Nikolai Vetr
University of California, Davis

nlashinsky@ucdavis.edu — https://github.com/NikVetr/

EDUCATION

PhD *2020*
University of California, Davis
Anthropology; Population Biology; Data Science & Informatics

BA *2013*
Vanderbilt University
Earth & Environmental Sciences (EES); Ecology, Evolution & Organismal Biology (EEOB)
Departmental Honors in EES, *summa cum laude* overall

ADDITIONAL EDUCATION

Discussion Groups *2013 - 2020*
Active participant in groups on Computational Molecular Evolution, Paleoanthropology, Quantitative Genetics, and Machine Learning. Founder and lead in groups on Python, Applied Linear Algebra, Bayesian Data Analysis, and Deep Learning

Workshops *2013 - 2020*
Have attended multiple workshops ranging in duration from 2h to 9d on Universal Design, Science Communication, Bayesian Phylogenetics, Database Management, Natural Language Processing, Carpentries Instructor Training, and other topics.

Online Courses *2013 - 2020*
In addition to reading textbooks through discussion groups or on my own, I've worked through several online courses, such as Gilbert Strang's *Linear Algebra* (MIT OpenCourseWare 18.06) or Andrew Ng's *Deep Learning Specialization* (Coursera)

LANGUAGES

Fluent: Russian, English
Proficient: Spanish, R, RevBayes, Stan, BASH
Beginner: Python, C++

PROJECTS

DISSERTATION WORK

My dissertation work has primarily centered on Bayesian computational phylogenetics methods development, specifically focused on extending and exploring multivariate Brownian (mvBM) diffusion models to novel applications and contexts. Projects in association with this work have involved:

- through simulation **exploring the informativeness and reliability of continuous character**

alignments at retrieval of simulating tree topology, with subsequent application to a craniometric dataset
- the **effects of model misspecification** in the above, both with respect to the evolutionary process and with respect to accommodating missing data and measurement error
- **comparing model-based phylogenetic methods to more common heuristic (maximum parsimony, distance) methods**, both in simulation and empirically (using ape and old-world monkey datasets), and contrasting their performance to methods that use molecular (e.g. nucleotide sequence) data
- **comparison of different relaxed clock models** in a species-rich fish dataset whose tree height straddles the K-Pg boundary
- **approximating truncated biogeographic diffusions using data augmentation** on phylogenies to sample migration histories while respecting (land, water) barriers and explore latitudinal / climactic drivers of morphological differentiation in modern humans
- **joint inference of ancestral character states** in a monkey dataset, with additional theory development on imputing unobserved tip characters in a phylogenetically sensible way
- layering a **high-dimensional probit model atop mvBM to accommodate discrete traits** and comparing it with more traditional CTMC (*Continuous Time Markov Chain*) models of discrete trait substitution, with empirical application to hominin dental variation

SIDE PROJECTS

Throughout my PhD, I've undertaken a number of side projects, some of which directly extended my dissertation work to non-evolutionary contexts and some of which used more distantly related methods. These have often been self-motivated, but also sometimes done in collaboration with or response to requests by third parties (e.g. as statistical co-author). These side-projects have included developing, fitting, interpreting, describing, and visualizing results from:

- high-dimensional multilevel generalized linear mixed models (GLMMs; Poisson, in this case) in Stan to **heart transplant patient immune response** data both before and after an acute rejection episode
- multilevel GLMMs in Stan to **consumer dietary and attitudinal response datasets** (zero-inflated Poisson and ordered logit, respectively) measuring the effects of exposure to a relevant advertisement urging dietary change
- a hierarchical univariate Ornstein-Uhlenbeck model in Stan to the 'evolution' of **nitrogen concentrations in manure ponds** across California to predict values after an arbitrary amount of time has passed in an arbitrary pond
- an efficient, conditional multivariate normal model in R that provides **personalized movie rating predictions** (given ratings for some movies and a large-but-sparse dataset of user reviews, what are the expected ratings and variances of unrated movies?) using probit and logit-normal transforms, exploiting basic properties of Schur complements

MINOR PROJECTS

I've written many dozens of minor scripts (< 100 LOC) devoted to data visualization, text mining, web scraping, replicating published analyses, and exploring off-the-cuff ideas proposed by myself or colleagues.

I've also consulted individuals in a minor role (e.g. over a $\sim$ 1h meeting) on many additional projects throughout graduate school.

Field & Labwork

Archaeological / Paleontological Excavator   *2013*
During the Summer 2013 field season I helped in excavation efforts ongoing at La Ferrassie, a Neandertal archaeological and paleontological site in Savignac-de-Miremont, France

Water Quality Analyst   *2011*
I performed water quality analyses (of salinity, reactive nitrogen concentrations, etc.) of a large set of streams distributed across the North Island of New Zealand

Stable Isotope Ecologist   *2012 - 2013*
As an undergraduate researcher, I examined how environmental and ecological conditions were recorded in enamel stable isotope ratios (carbon and oxygen) and dental microwear textures for several marsupial taxa throughout Australasia

Teaching

Associate Instructor*, University of California, Davis   *2015 - 2020*
As an associate instructor, I've taught three quarters of an upper-division paleoanthropology course, two quarters of an upper-division evolutionary primatology course, and one quarter of an introductory course on human evolution. In these roles, I created or modified all the lab and lecture materials, designed and graded the assignments and tests, delivered lectures, and either supervised teaching assistants or administered and oversaw the labs.

Teaching Assistant*, University of California, Davis & Vanderbilt University   *2012 - 2020*
I have taught a total of 14 discussion sections for the aforementioned human evolution and paleoanthropology courses (see above), as well as the lab for an introductory cell biology course at Vanderbilt University. As a TA, I occasionally gave guest lectures on various subjects (mostly statistical phylogenetics).

Outreach Lecturer, University of California, Davis   *2013 - 2020*
I've given multiple yearly talks on human evolution to elementary and middle school students on location and visiting UC-Davis, as well as during campus-wide events (e.g. *Picnic Day*).

Course Coordinator, Workshop in Applied Phylogenetics   *2019*
I served as a course coordinator for world-renowned, widely-attended, week-long workshop in applied computational Bayesian phylogenetics held in May 2019 at the Bodega Marine Laboratory.

Carpentries Instructor   *2019*
I have completed Instructor Training and Checkout for The Carpentries organization and am fully certified to teach courses in the Data and Software Carpentries.

**student evaluations available upon request*

Leadership

Founder & Lead, Applied Bayesian Statistics Research Cluster   *2019 - 2020*
I founded, coordinated, and supervised an interdisciplinary research cluster of 70+ scientists and statisticians across numerous career stages (PhD Student, Postdoc, PI, Industry Researcher), sponsored by the UC-Davis *Data Science & Informatics* Unit.

Coordinator, Various Reading Groups   *2014 - 2020*

I've founded and coordinated reading groups on Python, Applied Linear Algebra, Bayesian Data Analysis, and Deep Learning.

Undergraduate Student Mentor, UC-Davis   *2014 - 2020*
While serving as a college instructor, I've mentored several undergraduate students, meeting frequently to advise them on their career and research aspirations.

Publications

Papers

[1] **Vetr, N.** and Weaver, T. 2020. *Phylogenetic Inference under a Blocked Multivariate Threshold Model: Exploring Performance through Simulation and Application to Hominin Dentition.* In prep for PLOS ONE.

[2] **Vetr, N.** and Weaver, T. 2020. *Primate Phylogenetics with Landmark Data: A Model-Based, Multivariate Brownian Approach.* In prep for The Journal of Human Evolution.

[3] Gates, K., Biendarra-Tiegs, S., **Vetr, N.**, Higuita, M. Nelson, T., Pereira, N, and Griffiths, L. 2019. *Non-HLA antigens modulate heart transplant dysfunction.* In review by Circulation Research.

[4] **Vetr, N.**, May, M., Moore, B., and Weaver, T. 2019. *Bayesian Phylogenetic Inference Under Multivariate Brownian Motion.* Submitted to Systematic Biology.

[5] **Vetr, N.** 2017. *Non-Synapomorphic Traits are Phylogenetically Informative.* In prep.

[6] **Vetr, N.**, DeSantis, L., Yann, L., et al. 2013. *Is Rapoport's Rule a Recent Phenomenon? A Deep Time Perspective on Potential Causal Mechanisms.* Biology Letters 9(5): 1-5. DOI: 10.1098/rsbl.2013.0398

Conference Presentations

[1] **Lashinsky, N**. and DeSantis, L. 2012. *Assessing how the extant lesser forest-wallaby Dorcopsulus vanheurni and Lumholtz's tree kangaroo Dendrolagus lumholtzi record their environment: establishing a baseline for studies of extinct marsupials.* Poster presented at the 5th Annual Meeting of the Southeastern Association of Vertebrate Paleontology.

[2] Desantis, L., **Lashinsky, N**., Romer, J., Greshko, M., and Loffredo, L. 2013. *University Students' Acceptance of Climate Change and Evolution: Are Skeptics just Anti- Science?* Poster presented despite second authorship at the 73rd Annual Meeting of the Society of Vertebrate Paleontology.

[3] **Lashinsky, N**., DeSantis, L. Yann, S. Donohue, and R. Haupt. 2013 *Is Rapoport's Rule a Recent Phenomenon? A Deep Time Perspective on Potential Causal Mechanisms.* Oral session presented at the 73rd Annual Meeting of the Society of Vertebrate Paleontology.

[4] **Vetr, N**., May, M. Moore, B., and R. Weaver. 2018 *Adapting multivariate Brownian diffusion models to Bayesian inference of human population history and phylogeny.* Oral session presented at the 87th Annual Meeting of the American Association of Physical Anthropologists.

Grants & Awards

Crook Travel Award	*2018*
1st Place Picnic Day Exhibit Award in "Secrets of Nature" Category	*2017*
Outstanding Graduate Student Teaching Award Nomination	*2016, 2019*
Summer Research Grant	*2015*
NSF Graduate Research Fellowship	*2015*
Summer Research Grant	*2014*
Graduate Scholars Fellowship	*2013*
Departmental Honors in Earth and Environmental Sciences	*2013*
Eugene H. Vaughan Undergraduate Research Assistantship in Geology	*2012*
Geology Travel Grant	*2012*
Vanderbilt Undergraduate Summer Research Grant	*2012*
Ross Family Scholarship	*2012*
National Merit Scholarship	*2019*

Service

Journal Article Review

Science Communications (2018)
Evolution (2017)

Other

Computer Hardware
Familiar with computer assembly and repair. Constructed and networked an 8-node (64C/128T) Ryzen-based Beowulf cluster running macOS for dissertation work, with analyses parallelized using *GNU Parallel*, as well as a portable 5-node Raspberry Pi cluster running Linux. Have designed and assembled machines for colleagues after consultation on their computational needs.

Outdoors
Avid and experienced outdoorsman, having coordinated and lead many hiking, backpacking, and climbing trips of varying group size over distances typically between 50-150km, but ranging upwards to 1,300km, and have also backpacked and hitchhiked through several countries. Sporadically partake in other adventure sports (snowboarding, kayaking, cycling, etc.)

Photography
Hobbyist photographer, graphic designer, and illustrator focusing on portrait, landscape, and wildflife photography, with post-processing typically done in Capture One and Adobe Photoshop and recently augmented with more sophisticated computational tools (e.g. ANNs).

Charity
Active participant in the *Effective Altruism* movement since it's inception; donate regularly to meta-charity approved organizations, and also occasionally serve as a volunteer statistical consultant.