www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 17(1), Jan-April 2026, 332-340. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2026.v17i1.1156

A Neural Network-Based Predictive Model for Forest Fire Management in Michoacán

José Luis Martínez Alcántar
Tecnológico Nacional de México / Instituto Tecnológico de Zitácuaro
Laboratorio Nacional CONAHCYT en Inteligencia Artificial y Ciencia de Datos
ORCID: 0000-0002-4540-4231, ResearcherID: KSL-5122-2024
jose.ma@zitacuaro.tecnm.mx
Jhacer Kharen Ruiz Garduño
Tecnológico Nacional de México / Instituto Tecnológico de Zitácuaro
Laboratorio Nacional CONAHCYT en Inteligencia Artificial y Ciencia de Datos
ORCID: 0000-0003-3353-7966, ResearcherID: KSL-5132-2024
jhacer.rg@zitacuaro.tecnm.mx
Claudia Teresa González Ramírez
Tecnológico Nacional de México / Instituto Tecnológico de Zitácuaro
Laboratorio Nacional CONAHCYT en Inteligencia Artificial y Ciencia de Datos
ORCID: 0000-0002-4106-4583, ResearcherID: G-6313-2019
claudia.gr@zitacuaro.tecnm.mx
Samuel Efrén Viñas Álvarez
Tecnológico Nacional de México / Instituto Tecnológico de Zitácuaro
Laboratorio Nacional CONAHCYT en Inteligencia Artificial y Ciencia de Datos
ORCID: 0000-0001-5891-2801, ResearcherID: KSL-5107-2024
samuel.va@zitacuaro.tecnm.mx

Abstract. This study presents a predictive model based on neural networks to estimate the area affected by forest fires in Michoacán, aiming to optimize resource management. A total of 929 historical fire records (2015–2024) were analyzed from 18 localities, including six within the Monarch Butterfly Biosphere Reserve. Using variables such as operational cost, fire duration, impact, and firefighting personnel, several machine learning models were evaluated: linear regression, decision trees, random forest, and a neural network. The neural network achieved the best performance (R^2 = 0.84), identifying firefighting personnel, impact, and duration as the most influential factors. While the dataset lacked key environmental variables, the neural network demonstrated strong predictive capacity, suggesting its potential for future applications with richer datasets. The findings offer a methodological foundation to improve planning and decision-making in wildfire prevention and response under real-world data constraints.
Keywords: forest fire management, predictive modeling, artificial neural networks, Michoacán.

Article Info
Received June 24, 2025
Accepted August 13, 2025

1 Introduction

This study does not aim to achieve the most accurate possible prediction with current data, but rather to evaluate the performance of different regression models (particularly artificial neural networks) under real-world data limitations. The intention is to assess the feasibility and potential of these models for future implementations with more robust datasets. Rather than achieving maximum predictive precision with the current data, this research aims to assess the relative performance of different regression models (especially neural networks) under practical constraints, providing a methodological basis for future studies with improved datasets.

In Mexico, as in many other countries, vast forested areas are highly prone to wildfires due to various causes, with geographic factors playing a key role in their spread and characteristics. Chuvieco et al. (1998) assert that “the phenomenon of forest fires has a clear territorial manifestation, and both its causes and effects are spatially distributed and influenced by geography” (as cited in

Neger et al., 2022, p. 2). Historically, between 2016 and 2023, Mexico recorded an annual average of 7,409.75 wildfires. However, since 2020, there has been a steady increase in the area affected by these events; in the past four years, fires have consumed larger forest areas and show an alarming upward trend.

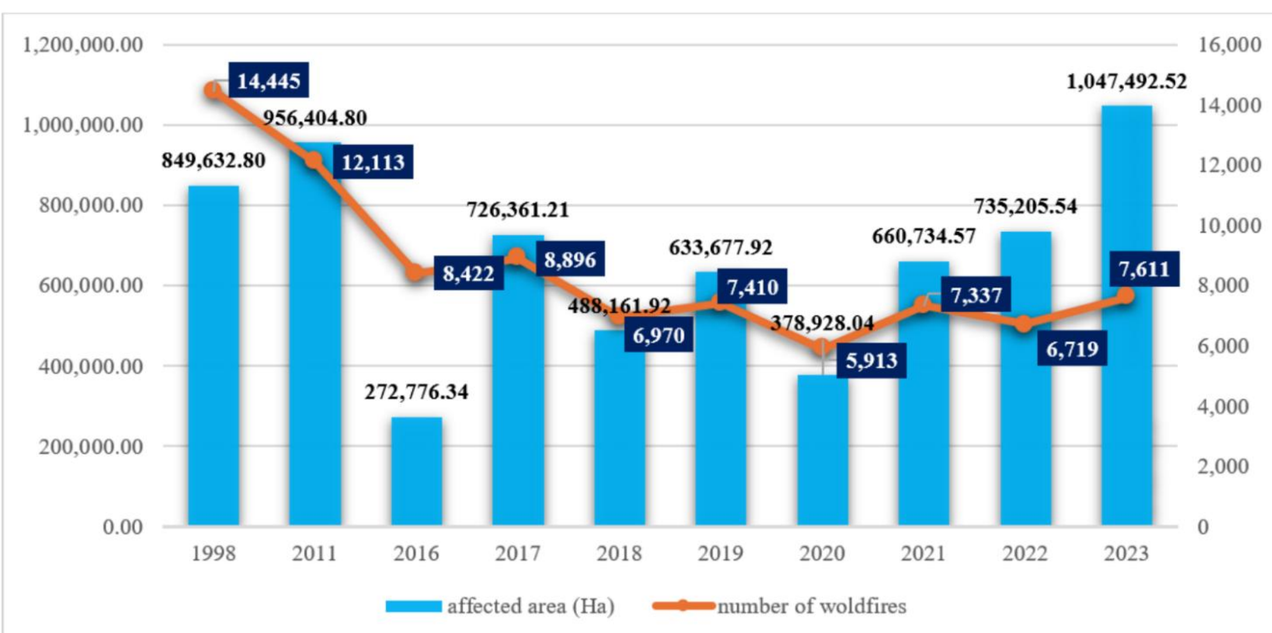
Wildfires can cause damage ranging from mild to severe, leaving significant environmental impacts in both the short and long term. These include disruptions to food chains, destruction of flora, soil erosion, loss of wildlife habitats, climate alterations, air and water pollution, and substantial economic losses. The phenomenon has been studied from various perspectives, particularly in terms of its causes—such as agricultural burns (Bassaber et al., 2024; Vilchis et al., 2018), including sugarcane field burning, land clearing, and intentional agricultural fires (España et al., 2024; Vilchis et al., 2015; Huerta et al., 2024).

Several researchers have pointed out that anthropogenic causes (i.e., those related to human activity, whether accidental or intentional) are the primary triggers of wildfires (Román & Martínez, 2006; Ibarra & Huerta, 2016; Martínez & Ezequiel, 2023; Neger et al., 2022). More than 99% of wildfires have been attributed to human intervention, confirming the findings of the aforementioned authors. While significant progress has been made in the development of techniques and tools for wildfire prevention, prediction, and monitoring, limited advances have been made in leveraging information technologies for the effective management of human, financial, and material resources during wildfire events.

Moreover, authors such as Bassaber et al. (2024), Ibarra and Huerta (2016), and Flores (2021) have studied the consequences of wildfires, highlighting issues such as vegetation degradation, habitat loss, disruption of natural regeneration cycles, wildlife mortality, groundwater contamination, adverse effects on the health of nearby populations, and increased soil erosion. Another recurring research theme involves wildfire management strategies, including prevention, coordination, mitigation, monitoring, restoration (Bassaber et al., 2024; González et al., 2023; Pérez et al., 2024), and funding (Román & Martínez, 2006). However, few studies address regulatory policies and post-fire recovery measures (Cruz & Bulnes, 2019; Bassaber et al., 2024), while others focus on the analysis of fire impacts using indicators such as frequency, ambient temperature, duration, and burned area (Santelices et al., 2022; Perello et al., 2024; González & Ortiz, 2022).

Historically, the number of wildfires in Mexico between 2016 and 2023 has remained stable, with an average of 7,409.75 events per year. Nevertheless, since 2020, there has been a notable increase in the number of hectares affected annually, reflecting a rising trend (Fig. 1). Table 1 presents the ten federal entities with the highest number of wildfires and affected surface area. Additionally, there are regions with minimal human impact that have been designated as protected natural areas, totaling 182 and covering 908,395.20 km² (CONABIO).

Fig 1. Comparison of the number of wildfires and affected area (in hectares) for the years 1998, 2011, and from 2016 to 2023.



Note. Source: National Forestry Commission of Mexico (CONAFOR), Statistical Report 2023 (2024).

Table 1. Mexican States with the Highest Incidence of Wildfires and Burned Area.

Rank	State (by Number of Fires)	Number of Fires	Rank	State (by Affected Area)	Affected Area (ha)
1	Jalisco	297	1	Sonora	38,791.74
2	Guerrero	94	2	Guerrero	22,058.76
3	Mexico City	79	3	Jalisco	16,338.46
4	Morelos	74	4	Coahuila	9,789.11
5	Chiapas	65	5	Nayarit	6,038.18
6	**Michoacán**	**54**	**6**	**Michoacán**	**4,347.75**
7	Sonora	53	7	Zacatecas	3,027.64
8	State of Mexico	48	8	Chiapas	2,584.67
9	Colima	42	9	Tamaulipas	2,100.56
10	Zacatecas	37	10	Oaxaca	1,864.40

Note. Adapted from the National Forestry Commission, 2023 Statistical Report (CONAFOR, 2024). The number of wildfires in sensitive vegetation across these 10 states accounts for 73% of all wildfires and 89% of the total affected area nationwide.

Despite efforts in wildfire prevention and mitigation, there is a lack of studies that comprehensively assess the impact of human, material, and financial resource management on the efficiency of wildfire response. In this regard, the present study aims to design a management model based on Information and Communication Technologies (ICT) to predict the impact of a wildfire (specifically, the number of forest hectares affected) using historical data. This approach is intended to support improved decision-making in wildfire control centers. It is expected that this model will help reduce response times, enhance resource allocation, and minimize both operational costs and environmental impact.

The use of ICT in wildfire management offers significant advantages, such as improved coordination among response teams, enhanced real-time monitoring, and the application of predictive tools based on historical data. By integrating statistical methodologies and machine learning models, this research seeks to answer the following question: How can an ICT-based management model improve the efficiency of wildfire response by optimizing resource use and reducing environmental impact?

Neger, Manzo, and Galicia (2022) compiled advances from geographical studies on wildfires, highlighting key findings, their relevance, and the identification of future challenges and opportunities in Mexico. Their work emphasizes a literature review using bibliometric analysis and thematic integration of publications by geographers and geographic institutions. They concluded that, although geographic contributions to this field have increased in recent years, they remain limited in terms of indexed publications (underscoring the need to further strengthen geographic research on wildfires in Mexico).

2 Experimental procedures

This research follows a quantitative approach with a predictive design, aiming to develop a machine learning model to estimate the surface area affected by wildfires in the eastern region of the state of Michoacán. The analysis was conducted using a dataset of 929 historical wildfire records from 2015 to 2024, covering 16 municipalities: Angangueo, Aporo, Benito Juárez, Contepec, Epitacio Huerta, Hidalgo, Irimbo, Jungapeo, Maravatío, Ocampo, Senguio, Susupuato, Tlalpujahua, Tuxpan, Tuzantla, and Zitácuaro. This area is of high environmental interest, as several municipalities are part of the Monarch Butterfly Biosphere Reserve.

The data were collected from the Regional Wildfire Command Center in the eastern region of Michoacán. Each record includes relevant information such as date, location, cause, and the resources deployed to combat the fire. To predict the affected surface area (target variable), the following variables were selected: month of the fire, operational cost of firefighting, duration of the fire (in hours), number of deployed firefighting personnel, cause of the fire (encoded), specific cause of the fire (encoded), and type of impact (encoded). While the dataset lacks some critical variables such as meteorological factors or high-resolution geospatial indicators, the aim of this study is to evaluate model performance under real-world data availability constraints, which is a common challenge in environmental management contexts.

The data analysis and predictive modeling followed this procedure:

1. **Data cleaning and transformation**: Individual fire records in Excel format were reviewed and organized to retain only relevant data.

2. **Variable encoding**: Categorical variables (Cause, Specific Cause, and Impact) were encoded for compatibility with machine learning algorithms.

3. **Feature selection**: A feature importance analysis was conducted to identify the variables contributing most to the prediction of the affected area.

4. **Model training and evaluation**: Three regression models were tested—Linear Regression, Random Forest, and Decision Tree—and their performance was evaluated using Root Mean Squared Error (RMSE).

5. **Neural network implementation**: A Multilayer Perceptron (MLP) neural network was developed using TensorFlow, achieving a coefficient of determination of $R^2 = 0.84$.

6. **Interpretation**: Modeling and analysis were performed in JupyterLab using Python, with libraries such as Scikit-learn and TensorFlow.

2.1 Model Architecture

This study proposes a deep neural network (MLP) designed for the prediction of a continuous variable, formulated as a supervised regression problem. The architecture consists of a feature input layer, three fully connected hidden layers, and a single output layer with a linear activation function. The hidden layers incorporate L2 regularization and nonlinear activation functions of the ReLU type. Additionally, dropout mechanisms are applied in the first two hidden layers to mitigate overfitting.

2.1.1 Formal Mathematical Representation

Let $x \in R^d$ be the input vector. The complete model can be represented as the following composite function:

$$\hat{y} = f(x;\theta) = W^{(4)} \cdot \emptyset^3(W^{(3)} \cdot \emptyset^2(W^{(2)} \cdot \emptyset^1\left(W^{(1)} \cdot x + b^{(1)}\right) + b^{(2)}) + b^{(3)}) + b^{(4)} \quad \textbf{(1)}$$

Where:

- $x \in R$ is he model's predicted output.
- $\theta = \left\{W^{(l)}, b^{(l)}\right\}_{l=1}^{4}$ represents the model's trainable parameters.
- $\emptyset^1, \emptyset^2, \emptyset^3$ are composite functions that apply ReLU activation and dropout, as applicable.
- $W^{(l)}$ and $b^{(l)}$ denote the weight matrices and bias vectors for layer l, respectively.

Layer dimensions:

$$W^{(1)} \in \mathbb{R}^{512\times d}, W^{(2)} \in \mathbb{R}^{256\times 512},\ W^{(3)} \in \mathbb{R}^{128\times 256}, W^{(4)} \in \mathbb{R}^{1\times 128} \quad \textbf{(2)}$$

The dropout function randomly deactivates 10% of units during training:

$$\text{Dropout}(h,p) = h\ o\ m, \text{ where } m \sim \text{Bernoulli}(1-p), p = 0.1 \quad \textbf{(3)}$$

The activation function used is the Rectified Linear Unit:

$$\text{Relu}(z) = \max\,(0, z) \quad \textbf{(4)}$$

2.1.2 Loss Function

The model is trained using the mean squared error loss function with L2 regularization:

$$\mathcal{L}(\theta) = \frac{1}{n}\sum_{i=1}^{n}(\mathcal{Y}i - \hat{\mathcal{Y}}i)^2 + \lambda\sum_{l=1}^{3}||W^{(l)}||^2 \quad \textbf{(5)}$$

Where:
- $\lambda = 0.0005$ is the regularization coefficient.
- n is the number of training samples.
- $|| \cdot ||$ denotes the Euclidean norm (Frobenius norm for matrices).

2.1.3 Model Properties

- **Representation Capacity**: The use of multiple hidden layers allows the modeling of complex nonlinear relationships between input and output variables.
- **Combined Regularization**: The use of both L2 and dropout contributes to robustness against overfitting, particularly in high-dimensional or limited-data contexts.
- **Extensibility**: This architecture can be easily scaled to multivariable problems or multi-output regression tasks.

This mathematical model enables direct implementation in frameworks such as Keras/TensorFlow or PyTorch, with tunable parameters depending on the application domain. Its predictive and generalization performance is evaluated in the experimental section, in comparison with traditional regression approaches.

3 Results

The model developed in this study enabled the prediction of the surface area affected by forest fires in municipalities located in the eastern region of the state of Michoacán. To evaluate its performance, three regression algorithms and a neural network based on a Multi-Layer Perceptron (MLP) architecture were compared, implemented using the TensorFlow and Keras libraries.

Three conventional machine learning models were trained and evaluated (Table 2). The linear regression model, which assumes a linear relationship between the input variables and the target variable (affected surface area), showed the lowest predictive accuracy among the evaluated models, but serves as a baseline for comparison. The decision tree model, which segments the prediction space through simple decision rules, performed the worst, indicating limited capacity to generalize patterns from the dataset. In contrast, the model based on an ensemble of decision trees (Random Forest) achieved a better balance between bias and variance, with a Root Mean Squared Error (RMSE) of 20.08.

Table 2. Root Mean Squared Error (RMSE) of Regression Models

Model	RMSE (lower is better)
Linear Regression	25.13
Random Forest	20.08
Decision Tree	43.56

The Random Forest model achieved the lowest Root Mean Squared Error (RMSE = 20.08), outperforming both Linear Regression (RMSE = 25.13) and the Decision Tree model (RMSE = 43.56). This suggests that tree-based models are better suited to capturing the relationship between the predictor variables and the affected surface area.

To further improve the accuracy in predicting wildfire-affected surface area, an artificial neural network with a multilayer architecture (MLP) was implemented. This network consists of an input layer, three fully connected hidden layers, and an output layer. The hidden layers use ReLU activation functions, L2 regularization to prevent overfitting, and Dropout layers with a rate of 10% to enhance generalization. The final network structure was as follows:

- Input layer: `Input(shape=(X_train.shape[1],))`
- 1st hidden layer: `Dense(512, activation='relu', kernel_regularizer=l2(0.0005)) + Dropout(0.1)`
- 2nd hidden layer: `Dense(256, activation='relu', kernel_regularizer=l2(0.0005)) + Dropout(0.1)`
- 3rd hidden layer: `Dense(128, activation='relu', kernel_regularizer=l2(0.0005))`
- output layer: `Dense(1, activation='linear')`

This model achieved a Mean Absolute Error (MAE) of 0.2235, a Mean Squared Error (MSE) of 0.5657, and a coefficient of determination (R^2) of 0.84, outperforming traditional regression models (Table 3). These results indicate that the neural network is capable of capturing complex nonlinear relationships between the variables, making it a powerful tool for prediction and planning in wildfire management.

Therefore, the neural network explains 84% of the variance in the affected surface area, surpassing traditional models in terms of predictive accuracy.

Table 3. Neural Network Performance Metrics

Metric	Value
Mean Absolute Error (MAE)	0.2235
Mean Squared Error (MSE)	0.5657
Coefficient of Determination (R^2)	0.84

Fig 2 shows the behavior of the Mean Absolute Error (MAE) during the training and validation process. The graph reveals a progressive decrease in error as training advances, stabilizing at a final value of approximately 0.2235. This result supports the model's efficiency, demonstrating not only a good level of accuracy but also a solid generalization capability.

Fig 2. Training and Validation MAE of the Neural Network Model.

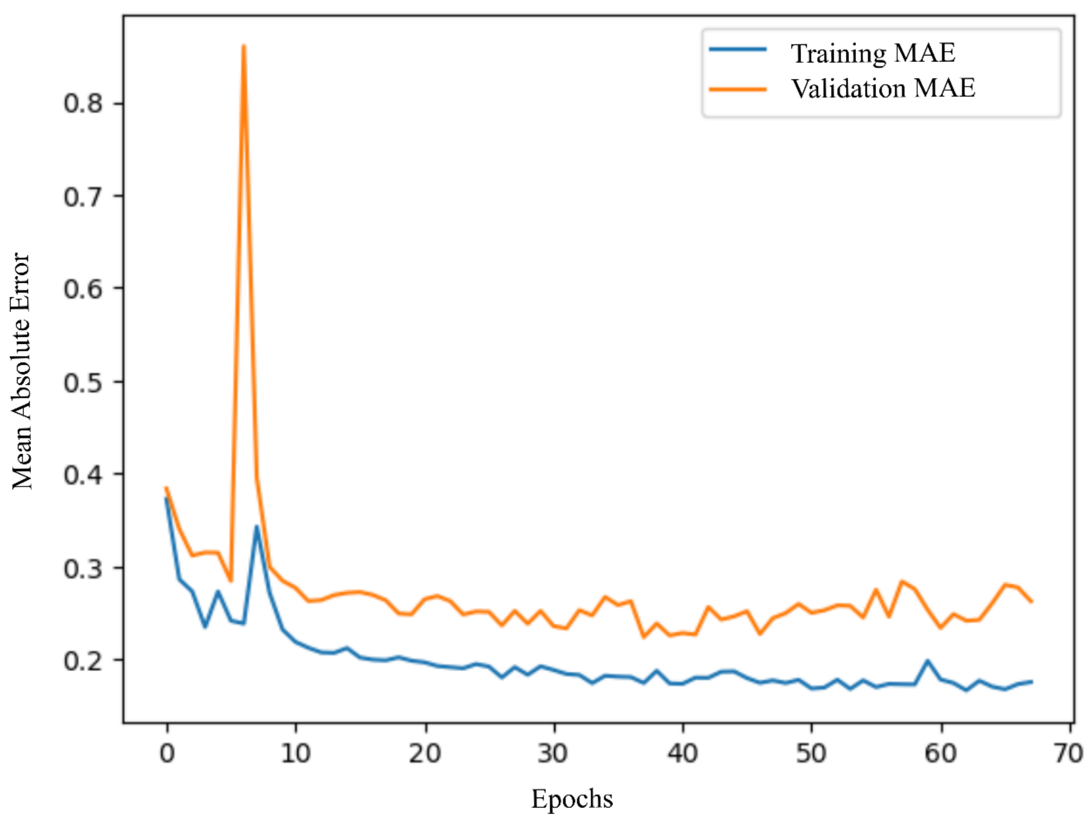
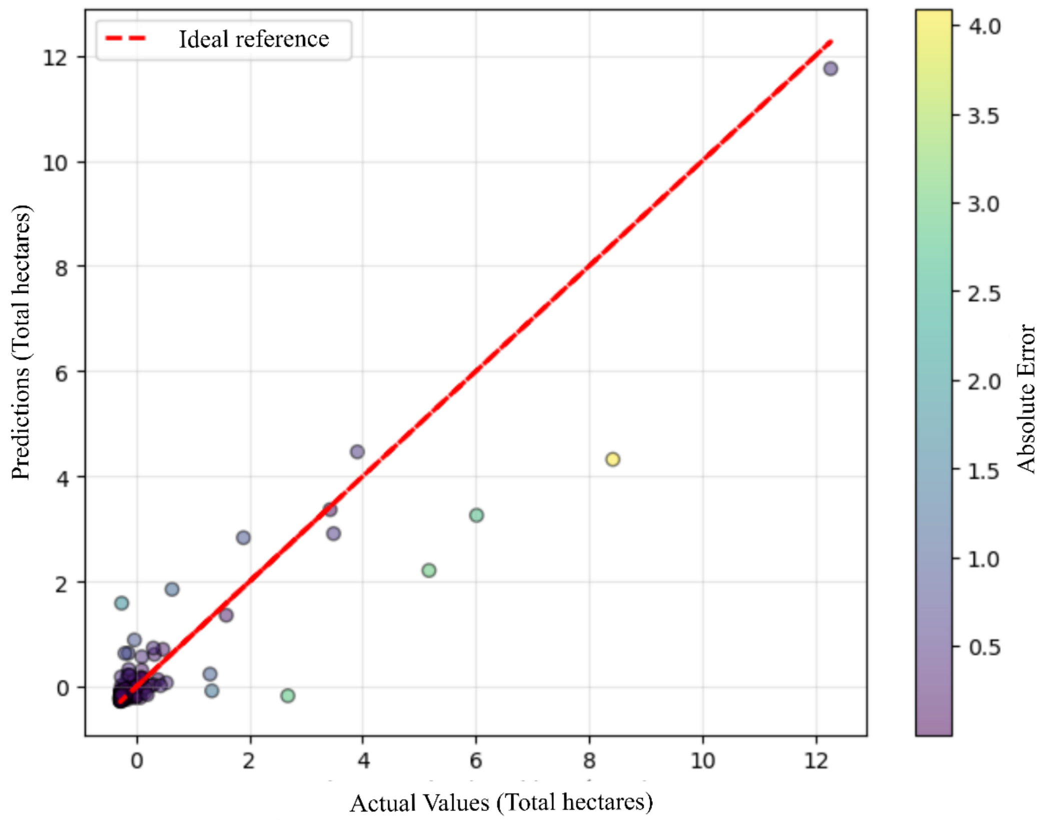


Fig 3 presents a comparison between the predictions generated by the neural network and the actual values of the affected surface area. The graph shows a high level of agreement between the two, confirming the model's ability to accurately reproduce the patterns observed in the historical data.

Fig 3. Comparison of Neural Network Predictions and Actual Values.



Note. The red dashed line represents the ideal reference (perfect prediction). The color gradient indicates the absolute error for each prediction, with lighter colors representing larger deviations from actual values.

The use of neural networks demonstrated a significant improvement in predictive accuracy compared to traditional regression models. While the Random Forest model achieved an RMSE of 20.08, the neural network outperformed it by capturing more complex, nonlinear relationships between the input variables. This suggests that deep learning models can serve as powerful tools for wildfire management and response, enabling more efficient resource planning and mitigation strategies. Moreover, the ability to predict the affected area from the early stages of a fire supports timely, evidence-based decision-making using operational and historical data. Despite the limited feature set, the neural network demonstrated a strong ability to model the underlying patterns, which supports its potential application in future models using richer datasets. These findings position neural networks as a promising methodological approach in wildfire impact prediction and resource optimization.

4 Conclusions

The allocation of resources for combating forest fires—provided by institutions such as CONAFOR, municipalities, and private entities—shows significant variability. Although detailed data on the distribution of these resources was not available, the information analyzed reflects a consistent annual increase in their use, including the deployment of firefighting personnel and specialized equipment. This pattern aligns with the findings of the predictive model, where the number of combat personnel emerged as the most influential variable in predicting the affected surface area. These results highlight the importance of improved resource planning based on historical data and artificial intelligence models, which can enhance wildfire management and response strategies.

Based on the analysis of 929 historical wildfire records (2015–2024), various regression models and a neural network with a Multi-Layer Perceptron (MLP) architecture were evaluated, yielding significant results in terms of accuracy and applicability. The variable importance analysis revealed that the number of deployed firefighting personnel had the greatest impact on the prediction of the affected surface area (34%), suggesting that the availability of human resources plays a key role in wildfire control. Other important variables included the type of fire impact (22%), event duration (20%), and cause of the fire (8%), underscoring the need to evaluate both fire characteristics and response efforts.

The neural network model outperformed traditional regression models in predictive accuracy. Although the Random Forest model achieved a solid RMSE of 20.08, the neural network reached an R^2 of 0.84, explaining 84% of the variability in the affected surface area. It also reduced the MAE to 0.2235, significantly improving prediction precision. While the influence of weather conditions on wildfires is undeniable, the lack of reliable climate data prevented their inclusion in the model. This highlights the need to improve the quality and availability of such data in future studies. Nevertheless, the implementation of neural networks demonstrates the potential of deep learning to optimize resource management in wildfire prevention and control, providing a robust foundation for informed and timely decision-making.

Although the current dataset imposes certain predictive limitations, this work provides a methodological foundation for evaluating and selecting machine learning models (particularly neural networks) in the context of wildfire impact estimation. The results validate the approach and justify future efforts using more comprehensive datasets. Despite the current dataset limitations, the findings support the feasibility of neural networks for wildfire impact prediction, encouraging future research with expanded variables (including climatic and geospatial data) to enhance model performance and real-world applicability.

References

Bassaber, A. I., Maldonado, M. M., Curiel, S., Rojas, M. L., & Valdez, J. S. (2024). *Revisión de causas, consecuencias y medidas de respuesta frente a los incendios forestales: Un enfoque en el estado de Jalisco*. *Revista Científica e-CUCBA*, (21), 100–114. https://doi.org/10.32870/e-cucba.vi21.327

Bassaber, A. I., Maldonado, M. M., Curiel, S., Rojas, M. L., & Valdez, J. S. (2024). *Revisión de causas, consecuencias y medidas de respuesta frente a los incendios forestales: Un enfoque en el estado de Jalisco*. *Revista Científica e-CUCBA*, (21), 100–114. https://doi.org/10.32870/e-cucba.vi21.327

Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. (2024). *Áreas protegidas*. https://www.biodiversidad.gob.mx/region/areasprot

Cruz, X., & Bulnes, E. (2019). *Emission impact of wildfires: El Tepozteco 2016*. *Atmósfera, 32*(2), 85–93. https://doi.org/10.20937/ATM.2019.32.02.01

España, M., Champo, O., & Uribe, M. (2024). *Extensión y severidad de incendios forestales en Michoacán en 2021 a partir de imágenes Sentinel-2*. *Polibotánica*, (57), 125–144. https://doi.org/10.18387/polibotanica.57.7

Flores, J. G. (2021). *Antecedentes y perspectivas de la investigación en incendios forestales en el INIFAP*. *Revista Mexicana de Ciencias Forestales*, número especial. https://doi.org/10.29298/rmcf.v12iEspecial-1.981

González, I., Farfán, M., Morales, L. M., Pérez, D. R., Garza, J. J., & Medina, A. (2023). *Índices y modelos para la predicción de la ocurrencia de incendios forestales: Una revisión para México*. *Revista Geográfica de América Central, 71*(2), 193–219. https://doi.org/10.15359/rgac.71-2.7

González, A., & Ortiz, C. F. (2022). *Superficie forestal afectada por incendios en México: Apuntes iniciales hacia un modelo de manejo preventivo*. *Revista de Ciencias Ambientales, 56*(1), 1–27. https://doi.org/10.15359/rca.56-1.1

Huerta, M. H., Parra, M. R., & Ianni, E. (2024). *Relación entre pobreza y conservación del capital natural desde los medios de vida sustentables en Villaflores, Chiapas, México*. *Revista de Ciencias Ambientales, 58*(1), 1–19. https://doi.org/10.15359/rca.58-1.9

Ibarra, J. L., & Huerta, F. M. (2016). *Cambio climático y predicción de incendios al 2050 en el Bosque La Primavera, Jalisco*. *Revista Mexicana de Ciencias Forestales, 7*(37), 39–50.

Martínez, A., & Ezequiel, P. (2023). *Predicción de incendios forestales mediante modelos de machine learning*. En *Actas del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022)*. Red de Universidades con Carreras en Informática. http://sedici.unlp.edu.ar/handle/10915/149568

Neger, C., Manzo, D. L., & Galicia, S. L. (2022). *La investigación geográfica de los incendios forestales en México: Una perspectiva bibliométrica y territorial*. *Investigaciones Geográficas*, (108), e60488. https://doi.org/10.14350/rig.60488

Neger, C., León, J. F., Galicia, L., & Manzo, L. D. (2022). *Dinámica espaciotemporal, causas y efectos de los megaincendios forestales en México*. *Madera y Bosques, 28*(2), e2822453. https://doi.org/10.21829/myb.2022.2822453

Pérez, L., Vega, M., Díaz, J., & Guadamuz, D. (2024). *Determinación de los efectos ambientales causados por eventos hidrometeorológicos extremos que inciden en el manejo del fuego en sitios del Área de Conservación Guanacaste (ACG), Guanacaste, Costa Rica*. *Revista de Ciencias Ambientales, 58*(1). https://doi.org/10.15359/rca.58-1.12

Román, R. M., & Martínez, J. (2006). *Effectiveness of protected areas in mitigating fire within their boundaries: Case study of Chiapas, Mexico*. *Conservation Biology, 20*(4), 1074–1086. https://doi.org/10.1111/j.1523-1739.2006.00478.x

Santelices, R., Cabrera, A., Silva, P., & Navarro, R. M. (2022). *Assessment of a wildfire in the remaining Nothofagus alessandrii forests, an endangered species of Chile, based on satellite Sentinel-2 images*. *International Journal of Agriculture and Natural Resources, 49*(2), 85–96. https://doi.org/10.7764/ijanr.v49i2.2337

Vilchis, A. Y., Díaz, C., Magaña, D., Ba, K. M., & Gómez, M. A. (2015). *Modelado espacial para peligro de incendios forestales con predicción diaria en la cuenca del río Balsas*. *Agrociencia, 49*, 803–820.