CruiseAuto Project – Milestone 1B

ANSWER SHEET: Parameter Identification Brainstorming

Table of Contents

Part 1. Assignment Header

Section and Team ID: 011-03

Team Member Name	Purdue Career Account Login	Programmer Number	Detailed Description of the Work	Percent of Work
Justin Clark	Clar1062	2	Primary programmer on managing data noise and errors. Your code will ensure that the data are usable and will be accurate for parameter identification. This programmer helped with the graphing but mainly focused on writing Part 4 of this document	25%
Joao Soares	soaresj	3	Focus on programming the function to find the starting acceleration time and the time constant for each test. This programmer assisted programmer 1 with plotting the functions of the tests but mainly focused on working on part 2 of this document.	25%
Ethan Zhang	Zhan5173	1	Create the main function and create the data visualization models to display the data, analyze it and produce results. This programmer is also responsible for coordinating the subfunctions so they are organized and properly used in the main function. For M1B, this programmer worked on part 3 for examining the data of the tires and tests.	25%
Peter Teal	pteal	4	I am responsible for finding both the initial and final speeds and for finding the speed parameter. This programmer helped with organizing and graphing in part 3, but mainly worked on part 5 of this document.	25%

Part 2. Milestone 1A Feedback and Reflection

Strength: The main strengths in the our approach is the way we are splitting the larger data set up into smaller, more manageable data sets of the same category.

Limitation: The main weaknesses of the our approach could be the fact that we planned on using nine graphs, which could be a lot for the user to navigate and digest.

How could the feedback from M1A lead to improvements? Because our approach could cause the graphs to be hard to read, we plan on combining the graphs or creating different subplot combinations and using data cleaning to make the graphs more readable for the reader.

What concrete steps will you take to incorporate the M1A feedback to improve your data visualization and handling? We will find a method to clean the data to manage noise

Part 3. Examine the Data

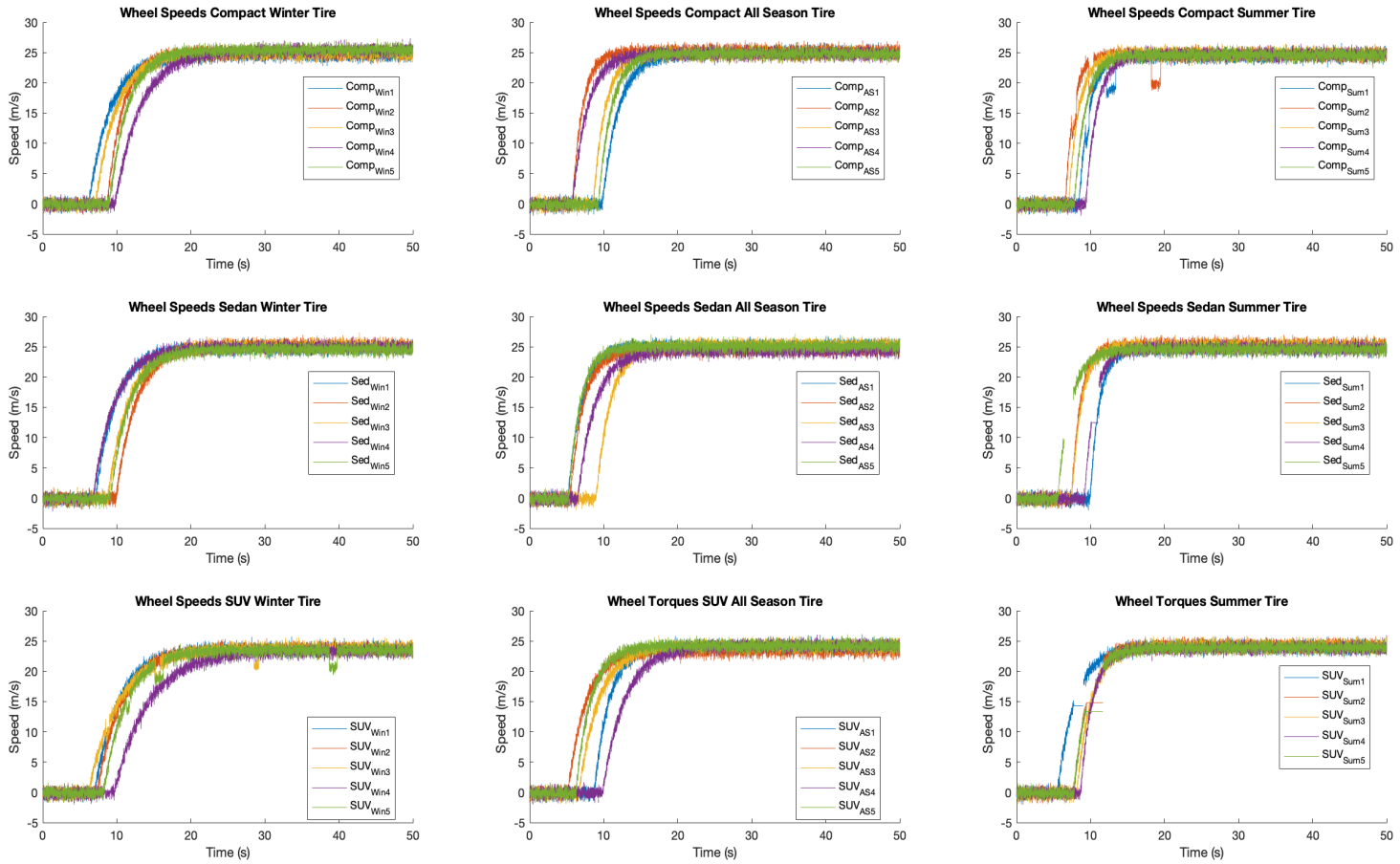

Figure 1: Nine separate graphs of wheel speeds comparing 5 tests per graph

Part 4. Brainstorm & Evaluate Coding Processes Using Experimental Data

Part 4a. Approaches to Managing Noise

Approach #1 to Managing Noise: To manage the noise in the data, we plan to make a loop that continuously takes an average of the data.

What is your evidence-based justification for your approach? If the data is averaged throughout it should be able to make a mostly linear model.

What is your expected difficulty for coding in MATLAB (low, medium, high)? medium

Why do you think it will be at the level of difficulty you indicated? Making the loop that averages the data repeatedly may be hard to create. The model should be easy to create so therefore the difficulty would be medium.

Approach #2 to Managing Noise: We can linearize the data using a nonlinear regression model (I think it's a logarithmic function) and then graph the regression model which should be a smooth line.

What is your evidence-based justification for your approach? The data is very noisy but if we were able to get a model that accurately represents the data it would completely remove the noise, and we could check this model's accuracy by finding its r^2 value.

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? We've already done many assignments on data linearization and regression, and we are all familiar with the process, though the function would be piecewise which makes it slightly more complicated, it still would not be very difficult

Part 4b. Approaches to Managing Error

Approach #1 to Managing Frozen Sensor Data Points: Using a for/if loop to evaluate the data and find all of the NAN and extra / meaningless data points (Wright, 2024) data points and then use a coefficient from the linear model to create a general model to find the missing data points at the given time.

What is your evidence-based justification for your approach? This way would be efficient as it runs through the whole data set and substitutes the NaN values with a value, which would be calculated the general model from the linearized model equation and inputting the x-value (time) at that certain point. This would be an appropriate method as it follows the same trend as the data model as a whole.

What is your expected difficulty for coding in MATLAB (low, medium, high)? Medium

Why do you think it will be at the level of difficulty you indicated? The difficulty would not be that high, however the identification of 'NaN' values and substituting them is not simple and does require a certain number of functions.

Approach #2 to Managing Dropped Data Points: We can linearize the data using a nonlinear regression model (logarithmic function) and then graph the regression model which should be a smooth line.

What is your evidence-based justification for your approach? The data is very noisy but if we were able to get a model that accurately represents the data it would completely remove the noise, and we could check this model's accuracy by finding its r^2 value.

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? We've already done many assignments on data linearization and regression, and we are all familiar with the process, though the function would be piecewise which makes it slightly more complicated, it still would not be very difficult

Approach #3 to Managing High/Low Reading Data Points: Using standard deviation, we can get rid of all high/low reading data points, outside of 2 standard deviations.

What is your evidence-based justification for your approach? In a data set usually about 95% of the data is within 2 standard deviations from the mean, so if we get rid of data points that are outside of 2 standard deviations we would ideally remove any outlier data points.

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? There are simple built-in equations in Matlab that can be used to find standard deviation

Part 5. Brainstorm Approaches to Parameter Identification

Part 5a. Approaches to Identifying Acceleration Start Time

Approach #1: Once the data is cleaned up, we can check the slope of the graph at various points and when the slope is greater than 0, we can assume that the car is accelerating.

What is your evidence-based justification for your approach? In a velocity/time graph the slope represents the acceleration of the object thus once we find a point with a nonzero slope, we know the car is accelerating

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? We can make a for loop that checks the slope of the line at every data point starting at 0 until we find a positive slope.

Approach #2: Use a moving average filter (Dey, 2018) to smooth the data, then scan from the beginning and locate the first index where the slope exceeds the smoothed baseline by say 5%

What is your evidence-based justification for your approach? The point at which the slope exceeds a certain negligible amount (in this case 5% though this number can be changed) is the point where the car would be accelerating i.e. the slope of a velocity time graph is acceleration

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? Once the data points are smoothened out by using a simple moving average formula for the velocity(Dey, 2018), the point where the slope or rate of change between two points would be very easy to spot since before the wheels accelerate there should be a slope very close to 0, so a simple while loop checking the slope of each point until a certain tolerance is exceeded could do this very easily.

$$\bar{p}_{\mathrm{SM}} = \frac{p_M + p_{M-1} + \cdots + p_{M-(n-1)}}{n}$$
$$= \frac{1}{n}\sum_{i=0}^{n-1} p_{M-i}$$

Figure 2: The math behind a simple moving average

Part 5b. Approaches to Identifying the Time Constant

Approach #1: Once the starting acceleration time is found, the velocity at that point can be used to solve for the time constant since the time constant enters the given differential equation at any time greater than or equal to the starting acceleration time

What is your evidence-based justification for your approach? The formula for velocity can be worked backwards using simple algebra once all of the other variables are known.

What is your expected difficulty for coding in MATLAB (low, medium, high)? Medium

Why do you think it will be at the level of difficulty you indicated? Making sure the algebra is correct may be difficult, also the other variables in the given equation need to be known as well

Approach #2: The initial equation given in the client memo pdf can be used to calculate the time constant as well, essentially once the data is smoothed out using the previously mentioned methods in part 4 a. and b. the formula given by $y_T = y_L + 0.632(y_H - y_L)$ to calculate the time constant is this the time that the wheels are traveling at "y_T" m/s. The actual time constant itself would be found by calculating the aforementioned y_T value then finding what time this speed is reached, and that is the time constant.

What is your evidence-based justification for your approach? The explanation given in the client memo defines the time constant as the amount of time it takes the wheel to reach the given (formula above) value for speed, thus we could find the final and initial speeds of the wheel, plug them into the formula and solve for the time constant

What is your expected difficulty for coding in MATLAB (low, medium, high)? medium/low

Why do you think it will be at the level of difficulty you indicated? This approach wouldn't be completely easy as we still must find the initial and final velocities, though it would be easier than the first approach as algebra is a lot simpler.

Part 5c. Approaches to Identifying Initial & Final Speed

Approach #1 – Initial Speed: Once the start time for the acceleration is calculated using a smoother/ averaged data set, we can just average the velocities before this time

What is your evidence-based justification for your approach? Theoretically before the wheels start accelerating, they should be travelling at roughly the same speed across the given time interval, obviously this is not the case but just averaging all the velocity point values from t = 0 to the acceleration time should give a rough estimate as to a singular value that can be used as the starting speed.

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? Once the starting time is calculated we can just use the mean function in Matlab to calculate the average velocity using all of the velocity data points that are before that given time.

Approach #2 – Final Speed: Once the data is roughly smoothed out using a moving average as previously mentioned, we can take the average velocity from say the last 5 seconds of the data set to generalize the final velocity

What is your evidence-based justification for your approach? The asymptotic nature of the velocity function results in the velocity reaching a theoretical maximum (barring any sensor irregularities) well before the end of the data set, thus we can safely assume that the velocity at the end of the data set is the highest (average) velocity the wheels will achieve

What is your expected difficulty for coding in MATLAB (low, medium, high)? low

Why do you think it will be at the level of difficulty you indicated? All we must do is average out the last 5 seconds of velocity data and that should give a fairly accurate result for the final velocity.

Part 6. References

Fiveable. (n.d.). 6.5 Adaptive cruise control. Fiveable autonomous vehicle systems review. https://library.fiveable.me/autonomous-vehicle-systems/unit-6/adaptive-cruise-control/study-guide/4bKoB3Zw9z1a1KCG

My Car Does What? (n.d.). Adaptive cruise control. My Car Does What? https://mycardoeswhat.org/deeper-learning/adaptive-cruise-control/

Wright, G. (2024, April 12). *What is noisy data?: Definition from TechTarget*. Search Business Analytics. https://www.techtarget.com/searchbusinessanalytics/definition/noisy-data

Brandon Dey, O. (2019, June 23). *A short summary of smoothing algorithms*. Open Data Science - Your News Source for AI, Machine Learning & more. https://opendatascience.com/a-short-summary-of-smoothing-algorithms/