Scotland C. Leman

CONTACT INFORMATION	Department of Statistics	*Cell:* (540) 449-9860
	Virginia Tech	*Office:* (540) 231-5441
	401A Hutcheson Bldg	*E-mail:* leman@vt.edu
	Blacksburg, VA 24061 USA	*WWW:* www.apps.stat.vt.edu/leman

Statistician, with expert level knowledge in Bayesian methodology and stochastic based inference.

RESEARCH INTERESTS

Bayesian statistics, visual analytics, large scale stochastic simulation, statistical genetics, molecular evolution, epidemiology , coalescence theory, branching processes, time series analysis, hierarchical modeling, data augmentation, Markov chain Monte Carlo methods, mixing theory, importance sampling methodology

EDUCATION

Duke University, Durham, North Carolina USA

Ph.D., Statistics, April 2007

- Dissertation Title: "On Evolutionary Theory, Inference, And Simulation: A Genealogical Perspective"

M.S., Statistics, January 2005

Stanford University, Palo Alto, California USA

M.S., Scientific Computing and Computational Mathematics, May 2003

University of California, at Davis, Davis, California USA

B.S., Mathematics, May, 2001 (Cum Laude, Honors)

CURRENT PROFESSIONAL POSITION

Virginia Tech, Blacksburg, Virginia USA
Associate Professor (Tenured)   **2013-current**
The Department of Statistics

Courses taught and developed:
- Statistical Inference: STAT 5114 (graduate)
  Years taught: 2014, 2015, 2016, 2017, 2018
- Bayesian statistics: STAT 5444 (graduate)
  Years taught: 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018
- Statistical simulation: STAT 5314 (graduate)
  Years taught: 2008, 2009, 2010, 2011, 2012, 2015, 2017
- Introduction to Data Analytics and Visualization: CMDA 3654 (undergraduate)
  Years taught: 2014
- Data Analytics: STAT 5525 (graduate)
  Years taught: 2013, 2015 ,2017
- Advanced Data Analytics: STAT 5526 (graduate)
  Years taught: 2012
- Stochastic Modeling and Inference: STAT 6224 (graduate)
  Years taught: 2013

Virginia Tech, Blacksburg, Virginia USA
Assistant Professor   **2007-2013**
The Department of Statistics.

REFEREED
PUBLICATIONS

2017 Publications

Hoegh A, Maiti D, Leman SC; The Multiset Model Selection Algorithm. *Journal of Computational and Graphical Statistics*, DOI: 10.1080/10618600.2017.1379408.

Hoegh A, Leman SC; Correlated Model Fusion. *Applied Stochastic Models in Bussiness and Industry*, DOI: 10.1002/asmb.2261.

Carzolio M, Leman SC; Weighted Particle Tempering, *Computational Statistics and Data Analysis*, 114:26–37.

Fry JT, Hoegh A, Leman SC, Montesano M; A Bayesian Hierarchical Model for Estimating the Cost of Postponing the Cyclo-Cross National Championships. *Journal of Applied Statistics*, 4(2):298–305.

Crandell I, Millican AJ, Leman S, Smith E, Alexander WN, Devenport WJ, Vasta R, and Gramacy R, Binois, M; Anomaly Detection in Large-Scale Wind Tunnel Tests Using Gaussian Processes, *American Institute of Aeronautics and Astronautics*, DOI: 10.2514/6.2017-4131.

Zeitz JS, Dowling M, Wenskovitch J, Crandell I, Wang M, House L, Leman SC, North C; Observation-Level and Parametric Interaction for High-Dimensional Data Analysis. *ACM Transactions on Interactive Intelligent Systems*, xx(x):xxx-xxx.

Chen X, Self JZ, House, L, Wenskovitch J, Sun, M, Wycoff N, Evia JR, Leman SC North C; Be the Data: Embodied Visual Analytics. *IEEE Transactions on Learning Technologies*, PP(99):1. DOI: 10.1109/TLT.2017.2757481.

Zeitz J, Self N, House L, Evia JR, Leman SC, North C; Bringing Interactive Visual Analytics to the Classroom for Developing EDA Skills. *Journal of Computing Sciences in Colleges*, 33(3):115–125.

2016 Publications

Hoegh A, Ferreira MAR, Leman SC; Spatiotemporal Model Fusion: Multiscale Modeling of Civil Unrest. *Journal of the Royal Statistical Society: Series C*, 65(4):529-545.

Han C, House L, Leman SC; Expert-Guided Generative Topographical Modeling with Visual to Parametric Interaction. *PLoS ONE*, 11(2): e0129122. doi:10.1371/journal.pone.0129122.

Leman SC, Hoegh A; Efficient Metropolis-Hastings Proposal Mechanisms for Bayesian Regression Tree Models: Good Trees from Bad Proposals (comments). *Bayesian Analysis*, 11(3):932-934

Chen X, House L, Self JZ, Leman SC, Evia JR, Fry JT, North C; Be the Data: An Embodied Experience for Data Analytics, *American Educational Research Association; AERA2016*.

Self, JZ, Hu, X, House, L, Leman, S, North, C; Designing Interactive Dimension Reduction Algorithms. *Human Centered Machine Learning; CHI2016*.

2015 Publications

Hoegh A, Carzolio M, Crandell I, Hu X, Roberts L, Song Y, Leman SC; Nearest-neighbor matchup effects: accounting for team matchups for predicting March Madness. *Journal of Quantitative Analysis in Sports* (2015), 11(1):29-37.

Leman S, House L, and Hoegh A; Developing a New Interdisciplinary Computational Analytics Undergraduate Program: A Qualitative-Quantitative-Qualitative Approach. *The American Statistician* (2015), 59(4):397-408.

House L, Leman SC, and Han C; Bayesian Visual Analytics (BaVA). *Journal of Statistical Analysis and Data Mining* (2015), 8(2):1-13.

Han C, Leman SC, and House L; Covariance-Guided Mixture Probabilistic Principal Component Analysis (C-MPPCA). *Journal of Computational and Graphical Statistics* (2015), 24(1):66-83.

Hoegh A, Leman SC, Saraf P, and Ramakrishnan N; Bayesian Model Fusion for Forecasting Civil Unrest. *Technometrics* (2015), 57(3), 332-340.
  ⋆ **Youden Award** recipient (2016) for best expository paper; established by ASQ/ASA.

Awe OO, Crandell I, Adepoju AA, and Leman SC; A Time Varying Parameter State-Space Model for Analyzing Money Supply Economic Growth Nexus. *Journal of Statistical and Econometric Methods* (2015), 4(1), pp.73-95.

Ramakrishnan N, Lu C, Marathe M, Marathe A, Vullikanti A, Eubank S, Leman SC, Roan M, Brownstein J, Summers K; Model-Based Forecasting of Significant Societal Events. *Intelligent Systems*, IEEE (2015), 30:86-90

Hoegh A, Leman SC; A Spatio-Temporal Model for Assessing Winter Damage Risk to East Coast Vineyards. *Journal of Applied Statistics* (2015), 42(4):834-845.

House L, Leman SC; Han C, Bayesian visual analytics: BaVA. *Statistical Analysis and Data Mining*: The ASA Data Science Journal (2015), 8: 1-13.

2014 Publications
Leman SC, House L, Szarka J, Nelson H; Life on the bubble: Who's in and who's out of March Madness? *Journal of Quantitative Analysis in Sports* (2014), 10(3): 315-328

Bradel L, North C, House L, Leman SC; Multi-Model Semantic Interaction for Text Analytics. *Visual Analytics Science and Technology* (VAST) (2014), IEEE.

Nsoesie E, Leman SC; A Dirichlet Process Model for Prediction of Epidemic Curves, *BMC Infectious Diseases* (2014), 14:12.

Marakeby H, Badr E, Torkey H, Song Y, Leman SC, Monteil CL, Heath LS, Vinatzer BA; A System to Automatically Classify and Name Any Individual Genome-Sequenced Organism Independently of Current Biological Classification and Nomenclature. *PLoS ONE* (2014), 9(2): e89142.

2013 Publications
Leman SC, House L, Maiti D, Endert A, North C; Visual to Parametric Interaction. *PLoS ONE* (2013), 8(3): e50474. doi:10.1371/journal.pone.0050474

Hu X, Bradel L, Maiti D, House L, North C, Leman SC; Semantics of Directly Manipulating Spatializations, *IEEE Transactions on Visualization and Computer Graphics* (2013), 19(12): 2052-2059.

Clarke CR, Chinchilla D, Hind SR, Taguchi F, Miki R, Ichinose Y, Martin GB, Leman SC, Felix G, Vinatzer B; Allelic Variation in FLS2-Dependent Detection of Two Distinct Epitopes of Pseudomonas Syringae Flagellin by Plants. *New Phytologist* (2013), 200(3): 847-60.

Monteil CL, Studholme D, Leman SC, Moris C, Vinatzer B; Non-agricultural reservoirs contribute to emergence and evolution of Pseudomonas syringae crop pathogens. *New Phytologist* (2013),199(3): 800-11

2012 Publications
Velasco-Cruz C, Leman SC, Smith EP; Assessing the Risk of Rising Temperature on Brook Trout: A Spatial Dynamic Linear Risk Model. *Journal of Agricultural, Biological, and Environmental Statistics (JABES)* (2012), DOI: 10.1007/s13253-012-0088-8

Leman SC, House L; Improving Mr. Myagi's Coaching Style: Teaching Data Analytics with Inter-

active Data Visualizations. *Chance* (2012), 25(2): 4-12

2011 Publications
Cai R, Yan S, Haijie L, Leman SC, Vinatzer BA; Reconstructing Host Range Evolution of Bacterial Plant Pathogens Using Pseudomonas Syringae pv. Tomato and Its Close Relatives as a Model. *Infection, Genetics and Evolution* (2011), 11(7): 1738-51

Cai R, Lewis J, Yan S, Liu H, Clarke SR, Campanile F, Almeida SF, Studholme DJ, Lindeberg M, Schneider D, Zaccardelli M, Setubal JC, Morales-Lizcano NP, Bernal A, Coaker G, Baker C, Bender CL, Leman SC, Vinatzer BA; The Plant pathogen Pseudomonas Syringae pv. Tomato is Genetically Monomorphic and Under Strong Selection to Evade Tomato Immunity, *Public Library of Science (PLoS) Pathogens* (2011), 7(8):e1002130.

Endert A, Fox S, Maiti D, Leman SC, North C; The Semantics of Clustering: Analysis of User-Generated Spatializations of Text Documents. *Advanced Visual Interfaces 2012 (AVI2012)* (2011), 978-1-4503-1287-5

Endert A, Han C, Maiti D, House L, Leman SC, North C; Observation-level Interaction with Statistical Models for Visual Analytics. *IEEE Visual Analytics Science and Technology (VAST)* (2011), 978-1-4673-0015-5

Sharakhova MV, Xia A, Leman SC, Sharakhov IV; Arm-Specific Dynamics of Chromosome Evolution in Malaria Mosquitoes, *BioMed Central (BMC) Evolutionary Biology* (2011) 11:91.

2010 Publications
Sharakhova MV, George P, Brusentsova IV, Leman SC, Bailey JA, Smith CD, Sharakhov IV; Genome mapping and characterization of the Anopheles gambiae heterochromatin, *BMC Genomics*, (2010), 11:459

Xia A, Sharakhova MV, Leman SC, Tu Z, Bailey JA, Smith CD, Sharakhov IV; Genome Landscape and Evolutionary Plasticity of Chromosomes in Malaria Mosquitoes, *PLoS One* (2010), 5(5): e10592

Almeida NF, Yan S, Cai R, Clarke CR, Morris CE, Schaad NW, Lacy GH, Jones JB, Castillo JA, Bull CT, Leman SC, Guttman DS, Setubal JC, Vinatzer BA; A Multilocus sequence typing & analysis database and website for plant-associated and plant-pathogenic microorganisms. *Phytopathology*, (2010), 100(3), pp. 208–215

2009 Publications
Leman S, Levy F, Walker E; Modeling The Spread Of Infectious Disease Using Genetic Information Within A Marked Branching Process. *Statistics in Medicine*, (2009), 28(29): 3581-3717

Leman SC, Chen Y, Lavine M; The Multiset Sampler. *Journal of the American Statistical Association (JASA)* (2009), 104(487): 1029-1041

Levy F, Leman SC, Walker E; Nosocomial Transmission Clusters and Risk Factors in Moraxella Catarrhalis. *Epidemiology and Infection* (2009), 137(4):581-90.

2007 Publications
Leman SC, Uyenoyama MK, Chen Y, Lavine M; The Evolutionary Forest Algorithm. *Bioinformatics*, (2007), 23(15):1962-1968,

2005 Publications
Leman SC, Chen Y, Stajich JE, Noor, MAF, Uyenoyama MK; Likelihoods from summary statistics: A Recent Divergence Between Species. *Genetics*, (2005), 171(3):1419-1436.

PUBLICATIONS IN PRESS	Updating.
TECHNICAL REPORTS	Leman SC, House L, Han C, Bayesian Visual Analytics (BAVA), FODAVA Technical Report (#FODAVA-10-02).

FUNDING SUMMARY	• Funding total (Gross): $28,148,148 • Total after allocated percent credit (Net): $4,236,731
CURRENT FUNDED RESEARCH	Computer Human Interactive Text Analytics (CHITA), General Dynamics Award Amount: $299,519 (20% credit) PIs: Chris North, Scotland Leman, Nicholas Polys, Leanna House
	UrbComp: Data Science for Modeling, Understanding, and Advancing Urban Populations, National Science Foundation (NSF), NRT-DTSE, 1545362. Award Amount: $2,999,128.00 (5% credit) PIs: Naren Ramakrishnan, Leanna House, Layne Watson, Mark Embree, Scotland Leman
	Data Analytics for Large Sensor Systems, Office of Naval Research (ONR). Award Amount: $260,752.00 (17% credit) PIs: Eric Smith, William Devenport, Leanna House, Scotland Leman, William Alexander
	Usable Big-Data Analytics via Multi-scale Visual To Parametric Interaction, National Science Foundation (NSF), III, 1447416 . Award Amount: $998,914.00 (25% credit) PIs: Chris North, Scotland Leman, Leanna House, Yong Cao
	Critical Thinking with Data Visualization, National Science Foundation (NSF), TUES, 1141096. Award Amount: $199,997 (33% credit) PIs: Leanna House, Scotland Leman, Chris North
PREVIOUS FUNDED RESEARCH	Human Machine Collaboration for Data Discovery, General Dynamics Award Amount: $250,000 (25% credit) PIs: Chris North, Scotland Leman, Nicholas Polys, Naren Ramakrishnan
	Early Model Based Event Detection Using Surrogates, Intelligence Advanced Research Projects Activity (IARPA), OSI (4th year extension). Award Amount: $3,000,000.00 (15% credit) PIs: Naren Ramakrishnan, Tien Chang Lu, Scotland Leman, Madhav Marathe, Michael Roan
	Early Model Based Event Detection Using Surrogates, Intelligence Advanced Research Projects Activity (IARPA), OSI. Award Amount: $19,494,495.00 (15% credit) PIs: Naren Ramakrishnan, Scotland Leman, Tien Chang Lu, Madhav Marathe, Michael Roan
	Be The Data, Institute for Creativity, Arts, and Technology (Virginia Tech). Award Amount: $50,000.00 (25% credit) PIs: Leanna House, Scotland Leman, Chris North, Wu Feng
	Data Analytics for Large Acoustic Array Data, Office of Naval Research (ONR). Award Amount: $44,163.00 (17% credit) CoPIs: Eric Smith, William Devenport, Leanna House, William Alexander
	First Year Experiences for Statistics Majors, Virginia Tech Quality Enhancement Plan (QEP) Award Amount: $32,000 (10% credit) CoPIs: Jane Robertson, Leanna House, Scotland Leman, Chris Franck, Eric Vance, Eric Smith
	Bayesian Analysis in Visual Analytics, National Science Foundation (NSF), CCF, 0937071. Award Amount: $499,307 (40% credit) PIs: Scotland Leman, Chris North, Leanna House
	User-Guided Spatialization for Visualizing NSF Award Portfolios. NSF 10222504. Award Amount: $24,900 (33% credit) PIs: Scotland Leman, Chris North, Leanna House

Computational Modeling and Data Analytics (CMDA)

Organizer/Curriculum Developer

Instrumental in developing a new major at Virginia Tech called Computational Modeling and Data Analytics (CMDA). CMDA is a relatively unique blend of statistical, computational, and mathematical approaches for analyzing data and complex scenarios which require integrated techniques, teamwork, and an understanding of real application areas. Starting in 2011, a preliminary meeting on an integrative computational curriculum was help which became the platform for the CMDA program. A brief summary of the timeline for creating the CMDA major follows.

- Curriculum drafted and revised (2011-2012).
- Degree accreditation proposal drafted (2012-13).
- Selected CMDA courses piloted (2013).
- CMDA program approved (2014).
- First incoming CMDA class (2016).

Course Development

While at Virginia Tech, I have developed several courses at both the graduate and undergraduate levels in both the Statistics and CMDA departments. A sampling of developed courses, which include lesson planning, design, and governance (and associated approvals).

Graduate Statistics Course Summaries

- Bayesian Statistics (STAT 5444): Graduate level course on Bayesian inferential principles and methodologies.
- Statistical Simulation (STAT 5314): Graduate level course on computational and simulation based methods for performing statistical inference.
- Stochastic Simulation and Inference (STAT 6124): Graduate level course on performing statistical inferences using complex stochastic models and their associated computational considerations.

Cross-listed Statistics/Computer Science Course Summaries

- Data Analytics I (STAT/Comp Sci. 5225): Graduate level course on contemporary statistical methodologies that are both algorithmically and computationally oriented and especially useful for analysis of high dimensional data. T
- Data Analytics I (STAT/Comp Sci. 5226): Graduate level course focusing on more theoretical concepts and techniques arising in analytics.

Undergraduate CMDA Course Summaries

- Introductory Data Analytics and Visualization (CMDA 3654): Undergraduate level course on the basic principles and techniques in data analytics and visualization.
- Intermediate Data Analytics and Machine Learning (CMDA 4654): Undergraduate level course on the technical aspects of analytical methods and algorithms.
- Computational Stochastic Modeling (CMDA 4664): Undergraduate level course on model development and the associated computational tasks.
- Integrated Quantitative Sciences (CMDA 2005/2005): Undergraduate level course(s) on the integration of basic and advanced analytics skills. This sequences replaces the traditional multivariate calculus, probability, statistics, matrix algebra, etc. courses with a single sequence that helps contextualize and motivate previously unintegrated techniques and theory.

Title: *Analytics for Large Scale, Diverse Sensor Systems*,
Location: Bethesda, MD
Venue: Carderock Naval Base, Feb. 2018.

Title: *Extensions of the Multiset Sampler*,
Location: Boulder, CO,
Venue: The University of Colorado, Department of Applied Mathematics, Sept. 2017.

Title: *Large Scale Anomaly Detection in Multi-Sensor Systems*,
Location: Denver, CO,
Venue: The American Institute of Aeronautics and Astronautics, Aviation Forum. June 2017.

Title: *Big Data Analytics for Large Sensor Systems*,
Location: Washington, DC,
Venue: Office of Naval Research Program Review, Aug. 2017.

Title: *Extensions of the Multiset Sampler*,
Location: Dunedin, NZ,
Venue: The University of Otago, Department of Mathematics and Statistics Oct. 2016.

Title: *Data Analytics for Large Sensor Systems*,
Location: State College, PA (Naval water tunnel test site),
Venue: Office of Naval Research Program Review, Aug. 2016.

Title: *Multi-scale Visual to Parametric Interaction*,
Location: Paris, France,
Venue: Visual Analytics Science and Technology, Nov. 2014.

Title: *Visual To Parametric Interaction*,
Location: Montreal, QC,
Venue: Joint Statistics Meetings, Aug. 2013.

Title: *Semantics of Directly Manipulating Spatializations*,
Location: Altalnta, GA,
Venue: Visual Analytics Science and Technology, Oct. 2013.

Title: *Why MCMC Has Changed Everything*,
Location: Honolulu, HI,
Venue: The University of Hawaii, Department of Mathematics, Jan. 2013.

Title: *Bayesian Visual Analytics*,
Location: Harvard University, Cambridge, MA,
Venue: Invited Colloquium, Department of Statistics, Nov. 2012.

Title: *Particle Filtered Tempering: De-Correlating the Ladder*,
Location: Banff, Alberta, Canada,
Venue: Challenges and Advances in High Dimensional and High Complexity Monte Carlo, Mar. 2012.

Title: *Bayesian Visual Analytics*,
Location: Atlanta, GA,
Venue: Foundations of Data and Visual Analytics Workshop, Dec. 2011.

Title: *MultiSet Model Selection*,
Location: Pittsburgh, PA,
Venue: Case Studies in Bayesian Statistics and Machine Learning, Nov. 2011.

Title: *March Madness: Who's In and Who's Out*,
Location: Miami Beach, FL,
Venue: Joint Statistical Meetings (Invited Session on Sports Statistics), Aug. 2011.

Title: *Bayesian Visual Analytics (BaVA)*,
Location: Valencia, Spain,
Venue: International Society for Bayesian Analysis, Jun. 2010.

Title: *Bayesian Visual Analytics*,
Location: Vancouver, BC, Canada,
Venue: Joint Statistical Meetings Session (Invited Session, organizer and contributor), Aug. 2010.

Title: *Modeling The Spread Of Infectious Disease Using Genetic Information Within A Marked Branching Process*,
Location: Georgia Southern University, Statesboro, GA,
Venue: Invited colloquium, School of Public Health, Oct. 2009.

Title: *Bayesian Philosophies and Practicalities*,
Location: Washington, D.C.
Venue: Joint Statistical Meetings (Topic Contributed Session, organizer and contributor, Aug. 2009.

Title: *The Multiset Sampler*,
Location: Hamilton Island, AU,
Venue: International Society For Bayesian Analysis, Jun. 2008.

Title: *Speciation in Coalescence Models*,
Location: Carnegie Mellon, Pittsburgh, PA,
Venue: Bayesian Case Studies, Oct. 2008.

Title: *The Multiset Auxiliary Variable Sampler*,
Location: Durham, NC,
Venue: Statistical and Applied Mathematical Sciences Institute, Sep. 2008.

Title: *Modeling The Spread Of Infectious Disease Using Genetic Information Within A Marked Branching Process*,
Location: Blacksburg, VA,
Venue: National Dynamics Simulation Science Laboratory, Nov. 2008.

Title: *The Evolutionary Forest Algorithm*,
Location: Hamilton Island, Australia,
Venue: International Society For Bayesian Analysis, Jun. 2006.

Miscellaneous others: Joint Statistical Meetings (2004-2007, Graduate student years: various *topic contributed* presentations).

INVITED SESSION ORGANIZER

Roundtable discussion leader at the Joint Statistical Meetings: "Don't Forget: Quantitative Statistics Courses are Very Qualitative.", Seattle, WA, August, 2015.

JSM organizer (invited session), The I.J. Good Memorial session: "Remembering a Genius", Vancouver, BC, August, 2010. Speakers: Jim Berger, Adrian Raftery, Steve Fienberg; Discussant: David Banks; Chairman: Golde Holtzman

JSM organizer, Bayesian Philosophies and Practicalities, Washington, D.C., August, 2009.

IBM, San Jose (Almaden), California USA
Developer/Tester/Technical Author **2003**
Worked with a team to develop and test current software in the web mining field.
- Wrote two published books describing the use of developed software and tools. These books were the core sources of information for the client to follow while integrating the software into their systems. These provided detailed methodology and examples through mock company scenarios; see:

  Roland Tretau, David Chiang, Daniel Greisokh, Scotland Leman, Roman Shekhtmeyster; WebFountain Application Development Guide. *International Business Machines*, first addition, ISBN: 0-7384-9992-7.

- Developed and tested software critical for end user consumption of the products.

University of Chicago, Medical Center, Chicago, Illinois USA
Researcher/Developer/Consultant **2001-2003**
Responsible for developing a software interface to access experimental data from a large relational database and transport and process it into a statistical analysis tool.
- Developed Graphical User Interface in Java to access data and port it to existing software.
- Redesigned portions of the database to aid in efficient queries. Ultimately enhanced query times by a factor of 1000.
- Analyzed data and performed inference.
- Algorithmic development for use in simulation and inference procedures.

Institute of Theoretical Dynamics, Davis, California USA
Research Assistant/Developer/Analyst **1999-2001**
Worked closely with scientists and mathematicians. In charge of analyzing simulation routines and results. Performed fluid flow experiments critical to answering fundamental scientific questions.
- Simulation development for three dimensional Navier-Stokes equations.
- Three dimensional mesh reconstruction used for simulating physical bodies.
- Analysis of simulations and data.

National Bioinformatics Network, Cape Town, South Africa

Instructor **September, 2006**

Taught graduate level students an intensive course on the topics of: generalized linear models, survival analysis, multi-dimensional scaling, cluster analysis, principal component analysis and data visualization.

- Class size of 25 students.
- Emphasis was placed on applications related to bioinformatics and genetics.
- A mixture of theory and software was used to demonstrate concepts.

The Duke School, Durham, North Carolina

Instructor **September, 2005 - May, 2006**

Taught an after school class on *algorithmic thought and computer programming* to middle school students.

- Class sizes of approximately 15 students.
- Object oriented computer programming in C++.
- Debugging emphasized using the integrated development environment Xcode.

Duke University, Durham, North Carolina

Teaching Assistant **August, 2003 - 2007**

Duties at various times have included office hours, consulting and leading weekly computer lab exercises.

Adjunct Instructor **May-August, 2007**

Taught undergraduate non statistics students (primarily; see Stat 103 in the Duke catalog).

- Class size of 28 students.
- Probability and statistical inference.
- Emphasized real data analyses (class project based).
- Team work encouraged.

Below is a sampling of notable media coverage that some of my (recent) work has featured in.

- WSLS, Channel 11 News, Jan 14, 2015, "Former Virginia Tech professor portrayed in WWII film, The Imitation Game" by Dawn Jeffries;
http://www.wsls.com/story/27849889/former-virginia-tech-professor-portrayed-in-wwii-film-the-imitation-game
- The Oklahoman, Feb 27, 2015, "NCAA Tournament: Evidence says name brand programs will get the selection committee's nod every time", by Barry Tramel;
http://newsok.com/ncaa-tournament-evidence-says-name-brand-programs-will-get-the-selection-committees-nod-every-time/article/5397173
- The Daily Bruin, Mar 17, 2015, "Q&A: Statistics analyst talks prestige bias in March Madness", by Andrew Erickson;
http://dailybruin.com/2015/03/17/qa-statistics-analyst-talks-prestige-bias-in-march-madness
- The American Statistical Association News, Feb 24, 2015, "A ?MARQUEE? BIAS CAN INFLU-ENCE WHICH ?BUBBLE? TEAMS GET INTO MARCH MADNESS", by Jeffrey Meyers
http://www.amstat.org/newsroom/pressreleases/2015-MarqueeFactorHelpsTeamsGetintoMarchMadness
- Sports Are 80 Percent Mental, Mar 3, 2011, "Is There Bias In Selection Of March Madness Teams?" by DanPeterson;
http://blog.80percentmental.com/2011/03/is-there-bias-in-selection-of-march.html
- fromtheeditr, March 2, 2011, "The Bias Behind 'Bracketology'; A Study" by Dan Smith; http://fromtheeditr.blogspot.com/2011/03/bias-behind-bracketology-study.html
- WSLS, Channel 10, Roanoke, VA, March 2, 2011 "Stats professors: Virginia Tech up against odds for NCAA tourney bid" by Ken Heineck
- WDBJ, Channel 7, Roanoke, VA, March 1, 2011 "Calculating VT's odds of making the NCAA Tournament Could biases have affected VT getting a bid in 2010?"
- Virginia Tech News, March 1, 2011, "March Madness: Statisticians quantify entry biases," by Catherine Doss
  Also appeared:
  - MSN Fox Sports,
  msn.foxsports.com/collegebasketball/story/MARCH-MADNESS-STATISTICIANS-QUANTIFY-ENTRY-BIASES-69543943
  - Science Daily,
  www.sciencedaily.com/releases/2011/03/110301111259.htm
  - Science Blog,
  scienceblog.com/43241/march-madness-statisticians-quantify-entry-biases/
  - Science Newsline Medicine,
  www.sciencenewsline.com/medicine/2011030112000067.html
  - Red Orbit,
  www.redorbit.com/news/sports/2005304/march_madness_statisticians_quantify_entry_biases/
- Tech Talk Live, February 28, 2011, Mentioned by Coach Seth Greenberg during an interview with Bill Roth
- College of Science Magazine, Virginia Tech, September 13, 2010, "Statisticians help researchers see their data in a new way," by Catherine Doss
- College of Science Magazine Video, Virginia Tech;
http://www.science.vt.edu/media/statistics-house-leman-video.html
- The Washington Post, March 17, 2010, "Hokies turn to statistic for answers to NCAA snub," by Mark Viera; http://voices.washingtonpost.com/hokies-journal/2010/03/hokies_turn_to_statistic_for_a.html#more
- Richmond Times Dispatch, Mar 16, 2010, "Virginia Tech stats professors doing NCAA tourna-ment study for Greenberg," by Darryl Slater; http://www.mytimesdispatch.com/index.php/sports/comments/virginia_tech_stats_professors_doing_ncaa_tournament_study_for_greenberg/11450/

ADVISING	Students advised that have led to published research: • Total Ph.D. disertations advised: 10 • Total Masters theses advised: 4 • Total Undergraduate students advised: 2
GRADUATED STATISTICS PH.D. STUDENTS	Dipayan Maiti: *Multiset Model Selection and Averaging, and Interactive Storytelling* Year completed: 2012 Ciro Velasco-Cruz: *Spatially Correlated Model Selection (SCOMS)* Year completed: 2012 Co-advisor: Eric P. Smith Chao Han: *Bayesian Visual Analytics: Interactive Visualization for High Dimensional Data* Year completed: 2012 Co-advisor: Leanna House Xinran Hu: *On Grouped Observation Level Interaction and Big Data Monte Carlo Sampling Algorithms* Year completed: 2014 Lucas Roberts: *Variable Selection and Decision Trees: The DiVaS and ALoVaS Methods* Year completed: 2014 Andrew Hoegh: *Predictive Model Fusion: A Modular Approach to Big, Unstructured Data* Year completed: 2016 Marcos Carzolio: *Selection of Advanced Markov Chain Monte Carlo Algorithms for Everyday Use: Weighted Particle Tempering, Practical Reversible Jump, and Extensions* Year completed: 2016 Yuhyun Song: *Linkage Based Dirichlet Processes* Year completed: 2016
GRADUATED OUTSIDE PH.D. STUDENTS	Rongman Cai: *New hypotheses about the origin of Pseudomonas syringae crop pathogens* Year completed: 2012 Co-advisor: Boris Vinatzer (Plant Pathology) Elaine Nsoesie: *Sensitivity Analysis and Forecasting in Network Epidemiology Models* Year completed: 2012 Co-advisor: Madhav Marathe, Richard Beckman (Virginia Bioinformatics Institute)
NON-THESIS ADVISING	Masters: Elaine Nsoesie (2011), Ho Cho (2012), Han Li (2012), Hayley Nelson (2012) Undergraduate: Robert Vasta (2016), Nathan Wycoff (2015).
UNDERGRADUATE ACTIVITIES	I have a strong interest in undergraduate curriculum development, mentoring, and teaching. Below are a (brief) sampling of my miscellaneous undergraduate activities. *Invited Lecturer: First year curriculum (FYE)* • Since 2012, I have contributed a visualization sequence (1 week of materials) to the FYE. • Roughly 20-40 students per year. • Tufte's rules, interactive visualization, exploratory analytics emphasized. *CMDA/Stat extracurricular involvements* • Have provided dozens of non-technical presentations. • Faculty judge for DataFest (ASAs undergraduate data competition). • Student project mentoring. *Undergraduate Revamp! committee* Co-member in charge of modernizing the undergraduate curriculum.

- Curriculum development (revamping) for the entire VT class directory.
- Weekly meetings (2015-2016)
- Course design and teaching materials constructed.

Departmental and University Service	Below I describe a sampling of various committees (both departmental and University wide) for which I've served.

Personnel/Departmental committee
- 4 member elected committee (2013-Present)
- Chairman (Elected: 2015,2016)
- Writes departmental letters concerning: *promotions, dismissals, faculty evaluations, discretionary issues.*

Teaching evaluation committee
- 2 members (2016)
- Responsible for teaching evaluations of un-tenured faculty
- Writes teaching critiques and mentors faculty on educational issues.

Computing committee
- 4 members (2007-2016)
- Chairman (2007-2016)
- Writes documentation requesting yearly allocations from state accreditation board.

Qualifying exam committee
- Evaluates and assigns awards to outstanding students.
- Writes and evaluates yearly qualifying exams (primarily inference section).
- Provides critiques of other QE sections.

Faculty search committees
I have served and led numerous faculty and staff searches (about 10).

Honor Societies and Academic Awards
- Mu Sigma Rho, National Honor Society, 2008
- Phi Kappa Phi, National Honor Society, 2000
- Pi Mu Epsilon, National Mathematics Honor Society, 2000
- Golden Key, National Honor Society, 2000
- Research Training Program (RTP) graduate, UC Davis, 2000
- Graduated Cum Laude, Honors in Mathematics, 2001

Professional Affiliations
- International Society for Bayesian Analysis (ISBA).
- The American Statistical Association (ASA).

Societal Awards
- **Youden Award** recipient for best expository paper; established by the American Society of Quality and the American Statistical Association, 2016.

Technical Skills
- Software: Matlab, R, CodeWarrior, Office.
- Languages: C/C++, Java, Python, Perl, Fortran, Matlab, R, SQL, some use of Unix shell scripts.
- Technical Skills: Probability and statistics. Ability to develop and deploy Bayesian models.
- Operating Systems: Macintosh, *nix Flavors, Windows.