Journal of Infrastructure, Policy and Development 2024, 8(15), 8226.
https://doi.org/10.24294/jipd8226

Article

Prediction of telecommunications market behavior based on LSTM models

Diego Armando Giral-Ramirez*, Cesar Augusto Hernández-Suarez*, Luis Fernando Pedraza-Martínez

Universidad Distrital Francisco José de Caldas, Bogota 111611-111611537, Colombia
*** Corresponding authors:** Diego Armando Giral-Ramírez, dagiralr@udistrital.edu.co; Cesar Augusto Hernández-Suarez, cahernandezs@udistrital.edu.co

CITATION

Giral-Ramirez DA, Hernández-Suarez CA, Pedraza-Martínez LF. (2024). Prediction of telecommunications market behavior based on LSTM models. Journal of Infrastructure, Policy and Development. 8(15): 8226. https://doi.org/10.24294/jipd8226

ARTICLE INFO

Received: 26 July 2024
Accepted: 20 August 2024
Available online: 16 December 2024

Abstract: The telecommunications services market faces essential challenges in an increasingly flexible and customer-adaptable environment. Research has highlighted that the monopolization of the spectrum by one operator reduces competition and negatively impacts users and the general dynamics of the sector. This article aims to present a proposal to predict the number of users, the level of traffic, and the operators' income in the telecommunications market using artificial intelligence. Deep Learning (DL) is implemented through a Long-Short Term Memory (LSTM) as a prediction technique. The database used corresponds to the users, revenues, and traffic of 15 network operators obtained from the Communications Regulation Commission of the Republic of Colombia. The ability of LSTMs to handle temporal sequences, long-term dependencies, adaptability to changes, and complex data management makes them an excellent strategy for predicting and forecasting the telecom market. Various works involve LSTM and telecommunications. However, many questions remain in prediction. Various strategies can be proposed, and continued research should focus on providing cognitive engines to address further challenges. MATLAB is used for the design and subsequent implementation. The low Root Mean Squared Error (RMSE) values and the acceptable levels of Mean Absolute Percentage Error (MAPE), especially in an environment characterized by high variability in the number of users, support the conclusion that the implemented model exhibits excellent performance in terms of precision in the prediction process in both open-loop and closed-loop.

Keywords: artificial intelligence; deep learning; long-short term memory; market prediction; telecommunications markets

1. Introduction

1.1. General context

The synergy of information and communications technologies over the last decade has profoundly impacted all aspects of society (Adetunji and Moses, 2022; Qazi and Al-Mhdawi, 2024). Information technologies have allowed greater transparency and administrative efficiency, significantly improving information management (Almaghrabi et al., 2022; Poobalan et al., 2025). This impact has optimized processes and services, democratized access to information, and empowered citizens, reducing inequality gaps and promoting more inclusive development (Almaghrabi et al., 2024; Tiwari et al., 2023).

The telecommunications services market develops in increasingly flexible environments adaptable to customer needs, increasing the volume of data sales and becoming a solid source of income for service providers (Bensalah et al., 2019). This growth has generated fierce competition between suppliers to maintain and attract new customers (Nguyen Chau et al., 2024). Telecommunications companies are investing in improving their infrastructure and diversifying their offerings, including personalized packages and value-added services, to differentiate themselves in a saturated market (Hajar et al., 2024).

Telecom service providers are at the forefront, focusing on continuous innovation and customer service quality to consolidate their market position and respond to changing consumer demands (Agha et al., 2021; Sheth et al., 2020). However, the urgency for governments to establish regulatory policies that allow for fair competition is paramount. These policies ensure that all companies, regardless of size, can compete on equal terms and that consumers receive high-quality services at fair prices (Beckert and Siciliani, 2022). This need for fair competition and consumer protection is not just a matter of policy, but a fundamental aspect of the telecom market's sustainability and growth.

In order to ensure that competition in the supplier market is fair and to reduce structures based on monopolies, it is necessary to promote research that allows governments to implement such regulatory policies effectively (Beckert and Siciliani, 2022; Myers and Tauber, 2011). Research has shown that spectrum monopolization by one operator not only reduces market competition but also has a direct negative impact on users and the overall dynamics of the sector (Jung and Katz, 2022). This impact is not to be underestimated, as it affects the very fabric of the telecom market and the experiences of its users. Therefore, spectrum management strategies should avoid unnecessary spectrum accumulation, seeking to balance the market power of telecommunications services. These elements are essential to improve competitiveness and reduce the digital divide, thus facilitating regional development and identifying possible investments (Czaplewski and Zakrzewska, 2023).

In order to analyze the provider market to reduce monopolies, it is necessary to promote research that allows governments to predict the behavior of the telecommunications market (Eshbayev et al., 2023; Jung and Katz, 2022). Market analysis and prediction are challenging to model due to their scale, multidimensional nature, and complexity. The uncertainties derived from the evolution of demand, prices, and user needs make it necessary to establish solid methodologies to address these challenges (Fisher et al., 2018; Glass and Tardiff, 2023).

Various techniques have been proposed and implemented to predict financial market movement. The most renowned can be classified into four categories: Statistics, machine learning, pattern recognition, and hybrid (Berradi et al., 2021). Artificial Intelligence (AI) is an interdisciplinary field that combines different branches of science, such as statistics, programming, and philosophy. Its primary goal is to design and develop artificial entities capable of solving problems using algorithms inspired by human behavior. The discussion about AI becomes increasingly intricate when it is often used in conjunction with (or even interchangeably) the terms machine learning (ML) and deep learning (DL) (Giral-Ramírez, 2022). This intersection of AI and ML plays a crucial role in predicting the behavior of the telecommunications market, a topic that is sure to pique your interest.

DL is a subset of ML, which itself is a subset of AI. ML is an essential field in IA; it allows systems to learn to perform tasks from experience (i.e., training data). ML focuses on developing algorithms capable of accessing data and using it to learn autonomously. DL is a branch of machine learning that uses artificial neural networks to model and solve problems. There are two main types of DL: Convolutional Neural

Networks (CNN) and Recurrent Neural Networks (RNN). **Figure 1** presents the subset of techniques based on ML and DL (Giral et al., 2021).

An RNN model allows processing and transforming a sequential data input into a specific sequential output. A Long-Short-Term Memory (LSTM) processes input data by forming a loop with time units and updating the state of the RNN. It is an RNN that extends memory to learn from essential experiences that happened long ago. The ability of LSTMs to handle temporal sequences, long-term dependencies, adaptability to changes, and complex data management makes them an excellent strategy for telecom market prediction and forecasting.

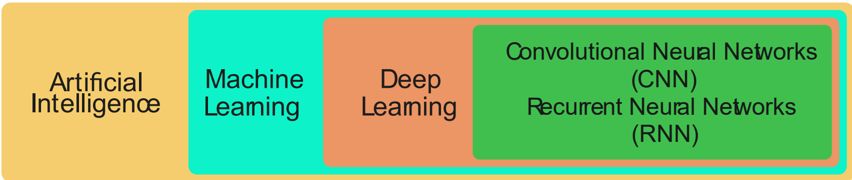

Figure 1. Subset of ML and DL.

Implications for the telecommunications market

Recent studies show a high concentration in the telecommunications market, both in fixed and mobile services, which directly impacts higher prices for users and low competitiveness (Bardey et al., 2020). Prediction techniques for the number of users, traffic level, and operators' revenues in the telecommunications market have direct and tangible applications, helping operators and regulators make informed decisions and adapt to a constantly evolving market environment. These predictions can be applied in real-world situations, such as infrastructure planning and resource enhancement, dynamic tariff management, and regulatory decision-making.

The recommendations from this research study have the potential to significantly impact the Colombian sector. By fostering greater competition in the telecommunications market, these recommendations could lead to increased use of ICT and a subsequent enhancement of the economy's competitiveness. This potential impact is underscored by the alignment of these recommendations with the OECD's findings in its Doing Business report for Colombia (World Bank Group, 2020).

1.2. Contributions and scope

Several questions must be resolved before information can be predicted. Various research projects are underway in different areas of information prediction. Continued work should focus on providing cognitive engines to address more challenges for different applications. This article aims to present a proposal based on artificial intelligence to predict the number of users, the traffic level, and the operators' income in the telecommunications market. The purpose of this prediction proposal is to analyze competition in telecommunications markets.

This paper aims not to propose or develop a new prediction strategy but to apply LSTM networks, which have demonstrated excellent results in various applications, especially in the prediction field. This work uses this proven technique to make predictions in the telecommunications market, providing a reliable and robust approach. Since the approach is based on applying an already validated methodology, no comparisons are made with other prediction strategies. Rather than innovating in the technique, we seek to leverage the effectiveness of LSTM networks to address the specific problem of analyzing the telecommunications market.

DL is implemented through an LSTM network as a prediction technique. The database used corresponds to the number of users, revenue, and traffic of 15 network operators. The data is taken from the Communications Regulation Commission of the Republic of Colombia from 2012 to 2022.

1.3. Limitations

Although the dataset used in this study is based on real data, publicly and freely available and provided by a Colombian state entity, it is essential to point out some inherent limitations. First, the data correspond to a specific region and cover a specific period, which could limit the generalizability of the results obtained. Furthermore, the characteristics of the region and the period analyzed differ from other areas or times, which could influence the applicability of the conclusions in different contexts. By pointing out these limitations, we seek to add transparency and clarity to the results of our research, helping to contextualize the findings better and allowing readers to evaluate how these factors could affect the extrapolation of our results to other scenarios.

1.4. Literature review

Regarding previous works, a search related to LSTM and competition prediction was performed. No study was identified that specifically related the keywords LSTM and Telecommunication Market. Consequently, the works described below correspond to identified applications associated with data prediction using LSTM for telecommunication systems in general. It is essential to highlight that (Abdoli, 2020; Alzheev and Kochkarov, 2020; Latif et al., 2021) support the claim presented in the Contributions and Scope section of this work: LSTM has superior accuracy, making it one of the most precise prediction techniques.

Regarding previous works, no study was identified that specifically related the keywords LSTM and Competition Prediction. Consequently, the works described below correspond to the identified applications associated with data prediction using LSTM for telecommunications systems.

Lu et al. (2024) propose a model based on LSTM and the incremental attention mechanism for chaotic time series prediction. The proposed model is validated using logistic chaotic, Lorenz, and sunspot time series and compared with LSTM, recurrent neural network (RNN), and support vector regression (SVR). The results show that the proposed model is better in terms of prediction accuracy.

In Ding et al.'s research (2022), a hybrid spectral occupancy prediction model is proposed based on a combination of LSTM and autoregressive moving average (ARMA). The performance of the proposed LSTM-ARMA model is validated through experimental data measured by the Tiantong-1 satellite. The results show that the proposed LSTM-ARMA model surpasses the conventional LSTM, the ARMA, and the convolutional neural network in terms of mean absolute error (MAE). Similarly, Wang et al. (2023) propose a hybrid model based on LSTM and ARIMA with the objective of better predicting the number of software users. Through a Bayesian combination model, the weights of each algorithm are determined to develop the proposed LSTM-ARIMA model. The results show that the proposed model has a better level of adaptation when predicting the target variable.

In Latif et al.'s research (2021), the utility of the stock market index as a tool for investors, public companies, and governments in financial decision-making is explored. It focuses on time series prediction to improve the accuracy in forecasting stock market indices, focusing on the Karachi Stock Exchange. The data is transformed into stationary time series using Dicky-Fuller statistics, and techniques such as data decomposition and seasonal ARIMA predictor along with autocorrelation functions are applied to make the predictions.

Alzheev and Kochkarov (2020) implemented the Keras, Theano, and Statsmodels libraries for the stock quotes of several Russian companies. The research aims to identify the most effective time series prediction model by comparing the ARIMA econometric and LSTM models. The results showed that the LSTM model is superior, with a 65% lower RMSE error than ARIMA, making it the most accurate prediction technique.

Abdoli (2020) addresses the complexity of creating prediction models in countries with unstable economies, where fluctuations in historical data are frequent. The paper uses LSTM networks to maintain data integrity when predicting the next two months using the Tehran Stock Exchange (TSE) data. The results show that although the accuracy of the LSTM and ARIMA models decreases in the long run, LSTM offers significantly higher accuracy.

Wang et al. (2020) seek to predict the spectral occupancy of the ISM band through an LSTM model. The proposed model can establish various classification predictors based on the characteristics of Wi-Fi data mining and practical classification algorithms. The results show an improvement in spectrum prediction.

1.5. Organization of the document

This research article is organized into five sections, including the introduction (section 1). Section 2 presents the materials and methods. Section 3 shows the results obtained. Section 4 discusses the results. Section 5 presents the general conclusions obtained.

2. Materials and methods

Figure 2 presents the flowchart for predicting telecommunications markets using the number of users, revenue, and traffic level. The methodology is divided into six stages. The first three stages correspond to the preprocessing and processing of the database for the training and validation of the LSTM network. The fourth stage involves the implementation of the LSTM network architecture, while the fifth stage addresses the open and closed-loop strategies used for data forecasting with the LSTM network. Finally, the sixth stage focuses on the metrics obtained. The following sections describe each of these stages in detail.

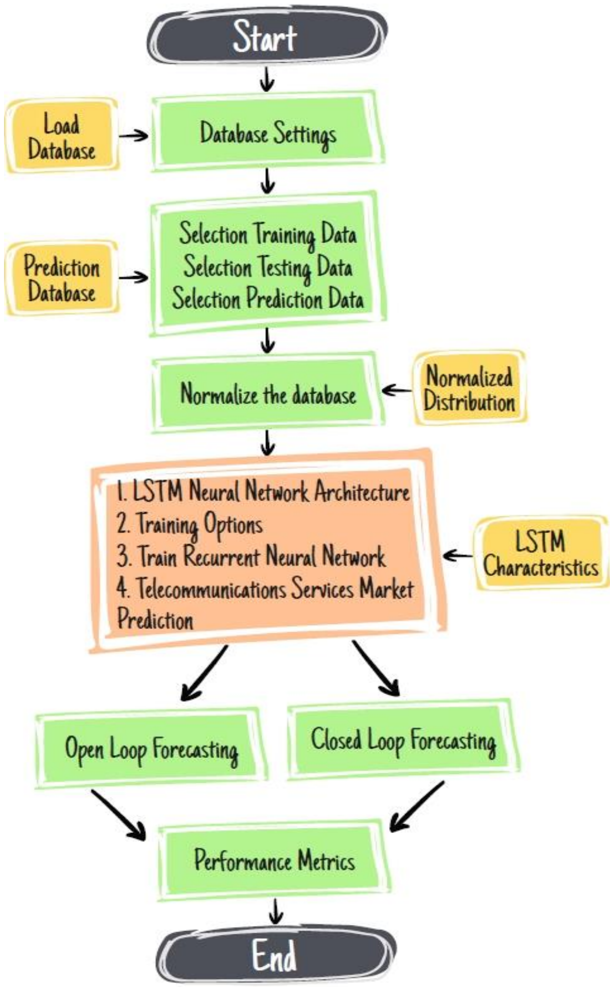

Figure 2. Flowchart implemented methodology.

2.1. Database settings

The database corresponds to the number of users, revenue, and traffic of 15 network operators. The data is taken from the Communications Regulation Commission of the Republic of Colombia for 2012–2023. **Figure 3** presents the characteristics of the database: Comisión de Regulación de Comunicaciones (2024).

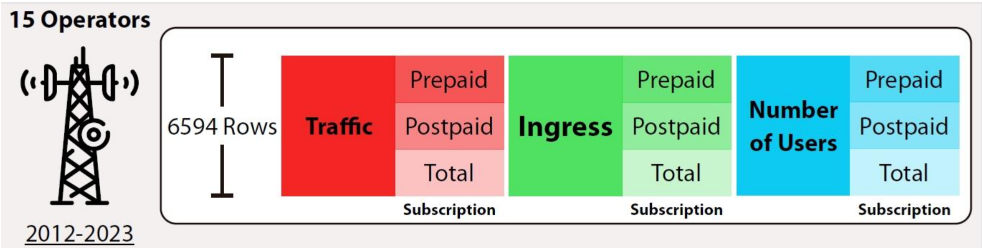

Figure 3. Database characteristics.

2.2. Selection data

The Test-Validation technique, a crucial component of our data analysis process, is used for training, validation, and testing, with a ratio of 70%, 20%, and 10%, respectively. The information we present is based on actual data from the Colombian telecommunications market, ensuring the accuracy and reliability of our findings.

It is important to note that the training, validation, and testing were all carried out concurrently, using data from all fifteen companies. While a common approach involves training a separate network for each company, we opted to exploit a key feature of deep learning: its ability to handle large datasets. As a result, a single network was trained for all fifteen companies.

2.3. Normalize database

The database is normalized to ensure the training process does not diverge and the predictors do not fail. The criterion is based on the normal distribution, with zero mean and unit variance (Equation (1)).

$$X_{\text{Normalized}} = \frac{X_{\text{Database}} - \mu}{\sigma} \quad (1)$$

2.4. LSTM neural network

Figure 4 presents the LSTM architecture. LSTM networks incorporate memory units that allow them to learn when to explicitly forget previous hidden states and when to update hidden states in the face of new information.

Figure 4 illustrates the data in the input gate, forget gate, output gate, memory cell internal state, and hidden state. The state updates satisfy the operations described in Equation (2), where $X_t \in R^{n\times d}$ and the hidden state of the previous time step is $H_{t-1} \in \mathrm{R}^{n\times h}$. Correspondingly, the gates at time step t are defined as follows: the input gate is $I_t \in \mathrm{R}^{n\times h}$, the forget gate is $F_t \in R^{n\times h}$, and the output gate is $O_t \in R^{n\times h}$. W_{xi}, W_{xf}, $W_{xo} \in R^{d\times h}$, W_{hi}, W_{hf}, $W_{ho} \in R^{h\times h}$, $\mathrm{W}_{xc} \in \mathrm{R}^{d\times h}$, and $W_{hc} \in R^{h\times h}$ are weight parameters, and b_i, b_f, b_o, $b_c \in R^{1\times h}$ are bias parameters.

$$\begin{aligned} F_t &= \sigma\left(W_{xf}X_t + W_{hf}H_{t-1} + b_f\right) \\ I_t &= \sigma(W_{xi}X_t + W_{hi}H_{t-1} + b_i) \\ \tilde{C}_t &= \tanh(W_{xc}X_t + W_{hc}H_{t-1} + b_c) \\ O_t &= \sigma(W_{xo}X_t + W_{ho}H_{t-1} + b_o) \\ C_t &= F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \\ H_t &= \tanh(O_t \odot C_t) \end{aligned} \quad (2)$$

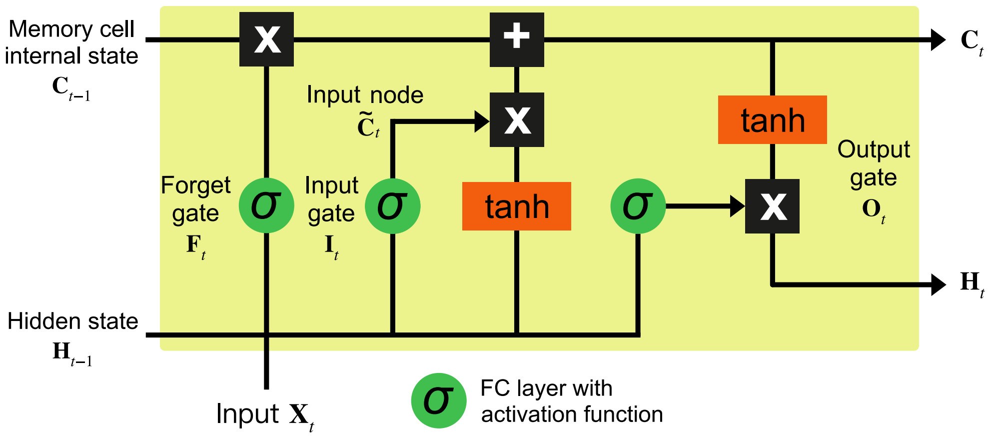

Figure 4. LSTM architecture.

The programming logic for an LSTM network in open-loop or closed-loop prediction is intrinsically linked to the simulator selected for its implementation. MATLAB was chosen as the primary tool in this research because the Universidad Distrital Francisco José de Caldas has an activated campus license. However, it is worth mentioning that there are currently multiple freely accessible tools that also allow these strategies to be implemented efficiently. MATLAB offers a specific LSTM network toolbox that facilitates configuration and adjustment. Algorithm 1 describes the hyperparameters configured for the MATLAB LSTM model and an overview of the implemented architecture. It is essential to highlight that the selection of these parameters was not only based on the literature review but was refined through a rigorous trial and error process. This process involved successive iterations in which critical variables such as the number of neurons, the learning rate, and the layer structure were adjusted. Each setting was evaluated against specific performance metrics, allowing the model to be progressively fine-tuned until reaching a configuration that maximizes its prediction capacity.

Algorithm 1 LSTM architecture implementation

```
numChannels = size (DataBase)
 layers = [
   sequenceInputLayer (numChannels)
   lstmLayer (128)
   fullyConnectedLayer (numChannels)
   regressionLayer]
 options = trainingOptions ("adam", ...
   MaxEpochs = 600, ...
   SequencePaddingDirection = "left", ...
   Shuffle = "every-epoch", ...
   Plots = "training-progress", ...
 ExecutionEnvironment = "gpu", ...
   Verbose = 0)
 net = trainNetwork (XTrain, TTrain, layers, options)
```

2.5. Open and closed loop forecasting

As data scientists, deep learning practitioners, and individuals interested in data prediction techniques, you're understanding and application of the concepts of 'Open and Closed Loop Forecasting' are crucial. These methods involve making predictions based on input parameters, where previous predictions are used as inputs to generate future predictions. Your active engagement in understanding and applying these methodologies is essential (Wen et al., 2024).

The importance of open and closed loop forecasting in deep learning cannot be overstated. The right approach can profoundly impact the model's accuracy, efficiency, and applicability, leading to better data-driven decisions and more reliable results. Each approach has its own set of advantages and disadvantages, and your careful consideration in choosing the right method can make a significant difference in your deep learning projects.

The choice between closed-loop and open-loop forecasting methods is a crucial aspect of predictive modeling, and their applicability varies depending on the context of the system being studied. Closed-loop prediction is a technique in which a deep learning model generates predictions used as inputs to generate future predictions.

Each prediction is fed back to the model to predict the next step in the time series. This method, which incorporates feedback from predictions in subsequent iterations, is particularly advantageous in dynamic systems where conditions change over time, allowing the model to adjust its predictions based on the most recent results continuously. This approach is preferred in scenarios where long-term accuracy is critical, such as industrial process management or autonomous system control (Fander and Yaghoubi, 2022; Gao et al., 2022).

On the other hand, open-loop prediction is a technique in which a deep learning model uses only historical data (not previous predictions) to generate predictions one or more steps into the future. Once the prediction has been made, the model does not reuse these predictions as input data. The open-loop method, which makes predictions without integrating previous outputs into the modeling process, may be more suitable in situations where historical data is stable and future conditions are assumed not to diverge significantly from past ones. This method is frequently used in short-term financial forecasting applications or in scenarios where processing speed is prioritized over accuracy over long cycles (Berret and Jean, 2020; Petropoulos et al., 2022).

2.5.1. Closed-loop forecasting

Since the value $t(n-1)$ is required to predict $t(n)$, and $t(n-1)$ is unknown, a recursive process is applied in which each new prediction is based on the previous one. This process is known as recursive or multi-step forecasting (**Figure 5**).

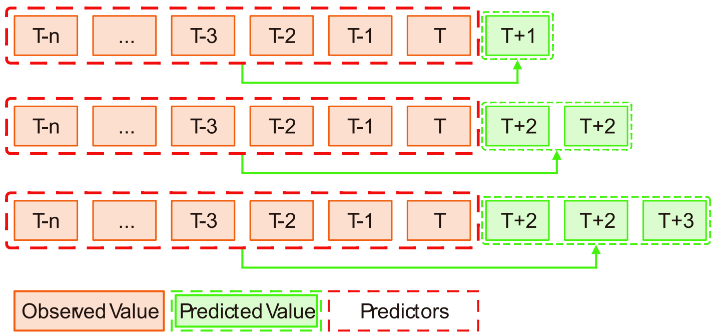

Figure 5. Closed-loop forecasting.

2.5.2. Open-loop forecasting

Direct multi-step forecasting consists of training a different model for each step of the forecast horizon. For example, five models are trained to predict the following five values of a time series, one for each step. As a result, the predictions are independent (**Figure 6**).

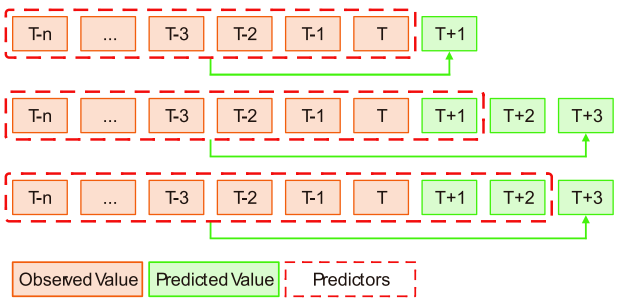

Figure 6. Open-loop forecasting.

2.6. Performance metrics

The errors of the strategies used in the dataset were evaluated using three performance criteria. These are MAPE and RMSE, which are frequently used to analyze model errors (Doğan, 2021). The respective equivalents to obtain the errors are presented in Equations (3) and (4).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x_n)^2} \tag{3}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - x_n}{x_i}\right)^2 \tag{4}$$

3. Results

This section meticulously presents and analyzes the results of predicting the number of users, the level of traffic, and the income of 15 operators in the Colombian telecommunications market. The results are methodically classified into three parts, each addressing a crucial aspect of the prediction process.

In keeping with the scope of this paper, the goal is to show the behavior of the prediction process in the time domain. Given this approach, visualizations that best capture the temporal evolution of predictions and their relationship to observed data were prioritized. Although confusion matrices and heat maps are powerful tools to evaluate a model's performance, especially regarding accuracy and classification errors, their inclusion in this phase would divert attention from the temporal analysis, which is central to this work.

As a simulation tool, R2024a from MATLAB—MathWorks was used under the Microsoft Windows 10 64-bit operating system; the hardware corresponds to a computer with an Intel (R) Core (TM) i7-7700HQ 2.8 GHz processor and 24 GB RAM.

3.1. Database

Figures 7–10 present information on the total number of users, total revenue, and total traffic for the operators MOVISTAR, TIGO, CLARO, ETB, EPM, AVANTEL, ÉXITO, VIRGIN, WOM, SETROC, UFF, CELLVOZ, FLASH, LOV, and SUMA for prepaid, postpaid, and total service (total = prepaid + postpaid). As previously described, the data is taken from the Communications Regulation Commission of the Republic of Colombia.

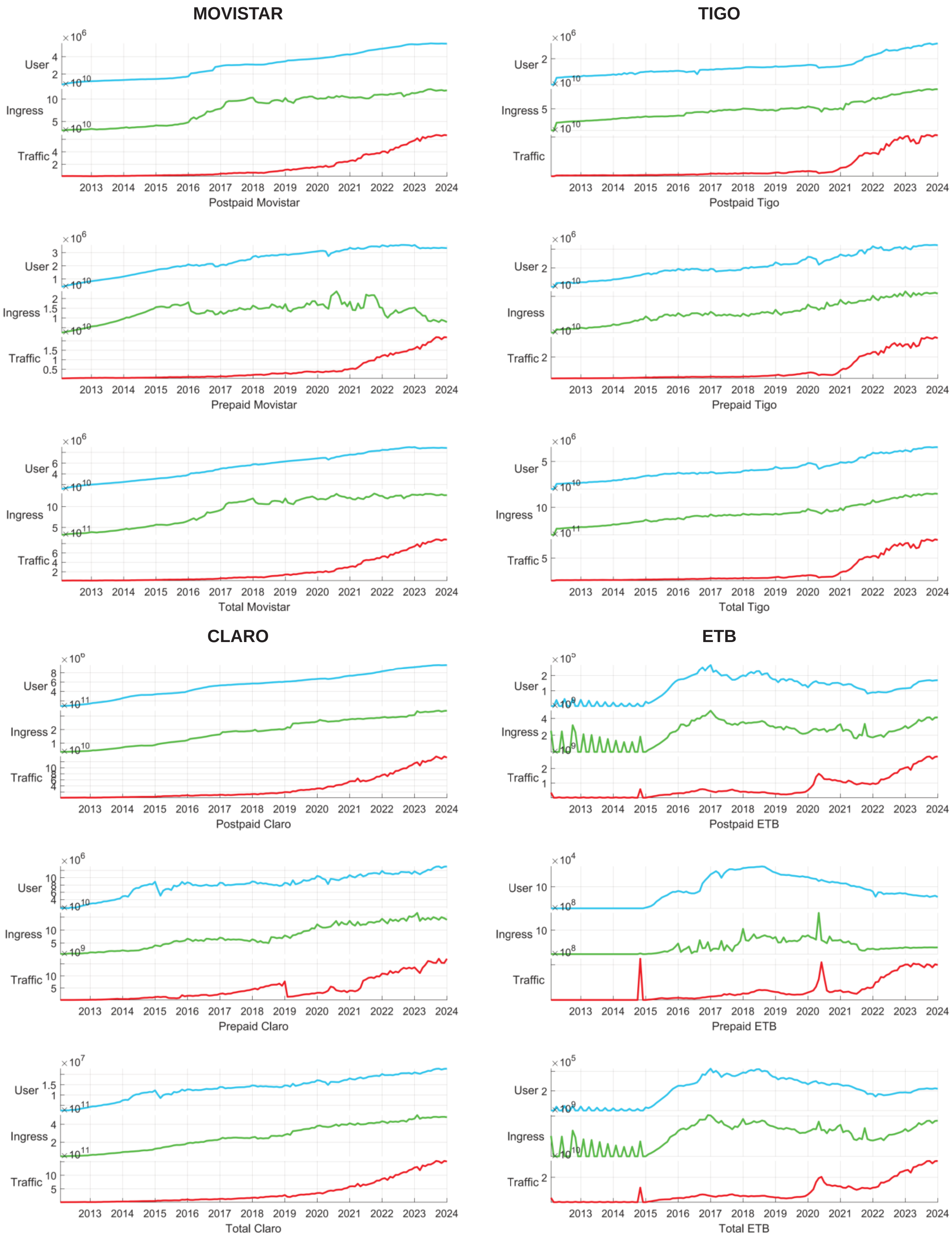

Figure 7. Number of users, total revenue, and total traffic for the operators MOVISTAR, TIGO, CLARO, and ETB.

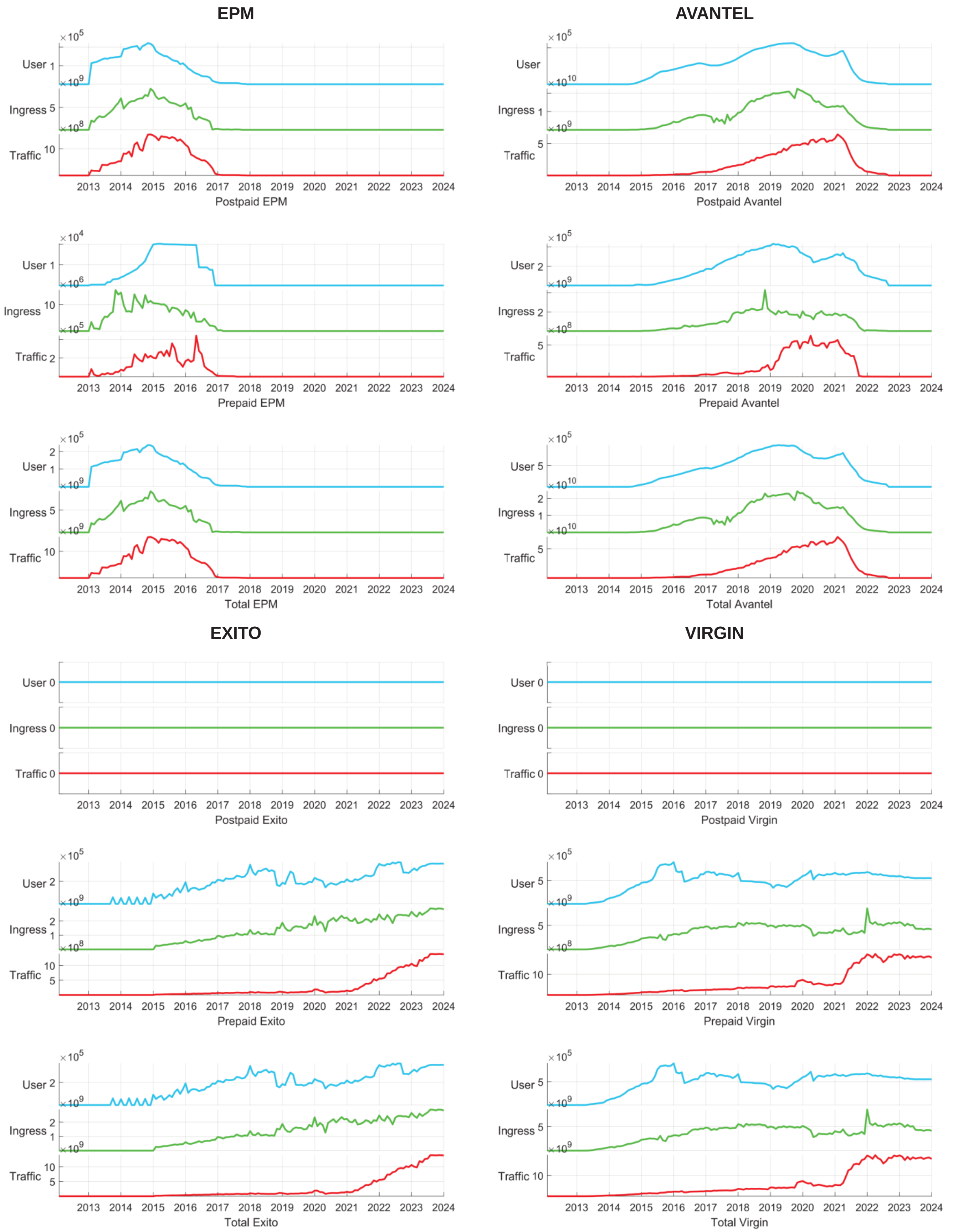

Figure 8. Number of users, total revenue, and total traffic for the operators EPM, AVANTEL, EXITO, and VIRGIN.

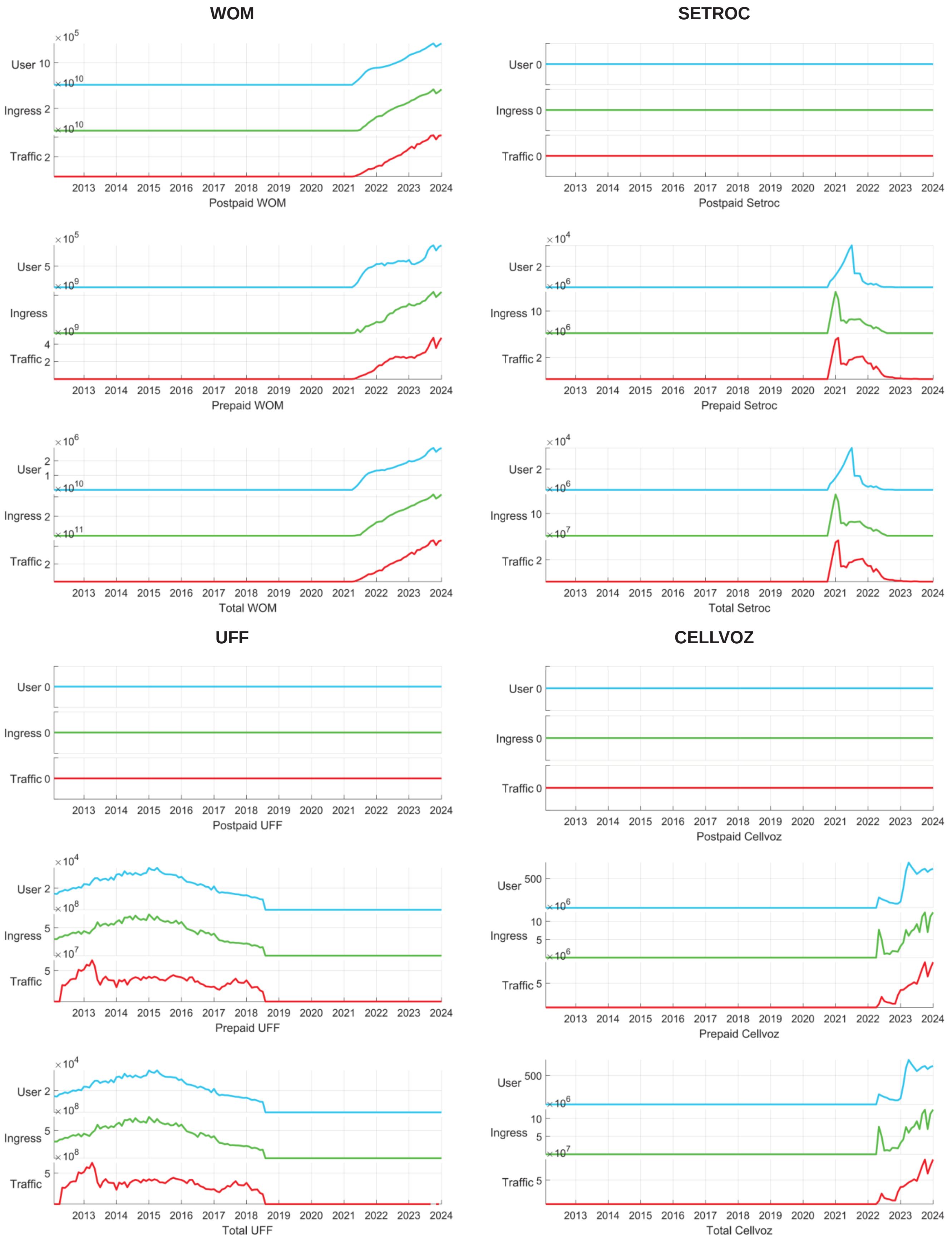

Figure 9. Number of users, total revenue, and total traffic for the operators WOM, SETROC, UFF, and CELLVOZ.

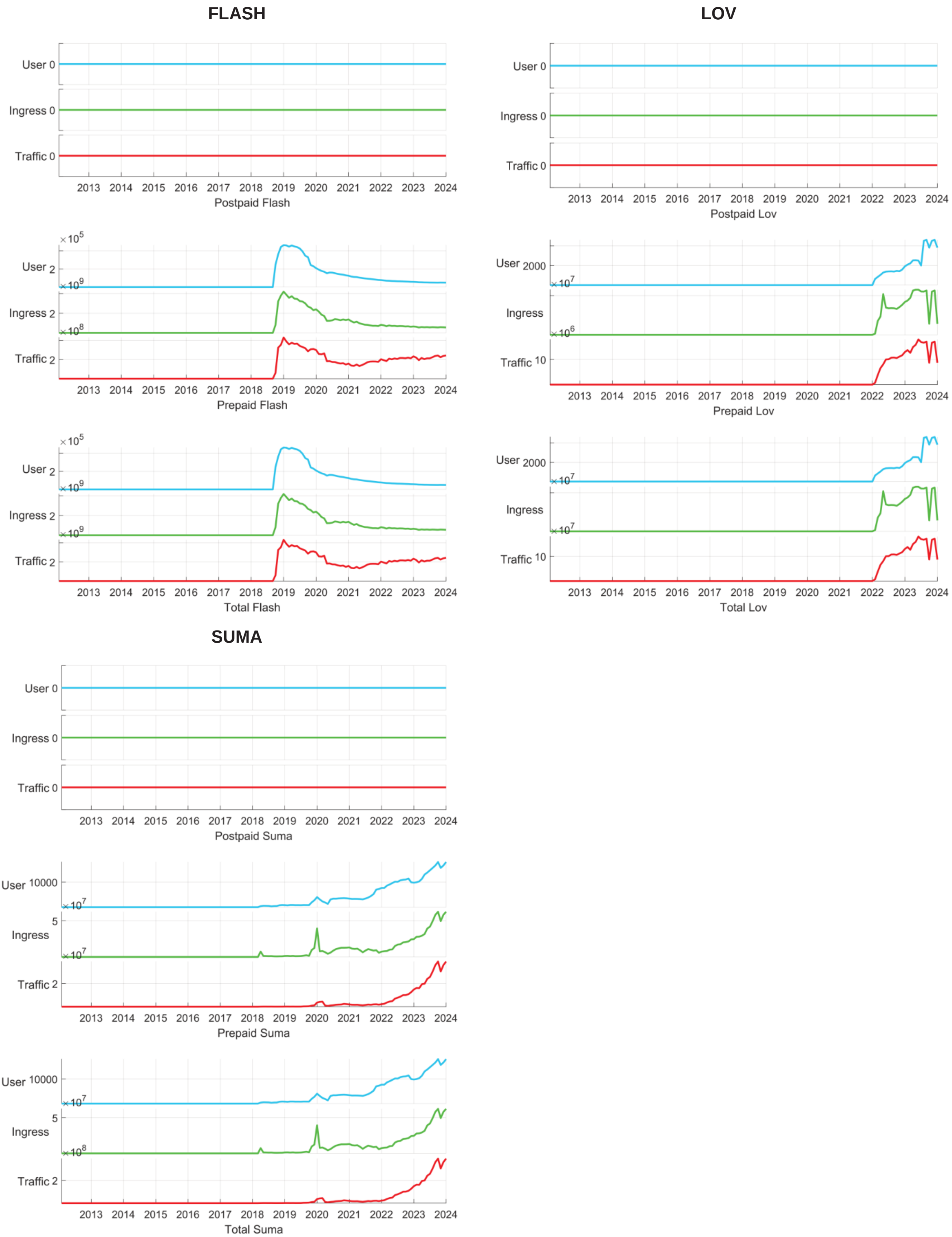

Figure 10. Number of users, total revenue, and total traffic for the operators FLASH, LOV, and SUMA.

3.2. Data prediction using LSTM network in open loop and closed loop

This section presents a comprehensive set of predictions obtained using the LSTM network in open and closed loops. The results cover the total number of users, total revenue, and total traffic of the operators MOVISTAR, TIGO, CLARO, and ETB for prepaid, postpaid, and total services (total = prepaid + postpaid). To ensure a focused presentation, the predictions of the 15 operators are not included to avoid duplication and redundancy.

3.2.1. Prediction using open-loop LSTM network

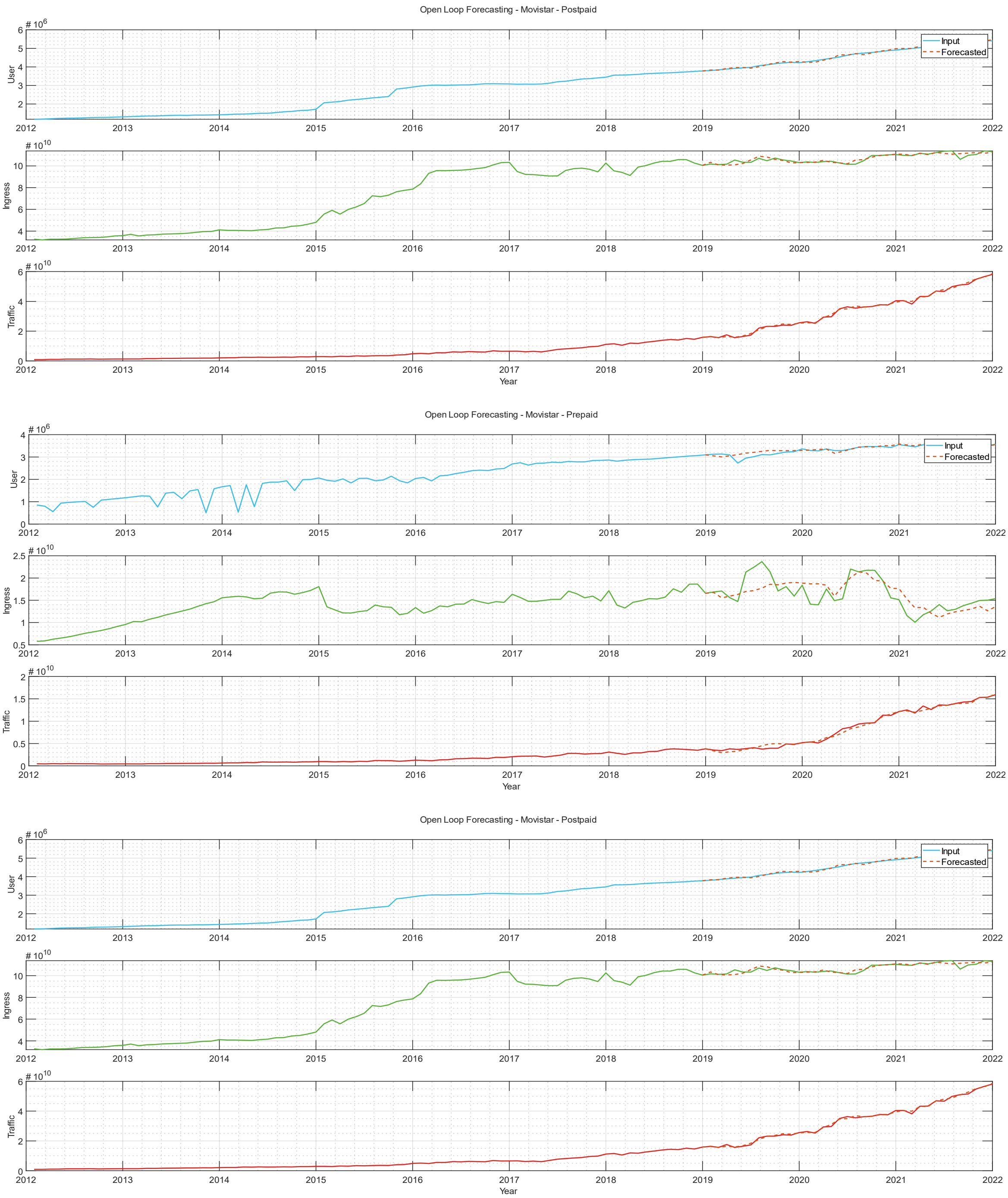

Figure 11. Open-loop prediction for the total number of users, total revenue, and total traffic of the MOVISTAR operator for prepaid, postpaid, and total services.

Figures 11–14 present the results of open-loop prediction for the total number of users, total revenue, total traffic of prepaid and postpaid services, and total operators. MOVISTAR, TIGO, CLARO and ETB. Considering the open-loop characteristics and the cross-validation process, the prediction is made from 2019 to 2022.

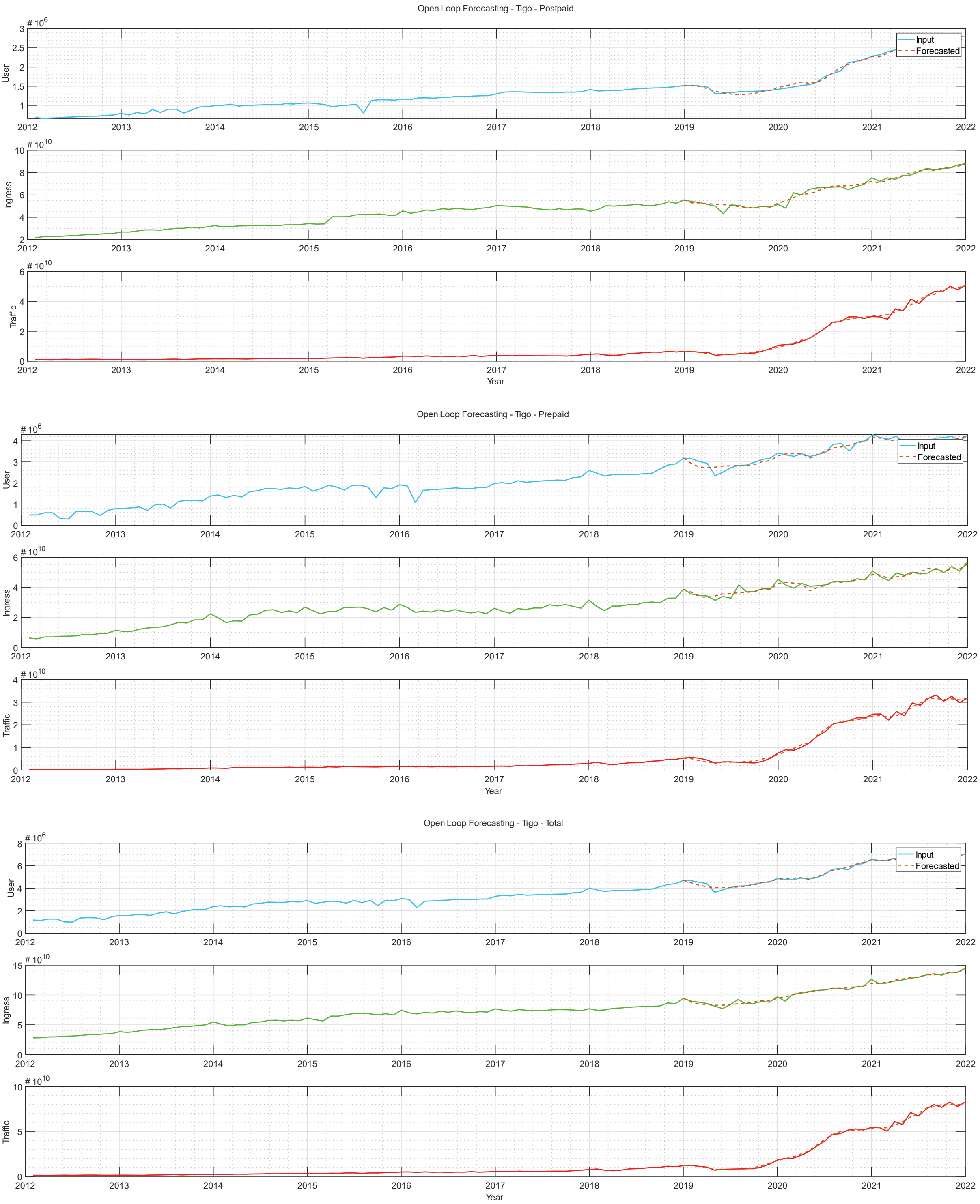

Figure 12. Open-loop prediction for the total number of users, total revenue, and total traffic of the TIGO operator for prepaid, postpaid, and total services.

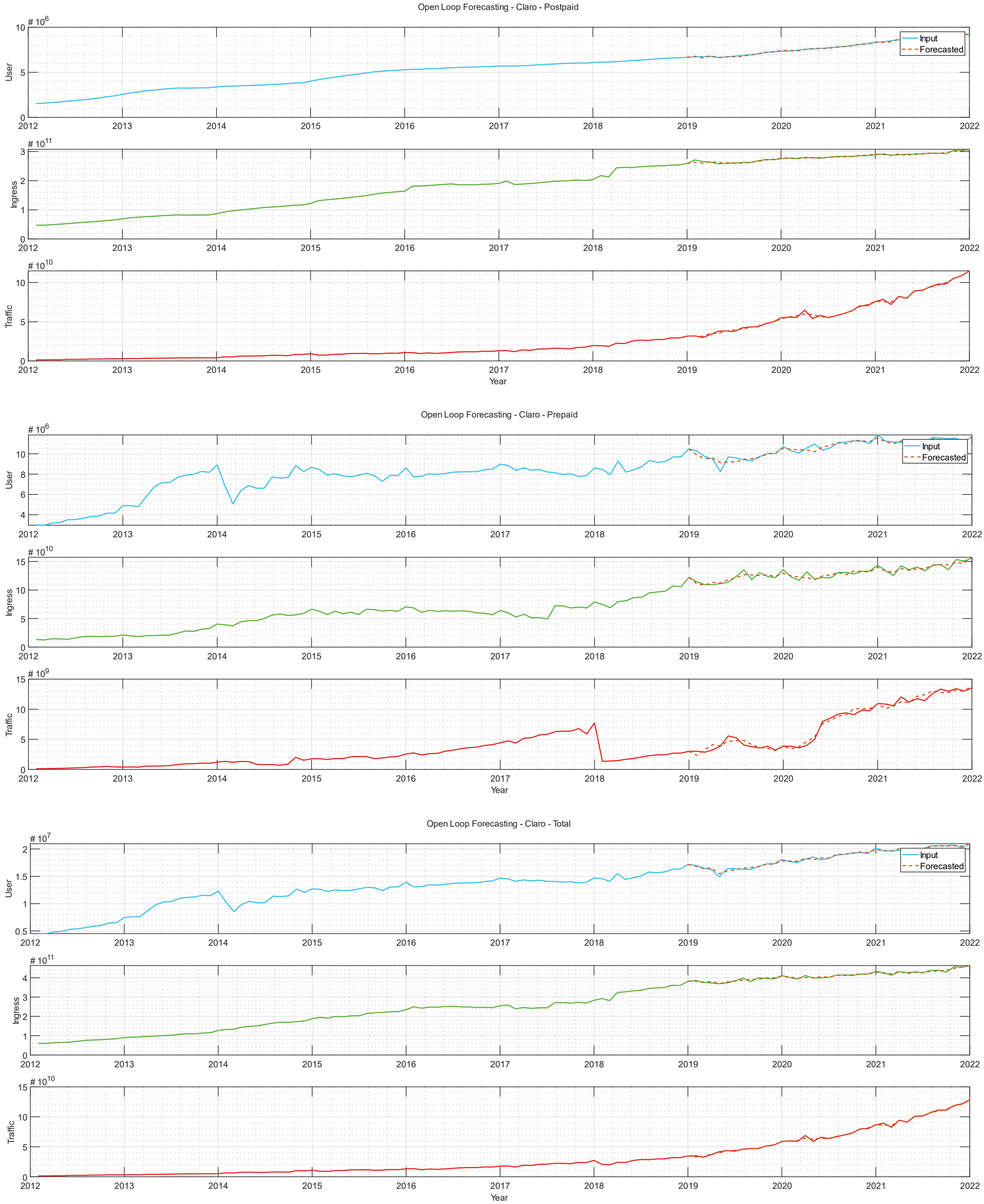

Figure 13. Open-loop prediction for the total number of users, total revenue, and total traffic of the CLARO operator for prepaid, postpaid, and total services.

Figure 14. Open-loop prediction for the total number of users, total revenue, and total traffic of the ETB operator for prepaid, postpaid, and total services.

3.2.2. Prediction using closed-loop LSTM network

Figures 15–18 present the results of closed-loop prediction for the total number of users, total revenue, and total traffic of prepaid and postpaid services and total operators. MOVISTAR, TIGO, CLARO and ETB. Considering the closed-loop characteristics and the cross-validation process, the prediction is made from 2019 to 2022.

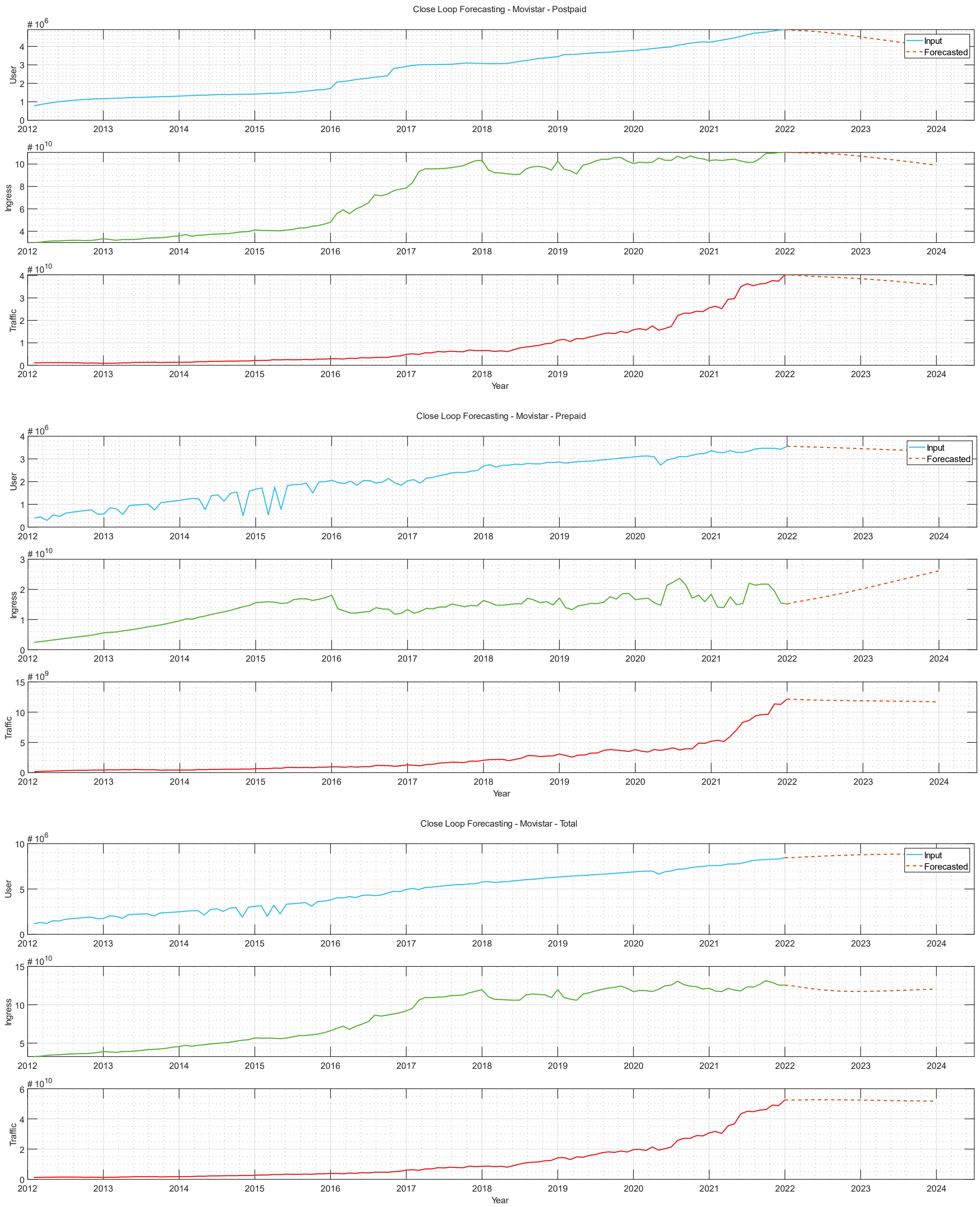

Figure 15. Closed-loop prediction for the total number of users, total revenue, and total traffic of the MOVISTAR operator for prepaid, postpaid, and total services.

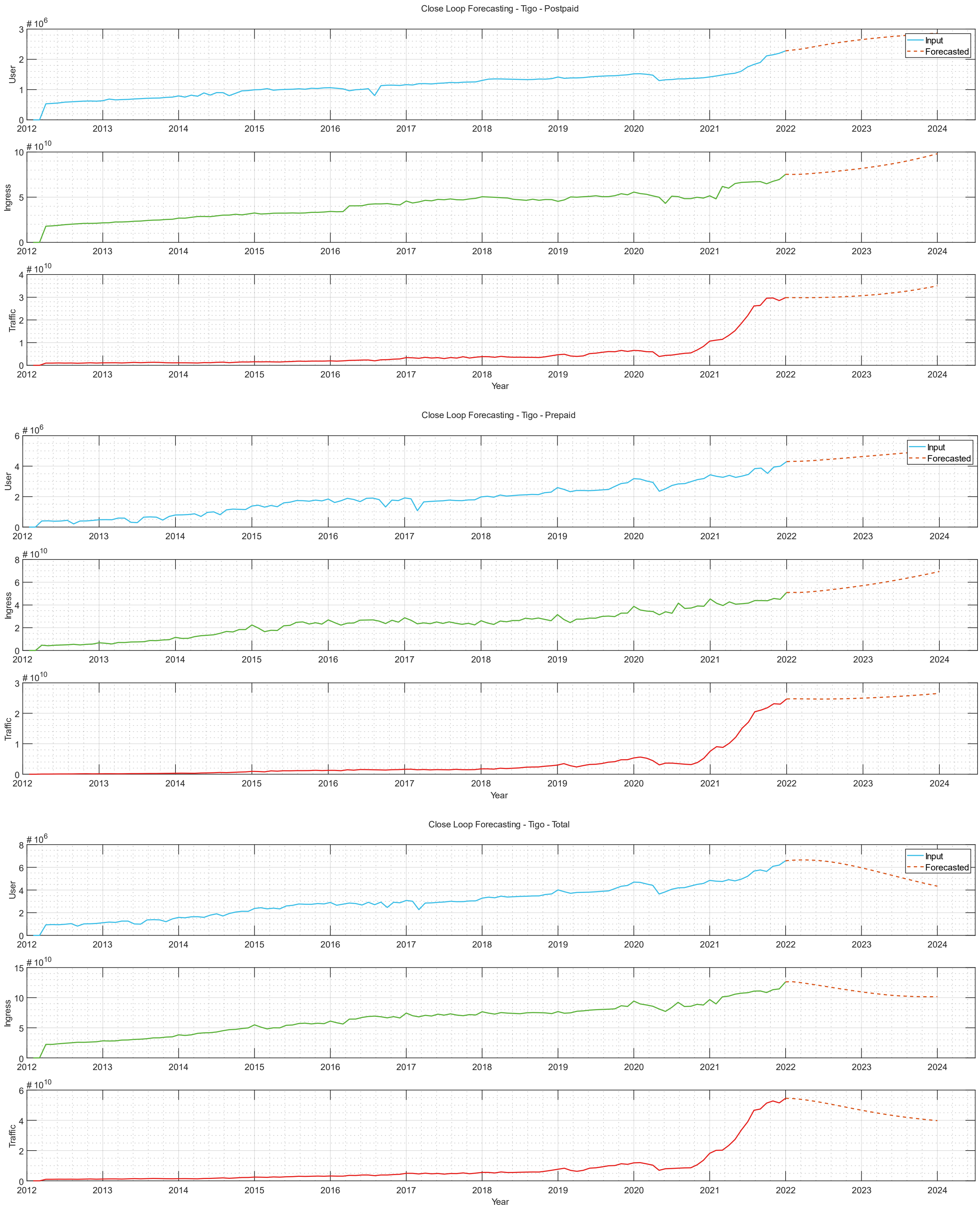

Figure 16. Closed-loop prediction for the total number of users, total revenue, and total traffic of the TIGO operator for prepaid, postpaid, and total services.

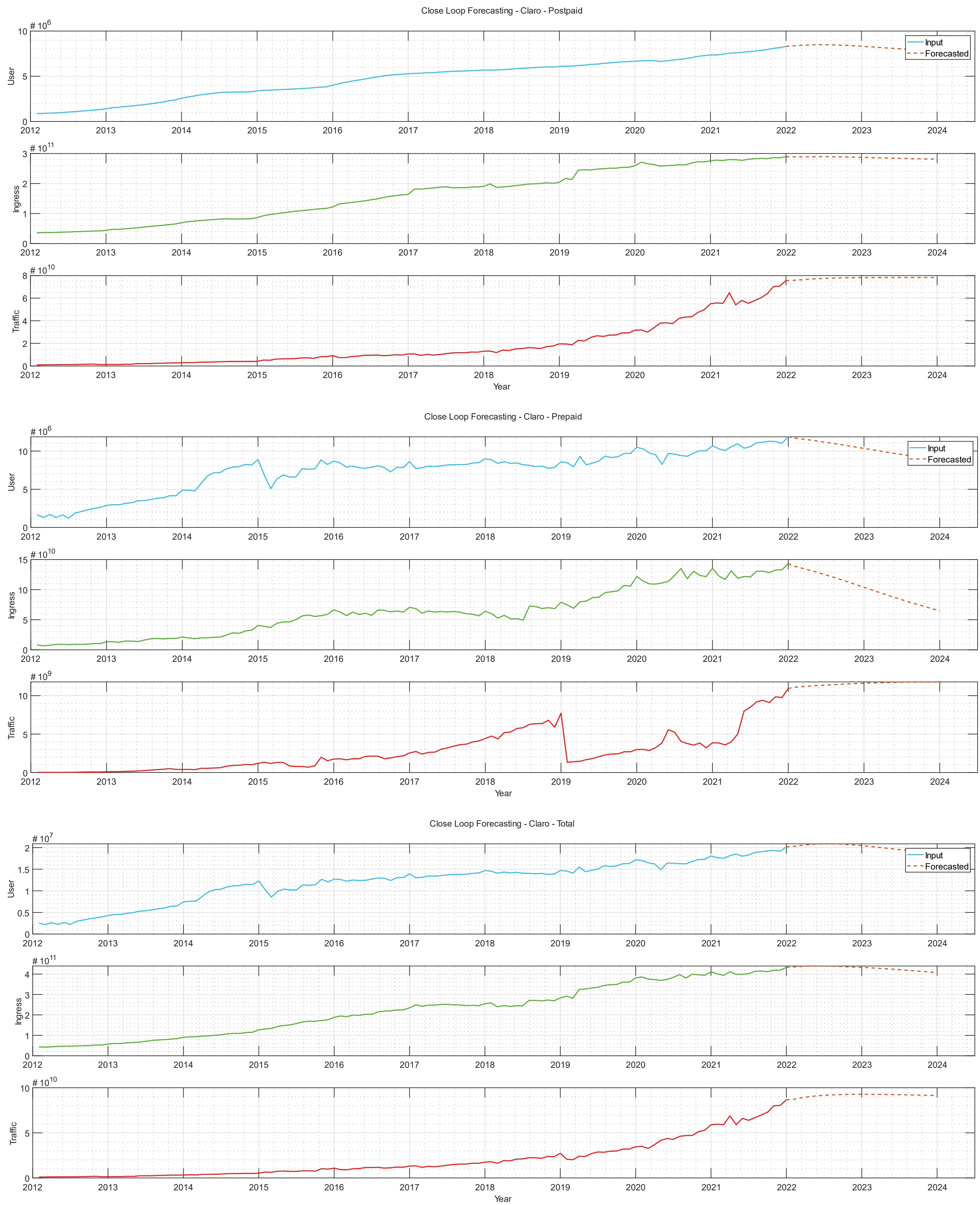

Figure 17. Closed-loop prediction for the total number of users, total revenue, and total traffic of the CLARO operator for prepaid, postpaid, and total services.

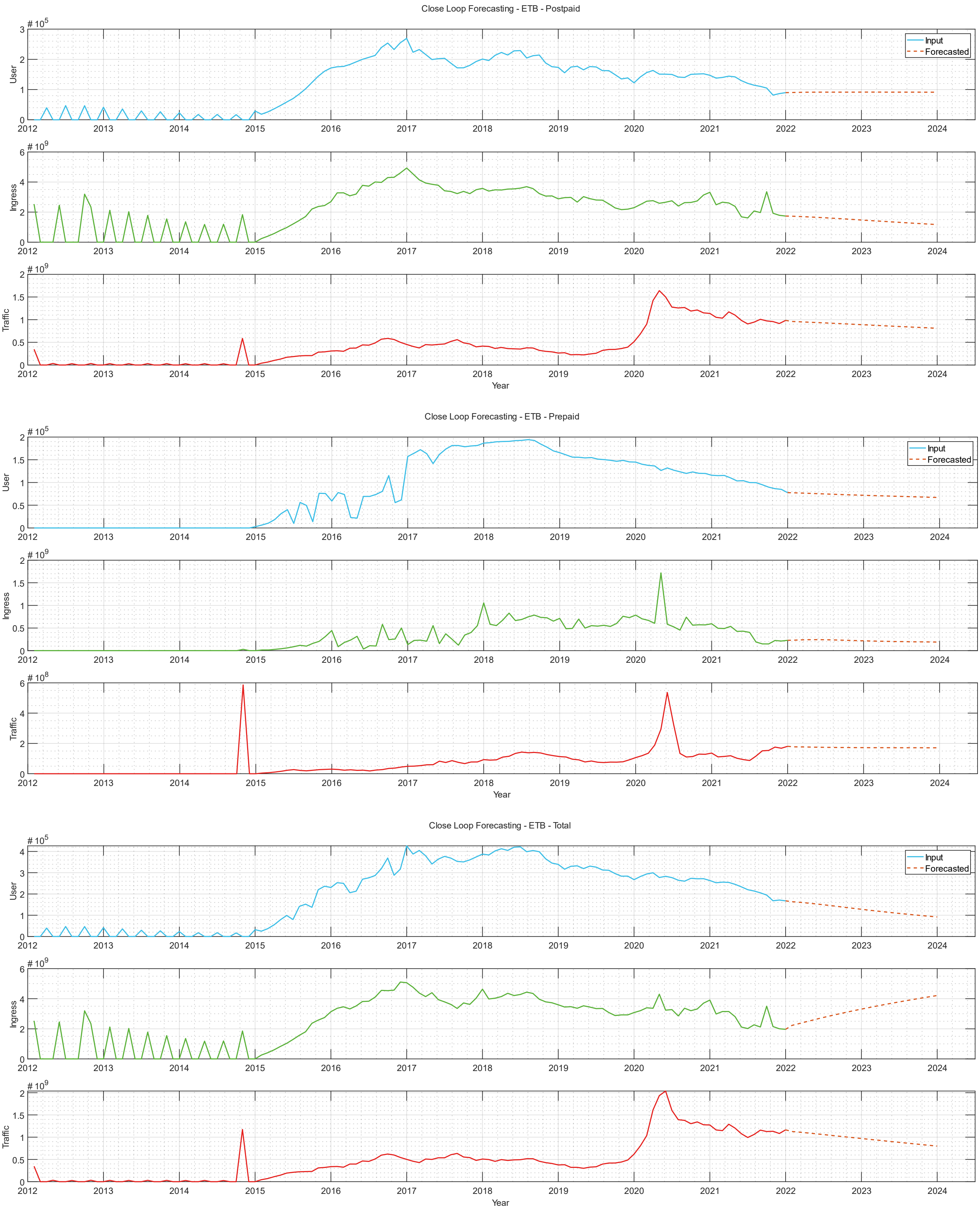

Figure 18. Closed-loop prediction for the total number of users, total revenue, and total traffic of the ETB operator for prepaid, postpaid, and total services.

4. Discussion

This section delves into the results obtained using the LSTM network in both open and closed loops. The analysis covers the total number of users, total revenue, and total traffic of the operators MOVISTAR, TIGO, CLARO, and ETB.

The discussion is divided into two sections. The first section corresponds to the analysis of the results of the open-loop LSTM network. **Figures 19–24** show, through bar charts, the error for prepaid, postpaid, and total services (total = prepaid + postpaid).

The second section thoroughly analyzes the results of the closed-loop LSTM network. **Figures 25–30** are present, and the errors for prepaid, postpaid, and total services are shown in the bar charts (total = prepaid + postpaid).

4.1. Analysis of prediction results using the open-loop LSTM network

For each test sequence, the RMSE between the predictions and the targets is calculated during the training process. **Figures 19**, **21** and **23** present, through a bar diagram, the RMSE in percentage for the postpaid, prepaid, and total open loop service for the operators MOVISTAR, TIGO, CLARO, and ETB. Each bar diagram shows that the RMSE obtained is low, below 0.12%.

RMSE is a metric used to evaluate the accuracy of the predictive model by comparing the predicted values with the actual observed values. A low RMSE suggests that the average difference between the model-predicted and actual values is slight. Therefore, according to the low RMSE values obtained, the implemented model performs excellently in the open-loop prediction process.

MAPE is used to analyze the prediction regarding the evaluation of the model in terms of the number of users, total revenue, and traffic of the network operators MOVISTAR, TIGO, CLARO, and ETB. This statistical indicator, like the RMSE, allows us to evaluate the accuracy of the prediction model. **Figures 20**, **22** and **24** present, through a bar diagram, the MAPE in percentage for the postpaid, prepaid, and total open loop service for the operators MOVISTAR, TIGO, CLARO, and ETB.

In many contexts, a MAPE of 10% or less is considered good, 10% to 20% acceptable, and more than 20% indicates that the model has poor accuracy. For example, a value of 40% indicates that, on average, the model predictions deviate by 40% from the actual values. However, the acceptability of a 40% MAPE depends on the specific context and domain. In some fields with high variability or inherent noise, a MAPE of 40% could be considered tolerable, while in other fields where high precision is required, this value could be unacceptable.

As previously described, the open loop prediction for the postpaid service shows that the prediction of the number of users is the least precise. This is an expected result due to the high variability of users. However, the prediction indicator falls within the acceptable range for revenue and traffic levels, which is of practical importance.

For the prepaid service, the model presented poor predictions for the ETB operator; For the other operators, the prediction is in the acceptable range. As for the total, the model presented poor predictions for the ETB operator and the number of users in the other operators; the other predictions are in the excellent range.

4.1.1. Postpaid service

Figures 19 and **20** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the postpaid service using the open loop of the operators Claro, ETB, Movistar and Tigo.

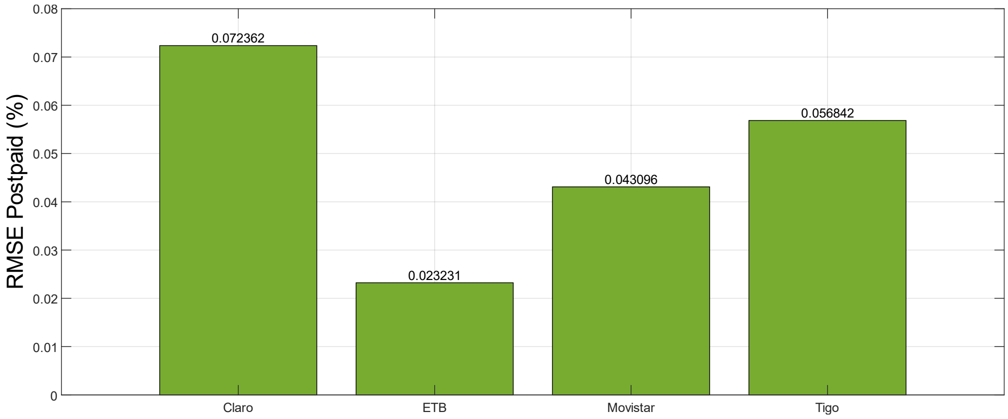

Figure 19. Open-loop RMSE for postpaid service between predictions and targets.

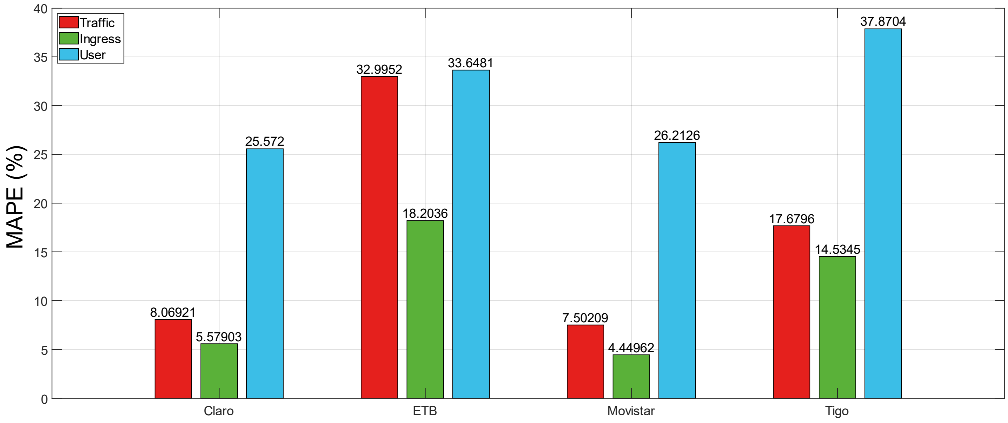

Figure 20. MAPE for open-loop prediction for postpaid service.

4.1.2. Prepaid service

Figures 21 and **22** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the prepaid service using the open loop of the operators Claro, ETB, Movistar and Tigo.

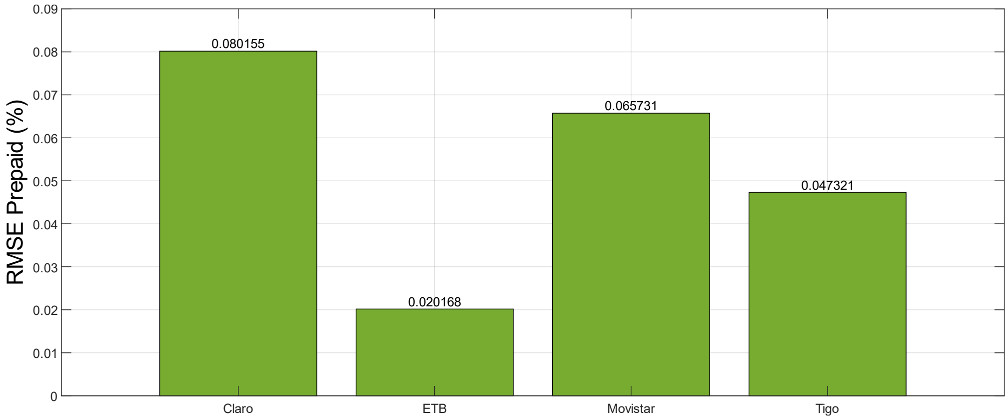

Figure 21. Open-loop RMSE for prepaid services between predictions and targets.

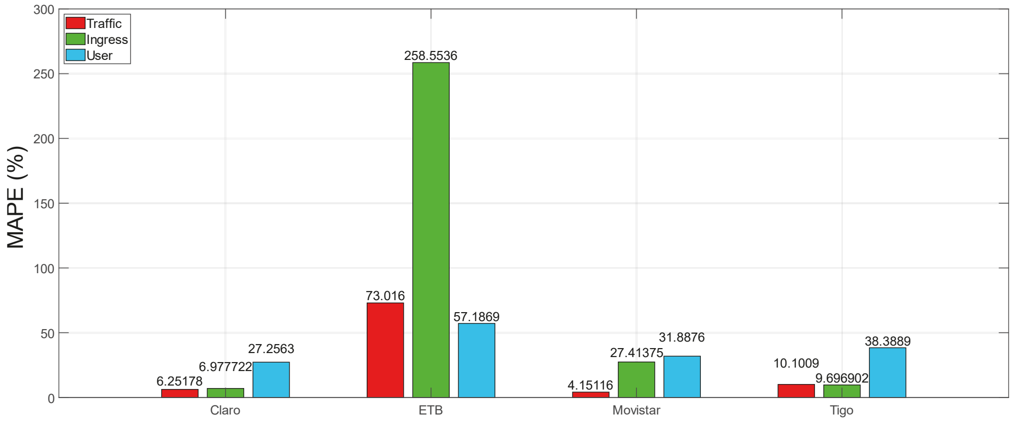

Figure 22. MAPE for open-loop prediction for prepaid service.

4.1.3. Total service

Figures 23 and **24** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the total service using the open loop of the operators Claro, ETB, Movistar and Tigo.

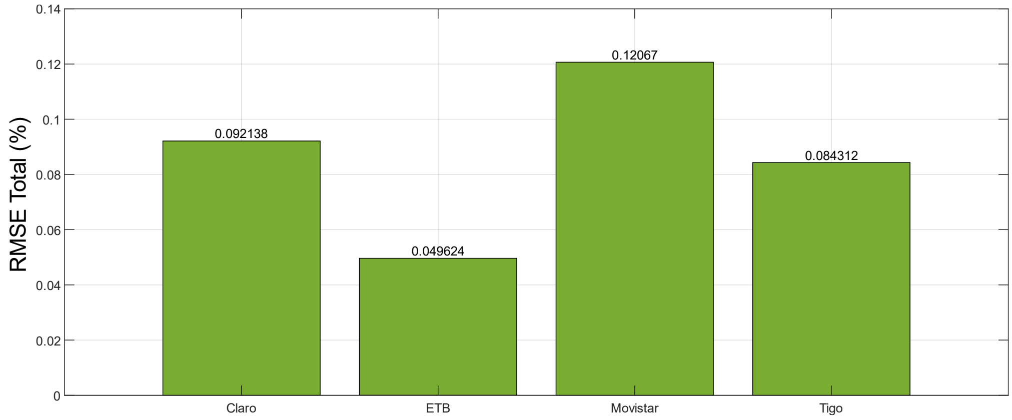

Figure 23. Open-loop RMSE for total service between predictions and targets.

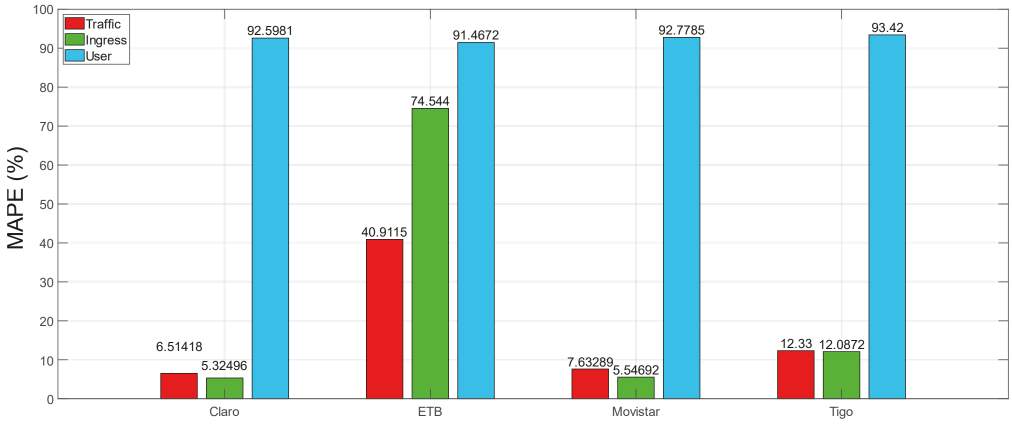

Figure 24. MAPE for open-loop prediction for total service.

4.2. Analysis of prediction results using the closed-loop LSTM network

The RMSE between the predictions and the targets is calculated for each test sequence. **Figures 25**, **27** and **29** present, through a bar diagram, the RMSE in percentage for the postpaid, prepaid, and total closed-loop service for the operators MOVISTAR, TIGO, CLARO, and ETB. As can be seen in each bar chart, the RMSE obtained is low, below 0.28%.

The consistently low RMSE values, all below 0.28%, reaffirm the model's exceptional accuracy in the closed-loop prediction process, bolstering confidence in its performance.

MAPE is used to analyze the prediction regarding the evaluation of the model in terms of the number of users, total revenue, and traffic of the network operators MOVISTAR, TIGO, CLARO, and ETB. This statistical indicator, like the RMSE, allows us to evaluate the accuracy of the prediction model. **Figures 26**, **28** and **30** present, through a bar diagram, the MAPE in percentage for the postpaid, prepaid, and total open loop service for the operators MOVISTAR, TIGO, CLARO, and ETB.

In many contexts, a MAPE of 10% or less is considered good, 10% to 20% acceptable, and more than 20% indicates that the model has poor accuracy. For example, a value of 40% indicates that, on average, the model predictions deviate by 40% from the actual values. However, the acceptability of a 40% MAPE depends on the specific context and domain. In some fields with high variability or inherent noise, a MAPE of 40% could be considered tolerable, while in other fields where high precision is required, this value could be unacceptable.

According to the previously described, the number of users has the lowest precision for the postpaid service's closed-loop prediction. However, this is expected due to the high variability of users. The prediction indicator is in the acceptable range for revenue and traffic levels.

For the prepaid service, the model's predictions are generally accurate, except for the operator MOVISTAR's total revenue and the operator ETB's number of users. Despite these specific instances of poor predictions, the model's overall accuracy is reinforced by its excellent predictions for other factors, particularly the revenue and traffic levels.

4.2.1. Postpaid service

Figures 25 and **26** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the postpaid service using the closed loop of the operators Claro, ETB, Movistar and Tigo.

Figure 25. Closed-loop RMSE for postpaid service between predictions and targets.

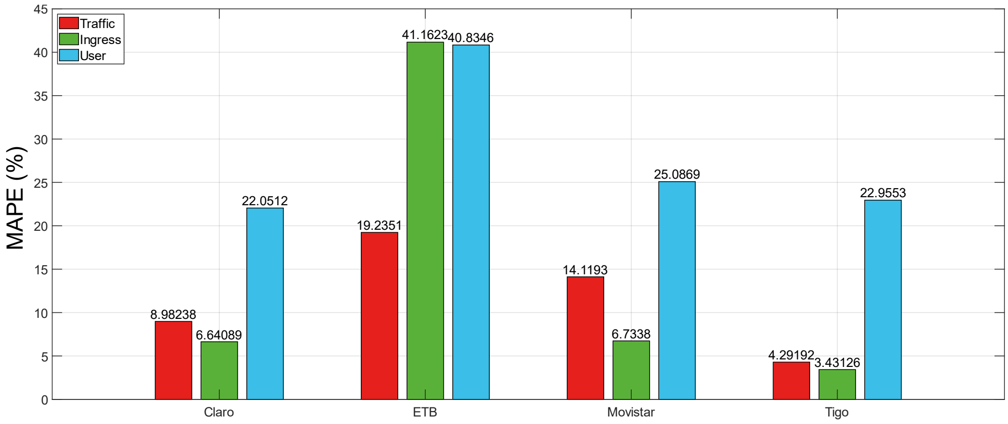

Figure 26. MAPE for open-loop prediction for postpaid service.

4.2.2. Prepaid service

Figures 27 and **28** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the prepaid service using the closed loop of the operators Claro, ETB, Movistar and Tigo.

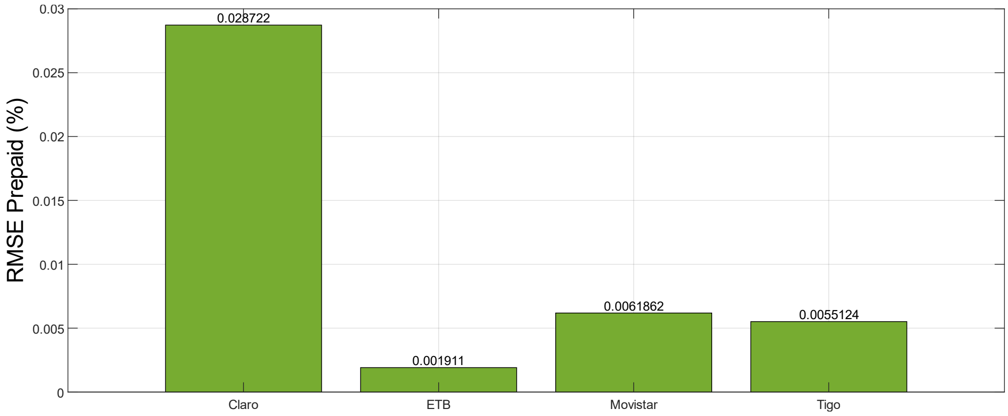

Figure 27. Closed-loop RMSE for prepaid services between predictions and targets.

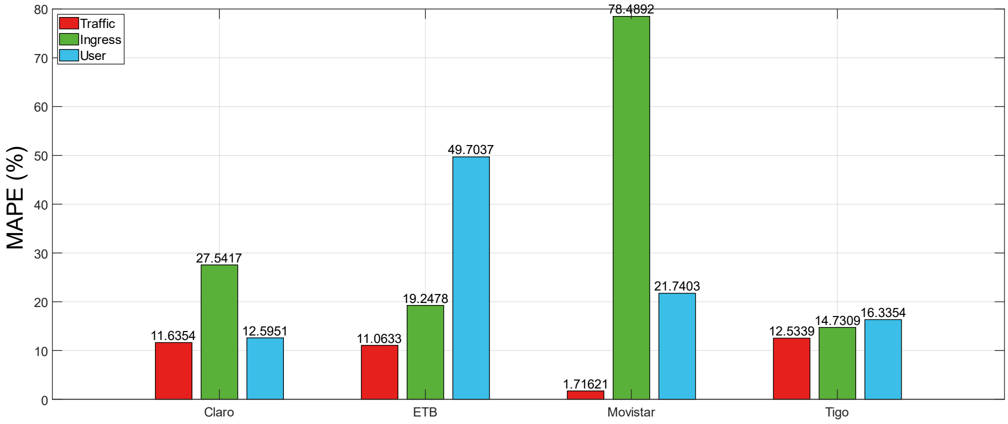

Figure 28. MAPE for open-loop prediction for prepaid service.

4.2.3. Total service

Figures 29 and **30** present, through a bar diagram, the RMSE and the MAPE, respectively, corresponding to the total service using the closed loop of the operators Claro, ETB, Movistar and Tigo.

Figure 29. Closed-loop RMSE for total service between predictions and targets.

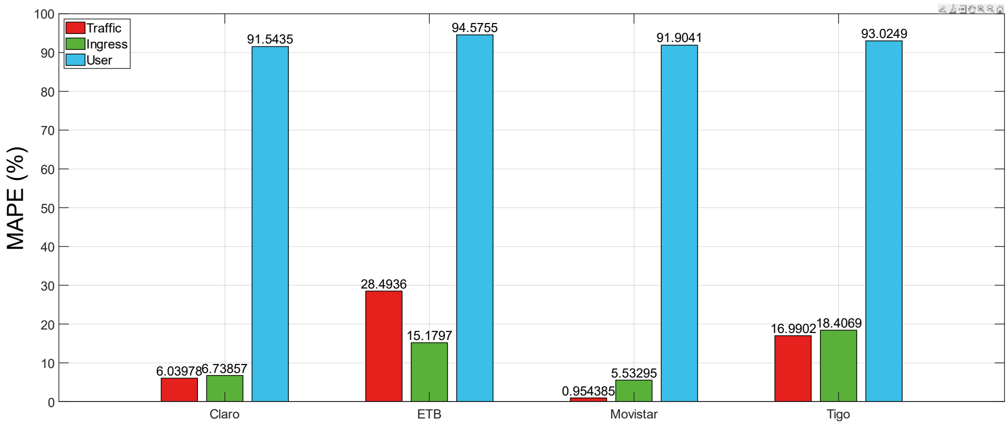

Figure 30. MAPE for open-loop prediction for total service.

As seen in each of the figures in the discussion on open-loop and closed-loop prediction, the lowest prediction accuracy was obtained for the number of users, which indicates that this area deserves more attention. This phenomenon could be due to several reasons.

Based on the description, possible explanations for this behavior will be explored. The first possible explanation is in terms of the data; it is likely that the data used is not sufficiently representative or contains noise that affects its performance. A possible solution would be to incorporate more historical data and other sources to improve the model's generalization.

The second possible explanation is due to the telecommunications market's behavior, which can present complex and non-linear dynamics that the current model fails to capture fully. A possible solution would be to explore the inclusion of additional variables that better reflect the sector's specific trends and patterns.

Finally, the third possible explanation is related to adjusting the model's hyperparameters. In this case, a possible solution would be to consider hybridizing the model with other approaches, such as time series or machine learning methods.

5. Conclusions

This research implemented an open-loop and closed-loop LSTM network to predict the number of users, the level of traffic, and the revenue of operators in the telecommunications market in Colombia. Using the LSTM network, with its ability to capture complex temporal behaviors, was revealed as a highly effective strategy for predicting market behavior. The low values of RMSE and the acceptable levels of MAPE, especially in an environment characterized by high variability in the number of users, support the conclusion that the implemented model exhibits excellent accuracy in terms of accuracy in the prediction process both in loop open and closed loop.

The implementation of strategies based on deep learning, such as LSTM networks, is presented as an invaluable tool for predicting and adjusting to the dynamic fluctuations of the communications market. This perspective not only promises to open new avenues of research but also to significantly transform how companies and researchers address ever-changing market challenges. By integrating advanced predictive capabilities with the ability to adapt quickly, these techniques could provide a robust framework for continued development and innovation in the telecommunications industry.

Future work

While the advances in prediction are promising, they also underscore the urgent need for further research. Future work should utilize advances in artificial intelligence and optimization to achieve even more accurate and reliable results. Additionally, it is crucial to address other prediction error criteria. This involves the development of scalable strategies that can handle more extensive computational loads and effectively solve problems of greater complexity. The importance of this work cannot be overstated in advancing predictive modeling in the telecommunications market in Colombia, and it presents a significant opportunity for growth and development in the field.

Adapting and testing the proposed model based on LSTM networks in other markets or industries is essential to evaluate its robustness in different scenarios. It is necessary to address the prediction inaccuracies identified in this study, proposing improvements to the model that integrate hybridization techniques. Combining LSTM with other prediction methods, such as time series or machine learning techniques, is a proven efficient strategy.

Author contributions: Conceptualization, DAGR, CAHS and LFPM; methodology, DAGR, CAHS and LFPM; software, DAGR, CAHS and LFPM; validation, DAGR, CAHS and LFPM; formal analysis, DAGR, CAHS and LFPM; investigation, DAGR, CAHS and LFPM; writing—original draft preparation, DAGR, CAHS and LFPM; writing—review and editing, DAGR, CAHS and LFPM; visualization, DAGR, CAHS and LFPM; supervision, DAGR, CAHS and LFPM. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science, Technology and Innovation (Minciencias) and the Communications Regulatory Commission (CRC) - Colombia. Under the framework of Contract 80740 082 of 2022 signed between Fiduprevisora and the Universidad Distrital Francisco José de Caldas, which was awarded in Call 908 of 2021 of MinCiencias.

Acknowledgments: The authors thank the Research Office (ODI) of the Universidad

Distrital Francisco José de Caldas, Ministry of Science, Technology and Innovation (Minciencias), and the Communications Regulatory Commission (CRC).

Data availability statement: The database corresponds to the number of users, revenue, and traffic of 15 network operators. The data is taken from the Communications Regulation Commission of the Republic of Colombia. The data set is archived and published at Comisión de Regulación de Comunicaciones (2024).

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

References

Abdoli, G. (2020). Comparing the Prediction Accuracy of LSTM and ARIMA Models for Time-Series with Permanent Fluctuation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3612487

Adetunji, A. J., & Moses, B. O. (2022). The Role of Network Technologies in the Enhancement of the Health, Education, and Energy Sectors. Network and Communication Technologies, 7(1), 39. https://doi.org/10.5539/nct.v7n1p39

Agha, A. A., Rashid, A., Rasheed, R., et al. (2021). Antecedents of customer loyalty at telecom sector. Turkish Online Journal of Qualitative Inquiry, 12(9).

Almaghrabi, H., Li, A., & Soh, B. (2022). IiCE: A Proposed System Based on IoTaaS to Study Administrative Efficiency in Primary Schools. In: IoT as a Service. Springer International Publishing. pp. 121–138.

Almaghrabi, H., Soh, B., & Li, A. (2024). Using ML to Predict User Satisfaction with ICT Technology for Educational Institution Administration. Information, 15(4), 218. https://doi.org/10.3390/info15040218

Al-Sharafi, M. A., Al-Emran M., Tan, G. W.-H., & Ooi, K.-B. (2024). Current and Future Trends on Intelligent Technology Adoption. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-61463-7

Alzheev, A. V., & Kochkarov, R. A. (2020). Comparative Analysis of ARIMA and lSTM Predictive Models: Evidence from Russian Stocks. Finance: Theory and Practice, 24(1), 14–23. https://doi.org/10.26794/2587-5671-2020-24-1-14-23

Bardey, D., Sáenz, B., Aristizábal, D., & Gómez, S. (2020). Impact of the concentration of the mobile market in Colombia on Competitiveness (Spanish). Available online: https://economia.uniandes.edu.co/sites/default/files/eventos/Estudio-movistar.pdf (accessed on 3 May 2024).

Beckert, W., & Siciliani, P. (2022). Protecting Sticky Consumers in Essential Markets. Review of Industrial Organization, 61(3), 247–278. https://doi.org/10.1007/s11151-022-09880-z

Bensalah, F., Bahnasse, A., & El Hamzaoui, M. (2019). Quality of Service Performance Evaluation of Next-Generation Network. In: Proceedings of the 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS). https://doi.org/10.1109/cais.2019.8769576

Berradi, Z., Lazaar, M., Mahboub, O., et al. (2021). A Comprehensive Review of Artificial Intelligence Techniques in Financial Market. In: Proceedings of the 2020 6th IEEE Congress on Information Science and Technology (CiSt). https://doi.org/10.1109/cist49399.2021.9357175

Berret, B., & Jean, F. (2020). Efficient computation of optimal open-loop controls for stochastic systems. Automatica, 115, 108874. https://doi.org/10.1016/j.automatica.2020.108874

Comisión de Regulación de Comunicaciones. (2024). Postscript - Beyond the data (Spanish). Available online: https://postdata.gov.co/search/type/dataset (accessed on 1 July 2024).

Czaplewski, M., & Zakrzewska, M. (2023). Frequency Spectrum Management within the Regulatory Framework of the European Union's Telecommunications Market. European research studies journal, xxvi, (4), 382–390. https://doi.org/10.35808/ersj/3291

Ding, X., Lv, Q., Zou, Y., et al. (2022). Spectrum Prediction for Satellite based Spectrum-Sensing Systems Using Deep Learning. IEEE. https://doi.org/10.1109/globecom48099.2022.10000832

Doğan, E. (2021). Performance analysis of LSTM model with multi-step ahead strategies for a short-term traffic flow prediction. Scientific Journal of Silesian University of Technology. Series Transport, 111, 15–31. https://doi.org/10.20858/sjsutst.2021.111.2

Eshbayev, O., Rakhimova, S., Mirzaliev, S., et al. (2022). A systematic mapping study of effective regulations and policies against digital monopolies: visualizing the recent status of anti-monopoly research areas in the digital economy. In:

Proceedings of the 6th International Conference on Future Networks & Distributed Systems. https://doi.org/10.1145/3584202.3584205

Fander, A., & Yaghoubi, S. (2022). Dynamic and stochastic modeling for a closed-loop automotive supply chain under fuel issue and government intervention: A case study. Computers & Industrial Engineering, 174, 108765. https://doi.org/10.1016/j.cie.2022.108765

Fisher, M., Gallino, S., & Li, J. (2018). Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments. Management Science, 64(6), 2496–2514. https://doi.org/10.1287/mnsc.2017.2753

Gao, N., Qu, L.-C., & Jiang, Y.-T. (2022). A System Dynamic Model of Closed-Loop Supply Chain considering Recovery Strengthening and Product Differentiation. Mathematical Problems in Engineering, 2022, 1–15. https://doi.org/10.1155/2022/3778597

Giral, D., Hernández, C., & Salgado, C. (2021). Spectral decision in cognitive radio networks based on deep learning. Expert Systems with Applications, 180, 115080. https://doi.org/10.1016/j.eswa.2021.115080

Giral-Ramírez, D. (2022). Intelligent Fault Location Algorithms for Distributed Generation Distribution Networks: A Review. Przegląd Elektrotechniczny, 1(7), 139–146. https://doi.org/10.15199/48.2022.07.23

Glass, V., & Tardiff, T. (2023). Analyzing Competition in the Online Economy. The Antitrust Bulletin, 68(2), 167–190. https://doi.org/10.1177/0003603x231163001

Hajar, M. A., Al-Sharafi, M. A., Ibrahim, N., et al. (2024) Innovation Practices and Challenges in the Telecommunication Industry: A Roadmap for Value Creation and Sustainable Growth BT. In: Current and Future Trends on Intelligent Technology Adoption. Springer Nature Switzerland. Springer. pp. 137–170. https://doi.org/10.1007/978-3-031-61463-7_8

Hussain, W., & Jan, M. A. (editors). (2022). IoT as a Service. Springer International Publishing. https://doi.org/10.1007/978-3-030-95987-6

Jung, J., & Katz, R. (2022). Spectrum flexibility and mobile telecommunications development. Utilities Policy, 75, 101351. https://doi.org/10.1016/j.jup.2022.101351

Latif, R. M. A., Naeem, M. R., Rizwan, O., et al. (2021). A Smart Technique to Forecast Karachi Stock Market Share-Values using ARIMA Model. In: Proceedings of the 2021 International Conference on Frontiers of Information Technology (FIT). https://doi.org/10.1109/fit53504.2021.00065

Lu, X.-Q., Tian, J., Liao, Q., et al. (2024). CNN-LSTM based incremental attention mechanism enabled phase-space reconstruction for chaotic time series prediction. Journal of Electronic Science and Technology, 22(2), 100256. https://doi.org/10.1016/j.jnlest.2024.100256

Myers, J. H., & Tauber, E. (2011). Market structure analysis. Marketing Classics Press.

Nguyen Chau, T., Vu Thi Hong, N., Pham Thi Thu, T., et al. (2024). Re-examining the effects of information and communication technology on economic growth. Technology in Society, 78, 102646. https://doi.org/10.1016/j.techsoc.2024.102646

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2020). Economy Profile of Colombia - Doing Business 2020 Indicators. Available online: https://www.doingbusiness.org/content/dam/doingBusiness/country/c/colombia/COL.pdf (accessed on 2 March 2024).

Petropoulos, F., Apiletti, D., Assimakopoulos, V., et al. (2022). Forecasting: theory and practice. International Journal of Forecasting, 38(3), 705–871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Poobalan, A., Ganapriya, K., Kalaivani, K., et al. (2025). A novel and secured email classification using deep neural network with bidirectional long short-term memory. Computer Speech & Language, 89, 101667. https://doi.org/10.1016/j.csl.2024.101667

Qazi, A., & Al-Mhdawi, M. K. S. (2024). Exploring critical drivers of global innovation: A Bayesian Network perspective. Knowledge-Based Systems, 299, 112127. https://doi.org/10.1016/j.knosys.2024.112127

Quinn, J., McEachen, J., Fullan, M., et al. (2019). Dive into deep learning: Tools for engagement. Corwin Press.

Sheth, J., Jain, V., & Ambika, A. (2020). Repositioning the customer support services: the next frontier of competitive advantage. European Journal of Marketing, 54(7), 1787–1804. https://doi.org/10.1108/ejm-02-2020-0086

Tiwari, A., Chugh, A., & Sharma, A. (2023). Uses of artificial intelligence with human-computer interaction in psychology. Innovations in Artificial Intelligence and Human-Computer Interaction in the Digital Era, 173–205. https://doi.org/10.1016/b978-0-323-99891-8.00003-6

Wang, X., Peng, T., Zuo, P., et al. (2020). Spectrum Prediction Method for ISM Bands Based on LSTM. In: Proceedings of the 2020 5th International Conference on Computer and Communication Systems (ICCCS). https://doi.org/10.1109/icccs49078.2020.9118535

Wang, Y., Zhou, Y., Yan, K., et al. (2023). Software Usage Prediction Based on Hybrid LSTM-ARIMA Algorithm. In: Proceedings of the 2023 IEEE 5th International Conference on Power, Intelligent Computing and Systems (ICPICS). https://doi.org/10.1109/icpics58376.2023.10235473

Wen, X., Liao, J., Niu, Q., et al. (2024). Deep learning-driven hybrid model for short-term load forecasting and smart grid information management. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-63262-x

World Bank Group. (2020). Economy Profile of Colombia - Doing Business 2020 Indicators. Available online: https://www.doingbusiness.org/content/dam/doingBusiness/country/c/colombia/COL.pdf (accessed on 3 May 2024).