Problem Title	2023 MathWorks M3 Challenge Summary Sheet	Team Number
The Growth of E-Bike Use		**16623**

Executive Brief

In the past few years, electric bicycles, also known as e-bikes, have become increasingly popular. E-bikes, especially in urban areas, have provided users with more efficient and sustainable means of transportation. To further understand what the future of e-bike use will look like, we sought to answer three main questions: (1) what will e-bike sale look like in the future in two and five years? (2) what factors do e-bike users take into account and find most important when deciding to purchase an e-bike?, and (3) how does the increased use of e-bikes impact carbon emissions, traffic congestion, and health and wellness?

To predict the growth in e-bike sales, we created a model that uses the various states of the economy, such as the national inflation rate, the price of gasoline, and the growing awareness of the benefits of e-bike usage, to predict future sales of e-bikes. This is also known as an aggregate demand model. Our model predicts an exponential increase in e-bike sales. For example, we predict that 914,930 will use an e-bike in 2 years and 25,244,255 in 5 years. In the United Kingdom, the number of people is 1,676,209 and 24,942,993 in 2 and 5 years respectively.

Next, we aimed to determine the underlying causes of the growth of e-bike sales. We already have a general understanding of the characteristics that impact transportation choices and environmental sustainability; however, we do not know which characteristics influence e-bike buyers' decisions the most. To approach this, we created a model that predicts how gas prices, income levels, sustainability, and the ability to safely bike in a local region, also known as *bikeability*, affect e-bike sales. Our model showed that gas prices affected e-bike sales by 27.7%, income levels by 19%, sustainability by 13.8%, and *bikeability* by 78%. Because bikeability is a significant contributor, we believe that adapting policy to improve biking infrastructure would sway additional people to purchase e-bikes.

As people switch from unsustainable modes of transportation to e-bikes, larger societal impacts may occur. In order to assess this, we created three distinct models to understand how carbon dioxide emissions, transportation, and health and wellness are affected by e-bike usage using our first model for predicted e-bike sales. We found that carbon dioxide emissions will be reduced by 400 metric tons from 2022 to 2025 and average wait times in traffic will decrease significantly. In addition to the impacts on our roads, e-biking will also positively impact the health of everyday citizens. Our model predicts that over time, 61% of e-bike users will be within the age group of 50 to 65 years. This age group is usually unable to use regular bikes due to their physical health, so the accessibility and simplicity of e-bikes will allow these people to exercise more.

Moving forward, we hope that e-bikes can be adopted to improve the quality of life and sustainability in our world.

Contents

1 Introduction

1.1 Problem Restatement

1.1.1 Question 1

Recent economic trends have shown significant growth for the sales of electric bikes in the United States and the United Kingdom. Question 1 includes predicting the growth of e-bike sales in the United States and the United Kingdom and calculating the predicted sales two and five years in the future.

1.1.2 Question 2

With the rising sales of of e-bikes, its necessary to learn the causes of growth, Question 2 asks to identify the underlying factors attributed to the growth of e-bikes and pinpoint which factors do or do not have significant role in the growth of e-bikes.

1.1.3 Question 3

The growth of electric bikes could potentially impact the usage of other modes of transportation. Question 3 includes computing these impacts on carbon emissions, traffic congestion, and health and wellness.

1.2 Assumptions

1. **The majority of e-bike users live in urban areas.**

   According to the United States Census Bureau, although 97% of land is rural, only 19% of Americans reside in those locations. On the other hand, 3% of urban land holds 81% of the population [1] . Therefore, the majority of the population–and those who use e-bikes–live in urban areas.

2. **Awareness of the benefits of e-biking is growing exponentially.**

   In the past few years, e-bike sales have been growing exponentially, showing that awareness of e-biking is similarly increasing as well [2].

3. **Gas prices and e-bikes are substitute goods.**

   According to Google Search Trends for the search term "electric bicycle", the number of times this phrase has been searched is increasing [3]. It is assumed that this increase in interest of e-bikes is partly due to the recent increase in gas prices [2].

4. **Aggregate results are representative of local results.**

   To make the model more robust, we assume that characteristics of local regions mimic those of the entire United States.

5. **As demand of e-bikes increases, the supply chain will keep up.**

   It is assumed that as e-bikes become increasingly popular, there will be sufficient supplies to manufacture e-bikes and meet demands.

6. **Those who own e-bikes use them as transportation for about half of their trips.**

   About half of e-bike trips in the United States replaced trips formerly made by cars [4].

7. **Each e-bike user uses at most one e-bike.**

   It is assumed that e-bike owners only own and use one e-bike.

8. **The age distribution of e-bike users will remain constant over time.**

   Each age group each have their own benefits to using e-bikes. Thus, age groups will maintain the same distribution over time.

2 Question 1: The Road Ahead

2.1 Problem Analysis

There are many factors that have an effect on e-bike sales and consumerism. In order to fully model future e-bike sales, we decided to create an aggregate demand model. Demand models aid in understanding various states of the economy and the need for certain products [5]. As part of our model parameters, we take into account the per-capita disposable personal income growth rate, gas prices, inflation rates, and the increasing awareness of e-bike benefits in both the United States and the United Kingdom. We believe that these factors will have the greatest impact on the demand and sales of e-bikes.

From our research, we determined that gross national income growth rate plays a key role in global consumerism, meaning that as income rates change, e-bike sales will also change. Because our model represents temporal data, we included general information about inflation, as changes in value of the economy can affect consumerism and the purchasing of products such as e-bikes. Furthermore, online research suggests that e-bike sales and gasoline prices are substitute goods, so we ensured that our model accounts for the changes in gasoline prices. Therefore, as gas prices increase, e-bike sales should increase as well. Lastly, we considered the fact that the awareness of e-bike benefits is increasing exponentially, which further affects e-bike sales as more people will be purchasing them.

2.2 Variables

Variable	Symbol	Description
Financial Accounting Period	k	One accounting period lasts one year. Number of years after 2010.
National Disposable Income Probability Factor	ρ	Sampled from a normal Gaussian distribution.
National Disposable Income Growth Rate	a	The national disposable income growth rate in the US and UK is 1.017 and 1.008 respectively [2].
National Inflation Rate	b	The average inflation rate between 2010 and 2020 is 1.03 in the US and 1.02 in the UK [6].
Initial Gas Cost	G_0	The cost of gas in 2010 in the US and UK was 2.78 and 5.33 USD per gallon, respectively [2].
Initial Gross National Income	Y_0	The initial GNI (Gross National Income) in the US and UK in 2010 was 15,000 and 2,600 billion USD, respectively [2].

2.3 Aggregate Demand Model

Our model predicts the growth in e-bike sales. As with most products, we assume that the sale of e-bikes follows the standard supply and demand model.

For modeling demand, we take into account characteristics that affect the desire for e-bikes. Additionally, because national income affects product demand, we decided to model income growth in the United States.

2.3.1 Demand Curve

The demand curve of e-bikes is affected by three primary components: the nation's gross national income, gas prices, and overall sustainability opinions.

First, we developed a linear difference equation to model the national income of the United States where a represents the growth factor and k represents the accounting period relative to the initial national income. It models the recursive nature of income growth, where future values are influenced by past values.

$$Y_k = aY_{k-1} + b \tag{1}$$

To account for stochastic, or random, variability, we introduced a factor of ρ to represent the probability of negative events, such as a natural disaster or a global pandemic. The solution to the above difference equation is modeled below with a stochastic factor.

$$Y_k = \left(1 - \frac{1}{3^\rho}\right) a^k \left(Y_0 - \frac{b}{1-a}\right) + \frac{b}{1-a} \tag{2}$$

Next, we defined a gas price function to model the impact of gas prices on e-bike demand. Because we assume that gas prices and e-bikes are substitute goods, increasing gas prices would increase e-bike demand. Thus, the model is defined where G_k is the gas price influence at accounting period k. To account for seasonal and non-linear gas price trends, we used a damped trigonometric scaling factor to moderate variability. Gas prices are seasonal due to different blends of gasoline between winter and summer months [7].

$$G_k = G_0 \left(k^2 \sin \left(\frac{1}{2\pi} k \right) \right) \tag{3}$$

Lastly, we combined environmental perceptions and bikeability scores of the respective country to an exponential growth function. Let E_k be the environmental and bikeability function with respect to accounting period k with growth factor R.

$$E_k = E_0 e^{Rk} \tag{4}$$

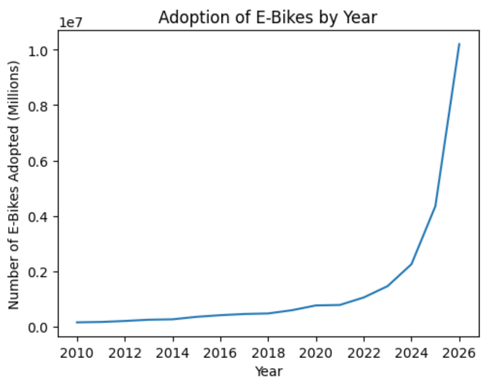
To create an encompassing aggregate demand curve with the three factors defined above, we created a weighted average function with scaling constants ψ and λ. Let D_x be the demand curve with respect to accounting period k. The aggregate demand curve represents the nation's quantity demanded of e-bikes with respect to price per e-bike.

$$D_k = \lambda Y_k + \psi G_k + E_k \tag{5}$$

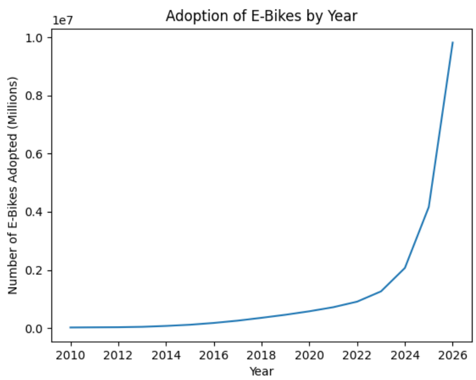
2.4 Results

Using initial values for the U.S. and the U.K. as outlined in the variables table above, the number of e-bikes adopted each year is graphed in Figure 1. In simulating U.S. e-bike adoption, we set $\lambda = 10^{-5}$ and $\psi = 10^3$ to scale the function appropriately.

The shape of the prediction graph is most similar to an exponential growth model. In years 2015-2022, the number of adopted e-bikes closely resembles empirically collected data [2]. Future predictions for adopted e-bikes will thus grow at a rate proportional to the current adoption of e-bikes.



(a) The Adoption of E-Bikes by Year in the US



(b) The Adoption of E-Bikes by Year in the UK

Figure 1: **Adoption of E-Bikes by Year**

From the defined e-bike adoption function, we predicted the following results for the **U.S. e-bike market**.

- In two years (2025), there will be **1,676,209 new e-bikes** will be adopted.
- In five years (2028), there will be **24,942,993 new e-bikes** will be adopted.

From the defined e-bike adoption function, we predicted the following results for the **U.K. e-bike market**.

- In two years (2025), there will be **914,950 new e-bikes** will be adopted.
- In five years (2028), there will be **25,244,755 new e-bikes** will be adopted.

2.5 Sensitivity Analysis

We conducted sensitivity analysis to determine which factors is most influential the adoption of e-bikes. We varied the factors of income growth, inflation rates, and initial gas costs by 5, 10, -5, and -10%. Finally, we recorded the percent change in the function output, $\vec{D}$, over interval from 2010 to 2027. The average percentage change was calculated using the below formula.

$$\overline{\text{Percentage Change}} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\vec{D_{original}} - \vec{D_{new}}}{\vec{D_{new}}}\right| \tag{6}$$

Testing Date Range: 2010 - 2027 (USA Region)						
% Change	**Income Growth**	*% Change*	**Inflation**	*% Change*	**Initial Gas Cost**	*% Change*
0	1.02	**0**	1.03	**0**	2.78	**0**
5	1.07	**19.83**	1.08	**-1.47**	2.92	**6.84**
10	1.12	**37.16**	1.13	**1.4**	3.06	**13.23**
-5	0.97	**-14.59**	0.98	**2.32**	2.64	**4.34**
-10	0.92	**-19.69**	0.93	**4.58**	2.5	**-11.2**

Figure 2: **Percentage Change Sensitivity Analysis Table**.

From our percentage change sensitivity analysis, it was determined that **income growth** and **initial gas costs** influence the adoption of e-bikes the most, respectively.

2.6 Model Discussion

2.6.1 Strengths

1. Our e-bike adoption aggregate demand model uses **country-wide parameters**.

   With the understanding of how certain factors can affect e-bike sales, we were able to create an overall model that can be applied to any country with the addition of statistics specific to that country. In this case, we used our model as well as specific statistics in both the Unites States and the United Kingdom in order to model e-bike sales for both countries.

2. Our e-bike adoption model can generalize with **numerous factors**.

   By using numerous factors such as inflation rates, disposable income rate growth rate, and more, we are able to account for many components in forming our analysis on e-bike adoption. Specifically, by considering the fact that awareness of e-bike benefits is currently growing exponentially, we were also able to encompass people's views on personal well-being and environmental benefits of purchasing e-bikes.

2.6.2 Limitations

1. Our model begins with accounting period 0 **in year 2010**.

   Since our model begins with initial value in 2010, our model represents what e-bike sales would look like using only data from 2010 and later. However, using data from before 2010 could make our model an even better representation .

2. Our model **does not account for every possible factor** that could affect e-bike sales.

   Although not included, there are certain factors that could still play a role in e-bike sales such as certain government policies as well as factors caused by outside countries. Some examples of this include interest rate, fiscal policy, international trade, and supply shortages. With the addition of these factors, our model could be more accurate.

3. Our model **does not have an asymptotic limit**.

   Our model does not include an asymptotic limit and would approach infinity. This makes our model more accurate in the short-term and would make inaccurate predictions in the distant future.

3 Question 2: Shifting Gears

3.1 Problem Analysis

The decision to purchase an e-bicycle is one that is impacted by many factors which vary in importance from person to person. Although each e-bike user purchases their bike for a different reason, we wanted to create a model that can convey and compare how much each factor plays a role in the choice of purchasing an e-bike.

In order to understand which factors affect a persons decision to buy an e-bike as well as how the factors can affect each other, we decided to use past studies that included surveys asking respondents to rate which factors were most important to them when making their decision [8].

Using this information, we decided to create a Hidden Markov Model in order to calculate which factors were most important in e-bike riders; therefore, indicating which factors most affect overall e-bike sales.

3.2 Variables

Variable	Symbol	Description
States	P_n	The internal "hidden" parameters of the model which are *Gas Prices*, *Income Levels*, *Bikeability*, and *Sustainability*.
Symbols	E_m	The external final parameters of the model which are **Buying** or **Not Buying** an e-bike.
Number of states	n	The number of hidden states or parameters, determined to be 4.
Number of symbols	m	The number of final outputs that could be made, which was 2 in this model.
Transition Matrix	A	A nxn probability matrix correlating hidden states.
Emissions Matrix	B	An nxm probability matrix correlating hidden states with symbols.

3.3 Hidden Markov Model

When approaching this problem, we wanted a method to consider individual decisions and their impact on the growth of e-bike adoption. Thus, we decided to consider a Hidden Markov Model (HMM) to be able to consider state transitions such as making the decision in buying or not buying an e-bike. The HMM differs from traditional Markov Chain models because they incorporate hidden states, or intermediate steps, to obtaining the final states known as symbols. Rather than using deep learning technologies, which are computationally expensive and do not allow internal analysis, we used HMMs to make an adaptable and highly scalable probabilistic model.

(Note: As an extension of a Markov Chain Model, the HMM follows the Markov Assumption, such that each future symbol will only rely on the previous state.)

When considering any Markov model, the probabilities correlating factors is the most crucial component. We determined four major variables that most heavily impacted the decision to buy an e-bike: (1) *Local Gas Prices*, (2) the *Bikeablility* of the region, (3) the consumer's *Income Levels*, (4) the *Sustainability* benefits. For example, one must consider whether their local region is well suited for biking or if it is even worth purchasing an e-bike given external factors.

3.3.1 Hidden State Definitions

The hidden state parameters are as listed below.

- *Gas Prices* — the local gas prices for a consumer.
- *Income Levels* — the amount of expendable money a consumer can use after accounting for cost of living.
- *Bikeability* — how safe and comfortable a consumer feels when biking through their local areas.

- *Sustainability* — the significance a consumer places on environmental concerns.

3.3.2 Matrix Parameterization

The first step of implementing a Hidden Markov Model is determining the probability, or relationships, between the variables in our nxn transition matrix A. After investigating cyclist behavior and the adoption of e-bikes, the matrix was filled such that for every row i and column j, a_{ij} represents the probability of moving from state i to state j. Factors such as *Gas Prices* and *Income Levels* are highly associated, thus their associated probability is greater. Our transition matrix A is defined below. Note that the sum of the probabilities of each row sum to 1.

$$\mathbf{A} = \begin{array}{c c} & \begin{array}{cccc} G & I & B & S \end{array} \\ \begin{array}{c} G \\ I \\ B \\ S \end{array} & \begin{pmatrix} 0 & 0.5 & 0.2 & 0.3 \\ 0.5 & 0 & 0.4 & 0.1 \\ 0.2 & 0.4 & 0 & 0.4 \\ 0.3 & 0.1 & 0.4 & 0 \end{pmatrix} \end{array} \tag{7}$$

Next, we determined the relationships between hidden states of the transition matrix A. Because transition matrix A is symmetric because the probability of one state on another is equivalent to the probability of the second state on the first. The values along the diagonal of A such that a_{ij} where $i = j$ are evaluated as 0 because one factor can not impact itself.

— Between G and I, the amount of gas money impacts the amount of expendable money significantly, with gas prices taking a greater portion of lower incomes.

— Between G and B, the price of gas is least impacted by the bikeability of a region, leading to a probability value of 0.2.

— Between G and S, the price of gas is somewhat impacted by the sustainability of the user, as leading to a probability value of 0.2.

— Between I and B, the amount of expendable money does have an impact on the bikeability, with a greater income leading to better infrastructure where one lives.

— Between I and S, there is a small correlation, with greater expendable money meaning one can spend more on being sustainable.

— Between S and B, sustainability has a considerable impact on bikeability, with sustainable people seeking out places to live with a more sustainable infrastructure like bike lanes.

$$\mathbf{B} = \begin{array}{c c} & \begin{array}{cc} Y & N \end{array} \\ \begin{array}{c} G \\ I \\ B \\ S \end{array} & \begin{pmatrix} 0.60 & 0.40 \\ 0.50 & 0.50 \\ 0.80 & 0.20 \\ 0.55 & 0.45 \end{pmatrix} \end{array} \tag{8}$$

Emission matrix B represents the probabilities of moving from state G, I, B, or S to purchasing or not purchasing an e-bike. The values for probability in Matrix B were obtained by considering survey results where respondents ranked their criteria for buying an electric bike. When the difference between the probability of buying a bike and not buying a bike is zero (0.5 vs 0.5), the criteria does not significantly impact the decision. 79% of respondents indicated that bikeability was a major factor that influenced their decision, hence the probability split was 0.8 to 0.2.

A graphical representation of our HMM is shown in Figure 3. We used the *PyDTMC* Python library to implement the transition and emission matrices in the HMM model as shown in Appendix 2. With the *PyDTMC* library, we were able to visualize, simulate, and make predictions using the HMM. The library allowed us to find eigenvectors and orthogonally diagonalize the transition matrix to make predictions. Our HMM initializes with an initial probability state equal to the transition matrix.

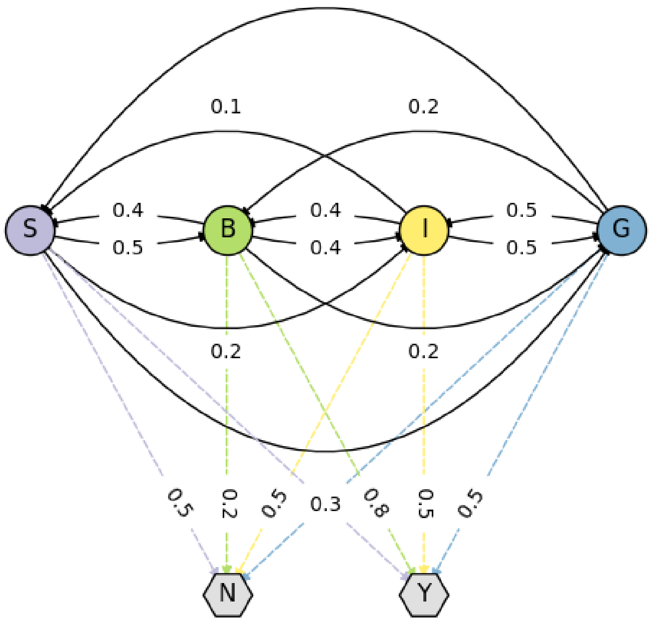


Figure 3: **A graph representation of the HMM model.** Edges between nodes are relationships with probability weights. The G, I, B, and S nodes are hidden states. N and Y are output symbols.

3.4 Sensitivity Analysis

In order to conduct sensitivity analysis on our developed HMM, we used the Sobol sensitivity analysis technique to determine which factors are most influential on people's decisions when adopting e-bikes in their daily lives. We chose Sobol analysis techniques rather than traditional percentage change analysis to better account for larger sample sizes and greater changes in the transition matrices [9]. Traditional percentage change analysis using tables and evaluating outputs does not take variance of inputs and outputs into account, thus, the Sobol analysis is favorable for this task.

In Sobol analysis, we used the *SAlib* Python library to create samples and run tests on the generated samples [10]. First, we generated $1,000$ samples of new transition matrices where each a_{ij} could be varied from 0 to at most 0.1. Thus, a change in a_{ij} would represent the sensitivity in parameter of row i.

Then, we simulated our HMM model for 5,000 iterations on each sample transition matrix A in order to reach the equilibrium state (A^{5000}). As states are passed recursively through the HMM, the state eventually converges to a state known as the equilibrium state. The equilibrium state is represented The number of times that the HMM model reached the final symbol that one would purchase an e-bike ("Yes" symbol) was recorded for each sample A in output vector Y^T.

The first-order Sobol sensitivity index S_1, is defined as followed for sample X_i, decomposition of variance VAR function, and output vector Y^T. Each parameter has its own Sobol index, in which a greater value indicates the parameter has a higher impact on the output of the HMM model.

$$S_1 = \frac{\text{VAR}_{X_i}}{\text{VAR}(Y^T)} \tag{9}$$

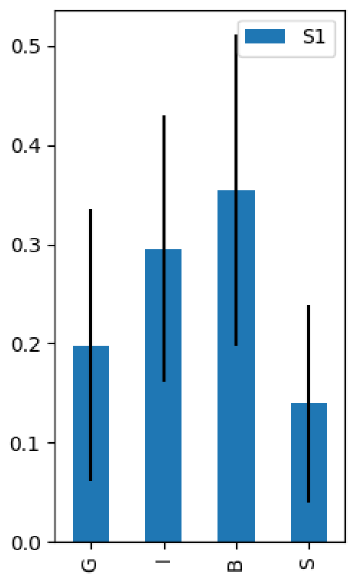


Figure 4: S_1 **Sobol Index for G, I, B, and S hidden parameters.**

This index signifies the direct impact the change of sample X_i has on output Y^T averaged over the combined variations of other samples. In Figure 4, the S_1 Sobol Index was calculated for each of the 4 hidden parameters. It was determined that since the *Bikeability* parameter had the greatest Sobol index, it has the greatest impact on individual decisions on purchasing e-bikes.

The ranked order (from greatest impact to least) of the four hidden parameters is: *Bikeability*, *Income Level*, *Gas Prices*, and *Sustainability*. The relative impact of these parameters can also be visualized in Figure 5.

Thus, the **most vital factor in impacting e-bike adoption is the *Bikeability* and *Income Levels***. The **least impactful factor is *Sustainability***.

3.5 Discussion

The factors determined to most affect e-bike adoption growth from the HMM model closely resembles that of the aggregate demand model from the previous question. Further strengths and limitations of our HMM model and sensitivity analysis are discussed below.

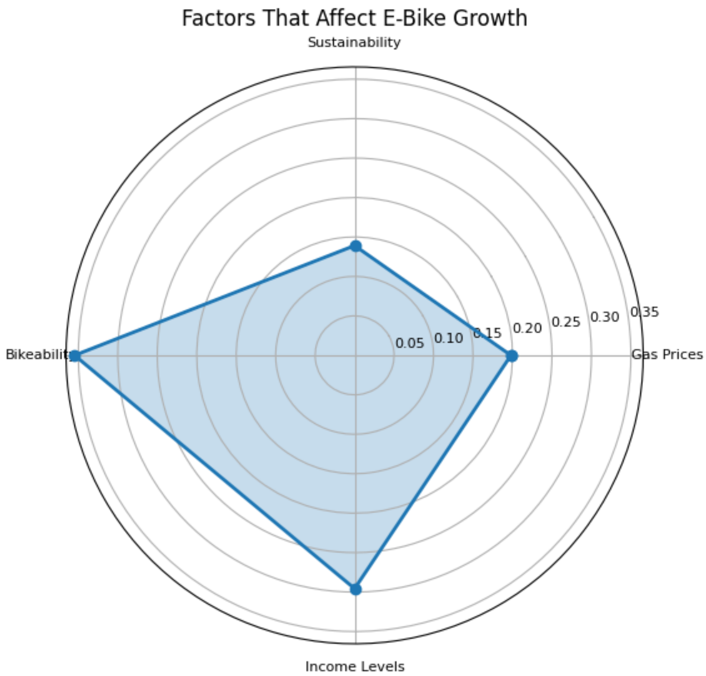


Figure 5: **Radar Chart of our HMM model.**

3.5.1 Strengths

1. Our HMM model is **location independent**.

   Our model does not depend on location, despite utilizing factors like gas prices and bikeability. Since the model does not pass in values for those parameters, instead calculating the equilibrium state by analyzing the relationships between the factors and the final decision, the actual location does not impact the output.

2. Our HMM model uses **probabalistic modeling**.

   Probabalistic modeling introduces variability that is similar to the real world. It allows the model to capture and incorporate in a structured way their insights into the risks and uncertainties they face.

3. Our HMM model is **individualized** and **highly scalable**.

   Our HMM model emphasizes individualistic decision-making by using a probabalistic transition matrix. Furthermore, this HMM can be scaled to increase depth by incorporating more variables and relationships.

4. The Sobol sensitivity analysis methodology utilized has **greater accuracy**.

   The Sobol method of sampling and sensitivity analysis requires a smaller number of iterations to attain similar accuracy levels to the Monte Carlo and Latin HyperCube sampling methods. Compared to Monte Carlo sampling, the Sobol methods require $1/20^{\text{th}}$ of the number of iterations, making it more computationally efficient [11].

3.5.2 Limitations

1. Our HMM model does not consider **health and personal interest factors**.

   There was a lack of data involving the factor of demand and interest in e-bikes due to the novelty and "coolness". We also could not include factors such as psychological health due to a lack of quantifiable information [12].

2. Our HMM model may not have the most **precise probability states**.

   Since the most up-to-date survey information is not available, the probability states that were used in our matrices may not be perfect.

4 Question 3: Off the Chain

4.1 Problem Analysis

As the number of e-bikes increases in the United States and the United Kingdom, the impact of e-bikes on traffic, carbon emissions, and personal well-being must be considered as they could play a role in e-bike sales and the decision to purchase an e-bike. In order to analyze how these components of everyday life are affected by the use of e-bikes, we decided to create three different agent-based models for each factor.

Specifically, for carbon emission, an important thing we kept in mind was that e-bikes do not omit at much carbon dioxide as other forms of transportation. For our model of traffic, we acknowledged that the use of e-bikes will affect the use of other vehicles on the road, which will further affect traffic. Lastly, our health and wellness model took into account specific age groups and how each age group has different wellness goals.

4.2 Carbon Emission Model

Given the alarming increase of CO_2 concentration in the atmosphere, with scientists from the National Oceanographic and Atmospheric Administration (NOAA) and Scripps Institution of Oceanography (SIO) noting peak concentration levels of 421 ppm in May 2022 and the OECD predicting CO_2 levels to reach 685 ppm by 2050, it is essential to cut carbon emissions, starting from where it is the greatest problem: cars.

In the United States, cars are used as a primary mode of transportation for most. Compared to a car which emits about 404 grams of CO_2 per mile, e-bikes emit around 22 grams of CO_2 per mile [13] [14]. Thus, the benefits of switching to e-bikes are unparalleled.

In order to model the the loss in carbon emissions with people transitioning to e-bikes in the United States, we modeled the total number of e-bikes using the model from Question 1.

By 2025, the total number of e-bikes adopted from 2022 to 2025 in the U.S. can be modeled by the following expression.

$$\text{Number of E-Bikes Adopted} = \int_{12}^{17} D_k dk \approx 17,899,931 \tag{10}$$

Thus, we will have decreased a total of

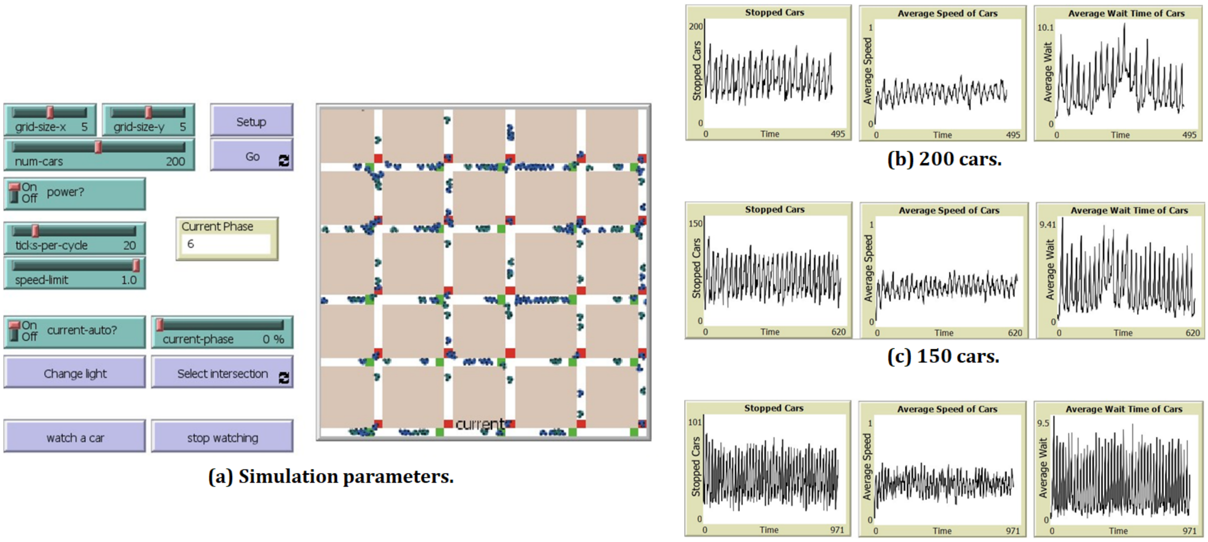
$$17,899,931 \text{ E-bikes} \cdot \frac{22\text{g}}{1\text{E-bike}} \cdot \frac{1\text{kg}}{1000\text{g}} \approx 393,798 \text{ kg of carbon emissions} \tag{11}$$

This reduction in carbon emissions is equivalent to **over 400 metric tons of carbon dioxide released** over a three year span in the U.S. This impact can greatly improve the health of our environments and our ocean, which we depend so much on.

4.3 Traffic Model

The overall goal of transferring ridership to e-bikes to reduce the number of cars on the road. Thus, we hypothesized that decreased car ridership would decrease traffic on the roads.

To test this hypothesis, we developed a traffic grid simulation using NetLogo agent-based modeling software [15]. Using a grid like traffic pattern, we varied the number of cars in the traffic scenario to observe effects on average wait times and car traffic within the simulation. We set the number of cars in the grid-like environment to 100, 150, and 200. Then, we measured some of the key traffic indicators, such as wait times, of the simulation in the NetLogo environment. The simulation environment can be seen in Figure 6.



(a) Simulation parameters.

(b) 200 cars.

(c) 150 cars.

(d) 100 cars.

Figure 6: **NetLogo simulation of traffic conditions.**

It can be seen from Figure 6 (b), (c), and (d), that when the number of cars decreases, the average wait time of cars also decreases. This signifies that a decrease in car ridership lead to **lesser traffic conditions**. By decreasing the number of cars in the grid simulation by 50%, we were able to decrease the average waiting times by around 20%.

4.4 Health and Wellness Model

Older people and those with disabilities have a difficult time finding ways to stay active. Strenuous activities such as biking are not possible for people who fall into these categories. Common motivations for e-bike users are to facilitate mobility in the presence of medical conditions or disabilities [16].

With the total number of e-bikes adopted from 2022 to 2025 estimated to be around 17,899,931, the age distribution curve is shown in Figure 7.

Studies have shown that e-bike usage has many health benefits: it could address diabetes, obesity, and facilitate cardiovascular exercise [8]. E-bike usage at their highest power setting were equivalent to the the metabolic exercise as walking. E-bike usage in the standard power setting had slightly lower metabolic exercise as walking. With the majority of current and future users from the older generations, more e-bike usage will benefit these users who cannot use conventional bicycles due to physical limitations.

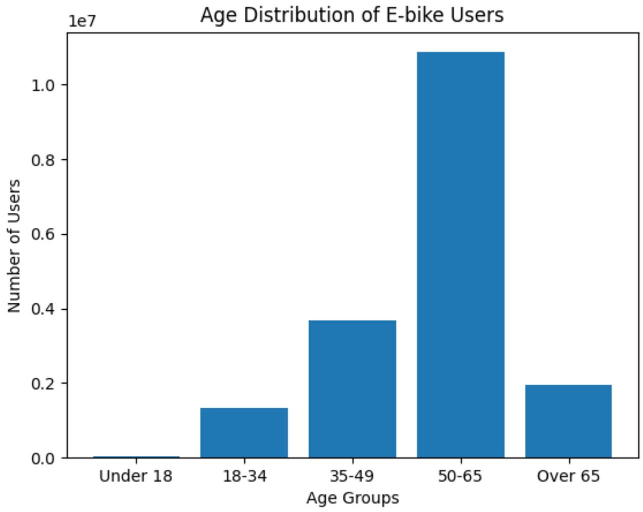


Figure 7: **The age distribution of e-bike users in the United States [17]**.

5 Conclusion

5.1 Our Model

Electric bikes, also known as e-bikes, are an affordable and accessible method of transportation. Their role in clean transportation has been solidified by their growing popularity in both the United States and the United Kingdom. Our predictive aggregate demand model uses the parameters of gross national income (GNI), nationwide gas prices, and sustainability related concerns. After 2 years, the model predicted 1,676,209 new e-bike sales and after 5 years, the model predicted 24,942,993 new e-bike sales in the United States. In the United Kingdom, the model predicted 914,950 e-bike sales after 2 years and 25,244,755 e-bike sales after 5 years from 2023. The parameters for the U.K. were adjusted to accounts for any regional differences, however the results were remarkably similar with very little variance between the graphs.

In order to evaluate the impact of certain factors on the ultimate decision to buy an electric bicycle, we implemented a Hidden Markov Model (HMM) that accepted the parameters from our previous model. Since the decision could vary from person to person, we decided to modify our variables to ensure each variable was more individualized and pertinent. Since we used near identical criteria for both models, we anticipated similar results for the models. The major difference between our models, aside from methodology was the output. The first model was a predictive model that projected the number of e-bikes produced in a given number of years, while the second model aimed to maximize the accuracy of the decision to buy an e-bike given the probability of parameters. Despite the difference in expected model, the outputs were remarkably similar. Both models indicated that the income factor was highly significant both on the individual and national representation of the decision to buy an e-bike.

To model the impacts increased e-bike demand might have, we evaluated the effects it would have on health, the environment, and traffic. Using the growth model constructed for the first question, we calculated the total amount of e-bike sales from 2022 to 2025. Using data about emissions, we discovered that e-bike sales would reduce global emissions by 400 metric tons of CO_2. We also analyzed the age distributions of users of e-bikes, and showed that increasing e-bike usage would lead to better health for people with trouble using conventional bikes due to age or disabilities considerably. We then used an agent-based model to simulate e-bike growth on traffic conditions and found out that when the number of cars is reduced by e-bikes, the wait times due to traffic reduced by 20%. Overall, e-bikes have a net positive influence on the world, and the increasing growth and popularity of these green vehicles will be better for our daily lives and our environment.

5.2 Future Extensions

Although these models accurately predicted responses that relatively neared anticipated results, each model was far from perfect. In the future, we would detail an aggregate supply curve, much similar to an aggregate demand curve, that encompasses factors such as supply-chain shortages, materials costs, and more. Furthermore, we would like to add further parameters to our growth model, such as an asymptotic plateau for e-bike sales and how other forms of public transportation would contribute to demand.

In addition to considering Markov models for exploring decision problems, additional factors for determining e-bike purchases should be considered such as battery life and health and wellness. Furthermore, more novel prediction architectures such as decision trees, deep learning classifiers, and temporal-based predictors should be explored.

6 Appendices

6.1 Appendix 1. Supply Demand Model Python Script

./demandsupply.py

```python
#Model 1 - Simulation Program for Adoption of E-Bikes

#Import Functions
import matplotlib.pyplot as plt
import random
import numpy as np

# Income Factor Function
def income_factor(k, y0, a, b):
    return (1 + 0.1 * random.randint(0, 3)) * (a ** k) * (y0 - b / (1 - a)) + b / (1 - a)

# Gas Factor Function
def gas_factor(k, g0):
    return g0 * (k**2 * np.sin(1 / (2 * np.pi) * k))

# Bikeability Factor Function

def bike_factor(k, e0):
    return e0 * np.exp(1.001 * k)

# Gross National Income Constant
USA_GNI = 2e9

# CHANGE THESE AND CHECK PERCENTAGE DIFFERENCE
USA_Income_Growth = 1.017
USA_Inflation = 1.03
INITIAL_GAS_COST = 2.78

# Get total adoption of cyclists with scaling factors applied
def get_total(i):
    return income_factor(i, USA_GNI, USA_Income_Growth, USA_Inflation) / 1e5 + gas_factor(i, INITIAL_GAS_COST) * 1e3 + bike_factor(i, 1)

#Simulate starting from 2010 to 2010 + YearsToSimulate
YearsToSimulate = 17
numBikesSold = []
for i in range(YearsToSimulate):
    numBikesSold.append(get_total(i))

print(sum(numBikesSold[-12:]))
#Graph plot
plt.plot([(2010 + i) for i in range(YearsToSimulate)], numBikesSold)
plt.xlabel("Year")
plt.ylabel("Number of E-Bikes Adopted (Millions)")
plt.title("Adoption of E-Bikes by Year")
plt.show()

#Sensitivity Analysis
# CHANGE THESE AND CHECK PERCENTAGE DIFFERENCE
# USA_Inflation = 2.502
```

```

# YearsToSimulate = 17
# numBikesSoldNew = []
# for i in range(YearsToSimulate):
#     numBikesSoldNew.append(get_total(i))

# percentage_difference = [(numBikesSoldNew[i] -
numBikesSold[i])/(numBikesSold[i]) * 100 for i in range(YearsToSimulate)]
# print(np.mean(percentage_difference))
```

6.2 Appendix 2. Hidden Markov Model Python Script

./HMMModel.py

```python
# Import related libraries
import numpy as np
import pydtmc as dt
import matplotlib.pyplot as plt
from SALib.sample import saltelli
from SALib.analyze import sobol
import random

# Set interactive mode in matplotlib
plt.ion()

# Initialize problem parameters
problem = {

    'num_vars': 4,

    'names': ['G', 'I', 'B', 'S'],

    'bounds': [[-0.1, 0.1],

               [-0.1, 0.1],

               [-0.1, 0.1],

               [-0.1, 0.1]

               ]

}

# Sample values using Sobol sampling method
param_values = saltelli.sample(problem, 100)

# Initialize output vector of 0s
Y = np.zeros([param_values.shape[0]])

# Loop through sampled values
for i, X in enumerate(param_values):

    # Initialize transition matrix
    pMatrix = np.array([[0, 0.5, 0.2, 0.3],

                        [0.5, 0, 0.4, 0.1],

                        [0.2, 0.4, 0, 0.4],

                        [0.3, 0.2, 0.5, 0]])

    # Initialize emission matrix with altered parameters for Sobol analysis
    eMatrix = np.array([[0.6+X[0], 0.4-X[0]], [0.5+X[1], 0.5-X[1]],
                       [0.8+X[2], 0.2-X[2]], [0.55+X[3], 0.45-X[3]]])

    # Initialize states
    states = ['G', 'I', 'B', 'S']
```

```

    # Initialize output symbols
    symbols = ['Y', 'N']

    # Initialize hidden markov model
    mc = dt.HiddenMarkovModel(pMatrix, eMatrix, states, symbols)

    # Calculate possible symbols and states after convergence to equillibrium state
    sim_states, sim_symbols = mc.simulate(5000, seed=1488)

    # Count number of Y for Solbol analysis
    Y[i] = sim_symbols.count('Y')

# Sobol analysis
Si = sobol.analyze(problem, Y)
print(Si)
Si.plot()
plt.pause(1000)
plt.show()

# S1 Yes: 0.19805434, 0.29564655, 0.35477548, 0.13911498
```

References

[1] U. C. Bureau, "One in Five Americans Live in Rural Areas." Section: Government.

[2] MathWorks Math Modeling Challenger, "Ride Like the Wind," 2023.

[3] "Google Trends."

[4] R. Buehler and J. Pucher, *Cycling for Sustainable Cities*. MIT Press, Feb. 2021. Google-Books-ID: fM_tDwAAQBAJ.

[5] "What Factors Affect the Increase in Aggregate Demand? | Indeed.com Canada."

[6] "U.K. Inflation Rate 1960-2023."

[7] "Why Gas Prices Rise Nearly Every Spring," Apr. 2021.

[8] A. Mayer, "Motivations and barriers to electric bike use in the U.S.: views from online forum participants," *International Journal of Urban Sustainable Development*, vol. 12, pp. 160–168, May 2020. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/19463138.2019.1672696.

[9] I. M. Sobol, "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Mathematics and Computers in Simulation*, vol. 55, pp. 271–280, Feb. 2001.

[10] W. Usher, J. Herman, C. Whealton, D. Hadka, Xantares, F. Rios, Bernardoct, C. Mutel, and J. V. Engelen, "Salib/Salib: Launch!," Oct. 2016.

[11] "Pi Day Comparison: Monte Carlo vs Latin HyperCube vs Sobol Sampling."

[12] E. Fishman and C. Cherry, "E-bikes in the Mainstream: Reviewing a Decade of Research," *Transport Reviews*, vol. 36, pp. 72–91, Jan. 2016. Publisher: Routledge _eprint: https://doi.org/10.1080/01441647.2015.1069907.

[13] O. US EPA, "Greenhouse Gas Emissions from a Typical Passenger Vehicle," Jan. 2016.

[14] M. McQueen, J. MacArthur, and C. Cherry, "The E-Bike Potential: Estimating regional e-bike impacts on greenhouse gas emissions," *Transportation Research Part D: Transport and Environment*, vol. 87, p. 102482, Oct. 2020.

[15] U. Wilensky, "NetLogo," 1999.

[16] J. MacArthur and C. Cherry, "A North American Survey of Electric Bicycle Owners | National Institute for Transportation and Communities," 2017.

[17] Z. Ling, C. Cherry, J. MacArthur, and J. Weinert, "Differences of Cycling Experiences and Perceptions between E-Bike and Bicycle Users in the United States," *Sustainability*, 2017.