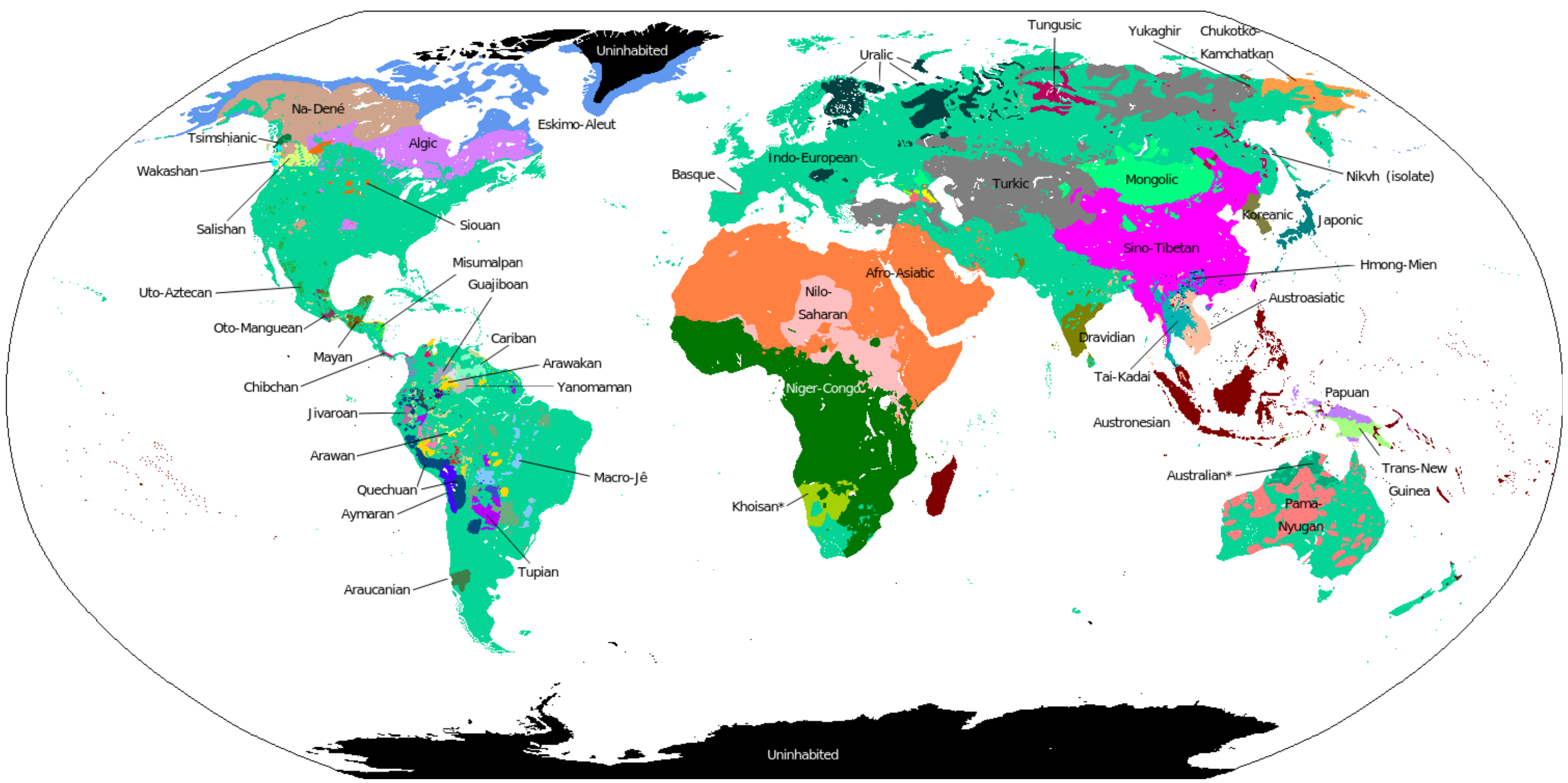
Using Global Hotspots to assess the Effects of Climate Change on Linguistic Diversity



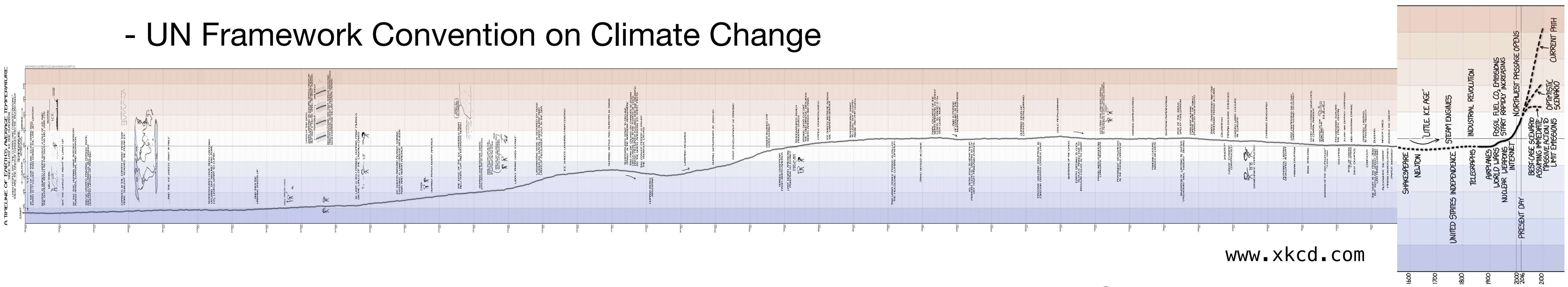
Ethan Gotlieb Wilcox
Harvard University, Department of Linguistics

CPL Lab Meeting
October 2019

Intro: Language & Climate Change

- Climate Change: “Change of climate that is attributed directly or indirectly to human activity that alters the composition of the global atmosphere”

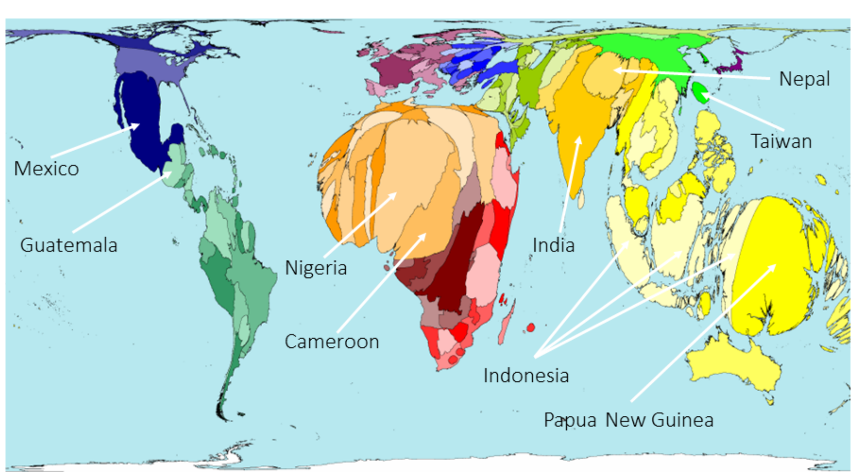
  - UN Framework Convention on Climate Change



www.xkcd.com

- Impact on Biological Ecosystems & Human Life

World’s Languages



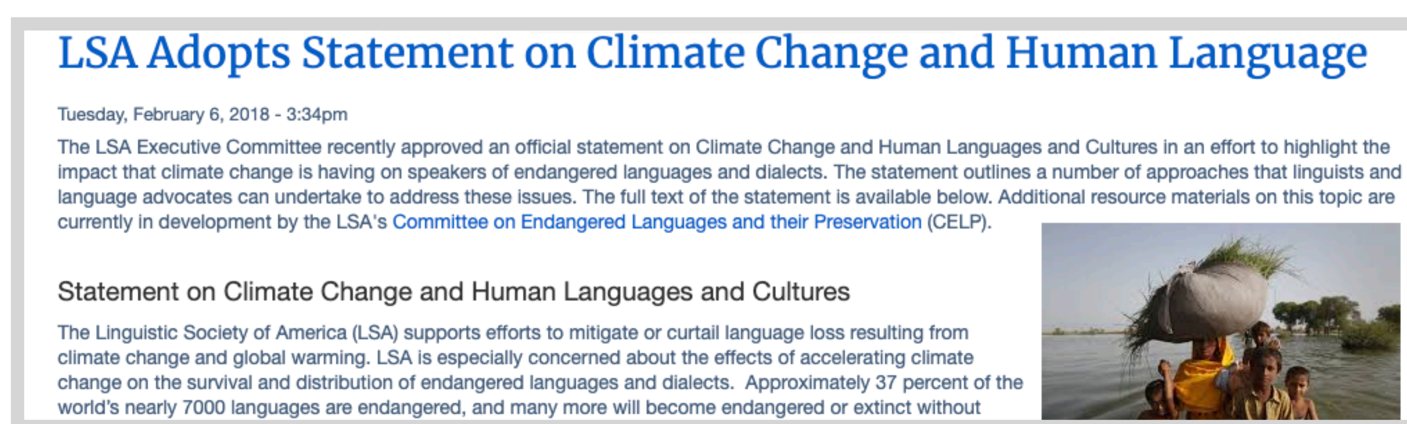
worldmapper.org

- ~6,500 languages
- 148 (+) top-level language families
- Up to 1/2 of languages may go extinct by end of century
- Major threats: economic and cultural globalization, industrialization

How does Climate Change affect this dynamic?

Motivating the Study of Climate Change on Languages

- **Linguists:**
  - Diverse data is the basis of good linguistic theory
  - Models of climate impact can transfer to other subfields, e.g. historical linguistics & typology
- **Policymakers:** Where should we fund documentation and conservation efforts?
- **Everyone:**
  - Understand the scope of the changing climate
  - Impact on linguistic diversity as a window into impact on cultural diversity



LSA Adopts Statement on Climate Change and Human Language

Tuesday, February 6, 2018 - 3:34pm

The LSA Executive Committee recently approved an official statement on Climate Change and Human Languages and Cultures in an effort to highlight the impact that climate change is having on speakers of endangered languages and dialects. The statement outlines a number of approaches that linguists and language advocates can undertake to address these issues. The full text of the statement is available below. Additional resource materials on this topic are currently in development by the LSA's Committee on Endangered Languages and their Preservation (CELP).

Statement on Climate Change and Human Languages and Cultures

The Linguistic Society of America (LSA) supports efforts to mitigate or curtail language loss resulting from climate change and global warming. LSA is especially concerned about the effects of accelerating climate change on the survival and distribution of endangered languages and dialects. Approximately 37 percent of the world's nearly 7000 languages are endangered, and many more will become endangered or extinct without

https://www.linguisticsociety.org/news/2018/02/06/lsa-adopts-statement-climate-change-and-human-language

Motivating Intuition: Climate & Linguistic Diversity

- **Intuition**: Climate change will impact linguistic diversity by putting areas of high linguistic diversity under outsized environmental stress

Importance of Islands

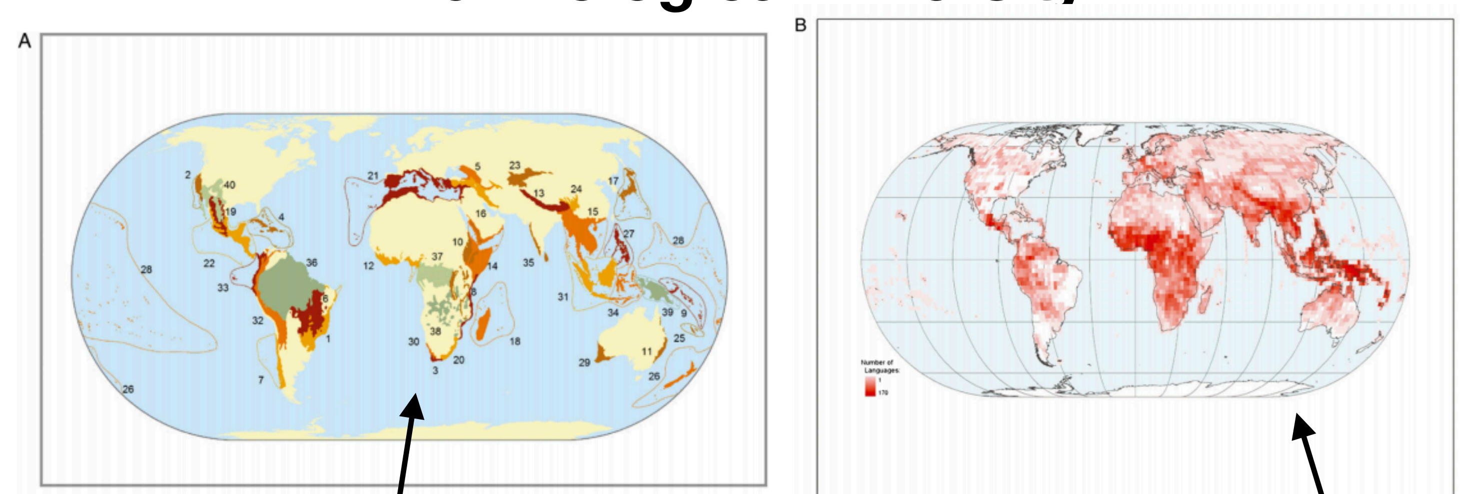
Table 1. The number of species and languages in each category and their threat status.

		Insular	Continental	Total	Percentage insular
Area (in 1000 square kilometers)		7821	141,118	140,939	5.3
Total diversity	Magnoliophyta	48,331	241,669	290,000	17
	Aves	1947	8117	10,064	19
	Rodentia	[illegible]	[illegible]	[illegible]	[illegible]
	Languages	2551	6885	9436	27
Extinct diversity	Plantae	71	43	114	62.3
	Animalia total	461	296	757	60.9
	Actinopterygii	4	99	103	3.9
	Amphibian	21	18	39	53.8
	Arthropoda	38	31	69	55.1
	Aves	127	10	137	95.3
	Mammalia	42	36	78	53.8
	Mollusca	212	97	309	68.6
	Platyhelminthes	0	1	1	0
	Reptilia	17	4	21	81.0
	Languages	43	336	379	11.4
Critically endangered diversity	Plantae	794	781	1575	50.4
	Animalia total	641	1100	1741	36.8
	Actinopterygii	49	240	289	16.9
	Amphibian	102	382	484	21.1
	Arthropoda	94	85	179	52.5
	Aves	113	79	192	58.8
	Chondrichthyes	0	25	25	0
	Mammalia	85	103	188	45.2
	Mollusca	154	137	291	52.9
	Reptilia	44	49	93	47.3
	Languages	113	347	460	24.6

Note: See the "Methods" section for references.

Tershy et al., 2015 "The Importance of Islands for the Protection of Biological and Linguistic Diversity"

Correlation of Linguistic Diversity & Biological Diversity



Corenflo et. al., 2012, "Co-occurrence of linguistic and biological diversity in biodiversity hotspots and high biodiversity wilderness areas"

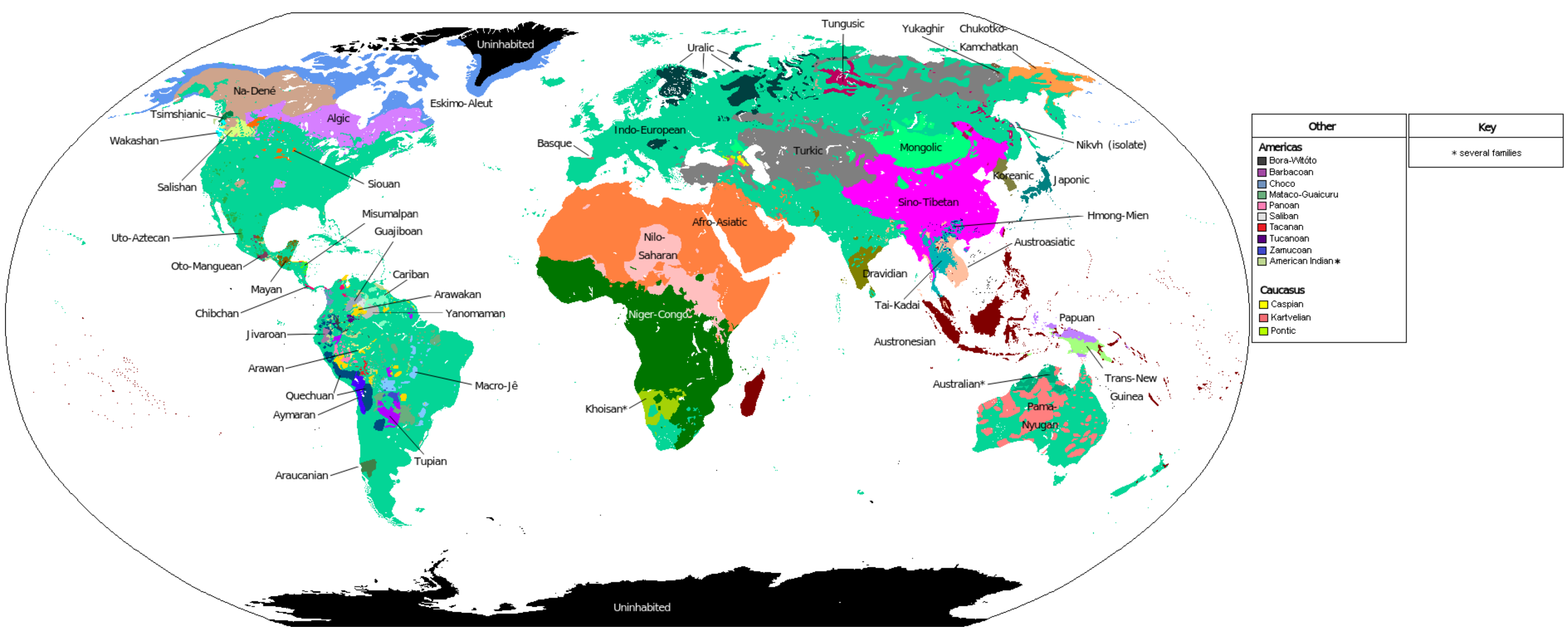
Global biodiversity hotspots

Heatmap of world's languages

Goal: Formalize this intuition by looking at global "hotspots" a state-of-the-art climate model

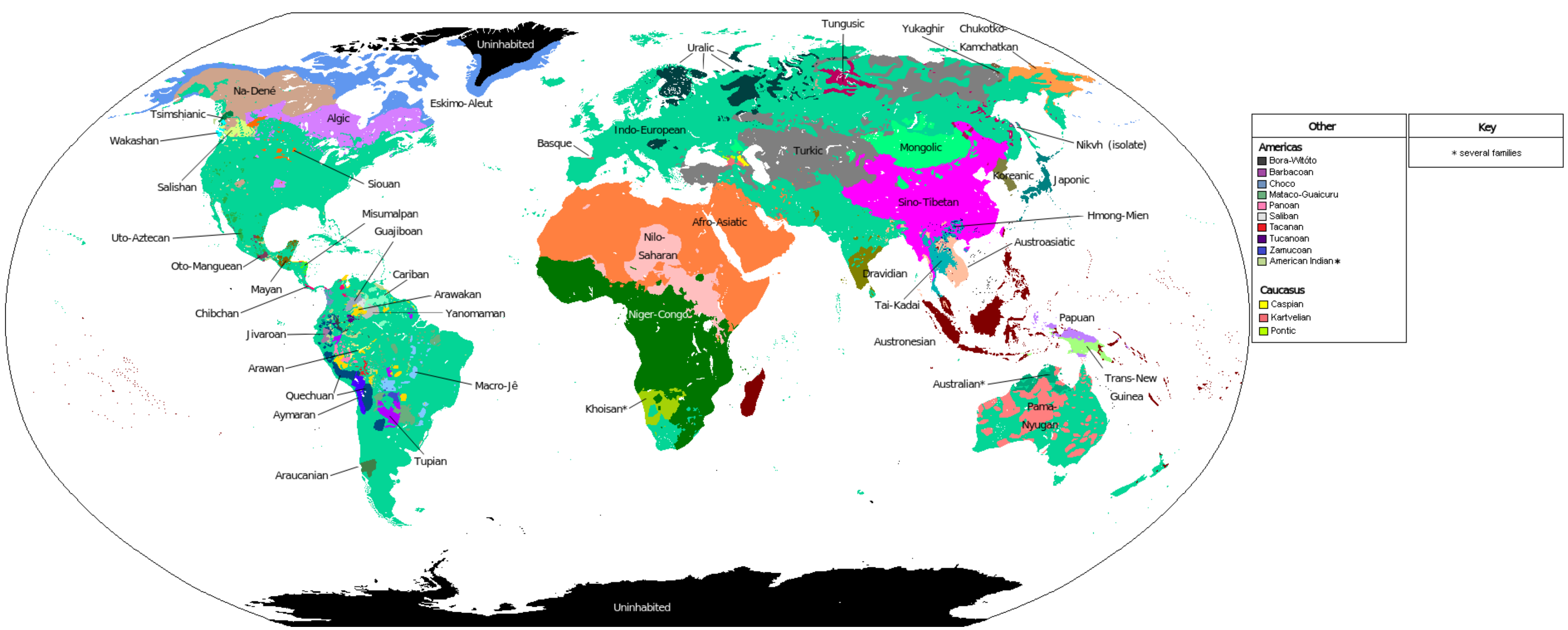
Outline: Desiderata / Questions

- Finding Climate Hotspots
- Will climate change have an outsized impact on linguistic diversity?
- What languages / language families will be impacted the most?
    - Do different climate pathways affect the same or different languages?
- Are endangered languages at greater risk from climate change?



Outline: Desiderata / Questions

- Finding Climate Hotspots
- Will climate change have an outsized impact on linguistic diversity?
- What languages / language families will be impacted the most?
    - Do different climate pathways affect the same or different languages?
- Are endangered languages at greater risk from climate change?



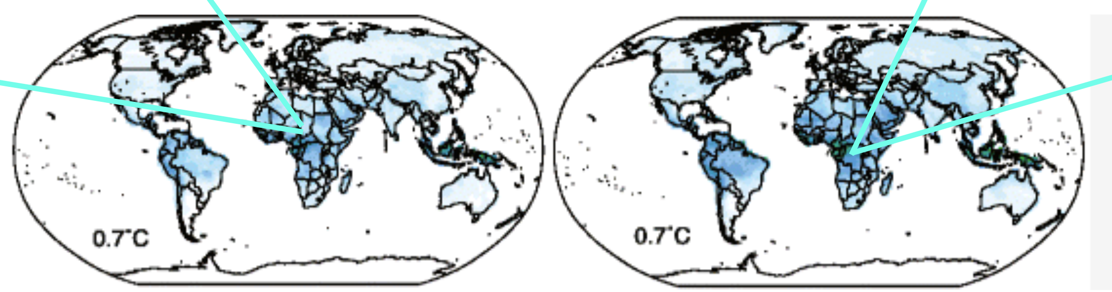
Finding Climate Hotspots: Previous Models

“Climate Change Hotspots in the CMIP5 Global Climate Model Ensemble” Diffenbaugh & Giorgi, 2012

1. Mean temperature
2. Mean precipitation
3. Inter-annual temp var
4. Inter-annual precip var
5. # Months extreme heat
6. # Months extreme wet
7. # Months extreme dry

Climate Risk = Euclidian Distance between Historical & Future matrix

1. Mean temperature
2. Mean precipitation
3. Inter-annual temp var
4. Inter-annual precip var
5. # Months extreme heat
6. # Months extreme wet
7. # Months extreme dry



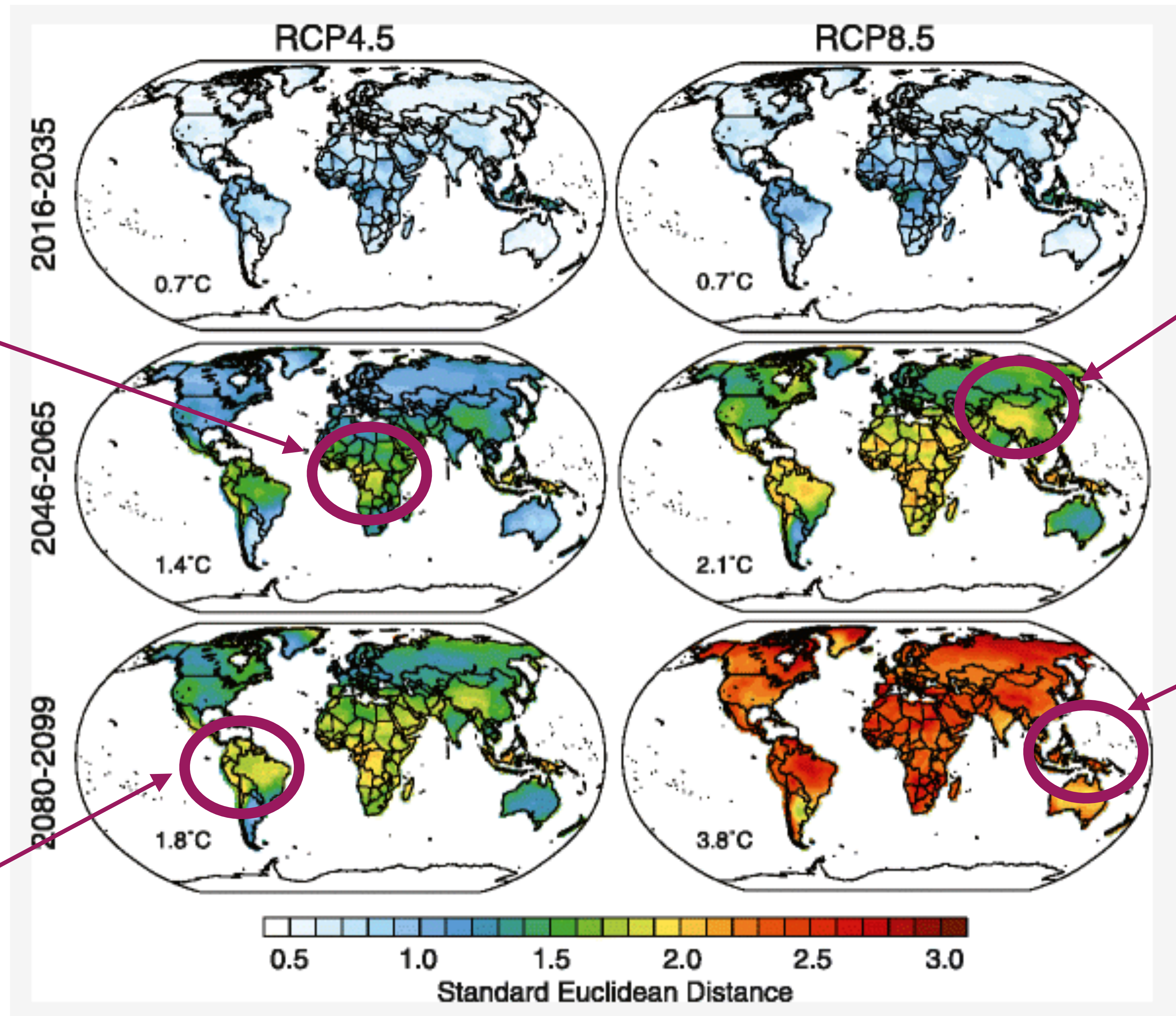
Historical Period **Future Period**

- *Historical Period:* 1986-2005 & 3 *Future* periods (2016-2035, 2046-2065, 2080-2090)
- 20 Models - CMIP5 Framework
- Use forcing function that corresponds to predicted carbon outputs:
    - RCP4.5 (650ppm/2.4 degrees warming)
    - RCP8.5 (>1370ppm co2 / 4.9 degrees warming)

Finding Climate Hotspots: Previous Results

"Climate Change Hotspots in the CMIP5 Global Climate Model Ensemble" Diffenbaugh & Giorgi, 2012

Global Hotspots:



Central Africa

Tibetan Plateau

Indonesia

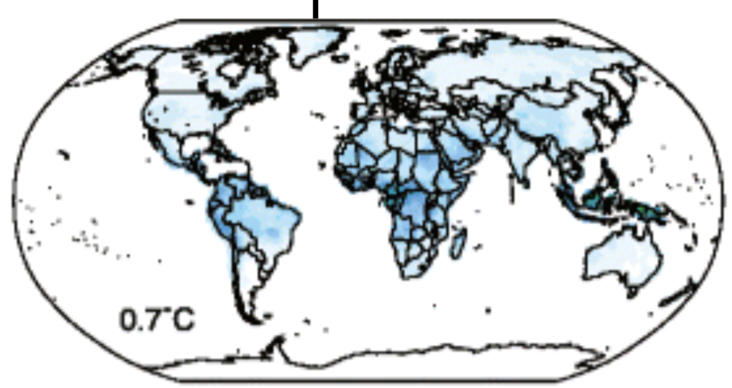
Western Amazonia

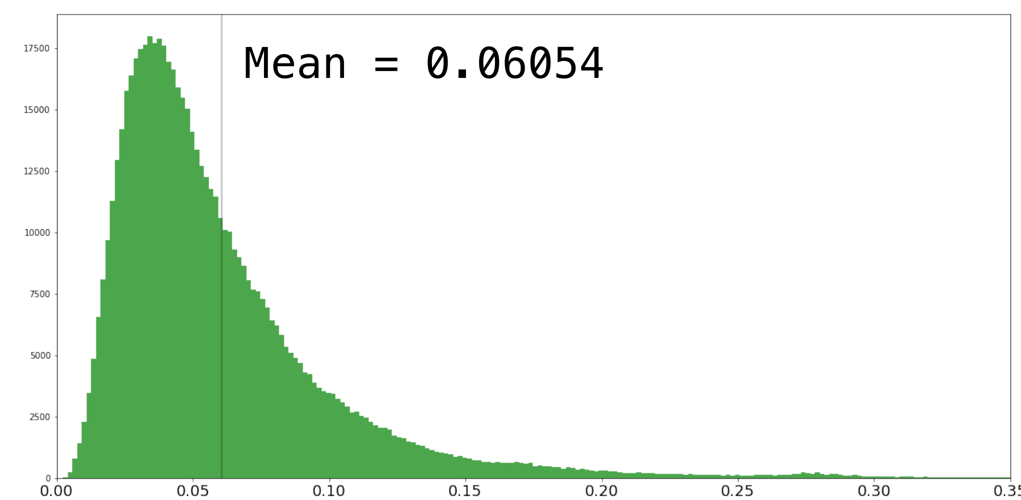
Finding Climate Hotspots: The Model

1. Normalized abs(mean $temp_p$ − mean $temp_b$)
2. Normalized abs(mean $precip_p$ − mean $precip_b$)
3. Normalized abs(mean SD($temp_p$) − SD($temp_b$))
4. Normalized abs(mean SD(precip) − SD(precip))
5. Normalized # of months hotter than hottest in baseline
6. Normalized # of months wetter than wettest in baseline
7. Normalized # of months drier than driest in baseline

<u>"Climate Risk Factor"</u>

- Mean of all seven predictors, from 0-1
- Mean normalized deviation from baseline across all 7 predictors
- Gives relative distribution of climate change across all globe points and future pathways





Histogram of Climate Risk Factor

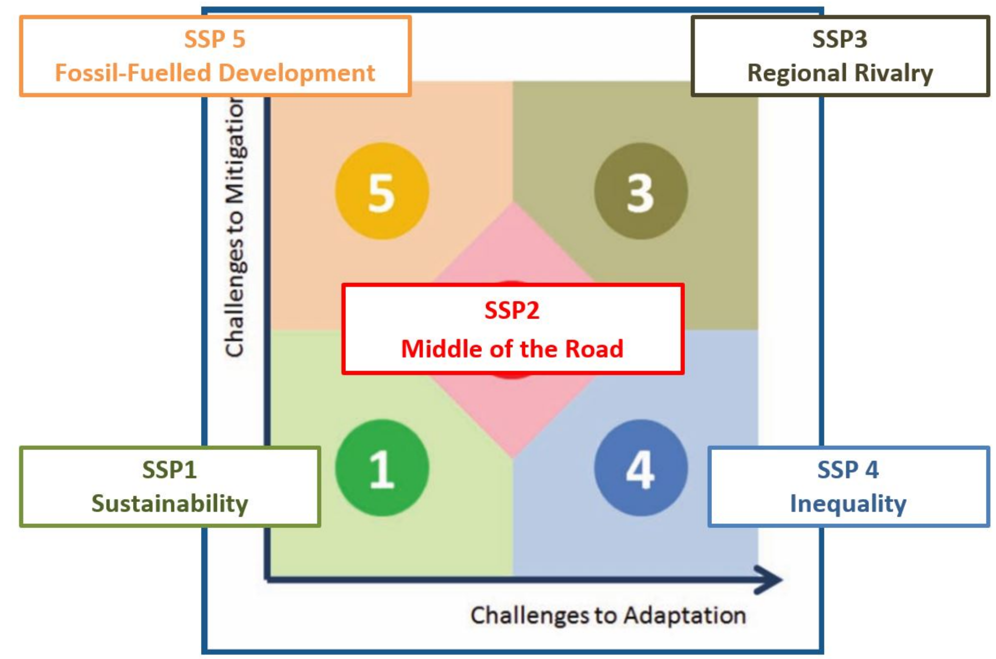
- CMIP 6 Framework
- Model: GDFL-NOAA Earth Systems Model 4 (results from this year)
- *Baseline Period:* 2005-2020 (model predictions) & 3 *Future* periods (2040-2060; 2060-2080; 2080-2100)
- Use "Shared Socioeconomic Pathways" instead of RCP Forcing Functions

Shared Socioeconomic Pathways (SSPs)

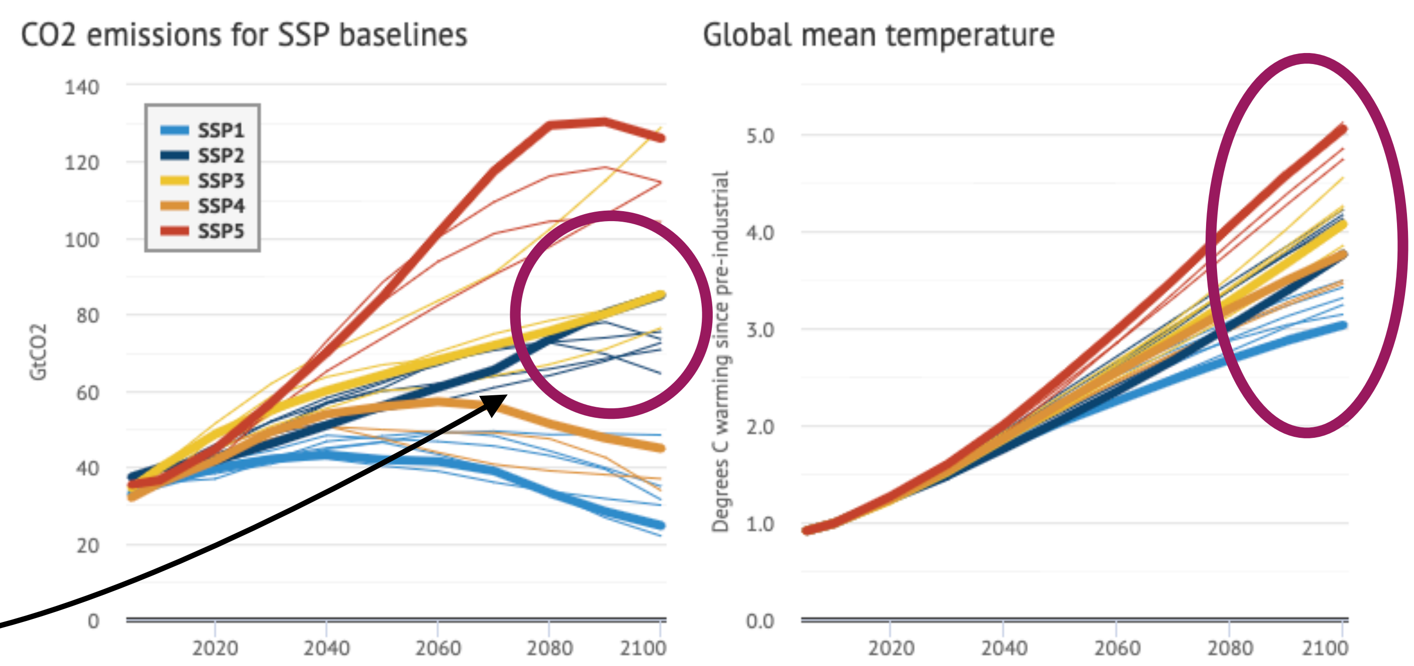
- Late 2010s: Switch from *Carbon* Output (RCPs) to **Climate Narratives** (SSPs)

Future is an interplay between…

- Callenges to **mitigation**
  - Can we reduce carbon output
- Challenges to **adaptation**
  - Shared response & coordination



- SSPs aren't strictly "better or worse"
- E.g. SSP2 & SSP3 predict more emissions than SSP4



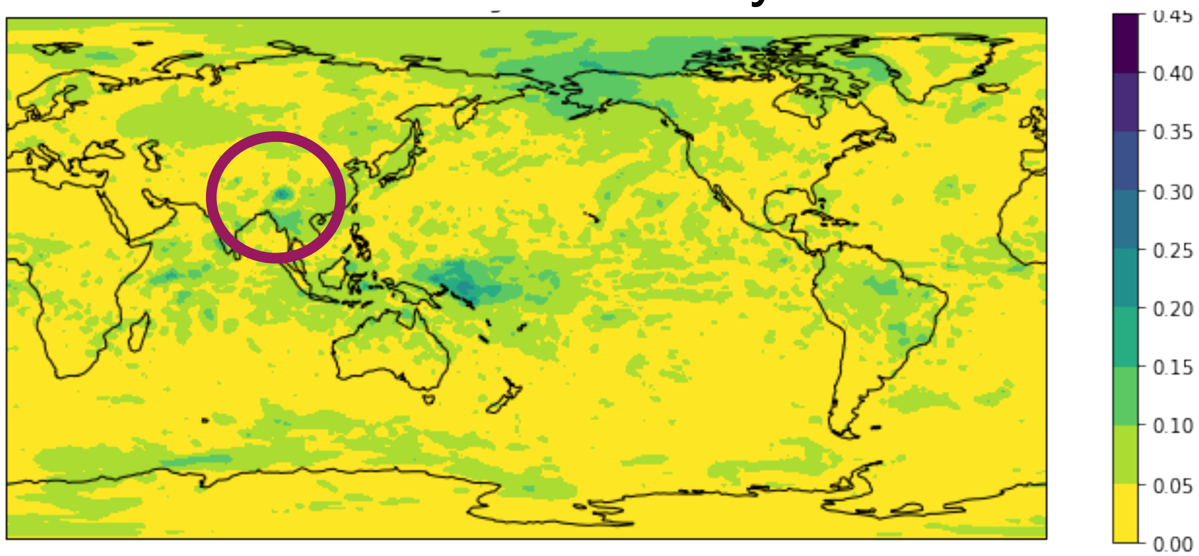
From carbonbrief.org

- Largest effects between pathways > 2060
- This work: 2015 - 2090 difference

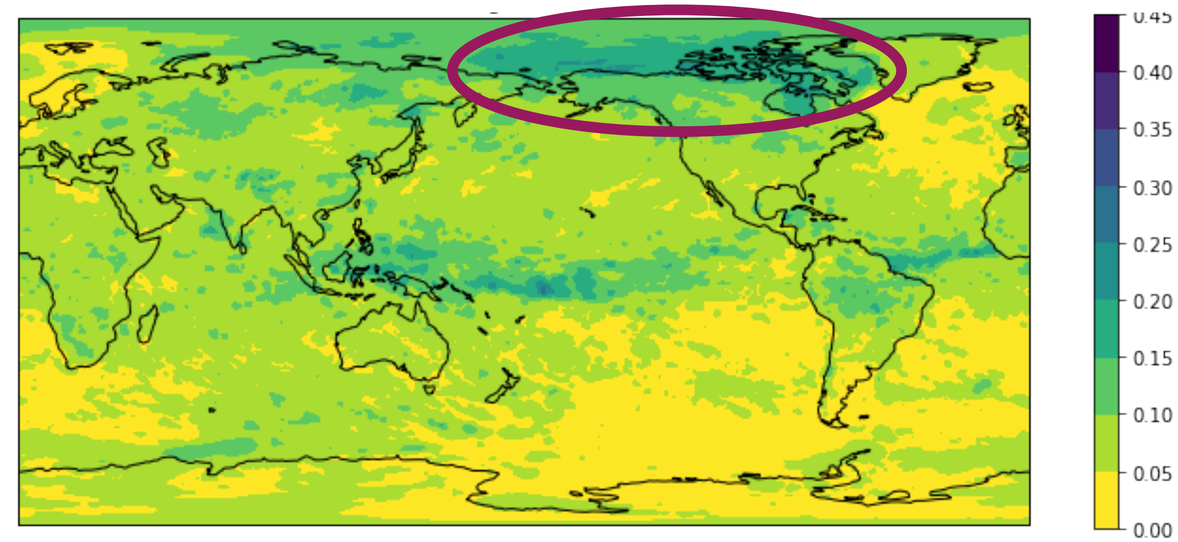
Finding Climate Change Hotspots

- Results: **Climate Risk Factor** (darker) = Deviation from Baseline Period
- Mean climate risk factor across all pathways = 0.06

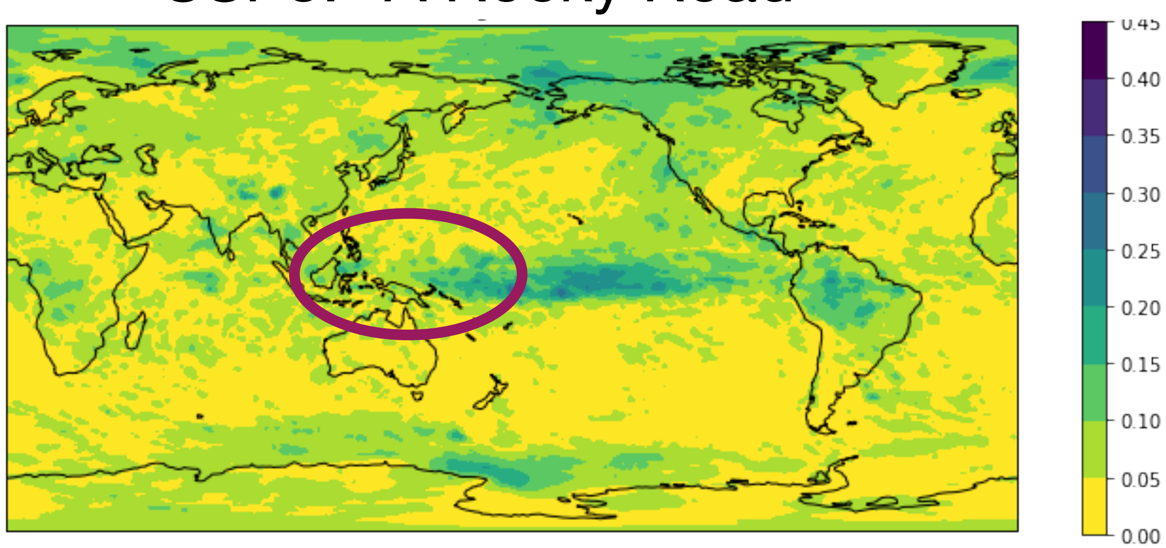
SSP1: "Sustainability"



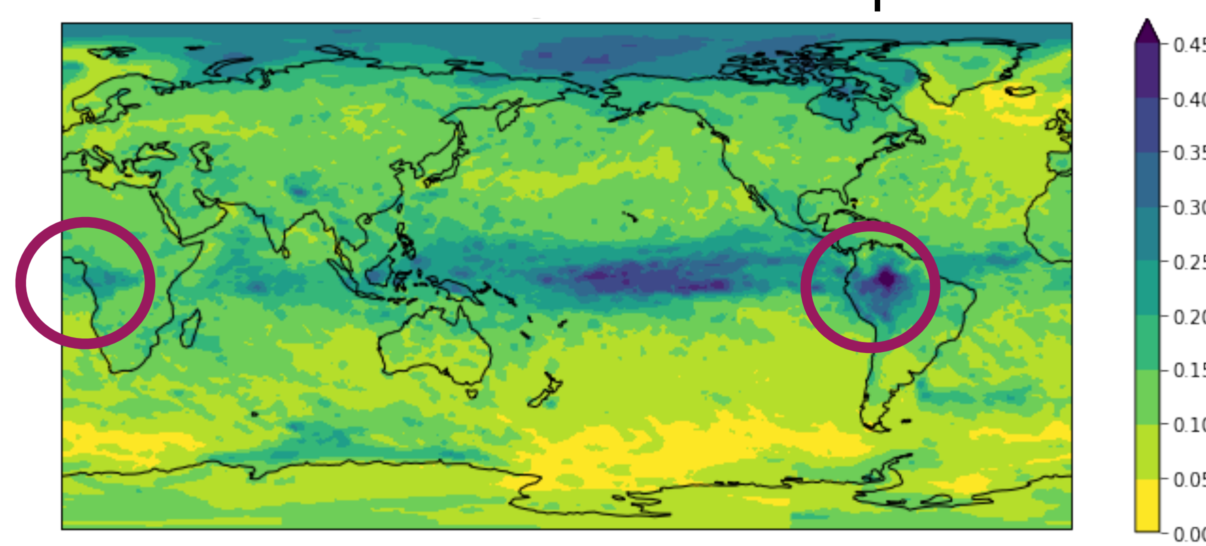
SSP2: "Middle of the Road"



SSP3: "A Rocky Road"



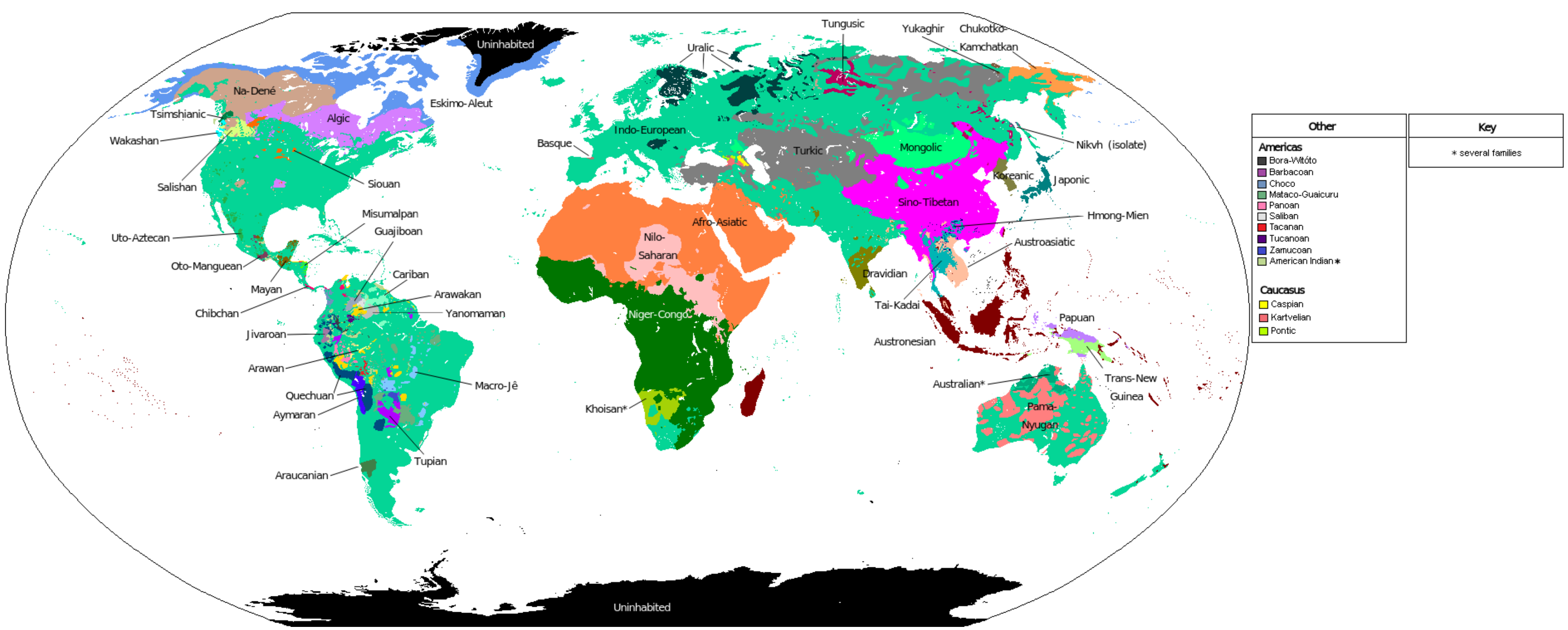
SSP5: "Fossil-Fueled Development"



Global Hotspots: Indonesia, Western Amazon, East Africa, Tibet & Arctic

Outline: Desiderata / Questions

- Finding Climate Hotspots
- Will climate change have an outsized impact on linguistic diversity?
- What languages / language families will be impacted the most?
    - Do different climate pathways affect the same or different languages?
- Are endangered languages at greater risk from climate change?

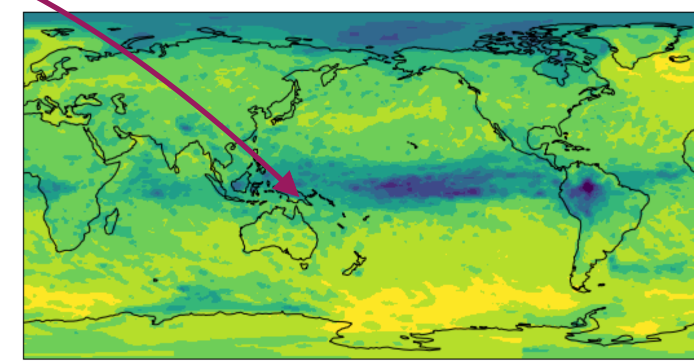


Language Dataset

- **Language Dataset:** Ethnologue 2019

	ISO_639	Language_Name	Region_Name	L1_Users	Digits	Countries	Family	Latitude	Longitude	EGIDS	Status
0	aaa	Ghotuo	Western Africa	9000.0	4.0	1	Niger-Congo	7	5	6a	Vigorous
1	aab	Alumu-Tesu	Western Africa	7000.0	4.0	1	Niger-Congo	9	8	6a	Vigorous
2	aac	Ari	Melanesia	50.0	2.0	1	Trans-New Guinea	-8	141	6b	In_Trouble
3	aad	Amal	Melanesia	830.0	3.0	1	Sepik	-5	141	6a	Vigorous
4	aae	Albanian, Arbëreshë	Southern Europe	100000.0	6.0	1	Indo-European	38	16	6b	In_Trouble
5	aaf	Aranadan	Southern Asia	200.0	3.0	1	Dravidian	11	75	6b	In_Trouble
6	aag	Ambrak	Melanesia	290.0	3.0	1	Torricelli	-4	142	6a	Vigorous
7	aai	Miniafia Oyan	Melanesia	3470.0	4.0	1	Austronesian	-10	148	5	Developing
8	aak	Ankave	Melanesia	1600.0	4.0	1	Trans-New Guinea	-8	145	5	Developing
9	aal	Afade	Middle Africa	5000.0	4.0	2	Afro-Asiatic	12	14	6a	Vigorous

SSP5



Desiderata 1: Will climate change have an outsized impact on linguistic diversity?

`Geographic Distribution`

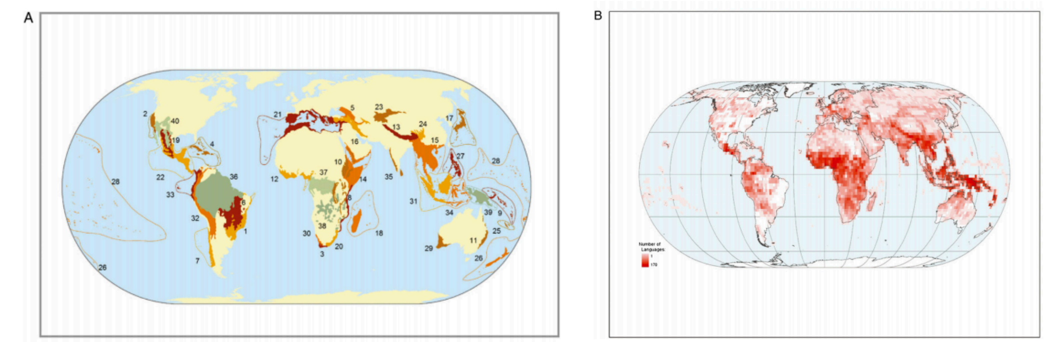
Analysis 1: Correlation between climate risk and linguistic diversity

Analysis 3: Total exposure to risk across 4 pathways.

Analysis 2: Counterfactual linguistic distribution vs. actual linguistic distribution

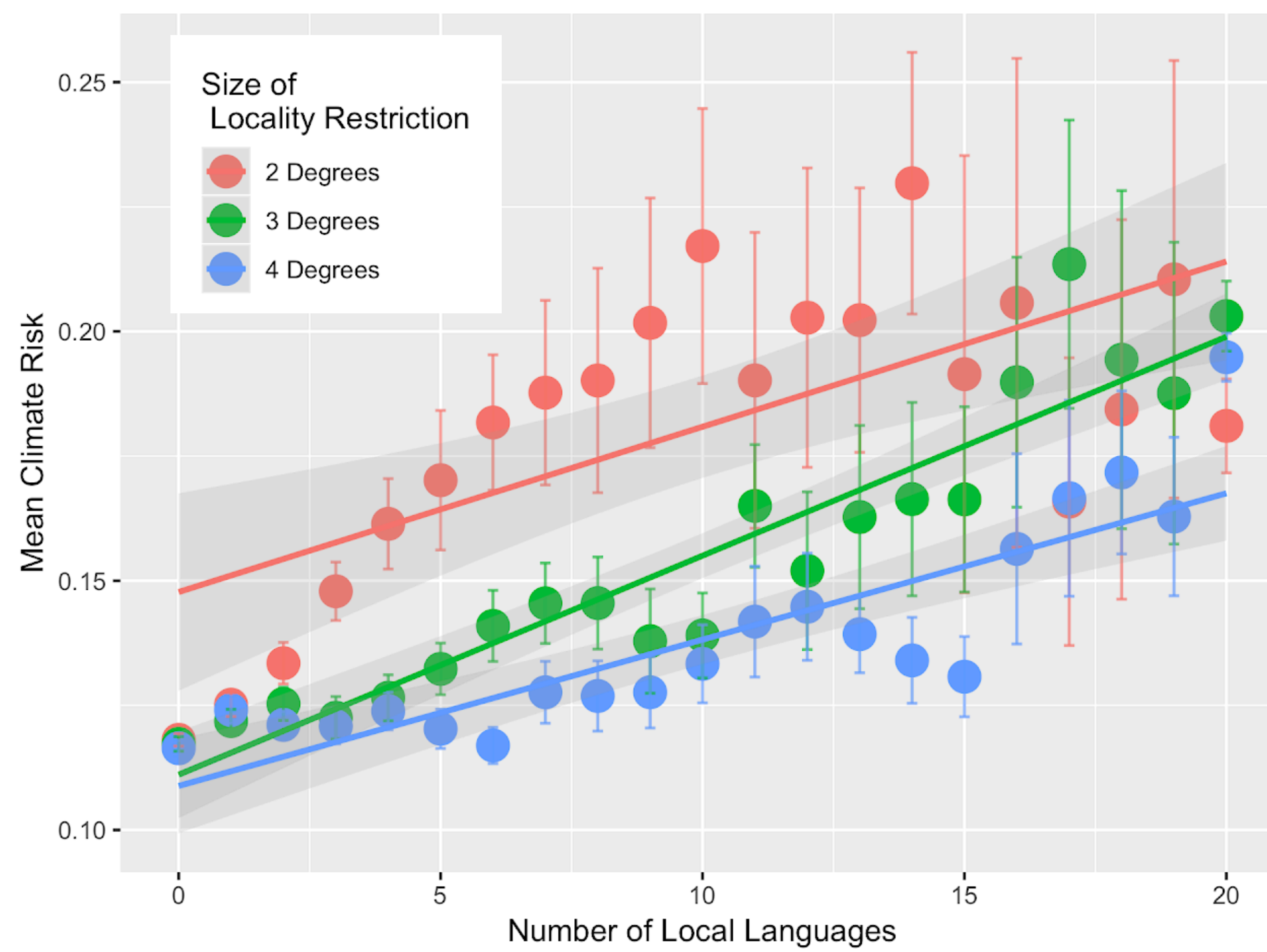
Correlation Between Linguistic Diversity & Climate Risk

- **Motivating Intuition:** Correlation between biological diversity, climate change and linguistic diversity.



- **Linguistic Diversity vs. Climate Risk**

1. Sample 10,000 Land Points
2. Count # of Langs within *n* degrees lat & long
3. Plot against mean climate risk at point *across all pathways*.
4. For 3 parameters of *n*



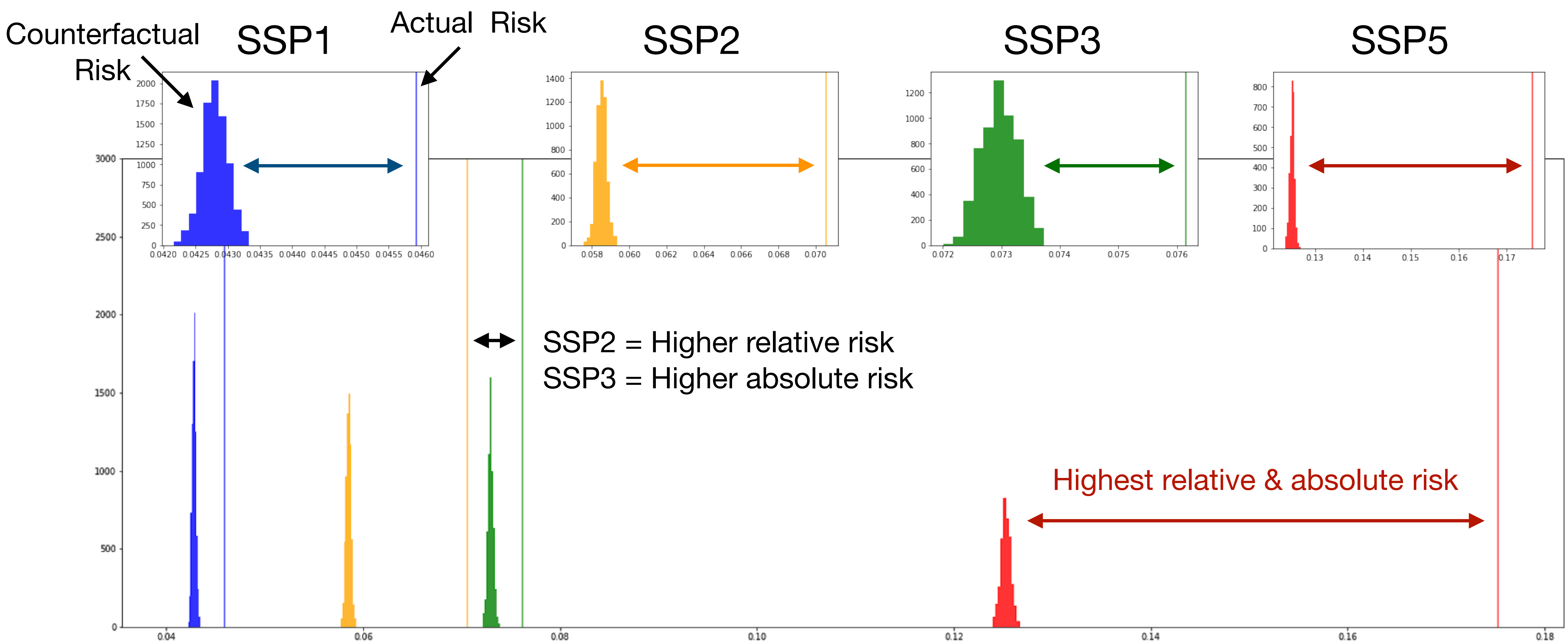
<u>Results:</u>

- Positive correlation ($p<0.001$) for all tested values of n
- Smaller locality restriction = higher average climate risk, indicating more risk in super diverse regions.

Language Distribution & Climate Risk

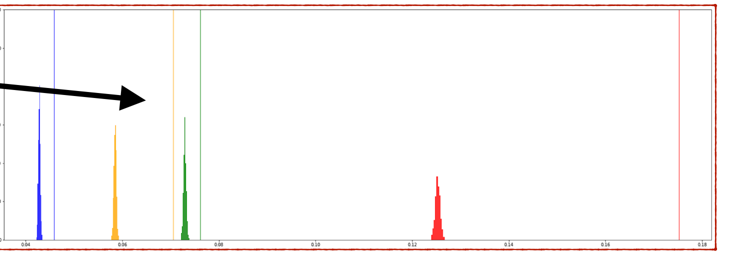
- Does the geographic distribution of languages expose them to outsize risk?
- Method:
  - Sample ~6,400 land points at random
  - Calculate mean climate risk across all 6,400 points
  - Do this 500 times to get histogram of risk across all languages (pyramid)
  - Compare to risk across actual languages (line)

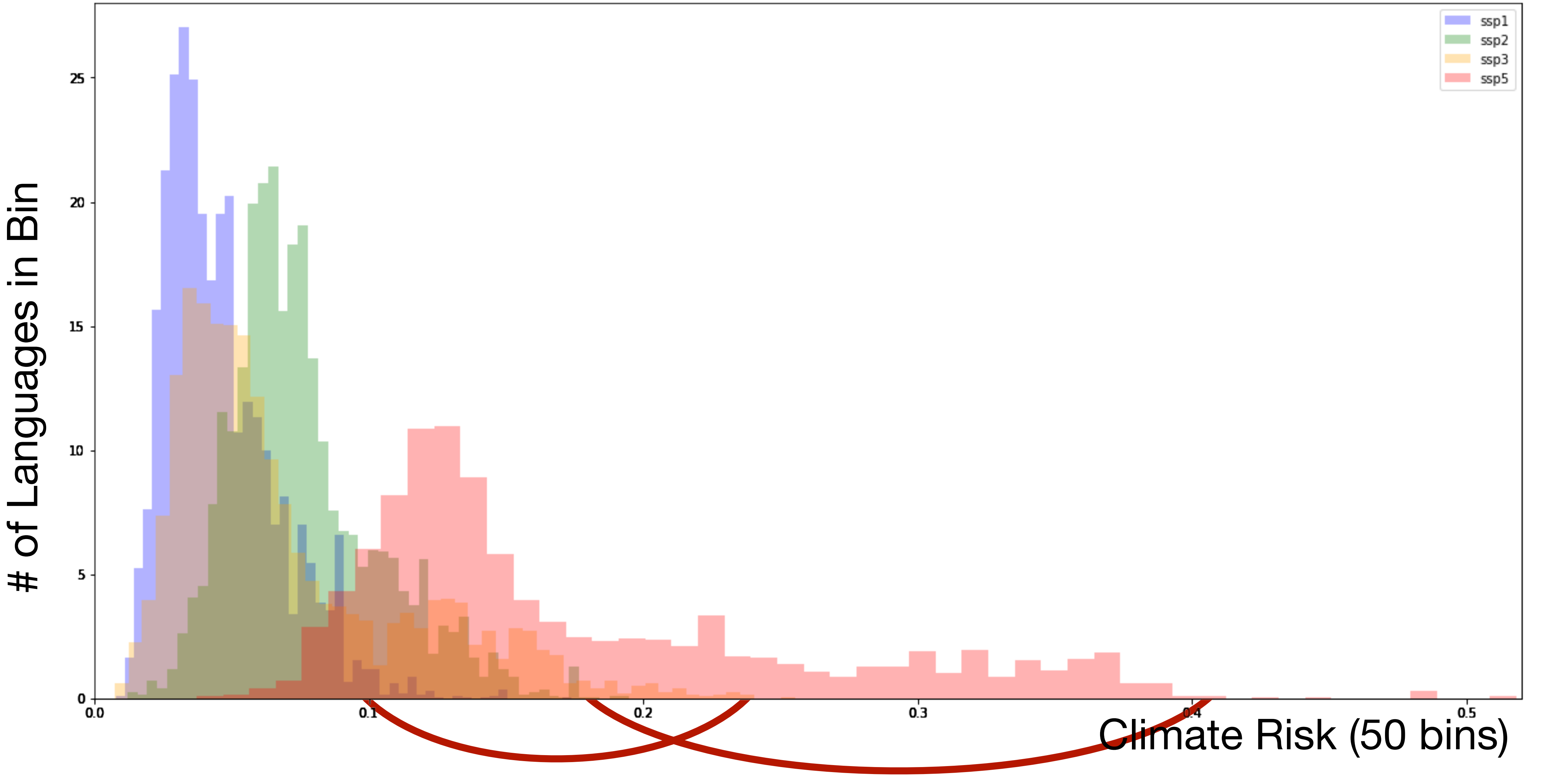
Counterfactual language distributions



Risk Distribution Across Pathways

- Shows mean risk of language points
- What is the spread of risk across pathways?





More change = more risk
The long tail of risk grows as pathways become more serious

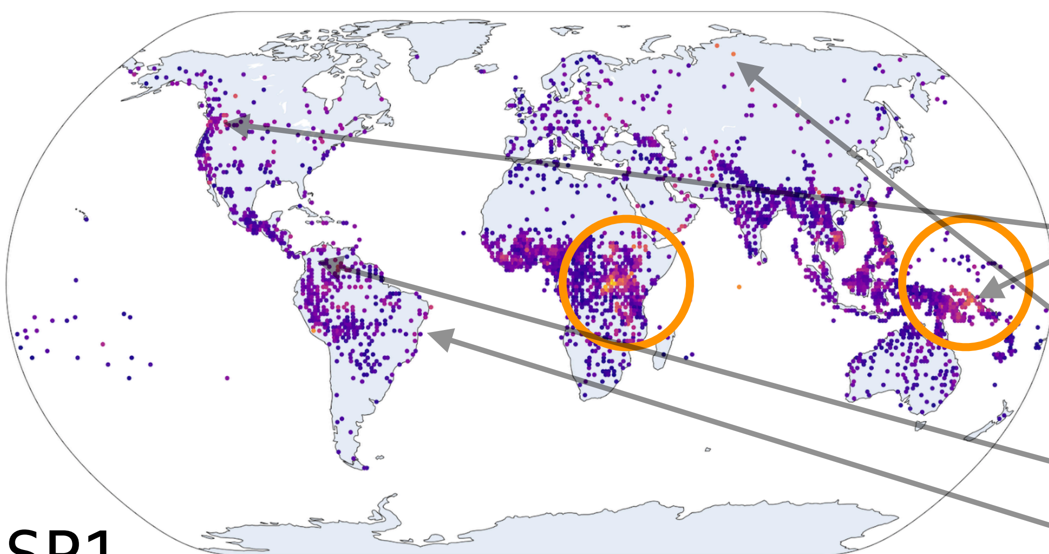
Outline: Desiderata / Questions

- Finding Climate Hotspots
- Will climate change have an outsized impact on linguistic diversity?
- What languages / language families will be impacted the most?
  - Do different climate pathways affect the same or different languages?
- Are endangered languages at greater risk from climate change?

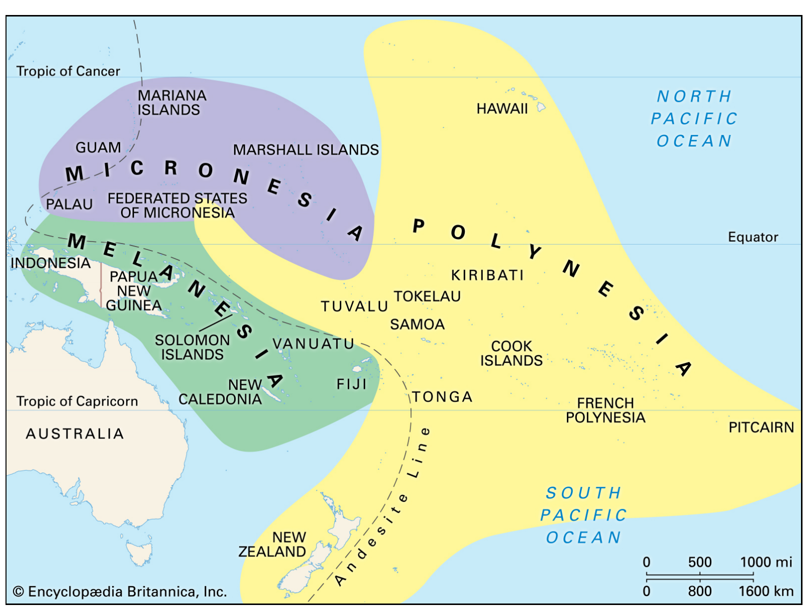
- Yes, will have outsize impact on languages due to the way they're distributed on the planet.
- More change = more impact

By-Language Impact in SSP1

- Plot each language w/ its climate risk (colors not normed across SSPs)
- Language family & region information from Enthnologue



SSP1



Top Impacted Families

East New Britain
Chimakuan
Yele-West New Britain
North Bougainville
Yeniseian
Sálivan
Kamakanan
Muran
Yokutsan
South Bougainville

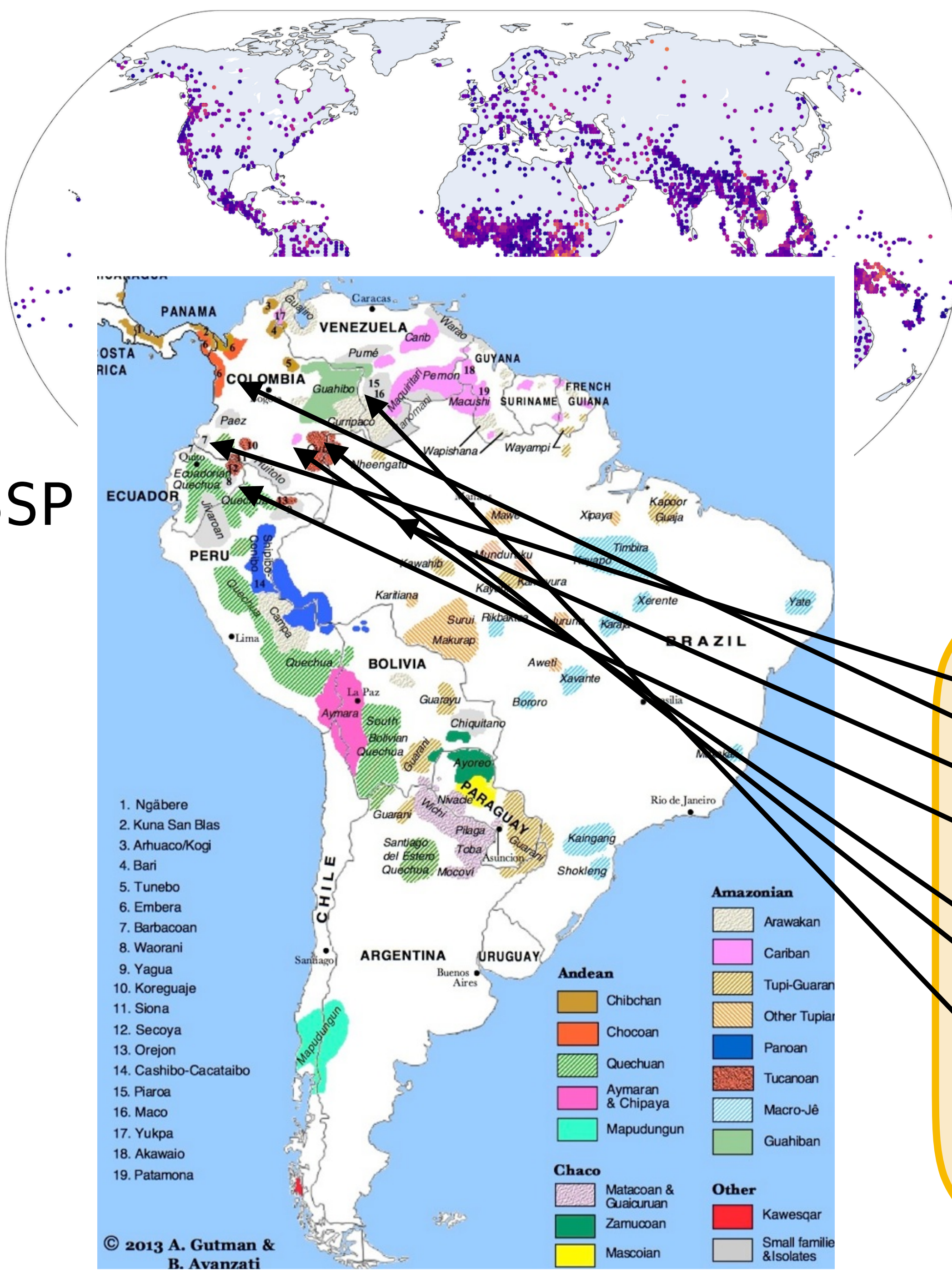
Top Impacted Regions

Melanesia
Eastern Africa
Central Asia
Northern Africa
Western Africa

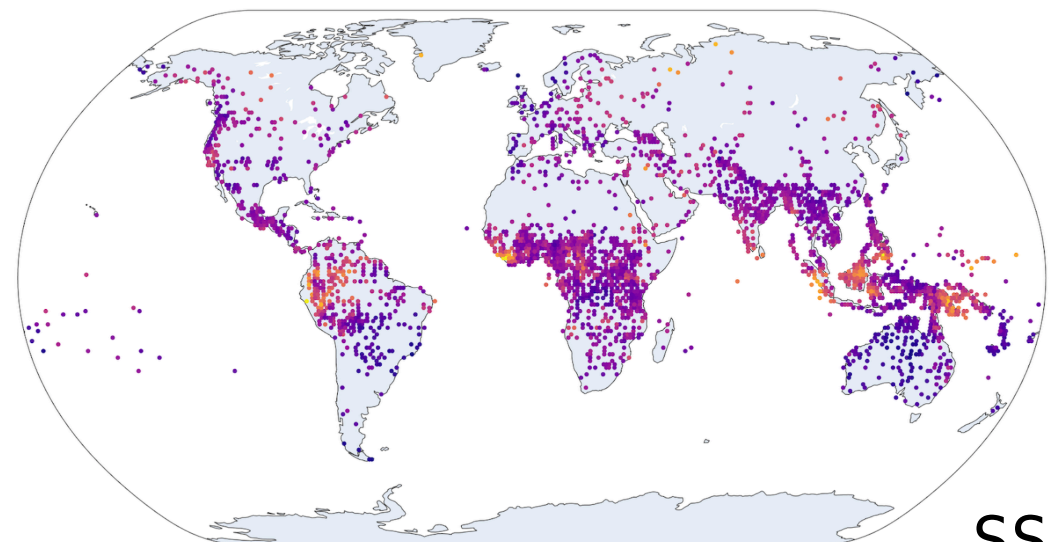
- Top impacts in Melanesia and central / Eastern Africa
- Impacted language families come from a diverse set of locations
- Most impacted families come from New Guinea ( )

By-Language Impact in SSP2

- Risk is spread out more evenly across the globe
- More impact in Western Africa, South America and South-East Asia / Micronesia



SSP



SSP2

Top Impacted Families

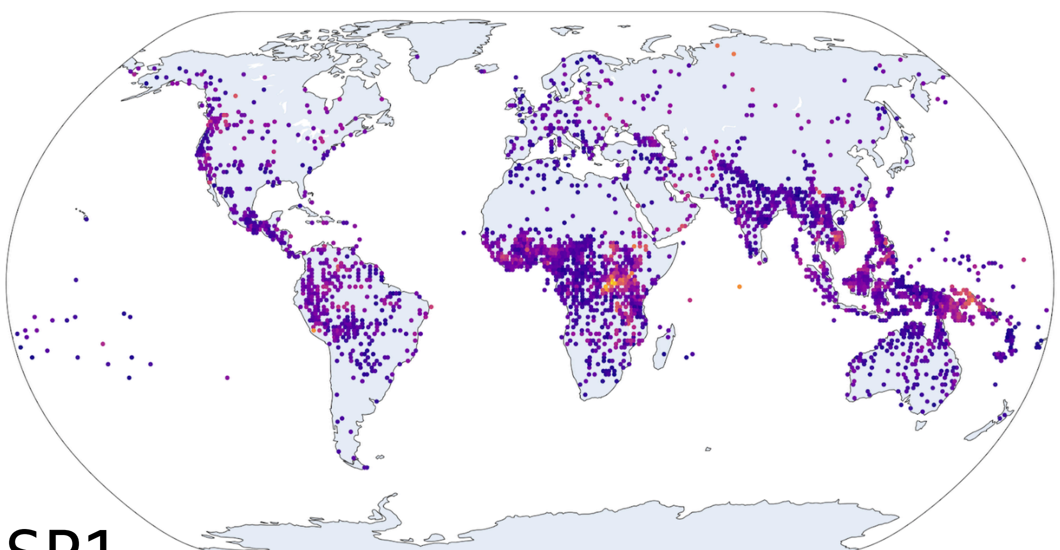
Barbacoan
Paezan
Yaguan
Jivaroan
Mongol-Langam
Cholonan
Yanomaman
Piawi
Puinavean
Yele-West New Britain

Top Impacted Regions

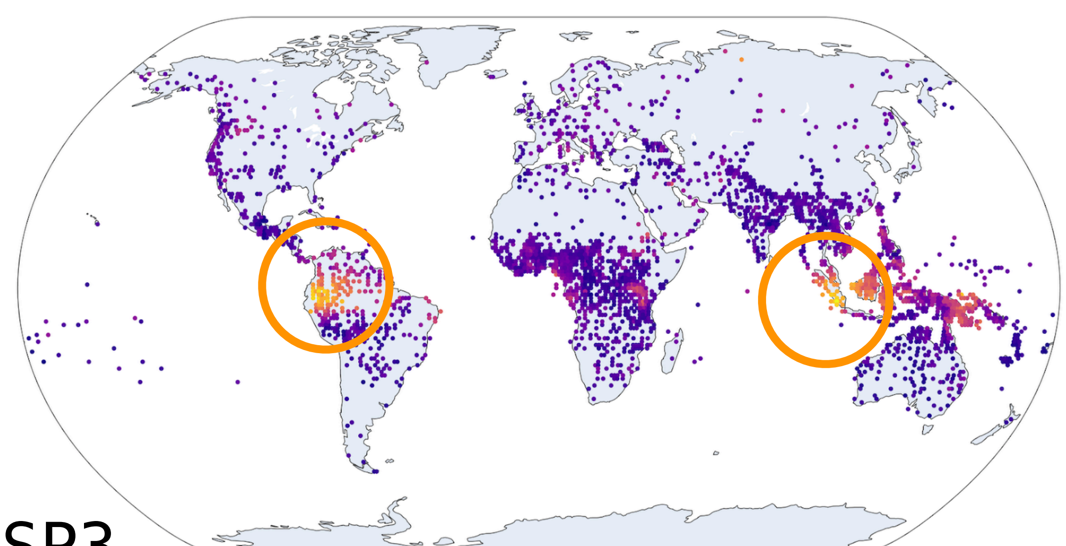
Melanesia
Micronesia
South America
Eastern Europe
Western Asia

By-Language Impact in SSP3

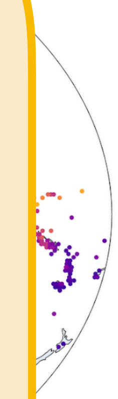
- Risk is concentrated in SE Asia / Melanesia + South America
- Greater impacts in Sumatra, Java and Borneo than previous pathways



SSP1



SSP3



Top Impacted Regions

Melanesia
South America
South-Eastern Asia
Caribbean
Southern Europe

Top Impacted Families

Zaparoan
Barbacoan
Witotoan
Jivaroan
Yaguan
Tucanoan
Katukinan
Cahuapanan
Paezan
Puinavean

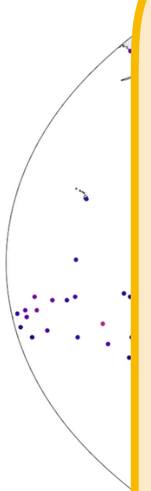
All of these are in Columbia, Venezuela, Equador and Peru

By-Language Impact in SSP5

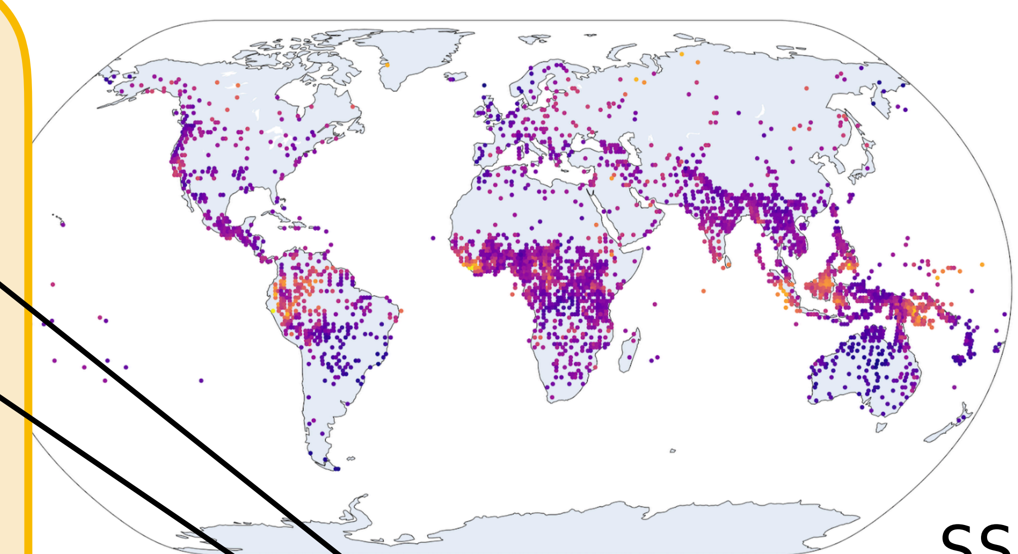
- Similar impacts to SSP3, with less impact in Sumatra, Java and Borneo
- Impacted Families: Either on New Guinea or in Northern Amazonia

Top Impacted Regions
Melanesia
South America
S.E. Asia
Micronesia
Eastern Africa

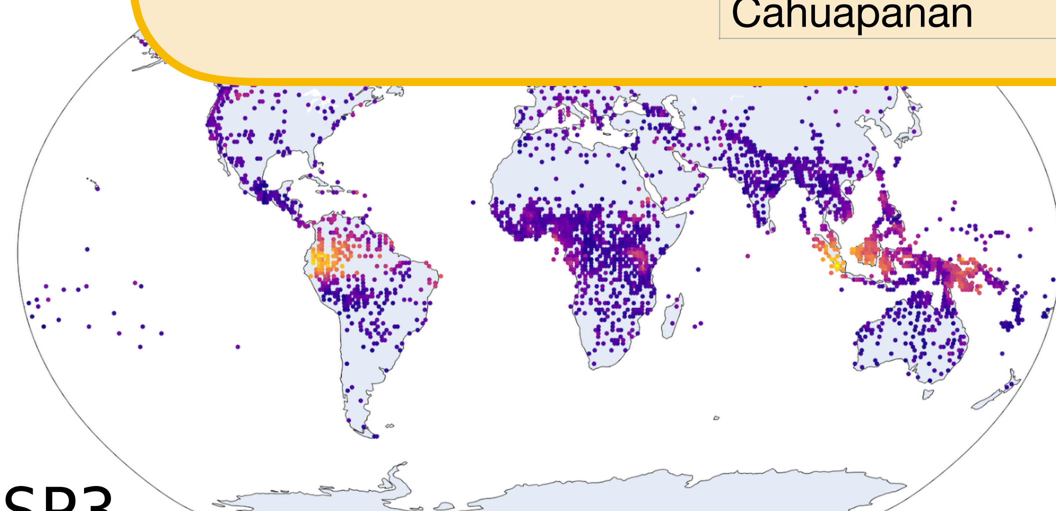
Top Impacted Families
Mongol-Langam
Barbacoan
Yuat
Witotoan
Yaguan
Zaparoan
Piawi
Ramu-Lower Sepik
Arafundi
Cahuapanan



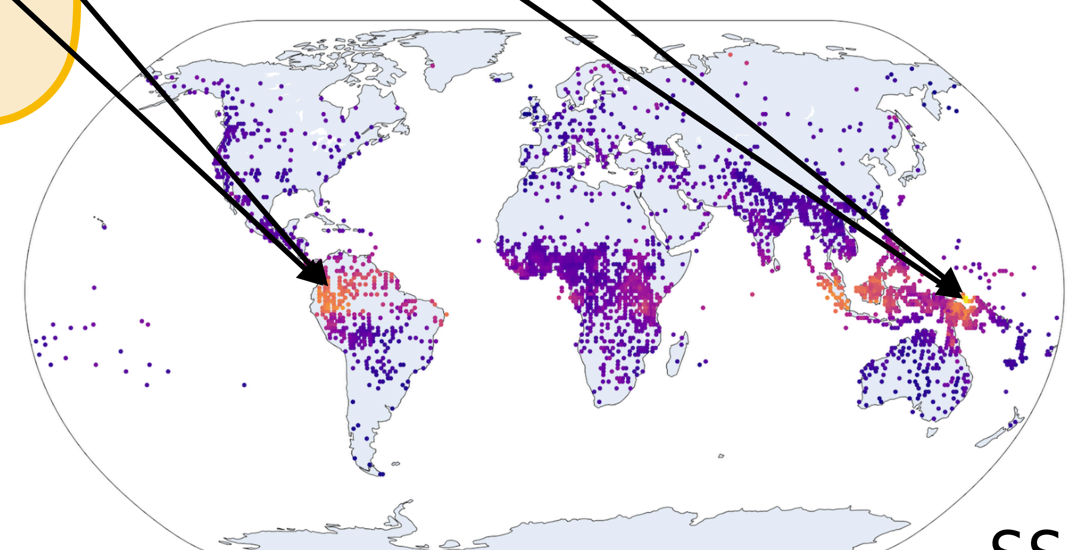
SSP1



SSP2



SSP3



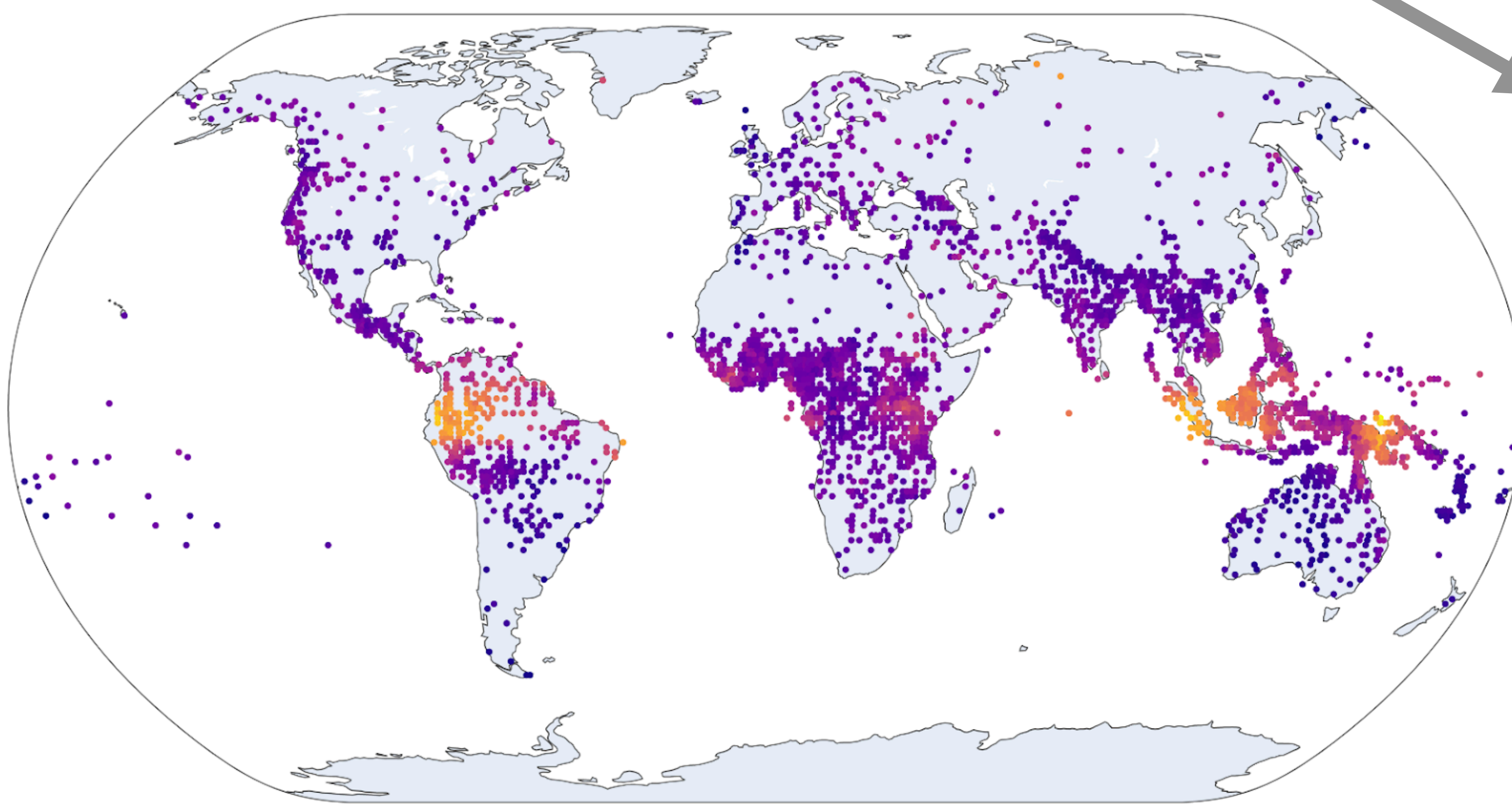
SSP5

By-Language Impact across all SSPs

- Risk averaged across all pathways
- Top Impacted Families are either in South American or Melanesia

Top Impacted Regions

Melanesia
South America
S.E. Asia
Micronesia
Eastern Africa
Caribbean
Western Africa
Northern Africa



Top Impacted Families

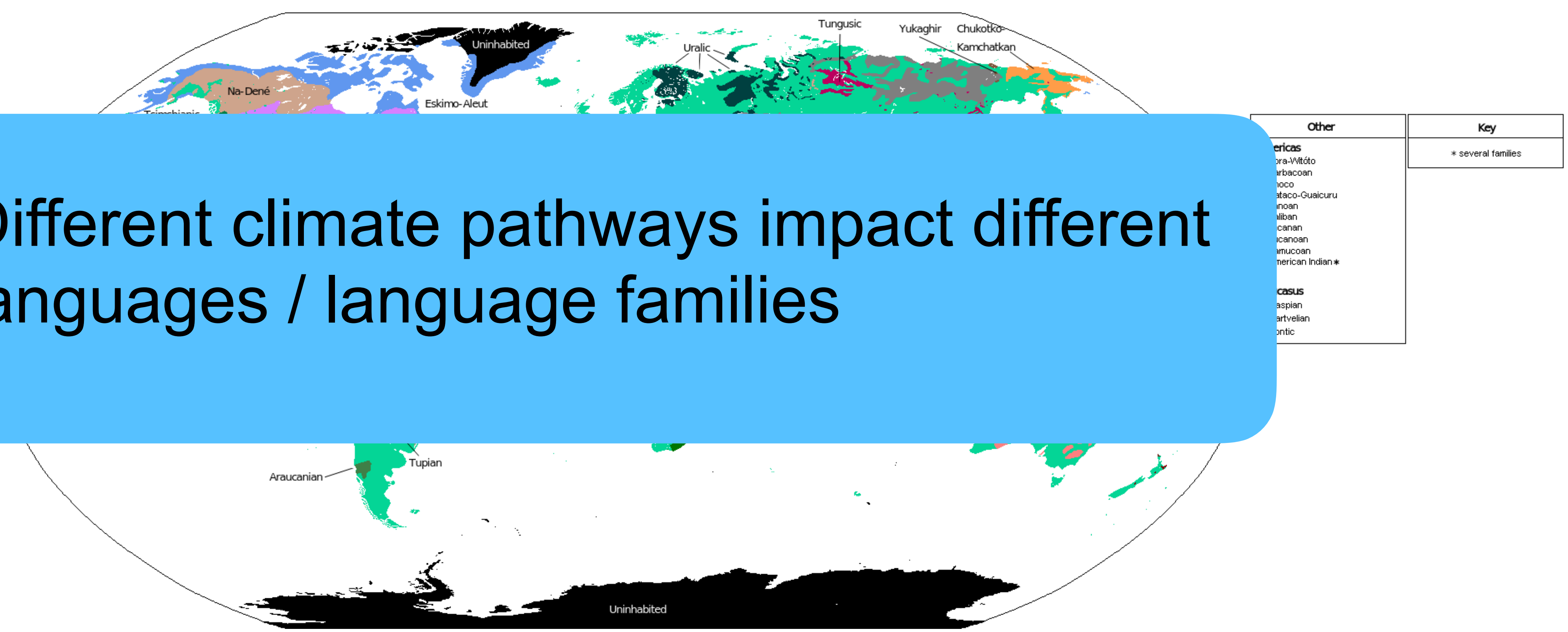
Barbacoan
Witotoan
Jivaroan
Yaguan
Zaparoan
Tucanoan
Mongol-Langam
Paezan
Cahuapanan
Puinavean
Piawi
Katukinan
Ramu-Lower Sepik
Yuat
Arafundi
Sálivan
Yanomaman
Trans-New Guinea
Torricelli
Sepik

Key Takeaways:

- Western South America & Melanesia remain high risk **throughout all pathways**
- Risk to other regions: Africa, S.E. Asia & Central Asia **varies with pathway**

Outline: Desiderata / Questions

- Finding Climate Hotspots
- Will climate change have an outsized impact on linguistic diversity?
- What languages / language families will be impacted the most?
  - Do different climate pathways affect the same or different languages?
- Are endangered languages at greater risk from climate change?

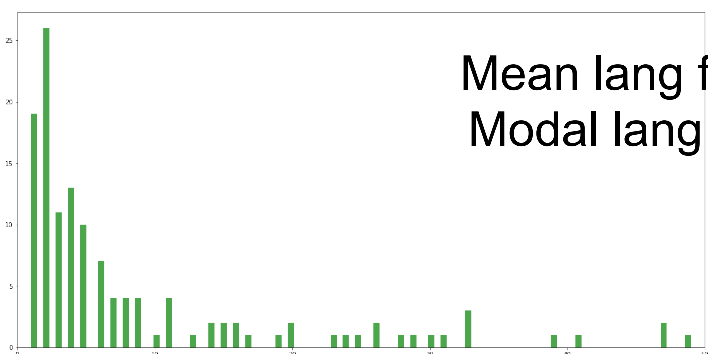


- Different climate pathways impact different languages / language families

Climate Risk to Language Family by Size

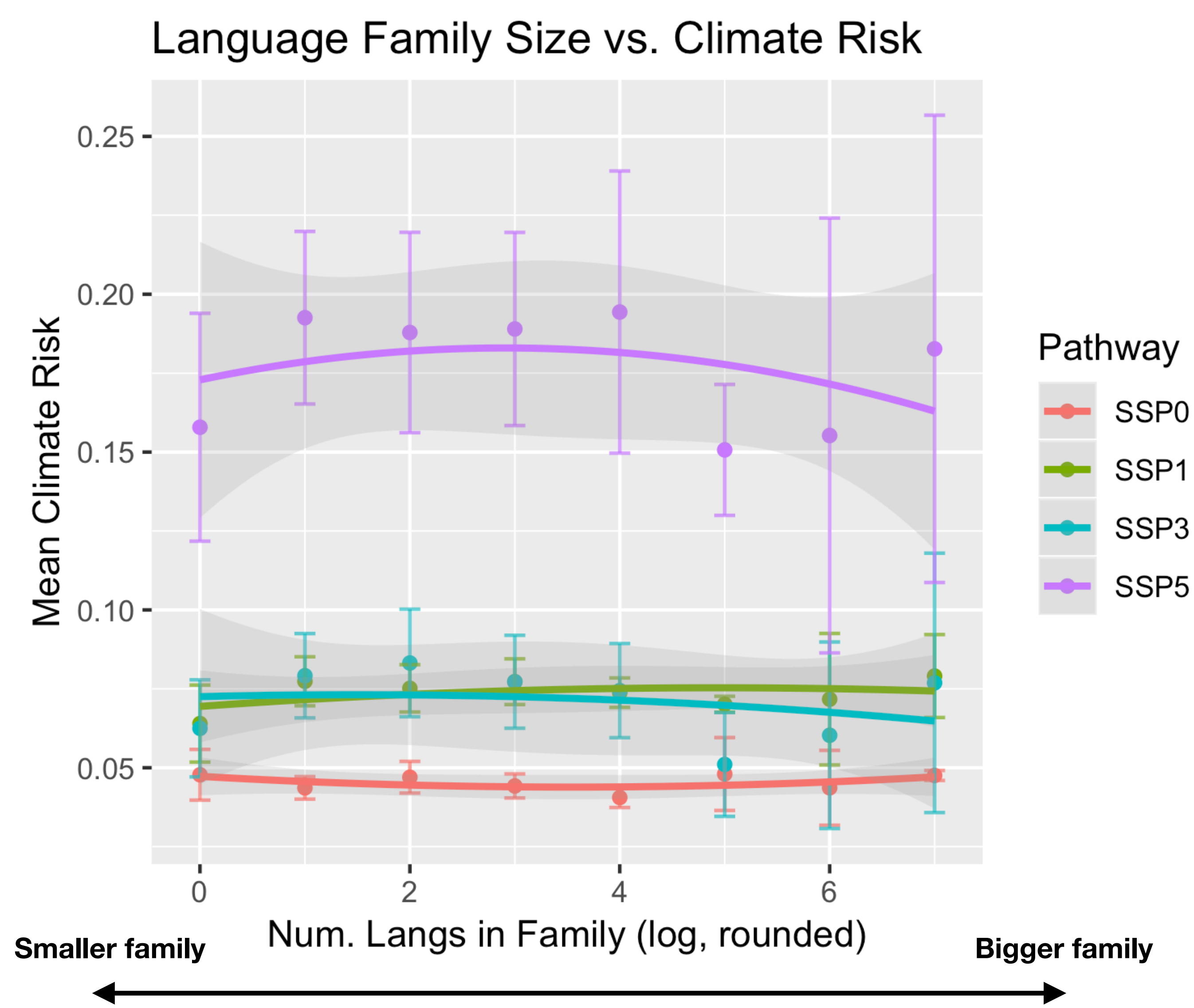
- Each language family represents a unique vantage point into the human language faculty.
- Families with fewer languages are at a higher risk of going extinct
- Does climate change differentially impact smaller language families?

Histogram of language family size



Mean lang family size: 42
Modal lang family size: 2

- Group by number of languages in family (log, rounded); take mean climate risk for each group
- Top-level language families in Ethnologue — 148 families
- **No correlation** between number of languages in family an mean climate risk (p>0.05)

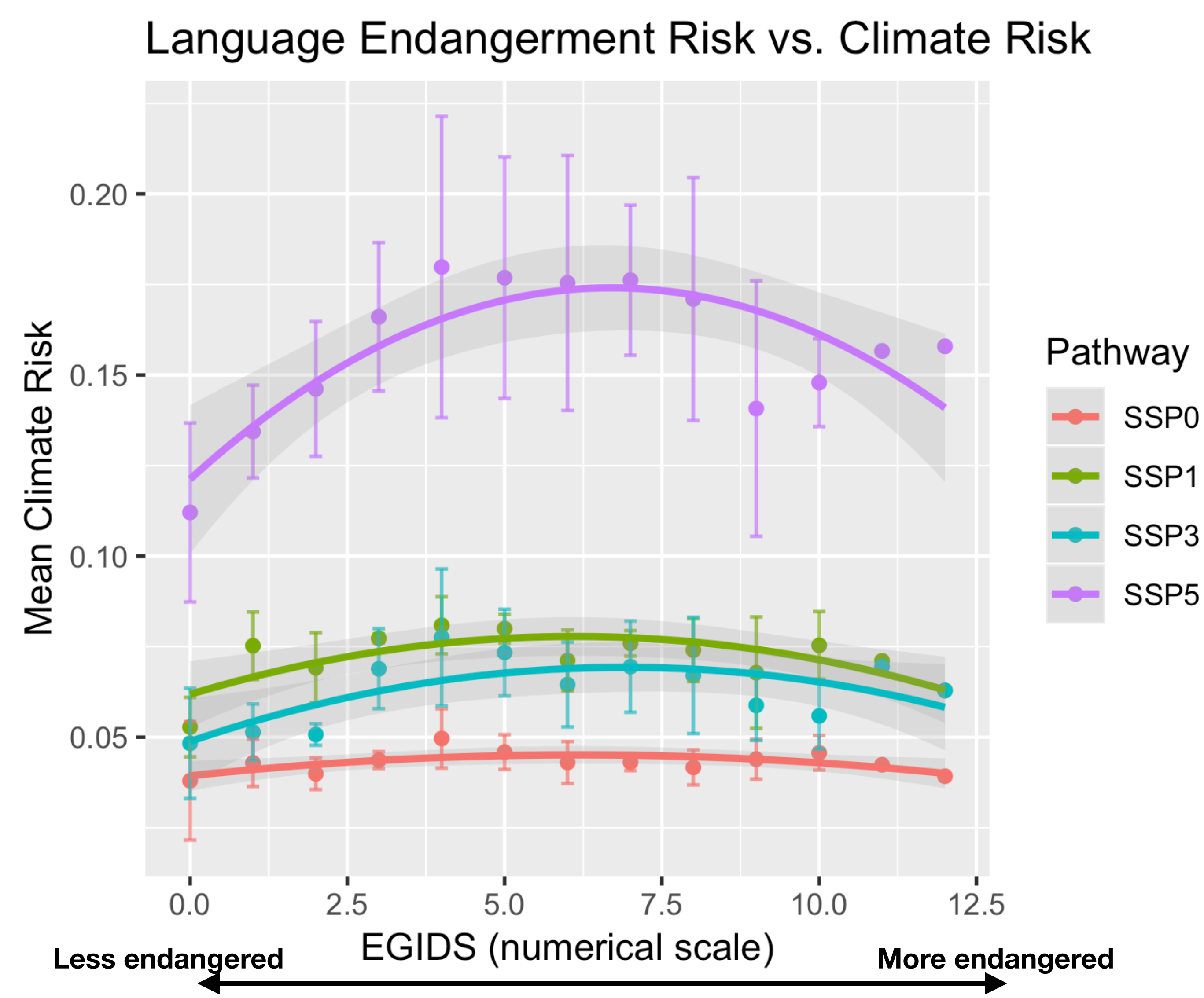


Climate Impact on Endangered Languages

- Does climate risk differentially impact endangered languages?
- Use numeric version of EGIDS endangerment scale from Ethnologue

	Level	Label	Description
0	0	International	The language is widely used between nations in trade, knowledge exchange, and international policy.
1	1	National	The language is used in education, work, mass media, and government at the national level.
2	2	Provincial	The language is used in education, work, mass media, and government within major administrative subdivisions of a nation.
3	3	Wider Communication	The language is used in work and mass media without official status to transcend language differences across a region.
4	4	Educational	The language is in vigorous use, with standardization and literature being sustained through a widespread system of institutionally supported education.
5	5	Developing	The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
6	6a	Vigorous	The language is used for face-to-face communication by all generations and the situation is sustainable.
7	6b	Threatened	The language is used for face-to-face communication within all generations, but it is losing users.
8	7	Shifting	The child-bearing generation can use the language among themselves, but it is not being transmitted to children.
9	8a	Moribund	The only remaining active users of the language are members of the grandparent generation and older.
10	8b	Nearly Extinct	The only remaining users of the language are members of the grandparent generation or older who have little opportunity to use the language.
11	9	Dormant	The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.
12	10	Extinct	The language is no longer used and no one retains a sense of ethnic identity associated with the language.

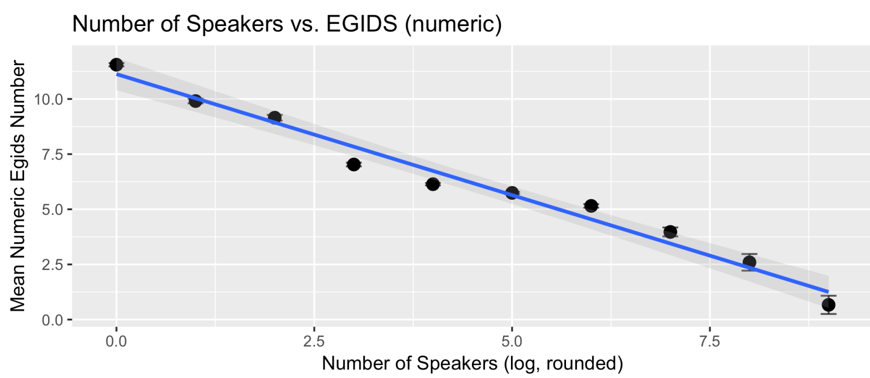
From ethnologue.com



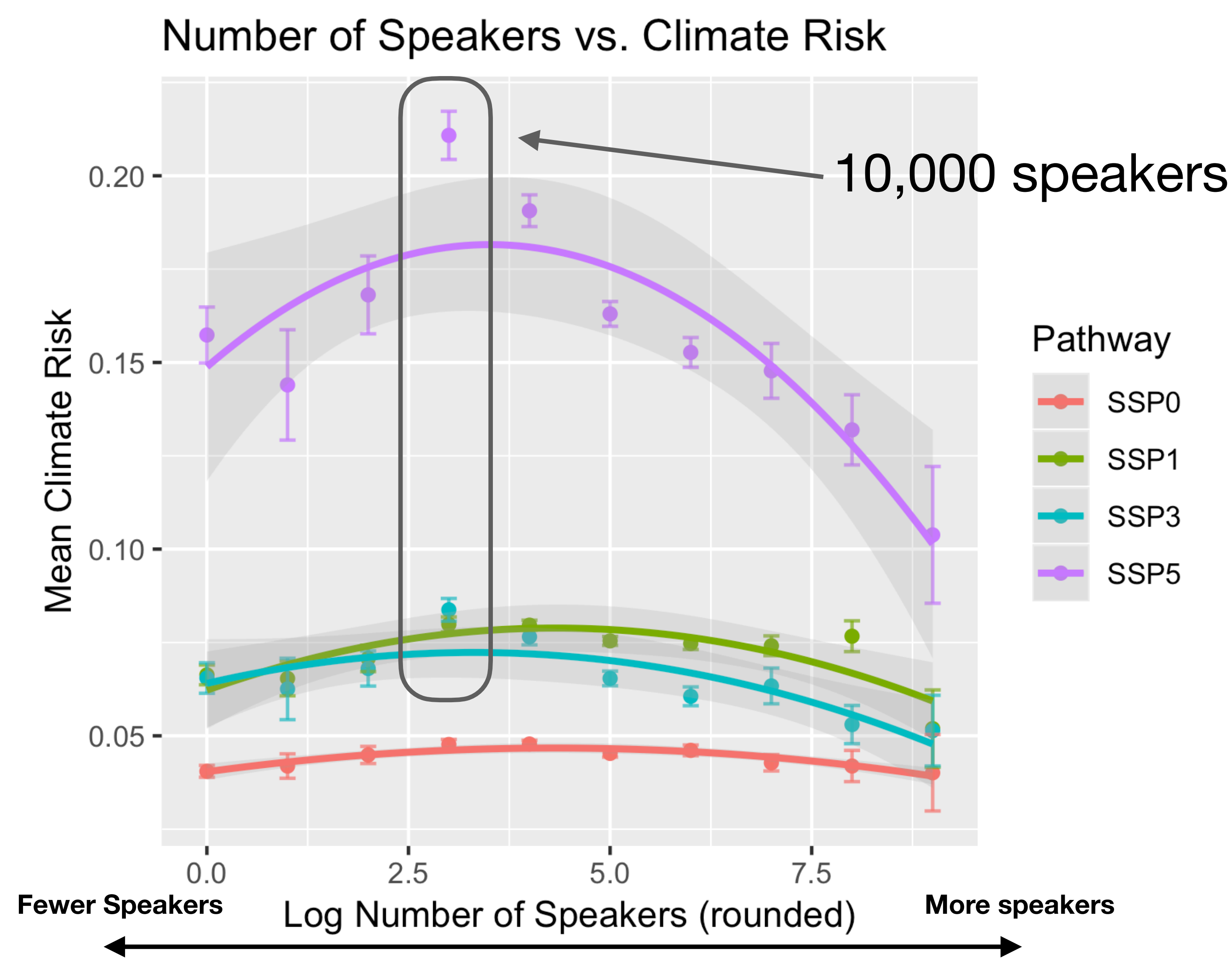
- Vigorous, threatened and shifting languages are at greatest risk
- Good on one hand, but also may target critical point on endangerment pathway

Climate impact on languages by Speaker Count

- EGIDS may be an imperfect way to measure endangerment
- Use # of Speakers as a different metric (the two are correlated)

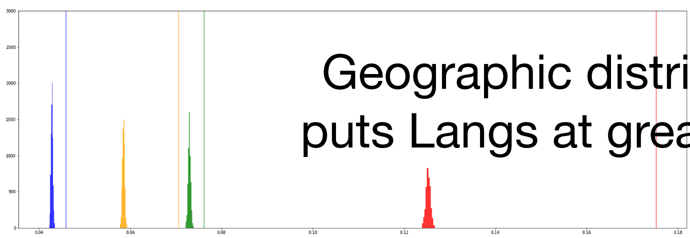


- Group by log normalized number of speakers and find mean climate risk per group
- Risk highest for languages with between 1,000 - 100,000 speakers
- Again, hard to interpret. May represent critical inflection point on endangerment pathway.



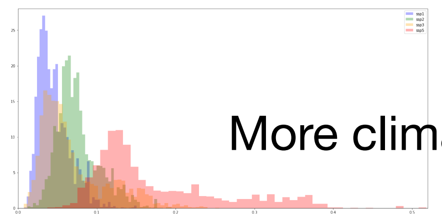
Conclusion

- <u>Will climate change have an outsized impact on linguistic diversity?</u>



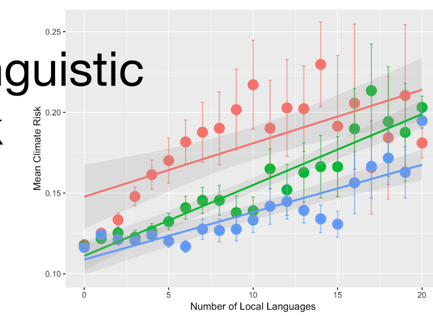
Geographic distribution puts Langs at greater risk

Yes!



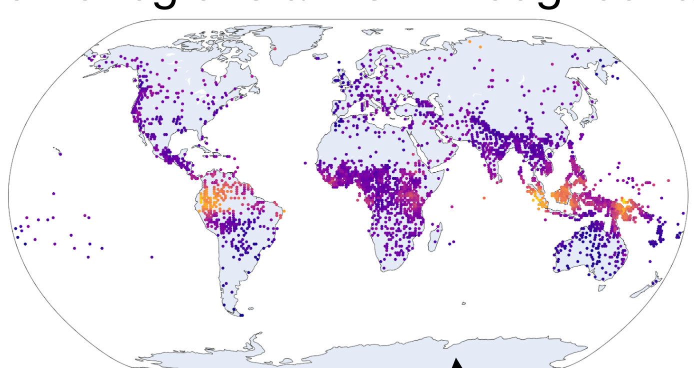
More climate change = more risk

Correlation between regional linguistic Diversity and climate risk



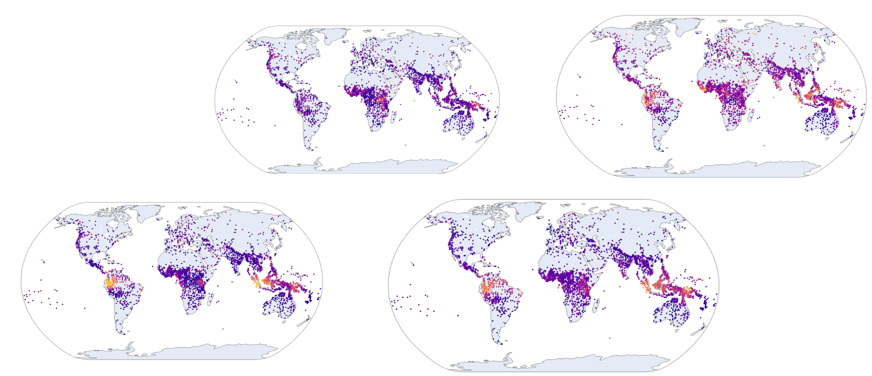
- <u>What languages / language families will be impacted the most?</u>

Some regions at risk throughout all pathways

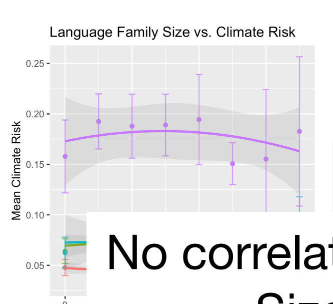


Melanesia
South America
S.E. Asia
Micronesia
Eastern Africa
Caribbean
Western Africa
Northern Africa

For some regions, risk varies with pathway

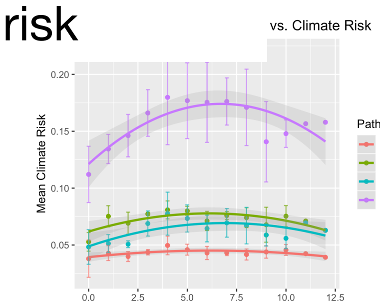


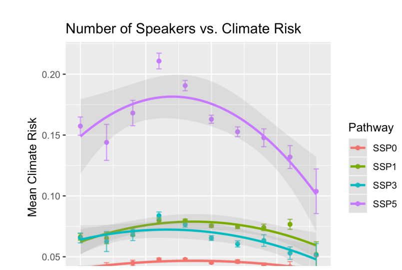
- <u>Are threatened languages at greater risk from climate change?</u>



No correlation between lang family Size and climate risk

Languages at **medium risk** (EGIDS) are At the most climate risk



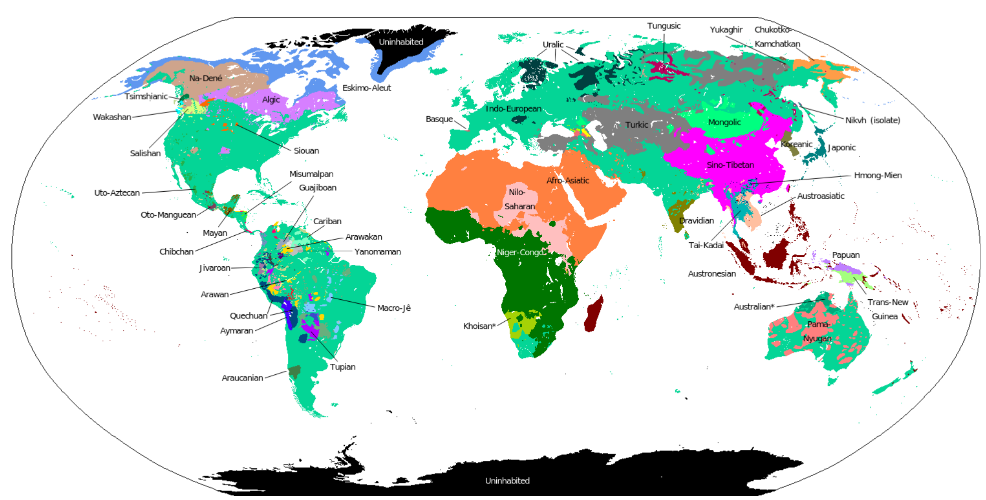


Languages with **medium # of speakers** Are at the most climate risk

Future Directions and Discussion

Limitations of Current Study

- Doesn’t take into account **sea level rise**
- Doesn’t take into account **population dynamics**
- Is a **relative model**, not an absolute model
    - Doesn’t make predictions about what will happen to the languages.
    - Need good historical data for this -- it’s hard to come by!



Thanks for listening!

Discussion

- How to connect the linguistics & climate modeling community!
    - Applied for funding for a workshop in February
- What goals should linguists & climate scientists work towards?