

Proposta de análise bayesiana não linear do desenvolvimento fonético-fonológico de línguas não nativas: o caso das vogais médias do PB de um aprendiz argentino

Ronaldo Lima Jr. (UnB-CNPq)

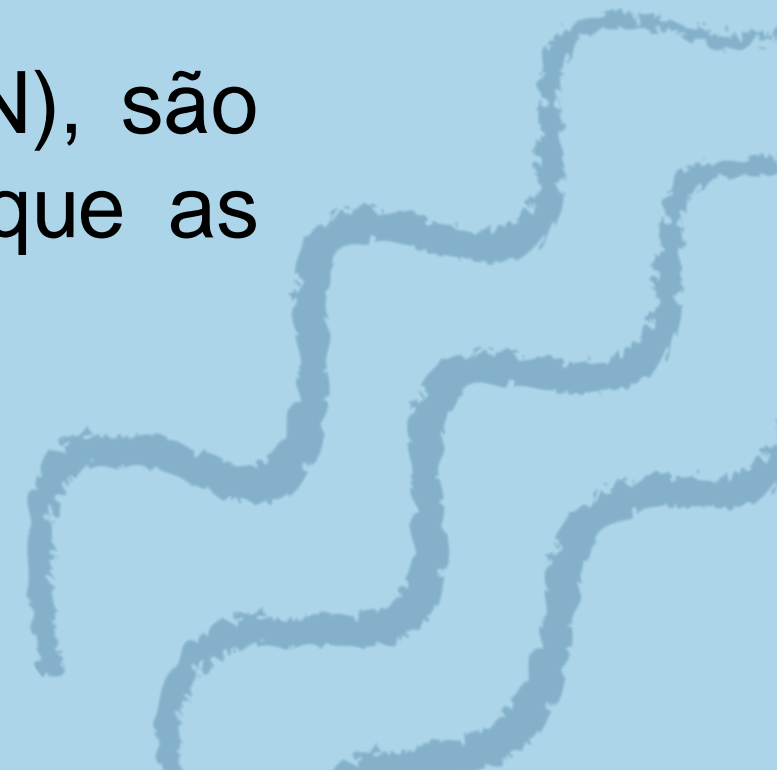
Ubiratã Kickhöfel Alves (UFRGS-CNPq)

Introdução

Na última década, a Teoria dos Sistemas Dinâmicos Complexos tem desafiado os pesquisadores a verificar novas abordagens e modelos de obtenção e análise de dados.

Ao considerarmos o desenvolvimento dos sistemas vocálicos, ao invés de questionarmos apenas se uma dada categoria foi formada, verificamos as flutuações dinâmicas e interações entre todas as categorias.

Com o desenvolvimento de novas categorias da Língua Não Nativa (LNN), são esperadas alterações em todo o espaço acústico não nativo, uma vez que as categorias interagem entre si.



Introdução

Neste trabalho, temos por objetivo analisar longitudinalmente o desenvolvimento das vogais tônicas do PB-L3 por um aprendiz argentino (inglês-L2) residente no Brasil há três anos no início das 24 coletas quinzenais de dados.

Propomos uma metodologia de análise de desenvolvimento fonético-fonológico de línguas não nativas sob as premissas da Teoria de Sistemas Dinâmicos Complexos (TSDC), partindo do princípio do desenvolvimento dinâmico e não linear de línguas (Larsen-Freeman, 1997; Lowie, 2017).

Para tanto, comparamos:

- (i) as análises dos dois primeiros formantes em Hertz e em Bark
- (ii) as distâncias Euclidiana e de Mahalanobis entre as vogais
- (iii) modelos de Regressão Linear e Aditivo Generalizado (Misto) bayesiano

Questão Norteadora

De que modo um Modelo Aditivo Generalizado Misto Bayesiano (GAMM) pode modelar as flutuações dinâmicas inerentes ao desenvolvimento do PB do aprendiz ao longo do tempo?



Metodologia



Participante

Natural da província de Buenos Aires, tinha 36 anos de idade e trabalhava como professor universitário no Brasil (Porto Alegre) há 3 anos e 7 meses quando do início da coleta de dados.

Possuía um nível A2 de proficiência em inglês e um nível “Avançado Superior” em Português Brasileiro (exame CELPE-BRAS – exame de proficiência oficial no Brasil);

Participou de 24 sessões de coleta de dados, realizadas quinzenalmente (Out 2018 – Set 2019);

Recebeu instrução de pronúncia (de acordo com o manual de Alves, Brisolara e Perozzo [2017]) acerca dos sistemas consonantal e vocálico do PB (entre o período das coletas 10 e 16). As aulas eram semanais e transcorreram ao longo de 12 semanas.

Instrumento de leitura: o mesmo aplicado em Pereyron (2017): Frases-veículo: “Diga_____”
6 palavras para cada vogal (3 repetições cada)

Análise acústica - F1, F2 e duração
(*Praat* – v. 6.1.16, Boersma & Weenink, 2019)

Plotagem das Vogais: Pacote PhonR

Resultados



Comparação (i) Formantes:

Hertz vs Bark



Hertz

Bark

F2

F2

3000 2500 2000 1500 1000 500

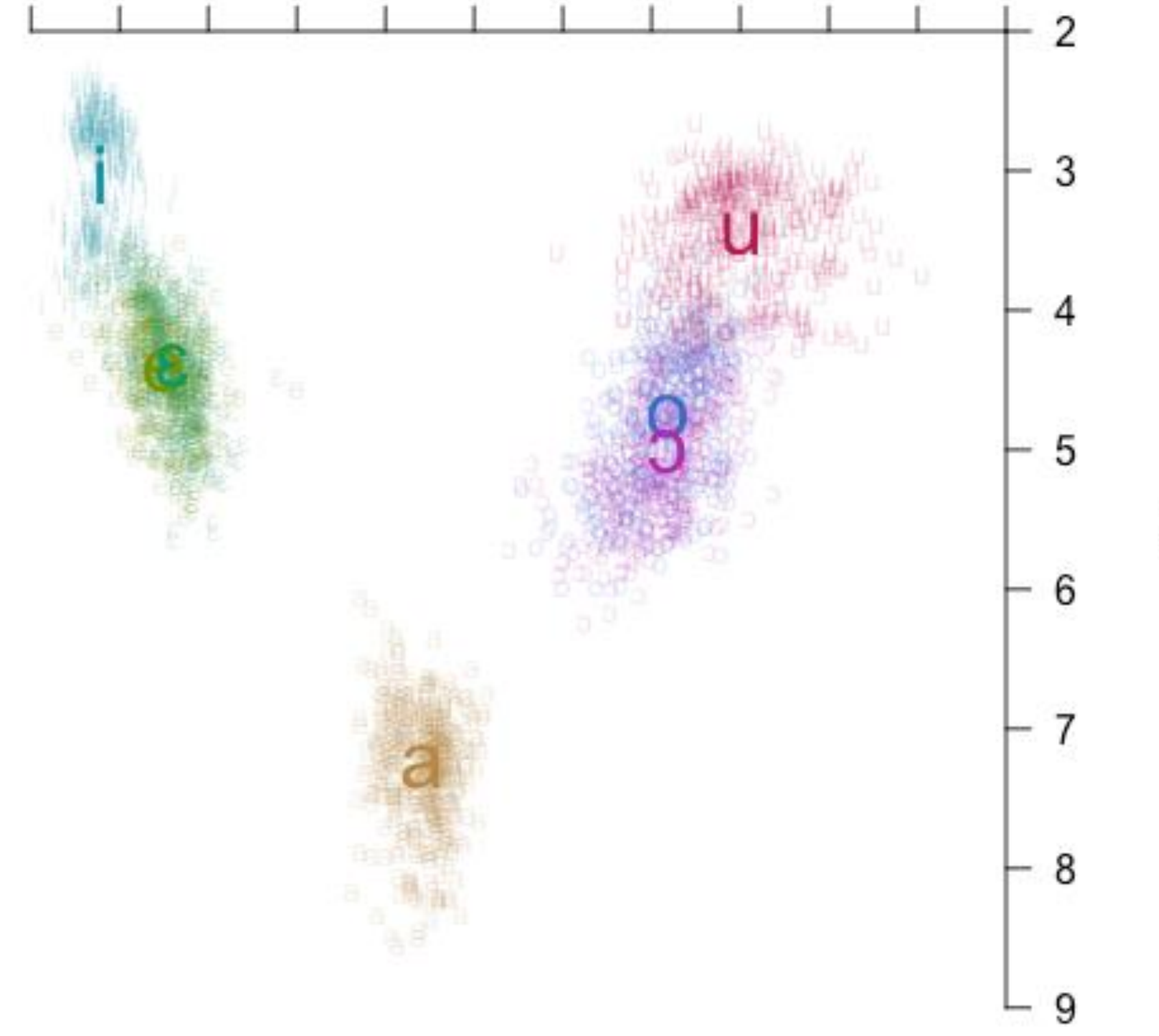
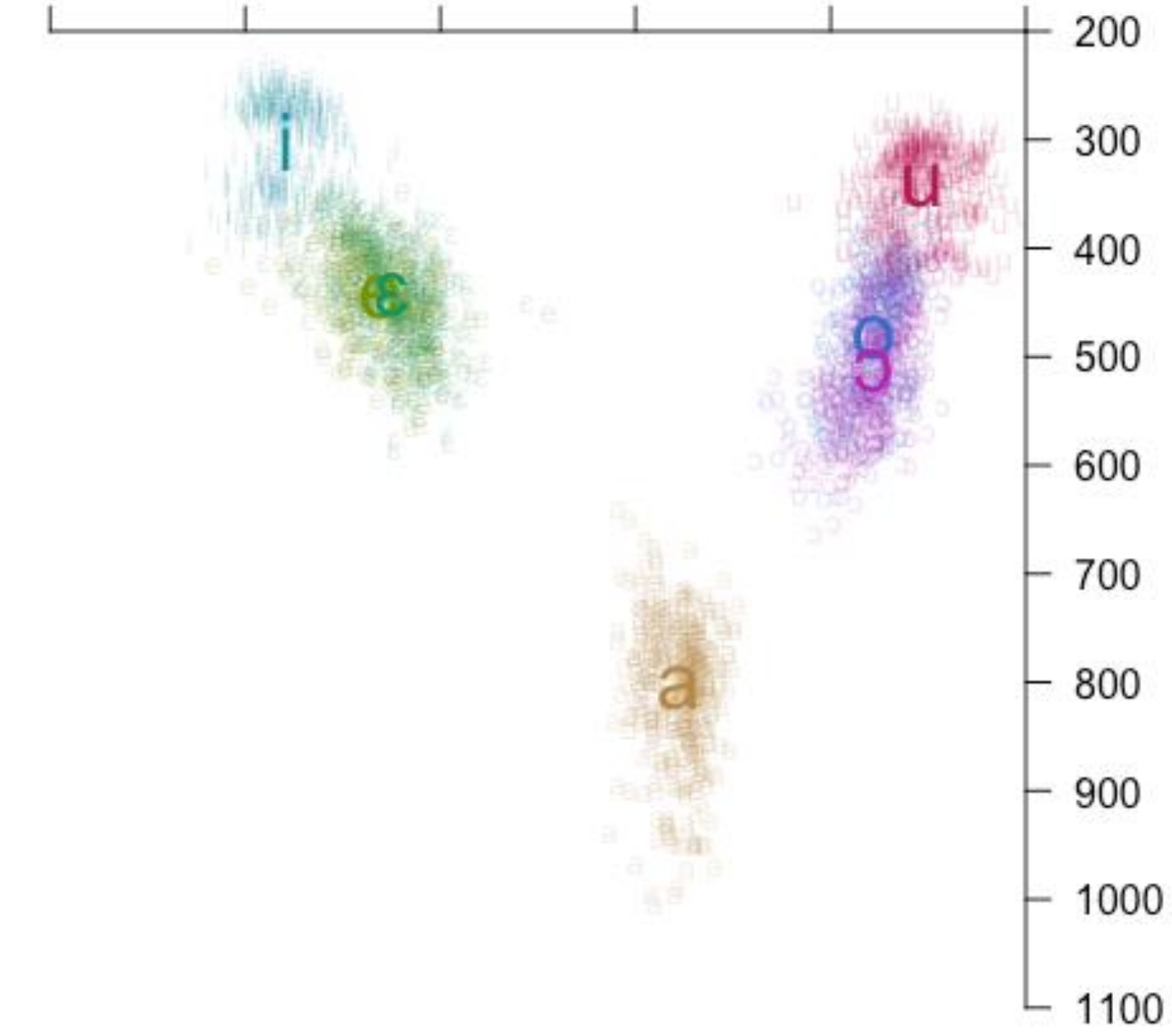
15 14 13 12 11 10 9 8 7 6 5 4

200
300
400
500
600
700
800
900
1000
1100

2
3
4
5
6
7
8
9

F1

F1



Hertz

Bark

F2

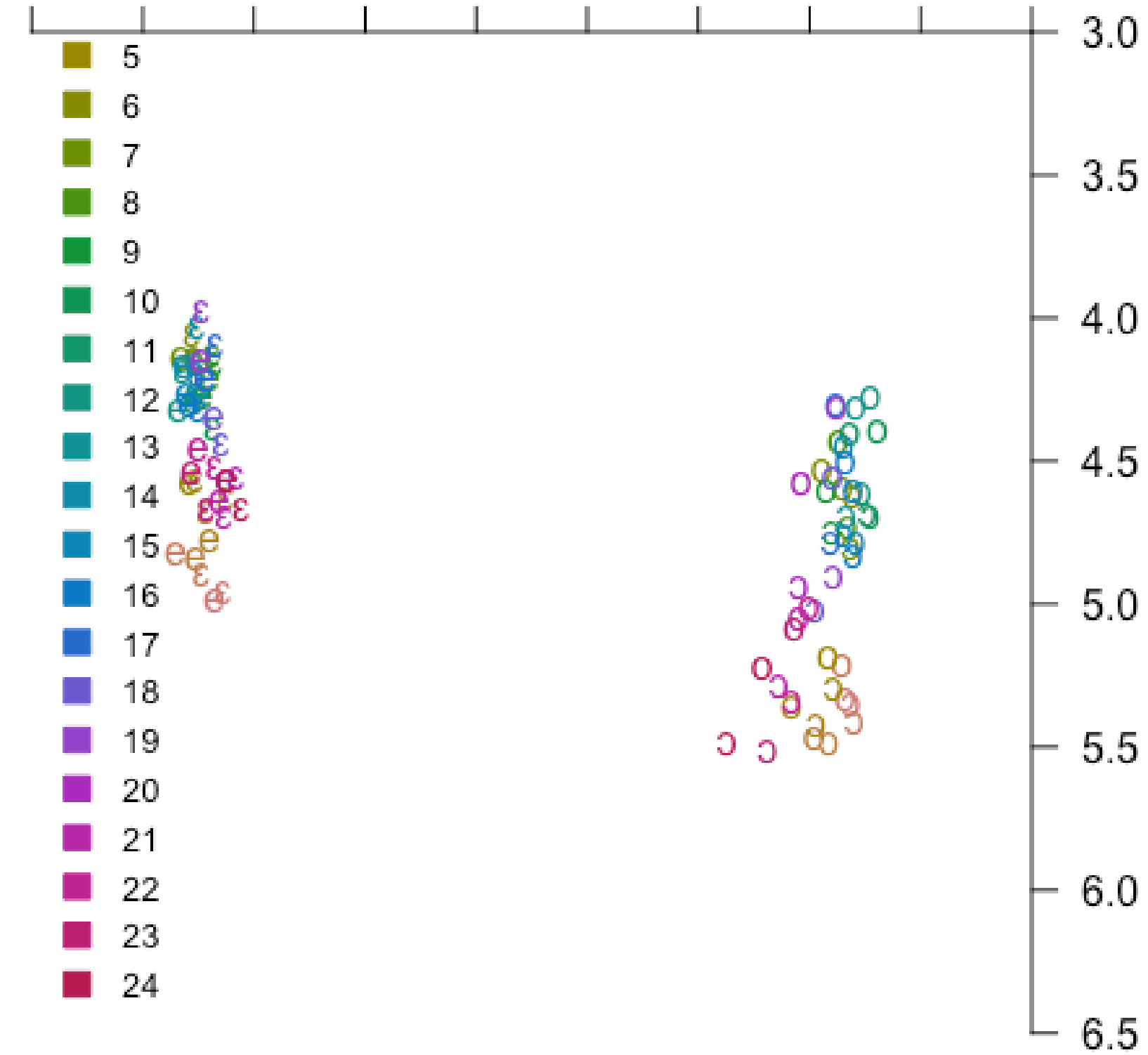
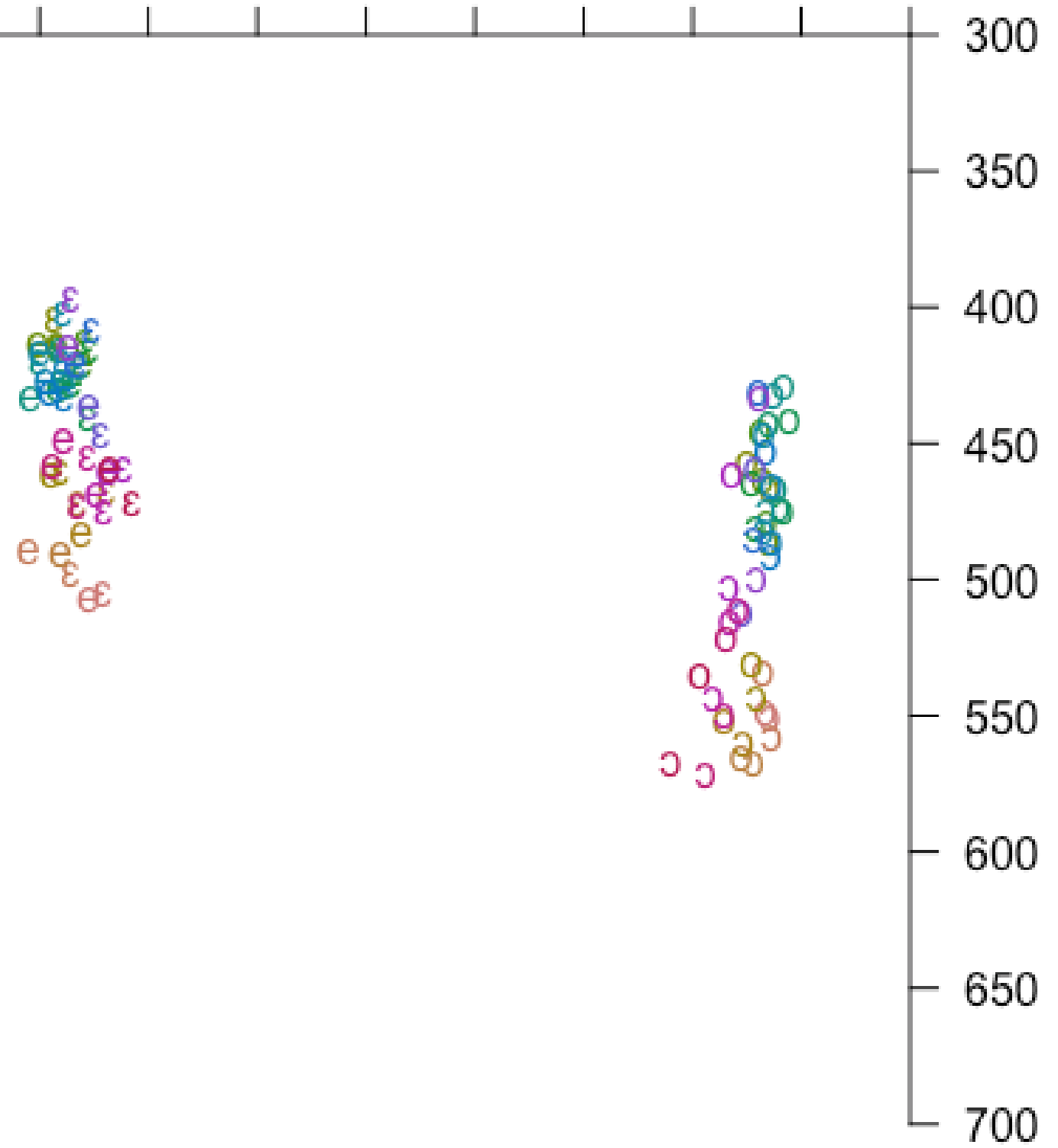
F2

2600 2200 1800 1400 1000 600

15 14 13 12 11 10 9 8 7 6

- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24

- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24



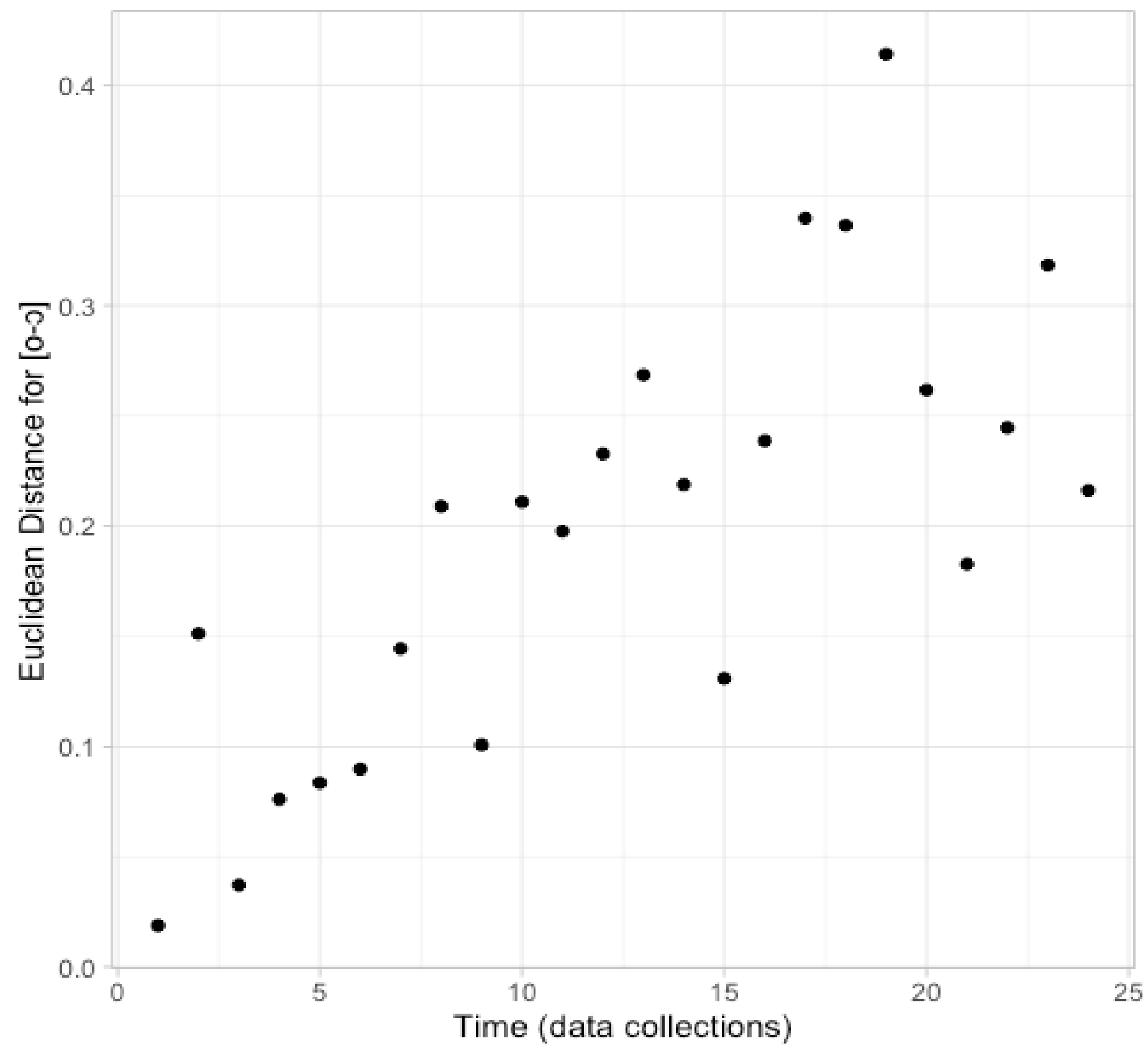
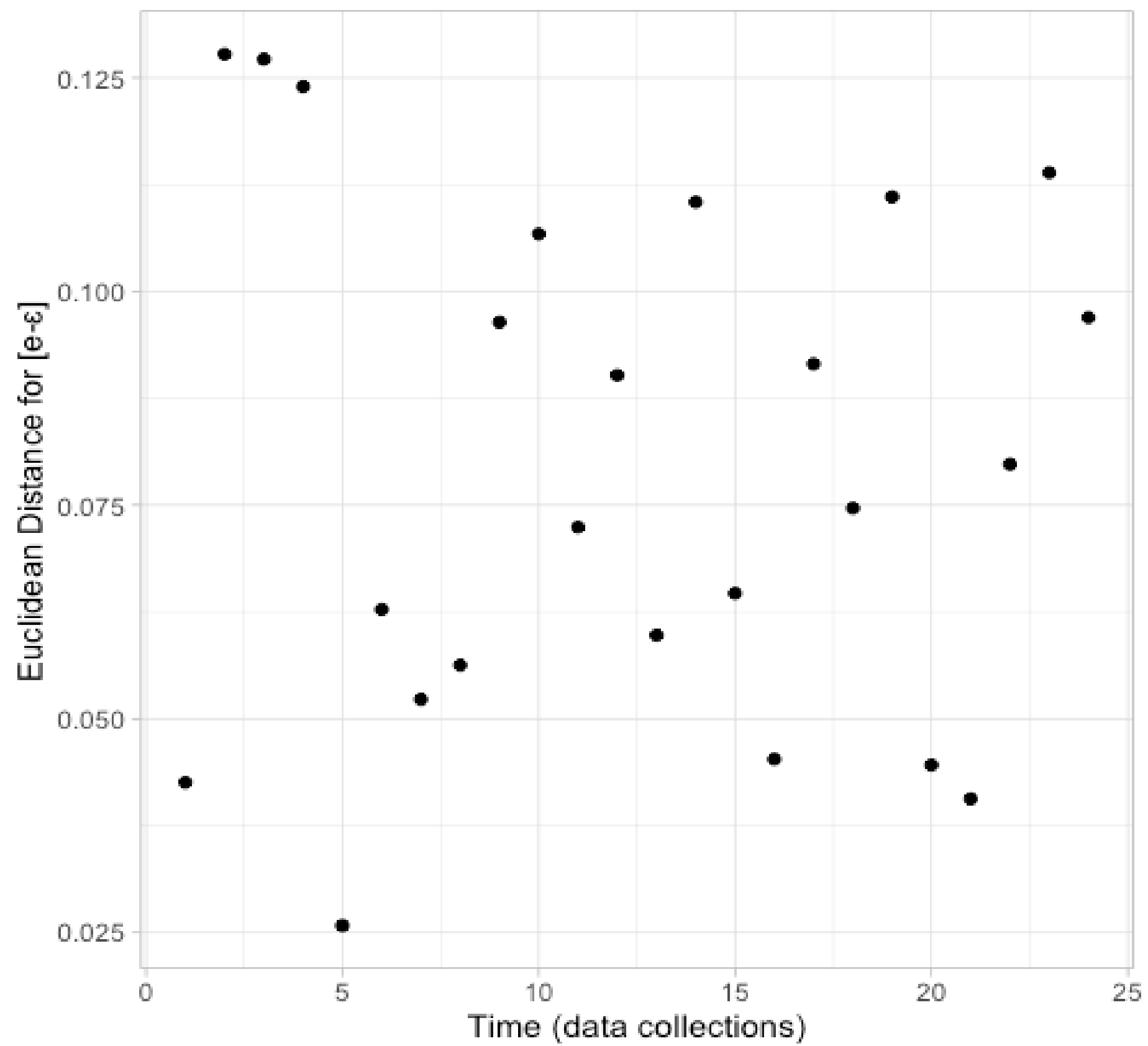
F1

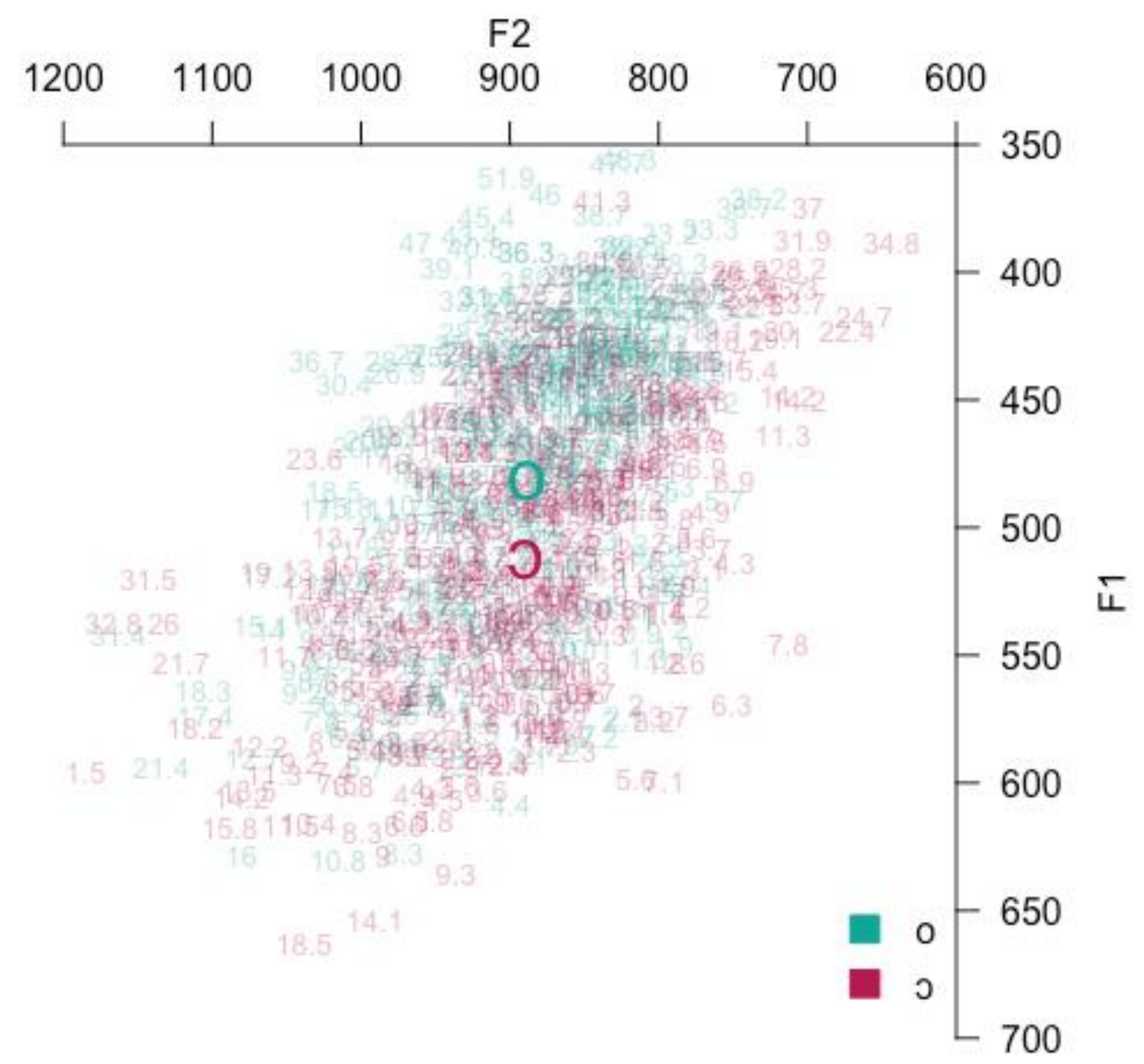
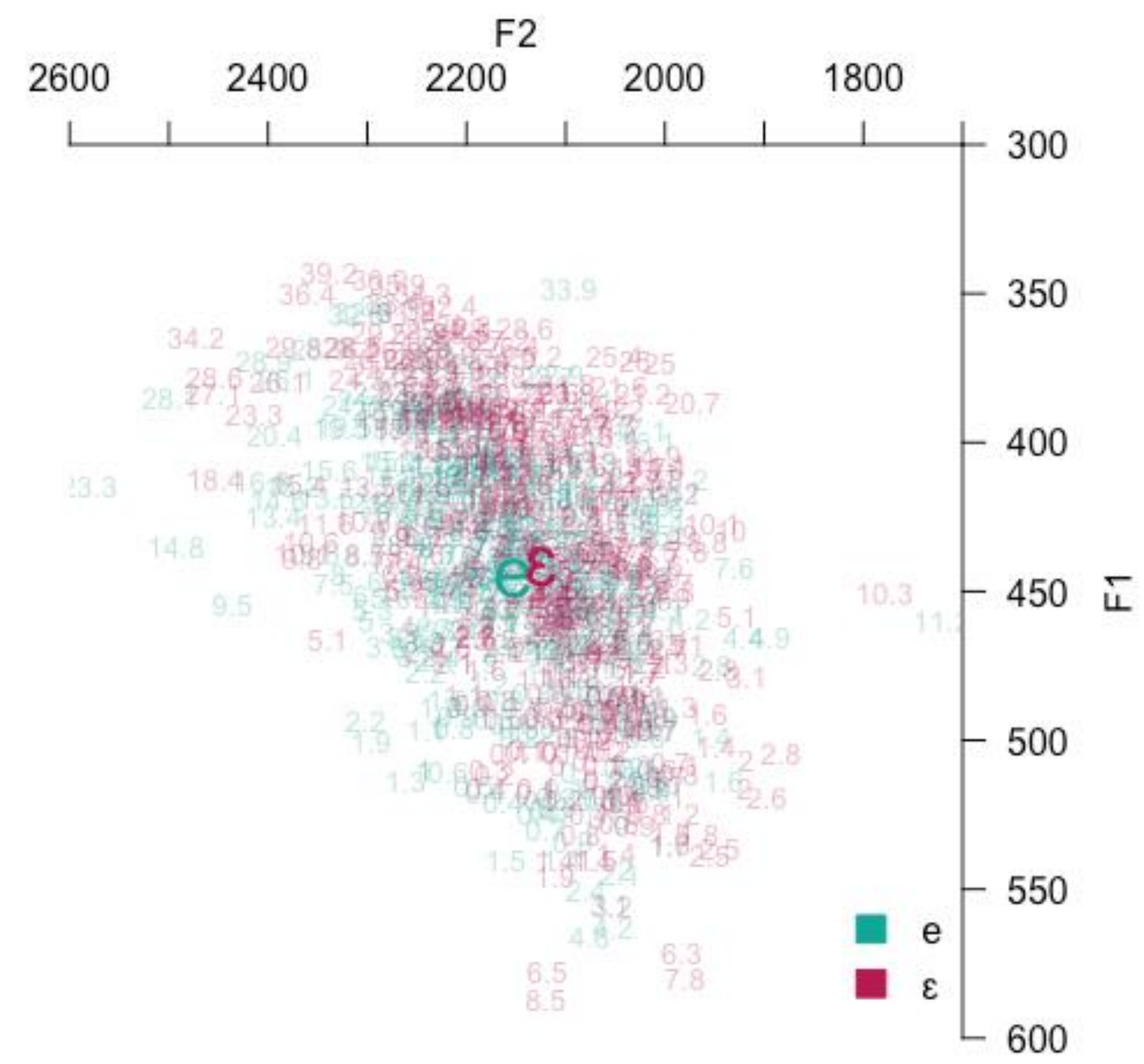
F1

Comparação (ii) Distância:

Euclidiana vs Mahalanobis



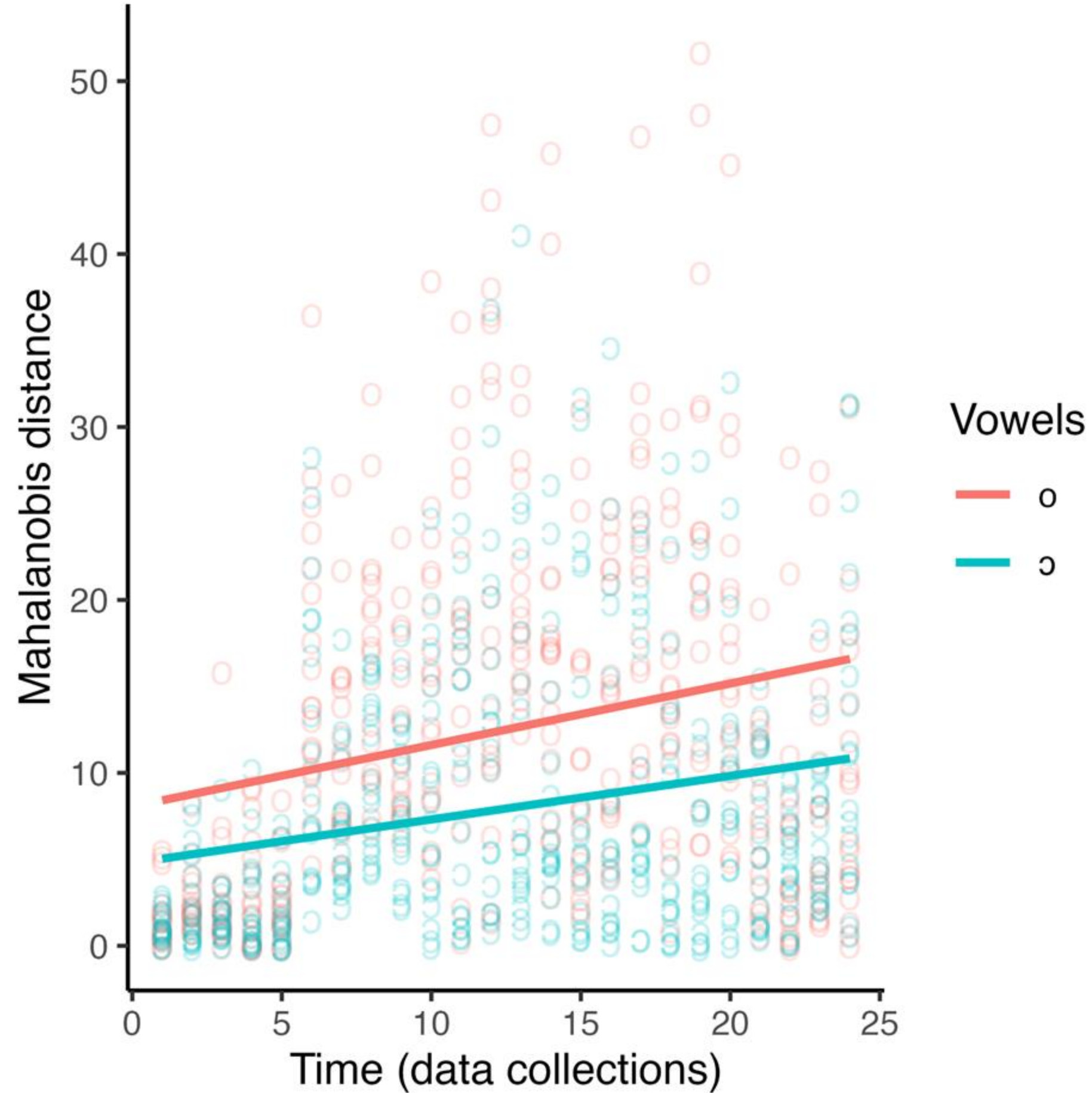
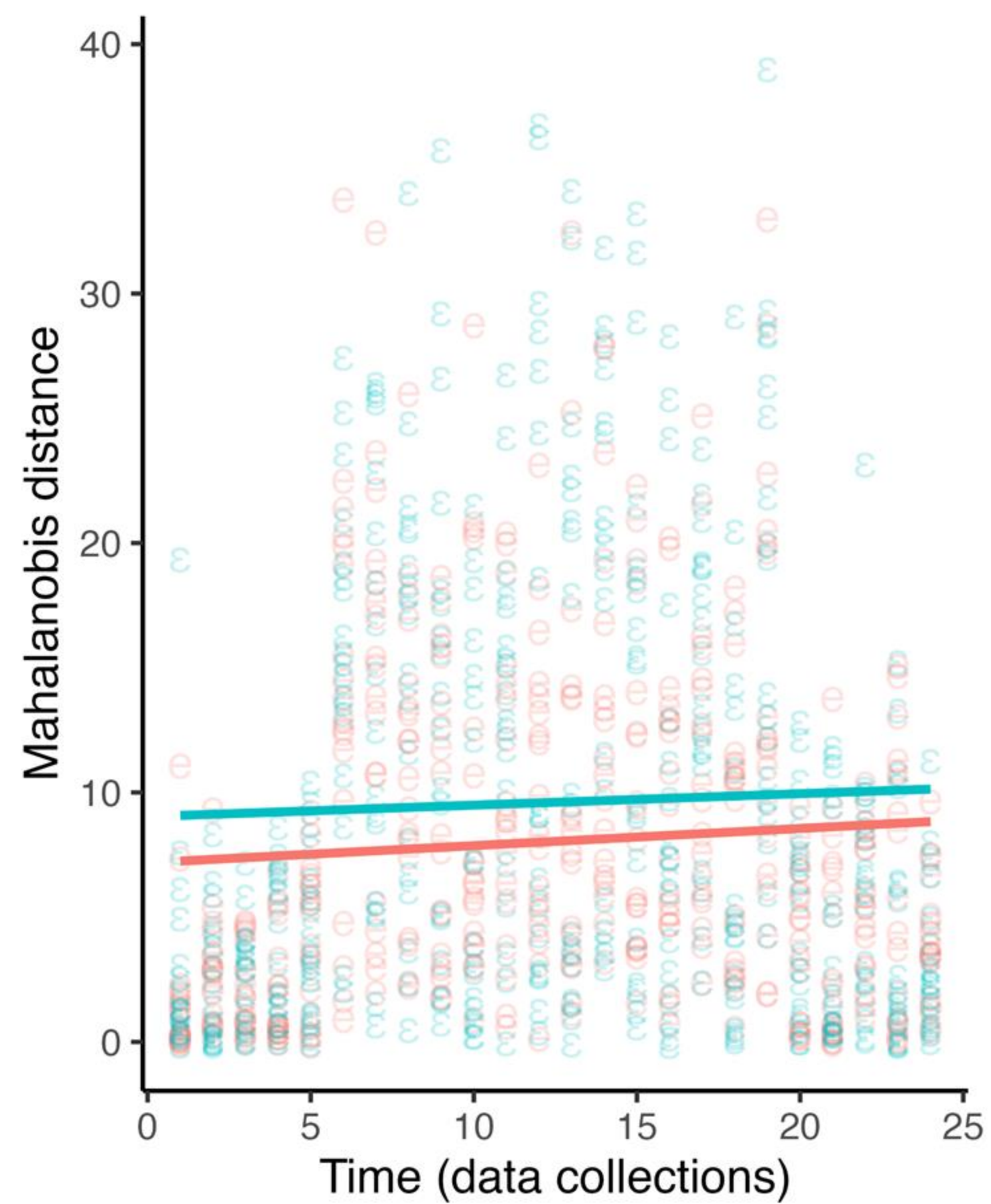


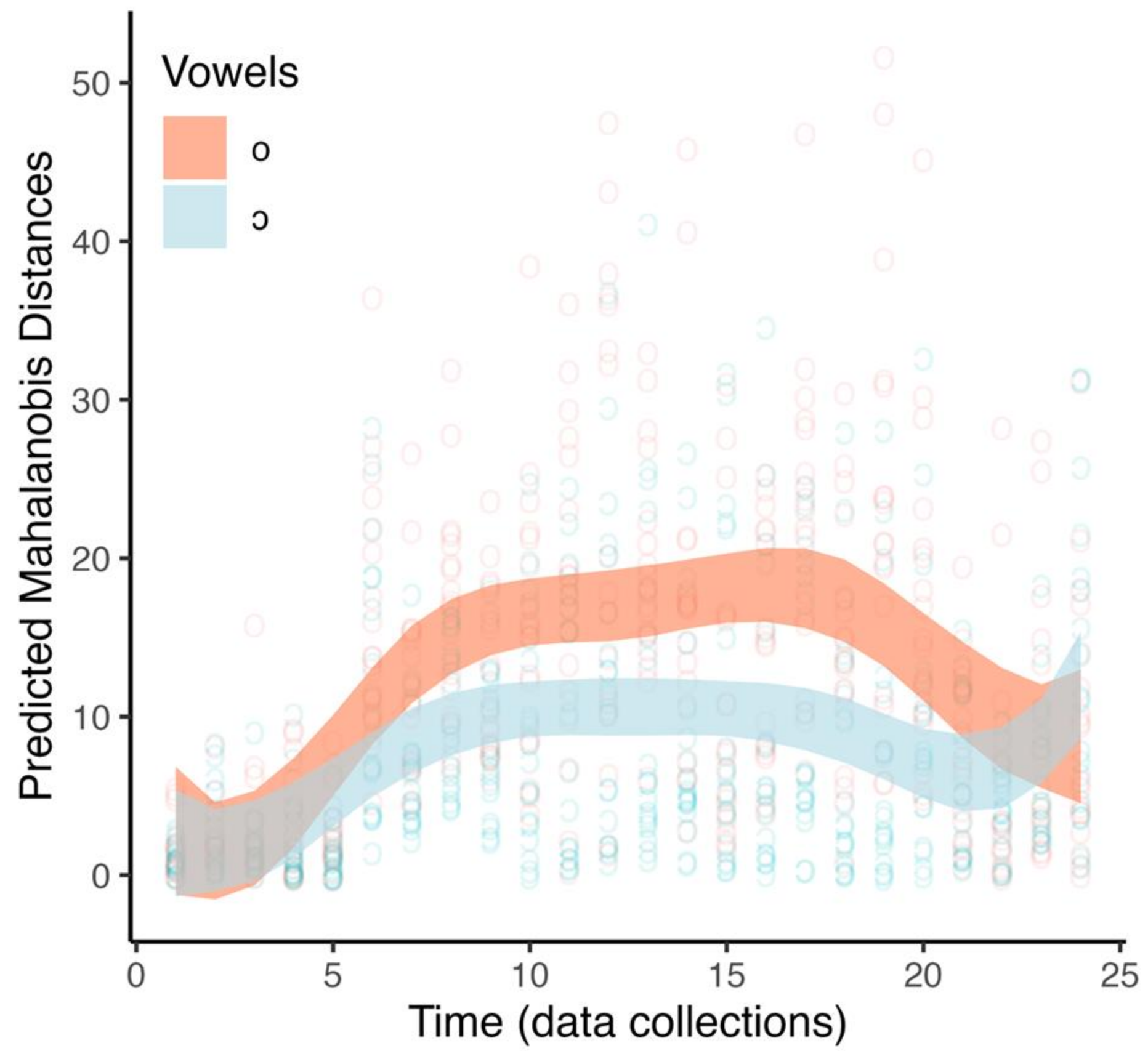
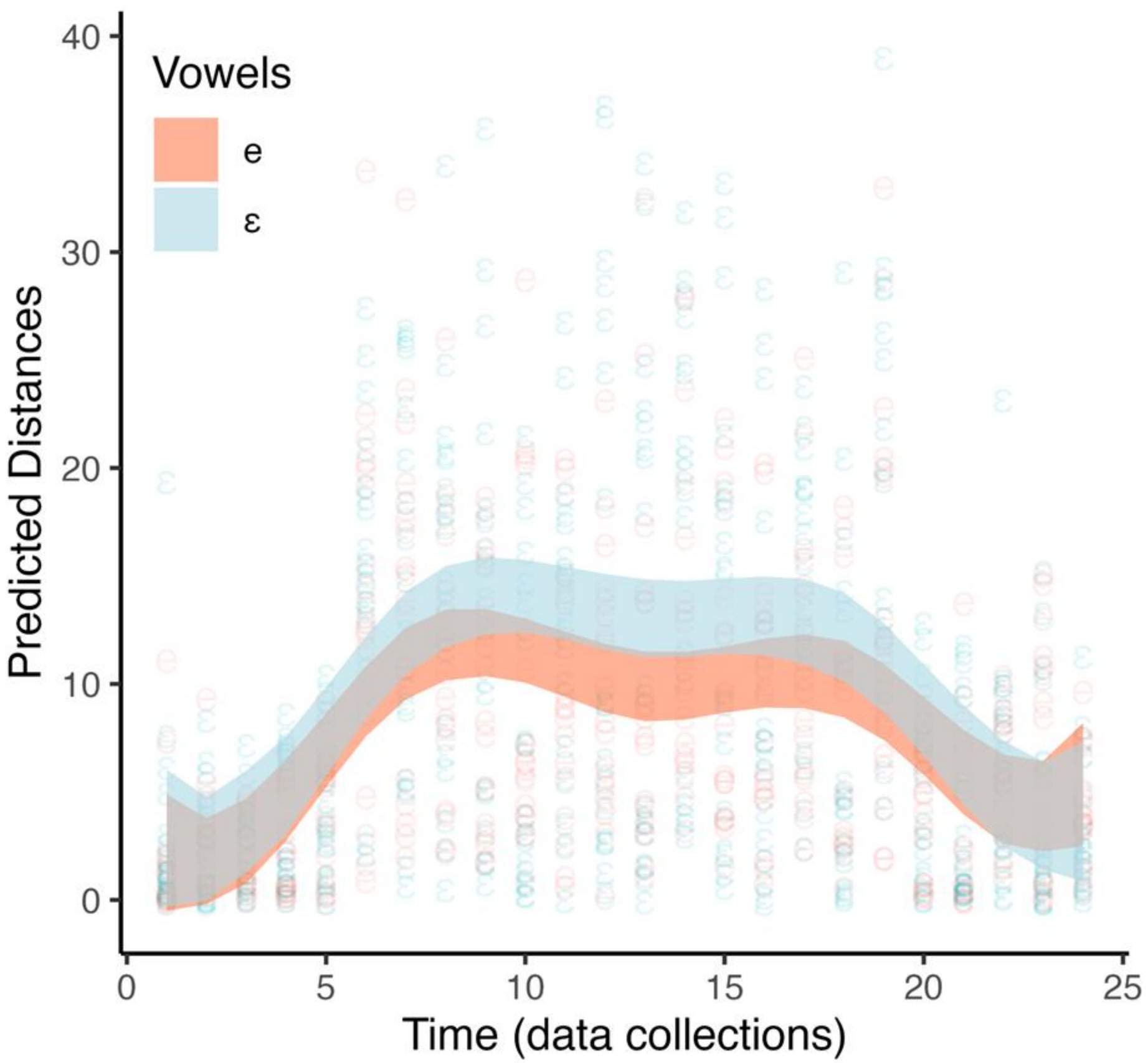


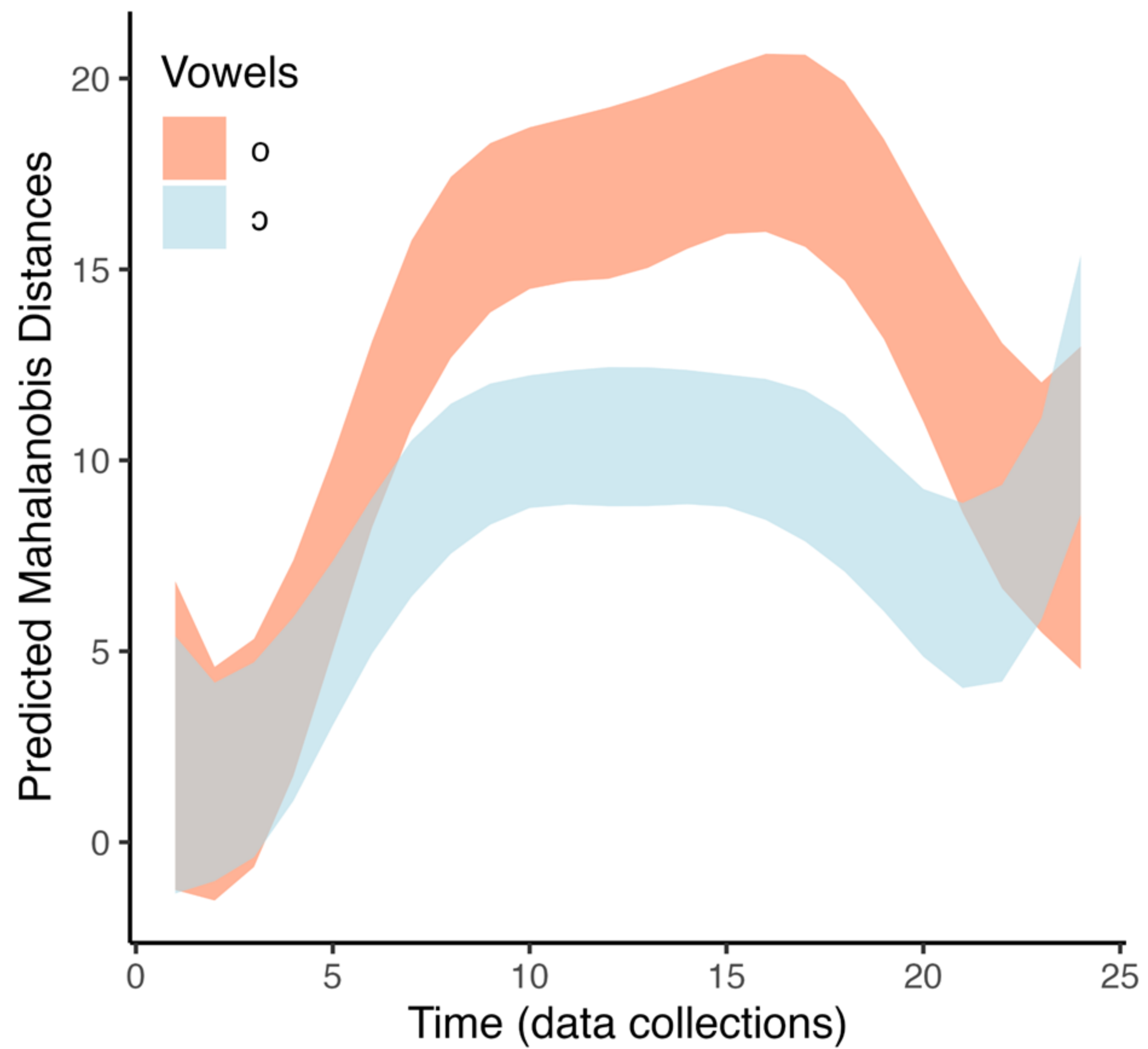
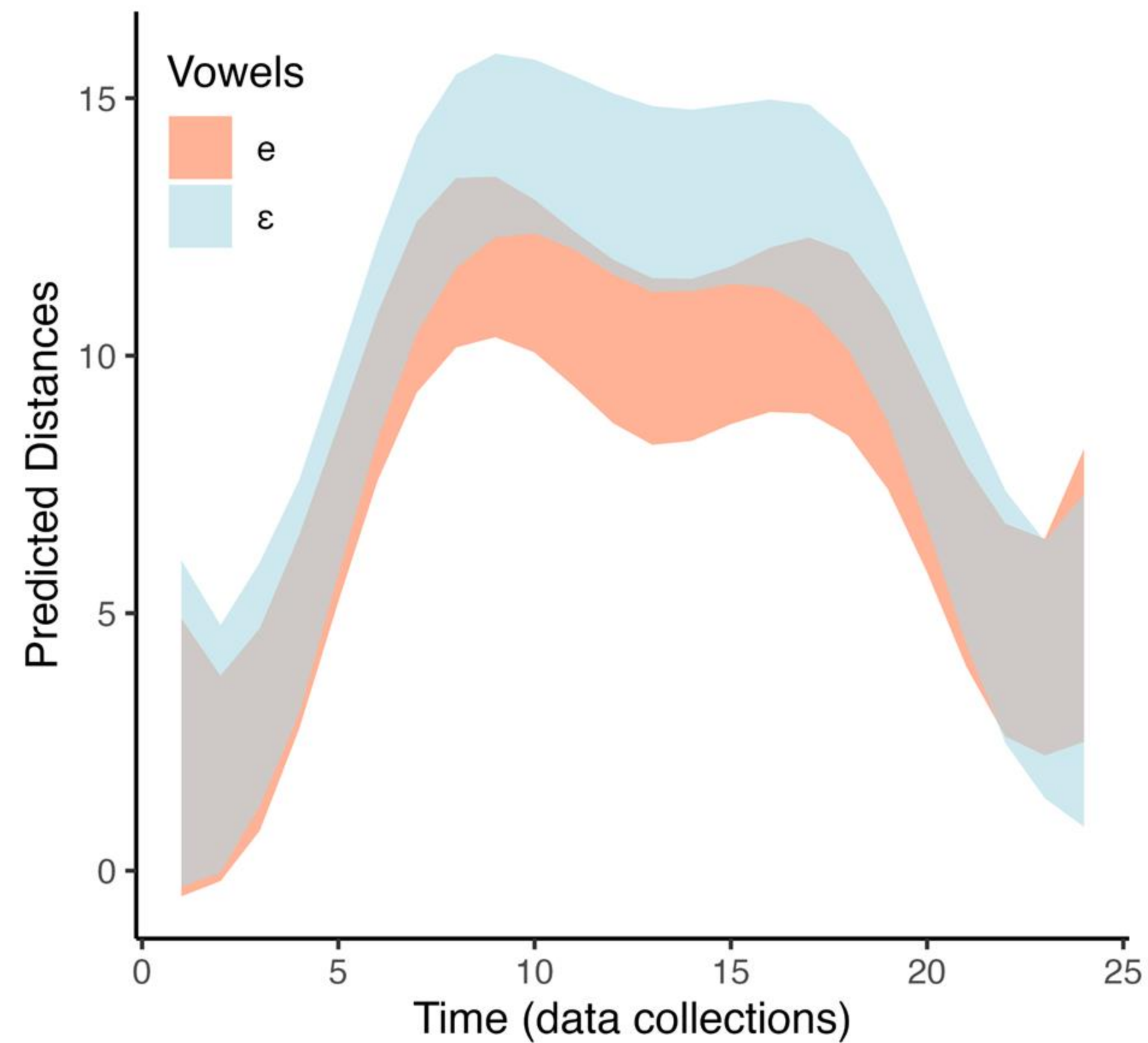
Comparação (iii) Modelo:

Linear vs GAMM Bayesiano









Discussão



Modelo Aditivo Generalizado Misto Bayesiano

Modela a probabilidade dos parâmetros a partir dos dados, e não a probabilidade dos dados a partir de uma hipótese nula (sem valor de p)

Fornece distribuições de probabilidade referentes aos coeficientes, ao invés de *point estimates*

Fornece intervalos de credibilidade, os quais são mais intuitivos do que intervalos de confiança

Permite ao pesquisador incluir o conhecimento prévio referente às prováveis distribuições dos dados

Na tentativa de formação de novas categorias para /ε/ and /ɔ/, todas as vogais interagem entre si; GAMMS Bayesianos são capazes de dar conta dessas alterações.

Vantagens dos GAMs:

- todos os dados incluídos
- todas as vogais em um mesmo modelo
- previsão de linhas curvas

Vantagens dos GAMs Bayesianos:

- probabilidade dos parâmetros a partir dos dados
- densidades de probabilidade em vez de *point estimates*
- conhecimento prévio adicionável

Referências bibliográficas

Alves, U. K.; Brisolara, L. B.; Perozzo, R. V. (2017). Curtindo os sons do Brasil: fonética do Português do Brasil para hispanofalantes. Lisboa: LIDEL Editorial.

Boersma, P.; Weenink, D. (2019) Praat: Doing phonetics by computer – Version 6.1.16. <https://www.fon.hum.uva.nl/praat/>.

Pereyron, L. (2017). A produção vocálica por falantes de Espanhol (L1), Inglês (L2) e Português (L3): uma perspectiva dinâmica na (multi) direcionalidade da transferência linguística. Unpublished Doctoral Dissertation. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2017.

Taylor, W. A. (2000). Change-Point Analysis: a powerful new tool for detecting changes. Retrieved from <https://variation.com/change-point-analysis-a-powerful-new-tool-for-detecting-changes/>. Access on September 6, 2021.

Van DIJK, M.; Verspoor, M.; Lowie, W. (2011). Variability and DST. *In: VERSPOOR, M.; de BOT, K.; LOWIE, W. (Eds.), A Dynamic Approach to Second Language Development: methods and techniques*, p. 55-84. Amsterdam: John Benjamins Publishing Company.

Wood, S (2017). Generalized Additive Models: An Introduction with R, 2 edition. Chapman and Hall/CRC.

Obrigado!

ronaldo.junior@unb.br
(@ronaldolimajunior)

ukalves@gmail.com
(@bira_alves_prof)

Slides:

