IJTE, Vol. 11, No. 4, Spring 2024, pp. 1595-1616
Research Paper

Designing and Developing a Model for Detection of Unusual Traffic Condition at Intelligent Signalized Intersection Equipped with SCATS

Mani Hazeghi[1], Mahmoud Saffarzadeh[2,*], Babak Mirbaha[3]

Received: 2023/03/25 *Accepted: 2023/06/28*

Abstract

Today, one of the most significant points of interference in streets is the signalized intersections; therefore, solving problems of traffic at intersections can increase the capacity of urban transportation. Inability to diagnose the traffic conditions results in the lack of proper timing, phasing and cycle length, all of which are attributed to the abnormal factors concerning intelligent control systems. In this paper, in addition to the introduction of abnormal traffic conditions at signalized intersections, an attempt has been made to intelligently diagnose anomalies for both an approach and its entire intersection. For this purpose, by making use of the data based on the GPS of users' cell phones extracted from NESHAN Application, which consists of 10-minute average speed in streets ending to an intersection, and by behavioral matching with the data concerning the volume and saturation rate in SCATS, and meantime, by analyzing the fundamental traffic relations, an attempt has been made to diagnose the abnormal traffic conditions through SCATS at Toos-Danesh intersection of Mashhad, in which abnormal conditions including the detecting of heavy traffic conditions when the traffic is light and vice versa. To achieve more accuracy, the method was built based on both quartiles and percentiles of DS, degree of saturation, and ADS, average degree of saturation, in SCATS. Finally, anomaly detection based on the 10th and 90th percentiles had 100 percent accuracy and the one based on the 1st and 3rd quartiles had between 57 to 80 percent accuracy, which have been checked by two real datasets.

Keywords: Anomaly, Degree of saturation, Fundamental Traffic relations, SCATS, Signalized intersection

---

[*] Corresponding author. E-mail: Saffar_m@modares.ac.ir

[1] Dept. of Civil Engineering Central Tehran Branch, Islamic Azad University, Tehran, Iran

[2] Dept. of Civil & Environmental Engineering, Tarbiat Modares University, Tehran, Iran

[3] Dept. of Civil Engineering Faculty of Imam Khomeyni, Ghazvin, Iran

1. Introduction

Intersections which are one of the most important parts of urban transportation network are especially significant as the nodes of the network and the meeting points of streets. They are the points where traffic flows interfere, the most essential factor in increasing trip duration, and the range of delays in traffic flows. Considering the limited finance sources, administering projects like constructing non-level intersections, purchasing and equipping the signalized intersections with

Other kinds of intelligent control systems, and the like, cost a lot. Therefore, under the present conditions, the best method to optimize the performance of the existing systems of intelligent control at the signalized- intersections, so that the best control conditions for traffic lights can be offered. SCATS is one of the intelligent systems for the management of traffic lights. In contrast to the other methods of traffic management, this system does not utilize the mathematical models and theories for the betterment of timing. In SCATS, an index called degree of saturation 'DS.' has a vital role in the selection of appropriate timing and phasing, the duration of the cycle at intersections, offsets, and coordination. Unfortunately, because drivers neither observe space time, nor drive between the lines, nor observe the speed limits to pass the intersections, the data acquired from their detectors differ from the actual conditions, which will drastically decrease the efficiency of the system. The resulting error of detectors or the incorrect planning of SCATS will lead to an error in the estimation of D.S for various approaches of intersection and hence, result in the wrong decision making of the system. Not diagnosing the abnormal conditions for SCATS, in due time, can be followed by several delays at signalized intersections and, considering the resulting effects on the other equipment of controlling the traffic, new delays will occur in other parts of the street or intersection. This will remain a defective cycle for a rather long time. Therefore, the punctual diagnosis of unstable or abnormal traffic conditions in the network of urban streets or at signalized intersections can enable the experts and city directors to make preventive decisions.

Most of the studies related to traffic flow prediction have been proposed and tested under either normal, recurrent, or non-incident conditions, and hence limit their applicability to other scenarios. A robust traffic flow prediction method should produce accurate prediction results across a wide range of traffic conditions, especially for unexpected incidents, which cause traffic flow breakdown in a short period. Those abnormal traffic conditions can be classified into two categories: (i) long-term and planned incidents, e.g. work zones, and holidays; and (ii) short-term unexpected incidents, e.g. accidents, adverse weather, and temporary traffic control. As the road traffic is a complex, non-linear system involving people, vehicles, and transport facilities, its characteristics make the short-term traffic flow prediction difficult to implement. In this complex huge system, the road traffic is often disturbed by abnormal incidents. However, most of the existing traffic flow forecasting methods are trained under normal conditions, which may generate a large prediction error during the abnormal conditions.

In addition to the above-mentioned cases, limited availability of human resources as the controlling factors for these systems is also one of the cases that make it a must to plan an intelligent system aimed at responding as soon as possible to errors of decision-making systems. Offering a method for the detection of abnormal traffic conditions and, at the same time, taking into account behavioral characteristics of the citizens include the significant goals of this paper. The decreasing indices such as the delay and travel time, queue length, likewise, increasing the level of citizens' consent, together with shortening the time of response in emergencies are among the cases that show the significance of dealing with such research. It must be added that the use of data based on users' cell phones in a part of street to diagnose a traffic state and matching it with the performance of SCATS at a signalized intersection (especially in conditions under which SCTAS make decisions based on gathered indices at the entrance of intersections or eventually in a part of the street) are among the parameters that can be very helpful in diagnosing proper or improper decisions of the intelligent system at intersections. Using suggested algorithms can be a modern method of calibrating the performance of urban signalized intersections. By using this method, it can be offered offer the fastest response to the wrong decisions of local and non-local intelligent systems. Because the decision made by systems like SCATS depend on numerous factors like traffic culture, the geometry of the intersection, and approaches leading to it; all these cases can cause errors in decision making of these systems. Suggesting an algorithm that can attribute the traffic condition of the street to the intersection and, on this basis, can manifest the abnormal traffic condition in the street is often important in the management and as quick as possible response to various traffic conditions; especially when the system is unable to respond.

The expected benefit from SCATS comes from its ability to constantly modify signal timing patterns to most effectively accommodate changing abnormal traffic conditions. While the potential benefits from this control structure may be significant, few research studies have documented the effect of implementing this method of signal control. The South Lyon signal improvement project is a small segment of the FAST-TRAC (Faster and Safer Travel-Traffic Routing and Advanced Controls) project [RCOC, 1993]. FAST-TRAC will involve the conversion of more than 1000 pretimed and actuated signalized intersections in Oakland County to SCATS control and will establish a county-wide, real-time route navigation system. While FAST-TRAC has been managed primarily by the Road Commission for Oakland County (RCOC), it has been a cooperative effort between federal, state, county, and local government agencies; as well as private corporations and universities.

SCATS was developed in the 1970s by the Roads and Traffic Authority of New South Wales, Australia. It functions by modifying traffic signal timings in real-time in response to the variations in traffic demand and system capacity. It operates by using traffic sensors to monitor flow conditions and thus coordinate signal timings to minimize stops and delay time when the system is at or near capacity. SCATS attempts to maximize the system capacity and minimize the possibility of traffic jams by controlling the formation of queues Lowrie, 1990, Lowrie, 1982.

The traffic information collected in the field involves the discharge characteristics (i.e., flow and occupancy during the green phase) on each intersection approach. This data is processed by the local controller or transmitted to a regional control center where the SCATS control program attempts to most effectively maintain progression with the intersection downstream of the collected traffic data. Other differences that separate SCATS from conventional fixed timed systems are its ability to modify timing strategies to fit various control philosophies and to collect, process, and maintain a history of traffic statistics for an area. Signal phases can be set to equalize saturation on all approaches or they can be arranged to give priority to a particular direction of importance. Since the SCATS system requires the use of certain traffic data information it can record and store these statistics to monitor the strategic performance of the system, detect signal faults, and allow manual overrides of the signals under special operating circumstances.

2. Literature Review

The former researches are categorized into numerous groups. The classification of subject studies is mainly divided into classification related to detecting traffic conditions at signalized intersection and researches related to the macroscopic fundamental diagram (MFD) to identify level of service or abnormal conditions. In category for detecting traffic conditions, Yi et al partitioned the urban network into several traffic clusters by using a traffic state index and network topology. They obtained a traffic state index with license plate recognition data and network topology information, considering the heterogeneity and direction of segments. They modified the Ratio-Cut algorithm and turn hyper-parameters automatically to identify traffic clusters in the road graph. The visualization and comparison results with conventional spectral clustering, DBSCAN, and K-means demonstrated the superiority and application value of the proposed approach [Yi et al, 2022]. Jiang et al tried to partitioning city network into homogeneous subzones according to network macroscopic fundamental diagram (NMFD). In their study six-step partitioning algorithm was proposed. The framework included graph definition, data preprocessing, feature handling, clusters and partitions identification, and boundaries reshaping. Tests on a simplified grid network and the city of the Melbourne Road network demonstrated the suitability of the framework for characterizing the traffic states by the partitions [Jiang et al, 2022]. Shafiei et al by integrating data-driven and traffic simulation modelling approaches, instead of directly predicting the traffic states, by using limited historical data, tried to predict a traffic states affected by road accidents. In this regard, they used a traffic Simulation reinforced with data-driven models. They showcased the capability of the proposed data-driven enforced traffic simulation platform for incident impact analysis in a real-life network in Sydney, Australia. [Shafiei et al, 2022]. Zheng et al proposed a multi agent reinforcement learning (MARL) traffic control strategy for macroscopic fundamental diagram region. In their proposed MARL strategy, a reinforcement learning (RL) agent adjusted the traffic signal for each intersection in the MFD region. The reward is defined as the unweighted sum of rewards for the level of services (LOS) of the intersection and for the weighted average density within the MFD region [Zheng et al, 2021].

Piyapong examined how to optimize traffic management under flood conditions. He proposed a method to better understand the impact of urban flood situations by expressing traffic conditions in specific ranges using the concept of a Macroscopic Fundamental Diagram (MFD). Using MFD analysis, he identified the traffic flow density and density-speed relationships by using the shape of the estimated MFD travel time series graphs and applied them to the Dynamic Traffic Allocation (DTA) model as traffic flow parameters to propose a method for road network performance [Piyapong, 2021].

Ranaweera et al used traffic flow theory for detecting anomalous data in urban network, by evaluating the consistency of microscopic parameters which are derived by traffic flow theory (i.e., speed and space-headway). Numerical results showed that in congested segments, higher accuracy belongs to space-headway detection and the anomalous speed detection performs well across all the traffic conditions [Ranaweera, 2021]. Mihaita et al presented an advanced deep learning framework for traffic flow prediction and defined high noise spots as outlier data and generation of spatial and temporal features to detect anomalies in Sydney Australia, in the following, they discussed that the hybrid CNN-LSTM, compared individual LSTMT, doesn’t improve the forecast accuracy for highway flow prediction. They further emphasized that removing outlier data is highly effective in detecting abnormal events [Mihaita et al, 2020].

Kinane et al, analyzed a number of heterogeneous datasets to determine abnormal conditions on the M50 motorway of Dublin with the title of the research program called INSIGHT. The dataset involved monitoring real-times SCATS, Dublin bus travel times and Twitter data. The process of detecting anomalies was done by comparing the dataset, such as travel time, in accident conditions with baseline reference (normal conditions) [Kinane et al, 2018]. Dakic et al developed flow-density relationship by using degree of saturation (D.S), a density-like measure from SCATS (Sydney Coordinated Adaptive Traffic System) at signalized intersections. The findings showed that D.S does not have to be a poor estimator of traffic conditions, but when it is combined with SCATS-measured traffic flow it gives a false representation of near-capacity and over-saturated conditions. They went on to discuss possible errors in the use of this type of data. Erroneous data at the time of oversaturation or sensor counting errors, were considered among these faults [Dakic et al, 2017].

In January 2017, Ken Michael et al and his team referred to the high potential of using big data in the control of congestion on urban roads. The uses of such systems as SCATS, SCOOT, OPAC, and the like are among these systems. They introduced a suitable tool in Australia called NETPRES (Network Performance Report System) which can combine different data from different sources and considered the performance index of the network from different perspectives. Moreover, the mere use of spatial and temporal data without using other informative sources will probably bring about errors in the used algorithms of diagnosing abnormal conditions, because the concept of abnormal traffic conditions is entirely different from that of congestion, the diagnosis of which requires a complete understanding of the numerous indicators affecting the abnormal traffic situation [Ken Michael et al, 2017].

Polson and Sokolov tried to predict the flow within a short time through Deep Learning. In this research, they tried to predict the traffic flow by using the relation of data collected from inductive sensors with the data collected from GPS; in the meantime, extreme discontinuity in the network was analyzed. At the end, the presentation of a model for the nonlinear spatiotemporal effects on repeating and non-repeating congestions was made [Polson and Sokolov, 2017].

Accurate traffic flow prediction under abnormal conditions, such as accidents, adverse weather, work zones, and holidays, is significant for proactive traffic control. Four different model structures whether considering the feature selection are proposed and tested for multi-step-ahead prediction under both normal and abnormal conditions. The results indicate that the proposed multi-model ensemble models are superior to the benchmark algorithms, i.e., support vector regression, and random forests, the GBRT model outperforms the Lasso model under normal traffic conditions, and the Lasso model has a better prediction accuracy under abnormal traffic conditions. In addition, the Lasso model with the feature selection is superior to the full feature model under either normal or abnormal conditions, while the GBRT model is not always better under normal conditions. The proposed integration framework is general and flexible to assemble various traffic prediction algorithms [Chen et al., 2019].

The prediction of abnormal traffic flows has always been a primary concern in traffic management. If the management unit can predict the occurrence of abnormalities, it can manage and control transportation in advance to avoid abnormal traffic flows and enhance the service quality. While previous researchers generally predicted abnormal conditions in traffic flows using a single time series prediction model, such methods might result in inaccuracy in the prediction of abnormal traffic flows in some areas. As a result, in terms of practicality, the traditional methods fail to produce satisfactory results. Moreover, with the traditional methods, researchers often find it difficult to obtain the best time-delay item in the model and can only do so by using the trial-and-error method, which is nevertheless cost-prohibitive. To solve this problem, they predicted abnormal traffic flows by using the Ensembling-mRBF-LSTM framework. With the aforementioned three major concepts, in this study, they sought to vastly improve the existing method for abnormal traffic flow prediction. Lastly, their study provided the traffic flow data of Taipei Mass Rapid Transit (MRT) in Taiwan to verify the effectiveness of the proposed method [Chen et al., 2020].

Most approaches are developed or trained under normal traffic conditions, but traffic states under abnormal conditions have not been well studied. Since unexpected incidents will cause rapid changes in traffic flow, the uncertainty evolution of traffic states under abnormal conditions needs further research. Hence, there have been some related studies that attempt to develop prediction models for abnormal conditions over the past decade. Guo et al. compared the prediction accuracy of three models with different input structures, and the results indicated that k -NN had the best prediction accuracy for a given model structure during non-recurring congestion conditions. However, only the one-step-ahead experiments were conducted [Guo et al., 2010]. Wu et al. proposed an online boosting approach model which composed of two parts: the base part dealt with the normal prediction, while the boosting part adapted for abnormal conditions. A supervised statistical learning technique called the online support vector machine for regression (OL-SVR) was presented for traffic flow prediction on either normal conditions or days with traffic incidents on freeways. OL-SVR was found to be the best performer under non-recurring traffic conditions. An NN model being incorporated with the backpropagation algorithm was developed to forecast the impact of adverse weather conditions on traffic states [Wu et al., 2012].

In 2014. Gausch & Smith tried to offer algorithms for the diagnosis of incidents in signalized road networks in Dublin. For this purpose, the writer made use of four types of incident diagnosis. Three of them were based on Multi-Layer Feed Forward Neural Network (MLF), Probabilistic Neural Network (PNN), and Fuzzy-Wavelet Radial Base Function Neural Network (FWRBFNN). The fourth one (SVM) was chosen as the model of Support Vector Machine (SVM). The other new approach referred to in this research paper was the collection of data to be used in suggested algorithms, the built-up detectors at signalized intersections aimed at the intelligent control of traffic lights. This makes the diagnosis of incidents even more intricate because the collected detectors the incident-diagnosing algorithms are commonly used at 100 meters from the signalized intersections, so that the collected data are not affected by traffic lights [Gausch & Smith, 2014].

In March 2010, Susan McMillan, carried out a minute research on the performance of intelligent managing) system of Sydney Coordinated Adaptive Traffic System (SCATS) as an instrument for the diagnosis of abnormal traffic conditions. In this research, possible controlling plans that can be carried out to control accidents have been introduced as follows:

a. Recognition and diagnosis using surveillance cameras or data of detectors.

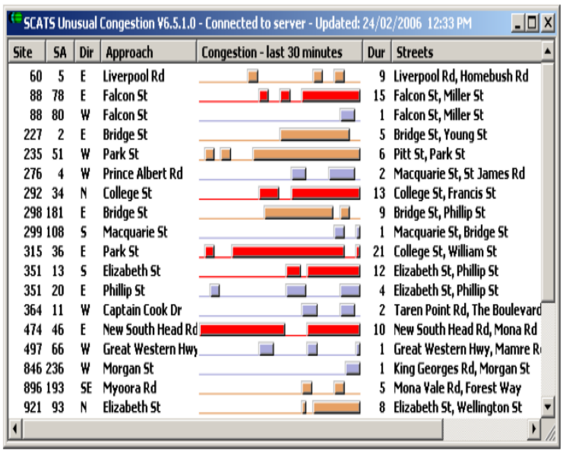
b. Passengers' information (Website or Variable Message Signs" VMS")

c. Response

d. The management of the scene and control the traffic

e. Clearance of the road and restoration of the normal condition to it [Susan McMillan, 2010].

The software of SCATS has also introduced a system called Unusual Congestion Monitor, in which the congestion occurs when the degree of saturation of detectors is high and the rate of passing flow is lower than what has been defined as acceptable. Considering the location of detectors at stop lines, it is evident that SCATS will not have the capacity of determining the queue length; however, the intended system can estimate the time of the congestion using the above-mentioned method. Figure 1 represents a sample of the performance of this system.


SCATS Unusual Congestion V6.5.1.0 - Connected to server - Updated: 24/02/2006 12:33 PM

Site	SA	Dir	Approach	Congestion - last 30 minutes	Dur	Streets
60	5	E	Liverpool Rd		9	Liverpool Rd, Homebush Rd
88	78	E	Falcon St		15	Falcon St, Miller St
88	80	W	Falcon St		1	Falcon St, Miller St
227	2	E	Bridge St		5	Bridge St, Young St
235	51	W	Park St		6	Pitt St, Park St
276	4	W	Prince Albert Rd		2	Macquarie St, St James Rd
292	34	N	College St		13	College St, Francis St
298	181	E	Bridge St		9	Bridge St, Phillip St
299	108	S	Macquarie St		1	Macquarie St, Bridge St
315	36	E	Park St		21	College St, William St
351	13	S	Elizabeth St		12	Elizabeth St, Phillip St
351	20	E	Phillip St		4	Elizabeth St, Phillip St
364	11	W	Captain Cook Dr		2	Taren Point Rd, The Boulevard
474	46	E	New South Head Rd		10	New South Head Rd, Mona Rd
497	66	W	Great Western Hwy		1	Great Western Hwy, Mamre R
846	236	W	Morgan St		1	King Georges Rd, Morgan St
896	193	SE	Myoora Rd		5	Mona Vale Rd, Forest Way
921	93	N	Elizabeth St		8	Elizabeth St, Wellington St

Figure 1. SCATS unusual congestion monitor

It must be mentioned that the diagnosis of abnormal conditions in the present research, in comparison with the diagnostic system of abnormal congestions in SCATS., has an entirely different performance because SCATS in this system attributes the abnormal congestion merely to the statistics of lower traffic obtained from the existing detectors at the intersection, as well as the high degree of saturation in the period of the time that has been studied; hence, it tries to introduce limits in the diagnosis of congestion using his method that he has already introduced. Whereas the present research does not necessarily define the abnormal condition as a congestion. The incompatibility of real conditions of the intersection with the offered data in SCATS under abnormal circumstances is defined [SCATS user manual, 2006].

According to the previous studies, it is concluded that most of the relevant studies have been conducted to detect abnormal conditions in the network of freeways and highways, and in addition, abnormal conditions are mainly defined for road traffic. Differences with previous studies and innovations of the present study are presented below. Since a major share of traffic in the urban road network is allocated to non-highway roads and urban traffic lighted intersections, and nowadays intelligent control systems play an important role in road traffic management, providing a method that can be used to detect abnormal conditions Traffic also observed the detection errors in these systems, it seems very necessary that it has not been addressed much. For this purpose, by combining the data available in the SCATS system and using the Spatial-temporal data of users' mobile phones (based on GPS and taken from the Neshan application) and using the fundamental parameters of traffic, it is tried to reach the goals stated above.

3. Methodology

This section presents the methodological framework of detecting the anomalies in intersections using SCATS data. Based on Figure 2 the framework consists of the following three major steps: data preprocessing, determining critical values and anomaly detection.

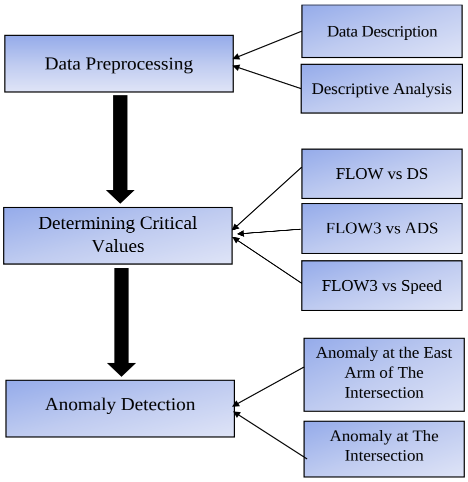


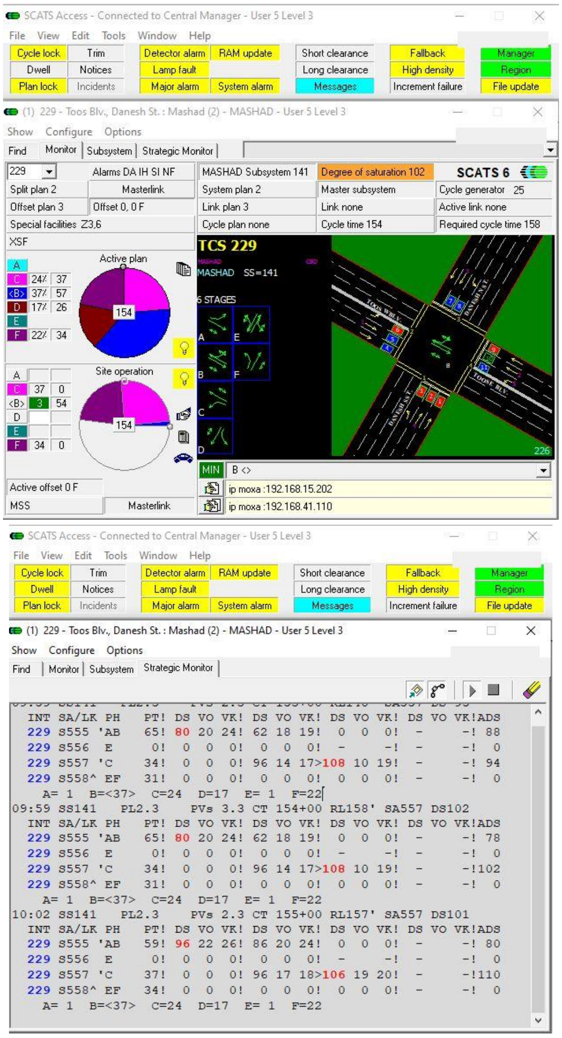
Figure 2. The methodological framework for traffic anomalies detection by SCATS data

3.1. Data Preprocessing

Before studying the performance of data and building models, dataset should be understood and prepared, which is done and explained in this section.

3.1.1. Data Description

Here, after describing the dataset and SCATS traffic control system, the short description of the variables consists of the definitions and their calculation formula is given. As it can be seen in Figure 3, the dataset applied for training and building the model consists of the main traffic information collected by SCATS[4] system, which is illustrated below and the 10-minute averaged speeds extracted from Neshan[5] application.


SCATS Access - Connected to Central Manager - User 5 Level 3

(1) 229 - Toos Blv., Danesh St. : Mashad (2) - MASHAD - User 5 Level 3

229 | Alarms DA IH SI NF | MASHAD Subsystem 141 | Degree of saturation 102 | SCATS 6
Split plan 2 | Masterlink | System plan 2 | Master subsystem | Cycle generator 25
Offset plan 3 | Offset 0, 0 F | Link plan 3 | Link none | Active link none
Special facilities Z3,6 | Cycle plan none | Cycle time 154 | Required cycle time 158
XSF

Active plan
A
C 24% 37
<B> 37% 57
D 17% 26
E
F 22% 34
154

Site operation
A
C 37 0
<B> 3 54
D
E
F 34 0
154

TCS 229
MASHAD SS=141
6 STAGES

Active offset 0 F
MSS | Masterlink
MIN | B <>
ip moxa :192.168.15.202
ip moxa :192.168.41.110

SCATS Access - Connected to Central Manager - User 5 Level 3

(1) 229 - Toos Blv., Danesh St. : Mashad (2) - MASHAD - User 5 Level 3

```
  INT SA/LK PH    PT! DS VO VK! DS VO VK! DS VO VK! DS VO VK!ADS
  229 S555 'AB    65! 80 20 24! 62 18 19!  0  0  0!  -       -! 88
  229 S556  E      0!  0  0  0!  0  0  0!  -       -!  -       -!  0
  229 S557 'C     34!  0  0  0! 96 14 17>108 10 19!  -       -! 94
  229 S558^ EF    31!  0  0  0!  0  0  0!  0  0  0!  -       -!  0
    A= 1  B=<37>  C=24  D=17  E= 1  F=22
09:59 SS141   PL2.3    PVs 3.3 CT 154+00 RL158' SA557 DS102
  INT SA/LK PH    PT! DS VO VK! DS VO VK! DS VO VK! DS VO VK!ADS
  229 S555 'AB    65! 80 20 24! 62 18 19!  0  0  0!  -       -! 78
  229 S556  E      0!  0  0  0!  0  0  0!  -       -!  -       -!  0
  229 S557 'C     34!  0  0  0! 96 14 17>108 10 19!  -       -!102
  229 S558^ EF    31!  0  0  0!  0  0  0!  0  0  0!  -       -!  0
    A= 1  B=<37>  C=24  D=17  E= 1  F=22
10:02 SS141   PL2.3    PVs 2.3 CT 155+00 RL157' SA557 DS101
  INT SA/LK PH    PT! DS VO VK! DS VO VK! DS VO VK! DS VO VK!ADS
  229 S555 'AB    59! 96 22 26! 86 20 24!  0  0  0!  -       -! 80
  229 S556  E      0!  0  0  0!  0  0  0!  -       -!  -       -!  0
  229 S557 'C     37!  0  0  0! 96 17 18>106 19 20!  -       -!110
  229 S558^ EF    34!  0  0  0!  0  0  0!  0  0  0!  -       -!  0
    A= 1  B=<37>  C=24  D=17  E= 1  F=22
```

Figure 3. The picture of Toos-Danesh intersection data in SCATS system

After merging two above-mentioned files, the time

---

[4] . Sydney Coordinated Adaptive Traffic System

[5] . Neshan is one of the main Iranian navigations softwares, which works on smartphones that have GPS support.

duration of data is shortened to the duration between August 15th and August 20th, 2020. Also, the data is related to the intersection named Toos-Danesh with the SCATS code of 269 of Toos motorway of Mashhad. Since Neshan mobile data is generated 10-minutely, for the combination of two datasets, by aggregation, the data frequency of SCATS changes to 10-minute. Also, the SCATS data is joint to six segments before Toos-Danesh intersection with the length of 610.29 meters. The variables given in Table 1, which are recorded by SCATS traffic control system, are satisfied with the fundamental parameters in traffic engineering.

Table 1. Variables Description

Variables	Description
DS	Degree of saturation of the intersection
VO	Original volume of the intersection
Phase Time	The time duration of green light of the intersection
FLOW	Hourly flow rate in green time; $FLOW = \frac{3600}{Phase\ Time} \times VO$
ADS	Average degree of saturation collected by sensors of an east arm of the intersection, averaged in 10-minute period
VO3	Original volume in phase time of the east arm of the intersection, summed in 10-minute period
Phase Time3	The time duration of green light of the east arm of the intersection, summed in 10-minute period
FLOW3	Hourly flow rate based on VO3; $FLOW3 = \frac{3600}{Phase\ Time3} \times VO3$
Speed	Average of 10-minute averaged speeds of segments before intersections

To confirm, in the following sections, the index 3 refers to the east arm of the intersection. Also, note that the anomaly detection will be done for both the entire intersection and the east approach of the SCATS system.

3.1.2. Descriptive Analysis

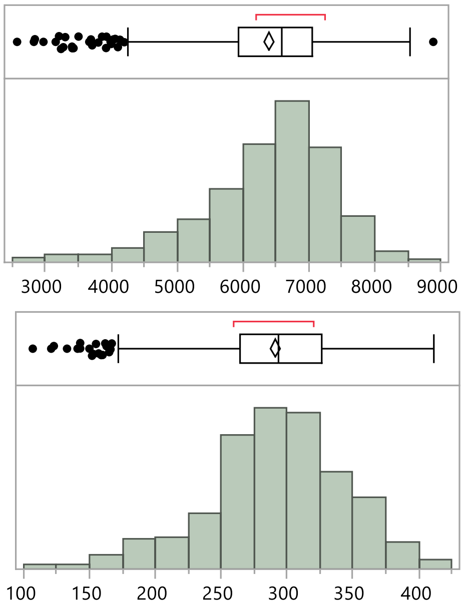
In Table 2, the descriptive analysis of the variables is given. Quantiles and percentiles of variables will be applied later for calculating their critical values.

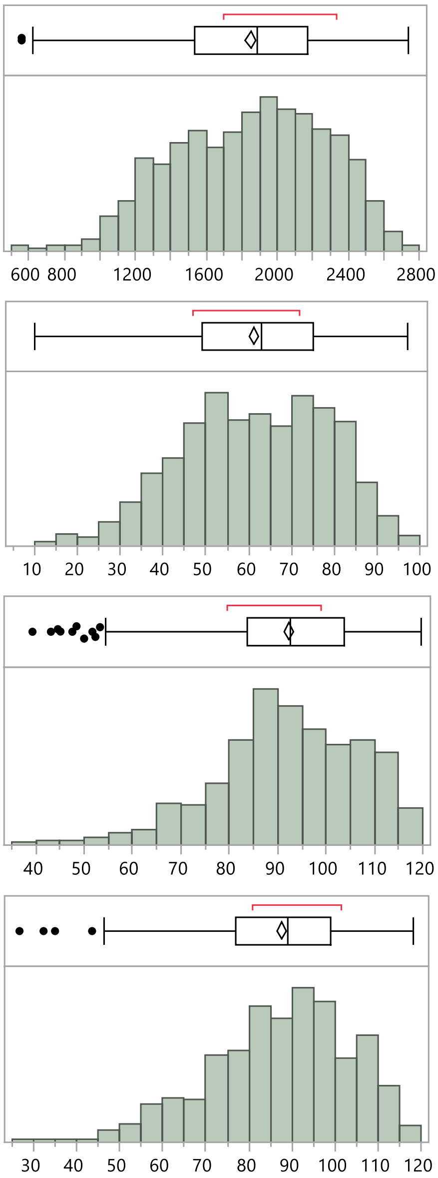
Table 2. Descriptive analysis of the variables

Variable	Min	Mean	Max	Std Dev
DS	39.3	92.3	119.6	14.6
VO	107	291.5	358	51.8
FLOW	2574.2	6396.2	8886.5	981.0
ADS	26.8	87.5	118	15.8
VO3	10	61.2	97	16.9
FLOW3	559	1850.7	2739.7	421.2
Speed	14	37.1	72	8.66

The variables distributions are visualized by the combination of histogram and boxplots in Figure 4, which are from top to bottom VO, FLOW, VO3, FLOW3, ADS, DS and Speed.

As it is seen, there are outliers for most of the variables, which can overlap anomalies. For example, one outlier of the variable Flow is related to its maximum value, 8887. Also, Flow3 has one outlier, which is its minimum one. Obviously, the distributions of DS and Flow are equal. Most of the plots except speed, which is totally left skewed, are right skewed. The frequency of low speeds is high compared to other values.





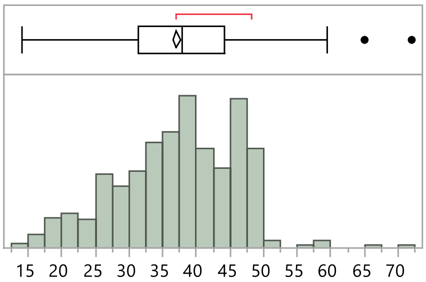


Figure 4. Boxplots of the variables; VO, FLOW, VO3, FLOW3, ADS, DS and SPEED

4. Determining Critical Values

In this section, it is tried to calculate critical values based on the traffic fundamental parameters. To achieve this goal, linear models (quadratic, cubic and the 4th degree of polynomial) are fitted on FLOW by DS, FLOW3 by ADS and FLOW3 versus speed.

4.1. FLOW vs DS

Linear models are fitted on FLOW by DS (Table 3). The best model is selected based on the least error and the most R square. As it is clear in the figure 5, three linear models, polynomial models of degree 2, 3 and 4 are fitted on the points of the scatter plot between DS and FLOW.

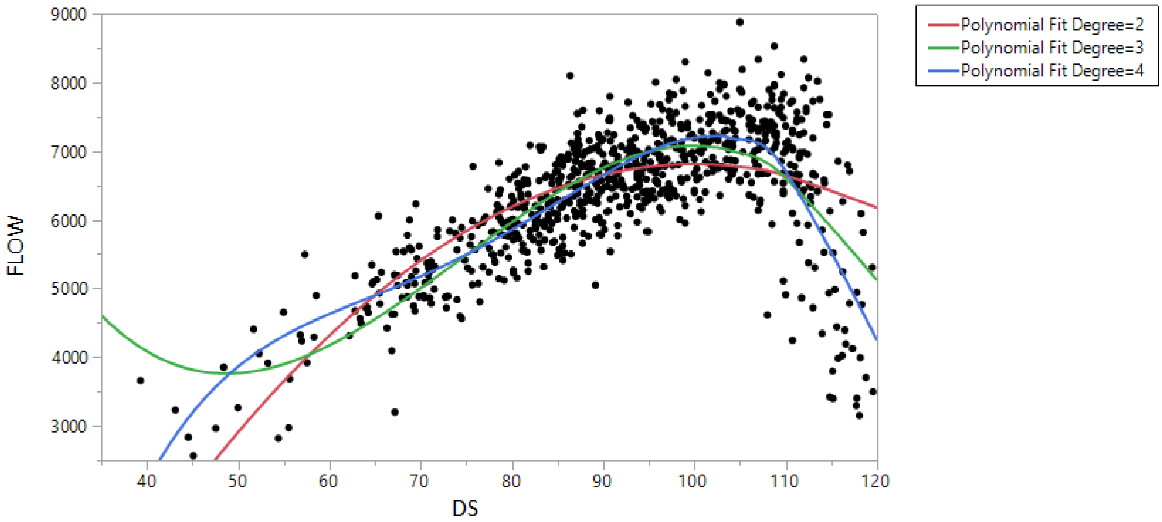


Figure 5. Fitted models on the scatter plot of FLOW (Veh/h) and DS

Table 3. Summary of Fits

Model	R^2	$Adj\ R^2$	RMSE	Parameter Estimates					
				Intercept	DS	$(DS-\mu)^2$	$(DS-\mu)^3$	$(DS-\mu)^4$
Polynomial with degree 2	0.486	0.485	703.97	4528.45	23.82	-1.57		
Polynomial with degree 3	0.579	0.578	637.64	2373.65	48.92	-2.69	-0.05	
Polynomial with degree 4	0.616	0.614	609.77	416.48	69.45	-1.79	-0.11	-0.001

The best fitted model based on the least Root Mean Square Error is polynomial fit with degree 4. Its equation equals to:

$$\mathrm{Flow} = 416.48 + 69.45\mathrm{DS} - 1.79(\mathrm{DS} - 92.31)^2 - 0.11(\mathrm{DS} - 92.31)^3 - 0.001(\mathrm{DS} - 92.31)^4 \quad (1)$$

Note that, increasing the degree of a model increases the model complexity and the overfitting probability. So, the polynomial models with degrees more than 4 with negligible differences are not consider in model selection process.

According to Figure 5, the maximum of Flow is 8887, for which the corresponding DS equals to 105. Furthermore, using the “Optimize” function of R software, the maximum point (peak) of the fitted polynomial is calculated. The maximum point for Flow equals to 7223.41, for which the DS is obtained as 102.01. In the appendix 1, the corresponding written code, named as “Code 1”, is given.

As it is obvious, the corresponding value for the maximum point of Flow is close to the one for the peak of the optimal fitted polynomial. Also, it is expected that no anomaly happens in the neighborhood of the maximum value.

4.2. FLOW3 vs ADS

Here, like the previous section, fitting model on FLOW3 by ADS is discussed and the maximum values are calculated (Figure 6, Table 4). Polynomial with degree 4 which is blue color has the highest value of R^2 and provides a better fit of the considered data. The RMSE in this model is also the lowest.

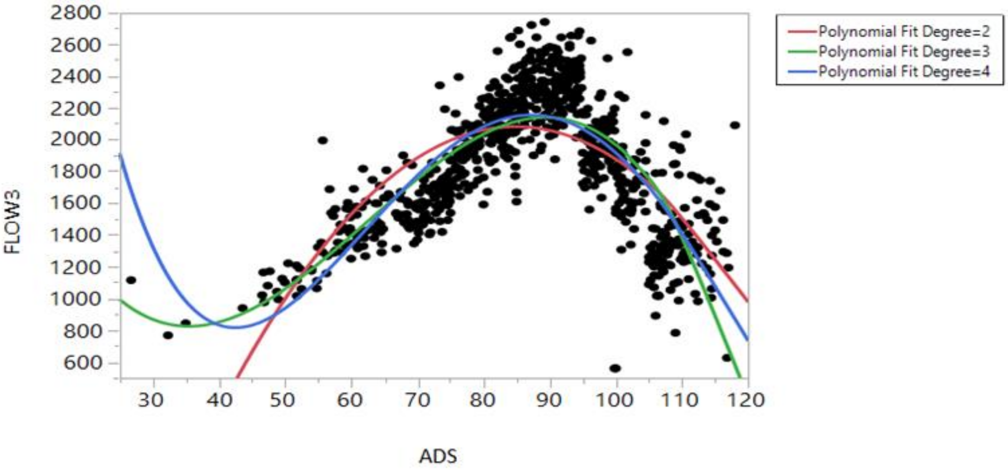


Figure 6. Fitted models on the scatter plot of FLOW3 (Veh/h) and ADS

Table 4. Summary of Fits

Model	R^2	$Adj\ R^2$	RMSE	Parameter Estimates				
				Intercept	ADS	$(ADS-\mu)^2$	$(ADS-\mu)^3$	$(ADS-\mu)^4$
Polynomial with degree 2	0.498	0.497	298.62	2486.72	-4.71	-0.89		
Polynomial with degree 3	0.600	0.599	266.80	1701.17	4.96	-1.26	-0.02	
Polynomial with degree 4	0.617	0.615	261.17	2207.90	-0.61	-1.46	-0.005	0.0003

According to the Root Mean Square Error, the best fitted model is polynomial fit with degree 4. Its equation equals to:

$$\begin{aligned}\mathrm{Flow3} = 2207.90 - 0.61\mathrm{ADS3} \\ - 1.46(\mathrm{ADS}-87.47)^2 \\ - 0.005(\mathrm{ADS}-87.47)^3 \\ + 0.0003(\mathrm{ADS}-87.47)^4\end{aligned} \quad (2)$$

According to Figure 3, the peak of Flow3 is 2739.7, for which the corresponding value of ADS is 89.3. Also, the maximum value of Flow3 is obtained as 2154.61 by optimize function in R, whose corresponding ADS is 87.26. This code is given in the appendix as “Code 2”.

Obviously, the value of ADS for the maximum point of Flow is almost equal to ADS value for peak of the optimal fitted polynomial.

4.3. FLOW3 vs Speed

Based on Figure 7 and Table 5, polynomial fit is done on FLOW3 by Speed (average 10-minutes speed). Polynomial with degree 4 which is blue color has the highest value of R^2 and provides a better fit of the considered data. The constant value of this model is also higher than the others.

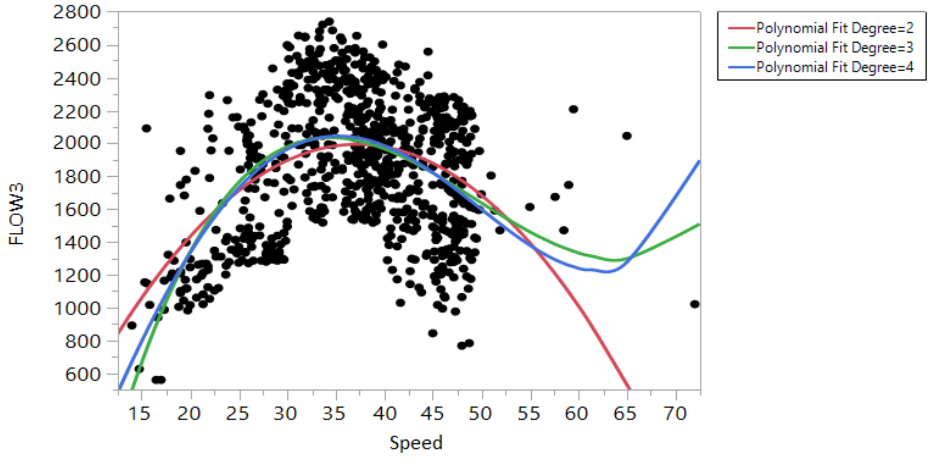


Figure 7. Fitted models on the scatter plot of FLOW3 (Veh/h) and Speed (Km/h)

The Figure shows three polynomial fits of different degrees on data.

Table 5. Summary of Fits

Model	R^2	$Adj\ R^2$	RMSE	Parameter Estimates				
				Intercept	ADS	$(ADS-\mu)^2$	$(ADS-\mu)^3$	$(ADS-\mu)^4$
Polynomial with degree 2	0.215	0.213	373.71	1975.38	0.46	-1.89		
Polynomial with degree 3	0.270	0.267	360.58	2468.27	-12.18	-2.07	0.057	
Polynomial with degree 4	0.274	0.270	359.85	2398.45	-9.82	-2.56	0.041	0.001

According to the least RMSE and the most R Square, the best fitted model is polynomial fit with degree 4. Its equation equals to:

$$\begin{aligned}\mathrm{Flow3} = 2398.45 - 9.82\mathrm{Speed} \\ - 2.56(\mathrm{Speed} - 37.09)^2 \\ + 0.041(\mathrm{Speed} - 37.09)^3 \\ + 0.001(\mathrm{Speed} - 37.09)^4\end{aligned} \quad (3)$$

The maximum value for Flow3 equals to 2739.7, whose speed is 34.3, while the maximum values for Flow3 via optimize function in R equals to 2043.38, whose speed is 35.25. As it is clear the corresponding ADS values are close.

5. Anomaly Detection

Here, both anomaly in the SCATS decisions and anomaly in the traffic situation of the intersection are considered as anomaly. Also, both anomalies at the intersection and at the east arm of the intersection are detected and studied under two subsections. Criteria or threshold of anomaly detection is defined according to the quantiles (quartiles and percentiles) of speed, ADS and DS.

5.1. Anomaly at the East Arm of the Intersection

It is expected that at the very low speed values, SCATS shows high values for ADS of the approach. On the other side, for the very high-speed values, the low values of ADS is expected. Totally, for the east arm of the intersection (the approach), obviously, if there is a time with too small speed and too small ADS or with too large values for both, it is defined as an abnormal condition. In this section, based on quantiles (quartiles and percentiles) values, given in Table 6, it is defined that how small or how large the measures should be.

Table 6. Quantiles of the variables

Variable	10th percentile	1st Qu.	Median	3rd Qu	90th percentile
DS	71.9	83.6	92.8	103.6	111.2
VO	224	264	294	358.3	326
FLOW	5063.9	5927.0	6576.6	7047.7	7466.0
ADS	64.9	77	89.1	99	107.3
VO3	38	49	63	75	83
FLOW3	1284.4	1535.0	1885.5	2172.5	2392.4
Speed	25.2	31.3	37.9	44.2	47.5

It should be noted that the results of the quartiles could be used to abnormal conditions detection, however, for higher accuracy 10th and 90th quantiles were used. Both will be reported in this section.

5.1.1. Based on 10th and 90th Percentiles

The 10th and 90th percentiles of ADS are 64.9 and 107.3 respectively. Based on the equation (2), the corresponding Flow3 values are calculated as shown in Appendix 2.

In Figure 8 the percentiles bounds are detected by colored lines, which divides the scatter plots of ADS and Flow3 into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2 and 5, and the least in zones 1 and 3. Therefore, ADS=107.3 and Flow3=1575.73 occur in zone 2.

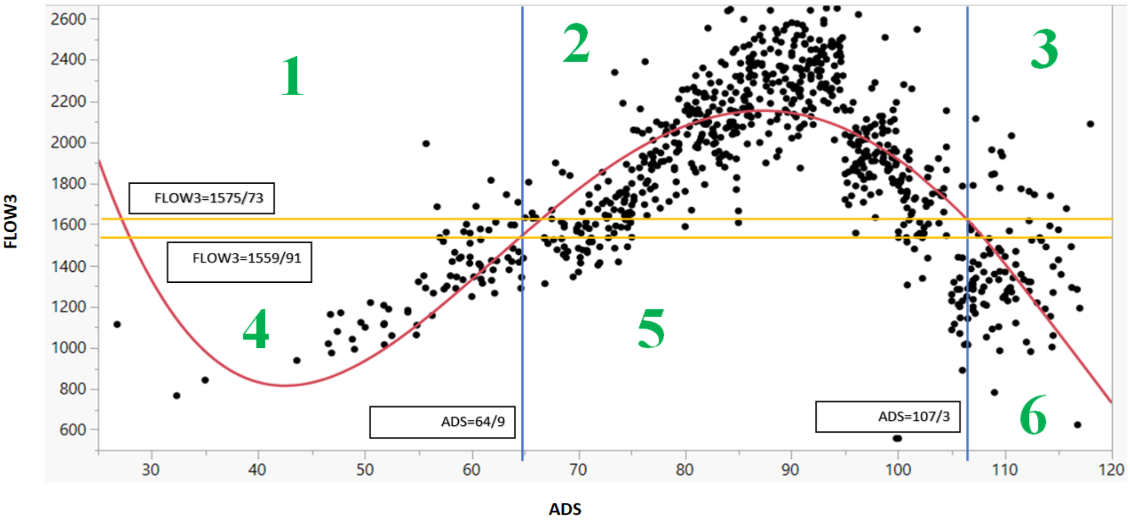


Figure 8. The scatter plot of FLOW3 (Veh/h) versus ADS with 6 zones indicated by percentiles bounds

Similarly, the 10th and 90th percentiles of Speed are 25.2 and 47.5 respectively. Based on the equation (3), the corresponding Flow3 values for Speed percentiles are calculated in Appendix 2. In Figure 9 the percentiles bounds are detected by colored lines, which divides the scatter plots of Speed and Flow3 into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2 and 5, and the least in zones 6 and 3. Therefore, ADS=47.5 and Flow3=1740.14 occur in zone 2.

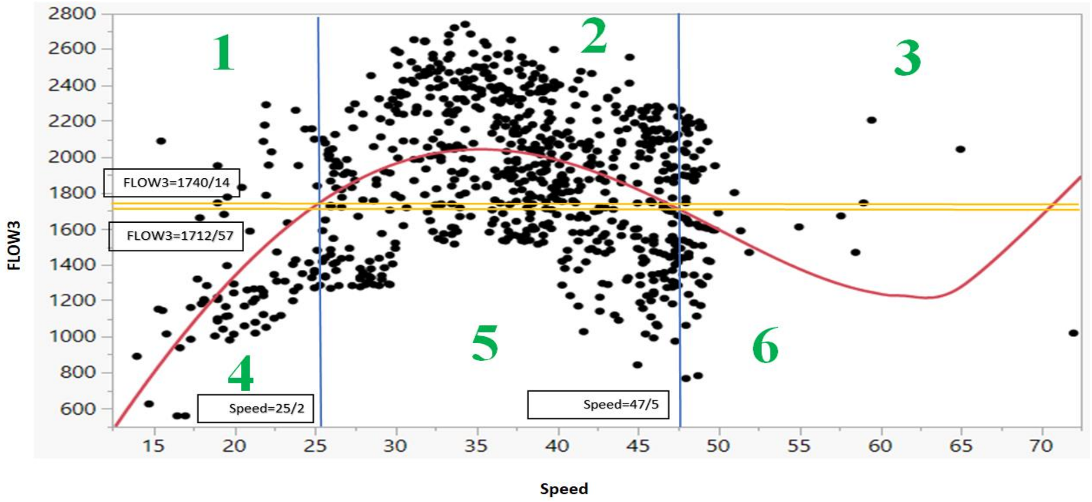


Figure 9. The scatter plot of FLOW3 (Veh/h) versus Speed (Km/h) with 6 zones indicated by percentiles bounds

According to the traffic fundamental relations, the points inside the zones number 1 and 4 at both graphs, and inside the zones number 3 and 6 at both graphs, Figures 8 and 9, are indicated as anomalies. Detected anomalies are shown by colored points in Figures 10. It can be seen that the majority of these colored points are in low Flow3 values. Therefore, this anomalies is not seen in high values of Flow3. Also, colored points generally occur in lower amounts of speed and ADS. This category indicates that anomalies do not occur in high values of Flow3. Most of these cases are associated with low Flow3.

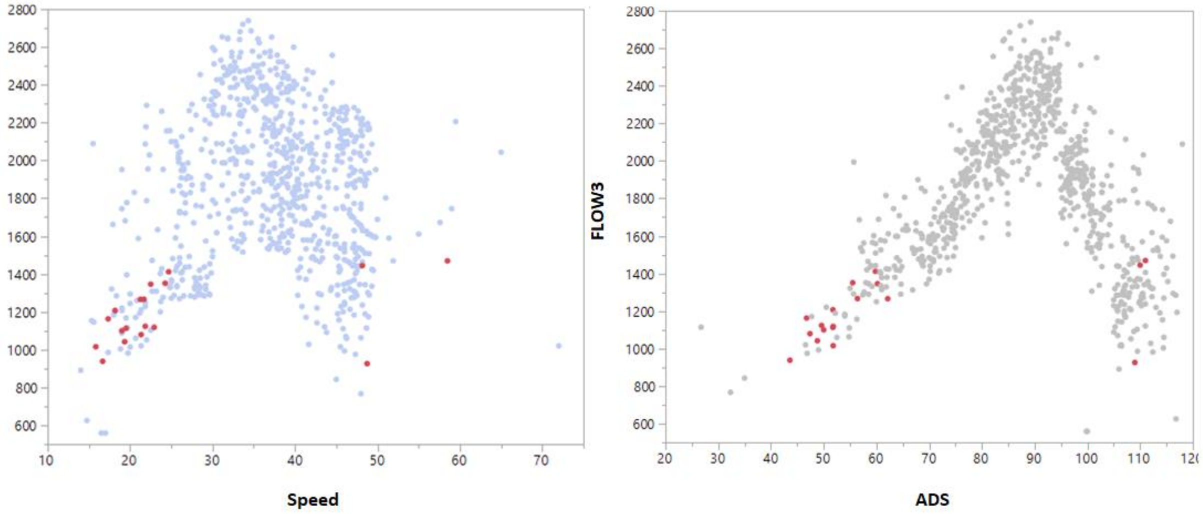


Figure 10. East Arm anomalies by 10th and 90th percentiles colored in scatter plots of FLOW3 (Veh/h) and ADS and Speed (Km/h)

For making comparison with a real data set, the date and the exact time of 18 detected anomaly points of the approach (the east arm of the intersection) by 10th and 90th percentiles are given in Figure 11. Most of the anomalies occurred on 2020-08-16 and 2020-08-22.

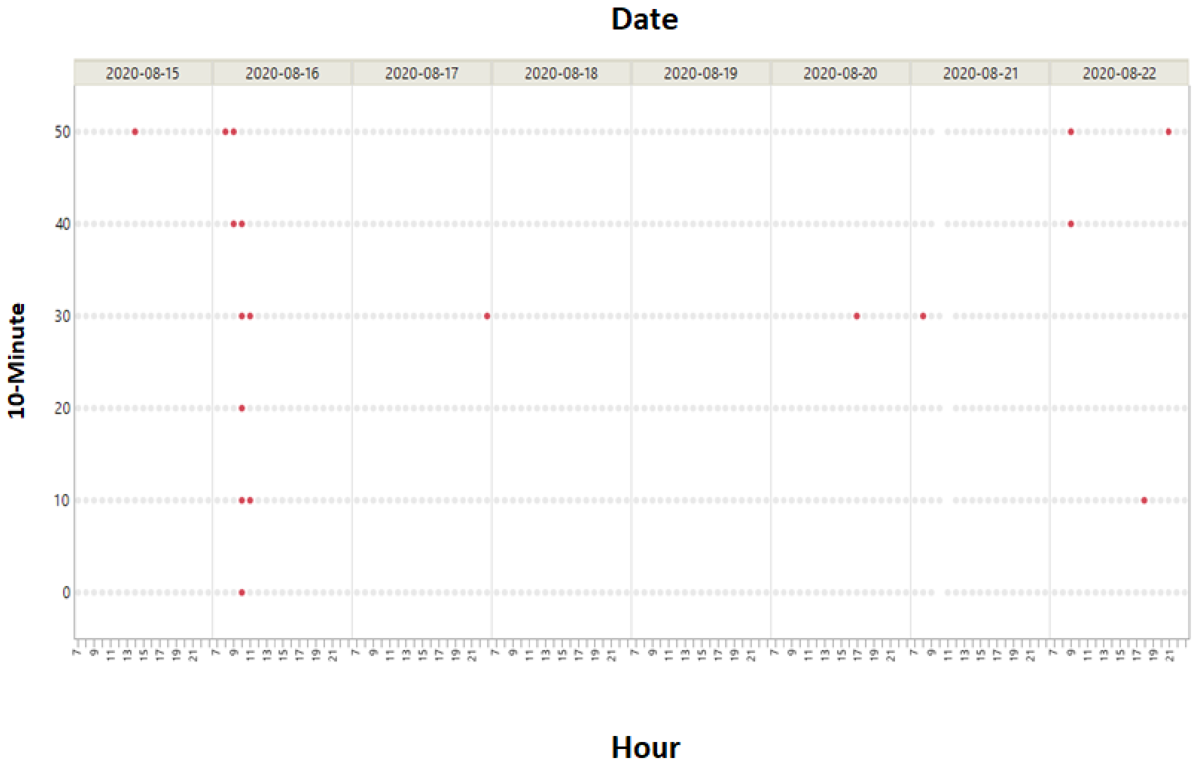


Figure 11. Time of East arm anomalies by 10th and 90th percentiles

5.1.2. Based on 1th and 3th Quartiles

The 1th and 3th quartiles of ADS are 77 and 99 respectively. Based on Equation 2, the corresponding Flow3 values are calculated in Appendix 2.

In the following scatter plot (Figure 12), the quartiles bounds are detected by colored lines, which divides the scatter plots of ADS and Flow3 into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2 and 4, and the least in zones 1 and 3. Therefore, ADS=99 and Flow3=2010.23 occur in zone 2.

That is, the accumulation of data is observed in the range of ADS =77 to ADS =99 and Flow3 more than 2010.23.

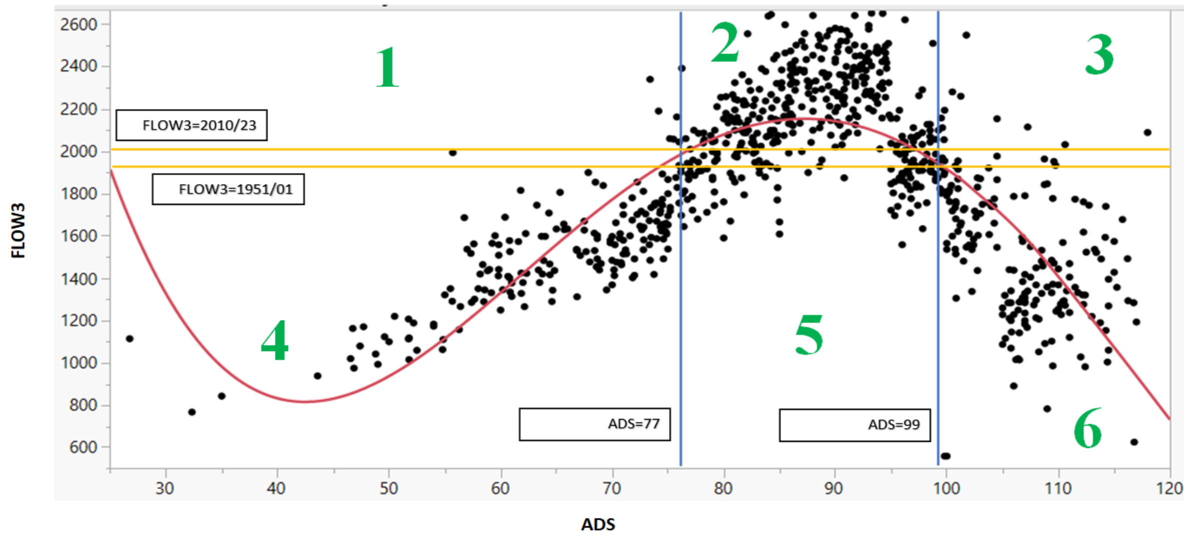


Figure 12. The scatter plot of FLOW3 (Veh/h) and ADS with 6 zones indicated by quartiles bounds

Similarly, the 1th and 3th quartiles of Speed are 31.3 and 44.2 respectively. Based on Equation 3, corresponding Flow3 values for speed quartiles are calculated in Appendix 2.

In Figure 13, the quartiles bounds are detected by colored lines, which divides the scatter plots of Speed and Flow3 into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2, 4, 5, and the least in zone 1. Therefore, Speed=31.3 and Flow3=1998.43 occur in zone 2. That is, the accumulation of data is observed in the range of Speed =31.3 to Speed =44.2 and Flow3 more than 1998.43.

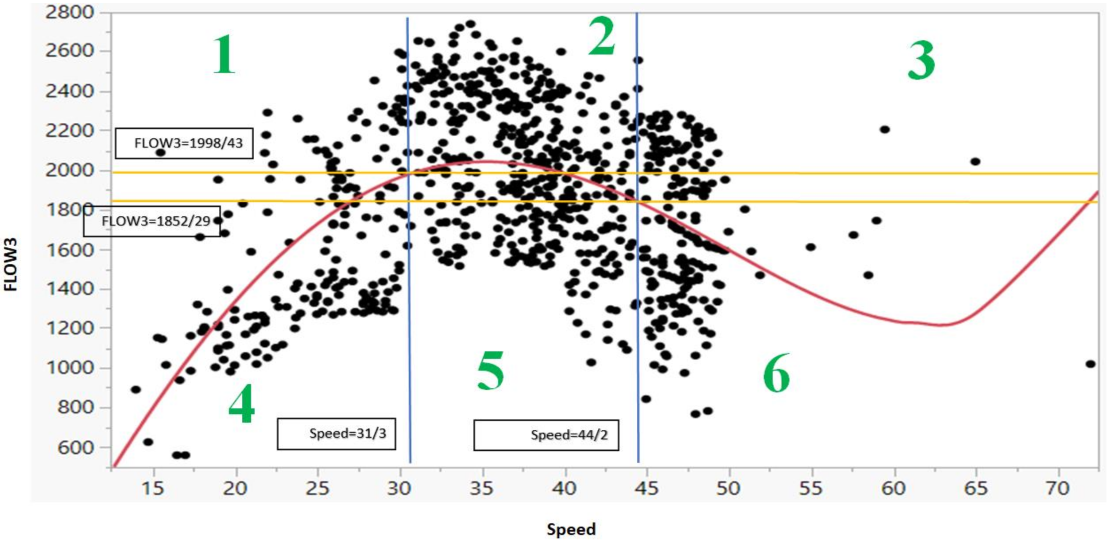


Figure 13. The scatter plot of FLOW3 (Veh/h) and Speed (Km/h) with 6 zones indicated by quartiles bounds

The points inside the zones number 1 and 4 at both graphs, and inside the zones number 3 and 6 at both graphs, Figures 14 and 15, are indicated as anomalies. Anomalies are shown by colored points in the following scatter plots. It can be seen that the majority of these colored points are in low Flow3 values. Therefore, this anomalies is not seen in high values of Flow3. Also, colored points generally not occur in medium amounts of speed and ADS. According to Figure 15, the data are widely distributed in the desired hours and days. At speeds less than 30 km/h and more than 45 km/h, as well as at ADS less than 75 and more than 100, we see this anomalies. This means that as the speed increases, the ADS values are reported to be lower.

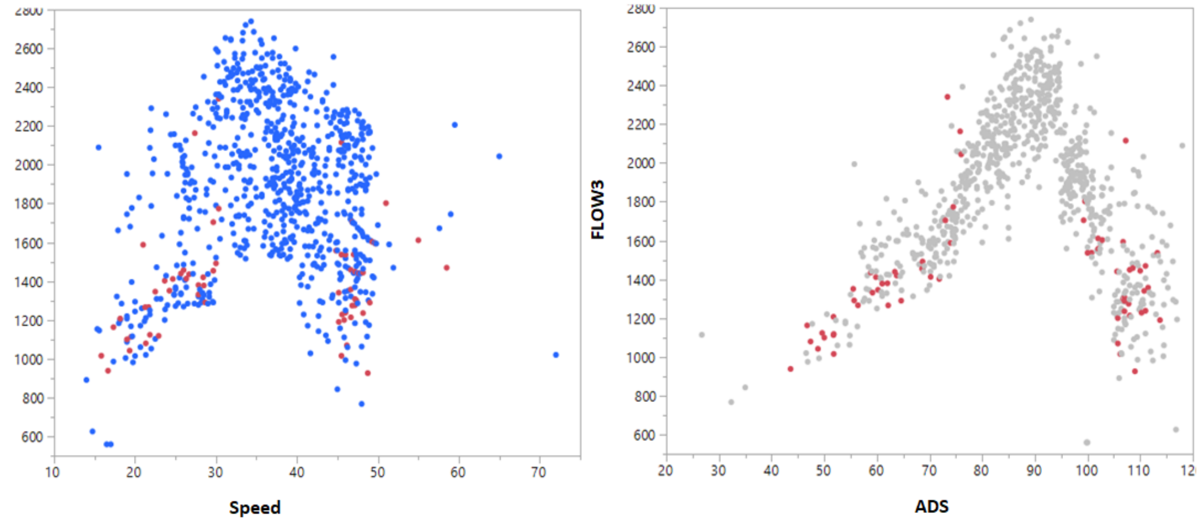


Figure 14. East arm anomalies by 1th and 3th quartiles colored in scatter plots of FLOW3 (Veh/h) and ADS and Speed (Km/h)

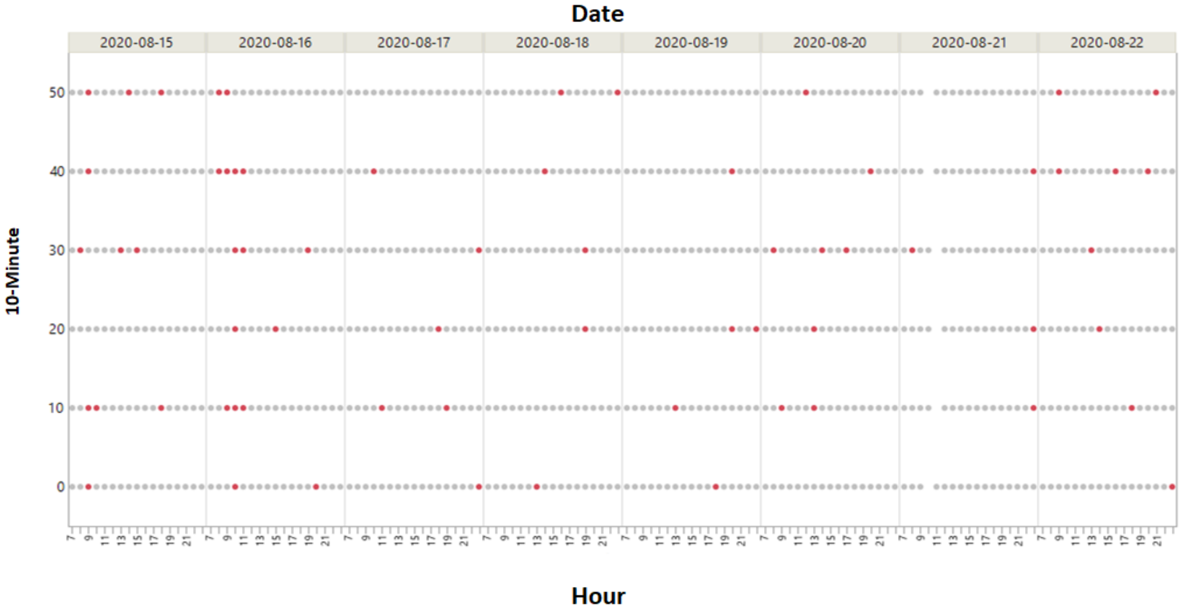


Figure 15. Time of East arm anomalies by 1th and 3th quartiles

5.2. Anomaly at the Intersection

In this section for anomaly detection of the intersection (Fraud decision of SCATS system), only the parameters Flow and DS were used to draw the fundamental Traffic diagrams. The reason is drawing the speed diagram for an intersection does not make sense.

5.2.1. Based on 10th and 90th Percentiles

The 10th and 90th percentiles of DS are 71.9 and 111.2 respectively. Based on Equation 1, the corresponding Flow values are calculated in Appendix 2.

In the following scatter plot (Figure 16), the percentiles bounds are detected by colored lines, which divides the scatter plots of DS and Flow into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2, 5, and the least in zone 1. Therefore, DS=111.2 and Flow3=6631.8 occur in zone 2. That is, the accumulation of data is observed in the range of DS =71.9 to DS =111.2 and Flow3 more than 6631.8.

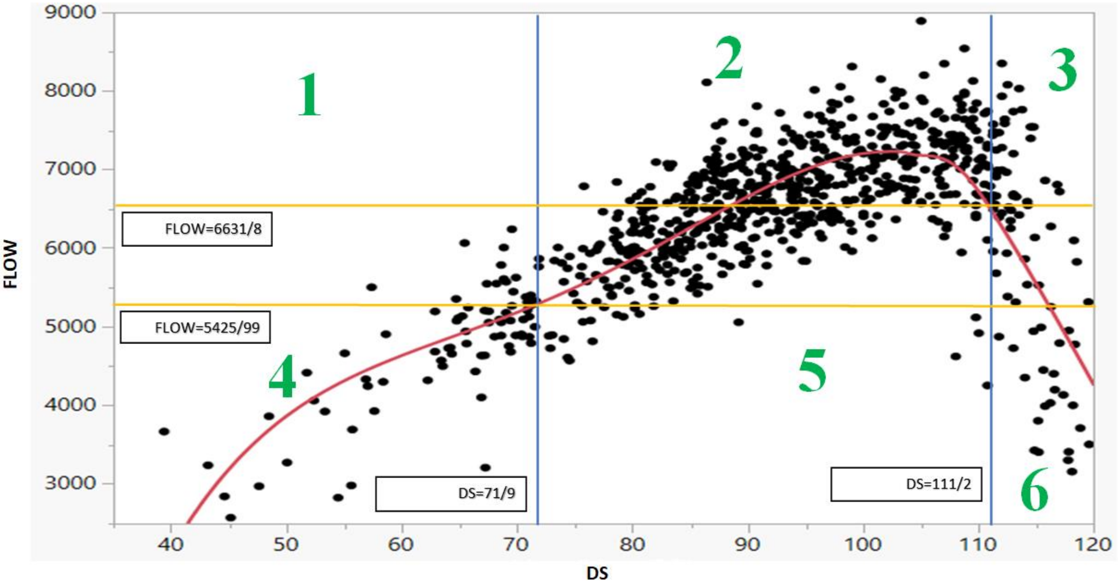


Figure 16. The scatter plot of FLOW (Veh/h) and DS with 6 zones indicated by percentile bounds

Thus, if a point inside the zones number 1 and 4 at Figure 16 is also inside the zones number 3 and 6 at Figures 8, i.e., common points of the zones 1 and 4 of Figure 16 and the zones 3 and 6 of Figure 8, then they can be consider as anomaly of SCATS system. In the other word, even though an intersection approach showed very high ADS values (the points in the zones number 3 and 6 of Figure 8), that approach was not taken into account and SCATS presents low values for DS (the points in the zones number 1 and 4 of Figure 16) of the intersection. The other way around of this fact is not necessarily true, because the DS values of SCATS may be calculated based on other arms of intersection. Detected anomalies are shown by colored points in Figures 17. As it can be seen, this type of anomalies occurs in very rare cases, which we see only one case.

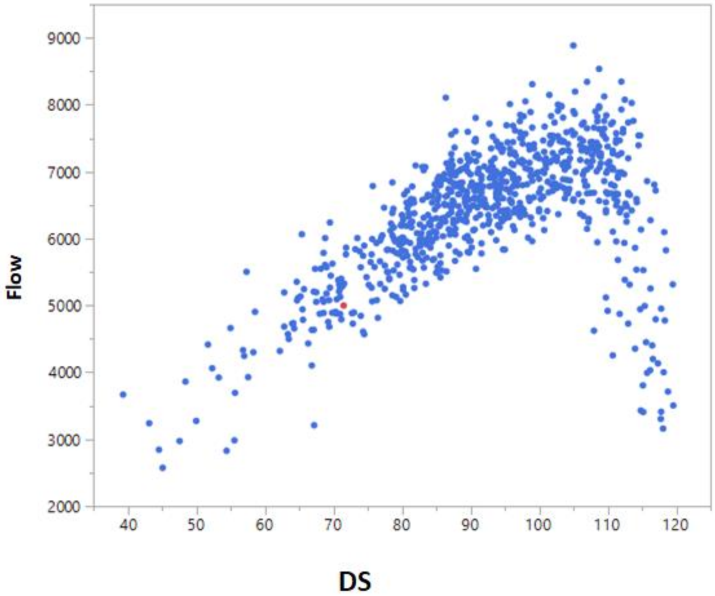


Figure 17. Intersection anomalies by 10th and 90th percentiles colored in scatter plots of FLOW (Veh/h) and DS

The time of 7:20 of 2020/08/15 is indicated as anomaly and is shown by colored point in the following scatter plot.

5.2.2. Based on 1th and 3th Quartiles

The 1th and 3th Quartiles of DS are 83.6 and 103.6 respectively. The corresponding Flow values are calculated in Appendix 2.

In the following scatter plot (Figure 18), the quartiles bounds are detected by colored lines, which divides the scatter plots of DS and Flow into 6 parts or 6 zones. As can be seen, the majority of points are seen in zones 2, 4, 5, and the least in zone 1. Therefore, DS=103.6 and Flow3=6153.63 occur in zone 2. That is, the accumulation of data is observed in the range of DS =83.6 to DS =103.6 and Flow3 more than 6153.63.

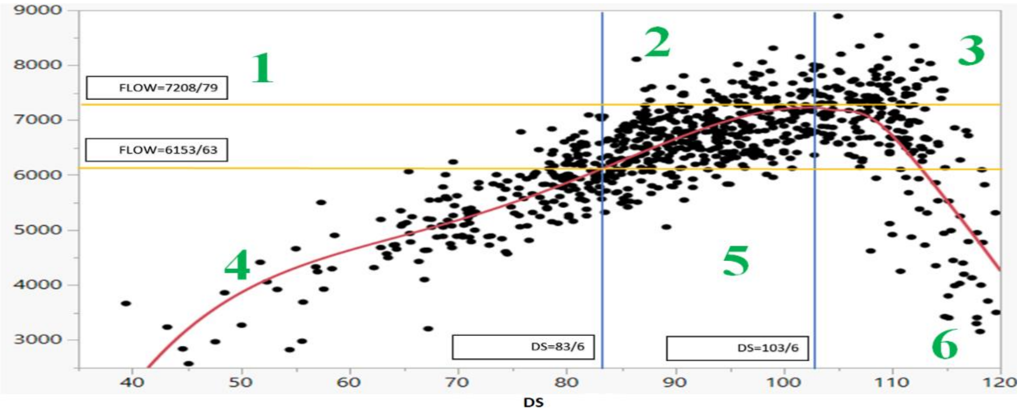


Figure 18. The scatter plot of FLOW (Veh/h) and DS with 6 zones indicated by quartile bounds

As above mentioned, if a point inside the zones number 1 and 4 at Figure 18 is inside the zones number 3 and 6 at Figures 12, it is anomaly. The Anomalies are shown by colored point in the following scatter plot (Figure 19, 20). Also, colored points generally occur in average amounts of DS. And their number is very limited. This type of anomalies occurs in DS values between 65 and 83. Therefore, by managing this range, this defect can be reduced.

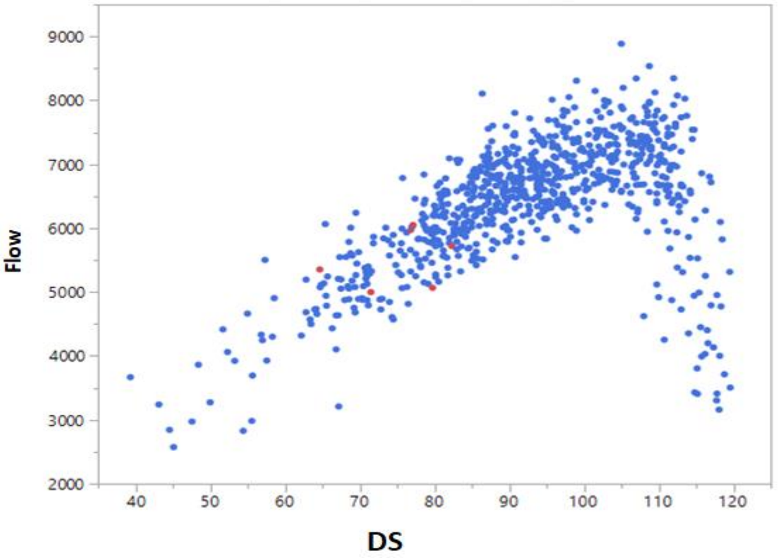


Figure 19. Intersection anomalies by 1st and 3th quartiles colored in scatter plots of FLOW (Veh/h) and DS

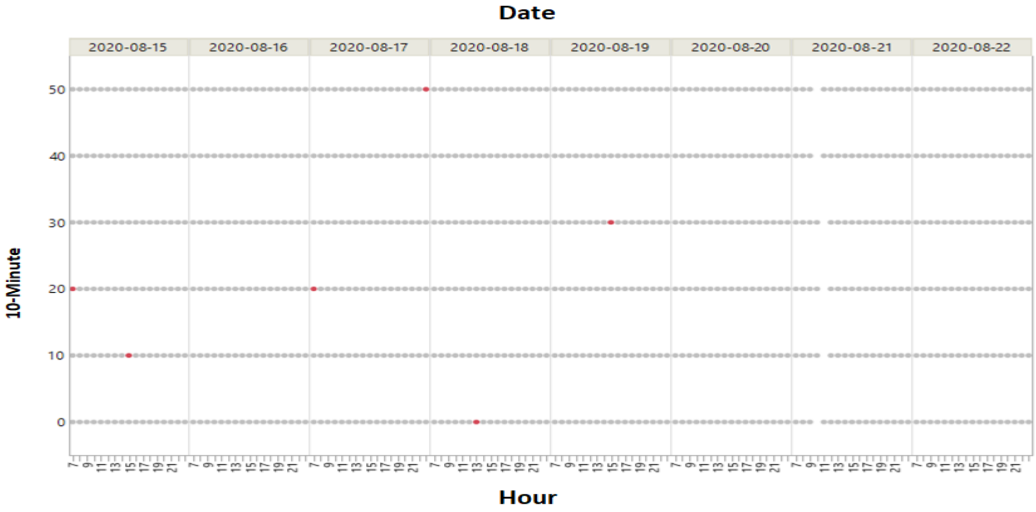


Figure 20. Time of intersection anomalies by 1st and 3rd quartiles

5.3. Method Evaluation

As it is seen in SCATS LOG, at the given times in Figure 11 (10th and 90th percentiles), which contains the 18 anomaly points of the arm, at all the detected times, intersection was manually controlled by an operator, it means that it was out of smart mode. On the other side, for the highlighted points in Figure 15 (1st and 3th quartiles), which contains the 65 anomaly points of the arm, at 37 times, around 57 percent of the detected points, the decisions were made by an operator. As the same way, for the intersection, based on the percentiles, the time on Figure 14 was indicated correctly. However, based on the quartiles, 80 percent or 4 points of 5 times on Figure 20 were detected correctly.

Also, as the results of Mashhad City Traffic Event Registration System shows, two congestions were detected correctly at the mentioned times in 100 meters from the intersection

At the end a comparison between previous studies and the current research is done and presented in table 7.

Table 7. Comparison table of a number of previous studies with the current study

Author Name, Year	Goodness of fit R^2	anomaly distribution	Type of anomaly	Model type
Ken Michael et al, 2017	0.91	Uniform	Traffic situation of density and flow	Data Mining-Black box
Polson and Sokolov, 2017	0.93	Scattered	Flow and congestion	Deep learning- Black box
Chen et al., 2019	0.78	Discrete	Traffic prediction under abnormal conditions, such as accidents, adverse weather, work zones.	multi-step-ahead prediction model support vector regression GBRT model
Chen et al., 2020	-	Separated	Traffic flow, service quality and speed	Ensembling-MRBF-LSTM framework
This Study	0.616	Separated	Flow-speed- saturation	Polynomial regression-White box

6. Conclusion

In this paper, SCATS features such as degree of saturation (DS), original volume (VO), hourly flow rate in green time (FLOW) were studied to detect their critical values. Also, adding 10-minute averaged speed time series data of vehicles passing the segments of Toos motorway in Mashhad, extracted from Neshan application to SCATS data helped to detect anomalies at the indicated arm of the intersection, the East arm, which leads to anomaly detection of the entire intersection. Using quantiles of DS, ADS and speed, the scatter plots of FLOW versus these variables are divided to six zones. Points located in two zones are defined as abnormal situations. Matching the scatter plot of Flow3 and ADS with the one of Flow3 and Speed indicated as anomalies. Results indicates that anomalies do not occur in high values of Flow3. Most of these cases are associated with low Flow3. At speeds less than 30 km/h and more than 45 km/h, as well as at ADS less than 75 and more than 100, we see this anomalies. This means that as the speed increases, the ADS values are reported to be lower. Intersection type of anomalies occurs in very rare cases, which we see only one case. Some anomalies occurs in DS values between 65 and 83. Therefore, by managing this range, this defect can be reduced. The labeled points have been evaluated by real dataset, SCATS LOG and Traffic Event Registration Software of Mashhad City. It represented that the method has 100 percent accuracy for 10th and 90th percentiles and between 57 and 80 percent accuracy for 1st and 3th quartiles. Furthermore, among the anomalies of the intersection, 5 points, two times were detected as congestion correctly. In this regard, it is suggested to design a software which announce if a 1st and 3th quartiles-based situation happens and alarm as anomaly when a 10th and 90th percentiles-based situation occurs.

7. ORCID

Mani Hazeghi https://orcid.org/ 0000-0003-3801-4056
Mahmoud Saffarzadeh https://orcid.org/0000-0002-9713-2776
Babak Mirbaha https://orcid.org/ 0000-0003-1004-7673

8. References

- Chen, X., Zhang, S., & Li, L. (2019). Multi-model ensemble for short-term traffic flow prediction under normal and abnormal conditions. IET Intelligent Transport Systems, 13(2), 260-268.

- Chen, Y. C., Liu, S. C., Chen, B. X., Loh, C. H., & Ying, J. J. C. (2020, December). Ensembling-mRBF-LSTM framework for prediction of abnormal traffic flows. In 2020 International Conference on Pervasive Artificial Intelligence (ICPAI) (pp. 206-213). IEEE.

- Dakic, Igor; Stevanovic, Aleksandar (2017) “On development of arterial fundamental diagrams based on surrogate density measures from adaptive traffic control systems utilizing stop-line detection”, Transportation Research Procedia 23, 942-961.

- G.Polson, Nicholas; O, Sokolov and Vadim (2017) “Deep learning for Short-term traffic flow Prediction”, Transportation Research Part C 79, 1-17, 25 February 2017.

- Ghosh, Bidisha; P.Smith and Damien. “Customization of Automatic Detection Algorithms for Signalized Urban Arterials”, Journal of Intelligent Transportation Systems, Volume 18, Issue 4, July 2014.

- Guo F. Polak J. Krishnan R.: ‘ Comparison of modelling approaches for short term traffic prediction under normal and abnormal conditions’. Proc. of the 13th Int. IEEE Annual Conf. on Intelligent Transportation Systems, Madeira Island, Portugal, September 2010, pp. 1209– 1214.

- Jiang, Shang; Ekbatani, Keyvan; Ngoduy, Dong (2022) “Partitioning of urban networks with polycentric congestion pattern for traffic management policies: identifying protected networks” Computer-Aided civil and infrastructure engineering, July 2022.

- Ken, Michael; Hompson, Keith; Newman, Peter; Hargrovs, Charlie; Stantic, Bela; Weeratunga, Kamal; Haque, Jannatun and Thomas, Kim(2017) ” Mining the Datasphere:

Big Data, Technologies and Transportation” Disaster Management, Version 1, January 2017.

- Kinane, Dermot; Hofmann, Dr.markus (2018) “Insights into Unusual Traffic Events in Dublin City, Ireland: Data Mining Delivering Business Intelligence” Civil engineering Conference, Dublin, Ireland, October 2018.

- Mihaita, Adriana-Simona; Haowen Li and Marian-Andrei Rizoiu (2020) "Traffic congestion anomaly detection and prediction using deep learning." arXiv: 2006.13215v1 Jun 2020.

- Piyapong, Suwanno(2021) “Study on Evaluation of Traffic Management Measures Using Macroscopic Fundamental Diagram (MFD) under Flooding Situation” Transportation Systems Engineering Major Graduate School of Science and Technology, Doctoral Course, Nihon University, January 2021.

- Ranaweera, Malith; A.Seneviratne, Rey,David; Saberi,Meead; V.Dixit, Vinayak(2021)” Detection of anomalous vehicles using physics of traffic” Vehicular Communications, Volume 27,100304, January 2021.

- SCATS Unusual Congestion Monitor, 6.5.1 user manual (2006), RTA-TC-335, Issue B, November 2006.

- Shafiei, Sajjad; Mihaita, Adriana-Simona; Nguyen, Hoang; Cai,Chen(2022) “Integrating data-driven and simulation models to predict traffic state affected by road incidents” , Transportation Letters, Volume 14, Issue 6, pages 629-639, August 2022.

- McMillan, Susan; Koorey, Glen and Dr.Alan Nicholson (2010)” Rechniques for using SCATS as an Incident Management Tool”, IPENZ Transportation Group Technical Conference, Christchurch, Newzeland, March 2010.

- Wu T. Xie K. Dong X. et al.: ‘ An online boosting approach for traffic flow forecasting under abnormal conditions’. The 9th Int. Conf. on Fuzzy Systems and Knowledge Discovery, Sichuan, China, May 2012, pp. 2555– 2559.

- Yi, Yu; Yanlei, Cui; Jiaqi, Zeng; Chunguang, He; Dihanhai, Wang(2022) “Identifying traffic clusters in urban networks based on graph theory using license plate recognition data” Physica A: Statistical Mechanics and its applizations , Volume 591, April 2022.

- Zheng, Lulu; Chen, Jiarui; Jianho, Wang; He, Jiamin; Hu, Yujing; Chen, Yingfeng; Fan, Changjie; Gao, Yang; Zhang, Chongjie(2021) “Episodic Multi-agent Reinforcement Learning With Curiosity-driven Exploration” arXiv:2111.11032v1 [CS.LG] 22 Nov 2021.

9. Appendix

Appendix 1: Codes

Code 1. Maximizing the polynomial fitted on FLOW vs DS

```
f <- function (x, a) 416.48+69.45*x-
1.79*(x - a)^2-0.11*(x-a)^3-0.001*(x-a)^4
xmax <- optimize (f, c(40, 120), tol =
0.0001, a = 92.31,maximum=TRUE)
print(xmax)
```

Code 2. FLOW3 vs ADS3; polynomial degree 4

```
f <- function (x, a) 2207.90-0.61*x-1.46*(x
- a)^2-0.005*(x-a)^3+0.0003*(x-a)^4
xmax <- optimize (f, c(30, 120), tol =
0.0001, a = 87.47,maximum=TRUE)
xmax
```

Code 3. FLOW3 vs speed; polynomial degree 4

```
f <- function (x, a) 2398.45-9.82*x-2.56*(x - a)^2+0.041*(x-a)^3+0.001*(x-a)^4
xmax <- optimize (f, c(10, 80), tol = 0.001, a = 37.09,maximum=TRUE)
xmax
```

Appendix 2: Flow3 values

$$\begin{aligned}For\ ADS = 64.9{:}\ Flow3 &= 2207.90 - 0.61(64.9) \\ &- 1.46(64.9 - 87.47)^2 \\ &- 0.005(64.9 - 87.47)^3 \\ &+ 0.0003(64.9 - 87.47)^4 \\ &= 1559.91\end{aligned}$$

$$\begin{aligned}For\ ADS = 107.3{:}\ Flow3 &= 2207.90 - 0.61(107.3) \\ &- 1.46(107.3 - 87.47)^2 \\ &- 0.005(107.3 - 87.47)^3 \\ &+ 0.0003(107.3 - 87.47)^4 \\ &= 1575.73\end{aligned}$$

$$\begin{aligned}For\ Speed = 25.2{:}\ Flow3 &= 2398.45 - 9.82(25.2) \\ &- 2.56(25.2 - 37.09)^2 \\ &+ 0.041(25.2 - 37.09)^3 \\ &+ 0.001(25.2 - 37.09)^4 \\ &= 1740.14\end{aligned}$$

$$\begin{aligned}For\ Speed = 47.5{:}\ Flow3 &= 2398.45 - 9.82(47.5) \\ &- 2.56(47.5 - 37.09)^2 \\ &+ 0.041(47.5 - 37.09)^3 \\ &+ 0.001(47.5 - 37.09)^4 \\ &= 1712.57\end{aligned}$$

$$\begin{aligned}For\ ADS = 77{:}\ Flow3 &= 2207.90 - 0.61(77) \\ &- 1.46(77 - 87.47)^2 \\ &- 0.005(77 - 87.47)^3 \\ &+ 0.0003(77 - 87.47)^4 \\ &= 2010.23\end{aligned}$$

$$\begin{aligned}For\ ADS = 99{:}\ Flow3 &= 2207.90 - 0.61(99) \\ &- 1.46(99 - 87.47)^2 \\ &- 0.005(99 - 87.47)^3 \\ &+ 0.0003(99 - 87.47)^4 \\ &= 1951.01\end{aligned}$$

$$\begin{aligned}For\ Speed = 31.3{:}\ Flow3 &= 2398.45 - 9.82(31.3) \\ &- 2.56(31.3 - 37.09)^2 \\ &+ 0.041(31.3 - 37.09)^3 \\ &+ 0.001(31.3 - 37.09)^4 \\ &= 1998.43\end{aligned}$$

$$\begin{aligned}For\ Speed = 44.2{:}\ Flow3 &= 2398.45 - 9.82(44.2) \\ &- 2.56(44.2 - 37.09)^2 \\ &+ 0.041(44.2 - 37.09)^3 \\ &+ 0.001(44.2 - 37.09)^4 \\ &= 1852.29\end{aligned}$$

$$\begin{aligned}For\ DS = 111.2{:}\ Flow &= 416.48 + 69.45(111.2) \\ &- 1.79(111.2 - 92.31)^2 \\ &- 0.11(111.2 - 92.31)^3 \\ &- 0.001(111.2 - 92.31)^4 \\ &= 6631.8\end{aligned}$$

$$\begin{aligned}For\ DS = 71.9{:}\ Flow &= 416.48 + 69.45(71.9) \\ &- 1.79(71.9 - 92.31)^2 \\ &- 0.11(71.9 - 92.31)^3 \\ &- 0.001(71.9 - 92.31)^4 \\ &= 5425.99\end{aligned}$$

$$\begin{aligned}For\ DS = 83.6{:}\ Flow &= 416.48 + 69.45(83.6) \\ &- 1.79(83.6 - 92.31)^2 \\ &- 0.11(83.6 - 92.31)^3 \\ &- 0.001(83.6 - 92.31)^4 \\ &= 6653.8\end{aligned}$$

$$\begin{aligned}For\ DS = 103.6{:}\ Flow &= 416.48 + 69.45(103.6) \\ &- 1.79(103.6 - 92.31)^2 \\ &- 0.11(103.6 - 92.31)^3 \\ &- 0.001(103.6 - 92.31)^4 \\ &= 7208.79\end{aligned}$$