RESEARCH ARTICLE   OPEN ACCESS

KASSANDRA Model: Detecting Dangerous Traffic Conditions By Modeling Drivers' Internal Stress Energy

Nikolaos Nagkoulis[1,2] (ID), Dimitrios Nalmpantis[1,*] (ID) and Nick Bearman[2] (ID)

[1]*School of Civil Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, PO Box 452, 541 24 Thessaloniki, Greece*

[2]*Department of Geography, Faculty of Social and Historical Sciences, University College London, Pearson Building, Gower Street, London WC1E 6BT, United Kingdom*

Abstract:

Introduction: This paper introduces an innovative method to reduce car accidents by employing mechanical concepts and energy conservation to model drivers' reactions in unexpected scenarios.

Methodology: The approach involves formulating equations to define drivers' "internal stress energy," indicative of their propensity for aggressive driving under time pressure. A spatiotemporal model was developed using traffic data from Highways England and accident data from Transport for London, analyzing around 200 car accidents with data from 80 cameras over two years.

Results and Discussion: Findings suggest a correlation between drivers' internal stress energy and car accidents, highlighting the predictive value of the proposed equations in assessing road segment dangers. More specifically, using the proposed model with 15-minute timeframes increased car accident prediction four (4) times compared to the evenly spatiotemporal car accident distribution. With smaller timeframes, *e.g.*, two (2) minutes, or with real-time data, its predictive power would be significantly higher!

Conclusion: The equations developed offer a promising tool for estimating and preventing car accidents by modeling the influence of drivers' stress on driving behavior.

Keywords: Crash prediction, Spatiotemporal dynamics, Traffic psychology, Internal stress energy, Road safety, Behavior modeling.

CrossMark

*Address correspondence to this author at the School of Civil Engineering, Faculty of Engineering, Aristotle University of Thessaloniki, PO Box 452, 541 24 Thessaloniki, Greece; Tel: +30-2310-995811; E-mail: dnalba@civil.auth.gr

1. INTRODUCTION

Car accident numbers remain high in the 21st century. According to the European Transport Safety Council (ETSC) [1], there was an average of 50 road deaths per 106 inhabitants in the European Union (EU) in 2017. However, only a tiny amount of effort is put into deeply understanding this social problem. It seems that a "normalization" process has taken place [2], and car accidents have been included in societies' everyday lives. This could be explained using Freud's repression theory [3] through a normalization process (to feel safe and continue driving).

However, transportation theorists often ignore these parameters and overestimate the need for restricting illegal driving. An illustrative example is the educational

campaigns targeting young drivers. These campaigns are usually scientifically grounded on research explaining the factors that lead young drivers to car accidents [4-6], and often indicate that there are psychological factors that result in irresponsible driving. These methods sometimes overestimate drivers' freedom and personality and, at other times, represent the drivers as deterministically functioning psychological objects [7-10]. This paper highlights that arriving safe depends on how the whole system reacts in unexpected situations.

Quantitative road safety modeling is focused mainly on crash injuries' severity [11-18] and accident analysis focused on the development and application of advanced econometric and statistical techniques [19]. On the other hand, driver behavior modeling is mainly qualitative [20-23]. Although there were some attempts towards quantitative driver behavior modeling, these were mainly partial models about fatigue, gender, and distracted driving [24-28], stress and risk [29-34], aggressive driving behavior [35-38], dilemma zones at intersections [39-47], car-following [48-57], *etc*. This paper is a modest proposal toward quantifying driver's behavior, trying to fill the gap between qualitative and quantitative modeling.

This paper aims to approach drivers' aggressive behavior in a way that makes it possible to reduce traffic accidents using live data. To do so, a theoretical tool using the energy conservation concept was developed and tested in London Highways. Specifically, drivers are approached as rational agents. Each agent makes an assumption for their arrival time and an expectation for the average velocity they will have during different parts of the journey. Even though this expectation might not be conscious in real life, we believe it determines drivers' behavior. Our model will suppose that whenever an agent fails to fulfill their expectations, their "internal stress energy" increases, making them drive aggressively. This paper will expand and apply this theoretical tool to London Highways, proving that high spatiotemporal concentrations of internal stress-energy are connected to car accidents. Thus, compared to all the previous approaches mentioned above, the novelty of this research is that it introduces the concept of "internal stress energy" and proves it to be an effective car accident prediction indicator.

We named our approach the "KASSANDRA model" after Cassandra, the daughter of Priam, the King of Troy, whose warnings for disastrous future events were not taken seriously. KASSANDRA is a spatiotemporal model that consists of a set of equations that define drivers' "internal stress energy," indicative of their propensity for aggressive driving under time pressure. The findings from the testing of the proposed model suggest a correlation between drivers' internal stress energy and car accidents, highlighting the predictive value of the proposed model in assessing road segment dangers. The model, its testing, and the findings are presented in the following chapters.

2. METHODOLOGY AND MATHEMATICAL FORMU-LATION

In this paper, we assume that before the start of a journey, each driver has some expectations about its duration. A driver arrives on time at their destination when their location at the expected time is identical to their actual location $x_{(duration)actual} = x_{(duration)expected}$. However, the drivers' initial expectations might not be accurate, either because they misinterpreted the conditions before starting or because of unexpected events that occurred during the journey. In this case, they can either change their expectations or try to fulfill them, changing their driving behavior. The main assumption of this paper is that when unexpected events occur, a number of drivers will change their behavior to arrive on time, resulting in car accidents. Therefore, for driver j, we assume that her target is to minimize the difference between her expected and actual journey duration. Equivalently, keeping the duration constant and equal to the expected duration to minimize the difference between her hypothetical location and her actual location. U_j in Eq. (**1**) represents the utility of the driver at the end of the journey. In Eq. (**1**), t_d represents the expected duration of the journey, and v is the velocity of the driver. Throughout the following sections, *exp* is used to indicate a variable of expected values. If there is no *exp* indicator, the variable represents the actual conditions.

$$U_{j,td} = \Delta x_{j,td} = x_{td} - x_{(exp)td}$$
$$= \int_0^{td} v\, dt - \int_0^{td} v_{(exp)}\, dt = \int_0^{td} \left(v - v_{(exp)}\right) dt \tag{1}$$

If the driver arrives at the destination later than the expected time (longer journey than expected), U_j takes negative values. If the driver arrives at the destination before the expected time U_j takes positive values. Even though it is possible to have positive U_j values, this happens when the drivers are free to move with higher velocities than expected. These conditions are not of interest in this paper, which emphasizes drivers who are stressed because they fail to arrive on time. Without loss of generality, we assume that the target of the driver is to maximize U_j.

Eq. (**1**) represents the utility of the driver j at the end of the journey ($t = t_d$). However, throughout the journey, the drivers make estimates and assumptions about the duration of their journey. They often try to assume (consciously or not) if they are going to arrive on time or not. This allows us to define a utility for each driver for every possible time t value (Eq. **2**).

$$U_{j,t} = \int_0^t \left(v - v_{(exp)}\right) dt \tag{2}$$

Next, we move from a temporal to a spatial analysis. Using the derivative of v_{exp}, we get Eq. (**3**). This is necessary because using road spatial segments instead of temporal sections is more practical.

$$dt = \frac{1}{v_{exp}}dx_{exp} \rightarrow U_{j,x_{exp}} = \int_0^{x_{exp}} \left(\frac{v}{v_{exp}} - 1\right)dx_{exp} \quad (3)$$

The journey can be divided into n consecutive segments $(S_1, S_2,..., S_n,)$, each with a starting and an ending point. Suppose that x_a is the starting point and x_b the final point of the segment S_i. Then, we can define the variable $\Delta U_{j,S_i}$ which represents the utility that the driver gained (or lost) during that segment S_i (Eq. **4**).

$$\Delta U_{j,S_i} = U_{j,x_b} - U_{j,x_a} = \int_{x_a}^{x_b} \left(\frac{v}{v_{exp}} - 1\right)dx_{exp} \quad (4)$$

A negative $\Delta U_{j,S_i}$ means that during that segment, the driver was driving slower than their initial expectations, increasing the actual duration of their journey. It goes without saying that $U_{j,td}$ is the sum of the $\Delta U_{j,S_i}$ values (Eq. **5**).

$$x_{(exp)td} = \sum S_i \rightarrow U_{j,td} = \sum_{i=1}^{i=n} \Delta U_{j,S_i} \quad (5)$$

From the above relationships, we understand that the drivers can increase their utility by adjusting their speed. However, drivers usually cannot unilaterally adjust their speed throughout their journey, as the existence of other drivers affects their behavior. Therefore, j also takes into consideration the velocities of the nearby drivers j. Therefore driver's j target can be represented using Eq. (**6**).

$$\underset{v_j}{\text{Max}}\, U_{j,td}(v_j, v_{-j}) \quad (6)$$

In this paper, we assume that all drivers perform the same estimations about their journey. This does not indicate that they perform the same journey but that all the drivers expect to find the same driving conditions in each road segment. Even though this assumption limits the findings of this research, it could be considered valid when the drivers are experienced. The reader can find an indicative figure about drivers' expectations in the appendix. Calculating the hourly average velocities, we can see some spatial and temporal patterns that can be easily guessed/expected by the drivers. The aforementioned assumption makes it possible to create a "hypothetical" driver whose velocity is identical to the hourly average velocity of the drivers of each segment.

We will define the stress-energy of the flow for a road segment $\{x_a, x_b\}$ using the classic definition of dynamic energy *Energy* $= mgh$, as in Eq. (**7**):

$$Energy = NgU_{j,td}(v_j, v_{-j}) \quad (7)$$

where:

- N: is the number of vehicles in the examined road segment, and
- g: is a constant similar to the gravitational constant representing the inclination of the drivers to change their velocity in order to reach their destination in time.

In this paper, we will assume that every agent has the same g constant, meaning they have the same motivation to reach their destination in time. To avoid examining g, we will define the parameter e in Eq. (**8**). Examining e is equivalent to examining *Energy*, as g is a constant. The number of vehicles can be expressed using a density factor: $N = \rho A$, where ρ(vehicles/(m*lanes)) is the car density and A(m*lanes) is the area examined. The number of lanes is expressed using the parameter λ(lanes):

$$e = \frac{Energy}{g} = NU_{j,td}(v_j, v_{-j}) = \rho A \int_{x_a}^{x_b} \left(\frac{v_t}{v_{exp}} - 1\right)dx \quad (8)$$

With the assumption that drivers have similar Δv in a small part of the road path dx (homogeneity), the integral is simplified for a specific road area (between x_a and x_b), as follows. The wider the area examined, the more problematic the assumption of homogeneity becomes. Naming L the distance between x_a and x_b, we get $A = \lambda L$ and from the Eq. (**8**), we get Eq. (**9**):

$$e = L^2\lambda\rho\left(\frac{v_t}{v_{exp}} - 1\right) \quad (9)$$

The higher the e parameter, the more stressed the driver is when she leaves the examined segment (S). We presented a theoretical example to improve the understanding of e. Let us suppose that an unexpected congestion takes place. "Unexpected" means that it does not normally appear at that road segment during that time of the day (visit the appendix for more information about expectations). This congestion will result in lower actual velocity values than expected, resulting in negative internal stress energy values $(v_t < v_{exp} \rightarrow e < 0)$. The density of the flow (ρ) and the length of the congested segment (L) can worsen the situation, resulting in even more negative internal stress energy values. That stress should be considered an indicator associated with high car accident possibilities. When drivers go through road segments that have high e values, they "collect" energy. However, a car accident is not expected to happen during these parts. We assume that a car accident is expected to occur when energy radically increases or decreases. To find these road parts, we calculate the drivers' internal "power" (a derivative of e with respect to time t, for a time period $\{t_a, t_b\}$) as shown in Eq. (**10**):

$$\beta = \frac{\partial e}{\partial t} \approx \frac{\Delta e}{\Delta t} = L^2\lambda\frac{\rho_b\left(\frac{v_b}{v_{exp}}-1\right) - \rho_a\left(\frac{v_a}{v_{exp}}-1\right)}{t_b - t_a} \quad (10)$$

The energy of a road segment decreases in time when

$\beta < 0$ and increases when $\beta > 0$. This means that for a road segment $\{x_1, x_2\}$, $\beta_{t_A} < 0 \rightarrow e_{t_A} > e_{t_B}$ for $t_A < t_B$ and vice versa. Therefore, supposing that $\beta_{t_0} < 0$ for a road segment $\{x_1, x_2\}$ and $\beta_{t_0} > 0$ for a road segment $\{x_3, x_4\}$, for continuous functions e and β there will be a road segment $\{x_2, s_3\}$ in which $\beta_{t_0} = 0$ will hold, based on the Bolzano's theorem.

In Fig. (**1**), we see a group of vehicles with high energy moving from $x = 0$ to $x = 10$. As the time variable increases, the energy values of the left road segment decrease ($\beta < 0$), while the energy of the right road segment increases ($\beta > 0$). Therefore, there will be an area where $\beta = 0$, which means that a local optimum of e will be found there.

For two consecutive road parts $\{x_1, x_2\}$ and $\{x_2, x_3\}$ (or $S1, S2$), we define the parameter B, as the overall internal power of the road segments in Eq. (**11**):

$$B_{x_1,x_3} = \left|\beta_{x_1,x_2}\right| + \left|\beta_{x_2,x_3}\right| = \left|\beta_{S1}\right| + \left|\beta_{S1}\right| \quad \textbf{(11)}$$

Summing up, in this paper, we will use the highest

25% of B values, subject to $\beta_{S1} < 0$ & $\beta_{S2} > 0$. In further research, more assumptions can be tested to improve the accuracy of the prediction.

3. APPLICATION

The relationships were applied using traffic flow data from the Highways England's Application Programming Interface (API) [58] and car accident data from the Transport for London (TfL) API [59]. For both datasets, Rstudio was used to download data in JSON formats [60]. Highway cameras' locations, together with car accidents, were inserted, projected, and plotted using their coordinates in Rstudio [61] and the "qtm" function for plotting [62]. As the car accidents dataset includes car accidents only from some areas of the highways, all cameras that do not intersect with these areas were excluded. The cameras left can be divided into three (3) groups: G1, G2, and G3 (Fig. **2**).

Each of the three groups contains cameras capturing data from different directions. Only "A" and "B" cameras refer to the main road. The direction of "A" ("Away from London" or "clockwise") cameras is opposite from "B" cameras (indicating "Back to London" or "anti-clockwise")
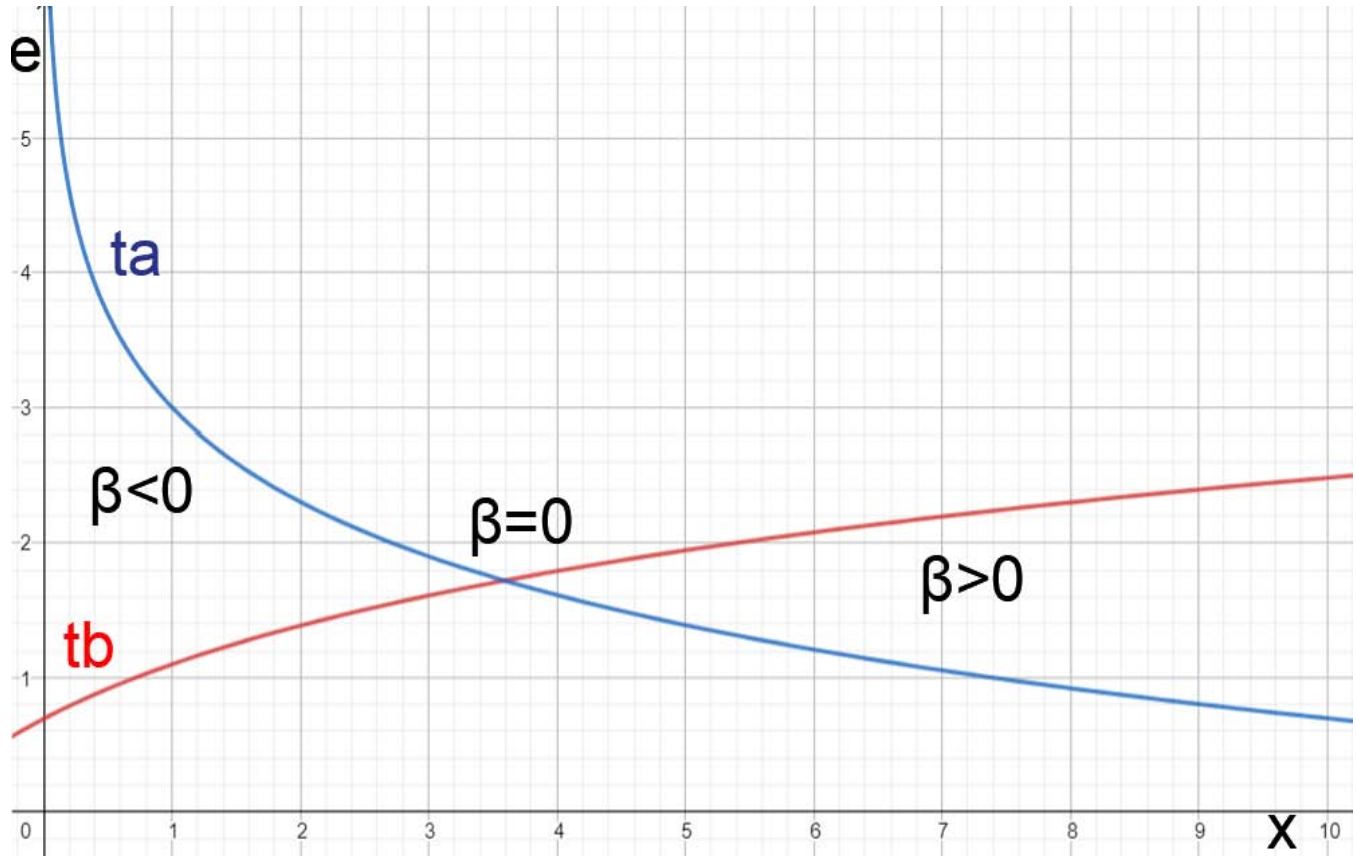

Fig. (1). Hypothetical energy profiles e (for t_a and t_b) along the path $x = 0$ to $x = 10$; t_a and t_b represent time sections with $t_a < t_b$.
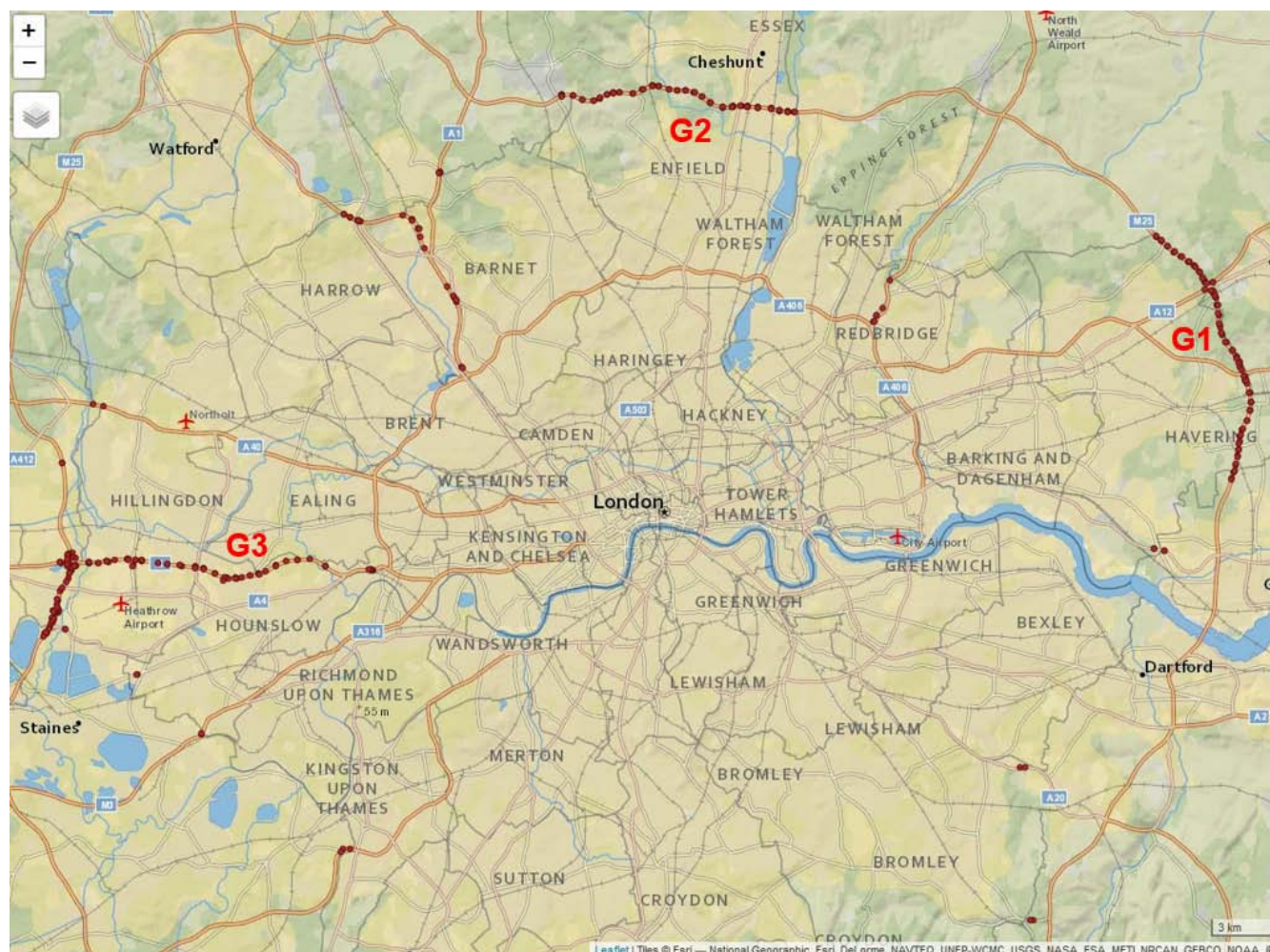
Fig. (2). The areas G1, G2, and G3 are the parts of the England Highways that intersect with the car accidents dataset.

[63]. These cameras are going to be analyzed separately. This way, we have the following semi-groups: G1A and G1B, G2A and G2B, and G3A and G3B. The consecutive cameras were connected manually in ArcGIS with line segments. The characteristics of each camera were used to make it possible to understand which camera is first and which follows while driving. The line segments were then manually imported to Rstudio to use their lengths to calculate β and also to divide the car accidents into two groups (A and B). The car accidents closer to the left part of the road were considered to belong to the A group, and the car accidents closer to the right part of the road were considered to belong to the B group. This division was necessary because TfL does not give details regarding the direction of the cars in a convenient way.

Table 1. Data availability per year for the first five (5) cameras of G1A.

Cam\year	2014	2015	2016	2017
1	8,711	28,219	31,868	31,588
2	8,735	28,266	31,868	31,588
3	8,711	28,242	31,868	31,588
4	8,711	28,266	31,868	31,588
5	8,735	28,218	31,868	31,588
6	...	...	31,868	31,588

Table 2. Necessary data for each road segment. For each specific camera, speed values and volumes were needed.

Site Name	Report Date	Time Period Ending	Time Interval	Average V (mph)	Total Volume
M25/5683A	2016-01-01T00:00:00	0:14:00	0	71	79
M25/5683A	2016-01-01T00:00:00	0:29:00	1	70	41
M25/5683A	2016-01-01T00:00:00	0:44:00	2	69	102
M25/5683A	2016-01-01T00:00:00	0:59:00	3	70	173

The Highways England's API was used to get each camera's available traffic flow characteristics. To get the traffic data from the API, the ID of the camera was needed, as well as the target time period. Finally, all years except 2016 and 2017 were excluded from our analysis because some of the cameras were not operating during the whole year in all years, and these years were excluded because they could result in errors. For example, in the case of G1A, only in the years 2016 and 2017 we could find the same number of flow data for every camera (Table 1).

The matrices give basic information about the traffic and vehicle characteristics. Three (3) sets of information were kept: average speed, the volume of cars, and the date and time that the data were recorded. A sample of these characteristics is presented in Table 2.

Using these data, some transformations were necessary to calculate the internal power β parameter. These transformations can be found in the Appendix, along with information about the calculation of the expected velocities. This way, we could calculate the β parameter for each road segment (Table 3).

To apply the methodology proposed, we calculated the B values only for those consecutive road segments $\{x_i, x_{i+2}\}$ with $\beta_{x_i,x_{i+1}} < 0 \, \& \, \beta_{x_{i+1},x_{i+2}} > 0$, based on Eq. (10) (Table 4).

Finally, when these values appeared, we compared the number of car accidents in the road segments with the highest B values (top 25%) with the number of car accidents in all the other road segments (Fig. 3).

We expected to find car accidents in the examined areas. To test that assumption, we chose a smaller area inside the area $\{x_i, x_{i+2}\}$ close to the middle camera x_{i+1}. This area's diameter was considered equal to L (the length of the road segments). We examined if a car accident occurred in that area (approximately 600 m) for a duration of 15 minutes. If an accident appeared under these circumstances, we would consider that accident to have been predicted. Finally, even though the methodology was applied to all six (6) groups: G1A, G1B, G2A, G2B, G3A, and G3B, the B areas were all excluded because they contained cameras that were not operating properly.

In Fig. (3), the accidents were represented using solid circles. If an accident occurred inside the circle, it was thought to have been predicted. In this specific case, using an area that is ¼ of the whole area and a time frame that is ½ of the whole time, we have predicted 50% of the car accidents. This means that a specific spatiotemporal section, which represents 1/8 of the overall, contains 50% of the car accidents (>> 1/8); therefore, the assumptions of this paper have been verified.

Table 3. Calculating β values for the first 15-minute timeframes for road segment 1 in the G1A group with coordinates (x_1, y_1), (x_2, y_2) = (24538,6700698),(24855,6700337).

-	Date & Time	L (m)	v (mph)	N (vehicles)	v_{exp} (mph)	Δt (sec)	βs
1	1/1/2016 0:14	481	32.18	78	28.77	840	-0.051
2	1/1/2016 0:29	481	31.73	40.5	28.77	900	0.047
3	1/1/2016 0:44	481	31.51	98	28.77	900	0.064
4	1/1/2016 0:59	481	31.51	172.5	28.77	900	0.058
5	1/1/2016 1:14	481	31.29	197	28.05	900	0.020

Table 4. Calculating power values for the first 15-minute temporal-segments for the first five (5) road segments in the G2A group.

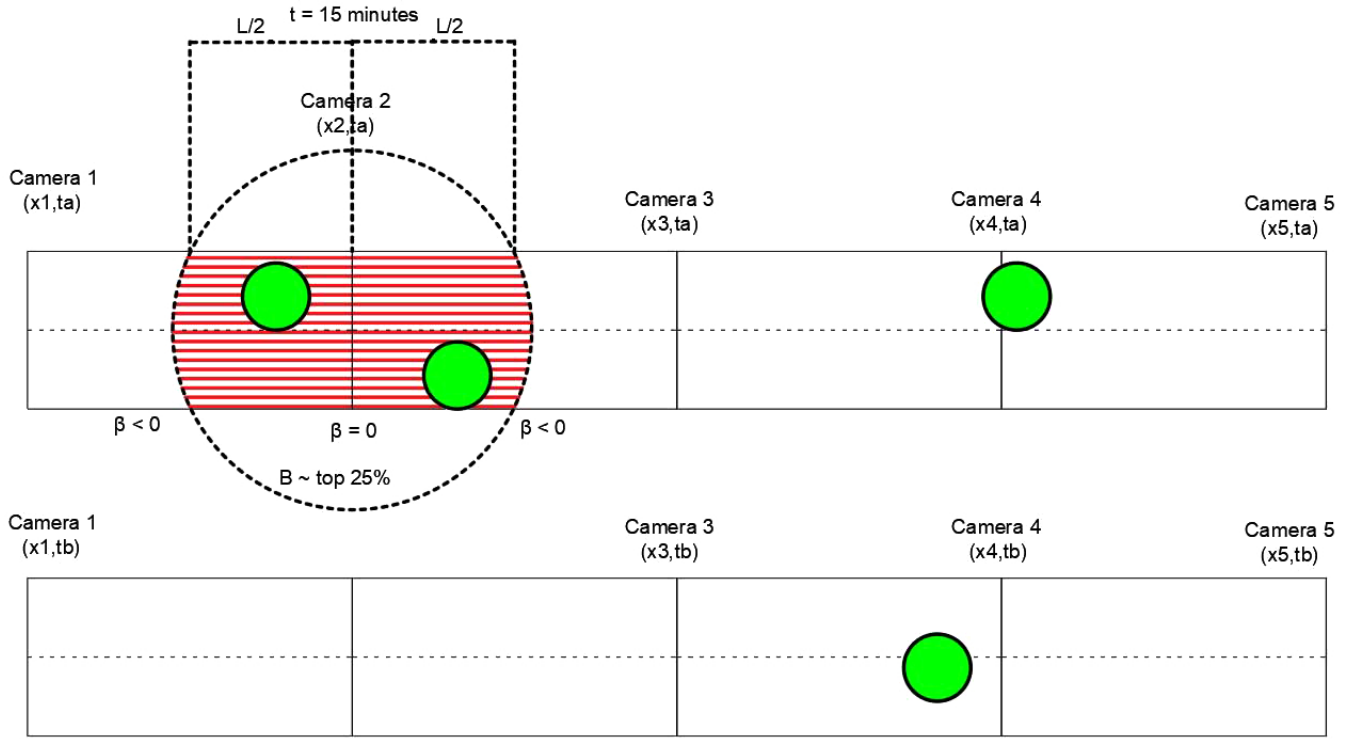| Date & Time	$B(x_0, x_2)$	$B(x_1, x_3)$	$B(x_2, x_4)$	$B(x_3, x_5)$
1/1/2016 0:14	-	0.082	-	-
1/1/2016 0:29	-	-	-	0.019
1/1/2016 0:44	-	-	-	-
1/1/2016 0:59	-	-	-	-
1/1/2016 1:14	-	-	-	-

Fig. (3). Four (4) consecutive hypothetical road segments for two (2) 15-minute temporal segments (t_a, t_b).

4. RESULTS AND DISCUSSION

The car accidents in the three areas examined during a period of two (2) years are 214. Based on data availability, the sum of the road segments examined in the areas G1A, G2A, and G3A is 77, each containing a camera in the middle. The available data for each camera corresponds to 63,456 15-minute entries. If these accidents were evenly distributed in space, then there would be 214/77=2.78 accidents in every road segment. Similarly, if the accidents were evenly distributed in time, there would be a probability of 2.78/63,456=4.3·10^{-5}=0.0043% of having a car accident in every 15-minute entry. Equivalently, we get this percentile by diving the number of accidents by the time entries, multiplied by the number of cameras (spatial entries): 214/(63,456.77)=4.3·10^{-5} car accidents/ (approximately 600 m .15 minutes).

Nevertheless, car accidents are not evenly distributed in space and time. This paper shows that the areas with high B values correspond to a higher probability of car accidents. Using the highest 25% of B values of each road segment $\{x_i, x_{i+2}\}$, subject to $\beta_{x_i,x_{i+1}} < 0 \,\&\, \beta_{x_{i+1},x_{i+2}} > 0$, we get 11 car accidents, which are considered as "predicted." The condition $\beta_{x_i,x_{i+1}} < 0 \,\&\, \beta_{x_{i+1},x_{i+2}} > 0$ applies to 249,420 spatiotemporal entries. We used the top 25% of them to search for car accidents around them, resulting in 62,355 spatiotemporal entries. Finally, the probability we get is 11/62,355=1.76·10^{-4}=0.0176%, which is considerably higher than the previous one. This

probability is lower in the area G1A and higher in G2A and G3A.

By dividing the probabilities, *i.e.*, 1.76·10^{-4}/4.3·10^{-5} = 4.09, we find that by using the proposed model, we can find four (4) times more car accidents compared to what we would have found if they were evenly distributed. As explained above, if we use lower time frames, this number will significantly increase; the lower the time frame, the more significant the increase and the same goes for real-time data.

Finally, we compared these results with the probability we get when we use 25% of the highest car volumes. The probability that we get is 2.14·10^{-5}. This means that internal stress energy is a more effective indicator than high car volumes regarding car accident prediction.

The overall number of accidents (214) also includes accidents that have taken place in B parts because of inaccurate spatial exclusion. These accidents cannot be predicted using traffic data from the A parts. If the accidents in B areas were removed from the datasets, the probability of predicting a car accident would have been even higher than the even car accident distribution. We believe that the accuracy of a prediction can be radically improved in case more frequent data are available (*e.g.*, every 2 minutes instead of 15 minutes).

CONCLUSION, PROPOSALS, AND ORIENTATIONS FOR FUTURE RESEARCH

In this paper, we propose a method that can be used to predict car accidents inspired by mechanics and the conservation of energy. The basic assumption is that drivers change their behavior during their trip to arrive on time at their destination. The drivers' inability to fulfill their expectations during the trip is counted using an "internal stress energy" parameter. The more they feel they will be late, the more nervous they become, increasing their energy. Each driver behaves like an agent aiming for energy minimization. The model is applied in London Highways using traffic data from London Highways and car accidents data from TfL. Using two years' data, we prove spatial variations of drivers' internal stress "power" are connected with an increased probability of a car accident. More specifically, using the proposed model with 15-minute timeframes increased car accident prediction four (4) times compared to the evenly spatiotemporal car accident distribution. The model can be easily applied when live traffic data are available. With such data, we expect a dramatic improvement of the model as its accuracy increases and the time frame decreases. The proposed model can be tested in other areas to estimate its accuracy and potential uses. Overall, the model can be expanded to analyze real-time data to make predictions and help stakeholders act in dangerous conditions.

The novelty of this research is that it introduces the concept of "internal stress energy" and proves it to be a more effective car accident prediction indicator than high car volumes, given there are data of small timeframes or real-time data available. Nevertheless, the main limitation of the research is that the concept was tested in timeframes of not less than 15 minutes, a case in which even better results are expected.

In further research, the assumptions of the model can be studied. For example, it is interesting to find out if the traffic conditions become less dangerous when the drivers change their expectations and adapt to their environment. It is also interesting to find out if the probability of a car accident increases when the energy is radically released. Finally, neural networks can be used to find more accurate correlations between car accidents and the drivers' internal stress energy, which might not be easily noticeable.

LIST OF ABBREVIATIONS

API	=	Application Programming Interface
ETSC	=	European Transport Safety Council
EU	=	European Union
TfL	=	Transport for London

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

FUNDING

CONFLICT OF INTEREST

Dr. Dimitrios Nalmpantis is the Editorial Advisory Board Member for the journal The Open Transportation Journal.

ACKNOWLEDGEMENTS

APPENDIX

This part describes how the main equations are transformed to match the examined dataset. Traffic flow can be defined as the number of vehicles passing through a point for a period of time $Q = \frac{N}{\Delta t}$. At the same time, flow is also the density of the cars in a road segment, multiplied by the average velocity of these cars $Q = \rho v \lambda$. In this specific case, λ is used to indicate the number of lanes, as density is defined using the number of lanes. Finally, combining these two relationships, we get the density $\rho = \frac{N}{\Delta t \, v \, \lambda}$.

In these relationships, Δt is predefined and depends on data availability. Assuming that the difference Δt between t_a and t_b is the same as the time difference needed for the data "Ndata" to be collected, the relationship created in the methodology becomes Eq. (**A.1**):

$$\beta = L^2 \, \lambda \, \frac{\rho_b\left(\frac{v_b}{v_{exp}}-1\right) - \rho_a\left(\frac{v_a}{v_{exp}}-1\right)}{\Delta t}$$
$$= L^2 \frac{N_{data,b}\left(\frac{1}{v_{exp}}-\frac{1}{v_b}\right) - N_{data,a}\left(\frac{1}{v_{exp}}-\frac{1}{v_a}\right)}{(\Delta t)^2} \quad \text{(A.1)}$$

Parameter β refers to a road segment $\{x_1, x_2\}$. Therefore, we get the values v and ρ using the following two cameras: the camera at the beginning of the road segment and the camera at the end. The average of the values of the two cameras is used to get the characteristics of the segment as shown in Eqs. (**A.2** and **A.3**):

$$\rho = 0.5\left(\rho_{x_1} + \rho_{x_2}\right) \quad \text{(A.2)}$$

$$v = 0.5\left(v_{x_1} + v_{x_2}\right) \qquad \textbf{(A.3)}$$

To better understand these relationships, the most important part is to remember that they refer to the same road segment for consecutive time periods (Fig. **4**).
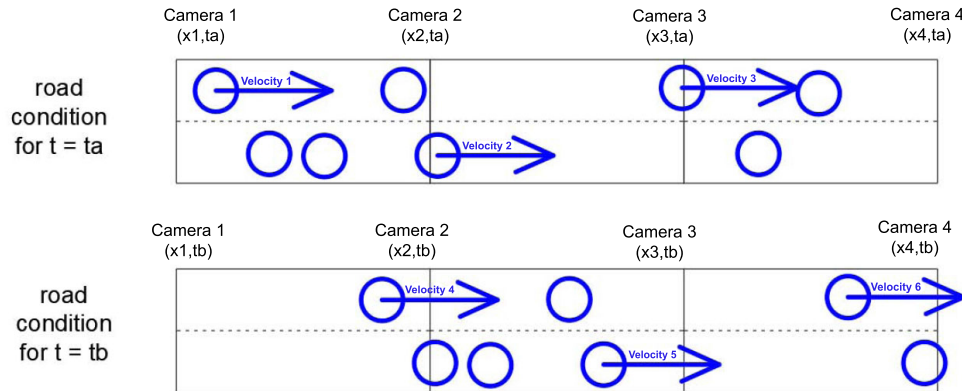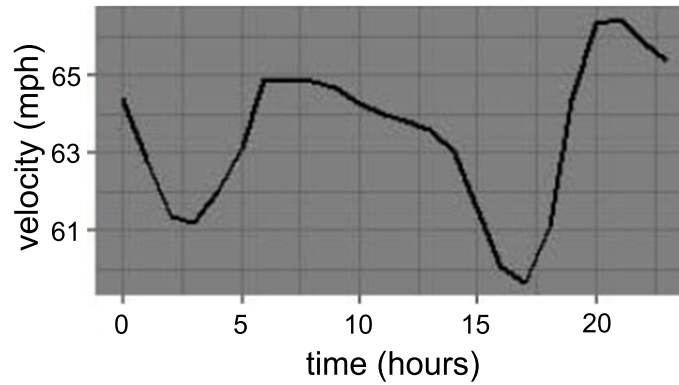

Fig. (4). A hypothetical road divided into three (3) groups using four (4) cameras. Flow density, as well as cars' velocity, change over time. To get the energy *e* and the derivative *β* of each part, consecutive cameras are used.
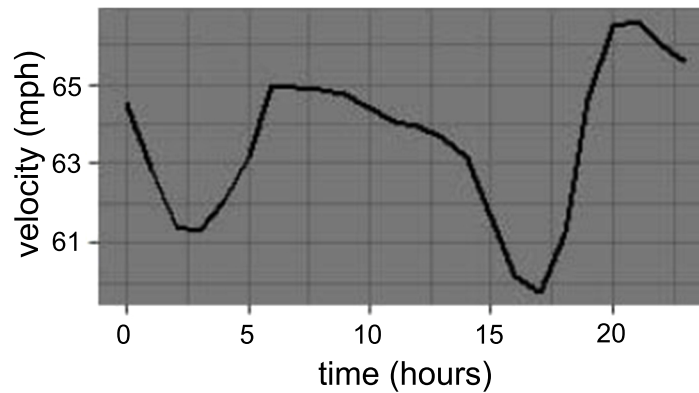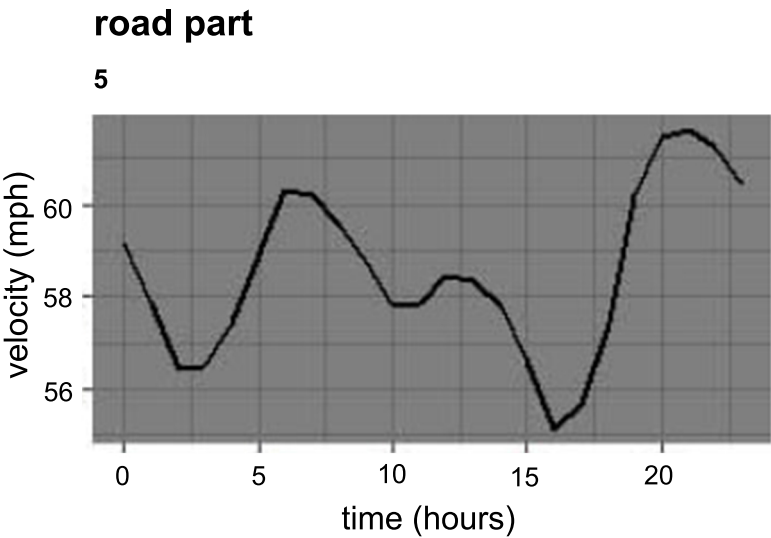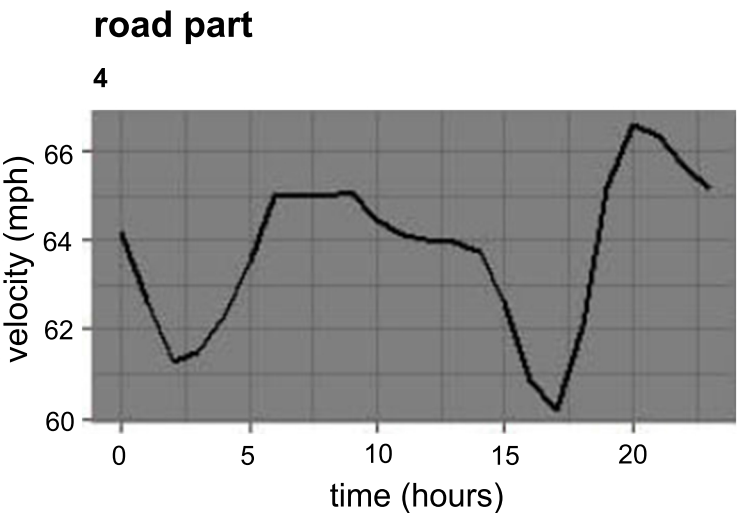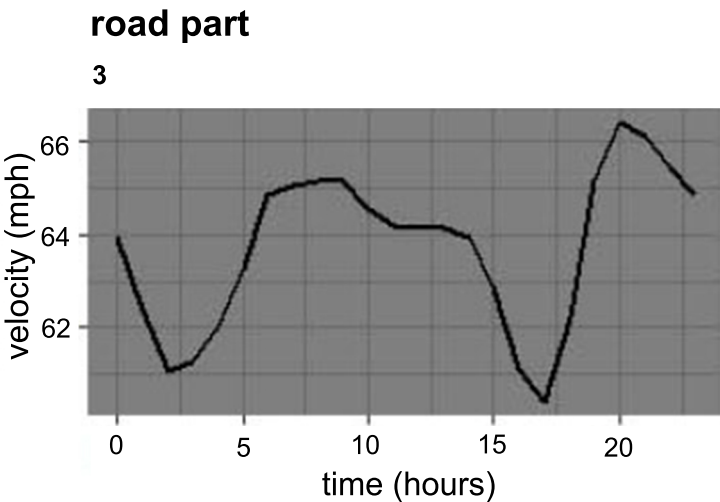
road part

3

road part

4

road part

5

road part

6
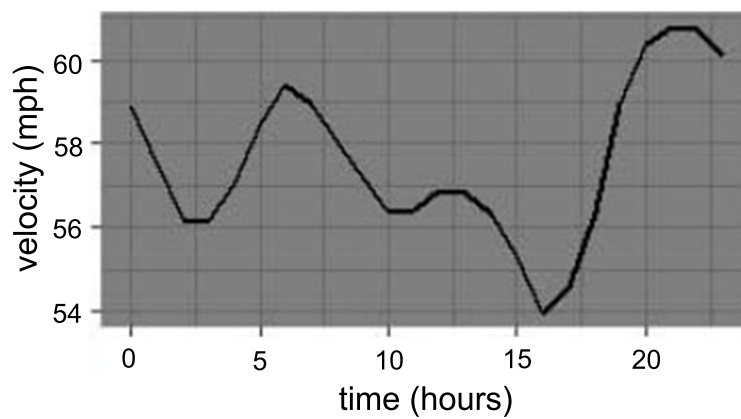
road part

7

road part

8

road part

9
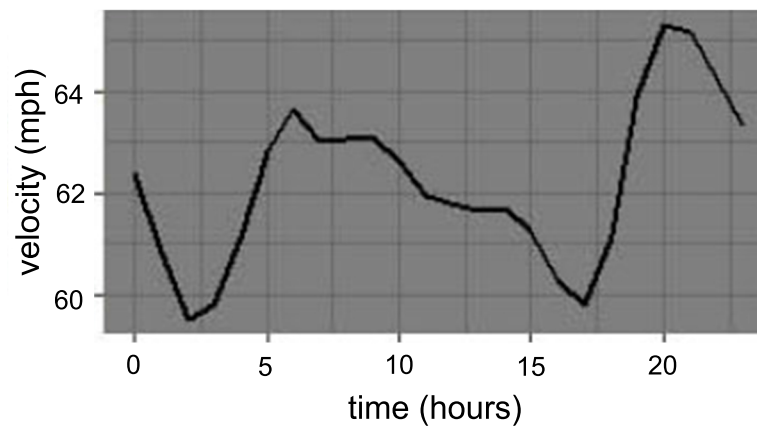
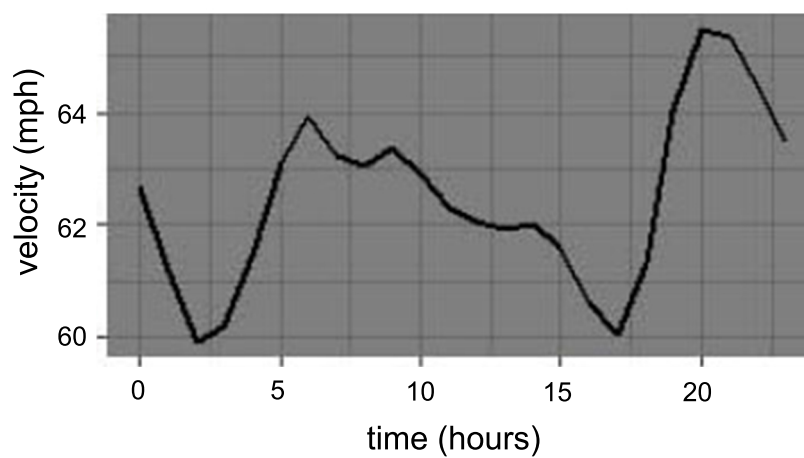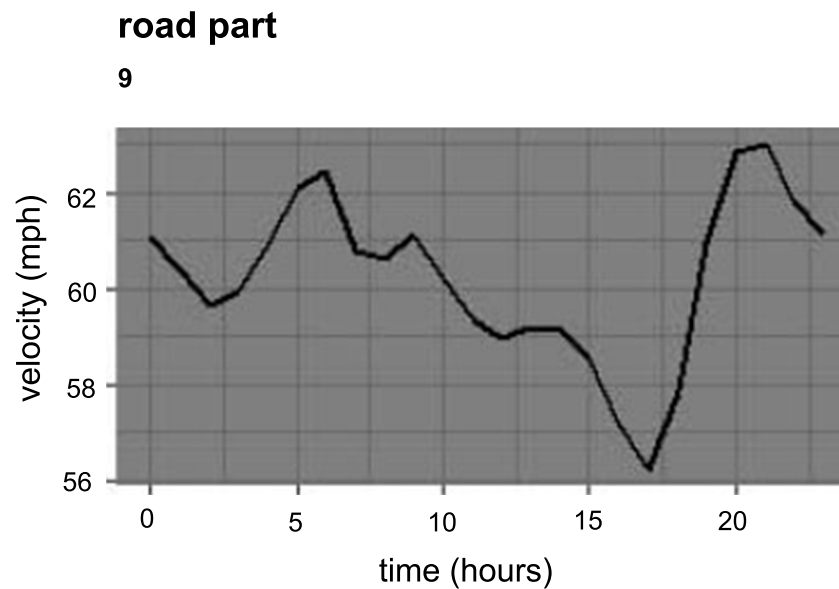**Fig. (5).** Average velocities for the first nine (9) road segments of G1A. For two (2) years of data, the average hourly vehicles' velocities are calculated using camera data. Y-axis: velocity (mph). X-axis: time (24 hours). Average velocities in consecutive parts change steadily, indicating that the utility's function continuity assumption is accurate. In addition, there is no high velocity deviation during the daytime (in these road segments), which means that the drivers can create (subconsciously) an imaginative profile of the average velocities for the whole path.

This way N_a refers to the volume data recorded at t_a and N_b refers to the volume data recorded at t_b. The same applies to v_a and v_b. Δt is 15 minutes based on the API data availability. In any case, all N and v values are the result of getting the average values from the road segment's starting and ending camera examined. Finally, v_{exp} is the average hourly velocity of a specific road segment, calculated using average yearly values (Fig. **5**) for an extended period of time. Drivers expect to find the road conditions they usually find at each road segment. We get these values using the 2-year data downloaded from the London Highways API.

REFERENCES

[1]   A.D. Fodor, C. Heilpern, and G. Jost, *Ranking EU Progress on Road Safety: 13th Road Safety Performance Index Report,* European Transport Safety Council: Brussels, Belgium, 2019.

[2]   D. Taylor, "Normativity and normalization", *Foucault Stud.,* no. 7, pp. 45-63, 2009.
[http://dx.doi.org/10.22439/fs.v0i7.2636]

[3]   S. Freud, "Repression", In: J. Strachey, Ed., *The Standard Edition of the Complete Psychological Works of Sigmund Freud.,* vol. Vol. 14. Hogarth: London, UK, 1957, pp. 141-158.

[4]   J.A. Groeger, "Youthfulness, inexperience, and sleep loss: The problems young drivers face and those they pose for us", *Inj. Prev.,* vol. 12, no. S1, pp. i19-i24, 2006.
[http://dx.doi.org/10.1136/ip.2006.012070] [PMID: 16788107]

[5]   Å. Murray, "The home and school background of young drivers involved in traffic accidents", *Accid. Anal. Prev.,* vol. 30, no. 2, pp. 169-182, 1998.
[http://dx.doi.org/10.1016/S0001-4575(97)00088-2]   [PMID: 9450121]

[6]   P. Ulleberg, and T. Rundmo, "Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers", *Saf. Sci.,* vol. 41, no. 5, pp. 427-443, 2003.
[http://dx.doi.org/10.1016/S0925-7535(01)00077-7]

[7]   A. af Wåhlberg, L. Dorn, and T. Kline, "The Manchester driver behaviour questionnaire as a predictor of road traffic accidents", *Theor. Issues Ergon. Sci.,* vol. 12, no. 1, pp. 66-86, 2011.
[http://dx.doi.org/10.1080/14639220903023376]

[8]   J.E. Chliaoutakis, P. Demakakos, G. Tzamalouka, V. Bakou, M. Koumaki, and C. Darviri, "Aggressive behavior while driving as predictor of self-reported car crashes", *J. Safety Res.,* vol. 33, no. 4, pp. 431-443, 2002.
[http://dx.doi.org/10.1016/S0022-4375(02)00053-1]   [PMID: 12429101]

[9]   T. Ehring, A. Ehlers, and E. Glucksman, "Do cognitive models help in predicting the severity of posttraumatic stress disorder, phobia, and depression after motor vehicle accidents? A prospective longitudinal study", *J. Consult. Clin. Psychol.,* vol. 76, no. 2, pp. 219-230, 2008.
[http://dx.doi.org/10.1037/0022-006X.76.2.219] [PMID: 18377119]

[10] L.R. Hartley, and J. El Hassani, "Stress, violations and accidents", *Appl. Ergon.,* vol. 25, no. 4, pp. 221-230, 1994.
[http://dx.doi.org/10.1016/0003-6870(94)90003-5]   [PMID: 15676972]

[11] K.A. Abay, R. Paleti, and C.R. Bhat, "The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity", *Transp. Res., Part B: Methodol.,* vol. 50, pp. 74-89, 2013.
[http://dx.doi.org/10.1016/j.trb.2013.01.007]

[12] N. Alnawmasi, and F. Mannering, "A statistical assessment of temporal instability in the factors determining motorcyclist injury severities", *Anal. Methods Accid. Res.,* vol. 22, p. 100090, 2019.
[http://dx.doi.org/10.1016/j.amar.2019.100090]

[13] A. Behnood, and F.L. Mannering, "The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence", *Anal. Methods Accid. Res.,* vol. 8, pp. 7-32, 2015.
[http://dx.doi.org/10.1016/j.amar.2015.08.001]

[14] G. Fountas, P.C. Anastasopoulos, and F.L. Mannering, "Analysis of vehicle accident-injury severities: A comparison of segment-versus accident-based latent class ordered probit models with class-probability functions", *Anal. Methods Accid. Res.,* vol. 18,

pp. 15-32, 2018.
[http://dx.doi.org/10.1016/j.amar.2018.03.003]

[15] J.C. Milton, V.N. Shankar, and F.L. Mannering, "Highway accident severities and the mixed logit model: An exploratory empirical analysis", *Accid. Anal. Prev.,* vol. 40, no. 1, pp. 260-266, 2008.
[http://dx.doi.org/10.1016/j.aap.2007.06.006] [PMID: 18215557]

[16] M.G. Mohamed, N. Saunier, M.L.F. Moreno, and S.V. Ukkusuri, "A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada", *Saf. Sci.,* vol. 54, pp. 27-37, 2013.
[http://dx.doi.org/10.1016/j.ssci.2012.11.001]

[17] J. Mooradian, J.N. Ivan, N. Ravishanker, and S. Hu, "Analysis of driver and passenger crash injury severity using partial proportional odds models", *Accid. Anal. Prev.,* vol. 58, pp. 53-58, 2013.
[http://dx.doi.org/10.1016/j.aap.2013.04.022] [PMID: 23702439]

[18] R. Yu, and M. Abdel-Aty, "Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data", *Saf. Sci.,* vol. 63, pp. 50-56, 2014.
[http://dx.doi.org/10.1016/j.ssci.2013.10.012]

[19] F.L. Mannering, and C.R. Bhat, "Analytic methods in accident research: Methodological frontier and future directions", *Anal. Methods Accid. Res.,* vol. 1, pp. 1-22, 2014.
[http://dx.doi.org/10.1016/j.amar.2013.09.001]

[20] E. Bekiaris, A. Amditis, and M. Panou, "DRIVABILITY: A new concept for modelling driving performance", *Cogn. Technol. Work,* vol. 5, no. 2, pp. 152-161, 2003.
[http://dx.doi.org/10.1007/s10111-003-0119-x]

[21] M. Hatakka, E. Keskinen, N.P. Gregersen, A. Glad, and K. Hernetkoski, "From control of the vehicle to personal self-control; broadening the perspectives to driver education", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 5, no. 3, pp. 201-215, 2002.
[http://dx.doi.org/10.1016/S1369-8478(02)00018-9]

[22] D. Nalmpantis, A. Naniopoulos, E. Bekiaris, M. Panou, N.P. Gregersen, T. Falkmer, G. Baten, and J.F. Dols, "'Trainer' project: Pilot applications for the evaluation of new driver training technologies", In: G. Underwood, Ed., *Traffic and Transport Psychology: Theory and Application.,* Elsevier: Oxford, UK, 2005, pp. 141-156.

[23] T.A. Ranney, "Models of driving behavior: A review of their evolution", *Accid. Anal. Prev.,* vol. 26, no. 6, pp. 733-750, 1994.
[http://dx.doi.org/10.1016/0001-4575(94)90051-5] [PMID: 7857489]

[24] E. Beli, and D. Nalmpantis, "Estimation of the willingness to pay for road safety improvements and its correlation with specific demographic, psychological, and behavioral factors", In: E.G. Nathanail, G. Adamos, I. Karakikes, Eds., *Advances in Intelligent Systems and Computing.,* vol. Vol. 1278. Springer: Cham, Switzerland, 2021, pp. 485-494.
[http://dx.doi.org/10.1007/978-3-030-61075-3_48]

[25] G. Fountas, S.S. Pantangi, K.F. Hulme, and P.C. Anastasopoulos, "The effects of driver fatigue, gender, and distracted driving on perceived and observed aggressive driving behavior: A correlated grouped random parameters bivariate probit approach", *Anal. Methods Accid. Res.,* vol. 22, p. 100091, 2019.
[http://dx.doi.org/10.1016/j.amar.2019.100091]

[26] M. Saffarian, J.C.F. de Winter, and J.W. Senders, "Measuring drivers' visual information needs during braking: A simulator study using a screen-occlusion method", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 33, pp. 48-65, 2015.
[http://dx.doi.org/10.1016/j.trf.2015.07.001]

[27] J.W.C. van Lint, and S.C. Calvert, "A generic multi-level framework for microscopic traffic simulation—Theory and an example case in modelling driver distraction", *Transport. Res. B-Meth.,* vol. 117, pp. 63-86, 2018.
[http://dx.doi.org/10.1016/j.trb.2018.08.009]

[28] K.L. Young, and P.M. Salmon, "Examining the relationship between driver distraction and driving errors: A discussion of theory, studies and methods", *Saf. Sci.,* vol. 50, no. 2, pp. 165-174, 2012.

[http://dx.doi.org/10.1016/j.ssci.2011.07.008]

[29] S.H. Hamdar, H.S. Mahmassani, and M. Treiber, "From behavioral psychology to acceleration modeling: Calibration, validation, and exploration of drivers' cognitive and safety parameters in a risk-taking environment", *Transp. Res., Part B: Methodol.,* vol. 78, pp. 32-53, 2015.
[http://dx.doi.org/10.1016/j.trb.2015.03.011]

[30] Y. Kuang, X. Qu, and S. Wang, "Propagation and dissipation of crash risk on saturated freeways", *Transportmetrica B: Transport Dynamics,* vol. 2, no. 3, pp. 203-214, 2014.
[http://dx.doi.org/10.1080/21680566.2014.930675]

[31] E. Paschalidis, C.F. Choudhury, and S. Hess, "Combining driving simulator and physiological sensor data in a latent variable model to incorporate the effect of stress in car-following behaviour", *Anal. Methods Accid. Res.,* vol. 22, p. 100089, 2019.
[http://dx.doi.org/10.1016/j.amar.2019.02.001]

[32] S. Peeta, P. Zhang, and W. Zhou, "Behavior-based analysis of freeway car–truck interactions and related mitigation strategies", *Transp. Res., Part B: Methodol.,* vol. 39, no. 5, pp. 417-451, 2005.
[http://dx.doi.org/10.1016/j.trb.2004.06.002]

[33] D. Tao, R. Zhang, and X. Qu, "The role of personality traits and driving experience in self-reported risky driving behaviors and accident risk among Chinese drivers", *Accident Anal. Prev.,* vol. 99, pp. 228-235, 2017.
[http://dx.doi.org/10.1016/j.aap.2016.12.009]

[34] S. F. Varotto, H. Farah, T. Toledo, B. van Arem, and S. P. Hoogendoorn, "Modelling decisions of control transitions and target speed regulations in full-range adaptive cruise control based on risk allostasis theory", *Transport. Res. B-Meth.,* vol. 117, pp. 318-341, 2018.
[http://dx.doi.org/10.1016/j.trb.2018.09.007]

[35] Y. Gao, Y. Liu, H. Hu, and Y.E. Ge, "Signal optimization for an isolated intersection with illegal permissive left-turning movement", *Transportmetrica B: Transp. Dyn.,* vol. 7, no. 1, pp. 928-949, 2019.
[http://dx.doi.org/10.1080/21680566.2018.1518734]

[36] Y. Huang, D.J. Sun, and L.H. Zhang, "Effects of congestion on drivers' speed choice: Assessing the mediating role of state aggressiveness based on taxi floating car data", *Accid. Anal. Prev.,* vol. 117, pp. 318-327, 2018.
[http://dx.doi.org/10.1016/j.aap.2018.04.030] [PMID: 29753220]

[37] M.T. Sarwar, P.C. Anastasopoulos, N. Golshani, and K.F. Hulme, "Grouped random parameters bivariate probit analysis of perceived and observed aggressive driving behavior: A driving simulation study", *Anal. Methods Accid. Res.,* vol. 13, pp. 52-64, 2017.
[http://dx.doi.org/10.1016/j.amar.2016.12.001]

[38] M.J.M. Sullman, "The expression of anger on the road", *Saf. Sci.,* vol. 72, pp. 153-159, 2015.
[http://dx.doi.org/10.1016/j.ssci.2014.08.013]

[39] W. Chen, X. Zhuang, Z. Cui, and G. Ma, "Drivers' recognition of pedestrian road-crossing intentions: Performance and process", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 64, pp. 552-564, 2019.
[http://dx.doi.org/10.1016/j.trf.2019.07.004]

[40] Y.S. Chung, "Multilevel modeling of gap-selection behaviors: Effects of vehicle arrival time and personal walking time projection abilities in mixed traffic conditions", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 60, pp. 783-797, 2019.
[http://dx.doi.org/10.1016/j.trf.2018.09.010]

[41] Y.S. Chung, and C.Y. Chang, "Farther and safer: An illusion engendered by incapability?", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 54, pp. 110-123, 2018.
[http://dx.doi.org/10.1016/j.trf.2018.01.020]

[42] J.I. Creaser, M.E. Rakauskas, N.J. Ward, J.C. Laberge, and M. Donath, "Concept evaluation of intersection decision support (IDS) system interfaces to support drivers' gap acceptance decisions at rural stop-controlled intersections", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 10, no. 3, pp. 208-228, 2007.
[http://dx.doi.org/10.1016/j.trf.2006.10.004]

[43]   S.M. Lavrenz, D.V. Pyrialakou, and K. Gkritza, "Modeling driver behavior in dilemma zones: A discrete/continuous formulation with selectivity bias corrections", *Anal. Methods Accid. Res.,* vol. 3-4, pp. 44-55, 2014.
[http://dx.doi.org/10.1016/j.amar.2014.10.002]

[44]   M.A. Pollatschek, A. Polus, and M. Livneh, "A decision model for gap acceptance and capacity at intersections", *Transp. Res., Part B: Methodol.,* vol. 36, no. 7, pp. 649-663, 2002.
[http://dx.doi.org/10.1016/S0191-2615(01)00024-8]

[45]   Y. Ren, Y. Wang, X. Wu, G. Yu, and C. Ding, "Influential factors of red-light running at signalized intersection and prediction using a rare events logistic regression model", *Accid. Anal. Prev.,* vol. 95, no. Pt A, pp. 266-273, 2016.
[http://dx.doi.org/10.1016/j.aap.2016.07.017] [PMID: 27472815]

[46]   Y.V. Tan, M.R. Elliott, and C.A.C. Flannagan, "Development of a real-time prediction model of driver behavior at intersections using kinematic time series data", *Accid. Anal. Prev.,* vol. 106, pp. 428-436, 2017.
[http://dx.doi.org/10.1016/j.aap.2017.07.003] [PMID: 28735178]

[47]   Y. Wu, A.M. Aty, Y. Ding, B. Jia, Q. Shi, and X. Yan, "Comparison of proposed countermeasures for dilemma zone at signalized intersections based on cellular automata simulations", *Accid. Anal. Prev.,* vol. 116, pp. 69-78, 2018.
[http://dx.doi.org/10.1016/j.aap.2017.09.009] [PMID: 28911878]

[48]   D. Chen, J. Laval, Z. Zheng, and S. Ahn, "A behavioral car-following model that captures traffic oscillations", *Transp. Res., Part B: Methodol.,* vol. 46, no. 6, pp. 744-761, 2012.
[http://dx.doi.org/10.1016/j.trb.2012.01.009]

[49]   P.G. Gipps, "A behavioural car-following model for computer simulation", *Transp. Res., Part B: Methodol.,* vol. 15, no. 2, pp. 105-111, 1981.
[http://dx.doi.org/10.1016/0191-2615(81)90037-0]

[50]   Z. He, L. Zheng, and W. Guan, "A simple nonparametric car-following model driven by field data", *Transp. Res., Part B: Methodol.,* vol. 80, pp. 185-201, 2015.
[http://dx.doi.org/10.1016/j.trb.2015.07.010]

[51]   R. Jiang, M.B. Hu, H.M. Zhang, Z.Y. Gao, B. Jia, and Q.S. Wu, "On some experimental features of car-following behavior and how to model them", *Transp. Res., Part B: Methodol.,* vol. 80, pp. 338-354, 2015.
[http://dx.doi.org/10.1016/j.trb.2015.08.003]

[52]   T.H. Law, M. Ghanbari, H. Hamid, A. Abdul-Halin, and C.P. Ng,

"Examining the effect of visual treatments on truck drivers' time-to-arrival judgments of motorcycles at T-intersections", *Transp. Res., Part F Traffic Psychol. Behav.,* vol. 33, pp. 66-74, 2015.
[http://dx.doi.org/10.1016/j.trf.2015.07.007]

[53]   C. Lee, and M. Abdel-Aty, "Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida", *Accid. Anal. Prev.,* vol. 37, no. 4, pp. 775-786, 2005.
[http://dx.doi.org/10.1016/j.aap.2005.03.019] [PMID: 15869737]

[54]   G. Lu, B. Cheng, Q. Lin, and Y. Wang, "Quantitative indicator of homeostatic risk perception in car following", *Saf. Sci.,* vol. 50, no. 9, pp. 1898-1905, 2012.
[http://dx.doi.org/10.1016/j.ssci.2012.05.007]

[55]   K. Long, Y. Liu, L.D. Han, and Y. Wang, "Impact of countdown timer on driving maneuvers after the yellow onset at signalized intersections: An empirical study in Changsha, China", *Saf. Sci.,* vol. 54, pp. 8-16, 2013.
[http://dx.doi.org/10.1016/j.ssci.2012.10.007]

[56]   D. Ngoduy, "Platoon-based macroscopic model for intelligent traffic flow", *Transportmetrica B: Transp. Dyn.,* vol. 1, no. 2, pp. 153-169, 2013.
[http://dx.doi.org/10.1080/21680566.2013.826150]

[57]   J. Sau, J. Monteil, R. Billot, and N.E. El Faouzi, "The root locus method: application to linear stability analysis and design of cooperative car-following models", *Transportmetrica B: Transp. Dyn.,* vol. 2, no. 1, pp. 60-82, 2014.
[http://dx.doi.org/10.1080/21680566.2014.893416]

[58]   Highways England, *Data from: 2016 and 2017,* 2019. Available from: http://trishighwaysengland.co.uk/detail/trafficflowdata

[59]   "Transport for London Unified API", Available from: https://api.tfl.gov.uk/swagger/ui/index.html?url=/swagger/docs/v1%23!/AccidentStats/AccidentStats_Get#!/AccidentStats/AccidentStats_Get

[60]   J. Ooms, "The JSONlite package: A practical and consistent mapping between JSON data and R objects", *arXiv:1403.2805,* 2014.

[61]   E. Pebesma, "Sp (version 1.3-2) classes and methods for spatial data", Available from: https://rdocumentation.org/packages/sp/versions/1.3-2

[62]   M. Tennekes, "Tmaptools: Thematic map tools", Available from: https://cran.r-project.org/web/packages/tmaptools/index.html

[63]   *Know your traffic signs,* 5th ed.The Stationery Office: London, UK, 1975.