HIGH SCHOOL MATHEMATICAL CONTEST IN MODELING

OUTSTANDING PAPERS

HiMCM

November 2011

The contest offers students the opportunity to compete in a team setting using applied mathematics in the solving of real-world problems.

Additional support provided by the National Council of Teachers of Mathematics (NCTM),
the Mathematical Association of America (MAA),
and the Institute for Operations Research and Management Sciences (INFORMS).

Editor's Comments

This is our fourteenth HiMCM special issue. Since space does not permit printing all eight National Outstanding papers, this special section includes abridged versions of two papers and summaries from the other six. We emphasize that the selection of these two does not imply that they are superior to the other outstanding papers. We also wish to emphasize that the papers were not written with publication in mind. Given the 36 hours that teams have to work on the problems and prepare their papers, it is remarkable how much they accomplished and how well written many of the papers are. The unabridged papers from all National and Regional Outstanding teams are on the 2011 HiMCM CD-ROM, which is available from COMAP ❑

Contest Director's Article

William P. FoxDepartment of Defense Analysis

Naval Postgraduate School
Monterey, CA 93943
wpfox@nps.edu

It is hard to believe that the High School Mathematical Contest in Modeling (HiMCM) completed its fourteenth year. It is and continues to be a fantastic endeavor. The mathematical and modeling ability of students, and faculty advisors, is very evident in the professional submissions and work being done. The contest is still moving ahead, growing with a positive first derivative, and consistent with our positive experiences from previous HiMCM contests. We hope that this contest growth continues. The plot of the growth over time is shown in Figure 1. The current trend is an exponential increase.

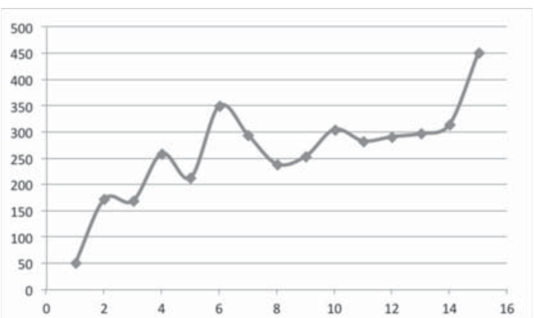


Figure 1: Number of HiMCM teams vs. contest year

This year the contest had 450 teams consisting of 1721 students from 24 states and 4 foreign countries. We had 266 U.S. teams and 169 foreign teams, representing 35% and 58% growth, respectively. In the United States these teams represented 49 schools. China represented about 64% of the foreign entries. Of the 1721 students, almost 35% were female students. The breakdown was 608 female, 1023 male students, and 90 unspecified genders. Since the beginning we have had 12,731 total participants, of which 36% have been female. This year, we again charged a registration fee of $75 per team.

The teams accomplished the vision of our founders by providing *unique* and *creative mathematical* solutions to complex open-ended real-world problems. This year the students had a choice of two problems both of which represent real-world issues.

Commendation: All students and advisors are congratulated for their varied and creative mathematical efforts. Of the 450 registered teams, 435 submitted solutions. These were broken down as 222 doing problem A and 213 doing Problem B. The thirty-six continuous hours to work on the problem provided for quality papers; teams are commended for the overall quality of their work.

Many teams had female members. There were 608 female participants on the 435 teams. There were 1721 total participants, so females made up over 35.3% of the total participation, showing that this competition is for both genders. This percent is almost triple the percent of woman in many other math competitions.

Teams again proved to the judges that they had "fun" with their chosen problems, demonstrating research initiative and creativity in their solutions. This year’s effort was a success!

Judging: We ran three regional sites in December 2011. The regional sites were:
Naval Postgraduate School in Monterey, CA
Francis Marion University in Florence, SC
Carroll College in Helena, MN.

Each site judged papers for problems A and B. The papers judged at each regional site may or may not have been from their respective region. Papers were judged as Regional Outstanding, Meritorious, Honorable Mention, and Successful Participant. All finalist papers from the Regional competition including all Outstanding awards were sent to the National Judging in Boston. For example, eight papers may be discussed at a Regional Final and only four selected as Regional Outstanding but all eight papers are judged for the National Outstanding. Papers receive the higher of the two awards. The national judging chooses the "*best of the best*" as National Outstanding. The National Judges commended the regional judges for their efforts and found the results were very consistent. We feel that this regional structure provides a good structure for the future as the contest grows.

Judging Results:

Problem	National Outstanding	Outstanding	Meritorious	Honorable Mention	Participant	Total
A	4	25	63	78	52	222
%	2%	11%	28%	35%	24%	
B	4	28	59	67	55	213
%	2%	13%	28%	31%	26%	
Total	8	53	122	145	107	435
%	2%	12%	28%	33%	25%	

NATIONAL OUTSTANDING TEAMS

Eastside High School, Gainesville, FL
Winchester Thurston High School, Pittsburg, PA
Maggie L. Walker Governor’s School, Richmond, VA.
Illinois Mathematics and Science Academy, Aurora, IL
The Charter School of Wilmington, Wilmington, DE
Mills Godwin High School, Richmond, VA
Hong Kong International School, Hong Kong (2 Teams)

REGIONAL OUTSTANDING TEAMS

The Charter School of Wilmington, Wilmington, DE
Chesterfield County Mathematics and Science High School, Midlothian, VA
China Welfare Institution, Shanghai, China
Eastside High School, Gainesville, FL (2 Teams)
Evanston Township High School, Evanston, IL (3 Teams)
Hangzhou Foreign Languages School, Hangzhou, China
Hanover High School, Hanover, NH (2 Teams)
Hanyoung Foreign Languages High School, Seoul, Korea (2 Teams)
Hong Kong International School, Hong Kong (6 Teams)
Illinois Mathematics and Science Academy, Aurora, IL (2 Teams)
Maggie Walker Governor's School, Richmond, VA (4 Teams)
Mississippi School for Mathematics and Sciences, Columbus, MS
Mills Godwin High School, Henrico, VA (3 Teams)
NO.2 High School of East China Normal University, Shanghai, China
North Springs Charter School, Atlanta, GA
Shanghai Foreign Language School, Shanghai, China (10 Teams)
Shanghai High School, Shanghai, China
Shenzhen High School, Shenzhen, China
Stanford University EPGY Online High School, Stanford, CA
The Ellis School, Pittsburgh, PA (2 Teams)
University High School, Irvine, CA (3 Teams)
Winchester Thurston High School, Pittsburgh, PA
Woodbridge High School, Irvine, CA (2 Teams)

Common Core State Standards: The director and the judges asked that we add this paragraph. Many of us have read the Common Core Standards and clearly realize the mapping of this contest to the Common Core mathematics standards. This contest provides a vehicle for using mathematics to build models to represent and to understand real world behavior in a quantitative way. It enables student teams to look for patterns and think logically about mathematics and its role in our lives. Perhaps in a future *Consortium* article we will dissect a problem (paper) and map the standards into it.

General Judging Comments: The judge's commentaries provide specific comments on the solutions to each problem. As contest director and head judge, I would like to speak generally about solutions from a judge's point of view. Papers need to be coherent, concise, clear, and well written. Students should use both spell and grammar checker before submitting a paper. Papers should use at least 12-point font. Students need to restate the problem in their own words so that the judges can determine the focus of the paper. Papers that explain the development of the model, assumptions with justifications, and its solutions and then support the findings mathematically generally do quite well. Modeling assumptions need to be listed and justified, but only those that come to bear on the solution (that can be part of simplifying the model). Laundry lists of assumptions that are never referred to in the context of model development are not considered relevant and deter from a paper's quality. The mathematical model needs to be clearly developed, and all variables that are used need to be well defined. Thinking outside of the "box" is also considered important by judges. This varies from problem to problem but usually includes model extensions or sensitivity analysis of the solution to the team's inputs. Students need to attempt to validate their model even if by numerical example or intuition if applicable. A clear conclusion and answers to specific scenario questions are all key components. The strengths and weakness section is where the team can reflect on their solution and comment on the model's strengths and weaknesses. Attention to detail and proofreading the paper prior to final submission are also important since the judges look for clarity and style. Citations are also very important within the paper as well as either a reference or bibliography page at the end. We encourage citations within the paper in sections that deal directly with data and figures, graphs, or tables. We have noticed an increase in use of Wikipedia. Teams need to realize that although useful, the information might not be accurate. Teams need to acknowledge this.

FACTS FROM THE 14TH ANNUAL CONTEST:

- A wide range of schools/teams competed including teams from Finland, Hong Kong, Korea, and China.
- The 435 teams from U.S. and International institutions represent a 47.4%% increase in participation.
- There were 1721 student participants, 1023 (62.7%) male and 608 (37.3%) female. There were 90 team members whose gender was not specified.
- Schools from only twenty-four states participated in this year's contest.

THE FUTURE:

The contest, which attempts to give the under-representative an opportunity to compete and achieve success in mathematics, appears well on its way in meeting this important goal.

We continue to strive to improve the contest, and we want the contest to grow. Any school/team can enter, as there are no restrictions on the number of schools or the numbers of teams from a school. A regional judging structure is established based on the number of teams.

These are exciting times for our high school students. Mathematics continues to be more than learning skills and operations. Mathematics is a language that involves our daily lives. Applying the mathematical principles that one learns is a key to future success. The abilities to recognize problems, formulate a mathematical model, use technology, and communicate and reflect on one's work are keys to success. Students gain confidence by tackling ill-defined problems and working together to generate a solution. Applying mathematics is a team sport!

Advisors need only be motivators and facilitators. They should encourage students to be creative and imaginative. It is not the technique used but the process that discovers how assumptions drive the techniques that is fundamental. Let students practice to be problem solvers. Let me encourage all high school mathematics' faculty to get involved, encourage your students, make mathematics relevant, and open the doors to success.

Mathematical modeling is an art and a science. Teach your students through modeling to think critically, communicate effectively, and be confident, competent problem solvers for this new century.

INTERNATIONAL FLAVOR OF THE CONTEST:

Next year's award format will differ as the contest continues to grow internationally.

Current Designation and proposed New Designation:
- Successful Participant still will be Successful Participant.
- Honorable Mention still will be Honorable Mention.
- Meritorious still will be Meritorious.
- Regional Outstanding Winner will be designated a Finalist.
- National Outstanding Winner will be designated as Outstanding Winner.

CONTEST DATES:

Mark your calendars early: the next HiMCM will be held in November 2012. Registrations are due in October 2012. Teams will have a consecutive 36-hour block within the contest window to complete the problem. Teams can register via the Internet at www.comap.com.

Math Models.org

It is highly recommended that participants in this contest as well as prospective participants take a look at the new modeling web site, www.mathmodels.org, which has a wealth of information and resources.

HiMCM Judges' Commentary

Professor William P. Fox, Naval Postgraduate School

Problem A: Space Shuttle Problem: No More Space Shuttles

On July 21, 2011, the 135th and final US Space Shuttle landed in Florida after its 13-day mission into orbit complete with a docking at the International Space Station (ISS). NASA will now have to rely on other nations or commercial endeavors to travel into space until a replacement vehicle is developed and constructed. Develop a comprehensive ten-year plan complete with costs, payloads, and flight schedules to maintain the ISS.

Some interesting facts possibly worthy of your consideration:

- The ISS is at full capacity with 6 astronauts, but can surge during shuttle docks to as high as 13.

- The ISS is scheduled to remain in service until at least the year 2020.

- Historically, it has cost between $5000-10,000 per pound to transport to the ISS using the US Shuttles. Shuttle missions have lasted approximately 10-14 days on orbit. Missions onboard the ISS are typically around six months.

- Recently, progress has been made within the private industry to launch unmanned rockets into space.

- Russia is willing to launch US astronauts into space for about $60 million each.

Judge's Comments:

Author: Jack Picciuto, United States Military Academy
Professor William P. Fox, Contest Director

This problem was of interest to the author who was an army aviator. First, the problem statement explicitly called for a plan that presented costs, payloads, and flight schedules. Many teams failed to provide a schedule. Addressing all three greatly increases the chance of recognition.

Although not explicitly asked for, it would be hard to address these three issues without considering schedule slippage. Have you ever seen or heard of a space flight taking off and landing on time? Very few teams considered or mentioned schedule slippage, nor did they make an assumption that there would not be schedule slippage due to any factors.

Many papers started with an equation for costs and then tried to explain it. Modeling says we start with variables and assumptions that lead to a model. We might assume linear or even nonlinear relationships but only after we examine the variables and assumptions.

The better papers this year attempted to present frameworks for choosing solutions. The mathematics required to do this are very accessible at the high school level.

There were many strengths in this year's papers. Almost all the papers did a reasonable job of estimating the costs, but few, if any, discussed the issue of bad weather and how weather related delays might impact the costs of the mission.

There were a wide variety of approaches used including simple algebra, statistics, and regression techniques. For those using regression techniques very few examined the residuals to insure the regression was useful. The R^2 value is not always a good indicator.

We provide an example of regression that shows why examining residuals is important.

Consider the following 4 sets of data:

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

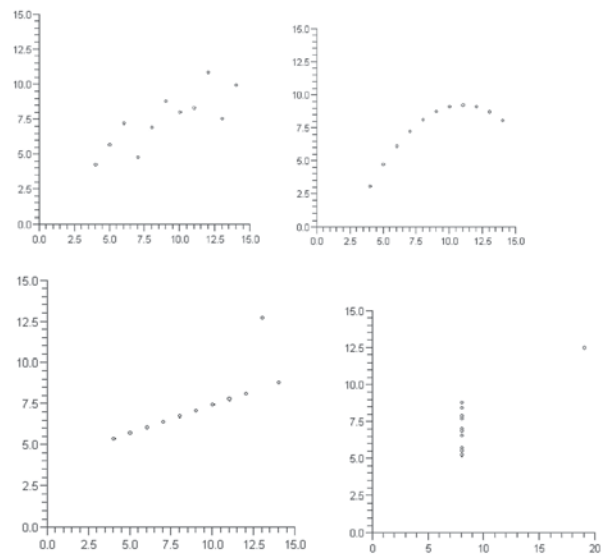
Suppose we fit the model $y = ax + b$ to each data set using the least-squares criterion. In each case the following model results:

$$y = 3 + 0.5x$$

The correlation coefficient in each case is 0.82, and $r^2 = 0.67$. The sum of the squared deviations between observed and predicted values is also the same. In particular,

$$\sum_{i=1}^{11} [y_i - (3 + 5x)]^2 = 13.75$$

These two numerical measures imply that for each case $y = 3 + 0.5x$ does about the same job explaining the data, and that it is a reasonable fit ($r^2 = 0.67$). However, the following scatter plots convey a different story:



A point to consider is how well the model $y = 3 + 0.5x$ captures the trend of the data. (This example is adapted from F. J. Anscombe, "Graphs in Statistical Analysis," Amer. Stat., 27, 1973, 17-21.)

There were some notable patterns of weakness. Many papers never considered foul weather, like storms in Russia or other countries. Few teams thought about returning to earth or delays in getting back. Schedule slippage was not addressed.

Some teams did consider training and costs to get space travelers to the country of departure.

Student groups should remember that the problems posed in this contest are not going to have a unique solution--in fact, they are designed **not** to have one. Students should remember that general high school mathematics are adequate to the task at hand--what we are looking for is evidence of good modeling of the problem with these tools, and then discussion of the implications of the model and its solution(s). We are looking for creative modeling and a thorough job of implementing the modeling process.

Problem B: Search and Find

Finding lost objects in not always an easy task, even when you have knowledge of a general location. Consider the following scenario: you have lost a small object, such as a class ring, in a small park, see map 1. It is getting dark and you have your pin light flashlight available. If your light shines on the ring, you assume that you see it. You cannot possibly search 100% of the region. Determine how you will search the park in minimum time. An average person walks approximately 4 mph. You have about 2 hours to search. Determine the chance you find the lost object. Then assume using map

2, a jogger is lost who was going on a 5-mile run. Determine how you search the region to have a good chance of finding the lost jogger (who might be not conscious). Assume it is night and you still only have your pin light.

Two maps were provided to the student.

Judge and Author's Comments:

William P. Fox, HiMCM Contest Director

The judges were amazed at the mathematics applied for this problem. They were amazed because the use of sophisticated mathematical concepts was unexpected. A majority of teams used graph theory to search for the objects, with many teams trying to define nodes and then minimize the circuit length to traverse. Most judges felt that graph theory was inappropriate and that teams using it should have maximized the coverage over time not minimized a circuit. With that said, we allowed the approach and read the papers that used it.

Students also were expected to randomly place the lost object and then try to find it through some mathematical plan or model. Very few teams considered the randomness of the lost object's location.

Students who compared areas to obtain the chance of finding the object were on a good track. The object is in an area of size X and the search area is size Y. The ratio of Y/X provides a good first estimate for the chance of finding the object. Some teams held the light vertically to get a circular search pattern on the ground; other teams tilted the light to maximize the coverage with an elliptical shape.

Some teams only solved one problem, but the problem statement asked for both to be solved.

The executive summaries for the most part are still poorly written although getting a little better. This has been an ongoing issue since the contest began. Faculty advisors should spend some time with their teams and advise them to write a good summary. Many summaries tend to be written before the teams start and only state how they will solve the problem. Summaries need to be written last and should contain the **results** of the model as well as a brief explanation of the problem. The executive summary should entice the reader, in our case the judge, to read the paper.

Few teams, if any, did sensitivity analysis or error analysis on their model. With the randomness of a lost item, this is a critical element.

We found that most teams did not do research to see if there were search and find methods to get them started. The references were generally weak.

We expected to see a wide variety of approaches from simple algebra through simulation models, but certainly not graph theory. We found very few simulation models, and they were never well explained nor were flow charts used. It was as if these techniques were a black box. As models, they should be explained as to what they do and why they could be used in the scenario.

Issues with graph theory included teams using Dykstra's algorithm for a minimal spanning tree. This does not insure that you will find the lost item. You need maximum coverage in minimal time.

We expect to draft a sample solution to Part I of this problem using typical high school mathematics and publish it in the next issue of *Consortium* for students and advisors to review.

General Comments from Judges:

Variables and Units: Teams must define their variables and provide units for each of them.

Computer generated solutions: Many papers used extensive computer code--especially the A* code. Computer code used to implement mathematical expressions can be a good modeling tool. However, judges expect to see an algorithm or flow chart from which the code was developed. Successful teams provided some explanation or guide to their algorithm(s)--a step-by-step procedure for the judges to follow. Code may only be read for papers that reach the final rounds, but not unless the code is accompanied by a good algorithm written in words. The results of any simulation need to be well explained and sensitivity analysis preformed. For example, consider a flip of a fair coin. Here is a general algorithm:

INPUT: Random number, number of trails
OUTPUT: Heads or tails
Step 1. Initialize all counters
Step 2. Generate a random number between 0 and 1.
Step 3. Choose an interval for heads, like [0.0.5]. If the random number falls in this interval, the flip is a heads. Otherwise the flip is a tails.
Step 4. Record the result as a heads or a tails.
Step 5. Count the number of trials and increment:
Count = Count + 1

An algorithm such as this is expected in the body of the paper with the code in an appendix.

Graphs: Judges found many graphs that were not labeled or explained. Many graphs did not appear to convey information used by the teams. All graphs need a verbal explanation of what the team expects the reader (judge) to gain (or see) from the graph. **Legends, labels, and points of interest** need to be clearly visible and understandable, even if hand written. Graphs taken from other sources ***should be referenced and annotated***.

Summaries: These are still, for the most part, the weakest parts of papers. These should be written after the solution is found. They should contain results and not details. They should include the "bottom line" and the key ideas used in obtaining the solution. They should include the particular questions addressed and their answers. Teams should consider a brief three paragraph approach: a *restatement of the problem* in their own words, a short description of *their method and solution* to the problem (without giving any mathematical expressions), and the *conclusions* providing the numerical answers in context.

Restatement of the Problem: Problem restatements are important for teams to move from the general case to the specific case. They allow teams to refine their thinking to give their model uniqueness and a creative touch.

Assumptions/Justifications: Teams should list only those assumptions that are vital to the building and simplifying of their mathematical model. Assumptions should not be a reiteration of facts given in the problem statement. Assumptions are variables (issues) acting or not acting on the problem. Every assumption should have a justification. We do not want to see "smoke screens" in the hopes that some items listed are what judges want to see. Variables chosen need to be listed with notation and be well defined.

Model: Teams need to show a **clear link** between the assumptions they listed and the building of their model or models. Too often models and/or equations appeared without any model building effort. Equations taken from other sources should be referenced. It is required of the team to show how the model was built and why it is the model chosen. Teams should not throw out several model forms hoping to WOW the judges, as this does not work. We prefer to see sound modeling based on good reasoning.

Model Testing: Model testing is not the same as testing arithmetic. Teams need to compare results or attempt to verify (even with common sense) their results. Teams that use a computer simulation must provide a clear step-by-step algorithm. Lots of runs and related analysis are required when using a simulation. Sensitivity analysis should be done in order to see how sensitive the simulation is to the model's key parameters. Teams that relate their models to real data are to be complimented.

Conclusions: This section deals with more than just results. Conclusions might also include speculations, extensions, and generalizations. This is where all scenario specific questions should be answered. Teams should ask themselves what other questions would be interesting if they had more time and then tell the judges about their ideas. **Strengths and Weaknesses**: Teams should be open and honest here. What could the team have done better?

References: Teams may use references to assist in their modeling, but they must also ***reference the source*** of their assistance. Teams are reminded that only ***inanimate resources*** may be used. Teams cannot call upon real estate agents, bankers, hotel managers, or any other real person to obtain information related to the problem. References should be cited where used and not just listed in the back of the paper. Teams should also have a reference list or bibliography in the back of the paper.

Adherence to Rules: Teams are reminded that detailed rules and regulations are posted on the COMAP website. Teams are reminded that they may use only ***inanimate sources*** to obtain information and that the ***36-hour time limit is a consecutive 36 hours.***

Problem A Summary: Hong Kong International School

Advisor: Kevin Charles Mansell

Team Members: Brendan Hung, Kristofer Siy, Jessica Tan, Daniel Zhu

The International Space Station (ISS) is a satellite in low Earth orbit, where scientists since 1998 have been investigating the long-term effects of a zero gravity environment on humans. On July 21, 2011, the final US Space Shuttle landed in Florida. Historically, these space shuttles have been used to maintain the ISS, by bringing up additional modules and by having astronauts conduct routine spacewalks and repairs. Without the space shuttles, we need to find a way to continue maintaining the ISS. We will consider the funds appropriated by the US government and the various methods of taking astronauts up into space in order to develop a comprehensive plan to maintain the ISS until 2021.

Problem A Summary: Hong Kong International School

Advisor: Kevin Charles Mansell

Team Members: Sau Man Cheng, Christopher Huie, Tara Lorimer, Martin Man

The United States Space Shuttle Program officially ended with the landing of Atlantis in Florida July this year. Without the Space Shuttle, National Aeronautics Space Administration (NASA) will be required to find other spacecraft to transport its astronauts to the International Space Station (ISS) for research and maintenance operations. Our group has developed a comprehensive, ten-year program for NASA that outlines costs, payloads, and flight schedules to maintain the ISS from 2012 to 2021.

We first determined the areas that we needed to allocate NASA's budget to. The four categories we decided on were personnel, operation, research, and resupply. Personnel includes the cost of the astronaut crew (NASA will be reliant on the Russians for transport to the ISS for the duration of the plan and each seat on the Soyuz spacecraft will cost $62.75 million) and the salaries of all of NASA's astronauts and ground operations crew. Operation includes the operation and maintenance costs for all of the ISS during the plan. As hardware on the ISS will soon exceed their certification limit, total operation costs include equipment recertification and replacement costs. As it is difficult to determine the cost of the research equipment needed to be transported to the ISS for as yet undetermined research programs, a model of future cost-per-kilogram rates on Commercial Orbital Transportation Services (COTS) was developed to allow calculation of cost by research material mass. Expenses for procurement of research materials was projected using data from NASA's budget estimates. Resupplying of the ISS with food, water, air, dry cargo etc. by NASA through unmanned commercial spacecrafts were determined to be necessary 3 times a year based upon the regular resupply missions of the Russian Progress spacecraft for the past decade.

Flight schedules for the crew and resupplying cargo were determined based upon NASA's existing flight schedules. This ten-year plan will allow NASA to continue maintaining and operating the ISS at least until the year 2021.

Problem A Summary: Illinois Mathematics and Science Academy

Advisor: Steven M. Condie
Team Members: Henry Deng, Matt Gietl, Andrew Ta, Matt Yang

In this paper, we model the costs and flight schedules necessary by NASA to maintain the ISS for the next ten years. The solution is divided into two primary components: Budgeting and flight schedule. For costs, we examined NASA contracts with other aerospace agencies, costs to sustain astronauts, as well as repair costs and contingency plans, while taking into account the effect of yearly inflation. We searched for consistent patterns in manned flights to the ISS, and extrapolated to create a flight plan for the following decade. In relation to scheduling and payload delivery, we analyzed trends in the masses of previous resupply deliveries.

Combined altogether, our model provides a comprehensive timetable for manned spacecraft launches while also providing an expected budget that contains a built-in cushion to overcompensate for inflation and emergence situations. NASA will continue to regularly send astronauts on expedition missions and astronauts will stay on the ISS for 6-month periods. Routine supply missions carrying oxygen, water, propellant, research equipment, and other hard cargo will continue unabated throughout all 10 years with a shift of dependence on Russian rockets to private contractors. A major section of the ISS budget, station repair and maintenance costs, are expected to slowly increase over time with a budget contingency plan provided in the event of critical damage necessitating module replacement. Two simple summary tables are provided below:

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Supply Payload (lb)	62197	48132	45027	55798	44602	62121	44356	62188	38533	44994
Expeditions	4	4	4	4	4	4	4	4	4	4
Supply Flights	5	4	4	4	4	5	4	5	3	4
Total Flights	9	8	8	8	8	9	8	9	7	8

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Total Spending (in million USD)	3169	3115	3185	3437	3571	3893	3769	4179	4206	4508
ISS Maintenance and Repair	1676	1803	1939	2085	2241	2409	2589	2780	2987	3208
Research Materials	235	238	255	281	302	327	355	384	416	450
Crew and Cargo	1258	1074	991	1071	1028	1157	825	1015	803	850
Seats on the Soyuz	255	271	287	305	324	0	0	0	0	0
Cargo Transportation	1003	804	704	766	704	1042	704	885	666	703.5
Crew Transportation	0	0	0	0	0	115	122	129	137	146

Problem A Paper: Eastside High School

Advisor: Carl Henriksen
Team Members: Xingchen Li, Alexandra Sourakov, Yuxin Zhang, Wenli Zhao

1. INTRODUCTION

Earlier this year, NASA's 135th shuttle mission marked the end of an era. The shuttle pioneered space exploration, as the first reusable spacecraft. It played a crucial role in building and maintaining the International Space Station (ISS) and as a facility for unique research. Space shuttles may now be retired, but it is still necessary to maintain the ISS and send American astronauts there. As the economy declined in the past decade, funding for space exploration decreased, making it necessary to find a cost effective alternative to shuttles. We propose a 10-year plan for ISS maintenance.

2. ASSUMPTIONS AND JUSTIFICATIONS

Assumption: NASA's ISS budget will follow the projecte linear trend.

Justification: NASA has projected budgets from 2011 to 2016 that follow a linear pattern with an r^2 value of 0.992

Assumption: The amount of money for any action remain- about the same.

Justification: For the purpose of our model, the value of the dollar does not change due to inflation or other factors.

Assumption: There will be no major spacecraft malfunctions that affect the monetary distribution of our plan.

Justification: Spacecraft have been tested many times to ensure their safety.

Assumption: Astronauts serve a term of 6 months at the ISS.

Justification: In the past, astronauts have remained at the ISS for periods ranging from 136 days to 215 days.

Assumption: The ISS will need about the same supplies per year until 2020.

Justification: The ISS goals and life support needs will not change until 2020.

Assumption: Cargo must be delivered to the ISS about once every 4 months.

Justification: This is how it has been done for the past few years.

Assumption: $60 million will buy a U.S. astronaut a trip to the ISS and back.

Justification: Otherwise, our brave men and women will have no way to return. Also, it is unreasonable Russia would waste space on return trips.

Assumption: A spacecraft that can fly both manned and unmanned costs $3.5 billion to build. The technology to make this spacecraft is available.

Justification: The spacecraft *Orion* has this cost and similar functions.

Assumption: There will be no new technological advances to significantly decrease the maintenance cost of the ISS up to 2020.

Justification: Since we are in an economic crisis, innovations are unlikely.

Assumption: Russia can transport U.S. astronauts.

Justification: Russia took U.S. astronauts to the ISS in their Soyuz spacecraft in 2011.

Assumption: We have access to NASA's ISS budget.

Justification: According to NASA's website, it is reasonable to assume the ISS budget would be used for the ISS.

3. PARAMETERS

- Projected total ISS budget (2011–2020): $31,360.5 million
- Supplies needed for the ISS per year: 12,000–16,000 lbs.
- Cost to build a new spacecraft: $3.5 billion
- Money paid to Commercial Orbital Transportation Services: $3.5 billion
- Average cost to launch a spacecraft: $450 million
- Fee of Russia transporting an U.S. astronaut to the ISS: $60 million
- Average cost per pound of payload: $5000–10000 (we use $7500)
- Average annual ISS supply cost: $1,500 million
- Amazing 10-year maintenance plan: *priceless*

4. VARIABLES AND EQUATIONS

C_T – Total cost
C_1 – Cost 2011-2016
C_2 – Cost 2017-2020
B – Total budget for space shuttles and ISS maintenance (2011-2020)
A – Number of astronauts sent to the ISS by Russian spacecraft
Y – Year from 2010
L – Number of *Ophiucus* launches
W – Weight of cargo and passengers (lbs.) on *Ophiucus*

5. OBJECTIVES

Our task is to create a 10-year plan for ISS maintenance. After extensive research of NASA's budget, contracts, and ISS maintenance, we divided our plan into two time periods. From 2011 to 2016, we uphold our contract with private companies to deliver supplies to the ISS. We also pay Russia to launch U.S. astronauts into space. We will begin to build a new spacecraft, *Ophiucus*, in 2011, which will be ready by 2017 and can then be used for both manned and unmanned missions.

To model our plan's cost, we create a cost equation based on number of astronauts and cargo weight. We then find the average or projected cost based on parameters we do not use in our initial model. Lastly, we create a detailed flight schedule for the next ten years.

TASK 1: ESTABLISHING BUDGET

1.1 INTERNATIONAL SPACE STATION BUDGET

NASA has projected the shuttle program's budget from 2011 to 2016. Starting in 2013, annual budgets drop to about $800,000 a year. We predicted budgets starting in 2017 to remain at $800,000. As shuttle funding has decreased greatly, building another one is unreasonable. Therefore, our missions to the ISS will be funded only by NASA's ISS budget, which in 2016 is estimated at $3.17 billion. Using linear regression, we projected budgets for 2017–2021.

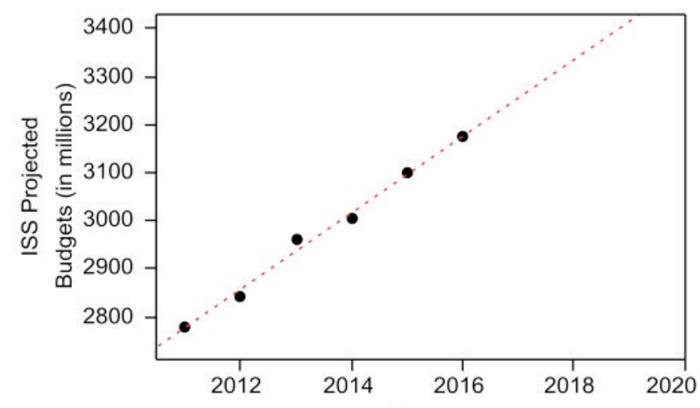


Figure 1: Linear Regression for ISS Projected Budgets

$$\textit{ISS Projected Budgets}\ (\text{in millions}) = -157493.8 + 79.7(\textit{Year})$$

NASA's projected ISS budget follows a linear model (Figure 1) with $r^2 = 0.992$, indicating a strong correlation in which 99.2% of budget variability is accounted for by the year. Values for 2017–2020 were predicted using the regression equation. (See Table 1 for precise values.)

A residual plot (Figure 2) shows random scatter, so a linear model is appropriate.

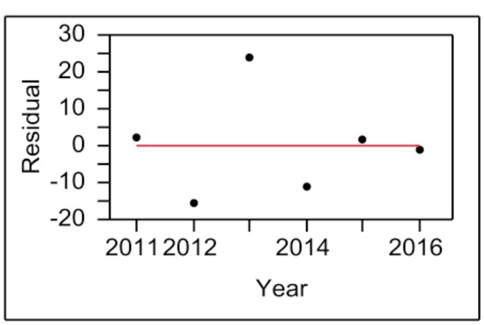


Figure 2: Residuals Plot for ISS Projected Budget

Year	ISS Projected Budget (in millions)	Residuals
2011	2779.8	2.37619048
2012	2841.5	-15.620952
2013	2960.5	23.6819048
2014	3005.4	-11.115238
2015	3098	1.78761905
2016	3174.8	-1.1095238
2017	3255.6	–
2018	3335.3	–
2019	3415	–
2020	3494.7	–

Table 1: Data table for ISS projected budgets

1.2 TOTAL PROJECTED BUDGET (2011-2020)

By adding yearly ISS budgets, we got a total of $31,360,500,000. This value is not used in our model. It is only considered when analyzing predicted total mission and maintenance cost. That is, we must be sure our cost does not exceed this total.

TASK 2: CREATING COST-EFFICIENT MODEL (2011–2016)

2.1 MANNED FLIGHTS PROVIDED BY RUSSIA

Without a shuttle, NASA needs the help of other nations and private companies. Russia offers to take U.S. astronauts on their shuttles for $60 million per person, and we accept this offer for the first seven years (2011-2016).

2.2 UNMANNED MISSIONS USING COMMERCIAL RESUPPLY SERVICES

Commercial Orbital Transportation Services (COTS) is a NASA program that uses private corporations to deliver of crew and cargo to the ISS. COTS relates to vehicle development, and Commercial Resupply Services (CRS) to deliveries.

NASA has contracts with Orbital Sciences Corp. and Space Exploration Technologies (SpaceX) for ISS resupply. NASA has ordered 8 flights from Orbital and 12 from SpaceX. Contracts run through 2016 and call for delivery of a minimum of 40 tons of cargo. The contracts' total value is about $3.5 billion.

These agreements will fulfill NASA's need for cargo delivery to the ISS after shuttle retirement. 2 tons of cargo will be carried on each flight.

2.3 CARGO TRANSPORT

Cargo itself costs an average of $1.5 billion per year, which includes everything needed to maintain the ISS. These estimates are based on a 2005 ISS mission:

$70 million will be used to develop new hardware (e.g. navigation, data support, environmental controllers).

$800 million will be used for spacecraft operations.

$350 million will be spent on software to maintain the integrity of the ISS design and its continuous, safe operation.

$140 million will be spent for purchase supplies, cargo and crew.

2.4 TOTAL COST OF CARGO AND MISSIONS (2011–2016)

The total cost of these arrangements can be modeled by:

$$C_1 = 60A + 1500Y + 3500,$$

where C_1 = Total costs 2011–2016 (in millions)
A = Number of Astronauts brought to ISS by Russian spacecraft
Y = Number of years from 2010
3500 = Contract fee for CRS by private companies (in millions)

TASK 3: CREATING COST EFFICIENT MODEL (2017–2020)

In 2011, we propose to start building *Ophiucus*, which can fly manned or unmanned. It will be functional by 2017 and be retired after 12 launches. The estimated cost is $3.5 billion, based on similar spacecraft built in the past.

After building *Ophiucus*, NASA will not need to buy flights from Russia or corporations. However, each launch will cost $450 million. We propose three cargo drops per year of 2 tons each (*Ophiucus* can carry astronauts and 2 tons in the same trip). In our model, we use the average of $7500 per pound of cargo.

Therefore, we can model the total cost for 2017–2020 by:

$$C_2 = 3500 + (0.0075W + 450)L,$$

where C_2 = Total costs 2017–2020 (in millions)
L = Number of *Ophiucus* launches
W = Number of pounds (payload) on *Ophiucus*
450 = Cost per launch (in millions)
3500 = Cost to build *Ophiucus* (in millions)

TASK 4: COMBINED COST EFFICIENT MODEL

Adding equations for 2011–2016 and 2017–2020 gives us a single cost equation:

$$C_T = L(450 + 0.0075W) + 60A + 1500Y + 7000.$$

TASK 5: PREDICTING COST USING ESTIMATES

For the equation derived in Task 4, our estimates are:

$L = 12$: This is based on three COTS supply launches from an earlier year. Since astronauts and cargo can fly together, we need three launches per year

$W = 4310$: This includes a 4000 lb. average payload per *Ophiucus* flight and a 310 lb. passenger weight based on the average human weight (155 lb.) used for elevator and aircraft carrying capacity. On each flight, two astronauts are present.

$Y = 10$: This is the number of years for our model (2011-2020).

$A = 24$: We propose to deliver two astronauts per six months to the ISS. This means we need to purchase 4 flights per year from Russia, for six years.

$$C_T = 12(450 + 0.0075 \times 4310) + 60 \times 24 + 1500 \times 10 + 7000$$
$$C_T = 29227.8$$
$$\text{Total Cost} = \$29{,}227{,}800{,}000$$
$$B - C_T = \$2{,}132{,}700{,}000.$$

The $2,132,700,000 that remains after total cost is subtracted from budget can be used for equipment malfunction or other such emergencies and needs.

5.2 SENSITIVITY ANALYSIS

To test sensitivity, analyzed each parameter (or variable) that is subject to change. We found the percentage of change of total cost when increasing each parameter by 5%, with the exception of cost of transport per pound, for which we used the $7500 average and 33% change. To ensure total cost did not exceed budget, we also found the percentage of change to total cost when parameters are

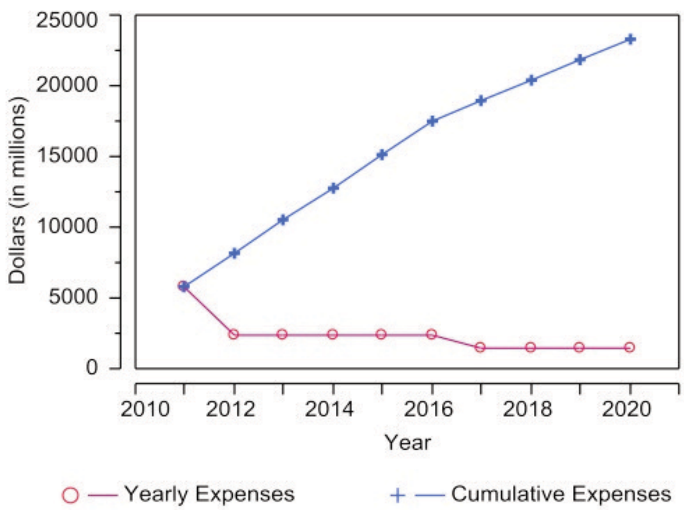
Parameters/ Estimates	Original	% Change to Parameter	Modified Parameter	Result	% Change Result
Cost per *Ophiucus* Launch	$450,000,000	5	$472,500,000	$29,497,900,000	0.923
Cost per pound of *Ophiucus* Launch	$7,500	33	$10,000	$29,357,200,000	0.44
		–33	$5,000	$29,098,600,000	–0.44
Average Weight per *Ophiucus* Launch	4,310 lbs.	5	4,525.5 lbs.	$29,111,530,000	0.07
Cost per Astronaut with Russia	$60,000,000	5	$63,000,000	$29,299,900,000	0.25
Cargo Cost per year	$1,500,000,000	5	$1,575,000,000	$29,977,900,000	2.57
Cost to Build *Ophiucus*	$3,500,000,000	5	$3,675,000,000	$29,402,900,000	0.60
All Parameters	Varies	Varies	Varies	$30,650,060,000	4.87

Table 2: Sensitivity Analyses

increased 5% and cost of cargo per pound is increased 33%. Table 2 shows our results.

The change when any parameter is increased is less than 1%, with the exception of the cost of cargo per year. This is expected because cargo comprises nearly half of our annual budget. This shows that our model is very robust. When all parameters are increased simultaneously, the change is less than 5%.

5.3 COSTS OVER TIME



The graph in Figure 3 shows yearly and cumulative expenses. After construction of *Ophiucus* in 2017, there is a significant decrease in yearly expenses.

Figure 3: Estimated Cumulative and Yearly Expenses

TASK 6: DEVELOPING THE FLIGHT SCHEDULE

We made flight schedules for 2011–2016 and for 2017–2020. The first uses CRS and Russia, and the second uses *Ophiucus*. All planned dates are subject to change due to environmental factors or malfunctions.

Flight Schedule (2011–2016)

CRS missions occur about every three months, sometimes four. These spacecraft collect ISS debris and are incinerated on reentry into Earth's atmosphere. Since there are 20 missions contracted and 19 scheduled, we reserve one for an emergency. Each mission lasts 12 days, the average duration of a mission.

Launch Date	Return Date	Crew	Travel Means
July 8, 2011	July 21, 2011	2	Russia
Oct 30, 2011	N/A	0	COTS
Jan. 21, 2012	N/A	0	COTS
Jan. 27, 2012	Feb. 8, 2012	2	Russia
Apr. 26, 2012	N/A	0	COTS
July 21, 2012	Aug. 2, 2012	2	Russia
Aug. 2, 2012	N/A	0	COTS
Nov. 9, 2012	N/A	0	COTS
Jan. 21, 2013	Feb. 2, 2013	2	Russia
Feb. 6, 2013	N/A	0	COTS
May 13, 2012	N/A	0	COTS
July 21, 2013	Aug. 2, 2013	2	Russia
Sept. 30, 2013	N/A	0	COTS
Dec. 27, 2013	N/A	0	COTS
Jan. 21, 2014	Feb. 2, 2014	2	Russia
Apr. 4, 2014	N/A	0	COTS
July 11, 2014	N/A	0	COTS
July 21, 2014	Aug. 2, 2014	2	Russia
Oct. 18, 2014	N/A	0	COTS
Jan. 21, 2015	Feb. 2, 2015	2	Russia
Jan. 25, 2015	N/A	0	COTS
June 2, 2015	N/A	0	COTS
July 21, 2015	Aug. 2, 2015	2	Russia
Nov. 9, 2015	N/A	0	COTS
Jan. 21, 2015	Feb. 2, 2015	2	Russia
Feb. 16, 2016	N/A	0	COTS
May 23, 2016	N/A	0	COTS
July 21, 2016	Aug. 2, 2016	2	Russia
Aug. 30, 2016	N/A	0	COTS
Dec. 7, 2016	N/A	0	COTS

Table 3: Flight Schedule (2011–2016)

Flight Schedule (2017–2020)

Ophiucus missions occur every six months to deliver astronauts to the ISS, and an unmanned flight occurs between manned missions. That is, the missions occur each January, April, and July. All manned missions occur over 12 days.

Launch Date	Return Date	Crew	Travel Means
Jan. 21, 2017	Feb. 2, 2017	2	*Ophiucus*
Apr. 21, 2017	TBD	0	*Ophiucus*
July 21, 2017	Aug. 2, 2017	2	*Ophiucus*
Jan. 21, 2018	Feb. 2, 2018	2	*Ophiucus*
Apr. 21, 2018	TBD	0	*Ophiucus*
July 21, 2018	Aug. 2, 2018	2	*Ophiucus*
Jan. 21, 2019	Feb. 2, 2019	2	*Ophiucus*
Apr. 21, 2019	TBD	0	*Ophiucus*
July 21, 2019	Aug. 2, 2019	2	*Ophiucus*
Jan. 21, 2020	Feb. 2, 2020	2	*Ophiucus*
Apr. 21, 2020	TBD	0	*Ophiucus*
July 21, 2020	Aug. 2, 2020	2	*Ophiucus*

Table 4: Flight Schedule (2017–2020)

6. DISCUSSION

6.1 STRENGTHS

- Our model is simple and understandable.
- Our model makes space travel independent of other nations and private companies.
- Our model allows for possible future malfunctions.
- Our model predicts projected cost in addition to addressing the cost plan.
- As compensation for many assumptions, we did many sensitivity analyses.

6.2 WEAKNESSES

- We made many assumptions.
- Parameters are likely to be subject to change.
- Our ISS projected budget was obtained by extrapolation, which could not be avoided since our plan projects cost and financial situations.
- Our model does not account for inflation.
- Our sources were sometimes vague or contradictory, which made further assumptions necessary on which source was most reliable.

6.3 TOPICS FOR FUTURE STUDY

- A future model should be in accord with projected economic conditions. For example, inflation should be taken into account.
- Cooperation with other nations should be considered.
- Technological advances that increase maintenance effi ciency should be predicted and accounted for.
- Future studies should create a plan that considers NASA's complete budget and reallocate money into ISS maintenance as needed.
- They should also divide the larger overall budgets, such as ISS maintenance, into smaller individual budgets such as materials that ensure the integrity of the ISS structure and those needed for life support.

Problem B Summary: Winchester Thurston School

Advisor: Stephen J. Miller

Team Members: Sonu K. Bae, J. Aaren Barge, Morgan C. Culbertson, Jesse S. Lieberfeld

I have lost my ring! We all lose things, and the hardest part is always finding where we left them. In Problem B we were assigned the intimidating task of finding an object in a large park, and a lost jogger in an even larger park using only a penlight. Needless to say, without the mathematical and computer models contained herein, it would be like finding a needle in a haystack. However, this task is not the impossibility that it may seem at first.

When searching for a lost object, it is important to first determine the locations where there is the greatest probability of finding it. Once this has been accomplished, the next step is to develop a method for searching those locations efficiently. In both problems, we are asked to search for an object with few initial clues as to its location. However, we can make several assumptions about how the objects, and those searching for them, behave. A lost ring, for example, will not arbitrarily change its location. We therefore base our method of search for the ring off of our subject, the person who lost the ring, who is far more likely to have changed his location, and whose location changes may be predictable. To find where the ring is located, we must focus our model on where our subject would have spent the most time because he could have only dropped the ring in a place where he was present. If we are able to determine the areas where our subject spent the majority of his time, or was likely to have spent the majority of his time, those areas will be where we are most likely to find the ring. In order to determine where these areas of high probability are, we need to answer two questions: First, where must our subject have gone? And second, where could he have remained for a period of time? Because the park can only be accessed by car, we can therefore deduce that Tim must have parked his car at one of the fifteen parking lots. The areas near parking lots, therefore, are much more likely to have been places that Tim walked than other stretches of like Aikens road. In addition, Tim is likely to have stopped for a period of time at picnic tables, restrooms, or the Swimming Pond. If he stopped at these areas, the area of trail surrounding them is also likely to have been covered by our subject walking to or from the location. We have now established where the ring is most likely to be found, it is most likely found at or around one of the locations mentioned previously: parking lots, restrooms, and picnic tables.

Having determined the criteria necessary to determine whether an area is beneficial to search, being close to the attractions of the park, we can construct a probability density map that for any point in the park will determine how likely it is that the ring is located at that point. Using this map, we must construct a path that allows us to search all of the highest density, highest probability, areas as quickly as possible. We start at each high-probability location, and attempt to find the shortest route from each location that passes through all the others. We then select the shortest of these routes, which is about 3.9 miles. As Tim can walk 8 miles in the allotted 2 hours, this leaves him with 5.1 miles worth of walking to pass by as many other facilities as possible and return to the starting point. This will enable him to cover the 80% of the road that gives him the highest probability of finding the ring.

In the second scenario, the search for Sam, our subject, presents a different problem: how to find an object that may or may not be moving. Because we cannot assume anything about Sam's behavior as we search for him (we assume that he will not go in closed-off areas), we are forced to gain our information from his natural inclinations, previous behavior studies, and from the park's layout.

As in the first problem, we can assume that Sam must have accessed the park by car. Therefore, there are five possible places where his run could have begun. We also assume that, as he planned to run for five miles, so at most he would have intended to turn around after 2.5 miles and then turn around. It is therefore safe to assume that at the time Sam became lost, he was within a 2.5-mile radius of one of the parking lots. He therefore is most likely to be found in an area that is within 2.5 miles of multiple parking lots. We may also assume that Sam is not lost if he is in a parking lot or at another recognizable landmark. He is therefore is most likely to be found at the edge of this 2.5-mile range, where he is farthest away from these landmark.

Sam does not, however have the same probability of being lost when starting from each of the five parking lots. There are more forks in the road in the vicinity of some parking lots than in others, giving Sam a greater probability of taking a wrong turn and becoming lost if he has started from one of those parking lots. It therefore makes sense to start our search in the area with the highest concentration of forks. In effect, Sam's probability of being lost in a given area is determined by three factors: number of 2.5-mile radii within which the location falls, distance from landmarks, and likelihood of making a wrong turn as determined by the number of forks in the area. The third of these factors gives us our starting point; the first and second give us the order in which we should search the park making sure to hit as many forks in the road (referred to herein as decision points) because research suggests that Sam, conscious or unconscious, will be located near the decision point where he make a wrong turn. We begin at the parking lot from which Sam has the greatest probability of making a wrong turn, lot 4, and then pass through the areas within the range of the most parking lots and with the highest concentrations of wrong turns from among the areas through which we have not yet searched.

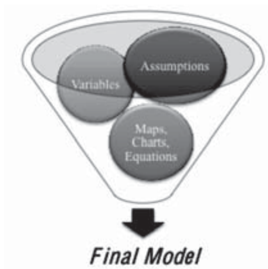
The search for a lost item in a vast area can seem like an impossible task, but with careful and methodical planning, the most likely areas in which the object could have been lost can be quickly and easily found. If these areas are searched in an efficient manner, a seemingly lost cause can be turned into a purposeful and systematic endeavor in which success is highly likely.

Problem B Summary: The Charter School of Wilmington

Advisor: Regina Fody

Team Members: Steven J. Burcat, Christopher X. Deng, Byron B. Fan, Martin J. Kurian

In this solution, we created innovative models that can be used to solve two common problems that were presented to us: computing the probability of finding a lost item in a small park, and developing the most effective method for searching a larger park for a lost jogger. Since the two problems had numerous similarities, we had similar solutions to both of the problems. To create these solutions, we had to first identify the variables that had to be considered in our models. Many of the variables were similar for both problems, and some of them included, but were not limited to: distance of trails, points of interests in the park, weather at the time of the event, parking station locations, the condition of the jogger, and some things even as detailed as the strength of the light of the searchers. Our model took into account all of these variables, and used maps, graphs, flow charts, and equations to come up with the optimum solution to the problem. The strengths of this model are that it accounts for a multitude of variables and the use of many visuals (such as maps, graphs, and flow charts) which help to visualize and allow for a more precise solution. Moreover, the equation that we developed for part B of the problem (the lost jogger) effectively finds the probability of finding the lost jogger while taking into account a myriad of variables and also helps to find the jogger. One weakness to our overall model is that many assumptions had to be made with limited data. Some of these assumptions include the general starting point of the jogger, ruling out areas where the jogger could not possibly be, and assuming that the lost object had to be on a trail. The way that our model will be tested is through actual values being plugged in (for the case of our equation) and real life experiences in similar situations to the given problem (lost items and lost joggers). The algorithm right shows the mixture of steps that led to our model:



Problem B Summary: Maggie L. Walker Governor's School

Advisor: Dickson Benesh
Team Members: Judy Hou, William Thomason, Melody Wang, Arthur Wu

Losing objects is an inevitability of human existence. We lose objects, and occasionally, we lose ourselves.

In the first section of the problem, we were tasked with finding a lost object in Hopkinton State Park with only a pen flashlight at our disposal and the sun quickly setting. We constructed four different scenarios that tailor our model to an individual's search, depending on what the pre-existing conditions exist. Our four models take into account various contingencies, considering whether or not the individual knows generally where the object is, and whether or not the individual stayed on the path. In situations where the individual stayed on the path, we aimed to maximize efficiency by minimizing redundant edge traversal through the use of Euler Circuits and Paths. If the individual knew a general area in which the object may be, we instead maximized area, increasing the probability that the object would be found by using an Archimedean Spiral, which has been shown to be the most efficient search pattern for an open area. We established 100%, 100%, 70.46%, and 2.21% certainty that our object would be found for each of the four scenarios. In the scenario with the lowest probability, the low level of certainty is due to the complete lack of parameters for search; in this case, we maximized the area covered.

The second part of the problem asked us to devise a model for finding a jogger who has gotten lost at night in Fort Ord, using the pen flashlight as our only light source. We determined that the jogger's intention of going on a 5.0 mile run meant that he or she is no more than 5.5 miles deep into the park, with 0.5 miles accounting for any wandering on the part of the jogger. For this scenario, we constructed two models: one that accounts for when the jogger is stationary and another for when the jogger is mobile. If the jogger is unconscious (and therefore stationary), the searcher must travel along the Euler path that extends a maximum of 5.5 miles in trails into the park depending on the jogger's point of origin. If conscious, the jogger could rationally decide to remain in the same position, wander frantically, or wander out and return to the same starting point. Because we did not have any basis to accurately predict the behavior of the jogger, we made certain assumptions about the jogger's path of travel. We once again used Eulerization to calculate the optimal search path to find a jogger moving in the park, based on his or her starting area. If this area-specific search along the Euler path is unsuccessful, then we decided that the searcher should travel along the most popular and prominent paths in the entire park. The combination of both methods greatly increases the efficiency and likelihood of finding the lost jogger, regardless of whether the jogger is conscious or unconscious.

While there were several limitations, such as the exclusion of terrain from our consideration, we believe that we have developed an optimal model with the given information that satisfies our logical assumptions. Our plan encompasses a variety of plausible scenarios, while minimizing total search time and maximizing probability of success.

Problem B Paper: Mills Godwin High School

Advisor: Todd A. Phillips
Team Members: Pascal Dangtran, Shakthi Ganesan, Jeffery Holste, Stephen Thompson

INTRODUCTION

Many searches begin at night. Without artificial lighting, it would be nearly impossible to find a lost person. In this problem, the lone searcher has only a penlight. The question arises whether or not the searcher can see the whole trail. The answer is, with a few exceptions, yes. Because the penlight gives off 20 lumens, through a calculation of the working distance (the distance in which one can see at a certain level of luminance), the searcher could see 19.09 feet away with the luminance level of a full moon. Any ambient light would only increase this. If he rotates the flashlight from 45° to the left of his head to 45° to the right each second and then the opposite direction in one second, he will periodically see the entire pathway. Exceptions occur with short paths or highly curved paths. In the first case there will be small segments on either side that he misses; in the second he will miss small patches on the side of the curvature. However, in either case the distance of the path would be so small that the probability of the lost object or person being there is insignificant enough that any effect due to curvature or length is negligible. Thus, it is reasonable to assume that he will see the entire path in front of him.

In this paper we seek to solve two problems. First, we develop a search model to find a lost object, in which we present three models: one taking into account the connectivity of each path, another taking into account the distance of the path from parking lots, and a third doing both. Through these models we seek to develop and improve upon a search method that would both improve our chance of finding the lost object and improve the accuracy of the calculation of the probability of finding it. Second, we develop a model to optimize a search pattern given the possible movement of a lost jogger. To accomplish either task, it is necessary to find the length of each segment. To do this, we use a piece of string to follow a segment on the map and then measure the piece of string and use the scale to find the length. Segments are any subsections of a path that are between two intersections of trails and/or paved roads.

ASSUMPTIONS FOR FINDING SMALL OBJECTS (FOR ALL MODELS)

1. The searcher walks at a constant 4mph.
2. The penlight is 20 lumens.
3. If the light touching the object has a lux of X, the object will be found.
4. The searcher has no memory of where he has been.
5. The searcher walks down the middle of paths.
6. The path is nine feet wide (Virginia state law for maximum for vehicles).
7. The searcher looks where the light shines and has a vision span of the radiation angle of the penlight.

ASSUMPTIONS SPECIFIC FOR SMALL OBJECT MODEL #1

1. The searcher cannot walk on paved roads, except to cross them.
   a. The distance walked on a paved road is negligible.
2. The searcher has a vehicle that, when driven, main tains a constant speed of 15 mph.
3. The searcher must leave the park two hours after entering.
4. The searcher always enters and exits the park through the main entrance.
5. The object is on a path.
6. The number of paths connected to the s^{th} segment's trail is directly proportional to the probability that the object is in that segment.

SMALL OBJECTS MODEL # 1 EQUATION AND CALCULATION

$$P_{k_1} = \sum_{s=1}^{z}\left[\frac{l_s C_s}{\sum_{i=1}^{n} l_i C_i}\right]$$

k is the union of all segments in a given search path
P_{k_1} is the probability of finding the missing object for a given k
z is the total number of segments accounted for in the search
l_s is the length of the s^{th} segment of K
Cs is the number of connected paths to the s^{th} segment trail,
l_i is the length of the i^{th} segment of the map
C_i is the number of connected paths to any trail.

$$\frac{(l_s)(C_s)}{\sum_{i=1}^{n}(l_i)(C_i)}$$

The probability of the object being on a specific segment.

EQUATION DERIVATION

Google developed a ranking system for web pages. They correlate the number of pages linking to a website to that site's popularity. We used a similar method. The more populated roads should be those with more linking trails. For instance, a trail connected to three others is more popular than one with one connection. Let C_s be the number of paths connected to the s^{th} trail.

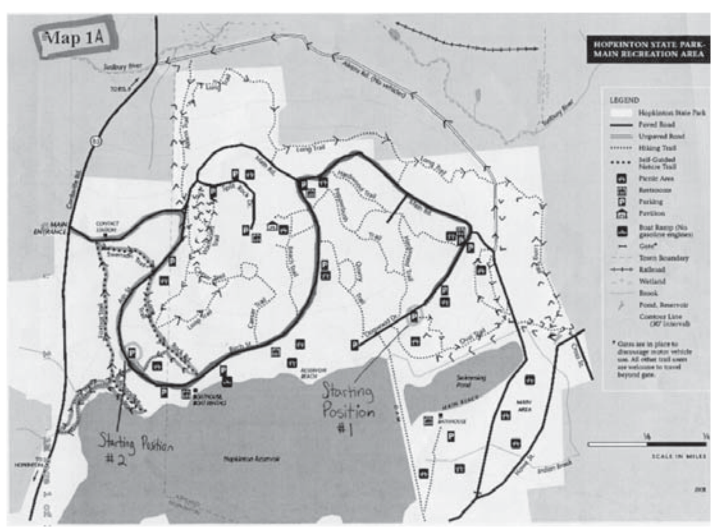
We can use this to weight each trail *segment* with the popularity measure, C_s. Since without this weight, the probability of finding the missing object in the s^{th} segment would be equal to its length, , divided by the total length of all trails. We can multiply by the popularity measure C_s to get the weighted measure. Furthermore, the probability of the ring being on the s^{th} segment is multiplied by the popularity measure C_s divided by the sum of the weighted measures of every segment in the park. If we define K to be the set of all segments in a given search path, then the probability, P_k, of K containing the ring is equal to the sum of the aforementioned probabilities for every segment in K.

MAXIMIZING THE PROBABILITY OF FINDING OBJECT

STEP 1:

$\sum_{i=1}^{n}(l_i)(C_i)$ is calculated by taking the lengths of each segment of each trail and multiplying it by the number of paths connected to that trail (we did this for all 58 segments and all 17 possible trails). This gave the weight of each segment. The "weight" is a numerical scale that determines how likely it is to find the object on a specific segment. Once the weight of each segment was found, the total weight of each path was found by adding the weights of all the segments in one path. The total weight (of all the paths) gave us the total weight of map 1, 15.45.

STEP 2: DEVELOPING THE BEST SEARCH PATH



After finding the sum of the weights of every segment on the map, we needed a way to search the park that allows for the greatest probability of finding the object. Our search method allows for visiting paths with the highest total weight, which increases the probability of the object being located, given our formula. However, there are no parking lots near the two most likely trails (those with the greatest weight), so we constructed the model to include other trails.

We picked the parking lot on Dogwood Drive that touches Oval Trail because this lot allows access to Aikens Road and Long Trail, the paths with the highest weights. The model begins with the searcher driving from the main entrance to this lot. The searcher goes past Contact Station, up Main Road, and then down Dogwood Drive to the lot, which takes 5.111 minutes. Once the searcher reaches the lot (Starting Position #1 on Map 1A), he begins the search. This is the portion of Map 1A highlighted in yellow, with direction marked by black arrows. In this part of the model, the searcher searches all of Aikens Road, Long Trail, Aikens Trail, Wachusett Trail, and some of Oval Trail, Caesar Trail, and Loop Trail. He returns to his vehicle and drives to Starting Position #2, which takes about 5.556 minutes. He then follows the blue path, with directions shown by black arrows. This route begins going towards Arborvitae Trail. By the end of this route, the searcher has looked through all portions of Arborvitae Trail, Vista Trail, Foxtrot Trail, Swenson Trail, Rock Road, and some of Loop Trail. Once finished, he returns to his vehicle and exits the park through the main entrance, along the red highlighted road. His exit takes about 2.64 minutes. This model results in an 82.091% chance of finding the object, because the routes mentioned above have a total weight of 12.683, which is 82.091% of the total weight. The total walking distance is 7.044 miles. Given that he walks at 4 mph, the walking time is 105.667 minutes. Adding the commute times gives 118.978 minutes, which is under the 2-hour limit.

SMALL OBJECTS MODEL # 1 CRITIQUE

PROS

- This model is based on a successful Internet ranking sys tem, so it should give a reasonable approximation for the popularity of paths.
- Relative to our other models, the numbers and vari ables are easy to calculate.

CONS

- This model does not take into account the crossing of roads.
- This model is only an approximation for trail popularity, without actual traffic data.

SMALL OBJECTS MODEL # 2

ASSUMPTIONS SPECIFIC FOR SMALL OBJECT MODEL #2

1. The searcher cannot walk on paved roads, except to cross them.
2. The searcher has a vehicle that, when driven, maintains a constant speed of 15 mph.
3. The searcher must leave the park two hours after entering.
4. The searcher always enters and exits the park through the main entrance.
5. The object is on a path.
6. The number of intersections between the segment's nearest parking lot and the segment is inversely pro portional to the probability that the object is in that segment.

SMALL OBJECTS MODEL # 2 EQUATION AND CALCULATION:

$$P_{k_i} = \sum_{s=1}^{z}\left[\frac{l_s}{S_s\sum_{i=1}^{n}\frac{l_i}{s_i}}\right]$$

P_{k_2} is the probability of finding the missing object for any given k

z is the total number of segments accounted for in the search

l_s is the length of the s^{th} segment of k

S_s is the number of intersections between the nearest parking lot of the s^{th} segment and the s^{th} segment.

li is the length of the i^{th} segment of the map

Si is the total number of interconnections that occur between the nearest parking lot to the i^{th} segment and the i^{th} segment

EQUATION DERIVATION

While some joggers run through an entire park, most shy away from long excursions. As the s^{th} segment gets farther away, the probability that such a person is there decreases. This distance can be approximated by S_s, or the minimum number of path intersections before the s^{th} path from the nearest parking lot. So, the probability that the s^{th} segment contains the ring decreases as S_s increases. Thus, we weight a given path with $1/S_s$.

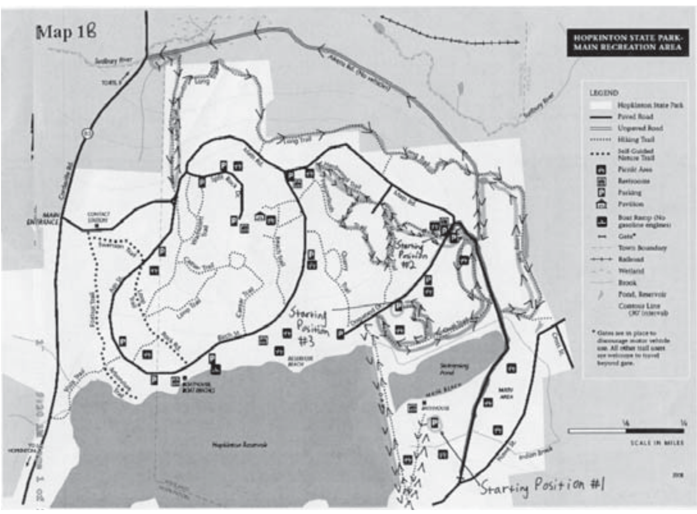
So the probability that the s^{th} segment contains the ring is proportional to the segment's length and inversely proportional to S_s and the sum of the weighted lengths of all segments. Thus, the probability of the ring being found on any search path K is equal to the sum of the previous probabilities for any given sub-segment of K.

MAXIMIZING PROBABILITY OF FINDING LOST OBJECT

STEP 1

$S_s\sum_{i=1}^{n}\frac{l_i}{s_i}$ is calculated by taking the number of intersections with a paved road that a segment has, and dividing it by the number of intersections it makes with a paved road.

This calculates the segment's distance from a parking lot. This was done for all segments in all paths used for walking. This calculation gave the weight of each specific segment. In this model, the "weight" is a numerical scale that gives the chance of finding the object on a given segment. Once the weight of each segment was found, the weights of all seventeen paths were found by taking the sum of the segments in each path. Adding these together, the total weight of map 1B is 5.45.



STEP 2: DEVELOPING THE BEST PATH

After finding the sum of the weights of every segment, we needed a way to search the park that gives for the greatest probability of finding the object. Our search method allows for visiting the trails with the highest total weight, which increases the probability of the object being found, given our formula.

Given that the trails with the highest weights are on the upper right part of the map, we developed a route that allows the searcher to go through these trails. This method begins on Starting Position #1 (map 1B) and continues down the yellow highlighted path, in the direction of the black arrows. The searcher then travels to Starting Position #2 by car, and follows the blue path. Lastly, the searcher drives to Starting Position #3, follows the purple path, and then exits the park. This route took 119.374 minutes and gave a 58.716% probability of finding the object.

SMALL OBJECTS MODEL # 2 CRITIQUE

PROS

- This model considers the place at which the searcher begins his hike, thus making it more likely that he will find the object.
- The numbers used in this model are fairly easy to calculate.
- The maximum time limit is not exceeded.
- A large portion of the map is searched, and over half of the map's total weight is searched.

CONS

- This model does not account for the distance or time it takes to cross the road on foot.
- This model is an approximation, and does not use the actual measurements of the distance of a segment to a parking lot.

SMALL OBJECTS MODEL # 3 (FOR FUTURE MODELS)

MODEL 3 DERIVATION AND EXPLANATION

While the first two models both give reasonably good probabilities for finding the object, perhaps taking into account both of their weights would improve effectiveness. Thus, both the number of connections to the segment's trail and the number of intersections that occur before the segment from the nearest parking spot are taken into account. The new weight, then, is equal to the division of the number of connected paths by the number of intersections before the segment. This weight is multiplied by the length of the segment and divided by the number of intersections occurring before the segment. This, when divided by the sum of these values for every segment on the map gives the probability that the ring will be found in that segment. Then, the probability of finding the ring on any given path *K* is equal to the sum of these individual probabilities such that each segment is a member of *K*. In variable from, we have:

$$P_{k_3} = \frac{\sum_{s=1}^{z} \frac{l_s C_s}{S_s}}{\sum_{i=1}^{n} \frac{l_i C_i}{S_i}}$$

This model may or may not give a better probability of finding the object, but it should give a more accurate probability that the searcher will actually find the object. We can infer from the previous two models that the optimum path according to this third model would incorporate the similar paths of the first two. This combination should yield more accurate information about the popularity of separate paths and thus the chance that he would have dropped the ring along any one of these paths.

MISSING RUNNER MODEL ASSUMPTIONS

1. The jogger understands that he is lost and tries to go to a parking lot or he is unconscious.
2. The jogger is lost on the path and not in the woods.
3. The jogger could be anywhere in the park, according to a distribution, *D*.
4. The jogger's initial lost position is not at an intersec tion, but rather on a path.
5. At the start of the search, the jogger has a 1/3 probabil ity of being unconscious and a 2/3 probability of mov ing.
6. The jogger is moving at a constant velocity.
7. After two hours of searching for the jogger, we will call 911.

RUNNER MODEL DERIVATION, EXPLANATION, AND TESTING

There are a couple of ways to predict a runner's location. In general, the greater the length of a trail segment, the greater the possibility of discovery becomes. If we take the length, L_s, of a segment and divide it by the length of all trails, we get the probability that the runner will be found on that segment *if* the runner would be lost on larger roads. However, this is the opposite of our initial assumption. So, we adjust this formula to give greater weights to smaller roads. Given the highest initial probability of any given segment, we subtract the segment's probability. This, then, gives an unadjusted probability, P_t, that the runner will be found on the t^{th} trail. Note that under this scenario, there is a 0% chance that the runner will be lost on the longest trail. However, there still is the issue that these numbers do not sum to 1. This can be accomplished with this equation:

$$a\sum_{s=1}^{n} P_s = 1,$$

where a is a scalar quantity by which we must multiply each P_t and n is the number of segments in the park. This final probability, $a^* P_t$, is the probability that the initial position of the runner is at the t^{th} segment. As there are a total of 112.09 miles of trails, it is necessary to look over a small portion of the trails. Given that we only have two hours to search, we can cover at maximum 8 miles by walking. This, however, is reduced as we drive to various drop-off points. Thus, we want to look at as few groups of trails as possible to minimize travel time. We then introduce the idea of clusters. Clusters are groups of segments that are in the same general area and have large probability densities (P_t/mile). The probability of the runner's initial position being in one of these clusters is equal to the sum of each of the P_ts such that the t^{th} segment is in one of the clusters. This, then, would be a simple model if not for the fact that the runner can move. Given a perfectly straight path with no curvature attached on either end by another road, the runner, if conscious, would have a 100% chance of reaching one of the roads. However, as the pathways become more convoluted and intertwined, the lower the chance that the runner escapes the cluster. We can approximate this notion of curvature by dividing the total length of all of the paths in the cluster, including its exterior paths (the set of paths that surround the interior portion) by the length of its exterior paths. In general, the greater this ratio, the greater the approximation of curvature, and so the probability that the runner will leave decreases. Given the original assumption that if the runner is in the middle of a completely straight road, s, he would have to travel a distance of $L_s/2$ to leave, and he would have a 100% chance of leaving. He would have a high chance of leaving the cluster after traveling the product of $L_s/2$ and the approximation of curvature, or:

$$\frac{L_I}{2}\left[\frac{L_I}{L_C}\right],$$

where L_1 is the length of the interior of the cluster, including its circumscriptive length, and L_C is the length of the circumscriptive path. Thus, if the jogger moves at 4 miles per hour, it takes him $\frac{L_I}{8}\left[\frac{L_I}{L_C}\right]$ hours to leave. Once this time is reached for a given cluster, the initial position probabilities would change. At this point, the probabilities become a function of time.

FINDING MISSING RUNNER MODEL CRITIQUE

PROS

- The model limits the large area to a few subsections and then restricts the subsections to an optimized search path
- It takes into account an approximated notion of curva ture in finding the amount of time it takes the runner to leave his position

CONS

- Only takes into account the notion of length and deter mining where the runner would have gotten lost
- It doesn't take into account the psychological decisions of a person who is lost

CONCLUSION

Although each of our models approximates the possible locations of the missing person or ring, the first model gives the best probability of finding the object. However, this does not necessarily imply that this is the best model, as the actual probability of finding the object may be lower. In this case, though the second model has a lower probability, the result may more accurately resemble reality. In either case, the third model combines these two and should, when tested, give a more accurate probabilities.

While there is only one model for the runner, this model successfully restricts the area that needs to be searched and allows for an optimized search path. Though this produces a low probability, it is likely that this resembles reality, as the total area is large. This model illustrates the sheer improbability of finding a missing person using a penlight in such a large park. Further traffic analysis of each path should confirm both of these models to have a certain degree of accuracy.