

NCCC-170 / USSES Joint Annual Meeting June

23 – 24, 2016

University of Maryland, College Park, MD

Technical Program

Attendees: William Bridges, Jerry Davis, Xuelin Luo, Bahram Momen

Thursday, June 23, 2016

8:00 – 8:30 AM

Registration and Check-In

8:30 – 8:45 AM

Welcome and Introduction

Dr. Bill Bowerman, Chair Department of Environmental Science and Technology University of Maryland

8:45 – 10:15 AM

“Predictive Modeling in an $n < p$ World”

Elizabeth Claasen, JMP Analytical Development

Abstract: There are many times where due to budget limitations, the number of parameters or factors of interest in an experiment outnumber the number of runs available. Supersaturated designs from the design of experiments arena and genomics data both exhibit the $n < p$ problem. Traditional least squares analysis is not sufficient to the task of modeling these data. Modern variable selection methods such as LASSO or Elastic Net, Partial Least Squares techniques, and partition methods such as the Bootstrap Forest are needed. In this presentation we will show some existing and new in JMP 13 techniques for working with these types of data.

10:15 – 10:30 AM

Morning Break

10:30 – 11:00 AM

“Where is Research Going: Long Term Agro-Ecosystem Research Initiative”

Sara Duke, USDA ARS

Abstract: The LTAR (Long Term Agro-ecosystem Research) initiative was announced in 2014 and initial funding was provided for 10 of 18 sites to participate in a nationwide network. These studies are expected to be on going for 30+ years. The goal is to evaluate alternative agricultural production systems that increase the sustainability and productivity of soils and food production at a field scale while also increasing ecosystem services. The common experiment is a simple two treatment comparison of Business-As-Usual (BAU) for each location and an alternative or aspirational (ASP) cropping system. The major statistical challenge is that of replication within each location. It was decided that there will be field scale as well as plot scale experiments conducted at each location but the field scale replication was limited to one for several reasons. A major challenge in this program is that many locations do not own the land on which they conduct research hence the security for long-term research is limited. The statistical challenges that this network will face is how to integrate research un-replicated field level results from 18 locations across many different ecological, temporal and spatial gradients. This LTAR project is the cornerstone for the new ARS Grand Challenge proposed by ARS in 2015 and also the natural resources component of the Big Data Initiative. The data management and statistical support for this project will be challenging.

11:00 – 12:00 PM

“Evaluating and Improving Eagle Monitoring and Cotton Variety Testing”

*W. Bowerman, H. Pittman, L. Fuentes, Department of Environmental Science and Technology
University of Maryland*

*K. F. Leith, M. C. Wierda, Department of Forestry and Natural Resources, Clemson University
F. Bourland, University of Arkansas – NEREC*

D. Jones, Cotton Incorporated

W. Bridges, Clemson University

Abstract: The objective of this presentation is to get comments and ideas about approaches we have taken to analyzing data sets from two interesting research programs.

The first program is eagle population monitoring. Due to the decline in bald eagle numbers (*Haliaeetus leucocephalus*) beginning in the 1950s, the National Audubon Society designed the Continental Bald Eagle Survey and Michigan joined in and began a statewide survey during the inaugural year of 1961. Since 1961 bald eagle reproduction has been counted statewide using two aerial surveys. A third survey occurs for a subset of the nest sites by climbing to the nests and banding the nestlings between 6-9 weeks of age. These surveys produce measures on nesting success and productivity. In 1999, the State of Michigan initiated the Wildlife Monitoring Program using concentrations of bioaccumulative compounds of concern (BCCs) to indicate the health of the state's aquatic environment. The study used a five-year rotation of

National Pollution Discharge Elimination System (NPDES) watersheds to evaluate the spatial and temporal trends in BCCs. We will discuss several analysis approaches we have tried to determine temporal trends and spatial differences (at scales ranging from State-wide to individual watershed levels) in nesting success, productivity, and BCCs. We will also discuss design ideas that have been used to decrease the number of samples required in the surveys.

The second program is long term cotton breeding effort lead by Fred Bourland of the University of Arkansas. Two specific cotton strain testing data sets are being evaluated. In the first data set, six strain tests (of 20 strains each) were tested for nine years with four locations per year. The objectives for this data set were to evaluate any location by variety interaction, and then use some design methods to determine if a reduced set of locations could be used in future testing of specific traits. In the second data set, 20 strains were tested for nine years, but now there are up to nine locations per year (and the strains and locations are not consistent from year to year), and this data set focuses on different traits. The objectives for this data set were similar to those in the first data set, but the unbalanced data issues lead to some interesting approaches.

12:00 – 1:00 PM

Lunch

1:00 – 1:30 PM

“Quantify remote sensing phenology using a nonlinear mixed model”

Yu Mo, Bahram Momen, & Michael S. Kearney, University of Maryland

Abstract: Satellites monitor global vegetation on a regular base, gathering unparalleled records that enable the investigation of temporal (among years) and spatial (across latitude and longitude) variations of plant phenologies. However, the current literature does not provide statistical test for comparisons of different vegetation units and of different phenological models. This work used a nonlinear mixed model that enables such comparisons to study the phenology of coastal marshes in Louisiana under different climatic conditions. We compared the phenological parameters of four marsh types (i.e. the freshwater, intermediate, brackish, and saline marshes) and tested the goodness-of-fit of three phenological models (i.e. the stepwise logistic, Gaussian, and stepwise Gaussian functions). The Gaussian function was found to be the best phenological model (pseudo R^2 0.56–0.85) in 2007 (a normal year), showing that: (1) the Normalized Difference Vegetation Index (NDVI) of the marshes peaked in late July to mid-August; (2) freshwater marshes had the highest peak NDVI, followed by intermediate, brackish, and saline marshes; and (3) saline marshes had the longest growth duration, followed by brackish, and then intermediate and freshwater marshes. The stepwise Gaussian function was found to be the best model (pseudo R^2 0.83–0.94) in 2005 (a hurricane year), showing a shortening in growth duration of all marshes by approximately half. This work shows that the nonlinear mixed model is able to quantify remote sensing phenology of Louisiana coastal

marshes, helping to analyze and predict the marshes' responses to current and future climatic stresses.

1:30 – 2:15 PM

“Rethinking Causal Inference from Observational Data: a structural equation modeling approach”

Nora M. Bello^{1,2}, Abigail Jager¹, Elva Cha², Mike Sanderson², Natalia Cernicchiaro² and David Renter²*

¹Department of Statistics, ²Department of Diagnostic Medicine/Pathobiology and Center for Outcomes Research and Education. Kansas State University, Manhattan, KS.

Abstract: Basic principles of experimental design and analysis, as developed by Sir Ronald A. Fisher, point to randomization as the gold standard and sole gateway for valid causal inference. Yet randomized experiments are not always feasible, appropriate or even reflective of realistic conditions. In this context, observational studies become not only inevitable necessities but actually, valuable inferential assets. In agriculture, a wealth of observational data consisting of livestock production, health, well-being and product quality is routinely collected by commercial operations for internal decision making. Such observational data is traditionally used to assess non-directional linear associations (i.e. correlations) between variables, though inferring causality from such correlations is problematic due to potential biases from confounding effects. Nevertheless, non-zero correlations are arguably necessary, though not sufficient, conditions for causality. One may then ask: is there anything to be learnt about causal relationships on the basis of observational data? Here we explore the concept of directed-separation to assess patterns of partial correlation and conditional independence in observational data, thereby allowing insight into plausible causal pathways between outcome variables. Further, we discuss the technique of structural equation models embedded within a multivariate mixed modeling framework as a methodological approach for the functional characterization of causal networks between outcome variables in animal agriculture. This has important implications for our understanding and management of multivariate systems, particularly in response to external interventions with potential repercussions extending beyond individual outcomes to the system as a whole.

2:15 – 2:30 PM

Afternoon Break

2:30 – 4:30 PM

“P-value Controversy – in preparation for JSM”

Walt Stroup, University of Nebraska

Abstract: In Spring, 2015, the *Journal of Basic and Applied Social Psychology* banned P -values in their publications. This prompted the American Statistical Association to convene a Roundtable Lunch at JSM 2015, at which I was a participant. Somewhat independently, ASA prepared a policy statement, which was formally posted in Spring, 2016. As a follow-up, an invited panel discussion session will be held at JSM 2016. I was invited because I was the only “Ag and Natural Resources” person at the 2015 Roundtable. So I’m taking the approach that when I speak at JSM 2016, I’m speaking for NCCC-170, and for statistics at land grant universities and ARS. The discussion questions participants have been given go beyond P -values per se. They concern the culture of collaboration, how early (or not) statisticians are involved in research projects, the extent to which analysis – specifically the technical aspects of analysis vs. the scientific big picture – is over-emphasized, whereas scientific thinking, planning and design are under-emphasized, and the ill-effects of the “democratization” of statistics, e.g. via black-box software. These questions raise several issues, from the integration of scientific thinking and statistical thinking to the content of stat methods courses, both for stat and for non-stat majors, to the promotion and tenure culture at land grant universities. I’ll give my take on all of these, and invite the NCCC-170 group to brainstorm about our potential role(s) in addressing these issues.

Matt Kramer, USDA-ARS

Abstract: The use of p -values has become firmly ingrained in many of the sciences, including agriculture and natural resources. They offer a simple, though not necessarily correct, way of deciding whether results are important, and require additional information to be meaningfully interpreted. One tests a 'null hypothesis', which is usually not the hypothesis of interest. An alternative is suggested for situations where one might use t -tests (easily extended to ANOVA type designs). In this framework, one first determines a 'key difference' (with help from a subject matter expert---requiring collaboration at the design stage), and then establishes the sample size necessary to demonstrate that an observed difference in means is at least as large as this key difference with a pre-specified probability. Once data are collected, one calculates the probability that the actual observed difference in means is at least as large as the key difference. In this framework, one directly calculates the probability that the hypothesis of interest is true, more interpretable for researchers in the biological sciences.

4:30 – 5:00 PM

Group Discussion/Adjourn

**NCCC-170 / USSES Joint Annual Meeting June
23 – 24, 2016
University of Maryland, College Park, MD
Business Meeting**

Friday, June 24, 2016

9:00 – 10:00 AM

Business Meeting

10:00 – 11:00 AM

Group Discussion Topics

How to train and advance Applied Statisticians in Professional and Academic settings.

Other topics

Reports:

Clemson

We successfully hired a candidate in the Applied Statistics subfaculty last summer. The faculty member has not yet engaged in collaborative research. The university reorganized effective July 1, 2016 so that the Department of Mathematical Sciences was moved to the College of Science from the College of Engineering and Sciences. There will be a 3% cost of living raise effective in September. The department head is considering hiring at least 2 faculty in the Applied Statistics subfaculty.

University of Georgia

Dr. Lew Hunnicutt was hired as the Assistant Provost and Director of the Griffin Campus. Dr. Hunnicutt is responsible for guiding the research and extension programs as well as the multidiscipline academic programs. Dr. Sam Pardue was named Dean and Director of the College of Agricultural and Environmental Sciences. Dr. Pardue was formerly associate dean and director of academic programs at N. C. State's College of Agriculture and Life Sciences. The state legislature allocated a 3 percent cost of living raise effective July 1.

University of Maryland

Bahram Momen is the UMD representative. Applied statisticians at UMD still experience difficulties in getting tenure or being promoted because their consultation efforts and publications are not perceived as scientific work. The Biometrics program at UMD continues to weaken.

Presentations:

Villena OC, Landa E, **Momen B**, Leishnam PT, Sullivan JH. **2015**. The effects of UV radiation on tire deterioration and pollutant release on competition between *Aedes albopictus* and *Culex pipiens* mosquitoes. *Ecological Society of America Annual Meeting*. Baltimore, MD.

Yu M, **Momen B** Kearney M. **2016**. Modeling remote sensing phenology of coastal marshes using a nonlinear mixed model. *USSES/NCCC-170 Joint Annual Meeting*. College Park, MD.

Publications:

Momen B, Behling SJ, Sullivan JH, Lawrence GB. **2015**. Photosynthetic and growth response of sugar maple (*Acer saccharum*, Marsh) mature trees and seedlings to calcium, magnesium, and nitrogen additions in the Catskill Mountains, NY, USA. *PLOS ONE*, 10:1-14. DOI:10.1371/journal.pone.0136148.

Khan S, Jackson RT, **Momen B**. **2016**. Impact of acculturation and its factors on metabolic syndrome on South Asian American adults in community health centers. *Journal of Public Health*. DOI: 10.1007/s10389-016-0726-7.

Khan S, Jackson RT, **Momen B. In Press.** The relationship between diet quality and acculturation of immigrated South Asian American adults and their association with metabolic syndrome. *PLOS ONE*.

Cartwright J, Dzantor EK, **Momen B. In Press.** Soil microbial community profiles and diversity in limestone cedar glades. *CATENA*.

Olsen VBK, **Momen B,** Langsdale S, Galloway G, Link E, Brubaker K, Matthias R, Hill R. **In Press.** An approach for improving flood risk communication using realistic interactive visualization. *Journal of Flood Risk Assessment*.

The USSES meeting for 2017 has not been planned. NCCC-170 plans to meet in Elk River, MN.

Attendees and email addresses:

William Bridges wbrdgs@clemsun.edu

Jerry Davis jwd@uga.edu

Xuelin Luo xuelin@uga.edu

Bahram Momen bmomen@umd.edu