The Kennesaw Journal of Undergraduate Research

Volume 11 | Issue 1 Article 2

2024

Using a Degree-based Network Model to Understand and Control Traffic Jams in the Atlanta Metropolitan Area

Ian Salamone-Lent
Kennesaw State University, isalamon@students.kennesaw.edu

Theresa Washington
Kennesaw State University, twashi63@students.kennesaw.edu

Asma Azizi
Kennesaw State University, aazizi@kennesaw.edu

Recommended Citation
Salamone-Lent, Ian; Washington, Theresa; and Azizi, Asma (2024) "Using a Degree-based Network Model to Understand and Control Traffic Jams in the Atlanta Metropolitan Area," *The Kennesaw Journal of Undergraduate Research*: Vol. 11: Iss. 1, Article 2.
DOI: https://doi.org/10.62915/2474-4921.1300
Available at: https://digitalcommons.kennesaw.edu/kjur/vol11/iss1/2

Using a Degree-Based Network Model to Understand and Control Traffic Jams in the Atlanta Metropolitan Area

Ian Salamone-Lent[1], Theresa Washington[2], Asma Azizi[1] (Faculty Advisor)

[1] Department of Mathematics, Kennesaw State University, Marietta, GA 30060

[2] Department of Mechanical Engineering, Kennesaw State University, Marietta, GA 30060

ABSTRACT. Traffic congestion is an enduring problem for major urban areas such as the metropolitan area of Atlanta, GA. Our goal was to understand the nature of traffic congestion patterns in the highway system of Cobb County in the metro area. We created a road network representative of the Cobb County highway system and then superimposed a degree-based Susceptible-Infectious-Recovered (SIR) model to simulate traffic congestion on that network. The model's parameters, propagation, and dissipation rates were estimated using empirical traffic data, which are vehicles' speed time series and speed limit of each road in the network. We then conducted a local sensitivity analysis of the model's key parameters, confirming each parameter's modest impact on the cumulative number of congested roads and the number of congested roads at peak time. Then, we used optimal control theory to identify the most effective control function in reducing traffic jams in Cobb County. Our findings showed that low levels of control did lower total congestion but mimicked the uncontrolled congestion behavior. However, increasing the level of control dropped congestion and changed its behavior by diminishing the peak. That is, the congestion time series monotonically decreased to zero. This paper provides additional evidence that traffic behavior can be accurately predicted by SIR modeling and suggests that there exists a precise level of traffic control that eliminates traffic congestion propagation behaviors, given enough resources.

KEYWORDS. SIR model, ODEs, traffic congestion, degree-based network, Cobb County

1. Introduction

Traffic congestion is a common problem for high population density American cities, which historically have higher car ownership rates and distance driven per capita than other western cities. Among American cities, Atlanta, GA, and its immediate surrounding area have rated exceptionally high in these metrics [1]. This metropolitan area has a history of sprawling suburban development based on roads and interstates. Efforts to improve public transportation have historically been met with strong opposition from residents and have entrenched the city into further relying on its road system as it grew in population and economic activity [2]. Economic shifts in the last few decades have especially strained the transportation network of metro Atlanta, with travel volume vastly outpacing the urban development needed to support it and as a result increasing commute times. Even more, as the economy further relies on timely shipments and employee access, the penalty for traffic gridlock worsens. As traffic becomes a more significant problem with more and more potential to hamper the city's prosperity, solutions and the tools to find them are becoming more urgently needed.

After periods of reduced daily traffic like the COVID-19 pandemic, traffic can rebound in a way that produces higher than regular congestion, especially in areas with large proportions of transit and carpool commuters [3]. Wu et al. have previously shown that an epidemiological Susceptible-Infectious-Recovered (SIR) model can describe traffic congestion over a road system, sensitive to rates of traffic propagation, recovery from congestion, as well as changes in network topology, provided that the road network could be described as a complex road system [4]. The result indicated that developed, non-stop road networks behave like complex scale-free networks [5]. On the other hand, Jiang et al. have suggested that road networks have small-world characteristics; that is, the average road meets only a few roads, but the path be- tween any two roads is relatively short [6], which can result in a fast spread in SIR-type models superimposed on such networks [7]. Recovery from congestion causes instability of traffic flow [8], disrupting flow to lower levels than expected across the road network, with an aspect of this being unpredictability of congestion behavior depending on individual network structure [9, 10], which suggests the importance of superimposing models on real road networks. Saberi et al. refined the model and simulated real-world traffic congestion in multiple metropolitan areas [11]. Their model requires a large amount of traffic speed data, typically requiring the wide-scale ability to track individual vehicles. It is possible to simulate live traffic data for the city adequately using survey data detailing the travel habits of drivers. However, this simulated data has significant inaccuracies compared to recorded speed data [12].

Here, we used a SIR degree-based model to study traffic jams and ways to control them in Cobb County, which is one of the largest and fastest-growing counties in the Atlanta metropolitan area. We generated the road network of the county and then used hourly recorded speeds on each road in Cobb County's highway system [13] to estimate the parameters of our model. Then we used the model to predict how traffic proliferation behaves in a road network representing the Cobb County highway system. We conducted local sensitivity analysis of the model parameters and then used optimal control theory to control the propagation rate to minimize the impact of congestion. Our result confirmed the importance of controlling propagation rate by different means, such as improved signal timing and capacity increases.

DOI: 10.62915/2474-4921.1300

2. Methods

Road Network Structure and Data Collection

We have a network of N nodes, with every node i representing a road, and E edges, where every edge ij represents a connection between roads i and j, such as an intersection or highway ramp. The degree of node i, shown by d_i, is the number of nodes connected to node i, and $P(k)$ is the probability that a randomly drawn node from the network has degree k for $k = 1, ..., D$ where D is the maximum degree of the network: the degree of the most connected node in the network. To construct this road network, we focused on 133 roads (highways in Cobb County, namely Cobb Parkway, I-75, I-285, and I-575, and major/minor roads connected to these highways) as the nodes of the network, and we recorded all connected roads for each of these 133 roads. Figure 1 shows the map of the studied area (left panel) and the giant component of the generated network (right panel).

We gathered recorded speed data manually from a television station in Atlanta, Georgia, called WSB-TV [13]. This station periodically posts the traffic speeds of major city roads via traffic cameras. We recorded the speeds every 10 minutes during the interval [3 pm, 8 pm] for two weekdays, resulting in two speed data points for each road for each 10 minute subinterval. Then we averaged over these points to report the speed at each given time interval on each of the 133 roads. We also recorded speed limits of the relevant roads from documents posted by the Cobb County Board of Commissioners.

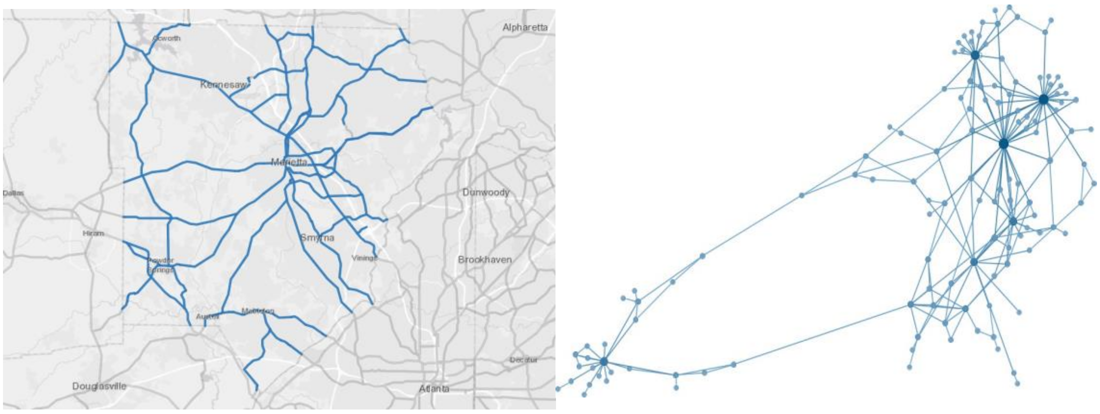

FIGURE 1. (Left) Map of the Cobb County highway system, (Right) Network representing 133 highways and connected roads in Cobb County, Georgia. Nodes represent individual roads, and edges represent connections between roads, including intersections and ramps.

Degree-Based Network Model

In order to classify a road at a given time t as congested or not, we recorded each road's registered speed limit and actual speed dependent on time. For node i corresponding to the i^{th} road, we represent its speed at time t as $v_i(t)$, and the speed limit as $\overline{v_i}$. We then defined $\lambda_i(t)$ as the ratio between speed at time t and registered speed limit, as shown in Equation 1 [11]:

$$\lambda_i(t) = \frac{v_i(t)}{\overline{v_i}} \tag{1}$$

The congestion threshold $\rho \in [0, 1]$ classifies a given node as either congested or not by the following rule in Equation 2:

$$\begin{aligned} &\lambda_i(t) < \rho \rightarrow \textit{the road i is congested} \\ &\lambda_i(t) \geq \rho \rightarrow \textit{the road i is not congested} \end{aligned} \tag{2}$$

In an ideal world, $\lambda_i(t) \leq 1$, but because $\overline{v_i}$ is the registered speed limit of the road i, not its maximum attainable speed, it is possible for $\lambda_i(t)$ to exceed 1 in the real world. Nevertheless, we assumed the threshold value of ρ to be less than one. The definition in Equation 2 means that if $\lambda_i(t) < \rho$, then the vehicles in the road i are moving less than ρ fraction of speed limit $\overline{v_i}$, which can be a sign of congestion in the road i.

To incorporate the heterogeneity of the road network structure into the SIR model, we used a degree-based network model [14]. The state variable $I_k(t)$ represents the number of congested roads with degree k at time t, $S_k(t)$ is the number of roads not yet congested with degree k at time t, and $R_k(t)$ is the number of recovered roads with degree k at time t. These integer-valued state variables are approximated by real-valued variables that are the solution of the following ordinary differential equation model in Equation 3 [15]:

$$\begin{aligned} \frac{dS_k}{dt} &= -\beta k \theta(t) S_k(t) \\ \frac{dI_k}{dt} &= \beta k \theta(t) S_k(t) - \gamma I_k(t) \\ \frac{dR_k}{dt} &= \gamma I_k(t) \end{aligned} \tag{3}$$

For $k = 1, ..., D$ the parameter β is the propagation rate that measures the rate at which traffic on one road causes congestion on another connected road, and γ is the dissipation rate that measures the average time needed for a congested road to become uncongested. In Equation 4 The function $\theta(t)$ describes the probability of being in contact with a congested road for any randomly selected road on the network:

$$\theta(t) = \frac{\sum_{k'=1}^{D} k' P(k') I_{k'}}{\sum_{k'=1}^{D} k' P(k') N_{k'}} \tag{4}$$

$N_{k'}$ is the total number of roads with degree k', and $P(k')$ is the probability that a randomly drawn road from the network has degree k'.

Controlling Traffic Using Optimal Control Theory

The modeling framework in Equation 4 can be used to identify propagation-optimal control strategies to minimize the total number of congested roads. We improved the modeled system in Equation 4 by controlling the parameter β such that demand for travel in the network is metered by different means. That is, since β represents the rate at which traffic spreads from congested roads to uncongested roads, we improved the model by accounting for strategies meant to prevent or curtail traffic once it has occurred. The exact strategy represented by the controlled model is abstract. However, examples of real-world propagation controls could be optimization of traffic signaling, increased road capacity, building additional connected roads for traffic to overflow onto, or greater availability of public transport to lower the demand for personal automobiles. Defining $U(t) \in [U_{min}, U_{max}]$ as a control function, we have the revised model in Equation 5:

$$\begin{aligned} \frac{dI_k}{dt} &= (1 - U(t))\beta k\theta(t)(N_k - I_k(t) - R_k(t)) - \gamma I_k(t) \\ \frac{dR_k}{dt} &= \gamma I_k(t) \end{aligned} \tag{5}$$

A successful control scheme minimizes the cost of the controls while reducing the number of congested roads until the final time τ. In Equation 6 the control function $U(t)$ is optimal if it minimizes the objective functional defined as:

$$J(U) = \int_0^\tau f(t)dt = \int_0^\tau C_1 I(t) + \frac{C_2}{2} U(t)^2 dt \tag{6}$$

As given, $I(t) = \sum_{k=1}^{D} I_k(t)$ is total congested roads, and constants C_1 and C_2 are the costs of controls in time $[0, \tau]$. The objective function J accounts for the total cost of congestion. Total number of congested roads times cost of one congested road is $\int_0^\tau C_1 I(t)$, and total cost of controlling congestion is $\int_0^\tau \frac{C_2}{2} U(t)^2$. The power 2 in the second integrand guarantees the existence of the minimized function $U(t)$. Minimizing $J(U)$ subject to the system in Equation 5, we find the optimal solution in Equation 7:

$$U^*(t) = max\{U_{min}, min\{\frac{\beta}{C_2}\sum_{k=1}^{D}\frac{k\lambda_{I_k}(N_k - I_k^* - R_k^*)}{\sum_{j=1}^{D} jP_jN_j}, U_{max}\}\} \tag{7}$$

λ_{I_k} for $k = 1, \ldots, D$ are the solutions of the adjoint system [16]. In the Results section, we will examine the dynamics of the controlled model in Equation 5 for various values of relative cost, with $\frac{C_2}{C_1}$, representing the ratio of resources spent implementing the control.

3. Results

Parameter Values and Parameter Estimation

The definitions of all the variables and parameters of the proposed model are given in Table 1. While most parameter values have been taken from data, we must estimate the pair (β, γ): propagation and dissipation rates. We did this by finding the pair that the model best fits the real-world congestion data for a given ρ. To find the real-world congestion data, we calculated the ratio of the speed of the road at a given time t to the maximum speed of the road, both taken from the collected speed data. We then compared this ratio to the described congestion threshold ρ. If the ratio was less than ρ, we flagged the road as congested. That way, we could count the number of congested roads at a given time t, presented as cross data points in Figure 2. Then, we used MATLAB's fminsearch function [17] to minimize the error between the model's simulation and the data, as shown in Figure 2. Table 2 shows the estimated pair (β, γ) alongside their corresponding ρ threshold.

TABLE 1. Parameters of model, definitions, and baseline values.

Parameter	Definition	Value	Ref.
N	Number of roads	133	Data
D	Maximum degree of the road network	26	Network
$\bar{k}$	Average degree of the road network	2.647	Network
ρ	Congestion threshold	0.7, 0.8, 0.9, 1.0	Arbitrary
β	Propagation rate	Depends on ρ	Estimated
γ	Dissipation rate	Depends on ρ	Estimated

TABLE 2. Estimated model parameters β and γ for different values of congestion threshold ρ for 5-hour simulation results on the network.

ρ	β	γ
0.7	0.055	0.115
0.8	0.062	0.088
0.9	0.063	0.064
1	0.072	0.053

Sensitivity Analysis

We conducted a local sensitivity analysis to quantify the robustness and resilience of the model result (total number of congested roads and the number of congested roads at the peak time) to our assumptions about model parameters (the propagation rate β, dissipation rate γ, and congestion threshold ρ). This local sensitivity analysis measures the relative derivative of the model output q with respect to the model input p nearby its particular reference (baseline) values [18, 19]. We measured the relative local sensitivity by the sensitivity index of output q, with respect to input p defined as $S_p^q = \frac{p}{q} \times \frac{\partial q}{\partial p}\big|_{p=p^*}$ where p^* is the baseline value for parameter p. Sensitivity index measures the percentage change in an output q given the percentage change in an input p, that is, if p changes by x% then q will change by $S_p^q \times x$%.

To determine the effectiveness of input parameters β, γ or ρ, we compared the total number of congested roads (TI) and the number of congested roads at peak time (I_{max}) by varying each input parameter around their baseline values ($\beta^* = 0.055$, $\gamma^* = 0.115$ and $\rho^* = 0.8$) at a time while freezing other parameters at their baseline value in Table 1.

Figure 3 shows change in TI as propagation rate β or dissipation rate γ change from 0 to 0.25. The sensitivity indices $S_\beta^{TI} = 0.48$ and $S_\gamma^{TI} = -0.49$ suggest that TI will increase by around 0.5% for every additional 1% increase in β, and decrease by around 0.5% for every additional 1% increase in γ. We also observe a similar trend and result for the sensitivity of I_{max} to propagation and dissipation rates β and γ in Figure 3.

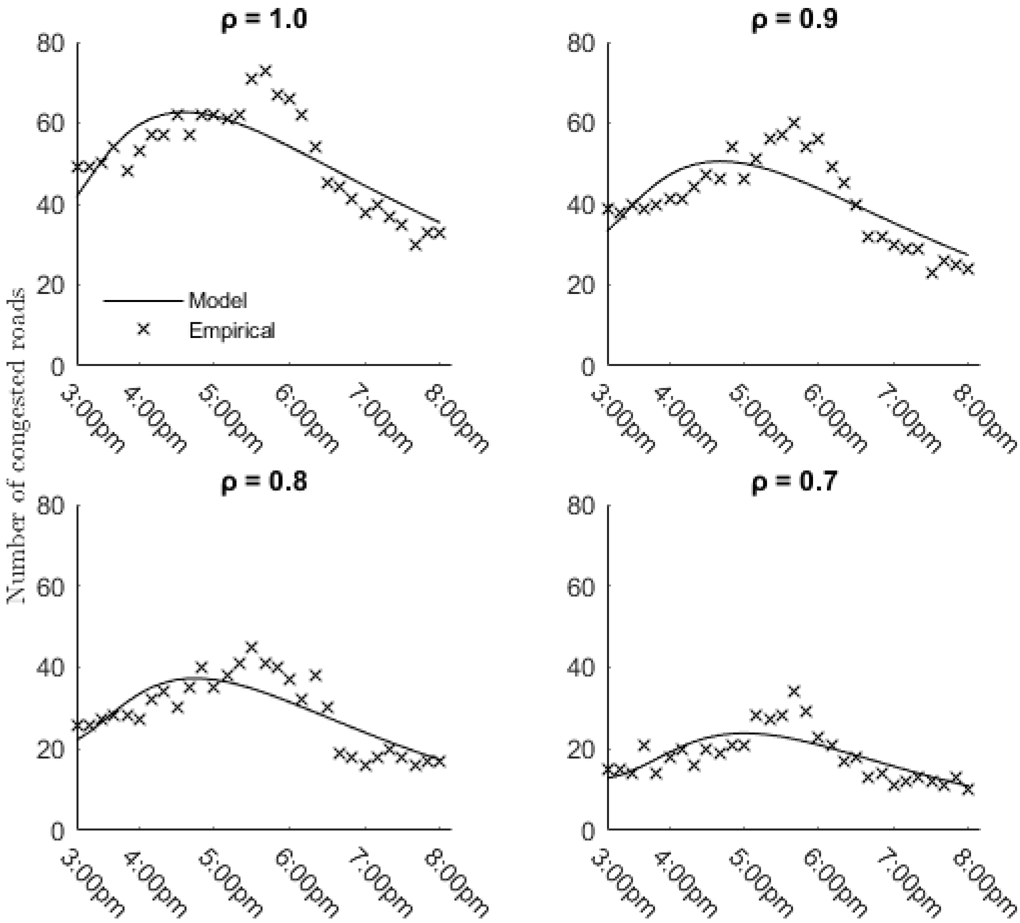

FIGURE 2. The number of congested roads (y-axis) with respect to time (x-axis) for the four different values of congestion threshold ρ. Curves indicate the model's result when given an initial condition dependent on the relevant value of ρ, and crosses indicate the calculated number of congested roads using speed data.

To observe how changing the congestion threshold ρ impacts the result, we compared TI and I_{max} by varying the threshold ρ between 0.7 and 1. Figure 3 shows a slight increment in TI as the threshold ρ increases from 0.7 to 1. The result suggests a sensitivity index of $S_{\rho}^{TI} = S_{\rho}^{Imax} \sim 0.4$ for $\rho^* = 0.8$, that is, every additional 1% of congestion threshold ρ causes a 0.4% increase in the total number of congested roads and the peak of congestion, which is not a significant increase, so the result would be robust with congestion definition as long as $\rho \in [0.7, 1]$.

Controlling Traffic Using Optimal Control Theory

Using the control model in Equation 5 and assuming ρ = 0.7, and therefore β = 0.0555 and γ = 0.115, we tested the controlling of propagation rate β for four different scenarios; a) baseline scenario no control (when relative cost $\frac{C_2}{C_1} = 0$), b) low control scenario when relative cost $\frac{C_2}{C_1}$ is high, so that controlling the system is costly ($\frac{C_2}{C_1} = 500$) and therefore the amount of implemented control would be low, c) medium control scenario when relative cost $\frac{C_2}{C_1}$ is medium, so that controlling the system is of moderate cost ($\frac{C_2}{C_1} = 50$) and thus a mid-level control can be implemented, and d) high control scenario when relative cost $\frac{C_2}{C_1}$ is small, so that controlling the system is inexpensive ($\frac{C_2}{C_1} = 5$) and therefore the amount of implemented control would be high. Figure 4 shows the number of congested roads over the selected time interval for the four scenarios. If we have finite resources and the cost of controlling β is relatively high, then our control has a relatively weak result, but as the cost of implementation lowers, we can afford a stronger control, which eventually results in propagation taking on a more controlled behavior; congestion monotonically decreases rather than increasing to some peak before dissipating, as represented by the medium control scenario. Any further decrease in cost allows us to strengthen the control and, thus, will cause a more drastic initial dissipation of congestion, represented by the high control scenario.

4. Discussion

The Atlanta metropolitan area is found to be one of the most congested urban areas in the United States, where drivers spent an average of 72 hours stuck in traffic in 2022, with a likely chance that that time stuck on the road will get even worse in 2023 [20]. To understand traffic in metro Atlanta, we collected traffic speed data from Cobb County's highway system, an area we selected for manageability. Data included traffic speed at a given time (average of vehicle's speed) and speed limit of each road, in the time interval of 3:00pm-8:00pm. Using this data, we calculated the ratio of current speed to the speed limit at a given time for each road, which allowed us to label streets as either congested or uncongested depending on if the ratio is below or above an arbitrary threshold $\rho \in [0, 1]$. We also generated a network representative of the Cobb County highway system, enabling us to develop a degree-based SIR model and simulate congestion propagation over the network. We calibrated the model's parameters, congestion, propagation, and congestion dissipation rates, to the observed data over the given time frame, resulting in an optimal set of parameters for each p threshold.

Conducting local sensitivity analysis on propagation and dissipation rates and congestion threshold confirmed that each parameter, when scaled, resulted in modest changes in the model's output, the cumulative number of congested roads, and the number of congested roads at peak time; see Figure 3. Finally, with the calibrated model, we implemented a control on propagation rate β. We showed that given a sufficiently large amount of resources, it could be made such that congestion dissipates monotonically over time. However, if resources are sparse, congestion propagation is lessened but maintains qualitative behavior; see Figure 4.

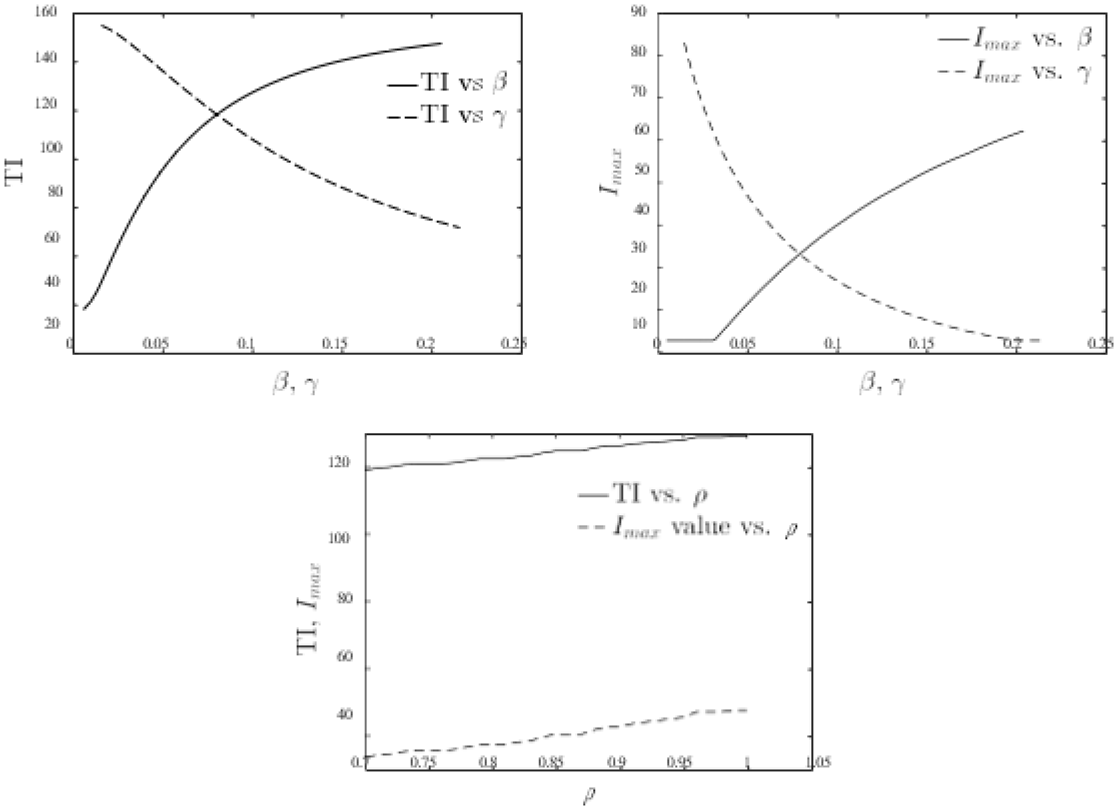

FIGURE 3. Local sensitivity plot of output q in y-axis with respect to input p in x-axis. Top left panel: Changing TI as function of propagation rate β (solid line), and dissipation rate γ (dashed line). Top right panel: Changing I_{max} as function of propagation rate β (solid line), and dissipation rate γ (dashed line). Bottom panel: Changing TI (solid line) or I_{max} (dashed line) as function of congestion threshold ρ.

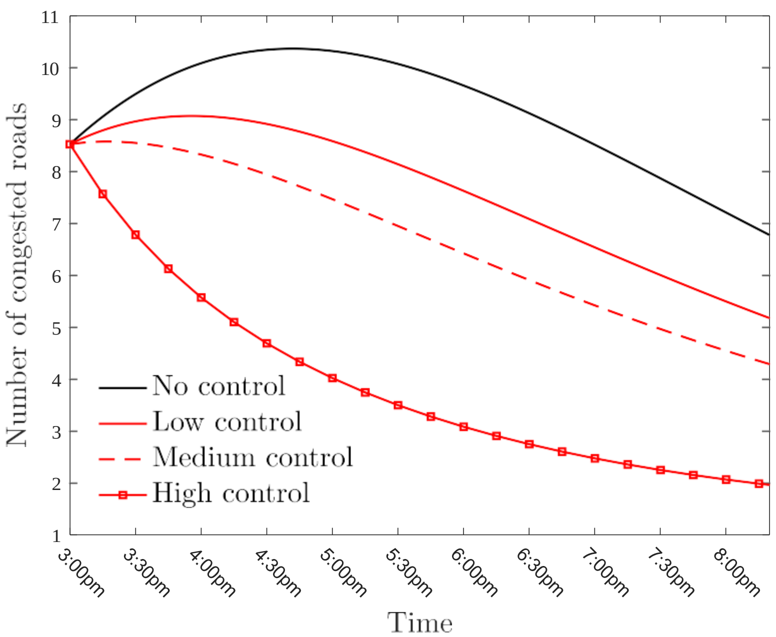

FIGURE 4. Congestion dynamic for the absence (black curve) and presence (red curves) of optimal control and for parameter values $\rho = 0.7$, and thus $\beta = 0.055$ and $\gamma = 0.115$. Results indicate that moderate or high levels of control implementation are able to reduce the congestion without first reaching any peak value.

Although our goal was to shed some light on the issue of traffic jams in one county of metro Atlanta, we can improve the work by studying the problem on a larger portion of the metro area. Our parameter estimation method did not sufficiently predict the timing and value for the peak of congestion; see Figure 2. More robust data collection by collecting finer data in a larger time interval and using a more detailed estimation method such as Weighted Least Squares (WLS) [21] would yield more accurate results, which could be the future direction of this work.

5. References

[1] J. Kenworthy and P. Newman, *The end of automobile dependence: How cities are moving beyond car-based planning*, Island Press. (2015).

[2] K.M. Kruse, *A traffic jam in Atlanta would seem to have nothing to do with slavery. but look closer*, The New York Times Magazine. (2019).

[3] Y. Hu, W. Barbour, S. Samaranayake, and D. Work, *Impacts of Covid-19 mode shift on road traffic*, arXiv preprint arXiv:2005.01610 (2020).

[4] J. Wu, Z. Gao, and H. Sun, *Simulation of traffic congestion with sir model*, Modern Physics Letters B. **18** (2004), no. 30, 1537–1542.

[5] J. Wu, Z. Gao, H. Sun, and H. Huang, *Urban transit system as a scale-free network*, Modern Physics Letters B. **18** (2004), no. 19n20, 1043–1049.

[6] B. Jiang and C. Claramunt, *Topological analysis of urban street networks*, Environment and Planning B: Planning and design. **31** (2004), no. 1, 151–162.

[7] D.J. Watts and S.H. Strogatz, *Collective dynamics of 'small-world'networks*, Nature. **393** (1998), no. 6684, 440–442.

[8] C.F. Daganzo, *On the macroscopic stability of freeway traffic,* Transportation Research Part B: Methodological. **45** (2011), no. 5, 782–788.

[9] H.S. Mahmassani, M. Saberi, and A. Zockaie, *Urban network gridlock: Theory, characteristics, and dynamics*, Procedia-Social and Behavioral Sciences. **80** (2013), 79–98.

[10] S. Colak, A. Lima, and M.C. Gonzalez, *Understanding congested travel in urban areas*, Nature Communications. **7** (2016), no. 1, 10793.

[11] M. Saberi, H. Hamedmoghadam, M. Ashfaq, S.A. Hosseini, Z. Gu, S. Shafiei, D.J. Nair, V. Dixit, L. Gardner, S.T. Waller, et al., *A simple contagion process describes spreading of traffic jams in urban networks*, Nature Communications. **11** (2020), no. 1, 1–9.

[12] C. Hofer, G. Jager, and M. Fullsack, *Including traffic jam avoidance in an agent-based network model*, Computational Social Networks. **5** (2018), no. 1, 1–12.

[13] WSB-TV, *Atlanta traffic and road conditions*. https://www.wsbtv.com/traffic/.

[14] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, *Epidemic outbreaks in complex heterogeneous networks*, The European Physical Journal B-Condensed Matter and Complex Systems. **26** (2002), no. 4, 521–529.

[15] F. Brauer and C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*, Springer. **2** (2012).

[16] L.D. Berkovitz, *Optimal control theory*, Springer Science & Business Media. **12** (2013).

[17] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright, *Convergence properties of the neldermead simplex method in low dimensions*, SIAM Journal on optimization. **9** (1998), no. 1, 112–147.

[18] L. Arriola and J.M. Hyman, *Sensitivity analysis for uncertainty quantification in mathematical models*, Mathematical and statistical estimation approaches in epidemiology. (2009) pp. 195–247.

[19] L.M. Arriola and J.M. Hyman, *Being sensitive to uncertainty*, Computing in Science & Engineering. **9** (2007), no. 2, 10–20.

[20] T. Fingert, *Study projects how much time, money is lost being stuck in Atlanta traffic*. https://www.fox5atlanta.com/news/study-projects-how-much-time-money-is-lost-being-stuck-in-atlanta- traffic#:˜:text=The%20report%20says%20drivers %20in,of%20staring%20at%20brake%20lights.

[21] R. de Levie, *When, why, and how to use weighted least squares*, Journal of Chemical Education. **63** (1986), no. 1, 10.