

IFAS Department of Statistics' Research Roadmap

IFAS Department of Statistics' Research Roadmap

On October 10, 2008, the IFAS Department of Statistics held a retreat with the faculty and consulting statisticians in attendance. Charles Goretsky served as the facilitator. The group participated in several exercises that led to the development of a vision statement:

To become internationally recognized as leaders of statistical methodologies and applications in all areas of relevance to IFAS research.

As a consequence of the process, the ten Research Roadmap questions were also addressed. The responses, which represent a consensus opinion, are provided below.

1. What areas of research is your department best known for by others?

Historically, our department has been well known for strengths in linear mixed models and design of experiments. Newer faculty are known for their expertise in statistical methods applicable in ecological and environmental studies, spatial analysis, statistical modeling, time series and statistical genetics, while the reputation for mixed models and design of experiments has been maintained.

2. What truly are your current areas of excellence, your research strengths currently in the discipline including your entire statewide faculty?

In terms of external perception, our department is ranked 9th nationally by *U.S. News and World Report*. Our faculty/staff are approachable and knowledgeable in their subject areas. The faculty publish in well respected journals, with papers appearing both in the statistical literature and in subject-matter specific journals.

The research strengths of our department lie in two directions due to our unique role within IFAS. The faculty are research statisticians, and both faculty and our MS consulting statisticians are collaborative scientists working with other researchers within and without IFAS. We address both of these in this response.

In statistical research, topical areas of strength include statistical modeling, including linear, non-linear mixed, and generalized linear mixed models; statistical genetics and breeding analysis; spatial analysis; methods relevant to ecology and the environment, and human-environment interaction. We also have faculty who work in developing statistical methodology for research in animal movement and use of space. Finally, our unit has the benefit of affiliated faculty who collaborate with departmental faculty and contribute substantially to the statistical literature. The collaborations with IFAS researchers and with scientists in state agencies provide a rich source of important statistical problems that feed further statistical research.

In collaborative research, we have both a collaborative and a consulting role. On the collaborative side, we have strengths in applying the newest statistical methodology in new ways using both existing statistical methods and new variants of those methods. Through consulting,

we provide the statistical expertise to ensure that the research conducted in IFAS is based on the best experimental approach and analysis.

Our consulting service has grown and become better known throughout UF and the state for its ability to do a comprehensive statistical plan and to provide broad experience on several theoretical and applied statistical topics.

Strong collaborative research and consulting require scientists with whom statisticians can collaborate and a group of statisticians who can effectively interact with those scientists. All aspects of our research work, from consulting to statistical research, require an infrastructure that supports these endeavors. We believe that we have these strengths as research strengths, as we describe next.

Faculty and staff provide software training, workshops and courses oriented on applied and/or theoretical topics to graduate students, faculty, and staff as well as individuals outside of the University. We also teach non-major graduate courses, such as STA 6166 and STA 6167. They have increased our visibility on campus and throughout the state and should eventually lead to more collaborative research opportunities.

Our faculty and staff have a positive attitude and diverse interests and participate on collaborative efforts with other units within the University as well as organizations outside the University.

Potential for actual and future development is supported by our hardware and software infrastructure. The recently added cluster supports research, including work on statistical genetics and spatial analysis. The addition of a polycom to our training facilities allows us to reach a larger audience for academic courses, short courses and software workshops. Our department supports a variety of statistical software, such as SAS, JMP and R.

In summary, the statistical research being conducted by IFAS Statistics faculty and collaborative research and consulting between IFAS Statistics and scientists within UF and in state agencies have given the unit a presence within the university, throughout the state, and nationally.

3. What are **your research weaknesses**, gaps that are not covered now, yet you deem essential for the future directions and scientific impact of your discipline?

We are a small unit consisting of only four faculty, one of which had no research appointment and another who spends 40% time outside of IFAS. Thus, we are not able to meet all of IFAS's collaborative research needs. Although collectively we have a wide range of statistical expertise, we lack a depth of experience in some critical areas of statistical methodology. The most critical areas that are missing include Bayesian inference and quantitative genetics and bioinformatics, followed by statistical methodologies for social/human dimensions research (design and analysis of sampling surveys) and multivariate methods.

Bayesian models have come into prominence during the past few years as the software and hardware needed to fit them have become more widely available. Bayesian methods are

becoming increasingly popular in a number of application areas, including ecology and spatial-temporal modeling. This approach to modeling is very different from the frequentist approach. All of the current IFAS faculty have strong training in frequentist methods and have only limited experience in the Bayesian realm. With current software, Bayesian models are easily fit, but they can also be easily misapplied. We need a Bayesian statistician who can ensure that the models are being properly developed and applied by the subject-matter scientists as well as develop new approaches to non-standard problems.

Quantitative genetics and bioinformatics are research areas in the intersection of statistics and other biological disciplines, and they form a crucial direction in statistics as an interdisciplinary science. Departments throughout IFAS are tapping into genetic markers, microarrays, and other genetic data structures in their research, and our unit lacks sufficient expertise and resources to address the increasing demand. Increased funding from national agencies including the NIH, NSF, and USDA further indicate a continued trend towards these research areas. The increased interest on the use of molecular tools brings new challenges on statistical genetics and bioinformatics that need to be addressed with multidisciplinary teams consisting on statisticians, breeders, molecular geneticists and computer scientists.

Several IFAS departments and schools have faculty who conduct research on social or human dimensions, such as the impact of environmental modifications on farming practices or the effect of diet modifications on health. Our unit currently lacks in-depth statistical expertise in this challenging area. As a consequence, collaborative research in these areas is not being developed, and we are unable to help make the needed advances in survey design and related research for such applications. Establishing connections with social science researchers in areas such as anthropology, sociology, human-environment interaction, survey sampling and design, will bring an added dimension to our department and provide new opportunities for research topics on theoretical and applied statistical methods.

Multivariate statistics is also important due to the rapidly increasing capability of collecting large and complex data more efficiently than in the past. As a result, the ability to model multiple responses simultaneously is becoming more popular, especially with the increase in computing capabilities that has occurred in the last ten years.

4. Where is your discipline moving to in the future? What are the major trends in your field of science? What do the very best departments in the country look like in this discipline? Think of these as **your research opportunities**.

The field of Statistics has been growing exponentially in the last 20 years, ever since the advent of increased computing capabilities. As a result, the areas toward which the field is moving that also intersect with the interests of IFAS are many but include the following (no specific order):

- Computationally intensive methods
- Analysis or mining of massive datasets
- Statistical genetics
- Spatial-temporal modeling
- Multivariate analysis

- Bioinformatics
- Generalized linear mixed models
- Bayesian statistics
- Semi-parametric methods
- Model diagnostics
- Stochastic processes and models
- Nonlinear modeling

The major trends in statistics are in Bayesian statistical methods, methods for large datasets and computationally intensive analyses, spatial-temporal modeling, statistical genetics, and modeling of non-standard data. One example of the latter is the analyses of movement and use of space by animals with GPS collars.

The very best statistics departments in the US that we would like to be more like have a critical mass of strong statistical researchers; statisticians well positioned to work with subject-matter scientists; a pool of talented graduate students; a strong funding base; areas of statistical expertise that cut across disciplines; and a general appreciation of the role of statistics within the University community. A key is the integration of statistics throughout the research community. Examples include Stanford University where every tenure-track faculty member has an appointment both within statistics and within another department as well: Bradley Efron is a GCRC statistician; Paul Switzer has a joint appointment with Geography; etc. At North Carolina State University and Iowa State University, students in other disciplines tend to take more statistics classes than our students generally do. As a result, these statistics departments have a more productive and synergistic interaction with researchers in the applied sciences.

5. Aside from times of limited resources, what/who are the **major research threats** to fulfilling your vision?

Perhaps the primary threat to our research is the varied demands made on our departmental members. Teaching and routine consulting are time consuming, and methodological and collaborative research is often done on the time that is left. This is a threat to the careers of our faculty who must help advance statistical knowledge if they are to be recognized by their peers. It is also a threat to IFAS research because the field of statistics is rapidly changing and statisticians who are not helping to move the field forward rapidly become obsolete. The faculty is increasingly stressing collaborative research, but many IFAS researchers are unaware of this change in focus. They are reluctant to work with our consulting statisticians on routine analyses and do not think to work with a statistician in proposal development.

Because statistics is integral to most scientific research efforts, several units have begun to teach statistics courses and to provide consulting within their departments. This represents a threat both to the quality of the science emanating from IFAS as well as a threat to the research of the statistics unit. Although there are some exceptions, the courses taught outside of statistics are often of the “cookbook nature,” providing students with only a superficial understanding of the material. Consequently, the methods are often inappropriately applied and, even if the analysis is appropriate, the results are often misinterpreted. Further, as new methods become available,

the students lack the background on how to apply the new methods correctly. This impacts their ability to be effective collaborators with statisticians.

Statistical software is constantly being developed for specialized analyses. As they become popular in a small application area, researchers within that area begin using it, but often encounter problems when using the software for their data. They then come to the statistics unit thinking that we should be able to run any statistical package. When we are unaware of the package or its use, they begin to question the credibility of the unit. This is further exacerbated because many of these new statistical software packages lack documentation, have not been fully evaluated from a statistical perspective, and require significant investments in time to understand. Further, if a software package is good for a specialized application, we may not be aware of it and thus are unable to use it effectively in research applications.

The ready availability of statistical software provides an additional challenge to the quality of the science coming from IFAS and to our department's research. Researchers with little background in statistics assume that, if they learn to run the program, their analyses must be correct. Often the analyses conducted are inappropriate and, even if done correctly, the results may be misinterpreted. The research proposals are less competitive, and our department misses the opportunity for collaborative research.

Although the younger faculty tend to have a solid statistical foundation, a surprising proportion of researchers lack the statistical foundation to be able to collaborate effectively with statisticians. They often seek non-statistical input or want to simply "give the data to a statistician." Effective collaborative research requires enough statistical background on the part of the researcher and enough subject-matter background on the part of the statistician for effective communication.

With the diversity of IFAS research, it is difficult, if even possible, to have the expertise needed for all projects. IFAS Statistics loses credibility if a researcher has statistical needs that we are unable to meet. Surprisingly, some IFAS researchers do not seem to be aware of the existence or the services offered by IFAS Statistics. Both of these are threats to the opportunity for effective research.

6. What current areas of research in your department will need to be enhanced to be the leading department in your field? What new areas of research in your department will need to be added to be the leading department in your field? And which areas of your discipline are less likely to be essential in your department in 10 and 20 years?

Currently, our areas of emphasis relate to agriculture, ecology, environment, and genetics. To fully develop these areas, we need to have people trained specifically in Bayesian methods and statistical genetics at the molecular level. Because of their pervasiveness, additional expertise in multivariate methods, generalized linear mixed models, semiparametric models would be very helpful. Expertise in computationally-intensive methods, especially those important to the analysis of massive data sets, is increasingly important.

IFAS's move to strengthen social science research has led to a need for statistical expertise in that area. Specifically, we need to enhance our research in survey methodology and categorical multivariate methods.

Although the ability to design experiments will continue to be useful, research in this area is likely to be relatively unimportant in 10 to 20 years. Similarly, some of the basic analyses for fixed effects are well understood, and the need for further research will be questionable.

7. What are the cross-cutting research topics that need to be addressed through partnerships with other disciplines in UF, at our universities, or with other agencies? How would your department benefit from partnerships/interaction with other units?

With the broad variety of and interest in the natural resources of Florida, topics of relevance are environmental statistics, ecological statistics, spatial-temporal analysis, non-linear modeling, Bayesian analysis, and human-environment interaction.

The increased interest on the use of molecular tools brings new challenges in statistical genetics and bioinformatics that need to be addressed with multidisciplinary teams consisting on statisticians, breeders, molecular geneticists and computer scientists.

IFAS is building research strength in the social sciences. Expertise in the relevant statistical methods, such as survey sampling and design and multivariate categorical models, needs to be strengthened.

Interaction with other units will increase the awareness of IFAS Statistics within the university, will bring new collaborative partners, and will provide new research topics to develop applied statistical tools and methodology. This interaction will lead to new areas for statistical research on theoretical issues.

8. Knowing the faculty that you must have in place to accomplish your goals, what critical hires in order of importance in your discipline will be necessary to position your department as the leader in its discipline? Create short job descriptions.

Due to the small size of our unit, we believe it is best to diversify our knowledge base as opposed to building up our current research areas. Our research weaknesses precisely highlight the areas of expertise that our unit most greatly needs. The most critical needs are for experts in statistical genetics and Bayesian inference. Statistical researchers in methods for social statistics and multivariate analysis would not only enhance the research expertise of the department, but also allow us to branch out and collaborate with other IFAS departments.

Short Job Description: 50% teaching, 50% research. Ph.D. in statistics or related field. Statistical expertise in statistical genetics or Bayesian modeling. Conduct statistical research leading to refereed publication in statistical journals, collaborate effectively with IFAS researchers leading to joint publications in statistics and subject-matter journals, teach courses of relevance to IFAS and statistics graduate students.

9. Would reaching your research goals be helped by key research hires in other IFAS departments? Who? What? (Please contact those departments if you list them here.)

As younger faculty are hired, they often have a sufficient statistical foundation to become excellent collaborators. It would be helpful to have a college-level person who is able to help people seeking collaborations identify common interests.

10. Are there mechanisms of research administration that you see as needing to change to assist you in attaining your department's goals? How can FAES administration change and thereby help your department meet its goals?

The FAES administration has been proactive in promoting the use of statistics more broadly within the IFAS community. It takes time to change traditions, and sometimes it helps to enforce some changes. As an example, all NIH grant proposals with a statistical component must have a named statistician as a part of the research team at the time of submission. IFAS would benefit by requiring all proposals with a significant statistical component to have a named statistician as a part of the research team, preferably as a co-PI, at the time of submission. A less favored, but alternative model, would be to have statistical support as a specified line item in the budget of all proposals with a statistical component.

IFAS STATISTICS CONTEXT MAP

RESEARCH TRENDS

- Computationally-intensive methods
- Massive data sets
- Statistical genetics
- Impact of Ag on environment
- Spatial/temporal modeling
- Social environ/human aspects
- Multi-variate analysis
- Bio-informatics
- Generalized/GLMM
- Bayesian statistics
- Semi-parametrics
- Diagnostics
- Animal movement/stochastic models
- Non-linear models

POLITICAL CLIMATE

- Focus on Global Warming
- Agriculture budget (farm bill) issues
- Focus on Bio-fuels
- "Bright Futures" program in Florida
- State budget cuts
- Federal agency budget cuts

ECONOMIC CLIMATE

- Stock/financial market-related woes
- Adequate share of (current) state funding, grants, NOT for positions

OTHER TRENDS

- Better stat methods across sciences
- Doing own analysis w/out understanding
- Heightened need for stat expertise (but don't know it)
- Increased recognition for line-item statistician support in proposals (non-IFAS)
- Data collected/secondary data analysis

CUSTOMERS

- State & Fed Agencies (WMD, FWC, DEP, DOH, etc.)
- Faculty, students
- IFAS
- Commodity groups (industry, extension)
- Other scientists (journal readers, statisticians)
- Policy makers

UNCERTAINTIES

- "Pay as you go" funding
- Grants with user fees or with identified statistician

COMPETITION

- Cal-Berkeley
- Stanford
- Texas A&M
- NC State
- Cornell
- Purdue
- Iowa State

TECHNOLOGY FACTORS

- Storing data for the future
- Ease of data collection
- Wide variety of statistical packages and inconsistent results

CUSTOMER NEEDS

- Vetting of stats in research
- Consulting
- Collaboration with other scientists
- Training

SWOT Analysis – IFAS Statistics October 10, 2008

<p>Strengths</p> <ul style="list-style-type: none"> • National ranking, best in state • Great executive support • Quality of teaching graduate-level (6166/6167) <ul style="list-style-type: none"> • Exposure with other academic depts. • Consulting – Mehgan & James • Computation and software packages • Diverse research, interests, resources • Links with applications strong • Collaborative research excellent • Modeling • Mixed models/GLMM • Spatial analysis • Ag/ecology/environment/genetics • Ability to “plug-in” with other academic Departments • Willingness to “plug-in” with other departments • Short-courses & workshops • Positive attitudes, collegial, strong team • Design of experiments • Affiliate faculty (USGS, FWC) • User-friendly software available on cluster 	<p>Weaknesses</p> <ul style="list-style-type: none"> • Lack of \$/funding • Staff size • Faculty • (Lack of) Dedicated grant specialist • Bayesian analysis • Ability to interact with Stat Grad students • Awareness of offerings (outside of dept) • Multi-variate analysis • Molecular genetics • Survey design • Social sciences
<p>Opportunities</p> <ul style="list-style-type: none"> • Reverse trend of other departments hiring “stat” instructors • Increased concern for environmental issues • Journal requirements for improved statistical treatment • Statistical package software training • Local industry training (in conjunction with other departments) • Everglades and southern Water Management District (clean and restore) efforts • State-wide animal and plant environment (natural resources) efforts • Range of climates and gradients in state of Florida • NGO/commercial training • PES • Water • Genetic • Emerging pathogens • Institutes 	<p>Threats</p> <ul style="list-style-type: none"> • Budget cuts • Financial • Unawareness of customers – lack of understanding (of Stat offerings) • Statistical s/w packages not accurately creating results • Lack of required or basic training (in stats for other researchers) • Stats taught by non-Stats professors in other departments

Key Themes: Options/flexibility to work in different application areas. Teaching vs. Research vs. Collaboration mix. Quantitative marketing has strengths where Statistics has weaknesses. Resource limitations.

Headlines - "Statistics Today" Issue - 2018

IFAS – Internationally renowned and the place to go for:

1. Statistical expertise related to environmental, agricultural and ecological and genetic issues and applications.
2. New genetic statistical models and analyses, including
 - a. Science of nucleotides, DNA sequencing and breeding
 - b. Plants and animals
 - c. Populations genetics
 - d. Genetically-modified organisms
 - i. Transgenics
3. Social and human dimensions of IFAS
4. Training of both future scientists and consumers of statistics
5. Applied Statistics
6. A well-established statistical consulting Center
7. The latest and greatest in statistical software and computing
8. Approachable statisticians
9. Faculty knowledgeable with areas of application across IFAS

Vision Statement:

To become internationally recognized as leaders of statistical methodologies and applications in all areas of relevance to IFAS research.

Customer & Collaborator Quotes/Testimonials - "Statistics Today" Issue – 2018

“Advanced both the science of statistics and scientific areas of applications (in IFAS)”

“Made a dramatic impact on both the problem and improved management of natural resources in the state of Florida, nationally and internationally”

“More rapid development and accelerated gains in breeding”

“Great example of how to develop new statistical theory that can lead of advances in applied statistics (via new models for statistical design)”

“New designs yield greater efficiencies in research outcomes”

“Statistics consulting and research that is modern and comprehensive”

THEMES – CONTEXT MAPPING

State budget cut – impacts both grants and positions

(Moving from) #9 to #1 national ranking requires more positions

Funding increases will come from outside the state of Florida

Trying to be too much to too many

Vibrant science with rapid changes