THE CONSORTIUM FOR
MATHEMATICS AND
ITS APPLICATIONS

2019 MCM Problem C
Triage Judging Guidelines

Purpose and Background

The Mathematical Contest in Modeling (MCM) and ***The Interdisciplinary Contest in Modeling (ICM)*** both rely on a Triage Judging and Final Judging process to identify seven classes of participant performance – Disqualified, Unsuccessful Participant, Successful Participant, Honorable Mention, Meritorious, Finalist, and Outstanding – based on the technical reports submitted under the MCM/ICM contest rules (http://www.comap.com/undergraduate/contests/mcm/instructions.php).

Final Judging is the culminating process whose purpose is to identify the papers that will be designated as Finalist or Outstanding. Prior to this event, Triage Judging in the United States and China efficiently ranks and categorizes each MCM/ICM submission based on desirable characteristics and content established over years of MCM/ICM operations, and upon the peculiarities of each problem.

Overall, the MCM/ICM supports and advocates an iterative mathematical modeling process consisting of major elements that include:

- Problem Restatement,
- Assumptions & Justifications,
- Model Construction and Application,
- Model Testing and/or Sensitivity Analysis,
- Analysis of Strengths & Weaknesses.

Papers that contain major elements in sufficient detail to address the problem posed receive higher recognition from judges.

The triage process relies on the professional expertise, experience, and judgment of academic faculty and industry professionals supporting the administration of the MCM/ICM to read and recognize key quality indicators in team papers – proper applications of mathematics and science, depth of exploration, completeness of a recognized modeling process, proper reliance upon and documentation of supporting research, innovative and insightful modeling approaches, and clear and concise exposition, among others. As noted frequently in UMAP Journal articles, elements such as these are universally valued among modelers, and are hallmark ingredients expected of top papers,.

While it is impossible to list all potential contributors to such quality, it is possible to note items that, if not present, will limit a paper's quality from the viewpoint of the MCM/ICM. While not claiming to be all-encompassing in its attempt to identify such limiting criteria, this memo will hopefully provide enough information to create a consistency in judgment despite extreme geographical and temporal separation of triage sites.

Triage Judging

In the Triage round we seek to cull out papers that do not have a chance at being Meritorious or Outstanding. In general we have percentage target levels we are planning to keep for final judging, but we always have to make some adjustments as we go through the process.

The judging of student team submissions uses seven paper classifications: Disqualified, Unsuccessful Participant, Successful Participant, Honorable Mention, Meritorious, Finalist, and Outstanding. The classification for each MCM/ICM paper is relative to the pool of papers received each year.

General guidelines and percentages for each category are as follows.

Disqualified (DQ) (% as warranted): the team's paper was found to be in violation of the contest rules. The rule violation should be noted in the comment column (e.g. "plagiarism," "same as paper xxx") and scored as 0. The MCM/ICM contest directors will review and verify all papers designated as disqualified (DQ).

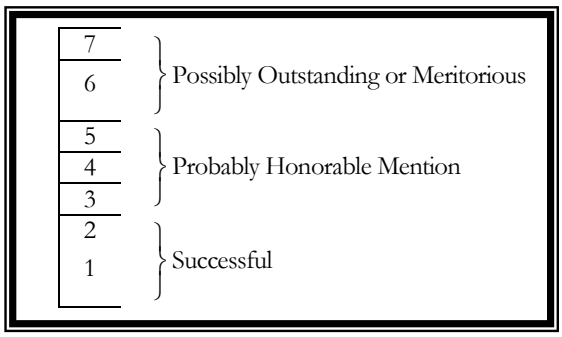
Unsuccessful Participant (UN) (% as warranted): the team's paper did not respond to any of the requirements of the contest problem, but did not violate any of the contest rules. A simple explanation of the failings should be noted in the comment column (e.g. "No modeling") and scored as 0. The MCM/ICM contest directors will review and verify all papers designated as unsuccessful (UN).

Successful Participant (P) (% as warranted): the team made an attempt at the problem and successfully submitted their paper. However, their overall paper is best described as fair to average, and possibly contains an incomplete modeling process or solution, and mathematical or logical errors. Scored as a 1 or 2.

Honorable Mention (HM) (% as warranted): the team submitted a complete, acceptable modeling approach and solution, but their solution contains at least one detractor, deficiency, or error that prevents it from being classified as Meritorious or Outstanding. Scored as 3, 4 or 5.

Meritorious (M) (10%): the team's paper represents an exemplary modeling approach, but their solution may contain minor errors or issues in logic, calculation, modeling, or assumptions. Scored as a 6 or 7.

Finalist (F)/Outstanding (O)(<1%): the team's paper represents an excellent modeling approach and a solid solution, including demonstrating an ability to clearly and concisely communicate their process, results, and conclusions. Scored as a 7.



Triage judging sessions are designed to accomplish a crude categorization of the papers. In the time allotted to each paper (no more than 15-20 minutes and 10-15 minutes on average), judges assess whether the required elements of the modeling process are addressed, and whether the teams have answered the questions posed in the problem statement. Judging during the triage sessions use a 7-point scale shown to the left to achieve the desired categorization. What makes this possible is the mindset of triage judges to primarily look for the very top papers and not try to fully evaluate every part of every paper. Don't be afraid to establish separations between quality levels with these numerical scores. Each paper is given 2 reads during the triage judging, so each judge is responsible for half the triage decision to have the paper forwarded to final judging or not.

The head judge culls approximately 60-80% of the papers after the triage round and take the remaining papers to final judging.

Triage Judging Notes

If you find a paper you are assigned to read is missing, damaged or incorrect, note the paper number and notify your head judge so that COMAP can check for the correct paper.

If you find that a team included any distinguishing information such as school name or student names, read the paper as normal and grade as normal, but add a note to the comment column (e.g. "includes school name on page xxx," "includes student name on page yyy").

If you find that a paper has gone over the assigned paged limit, read the paper as normal and grade as normal, but add a note to the comment column (e.g. "paper exceeded the assigned page limit").

Triage judges are encouraged, but not required, to include comments on their grading sheet. It could be as simple as a few words (e.g. "great assumptions"), or a sentence justifying the papers score (e.g. "fatal logic flaw on page zzz").

Problem Specific Guidance: Problem C (2019):

The core of the 2019 Problem C is a modeling problem seeking to characterize and understand the spread of opioid use in space and time as a behavior threatening to become a major epidemic among the U.S. population. The intent of the problem is to have teams build models capable of rolling forward and back in time to potentially identify future opioid conditions (if left unabated) and potential locations in space and time where such use/abuse originated in the targeted five (5) U.S. states, respectively. The NFLIS data is provided for precisely this modeling building effort as it contains evidence of evolving use/abuse for each county in all five states. For the sake of triage guidance, we'll refer to this model as the "Spread Model."

Once such a model is constructed and verified, teams are asked to identify potential socio-economic factors from data collected by the U.S. Census Bureau that exhibit, in some mathematically defendable manner, associations with behavior(s) noted in The Spread Model, or demonstrate that no such associations exist. Most likely, teams will attempt to do this using statistical methods, but judges should be open to other approaches as well.

If and when such factors are identified from the U.S. Census data provided, teams are asked to modify their Spread Model appropriately with the purpose of exploring the possibility that adding these factors might shed insights into potential solution strategies to the crisis. Call this model the "Complete Model." There is no predetermined number or type of factors that could be used in this manner. It is up to MCM teams to identify and support those used.

It is this Complete Model that is intended to be used by teams to generate and test potential solution strategies for the opioid crisis as described in this problem. Papers that do so are candidates for higher MCM award classification on this problem as they would represent the purpose of the Data Insight MCM problem.

There is an extensive body of literature on various types of deterministic and stochastic models that could be used to construct a Spread Model in support of the questions posed in Part 1. Diffusion models, epidemic models, network graphs, and other mathematical models that effectively describe space/time associations between and among geospatially organized data are likely candidates to appear. The problem does require significant data preparation and integration, as data from different databases are provided for exclusive use by the teams. Selected elements and guidance for triage:

Part 1: Developing the Spread Model

"Using the NFLIS data provided, build a mathematical model to describe the spread and characteristics of the reported synthetic opioid and heroin incidents (cases) in and between the five states and their counties over time."

- Teams that fail to build this Spread Model can score no higher than a 2 in triage.
- Teams that fail to address both time and geographical dimensions can score no higher than a 3 in triage.
- Teams that present good graphical results should be scored higher than teams that do not.

"Using your model, identify any possible locations where specific opioid use might have started in each of the five states"

- Teams that fail to use their Spread Model to address this element can score no higher than a 2 in triage.
- This is the problem requirement that is asking teams to 'roll their model backward over time" as described earlier. Answers to this question can be based on model output or behavior analysis or on descriptive statistics relating geospatial location-specific opioid use data to time.

"If the patterns and characteristics your team identified continue, are there any specific concerns the U.S. government should have? At what drug identification threshold levels do these occur? Where and when does your model predict they will occur in the future? "

- Teams that fail to use the output and/or behavioral characteristics of their Spread Model to answer these questions can score no higher than a 2 during triage.
- This is the problem requirement that is asking teams to 'roll their model forward in time' to identify or predict critical information that might come to pass in the future. The patterns identified by the teams can be at county, state, or multi-state levels of aggregation because patterns and thresholds that teams consider important here are subjective modeling options, obligating teams to mathematically support.
- Teams that include an uncertainty analysis in their predictions should receive higher scores than those that do not.
- Teams that fail to address any of these elements (patterns, concerns, thresholds, predictions) can score no higher than a 2 in triage.
- Teams that fail to address all these elements (patterns, concerns, thresholds, predictions) can score no higher than a 1 in triage as this represents a major use purpose for their Spread Model.

Part 2: Developing the Complete Model

"Is use or trends-in-use somehow associated with any of the U.S. Census socio-economic data provided?"

- A "yes" or "no" answer supported by statistical or mathematical analysis is required.
- Teams that select U.S. Census data elements without mathematical or statistical support can score no higher than a 3 during triage.

"If so, modify your model from **Part 1** to include any important factors from this data set."

- Implicit in this Part 2 is a requirement to compare model behavior and results of this Complete Model with those of the Spread Model developed in Part 1. In doing so, the same guidance supplied for Part 1 applies for Part 2 with the addition of socio-economic explanatory variables.
- Teams recognizing that the problem answers previously supplied using only the Spread Model might have changed when socio-economic factors are included in the Complete Model should score higher than those that do not.
- The choice of important socio-economic factors should be justified based on previous mathematical or statistical analysis.
- Teams that fail to support their choice of important socio-economic factors can score no higher than a 4 during triage.

Part 3:

"Finally, using a combination of your **Part 1** and **Part 2** results, identify a possible strategy for countering the opioid crisis."

- This element requires teams to synthesize previous elements of the problem, blending the results of both the Spread Model and the Complete Model as appropriate. Any mitigation or solution strategy suggested by teams must be connected to model analysis and results, be plausible, and be feasible.
- Teams that provide model-based rationales will score higher in triage.
- Teams that suggest mitigation or solution strategies not related to their model analysis or results should be scored lower than teams that do.

"Use your model(s) to test the effectiveness of this strategy; identifying any significant parameter bounds that success (or failure) is dependent upon."

- Teams that fail to analyze the effectiveness of their proposed strategy using their Complete Model can score no higher than a 4 during triage.

Memo to the Chief Administrator: "In addition to your main report, include a 1-2 page memo to the Chief Administrator, DEA/NFLIS Database summarizing any significant insights or results you identified during this modeling effort."

- This element requires the students to summarize and interpret the results of their modeling effort for a relatively non-technical audience. It implies the use of the NFLIS data and the demographic and economic variables included in the problem, and some general familiarity with the five states.
- The insights or results presented in the memo should follow from the team's answers to Parts 1 and 2.
- Teams that fail to include this Memo can score no higher than a 2 in triage.
- Teams that fail to include their major results or recommended strategy can score no higher than a 3 in triage.
- Teams that suggest actions not related to their model should be scored lower than teams whose suggestions are supported by their modeling and analysis.
- Poorly written summaries should be scored lower than ones that are well written.

Required elements from Contest Instructions:

- One-page Summary Sheet

The contest instructions say: "…a summary should clearly describe your approach to the problem and, most prominently, your most important conclusions. Summaries that are mere restatements of the contest problem or are a cut-and-paste boilerplate from the introduction are generally considered to be weak."

Besides the summary sheet as described each paper should contain the following mathematical modeling elements:

- Problem Restatement
- Assumptions & Justifications
- Model Construction and Application
- Model Testing and/or Sensitivity Analysis (note for Probem C: statistical confidence intervals or hypotheses tests may satisfy this)
- Model Revision
- An analysis of Strengths & Weaknesses.

These need not appear in stand-alone sections, but must be addressed somewhere in the paper for the team submission to be considered a complete modeling effort. Papers that omit any of the five elements above can be scored no higher than a 4 during triage.

The Final Judging Sessions (for information only for triage graders)

The final judges develop a rubric for each problem and customize it to the problem being judged and the set of papers present. After the triage event, judges have a better idea of how the top papers they have read are addressing the problem and what elements are evolving to set papers apart from each other. This knowledge provides the basis for refining the rubric prior to the last judging session to pick the Outstanding papers. We usually have 4 or 5 rounds of final judging where approximately 50% of the papers are culled in each round.

Patrick Driscoll
MCM Director