MENDEL
Soft Computing Journal

Predictive Model of the ENSO Phenomenon Based on Regression Trees

Indalecio Mendoza Uribe

Subcoordinación de Eventos Extremos y Cambio Climático, Instituto Mexicano de Tecnología del Agua, México

indalecio_mendoza@tlaloc.imta.mx

Abstract

In this work, the supervised machine learning technique was applied through regression trees to develop a predictive model of the phase of the El Niño-Southern Oscillation (ENSO) phenomenon. Data from the period 1950-2022 were used as training and test. The performance of the predictive model was validated using three continuous type error measurement metrics: Mean Absolute Error, Maximum Error, and Root Mean Square Root. The results indicate that with a greater number of training data the model improves its performance, with a tendency to decrease the error in forecasts. Which starts for the year 1953 with errors of 0.77, 1.41 and 0.75 for MAE, ME and RMSE respectively, ending for the year 2022 with errors of 0.28, 0.72 and 0.13 for the same metrics. It is concluded that, based on the results, the developed model is consistent and reliable for ENSO phase forecasts in a 12-month window.

1 Introduction

Climate is a complex and changing system [7, 10], that presents two behaviors: the continuous and the discontinuous. The first behaves predictably because of how its elements are ordered, for example, the changes associated with the translational motion of the Earth, that give rise to the different seasons of the year. While the second operates when the occurrence of some infrequent event or extreme is recorded, how can a tropical depression be in its different scales (tropical depression, tropical storm, or hurricane). Both processes allow us to observe that the more complex the system is, the less reliable are their samples to predict the possible effects of a disturbance in the system [24].

One of the great challenges for the scientific community is the development of climate forecast models that allow us to know in advance, and with a high degree of reliability, the prevailing climatic conditions at different geographical and temporal scales.

Analysis of climate variability has recently been studied with different variables and approaches. One of them corresponds to the variability of Sea Surface Temperature (SST) in the Pacific Basin can substantially impact the general circulation of the atmosphere on short (seasonal) and long time (inter-annual) scales [4]. Specifically, coupled ocean-atmosphere events such as the El Niño-Southern Oscillation (ENSO) have been linked to significant inter-annual variations in local and regional climates [6, 19, 8, 11, 14, 15].

The ENSO is a natural climatic pattern characterized by significant changes in the SST and the surface pressure of the air of the Tropical Pacific Ocean between the east and west. The ENSO fluctuates between three phases: 1) Niño (warmer than normal); Niña (colder than normal); and Neutral (normal conditions).

During warm events, there is a weakening of atmospheric winds along the equator, so the atmosphere presents a displacement, from its normal position, of the great cloud formation and humidity elevation systems from the Indonesian Tropical Pacific to the American Pacific; its effects are maximized in the winter Northern Hemisphere due to ocean temperatures reaching their maximum value; during hurricane season (June to November), the jet stream is aligned so that the vertical wind increases in the Caribbean and the Atlantic. On the other hand, during cold events, when there is an intensification of the trade winds and, therefore, an anomalous cooling of the waters of the Equatorial Pacific, convection decreases and convective systems move to the opposite side, causing, during the hurricane season, the upper winds are lighter, a condition that favors the development of meteorological phenomena in the Caribbean and the Atlantic [25].

It has been shown that during ENSO events, both in its warm or cold stage, major fluctuations are triggered, as far as magnitude is concerned, in the rainfall patterns of multiple areas of the planet, both northern and southern hemispheres, thus causing droughts, floods, poor harvests and diseases with strong and extensive consequences for society and the environment throughout the world [16, 23, 26, 27, 30]. The duration of the ENSO condition is variable, but various authors have identified that it occurs in periods ranging from three to seven years [17] and with a duration of months to two years.

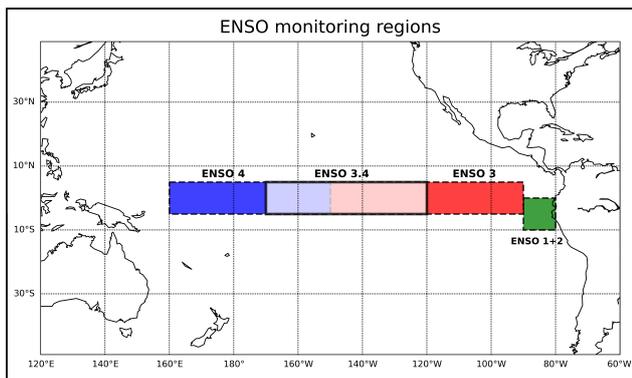In the area of artificial intelligence, various machine-
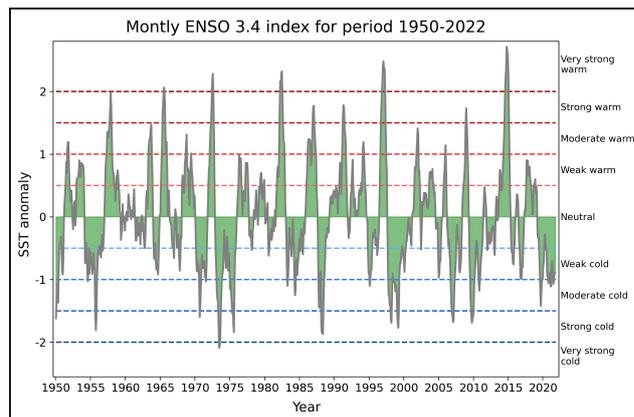
Figure 1: ENSO monitoring regions.

Figure 2: Monthly variability of the ENSO index for the period 1950-2022.

learning techniques have been developed with multiple applications in the world. For example, in the climate [28], food industry [18], and health [2, 3], among others. They correspond to data analysis techniques that provide computers the ability to learn from experience without relying on a given equation as a model. These algorithms find natural patterns in the data that generate knowledge. Learning techniques are classified in a general way as supervised and unsupervised.

A supervised learning algorithm takes a set of known data (inputs) and known answers for this data (outputs) to train a model that can generate reasonable predictions in response to new data. Supervised learning uses classification and regression techniques to develop predictive models. For his part, unsupervised learning looks for hidden patterns or intrinsic structures in data. Used to infer information from data sets consisting of input data with no tagged responses. Clustering, k-means, and neural networks are the most common unsupervised learning technique.

The objective of this work is to develop a predictive model of the monthly phase of the ENSO based on the supervised machine learning method, applying the regression tree technique [5], using as training characteristics the monthly ENSO records prior to the forecast target month.

2 Materials and Methods

2.1 Datasets

The most commonly used SST monitoring regions are: ENSO 1+2 (0°-10°S, 90°-80°S), ENSO 3 (5°N-5°S, 90°-150°W), ENSO 3.4 (5°N-5°S, 120°-170°W) and ENSO 4 (5°N-5°S, 160°E-150°W). However, an index of SST anomalies applied in the ENSO 3.4 region is highly appropriate as a general index of the state of the ENSO cycle [13]. Figure 1 shows the ENSO monitoring regions.

ENSO intensity is classified according to the anomaly threshold: weak (0.5° - 0.9° C), moderate (1.0° - 1.4° C), strong (1.5° - 1.9° C), and very strong (greater than or equal to 2° C). For an event to be classified as weak, moderate, strong, or very strong, must have met or exceeded the threshold for at least

Table 1: Frequency by range of values of the monthly ENSO phase for the period 1950-2022.

Phase ENSO	Range of values	Absolute frequency
Warm very strong	≥ 2.0	20
Warm strong	$\geq 1.5°$ and $< 2.0°$	24
Medium warm	$\geq 1.0°$ and $< 1.5°$	55
Warm weak	$\geq 0.5°$ and $< 1.0°$	125
Neutral	$>$-0.5 and < 0.5	397
Weak cold	$> -1.0°$ and $\leq$ -0.5°	169
Medium cold	$> -1.5°$ and $\leq$ -1.0°	56
Strong cold	$> -2.0°$ and $\leq$ -1.5°	28
Very cold	$\leq 2.0°$	2

three consecutive, overlapping three-month periods. For this study, data Monthly ENSO-3.4 index from the period 1950-2022 were downloaded from the NOAA [1].

2.2 Initial Exploratory Analysis

As a first exploratory analysis, the data series was reviewed to observe the monthly variability of the ENSO in the study period (Figure 2).

As a second descriptive phase of the data, the frequency distribution was elaborated according to the intensity classification of the ENSO phase according to: neutral, weak, moderate, strong, and very strong. Table 1 shows the absolute frequencies obtained for the entire study period.

2.3 Classification and Regression Trees

Quinlan [21, 22] developed the Classification and Regression Trees (CART), which are a non-parametric supervised learning method [12]. Tree models where the target variable can take a finite set of values are called classification trees. On the other hand, trees, where the target variable can take continuous values, are named regression trees. Table 2 describes the main differences between the two types of trees. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from

Table 2: Differences between classification and regression trees.

Classification	Regression
The dependent variable is categorical	The dependent variable is continuous
The value at the terminal node reduces to the mode of the training set observations that have fallen in that region	The values of the terminal nodes are reduced to the mean of the observations in that region
Use entropy as a division criterion	Use the Gini index as a division criterion

the data features. A tree can be seen as a piecewise constant approximation.

The Decision Trees (DTs) are a graphical representation of possible solutions to a decision based on certain conditions, they have a first node called root and then the rest of the input attributes are decomposed into two branches, posing a condition that can be true or false. Each node is bifurcated in two, and they are subdivided again until reaching the leaves which are the final nodes and that are equivalent to answers to the solution or classification. The leaves correspond to the values of the target variable (labels) given the values of the input variables (features) represented by the path from root to leaf.

In general, these algorithms divide the population or sample into homogeneous sets based on the most significant input variables by a search strategy. The data come from records of the form expressed in Eq. (1)

$$(x, Y) = (x_1, x_2, x_3, \ldots, x_k, Y), \qquad (1)$$

where:

Y: Dependent variable (label)

x: Vector of input variables (features)

k: Number of input variables

The dependent variable Y, is the target variable that you are trying to understand, classify or generalize. The input and output variables of the trees can be categorical or continuous. Divide the predictor space (independent variables) in distinct and non-overlapping regions. To obtain the optimal tree, evaluate each subdivision among all possible trees, and get the root node and the subsequent ones, the algorithm must measure the predictions achieved and evaluate them to select the best.

A tree can be learned by dividing the initial set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive way called recursive partitioning. Recursion ends when the subset in a node has all the same value of the target variable, or when partitioning no longer adds value to predictions. This top-down induction process of decision trees is an example of a greedy algorithm [21]. That is, the heuristic consisting of choosing the optimal option is followed in each local step in the hope

of reaching an optimal general solution. The combinations to decide the best tree can be hundreds or thousands.

The mathematical formulation is as follows, given training vectors $x_i \in \mathbb{R}^n$, $i = 1, \ldots, I$ and a label vector $Y \in \mathbb{R}^l$, a decision tree recursively partitions the feature space such that the samples with the same labels or similar target values are grouped together.

Let the data at node m be represented by Q_m with N_m samples. For each candidate split $\theta = (j, t_m)$ consisting of a feature j and threshold t_m, partition the data into $Q_m^{\text{left}}(\theta)$ and $Q_m^{\text{right}}(\theta)$ subsets.

$$Q_m^{\text{left}}(\theta) = \{(x, y) \mid x_j \leq t_m\} \qquad (2)$$

$$Q_m^{\text{right}}(\theta) = Q_m / Q_m^{\text{left}}(\theta) \qquad (3)$$

The quality of a candidate split of node m is then computed using an impurity function or loss function $H()$.

$$G(Q_m, \theta) = \frac{N_m^{\text{left}}}{N_m} H(Q_m^{\text{left}}(\theta)) + \frac{N_m^{\text{right}}}{N_m} H(Q_m^{\text{right}}(\theta)) \qquad (4)$$

Select the parameters that minimize the impurity (5).

$$\theta^* = \text{argmin}_\theta \, g(Q_m, \theta) \qquad (5)$$

Recurse for subsets $Q_m^{\text{left}}(\theta^*)$ and $Q_m^{\text{right}}(\theta^*)$ until the maximum allowable depth is reached, $N_m < \text{min}_{\text{samples}}$ or $N_m = 1$.

2.4 Evaluation Metrics

To consider that the predictive model of ENSO phase generates good estimates, the following continuous error measurement metrics were applied: Mean Absolute Error, Maximum Error, and Mean Square Root. These accuracy metrics are described below.

The Mean Absolute Error (MAE) measures the magnitude of the errors of a set of predictions, regardless of your address. It corresponds to the average of the same absolute differences between the prediction and the actual observation where all the individual differences have the weight (6).

$$\text{MAE} = \sum_{i=1}^{n} \frac{|F_i - O_i|}{n} \qquad (6)$$

where:

F_i: Forecast value at position i

O_i: Observed value at position i

n: Sample size

The Maximum Error test (ME) allows to identify the largest absolute value of the observed error between the prediction and the actual observation (7).

$$\text{ME} = \sum_{i=1}^{n} \max\{|F_i - O_i|\} \qquad (7)$$

The Root Mean Square Root (RMSE) measures the mean magnitude of the error. Corresponds to the

Table 3: Percentage of importance by characteristic for the monthly predictive model. The shaded cells indicate the characteristics with relevance to the predictive model.

Month	Features											
	1	2	3	4	5	6	7	8	9	10	11	12
Jan	0	0	0	0	1	0	0	0	1	1	0	96
Feb	1	0	0	0	0	1	0	0	2	1	15	80
Mar	0	0	1	1	0	0	0	3	0	20	1	73
Apr	3	1	1	2	4	2	2	0	11	0	2	73
May	2	0	3	0	3	3	0	0	2	0	1	86
Jun	5	1	0	2	2	1	1	4	1	1	1	81
Jul	1	2	1	1	0	4	0	0	1	2	4	84
Aug	1	0	4	0	0	1	0	1	1	1	9	83
Sep	0	0	1	1	0	0	1	0	1	0	0	96
Oct	1	0	0	0	0	1	2	1	0	1	2	92
Nov	0	0	0	0	1	0	0	0	1	10	2	86
Dec	0	0	1	0	0	0	0	0	7	1	4	86

square root of the average of the squared differences between the forecast and the observation (8). RMSE amplifies and penalizes with greater force those errors of greater magnitude.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(F_i - O_i)^2} \qquad (8)$$

3 Results

3.1 Training and Test Data

Machine learning is about learning some properties of a data set and then testing those properties against another data set. A common practice in machine learning is to evaluate an algorithm by splitting a data set into two. We call one of those sets the training set, on which we learn some properties; we call the other set the testing set, on which we test the learned properties.

Due to the nature of the ENSO phenomenon, with monthly variations of the SST abnormality, it was proposed to generate a predictive model for each of the months of the year, 12 in total, in this way, the interannual variation intrinsic to the data is considered.

For each monthly predictive model, annual updates were made. Initially, the first predictive model was trained using the monthly data from 1950-1952, and as test data the monthly values of 1953. In a second iteration, the model was retrained with the data from 1950-1953, and as proof the monthly values of 1954. This process was carried out cyclically until the year 2022.

To feed and test the predictive model, the data were grouped into twelve files in CSV format (plain text separated by commas), one per target month. Each file contains 13 columns, the first 12 correspond to the 12 monthly ENSO values prior to the target month (features), and the last one, to the value of the ENSO in the target month (label). Characteristic number 12 corresponds to the ENSO value of the month immediately preceding the target month, while characteristic number 1 corresponds to the ENSO value of the previous

Table 4: Parameters applied in the training of the predictive model

Parameter	Value	Description
Criterion	mse	The function to measure the quality of a split.
Splitter	best	The strategy used to choose the split at each node.
Max depth	None	The maximum depth of the tree. None indicates that nodes are expanded until all sheets are pure or until all sheets contain less than min samples.
Min samples	2	The minimum number of samples required to split an internal node.
Min samples leaf	1	The minimum number of samples required to be at a leaf node.
Max features	7	The number of features to consider when looking for the best split.
Random state	2	Controls the randomness of the estimator. To obtain a deterministic behavior during fitting, the random state must be fixed to an integer.

month furthest from the target month. The features and labels were stored in lists, data structures implemented natively in Python.

3.2 Predictive Model Training

To develop the predictive model, machine learning was applied through the Scikit-Learn library of the Python programming language that integrates a wide range of state-of-the-art machine learning algorithms for supervised and unsupervised problems [20], including the CART algorithm. This package emphasizes ease of use, performance, documentation, and consistency of the API. It has minimal dependencies and is distributed under the simplified Berkeley Software Distribution license (BSD), thus promoting its use in both academic and commercial settings.

The DecisionTreeRegressor class was used for the creation of the monthly predictive models based on regression trees. As part of the training, the algorithm identifies the impact on the forecast of each of the features. As shown in Table 3, on average, the highest percentage of importance for the predictive model is concentrated in 7 features. Above all, in the values of the months closest to the target month.

When creating the predictive model, it is possible to define different configuration parameters. Table 4 shows the main parameters with which the best results were obtained in the proposed predictive model.

Once the training was applied, the trees of the predictive models were graphed using the plot_tree function. As an example, Figure 3 shows the tree corresponding to the predictive model for the month of January.
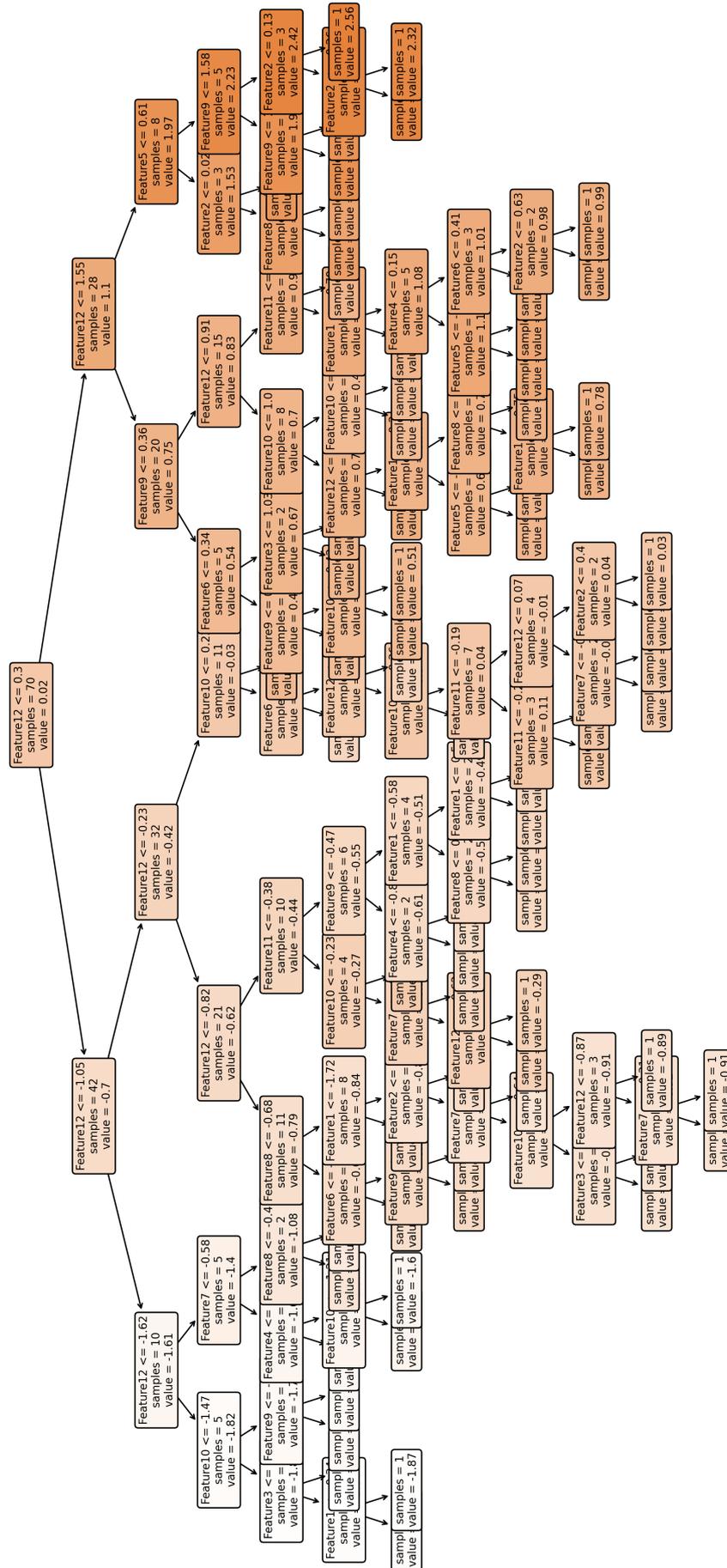
Figure 3: Example regression tree for the month of January. The predictive model was trained with data from the 1950-2021 period. The strongest fill color indicates the majority class for classification.
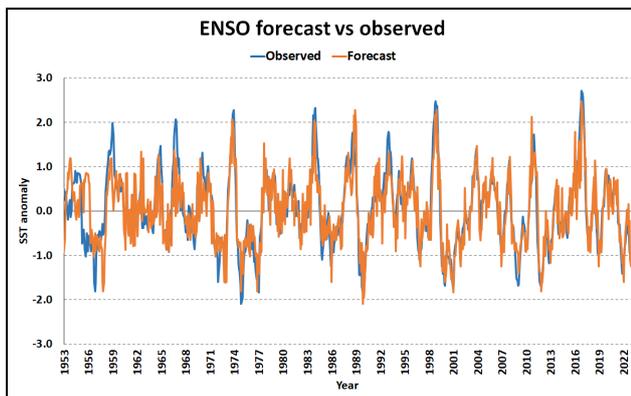
Figure 4: Values of the observed and predicted monthly ENSO index for the period 1953-2022.
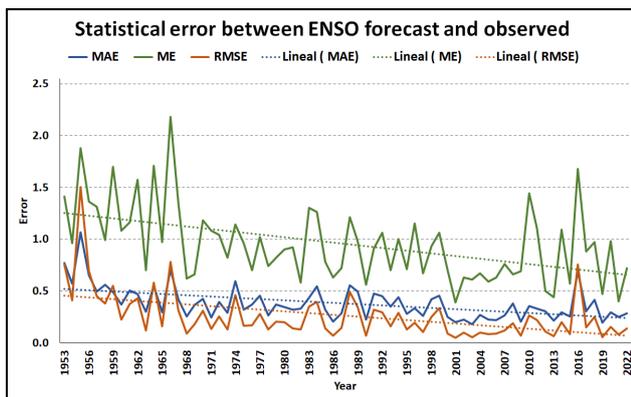

Figure 5: Statistical error between ENSO forecast and observed.

3.3 Validation of the Predictive Model

Figure 4 shows the comparison between the observed data and those predicted by the model for the period 1953-2022. The bias presented mainly in the first years of forecasting is notorious; this is because the regression tree algorithm adaptively improves its performance according to the number of samples provided during training, and in the first training iterations there were few.

Error statistics were calculated between the observed and predicted data year by year. As seen in Figure 5, the three types of errors measured show a tendency to decrease as new samples are incorporated into the training of each of the models. Which start for the year 1953 with errors of 0.77, 1.41 and 0.75 for MAE, ME and RMSE respectively, ending for the year 2022 with errors of 0.28, 0.72 and 0.13 for the same metrics.

These results allow us to determine that the developed model is consistent and reliable for ENSO phase forecasts in a 12-month window.

The robustness, adaptation, and optimization of decision trees are still being discussed. However, in contrast to other methods of data classification, decision trees create an efficient rule collection that is simple to understand. Taha and Mohsin [9] review the most recent research that is conducted in many areas, such

as analysis of medical diseases, classification of texts, classification of user smartphones and images, etc. Determining that, the best accuracy achieved for the decision tree algorithm is 99.93% when it uses a machine learning repository as a dataset.

Another successful case of the application of the CART technique corresponds to the work presented by Xiang et al. [28], based on the summer precipitation data and 130 circulation indexes of 34 national meteorological observation stations in Chongqing, China from 1961 to 2010, created a prediction model of Chongqing summer precipitation. The model carried out the prediction test from 2011-2018 independently. Compared with the results of the single-factor prediction model, the trend consistency rate increased by 37.5% and 12.5% respectively. In addition, when using the random forest model to predict summer precipitation in Chongqing from 2014 to 2018, the 5-year average Ps, Cc and PC scores were 84.6, 0.27 and 67.1, respectively, which were significantly improved compared with 72.4, -0.12 and 52.9 of the current climate forecasting methods, and the forecast quality of the random forest was relatively stable. The multi-system collaborative impact model based on decision tree and random forest algorithms can achieve high accuracy and stability. Thus, this method can not only be an effective means for the diagnosis and prediction of climate causes but also has a good theoretical and practical value for the prediction of extreme disasters.

On the other hand, Xu et al. [29] analysed the influences of climate warming on fire risk. By data joining and processing, a dataset was born which includes 20,622 fire incidents and the related weather data from 2011 to 2017 in Changsha, China. Predictive models of fire frequency were established based on different regression methods of machine learning (random forest, support vector machine and polynomial). Among them, random forest regression models had the best fitting performance and were selected to predict the fire frequency under climate warming scenarios. Under the current warming rate in Changsha, the annual fire frequency in 2067 (50 years after 2017) will increase by 0.69% to 0.89%. By rebuilding predictive models for other cities based on the proposed methods in this study, the influences of climate warming on their fire frequencies can also be analysed.

4 Conclusion

Derived from the analysis of frequencies of the ENSO values for the entire study period, it was observed that the months with neutral values of the index are predominant, with 45% of the total data. Followed by the weak cold and weak warm phases, with values of 19% and 14% respectively. These frequencies were relevant in the evaluation period of the predictive model, where most of the values, both observed and predicted, remained in the range [-1.0, 1.0].

The predictive model developed had a high capacity to estimate the monthly ENSO phase for 12 months

of forecasting. With a tendency to decrease errors as the model considers a greater number of samples in the training stage.

It was noticeable that for each month of the year, the regression tree was different. However, the predominant feature was that the root node of each tree was feature number 12, corresponding to the ENSO value of the month immediately preceding the target month.

Of the 12 features considered, it was identified that on average with seven of them, a relevance close to 100% is obtained for an adequate prognosis, mainly in the feature prior to the target month.

Machine learning through regression trees proved to be effective in developing a predictive model of the ENSO phase, so it can constitute a reliable, but not perfect, forecasting. For an extended forecast (greater than 12 months), the model can predict the shape of the inter-annual fluctuation at the cost of lower forecast accuracy.

Regarding the computational cost required, both in the training process of the predictive model and for the forecast generation, in a personal computer (with Intel i7 2.60 GHz processor and 16 GB RAM) it was less than 10 seconds.

References

[1] National oceanic and atmospheric administration (noaa). monthly niño-3.4 index. https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/detrend.nino34.ascii.txt [accessed on January 6, 2023].

[2] Amami, R., Al Saif, S. A., Amami, R., Eleraky, H. A., Melouli, F., and Baazaoui, M. The use of an incremental learning algorithm for diagnosing covid-19 from chest x-ray images. *MENDEL Journal 28*, 1 (2022), 1–7.

[3] Beaulieu-Jones, B., Finlayson, S. G., Chivers, C., Chen, I., McDermott, M., Kandola, J., Dalca, A. V., Beam, A., Fiterau, M., and Naumann, T. Trends and focus of machine learning applications for health research. *JAMA network open 2*, 10 (2019), e1914051–e1914051.

[4] Birk, K., Lupo, A. R., Guinan, P., and Barbieri, C. The interannual variability of midwestern temperatures and precipitation as related to the enso and pdo. *Atmosfera 23*, 2 (2010), 95–128.

[5] Breiman, L. *Classification and regression trees.* Routledge, 2017.

[6] Cane, M. A. Oceanographic events during el nino. *Science 222*, 4629 (1983), 1189–1195.

[7] Carrera, E. J. B. Sobre las variaciones climáticas en méxico. *Ingeniería de Recursos Naturales y del Ambiente*, 11 (2012), 117–127.

[8] Cerano-Paredes, J., Villanueva-Díaz, J., Valdez-Cepeda, R. D., Arreola-Ávila,

J. G., and Constante-García, V. El niño oscilación del sur y sus efectos en la precipitación en la parte alta de la cuenca del río nazas. *Revista Chapingo serie ciencias forestales y del ambiente 17*, SPE (2011), 207–215.

[9] Charbuty, B., and Abdulazeez, A. Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends 2*, 01 (2021), 20–28.

[10] Conde, C. *México y el cambio climático global.* Universidad Nacional Autónoma de México, Dirección General de Divulgación de . . . , 2006.

[11] Diaz, H. F. *El Niño and the Southern Oscillation: multiscale variability and global and regional impacts.* Cambridge University Press, 2000.

[12] Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.

[13] Kousky, V., and Higgins, R. An alert classification system for monitoring and assessing the enso cycle. *Weather and Forecasting 22*, 2 (2007), 353–371.

[14] Kung, E. C., and Chern, J.-G. Prevailing anomaly patterns of the global sea surface temperatures and tropospheric responses. *Atmósfera 8*, 3 (1995).

[15] Lupo, A., Kelsey, E., Weitlich, D., Woolard, J., Mokhov, I., Guinan, P. E., and Akyüz, F. Interannual and interdecadal variability in the predominant pacific region sst anomaly patterns and their impact on climate in the mid-mississippi valley region. *Atmosfera 20*, 2 (2007), 171–196.

[16] McPhaden, M. J. Toga-tao and the 1991-93 el nino-southern oscillation event. *Oceanography 6*, 2 (1993), 36–44.

[17] McPhaden, M. J. Genesis and evolution of the 1997-98 el niño. *Science 283*, 5404 (1999), 950–954.

[18] Minh, H. T., Anh, T. P., et al. A novel lightweight dcnn model for classifying plant diseases on internet of things edge devices. *MENDEL Journal 28*, 2 (2022), 41–48.

[19] Molina, J. J. C. El fenómeno enso (el niño. oscilación del sur) en 1997-1998: alteraciones climáticas inducidas en el mundo. *Nimbus: Revista de climatología, meteorología y paisaje*, 3 (1999), 37–62.

[20] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research 12* (2011), 2825–2830.

[21] Quinlan, J. R. Induction of decision trees. *Machine learning 1* (1986), 81–106.

[22] Quinlan, J. R. *C4. 5: programs for machine learning.* Elsevier, 2014.

[23] Ramírez Solano, A. M., Chamizo García, H. A., and Fallas Sojo, J. C. Fenómeno enos y el dengue, regiones pacífico central y huetar atlántico, costa rica, 1990 a 2011. *Población y Salud en Mesoamérica 15*, 1 (2017), 6–25.

[24] Ritter, W., Guzmán, S., Sánchez-Santillán, N., Suarez, J., Corona, C., Muñoz, H., Ramos, A., Rodríguez, R., and Pérez, T. El clima como sistema complejo adaptativo en coevolución. *Ciencia y Mar 6*, 17 (2002), 23–25.

[25] Rodríguez-Moreno, V. M., Ruíz-Corral, J. A., Medina-García, G., Padilla-Ramírez, J. S., and Gunter Kretzschmar, T. Efecto de la condición enso en la frecuencia e intensidad de los eventos de lluvia en la península de baja california (1998-2012). *Revista mexicana de ciencias agrícolas 5*, SPE10 (2014), 1923–1937.

[26] Sánchez-Santillán, N., Lanza-Espino, G., Garduño, R., and Sánchez-Trejo, R. La influencia antropogénica en el cambio climático bajo la óptica de los sistemas complejos. *Rev. Iberoam. Ciencias 2*, 6 (2015), 69–84.

[27] Valiente, Ó. M. Evolución en el estudio del fenómeno enso (el niño-oscilación del sur): de anomalía local a la predicción de variaciones climáticas globales. *Investigaciones Geográficas (Esp)*, 21 (1999), 5–20.

[28] Xiang, B., Zeng, C., Dong, X., and Wang, J. The application of a decision tree and stochastic forest model in summer precipitation prediction in chongqing. *Atmosphere 11*, 5 (2020), 508.

[29] Xu, Z., Liu, D., and Yan, L. Temperature-based fire frequency analysis using machine learning: A case of changsha, china. *Climate Risk Management 31* (2021), 100276.

[30] Zhang, Y., Wallace, J. M., and Battisti, D. S. Enso-like interdecadal variability: 1900–93. *Journal of climate 10*, 5 (1997), 1004–1020.