Katherine **Goode**

Research and Development Statistician

Sandia National Laboratories, P.O. Box 5800 MS 0829, Albuquerque, NM 87185

+1 505-844-1998 | kjgoode@sandia.gov | goodekat.github.io | github.com/goodekat

Education

Iowa State University — *Ames, IA*

Doctor of Philosophy, Statistics — *2016-2021*

- Dissertation: Visual Diagnostics for Explaining Machine Learning Models
- Major Professor: Dr. Heike Hofmann

University of Wisconsin, Madison — *Madison, WI*

Master of Science, Statistics — *2013-2015*

-

Lawrence University — *Appleton, WI*

Bachelor of Arts, Mathematics — *2009-2013*

- Graduated Magna Cum Laude
- Senior Capstone: An Explanation of Double-Error-Correcting BCH Codes

Awards

Midwest Statistical Machine Learning Colloquium Poster Award

Awarded for poster Using LIME to Interpret a Random Forest Model with an Application to Bullet Matching Data — *2019*

ISU Department of Statistics Dan Mowrey Consulting Excellence Award

Awarded in recognition of outstanding contributions in the area of statistical consulting while working toward a graduate degree. — *2018*

ISU Department of Statistics Award for Experiential Development

Presented to a graduate student for excellent performance in multiple statistical efforts (teaching and consulting) as part of the graduate program. — *2017*

Experience

Postdoctoral Researcher — *Sandia National Laboratories*

Statistical Sciences Department — *Sep 2021 - Current*

- Performed research on inverse models with functional data
- Implemented shape analysis methods with national security data
- Prepared manuscript for submission to peer reviewed journal

Research and Development Intern — *Sandia National Laboratories*

Statistical Sciences Department — *Dec 2019 - Sep 2021*

- Performed research on neural networks explainability with functional data
- Applied explainability methods to machine learning models
- Presented on work at internal and external events

Graduate Research Assistant — *Iowa State University*

Natural Resource Ecology and Management — *Jan 2021 - June 2021*

- Developed R Shiny application to predict taxonomy of fish eggs using random forests
- Assisted in writing manuscript to present the application
- Advised by Dr. Michael Weber and Dr. Philip Dixon

Statistical Consultant — *Iowa State University*

Agriculture Experiment Station — *May 2016 - Dec 2020*

- Senior consultant from May 2018 to May 2020
- Helped with administrative decisions and trained new consultants
- Provided statistical support on research projects for graduate students, professors, and staff from the colleges of agriculture and life sciences, engineering, human sciences, liberal arts and sciences, and veterinary medicine
- Assisted with the implementation of analyses in R, SAS, JMP, and SPSS

Graduate Research Assistant — *Iowa State University*

Department Natural Resource Ecology and Management — *May 2019 - Aug 2019*

- Assisted with analysis of toxicology study of monarch butterfly larvae exposed to insecticides
- Wrote R code to compute profile confidence intervals for dose response curve models
- Collaboration with Dr. Steven Bradbury and PhD Student Niranjana Krishnan

Data Analyst — *Lawrence University*

Research Administration Office — *2015*

- Analyzed data from a study to compare the academic success and mood towards the university of undergraduates from freshman to sophomore years
- Performed statistical analyses using SPSS

Data Collection Assistant — *Lawrence University*

Research Administration Office — *Sep 2014 May 2015*

- Assisted with the data collection for a study on the evaluation of warning lights installed at a busy crosswalk on the university campus
- Used Tracker software to determine the deceleration rate of vehicles from videos taken of cars approaching the crosswalk

Research Interests

Model Assessment, Explainable Machine Learning, Data Visualization, Random Forests, R Package Development

Publications

1. English, L., J., N., B., W., K., G., & M., L. (2023). Understanding the variation in vegetation composition of prairie restorations within crop yields. *Submitted to Ecological Restoration*.
2. Goode, K., Weber, M. J., & Dixon, P. M. (2023). WhoseEgg: Classification software for invasive carp eggs. *Submitted to PeerJ Life and Environment*.
3. Ausdemore, M. A., McCombs, A., Ries, D., Zhang, A., Shuler, K., Tucker, J. D., Goode, K., & Huerta, J. G. (2022). A probabilistic inverse prediction method for predicting plutonium processing conditions. *Frontiers in Nuclear Engineering*, *1*. https://doi.org/10.3389/fnuen.2022.1083164
4. Goode, K., Weber, M. J., Matthews, A., & Pierce, C. L. (2022). Evaluation of a random forest model to identify invasive carp eggs based on morphometric features. *North American Journal of Fisheries Management*. https://doi.org/https://doi.org/10.1002/nafm.10616
5. Goode, K., & Hofmann, H. (2021). Visual diagnostics of an explainer model: Tools for the assessment of LIME explanations. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *14*(2), 185–200. https://doi.org/https://doi.org/10.1002/sam.11500
6. Ball, E. E., Goode, K. J., & Weber, M. J. (2020). Effects of transport duration and water quality on age-0 walleye stress and survival. *North American Journal of Aquaculture*, *82*(1), 33–42. https://doi.org/https://doi.org/10.1002/naaq.10114
7. Dixon, P. M., Goode, K., & Lay, C. (2020). *Profile likelihood confidence intervals for ECx*. https://dr.lib.iastate.edu/entities/publication/7e0d7d0a-f514-4642-9814-c3b7bd821cc0
8. Goode, K., Ries, D., & Zollweg, J. (2020). Explaining neural network predictions for functional data using principal component analysis and feature importance. *AAAI FSS-20: Artificial Intelligence in Government and Public Sector*. https://doi.org/10.48550/ARXIV.2010.12063

Talks

1. Goode, K., Ries, D., & Zollweg, J. (2020, November). Explaining neural networks with functional data using PCA and feature importance. *AAAI 2020 Fall Symposium on AI in the Government and Public Sector*.
2. Goode, K., & Hofmann, H. (2019, July). Visual diagnostics of a model explainer: Tools for the assessment of LIME explanations from random forests. *Joint Statistical Meetings*.
3. Goode, K. (2019, June). A review and discussion of residuals for mixed models. *NCCC-170 Meeting*.

Posters

1. Goode, K., & Hofmann, H. (2019, May). Using LIME to interpret a random forest model with an application to bullet matching data. *Midwest Statistical Machine Learning Colloquium*.
2. Goode, K., & Hofmann, H. (2019, April). Using LIME to interpret a random forest model with an application to bullet matching data. *Iowa State University Graduate and Professional Student Research Conference*.

3. Goode, K., & Rey, K. (2018, May). Introducing ggResidpanel: An r package for easy visualization of residuals. *Kansas State University Conference on Applied Statistics in Agriculture*.

Software

1. Goode, K. (2022). *TreeTracer: Trace plots using ggplot2*. `https://github.com/goodekat/TreeTracer/`
2. Goode, K. (2022). *Limeaid: Diagnose LIME explanations*. `https://github.com/goodekat/limeaid`
3. Goode, K., McClernon, K., Zhao, J., Zhang, Y., & Huo., Y. (2022). *Redres: Residuals and diagnostic plots for mixed models*. `https://github.com/goodekat/redres.git`
4. Goode, K., & Rey, K. (2019). *ggResidpanel: Panels and interactive versions of diagnostic plots using 'ggplot2'*. `https://CRAN.R-project.org/package=ggResidpanel`