ASA

AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

732 North Washington Street, Alexandria, VA 22314-1943
(703) 684-1221 ■ www.amstat.org ■ asainfo@amstat.org
f www.facebook.com/AmstatNews **◻** www.twitter.com/AmstatNews

October 23, 2017

James W. Hurrell
Director, National Center for Atmospheric Research
P.O. Box 3000
Boulder, CO 80307-3000

Dear Dr. Hurrell,

We write to express our deep concern for the elimination of NCAR's Institute for Mathematics Applied to the Geosciences (IMAGe) and its research groups, especially the Statistics and Data Science (SDS) group. While we understand the decision was a difficult one brought on by budget cuts, we believe it undermines NCAR's mission "to understand the behavior of the atmosphere and related Earth and geospace systems" and urge the decision be reversed. Indeed, the decision runs counter to the strong and growing consensus that scientific challenges—like better understanding of our climate—are best tackled by multi-disciplinary teams. The decision also seems incongruent with the National Science Foundation's Harnessing the Data Revolution initiative and the emergence of data science over the last several years.

In this letter, we focus on the importance of statistics to climate science and what statisticians bring to bear on this challenging and important issue. We will also speak to the important role NCAR has played in engaging statisticians and training them to work in climate and atmospheric science.

As the "science of learning from data, and of measuring, controlling, and communicating uncertainty,"[1] statistics is core to atmospheric science. Statistics' role in climate science is discussed in the enclosed 2014 whitepaper[2], issued by the American Statistical Association's Advisory Committee on Climate Change Policy (ACCCP), a paragraph from which follows:

> Data are fundamental to all of science. Data enhance scientific theories and their statistical analysis suggests new avenues of research and data collection. Climate science is no exception. Earth's climate system is complex,

---

[1] "Why Statistics?", Marie Davidian and Thomas Louis, *Science*, **336,** April 6 2012, p. 12.
[2] https://www.amstat.org/asa/files/pdfs/POL-ClimateStatisticsApril2014.pdf

involving the interaction of many different kinds of physical processes and many different time scales. Thus this area of science has a critical dependence on the examination of all relevant data and the application of statistics for its interpretation. Climate datasets are increasing in number, size, and complexity and challenge traditional methods of data analysis. Satellite remote sensing campaigns, automated weather monitoring networks, and climate-model experiments have contributed to a data explosion that provides a wealth of new information but can overwhelm standard approaches. Developing new statistical approaches is an essential part of understanding climate and its impact on society in the presence of uncertainty. Experience has shown that rapid progress can be made when "big data" is used with statistics to derive new technologies. Crucial to this success are new statistical methods that recognize uncertainties in the measurements and the scientific processes but are also tailored to the unique scientific questions being studied.

IMAGe/SDS has been instrumental in developing statistical methods for model comparison and validation, for combining information from varied sources often of varied spatial and temporal resolution, and for developing methods to analyze climate data/output whose size does not allow for standard approaches. If NCAR is to be an organization without a statistical research presence, it is our firm view that fundamental scientific progress and innovative research in Earth's climate system will be impeded.

Because of its renowned statisticians, its exclusive mission to answer questions that arise from atmospheric science, and its proximity to vanguard climate science work, NCAR's statistical research group has served as the nation's epicenter of statistical climatology since its inception as the Geophysical Statistics Project more than 20 years ago. NCAR has been, and continues to be, uniquely suited for this role, and this interdisciplinary work done by IMAGe/SDS cannot be easily replicated at universities. IMAGe/SDS has served as the liaison between the atmospheric science and statistics communities, conveying the important science questions to the statisticians who can develop statistical methodologies to help answer them. Additionally, IMAGe/SDS has been instrumental in providing postdoctoral and PhD-internship training for many of today's leading researchers in statistical climatology. The cuts to IMAGe/SDS will result in these capabilities being lost to the United States' atmospheric science community.

We in statistics are excited about the opportunities presented by vast new data sources and emerging methods in data science. However, particularly in atmospheric science, there is reason to be concerned about pursing data science without guidance from statisticians. Because machine learning methods tend to be algorithmic, they often do not adequately quantify uncertainty of the predictions they produce. Furthermore, because these algorithms act as black box models, they often do not provide insight about the physical processes that drive the outcomes.

Statisticians have expertise in both developing models that use data to inform about underlying processes and in accounting for and describing uncertainty in estimates and predictions, and the discipline is integrating this expertise into the field of data science and its methods. Indeed, we believe the important role of statistics to data science is implicit in NSF's Transdisciplinary Research in Principles of Data Science (TRIPODS) program, which brings together the statistics, mathematics, and theoretical computer science communities. The American Statistical Association (ASA) and Computing Research Association statements on data science also emphasize the important role of statistics in data science.[3,4]

Thank you for your consideration. We in the statistical community would be happy to be of assistance as you consider how to ensure NCAR can maintain its quality work toward its mission through robust integration of statisticians.

Sincerely,

Daniel Cooley
Chair, ASA Advisory Committee on Climate Change Policy

Barry D Nussbaum
President

---

[3] ASA Statement on the Role of Statistics in Data Science, http://magazine.amstat.org/blog/2015/10/01/asa-statement-on-the-role-of-statistics-in-data-science/, 2015.

[4] Computing Research and the Emerging Field of Data Science, http://cra.org/data-science/, 2016.