The ‘Out-of-Africa’ hypothesis

Peter McCullagh

Department of Statistics
University of Chicago

South Africa, November 2013

Outline

Atkinson paper

- Phonemic inventories and linguistic diffusion hypothesis
- WALS data
- Fitting Atkinson's cline model

Spatial and variance-component models

Controversies: Feb/Mar 2012

Map No. 4045 Rev. 7 UNITED NATIONS
November 2011

Department of Field Support
Cartographic Section

Quentin Atkinson paper: Science 2011

Linguistic diffusion hypothesis:
- Single point of origin (where?)
- diversity process accumulates like genetic mutation
- language spreads by migration/diffusion

At migration/fragmentation:
- the ancestral population maintains linguistic inventory
- the migrating group loses inventory

Implications of the diversity/inventory hypothesis:
- older languages have greater vowel inventory
- older languages have greater consonant inventory
- older languages have greater tone inventory
- older languages have greater phoneme inventory

Greater phoneme inventory $\leftrightarrow$ closer to geographic source
Less phoneme inventory $\leftrightarrow$ more distant from source

$\Rightarrow$ Phoneme inventory gradient

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

Atkinson's WALS data

From *World Atlas of Language Structures*

1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

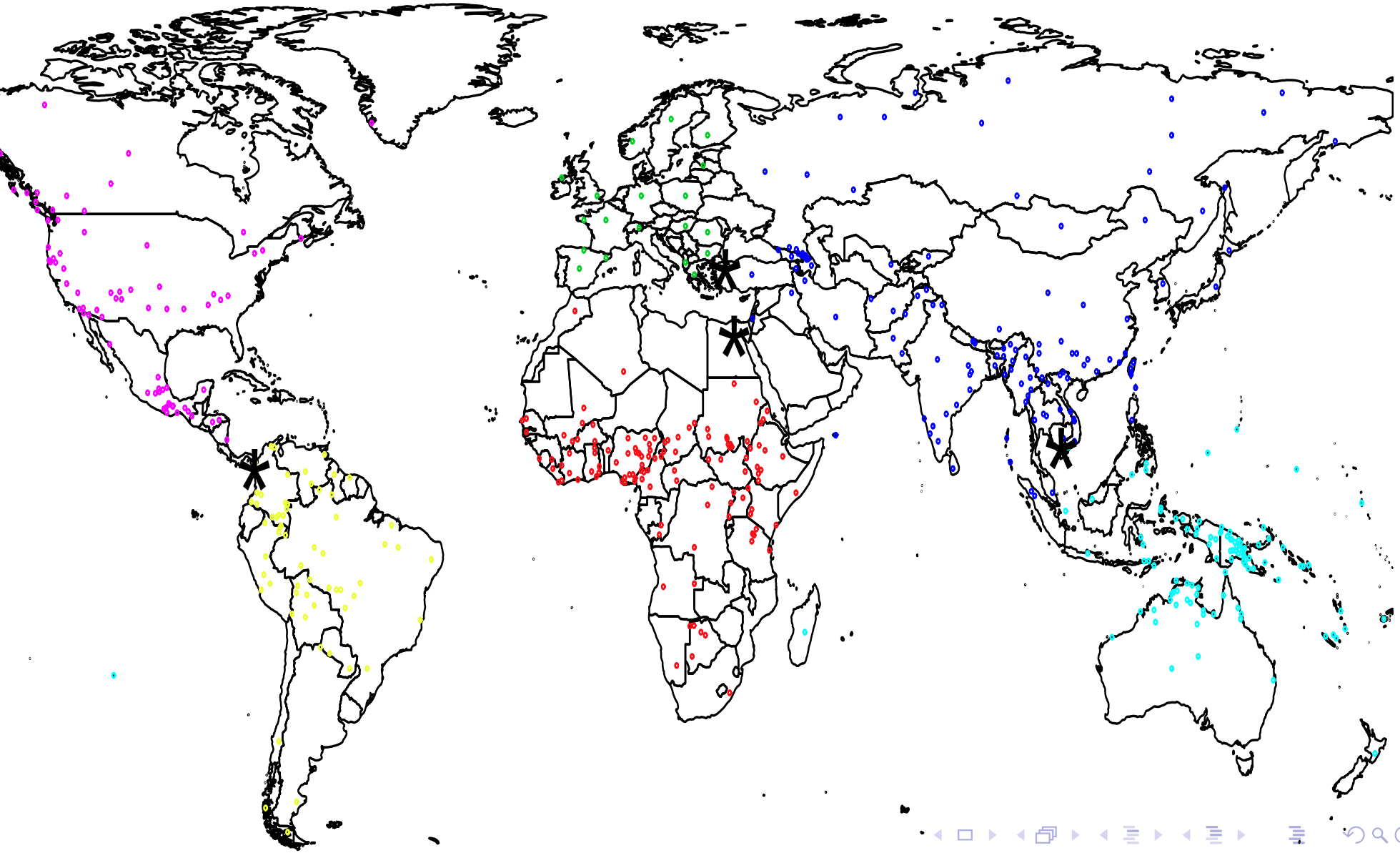
Atkinson's WALS data

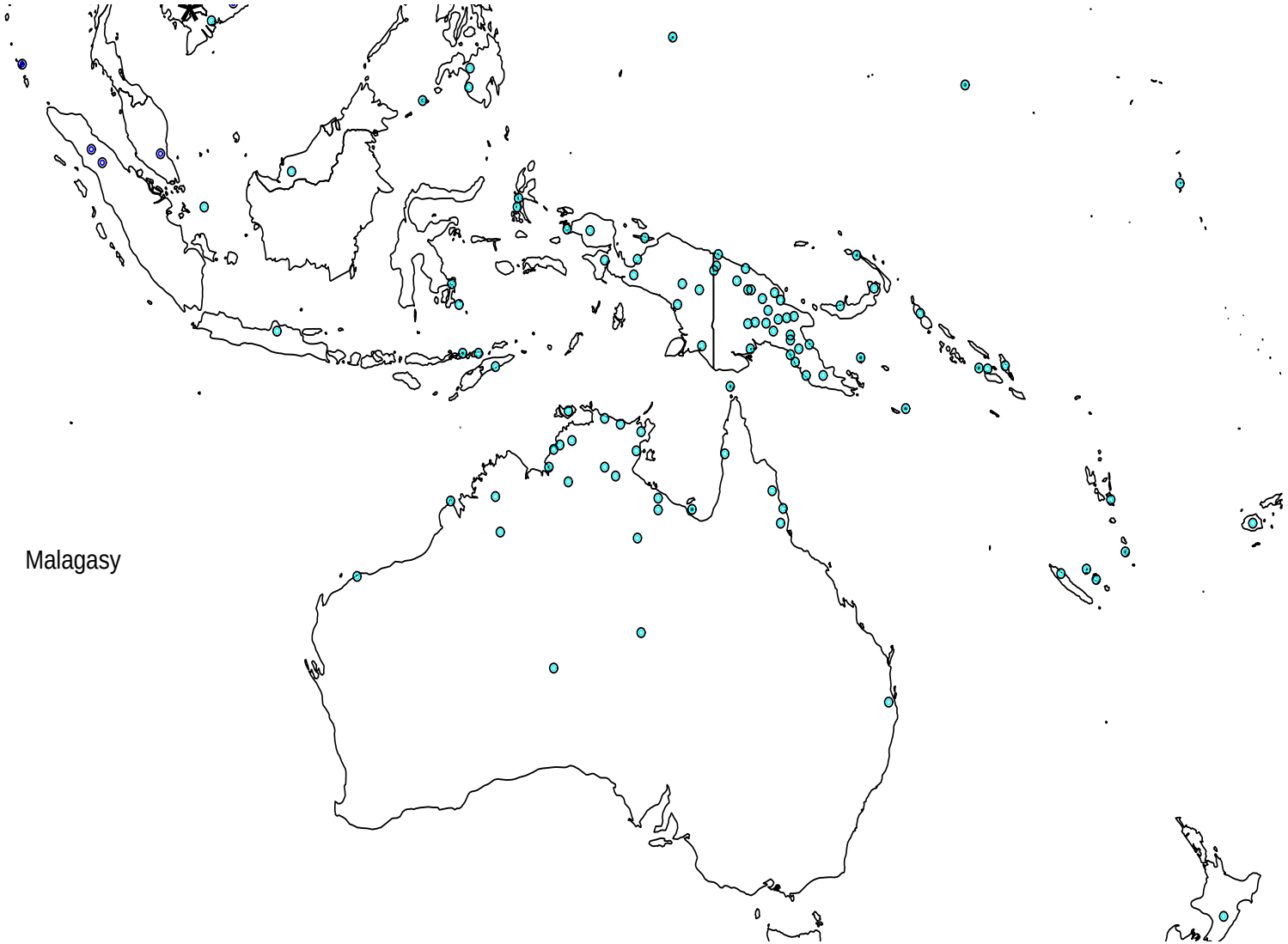
From *World Atlas of Language Structures*

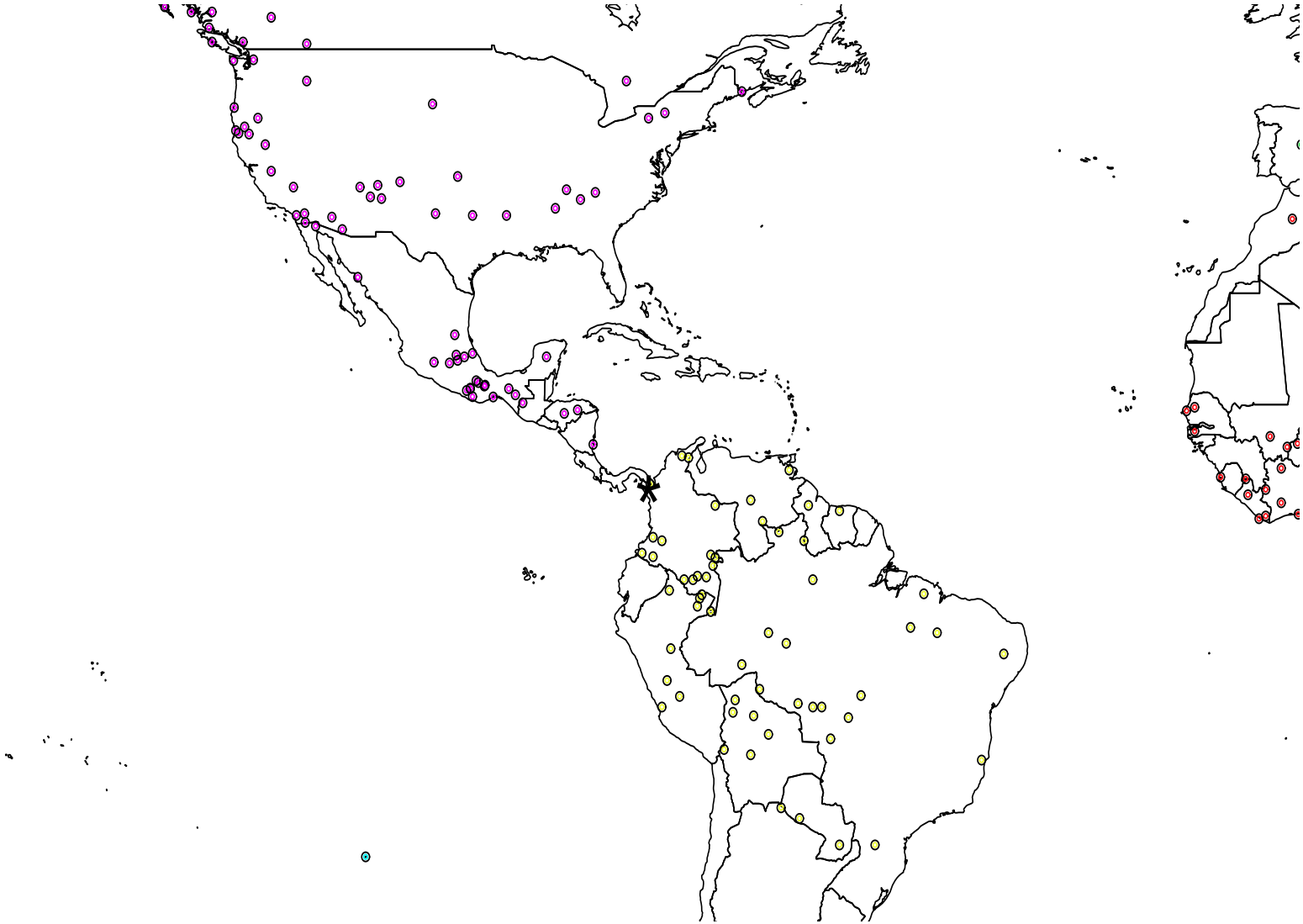
1. `Lname` Abkhaz, Aikan??, B?ⓒt?ⓒ, . . ., Zuni
2. `Fam` Arawakan, Indo-European, Sino-Tibetan,...
3. `Lat` as a decimal number: -12.67
4. `Long` e.g. -60.67 (meaning $60^\circ\ 40'$W)
5. `Nvd` Normalized vowel diversity: WALS feature No. 2
6. `Ncd` Normalized consonant diversity: WALS feature No. 1
7. `Ntd` Normalized tone diversity: WALS feature No. 13
8. `Tnpd` Total normalized phoneme diversity
9. `Popn` Estimated speaker population: integer 1–873 014 298
10. `Dbo` Distance in km. from Atkinson's best-fit origin

504 languages; 109 families Niger-Congo 62; Austronesian (42); Indo-European (42); Afro-Asiatic (??)

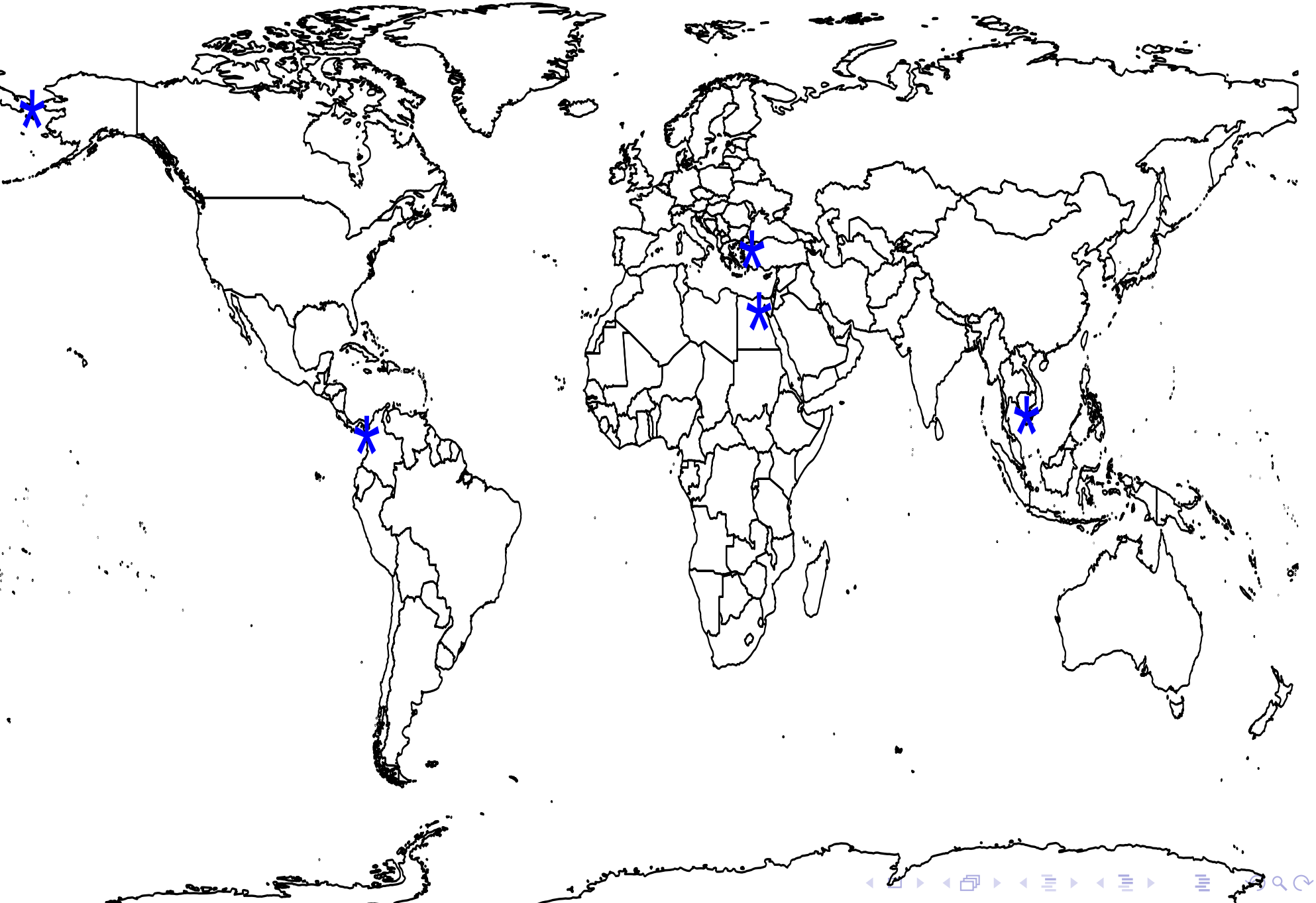
Worldwide language distribution



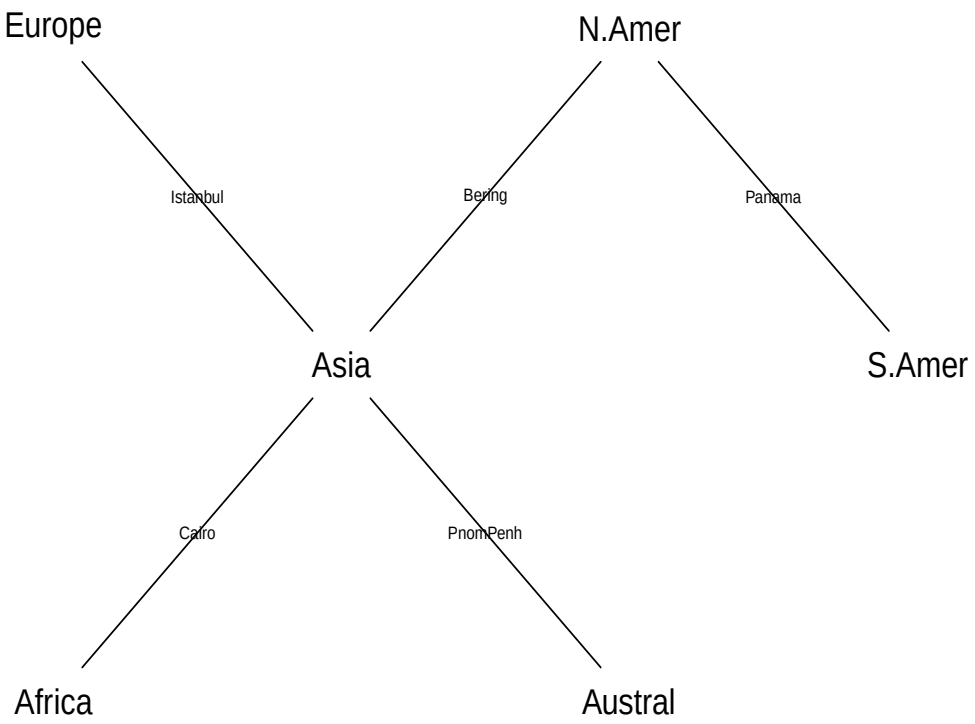

Malagasy



Restrictions on migration routes: choke points



Restrictions on migration routes: choke points



A sample of Atkinson's WALS data

Lname	WALS	Fam	Lat	Long	Tnpd	ISO	Popn
Abkhaz	abk	Nw Caucasian	43.08	41	-1.186	abk	105952
Acoma	aco	Keresan	34.92	-107.58	0.220	kjq	3391
Andoke	adk	Andoke	-0.67	-72	0.127	ano	619
Aleut E	aea	Eskimo-Aleut	54.75	-164	-0.348	ale	490
Arabic (Egyptian)	aeg	Afro-Asiatic	30	31	-0.377	arz	46321000
Aghem	agh	Niger-Congo	6.67	10	0.965	agq	26727
Ahtna	aht	Na-Dene	62	-145	-0.936	aht	80
Aikan??	aik	Arawakan	-12.67	-60.67	0.061	tba	90
Ainu	ain	Ainu	43	143	0.061	ain	15
Akan	akn	Niger-Congo	6.5	-1.25	0.127	aka	8300000
Akawaio	akw	Cariban	6	-59.5	-0.311	ake	5000
Alamblak	ala	Sepik	-4.67	143.33	0.247	amp	1527
Albanian	alb	Indo-European	41	20	0.247	aln, als	5823075
Alladian	ald	Niger-Congo	5.17	-4.33	0.127	ald	23000
Alawa	alw	Australian	-15.17	134.25	-0.348	alh	17
Amele	ame	Trans-New Guinea	-5.25	145.58	0.181	aey	5300
Amharic	amh	Afro-Asiatic	10	38	0.526	amh	17417913
Amo	amo	Niger-Congo	10.33	8.67	0.686	amo	12263
Arrernte	amp	Australian	-24	136	-0.348	aer	2175
Amuesha	amu	Arawakan	-10.5	-75.42	-0.348	ame	9831
Amuzgo	amz	Oto-Manguean	16.83	-98	0.845	amu	23000
Araona	ana	Tacanan	-12.33	-67.75	-0.069	aro	81
Angas	anc	Afro-Asiatic	9.5	9.5	0.566	anc	400000
//Ani	ani	Khoisan	-18.92	21.92	-0.218	hnh	1000
Angaatiha	ant	Trans-New Guinea	-7.22	146.25	0.339	agm	2100
Yulu	yul	Nilo-Saharan	8.5	25.25	1.682	yul	7000
Yurok	yur	Algic	41.33	-124	-0.098	yur	12
Yupik	yus	Eskimo-Aleut	65	-173	-0.348	ess	1350
Zande	zan	Niger-Congo	4	26	0.685	zne	1142000
Zoque	zqc	Mixe-Zoque	17	-93.25	-0.657	zoc	10000
Zulu	zul	Niger-Congo	-30	30	0.339	zul	9563422
Zuni	zun	Zuni	35.08	-108.83	-0.377	zun	9651

Statistical modelling: Atkinson style

Response: Total normalized phoneme inventory Y_i
Geographic coords: $x_i = (x_1, x_2)_i$ of language i
No of speakers: s_i
Language origin coordinates: $\tau = (\tau_0, \tau_1)$

$$E(Y_i) = \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i)$$
$$\mathrm{cov}(Y_i, Y_j) = \sigma^2 \delta_{ij}$$

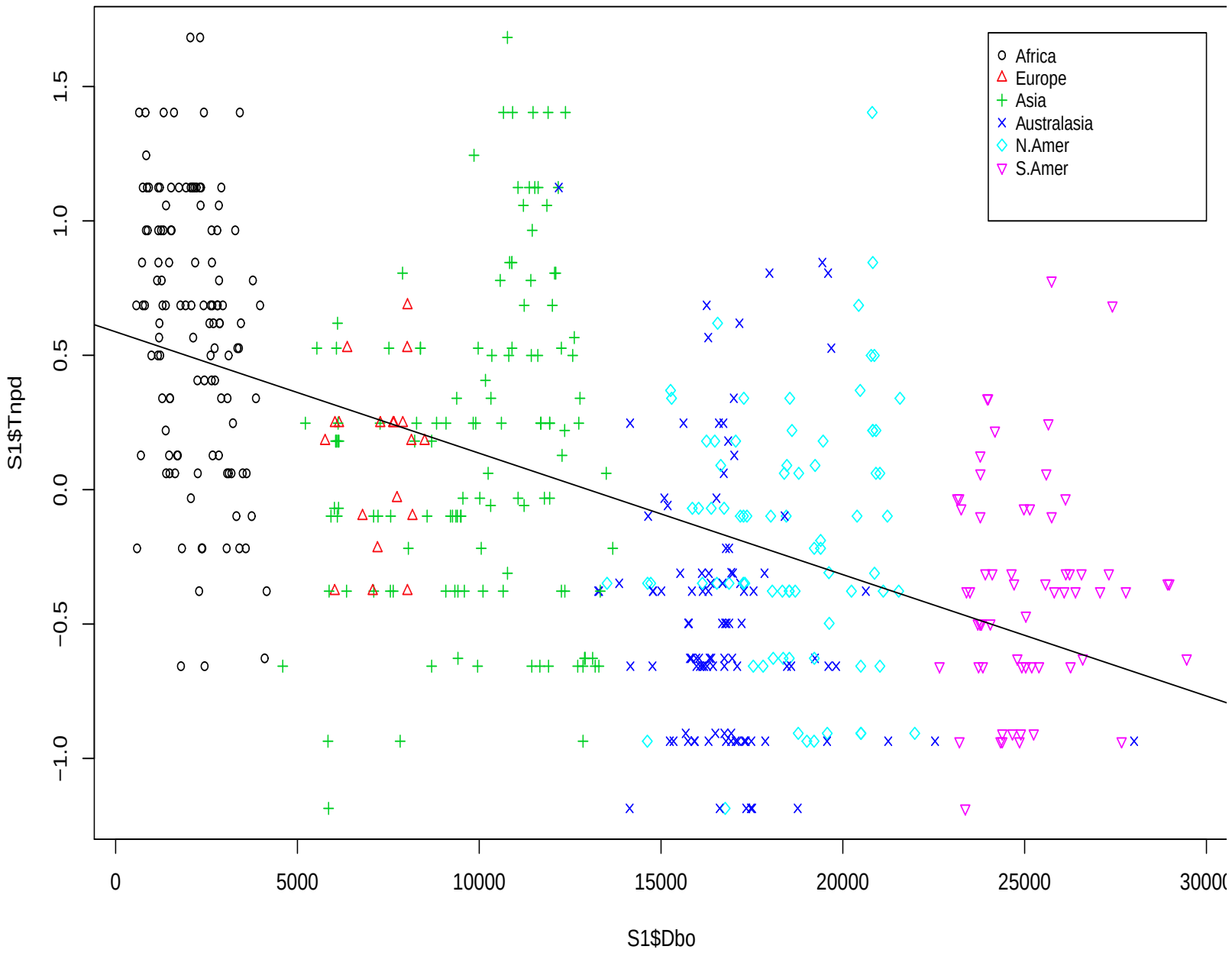
Parameters: $\beta_0, \beta_1, \gamma, \tau, \sigma^2$

distance $\|x_i - \tau\|$ versus log distance

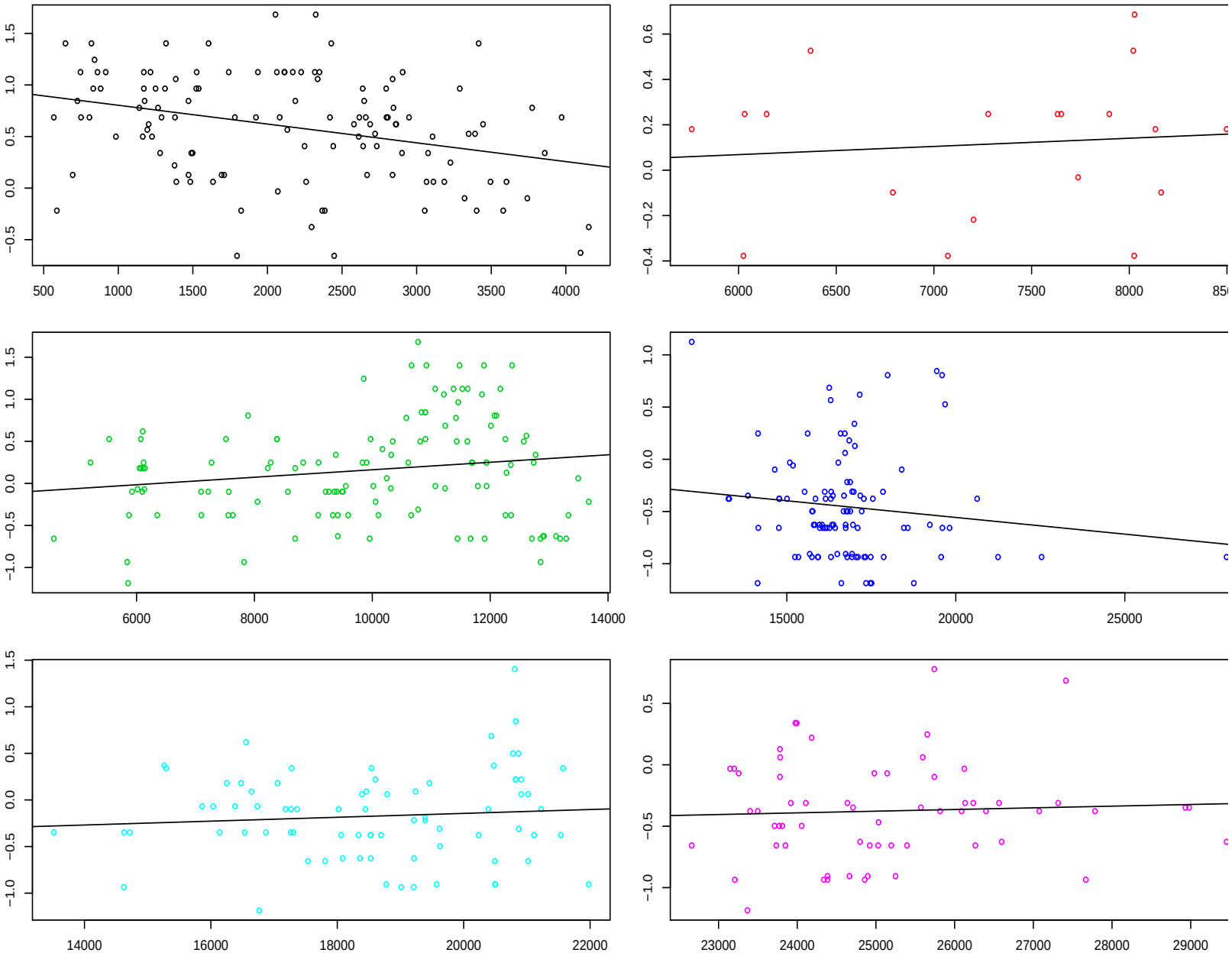
RSS as a function of τ or F-ratio

$$F(\tau) = \frac{(\mathrm{RSS}(\tau) - \mathrm{RSS}(\hat{\tau}))/2}{\mathrm{RSS}(\hat{\tau})/(n-5)}$$

Phonemic inventory versus distance from $\hat{\tau}$



Phonemic inventory versus distance (by region)



Consistency of pattern?

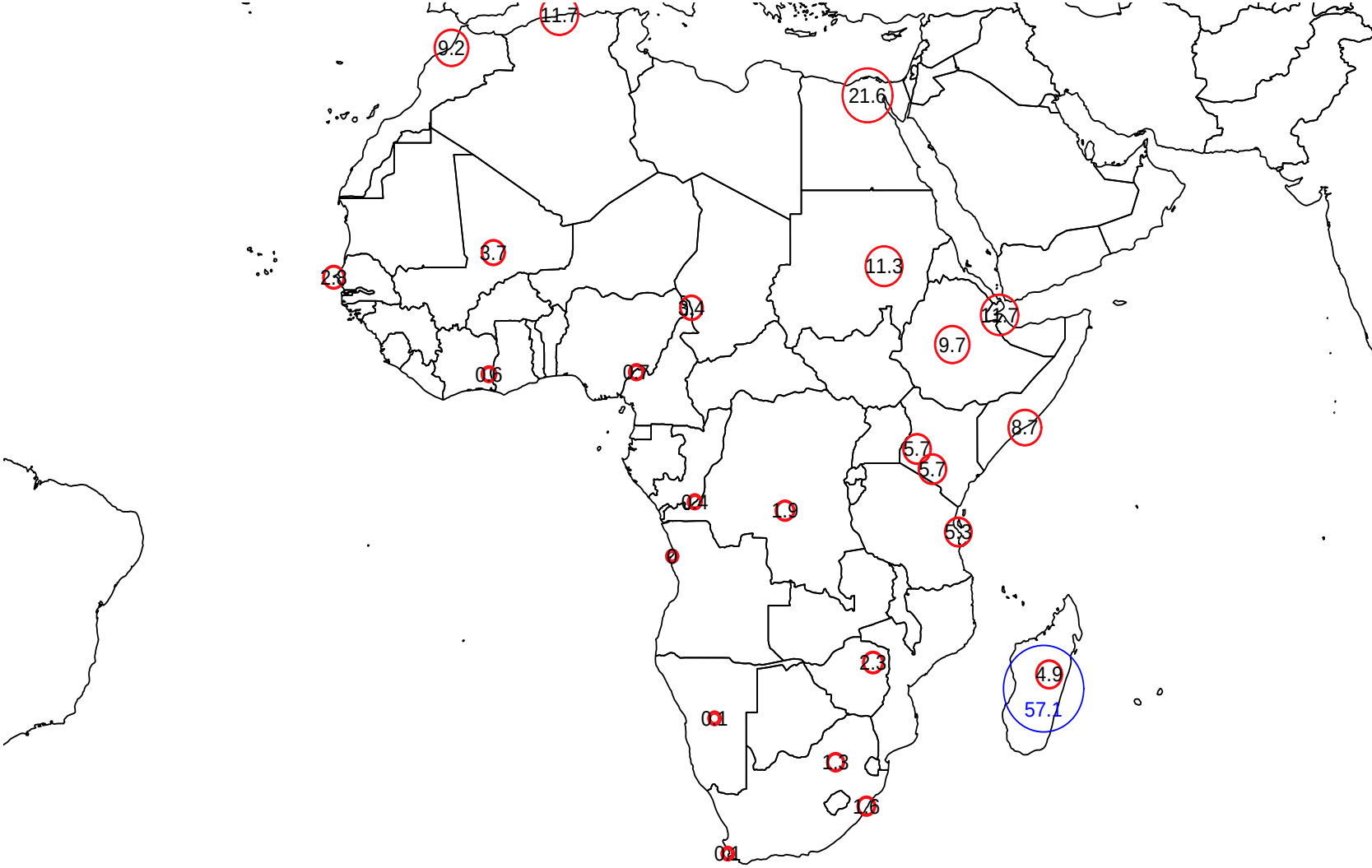
Test for consistency of pattern across continents:

$$
\begin{aligned}
\text{Null:} \quad E(Y) &= \text{dist} + \log(\text{pop}) \\
\text{Alt:} \quad E(Y) &= \text{Reg} + \text{Reg.dist} + \log(\text{pop}) \\
F &= \frac{\text{Reduction in RSS}/(\text{Reduction in df})}{\text{Resid SS}/\text{Resid df}}
\end{aligned}
$$

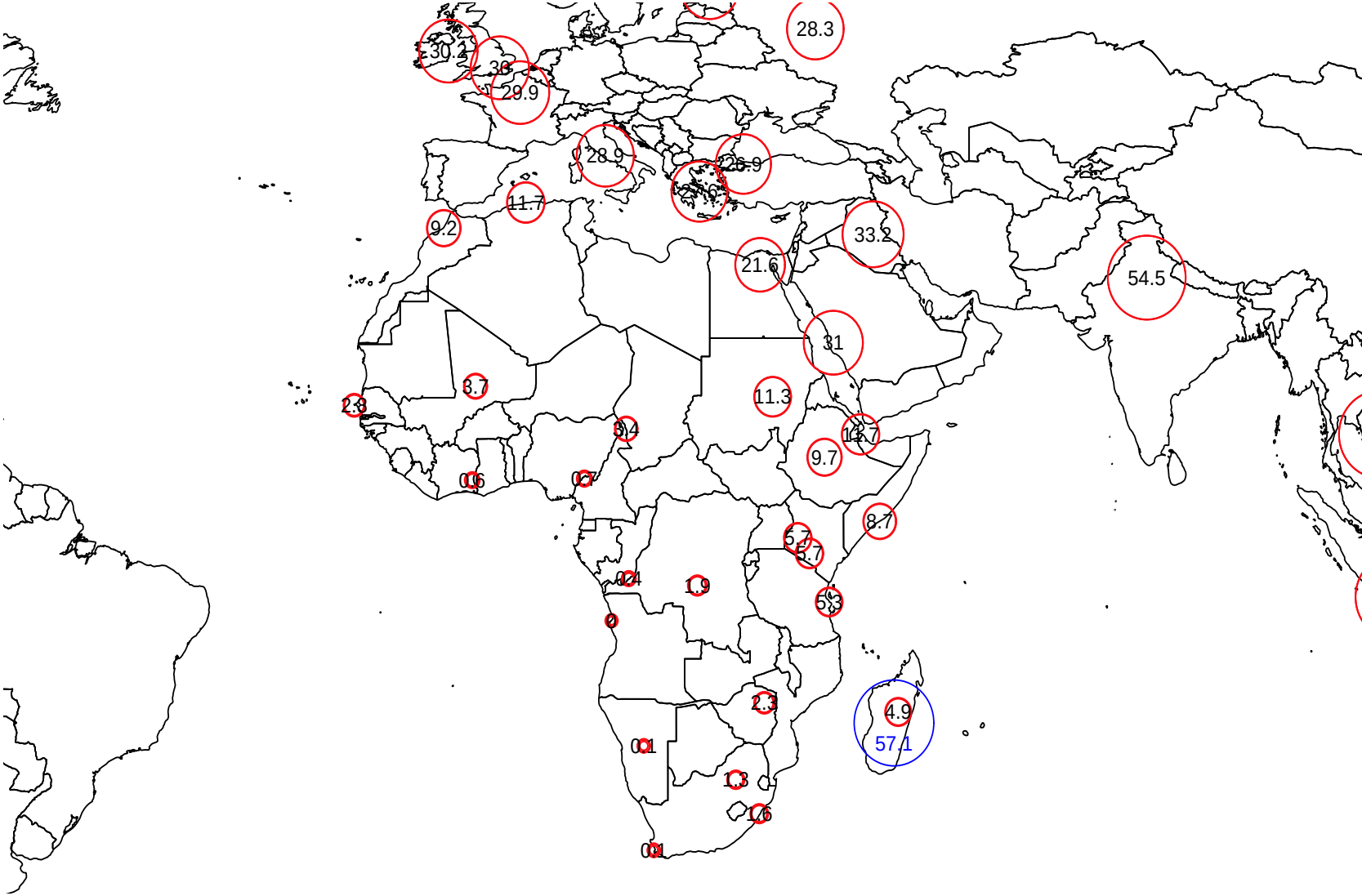
F-ratio = 8.6 on (10, 491) df. (should be < 2)

No clear evidence of a consistent cline

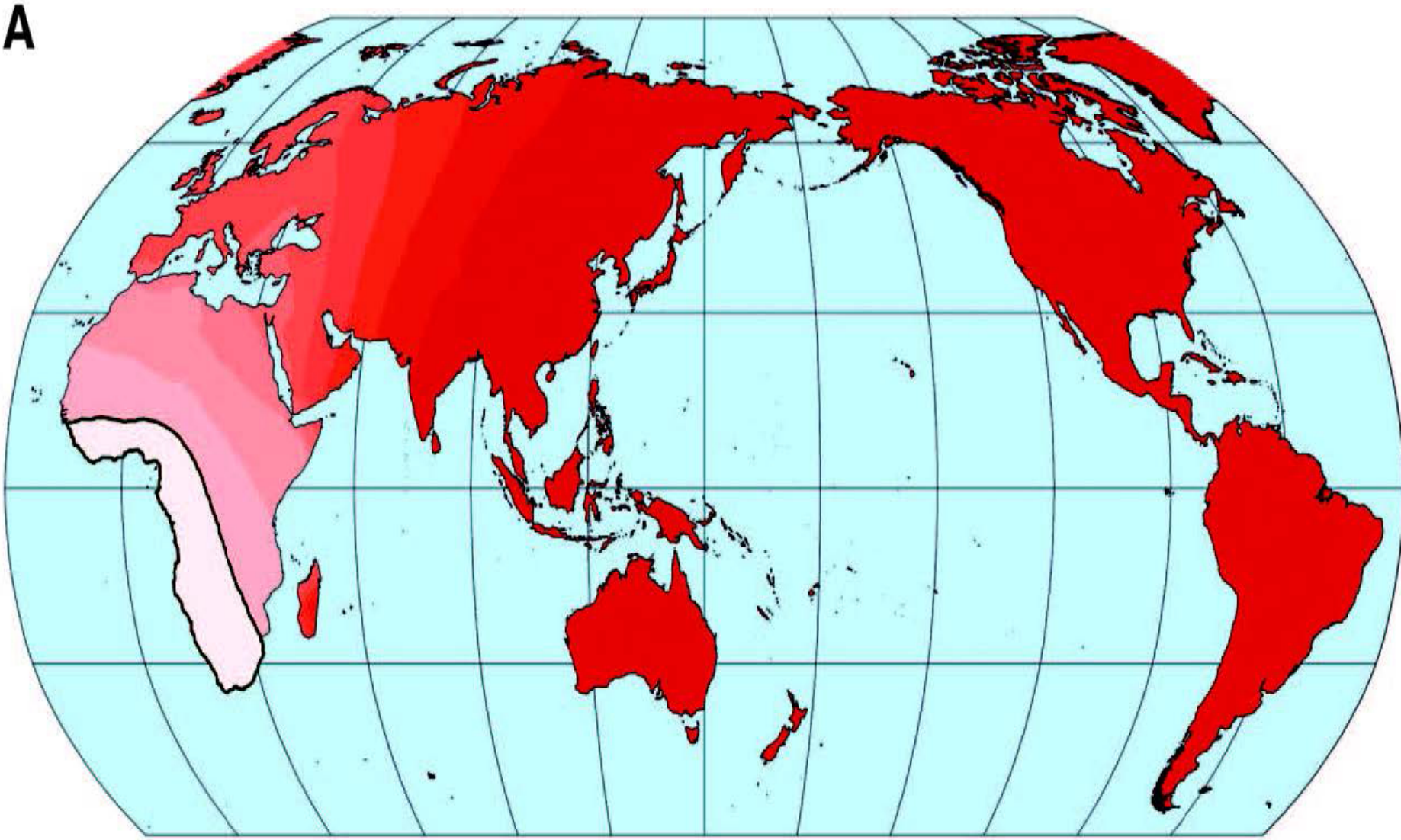
RSS (F-ratio) as a function of origin τ



RSS (F-ratio) as a function of origin τ



Atkinson's posterior density for τ



Confidence regions, HPD regions

Fisher's F-ratio and the excess variance criterion

$$F(\tau) = \frac{(\mathrm{RSS}(\tau) - \mathrm{RSS}(\hat{\tau}))/2}{\mathrm{RSS}(\hat{\tau})/(n-5)}$$

Distributed as $F_{2,n-5}$ approx; 95% quantile $F < 3.01$

BIC $\equiv$ LR:

$$n \log \mathrm{RSS}(\tau) - n \log(\mathrm{RSS}(\hat{\tau}) = n \log(1 + 2F/(n-5)) \simeq 2F$$

distributed as $2F_{2,n-5} \sim \chi^2_2$ approx

95% frequentist region LR$(\tau) < 6$
Atkinson's region as illustrated: BIC < 4
whence 4 BIC units?
coverage prob: 86%

Further modelling considerations

$$E(Y_i) = \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i)$$
$$\mathrm{cov}(Y_i, Y_j) = \sigma^2 \delta_{i,j}$$

- Diversity implies $\beta_1 \leq 0$: F-ratio does not distinguish
  In fact: $\hat{\beta}_1 < 0$ for Africa, Europe, Asia
  $\hat{\beta}_1 > 0$ for America, Oceania (incl Malagasy)
- Covariates versus relationships
  covariate $i \mapsto s_i$ (number of speakers)
  $(i, j) \mapsto R_1(i, j)$ same linguistic family
  $(i, j) \mapsto \|x_i - x_j\|$ geographic or linguistic distance
- Independence
  Possible correlation within linguistic families
  Possible correlation due to geographic proximity
- Covariates: $E(Y_i) = \beta_0 + \beta_1 x_i + \cdots$
  Relationships: $\mathrm{cov}(Y_i, Y_j) = \sigma_0^2 \delta_{i,j} + \sigma_1^2 R_1(i, j) + \cdots$

Further modelling considerations

$$E(Y_i) = \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i)$$
$$\mathrm{cov}(Y_i, Y_j) = \sigma^2 \delta_{i,j}$$

- Diversity implies $\beta_1 \leq 0$: F-ratio does not distinguish
  In fact: $\hat{\beta}_1 < 0$ for Africa, Europe, Asia
  $\hat{\beta}_1 > 0$ for America, Oceania (incl Malagasy)
- Covariates versus relationships
  covariate $i \mapsto s_i$ (number of speakers)
  $(i, j) \mapsto R_1(i, j)$ same linguistic family
  $(i, j) \mapsto \|x_i - x_j\|$ geographic or linguistic distance
- Independence
  Possible correlation within linguistic families
  Possible correlation due to geographic proximity
- Covariates: $E(Y_i) = \beta_0 + \beta_1 x_i + \cdots$
  Relationships: $\mathrm{cov}(Y_i, Y_j) = \sigma_0^2 \delta_{i,j} + \sigma_1^2 R_1(i, j) + \cdots$

Further modelling considerations

$$E(Y_i) = \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i)$$
$$\mathrm{cov}(Y_i, Y_j) = \sigma^2 \delta_{i,j}$$

- Diversity implies $\beta_1 \leq 0$: F-ratio does not distinguish
  In fact: $\hat{\beta}_1 < 0$ for Africa, Europe, Asia
  $\hat{\beta}_1 > 0$ for America, Oceania (incl Malagasy)
- Covariates versus relationships
  covariate $i \mapsto s_i$ (number of speakers)
  $(i, j) \mapsto R_1(i, j)$ same linguistic family
  $(i, j) \mapsto \|x_i - x_j\|$ geographic or linguistic distance
- Independence
  Possible correlation within linguistic families
  Possible correlation due to geographic proximity
- Covariates: $E(Y_i) = \beta_0 + \beta_1 x_i + \cdots$
  Relationships: $\mathrm{cov}(Y_i, Y_j) = \sigma_0^2 \delta_{i,j} + \sigma_1^2 R_1(i, j) + \cdots$

Further modelling considerations

$$E(Y_i) = \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i)$$
$$\mathrm{cov}(Y_i, Y_j) = \sigma^2 \delta_{i,j}$$

- Diversity implies $\beta_1 \leq 0$: F-ratio does not distinguish
  In fact: $\hat{\beta}_1 < 0$ for Africa, Europe, Asia
  $\hat{\beta}_1 > 0$ for America, Oceania (incl Malagasy)
- Covariates versus relationships
  covariate $i \mapsto s_i$ (number of speakers)
  $(i, j) \mapsto R_1(i, j)$ same linguistic family
  $(i, j) \mapsto \|x_i - x_j\|$ geographic or linguistic distance
- Independence
  Possible correlation within linguistic families
  Possible correlation due to geographic proximity
- Covariates: $E(Y_i) = \beta_0 + \beta_1 x_i + \cdots$
  Relationships: $\mathrm{cov}(Y_i, Y_j) = \sigma_0^2 \delta_{i,j} + \sigma_1^2 R_1(i, j) + \cdots$

A linear Gaussian variance-components model

$$
\begin{aligned}
E(Y_i) &= \beta_0 + \beta_1 \|x_i - \tau\| + \gamma \log(s_i) \\
\mathrm{cov}(Y_i, Y_j) &= \sigma^2 \delta_{i,j} + \sigma_1^2 F_{ij} + \sigma_2^2 \exp(-d_{ij}/\rho)
\end{aligned}
$$

$\tau = (9.5, -1.25)$; $\rho = 1000$ km fixed

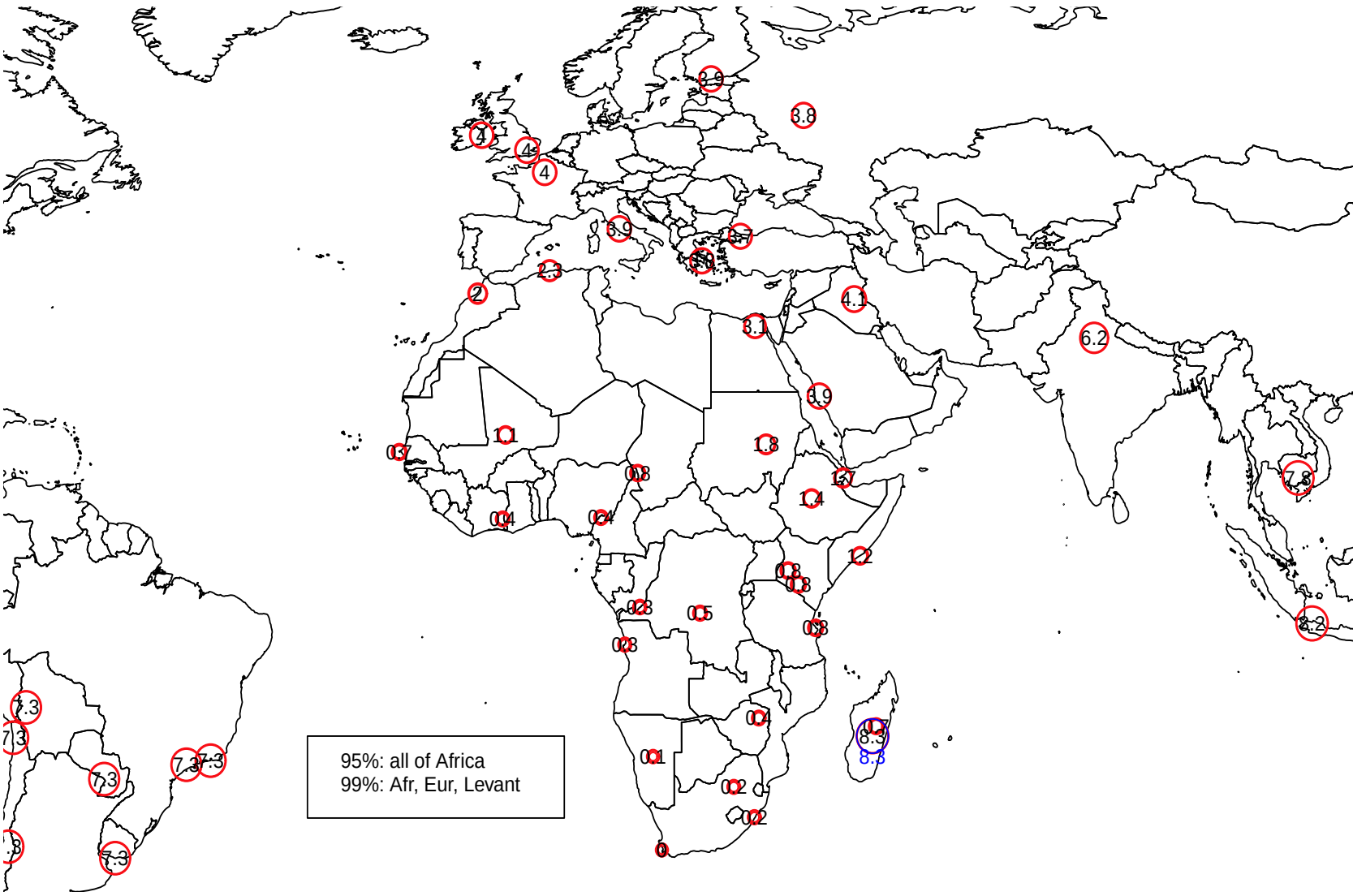
Model comparison via likelihood-ratio statistics

	Covariance model				
	δ	F	V	$F+V$	$+F.V$
LR $*$ 2	0.00	88.80	127.17	143.69	147.70
$\hat{\beta}_1 \times 10^4$	-0.391	-0.313	-0.351	-0.331	-0.334
SE	0.036	0.057	0.077	0.079	0.074

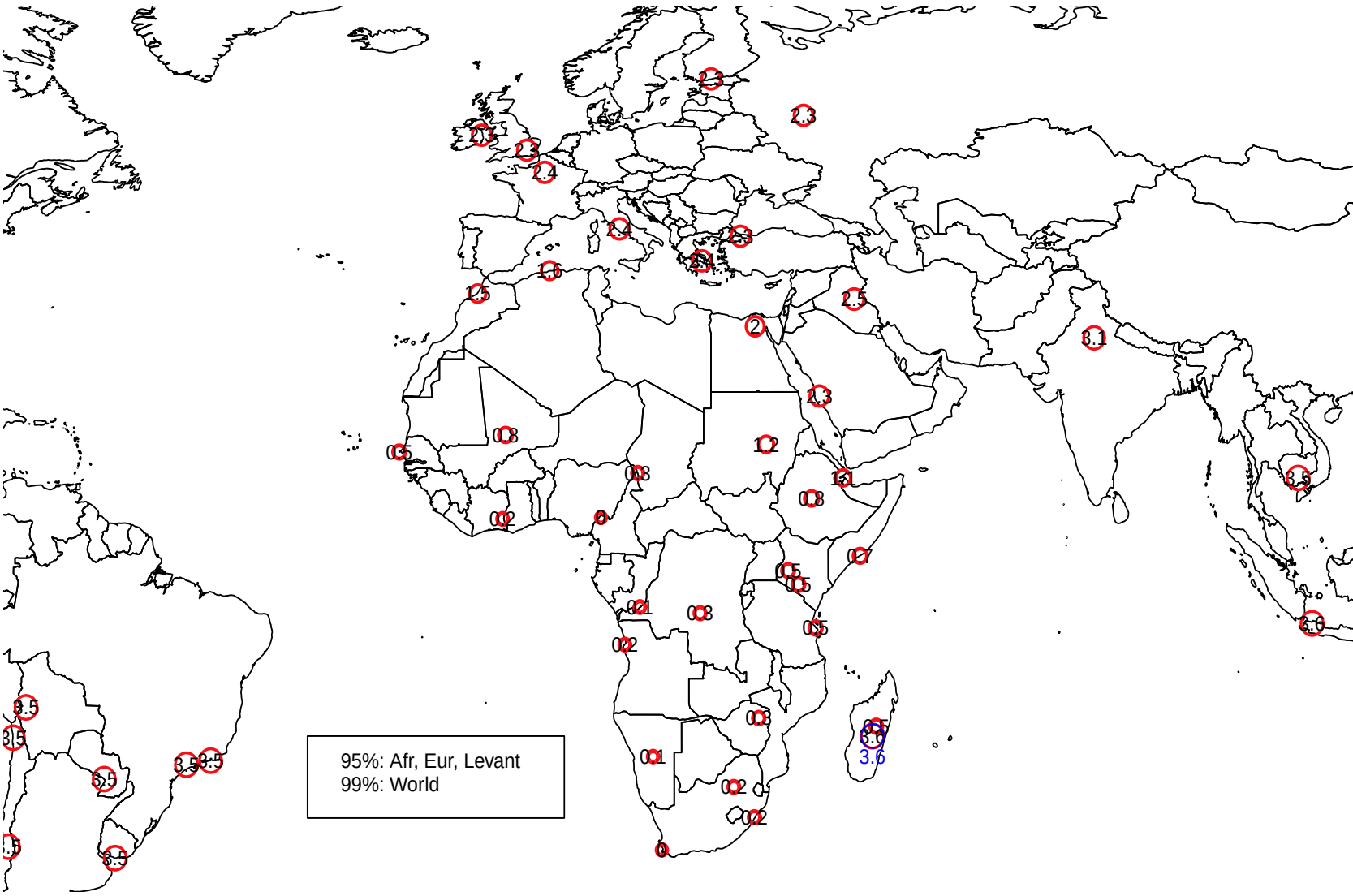
Conclusions:

- Evidence for both family and geographic correlatons
- Dbo coefficient $\hat{\beta}_1$ not much affected
- Ordinary least-squares SE is misleading

Confidence regions for source (exponential model)



Confidence regions for source (Bessel0 model)



Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Summary

- Distinction between covariates and relationships
- Rel_1: family as a block factor F_{ij}
- Rel_2: geographic separation D_{ij}
- Spatial covariance models: Matérn's Bessel family
- Effect of correlation on OLS coefficients
- Effect of correlation on SE of OLS coefficients
- Selection of languages for inclusion in analysis
- Choice of total phoneme inventory as response
- Alternative forms of data using specific phonenes

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

Published commentary: *Science* Feb/Mar 2012

Four commentaries: Munich, Shanghai, Stanford, Rochester/MIT

Summary of criticisms

- WALS data: vowels and tones exaggerated: (M, Sh, St, R)
  (1 tone $\equiv$ 2.2 vowels $\equiv$ 5.7 consonants)
- Inconsistency of pattern: cline in Africa but not elsewhere (St)
- Inconsistency of pattern for v/c/t: (Sh, St)
- BIC-4 allowance: (M)
- Neighbour influence (St)
- Global geography (R)

References

Atkinson, Q.D. (2011) *Science Mag* 332, 346.
Cysouw, M., Dediu, D. and Moran, S. (2012) Technical comment. *Science* 335, 657-b.
Wang, C.-C, Ding, Q-L, Tao, H. and Li, H. (2012) Technical comment. *Science* 335, 657-c.
van Tuyl, R. and Pereltsvaig, A. (2012) Technical comment. *Science* 335, 657-d.
Atkinson, Q.D. (2012) Response. *Science Mag* 335, 657-e
Jaeger, T. Pontillo, D. and Graff, P. (2012) Technical comment. *Science* 335, 1042-ad.
Atkinson, Q.D. (2012) Response. *Science Mag* 335, 1042-b
`http://www.en.uni-muenchen.de/news/newsarchiv/2012/2012_`
`http://www.stat.uchicago.edu/~pmcc/prelims/2011/`
McC and Clifford, D. (2006) Evidence for conformal invariance of crop yields. *Proc. Roy. Soc. Lond. A*.
Clifford, D. and McC (2006). The regress function. *R Newsletter* **6**, 6-10.