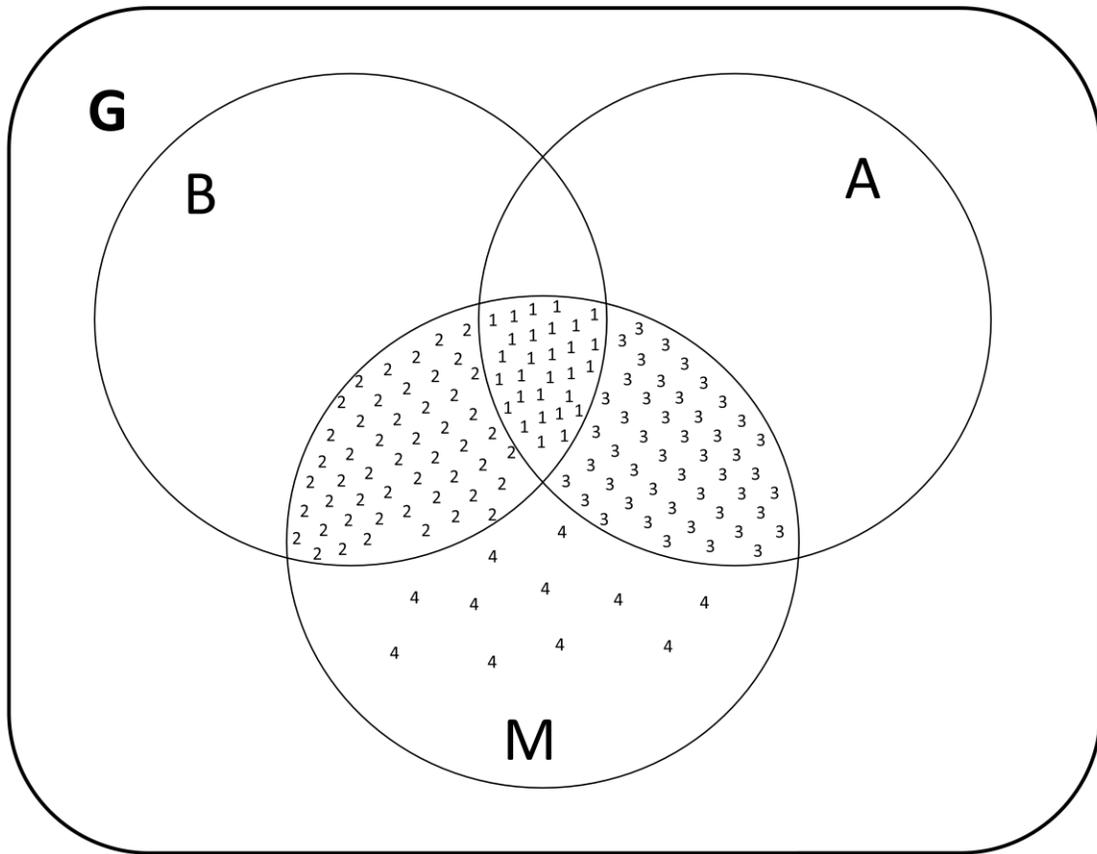


Figure 1 – Heuristic BAM diagram depicting the interaction between *biotically-suitable environments* (B), *abiotically-suitable environments* (A), and *mobility* (M) interplay to constitute species ecological niches. One's (1) indicate observed species that exist in both *biotically-suitable & abiotically-suitable* environments. Two's (2) indicate observed species that exists outside of *abiotically-suitable* environments, but still persist. Three's (3) represent observed species that exists outside of *biotically-suitable* environments, but still persist. Four's represent observed species that exists outside of both *biotically-suitable & abiotically-suitable* environments, but still persist. These species observations ultimately influence the resultant ecological niche model. Following (Peterson and Soberón 2012; Zimmermann, Niklaus E. Thuiller, Wilfried Guisan 2017), who followed (Soberón, 2007).

Individuals or communities *outside* this middle area of the BAM diagram could exist for theoretically innumerable reasons. Migration, for instance, could result in sink populations in abiotically unsuitable environments ('2s' in Figure 1). Alternatively, individuals within abiotically suitable environments but outside of a biotically suitable environment might exist in an ecosystem with strong competition ('3s' in Figure 1). Individuals outside of both biotically- and abiotically-suitable environments might exist as historical remnants of previous populations ('4s' in Figure 1).

Within the scope of maize distribution and diversity, modeling the influences of both biotic and abiotic variables on the realized niches of traditional Mexican maize landrace are primary goals of this research. Abiotic influences are especially explored through this thesis, as biotic interactions are typically only discernable at very large geographic scales/small spatial scales and require more intensive methodologies.

VI. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Maize Domestication and History

Maize was first domesticated from the wild varieties of the Mexican annual teosinte (*Z. mays* ssp. *Parviglumus* and ssp. *mexicana*) via artificial selection by indigenous cultures in southern Mexico at least 9,000 years ago (Matsuoka et al. 2002; Kato et al. 2009). Although profoundly different morphologically, early research by G.W. Beadle note that maize and teosinte varieties are so genetically similar that they can

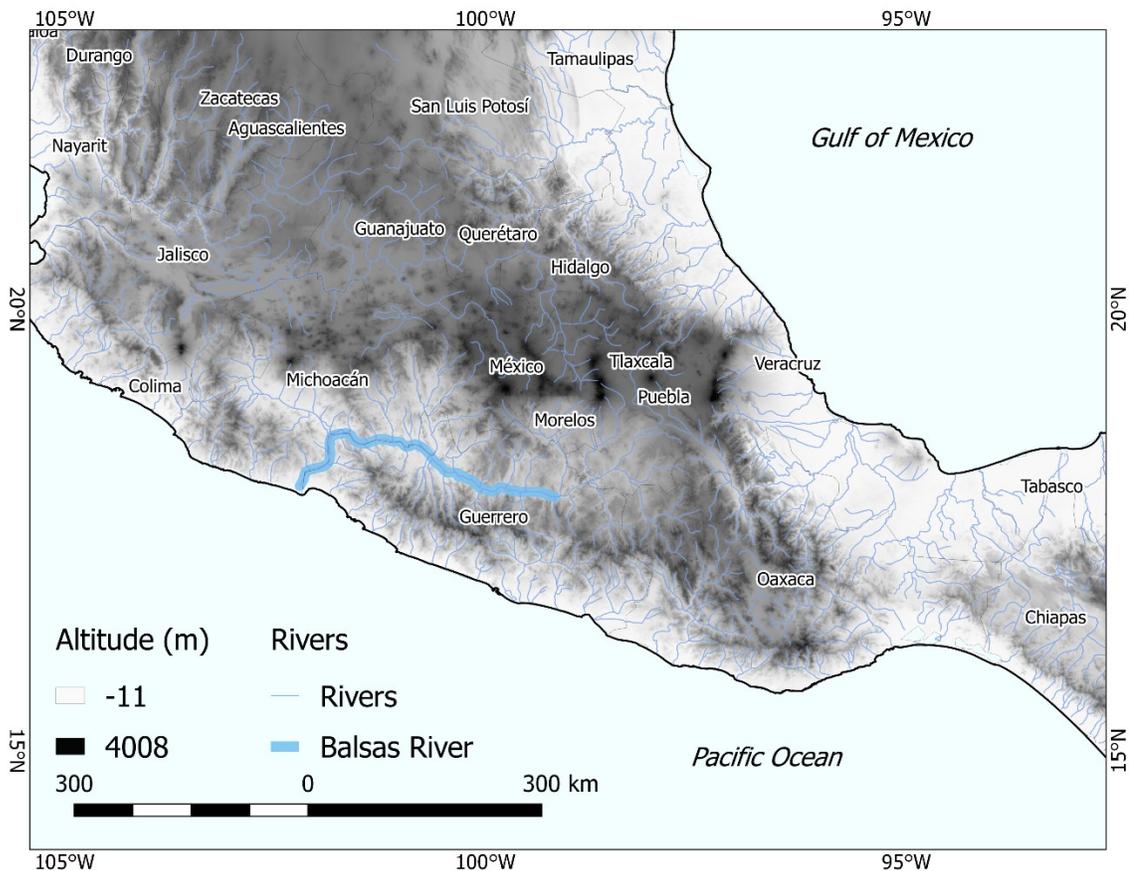


Figure 2 - Location of the Balsas River, overlaid the 30 arc-second WorldClim altitude raster to depict the profound drainage basin, with scalebar.

hybridize to produce fertile hybrids (Matsuoka 2005), which was later supported by molecular systematics studies (Doebley 1990). Although the specific origin of domestication is debated, genetic microsatellite data by Matsuoka et al. (2002) constitutes the majority consensus that maize had a single domestication origin at high altitudes in the Balsas River drainage basin in southwest Mexico (Fig. 2). From here, maize varieties then migrated in two major pathways into North and South America (Matsuoka et al. 2002).

The term “maize landrace” was first formalized by Edgar Anderson and Hugh Cutler in 1942 to describe maize diversity in Mexico (Cutler 1942; Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012). Anderson and Cutler’s systematic classification for maize diversity was later adapted by Wellhausen and colleagues (E.J. Wellhausen et al., 1952) to incorporate not only morphological characteristics, but also genetic, cytological, and agronomic characteristics to enforce natural, rather than artificial, systematic taxonomy (Perales and Golicher 2014). Wellhausen et al. (E.J. Wellhausen et al., 1952) loosely defined maize races as “a group of related individuals with enough characteristics in common to permit their recognition as a group,” well aware of the continuous variation in many maize landraces (Perales and Golicher 2014). It should be noted that morphological characteristics used by farmers to organize maize “types” does not correlate with genetic diversity (Perales and Golicher 2014). While formal recognition of maize landraces are sometimes contentious or ambiguous (Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012; Perales and Golicher 2014), few advancements have been made in refining this system. Despite this, maize races nevertheless have been shown to exhibit

different ecological spaces (Anderson 1947; Perales, Benz, and Brush 2005; Ruiz Corral et al. 2008).

Both historically and presently grown with traditional, rain-fed agrotechnology, traditional maize landrace cultivation is largely dependent upon climate and weather. As such, maize landraces have evolved to adapt to an incredible array of conditions, and the crop is now grown globally in a variety of climates. Maize landraces have been grown in altitudes ranging from sea-level to 2,900 m a.s.l., with growing season (two seasons: Spring/Early Summer and Late Summer/Fall) mean temperatures ranging from 12.0 °C to 29.1 °C, and with average growing season precipitation ranging from 400 mm to 3555 mm (Ruiz Corral et al. 2008). Topographic variables, particularly elevation, greatly influence maize landrace distributions (Ruiz C., Sánchez G., and Goodman 1998; Perales R., Brush, and Qualset 2003; Perales and Golicher 2014; Ureta et al. 2015). Maize landraces, in contrast with commercial cultivars, have also been noted to grow in highlands with steep slopes (Perales and Golicher 2014). Other early works further established the biogeography of maize in Mesoamerica (Cutler 1942; Anderson 1947; E.J. Wellhausen, L.M. Roberts, E.Hernandez X., Paul C 1952; Dyer and López-Feldman 2013), indicating continuous variation in phenotypes of maize landraces, but with regional clusters of landraces that share greater similarity than with other landraces in other regions. While environmental differences have been noted as the primary predictors of maize diversity and distribution, ethnic diversity appears to have an influence on maize diversity at finer geographic scales (Brush and Perales 2007).

Maize as a Mirror of Humanity

All domesticated crops reflect, through their morphology, fine adaptations in response to their agro-ecological environments and artificial selective practices experienced through their evolution. This is especially true of maize, whose selective breeding patterns have contributed greatly to its diverse array of phenotypes (Fig. 3).

Some of the first research on maize landraces by Edgar Anderson (Anderson, 1947) note “maize [as] a sensitive mirror of the people who grow it,” establishing in academia the theory of fluid interactions between maize and peoples. Since then, associations between seed management practices (Anderson 1947; Dyer and López-Feldman 2013), indigenous religious practices (Rendón-Aguilar et al., 2015), ethnolinguistic diversity (Perales, Benz, and Brush 2005; Brush and Perales 2007; Rivero-Romero et al. 2016), indigenous agricultural knowledge (García-Martínez et al. 2016; Rivero-Romero et al. 2016), and agro-technological practices (García-Martínez et al. 2016; Rivero-Romero et al. 2016; Toledo and Barrera-Bassols 2017) with maize diversity have been evaluated.

As a domesticated crop, seed management techniques greatly influence maize diversity. Maize farmers tend to use different outcropping methods with multiple maize varieties and wild ancestors with open-pollination to increase genetic diversity, to promote, for example, drought tolerance (Taba 1995; Wen et al. 2012). These artificial selective practices are educated according not only to agronomic characteristics of maize varieties, but also to the intended use and quality (e.g.: flavorful maize for tortillas, dry maize for storage, poor quality maize for feedstock, etc.) (Bellon and Brush 2017).

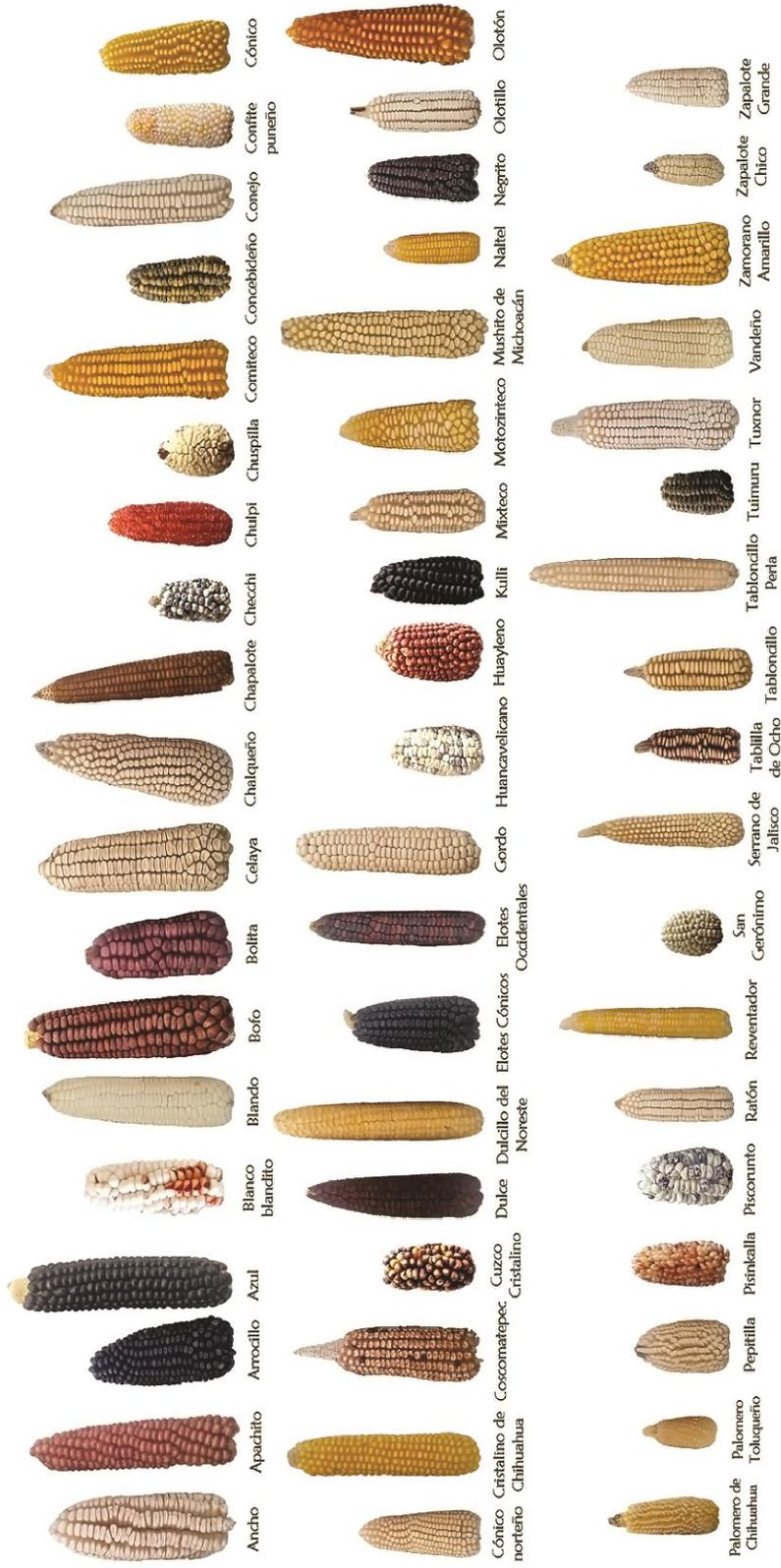


Figure 3 - A visual representation of phenotypic diversity of maize landraces in Mexico. Photo courtesy of the University of Florida & the International Maize and Wheat Improvement Center.

Today, many indigenous Mexican populations grow numerous landraces and varieties of maize with traditional agrotechnology. While indigenous maize production is seemingly insignificant to global maize production at the outset, indigenous knowledge has been proven useful in assessing climate change (Dewalt, 1994; Alexander, 2011), ecology (Hernandez-Stefanoni et al., 2006), and agro-technology at localized-levels. Indigenous knowledge can also contribute to global scientific knowledge and understanding of climate systems (Alexander et al, 2011). Traditional indigenous knowledge of local climate and weather can even predate climatological records, supplementing intellectual voids of a local climate's history. For instance, knowledge of weather, climate, and soil by the Zoque peoples of Mexico can go back many generations, and have been noted to support scientific and technical planning (J. Luis Hernandez-Stefanoni, Pineda, and Valdes-Valadez 2006; Sánchez-Cortés and Chavero 2011). It is also noted that indigenous communities practice different maize seed management techniques using traditional knowledge. Tzotzil farmers in Chiapas, Mexico can associate specific maize varieties with specific habitats and climates (Benz, Perales, and Brush 2007). These seed management techniques also differ drastically between indigenous and Latino/Mestizo populations (Hugo, Brush, and Qualset 2003; Perales, Benz, and Brush 2005; Dyer and López-Feldman 2013). In Chiapas, Latino communities are reportedly more inclined to sow commercial maize varieties, while indigenous producers typically maintain traditional varieties (Brush and Perales 2007). In these ways, indigenous communities in Mexico contribute positively to global and regional climate change planning and mitigation by offering unique knowledge of local climates, seed management, and sustainable agrotechnology both globally and within Mexico.

The evolution of traditional maize landraces is inextricably entwined with indigenous peoples of Mexico, who domesticated maize from wild teosinte varieties thousands of years prior to colonization of the New World. Today, 68 distinct indigenous macro-languages from 11 different language families are spoken in Mexico, for a total of 364 dialects (Table 2). While the ethnolinguistic diversity of Mexico is quite remarkable, its stability is dubious. Mexican legislation such as the Law of Linguistic Rights, promulgated in 2003, requires *federales* (states) to offer services to indigenous communities in their native languages, yet many indigenous languages and cultures in Mexico are still in danger of extinction. While roughly 12% of the nation's population identifies as indigenous, accumulating to nearly 6 million people, only 6% of the population over age 5 speaks an indigenous language (INEGI, 2015). Nahuatl, the language group of the Aztecs, is the only language group that is spoken by more than 1 million people (Table 2). These 68 language groups, and roughly 364 language dialects, are spread thinly over nearly 6 million indigenous peoples in Mexico, and declining (INEGI, 2015). The proportion of those who speak an indigenous language to those who identify as indigenous in Mexico has steadily declined through the 20th and 21st centuries. This trend will likely continue and may further decrease due to the increased effects of social vulnerability associated with climate change.

Table 2 – Indigenous language groups in Mexico by language family, including number of dialects per language group and estimated number of speakers above the age of 5.

Language family	Language Group	Number of Dialects	Estimated Number of Speakers above age 5
<i>Algica</i>	Kickapoo	1	157
<i>Chontal de Oaxaca</i>	Chontal de Oaxaca	3	3,453
<i>Cochimí-Yumana</i>	Cucapá	1	116
	Kiliwa	1	36
	Kumiai	1	298
	Ku'ahl	1	
	Paipai	1	200
	<i>Huave</i>	Huave	2
<i>Maya</i>	Akateko	1	532
	Awakateko	1	21
	Chontal de Tabasco	4	32,584
	Chuj	1	2,180
	Ch'ol	2	185,299
	Huasteco	3	149,532
	Ixil	2	77
	Jakalteko	1	400
	Kaqchikel	1	105
	K'iche'	3	251
	Lacandón	1	734
	Mam	5	7,492
	Maya	1	758,310
	Qato'k	2	110
	Q'anjob'al	1	8,526
	Q'eqchí'	1	1,070
	Teko	1	61
	Tojolabal	1	43,169
	Tseltal	4	371,730
	Tsotsil	7	329,937
<i>Mixe-Zoque</i>	Ayapaneco	1	2
	Mixe	6	115,824
	Oluteco	1	63
	Popoloca de la Sierra	1	28,194
	Sayulteco	1	2,583

Table 2 (continued) – Indigenous language groups in Mexico by language family, including number of dialects per language group and estimated number of speakers above the age of 5.

Language family	Language Group	Number of Dialects	Estimated Number of Speakers above age 5
<i>Mixe-Zoque</i>	Texistepequeño	1	238
	Zoque	8	54,002
<i>Oto-Mangue</i>	Amuzgo	4	43,761
	Cuicateco	3	12,610
	Chatino	6	42,791
	Chichimeco Jonaz	1	1,625
	Chinanteco	11	125,706
	Chocholteco	3	616
	Ixcateco	1	213
	Matlatzinca	1	1,134
	Mazahua	2	111,840
	Mazateco	16	206,559
	Mixteco	81	423,216
	Otomí	9	239,850
	Pame	2	9,720
	Popoloca	4	16,163
	Triqui	4	23,846
	Tlahuica	1	842
	Tlapaneco	9	98,573
	Zapoteco	62	410,906
	<i>Seri</i>	Seri	1
<i>Tarasca</i>	Tarasco	1	105,556
<i>Totonaco-Tepehua</i>	Tepehua	3	8,321
	Totonaco	7	230,930
<i>Yuto-Nahua</i>	Cora	8	17,086
	Guarijío	2	1,648
	Huichol	4	35,724
	Mayo	1	32,702
	Náhuatl	30	1,376,026
	Pápago	1	116
	Pima	3	738
	Tarahumara	5	75,371
	Tepehuano del Norte	1	6,809
	Tepehuano del Sur	3	24,782
Yaqui	1	14,162	

Source: Instituto Nacional de Lenguas Indígenas (INALI) Diversidad Lingüística. Categorías Lingüísticas por agrupación lingüística 2005.

Hutchinson's Niche

The approach to this research is strongly guided by developments of the concept 'ecological niche' in the 20th century. Very broadly, Joseph Grinnell and Charles Elton pioneered early conceptualizations of biological niches by systematically associating environmental attributes to specific niches, thereby regulating niches to particular environments (e.g.: tropical rainforest, boreal forest, coral reef, etc.) or roles within their environments (e.g.: predator of small mammals, parasite, etc.), respectively. George Evelyn Hutchinson revolutionized this definition by claiming ecological niches as the roles a community or species performs at any particular time within its habitat (Colwell and Rangel 2009). The ecological niche, as defined by Hutchinson, is an attribute of a species in relation to its biotic and abiotic environment to n -level of ecological space (Colwell and Rangel 2009; Holt 2009). Under this definition, a species' ecological niche defines its role within its habitat but is constituted neither solely by its habitat nor by its biotic interactions.

While Grinnell's and Elton's work maintained abstracted conceptions relating ecological variables to hyper-dimensional n -space, Hutchinson's modification provides a necessary duality between niche space and habitat space that allow for statistical and theoretical linkages between distributions in those spaces (Colwell and Rangel 2009). Habitat locations, parameterized with values of n environmental variables corresponding to n niche axes can be projected into "ordinary physical space" with this duality (Hutchinson 1957; Colwell and Rangel 2009). Conceptually, and mathematically, this relationship is not reciprocal. While a point in geographic space corresponds to a single point in ecological niche space, a point in n -dimensional niche space may represent

multiple or zero spaces in “ordinary” space. The non-reciprocal nature between niche space and geographic space can be attributed to three causes: 1) the geographic space corresponding to n -dimensional niche space does not actually exist (e.g.: habitat disruption due to deforestation), 2) the species is prevented from occupying that habitat space due to dispersion limitations (e.g.: geographic features, such as a mountain ridge, limiting migration), 3) the species’ geographic distribution is restricted due to biotic interactions (e.g.: predation, disease, competition, mutualism, etc.) (Colwell and Rangel 2009).

Indeed, Hutchinson’s niche is pivotal in today’s statistical modeling of n -dimensional ecological space to geographic space. It provides the conceptual backbone behind statistical niche modeling, but also provides an avenue with which to define two key nuances of niche space: fundamental and realized niches.

Fundamental and Realized Niches

The fundamental niche is a hyper-dimensional niche volume in which “every point ... corresponds to a state of the environment that would permit [a] species ... to exist indefinitely” (Hutchinson 1957; Holt 2009). This definition obviates migration for clarity. The term “fundamental niche,” therefore, refers to a species’ potential geographic distribution in true 4D space, constrained by its n -dimensional bioclimatic envelope. It is a theoretical construct that cannot be directly observed or modeled (Phillips and Dudík 2008).

Hutchinson argued that two species that share resources cannot coexist in a “limited system” (Pulliam 2000), therefore requiring a true “realized niche” smaller than the fundamental niche. A species’ niche is inherently limited not only by its abiotic

environment, but also by its biotic interactions with other species. Therefore, a species' realized niche represents the portion of a species' fundamental niche that exists in reality and that does not overlap with other species niches (Griesemer 1992). This relationship between niche and distribution becomes increasingly complex when recognizing species interaction, including species' niches that are altered by other species, or species that are other species' niches (e.g.: parasites) (Pulliam 2000). These discussions are even further complicated when introducing metapopulations and source-sink dynamics (Pulliam 2000); however, this research does not develop these distinctions further.

Ecological Niche Modeling (ENM)

Hutchinson's theoretical development of ecological niches provides the necessary terminology, as well as theoretical and conceptual framework, with which to compartmentalize and analyze complex ecological niches.

In general, ecological niche modeling (ENM, also: environmental niche modeling, habitat suitability modeling, species distribution modeling, and bioclimatic envelope modeling) refers to the use of georeferenced species survey data together with environmental data collected at survey locations to map species distributions in geographic space (Booth et al. 2014). Ecological niche models utilize species presence and/or absence observation data and ecological data at those observation sites to estimate complex relationships of species niche (Elith et al. 2011). ENM has been shown to accurately model species' potential geographic distributions (Tsoar et al. 2007; Miguel B Araújo 2008; Elith, Kearney, and Phillips 2010).

Presence with true absence observation data require extensive, costly sampling (Phillips, Dudík, and Schapire 2004); therefore, presence-only observation data tend to be

more commonly available. For these reasons, environmental niche model algorithms using presence-only data have been vigorously researched in the 21st century.

Numerous ecological niche modeling algorithms have been developed to project ecological niche spaces into geographic spaces. The machine learning algorithm MaxEnt (Phillips, Dudík, and Schapire 2004), short for *maximum entropy*, is the most commonly used presence-only software/algorithm utilized for ecological niche modelling; however, numerous other algorithms have been used for ENM including Genetic Algorithm for Rule-Set Production (GARP) (Stockwell and Peters 1999), BIOCLIM (Booth et al. 2014), DOMAIN (Carpenter, Gillison, and Winter 1993), Environmental Niche Factor Analysis (ENFA) (Hirzel et al. 2002), Artificial Neural Networks, Random Forest (Breiman 2001; Cutler et al. 2007), Boosted Regression Trees (Elith, Leathwick, and Hastie 2008; Friedman 2011), and Support Vector Machines (A. R. J. Hijmans et al. 2017).

MaxEnt

MaxEnt, a stand-alone machine-learning Java application, is one of the most widely used programs for ENM and SDM with presence-only data, due largely to its high predictive performance (Elith et al. 2006, 2011) and low number of tuning parameters. The software has received criticism as a ‘black-box’ algorithm in the past. Indeed, very few ENM applications published between 2013 – 2015 show that authors utilized non-default parameters (Morales, Fernández, and Baca-González 2017). The recent open-sourcing of the MaxEnt Java application¹, as well as an open-source R package², hopes

¹ <https://github.com/mrmaxent/Maxent>

to address these qualms by inviting developers to positively contribute to any theoretical or methodological issues.

The algorithm utilizes a sample of data across the landscape (called background data) to contrast with samples extracted from observation locations. MaxEnt leans on the principle of Maximum Entropy derived from (Jaynes 1957) to make predictions from incomplete information (Phillips, Dudík, and Schapire 2004). It estimates the target probability distributions of maximum entropy – the most spread out or closest to uniform distribution (i.e.: default or *a priori* species prevalence, *tau*) – subject to the constraints imposed by the incomplete data, and subsequently relaxed by user-defined regularization parameters (Phillips, Dudík, and Schapire 2004).

As indicated by the open-sourcing of MaxEnt, the algorithm is still actively being developed. Originally, MaxEnt sought to estimate the density of presences across a landscape, which inherently assumes a randomly-sampled landscape (Merow, Smith, and Silander 2013). As such, the defined ecological/geographic landscape utilized in modeling has great effects on model output (Phillips, Dudík, and Schapire 2004; Phillips and Dudík 2008; Elith et al. 2011; Song et al. 2013; Phillips et al. 2016). Because population size is generally unknown in presence-only datasets, the resultant density estimates are relative, giving Relative Occurrence Rates (ROR) -- the ‘raw’ MaxEnt output. These RORs ($P^*(z(x_i))$) take the form,

$$P^*(z(x_i)) = \exp(z(x_i) \lambda) / \sum \exp(z(x_i) \lambda)$$

² <https://cran.r-project.org/web/packages/maxnet/maxnet.pdf>

where z is a vector of J environmental variables at location x_i with λ representing the vector regression coefficients, with $z(x_i) \lambda = z_1(x_i) \lambda_1 + z_2(x_i) \lambda_2 \dots + z_j(x_i) \lambda_j$ (Phillips, Dudík, and Schapire 2004; Elith et al. 2011; Merow, Smith, and Silander 2013).

In order to translate the RORs into estimates of species observation probabilities, MaxEnt uses a logistic transformation that relies heavily on assumptions of spatial sampling bias, background sample landscape, and *a-priori* species prevalence, τ , the theoretical backings of which are still contentious (Royle et al. 2012; Hastie and Fithian 2013; Merow, Smith, and Silander 2013).

The logistic transformation of RORs into relative predicted probability relies heavily on the assumed *a-priori* prevalence to minimize relative entropy in environmental space between the predicted probability estimation and the *a-priori* probability estimation. Minimizing relative entropy in ecological spaces equates to maximizing entropy in geographic space, which is theoretically upheld regarding ecological thermodynamics.

In order to achieve such a solution, the algorithm constructs ‘features’ for each environmental covariate or combination of covariates, which represent a mathematical transformation of said covariate(s). MaxEnt currently implements 5 feature classes: 1) linear, 2) quadratic, 3) product, 4) threshold, and 5) hinge. MaxEnt achieves the final model by selecting the model that maximizes a gain function that balances model overfitting with performance with L1 regularization, the latter of which shrinks unimportant feature coefficients to zero. Maximizing the gain function results in selecting the model which best discriminates between background samples and presences (Merow, Smith, and Silander 2013). This method also provides a generative rather than a

discriminative approach (e.g.: Generalized Additive Models [GAMS]) to modeling, indicating the algorithm always converges to a solution.

The features, regularization multiplier, *a-priori* prevalence τ , and background sample landscape constitute the primary parameters required to tune in order to increase model performance significantly from default settings (Anderson and Gonzalez 2011; Merow, Smith, and Silander 2013; Radosavljevic and Anderson 2014). While the default options provided by MaxEnt generally work well across numerous taxonomic groups collectively, making them useful for modeling numerous species together, their performance may not be optimal for an individual species or taxon (Phillips and Dudík 2008).

VII. DATA AND METHODS

An overview of this research’s workflow can be found in the figure below (Figure 4). The individually labeled A, B, C, and D sub-processes are also presented as diagrammatic processes later. This research utilizes the MaxEnt algorithm to perform Ensemble Niche Modeling (ENM) in R with the ‘biomod2’ package for traditional maize landraces to forecast distributions for 3 time-periods: 1970 – 2000, 2041 – 2060, and 2061 – 2080.

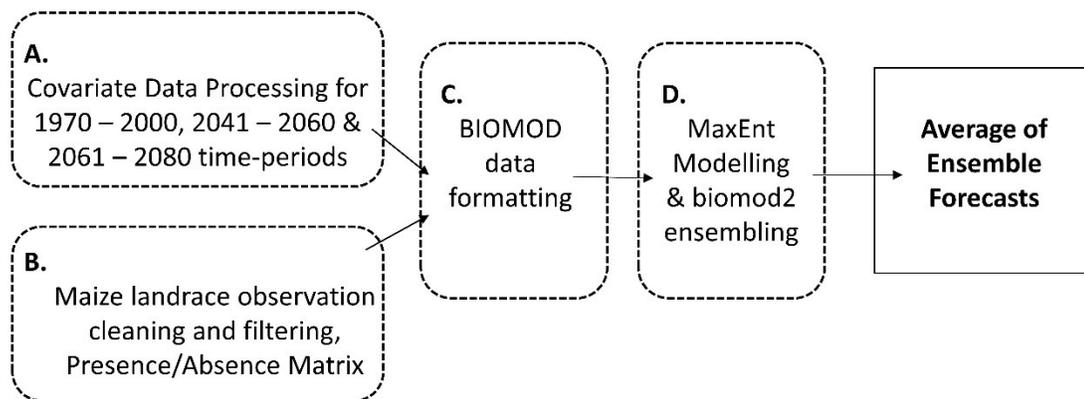


Figure 4 – Workflow overview, generally following the workflow of Ecological Niche Modeling as a whole, but with ensembling.

Site and Situation

Mexico's national border delineates this project's study area. Mexico's governmental organization, the *National Commission for the Knowledge and Use of Biodiversity* (CONABIO, Spanish acronym), provides easily accessible georeferenced maize landrace observations data from systematic sampling expeditions. Furthermore, the influence of social factors on maize niches may be obscured by studies with too large of a spatial extent.

WorldClim Climatologies

The publicly-accessible WorldClim database (Hijmans et al. 2005; Fick 2017) hosts gridded climatologies for past (Mid Holocene and Last Glacial Maximum), 'present' (1960 – 1990 & 1970 – 2000), and future time-periods (2041 – 2060 & 2061 – 2080). 'Present' climate observation data were interpolated using the thin-plate smoothing spline implemented in the ANUSPLIN software package (version 4.3), which has performed well in comparative tests (Hijmans et al. 2005) with other methods. WorldClim data is available at four spatial scales: 30 arc-second, 2.5 minutes, 5 minutes, and 10 minutes. Using the most recent 19 General Circulation Model climate projections (GCMs) from the Fifth Assessment IPCC report (CMIP-5), WorldClim provides future climate projections, downscaled and bias corrected using baseline 'current' climatological periods (i.e.: 1960 – 1990 for WorldClim 1.4 and 1970 – 2000 for WorldClim 2.0) for four Representative Concentration Pathways (RCP. 2.6, RCP 4.5, RCP 6.0, and RCP 8.5) and two time-periods (2041 – 2060 avg. & 2061 – 2080 avg.). These Representative Concentration Pathways (RCPs) represent the proposed greenhouse

gas trajectories established by IPCC's Fifth Assessment Report (AR5), superseding AR4 (Intergovernmental Panel on Climate Change 2007). The RCP names are derived from the respective possible ranges of radiative forcing values in the year 2100, relative to pre-industrial values (+2.6, +4.5, +6.0, and +8.5) (W/m^2) (Weyant et al. 2009). All of these RCPs are considered equally valid at present, though evidence suggests the world is currently on-track for the unabated emissions scenario -- RCP 8.5 W/m^2 .

Maximum temperature ($^{\circ}\text{C}$), mean temperature ($^{\circ}\text{C}$), and precipitation (mm) data were downloaded from WorldClim 2.0 for 'present' conditions (1970 – 2000) and from WorldClim 1.4 for 2041 – 2060 & 2061 – 2080 time-periods (RCP 8.5) at 30 arc-second resolution. This data was subsequently clipped to Mexico's national boundary (x-min: -117.625 $^{\circ}$, x-max: -86.20833 $^{\circ}$, y-min: 14.025 $^{\circ}$, y-max: 33.225 $^{\circ}$) and stacked. WorldClim 2.0 solar radiation monthly means were utilized for both present and future time-periods for subsequent data processing due to data availability constraints with the WorldClim 1.4 database.

As stated, WorldClim provides future downscaled climate interpolations for numerous GCMs. This research utilized future monthly means, mean-ensembled (Fig. 5) across 5 GCMs: CCSM4 (Community Climate System Model from UCAR) (Gent et al. 2011), MIROC5 (Model for Interdisciplinary Research on Climate) (Watanabe et al. 2010), MPI-ESM-LR (from the Max-Planck Institute) (Giorgetta et al. n.d.), HADGEM2-ES (from the Met Office Hadley) (Jones et al. 2011), and GFDL-CM3 (from the Geophysical Fluid Dynamics Laboratory) (Griffies et al. 2011). These GCMs represent the updated analogues used in (CONDE et al. 2011) & the 5th National Communication of Mexico for the United Nations Framework Convention on Climate Change (2012).

The CMIP-3 analogue GCMs utilized in (CONDE et al. 2011) perform singularly better than a mean-ensemble of all CMIP-3 GCMs in predicting historical climatic conditions. The ensembling of future climate models aims to reduce uncertainty and bias inherent in each individual model. By testing a GCM's predictive performance in hindcasting for a given study area, it follows that this GCM will perform well in this study area in the future.

WorldClim and ENVIREM Bioclimatic Variable Processing

Using 1970 – 2000 climatologies and mean-ensembles for 2041 – 2060 and 2061 – 2080 time-periods, ecologically important ‘bioclimatic’ variables were calculated using the R packages ‘dismo’ (A. R. J. Hijmans et al. 2017) and ‘ENVIREM’ (Pascal O. Title and Bemmels 2017) (Figure 5). Bioclimatic variables aid in ecological niche modeling performance by better representing species bioclimatic envelopes and biological requirements. Bioclimatic variables were first utilized for ecological niche modeling and species distribution modeling within the ‘BIOCLIM’ algorithm (Booth et al. 2014, Nix, 1986). Many of the original 16 bioclimatic variables originally devised are still commonly used today. Using monthly mean temperature (T_{mean}), maximum temperature (T_{max}), minimum temperature (T_{min}), and precipitation (mm) grids, 19 bioclimatic variables, corresponding to those found in the WorldClim database, were calculated using the ‘dismo’ R package (A. R. J. Hijmans et al. 2017; R Core Team 2017b) (Table 3). Using WorldClim 2.0 solar radiation monthly means and T_{mean} , T_{max} , T_{min} , and precipitation grids for present and future time-periods, 15 newly-proposed complementary variables to the WorldClim database were calculated using the ‘ENVIREM’ package in R (Pascal O. Title and Bemmels 2017) (Table 4).

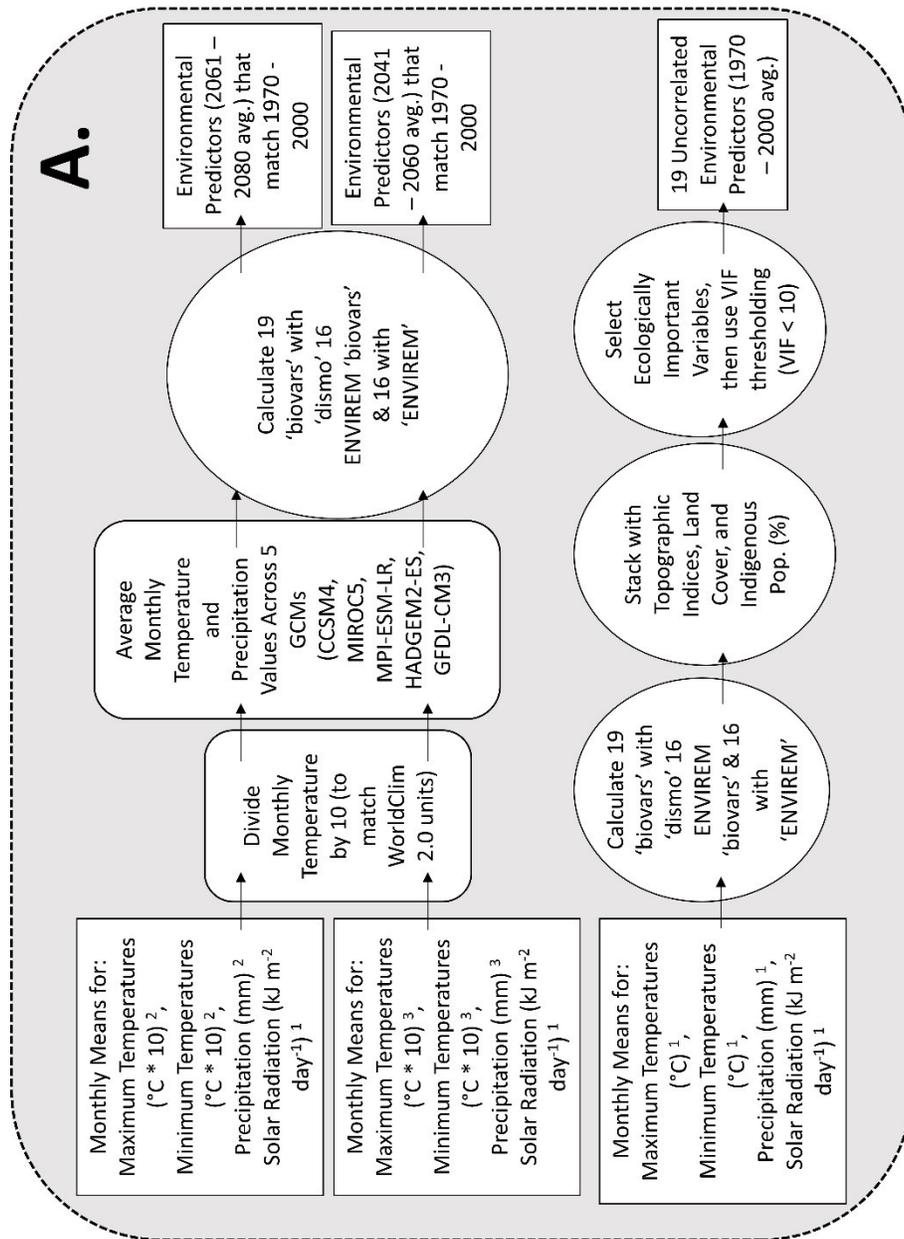


Figure 5 – Workflow (Part A), bioclimatic variable processing and variable selection by VIF thresholding.

Table 3 - Nineteen bioclimatic variables calculated with 'dismo' in R, corresponding to those found in the WorldClim database.

Variable	Description and units
BIO1	Annual T_{mean} (°C)
BIO2	Mean Diurnal Range (Mean of monthly (T_{max} - T_{min})) (°C)
BIO3	Isothermality (Mean Diurnal Range / Temperature Annual Range) (* 100) (°C)
BIO4	Temperature Seasonality (standard deviation *100) (°C)
BIO5	T_{max} of Warmest Month (°C)
BIO6	T_{min} of Coldest Month (°C)
BIO7	Temperature Annual Range (T_{max} of Warmest Month - T_{min} of Coldest Month) (°C)
BIO8	T_{mean} of Wettest Quarter (°C)
BIO9	T_{mean} of Driest Quarter (°C)
BIO10	T_{mean} of Warmest Quarter (°C)
BIO11	T_{mean} of Coldest Quarter (°C)
BIO12	Annual Precipitation (mm)
BIO13	Precipitation of Wettest Month (mm)
BIO14	Precipitation of Driest Month (mm)
BIO15	Precipitation Seasonality (Precipitation Coefficient of Variation [COV])
BIO16	Precipitation of Wettest Quarter (mm)
BIO17	Precipitation of Driest Quarter (mm)
BIO18	Precipitation of Warmest Quarter (mm)
BIO19	Precipitation of Coldest Quarter (mm)

Table 4 – Fifteen bioclimatic variables calculated using ‘ENVIREM’ in R, newly-proposed complementary extensions to the WorldClim’s bioclimatic variables.

Variable	Description and units
EVMannPET,	Annual Potential Evapotranspiration (PET)
EVMthornthwaiteAI	Thornthwaite Aridity Index
EVMclimaticMI	Climatic Moisture Index
EVMcontinentality	Continentalty
EVMembergerQ	Emberger Q
EVMgrowingDegDays0	Growing Deg Days > 0 °C
EVMgrowingDegDays5	Growing Deg Days > 5 °C
EVMminTempWarmest	T _{max} for Coldest Quarter
EVMmonthCountByT10	Month Count by Temp 10 (°C)
EVMPETColdestQ	PET Coldest Quarter
EVMPETDriestQ	PET Driest Quarter
EVMPETseas	PET seasonality
EVMPETWarmestQ	PET Warmest Quarter
EVMPETWettestQ	PET Wettest Quarter
EVMthermicityIndex	Thermicity Index

CONABIO Maize Survey Records

Maize landrace observation data was downloaded from CONABIO’s geospatial portal³. CONABIO hosts an amalgamation of species observation databases from over 168 national and international datasets. This includes observations for 62 Mexican maize landraces collected between 1940 – 2010 from numerous systematic sampling accessions. Aside from observation location, these data also include numerous attributes regarding the collection and sampling parameters including collector name, collector organization, date, and consistency, among other variables. Using this information, maize observation records were systematically cleaned of erroneous location data (i.e.: Longitude = 00° 00’00’’), data flagged “inconsistent”, and missing data (i.e.: altitude and primary maize landrace observed) (Fig. 6).

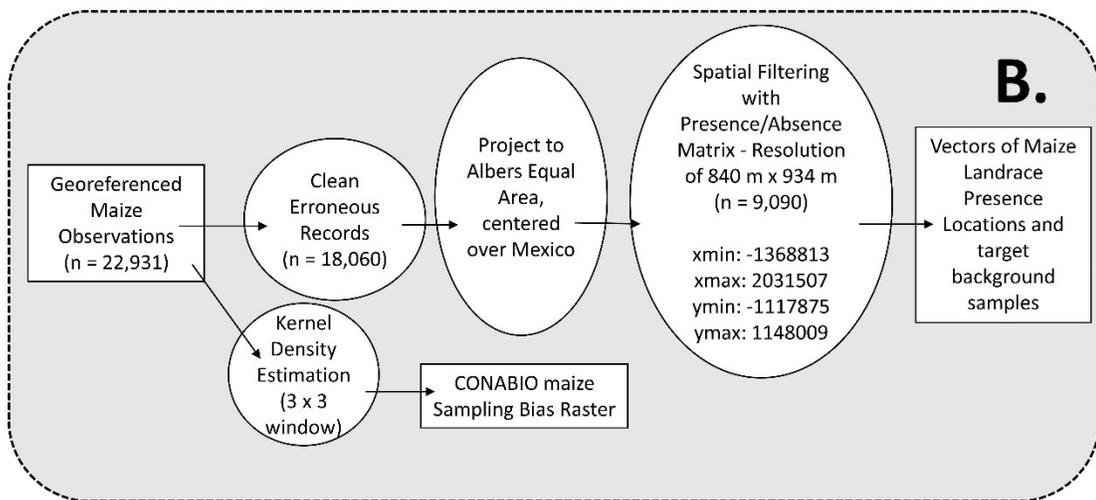


Figure 6 – Workflow (Part B), cleaning of georeferenced maize landrace observation to create Presence/Absence Matrix (PAM) for variety observation and target background sampling.

³ <http://www.conabio.gob.mx/informacion/gis/>

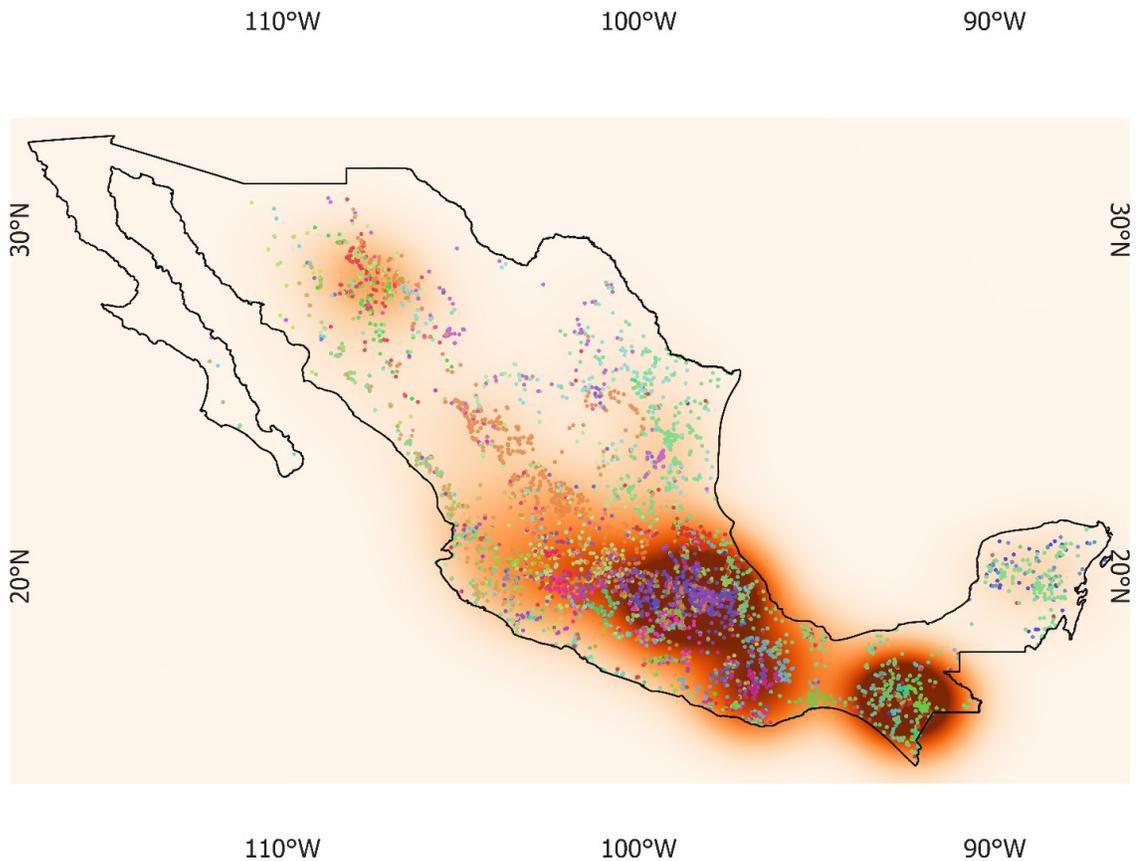


Figure 7 – Maize landrace observations, after cleaning, with each landrace assigned a different color, overlaid the Kernel Density Estimation of sampling density bias present in CONABIO maize observations.

To help reduce sampling bias present (Fig. 7, spatial bias found using Kernel Density Estimation with 3 x 3 window) in the CONABIO dataset of traditional maize landrace observations, a presence/absence matrix (PAM) was created using the ‘letsR’ package in R (Bruno Vilela and Fabricio Villalobos 2015; R Core Team 2017b) (Elith et al. 2011). The PAM utilizes the same spatial resolution and extent as the predictor raster data projected to a modified version of the North American Albers Equal Area Projection

(EPSG: 102008) but centered on Mexico⁴. PAMs provide not only a useful format with which to record species observation data, but also allow for spatial filtering of observation data to reduce spatial sampling biases: a violation of the assumptions of most ecological niche modeling methods. Projecting the data to an equal-area projection also ensures that each pixel has an equal chance of sampling, further reducing any potential violations of sampling bias, as recommended by Elith et al. (2011). Each column of the final presence/absence matrix used in modeling represents observations of maize landrace for which at least 15 observations exist following Phillips and Dudík (2008). Values of '0' can be included in the presence/absence matrix if researchers have access to reliable presence/absence data obtained from systemic surveys in classic logistic regression of species presence/absence; however, without absence data, these rows were reformatted as 'NA' to be utilized in target background sampling as 'pseudo-absences.'

DEM-derived Topographic Indices

To incorporate topographic variables as predictors for maize landrace distributions, terrain indices were derived from 30 arc-second resolution WorldClim elevation data. The WorldClim database derives its elevation data from the Shuttle Radar Topography Mission (SRTM), aggregated to 30-arc seconds. Aspect (radians) and slope (degrees) surface rasters were derived using the 'raster' package in R (Cheng et al. 2016; R Core Team 2017b). The aspect raster was cosine-transformed to achieve a linear, non-circular predictor of 'northness.' Using the 'spatialEco' package in R, a Vector Rugosity

⁴ proj4string: "+proj=aea +lat_1=14.5 +lat_2=32.5 +lat_0=24 +lon_0=-105 +x_0=0 +y_0=0 +ellps=GRS80 +datum=NAD83 +units=m +no_defs"

Measure (VRM) surface grid was calculated using the slope raster as per Sappington et al. (2007) to achieve a topographic index not correlated with slope.

Harmonized Soil Database v 1.2 Land Cover

This research utilized categorical continuous landcover raster data from the Food and Agricultural Organizations Harmonized Soil Database v1.2 (Fischer et al. 2012; Bocinsky and Kohler 2014). The landcover data include continuous percentages of (1) rain-fed cultivated land, (2) irrigated cultivated land, (3) total cultivated land, (4) forest land, (5) grass/scrub/woodland, (6) built-up land (residential and infrastructure), (7) barren/very sparsely vegetated land, and (8) mapped water bodies per 5-arc-minute resolution grid cell. This data was resampled using bilinear interpolation to match the 30 arc-second resolution of the rest of the data.

CONABIO Ethnolinguistic Data

While the relationship between climatic variables and maize landrace diversity is evident, the relationship between maize landrace diversity and social factors is more complex. Social influences on maize landrace diversity and distributions have long been theorized, (E.J. Wellhausen, L.M. Roberts, E.Hernandez X., Paul C 1952; Perales R., Brush, and Qualset 2003; Dyer and López-Feldman 2013; York and Garden 2017), but only quantified relatively recently at the national level (Ureta et al. 2013).

Mexico's 2010 Population and Housing Census includes different data regarding for indigenous demographic. While manipulation of such data (e.g.: using Factor Analysis to describe the primary spatial patterns of indigeneity in Mexico) would likely be useful in follow-up studies, this research utilized a simple indigenous population

percentage rasterized to 30 arc-second resolution data using data at the *municipio* geographic level (INEGI, 2010). This was chosen as an exploratory device to determine if any broad influence of indigenous population percentage (Fig. 8) on maize landrace distributions can be discerned at this scale. Follow-up research might include further evaluating associations of risk, climate change, and indigeneity in Mexico, perhaps utilizing maize distribution model outputs derived from this research. To identify the ethnolinguistic groups that might be most affected by maize distributional change, simple geo-visualization of distributional change over the 1st, 2nd, 3rd, and 4th (Figs. 9 - 12) most commonly spoken indigenous languages at the *municipio* geographic-level may be utilized in the future.

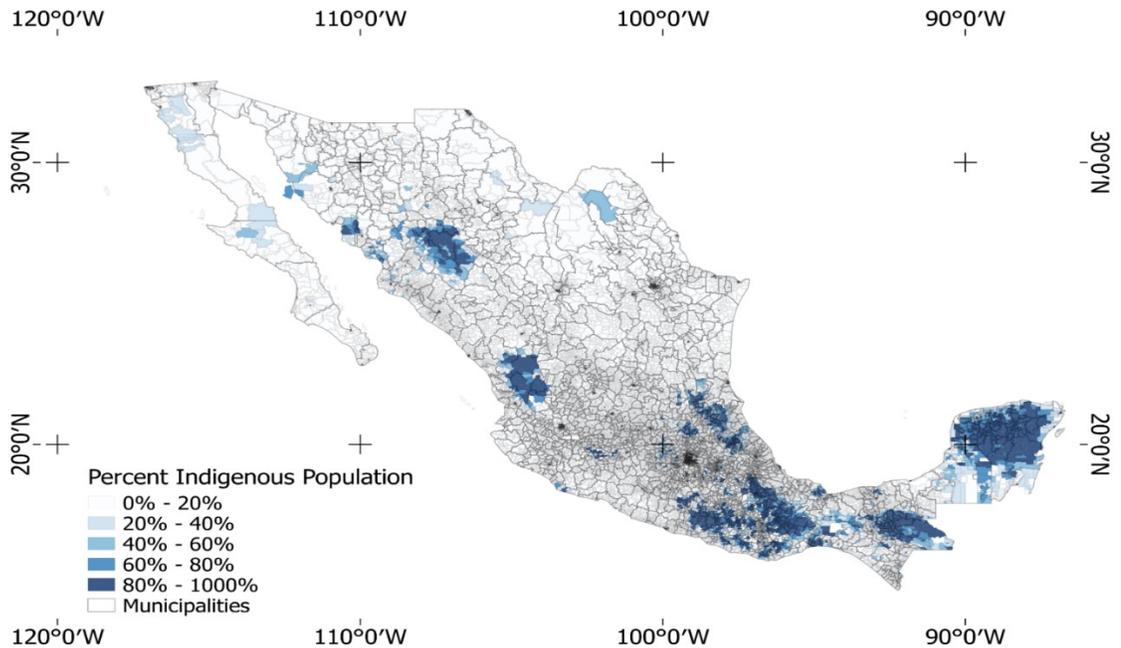


Figure 8 – Percent indigenous population by *municipio* (municipality). Data via *el Comision Nacional para el Conocimiento y Uso de la Biodiversidad* (CONABIO). 2010.

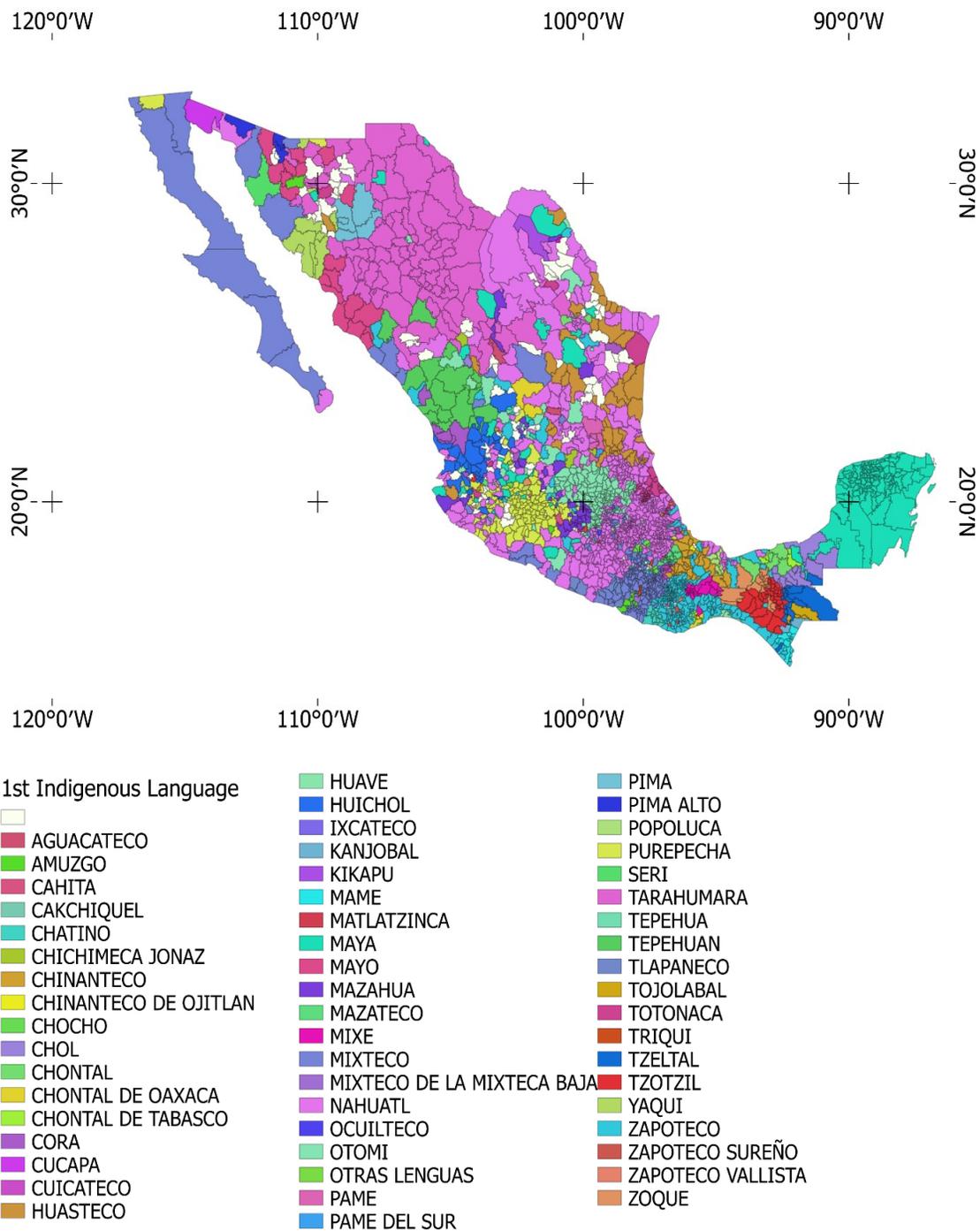


Figure 9 – Spatial distribution of the majority indigenous language by *municipio*. Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por *municipio*, 2010'.

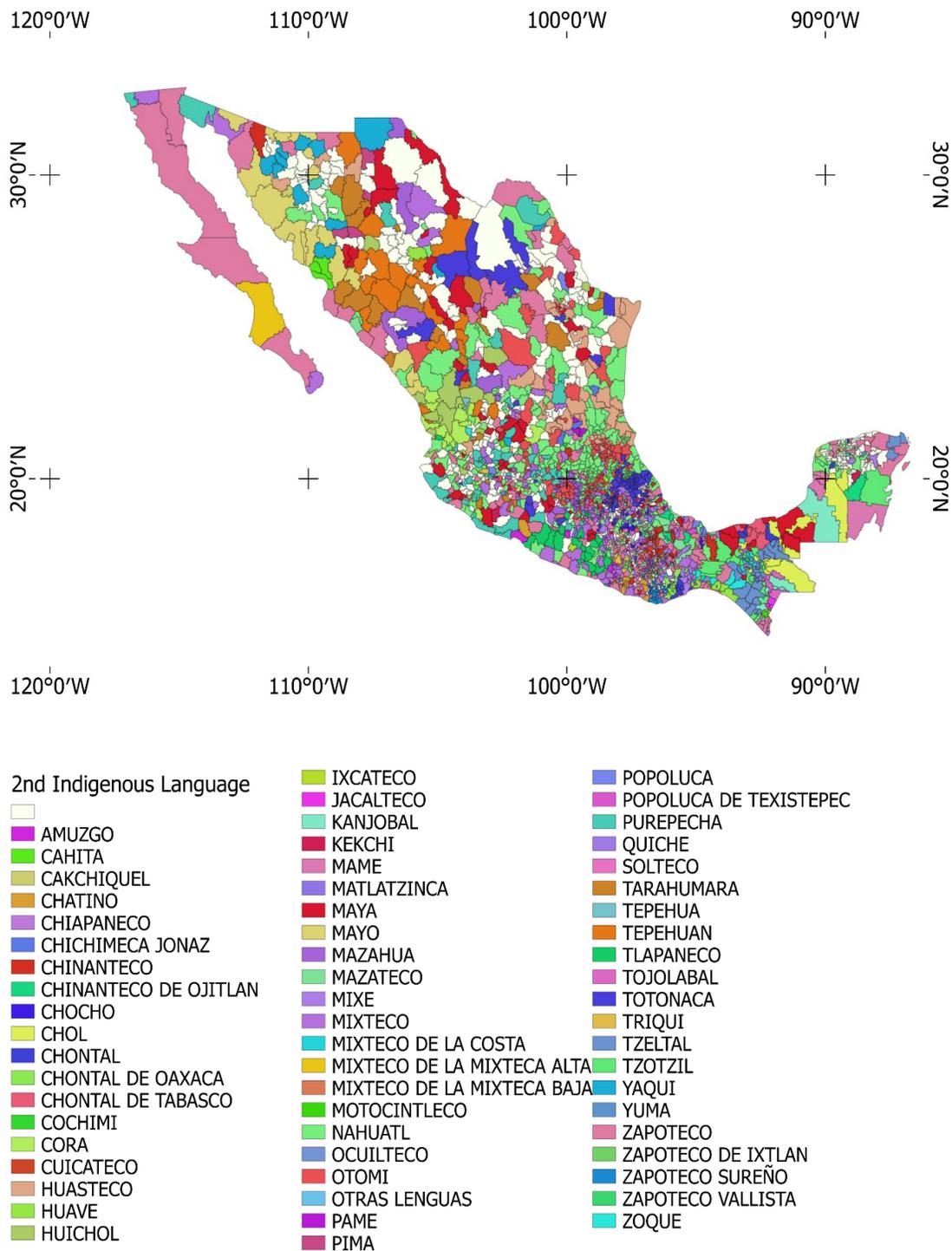


Figure 10– Spatial distribution of the second major indigenous language by *municipio*. Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por *municipio*, 2010'.

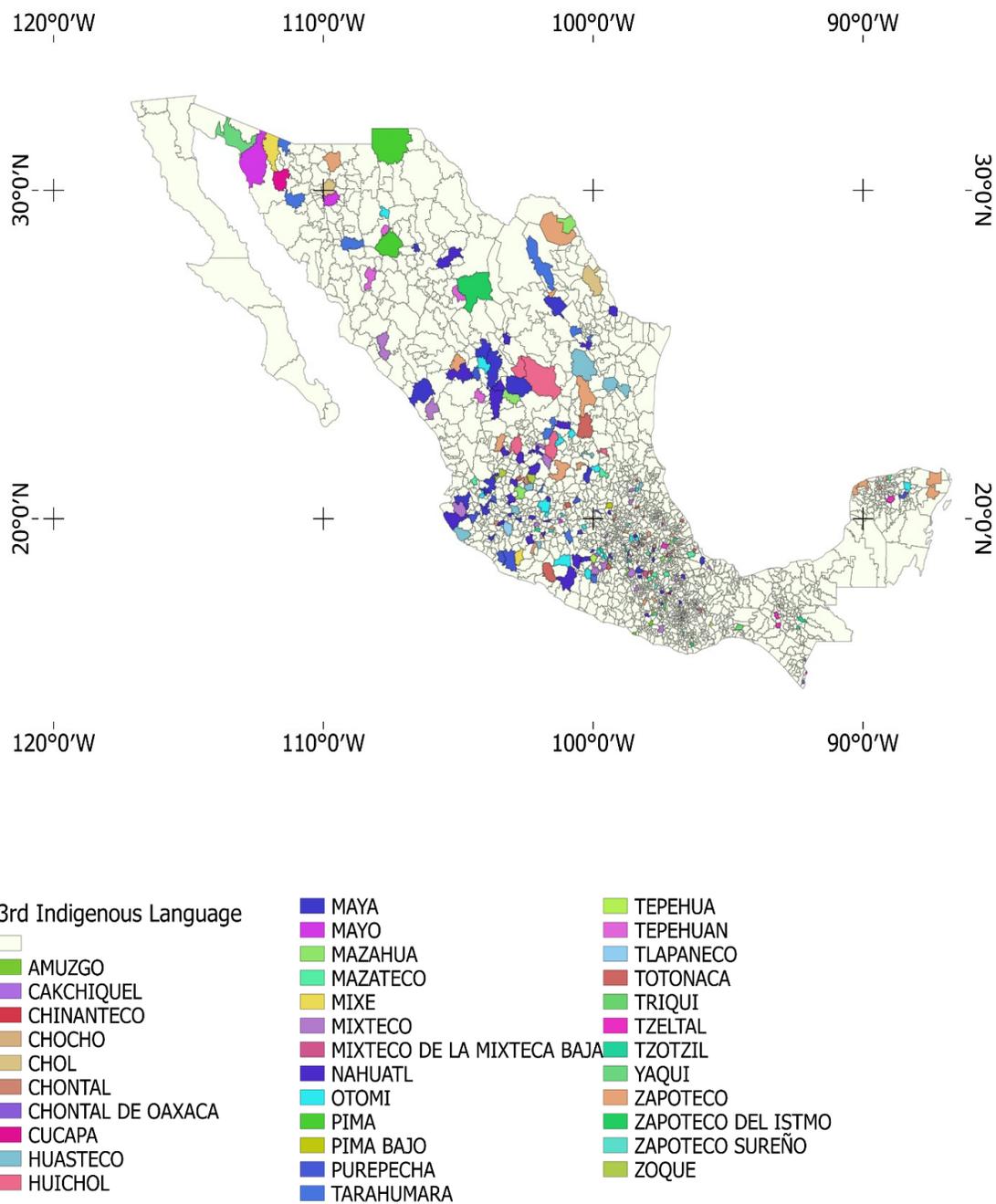


Figure 11 – Spatial distribution of the third major indigenous language by *municipio*. Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por *municipio*, 2010'.

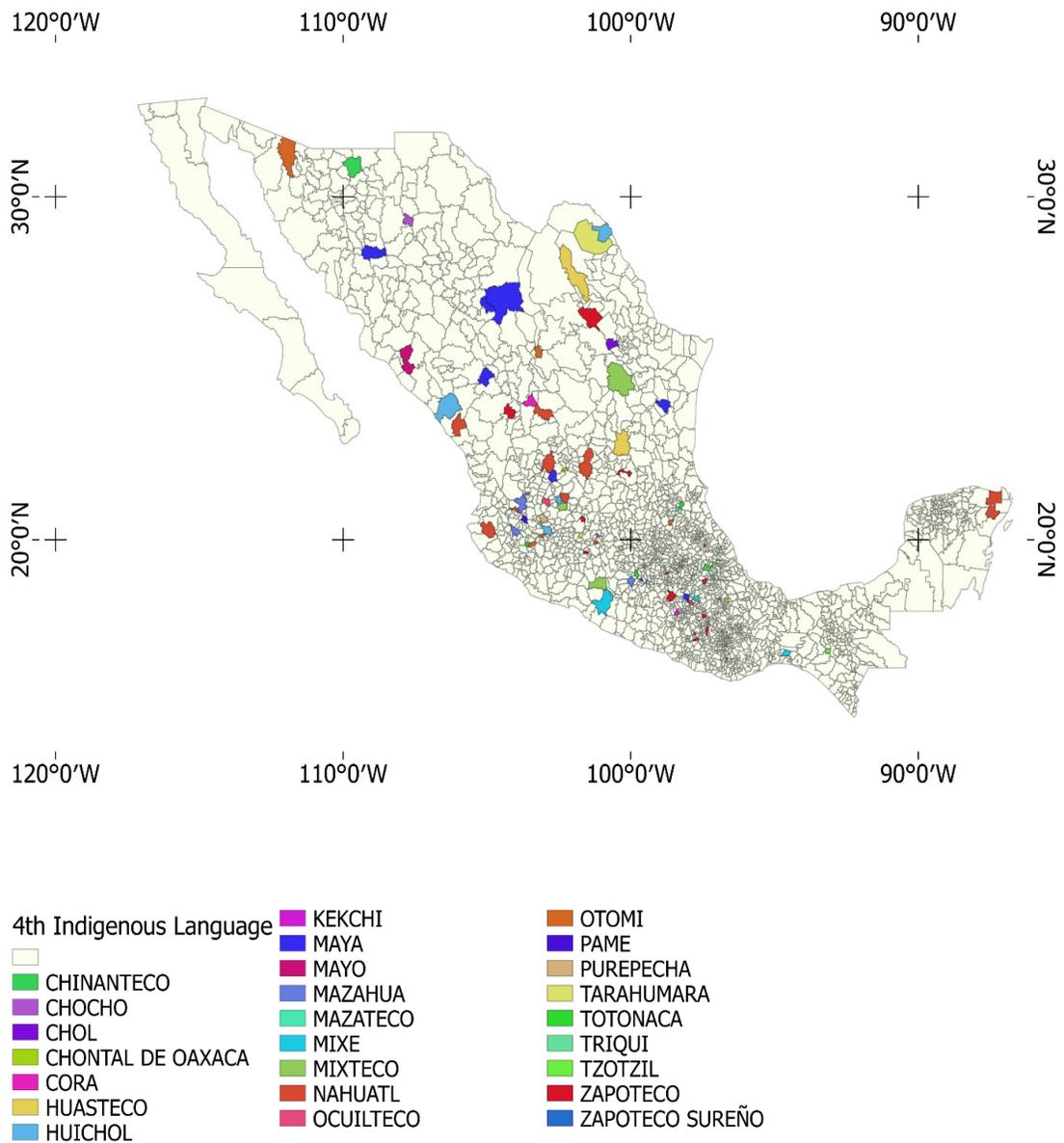


Figure 12 – Spatial distribution of the fourth major indigenous language by *municipio*. Data via the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). 'Vivienda de la población indígena por *municipio*, 2010'.

VIII. ANALYSIS AND TECHNIQUE

MaxEnt Variable Selection

Although the MaxEnt algorithm effectively reduces multicollinear variables into composite ‘features,’ and transforms certain multicollinear variable beta values to zero, the calculated variable contribution in the resultant model will be skewed. While some research reduces complex spatial data with Principle Components Analysis (PCA), these techniques rely on assumptions of variable correlation structures, which can be problematic when projecting to new geographic or temporal extents. Further, PCA complicates interpretability of variable importance in modeling. In order to maintain the highest amount of data, while eliminating multicollinearity amongst variables (to ensure variable importance interpretability) (Fig. 13), this research utilized a VIF-thresholding technique to achieve the subset of variables included in modeling. VIF-thresholding for variable selection is preferred over Pearson’s R thresholding, especially when projecting to future time-periods, as correlation structures amongst variables may change over time (Zimmermann, Niklaus E. Thuiller, Wilfried Guisan 2017). Variables were iteratively selected for ecological importance until automating final selection using *vifstep* in the ‘usdm’ R package (Naimi et al. 2014). The final subset of variables (Table 5) was achieved by selecting the variables with VIF values < 10 as suggested by (Zimmermann, Niklaus E. Thuiller, Wilfried Guisan 2017).

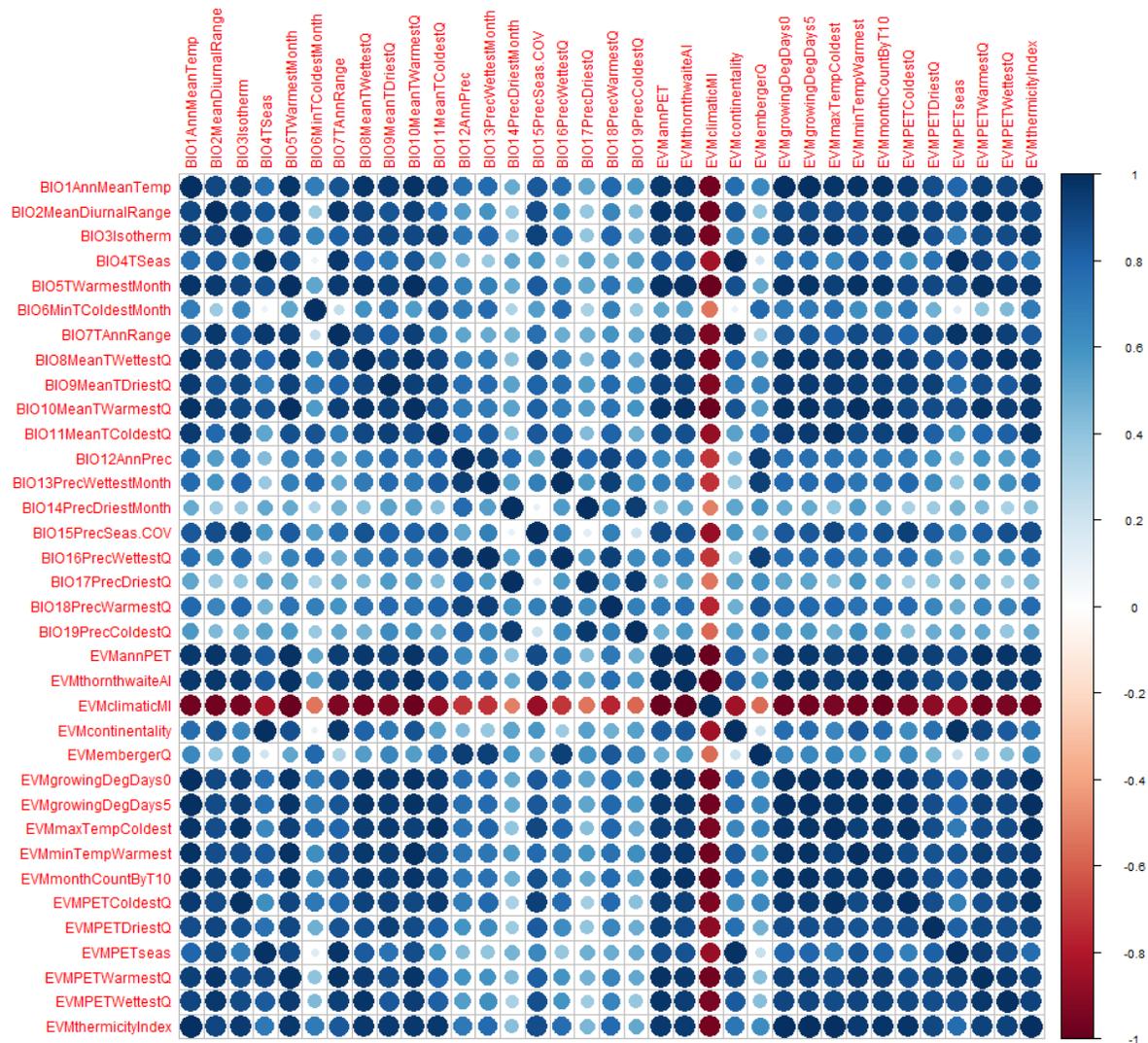


Figure 13 – Correlation matrix of bioclimatic variables. WorldClim bioclimatic variables are prefixed with ‘BIO’, whereas ENVIREM bioclimatic variables are prefixed with ‘EVM’.

Table 5 – Final 20 uncorrelated variables selected for modeling, with variable name, Variance Inflation Factor, source or R-package used for calculation, and a brief description.

Variable	VIF	Source	Description
BIO8MeanTWettestQ	6.60	dismo	Mean Temperature of Wettest Quarter (°C)
BIO13PrecWettestMonth	8.1	dismo	Precipitation of Wettest Month (mm)
BIO15PrecSeas.COV	3.23	dismo	Precipitation Seasonality (Coefficient of Variation)
BIO18PrecWarmestQ	4.66	dismo	Precipitation (mm month ⁻¹) in the warmest quarter
BIO19PrecColdestQ	3.29	dismo	Precipitation (mm month ⁻¹) in the coldest quarter
EVMminTempWarmest	6.36	ENVIREM	Minimum temperature in the warmest month (°C * 10)
EVMmonthCountByT10	2.96	ENVIREM	Count of the number of months with $T_{avg} > 10$ °C
EVMPETColdestQ	4.09	ENVIREM	Potential Evapotranspiration (PET) (mm month ⁻¹) in coldest quarter of year
EVMPETDriestQ	4.03	ENVIREM	Potential Evapotranspiration (PET) (mm month ⁻¹) in driest quarter of year
EVMPETWarmestQ	5.05	ENVIREM	Potential Evapotranspiration (PET) (mm month ⁻¹) in warmest quarter of year
IrrCult	1.25	FAO	0 – 100% of pixel classified as irrigated agriculture
Rain.fedCult	1.91	FAO	0 – 100% of pixel classified as rainfed agriculture
Grass.Woodland	2.59	FAO	0 – 100% of pixel classified as grassy woodlands
Barren	2.15	FAO	0 – 100% of pixel classified as barren
Urban	1.07	FAO	0 – 100% of pixel classified as an urban landscape
Water	1.06	FAO	0 – 100% of pixel classified as water
Cos.asp.rad	1.03	raster	Cosine-transformed Aspect (radians), a measure of ‘northness’
Slope	1.47	raster	Slope of terrain (°)
VRM	1.17	SpatialEco	Vector Ruggedness Measure, a measure of variability in slope ranging from 0 to 1; calculated with 3 x 3 window
Ind_pob_pct	1.33	INEGI	Indigenous population percentage by the <i>municipio</i> geographic level

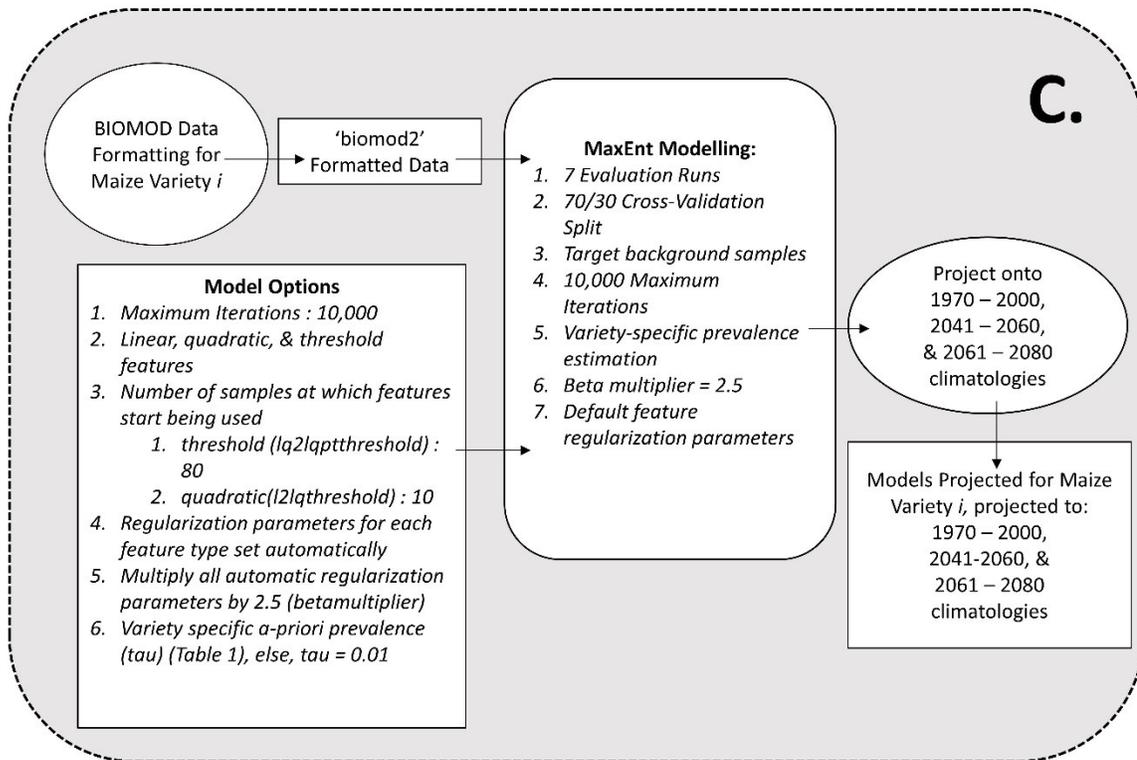


Figure 14 – Workflow (Part C), 'biomod2' Data Formatting and MaxEnt Modeling

MaxEnt in biomod2

To evaluate current ecological niche distributions of Mexican maize landraces, uncorrelated variables were utilized as covariates (or the respective transformed 'features' in MaxEnt) in ensemble modelling. Using the R package 'biomod2' (Thuiller et al. 2009), ensemble ecological niche models were run using the MaxEnt algorithm (Phillips, Dudík, and Schapire 2004) with 7 replication runs. Seventy percent of the observation data were used for model training and calibration, leaving the remaining 30 percent for model evaluation in lieu of independent calibration data. MaxEnt features were limited to hinge, linear, and quadratic features to yield response curves that minimize model over-

fitting (Merow, Smith, and Silander 2013; Radosavljevic and Anderson 2014). In order to further ensure generalizable models, the beta regularization multiplier parameter increased from 1 to 2.5, which has been shown to perform well for different species in systematic studies (Anderson and Gonzalez 2011; Radosavljevic and Anderson 2014). Regularization values for each type of feature was automatically set within MaxEnt.

Target-group background samples were adopted as background samples for MaxEnt modeling by selecting records for all maize landraces *except* the landrace being modeled. These target background samples were derived from the 5,703 coordinates in the PAM which contained at least one observations for any maize landrace. For each maize landrace *i*, the target background sample size equals 5,703 minus the number of coordinates in which maize landrace *i* was observed. These methods further aid in reducing the effects of sampling bias since the sampling bias across background locations is theoretically equal to observation bias (Costa et al. 2010; Anderson and Gonzalez 2011; Phillips et al. 2017).

A-priori prevalence values required for MaxEnt modeling were estimated for each maize landrace using potential distributions for each maize landrace provided by CONABIO (Table 1). These data are assumed to correspond with model outputs from (Perales and Golicher 2014). While incorporating sampling bias layer into model weights is possible within the MaxEnt program, this implementation is not currently available within the 'biomod2' R package. However, sampling bias is minimized due to robust estimations of default prevalence, target-group background sampling, spatial-filtering of observation data, and equal-area projected observations and gridded raster.

Each model run was evaluated according to the following evaluation metrics : (1) ‘ROC’ : Relative Operating Characteristic (Area Under the ROC Curve – ‘AUC’), (2) ‘KAPPA’ : Cohen’s Kappa (Heidke skill score), (3) ‘TSS’ : True skill statistic (Hanssen and Kuiper’s discriminant, Peirce’s skill score⁵).

The evaluation scores derived from each model run allow combining of individual models into meta ‘ensemble’ models (Fig. 15). Weighted mean ensembles were calculated according to each evaluation metric across the seven model replication runs. The weighted sum of probabilities was calculated using proportional mean weight decays, whereby attributed weights are proportional to evaluation scores.

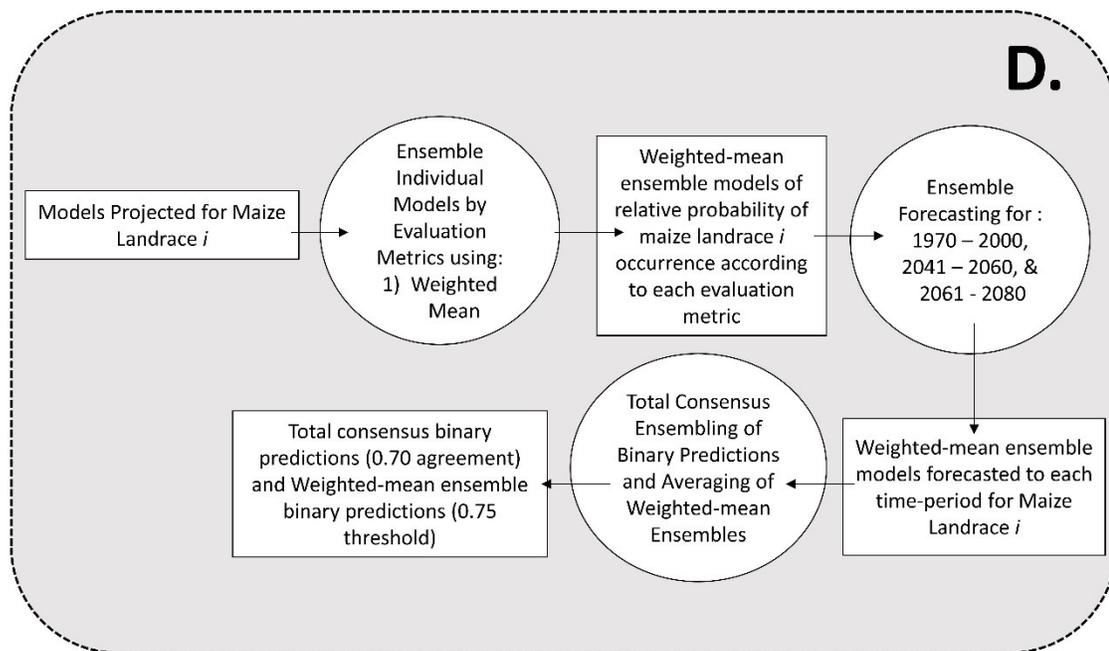


Figure 15 – Workflow (Part D), Ensemble Forecasting and Total Consensus Ensembling

⁵ Detailed descriptions of each metric can be found at CAWRC website: http://www.cawcr.gov.au/projects/verification/#Methods_for_dichotomous_forecasts

The weighted-mean models across each evaluation metric were averaged for each maize landrace to achieve a final model output of predicted relative probability (Fig. 14). Averaging the mean-weighted ensemble models allows for a single forecast projection for each maize landrace. Binary forecasts of maize landrace distributions (0 for predicted not to be present, 1 for predicted to be present) were calculated with threshold classification using a threshold value of 0.75 from the average of the weighted-mean ensemble models (i.e.: forecasted values of relative probability of presence greater than or equal to 0.75 were classified as 1, with values less than 0.75 were classified as 0). Binary predictions were also calculated by averaging the binary predictions of all ensemble models and then by classifying each pixel as 1 if the average of the binary predictions were greater than or equal to 0.70; however, these predictions were nearly identical to the former method, and so the former was utilized.

The resultant binary forecasts for each maize landrace were added together to calculate overall maize landrace alpha-diversity, as well as calculate distribution change for each maize landrace for the 2041 – 2060 and 2061 – 2080 time-periods.

Technical Details

All analyses were run on the LEAP High-Performance Computing (HPC) cluster at Texas State University in San Marcos, TX. This research utilized the ‘Rmpi’ package in R (v 3.4.1) built on OpenMpi with CentOS 6.8. GDAL v.1.7.3 and PROJ4 v. 4.6.0 were built from source for use with the ‘rgdal’ R package. Ensemble niche modeling was run using biomod2 v 3.3-15 in parallel with the ‘snowfall’ package in R using 4 parallel nodes with 28 CPUs and 128 Gb RAM each. Model projections and forecasts were run in parallel on 1 node using the ‘parallel’ package in R due to issues with snowfall or biomod2 in this stage.

All code utilized in this thesis may be found at the following web address:

www.github.com/stevenconnorg/maices-enm.

IX. RESULTS

Ensemble model performance

Using the dataset derived from the 30% data holdout, model accuracy for each weighted-mean ensemble model was calculated according to the three evaluation metrics used to ensemble the models: KAPPA, TSS, and ROC (i.e. Area Under the ROC curve – AUC).

All metrics are scaled from 0 – 1 for clarity and comparability. AUC values less than 0.7 can be considered poor, while KAPPA and TSS values below 0.4 are considered poor. KAPPA and TSS scores between 0.4 – 0.8 indicate useful models, while scores >0.8 indicate good to excellent model performance (Allouche, Omri Tsoar, Asaf Kadmon n.d.; McHugh 2012). Each metric provides a different weight to different prediction errors, such as omission (Type II errors), commission (Type I errors), etc. For example: as AUC is threshold-independent, this metric is largely independent of species prevalence, while on the contrary, both KAPPA and TSS scores are dependent upon threshold-dependent measures of model accuracy. In this way, applying three different metrics to individual models allows for a more robust evaluation and ensembling of individual models.

The average evaluation metric scores across the weighted-mean ensembled models performed generally well (Table 6). The average AUC values across the maize landrace varieties is 0.939 ± 0.08 , with the highest AUC value (0.990) and the lowest score (0.835) derived from the ensemble models of the Apachito and Celaya maize

landrace, respectively. A ROC plot for all ensemble models together can be found in Figure 16.

True Skill Statistic (TSS) values were relatively high with a mean of 0.785 ± 0.216 . Similarly, the highest and lowest TSS scores derived from the Apachito (0.943) and Celaya (0.511) ensemble models. As expected, KAPPA values were relatively low (mean = 0.444 ± 0.294), indicating negative effects of low landrace prevalence on KAPPA scores.

Further investigations of the average specificity and sensitivity across the weighted-mean ensemble models for each maize landrace (Appendix A, Section A) reveal no patterns with regard to maize landrace. Indeed, no significant correlation exists between evaluation scores and species prevalence or niche size ($R^2 = 0.03$). However, it appears that in general these models are more specific (mean = 0.91) than sensitive (mean = 0.78), thereby regulating more conservative model forecasts.

Further evaluation statistics of ROC (AUC) weighted-mean ensemble can be found in Table 7, including AUC values, threshold cutoff with the minimum difference between sensitivity and specificity (SS) (SS Min Diff), the threshold cutoff with the maximum SS difference (SS Max), the minimum error rate (Min Error), and the threshold cutoff resulting in the minimum error (Min Error Cutoff).

Table 6 - Mean evaluation metric scores on 30% holdout data across 3 weighted-mean ensemble models. Each weighted-mean ensemble model performed similarly for each metric.

Landrace	KAPPA	AUC	TSS
Ancho	0.465	0.938	0.727
Apachito	0.524	0.990	0.943
Arrocillo Amarillo	0.559	0.964	0.799
Azul	0.547	0.987	0.924
Blando	0.446	0.989	0.936
Bofo	0.439	0.976	0.883
Bolita	0.485	0.843	0.551
Cacahuacintle	0.228	0.948	0.812
Celaya	0.394	0.835	0.511
Chalqueño	0.467	0.911	0.658
Chapalote	0.371	0.995	0.983
Chiquito	0.306	0.968	0.845
Comiteco	0.678	0.967	0.811
Conejo	0.352	0.943	0.781
Cónico Norteño	0.695	0.944	0.761
Cónico	0.719	0.963	0.813
Coscomatepec	0.598	0.984	0.880
Cristalino de Chihuahua	0.646	0.986	0.918
Cubano Amarillo	0.345	0.956	0.798
Dulce	0.155	0.909	0.687
Dulcillo del Noroeste	0.245	0.982	0.911
Dzit Bacal	0.480	0.973	0.840
Elotero de Sinaloa	0.381	0.971	0.863
Elotes Cónicos	0.367	0.896	0.656
Elotes Occidentales	0.349	0.893	0.655
Gordo	0.444	0.982	0.922
Jala	0.270	0.979	0.895
Mushito	0.470	0.919	0.668

Table 6 (cont.)- Mean evaluation metric scores on 30% holdout data across 3 weighted-mean ensemble models. Each weighted-mean ensemble model performed similarly for each metric.

Landrace	KAPPA	AUC	TSS
Olotillo	0.565	0.925	0.690
Olotón	0.568	0.961	0.828
Onaveño	0.483	0.965	0.847
Palomero Toluqueño	0.392	0.979	0.892
Pepitilla	0.536	0.933	0.721
Ratón	0.492	0.916	0.714
Reventador	0.281	0.959	0.847
Tablilla de Ocho	0.175	0.929	0.739
Tabloncillo Perla	0.565	0.938	0.746
Tabloncillo	0.377	0.962	0.826
Tehua	0.238	0.964	0.843
Tepecintle	0.416	0.940	0.758
Tuxpeño Norteño	0.543	0.886	0.604
Tuxpeño	0.550	0.935	0.745
Vandeno	0.347	0.893	0.652
Zamorano Amarillo	0.452	0.965	0.805
Zapalote Chico	0.802	0.842	0.685
Zapalote Grande	0.214	0.931	0.733
Mean	0.444	0.944	0.785
Standard Deviation	0.147	0.040	0.108

Table 7 – AUC values of weighted-mean ensemble by ROC (AUC) with the threshold cutoff with the minimum difference between sensitivity and specificity (SS) (SS Min Diff), the threshold cutoff with the maximum SS difference (SS Max), the minimum error rate (Min Error), and the threshold cutoff resulting in the minimum error (Min Error Cutoff).

Landrace	AUC	SS Min	SS_Max	Min. Error	Min. Error Cutoff
Ancho	0.94	0.25	0.22	0.03	1.00
Apachito	0.99	0.16	0.06	0.01	0.97
Ancho	0.94	0.25	0.22	0.03	1.00
Apachito	0.99	0.16	0.06	0.01	0.97
Arrocillo Amarillo	0.96	0.12	0.12	0.02	0.99
Azul	0.99	0.08	0.02	0.01	0.90
Blando	0.99	0.16	0.14	0.00	1.00
Bofo	0.98	0.26	0.20	0.00	1.00
Bolita	0.84	0.17	0.25	0.03	1.00
Cacahuacintle	0.95	0.34	0.28	0.01	0.98
Celaya	0.84	0.44	0.45	0.09	0.98
Chalqueño	0.91	0.32	0.39	0.06	0.94
Chapalote	0.95	0.03	0.33	0.00	Inf
Chiquito	0.97	0.31	0.31	0.00	Inf
Comiteco	0.96	0.03	0.01	0.02	0.91
Conejo	0.94	0.34	0.34	0.02	1.00
Cónico Norteño	0.96	0.12	0.12	0.06	0.45
Cónico	0.95	0.11	0.15	0.07	0.66
Coscomatepec	0.98	0.07	0.06	0.01	0.98
Cristalino de	0.99	0.04	0.02	0.02	0.87
Cubano Amarillo	0.93	0.43	0.57	0.01	Inf
Dulce	0.91	0.53	0.38	0.01	Inf
Dulcillo del	0.98	0.32	0.21	0.00	Inf
Dzit Bacal	0.96	0.06	0.26	0.01	0.94
Elotero de Sinaloa	0.95	0.20	0.19	0.01	1.00
Elotes Cónicos	0.90	0.39	0.25	0.07	0.97
Elotes Occidentales	0.89	0.42	0.48	0.05	0.97

Table 7 (cont.) – AUC values of weighted-mean ensemble by ROC (AUC) with the threshold cutoff with the minimum difference between sensitivity and specificity (SS) (SS Min Diff), the threshold cutoff with the maximum SS difference (SS Max), the minimum error rate (Min Error), and the threshold cutoff resulting in the minimum error (Min Error Cutoff).

Landrace	AUC	SS Min	SS_Max	Min. Error	Min. Error Cutoff
Gordo	0.98	0.38	0.16	0.01	0.98
Jala	0.98	0.55	0.56	0.00	Inf
Mushito	0.92	0.22	0.26	0.03	0.99
Olotillo	0.92	0.21	0.18	0.06	0.78
Olotón	0.96	0.13	0.15	0.01	0.93
Onaveño	0.97	0.17	0.13	0.01	0.97
Palomero Toluqueño	0.98	0.39	0.30	0.01	0.99
Pepitilla	0.93	0.25	0.36	0.03	0.96
Ratón	0.92	0.24	0.29	0.06	0.96
Reventador	0.94	0.26	0.19	0.01	Inf
Tablilla de Ocho	0.93	0.54	0.39	0.01	Inf
Tabloncillo Perla	0.92	0.09	0.06	0.02	1.00
Tabloncillo	0.94	0.19	0.13	0.05	0.93
Tehua	0.97	0.61	0.58	0.00	Inf
Tepecintle	0.93	0.23	0.16	0.04	0.98
Tuxpeño Norteño	0.94	0.19	0.26	0.03	0.96
Tuxpeño	0.88	0.36	0.36	0.16	0.69
Vandeno	0.89	0.41	0.41	0.04	1.00
Zamorano Amarillo	0.97	0.24	0.21	0.01	1.00
Zapalote Chico	0.84	0.05	0.26	0.00	0.62
Zapalote Grande	0.93	0.39	0.24	0.01	1.00
mean	0.94	0.25	0.24	0.03	0.93

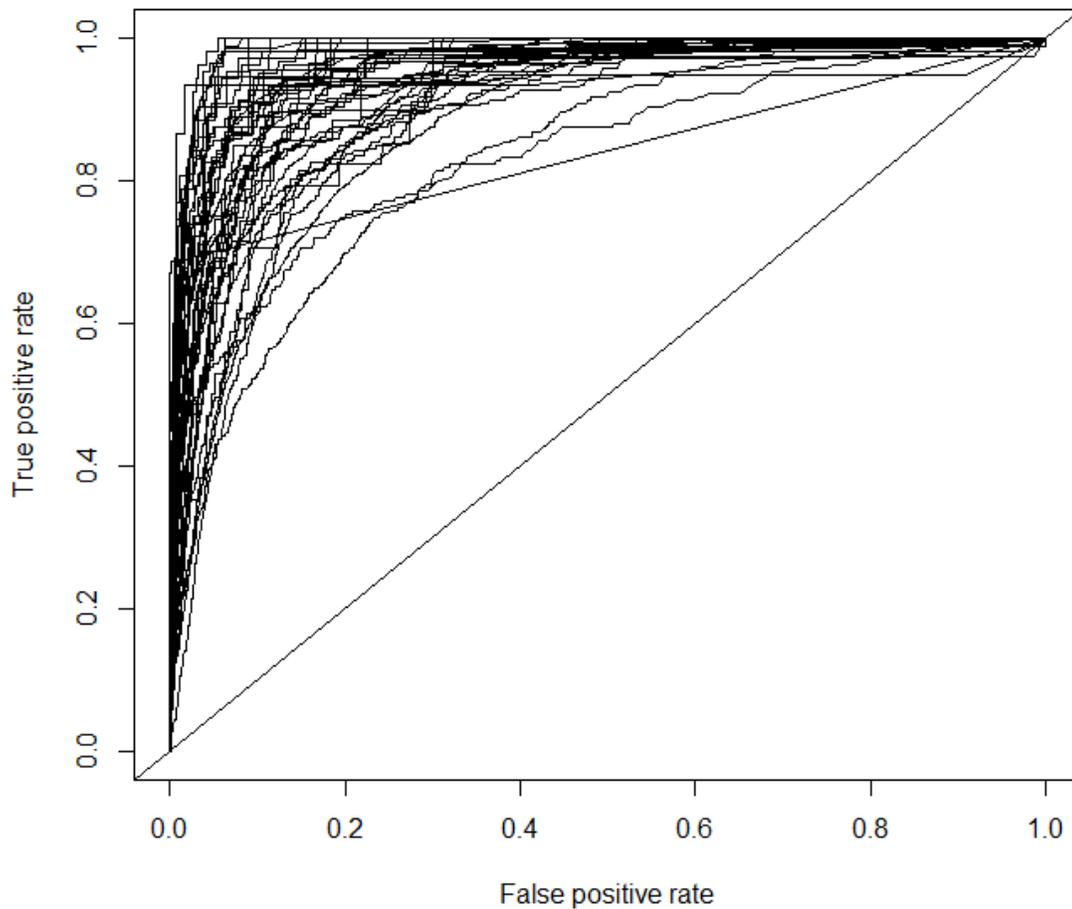


Figure 16 – ROC curves for the average weighted-mean ensemble predictions according to each maize landrace used to calculate AUC values.

Niche Overlap

Geographic Overlap (I similarity statistic)

Geographic niche overlap statistics were calculated using the *I* similarity statistic (Warren, Glor, and Turelli 2008) with ENMeval (Muscarella et al. 2014). The *I* statistic is similar to Schoener's *D* similarity statistic, but it overcomes unwarranted assumptions of prior species densities using a Hellinger distance-based method. It ranges in value from 0 (no overlap) to 1 (complete overlap). This statistic has been used to quantify evolutionary niche evolution for mammals in Mexico (Warren, Glor, and Turelli 2008), and has proven useful for niche equivalency and overlap testing, typically related to niche evolution research. Using the average weighted-mean ensemble model forecasts for the 1970 – 2000 time-period, *I* similarity statistics were calculated across each pair of species.

Ecological Overlap of Variable Importance

To evaluate niche overlap in ecological space, as well as to assess the effects of indigenous population percentage on maize landrace distributions, the 'variables_importance' function in the R package biomod2 (Thuiller et al. 2017) was applied to the weighted-mean ensemble models for each maize landrace. Similar to the random forest algorithm, this function estimates variable importance through use of correlative methods on artificially shuffled data sets.

While variable importance (scaled from 0 to 1) of indigenous population percentage was generally low across maize landraces (mean = 0.11 ± 0.08), a number of landraces exhibit high to very high influence of the indigenous population percentage

predictor variable. The maximum variable importance value for indigenous population percentage derived from the Zapalote Chico variety with a value of 0.52. Three other varieties – Apachito, Dzit Bacal, and Tehua – exhibited variable importance of indigenous population percentage greater than 0.15, with values of 0.24, 0.173, and 0.178, respectively.

Hierarchical cluster analysis using average ensemble model variable importance was calculated using Ward's method with a correlative distance measure in the *pvclust* package in R (Suzuki and Shimodaira 2015). This package provides two p-values, AU (Approximately Unbiased) and BP (Bootstrap Probability) to indicate strength of clusters identified in the data, with the AU statistic being the preferred method. Clusters with AU larger than 95% are highly supported by correlation matrix. The resultant dendrogram was cut into five groups as indicated in Fig. 17.

The correlation matrices derived from both the variable importance hierarchical clustering and the correlation matrix of variable importance distances were combined to create a square matrix using the 'psych' package in R (Revelle 2017). A correlation matrix is visualized for each maize landrace pair in this matrix (Figure 18) with help from the 'corrplot' package in R (Wei and Simko 2017). The rectangles in the figure delineate the groups detected in full correlation matrix via Ward's method. Though these clusters do not perfectly correspond to the clusters identified above using the cluster of variable importance distances alone, similar patterns exist between the two. Due to the relationship between ecological and geographic space in ecological niches, clusters derived from variable importance clusters were adopted for subsequent analysis.

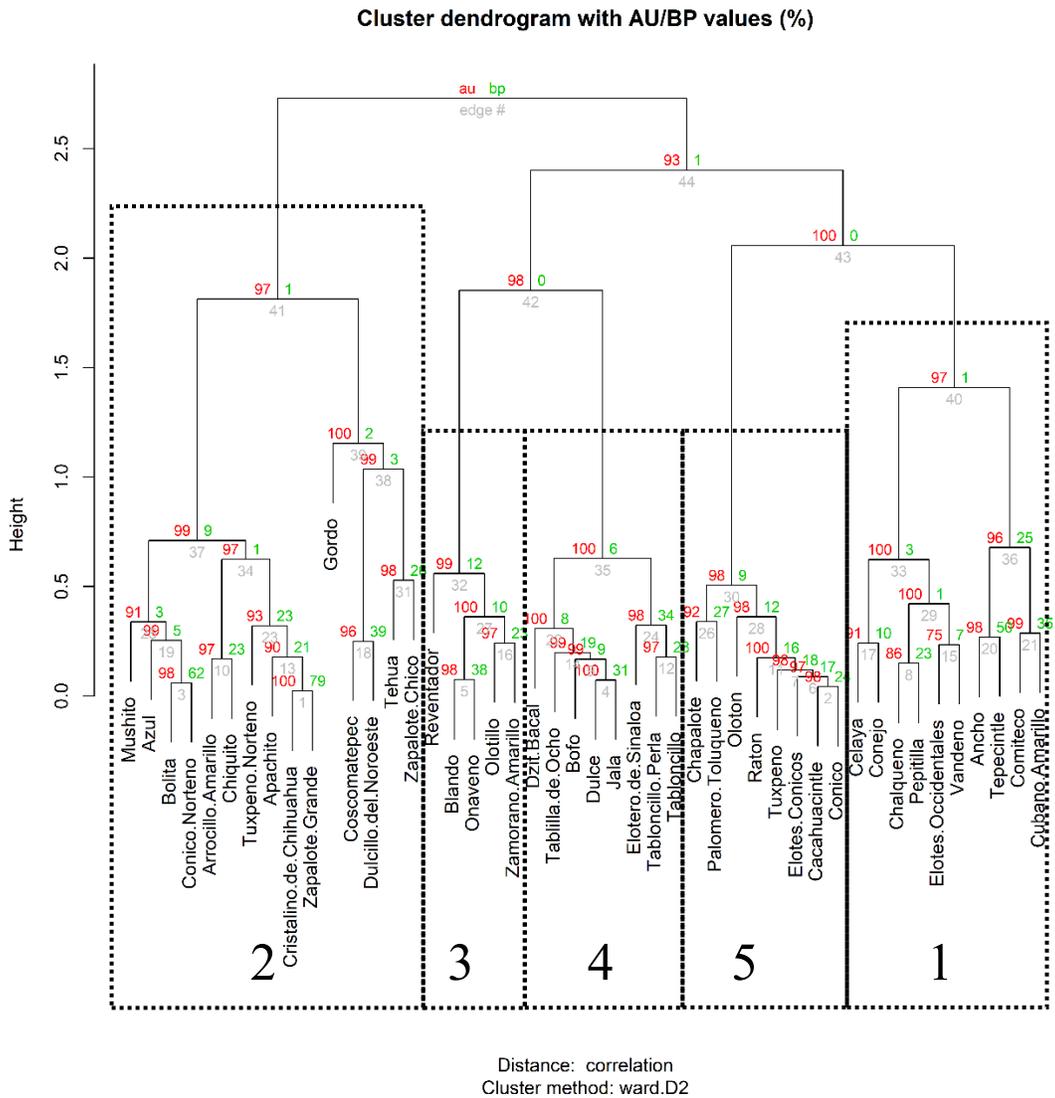


Figure 17 – Cluster dendrogram using the correlation matrix of distances of variable importance between each pair of maize landraces using Ward’s hierarchical clustering method with correlative distances. Approximately unbiased (AU) values are highlighted in red, with values > 95 indicating strong support by the data. Five groups identified from the cluster are enclosed in dashed boxes.

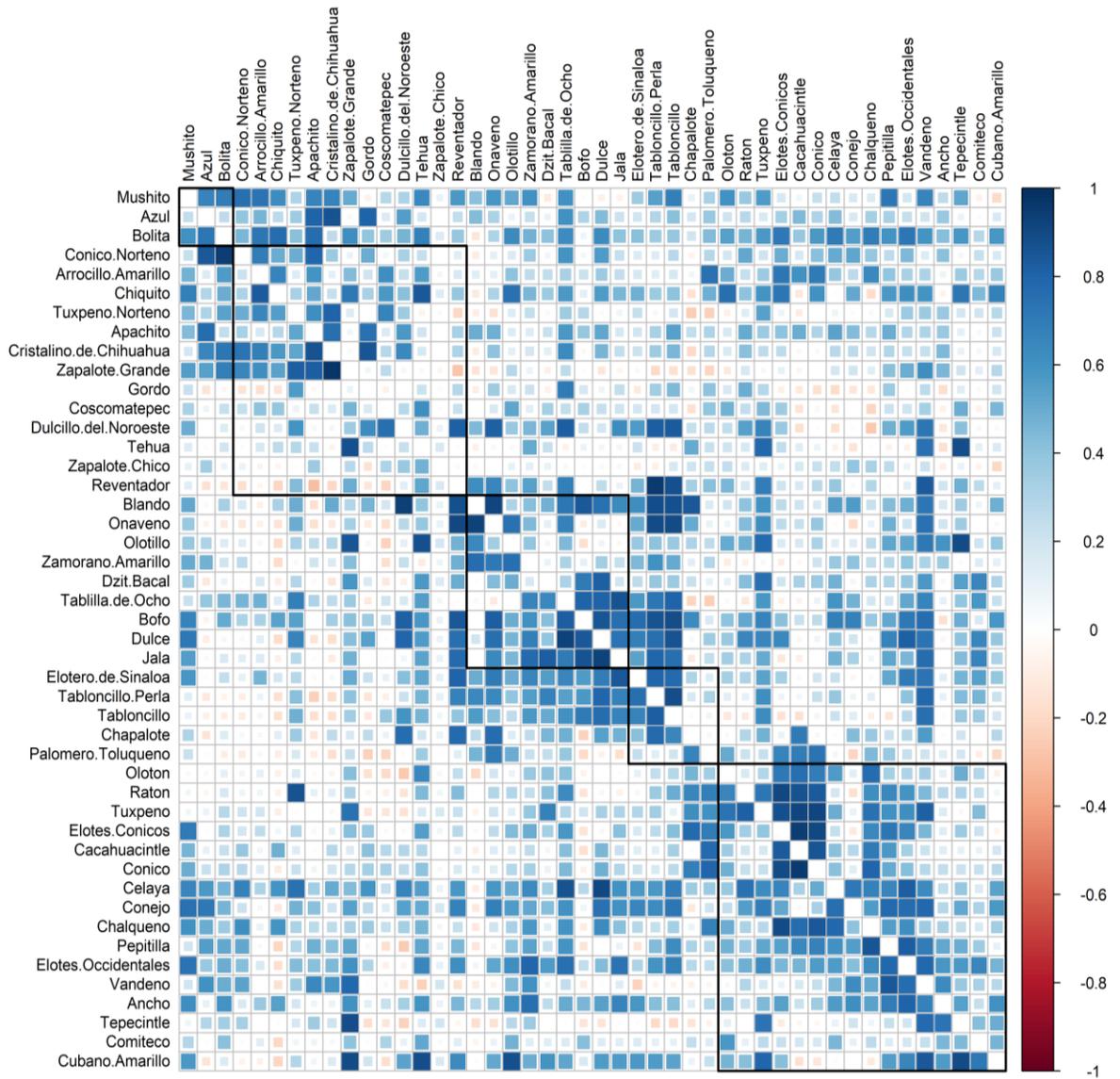


Figure 18 – Combined correlation matrix of I similarity statistic (upper triangle) and variable importance distances (lower triangle) with 5 different groups identified than when using Ward's hierarchical clustering method on variable importance alone.

Multi-variate Environmental Similarity Surfaces (MESS)

Due to the inherent stochasticity in forecasting to future periods or geographic regions, it is important to recognize the differences in predictor variables used between modeling and forecasting. Multivariate Environmental Similarity Surfaces (MESS) (Elith, Kearney, and Phillips 2010) provide an index to evaluate the multivariate, as well as univariate, differences between two different sets of environmental predictors. In doing so, MESS raster layers aid in visualizing not only the magnitude of potential, problematic extrapolation, but also the locations of such dissimilarity. The algorithm works similarly to the BIOCLIM niche modelling algorithm but provides the index with positive values indicating similarity and negative values indicating dissimilarity.

Because this research utilizes target-background sampling methods, MESS surfaces were calculated using the full training coordinates dataset (comprised of the combination of background samples and observation samples, i.e.: all observation coordinates for the full Presence/Absence Matrix)—for present (1970 – 2000) and future (2041 – 2060 & 2061 – 2080) time-periods—using both the full raster stack of predictors and only the stack’s bioclimatic variables using the ‘dismo’ package in R (Figs. 19 & 20) (A. R. J. Hijmans et al. 2017).

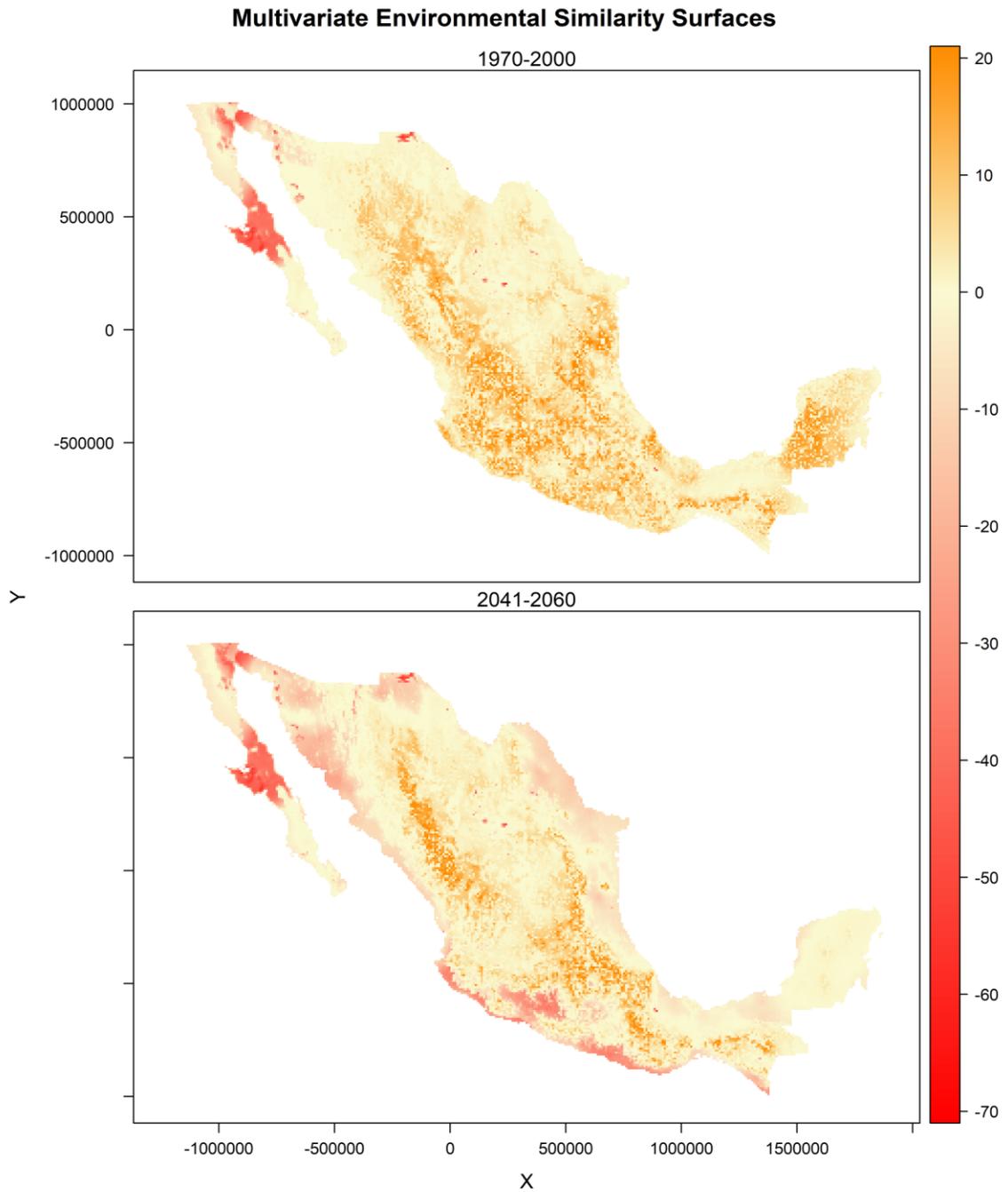


Figure 19 – Multivariate Environmental Similarity Surface (MESS) maps using all predictor variables for 1970 – 2000 (top) and 2041 – 2060 (bottom) time periods.

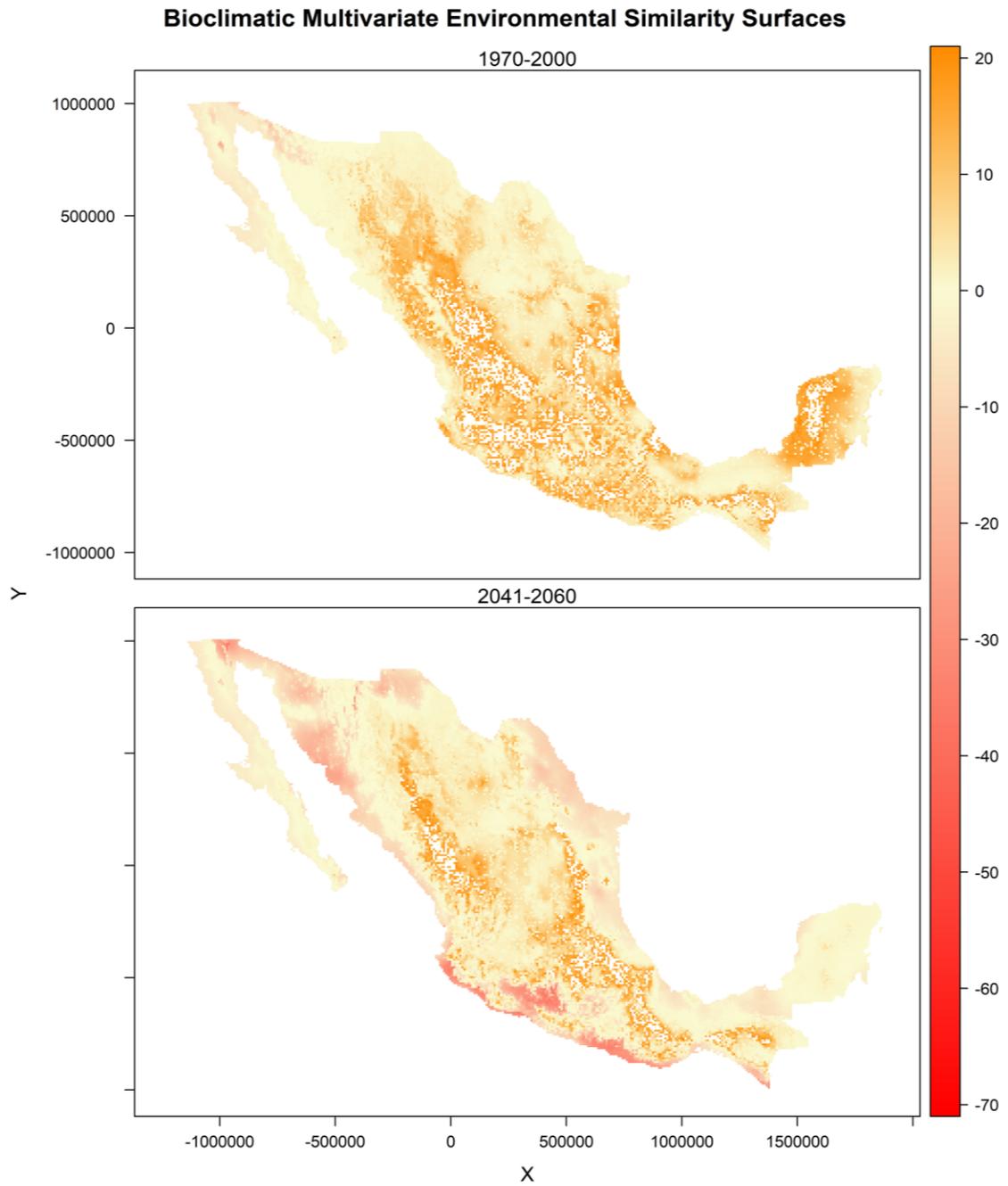


Figure 20 – Multivariate Environmental Similarity Surface (MESS) maps using only bioclimatic predictor variables for 1970 – 2000 (top) and 2041 – 2060 (bottom) time periods.

Both the full and bioclimatic-only MESS layers indicate consistent spatial patterns of dissimilarity. In particular, areas on the Baja Peninsula are significantly dissimilar for both current and future (2041 – 2060) time periods.

Aside from the Baja region, the target background sampling method appears to provide a good representation of 1970 – 2000 environmental space throughout the study area, with high levels of similarity at locations more heavily sampled in the CONABIO dataset. Even MESS dissimilarity values in the less sampled areas (e.g. mountainous regions in Northern Mexico) are not concerning. Further, where the current dataset did not have high similarity or dissimilarity values, the data is generally no different than the target-background reference data (values near zero).

The 2041 – 2060 MESS layer, on the other hand, indicates that various locations along Mexico's west coast have high multivariate dissimilarity, including coastal areas in the state of Sonora. This pattern appears especially pronounced within the Balsas River basin, in which maize is theorized to have evolved. Less pronounced, but nonetheless apparent, dissimilarity values also occur along the Gulf Coast, fanning out into a broader region the width of the southern Texas border.

Maize Landrace Niche Size

Using the binary forecasts for each maize landrace, niche size was calculated by summing the cells equal to 1 and converting to area (km²). Niche size is often a basic indicator of whether a species is a generalist or specialist and are listed in decreasing order in Table 8.

Table 8 – Ensemble forecasted 1970 – 2000 realized niche size (km²), ranked from most generalist to most specialist species. Landrace groups are exponentiated at the end of each landrace name.

	Landrace	km²		Landrace	km²
1	Ratón ⁵	420,312	24	Elotes Cónicos ⁵	44,395
2	Tuxpeño ⁵	347,557	25	Zapalote Grande ²	43,107
3	Tuxpeño Norteño ²	331,200	26	Cónico ⁵	41,115
4	Tabloncillo ⁴	196,877	27	Bolita ²	37,690
5	Onaveño ³	180,530	28	Cubano Amarillo ¹	37,313
6	Tablilla de Ocho ⁴	151,136	29	Mushito ²	33,788
7	Bofo ⁴	129,146	30	Chapalote ⁵	31,344
8	Conejo ¹	128,097	31	Jala ⁴	30,876
9	Pepitilla	124,314	32	Gordo ²	25,393
10	Reventador ³	119,148	33	Chiquito ²	23,522
11	Dulce ⁴	116,741	34	Olotón ⁵	22,314
12	Dulcillo del Noroeste ²	115,807	35	Ancho ¹	21,186
13	Cónico Norteño ²	115,151	36	Chalqueno ¹	19,223
14	Dzit Bacal ⁴	111,862	37	Zamorano Amarillo ³	18,803
15	Celaya ¹	108,707	38	Apachito ²	18,102
16	Blando ³	100,308	39	Cristalino de Chihuahua ²	16,352
17	Vanden ¹ o	95,969	40	Cacahuacintle ⁵	16,089
18	Elotes Occidentales ¹	88,640	41	Azul ²	15,688
19	Tehua ²	61,400	42	Zapalote Chico ²	13,580
20	Tepecintle ¹	57,677	43	Comiteco ¹	11,764
21	Elotero de Sinaloa ⁴	52,964	44	Palomero Toluqueño ⁵	8,566
22	Olotillo ³	44,794	45	Arrocillo Amarillo ²	6,268
23	Tabloncillo Perla ⁴	44,405	46	Coscomatepec ²	3,844

Geographic Distribution of Mexican Maize Diversity

In order to answer the first research questions, which ask “Where do current maize landraces exist and what are the spatial patterns of Mexico’s maize landrace diversity?”, this thesis calculated maize landrace diversity across Mexico for the 46 maize landraces successfully modeled.

Alpha diversity forecasts are calculated by summing the binary forecasts for each species for the 1970 – 2000 time-periods for all varieties together, as well as by the groups identified in hierarchical clustering of variable importance. Binary forecasts for each landrace were calculated by averaging the weighted mean ensemble forecasts, then classified all pixels $\geq .75$ as 1 and $< .75$ as 0 for consistency across landraces. Current distributions appear to conform with previous literature maize diversity in Mexico (Perales and Golicher 2014), but in a seemingly more conservative manner. Perales and Golicher (2014) show the highest levels of alpha diversity to be 11 landraces per pixel, while this research shows a maximum of 9 landraces per pixel, though this is most likely due to increased spatial resolution (30 arc second vs 5 arc minutes). Further, reviewing the specificity and sensitivity values of the ensemble models (Table 6 & Appendix Section A) reveals a general trend of increased specificity over sensitivity, which may also contribute to the discrepancy between maximum alpha diversity levels between this research and (Perales and Golicher 2014).

These binary forecasts show that the highest levels of alpha diversity occur in the northwestern coastal regions, in south Sonora and Sinaloa on the western side of the Sierra Madre Occidental mountain range (Fig. 21). Elevated levels of maize diversity also appear to exist in the Balsas River drainage basin on the border between Michoacán and

Guerrero, as well as in the Santiago River drainage basin in the states Nayarit and Jalisco, but in more compact areas.

Low levels of overall alpha diversity (Fig. 21) exist in the higher altitude plateaus between the Sierra Madre Occidental and Sierra Madre Oriental mountain ranges in the states of Zacatecas, Coahuila de Zaragoza, Tamaulipas, San Luis Potosí, Durango and Nuevo León. Similarly, low levels of diversity are forecast to occur at high elevations in the Sierra Nevada mountain range.

Mapping alpha diversity by the five variable importance clusters detected above illustrates that these ecological groups also tend to cluster in geographic space. These spatial clusters help inform not only unique biogeographic regions of Mexico, but perhaps also the genealogy of Mexican maize. Quantitatively, global Moran's I statistics for the mean diversity variable measured at the *municipio* level are high and statistically significant for all landrace groups extracted— 0.82, 0.79, 0.85, 0.85, and 0.78 for groups 1 through five, respectively—when using a spatial weights matrix (SWM) with 8 k-Nearest Neighbors. In other words, the distributions of the groups are all spatially clustered. The 8 k-Nearest Neighbors SWM was used to ensure that each *municipio* was linked to an adequate number of neighbors during the analyses. This is important for the present case, as several polygons from the dataset contain one or no neighbors when using more common connectivity-based SWMs such as queen contiguity.

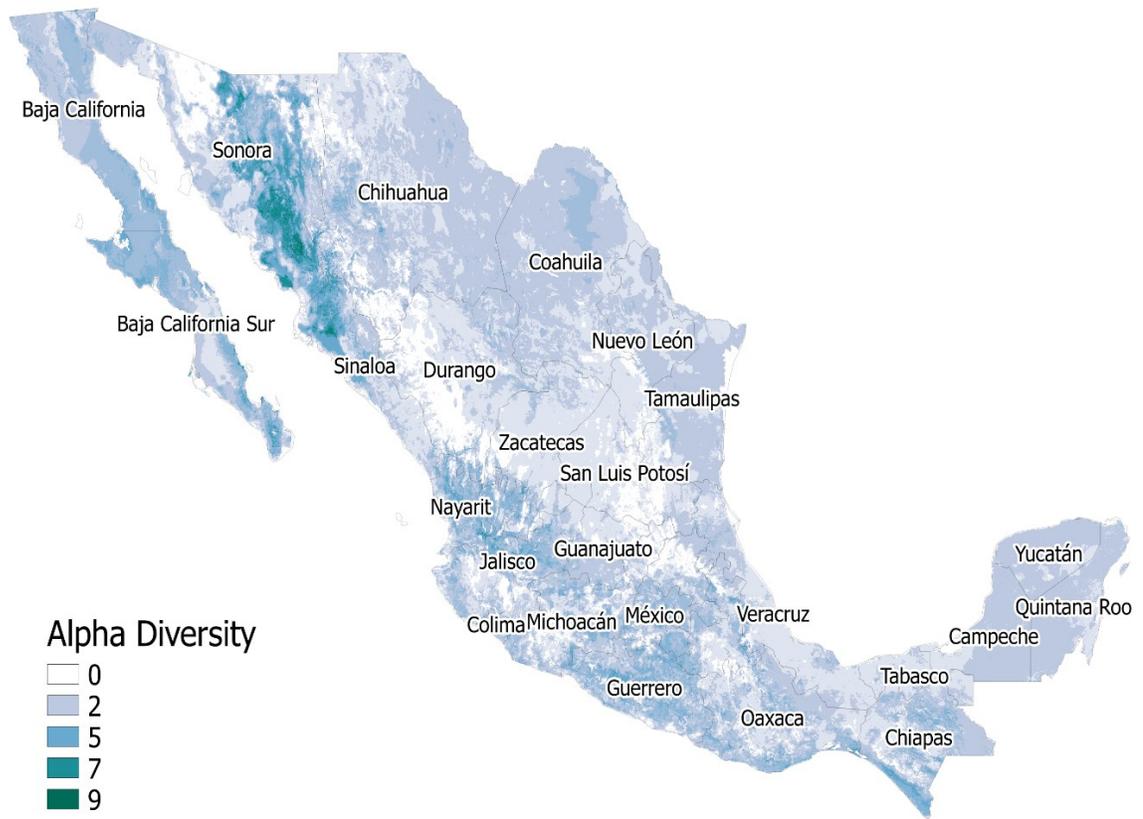


Figure 21 – Ensemble forecast of alpha diversity of all maize landraces for 1970 - 2000

Cluster Group 1 landraces exhibit the highest levels of diversity in areas encompassing the Balsas River Basin in Guerrero, Morelos, and Michoacán (Fig. 22). Group 1 also exhibits unique clusters of landraces slightly northward in Jalisco Guanajuato, as well as isolated clusters in Baja California, Oaxaca, and Chiapas/Tabasco. The highest diversity (3.56 – 4.58 landraces) by *municipio* for Group 1 occurs almost exclusively in Morelos. Relatively high pockets of diversity (1.63 – 3.56 landraces) also appear to exist along the Balsas River Basin, extending toward the Pacific Ocean, and north of the Sierra Madre del Sur mountain range in Jalisco, Michoacán, and Guanajuato.

On visual inspection, Group 2 landraces appear to exhibit lower spatial clustering in diversity, but the global Moran's I value (0.79) is still relatively high. This may be partly due to the fact that the k-nearest neighbors spatial weights matrix does not consider the size of each *municipio*, since Group 2 exhibits high diversity in Northern Mexico, which tend to have larger *municipio* tracts. As such, future research should conduct sensitivity analyses using alternative conceptualizations of spatial relationships. With that caveat in mind, in the present context, the highest diversity in Group 2 (2.6 – 3.7 landraces per km²) tends to occur almost exclusively in Oaxaca (Fig. 23). Further, a unique pattern of relatively high mean alpha diversity (1.5 – 2.6) appears to occur on the west portion of the beginnings of the Sierra Madre Oriental mountain range on the borders of Veracruz, Puebla, and Oaxaca. Similarly high diversity pockets also exist in northern Mexico in the state of Colima, between Chihuahua and Sonora.

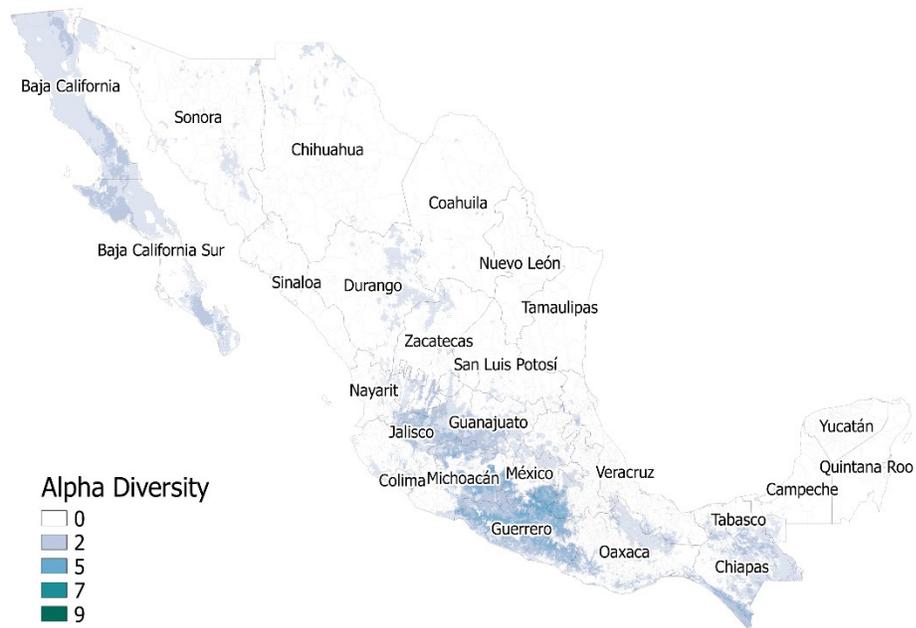


Figure 22 – Ensemble forecast of alpha diversity of Group 1 maize landraces for 1970 – 2000.

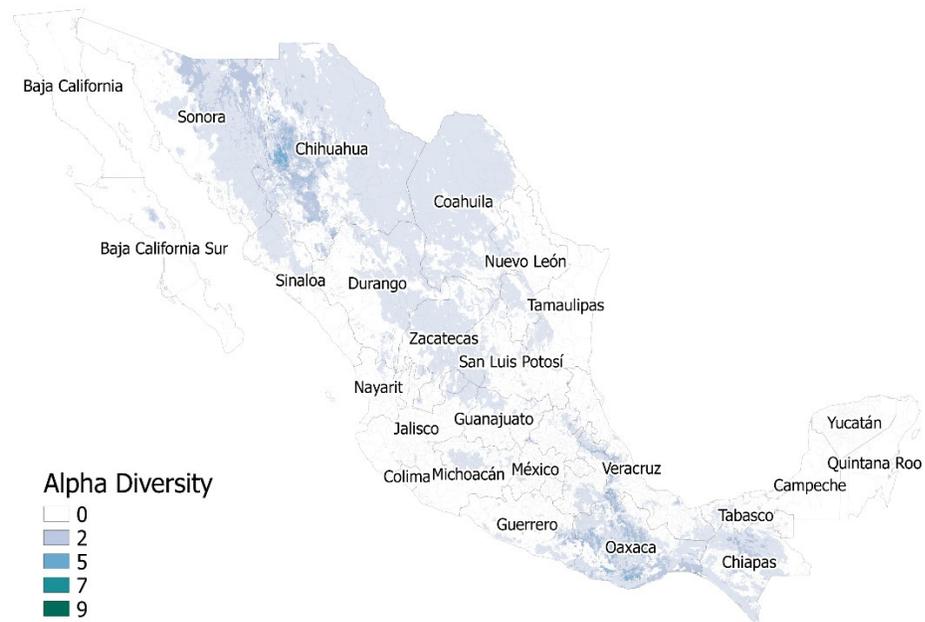


Figure 23 -- Ensemble forecast of alpha diversity of Group 2 maize landraces for 1970 – 2000.

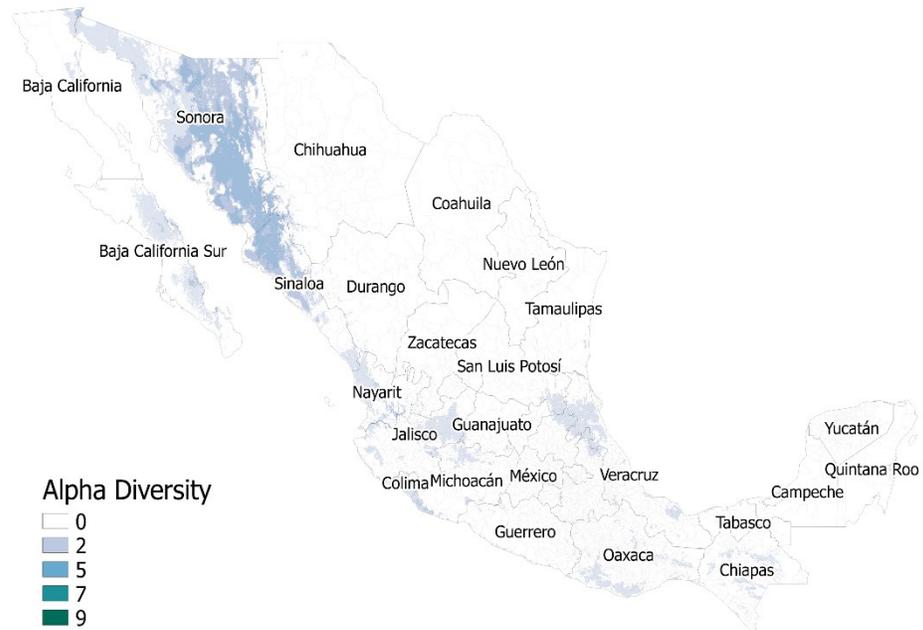


Figure 24 – Ensemble forecast of alpha diversity of Group 3 maize landraces for 1970 – 2000.

Group 3 exhibits one of the most spatially clustered distributions of mean diversity, with a global Moran’s I statistic of 0.85. Group 3 exhibits its highest levels of mean diversity (greater than 2.03) throughout the river basins of Sonora in northwestern Mexico (Fig. 24). Isolated pockets of high diversity (from 0.67 - 2.03) also appear to exist in distinct regions of Oaxaca, Nayarit, Jalisco/Guanajuato, Colima, and north Veracruz. The Group 3 landrace helps reaffirm the non-reciprocal relationship between environmental and geographic niches, proposing that distinct ecological niches may exist in disparate geographic locations, isolated by human, physiographic, and climatological conditions.

Spatial patterns of diversity for Group 4 are similar to Group 3: clustered along the west coast of Mexico and up into Sonora, but extending southward into Sinaloa, Nayarit, Jalisco, Colima, and Guanajuato (Fig. 25). Additionally, Group 4 shows relatively high diversity on the southern portion of the Baja Peninsula. Another unique band of Group 3 diversity appears to form on the eastern side of the Sierra Madre Occidental mountain range, east of Sonora, where Group 2 also exhibited anomalously high diversity. This pattern appears to get disrupted in Durango before connecting with higher diversity in the south. In contrast with Group 3, Group 4 also shows elevated diversity in the Yucatan Peninsula, in the states Campeche, Yucatan and Quintana Roo.

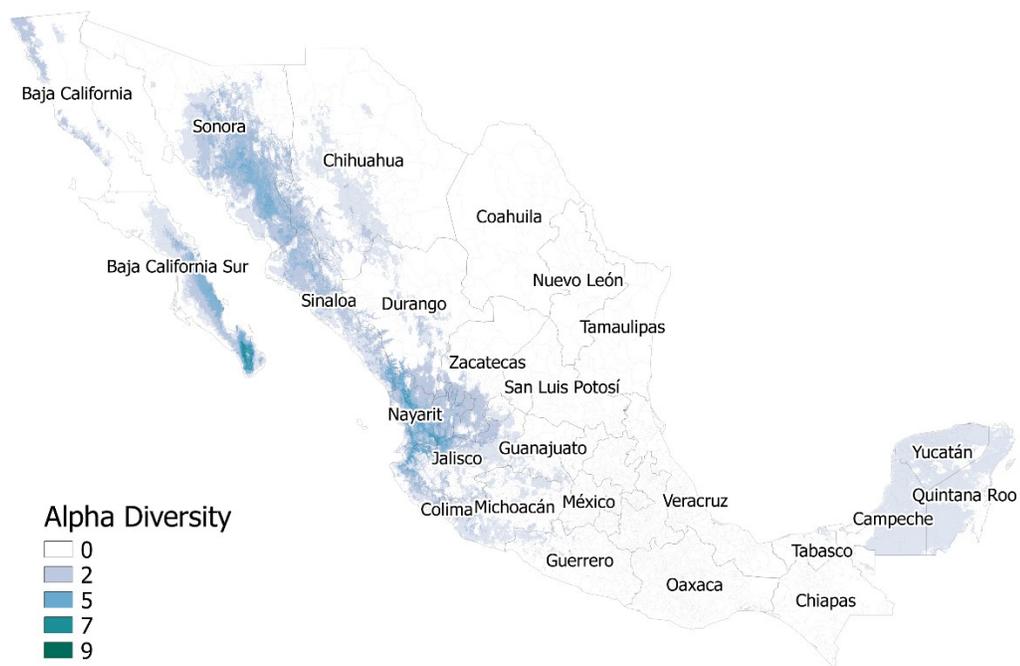


Figure 25 – Ensemble forecast of alpha diversity of Group 4 maize landraces for 1970 – 2000.

Unlike Group 4, ecological suitability for Group 5 appears to be highest along the eastern coasts of Mexico, with the greatest diversity levels in Tamaulipas and Mexico states (Fig. 26). The most diverse regions of Group 5 occur more inland toward the states of Mexico, Tlaxcala, and Puebla, just west of Veracruz. Interestingly, high levels of diversity for Group 3 also exist here, but not in Morelos, just south of this location, where Group 1 landraces are forecasted to have the highest levels of diversity—a further indication of the complex ecological dynamics that contribute to maize landrace distributions.



Figure 26 – Ensemble forecast of alpha diversity of Group 5 maize landraces for 1970 – 2000.

Influences of Indigeneity on Maize Landrace Diversity

Evaluating the influences of indigeneity on crop diversity is highly studied; however, the statistical associations of which are rarely found to be significant. Themes addressed in this thesis suggest that indigenous populations influence maize diversity at localized scales. To answer the research question “How does indigeneity covary with maize landrace distributions and diversity?”, this section analyzes differences of forecasted maize landrace diversity by differing levels of indigeneity by the *municipio* geographic level.

Calculating median alpha diversity by *municipio* indicates significant diversity differences between different types of *municipios*, categorized according to their levels of indigenous presence. The classes of *municipios* by indigenous presence were taken directly from CONABIO (2000), and include the following four categories, listed in ascending order from no indigenous presence to substantial indigenous presence: (1) Without indigenous population, (2) Dispersed indigenous population, (3) Population with indigenous presence, and (3) Indigenous populations. Diversity also exhibits *municipio*-level differences across levels of indigenous marginality, which CONABIO (2000) data classify as: (1) very low, (2) low, (3) medium, (4) high, and (5) very high. More specifically, due to the skewed distributions of the diversity variables and non-constant between-group variances, Kruskal-Wallis tests demonstrated that overall alpha diversity significantly differs between levels of indigenous population ($\chi^2 = 50.257$, $df = 3$, $p\text{-value} = 7.044e-11$) and marginality ($\chi^2 = 77.04$, $df = 4$, $p\text{-value} = 7.377e-16$). Post-hoc analysis to identify significant pairwise differences were conducted using the Games-Howell follow-up (Tables 9 and 10), which is a non-parametric analog to Tukey’s HSD.

	groups	Mean Difere nce	Stand ard Error	t	df	p	upp er limi t	low er limi t
1	Población con presencia indígena : Población indígena dispersa	-0.255	0.047	3.8 59	336.46 2	0.0 01	- 0.0 84	- 0.4 26
2	Población con presencia indígena : Sin población indígena	0.355	0.224	1.1 22	33.452	0.6 79	1.2 1	-0.5
3	Población con presencia indígena : Población indígena	-0.059	0.048	0.8 64	366.28 8	0.8 24	0.1 17	- 0.2 34
4	Población indígena dispersa : Sin población indígena	0.61	0.22	1.9 6	31.411	0.2 25	1.4 54	- 0.2 34
5	Población indígena dispersa : Población indígena	0.197	0.027	5.0 84	1576.5 11	0	0.2 96	0.0 97
6	Sin población indígena : Población indígena	-0.413	0.22	1.3 27	31.558	0.5 53	0.4 32	- 1.2 59

Table 9 - Games-Howell post-hoc analysis for Kruskal Wallis test indicating significant paired differences of median diversity per *municipio* (n = 2,511) by indigenous population presence group.

Indigenous Presence Pairs	Mean Difference	Standard Error	df	p	upper	lower
Population with Indigenous Presence : Dispersed Indigenous Population	-0.255	0.047	336.462	0.001	-0.084	-0.426
Population with Indigenous Presence : Without Indigenous Population	0.355	0.224	33.452	0.679	1.21	-0.5
Population with Indigenous Presence : Indigenous Population	-0.059	0.048	366.288	0.824	0.117	-0.234
Dispersed Indigenous Population : Without Indigenous Population	0.61	0.22	31.411	0.225	1.454	-0.234
Dispersed Indigenous Population : Indigenous Population	0.197	0.027	1576.511	0	0.296	0.097
Without Indigenous Population : Indigenous Population	-0.413	0.22	31.558	0.553	0.432	-1.259

Table 10 – Games-Howell post-hoc analysis for Kruskal-Wallis and Brown-Forsythe tests indicating differences of median diversity per *municipio* (n = 2,511) by indigenous marginality level.

Marginality Level	Mean Difference	df	p	upper limit	lower limit
Very Low : Low	-0.113	600.896	0.617	0.105	-0.331
Very Low : Moderate	-0.410	416.043	≤ 0.001	-0.225	-0.595
Very Low : High	-0.377	550.837	≤ 0.001	-0.170	-0.584
Very Low : Very High	-0.107	495.904	0.575	0.091	-0.304
Low : Moderate	-0.297	693.874	≤ 0.001	-0.132	-0.462
Low : High	-0.264	804.211	0.001	-0.074	-0.454
Low : Very High	0.006	774.333	1.000	0.185	-0.173
Moderate : High	0.033	792.549	0.974	0.184	-0.117
Moderate : Very High	0.303	975.247	≤ 0.001	0.440	0.167
High : Very High	0.270	833.301	≤ 0.001	0.435	0.104

With respect to indigenous presence, *municipios* classified as “Populations with Indigenous Population Presence” are flagged for having statistically significantly different median alpha diversity relative to *municipios* classified as “Dispersed Indigenous Population.” Overall maize alpha diversity, on average, is 0.255 higher in the former compared to the latter categories ($p \leq 0.001$). Further, a significant difference exists between *municipios* with “Dispersed Indigenous Populations” and *municipios* with “Indigenous Populations”, with overall alpha diversity 0.197 units lower, on average, in the former relative to the latter ($p \leq 0.001$).

In contrast, evaluating the differences of alpha-diversity by levels of indigenous produces more identifiable patterns. Explicitly, for all significant pairwise differences, the lower the marginality, the higher the alpha diversity, with very high levels of

indigeneity being the only exception. By extension, it appears that moderately and highly marginalized indigenous communities exhibit significantly low levels of diversity.

For Group 1, significant differences in mean alpha diversity were detected between *municipios* classified as “Indigenous Population” relative to those with “Dispersed Indigenous Populations,” increasing from the former to the latter by 0.120 landraces per *municipio* on average ($p = 0.002$). Areas characterized by “Dispersed Indigenous Populations” were further found to have diversity levels significantly different from *municipios* “Without Indigenous Populations,” where the former is linked to higher diversity of 0.303 on average ($p = 0.046$). With respect to marginality, diversity is highest in *municipios* with “Very High” marginality of indigenous populations (0.70 landraces per *municipio*). This group is significantly different from all other groups ($p < 0.02$), with a trend whereby alpha diversity is lowest in Highly and Moderately marginalized *municipios*.

In Group 2, diversity appears to be highest in *municipios* “With Indigenous Populations”, and the high levels of diversity in this group are significantly different from comparably lower levels of diversity in *municipios* “Without Indigenous Presence” ($p < 0.0001$). Mean diversity appears to increase with increasing levels of indigenous marginality for Group 2 landraces. High and Very High levels of marginality have highly significant ($p < 0.001$) elevated levels of Group 2 diversity when compared with all other levels of marginality.

Only marginal mean diversity differences exist in Group 3 between “Dispersed Indigenous Population” and “Indigenous Population” *municipios* (0.068 ± 0.011 , $p < 0.0001$), increasing by 0.068 ± 0.011 landraces per *municipio* from the former to the latter

group on average. Mean diversity increases relatively highly (0.531 ± 0.138) when going from *municipios* with Indigenous Presence to those Without Indigenous Populations ($p = 0.049$). Similarly, mean diversity is only negligibly different between *municipios* based on their indigenous marginality classification assigned by CONABIO, with High and Very Highly marginalized *municipios* exhibiting slightly higher levels of mean Group 3 alpha diversity ($+ 0.07 \pm 0.01$) than Moderately marginalized communities. Moderately marginalized communities also exhibit lower levels of diversity than in *municipios* with Very Low levels of marginalization, decreasing slightly by 0.085 ± 0.021 landraces per *municipio* ($p = 0.037$).

The only significant pairwise difference for mean diversity in Group 4 by indigenous presence exists between “Dispersed Indigenous Populations” and “Indigenous Populations”, with the former being higher than the latter ($p < 0.0001$) by 0.093 ± 0.015 landraces per *municipio* on average. Instead, Group 4 mean diversity exhibits more significant differences when comparing levels of indigenous marginality, which tends to decrease with increasing levels of marginality. Significant differences occur between Very Low and High (former is greater by 0.211 ± 0.029), Very Low and Very High (former is greater by 0.202 ± 0.029), Low and High (former is greater by 0.298 ± 0.027), Low and Very High (former is greater by 0.289 ± 0.028), Moderate and High (former is greater by 0.201 ± 0.017), and Moderate and Very High (former is greater by 0.192 ± 0.018) ($p < 0.0001$). In other words, unlike Groups 1 and 2, increasing indigenous marginality is linked with lower levels of alpha diversity in Group 4.

Group 5 exhibits nuanced effects of indigenous presence. Mean diversity for Group 5 landraces are highly significantly different between those *municipios* “With

Indigenous Presence” and all other categories. In this case, areas “With Indigenous Presence” tend to have higher diversity than all other *municipio* types, thereby supporting the notion that negative climate change impacts on diversity can have a disproportionate effect on indigenous populations in these spaces. In addition, like Group 4, Group 5 mean diversity is higher with lower levels of marginality. Significant differences occur between Very Low and Moderate (former is greater by 0.271 ± 0.041), Very Low and High (former is greater by 0.351 ± 0.042) ($p < 0.0001$), Very Low and Very High (former is greater by 0.368 ± 0.041) ($p < 0.0001$), Low and High (former is greater by 0.173 ± 0.034) ($p < 0.01$), Low and Very High (former is greater by 0.190 ± 0.033) ($p < 0.0001$), and Moderate and Very High (former is greater by 0.098 ± 0.023) ($p < 0.5$).

Geographic Clusters of Group 2 Diversity and Indigenous Presence

In order to answer the research question ‘Where might climate change, through changing maize’s distribution and diversity, have the greatest impacts on indigenous communities?’, this section presents results from a smaller-scope, zoomed-in case study of Group 2, as this group has the highest levels of diversity in *municipios* with high levels of indigenous presence and with highly marginalized indigenous populations. Indeed, a simple linear model predicting Group 2 mean diversity by *municipio* indigenous marginality level, indigenous presence level, and indigenous population percent yields a non-significant ($\alpha < 0.01$) R^2 value of 0.23. When incorporating an 8-kNN SWM into a spatial lag regression model, this R^2 increases to 0.70 (but with unequal variances).

To unpack the nature of the spatial relationship between diversity and indigeneity, Local Getis-Ord G_i^* statistics were calculated using Group 2 mean diversity with the maximum permutation number (99,999 iterations) using False Discovery Rate (FDA)

with significance level of 0.0001. As a first step, this analysis revealed that high Group 2 diversity clusters are situated in three states – Oaxaca, Veracruz, and Puebla – and in five primary areas, with two isolated *municipios* in Veracruz and Chiapas, consisting of a total of 121 *municipios* (Figures 27 & 28). Three-hundred forty-eight (348) *municipios* are classified as significant clusters with low-diversity for Group 2 maize landraces.

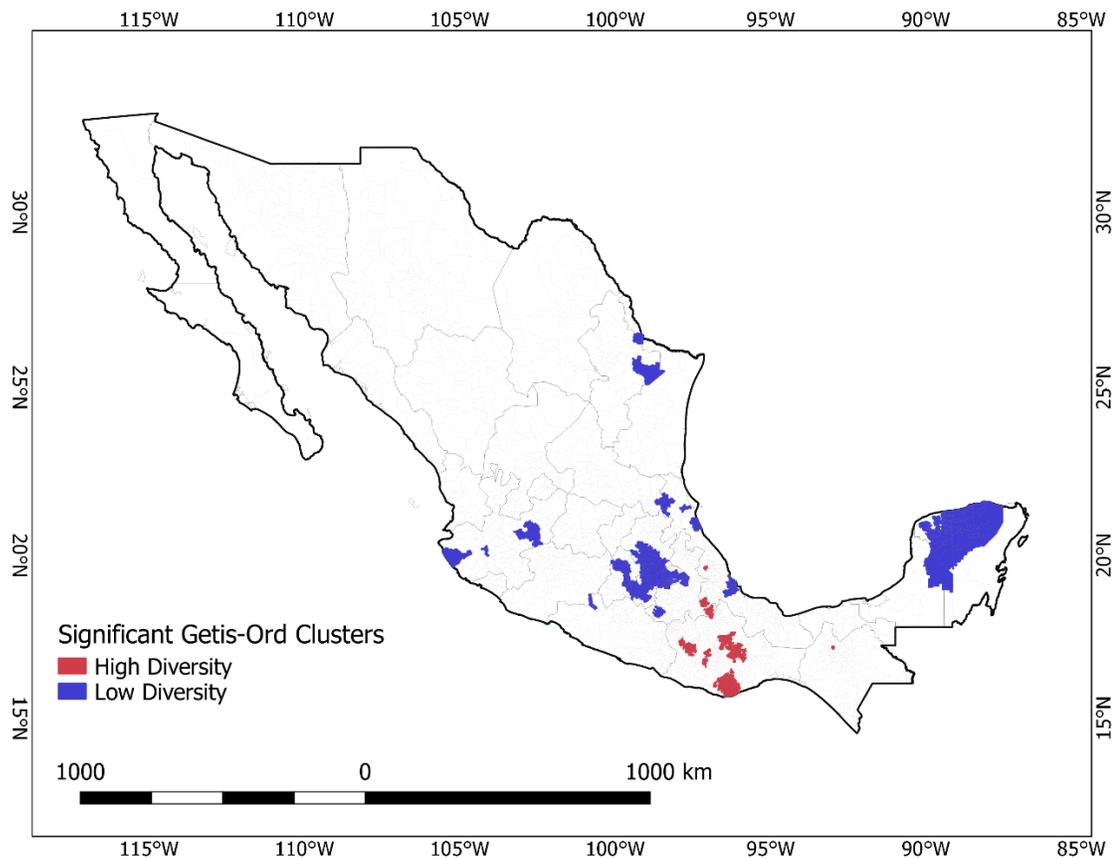


Figure 27 – Significant clusters of Group 2 mean diversity by *municipio* using Getis-Ord with 99,999 permutations using False Discovery Rate at $p = 0.001$.

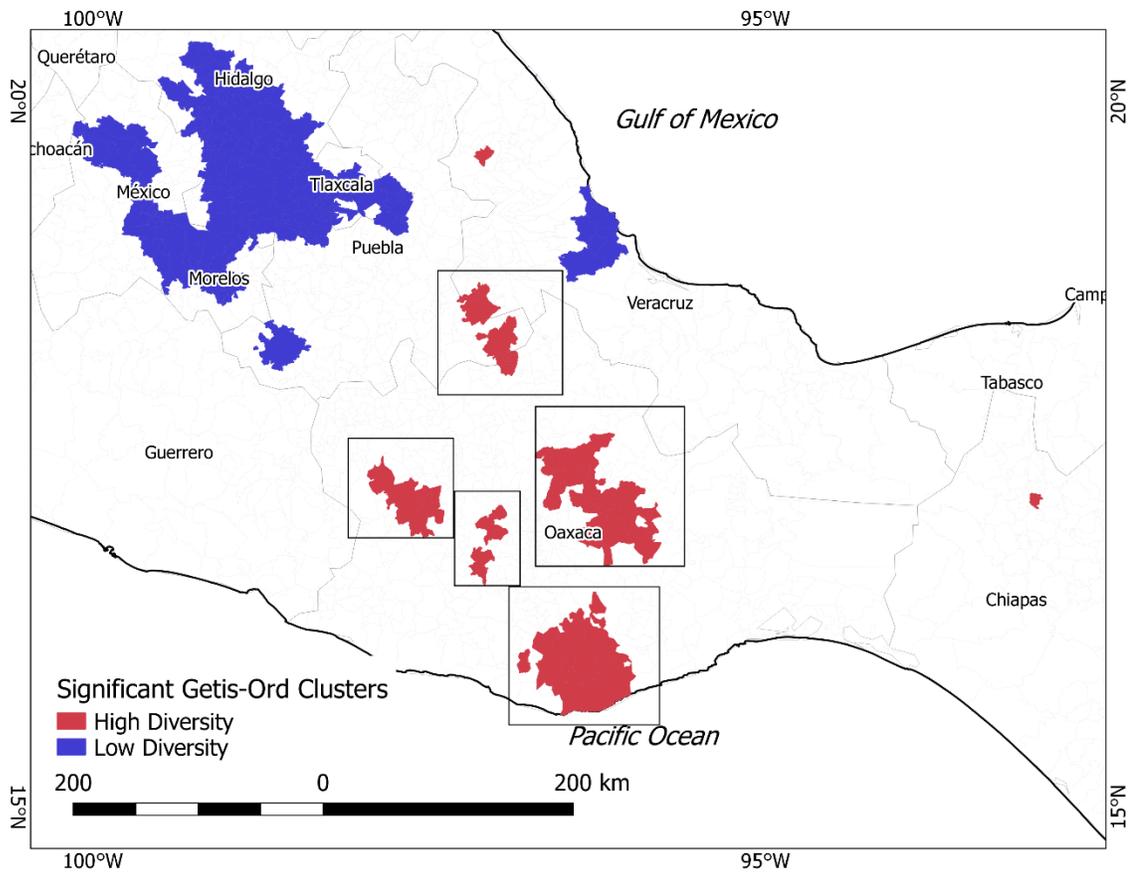
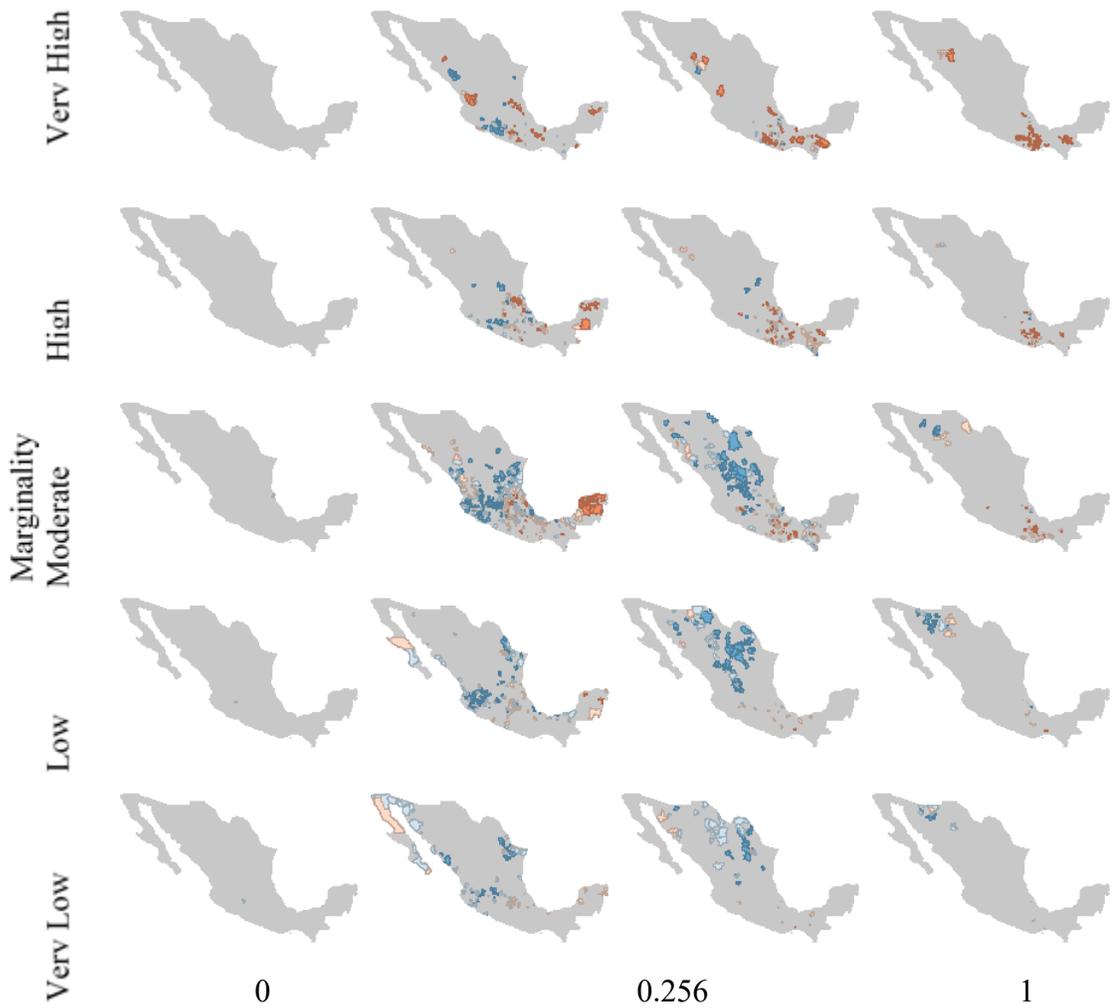


Figure 28 – Detail of figure 27, focusing on the five geographic clusters of significantly high Group 2 mean diversity by *municipio* using Getis-Ord with 99,999 permutations using False Discovery Rate at $p = 0.001$.

Next, mapping a conditional plot (Fig. 29) of *municipio* indigenous marginality by Group 2 mean diversity, colored by indigenous population percent with a hinge of 1.5, shows that the most highly marginalized *municipios* with high indigenous population percentages are located in the southern regions of Mexico. These regions are located primarily in Oaxaca and Chiapas, largely coincident with the highly significant clusters of high Group 2 alpha diversity. Other, more isolated *municipios* of highly marginalized, highly indigenous exist in portions of northwestern Mexico. Viewing the conditional plot further shows moderately marginalized *municipios* with high indigenous populations throughout the Yucatan Peninsula.

These findings help to support the hypothesis that the spatial distribution of maize diversity is contingent upon indigenous population percentage. The Getis-Ord statistic reinforces the understanding that maize diversity is spatially clustered in areas with high percentages of indigenous populations. In practical terms, these findings show that Group 2 maize diversity is significantly high in relation to national Group 2 maize diversity in specific *municipios* within southern Mexico, which likely group in areas with suitable bioclimatic, soil, or societal conditions for this landrace group. The conditional mapping of these results by indigenous marginality levels further shows that these *municipio* clusters of high Group 2 mean diversity also tend to be highly marginalized.



Group 2 Quantiles of *municipio* mean diversity

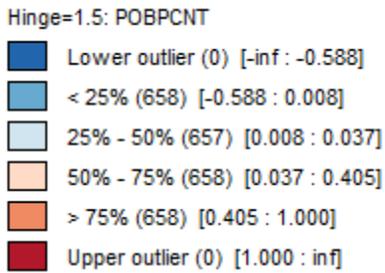


Figure 29– Conditional plot of Group 2 quantiles of mean diversity by levels of marginality, with *municipios* colored by indigenous population percent with a 1.5 hinge (POBPCNT).

Changes in Maize Landrace Diversity with Climate Change

To further engage with the research question ‘How is climate change expected to alter current maize landrace distributions and diversity?’, this section presents the forecasted changes of maize landrace alpha diversity in Mexico to the future time-periods 2041 – 2060 and 2061 – 2080.

Subtracting the forecast of maize diversity for future time periods from the current period allows for mapping of projected alpha diversity change in response to climate change. Evaluating the distributional changes of maize landrace diversity to the future time-periods for RCP 8.5 reveals that maize diversity is forecasted to change in a

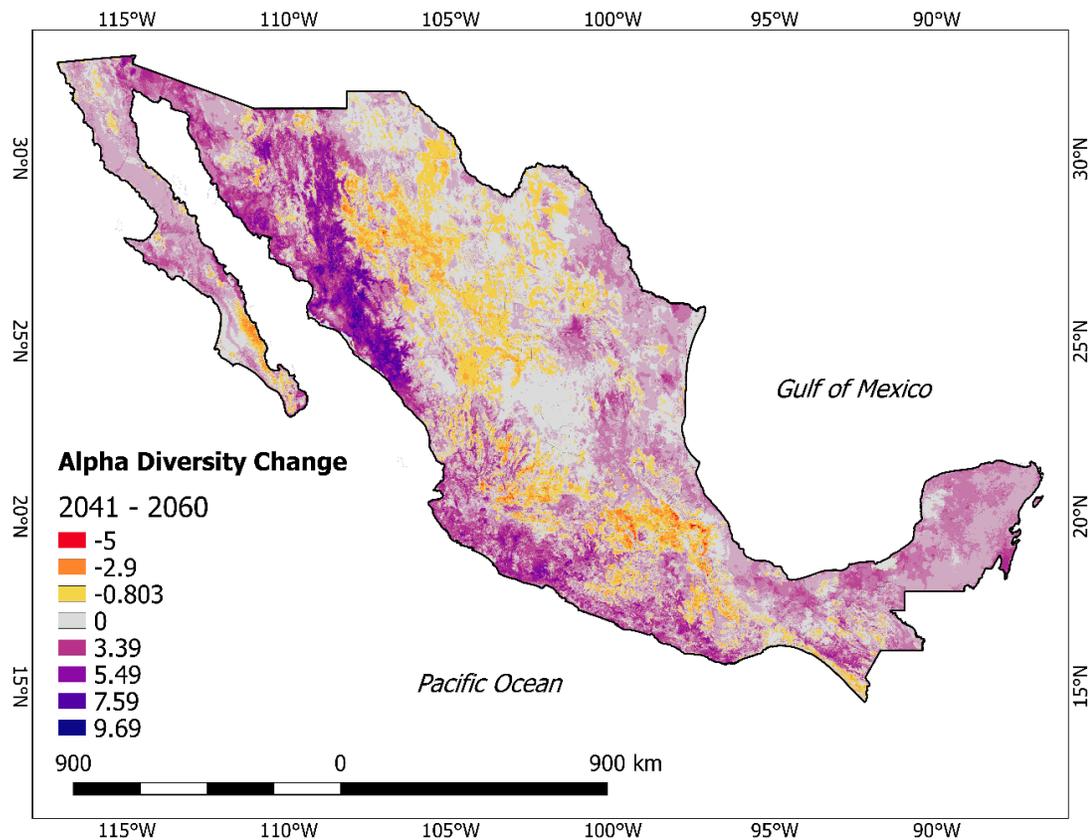


Figure 30 – Overall maize diversity change, calculated by subtracting forecasted 2041 – 2060 alpha diversity from 1970 – 2000 alpha diversity.

non-uniform manner.

Overall, maize landrace diversity change is similar between both 2041 – 2060 and 2061 – 2080 time-periods (Figs. 30 and 31). In general, these forecasts suggest that overall maize landrace diversity may *increase* throughout Mexico, but primarily in coastal regions, with the highest increases in forecasted diversity primarily in coastal regions, with the highest increases in forecasted diversity occurring in the states of Sonora and Sinaloa. Other areas of high increases in diversity occur along the southwestern coastal region of Mexico, extending from the Balsas River basin northwestward. It is worth noting that these areas with projected increases in diversity

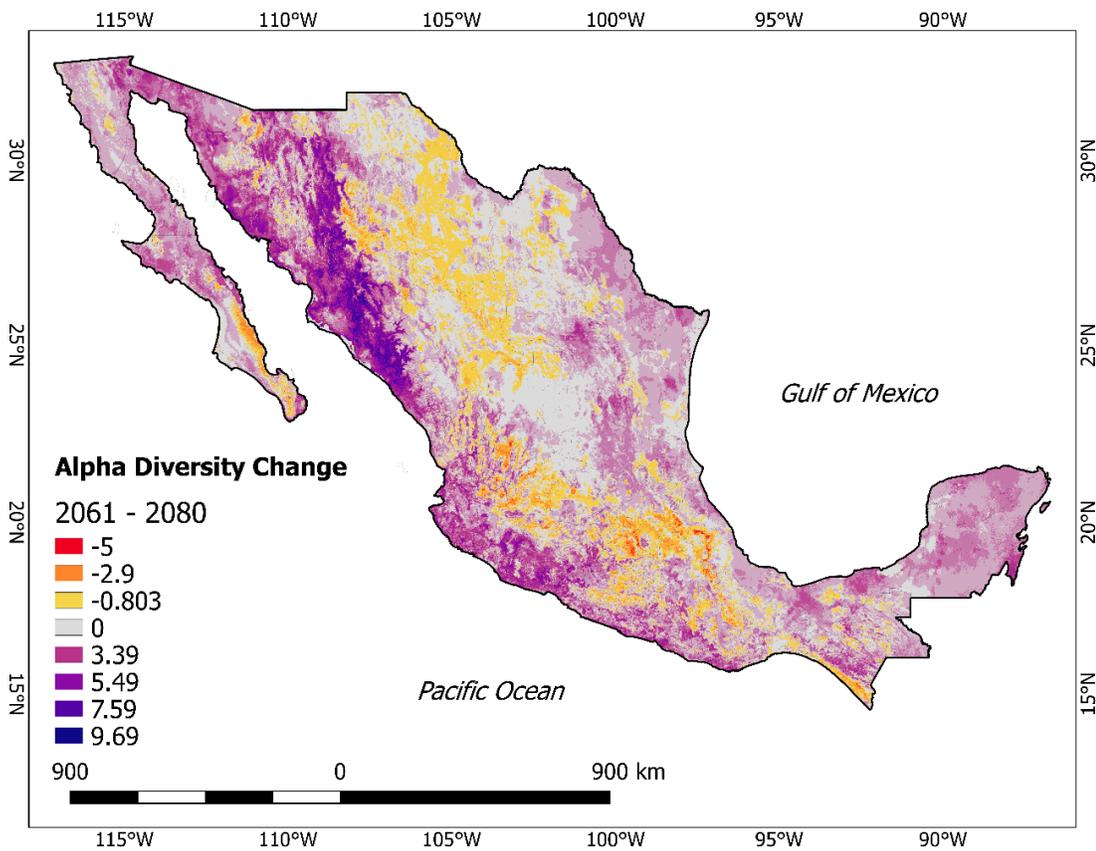


Figure 31 – Overall maize diversity change, calculated by subtracting forecasted 2061 – 2080 alpha diversity from 1970 – 2000 alpha diversity.

along the west coast of Mexico also coincide largely with low MESS values, indicating that extrapolation to these areas occurred and should thus be interpreted carefully. Maize diversity decreases are forecasted in primarily interior regions of Mexico, with the highest decreases in maize diversity occurring in areas through Central Mexico.

These forecasted diversity changes don't appear to conform strictly with other research (Ureta, Martínez-Meyer, Hugo R Perales, et al. 2012) that report negative changes in maize landrace diversity across Mexico; however, this thesis does support the findings (Ureta, Martínez-Meyer, Hugo R Perales, et al. 2012) that show that several landraces show new potential distribution areas in the future. Although the findings here do not follow previous literature, the potential reasons for these differences are discussed later in 'Limitations of Proposed Data and Methods.'

With that caveat in mind, further evaluating the preceding representation of maize landrace diversity change by landrace groups can help explicate the spatial patterns of overall maize landrace diversity change. Maps of maize landrace diversity change are depicted for only the 2041 – 2060 time-periods because the projections between the two across landrace groups are largely similar.

Group 1 maize landrace diversity is projected to be largely unchanged or increase by 2041 – 2060 (Fig. 32), with exceptions of maize diversity decreases in particular pockets through central Mexico. Maize landrace diversity within Group 1 appear to increase in lower-elevation areas, particularly along Pacific coastal regions. In particular, maize landrace diversity through the western of Balsas River basin are expected to increase or remain unchanged, while areas to the southeast of the Balsas River basin indicate spatially-diffuse areas of decreased of maize diversity, interspersed with areas projected to have no change. In general maize diversity appears to increase in low-elevation river basins, with decreases occurring in varied locations with no clearly-

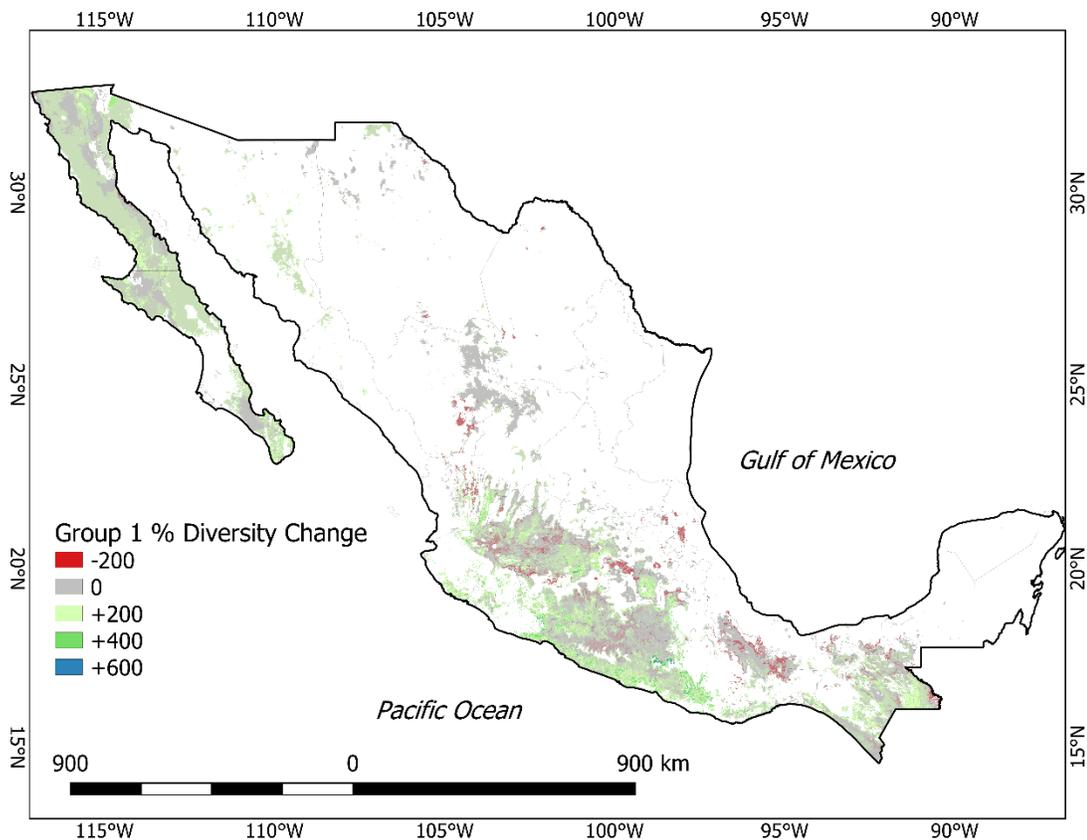


Figure 32 – Group 1 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.

emerging patterns.

Upon visual inspection, Group 2 maize landrace diversity appears to be on course for the most drastic decreases in diversity (Fig. 33), due to the forecasted *local* extinctions in maize landraces across large swaths of Mexico, especially in the northern to north-central regions. A handful of exceptions exist, including in areas in Sonora where increases in diversity are forecasted along lower-elevation river basins. In the southern portion of Mexico, maize landrace diversity change is forecasted to be largely negative, particularly in the eastern side of the Sierra Oriental mountain range in

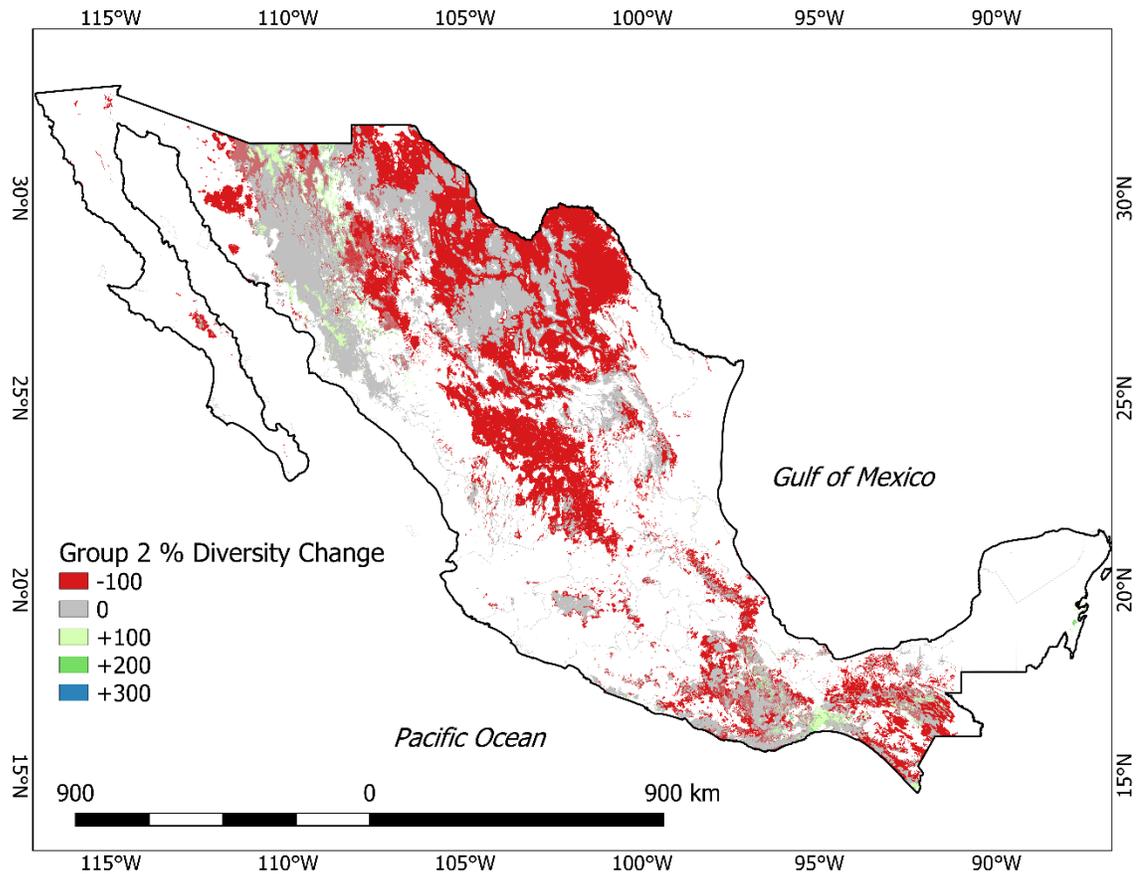


Figure 33– Group 2 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.

Veracruz. Small, speckled portions of the maize diversity forecast maps appear to have increase maize diversity forecasted, with no obvious geographic relationship when overlain an altitude DEM. A significant cluster of positive maize change is forecasted for Group maize landraces in Lake Superior, near the southern border between Chiapas and Oaxaca.

Examining the future forecasts of Group 3 maize landrace diversity changes (Fig. 34) shows fairly clear geographic patterns. Along the east coast, maize diversity is forecasted to largely decrease, with portions forecasted to remain unchanged in the lower

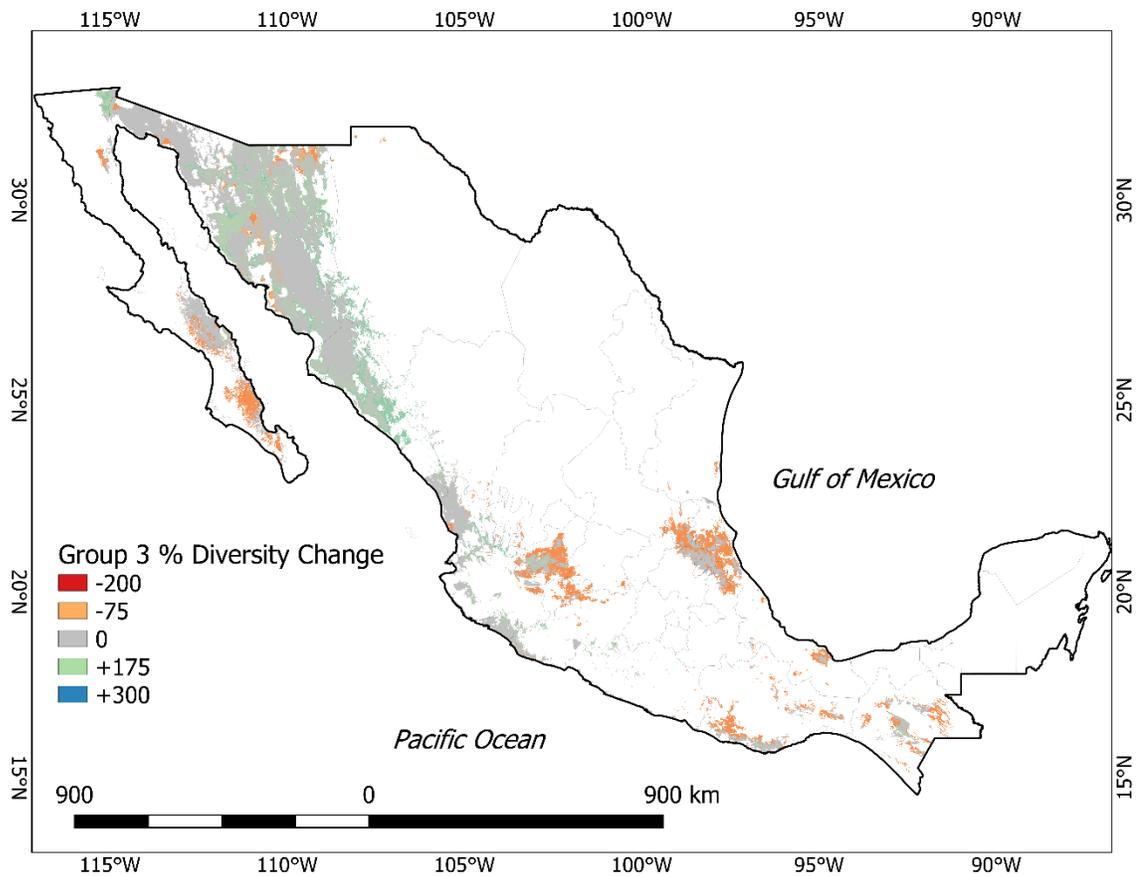


Figure 34 – Group 3 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.

elevation river basins along the western side of the Sierra Oriental mountain range. In the south, diversity is forecasted only to increase or remain unchanged along higher elevation portions (in Chiapas) or along coastal regions (in Oaxaca). Overall, maize landrace diversity is forecasted to decrease across central and southern Mexico, while the largest increases in Group 3 maize landrace diversity are forecasted along the west coast in Sonora and Sinaloa.

Similar to Group 3 diversity change, Group 4 maize landrace diversity is forecasted to primarily increase along the western river basins (Fig. 35); however, in contrast, Group 4 landrace diversity is forecasted to decrease substantially in Sonora and Southern Baja California. Along the eastern portion of the Sierra Occidental mountain

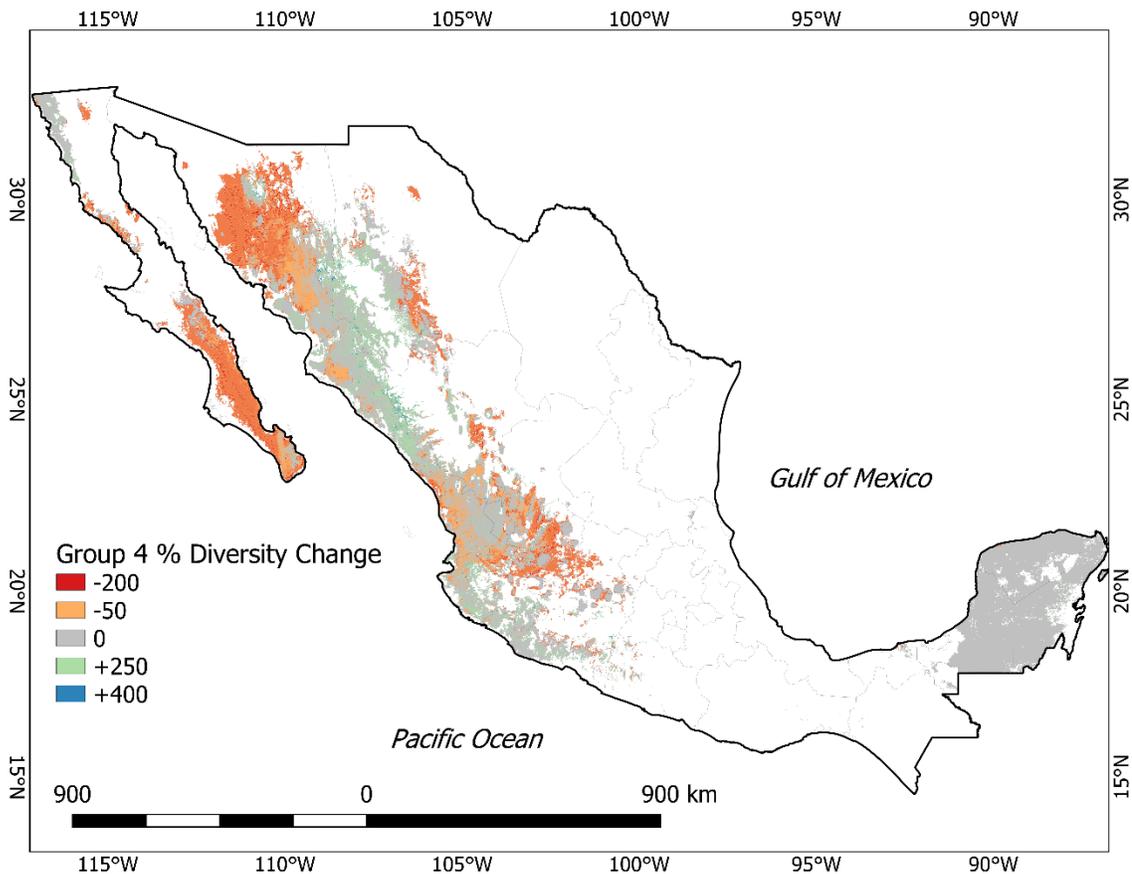


Figure 35 – Group 4 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.

range into Chihuahua, maize landrace diversity seems to decrease with decreasing elevation. Landrace diversity is further forecasted to decrease substantially along the west-central portions of Mexico. Forecasted increases in diversity appear to coincide with lower-elevation river basins along the western coast, until Sonora. In the Yucatán peninsula, maize landrace diversity is forecasted to remain largely unchanged.

For Group 5 maize landraces, diversity is forecasted to increase substantially along pockets of the western coast, through Veracruz and Tabasco, with the majority of the eastern and northeastern portions to have unchanged levels of diversity (Fig. 36). Through central Mexico, maize diversity is forecasted to primarily decrease, with small

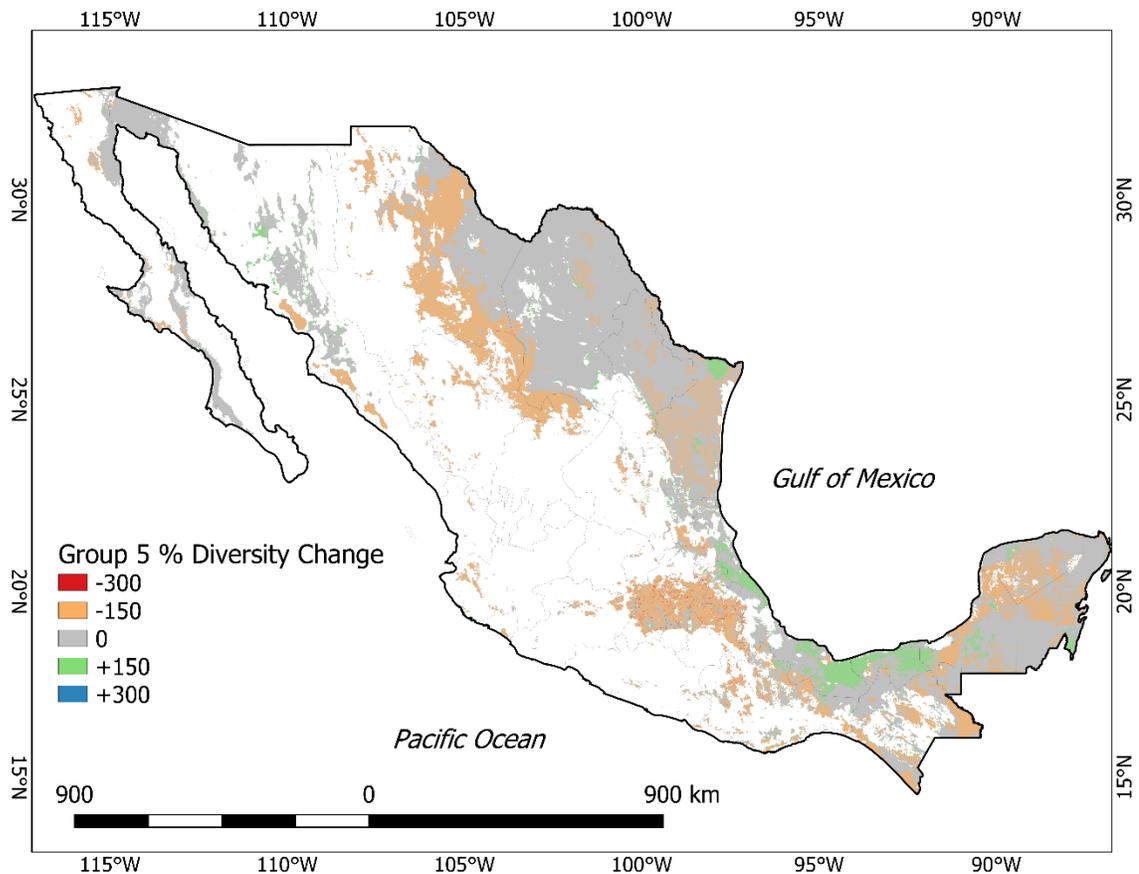


Figure 36 – Group 5 alpha diversity change (%) to the 2041 – 2060 time-period compared with 1970 – 2000 forecasted diversity levels.

pockets of forecasted increases at high elevations. Similar patterns emerge in north central Mexico, with diversity forecasted to decrease at lower elevations. Other pockets of forecasted increases exist through the Yucatan states of Quintana Roo, Campeche, and Yucatan, with no clear geographic reason.

Utilizing the 'Range_Size' function in biomod2, changes in niche size were calculated for each maize landrace for the 2041 – 2060 time-period (Table 11). This function calculates range change size by subtracting percent loss from percent gain. Percent loss is calculated by dividing the number of pixels to be lost by species by the number of pixels lost by species plus the number of pixels retained by species. Percent gain is similarly calculated as the number of pixels predicted to be gained divided by the number of pixels to be lost by species plus number of pixels to be retain by species.

The median landrace size change across all landraces is -14.49%, while median group landrace diversity percent change values are +66.9%, -77.9%, +74.54%, +3.1%, and -64.5% for groups 1 thru 5, respectively. As shown, only Groups 2 and 5 median maize landrace diversity change is forecasted to decrease, with all other Groups forecasted to have increased diversity levels by median-inference. Though, overall, maize landrace suitability is forecast to decrease.

While these statistics are alarming, these range size calculations also assume completely unbridled migration amongst all landraces. That is, the 'Percent Gain' portion of the calculation assumes that migration of maize landraces is wholly unhindered. Instead, separating the Percent Loss and Percent Gain portions of this equation (Table 12) allows a better look at the distributional changes of maize landraces. Indeed, the average percent loss across all maize landraces is 54.98%, indicating that more than 50% of all

currently suitable area for maize nationally will be lost, while 34% will be gained, on average. The median percent loss values by maize landrace group are 31.7%, 78.6%, 10.5%, 39.9%, and 74.5% for maize landrace groups 1 thru 5, respectively. Similarly, median percent gain values by maize landrace group are 86.5%, 4.8%, 85.0%, 43.0%, and 9.36% for maize landrace groups 1 thru 5, respectively.

These statistics should have greater prominence in interpretation because the Percent Loss alone better indicates levels of *local* extinction, while Percent Gain alone explains unbridled migration into novel locations. Given that maize landrace diversity is contingent upon indigeneity in Mexico, assuming uninhibited migration is problematic. For instance, local communities may perform specific agricultural techniques for specific landraces not known or practiced elsewhere, so while a novel area may be bioclimatic suitable for a specific landrace, the local food networks and agricultural practice may prohibit biologically-suitable areas.

This facet is of particular concern when considering Group 2 maize landraces distributional change, whose diversity appears highly contingent upon marginalized indigenous communities. Group 2 shows a forecasted decrease in suitable area by greater than three-quarters, with only a 4.8% increase in novel, suitable area, assuming full migration. Within Group 2 maize landraces, roughly half of the landraces have a forecasted Percent Loss >90%, with Gordo, Azul, and Arrocillo Amarillo landraces forecasted to also have 0% Percent Gain. Further, 60% of Group 2 maize landraces show a forecasted Percent Gain less than 7%. While the use of ecological niche modeling to quantify species extinctions is contentious, these results are nonetheless alarming.

Table 11 – Forecasted species range percent change by maize landrace for 2041 – 2060. Landrace groups are exponentiated at end of landrace name.

Landrace	% Change	Landrace	% Change
Ancho ¹	4,945.256	Elotes Cónicos ⁵	-67.152
Apachito ²	-93.614	Elotes Occidentales ¹	-3.651
Arrocillo Amarillo ²	-96.554	Gordo ²	-97.503
Azul ²	-100	Jala ⁴	-76.687
Blando ³	94.785	Mushito ²	53.137
Bofo ⁴	11.687	Olotillo ³	-52.052
Bolita ²	126.511	Olotón ⁵	-61.759
Cacahuacintle ⁵	-80.229	Onaveño ³	74.538
Celaya ¹	73.011	Palomero Toluqueño ⁵	-99.626
Chalqueno ¹	-30.552	Pepitilla ¹	39.924
Chapalote ⁵	273.089	Ratón ⁵	12.64
Chiquito ²	-77.863	Reventador ³	151.354
Comiteco ¹	-23.478	Tablilla de Ocho ⁴	-48.552
Conejo ¹	60.786	Tabloncillo Perla ⁴	229.197
Cónico Norteño ²	-98.81	Tabloncillo ⁴	-5.508
Cónico ⁵	-91.113	Tehua ²	-52.107
Coscomatepec ²	-93.809	Tepecintle ¹	88.585
Cristalino de Chihuahua ²	-98.905	Tuxpeño Norteño ²	-31.618
Cubano Amarillo ¹	369.134	Tuxpeño ⁵	-39.486
Dulce ⁴	-47.83	Vandeno ¹	269.085
Dulcillo del Noroeste ²	41.83	Zamorano Amarillo ³	-80.524
Dzit Bacal ⁴	46.071	Zapalote Chico ²	1,221.377
Elotero de Sinaloa ⁴	100.623	Zapalote Grande ²	29.341

Table 12 – Forecasted species range change by maize landrace for 2041 – 2060. Landrace groups assigned to each maize landraces are exponentiated at end of landrace name. Percent (%) Loss values greater than .90 are in bold.

Landrace	% Loss	% Gain	Landrace	% Loss	% Gain
Ancho ¹	54.3	4,999.6	Elotes Cónicos ⁵	76.0	8.902
Apachito ²	98.3	4.751	Elotes Occidentales ¹	71.9	68.331
Arrocillo Amarillo ²	96.5	0	Gordo ²	97.5	0
Azul ²	100	0	Jala ⁴	79.3	2.691
Blando ³	1.88	96.668	Mushito ²	46.4	99.595
Bofo ⁴	30.0	41.732	Olotillo ³	67.1	15.062
Bolita ²	45.2	171.72	Olotón ⁵	72.9	11.208
Cacahuacintle ⁵	84.1	3.959	Onaveño ³	10.4	85.026
Celaya ¹	60.0	133.02	Palomero Toluqueño ⁵	99.6	0
Chalqueno ¹	55.6	25.053	Pepitilla ¹	13.7	53.71
Chapalote ⁵	0	273.08	Ratón ⁵	24.4	37.06
Chiquito ²	78.5	0.701	Reventador ³	1.55	152.91
Comiteco ¹	47.1	23.648	Tablilla de Ocho ⁴	63.0	14.471
Conejo ¹	5.09	65.878	Tabloncillo Perla ⁴	4.04	233.24
Cónico Norteño ²	99.1	0.316	Tabloncillo ⁴	49.8	44.324
Cónico ⁵	92.7	1.661	Tehua ²	58.8	6.739
Coscomatepec ²	95.5	1.743	Tepecintle ¹	16.0	104.62
Cristalino de	99.5	0.63	Tuxpeño Norteño ²	62.5	30.961
Cubano Amarillo ¹	13.0	382.16	Tuxpeño ⁵	49.3	9.815
Dulce ⁴	64.5	16.728	Vandeno ¹	16.3	285.43
Dulcillo del Noroeste ²	4.09	45.926	Zamorano Amarillo ³	91.5	11.046
Dzit Bacal ⁴	0.32	46.392	Zapalote Chico ²	40.5	1261.9
Elotero de Sinaloa ⁴	19.2	119.86	Zapalote Grande ²	38.1	67.534

Clusters of Group 2 Maize Diversity Changes and Indigenous Communities

To answer the research question “Where might climate change, through changing maize’s distribution and diversity, have the greatest impacts on indigenous communities?”, this section presents results from a case study of using Group 2 mean diversity change by *municipio*. Group 2 was chosen because this cluster exhibits the highest levels of diversity in highly marginalized indigenous *municipios*. This section is further targeted on Group 2 mean diversity change because Group 2 maize landraces are forecasted to have the most highly negative distributional change across maize landrace groups. These maps help identify the geographic locations of the nuanced changes of Group 2 diversity change in Mexico and the particular communities most likely to be affected.

Bivariate local indicator of spatial association (BiLiSA) maps were conducted using Group 2 maize landrace diversity change and indigenous population percentages by *municipio* (Fig. 37). The BiLiSA utilized an 8 K-nearest neighbors SWM to ensure that all *municipios* had an equal number of neighbors. Further, the BiLiSA was run using the maximum number of permutations (99,999) and significance filter of 0.025 pseudo p-value with False Discovery Rate to help identify only significant clusters

BiLiSA Cluster Map: munidiv_time, g250d w/ INDOBPCT (99999 perm)

- Not Significant (1231)
- High-High (248)
- Low-Low (254)
- Low-High (249)
- High-Low (529)

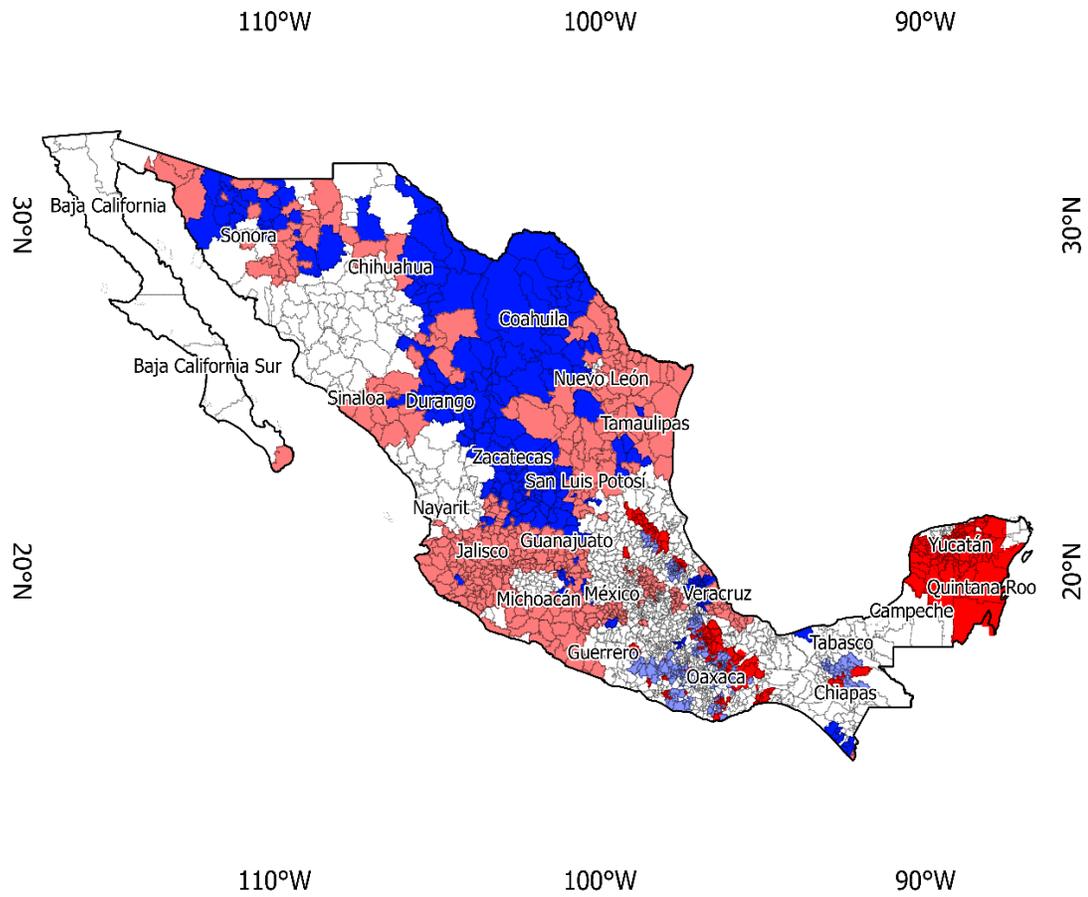


Figure 37 – Bivariate local indicator of spatial association (BiLiSA) map of Group 2 mean diversity change and indigenous population percentage by *municipio*.

Low-low clusters refer to *municipios* where diversity losses are larger than what would be expected by chance *and* where the local average percentage of indigenous population are lower than expectations. Significant high-high clusters have the opposite interpretation: forecasted diversity increases are atypically high in areas characterized by high local averages of percent indigenous population. Meanwhile, Low-High clusters are *municipios* forecasted to have substantial decreases in maize diversity that are situated in areas with high average indigenous population percentages. Lastly, High-Low clusters indicate significant *municipios* with high forecasted increases in Group 2 diversity that are situated in areas with low indigenous population percentages.

From a social justice perspective, the “Low-High” clusters are perhaps the most critical, insofar as they represent areas where indigenous populations are projected to experience substantial maize diversity loss. Such areas exist largely in the 5 geographic regions of high Group 2 maize diversity identified previously in the states of Puebla, Oaxaca, and Chiapas; however, other significant clusters exist along the southwestern sides of the Sierra Madre Oriental mountain range, west of Veracruz.

As another matter of equity, the BiLiSA map detects numerous areas where maize diversity is likely to increase, but where indigenous populations are low (High-Low clusters). If gains in diversity area a positive outcome—or, at minimum, a less negative outcome than diversity loss—then these areas represent *municipios* where non-indigenous populations stand to experience better outcomes than the indigenous populations in the “Low-High” clusters described above. In this case, the High-Low areas occur through from northeast Mexico in Tamaulipas and Nuevo Leon, spreading southwest to Michoacán, Colima, Jalisco, southwest Guanajuato, west Guerrero, and

southeast Mexico state. These High-Low clusters area bisected with clusters of significant decreases in Group 2 maize diversity in areas with low indigenous population percentages, particularly in the western portion of the Northern Plains in eastern Durango, eastern Durango, western San Luis Potosi, the majority of Coahuila, and northeastern portions of Chihuahua. Though, these values through western Guerrero should be evaluated with skepticism due to the high dissimilarity values shown in the MESS maps for this area.

X. CONCLUSIONS

Maize Landrace Diversity

The ensemble forecasted alpha diversity data presented here provide novel insight into the spatial distribution of maize diversity in Mexico at a relatively high resolution. These projections reveal that the highest levels of maize alpha diversity exist in Sonora and Sinaloa in northwestern Mexico, on the western river basins of the Sierra Madre Occidental mountain range.

Because most of Mexico appears to be suitable for at least one type of maize landrace, except for at very high elevations, hierarchical clustering of landraces by variable offers a useful generalization of maize diversity. Although the groups created here do not relate strictly to any *a-priori* physiographic (Perales and Golicher 2014), phenotypic (CONABIO), or phylogenetic (Kato et al. 2009) groups of maize diversity, certain patterns emerge in ecological and geographic space. While qualitative examinations show that though these groups do not conform completely with other *a-priori* clusters, some agreement certainly exists between these groups and CONABIO's 7 maize landrace groups⁶, which is based on geographic and phenotypic indicators.

From this research, we find that maize landrace diversity, by group, exhibit high spatial clustering (Moran's I all > 0.75). Although the specific clusters of maize landrace groups do not appear entirely compact, they do appear to exhibit starkly different biogeographic and physiographic zones.

⁶ Found here: <http://www.biodiversidad.gob.mx/usos/maices/razas2012.html>

Group 1 maize landraces appear most often in the Western Coastal Mountain Range, Central Plateau and Bajío biogeographic regions utilized in (Perales and Golicher 2014)⁷, created using the Biodiverse software, and redrawn to conform with INEGI's physiographic sub-provinces. Further, this group appears to encapsulate and spread from the Balsas River drainage basin in southwestern Mexico, where maize is hypothesized to have evolved. From this vantage point, Group 1 may correspond to the most ancestral maize landraces, from which other maize landraces evolved, and help inform maize landrace evolution.

In contrast, Group 2 landraces are modeled to exist largely in the Northern Plateau, Baja California and Northwest, but with the highest diversity levels in the Oaxacan Valleys and Sierras and Chiapas Complex biogeographic regions in the south, as well as the Chihuahua Canyons biogeographic regions in the north.

Group 3 landraces also appear to exist largely in the Baja California and Northwest biogeographic regions, with more isolated groups throughout portions of the northern portion of the Gulf and Isthmus Plans and Bajío biogeographic regions. Smaller, more isolated clusters of Group 3 appear also in western coastal regions in the Northwest Sierras, Western Coastal Mountain Range and Oaxacan Valley and Sierras biogeographic regions.

As stated, Group 4 largely overlaps the Group 3 maize landraces in the Baja California and Northwest biogeographic region but appears more spatially spread southward into the Northwest Sierras and into the lower elevation areas within the

⁷ See figure 1 for biogeographic regions:
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114657>

Western Coastal Mountain Range biogeographic regions on the western coast. Group 4 also exhibits spatial clusters within the Yucatan Peninsula biogeographic region.

Lastly, Group 5 landraces tend to exist largely along the Gulf coastal biogeographic regions of the Yucatan Peninsula, Gulf and Isthmus Plans, and Northern Plateau, with the highest levels of diversity concentrated in the Central Plateau region near Mexico City.

Associations between Indigeneity and Maize Landrace Diversity

Given the spatial distributions of mean maize landrace diversity by group, this research finds that maize diversity tends to be highest in *municipios* with dispersed indigenous population or those with indigenous presence. Indeed, in *municipios* classified as indigenous populations, maize diversity decreases significantly between groups classified as dispersed indigeneity or an indigenous presence for Groups 1, 4, and 5. In contrast, Group 2 diversity is highly contingent upon indigenous population presence, marginality, and percentage. Difference of means tests reveal that Group 2 diversity increases significantly highly with increasing levels of indigeneity. Using linear regression, this research further finds Group 2 diversity increases with both increasing marginality and indigenous population percentages.

While maize diversity tends to be lower for Groups 1, 3, 4, and 5 with increasing levels of indigeneity, this research finds that only Groups 4 and 5 show an inverse link between diversity and marginality, with higher diversity found in areas with lower marginality. Meanwhile, Group 1 diversity appears to increase with increasing levels of indigenous marginality, with the highest levels of diversity found in *municipios* with Very High levels of marginality. These findings show that while maize diversity appears

continent upon indigeneity, the direction and magnitude of these relationships vary by individual landrace and landrace group. Further, the interactions between indigenous presence, indigenous marginality, and maize diversity is, understandably, highly complex. We can infer that maize diversity is most influenced by highly marginalized *municipios* with high indigenous populations within the Group 2 cluster, but that maize landraces diversity within the Group 1 is influenced heavily by marginalized communities within *municipios* with both indigenous and non-indigenous ethnic groups. In general, we also find that alpha diversity within Group 5 maize landraces is associated not with marginalized indigenous communities, but with, perhaps, rural, Latino farmers. Similarly, Group 4 diversity is highest in communities with (dispersed) indigenous presence, but that these communities tend to have low levels of marginality.

Climate Change and Maize Landraces: Model Performance and Forecasts

Overall, ensemble model performance for the current time-period performed fairly well on evaluation data across maize landraces, with ensemble models tending toward specificity over sensitivity, thereby forecasting relatively conservative models.

Evaluation of individual MaxEnt response curves (not presented) generated in biomod2 show that models for some landraces do not capture the upper threshold of many bioclimatic predictors, particularly those relating to temperature. This is most likely due to the lack of threshold feature classes used during MaxEnt modeling, in combination with target-background sampling. Because target background sampling was restricted to areas within Mexico at locations corresponding to the Presence/Absence Matrix, the MaxEnt algorithm does not appear to capture the upper thresholds. If target sampling had been executed along with threshold features in MaxEnt modeling, ensemble forecasts

might perform better. While target background sampling appears to perform well for current distributions, including a background sample from a larger study area (e.g.: global, continental, etc.) may also aid in more realistic forecasts to future time-periods.

While considerable caution should be granted with regard to all ecological niche models that forecast to novel spaces or times, this research notes that these forecasts should be especially questioned in areas that exhibit negative Multivariate Environment Similarity Surfaces (MESS) values of high magnitude. While the MESS maps for all variables between current and future scenarios appear non-critical, these statistics are also inherently distorted with regard to indigenous population percentages since the future projections of indigenous population percentages by *municipio* assume no change. That is, both current and future predictors of indigenous population percentages are equal due to a lack of data, while this is recognized as not being the case. Further considering the lack of threshold features and target background sampling methods suggests that these forecasts might be somewhat biased.

Despite this, future forecasts of maize landrace distributions appear nonetheless highly influenced by climate change, whether in the positive or negative direction. From these forecasts, 9 of the maize landraces modeled are expected to range size decreases of -90% to -100% by 2041 – 2060, assuming full migration (Table 11). Though extinction rates from future forecasts should be interpreted cautiously, these results nevertheless propose large decreases in maize diversity on a national-level in Mexico, with a median decrease in current maize landrace suitability of 51% being lost by 2041 – 2060 across all landraces. While true across all landrace groups, diversity in Group 2 landraces may

disproportionately affect indigenous populations in 5 specific regions in southern Mexico.

On the flip side, maize landraces that are expected to increase have *drastic* percent increases. This research hypothesizes that if the maize landrace responses had incorporated threshold features, then the upper limits of the bioclimatic predictors of maize landraces, the direction of range change might change. While the forecasts are dubious, the magnitude of change is nevertheless apparent. Another way of rationalizing these ensemble forecasts is by suggesting that the maize landraces with positive increases might be those most resilient to climate change, while those with negative niche size changes are those least resilient to climate change.

XI. CONTRIBUTIONS

By way of ensemble, maximum entropy species niche modeling, this research presented ecological niches for 46 traditional Mexican maize landraces. Using ensemble model information, maize landraces are clustered into 5 groups according to variable importance. Mapping alpha diversity of maize landrace by hierarchical groups assigned indeed shows that maize landrace group niches exist in unique biogeographical clusters.

This research helps to unpack the spatial and statistical relationships between maize diversity and marginalized, indigenous communities at a national-level in Mexico. As supported by local research, this thesis finds that maize landrace diversity is contingent upon indigenous population percentage, particularly for Group 2 landraces, whose diversity is correlated with indigenous communities with elevated levels of marginality.

Situated on these findings, there is clear opportunity for future work that focuses on the negative impacts of climate change on maize diversity within Group 2 maize landraces, and specifically within the 241 *municipios* that exhibit significantly high levels of indigenous population percentages with significantly high decreases in Group 2 maize landrace diversity. These *municipios* could be further limited to include only the 121 *municipios* within the 5 geographic regions in Oaxaca, Veracruz, and Puebla that exhibit highly significant clusters of Group 2 maize diversity using Getis-Ord statistics with very low pseudo p-values (0.001) with False Discovery Rates.

As climate change is expected to increase marginality of communities worldwide (UNU-IAS 2008; Alexander, Chief Clarence, Nora Bynum, Liz Johnson, Ursula King, Tero Mustonen, Peter Neofotis, Noel Oetllé, Cynthia Rosenzweig, Chie Sakakibara,

Chief Vyacheslav Shadrin, Marta Vicarelli, Jon Waterhouse 2011; Sánchez-Cortés and Chavero 2011), particularly those of indigenous ethnicity (European Parliament n.d.; Thrupp 2000; UNU-IAS 2008), these findings may help prioritize future agricultural climate change mitigation programs regarding sustainable maize cultivation in Mexico.

As well as identifying the specific indigenous communities that might be most affected, this research has shown that the distributions of specific maize varieties -- Zapalote Chico, Apachito, Dzit Bacal, and Tehua in particular – are highly influenced by indigenous communities. In this way, this research aids in outlining the deep, underlying food networks that contribute to maize diversity in Mexico.

Further extending these conclusions, this thesis hopes to instigate research to examine intersectionality of marginalization, indigeneity, and climate change within indigenous communities at local scales both within Mexico and globally, given the findings presented herein. Though this research provides broad strokes with which to consider the impending effects of climate change on maize, only personal fieldwork will fully capture the local-level effects of climate change on individual families, farms, and farmers.

Given the GHG emission scenario implemented here, this research intends to focus the audiences' attention toward realistic, yet devastating, implications of climate change. These forecasts into the 2041 – 2060 time-period reveals drastic, though varied, changes in traditional Mexican maize suitability nationally. As a staple crop, indigenous to Mexico, maize provides not only sustenance, but also linkages to traditions, ancestry, and community. Though these forecasts should be considered with care, this research nonetheless hopes to encourage climate change mitigation research within

indigenous communities in Mexico not only to sustain a globally important food crop, but also to support the communities that continue these ancestral networks.

XII. LIMITATIONS OF PROPOSED DATA AND METHODS

A major limitation of this research lies in its bias toward secondary data and quantitative methods. Due to budgetary and time limitations, detailed qualitative analyses of community characteristics are not feasible. While this thesis suggests that marginalized, indigenous communities covary with maize diversity at the national-level, this study incorporates only easily accessible ethnic data. No data regarding seed management practices, indigenous knowledge, or agricultural techniques are examined in this research, though this helps provide an outline of future research.

Another limitation of this study is the limitedness of ecological scope. Numerous other ecological factors contribute to delimiting the realized niches of maize landraces, including pests, herbivory, diseases, etc. The methods proposed here cannot accommodate extensive biotic interactions, largely due to the of the scale of the study.

Another drawback of this data analysis derives from the rasterization of vector data for predictor variables in the ecological niche modelling function. While the vector data derived from Mexico's 2010 Housing and Population census is derived a relatively low geography-level, errors derived from the Modifiable Area Unit Problem are recognized.

Lastly, as stated previously, more research surrounding the effects of MaxEnt feature selection, species-specific model tuning, and background sampling strategies on maize landrace niche modeling in Mexico. While the forecasts presented here provide an insight into potential changes in maize diversity nationally, some of the individual landrace forecasts exhibit non-practical distributional changes. In particular, including threshold features within MaxEnt should be incorporated to better detect the upper

thresholds of maize landrace bioclimatic envelopes. In addition, species-specific tuning using ENMeval in R would no doubt increase model performance and allow for increased transferability to novel environments, particularly when tuning using AICc. Furthermore, future research should carefully evaluate the impacts of background sampling methods on transferability of MaxEnt models.

XIII. CONCLUDING REMARKS

While the methods employed in this thesis draw on prominent research (Ureta, Martínez-Meyer, Hugo R. Perales, et al. 2012; Ureta et al. 2013, 2015; Perales and Golicher 2014), it is worth noting that the analyses contained hereinbefore are relatively novel in their scale and magnitude. Much of the national-level ecological niche models for maize landraces in Mexico have used much larger spatial resolutions (5 km x 5 km or greater), as well as only one modeling algorithm (Ureta et al. 2013; Perales and Golicher 2014). In addition, whereas some previous research indicates low influence of social factors on maize at a national scale, this research uncovered an unambiguous statistical association between indigeneity and the distribution of maize diversity across Mexico. While the nature of that association is nuanced and varies based on maize groupings, there is evidence to suggest that at least some indigenous cultures are at higher risk of being impacted by maize diversity loss via climate change relative to non-indigenous populations in Mexico.

APPENDIX SECTION

A. Ensemble Model Evaluation Performance by KAPPA, ROC, TSS

Mean evaluation metric (scaled 0 – 1), Specificity, and Specificity scores averaged across 3 weighted-mean ensemble models

Landrace	Metric	Score	Sensitivity	Specificity
Ancho	KAPPA	0.465	48.023	98.323
	ROC	0.938	89.266	83.757
	TSS	0.727	81.356	91.288
Apachito	KAPPA	0.524	68.554	99.141
	ROC	0.990	100.00	94.440
	TSS	0.943	100.00	94.299
Arrocillo Amarillo	KAPPA	0.559	59.821	98.971
	ROC	0.964	91.071	89.038
	TSS	0.799	91.071	88.733
Azul	KAPPA	0.547	67.064	98.881
	ROC	0.987	98.810	93.641
	TSS	0.924	98.810	93.374
Blando	KAPPA	0.446	61.539	99.493
	ROC	0.989	100.00	93.677
	TSS	0.936	100.00	93.495
Bofo	KAPPA	0.439	52.174	99.658
	ROC	0.976	100.00	88.604
	TSS	0.883	100.00	88.227
Bolita	KAPPA	0.485	41.429	99.056
	ROC	0.843	70.476	85.017
	TSS	0.551	74.921	80.102
Cacahuacintle	KAPPA	0.228	43.860	97.802
	ROC	0.948	94.737	86.621
	TSS	0.812	94.737	86.420
Celaya	KAPPA	0.394	43.110	94.665
	ROC	0.835	74.912	76.536
	TSS	0.511	74.912	76.164

Appendix A (continued) - Mean evaluation metric (scaled 0 – 1), Specificity, and Specificity scores averaged across 3 weighted-mean ensemble models

Landrace	Metric	Score	Sensitivity	Specificity
Chalqueño	KAPPA	0.467	50.091	96.571
	ROC	0.911	80.874	84.898
	TSS	0.658	80.874	84.847
Chapalote	KAPPA	0.371	92.857	99.241
	ROC	0.995	100.00	98.353
	TSS	0.983	100.00	98.342
Chiquito	KAPPA	0.306	38.889	99.635
	ROC	0.968	94.444	90.380
	TSS	0.845	94.444	89.980
Comiteco	KAPPA	0.678	60.694	99.492
	ROC	0.967	94.798	86.969
	TSS	0.811	88.439	92.498
Conejo	KAPPA	0.352	38.314	98.862
	ROC	0.943	89.655	88.641
	TSS	0.781	89.655	88.474
Cónico Norteño	KAPPA	0.719	78.415	95.698
	ROC	0.963	90.756	90.702
	TSS	0.813	90.210	91.032
Cónico	KAPPA	0.695	73.983	95.587
	ROC	0.944	85.450	90.743
	TSS	0.761	85.450	90.729
Coscomatepec	KAPPA	0.598	60.000	99.603
	ROC	0.984	96.364	91.839
	TSS	0.880	96.364	91.536
Cristalino de Chihuahua	KAPPA	0.646	66.892	99.000
	ROC	0.986	98.649	93.550
	TSS	0.918	97.973	93.725
Cubano Amarillo	KAPPA	0.345	45.299	99.215
	ROC	0.956	87.179	92.759
	TSS	0.798	87.179	92.529

Appendix A (continued) - Mean evaluation metric (scaled 0 – 1), Specificity, and Specificity scores averaged across 3 weighted-mean ensemble models

Landrace	Metric	Score	Sensitivity	Specificity
Dulce	KAPPA	0.155	31.373	98.436
	ROC	0.909	100.000	69.033
	TSS	0.687	100.000	68.679
Dulcillo del Noroeste	KAPPA	0.245	52.381	98.988
	ROC	0.982	100.000	91.169
	TSS	0.911	100.000	91.052
Dzit.Bacal	KAPPA	0.480	70.892	98.485
	ROC	0.973	87.324	96.668
	TSS	0.840	87.324	96.644
Elotero de Sinaloa	KAPPA	0.381	53.234	98.564
	ROC	0.971	95.522	90.889
	TSS	0.863	95.522	90.729
Elotes Cónicos	KAPPA	0.367	50.583	93.097
	ROC	0.896	91.833	73.929
	TSS	0.656	91.833	73.696
Elotes Occidentales	KAPPA	0.349	65.940	91.340
	ROC	0.893	81.091	84.510
	TSS	0.655	81.091	84.436
Gordo	KAPPA	0.444	44.828	99.437
	ROC	0.982	98.276	93.967
	TSS	0.922	98.276	93.967
Jala	KAPPA	0.270	59.649	99.047
	ROC	0.979	94.737	94.880
	TSS	0.895	94.737	94.703
Mushito	KAPPA	0.470	55.165	97.781
	ROC	0.919	81.871	85.214
	TSS	0.668	80.702	86.037
Olotillo	KAPPA	0.565	57.810	96.712
	ROC	0.925	87.221	81.940
	TSS	0.690	87.356	81.625

Appendix A (continued) - Mean evaluation metric (scaled 0 – 1), Specificity, and Specificity scores averaged across 3 weighted-mean ensemble models

Landrace	Metric	Score	Sensitivity	Specificity
Olotón	KAPPA	0.568	53.933	99.440
	ROC	0.961	91.011	91.956
	TSS	0.828	89.888	92.861
Onaveño	KAPPA	0.483	38.710	99.769
	ROC	0.965	95.161	89.765
	TSS	0.847	95.161	89.444
Palomero Toluqueño	KAPPA	0.392	46.296	99.434
	ROC	0.979	97.222	92.049
	TSS	0.892	97.222	91.937
Pepitilla	KAPPA	0.536	61.806	97.839
	ROC	0.933	81.945	90.190
	TSS	0.721	82.639	89.407
Ratón	KAPPA	0.492	61.418	95.208
	ROC	0.916	83.425	88.029
	TSS	0.714	82.320	89.025
Reventador	KAPPA	0.281	48.148	98.224
	ROC	0.959	98.148	86.753
	TSS	0.847	98.148	86.504
Tablilla de Ocho	KAPPA	0.175	33.333	98.536
	ROC	0.929	97.143	77.160
	TSS	0.739	93.333	80.501
Tabloncillo Perla	KAPPA	0.377	32.609	99.398
	ROC	0.962	97.826	85.071
	TSS	0.826	97.101	85.304
Tabloncillo	KAPPA	0.565	70.811	95.363
	ROC	0.938	91.622	83.186
	TSS	0.746	91.622	82.947

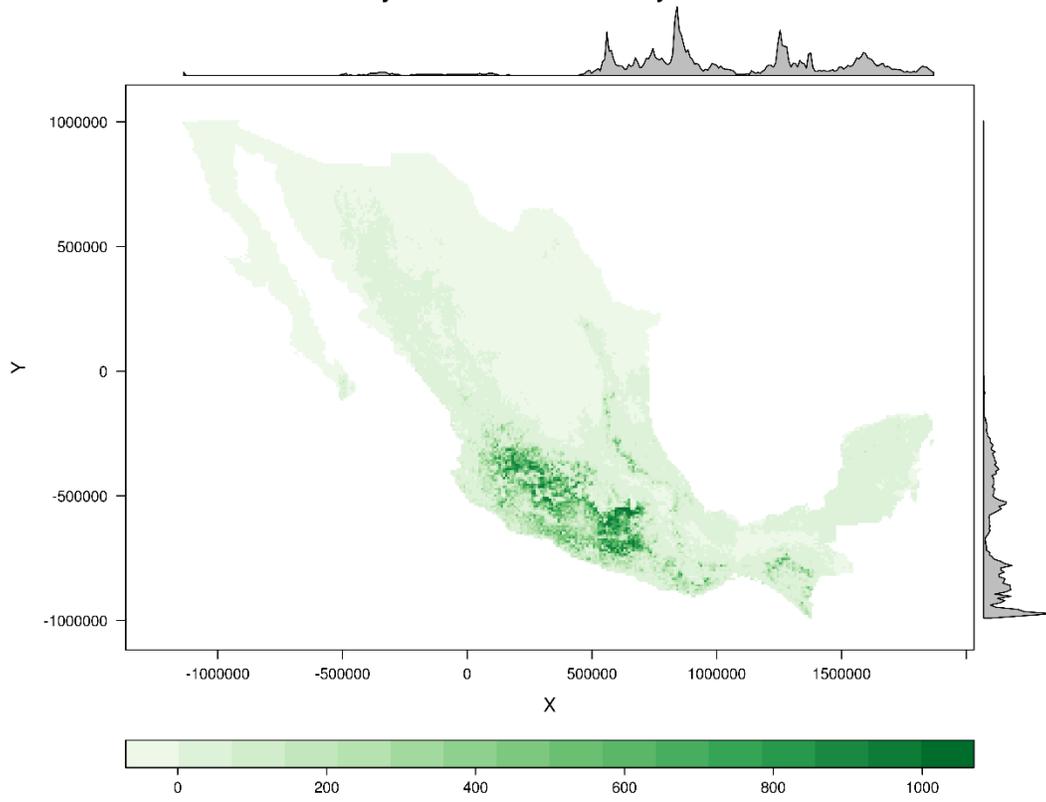
Mean evaluation metric (scaled 0 – 1), Specificity, and Specificity scores averaged across 3 weighted-mean ensemble models

Landrace	Metric	Score	Sensitivity	Specificity
Tehua	KAPPA	0.238	36.364	99.352
	ROC	0.964	95.455	89.201
	TSS	0.843	95.455	88.818
Tepecintle	KAPPA	0.416	51.038	96.087
	ROC	0.940	93.773	82.
	TSS	0.758	93.529	82.384
Tuxpeño Norteño	KAPPA	0.550	53.054	98.780
	ROC	0.935	86.388	88.074
	TSS	0.745	86.388	88.031
Tuxpeño	KAPPA	0.543	66.978	88.505
	ROC	0.886	81.257	79.132
	TSS	0.604	81.360	78.995
Vandeño	KAPPA	0.347	38.406	97.194
	ROC	0.893	83.043	82.175
	TSS	0.652	83.043	82.047
Zamorano Amarillo	KAPPA	0.452	59.375	98.873
	ROC	0.965	92.188	88.438
	TSS	0.805	91.667	88.806
Zapalote Chico	KAPPA	0.802	67.164	100.000
	ROC	0.842	68.657	99.893
	TSS	0.685	68.657	99.893
Zapalote Grande	KAPPA	0.214	45.503	97.111
	ROC	0.931	95.238	78.470
	TSS	0.733	92.593	80.747

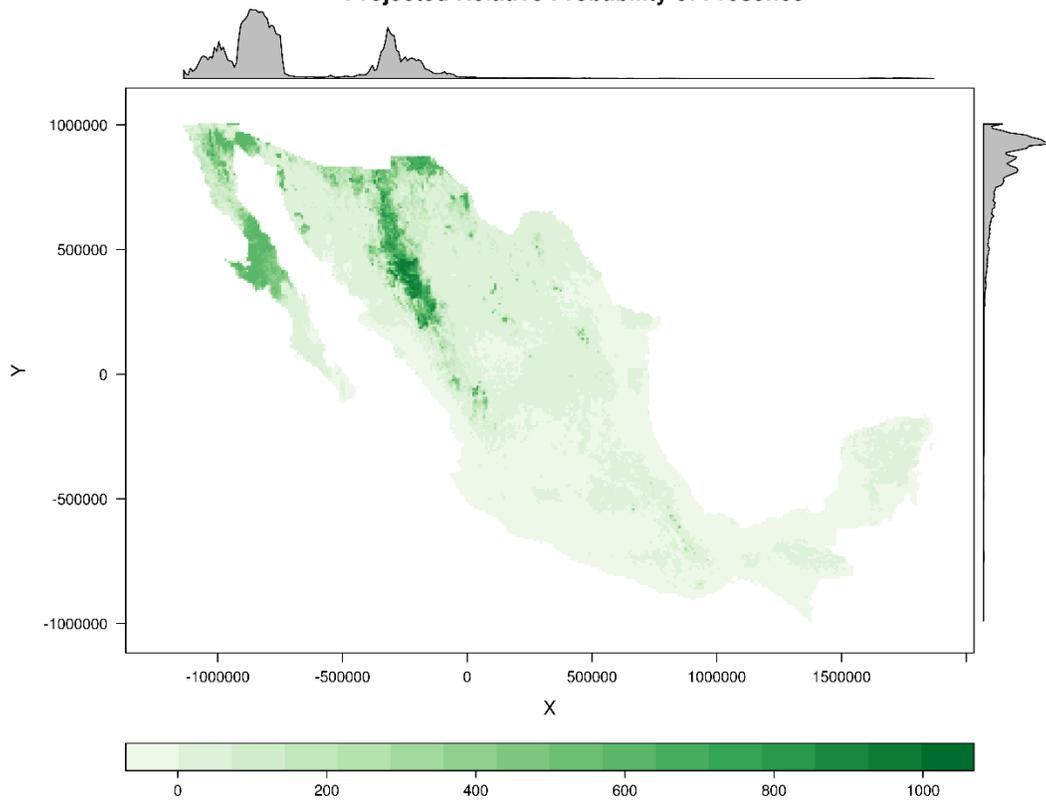
B. 1970 – 2000 Ensemble Forecasts by Landrace

The following maps depict the average relative probability of presence values across the weighted-mean ensemble models for the 1970 – 2000 time-period. These maps were utilized later to classify binary predictions of maize landrace existence with a 0.75 threshold value.

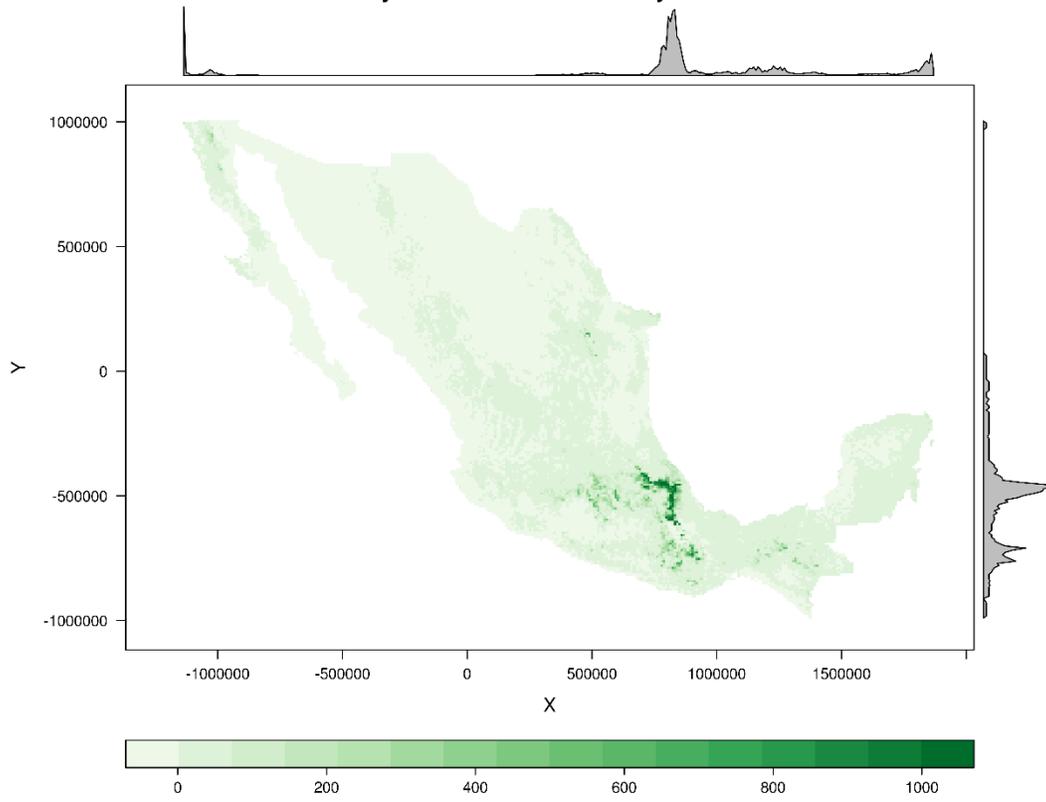
Ancho : 1970 - 2000
Projected Relative Probability of Presence



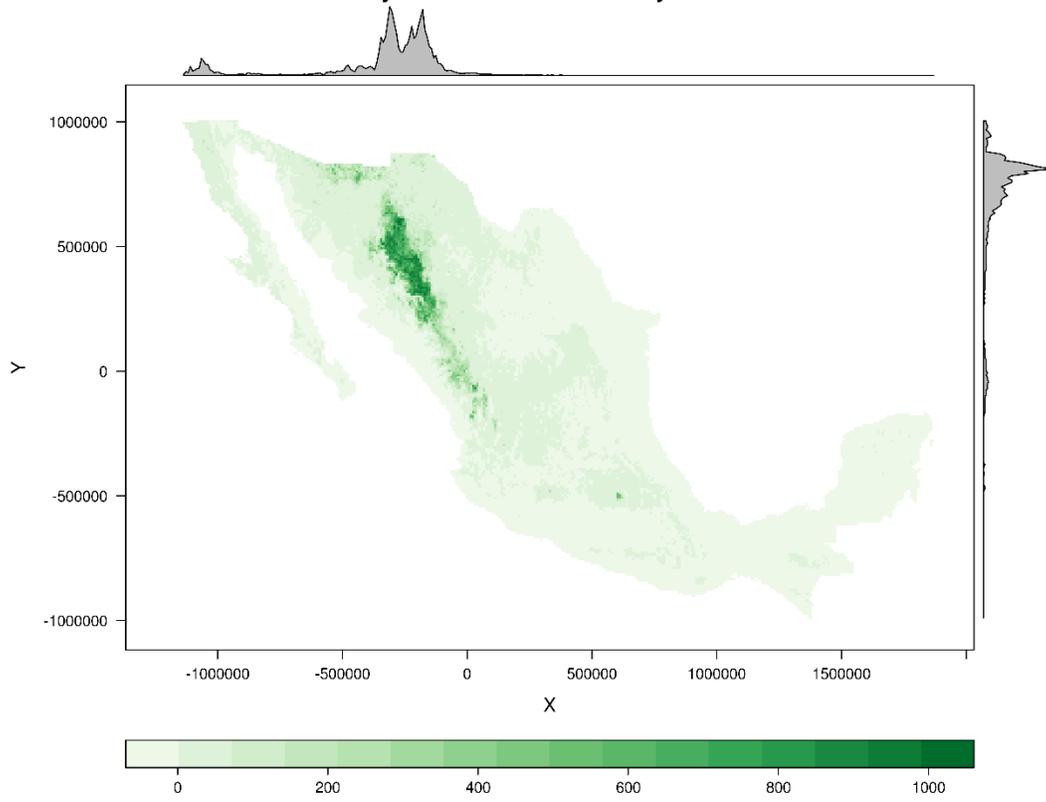
Apachito : 1970 - 2000
Projected Relative Probability of Presence



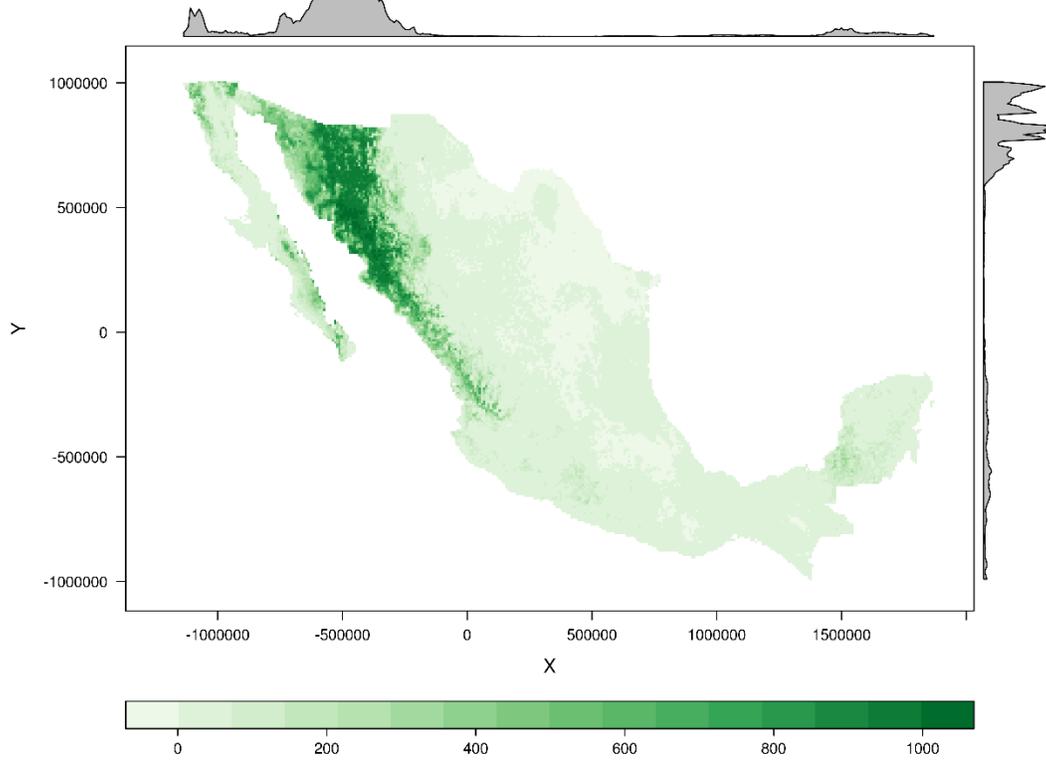
Arrocillo.Amarillo : 1970 - 2000
Projected Relative Probability of Presence



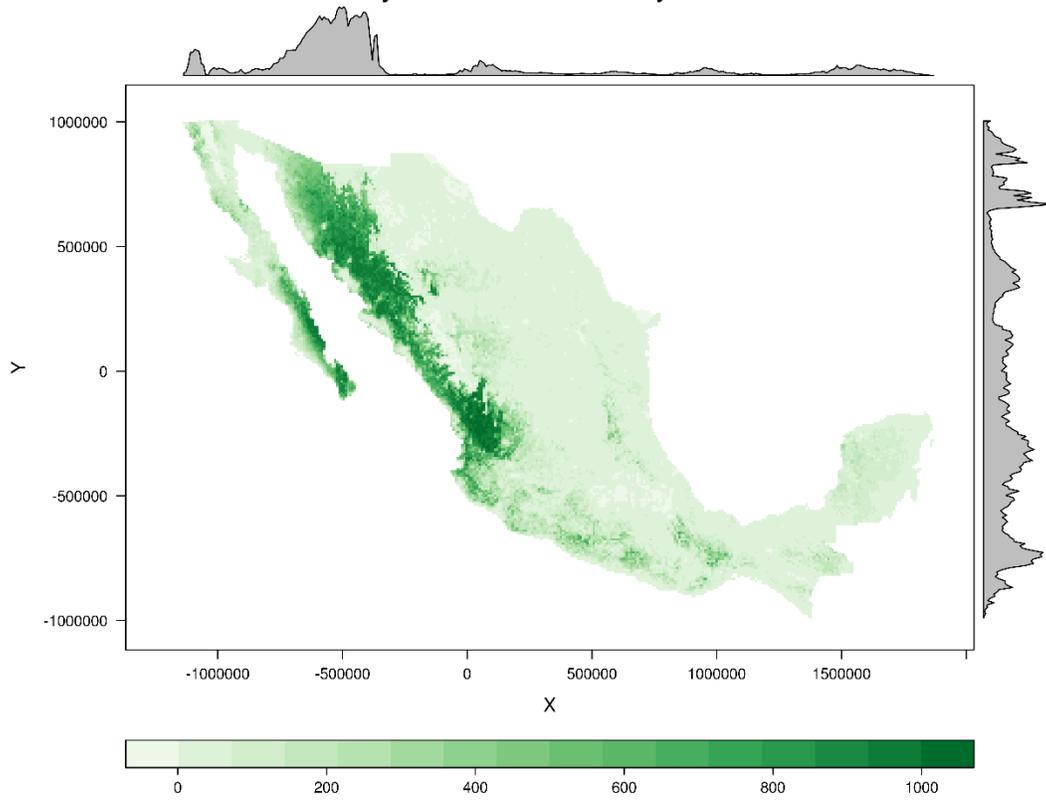
Azul : 1970 - 2000
Projected Relative Probability of Presence



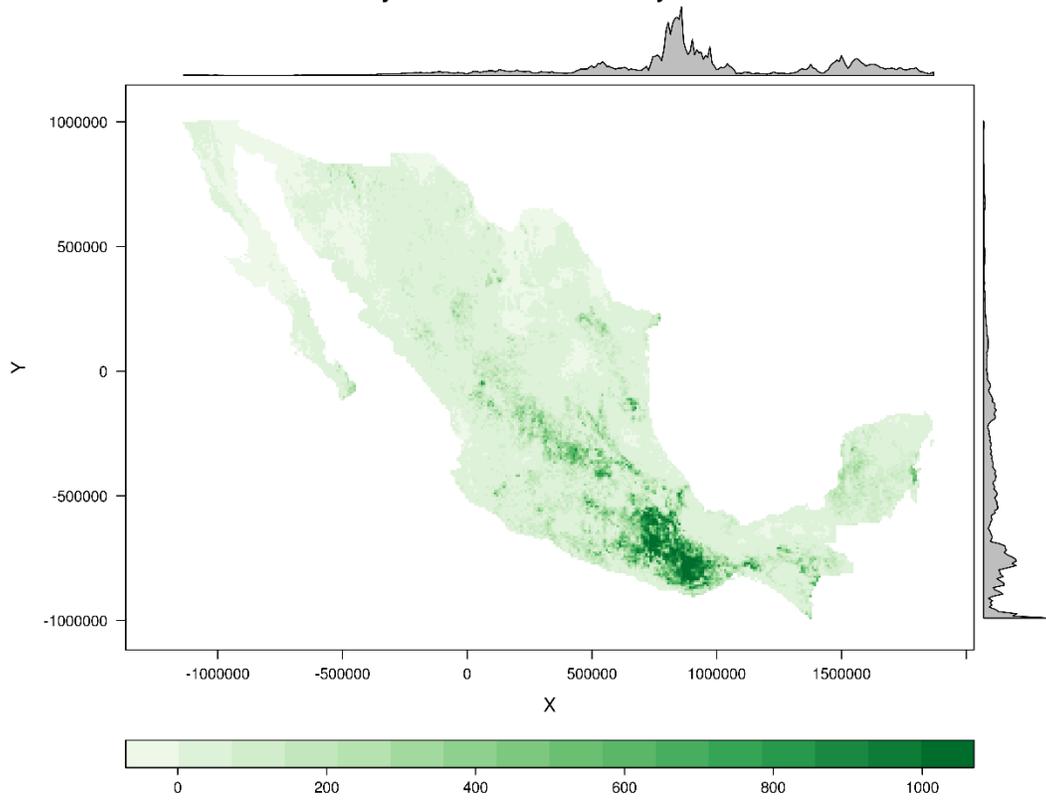
Blando : 1970 - 2000
Projected Relative Probability of Presence



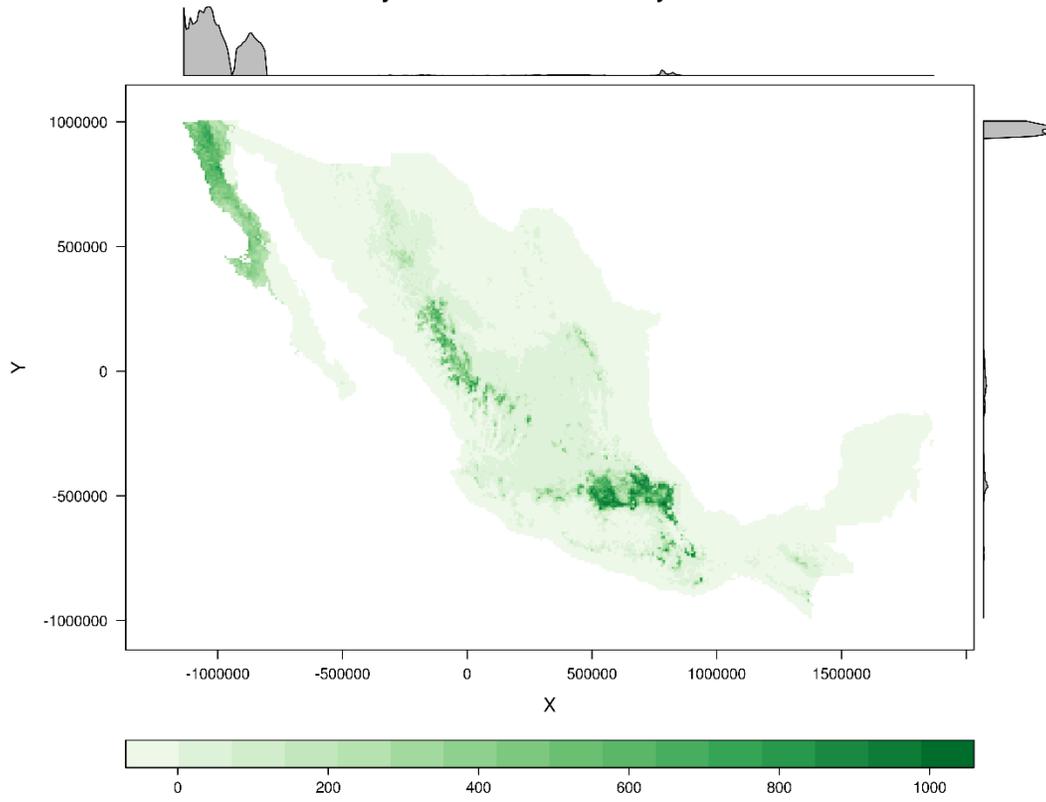
Bofa : 1970 - 2000
Projected Relative Probability of Presence



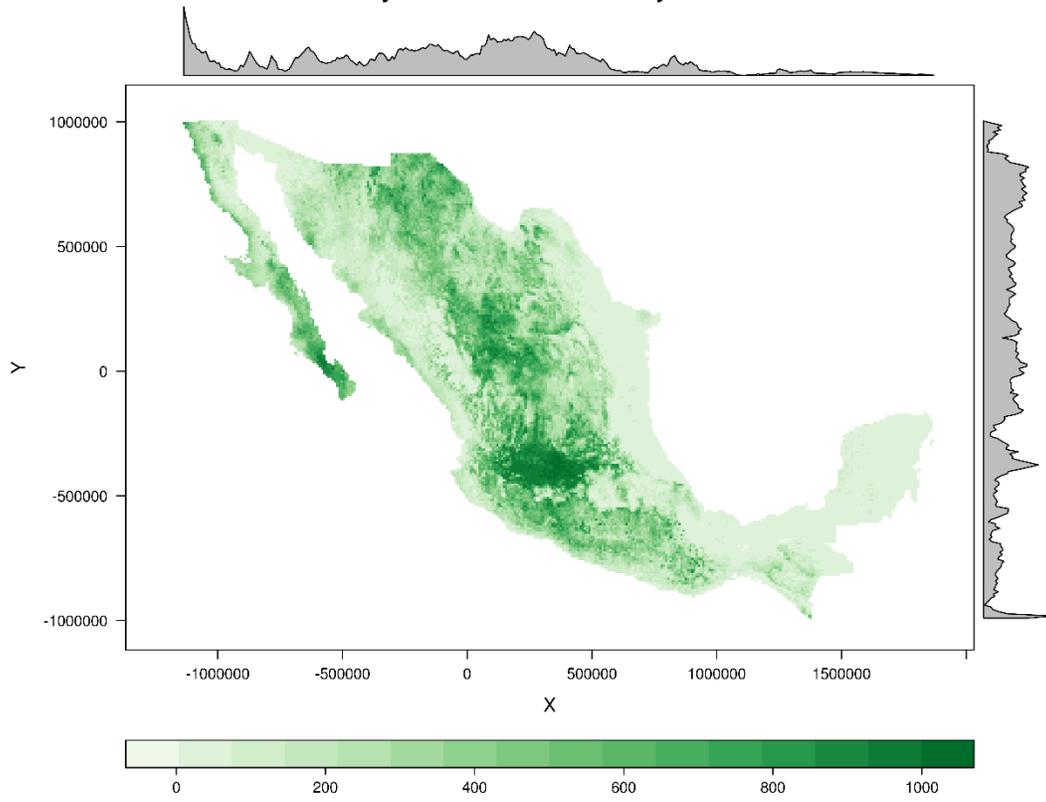
Bolita : 1970 - 2000
Projected Relative Probability of Presence



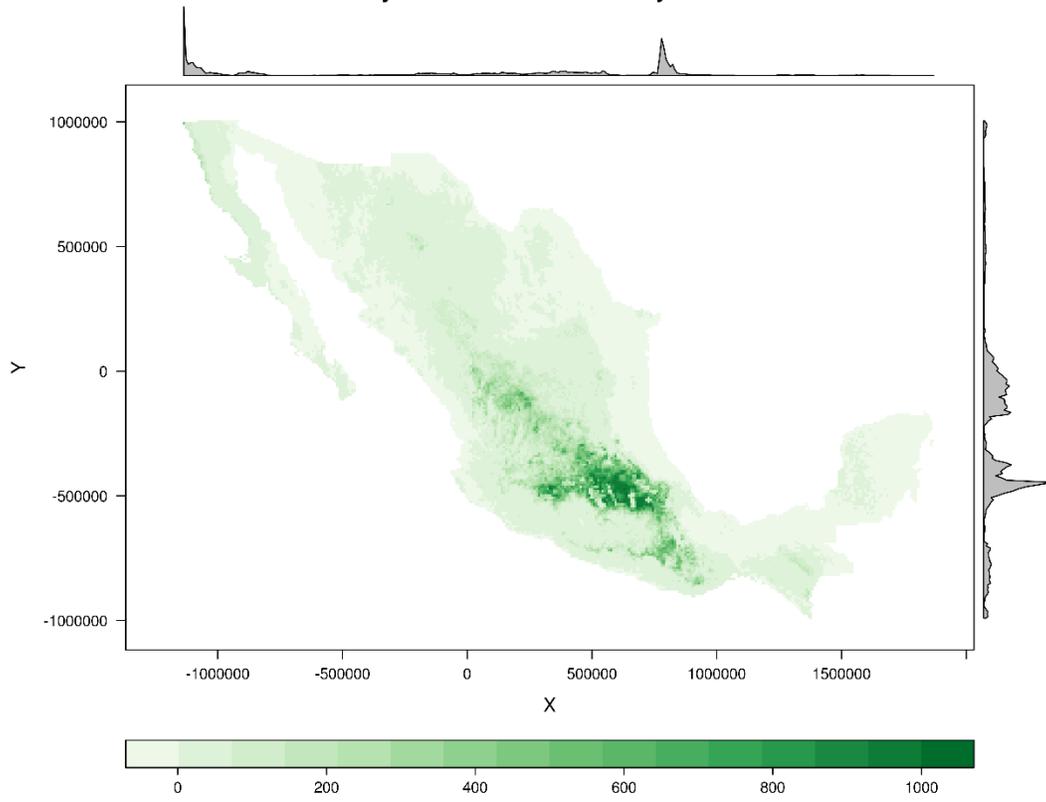
Cacahuacintle : 1970 - 2000
Projected Relative Probability of Presence



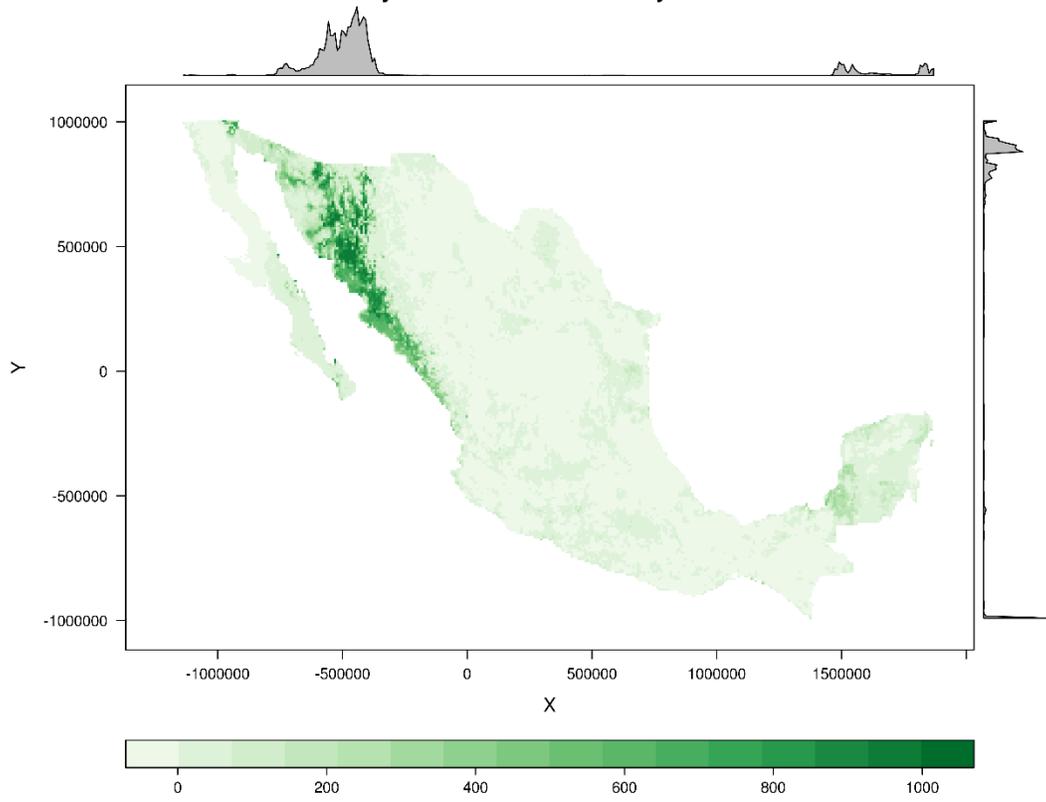
Celaya : 1970 - 2000
Projected Relative Probability of Presence



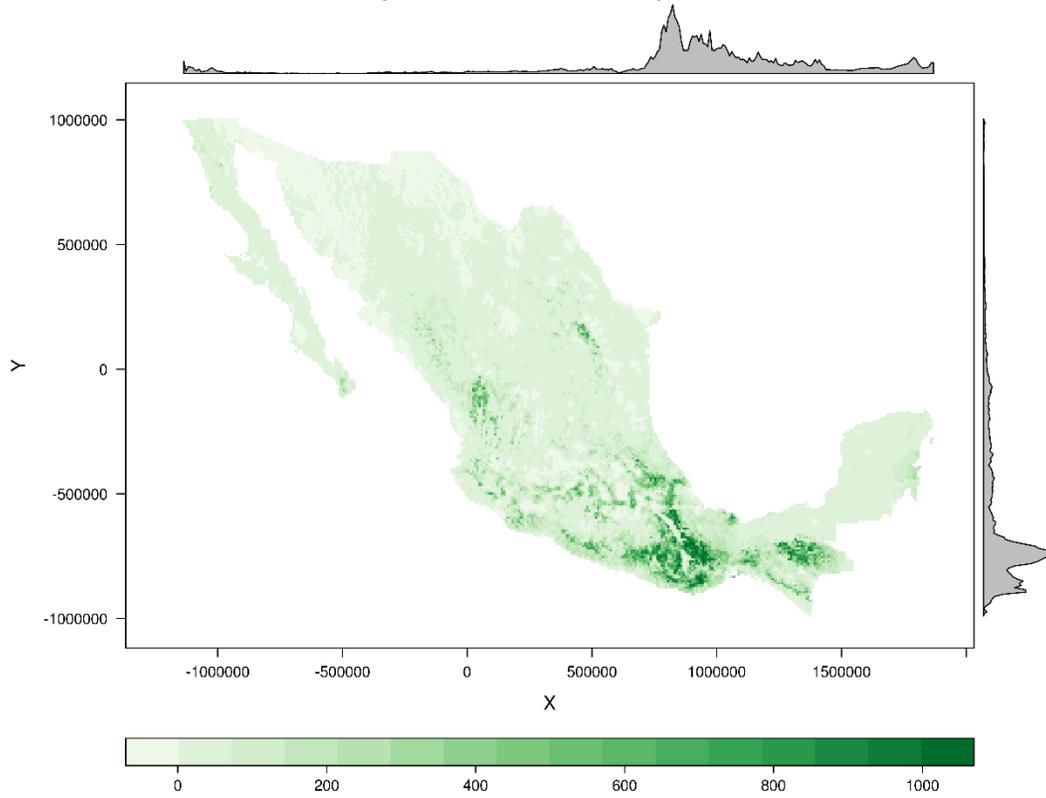
Chalqueno : 1970 - 2000
Projected Relative Probability of Presence



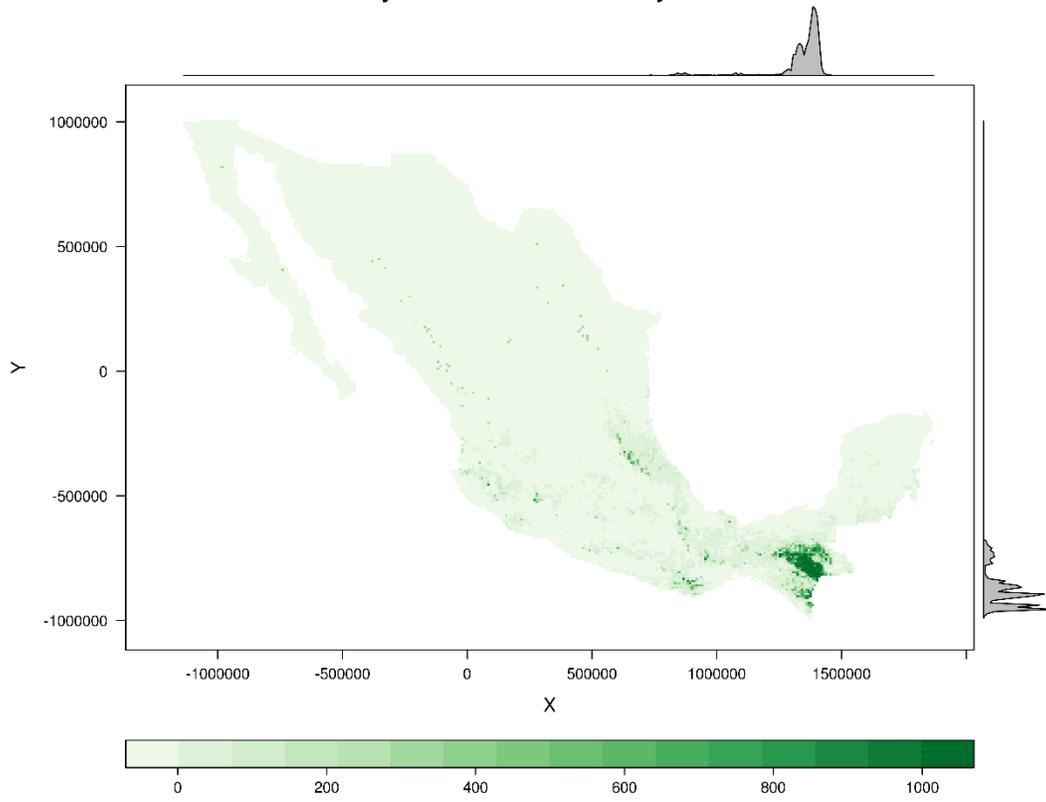
**Chapalote : 1970 - 2000
Projected Relative Probability of Presence**



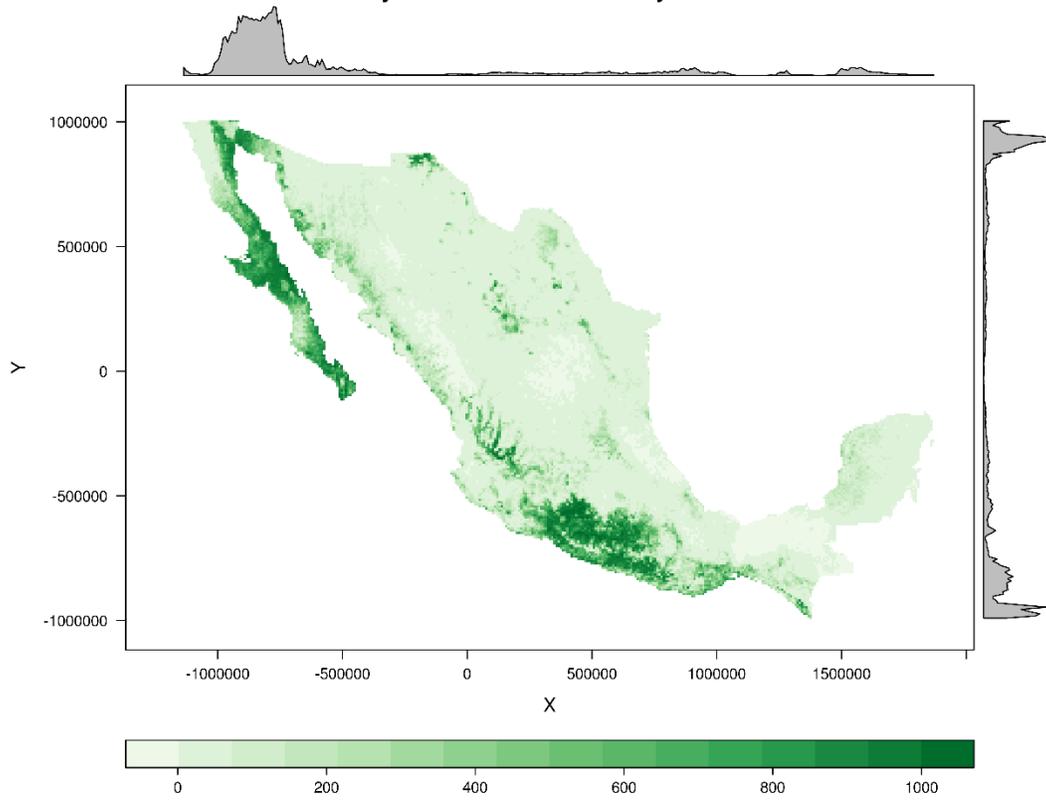
Chiquito : 1970 - 2000
Projected Relative Probability of Presence



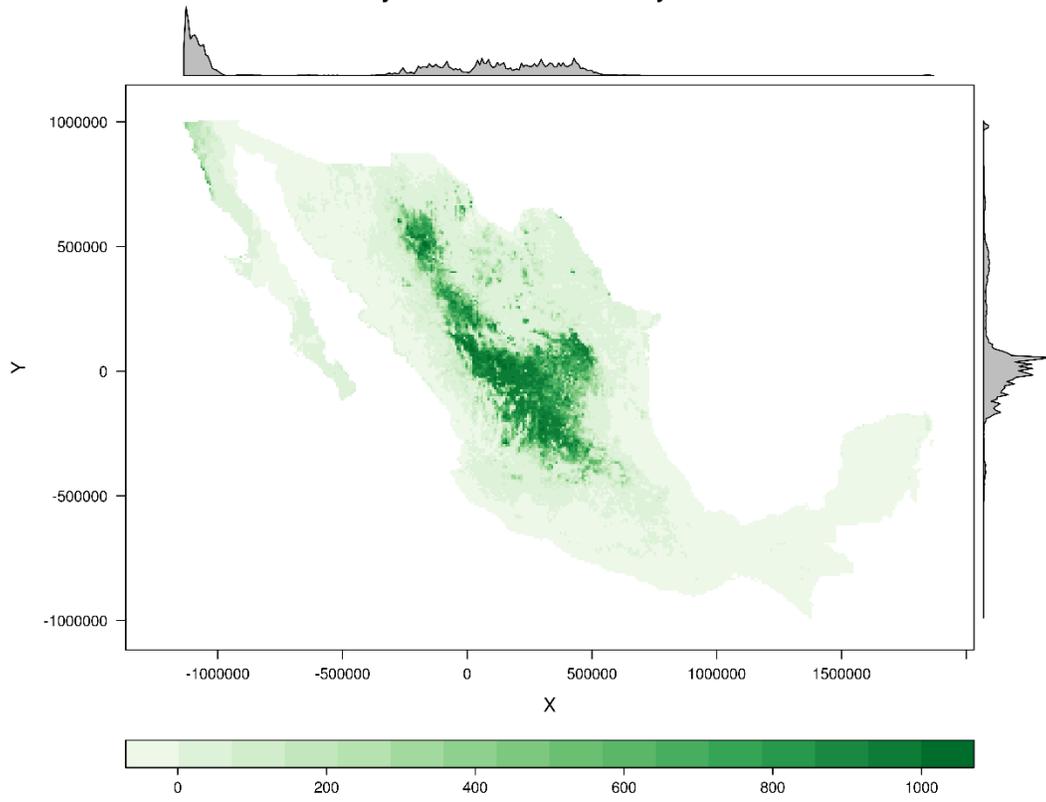
Comiteco : 1970 - 2000
Projected Relative Probability of Presence



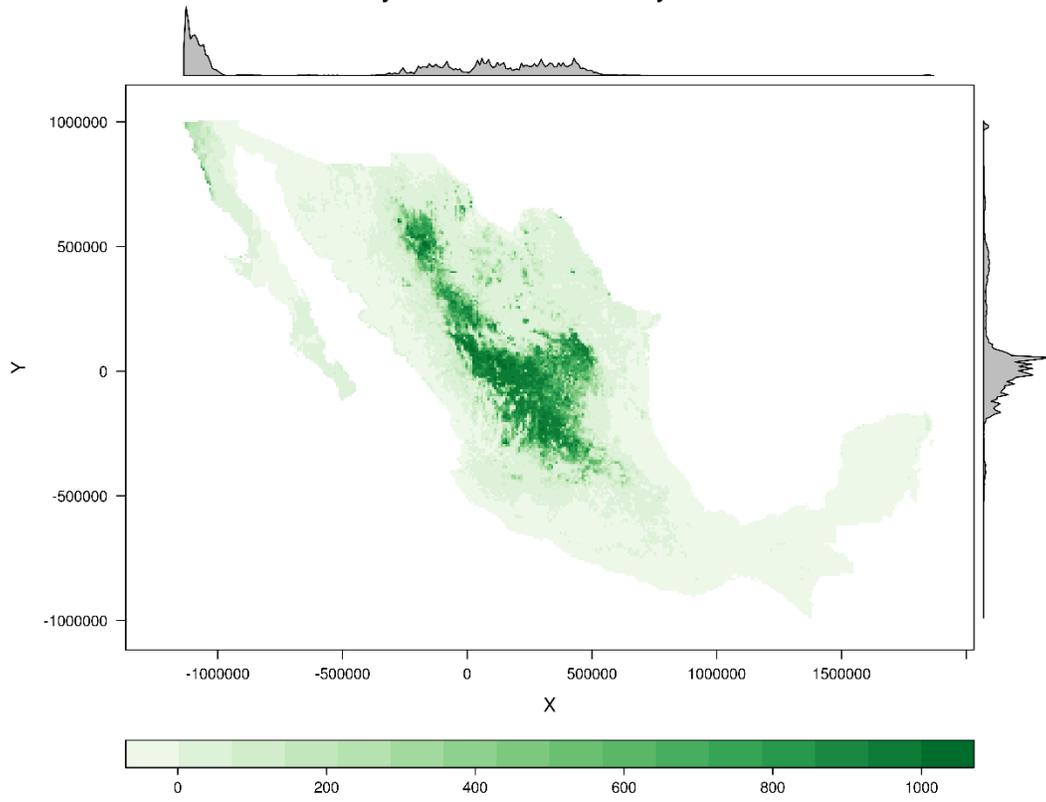
Conejo : 1970 - 2000
Projected Relative Probability of Presence



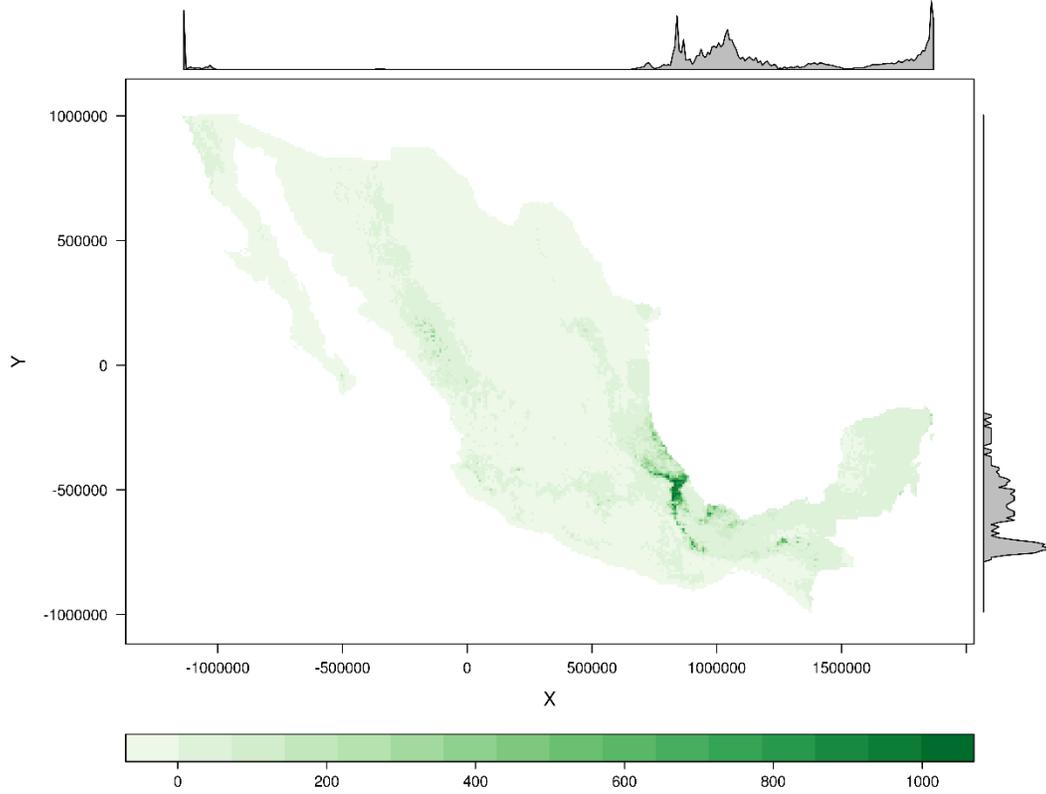
**Conico.Norteno : 1970 - 2000
Projected Relative Probability of Presence**



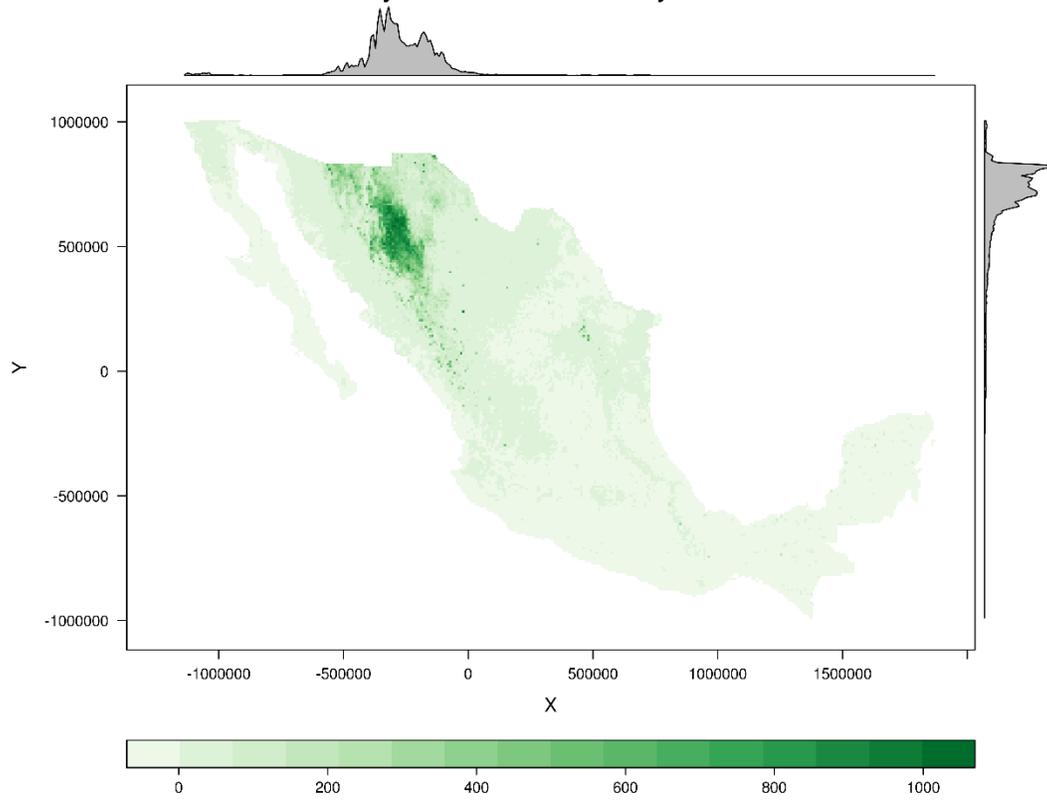
**Conico : 1970 - 2000
Projected Relative Probability of Presence**



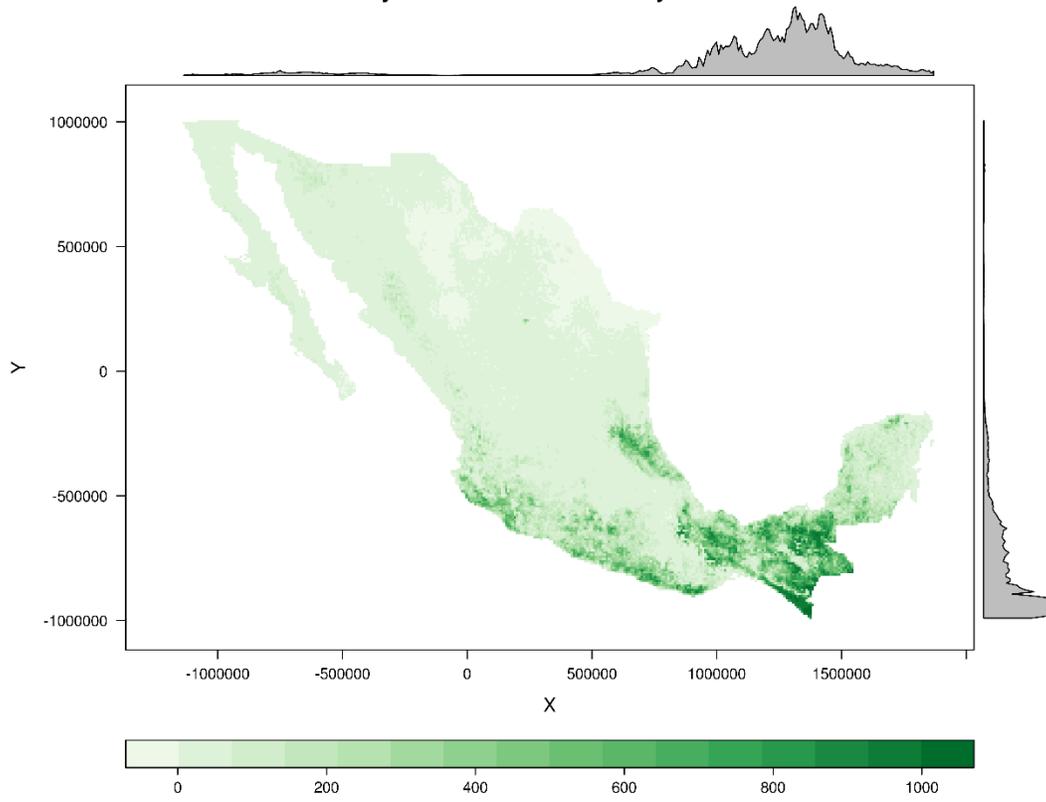
Coscomatepec : 1970 - 2000
Projected Relative Probability of Presence

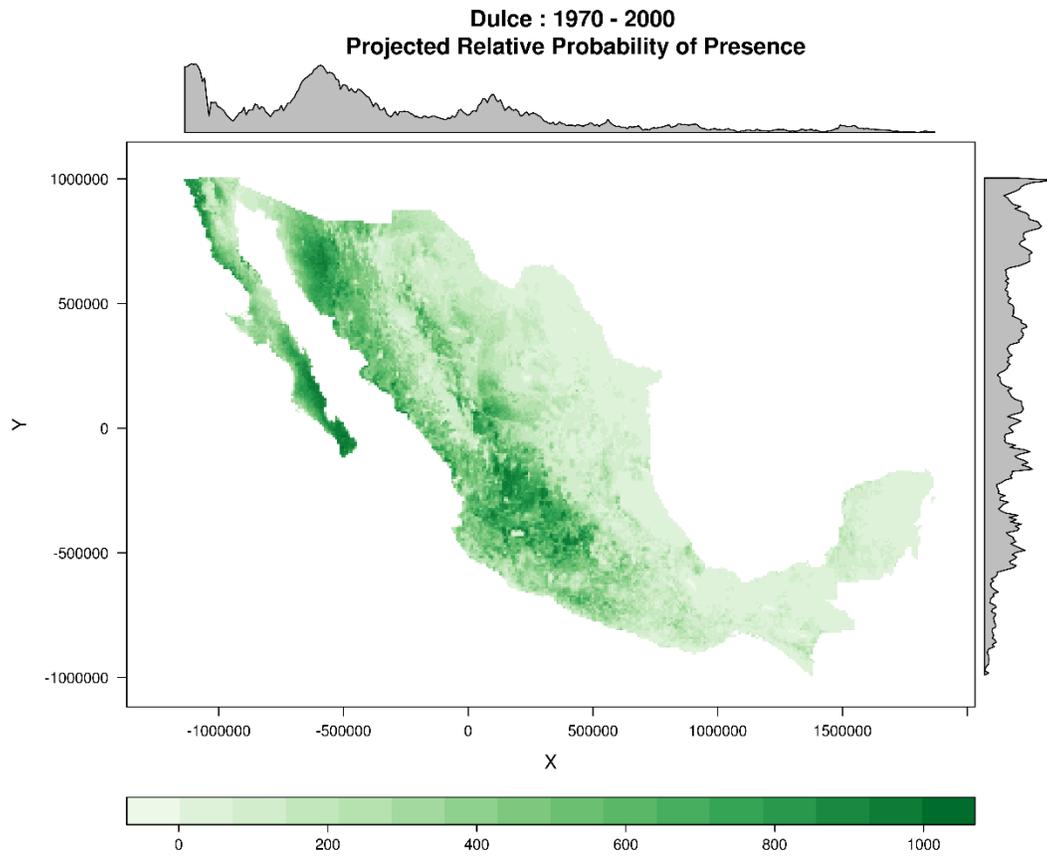


**Cristalino.de.Chihuahua : 1970 - 2000
Projected Relative Probability of Presence**

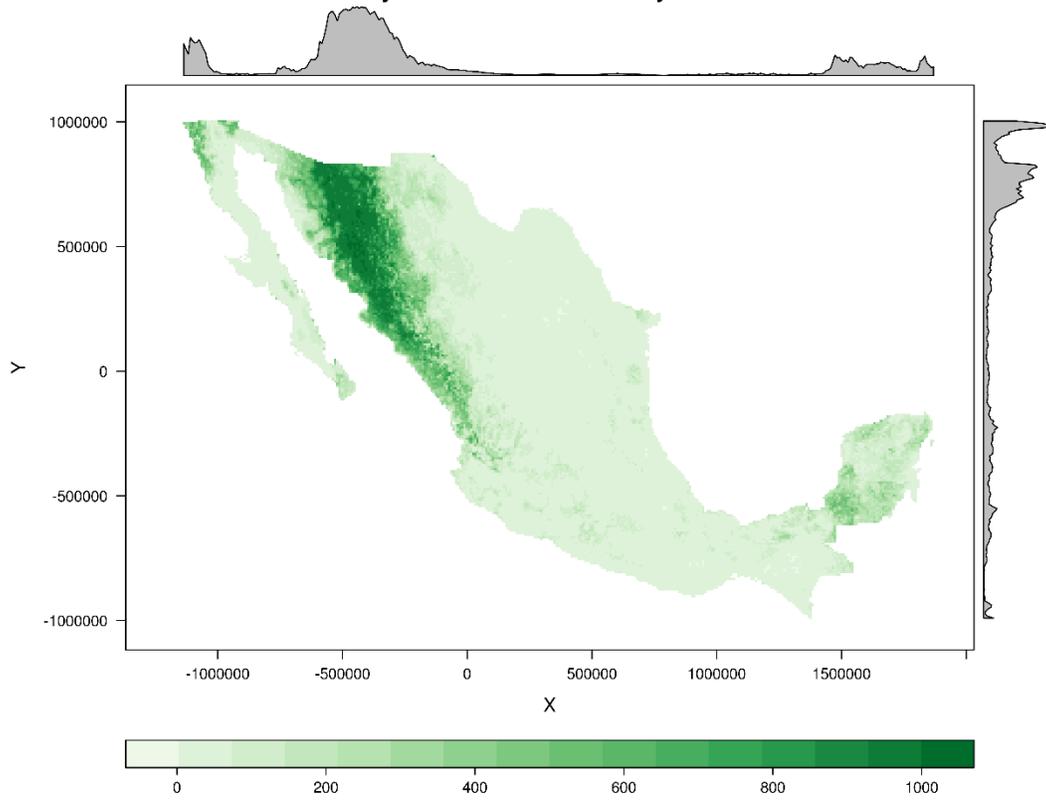


Cubano.Amarillo : 1970 - 2000
Projected Relative Probability of Presence

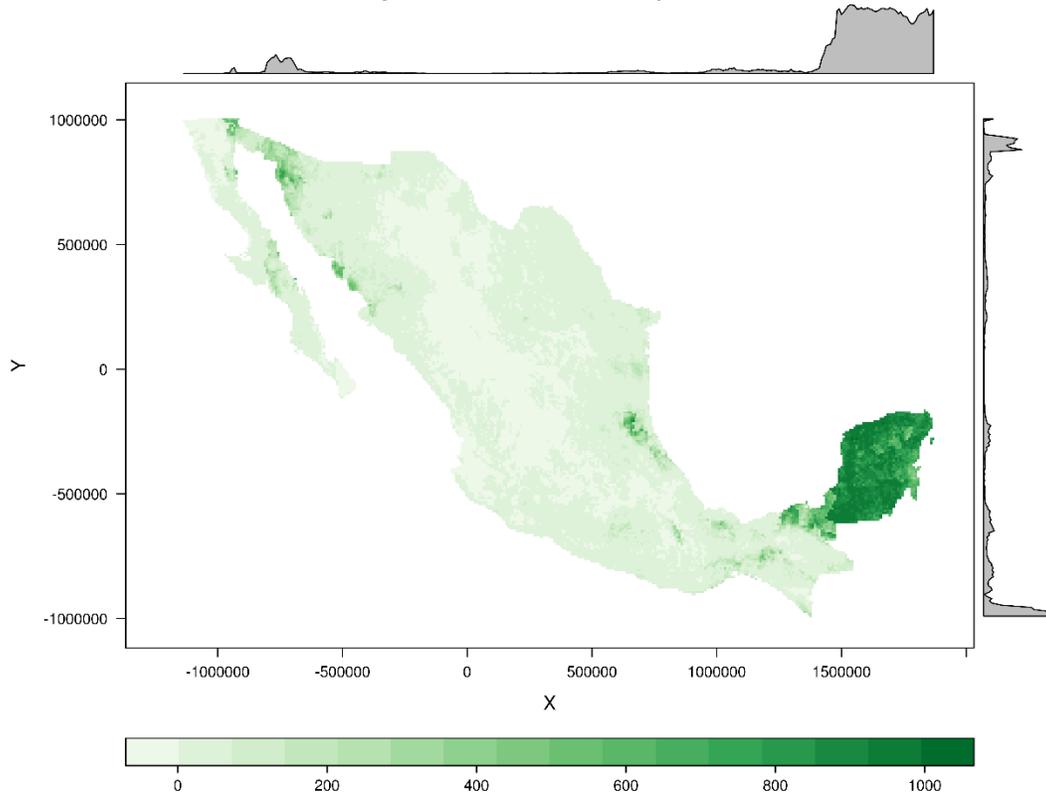




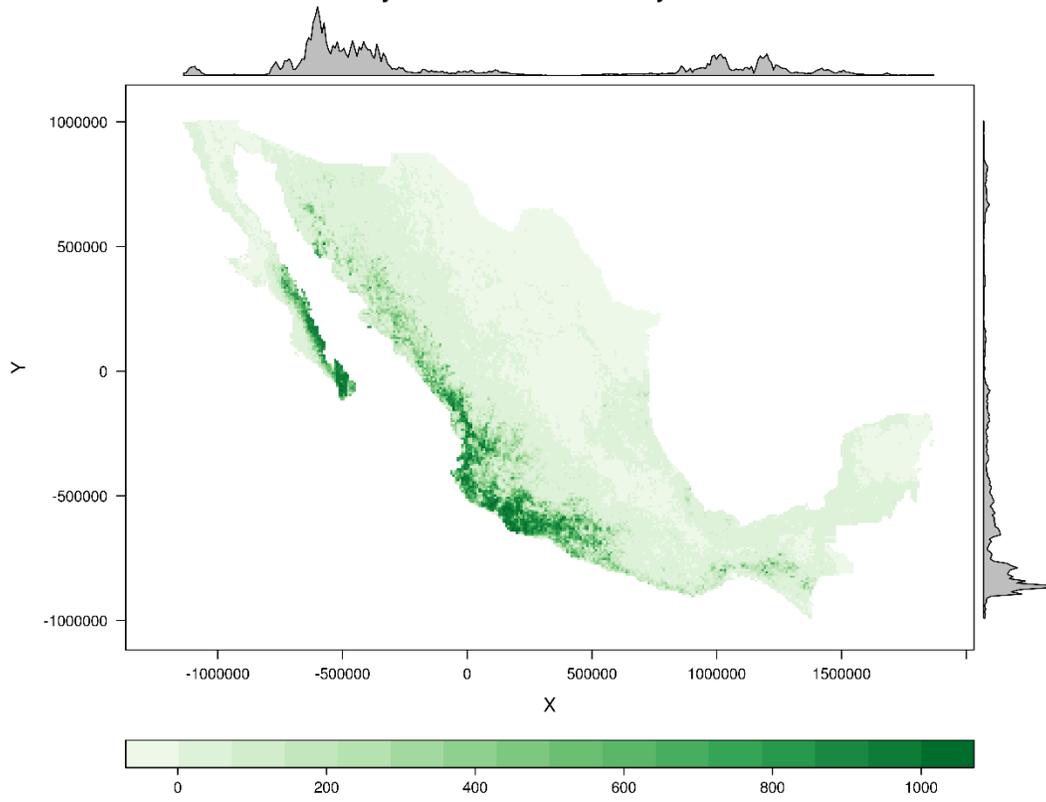
Dulcillo.del.Noroeste : 1970 - 2000
Projected Relative Probability of Presence



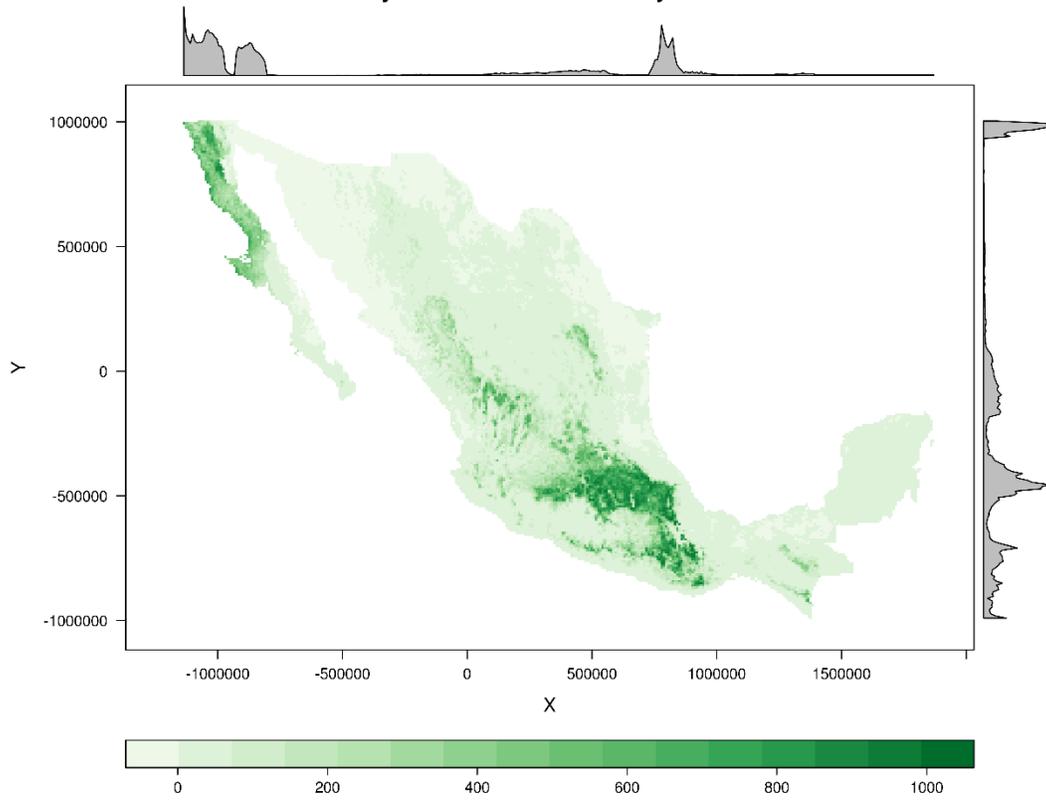
Dzit.Bacal : 1970 - 2000
Projected Relative Probability of Presence



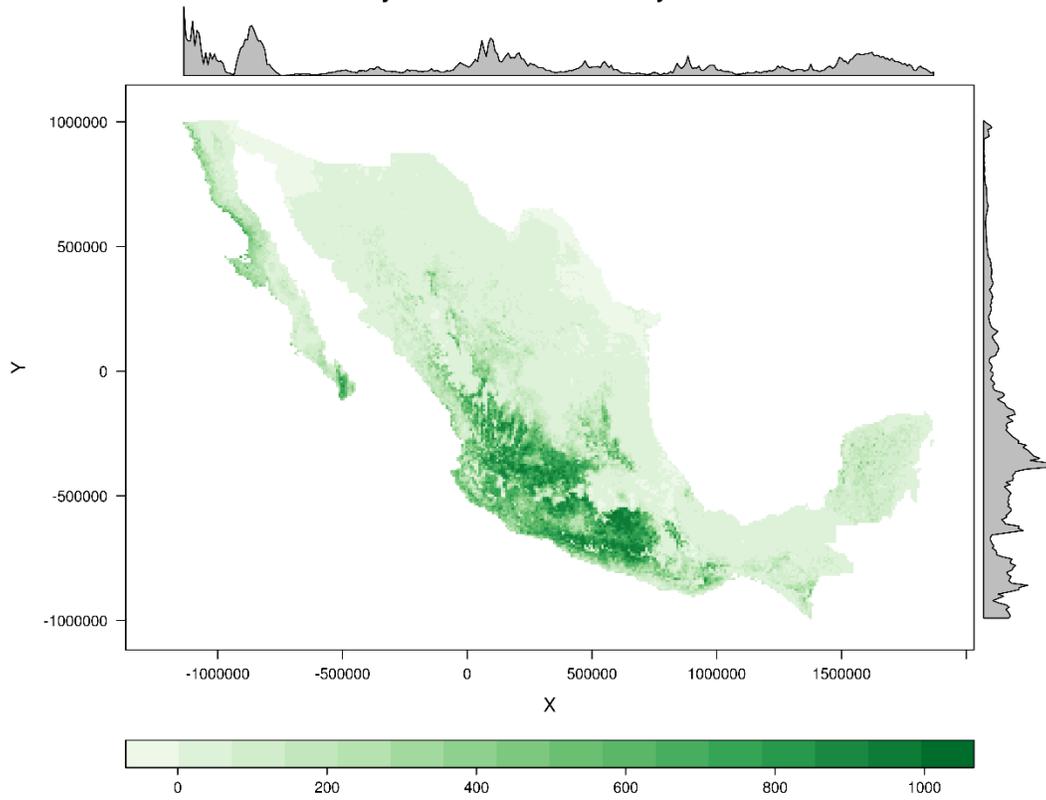
Elotero.de.Sinaloa : 1970 - 2000
Projected Relative Probability of Presence



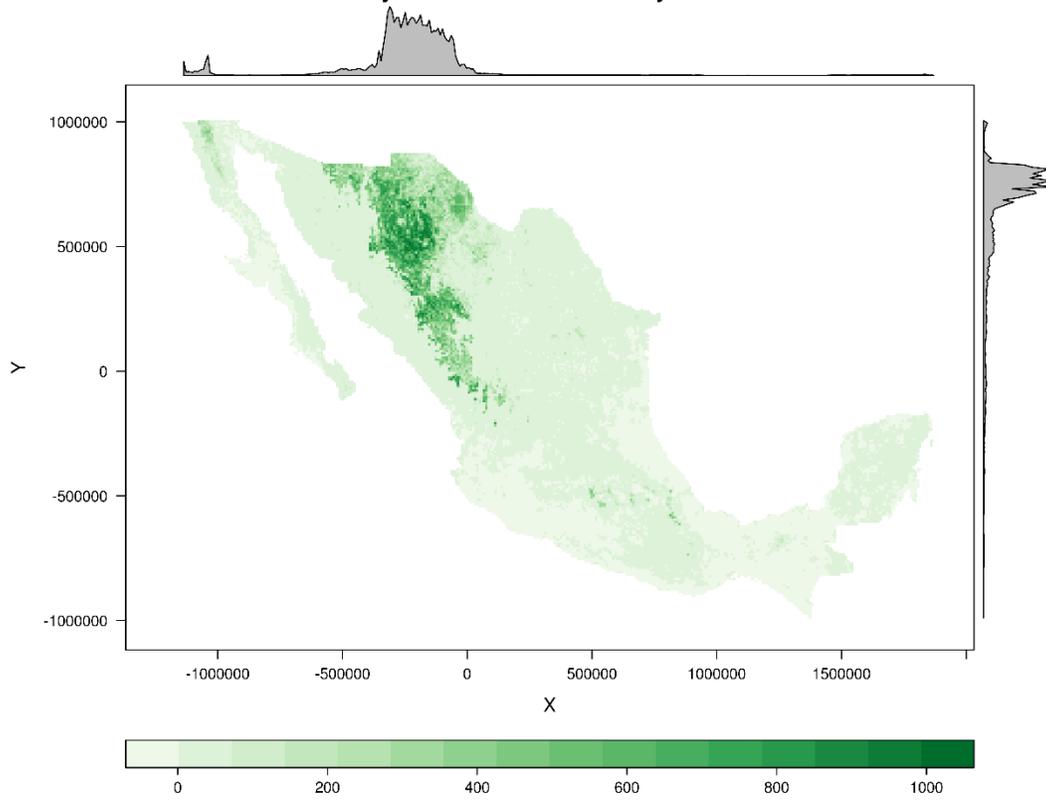
Elotes.Conicos : 1970 - 2000
Projected Relative Probability of Presence



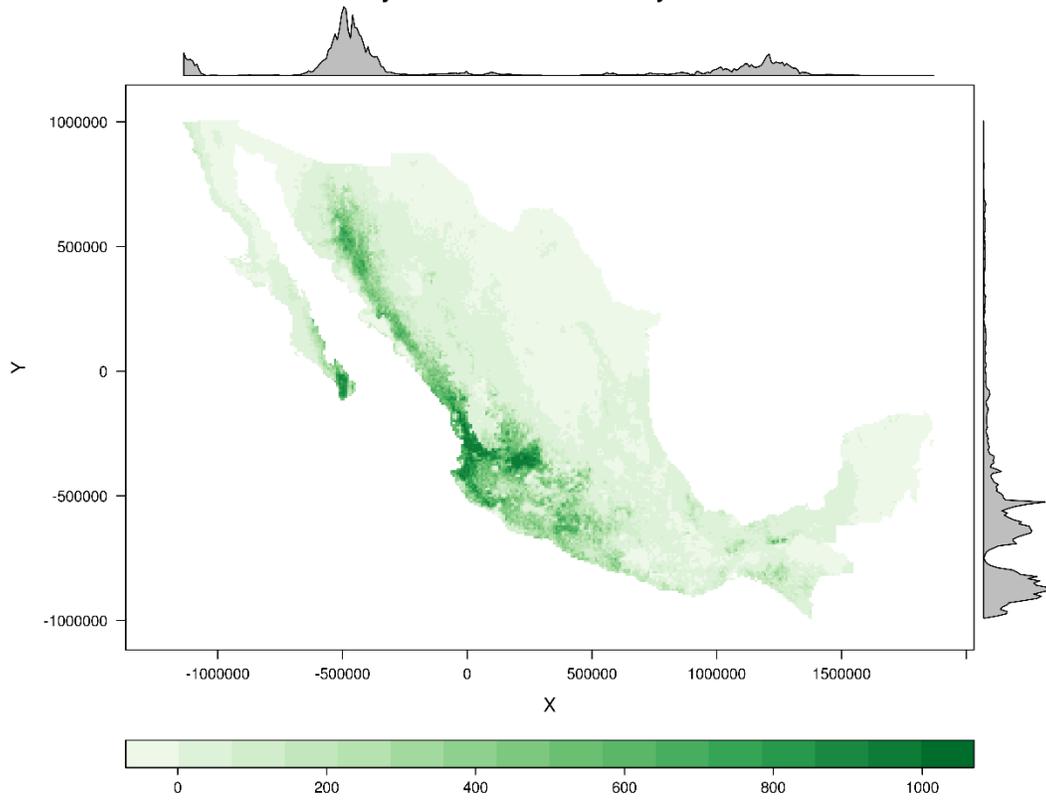
Elotes Occidentales : 1970 - 2000
Projected Relative Probability of Presence



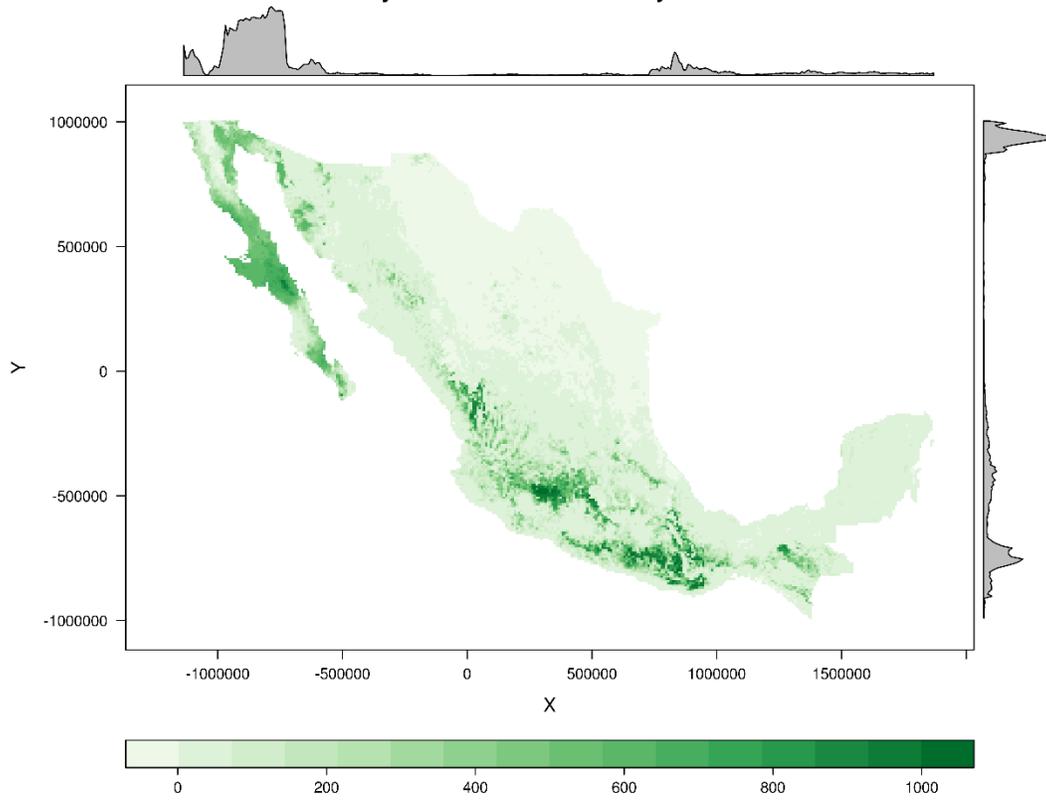
Gordo : 1970 - 2000
Projected Relative Probability of Presence



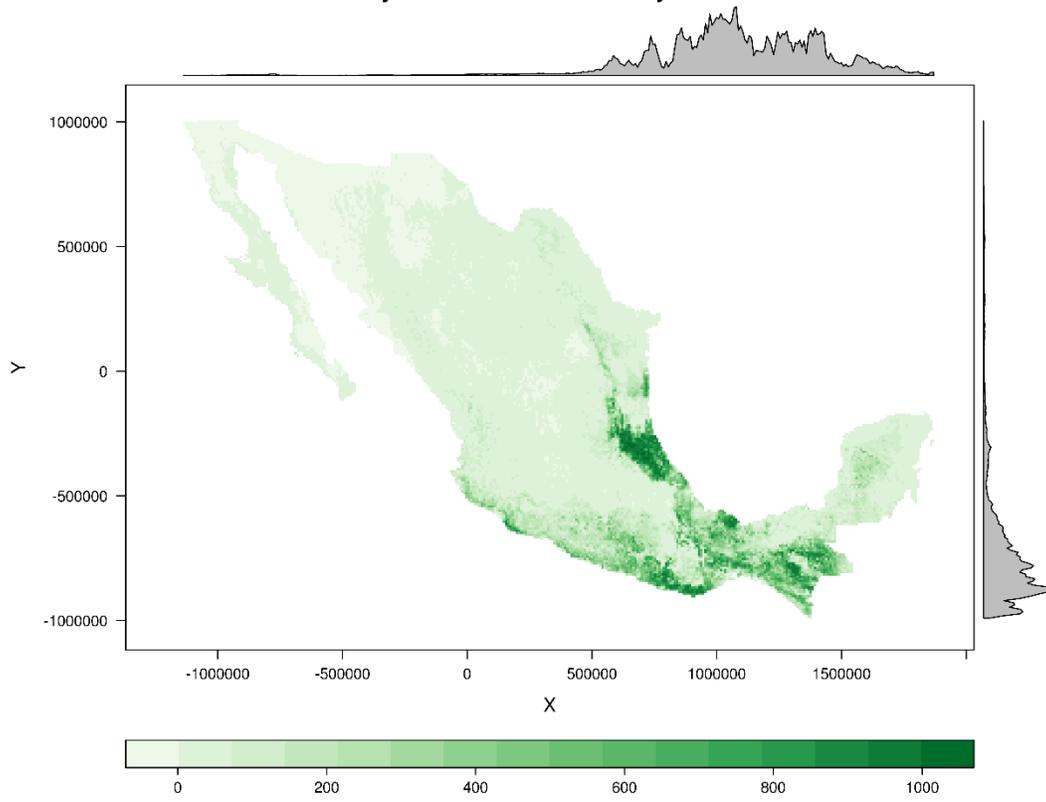
Jala : 1970 - 2000
Projected Relative Probability of Presence



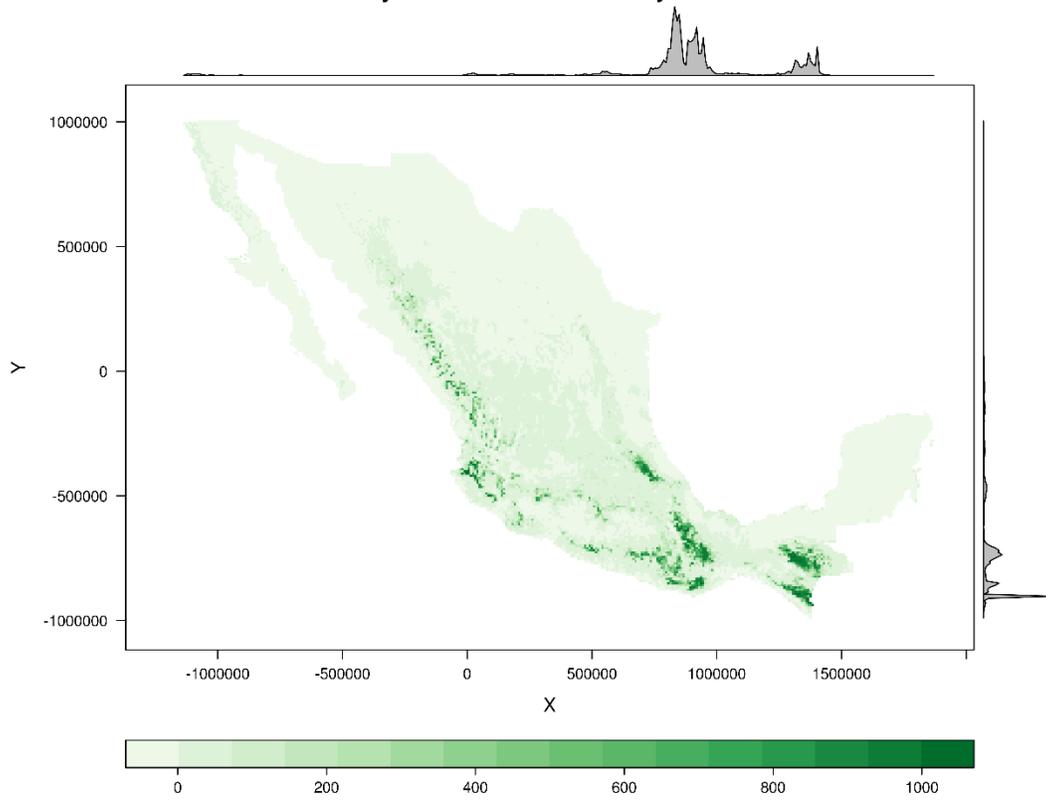
Mushito : 1970 - 2000
Projected Relative Probability of Presence



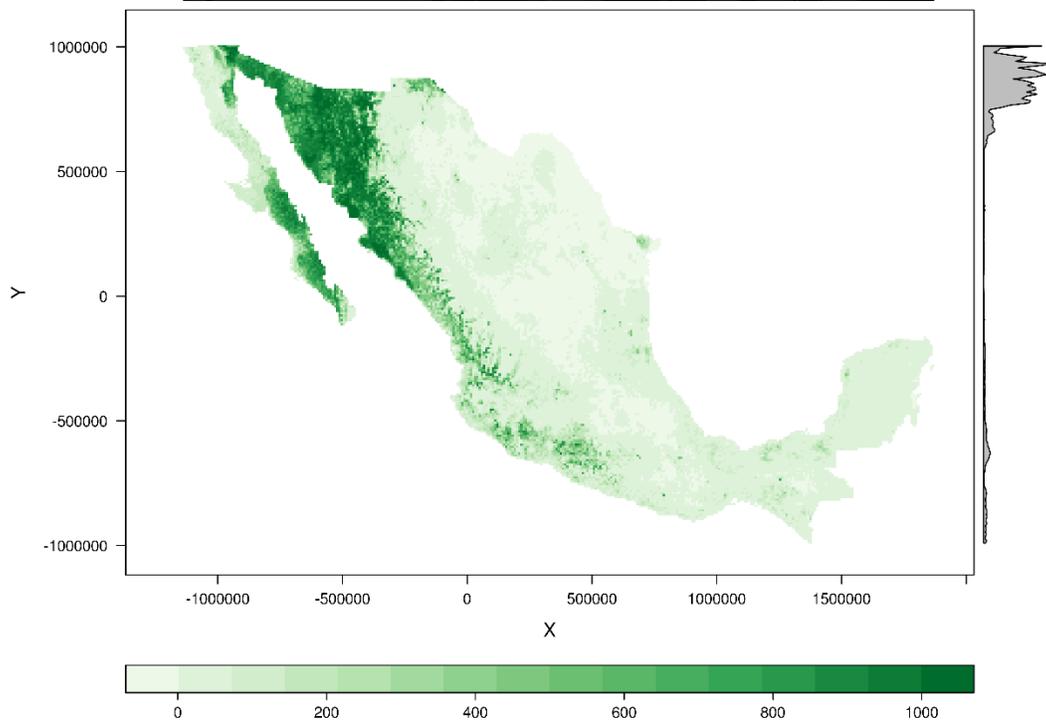
Olotillo : 1970 - 2000
Projected Relative Probability of Presence



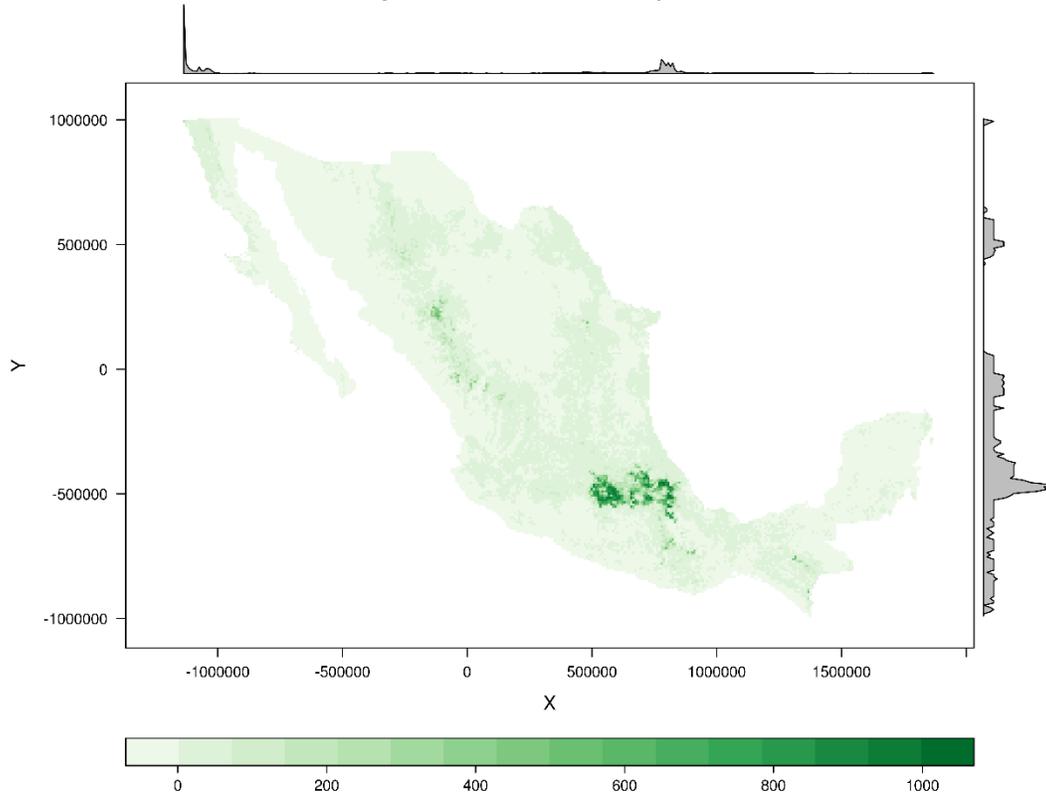
**Olton : 1970 - 2000
Projected Relative Probability of Presence**



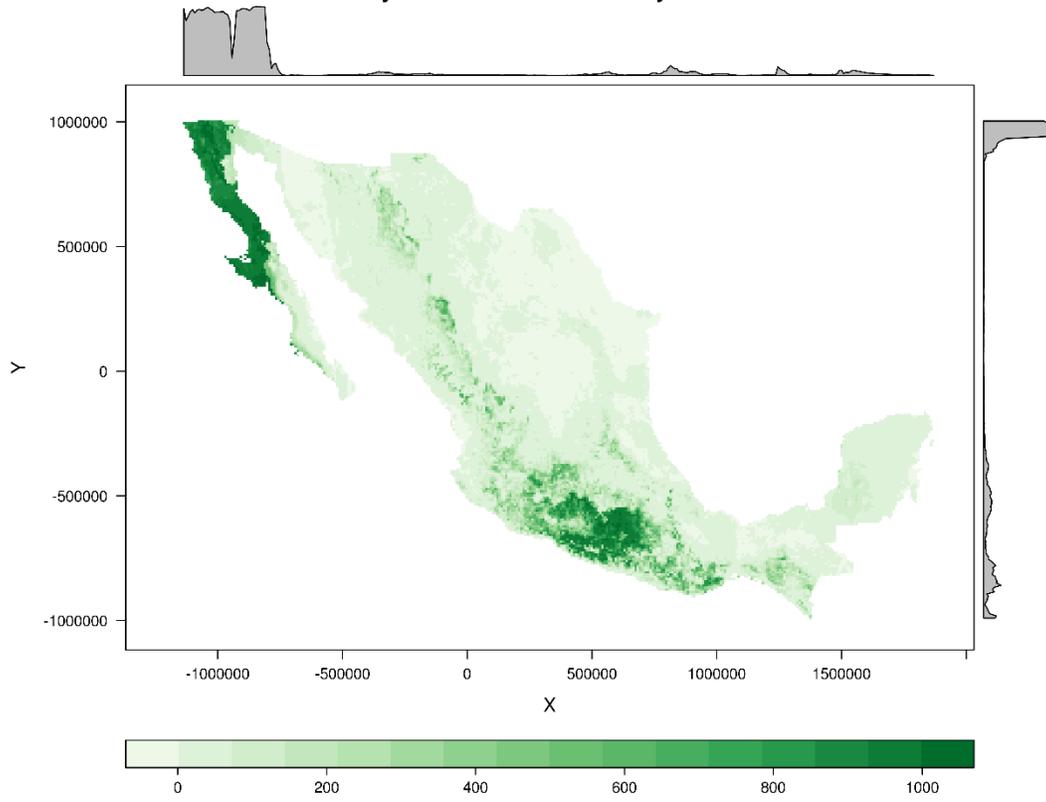
Onaveno : 1970 - 2000
Projected Relative Probability of Presence



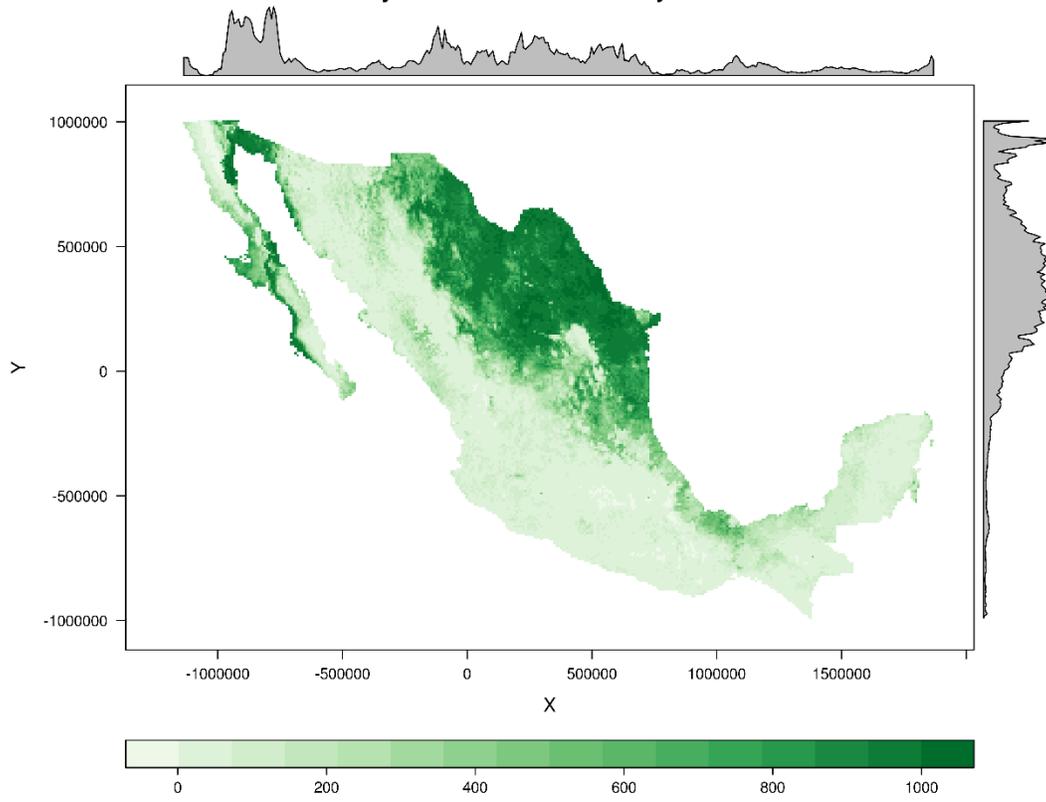
Palomero.Toluqueno : 1970 - 2000
Projected Relative Probability of Presence



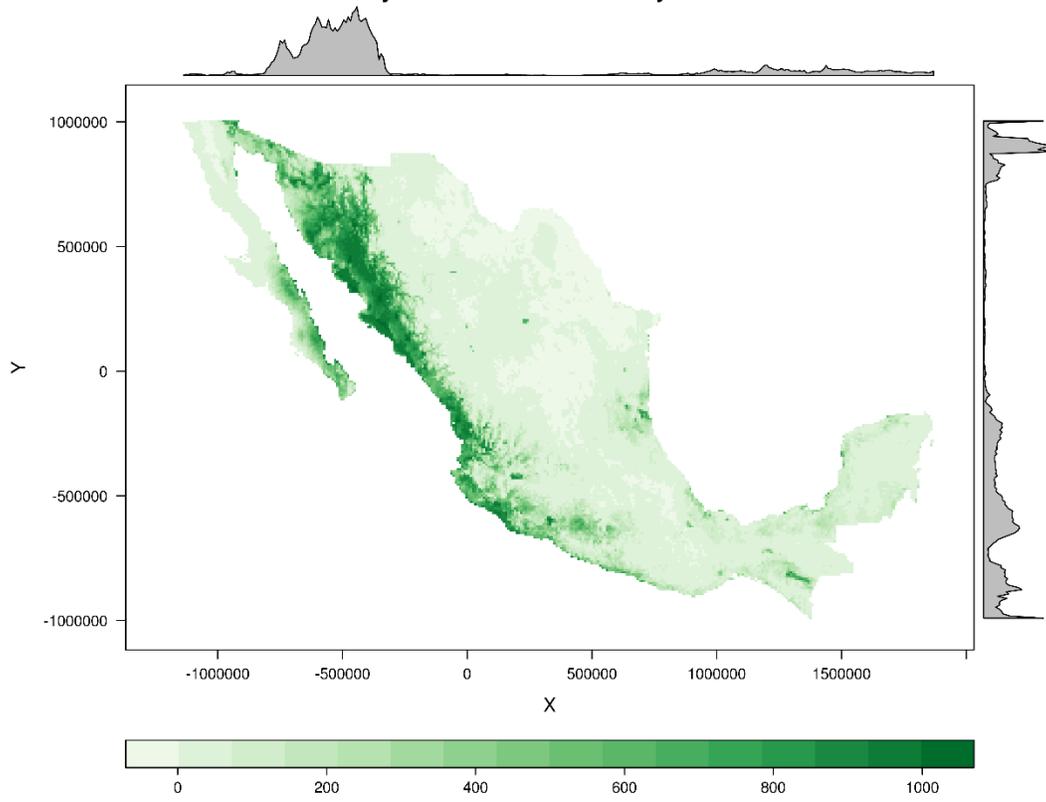
Pepitilla : 1970 - 2000
Projected Relative Probability of Presence



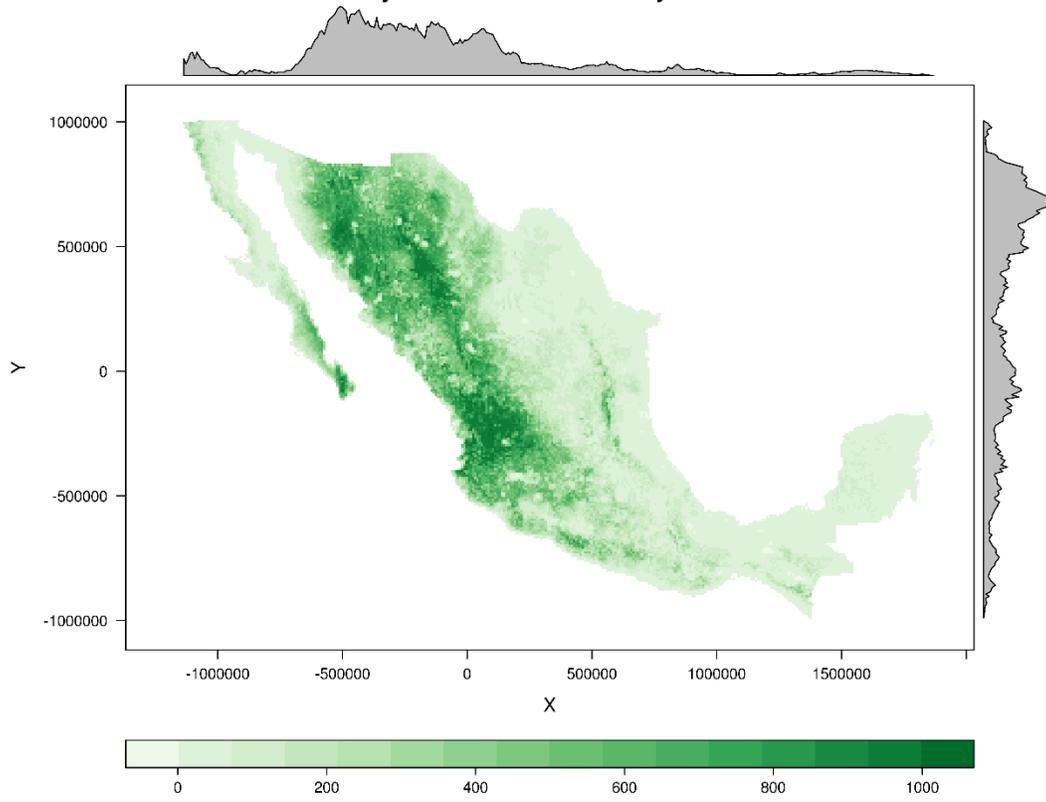
Raton : 1970 - 2000
Projected Relative Probability of Presence



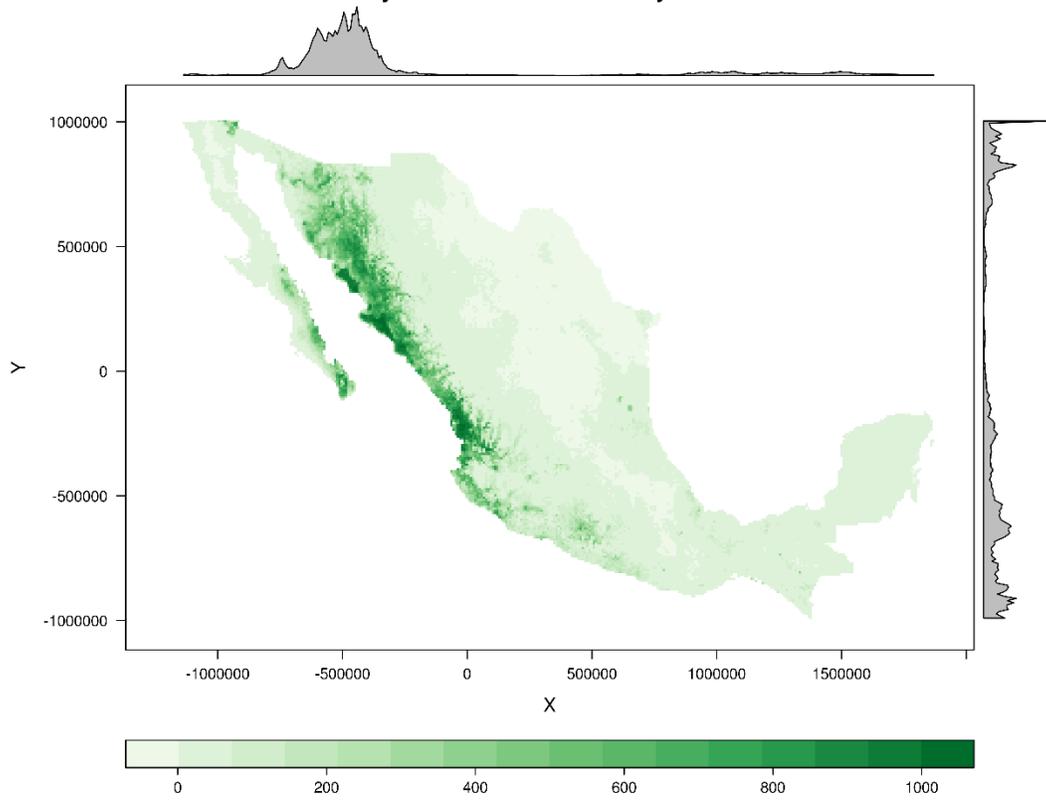
Reventador : 1970 - 2000
Projected Relative Probability of Presence



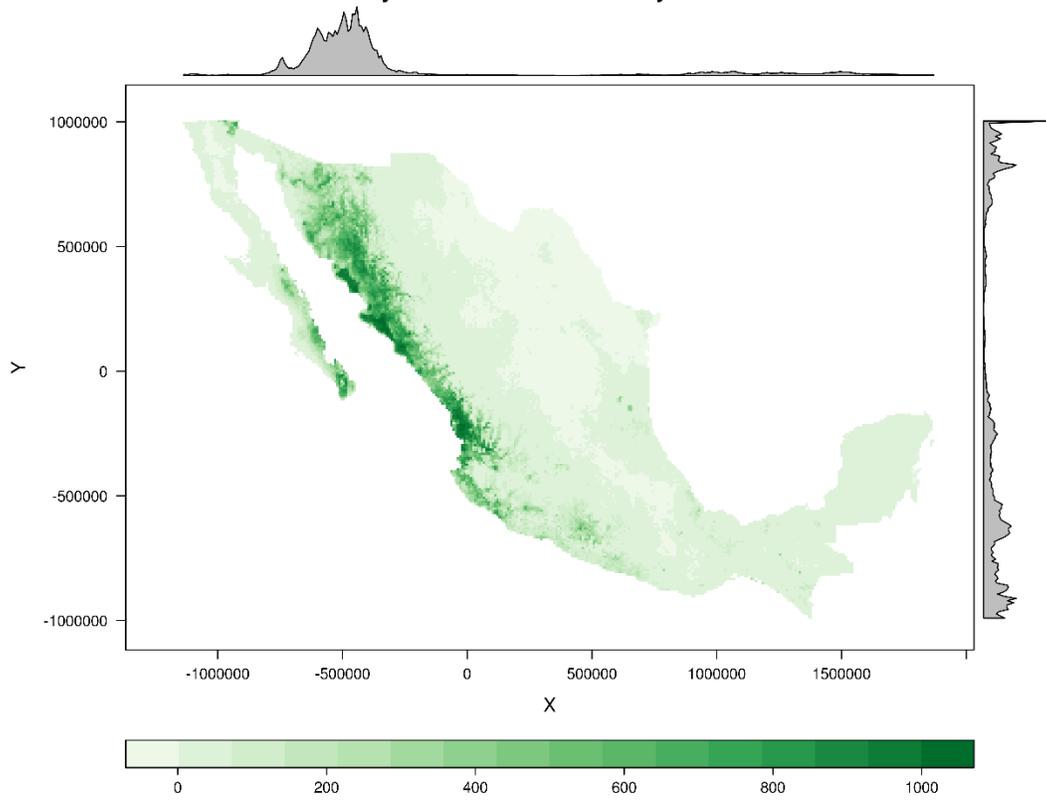
Tablilla.de.Ocho : 1970 - 2000
Projected Relative Probability of Presence



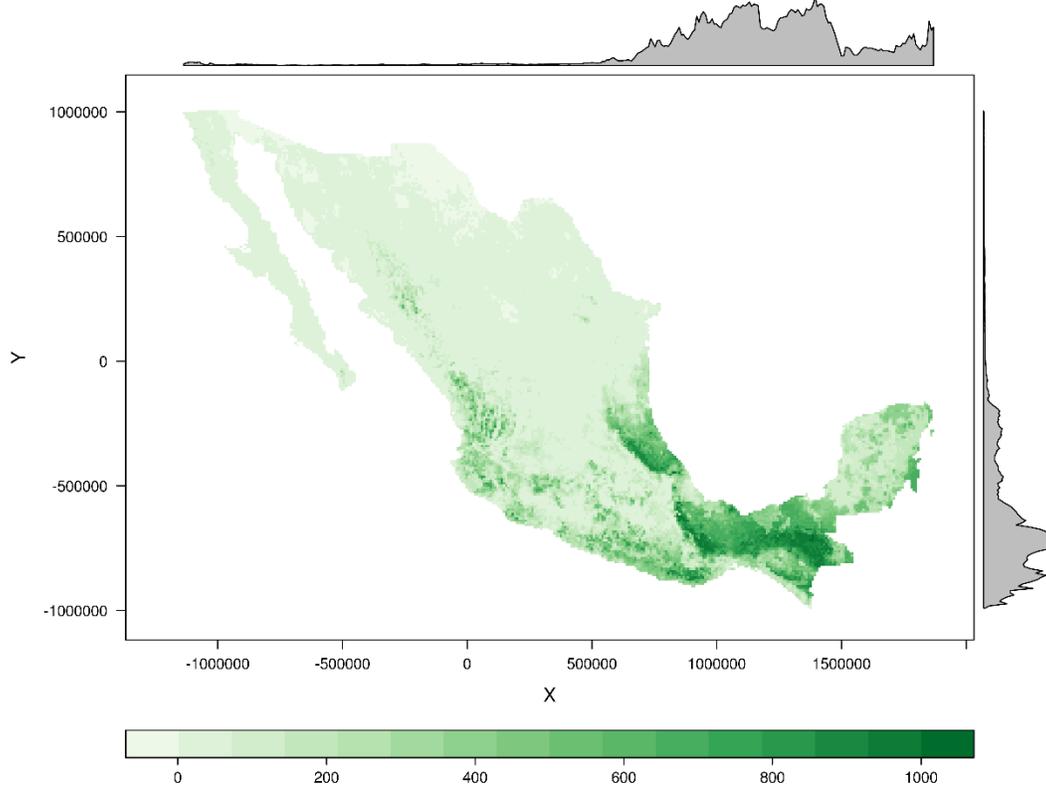
Tabloncillo.Perla : 1970 - 2000
Projected Relative Probability of Presence



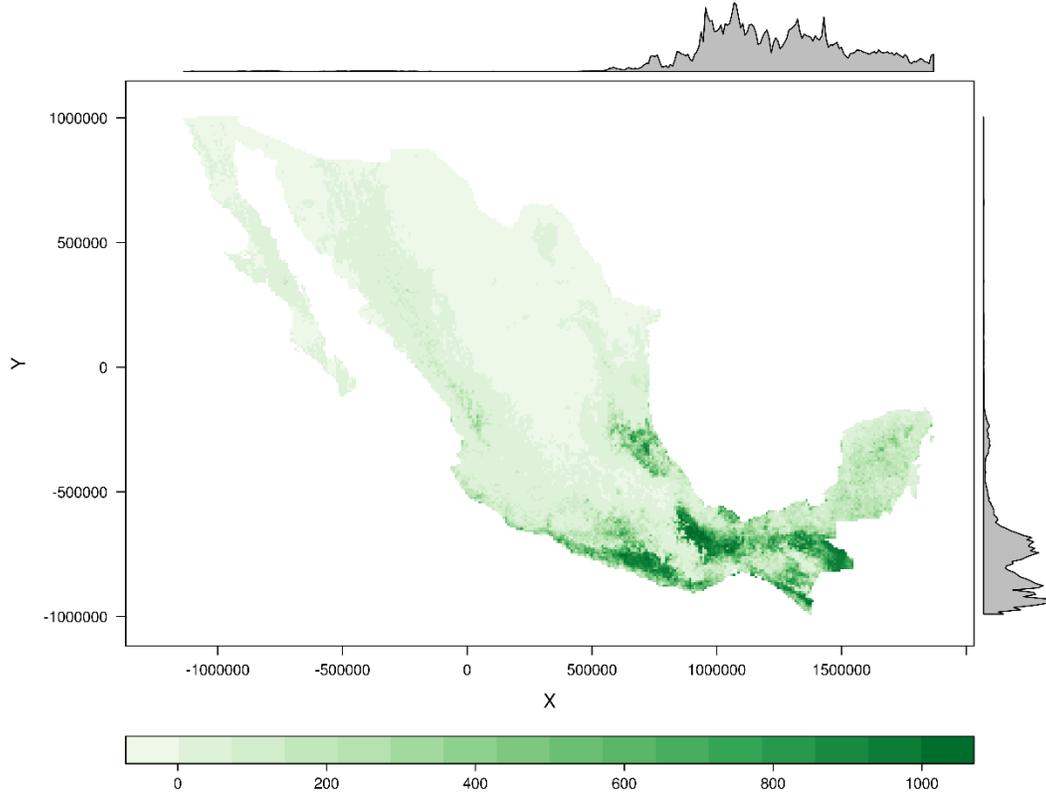
Tabloncillo : 1970 - 2000
Projected Relative Probability of Presence



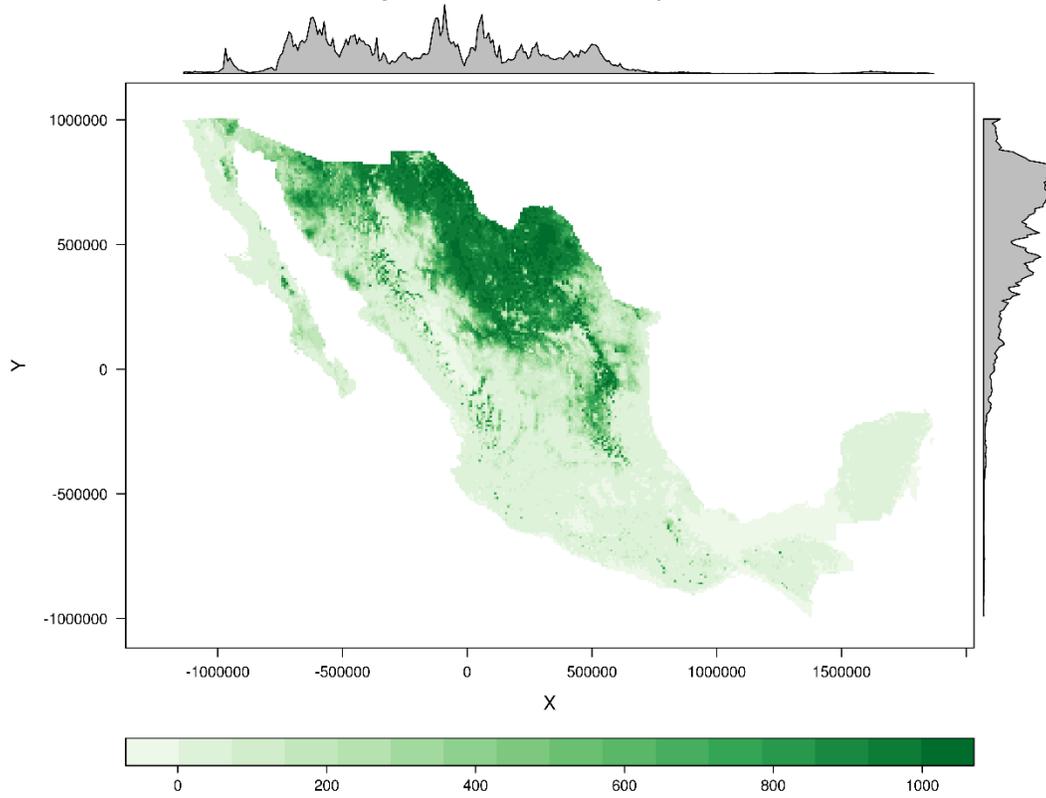
Tehua : 1970 - 2000
Projected Relative Probability of Presence



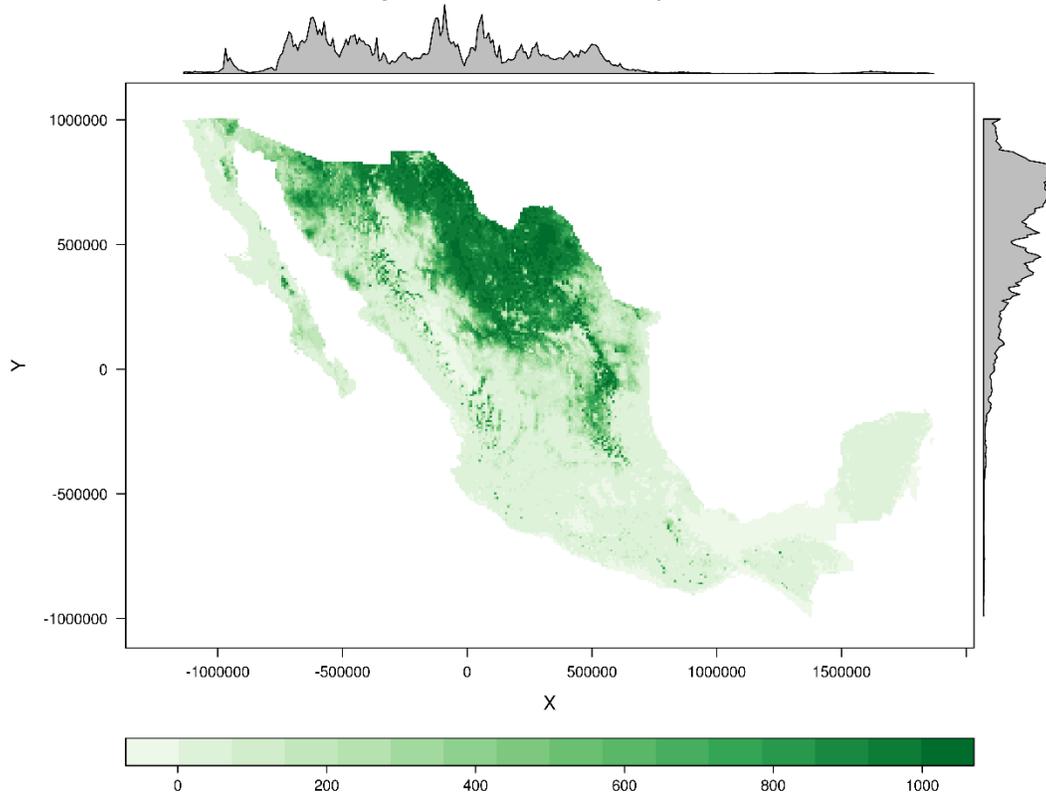
**Tepecintle : 1970 - 2000
Projected Relative Probability of Presence**



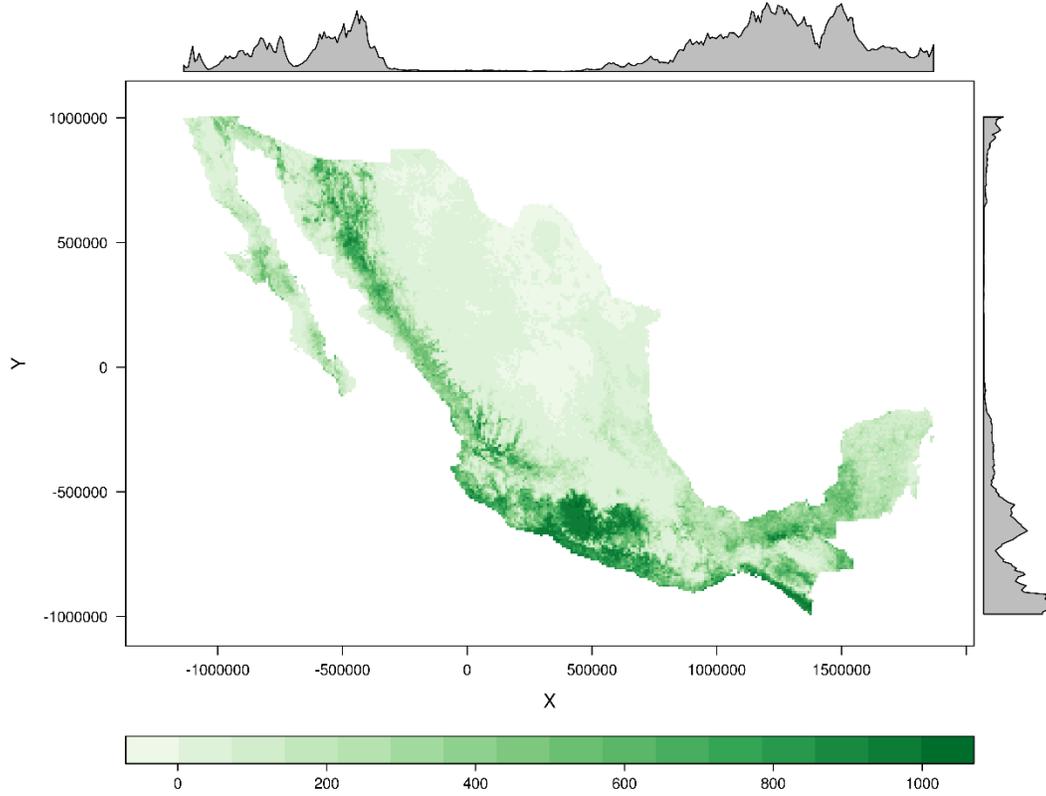
Tuxpeno.Norteno : 1970 - 2000
Projected Relative Probability of Presence



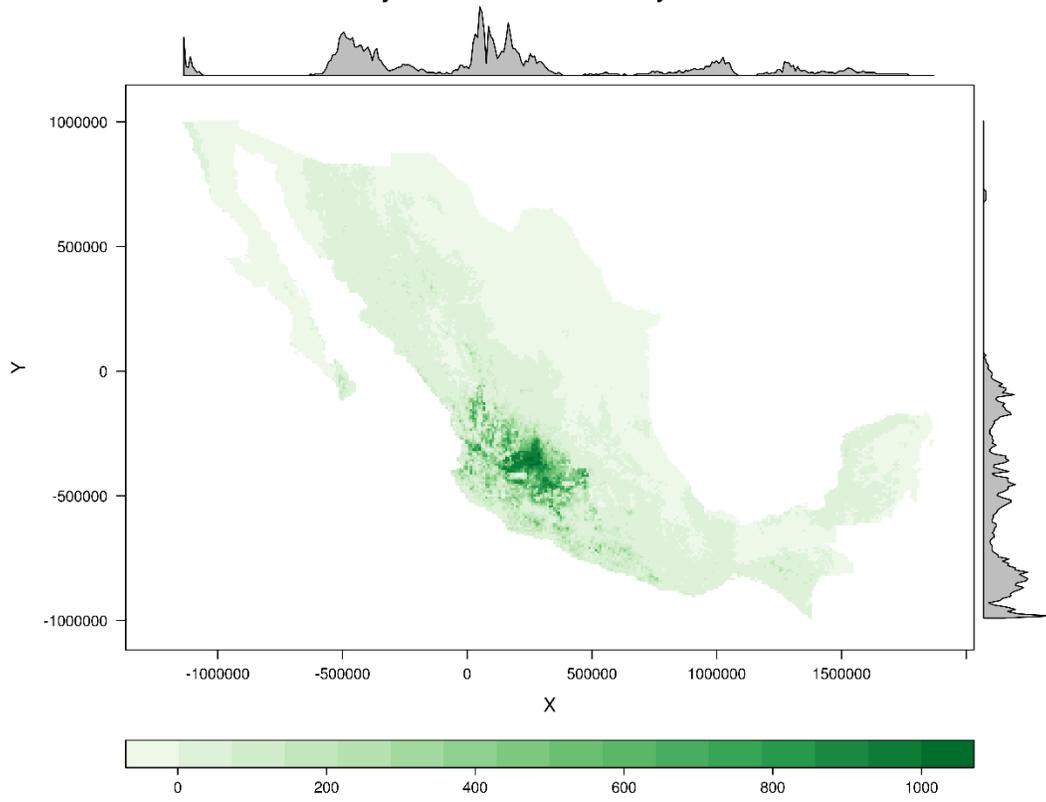
Tuxpeno : 1970 - 2000
Projected Relative Probability of Presence



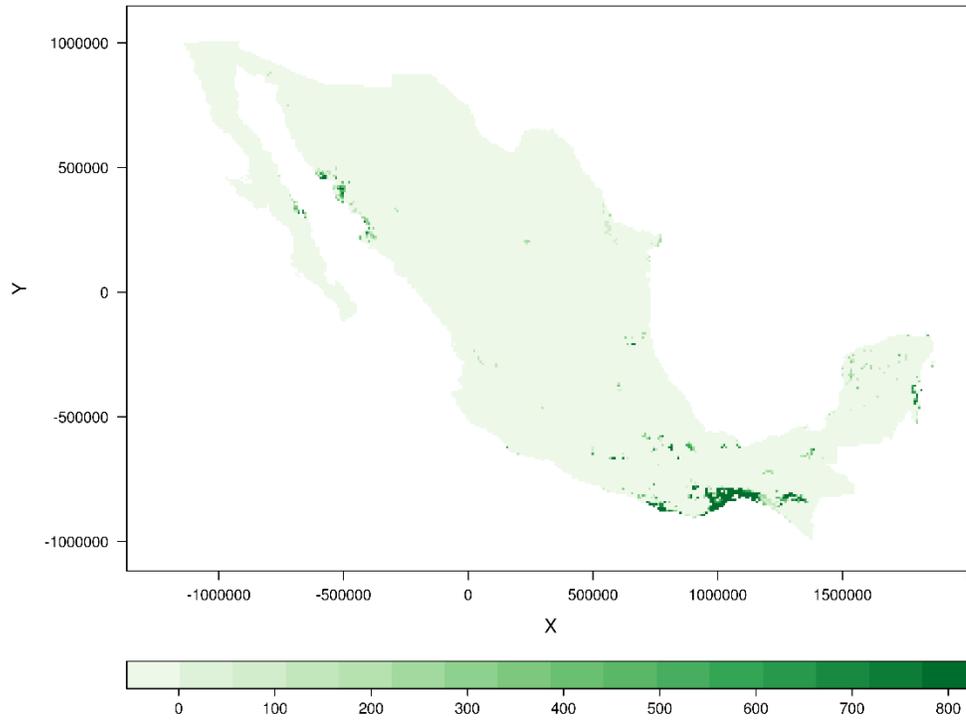
Vandeno : 1970 - 2000
Projected Relative Probability of Presence



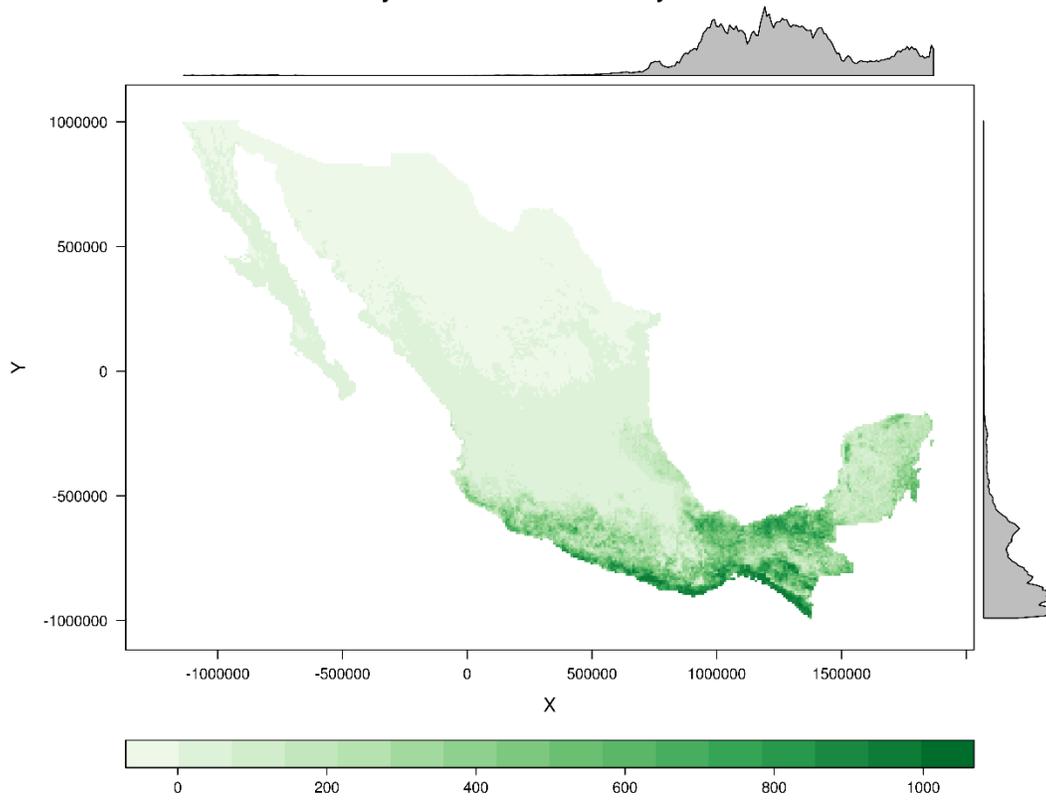
Zamorano.Amarillo : 1970 - 2000
Projected Relative Probability of Presence



Zapalote.Chico : 1970 - 2000
Projected Relative Probability of Presence



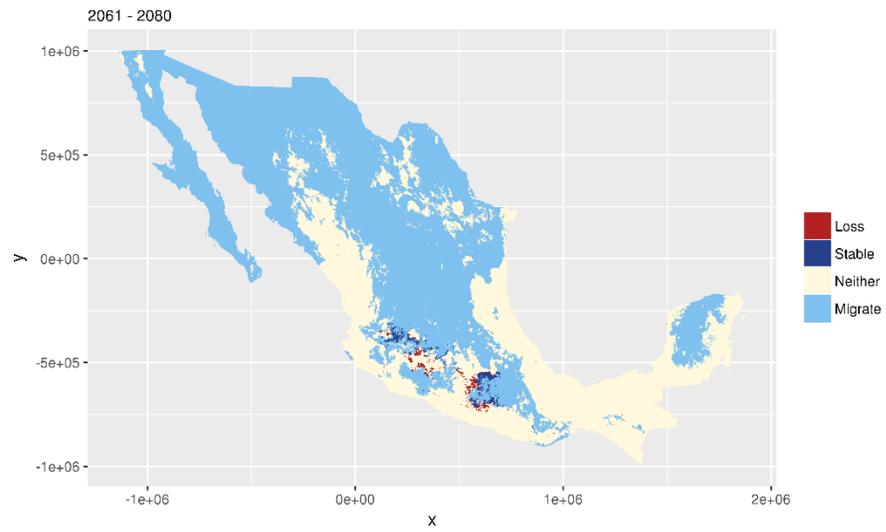
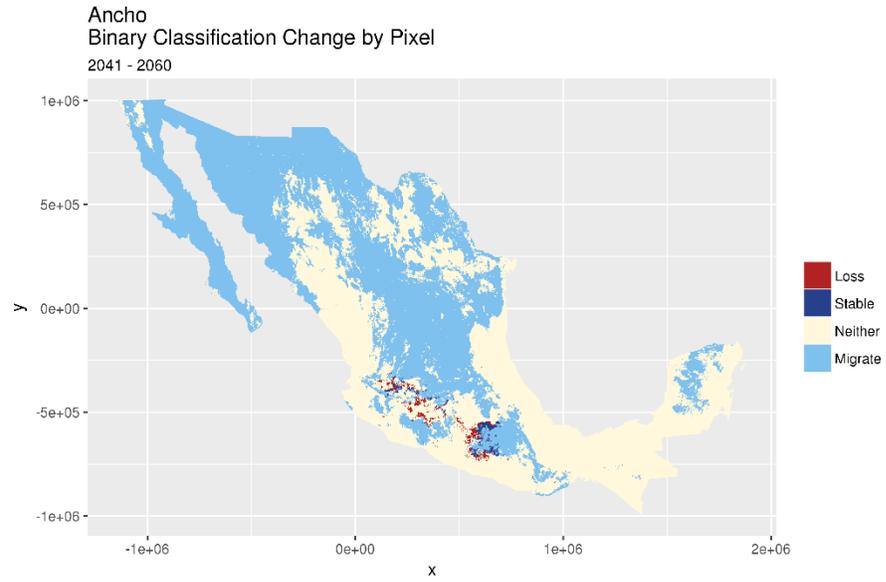
Zapalote.Grande : 1970 - 2000
Projected Relative Probability of Presence

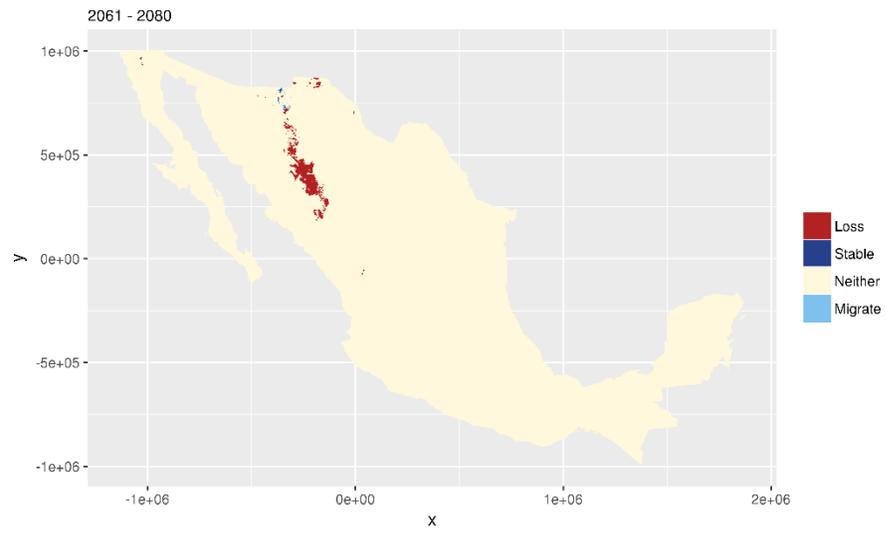
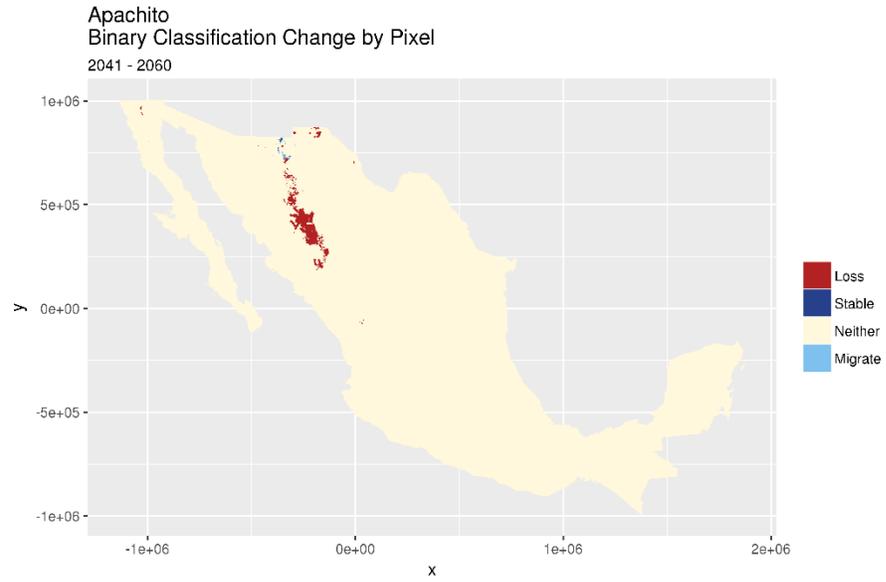


C. Projected Range Change by Landrace

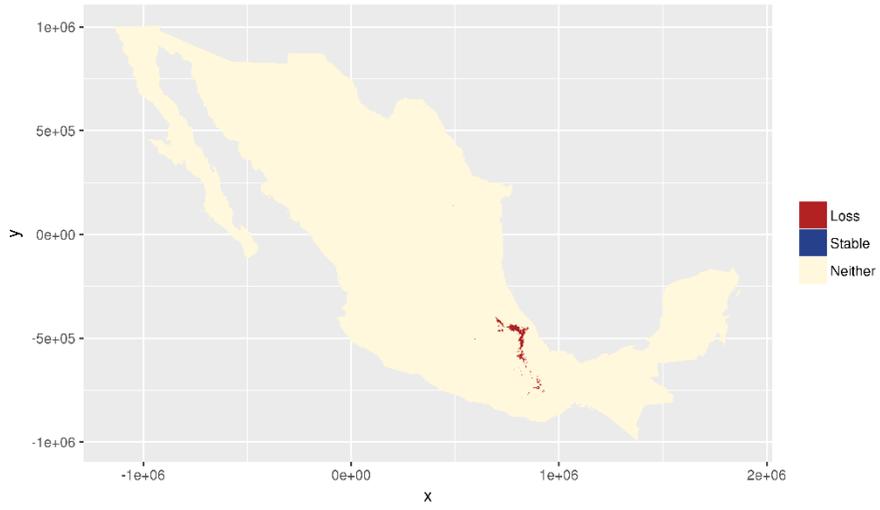
The following maps depicted the projected range changes calculated by using the binary forecasts of each maize landrace in future time-periods compared with 1970 – 2000 binary forecasts. Each map is classified into four discrete categories: Loss, Neither, Migrate or Stable.

‘Loss’ indicates pixels in which the landrace was forecasted to exist in the 1970 – 2000 time-period, but is not forecasted to exist in the future time-period. ‘Neither’ reflects pixels in which neither future or reference time-periods had the landrace forecasted to be present. ‘Migrate’ indicates that a certain pixel did not forecast the landrace to exist in the reference time-period, but is forecasted to migrate to in future time-periods. ‘Stable’ indicates that a pixel is forecasted to have species exist in both future and reference time-periods.

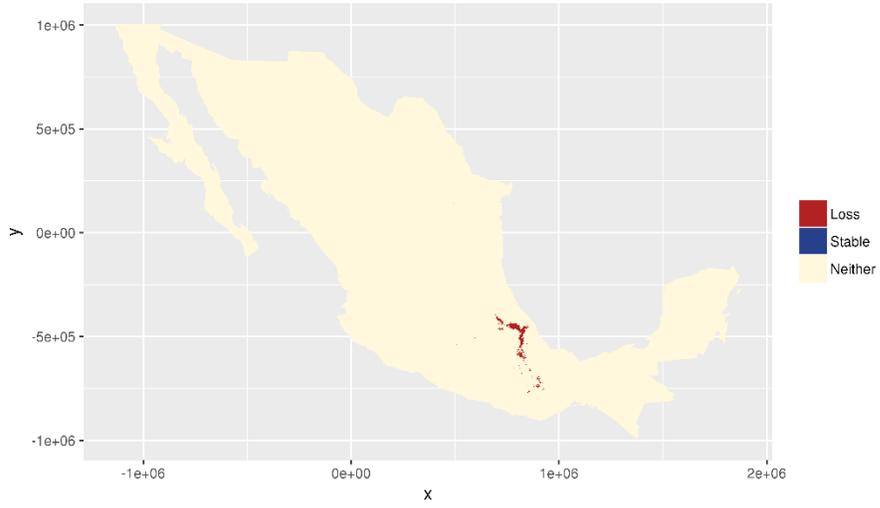


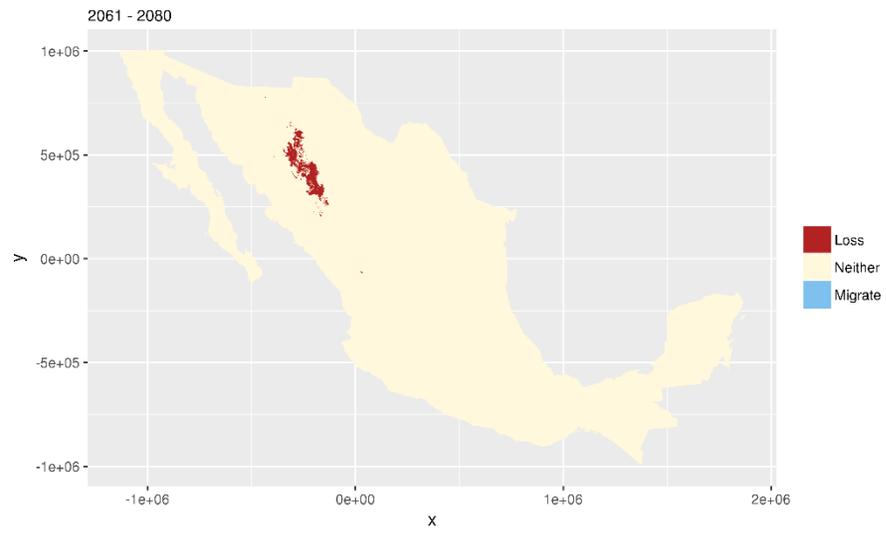
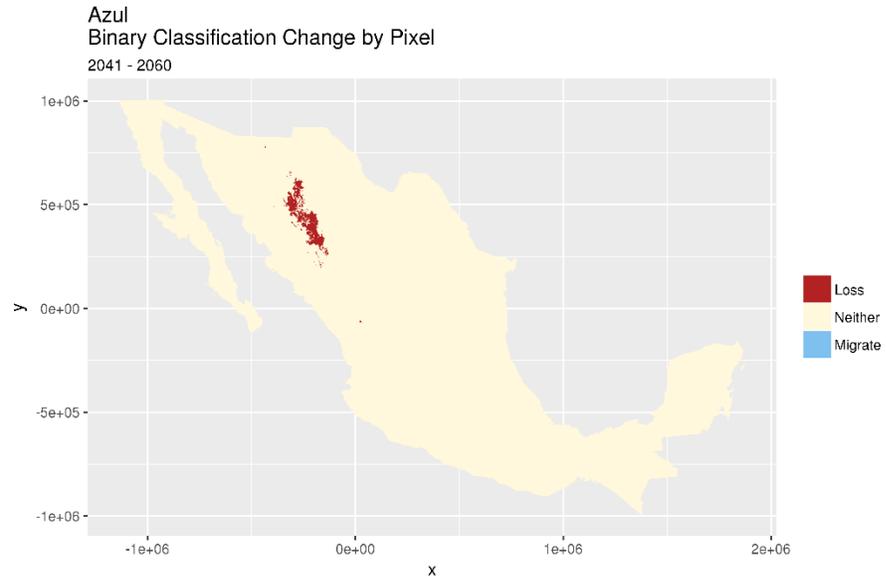


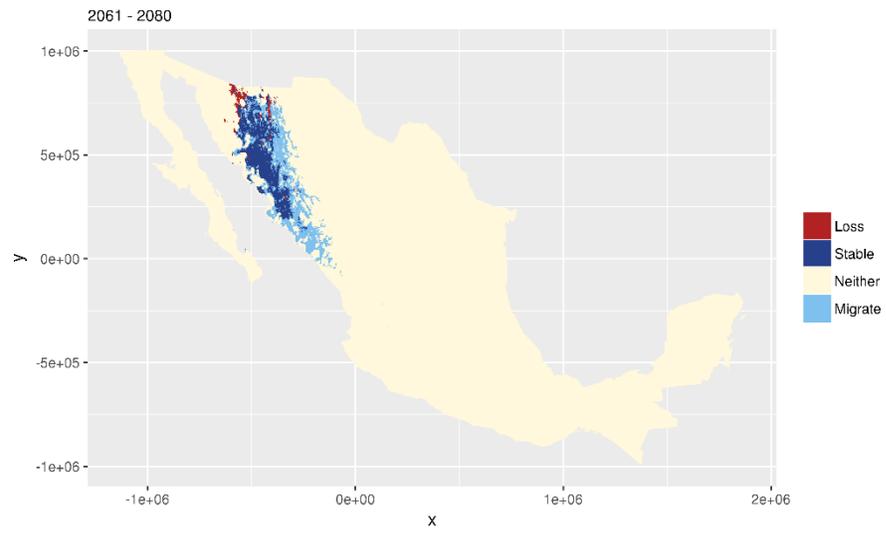
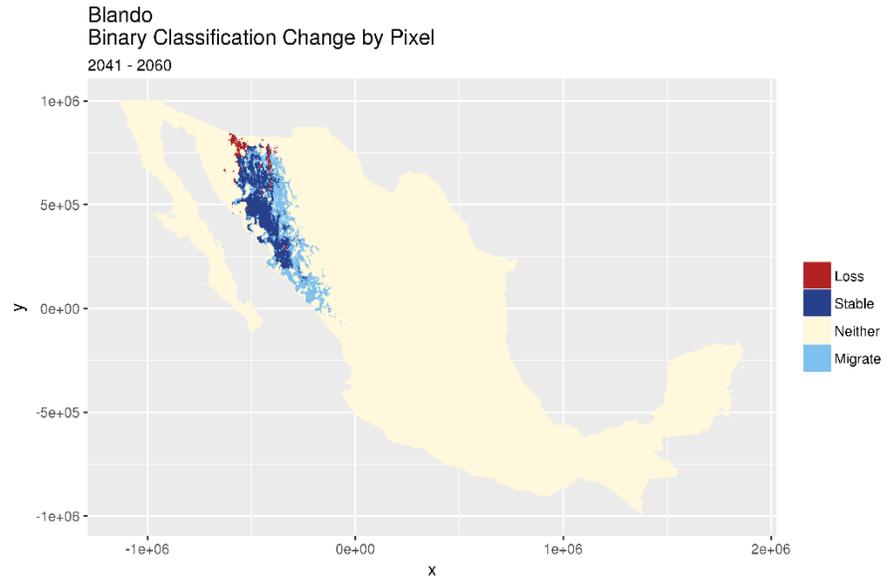
Arrocillo.Amarillo
Binary Classification Change by Pixel
2041 - 2060

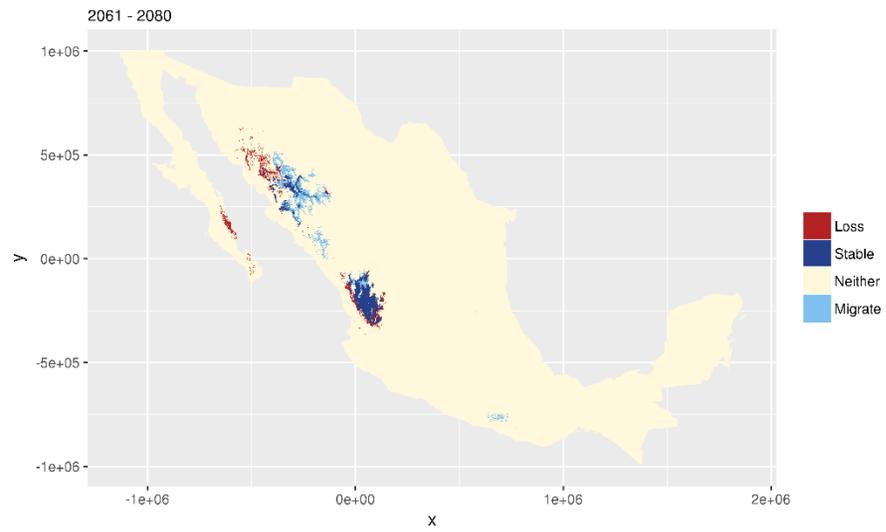
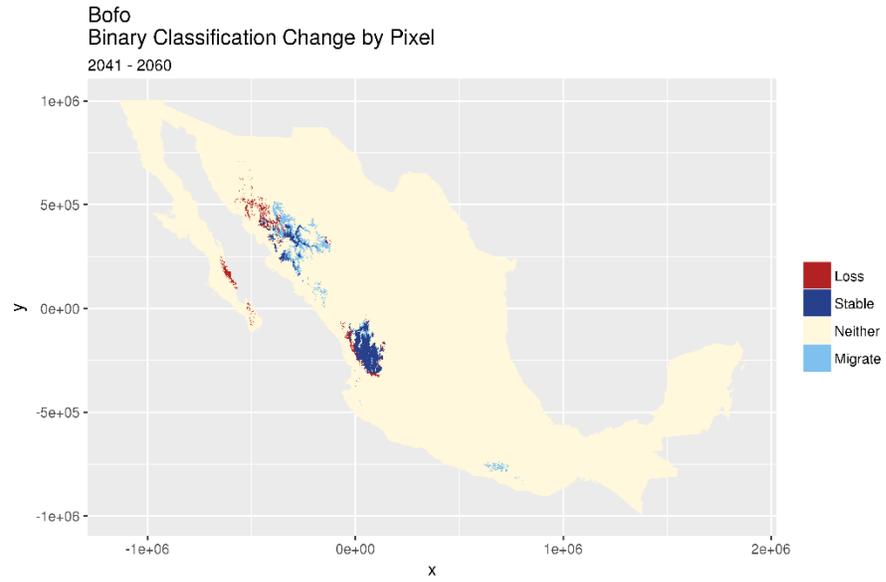


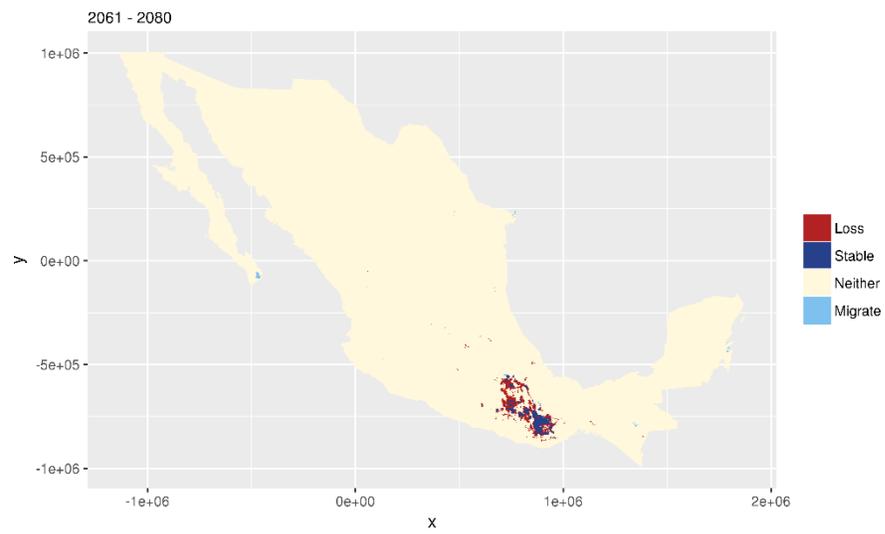
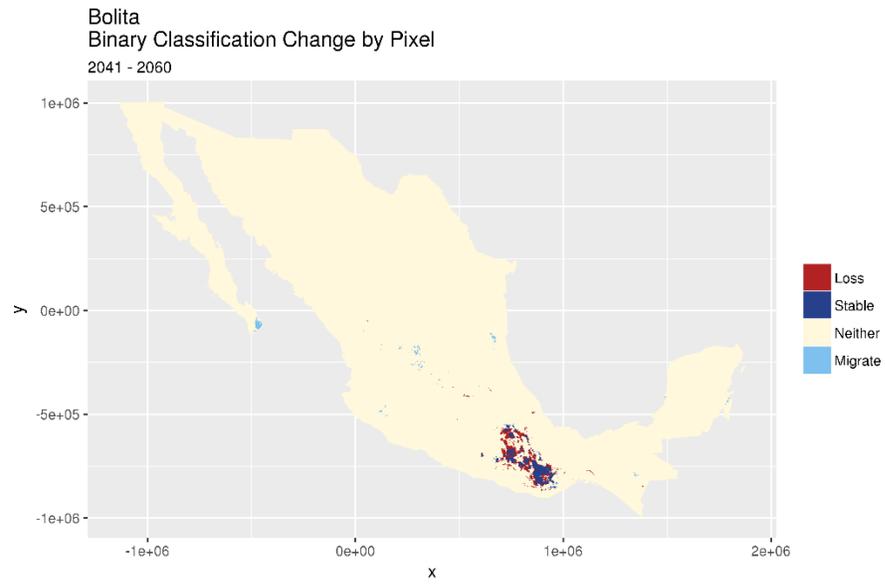
2061 - 2080

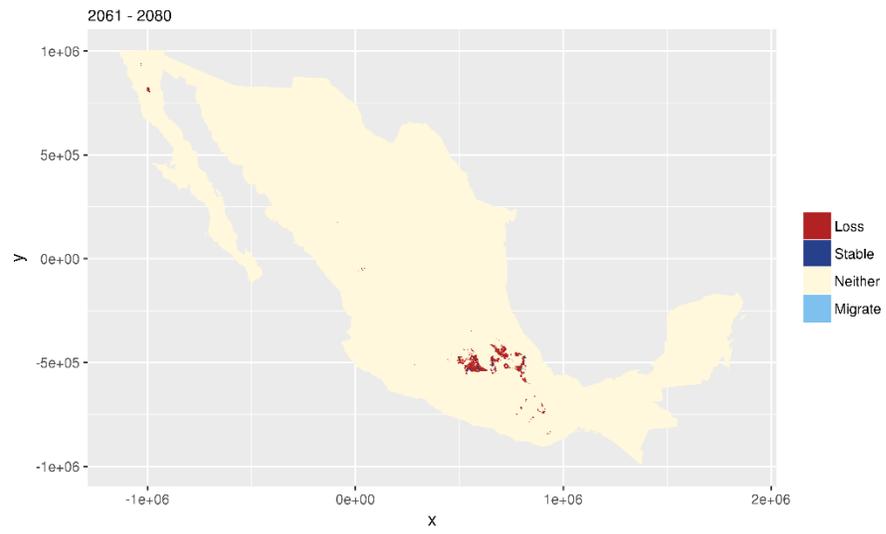
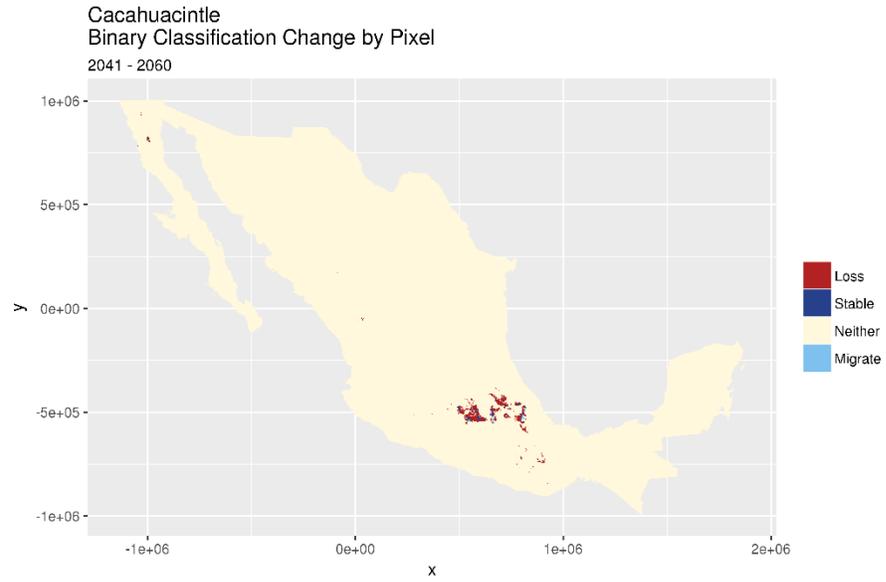


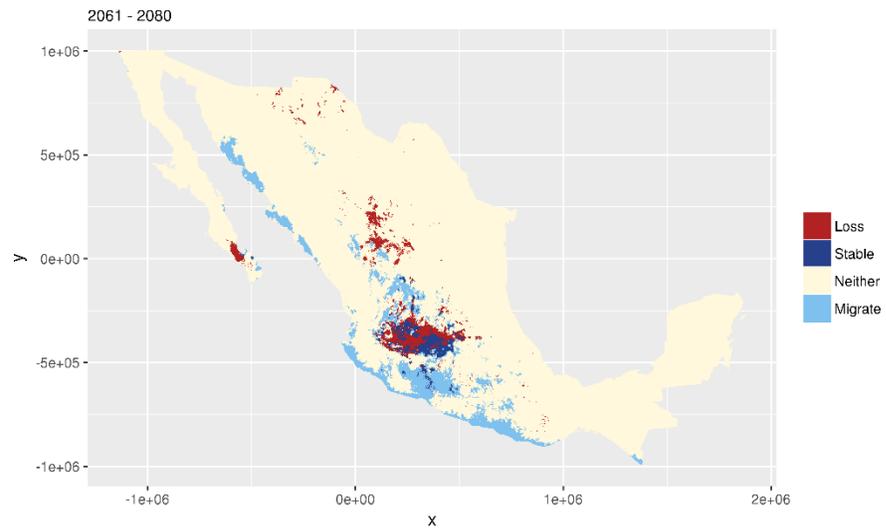
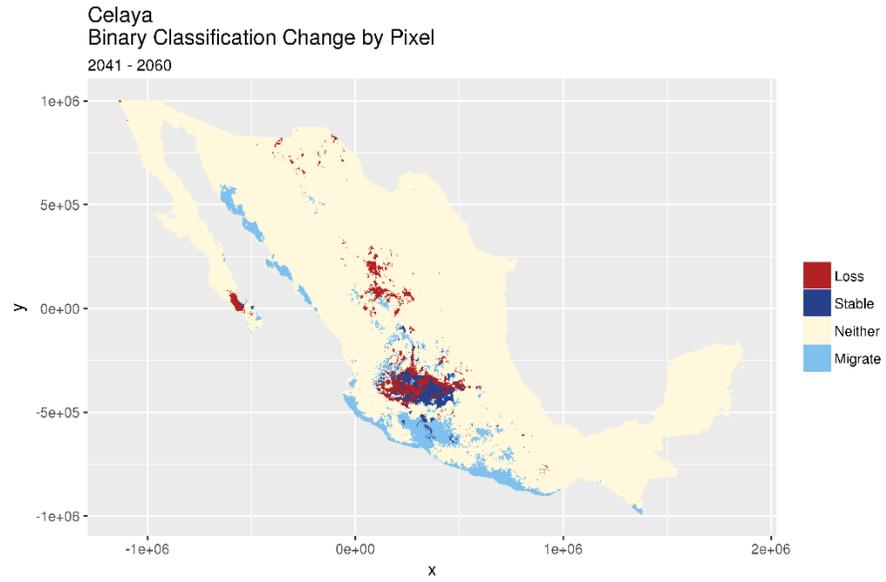


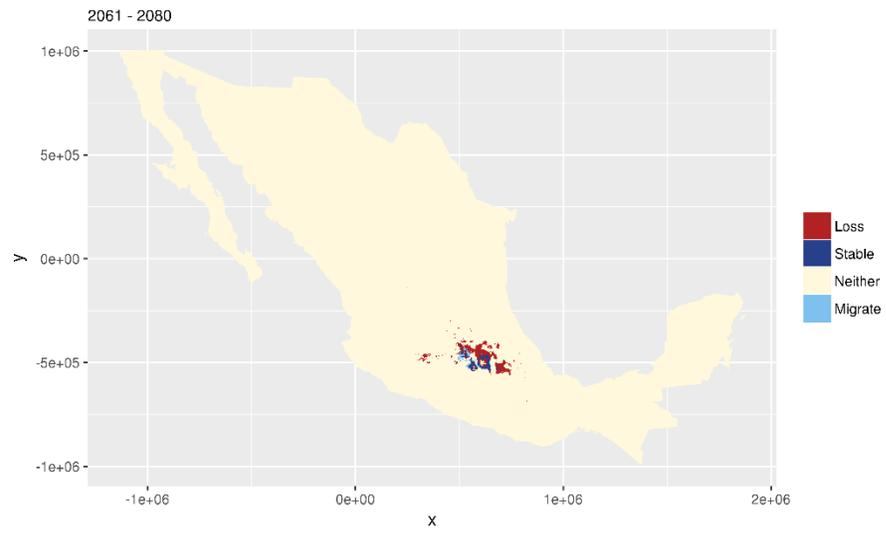
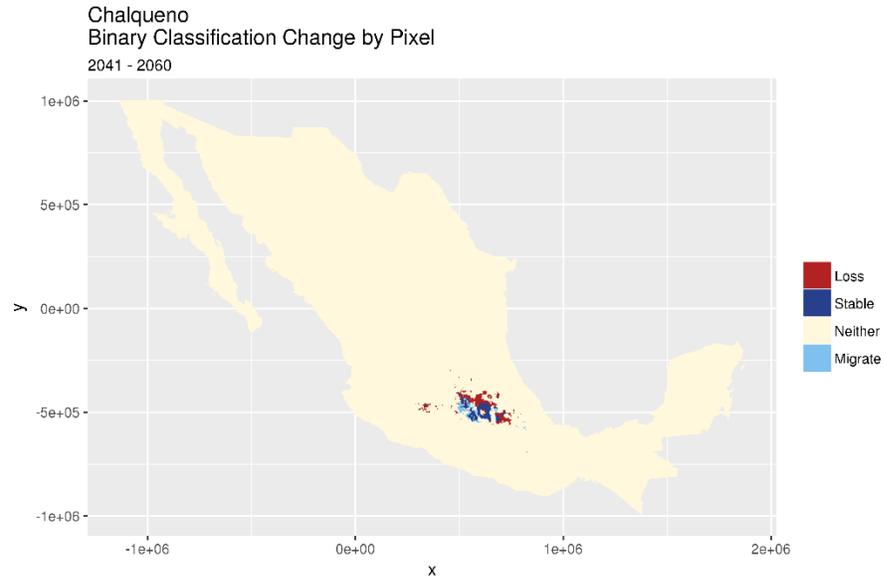


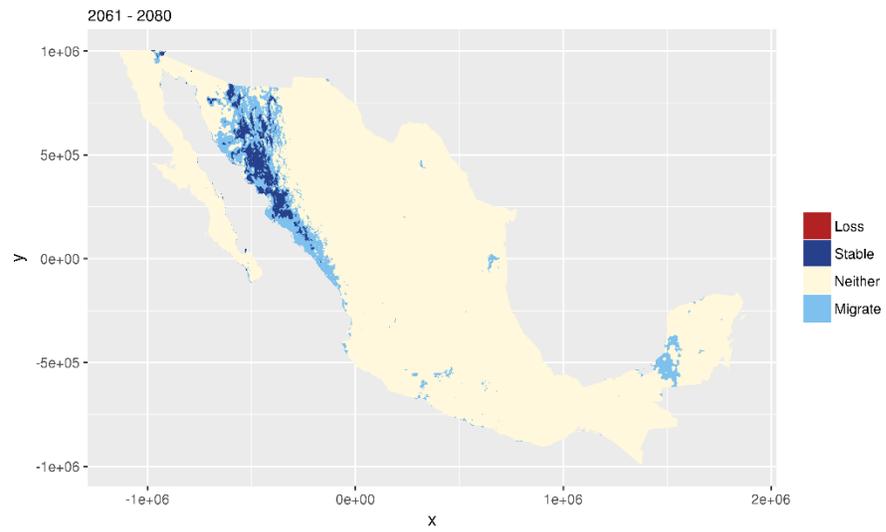
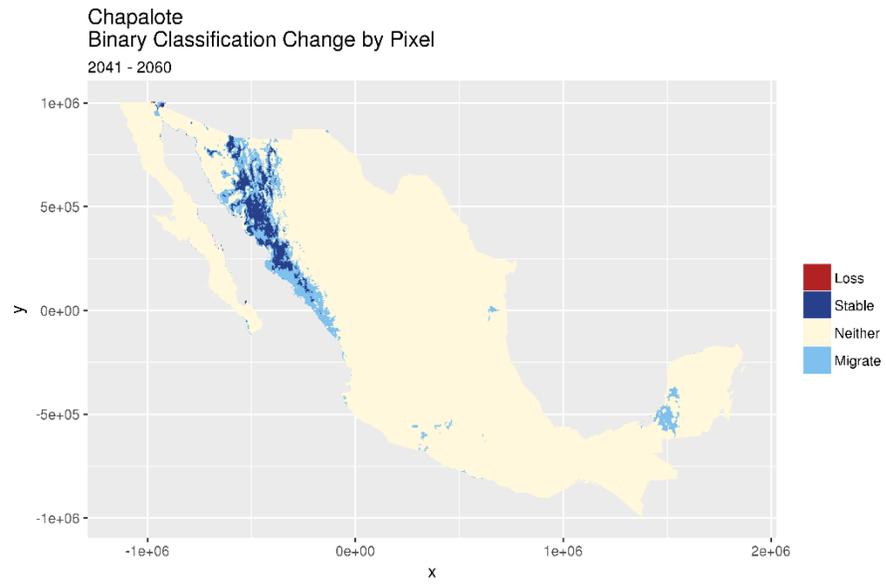


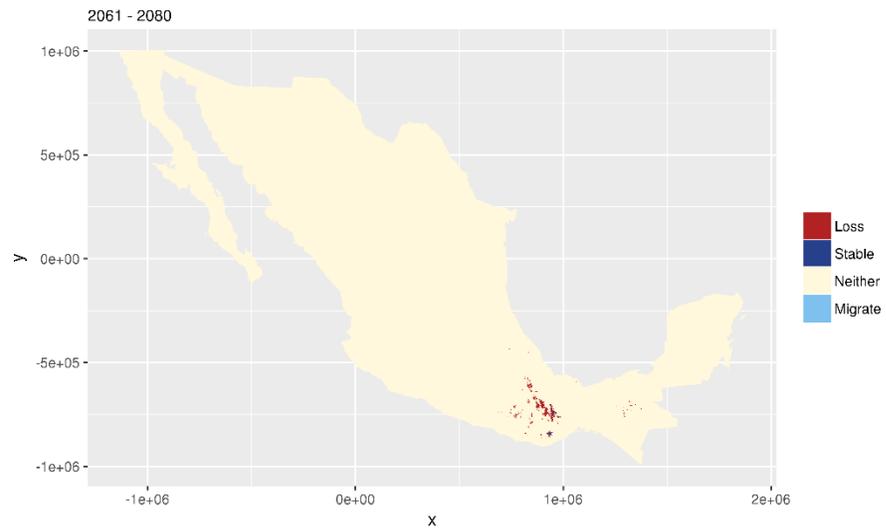
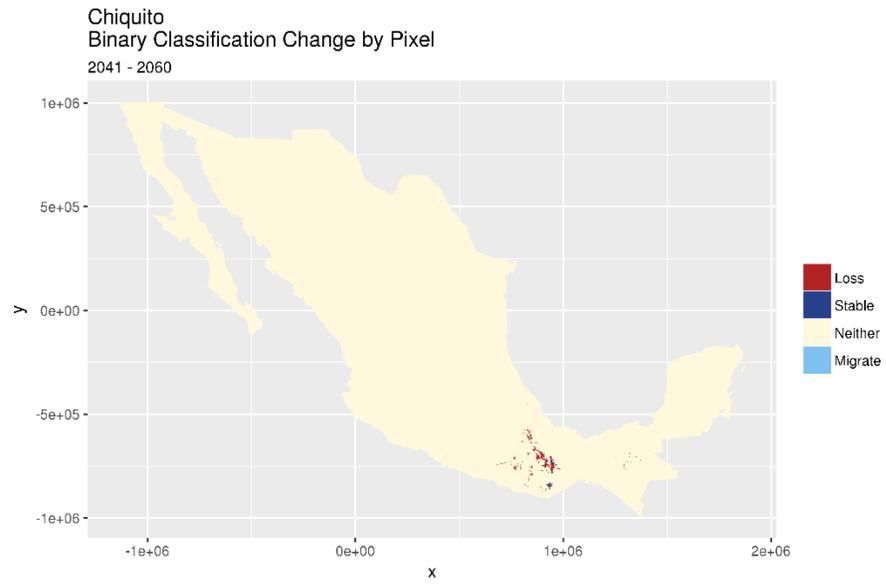


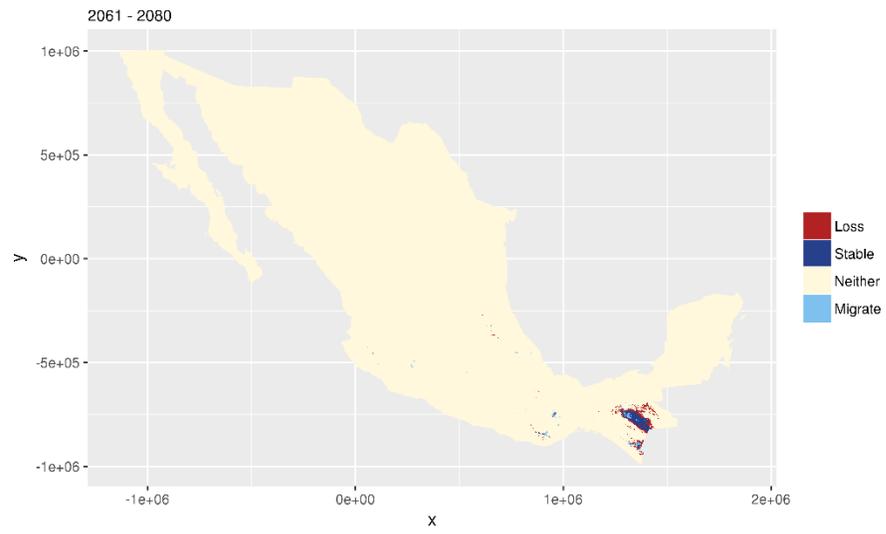
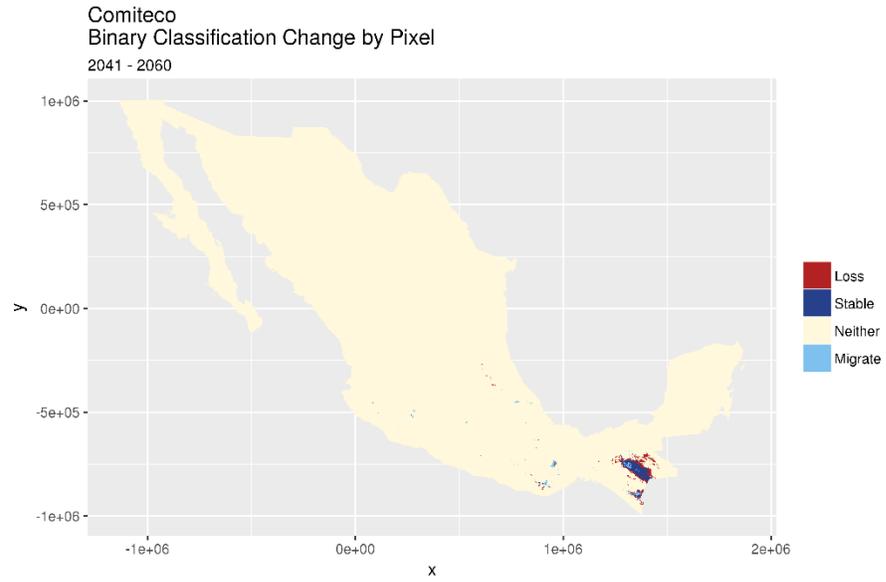


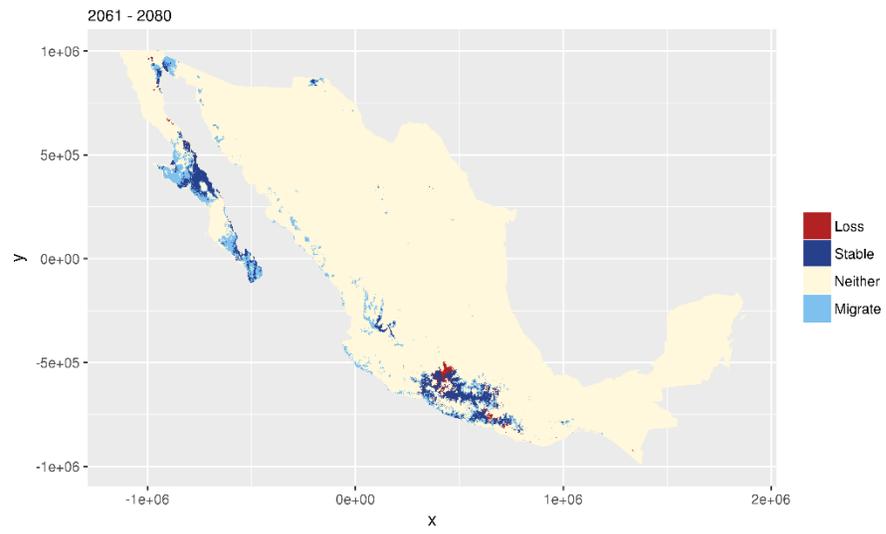
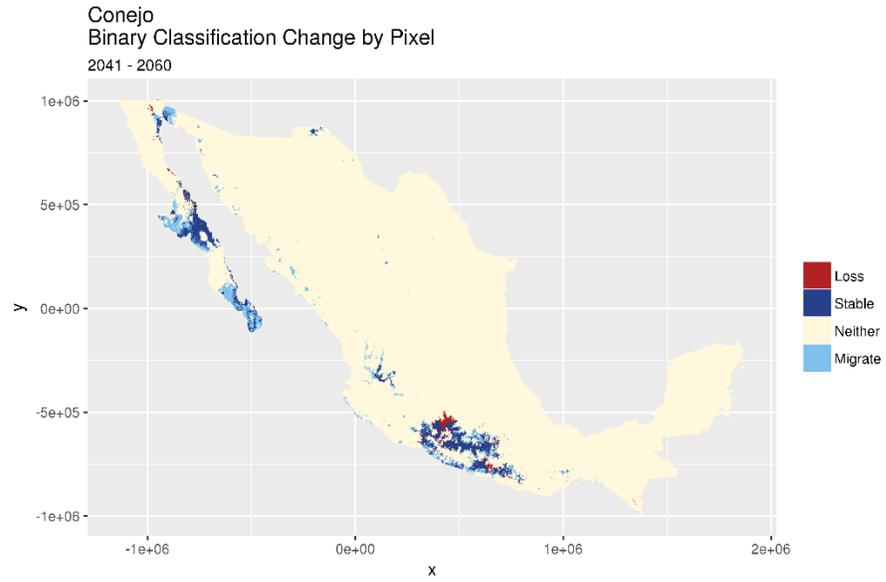


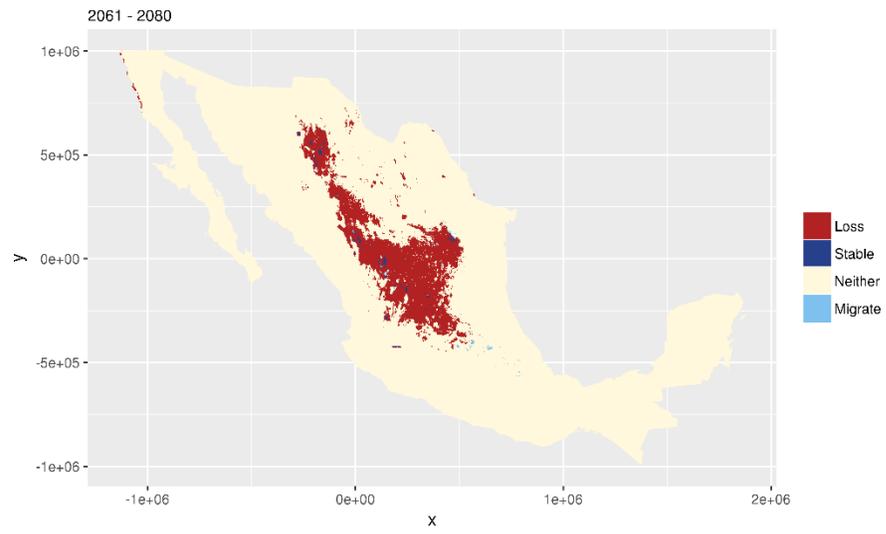
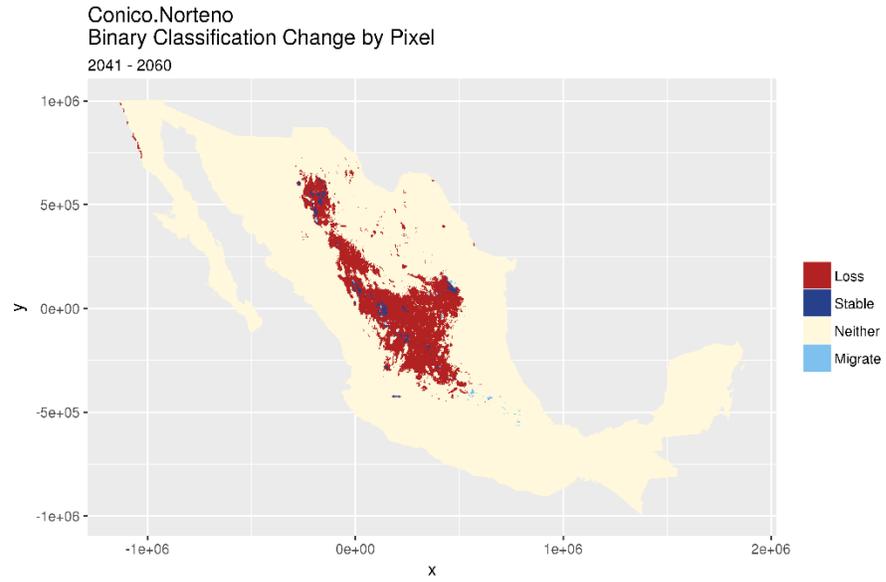


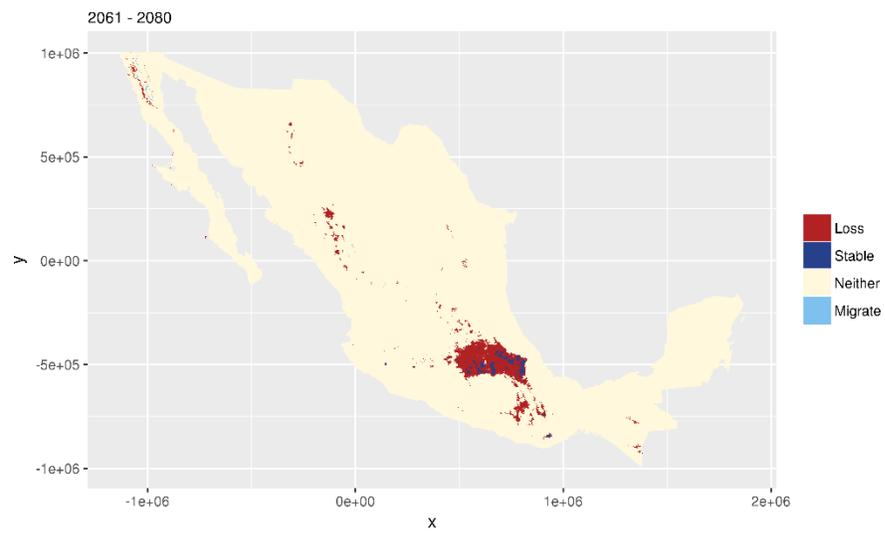
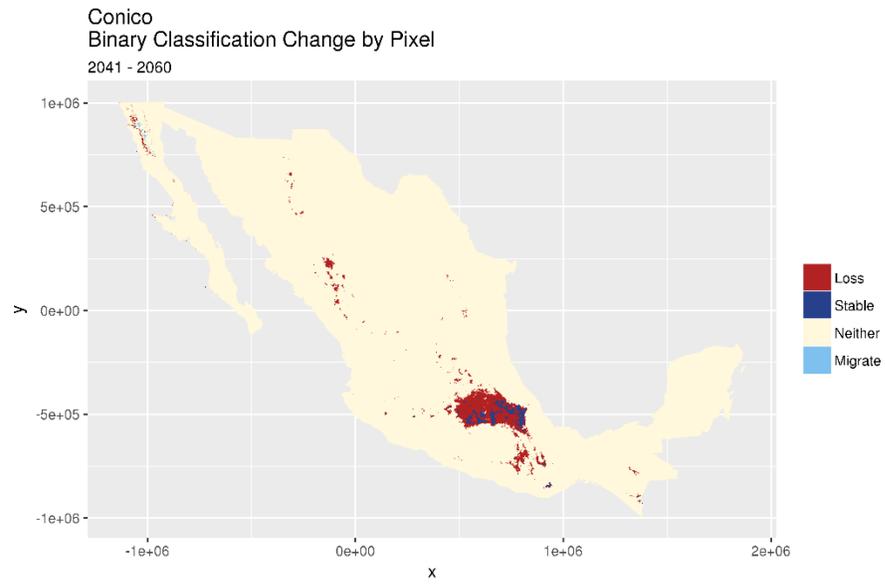


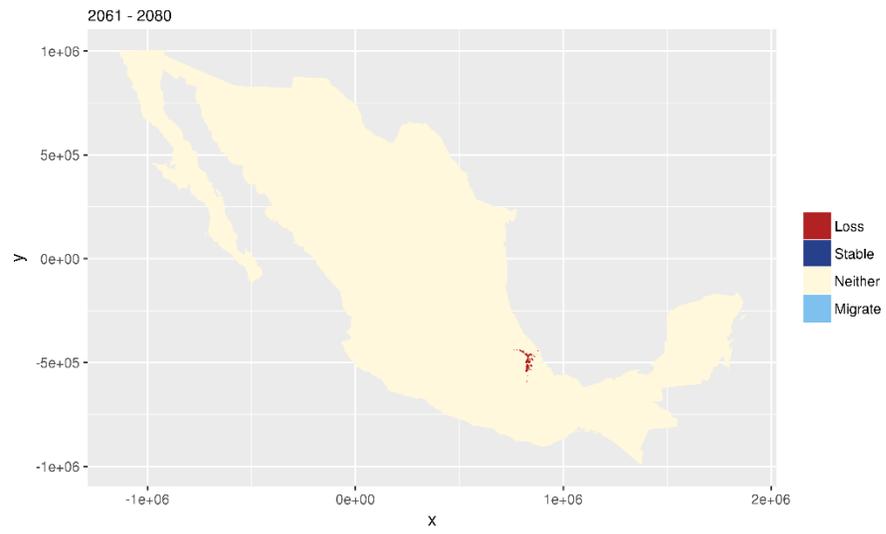
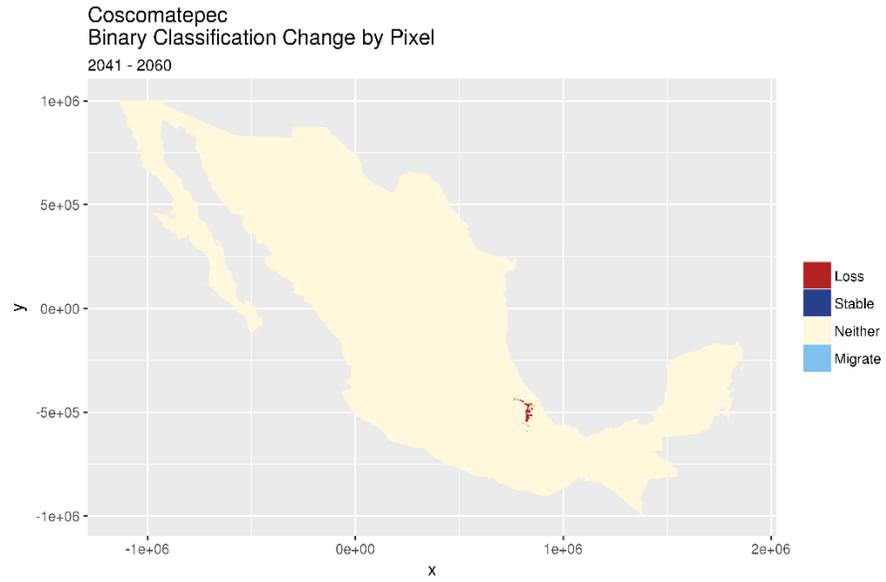


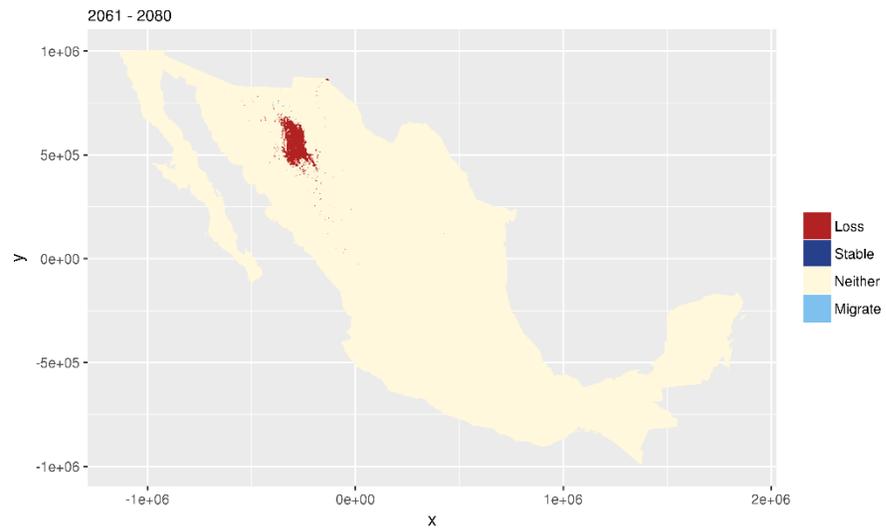
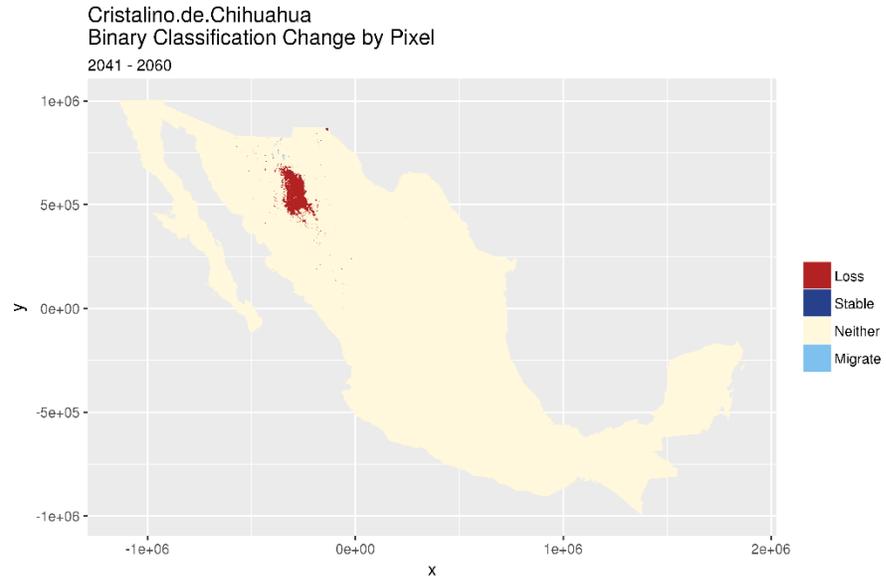


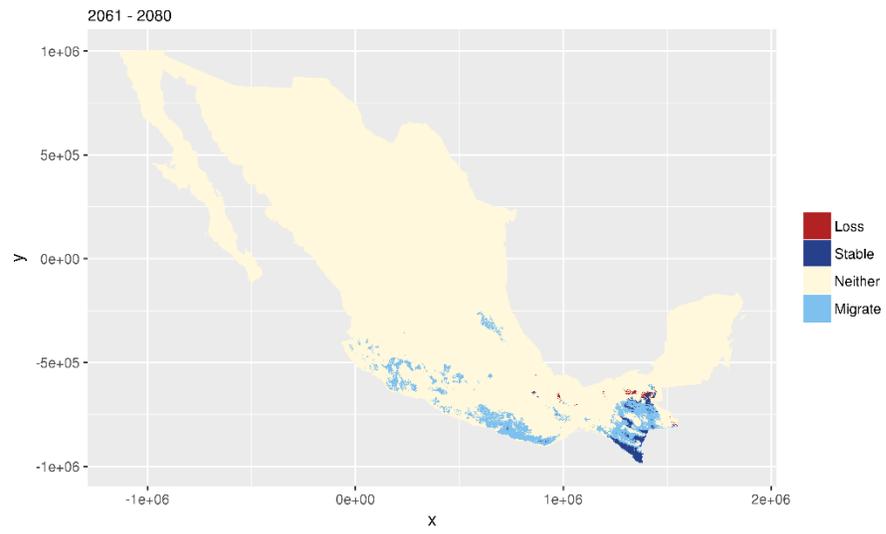
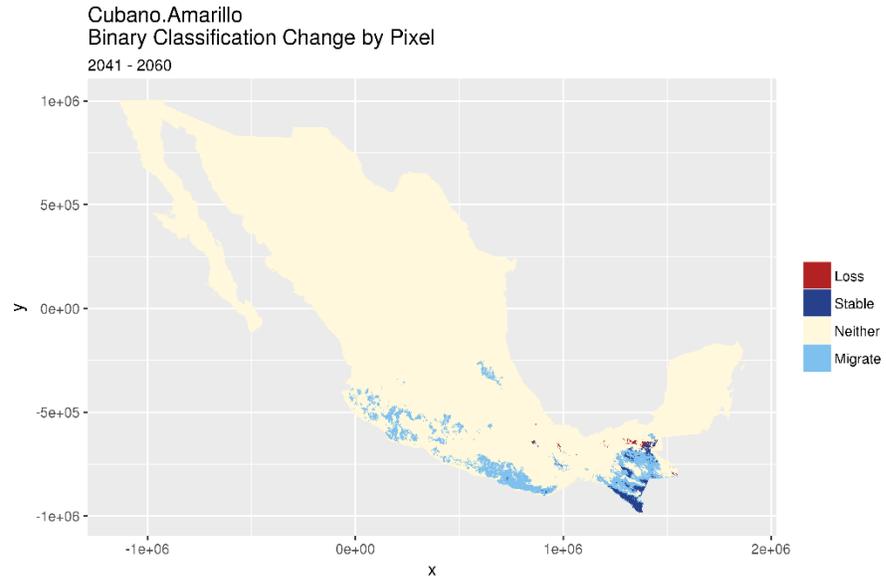


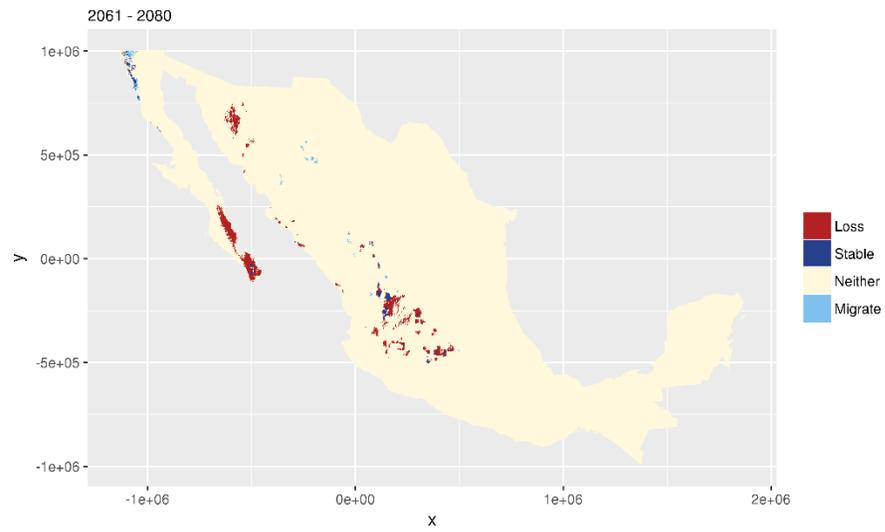
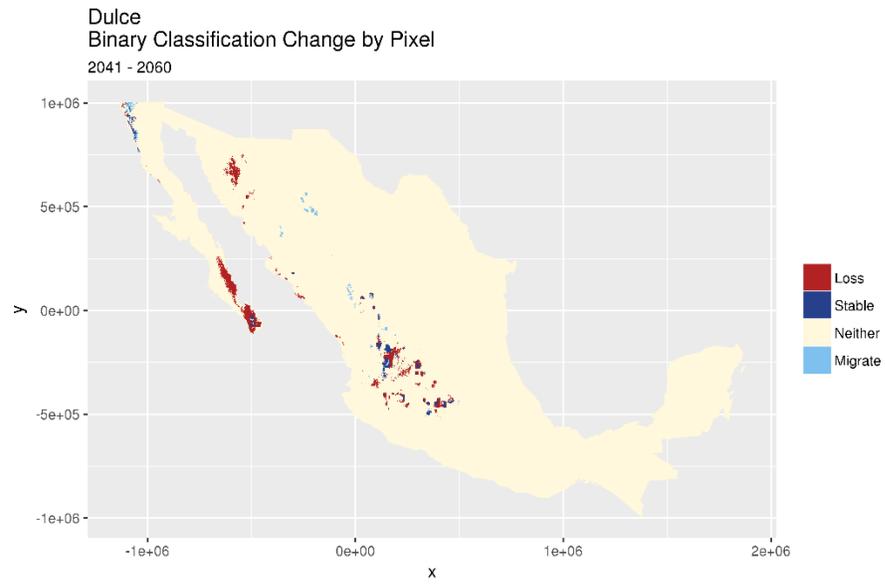




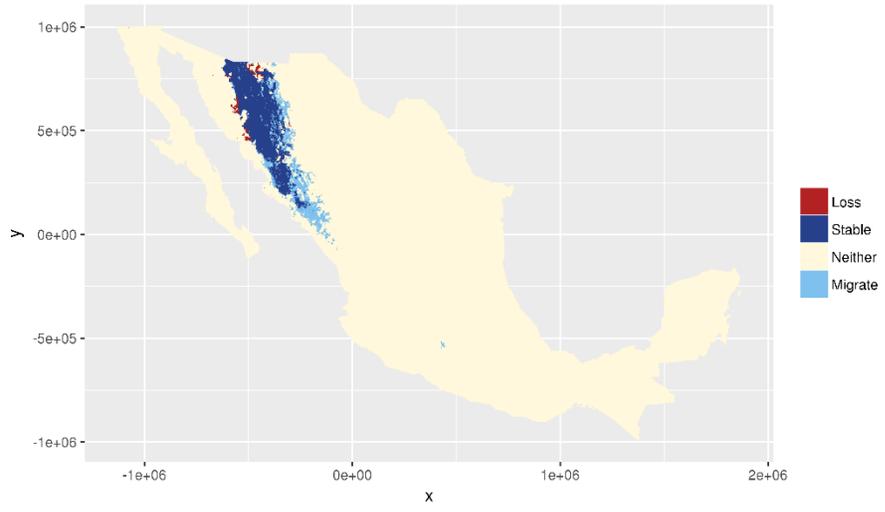




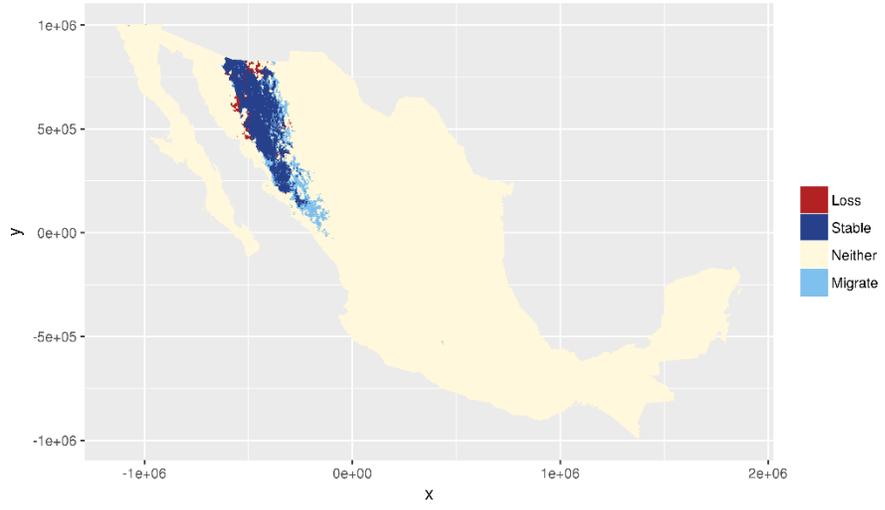


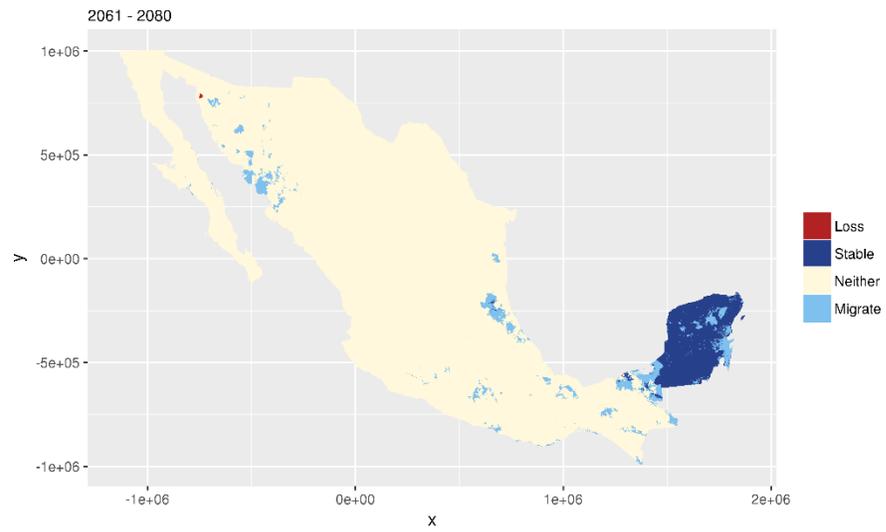
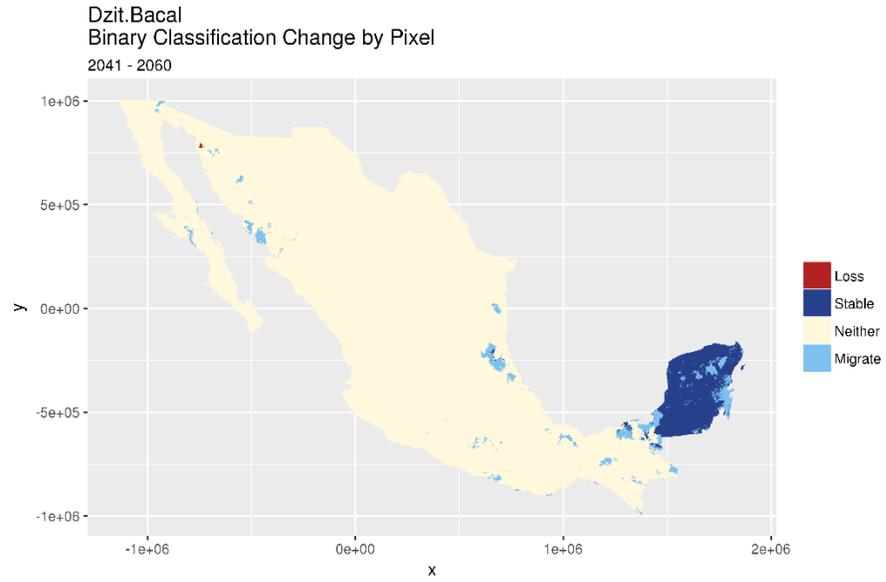


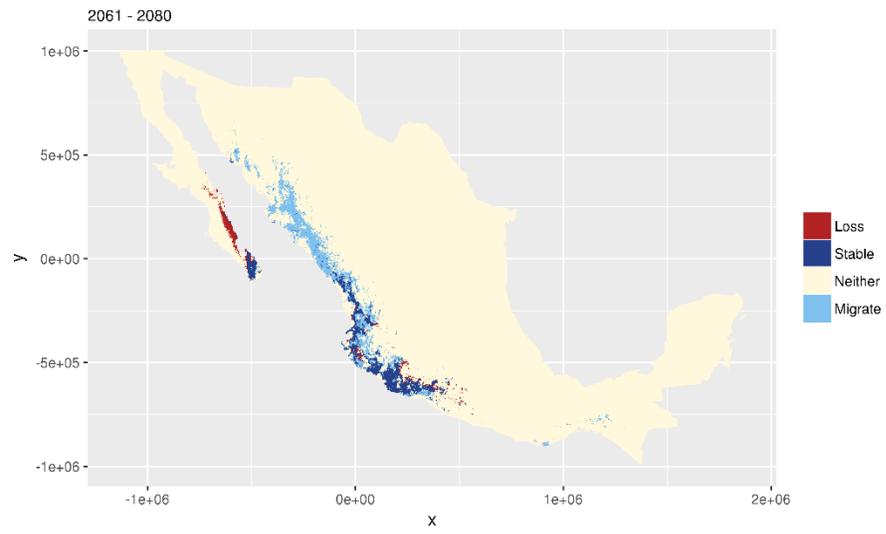
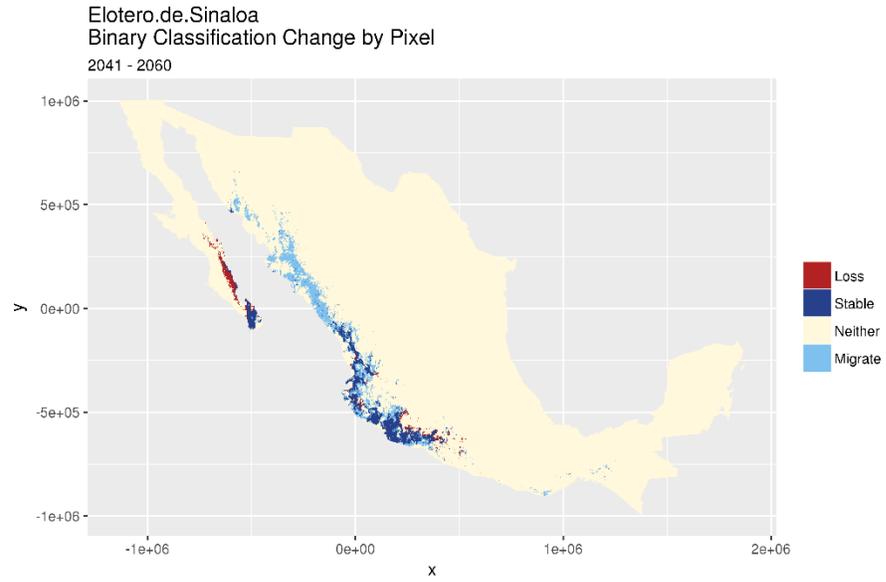
Dulcillo.del.Noroeste
Binary Classification Change by Pixel
2041 - 2060

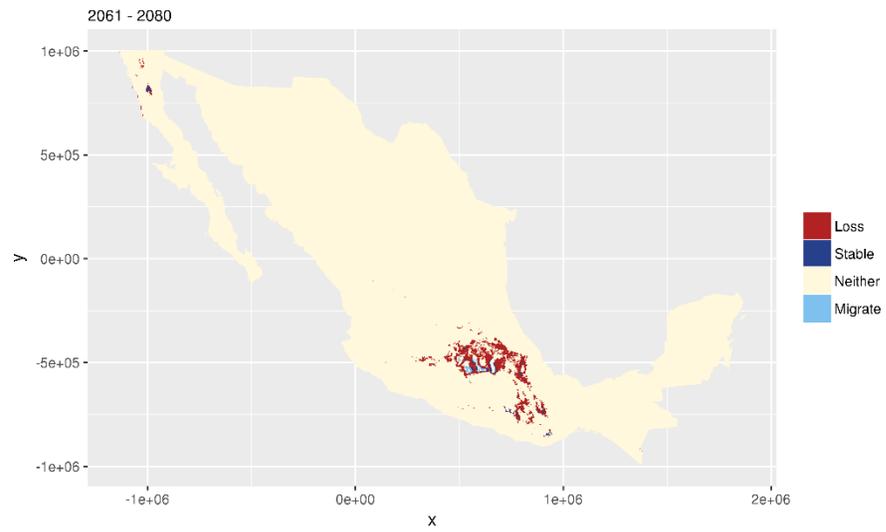
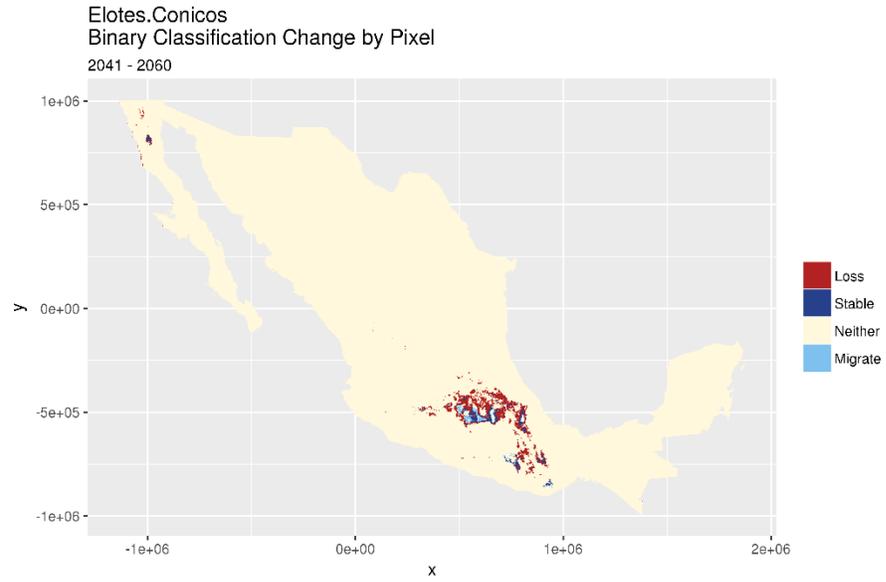


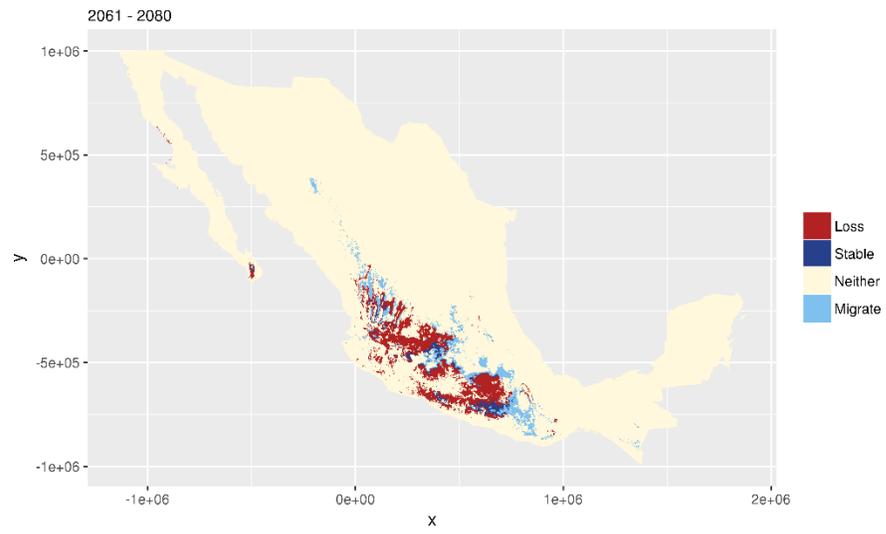
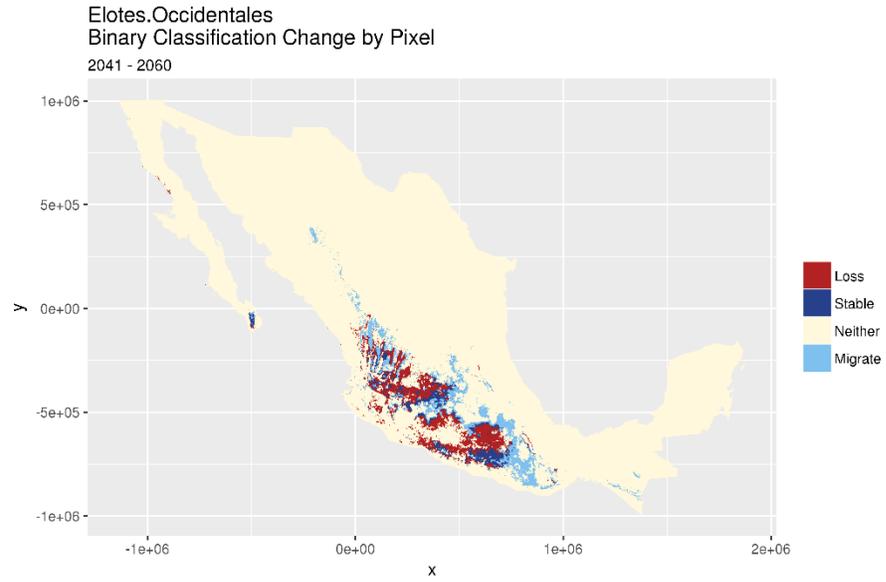
2061 - 2080



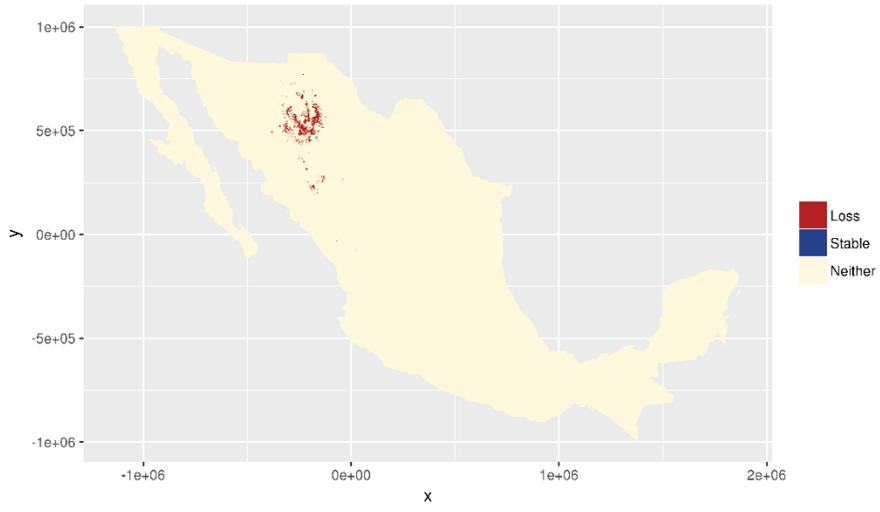




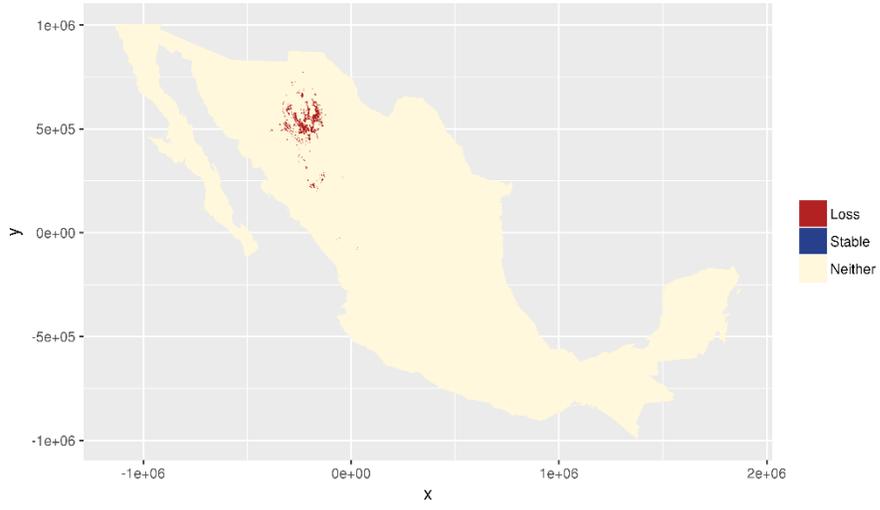


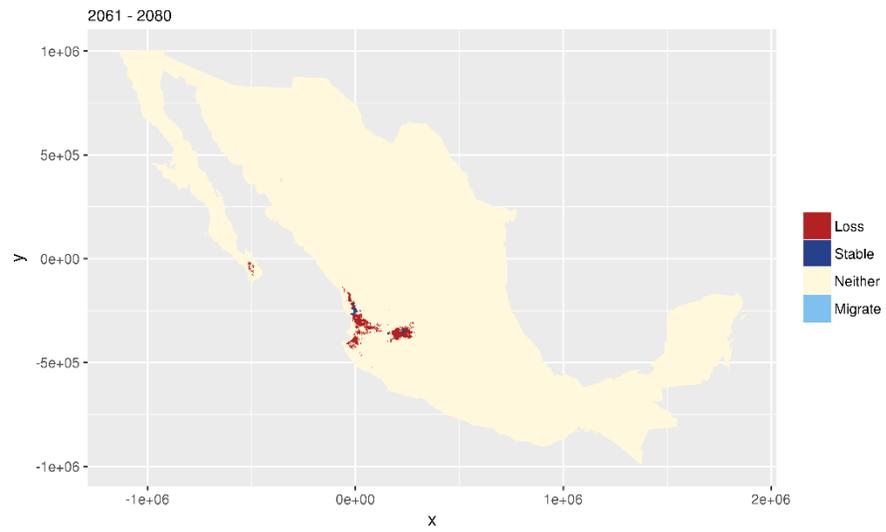
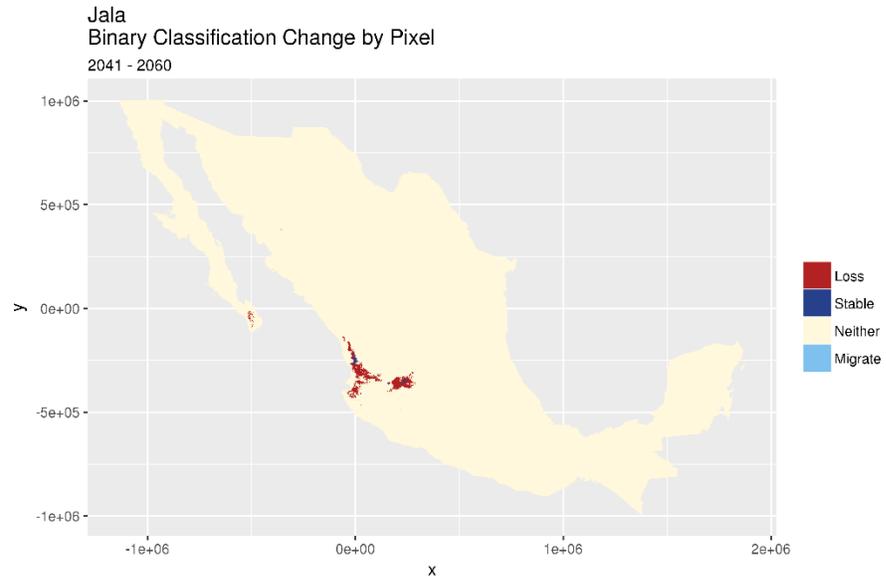


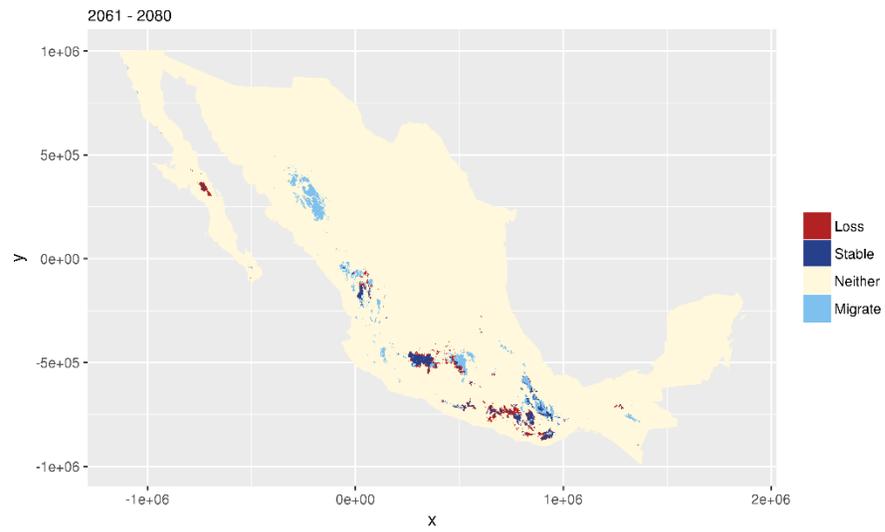
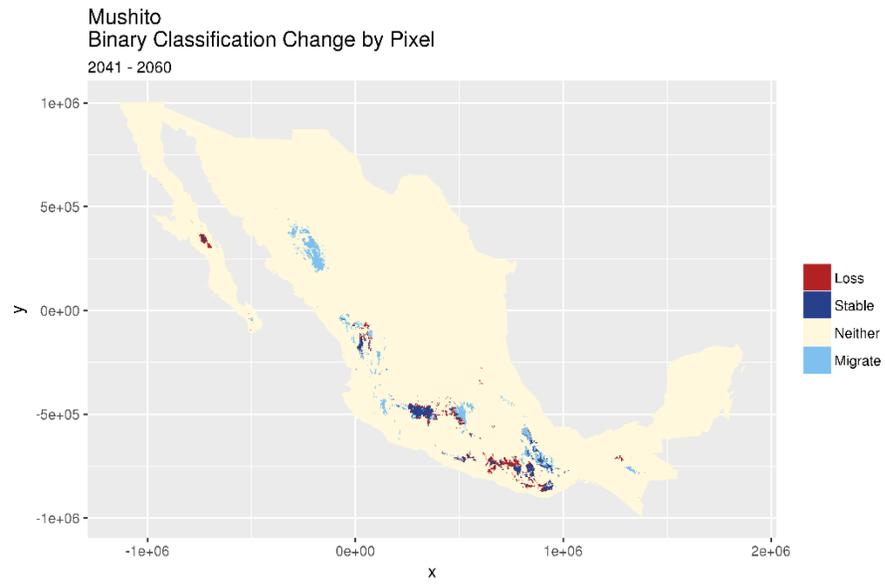
Gordo
Binary Classification Change by Pixel
2041 - 2060

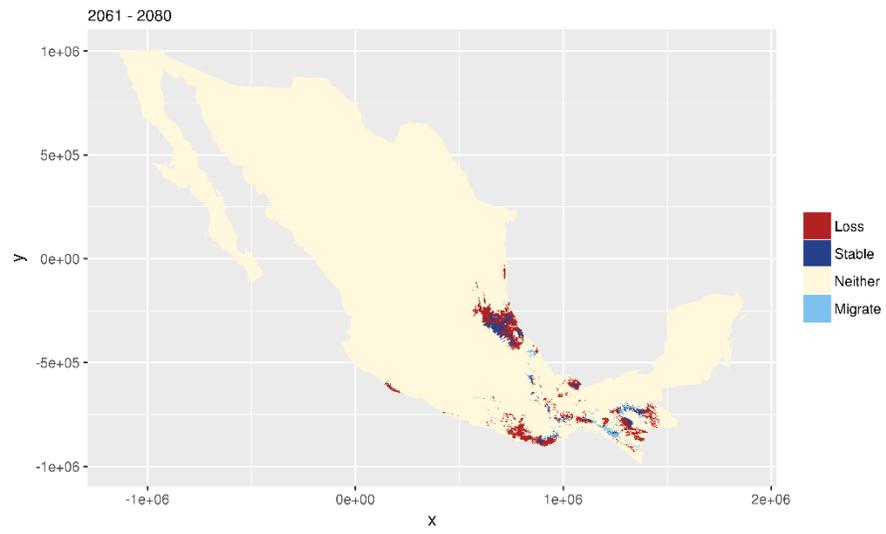
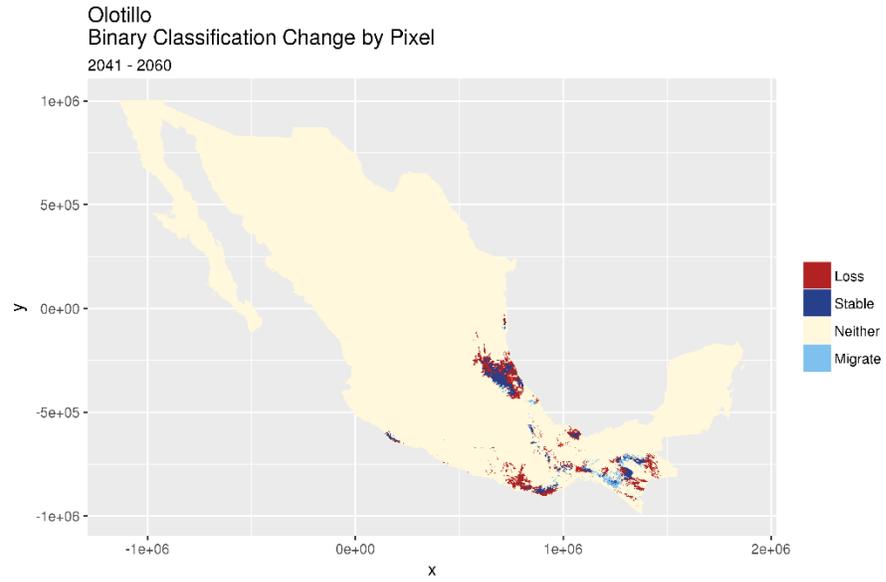


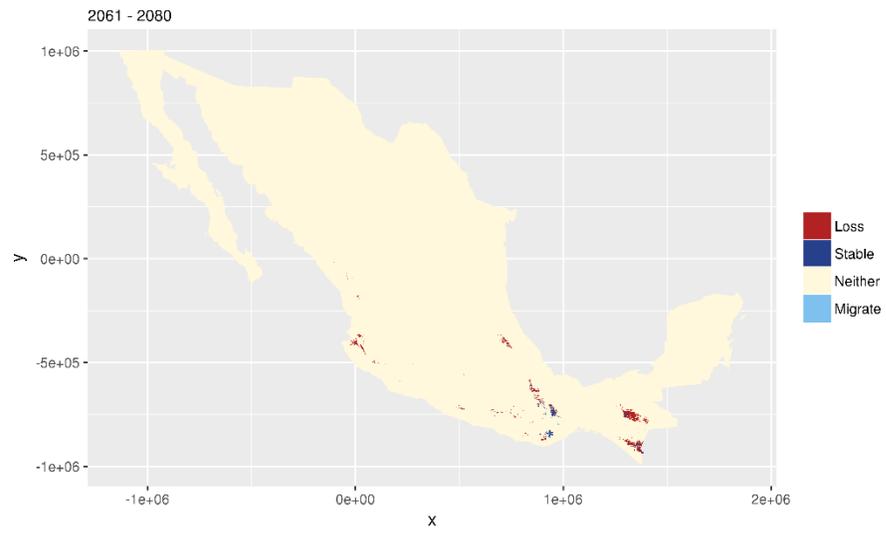
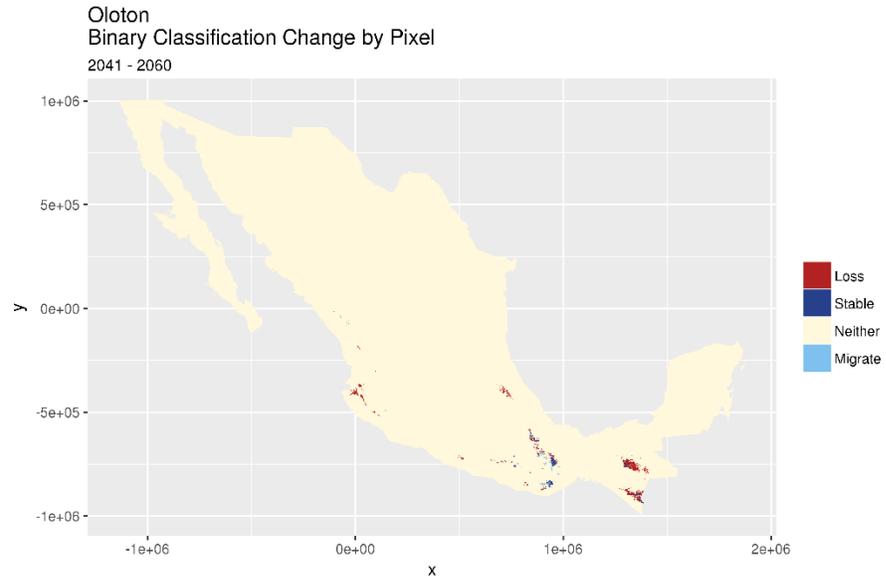
2061 - 2080

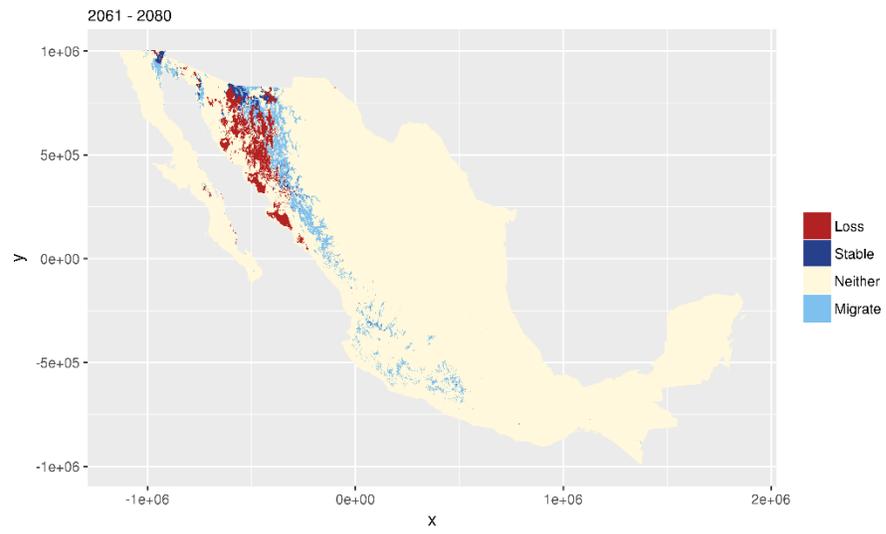
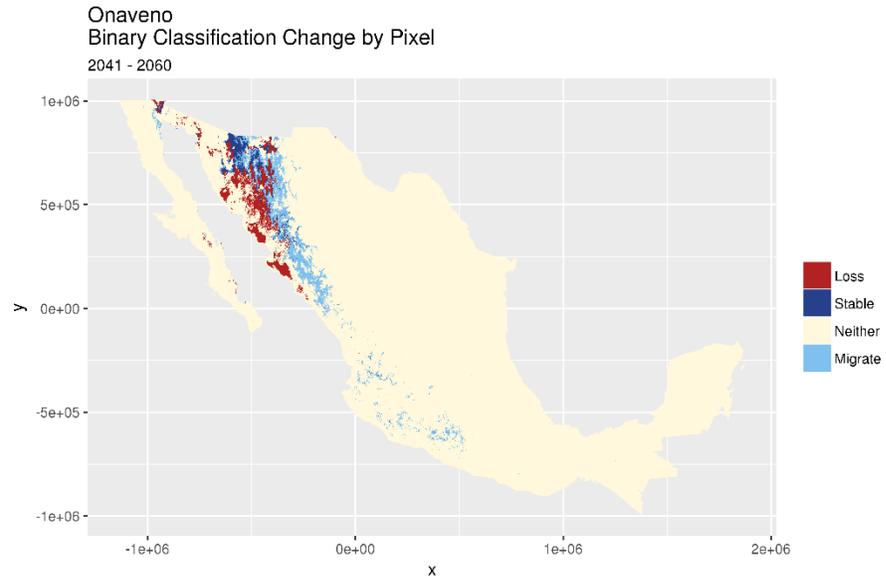




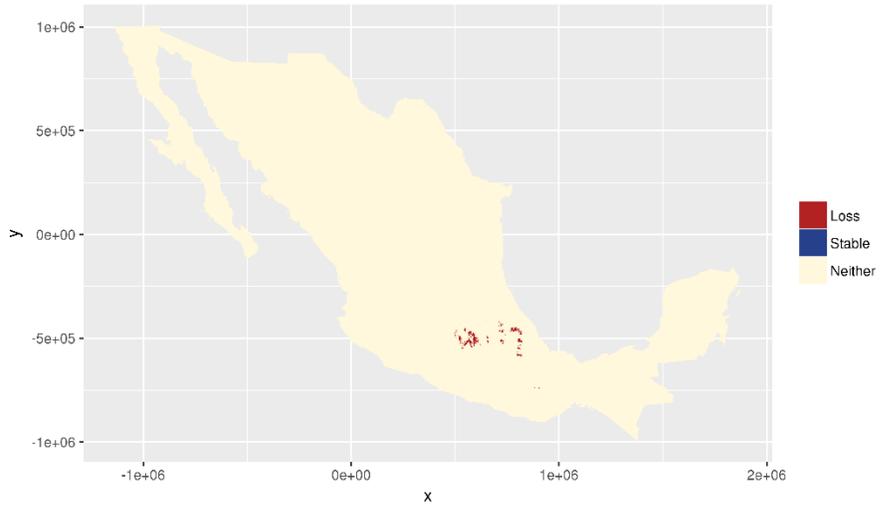




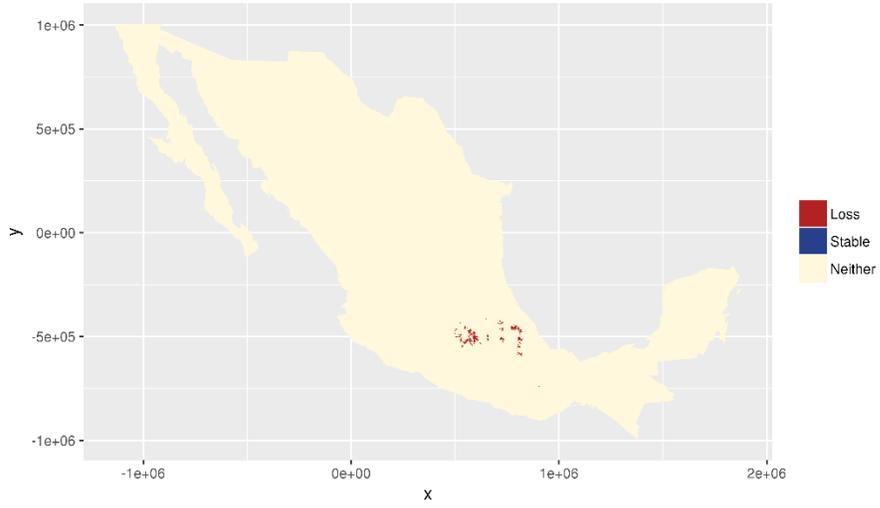


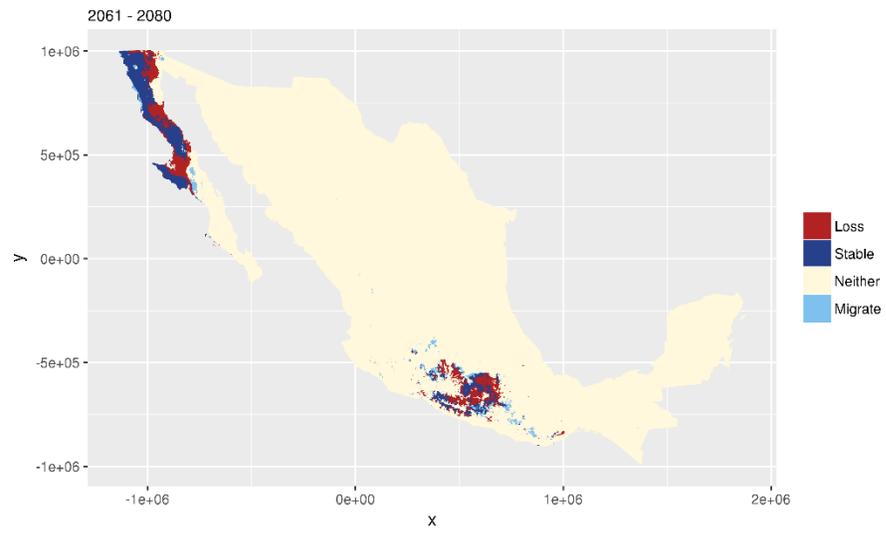
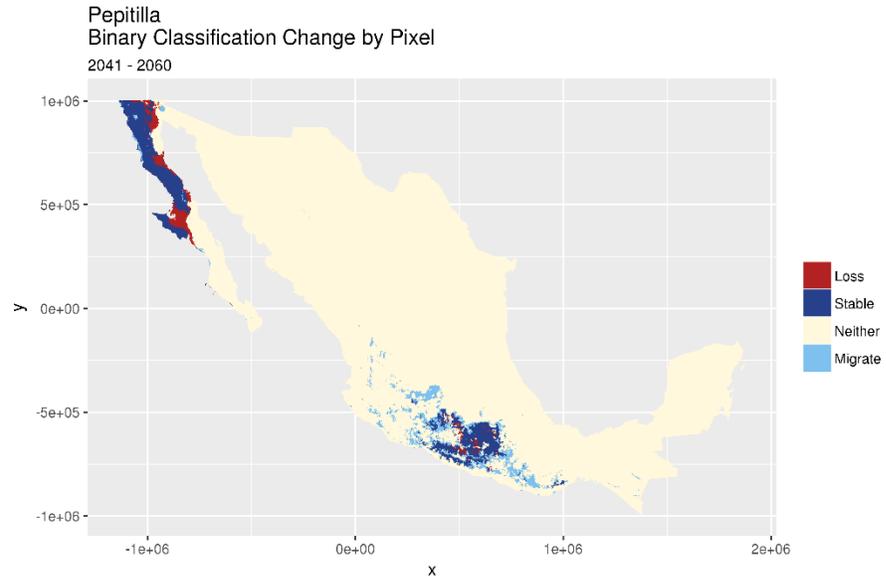


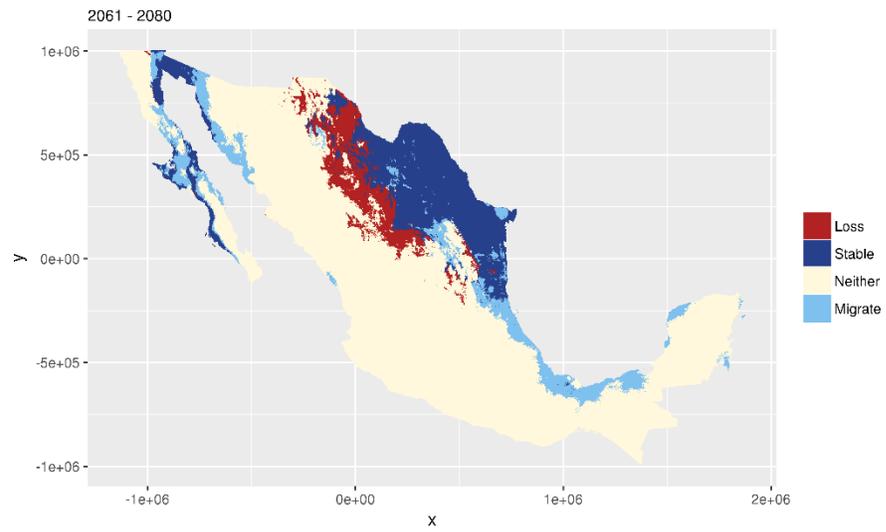
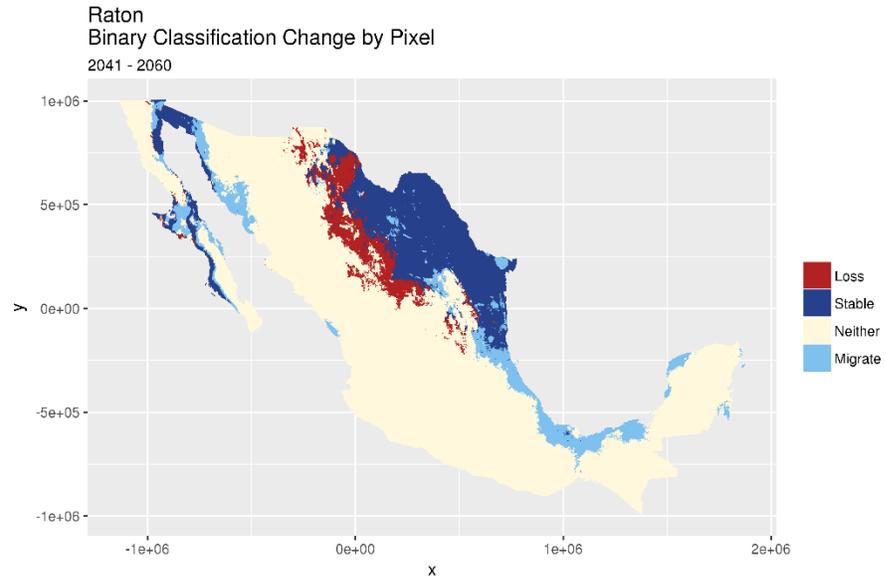
Palomero.Toluqueno
Binary Classification Change by Pixel
2041 - 2060

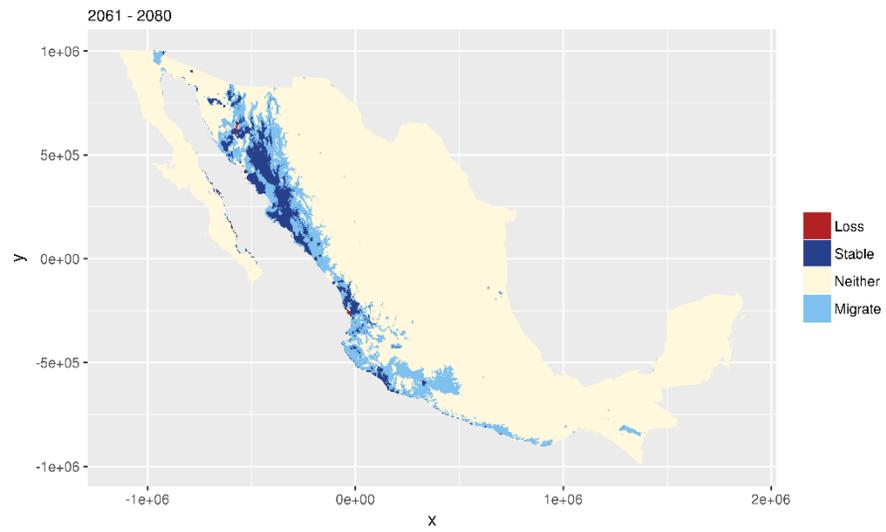
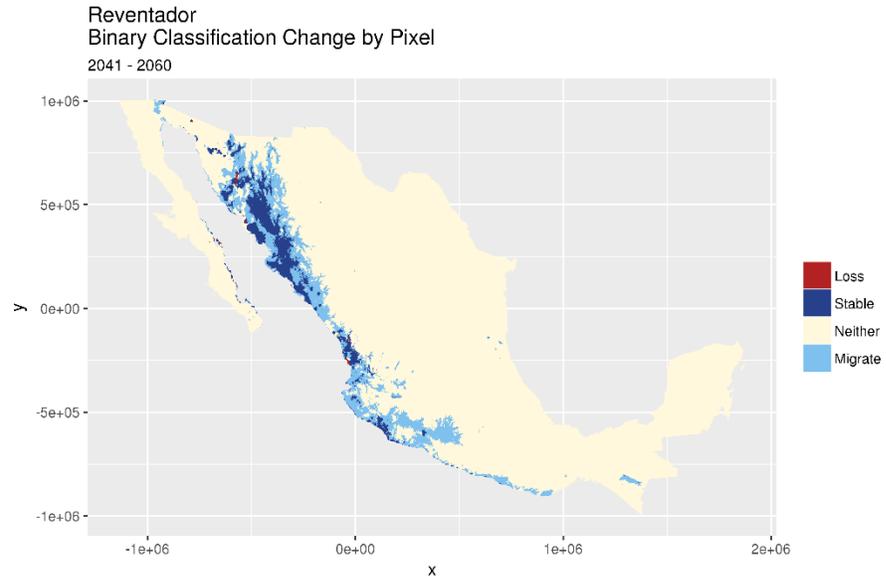


2061 - 2080

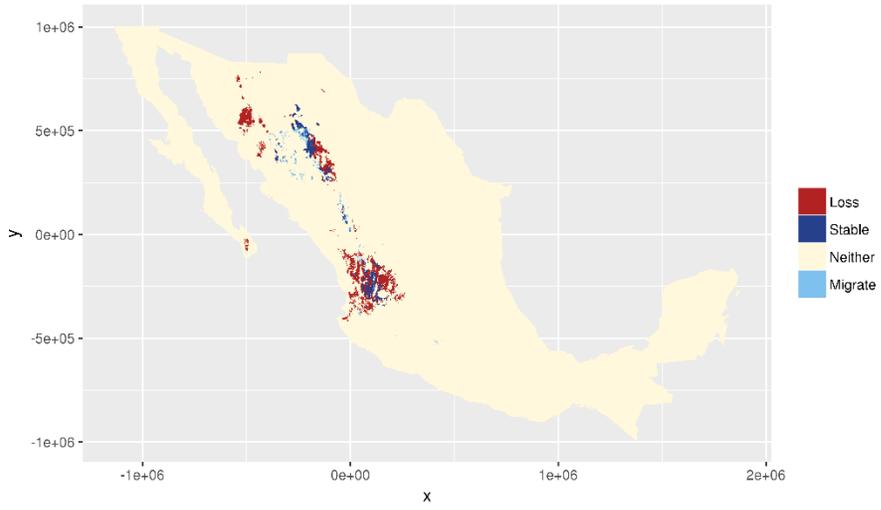




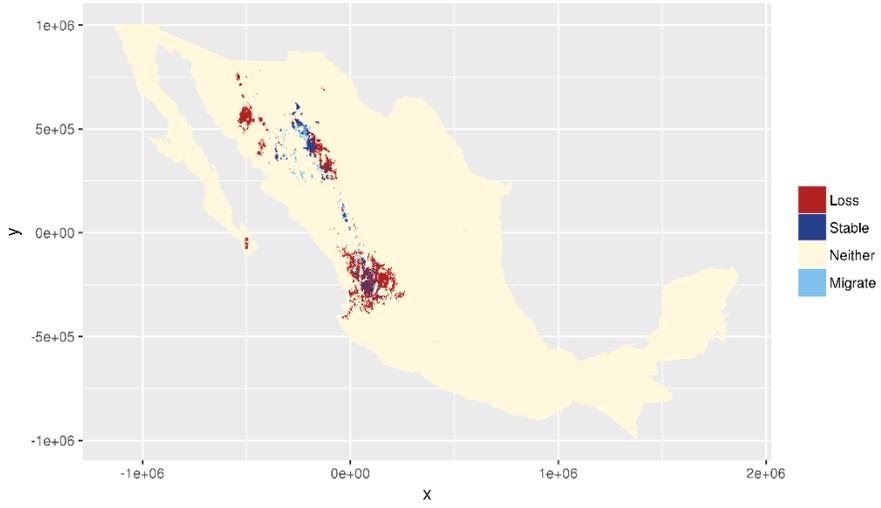


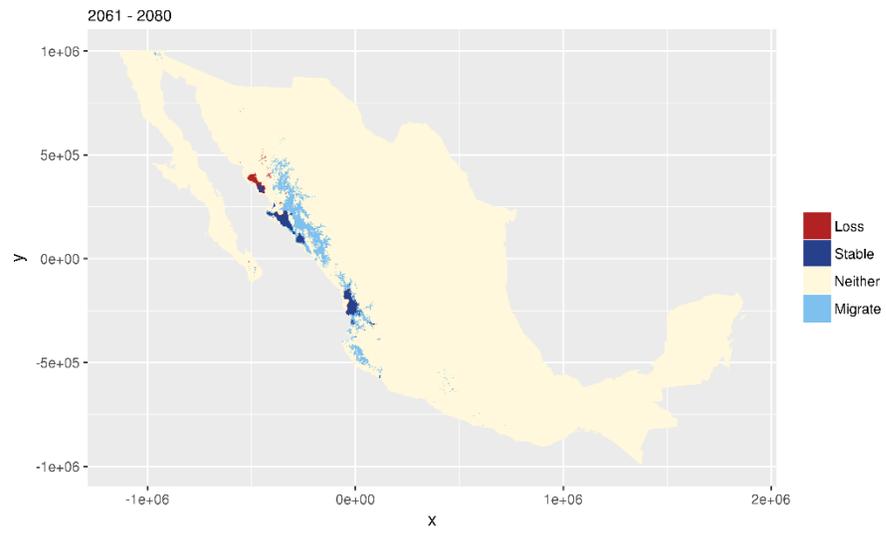
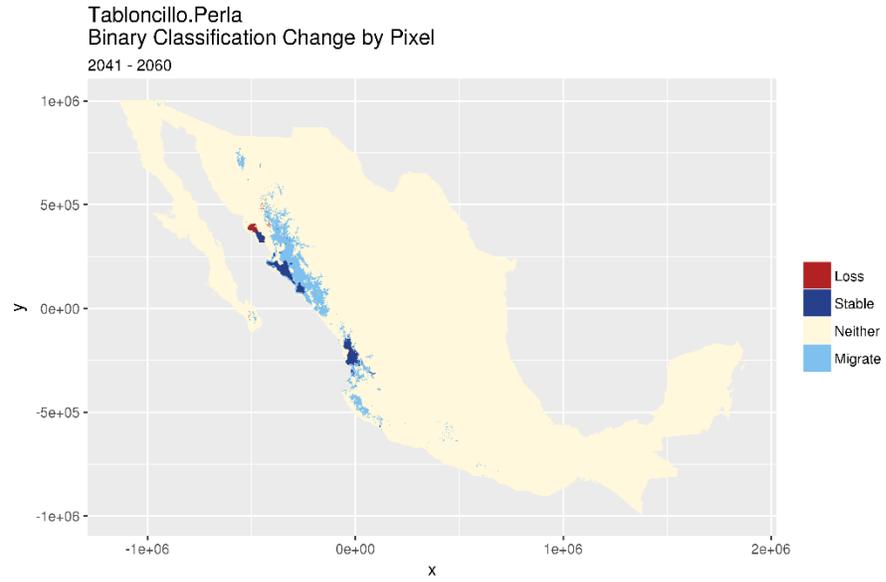


Tablilla.de.Ocho
Binary Classification Change by Pixel
2041 - 2060

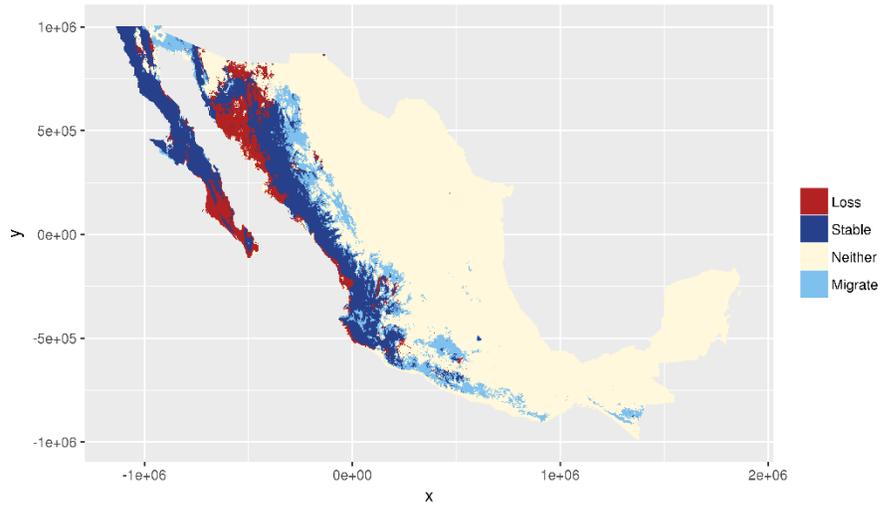


2061 - 2080

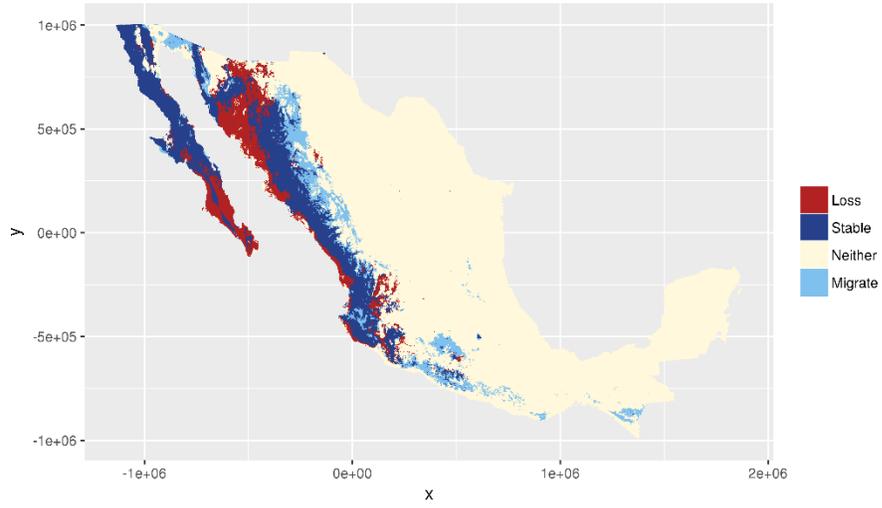


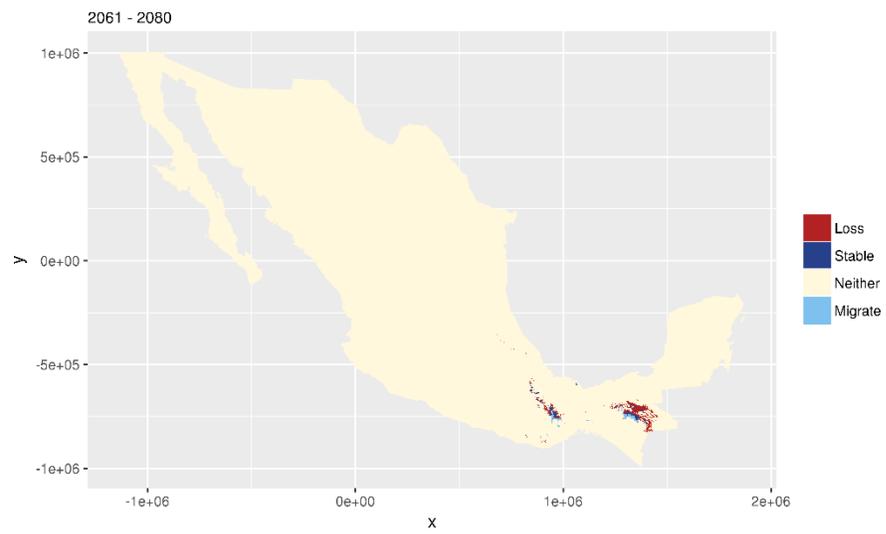
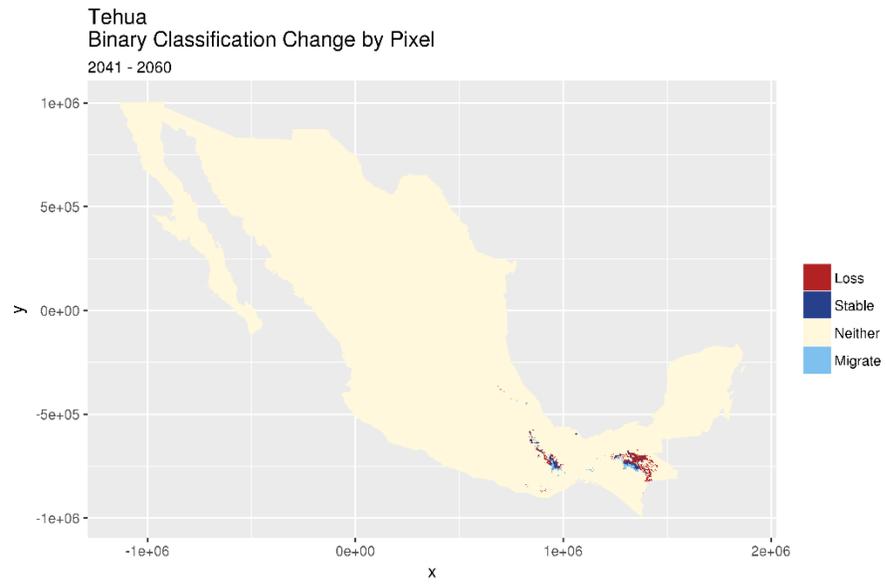


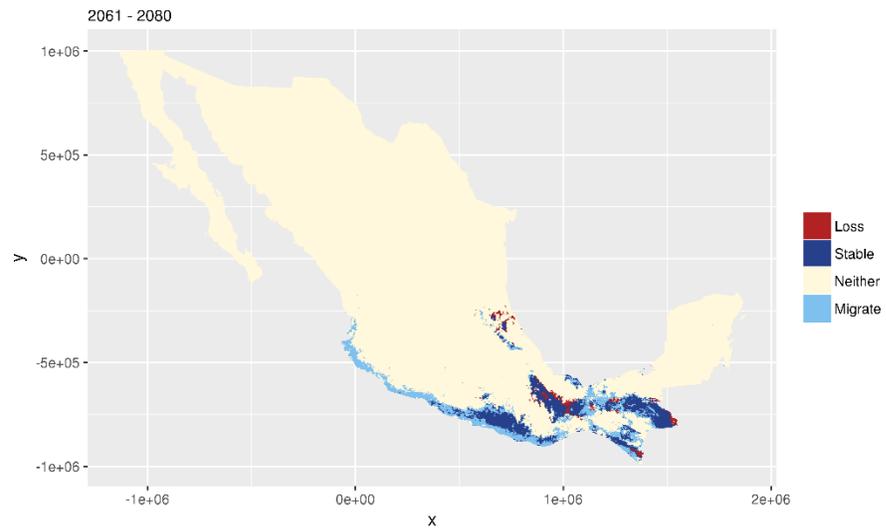
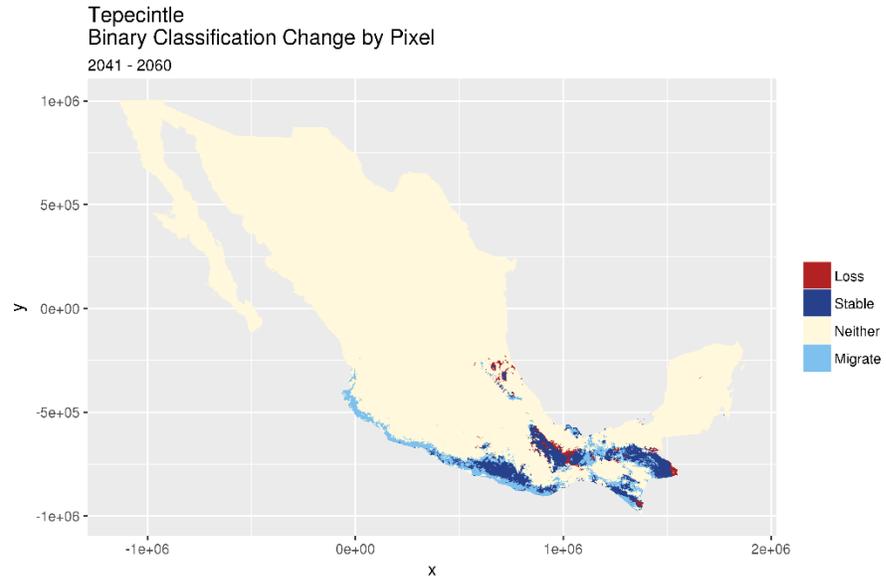
Tabloncillo
Binary Classification Change by Pixel
2041 - 2060



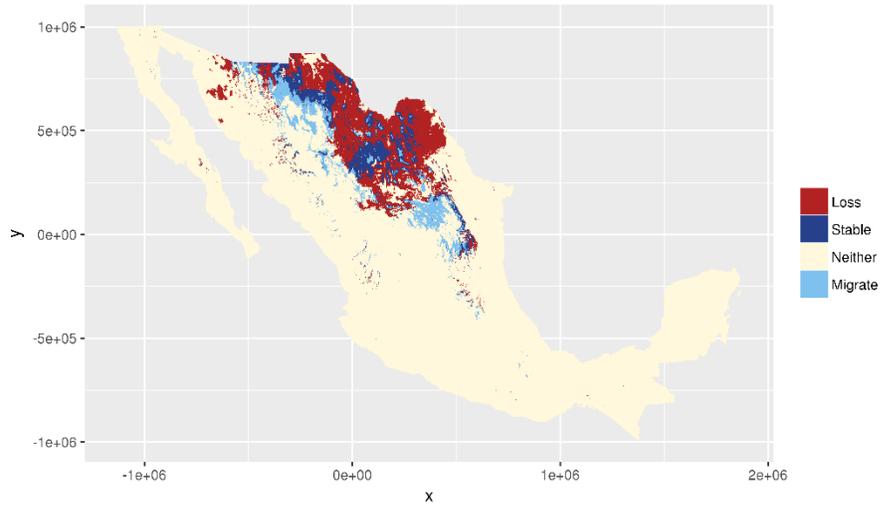
2061 - 2080



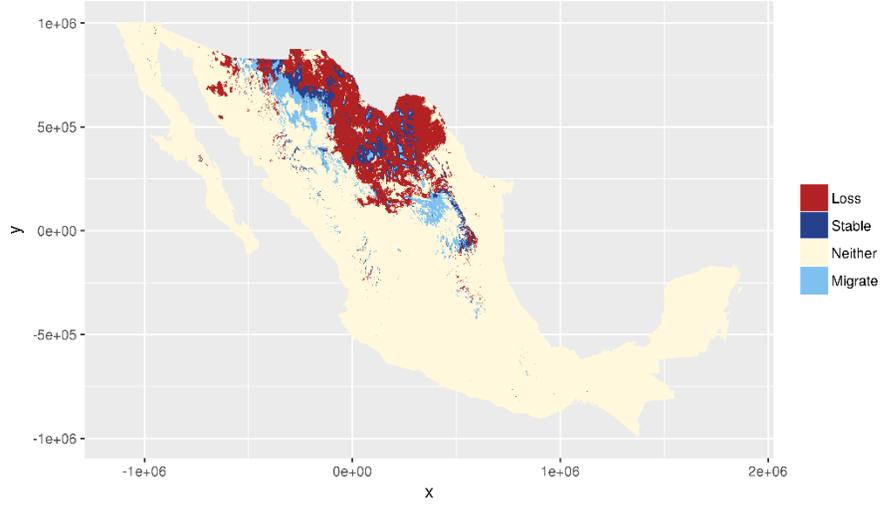


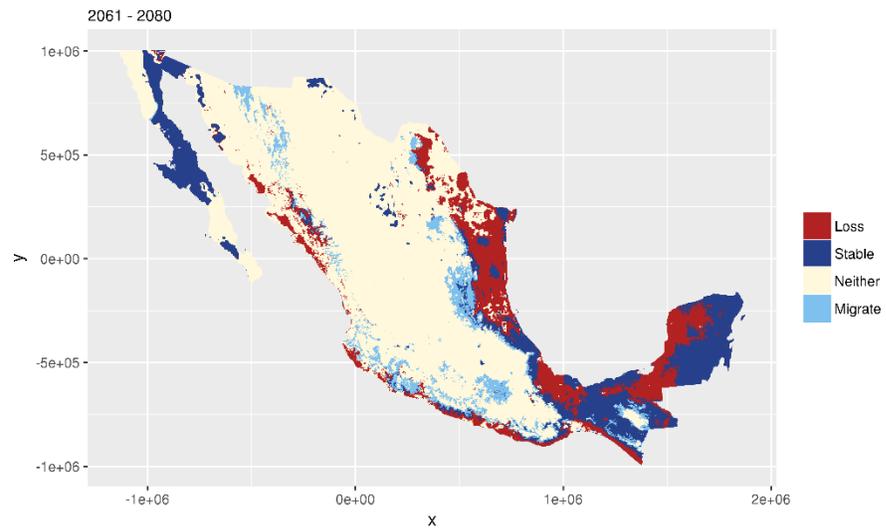
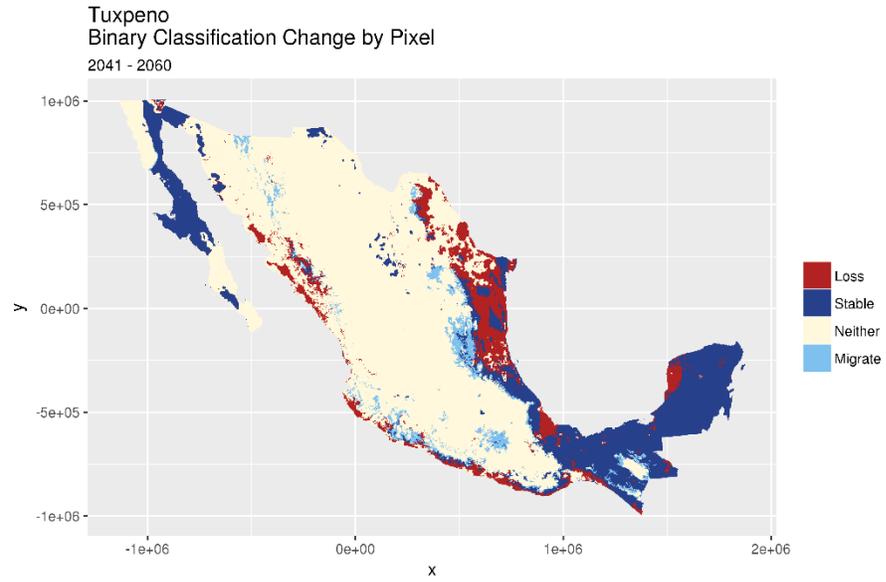


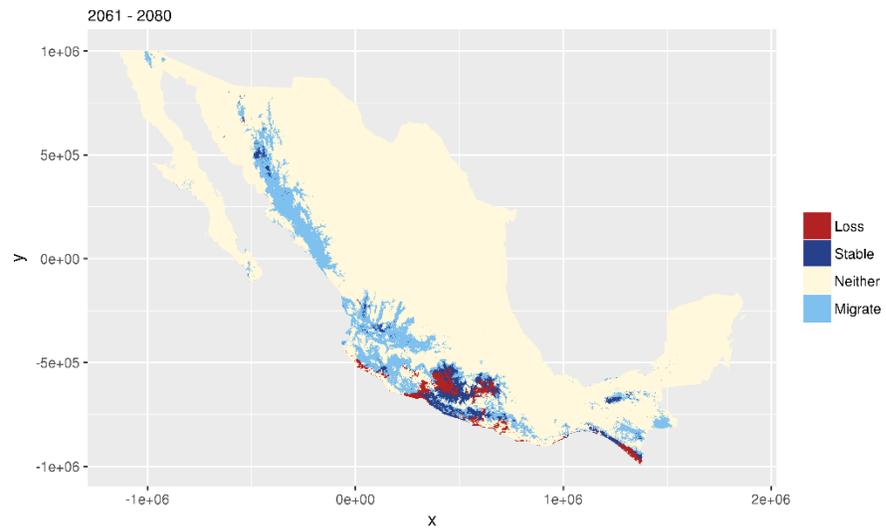
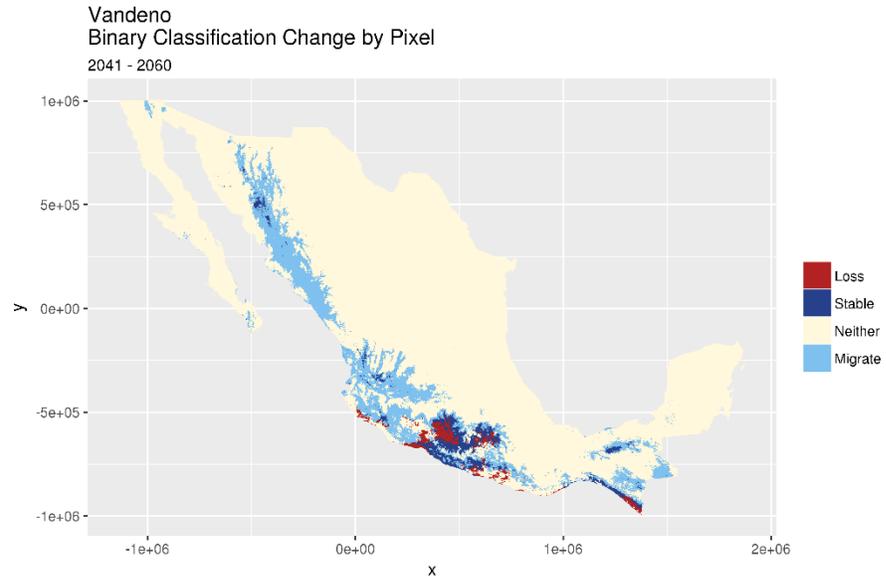
Tuxpeno.Norteno
Binary Classification Change by Pixel
2041 - 2060



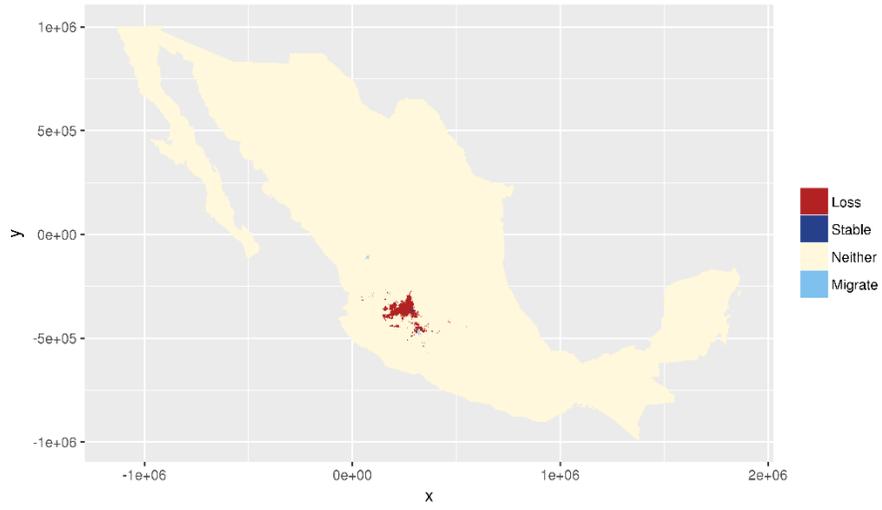
2061 - 2080



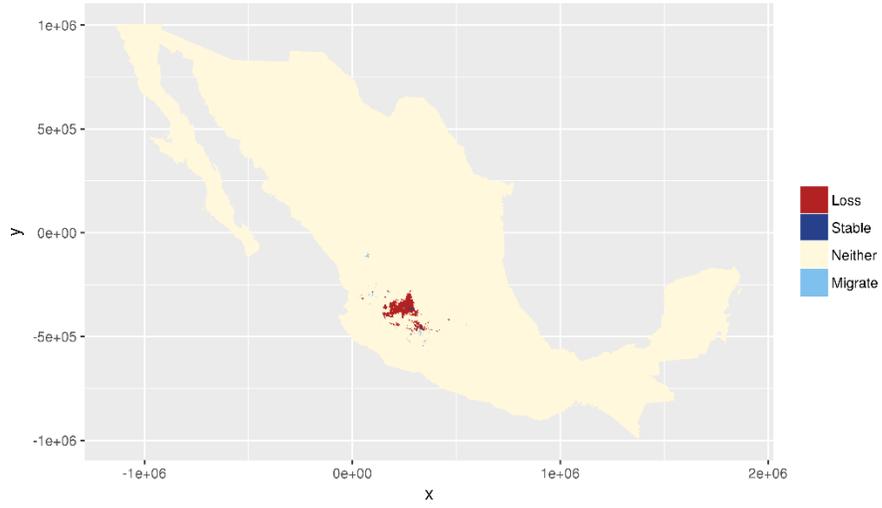




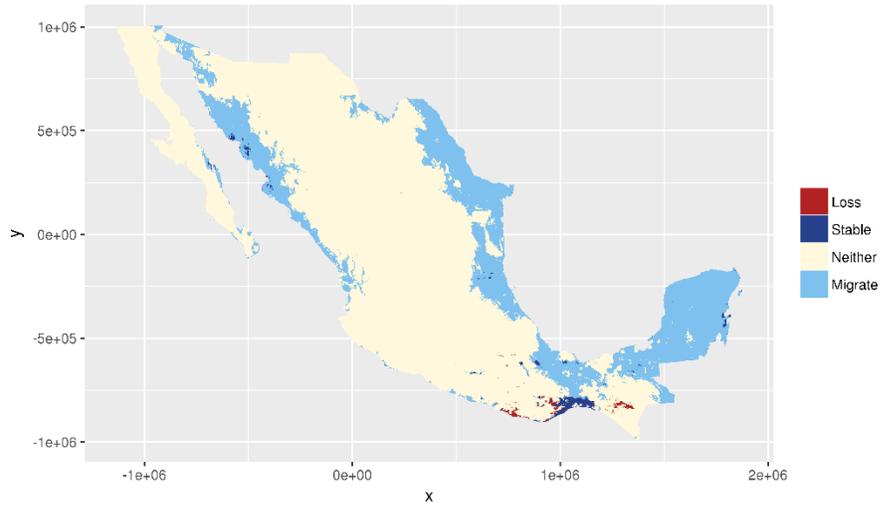
Zamorano.Amarillo
Binary Classification Change by Pixel
2041 - 2060



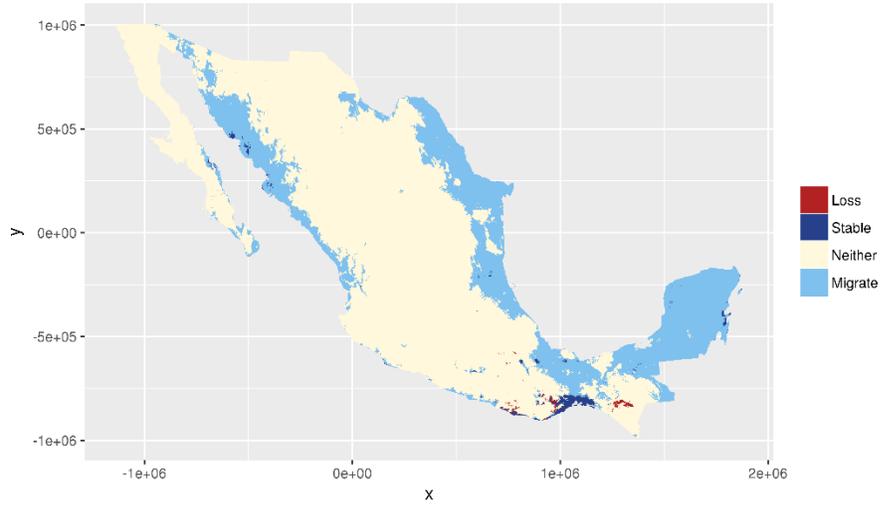
2061 - 2080

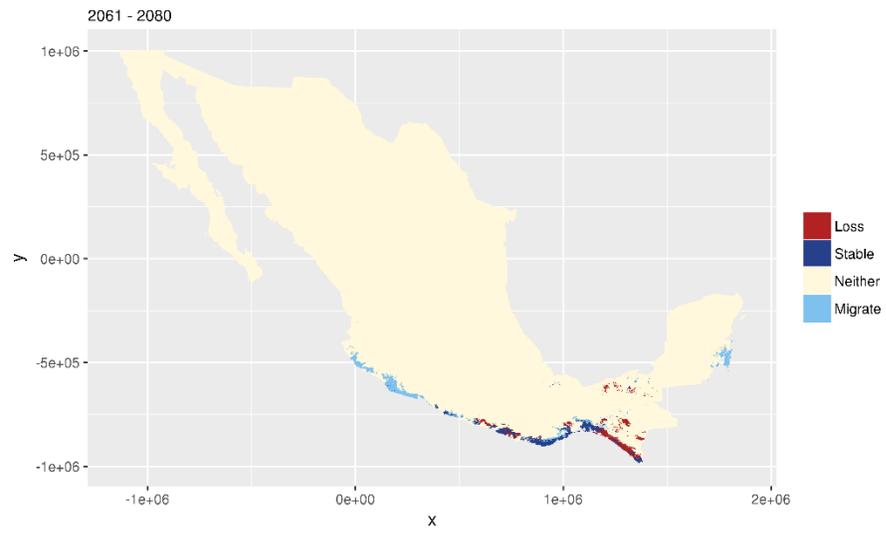
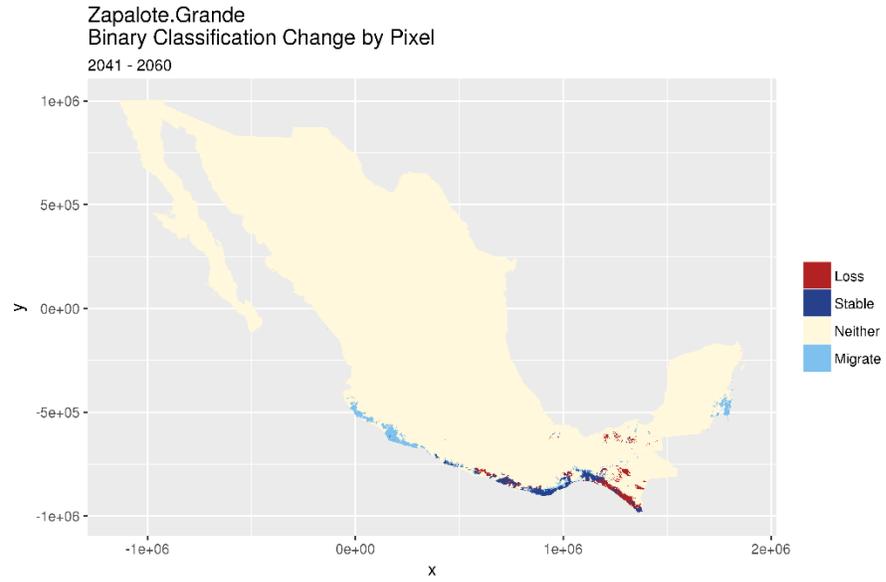


Zapalote.Chico
Binary Classification Change by Pixel
2041 - 2060



2061 - 2080





D. R Packages Citations

(Allouche, Omri Tsoar, Asaf Kadmon n.d.; Davison and Hinkley 1997; Chessel, Dufour, and Thioulouse 2004; Sing et al. 2005; Calenge 2006; Wickham 2007, 2016, 2017; Dray and Dufour 2007; Dray, Dufour, and Chessel 2007; Tierney 2011; McHugh 2012; Stabler 2013; Bache and Wickham 2014; Grosjean and Ibanez 2014; Muscarella et al. 2014; Naimi et al. 2014; Neuwirth 2014; VanDerWal et al. 2014; Knaus 2015; Suzuki and Shimodaira 2015; Perpiñán and Hijmans 2016; Sarkar and Andrews 2016; Simpson 2016; Tierney et al. 2016; Urbanek 2016; Warnes et al. 2016; Zhang 2016; Francois 2017; Giraudoux 2017; Gordon 2017; Hijmans 2017; Kassambara 2017; Lang and the CRAN Team 2017; McIntire and Chubaty 2017; Ogle 2017; Oksanen et al. 2017; P O Title and Bemmels 2017; R Core Team 2017a; R. J. Hijmans et al. 2017; Canty and Ripley 2017; Thuiller et al. 2017; Wei and Simko 2017; Wright 2017; Evans 2017; Dag, Dolgun, and Konar 2018)

REFERENCES

- Alexander, Chief Clarence, Nora Bynum, Liz Johnson, Ursula King, Tero Mustonen, Peter Neofotis, Noel Oettlé, Cynthia Rosenzweig, Chie Sakakibara, Chief Vyacheslav Shadrin, Marta Vicarelli, Jon Waterhouse, and B. W. 2011. Linking Indigenous Knowledge and Observed Climate Change Studies. *BioScience* Vol. 61 (6):477–484.
- Allouche, Omri Tsoar, Asaf Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43 (6):1223–1232.
<https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2664.2006.01214.x>.
- Anderson, E. 1947. Field Studies of Guatemalan Maize. *Annals of the Missouri Botanical Garden* 34 (4):433–467. <http://www.jstor.org/stable/2394775>.
- Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling* 222 (15):2796–2811.
<http://dx.doi.org/10.1016/j.ecolmodel.2011.04.011>.
- Bache, S. M., and H. Wickham. 2014. magrittr: A Forward-Pipe Operator for R.
<https://cran.r-project.org/package=magrittr>.
- Beatriz Rendón-Aguilar, Verónica Aguilar-Rojas, María del Consuelo Aragón- Martínez, José Francisco Ávila-Castañeda, Luis Alberto Bernal-Ramírez, David Bravo-Avilez, Guadalupe Carrillo-Galván, Amelia Cornejo-Romero, Ernesto Delgadillo-Durán, Gilberto Hern, R. O.-P. 2015. DIVERSIDAD DE MAÍZ EN LA SIERRA SUR DE OAXACA, MÉXICO: CONOCIMIENTO Y MANEJO TRADICIONAL.

- Bellon, M. R., and S. B. Brush. 2017. Keepers of Maize in Chiapas , Mexico Author (s): Mauricio R . Bellon and Stephen B . Brush Published by : Springer on behalf of New York Botanical Garden Press Stable URL : <http://www.jstor.org/stable/4255617> JSTOR is a not-for-profit service that helps. 48 (2):196–209.
- Benz, B., H. Perales, and S. Brush. 2007. Tzeltal and Tzotzil Farmer Knowledge and Maize Diversity in Chiapas, Mexico. *Current Anthropology* 48 (2):289–300. <http://www.journals.uchicago.edu/doi/abs/10.1086/512986>.
- Biodiversidad, S. N. de I. S. Portal de Geoinformacion. *CONABIO*. <http://www.conabio.gob.mx/informacion/gis/>.
- Bird, R. M. 2010. Maize Evolution from 500 B.C. to the Present. *Missouri Botanical Garden* :3–4. <ftp://169.158.189.34/pub/Biotropica/1980s/1980/12?1/Biotropica?1980?12?1?p30.pdf>.
- Bocinsky, R. K., and T. A. Kohler. 2014. A 2,000-year reconstruction of the rain-fed maize agricultural niche in the US Southwest. *Nature Communications* :1–12.
- Booth, T. H., H. A. Nix, J. R. Busby, and M. F. Hutchinson. 2014. Bioclim: The first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Diversity and Distributions* 20 (1):1–9.
- Breiman, L. 2001. Randomforest2001. *University of California at Berkeley, Statistics Department* :1–33.
- Brown, M. E., and C. C. Funk. 2008. Climate. Food security under climate change. *Science (New York, N.Y.)* 319:580–581.
- Bruno Vilela, and Fabricio Villalobos. 2015. letsR: a new R package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*. <http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12401/abstract>.
- Brush, S. B., and H. R. Perales. 2007. A maize landscape: Ethnicity and agro-biodiversity in Chiapas Mexico. *Agriculture, Ecosystems and Environment*.

- Calenge, C. 2006. The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling* 197:1035.
- Canty, A., and B. D. Ripley. 2017. boot: Bootstrap R (S-Plus) Functions.
- Carpenter, G., A. N. Gillison, and J. Winter. 1993. Domain - a Flexible Modeling Procedure for Mapping Potential Distributions of Plants and Animals. *Biodiversity and Conservation* 2 (6):667–680.
- Cheng, J., M. Mattiuzzi, M. Sumner, J. A. Greenberg, A. Bevan, A. Shortridge, and M. R. J. Hijmans. 2016. Package “ raster .”
- Chessel, D., A. B. Dufour, and J. Thioulouse. 2004. The ade4 package- {I}- {O}ne-table methods. *R News* 4:5–10.
- Chung, U., S. Gbegbelegbe, B. Shiferaw, R. Robertson, J. I. Yun, K. Tesfaye, G. Hoogenboom, and K. Sonder. 2014. Modeling the effect of a heat wave on maize production in the USA and its implications on food security in the developing world. *Weather and Climate Extremes*.
- Colwell, R. K., and T. F. Rangel. 2009. Hutchinson’s duality: The once and future niche. *Proceedings of the National Academy of Sciences*.
- CONDE, C., C. CONDE, F. ESTRADA, B. MARTÍNEZ, O. SÁNCHEZ, and C. GAY. 2011. Regional climate change scenarios for México. *Atmósfera* 24 (1):125–140. <http://www.revistascca.unam.mx/atm/index.php/atm/article/view/23806>.
- Costa, G. C., C. Nogueira, R. B. Machado, and G. R. Colli. 2010. Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiversity and Conservation* 19 (3):883–899. <http://dx.doi.org/10.1007/s10531-009-9746-8>.
- Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random Forests for Classification in Ecology. *Ecology* 88 (11):2783–2792. <http://doi.wiley.com/10.1890/07-0539.1>.

- Cutler, E. A. and H. C. . 1942. Races of Zea Mays : I . Their Recognition and Classification. *Annals of the Missouri Botanical Garden* 29 (2):69–86+88.
<http://www.jstor.org/stable/2394331>.
- Dag, O., A. Dolgun, and N. M. Konar. 2018. onewaytests: An R Package for One-Way Tests in Independent Groups Designs. *The R Journal*.
- Dai, A. 2011. Drought under global warming: A review. *Wiley Interdisciplinary Reviews: Climate Change* 2 (1):45–65.
- Davison, A. C., and D. V Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press. <http://statwww.epfl.ch/davison/BMA/>.
- Doebley, J. F. 1990. Molecular Evidence and the Evolution of Maize. *Economic Botany* 44 (1):6–27. <http://link.springer.com/10.1007/BF02860472>.
- Dray, S., and A. B. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22 (4):1–20.
- Dray, S., A. B. Dufour, and D. Chessel. 2007. The ade4 package- $\{II\}$: $\{T\}$ wo-table and $\{K\}$ -table methods. *R News* 7 (2):47–52.
- Dyer, G. A., and A. López-Feldman. 2013. Inexplicable or Simply Unexplained? The Management of Maize Seed in Mexico. *PLoS ONE*.
- E.J. Wellhausen, L.M. Roberts, E.Hernandez X., Paul C, M. 1952. Races of Maize in Mexico. *The Bussey Institution of Harvard University*.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29 (2):129–151.

- Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. *Methods in Ecology and Evolution* 1 (4):330–342.
<http://dx.doi.org/10.1111/j.2041-210X.2010.00036.x>
<http://onlinelibrary.wiley.com/store/10.1111/j.2041-210X.2010.00036.x/asset/j.2041-210X.2010.00036.x.pdf?v=1&t=hrlhwru3&s=48e17a701a52d91fbecae111ba2dfe a8a8912197>.
- Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology* 77 (4):802–813.
- Elith, J., T. Phillips, StevenHastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17 (1):43–57.
- Esperón-Rodríguez, M., M. Bonifacio-Bautista, and V. L. Barradas. 2016. Socio-economic vulnerability to climate change in the central mountainous region of eastern Mexico. *Ambio*.
- Esquinas-Alcázar, J. 2005. Science and society: protecting crop genetic diversity for food security: political, ethical and technical challenges. *Nature reviews. Genetics* 6 (12):946–53.
<http://www.ncbi.nlm.nih.gov/pubmed/16341075>
<http://dx.doi.org/10.1038/nrg1729>
<http://www.nature.com/nrg/journal/v6/n12/abs/nrg1729.html>.
- European Parliament. Indigenous peoples and climate change. *Tyndall Center for Climate Change Research, Oxford*. <http://www.culturalsurvival.org/publications/cultural-survival-quarterly/none/indigenous-peoples-and-climate-change>
<http://www.tyndall.ac.uk/sites/default/files/Indigenous>.
- Evans, J. S. 2017. spatialEco. <https://cran.r-project.org/package=spatialEco>.
- Fick, S. E. and R. J. H. 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*.

- Fischer, G., F. O. Nachtergaele, S. Prieler, E. Teixeira, G. Toth, H. van Velthuisen, L. Verelst, and D. Wiberg. 2012. Global Agro-ecological Zones (GAEZ): Model Documentation. :1–179.
- Francois, R. 2017. bibtex: Bibtex Parser. <https://cran.r-project.org/package=bibtex>.
- Friedman, J. H. 2011. Greedy function machine: A gr. *Statistics* 29 (5):1189–1232.
- García-Martínez, Y. G., C. Ballesteros, H. Bernal, O. Villarreal, L. Jiménez-García, and D. Jiménez-García. 2016. Traditional agroecosystems and global change implications in Mexico. *Bulgarian Journal of Agricultural Science*.
- Gent, P. R., G. Danabasoglu, L. J. Donner, M. M. Holland, E. C. Hunke, S. R. Jayne, D. M. Lawrence, R. B. Neale, P. J. Rasch, M. Vertenstein, P. H. Worley, Z.-L. Yang, and M. Zhang. 2011. The Community Climate System Model Version 4. *Journal of Climate* 24 (19):4973–4991. <https://doi.org/10.1175/2011JCLI4083.1>.
- Giorgetta, M. A., J. Jungclaus, C. H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, K. Glushak, V. Gayler, H. Haak, H. Hollweg, T. Ilyina, S. Kinne, L. Kornblueh, D. Matei, T. Mauritsen, U. Mikolajewicz, W. Mueller, D. Notz, F. Pithan, T. Raddatz, S. Rast, R. Redler, E. Roeckner, H. Schmidt, R. Schnur, J. Segschneider, K. D. Six, M. Stockhause, C. Timmreck, J. Wegner, H. Widmann, K. Wieners, M. Claussen, J. Marotzke, and B. Stevens. Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems* 5 (3):572–597. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jame.20038>.
- Giraudoux, P. 2017. pgirmess: Data Analysis in Ecology. <https://cran.r-project.org/package=pgirmess>.
- Gordon, M. 2017. htmlTable: Advanced Tables for Markdown/HTML. <https://cran.r-project.org/package=htmlTable>.
- Griesemer, J. R. 1992. Niche: Historical Perspectives. In *Keywords in Evolutionary Biology*, 231–240.

- Griffies, S. M., M. Winton, L. J. Donner, L. W. Horowitz, S. M. Downes, R. Farneti, A. Gnanadesikan, W. J. Hurlin, H.-C. Lee, Z. Liang, J. B. Palter, B. L. Samuels, A. T. Wittenberg, B. L. Wyman, J. Yin, and N. Zadeh. 2011. The GFDL CM3 Coupled Climate Model: Characteristics of the Ocean and Sea Ice Simulations. *Journal of Climate* 24 (13):3520–3544. <https://doi.org/10.1175/2011JCLI3964.1>.
- Grosjean, P., and F. Ibanez. 2014. pastecs: Package for Analysis of Space-Time Ecological Series. <https://cran.r-project.org/package=pastecs>.
- Hastie, T., and W. Fithian. 2013. Inference from presence-only data ; the ongoing controversy. 1 (February):864–867.
- Hernandez-Stefanoni, J. L., J. B. Pineda, and G. Valdes-Valadez. 2006. Comparing the use of indigenous knowledge with classification and ordination techniques for assessing the species composition and structure of vegetation in a tropical forest. *Environmental Management*.
- Hernandez-Stefanoni, J. L., J. B. Pineda, and G. Valdes-Valadez. 2006. Comparing the use of indigenous knowledge with classification and ordination techniques for assessing the species composition and structure of vegetation in a tropical forest. *Environmental Management*.
- Hijmans, A. R. J., S. Phillips, J. Leathwick, J. Elith, and M. R. J. Hijmans. 2017. Package “dismo.”
- Hijmans, R. J. 2017. raster: Geographic Data Analysis and Modeling. <https://cran.r-project.org/package=raster>.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, G. Jones, and A. Jarvis. 2005. VERY HIGH RESOLUTION INTERPOLATED CLIMATE SURFACES FOR GLOBAL LAND AREAS. 1978:1965–1978.
- Hijmans, R. J., S. Phillips, J. Leathwick, and J. Elith. 2017. dismo: Species Distribution Modeling. <https://cran.r-project.org/package=dismo>.

- Hirzel, A. H., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology* 83 (7):2027–2036.
- Hoegh-Guldberg, O., and J. F. B. 2010. The Impact of Climate Change on the World's Marine Ecosystems. *Science* 328 (5985):1523–1528.
- Holt, R. D. 2009. Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences* 106 (Supplement 2):19659–19665.
<http://www.pnas.org/content/106/suppl.2/19659.abstract>
<http://www.pnas.org/content/106/suppl.2/19659.full.pdf>.
- Houghton JT, D. Y, G. DJ, N. M, van der L. PJ, D. X, M. K, and J. C. 2001. Climate Change 2001: The Scientific Basis. *Climate Change 2001: The Scientific Basis* :881.
- Hugo, P. R., S. B. Brush, and C. O. Qualset. 2003. Dynamic Management of Maize Landraces in Central Mexico. *Economic Botany* 57 (1):21–34.
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427.
- INEGI, I. N. de E. y G. 2010. Mexico's 2010 Population and Housing Census.
<http://gaia.inegi.org.mx/geoelectoral/viewer/html#> (last accessed 20 March 2017).
- Intergovernmental Panel on Climate Change. 2007. *Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report*. Geneva: IPCC.
<http://www.ipcc.ch/ipccreports/ar4-wg1.htm>.
- IPCC. 2007. *Mitigation of climate change: Contribution of working group III to the fourth assessment report of the Intergovernmental Panel on Climate Change*.
- . 2013. *Climate Change 2013*.
- Jablonski, L. M., X. Wang, and P. S. Curtis. 2002. Plant reproduction under elevated CO₂ conditions: A meta-analysis of reports on 79 crop and wild species. *New Phytologist* 156 (1):9–26.

- Jaynes, E. T. 1957. Information theory and statistical mechanics. *Physical Review* 106:620–630.
<http://libproxy.txstate.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=515974964&site=eds-live&scope=site>.
- Jones, C. D., J. K. Hughes, N. Bellouin, S. C. Hardiman, G. S. Jones, J. Knight, S. Liddicoat, F. M. O'Connor, R. J. Andres, C. Bell, K.-O. Boo, A. Bozzo, N. Butchart, P. Cadule, K. D. Corbin, M. Doutriaux-Boucher, P. Friedlingstein, J. Gornall, L. Gray, P. R. Halloran, G. Hurtt, W. J. Ingram, J.-F. Lamarque, R. M. Law, M. Meinshausen, S. Osprey, E. J. Palin, L. Parsons Chini, T. Raddatz, M. G. Sanderson, A. A. Sellar, A. Schurer, P. Valdes, N. Wood, S. Woodward, M. Yoshioka, and M. Zerroukat. 2011. The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geoscientific Model Development* 4 (3):543–570.
<https://www.geosci-model-dev.net/4/543/2011/>.
- Kang, Y, S. Khan, X. M. 2009. Climate Change Impacts on Crop Yied, Crop Water, Productivity and Food Secury - A Review. *Progress in Natural Science* 19 (12):1665–1674.
- Kassambara, A. 2017. ggpubr: “ggplot2” Based Publication Ready Plots. <https://cran.r-project.org/package=ggpubr>.
- Kato, T. a., C. Mapes, L. . M. Mera, J. a. Serratos, and R. a. Bye. 2009. *Origen y diversificación del maíz*.
- Kehoe, T. & A. W. 2003. *Corn and Capitalism: How a Botanical Bastard Grew to Global Dominance*. Chapel Hill: University of North Carolina.
- Knaus, J. 2015. snowfall: Easier cluster computing (based on snow). <https://cran.r-project.org/package=snowfall>.
- Lang, D. T., and the CRAN Team. 2017. XML: Tools for Parsing and Generating XML Within R and S-Plus. <https://cran.r-project.org/package=XML>.
- Liu, J., C. Folberth, H. Yang, J. Röckström, K. Abbaspour, and A. J. B. Zehnder. 2013. A Global and Spatially Explicit Assessment of Climate Change Impacts on Crop Production and Consumptive Water Use. *PLoS ONE* 8 (2).

- Matsuoka, Y. 2005. Origin Matters: Lessons from the Search for the Wild Ancestor of Maize. *Breeding Science* 55 (4):383–390.
<http://japanlinkcenter.org/JST.JSTAGE/jsbbs/55.383?from=Google%5Cnpapers2://publication/uuid/3EFFEC06-005F-43CD-ACCC-4A4F4ACA3C86%5Cnhttp://joi.jlc.jst.go.jp/JST.JSTAGE/jsbbs/55.383?from=CrossRef>.
- Matsuoka, Y., Y. Vigouroux, M. M. Goodman, J. Sanchez G, E. Buckler, and J. Doebley. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences of the United States of America* 99 (9):6080–6084.
- McHugh, M. L. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica* 22 (3):276–282.
- McIntire, E. J. B., and A. M. Chubaty. 2017. quickPlot: A System of Plotting Optimized for Speed and Modularity. <https://cran.r-project.org/package=quickPlot>.
- Mendelsohn, R. 2008. The Impact of Climate Change on Agriculture in Developing Countries. *Journal of Natural Resources Policy Research* 1 (1):5–19.
- Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* 36 (10):1058–1069.
- Miguel B Araújo, A. T. P. 2008. Uses and misuses of bioclimatic envelope modeling. *Ecology* 89 (10):2712–2724.
- Monterroso Rivas, A. I., C. Conde Álvarez, G. Rosales Dorantes, J. D. Gómez Díaz, and C. Gay García. 2011. Assessing current and potential rainfed maize suitability under climate change scenarios in Mexico. *Atmosfera* 24 (1):53–67.
- Morales, N. S., I. C. Fernández, and V. Baca-González. 2017. MaxEnt's parameter configuration and small samples: are we paying attention to recommendations? A systematic review. *PeerJ* 5:e3093. <https://peerj.com/articles/3093>.

- Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. Kass, M. Uriarte, and R. P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for ecological niche models. *Methods in Ecology and Evolution* 5 (11):1198–1205.
- Naimi, B., N. a.s. Hamm, T. A. Groen, A. K. Skidmore, and A. G. Toxopeus. 2014. Where is positional uncertainty a problem for species distribution modelling. *Ecography* 37:191–203.
- Neuwirth, E. 2014. RColorBrewer: ColorBrewer Palettes. <https://cran.r-project.org/package=RColorBrewer>.
- Nuss, E. T., and S. A. Tanumihardjo. 2010. Maize: A paramount staple crop in the context of global nutrition. *Comprehensive Reviews in Food Science and Food Safety* 9 (4):417–436.
- Ogle, D. H. 2017. FSA: Fisheries Stock Analysis.
- Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2017. vegan: Community Ecology Package. <https://cran.r-project.org/package=vegan>.
- Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annual of Ecology, Evolution and Systematics* 37 (May):637–669.
- Perales, H., and D. Golicher. 2014. Mapping the diversity of maize races in Mexico. *PLoS ONE* 9 (12):1–20.
- Perales, H. R., B. F. Benz, and S. B. Brush. 2005. Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proceedings of the National Academy of Sciences of the United States of America* 102 (3):949–954.
- Perales R., H., S. B. Brush, and C. O. Qualset. 2003. Landraces of Maize in Central Mexico: An Altitudinal Transect. *Economic Botany* 57:7–20.
- Perpiñán, O., and R. Hijmans. 2016. rasterVis. <http://oscarperpinan.github.io/rastervis/>.

- Peterson, A. T., and J. Soberón. 2012. Species distribution modeling and ecological niche modeling: Getting the Concepts Right. *Natureza a Conservacao*.
- Phillips, S. J., and M. Dudík. 2008. Modeling of species distribution with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31 (December 2007):161–175.
- Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, J. Leathwick, S. Ferrier, S. E. Applications, N. Jan, S. J. Phillips, M. Dud, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2016. Sample Selection Bias and Presence-Only Distribution Models : Implications for Background and Pseudo-Absence Data Stable URL : <http://www.jstor.org/stable/27645958> REFERENCES Linked references are available on JSTOR for this article : Sample selection bia. 19 (1):181–197.
- Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, J. Leathwick, S. Ferrier, S. J. Phillips, M. Dud, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2017. Sample Selection Bias and Presence-Only Distribution Models : Implications for Background and Pseudo-Absence Data Published by : Wiley on behalf of the Ecological Society of America Stable URL : <http://www.jstor.org/stable/27645958> REFERENCES Linked refer. 19 (1):181–197.
- Phillips, S. J., M. Dudík, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. *Twentyfirst international conference on Machine learning ICML 04* 69:83.
<http://portal.acm.org/citation.cfm?doid=1015330.1015412>.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. *Ecology Letters* 3 (4):349–361.
- R Core Team. 2017a. foreign: Read Data Stored by “Minitab”, “S”, “SAS”, “SPSS”, “Stata”, “Systat”, “Weka”, “dBase”, ... <https://cran.r-project.org/package=foreign>.
- . 2017b. R: A Language and Environment for Statistical Computing. <https://www.r-project.org/>.

- Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography* 41 (4):629–643.
- Revelle, W. 2017. psych: Procedures for Psychological, Psychometric, and Personality Research. <https://cran.r-project.org/package=psych>.
- Rivero-Romero, A. D., A. I. Moreno-Calles, A. Casas, A. Castillo, and A. Camou-Guerrero. 2016. Traditional climate knowledge: a case study in a peasant community of Tlaxcala, Mexico. *Journal of Ethnobiology and Ethnomedicine* 12.
- Royle, J. A., R. B. Chandler, C. Yackulic, and J. D. Nichols. 2012. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods in Ecology and Evolution* 3 (3):545–554.
- Ruiz C., J. A., J. J. Sánchez G., and M. M. Goodman. 1998. Base temperature and heat unit requirement of 49 Mexican maize races. *Maydica* 43 (4):277–282.
- Ruiz Corral, J. A., N. Durán Puga, J. D. J. Sánchez González, J. Ron Parra, D. R. González Eguiarte, J. B. Holland, and G. Medina García. 2008. Climatic adaptation and ecological descriptors of 42 Mexican maize races. *Crop Science* 48 (4):1502–1512.
- Sala, OE, Chapin, FS, Armesto, JJ, Berlow, E, Bloomfield, J, Dirzo, R, Huber-Sanwald, E, Huenneke, LF, Jackson, RB, Kinzig, A, Leemans, R, Lodge, DM, Mooney, HA, Oesterheld, M, Poff, NL, Sykes, MT, Walker, BH, Walker, M, Wall, D. 1999. Global Biodiversity Scenarios for the Year 2100. *Science* 287 (5459).
- Sánchez-Cortés, M. S., and E. L. Chavero. 2011. Indigenous perception of changes in climate variability and its relationship with agriculture in a Zoque community of Chiapas, Mexico. *Climatic Change*.
- Sarkar, D., and F. Andrews. 2016. latticeExtra: Extra Graphical Utilities Based on Lattice. <https://cran.r-project.org/package=latticeExtra>.

- Sealing, K. 2003. Indigenous Peoples, Indigenous Farmers : NAFTA's Threat to Mexican Teosinte Farmers and What Can be Done About It. *American University International Law Review* 18 (6):1383–1398. <http://digitalcommons.wcl.american.edu/cgi/viewcontent.cgi?article=1204&context=auilr>.
- Simpson, G. L. 2016. permute: Functions for Generating Restricted Permutations of Data. <https://cran.r-project.org/package=permute>.
- Sing, T., O. Sander, N. Beerenwinkel, and T. Lengauer. 2005. ROCr: visualizing classifier performance in R. *Bioinformatics* 21 (20):7881. <http://rocr.bioinf.mpi-sb.mpg.de>.
- Song, W., E. Kim, D. Lee, M. Lee, and S. W. Jeon. 2013. The sensitivity of species distribution modeling to scale differences. *Ecological Modelling* 248:113–118. <http://dx.doi.org/10.1016/j.ecolmodel.2012.09.012>.
- Stabler, B. 2013. shapefiles: Read and Write ESRI Shapefiles. <https://cran.r-project.org/package=shapefiles>.
- Stockwell, D., and D. Peters. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *int. j. geographical information science* 13 (2):143–158.
- Suzuki, R., and H. Shimodaira. 2015. pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. <https://cran.r-project.org/package=pvclust>.
- Taba, S. 1995. *Maize Genetic Resources. Maize Program Special Report. Mexico, D.F.: CIMMYT*. http://link.springer.com/chapter/10.1007/978-3-540-68922-5_14.
- Thrupp, L. A. 2000. Linking Agricultural Biodiversity and Food Security: The Valuable Role of Sustainable Agriculture. *International Affairs (Royal Institute of International Affairs 1944-)* 76 (2):265–281. <http://www.jstor.org/stable/2626366>.
- Thuiller, W., D. Georges, R. Engler, and F. Breiner. 2017. biomod2: Ensemble Platform for Species Distribution Modeling. <https://r-forge.r-project.org/projects/biomod/>.

- Thuiller, W., B. Lafourcade, R. Engler, and M. B. Araújo. 2009. BIOMOD - A platform for ensemble forecasting of species distributions. *Ecography* 32 (3):369–373.
- Tierney, L. 2011. tkrplot: TK Rplot. <https://cran.r-project.org/package=tkrplot>.
- Tierney, L., A. J. Rossini, N. Li, and H. Sevcikova. 2016. snow: Simple Network of Workstations. <https://cran.r-project.org/package=snow>.
- Title, P. O., and J. B. Bemmels. 2017. ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* (January):1–16.
- Title, P. O., and J. B. Bemmels. 2017. ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*.
- Toledo, V., and N. Barrera-Bassols. 2017. Political Agroecology in Mexico: A Path toward Sustainability. *Sustainability*.
- Tsoar, A., O. Allouche, O. Steinitz, D. Rotem, and R. Kadmon. 2007. A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions* 13 (4):397–405.
- Tubiello, F. N., and G. Fischer. 2007. Reducing climate change impacts on agriculture: Global and regional effects of mitigation, 2000-2080. *Technological Forecasting and Social Change* 74 (7):1030–1056.
- UNU-IAS, U. N. U. I. of A. S. 2008. *Indigenous Peoples and Climate Change in International Processes*.
- Urbanek, S. 2016. rJava: Low-Level R to Java Interface. <https://cran.r-project.org/package=rJava>.
- Ureta, C., C. González-Salazar, E. J. González, E. R. Álvarez-Buylla, and E. Martínez-Meyer. 2013. Environmental and social factors account for Mexican maize richness and distribution: A data mining approach. *Agriculture, Ecosystems and Environment*.

- Ureta, C., E. Martínez-Meyer, E. J. González, and E. R. Álvarez-Buylla. 2015. Finding potential high-yield areas for Mexican maize under current and climate change conditions. *Journal of Agricultural Science* :1–13.
- Ureta, C., E. Martínez-Meyer, H. R. Perales, and E. R. Álvarez-Buylla. 2012. Projecting the effects of climate change on the distribution of maize races and their wild relatives in Mexico. *Global Change Biology*.
- Ureta, C., E. Martínez-Meyer, H. R. Perales, and E. R. Álvarez-Buylla. 2012. Projecting the effects of climate change on the distribution of maize races and their wild relatives in Mexico. *Global Change Biology*.
- VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2014. SDMTTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. <https://cran.r-project.org/package=SDMTTools>.
- Walther, G., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. Fromentin, O. H. I, and F. Bairlein. 2002. Ecological responses to recent climate change. *Nature* 416.
- Warnes, G. R., B. Bolker, L. Bonebakker, R. Gentleman, W. H. A. Liaw, T. Lumley, M. Maechler, A. Magnusson, S. Moeller, M. Schwartz, and B. Venables. 2016. gplots: Various R Programming Tools for Plotting Data. <https://cran.r-project.org/package=gplots>.
- Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution* 62 (11):2868–2883.
- Watanabe, M., T. Suzuki, R. O’Ishi, Y. Komuro, S. Watanabe, S. Emori, T. Takemura, M. Chikira, T. Ogura, M. Sekiguchi, K. Takata, D. Yamazaki, T. Yokohata, T. Nozawa, H. Hasumi, H. Tatebe, and M. Kimoto. 2010. Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *Journal of Climate* 23 (23):6312–6335.

- Wei, T., and V. Simko. 2017. R package “corrplot”: Visualization of a Correlation Matrix. <https://github.com/taiyun/corrplot>.
- Wen, W., J. Franco, V. H. Chavez-Tovar, J. Yan, and S. Taba. 2012. Genetic characterization of a core set of a tropical maize race Tuxpeño for further use in maize improvement. *PLoS ONE* 7 (3):1–10.
- Weyant, J., C. Azar, M. Kainuma, and J. Kejun. 2009. *Report of 2.6 versus 2.9 Watts/m² RCPP evaluation panel*. <http://ipcc.ch/meetings/session17/doc3d.pdf%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Report+of+2.6+Versus+2.9+Watts/m2+RCPP+Evaluation+Panel#0>.
- Wickham, H. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21 (12). <http://www.jstatsoft.org/v21/i12/paper>.
- . 2016. gtable: Arrange “Grobs” in Tables. <https://cran.r-project.org/package=gtable>.
- . 2017. stringr: Simple, Consistent Wrappers for Common String Operations. <https://cran.r-project.org/package=stringr>.
- Wright, K. 2017. corrram: Plot a Correlogram. <https://cran.r-project.org/package=corrram>.
- York, N., and B. Garden. 2017. Relationships among the Mexican Races of Maize
Author (s): J. J. Sanchez G. and M. M. Goodman Published by: Springer on behalf of New York Botanical Garden Press Stable URL : <http://www.jstor.org/stable/4255410> Springer, New York Botanical Garden. 46 (1):72–85.
- Zhang, J. 2016. spaa: SPecies Association Analysis. <https://cran.r-project.org/package=spaa>.

Zimmermann, Niklaus E. Thuiller, Wilfried Guisan, A. 2017. *Habitat Suitability and Distribution Models with Applications in R* 1st ed. ed. J. Usher, Michael Saunders, Denis Peet, Robert Dobson, Andrew Adam, Paul Birks, H. J. B. Gustafsson, Lena McNeely, Jeff Paine, R. T. Richardson, David Wilson. Cambridge, United Kingdom: Cambridge University Press.
www.cambridge.org/9780521765138.