The UMAP Journal

Vol. 38, No. 3

Vol. 38, No. 3 2017

Table of Contents

Publisher's Editorial

The Doug Faires Awards for 2017

Solomon A. Garfunkel
Executive Director
COMAP, Inc.
175 Middlesex Turnpike, Suite 3B
Bedford , MA 01730–1459
s.garfunkel@comap.com

Introduction

COMAP is proud to announce the winners of the second annual Doug Faires Award. The purpose of the award is to encourage and recognize efforts to start modeling teams at both the high school and college levels.

COMAP wishes to encourage current faculty advisors to reach out, recruit, and mentor new faculty advisors at either the college or high school level, particularly at nearby schools. The goal is to form local groups with a common interest in mathematical modeling.

We dedicate this award to Doug Faires, who provided us with the perfect example of the goals we wish to attain. First, a snapshot of Doug Faires:

About Doug Faires

Doug served first as a faculty advisor for the undergraduate Mathematical Contest in Modeling (MCM)™. He also gave talks to local high schools, inviting them to form modeling teams to compete in the High School Mathematical Contest in Modeling (HiMCM)™. He recruited and mentored high school faculty advisors and invited them and their teams to Youngstown State University, where the teams met one another and the experienced members of the college teams mentored the high school students. Long-term bonds were formed, and each year college and high school teams were encouraged to participate in the modeling contests, including the Interdisciplinary Contest in Modeling (ICM)™. Additionally, teams were given feedback by Doug and others after the contest was over. Later, Doug served as a Final Judge for the MCM, where he again was a true leader.

The UMAP Journal 38 (3) (2017) 241–250.

Our Goal for the Award

Our goal is to emulate Doug's success at the local level.

The Doug Faires Award will be given for individuals who achieve great results in a particular year or cumulative excellent results over a period of time.

This Year's Awardees

This year's **Lifetime Achievement Award goes to Daniel J. Teague of the North Carolina School of Science and Mathematics (NCSSM)**. In 2017, he had two Outstanding teams in the ICM, in the Airport Security and Migration to Mars problems; his past achievements are detailed below.

The award for a **Single Year's Achievement goes to Prof. Xiaofeng Gao of Shanghai Jiao Tong University** for being the advisor to 52 teams in the 2017 MCM/ICM, 6 of which were Outstanding.

Both of this year's Doug Faires Award recipients have extremely interesting stories to tell about their experiences in mathematical modeling, as students and as mentors. We note below their achievements; we also asked them to describe a little of their backgrounds and practices.

Lifetime Achievement: Daniel J. Teague

His Background

Dan is an Instructor of Mathematics at the North Carolina School of Science and Mathematics (NCSSM), where he has taught since 1983. He received his undergraduate degree from the University of North Carolina at Chapel Hill, a Master's of Education from Springfield College, and a Ph.D. in Mathematics Education from North Carolina State University.

Dan has served as Second Vice President of the Mathematical Association of America (MAA) and twice as the Governor-at-Large for Secondary Teachers on the MAA Board of Governors. He is currently At-Large Delegate for the Board of Directors of the National Council of Teachers of Mathematics (NCTM). His committee work includes the National Research Council's Committee on Programs for Advanced Study of Mathematics and Science in American High Schools, two terms as a member of the U.S. National Commission on Mathematics Education, and the Advanced Placement Statistics Test Development Committee.

Dan was recognized with the Presidential Award for Excellence in Mathematics Teaching, the W.W. Rankin Memorial Award for Excellence in Mathematics Education from the North Carolina Council of Teachers of Mathematics, the University of North Carolina Board of Governor's Award for

Excellence in Teaching, and was twice recognized with the Distinguished Educator and Mentor Award by the College of Physical and Mathematical Sciences at North Carolina State University.

Dan was the section editor of "Everybody's Problems" in COMAP's newsletter for high school mathematics teachers, *Consortium*; is a co-author of the texts *Contemporary Precalculus through Applications* and *Contemporary Calculus Through Applications*; and recently was a member of the writing team for the *GAIMME: Guidelines for Assessment and Instruction in Mathematical Modeling Education*.

My Path to Mathematical Modeling

The North Carolina School of Science and Mathematics (NCSSM) takes pride in having a strong focus on mathematical modeling in its mathematics curriculum. The mathematical modeling emphasis began early in the school's history. NCSSM opened in the fall of 1981 with 134 high school juniors from across North Carolina. The curriculum was quite standard for the early 1980s when I was hired in January of 1983. NCSSM had the great fortune to have Henry Pollak, then head of the Mathematics Division at Bell Labs, as a member of its Board of Trustees. Dr. Pollak talked with the department about mathematical modeling; but since none of us had any experience with modeling, his excellent advice fell on willing but uninformed ears.

In the summer of 1984, I had the privilege of attending the first Woodrow Wilson Summer Institute in Mathematics, which was focused on Statistics and Quantitative Literacy. This summer changed my life and—after a few years—the mathematics program at NCSSM. One of the highlights of the summer was an evening talk by Henry Pollak. The combination of the sessions in the Institute by statisticians Dick Scheaffer and Stu Hunter and the presentation by Henry opened my eyes to what modeling could mean in a high school curriculum.

The final cog fell into place at the Joint Mathematics Meetings in 1986. Frank Giordano ran an MAA Mini-Course on mathematical modeling that I attended. I also attended the presentation by one of the 1985 MCM Outstanding teams on their solution to the Animal Population Problem, which presentation introduced me to the MCM. I finally understood what Henry had been saying! When I returned to school, I asked if I could teach a course in Mathematical Modeling as one of the first elective courses offered at NCSSM. In 1988, we had an Outstanding paper in the MCM ourselves, and the students presented their work at the Operations Research Society of America meeting that spring.

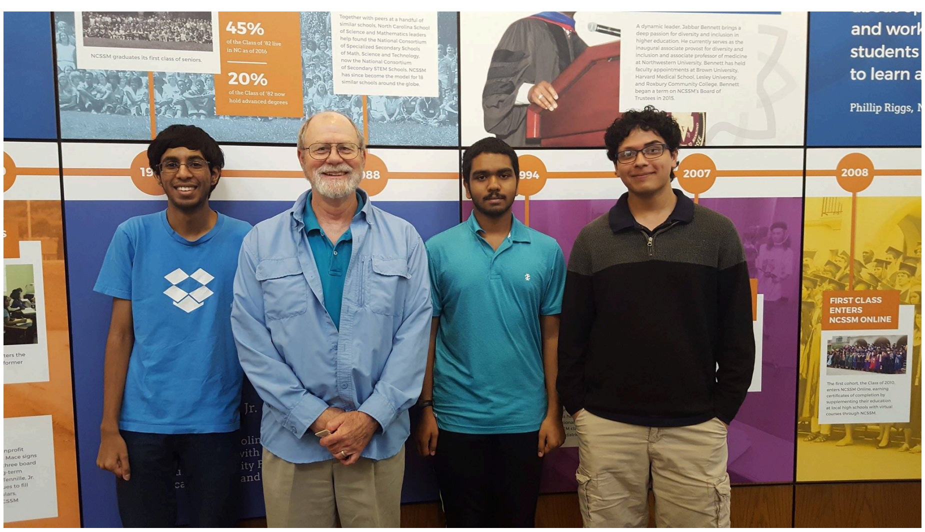
My understanding of the explanatory power of mathematical models was greatly enhanced by the opportunity to become part of the planning and instruction team for the Woodrow Wilson Summer Institutes in Mathematics for the next nine years. This allowed me to work with Henry for a month each summer over a six-year span and to interact with mathematicians such as Sol Garfunkel, Steven Brahms, Peter Hilton, Jean Pederson, Bill Lucas, Tom Banchoff, and Joe Malkevitch, and mathematics educators such as Kay Merseth, Kathy Heid, Landy Godbold, Alan Schoenfeld, and many others. I have been extraordinarily fortunate to have had this unique opportunity to learn from master practitioners.

Since that first course, my colleagues and I have taught a course in mathematical modeling to over 1,500 high school students and had over 100 teams competing in MCM/ICM, HiMCM, and Moody's Mega Math Challenge. In 2016, we offered four sections of Mathematical Modeling to more than one-fifth of the senior class.

In addition to the formal course in modeling, we offer complementary electives in Modeling with Differential Equations, Introduction to Complex Systems (where students are introduced to agent-based models), the Structure and Dynamics of Modern Networks, and Advanced Probability Models.

An NCSSM team wrote its first MCM Outstanding paper in 1988, and teams from NCSSM have participated every year since then. In the past five years, my students have written seven HiMCM Outstanding papers and six MCM/ICM Outstanding papers (including two INFORMS winners, an MAA winner, and two Vilfredo Pareto Award winners), eight top-6 finalists in the Moody's Mega Math Challenge (including two First, three Second, and one Third Place teams), and one Outstanding team in the International Mathematical Modeling Challenge.

NCSSM students have many opportunities to develop mathematical models in diverse areas of interest. COMAP, through the materials produced and interactions with the talented staff, has played a significant role in the successes of our program.



Dan Teague with team members Nikhil Reddy, Sreeram Venkat, and Nikhil Milind of the NCSSM team that was Outstanding in the 2017 ICM Migration to Mars Problem.

Single-Year Achievement: Xiaofeng Gao

Her Background

Xiaofeng Gao received the B.S. degree in information and computational science from Nankai University, China, in 2004; the M.S. degree in operations research and control theory from Tsinghua University, China, in 2006; and the Ph.D. degree in computer science from the University of Texas in 2010. She is currently an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. Her research interests include wireless communications, data engineering, and combinatorial optimization. She has published more than 110 peer-reviewed papers and 7 book chapters in related areas.. She has served on the editorial board of *Discrete Mathematics, Algorithms and Applications*, and as peer reviewer for a number of international conferences and journals.

My Experience

I have been participating in MCM/ICM for the last 15 years. In 2003 and 2004, I entered the competition as an undergraduate student; in 2004, my team won a Meritorious Winner award for our solution to the Quick Pass Problem, which was the highest award among all Chinese teams. Earlier, I related the details of that experience and the next several years of competitions in this *Journal* [Gao, Wu, and Wu 2013].

Due to this experience, many of my friends came to me for advice about MCM/ICM, which pushed me to follow the MCM and ICM problems over years. In 2011, I became an assistant professor in Shanghai Jiao Tong University. From then on, I took on the role as a team advisor supporting students in MCM/ICM. In the first year, I had only one team, with three students from my lab; but gradually more students sought my guidance. By now, I have been an advisor for MCM/ICM for six years.

This year, among the teams that I advised were 6 Outstanding Winners and 1 Finalist (including two INFORMS Awards and one Leonhard Euler Award).

It is a great honor to receive the Doug Faires Award from COMAP, since it serves as an affirmation of my previous endeavor and also as encouragement for future work. Thus, in the following section, I would like to share some of my experience in guiding MCM/ICM teams, in the hope of benefitting advisors and students in the future.

Xiaofeng Gao with team members of her Outstanding and Finalist teams in 2017.

How to Guide a Team for the MCM/ICM?

Frankly, the secret of our success consists of only three words:
be *novel*, be *professional*, and be *beautiful*.

Be Novel, Like a Scientist

At the very beginning of the MCM in 1984, when Ben Fusaro proposed a new competition parallel to the William Lowell Putnam Mathematical Competition, he hoped that students would use mathematical tools to explore real-world problems.

Grasping the essence of MCM/ICM, contestants should view themselves as scientists searching for the solution to open-ended application problems rather than as students sitting through an exam. Hence, in considering the problem and writing the paper, students should show their understanding of the problem, their line of thought, and how they tailored their solution to the specific problem, instead of introducing some general solution and then explaining the results.

Specifically, undergraduate students in Shanghai Jiao Tong University usually have the opportunity to join research labs or join the Participation in Research Program to learn how to become a researcher. They are trained to face unsolved problems, read contemporary literature, discuss state-of-the-art research, and think individually and independently. Such training helps a lot to improve the novelty of their solutions.

For instance, Yiming Zhang, member of the Outstanding team on the Zambezi River Dam Problem (also the INFORMS Award Winner) said:

> I never thought that laboratory research could play a role in this contest. However, my actual experience proved that modeling and researching can be interlinked. When our team was designing the framework of the model, a paper by some researchers at Microsoft Research came to mind [Yuan et al. 2016]. I had read it earlier when working on the topic prediction problem for social networks. Its hierarchical and modular structure seemed quite suitable for this modeling problem. Thus, during the design of our model, I learned from its ideas and also from the authors' schematic style.

Another example is Zhiying Xu, member of an Outstanding team on the ICM Airport Security Problem (also the Leonhard Euler Award Winner). She said:

> Most students easily think of cellular automata or a queuing model for the Airport Security Problem, which are dreary, clichéd, and lack novelty. However, I did research on multicast in ad-hoc networks in my sophomore year in the lab of Industrial Internet of Things (IIOT) at Shanghai Jiao Tong University. I read a lot of papers about scheduling optimization during that period. Although it was a year ago and my research focus has changed since then, the idea of Lyapunov optimization and the back-pressure algorithm [Neely 2010] dawned on me. Although the background was very different, the problem was essentially similar. Thus, I adapted these models and got an excellent result.

Novelty is one of the most important issues that makes a team stand out above the rest.

Be Professional, Like an Analyst

After the model construction, you've just finished only one-third of the task! The next step should be explaining the "superiority" of your design. Problem clarification, assumption justification, and model explanation are the first and foremost things to discuss. Next, theoretical proofs, numerical tests, visual analysis, sensitivity tests, case studies. . . : Many methods could be adopted to support your idea. The bottom line is to provide evidence for your approach.

At this point, you should be professional, like an analyst: Prove the efficiency, effectiveness, and priority of your approach, using any available means. Also, the negative points—including the things that you cannot accomplish, the future work that you may consider, and the weaknesses of your model—should all be carefully evaluated in your solution paper.

For instance, Yisen Yao, member of the Outstanding team on the Zambezi River Dam Problem (also the INFORMS Award Winner) said:

> Before entering the MCM/ICM Contests, I considered building the model as the most difficult step. However, after the competition, I

> realized that there is not a single gold standard for the model, as long as you can make a strong case for your solution. It is the process of collecting data, solving model formulas, and running simulations that makes all the difference. Facing great quantities of formulas and numbers, it is a huge challenge for participants to summarize their models with concise words and reasonable formulas. For example, in this contest, we used a wide range of data sources, such as sea-level maps from Google Earth and the data of average precipitation from the Zambezi River Authority. Additionally, we used MATLAB for our simulation, which makes our result more accurate and credible.

In other words, after the model construction, the most important thing is analysis and justification.

Be Beautiful, Like an Artist

Third, a good solution paper should be treated like a work of art. You should write the paper like an artist, paying attention to every possible aspect, especially the organization, the visualization, and the presentation.

For example, Duxing Hao, member of the Outstanding team on the Migration to Mars Problem (also the INFORMS Award Winner), said:

> When it comes to the highlights of our paper, illustration should surely be the first. Given our paper's title ("Society Planning: Model, Simulation and Visualization"), we spent lots of time on "visualizing" numerical results [see **Figure 1**.]. Our principle is that by reading the summary part and just the images, readers should be able to grasp the idea, the model, and all essential results of our paper. To the basic caption for a figure, we added direct conclusions from the figure and pointers to corresponding paragraphs giving detailed discussion. These reader-centered features really helped in organizing our paper and clarifying the structure. Besides, one should always read published articles in the relevant research area and imitate their figure styles and paper organization as much as possible.

Every part matters in the contest: not just the main body, but the title, the abstract, the outline, the references, and if required, the letter, advertisement, or nontechnical report. In the Judges' Commentary on the Space Junk Problem in 2016, Catherine A. Roberts said: "Meritorious paper 47676 from Shanghai Jiao Tong University was the exemplar in regard to citations" [Roberts 2016] (Team 47676 was also one of my teams that year). This shows that every detail may leave a deep impression on the judges, bringing more possibilities of the Outstanding award.

Be novel, be professional, and be beautiful. I hope that in the future, more students and advisors will participate in this intellectually challenging yet rewarding competition and enjoy their journey.

Sample in database			
Column	data	Explaination	Unit
1	10976	ID	No unit
2	20	AGEP*	(Years old)
3	2	SEX*	1=M, 2=F
4	18	SCHL*	(1,2,...24)
5	50000	Productivity*	($/year)
6	42	WKHP*	Hrs/week
7	5	RAC1P*	No unit
8	3	MAR*	No unit
9	2540	OCCP*	No unit
10	48000	Income	($/year)
11	0	Flag Of Death(FOD)	0=no risk, 1=in risk, 2=dead
12	1	Flag Of Need of spouse(FON)	0=no need, 1=in need
14	2	Flag Of Working status(FOW)	0=laid off, 1=employed, 2= parental leave
15	11003	Spouse's ID	No unit
17	1	Flag Of Marriage status(FOM)	0=single, 1=married
18-20	11253	Child's ID	No unit
21	0.6	Happiness Index	No unit
22	2	Number of Children	No unit
24	8992	Father's ID	No unit
25	9001	Mother's ID	No unit

*: Corresponding data of Population Zero can be found in PUMS database.

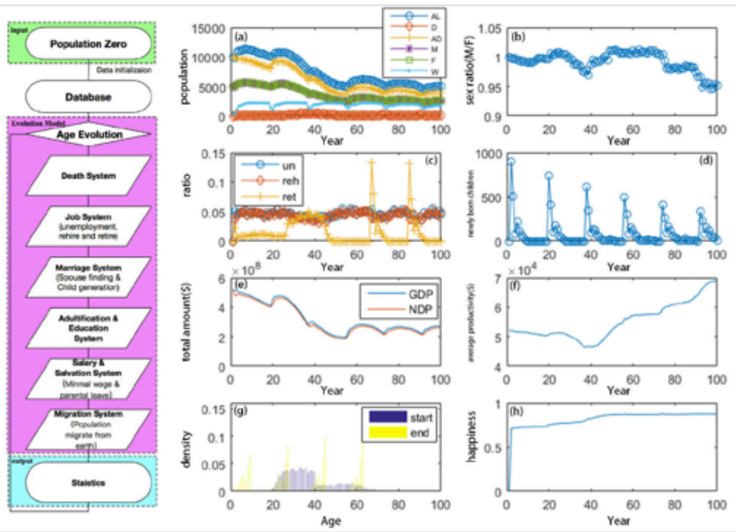


Figure 12: Database structure, algorithm and a sample result of the model.(a)population of certain kind of people each year(AL=alive,D=dead,AD=adult,M=male,F=female,W=people whose income under minimal wage); (b)sex ratio of male/female each year; (c)rate of unemployment(un), rehire(reh) and retire(ret) each year; (d)population of newly born children each year; (e)Gross Domestic Product(GDP) and Net Domestic Product(NDP) each year; (f)GDP per capita each year; (g)histogram of age distribution of Population Zero(start) and population after 100 years(end); (h) Average happiness index each year; All person's information in Population Zero corresponds to a person in reality(Data source:[1]); Other parameters: minimal wage 5000$/year; Parental leave pay is 40000$/year;retirement age is 65; rate of natural unemployment is 5 %; NO migration from earth after year 0;

Figure 1. The visualization example from Duxing Hao's team for the Migration to Mars Problem.

Acknowledgement

I give my sincere thanks to all Outstanding and Finalist teams from Shanghai Jiao Tong University, including:

Team 55280: Yiming Zhang, Jiachen Sun, Yisen Yao (Outstanding Winner and INFORMS Award Winner for the Zambezi River Dam Problem)

Team 57659: Xiangyu Liu, Zhenghui Wang, Xiwei Hu (Finalist for the Zambezi River Dam Problem, with advisor Prof. Yishuai Niu)

Team 55583: Junhao Xu, Tianling Bian, Qinghao Liu (Outstanding Winner and INFORMS Award Winner for the Self-Driving Cars Problem)

Team 55278: Jiefeng Chen, Qi Li, and Yu Shi (Outstanding Winner for the Self-Driving Cars Problem)

Team 55585: Qinghao Mao, Zhanghao Wu, Junpeng Hu (Outstanding Winner for the Self-Driving Cars Problem)

Team 55295: Zhiying Xu, Peitian Pan, Xiaoxing Wang (Outstanding Winner and Leonhard Euler Award Winner for the Airport Security Problem)

Team 55285: Yikai Huo, Zhiyu You, Kan Chang (Outstanding Winner for the Airport Security Problem)

Team 55491: Yunyi Yang, Xiaoqing Geng, Baicheng Xiang (Finalist for the Sustainable Cities Problem)

Team 64486: Duxing Hao, Hao Yu, Yitong Ou (Outstanding Winner and INFORMS Winner for the Migration to Mars Problem, with advisor Prof. Yishuai Niu)

Solomon Garfunkel of COMAP presenting the Doug Faires award to Xiaofeng Gao at Shanghai Jiao Tong University, June 15, 2017.

The Future

The Doug Faires Award will be given to individuals who achieve great results in a particular year or attain cumulative excellent results over a period of time. Xiaofeng Gao and Dan Teague exemplify the goals of the Doug Faires award. Awardees receive a certificate of appreciation that expresses our enduring gratitude.

We would like you to nominate someone in your community (including possibly yourself!) who is promoting mathematical modeling at the local level.

References

Gao, Xiaofeng Guoliang Wu, and Meng Wu. 2013. My decade with the MCM. *The UMAP Journal* 34 (2–3): 151–161.

Neely, Michael J. 2010. Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks* 3 (1): 1–211.

Roberts, Catherine A. 2016. Judges' commentary: The space junk papers. *The UMAP Journal* 37 (3): 301–304.

Yuan, Nicholas Jing, Yuan Zhong, Fuzheng Zhang, Xing Xie, Chin-Yew Lin, and Yong Rui. 2016. Who will reply to/retweet this tweet? The dynamics of intimacy from online social interactions. In Proceedings of the 9th ACM International Conference on Web Search and Data Mining (WSDM'16), San Francisco, CA, U.S.A., February 22–25, 2016. `https://pdfs.semanticscholar.org/167d/a2ad2b01a43566b21a4715d1315ff7ac45ef.pdf`.

MCM Modeling Forum

Results of the 2017 Mathematical Contest in Modeling

Patrick J. Driscoll, MCM Director
United States Military Academy
West Point, NY
chi.decisions@gmail.com

Introduction

A total of 8,843 teams of undergraduates from hundreds of institutions and departments in 14 countries spent a weekend working on applied mathematics problems in the 33rd Mathematical Contest in Modeling (MCM)®.

The 2017 MCM began at 8:00 P.M. EST on Thursday, January 19, and ended at 8:00 P.M. EST on Monday, January 23. During that time, teams of up to three undergraduates researched, modeled, and submitted a solution to one of two open-ended modeling problems. Students registered, obtained contest materials, downloaded the problems and data, and entered completion data through COMAP's MCM Website. After a weekend of hard work, solution papers were sent to COMAP on Monday. Three of the top papers appear in this issue of *The UMAP Journal*, together with commentaries from the contest judges.

A companion contest, the Interdisciplinary Contest in Modeling (ICM)®, took place concurrently over the same weekend. The ICM offers modeling problems involving network science, human-environment interactions, and policy modeling. Details about the 2017 ICM Contest and its results are in Volume 38, Number 2 of this *Journal*.

The 2018 MCM/ICM Contests will take place February 8–12, 2018.

COMAP is pleased to announce support from the American Statistical Association (ASA) for the MCM/ICM, joining the other major mathematics professional societies. For the past two years, the MCM has included a Data Insights Problem (Problem C) that should be of particular interest to students of statistics. Given the support of ASA for our efforts, we look forward to increased participation from students and faculty in departments of statistics here and abroad.

The UMAP Journal 38 (3) (2017) 251–268.

Description of the Problems

This year, the three MCM problems represented interesting scenarios for contestants, each offering a dimension of mathematical modeling that was unique. The authors of Problems A, B, and C are Dr. Kelly Black, Dr. Michael Tortorella, and Prof. David Olwell, respectively.

Problem A addressed options associated with a dam in the Kariba Gorge of the Zambezi river basin between Zambia and Zimbabwe that is suffering from aging effects that potentially threaten the safety of downriver populations. Teams were challenged to provide an overall assessment of three options and then delve into a detailed modeling and analysis of a replacement scheme in which multiple smaller dams would be constructed to replace the existing dam. Lifecycle costs, electrical power generation, water management policies, and environmental impacts were among the problem characteristics that teams were required to address. This was a particularly thought-provoking problem with multiple avenues along which teams could become lost if not careful.

Problem B revisited an issue that time, technology, and engineering have yet to resolve to satisfaction—diffusion and convergence of multi-lane traffic on two sides of a highway toll collection facility. However, rather than posing a situation in which performance assessment of a particular configuration was sought, Problem B instead asked teams to consider if there existed some yet unidentified design that offered improved performance in comparison to existing designs. While challenging in its own right, added complexity was introduced by asking teams to consider the impact of self-driving vehicles, human versus automatic toll collectors, and surges of heavy traffic.

Finally, Problem C asked teams to assess the impact of self-driving vehicles on a particularly dense traffic network in the Greater Seattle area. As with previous C problems, a significant amount of real-world data was provided for team use. Teams were to determine patterns and insights upon which to construct a mathematical model to analyze the effects on traffic flow of the number of lanes, peak and/or average traffic volume, and percentage of vehicles using self-driving, cooperating systems.

As in past years, teams submitted clever and insightful approaches that judges found fascinating.

An Outstanding paper from each of the problems is featured in this issue of the *Journal*, together with commentaries by the judges.

All of the competing teams are to be congratulated for their excellent work and enthusiasm for mathematical modeling and interdisciplinary problem solving.

Resources for Mathematical Modeling

COMAP, whose educational philosophy is centered on mathematical modeling, supports the use of mathematical concepts, methods, and tools to explore real-world problems. COMAP serves society by developing students as problem solvers in order to become better informed and prepared as citizens, contributors, consumers, workers, and community leaders. The MCM is an example of COMAP's efforts in achieving these goals.

In addition to this special issue of *The UMAP Journal*, COMAP offers at `www.mathmodels.org` the press releases for the 2017 contests, their results, their problems, unabridged versions of all the Outstanding papers, and judges' commentaries.

Results and winning papers from previous contests were published in special issues of *Mathematical Modeling* (1985–1987) and *The UMAP Journal* (1985–2015). The 1994 volume of *Tools for Teaching*, commemorating the tenth anniversary of the contest, contains the 20 problems used in the first 10 years of the contest and an Outstanding paper for each year. That volume and the special MCM issues of the *Journal* for the last few years are available from COMAP. The 1994 volume is also available on COMAP's special *Modeling Resource* CD-ROM. Also available is *The MCM at 21* CD-ROM, which contains the 20 problems from the second 10 years of the contest, an Outstanding paper from each year, and advice from advisors of Outstanding teams. These CD-ROMs can be ordered from COMAP at

`http://www.comap.com/product/cdrom/index.html.`

Contest problems and results of the MCM/ICM contests are on the COMAP Website at

`http://www.comap.com/undergraduate/contests.`

The volume *Mathematical Modeling for the MCM/ICM Contests Volume 1* is an exposition of the ideas, background knowledge, and modeling methodologies for solving problems in the MCM/ICM contests.

That volume also presents a brief history of the MCM/ICM contests, offers ideas to help students prepare for the MCM/ICM contests, presents general modeling framework and methodologies, describes the judging procedure of the MCM/ICM papers, explains how to write successful MCM/ICM papers, and presents a sample scheduling of tasks during the contest. A number of exercise problems are included to help students understand the materials presented in the book.

Details and ordering are at

`http://216.250.163.249//product/?idx=1465.`

Finally, COMAP also makes available three volumes of *The Mathematical Modeling Handbook* on CD-ROM, with the first volume also available in print.

Details and ordering are at

`http://216.250.163.249//product/?idx=1467,` http://216.250.163.249//product/?idx=1348.

COMAP also sponsors:

- The MCM/ICM Media Contest (see p. 268).
- The Interdisciplinary Contest in Modeling (ICM), noted above.
- The High School Mathematical Contest in Modeling (HiMCM)®, which offers high school students a modeling opportunity similar to the MCM. Further details are at

  http://www.comap.com/highschool/contests.

2017 MCM Statistics

- 8,843 teams participated (with 8,085 more in the ICM)
- 311 U.S. teams (4%)
- 7,532 foreign teams (96%), from Australia, Canada, China, Hong Kong SAR, Indonesia, Macau SAR, Mexico, Scotland, Singapore, South Africa, South Korea, and the United Kingdom
- 13 Outstanding Winners (<1%)
- 23 Finalist Winners (<1%)
- 651 Meritorious Winners (7%)
- 3,540 Honorable Mentions (40%)
- 4,468 Successful Participants (49%)
- 28 Unsuccessful Participants (<1%)
- 120 Disqualified Teams (2%)

A Caution

A relatively large number of student teams (120) were categorized as Disqualified because of plagiarism and copying. The MCM expects contestants to be honest about their work. Submitted papers are expected to be the team's own effort; and when data, methodology, or ideas are used from others, which unto itself is good scholarship, credit must be carefully and clearly given to the other sources of those data, methods, and ideas. The MCM requires teams to be scrupulously honest in their research and presentation.

Problem A: Managing the Zambezi River

The Kariba Dam on the Zambezi River is one of the larger dams in Africa. Its construction was controversial, and a 2015 report by the Institute of Risk Management of South Africa included a warning that the dam is in dire need of maintenance. A number of options are available to the Zambezi River Authority (ZRA) that might address the situation. Three options in particular are of interest to ZRA:

- (Option 1) Repairing the existing Kariba Dam,
- (Option 2) Rebuilding the existing Kariba Dam, or
- (Option 3) Removing the Kariba Dam and replacing it with a series of ten to twenty smaller dams along the Zambezi River.

There are two main requirements for this problem:

- **Requirement 1** ZRA management requires a brief assessment of the three options listed, with sufficient detail to provide an overview of potential costs and benefits associated with each option. This requirement should not exceed two pages in length, and must be provided in addition to your main report.
- **Requirement 2** Provide a detailed analysis of Option 3—removing the Kariba Dam and replacing it with a series of 10 to 20 smaller dams along the Zambezi river. This new system of dams should have the same overall water management capabilities as the existing Kariba Dam while providing the same or greater levels of protection and water management options for Lake Kariba that are in place with the existing dam. Your analysis must support a recommendation as to the number and placement of the new dams along the Zambezi River.

For Requirement 2, you should include a strategy for modulating the water flow through your new multiple-dam system that provides a reasonable balance between safety and costs. In addition to addressing known or predicted normal water cycles, your strategy should provide guidance to the ZRA managers that explains and justifies the actions that should be taken to properly handle emergency water flow situations (i.e., flooding and/or prolonged low-water conditions). Your strategy should provide specific guidance for extreme water flows ranging from maximum expected discharges to minimum expected discharges. Finally, your recommended strategy should include information addressing any restrictions regarding the locations and lengths of time that different areas of the Zambezi River should be exposed to the most detrimental effects of the extreme conditions.

Your MCM submission should consist of three elements: a standard 1-page MCM Summary Sheet, a 1–2-page brief assessment report (Requirement 1), and your main MCM solution (Requirement 2) not to exceed 20 pages, for a maximum submission of 23 pages. Note: Any appendices or reference pages you include will not count towards the 23-page limit.

Problem B: Merge After Toll

Multi-lane divided limited-access toll highways use "ramp tolls" and "barrier tolls" to collect tolls from motorists. A ramp toll is a collection mechanism at an entrance or exit ramp to the highway and these do not concern us here. A barrier toll is a row of tollbooths placed across the highway, perpendicular to the direction of traffic flow.

There are usually (always) more tollbooths than there are incoming lanes of traffic (see the 2005 MCM Tollbooth Problem at `http://www.comap.com/undergraduate/contests/mcm/contests/2005/problems/`). So when exiting the tollbooths in a barrier toll, vehicles must "fan in" from the larger number of tollbooth egress lanes to the smaller number of regular travel lanes.

A toll plaza is the area of the highway needed to facilitate the barrier toll, consisting of the fan-out area before the barrier toll, the toll barrier itself, and the fan-in area after the toll barrier. For example, a three-lane highway (one direction) may use 8 tollbooths in a barrier toll. After paying toll, the vehicles continue on their journey on a highway having the same number of lanes as had entered the toll plaza (three, in this example).

Consider a toll highway having L lanes of travel in each direction and a barrier toll containing B tollbooths ($B > L$) in each direction. Determine the shape, size, and merging pattern of the area following the toll barrier in which vehicles fan in from B tollbooth egress lanes down to L lanes of traffic.

Important considerations to incorporate in your model include accident prevention, throughput (number of vehicles per hour passing the point where the end of the plaza joins the L outgoing traffic lanes), and cost (land and road construction are expensive). In particular, this problem does not ask for merely a performance analysis of any particular toll plaza design that may already be implemented. The point is to determine if there are better solutions (shape, size, and merging pattern) than any in common use.

Determine the performance of your solution in light and heavy traffic. How does your solution change as more autonomous (self-driving) vehicles are added to the traffic mix? How is your solution affected by varying the proportions of conventional (human-staffed) tollbooths, exact-change (automated) tollbooths, and electronic toll-collection booths (such as electronic toll-collection via a transponder in the vehicle)?

Your MCM submission should consist of a 1-page Summary Sheet, a 1–2 page letter to the New Jersey Turnpike Authority, and your solution (not to exceed 20 pages), for a maximum of 23 pages. Note: The appendix and references do not count toward the 23-page limit.

Problem C: "Cooperate and Navigate"

Traffic capacity is limited in many regions of the United States due to the number of lanes of roads. For example, in the Greater Seattle area, drivers experience long delays during peak traffic hours because the volume of traffic exceeds the designed capacity of the road networks. Delays are particularly pronounced on Interstates 5, 90, and 405, as well as State Route 520, the roads of particular interest for this problem.

Self-driving, cooperating cars have been proposed as a solution to increase capacity of highways without increasing number of lanes or roads. The behavior of these cars interacting with the existing traffic flow and one another is not well understood at this point.

The Governor of the State of Washington has asked for analysis of the effects of allowing self-driving, cooperating cars on the roads listed above in Thurston, Pierce, King, and Snohomish counties. See the provided map in **Figure 1** and the Excel spreadsheet at

`http://www.comap.com/undergraduate/contests/mcm/contests/2017/problems/2017_MCM_Problem_C_Data.xlsx`.

In particular, how do the effects change as the percentage of self-driving cars increases from 10% to 50% to 90%? Do equilibria exist? Is there a tipping point where performance changes markedly? Under what conditions, if any, should lanes be dedicated to these cars? Does your analysis of your model suggest any other policy changes?

Your answer should include a model of the effects on traffic flow of the number of lanes, peak and/or average traffic volume, and percentage of vehicles using self-driving, cooperating systems. Your model should address cooperation between self-driving cars as well as the interaction between self-driving interest, provided in the Excel spreadsheet.

Your MCM submission should consist of a 1-page Summary Sheet, a 1–2-page letter to the Governor's office, and your solution (not to exceed 20 pages), for a maximum of 23 pages. Note: The appendix and references do not count toward the 23-page limit.

Some useful background information:

- On average, 8% of the daily traffic volume occurs during peak travel hours.
- The nominal speed limit for all these roads is 60 miles per hour.
- Mileposts are numbered from south to north, and west to east.
- Lane widths are the standard 12 feet.
- Highway 90 is classified as a state route until it intersects Interstate 5.
- In case of any conflict between the data provided in this problem and any other source, use the data provided in this problem.

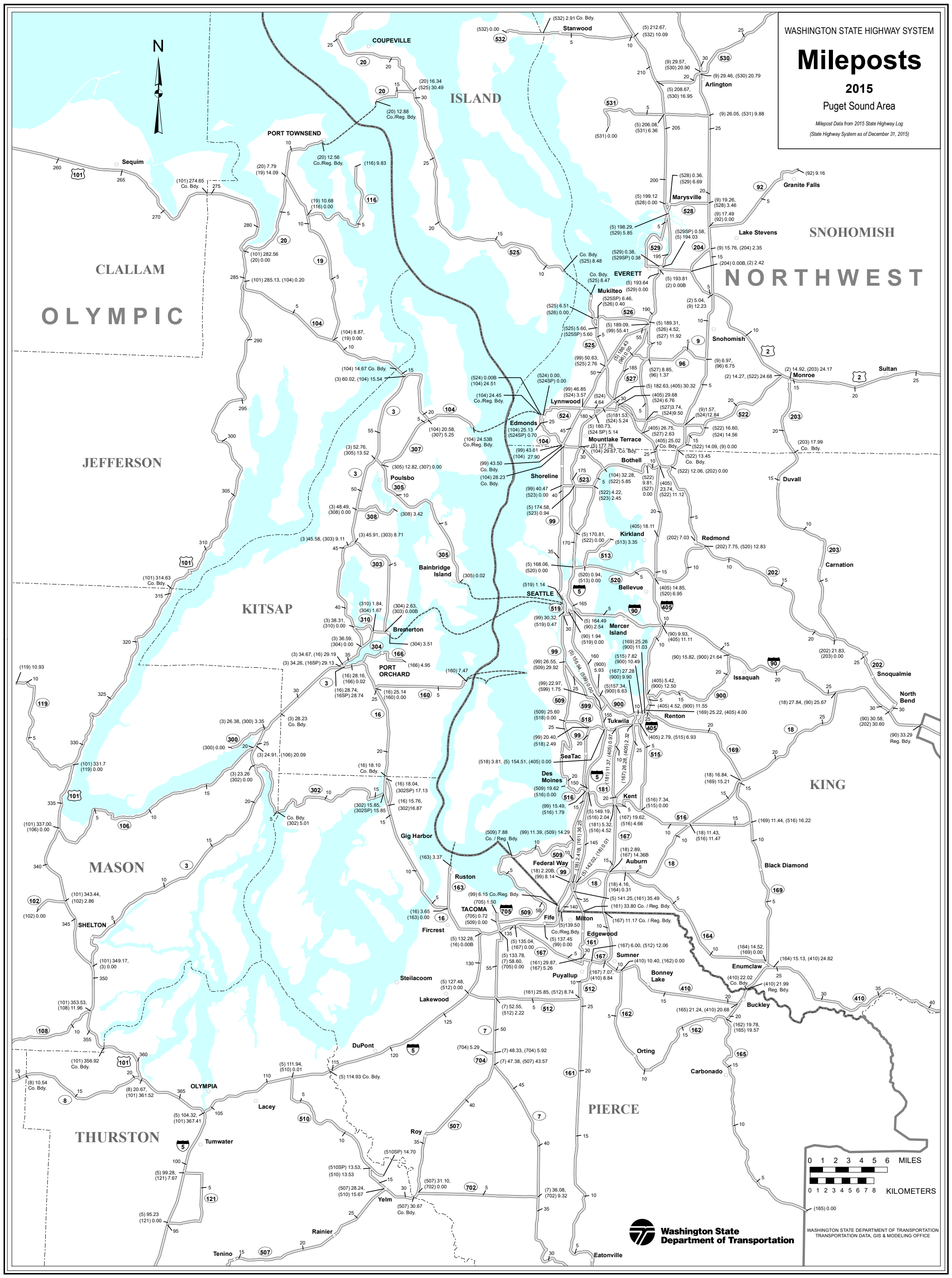


Figure 1. Map for Problem C.

Definitions:

- **milepost:** A marker on the road that measures distance in miles from either the start of the route or a state boundary.
- **average daily traffic:** The average number of cars per day driving on the road.
- **interstate:** A limited-access highway, part of a national system.
- **state route:** A state highway that may or may not be limited access.
- **route ID:** The number of the highway.
- **increasing direction:** Northbound for N-S roads, Eastbound for E-W roads.
- **decreasing direction:** Southbound for N-S roads, Westbound for E-W roads.

The Results

The solution papers were coded at COMAP headquarters so that names and affiliations of the authors would be unknown to the judges. Each paper was then read preliminarily by two "triage" judges at a triage session site. At each triage site, two time-limited comprehensive readings of each paper take place in comparison to a set of scoring criteria that establish minimum and maximum point allocations against a standard baseline. The scoring criteria are created by each problem's Head Judge in collaboration with the Contest Director and subsequently distributed to each triage site prior to the start of judging. These point allocations, in concert with established MCM quality elements, determine a paper's positioning within the overall pool of submissions for a particular problem. In the case that two triage judges' scores on a paper differ by more than 3 points, a third judge scores the paper.

Final judging took place in Carmel, CA.

The judges classified the papers as follows:

	Outstanding	Finalist	Meritorious	Honorable Mention	Successful Participation	Total
Zambezi River Problem	4	8	199	822	1,311	2,409
Merge After Toll Problem	5	8	301	2,179	2,357	4,907
Self-Driving Car Problem	4	7	151	539	800	1,527
	13	23	651	3,540	4,468	8,843

We list here the 13 teams that the judges designated as Outstanding; the list of all participating schools, advisors, and results is at the COMAP Website.

Outstanding Teams

Institution and Advisor	**Team Members**
Zambezi River Problem	
Southwestern University of Finance and Economics Chengdu, Sichuan, China Zekai He	Dan Li Junjun Luo Xinyu Liu
Shanghai Jiao Tong University Shanghai, China Xiaofeng Guo	Yiming Zhang Yiachen Sun Yisen Yao
Nanjing University Nanjing, Jiangsu, China Hong-Jun Fan	Yue Sun Qian Xu Qi-Zhi Cai
Nanjing University of Posts and Telecommunications Nanjing, Jiangsu, China Qinglun Yan	Jiacheng Li Han Cao Biao Huang
Merge After Toll Problem	
Shandong Normal University Jinan, Shandong, China Hui Ji	Jiahua Zhang Yang Liu Rui Liu
Chongqing University of Posts and Telecommunications Chongqing, China Qinghua Zhang	Guanyu Fu Liucheng Liao Yiqing Feng
Neijiang Normal University Neijiang, Sichuan, China Lian-Ming Mou	Sheng-Bing Hu Du Yan Yan Yu
North Carolina State University Raleigh, North Carolina, USA Ralph C. Smith	Graham Pash Jaye Sudweeks Davis Atkinson
University of Hong Kong Hong Kong SAR Kwok Fai Lam	Lexiao Lai Shaoxiong Zheng Shuyao Li

Self-Driving Cars Problem

University of California at Berkeley Berkeley, California, USA John Harte	Katherine Latimer Lianyuan Na Ruta Jawale
Shanghai Jiao Tong University Shanghai, China Xiaofeng Gao	Jiefeng Chen Qi Li Yu Shi
Shanghai Jiao Tong University Shanghai, China Xiaofeng Gao	Junhao Xu Tianling Bian Qinghao Liu
Shanghai Jiao Tong University Shanghai, China Xiaofeng Gao	Qinghao Mao Zhanghao Wu Junpeng Hu

Awards and Contributions

Each participating MCM advisor and team member received a certificate signed by the Contest Director and the appropriate Head Judge.

INFORMS, the Institute for Operations Research and the Management Sciences, recognized three Outstanding teams as INFORMS Outstanding teams: the teams from Shanghai Jiao Tong University (Zambezi River Problem), Neijiang Normal University (Merge After Toll Problem), and Shanghai Jiao Tong University (Self-Driving Cars Problem). INFORMS provided the following recognition:

- a letter of congratulations from the current president of INFORMS to each team member and to the faculty advisor;
- a one-year complimentary access to the full suite of INFORMS journals online for the faculty advisor;
- a crystal trophy for display at the team's institution, commemorating the team members' achievement;
- individual crystal trophies for the team members, as a personal commemoration of this achievement; and
- a one-year student membership in INFORMS for each team member, which includes their choice of a professional journal plus the *OR/MS Today* periodical and the INFORMS newsletter.

The Society for Industrial and Applied Mathematics (SIAM) designated two Outstanding teams as SIAM Winners. The SIAM Award teams were from Nanjing University of Posts and Telecommunication (Zambezi River Problem) and North Carolina State University (Merge After Toll Problem).

Each team member was awarded a $300 cash prize. Their schools were given framed hand-lettered certificates in gold leaf.

The Mathematical Association of America (MAA) designated one North American Outstanding team as an MAA Winner. The team was from the University of California at Berkeley (Self-Driving Cars Problem)). Each team member was presented a certificate by an official of the MAA Committee on Undergraduate Student Activities and Chapters.

Ben Fusaro Award

One Meritorious, Finalist, or Outstanding paper is selected for the Ben Fusaro Award, named for the Founding Director of the MCM and awarded for the 13th time this year. It recognizes an especially creative approach; details concerning the award, its judging, and Ben Fusaro are in Vol. 25 (3) (2004): 195–196. The Ben Fusaro Award Winner was the Outstanding team from Nanjing University (Zambezi River Problem).

Frank Giordano Award

For the 5th time, the MCM is designating a paper with the Frank Giordano Award. This award goes to a paper that demonstrates a very good example of the modeling process in a problem featuring discrete mathematics—this year, the Merge After Toll Problem. Having worked on the contest since its inception, Frank Giordano served as Contest Director for 20 years. The Frank Giordano Award for 2015 went to the Outstanding team from Shandong Normal University.

Two Sigma Scholarship Award

The Two Sigma Scholarship Award went to the team from the University of California at Berkeley (Self-Driving Car Problem). They received a total of $10,000. This was the third year of this award, for which the 472 U.S. teams competing in the MCM and ICM were eligible for one of two such scholarship prizes. We thank Two Sigma Investments for making this award possible.

Judging

Director
Patrick J. Driscoll, Dept. of Systems Engineering, U.S. Military Academy, West Point, NY
Associate Director
William C. Bauldry, Chair-Emeritus, Dept. of Mathematical Sciences, Appalachian State University, Boone, NC

Zambezi River Problem

Head Judge
Kelly Black, Dept. of Mathematics, Clarkson University, Potsdam, NY
Associate Judges
Karen Bolinger, Dept. of Mathematics, Clarion University, Clarion, PA (MAA Judge)
Tim Elkins, Dept. of Systems Engineering, U.S. Military Academy, West Point, NY (INFORMS Judge)
Jerry Griggs, Dept. of Mathematics, University of South Carolina, Columbia, SC
Paul Heiney, U.S. Military Academy Preparatory School, West Point, NY
Jack Picciuto, Director of Operations Analysis and Planning at IT Cadre, Ashburn, VA (SIAM Judge)
Elizabeth "Libby" Schott, Dept. of Mathematics, Florida SouthWestern State College, Fort Myers, FL
Jie "Jed" Wang, Dept. of Computer Science, Univ. of Massachusetts Lowell, Lowell, MA
Bill Wilhelm, Lockheed Martin Corporation, Huntsville, AL

Merge After Toll Problem

Head Judge
Maynard Thompson, Mathematics Dept., University of Indiana, Bloomington, IN
Associate Judges
Thomas Fitzkee, Dept. of Mathematics, Francis Marion University, Florence, SC
William P. Fox, Dept. of Defense Analysis, Naval Postgraduate School, Monterey, CA
Michael Jaye, Dept. of Defense Analysis, Naval Postgraduate School, Monterey, CA (SIAM Judge)
Greg Mislick, Dept. of Operations Research, Naval Postgraduate School, Monterey, CA
Dan Solow, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH (INFORMS Judge)
Michael Tortorella, Assured Networks, LLC, Middletown, NJ
Marie Vanisko, California State University, Stanislaus, Turlock, CA (MAA Judge)
Rich West, Emeritus Professor of Mathematics, Francis Marion University, Florence, SC (Giordano Award Judge)
Owen Xu, East China University of Science and Technology, Shanghai, China

Self-Driving Cars Problem

Head Judge
David H. Olwell, Professor and Dean, Hal and Inge Marcus School of Engineering, Saint Martin's University, Lacey, WA
Associate Judges
Robert Burks, Dept. of Defense Analysis, Naval Postgraduate School, Monterey, CA (INFORMS Judge)
James Enos, Dept. of Systems Engineering, U.S. Military Academy, West Point, NY
Richard Marchand, Mathematics Dept., Slippery Rock University, Slippery Rock, PA
Veena Mendiratta, Lucent Technologies, Naperville, IL
Scott T. Nestler, Dept. of Operations Research, Naval Postgraduate School, Monterey, CA
Katie Oliveras, Mathematics Dept., Seattle University, Seattle, WA
Rodney Sturdivant, Dept. of Mathematics and Physics, Azusa Pacific University, Azusa, CA
Catherine Roberts, Dept. of Mathematics and Computer Science, College of the Holy Cross, Worcester, MA
Wenbo Zhang, Beijing University of Posts and Telecommunications, Beijing, China

Judges at Triage Sessions in the U.S.

Zambezi River Problem, at Appalachian State University
Head Judge
William Bauldry, Appalachian State University, Boone, NC
Associate Judges
Bill Cook, Ross Gosky, Jeff Hirst, Lisa Maggiori, Greg Rhoads, René Salinas, and Jose Sanqui,
all from Appalachian State University, Boone, NC
Heidi Berger, Simpson College, Indianola, IA
Amy H. Erickson, Keith Erickson, Junkoo Park, and Kathy Pinzon,
all from Georgia Gwinnett College, Lawrenceville, GA
Steve Kaczkowski, South Carolina Governor's School for Science and Mathematics, Hartsville, SC
Harrison Schramm, CANA Advisors, Monterey, CA

Zambezi River Problem, at the U.S. Military Academy
Head Judge
Patrick J. Driscoll, US Military Academy, West Point, NY
Associate Judges
Tim Elkins, James Enos, Jacqueline Harris, Eugene Lesinski, Dan McCarthy, Ken McDonald, and Russell Schott,
all from the U.S. Military Academy, West Point, NY

Paul Heiney, U.S. Military Academy Preparatory School, West Point, NY
Steven Henderson, Citi Group
Jack Picciuto, IT Cadre, Ashburn, VA
Marie Samples, NYC Office of Chief Medical Examiner
Elizabeth Schott, Florida SouthWestern State College, Fort Myers, FL
Michael Tortorella, Assured Networks, LLC, Middletown, NJ

Merge After Toll Problem, at the Naval Postgraduate School
Head Judge
William Fox, Naval Postgraduate School, Monterey, CA
Associate Judges
Terry Cline, Terry Mullen, Marie Vanisko, and Ted Wendt,
all from Carroll College, Helena, MT
Rob Burks, Michael Jaye, and Greg Mislik,
all from the Naval Postgraduate School, Monterey, CA
Ahula Herat, Manuel Valera, and Tom Wakefield,
all from Slippery Rock University, Slippery Rock, PA
Brian Albright, Concordia University, Seward, NE
Jeremiah Bartz, North Dakota State University, Fargo, ND
Ethan Berkove, Lafayette College, Easton, PA
Jay Belanger, Truman College, Chicago, IL
David Hoeflin, Applied Research Data Mining, AT&T (retired)
Tom Fitzkee, Francis Marion University, Florence, SC
Rich Marchand, Salisbury University, Salisbury, MD
Tim McDevitt, Elizabethtown College, Elizabethtown, PA
Jack Picciuto, IT Cadre, Ashburn, VA
Karen Richey, US Government Accountability Office
Thomas Smoltzer, Youngstown State University, Youngstown, OH
Michael Tortorella, Assured Networks, LLC, Middletown, NJ
Rich West, Francis Marion University, Florence, SC
Bill Wilhelm, Lockheed Martin Corporation, Huntsville, AL

Self-Driving Cars Problem, at St. Martin's University
Head Judge
David Olwell, St. Martin's University, Lacey, WA
Associate Judges
Katie Oliveras, and J. McLean Sloughter, Seattle University, Seattle, WA
Mike Spivey, University of Puget Sound, Tacoma, WA
Edoh Amiran and Tilman Glimm, Western Washington University,
Bellingham, WA
Wai Lau, Seattle Pacific University, Seattle, WA
James Bisgard, Central Washington University, Ellensburg, WA
Carol Overdeep, St. Martin's University, Lacey, WA
Jakob Kotas, University of Portland, Portland, OR
Ryan Card and Olga Shatunova, University of Washington–Tacoma,
Tacoma, WA

Judges at Triage Session in China

Director
Jinxing Xie, Tsinghua University, Beijing
Associate Judges
Fengshan Bai, Tsinghua University, Beijing
Kangsheng Liu, Zhejiang University, Hangzhou
Yejin Lu, Guangxi University, Nanning
Xiaoyin Wang, Tianjin Polytechnic University, Tianjin
Hongli Yang, Shandong University of Science and Technology, Qingdao

Acknowledgments

Major funding for the MCM is provided by COMAP. Additional support is provided by the Institute for Operations Research and the Management Sciences (INFORMS), the Society for Industrial and Applied Mathematics (SIAM), and the Mathematical Association of America (MAA). We are indebted to these organizations for providing judges and prizes.

We also thank Two Sigma Investments LLC for providing its prize. "This group of experienced, analytical, and technical financial professionals based in New York builds and operates sophisticated quantitative trading strategies for domestic and international markets. The firm is successfully managing several billion dollars using highly-automated trading technologies. For more information about Two Sigma, please visit `http://www.twosigma.com`."

Finally, we thank for their involvement and unflagging support the MCM judges and MCM Board members, as well as the advisors to the competing teams.

Cautions

To the reader of research journals:

Usually a published paper has been presented to an audience, shown to colleagues, rewritten, checked by referees, revised, and edited by a journal editor. Each paper here is the result of undergraduates working on a problem over a weekend. Editing (and usually substantial cutting) has taken place; minor errors have been corrected, wording has been altered for clarity or economy, and style has been adjusted to that of *The UMAP Journal*. The student authors have proofed the results. Please peruse these students' efforts in that context.

To the potential MCM advisor:

It might be overpowering to encounter such output from a weekend of work by a small team of undergraduates, but these solution papers are highly atypical. A team that prepares and participates will have an enriching learning experience, independent of what any other team does.

COMAP's Mathematical Contest in Modeling and Interdisciplinary Contest in Modeling are the only international modeling contests in which students work in teams. Centering its educational philosophy on mathematical modeling, COMAP serves the educational community as well as the world of work by preparing students to become better-informed and better-prepared citizens.

About the Author

Patrick J. Driscoll is Professor of Operations Research in the Dept. of Systems Engineering at the U.S. Military Academy. Formerly an Academy Professor in the Dept. of Mathematical Sciences, he has also served as the Director of the Mathematical Sciences Center of Excellence and the Associate Dean of Information and Educational Technology. He received both an M.S. in Operations Research and an M.S. in Engineering Economic Systems from Stanford University, and a Ph.D. in Industrial and Systems Engineering from Virginia Tech. He is a member of the Operational Research Society (ORS) of the United Kingdom, the Institute for Operations Research and the Management Sciences (INFORMS), the Military Operations Research Society (MORS), and the honor societies Phi Kappa Phi and Pi Mu Epsilon. He is the Director for the Mathematical Contest in Modeling (MCM) and one of the designers of the High School Mathematical Contest in Modeling (HiMCM). He serves on the Board of Directors for the Hudson Valley Shakespeare Festival, is a partner in Winemates & Company, LLC, has three cats, and is continuing to have more fun than should be legally allowed.

Media Contest

Over the years, contest teams have increasingly taken to various forms of documentation of their activities over the grueling 96 hours—frequently in video, slide, or presentation form. This material has been produced to provide comic relief and let off steam, as well as to provide some memories days, weeks, and years after the contest. We *love* it, and we want to encourage teams (outside help is allowed) to create media pieces and share them with us and the MCM/ICM community.

The media contest is *completely separate* from MCM and ICM. No matter how creative and inventive the media presentation, it has *no* effect on the judging of the team's paper for MCM or ICM. We do not want work on the media project to detract or distract from work on the contest problems in any way. This is a separate competition, one that we hope is fun for all.

Further information about the contest is at

`http://www.comap.com/undergraduate/contests/mcm/media.html.`

There were 10 entries, 5 of them from Dalian Maritime University. (Come on, you other schools!)

Outstanding Winners of the 7th MCM/ICM Media Contest:

- Shandong University (Weihai), Weihai, Shandong, China
  (Mengran Wang, Yunxiao Jia, and Junxian Yin)
- Virginia Military Institute, Lexington, Virginia, USA
  (Jonathan Chu, Thayer Meyer, and James Chapman)

Finalists:

- Wuhan University, Wuhan, Hubei, China
  (Xinyu You, Mengzhen Jian, and Bei Zhou)
- Dalian Maritime University, Dalian, Liaoning, China
  (Yifan Liu, Haolin Han, and Hongyu Huang)

The remaining entries were judged Successful Participants. Complete results, including links to the Outstanding videos, are at

`http://www.comap.com/undergraduate/contests/mcm/contests/2017/solutions/index.html.`

Replacing the Kariba Dam

Yiming Zhang
Jiachen Sun
Yisen Yao

Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai 200240
China
Advisor: Xiaofeng Gao

Abstract

The Kariba Dam on the Zambezi River, built in the 1950s, holds back the world's largest reservoir. After so many years, despite efforts to slow its structural problems, the Kariba Dam is in great danger.

We provide a detailed analysis of the option of replacing the dam with a series of 10–20 smaller dams. We analyze different aspects, including the number, placement, and height of the new dams. We set the water management capabilities to be the same as the existing dam, providing protection and water management options for Lake Kariba.

We first analyze the water flow, using the Manning formula, for a single dam. After that, we simulate a simplified dam with the three-dimensional software 3ds MAX and derive its storage capacity and building cost.

Our overall model consists of three submodels: Risk Cost Model, Power Supply Ability Model, and Series Safety Model.

The Risk Cost Model implements risk analysis, using the acceptable risk ratio, to achieve a balance between cost and safety.

The Power Supply Ability Model measures the benefits that the dam system can produce.

The Series Safety Model measures how safe the whole system is, based on an ideal distribution of dams along a river.

We use the Analytic Hierarchy Process (AHP) to prioritize these three aspects and propose the overall model.

We use a genetic algorithm to determine the number and placement of dams, and we further determine the height of each dam.

We further discuss strategies for addressing several situations, including the balance between safety and costs, protection for Lake Kariba, guidance for emergency water flow situations, and extreme water flows.

Finally, we do sensitivity analysis and discuss strengths and weaknesses.

The UMAP Journal 38 (3) (2017) 269–288.

Introduction

Problem Background

The Kariba Dam (see **Figure 1**) provides the countries of Zimbabwe and Zambia with electric power.

Figure 1. Kariba Dam. Source: [1].

However, the dam is in dire need of maintenance or replacement [8]. Droughts have lowered the reservoir's volume to 12% of capacity; but if the reservoir were to fill again, the dam would very likely collapse.

Our Work

The three options for the dam are repair, rebuild, or replace (with 10–20 smaller dams).

We focus on the last option. We recommend a specific number and precise placement of replacement dams, balancing safety and costs. Our system has the same overall water management capabilities as the existing dam while providing greater levels protection and water management options for Lake Kariba.

We state several basic assumptions, provide a table of symbols and nomenclature, and then give details about our model. We then analyze aspects of our proposed model, point out strategies dealing with several situations, and reach some conclusions.

Assumptions

Our model makes the following assumptions:

- We don't consider extreme conditions on the river, such as waterfalls or water-cutoff seasons. We discuss emergency water flow situations later.
- Other existing dams on the river are ignored. They are not as big as Kariba Dam, and few data about them are accessible.
- We simplify the water flow as open-channel flow, which lets us estimate the average velocity of the river's flow.

Nomenclature

We use the nomenclature in **Table 1** to describe our model. Other symbols that are used only once are described later.

Table 1.
Table of symbols.

Symbol	Definition
i	the ith dam in a series of small dams
X_i	Distance along the river to dam i
h_i	Height of dam i
k_i	Slope of the hydraulic grade line at dam i
d_i	Distance from dam i to dam $(i+1)$
$Q_i(t)$	Volumetric flow rate at time t for dam i
$Series(X)$	Value of safety evaluation for a series X of dams
$Supply(X_i)$	Power supply ability at dam i
$Risk(X_i)$	Risk Cost of dam i
P_f	Acceptable risk ratio
N	Planned working years for a dam
n	Number of small dams
M	Overall supply of water to be provided
L	Length of the Zambezi River
w	Width of the Zambezi River
α	Extreme condition ratio
λ	Volume-to-expense coefficient for building a dam
γ	Cost of indirect damage if the dam fails, as a fraction of the dam cost

Description of the Model

Our model takes several dimensions into consideration, ranging from liquid flow theory to economic considerations.

We first investigate water flow.

Behavior of Water Flow

We need to calculate the ability of each dam to supply electric power. A key variable is the average velocity of water when it reaches a dam. We assume channel flow so can employ an empirical formula, the Manning formula, to estimate this velocity [3]:

$$v = \frac{k}{n} {R_h}^{2/3} S^{1/2}, \qquad \textbf{(1)}$$

where:

- v is the cross-sectional average velocity;
- n is the Gauckler-Manning coefficient, empirically derived for the river channel;
- R_h is the hydraulic radius;
- S is the slope of the hydraulic grade line, or the linear hydraulic head loss; and
- k is a conversion factor between SI and English units.

This formula can be obtained by use of dimensional analysis.

According to the formula, when water reaches a dam, its average velocity depends only on the channel slope S and hydraulic radius R_h.

Analysis of a Single Dam

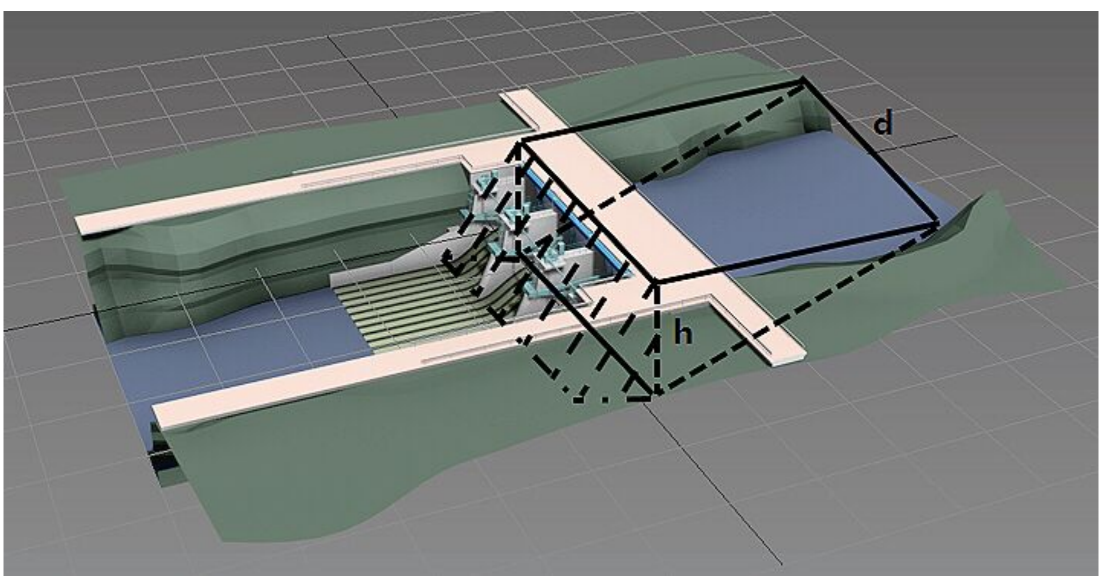


Figure 2. Simulation of a dam via the software 3ds MAX.

Figure 2 shows our model of a dam of height h and width d, with hydraulic grade having slope k.

Reservoir Capacity

For simplicity, we consider that a reservoir has the areas enclosed by contours in **Figure 2**, outlined in black. Thus, the reservoir capacity is the volume of the wedge shown, which is the area of the right triangle times the width w:

$$V = \frac{1}{2} h \cdot \frac{h}{k} \cdot d = \frac{h^2 d}{2k}.$$

Cost

Another key consideration is a dam's building cost. There are two parts: construction cost dependent on dam size and fixed costs. Construction cost is linear in the volume of the dam. To calculate this volume, we need the width of the bottom of the dam, which is linear in the pressure $P_a = \rho g h$ [4].

We need to take into account a safety factor for compressive stress σ, as follows [4]:

$$C = \lambda \frac{(\rho g h)(dh)}{2\sigma} + \beta = \lambda \frac{\rho g d h^2}{2\sigma} + \beta, \tag{2}$$

where λ is the volume-to-expense coefficient and we simplify the fixed costs to a constant β.

Dams in Series

Our integrated model consists of three main parts: Risk Cost Model, Power Supply Ability Model, and Series Safety Model. As illustrated in **Figure 3**, our model works as an integrated machine that evaluates both a single dam and the whole group of dams; basically, we can regard it as a nonlinear programming method.

The three main parts will be assigned weights according to their contributions to the whole system.

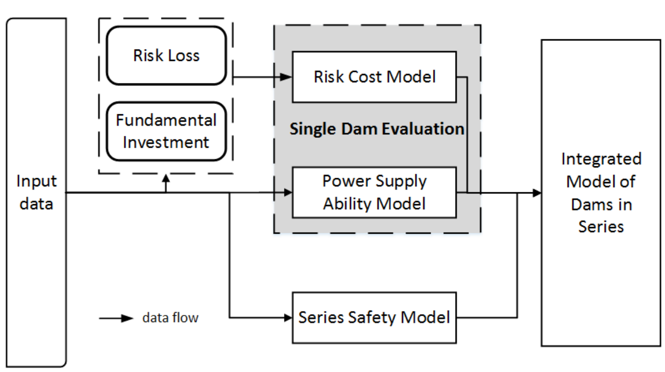


Figure 3. Overview of the integrated model of dams in series.

Prerequisites

- V_i, C_i, and k_i are the reservoir capacity, cost, and slope of the hydraulic grade at dam i, which is located distance at X_i down the river.
- $Q_i(t)$ is the volumetric flow rate for dam i at time t. For simplicity, we replace $Q_i(t)$ by $Q_i(\bar{t})$, using the annual average volumetric flow rate $\bar{t}$. We assume that the width d of the river is constant throughout the length of the river.
- We set the slope k_i of the hydraulic grade line at dam i to be

$$k_i = \left.\frac{\mathrm{d}\,\text{Elevation}(x)}{\mathrm{d}x}\right|_{x=X_i}.$$

Risk Cost Model

Two key factors for a dam are cost and safety. We combine them based on principles of risk analysis: We determine the Risk Loss and Fundamental Investment based on an acceptable risk ratio P_f, which measures the extent that a person focuses more on safety than merely on cost.

Cost: Fundamental Investment More safety costs more money. Thus, P_f is negatively related to the cost; it factors in with an exponent that is set by experiment. We slightly modify **(2)** to get

$$\text{Fundamental Investment} = \frac{C}{P_f^{1.1}}. \tag{3}$$

Safety: Risk Loss Projects that pay less attention to safety will result in a higher likelihood of collapse of the dam. We call it Risk Loss. We consider P_f to be the probability of a dam collapse in any one year. Then, over a service lifetime of N years, the probability of a dam collapse is $[1-(1-P_f)^N]$. We define the loss at a dam to be:

$$\begin{aligned}\text{Risk Loss} &= \text{collapse probability} \times (1 + \text{indirect damage index}) \times \\ &\qquad \text{volumetric flow rate} \\ &= [1-(1-P_f)^N]\alpha(1+\gamma)Q_i(\bar{t}),\end{aligned}$$

where N is the planned years of service time for the dam; α is the extreme condition ratio (a coefficient that reveals how likely and serious a river is to suffer from extreme conditions); and γ is an estimate of the cost of indirect damage if the dam fails, as a fraction of the dam cost.

This method contributes to the evolution of risk-based dam safety assessment methods. For a series X of n dams at locations X_i, we have

$$Risk(X) = \text{Risk Loss} + \text{Fundamental Investment}.$$

Figure 4 shows $Risk(X)$ as a function of P_f, for values of α varying from 0.5 to 1.0 in rising steps of 0.05.

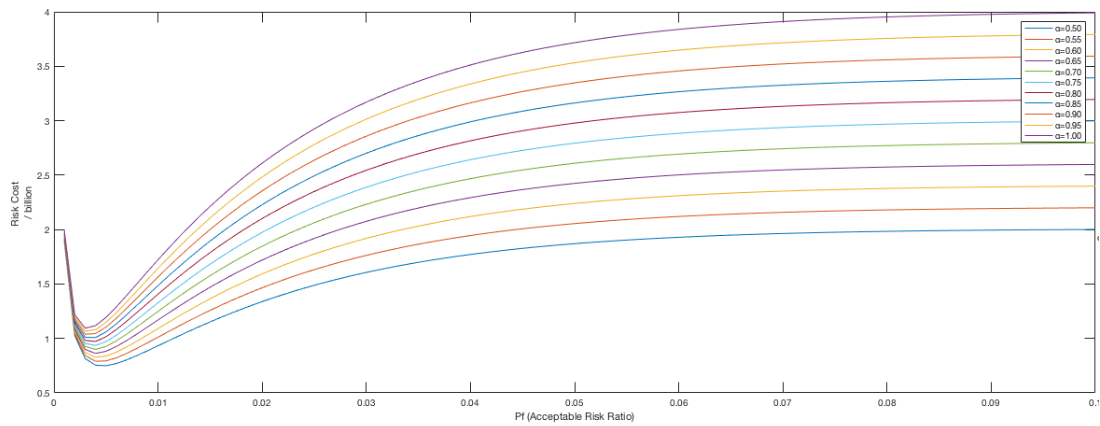


Figure 4. Risk Value vs. P_f for values of α from 0.5 (bottom curve) to 1.0 (top curve).

Observation 1 Each curve decreases to a minimum, then rises. Risk Value is not a monotonic function because **Risk Loss** is monotonically increasing in P_f while Fundamental Investment is monotonically decreasing in P_f.

Observation 2 The larger α, the higher the Risk Value, with the minimizing value of P_f decreasing slightly. A larger α means that for the same P_f, the government has to handle higher potential losses. Thus, to maintain the balance between cost and safety and minimize Risk Value, the government will have to use a lower value of P_f.

Series Safety Model

Our model removes the Kariba Dam and replaces it with 10–20 smaller dams. Research shows that if smaller dams are built uniformly along a river, the resulting system can produce greater benefits than operating a single large dam.

To quantify safety based on locations, we consider two factors: the average distance between two neighboring dams, and the degree of even distribution of the dams along the river. We use a modification of the Pearson chi-square formula to represent evenness of the distribution. The more even this distribution is, the safer we consider the system to be. Thus, for a system X consisting of n dams, with dam i located at distance X_i downstream, we use the following formula:

$$Series(X) = \frac{\overline{d}}{e^{\sum_{i=1}^{n-1} \frac{(d_i - \overline{d})^2}{\overline{d}^2}}},$$

where d_i is the distance from dam i to dam $i+1$ and $\overline{d}$ is the average distance between two neighboring dams.

Power Supply Ability Model

Kariba was designed as a hydropower project. We need to estimate the hydropower supply of our proposed system of dams. We illustrate the calculation for dam i.

How much work a unit of water can do depends on its weight times the head. And the power available from a dam can be measured by the density of water, acceleration due to gravity, the average flow rate $Q_i(\overline{t})$, and the height h_i of fall:

$$Supply_i = \eta \rho g h_i Q_i(\overline{t}), \tag{4}$$

where

- η is the dimensionless efficiency of the turbine,
- ρ is the density of water, and
- g is the acceleration due to gravity.

Then

$$Supply(X) = \sum_{i=1}^{n} \eta \rho g h_i Q_i(\overline{t}).$$

Integration

- $Series(X)$ is *positively* related to the safety level. The higher $Series(X)$, the better our system is.
- $Risk(X)$ is *negatively* related to the safety level. There is an equilibrium between investment and loss that reaches the lower bound of this model. We always want to save money while ensuring safety. So, the lower $Risk(X)$, the better our system is.
- $Supply(X)$ is *positively* related to power supply. The higher $Supply(X)$, the better our system is.

To put all features on an equal basis, we need scaling, or data normalization, because these features all have a broad range of values. We rescale the ranges to $[0,1]$, using the general formula to rescale an original value x to its rescaled value x':

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}.$$

We denote the rescalings with a prime ($'$) after the quantity. Since $Risk(X)$ is *negatively* related to the safety level; we modify it to be *posi-*

tively related by a modified rescaling:

$$Risk(X)' = \sum_{i=1}^{n} \frac{\min_i\big(Risk(X_i)\big)}{Risk(X)}.$$

We also have:

$$Series(X)' = Series(X) \times \frac{n-1}{L},$$
$$Supply(X)' = Supply(X) \times \frac{n-1}{L},$$

where the factor $(n-1)/L$ rescales the ranges of $Series$ and $Supply$, based on the $(n-1)$ distances between dams along the total river length L.

Our goal is to maximize Final Value:

$$\begin{aligned} \textbf{max} \quad & \text{Final Value} = Series(X)' + Risk(X)' + Supply(X)' \\ \textbf{such that} \quad & 0 \le X_i \le L, \qquad i \in [1, n] \\ & \sum_{i=1}^{n} V_i \ge M, \\ & 10 \le n \le 20, \qquad n \in \mathbb{Z}, \end{aligned}$$

where:

$$\begin{aligned} \text{Final Value} &= Series(X)' + Risk(X)' + Supply(X)' \\ &= \frac{\overline{d}}{e^{\sum_{i=1}^{n-1} \frac{(d_i - \overline{d})^2}{\overline{d}^2}}} \times \frac{n-1}{L} \\ &\quad + \sum_{i=1}^{n} \frac{\min_i(Risk(X_i))}{\frac{Cost(X_i)}{P_f^{1.1}} + [1 - (1 - P_f)^N]\alpha(1+\lambda)Q_i(\overline{t})} \\ &\quad + \sum_{i=1}^{n} \rho g \eta L k_i Q_i(\overline{t}) \times \frac{n-1}{L}. \end{aligned}$$

Ranking Submodels with AHP

The Analytic Hierarchy Process (AHP) is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. Our goal is to assign weights to the quantities arising in the three submodels.

Table 2 shows the result of AHP. [EDITOR'S NOTE: The paper omits the details of the calculation.] We then modify Final Value in our programming

Table 2.
Results of the Analytic Hierarchy Process.

Feature	Weight
Risk Cost	0.37
Series Safety	0.28
Power Supply Ability	0.35

problem to become

$$\text{Final Value} = 0.28\,Series(X)' + 0.37\,Risk(X)' + 0.35\,Supply(X)'.$$

Implementation

Data

Google Earth provides the elevation profile of the Zambezi River, and we used MATLAB to fit a curve to it, as shown in **Figure 5**. The dots are the raw data, while the curve (an eighth-degree polynomial) is the fit.

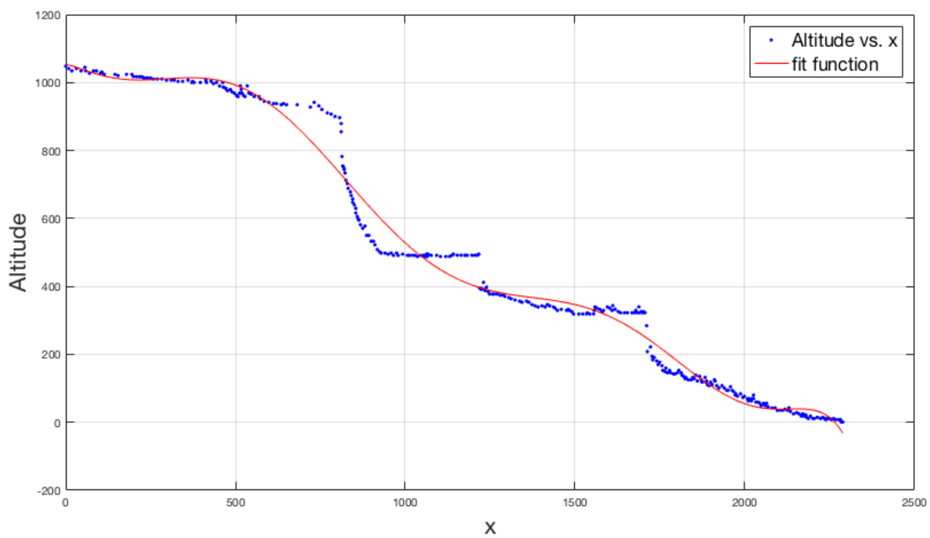


Figure 5. The river's elevation (blue dots) with approximating red curve. The horizontal scale gives the distance from the river's source, while the vertical scale is altitude.

Thanks to the basin analysis in [5], we are able to analyze the river's flow data. **Figure 6a** and **Figure 6b** show the Zambezi River's mean annual river flow.

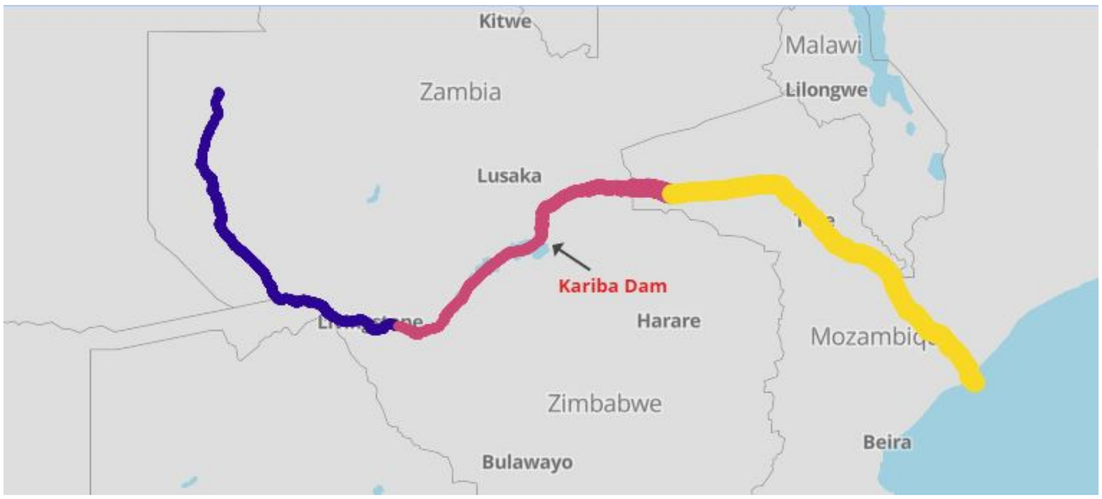


Figure 6a. Two-dimensional rendering of the Zambezi River basin, indicating surrounding countries and cities. The colored bands denote stretches of the river; the width of each band measuring mean annual flow.

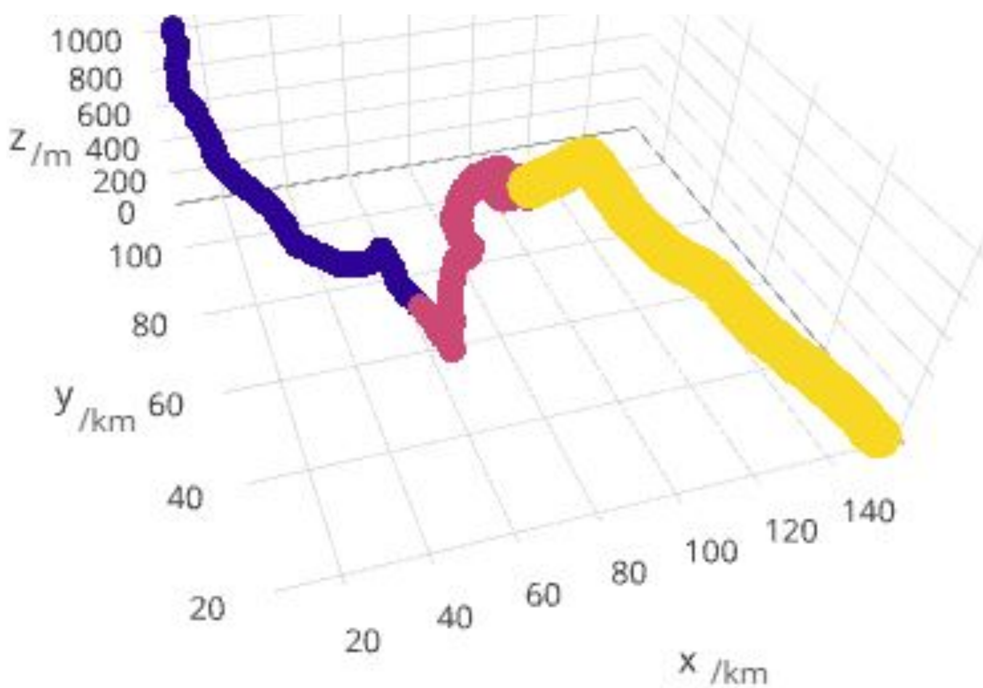


Figure 6b. Two-dimensional rendering of the Zambezi River basin, showing elevation changes.

Number and Placement of Dams

Use of a Genetic Algorithm

We want to give analysis and a recommendation about the number and placement of dams. However, if the height of each dam is allowed to vary, the problem becomes one of more than 20 dimensions. For simplification, we assume that each dam is the same height.

Because of the simplification, h_i is a constant. The whole evaluation is sensitive to the variables $\{Xi\}$, P_f and n. The optimization of a dam group is high-dimensional, multilevel, and nonlinear problem; classical linear programming algorithms cannot find the global optimum. A heuristic algorithm, however, can generate a near-optimal solution. We use a genetic algorithm, an algorithm based on a natural selection process that mimics biological evolution. [EDITOR'S NOTE: The paper does not describe the details of the algorithm but provides in an appendix computer code that implemented it.]

The AHP weights of $Risk(X)$ and $Supply(X)$ are roughly equal. Both are larger than that of $Series(X)$, whose value favors an even distribution of dams. To maximize $Risk(X)$ and meet the constraint for total water management M, we select P_f to an optimal value, which means locating dams to relatively gentle slopes.

Figure 7 shows that with $Risk(X)$ and $Supply(X)$ weighted equally, Final Value reaches the optimal value for $n = 13$ and $P_f = 0.0024$.

This result is convincing. If the number of dams n increases, then:

- $Risk(X)$ decreases. Although the height of each dam is lower, the total $n\beta$ of the initial costs increases.
- $Supply(X)$ increases and then decreases. With increased n, hydropower supply goes higher. But when n gets large, $Series(X)$ restricts it to lower values.
- $Series(X)$ is not relevant to the number n of dams.

Thus, in this decision-making process, there is an optimal value between 10 and 20. We find that it is $\mathbf{n = 13}$ with $\mathbf{P_f = 0.0024}$.

We provide a table and a figure showing the optimal locations of this new dam system. [EDITOR'S NOTE: We omit the table and the figure.]

Determining the Height of Each Dam

Here we eliminate the assumption that each dam has the same height h.

The height of the dams is a key factor in Final Value, since it influences the system's $Risk(X_i)$ and $Supply(X_i)$ that these dams together provide. We determine the height that optimizes the system's Final Value.

Given P_f, a dam's contribution to the system depends only on its height h_i, so we can ignore $Series(X_i)$. To optimize the contribution, we differ-

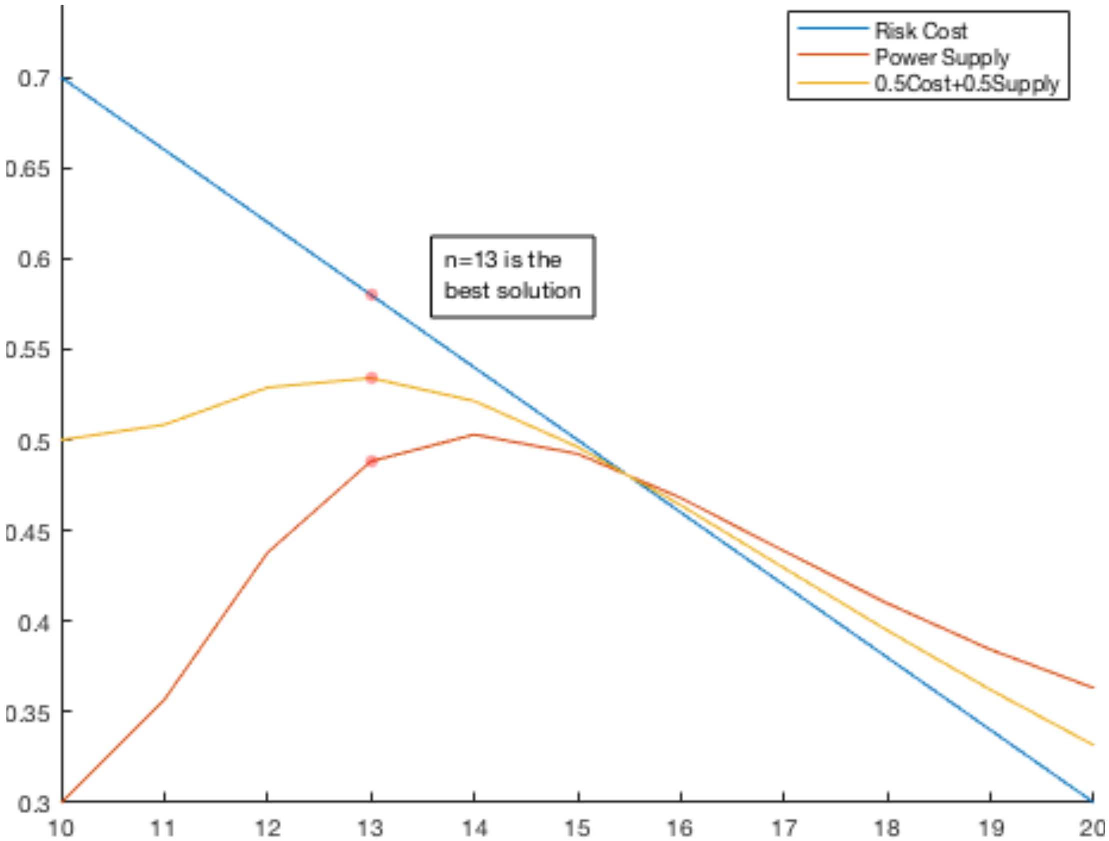


Figure 7. Trend of Final Value for increasing number n of dams.

entiate the ith component of Final Value with respect to h_i and set it to 0:

$$\frac{d}{dh_i}\left[Risk(X_i)+Supply(X_i)\right]$$

$$=\frac{d}{dh_i}\left[\frac{\min_i\big(Risk(X_i)\big)}{\dfrac{Cost(X_i)}{P_f^{1.1}}+\left[1-(1-P_f)^N\right]\alpha(1+\gamma)Q_i(\bar{t})}+\rho g\eta Lk_iQ_i(\bar{t})\right]$$

$$\frac{d}{dh_i}\left[\frac{\min_i\big(Risk(X_i)\big)}{\dfrac{\dfrac{\lambda\rho g d h_i^2}{2\sigma}+\beta}{P_f^{1.1}}+\left[1-(1-P_f)^N\right]\alpha(1+\gamma)Q_i(\bar{t})}+\rho g\eta h_iQ_i(\bar{t})\right]$$

$$=\lambda\frac{-2\rho g d h_i}{2\sigma P_f^{1.1}}+\rho g\eta Q_i(\bar{t})=0,$$

which gives

$$h_i = \frac{\eta \sigma P_f^{1.1}}{\lambda d} Q_i(\bar{t}).$$

So the height of the dam should be proportional to the average flow. Hence, as $Q_i(\bar{t})$ increases, the best h_i also increases. However, there is a constraint for all of the h_i taken together:

$$\sum_{i=1}^{n} V_i = \sum_{i=1}^{n} \frac{h_i^2 d}{2k_i} \geq M.$$

If our optimal $\{X_i\}$ and $\{h_i\}$ do not meet the constraint, we can slightly modify X_i or h_i to meet the constraint.

Thus, the optimal value of Final Value and each dam's height are determined by $\{X_i\}$.

Figure 8 shows our final result for number, placement, and height of the dams.

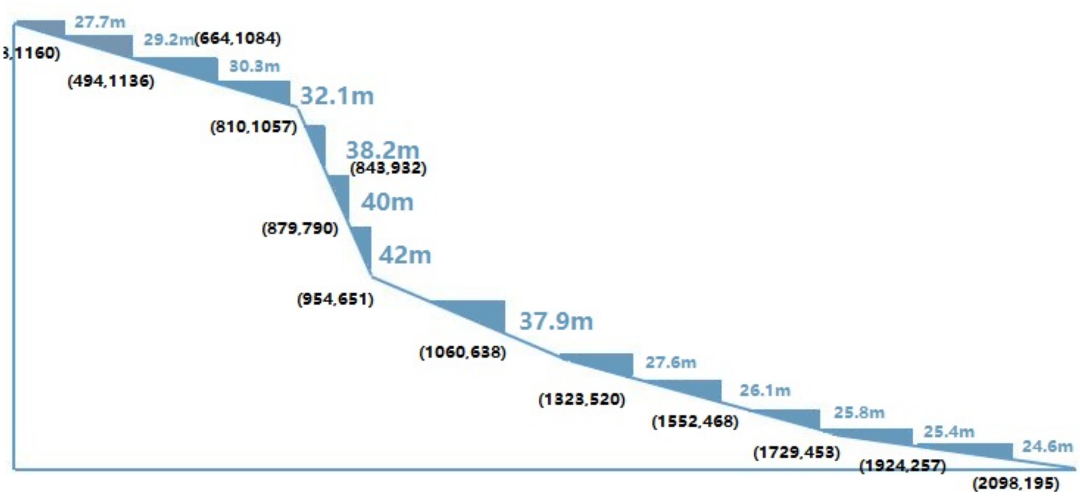


Figure 8. Placement of 13 (the optimal number) of dams, with each dam labelled with its height. The triangular areas (not to scale) are the dams' corresponding reservoirs.

Strategies

A Balance Between Safety and Costs

Our dam system is based on three key factors: Risk Cost, Power Supply, and Series Safety. Risk Cost fully considers the balance between safety and costs through the acceptable risk ratio P_f. Series Safety further investigates the safety of a dam group. Thus, our system is well-balanced; it also takes benefits into consideration.

Protection for Lake Kariba

Now we take protecting Lake Kariba into consideration. In the earlier analysis, we considered only the total volume of the reservoirs. However, we find it is reasonable to consider the distance between a reservoir and the lake. We can weight the volume of each dam by a function of its distance from Lake Kariba, and build dams that are near the lake with stronger ability to manage water. We also need to consider the ecosystem of the lake; so dams near the lake need to be designed carefully to minimize damage to the environment.

In the dry season, the dams below Lake Kariba should start to reserve water and dams above should start to discharge water, until the season passes, with the reverse behavior during the wet season.

Guidance for Emergency Water Flow Situations

Emergency water flow situation (i.e., flooding and prolonged low-water conditions) can generate a catastrophe. Flood water could wash away cities downstream and cause considerable loss of life. Extreme low-water conditions could destroy many fields.

To ensure the safety of our dam system, we provide the managers with guidance about preventing an uncontrolled release of water from reservoirs.

Flood Control

For flood control, we want to maintain as much control over flows entering the system above a critical flood-prone reach as possible [6]:

Fill the upper reservoirs first and empty the lower reservoirs first.

An illustration of this method has been studied [7]. Flood storage at the reservoir close to upstream provides greater flood control than the lower reservoirs.

However, an exception is that if the lower reservoir's outflow capacity is restricted, then it is better to fill the lower reservoir first to increase head on the outlet. In this way we can increase the release capacity from the entire system to the downstream channel capacity. This exception is prepared for some extremely serious floods.

Low-Water Control

For prolonged low-water conditions, our objective is to maintain the river runoff.

Empty the reservoirs from lower ones to higher ones.

During low-water periods, the river's natural runoff is considerably smaller than the average water runoff. We must take advantage of the dam system's capability of dynamic scheduling. In other words, we maintain the river's minimum runoff by compensative regulation. Emptying the reservoirs downstream first can avoid people's water shortage while enabling reservoirs upstream to store water.

Guidance for Extreme Water Flows

In our analysis, we used average annual water flow $Q(\bar{t})$ to denote the annual water flow. However in real life, water flow varies; there is a dry season and a wet season.

Dry Season

The dam system needs to discharge water during the dry season. The water in the upper reservoirs is eventually distributed along all the terrain beside the river; but the water in the lower reservoirs downstream is used by a relatively smaller terrain. Hence the upper reservoirs will discharge water at a greater rate than the lower reservoirs.

Wet Season

During the wet season, we need to charge the reservoirs to prepare for the next dry season. We can assume that rainfall is uniformly distributed along the river. The upper reservoirs must start to store water first, since if a reservoir downstream starts first, it may be forced to release water, leading to risk and waste of water.

Model Analysis

Sensitivity Analysis

We determined the value of some of the model parameters through least-squares fitting, some from values in the literature, and others by still other methods.

In this section, we produce a sensitivity analysis to show whether our model is sensitive to different values of the parameters. In particular, we study the impact of two important parameters: the lifetime N of the dams, and the extreme condition ratio α.

Impact of Planned Working Years N for a Dam

In the literature, a dam over 50 years old is defined to be dangerous. The Kariba Dam is such a dam. **Figure 9** shows that the trend of Risk Loss remains the same for dam lifetimes of 50 to 100 years, and the optimizing value of P_f does not change appreciably. Therefore, we conclude that our model is not sensitive to the value of N.

Impact of Extreme Condition ratio α

Figure 4 on p. 275 shows the sensitivity analysis of α. We see little sensitivity to α of Risk Value near the optimal value of P_f.

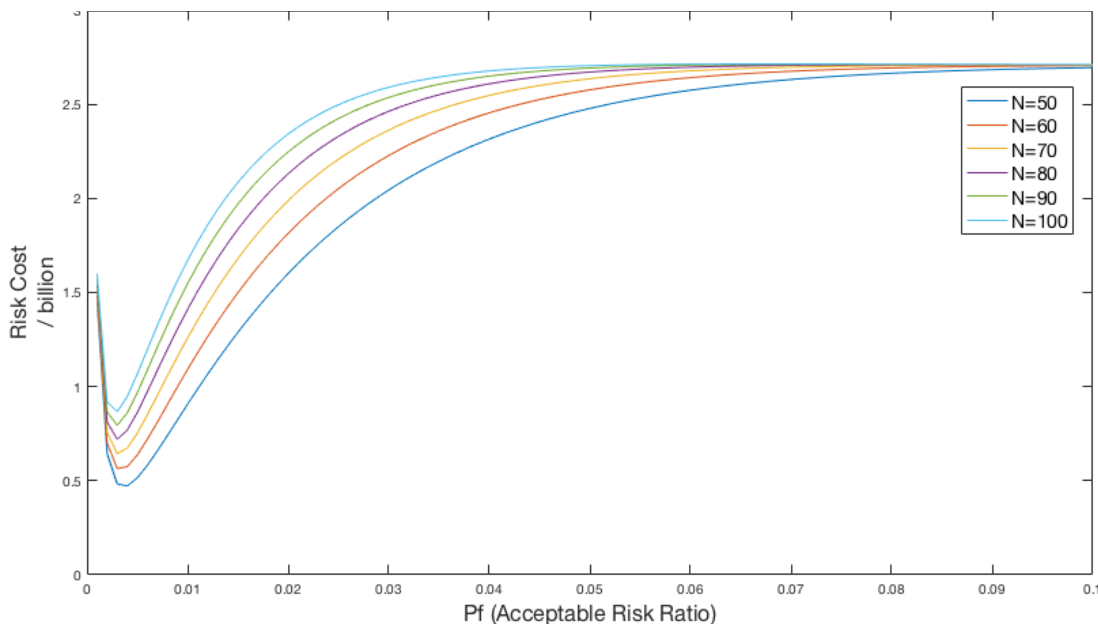


Figure 9. Risk Value vs. P_f for values of N from 50 to 100 years, for fixed locations of the dams.

Strengths and Weaknesses

Strengths

- We use accurate data about the Zambezi River.
- Our design of the new dam system uses novel and innovative methods across disciplines to give an integrated evaluation function of dams, including the Analytic Hierarchy Process and a genetic algorithm.
- We use effective visualization to display our model.

Weaknesses

- We may not realize the optimal situation. Because of the high-dimensional constraint, we used a genetic algorithm as our optimization algorithm. It is a heuristic algorithm that may not get the optimal results for site selection for the dams.
- Some real water conditions are omitted. We simplified the river as an open channel. However, the real condition is much more complex and will also influence the construction cost and safety evaluation.
- We lack some details of the trade-off between cost and safety. Power Supply is included in our model, but it still lacks other benefits of dam construction.

Conclusion

We provide a detailed analysis of replacing the Kariba Dam with 10 to 20 smaller dams. We start with analysis of water flow at a single dam, then extend our model to multiple dams. The model takes safety, cost, and benefits into consideration and fashions a balance among them. Three submodels, Risk Cost Model, Power Supply Ability Model, and Series Safety Model, comprise the integrated model.

We use real data to furnish a recommendation of the number, placement, and height of the new dams. Because this is a high-dimensional problem, we implement a genetic algorithm to find a global optimum result.

We discuss strategies for protection of Lake Kariba and offer guidance for periods of high water and low water.

Brief Assessment Report

The Kariba Dam, which was opened in 1959 and holds back the world's largest reservoir, has been working for over 50 years and faces the possibility of collapse. Also serving as a major supply of Zambia's and nearby countries' electricity, it is in dire need of maintenance. To deal with this situation, there are three major options. For each, we will assess its potential cost and provide other information.

Option 1: Repair the Existing Kariba Dam

Potential Cost

[9] says that the expected cost will be $294 million this time. However, additional repairs will no doubt be needed in future as the dam ages further.

Benefits

Repair saves the situation for the time being, and construction during the repair period does not affect the hydropower supply, which is generated from the reservoir.

Shortcomings

There is still danger of collapse, plus a high repair rate with high cost each time.

Option 2: Rebuild the Existing Kariba Dam

Potential Cost

Rebuilding such a giant dam is extremely expensive. According to [2], the project was built in two stages at a total cost of $1.8 billion. To update the cost to 2017 dollars, we use the Consumer Price Index (CPI) [10]. The

CPI for the 1950s was 27.2, while the current CPI is about 241.4. Thus, the total cost of rebuilding can be estimated as $\$1.8 \times 241.4/27.2 \approx \$16$ billion.

Also, there will still be future repair and maintenance costs.

Benefits

No severe potential danger any more. No need to repair very often.

Shortcomings

A long period for construction; meanwhile, no dam, affecting hydro power supply, and no protection for Lake Kariba.

Option 3: Replace the Kariba Dam with Smaller Dams

Potential Costs

The construction cost of a dam can be described in terms of a fixed base cost plus a term proportional to the capacity:

$$\text{Construction cost} = (\text{base cost}) + C \times (\text{dam capacity}).$$

The total capacity of n small dams should equal the capacity of the Kariba Dam. So the construction cost of Option 3 is close to n times the base cost plus the cost of Option 2.

We suppose that the base cost is 10% of the total construction cost and the cost of removing the Kariba Dam is about 15% of the total construction cost. Then the total construction cost is about

$$\$16 \times (1 + 10\% \times n + 15\%) \approx \$34 \text{ billion for 10 dams, } \$50 \text{ billion for 20.}$$

For $n = 13$, the number of dams that we prescribe, the cost would be about \$39 billion.

Benefits

The system would strengthen the abilities to handle emergency water flows, protect Lake Kariba, and adjust water flow between seasons.

Shortcomings

This option costs the most of the three options, and it brings in new problems of coordinating and control of a dam system.

References

[1] Wikipedia. 2017. Kariba Dam. `https://en.wikipedia.org/wiki/Kariba_Dam`.

[2] Soils Incorporated (Pty) Ltd and Chalo Environmental and Sustainable Development Consultants. 2000. Kariba Dam Case Study, prepared as an input to the World Commission on Dams, Cape Town. https://archive.li/ngwMK#selection-649.0-649.200 .

[3] Wikipedia. 2017. Manning formula. https://en.wikipedia.org/wiki/Manning_formula .

[4] YANG, Qiang, WU Hao, and ZHOU Wei-yuan. 2006. Stress evaluation in finite element analysis of dams [in Chinese]. *Engineering Mechanics* 23(1): 69–73. http://en.cnki.com.cn/Article_en/CJFDTOTAL-GCLX200601013.htm .

[5] World Bank. 2010. The Zambezi River basin: A multi-sector investment opportunities analysis—basin development scenarios. https://openknowledge.worldbank.org/handle/10986/2959 .

[6] Lund, Jay R., and Joel Guzman. 1999. Some derived operating rules for reservoirs in series or in parallel. *Journal of Water Resources Planning and Management* 125 (3) (May–June): 143–153. http://cadswes2.colorado.edu/~zagona/CVEN5838/JournalArticles/Lund%202000%20op%20rules%20for%20res%20in%20series%20and%20parallel.pdf .

[7] Kelman, J., J.M. Damazio, J.L. Mariën, and J.P. da Costa. 1989. The determination of flood control volumes in a multireservoir system. *Journal of Water Resources Planning and Management* 25 (3): 337–344. http://kelman.com.br/pdf/the%20determination%20of%20flood%20control%20volumes/the%20determination%20of%20flood%20control%20volumes.pdf .

[8] Leslie, Jacques. 2016. One of Africa's biggest dams is falling apart. *The New Yorker*. http://www.newyorker.com/tech/elements/one-of-africas-biggest-dams-is-falling-apart .

[9] Reuters Staff. 2016. Zambia, Zimbabwe to start Kariba Dam repairs in early 2017. http://af.reuters.com/article/zambiaNews/idAFL5N17T1OI .

[10] Mankiw, N. Gregory. 2014. *Principles of Macroeconomics*. 7th ed. Boston, MA: Cengage Learning.

Judges' Commentary: The Dam Problem

Bill Wilhelm
Lockheed Martin
Huntsville, AL
billwilhelm1@comcast.net

Overview of the Problem

Students who explored Problem A in this year's Mathematical Contest in Modeling examined options to address concerns over the potential instability of the Kariba Dam on the Zambezi River basin between Zambia and Zimbabwe. This potential instability was brought about by the erosion of bedrock at the base of the dam caused by years of water spilling over the spillway.

This commentary includes a brief overview of the problem statement. That is followed by a list of some of the common themes and more detailed points that emerged as the judging proceeded, including a list of some of the common approaches adopted by the students. Finally, we offer a summary of the key discriminators used by judges to compare student submissions.

The Problem Requirements

There were two main requirements:

- To provide a brief assessment of three options to address the deteriorating conditions of the Kariba Dam: repairing the dam, rebuilding the dam, or removing and replacing the dam with a series of 10 to 20 smaller dams. This assessment was to be limited to one or two pages and was expected to include an overview of the potential costs and benefits of each option.
- To provide a detailed analysis of the option to remove and replace the dam with a number of smaller dams. This analysis was not to exceed 20 pages.

The UMAP Journal 38 (3) (2017) 289–296.

In addition to the written analysis addressing the two requirements, a single-page summary sheet was expected.

The Option Analysis Requirement

The goal for this requirement was to identify the costs and benefits associated with each option for addressing the worsening situation with the Kariba Dam, and to combine them in a meaningful way to provide a supportable recommendation.

Successful teams considered not only dam demolition, construction, and repair costs, but also costs of lost revenue related to electricity production, drinking water, irrigation, fishing, and recreation. Successful teams also incorporated the time required to implement each option into their analysis, and provided a relative comparison of the safety aspects of each option, as well. Team recommendations varied based on the relative importance each team placed on the costs versus the benefits, and whether they focused on the long-term or short-term costs and benefits.

In evaluating this requirement, judges were particularly looking for consistency between statements that teams made about the relative importance of the factors affecting their decision and their ultimate conclusion. Teams that simply provided a discussion of costs and benefits without a recommendation, or a recommendation without supporting rationale, were not considered successful with this requirement.

The Remove and Replace Option Requirement

As indicated by the page limit imposed on the analysis of the two requirements, the primary focus of the problem was to provide an analysis of the option to remove the Kariba Dam and replace it with smaller dams. Successful teams clearly addressed the following questions:

- How many dams are required?
- Where should the dams be placed?
- How should the flow of water be modulated through the dams to balance safety and costs during prolonged flooding?
- How should the flow of water be modulated through the dams to balance safety and costs during prolonged drought?
- What are the restrictions on locations that should be exposed to detrimental effects of extreme water flow conditions and the lengths of time of that exposure?

For this requirement, teams typically identified a number of criteria to be applied to identify candidate locations for dam placement. These criteria often included

- the width of the river,
- the slope of the neighboring terrain,
- the proximity of population centers, and
- the geology of both the dam site itself as well as the area to be occupied by the reservoir.

Some teams simply chose locations equally spread out along either the length or the elevation of the river.

After identifying candidate sites, teams typically applied constraints to ensure that the replacement dams provide the same capability and functionality as the Kariba Dam. Such constraints included overall water storage capacity as well as electrical energy-generation potential. Teams then mathematically modelled these constraints and the overall costs associated with each dam in order to minimize the overall cost of the Zambezi River water management system.

Most good papers offered logically constructed models with concise and understandable descriptions. The judges expected the students to clearly describe the objectives and evaluate their models in a manner consistent with the stated objectives. Specific results and conclusions were presented in higher-ranked papers in a clear and concise manner.

The Water Modulation Requirement

In addition to identifying the recommended number and location of dams to support efficient management of water in the Zambezi River Basin, teams were asked to include a strategy for modulating the water flow through the multiple dam system that balanced safety and costs.

To achieve this, teams had to model ways to quantify both costs and safety, and develop a simulation or analytical model to "balance" the two. These models needed to consider

- the water held in the reservoir behind each dam,
- the amount of water flowing from one dam to another,
- the amount of water from rain and runoff,
- the amount of water siphoned off for human consumption and irrigation, and
- water loss due to evaporation.

These models needed to be applied to the extremely high and extremely low values of water flow due to flooding and drought.

The Summary Requirement

An essential requirement was to write a standard one-page summary sheet. Successful papers included a brief overview of the technical approach to satisfying the problem requirements, along with a summary of the results. At a minimum, successful papers included the recommended option for addressing the Kariba Dam foundation issues and the number and location of recommended replacement dams.

Condensing the substance of the team's analysis into a single page was a difficult task for the teams. The student teams that managed to convey a sense of the basic models, the underlying assumptions, and the limitations of their models tended to make a stronger impression.

The Modeling Effort

With the quantity and availability of information available on the Internet today, many teams were able to find an example of a solution to this problem or one similar to it and modify it for use in this competition. The most successful teams, however, built cohesive mathematical models and described their derivation in sufficient detail to explain both the underlying logic as well as the results and conclusions drawn from them.

The organization and consistency of papers was also a discriminator. Successful teams

- clearly defined the variables that impacted the problem,
- documented their data collection effort,
- stated any assumptions made to either compensate for the lack of data availability or to support the development of the model,
- expressed the relationships between the variables with mathematical equations, and then
- used suitable approaches or appropriate techniques to optimize or analyze these relationships to support a conclusion and recommendation.

Variable Definition

Key to the modeling effort was the definition of variables used: Judges expected each variable used in the model to be explained. Additionally, each variable should be associated with some type of unit, if appropriate. Units of variables should be chosen to ensure that a dimensional analysis of all derived equations indicates the same fundamental quantities on both sides. Papers that included models involving generic "factors," "constants," or "coefficients" without a detailed explanation of how the variable was derived, or specifically why it was used, were not as successful as those that provided more detailed explanations. This was very often the case when a team leveraged work from an Internet source without adequate description. Many teams chose to define all of the variables that they used up front in their paper, in a table dedicated to just that purpose. Some of those teams either failed to include all of the variables they used, or included variables that were not used elsewhere in the paper.

The more successful teams defined each variable as it was introduced in the paper. This way, the variables were described within the context of the modeling effort, making it much easier for the judges to follow both what the variable was meant to represent and how it was to be used in the model. Another observation concerned the number of variable definitions used. Statistician George Box is often quoted for his saying, "All models are wrong but some are useful" [Box 1979, 202]. What he meant by this is that, try as we may, models cannot exactly represent a system in the real world (with emphasis on the word "exactly"). Given that we cannot precisely model such things, Box felt that "overparameterization is often the mark of mediocrity" [Box 1976].

Many of the models presented to address the requirements of the Kariba Dam problem had a large number of variables, were overly complex, and provided the teams much difficulty with conducting meaningful sensitivity analysis. The most effective papers constrained their models to the variables most pertinent to the problem at hand, and lent themselves well to sensitivity analysis.

Data Collection

In some cases, data availability was a driving force in the derivation of the model itself. Clearly, it makes no sense to develop a mathematical model that requires input data that are unobtainable. On the other hand, the source of the data also affects the credibility of the modeling effort. Successful teams not only developed models, but collected data to support them and documented the sources of the data, as well.

Assumptions

Most teams included a section about the assumptions made during the course of model development. Successful teams also included some rationale as to why they were making the assumption, how the assumption ultimately influenced the model, and why the assumption was reasonable to make. The most successful teams clearly limited the assumptions that they documented to ones that were necessary. In other words, the most successful teams only made assumptions pertinent to the model—ones that if, in the event that they turned out to be wrong, would likely alter the model itself or any conclusions drawn from it.

As with variable definitions, some teams chose to dedicate a section of their paper to assumptions and justifications. While this technique ensured that judges could clearly see what assumptions were made by the team, it did not always convey the rationale or the need for the assumptions in the model development. Teams that included their assumptions in-line with their model derivation were generally more successful in conveying not only why the assumption was being made, but how it was applied in the model as well.

Solution Approaches

Once variables were identified, then data were collected, assumptions were made, and mathematical models were presented. Approaches to satisfy the requirements of the problem varied. The time to address this problem was limited, and nothing in the problem statement suggested any particular direction with respect to solution approaches.

So the judges did not expect extensive data collection efforts and elaborate sophisticated models. They did expect a strong analysis that was consistent with any models that were developed. Below is a summary of some of the common approaches used. If standard methods developed by others for dealing with certain aspects of the problem were used, appropriate citations were expected. Proper validation of the models used against similar problems or situations that are documented in the literature was also a discriminator.

Analytic Hierarchy Process

Many teams derived costs and benefits for the three options for addressing the Zambezi River situation and then used an Analytic Hierarchy Process for rank-ordering the options.

While the Analytic Hierarchy Process is widely used to support group decision-making, it is also extremely useful in solving problems like the Kariba Dam where alternatives are to be rank-ordered based on a set of evaluative criteria.

Genetic Algorithm

Many teams used a heuristic genetic algorithm to optimize the number and location of dams required for the option to replace the Kariba Dam with smaller dams. The genetic algorithm is a common technique for solving optimization and search problems, particularly if the modeled objective function is not linear. This approach typically required the use of some sort of computer program to iteratively “evolve” a solution.

Integer Linear Programming

In addressing the problem of replacing the Kariba Dam with multiple smaller ones, some teams expressed restrictions on water storage capacity, electricity generation, and other economic factors as linear constraint equations and then used linear integer programming methods to determine the least cost set of dam sites to satisfy the constraints. Teams that took this approach also made use of existing commercial integer programming software packages.

Simulation

Many teams used simulation in their analysis, particularly to address the water modulation requirement. Successful teams provided an explanation of the statistical significance of the simulation results. Most teams that used simulation included the source code used in an appendix to the paper. Yet the most successful teams explained their simulation process in detail; some even included flow diagrams.

Sensitivity Analysis

Not all teams had time to perform sensitivity analysis. Of those that did, some simply varied selected variables and produced graphs indicating the sensitivity of their model results to those particular variables. More successful papers not only labeled their graphs appropriately, but actually explained the implications of these graphs, along with recommendations for further analysis or data collection if warranted.

The most successful teams, however, focused their sensitivity analysis on the assumptions that they made during the course of modeling. This provided consistency between the sensitivity analysis and the modeling process, and either enhanced the credibility of the team's recommendations (if the model results were not sensitive to the assumptions made) or suggested areas for further analysis.

Summary

This problem was a difficult one to address in the time given in the competition. Many teams were sidetracked by delving too deeply into one aspect of the problem or another. The most successful teams stayed focused, managed their time wisely, addressed all of the requirements of the problem and included:

- pertinent variables that were well defined (with units),
- data collection of the pertinent variables from reliable sources,

- assumptions made either to compensate for the lack of data, or to explain mathematical relationships between the pertinent variables,
- an appropriate mathematical technique or approach to analyze the mathematical relationships and form a recommendation, and
- sensitivity analysis to address any assumptions made, or address anticipated changes in the pertinent variables used.

References

Box, George E.P. 1976. Science and statistics. *Journal of the American Statistical Association* 71: 791–799.

________. 1979. Robustness in the strategy of scientific model building. In *Robustness in Statistics*, edited by Robert L. Launer and Graham N. Wilkinson, 201–236. New York: Academic Press.

About the Author

Bill Wilhelm is a Systems Engineer and a Certified Systems Architect in the Battle Management Technical Center of Lockheed Martin's Space Systems Company in Huntsville, Alabama. Prior to joining Lockheed Martin, he served for twenty years as an officer in the United States Army, with multiple assignments as an operations research analyst and an assistant professor. He received his undergraduate degree from the U.S. Military Academy at West Point and master's degrees in Operations Research and Engineering Economic Systems from Stanford University. He has wide-ranging research interests, including simulation, computer algorithm optimization, and anything related to integrated air and missile defense.

Honeycomb Toll Plaza

Jiahua Zhang
Yang Liu
Rui Liu

Information Science and Engineering
Shandong Normal University
Jinan, Shandong
China
Advisor: Hui Ji

Abstract

We analyze the performance of a common toll plaza design compared to our proposed new improved toll plaza. The new design would reduce the cost, decrease the probability of collision at merge points, and increase throughput.

Our proposed cellular toll plaza resembles a honeycomb. Each hexagon contains two toll booths, which serve two separated vehicle streams that are merged in advance before they re-enter the highway. The total area of the proposed plaza is reduced significantly. Also, the average waiting time in queue is diminished, which means that throughput is increased . Additionally, by splitting the merging into two stages, the probability of accidents is decreased.

Our main contributions are:

- The new cellular architecture can greatly reduce the construction area compared with traditional linear distributed toll booths.
- We analyze the throughput of toll plazas by means of queuing theory. To verify our theory, we simulate the behavior of a large number of vehicles passing the toll plaza, with the help of PTV-VISSIM. Simulation results show that our cellular toll plazas have better results than traditional toll plazas, especially when the traffic flow is heavy; the average travel time is reduced by about 55% and the average delay time is reduced by about 70%.
- We analyze the influence of the proportions of varied types of toll booths to our design. According to sources, the impact of exact-change toll booths is similar to manual toll booths; so we consider only two kinds of toll booths: human-staffed (MTC) toll booths and E-ZPass (ETC) toll booths. Simulation results using PTV-VISSIM show that an ETC toll booth is 8 times as fast as an MTC toll booth.

The UMAP Journal 38 (3) (2017) 297–318.

- We simulate the performance of the cellular toll plaza under varied traffic throughput. Simulation results show that the average transit time remains at about 11 s under throughputs up to 2,000 vehicles/hr). We conclude that our model is not sensitive to traffic flow variations and has strong robustness.
- To further reduce the probability of accidents, we make the transition zone smoother and rearrange the different kinds of toll booths.
- For self-driving vehicles, we reserve special E-ZPass toll booths in the center of the toll plaza, which match the characteristics of autonomous vehicles: safer and faster.

Electronic toll collection and autonomous vehicles are the trends of modern transportation. Our new design can improve the performance of toll collection in terms of cost, throughput, and accident prevention.

Introduction

Problem Background

Research found that 36% of total car travel time in China is delay time caused by tolling [1]. In addition, as a vehicle-intensive place, the toll plaza has become an accident-prone section [2]. The congestion problem at toll plazas becomes more and more serious due to the outdated design.

With the widespread use of Electronic Toll Collection (ETC) (such as the E-ZPass system in the U.S.) to replace Manual Toll Collection (MTC), the efficiency of toll collection has improved significantly and relieved some congestion at the toll plaza. However, due to the higher speed of vehicles passing through the toll plaza, the probability of collision in the merging zone is increased.

We design a toll plaza based on bionics: a honeycomb hexagonal tiling creates equal-sized cells while minimizing the total perimeter of the cells [3]. Such cellular structure is widely used; for example, the base stations of mobile communications are distributed in honeycomb-like fashion. In our toll plaza, the toll booths are located in regular hexagons.

Description of Terminology

- **Total time cost:** The average time interval for a vehicle from the beginning point of the detection area to the ending point of the detection area.
- **Theoretical time cost:** If there is only one vehicle in the system and that vehicle is not limited by a control signal, the time interval for that vehicle from the beginning point of the detection area to the ending point of the detection area is the theoretical time cost.

- **Time delayed:** Total time cost minus theoretical time cost.
- L: the number of lanes in each direction of the highway.
- B: the total number of toll booths in each direction.

Our Work

With the popularization of ETC equipment and autonomous vehicles, MTC lanes will be totally replaced by ETC lanes in the next 20 years; that will increase the road capacity and decrease the time cost for each car.

Present toll plazas of traditional design occupy a large area, and the cost of construction is high. With an increase in vehicles' speed, congestion at the merge points may increase the probability of accidents.

- We design a model of a cellular toll plaza, including its shape, size, and merge.
- We quantify how much we could reduce the area of the toll plaza.
- We use the VISSIM software to simulate a cellular toll plaza.
- We use VISSIM to analyze the traffic capacity of a cellular toll plaza with a mixture of toll lanes (MTC, ETC, etc.), for both heavy traffic and light traffic.
- Based on our results, we improve our design in three particular aspects.
- We analyze the effect of traffic flow on the capacity of the toll plaza.
- We investigate whether the toll plaza can meet the needs of autonomous vehicles.

Figure 1 shows the process of evolution of our design.

General Assumptions

- The arrival of vehicles follows a Poisson distribution.
- In general, the traffic at ETC toll booths should be much heavier than the traffic at other types of toll booth.
- All toll booths are ETC or E-ZPass unless otherwise specified.
- There are no entrance or exit ramps near the toll booths.
- The service procedure at the toll booths, and the merging procedure of the vehicles after the toll booths, first-come, first-served queuing systems.

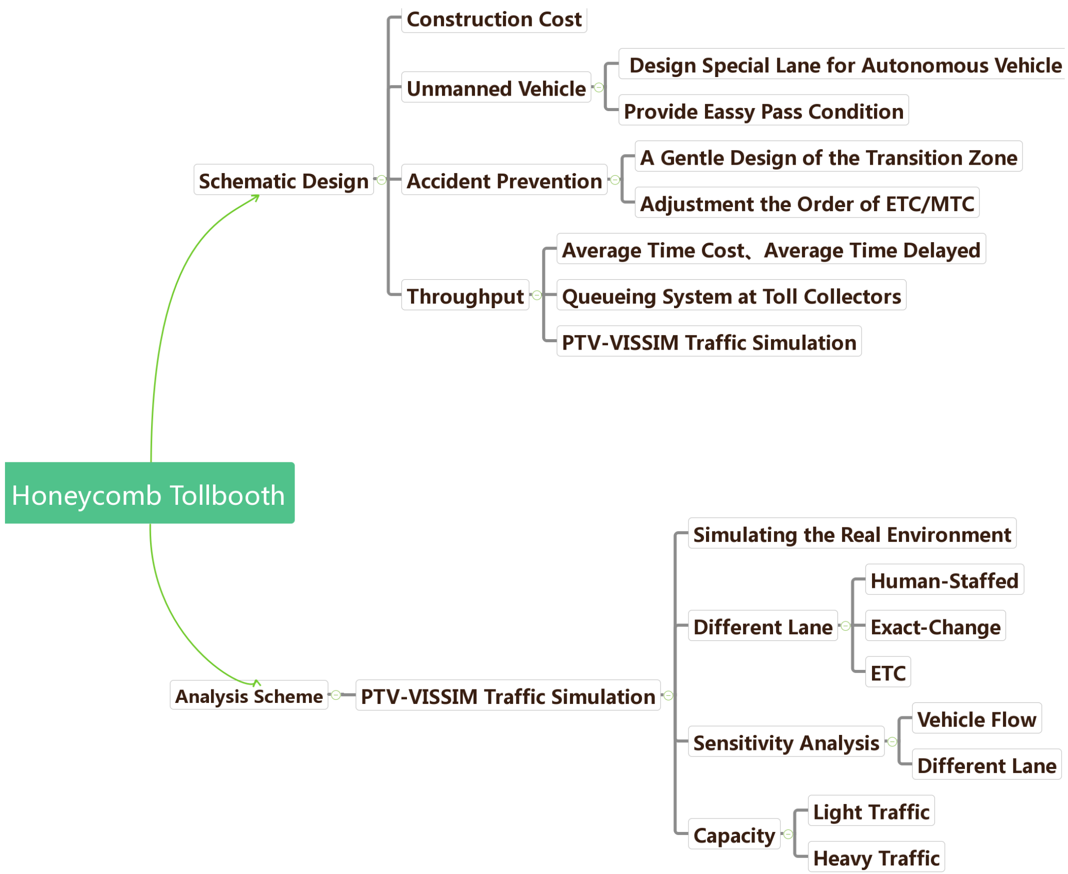


Figure 1. Evolution of our design.

Design of a Honeycomb-Like Toll Plaza

In traditional toll plazas, there are more toll booths than lanes of incoming traffic. A toll plaza consists of

- the fan-out area before the toll booths,
- the toll booths themselves, and
- the fan-in area after the toll booths.

The toll booths are often constructed in a straight line across the highway, perpendicular to the direction of traffic flow. Therefore, the area of the toll plaza is large. To reduce the area and save on the construction cost, we design a new kind of toll plaza based on the structure of the honeycomb. In addition, by splitting the merging procedure into two stages, our new design can reduce the probability of collision, in contrast to the traditional merging procedure, where a large number of vehicles concentrate onto the highway simultaneously. The evolution of our design is shown in **Figure 2**, where we smooth the transition zone to avoid sharp turns and add some reserved toll booths in the middle for autonomous vehicles.

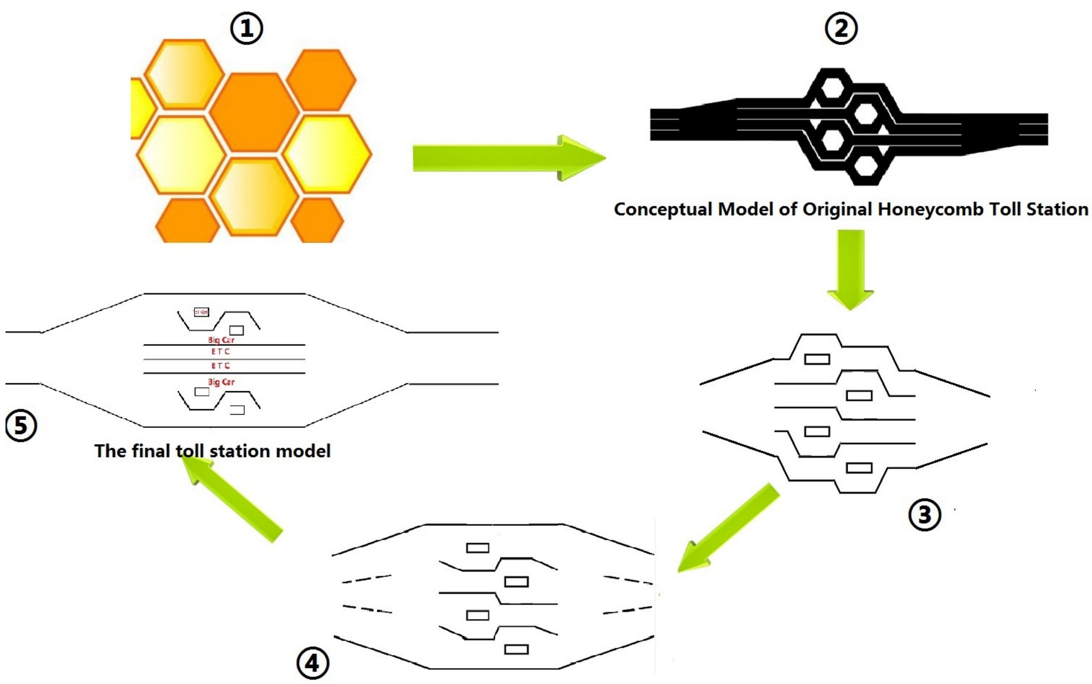


Figure 2. Evolution of our design.

Model Design

Estimated Cost of the Toll Plaza

The main cost of building a toll plaza includes the construction costs of the road surface and of the toll booths. We assess the land area and try to minimize it. The toll plaza's total area can be divided into zones as shown in **Figure 3**.

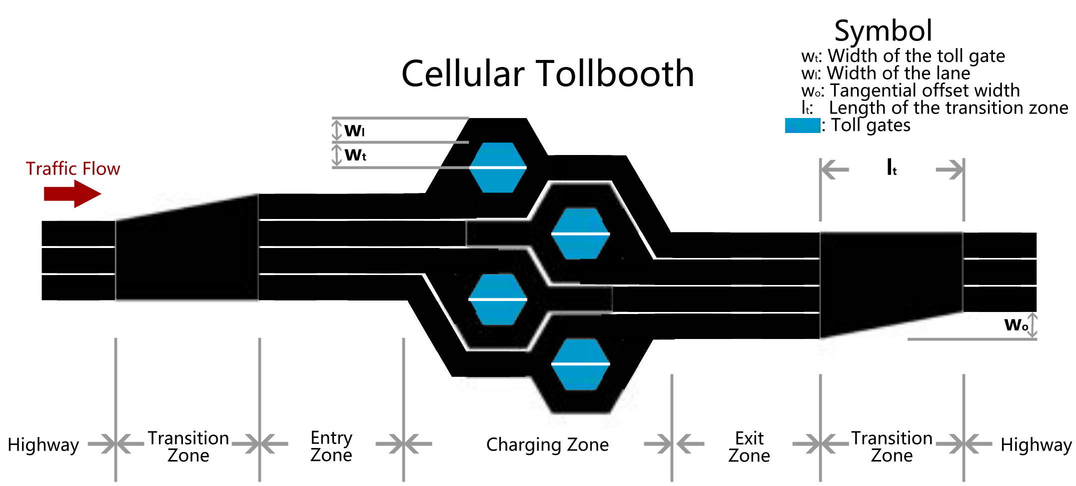


Figure 3. Design and design parameters for a cellular toll plaza.

We establish some parameters:

- n_t = number of toll booths,
- w_t = width of a toll booth,
- n_l = number of the lanes of the highway,
- w_l = width of a lane,
- w_o = tangential offset width,
- l_t = length of the transition zone, and
- v = design speed.

Comparison of the Areas of the Toll Plazas

[EDITOR'S NOTE: We omit the authors' calculation of the geometric areas of the traditional toll plaza and of the cellular toll plaza. **Figure 4** serves well to illustrate their point.]

The cellular toll plaza can significantly save space compared with the traditional toll plaza. The effect can be seen in **Figure 4**.

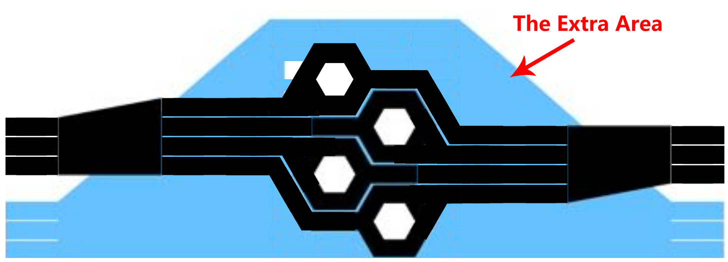


Figure 4. Comparison of the footprints of the cellular toll plaza and the traditional toll plaza.

Analysis of the Throughput of Toll Plazas

We consider the entire process of tolling as the operation of two serially-connected queuing systems:

- vehicles passing through the toll booths and queuing in front of the toll booths, and
- vehicles passing through the merge points at the exit of the toll plaza.

Queuing System at Toll Collectors

When a vehicle enters the toll plaza, the driver will head to a toll booth according to certain principles, such as the distance to each toll booth and the number of vehicles waiting in each queue. In our model:

- The arrival of vehicles at each toll booth is a Poisson process, so the interval between vehicles arriving at a toll booth follows an exponential distribution.

- The time cost for each vehicle at the toll booth also follows an exponential distribution.

Each toll booth can handle only one lane at each time. In our model there are two toll booths, on each toll island, one on each side, each serving a single lane.

In summary, we believe that each toll booth can be considered an M/M/1 queuing system, meaning that the time between arrivals is exponential ("M" for "Markovian"), the service time is exponential, and there is a single server.

Queuing at Merge Points

We note Burke's theorem:

If the arrival time and service time of a M/M/1 queuing model is a Poisson process with parameter λ, the departure process of the queuing model is also a Poisson process with parameter λ [5].

Since the output of a toll booth is a Poisson process, arrivals at the merge points are also Poisson processes.

Current U.S. road design guidelines stipulate that lane merging can only be started from the right side of the vehicle's driving direction and only one lane can be merged at a time [6]. According to this provision, and also to simplify the model, we divide the lanes into two types:

- A Type I lane doesn't pass through any merge point, but
- a Type II lane does.

Recall that there are L highway lanes and B toll booths; Type II lanes all merge into a single highway lane, so there are $(L-1)$ Type I lanes and $(B-L+1)$ Type II lanes. Type II lanes need to merge with all lanes on their right; for example, in **Figure 6**, lane 4 needs to merge with cars from lanes 1, 2, and 3, while lane 5 is of Type I and does no merging.

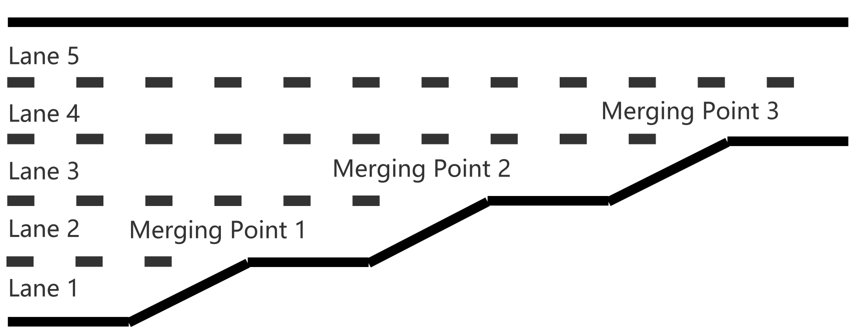


Figure 6. Merge points in the transition zone.

For vehicles in Type I lanes, which can pass through without merging, we consider the total time cost to be the number of merge points times the time cost for passing each of them, that is,

$$\begin{cases} (B-L)\times\dfrac{1}{\mu_0}, & \text{for the traditional toll plaza; and} \\ \left(\dfrac{B}{2}-L\right)\times\dfrac{1}{\mu_0}, & \text{for the cellular toll plaza,} \end{cases}$$

where μ_0 is the service rate at a merge point when there is no merge conflict.

The probability of a vehicle being in a Type I lane is $(L-1)/B$ for the traditional toll plaza and $(L-1)/(B/2)$ for our cellular toll plaza.

The probability of a vehicle being in a Type II lane is $(B-L+1)/B$ for the traditional toll plaza and $(B-2L+2)/B$ for our cellular toll plaza. For a vehicle in a Type II lane, we must take into account not only the number of merge points that it must pass through but also the probability of merging at a merge point.

At the kth merge point, the probability of arrival of two vehicles simultaneously is the sum of the probability of one lane plus the probability of a vehicle from the $(k-1)$st merge point, that is, $(k+1)/B$ and $2(k+1)/B$, respectively. Take the merging pattern shown in **Figure 6** as an example, the probability of merging at merge point 1 equals the probability of a vehicle in lane 1 plus the probability of a vehicle in lane 2, that is, $1/B+1/B=2/B$; similarly, at merge point 2, the probability is the probability of a vehicle in lane 3 plus the probability of a vehicle in merge point 1, that is, $1/B+2/B=3/B$. The traffic flow at merge point k is thus $(k+1)\Phi/B$, where Φ is the total traffic flow.

To simplify the model, we don't distinguish between the two lanes that merge at the same merge point; that is, the time for a vehicle to pass through a single merge point is independent of which of the two lanes it is in. If there is no vehicle in the other lane, or a vehicle in the other lane that slows or stops to avoid the merging conflict, this vehicle can complete the merging process without deceleration, that is, with time cost $1/\mu_0$. Otherwise, this vehicle needs to slow or stop and wait, with time cost defined as $1/\mu_1$.

In summary, the arrival rate of the queuing system at merge points follows an exponential distribution (i.e., it is a Poisson process), but the service rate is a more general function; that is, we have an M/G/1 queuing model (with "G" for "General").

Calculation

Parameter Assignment

- B**: Number of toll booths** In reality, the number of toll booths depends on the traffic flow, types of vehicles, etc. However, to simplify our model, we set $B=8$.
- L**: Number of lanes of highway** We set $L=3$.

- μ_T: **Service rate of a toll booth** We take the service rate to be 1200 vehicles/hr, corresponding to that of a widely-installed current electronic toll system [7].

- μ_0: **Service rate at a merge point when no merging conflict occurs, or conflict occurs but the other vehicle slows or stops to wait** We also take μ_0 as the service rate when a vehicle directly heads for a lane of the highway without passing any merge point. The average speed of vehicles on the highway is 60 mph [7]. Therefore, the length of the merge point should be the length of a normal vehicle (15 ft) plus a safe distance, which is six times the length of a normal vehicle [7]; so the length of the merge point is 105 ft. Then the average time for a vehicle to pass through a merge point is 105 ft / 60 mph $\approx$ 1.2 s. We take the reciprocal of this as the value of μ_0, converted to an hourly basis:

$$\mu_0 = \frac{3600 \text{ s}}{\text{hr}} \times \frac{1}{1.2 \text{ s}} \approx 3017 \text{ vehicles/hr.}$$

- μ_1: **Service rate at a merge point when a merging conflict occurs** This rate applies to a vehicle that stops to avoid another vehicle when two vehicles reach the merge point at the same time. When starting again, the vehicle will start at 0 m/s, and the (additional) vehicle safe distance is only one times the length of the vehicle [7]. From the displacement formula $s = \frac{1}{2}at^2$, we can derive $t = \sqrt{2s/a}$. The average acceleration of such a vehicle is 6.5 ft/s^2 [1]. Substituting, we find

$$\sqrt{\frac{15 \text{ ft} + 15 \text{ ft}}{6.5 \text{ ft/s}^2}} \approx 3.0 \text{ s.}$$

  As before, we take the reciprocal and convert to an hourly basis, getting $\mu_1 = 1185$ vehicles/hr.

Time Cost at Toll Booths for the Cellular Model

Based on the arrival rate of a lane as given above, we can calculate the average time spent by a vehicle at the toll booth according to the formula below given in [8]. The formula applies to both the traditional toll plaza and our new design, since in the cellular toll plaza, each toll booth faces the same traffic flow as one lane of the traditional toll plaza.

$$W_T = \frac{1}{\mu_T - \dfrac{\Phi}{B}},$$

where

- W_T is the time cost passing through a toll booth,

- μ_T is the service rate of a toll booth,
- Φ is the traffic flow, and
- B is the number of toll booths.

Time Cost at Merge Points for the Cellular Model

The merge process at each merge point is essentially a birth-death process. **Figure 7** describes the state transition of this process in the form of a Markov chain.

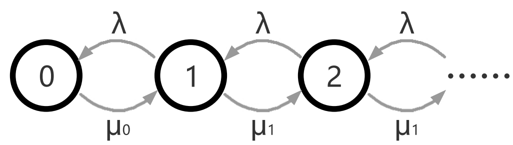


Figure 7. State transitions of a birth-death process.

In this process, each state follows the rule that the sum of the transit-in probabilities equals to the sum of the transit-out probabilities [5], and the probability sum of all events is 1. We let P_i be the probability of being in state i. Then we have the equations below:

$$\begin{aligned} \lambda P_0 &= \mu_0 P_1; \\ \lambda P_1 + \mu_0 P_1 &= \lambda P_0 + \mu_1 P_2; \\ \lambda P_n + \mu_0 P_n &= \lambda P_{n-1} + \mu_1 P_{n+1}, \qquad n \geq 2; \\ \sum_{i=0}^{\infty} P_i &= 1, \end{aligned}$$

where our interpretations in the toll plaza setting are:

- P_n is the probability of n vehicles in the system,
- λ is the arrival rate at a merge point,
- μ_0 is the service rate at a merge point when there is no merging conflict,
- μ_1 is the service rate at a merge point when merging conflict occurs.

Solving the equations above for the values of the P_i, we obtain:

$$\begin{aligned} P_0 &= \left(1 + \frac{\lambda}{\mu_0} + \frac{2\lambda\mu_1}{\mu_0^2\mu_1 - \lambda\mu_0^2 + \lambda\mu_0\mu_1 - \lambda^2\mu_0}\right)^{-1}, \\ P_1 &= \frac{\lambda}{\mu_0} P_0, \\ P_n &= \frac{2\lambda^2}{\mu_0^2 + \lambda\mu_0} \left(\frac{\lambda}{\mu_1}\right)^{n-2} P_0, \qquad n \geq 2. \end{aligned}$$

From the probabilities obtained above, we can calculate the expected number of vehicles L_s in the whole queuing system:

$$L_s(\lambda) = \sum_{i=1}^{\infty} iP_i = \frac{\lambda}{\mu_1 - \lambda} + \frac{\lambda\mu_1 - \lambda\mu_0}{\lambda\mu_1 - \lambda\mu_0 + \mu_0\mu_1}.$$

L_s is also called the average queue length. According to Little's Law [5], the average waiting time W_s for a vehicle at a merge point is

$$W_s = \frac{L_s}{\lambda}.$$

Total Time Costs for the Two Models

According to our assumptions and calculations, the traffic flow at the kth merge point in a traditional toll plaza is

$$\frac{(k+1)\Phi}{B}, \quad k = 1, \ldots, (B - L + 1).$$

Hence the probability of arrival of a vehicle at the kth merge point is $(k+1)/B$.

According to the formulas above, the total average time cost at the merge point is

$$W_{MT} = \frac{L-1}{B} \cdot \frac{B-L}{\mu_0} + \frac{B-L+1}{B} \sum_{k=1}^{B-L} \frac{k+1}{B} W_s \left(\frac{k+1}{B} \Phi \right).$$

Adding the time cost passing through each toll booth obtained above, we can calculate the average time cost passing through the whole *traditional toll plaza*:

$$\begin{aligned} W_{AT} &= W_T + W_{MT} \\ &= \frac{1}{\mu_T - \dfrac{\Phi}{B}} + \frac{L-1}{B} \cdot \frac{B-L}{\mu_0} + \frac{B-L+1}{B} \sum_{k=1}^{B-L} \frac{k+1}{B} W_s \left(\frac{k+1}{B} \Phi \right). \end{aligned}$$

But in our design, since the traffic flow merges in advance, the traffic flow of each lane becomes twice that of the previous lane, and the number of lanes is reduced by half. To simplify the calculation, we may assume that B is always even, so that the traffic flow at the kth merge point is

$$\frac{2(k+1)\Phi}{B}, \quad k = 1, \ldots, \left(\frac{B}{2} - L + 1 \right),$$

so the probability of an arrival at the kth merge point is

$$\frac{2(k+1)}{B}, \quad k=1,\ldots,\left(\frac{B}{2}-L+1\right).$$

The total average time spent by a vehicle at merge points in the cellular toll plaza is is

$$W_{MI}=\frac{2(L-1)}{B}\cdot\frac{B-2L}{2\mu_0}+\frac{B-2L+2}{B}\sum_{k=1}^{\frac{B}{2}-L}\frac{2(k+1)}{B}W_s\left(\frac{2(k+1)}{B}\Phi\right).$$

In a cellular toll plaza, all lanes are merged in advance. So we need to calculate the additional time cost of pre-merging process:

$$W_{Ex}=W_s\left(\frac{2\Phi}{B}\right).$$

Adding together all the time costs, we find that the average time cost for each vehicle passing through the *cellular toll plaza* is

$$\begin{aligned}W_{AI}&=W_T+W_{MI}+W_{Ex}\\&=\frac{1}{\mu_T-\dfrac{\Phi}{B}}+\frac{L-1}{B}\cdot\frac{B-L}{\mu_0}+\frac{B-L+1}{B}\sum_{k=1}^{B-L}\frac{k+1}{B}W_s\left(\frac{k+1}{B}\Phi\right)\\&\quad+W_s\left(\frac{2\Phi}{B}\right).\end{aligned}$$

Substituting the specific values of the parameters and plotting, we see the comparison results shown in **Figure 8**.

Improve the Accident Prevention Ability

Hierarchical Merge Pattern

A toll booth in a cellular toll plaza has only one merge point at the end of the transition zone, increasing the possibility of it being overcrowded. However, in our cellular toll plaza, there are bends among the cells, keeping the speed of vehicles within a safe range. Therefore, we decrease the possibility of accidents caused by traffic congestion as well as those caused by excessive speed.

Figure 9 and **Figure 10** illustrate the contrast.

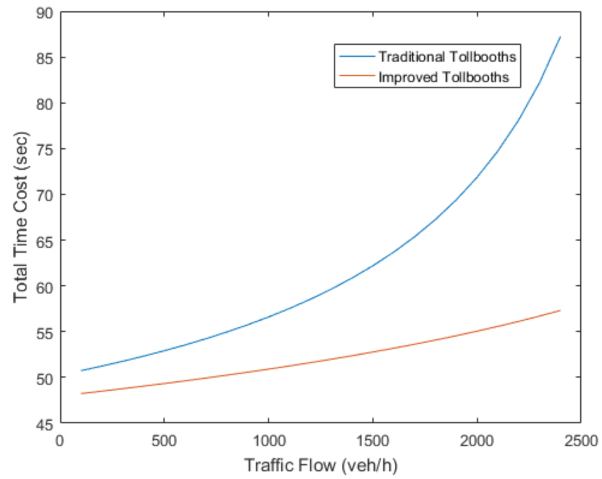


Figure 8. Comparison of total time cost to pass through a traditional toll plaza vs. through a cellular toll plaza, for different values of traffic flow.

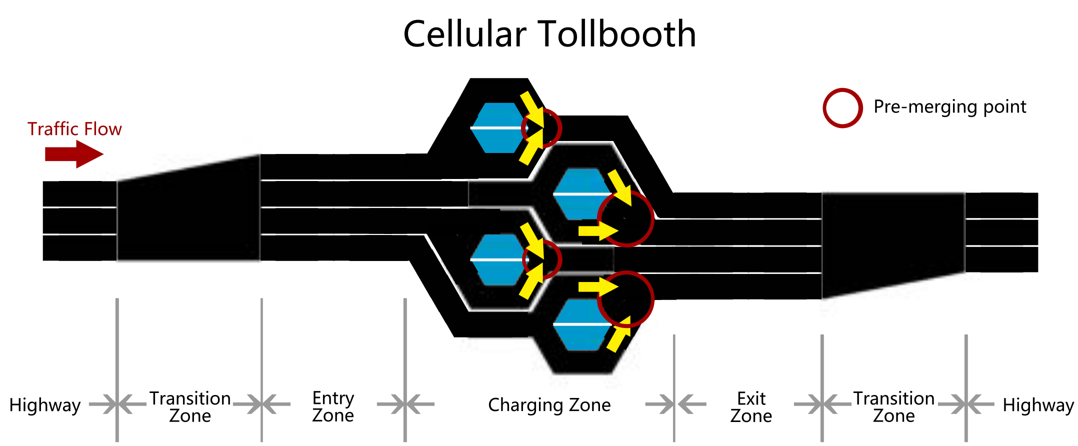


Figure 9. Merging in a cellular toll plaza: low congestion, low speed.

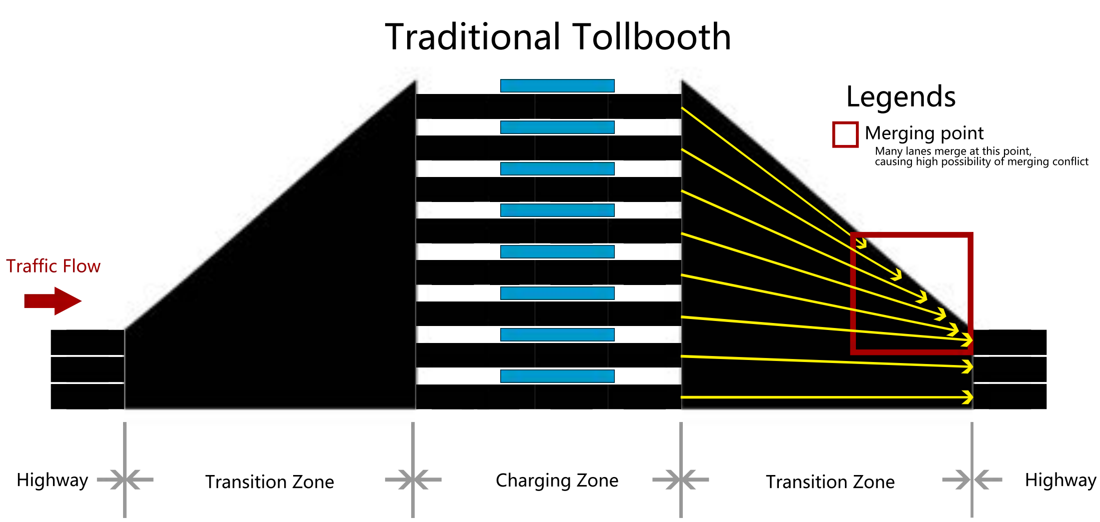


Figure 10. Merging in a traditional toll plaza: high possibility of conflict, potentially higher speed.

A Gentle Design of the Transition Zone

The gradient of the transition zone is set according to the design speed and the tangential offset width; different countries have different standards: The maximum ratio of the U.S. standard is 1:20, and the minimum is 1:8 [9].

Therefore, we further improve the model to improve safety, by changing the gradient rate of cross-section of the toll plaza, as shown in **Figure 11**.

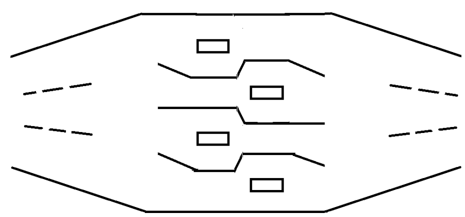


Figure 11. Improved design of the cellular toll plaza.

More Suitable for ETC Technology

Toll plazas typically include a different set of charging models: conventional (human-staffed) toll booths, exact-change (automated) toll booths, and electronic toll collection (ETC) toll booths. Vehicles near the entrance to the toll plaza often encounter traffic accidents due to the choice of different access routes; so the locations of the different types of toll booths is also critical to safety.

With ETC technology, the ways vehicles enter, drive through, and exit the toll plaza are different from the traditional charging pattern. A computational experiment showed that booths associated to higher-risk traffic flows—e.g., traffic such as that directed to ETC toll booths, which approaches the toll plaza at higher speeds—should be located in a central position with respect to other booth types [10]. So, we add two ETC lanes in the middle of the new toll plaza (see **Figure 12**).

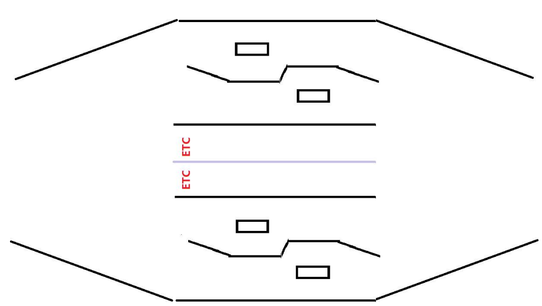


Figure 12. Further improved design of the cellular toll plaza.

The Influence of Autonomous Vehicles

Compared to the traditional toll plaza, a cellular toll plaza can better meet the needs of autonomous ("smart") cars, since they do not pay in cash

We first analyze the principle and characteristics of autonomous vehicles, then optimize the cellular toll plaza model.

Characteristics of Autonomous Vehicles

- Autonomous cars need to be equipped with an automatic payment system, that is, ETC equipment, which means that they can pass through the toll plaza quickly.
- Self-driving vehicles have better control, which can reduce the possibility of accidents in the toll plaza. Besides, a driver's own factors (such as bad mood, disputes with the toll plaza service staff, etc.) will not affect the vehicle's safety of an autonomous vehicle [11].
- Vehicles at the junction of the toll plaza convergence can be more orderly, thus avoiding congestion and maximizing the efficiency of cellular toll plazas.

Our Solution

- Since in the future autonomous vehicles will be more numerous, to maximize efficiency, cellular toll plazas must increase the number of automatic toll booths and reduce the number of manual toll booths . The remaining MTC lanes would be located at the sides of the toll plaza, with straight lanes for large vehicles set in the middle. Since autonomous vehicles are easy to steer, the remaining lanes would all be ETC lanes.
- According to the queuing model and the simulation result from VISSIM, the throughput of the toll plaza increases with a greater proportion of ETC lanes. Since autonomous vehicles are all non-cash payment, when compared with the traditional toll plaza, the cellular design is more suitable.

Analysis of our Design

Using Simulation

Basic Data for the Simulation

We use PTV-VISSIM 4.3 to do this simulation [12]. We set the speed of the vehicle through the ETC deceleration belt at 24 km/h [13] and deceleration is 2 m/s^2.

We did two pairs of simulations comparing the cellular toll plaza to a traditional one. Both sets had 3 highway lanes and 8 toll booths. The difference was in the number of ETC lanes: 8 in the first simulation, and 2 in the second.

We show the results of the simulations in **Figure 13** and **Figure 14**.

Here we address some questions that may arise in the mind of the reader:

- Q: Why consider ETC tolling but not exact-change toll booths?

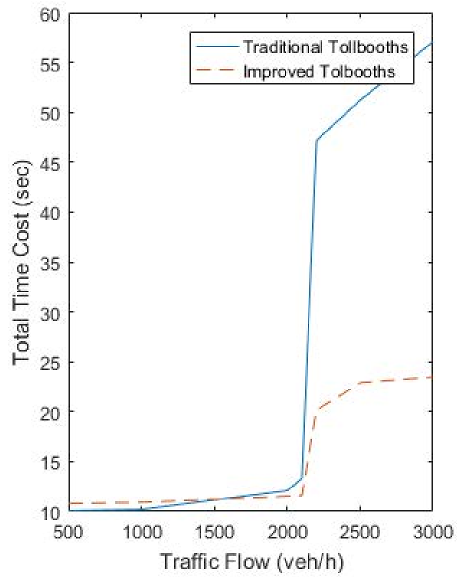


Figure 13. Total time cost vs. traffic flow, for traditional toll plaza (solid curve) and cellular toll plaza (dotted curve) with 8 lanes, all ETC.

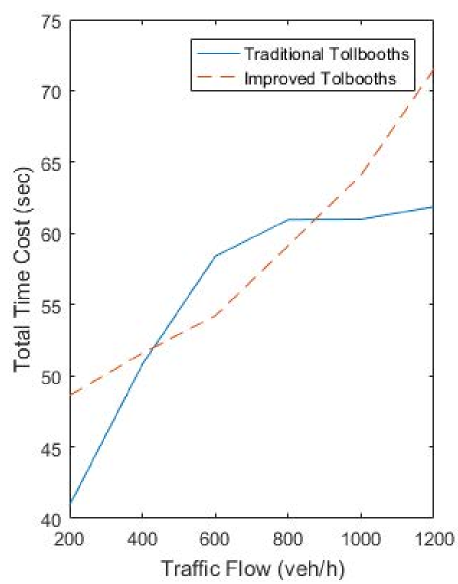


Figure 14. Total time cost vs. traffic flow, for traditional toll plaza (solid curve) and cellular toll plaza (dotted curve) with 2 ETC and 6 MTC lanes.

  - A: If we considered all possible charging patterns, our model would be too complicated. According to [14], there is little difference between the performance of MTC toll booths and that of exact-change toll booths: 425 vehicles/hr vs. 500 vehicles/hr. So we can combine the MTC method and the exact-change method into a single "cash" method. ETC systems can achieve between 1,200 vehicles/hr and 1,800 vehicles/hr.

- Q: Why not consider autonomous vehicles?
  - A: Because autonomous vehicles don't need a driver. Just install an ETC device on the car.

- Q: How to explain the great difference between the result of the queuing model and the result of the VISSIM simulation?
  - A: The VISSIM software has taken a lot of factors into consideration. Therefore, compared with the pure theoretical derivation, VISSIM is more practical.

- Q: How to explain the great change at a traffic flow of 2,000 vehicles/hr.
  - A: Both kinds of toll plaza have a maximum capacity and throughput.

Simulation Conclusions

- **All lanes ETC** (**Figure 13**) When the toll plaza is configured to all lanes ETC, the traditional toll plaza and the cellular plaza have about the same throughput for light traffic. But the cellular plaza is better than the traditional plaza in heavy traffic (over 2,200 vehicles/hr), with less than half as much total time cost. The simulation results are in good agreement with the results of queuing theory.

- **2 ETC lanes, 6 MTC lanes** (**Figure 14**) The cellular plaza produces less total time delay for traffic flows between 400 vehicles/hr and 900 vehicles/hr.

The Influence of Different Kinds of Toll Booths

Results of the Analysis

Figure 15 shows a comparison of total time cost (and also time delayed) for 0 through 8 ETC lanes (and correspondingly 8 through 0 cash lanes), for a traffic flow of 2,400 vehicles/hr.

The capacity of a cellular toll plaza is more sensitive to the proportion of ETC lanes than a traditional toll plaza. The smaller the proportion of cash lanes, the shorter the average transit time and average delay.

With 8 ETC lanes, all cars pass through without any time delay. Moreover, we have:

The Rule of 8: 8 ETC lanes are 8 times as fast as 8 cash lanes.

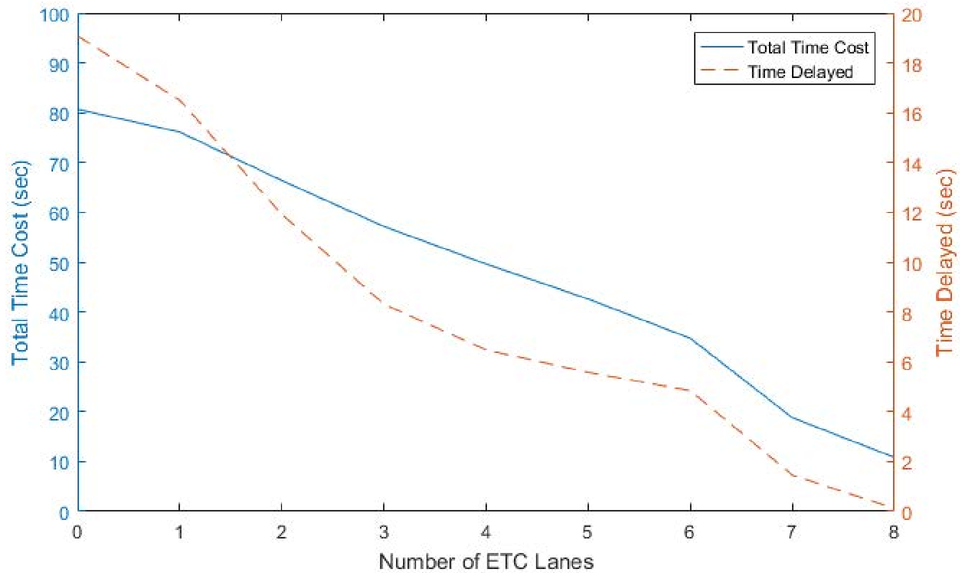


Figure 15. Total time cost and time delayed vs. number of ETC lanes (out of 8), for a cellular toll plaza with 8 lanes.

Strengths and Weaknesses

Strengths

- Cellular toll plazas would save land area and reduce construction costs.
- The pre-merging in cellular toll plazas would prevent congestion.
- Cellular toll plazas would force deceleration and slower speeds through the toll plaza, preventing accidents caused by the speed difference between ETC lanes and MTC lanes.
- Our simulations confirm results of our queuing model. Simulation results show that a cellular toll plaza with ETC lanes would have better results than a traditional toll plaza, especially when traffic is heavy.
- Our simulations produced results not directly obtainable from the queuing model, for combinations of ETC and MTC.

Weaknesses

- The length of vehicles should be taken into consideration, since some long vehicles could not pass through the cellular toll plaza.
- Because actual data on traffic composition, speed, and acceleration are not easy to obtain, the traffic simulations may not be close enough to reality. We had to consult a large number of references to determine the average speed of a vehicle passing through a toll plaza, and we then needed to simulate its progress through the deceleration zone.

- In the simulations, we ignore the difference between exact-change lanes and MTC lanes.

Conclusion

With the popularization of ETC equipment and autonomous vehicles, MTC lanes will be totally replaced by ETC lanes in the next 20 years. That change will increase the road capacity and decrease the time cost for each car.

A traditional toll plaza covers a large area, and the cost of construction is high. There is congestion at merge points, which costs time and can lead to accidents.

To solve these problems, we put forward a new design, that of a cellular toll plaza, inspired by the honeycomb. Toll booths are located on hexagons. Thanks to this special structure, the cost of construction can be significantly reduced. Meanwhile, by pre-merging the traffic inside the toll plaza, the expected time cost for vehicles in merging can be greatly diminished. In addition, by appropriate design, the probability of collision at merge points can be decreased.

References

[1] Wang, Dianhai. 2002. *Traffic Flow Theory*. Beijing: China Communications Press.

[2] Wu, Xiaowu. 2004. Study on traffic safety of toll station of expressway. Chang'an University.

[3] Wikipedia. 2017. Honeycomb. `https://en.wikipedia.org/wiki/Honeycomb`.

[4] Wikipedia. 2017. Queueing theory. `https://en.wikipedia.org/wiki/Queueing_theory`.

[5] Gross, D. 2008. *Fundamentals of Queueing Theory*. New York: John Wiley & Sons.

[6] Washington State Department of Transportation. 2017. *Design Manual*, Chapter 1210, Geometric Plan Elements. `http://www.wsdot.wa.gov/publications/manuals/fulltext/M22-01/1210.pdf`.

[7] Transportation Research Board. 2017. *Highway Capacity Manual*. 6th ed. Washington, DC: Transportation Research Board.

[8] Hock, Ng Chee. 1996. *Queueing Modeling Fundamentals*. New York: Wiley.

[9] CHENG, Jianxing, and PING Jiang. 1995. Design of toll collection station for Quanzhou to Xiamen freeway. *Journal of Highway & Transportation Research and Development*.

[10] Pratelli A., and F. Schoen. 2006. Multi-toll-type motorway stations optimal layout. In *Urban Transport XII: Urban Transport and the Environment in the 21st Century*, 911–921. https://www.researchgate.net/profile/Antonio_Pratelli/publication/271450547_Multi-toll-type_motorway_stations_optimal_layout/links/574308e408ae9ace8418becd.pdf .

[11] Sivak, Michael, and Brandon Schoettle. 2015. Road safety with self-driving vehicles: General limitations and road sharing with conventional vehicles. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/111735/103187.pdf?sequ .

[12] PTV Group. n.d. Vissim. http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/ .

[13] Liu, Lili, Jiancheng Weng, and Jian Rong. 2012. Simulation based mixed ETC/MTC freeway toll station capacity. 19th ITS World Congress.

[14] Ding, Chuangxin. 2005. Study on capacity of high-grade highway toll station. Kunming University of Science and Technology.

Dear New Jersey Turnpike Administration:

It is our pleasure to offer policy recommendations regarding tolling on the New Jersey Turnpike. We have developed a new toll plaza design that can solve many problems caused by the increase of ETC service and autonomous cars.

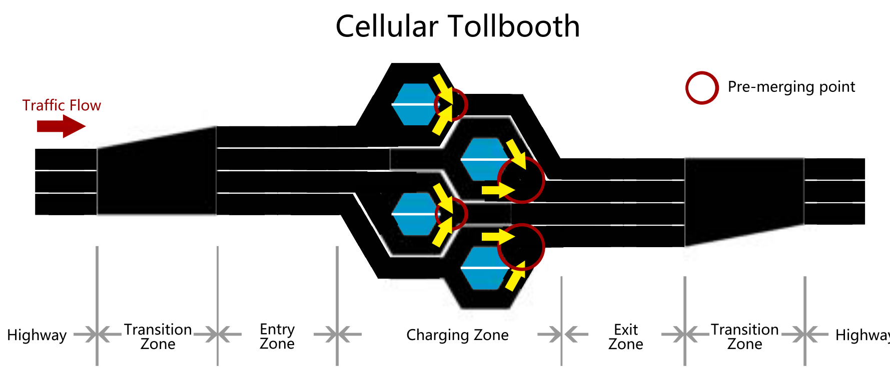
The New Jersey Turnpike is about 200 miles long, the fifth longest of U.S. tollroads. The ETC utilization rate of toll booths of New Jersey is over 80%. In the future, the use of ETC is likely to increase. Meanwhile, progress in autonomous vehicles is accelerating; so in the next 20 years there may be a large number of autonomous vehicles on the Turnpike, which could pass through toll plazas at a faster speed. Toll plazas need to keep up with increased traffic flow and ensure the safety of vehicles.

We have come up with a cellular toll plaza design, whose conceptual model is shown below.

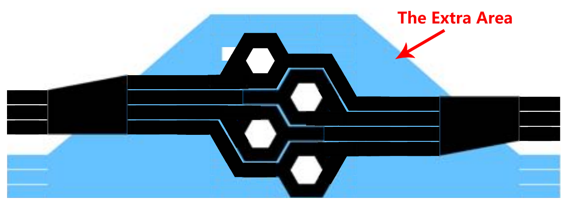
This design has four main advantages:

1. The cellular architecture can reduce plaza area about 50% compared to a traditional toll plaza design. In the figure below, the black part is our design for 8 toll booths, while the blue part is the traditional design.

2. Simulation results show that cellular toll plazas have better performance than traditional toll plazas, especially under heavy traffic, reducing average travel time through the plaza by about 50%.



Conceptual design of cellular toll plaza.



Comparison of the footprints of the cellular toll plaza and a traditional toll plaza.

3. The Rule of 8: A plaza with 8 ETC toll booths is 8 times as fast as one with 8 MTC toll booths.
4. We simulated the performance of a cellular toll plaza under different traffic levels. The average transit time remains about 11 seconds for up to 2,000 vehicles/hr.

To give you a more intuitive view of the advantages of our cellular toll plaza, we have made a comparison diagram for you, shown above.

Furthermore, to help you understand the advantages of the cellular toll plaza with more ETC lanes, we offer a graph on p. 312 of our technical report, where the blue solid line stands for the performance of the traditional design and the red dashed line is for our design. You can see that with heavy traffic (but below the absolute capacity of the toll plaza), the time cost passing through the cellular toll plaza is lower than for the traditional one.

Thank you again for taking the time to read our suggestions. We sincerely hope that the cellular toll plaza can solve the congestion problem on the New Jersey Turnpike.

Sincerely,

The Toll Team

Team members Jiahua Zhang, Yang Liu, and Rui Liu, and advisor Hui Ji.

Author's Commentary: The Toll Plaza Problem

Michael Tortorella
Managing Director
Assured Networks, LLC
Middletown, NJ
assurenet@verizon.net

Introduction

Entrants choosing the 2017 Toll Plaza Problem presented judges with an interesting conundrum. While some papers were quite good, none of the papers reaching final judging contained all of the features that are expected of Outstanding papers.

Every paper had something missing, and different papers were missing different things. Some papers considered the required extensions but did not include any sensitivity analysis. A few papers incorporated a sensitivity analysis but did not include a realistic assessment of strengths and weaknesses.

As a rule, the first thing judges look for is whether a paper satisfies all the clearly-stated requirements of the problem. The problem-setting committee felt that this year's problem statement was especially clear about required elements; indeed, considerable pains were taken by the committee to make sure that the problem statement was clear and simple.

It was surprising, then, that so many papers failed to address the requirements. Many papers did address most of the requirements, but none addressed all of the requirements.

Remember, also, that there are also general MCM requirements of all entrants. These are not usually stated again in the problem statement; but good solutions will always address sensitivity analysis, assessment of model strengths and weaknesses, and explicit model construction and execution.

We urge coaches and faculty advisers to train teams to list for themselves all the requirements that can be identified from a problem statement, as

The UMAP Journal 38 (3) (2018) 319–322.

well as the general MCM requirements, and from time to time during the solution process to re-examine the list to make sure everything is being considered appropriately.

Read the Problem Statement!

There are also specific issues concerning this particular problem. A prominent one is that many teams did not read the problem statement carefully. They appeared to think that the problem asked them to solve for the proper number B of tollbooths. This was in fact an MCM problem in 2005. This year's problem statement clearly indicates that B (and the number L of highway lanes) are given, and the solution (the optimal merge-area design) is to be expressed as a function of B and L. There is no need to determine B (indeed, there is not enough information given in the problem to determine B), and many teams—even some who went on to deal with the real problem at hand—wasted time on this.

The takeaway for coaches and faculty advisers is to train their teams to read the problem carefully and thoroughly so that mistakes like this do not happen.

Similarly, even though the problem statement explicitly cautions against this, many teams engaged in extensive performance analysis of existing merge-area designs when the problem instead calls for design improvements. Indeed, this is fundamentally a multi-criterion optimization problem: Given B, L, and the various other constraints, what is the best merge-area shape and size, considering safety, throughput, and cost?

Perhaps it is too much to expect that undergraduates provide a full solution to this optimization in detail. However, a crucial point for the teams is that they need to penetrate the problem enough to be able to determine what the linchpin of the problem is—in this case, the multi-criterion optimization.

A major purpose of the MCM is to foster model-building skill in the participants. One of the key components of that skill is the ability to discern what approaches are germane to the problem and may lead to fruitful results, and what approaches are simply not aligned with the problem needs and should be avoided.

Cite References Appropriately

Teams need to learn to cite references appropriately. The judges found many reports that used the same illustration (an aerial photo of a toll plaza on the Garden State Parkway) without any attribution. Similarly, many reports included pieces of mathematics that were drawn from other sources without citation of those sources.

There is no shame in acknowledging in print that you are using the work of others to build your own solution on. Indeed, the opposite is true: It is unethical to take the work of others and incorporate it into your own work without giving proper credit through a formal and focused citation.

Write a Succinct Summary

The summary or abstract must contain a capsule summary of the results. Most teams use the summary as a brief description of the work that they did or the methods that they used. *This is not satisfactory*, and reports whose summary did not include any results were downgraded.

The purpose of the summary is to entice the reader into wanting to read more of your report. This is better done by including some results from the solution, because doing so arouses the reader's curiosity: How did they get that? Does this make sense? A summary with some results and no description of method is far superior to a summary that describes methods used but no results obtained. Coaches and faculty advisors, please take note of this important point. The vast majority of papers did not ever get to the point of saying what their solution was in specific, numerical terms: "This is the shape of the merge area, and it is so many meters long and so many meters wide" (or other appropriate size description). Someone who is going to build a merge area based on the recommendations in a report needs these dimensions; without them, no building can take place.

Coaches and faculty advisers are urged to train their teams to ask what a solution should look like. In this case, specific numbers are required. Many teams used a cellular automaton (CA) model. However, a CA model necessarily operates on one pair (B, L) at a time. Solution of this problem required consideration of a whole space of possible (B, L) values to find the best merging area shape and size for each B and L.

Indeed, a significant challenge was contained in the fact that once a merge-area shape and size is chosen for a particular B and L, the merge area must remain suitable under changing conditions of traffic flow, weather, and other variables—here is where sensitivity analysis comes into play. It's not clear how you would perform this kind of analysis with a single CA solution. Many teams considered one or two of the three performance criteria (throughput, safety, and cost); but very few considered all three together, as was required. The problem statement was explicit about this, as well as about the need to avoid simply doing a performance analysis of an existing merge-area shape, and rather determining a better or best shape considering the requirements. We again urge coaches and faculty advisers to train their teams to read problem statements closely and summarize the conditions and requirements for themselves, as an aid to staying on track during the solution and report-writing processes.

Smaller Issues

Some smaller issues include:

- Sloppy notation: Many teams used L for length (of something). But L is specified in the problem statement as the number of travel lanes exiting the merge area—you can't use it for anything else.
- Too few teams used diagrams appropriately in their report. A solution to the problem can be described without a diagram, but it is immeasurably easier for a reader if some simple diagram is included. Teams are advised to make liberal use of diagrams not only for their own understanding but also for clearer communication in their reports.

Conclusion

The Toll Plaza Problem posed interesting challenges for teams and judges alike. I hope that coaches and faculty advisers can take away some useful advice from this summary, while judges continue to concentrate on fairness and the difficult task of discerning the true meaning of reports written by nonprofessionals under time pressure.

About the Author

Mike Tortorella is Managing Director of Assured Networks, LLC, and Adjunct Professor of Systems Engineering at Stevens Institute of Technology. He retired from Bell Laboratories as a Distinguished Member of the Technical Staff after 26 years of service. He holds the Ph.D. degree in mathematics from Purdue University. His current interests include stochastic flow networks, network resiliency and critical infrastructure protection, and stochastic processes in reliability and performance modeling. Mike has been a judge at the MCM since 1993 and particularly enjoys the MCM problems that have a practical flavor of mathematical analysis of social policy. Mike enjoys amateur radio, playing the piano, and cycling.

The Effects of Self-Driving Vehicles on Traffic Capacity

Yu Shi
Jiefeng Chen
Qi Li

Computer Science and Engineering
Shanghai Jao Tong University
Shanghai
China
Advisor: Xiaofeng Gao

Abstract

We propose an innovative model based on cellular automata to analyze the effects on highway traffic capacity of self-driving cooperating vehicles. We extend the traditional cellular automata used for traffic analysis to increase the flexibility of our model, so that it can deal with the complex dynamics of self-driving vehicles.

We first model the car generation process using Poisson distributions. This model generates a specified volume of vehicles on a segment of highway. Our car-generation model distinguishes the traffic volumes during peak hours and non-peak hours. Then we model the behaviors of vehicle-following and lane-changing. In this part, we consider cooperation between self-driving vehicles and interaction between self-driving and human-driven vehicles. Finally, we propose a probability model to analyze the behavior of traffic flow at highway intersections.

We verify the feasibility of our model by running various simulations. We test our model with varied percentages of self-driving cars, different numbers of lanes, and varied traffic volumes. We find that self-driving cars can stabilize the performance of traffic system when the number of lanes and the traffic volume change.

Then we apply our model to the data of four highways in the Greater Seattle area. Our results show that the traffic capacity increases with a rising percentage of self-driving cooperating vehicles. The average speed of traffic flow doubles when the percentage increases from 10% to 50%, and in most cases an equilibrium exists after 60%. We also test the effects of dedicated lanes for self-driving cars. When the percentage of self-driving vehicles is higher than 70% percent, dedicated lanes can be helpful. We also compare

The UMAP Journal 38 (3) (2017) 323–346.

performance during peak hours with average performance during the day, and find that self-driving cars significantly speed up peak-hour traffic flows.

Finally, we analyze the effect of potential errors of self-driving systems. We find that defects in self-driving systems can degrade the benefit to traffic capacity. We also do sensitivity analysis of various parameters, including the probability of lane-changing and parameters in the vehicle dynamics, to prove the robustness of our model.

Introduction

Self-driving is a technology with growing importance. The rise of self-driving cooperating vehicles has long been a hope for increasing traffic capacity without enlarging the roads, since these vehicles have more accurate control, better sense, and greater prediction capability of the behavior of other vehicles, than human-driven vehicles have.

We make a quantitative and qualitative analysis of the impact on traffic capacity of such vehicles.

We establish a traffic flow model based on cellular automata to simulate highway traffic flow. We extend the traditional model to taking into account acceleration of vehicles. This extension brings flexibility in modeling the traffic flow, and enables us to test the impact of more-detailed control rules for self-driving cooperating cars.

Our model can be divided into four parts.

- We propose a traffic generation model based on Poisson distributions. This model captures the variation in traffic during peak travel hours and non-peak hours.
- We model the behavior of vehicle-following, which defines how a car acts according to what the car ahead does. We apply different rules for the cooperation between two self-driving cars than for interactions between self-driving and human-driven cars.
- We design a lane-changing model for vehicle movement across lanes.
- Finally, we build an intersection model to describe how vehicles change from one highway to another through interchanges.

We analyze the feasibility of our model by running simulations with varied percentage of self-driving vehicles, different numbers of lanes, and varied daily traffic volumes. Under all these simulations our model produces reasonable results, which is consistent with the properties of self-driving vehicles.

After that, we apply the provided data for Greater Seattle. We analyze how the highway capacity changes with the percentage of self-driving cars. Results show that an equilibrium exists at around 50% to 70% self-driving cars. We also study the effect of dedicated lanes for self-driving vehicles, and propose a strategy to allocate such lanes.

Finally, we do sensitive analysis on some parameters in our model. We also consider what if self-driving systems are not perfect. In other words, autonomous vehicles may introduce new risks. We simulate the inaccuracy of a self-driving system by a small probability of random acceleration or deceleration. Results shows that such inaccuracy can degrade the benefit of self-driving vehicles.

Literature Review

Using cellular automata to analyze traffic flow is widely used by many researchers [1; 2]. This approach has been proven effective for analyzing real traffic flow, and it's easy to implement. However, these researchers did not consider the scenario with self-driving cars. Their human-driven models lack flexibility to adapt to self-driving vehicles.

Most studies of self-driving vehicles in recent years use continuous models with complex simulation software and focus on the dynamics of self-driving vehicles and the interactions of several vehicles [3; 4]. But how self-driving technology will influence the whole highway traffic flow is still not answered.

Some researchers in civil and environmental engineering believe that self-driving technology can increase highway capacity and reduce air pollution [5]. But they don't provide quantitative analysis.

We propose a new cellular automata system that models the impact of self-driving vehicles on highway traffic flow. We extend the traditional traffic flow methods by adding the dynamics of self-driving vehicles. We analyze our model with simulations and verify its feasibility. Finally, we apply our model to real data and provide constructive advice on traffic management for self-driving cars.

Assumptions

- Each highway is a straight line without curves.
- Self-driving cars can communicate real-time data with one another, including speed and acceleration.
- Self-driving cars do not need response time to take action as human drivers do.
- Human-driven vehicles do not communicate with any other vehicle.
- The width of a lane is only enough for a single vehicle.
- All vehicles are the same size.
- Cars change to another highway through interchanges in intersections, without crossing any traffic flow.
- All human-driven cars are identical, and all self-driving cars are identical.

The Models

Basic Settings of cellular Automata

Cellular automata has been a widely-used model for traffic flow analysis [1]. The road is divided into cells, with each cell having one car or no car in it. A cell also records the current speed of the car in it, if any. In cellular automata, positions, time, speed, and acceleration are all discrete. The state of a cell is updated according to its current state and the current states of surrounding cells.

In our models, we apply different updating rules to distinguish the behaviors of self-driving and human-driven cars.

Notations

Table 1.

Notation.

Symbol	Meaning
s	An indicator representing the car type in a cell
v	Current speed of a car
v_{new}	The newly-updated value of v after each iteration
a	Current acceleration of a car
a_{new}	The newly-updated value of a after each iteration
N	Daily traffic count of a highway segment
n	Newly arriving vehicles in a highway segment in one second
p	Number of peak hours.
q	Percentage of self-driving cars
G	Distance between two cars in the vehicle-following model
v_p	Speed of the preceding vehicle in the vehicle-following model
a_p	Acceleration of the preceding car in the vehicle-following model
T_r	Response time for a human driver to take action
V_{min}	Minimum speed allowed on a highway
V_{max}	Maximum speed allowed on a highway
λ_{peak}	Rate parameter of Poisson distribution for car generation during peak hours
$\lambda_{\text{non-peak}}$	Rate parameter of Poisson distribution for car generation during non-peak hours
T_{rh}	Response time for a human driver
T_{rs}	Response time for a self-driving car
T_{inter}	Time needed to transfer from one highway to another through an interchange
μ_{ij}	Probability of crossing interchanges in intersection Model
P_a	Acceleration probability when $G > G_S$
P_d	Deceleration probability when $G > G_S$
P_c	Probability of lane-changing
α, β, γ	Coefficients for vehicles following rules

Introducing Acceleration into cellular Automata

Most cellular automata used for traffic analysis consider the speed of vehicles without considering acceleration [1; 2; 6]. In such models, when a speed changes, it can increase or decrease by only a fixed value. A human driver has no accurate perception of acceleration, so for a cellular automata model for human-driven cars, considering only the speed of cars is enough.

However, researchers of self-driving cars use acceleration as an important variable in their analysis [3; 7; 8]. For such vehicles, a fixed acceleration increment is too limiting, given their capabilities .

- **Self-Driving** Autonomous vehicles are often equipped with accurate sensors[9], a complex control system, and even artificial intelligence technology. They can control acceleration accurately. Thus, they can steer more flexibly and fearlessly than human-driven ones.
- **Cooperating** These vehicles can share data including their current acceleration [4]; and with these data, they can take more accurate actions to avoid collision as well as make full use of space on the highway. Considering acceleration in our model makes more-complex cooperating rules possible.

Thus, we extend the traditional cellular automata traffic flow model by adding acceleration as an attribute of a cell.

Attributes of a Cell

In our model, a cell represents a 4-meter long space on a single lane, and each cell can contain only one car at a time. A cell has three attributes:

1. An integer s to represent the car type: 0 for no car, 1 for human-driven, and 2 for self-driving.
2. Current speed v of a car.
3. Current acceleration a of a car.

Only when s is nonzero do the other two attributes make sense.

In each step of the simulation, a car moves forward by v cells, and v is updated by adding a; v is bounded by $V_{\min}$ and $V_{\max}$, the speed limitations of highway. Rules are applied to update a according to the states of surrounding cells. For human-driven and self-driving vehicles, we have different rules, introduced in following sections.

Generation Model

We need a model to generate cars for each highway segment; a Poisson distribution is appropriate for this task [10]. It expresses the probability of a number of events occurring in a fixed interval of time. Instead of using one Poisson distribution for all time during the day, we use one for peak hours and another for non-peak hours. Our generation model also controls the percentage of self-driving cars.

We assume that there are p peak hours in a day and 8% of the daily traffic volume occurs during peak hours. With the data of daily vehicle counts, we can calculate the average number of newly-arriving cars in one second in each highway segment:

$$\lambda_{\text{peak}} = \frac{0.08N}{3600p}, \qquad \lambda_{\text{only non-peak}} = \frac{0.92N}{3600(24-p)}$$

Supposing that the arrival of vehicles follows Poisson distributions for peak and for non-peak hours, then λ_{peak} and $\lambda_{\text{non-peak}}$ are the corresponding rate parameters. Thus, each second we can sample from the Poisson distributions

$$P_{\text{peak}}(n) = \frac{\lambda^n_{\text{peak}}}{n!}\, e^{-\lambda_{\text{peak}}}, \qquad P_{\text{non-peak}}(n) = \frac{\lambda^n_{\text{non-peak}}}{n!}\, e^{-\lambda_{\text{non-peak}}}$$

for the number of newly-entering vehicles onto a highway segment in our simulation, during peak hours and non-peak hours, respectively.

By adjusting p, we can control the average frequency of vehicles arriving at peak hours, i.e., λ_{peak}. Since the proportion of peak hours is fixed at 8%, when p is larger, these 8% of vehicles can arrive at a relatively low frequency. We will simulate with different values of p to test the influence of peak and average traffic volume on our model. **Figure 1** shows how λ_{peak} and $\lambda_{\text{non-peak}}$ vary with respect to p, when the daily traffic count is 200,000.

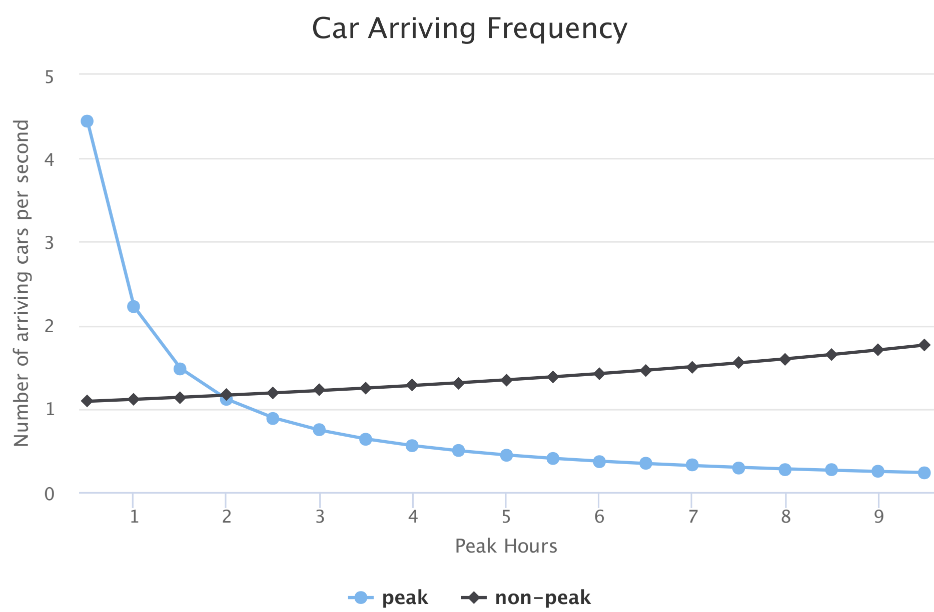


Figure 1. Car-arriving frequency at peak and at non-peak hours for a daily traffic count of 200,000.

From **Figure 1**, when $p = 1$ we have about 2 newly-arriving cars each second on average during peak hours, and about 1 during non-peak hours. As the number of peak hours increases, the traffic density during peak hours decreases and that during non-peak hours increases.

For each car generated from the Poisson distributions, it will be a self-driving car with probability q and a human-driven one with probability $1 - q$. We can adjust q to vary the percentage of self-driving cars.

We set the initial speed v of a newly-arriving vehicle to be 20 m/s (45 mph), which is a relatively low value. Setting the initial speed low does not hurt the following simulation, because if the traffic density is low, a higher speed is allowed, and the vehicle will speed up soon according to our vehicle-following model. The initial value of acceleration a is calculated according to the state of nearby vehicles, using the rules introduced in the vehicle-following model.

Vehicle-Following Model

The vehicle-following model defines the behavior of two cars in the same lane, one directly behind the other. We call the car in front the *preceding car*, and the other car the *follower car*.

The vehicle-following model defines how the follower car changes its speed and acceleration according to the status of the preceding car. We consider three different cases:

1. The follower car is human-driven.
2. The follower car is self-driving and the preceding car is human driven.
3. Both cars are self-driving.

In each case, we define a rule to update the speed and acceleration of the follower car.

Follower Car is Human-Driven

When the follower car is human-driven, the driver can only try to keep a safe distance from the preceding car; a human driver needs longer response time than self-driving cars to react when the preceding car decelerates suddenly.

The safe distance of a human-driven follower car is

$$G_S = T_{\text{rh}} v,$$

where T_{rh} is response time of a human driver. According to the Two Second Rule [11], $T_{\text{rh}} = 2s$ is an appropriate setting, which we use in our simulation.

Human drivers tend to be careful, due to potential human error, when changing speed. Thus, their acceleration is a fixed value: In every second, they can either increase or decrease the speed by one unit, which is also a convention in previous cellular automata-based traffic flow analysis.

Thus, the behavior of a human-driven car can be divided into two cases:

- If $G \leq G_S$, the driver decelerates.
- If $G > G_S$, the driver can accelerate with probability P_a and decelerate with probability P_d, and keep the speed unchanged otherwise. Here $P_a + P_d \leq 1$. We specify values for P_a and P_d later.

When decelerating, the speed decreases by 1 per unit time, i.e., $v_{\text{new}} = \max(V_{\min}, v-1)$. When accelerating, the speed increases by 1 per unit time and is bounded by the maximum speed for the highway, i.e., $v_{\text{new}} = \min(v+1, V_{\max})$.

Finally let's specify values of P_a and P_d. When a car is running very fast, the human driver is more likely to tend to slow down. Thus, as the current speed v increases, P_a decreases and P_d increases. Each cell is 4 m long and the maximum speed allowed is 60 mph (about 7 cells/s). **Table 2** shows values that we set for P_a and P_d for different values of v.

Table 2.

Acceleration and deceleration probabilities when $G > G_S$.

v	1	2	3	4	5	6	7
P_a	1	0.9	0.8	0.7	0.5	0.2	0
P_d	0	0.1	0.2	0.2	0.3	0.4	0.6
no change	0	0	0	0.1	0.2	0.4	0.4

Self-Driving Vehicle Following a Human-Driven Vehicle

Even if a self-driving vehicle cannot communicate with the preceding human-driver, its sensors can still detect the speed of the human-driven car and react immediately. Thus, the response time is much shorter than for a human driver. The safe distance in this case is measured by $G_S = T_{\text{rs}} v$, where T_{rs} is the response time for a self-driving vehicle; but $T_{\text{rs}} < T_{\text{rh}}$. We use $T_{\text{rs}} = 1$ s in our simulation.

Unlike human drivers, self-driving cars have greater flexibility in the control of acceleration; they can adjust their acceleration value constantly. In every second, we have the updating rules for its speed and acceleration:

$$
\begin{aligned}
v_{\text{new}} &= \begin{cases} \min(V_{\max}, v+a), & a \geq 0; \\ \max(V_{\min}, v+a), & a < 0, \end{cases} \\
a_{\text{new}} &= \alpha(G - G_S) + \beta(v_p - v).
\end{aligned}
$$

The acceleration of a self-driving car depends on the current distance G, safe distance G_S, current speed v, and speed v_p of the preceding vehicle. The coefficients α and β determine how much the acceleration depends on $(G - G_S)$ and the speed difference $(v - v_p)$. When $G > G_S$, the distance is too large, thus acceleration is favorable. Conversely, when $G < G_S$, the car should slow down. The second term $\beta(v_p - v)$ favors acceleration when the self-driving car is slower than preceding human-driven car, and deceleration when it is faster.

Following Model for Two Self-Driving Vehicles

When both vehicles are self-driving, they can communicate speed and acceleration constantly. They still need a safe distance $G_S = T_{\text{rs}} v$. Since two cooperating cars can communicate, the follower car can get accurate real-time acceleration of the preceding car a_p. Taking a_p into account when updating the acceleration provides further flexibility for a self-driving car.

The updating rules [3] for the follower car are

$$v_{\text{new}} = \begin{cases} \min(V_{\max}, v + a), & a \geq 0; \\ \max(V_{\min}, v + a), & a < 0, \end{cases}$$
$$a_{\text{new}} = \alpha(G - G_S) + \beta(v_p - v) + \gamma a_p.$$

Because of the term γa_p, when the preceding car wants to accelerate ($a_p > 0$) or decelerate ($a_p < 0$), the follower car will immediately take corresponding action. Thus, this rule makes use of the communication between two self-driving cooperating cars.

Lane-Changing Model

Besides adjusting its speed and acceleration, each car has a chance to move to the neighboring lane. A lane change can happen only when

- the current distance to the preceding car is less than the safe distance and
- the cell in the neighboring lane is empty and distances to the preceding and follower cars in the neighboring lane are safe.

Let G_{XY} be the actual distance between car X and car Y; and let $G_{S_{XY}}$ be the safe distance between car X and car Y, defined in the same way as in the vehicle-following model, according to the categories of the cars.

Formally, as shown in **Figure 2**, if $G_{AB} \leq G_{S_{AB}}$, then car A can change to lane 1 when $G_{AC} > G_{S_{AC}}$ and $G_{AD} > G_{S_{AD}}$.

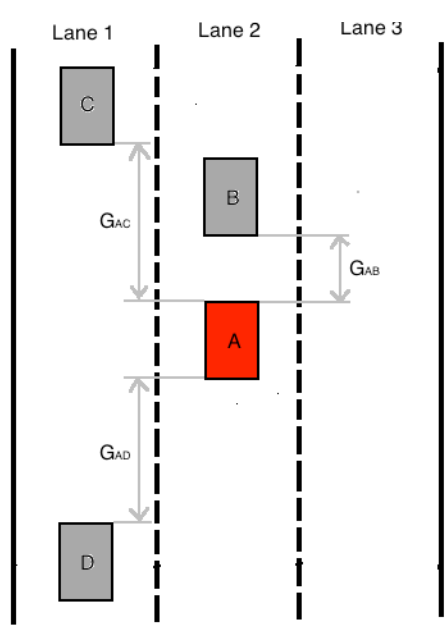


Figure 2. Conditions for lane-changing.

When only one neighboring lane satisfies the condition, the car changes to that lane with probability P_c. When both left and right lanes satisfies the condition, the car changes to either lane with probability $P_c/2$.

Lane-changing happens before the speed and acceleration update in the vehicle-following model. After lane-changing, the speed and acceleration of a vehicle are updated according to the rules in the vehicle-following model.

Intersection Model

Most highway intersections are based on interchanges, a structure that allows traffic to pass through the junction without directly crossing any traffic stream. Thus, in such an intersection, a car on one highway can move to another highway without waiting at a traffic light.

We use a cloverleaf interchange in our model, since it is the most widely-used one. **Figure 3** shows the traffic flow on a cloverleaf interchange. A car in Position A on Lane 2 can move to Position B on Lane 3 or Position C on Lane 4. Similarly, a car can enter Lane 2 from Position D on Lane 3 and Position E on Lane 4. The traffic flow on Lane 1 has a similar interaction with traffic flow on Lane 3 and Lane 4.

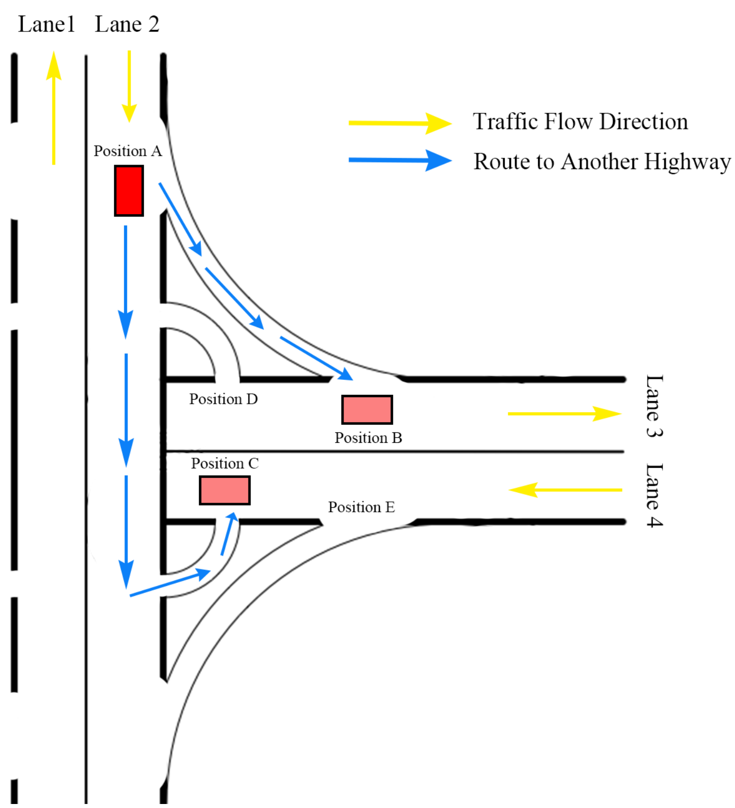


Figure 3. Cloverleaf interchange.

Because there's no crossing through a traffic stream, we can assume that the car disappears from one highway and appears on another one after a period of time.

With this assumption, we introduce eight parameters μ_{ij} and μ_{ji}, for $i \in \{1, 2\}, j \in \{3, 4\}$, to represent the probabilities of cars moving from lane i to lane j, and from lane j to lane i, respectively. We assume that each of these movements across the interchange takes time T_{inter}. We summarize the rules for intersection crossing as follows:

- When a car arrives at the intersection, with probability μ_{ij} it starts moving from lane i to lane j.
- If the lane-change starts and T_{inter} seconds pass after that, check if the destination position on lane j is empty. If so, put the car there with its old speed and update its acceleration with the rules of the vehicle-following model. Otherwise, the car waits until the position becomes empty.

The intersection model is simulated by running four cellular automata using vehicle-following rules and lane-changing rules, with the intersection rules defined above. Details are deferred to the simulation part.

Analyzing the Model by Simulation

We analyze how the highway capacity varies in our model when the percentage of self-driving vehicles, number of lanes, and traffic volume change. We run the simulation on a highway segment 250 cells long (1 km). Each simulation runs for a traffic flow of 24 hours, using the generation model to generate vehicles.

The simulation results show that self-driving vehicles increase highway capacity. We also compare the effects of self-driving vehicles with other factors.

Criteria

To evaluate the traffic capacity, we need some criteria. In our simulations, we mainly use average passing time and traffic density.

- **Whole-Day Average Passing Time** The average time for a vehicle to pass through the 250-cell highway segment.
- **Peak-Hour Average Passing Time** The average time for a vehicle to pass through the 250-cell highway segment during peak hours.
- **Traffic Density** Number of vehicles on the 250-cell highway segment. This value will change with time as the cars enter and leave the highway. So for each combination of model parameters, we can obtain a curve of traffic density varying with time. Since the traffic load is time-dependent, we can further use it to analyze the stability of a traffic flow.

Effects of Percentage of Self-Driving Vehicles

We run the cellular automata on a highway segment, varying the percentage of self-driving vehicles to observe the change in highway capacity.

The highway segment has 3 lanes in both directions. We fix the daily traffic volume at 100,000, with 50,000 in each direction. We set the lane-changing probability $P_c = 0.5$ and the number of peak hours in the generation model to be $p = 1$. We calculate the average passing (transit) time to pass through the highway segment and also the traffic load. **Figure 4** and **Figure 5** show the results.

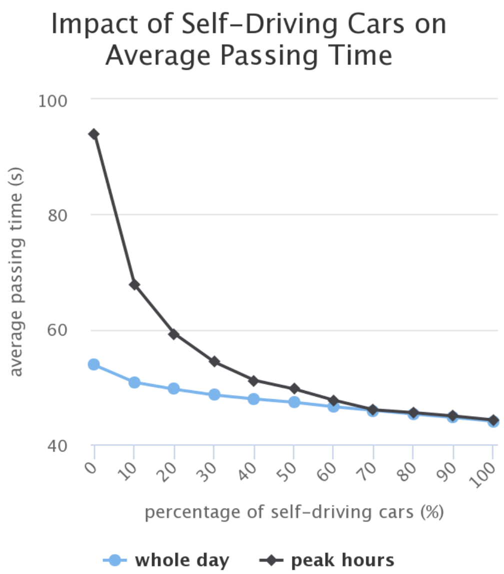


Figure 4. Average passing time vs. percentage of self-driving cars, over the whole day and during peak hours.

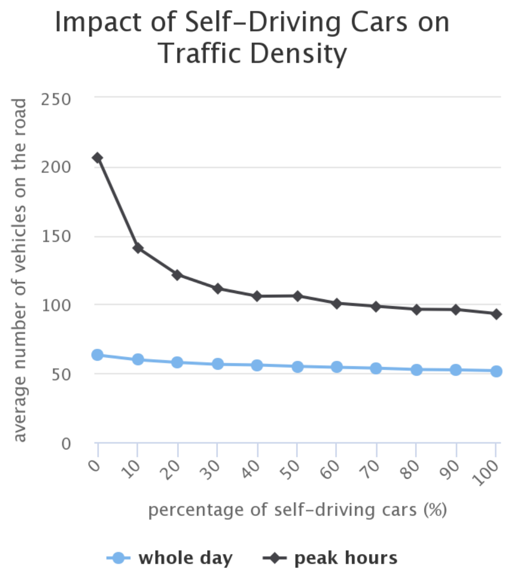


Figure 5. Traffic density vs. percentage of self-driving cars, over the whole day and during peak hours.

When the self-driving percentage increases:

- **Average passing time goes down.** As seen in **Figure 4**, the average passing time during peak hours decreases sharply at first, and reaches an equilibrium at around 70%. Self-driving vehicles can increase the traffic capacity by more-accurate control and cooperating. The benefit is tiny when the percentage of self-driving cars is low, since cooperation is difficult with few self-driving cars on the road.
- **Peak-hour pressure is alleviated.** Self-driving cars reduce the difference between peak-hour passing time and whole-day passing time. In **Figure 4**, as the percentage of driving cars increases, the two curves get closer. This feature reveals an important advantage of self-driving cars, that they are flexible enough to deal with peak hours.
- **The capacity of the highway increases.** As seen in **Figure 5**, the average number of vehicles on the 250-cell highway segment goes down. When the self-driving percentage is low, cars on the road move slowly, as indicated in **Figure 4**. Thus, traffic jams more easily arise; but with more self-driving cars, traffic jams can be reduced. Again, the impact of self-driving cars is more obvious during peak hours.

Effects of the Number of Lanes

We run simulations with 3, 4, and 5 lanes in both directions. The daily traffic volume is 100,000, with 50,000 in each direction, and the number of peak hours is $p = 1$. The lane-changing probability $P_c = 0.5$, and the self-driving percentage is 50%.

We also test the effects of the number of lanes with different self-driving percentage. We simulate 3, 4, and 5 lanes, with self-driving percentage 30% and 70%.

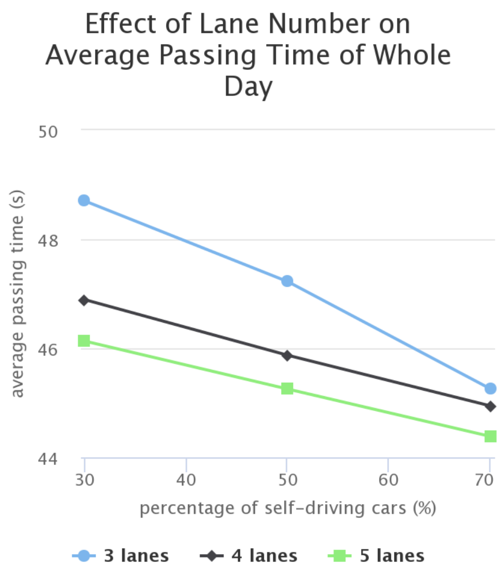


Figure 6. Impact of number of lanes on passing time, over the whole day.

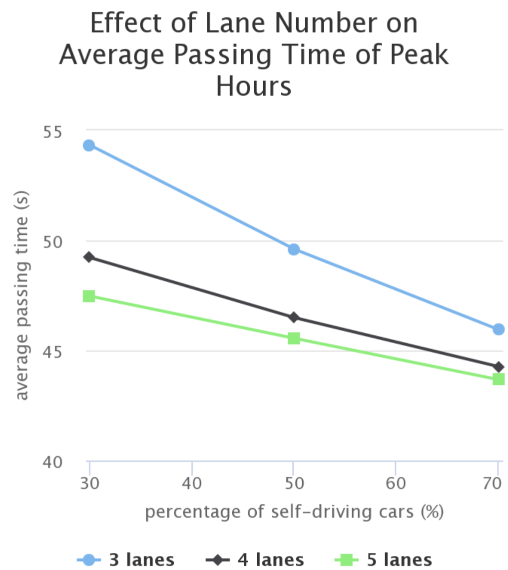


Figure 7. Impact of number of lanes on passing time, during peak hours.

We have two important observations based on **Figure 6** and **Figure 7**.

- **More lanes can speed up traffic flow:** Both over the whole day and during peak hours, adding lanes shortens the average passing time. This is consistent with our intuition, since with more lanes, cars have greater freedom to change lanes when the current lane is too crowded.
- **Self-driving reduces the impact of the number of lanes:** In both **Figure 6** and **Figure 7**, as the self-driving percentage increases, the three curves get closer to each other. When most cars are self-driving, adding more lanes makes little difference.

Effects of Dedicated Lanes for Self-Driving Cars

To test the effects of dedicated lanes, we modify the generation model and lane-changing model.

- **Generation Model** When a self-driving car is generated, we put it into one of the dedicated lanes.

- **Lane-Changing Model** A self-driving car cannot change to a non-dedicated lane.

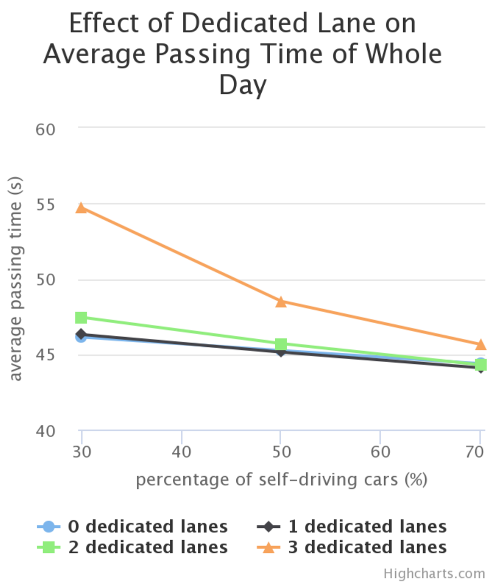


Figure 8. Impact of dedicated lanes on passing time, over the whole day.

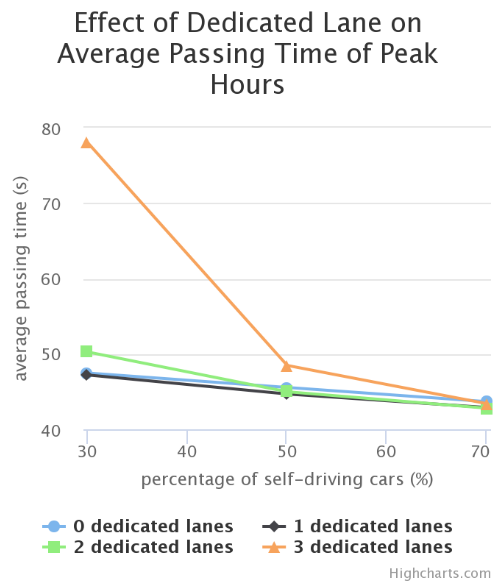


Figure 9. Impact of dedicated lanes on passing time, during peak hours.

With these modifications, we ran the simulation with 5 lanes in both directions. The daily traffic volume is 100,000, with 50,000 on each direction. The number of peak hours is $p = 1$. The lane-changing probability is $P_c = 0.5$. We vary the dedicated lanes from 1 to 3 and calculate the average passing time, for self-driving percentages of 30%, 50%, and 70%.

From **Figure 8** and **Figure 9**, when the percentage of self-driving vehicles is low, dedicated lanes may hurt traffic capacity, especially during peak hours. For example, in **Figure 9** when only 30% are self-driving cars, 3 dedicated lines results in an average passing time of 80 s (corresponding to 28 mph), which is much slower than in the other cases.

However, as the percentage of self-driving cars increases, dedicated lanes can improve performance a little. For example, in **Figure 9**, when 70% are self-driving, any number of dedicated lanes outperforms the no-dedicated-lanes strategy.

Effects of Traffic Volume

We test daily traffic volume from 100,000 to 500,000, distributed equally in both directions. Each direction has 5 lanes. The lane-changing probability is $P_c = 0.5$, and we fix the number of peak hours at $p = 1$. We do simulations for 30%, 50%, and 70% self-driving cars.

The impact of traffic volume on passing time declines as the percentage of self-driving cars increases. Thus, self-driving cars can maintain the highway capacity well even when traffic is quite heavy.

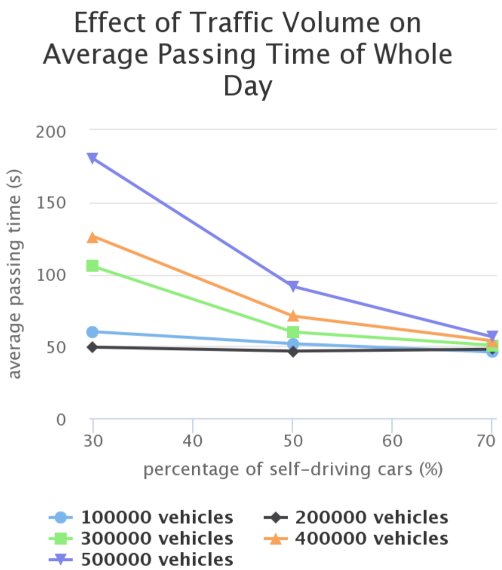


Figure 10. Impact of traffic volume on passing time, over the whole day.

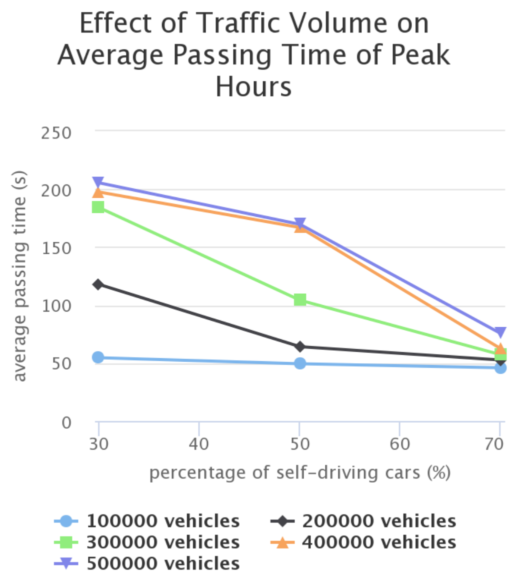


Figure 11. Impact of traffic volume on passing time, during peak hours.

Applying Our Model to Real Data

We apply our model to the provided data of Interstates 5, 90, and 405, as well as State Route 520. We test how the traffic capacity of these roads changes with the percentage of self-driving cars. We also analyze the feasibility of dedicated lanes, and give suggestions based on our analysis.

Percentage of Self-Driving Vehicles

We run simulations for each segment on the data sheet, and study the impact of self-driving vehicles on highway capacity during the whole day and during peak hours. Though self-driving vehicles do not increase the highway capacity a lot on average during the day, they have a significant impact during the peak hours.

Average Speed During a Whole Day

The data sheet has 224 highway segments, belonging to four highways, Interstates 5, 90, and 405 and State Route 520. For each segment and its corresponding daily traffic count, we test the average speed $\overline{v_i}$ of cars passing through the segment during the day. Then we calculate the average speed for each road, by averaging $\overline{v_i}$ of the segments belonging to the road. **Figure 12** shows the results.

On each highway, the average speed increases with the percentage of self-driving vehicles. As the percentage of self-driving cars increases from 10% to 50%, the average speed increases rapidly. However, when the percentage changing from 50% to 90%, the average speed increases only a little. An equilibrium is reached at around 80%.

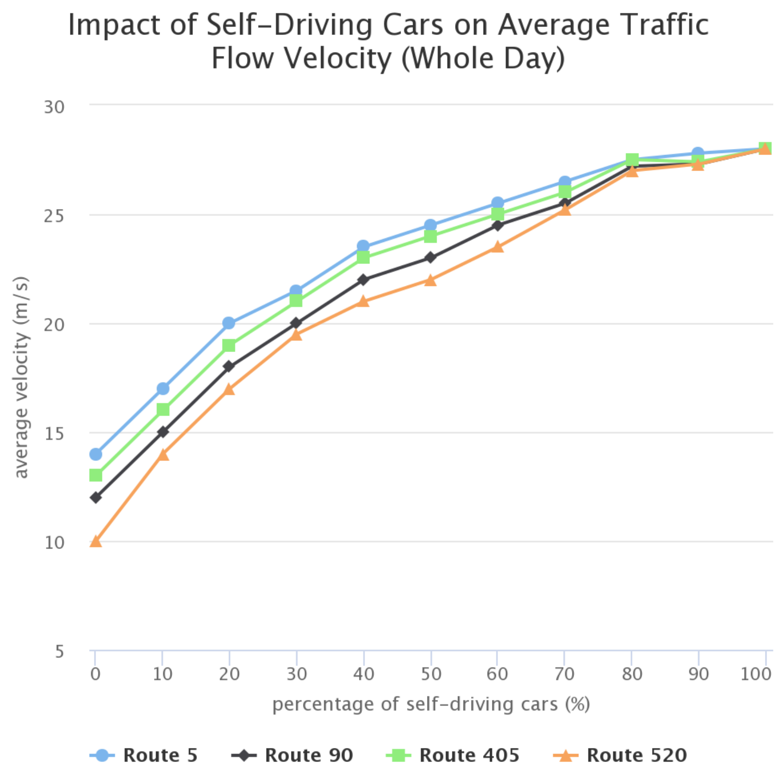


Figure 12. Average traffic speed over a whole day, for each of the four highways. The average speed varies between 10 m/s (22 mph) and 28 m/s (62 mph).

The average speed increases most markedly when the percentage of self-driving vehicles changes from 0% to 10%, and from 10% to 20%.

Average Speed During Peak Hours

Average speeds on each highway during peak hours are calculated in the same way and shown in **Figure 13**.

Compared with **Figure 12**, we find that self-driving has a more significant impact for peak-hour traffic flow. The average speeds on all four highways experience a huge increase as the self-driving percentage increases. For example, when the percentage increases from 10% to 50%, the average speed on Route 5 increases by about 10 m/s (22 mph), and by about 5 m/s (11 mph) when the percentage increases from 50% to 90%. An equilibrium is reached at around 70%.

Traffic Density

From **Figure 14** and **Figure 15**, we see that self-driving cars help reduce average traffic density. Again, the impact is more significant during the peak hours. In **Figure 15**, the traffic density decreases more rapidly when the percentage of self-driving is low. From 10% to 50%, the traffic density reduces by half. While from 50% to 90%, there's no significant change. An equilibrium is reached at around 50%.

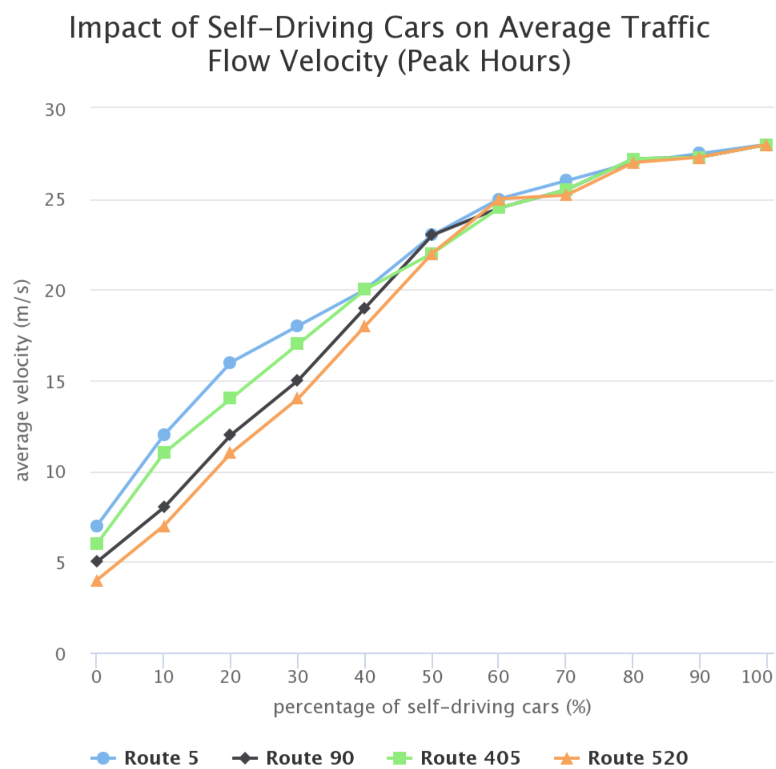


Figure 13. Average traffic speed during peak hours, for each of the four highways. The average speed varies between 5 m/s (11 mph) and 28 m/s (62 mph).

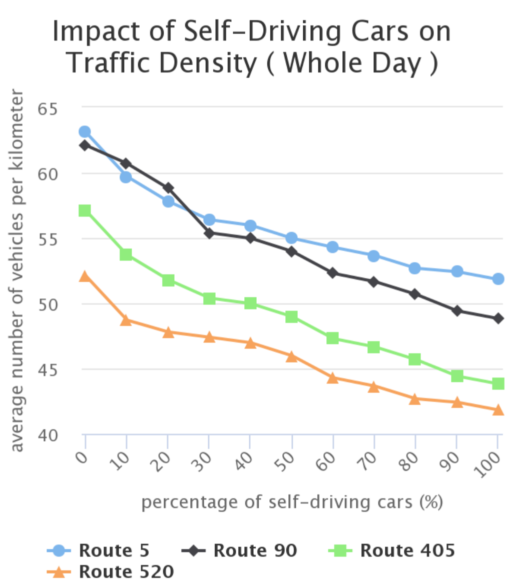


Figure 14. Impact of self-driving cars on traffic density over the whole day, for each highway.

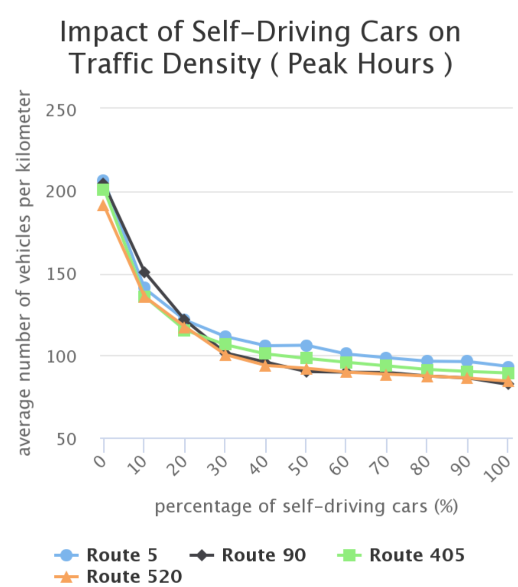


Figure 15. Impact of self-driving cars on traffic density during peak hours, for each highway.

Average Speed at Intersections

We apply the intersection model to simulate traffic flow from milepost 152.72 to milepost 155.18 of Interstate 5, where the highway intersects Interstate 405 (at its mileposts 0 to 0.09). We calculate the average speed of vehicles over the whole day. Each vehicle traveling through milepost 152.72 to 155.18 on Interstate 5, or through milepost 0 to 0.09 on Interstate 405, may change highways, with all the probabilities in the intersection model being 50%. We set T_{inter}, the time to cross interchange, as 10 s. We get similar results as for our straight-line highway results.

Dedicated Lanes

We pick the data for Interstate 5 to analyze effects of dedicated lanes. **Figure 16** and **Figure 17** show the average speed of Interstate 5 with different number of dedicated lanes for self-driving cars.

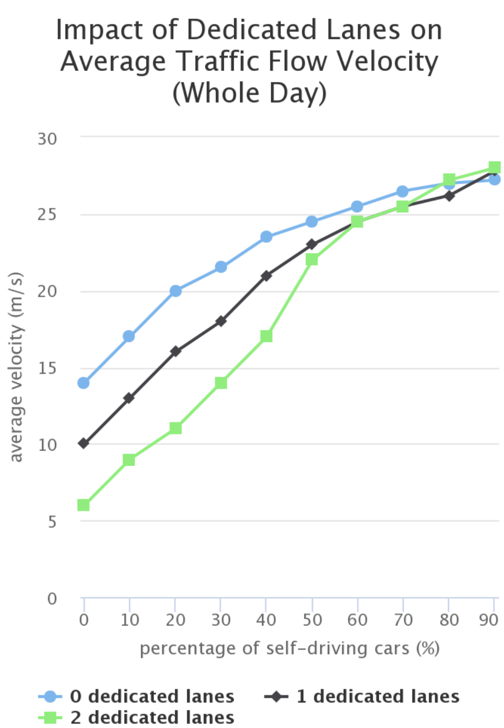


Figure 16. Impact of number of dedicated lanes on average speed over the whole day.

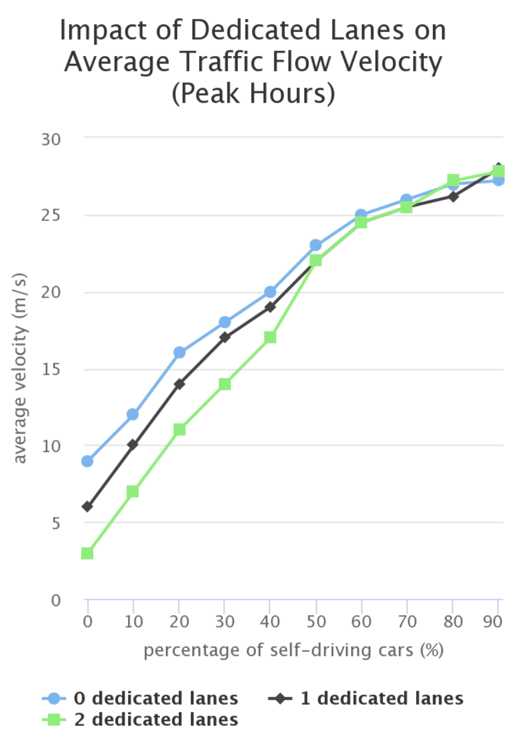


Figure 17. Impact of number of dedicated lanes on average speed during peak hours.

With a small proportion of self-driving cars, dedicated lanes do not help much to increase average speed. However, when most vehicles are self-driving, for example 70%, dedicated lanes become necessary.

Sensitivity Analysis

We analyze the sensitivity of some parameters in our model. We run simulations on a 250-cell highway segment with 5 lanes (but no dedicated lane). The daily traffic count is 100,000. With this setting, we analyze how the parameters in our model influence our conclusions. The results show that our model is quite robust.

Lane-Changing Probability P_c

In our previous experiments, the value of P_c was 0.5. So when the distance G between two cars is less than the safe distance G_S, the follower car will change to another lane with probability 50%. Here we test the impact of self-driving cars on highway capacity with different P_c's.

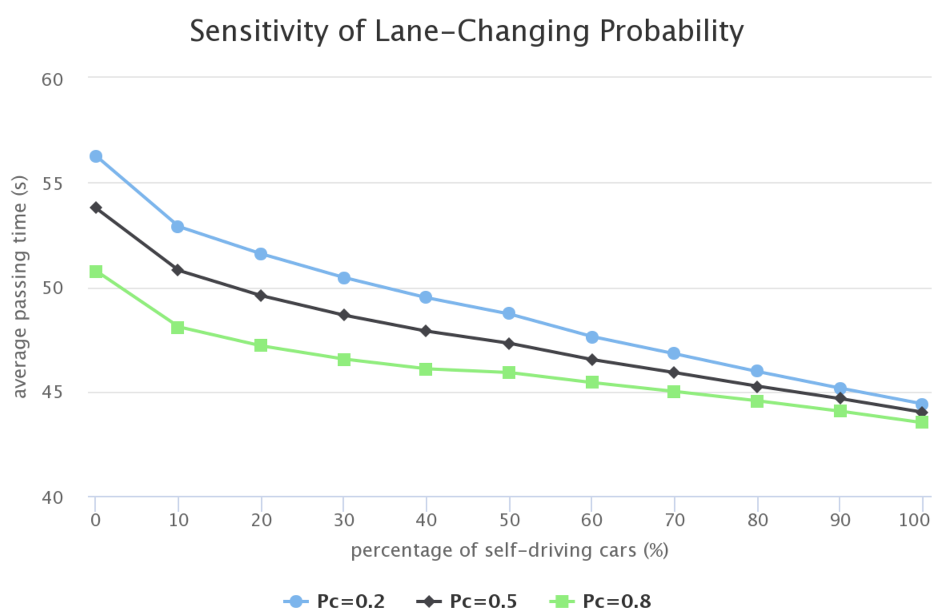


Figure 18. Sensitivity of passing time to lane-changing probability P_c.

When $P_c = 0.2$, cars tend to stay in the same lane even if it is crowded. Thus, the average passing time for a car is longer. When $P_c = 0.8$, cars change lanes frequently, so the average passing time is shorter. However, the differences from $P_c = 0.5$ are small.

Coefficients α,β, and γ in Vehicle-Following Model

Recall the updating rule for acceleration of self-driving vehicles:

$$a_{\text{new}} = \alpha(G - G_S) + \beta(v_p - v) + \gamma\alpha_p,$$

where

- α controls how much the acceleration depends on distance,
- β controls the impact of relative speed, and
- γ controls the impact of acceleration of the preceding car.

Our settings were $\alpha = \beta = \gamma = 1$.

When α is larger, the follower car tends to keep a safe distance. When β is larger, the follower car tries to catch up to the speed of the preceding car. And when γ is large, the follower car will accelerate synchronously with the preceding car.

We test different combinations of α, β and γ. The results are shown in **Figure 19**. We see only a tiny speed up of traffic flow when the values are larger.

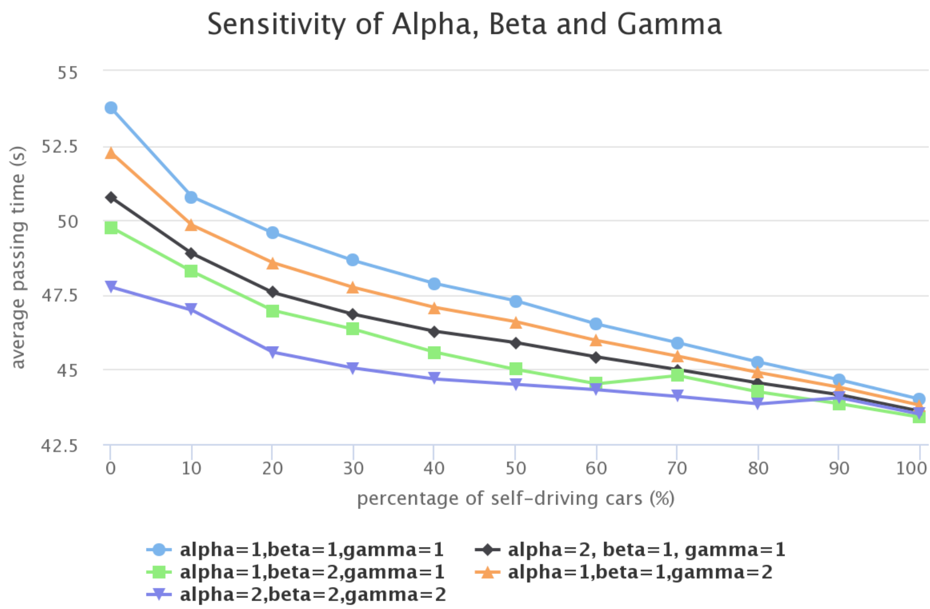


Figure 19. Sensitivity of passing time to combinations of values of α, β, and γ.

Error of Self-Driving Systems

In our model, we assume that self-driving cars are perfect; they don't introduce any new risk. However, as a newly-emerging technology, self-driving today is not mature enough. On June 30, 2016, a fatal crash of a self-driving Tesla occurred. Thus, it is necessary to test how system error will affect the performance of self-driving cars.

We introduce self-driving system error by making self-driving cars accelerate or decelerate by 1 unit, each with a probability of 10%. The result is shown in **Figure 20**. As we can see, such inaccurate acceleration and deceleration degrades performance, resulting in a longer passing time compared with perfect self-driving systems.

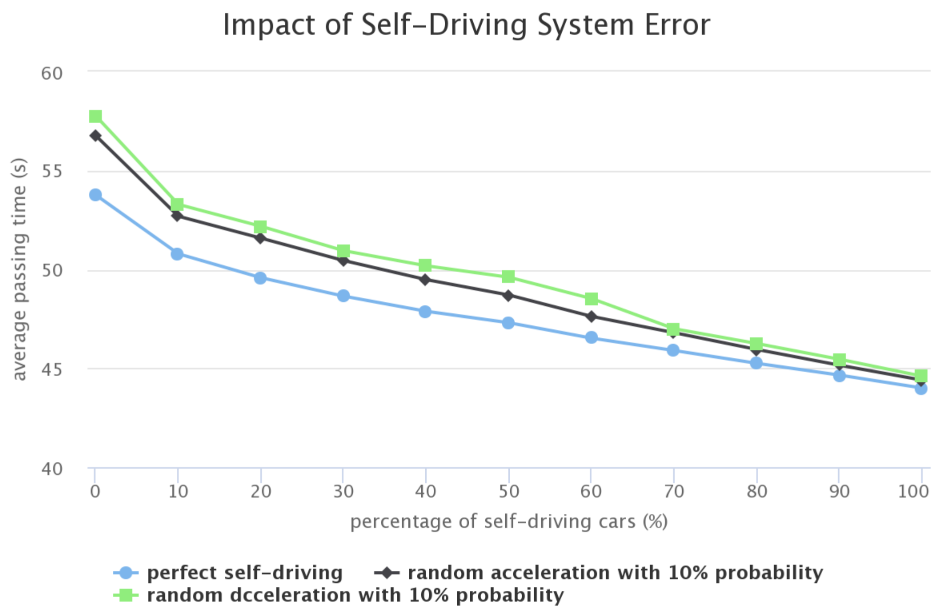


Figure 20. Impact of self-driving system errors on passing time, compared to perfect self-driving.

Conclusion

We build a model to analyze the impact of self-driving cars on highway capacity, perform many simulations, and apply our model to the data of four real highways. Here we note our interesting findings.

- Self-driving vehicles can increase highway capacity. As the percentage of self-driving cars increases, the traffic proceeds at higher speed and lower density.
- Dedicated lanes for self-driving cars are not necessary when the percentage of self-driving cars is low, while dedicated lanes can increase the capacity by a little when the percentage is higher than 70%.
- Self-driving vehicles can reduce the impact of the number of lanes and traffic volume. When most cars are self-driving, the speed of traffic does not change a lot with these factors.

References

[1] Nagel, Kai, and Michael Schreckenberg. 1992. A cellular automaton model for freeway traffic. *Journal de Physique I* 12: 2221–2229. `https://hal.archives-ouvertes.fr/file/index/docid/246697/filename/ajp-jp1v2p2221.pdf` .

[2] Ren, D., Y. Du, Y. Wang, and L. Huang. 1997. Analysis of traffic flow

based on cellular automata (in Chinese). http://d.wanfangdata.com.cn/Periodical/shgcjsdxxb201402008.

[3] van Arem, Bart, Cornelie van Driel, and Ruben Visser. 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on Intelligent Transportation Systems* 7 (4): 429–436. https://research.utwente.nl/en/publications/c1124b65-89dd-418d-b0b9-c60d9cb3699d.

[4] Baskar, L.D., B. De Schutter, and H. Hellendoorn. 2008. Model-based predictive traffic control for intelligent vehicles: Dynamic speed limits and dynamic lane allocation. In *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium (IV'08)*, 174–179. https://pdfs.semanticscholar.org/c288/87da850d3ca83c5db7c8c7bcfad293f70b2a.pdf.

[5] Pinjari, Abdul Rawoof, Bertho Augustin, and Nikhil Menon. 2013. Highway capacity impacts of autonomous vehicles: An assessment. https://pdfs.semanticscholar.org/713c/da5859535ca41986413e8c77c105f4898f2a.pdf.

[6] Benjaafar, Saifallah, Kevin Dooley, and Wibowo Setyawan. 1997. Cellular automata for traffic flow modeling. http://hdl.handle.net/11299/155096.

[7] Tientrakool, Patcharinee, Ya-Chi Ho, and Nicholar F. Maxemchuk. 2011. Highway capacity benefits from using vehicle-to-vehicle communication and sensors for collision avoidance. In *Vehicular Technology Conference (VTC Fall), 2011 IEEE*, 1–5. http://safevehicle.ee.columbia.edu/docs/Tientrakool11.pdf.

[8] Zhao, Li, and Jian Sun. 2013. Simulation framework for vehicle platooning and car-following behaviors under connected-vehicle environment. *Procedia Social and Behavioral Sciences* 96: 914–924. http://ac.els-cdn.com/S1877042813022313/1-s2.0-S1877042813022313-main.pdf?_tid=f6c3a3c2-7954-11e7-b5a1-00000aab0f6b&acdnat=1501879328_1fe5f7d15ef952a228f6194a0b259f4f.

[9] Yin, Yafeng, Mark A. Miller, Steven E. Shladover, F. Gerald Rawling, and Claire E. Bozic. 2004. Assessment of the applicability of cooperative vehicle-highway automation systems to freight movement in Chicago. In *Transportation Research Board 2004 Annual Meeting*. https://pdfs.semanticscholar.org/288b/d581e1c84868d0794d404e92856b07872f18.pdf.

[10] Gerlough, Daniel L. 1955. *Use of Poisson Distribution in Highway Traffic.* Saugatuck, CT: Eno Foundation for Highway Traffic Control. https://ntl.bts.gov/lib/26000/26800/26814/use_of_poisson_distribution_in_highway_traffic.pdf.

[11] Seymour, William K., and Ernest J. Bertothy. 2016. *Driver's Manual.* Connecticut Department of Motor Vehicles. `http://www.ct.gov/dmv/lib/dmv/20/29/r12eng.pdf`.

Dear Officials:

It is our honor to help you with analyzing the effects of allowing self-driving cooperating vehicles on the capacity of highways in the Greater Seattle area. We are writing this letter to report our findings.

We have built a novel cellular automata model to analyze the effects of allowing self-driving cooperating cars on Interstates 5, 90, and 405, as well as State Route 520 in Thurston, Pierce, King and Snohomish counties. By running simulations with varied percentages of self-driving vehicles, different numbers of lanes, and varied daily traffic volumes, we find that that self-driving cooperating cars can increase highway capacity. In other words, they can relieve the stress of traffic jams.

The capacity of the highways increases markedly when there are 10% self-driving vehicles, and reaches an equilibria at around 70%. Specifically, the average traffic speed flow increases, and traffic density decreases, with an increasing percentage of self-driving vehicles.

These vehicles are equipped with accurate sensors and intelligent control systems. So they require a shorter safe distance from other vehicles and have higher flexibility in acceleration and deceleration. These properties enable them to make full use of the space on highways and accelerate overall traffic flow.

What's more, these vehicles can communicate with each other and share data, including speed and acceleration. Such communication makes synchronous actions between two vehicles possible and eliminates potential collisions even when two cars are very close to each other. Such vehicles measure the speed of nearby vehicles and adjust their own speeds accordingly.

Based on our analysis, we provide some suggestions to better manage and exploit self-driving vehicles:

- **Popularize self-driving vehicles, even if only a little bit.** Our analysis shows most benefit to traffic capacity happens when the percentage of self-driving increases from 0% to 50%. Since self-driving vehicles are not very common yet, a huge improvement can be achieved by even a small increase in the number of these.

- **Do not hurry to build dedicated lanes for self-driving cars.** Our simulations suggest that dedicated lanes can be helpful only when the percentage of self-driving cars is high enough, for example 70%. Allocating dedicated lanes can increase overhead and be expensive. Since few self-driving vehicles are on the road today, such dedicated lanes may even hurt the traffic system by making other lanes for human-driven cars too crowded.

We believe that our model is useful for management of self-driving cars in the future. You are welcome to contact us at any time for further cooperation.

Yours sincerely,

Team # 55278

Team members Jiefeng Chen, Qi Li, and Yu Shi.

Judges' Commentary: The Self-Driving Cars Problem

Carol Overdeep
Mathematics Dept.
Saint Martin's University
Lacey, WA

Katie Oliveras
Mathematics Dept.
Seattle University
Seattle, WA

David H. Olwell
Dean and Professor
Hal and Inge Marcus School of Engineering
Saint Martin's University
Lacey, WA
dolwell@stmartin.edu

Introduction

2017 marked the second year for an MCM Problem C—a problem designed to elicit data insights in the context of a large, potentially messy data set. Problem C is not intended to be a "big data" problem requiring specialized computer science-based tools and techniques; but rather an opportunity to encounter real-world, challenging data sets that have interesting characteristics.

The 2017 problem provided participants the opportunity to analyze the effects of allowing self-driving cars on various high-volume roads in the Seattle-Tacoma metropolitan area. Specifically, participants were asked to determine if there were any *equilibria* or *tipping points* as the percentage of self-driving cars was varied—terms that the participants had to define as part of the modeling process. The associated data set provided participants with average daily traffic at various mileposts along the roads of interest.

The problem-writing team continues to refine its expectations for Problem C. This year's problem and associated data set shifted the focus slightly from statistical techniques to model-based data insights.

The UMAP Journal 38 (3) (2017) 347–353.

The Problem: "Cooperate and Navigate"

We repeat the problem verbatim:

Traffic capacity is limited in many regions of the United States due to the number of lanes of roads. For example, in the Greater Seattle area drivers experience long delays during peak traffic hours because the volume of traffic exceeds the designed capacity of the road networks. This is particularly pronounced on Interstates 5, 90, and 405, as well as State Route 520, the roads of particular interest for this problem.

Self-driving, cooperating cars have been proposed as a solution to increase capacity of highways without increasing number of lanes or roads. The behavior of these cars interacting with the existing traffic flow and each other is not well understood at this point. The Governor of the State of Washington has asked for analysis of the effects of allowing self-driving, cooperating cars on the roads listed above in Thurston, Pierce, King, and Snohomish counties (see the provided map and Excel spreadsheet). In particular, how do the effects change as the percentage of self-driving cars increases from 10% to 50% to 90%? Do equilibria exist? Is there a tipping point where performance changes markedly? Under what conditions, if any, should lanes be dedicated to these cars? Does your analysis of your model suggest any other policy changes?

Your answer should include a model of the effects on traffic flow of the number of lanes, peak and/or average traffic volume, and percentage of vehicles using self-driving, cooperating systems. Your model should address cooperation between self-driving cars as well as the interaction between self-driving and non-self-driving vehicles. Your model should then be applied to the data for the roads of interest, provided in the attached Excel spreadsheet.

Your MCM submission should consist of a 1-page Summary Sheet, a 1–2 page letter to the Governor's office, and your solution (not to exceed 20 pages) for a maximum of 23 pages. Note: The appendix and references do not count toward the 23-page limit.

Summary

Once again, this year's problem provided a challenge for the teams. The teams had no difficulty identifying existing models on traffic flow (e.g., cellular automata, microscopic/macroscopic models, etc.). Incorporating the given data into the models, however, is where many teams struggled. A surprising number of teams either did not mention the given data set in their papers or summarized the given data set but did not apply their models/results to the data. Regardless of how well these papers were written, not including the given data in a meaningful way prevented them from further consideration beyond the triage phase of judging. The most successful teams provided detailed descriptions of both their model assumptions and how their model was constructed. They used the data to gain insight into regional traffic patterns and then applied their traffic model to the traffic data. Once the traffic data were applied to their model, they carefully investigated the impact of both self-driving cars along with dedicated lanes to the overall traffic flow throughout the entire system.

The top papers discussed both equilibrium and tipping points within this context, and used the ensuing results to provide policy recommendations in a clear, concise manner.

As always, judges value well-written, well-illustrated papers that carefully followed the contest directions. Judges valued papers that described and motivated the equations used, especially those adapted from other sources. Judges paid close attention to citations, both proper use of inline citations and using reputable sources. Repurposing images from external sources without credit detracted from a paper's overall rating.

There were four Outstanding papers. One was from the University of California–Berkeley (advised by John Harte) and the other three were from Shanghai Jiao Tong University (all advised by Xiaofeng Gao). Shanghai Jiao Tong University had an exceptional performance on Problem C, with 3 Outstanding papers, 8 Meritorious papers, 22 Honorable Mention papers, and 10 Successful Participants.

The Judging Process

The judging process followed the usual scheme of triage, screening rounds, and final judging as described in previous Judges' Commentaries [Black 2009; Black 2011; Black 2013]. In 2016, the judging team met in early March near the end of the triage phase; this year, the judging team (a mix of judges from the previous year and new judges) met in mid-February at the start of the triage phase. Prior to the meeting, the judging team had been asked to read and evaluate five papers. This facilitated discussion about scoring criteria used in the evaluation process, leading to increased consistency among judges, as well as providing hands-on training for the new judges.

Defining the Problem

Teams were asked to model the effects of introducing self-driving cars on traffic flow during peak and/or average volumes. In particular, teams were asked to analyze the impacts of self-driving cars as the percentage of these vehicles increased from 10% to 50% to 90%.

The teams had to determine how to measure the impacts of self-driving cars; unlike last year's problem, this was not a critical element in the development of the model but rather a byproduct of the type of model chosen. The best papers clearly outlined how they would measure the impacts to the traffic systems and what both an equilibrium and a tipping point would represent with regards to their model.

The problem also required that teams address the "cooperation between self-driving cars as well as the interaction between self-driving and non-self-driving vehicles." While many papers described the interaction between self-driving and non-self-driving vehicles, very few papers described or modeled cooperation between self-driving cars; those that did were rated higher.

The best papers also included a literature review that discussed both self-driving cars and various co-operation models. The best teams described how their models compared to the existing literature, and where they built upon it; the judges found this approach helpful.

Insights from the Data

The data set came from the Washington State Dept. of Transportation and contained seven fields:

- route ID,
- start milepost,
- end milepost,
- average daily traffic counts for 2015,
- number of lanes in the milepost-decreasing direction,
- number of lanes in the milepost-increasing direction,
- the route type (state route or interstate), and
- comments.

One of the challenging aspects of the data is that it contained average daily counts, not hourly counts. The problem asked teams to address the effects on traffic flow during peak and/or average traffic volume. In the problem description, teams were told that on average, 8% of the daily traffic volume occurs during peak travel hours. Teams had to identify peak travel hours, and most teams chose two 2-hour windows (one for the morning commute, one for the evening commute) while others chose two 3-hour windows. The judges gave no preference to either selection. The best papers used the data to gain insight into the overall traffic flow and potential locations for bottlenecks in the overall traffic system. They then applied their traffic models to the data set and analyzed the resulting changes throughout the entire region. They then compared their metrics against the given data to provide evidence for their proposed policy suggestions.

As noted earlier, papers that did not use the data set or did not clearly communicate how they used the data set were not considered after the triage phase.

Building the Models

There is a wide variety of existing traffic models. Judges had no preferred model; rather, they looked for

- clarity in the description of the model(s),
- a complete list of assumptions used in the model, and
- how the given data set was incorporated into the model.

The key assumptions in the models included

- the reaction time of a self-driving vehicles,
- the reaction time of human-driven vehicles,
- the following distance of self-driving cars, and
- the following distance of human-driven vehicles.

Of paramount importance was the use of the given data set! A somewhat surprising percentage of papers were either "Here are several traffic flow models and here are our results" or "Here are several traffic flow models, here is a summary of the data, and here are our results" without connecting the data to the model and/or results.

Some of the models that fell into one of these two categories had another shortcoming: They included modeling of traffic lights. Traditional traffic lights are not part of the interstate system.

Similarly, many papers built on existing traffic flow models without clearly providing either motivation or the associated assumptions corresponding to these base models. Papers that simply presented equations or models without motivation or justification were not evaluated highly by the judges.

The highest-ranked papers included valid mathematics with sufficient details justifying the model's choice. They also defined interactions between vehicles and applied their model to the data. Both continuous and discrete models were successfully applied to the problem by the Outstanding papers. The paper by the University of California–Berkeley team included both a macro model based on a PDE flux-flow-density model and a discrete model to analyze lane-changing behavior. The three papers from Shanghai Jiao Tong University used a cellular automata approach, each with a different emphasis.

Model Analysis and Validation

By design, there was no one "correct solution" to the problem, and this feature was borne out in the papers. Policy recommendations ran the gamut from not supporting self-driving cars up to advocating that 90% of the cars be self-driven. Consequently, judges did not read the papers in search of a "correct" answer; instead, judges looked for keen insights into the results generated by the model. Most teams justified their models by citing existing literature. Papers that also contained sensitivity analysis were viewed more favorably by the judges. Papers whose sensitivity analysis focused on the assumed values of the parameters (e.g., reaction times and following distances) and demonstrated an understanding of the purpose of sensitivity analysis were viewed even more favorably.

Communicating Results

Teams produced two documents: the paper and a letter to the Governor of Washington outlining policy recommendations with respect to self-driving vehicles. For the paper itself, highly-ranked papers were well organized, contained proper in-line citations, and included all the required elements. For the letter, again, organization and the inclusion of all the required elements were hallmarks of highly-ranked papers. In addition, letters that were written in "plain English" for a non-technical audience were valued.

Conclusions

The first Problem C, in 2016, attracted 1,875 student teams, with 4 recognized as Outstanding and 8 recognized as Finalists. This year's Problem C had 1,527 submissions, with 4 recognized as Outstanding and another 7 as Finalists. The judges continue to be delighted with the quality of submissions to the contest.

References

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

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

________. 2013. Judges' Commentary: The Ultimate Brownie Pan papers. *The UMAP Journal* 34 (2–3): 141–149.

Olwell, David H. 2013. Judge's Commentary: The Outstanding National Water Strategy papers. *The UMAP Journal* 34 (2–3): 189–195.

________, Carol Overdeep, and Katie Oliveras. 2016. Judges' Commentary: The Goodgrant Challenge papers. *The UMAP Journal* (3): 325–331.

About the Authors

Carol Overdeep is an Associate Professor of Mathematics at Saint Martin's University. She earned a B.S. in Mathematics at the University of Puget Sound, then began her professional career in the ADCAP MK-48 torpedo software validation group at Hughes Aircraft Company Ground Systems Group in Buena Park, CA. While working for Hughes, Dr. Overdeep earned an M.A. in applied mathematics from California State University–Fullerton. After 15 years as a torpedo systems engineer for Hughes Aircraft Company and the Naval Undersea Warfare Center Division Newport, Rhode Island, Dr. Overdeep earned an M.B.A. and then a Ph.D. in Applied Mathematics at the University of Rhode Island.

Katie Oliveras is an Assistant Professor of Mathematics at Seattle University. She earned a B.S. in Systems Science and Engineering from Washington University in St. Louis, and then M.S. and Ph.D. degrees in Applied Mathematics from the University of Washington. As an undergraduate, she participated in the MCM competition; and she has served as a team advisor for MCM teams at Seattle University.

David H. Olwell is Professor and Dean, Hal and Inge Marcus School of Engineering, at Saint Martin's University. He earned a B.S. at the United States Military Academy, where he studied mathematics, and an M.S. in Mathematics, M.S. in Statistics, and Ph.D. in Statistics from the University of Minnesota. He has previously been on the faculty of the U.S. Military Academy and the Naval Postgraduate School. Dr. Olwell has been a problem author for both the MCM and HiMCM, and has been an MCM judge for a decade.