Judge’s Commentary

MathWorks Math Modeling Challenge

Kelly Black, Ph.D.

Department of Mathematics, University of Georgia

Introduction

The focus of MathWorks Math Modeling Challenge 2019 was the use and abuse of substances such as nicotine, marijuana, alcohol, and unprescribed opioids. The specific questions were about predicting the spread of nicotine use due to vaping, modeling and predicting the likelihood that certain individuals will use certain substances, and assessing the broader impacts of substance abuse. As we have come to expect, the teams did a tremendous job of exploring a complex topic and providing insightful analyses.

The primary goal of M3 Challenge is to provide students with a context to work in small teams, explore an open ended question, engage in the full process of bringing together disparate mathematical ideas, and share these ideas in written form. The judges recognize that it is the teachers and coaches who help prepare teams for this event and provide the tools, structure, and support that make this a meaningful experience. Thank you!

This event has grown, and it is because of the dedication of the coaches and teachers that it is also maturing. We initially saw large gains in the quality of the students’ submissions. We continue to see improvements, but they are more subtle and in some ways more substantive. These continued gains are the direct result of the efforts of the teachers and coaches.

This commentary is provided as a way to give insight into some of the things that judges noticed and wish to share. It is our hope it can be used as one of many resources. In this commentary, I will discuss my personal observations about each of the three questions and the role of scientific computing, and I will make some general comments about modeling and technical writing. Each of these topics is discussed in the following sections in order.

Question One

The first question has two parts, and the first part required teams to predict how nicotine use will spread over the next ten years due to vaping. The second part required teams to compare their results to nicotine use associated with cigarettes. In this section I will first discuss how teams reacted to the two parts of the question. I will then discuss the differences between modeling percentages of the population versus numbers of people especially with respect to regression techniques. Lastly, I will discuss a more advanced approach—the use of SIR models—that a number of teams adapted for this question.

The first part of the question required the teams to predict how nicotine use will spread. Few teams explicitly defined their interpretation of this part of the question. The word “spread” is ambiguous, and teams addressed this aspect in different ways. One common approach was to predict either the number of people or the percentage of people who would use a product that contained nicotine *at least once* in the ten year time span. Another common approach was to predict the number of people or the percentage of people who would use a product containing nicotine *regularly* during the ten year time span. It was not always clear if a team examined the total number of people or a percentage of people, but most teams were careful to state whether or not they examined one-time or long-term users.

To answer the second part of the question, teams were asked to compare cigarette use to current vaping trends. There was a wide variation in how teams reacted to this part of the question. A large number simply did not address this part. Many explored the historical trends of cigarette use without directly comparing their predictions about vaping. Those teams that did make an explicit comparison were split between those that made vaping predictions based on past cigarette trends, and those that made predictions using current cigarette trends. It is notable that several teams stated assumptions that future vaping trends will follow past cigarette trends, but did not discuss this assumption when making a comparison.

As stated above, nearly every team focused on either the total number of people or the percentage of the population who met their definition of nicotine users. Unfortunately, it was not always clear which quantity a team used for its prediction. The distinction is important. A large number of teams built a model based on regression of a model given data. This is a topic that arises almost every year. Simply determining the parameters of a function that appears to fit the data is not modeling, especially when the task requires a team to extrapolate to times outside of the given data.

For example, it was not uncommon for a team to use a linear regression model or a logarithmic model to predict the percentage of people who use nicotine. This is problematic in that a linear model with non-zero slope will definitely go outside the bounds of 0% to 100% at some point in the future. All models are only good within a limited scope, and a team that uses this kind of model should state why they chose it, provide motivation as to why it is appropriate, and be careful to note what limits should be used when predicting the future. Simply noting that it is consistent with the data over a given time interval is not good enough.

A better example of model construction is to first note the general trend. For example, one common assumption was that the percentage of people using nicotine due to vaping would rapidly increase, but the increase would eventually slow with the percentage of nicotine users slowly approaching an equilibrium. Based on this assumption, a logistic model is a good choice to approximate the percentage of nicotine users. Once the model is in place then the next step is to use the data to approximate the values of the parameters.

Using the data to help construct a model requires that a team be careful in interpreting the resulting approximation of the model's parameters. In this year's event the data associated with vaping is limited. The phenomena of interest is relatively new and few long term studies have been conducted. Many of the data sets that teams used were small, and many contained points that have the potential to greatly impact the calculations. Teams should be careful to explore the impact of influential points and examine how their predictions could change if those points were not used or were changed by a small amount.

This exploration of influential points is related to sensitivity analysis but is not the same thing. Both types of analyses play important roles in evaluating how a prediction can change based on a change in the underlying calculations. A sensitivity analysis would examine how the predictions differ if one or more parameters is changed by some small amount. The role of looking at influential points is to determine how the predictions change if there is a change in the underlying data. The resulting analyses can then determine which parameters have the greatest impact on the predictions as well as how the results may be influenced by the data.

The last topic examined is a modeling approach employed by many teams that requires the construction of a system of differential equations. This is an advanced technique and is more generally associated with the undergraduate mathematics curriculum. In particular, a number of teams made use of an SIR model [1] to predict how nicotine use will change over time. The idea is that you divide a population of people into three separate groups. The first group, denoted S, is the set of people who are susceptible to using a nicotine product but have not yet used nicotine. The second group, denoted I, is the set of people who are "infected," that is they are current nicotine users. The final group, denoted R, is the group of "recovered" people, that is, the people who once used nicotine but no longer use nicotine.

This approach is an interesting and appropriate path to gain insight into the relevant dynamics. It also makes use of an advanced topic. Many simpler approaches yield similar results and can also help illuminate the relevant dynamics. The SIR model is normally used to model the dynamics of the spread of a disease. In this case, though, the interactions between the groups differ from the interactions associated with disease. For example, it is not clear what the distinction is between people who are susceptible and those who have recovered. The people in both classes can move into the infected class, and it is not clear how that movement should differ between the two groups. Also, if a team defined their notion of the spread of nicotine use to be the total number of people that have tried a nicotine product, then determining how to add those people up over time is a non-trivial task.

Judges who read the entries tend to have a detailed knowledge of a wide range of mathematical models. A judge reading a paper likely knows that there are important nuances associated with an approach such as the SIR model. A team that decides to pass over a more straight forward approach may not necessarily elicit a favorable reaction if they overlook a subtle assumption implicit in the model they choose.

Question Two

The second question also has two parts. The first part of the question is to create a model to determine the likelihood that a given individual will use a particular substance based on various characteristics of the person and the substance. The second part is to demonstrate how the model works by predicting how many students in a high school class of 300 will use four specific substances. The two parts are closely linked, and I will discuss three different aspects of the teams' submissions for question two. The first aspect is the set of factors teams decided influenced an individual's long term behavior, the second is the role of social influence, and the third is the ways in which the different factors were brought together.

The first thing students had to do for question two was determine which factors impacted whether or not a person will use a given substance and create a model that approximates the likelihood that a given individual will use a given substance based on those factors . The majority of teams chose factors that were found in a data set they obtained. Common factors included income, race, and socio-economic status (SES). A small number of teams examined other important factors such as the difference between rural, suburban, and urban groups. Teams that incorporated other factors impressed the judges by demonstrating that they had thought about the problem and explored a wide range of situations that impact people's habits.

Part of the difficulty, though, was to find data sources that could be used to determine the relative strengths of different factors. Teams that made use of a wider array of data sources generally tailored their factors to make use of the information available to them. This is a common limit for anybody developing mathematical models. It is important, though, to recognize this limit and acknowledge that the data available put limits on what traits a team was able to consider. Judges do not expect teams to find many data sources, but a team that discusses the limitations in the factors they considered while suggesting other factors tends to make a stronger impression about their insights into a phenomena.

One last issue about the variety of factors examined is the way in which a team described the factors. Many teams simply listed the factors and then made assumptions about the interactions between their effects. It is good to be explicit in discussing which factors are examined, but it is also important to discuss the role the different factors have and how they impact the situation.

In addition to identifying different factors that influence a person's decision to try given substances, part of the question asks that teams reflect on the role of social influences. Prior to the start of the event, judges predicted that this would be an important discriminator between papers. The underlying assumption was that students would have good insights into how social networks influence people, and this part of the question lent itself to many creative approaches. In the end, though, few teams considered this part of the question, and judges had to change their expectations and downplayed the weight given to this part of the question.

The third aspect of this part of the question required teams to examine the different factors they considered in the first part and decide how to bring them together. They had to think about the relationships between the different factors, figure out their relative weights, and decide how likely it would be for a given a person to use the given substances.

The two most common approaches were to either add the impacts or multiply the impacts; however, both of these approaches are problematic. Given the restrictions of a mathematical model, the limited information, and the difficult time restrictions of the event, either approach can be justified. It is important, therefore, that teams recognize and discuss the potential issues. For example, many teams used the data to determine the proportion of people sharing a given trait who will also use a given substance. Then, given all the traits that a person shares, many teams multiplied the proportions to get a final result. One disadvantage to this approach is that it assumes the traits are independent. Some factors, however, such as race and socioeconomic status have strong links. Additionally, every time a number is multiplied by a probability, its value will decrease. This implies that if you use enough traits the probability calculated can be made to be quite small. The second approach, adding the proportions, is also problematic. Adding assumes that the traits are disjoint. Clearly it is possible for a person to embody more than one trait, so it is not appropriate to simply add the proportions.

In the second part of question two, teams were expected to use their model to predict how many high school students in a class of 300 will use each of the given substances. The large majority of teams took the proportions obtained for their different factors and multiplied by 300. This is a good first step, but it assumes that the demographics of a given high school population are the same as the demographics of the people used to obtain the original data. Teams that noted this underlying assumption made a positive impression and demonstrated they understood an important limitation of their model. Some teams went further, and examined the demographics of a particular location for the relevant age group, and they used those statistics to create a class of students.

Question Three

The third question was open ended and required students to examine the broader impacts of the use of nicotine, marijuana, alcohol, and unprescribed opioids. Students had to determine which aspects play a role in the societal costs associated with the use of each substance. Teams were specifically asked to develop a metric that can be used to measure the impact of substance use. This is a difficult task, and many teams struggled to determine and describe a metric.

The question statement says that the metric should be "robust." Very few teams addressed this part of the question, or shared their interpretation of what it means for a metric to be robust. This is something that many professional modelers tend to struggle with, but it is an important consideration. Because of its imprecise nature, though, it is not uncommon for this part of a model to be the subject of a great deal of debate and disagreement.

Once a team determined the different aspects to use, the next task was to find a way to combine them. For example, substances like opioids can take a heavy toll with respect to lives lost, money lost, and the long term social strain on a given community. Teams had to take these different aspects and make a direct comparison between them.

One common approach was to simply state a weighted sum of the impacts of the different factors. This is problematic in that adding values with different units does not necessarily make sense. Also, the weights that were used were generally not well motivated and were often simply stated without much discussion. Teams that made an effort to convert the different impacts into a single, common unit tended to make a more positive impression. For example, the different aspects may be measured in terms of money, lives lost, time while a person was incapacitated, or numbers of families experiencing a great deal of stress.

Technical Computing

This is the second year that the event featured special consideration of the use of technical computing as a separate award. Judges have a better feel for what to expect and what is appropriate compared to previous years. The difficulty is to ensure that our expectations are realistic and are also consistent with our goal to have students work together to develop mathematical models as a way to gain insight into the world around us.

Computing is a vital part of the modeling process. It helps us explore and understand a mathematical model which then allows us to adapt and refine a model. Given the tight time constraints of this event we do not necessarily expect teams to be able to adapt their models. The results of an approximation can be used, though, to help identify potential strengths and weaknesses of a team's approach.

When trying to decide how to compare the use of computation and approximation by a given team, judges generally look at a few different practices employed by a team. For a given team that has been designated as making use of scientific computation judges ask the following questions:

- Did the team discuss their computational approach as well as their results in the narrative?
- Did the team interact with the results and use the results to inform their decisions?
- Were the computational approaches consistent with their model?
- Was the code clear, efficient, well organized, and commented?

With respect to the first question, teams should not simply state that they wrote a program and then refer to an appendix that contains their code. Teams should briefly discuss the code and discuss the decisions that were made that led to the final code. Moreover, a broad overview of the code should be available within the narrative that makes it clear what the code does and how it answers a question. After reading the narrative, there should be minimal difficulty in reading the code, and there should not be any surprises.

Carefully integrating the code into the model development should keep the reader well informed of the goals and the approach employed in the code. The code should be an integral part of the model. Just like with the mathematical model, a team should be critical about the limitations of their code. The team should be able to identify what is good about the code, but they should also make it clear that they recognize shortcomings that should be addressed.

With respect to the second question, it is expected that the team goes beyond simply presenting their results. The role of scientific computing is to help inform and guide the modeling process. The team should examine their results and make decisions about the efficacy of their model. Given the difficult time constraints they likely will not have time to change or update their model, but they should be able to state how their results influence changes that should be made in a future refinement.

With respect to the third question the programs and the mathematics should be tightly interwoven. The mathematics should guide the process and decisions about the programming approaches, and the results of an approximation should provide important insights that in turn help a team to adapt and refine the mathematics. Judges understand that this event is a difficult challenge, and it is not fair nor reasonable to expect a complete paper. We do hope, however, that the computational tools be consistent with the mathematics as well as the particular questions raised by a team. A wonderful program that provides insight into a different question is not helpful within the given context.

Finally, the fourth question is focused more on the usability and sharing of the code itself. To be able to go back and work on an existing project, the team's work should be well documented and their intent should be clearly expressed. With respect to the computational tools, the code should be neat, efficient, well organized, and commented. Multiple people generally work on large projects and have to interact. Code that is not neat and commented impedes progress. At the same time, as computational tools tend to expand, it is common to push hardware to its limits. Code that is efficient and well organized tends to aid in the kinds of improvements that allow a program to make the most of the architectural limits imposed on it.

General Modeling

In this final section five general considerations are discussed. The first is how to express equations. The second is references for a URL. The third is citing software and the results associated with a software package. The fourth is the relationship between assumptions and the resulting mathematical model. The fifth and final topic is how assumptions are shared with a reader.

The first topic to discuss is how equations and text can be integrated within a narrative. It was not uncommon to read papers during the early rounds with equations that were disjoint from the text. An equation should be part of a sentence and proper punctuation should be used. A person reading the narrative does not know what factors were balanced to define a model, nor does the reader know what phenomena a team is trying to approximate. A team's report should discuss their insights and thoughts about the model itself. The report should include an explanation of an expression's meaning and how it is consistent with the phenomena of interest. [2]

The expressions developed by a team are usually motivated by the work of others, and it is common to use the work of others to gain insights into a problem. It is a good thing to make use of the work of others! The teams taking part in M3 Challenge continue to get better in providing both citations within the narrative as well as a list of references at the end of the discussion. However, a number of teams still struggle with how to provide references for URLs. Many teams provided a list of references that were simply a long list of URLs. When providing a reference for a URL, teams should give the name of the organization hosting the site, the title of the page, the date that the webpage was retrieved, and the URL, at a bare minimum. [3]

Another small detail associated with citations is how to cite software. Citing software was an especially acute issue this year due to the large number of teams that employed some manner of regression to fit the parameters associated with their model. Under both the APA [4] and the IEEE style guide [5], standard software packages do not require a citation, but you should state within the narrative both the name and the version number of the software package used. When in doubt, though, giving a citation for software is a good thing. Also, if you use a website such as Desmos [6], a citation for the URL should be given.

Another area where we continue to see improvements is in the way teams share their assumptions with the reader. More teams are providing specific details about the motivations for their assumptions as well as more details about the implication of their assumptions. As teams state more details they should be careful to differentiate between the assumptions and the model itself. For example, most teams did a great job of discussing their assumption that the percentage of people using nicotine will rise and eventually level off. This assumption implies that a logistic model may be a good first choice for a model. A small number of teams started with the assumption that the percentage follows a logistic function which implies that the percentage eventually gets close to a constant carrying capacity. The former approach is a better starting point in that the assumptions describe a team's expectations about the physical situation which in turn impacts the choices in deciding which models to employ.

Finally, as teams continue to improve in the way they state and use their assumptions, it was more common this year to see reports in which each section started with its own list of assumptions. Teams were more likely to provide motivations for their assumptions and the implications of each assumption within the relevant section. Reading the relevant assumptions within each section results in a paper that is much easier to read and follow. The teams that do so demonstrate that they understand the important role of assumptions, and they also demonstrate that they know how their assumptions impact which parts of their modeling efforts.

Conclusions

MathWorks Math Modeling Challenge continues to mature and the results of the teams' efforts continue to improve. These improvements are more subtle than we have seen in previous years, but the teams continue to impress the judges. The efforts of the students, coaches, and teachers are inspirational, and you all deserve our highest praise. Thank you again for making this event possible and creating an environment that allows your students to excel and take part in a formative experience.

The topic of this year's M3 Challenge was the use and abuse of substances such as nicotine, marijuana, alcohol, and unprescribed opioids. One of the difficult challenges this year was to construct a prediction based on minimal data with regard to the vaping/nicotine question. The vaping phenomena is still relatively new, and few data sources are available. Despite the lack of data, important questions are raised, and mathematical models can still be used to provide insight and help understand important questions.

With respect to generating mathematical models, the role of regression techniques is a recurring topic. Due to the nature of the questions posed this year, it was a central topic. Careful consideration about the meaning and limitations of regression are necessary when deciding on a particular model, and then the parameters of that model can be estimated using regression techniques. This is especially true for this year's event where the central questions explicitly ask students to extrapolate beyond the available data.

Acknowledgments

I would like to thank a number of people who helped to bring the ideas discussed here together. First, thanks are extended to Erin Byrne from MathWorks. She explicitly identified and graciously shared her insights into the role of computing and modeling. Her point of view was a great aid in assembling the section on the role of technical computing. The same is true of Pat Gorman of the University of Nevada, Las Vegas. He also provided insights and feedback that were helpful.

With respect to the writing and editing, special thanks go to Kathleen LeBlanc at SIAM. Her help in identifying passages that were not clear was invaluable. Her work was a direct aid in making this a better document.

Finally, a huge thank you also goes to Michelle Montgomery, M3 Challenge Project Director at SIAM. Her insights and guidance into the document were warmly received, and she also played a special role in improving this document. Additionally, her unwavering, guiding spirit and forceful efforts in pushing this whole enterprise is a testament to her indomitable spirit. She provides a constant push forward that impacts students nationwide.

Bibliography

1. Morris W. Hirsch, Stephen Smale, and Robert L. Devney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos, 2nd Edition*, 2004, Elsevier.

2. Nicholas J. Higham*, Handbook of Writing for the Mathematical Sciences, Second Edition*, 1998, SIAM, Philadelphia.

3. Columbia College, Vancouver BC (Canada). Websites – APA Citation Guide (6th edition). Retrieved 4 April 2019 from http://columbiacollege-ca.libguides.com/apa/websites.

4. American Psychological Association, APA Style Bold: How to Cite Software in APA Style. Retrieved 4 April 2019 from https://blog.apastyle.org/apastyle/2015/01/how-to-cite-software-in-apa-style.html.

5. Naval Postgraduate School, IEEE – Citation Guide. Retrieved 4 April 2019 from https://libguides.nps.edu/citation/ieee#computer-program.

6. Desmos, Desmos | Beautiful Free Math. Retrieved 4 April 2019 from https://www.desmos.com/.