BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.

Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Dahl, David B.

eRA COMMONS USER NAME (credential, e.g., agency login): DavidDahl

POSITION TITLE: Professor and Chair of Statistics

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Brigham Young University	B.S.	04/1997	Statistics
Brigham Young University	M.S.	08/1998	Statistics
University of Wisconsin, Madison	Ph.D.	08/2004	Statistics / Biostatistics

A. Personal Statement

The proposed project requires an interdisciplinary team of individuals who are both experts in their field and capable of working productively with those in other disciplines. I have a strong research record in Bayesian nonparametric statistics and statistical computing and their application to the life sciences. I have a history of successful collaborations with nonstatisticians. Starting with my Ph.D. dissertation at the University of Wisconsin-Madison, my statistical research has been solidly rooted in applications for biology, genetics, and biochemistry. Since then, I have collaborated in a variety of interdisciplinary teams. Of note is a long-term collaboration with Dr. Jerry Tsai (biochemist) and Dr. Marina Vannucci (statistician). I was PI on an NIH R01 grant in which we worked intensively on rigorous statistical approaches for protein structure prediction. Together we identify opportunities to further the field and formulate statistical models to address these challenges. I have applied my skills in statistical computing to lead the efforts to efficiently implement statistical methods. My previous experience and expertise make me well-suited to conduct the interdisciplinary research that we propose in this grant application.

B. Positions, Scientific Appointments, and Honors

Positions and Scientific Appointments
2004-2010   Assistant Professor, Department of Statistics, Texas A&M University, College Station, TX
2007-2010   Adjunct Assistant Professor, Department of Biostatistics, University of Texas
MD Anderson Cancer Center, Houston, TX
2010-2012   Adjunct Associate Professor, Department of Biostatistics, University of Texas
MD Anderson Cancer Center, Houston, TX
2010-2012   Associate Professor, Department of Statistics, Texas A&M University, College Station, TX
2012-2015   Associate Professor, Department of Statistics, Brigham Young University, Provo, UT
2015-present Professor, Department of Statistics, Brigham Young University, Provo, UT
2022-present Chair, Department of Statistics, Brigham Young University, Provo, UT

Other Experience and Professional Memberships
1996-present Member, American Statistical Association
2003-present Member, Institute of Mathematical Statistics
2003-present Member, International Society for Bayesian Analysis

2009-2012   Associate Editor, Bayesian Analysis
2013-2015   Co-Editor, Bayesian Analysis
2021-2023   Member, Board of Directors, International Society for Bayesian Analysis

Honors
1997   Ellis R. Ott Scholarship from the American Society for Quality
1999-2004   National Eye Institute Traineeship in Biostatistics at UW-Madison

C. Contributions to Science

1. One of my major research areas in statistics is <u>Bayesian modeling of complex structures, such as random partitions and feature allocations</u>. These models permit the borrowing of information across subjects to improve prediction and interpretation and to allow inference in situations that are otherwise too sparse. A few of my key publications in this area are:
   a. **D. B. Dahl**, R. Day, J. Tsai (2017), Random Partition Distribution Indexed by Pairwise Information, *Journal of the American Statistical Association*, 112, 721-732. DOI:10.1080/01621459.2016.1165103.
   b. G. L. Page, F. A. Quintana, **D. B. Dahl** (2022), Dependent Modeling of Temporal Sequences of Random Partitions, *Journal of Computational and Graphical Statistics*, 31(2), 614-627. DOI:10.1080/10618600.2021.1987255.
2. Typically, in Bayesian data analysis, a great deal effort is spent on "fitting the model" such as sampling from the posterior distribution of the model (like those cited previously). It is also necessary, however, to <u>summarize the posterior distribution</u> to convey meaningful results. Another area of expertise is the estimation of parameters, partitions, and feature allocations based on posterior samples. A few of my key publications in this area are:
   a. **D. B. Dahl**, M. A. Newton (2007), Multiple Hypothesis Testing by Clustering Treatment Effects, *Journal of the American Statistical Association*, 102, 517-526.
   b. **D. B. Dahl**, D. J. Johnson, P. Müller (2022), Search Algorithms and Loss Functions for Bayesian Clustering, *Journal of Computational and Graphical Statistics*, accepted. DOI:10.1080/10618600.2022.2069779.
3. I have also made contributions to the field of statistical models for protein structure prediction. I was PI on an NIH R01 grant with Dr. Jerry Tsai (biochemist) and Dr. Marina Vannucci (statistician) in which built rigorous statistical approaches for protein structure prediction. Many publications are with my Ph.D. student Kristin Lennox. This collaboration led to 10 publications, the most significant from a statistical perspective are:
   a. K. P. Lennox, **D. B. Dahl**, M. Vannucci, J. W. Tsai (2009), Density Estimation for Protein Conformation Angles Using a Bivariate von Mises Distribution and Bayesian Nonparametrics, *Journal of the American Statistical Association*, 104, 586-596.
   b. K. P. Lennox, **D. B. Dahl**, M. Vannucci, R. Day, J. W. Tsai (2010), A Dirichlet Process Mixture of Hidden Markov Models for Protein Structure Prediction, *Annals of Applied Statistics*, 4, 916-942.
   c. Q. Li, **D. B. Dahl**, M. Vannucci, H. Joo, J. W. Tsai (2016), KScons: A Bayesian Approach for Protein Residue Contact Prediction using the Knob-socket Model of Protein Tertiary Structure, *Bioinformatics*, 32(24): 3774-3781.
4. I also work in the area of efficient algorithms to fit Bayesian nonparametric models. Most notable are my merge-split samplers for random partition models, including these articles:
   a. **D. B. Dahl**, S. Newcomb (2022), Sequentially-Allocated Merge-Split Samplers for Conjugate Bayesian Nonparametric Models, *Journal of Computational Statistics and Simulation*, 92(7), 1487-1511. DOI:10.1080/00949655.2021.
   b. **D. B. Dahl** (2007), Invited Discussion of Jain and Neal's "Splitting and Merging Components of a Nonconjugate Dirichlet Process Mixture Model," *Bayesian Analysis*, 2, 473-478.
5. Nearly all of my papers provide a reference software implementation. To aid other researchers in also providing software for their methods, I have worked in the area of statistical software integrations. For example, I have a paper under review that provides a framework for developing R packages using fast and safe Rust code.