

# Annual Report Executive Summary, 2007-08

## Department of Statistics

<b>Learning: Undergraduate</b>
Number of undergraduates 45 majors, 22 minors
Graduated 10
<b>Learning: Graduate</b>
Number of graduate students 70
Number of PhDs awarded: 6
Number of Masters awarded: 21
Funded on grants 9
Post-doctoral students funded 1
Funded on fellowships 2
<b>Engagement</b>
Consulting center activities 173 clients
Corporate partners program – 8 members
Secured funding for graduate student support from MINITAB, DuPont, RJ Reynolds
Five faculty are on editorial boards of journals, one is a major journal

## Details

### Grants (continuing):

Hoeschele (Co-PI), National Science Foundation DBI-0211863, (2002-2008), “Dissecting soybean resistance to Phytophthora by QTL analysis of host and pathogen expression profiles”, B.M. Tyler (PI), M.A. Saghai Maroof, I. A. Dorrance, S. St. Martin.

Hoeschele, I. (PI).Comprehensive Cancer Center Of Wake Forest University, (2006-2010), “Biostatistics Consulting”

Morgan, J.P. NSF DMS 06-04997. Symmetry and Asymmetry in Experimental Design. Award period: September 2006 -September 2009.

Smith, E. P., Co-PI (with S. Prisley, L. Carstensen, Jr. and K. Ye. Application of spatial uncertainty models to automate and enhance data fusion. National Geospatial Intelligence Agency. (8/1/05-7/31/08).

Smith, E.P., Samantha Bates, Saied Mostaghimi, Gene Yagow, Kevin Brannan, Donald Orth, Jim Berkson Model-based Clustering for Classification of Aquatic Systems and Diagnosis of Ecological Stress.

Vining, G. PI. Becton – Dickinson, 2006. Project Title: BD Assistantship. Period: 10/15/2006 – 8/15/2007.

### Grants (new):

Guo, F. Co-PI , “Modeling accident data for the 100-car naturalistic study”, sponsored by The National Surface Transportation Safety Center for Excellence.

Smith, E.P. Principal Investigator (with Penelope Pooler). 2006 renewal. Cooperation And Collaboration On Vital Signs Monitoring Program Of The Northeast Coastal And Barrier Network And Eastern Rivers And Mountain Network Of The National Park Service.

Smith, E. P. Principal Investigator, 2006-7. Ecological thresholds and structured decision analysis using aquatic macroinvertebrate as indicators of stream health in the mid Atlantic Highlands. US Geological Survey.

Y. Jiao, D. Orth, and E.P. Smith MARFIN two-year proposal 08MF004(RF) "Model complexity and stock assessment quality: an investigation of the performance of models of different complexity and implications for model selection in fisheries".

G. Gao (PI), I Hoeschele (Co-PI). National Institutes of Health (2007-2011), “Haplotyping and QTL mapping in pedigrees with missing data”.

K. Davy, I. Hoeschele (Co-PI).National Institutes of Health (2007-2009), “Angiotensin Receptor Blockade and Adipose Tissue Inflammation in Obesity”.

Kim, D-Y. Dynamic Interactions Among People, Livestock, and Savanna Ecosystems Under Climate Change (NSF CNH grant BCS-0709671). Role: Sole P.I. Subcontract with Michigan State University. Period: 2007-2010.

Kim, D-Y. Prototyping a Rangeland Decision Support System (NASA-SBR grant). Role: Co-investigator. Sponsor Applied Geosolutions, LLC with Michigan State University. Period: June-August 2007.

Vining, G., P.I., Pratt and Whitney, 2007. Project Title: Monte Carlo Simulation of the Engine Development Process. Period: 1/1/2008 – 12/31/2008.

Vining, G., P.I., Pratt and Whitney, 2007. Project Title: Investigation of Maximum Likelihood versus Rank Regression for Modeling of Life Time Data. Period: 11/1/2007 – 8/15/2008.

Vining, G., P.I., Becton – Dickinson, 2007. Project Title: BD Assistantship. Period: 10/1/2007 – 9/30/2008.

Vining, G., P.I., National Institute of Aerospace, 2007. Project Title: Rapid Prototyping Studies. Period: 8/15/2007 – 8/15/2008.

Vining, G., P.I., Pratt and Whitney, 2007. Project Title: Various smaller projects. Period: 1/1 – 12/31/2007

**Books (published or in-press):**

Banks, D. and Smith, EP (editors). *The Good Book: IJGood's comments, criticisms and conclusions*. Rice University Press (to appear)

**Book Chapter (published or in-press):**

Hoeschele I. (2007) Mapping quantitative trait loci in outbred populations. *Handbook of Statistical Genetics*, DJ Balding, M Bishop & C Cannings (eds.), Wiley, p. 623-677

Huebner, M., Kim, D., Ewart, S., Karmaus, W., Sadeghnejad, A., Arshad, SH. (2007), Patterns of GATA3 and IL13 Gene Polymorphisms Associated with Childhood Rhinitis and Atopy in a Brith Cohort. *J. Allergy and Clinical Immunology*. Available through E-pub.

Pang, H., Kim, I. and H. Zhao (2007). *Pathway-Based Methods for Analyzing Microarray Data.*, *Analysis of Microarray Data: A Network-Based Approach*, Edited by M. Dehmer and F. Emmert-Streib, John Wiley & Sons, Ltd, in press.

Vining, G.G. (2007). "Adapting Response Surface Methodology for Computer and Simulation Experiments," *The Grammar of Technology Development*, Tsubaki, H., Nishina, K., Yamada, S. eds., pp. 127-134.

**Papers Published (since last FAR):**

Bates Prins, S.C. and Smith, E.P. (2007). Using Biological Metrics to Score and Evaluate Sites: A Nearest-Neighbour Reference Condition Approach. *Freshwater Biology*, 52, 98-111

Gao, G. and I. Hoeschele (2007). A note on a haplotyping method in pedigrees. *Genetics, Selection, Evolution* 40: 25-36.

Guo, F., Dey, D., Holsinger, K., (2007). A hierarchical Bayesian approach for estimating origin of a mixed population. *IMS, Lecture Notes Monograph Series*

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2007), High Breakdown Estimation for Phase I Multivariate Control Charts. *Quality and Reliability Engineering International* 23, 615-629.

Kim, I., Liu, Y. and H. Zhao (2007). Bayesian methods for predicting interacting protein pairs using domain information, *Biometrics*, 63, 824-833. (Editor's invited paper in *Biometrics*, JSM 2007)

Kim, B., Kim, I., Lee, S., Rha, S. and H. Chung (2007). Ranking candidate genes for the biomarker development in a cancer diagnostics, *Korean Communication in Statistics*, 14, 169-182.

Kowalski, S.M., Parker, P.A., and Vining, G.G. (2007). "Tutorial on Split-Plot Experiments," *Quality Engineering*, 19, pp. 1-15.

- Leman, S.C. et al. (2007) The Evolutionary Forest Algorithm. *Bioinformatics*. 23 (15) 1962-1968.
- Morgan, J. P. and Reck, R. (2007). Resolvable designs with large blocks. *Annals of Statistics*, 35, 747-771.
- Morgan, J. P. (2007). Optimal incomplete block designs. *Journal of the American Statistical Association*, 102, 655-663.
- Morgan J. P. and Jin, B. (2007). Optimal experimentation in two blocks. *Journal of Statistical Theory and Practice*, 1, 357-375.
- Parker, P.A., Kowalski, S.M., and Vining, G.G. (2007). "Construction of Balanced Equivalent Estimation Second-Order Split-Plot Designs," *Technometrics*, 49, pp 56-65.
- Parvu, V. and Morgan, J. P. (2008). E-optimal designs for three treatments. *Journal of Statistical Planning and Inference*, 138, 642-653.
- Mahmoud, M. A., Parker, P. A., Woodall, W. H., and Hawkins, D. M. (2007), A Change Point Method for Linear Profile Data. *Quality & Reliability Engineering International*, 23(2), 247-268.
- Marshall, J. B., Spitzner, D. J., and Woodall, W. H. (2007), Use of the Local Knox Statistic for the Prospective Monitoring of Disease Occurrences in Space and Time. *Statistics in Medicine* 26 (7), 1579-1593.
- Parker, P.A., Kowalski, S.M., and Vining, G.G. (2007). "Unbalanced and Minimal Point Equivalent Estimation Second-Order Split-Plot Designs," *Journal of Quality Technology*, 39, pp. 376-388.
- Reynolds, M.R., Jr. and Kim, K. (2007) Multivariate Control Charts for Monitoring the Process Mean and Variability Using Sequential Sampling. *Sequential Analysis*, 26, 283-315. (Invited paper in the special volume of *Sequential Analysis* in honor of Walter Shewhart).
- Reynolds, M. R. Jr. and Stoumbos, Z. G. (2007). Variable Sampling Rate Control Charts, in *Encyclopedia of Statistics in Quality and Reliability*, Ruggeri, F., Kennett, R. and Faltin, F. W. (eds). John Wiley & Sons Ltd, Chichester, UK, pp. 2045-2050.
- Runger, G. C., Barton, R. R., Del Castillo, E., and Woodall, W. H. (2007), Optimal Monitoring of Multivariate Data for Fault Detection. *Journal of Quality Technology*, 39(2), 159-172.
- Stock, K.F., I. Hoeschele and O. Distl (2007) Bayesian estimation of genetic parameters for multivariate threshold and continuous phenotypes and molecular genetic data in simulated horse populations using Gibbs sampling. *BMC Genetics* 8:19.
- Stock, K.F., I. Hoeschele and O. Distl (2007) Estimation of Genetic parameters and prediction of breeding values for multivariate threshold and continuous data in a simulated horse population using Gibbs sampling and residual maximum likelihood. *J. Anim. Breed. Genet.* 124:308-319.
- Terrell, G. R. (2007) Finite element methods for density estimation. *Proceedings of the Statistical Computing Section, Joint Statistical Meetings*, Seattle, Washington
- Tyler, B.M., R.J.Y. Jiang, L. Zhou, S. Tripathy, D. Dou, T. Torto-Alalibo, H. Li, Y. Mao, B. Liu, M. Vega-Sanchez, S.X. Mideros, R. Hanlon, B.M. Smith, K. Krampis, K. Ye, S. St. Martin, A.E. Dorrance, I. Hoeschele and M.A. Saghai Maroof (2007) Functional genomics and bioinformatics of the Phytophthora-soybean interaction. *Stadler Genetics Symposium* book.
- Williams, J.D., J.B. Birch, W.H. Woodall, and N. M. Ferry. 2007. Statistical Monitoring of Heteroscedastic Dose-Response Profiles from High-throughput Screening. *Journal of Agricultural, Biological and Environmental Statistics*. 12, 2, pp 216-235.

Willis A. Jensen, Jeffrey B. Birch, and William H. Woodall. 2007. High Breakdown Estimation Methods for Phase I Multivariate Control Charts. *Quality and Reliability Engineering International* 23 (5), pp. 615-629.

Williams, J. D., Woodall, W. H., and Birch, J.B. 2007. Statistical Monitoring of Nonlinear Products and Processes Quality Profiles. *Quality and Reliability Engineering International* . 23(7), pp 925-941.

Woodall, W. H. (2007), Current Research in Profile Monitoring. *Revista Produção* 17(3), 420-425. (Invited paper for the Brazilian industrial engineering organization)

#### Papers In Press:

Bao, L and I. Hoeschele (2008) Quality assessment of a very large microarray experiment. *Technometrics* (in press).

Fraker, S. E., Woodall, W. H., and Mousavi, S. (2008), Performance Metrics for Surveillance Schemes. To appear in *Quality Engineering*.

Fraker, S. E., Woodall, W. H., and Burkom, H. S. (2008), A Note on the Poisson Likelihood Ratio Test Statistic for Kulldorff's Scan Methods. To appear in *Communications in Statistics – Theory and Methods*.

Hagerthey, S., Newman, S., Rutchey, K., Godin, J. and Smith, EP. The regime shift boundary in a subtropical wetland: Establishing ecological thresholds to low-level phosphorus enrichment. To appear in *Ecology*

Jensen, Willis A., Jeffrey B. Birch, and William H. Woodall. Monitoring Correlation within Linear Profiles Using Mixed Models. *Journal of Quality Technology* to appear in 2008.

Jensen, Willis A. and Jeffrey B. Birch. Profile Monitoring Via Nonlinear Mixed Models. *Journal of Quality Technology* to appear in 2008.

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2008), Monitoring Correlation within Linear Profiles Using Mixed Models. To appear in *Journal of Quality Technology*.

Jin, B. and Morgan, J. P. (2007). Optimal saturated block designs when errors are correlated. *Journal of Statistical Planning and Inference*, to appear.

Jensen, W. A., Bryce, G. R., and Reynolds, M.R., Jr. (2008). Design Issues for Adaptive Control Charts. To appear in *Quality and Reliability Engineering International*.

Joner, M. D., Jr., Woodall, W. H., and Reynolds, M.R., Jr. (2008), Detecting a Rate Increase Using a Bernoulli Scan Statistic. To appear in *Statistics in Medicine*.

Lipkovich, I., Smith, E.P., Ye, K. Detecting pattern in biological stressor response. *Ecological and Environmental Statistics*.

Liu, Y., Kim, I., and H. Zhao (2007). Protein interaction prediction from diverse sources, *Drug Discovery Today*, in press.

Liu, B., A. de la Fuente and I. Hoeschele (2008) Gene network inference via structural equation modeling in genetical genomics experiments. *Genetics* 178 (to appear in the March '08 issue and chosen as the issue's highlight article).

Mahmoud, M. A., Woodall, W. H., and Davis, R. E. (2008), Performance Comparison of Some Likelihood Ratio-Based Statistical Surveillance Schemes. To appear in *Journal of Applied Statistics*.

Mohammed, M. A., Worthington, P., and Woodall, W. H. (2008), Tutorial Notes on How to Plot Some Basic Control Charts for Health Care Practitioners. To appear in *Quality and Safety in Health Care*.

Morgan, J. P. and Parvu, V. (2007). Most robust BIBDs. *Statistica Sinica*, to appear.

Morgan, J. P. (2008). Latin squares and related designs. A chapter in *Encyclopedia of Statistics in Quality and Reliability*, to appear.

Park, C. and Reynolds, M.R., Jr. An Economic Design of an Integrated Process Control Procedure with Repeated Adjustments and EWMA Monitoring. To appear in *Journal of the Korean Statistical Society*.

Pickle, S. M., T. J. Robinson, J. B. Birch, and C. Anderson-Cook. A Semi-Parametric Approach to Robust Parameter Design. *Journal of Statistical Planning and Inference* to appear in 2008.

Reynolds, M.R., Jr. and Stoumbos, Z.G. (2008). Combinations of Multivariate Shewhart and MEMWA Control Charts for Monitoring the Mean Vector and Covariance Matrix. To appear in *Journal of Quality Technology*.

Rose, K.A., Roth, B.M., Smith, E.P. Skill assessment of spatial maps. To appear in *Journal of Marine Science*

Sego, L. H., Woodall, W. H., and Reynolds, M. R., Jr. (2008), A Comparison of Surveillance Methods for Small Incidence Rates. To appear in *Statistics in Medicine*.

Wang Y. and I. Kim (2007). Profiling user behavior for intrusion detection using item response modeling, *Computer Security Journal*, in press.

Woodall, W. H. and Borror, C. M. (2008), Some Relationships between Gage R&R Criteria. To appear in *Quality and Reliability Engineering International*.

Woodall, W. H., Marshall, J. B., Joner, M. D., Jr., Fraker, S. E., and Abdel-Salam, A. G. (2008), On the Use and Evaluation of Prospective Scan Methods in Health-Related Surveillance. To appear in the *Journal of the Royal Statistical Society, Series A*.

Vining, G.G. (2008). "Geoff Vining's Discussion of 'Must a Process Be in Statistical Control before Conducting Designed Experiments,'" *Quality Engineering*, to appear.

Vining, G.G. and Kowalski, S.M. (2008). "Exact Inference for Response Surface Designs within a Split-Plot Structure," *Journal of Quality Technology*, to appear.

Zhang, T. Thieling, S. Prins, M. Hudy, EP Smith. Model-Based Clustering in a Brook Trout Classification Study within the Eastern United States. *Transactions of the American Fisheries Society*

#### Presentations at Scientific/Professional Conferences:

Birch, J. "pRIPPLE: a Parallel Version of a Polynomial-time Piecewise Linear Estimation Algorithm." With M. A. Iyer\*, L. T. Watson. Presented at the Proc. 2007 Spring Simulation Multiconf., Business and Industry Symp.

Birch, J. "An Improved Genetic Algorithm Using a Directional Search." With Wen Wan\*. Presented at the Spring Eastern Region meeting of The Biometric Society, 2007, Atlanta, GA.

Birch, J. "Model Robust Regression: Development and Applications." With James Mays. Presented at the Southern Regional Council on Statistics, 2007 Summer Research Conference. Richmond, VA. Invited presentation.

Birch, J. "Optimal Smoothing and Mixing in Model Robust Regression." With B. Alden Starnes\*. Presented at the Southern Regional Council on Statistics, 2007 Summer Research Conference. Richmond, VA.

Birch, J. "Model Robust Regression for Instrument Calibration." With James Mays\*. Presented at the Southern Regional Council on Statistics, 2007 Summer Research Conference. Richmond, VA.

Birch, J. "Monitoring Correlation within Nonlinear Profiles Using Mixed Models" with Willis A. Jensen\*. Joint Statistical Meetings. August, 2007. Salt Lake City, UT. Invited Presentation.

Birch, J. "Monitoring Correlation within Nonlinear Profiles Using Mixed Models" with Willis A. Jensen. Birch\*. National INFORMS Conference. November, 2007. Seattle. Invited Presentation.

Birch, J. "On Detecting Stabilizing or Divergent Selection Using Patterns of Variation at SNP Loci" Joint Statistical Meeting, Salt Lake City, July, 2007

Carlson, J., Lemons, M., (2007), The Role of Family Support in Female Students' Transitions to Graduate School. Presented at the 2008 Virginia Tech SOESA Conference.

Du, P. "Penalized likelihood frailty model with smooth baseline hazard function". *The Tenth Meeting of New Researchers in Statistics and Probability*, Salt Lake City, Utah, July 2007.

Du, P. "Nonparametric smoothing spline model for gap time hazard function in recurrent event data". *Department of Statistics, George Mason University, Fairfax*, November 2007.

Du, P. Topic contributed talk: "Penalized likelihood frailty model with smooth baseline hazard function". *Joint Statistical Meeting 2007*, Salt Lake City, Utah, August 2007.

Du, P. Contributed talk: "Smoothing spline frailty model". *The 85<sup>th</sup> Annual Meeting of The Virginia Academy of Science*, Harrisonburg, Virginia, May 2007.

Du, P. Contributed talk (session chair): "Smoothing spline frailty model". *The Third Erich L. Lehmann Symposium*, Houston, Texas, May 2007.

Du, P. Contributed talk: "Frailty model with spline estimated baseline hazard function". *The 2007 ENAR Spring Meeting*, Atlanta, Georgia, March 2007.

Kim, I. Interface 2007 Conference, 39th Symposium on the Interface: Computing Science and Statistics 2007 (Invited talk)

Leman, S. Joint Statistical Meetings, in Salt Lake City, UT: The Evolutionary Forest Algorithm

Leman, S. Network Dynamics and Simulation Science laboratory (NDSSL): Modeling the Spread of Infectious Disease Using Genetic Information Within a Marked Branching Process.

Leman, S. Institute For Operations Research And The Sciences (INFORMS) at Virginia Tech: The Mutiset Sampler

Lemons, M., (2007), Analyzing Student Attitudinal Levels of Undergraduate Engineering Majors. Presented at the 2008 Virginia Tech SOESA Conference.

Terrell, G. Kim Parsimony and POP Modeling. Department of Statistics Colloquium, VPI&SU, April 12, 2007.

Terrell, G. Elementary Chis Squared. INFORMS Seminar, VPI&SU, September 19, 2007.

Terrell, G. Co-residual Analysis. Department of Statistics Colloquium, VPI&SU, November 29, 2007.

*International Invited*

Woodall, W. “Research Issues and Ideas on Health-Related Surveillance”, presented at the IX<sup>th</sup> International Workshop on Intelligent Statistical Quality Control, Beijing, China, September 2007.

Vining, G.G. XXII Foro Nacional de Estadística y XI Seminario de Estadística Aplicada del IASI (2007) – Jurica, Querétaro, México. The Future of the Design and Analysis of Industrial Experiments.

Vining, G.G. International Symposium on Business and Industrial Statistics (2007) – Ponta Delgado, Azores. Weighting Issues for Comparing Split-Plot Experimental Designs.

Vining, G.G. Joint European Network for Business and Industrial Statistics – Design of Industrial Experiments Conference (2007) – Turin, Italy, “Comments on the Use of Computer Experiments for Robust Design”.

Reynolds, M.R., Jr. and Stoumbos, Z. G. (2007). Multivariate Monitoring of the Process Mean and Variability Using Combinations of Shewhart and MEWMA Control Charts. Invited presentation at the IX<sup>th</sup> International Workshop on Intelligent Statistical Quality Control, Beijing, China.

*Invited National Meetings*

Morgan, J.P. Optimal Experimentation in Two Blocks. International Conference on Advances in Interdisciplinary Statistics and Combinatorics. Greensboro, October 13, 2007.

Morgan, J.P. Multi-criteria optimality in discrete design. Design and Analysis of Experiments DAE 2007. Mobile, November 2, 2007.

Vining, G.G. Designs and Analysis of Experiments (2007) – Memphis, Tennessee. Weighting Issues for Comparing Split-Plot Experimental Designs.

Vining, G.G. Q&P Research Conference (2007) – Santa Fe, New Mexico. The Beauty of Classical Designs

Reynolds, M.R., Jr. (2007). Control Charts for Multivariate Monitoring the Mean and Variability of Multivariate Processes with Sequential Sampling. Invited Plenary Lecture at the First International Workshop in Sequential Methodologies, Auburn, Alabama.

**Technical Reports:**

Williams, James D., Woodall, William H., Birch, Jeffrey B. 2007. Statistical Monitoring of Nonlinear Product and Process Quality Profiles. Virginia Tech Department of Statistics, Technical Report #07-2, 39 pages.  
[http://www.stat.org.vt.edu/dept/web-e/tech\\_reports/TechReport07-2.pdf](http://www.stat.org.vt.edu/dept/web-e/tech_reports/TechReport07-2.pdf)

Waterman, Megan J., Birch, Jeffrey B., Schabengerger, Oliver. 2007. Linear Mixed Model Robust Regression. Virginia Tech Department of Statistics, Technical Report #07-3, 39 pages.  
[http://www.stat.org.vt.edu/dept/web-e/tech\\_reports/TechReport07-3.pdf](http://www.stat.org.vt.edu/dept/web-e/tech_reports/TechReport07-3.pdf)

Fall 2007 Graduates

Chaturvedi, Garima	MS
Chen, Yunjie	MS
Gray, Travis	MS
Han, Chao	MS

Howard, Joshua	MS
Howard, Lauren	MS
Huang, Wandi	MS
Lin, Kuan-Chin	MS
Ryan, Anne	MS
Szarka, John	MS
Wang, Heng	MS
Wang, Sai	MS
Wang, Yadan	MS
Wentz, Pamela	MS
Zhang, Huan	MS
Zhang, Yanmei	MS
Zhao, Xin	MS

Fraker, Shannon	PHD
Lee, Mi Hyun	PHD
Li, Zhengrong	PHD
Love, Kimberly	PHD
Wan, Wen	PHD

Spring 2007 Graduates

Graham, Jeffrey	MS
Packard, Kevin	MS
Potts, Jonathan	MS
Wang, Lu	MS
Joner, Michael	PHD