Shockwave identification and critical parameter estimation using a detailed dataset from a swarm of drones

Emmanouil Barmpounakis*[1], Nicolas Sébastien Richter[2], Nikolas Geroliminis[3]

[1] Postdoctoral Researcher, Laboratory of Urban Transport Systems, EPFL, Switzerland

[2] MSc Student, EPFL, Switzerland

[3] Professor, Laboratory of Urban Transport Systems, EPFL, Switzerland

SHORT SUMMARY

In the era of big data, new transportation-related concepts and methodologies need to be proposed for understanding how congestion propagates. pNEUMA, a unique dataset that was acquired during a first-of-its-kind experiment with a swarm of drones over a dense city center, has uncovered new opportunities for revisiting and evaluating existing concepts but also for new ways to describe significant traffic related phenomena. This dataset is part of an open science initiative shared to the research community and consists of more than half million detailed trajectories of almost every vehicle that was present in the study area. The aim of this paper is to provide a first methodological approach on how to extract information regarding shockwaves. First, we identify the critical start and stop points for every vehicle and using two different approaches (a rule-based and an unsupervised clustering technique) shockwaves are identified for multilane arterials in space and time. It is seen that the rule-based approach is based on shockwaves characteristics according to traffic flow theory and can successfully identify shockwaves and spillbacks when the location of traffic signals is known, while an unsupervised clustering technique is used to identify stop-and-go conditions. Finally, the different characteristics per lane and per shockwave are calculated and compared.

Keywords: drone dataset, traffic flow analysis, shockwave analysis, spillback

1. INTRODUCTION

Shockwave analysis has been extensively used in traffic flow analysis since they were first introduced by (Lighthill and Whitham, 1955) and (Richards, 1956). Most of the studies until now focus on shockwaves' characteristics to study different measures, such as delays, using microscopic traffic simulation software (Dion et al., 2004). In (Cho and Tseng, 2007), authors use the stopped time of vehicles to detect shockwaves and evaluate their methodology utilizing simulated data. In (Izadpanah et al., 2009) authors use an iterative two-phase piecewise regression to identify the "inflection points" of shockwaves and then a data filtering and linear-clustering algorithm to detect shockwaves, which was the main challenge. Then, their methodology was tested using both simulated and real data from the NGSIM dataset (NGSIM, 2006). Similar methodologies by identifying the "critical points" and then evaluated using both simulated data and real data from the NGSIM dataset is utilized in other studies, like (Wang et al., 2019) who combined shockwave analysis with Bayesian Networks to estimate traffic parameters at signalized intersections or (Cheng et al., 2010) to detect signal timing and estimate queue length. One of the limitations of the latest study is the identification of stop-and-go conditions. In (Elfar et al., 2018) shockwaves were analyzed in an assumed connected or partially connected environment using NGSIM data. Other previous works emphasize on the significance on shockwave analysis regarding travel times and cause of traffic delays (Ramezani and Geroliminis, 2015; Yildirimoglu and Geroliminis, 2013). One of the first studies that focused on congested arterial networks using the shockwave profile model (SPM) is (Wu and Liu, 2011). While their methodology was replicated using data from a field test on an arterial corridor, authors state that queuing dynamics should be further investigated using better data, such as from video recordings. Another limitation in studying shockwaves has been reported in (Li et al., 2017), where authors estimate dynamic traffic shockwave speeds focusing in a connected vehicle environment using high temporal and spatial resolution data, concluding that their methodology should be also extended to multi-lane scenaria. As increased automation is expected to prevent shockwave formation and propagation of congestion, new mechanisms should be developed to better understand how shockwaves move through traffic (Talebpour and Mahmassani, 2016).

In October 2018, a large scale experiment, nicknamed pNEUMA, was conducted to record traffic streams over an urban setting using drones and to provide significant insight on how their unique characteristics can overcome existing limitations in traffic monitoring and recording traffic streams (Barmpounakis and Geroliminis, 2020). A swarm of drones was flying over the congested city of Athens, Greece and produced a unique dataset of over half a million trajectories (Figure 1). This massive dataset contains trajectories of every vehicle that was present in the study area, calibrated in the WGS-84 system, every 0.04 seconds, as this is the maximum frequency allowed by the video's frame rate. The advantages of extracting trajectories from drone videos and their potential in shockwave analysis have been described in (Khan et al., 2018), although authors do not describe how this methodology performs in such data.

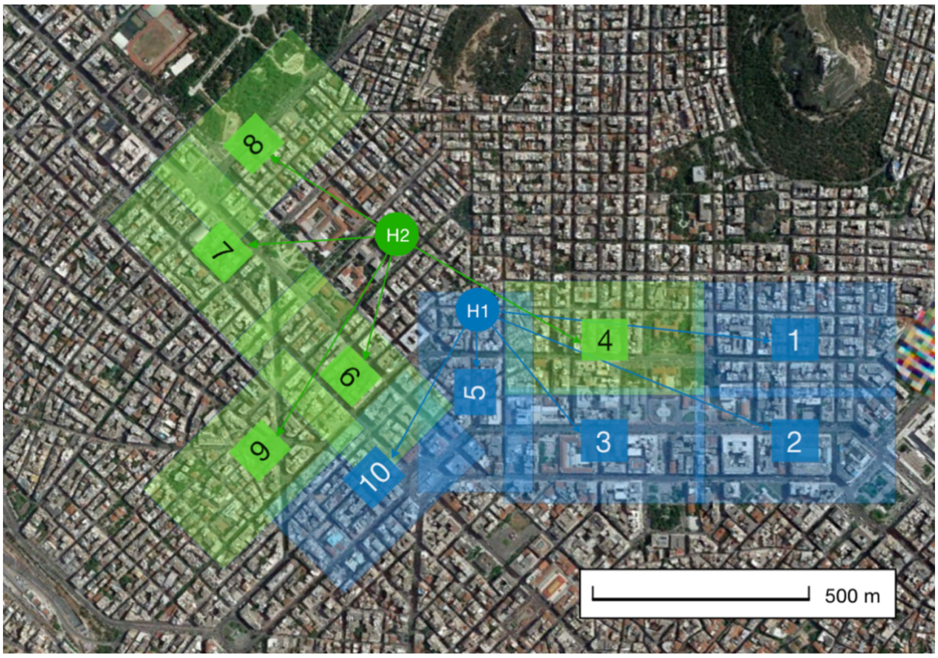

Figure 1: The study area of pNEUMA experiment

The aim of this study is to provide a first methodological approach on how such data as the pNEUMA dataset, can be utilized to conduct shockwave information from this new kind of data that is available and set the benchmark for possible future approaches.

2. METHODOLOGY

2.1. Data Collection

More details on the design of the experiment can be found in (Barmpounakis and Geroliminis, 2020). For the specific study, the sample dataset that can be found in *https://open-traffic.epfl.ch* is used. The sample dataset was collected by one drone during 10:00-10:30 from the intersection of the Alexandras and 28th Oktovriou avenue (Figure 2). This 3-lane arterial includes three intersections and the positions of the three traffic signals are at 110m, 260m and 350m starting from upstream point (0,0) as indicated in the figure. It includes 2478 trajectories from cars, taxis, motorcycles, buses, heavy and medium vehicles. However, since motorcycles follow chaotic trajectories that would infer noise in our analyses, they are not taken into account for the current study.

Figure 2: The study area is a 3-lane arterial including three signalized intersections

2.2. Identification of critical points

The identification of the start and stop points of the vehicles is the first step to analyze shockwaves. It is assumed that a vehicle travelling at time $i-1$ at a higher speed than a critical value parameter and at time $i+1$ at a lower speed than the critical value, is considered to be a vehicle about to stop. Similarly, it is assumed that a vehicle travelling at time $i-1$ at a speed lower than the critical value and at time $i+1$ at a speed higher than the parameter, is considered to be a vehicle about to accelerate again. Thus, $\{A,\ S\} = V_{t-1} \lesseqgtr 1 \lesseqgtr V_{t+1}$, where A are the acceleration points and S are the stopping points at time t, and the critical value is equal to 1 km/h. Then, the critical points are filtered based on the lanes' coordinates (Figure 3: Critical points during acceleration (green) and deceleration (red) for all lanes in the time space diagramFigure 3). It should be noted that ongoing work of the authors aims to automatize the process of lane assignment of each trajectory or point in the network.

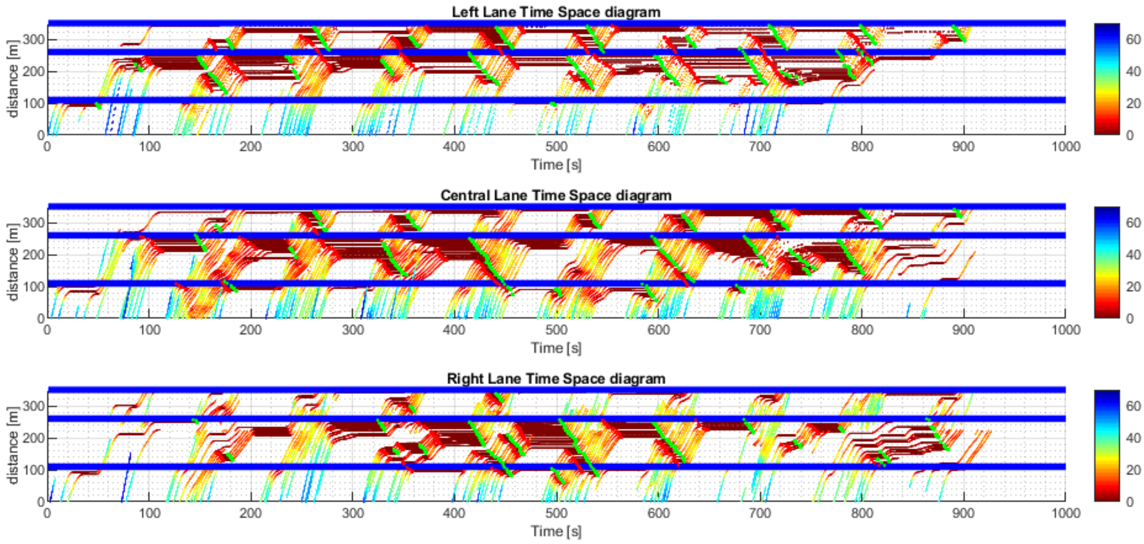

Figure 3: Critical points during acceleration (green) and deceleration (red) for all lanes in the time space diagram

Then data are filtered to deal with noisy data around the critical value. Specifically, two points from the same vehicle separated by less than 5 sec are considered to be as one point. The final result of this process is illustrated in a time-space diagram in Figure 4, for a specific lane. It can be seen that the form of the shockwaves is easily recognizable while the cloud of points representing stop-and-go conditions can also be distinquished.

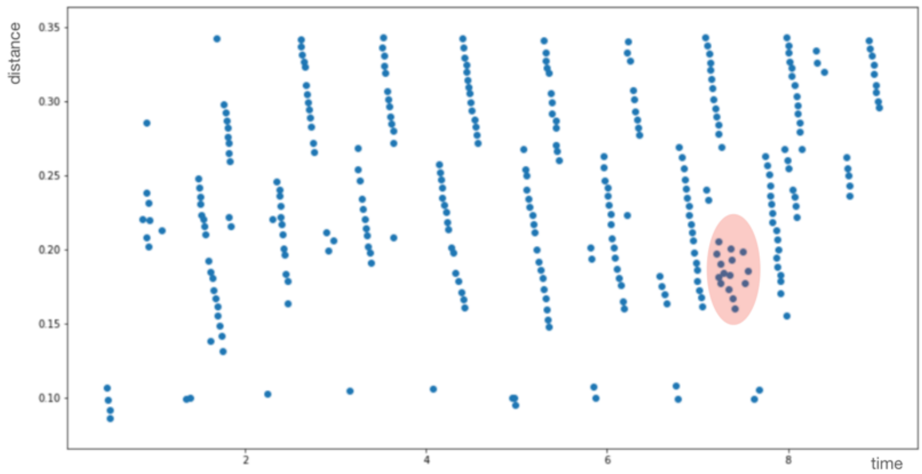

Figure 4: Critical points during acceleration for the left lane

When it comes to spillback phenomena, that concern the blockage of an intersection due to the upstream traffic light, they are examined using the stopping points. Specifically, as the position of all traffic signals in the study area is known, a spillback will be defined as the intersection point between the stopping shockwaves and for distance equal to position of the traffic signal (Figure 5).

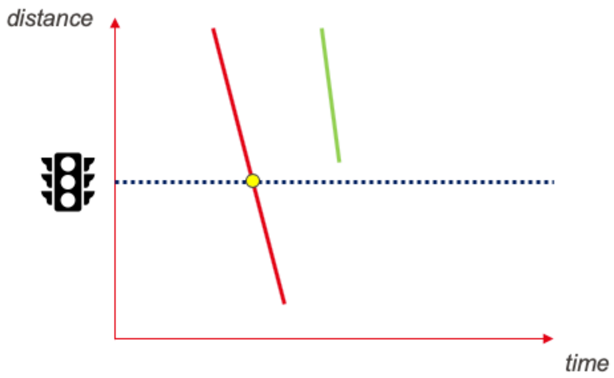

Figure 5: Spillback illustration in the time space diagram

2.3. Rule-based approach

The set of rules that define a shockwave are described below and illustrated in Figure 6. The first condition sets a maximum time interval (10000 msec) between points j and i. The second condition is based on the assumption that a shockwave moves through traffic not faster than 20 km/h. The third condition is that point j must be closer to the origin than point i, ensuring that the slope of the shockwave is negative. By combining these four conditions in order to identify shockwaves, the search perimeter of points becomes a surface (shown in blue in Figure 6). If points i, j and k meet the four conditions then the three points will belong to the same shockwave and therefore will have the same shockwave's ID.

$$time_j - time_i < 10000 \ (1)$$
$$distance_i - distance_j < 20 \cdot \frac{time_j - time_i}{3600 \cdot 1000} \ (2)$$
$$distance_j - distance_i < 0 \ (3)$$
$$time_j - time_i > 0 \ (4)$$

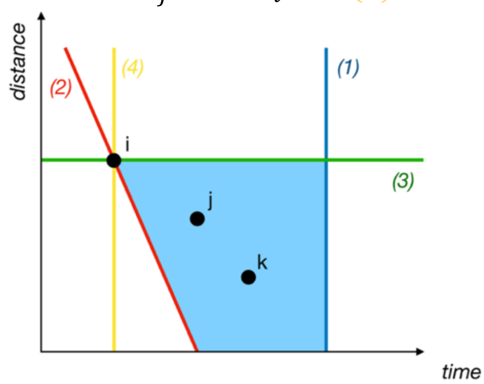

Figure 6: Illustration of the search area for the rule-based identification of shockwaves

2.4. Unsupervised Clustering approach

For the specific study, the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm is applied to the starting and ending points. Unlike other clustering algorithms like k-means, DBSCAN figures out the number of cluster while two parameters are required to be determined. The first parameter, *epsilon* corresponds to the maximum distance between points. The value of *epsilon* is set by calculating the distance to the nearest n points for each point, sorting and plotting the results.

Then, the optimal value for *epsilon* is set as the point of maximum curvature in the plot. An automated method to calculate the optimal value *epsilon* based on the abovementioned process can be found in (Rahmah and Sitanggang, 2016). The second parameter, *minimum_samples* sets the minimum number of points that can form a cluster.

3. RESULTS AND DISCUSSION

3.1. Rule based approach

In Figure 7, it can be seen that the deterministic way of the rule based approach, results in the loss of some of the shockwaves and does not perform well in identifiying smoother shockwaves or stop-and-go conditions. It should be noted that while different thresholds were examined, the ones that are described above are the ones that produce the best results. It is seen that for the left and central lanes shockwaves are identified during almost all periods. However, for the right lane, unlike the two previous ones, shows shockwaves only during a period of about 350-600 sec.

Despite the uncertainty of the rule-based approach, it can be seen that shockwaves appear mostly in the left and central lane, while for the right lane shockwaves appear between the first and second traffic signal, where a bus stop is present. Moreover, upstream the second traffic signal an extra lane is added that increases the capacity as only vehicles from the right lane can reach the fourth lane).

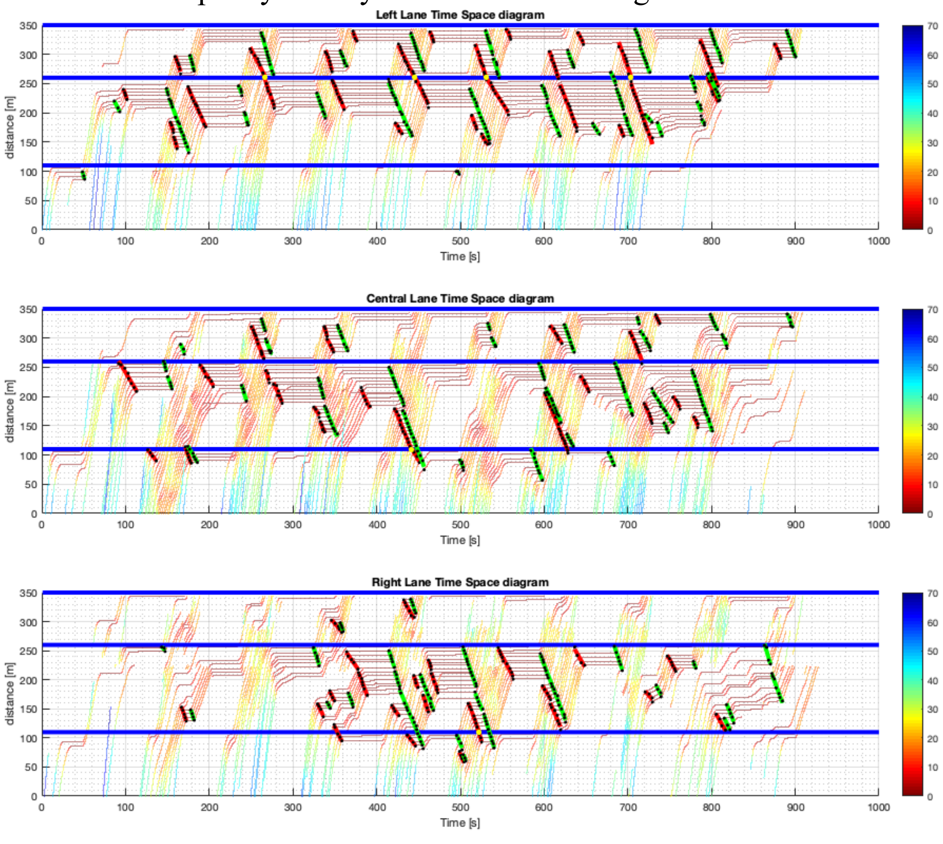

Figure 7: Illustration of identified shockwaves (red lines illustrate stopping waves, green lines illustrate start shockwaves, blue horizontal lines represent the three traffic signals and yellow points illustrate spillbacks)

3.2. Unsupervised Clustering

For the specific paper, only the results from the left lane are presented. One of the issues that emerged was that DBSCAN is very sensitive to scale since *epsilon* is a fixed value for the maximum distance between two points. The arithmetic values of the spatial and temporal information of the critical points have different ranges which then affects the calculation of the *epsilon* value in favor of the parameter with the larger values. For this case, the range of distance is [0.08, 0.343] in km while the time's range is [47120-900440] in msec. It should be noted that different normalization techniques were first applied that led to misclasification of shockwaves as the physical meaning of time and space was not taken into account. Thus, it was found that when the values of time are divided by 10^6, the algorithm's sensitivity is "manipulated" so that the clusters that are formed have a physical meaning. In Figure 8, the result for the estimation of the optimal *epsilon* value is illustrated.

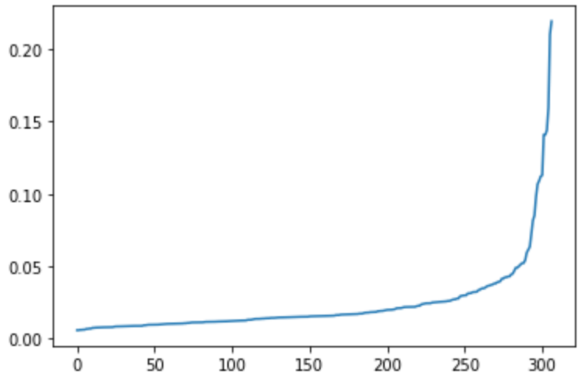

Figure 8: Estimation of the optimal epsilon value for the DBSCAN algorithm.

Then, our model is trained selecting 0.05 for *epsilon* and setting *minimum_samples* equal to 3. In Figure 9 and Figure 10 the results of the DBSCAN algorithm are presented. It is seen that the model has classified the densely populated areas, and shockwaves are presented better than the rule-based approach. The noise (critical points that are not part of a shockwave) is categorized with grey color.

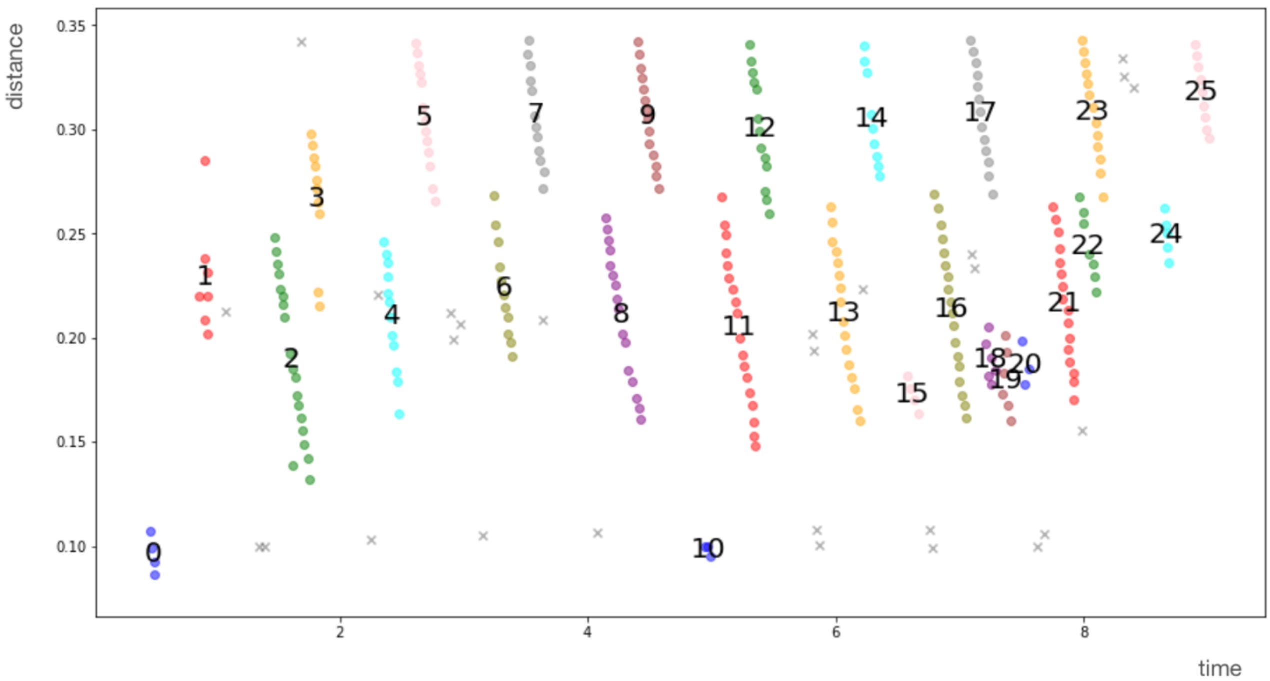

Figure 9: Identified starting shockwaves during acceleration

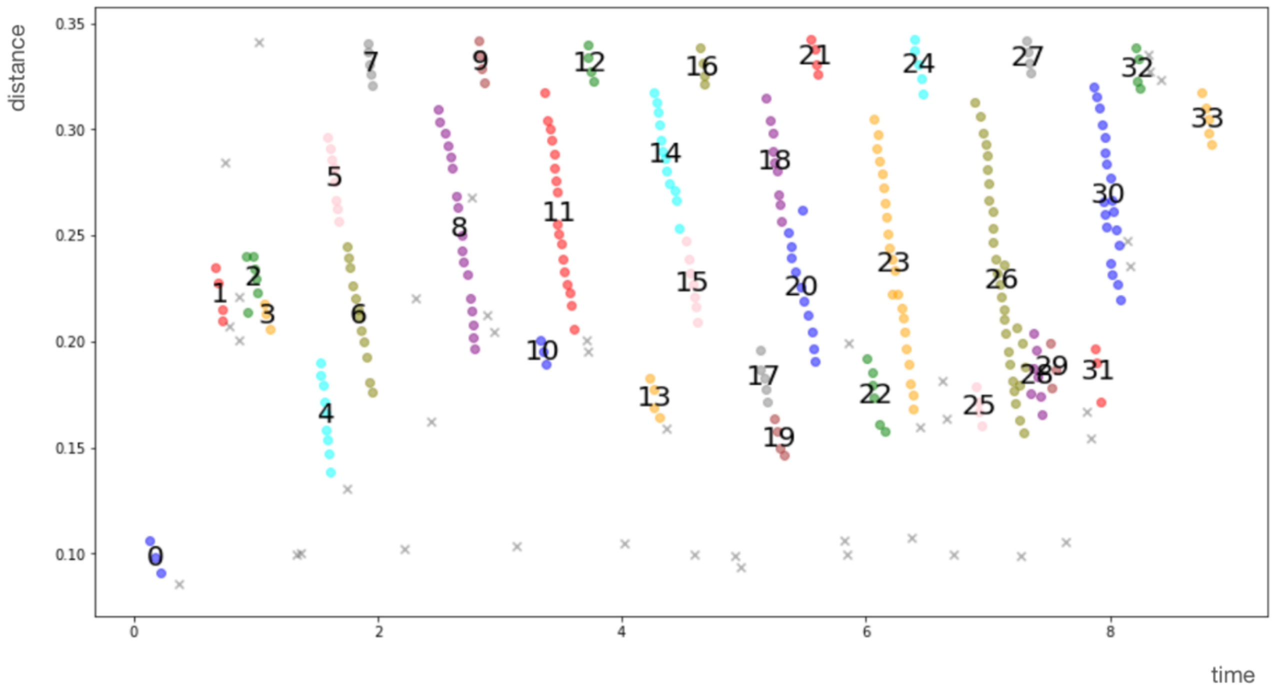

Figure 10: Identified starting shockwaves during deceleration

After the clusters/shockwaves have been formed, they can be represented as a first degree polynomial, which allows the identification of their key characteristics, such as the speed and duration of the shockwave (Figure 11).

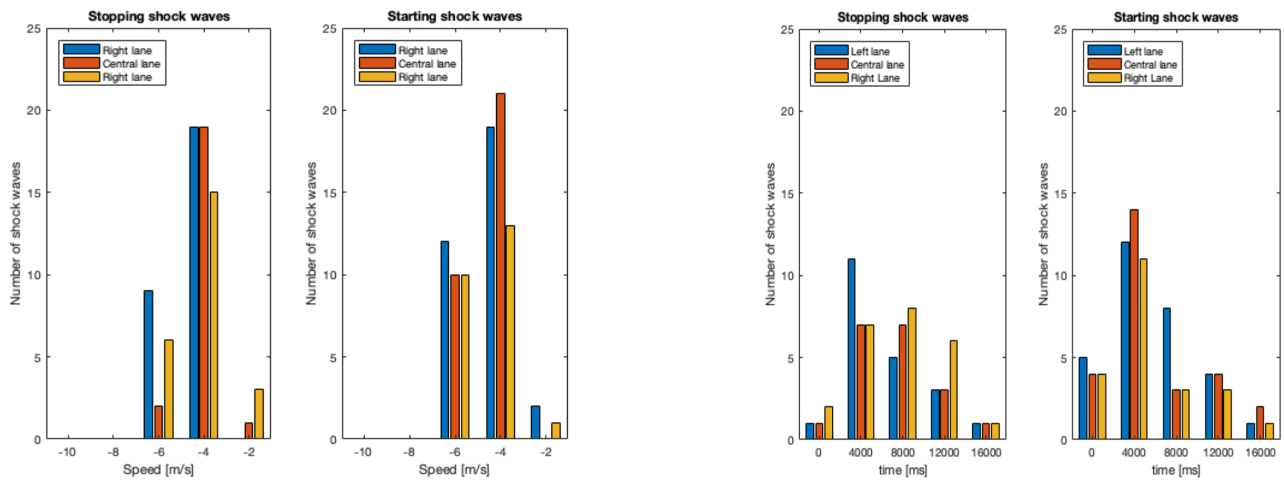

Figure 11: Histograms for (a) speed and (b) duration of shockwaves

4. CONCLUSIONS

In this paper, the pNEUMA dataset is utilized to identify shockwaves using two different approaches. It is seen that the DBSCAN algorithm, after the necessary modifications and parameter estimation performs better than the rule-based approach. Shockwaves can be detected by first identifying the critical points during acceleration and deceleration and their key characteristics such as speed and duration can be extracted.

In our future research attempts, we aim to use more data and examine how well the DBSCAN in identifying shockwaves for different arterials and congestion levels. In addition, we aim to further study and model the propagation of congestion in terms of shockwave formation and associate it with temporal and spatial characteristics and the formation of spillbacks.

ACKNOWLEDGMENT

Data source: pNEUMA - open-traffic.epfl.ch
This research was partially funded by Swiss National Science Foundation (SNSF) grant (200021_188590) "pNEUMA: On the new era of urban traffic models with massive empirical data from aerial footage".

REFERENCES

Barmpounakis, E., Geroliminis, N., 2020. On the new era of urban traffic monitoring with massive drone data: The pNEUMA large-scale field experiment. Transp. Res. Part C Emerg. Technol. 111, 50–71.

Cheng, Y., Qin, X., Jin, J., Ran, B., 2010. An Exploratory Shockwave Approach for Signalized Intersection Performance Measurements Using Probe Trajectories. Environ. Eng. 1–23.

Cho, H.J., Tseng, M. Te, 2007. Shockwave detection for electronic vehicle detectors, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 275–282. https://doi.org/10.1007/978-3-540-72590-9_39

Dion, F., Rakha, H., Kang, Y.S., 2004. Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections. Transp. Res. Part B Methodol. 38, 99–122. https://doi.org/10.1016/S0191-2615(03)00003-1

Elfar, A., Xavier, C., Talebpour, A., Mahmassani, H.S., 2018. Traffic shockwave detection in a connected environment using the speed distribution of individual vehicles. Transp. Res. Rec. 2672, 203–214. https://doi.org/10.1177/0361198118794717

Izadpanah, P., Hellinga, B., Fu, L., 2009. Automatic Traffic Shockwave Identification Using Vehicles ' Trajectories, in: Transportation Research Board 88th Annual Meeting.

Khan, M.A., Ectors, W., Bellemans, T., Janssens, D., Wets, G., 2018. Unmanned Aerial Vehicle-Based Traffic Analysis : A Case Study for Shockwave Identification and Flow Parameters Estimation at Signalized Intersections. Remote Sens. https://doi.org/10.3390/rs10030458

Li, J., Wang, C., He, S., Qiu, T.Z., 2017. Dynamic traffic shockwave speed estimation in connected vehicle environment, in: 2017 4th International Conference on Transportation Information and Safety, ICTIS 2017 - Proceedings. pp. 54–59. https://doi.org/10.1109/ICTIS.2017.8047742

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II . A theory of traffic flow on long crowded roads. Proc. R. Soc. London, Ser. A 229, 317–345.

NGSIM, 2006. Next generation simulation [WWW Document]. URL http://ngsim.fhwa.dot.gov/

Rahmah, N., Sitanggang, I.S., 2016. Determination of Optimal Epsilon (Eps) Value on DBSCAN Algorithm to Clustering Data on Peatland Hotspots in Sumatra, in: IOP Conference Series: Earth and Environmental Science. https://doi.org/10.1088/1755-1315/31/1/012012

Ramezani, M., Geroliminis, N., 2015. Queue Profile Estimation in Congested Urban Networks with Probe Data. Comput. Civ. Infrastruct. Eng. 30, 414–432. https://doi.org/10.1111/mice.12095

Richards, P.I., 1956. Shock Waves on the Highway. Oper. Res. 4, 42–51. https://doi.org/10.1287/opre.4.1.42

Talebpour, A., Mahmassani, H.S., 2016. Influence of connected and autonomous vehicles on traffic

flow stability and throughput. Transp. Res. Part C Emerg. Technol. 71, 143–163. https://doi.org/10.1016/j.trc.2016.07.007
Wang, S., Huang, W., Lo, H.K., 2019. Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian Network. Transp. Res. Part C Emerg. Technol. 104, 22–37. https://doi.org/10.1016/j.trc.2019.04.023
Wu, X., Liu, H.X., 2011. A shockwave profile model for traffic flow on congested urban arterials. Transp. Res. Part B Methodol. 45, 1768–1786. https://doi.org/10.1016/j.trb.2011.07.013
Yildirimoglu, M., Geroliminis, N., 2013. Experienced travel time prediction for congested freeways. Transp. Res. Part B Methodol. 53, 45–63. https://doi.org/10.1016/j.trb.2013.03.006