Evaluating Impacts of Climate Change on Traditional Mexican Maize

M.S. Thesis Proposal
Steven C. Gonzalez
Department of Geography
December 6th, 2017

Presented to:
Dr. Russell Weaver (Chair)
Dr. Jennifer Jensen
Dr. Thomas Ballinger

Thesis Proposal Statement

Climate change will have substantial effects on human and natural systems worldwide without significant intervention. Climate change's nuanced effects on historical temperature and precipitation patterns are expected to particularly influence agriculture. Maize, a globally important crop, is expected to be negatively influenced by climate change, especially within the tropics. Evaluating the impacts of climate change on maize requires interdisciplinary research to ensure global food security.

As a domesticated crop -- cultivated over 9,000 years ago by indigenous communities in the highlands of Mexico -- maize diversity and distribution is inherently contingent upon human influences. Although maize productivity and suitability is largely defined by climatic conditions, a more full representation of maize niches should incorporate the influences of societies on maize distribution; especially as mounting literature indicates significant influences of ethnolinguistic diversity on maize diversity in Mexico. Mexico presently harbors sixty-eight declared indigenous macro-languages, with numerous unique dialects within each macro-language. Although Mexico's ethnolinguistic diversity is substantial, its stability is dubious.

This research is driven by the understanding that indigenous climatic knowledge can offer significant contributions to climatic research. Further, as agriculture contributes upwards of 25% of greenhouse gas (GHG) emissions; indigenous agro-technological knowledge may contribute to sustainable agriculture development, thereby reducing agricultural GHGs.

This thesis aims to model current and future traditional maize landrace niches to evaluate the magnitude of influences of climate change on maize distributions and diversity in Mexico. In doing so, this research hopes to identify the ethnolinguistic groups that may be most heavily affected by maize distribution changes.

Outline of Presentation

- Introduction
    - Climate Change Effects on Agriculture, Maize
    - Mexico: Maize's Center of Domestication
    - Maize Diversity and Evolution
    - Maize as a Mirror of Humanity
- Purpose Statements
- Research Questions
- Theoretical Background
    - Ecological Niche
        - Grinnellean, Eltonian and Hutchinsonian Niches
        - Fundamental v.s. Realized Ecological Niches
        - BAM Diagram
- Methodological Framework
    - Ensemble Ecological Niche Modelling
- Data
    - Georeferenced Maize Observations
    - Climatological Data and Processing
        - WorldClim
            - Current – WorldClim 2.0
            - Future – WorldClim 1.4
        - Mean Ensembling
        - 'Bioclimatic' Variables
            - 'dismo' & 'envirem'
        - Topographic Variables
        - Land-cover and Soils
    - Ethnolinguistic Diversity
- Proposed Methods
    - Variable Thinning with 'usdm'
    - Ecological Niche Modelling in 'biomo2'
        - Models Available
        - Evaluation Metrics
        - Pseudo-Absence Selection
    - LEAP HPC Cluster
    - Subsequent Analyses
- Novelty of Research
- Bibliography

Purpose Statements:

1. Model current relationships between maize landraces and their niches
2. Project maize niches into geographic space
3. Forecast these projections with future climate models
4. Identify patterns
5. Evaluate spatiotemporal relationships between indigeneity and maize diversity
6. Quantify maize landrace diversity and its projected change Identify similarities among species occurrence patterns
7. Geo-visualization
8. Provide the groundwork for future work

Research Questions:

1. How do social and environmental factors contribute to maize landraces' current distribution?
2. What are the characteristics of each maize landraces ''bioclimate envelope''?
3. Can the contributions of indigeneity and society to maize landraces distribution and diversity be discerned and modelled at the national level?
4. What is the nature of the relationship between maize landrace diversity and indigeneity in Mexico?
5. How is climate change projected to alter current maize landrace distributions?
6. Which maize landraces are susceptible of extinction/expansion?

Data

Ethnolinguistic Diversity

- 2000 indigenous population percentage by *municipio*
- 1990 1st – 4th major indigenous language by *municipio,* provided by CONABIO.

Georeferenced Maize Observations

- Roughly 18,000 cleaned maize observation records for 64 unique maize landraces, provided by CONABIO.

Climatological Data and Processing

Current (1970-2000) climatological data derived from WorldClim 2.0. Future (2041 – 2060; 2061 – 2080) climatological data derived from WorldClim 1.4.

- WorldClim 2.0 provides monthly mean Tmin, Tmax, Tmean, Precipitation, and Solar Radiation rasters, among other variables
- WorldClim 1.4 provides only monthly mean Tmin, Tmax, and Precipitation rasters

Future data was created by averaging RCP 8.5 data across 5 GCMs:

- CCSM4 (Community Climate System Model, UCAR)
- MIROC5 (Model for Interdisciplinary Research on Climate)
- MPI-ESM-LR (Max-Plank Institute)
- HADGEM2-ES (Met Office Hadley)
- GFDL-CM3 (Geophysical Fluid Dynamics Laboratory)

Currently, 19 bioclimatic variables have been processed with the 'dismo' package in R for 1970 – 2000, 2041 – 2060, and 2061 – 2080 from Tmin, Tmax, and Precipitation monthly means, including:

- BIO1 = Annual Mean Temperature
- BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
- BIO3 = Isothermality (BIO2/BIO7) (* 100), BIO4 = Temperature Seasonality (standard deviation *100)
- BIO5 = Max Temperature of Warmest Month
- BIO6 = Min Temperature of Coldest Month
- BIO7 = Temperature Annual Range (BIO5-BIO6)
- BIO8 = Mean Temperature of Wettest Quarter
- BIO9 = Mean Temperature of Driest Quarter
- BIO10 = Mean Temperature of Warmest Quarter
- BIO11 = Mean Temperature of Coldest Quarter
- BIO12 = Annual Precipitation
- BIO13 = Precipitation of Wettest Month
- BIO14 = Precipitation of Driest Month
- BIO15 = Precipitation Seasonality (Coefficient of Variation)
- BIO16 = Precipitation of Wettest Quarter
- BIO17 = Precipitation of Driest Quarter
- BIO18 = Precipitation of Warmest Quarter
- BIO19 = Precipitation of Coldest Quarter

Additional, complementary bioclimatic variables are available using 'envirem' R package; however, this dataset requires Tmin, Tmax, Tmean, Precipitation, and Solar Radiation monthly means. The future WorldClim 1.4 data sets provide only Tmin, Tmax, and Precipitation. Therefore, in order to use, must:

- Average Minimum Temperature and Maximum Temperature monthly means to derive Mean Temperatures monthly means for 2041 – 2060 & 2061 -2080 (.99 R^2 with values for 1960 – 2000)
- Must use 1970 – 2000 solar radiation monthly data for in future (2041 – 2060 & 2061 -2080) raster stacks

'envirem' variables include:

1. Annual PET
2. Thornthwaite Aridity Index
3. Climatic Moisture Index
4. Continentality
5. EmbergerQ
6. Growing Deg Days 0
7. Growing Deg Days 5
8. Growing Deg Days 10
9. Max Temp Coldest
10. Min Temp Warmest
11. Month Count By Temp 10
12. PET Coldest Quarter
13. PET Driest Quarter
14. PET seasonality
15. PET Warmest Quarter
16. PET Wettest Quarter
17. Thermicity Index

Terrain Indices

Aspect, slope and roughness (difference between the maximum and the minimum value of a cell and its 8 surrounding cells) calculated with 'raster' package in R.

Land-cover and Soils

FAO Harmonized Soil Database 1.2 (Fischer et al., 2008)

- Soil Quality Data (Ordered Factor 1-4; 'No or slight limitations' to 'Very severe limitations')
    1. Nutrient availability
    2. Nutrient retention capacity
    3. Rooting conditions
    4. Oxygen availability to roots
    5. Excess salts
    6. Toxicity
    7. Workability (constraining field management)
- Land cover (Continuous)
    1. rain-fed cultivated land
    2. irrigated cultivated land, according to GMIA 4.0
    3. total cultivated land
    4. forest land, calibrated to FRA2000 land statistics
    5. grass/scrub/woodland
    6. built-up land (residential and infrastructure)
    7. barren/very sparsely vegetated land
    8. Mapped Water Bodies

Proposed Methods

Variable Thinning with 'usdm'

Multicolinear predictor variables are problematic in ecological niche modelling, especially when projecting to future climatologies that may have altered correlation structures. In order to reduce collinearity amongst predictor variables, I plan to use the 'usdm' package in R to select variables with Variance Inflation Factors (VIFs) less than 10, as suggested by (Guisan et al., 2017).

Ecological Niche Modelling in 'biomod2'

I propose using the 'biomod2' package in R, an object-oriented package for ensemble niche modelling, to predict maize landrace niches. The 'biomod2' package supports 10 models:

1. GLM : Generalized Linear Model (glm)
2. GAM : Generalized Additive Model (with gam)
3. GBM : Generalized Boosting Model or usually called Boosted Regression Trees (gbm)
4. CTA: Classification Tree Analysis (rpart)
5. ANN: Artificial Neural Network (nnet)
6. SRE: Surface Range Envelop or usually called BIOCLIM
7. FDA: Flexible Discriminant Analysis (fda)
8. MARS: Multiple Adaptive Regression Splines (earth)
9. RF: Random Forest (randomForest)
10. MAXENT.Phillips: Maximum Entropy (http://www.cs.princeton.edu/~schapire/ maxent/)
11. MAXENT.Tsuruoka: low-memory multinomial logistic regression (maxent)

Evaluation Metrics Available

Each model is evaluated according to selected evaluation metrics. Evaluation metrics[1] available include:

1. 'ROC' : Relative Operating Characteristic
2. 'KAPPA' : Cohen's Kappa (Heidke skill score)
3. 'TSS' : True skill statistic (Hanssen and Kuipers discriminant, Peirce's skill score)
4. 'FAR' : False alarm ratio
5. 'SR' : Success ratio
6. 'ACCURANCY' : Accuracy (fraction correct)
7. 'BIAS' : Bias score (frequency bias)
8. 'POD' : Probability of detection (hit rate)
9. 'CSI' : Critical success index (threat score)
10. 'ETS' : Equitable threat score (Gilbert skill score).

Ensemble Modelling in 'biomod2'

The ensemble modelling function in 'biomod2' combines models to create ensemble predictions used with forecasting. Weighted mean ensembles will be created according to each selected evaluation

---

[1] More information on each evaluation metric available at:
http://www.cawcr.gov.au/projects/verification/#Methods_for_dichotomous_forecasts

metric using an evaluation threshold of 0.7 or 0.3 (0.7 for evaluation statistics with 1 being a perfect score, and 0.3 for evaluation statistics with 0 being a perfect score [i.e.: False Alarm Ratio]). A 'total consensus' model will be created for each maize landrace by averaging each weighted-mean model created for each evaluation metric.

Binary predictions will be created with the ensemble modelling forecasting function for each evaluation statistic. These binary predictions from the mean of ensemble forecasts will be used to create binary predictions for each maize landrace, used for subsequent analyses (e.g.: range change, alpha-diversity, etc.)

The binary predictions from the ensemble forecasts will be subsequently used to create 'Committee Averaging' maps, created by averaging binary predictions. Committee Averaging maps give a sense of both prediction and uncertainty across models. When the prediction is close to 0 or 1, all models agree to predict 0 and 1 accordingly. When the prediction is around 0.5, half of the models predict 1 and half predict 0, indicating uncertainty across models.