

# Predicción de cobertura de coral de los arrecifes de coral escleractinios de Santa Marta, Colombia aplicando técnicas de aprendizaje automático

Sebastián Zapata,  
Lizeth Blanco Alfonso  
Maestría en Ingeniería y Analítica de Datos  
Universidad Jorge Tadeo Lozano  
[lizethy.blancoa@utadeo.edu.co](mailto:lizethy.blancoa@utadeo.edu.co)  
Director: [sebastian.zapatar@utadeo.edu.co](mailto:sebastian.zapatar@utadeo.edu.co)

**Resumen**—Los arrecifes de coral son invertebrados coloniales que existen hace miles de años. Son ecosistemas que permiten mantener el equilibrio químico, bioquímico y biológico del océano. Sin embargo el cambio climático, transformaciones naturales y los procesos ejecutados por los seres humanos sobre el medio ambiente han provocado alteraciones en la dinámica poblacional y deterioro del ecosistema. Es así como esta investigación busca evaluar modelos como redes neuronales y procesos estocásticos para predecir el cambio de cobertura de coral de los arrecifes de coral escleractinios en Santa Marta, Colombia. Específicamente, se evaluaron dos técnicas de predicción de series temporales: proceso estocástico ARIMA y red neuronal LSTM sobre la serie temporal de cobertura de coral escleractinio en el Chengue Santa Marta Colombia, encontrando que la red neuronal LSTM tiene un mejor rendimiento respecto al modelo ARIMA.

**Palabras clave:** Arrecifes Coral, Red Neuronal, Series de Tiempo, Coral Escleractinio, Aprendizaje Automático.

**Abstract**—Coral reefs are colonial invertebrates that have existed for thousands of years. They are ecosystems that allow maintaining the chemical, biochemical and biological balance of the ocean. However, climate change, natural transformations and the processes carried out by human beings on the environment have caused alterations in population dynamics and deterioration of the ecosystem. This research seeks to evaluate models such as neural networks and stochastic processes to predict the change in coral cover of scleractinian coral reefs in Santa Marta, Colombia in order to contribute to the estimation of the behavior of these ecosystems. Two time series prediction techniques were evaluated: ARIMA stochastic process and LSTM neural network on the time series of scleractinian coral cover in Chengue Santa Marta Colombia, finding the LSTM neural network has a better performance compared to the ARIMA model.

**Keywords:** Coral reef, Neural Network, Time Series, Scleractinia, Machine Learning.

## I. INTRODUCCIÓN

Los arrecifes de coral son invertebrados coloniales que existen hace miles de años. Son uno de los ecosistemas más valiosos, siendo de los más productivos y diversos del planeta que mantienen el equilibrio químico, bioquímico y biológico del océano [1].

En el caso de Colombia, limitada por el Océano Atlántico y Océano Pacífico, este ecosistema posee una extensión del 77% y 1%, respectivamente [2]. Aclarando que la zona del Caribe está conformada por Islas del Rosario, Bajo Tortugas, Santa Marta, el Tayrona, Islas de San

Bernardo, Barú, Isla Fuerte, Santa Catalina, Islas de San Andrés, Providencia y el Urabá. Según un estudio realizado por Rangel [3], el Parque Nacional Natural Tayrona de Santa Marta de Colombia abarca el mayor porcentaje de coral en el Caribe, sin embargo, los procesos naturales y efectos de las actividades de los seres humanos sobre el medio ambiente [32] han generado alteraciones en la dinámica poblacional y deterioro del ecosistema.

Tal deterioro ha sido evidenciado con la reducción de la cobertura de corales vivos de hasta un 50% [30], debido a impactos por sedimentación, contaminación, cambio

climático y sobrepesca, generando enfermedades y un aumento en la mortalidad coralina.

Desde la perspectiva de manejo y gestión de información sobre la condición de los cambios de los ecosistemas marinos, el INVEMAR<sup>1</sup>, cuenta con un indicador denominado “Indicador de Condición-Tendencia (ICT)” que permite evaluar la condición general de la integridad biótica de áreas coralinas [4]. En el monitoreo de áreas coralinas se han aplicado técnicas de aprendizaje automático para detectar corales y facilitar su clasificación esta información no ha sido analizada desde un punto de vista predictivo [33].

Por lo anterior, esta investigación busca evaluar el desempeño de diferentes técnicas de predicción estocásticas y de aprendizaje automático para la búsqueda de la mejor predicción del cambio de cobertura de coral de arrecifes de coral escleractinios en Santa Marta, Colombia, con el fin de brindar un diagnóstico temprano y apoyar las medidas de manejo y conservación de estos ecosistemas.

El entrenamiento y evaluación del desempeño de los modelos se realiza a partir de los datos compartidos por el INVEMAR, que recopila información de tipo biológico y ambiental en diferentes estaciones de Santa Marta, Colombia.

Para comprender y predecir el índice de cobertura de coral en el Chengue Santa Marta, Colombia, el documento se estructura en 7 secciones, donde en las primeras secciones se realiza una revisión teórica de los principales fundamentos y antecedentes de series de tiempo e investigaciones realizadas en los arrecifes de coral.

En la sección 4 se plantea el problema de la investigación; en la sección 5 se detalla la metodología y alcance geográfico de la investigación; en la sección 6, se

presentan los resultados de la aplicación de la metodología y finalmente en la sección 7 se detallan las conclusiones del trabajo.

## II. ESTADO DEL ARTE

El presente estado del arte da cuenta de la revisión de textos académicos que abordan la temática principal de la presente investigación, de manera que, inicialmente se recopilan artículos sobre predicciones del cambio de cobertura coralina, cambio poblacional, cambio comunitario a nivel global, así como modelos de aprendizaje automático aplicados en arrecifes de coral y por último modelos estocásticos aplicados a la dinámica poblacional de los arrecifes de coral.

La inquietud por los arrecifes de coral y su impacto en los ecosistemas marinos llevó a Ward, Kim y Harvell [20], a analizar los umbrales de estrés de los corales por elevación de la temperatura en los Cayos de Florida, Estados Unidos. Para ello, los autores tomaron fracciones del coral *Gorgonia* (genus *Gorgonia*) sano y sometieron al coral a diferentes temperaturas, luego realizan un análisis de la relación de las variables mediante la generación de una matriz de correlación entre la abundancia de *zooxanthellae* y actividad de los patógenos (hongo *aspergillus*), encontrando que al aumentar la temperatura disminuye la resistencia del coral *Gorgonia*, haciéndolo más susceptible a la infección por hongos tipo “*Aspergillus sydowii*”, al tiempo que aumenta la tasa de crecimiento de estos patógenos, ocasionando así la muerte del coral *Gorgonia*.

Con estas características y bajo la noción de modelar las enfermedades de los arrecifes de coral en los ecosistemas marinos, Sooley, Foley y Richarson [21], simulan la dinámica poblacional y modelan estocásticamente las enfermedades de los arrecifes de coral de los Cayos de Florida, Estados Unidos.

Para modelar la dinámica poblacional de los microorganismos patógenos de corales, los autores adaptan los modelos epidemiológicos propuestos por Levy y Hess, asumiendo que no hay transmisión vertical

<sup>1</sup> Instituto de Investigaciones Marina y Costeras José Benito Vives de Andrés (INVEMAR).

entre los corales y tampoco existe recuperación individual. Estas características, permiten a los investigadores modelar a los microorganismos patógenos de coral como un sistema de enfermedad.

De esta investigación se denota que mediante ecuaciones diferenciales estocásticas se puede simular el comportamiento de los arrecifes de coral con respecto a enfermedades infecciosas en Florida, Estados Unidos; sin embargo, un escenario más realista podría considerar la relación entre los corales y algas con las enfermedades infecciosas de los arrecifes de coral.

En esta línea, González [22], representa en un modelo de ecuaciones diferenciales estocásticas, la dinámica poblacional de arrecifes de coral, incluyendo elementos importantes dentro de los arrecifes de coral como: la presencia de algas, peces (escaridos y peces loro) y el efecto de la alimentación de algunos peces, representada en el modelo como un movimiento browniano.

Si bien, esta investigación incluye elementos importantes en el modelo sobre los ecosistemas de arrecifes de coral, no incluye la influencia de factores externos como la temperatura o la importancia de ciertos parámetros, la cual se puede obtener aplicando modelos de regresión.

Es así como, William, Aaeby Cowie y Davy [23], buscan predecir la dinámica poblacional de los arrecifes de coral en Hawái, incluyendo factores ambientales para cuatro enfermedades de coral: anomalías de crecimiento (PorGA), pérdida de tejido (PorTL), trematodiasis (PorTrem) y síndrome de blanqueamiento (MWS). En la predicción de las enfermedades de arrecifes de coral de Hawái, los investigadores aplican técnicas de aprendizaje automático supervisado, puntualmente, Boosted Regression Trees (BRT), más no aplican técnicas de predicción como series temporales.

Para mejorar la precisión del monitoreo de arrecifes de coral, Corruz, Filipina, Santiago, Lanzana y Bandala en su investigación: “*BahurApp: Development And Implementation Of Coral Bleaching Monitoring*

*Application Using Convolutional Neural Network*” [34], desarrollaron una aplicación móvil que monitorea daños de blanqueamiento en arrecifes de coral de Filipinas. Tal monitoreo se realiza mediante el uso de redes neuronales convolucionales que clasifican la severidad del blanqueamiento en los corales.

Desde esta óptica, Sanz, Ser, Torres, Lopez y Figueras, desarrollaron un modelo de optimización de arrecifes de coral denominado CRO, que simula artificialmente la dinámica de crecimiento, reproducción y supervivencia de corales (lucha entre corales por espacios en el arrecife) [9], en una matriz o una cuadrada de  $N \times M$  cuadros, donde cada cuadro  $(i,j)$  de la grilla representa un coral o colonia de corales, optimizando el espacio del arrecife y propendiendo al crecimiento y reproducción de los corales.

Con base en la anterior investigación, Yang y Zhang [24], mejoran el modelo aplicando técnicas de aprendizaje profundo. En esta investigación, los autores desarrollan un modelo de evolución diferencial a los arrecifes de coral aplicando redes neuronales de una sola capa y así alcanzar un mejor rendimiento del modelo de optimización de arrecifes de coral. De esta investigación se resalta, la aplicación de técnicas de aprendizaje automático, para modelar la dinámica de los corales en los arrecifes y su respectiva evolución.

De este modo, Pandolfi [25], busca predecir la respuesta ecológica y evolutiva de los arrecifes de coral tropicales de Australia, con respecto al cambio climático global. En este modelo se analiza e interpreta la respuesta de los corales del trópico Australiano ante el estrés térmico y acidificación del océano.

Siguiendo esta línea, Sanguino, Ortinz, Thompson, Wolff, Ferari, Robson, Magno-Canro, Fabricius y Uthicke [26], desarrollan un modelo de regresión lineal con el fin de comprender la influencia de múltiples eventos estresantes acumulativos sobre el estado y desempeño de los arrecifes de coral en la Gran Barrera de Coral Central y Sur de Australia. Para el desarrollo del modelo, los

investigadores incluyen 25 métricas ambientales que consideran de alto impacto en Australia, concluyendo que estas métricas tienen una mayor influencia y predicen mejor el desempeño del arrecife de coral comparado con el estado del mismo arrecife de coral.

Teniendo en cuenta los diferentes eventos naturales que pueden afectar el deterioro de los arrecifes, Mumby, Hasting y Edwars [27], proponen un modelo estocástico que simula la dinámica y deterioro de las colonias de corales ante eventos como huracanes. Este modelo está compuesto por ecuaciones estocásticas. De esta manera en la investigación “A global coral reef probability map generated using convolutional neural networks” [35], plantean un modelo que busca estructurar el arrecife de coral que pueda controlar la erosión costera y ayudar a preservar el medio ambiente costero.

En el marco del diseño los protocolos del Sistema de Monitoreo de Arrecifes de Coralinos (Simac), el INVEMAR [4], establece un indicador de clasificación de condición-tendencia (ICT) de los ecosistemas marinos, para evaluar la situación general de integridad o salud de los arrecifes de coral a lo largo del tiempo. Este indicador integra cuatro variables, tal y como se evidencia a continuación:

$$ICT_{AC} = \frac{CCV+CMF+BPH+BPC}{4}$$

Donde:

*CCV = Cobertura de coral duro vivo que corresponde al porcentaje de superficie del fondo cubierta por coral duro*

*CMFT = Cobertura de macroalgas frondosas y tapete algal que corresponde al porcentaje de superficie del fondo cubierto por macroalgas frondosas*

*BPH = Biomasa de peces herbívoros – Loros y Citrujanos: se expresa en gramos*

*BPC = Biomasa de peces carnívoros como Pargos, Chernas y Meros: Se expresa en gramos*

De acuerdo al valor obtenido en el indicador ICT, la condición general de la integridad del área coralina puede clasificarse en una de las siguientes cinco (5) categorías: Deseable (4,21-5), Buena (3,41-4,2), Regular (2,61-3,4), Alerta (1,8 1-2,6), No deseable (1-1,8) [4].

Con la revisión de la literatura existente, se evidencia que a la fecha hay investigaciones que plantean indicadores

que permiten monitorear la dinámica de los arrecifes de coral, modelos de aprendizaje automático para la clasificación de corales en los arrecifes y modelos que simulan la dinámica poblacional en los arrecifes.

### III. MARCO TEÓRICO

A continuación, se realiza una revisión sobre los conceptos y principios teóricos relevantes para el desarrollo de la investigación.

#### A) ARRECIFES DE CORAL.

Los arrecifes de coral son hábitats marinos conformados por un conjunto de organismos tales como algas y corales. De acuerdo con Spalding [5], los arrecifes de coral son estructuras bio-génicas construidas a través de la acumulación de carbonato de calcio, corales y otros organismos marinos. Estos se definen como una gran estructura subacuática que está compuesta por múltiples individuos de phylum cnidaria<sup>2</sup>, los cuales presentan esqueleto y por esto son considerados como la base fundamental para la formación de los grandes arrecifes de coral [6].

La forma y estructura de los arrecifes es variable, existen desde pequeños parches coralinos hasta amplias barreras arrecifales. Se clasifican en atolones, banco o plataforma, arrecifes de parche, arrecifes bordeantes y de barrera. Los atolones son ecosistemas que se encuentran alrededor de superficies de agua donde existieron islas volcánicas; los arrecifes de banco o plataforma, son arrecifes semi circulares o lineales concentrados en el pasto marino; los arrecifes de parche son arrecifes similares a los de banco, se diferencian en la concentración y estos por lo general se encuentran dentro de una laguna; los arrecifes de barrera, poseen un mayor porcentaje de cobertura y, los arrecifes bordeantes, por su nombre, bordean la isla y se encuentran en profundidades bajas cerca a la costa [7].

<sup>2</sup> Grupo de organismos invertebrados acuáticos con presencia de células urticantes llamadas cnidoblastos presente en los tentáculos o boca de estos organismos.

### B) Corales

Los corales son invertebrados marinos y simbioses que pertenecen al Phylum Cnidaria, cuyo esqueleto está constituido de carbonato de calcio ( $\text{CaCO}_3$ ) [8]. Se caracterizan por estar compuestos de organismos denominados pólipos. Su estructura original es de color blanco; los diferentes colores que representa se deben a un proceso simbiótico entre el coral y unas microalgas denominadas zooxantellas que deben desarrollar en el proceso de fotosíntesis. Los corales son organismos sésiles, es decir que están incrustados a una superficie y no tienen movimiento, convirtiéndose en los constructores de estructuras arrecifales en las zonas tropicales del mundo [5].

Los corales se clasifican en corales octorales, también denominados, blandos o ahermatípicos y los corales pétreos también denominados duros o hermatípicos. Los primeros están compuestos de espículas (fragmentos de calcita) debido a la falta de esqueleto. Por otro lado, los corales duros, están compuesto por células caliciformes que extraen el calcio disuelto del mar conformando un exoesqueleto para sostenerse [9].

### C) Corales Escleractineos

Los corales escleractinios son considerados organismos sésiles y constructores, caracterizados por la capacidad de formar un exoesqueleto de carbonato de calcio que puede alcanzar grandes dimensiones. Pertenecen al Phylum Cnidaria, Clase Anthozoa y orden Scleractinia [10].

Para que una colonia crezca se llevan a cabo dos procesos básicos: la clonación y la calcificación. En el primer proceso, cada pólipo se divide generando uno o más organismos genéticamente iguales asociados entre sí, permitiendo que nuevos individuos se añadan a la colonia y que éste incremente su tamaño [11]. Por otro lado, el proceso de calcificación consiste en la formación de esqueletos duros mediante la relación simbiótica entre el pólipo con una población de microalgas tipo

dinoflagelado endosimbionte<sup>3</sup>, del género *Symbiodinium spp*, comúnmente llamado *zooxanthella* [12].

### D) Cambio de cobertura de coral

La cobertura de coral es un indicador del área cubierta por corales en el arrecife. De manera que permite medir la conformación de corales a lo largo del arrecife y en contraste con los demás organismos en el arrecife [13].

### E) Aprendizaje Automático

El aprendizaje automático hace parte de la inteligencia artificial, permitiendo a las máquinas (computadoras) aprender sin ser programadas explícitamente, es decir, no requieren de instrucciones específicas para resolver un problema; por el contrario, el aprendizaje automático consiste en encontrar patrones o tendencias mediante el desarrollo de algoritmos que aprenden de los datos recibidos, de ello resulta que para cada problema o contexto existe una estructura o arquitectura particular del algoritmo de predicción [14]. De acuerdo con la estructura taxonómica, los algoritmos se clasifican en tres grupos: aprendizaje supervisado, no supervisado y de refuerzo.

- Aprendizaje supervisado. Técnica donde los algoritmos aprenden de los datos que han sido previamente etiquetados. Estos se categorizan en algoritmos de clasificación, regresión y aprendizaje profundo.
- Aprendizaje no supervisado. Técnica donde los algoritmos aprenden por sí mismos sin supervisión o sin etiquetas; el algoritmo debe clasificar la información por sí mismo. Estos se categorizan en algoritmos de agrupación y reducción de dimensionalidad (PCA).
- Aprendizaje por refuerzo. Técnica donde los algoritmos se implementan como agentes que deben explorar un ambiente particular, aprender de él y tomar acciones con base en este [15].

---

<sup>3</sup> Reciben esta denominación por estar en el tejido gastrodermis del coral.

### F) Aprendizaje profundo

El aprendizaje profundo es un subgrupo de algoritmos de aprendizaje automático que se caracteriza por tener una estructura conformada por múltiples capas de procesamiento interconectadas que se comunican entre sí [16]; de aquí que estos algoritmos reciban el nombre de redes neuronales artificiales dado que según el nivel de abstracción o profundidad de cada capa de neuronas, se aplican diferentes transformaciones a los datos de entrada con el fin de obtener una salida. Para ilustrar este comportamiento, a continuación, se muestra una representación de la estructura lógica básica de una red neuronal artificial.

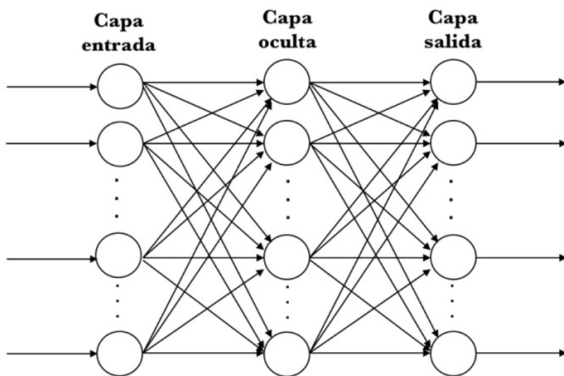


Figura 1. Arquitectura red neuronal básica

Fuente. OHLSSON [16].

En la Figura 1, se evidencia una red neuronal artificial con 3 capas: una de entrada, una capa oculta y una de salida. La primera capa, se encarga de recibir los datos de entrada; la segunda (capa oculta) conformada por diferentes neuronas artificiales, se encarga de aplicar las transformaciones lineales y no lineales necesarias y, entregarla a la capa de salida que devuelve la predicción realizada.

### G) Proceso estocástico.

Un proceso estocástico trata de estudiar el comportamiento de variables aleatorias que puede estar indexado por un conjunto no numerables (aleatorio), matemáticamente, se define como una colección de variables aleatorias definidas sobre un mismo espacio de probabilidad  $T$  (espacio parametral) y con valores en un

espacio medible, donde las variables toman valores en un conjunto  $S$ , llamado espacio de estados [17].

De acuerdo con el espacio parametral y el espacio de estados se clasifican en:

- El espacio parametral es un subconjunto de los números naturales, entonces el proceso estocástico recibe el nombre de sucesión aleatoria o proceso de parámetro discreto.
- El espacio parametral es un intervalo de los reales, entonces el proceso estocástico recibe el nombre proceso de parámetros continua.
- El espacio de estados es discreto, es decir finito o numerables, entonces el proceso estocástico recibe el nombre de proceso de estado discreto.
- El espacio de estados es no numerable, entonces el proceso estocástico recibe el nombre del proceso de estado continuo.

Teniendo en cuenta la anterior clasificación y el problema a estudiar, los procesos estocásticos pueden ser: procesos de ensayos independientes, procesos de Márkov, procesos con incrementos independientes, procesos estacionarios, procesos con incrementos estacionarios, martingala, procesos de Lévy, procesos Gaussianos, proceso de Poisson y procesos de autorregresión o series de tiempo [17].

### H) Series de Tiempo

Una serie de tiempo es un conjunto de datos que se registran en intervalos de tiempos regulares (diario, semanal, semestral o anual). En ese orden de ideas, es una secuencia ordenada de observaciones sobre una variable en particular y su principal objetivo es realizar previsiones o pronósticos a partir de los datos pasados de la serie [18]. Las series de tiempo se pueden clasificar en series estacionarias o no estacionarias: las series estacionarias, son estables a través del tiempo, es decir la media y varianza del conjunto de datos es constante en el tiempo, por el contrario, en las series no estacionarias la

tendencia es variable a través del tiempo, por lo que la serie no oscila en un valor constante.

Formalmente, un modelo de serie de tiempo para los datos observados es una especificación de las distribuciones conjuntas (o las medias y covarianzas) de una secuencia de variables aleatorias de la cual se postula como un valor observado [19].

#### IV. PLANTEAMIENTO DEL PROBLEMA

Los arrecifes de coral abarcan en promedio el 25% de la fauna marina y son imprescindibles para los ecosistemas marinos dado que mantienen el equilibrio químico, bioquímico y biológico del océano; al funcionar como almacenes de carbonato de calcio, reduciendo los efectos de la acidificación del océano. Si bien los arrecifes de coral es uno de los ecosistemas más productivos del planeta también es uno de los más frágiles [28].

En los últimos años los arrecifes de coral han sufrido drásticos deterioros a tal punto que en la actualidad se considera que el 60% de los arrecifes están amenazados [29], amenaza de la cual los arrecifes de coral del Caribe colombiano no son ajenos. Según un estudio realizado por el INVEMAR, un tercio de los arrecifes coralinos del Caribe está amenazado por actividades humanas como la sedimentación, contaminación y sobrepesca.

A esto se debe agregar que la capacidad de los arrecifes de coral para soportar agentes de destrucción se ha visto debilitado por estas amenazas que a su vez generan enfermedades y transformaciones, provocando así una reducción de la cobertura de corales vivos de hasta un 50% [30].

Con lo anterior, cabe preguntarse: cómo predecir el cambio de cobertura de coral de los arrecifes de coral escleractinios en Santa Marta, Colombia, aplicando modelos de aprendizaje automático y modelos estocásticos para brindar un diagnóstico temprano y apoyar las medidas de manejo y conservación de estos ecosistemas.

#### V. METODOLOGÍA

CRISP-DM (Cross Industry Standard Process for Data Mining), es un marco que brinda una serie de pasos ordenados para la ejecución de procesos de minería de datos [31]. En la Figura 2 se puede observar las fases del modelo: análisis del negocio, comprensión de los datos, preparación de los datos, modelamiento, evaluación e implementación. Aclarando que el alcance de la presente investigación está dado hasta la fase de evaluación de los modelos.

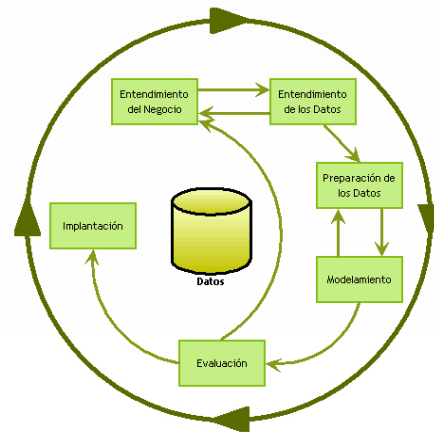


Figura 2. Metodología CRISP-DM.

Para esta investigación se toma como punto de partida la información histórica cobertura de coral escleractinios en los arrecifes de Santa Marta, Colombia. Esta información se recolectó inicialmente del portal público de datos “SiSMAC – Sistema de Soporte al Monitoreo de Áreas Coralinas” que recopila diferentes fuentes de datos sobre el estado de las especies marinas en Colombia luego el conjunto de datos fue completado en la variable fecha, latitud y longitud y código de la estación por el Instituto de Investigaciones del Mar - INVEMAR.

En cuanto al área de estudio, se selecciona un área geográfica de arrecifes en la zona oceánica de Caribe Colombiano (Figura 3), específicamente, estación del Chengue, Santa Marta.

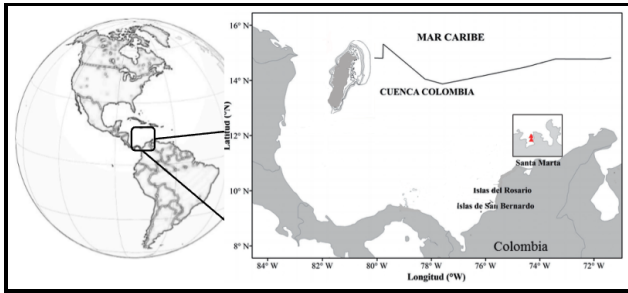


Figura 3. Área de estudio. Los triángulos señalan las estaciones del programa SIMAC- fuente de datos de arrecifes de coralinos.

## VI. RESULTADOS

A continuación se presenta la ejecución de las fases de la metodología CRISP-DM para la serie de tiempo de cobertura de coral escleractinio en el Chengue, Santa Marta, Colombia. Esta serie de tiempo comprende registros desde el año 1994 al año 2017.

### A. Comprensión del negocio

Con el fin de comprender la distribución de los organismos en los arrecifes de coral de las estaciones del chengue en Santa Marta, Colombia, se construyen los siguientes gráficos, donde en la Figura 4 se evidencia que la cobertura de los corales duros masivos oscila entre un 15% a un 25% en la estación número uno (1).

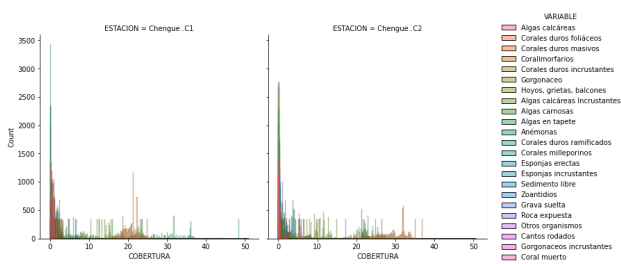


Figura 4. Recuento de organismos en estación 1 y estación 2, chengue, Santa Marta, Colombia.

### B. Comprensión de los datos

Teniendo en cuenta la información obtenida de las estaciones del Chengue de Santa Marta se analiza las variables del conjunto de datos:

TABLA I  
DESCRIPCIÓN VARIABLES

Variable	Descripción
Fecha	Fecha y hora de captura del registro
Área	Zona de observación y captura del registro. Para efectos de la investigación es Santa Marta.
Localidad	Lugar dentro del área de captura del registro. Para efectos de la investigación es Chengue.
Estación	Estación de captura del registro. Chengue 1 y Chengue 2.
Cod_Estacion	Identificador único de la estación.
Variable	Nombre de la especie que hace parte del arrecife de coral.
Cobertura	Porcentaje de cobertura de la especie en el arrecife de coral ( Corresponde a la variable objetivo de la investigación).
Año	Año de captura del registro.
Desv	Desviación estándar de la captura del registro.
Error	Error relativo de la precisión del instrumento que captura el registro.
Latitud	Distancia en grados, minutos y segundos del punto de captura del registro al paralelo del ecuador.
Longitud	Distancia en grados, minutos y segundos del punto de captura del registro al meridiano de Greenwich.

Inicialmente, se evidencia que el conjunto de datos recopila información sobre la cobertura de diferentes especies y microorganismos que conforman el arrecife de coral. Teniendo en cuenta que el alcance de la presente investigación es predecir la cobertura de coral escleractinio, resulta necesario filtrar el conjunto de datos sobre la especie correspondiente, que para el caso es: “coral duro masivo”.

Siguiendo con el análisis, se encontró que el conjunto de datos presenta datos faltantes por lo cual es necesario

evaluar si hay algún patrón o dependencia en estos. En las siguientes gráficas podemos observar que los datos faltantes corresponden al año 2007, 2012 y del 2014 al 2016 por lo que es necesario realizar una imputación de datos para el desarrollo del modelo (ver Figuras 5, 6 y 7).

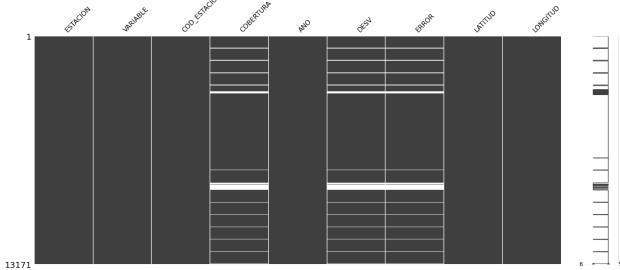


Figura 5. Valores perdidos o faltantes para la especie “Corales duros masivos”.

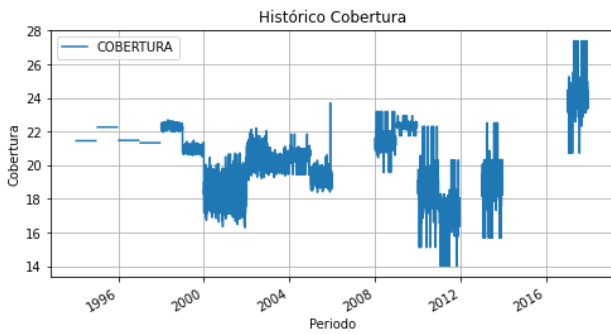


Figura 6. Serie de tiempo cobertura de coral escleractinio en la estación no. 1 Chengue, Santa Marta, Colombia.

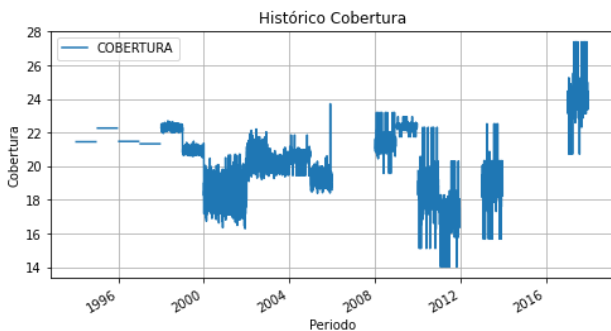


Figura 7. Serie de tiempo cobertura de coral escleractinio en la estación no. 2 Chengue, Santa Marta Colombia.

Finalmente, para determinar la existencia de datos atípicos, se genera el diagrama de cajas o bigotes para cada estación (Figura 8), evidenciando que existen valores fuera de la media para la estación uno (1) y dos (2).

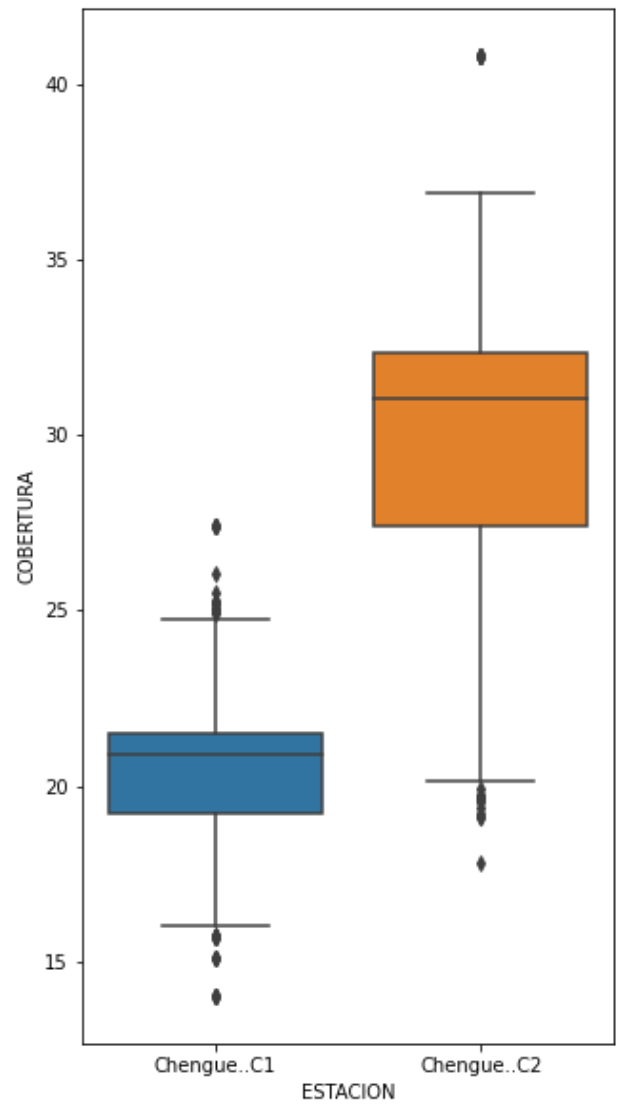


Figura 8. Diagrama de cajas de la cobertura de coral escleractinio en la estación no. 1 y estación no. 2 Chengue, Santa Marta, Colombia.

En cuanto a la densidad de los datos mostrados en las Figura 9, y 10 se identifica que en la estación uno (1), la cobertura de coral oscila entre el 12.5% al 27.5% con un pico en 27.5. En cuanto a la estación dos (2), la cobertura de coral oscila entre el 8.3% al 37.5% con un pico en 37.5.

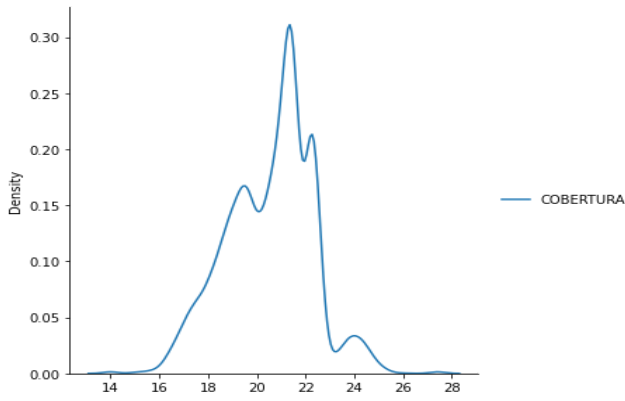


Figura 9. Densidad de cobertura de coral en la estación no. 1 del Chengue.

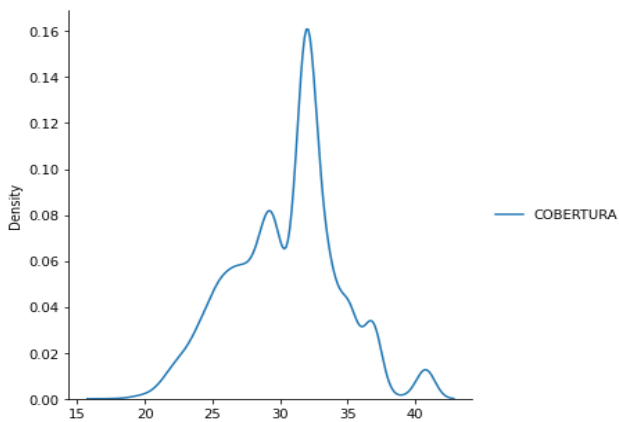


Figura 10. Densidad de cobertura de coral en la estación no. 2 del Chengue.

### C. Preparación de los datos

En el apartado anterior, se encontraron registros nulos o sin medición, por lo que es necesario aplicar un método de imputación para los datos faltantes.

Para seleccionar el método de imputación que permita modelar la serie de tiempo, se evalúan tres métodos de imputación para cada estación: interpolación lineal (Ver Figura 11 y 12), valor siguiente (forwardfill) (Ver Figura 12 y 15) e interpolación del valor observable más cercano (Ver Figura 13 y 16) donde se promedia el último valor válido antes del valor perdido y el primer valor válido después del valor perdido. En los siguientes gráficos se visualiza el comportamiento de cada uno de los métodos sobre la variable cobertura en la estación no.1 y estación no. 2 del Chengue.

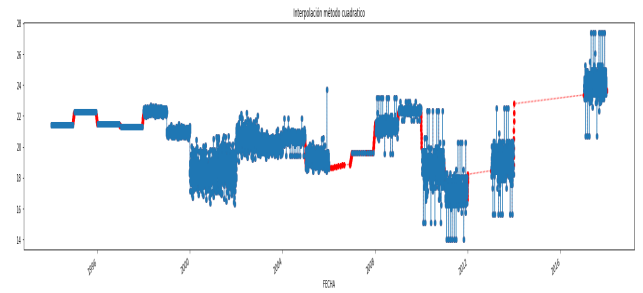


Figura 11. Interpolación lineal de la serie temporal de cobertura de coral en la estación 1 del Chengue, Santa Marta, Colombia.

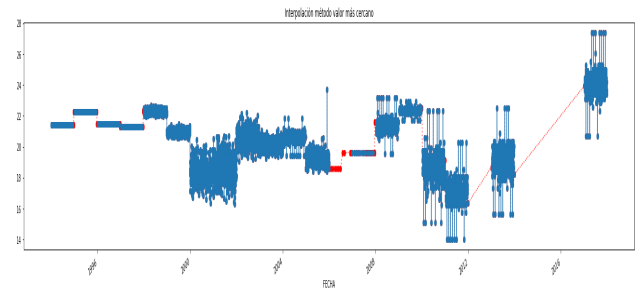


Figura 12. Imputación método forwardfill de la serie temporal de cobertura de coral en la estación 1 del Chengue, Santa Marta, Colombia.

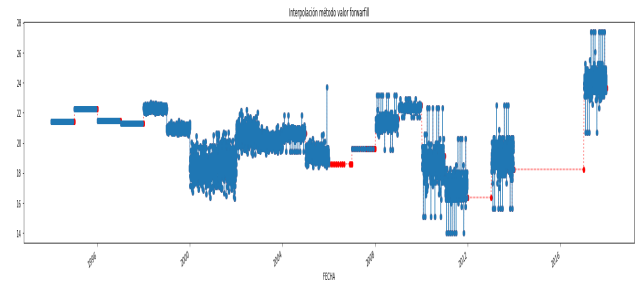


Figura 13. Imputación método nearest de la serie temporal de cobertura de coral en la estación 1 del Chengue, Santa Marta, Colombia.

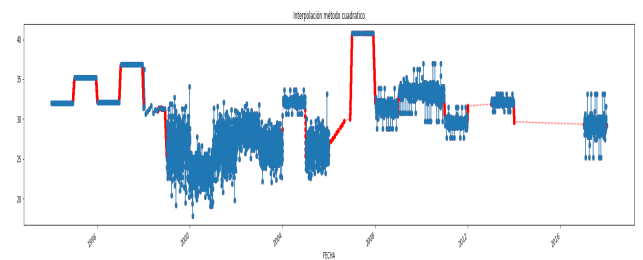


Figura 14. Interpolación lineal de la serie temporal de cobertura de coral en la estación 2 del Chengue, Santa Marta, Colombia.

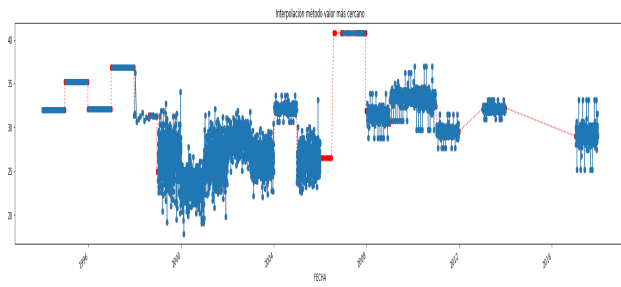


Figura 15. Imputación método forwardfill de la serie temporal de cobertura de coral en la estación 2 del Chengue, Santa Marta, Colombia.

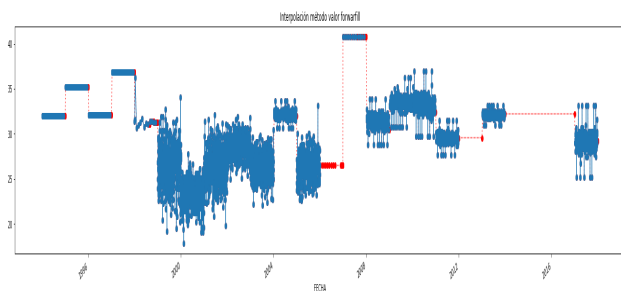


Figura 16. Imputación método nearest de la serie temporal de cobertura de coral en la estación 2 del Chengue, Santa Marta, Colombia

Teniendo en cuenta el factor de estacionalidad de los datos y que la cantidad de valores nulos sea la menor posible, se selecciona el método de imputación de interpolación lineal. Con este, se observa que los valores que reemplazan a los datos perdidos buscan mantener la estacionalidad de la serie de tiempo.

Con la serie de datos completa, se obtiene la media, valor mínimo y máximo para cada estación (Ver Tabla II), encontrando que para la estación uno (1) la media de la cobertura de coral es de 20.5%, el valor mínimo es de 14% y el máximo es 27%; en cuanto a la estación número dos (2) la media, valor mínimo y máximo son de 30.5%, 17.8% y 40,8%, respectivamente.

TABLA II

ESTADISTICAS SERIE TEMPORAL COBERTURA DE CORAL POR ESTACIÓN EN EL CHENGUE, SANTA MARTA, COLOMBIA

	ESTACION 1	ESTACION 2
<b>count</b>	6.690	6479
<b>mean</b>	20,53	30,5
<b>std</b>	1,76	3,981
<b>min</b>	14,0041	17,8
<b>25%</b>	19,26	27,69
<b>50%</b>	20,84	31,37
<b>75%</b>	21,47	32,603
<b>95%</b>	23,19	36,864
<b>99%</b>	24,45	40,805
<b>max</b>	27,4	40,805

Teniendo en cuenta los valores de la tabla anterior y las figuras 8, 9 y 10, donde se evidencia que la densidad y la distribución de la estación número uno (1) tiende a ser normal respecto a la estación número dos (2), se escoge la serie temporal de la estación número uno (1) para el desarrollo de las siguientes fases. A continuación, se presenta la serie temporal de la cobertura de coral escleractinio desde el año 1994 hasta el año 2017, para la estación número uno (1) del Chengue, Santa Marta, Colombia (Ver Figura 17).

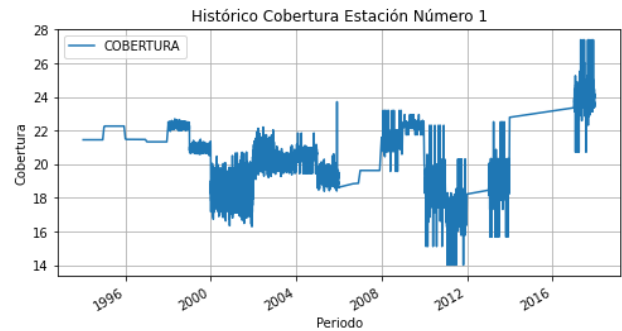


Figura 17. Serie de tiempo estación 1 después del proceso de imputación.

En la Figura 18 se descompone la serie temporal, evidenciando que esta tiende a crecer y también posee valores residuales con una media alrededor de cero (0).

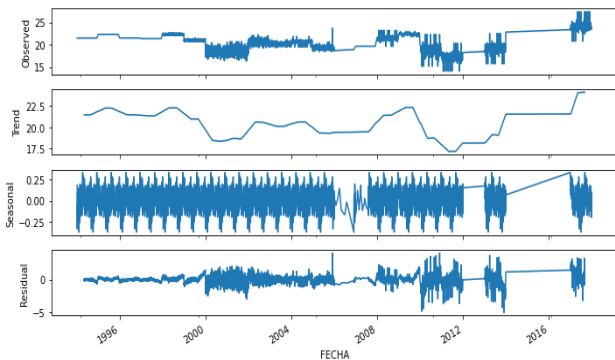


Figura 18. Descomposición detallada serie temporal

Fuente. Elaboración propia

#### D. Modelado y Evaluación

A partir del análisis y calibración de los datos, se seleccionan y aplican dos técnicas de modelado: ARIMA y Red Neuronal LSTM que tienen como característica principal pronosticar valores futuros con base a valores históricos. La técnica ARIMA al ser un proceso estocástico permite tratar a la serie de tiempo como un proceso aleatorio y no determinístico, para la predicción del futuro. Las redes neuronales LSTM, son un tipo de red neuronal recurrente especial porque tiene la capacidad de recordar información por periodos largos de tiempo.

Para modelar la serie de tiempo de la cobertura de coral escleractinio en el Chengue, Santa Marta, Colombia, es necesario validar la estacionalidad de la serie. Por lo que se aplica la prueba de Dickey-Fuller Aumentada (ADF), que consiste en determinar la estacionalidad de la serie de acuerdo al p-value obtenido y los niveles de confianza, para 1%, 5% y 10%. A continuación, se presentan los resultados obtenidos:

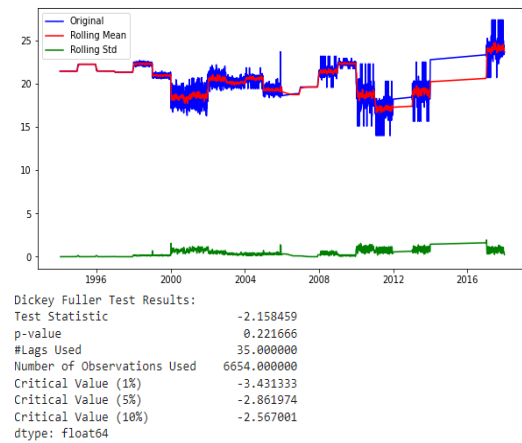


Figura 19. Aplicación prueba Dickey-Fuller a serie temporal de cobertura de coral duro masivo estación 1 - Chengue, Santa Marta, Colombia.

Fuente. Elaboración propia.

De la Figura 19 se observa que si bien la serie de tiempo es estacional el valor de p-value no es menor al 5% por lo que es necesario diferenciar la serie para corregir la estacionalidad en la varianza. En la siguiente figura se presentan los resultados de la diferenciación aplicada a un grado sobre la serie temporal.

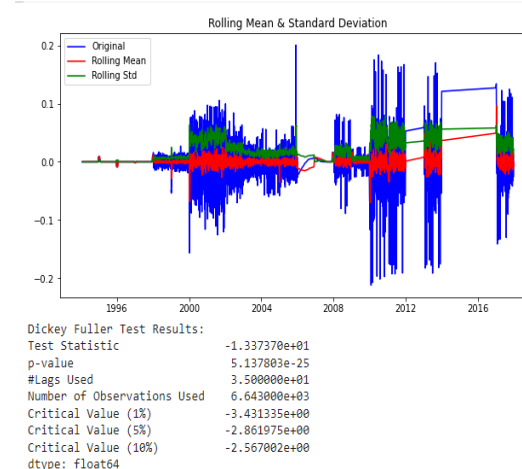


Figura 20. Serie temporal de cobertura de coral masivo diferenciada a un grado.

Fuente. Elaboración propia.

De la Figura 20, se identifica que se cumple el supuesto de estacionalidad dado que el p-value es menor al 5% y los valores críticos cercanos a los niveles de confianza. Una vez generada la estacionalidad de la serie se puede iniciar el modelamiento ARIMA para predecir la cobertura de coral escleractinio en el Chengue, Santa Marta, Colombia.

1. ARIMA

Tras validar la estacionalidad de la serie, se aplica un modelo tipo ARIMA desarrollado por Box y Jenkins. Este modelo está conformado por tres componentes: AR, I y MA, donde el componente AR valida la parte regular mientras que el componente MA evalúa la estacionalidad de la misma [28].

Para encontrar los valores p,d,q más apropiados, se valida la tendencia de la serie temporal, generando el gráfico de autocorrelación (Figura 22) y autocorrelación parcial (Figura 21).

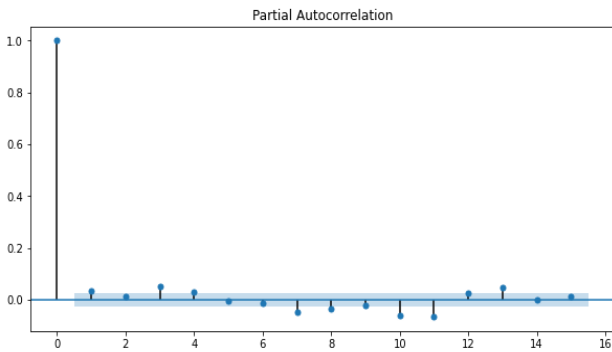


Figura 21. Autocorrelación parcial serie de tiempo cobertura de coral estación 1, Chengue, Santa Marta, Colombia.

Fuente. Elaboración propia.

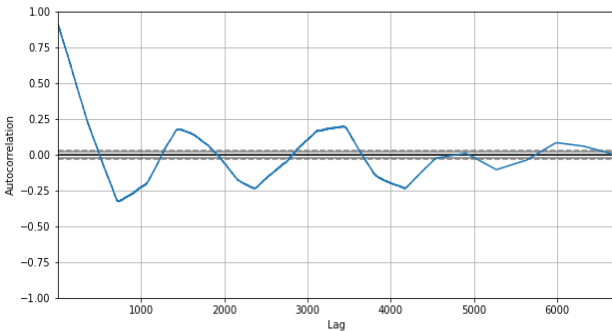


Figura 22. Autocorrelación serie de tiempo cobertura de coral estación 1, Chengue, Santa Marta, Colombia.

Fuente. Elaboración propia.

En la Figura 22, se evidencia que la función de autocorrelación de la serie disminuye sensiblemente a medida que aumenta el desfase temporal donde los en 4, 5 y 6 se empieza a notar la estacionalidad de la serie diferenciada.

Considerando los valores visualizados en las anteriores gráficas para el modelo AR y MA, se realiza la estimación de los parámetros que mejor se adecuan a la serie temporal de cobertura de coral. En la estimación se debe tener en cuenta que los modelos ARIMA consideran los p,d,q y P,D,Q,m, donde:

- p: número de rezagos del modelo o el orden del modelo.
- q: orden del promedio móvil.
- d: grado de diferenciación.
- m: número de periodos que tiene cada estación.

Para la estimación de parámetros AR se calibra el número de orden inicial del modelo “p” en cinco (5) dado que es el valor donde inicia la estabilización de ACF (Figura 23); el grado de diferenciación (d) se calibra en uno (1) teniendo en cuenta que el supuesto de estacionalidad realizado en el apartado anterior, se cumplió al aplicar un grado de diferenciación y, el orden del promedio móvil (q) se calibra en dos (2) dado que es el valor más alto.

ARIMA Model Results						
Dep. Variable:	COBERTURA	No. Observations:	6679			
Model:	ARMA(5, 3)	Log Likelihood	14745.881			
Method:	css-mle	S.D. of innovations	0.827			
Date:	Thu, 26 May 2022	AIC	-29471.763			
Time:	05:02:46	BIC	-29403.695			
Sample:	0	HQIC	-29448.252			

	coef	std err	z	P> z	[0.025	0.975]
const	8.025e-05	0.000	0.229	0.819	-0.001	0.001
ar.L1.COBERTURA	1.5717	1.111	1.415	0.157	-0.605	3.749
ar.L2.COBERTURA	-0.7689	1.856	-0.414	0.679	-4.407	2.870
ar.L3.COBERTURA	-0.0280	1.039	-0.027	0.979	-2.065	2.009
ar.L4.COBERTURA	-0.0341	0.076	-0.446	0.655	-0.184	0.116
ar.L5.COBERTURA	-0.0009	0.044	-0.021	0.983	-0.087	0.085
ma.L1.COBERTURA	-1.5465	1.111	-1.392	0.164	-3.724	0.630
ma.L2.COBERTURA	0.7356	1.828	0.402	0.687	-2.848	4.319
ma.L3.COBERTURA	0.0914	0.999	0.091	0.927	-1.867	2.050

Figura 23. Modelo ARIMA (5,1,3).

Fuente. Elaboración propia.

Luego de generar tres (3) modelos se obtiene el valor máximo log-likelihood y el coeficiente de AIC para cada uno:

ARIMA Model Results						
Dep. Variable:	D.COBERTURA	No. Observations:	6689			
Model:	ARIMA(5, 1, 3)	Log Likelihood	-5578.575			
Method:	css-mle	S.D. of innovations	0.557			
Date:	Fri, 27 May 2022	AIC	11177.150			
Time:	14:00:18	BIC	11245.233			
Sample:	1	HQIC	11200.664			

Figura 24. Modelo ARIMA (5,1,3).

Fuente. Elaboración propia..

```

-----
                ARIMA Model Results
-----
Dep. Variable:    D.COBERTURA    No. Observations:    6689
Model:           ARIMA(3, 1, 3)    Log Likelihood       -5579.264
Method:          css-mle          S.D. of innovations   0.557
Date:            Fri, 27 May 2022    AIC                  11174.527
Time:            13:13:44          BIC                  11228.993
Sample:          1                 HQIC                 11193.338
    
```

Figura 25. Modelo ARIMA (3,1,3).

Fuente. Elaboración propia.

```

-----
                ARIMA Model Results
-----
Dep. Variable:    D.COBERTURA    No. Observations:    6689
Model:           ARIMA(2, 1, 2)    Log Likelihood       -5579.919
Method:          css-mle          S.D. of innovations   0.557
Date:            Fri, 27 May 2022    AIC                  11171.838
Time:            13:13:38          BIC                  11212.688
Sample:          1                 HQIC                 11185.947
    
```

Figura 26. Modelo ARIMA (2,1,2).

Fuente. Elaboración propia.

De estos resultados, se observa que el modelo que maximiza el log-likelihood y minimiza el coeficiente AIC es el modelo ARIMA (5,1,3) calibrado con un orden (p) cinco (5) y orden del promedio móvil 2. De este modelo también se evidencia que los dos primeros resultados tuvieron un mejor rendimiento dado que el valor de z en la prueba P es menor a 0.05 y que la constante tiene mayor peso en las predicciones del modelo.

Finalmente, se analizan los residuos para el supuesto de aleatoriedad e independencia de los errores del modelo. A continuación se presentan los residuales del modelo:

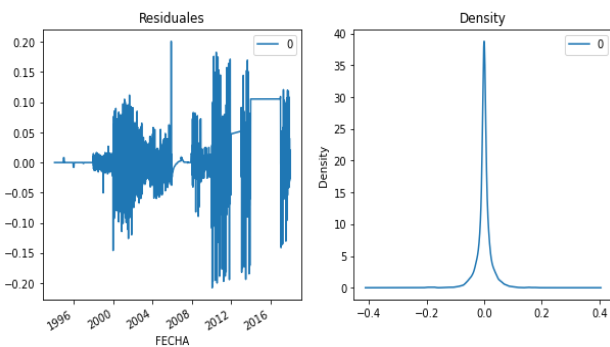


Figura 27. Residuales del modelo.

De la Figura 27 se observa que se cumple el supuesto de aleatoriedad de los residuales dado que la media de la densidad de los datos tiende a ser constante en 0.

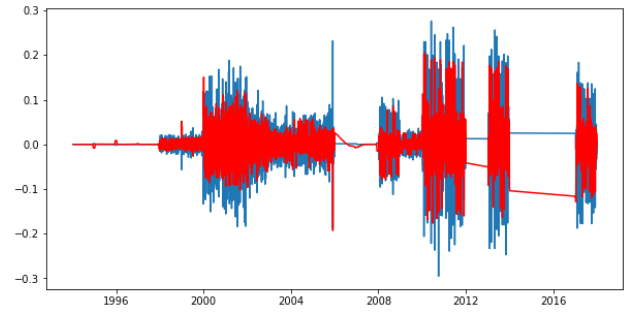


Figura 28. Predicción de la serie vs datos reales.

Las pruebas y predicciones del modelo se realizan sobre el conjunto de validación. En la Figura 28 se observa los valores originales comparados con los pronosticados obteniendo un error cuadrático medio MSE de 5,56 (RMSE).

## 2. Red Neuronal Recurrente LSTM

Se genera una Red Neuronal Recurrente LSTM que aprende de los datos anteriores, permitiendo pronosticar valores futuros. Para entrenar el modelo se crea el conjunto de datos de entrenamiento entre el año 1994 al 2010 y el conjunto de datos de validación desde el año 2010 hasta el año 2017.

En el modelo se busca predecir un (1) valor futuro con base a 60 registros consecutivos por lo que se genera un lote con bloques de 60 datos. En la construcción del modelo también se genera una lista vacía para almacenar cada valor predicho.

La red neuronal está compuesta por 50 neuronas y un total de 20 iteraciones (epoch). En cuanto al entrenamiento se usa el método RMPRP, el cual funciona de manera similar al método de gradiente descendiente ver Figura 29.

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 50)	10400
dense (Dense)	(None, 1)	51

---

Total params: 10,451  
Trainable params: 10,451  
Non-trainable params: 0

Figura 29. Arquitectura red neuronal LSTM.

Se generó la validación del modelo con el conjunto de datos validación. A continuación se presenta la Figura 30 en la cual se muestran los valores predichos comparados con los valores reales, evidenciando que los valores predichos siguen la tendencia de los valores reales.

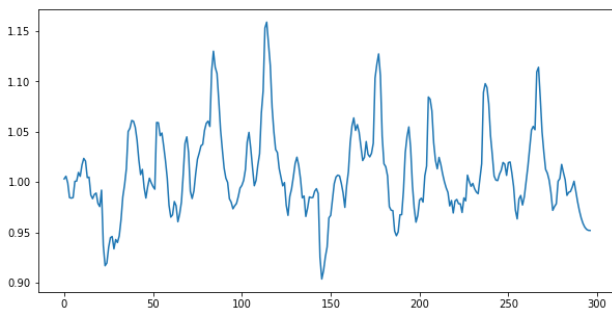


Figura 30 Predicciones del modelo con el conjunto de validación.

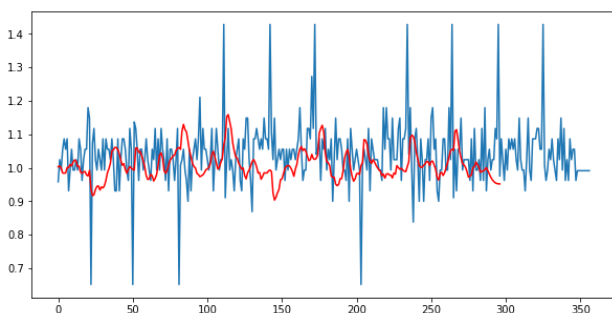


Figura 31. Predicción de la serie vs datos reales.

En las gráficas anteriores se observan los valores originales (Ver Figura 31, valores en color azul) comparados con los pronosticados (Ver figura 30, valores en color rojo) obteniendo un error cuadrático medio (RMSE) de 0,725.

## VII. CONCLUSIONES

La presente investigación realiza un análisis de los datos históricos de la cobertura de coral escleractinio en el Chengue, Santa Marta, Colombia, donde se encontraron datos faltantes para los años 2007, 2012, 2014, 2015 y 2016, por lo que se vio necesario la aplicación de métodos de imputación. De igual forma, se identificó la presencia de datos atípicos en la estación número dos (2) del Chengue.

En la validación de la estacionalidad, se encontró que la serie diferenciada a un grado se vuelve estacionaria dado que arrojó resultados satisfactorios para el p-value y los niveles de confianza, permitiendo así modelar los datos para predecir la variable objetivo de la investigación.

De igual manera, en la generación del modelo ARIMA se encontró que los parámetros que maximizan el log el log-likelihood y minimizan el coeficiente AIC son  $p=5$ ,  $d=1$  y  $q=2$ .

En la evaluación de los modelos generados, se validó con respecto al error cuadrático medio (RMSE) donde se evidenció que la red neuronal LSTM tuvo un mejor desempeño comparado con el modelo ARIMA, obteniendo un RMSE de 0.72 y 5.2, respectivamente. Por tal motivo, se escoge la red neuronal LSTM, para la predicción de cobertura de coral escleractinio en el Chengue, Santa Marta.

Si bien, los valores del RMSE para la red neuronal son significativamente más bajos que el modelo ARIMA, podría optimizarse aún más. De esta manera, la presente investigación da pie y son el punto de partida para nuevas investigaciones que busquen aplicar este tipo de técnicas en las diferentes especies que conforman los arrecifes de coral en las estaciones de monitoreo marino en Colombia.

## VIII. REFERENCIAS

- [1] J. Reyes, "Lista de los Corales (Cnidaria: Anthozoa: Scleractinia) de Colombia," *Biota Colomb.*, vol. 1, no. 2, pp. 164–176, 2000, doi: 10.21068/bc.v1i2.66.

- [2] A. Rodríguez-Ramírez et al., “Recent dynamics and condition of coral reefs in the Colombian Caribbean,” *Rev. Biol. Trop.*, vol. 58, pp. 107–131, 2010.
- [3] N. G. Rangel-Buitrago, G. Anfuso, and A. T. Williams, “Coastal erosion along the Caribbean coast of Colombia: magnitudes, causes and management,” *Ocean Coast. Manag.*, vol. 114, pp. 129–144, 2015.
- [4] A. M. Rodríguez Rincón, S. M. Navarrete Ramírez, D. I. Gómez López, and R. Navas Camacho, “Protocolo Indicador Condición Tendencia Áreas Coralinas (ICTAC). Indicadores de monitoreo biológico del Subsistema de Áreas Marinas Protegidas (SAMP).” 2014.
- [5] M. Spalding, M. D. Spalding, C. Ravilious, and E. P. Green, *World atlas of coral reefs*. Univ of California Press, 2001.
- [6] A. C. Castillo Serrano, “Descripción de microorganismos asociados a enfermedades coralinas con base en técnicas moleculares-Revisión Bibliográfica,” 2020.
- [7] D. R. Bellwood, R. P. Streit, S. J. Brandl, and S. B. Tebbett, “The meaning of the term ‘function’ in ecology: a coral reef perspective,” *Funct. Ecol.*, vol. 33, no. 6, pp. 948–961, 2019.
- [8] S. Schmidt-Roach, K. J. Miller, and N. Andreakis, “*Pocillopora aliciae*: a new species of scleractinian coral (Scleractinia, Pocilloporidae) from subtropical Eastern Australia,” *Zootaxa*, vol. 3626, no. 4, pp. 576–582, 2013.
- [9] J. L. Cornejo-Ortega and R. M. Chávez-Dagostino, “Implicaciones en la observación de la ballena jorobada,” *Investig. COSTERAS*, p. 143, 2014.
- [10] C. M. Yonge, “Coral Reef Project—Papers in Memory of Dr. Thomas F. Goreau. 1. The Nature of Reef-Building (Hermatypic) Corals,” *Bull. Mar. Sci.*, vol. 23, no. 1, pp. 1–15, 1973.
- [11] A. Harris, G. Manahira, A. Sheppard, C. Gouch, and C. Sheppard, “Demise of Madagascar’s once great barrier reef: changes in coral reef conditions over 40 years,” *Atoll Res. Bull.*, 2010.
- [12] N. Stambler, “Zooxanthellae: the yellow symbionts inside animals,” in *Coral Reefs: an ecosystem in transition*, Springer, 2011, pp. 87–106.
- [13] V. Fonseca González and D. A. García Bonilla, “Análisis geográfico de la pérdida de cobertura coralina en el Parque Nacional Natural Corales del Rosario y San Bernardo,” 2019.
- [14] J. Torres, “Deep learning,” *Introd. práctica con Keras*, Barcelona Lulu. com, 2018.
- [15] E. Alpaydin, *Machine Learning: The New AI*. Cambridge, MA: The MIT Press, 2016.
- [16] S. Ohlsson, *Deep Learning: How the Mind Overrides Experience*. Cambridge: Cambridge University Press, 2011.
- [17] S. M. Ross et al., *Stochastic processes*, vol. 2. Wiley New York, 1996.
- [18] J. Villavicencio, “Introducción a series de tiempo,” Puerto Rico, 2010.
- [19] P. J. Brockwell, P. J. Brockwell, R. A. Davis, and R. A. Davis, *Introduction to time series and forecasting*. Springer, 2016.
- [20] J. R. Ward, K. Kim, and C. D. Harvell, “Temperature affects coral disease resistance and pathogen growth,” *Mar. Ecol. Prog. Ser.*, vol. 329, pp. 115–121, 2007.
- [21] S. H. Sokolow, P. Foley, J. E. Foley, A. Hastings, and L. L. Richardson, “Editor’s choice: Disease dynamics in marine metapopulations: modelling infectious diseases on coral reefs,” *J. Appl. Ecol.*, vol. 46, no. 3, pp. 621–631, 2009.
- [22] C. L. G. González, “Un modelo Estocástico para arrecifes de coral,” *Pädi Boletín Científico Ciencias Básicas e Ing. del ICBI*, vol. 4, no. 8, 2017.
- [23] G. J. Williams, G. S. Aeby, R. O. Cowie, and S. K. Davy, “Predictive Modeling of Coral Disease Distribution within a Reef System,” *PLoS One*, vol. 5, no. 2, p. 9264, 2010, doi: 10.1371/journal.pone.0009264.
- [24] Z. Yang, T. Zhang, and D. Zhang, “A novel algorithm with differential evolution and coral reef optimization for extreme learning machine training,” doi: 10.1007/s11571-015-9358-9.
- [25] J. M. Pandolfi, “Incorporating uncertainty in predicting the future response of coral reefs to climate change,” *Annu. Rev. Ecol. Evol. Syst.*, vol. 46, pp. 281–303, 2015.

- [26] C. Castro-Sanguino et al., "Reef state and performance as indicators of cumulative impacts on coral reefs," *Ecol. Indic.*, vol. 123, p. 107335, 2021.
- [27] P. J. Mumby, A. Hastings, and H. J. Edwards, "Thresholds and the resilience of Caribbean coral reefs," *Nature*, vol. 450, no. 7166, pp. 98–101, 2007, doi: 10.1038/nature06252.
- [28] G. Díaz Pulido, "Informe nacional sobre el estado de la biodiversidad en Colombia: ecosistemas marinos y costeros," 1997.
- [29] L. Burke, K. Reytar, M. Spalding, and A. Perry, *Reefs at risk revisited*. World Resources Institute, 2011.
- [30] P. Jørgensen, "Amenazas y esperanzas para los arrecifes de coral," *Ecofronteras*, pp. 22–25, 2008.
- [31] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining," in *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, 2000, vol. 1.
- [32] E. O. P. Sánchez, "El principio antrópico," *Trib. Astron. Rev. Astron. astrofísica y ciencias del Espac.*, no. 16, pp. 68–73, 2000.
- [33] T. Thomas, P. Maurya, B. Manikandan and N. F. Desai. A Survey on ML-Based Techniques To Estimate Coral Reef Cover: state of the art and Challenges. 2022.
- [34] M. G. Corruz, E. Filipina, M. J. Santiago, S. M. Uy, C. Lazana and A. Bandala, "BahurApp: Development And Implementation Of Coral Bleaching Monitoring Application Using Convolutional Neural Network," 2021.
- [35] T. Kim, Y. Kwon, J. Lee, E. Lee, and S. Kwon, S. Wave attenuation prediction of artificial coral reef using machine-learning integrated with hydraulic experiment. *Ocean Engineering*, 248, 110324, 2022.