HIGH SCHOOL MATHEMATICAL CONTEST IN MODELING
OUTSTANDING PAPERS

HiMCM

November 2012

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 fifteenth 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 2012 HiMCM D-ROM, which is available from COMAP.

HiMCM Director's Article

William P. Fox

Department of Defense Analysis
Naval Postgraduate School
Monterey, CA 93943
wpfox@nps.edu

We should all celebrate since the High School Mathematical Contest in Modeling (HiMCM) completed its fifteenth year. It is and continues to be a fantastic endeavor for students, advisors, schools, and judges. We have even had schools ask for speakers to come in and discuss the modeling process so that their teams can improve and compete. The mathematical 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. Figure 1 is a plot of the growth over time. The trend over the last few years has been an exponential increase.

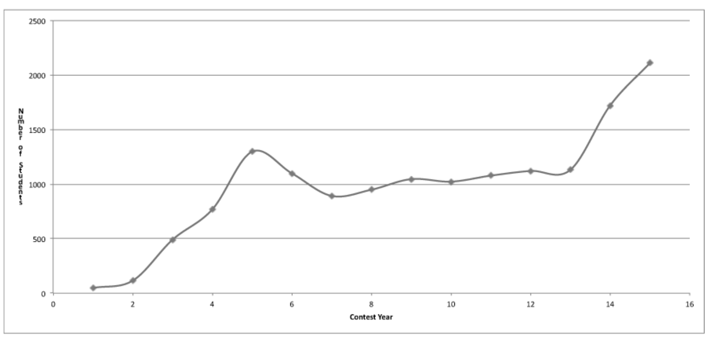


Figure 1: Number of HiMCM participants vs. contest year

This year the contest had 552 teams. This represents an increase of about 23% over last year. We had 2111 students from 23 states representing 58 schools and 5 foreign countries. We had 301 U.S. teams and 251 foreign teams, representing 13% and 49% growth, respectively. In the United States, these teams represented 58 schools. China represented about 93% of the foreign entries. Of the 2111 students, 777 or almost 37% were female students. The breakdown was 777 female, 1315 male students, and 19 unspecified genders. Since the beginning we have had 14,842 total participants, of which 36.34% have been female. We feel this is remarkable and that we hope they all continue on to some STEM education.

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 552 registered teams, 518 submitted solutions. These were broken down as follows: 213 doing problem A and 205 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.

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 2012. 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 Finalist, Meritorious, Honorable Mention, and Successful Participant. All finalist papers from the Regional competition were sent to the National Judging in San Diego. The national judging, consisting of eight judges from academia (high school and college) and industry, chose the "*best of the best*" as National Outstanding. We usually discuss between 8-10 papers in the final round so all these papers were awarded "National Finalist." 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	National Finalist	Finalist	Meritorious	Honorable Mention	Participant	Total
A	3	7	22	43	67	71	213
%	1%	3%	10%	20%	32%	34%	
B	5	3	25	68	97	107	305
%	2%	1%	8%	22%	32%	35%	
Total	8	10	47	111	164	178	518
%	1%	2%	10%	21%	32%	34%	

National Outstanding Teams

Hanover High School, Hanover, NH
High Technology High School, Lincroft, NJ
Hong Kong International School, Hong Kong
Illinois Mathematics and Science Academy, Aurora, IL
Maggie L. Walker Governor's School, Richmond, VA
Shanghai High School International Div., Shanghai, China
Shenzhen Middle School, Shenzhen, China
University High School, Irvine, CA

National Finalist Teams

Chesterfield County Math/Sci HS @ Clover Hill, Midlothian, VA (2 Teams)
China Welfare Institution, Shanghai, China (2 Teams)
Hanyoung Foreign Language High School, Seoul, Korea
Hong Kong International School, Hong Kong
Liaoning Province Shiyan High School, Shenyang, China
Maggie L. Walker Governor's School, Richmond, VA
Shenzhen Middle School, Shenzhen, China
Woodbridge High School, Irvine, CA
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)

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 spelling and grammar checkers 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 usually do 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. Teams that merely present a model to use without development do not generally do well. 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. 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. Most papers did not use references within the paper, yet we saw many tables or graphs that obviously came from websites. 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 fact.

Facts from the 15th Annual Contest:

- A wide range of schools/teams competed including teams from Finland, Hong Kong, Korea, Guam and China.
- The 552 registered teams from U.S. and International institutions represent a 22.66%% increase in participation.
- There were 2111 student participants, 1334 (63.2%) male & unspecified and 777 (36.8%) female.
- Schools from twenty-three 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 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.
The final teams in discussion will be designated as national Finalist.
National Outstanding Winner will be designated as Outstanding Winner.

Contest Dates: Mark your calendars early: the next HiMCM will be held in November 2013. Registrations are due in October 2013. Teams will have a consecutive 36-hour block within the contest window to complete the problem and electronically submit a solution. 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.

Judges Commentary, 15th Annual Contest

Problem A: American Elk

Prior to the arrival of European colonization on the North American continent, the ecological bio-diversity was much richer than we currently know in the 21st Century. Prior to the colonization animals such as the American Bison (Bison bison), Eastern Elk (Cervus canadensis canadiensis), Eastern Cougar (Puma concolor couguar), and Wolf (Canis lupus) were commonly seen across the North American continent. However, with the colonization from the old world came old world prejudices and practices. Within two hundred and fifty years all of these species were eradicated from the Eastern United States or extinct. Over the course of the last century an effort has been made to halt the loss of the North American fauna with the creation of national parks and animal preservation habitats, and in the last half of 20th century work was being done to reintroduce that fauna back to its natural habitat. This process has been used with several different species; however, no species has been given as much attention in the Eastern United States as that of the American Elk.

Prior to the reintroduction of Elk back into Eastern United States the only attempt at having these animals in their natural Eastern habitat was done by the owners of private exotic hunting preserves. However, in the latter half of the 20th century serious studies were conducted on the impact of a reintroduction program. The problems listed for this reintroduction program were as follows:

- What states to reintroduce the elk
- What would be the impact on agriculture
- Would the Elk adapt to the more densely populated Eastern U.S.

However, the most pressing question was the elk species themselves. The Elk native to the Eastern U.S. (Cervus canadensis canadiensis) was hunted to extinction sometime in the early 1800s. The Eastern Elk, while similar to the Manitoba Elk (Cervus canadensis manitobensis), are not the same. The Manitoba Elk were smaller in size than the Eastern subspecies and were adapted to living in the Western U.S and Canadian prairie. These adaptations included disease tolerances, foods, and environmental differences. How would these adaptations affect an introduction of Manitoba Elk into the East and more specifically in the Great Smokey Mountain National Park (GSMNP)?

Build a mathematical model to determine whether the elk survive or die out. Regardless, come out up with a plan to improve the growth of the elk over time. In addition to your one page summary sheet and complete project report prepare a one page letter describing your results to the Commissioner of the Department of Wildlife. The following are the reported numbers over the course of the study:

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Reintroduced	25	27					5				
Births	4	2	10	8	10	13	19	19	19	25	19
Death											
Poached				1					1	1	
Sickness	1	5	2	6	2			3	2	3	
Accident			1			1	1	1		2	
Predator	1		1	2	5	1	4		2		
Unknown				3				5	3	2	3
Population	27	51	57	53	56	67	86	96	107	124	140

Figure 1: This is the numeric data for the population growth and the corollary effects on the growth, culminating into the current approximate population of the herd in the GSMNP region.

This is an approximation. Some species were killed off faster than others, with the Eastern Elk being killed to extinction in the early 1800s and the Eastern Cougar being recorded as being killed off as late as 1930s. Ref: http://www.huffingtonpost.com/2011/03/02/eastern-cougar-extinct-mo_n_830181.html and http://www.rmef.org/Conservation/HowWeConserve/Restoration/FallRise/

Judge's Commentary

Professor William P. Fox, Naval Postgraduate School

Problem Author: Professor William P. Fox, Naval Postgraduate School

First, the problem statement explicitly called for a model to consider the Great Smokey Mountains National Park, and many teams failed to provide analysis of this specific region. Addressing additional regions increased the chance of recognition.

Many teams began directly with a logistic model. Judges expected teams to build or discuss the analysis that leads to a logistic model if they were using one. Building the model is a key component. Many teams began with a linear model and then after seeing the population goes to infinity modified the model to level at some form of a carrying capacity. Some teams found a model online: The VORTEX model. Use of models like this are allowed if the teams find some "value added" that they can add to the existing model. Simply extracting a model is of little merit in a contest such as this.

Teams often calculated results to seven to ten decimal places to model elk. At some point, the number of elk needed to be rounded to an integer.

Some papers ignored the life expectancy of elk, which judges felt could be critical to a *useful* model. It was interesting seeing all the methods for estimating carrying capacity used by teams. Just dividing the total area by the number of square feet an elk needs or may take up may be a first approximation may not be a reasonable carrying capacity—we would expect the capacity to be less than that.

This year's papers had many strengths. Almost all the papers did a reasonable job of estimating the growth of elk, but few discussed issues of weather, predators, and adequate space.

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

We provide an example of regression that shows why examining R^2 may not be the best indicator.

Consider the following four 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

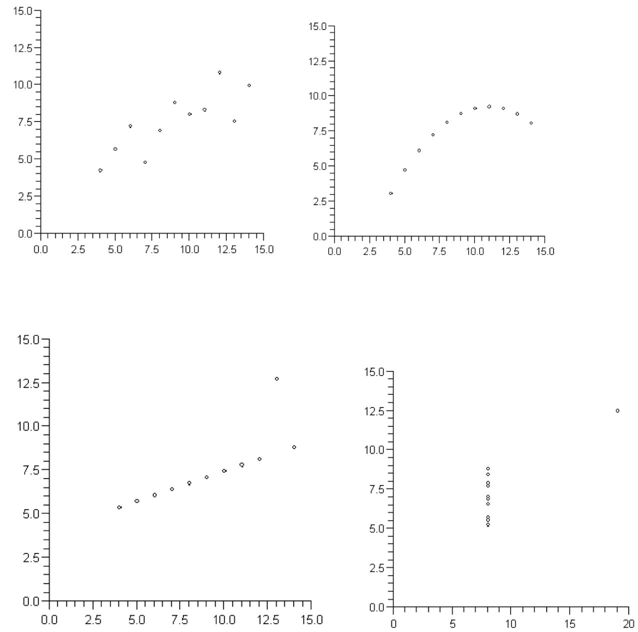
Suppose we fit the model $y = ax + b$ to each data set using the least-squares criterion. In each case the following model results:

$$y = 3 + 0.5x$$

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.)

Problem B: How Much Gas Should I Buy This Week?

Gas prices fluctuate significantly from week to week. Consumers would like to know whether to fill up the tank (gas price is likely to go up in the coming week) or buy a half tank (gas price is likely to go down in the coming week).

Consider the following cases:

- Consumer drives 100 miles per week
- Consumer drives 200 miles per week

Assume:

- Gas tank holds 16 gallons and average mileage is 25 miles/gallon => 400 miles/tank
- Consumer buys gas once a week

Therefore, the consumer can drive for 2 weeks or 1 week on half a tank of gas for cases (1) and (2) respectively. Thus the choice each week is whether to buy a full or half tank of gas or no gas.

Use the weekly gas price data available by region, state and city at: http://www.eia.gov/dnav/pet/pet_pri_gnd_a_epmr_pte_dpgal_w.htm

You can also use weekly crude price data http://www.eia.gov/dnav/pet/pet_pri_spt_s1_w.htm and any other publicly available data such as weather data, economic data, world events, etc.

- Develop a model that a consumer could use each week to determine how much gas – full tank or half tank – to purchase.
- Use the 2011 data to build/train your model and the 2012 data to test and validate your model.
- Is there an upper bound on "mileage driven" that changes the decision for buying weekly gasoline?
- Develop models for the following at least one large US city such as: Boston, Chicago, Denver, Dallas, Los Angeles, Houston, New York, or San Francisco.

In addition to the one page summary and complete project report prepare a short one page non-technical letter to your local paper describing how the average person might use your model.

Judge Comments:
William P. Fox, HiMCM Contest Director

Problem Author: Veena Menderitta, Lucent Technology

The judges were surprised with the amount of time and effort put into modeling the 2011 data. Many teams never got beyond this to move on the key modeling issues and questions posed in the contest. How accurately do we model money when it comes down to only 2 decimal places anyway? The number of decimals places used in modeling the price of crude and the cost of gasoline was entirely too many for most teams.

Often teams built many different models to compare crude to gasoline costs and most never choose one to use. Offering up many choices without picking one to use is not helpful in modeling.

Modeling what the driver should do took on many forms. Most teams that attempted this developed an algorithm. Most U. S. drivers are not as math literate as the team members. It is better to give a simple "rule of thumb" that a driver can easily use than to provide a complex algorithm with strange looking numbers. The results should have addressed the dollar savings that a driver would have using the algorithm.

Graphs were also an issue. Papers were judged in hard copy, not electronically. Thus, references to color in a graph were not helpful. Teams should restrict their graphs to various gray-scales and perhaps differing markings (line, dashes, dots, etc.). Legends and axes should be labeled and should make sense.

Bounding mileage: upper bounds for driving that affect the decision process were not given by most teams. This was a requirement of the problem and judges were unsure why most teams did not mention it.

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 appear 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. With the randomness of the cost of gasoline this is a critical element.

General Comments from Judges:

Assumptions with Justifications: All assumptions should have some justification of their importance in the modeling process. Assumptions should not be listed that do not impact on the model used or developed.

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

Computer generated solutions: Many papers used extensive computer 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 will not be read for papers that reach the final rounds unless the code is accompanied by an algorithm that is clearly and logically explained. An algorithm, if code is used, 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. It is requested that teams use some in-line reference to graphs, figures, and ideas taken from other sources.

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 Paper: University High School

Team Members: Thomas Gui, Sarah Sukardi, Ann Tao, Wendy Wei

Advisor: Stephanie Chang

Introduction

North America was once home to one of the world's most diverse ecosystems. Many animals such as the elk that once roamed the eastern U.S. disappeared after European colonization because of excessive hunting and loss of habitat. In the early 1900s, after elk had nearly disappeared from several eastern states, hunting and conservation groups begin to advocate for their protection.

Today, elk populations remain low in the U. S. Efforts to introduce the Manitoba elk in areas once home to the extinct Eastern elk are currently being considered. Despite the honorable intentions, simply introducing Manitoba elk could end in failure. A Manitoba elk is not only smaller than an Eastern elk, but also adapted to different diseases, foods, and ecosystems. In this paper we address the plausibility of sustaining the Manitoba elk in the Great Smoky Mountains National Park (GSMNP) through a mathematical model of population growth and sustainability.

Interpretation / Problem Restatement

We use our model to address issues encountered by previous reintroduction programs: where to reintroduce elk, the elk's impact on agriculture, and whether elk would adapt to the more densely populated Eastern U. S. (Pt. 1). In addition, we determine the survival success or failure of Manitoba elk in the GSMNP (Pt. 2). We address factors that may affect elk survival and incorporate data provided from the study (Pt. 3).

Assumptions/Justifications

Assumption 1: The Eastern U. S. is defined as the Appalachian Mountains and eastward (not including Florida).
Justification: Although the Eastern U. S. is often defined as east of the Mississippi River, we found that the areas east of the Appalachian Mountains had the most homogenous geography and that a larger area would contain too many geographical factors.

Assumption 2: The geographical regions where we introduce elk have enough resources to sustain the elk, at least to some carrying capacity.
Justification: Manitoba elk may be adapted to differing environments, but in order to identify where their integration will succeed we must determine a *K*-value for our logistic model obtained from either our own calculations or from other researchers.

Assumption 3: There will be no significant emigration or immigration of elk.
Justification: Migratory patterns are near impossible to model without field tracking.

Assumption 4: Differences between individuals such as sex, age, breeding status, state of health, etc. are ignored.
Justification: Individual elk are so varied that we found it impossible to incorporate every factor.

Assumption 5: Population cycles are disregarded.
Justification: Fluctuations in population throughout the year cannot affect the model because it is based on the population given at the time.

The Mathematical Models Pt. 1

A. Where to introduce the elk?

In order to determine locations for elk introduction, we decided logistic models should be based on ecoregions rather than state lines. Thus, we used a density-dependent (logistic) model for 3 ecoregions: North American Desert, Great Plains, and Eastern Temperate Forests. In calculating variables for a logistic model, however, we must calculate individual *K*-values for carrying capacity, as each ecoregion can sustain differing numers of elk. Variations per ecoregion, however, make it difficult to find a value for *K* given our inability to do field research. Even if we did attempt to calculate carrying capacity, this too would be implausible without data from previous studies. For these reasons, we have deemed where to place the elk as inconclusive.

B. What would be the impact on agriculture?

When elk herds expand their foraging to include agricultural products, we find that there are proportionate losses in alfalfa, wheat, and sunflower fields. As a result, farmer resentment tends to lead to poaching, and elk numbers fall.

C. Would the Elk adapt to the more densely populated Eastern U.S.?

The Manitoba elk's new environment in the eastern U.S. has a higher population density than in its original environment. Thus, carrying capacity in the east is slightly lower than in the west. This slight difference is insignificant because, by analyzing scientific data through our model, we found that the carrying capacity in eastern ecosystems is sufficient for elk survival.

The Mathematical Models Pt. 2
Density-Dependent (Logistic) Model

$$P(t)=\frac{KP_0e^{rt}}{K+P_0\left(e^{rt}-1\right)}\text{, where }\lim_{t\to\infty}P(t)=K\text{ .}$$

$P(t)$ = change in population over period of time
P_0 = number of individual elk in initial population
r = biotic potential or reproductive capacity of individual (probability of reproduction)
t = number of years passed since introduction of initial population
K = carrying capacity

Carrying Capacity Equation

$$K=\frac{50(A-C)}{a}$$

A = total available acreage
C = acres taken by native competitors = minimum acreage per individual × population
a = minimum acreage required per elk herd

We derived a carrying capacity formula based on amount of resources in a certain amount of land. In general, there are two ways to calculate carrying capacity: 1. Amount of biomass in an environment, and 2. Minimum acreage for survival of individual elk. Since calculating biomass of the environment would have to account for the varying biomass of each food source based on geographic density, we found it easier to create our equation based on the minimum acreage necessary for survival. After some research, we discovered the minimum acreage necessary for elk in the GSMNP: 4500 acres per elk. (For more on application of this equation to the GSMNP, see K-value below under specific variables/parameters for GSMNP.)

Analysis using Actual Data
Specific variables/parameters for GSMNP

P_0 = number of elk initially introduced into the park since 2001 = 140
r = 0.0924 (biotic potential or reproductive capacity of individual)

Assuming a 23:77 ratio of males to females and an average 12% of the females pregnant at any time, we calculate a probability of .0924 that any female elk (for purposes of simplification) would successfully bear offspring. (0.77 × 0.12= 0.0924).

The carrying capacity of the GSMNP is about 104 herds. This was found using a minimum of 4,500 acres per herd, as the park is 522,419 acres. We assume that competition (i.e., from white-tailed deer) takes up to 250 acres per adult and there are 500 deer. [522419 – (250 × 500)]/4500 = 88.315 herds × 50 elk per herd. Thus the K-value or maximum carrying capacity of GSMNP is about 4500 Manitoba elk.

t = number of years. t is the independent variable (x-axis) in this equation, since we are seeing the effects of the environment on the population over time (Figure 1).

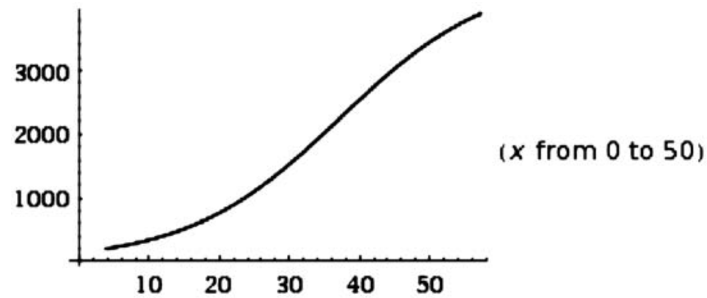


Figure 1: *P*(*t*) = (630000e^(0.0924*t*))/(4500+140 (-1+*e*^(0.0924*t*)), where *x*-axis represents t in increments of years and the *y*-axis represents the total number of Manitoba elks in the population

Proof of survival success
Since carrying capacity is more than initial population, the slope of the population function is positive. In application, the population's net change is constantly increasing. By using this model, we can conclude that the current population of 140 elk in the GSMNP will survive and eventually reach carrying capacity. We can further show that though the death rate may decrease the population, birth rate (b) is greater than death rate (d) so the difference between the two is ($bN - dN$ = positive value), where N = number of elk in the population at the time.

Number of Births
A typical adult elk population has a 23:77 ratio of males to females. Of the 77% females, an average 12% are pregnant at a given time. Thus, of the total population, 9.24% (0.77 × 0.12 = 0.0924) will bear offspring. Success, however, is found by multiplying birth rate by population by probability of calf survival. This gives our equation for births, $Y = (0.0924 \times N \times 0.85) = 0.07854N$.

Number of Deaths
We found that deaths from predation are 66% of all deaths, and used this ratio to calculate total number of deaths. Since 8% of a population is calves and their death rate is 15%, we can conclude that $0.012N$ represents the rate of calf death ($0.08N \times 0.15 / 0.66 = 0.01818N$). Adult elk, which are 92% of the population, have an average mortality rate of 4.5%. We can thus calculate that $0.0598N$ represents the mortality rate of adults. ($0.92N \times 0.03 / 0.66 = 0.041818N$). For the total mortality rate, we combine rates for calves and adults to get $0.43636N$. Thus, there is a gradual increase of 3.4904% per year.

From this positive growth rate, we conclude the elk population in the GSMNP will continue to rise and in some 50 years be near carrying capacity. (By 2060, there are about 4180 elk, 320 short of capacity.)

Reinforcement Measures: The key to maintaining the GSMNP elk population lies in prevention of what drove the Eastern elk to extinction: poaching. We can either eliminate elk predators from the region by relocating them or eliminate poaching. Though the GSMNP prohibits poaching, elk found outside the Park can be legally shot. Because the GSMNP cannot expand its boundaries wherever elk roam, we propose non-lethal physical barriers to keep elk in park and social media to increase public awareness while the elk are in their critical growth stage in the first few years. In addition, we acknowledge these possibilities for survival enhancement: increasing foraging supply by cultivating plants as food sources or decreasing competing species such as white-tailed deer.

These courses of action should be implemented only if either a drastic decrease in elk population is confirmed (not a mere fluctuation) or a prolonged slowing of population growth while the population is still far from carrying capacity is recorded. Doing this requires tracking some elk via radio transmitters and population surveys via helicopter.

Though we fully expect the elk to thrive, fluctuations will occur due to unaccounted factors such as weather and natural disaster. These should not play a vital role, and we thus expect the elk population to rebound quickly without fail.

The Mathematical Models Pt. 3

Endogenous variables: independent factors our study is designed to focus on.

1. *Change in population over a given period of time*
   Reasoning: Determining whether the elk population survives can only be done over a long period of time spanning perhaps multiple elk populations. Only after multiple cycles of birth and death will we be able to see a trend.

2. *Presence of other competitors for resources such as food availability.*
   Reasoning: The amount of food competitors consume is important in calculating carrying capacity since as the number of competitors increases, the number of elk the environment can sustain decreases.

3. *Presence of predators in contribution to mortality.*
   Reasoning: Since predation causes 2 out of 3 elk deaths, we consider this vital to our model: as the number of predators increases, so does the elk mortality rate.

Exogenous variables: independent factors our study acknowledges but does not take into account.

1. *Adaptations and innate characteristics of the Manitoba elk*
   Reasoning: The effects of these are insignificant. The more moderate climate of the eastern U. S. is not different enough to affect survival rate. In addition, the food available in the eastern U. S. is similar to that in native regions. Lastly, predators of the elk in their native region and in the eastern U. S. are similar.

2. *Discrepancies in abiotic factors of an environment including weather, soil, nutrients, etc.*
   Reasoning: if we were able to obtain data for each factor and given adequate time, we could account for each and its impact on the elk.

3. *Variations in type of food fit for consumption*
   Reasoning: The vegetation of the elk's native regions and the eastern U. S. are similar. In addition, the eastern U. S. offers a wide variety of plant material edible to the elk. We believe the elk would quickly adapt to a "new" diet similar to the old one.

4. *Possible disease transmission from domestic animals such as Bovine Tuberculosis (TB) or Chronic Wasting Disease (CWD)*
   Reasoning: TB is predominant on the west coast--it is near unheard of in the east. CWD is also negligible since previously recorded cases are not near the east coast. In addition, imported elk are assumed to have undergone veterinary inspection so as to prevent introduction of TB or CWD.

5. *Changes in population due to immigration and emigration*
   Reasoning: Lacking tools to track the movement of elk over a long period of time or existing data, we assumed that the elk in GSMNP are a closed population.

Sensitivity Testing
Our model's accuracy can be affected by outside factors that were not included. However, we are able to adjust for some factor changes. If the number of predators increases, the elk death rate can be adjusted accordingly. Similarly, if disease impacts the population, we can adjust the death rate. However, our model only addresses generalities of an ecosystem and is thus sensitive to individual events that affect the population significantly (i.e., an epidemic or a sudden increase in hunting.)

To assess our model's accuracy we compared our data to those from the first 10 years of elk introduction. Although we could not account for reintroduced elk, one can clearly

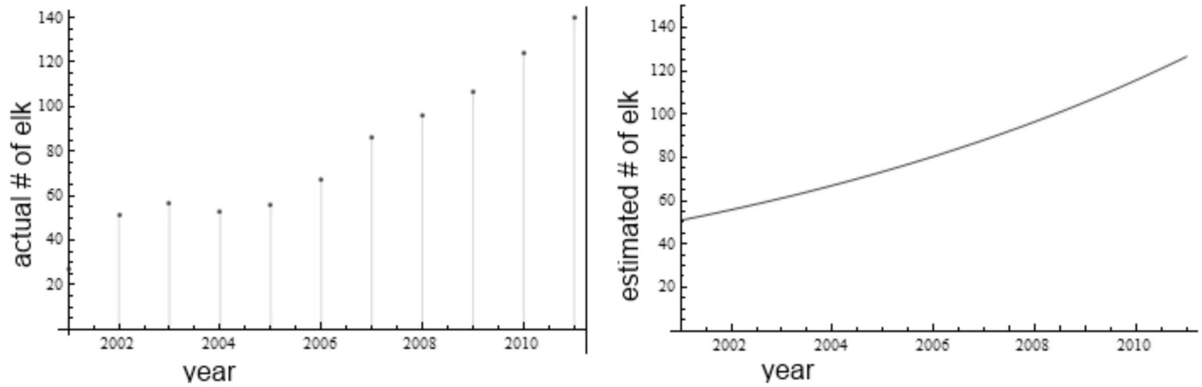


Figure 2: Model manipulated to start at 2001 in order to compare with data; the introduction of elk in 2001 and 2007 are not consider for the purposes of comparison.

see the consistency our model otherwise has with GSMNP data. (In the following simulations we changed carrying capacity to 5,200 in order to exemplify the sensitivity.)

If we start the model in 2001 (Figure 2), when most elk were reintroduced, it predicts about 110 in 2011--an underestimate. In fact, after about 2008, our model underestimates actual data, implying that the elk may do better than our model predicts.

However, this is only a small interval of actual data--if we had data for 2001 to 2060, we would have a much better idea of our model's error. Despite this drawback, we can see consistency between the data and our model--showing that our model may be viable for accurately predicting the population in question.

In our model we can adjust several factors that affect the population (Figure 3). In the following simulations we changed carrying capacity to 5,200 in order to show sensitivity.

Figure 3: Simulations with different initial populations. If we use the birth and death rates for the GSMNP, we see that these populations are stable and survivable.

When we consider an ecosystem with quadruple the area but half the resources of the GSMNP, our model works nicely. We simply adjust carrying capacity to 20,800 and then decrease it to 10,400 to account for fewer resources. In addition, if this ecosystem has say, 4,000 elk, we can input this into our model to see population growth (Figure 4).

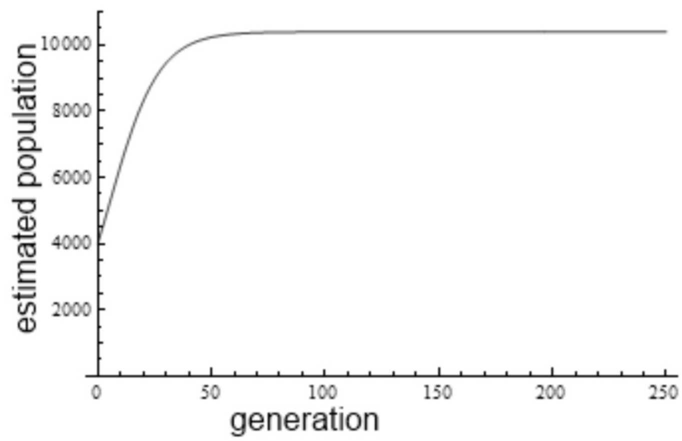
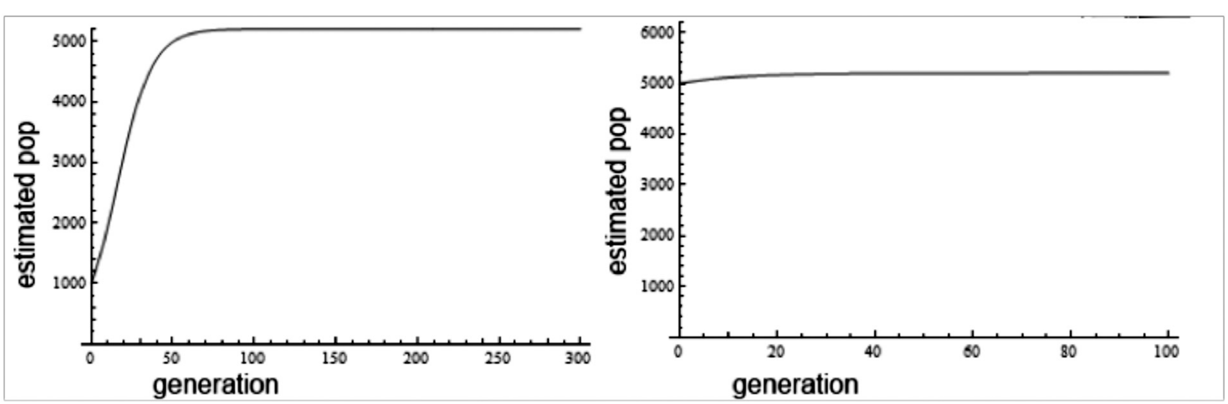


Figure 4: An ecosystem with quadruple area, but half the resources of the GSMNP.

We also can further address the differences in the ecosystem by adjusting variables such as birth rate or death rate. If this ecosystem has significantly more predators, our model can account for that by adjusting the death rate.



(Initial population 1000) (Initial population 4000)

Strengths of Our Model

1. The logistic model can calculate population growth in a wide variety of ecosystems to a high degree of accuracy.
2. Our model is highly adaptable in terms of the birth rate, death rate, and starting population. If there is an epidemic or a high level of predation, we simply adjust our numbers to model the new situation accurately. In addi tion, we cannot only adjust carrying capacity but also adjust starting population in order to better account for different ecosystems.

Weaknesses of Our Model

1. We do not take into account the possibility of emigration or migration.
2. The variability between individuals is not accounted for.
3. The model is sensitive to sudden introduction of factors that affect population.
4. Introduction of new elk following the initial group can not be accommodated in our model.
5. Our calculations for carrying capacity of ecosystems are estimated.

Conclusion

We found that the current elk population in the GSMNP is sustainable without introduction of more elk. The population will increase exponentially and taper off when it nears carrying capacity. Thus, our plan to increase the population is to allow the elk to continue to propagate naturally. To ensure that the elk do not experience new threats, we propose reinforcement measures, including prohibiting poaching and monitoring when the population is at risk so that park ecologists can alter resources or reduce interspecies competition.

In addition, we found that concerns regarding adaptation of the elk to be insignificant to the end goal of achieving a stable population. Changes in diet, climate, and geography are not large, so we believe that the Manitoba species would not have problems populating areas in the eastern U. S. The majority of the region is rural, and elk would avoid urban areas.

In the future, we wish to model elk populations outside of the GSMNP. With that said, we feel that we could use our model to implement a more widespread and comprehensive elk reintroduction plan on the east coast.

Also, further investigation on our model could include using more nuanced methodology in developing carrying capacity, births, and deaths of a population. Doing so would produce a more accurate and realistic model.

Problem B: Paper: Hanover High School

Team Members: Katherine Chen, William DeLucia, Melanie Subbiah, George Voight

Advisor: Greta Mills

Introduction

Problem Restatement

The problem we faced was how to predict gas prices so that consumers could decide when to buy gas. We wanted our model to be user-friendly while maintaining a high level of accuracy. We looked at two drivers, one who drives 100 miles/week, and one who drives 200 miles/week. To reduce variables, we set a single car with a range of 400 miles/tank, which means that the driver will empty the tank in four weeks or two. Also, we assumed that a driver only buys a half tank, a full tank, or no gas, and buys once a week. Using these parameters, we needed to determine when a driver should buy gas based on whether prices trend up or down. Therefore, we needed to predict future gas prices by using factors that correlate with gas price.

Assumptions and Justifications

Customers only buy gas in half and full tank increments.

This assumption simplifies our model and was given in the problem statement. By restricting choices, the model will not vary greatly based on small changes in miles driven per week. This makes our model for 200 miles/week and 100 miles/week relevant with only minor modifications to drivers who drive between 100 and 200 miles a week. Indeed, the actual number of miles driven is less important than the ratio between amount of miles driven in a week and number of potential miles per tank. (In our case, the tank holds 16 gallons, and the car gets 25 mpg.)

The historical correlation of crude oil and gas prices will continue.

The assumption is necessary given the nature of our model. In using prior crude oil prices to model future gas prices, we need to assume that the correlation will continue. This assumption is grounded in the basis that since 2011, the method of refining oil into gasoline has not changed dramatically. Moreover, we will later test our model to make sure it applies to 2012.

It takes time for the gas market to respond to changes in crude oil prices.

Because of the time lag due to transportation refinement, it takes one to two weeks for gas prices to respond to changing oil prices.

People would like to minimize their expenditure on gas while maintaining their desired amount of driving.
A consumer decides whether to buy gas based on price. However, they do not want to run out of gas, so the idea is to buy the same total amount of gas and run your car at the same level of operation that it is now, but buy gas at optimal times.

The trend in the average of the Brent and WTI crude oil prices is representative of the trends in the price of all crude oil relevant to U.S. gasoline.
From the data given us, we assumed that the trends in the prices of Brent and WTI crude oil are representative of the trends for all sources of crude oil used in major U.S. cities. Moreover, when we compared the graphs of Brent and WTI, we found that these two had similar trends. Therefore, we felt justified in saying that crude oil price trends are fairly consistent across the industry and do not depend heavily on the source. We used an average of the Brent and WTI prices as our crude oil price.

The Model

Model Building

In order to determine how much gas to buy each week, we had to find a way to predict gas price. Our model predicts gas price based on the past weeks' crude oil prices. We discovered this correlation by investigating the relationships between current gas prices and factors that affect the gas price. We found that time of year, temperature, and economic stability had low correlation with gas price. In contrast, the price of crude oil was very closely related to gas price (Figure 1). Additionally, its price takes into account economic and natural factors because a recession or natural disaster affects the price of crude as well as the price of gas. By using recent crude prices, we can accurately predict the price of gas two weeks in the future. We knew that due to transporting and refining, there would be a lag between a change in crude price and a corresponding change in gas price. Next we needed to determine the amount of lag.

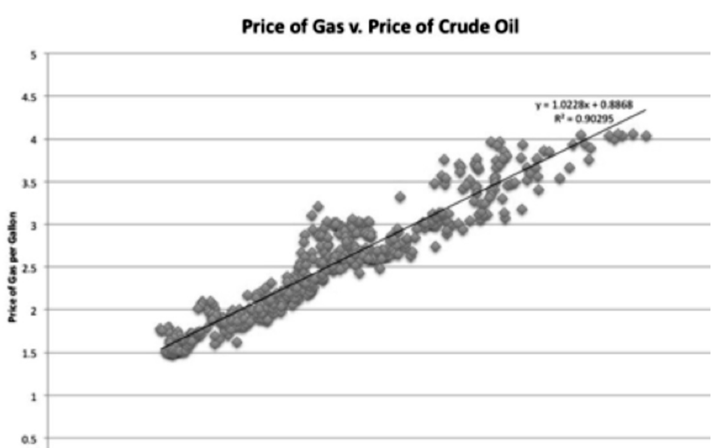


Figure 1: Price of Gas vs. Pride of Crude Oil

When we plotted crude price, shifted 1-6 weeks, against the un-shifted price of gas, we found the strongest correlation with a shift of 1 week. However, a 1-week shift only allowed us to predict 1 week into the future and we needed 2 weeks (for reasons explained later). Since a 2-week shift also had a strong correlation, we decided to use this relationship as the basis for our model. The recursive equation for new gas price then followed logically: $P_1 = (C_{-2} - C_{-1}) + P_0$ where P_1 is next week's gas price, $(C_{-2} - C_{-1})$ is change in crude oil price between one and two weeks ago, and P_0 is current gas price. We modeled this relationship in Excel and graphed our predicted values with the actual values. We found that since crude oil price had more noise than gas price, our predicted values also jumped around more, while still followed the correct overall trend.

We discovered we could smooth the graph by averaging. Instead of using actual crude oil price (C_0) for a given week, we averaged the values of the previous week (C_{-1}), the week itself (C_0), and the next week (C_1) to get a new value $C_a = (C_{-1} + C_0 + C_1)/3$). This reduced noise and allowed us to model gas prices with an average error of 1.63%.

We then moved on to the real question: how to decide when to buy gas. We made decision trees for 200 miles/week and 100 miles/week drivers (Figure 2). There are three options: buy none, buy a half tank, and buy a full tank. However, a driver can only buy a full tank if they have no gas. If they have no gas, they must buy some. Thus, a driver has at most two options.

Our model was then complete except for one detail. Since the 100-mile model had to predict gas prices two weeks in the future, it used crude oil prices from the current week. That meant that the averaged value for the current week needed a future crude price not available yet. We resolved this by using the actual crude price to predict two weeks out. Our one-week prediction used $P_1 = (C_{a-1} - C_{a-2}) + P_0$ and our two-week prediction used $P_2 = (C_0 - C_{a-1}) + P_1$. Another difference is that the one-week prediction is based on the difference in crude price added to the current gas price whereas the two-week prediction is based on the difference in crude oil price added to the predicted one-week gas price.

Model Testing

The first level of testing was to calculate average percent error between our predicted gas prices and actual prices. We first calculated percent error for data from Boston 2003-2011 (1.63%), and then for 2012 (0.88%). The small errors showed the accuracy of our model and that it could be applied to new data.

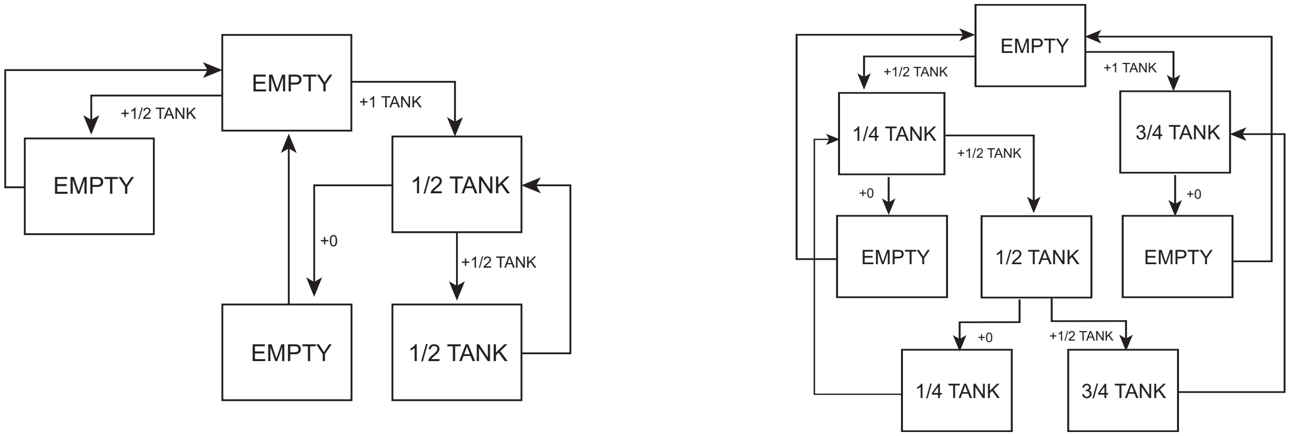


Figure 2: 100 mile/week (left) and 200 mile/week decision trees.

The next level of testing was comparison of how much a consumer would spend on gas with our model, how much they would spend if they knew future gas prices in advance, and how much they would spend if they did the opposite of what our model suggested. In all cases, our model was very close to the ideal. Our model also saved the consumer a significant amount of money over the worst case.

100 mile a week driver: Boston 2003-2011: 0.46% difference between model spending and omniscient spending

2.48 % saved from worst case spending

Boston 2012: 0.51 % difference between model spending and omniscient spending, 6.08 % saved from worst case spending

200 mile a week driver: Boston 2003-2011: 0.36% difference between model spending and omniscient spending

1.54% saved from worst case spending

Boston 2012: 0.26 % difference between model spending and omniscient spending, 3.60 % saved from worst case spending

The final testing was to apply our model to other cities: Houston, Los Angeles, San Francisco, New York, Chicago, and Denver. For almost all the cities, there was less than a 1% difference between our recommended spending and the ideal.

Reflections on the Model

Strengths

- **The cost of buying gas using our model is very similar to the most efficient mode of buying gas.** Since one of our goals was to create a model to predict future gas prices in an effort to reduce spending, and this model generates a buying plan similar to the ideal, we deem the model a success.
- **Our model can be used to predict prices throughout the U. S.** When testing our model on other cities, we saw that it predicts gas prices well.
- **Our model predicts gas prices with a 1.63% average error.** This precision let us accurately determine how much gas a driver should buy.
- **Our model is easy to use.** One only needs to enter crude oil prices from the last three weeks and current gas price, and our model predicts prices for the next two weeks and tells the user how much gas to buy and how much it will cost.

Areas for improvement

- **For some cities, our predicted values were slightly shifted from the actual values when graphed over a shorter time frame.** Longer transportation times and other factors may mean that the time lag between crude price and gas price varies among cities. With more time to investigate other cities, we could adjust our model for each city and correct this minor inaccuracy.
- **Separate models are needed for consumers who drive different amounts.** The more weeks a driver can go without buying gas, the more weeks they need to use in the calculation. While both of our models are accurate, it would be more convenient if a user could access both in one interface.
- **Our model does not accommodate drivers who do not fit the problem's profile.** If someone uses other than a quarter- or half-tank a week, the model does not apply.
- **If gas continues at a constant pace, there is no opportu nity for our model to save money.** The only opportuni ties to save money are when gas price reaches a local maximum or minimum, and these are the hardest points to predict. Past data repeatedly predicts continuation of a trend. Furthermore, the price of crude is more volatile than that of gas, often making the model noisier than reality. However, the model is accurate enough to save money. For example, if a 200 mile/week driver in Chicago used our model from the beginning of January 2012 to the end of October 2012 they would save about $52.50 over someone who filled up their tank at the worst times.

Conclusion

Results

Our model accurately tells a driver how much gas to buy weekly. Even with sudden changes in gas price (such as the recession of 2008, the largest price drop in our data), our model is consistent with real data. Tables 1 and 2 compare our model with one where the knowledge of future gas prices is omniscient and one where the most amount of money possible was spent on gas. Our model is consistently close to the ideal.

With minor modification, this model can be applied to weekly mileage between 100 and 200. Our model can also be applied to values below 100 miles, but we would need to predict gas prices further in advance, since a driver could drive longer on a single tank. If you drive more than 200 miles/week, the problem changes significantly. For example, suppose you travel 201 miles/week. The first week you must buy a full tank because if you put in a half tank, you could only travel 200 miles. The next week, you have enough gas left to go 199 miles so you need another half tank. Thus, people who drive more than 200 miles/week must refill each week. On the other hand, almost every

City	Total spent according to our model	Total spent ideally	Difference between predicted and ideal	% Difference predicted and ideal	Total spent with worst case scenario scenario	Saved (compared to worst case scenario)	% Saved (compared to worst case scenario)
Boston	$1203.15	$1200.11	$3.04	0.26%	1243.91	$40.76	3.28%
Houston	$1119.42	$1116.82	$2.60	0.23%	$1160.43	$41.02	3.54%
Los Angeles	$1332.53	$1323.10	$9.43	0.71%	$1384.06	$51.53	3.72%
San Francisco	$1323.20	$1315.13	$8.07	0.61%	$1374.14	$50.94	3.85%
New York	$1199.10	$1195.78	$3.32	0.28%	$1240.37	$41.27	3.33%
Chicago	$1259.82	$1252.46	$7.35	0.59%	$1312.40	$52.58	4.01%
Denver	$1133.24	$1128.38	$4.86	0.43%	$1171.55	$38.31	3.27%

Table 1: 1/2/12 - 10/22/12 for the 200 mile a week driver

City	Total spent according to our model	Total spent ideally	Difference between predicted and ideal	% Difference predicted and ideal	Total spent with worst case scenario scenario	Saved (compared to worst case scenario)	% Saved (compared to worst case scenario)
Boston	$587.38	$584.42	$2.96	0.51%	$622.23	$34.85	5.60%
Houston	$557.65	$555.84	$1.81	0.33%	$594.65	$37.00	6.22%
Los Angeles	$666.78	$660.01	$6.77	1.03%	$709.74	$42.96	6.05%
San Francisco	$659.42	$652.56	$6.86	1.05%	$702.64	$43.22	6.15%
New York	$599.25	$596.56	$2.69	0.45%	$635.15	$35.90	5.65%
Chicago	$627.54	$623.19	$4.35	0.70%	$668.55	$41.01	6.13%
Denver	$569.14	$564.84	$4.30	0.76%	$602.16	$33.02	5.48%

Table 2: 1/2/12 - 10/22/12 for the 100 mile a week driver

week, they would have some gas left, which precludes buying a full tank. Therefore, drivers would buy a half tank almost every week. They can only buy a full tank when they have nothing left going into the week. (For example, every 4th week someone who drove 250 miles/week could buy a full tank.) When they need more than half a tank, the driver's choices are so restricted that they have much less need for our model.

Extensions

With more time, we could create an Excel or Stella program that would take inputs of a particular car (mpg, miles per week, tank size) and describe the choices that the driver has in order to minimize spending. Similarly, if we had a larger pool of data, we could include more factors that are not already affecting crude oil.

We uncovered an interesting fact about the modeling system that we had created. Suppose we were a company buying gas in bulk to resell for profit. We realized that buying in bulk increases the amount one can save. This is true because of the way gas price fluctuates. There are peaks and valleys, and if a company were to buy all their gas in a valley, they would save a lot of money. We would have to expand our model to explore this option, since our prediction window does not extend more than a few weeks with reasonable accuracy.

Problem A Summary: Illinois Mathematics and Science Academy

Team Members: Zi-Ning Choo, Mary Do, Grace Li, Summer Wu

Advisor: Dr. Padmakar A. Patankar

In the winter of 2001, 25 Manitoba Elk, native to Canada and the Western United States, were reintroduced into the Eastern United States. However, because the elk were foreign inhabitants in this new region, there are many factors that would affect their survival in the new environment. Our task was to analyze the success of reintroducing the elk to the Great Smoky Mountains National Park and determine how to optimize the results. If the model predicts that the elk population will survive, we would have to consider how to improve the growth of the elk population over time, and if the model predicts that the elk population will die out, we would have to devise a plan to prevent it.

Our method consisted of a two-part solution: first, creating a model to predict the future of the elk population and determine whether or not the elk will survive in their new habitat, and second, proposing ways to improve the growth of the elk population based on that model. Using the data we were given, we formulated an equation to relate the birth rate and total death rate of a given year to the population size the previous year. The total death rate function was determined by adding up the equations for each of the different causes of death: predation, accidents, poaching, sickness, and unknown causes. The function for population growth was found by subtracting the total death equation from the birth equation. Population growth is the derivative of population with respect to time, so we were left with an unsolvable differential equation. Consequently, we used Euler's method and a step size of 1 year to derive a graph for population over time. We used a step size of 1 year because elk can only give birth once a year and the data was given in yearly increments.

From our population growth model, we determined that the elk would adjust well to their new surroundings and survive in their habitat. It projects that their population will increase until leveling off at a carrying capacity of about 1900 elk. From our model for death rate, we determined that the two causes that contributed most to the number of deaths--accidents/unknown and sickness--could be limited by 1) lowering the speed limit, 2) increasing driver awareness, 3) fencing off areas highly populated with disease-causing agents, and 4) administering the appropriate drugs as needed both to prevent and to treat diseases. In addition, to limit the harmful effects of inbreeding, we also recommend introducing five Manitoba elk from the West to the Great Smoky Mountains National Park population every ten years.

Problem A Summary: Shanghai High School International Division

Team Members: Yu Cheng Gu, Bill S Li, Michael Lingzhi Li, Andrew Siqi Tan

Advisor: Mingxin Zhang

The final goal of our model is to simulate the number of Manitoba elk at the end of a period of time and evaluate if the Manitoba elk can adapt to the environmental conditions in eastern America, more specifically the GSMNP (Great Smokey Mountains National Park). In going about this problem, we constructed a model that predicts the change in the elk population over time.

Our model calculated the change in population by first calculating the birth rate and death rate, then calculating the limit to the elk population in GSMNP: The carrying capacity.

We calculated the birth rate of elk by averaging the elk births in the given data to obtain the arithmetic mean, and calculated the death rate in a slightly more complicated manner of splitting it into two sub-models: An accidental death rate model and a natural death rate model. The accidental death rate model takes the arithmetic mean of the poaching rate, accident rate, predator attack rate, sickness rate, and an unknown factor rate based on the data given to obtain the probability of an elk dying from environmental factors in any given year. The natural death rate model, on the other hand, simply models the deaths of elk that reach their natural age limit.

We then estimated the carrying capacity of the elk population by rearranging the main function and calculating the arithmetic mean of all data inputted into it, and combined the above models to form our final elk population iterated model.

Next, we constructed a Monte Carlo Simulation Program to test our function. The program simulates the elk population's interaction with nature based on the functions derived in the iterated model for a large number of trials and outputs a statistical average of final population for comparison to our given data. We optimize the program by tweaking input rates and the sickness rate function in order to accurately portray the population growth function.

The original figures were derived from data for 2001- 2007. By applying the calculated rates to the Monte Carlo simulations, a stunning accuracy of elk was achieved. After optimization and modification of the constant rates into functions, the absolute uncertainty of elk was reduced to for any set of given statistics (including total poached, total killed by predators, total killed by accidents, total killed by sickness etc.)

The model predicts that the number of elk will increase steadily in the future until it peaks at a limit of 1101, then fluctuate and oscillate slightly around this limit.

Problem B Summary: Hong Kong International School

Team Members: Nicholas Chan, Alvin Luk, Joshua Silva, Edward Tian

Advisor: Kevin C Mansell

The problem here is to determine how much gas to buy on any given week. Essentially, we develop a model to predict gasoline prices based on fundamental variables and a recursive formula.

We are given two cases, one where the consumer drives 100 miles per week and one where the consumer drives 200 miles per week. If we were to have the option of adding as much gas we would like, then there would be a large number of choices and complicate the model to reduce accuracy. Instead, we deemed accuracy to be of more importance and limited the choices of the consumer to binary form every week.

We divided our solution into 5 parts:

Part I) Establishing Preliminaries

Our assumptions, definitions, variables, and general analysis; these are constantly referred in the model

Part II) The General Pricing Model

Here we determine quantitative factors as well as the laws of supply and demand to create a model that could predict future gas prices. We create theoretical models and then find the e ects of the actual data to create an accurate pricing model.

Component 1: Supply

Here we establish and prove a link between crude oil supply and gasoline prices. We also show that there is a response time such that we can use supply to anticipate gasoline prices.

We create a theoretical model (created entirely based on observations we made and reasoning).

The model here uses current gasoline price, supply of crude oil, and a constant "C" (a natural result of one of the integrals), which are the combined e ect of qualitative data (This constant is modeled in component 3).

Component 2: Individual Profiles

This component is simply a basic analysis and profile of the top five oil exporting countries to the United States. Important factors are discussed and analyzed.

Component 3: Qualitative Data

We determined that the global weather and political stability to be the most important qualitative data that a ects gasoline prices.

We categorize the di erent intensities of the two data, and each category is given a quantitative value that a ects the constant "C" and subsequently the price of gasoline in the future. Essentially, we are able to quantify qualitative data.

We now have a complete model that can predict future gasoline prices. The only unknown variable is , a constant value that is applied to calibrate the magnitude of the changes in supply to the changes in gas prices. We use the retail gasoline prices in 2011 to solve for   and the average value is used for the General Pricing Model.

In building the model, we accept that there exists uncertainties and near random movements. We express this concept and its e ects by introducing a random variable   and allow it to act like a simplified Wiener Process on the model. This accounts for random movements that we cannot see.

Our next step is to test the model. We used the model starting in January of 2012 until the most recently available data. This is then compared to the actual prices.

Part III) The Binary Tree Model
In this part, we plan out and model the actual decision making possibilities of the consumer. Initially, we established that the consumer has a binary choice each time he has the chance to buy gas. We analyze each given case separately.

By looking at the outcomes of the binary trees, we create a table that the consumer can refer to. The table contains a column that calculates the average amount paid for every mile driven for very outcome. The consumer only needs to use the General Pricing Model from Part II to find the price each week in the future and refer to the table in this part to find the outcome where he would be paying least per unit mile driven.

Part IV) The Boston Model
In this part, we take our General Pricing Model (which prices for the weekly United States Regular All Formulations Retail Gasoline Prices) and apply it to the city of Boston by adding in location-specific variables to recalculate the value "C" in the model.

This part is very similar to the testing of the General Pricing Model at the end of Part II.

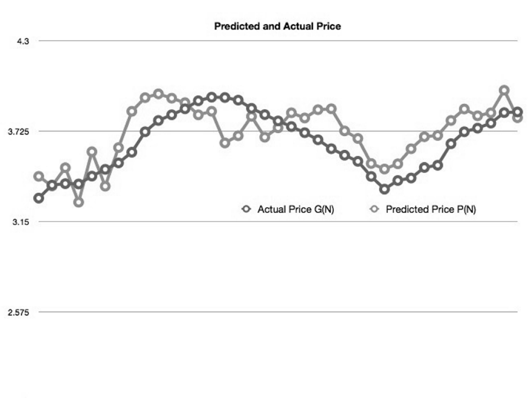
Part V) The Letter

As requested, we wrote a letter to our local newspaper (The South China Morning Post) and presented them with an article detailing the basics of our model and how the average person can use it to save money. This is written in a non-technical language.

Results
We obtained relatively positive results from our model. Here is a graph of our predicted price of 2012 and the actual price (from the testing of the General Pricing Model at the end of Part II):

Our percent error was 3.62%, which is very good.

Here is a graph of the results from the Boston Model: Here, the correlation is slightly less but still maintains it's positive results.



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

Team Members: Harrison Grinnan, William Kunkel, Lisa Li, Jonas Rogers

Advisor: Dickson Benesh

The average American family spent $368 per month, or $4416 per year, on gasoline in 2011. Gas has gone up by an average of 19 cents per year over the past 9 years. For our model city (New York), our model saved 1.75% over normal consumers when driving 200 miles a week. This would translate to $77.45 savings per year for the average American family just from optimizing the timing of gas purchases. These numbers explain the immediate importance of this problem to all Americans.

Our model is a multiple regression for the change in price of gas, given past weeks' change in gas, past change prices of crude oil, and the displacement of the current price of gas from the average. It uses all linear combinations between past weeks' change in gas, and all linear combinations of past change prices of crude oil. The R^2 of our model for our model city, or the percentage of variance in the output variable that is explained by variance in the input variables, is 0.784. More importantly, the model correctly forecasted the direction of change in price of gas in 40 out of 45 weeks in 2012 after being trained on data from 2011.

We also investigated other methods of calculating the change, finding that a simple Markov process could make many effective predictions. We included this model in our analysis, and used it in our letter to the editor, as it is easy to understand and simple to implement in actual gas buying routine.

We then calculated the savings garnered using the multivariate linear model in 2012 and compared it to the simple Markov model, the model that buys gasoline whenever the tank is empty, the model that fills the tank every week, and the optimal solution, found using Dijkstra's Algorithm. Our model saved more than the simple Markov model in most cases. Variance in savings can be attributed to the severity of errors in the model's predictions.

Our model is consistent in saving money and picking the correct direction of change. It is not as useful for predicting amount of change, and does not predict wild swings with great frequency, but rapidly adapts to them. The simplified Markov model in particular would be very easy and effective to implement in real life.

Problem B Summary: Shenzhen Middle School

Team Members: Yichen Li, Zhuoer Li, Guixing Lin, Zhepei Wang

Advisor: Wentao Zhang

Gasoline is the blood that surges incessantly within the muscular ground of city; gasoline is the feast that lures the appetite of drivers. “To fill or not fill?” That is the question flustering thousands of car owners. This paper will guide you to predict the gasoline prices of the coming week with the currently available data with respect to swift changes of oil prices. Do you hold any interest in what pattern of filling up the gas tank can lead to a lower cost in total?

By applying the Time series analysis method, this paper infers the price in the imminent week. Furthermore, we innovatively use the average prices of the continuous two weeks to predict the next two weeks' average price; similarly, employ the four-week-long average prices to forecast the average price of four weeks later. By adopting the data obtained from 2011and the comparison in different aspects, we can obtain the gas price prediction model:
$Gt{+}1 = 0.0398 + 1.6002gt{+}{-}0.7842gt-1 + 0.1207gt-2 + 0.4147gt - 0.5107gt-1 + 0.1703gt-2 + \hat{\varepsilon}$.

This predicted result of 2012 according to this model is fairly ideal. Based on the prediction model, we also establish the model for how to fill gasoline. With these models, we calculated the lowest cost of filling up in 2012 when traveling 100 miles a week is 637.24 dollars with the help of MATLAB, while the lowest cost when traveling 200 miles a week is 1283.5 dollars. These two values are very close to the ideal value of cost on the basis of the historical figure, which are \$635.24 and \$1253.5, respectively. Also, we have come up with the scheme of gas fulfillment respectively. By analyzing the schemes of gas filling, we can discover that when you predict the future gasoline price going up, the best strategy is to fill the tank as soon as possible, in order to lower the gas fare. On the contrary, when the predicted price tends to decrease, it is wiser and more economical for people to postpone the filling, which encourages people to purchase a half tank of gasoline only if the tank is almost empty.

For other different pattern for every week’s “mileage driven”, we calculate the changing point of strategies-changed is 133.33 miles.

Eventually, we apply the models to an analysis of New York City. The result of prediction is good enough to match the actual data approximately. However, the total gas cost for New York is a little higher than that of the average cost nationally, which might be related to the higher consumer price index in the city. Due to the limit of time, we are not able to investigate further the particular factors.

Problem B Summary: High Technology High School

Team Members: Angelica Chen, Derek Liu, Eric Schneider, Kevin Zhou

Advisor: Ellen LeBlanc

The cost of gasoline plays a major role in the everyday life of middle-class families. However, gasoline prices fluctuate significantly, increasing the need for wise financial decision-making.

The first part of our model identified possible predictors of gasoline price. We took into account both economic factors, such as gasoline company stock prices and crude oil price, as well as natural factors, including the weather. We constructed a multiple regression model from 2011 data and reduced it through AIC-selection, until only three significant predictor variables remained: the change in gasoline, crude oil price from the past week, and the value of ExxonMobil stock. Because we found that prices from two weeks ago were only weakly correlated with current prices, we decided not to model gasoline price more than one week into the future.

The second part of our model used these variables to generate a probabilistic prediction of the change in gasoline price in the following week. We split the possible percent changes in gasoline price into five quintiles, and then used the 2011 training data to fit the parameters of each quintile to a multivariate normal distribution. Testing the model on 2012 data, we obtained, for each week, the probabilities that the change in gas would be in each of the five quintiles. We then compared our predictions with actual 2012 gasoline prices and showed that our model was 1020 times more likely than the null hypothesis.

The final part of our model used the probabilities generated to decide what actions to take. Because it is more advantageous to buy gas right before a price increase, we used the idea of thresholds: if the predicted change in price is above a certain threshold, we advise the consumer to buy more gas (a full tank instead of a half tank, or a half tank instead of nothing). We found the optimal thresholds by training on 2011 data and then applied them to 2012 data using the second part's price predictions. We calculated the efficiency of our strategy by comparing the resulting savings to the best possible savings.

Applying this model to gasoline prices from New York City gave efficiencies of 35% and 53% for cases 1 and 2, respectively. These correspond to savings of \$3.84 and \$5.75 per year, respectively.