The 2022 HiMCM

Jack Picciuto, PhD, HiMCM Contest Director

Congratulations to our 2022 Outstanding team winners and all teams participating in our twenty-fifth International High School Mathematical Contest in Modeling (HiMCM)®. We are excited to again join the National Council of Teachers of Mathematics (NCTM) to designate two of our teams as NCTM Award winners. The HiMCM continues to be an amazing and rewarding experience for students, advisors, schools, and judges across the globe. A total of 854 teams, with up to 4 students each, representing 362 schools and 18 countries/regions, competed this year.

Outstanding Teams

- 12465 Nanjing Foreign Language School International Center, Jiangsu, China
- 12600 Shanghai Linstitute School, Shanghai, China (NCTM Winner)
- 12646 Hangzhou No.14 High School AP Center, Zhejiang, China
- 12678 Shanghai American School (Puxi Campus), China (NCTM Winner)
- 12821 Shanghai Pinghe School, Shanghai, China
- 12911 Ward Melville High School, NY, USA
- 13010 North Carolina School of Science and Mathematics, NC, USA
- 13014 North Carolina School of Science and Mathematics, NC, USA
- 13405 The Nueva School, CA, USA

The 2022 Contest

Once again, our 2022 participating teams' papers were a joy for the judges to read with some truly impressive work. COMAP seeks to continue the tradition of creating challenging real-world problems that are interesting for students. As in the past, students could choose from two problems. This year's problems challenged teams to investigate honeybee population growth and pollination capacity in Problem A: The Need for Bees or to determine a relationship between CO_2 levels and global temperatures in Problem B: CO_2 and Global Warming. The judges continue to be very impressed with the students' drive to compete and their mathematical abilities shown in this modeling contest. We understand the challenges of time and resources put on students and would like to thank all participants and advisors who competed in this year's HiMCM contest.

Overview

While COMAP has offered international modeling contests for over 40 years, HiMCM celebrated its 25th contest in 2022. As growing numbers of schools engage their students in mathematical modeling, we continue to see increasing participation in COMAP's modeling contests. Starting with 115 students in the first year of the HiMCM, over the course of 25 contests we have had 44,544 students apply their mathematical knowledge and skills as they modeled challenging problems in the HiMCM. COMAP recognizes the value and importance of the students' team advisors and teachers. These are the true champions who see the value of their students' participation in a math modeling contest. These are the educators and mentors who encourage their students to go beyond the standard curriculum. We see many dedicated team advisors who, year after year, support one or more teams in this math contest challenge. COMAP is truly thankful for these individuals.

The 2022 contest had 854 submissions. Of the 854 submissions, 413 completed Problem A: The Need for Bees, and 441 completed Problem B: CO_2 and Global Warming. **Table 1** on the following page shows the judging results of the 2022 HiMCM. We accept partial solutions and encourage all registered teams to submit a solution paper to experience the learning impact and satisfaction of fully participating in this challenging contest.

In total, 3,197 students participated in the 2022 HiMCM. A wide range of schools competed, including teams from Australia, Canada, Chile, China, Germany, Hong Kong (SAR) China, India, the Philippines, Singapore, South Korea, Taiwan China, Thailand, the Netherlands, the United Kingdom, the United Arab Emirates, the United States of America, and Vietnam.

Rules

Prior to (and during) the contest, teams and advisors should review the rules and any rule updates. One important rule is that students may only use the members of their team along with inanimate (non-living) sources to complete the contest problem. Students may not

Problem	Outstanding	%	Finalist	%	Meritorious	%	Honorable Mention	%	Successful Participant	%	Totals
A	4	1%	25	6%	62	15%	98	24%	213	51%	**402**
B	5	1%	29	6%	70	16%	109	25%	218	49%	**431**
Totals	9	1%	54	6%	132	15%	207	24%	431	50%	**833**

Table 1: *2022 HiMCM Judging Results*

Note: The table does not include teams that were not judged, unsuccessful or disqualified

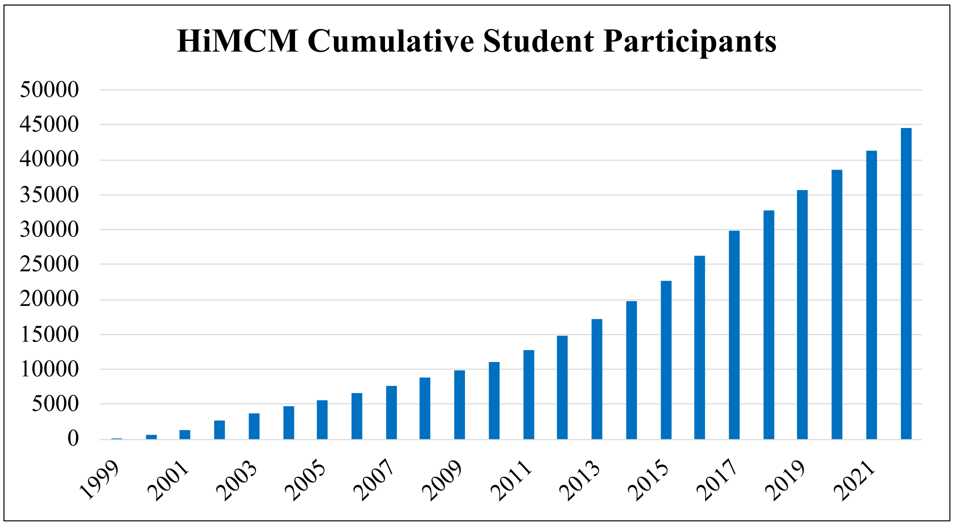


Figure 1: *Total HiMCM Student Participants 1999 – 2022*

use chat rooms, electronic communication, or social media sources. Each year we have some teams that do not understand this rule. To be clear, contacting an expert in a field or an author of one of the referenced sources is a violation of this rule. Gathering data from persons outside of your team using an interview or a survey or a questionnaire is a violation of this rule. Using solutions shared electronically by other teams or by organizations is a violation of this rule. Again, only the team members may contribute to the solution through their knowledge and work, and by using inanimate resources (e.g., research articles, websites, textbooks, journals, and publications). You will also need to cite these inanimate resources in your reference or works cited section. Additionally, COMAP will never require that you purchase additional materials or information to be successful in the HiMCM. The materials and information provided by COMAP, along with your own team's knowledge, skills, and perhaps a bit of research using allowed references, is all that you need for success.

COMAP uses Twitter and Weibo to provide contest information to participants. Follow us @COMAPMath on Twitter or COMAPCHINAOFFICIAL on Weibo for contest guidance and up to date contest information.

Judging

All contest submissions are electronic. This allows a high quality and diverse judging pool from academia and industry to simultaneously judge papers. Soon after the contest ends, we conduct our first round of contest judging. Each paper is read and scored by two preliminary judges. We thank these judges for their careful review of our HiMCM submissions.

All judging is blind with respect to any identifying information about the participants or their schools. Each year during the conduct of the contest, COMAP does get a few clarifying questions about the problems. In most cases, our response is the same – "These are open ended questions and based on your assumptions and approach your team will do its best to answer the questions posed." We do not compare your responses to an official answer key. Modeling problems rarely have a single solution approach or answer. The path(s) your team takes may use different techniques than other teams use and may lead to very different and equally valid solution(s). Be sure to explain and support your approach and critically analyze your solution. Judges score each submission on its own merits as we look for completeness, creativity, and the use of good math modeling techniques. Preliminary judges rank papers as Finalist, Meritorious, Honorable Mention, and Successful Participant. Judges send all papers ranked as "Finalist" to Final Judging. This year, 65 papers from the two problems went to Final Judging for a panel of twelve judges to consider. As these 65 papers were the best submissions from the preliminary round, at Final Judging the judges chose the "best of the best" as Outstanding papers. Nine papers earned the Outstanding award. The final judges commend the preliminary judges for their efforts in selecting the high-quality Finalist papers. We feel that the structure of preliminary and final judging provides a good process for identifying our top papers.

HiMCM

The Future

For 25 years, the HiMCM has sought to provide all high school students the opportunity to compete and achieve success in applying mathematics. Our efforts remain focused on meeting this important goal. Mathematical modeling continues to grow within the high school curricula across the globe, and we recognize that middle school students are now modeling too. In 2022, COMAP held its second international MidMCM, a middle school/level contest option. The MidMCM occurred concurrently with HiMCM. The MidMCM allows middle school/level students under the age of 14 ½ years old the opportunity to demonstrate their mathematics and modeling abilities. Please visit www.MidMCM.com for more details and the results of this new contest.

The MidMCM and the HiMCM provide a vehicle for using mathematics to build models that allow students to represent, and to understand, real world behavior in a quantitative way. Both contests enable student teams to look for patterns and think logically about mathematics and its role as a language in our daily lives. Students gain confidence by tackling ill-defined problems and working as part of a team to generate a solution. We are excited that in our modeling contests applying mathematics is a team sport.

Advisors and students often ask what level of mathematics is required, and what special programming or coding skills are needed for the contest. All HiMCM problems are accessible using high school level mathematics alone, and no programming or coding skills are required or necessary. Our new MidMCM problems require only middle school level mathematics. As in all our contests, each of our problems is accessible on multiple levels. Students should apply the mathematics they understand and are able to explain in their solution analysis. COMAP understands that there is not a standard body of knowledge for high school level mathematics in an international modeling competition. The judges see a wide range of abilities and skills. Teams should note that by properly using *lower-level* concepts of mathematics that they understand, their team can do just as well in the contest as students using higher levels of mathematics. Teams should not be focused on impressing the judges with overly complex techniques and models, but rather concentrate on using good modeling techniques in putting together their solution.

Advisors need only be motivators and facilitators to encourage students to be creative and imaginative. COMAP encourages all middle and high school mathematics faculty to get involved—encourage your students to be problem solvers, make mathematics relevant, and open the doors to future success. Any school can enter, and each school can enter as many teams as that school desires. MidMCM and HiMCM have no restrictions on the number of total schools or the numbers of total teams. Advisors should encourage student teams to review the COMAP website for resources and read judges commentary of past student solutions. More than just learning skills and operations, mathematics is both an art and a science. Through mathematical modeling, students learn to think critically, communicate effectively, and be confident, competent problem solvers. We want to partner with teachers as we continually strive to improve the contest and make it accessible and impactful to all students. Contact COMAP via www.comap.com with any questions or suggestions.

2023 Contest Dates

Mark your calendars for the next HiMCM, and the third annual MidMCM, to be held November 1 – 14, 2023. Registration for the 2023 MidMCM and HiMCM will open in September. Student teams may work at any time during the contest window. At the team members' convenience, teams download and choose their problem, complete their modeling solution, and electronically submit their solution document by the deadline on November 14th. Again in 2023, one team for each problem will receive the NCTM award. Teams can learn more about COMAP's contests and registration at www.comap.com.

MathModels.org

Powered by COMAP content, Mathmodels.org is a wonderful resource for students and teachers to make math modeling a year-round activity. Teachers and students may use the materials found on this site to enrich their classes and help prepare students for mathematical modeling competitions. We encourage you to visit www.mathmodels.org.

—AWARDS—

Outstanding
Finalist
Meritorious
Honorable Mention
Successful Participant

After final judging, HiMCM papers receive a designation in one of the categories above. Depending upon the quality of the papers, the top 20–25% of submitted papers receive a designation of Meritorious or above, with approximately the top 1% designated as Outstanding.

The International Mathematical Modeling Challenge, IM^{2}C

The IM^{2}C is held each spring and continues to grow. This unique contest is similar to an all-star game where each country administers and judges a modeling contest and then sends its top two teams to the IM^{2}C. The purpose of the IM^{2}C is to promote the teaching of mathematical modeling and applications at school level (high school and below) for all students around the world. It is based on the firm belief that students and teachers need to experience the power of mathematics to help better understand, analyze, and solve real world problems outside of mathematics itself – and to do so in realistic contexts. An international Expert Panel of final judges determines winners and selects teams to present their solutions at an international award ceremony. To learn more visit www.immchallenge.org for rules and country / region contacts.

COMAP invites selected teams from the United States, to include teams earning Meritorious or above in the HiMCM contest, to compete in the IM^{2}C U.S. Regional Round. Registration is free. From these participants, U.S. IM^{2}C judges select the two top teams to represent the U.S. in the IM^{2}C international round. See the U.S. Rules at
https://immchallenge.org/Pages/Rules/USA/USA-Rules.html

Problem Discussions and Judge's Commentary

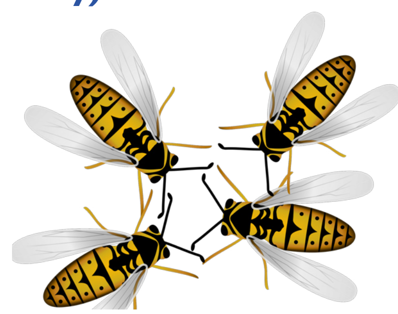
The following paragraph describes what our preliminary and final judges look for in our HiMCM papers:

> *Regardless of the problem chosen, competitive papers include a comprehensive summary, address all requirements through developing and applying a mathematical model(s). Better papers do all the above in an articulate, well-supported, well-organized, and well-presented manner. The best papers combine complete mathematical and logical analysis and explain their work in an organized presentation beyond simply addressing the requirements. These best papers are easy to read, flow logically, are creative, and they include sections that address assumptions with justifications, the modeling process(es), results of modeling and analysis, strengths and weaknesses, sensitivity, conclusions, and references.*

Our judges have asked that we continue to stress that all our HiMCM problems are accessible by students at any level using high school mathematics. Some teams attempt to use advanced concepts and tools found on the Internet that they do not understand, explain clearly, or use appropriately. Judges recognize this, and these papers do not do well. We are not looking for papers that use the most advanced mathematics. We have found that the best papers develop a mathematical model that incorporates high school level mathematical concepts and tools that the teams understand, are able to fully explain, use appropriately, and analyze subsequent results. The most important aspects of solutions are the model development and the clear use and analysis of the model toward addressing the requirements of the problem.

The specific problem discussions below provide comments on how teams addressed the requirements of each problem. Following this section, we provide judges' comments about the solutions and presentations by breaking down the various parts of a submission and providing exemplars. To view the complete problem statements, visit
www.mathmodels.org or
www.himcmcontest.com.

Problem A: The Need for Bees (and not just for honey)



In this problem, we asked teams to think about honeybees and their importance to human existence. Most of us see honeybees in everyday life. They are buzzing around and always on the move. We may think of honeybees as just a source of honey, but they play a vital role in the pollination of many plants, flowers, and trees.

The problem introduction mentioned factors that have led to a decline in the number of honeybee hives including viruses, pesticides, habitat destruction, and environmental conditions. The problem provided some helpful information to consider but also encouraged students to research other information using online or other inanimate sources. Student teams developed a model to determine the population of a honeybee colony over time. Additional problem requirements included:

- Conduct sensitivity analysis on your model to determine which factors (e.g., lifespans, egg laying rates, fertilized / unfertilized egg ratios, or other factors) have the greatest impact on honeybee colony size.

- Model and predict how many honeybee hives you will need to support pollination of a 20-acre (81,000 square meters) parcel of land containing crops that benefit from pollination.
- Create a non-technical, one-page blog or infographic for a website that presents the information you developed.

In past contests, we asked students to provide a 1-2 nontechnical report or news article to *publicize* their findings. Condensing an up to 25-page report into a abridged, nontechnical language remains important to the modeling process. COMAP decided this year to modify this requirement and instead had the students provide a blog or infographic. Judges were impressed by both the blogs and infographics this year. Teams presented a range of methodologies, from simple and straightforward to more complex, to address the challenges facing honeybees as well as their capacity to pollinate a large parcel of land. The judges enjoyed reading these thoughtful and creative solution papers.

Problem B: CO_2 and Global Warming

This problem addressed the relationship between the amount of CO_2 in the atmosphere and its relationship to increasing global temperatures. The problem included data to either support or refute claims made in a scientific journal. The problem prompt asked students to respond to the following:

- Do you agree that the March 2004 increase of CO_2 resulted in a larger increase than observed over any previous 10-year period? Why or why not?
- Fit various (more than one) mathematical models to the data to describe past, and predict future, concentration levels of CO_2 in the atmosphere.
- Use each of your models to predict the CO_2 concentrations in the atmosphere in the year 2100. Do any of your models agree with claims and predictions that the CO_2 concentration level will reach 685 ppm by 2050? If not by 2050, when do your models predict the concentration of CO_2 reaching 685 ppm?
- Which model do you consider most accurate? Why?

After agreeing or disagreeing with the claims from the journal, students investigated the relationship, if any, between CO_2 levels and changes in global temperatures based on any conclusions obtained from the first requirement. Specifically, the problem required students to:

- Build a model to predict future land-ocean temperatures changes. When does your model predict the average land-ocean temperature will change by 1.25°C, 1.50°C, and 2°C compared to the base period of 1951-1980?
- Build a model to analyze the relationship (if any) between CO_2 concentrations and land-ocean temperatures since 1959. Explain the relationship or justify that there is no relationship.
- Extend your model into the future. How far into the future is your model reliable? What concerns, if any, do you have with your model's ability to predict future CO_2 concentration levels and/or land-ocean temperatures?

Lastly, student teams prepared a one-page non-technical article to explain the team's findings and possible recommendations for the future.

Judges made many positive comments on the teams' work on this problem. Papers generally answered the given questions and provided mathematical models to analyze data and make predictions as required. The robustness of models varied widely. Most models focused on various types of regression, and some even used time series analysis, such as the autoregressive integrated moving average (ARIMA) model. The robustness of the model did not differentiate the quality of the paper as much as the explanation and analysis of the models' results. Judges realize there are stand-alone computational models available to students. The better papers explained the model process and evaluated the results for accuracy and common sense based on their own model and the data. The best papers also provided specific recommendations to lower CO_2 and land-ocean temperatures in the non-technical article. We remind students that they should address results, and not just methodologies, in both the paper's executive summary and accompanying blog, infographic, or nontechnical article.

Judges' Discussion

While the problem discussions above provide comments on the solutions to this year's problems, in the following paragraphs we examine the sections of a submission and provide comments about the solutions and the presentation of the solutions from our judges' point of view. At the end of the article, we have included excerpts from the top papers as exemplars from both problems. Mathmodels.org members can view all the unabridged versions of the the top papers online.

Overall

Student participants must ensure their papers follow the contest rules posted on the contest website. Every year, judges see papers that have exceeded the page limit, decreased their font, or narrowed their margins to present more material. While these papers are not necessarily disqualified, they are unlikely to score well. Papers that are complete, coherent, organized, clear, and well written impress the judges. The reasoning and mathematics of these papers are easy to follow. These best papers tell a story in a logical sequence. These papers clearly provide the background and justification of the developed model(s) and properly apply and analyze the model(s). Teams should present their entire submission in 25 pages or less, using at least 12-point font. These 25 pages should include your introduction/executive summary, your solution that addresses all requirements, a resource list, and any appendices. While students may want to include some background research on the problem topic, this information should be brief. It is not the number of pages, but the ability to complete all problem requirements and communicate the solution in a concise and articulate fashion that will earn recognition. Students should use spelling and grammar checkers before submitting a paper. Foreign papers should ensure that all symbols in tables and graphs are in English. Student and school names should not appear on solution papers.

Papers considered for Finalist and Outstanding start with a clear summary that describes the problem. These papers then preview their paper with an organized Table of Contents. They present assumptions with justifications, explain the development of their model and its solutions, apply their model, and support the results mathematically. These best papers communicate all the above clearly, conduct sensitivity analysis or other model testing, address model strengths and limitations, and finally, close by stating overall conclusions.

Lastly, please only submit your paper to COMAP one time and select the problem (A or B) that you are submitting. Multiple submissions and/or teams that do not identify which problem they chose slows the judging process.

Executive Summary

As the first page, or cover sheet, the Executive Summary provides a first impression of your paper. It offers the judge (or any reader) not only a synopsis of the paper, the modeling and the analysis process, but also the solution(s) to the problem. Judges see many well-detailed descriptions of the problem and the process but look for well-written and complete summaries that include the actual results and recommendations. Every year, the judges see many submissions that continue to fail to include results in their Executive Summary. Teams should write the Executive Summary *after they finish* their solution to summarize the entire contents of the paper. **Example 1** from Team #13014's submission includes a fine example of a summary. Note this team includes their results and not just a discussion of their approaches. An ideal summary can stand alone and give the reader a synopsis of the problem, the methods used, and solutions.

Assumptions with Justifications

Good models make a few necessary assumptions to help simplify the modeling process. These are sometimes called simplifying assumptions. A common mistake judges see is teams that make assumptions that are not needed nor relevant to developing their model. Students should consider asking themselves, "do we need this assumption to develop or support our model?" Good and relevant assumptions are difficult to identify and articulately state. Long lists of assumptions that do not play directly in the context of model development or its solution are not considered relevant and deter from a paper's quality. Assumptions that oversimplify the problem too much do not allow for a complete or useful solution. You should include a short justification to show each assumption is reasonable and necessary. These assumptions often directly relate to your model sensitivity analysis or impact your model strengths and weaknesses. For example, if you assumed a certain value or no impact by a certain exterior factor, how would changing this assumption impact the validity of your model conclusions? See **Example 2** of a concise and well written set of assumptions with detailed justifications for Problem B from Team #12465.

Definition and Use of Variables

Most mathematical models include several variables that teams must define for the reader. This list of variables should include the variable symbol, a short description of the variable, and the units of the variable. Judges often see lists of 10, 20, or even 30 or more unique variables in HiMCM submissions. There is no minimum or maximum number of variables, but rather students should choose an adequate set of appropriate variables to best and fully model the problem. Any complexity gained by adding an excessive number of variables is often offset by their practicality and usefulness. Using best practices, teams should focus on a manageable set of variables when modeling. Judges often see many variables that are defined but never used in the models. Additionally, as you use variables in your model, remind the reader of the variable definitions and units. This practice assists the reader in following the logic of your process. A nice list of variables for Problem A is shown by team #13405 in **Example 3**. Their list both defines the variables and includes units of measurement where applicable.

Mathematical Model

The development of the mathematical model is the most important part of your submission. There is always more than one appropriate solution method to our HiMCM problems and so teams should address the problem with the mathematics they know and understand. Papers should explain the development of the mathematical model(s) and/or algorithm(s) and define all variables in a logical manner. Teams should take the reader on a journey describing why they selected a particular model or decided to use one or more models. Better teams will explain why they choose their model and how they plan to use or modify it to fit this problem. Teams that merely present a model without explaining or showing the development of that model do not generally do well. Although during your HiMCM work you may develop several models, presenting multiple models without identifying the most appropriate model to answer the questions is detrimental to your paper's success. Judges are more impressed with a well thought out (and perhaps simple) model then with a very complex model that a team struggles to apply. To impress the judges, focus on applying sound principles to your model that you understand. Judges do value and reward creativity and thinking "outside of the box" in your modeling process but be sure to balance creativity with your level of expertise and modeling experience. A submission that searched the Internet for an existing model to apply may not do as well as a team who modified an existing model a team who or took the results from one model as inputs to another model. So, be creative and have fun.

Perhaps the most important step of the modeling process is the last one: explicitly present your final model in its full form. Do not make the judges have to look for your final model. Judges continue to see papers with an initial model mentioned in one section and then different models used in subsequent sections without the team connecting the models logically. Be consistent with your modeling process and guide the reader through your solution. Clearly identify your model(s) and then use your final model as you address the problem requirements and determine your results. Does your model use your variables and is it based on the assumptions you made earlier in your submission? Papers that do not flow well tend to not be judged as the best papers.

There are many ways to model and analyze the HiMCM problems. This year we saw a variety of appropriate, as well as creative, models to address both problems. The use of tables, graphs, and images is often helpful to show your modeling process. Judges appreciate a good mix of visual aids and quality writing. Papers that are pages and pages of text or pages and pages of graphs and charts do not do as well as those that have a good mix of both. The better papers will reference and discuss the impact of all included graphs and charts. We include several examples of the processes involved in model development for this year's problems. For Problem A, the Honeybee Problem, Team #13405 in **Example 4**, does a nice job laying the groundwork for their model development. Their discussion explains to the reader why they chose this modeling path. In Problem B, Team #12678, in **Example 5** started "simple" and expanded from there. Often a simpler approach will provide valuable insight into the problem. Many teams continue to seek out and apply overly complex models when a simpler approach can yield better results.

The second part of Problem A was to model and determine an appropriate number of hives to pollinate a 20-acre field. As stated previously, judges do not compare students' specific answers to a solution key and this requirement saw responses from a single hive to 50 or more hives. If properly justified by assumptions and other factors, most all solutions may be acceptable. In **Example 6**, Team 13104 shows how they started this portion of the problem. One of their assumptions mentions a circular radius for bee travel. This led to their model.

The second part of Problem B was to model the relationship between CO_2 concentrations and land-ocean temperatures. In **Example 7**, Team #12465, chose a different model than used in the previous parts of the problem.

Sensitivity Analysis

Sensitivity analysis is a very important part of the modeling process. A primary purpose of sensitivity analysis is to "test" some of the assumptions made to include parameter values. If your model yields a certain result based on these assumptions, then what happens if you modify an assumption? **Example 8** from Team #12600 shows how they modified a few parameters by 10%. The judges like to see some sensitivity analysis along with a brief discussion by the students on resulting impacts, if any, on the solution.

Strengths and Limitations

As teams only have a short time to develop their model, we expect their models to have both strengths and limitations. Teams should be critical of their proposed model and solution as they address strengths, limitations, and possible model extensions or improvements. Is your solution reasonable? Under what conditions will it perform best and where will it not? If you had more time or resources what else would the team like to know to improve their model? This is also an opportunity for teams to apply common sense to their model and check the reasonableness of their proposed solutions. In **Example 9**, Team #13010 lists some honest criticism of their model.

Conclusion

A clear conclusion and answers to the specific scenario questions are key components to an Outstanding paper. Many students confuse the conclusion with the Executive Summary. The conclusion is more about overtly and specifically stating the results of any requirements, whereas the Executive Summary is a stand-alone section that allows a reader to preview and quickly understand the entire problem, modeling processes applied, and solution. Team #12911 shown in **Example 10** has one of the better conclusions.

The Infographic, Blog, or Non-technical Article

The purpose of this section is to show the judges that regardless of the complexity of your mathematics or analysis, you can convey your work and solution in common terms and language. Often, the recipients and consumers of mathematical modeling and research are not as technically astute as the mathematicians and scientists doing the work. Teams must translate their hard work into a shorter descriptive format, such as a news article or graphic. This increases the likelihood your audience will understand and consider your solution and recommendations. This section should tell the same story as your technical write-up but using easier to follow text and graphics. From Problem A, a nice example of an infographic is shown in **Example 11** from Team #13014. From Problem B, Team #12505 presents a great non-technical article in **Example 12**.

Citations and References

Citations and references are very important within this paper or any other paper you write that uses outside sources. Teams that use existing models should cite their source(s) within the paper at the point they present the model and include a reference citation in the back of the paper. This is also true for all graphs and tables taken from the literature. Use "in line" documentation with footnotes or endnotes to give proper credit to outside sources. All data, figures, graphs, and tables that come from outside sources require documentation at the point in the paper where they appear. Lack of documentation will result in a lower designation. We have noticed an increase in the use of Wikipedia. Teams need to realize that although useful, information from Wikipedia might not be accurate. Teams should recognize and acknowledge this fact and look for primary resources.

Final Thoughts

On behalf of COMAP and the contest judges, we again thank all advisors and students for their participation in the HiMCM. Each year the quality and level of mathematics demonstrated by our high school student teams amazes and impresses our judges. We truly enjoy reading all solution papers. Successful teams use a wide variety and level of mathematics. While teams using post-high school/undergraduate level mathematics are in a league of their own, teams that use basic high school mathematics and much simpler approaches are often among our Outstanding designees when they understand and can explain their work. We encourage students of all levels to compete in future HiMCM competitions as well as our MCM/ICM contests targeted to undergraduates. To be successful read the comments and guidance provided in this article, see the TIPS article on the COMAP website, and visit

www.mathmodels.org

to review previous problems.

Follow us @COMAPMath on Twitter or COMAPCHINAOFFICIAL on Weibo, LinkedIn, and/or Facebook for information about all COMAP contests.

List of Examples:

1. **Summary** (Problem A, Team 13014, North Carolina School of Science and Mathematics, NC, USA)
2. **Assumptions and Justifications** (Problem B, Team 12465, Nanjing Foreign Language School International Center, Jiangsu, China)
3. **List of Variables** (Problem A, Team 13405, The Nueva School, CA, USA)
4. **Model Development** (Problem A, Team 13405, The Nueva School, CA, USA)
5. **Model Development** (Problem B, Team 12678, Shanghai American School (Puxi Campus), China)
6. **Model Development** (Problem A, Team #13014, North Carolina School of Science and Mathematics, NC, USA)
7. **Model Development** (Problem B, Team #12465, Nanjing Foreign Language School International Center, Jiangsu, China)
8. **Sensitivity Analysis** (Problem A, Team #12600, Shanghai Linstitute School, Shanghai, China)
9. **Strengths and Limitations** (Problem B, Team #13010, North Carolina School of Science and Mathematics, NC, USA)
10. **Conclusion** (Problem B, Team #12911, Ward Melville High School, NY, USA)
11. **InfoGraphic** (Problem A, Team# #13014, Team#13014, North Caro lina School of Science and Math ematics, NC, USA)
12. **Non-Technical Article** (Problem B, Team #12505, Wuhan Britain-China School, Hubei, China)

Honeybees are crucial to the survival of humanity. These insects fertilize the crops that feed billions of people worldwide. Given the significance of these insects' role and the recent decline of honeybee populations, it is critical to understand trends in bee population over time, determine which factors have the most significant impact on a honeybee colony's population, and model the number of hives needed to pollinate a given field.

To this end, we developed a bee population model that predicts the population of worker and drone bees in a colony at a given time. We constructed differential equations to model the change in worker and drone bee populations. To create these differential equations, we considered the following factors: the initial population of bees, ratio of total bees to drone bees, length of the active honeybee season, egg laying rate, average lifespan of bees, and gestational period of bees. Next, we utilized Euler's method to predict these trends over ten years. We determined that the population of a honeybee colony ranges from approximately 19,046 bees at the end of the inactive season to approximately 60,759 bees at the end of the active season. The worker bee population ranges from 19,046 to 60,053 bees, and the drone bee population ranges from 0 to 706 bees. These results are validated by online sources about typical bee populations ("Honey Bee Colony," 2021). We observed that the cycle of bee population repeats sinusoidally on a yearly period, so we also performed a sinusoidal regression, allowing farmers to estimate bee populations in the future.

When performing sensitivity analysis on the population model, we modified each factor considered by up to 40% in either direction. This analysis revealed that the egg-laying rate, worker bee lifespan during the active/inactive season, and length of the active season had the greatest impact on the population of bees over ten years. In contrast, modifying the initial population, the ratio of total bees to drone bees, drone lifespan, or any gestational period had little to no effect on the bee population.

We also created a model that utilizes the bee population model to predict the number of hives needed to pollinate a 20-acre field. In the pollination model, we consider the field area, the area required by each flower, the number of flowers pollinated by each worker bee, and the maximum distance a bee travels from its hive (called "bee range"). By calculating the total number of flowers on the field and the number of flowers worker bees pollinate daily, we discovered that it would be best to use 22 hives to pollinate a 20-acre field. Since we discovered the placements of hives did not matter, these 22 hives are assumed to all be placed at the center of the 20-acre field. This result was also confirmed by the University of Georgia, which recommends using approximately one beehive for every acre that needs pollination ("Managing bees for pollination," n.d.). However, using 21 or 20 can also have similar effects, and using one or two fewer hives results in less money spent on buying hives and maintenance for a minimal gain in pollination.

To assess the robustness of our model, we performed a sensitivity analysis on all four factors considered. By modifying each factor up to 40% in either direction, we determined that changing the honeybee range had the most significant effect on the number of hives needed. A greater bee range resulted in more hives needed. Decreasing the flowers visited per day per bee and decreasing the area needed per flower also increased the number of hives required. In contrast, modifying the field size by up to 40% had no impact on the number of hives needed to pollinate the field. Since bees typically stay within 6,000 m of their hive, the 20-acre field is relatively tiny compared to the circle with a radius of 6,000 m pollinated by the hive. Thus, whenever the side length of the field is as small as 285 m, changing the field's size by a relatively small amount has no impact on the number of hives needed to pollinate the field.

Lastly, we created an infographic that outlines all of our findings and relevant information for the general public. We hope this will be a valuable resource for the general public and policymakers to better understand the populations of honeybee colonies and their ability to fertilize crops within a given area.

Example 1: Summary Problem A
Team 13014, North Carolina School of Science and Mathematics, NC, USA

3 Assumptions

Assumption 1: A common disadvantage of empirical/statistical models is that in most cases they are applicable only for the conditions the data were collected, thus they are often not able to predict beyond this particular condition. In our case, we assume the statistical trend of change obtained from historical data applies to future changes. In other words, we simulate future changes under a business-as-usual scenario.

Assumption 1: A common disadvantage of empirical/statistical models is that in most cases they are applicable only for the conditions the data were collected, thus they are often not able to predict beyond this particular condition. In our case, we assume the statistical trend of change obtained from historical data applies to future changes. In other words, we simulate future changes under a business-as-usual scenario.

Justification: Since we don't know how exactly CO2 levels will change in the future, a plausible solution is to simulate different scenarios describing all kinds of possibilities. This approach is also taken by IPCC, in whose annual report many scenarios were created and simulated. One of the scenarios is a business-as-usual scenario where the current trend continues into the end of 21st century. However, we must avoid overfitting during the training stage. In addition, we validate model ability against the validation data set, which was specifically left out from historical data for validation purposes.

Assumption 2: The link between temperature and CO2 is very complex. We assume a rise in CO2 precedes a rise in temperature in a short period and the trend of change in CO2 primarily affects a long-term trend of land-ocean temperature. However, temperature fluctuations between years can be caused by many factors and thus expressed as an autoregressive process.

Justification: Considering physics behind the link, changes in temperature are roughly proportional to changes in radiative forcing as a function of changing concentration of CO2 [7]. Strong correlations present between the trends of temperature anomalies and CO2 levels also support this assumption.

Assumption 3: The CO2 concentration data provided come from the Mauna Loa site and may differ from other CO2 sampling sites. We assume these data have sufficient representativeness for the global average condition in our attempts to investigate the relationship between CO2 levels and land-ocean temperatures.

Justification: The team at Mauna Loa has confidence that CO2 measurements made at the Mauna Loa Observatory reflect truth about global atmosphere [8]. The site is located 3400 m high enough to represent very large areas. All measurements are rigorously calibrated with a very high accuracy.

Example 2: Assumptions and Justifications Problem B
Team 12465, Nanjing Foreign Language School International Center, Jiangsu, China

2.3 List of Variables

Variables	Meaning
L	number of eggs laid per day
s_H	survival function sensitivity to hive bees
s_p	survival function sensitivity to pollen
s_n	survival function sensitivity to nectar
r_p	recruitment function sensitivity to pollen
r_n	recruitment function sensitivity to nectar
r_I	recruitment function sensitivity to social inhibition
b	baseline recruitment rate
u_0	inverse of days as uncapped brood
c_0	inverse of days as capped brood
m_c	capped brood mortality
m_p	pollen forager mortality
m_n	nectar forager mortality
c	pollen gathered ($\frac{grams}{bee}$)
p_U	capped brood pollen consumption (grams)
p_H	hive bee pollen consumption (grams)
n_U	capped brood nectar consumption (grams)
n_W	adult bee nectar consumption (grams)

Example 3: List of Variables Problem A
Team 13405, The Nueva School, CA, USA

3 Bee Colony Model

In this section we present a mathematical model of a bee colony that calculates the number of different types of bees each day as a function of the hive's population the previous day and the food that was collected by the foragers throughout the day. Starting from the initial larva state, moving to pupa, then to capped brood, then to hive bees, and finally to a forager bee, a certain proportion of the bees in each state dies and the rest moves on to the next state after a certain period of time. We don't explicitly include the queen or the drones in the model. The queen is included implicitly though, through the eggs that are laid every day. The drones are also included only implicitly through the laid eggs. Thus the model explicitly captures the number of larva, pupa, capped brood, and worker bees. The dynamics of bee population are represented by differential equations, one equation per bee type. In particular, we are interested in the following vector:

$$\text{hive}(t) = [U(t), C(t), H(t), P(t), N(t), p(t), n(t)]$$

where

$U(t)$	number of uncapped brood bees at time t
$C(t)$	number of capped brood bees
$H(t)$	number of hive bees
$P(t)$	number of pollen foragers
$N(t)$	number of nectar foragers
$p(t)$	amount of pollen stored in the colony at time t
$n(t)$	amount of nectar stored in the colony at time t

There are two decisions embedded in these definitions:

1. Eggs and larvae are combined together into the uncapped brood variable $U(t)$ while capped larva/pre-pupal stage and pupae are combined into the capped brood variable $C(t)$. This is done to explicitly separate the brood into a stage where food is consumed, drawing on the resources of the hive and figuring into colony's allocation of food resources, and a stage where pupae are not consuming any food. Hive bees are required to attend to pupae but no food is spent on them.

2. Foragers are split into pollen foragers and nectar foragers. Because pollen and nectar play complementary roles in a colony's life we want to capture the dynamics that influence how hive bees are transformed into foragers and, possibly, back to hive bees. These decisions are made based on the colony's needs at a given time and, in turn, influence food collection. Recruitment rates create a feedback mechanism since different relative amounts of pollen and nectar foragers affect the nutrition provided to the uncapped brood. The self-regulating mechanisms that allow bees to adapt to different conditions are important in modeling a colony and require these feedback connections.

3.1 What is the simple model?

We first consider a simple time-series model of the annual CO_2 concentration without considering any potential impacting variables. We plot the annual CO_2 concentration versus year in the left panel of Figure 4. There is a clear nonlinear behavior in the data. Furthermore, as the annual changes of the concentration are quite small relative to the values of concentrations, directly fitting the concentrations may result in misleading results. Therefore, we consider the annual changes of the CO_2 concentration, defined as $\Delta C_t = C_t - C_{t-1}$, and plot them in the right panel of Figure 4.

```
                    OLS Regression Results
==============================================================================
Dep. Variable:    Relative Temperature  R-squared:                 0.924
Model:                     OLS  Adj. R-squared:               0.923
Method:             Least Squares  F-statistic:                743.0
Date:            Sun, 13 Nov 2022  Prob (F-statistic):          7.34e-36
Time:                  16:45:24  Log-Likelihood:              63.093
No. Observations:              63  AIC:                        -122.2
Df Residuals:                  61  BIC:                        -117.9
Df Model:                      1
Covariance Type:        nonrobust
==============================================================================
                     coef   std err        t    P>|t|    [0.025    0.975]
------------------------------------------------------------------------------
const             -3.3926     0.138  -24.618    0.000    -3.668    -3.117
Carbon Concentration  0.0105  0.000   27.259    0.000     0.010     0.011
==============================================================================
Omnibus:                  11.338  Durbin-Watson:                1.669
Prob(Omnibus):             0.003  Jarque-Bera (JB):             3.215
Skew:                     -0.034  Prob(JB):                     0.200
Kurtosis:                  1.895  Cond. No.                  4.34e+03
==============================================================================
```

Figure 3: Regression report of concentration to temperature

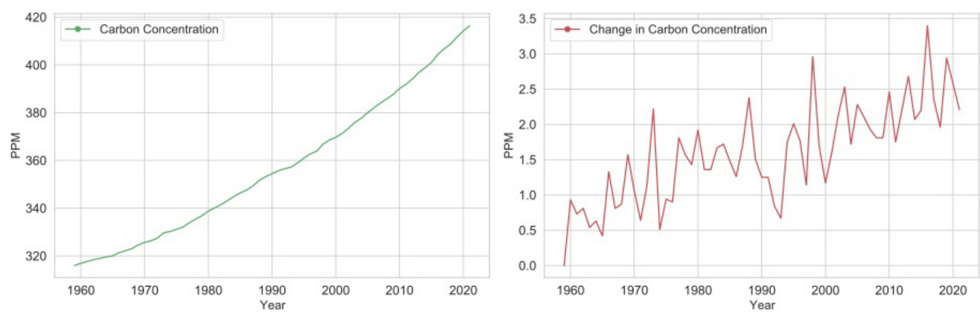


Figure 4: Carbon concentrations and their differences

Following the theory of Ockham's razor, we decide that a linear model is valuable as it effectively models the data with the fewest parameters. Therefore, we make the following assumption.

Assumption 1. *ΔC_t is linear in t, i.e., $\Delta C_t = \beta_0 + \beta_1 t$ for some β_0 and β_1.*

Model Development

When modeling the number of hives needed to pollinate a 20-acre field, we consider a square field with an area of 20-acres, or 81,000 m^2. Since the field is a square, the side length of this field is $\sqrt{81000\ m^2} = 284.6\ m$. Additionally, from the problem description, we are given that honeybees typically stay within 6,000 m of their hive. For simplicity, we assume that the bees are evenly pollinating the flowers within this 6,000 m. Figure 12 shows the area of one beehive compared to size of the field.

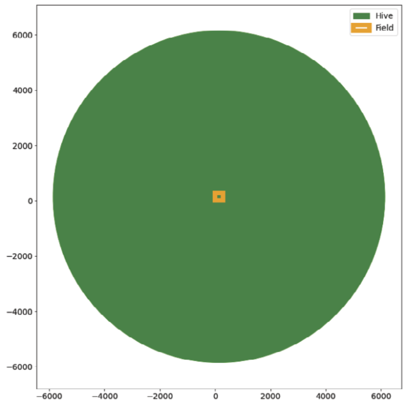


Figure 12: Size of the field compared to the coverage of a beehive when bees typically stay within 6,000 m of the hive. All axis labels are in meters.

From Figure 12, it is clear that the coverage of the beehive is far greater than the size of the field. However, it is important to note that this does not mean only one hive is enough to pollinate the entire field. Since bees will spread out evenly across the 6 km radius around the hive, many plants pollinated will be outside of the 20-acre field that needs to be pollinated. Thus, only a small fraction of the crops pollinated will be in the 20-acre field. Since the coverage of the hive is far greater than the size of the field, the placement of the hive within the field does not matter, as the hive will cover the entire field.

When developing our model to determine the number of hives needed to pollinate a 20-acre field, we considered four main factors: field area (A_{field}), the area needed by each flower (A_{flower}), the number of flowers pollinated per bee each day (D), and the typical max distance a bee travels from its hive (R). These variables and their default values are shown in Table 3. Additionally, our model is built off of the population model from the previous problem, so we also utilize the factors in Table 1.

Our model builds off the bee population model and predicts the percentage of plants on the 20-acre field that are pollinated on any given day. The number of hives that produces a result where the average number of plants pollinated per day is 100% is called the best model. Similar to the bee population model, the simulation begins at the start of the honeybee active season (April 1st), which we call $t = 0$ Recall from the bee population model that $P(t)$ represents the population of a honeybee colony at time t, where t is the number of days since the beginning of the active honeybee season.

4.5 Evaluating robustness of CO2 prediction models (Question 3b)

We have established a univariate model combining the seasonal-component-free Holt-Winters method and bootstrap (HW+Bootstrap) and a bivariate model of linear regression with ARIMA errors (Trend+ARIMA) for predicting future CO2 level changes. To answer Question 3b, we assessed the models' forecast ability from two aspects and figured out what factors may affect the ability.

1) Quantify the model ability based on scenarios configured with different training dataset and test set created from historical data, as listed in Table 1.

By comparing S1, S3 and S5, we can investigate changes of the performance of HW+Bootstrap univariate model as the forecast moves forward. Likewise, comparison among S2, S4 and S6 helps investigate that of Trend+ARIMA model.

Alternatively, comparing S1 and S2, S3 and S4, and S5 and S6, allows us to look into the performance distinction between the two models with the same configuration of training and test datasets.

Model performance is measured by correlation coefficient (R) and RMSE.

2) Comparing the forecast ability of the models for 2021-2100. Since no true data are present for 2021-2100, we examined predicted temperature changes estimated from both HW+Bootstrap and Trend+ARIMA models in terms of projected trend, evolution pattern over time and uncertainty propagation as time progresses.

Table 1 Scenarios for evaluating models' forecasting ability. 'HW+Bootstrap' represents the method combining the seasonal-component-free olt-Winters method and bootstrapped residuals and 'Trend+ARIMA' represents the method combining the linear regression and ARIMA errors.

# of scenario	Training set	Testing set	Method
S1	1959-1979	1980-2000	HW+Bootstrap
S2	1959-1979	1980-2000	Trend+ARIMA
S3	1959-1979	2001-2021	HW+Bootstrap
S4	1959-1979	2001-2021	Trend+ARIMA
S5	1980-2000	2001-2021	HW+Bootstrap
S6	1980-2000	2001-2021	Trend+ARIMA

4.4 Sensitivity Analysis

Lastly, we perform sensitivity analysis to ensure the stability and reliability of our model's outcomes. We alter the arbitrary variables r and daily increase in not pollinated flowers by:

- Changing the lower bound for r by $\pm 10\%$, so that $r \in [0.75 \pm 0.075, 1]$.
- Changing the factor of daily increase in not pollinated flowers by $\pm 10\%$, so that the increase becomes $(\frac{1}{50} \pm \frac{1}{500})(k_0 - k)$.

We randomly change the two variables in the $\pm 10\%$ range for 50 random combinations, then test and compare the outcomes on lillies, whose $k_0 = 180000, s = 157, \Delta t = 35$ and require 10 hives according to previous results in Table 4.2. The tested results are plotted as shown in Figure 4.5.

We notice that all tests evaluate an outcome of $H = 9$. Therefore, we consider the results to be stable and less prone to the arbitrary parameters mentioned above.

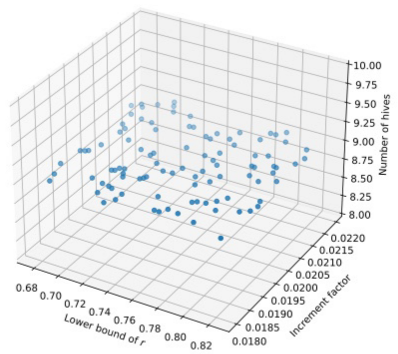


Figure 4.5: Sensitivity Analysis of HDM

Strengths & Weaknesses

Strength:
One strength of the model is that it relies on basic calculations that can be done on the average calculator or Desmos with data coming from basic data from a preexisting reputable source. The model utilizes concepts of regression to develop the fitted lines of the data, something most graphing calculators are capable of.

Strength:
Another strength of the model includes the relatively easy visualization of the model through the regression line. The overall line also provides a holistic representation of the growth of CO2 and temperature far into the future. It allows us to view if CO2 levels are ever stagnant, how fast it's growing, without complex visualization techniques.

Weakness:
One weakness of the exponential regression model is that it predicts that the data will progress in the same direction and method for the future data points. It's very unlikely that the CO2 levels will progress in exactly the same manner they did when the Industrial Revolution first began to this point. For one, top contributors of the CO2 levels had no care for the effect of their actions towards the environment. Though such is somewhat unfortunately still true today, it's at a much smaller scale than it once was in the beginning of human industrial development. On top of that, there are many other confounding factors that point to the fact that CO2 growth may not be as consistent and upward to be fitted nicely into a simple exponential equation.

Weakness
The model is only based off of only the past 53 years worth of data, even though the Industrial Revolution as early as the mid-19th century. Currently, the model is based on only the past 50 years, yet humans have had a large impact on CO2 levels for almost another century before. To sufficiently model the growth of CO2 and the relationship humanity has had on those levels over time, more data is simply needed. However, it is understandable that such data may not have been available as CO2 PPM levels were only measured starting from the mid-20th century.

Example 9: Strengths and Limitations Problem B
Team #13010, North Carolina School of Science and Mathematics, NC, USA

8 Conclusion

To better understand future global warming trends, we created several models to predict CO_2 concentration levels and global land-ocean temperatures.

We first analyzed the CO_2 levels data and discovered that the March 2003 increase of CO_2 resulted in a larger increase than observed over any previous 10-year period instead of the March 2004 increase that the NOAA claimed. We then fit four different models, Holt's linear trend, ARI(8, 2), IMA(2, 8), and ARIMA(3, 2, 3), on annual March averages of CO_2 levels and used these models to forecast the future CO_2 levels up to 2100. Holt's linear trend predicted that CO_2 levels will reach 685 ppm by 2132, ARI(8, 2) predicted by 2091, IMA(2, 8) predicted by 2082, and ARIMA(3, 2, 3) by 2083. Our results disagreed with the OECD's claim that the CO_2 concentration level will reach 685 ppm by 2050. After comparing the model performance statistics of all four models, the best model for describing patterns in the CO_2 levels was the ARI(8,2) model and hence we think this model is the most accurate one. Our sensitivity analysis verified that the ARI model is robust and model performance is not greatly affected by the amount of data or using data from different months.

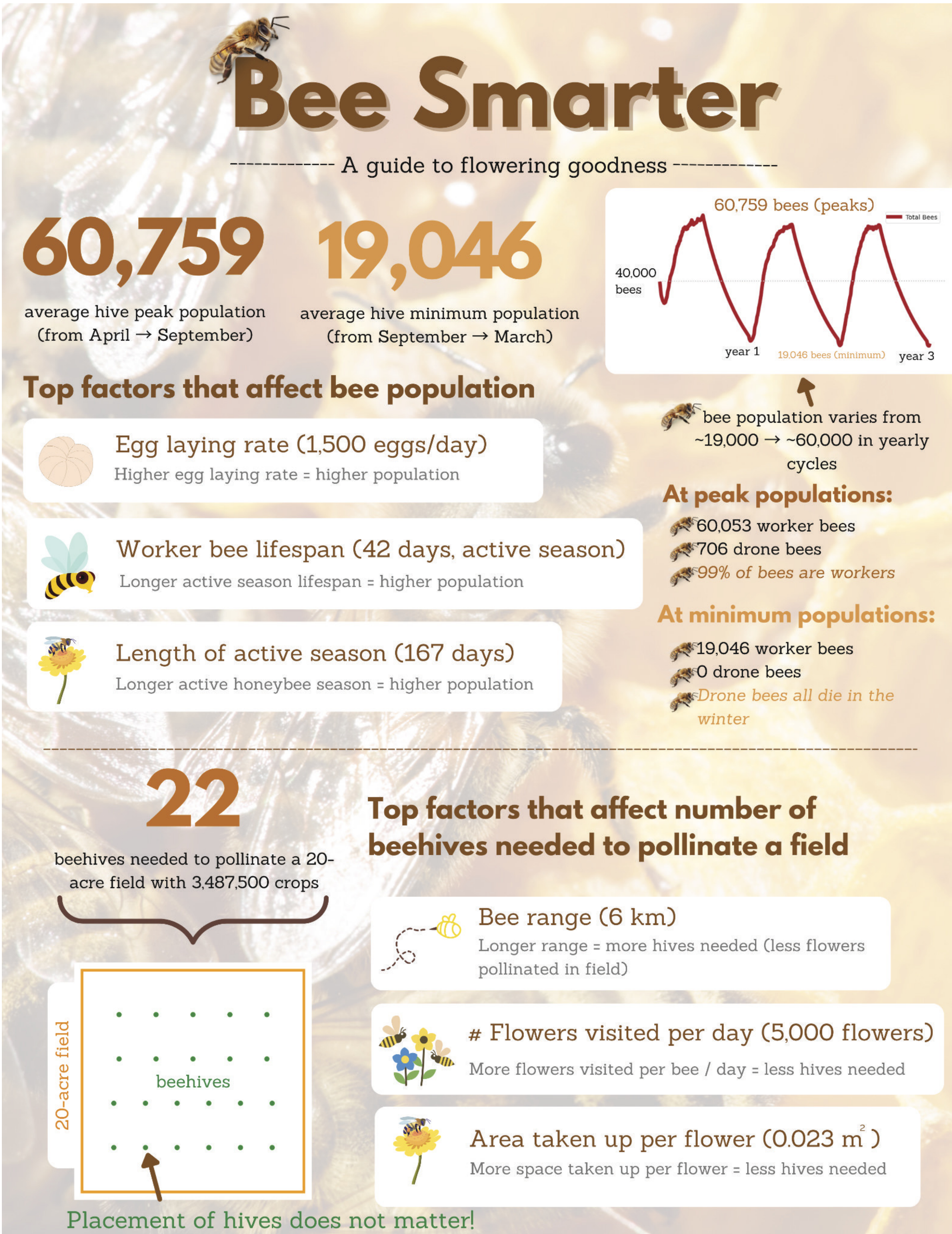
In order to forecast future global temperatures, we then fit an ARIMA(3, 1, 3) model on global annual mean temperature changes. ARIMA(3, 1, 3) forecasted that the global average temperature will change by 1.25°C in 2038, 1.5°C in 2052, and 2°C in 2081 when compared to the average temperature from 1951 to 1980.

Finally, to investigate the relationship between CO_2 and temperature, we first examined Pearson's correlation and determined that there's a strong positive relationship between CO_2 levels and temperature. We then used VAR(5) to model the temporal causal relationship between CO_2 concentration levels and global temperatures. After performing Granger causality tests with our VAR model, we found that there is a strong Granger causal relationship from CO_2 levels to global temperatures and a weak Granger causal relationship from global temperatures to CO_2 levels. VAR(5) forecasts that in 2050, the CO_2 level will be 512.85 ppm and the global temperature will be 2.01°C. We determined that all predictions up to 2100 from VAR(5) should be reliable because of the narrow width of corresponding 95% confidence intervals. The concerns of our VAR(5) model include a small training data size and that other factors besides temperature and CO_2 levels were not considered in the model. To test the seriousness of these concerns and the robustness of the VAR model, we performed several sensitivity analyses. We verified that the model is effective with smaller data sizes in terms of forecasting and identifying Granger causal relationships. In addition, the predictions of the model and strong Granger causal relationships are robust and not greatly affected by changing the sampling frequency to monthly or the inclusion of new factors such as CH_4, N_2O, and SF_6.

All our analysis and model results confirm that CO_2 levels and global temperatures are steadily increasing and verify that CO_2 levels greatly influence global warming. To protect the environment and lessen global warming, CO_2 emis-sions should be greatly reduced and new policies to reduce emissions should be enacted, such as restricting the use of fossil fuels.

Example 10: Conclusion Problem B
Team 12911, Ward Melville High School, NY, USA

Non-Technical Infographic



Example 11: InfoGraphic Problem A
Team#13014, North Carolina School of Science and Mathematics, NC, USA

Global Warming:

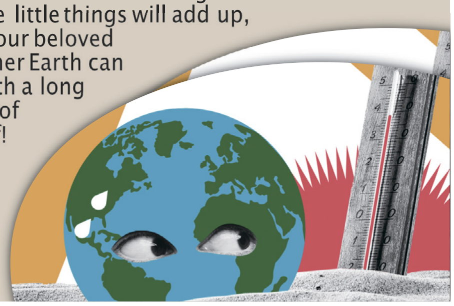
How Serious Is It, And What Should We Do?

Global warming and climate change just never seem to be out of the spotlight and things only seem to get worse. As the main cause of rising temperatures, greenhouse gas wraps around the Earth's atmosphere like a thick blanket, making our Earth 1.1°C warmer than it was in the late 1800s, triggering intense droughts, severe fires, flooding, and catastrophic storms. According to our recent study, things aren't exactly turning to the bright side: while our team has disproved the claim made by the Organization for Economic Co-Operations and Development (OECD) that the CO_2 level will double by 2050, our predictions still show that the CO_2 level will double by 2100, and our global temperature is likely to increase 2 degrees Celsius by 2076.

To better understand the current situation and find some possible solutions to it, we looked into the data on carbon dioxide levels from 1959 to 2021 in an attempt to predict the future. Trying out various mathematical models, we found that by the end of the 21st century, the carbon dioxide concentration in the atmosphere will reach about 685 PPM, which is a situation not seen for more than 50 million years! Additionally, the CO_2 level is undergoing a quadratic increase.

Then, we also looked at the increase in temperature. Fitting various models, we found that the global temperature is expected to increase 1 degree by 2058, 1.5 degrees by 2065, and 2 degrees by 2076. We also investigated the quantitative relationship between carbon dioxide levels and global temperature. The results show a direct, linear relationship: the global temperature is bound to increase when more CO_2 is released into the atmosphere.

Based on our findings, the situation is considerably stern. So, we will recommend a few solutions that could reduce carbon dioxide levels and let a fresh breeze in. Firstly, governments can introduce carbon dioxide emissions trading in industrial production, in which they allocate a limited amount of pollution permits to firms. Firms must buy permits to pollute and cannot pollute an amount exceeding the permit. This policy creates a price incentive for firms to reduce carbon emissions. Secondly, carbon sequestration could be employed to allow carbon dioxide to be absorbed naturally. Some areas on earth, known as "carbon sinks", such as soil and land covered by certain vegetation, can absorb carbon dioxide, However, a large amount of them is now used to plant crops, which do not have such ability. Therefore, we could redesign agricultural land use and grow more non-crop plants. Besides, trees are also a kind of carbon sink, so planting more trees may help as well. Finally, there are lots of little things we can do in our everyday lives: cycle to work instead of driving, finish your meals instead of throwing them away and producing more carbon release, wear a jacket instead of blasting the air con. All these little things will add up, and our beloved mother Earth can breath a long sign of relief!



Example 12: Non-Technical Article Problem B
Team #12505, Wuhan Britain-China School, Hubei, China