Discussion of "The statistical analysis of acoustic phonetic data: exploring differences between spoken Romance languages" by D. Pigoli, P.Z. Hadjipantelis, J.S. Coleman and J.A.D. Aston.

J.P. Meagher *[1], T. Damoulas[1], K. Jones[2], and M.A. Girolami[3]

[1]The University of Warwick
[2]Centre For Biodiversity and Environment Research, UCL
[3]Imperial College London

18 April, 2018

We offer our congratulations to the authors on this excellent contribution. The application of functional data analysis to a time-frequency representation of sound presents many exciting research directions in applied linguistics and beyond. A central motivation of the paper is to model the historical development of language through the statistical analysis of present day speech recordings. This process is presented by the authors as being analogous to biological evolution, and so we consider how phylogenetic Gaussian processes [2] [1] could be applied to these smooth spectrogram surfaces for evolutionary inference.

Spectrogram surface over a phylogeny

For each language L, consider the sample of smoothed spectrogram surfaces representing the digit 1. Denote the language phylogeny, representing the evolutionary relationships between languages, as $\mathcal{P}$. Each language corresponds to a position on the phylogeny and so denote $S_{1k}(\omega, t, \mathbf{p}) \equiv S^L_{1k}(\omega, t)$, for $\mathbf{p} \in \mathcal{P}$.

The spectrogram surfaces can now be thought of as function valued traits and modelled as phylogenetic Gaussian processes. This means that, subject to the simplifying assumptions detailed by [2] we can write

$$S_{1k}(\omega, t, \mathbf{p}) = \mu(\omega, t) + \sum_{q=1}^{Q} \phi_q(\omega, t) f_q(\mathbf{p}) + \phi_q(\omega, t) z_q$$

where $\mu(\omega, t)$ is the mean spectrogram surface, $f_q(\mathbf{p}) \sim \mathcal{GP}(0, k_q(\mathbf{p}, \mathbf{p}'))$, $z_q \sim \mathcal{N}(0, \sigma_q)$, and ϕ_q is the q^{th} basis function.

Assume an Ornstein-Uhlenbeck process prior for $f_q(\mathbf{p})$,

$$k_q(\mathbf{p}, \mathbf{p}') = \gamma_q \exp\left(-\frac{d_{\mathcal{P}}(\mathbf{p}, \mathbf{p}')}{\ell_q}\right)$$

where $d_{\mathcal{P}}(\mathbf{p}, \mathbf{p}')$ is the distance between points $\mathbf{p}$ and $\mathbf{p}'$ over $\mathcal{P}$

Given the set of hyperparameters $\theta_q = (\gamma_q, \ell_q, \sigma_q), q = 1, \ldots, Q$, we have a posterior distribution for the smooth spectrogram representation of the digit 1 for any $\mathbf{p} \in \mathcal{P}$

Application to bat echolocation

The model outlined above maps directly to a very different research question, that being, what did ancestral bat echolocation calls sound like? The model is illustrated in figure 1 where the analysis in [3] is repeated for smooth spectrogram surface representations of echolocation calls.

*Corresponding Author: J.Meagher@warwick.ac.uk

Looking ahead

In reconstructing ancestral acoustic phonetic states for language, it is not immediately clear what the appropriate unit for reconstruction should be. One approach would be to perform reconstructions on a word by word basis as outlined above, however this would seem a haphazard approach. Perhaps ancestral reconstruction of the covariance structure would be more informative. In any case, we look forward to seeing how this work develops and informs research in linguistics, evolutionary biology, or any domain where acoustic phonetic data is of interest.

References

[1] Pantelis Z Hadjipantelis, Nick S Jones, John Moriarty, David A Springate, and Christopher G Knight. Function-valued traits in evolution. *Journal of The Royal Society Interface*, 10(82):20121032, 2013.

[2] Nick S Jones and John Moriarty. Evolutionary inference for function-valued traits: Gaussian process regression on phylogenies. *Journal of The Royal Society Interface*, 10(78):20120616, 2013.

[3] JP Meagher, T Damoulas, KE Jones, and M Girolami. Phylogenetic gaussian processes for bat echolocation. *Statistical Data Science*, page 111, 2018.

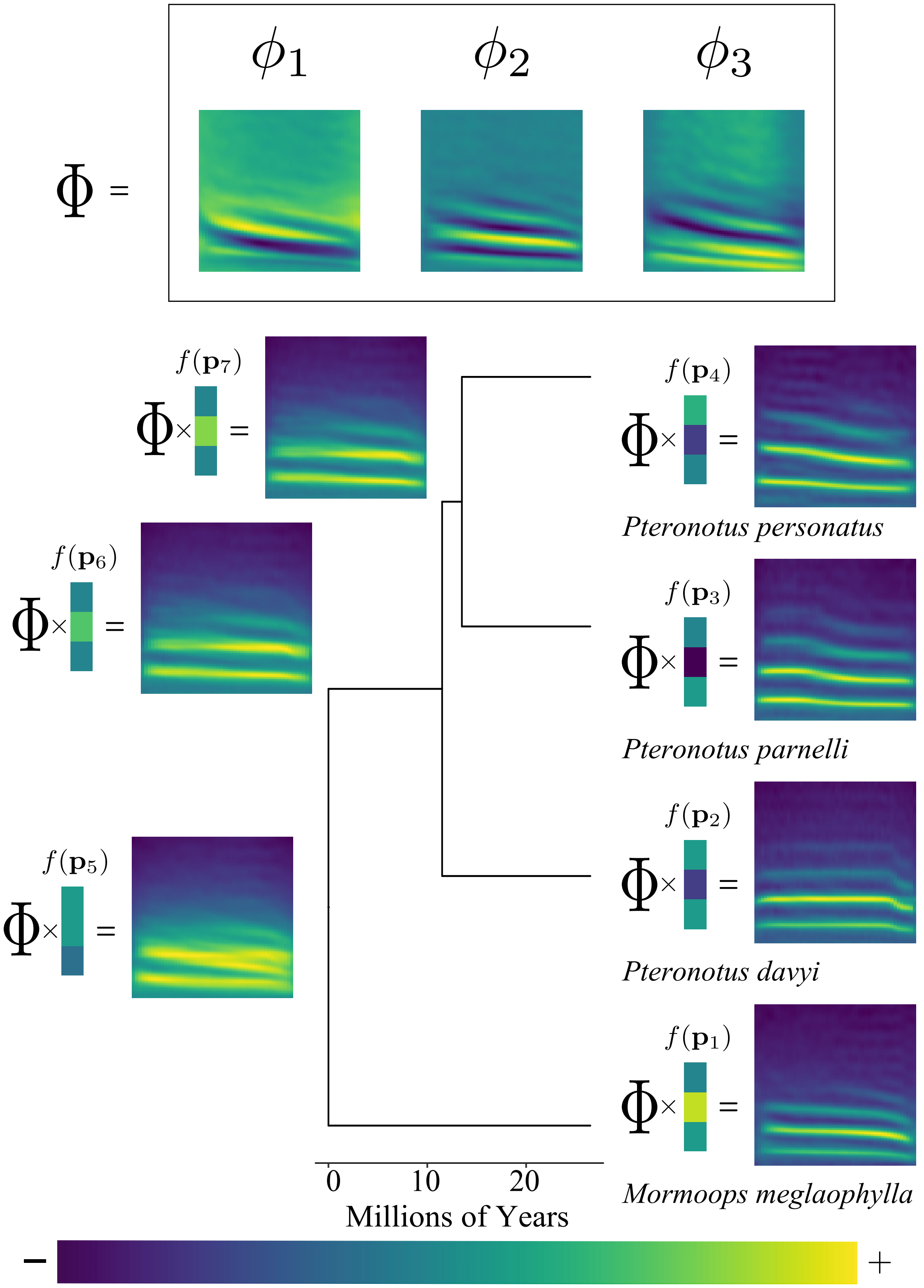

Figure 1: An ancestral reconstruction for the echolocation call spectrogram for 4 species of bat from the Molossidae family. The set of 'evolutionary basis', $\Phi = \{\phi_1, \phi_2, \phi_3\}$ have been inferred by a Principal Component Analysis of echolocation spectrogram surfaces. $f(\mathbf{p}_n)$ denotes the phylogenetic Gaussian process of basis weights for the species at point $\mathbf{p}_n$ on the phylogeny, where the weight of each basis is modeled as an independent, univariate Ornstein-Uhlenbeck process. Spectrogram surfaces at internal nodes are represent maximum-a-posteriori estimates of the ancestral echolocation call spectrogram.