Judges' Commentary: Eradicating Ebola

Kelly Black
Dept. of Mathematics
Clarkson University
P.O. Box 5815
Potsdam, NY 13699–5815
kjblack@gmail.com

Introduction

The Eradicating Ebola Problem required teams to estimate the optimal way to distribute a new medication to areas stricken by the Ebola virus, so as to eradicate the virus. The tasks included modeling the spread of the disease, determining how to ship and store the medications, and how to deliver the medications to people.

The majority of teams focused their efforts on modeling the spread of the disease. Many teams provided additional models to examine where to ship and store stockpiles of the medication as well as how to distribute the medication. A smaller number of teams were able to construct models that included the interactions between these two parts of their models.

We start the discussion of the teams' efforts with the modeling associated with the spread of Ebola. Next, we briefly indicate the distribution issues, followed by a brief note on how to examine the interactions between these two tasks. Once we have discussed the modeling, we examine the issue of parameter fitting, and then explore the practice of examining the sensitivity of the models. Finally, we offer a few notes about writing in general.

We do not provide here an overview of the judging process. It is an important consideration, described in previous commentaries [Black2009; Black 2011; Black 2013]. Faculty advisers and team members should be aware of the process, since it provides valuable insight into why particular parts of their document are important and how they are viewed at different stages in the judging process.

The UMAP Journal 36 (3) (2015) 237–244.

Disease Modeling

The vast majority of teams first discussed modeling the spread of Ebola. The most common model was a compartmental model, but a large minority of teams used cellular automata models. The compartmental models provide an excellent starting point for a whole population, but the cellular automata models make it easier to examine the spatial dimension of the spread of the disease.

Compartmental Models

Many teams with a compartmental model tended to start with an SIR (susceptible–infected–recovered) model [Murray 2008]. It was common then to follow with refinements such as an SEIR (susceptible–exposed–infected–recovered) model, or an SIQR (susceptible–infected–quarantined–recovered) model, or an SEIR model with a delay. The addition of an exposed class is motivated by the time required by the virus until the infected person is able to pass the disease to other people. The addition of the delay is motivated by the practices associated with local customs and how bodies are treated prior to burial.

One thing missing from many of the models that the teams examined is the impact of the spatial dynamics associated with the spread of the disease. The teams that did try to address the spatial component usually examined separate populations in different regions, and treated each population using their model. In some cases, the models were adapted to accommodate movement between the different regions.

Cellular Automata Models

In the cellular automata approach, a simple set of rules was determined, and the results of numerical experiments were provided. Just as in the compartmental models, it was common to divide the area in question into separate regions and then include terms to accommodate movement between the different regions.

The teams with a cellular automata approach had the added burden of formulating nontrivial rules to define the behavior of their computational model. Additionally, the numerical trials result in a random variable, so their results should be stated in the context of a statistical framework.

Distribution

Another part of the task included modeling the transport, storage, and the distribution of the medicines used to combat Ebola. This was one of the more difficult aspects of this year's event. The methods used to address this part of the question varied widely. The most common approach was simply to choose cities in areas with a high density of Ebola cases that also have the infrastructure to handle large shipments. The model was then constructed by treating the separate cities as nodes on a connected graph.

Most teams simply stated the locations with little or no motivation. Some teams provided a more careful analysis that motivated their choices, and some teams assembled models that calculated the impacts for different combinations of cities. The teams then devised schemes to search for the best choice of cities.

Once the locations were chosen, most teams built on the idea of a connected graph, with the cities the nodes of the graph. Once the model was constructed, an objective function was generally defined. Some teams created a weighted model that used the distances between nodes or the average distance from patients; others created a weighted model based on the need, which came from their model of the spread of Ebola; still others used a combination that tried to balance the distances with the need.

Once the locations were determined and the objective function defined, the teams then had to determine the best way to disseminate a fixed amount of medicine that would minimize/maximize their objective function. Because of the discrete nature of the models, the approaches generally included approaches such as a genetic algorithm, a clustering algorithm, or swarm optimization.

Connecting the Disease Model and the Distribution Model

The majority of teams were able to create, approximate, and describe their model for the spread of Ebola. A large number of teams were then able to create a model for the distribution of the medications that were available. Relatively few were able to bring the two models together and explore their interaction.

Most teams used their disease model to determine the level of needs in the various areas. Once the need was determined, they were then able to apply their distribution model to determine how to address the predicted need. The problem with this approach, though, is that the predicted need is closely tied to the ability to deliver medicines.

A small number of teams addressed this important aspect of the problem. Of those that did, many simply applied an iterative process in which they applied their disease model, determined the optimal way to deliver medicine, and then repeated the process to determine the need within the new context of the assumed availability of the medicines.

An even smaller number of teams were able to incorporate their disease model directly into the objective function that determined the level of need for a given region. This is an impressive accomplishment for a relatively short event. We certainly do not expect many teams to be able to complete such a difficult task, but it is important for a team at least to acknowledge this important aspect of the modeling effort.

One of the downsides of the MCM is that the idea of model refinement and the process of updating a model is skipped because of time constraints. It is important, though, for a team to at least recognize what should be changed and what should be retained when updating their model. This is why we expect the teams to include addressing the strengths and weaknesses of their model, and they should identify what they would consider changing or at least provide some idea of their priorities when refining their model.

Too many teams simply add an extra section titled "Strengths and Weaknesses" and then provide a bulleted list with brief sentences. This is certainly good; but to do better, the team should demonstrate that they have thought deeply about their model and can identify a deep understanding of the model that they created.

Parameter Fitting

Every year, teams develop models that include a large number of parameters. In this year's event, the teams were asked to examine a specific system and provide insight into the system. In this context, it is not appropriate to non-dimensionalize the system and provide an analysis of the overall behavior of the system.

A number of teams spent considerable time to determine the proper values of the parameters to use. Some of the choices came from examining a variety of papers whose focus is on the spread of the Ebola virus. In these cases, it could be difficult for the judges to determine how appropriate the choice of values are given that the original paper likely made use of a different model, and the values of the parameters are not necessarily the same. In these cases, though, the judges tended to give the teams the benefit of doubt considering the time constraints.

Other teams made use of various data sources and attempted to determine the best choice of the parameters, using a variety of approaches. Many teams used relatively straightforward least-squares estimation. This is a difficult task, and the judges made every effort to give the teams credit for their efforts. The primary issue is that the teams provide a strong written description of what they did and provide citations as well as references for other works they used.

Sensitivity

Every year, we say that sensitivity is important. Every year, this has been an area of the modeling process that tends to receive the least attention. This year was different. For the first time, we saw a large number of papers that provided an organized strategy to explore the sensitivity of the teams' models.

The majority of the teams' efforts focused on examining how the conclusions changed based on small changes in one or more parameters. This is an extremely important part of the modeling process, and the appearance of this simple exploration is an exciting development. We recognize that the time constraints make a complete sensitivity analysis difficult, but exploring the impact that different parameters have on the model is a relatively straightforward task.

In addition to the exploration of the parameters, for the first time we saw multiple teams that provided a brief analysis of the sensitivity of a model to basic assumptions. Several papers examined the impact of changing one or more assumptions, changing their model, and then examining how the conclusions change. This is an exciting development and represents a nontrivial leap in the modeling process.

The teams that were able to conduct this kind of analysis explicitly demonstrated a deep understanding of the modeling process. If this trend continues, then it will soon become vital that a team conduct some kind of basic sensitivity analysis in order to be recognized in the highest tier of the papers. This year's event required that teams make important predictions. It is important for the teams to provide some level of insight into how much trust can be placed in their predictions, and a sensitivity analysis is an important tool to determine how robust a model is.

Miscellaneous Comments

We survey a variety of unrelated but common issues that we observed. These issues include the use of figures and annotation of figures, citations and references, and the adaptation of existing models. The problems associated with these aspects of writing and modeling are important topics that represent stumbling blocks for a large number of teams. These are areas that advisers can discuss before the competition, and treatment of them can make a big difference in how a team's work will be received by a judge.

Figures

Every year we see large numbers of teams that provide a long list of figures but provide little to no discussion about the figures. The assumption is that a person reading the paper will understand what is important about a figure and will interpret the figures in the same way the teams interpret the figures.

Every figure should be described and discussed in the narrative of the report. The team should explain what is in the figure and why it is important. For example, this year it was common to see multiple approximations for the number of people infected by the Ebola virus from different regions, but it was not always clear what conclusion could be drawn from approximations that were not clearly labeled.

Additionally, every figure should have a title. Every figure should have a caption. The axes should be clearly labeled including units. A large number of teams provide plots without labels on the axes!

Use of References and Citations

Many teams include a list of references at the end of their report. A minority of teams provide citations within the narrative. To present ideas and not clearly provide the source of the motivation and where the idea originates from is plagiarism. A paper that provides clear citations and consistent references stands apart from many of the other reports.

How a Model is Adapted from Other Models

This is a source of concern every year, but this year it was a bigger issue than usual. Most of the models presented for the modeling of the spread of the Ebola virus made use of standard models and then adapted the model to accommodate the specific circumstances in this case.

A large number of teams simply assumed that the reader is aware of those standard models and that those models require little explanation. However, it can be difficult to understand a paper when a model is simply stated with little motivation or little discussion.

When a model is provided, the different parts of the system should be discussed, and the individual terms should be discussed. In doing so, the team has an opportunity to impress the judges and let us know that they understand the model. If a model is stated with no overview and no discussion, then we cannot determine if it is copied from some other source or has been adapted from other sources. We also cannot determine if the team understands the model or is simply restating something that they happened to find.

Discussing the model and providing motivation for how a model is adapted results in a paper that is much easier to read. When the paper is easier to read, and has proper citations, it immediately makes a strong impression on a judge. In the early stages of the judging, such a paper sends a message that the paper should be read in more detail. In the later stages of the judging, it sends a message that the team was respectful of the work of others, and they have a deeper understanding of the work.

Conclusion

The Eradicating Ebola Problem required teams to examine the connection between two models. Each topic on its own—disease modeling and distribution of resources—is a topic of intense interest in the research community. Many teams were able to build on the existing work and extend existing models. A smaller number of teams were able to bring the disparate models together into a cohesive whole.

In addition to the creation of the models, the teams had to determine values for the resulting parameters. This was done in a variety of ways, and teams that were able to discuss their approach and included particular details were more likely to make a stronger impression on the people reading their work.

Another important aspect of modeling is the analysis of the resulting model. In this year's event, we were pleasantly surprised to see that teams provided a much stronger analysis of their models than in the past. In particular, many teams provided a sensitivity analysis with respect to parameters, and a small number of teams went further with a sensitivity analysis with respect to their assumptions.

Finally, as is the case every year many teams had difficulties in using figures and graphs to share their insights about their results. Also, a large number of teams failed to provide citations within their narrative as well as consistent references. Additionally, many teams struggled with how to discuss their model and share how they adapted and extended other models.

References

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

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

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

Centers for Disease Control and Prevention (U.S.). 2014. Generic Ebola response (ER): Modeling the spread of disease impact & intervention. Version 2.3. `http://dx.doi.org/10.15620/cdc.24900`. 11 September 2014. Last accessed 27 June 2015.

Meltzer, Martin I., Charisma Y. Atkins, Scott Santibanez, Barbara Knust, Brett W. Petersen, Elizabeth D. Ervin, Stuart T. Nichol, Inger K. Damon, and Michael L. Washington. 2014. *Estimating the future number of cases in the Ebola epidemic—Liberia and Sierra Leone, 2014–2015. Morbidity and Mortality Weekly Report (MMWR)* Supplements September 26, 2014 / 63(03); 1-14. `http://www.cdc.gov/mmwr/preview/mmwrhtml/su6303a1.htm`. Last accessed 27 June 2015.

Murray, J.D. 2008. *Mathematical Biology: I: An Introduction*. 3rd ed. New York: Springer-Verlag.

About the Author

Kelly Black is a faculty member in the Department of Mathematics and Computer Science at Clarkson University. He received his undergraduate degree in Mathematics and Computer Science from Rose-Hulman Institute of Technology and his daster's and Ph.D. degrees from the Applied Mathematics program at Brown University. He has wide-ranging research interests including laser simulations, ecology, and spectral methods for the approximation of partial differential equations.