Judges' Commentary: The Outstanding Leaf Problem Papers

Peter Olsen, P.E.
Commander, US Coast Guard Reserve
Baltimore, MD
pcolsen@gmail.com

A manager would rather live with a problem he cannot solve than accept a solution he does not understand. —Robert E.D. "Gene" Woolsey [2003]

Introduction

Problem A of the 2012 Mathematical Contest in Modeling (MCM)™ was written by Lee Zia, who posed a challenging problem, "How can you measure the weight of leaves on a tree?" and several equally challenging subproblems. The problems were easy to state, but there were no traditional approaches. Successful teams would have to combine existing models, data, and new ideas in creative and original ways.

The results were gratifying. The judges were impressed by the variety of approaches submitted by the teams. The approaches were creative and the models showed each team's ability to use their own new ideas to refine and extend work that had gone before.

No two of the Outstanding papers shared the same model. Some share parts and data; but those are emphasized, combined, and used in different ways. Existing work, often quickly findable on Google, forms the scaffold on which each team built their own model. The final structures were a pleasure to behold.

Problem Statement

"How much do the leaves on a tree weigh?" How might one estimate the actual weight of the leaves (or for that matter any other parts of the tree)?

The UMAP Journal 33 (3) (2012) 223–229.

How might one classify leaves? Build a mathematical model to describe and classify leaves. Consider and answer the following:

- Why do leaves have the various shapes that they have?
- Do the shapes "minimize" overlapping individual shadows that are cast, so as to maximize exposure? Does the distribution of leaves within the "volume" of the tree and its branches effect the shape?
- Speaking of profiles, is leaf shape (general characteristics) related to tree profile / branching structure?
- How would you estimate the leaf mass of a tree? Is there a correlation between the leaf mass and the size characteristics of the tree (height, mass, volume defined by the profile)?

In addition to your one-page summary sheet, prepare a one-page letter to an editor of a scientific journal outlining your key findings.

Data

For some of the subproblems, such as the leaf classification, real data could be obtained easily. For others, such as the calculation of the mass of the leaves of the tree, it was difficult or impossible to obtain real data. In these latter cases, the teams showed creativity finding and using secondary sources.

Criteria for Judging

Here are some of the issues that kept papers from the final rounds:

- Errors in mathematics, which quickly took them out of further consideration.
- Including mathematics that didn't fit the flow of the presentation. In a few cases, mathematics appears to have been inserted to make a paper look more credible or to take the place of other work that had led to a dead end.
- Changing notation, sometimes even within a single section.
- Using undefined or poorly defined symbols, or using symbols before defining them.
- Incomplete expressions, either because the team made an error or because the expression did not survive the word-processor. (One of the Outstanding papers addressed in this commentary had a few incomplete expressions, probably because they didn't survive the word-processor,

but the coherence of its model and the strength of its presentation overcame that defect.)

Modeling Issues

This problem required two different types of model:

- *Increasing abstraction:* The leaf classification model abstracted from an immense number of natural leaf characteristics a set of artificial ones small enough to be useful for classification.
- *Decreasing abstraction:* The leaf mass problem took abstract models, applied them to data, and got concrete numerical results

Some models were difficult to understand; poor writing was the most common cause. Another cause was the use of inapposite mathematics. If the mathematics was a result of a "drive-by" insertion, fitting it into the model could be difficult.

Here are a few of the modeling issues that hurt some papers' chance of entering the final rounds:

- Questionable, conflicting, or unjustifiably speculative assumptions. Good papers did not assume any spherical cows ("a metaphor for highly simplified scientific models of complex real life phenomena" [Wikipedia 2012]).
- Dependence on *deus ex machina*: an assumption, equation, reference, or procedure invoked without explanation or context. Often the invocation would start with the phrase "It is well-known that..." It may be well-known to those who know it well, but that is unlikely to be the customer or client.
- Confusing, missing, or misplaced model definitions; model definitions are more complex and more important than mathematical ones, since they must not only name the *definiendum* but also specify what it is and what it is to be used for.
- Failure to reach a conclusion.
- Conflicting subproble models with unexplained conflicts between assumptions.
- Unexplained inconsistencies in data.
- An unclear, incomplete, or unrepresentative letter to the journal editor.
- A poor abstract:
  - too much detail, so much that it was difficult to see the overall structure of the model or the strategy for using it; or

  - too little detail, so that it was difficult for the reader to what was actually to be done; or
  - an incomplete abstract, presenting only part of the problem.

- Poor presentation, including bad prose style, poor vocabulary, and disorganized explanations. Good presentation won't get a bad paper into the finals, but poor presentation may keep a good one out. (The weight given to this criterion varies among the judges.)

Letter to a Journal Editor

The one-page letter to a journal editor was an important part of the problem. Its goal was to give insight into whether or not the teams could explain their results clearly, simply, and directly. The most important criterion of modeling is whether or not the models are used, either to increase understanding directly (through use) or indirectly (through publications, conferences, or professional tools such as software). A model that cannot be understood will not be used (see the quotation from Woolsey [2003] at the head of this commentary). A good letter should present an overview of the problem, technique, and results in a single page. The clarity of each team's letter is one indication of how their model might fare in the real world.

The Outstanding Papers

Hong Kong Baptist University

This team's entry was nicely laid out and easy to follow. The tree-classification models appeared to be traceable back to the first principles of physics.

Each model's development began with a clear description of the approach the team intended to follow. For example, in the leaf classification subproblem the approach was to reduce all leaf structures to a one of several polar coordinate functional shapes. These easily can be distinguished.

The team's solution to the problem of finding the mass of leaves on a tree was unique. The team used the structural properties of the tree, not properties of the tree canopy directly. The advantage of this approach is that the team did not need any information about the size or density of the canopy, the properties of individual leaves, or the number or distribution of the leaves. Knowing each branch's modulus of elasticity and its deflection under load provided enough information so that its leaf load could be inferred from the branch's deflection. Conceptually, this solution was much simpler than most of the others. As a practical matter, users of this solution might find difficulty in obtaining some of the data, such as the deflection of an unloaded branch; but if they could, this would an efficient and elegant solution.

The presentation was excellent for all models. The prose, graphics, and equations flowed seamlessly throughout the paper.

The team's letter to the editor was the paper's one weakness. The team employed a very high-level approach, laying out the overall goals for the problem, but without giving insight into the models' operational details.

Shanghai Foreign Language School

This paper had a particularly strong beginning. Within the space of three pages, the team

- reorganized the problem into four consolidated subproblems,
- stated their assumptions clearly and succinctly, and
- provided a table listing their model's parameters and their symbols.

The team's leaf classification model used seven simple measurement procedures involving 10 parameters, the most complicated of which is area. The measurements can be conducted on-site using only a sheet of fine-ruled graph paper. Only one parameter requires calculation: division of the area of a fractional leaf segment by the leaf's entire area. (In times past, this could have been done by eye with a simple nomograph. Now people will stop and key data into calculators.)

As with the team from Hong Kong Baptist University, the model for estimating leaf mass has an unusual approach. The model does not rely on direct measurements of leaf characteristics or tree size. This can be used to show that leaf-mass and tree size are correlated. The challenge in using this model is determination of the rate of sequestration of carbon-dioxide. The model uses sequestration data from a U.S. Department of Energy document.

The last section of the paper contained a clear and well-organized summary list of each problem's strengths and weaknesses.

The team's letter to the editor was clear and concise. It covered the high-level statement of the problem, then gave enough detail of the solution plan that an knowledgeable but non-expert reader could feel conversant with the approach.

National University of Singapore

This team's leaf classification algorithm is the simplest of the four described in this commentary. It has four steps:

- project the leaf onto a grid;
- determine the grid squares covered by the projection to determine if the leaf has convexities:

- If it convex, it is a palmate leaf, exit;
- if it is not convex, then perform further classification.

The leaf mass is calculated based on the team's vector tree model of tree-structure and their insolation model. The vector tree model represents a tree as three-dimensional vectors; daughter branches are obtained by applying a linear transform to the parents.

This entry made excellent use of graphics in presenting their models and results.

This team's letter to the editor successfully wove their research, their results, and their ideas about further research into a single clear narrative.

Zhejiang University

This paper presented a neural-net-based leaf classifier that was most sophisticated of all of the leaf classification schemes. The input layer had 4 nodes, the middle layer 10 nodes, and the output layer had 1 node.

The team divided a sample of leaves into four classes. They trained the network on 32 exemplars of each class, then tested the network on 8 other leaves drawn at random from the entire ensemble.

The network misclassified 1 of the 8. In general, it's impossible to tell how a back-propagation reaches its results; but it's reasonable to hypothesize that more training data might have corrected the one misclassification.

The leaf mass estimation was the most traditional of these four papers. It was based directly on the leaf mass constant, a known value that varies with tree species, and an estimate of the volume of an approximating regular solid.

Summary

These four solutions had strong similarities—importantly, not in the solutions themselves. Models work when they provide understandable bases for reasonable decisions. All four solutions met that criterion and several others:

- They were presented clearly.
  - The descriptive text was clear. There were comparatively few errors in grammar, vocabulary, or style; and these didn't interfere with the reader's understanding.
  - Graphics were appropriate and clear. They supported the argument being made. The appropriate text referred to them.
- The models were appropriate to the problem to be solved, in that
  - the assumptions and goals were clearly stated;

– the physics was correct and appropriate—there were no *dei ex machina* or spherical cows;
– there was no extraneous mathematics air-dropped into the model—the solution was organized in sections; and
– the graphics were easy to find.

Reference

Arney, Chris, and Kathryn Coronges. 2012. Modeling for crime busting. *The UMAP Journal* 33 (3) (2012) 291–302.

Wikipedia. 2012. Spherical cow. `http://en.wikipedia.org/wiki/Spherical_cow`.

Woolsey, Robert E.D. 2003. *Real World Operations Research: The Woolsey Papers*. Marietta, GA: Lionheart Publications.

Acknowledgments

This paper benefitted from insights in the Judges' Commentary by Chris Arney and Kathryn Coronges [2012] in this issue.

About the Author

A graduate of the U.S. Coast Guard Academy, Peter Olsen retired from the Coast Guard Reserve as a Commander in 1997, after 27 years service, active and reserve. His most challenging assignment was to build the quantitative model used to allocate resources for the *Exxon Valdez* oil-spill cleanup. Of the model, Vice Admiral Robbins, the on-scene coordinator, wrote that it was completed on time, it was used by the people who paid for it, and its predictions were borne out by events.