

Article

Performance Assessment and Modeling of Routing Protocol in Vehicular Ad Hoc Networks Using Statistical Design of Experiments Methodology: A Comprehensive Study

Souad Ajjaj ^{1,*}, Souad El Houssaini ², Mustapha Hain ¹ and Mohammed-Alamine El Houssaini ³

¹ National School of Arts and Crafts (ENSAM), Hassan II University, Casablanca 20670, Morocco; MUSTAPHA.HAIN@univh2c.ma

² Faculty of Sciences, Chouaib Doukkali University, El Jadida 24000, Morocco; elhoussaini.s@ucd.ac.ma

³ High School of Education and Training (ESEF), Chouaib Doukkali University, El Jadida 24000, Morocco; elhoussaini.m@ucd.ac.ma

* Correspondence: SOUAD.AJAJ-ETU@etu.univh2c.ma

Abstract: The performance assessment of routing protocols in vehicular ad hoc networks (VANETs) plays a critical role in testing the efficiency of the routing algorithms before deployment in real conditions. This research introduces the statistical design of experiments (DOE) methodology as an innovative alternative to the one factor at a time (OFAT) approach for the assessment and the modeling of VANET routing protocol performance. In this paper, three design of experiments methods are applied, namely the two-level full factorial method, the Plackett–Burman method and the Taguchi method, and their outcomes are comprehensively compared. The present work considers a case study involving four factors namely: node density, number of connections, black hole and worm hole attacks. Their effects on four measured outputs called responses are simultaneously evaluated: throughput, packet loss ratio, average end-to-end delay and routing overhead of the AODV routing protocol. Further, regression models using the least squares method are generated. First, we compare the main effects of factors resulted from the three DOE methods. Second, we perform analysis of variance (ANOVA) to explore the statistical significance and compare the percentage contributions of each factor. Third, the goodness of fit of regression models is assessed using the adjusted R-squared measure and the fitting plots of measured versus predicted responses. VANET simulations are implemented using the network simulator (NS-3) and the simulator of urban mobility (SUMO). The findings reveal that the design of experiments methodology offers powerful mathematical, graphical and statistical techniques for analyzing and modeling the performance of VANET routing protocols with high accuracy and low costs. The three methods give equivalent results in terms of the main effect and ANOVA analysis. Nonetheless, the Taguchi models show higher predictive accuracy.

Keywords: VANETs; NS-3; SUMO; AODV; performance assessment; design of experiments methodology; regression analysis; ANOVA



Citation: Ajjaj, S.; El Houssaini, S.; Hain, M.; El Houssaini, M.-A. Performance Assessment and Modeling of Routing Protocol in Vehicular Ad Hoc Networks Using Statistical Design of Experiments Methodology: A Comprehensive Study. *Appl. Syst. Innov.* **2022**, *5*, 19. <https://doi.org/10.3390/asi5010019>

Academic Editor: Luis Oliveira

Received: 22 December 2021

Accepted: 28 January 2022

Published: 2 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Vehicular ad hoc networks are a key component of intelligent transportation systems (ITS) [1,2], which one of the leading domains in smart cities used to enhance traffic efficiency and reduce transportation problems. VANETs (vehicular ad hoc networks) are a subclass of mobile ad hoc networks (MANETs), where mobile nodes are intelligent vehicles communicating in a self-organized manner without relying on any fixed infrastructure support. VANETs are subject to several challenges due to their special features: mainly, the highly dynamic topology due to high mobility, the intermittent connectivity and the high delay constraints [1]. Given these special characteristics, the performance of routing protocols in VANETs can be highly affected by the traffic conditions, the vehicular density, the mobility model, the security mechanisms [2], etc. Therefore, a reliable performance analysis study is

an essential prerequisite used to identify the way different factors influence the network performance, alongside giving more insights for optimization and improvement prospects. In that regard, various works in the literature have conducted performance analysis studies wherein the conventionally used approach is to study the effect of one factor at a time (OFAT). This approach, which varies only one factor at a time while keeping others constant, is demanding and time consuming [3]. Thus, in order to overcome this problem, we opted for statistically designed experiments that vary several factors simultaneously, known as the design of experiments methodology. The present work focuses on the comparison of the results of three DOE methods namely: the two-level full factorial method (FFM) [4,5], the Plackett–Burman method (PBM) [6] and the Taguchi method (TM) [7]. The details of each method were presented at international conferences [8–10]. The effects of four factors were assessed—node density (ND), number of connections (NC), black hole (BKH) and worm hole (WMH) attacks—on the following performance metrics: throughput (TH), packet loss ratio (PLR), average end-to-end delay (AE2ED) and routing overhead (RO) of the AODV routing protocol. For VANET simulations, we used NS-3 [11] and SUMO [12,13].

First, the comparison is performed in terms of the main effects of factors on the considered responses that resulted from each DOE method. Secondly, analysis of variance is conducted to determine the factors that are statistically significant and compare their percentage contributions. The adjusted R-squared measure along with the plots of measured versus predicted values of the response are employed to assess the goodness of fit of the regression models. Additionally, only significant estimates are included in the final regression models.

Our study demonstrates that the statistical design of experiments methodology is a practical approach to analyze the performance of any routing protocol in VANETs. The comparison results indicate that the three methods give practically identical outcomes with regard to the main effects and the ANOVA analysis. However, the Taguchi method provides higher descriptive and predictive qualities based on the adjusted R-squared and visual plots of measured versus predicted responses.

The rest of this paper is organized as follows: Section 2 outlines the previous works in the literature. The following section details the proposed methodology along with implementation tips. Then, an exhaustive comparative analysis of the findings obtained from the three DOE methods is given in Section 4. Sections 5 and 6 end the paper with concluding remarks and future research directions.

2. Related Work

Various researchers in the literature conducted performance analyses of various routing protocols in VANETs, with respect to several performance metrics. For instance, in the study [14], the performance results of seven different routing protocols were examined based on performance metrics such as packet delivery ratio, average throughput, delay and overall energy consumption. The authors highlighted that the performance of VANET routing protocol depends heavily upon the mobility model, the vehicle density and the propagation rate of data. The research [15] analyzed the throughput, packet delivery ratio (PDR), and normalized routing load (NRL) of AODV, DSR, and DSDV. The outcomes emphasize that AODV is more appropriate than DSR and DSDV in terms of PDR and throughput; however, the NRL for DSDV remains lower than AODV and DSR under different numbers of vehicles.

The study [16] examined the performance of several routing protocols, including AODV, DSR, and DSDV, using different performance parameters, such as data throughput, PDR, end-to-end delay, and network stability. The authors stated that it is difficult to select a routing protocol that suits all requirements. They outlined that the efficiency of AODV protocol is better than DSR and DSDV protocols. In Ref. [17], the authors investigated the effect of traffic types (UDP and TCP) on AODV delay and throughput. They concluded that the traffic type UDP is better than TCP for safety applications and highlighted the necessity of implementing an improved version of AODV routing protocol. The researchers

in [18] developed a new routing protocol labeled LENC (low-latency and energy-efficient routing based on network connectivity). Extensive simulations were conducted to compare AODV and the proposed routing protocol. Their findings outline the usefulness of LENC in providing more routing stability and the network connectivity. In [19], the authors proposed an enhanced version of a protocol based on cluster approach. The results prove that the proposed protocol outperforms CBCLR and AODV-CV protocols regarding throughput, packet delivery ratio, and end-to-end delay parameters while varying the number of vehicular nodes in the network. In the paper [20], researchers discussed the most important security features of the routing protocols in VANET and examined the effect of black hole attacks on AODV, along with attempting to introduce a mechanism to detect and prevent the network from such attacks. Similarly, the authors of paper [21] mitigated the impact of black hole attacks and vehicle density on the average throughput, packet delivery ratio, end-to-end delay, normalized routing load and average path length of AODV and ZRP routing protocols. They illustrated the vulnerability of these protocols to black hole attacks and recommended performing suitable solutions to mitigate the black hole attack in VANETs. In Ref. [22], the authors highlighted the high vulnerability of AODV to wormhole attacks. They proposed a practical approach for dealing with such attacks.

In all the aforementioned studies, the common adopted approach relies on varying each factor individually, referred to as the OFAT approach (one factor at a time) [3]. The main problem of OFAT is that the number of experiments goes up as the number of studied factors increases. Furthermore, the OFAT approach does not allow direct comparisons of the magnitudes of each factor's impact on the investigated responses. In addition, it can neither estimate interactions among factors nor determine the optimal settings of factors.

From the previous mentioned studies, we note that performance analysis studies based on the statistical design of experiments approaches have not gained much attention in the literature. Hence, the aim of the next section is to introduce its application to the assessment and the modeling of VANET routing protocols performances.

3. Materials and Methods

3.1. VANET Simulation

In the present study, we simulated a realistic VANET scenario based on the combination of the road traffic simulator SUMO (simulation of urban mobility) and the network simulator NS-3. SUMO is a free, open and microscopic simulator implemented in C++. It is destined to simulate unlimited network size and number of vehicles. It offers the possibility to configure vehicle types, traffic lights, vehicle speeds, multi-lane roads; it also supports the lane-changing model, as well as generating automatic time schedules for traffic lights. SUMO also supports the import formats, such as OpenStreetMap [12,13]. We executed a set of Python command lines on SUMO in order to generate realistic vehicle trace files that are then used as an input by the network simulator, NS-3 [11]. In our study, the simulation zone was extracted from OpenStreetMap, consisting of a map of El Jadida city, Morocco (Figure 1).

Additionally, the AODV (ad hoc on demand distance vector) is adopted as a routing protocol. It is a reactive protocol wherein routes are created only when demanded. AODV relies on the sequence numbers to determine fresher and newer paths. The most important mechanisms of this protocol are the route discovery and the route maintenance. This protocol is comprehensively detailed in Ref. [23].

Other simulation components are exhibited in Table 1.



Figure 1. Simulation zone from El Jadida city from Open Street Map and XML file edited by SUMO.

Table 1. Components of the simulation.

N	Parameter	Value
1	Network simulator	NS3.29
2	Mobility Simulator	SUMO-0.32.0
3	Propagation Model	TwoRayGroundPropagationLossModel
4	Wifi Channel	YansWifi
5	Mac and Physic Layer	IEEE 802.11p
6	Simulation Time	200 s
7	Traffic type	CBR
9	Data rate	50 Kbps
10	Packet Size	1024 bytes
11	Routing protocol	AODV

3.2. Design of Experiments Methodology

The theory of the design of experiments (DOE) methodology was first introduced by Ronald Aylmer Fisher in the early 20th century for agriculture [24]. Since then, it has evolved into many scientific areas, such as medicine, biochemistry, physics, engineering and computer science. DOE offers powerful mathematical, graphical and statistical tools for planning experiments, data visualization and analysis [4]. Its main feature is the possibility of simultaneous assessment of the effect of multiple input variables (factors) on output variables (responses) with the minimum number of experimental tests [25].

Unlike the one factor at a time (OFAT) approach, the application of design of experiments (DOE) [3,26] has numerous advantages. First, this methodology offers potentially useful tools for providing efficient estimates of the factor effect as well as the effect of possible interactions among factors. Another important aspect of this approach is the possibility of performing mathematical models that characterize the factor–response relationship. Moreover, the design of experiments (DOE) methodology has proven to be an extremely useful tool for conducting optimization studies by determining the optimal settings of factors that lead to optimal results.

Different types of DOE methods are generally used: the full-factorial designs, the fractional factorial designs, the screening designs, the Taguchi design, etc. The most commonly used full-factorial design is the two-level full factorial, where all factors are set at two levels (a low level coded (−) and a high level denoted (+)). A feature characteristic of this design is that all possible combinations for factors and their levels are examined. For studying k factors, a 2^k number of experimental runs is required. However, these designs become costly as the number of factors increases. Consequently, fractional factorial designs are the alternative, in which a selected fraction of the full factorial design is employed. One of the well-known designs of the two-level fractional factorial is the Plackett–Burman, used for screening experiments [6], where it is possible to investigate up to $N-1$ factors in a number N of experiments. Another popular method is the Taguchi method, developed

in the early 1980's by the Japanese, Dr. Genichi. It is widely used among engineers and scientists for quality and performance optimization [7]. Taguchi makes use of two concepts: (OA) orthogonal arrays (a library of predefined matrices to construct matrices of experiments) and the signal-to-noise ratio (SNR) used as a statistical measure of quality performance.

The steps involved in implementing the DOE methodology are outlined in the diagram below (Figure 2):

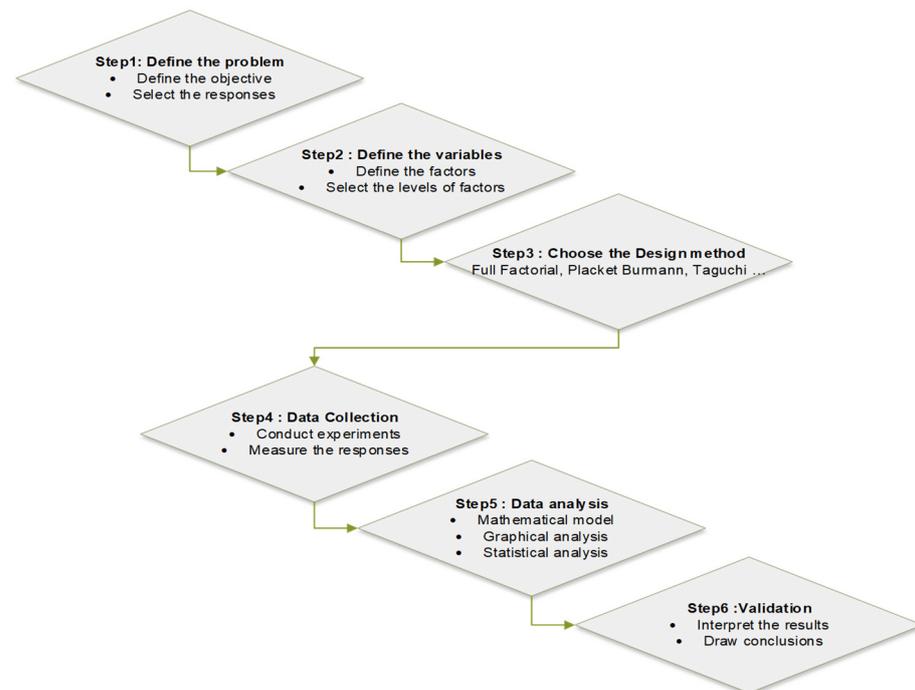


Figure 2. Design of experiments process.

The DOE methodology starts with stating the objectives of the study before identifying the responses or the measures of performance. After that, the studied factors are defined, and their levels are selected. The levels are the ranges of values the factors can assume within practical limits.

Using the determined factors and their levels, the corresponding design matrix is constructed, according to the adopted DOE method. This design matrix refers to the combinations of factor levels for each experimental run. For each combination, the responses are measured and reported in the so-called matrix of experiments. Further, mathematical models are built to establish a relationship between each response and the studied factors.

The next obvious step is the analysis of the collected data. First, the graphical analysis allows a more visual insight of the effects of the factors and their interactions on the response using the diagram of effects. The mathematical analysis consists of determining the coefficients of the mathematical models. The statistical analysis consists of determining the factors that are statistically significant and their percentage contribution by applying statistical methods, such as ANOVA (analysis of variance). Finally, validation of the acquired results and the developed models is achieved using statistical metrics and graphical tools, such as the R-squared coefficient or the adjusted R-squared coefficient, the normal probability plots of residuals, etc.

In the present study, we consider comparing the results from applying the following DOE methods:

1. Two-level full factorial method (FFM);
2. Plackett–Burman method (PBM);
3. Taguchi method (TM).

The description of each method is available in the previous works presented at international conferences [8–10]. The experimental setup and the implementation details are given in the next subsection.

3.3. Experimental Setup and Implementation

3.3.1. Definition of Objectives

The objective behind applying each method is to assess the effects of factors related to scalability, traffic conditions and routing security attacks on the performance metrics of the AODV routing protocol in VANETs. Secondly, we build regression models to model the performance metrics of AODV routing protocol. Finally, we analyze the obtained results using graphical, mathematical and statistical techniques.

3.3.2. Definition of Responses

In our study, we selected the following performance metrics as the responses: throughput (TH), packet loss ratio (PLR), average end-to-end delay (AE2ED) and routing overhead (RO), computed using the equations below.

Throughput is the total number of bits successfully received during simulation time. It is expressed in kilobits per second (Kbps).

$$TH(Kbps) = \frac{\text{totalRxBytes} * \text{PacketSize} * 8}{T * 1000} \quad (1)$$

where totalRxBytes and PacketSize represent the total received bytes and the packet size, respectively. T is the difference between the time when the last packet is received and the time when the first packet is sent.

Packet loss ratio represents the ratio of the number of lost packets to the total number of transmitted packets.

$$PLR (\%) = 100 * \frac{\text{totalTxPackets} - \text{totalRxPackets}}{\text{totalTxPackets}} \quad (2)$$

where totalTxPackets and totalRxPackets refer to the total amount of sent and received packets, respectively.

Average end-to-end delay is the average time needed to transmit data from the source to destination, and it is measured in milliseconds (ms).

$$AE2ED(ms) = \frac{\text{totalDelays}}{\text{totalRxPackets}} \quad (3)$$

where totalDelays represents the summation of all delays of received packets, which corresponds to the difference between the transmission time and reception time.

Routing overhead (RO) is the ratio between the total number of control packets generated by the AODV routing protocol during route discovery and route maintenance to the total number of transmitted packets.

$$RO = \frac{\text{totalControlPackets}}{\text{totalTxPackets}} \quad (4)$$

where totalControlPackets and totalTxPackets refer to the total number of control packets generated by AODV and the total amount of sent packets, respectively.

3.3.3. Definition of Factors and Their Levels

Vehicular ad hoc networks rely on wireless communication with no infrastructure support; this yields numerous vulnerabilities [1], such as black hole and wormhole attacks [27].

In the current study, black hole (BKH) and wormhole (WMH) are selected as input variables or factors. Further, this study aims at evaluating AODV under different density and traffic levels. Hence, the number of vehicles (ND) and the number of connections (NC),

which is the number of source–destination pairs, are also considered. Table 2 lists these factors and their levels.

Table 2. Values of low and high levels for factors.

Factor	Variable	Low Level (−1)	High Level (+1)
Black Hole attack	BKH	0	1
Wormhole attack	WMH	0	1
Node Density	ND	50	150
Number of Connections	NC	4%	20%

3.3.4. Selection of the Design Matrix

The design matrix for the full factorial method is composed of $24 = 16$ experiments since 4 factors are involved. Those for the PBM are based on HDAMARD matrices with 8 experiments. The Taguchi design relies on an L8 orthogonal array of 8 experiments (see Appendix A). Three repetitions are carried out for each experiment.

3.3.5. Calculation of the Main Effects of Factors

The main effect of a factor refers to how this factor influences the response when factor settings change from the low level to the high level. It is the difference between the mean responses at high level M_{Factor^+} and low level M_{Factor^-} (see Equation (5)):

$$E_{\text{Factor}} = M_{\text{Factor}^+} - M_{\text{Factor}^-} \tag{5}$$

where M_{Factor^+} and M_{Factor^-} are, respectively, the mean of the responses when Factor is at a high level (+1) and the mean of the responses when Factor is at a low level (−1).

In our study and for comparison purposes, these averages were normalized by dividing each mean value by the standard deviation σ of the given values of the response (Equations (6) and (7)).

$$\text{Norm}M_{\text{Factor}^+} = \frac{M_{\text{Factor}^+}}{\sigma} \tag{6}$$

$$\text{Norm}M_{\text{Factor}^-} = \frac{M_{\text{Factor}^-}}{\sigma} \tag{7}$$

where $\text{Norm}M_{\text{Factor}^+}$ and $\text{Norm}M_{\text{Factor}^-}$ are, respectively, the normalized main effects when Factor is set at a high level and low level. σ is the standard deviation of all values of the response.

The plots of effects graphically represent these main effects with lines connecting the normalized values of mean responses at the low level with those calculated at the high level. The direction of the line indicates the sign of the effect (positive or negative), while the inclination of the line reveals the magnitude of the effect. When the line is horizontal, no effect is revealed.

3.3.6. ANOVA Analysis

The statistical analysis of variance method (ANOVA) [28] is used to analyze the data from the three experimental designs. The ANOVA method employs a hypothesis test to determine the factors that are statistically significant (p -value < 0.05), by means of the F-test with a confidence level of 95%. The percentage contribution (P%) of each factor is then used to quantify the degree of influence of each factor on each response. The ANOVA method is detailed in Ref. [28]. We give below the formulas used for ANOVA calculations.

1. The total sum of squares:

$$SS_T = \sum_{i=1}^n (y_i - \mu) \tag{8}$$

where y_i is the response of the i -th experiment, μ represents the mean of all responses, and n is the number of all experimental runs.

2. The sum of squares of error:

$$SS_E = SS_T - \sum_{i=1}^k SS_i \tag{9}$$

where SS_i is the sum of squares of the i -th factor and k the number of factors.

3. Mean of squares or the variance:

$$MS = \frac{SS}{df} \tag{10}$$

where df is the degree of freedom. It is equal to number of levels minus 1 for a given factor X . The total degrees of freedom are equal to the number of all experimental runs minus 1.

4. F-ratio: the variance of a factor X to the variance of error:

$$F_X = \frac{MS_X}{MS_E} \tag{11}$$

5. The percentage of contribution of factor X :

$$P\% = \frac{SS_X}{SS_E} \tag{12}$$

The measured experimental responses are employed in both the PBM and the FFM, whereas the Taguchi method makes use of the SNR analysis. Indeed, the Taguchi method converts the measured responses to signal-to-noise ratios (SNR). The SNR measures how the response varies relative to a target function defined depending on the goal of the experiment. If the goal is to maximize the response, then the “larger-the-better” (LTB) formulation is selected, and when the objective is to minimize the response, “smaller-the-better” (STB) is adopted.

In our study, LTB is chosen for throughput and STB is assigned to the packet loss ratio, average end-to-end delay and routing overhead.

Smaller-the-better (STB),

$$SNR = -10 \log \left[\frac{1}{N} \sum_{u=1}^N (y_u)^2 \right] \tag{13}$$

Larger-the-better (LTB),

$$SNR = -10 \log \left[\frac{1}{N} \sum_{u=1}^N \frac{1}{(y_u)^2} \right] \tag{14}$$

where N is the number of repetitions for an experiment, and y_u is the corresponding measured value of a response, such as throughput, packet loss ratio, average end-to-end delay or routing overhead. In this study, three repetitions are performed for each experiment.

3.3.7. Regression Models

In our study, we adopted a linear regression model to express the relationship between the factors and the responses. The goal is to establish a relationship between each dependent variable (throughput, packet loss ratio, average end to end delay and routing overhead) and the independent variables (factors ND, NC, BKH and WMH). Models of this relationship or estimated regression equations are used to predict the value of the response when values for the independent variables are given, always within the defined ranges from each

factor levels. These models will be very useful in the context of our study, where VANET simulations generally require high processing and execution times.

This mathematical relationship between a response y and factors $x_1 \dots x_k$ is given in Equation (11).

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon \tag{15}$$

where y represents the response, $x_1 \dots x_k$ are the coded variables of the k factors, β_0 is the constant of the model, $\beta_1, \beta_2 \dots \beta_k$ are the linear model coefficients, and ε is the error term.

To find out the coefficients, we used the linear least squares method [29]. This relationship can be expressed in a matrix form as follows:

$$B = (X)^{-1} (X^T Y) \tag{16}$$

B is the matrix of coefficients of order $k + 1$.

X^T is the transposed matrix of X The order of this matrix is $N \times p$ with $p = k + 1$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix} \tag{17}$$

where x_{ij} represents the level of the j th factor at the i th experiment, the Y column matrix of the n measured responses.

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \quad B = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \vdots \\ \beta_k \end{pmatrix} \tag{18}$$

In order to assess the goodness of fit of these models, the adjusted R-squared coefficient is adopted. This coefficient is based on the R-squared coefficient that adjusts for predictors that are not significant in a regression model [30]. The values of adjusted R-squared are between 0 and 1. When their values are closer to 1, this suggests that the model has good predictive accuracy. The adjusted R-squared (Adj.R²) coefficient is measured by the following equations (Equations (19) and (20)):

$$R^2 = 1 - \frac{SS_T}{SS_E} \tag{19}$$

$$Adj.R^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - k - 1} \tag{20}$$

where SS_T is the total sum of squares and represents the total variation, SS_E is the residual sum of squares, n is the total sample size, and k the number of predictors in the regression model.

4. Results and Discussion

4.1. Comparison of the Main Effects

This section deals with the comparison of the main effects of the four factors on the four performance metrics or responses.

4.1.1. Main Effects of Factors on Throughput

Figure 3 illustrates the main effects of factors on the AODV throughput (TH) from the FFM, PBM and TM, respectively. First, we notice that the values of normalized effects from the FFM are closer to those from the Taguchi method, unlike the PBM, which shows quite different values. Secondly, the trends of the main effects for factors ND, BKH and WMH are almost the same from the three methods. They all have a negative effect on the AODV throughput. This finding reveals that when the number of vehicles grows, and when the black hole and the wormhole attacks are initiated, the AODV throughput degrades. These results are logical since in the black hole attack, the attacker node acts by dropping the entire data packet forwarded to it, pretending to have the shortest path to the destination. Similarly, the two colluding nodes in the wormhole attack create a tunnel between them, claiming that they are next to each other. Hence, neighboring nodes believe that the shortest path passes through this tunnel and thus, they begin the transmission. Consequently, the majority of the data packets are absorbed; they are either dropped or replayed in the network. Therefore, the amount of successfully forwarded packets decreases, and the AODV routing performance degrades.

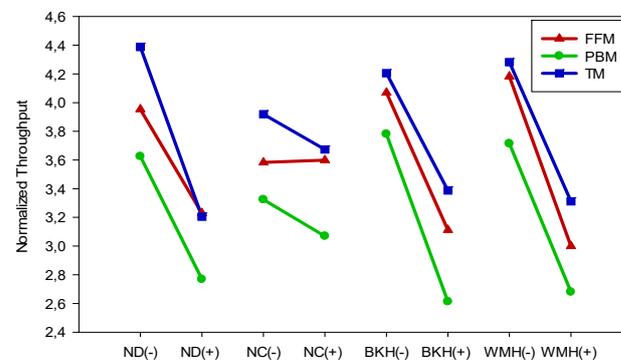


Figure 3. Comparison of main effects on throughput (TH) from FFM, PBM and TM.

Additionally, it is stated that the negative impact of the node density on throughput is important. This can be justified by the dynamic topology and the high mobility of vehicles. Indeed, the wireless link may fail in transmitting all the data packets and generate more route error packets (RERR).

As far as the factor NC, we notice that it has a negligible effect from the full factorial method, as the slope of the line is almost horizontal. This could be due to the ranges defined for each factor. Moreover, the magnitude of the effect on the throughput is quite important in both TM and PBM. Indeed, the increment in the number of communicating pairs of vehicles lowers the successfully transmitted packets. This decrease may be due to many factors, such as collision and transmission errors, so packets fail in reaching their destination resulting in the degradation of AODV throughput.

In summary, the graphical analysis of the results obtained from the three methods revealed three main influential factors (ND, BKH and WMH). This means that the three methods were able to efficiently estimate the effect of factors on the AODV throughput.

4.1.2. Main Effects of Factors on Packet Loss Ratio

The figure below (Figure 4) contrasts the main effects of factors on the AODV packet loss ratio (PLR) obtained from the three DOE methods. It is clear that the normalized effects of the FFM cope perfectly with those from the TM, while the values of normalized effects given by the PBM present quite small differences.

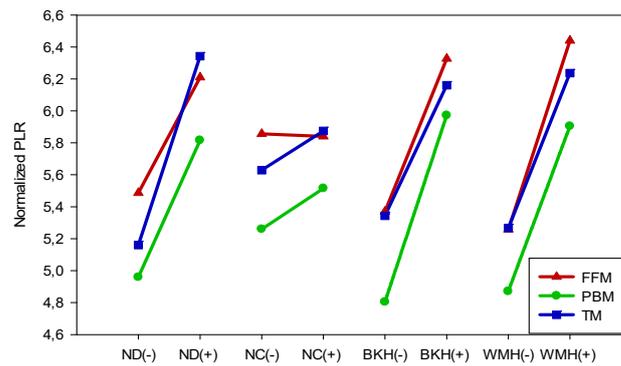


Figure 4. Main effects on the packet loss ratio (PLR) from FFM, PBM and TM.

The trends of the main effects for factors ND, BKH and WMH from the FFM and TM are closer to those obtained from the PBM. They all have a positive effect on the packet loss ratio. The difference lies in the effect of the factor NC. We note that it has a slight positive effect from the PBM and TM. This effect, however, is not highly important in the FFM. These outcomes point out that in scaled-up networks as well as when the network flows go up, the AODV packet loss ratio increases correspondingly. Moreover, the black hole and the wormhole attacks impact considerably the AODV performance and disrupt the correct execution of the routing operations.

Briefly, both the FFM and the TM produce closer results. These results show small differences with the PBM. We conclude that they all allow efficient estimates of the main effects within the defined ranges.

4.1.3. Main Effects of Factors on End-to-End Delay

Figure 5 shows the main effects of factors with regard to average end-to-end delay. The three methods, FFM, PBM and TM, produce similar results regarding the trends of the main effects. Indeed, the plots report that the change in the levels from a low to high level leads to an increase in average end-to-end delay for factors ND, NC and WMH. The conclusions with respect to ND and NC factors are justified by the mechanism of route discovery established by the AODV routing protocol that adds delays before data transmission. Likewise, the two colluding attackers in the wormhole attack create a distant link between them, claiming that all routes must be formed through this wormhole link, causing a failure in node connectivity and resulting in infinite delays.

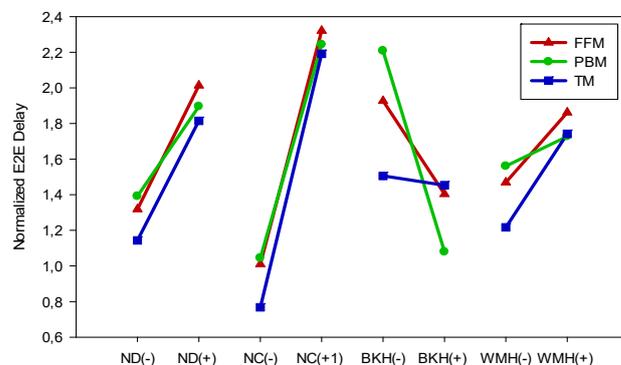


Figure 5. Main effects on AODV end-to-end delay (E2ED) from FFM, PBM and TM.

In summary, the effects of ND, NC and the WMH are the same from the three methods. The main difference is in the effect of BKH. The FFM and Taguchi method outline that the black hole attack has no direct impact on average end-to-end delay. In that regard, the findings of both the FFM and Taguchi method seem to be more logical due to the fact that the time spent in the AODV route discovery phase is reduced when a black hole attack is

initiated because the malicious node sends a fast RREP message to the source node in the AODV route reverse path.

4.1.4. Main Effects of Factors on Routing Overhead

Figure 6 summarizes the main effects of factors on AODV routing overhead derived from the three methods. It appears that the trends of the main effects produced by the FFM and the TM are equivalent for the four factors. Compared to those obtained by applying the PBM, the plots indicate that the only difference observed consists of the effect of the NC factor. This latter presents a more important negative effect on the routing overhead.

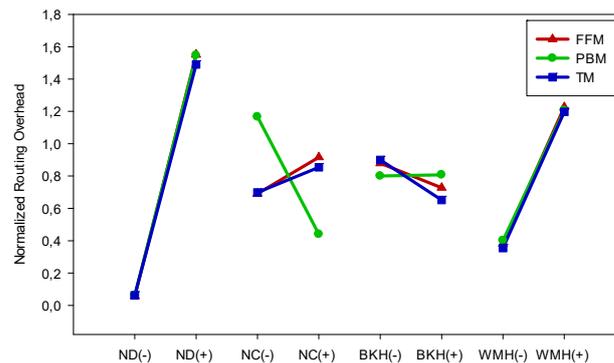


Figure 6. Main effects on routing overhead (RO) from FFM, PBM and TM.

Regarding the overall magnitude of the effects, it is clearly indicated that the factor ND has the biggest influence on the routing overhead. This observation is due to the fact that AODV routing protocol generates enormous quantities of control packets in scaled-up scenarios. This is merely due to the large number of RREQ messages produced by AODV and the specific features of VANETs, such as a highly dynamic environment and the frequent connections and disconnections of vehicles. Further, it is revealed that the routing overhead is not very sensitive to black hole attacks. This observation can be supported by the fact that the malicious node broadcasts a fake and fast route reply (RREP) to make other nodes believe that it owns the shortest path. Accordingly, the overhead is cut down because the route request (RREQ) packets are lesser.

These findings suggest that the Taguchi method gives closer results to the FFM, unlike the PBM that show quite different values. However, the three methods, including the PBM, are able to give an efficient estimate of the most influential factors within the defined ranges.

4.2. ANOVA Results

In this subsection, the analysis of variance (ANOVA) is conducted to assess the statistical significance of factors, along with their percentage contributions. First, the results of ANOVA based on the F test are exhibited in Tables 3–6. The critical F values, using a 5% significance level (95% confidence level) from the Fisher table are, respectively, $F_{0.05}(1, 11) = 4.84$, $F_{0.05}(1, 3) = 10.13$ and $F_{0.05}(1, 2) = 18.51$ for the full factorial, Plackett–Burman and Taguchi methods. Factors with an F-ratio greater than these critical values are considered statistically significant denoted by (*). Further, the percentage contributions (P%) of factors from the three methods are also given. The ANOVA tables show also the adjusted R-squared measure.

Table 3. ANOVA statistical significance and percentage contribution of the factors on throughput.

	Full Factorial				Plackett–Burman				Taguchi			
	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig
ND	1	7.0 *	14.5%	Significant	1	12.7 *	27.5%	Significant	1	134.9 *	35.5%	Significant
NC	1	0.1	0.3%	Insignificant	1	0.2	0.4%	Insignificant	1	8.2	2.2%	Insignificant
NDxNC	–	–	–	–	–	–	–	–	1	15.7	4.1%	Insignificant
BKH	1	11.7 *	24.4%	Significant	1	13.5 *	29.2%	Significant	1	140.1 *	36.9%	Significant
WMH	1	18.1 *	37.8%	Significant	1	16.8 *	36.3%	Significant	1	79.2 *	20.8%	Significant
Error	11	–	22.9%	–	3	–	6.5%	–	2	–	0.5%	–
Total	15	–	–	–	7	–	–	–	7	–	–	–
Adj.R ²	0.681				0.849				0.982			

Table 4. ANOVA statistical significance and percentage contribution of the factors on PLR.

	Full Factorial				Plackett–Burman				Taguchi			
	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig
ND	1	5.6 *	13.6%	Significant	1	19.2 *	21.7%	Significant	1	207.1 *	13.2%	Significant
NC	1	0.3	0.6%	Insignificant	1	8.7	9.8%	Insignificant	1	62.2 *	4.0%	Significant
NDxNC	–	–	–	–	–	–	–	–	1	10.4	0.7%	Insignificant
BKH	1	9.8 *	23.9%	Significant	1	25.0 *	28.2%	Significant	1	451.6 *	28.8%	Significant
WMH	1	14.6 *	35.6%	Significant	1	32.7 *	36.9%	Significant	1	836.2 *	53.3%	Significant
Error	11	–	26.9%	–	3	–	3.4%	–	2	–	0.1%	–
Total	15	–	–	–	7	–	–	–	7	–	–	–
Adj.R ²	0.634				0.821				0.981			

Table 5. ANOVA significance and percentage contribution of the factors on AE2ED.

	Full Factorial				Plackett–Burman				Taguchi			
	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig
ND	1	7.5 *	16.4%	Significant	1	10.4 *	13.1%	Significant	1	33.8*	31.7%	Significant
NC	1	18.9 *	41.4%	Significant	1	35.2 *	44.2%	Significant	1	29.7 *	27.9%	Significant
NDxNC	–	–	–	–	–	–	–	–	1	0.2	0.1%	Insignificant
BKH	1	7.4 *	16.3%	Significant	1	21.0 *	26.4%	Significant	1	34.9 *	32.8%	Significant
WMH	1	0.8	1.8%	Insignificant	1	10.0	12.5%	Insignificant	1	6.1	5.7%	Insignificant
Error	11	–	24.1%	–	3	–	3.8%	–	2	–	1.9%	–
Total	15	–	–	–	7	–	–	–	7	–	–	–
Adj.R ²	0.671				0.812				0.858			

Table 6. ANOVA significance and percentage contribution of the factors on routing overhead.

	Full Factorial				Plackett–Burman				Taguchi			
	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig	DoF	F	P%	Sig/Insig
ND	1	242.8 *	84.9%	Significant	1	93.0 *	83.0%	Significant	1	550.1 *	88.6%	Significant
NC	1	7.3 *	2.5%	Significant	1	6.8	6.1%	Insignificant	1	48.8 *	7.9%	Significant
NDxNC	–	–	–	–	–	–	–	–	1	7.1	1.1%	Insignificant
BKH	1	3.9	1.4%	Insignificant	1	0.8	0.7%	Insignificant	1	6.4	1.0%	Insignificant
WMH	1	20.9 *	7.3%	Significant	1	8.4	7.5%	Insignificant	1	6.8	1.1%	Insignificant
Error	11	–	3.8%	–	3	–	2.7%	–	2	–	0.3%	–
Total	15	–	–	–	7	–	–	–	7	–	–	–
Adj.R ²	0.633				0.791				0.788			

Examining the F-ratio column in Table 3, we note that the full factorial, the Plackett–Burman and the Taguchi methods give approximately identical results for the statistically significant effects with no obvious differences. With regard to throughput, the effects that are statistically significant include node density, black hole and wormhole attacks from the three methods, while the effect of the number of connections remains insignificant. Based on the percentage contributions of factors from the different methods, we observe that the ranks of factors are practically the same from the full factorial and the Plackett–Burman

methods. Indeed, the wormhole attack has the most impact, succeeded by the black hole attack and node density. According to the Taguchi method, the black hole attack has the largest effect, followed by node density and the wormhole attack. However, the interaction and the factor number of connections have the least contribution.

Based on the F-ratio exhibited in Table 4, we observe that the ANOVA results regarding the packet loss ratio of both Plackett–Burman and full factorial reveal three significant effects, which are node density, black hole and wormhole. Similarly, these results are close to the Taguchi method, except for the factor NC, that shows a significant effect but lesser contribution. The effect of the interaction remains insignificant as estimated by the Taguchi method. Based on the percentage contribution of each factor, it is seen that the most influential factors from the three methods are, respectively, wormhole (WMH: 35.6%, 36.9% and 53.3%), black hole (BKH: 23.9%, 28.2%, 28.8%), and node density (ND: 13.6%, 21.7% and 13.2%); however, the number of connections has the lowest contributions (0.6%, 9.8% and 4.0%).

Table 5 illustrates the ANOVA results with respect to average end-to-end delay. It is clear that the three methods produce equivalent results in terms of the factor's statistical significance. The number of connections, node density and the black hole attack are revealed to be statistically significant, whereas the wormhole attack has an insignificant effect and the lowest percentage contribution. The Taguchi method indicates an insignificant effect of the interaction between node density and number of connections.

The ANOVA results relative to the routing overhead show that the node density (ND) is the factor that produces the most important effect, with percentage contributions greater than 83.0% from the three methods.

To conclude, we note that ANOVA analysis based on the F-ratio statistic gave the same results for both full factorial and Plackett–Burman methods, and their results were close to those from the Taguchi method. The differences are observed regarding the percentage contributions and, subsequently, the ranks of the factors in affecting the four responses. In that regard, both the FFM and PBM present closer percentage contributions than with those computed from the Taguchi method. This is mainly due to the lowest percentage contribution of error produced by the Taguchi method, since ANOVA analysis was conducted using the measured experimental responses from both the full factorial and the Plackett–Burman methods and the signal-to-noise ratio (SNR) from the Taguchi method.

4.3. Regression Models

In this part of the paper, the resulted mathematical models relative to each method are given. Reduced models after eliminating statistically insignificant factors with respect to ANOVA are shown in the following equations (Equations (21)–(32)). Additionally, we plotted experimental responses versus the predicted ones in order to visualize the agreement between measured values of the responses and the ones calculated from the mathematical models (see Figure 7).

From the ANOVA tables, we note that Taguchi models have higher adjusted R-squared values: 0.982, 0.981, 0.858, and 0.788 for throughput, packet loss ratio, average end to end delay and routing overhead, respectively. The Plackett–Burman method comes in the second rank with adjusted R-squared values between 0.791 and 0.849, whereas, the full factorial method shows quite inferior values (0.633, 0.671, 0.634, 0.681). This states that the Taguchi method has better predictive qualities than the Plackett–Burman and the full factorial methods.

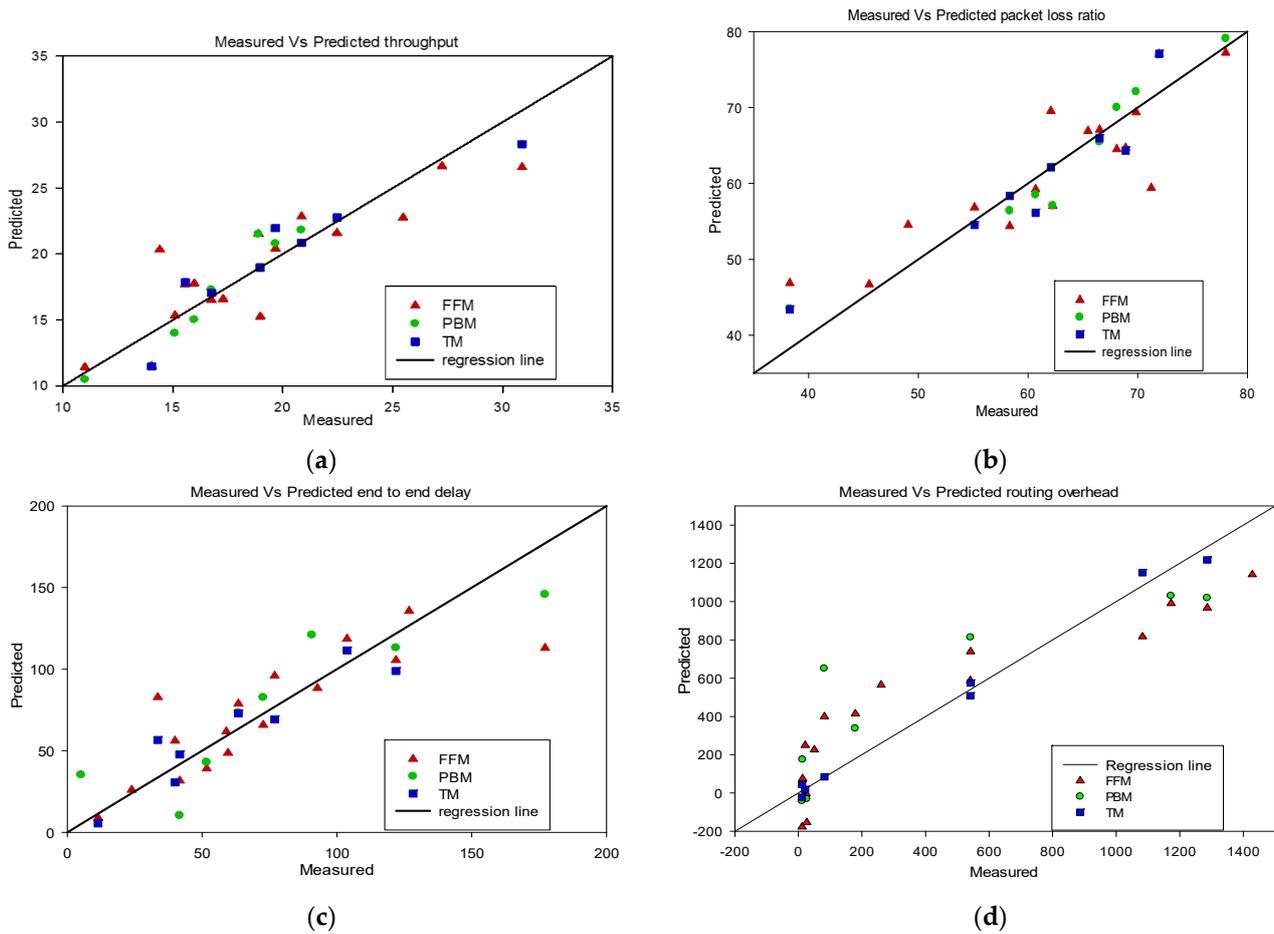


Figure 7. Measured versus predicted (a) throughput, (b) packet loss ratio, (c) average end-to-end delay, (d) routing overhead from FFM, PBM and TM.

Two-level full factorial models

$$\hat{y}_1(TH) = 28.4764 - 0.0384x_1 - 5.0745x_3 - 6.2686x_4 \tag{21}$$

$$\hat{y}_2(PLR) = 43.0880 + 0.0767x_1 + 10.1445x_3 + 12.5325x_4 \tag{22}$$

$$\hat{y}_3(EED) = 2.2868 + 0.3015x_1 + 355.4039x_2 - 22.6936x_3 \tag{23}$$

$$\hat{y}_4(RO) = -591.0738 + 7.4172x_1 + 1086.5367x_2 + 402.6649x_4 \tag{24}$$

Plackett–Burman models

$$\hat{y}_1(TH) = 31.1578 - 0.0499x_1 - 6.8054x_3 - 6.0242x_4 \tag{25}$$

$$\hat{y}_2(PLR) = 37.7360 + 0.0997x_1 + 13.6011x_3 + 12.0410x_4 \tag{26}$$

$$\hat{y}_3(EED) = 26.9985 + 0.1888x_1 + 280.6431x_2 - 42.2375x_3 \tag{27}$$

$$\hat{y}_4(RO) = -296.1545 + 8.3173x_1 \tag{28}$$

The Taguchi models

$$\hat{y}_1(TH) = 31.7435 - 0.0666x_1 - 6.6461x_3 - 5.3924x_4 \tag{29}$$

$$\hat{y}_2(PLR) = 26.4479 + 0.1331x_1 + 20.5388x_2 + 13.2862x_3 + 10.7835x_4 \tag{30}$$

$$\hat{y}_3(EED) = 33.8116 + 0.3171x_1 + 233.2547x_2 - 34.6554x_3 \tag{31}$$

$$\hat{y}_4(RO) = -7.6884 + 5.9135x_1 + 2381.3656x_2 \quad (32)$$

Figure 7 presented below outlines the scatterplots of the predicted values versus the measured values for each response obtained from the three DOE methods. It is observed that the plots of the Taguchi method show stronger fitting between the measured and predicted values, followed by the Plackett–Burman method. However, the models revealed by the two-level full factorial method show moderate adequation.

5. Conclusions

In this research, we investigated the reliability of statistical design of experiments (DOE) methodology in analyzing and modeling the performance of VANET routing protocols, by applying three DOE methods, mainly the full factorial method, the Plackett–Burman method and the Taguchi method. The analysis and interpretation of the results led to the following conclusions.

Methods based on the design of experiments methodology offer powerful mathematical, graphical and statistical techniques that allow accurate retrieval of information with low costs. This conveniently responds to the particular case of VANET simulations often requiring higher processing and execution times.

Both the Plackett–Burman and the Taguchi methods have the advantage of reducing the number of experiments by half in comparison with the full factorial method. Indeed, the effects of four factors (node density, number of connections, black hole and wormhole attacks) on four responses (throughput, packet loss ratio, average end-to-end delay and routing overhead) were simultaneously assessed using only 8 experiments in the case of the Plackett–Burman and Taguchi methods, while 16 experiments were performed for the full factorial method. Consequently, the Taguchi method might be considered a better alternative to the full factorial method when the number of involved factors is large. Additionally, it is important to note that the Taguchi method proposes powerful orthogonal arrays to design the experiments.

The findings indicate that the three methods allow efficient estimates of the most influencing factors. However, the overall trends of the main effects from the Taguchi method are much closer to those of the full factorial method than those obtained from the Plackett–Burman method.

The ANOVA analysis reveals that no noticeable differences are detected between the three methods regarding the factors that are statistically significant. The discrepancies are observed in the ranking of factors based on the percentage contributions. This is due to how the collected data are analyzed since the FFD and the PBM focus only on the measured values of the response, whereas the Taguchi method uses signal-to-noise ratios based on both the mean and the variance of the response.

Regarding the accuracy of the mathematical models, the results of the adjusted R-squared and the fitting plots show that the Taguchi method gives the best descriptive and predictive power followed by the Plackett–Burman and full factorial methods.

Finally, it is essential to mention that these conclusions are relatively valid to our case study, limited by the number of factors considered and the assigned levels. Thus, interested researchers can perform their scientific studies based on their specific conditions and potential limitations. The goal was to demonstrate the practicability of the implementation of some DOE approaches since their application to the performance assessment is less addressed in the literature. As far as we know, we believe that no such comparative study was presented before that applies different DOE methods together for the evaluation and the modeling of VANET routing protocol performance.

6. Future Work

A more comprehensive study is needed to explore other DOE approaches. Other factors should be involved, and their effects can be investigated on other metrics such as computational complexity, energy consumption, collision or the percentage of disconnectivity.

The next step in our work will be the optimization of VANET routing protocols based on statistical approaches to address the scalability and the security issues in VANETs, where a normalized measure of performance putting together multiple performance metrics can be performed.

Author Contributions: Conceptualization, S.A. and S.E.H.; methodology, S.A. and M.-A.E.H.; validation, M.H.; formal analysis, S.A., M.H. and S.E.H.; investigation, S.A. and S.E.H.; resources, S.A. and S.E.H.; data curation, S.A. and S.E.H.; writing—original draft preparation, S.A., M.H. and S.E.H.; writing—review and editing, S.A., M.H. and S.E.H.; visualization, S.A., M.-A.E.H. and S.E.H.; supervision, M.H.; project administration, M.H. and S.E.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The design matrix of the FFM.

N°	ND	NC	BKH	WMH	Y1(TH)	Y2(PLR)	Y3(AE2ED)	Y4(RO)
1	–	–	–	–	30.8856	38.2787	41.6846	11.356
2	+	–	–	–	25.4735	49.0847	58.9551	259.989
3	–	+	–	–	27.2436	45.5328	92.798	23.73
4	+	+	–	–	20.8482	58.3306	103.756	541.77
5	–	–	+	–	18.882	62.2541	11.4101	10.982
6	+	–	+	–	15.5586	68.9071	51.6864	179.034
7	–	+	+	–	22.459	55.1148	72.6775	26.227
8	+	+	+	–	15.9643	68.0929	76.912	541.073
9	–	–	–	+	14.397	71.2295	59.5812	49.842
10	+	–	–	+	16.7487	66.5301	63.4301	1286.464
11	–	+	–	+	19.6663	60.6967	121.866	81.667
12	+	+	–	+	17.2816	65.4645	126.736	1428.075
13	–	–	+	+	18.9702	62.0902	23.8954	13.046
14	+	–	+	+	10.9926	78.0328	39.9311	1082.724
15	–	+	+	+	15.0819	69.8607	33.5335	21.096
16	+	+	+	+	14.0337	71.9508	177.212	1172.566

Table A2. The design matrix of the PBM.

N°	ND	NC	BKH	WMH	Y1(TH)	Y2(PLR)	Y3(AE2ED)	Y4(RO)
1	+1	–1	–1	+1	16.7487	66.5301	63.4301	1286.464
2	+1	+1	–1	–1	20.8482	58.3306	103.756	541.77
3	+1	+1	+1	–1	15.9643	68.0929	76.912	541.073
4	–1	+1	+1	+1	15.0819	69.8607	33.5335	21.096
5	+1	–1	+1	+1	10.9926	78.0328	39.9311	1082.724
6	–1	+1	–1	+1	19.6663	60.6967	121.866	81.667
7	–1	–1	+1	–1	18.8882	62.2541	11.4101	10.982
8	–1	–1	–1	–1	30.8856	38.2787	41.6846	11.356

Table A3. Taguchi's OA L_8 with signal-to-noise ratios (SNR).

N°	ND	NC	BKH	WMH	SNR _{TH}	SNR _{PLR}	SNR _{AE2ED}	SNR _{RO}
1	1	1	1	1	29.795	−31.659	−32.400	−21.105
2	1	1	2	2	25.561	−35.860	−27.566	−22.310
3	1	2	1	1	28.705	−33.166	−39.351	−27.506
4	1	2	2	2	23.569	−36.885	−30.510	−26.484
5	2	1	1	2	24.480	−36.460	−36.046	−62.188
6	2	1	2	1	23.839	−36.765	−34.268	−45.059
7	2	2	1	2	24.752	−36.320	−42.058	−63.095
8	2	2	2	1	24.063	−36.662	−37.720	−54.665

References

- Dafalla, M.E.M.; Mokhtar, R.A.; Saeed, R.A.; Alhumyani, H.; Abdel-Khalek, S.; Khayyat, M. An optimized link state routing protocol for real-time application over Vehicular Ad-hoc Network. *Alex. Eng. J.* **2021**, *61*, 4541–4556. [\[CrossRef\]](#)
- Tripp-Barba, C.; Zaldívar-Colado, A.; Urquiza-Aguilar, L.; Aguilar-Calderón, J.A. Survey on Routing Protocols for Vehicular Ad Hoc Networks Based on Multimetrics. *Electronics* **2019**, *8*, 1177. [\[CrossRef\]](#)
- Czitrom, V. One-Factor-at-a-Time versus Designed Experiments. *Am. Stat.* **1999**, *53*, 126. [\[CrossRef\]](#)
- Jones, R. Design and Analysis of Experiments. *Qual. Reliab. Eng. Int.* **2002**, *18*, 163. [\[CrossRef\]](#)
- Mee, R.W. Introduction to Full Factorial Designs with Two-Level Factors. In *A Comprehensive Guide to Factorial Two-Level Experimentation*, 1st ed.; Mee, R., Ed.; Springer: New York, NY, USA, 2009; pp. 3–26.
- Plackett, R.L.; Burman, J.P. The design of optimum multifactorial experiments. *Biometrika* **1946**, *33*, 305–325. [\[CrossRef\]](#)
- Taguchi, G.; Chowdhury, S.; Wu, Y. *Taguchi's Quality Engineering Handbook*, 1st ed.; ASI Consulting Group, LLC: Livonia, MI, USA, 2005.
- Ajjaj, S.; El Houssaini, S.; Hain, M.; El Houssaini, M.-A. A New Approach for Evaluating the Performance of AODV Routing Protocol in VANETs Using the Plackett-Burman Method. In *Digital Technologies and Applications, Proceedings of International Conference on Digital Technologies and Applications (ICDTA 21), Fez, Morocco, 29–30 January 2021*; Motahhir, S., Bossoufi, B., Eds.; Springer: Cham, Germany; New York, NY, USA, 2021; pp. 1021–1032. [\[CrossRef\]](#)
- Ajjaj, S.; El Houssaini, S.; Hain, M.; El Houssaini, M.-A. Taguchi and ANOVA methods for the performance analysis of VANET routing protocols. In Proceedings of the 2nd international conference on Embedded Systems and Artificial Intelligence (ESAI'21), Fes, Morocco, 1–2 April 2021. *in press*.
- Ajjaj, S.; El Houssaini, M.-A.; Hain, M.; El Houssaini, S. Efficient Performance Evaluation of VANET Routing Protocol Using a 2k Full Factorial Design. In *Ubiquitous Networking, Proceedings of International Symposium on Ubiquitous Networking (UNet 2021), Marrakesh, Morocco, 19–22 May 2021*; Elbiaze, H., Sabir, E., Falcone, F., Sadik, M., Lasaulce, S., Ben Othman, J., Eds.; Springer: Cham, Germany; New York, NY, USA, 2021; pp. 31–41. [\[CrossRef\]](#)
- A Discrete-Event Network Simulator for Internet Systems. Available online: <https://www.nsnam.org/> (accessed on 18 November 2020).
- Krajzewicz, D.; Erdmann, J.; Behrisch, M.; Bieker-Walz, L. Recent Development and Applications of SUMO-Simulation of Urban MObility. *Int. J. Adv. Syst. Meas.* **2012**, *5*, 128–138.
- Documentation-SUMO Documentation. Available online: <https://sumo.dlr.de/docs/index.html> (accessed on 21 September 2021).
- Khoza, E.; Tu, C.; Owolawi, P.A. Comparative Study on Routing Protocols for Vehicular Ad-Hoc Networks (VANETs). In Proceedings of the 2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Durban, South Africa, 6–7 August 2018; pp. 1–6.
- Kandali, K.; Bennis, H. Performance Assessment of AODV, DSR and DSDV in an Urban VANET Scenario. In *Proceedings of International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2018), Tangier, Morocco, 12–14 July 2018*; Ezziyyani, M., Ed.; Springer: Cham, Germany; New York, NY, USA, 2019; Volume 915, pp. 98–109.
- Rehman, M.U.; Ahmed, S.; Khan, S.U.; Begum, S.; Ahmed, S.H. Performance and Execution Evaluation of VANETs Routing Protocols in different Scenarios. *EAI Endorsed Trans. Energy Web* **2018**, *5*, 154458. [\[CrossRef\]](#)
- Sadakale, R.; Bhosale, A.; Ramesh, N.V.K. Performance Analysis of Traffic Types in AODV Routing Protocol for VANETs. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019. [\[CrossRef\]](#)
- Wang, X.; Weng, Y.; Gao, H. A Low-Latency and Energy-Efficient Multimetric Routing Protocol Based on Network Connectivity in VANET Communication. *IEEE Trans. Green Commun. Netw.* **2021**, *5*, 1761–1776. [\[CrossRef\]](#)
- Sindhvani, M.; Singh, C.; Singh, R. Implementation of K-Means Algorithm and Dynamic Routing Protocol in VANET. *Comput. Syst. Sci. Eng.* **2022**, *40*, 455–467. [\[CrossRef\]](#)
- Tyagi, P.; Dembla, D. Performance analysis and implementation of proposed mechanism for detection and prevention of security attacks in routing protocols of vehicular ad-hoc network (VANET). *Egypt. Inform. J.* **2017**, *18*, 133–139. [\[CrossRef\]](#)
- Purohit, K.C.; Dimri, S.C.; Jasola, S. Mitigation and Performance Analysis of Routing Protocols Under Black-Hole Attack in Vehicular Ad-hoc Network (VANET). *Wirel. Pers. Commun.* **2017**, *97*, 5099–5114. [\[CrossRef\]](#)

22. Singh, P.K.; Gupta, R.R.; Nandi, S.K.; Nandi, S. Machine Learning Based Approach to Detect Wormhole Attack in VANETs. In *Proceedings of the Web, Artificial Intelligence and Network Applications, 27–29 March 2019*; Springer: Cham, Germany; New York, NY, USA, 2019; pp. 651–661.
23. Das, S.R.; Belding-Royer, E.M.; Perkins, C.E. Ad Hoc On-Demand Distance Vector (AODV) Routing. Available online: <https://tools.ietf.org/html/rfc3561> (accessed on 20 December 2020).
24. Yates, F. Sir Ronald Fisher and the Design of Experiments. *Biometrika* **1964**, *20*, 307–321. [[CrossRef](#)]
25. Sethuramiah, A.; Kumar, R. Statistics and Experimental Design in Perspective. In *Modeling of Chemical Wear*, 2016 ed.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 129–159. [[CrossRef](#)]
26. Wahid, Z.; Nadir, N. Improvement of One Factor at a Time through Design of Experiments. *World Appl. Sci. J.* **2013**, *21*, 56–61. [[CrossRef](#)]
27. Mishra, R.; Singh, A.; Kumar, R. VANET Security: Issues, Challenges and Solutions. In *Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, Chennai, India, 3–5 March 2016; pp. 1050–1055.
28. Yadolah Dodge. Analysis of Variance. In *The Concise Encyclopedia of Statistics*, 2008 ed.; Springer: New York, NY, USA, 2008; pp. 9–11.
29. Least Squares Method: Definition & Explanation. Available online: <https://www.britannica.com/topic/least-squares-approximation> (accessed on 18 September 2021).
30. Nagelkerke, N.J.D. A note on a general definition of the coefficient of determination. *Biometrika* **1991**, *78*, 691–692. [[CrossRef](#)]