Design a Model-Based on Nonlinear Multiple Regression to Predict the Level of User Satisfaction when Optimizing a Traditional WLAN Using SDWN

Leonel Hernández[a,*], Carlos Eduardo Uc Rios[b], Andri Pranolo[c]

a Faculty of Engineering, Institución Universitaria ITSA, Carrera 45#48-31, Barranquilla, Colombia
b Universidad Internacional Iberoamericana, Campeche, México
c Universitas Ahmad Dahlan, Yogyakarta, Indonesia
Corresponding author: *lhernandezc@itsa.edu.co

Abstract— **Higher education institutions' wireless networks have different roles and network requirements, ranging from educational platforms and informative consultations. Currently, the inefficient use of network resources, poor wireless planning, and other factors, affect having a robust and stable network platform. Different authors have investigated the various strategies for the optimization of wireless infrastructures. Still, most of the cases studied aim to improve traditional performance variables without considering maximizing the level of user satisfaction, which represents a flaw that this research paper hopes to solve through SDWN and a predictive model. The authors will determine an appropriate methodology to estimate the user's level of satisfaction through an algorithm or predictive model based on nonlinear multiple regression supported on network performance variables, making a characterization of the project's environment analyzing the wireless conditions. The investigation phases will follow the life cycle guidelines defined by the Cisco PPDIOO methodology (Prepare, Plan, Design, Implement, Operate, Optimize). As a result, it is expected that the project will be the beginning of academic research that will help create strategies to optimize the WiFi network of any educational institution to maximize user satisfaction. In short, the optimization process provides the network with differentiating factors through a modular design with variable modification of parameters according to the users' requirements and needs.**

Keywords— **Software-Defined Wireless Networks (SDWN); optimization; predictive model; wireless networks; PPDIOO.**

I. INTRODUCTION

Higher Education Institutions have technology platforms that offer students and teachers many services, facilitating the teaching and learning process. Among the services provided to students and teachers, there is total access to the internet, virtual classrooms, consultations in databases and specialized bibliographic resources, access to software and applications on the web, online talks of grades and activities to develop, etc. However, various authors have identified in some local Higher Education Institutions shortcomings in terms of the services at the technological level that is offered to students, teachers, and other members of the institutional community, due to several factors such as lack of resources aimed at improving network infrastructure, unexpected solving day-to-day connectivity problems, improper use of existing resources (traffic saturation, poorly defined security policies, non-use of quality of service for the traffic management network,

inadequate protocol configuration). Although the shortcomings mentioned above have been identified in the academic community, and proposals for improvement have been written, more research and execution are still needed. Other innovative alternatives are pending to strengthen the institutions' technological infrastructures [1]. Network infrastructure migration processes are complex, and SDN represents an excellent alternative to carry out this process [2].

To promote technological growth in Higher Education Institutions, it is necessary to optimize the wireless network platform or infrastructure to efficiently manage the institution's network services and maximize user satisfaction. In the optimization process, different authors have researched and written documentation but without considering user satisfaction. For example, Hernandez *et al.* [3] showed how to determine users' level of satisfaction with a wireless network through a multiple correspondence analysis. Therefore, it is necessary to use a new trend growing for this purpose to

optimize the network. Software-Defined Wireless Networking (SDWN) is an emerging approach based on decoupling radio control functions from the radio data plane through programmatic interfaces, which allows the development of hybrid networks [4].

In the literature, several studies have found that exciting models of optimization of traditional wireless networks (and SDWN) are exposed based on two or a maximum of 3 variables that measure network performance, such as Delay (Delay) and response time [5]. A point that does not exist in these studies is that they do not explain the predictive capacity that determines its reliability. Likewise, SDN allows efficient network traffic and assignment of tasks [6], [7].

The rest of the paper is structured as follows. The second part is the literature review about SDN and the predictive model. Also, the applied research methodology is shown. The third section describes the predictive model's development, leading to the variables' equations, assumptions, and transformation. Finally, the conclusions are exposed, in which the results obtained in the experiments are highlighted, and the future work that can be done from this research is described.

II. MATERIAL AND METHOD

A. SDWN and Predictive Model. A Literature Review.

Implementing a solution based on SDN or software-defined networks is a novel and innovative way to optimize the wireless network infrastructure. SDN enables organizations to accelerate application deployment and distribution by dramatically reducing IT costs through policy-based workflow automation [8]. SDNs converge the management of network services and applications into centralized and scalable coordination platforms that can automate the entire infrastructure's provisioning and configuration. In our academic environment, no research has been done related to this topic. Much less has a solution for network optimization being implemented, either wireless or wired, based on this new approach. [9]. SDN has many applications, such as in vehicular networks [10] and wireless networks in general [11], [12]. Fig. 1 shows the SDWN architecture based on a centralized controller:
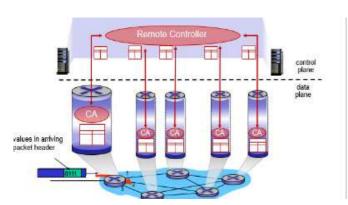

Fig. 1 Centralized controller-based SDN architecture

An example of the SDWN implementation is shown in work done by Sequeira *et al.* [13], in which the performance of an SDWN network based on virtual APs is demonstrated. The authors present a framework for optimizing the channel

management of virtual APs and centralizing all the administration and management of the network in an SDN controller.

To certify that SDWN is an excellent strategy to optimize wireless networks, several investigations have been developed supported by predictive models based on network performance variables. Several research studies have been carried out to describe and explore algorithms or statistical models to optimize network infrastructures considering user satisfaction. In this section, the work developed by Uc-Rios and Lara-Rodriguez [14], whose objective was to create an algorithm that would guarantee the satisfaction of connectivity of the users of a wireless network, seeking an optimal transmission rate for each user, based on the variable use of the channel and its quality of service. A fascinating investigation was developed by Rugelj *et al.* [15]. An experimental study focused on defining a predictive model presented to show the relationship between user perception and user satisfaction concerning technical parameters in data communication services.

The predictive model of the study is based on Markov chains. Cao *et al.* [16] present MCC-SDWN (Mobile Cloud Computing based Software Defined Wireless Network), a model aimed at the fifth generation of wireless networks. The study's main objective is to show the model described as a tool to optimize resource allocation in an SDWN and evaluate the variables delay and packet computing time and bandwidth. [17] The SAQ-2HN (Smart Adaptive QoS for Heterogeneous and Homogeneous Networks) model is defined as an SDN architecture for intelligent, dynamic, and adaptive QoS management in heterogeneous or homogeneous wireless networks. Other studies have shown statistically better values of network performance indicators, such as Delay, jitter, throughput, in SDWN networks than traditional networks based on standard protocols, such as the one performed by Hernandez *et al* [18] and through machine learning techniques [19]. Fig. 2 shows the SDWN topology emulated in Mininet for the study.
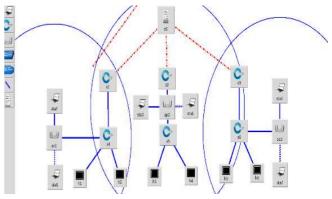

Fig. 2 Emulated SDWN Topology

B. Research Methodology

The guidelines for the development of the research are given by Sampieri *et al.* [20]. First, the study or exploratory research because the project is based on innovation. Therefore, some essential aspects of its development are unknown. Likewise, with the implementation of the exploratory study, it is sought to know strategies and methodologies that have already been used in projects to

optimize data networks through background research. The second type of study or research is applied research because it is proposed to solve a problem in the environment practically, in this case, organizational. Fig. 3 shows the research methodology. The phases of the project are aligned to what is established by the Cisco PPDIOO methodology [21], which are shown in Fig. 4.
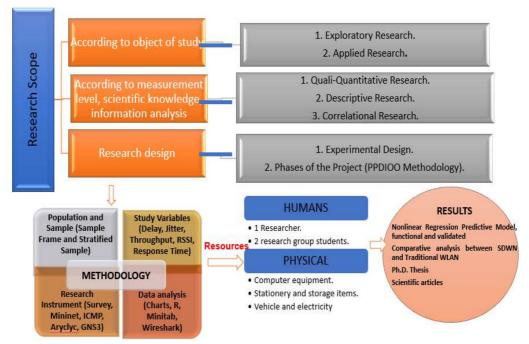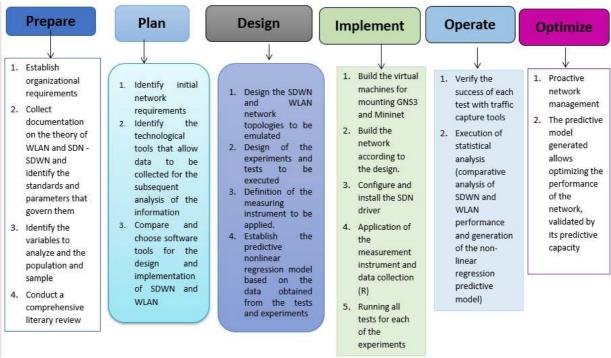

Fig. 3 Research Methodology

Fig. 4 PPDIOO Cisco Methodology

III. RESULTS AND DISCUSSIONS

The model is defined starting with the evaluation of the significance of potentially predictive continuous variables (Jitter, Throughput, RSSI, and packet size) on the response variable "Delay" in the SDWN (Software Defined Wireless Network) scheme. 4 pack sizes were considered (200B, 1000B, 10000B, and 15000B). To evaluate the significance of the variables, their p-values and F statistics were estimated. The variable is declared significant if the p-value is less than the alpha level ($\alpha = 0.05$). Its inclusion within the predictive model is suggested; otherwise, it is considered non-significant, and its use within the regression equation is discarded. The analysis highlights (shown in Table I) only

those statistically significant variables; in the interpretation, those discarded are mentioned along with their p-values. The report also shows the correlation coefficients R^2 and $R^2_{ajustado}$ how the fit of the data provided by the predictive model is evaluated.

Additionally, it presents the R^2_{pred} with which the predictive ability of the said model is verified. The contribution coefficients and T statistics are also listed to determine the response variable's direction and strength of significance. Next, the regression equation is stated through which the relationship between the response variable and the predictor variables is modeled. Said equation would be used to predict new observations of the response variable in the event of a right or excellent predictive ability. Finally, the assumptions of normality, independence, and homoscedasticity of the residuals derived from the model are checked to demonstrate their suitability for later uses. The significance hypothesis allows defining how much a predictor variable affects or impacts the response variable, which for the study, the Delay has been described.

C. Transformation of the Response Variable

The response variable "Delay" (SDWN scheme) did not follow normal behavior, which affected the generation of non-normal residuals. This was verified through an Anderson-Darling test with AD = 1.784 and p-value (<0.005) lower than the alpha level (Fig. 5). Given this adjustment, we proceed with the application of the Johnson transform (see Fig. 5). As a result, a normal data set is generated with AD = 0.328 and p-value (0.510) more significant than the error level. Such data were adjusted to a level Z = 0.7 with the type of transformation SB. The function (Eq. 1) that transforms the response variable "Delay" (Y) into a variable Y 'is shown below:

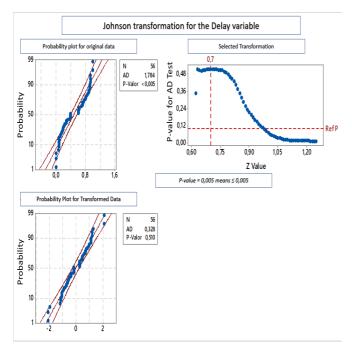$$y' = 0.0437394 + 0.472087 * \ln \frac{y + 0.00934588}{1.01257 - y} \quad (1)$$

Fig. 5 Johnson transformation for the response variable "Delay": data fit before and after modification.

TABLE I
ANALYSIS OF VARIANCE - ANOVA: SIGNIFICANCE OF VARIABLES ON "DELAY" (SDWN)

Source	DF	Adj SS	Adj MS	F-value	P-value
Regression	4	39.704	9.926	53.54	0.000
RSSI (X_1)	1	1.839	1.839	9.92	0.003
Packet Size (X_2)	1	5.652	5.652	30.48	0.000
Transf (Throughput) (X'_3)	1	2.896	2.895	15.62	0.000
Transf2 (Throughput) (X'^2_3)	1	4.333	4.332	23.37	0.000
Error	52	9.641	0.185		
Total	56	49.346			

D. Throughput Transformed Variable

According to the ANOVA, the variable "Throughput" was not considered significant since its p-value (0.235) is less than the level of error (0.05). However, the possibility that this variable's transformation could be incorporated into the model was explored to increase its adjustment capacity and predictive ability. Using nonlinear regression with Weibull growth expectation function, a variable X'_3 was found that responds to the following mathematical expression (Eq. 2).

$$X'_3 = 14.10473 * e^{-0.724689 * X_3^{0.119405}} \quad (2)$$

Most of the variables except for throughput (p-value = 0.235; F = 1.45) and Jitter (p-value = 0.478; F = 0.51) were highly correlated (p-values close to 0) with the variable response "Delay" (SDWN scheme) for which they can be classified as predictors of this variable (shown in Table II). It should be noted that a quadratic term of the Throughput transform (X'^2_3) was also significantly associated with the variable of interest.

E. Data Fit and Predictive Ability

As a next step, we evaluated the fit and the predictive capacity of the predictive model. Given that $R^2 = 80.46\%$ and $R^2_{ajustado} = 78.96\%$, it is concluded that the regression model (Eq. 3) provides a good fit of the variable "delay" (SDWN scheme) based on the set of significant predictors identified in the previous section. It should be noted that, although there is a good fit, 21.04% of the response variable is not explained by the predictors mentioned above. For this reason, future research should point towards identifying and evaluating such variables to increase the fit of the data. Furthermore, the low difference (1.50%) between R^2 and $R^2_{ajustado}$ rules out overfitting effects in the predictive model. This is corroborated by the standard deviation S (0.430595), which turns out to be close to 0. Finally, the regression equation (Eq. 3) allows a correct prediction (p-value = 0.000) of new observations of the variable of response ($R^2_{pred} = 75.76\%$); however, it is suggested to increase it to higher levels with the inclusion of new predictor variables. The list of coefficients and T statistics is provided in Table II:

$$\ln \frac{y + 0.00934588}{1.01257 - y} = $$
$$= \frac{-0.3088X_1 - 0.000172X_2 - 65.6X'_3 + 66.7X'^2_3 - 0.0437394}{0.472087} \quad (3)$$

TABLE II
CONTRIBUTION COEFFICIENTS AND T STATISTICS OF THE PREDICTOR
VARIABLES

Predictor Variable	Contribution Coefficient	T-value
RSSI (X_1)	-0.3088	-3.15
Packet Size (X_2)	-0.000172	-5.52
Transf (Throughput) (X_3')	-65.6	-3.95
Transf2 (Throughput) ($X_3'^2$)	66.7	4.83

Given the contribution coefficients shown in Table II, it can be concluded that the higher the RSSI, the response variable will tend to be lower in the SDWN scheme. A similar case occurs with the packet size but in a smaller proportion (-0.000172).

F. Verification of Assumptions

1) Assumption of normality: part of the following postulate, Ho (Null hypothesis), the residuals conform to a normal distribution, and Ha (Alternative hypothesis): The residuals are not normal.
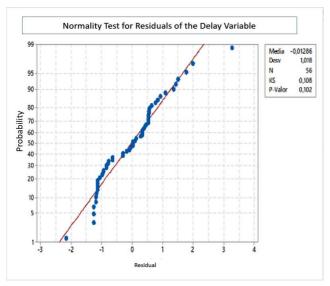

Fig. 6   QQ Plot for normality of residuals

The Kolmogorov-Smirnov test, as shown in Fig. 6, shows with a p-value of 0.102 (more significant than the alpha level), KS = 0.108, and a mean close to 0 (-0.01286), that the residuals of the response variable follow a normal distribution with mean 0 for which the assumption is fulfilled.

2) Assumption of independence: The following postulate: Ho (Null hypothesis): The residuals have a random behavior and Ha (Alternative hypothesis): The residuals do not behave randomly. Fig. 7 shows the behavior of the residuals over time. In this case, no point concentrations are observed on either side of the mean; neither positive nor negative trends are identified; facts by which the fulfillment of the assumption of independence in the residuals is concluded.
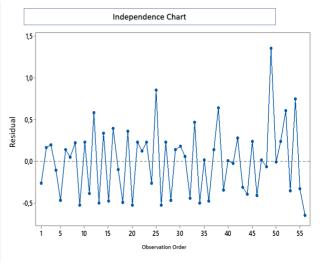

Fig. 7   Independence graph for residuals of the "Delay" transform.

3) Assumption of homoscedasticity: part of the following postulate: Ho (Null hypothesis): The residuals are homoscedastic and Ha (Alternative hypothesis): The residuals are not homoscedastic. Fig. 8 and Fig. 9 show the Bonferroni confidence intervals (95%) for the residuals of the regression model vs. the levels of the variables "RSSI" and "Packet size" correspondingly.
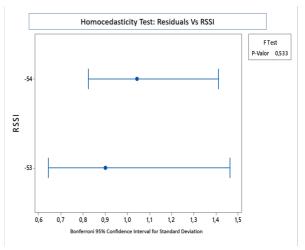

Fig. 8 Homoscedasticity test with Bonferroni intervals (95%) for the relationship "Residuals Vs. RSSI"
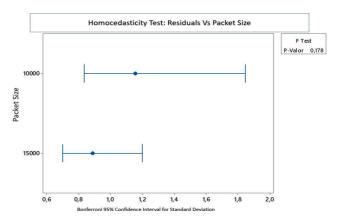

Fig. 9 Homoscedasticity test with Bonferroni intervals (95%) for the relationship "Residuals Vs. Packet size".

In both cases, the confidence intervals can be cut with the same perpendicular line. Thus, their homoscedastic character is concluded. This is confirmed with the p-values (0.533 and 0.178 for "RSSI" and "Packet size," respectively), which turn out to be greater than the alpha level (0.05). Given that the assumptions of normality, independence, and homoscedasticity are fulfilled in the residuals, it is concluded that the nonlinear regression model presented in Eq. 2 is valid and valuable for practical purposes. However, the inclusion of other predictor variables is recommended to increase its adjustment and prediction capacity. The residuals are the random error component (ε) of the response variable (Y). This is expressed according to Eq. 4:

$$Y = \mu + \varepsilon \qquad (4)$$

G. Analysis of Results

The predictive model based on nonlinear multiple regression previously explained it allowed the definition of a

reliable algorithm to maximize user satisfaction, with a predictive capacity greater than 75% (recommended more significant than 70%). The response variable was defined as the Delay or Delay of the network, verified in the different statistical tests explained in the previous sections. In the studies reviewed in the literature review, the Delay is related to the users' level of satisfaction. As stated in the previous point, the data that feeds the model is obtained from an SDWN solution, an acceptable solution to optimize wireless infrastructures. The statistical process showed that the predictive model could be improved by including a more significant number of variables. It should be noted that, although there is a good fit, 21.04% of the response variable is not explained by the predictors mentioned above. Fig. 10 summarizes the construction of the predictive model.
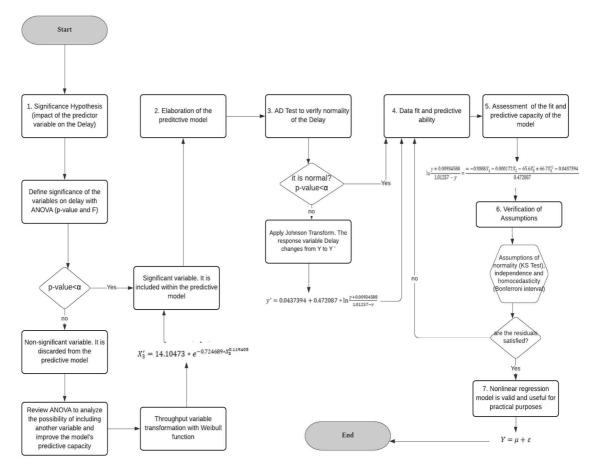

Fig. 10 General Scheme of the Predictive Model Based on Multiple Nonlinear Regression

IV. CONCLUSION

A process of optimizing the wireless network starts from designing and implementing this under international standards. It is necessary to analyze the wireless network state to verify if the design follows the standards and if its performance has considered a series of requirements and protocols needed for its operation. In this research work, variables such as Delay, Jitter, Throughput, response time,

among others, were analyzed to verify the status of the institution's wireless network. This analysis can be replicated to any entity. SDN is an innovative, on-premises approach to wired and wireless network optimization. Although there are several studies and implementations on the subject globally, in Colombia and especially in our city Barranquilla, little has been investigated and executed. SDN allows organizations to grow in their technological infrastructure, optimizing the internal network's behavior in a centralized, scalable, and reliable way.

A nonlinear multiple regression predictive model is built whose main objective is, considering the study variables, the maximization of the level of satisfaction of the users of any wireless infrastructure, taking as the source of the data the values obtained for an emulated architecture SDWN. As future work, the following can be mentioned: in the medium term: implement the SDWN architecture proposed in this document; the predictive model is capable of being improved, increasing its predictive capacity by including a more significant number of variables; the optimization of an SDWN infrastructure as such, analyzing the different existing controllers in the market similar to Zhu *et al* [22], defining policies for quality of service and user experience that are adjusted to technical and institutional goals.

REFERENCES

[1] L. Hernandez, G. Jimenez, and C. Baloco, "Characterization of the Use of the Internet of Things in the Institutions of Higher Education of the City of Barranquilla and Its Metropolitan Area," in *HCI International 2018 – Posters' Extended Abstracts*, 2018, vol. 852, pp. 17–24, doi: 10.1007/978-3-319-92285-0.

[2] Y. Al Mtawa, A. Haque, and H. Lutfiyya, "Migrating from Legacy to Software Defined Networks: A Network Reliability Perspective," *IEEE Trans. Reliab.*, pp. 1–17, 2021, doi: 10.1109/tr.2021.3066526.

[3] L. Hernandez *et al.*, "Optimization of a Wifi wireless network that maximizes the level of satisfaction of users and allows the use of new technological trends in higher education institutions," 2019, doi: 10.1007/978-3-030-21935-2_12.

[4] R. Amin, M. Reisslein, and N. Shah, "Hybrid SDN Networks : A Survey of Existing Approaches," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 4, pp. 3259–3306, 2018, doi: 10.1109/COMST.2018.2837161.

[5] A. Jimenez, J. F. Botero, and J. P. Urrea, "Admission control implementation for qos performance evaluation over SDWN," *2018 IEEE Colomb. Conf. Commun. Comput. COLCOM 2018 - Proc.*, 2018, doi: 10.1109/ColComCon.2018.8466339.

[6] P. B. Madhukrishna Priyadarsini, "Software defined networking architecture, traffic management, security, and placement: A survey," *Comput. Networks*, vol. 192, 2021, doi: https://doi.org/10.1016/j.comnet.2021.108047.

[7] J. -y. B. and B. S. C. Miranda, G. Kaddoum, "Task Allocation Framework For Software-Defined Fog v-RAN," *IEEE Internet Things J.*, 2021, doi: 10.1109/JIOT.2021.3068878.

[8] L. Hernandez and A. Prasetya, "SDN: A Different Approach for the Design and Implementation of Converged NetworksTitle," *2021 3rd East Indones. Conf. Comput. Inf. Technol.*, pp. 450–455, 2021, doi: 10.1109/EIConCIT50028.2021.9431937.

[9] J. F. Kurose and K. W. Ross, *Computer Networking. A Top-Down Approach*, 7th ed. New Jersey, 2017.

[10] T. Mekki, I. Jabri, A. Rachedi, and L. Chaari, "Software-defined networking in vehicular networks: A survey," *Trans. Emerg. Telecommun. Technol.*, 2021, doi: https://doi.org/10.1002/ett.4265.

[11] J. Chen, B. Liu, H. Zhou, Q. Yu, L. Gui, and X. S. Shen, "QoS-Driven Efficient Client Association in High-Density Software-Defined WLAN," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, 2017, doi: 10.1109/TVT.2017.2668066.

[12] B. Gomez, E. Coronado, J. M. Villalon, R. Riggio, and A. Garrido, "WiMCA: multi-indicator client association in software-defined Wi-Fi networks," *Wirel. Networks*, 2021, doi: https://doi.org/10.1007/s11276-021-02636-9.

[13] L. Sequeira, J. L. De La Cruz, J. Ruiz-Mas, J. Saldana, J. Fernandez-Navajas, and J. Almodovar, "Building an SDN enterprise WLAN based on virtual APs," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 374–377, 2017, doi: 10.1109/LCOMM.2016.2623602.

[14] C. E. Uc-Rios and D. Lara-Rodriguez, "A low complexity scheduling for maximizing satisfied users in wireless networks," *4th Int. Conf. Signal Process. Commun. Syst. ICSPCS'2010 - Proc.*, pp. 8–12, 2010, doi: 10.1109/ICSPCS.2010.5709660.

[15] M. Rugelj, M. Volk, U. Sedlar, J. Sterle, and A. Kos, "A novel user satisfaction prediction model for future network provisioning," *Telecommun. Syst.*, vol. 56, no. 3, pp. 417–425, 2014, doi: 10.1007/s11235-013-9853-4.

[16] B. Cao, Y. Li, C. Wang, G. Feng, S. Qin, and Y. Zhou, "Resource Allocation in Software Defined Wireless Networks," *IEEE Netw.*, vol. 31, no. 1, pp. 44–51, 2017, doi: 10.1109/MNET.2016.1500273NM.

[17] A. Khiat, A. Bahnasse, M. E. L. Khaili, and J. Bakkoury, "SAQ-2HN: A Novel SDN-Based Architecture for the Management of Quality of Service in Homogeneous and Heterogeneous Wireless Networks," *Int. J. Comput. Sci. Netw. Secur.*, no. March, 2017.

[18] L. Hernandez, G. Jimenez, A. Pranolo, and C. U. Rios, "Comparative Performance Analysis Between Software-Defined Networks and Conventional IP Networks," *Proc. 5th Int. Sci. Inf. Technol. ICSITech 2019*, 2019.

[19] J. Xie *et al.*, "A Survey of Machine Learning Techniques Applied to Software Defined Networking (SDN): Research Issues and Challenges," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 1, pp. 393–430, 2019, doi: 10.1109/COMST.2018.2866942.

[20] R. Hernandez Sampieri, C. Fernandez Collado, and P. Baptista Lucio, *Metodología de la Investigación Científica*, 6th ed. Mexico D.F., 2014.

[21] P. Oppenheimer, *Top-down Network Design*, 3rd ed. Indianapolis: Cisco Press, 2011.

[22] L. Zhu *et al.*, "SDN Controllers: A Comprehensive Analysis and Performance Evaluation Study," *ACM Comput. Surv.*, 2020, doi: https://doi.org/10.1145/3421764.