### SELECTED WORK EXPERIENCE

### Kite Pharma | Associate Data Scientist

Skills: Python, R, SQL, Git, Databricks, NLP, unsupervised learning, generative AI, XGBoost

- · Performing spatial and hierarchical unsupervised clustering of textual Kite Quality Control deviations data to inform Quality leadership on accuracy and consistency of human deviation labeling and identify higher quality documents ideal for deviation classification model fine-tuning use.
- Developed scheme for the trending of commercial invalid assay counts from Kite Laboratory Information System data with approximate and simulated contingency table statistical tests; execution of this scheme identified previously unknown data input errors in Kite's data warehouse, which were corrected by Kite Quality Control.
- Assisted Kite Global Supply Chain Packaging Team in designing and executing a study on label application characterization using contingency table tests of independence; study led to CAPA which recommended the increase of label cure time to  $\geq 2$  hours and label application to non-chilled cassettes for label adhesion performance improvement.
- Led an award-winning team in the 2024 Gilead Data Science Hackathon, in which participants were tasked with binary prediction
  of HIV status in individuals based on insurance claims and demographic data; implemented an ensemble model approach combining deep neural network, boosted decision tree, and Shapley feature selection methods to make explainable and interpretable
  predictions that resulted in a top three Creativity Award and People's Choice Award finish.

### Kite Pharma | Quality Engineer III - Statistician

Skills: Python, R, Git, AWS, Veeva Vault, GMP regulations, frequentist statistics

- Produced technical reports that informed Quality leadership decisions regarding product manufacturing and assay development; results communicated included statistical unfeasibility of universal Jurkat cell control limits and numerical evidence of positive approved treatment center performance.
- · Authored global Quality policy documentation guiding GMP data governance, trending, modeling, and analysis pursuant to agency guidance and regulations.
- · Quality participant of Gilead's Project GAIA initiative to evaluate large language models (LLMs); compared AWS and Azure LLM frameworks for Gilead use and proposed governance measures in anticipation of future agency regulations.

### University of California | Graduate Student Researcher

Skills: Python, R, Stan, Julia, Bash, Git, Bayesian statistics, mathematical modeling, deep learning Jun. 2016 - Dec. 2022

- · Developed cutting edge Bayesian deep learning variational inference algorithms for model fitting and parameter estimation.
- · Formulated Bayesian statistical frameworks rigorously assessing the ability of soil carbon models to predict global soil changes.
- · Parameterized and refined stochastic differential equation soil carbon models; resulted in more realistic soil system simulations.
- · Amalgamated and cleaned data from separate soil experiments to construct a meta-analysis data product ready for statistical and machine learning insight extraction.

# National Institute of Diabetes and Digestive and Kidney Diseases | Intramural Research FellowBethesda, MDSkills: Python, Julia, C, R, Stan, Mathematica, Bash, Bayesian statistics, mathematical modelingOct. 2013 - Jul. 2015

- · Performed computational neuroscience research investigating divisive normalization, a canonical neural phenomenon observed across multicellular organisms including humans.
- · Constructed and simulated biophysically realistic neuron population models using a variety of scientific computing tools.
- · Developed parallel, distributed Bayesian Markov chain Monte Carlo algorithms assimilating neuron models with empirical data.

## **RESEARCH PUBLICATIONS AND MANUSCRIPTS**

**Hua W. Xie**, Debora Sujono, Tom Ryder, Erik B. Sudderth, and Steven D. Allison. (2024). "A framework for variational inference and data assimilation of soil biogeochemical models using state space approximations and normalizing flows." *ESS Open Archive*. Submitted and pending publication.

**Hua W. Xie**, Steven D. Allison, Elizabeth Duan, and Brian Chung. (2024). "Chapter 16 - Advancing quantitative models of soil microbiology, ecology, and biochemistry." In Eldor A. Paul and Serita Frey (Eds.), *Soil Microbiology, Ecology and Biochemistry, Fifth Edition* (pp. 473-492), ed. Eldor A. Paul and Serita Frey. Elsevier.

Debora Sujono, Hua W. Xie, Steven D. Allison, and Erik B. Sudderth. (2022). "Variational Inference for Soil Biogeochemical Models." *ICML 2022 2nd AI for Science Workshop*.

Hua W. Xie, Adriana L. Romero-Olivares, Michele Guindani, and Steven D. Allison. (2020). "A Bayesian approach to evaluation of soil biogeochemical models." *Biogeosciences*. 17:4043 - 4057.

Santa Monica, CA Jan. 2024 - current

Santa Monica. CA

Jan. 2023 - Dec. 2023

Irvine, CA

Shashaank Vattikuti, Phyllis Thangaraj, Hua W. Xie, Stephen J. Gotts, Alex Martin, and Carson C. Chow. (2016). "Canonical cortical circuit model explains rivalry, intermittent rivalry, and rivalry memory." *PLoS Comput Bio.* 12:e1004903.

## SKILLS AND COMPETENCIES

- Regulatory Competencies: 21 CFR Parts 11, 210, 211, and 820; ICH Q8 (R2), Q9 (R1), and Q10
- Technical Skills: Python (PyTorch, Pandas, Scipy, Numpy), R (Tidyverse), Cloud Data Platforms (AWS, Databricks), SQL

## **EDUCATION**

University of California, Irvine Ph.D. in Mathematical, Computational, and Systems Biology Northwestern University, Evanston B.S. in Biological Sciences and Radio/Television/Film

Nov. 2022

Jun. 2012