

Katherine Goode

RESEARCH AND DEVELOPMENT STATISTICIAN

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

☎ 505-844-1998 | ✉ kjgoode@sandia.gov | 🏠 goodekat.github.io | 🐙 github.com/goodekat | 🌐 katherine-goode-363517a9

I'm a research and development statistician at Sandia National Labs. My research interests include model assessment, explainable machine learning, and data visualization. I enjoy consulting and collaborating with scientists on data analyses.

Education

Iowa State University

DOCTOR OF PHILOSOPHY, STATISTICS

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

[Ames, IA](#)

2016-2021

University of Wisconsin, Madison

MASTER OF SCIENCE, STATISTICS

[Madison, WI](#)

2013-2015

Lawrence University

BACHELOR OF ARTS, MATHEMATICS

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

[Appleton, WI](#)

2009-2013

Awards

Sandia National Laboratories Up and Coming Innovator Award

2023

Midwest Statistical Machine Learning Colloquium Poster Award

AWARDED FOR '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

Senior Member of Technical Staff

STATISTICAL SCIENCES DEPARTMENT

- Research and development statistician

[Sandia National Laboratories](#)

Dec 2021 - Current

Postdoctoral Researcher

STATISTICAL SCIENCES DEPARTMENT

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

[Sandia National Laboratories](#)

Sep 2021 - Dec 2021

Research and Development Intern

STATISTICAL SCIENCES DEPARTMENT

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

[Sandia National Laboratories](#)

Dec 2019 - Sep 2021

Graduate Research Assistant

NATURAL RESOURCE ECOLOGY AND MANAGEMENT

- 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

[Iowa State University](#)

Jan 2021 - June 2021

Statistical Consultant

AGRICULTURE EXPERIMENT STATION

- Senior consultant from May 2018 to May 2020
- Helped with administrative decisions
- Trained and mentored 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

[Iowa State University](#)

May 2016 - Dec 2020

Graduate Research Assistant

DEPARTMENT NATURAL RESOURCE ECOLOGY AND MANAGEMENT

- 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

[Iowa State University](#)

May 2019 - Aug 2019

Data Analyst

RESEARCH ADMINISTRATION OFFICE

- 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

[Lawrence University](#)

2015

Data Collection Assistant

RESEARCH ADMINISTRATION OFFICE

- 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

[Lawrence University](#)

Sep 2014 May 2015

Publications

1. Goode, K., Ries, D., & McClernon, K. (2023). Characterizing climate pathways using feature importance on echo state networks. *Submitted to Statistical Analysis and Data Mining*.
2. Goode, K., Weber, M. J., & Dixon, P. M. (2023). WhoseEgg: Classification software for invasive carp eggs. *PeerJ Life and Environment*, 11.
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., McClernon, K., & Shand, L. (2023). Characterizing climate pathways using feature importance on echo state networks. *Albuquerque Chapter of the American Statistical Association Annual Meeting*, Santa Fe, New Mexico.
2. Goode, K., Ries, D., McClernon, K., & Shand, L. (2023). Characterizing climate pathways using feature importance on echo state networks. *Joint Statistical Meetings: Topic-Contributed Session “Deep Learning for Climate Change: Forecasts, Mitigation, and Adaption”*, Toronto, Canada.
3. Goode, K., Ries, D., McClernon, K., & Shand, L. (2023). Characterizing climate pathways using feature importance on echo state networks. *Sandia National Labs Machine Learning/Deep Learning Workshop*, Albuquerque, New Mexico.
4. Goode, K., Ries, D., McClernon, K., & Shand, L. (2023). Feature importance with deep echo state models. *SIAM Conference on Mathematical and Computational Issues in the Geosciences*, Bergen, Norway.

- Goode, K., Tucker, J. D., & Ries, D. (2022). Functional inverse prediction with elastic shape analysis. *Joint Statistical Meetings*, Washington D.C.
- Goode, K., Tucker, J. D., Ries, D., & Hofmann, H. (2022). An explainable pipeline for machine learning with functional data. *8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2022)*, Oslo, Norway.
- 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*.
- Goode, K., & Hofmann, H. (2019). Visual diagnostics of a model explainer: Tools for the assessment of LIME explanations from random forests. *Joint Statistical Meetings*, Denver, Colorado.
- Goode, K. (2019). A review and discussion of residuals for mixed models. *NCCC-170 Meeting*, University of Illinois Champaign–Urbana.

Posters

- Goode, K., Ries, D., McClernon, K., & Shand, L. (2023). Validating climate pathways using feature importance on echo state networks. *Confernece on Data Analysis*, Santa Fe, New Mexico.
- Goode, K., Ries, D., Tucker, J. D., & Shand, L. (2022). Feature importance with deep echo state models for long-term climate forecasting. *Climate Informatics*, Asheville, North Carolina.
- Goode, K., & Hofmann, H. (2019). Using LIME to interpret a random forest model with an application to bullet matching data. *Midwest Statistical Machine Learning Colloquium*, Ames, Iowa.
- Goode, K., & Hofmann, H. (2019). Using LIME to interpret a random forest model with an application to bullet matching data. *Iowa State University Graduate and Professional Student Research Conference*, Ames, Iowa.
- Goode, K., & Rey, K. (2018). Introducing ggResidpanel: An r package for easy visualization of residuals. *Kansas State University Conference on Applied Statistics in Agriculture*, Manhattan, Kansas.

Software

- Goode, K. (2022). *TreeTracer: Trace plots using ggplot2*. <https://github.com/goodekat/TreeTracer/>
- Goode, K. (2022). *Limeaid: Diagnose LIME explanations*. <https://github.com/goodekat/limeaid>
- 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>
- Goode, K., & Rey, K. (2019). *ggResidpanel: Panels and interactive versions of diagnostic plots using 'ggplot2'*. <https://CRAN.R-project.org/package=ggResidpanel>

Workshops

Industrial Math/Stat Modeling (IMSM) Workshop for Graduate Students

Asheville, NC

THE STATISTICAL AND APPLIED MATHEMATICAL SCIENCES INSTITUTE (SAMSI)

July 2019

- Two week research workshop
- Worked in a research group mentored by senior statisticians from Rho Inc.
- Analyzed continuously monitored glucose data using functional data analysis
- Assisted with the writing of a report and presentation on the research analysis

Teaching

Workshops

Predictive plant phenomics graduate student statistics bootcamp

Iowa State University

PREDICTIVE PLANT PHENOMICS PROGRAM

2018-2019

- Led a one day statistics bootcamp
- Discussed randomization, confidence intervals, and design of experiments
- Prepared slides

Instructor

STAT 101: Introduction to statistics

Iowa State University

DEPARTMENT OF STATISTICS

Spring 2016

- Prepared and led lectures
- Wrote and graded exams
- Topics included summary statistics, visualization, normal distribution, hypothesis testing, confidence intervals, and JMP

MATH 107: Elementary statistics

Lawrence University

MATHEMATICS DEPARTMENT

Fall 2015

- Organized the curriculum
- Prepared and graded homework and exams
- Topics included summary statistics, visualizations, randomization tests, bootstrap, normal distribution, hypothesis testing, confidence intervals, and R

Teaching Assistant

BMI 552: Regression methods for population health graduate students

University of Wisconsin, Madison

DEPARTMENT OF BIOSTATISTICS & MEDICAL INFORMATICS

Spring 2016

- Taught labs
- Held office hours
- Topics included simple and multiple linear regression, logistic regression, survival analysis, and SAS

BMI 551: Introduction to biostatistics for population health graduate students

University of Wisconsin, Madison

DEPARTMENT OF BIOSTATISTICS & MEDICAL INFORMATICS

Fall 2014

- Taught labs
- Held office hours
- Topics included summary statistics, visualizations, probability, normal distributions, hypothesis testing, confidence intervals, and R

STAT 301: Introduction to statistical methods for nonstatistics majors

University of Wisconsin, Madison

DEPARTMENT OF STATISTICS

Summer 2014

- Prepared and led discussions
- Graded homework and exams
- Held office hours
- Topics included summary statistics, visualizations, probability, normal distributions, hypothesis testing, and confidence intervals

STAT 302: Accelerated introduction to statistical methods for statistics undergraduate majors

University of Wisconsin, Madison

DEPARTMENT OF STATISTICS

Spring 2014

- Prepared and led discussions
- Graded homework and exams
- Held office hours
- Topics included summary statistics, visualizations, randomization tests, bootstrap, normal distribution, hypothesis testing, confidence intervals, and R

STAT 371: Introductory applied statistics for the life sciences

University of Wisconsin, Madison

DEPARTMENT OF STATISTICS

Fall 2013

- Prepared and led discussions
- Graded homework and exams
- Held office hours and worked in the statistics help room
- Topics included summary statistics, visualizations, probability, normal distributions, hypothesis testing, confidence intervals, and R

Mentor

Academic Mentor for Minority and First Generation Undergraduates

University of Wisconsin, Madison

CENTER FOR ACADEMIC EXCELLENCE

Fall 2014 - Spring 2015

- Mentored minority and first generation undergraduate students enrolled in statistics courses
- Met weekly throughout the semester with individuals or small groups to review statistical concepts from class and make the material approachable
- Discussed and encouraged strategies for academic success

Tutor

Statistics Tutor

University of Wisconsin, Madison

DEPARTMENT OF STATISTICS

Fall 2014 - Spring 2015

- Tutored undergraduate students
- Various introductory statistics courses

Service

IOWA STATE STATISTICAL GRAPHICS WORKING GROUP WEEKLY SEMINAR ORGANIZER	Sep 2019 - May 2020
GRADUATE STUDENT REPRESENTATIVE TO ISU STATISTICS DEPARTMENT FACULTY MEETINGS	Sep 2018 - May 2019
RECYCLING COORDINATOR FOR ISU STATERS (STATISTICS GRADUATE STUDENT ORGANIZATION)	Sep 2017 - May 2019
TREASURER AND MEMBER OF STATCOM (STATISTICS IN THE COMMUNITY) AT IOWA STATE	Sep 2017 - May 2019

Skills

WORKING KNOWLEDGE: GITHUB, JMP, LATEX, PYTHON, R, R MARKDOWN, R SHINY, SAS

BASIC KNOWLEDGE: C, SPSS