

Pre-Trained Language Models Represent Some Geographic Populations Better Than Others

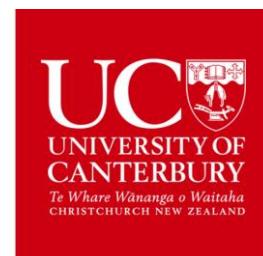
Jonathan Dunn

jedunn@illinois.edu

www.jdunn.name

Ben Adams

benjamin.adams@canterbury.ac.nz



Harish Tayyar Madabushi

htm43@bath.ac.uk

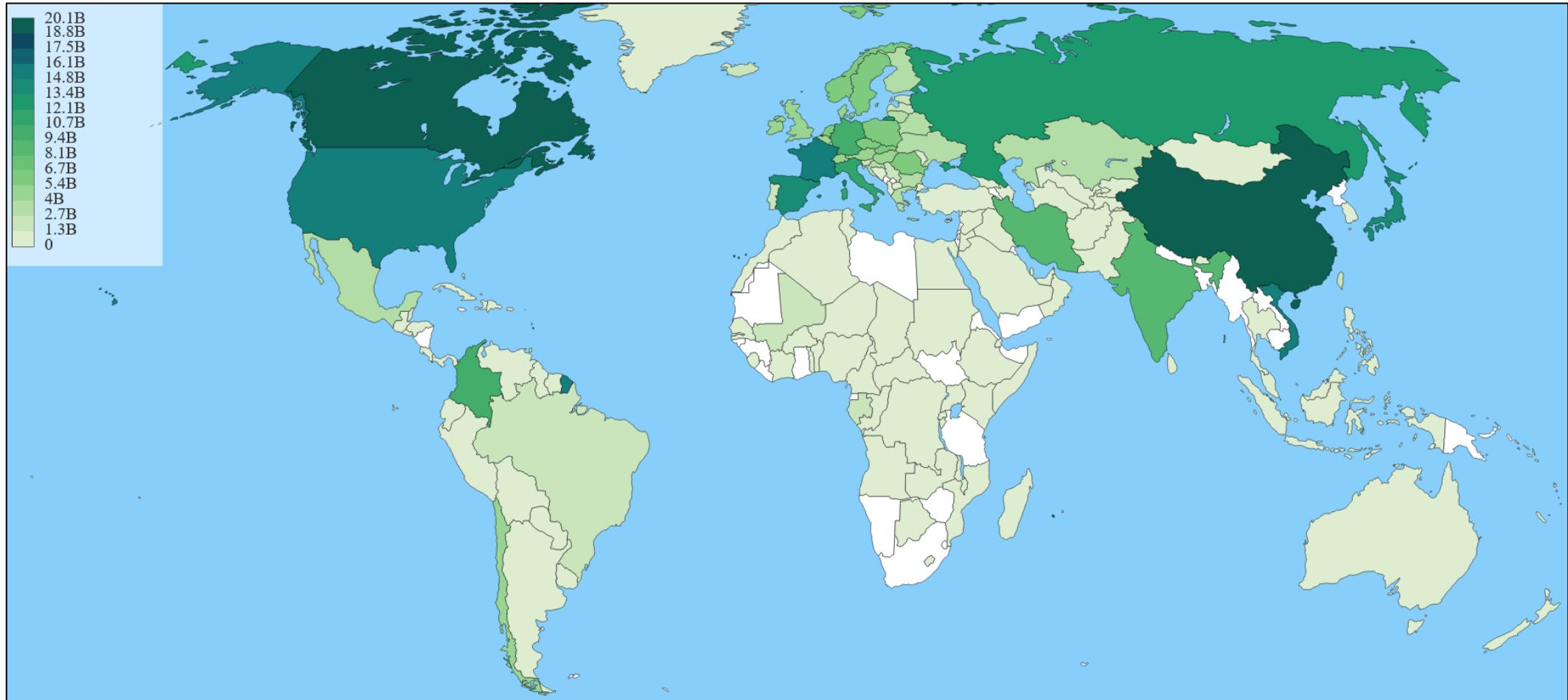
www.harishtayyarmadabushi.com



How does linguistic variation impact downstream models?

We know that training corpora have a geographic skew

www.earthLings.io



Population Skew/Bias

Who do large language models represent?

BigScience BLOOM (560 mil, 1.7 bil, 3 bil)

Facebook OPT (350 mil, 1.3 bil, 2.7 bil)

Population Skew/Bias

Populations represented

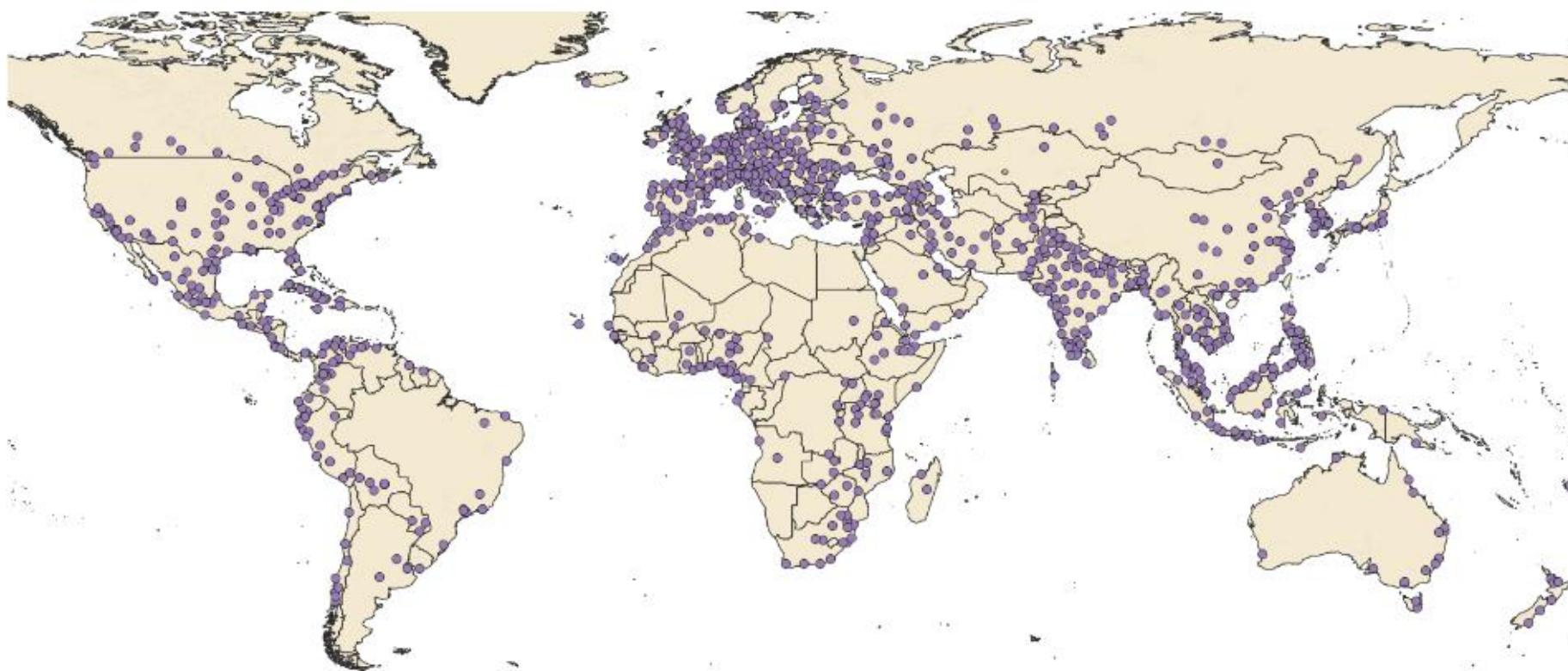


Figure 1: Map of Local Populations. Each point represents a data collection location.

Population Skew/Bias

Country Level

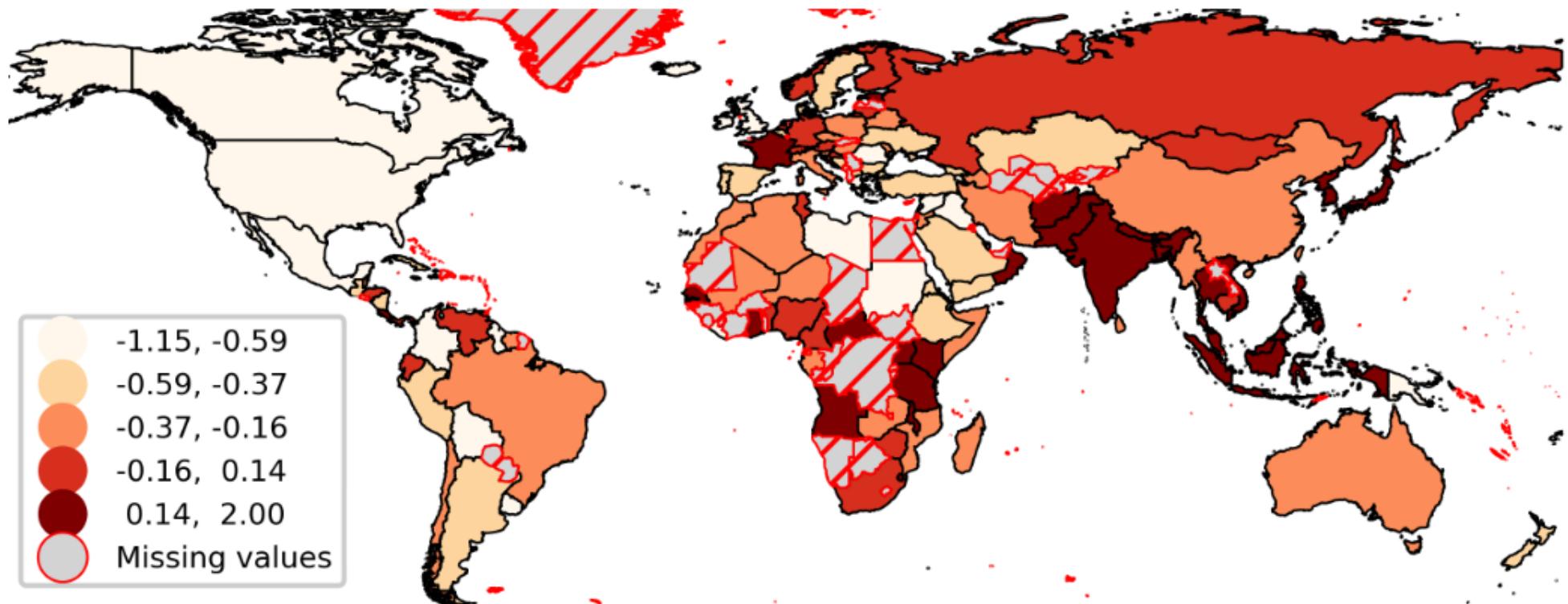
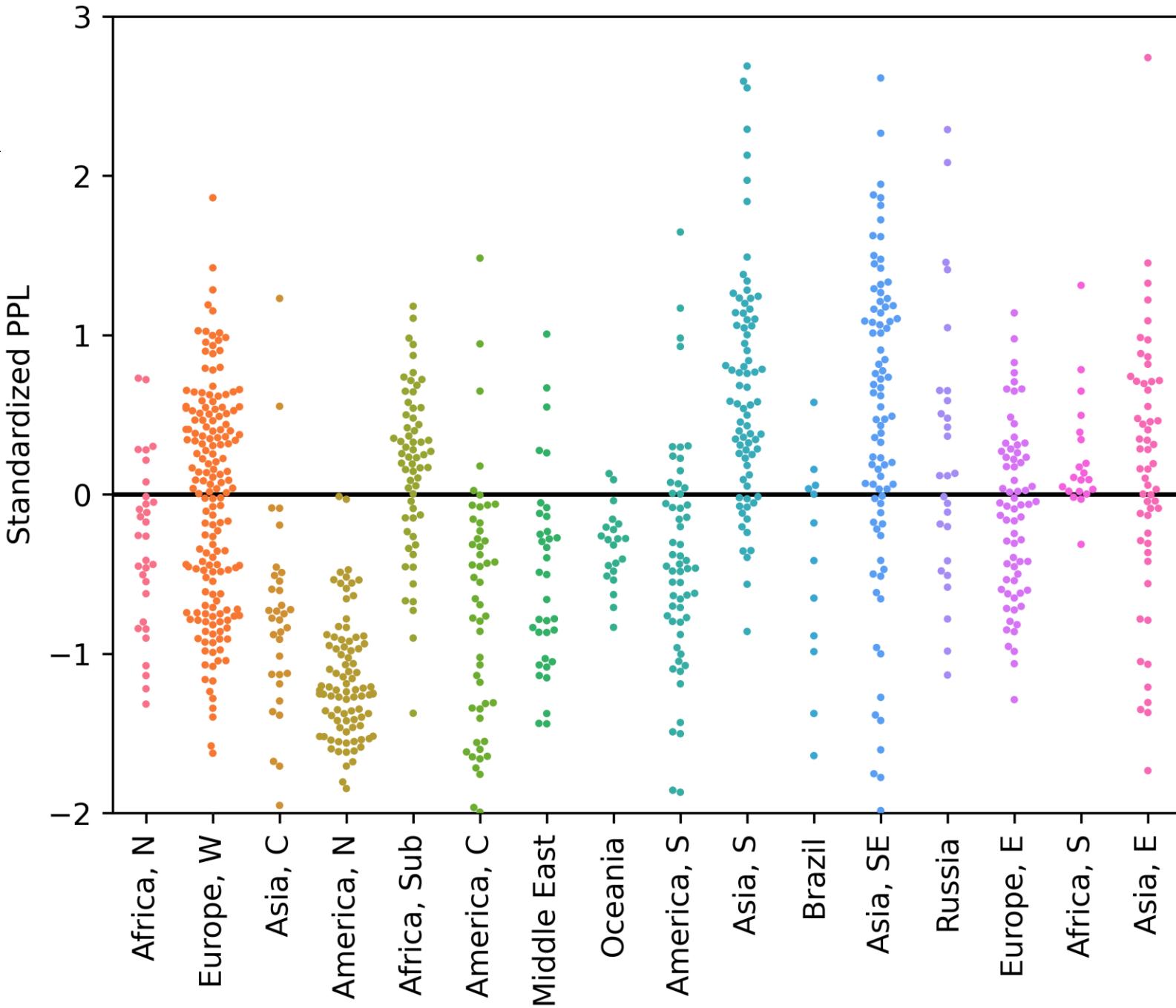


Figure 2: Countries by standardized perplexity scores across sub-corpora for BigScience BLOOM 3b. Because the z-score is used, 1 reflects one standard deviation above the mean and -1 below the mean.

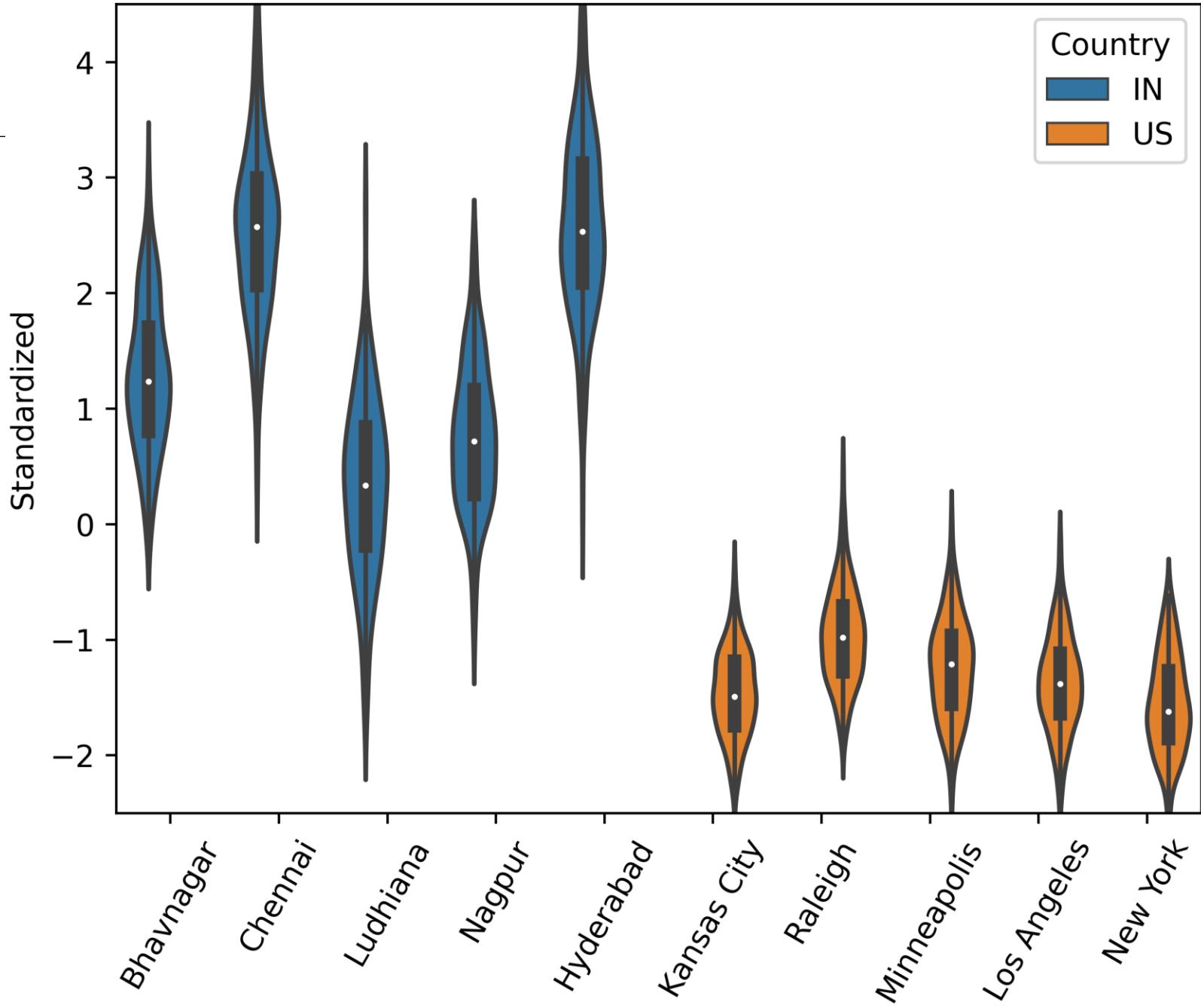
Population Skew/Bias

Local Areas by Region
(OPT 2.7b)



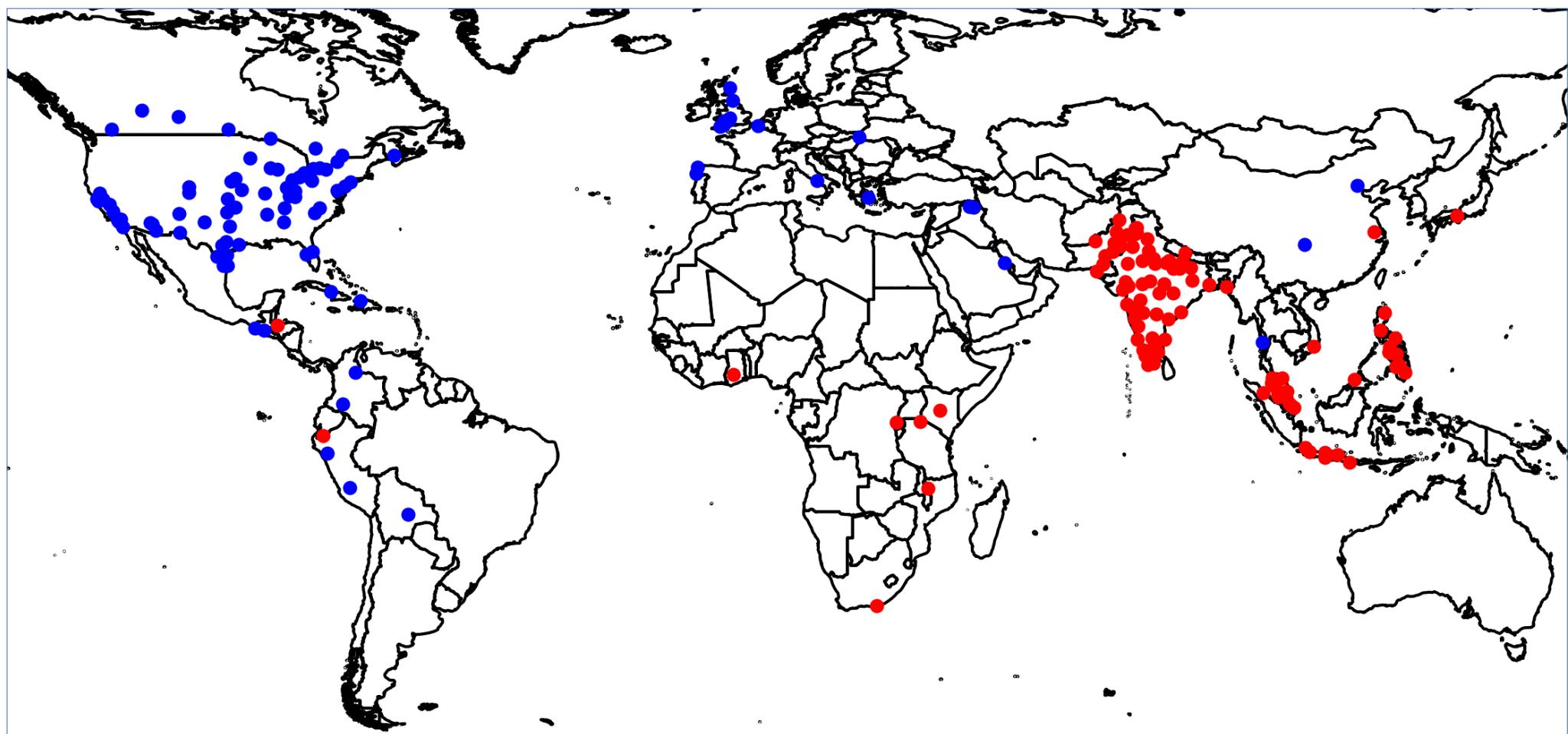
Population Skew/Bias

Distribution for Select Cities
(OPT 2.7b)



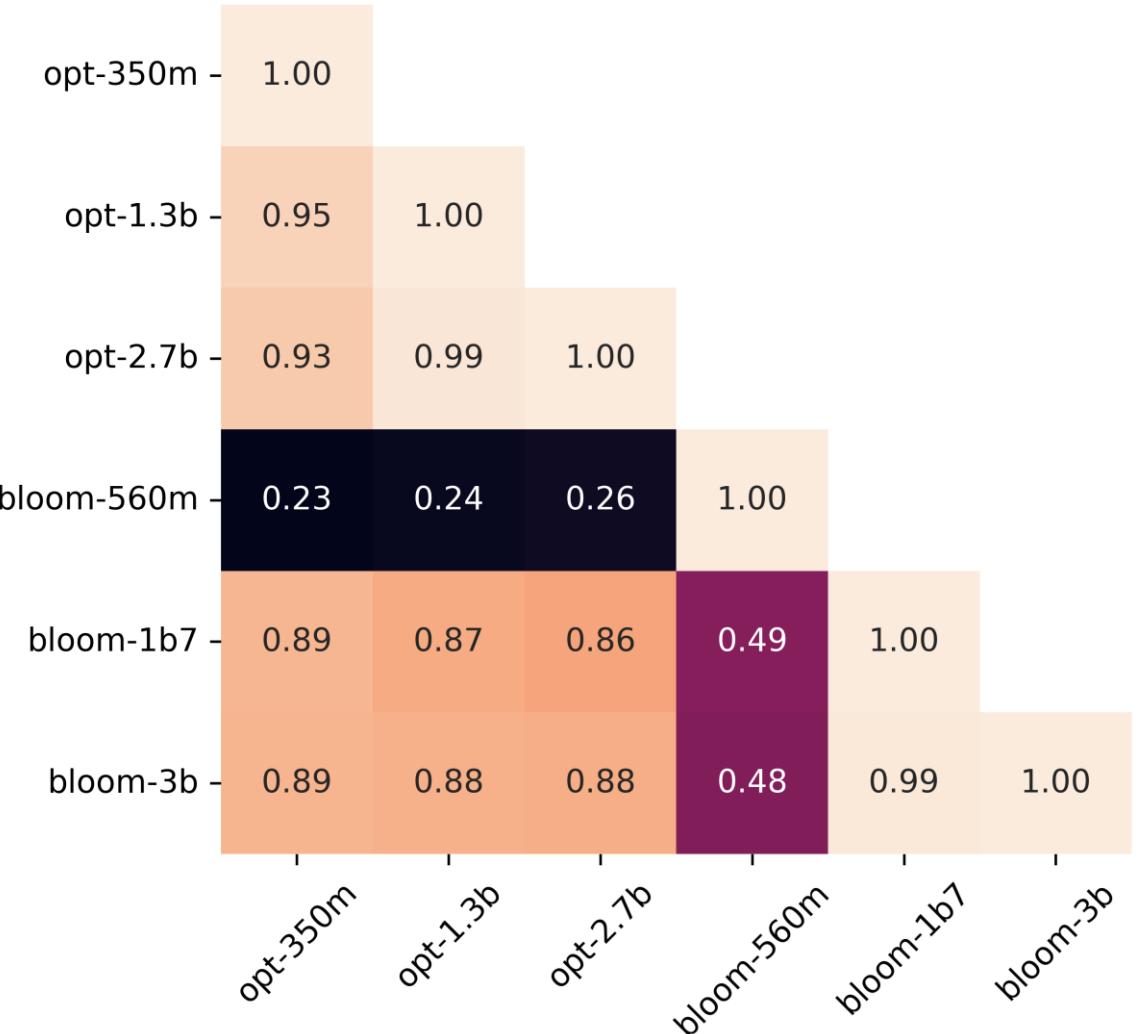
Population Skew/Bias

Best Cities (blue) and Worst Cities (red)
(BLOOM 3b)



Population Skew/Bias

Relatively high agreement across both model families and parameter sizes



Population Skew/Bias

Consistent across temporal partitions

OPT 2.7b		
<i>Country</i>	<i>City</i>	<i>P Value</i>
China	Hong Kong	0.02
Pakistan	Karachi	0.02
Pakistan	Lahore	0.03
Germany	Düsseldorf	0.01
United Kingdom	South London	0.03

BLOOM 3b		
<i>Country</i>	<i>City</i>	<i>P Value</i>
Afghanistan	Kabul	0.02
Pakistan	Lahore	0.04
Germany	Düsseldorf	0.03
Georgia	Tbilisi	0.04

Table 3: Local areas with a significant difference in samples by time. 97.4% of local areas are consistent across samples using a two-tailed t-test.

Conclusions

Geographic skew in training corpora persists
downstream in trained models

Thanks!

Jonathan Dunn

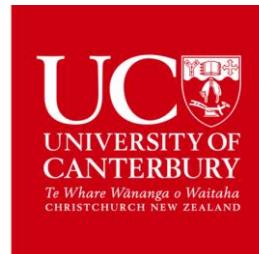
jedunn@illinois.edu

www.jdunn.name



Ben Adams

benjamin.adams@canterbury.ac.nz



Harish Tayyar Madabushi

htm43@bath.ac.uk

www.harishtayyarmadabushi.com

