Judges' Commentary: The Southwest States' Energy Compact

Katie Oliveras
Mathematics Dept.
Seattle University
Seattle, WA

Stacey Hancock
Dept. of Mathematical Sciences
Montana State University
Bozeman, MT

David H. Olwell
Professor and Dean
Hal and Inge Marcus School of Engineering
Saint Martin's University
Lacey, WA
dolwell@stmartin.edu

Introduction

This is the third year that Problem C has been offered by the MCM. Problem C focuses on mathematical modeling based on real-world data, and is best described as "Data Insights." While not a "big data" challenge, in the sense of teams needing to develop specialized data-handling algorithms and analysis techniques, or have access to high-performance computing platforms, the problem provides teams with an opportunity to encounter real-world, challenging data that have interesting characteristics. Naturally-occurring complicating factors, such as data-set size (but not big data), blend of data types, breadth of representation in data elements, cross-discipline sources, time-series dependencies, censored (or missing) data, and others present themselves, depending on the specifics of the modeling problem.

The UMAP Journal 39 (3) (2018) 343–350. ©Copyright 2018 by COMAP, Inc. All rights reserved.
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice. Abstracting with credit is permitted, but copyrights for components of this work owned by others than COMAP must be honored. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior permission from COMAP.

The problem authors continue to present one of two types of problems: those that require a model to be developed from a large data set, and those that require a model to be applied to a large data set. This year's problem, like the first year's [Olwell et al. 2016] is an example of the former, while last year's problem [Overdeep et al. 2017] was an example of the latter.

In both cases, we expect teams to analyze and discuss the uncertainties of their results in the context of the uncertainties inherent in both the data and the model. That explicit consideration of uncertainty due to random variation is what distinguishes Problem C from Problems A and B.

This year was one of explosive growth for Problem C, as 4,747 teams participated, involving 12,707 students.

Problem Summary

The 2018 Problem C considers energy production and usage based on 50 years of data (1960–2009) on 583 variables for four contiguous states in the southwestern U.S. For the problem, these four states—California, Arizona, New Mexico, and Texas—wish to form a "realistic new energy compact focused on increased usage of cleaner, renewable energy sources," and they are seeking mathematical models that can provide quantitative insights to assist with policy changes.

The problem is divided into three parts. Part I requires teams to:

- model an energy-usage profile for each of the four states that describes the status quo and includes discussion of renewable energy sources;
- describe the historical evolution of energy usage 1960–2009 in terms of that model, and analyze and interpret these results in a way that describes similarities and differences between the four states' usage of renewable energy sources in a way easily understood by policy makers;
- determine which of the four states has the "best" profile for use of cleaner, renewable energy in 2009; and
- use the model to predict the future energy profile in 2025 and in 2050 under current policies.

Part II asks teams to determine renewable-energy usage targets for each of the four states in 2025 and 2050, stated as goals in the new energy compact, and to identify and discuss at least three actions that the four states should take to meet these goals.

In Part III, teams are asked to write a summary of their profiles, predictions, and recommended goals and actions, in the form of a one-page memo addressed to the state governors.

Overview

This problem required teams to develop several components—current profile models, historical profile evolution and comparisons across states, determination of a "best" state, and energy forecast models—to suggest policy recommendations. Overall, judges valued well-written papers that carefully followed the contest directions. They highly valued papers that presented a cohesive narrative, where each step of the modeling and analysis process clearly built upon previous modeling/analysis steps.

The most successful teams first used descriptive statistics and selected plots to analyze differences among the states in both energy usage and energy production, and isolated key influencing factors such as population growth, climate, and the dominant industries for each state. Teams used valid contextual and statistical methods to handle missing data and perform data dimension reduction. They then judiciously used graphs to visualize the similarities and differences between states in terms of energy usage, production, and sector; these visualizations were also used as part of their justifications for their choice of energy profile.

Next, using their energy profile and forecast, the most successful teams made both renewable-energy targets and policy recommendations for each state, based on the entirety of their model and understanding of the data. The top papers modeled how their predicted energy profile would change, including measures of uncertainty, under various policy recommendations, and included some optimization to address questions along the lines of which policy changes would have the most significant impact. In the one-page memo, the strongest papers clearly communicated their results in a way easily understood by and applicable to the intended audience.

Judges looked for papers that clearly described and justified their modeling process. Many papers attempted to use too many modeling techniques. These papers suffered from a lack of clarity and were difficult to read and decipher. It is better to pick just one or two mathematical models that are properly suited to the modeling problem and use them well, rather than use 8–10 different models that may not be suited to the problem.

Additionally, teams generally struggled to describe how the different parts of the problems fit together. For example, many papers made policy recommendations with no clear connection to their mathematical model. Their recommendations appeared to have been made in isolation from their models and the content of the paper.

Judging Process and Categories

The judging followed the typical process of triage, screening rounds, and final judging. For specific details, see the process as described in previous Judges' Commentaries (e.g., [Black, 2013]).

An automated plagiarism check was added this year to the contest.

During the final judging rounds, the judges evaluated the papers based on the following six criteria:

- Data exploration and insights.
- Profiling, modeling, and defining "best."
- Forecasting, sensitivity, and uncertainty.
- Results and recommendations.
- Writing and communication.
- Overall connected, credible content.

Not all categories were weighted equally; the best papers received high marks in each category.

Data Exploration and Insights

As an initial part of the problem, teams were asked to create an energy profile for each of the four states. They were then asked to model the evolution of this profile for each state over time and profile the similarities and differences among the states, including state-specific influential factors as part of the discussion.

Simple descriptive statistics, informative data visualizations, and a clear understanding of the provided data set were critical for addressing this portion of the problem. The data set included more than 500 variables, with some entries either invalid or missing. Strong papers clearly described and justified how they handled missing data, and also clearly explained their process for reducing the dimension of the data set using both contextual and statistical techniques.

Profiling, Modeling, and Defining "Best"

Teams used a variety of techniques to develop their energy profile; many teams used descriptive statistics. The highest-ranked papers explicitly stated how their data exploration and insights justified their energy profile, and then referenced the energy profile in later policy recommendation discussions. They also used appropriate and meaningful data visualizations to communicate how their energy profiles demonstrated the similarities and differences among the four states, and how these profiles changed over time. The best graphs used effective data visualization principles; they facilitated comparisons, provided context, and displayed uncertainty [Cleveland 1994; Robbins 2005; Tufte 2001]. Profiling included clearly identifying state-specific influential factors; teams that failed to identify and discuss possible influential factors were not evaluated past the triage phase.

As with the 2016 Problem C, the question of which state had the "best" profile was deliberately left open-ended. Many metrics were offered by teams; better papers explicitly stated their reasoning for their choice of "best" state, and ranked the four states according to their chosen metric. Teams that failed to identify a "best" state were not considered after triage.

Forecasting, Sensitivity, and Uncertainty

To evaluate potential future impact of policy changes, teams had to develop a forecasting model to predict each state's energy profile in the future, specifically, in years 2025 and 2050. Teams that used insights from their data model were marked higher than those that simply extrapolated the data to provide a future forecast. Many of the strongest papers included separate modeling of predictor variables (such as population, GDP, etc.) to adjust their forecast.

A common mistake in many papers was a failure to acknowledge the uncertainty in predictions or to include any sort of sensitivity analysis. When using predictive models for forecasting, some type of error analysis, confidence intervals, or standard errors must be included with the predictions. Most statistical methods and forecasting strategies have standard techniques that can be used to estimate the uncertainty in the forecast. In particular, when forecasting into the future, predictions become more uncertain (less precise) the further in time you predict.

Judges also looked for papers that attempted to test the validity of their model. Some papers attempted to do this by fitting the forecasting model using a subset of the data (a training set), and used the remaining data as a test set. A predictive model that accurately predicts the training data on which it was fitted may not necessarily predict testing data well. Some teams found more-recent data online, from 2010–2014, and used these data for validation. These attempts at model validation were viewed favorably by the judges.

Results and Recommendations

Many teams made the mistake of providing policy recommendations without connecting the recommendation to the data, model, and results. Any policy recommendations need to be motivated by the model and results; these connections must be made explicit. Since states differ in energy profiles, recommendations should also be specific to each state. Additionally, any recommendations should be actionable. A policy recommendation to "increase renewable energy usage" is not actionable—by how much? what type of renewable energy? and how?

Strong papers made informed policy recommendations and then implemented these policy recommendations in their forecast model to test if their recommendation would be effective. Some papers incorporated an optimization mechanism to make better recommendations.

Writing and Communication

Many papers had interesting mathematical models and results. However, if the model or results were poorly communicated, the paper may not have even made it beyond triage. This was especially true of the memo, which should be a nontechnical, highly-readable letter that "sells" the policy recommendations. Since governors, the intended audience of the memo, may not have a strong background in modeling, the memo should communicate possibly complex models in a way that is readable and understandable to its audience.

Judges looked for papers, memos, and summaries that were clear, comprehensive, and coherent. Good papers managed to find the right balance between details and high-level descriptions, and did so in a way that formed an engaging narrative. With the page limitations, the strongest papers optimized their space usage by judiciously using graphs, charts, and tables that helped communicate their mathematical model and results.

Communication is important not only in the narrative, but also in figures and tables. Overall, judges were disappointed with the types of data visualizations chosen by the teams. For example, pie charts that display the distribution of different energy categories in each state do not allow comparison across states, nor do they even allow for easy comparison between categories. In developing energy profiles, many teams simply reported pages of graphs without explaining why the graphs were necessary, or how they informed their energy profile. Many teams used the original variable abbreviations provided in the data set rather than using descriptive labels, and some even failed to label axes or include captions. Effective data visualizations tell a stand-alone story and make important comparisons easy; they encode data in a way that allows viewers to quickly and easily decode the information [Robbins 2005]. Fancy graphics do not always tell the best story.

Overall Connected, Credible Content

Judges viewed papers that presented a unified narrative throughout the entire paper favorably. This was not limited to the writing, but included also the modeling effort itself. The strongest papers focused on the whole picture, where each component of the paper and model flowed together. For example, the best papers explicitly stated how their data exploration informed their profile, how their forecast informed their recommendations, and how their recommendations would change their forecast.

As always, the best papers provided references and explanations regarding formulas used in their model as well as a brief literature review. Sources of information, including graphics and charts, should always be properly cited.

Conclusion

Problem C attracted significantly more teams in 2018, with 4,747 submissions, compared to 1,527 in 2017 and 1,875 in 2016. Of this year's entries, 465 were recognized as Meritorious, and of those, 138 were selected to go to the final judging. After several rounds of readings, 14 papers remained in the final round. Of these, 6 were recognized as Outstanding and the other 8 were recognized as Finalists.

The Outstanding teams were from

- Beijing Forestry University, Beijing, China;
- Beijing University of Technology, Beijing, China;
- Shanghai Jiao Tong University, Shanghai, China;
- Shanghai University of Finance and Economics, Shanghai, China;
- Virginia Tech, Blacksburg, Virginia, USA; and
- Xi'an Jiaotong University, Xi'an, Shaanxi, China.

The Finalists were from

- Beihang University, Beijing, China
- Harbin Institute of Technology, Harbin, Heilongjiang, China
- Northeastern University of Qinhuangdao, Qinhuangdao, Hebei, China
- Peking University, Beijing, China
- Tianjin University, Tianjin, China
- Tongji University (two teams), Shanghai, China, and
- Wuhan University, Wuhan, Hubei, China.

The Outstanding team from Virginia Tech was recognized by the Mathematical Association of America (MAA). The one from Shanghai Jiao Tong University was recognized by the Institute for Operations Research and the Management Sciences (INFORMS). For the first time, the American Statistical Association (ASA) honored a paper: the Outstanding team from Xi'an Jiaotong University.

The judges were very pleased with the quality of the submissions and the increased interest in Problem C.

References

Black, Kelly. 2009. Judge's commentary: The Outstanding Traffic Circle papers. *The UMAP Journal* 30 (3): 305–311.

________. 2011. Judges' commentary: The Outstanding Snowboard Course papers.*The UMAP Journal* 32 (2): 123–129.

________. 2013. Judges' commentary: The Ultimate Brownie Pan papers. *The UMAP Journal* 34 (2–3): 141–149.

Cleveland, W.S. 1994. *The Elements of Graphing Data*. 2nd ed. Summit, NJ: Hobard Press.

Olwell, David H., Carol Overdeep, and Katie Oliveras. 2016. Judges' commentary: The Goodgrant Challenge papers. *The UMAP Journal* 37 (3): 32–331.

Overdeep, Carol, Katie Oliveras, and David H. Olwell. 2017. Judges' commentary: The Self-Driving Cars problem. *The UMAP Journal* 38 (3): 347–353.

Robbins, N.B. 2005. *Creating More Effective Graphs*. Hoboken, NJ.: Wiley.

Tufte, E.R. 2001. *The Visual Display of Quantitative Information*. 2nd ed. Cheshire, CT: Graphics Press.

About the Authors

Katie Oliveras is an Associate Professor of Mathematics at Seattle University. She earned a B.S. in Systems Science and Engineering at Washington University in St. Louis, and then M.S. and Ph.D. degrees in Applied Mathematics from the University of Washington. As an undergraduate, she participated in the MCM competition; and she has served as a team advisor for the MCM teams at Seattle University.

Stacey Hancock is an Assistant Professor of Statistics at Montana State University. She earned her B.A. in Mathematics and Music at Concordia College in Moorhead, MN, M.S. in Statistics at Montana State University, and Ph.D. in Statistics from Colorado State University. Her primary research interests lie in statistics and data-science education, and she serves as the American Statistical Association judge for the MCM competition.

David H. Olwell is Professor and Dean at the Hal and Inge Marcus School of Engineering of Saint Martin's University. He earned a B.S. at the U.S. Military Academy, where he studied mathematics, and an M.S. in Mathematics, an M.S. in Statistics, and a Ph.D. in Statistics from the University of Minnesota. He has previously been on the faculty of the U.S. Military Academy and of the Naval Postgraduate School. Dr. Olwell has been a problem author for both the MCM and HiMCM, and has been an MCM judge for almost two decades. He was the Head Judge for Problem C.