

Judge's Commentary: The Outstanding Zebra Mussel Papers

Gary Krahn

Dept. of Mathematical Sciences
United States Military Academy
West Point, NY 10996
ag2609@usma.edu

Introduction

The papers were assessed on

- their ability to transform the data into useful information;
- the application of an appropriate modeling process; and
- the integration of environmental science to render appropriate recommendations.

The judges appreciated the effort and valued the results of the papers. It was a very difficult problem that required a blend of science, mathematics, and conviction to solve a complex interdisciplinary problem during the four-day contest. It was clear that a solution was not going to jump out of the 40 pages of data; rather, it had to be pulled out skillfully.

The Problem

Zebra mussels were introduced to North America in 1980. They are an ecological “dead end,” since native fish do not eat them. Researchers are currently attempting to identify environmental factors that may influence the population of zebra mussels within our waterways. Zebra mussels are now spread throughout the eastern waterways of the United States, causing tremendous problems for the ecosystem and the regional economies.

The UMAP Journal 22 (4) (2001) 415–420. ©Copyright 2001 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 data in the problem statement are real: Prof. Nierzwicki-Bauer of Rensselaer Polytechnic Institute, a leading researcher of zebra mussels, provided data from several lakes in New York. Several population models appear in the literature; however, the collection of environmental factors that influence the rate of population growth of the zebra mussel is still unknown. This is a genuine interdisciplinary problem that confronts North America today.

The Data

The data appear to have created an “uncomfortable” feeling in the hearts and minds of the modelers. It was difficult for many to digest all of the data and either incorporate all of it into a model or else justify eliminating portions of the data. Often, teams did not address how they managed “missing” data or why they accepted or refuted data that appeared to be erroneous. In most cases, teams had done a significant amount of work in an attempt to understand the data. Most teams categorized the population data by month in order to synthesize the data into a more useful form. Similarly, they attempted to align the chemical data by averaging several time periods into a single data point. Many teams had difficulty describing their analysis and the interpretation of their results. The successful teams discussed how they transformed data and how they confronted missing or confusing data. **Tables 1** and **2** show portions of the data, the zebra population of one lake from 1994–2000 and the chemical information on the same lake for 1999. Confusing—yes, but real.

The entire set of data included the following categories: stratum, total phosphorus, dissolved phosphorus, calcium, magnesium, total nitrogen, temperature, chlorophyll, alkalinity chloride, iron, potassium, sodium, pH, secchi disk transparency, and population levels. It was essential to explain how the data would be organized for analysis. The judges expected teams to describe why they selected certain data to remain in their analysis and why other chemicals were eliminated. It was clear that contestants had to make several decisions to transform the data into a useful form. This problem, like last year’s problem, was not clear-cut. Once again, we found that as the contestants formulated and refined their assumptions, they confronted the complexities typically associated with an open-ended problem. Last year they had reasonably clean data, while this year they had some “dirty” data.

The characteristic of a strong paper was the ability to uncover the uncertainties of the population growth of zebra mussels due to chemical concentrations using science and mathematical models. In some cases, the incomplete data and large unexplainable fluctuations in the population obscured the effect of specific chemicals. The data alone cannot reveal the complete interaction among the chemicals affecting population growth. For that reason, successful teams had to take an interdisciplinary problem-solving approach.

Table 1.
Zebra mussel population of one lake.

Date	Population
7/1/94	100
8/1/94	70
9/1/94	50
10/1/94	248
11/1/94	1,045
7/12/95	222
8/1/95	50
9/1/95	70
10/1/95	40,000
11/1/95	200,385
7/1/96	39
8/1/96	4,843
9/1/96	30,033
10/1/96	949,433
11/1/96	49,333
7/1/97	0
8/1/97	20,456
9/1/97	44,678
10/1/97	345,555
11/1/97	98,789
7/1/98	605
8/1/98	84,132
9/1/98	599,432
10/1/98	454,932
11/1/98	49,332
7/1/99	93
8/1/99	45
9/1/99	83,962
10/1/99	539,229
11/1/99	30,012
7/1/00	0
8/1/00	50
9/1/00	9,483
10/1/00	592,339
11/1/00	467,876

Table 2.
Chemical profiles of the same lake.

Date	Ca mg/L	Mg mg/L	TN mg/L	Temp °C	Chl-a
4/15/99	23.20	5.07	0.44	8.50	4.72
5/17/99			0.32	15.50	7.27
5/18/99	27.50	6.71	0.45		
6/1/99			0.49	18.20	10.18
6/9/99			0.42		
6/14/99			0.47	21.00	11.64
7/1/99			0.52	20.80	9.45
7/19/99	26.80	6.72	0.86	21.00	10.18
7/20/99	27.20	6.61	0.56		
7/29/99			0.44	21.80	13.58
8/4/99			0.52		
8/11/99			0.51		6.30
8/23/99			0.44	21.00	5.09
8/25/99			0.41		
9/7/99			0.38	22.00	12.12
9/13/99			0.38		
9/24/99	24.80	5.67	0.89	16.00	1.09
10/7/99			0.73	13.50	3.64

The Science

If science is defined to be the knowledge and study of "what is," then most of the teams got half of the science—the knowledge part. Almost every team was able to find an enormous amount of information from the open literature by using the Internet. The stronger teams not only gathered information, but they also explained the impact of specific environmental conditions on the life-cycle process of the zebra mussel. If chemicals such as nitrate and magnesium were eliminated without explaining why, the grader immediately suspected that the student did not know why. Likewise, if variables such as chlorophyll, pH level, and calcium were kept in the model, the outstanding teams explained why, from both a modeling and a scientific perspective. An explanation of the model using both science and mathematics was a characteristic of an outstanding paper.

An understanding of the ecological fabric of the waterways was important in the design of an outstanding solution to this problem. Environmental science was the thread that related the data to the model and the model to a "realistic" solution.

The Model

It was important that the modeling process be well formulated and that the rationale of the selected model be clearly explained. The definition of variables, identification of simplifying assumptions, and a discussion of the

ramifications of these assumptions were important ingredients in the paper. Finally, it was important that the model developed was used to answer the question regarding the expected growth of the mussels in lakes B and C. An interdisciplinary discussion of the ramifications of the de-icing policy required in Part E was also directly tied to the model. Surprisingly, many teams did not take advantage of their model to address follow-on questions.

The explanations of the modeling process varied tremendously. Some papers contained models that were well designed with results that were analyzed and interpreted. The teams also recognized their models to be both predictive and descriptive. Unfortunately, other papers had wonderful models that utilized a commercial package or constructed models, but they never explained how the model functioned. It appeared that providing the details of the model's underpinnings impacted the entire paper. Groups who had good explanations of their models also related these models nicely to the environmental science of the zebra mussel rate of growth.

The analysis of the data tied in nicely to how the students performed their modeling. Some students saw the problem as fitting a growth differential equation, and others as fitting a multivariable regression. The approach did not affect the assessment of the paper. Furthermore, whether a team used a discrete dynamical system, curve fitting, or simulation, or adapted the logistic model, a correlation analysis was very important. The stronger papers tended to perform this analysis graphically. Those groups providing useful graphs and explanations of these graphs faired quite well.

The Analysis

The problem was an interdisciplinary endeavor. Teams that did great mathematics but revealed little knowledge on environmental science could not capture the relationships required to solve this problem. Since the data were not clean, it was impossible to use only the data to uncover the essential relations affecting population growth. Similarly, teams that had a tremendous knowledge of the science but little mathematics were not able to create an appropriate predictive model. A thorough explanation of the implication of each variable on the growth of the mussel population was essential. Good teams shared a modeling process that was well thought out and justified the rationale of the selected model. In Part C (adjustment of the model) a clear explanation of the process involved in modifying the model was important. Finally, in Part D it was important that the analysis of the model was used to answer the question regarding the expected growth of the mussels in Lakes B and C.

An interdisciplinary discussion of the ramifications of the de-icing policy was directly tied to the model. Those students who answered all the requirements had a significantly greater chance of going forward than those groups who either did not answer the requirements or who only addressed one or more requirements superficially.

Presentation

Some papers revealed tremendous analysis but lacked clarity in the presentation. The strong papers presented the problem, discussed the data and explained their analysis, and finally revealed the development of their mathematical methods/models. The big difference in papers was whether they informed the reader of what they did and, more important, how they did it. A clear presentation allowed the judge to comprehend their logic and reasoning. One judge noted that he wished he was a mind reader because there was clearly lots of outstanding work; however, only the result was revealed. The strong papers revealed their analysis, not just the results.

Very broadly, we saw two types of weak presentations. The first consisted of reports that had a significant narrative, but no support in the form of mathematical modeling or analysis. In these reports, the groups appeared to rely on qualitative observations and the information from the literature (web sites) to reach conclusions. The other type of poor-quality report was those that had a significant amount of mathematics in the form of tables and graphs, but no modeling or analysis to pull it together. These papers appeared to dump their computer runs into the report but did not really know what to do with them.

This year we noticed that the stronger teams clearly documented information they gathered from outside sources. When constructed models aligned very closely with models found in the open literature, it became difficult for judges to determine what was original work.

Conclusion

The effort and creativity of almost every team was inspiring. It appears, however, that most teams can reason better than they can communicate. Often, wonderful ideas were not revealed to the reader. The necessity to work with large data sets appeared much more difficult than anticipated. The top papers, however, did an amazing effort of blending and revealing the science, research, and mathematics. The best teams revealed the power of interdisciplinary problem solving.

About the Author

Gary Krahn received his Ph.D. in Applied Mathematics at the Naval Post-graduate School. He is currently the Head of the Dept. of Mathematical Sciences at the U.S. Military Academy at West Point. His current interests are in the study of generalized de Bruijn sequences for communication and coding applications. He enjoys his role as a judge and Associate Director of the ICM.