

# REVISTA INGENIERÍAS

UNIVERSIDAD DE MEDELLÍN

19(37), julio-diciembre de 2020



**Universidad<sup>®</sup>  
de Medellín**  
Ciencia y Libertad



# REVISTA INGENIERÍAS

UNIVERSIDAD DE MEDELLÍN

Volumen 19, número 37, julio/diciembre de 2020

ISSN (en línea): 2248-4094

ISSN (impreso): 1692-3324

## Fredy López Pérez

Editor

[jflopez@udem.edu.co](mailto:jflopez@udem.edu.co)

## Carlos Andrés Arredondo

Coeditor

[caaredondo@udem.edu.co](mailto:caaredondo@udem.edu.co)

## Laura Agudelo Calle

Asistente Editorial

[asist\\_ingenierias@udem.edu.co](mailto:asist_ingenierias@udem.edu.co)

Sello Editorial Universidad de Medellín

[selloeditorial@udem.edu.co](mailto:selloeditorial@udem.edu.co)

Carrera 87 N.º 30-65. Bloque 20, piso 2

Teléfono: (+57+4) 340 52 42

Medellín, Colombia

<https://www.udem.edu.co/index.php/servicios/sello-editorial>

## Solangy Carrillo-Pineda

Coordinación editorial

## Jorge I. Escobar

Corrección de estilo

## 2LA2 traducción y edición

Traducción al portugués

## Álvaro Hernán Cruz Mejía

Traducción al inglés

## Infolio. Corrección de textos

Revisión de prueba diagramada

## Leonardo Sánchez Perea

Diagramación

## Claudia Castrillón

Diseño de cubierta

## Foto de carátula

xxxxxx por Dr. Fredy López-Pérez

## COMITÉ EDITORIAL

Editor Dr. Fredy Lopez-Perez, Universidad de Medellín, Colombia

Coeditor, Dr. Carlos Andrés Arredondo, Universidad de Medellín, Colombia

Dra. Bell Manrique Losada, Universidad de Medellín, Colombia

Dra. Ximena Gaviria Gómez, Universidad de Medellín, Colombia

Dra. Diana Isabel Quevedo-Tejada, Universidad de Santiago de Chile, Chile

Dr. Paulo B. Lourenço, Universidad de Minho, Portugal

Dr. Harveth Hemán Gil Sánchez, Politécnico Colombiano Jaime Isaza Cadavid, Colombia

Dr. Alejandro Zuleta Gil, Universidad Pontificia Bolivariana, Colombia

Dr. Martín Solari Buella, Universidad ORT Uruguay, Uruguay

Dra. Marta Silvia Tabares Betancur, Eafit, Colombia

Dr. Mauricio González Betancourt, Tecnológico de Antioquia – Institución Universitaria, Colombia

Dra. Ana María Cárdenas Soto, Universidad de Antioquia, Colombia

Dr. Mauricio Hincapié, Institución Universitaria Salazar y Herrera, Colombia

Dr. Pedro M. Wightman, Universidad del Norte, Colombia

Dra. María Guadalupe Valladares-Cisneros, Universidad Autónoma del Estado de Morelos, México

Dr. Germán Urrego Giraldo, Universidad de Antioquia, Colombia

Dr. José Gallardo Arancibia, Universidad Católica del Norte, Chile

Dr. Jordi Morató, Universidad Politécnica de Cataluña, España

Dr. Oscar Pastor Lopez, Universidad Politécnica de Valencia, España

Dr. Damià Vericat, Universidad de Lleida, España

Dr. Juan Bosco Hernández, Universidad Autónoma de Querétaro, México

## COMITÉ CIENTÍFICO

Dra. Alice Elizabeth González, Universidad de la República, Uruguay

Dr. Claudio Meneses Villegas, Universidad Católica del Norte, Chile

Dr. José Luis Rosúa Campos, Universidad de Granada, España

Dr. Guillermo Franco, Catastrophe Risk Research - EMEA, Guy Carpenter, Reino Unido

Dr. Jorge A Ramirez, Colorado State University, Estados Unidos

Dr. André Pacheco de Assis, Universidade de Brasília, Brasil

Dr. Juan Diego Jaramillo Fernández, Universidad EAFIT, Colombia

Dr. Luis Joyanes Aguilar, Escuela Superior de Ingeniería y Arquitectura, Universidad Pontificia de Salamanca, España

Dra. Teresa López Lara, Universidad Autónoma de Querétaro, México

Canje, correspondencia y suscripciones: Biblioteca de Facultades “Eduardo Fernández Botero” Universidad de Medellín

Teléfonos: 340 5252- 340 5276. Fax (57-4) 345 5216

Correo electrónico: [rbarrientos@udem.edu.co](mailto:rbarrientos@udem.edu.co), [revistaingenierias@udem.edu.co](mailto:revistaingenierias@udem.edu.co)

Sitio web: <https://revistas.udem.edu.co/index.php/ingenierias>

Medellín, Colombia, Suramérica

Impresión

Xpress Estudio Gráfico y Digital

Bogotá, Colombia

Fecha de impresión: diciembre de 2020

Tiraje: 200 ejemplares

Nuestra publicación ha sido admitida en:

- Índice Bibliográfico Nacional de Publicaciones Seriadadas Científicas y Tecnológicas, Publindex, Minciencias, clasificación B
- Scientific Electronic Library Online –SciELO- Colombia
- Índice de Revistas Latinoamericanas en Ciencias-Periódica, Universidad Nacional Autónoma de México
- Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal-Latindex
- Bases de datos “Fuente Académica”, “Fuente Académica Premier” y “Academic Search Complete” de Ebsco Information Services
- Red de Revistas Científicas de América Latina, el Caribe, España y Portugal-RedAlyC- [redalyc.org](https://www.redalyc.org/)
- Base de datos “Scifinder” de Chemical Abstracts Service
- Base de datos “Informe Académico” de Gale Cengage Learning
- Dialnet. Red de revistas científicas de investigación de habla hispana
- Directory of Open Access Journals -DOAJ
- ISI-Web of Science, Scielo Citation Index

Queda autorizada la reproducción total o parcial de los contenidos de la revista con finalidades educativas, investigativas o académicas siempre y cuando sea citada la fuente. Para poder efectuar reproducciones con otros propósitos, es necesario contar con la autorización expresa del Sello Editorial Universidad de Medellín. Las ideas, contenidos y posturas de los artículos son responsabilidad de sus autores y no comprometen en nada a la institución ni a la revista.

## Presentación

*Revista Ingenierías Universidad de Medellín* es una publicación multidisciplinar para ingeniería. La revista se orienta a discutir problemas clave para el desarrollo regional y global desde la mirada de la ingeniería.

Por su historia, la revista ha cubierto fundamentalmente los intereses de las ingenierías ambiental, civil, de sistemas y financiera. De manera reciente los escritos asociados a las ingenierías industrial, administrativa, eléctrica y electrónica han tenido una presencia constante y relevante. Asimismo, la revista está abierta a la difusión en cualquier otro campo de la ingeniería, lo cual permite darles a muchos de sus números un enfoque temático específico.

La revista prioriza la recepción de producción derivada de proyectos de investigación para hacer su adecuada divulgación entre académicos, estudiantes y grupos de interés en la sociedad.

Desde *Revista Ingenierías Universidad de Medellín* declaramos la necesidad de que el conocimiento científico circule y se apropie en nuestra región y aporte en el análisis y solución de problemas específicos. No obstante, reconocemos también la necesidad de que nuestros contenidos circulen del modo más amplio posible entre comunidades amplias de académicos y por tal razón, a partir de 2021 nuestra revista publica sus contenidos exclusivamente en inglés.

La revista garantiza la ética en el tratamiento de sus contenidos y en sus procedimientos que involucran a autores, evaluadores y editores por los estándares que acoge tras su adhesión a COPE.





## Presentation

*Revista Ingenierías Universidad de Medellín* is a multidisciplinary publication in the Engineerings field. The journal is oriented towards discussing key problems for regional and global development from an engineering perspective. Because of its history, the journal has covered the interests of environmental, civil, systems and financial engineering. Increasingly, those pieces of writings on industrial, administrative, electric and electronic engineering have been constantly present and with evermore greater relevance. Likewise, the journal is open to broadcasting any other field in engineering, which allows for most of its issues a specific topic approach.

The journal prioritizes the reception of productions derived from research projects in order to perform adequate broadcasting among academics, students and other groups of interest.

As *Revista Ingenierías Universidad de Medellín* we declare the need for the scientific knowledge to be broadcasted and appropriated in our region to contribute to the analysis and solution of specific problems. Nonetheless, we also recognize the need for our contents to be broadcasted in the wider way possible among even wider academic communities and for that reason, from 2021 on, our journal will publish its contents exclusively in English.

The journal guarantees ethical conduct in the treatment of its contents and also within its procedures regarding authors, reviewers and editors by the standards applied after its adherence to the COPE.



# Contenido

Editorial .....	9
<i>Editorial (English)</i> .....	10
<i>Editorial (Português)</i> .....	11
Estimación de la captura de carbono en un relicto de bosque urbano mediante técnicas de teledetección.....	13
<i>Estimation of Carbon Capture in an Urban Forest Relict through Teledetection Techniques</i> <i>Claudia Marcela Cardona Lindo, Julián Garzón Barrero y Gonzalo Jiménez Cleves</i>	
Testbed for Sub-Nyquist Wideband Spectrum Monitoring.....	35
<i>Banco de pruebas para monitoreo Sub-Nyquist de espectro de banda ancha</i> <i>Juan Felipe Medina Lee, Alexander López-Parrado, Alexander Vera Tasamá y Duvier de Jesús Bohórquez Palacio</i>	
A Curriculum Design Approach by the Means of a General Morphological Analysis.....	59
<i>Aproximación al diseño curricular mediante análisis morfológico general</i> <i>Alexander Vera Tasama, Jorge Iván Marín Hurtado, Jaiber Evelio Cardona Aristizábal</i> <i>y Francisco Javier Ibarquien Ocampo</i>	
Evaluation of Clusters based on Systems on a Chip for High-Performance Computing: A Review .....	75
<i>Evaluación de clústeres basados en sistemas en un chip para computación de alto desempeño: una revisión</i> <i>Melissa Johanna Aldana, Jaime Alberto Buitrago y Julián Esteban Gutiérrez</i>	
Monitoreo de indicadores de valor a través de minería de datos, gestión de procesos de negocio y mejoramiento continuo con gestión del riesgo .....	93
<i>Monitoring Value Indicators by Applying Data Mining, Business Process Management, and Continuous Improvement with Risk Management</i> <i>Dario Enrique Soto Durán, Juan Camilo Giraldo Mejía, Fabio Alberto Vargas Agudelo, Jovani Jiménez Builes y Antonio Jesús Valderrama Jaramillo</i>	
Proceso y evolución de los métodos formales en la ingeniería de requisitos .....	119
<i>Process and evolution of formal methods in requirements engineering</i> <i>Jorge Hernán Suaza Jiménez y Gloria Amparo Lora Patiño</i>	
Definiendo un modelo de proceso de <i>software</i> para la práctica del modding .....	137
<i>Defining a Software Process Model for Modding Practice</i> <i>Andrés Felipe Ceballos, Wilson Libardo Pantoja Yépez y Julio Ariel Hurtado</i>	
Construcción de un juego serio como apoyo al aprendizaje de la física cinemática .....	159
<i>Construction of a Serious Game as Support for the Learning of Kinematics</i> <i>Carlos Fernando Aguilar Castrillón, Andrés Felipe Ojeda Rivera, Carol Julieth Aguilar Paz, María Isabel Vidal Caicedo, Marta Cecilia Camacho Ojeda y Gabriel Elías Chanchí Golondrino</i>	

Influencia de la posición de impresión y la densidad de relleno en las propiedades mecánicas de probetas fabricadas en ABS..... 179  
*Influence of the Printing position and the Infill Density on the Mechanical Properties of 3D Printed ABS Specimens*  
*Manuel José Carvajal Loaiza, Pablo González Díaz, Carlos Alberto Mejía Blandón,*  
*Liliana Marcejal Bustamante Góez y Junes Abdul Villarraga Ossa*

Desarrollo de una aplicación para la predicción de ingredientes y recetas de cocina por medio de TensorFlow y máquinas de soporte vectorial ..... 195  
*Development of an Application for the Prediction of Kitchen Ingredients and Recipes through TensorFlow and Support-Vector Machines*  
*Yeny Muñoz-Castaño, Luis Castillo-Ossa, Omar Castrillón-Gomez, Felipe Buitrago-Carmona*  
*y Santiago Loaiza Giraldo*

Current-Voltage Characteristics of a Self-Assembled DNA ..... 217  
*Curva característica de voltaje y corriente de un ADN autoensamblado*  
*Carlos José Páez-González, Jorge H Quintero-Orozco y Andrés Camilo García-Castro*

Estado de corrosión en tuberías y accesorios metálicos expuestos al aire, basado en conjuntos difusos ..... 227  
*Corrosion State in Metal Pipes and Fittings Exposed to Air Based on Fuzzy Sets*  
*Fernando Gutiérrez Fique, Milton Mena Serna y César Augusto García-Ubaque*

Web Service Adaptation for the Customization of Early Alerts in Agriculture..... 239  
*Adaptación de servicios web para la personalización de alertas tempranas en agricultura*  
*Oscar Ricardo Valencia Aguilar, Emmanuel Lasso y Juan Carlos Corrales*

Prototype and Method for Crops Analysis in the Visible and Infrared Spectrum from a Multirrotor Air Vehicle ..... 259  
*Prototipo y método para análisis de cultivos en el espectro visible e infrarrojo desde un vehículo aéreo multirrotor*  
*Julián Andrés Bolaños, Liseth Viviana Campo y Juan Carlos Corrales*

Instrucciones a autores ..... 283

<i>Revista Ingenierías Universidad de Medellín</i>	19(37)	Julio-Diciembre	2020	ISSN (en línea): 2248-4094 ISSN (impreso) 1692-3324
--	--------	-----------------	------	--

## Editorial

Respetado lector, en esta edición de *Revista Ingenierías Universidad de Medellín* estamos entregando 14 artículos de los cuales 6 están en inglés, esto se alinea con el propósito que nos hemos trazado hacia la internacionalización para llevar nuestro contenido hacia el público no hispanoparlante. Esto también va en conjunción con la meta que tenemos de continuamente mejorar y elevar la calidad de nuestros contenidos, y de aumentar la visibilidad en el público extranjero. Es por esta razón que se hemos tomado la decisión que, desde julio del año en curso, todos los artículos que se reciben y sean aprobados para publicación circularan en inglés. Esto representa una serie de retos para todos los actores que hacemos parte de *Revista Ingenierías Universidad de Medellín*: autores, evaluadores, comité editorial, colaboradores Universidad de Medellín; que nos exige romper la endogamia país para poder incluir dentro de nuestro contenido, publicaciones que incluyan una mayor cantidad resultados de trabajos e investigaciones internacionales y/o de trabajo conjunto de alta calidad entre instituciones e investigadores que nos permita también incrementar el impacto y la visibilidad de nuestra revista y los contenidos que entregamos a nuestro público, con el claro propósito de pertenecer y figurar en bases de datos internacionales de alto impacto.

Agradecemos a quienes de alguna forma u otra nos acompañan y hacen posible que continuemos creciendo y proyectándonos. Los invitamos a que asuman este reto con nosotros y que podamos entregar contenidos cada vez más relevantes para el público internacional.

## Editorial (English)

Dear reader, this issue of *Revista Ingenierías Universidad de Medellín* delivers 14 articles, 6 of which are in English, in accordance with the internationalisation goal necessary for delivering our content to the non-Spanish speaking public of our journal. This is why we have decided that since July of this ongoing year, every article received and approved for publication will circulate in English. This represents a series of challenges for all of the people that are part of this journal: authors, evaluators, the editorial committee, the University of Medellín collaborators; which demands from us breaking the country's endogamy and including into our published contents more international works and researches, which requires a high-quality joint work among institutions and researchers that allows us to increase the impact and visibility of our journal and the contents we deliver to our public with the clear purpose of belonging and being visible in high impact international data.

We thank those who in one way or another accompany us and make possible our continuous growth and projection. We invite you to assume this challenge with us in order to deliver ever more relevant contents for the international public

## Editorial (Português)

Prezado(a) leitor(a):

Nesta edição da *Revista Ingenierías Universidad de Medellín*, estamos publicando 14 artigos, dos quais seis estão em inglês, o que alinha a revista com o objetivo de caminhar à sua internacionalização para levar nosso conteúdo ao público não hispanofalante. Isso também vai em consonância com a meta que temos de melhorar e elevar continuamente a qualidade do nosso conteúdo e de aumentar a visibilidade no âmbito internacional. Por isso, tomamos a decisão de que, desde julho deste ano, todos os artigos recebidos e aprovados para a publicação circularão em inglês. Isso representa desafios para todos os atores que fazem parte da *Revista Ingenierías Universidad de Medellín*: autores, pareceristas, comissão editorial e colaboradores da Universidad de Medellín, porque exige romper a endogamia existente no país para poder incluir publicações com mais quantidade de resultados de trabalhos e pesquisas internacionais ou de trabalho conjunto de alta qualidade entre instituições e pesquisadores, o que nos permitirá também aumentar o impacto e a visibilidade da nossa revista e os conteúdos oferecidos ao público, com o claro objetivo de pertencer e figurar em bases de dados internacionais de alto impacto.

Agradecemos a todos que nos acompanham e tornam possível o crescimento e a projeção da revista. Convidamos a todos que assumam esse desafio conosco e que possamos oferecer um conteúdo cada vez mais relevante para o público internacional.





# Estimación de la captura de carbono en un relicto de bosque urbano mediante técnicas de teledetección \*

Claudia Marcela Cardona Lindo\*\*

Julián Garzón Barrero\*\*\*

Gonzalo Jiménez Cleves\*\*\*\*

Recibido: 23/11/2018 • Aceptado: 28/06/2019

<https://doi.org/10.22395/rium.v19n37a1>

## Resumen

El objetivo de este trabajo es calcular la capacidad de captura de CO<sub>2</sub> del relicto boscoso de la Universidad del Quindío que recibe el nombre de “Jardín Botánico Cedro Rosado” a través de técnicas que integran mediciones *in situ* y teledetección. En esta primera fase se obtendrán las imágenes multiespectrales, índices de vegetación diferencial normalizado (NDVI), mejorado (EVI) y ajustado al suelo, (SAVI), y la clasificación basada en objetos. En una segunda fase se medirán variables arbóreas y se estimarán las variables biofísicas índice de área foliar (LAI) y fracción de radiación activa fotosintéticamente absorbida (Fapar) con el instrumento Tracing Radiation and Architecture of Canopies (Trac) para correlacionarlas con los índices de vegetación. Esto definirá las constantes del modelo exponencial de regresión a través del establecimiento de la ecuación alométrica local, para interpolar la biomasa contenida en la zona de estudio y relacionarla con su capacidad de almacenamiento de CO<sub>2</sub>.

**Palabras clave:** captura de carbono; teledetección; ecuaciones alométricas; índices de vegetación; variables biofísicas; clasificación de imagen basada en objetos.

---

\* Artículo derivado de la investigación “Estimación del carbono capturado por un bosque urbano mediante herramientas de teledetección” (en curso).

\*\* Magíster en Ingeniería de Recursos Hídricos y Medio Ambiente (Universidad del Quindío), especialista en gestión ambiental, y química. Coordinadora del Sistema de Gestión Ambiental de la Universidad del Quindío, e investigadora del Grupo Geoides G62. Correo electrónico: cmcardona@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0003-1370-0869>.

\*\*\* Estudiante de Doctorado en Ingeniería Geomática, Universidad Politécnica de Madrid. Magíster en sistemas de información geográfica. Especialista en geomática. Profesor del Programa de Ingeniería Topográfica y Geomática e investigador del Grupo Geoides G62, Universidad del Quindío. Correo electrónico: juliangarzonb@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-4871-3726>.

\*\*\*\* Magíster en ingeniería de sistemas. Profesor del Programa de Ingeniería Topográfica y Geomática y líder del grupo de investigación Geoides G62, Universidad del Quindío. Correo electrónico: gjcleves@uniquindio.edu.co. Orcid: <http://orcid.org/0000-0002-0769-9729>

## Estimation of Carbon Capture in an Urban Forest Relict through Teledetection Techniques

### **Abstract**

The objective of this study is to calculate the capacity of CO<sub>2</sub> capture from the forest relict of the University of Quindío “Jardín Botánico Cedro Rosado” through the use of techniques that integrate in situ measurements with remote sensing. In the first phase, multispectral images, Normalized Differential Vegetation Index (NDVI), Improved Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), and object-based classification will be obtained. In the second phase, tree variables will be measured, and Leaf Area Index (LAI) and the Fraction of Absorbed Photosynthetically Active Radiation (Fapar) biophysical variables will be estimated with the Tracing Radiation and Architecture of Canopies (TRAC) optical instrument, in order to correlate them with the vegetation indexes. This will define the constants of the exponential regression model defining the local allometric equation, which will interpolate the biomass in the entire image.

**Keywords:** Carbon caption; remote sensing; allometric equation; vegetation indices; biophysical variables; OBIA.

## Estimativa da captura de carbono num relicto de floresta urbana mediante técnicas de sensoriamento remoto

### **Resumo**

O objetivo deste trabalho é calcular a capacidade de captura de CO<sub>2</sub> do relicto arborizado da Universidad del Quindío que recebe o nome de “Jardim Botânico Cedro Rosado” por meio de técnicas que integram medidas *in situ* e detecção remota. Nessa primeira fase, serão obtidas as imagens multiespectrais, índices de vegetação diferencial normalizada (NDVI), melhorado (EVI) e ajustado ao solo (SAVI), e a classificação baseada em objetos. Numa segunda fase, serão medidas variáveis arbóreas e serão estimadas as variáveis biofísicas índice de área foliar (LAI) e a fração de radiação fotossinteticamente ativa absorvida (Fapar) com o instrumento Tracing Radiation and Architecture of Canopies (Trac) para correlacioná-las com os índices de vegetação. Isso definirá as constantes do modelo exponencial de regressão por meio do estabelecimento da equação alométrica local, para interpolar a biomassa contida na zona de estudo e relacionar com sua capacidade de armazenamento de CO<sub>2</sub>.

**Palavras-chave:** captura de carbono; sensoriamento remoto; equações alométricas; índices de vegetação; variáveis biofísicas; classificação de imagem baseada em objetos.

## INTRODUCCIÓN

El calentamiento global es el aumento de la temperatura media de la superficie terrestre como resultado del aumento de concentraciones de gases de efecto invernadero (GEI). Es bien conocido que la liberación de CO<sub>2</sub> a través de diversas fuentes es uno de los causales de este calentamiento. A través del proceso de fotosíntesis los bosques capturan CO<sub>2</sub>, lo fijan en sus estructuras vivas, lo acumulan en su biomasa y lo transfieren al mantillo en descomposición, y así se constituyen en reservas de carbono [1]. Esta importancia mundial de los ecosistemas forestales insiste en la necesidad de determinar con precisión la cantidad de carbono almacenado en los bosques.

Existen diferentes métodos convencionales para cuantificar la captura de CO<sub>2</sub>, muchos de los cuales son complejos, costosos y de cobertura limitada [2]. A este respecto, la posibilidad de estimar la biomasa y carbono a través de técnicas de teledetección representa una alternativa importante para sortear estas limitaciones, y ha sido ampliamente reportada en estudios a diversas escalas y ambientes especiales [3-5]. Para estimaciones de almacenamiento de CO<sub>2</sub>, IPCC [6] sugiere aplicar una fracción del 50 % de la biomasa aérea.

La estimación de la biomasa es una tarea difícil, especialmente en zonas con estructuras de condiciones ambientales complejas [7]. Los modelos alométricos han sido desarrollados para aplicaciones en bosques naturales y plantaciones forestales, pero su uso reviste limitaciones debido a condiciones naturales locales —tales como forma del árbol y arquitectura del dosel—, lo que sugiere desarrollar modelos locales de estimación [8-9]. A este respecto, los métodos de estimación de biomasa y carbono a partir de sensores remotos necesitan la integración de información de campo, basada en un muestreo espacial suficiente, y el uso de las ecuaciones alométricas apropiadas para mejorar su estima y asegurar su robustez [10-11]

Los vehículos aéreos no tripulados (*unmanned aerial vehicles*, UAV) equipados con sensores espectrales ligeros facilitan el análisis de vegetación con procedimientos no destructivos [12]. En los últimos años, los sistemas UAV han llamado la atención de las geociencias debido a la posibilidad de capturar datos de forma rentable a altas resoluciones espaciales y temporales [13]. En los trabajos con UAV donde se busca discriminar entre especies individuales de vegetación, Torres-Sánchez *et al.* [14] recomiendan resoluciones espaciales inferiores a cuatro centímetros que se logran con alturas de vuelo inferiores a 100 metros. A fin de corregir el posicionamiento geográfico de los elementos capturados en la escena, es necesario establecer puntos de control terrestre (GCP) distribuidos de manera uniforme en la zona de trabajo [15-16].

La atmósfera es la fuente primaria de ruido para la medición precisa de la reflectancia superficial con sensores remotos ópticos [17-18]. Los efectos de la atmósfera sobre los índices de vegetación son relevantes: Agapiou *et al.* [19] encontraron diferencias significativas de hasta 15 % entre hacer o no correcciones atmosféricas en el cálculo del NDVI. Homem *et al.* [20], a su turno, evaluaron los efectos atmosféricos en el NDVI y el SAVI utilizando el algoritmo de corrección 6S: encontraron que para la obtención de valores ajustados de estos índices es necesario corregir atmosféricamente las imágenes; así mismo, concluyen que estos efectos tienen alta incidencia cuando se busca relacionar los índices de vegetación con variables biofísicas como LAI y biomasa, por lo cual recomiendan efectuar mediciones de campo para ajustar los modelos.

El dosel forestal es la frontera entre la atmósfera y la superficie terrestre; la comprensión de su estructura es importante para modelar la captura de gases atmosféricos como el CO<sub>2</sub> [21]. El LAI es un atributo estructural de los ecosistemas forestales a través del cual se han realizado investigaciones en estimativos de producción forestal y capacidad de captura de carbono [22], su correcta estimación es esencial para impactos de ecosistemas sobre el clima, el ciclo biogeoquímico y el balance energético [23]. Para medir de manera precisa y eficiente el LAI y Fapar se utiliza el instrumento óptico que recibe el nombre de Tracing Radiation and Architecture of Canopies (Trac): a través de la interceptación de la luz que pasa por el dosel, logra definir su estructura [24].

Con el trabajo presentado aquí se pretende calcular la capacidad de captura de CO<sub>2</sub> del relicto boscoso “Jardín Botánico Cedro Rosado” de la Universidad del Quindío mediante la integración de técnicas de teledetección con datos de campo. Para lograr este objetivo, en la fase inicial del trabajo se parte de la compactación del ortomosaico de la zona de estudio con los valores de reflectancia necesarios para calcular los índices de vegetación NDVI, EVI y SAVI. Será necesario discriminar las diversas especies vegetales a través de la clasificación de imagen basada en objetos (OBIA), lo que llevará a definir las áreas de mediciones *in situ* de variables arbóreas. En la segunda fase, para el cálculo de las variables biofísicas LAI y Fapar, se utilizará el instrumento óptico Trac, el cual registra la transmisión de luz directa del dosel a través de los transectos diseñados. Finalmente, para encontrar las constantes del modelo exponencial de regresión propias para calcular biomasa, se medirá el grado de correlación de las variables obtenidas con los datos de campo, de modo que se pueda hacer la interpolación de la variable dependiente para la totalidad de la zona de estudio.

## 1. MATERIALES Y MÉTODOS

### 1.1 Caracterización biológica de la zona de estudio

El Jardín Botánico Cedro Rosado es un bosque secundario ubicado en el campus principal de la Universidad del Quindío (Municipio de Armenia, Depto. del Quindío, Colombia), con posición 4°33'N , 75°39'W; comprende, principalmente, vegetación natural intervenida con manchas de bosque secundario y guaduales. De acuerdo con la clasificación ecológica de zonas de vida expuestas por Holdridge [25], el espacio que nos ocupa corresponde a bosque muy húmedo premontano (BMH - PM), cuya superficie aproximada es de 13 Ha y se encuentra a una altitud que oscila entre 1490 y 1530 msnm. Su temperatura promedio asciende a 19,5 °C; su precipitación, a 2436 mm/año; y su humedad relativa varía entre 65-75 %. Posee dos microcuencas que, en conjunto, forman la Quebrada del Cedro Rosado [26]. La diversidad biológica encontrada incluye 64 especies de aves; 53 de hongos macroscópicos; 3 de anfibios; 2 de reptiles; 7 de mamíferos; y 247 de plantas (230 nativas y 17 cultivadas)

### 1.2 Diseño del plan de vuelo

Esta labor definió la ruta a seguir por la aeronave desde su despegue hasta su aterrizaje. La misión se planificó con el software Map Pilot® instalado en la estación de cómputo terrestre. La línea de vuelo, a su turno, se diseñó sobre las imágenes del área de trabajo importadas desde la aplicación Google Earth®, y para planificarla fueron necesarios tres parámetros: zona de estudio, especificaciones de la cámara y tareas propias del UAV. La información del área de vuelo incluye el ancho y la longitud de la zona de estudio, el ángulo de dirección del lado principal y el traslape deseado en las imágenes, capturadas con 60 % de recubrimiento longitudinal y 30 % transversal. Se empleó la cámara Survey 2 fue empleada para registrar longitudes de onda del infrarrojo cercano en la vegetación, con una resolución espectral de 760 a 840 nm, y una resolución radiométrica de 24 bits; y con la cámara Phantom 3 Advanced se registraron los valores de RGB con una resolución radiométrica de 8 - 10 bits [27]. Las tareas del UAV son las acciones que el sensor tiene que realizar una vez llega a cada punto de captura de la imagen; aquí se incluyen el número de fotos y el tiempo de permanencia en cada punto. Ingresadas estas variables y la altura del vuelo, el aplicativo estableció de forma automática la ruta autónoma de vuelo y estimó su duración de acuerdo con el número total de imágenes planificadas.

### 1.3 Vuelo UAV y adquisición de imágenes

Una vez diseñada la misión de vuelo, se cargó la ruta en la plataforma Map Pilot. Luego, se realizó el montaje de las cámaras al vehículo que registraron información

de manera simultánea: una registraría radiación visible en el sensor RGB; mientras que la otra poseía un filtro que reflejaba la luz del infrarrojo cercano en el canal azul del sensor RGB. Mediante el transmisor de radiocontrol se ejecutó el lanzamiento de la aeronave activando la ruta de vuelo trazada para que el vehículo llegara hasta el punto de inicio de la primera fotografía y se desplazara por las líneas de vuelo hasta cubrir por completo la zona estudiada.

#### 1.4 Procesamiento de restitución para compactación del ortomosaico

Para iniciar el procedimiento de restitución por coincidencia espacial entre elementos de cada imagen se utilizó el *software* especializado Agisoft PhotoScan Pro®. A través de esa herramienta se alienaron imágenes para identificar puntos coincidentes entre ellas, de tal suerte que se compuso una nube de puntos para definir el modelo; y luego se definieron los puntos de control terrestre pertenecientes a la red topográfica de alta precisión que se encuentra en el campus de la universidad, a fin de georreferenciar el ortomosaico.

#### 1.5 Preprocesamiento digital de imágenes

El modelo de procesamiento digital de la imagen exige dos tareas fundamentales para obtener reflectancias corregidas, esto es, las calibraciones radiométrica y atmosférica. Los procedimientos de calibración radiométrica aplicados a la imagen tienen en cuenta los errores que afectan el valor del brillo de cada pixel de la escena debido a dos fenómenos fundamentales: el error del detector del sistema sensor y el de atenuación ambiental [28]. El sistema MAPIR contiene un aditamento especial con tres objetivos, de los cuales se conoce su curva de reflectancia. Antes del vuelo se toman fotografías con cada cámara al conjunto de blancos que se usarán para calibrar los valores de los pixeles en el *software* de posprocesamiento; ello permite mejorar el nivel de normalización de las mediciones y capturar imágenes en días de poca claridad [29]. La calibración aplica los modelos señalados en las ecuaciones (1) y (2).

$$L_{\lambda} = Gain \cdot ND + Bias \quad (1)$$

Donde:

- $L_{\lambda}$ : radiancia espectral en la apertura del sensor.
- ND: valor crudo del pixel.
- Gain y Bias son las constantes de calibración de la cámara.

$$R_T(\theta_z) = \frac{DN_T(t)}{DN_R(t)} R_{BaSO_4} \left( \frac{0^\circ}{\theta_z} \right) \quad (2)$$

Donde:

- $R_T$ : factor de reflectancia de un objetivo desconocido.
- $\theta_z$ : ángulo cenital solar en cualquier tiempo  $t$  medido con el radiómetro.
- $DN_T(t)$ : matriz digital producida por el UAV cuando el instrumento visa el objetivo.
- $DN_R(t)$ : matriz digital del panel de referencia blanco cuando la cámara está visando al panel de referencia entre los tiempos  $t_0$  y  $t_1$ .
- $R_{BaSO_4} [0^\circ/\theta_z]$ : factor de reflectancia del panel  $BaSO_4$  con ángulo  $0^\circ$  y ángulo cenital solar de  $\theta_z$ .

La interacción de la radiación electromagnética con las partículas de agua y polvo que se pueden encontrar entre el sensor de la cámara y el dosel causan perturbaciones en las señales captadas, debido a que actúan como medio absorbente y dispersante de la señal. La corrección atmosférica, que es la eliminación de estas alteraciones, constituye uno de los factores de más influencia al estimar con precisión los índices de vegetación.

Luego de realizar las calibraciones radiométrica y atmosférica se obtienen los valores de reflectancias necesarios para calcular los índices de vegetación NDVI, EVI y SAVI, señalados en las ecuaciones (3), (4) y (5).

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (3)$$

Donde:

$\rho_{NIR}$ : radiancia en unidades de reflectancia en la banda del infrarrojo cercano.

$\rho_{red}$ : radiancia en unidades de reflectancia en la banda del rojo.

$$EVI = G \cdot \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + C_1 \rho_{red} - C_2 \rho_{blue} + L} \quad (4)$$

Donde:

- $\rho$ : reflectancia atmosféricamente corregida en el NIR, regiones espectrales rojo y azul.  $C_1=6,0$  y  $C_2=7,5$ : coeficientes de resistencia atmosférica.

- $L=1,0$ : corrección al efecto del fondo del follaje.
- $G=2,5$ : factor de ganancia.

$$SAVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red} + L} \cdot (1 + L) \quad (5)$$

Donde:

- $\rho_{NIR}$ : reflectancia atmosféricamente corregida en la banda del infrarrojo cercano.
- $\rho_{red}$ : reflectancia atmosféricamente corregida en la banda del rojo.
- $L=0,5$ : constante de compensación promedio.

El cálculo de estos índices de vegetación mediante sensores instalados en UAV se ha convertido en una alternativa *costo-efectiva* más flexible y precisa que el procesamiento de imágenes satelitales [12] [30], además de facilitar la realización de un monitoreo periódico del área de estudio [31].

## 1.6 Clasificación de imagen basada en objetos (*object-based image analysis*, OBIA)

A través de este modelo se pretende superar los problemas de clasificación basada en píxeles mediante la definición de segmentos, permitiendo que la variabilidad de la reflectancia espectral se use como atributo para discriminar las características de la segmentación. Se realiza esta clasificación OBIA a fin de obtener el reconocimiento de patrones que permitan discriminar las distintas especies vegetales en la zona de estudio.

OBIA se ha usado durante la última década para definir patrones específicos en una imagen. Esta tecnología se orienta hacia la identificación de objetos a través de la información espectral, las formas geométricas, el análisis de vecindad, la textura y el tono de la imagen; y supera la clasificación basada en píxeles debido a que esta última no tiene en cuenta la forma, textura y estructura, clasificadores de baja eficiencia en el reconocimiento de patrones [32].

## 1.7 Mediciones *in situ* de variables arbóreas

Una vez hecha la clasificación, se harán mediciones *in situ* de especies para cuantificar los almacenamientos de carbono en las parcelas seleccionadas a través de la ecuación (6), mediante la medición de variables arbóreas como altura y diámetro a la altura del pecho (DAP). La densidad de la madera ( $\rho$ ) se tomará de la base de datos reportada por Zanne *et al.* [33], en la cual se encuentran las especies arbóreas identificadas en la zona de estudio, como se muestra en la tabla 1.



Tabla 1. Densidad de la madera de las especies arbóreas encontradas en el Sendero Cedro Rosado

Familia	Nombre científico	Densidad madera (g/c <sup>3</sup> ) masa secada en horno/ volumen fresco	Región
Fabaceae	<i>Bauhinia variegata</i>	0,606	Sureste de Asia (tropical)
Malvaceae	<i>Heliocarpus americanus</i>	0,200	América del Sur (tropical)
Urticaceae	<i>Cecropia peltata</i>	0,310	América del Sur (tropical)
Juglandaceae	<i>Juglans neotropica</i>	0,490	América del Sur (tropical)
Myrtaceae	<i>Syzygium jambos</i>	0,700	América del Sur (tropical)
Myrtaceae	<i>Psidium guajava</i>	0,629	América del Sur (tropical)
Sapindaceae	<i>Cupania Americana</i>	0,730	América del Sur (tropical)

Fuente: Adaptación de [33].

Para aplicar el modelo alométrico referido en la ecuación 6 se usarán las constantes B1 reportadas en el *Protocolo para la estimación nacional y subnacional de biomasa - Carbono en Colombia* [34]: allí se establece que para la zona de vida “bosque muy húmedo premontano” (BMH-PM) se puede utilizar la ecuación (6) asociada a ella: B1 = 0,932 y a = -2,289.

$$\ln BA = \alpha + B_1 \cdot \ln (D^2 \cdot H \cdot \rho) \quad (6)$$

Donde:

- BA: biomasa aérea en kg.
- D: diámetro a la altura del pecho medido a 1,30 m de altura sobre el suelo (cm).
- H: altura total del árbol (cm).
- $\rho$ : densidad de la madera (g/cm<sup>3</sup>).

Para las especies como el Otobo (*Otobo lehmannii*) y Guama (*Inga densiflora*), de las cuales no se encuentra información de la densidad de la madera en las bases de datos, se usará el promedio del nivel taxonómico superior (familia) [34].

## 1.8 Obtención de variables biofísicas

Una vez diseñados los transectos, se realizará el recorrido con el instrumento óptico Trac, que registrará la transmisión de luz directa a una frecuencia alta (32 Hz) a través del dosel como insumo fundamental para calcular la fracción de hueco (*gap fraction* P<sub>0</sub>(Θ)), lo que definirá las variables biofísicas LAI [35] y Fapar [36] a través de las ecuaciones (7), (8) y (9).

$$LAI = \int_0^H l(h) \cdot dh \quad (7)$$

Donde:

- $l(h)$ : densidad de área a la altura del dosel.

$$fAPAR_{Day}^{Black\ Sky} = \frac{\int_{sunset}^{sunrise} \cos(\theta) \cdot (1 - P_o(\theta)) \cdot d\theta}{\int_{sunset}^{sunrise} \cos(\theta) \cdot d\theta} \quad (8)$$

$$fAPAR^{White\ Sky} = \frac{1}{\pi} \cdot \int_0^{\frac{\pi}{2}} \int_0^{\frac{2\pi}{2}} (1 - P_o(\theta)) \cdot \cos\theta \cdot \sin\theta d\theta \cdot d\varphi = 2 \int_0^{\frac{\pi}{2}} (1 - P_o(\theta)) \cdot \cos\theta \cdot \sin\theta d\theta \quad (9)$$

Donde:

- $P_o(\theta)$ : fracción de hueco (*gap fraction*).

Existen diversos instrumentos que permiten capturar estas variables en campo, entre los cuales se encuentran el LAI-2200C Plant Canopy Analyzer, el cleptómetro, Trac y cámaras equipadas con lentes de tipo “ojo de pescado” (*fisheye*) que capturan fotografías hemisféricas. Tanto estas últimas como el LAI-2200C dependen de un cielo uniforme y cubierto ya que, cuando las hojas se encuentran muy iluminadas por el Sol, se producen efectos difíciles de medir. El cleptómetro trabaja bajo condiciones de cielo despejado y utiliza sensores para estimar la luz promedio bajo el dosel [37]; se ha encontrado que estos instrumentos ópticos tienden a subestimar el LAI de bosques donde la distribución de los elementos del follaje no es aleatoria.

Para minimizar la incidencia de estos efectos en la medición de las variables de campo, en la segunda fase de este trabajo se utilizará el mencionado Trac, instrumento óptico desarrollado por [24] y empleado para medir el LAI y la Fapar absorbida por la copa de los árboles. Esta tecnología utiliza la información de la “fracción de hueco” (*gap size*) del dosel para proporcionar un índice de aglomeración del follaje que cuantifica el efecto de distribución espacial no aleatoria del follaje. Tales elementos conducen a la definición de la arquitectura del dosel a través del modelo de Poisson [38].

## 1.9 Modelación estadística

Con esta se pretende medir el grado de relación entre las variables biofísicas relacionadas, los índices de vegetación y los datos híbridos de campo. Así mismo, se busca detectar el comportamiento de conglomerados de las variables a través del índice de Moran —ecuación (10)— estadístico, que mide la autocorrelación espacial a través

de la identificación de clases de distinta homogeneidad de la muestra; su objetivo es evaluar si el patrón se encuentra agrupado, disperso, o aleatorio.

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \cdot (Z(s_i) - \bar{Z}) \cdot (Z(s_j) - \bar{Z})}{\sum_{i=1}^n (Z(s_i) - \bar{Z})^2} \tag{10}$$

Donde:

- $Z(s_1), \dots, (Z(s_n))$ : variables medidas en las  $n$  áreas.

Al medir el grado de relación entre las variables se tendrá la representación gráfica de la correlación a través de diagramas de dispersión, así como la relación que presenta cada una de las variables calculando su correlación cruzada, la correlación de Pearson y la bondad del ajuste a partir del coeficiente de determinación  $R^2$ . Esto definirá los parámetros necesarios para hallar las constantes del modelo exponencial de regresión, de modo que se pueda hacer la interpolación de la variable dependiente (biomasa) para la totalidad de la imagen.

En este trabajo se aplicó la primera parte de la metodología: esto es, captura y procesamiento de las imágenes, cálculo de los índices de vegetación y OBIA, como se muestra en la figura 1.

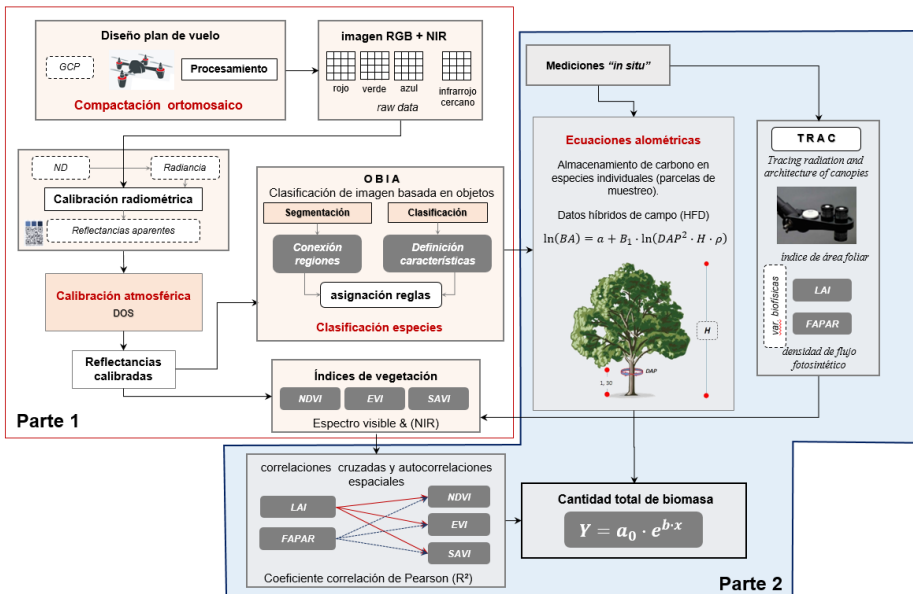


Figura 1. Metodología aplicada

Fuente: elaboración propia.

## 2. RESULTADOS

### 2.1 Vuelo UAV y adquisición de imágenes

Se realizaron tres vuelos para garantizar cobertura y autonomía de vuelo, a partir de los cuales se capturaron 326 fotografías a una altura de 90 m, con traslapes mínimos longitudinal de 80 % y transversal de 60 %. La duración de los vuelos osciló entre 8 y 14 minutos, y se cubrieron cerca de 38 Ha en total (tabla 2 y figura 3).

Tabla 2. Parámetros de realización de itinerarios de vuelo.

Parámetro	Itinerario 1	Itinerario 2	Itinerario 3
Área de cobertura	14.47 hectáreas	15.46 hectáreas	8.2 hectáreas
Altura de vuelo	90 m	90 m	90 m
Distancia recorrida	5,09 km	4,96 km	3,14 km
Velocidad	7 m/s	7 m/s	7 m/s
Cantidad de baterías	1	1	1
Duración del recorrido	13 m, 30s	13 m, 13 s	8 m, 56 s
Numero de imágenes	121	130	75
Espacio necesario en memoria SD	0,60 GB	0,65 GB	0,37 GB
Resolución	3,9 cm/ pixel	3,9 cm/ pixel	3,9 cm/ pixel

Fuente: elaboración propia.

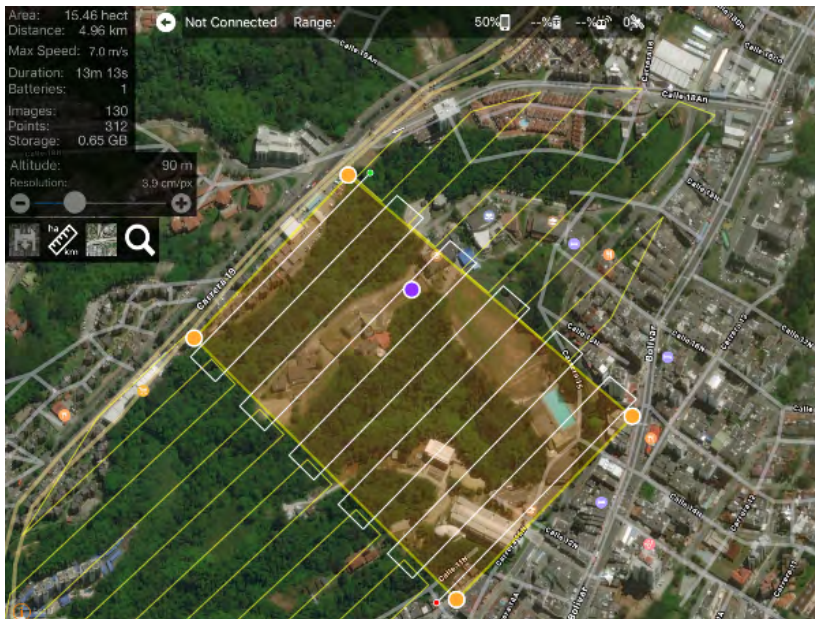


Figura 2. Itinerario de vuelo 2, Interfaz de parámetros Map Pilot for DJI®

Fuente: elaboración propia.

Antes de efectuarse los vuelos se realizaron tomas con la cámara NIR al *target* de calibración suministrado por el fabricante. Estas proporcionan parámetros de reflectancia conocidos y estandarizados para los colores y materiales en los que se encuentra fabricado dicho *target*, y son necesarios en el proceso de calibración de la ortofoto NIR.

## 2.2 Procesamiento de restitución para compactación del ortomosaico

Para la construcción de los ortomosaicos se empleó el *software* Agisoft Photoscan®, aplicación especializada que extrae información tridimensional de las fotografías aéreas, al tiempo que permite la elaboración de ortofotos georreferenciadas mediante la identificación y relacionamiento de características similares (homólogas) entre imágenes en áreas comunes o de traslape [39]. Para la orientación exterior del ortomosaico se utilizaron los puntos de control terrestre pertenecientes al campus de la Universidad del Quindío, de los cuales se estableció su posicionamiento con instrumental topográfico de alta precisión como marco de referencia para trabajos cartográficos y fotogramétricos. Este proceso se ilustra en la figura 3: el color indica el error en Z; el tamaño y la forma de la elipse representan el error en XY; y las posiciones estimadas de los puntos de control se indican con puntos negros.

Sumado a lo anterior, se realizó el modelo digital de elevaciones (MDE) presentado en la figura 4 con la información de la nube densa de puntos obtenida de las imágenes. Este tiene una resolución de 3,89 cm/pixel y una densidad de puntos de 662 puntos/m<sup>2</sup>. Con la nube de puntos y el MDE se generó el ortomosaico RGB con una resolución de 3 cm/pixel; y se exportó en formatos TIFF y JPG para su posterior procesamiento digital (figura 5).

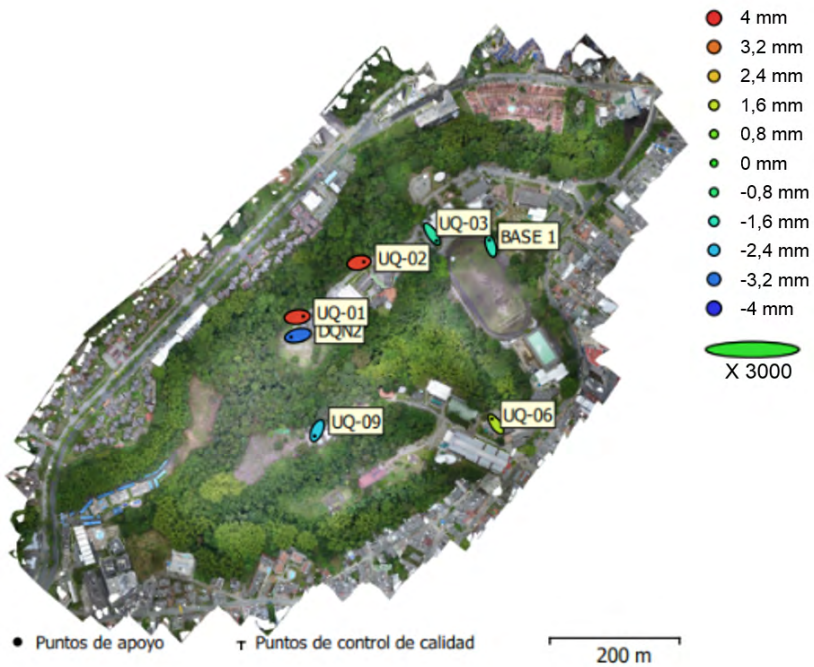


Figura 3. Posiciones de puntos de apoyo y estimaciones de errores, software Agisoft Photoscan®

Fuente: elaboración propia.

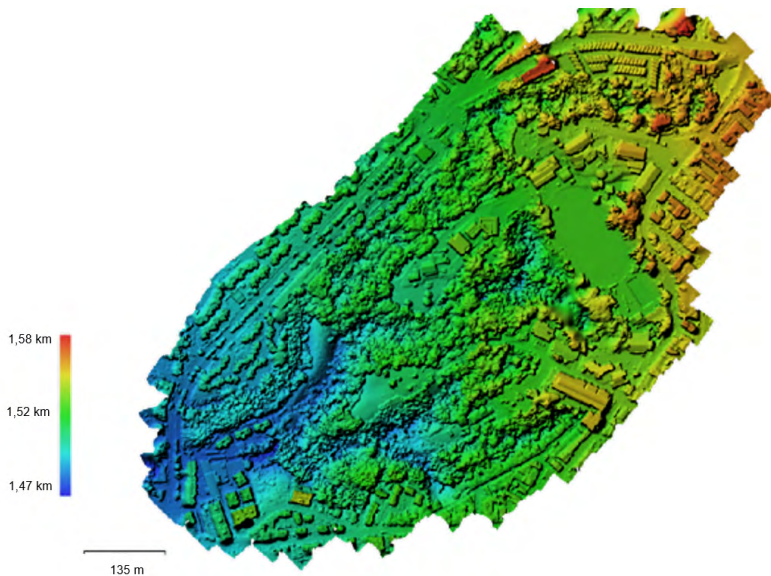


Figura 4. Modelo Digital de Elevaciones Zona de estudio, software Agisoft Photoscan®

Fuente: elaboración propia.





Figura 5. Ortofotografía RGB de la zona de estudio, *software* ArcGIS 10.4.1

Fuente: elaboración propia.

El ortomosaico NIR, representado en la figura 6, contiene valores de reflectancia en dos longitudes de onda —infrarrojo cercano a 850 nm y luz roja a 660 nm—, los cuales fueron calibrados con la curva de reflectancia de los objetivos de calibración propios del sistema Mapir.

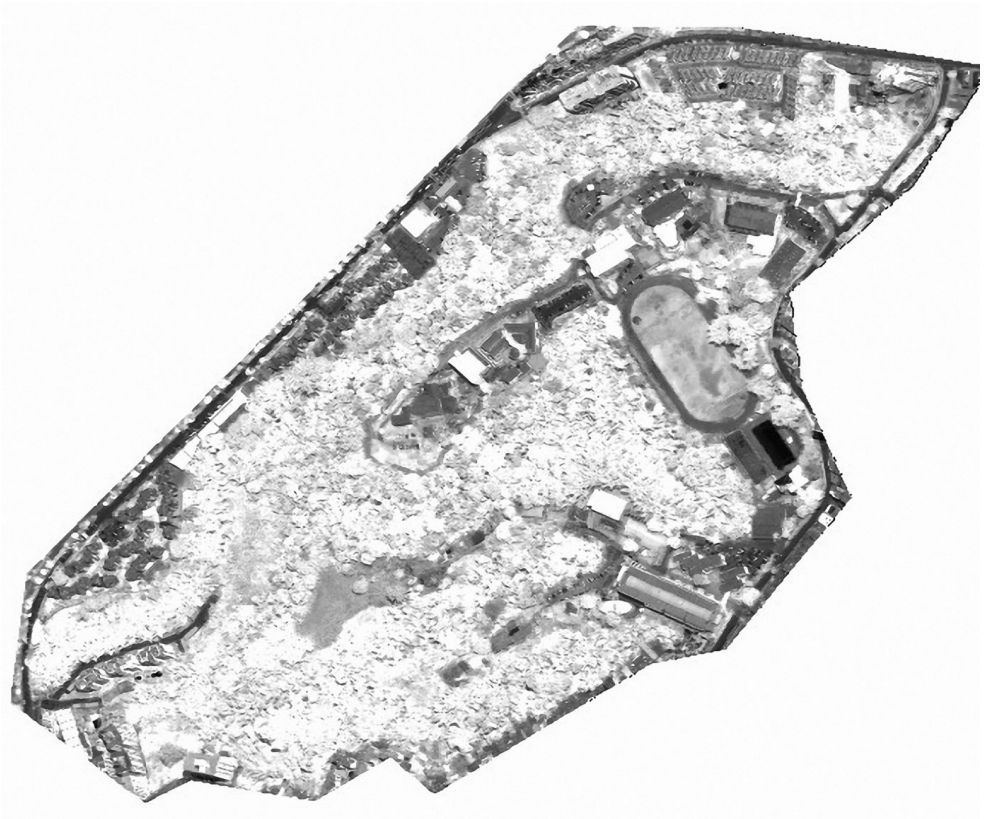


Figura 6. Ortofotografía NIR de la zona de estudio, *software* QGIS 2.18

Fuente: elaboración propia.

### 2.3 Estimación de los índices de vegetación

Una vez calibradas las reflectancias de las imágenes, se realizó el cálculo de los índices de vegetación NDVI, EVI y SAVI con el *software* QGIS 2.18. Los valores obtenidos en el NDVI se presentan en la figura 7. Se pudo observar que permite la diferenciación entre diferentes tipos de coberturas del suelo; no obstante, los valores obtenidos dentro del relicto boscoso son muy similares entre especies, lo que no permite realizar una diferenciación basada en estos índices.



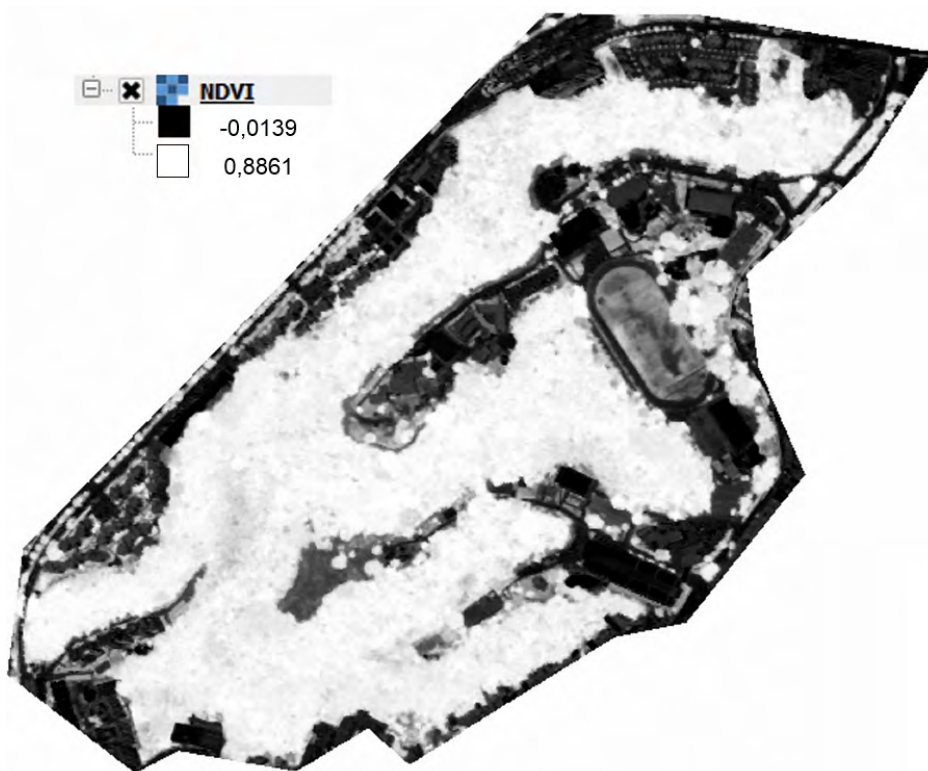


Figura 7. Estimación del NDVI de la zona de estudio, *software* QGIS 2.18

Fuente: elaboración propia.

#### 2.4 Clasificación de imagen basada en objetos (OBIA)

El primer paso del proceso de clasificación correspondió a la segmentación de la imagen. En el desarrollo de dicha labor se dio un peso superior a la reflectancia banda infrarroja, con lo cual se buscaba un mayor nivel de diferenciación debido a los atributos espectrales de la vegetación. Para llevar a cabo la clasificación, y con la colaboración del personal del Jardín Botánico, se llevó a cabo el entrenamiento que dio lugar a las siguientes clases: agua, arbustos, asbesto, asfalto, bosque mixto, caracolí, carbonero, casco de buey, caucho, cedro negro, cedro rosado, construcciones, guadual, gualanday, guayabo, mestizo, otobo, pasto, pavimento, suelo desnudo, vehículos y yarumo. Una vez seleccionadas las muestras de cada una de las clases —para lo cual se empleó el *software* eCognition® 9.0.1—, se realizó la OBIA de la imagen total, como se observa en la figura 8.

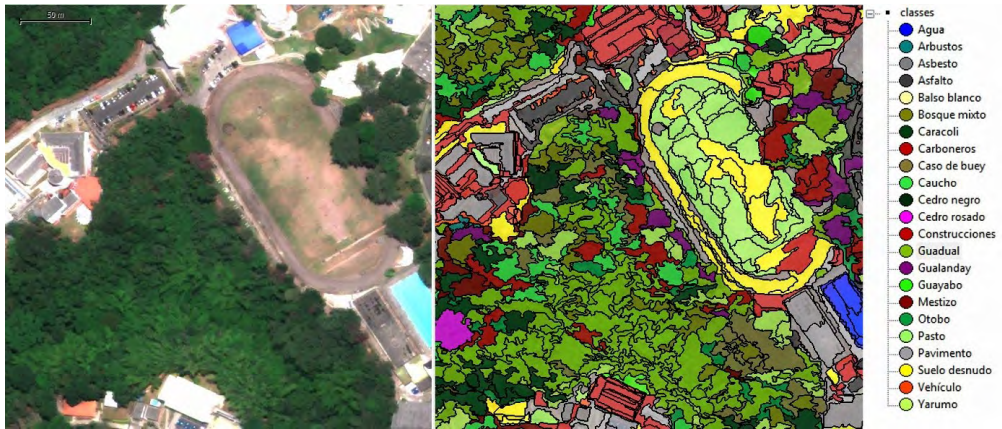


Figura 8. OBIA con 22 clases identificadas, *software eCognition® 9.0.1*

Fuente: elaboración propia.

### 3. DISCUSIÓN DE RESULTADOS

En esta fase del trabajo se obtuvieron dos imágenes, una en el espectro visible y otra en el infrarrojo cercano empelando un VANT, con las que se calcularon los índices de vegetación NDVI, EVI y SAVI. Estos no posibilitaron una disgregación entre tipos de vegetación, dada su similitud en coberturas boscosas; sin embargo, mediante la técnica OBIA, y a través del algoritmo multirresolución para definir la segmentación, se dio peso doble a los valores de reflectancia del infrarrojo cercano sobre las reflectancias de las otras bandas, con lo que se logró detectar diferentes especies. De acuerdo con los resultados obtenidos se concluye que, mediante el procesamiento digital de las imágenes obtenidas a través de sensores instalados en VANT, se pueden calcular diferentes índices de vegetación. Así mismo, la alta resolución obtenida (3 cm/pixel) hizo posible realizar una clasificación OBIA de relictos boscosos urbanos, con la que se consiguió una diferenciación ente tipos de vegetación y especies arbóreas; esto último no sería posible si se emplearan imágenes satelitales de resoluciones gruesas, dado que no bastan para discriminar entre especies.

### 4. CONCLUSIONES

Los resultados obtenidos constituyen un punto de partida importante para caracterizar relictos boscosos mediante el uso de sensores remotos, hecho que propone una metodología novedosa de estimación de variables forestales. El uso de técnicas de observación de la tierra en la estimación de biomasa arbórea disminuye costos en tiempo y dinero sin implicaciones destructivas del ecosistema; sin embargo, las mediciones de campo son necesarias para calibrar los modelos generados. Las variables derivadas NDVI, EVI,

SAVI obtuvieron valores con alto grado de similitud y no se logró discriminación de especies a través de estos índices. Como alternativa efectiva se utilizó la clasificación por objetos OBIA, dentro de la cual se asignó peso doble a la banda infrarroja y, en combinación con el algoritmo de segmentación multiresolución, se obtuvo la disgregación de especies o agrupaciones de especies con características similares. Los índices NDVI, EVI, y SAVI derivados a partir de la combinación de bandas multispectrales, capturadas con sistemas VANT, deberán ser objeto de análisis estadístico para hallar su correlación con las características estructurales del ecosistema.

## AGRADECIMIENTOS

Al topógrafo Jonnathan Castillo por su valiosa colaboración y continua asesoría. Al Centro de Investigaciones en Biodiversidad de la Universidad del Quindío (Cibuq) y la Oficina de Planeación y Desarrollo de la misma entidad, por la información suministrada para la caracterización de la zona de estudio

## REFERENCIAS

- [1] A. Vásquez y H. Arellano, “Estructura, Biomasa aérea y carbono almacenado en los bosques del Sur y Noroccidente de Córdoba”, en *Colombia Diversidad Biótica XII: La Región Caribe de Colombia*, Bogotá: Universidad Nacional de Colombia, 2012, pp. 923-962.
- [2] T. Y. Simegn, T. Soromessa y E. Bayable, “Forest Carbon Stocks in Lowland Area of Simien Mountains National Park: Implication for Climate Change Mitigation”, *Science, Technology and Arts Research Journal*, vol. 3, n.º 3, pp. 29-36, 2014. DOI: <http://dx.doi.org/10.4314/star.v3i3.5>
- [3] P. Vicharnakorn, R. P. Shrestha, M. Nagai, A. P. Salam y S. Kiratiprayoon, “Carbon Stock Assessment Using Remote Sensing and Forest Inventory Data in Savannakhet, Lao PDR”, *Remote Sensing*, vol. 6, n.º 6, pp. 5452-5479, 2014. DOI: <https://doi.org/10.3390/rs6065452>
- [4] S. M. Raciti, L. R. Hutyra y J. D. D. Newell, “Mapping carbon storage in urban trees with multi-source remote sensing data: Relationships between biomass, land use, and demographics in Boston neighborhoods”, *Science of the Total Environment*, vol. 1, pp. 72-83, 2014. DOI: <https://doi.org/10.1016/j.scitotenv.2014.08.070>
- [5] F. Maselli, F. P. Vaccari, M. Chiesi, S. Romanelli y L. P. D’Acqui, “Modelling and analyzing the water and carbon dynamics of Mediterranean macchia by the use of ground and remote sensing data”, *Ecological Modelling*, vol. 351, pp. 1-13, 2017. DOI: <https://doi.org/10.1016/j.ecolmodel.2017.02.012>
- [6] IPCC, “Forest Land”, *Guidelines for National Greenhouse Gas Inventories*, vol. 4, 2006.

- [7] D. Lu, “The Potential and Challenge of Remote Sensing-based Biomass Estimation”, *International Journal of Remote Sensing*, vol. 27, n.º 7, pp. 1297-1328, 2006. DOI: <https://doi.org/10.1080/01431160500486732>
- [8] G. Galindo García, E. Cabrera Montenegro, D. M. Vargas Galvis, H. R. Pabón Méndez, K. R. Cabrera Torres, A. P. Yepes Quintero, J. F. Phillips Bernal, D. A. Navarrete Encinales, Á. J. Duque Montoya, M. C. García Dávila y M. F. Ordóñez Castro, *Estimación de la biomasa aérea usando datos de campo e información de sensores remotos*, Bogotá: Ideam, 2011.
- [9] M. Segura, M. Kanninen y D. Suárez, “Allometric Models for Estimating Aboveground Biomass of Shade Trees and Coffee Bushes Grown Together”, *Agroforestry Systems*, vol. 68, n.º 2, pp. 143-150, 2006. DOI: <https://doi.org/10.1007/s10457-006-9005-x>
- [10] J. Chave, C. Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Fölster, F. Fromard, H. N., T. Kira, J. P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Riéra y T. Yamakura, “Tree allometry and improved estimation of carbon stocks and balance in tropical forests”, *Oecologia*, vol. 145, pp. 87-99, 2005. DOI: <https://doi.org/10.1007/s00442-005-0100-x>
- [11] A. Valerio, M. Herold, H. Matieu y C. Schullius, “Mapping Biomass with Remote Sensing: A Comparison of Methods for the Case Study of Uganda”, *Carbon Balance and Management*, vol. 6, n.º 1, p. 7, 2011. DOI: <https://doi.org/10.1186/1750-0680-6-7>
- [12] S. K. von Bueren, A. Burkart, A. Hueni, U. Rascher, M. P. Tuohy y I. J. Yule, “Deploying four optical UAV-based sensors over grassland: challenges and limitations”, *Biogeosciences*, vol. 12, pp. 163-175, 2015. DOI: <https://doi.org/10.5194/bg-12-163-2015>
- [13] N. Yastikli, I. Bageci y C. Beser, “The Processing of Image Data Collected by Light UAV Systems for GIS Data Capture and Updating”, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-7/W2 (noviembre), pp. 267-70, 2013. DOI: <https://doi.org/10.5194/isprsarchives-XL-7-W2-267-2013>
- [14] J. Torres-Sánchez, F. López-Granados, A. I. De Castro y J. M. Peña-Barragán, “Configuration and Specifications of an Unmanned Aerial Vehicle (UAV) for Early Site Specific Weed Management”, *PLoS ONE*, vol. 8, n.º 3, p. e5821, 2013. DOI: <https://doi.org/10.1371/journal.pone.0058210>
- [15] F. Remondino, L. Barazzetti, F. Nex, M. Scaioni y D. Sarazzi, “Uav Photogrammetry for Mapping and 3D Modeling – Current Status and Future Perspectives”, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII(1/C22), pp. 25-31, 2011. DOI: <https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-25-2011>
- [16] J. P. Dandois y E. C. Ellis, “High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision”, *Remote Sensing of Environment*, vol. 136, pp. 259-276, 2013. DOI: <https://doi.org/10.1016/j.rse.2013.04.005>
- [17] C. Song y C. E. Woodcock, “Monitoring forest succession with multitemporal Landsat images: factors of uncertainty”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, n.º 11, pp. 2557-2567, 2003. DOI: <https://doi.org/10.1109/TGRS.2003.818367>

- [18] E. Honkavaara, R. Arbiol, L. Markelin, L. Martinez, M. Cramer, S. Bovet, L. Chandelier, R. Ilves, S. Klonus, P. Marshal, D. Schlöpfer, M. Tabor, C. Thom y N. Veje, “Digital Airborne Photogrammetry—A New Tool for Quantitative Remote Sensing?—A State-of-the-Art Review On Radiometric Aspects of Digital Photogrammetric Images”, *Remote Sensing*, vol. 1, n.º 3, pp. 577-605, 2009. <https://doi.org/10.3390/rs1030577>
- [19] A. Agapiou, D. G. Hadjimitsis, C. Papoutsas, D. D. Alexakis y G. Papadavid, “The Importance of Accounting for Atmospheric Effects in the Application of NDVI and Interpretation of Satellite Imagery Supporting Archaeological Research: The Case Studies of Palaepaphos and Nea Paphos Sites in Cyprus.”, *Remote Sensing*, vol. 3, n.º 12, pp. 2605-2629, 2011. DOI: <https://doi.org/10.3390/rs3122605>
- [20] M. A. Homem Antunes, J. M. Gleriani y P. Debiasi, “Atmospheric Effects on Vegetation Indices of Tm and Etm + Images From a Tropical Region Using the 6S Model”, *IEEE International Geoscience and Remote Sensing Symposium*, pp. 6549-6552, 2012. DOI: <https://doi.org/10.1109/IGARSS.2012.6352099>
- [21] M.-L. Smith, J. Anderson y M. Fladeland, “Forest canopy structural properties Chapter 14”, en *Field Measurements for Forest Carbon Monitoring: A Landscape-Scale Approach*, Coeli M. Hoover, Ed., Nueva York: Springer Science, Business Media, 2008, pp. 179-196.
- [22] P. V. Bolstad y S. T. Gower, “Estimation of Leaf Area Index in Fourteen Southern Wisconsin Forest Stands Using a Portable Radiometer”, *Tree Physiology*, vol. 7, pp. 115-124, 1990. DOI: <https://doi.org/10.1093/treephys/7.1-2-3-4.115>
- [23] M. Sprintsin, S. Cohen, K. Maseyk, E. Rotemberg, J. Grünzweig, A. Karnieli, P. Berliner y D. Yakir, “Long term and seasonal courses of leaf area index in a semi-arid forest plantation”, *Agricultural and Forest Meteorology*, vol. 151, pp. 565-574, 2011. DOI: <https://doi.org/10.1016/j.agrformet.2011.01.001>
- [24] J. M. Chen y J. Cihlar, “Plant Canopy Gap-Size Analysis Theory for Improving Optical Measurements of Leaf-Area Index”, *Applied Optics*, vol. 32, n.º 27, p. 6211, 1995. DOI: <https://doi.org/10.1364/AO.34.006211>
- [25] L. R. Holdridge, *Ecología basada en zonas de vida*, Costa Rica: Agroamérica, 1987.
- [26] J. Acevedo, S. Acosta Arrubla, S. Aranzales, R. d. M. Bedoya, M. J. García, O. A. Jojoa, A. Osorio y S. Vázquez, *Plan de Manejo Ambiental para el Jardín Botánico de la Universidad del Quindío*, Armenia: 2014.
- [27] O. E. Peláez Martínez, *Análisis de la respuesta espectral de las coberturas vegetales de los ecosistemas de páramos y humedales a partir de los sensores aerotransportados Utracam D, DJI Phantom 3 Pro y MAPIR NIR. Caso de estudio Humedal “El Ocho”, Villamaría - Caldas*, tesis de maestría, Universidad Católica de Manizales, 2017.
- [28] N. Sánchez Martín, B. Arias Pérez, D. González Aguilera y J. Gómez Lahoz, “Análisis aplicado de métodos de calibración de cámaras para usos fotogramétricos”, en *VIII Congreso Nacional de Topografía y Cartografía TOPCART 2004*, 2004.

- [29] Mapir, *Camera Reflectance Calibration Ground Target Package*, San Diego: Mapir, 2017, pp. 1-2.
- [30] J. Tian, L. Wang, X. Li, H. Gong, C. Shi, R. Zhong y X. Liu, “Comparison of UAV and WorldView-2 imagery for mapping leaf area index of mangrove forest”, *International Journal of Applied Earth Observation and Geoinformation*, n.º 61, pp. 22-31, 2017. DOI: <https://doi.org/10.1016/j.jag.2017.05.002>
- [31] C. A. Rokhmana, “The potential of UAV-based remote sensing for supporting precision agriculture in Indonesia”, *Procedia Environmental Sciences*, n.º 24, pp. 245-253, 2015. DOI: <https://doi.org/10.1016/j.proenv.2015.03.032>
- [32] A. Lisita, E. E. Sano y L. Durieux, “Identifying potential areas of Cannabis sativa plantations using object-based image analysis of SPOT-5 satellite data”, *International Journal of Remote Sensing*, vol. 34, n.º 15, pp. 5409-28, 2013. DOI: <https://doi.org/10.1080/01431161.2013.790574>
- [33] A. E. Zanne, G. Lopez-Gonzalez, D. A. Coomes, J. Ilic, S. Jansen, S. L. Lewis, R. B. Miller, N. G. Swenson, M. C. Wiemann y J. Chave, “Data from: Towards a worldwide wood economics spectrum. Dryad Digital Repository”, 2009. [en línea]. Disponible en <https://doi.org/10.5061/dryad.234>.
- [34] Ideam, *Protocolo para la estimación nacional y subnacional de Biomasa - Carbono en Colombia*, Bogotá: 2011.
- [35] J. M. Chen y T. A. Black, “Defining Leaf Area Index for Non-Flat Leaves”, *Plan, Cell and Environment*, vol. 15, n.º 4, pp. 421-429, 1992. DOI: <https://doi.org/10.1111/j.1365-3040.1992.tb00992.x>
- [36] GCOS, “Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC”, *World Meteorologica Organization GCOS*, vol. 138, 2010.
- [37] LI-COR (2010, sept.). “LI-COR” [en línea]. Disponible en <https://www.licor.com/documents/jlhuprnmuu6arl10s1t8g4nr1nlfuhat>.
- [38] J. W. Chason, D. D. Bladocchi y M. A. Huston, “A Comparison of Direct and Indirect Methods for Estimating Forest Canopy Leaf Area”, *Agricultural and Forest Meteorology*, vol. 57, n.º 1-3, pp. 107-128, 1991. DOI: [https://doi.org/10.1016/0168-1923\(91\)90081-Z](https://doi.org/10.1016/0168-1923(91)90081-Z)
- [39] J. O. Escalante Torrado, J. J. Cáceres Jiménez y H. Porras Díaz, “Ortomosaicos y modelos digitales de elevación generados a partir de imágenes tomadas con sistemas UAV”, *Revista Tecnura*, vol. 20, n.º 50, pp. 119-140, 2016.

# Testbed for Sub-Nyquist Wideband Spectrum Monitoring\*

Juan Felipe Medina Lee\*\*

Alexander López-Parrado\*\*\*

Alexander Vera Tasamá\*\*\*\*

Duvier de Jesús Bohórquez Palacio\*\*\*\*\*

Recibido: 23/11/2018 • Aceptado: 07/12/2019

<https://doi.org/10.22395/rium.v19n37a2>

## Abstract

Radioelectric spectrum management is a concern for today's world, mainly due to the misuse that has been given to this resource through the years, especially on the UHF band. To address this problem, a testbed for sub-Nyquist Wideband Spectrum Monitoring was built, that includes a web interface to remotely measure occupancy of the UHF band.

To achieve the above, an RF interface that allows tuning UHF frequencies with an instantaneous bandwidth of 95 MHz was built. Afterwards, a Random Demodulator was connected, and then an embedded system performed sub-Nyquist sampling and spectrum recovery. The embedded system connected to an information system that serves a web page, through which remote users can perform UHF band monitoring.

Experimental results showed that spectrum sensing can be achieved by using different algorithms on certain sparse spectra. In addition, the aforementioned web interface allowed simultaneous user connections, in order to perform independent measurements by sharing a hardware subsystem.

**Keywords:** Compressed sensing; sampling methods; information systems; system-on-chip; analog processing circuits; web services.

---

\* Investigación en curso, Proyecto 842 "Banco de Pruebas de Sistemas de Sensado de Espectro de Banda Ancha Basado en Muestreo Sub-Nyquist" financiado por la Universidad del Quindío (18 meses).

\*\* Ingeniero electrónico (Universidad del Quindío) y magíster en ingeniería con énfasis en electrónica (Universidad del Valle). Correo electrónico: jfmedina@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0003-4489-4280>.

\*\*\* Ingeniero electrónico (Universidad del Quindío), magíster en ingeniería electrónica (Universidad del Valle) y doctor en ingeniería (Universidad del Valle). Áreas de investigación: procesamiento digital de señales, telecomunicaciones, computación de alto rendimiento. Correo electrónico: parrado@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-0274-6901>.

\*\*\*\* Ingeniero electrónico y especialista en radiocomunicaciones (Universidad del Quindío). Doctor en Ingeniería electrónica (Universidad del Valle). Correo electrónico: avera@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-4183-9545>.

\*\*\*\*\* Ingeniero Electrónico, Universidad del Quindío (2020). Correo electrónico: ddbohorquezp@uqvirtual.edu.co. Orcid: <https://orcid.org/0000-0002-9855-4110>.

## Banco de pruebas para monitoreo *Sub-Nyquist* de espectro de banda ancha

### **Resumen**

La gestión del espectro radioeléctrico es una preocupación en la actualidad, hecho derivado, ante todo, del mal uso que se ha dado a este recurso a través de los años, especialmente en la banda de UHF. Para afrontar este problema, se construyó un banco de pruebas para la supervisión del espectro de banda ancha a través de muestreo sub-Nyquist, el cual incluye una interfaz web para medir de forma remota la ocupación de la banda UHF. Para lograr esto, se construyó una interfaz RF que permitiría sintonizar frecuencias UHF con un ancho de banda instantáneo de 95 MHz. Después, se conectó un demodulador aleatorio; y luego, un sistema embebido realizaría el muestreo sub-Nyquist y la recuperación del espectro. Este se conectaría, a su turno, con un sistema de información que sirve una página web, a través de la cual los usuarios remotos pueden realizar la supervisión de la banda de UHF. Los resultados muestran que la detección del espectro se puede lograr mediante diferentes algoritmos en ciertos espectros dispersos. Además, la interfaz web permitió que existiesen conexiones de usuario simultáneas, de tal manera que se realizaran mediciones independientes compartiendo el subsistema de *hardware*.

**Palabras clave:** detección compresiva, métodos de muestreo, sistemas de información, sistema en un solo chip, circuitos de procesamiento analógico, servicios Web.

## Banco de testes para monitoramento *sub-Nyquist* de espectro de banda larga

### **Resumo**

O gerenciamento do espectro radioelétrico é uma preocupação na atualidade, fato derivado, inicialmente, do mau uso que se tem dado a esse recurso através dos anos, especialmente na banda de UHF. Para enfrentar esse problema, construiu-se um banco de testes para a supervisão do espectro de banda larga por meio de amostragem sub-Nyquist, a qual inclui uma interface web para medir de forma remota a ocupação da banda UHF. Para isso, construiu-se uma interface RF que permitiria sintonizar frequências UHF com uma largura de banda instantânea de 95 MHz. Em seguida, ligou-se um demodulador aleatório; logo, um sistema embebido realizaria a amostragem sub-Nyquist e a recuperação do espectro. Este se ligaria, por sua vez, com um sistema de informação que serve um site, através do qual os usuários remotos podem realizar a supervisão da banda de UHF. Os resultados mostram que a detecção do espectro pode ser conseguida mediante diferentes algoritmos em certos espectros dispersos. Além disso, a interface web permitiu que existissem conexões de usuário simultâneas, de tal maneira que se realizassem medidas independentes compartilhando o subsistema de hardware.

**Palavras-chave:** detecção compressiva, métodos de amostragem, sistemas de informação, sistema em um chip, circuitos de processamento analógico, serviços web.



## INTRODUCTION

The radioelectric spectrum is not being used efficiently. The main reason for this is that frequency bands are assigned to primary users (PU), which do not allow any other service to use their bands, even though they are only used occasionally [1]. An example of misuse of the spectrum are television bands: Those ones are not being used anymore for analog television transmission, therefore they are in a re-licensing process for digital TV [2] and WRAN services [3].

The main goal of cognitive radio (CR) is to provide an RF system capable of knowing its environment and adjusting its parameters accordingly [4]. This is a promising solution for the misuse of the radioelectric spectrum in the next years. One of the most important components of a CR is the spectrum sensing function, which detects spectral holes and primary users in the environment. This information is then used by the radio system to decide which band to use and transmit the data.

Preliminary work on cognitive radio (CR) focused on its technology modeling, simulation, and implementation. Its main issues consider three levels: Spectrum sensing [5][6][7], spectrum access [8], and hardware implementation [5]. So, the current trends in ultra-wideband wireless communications have increased the complexity of the CR systems, particularly its hardware receivers.

Some studies investigating CR have been carried out on hardware experiments and prototypes, focusing in spectrum sensing techniques such as analog-to-information converters (AIC), modulated wideband converters (MWC), and multi-coset samplers (MC).

Many approaches have been proposed on designing and implementing AIC subsystems, such as theories and simulations where the demodulation, as well as the mixer components, define the primary design specifications. A transistor-level simulation for an AIC is described in [9], which uses a mixer based on a Gilbert cell and CMOS 0.13 $\mu\text{m}$  fabrication technology, achieving a sampling frequency of 100 MSample/s. On the other hand, [10] presents a behavioral model which considers the dominant parameters for a detailed design to implement an AIC, with a bandwidth of 40 MHz. In [11], the authors present a working prototype of AIC, which is based on a commercial digital signal processor (DSP) board that has a clock frequency of 160 MHz; however, the AIC bandwidth is 400 kHz. Very-large-scale integration (VLSI) design level allows to research also for implementation parameters of AIC with this technology: In this regard, [12] presents a VLSI design of an AIC with CMOS 28 nm fabrication technology, which can sample RF signals up to 6 GS/s. In [13], the authors present an AIC built with integrated circuits (IC), which can sample signals with a

bandwidth of up to 236 MHz: Its design specifies a printed circuit board (PCB), a commercial IC, and a chipset designed with 0.45  $\mu\text{m}$  InP HBT (Indium-Phosphide heterojunction bipolar transistor) technology.

Regarding the MWC subsystems, there are three relevant works related to hardware implementation. A recent contribution of [14], concerning to a hardware prototype, proposes a system based on a FPGA, a commercial ADC, and environments such as *Matlab* and *Labview*; it has an unknown total bandwidth occupancy of up to 200 MHz and Nyquist rate of 6 GHz. On the other hand, a doctoral thesis developed at Stanford University [15] presents the VLSI implementation of an MWC receiver, which is able to detect four 1.4 MHz channels up to 900 MHz. Finally, in [16] the authors describe the *Xampling* platform based on MWC with support to sample signals with 2 GHz Nyquist rate, 120 MHz spectrum occupancy, and a sampling rate of 280 MHz.

Regarding MC-based sampling systems, the main findings show how it has gained popularity. A MC sampler prototype based on discrete integrated circuits is presented in [17], where an FPGA can sample signals with a bandwidth of 100 kHz and 10 sampling channels. In [18], the authors present an implementation of a MC-based sub-Nyquist sampling scheme for a surveillance radar with 1.1 GHz bandwidth; however, no details of the hardware implementation parameters are offered. Moreover, an algorithm for broadband spectrum sensing based on sFFT is presented [19], which uses two sampling channels: Its implementation was carried out by using the software-defined radio (SDR) platform USRP [15], with a bandwidth up to 900 MHz. In [20], an algorithm to perform cooperative wideband spectrum sensing (and its hardware implementation) is presented, based on MC sampling: It is robust to noise, overcoming the performance of the algorithm presented in [19]. On the other hand, another doctoral thesis is presented by the Georgia Institute of Technology [19], which describes an asynchronous multi-rate sampling system, similar to the MC sampling: Its hardware implementation can achieve a bandwidth up to 18 GHz.

Recent evidence suggests that CR receivers can monitor wide spectra by sampling at very low rates, such as [21], which the authors describe as “the first hardware implementation results”, achieving a sampling frequency of 200 MHz and resolving an instantaneous bandwidth of 2GHz by using two ADCs, an FPGA demo board, and other commercial integrated circuits.

As for complete testbed implementations for CR, all the studies reviewed so far suffer from the commercial availability of implementations about sub-Nyquist sampling systems, especially as testbeds for the academic and scientific community. Nevertheless, the embedded platform BEE2 (Berkeley Emulation Engine 2, developed at the University of California) for high performance computing is presented in [22] as a

framework with 5 high performance FPGA, for designing high-end reconfigurable and complex computers. This platform was used by [23] to implement a spectrum sensing system, however, this is not based on sub-Nyquist sampling. On the other hand, an FPGA-based platform, called CRUSH by the authors [24], was conceived as a testbed for spectrum sensing systems. CRUSH is also based on the SDR platform USRP [25], and only supports NSS (Narrowband Spectrum Sensing).

Considering that testbeds are very useful tools to assess the performance of different CR algorithms and techniques, and the fact that many authors offer no explanation about the detailed design for the hardware implementation and electronic support systems, the main contribution of the work presented in this paper is a configurable testbed with the necessary circuits to capture real wireless signals by using sub-Nyquist sampling. Hence, our system would be at the same level as the ones implemented at Israel Technology Institute and California University [22][26]. In addition, this testbed would allow students to make different experiments with cutting-edge technology communication devices, to change configuration parameters for tuning the system, or testing new algorithms. Finally, remote access to our testbed opens the door for collaborative work between various universities, and thus it could have a greater impact in the academic community.

The paper is organized as follows: Section 1 describes the testbed architecture, starting with the RF system and analog frontend, the hardware/firmware configuration for the SoC-FPGA, and the information system. Section 2 presents the experiments performed to the testbed: Verification results for both the analog frontend and the SOC/FPGA system, and tests for the web server. In Section 3, a conclusion is drawn, and acknowledgements are given in Section 4.

## 1. TESTBED ARCHITECTURE

In this section, different components of the sensing platform are described, starting with the RF components (antenna, and RF interface). After that, the analog front end between the RF sub-system and the SOC-FPGA will be presented. The chapter ends with the description of the information management service for the spectrum sensing platform.

### 1.1 RF System

An antenna that would allow to capture wideband signals in the UHF frequency range (between 300 MHz and 3 GHz) with a good gain was needed. The *M-Series Ultra Base Station* antenna [27] was a great option for our requirements, since it covers the frequency range from 25 MHz to 6 GHz in obstructed non-line-of-sight

(NLOS) environments, and it has omnidirectional detection. shows the Voltage Standing Wave Ratio (VSWR) for the Super-M Ultra Base Station antenna: It is lower than 2.0 from frequencies greater than 25 MHz, making detection possible for this frequency range [27].

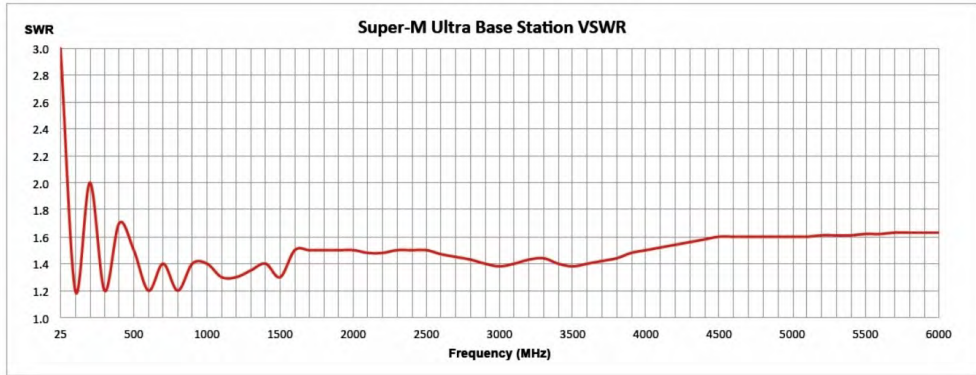


Figure 1. Super-M Ultra Base Station VSWR

Source: [27].

For the radio frequency (RF) interface, an analog system that did not limit our bandwidth with a built-in low pass filter (LPF), like most commercial devices, was needed. Hence, the Myriad RF board [28] was selected: It is a low cost universal radio development platform, based on LMS6002DFN transceiver (.). This board can receive incoming signals between 0.3 and 2.8 GHz, and down-convert them to baseband with a bandwidth of 14 MHz in normal operation. This board allows the developer to bypass the built-in lowpass-filters with a simple configuration, and to increase the bandwidth to 95 MHz.

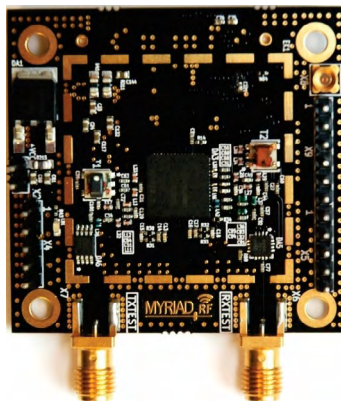


Figure 2. Myriad-RF board

Source: [28].

This board has two SMA connectors for the antenna interface, and it is equipped with pin headers for power supply, reference clock, analog I/Q input/output, and SPI interface connections. It contains everything needed to connect it with baseband chipsets or FPGAs, allowing the user to set parameters such as central frequency, gain, bandwidth, or sampling rates.

## 1.2 Architecture of Designed AIC

An Analog-to-Information Converter (AIC) based on Random Demodulator (RD) was designed by using block diagram in [29] [9].

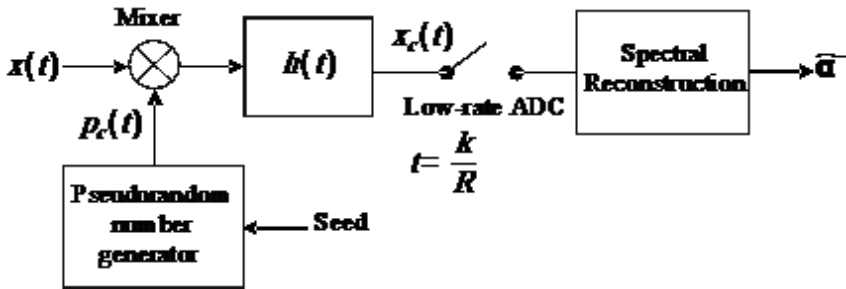


Figure 3. AIC based on RD

Source: Prepared by the authors.

In ,  $x(t)$  is a signal which is representable and compressible in Fourier basis;  $\psi_n(t) = e^{j\omega_n t}$ ;  $p_c(t)$  is a pseudo-random square wave with chipping rate of BW Hz, where BW is the bandwidth of  $x(t)$ ; and  $h(t)$  is the impulse response of an anti-alias low-pass filter. Then, Spectral Reconstruction block has to solve the Compressed Sensing (CS) problem stated in equation (1).

$$\min_{\tilde{\mathbf{a}} \in \mathbb{C}^N} \|\tilde{\mathbf{a}}\|_1 \text{ subject to } \mathbf{y} = \mathbf{\hat{E}} \tilde{\mathbf{a}} \quad (1)$$

Where  $\tilde{\mathbf{a}} \in \mathbb{C}^N$  is the estimated sparse vector,  $\mathbf{y} \in \mathbb{C}^M$  is the vector containing the sub-Nyquist samples, and  $\mathbf{\hat{E}} \in \mathbb{C}^{M \times N}$  is the measurement matrix which is computed according to equation (2).

$$\mathbf{\hat{E}}_{k,n} = \int_{-\infty}^{\infty} e^{j\omega_n \tau} p_c(\tau) h(k/R - \tau) d\tau \quad (2)$$

Designed AIC can recover signals with  $BW=100$  MHz by using a sub-Nyquist rate of  $R=4$  MHz, which leads to a downsampling ratio of  $N/M=25$ . It was implemented by using off-the-shelf analog components and a SoC FPGA-based board. In this case, mixer and anti-alias filter were constructed by using an integrated Gilbert cell and a first

order low-pass filter, respectively. On the other hand, generation of , analog to digital conversion, and spectral reconstruction were implemented by means of the ADC-SoC board, by taking advantage of both FPGA and HPS portions of Cyclone V chip [30][31].

Mixer circuit performs multiplication of  $x(t)$  and  $p_c(t)$ , and it was implemented by using a Gilbert cell built from the HFA-3101 high-speed transistor [32]. shows designed Gilbert cell.

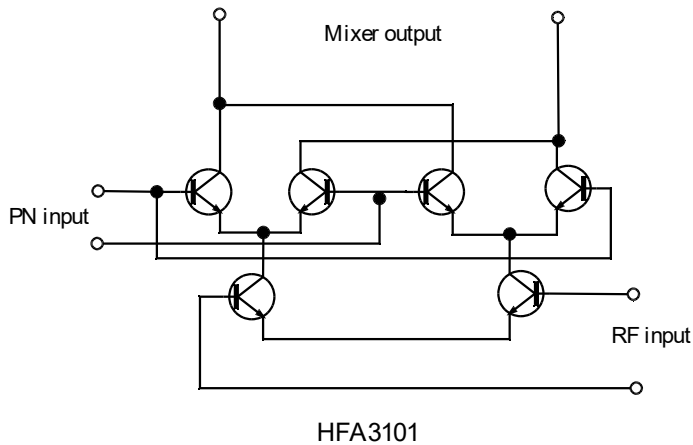


Figure 4. Designed Gilbert Cell

Source: Prepared by the authors.

In, biasing circuit was not included for the sake of simplicity [32]. Nonetheless, to perform mixing,  $x(t)$  and  $p_c(t)$  are required to be low voltage differential signals, since mixer has an approximate gain of  $12 \times$ . In addition, to feed compressed signal into ADC in ADC-SoC board, it is necessary to convert the differential signal from mixer output to single-ended, and to perform low-pass filtering. Both operations were accomplished by using circuit in.

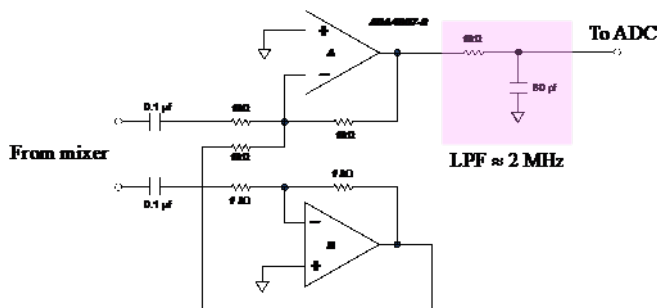


Figure 5. Circuit for Differential to Single-ended Conversion and Low-pass Filtering

Source: Prepared by the authors.

Circuit in was built from the operational amplifier ADA-4807 [33], where conversion from differential to single-ended is performed with unity gain, and the output stage is a first-order low-pass filter with transfer function given in equation (3).

$$H(s) = \frac{1}{8 \times 10^{-8} s + 1} \tag{3}$$

From equation (3) it can be observed that cut-off frequency of filter is  $f_c \approx 2$  MHz, therefore sub-Nyquist rate can be up to  $R=4$  MHz. Finally, DC-blocking capacitors were used to isolate bias from mixer.

SoC-FPGA chip was used to implement the digital sub-system of designed AIC. FPGA-side was used to manage ADC and to generate a PN sequence, and HPS-side was used to manage circuit on FPGA and to execute spectrum recovery algorithms. shows the implemented circuit on FPGA-side.

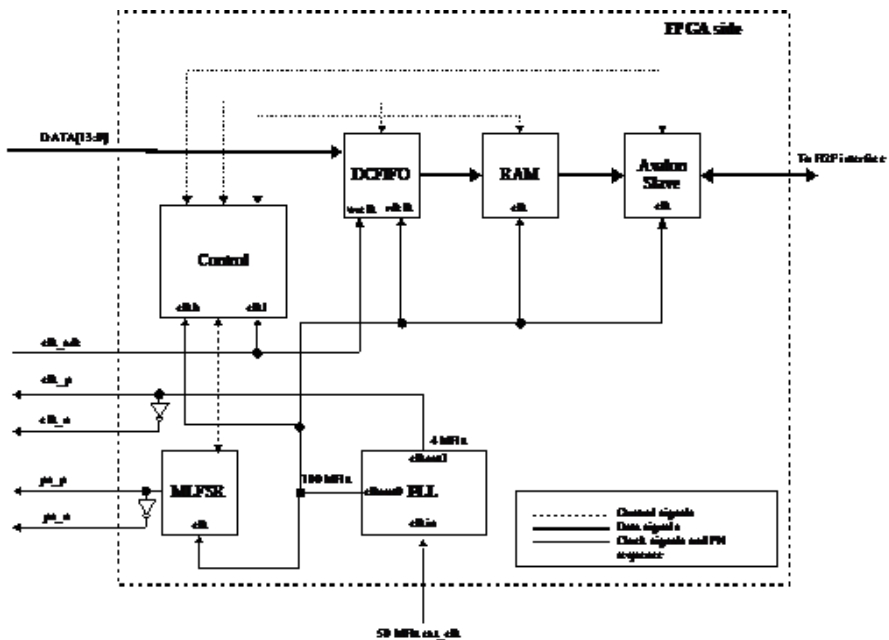


Figure 6. Circuit for AIC on FPGA-side of SoC-FPGA

Source: Prepared by the authors.

Circuit in is composed of six systems: Phase-Locked Loop (PLL), Maximal Length Linear-Feedback Shift Register (MLFSR), Dual Clock First In First Out (DCFIFO) circuit, on-chip Random Access Memory (RAM), Avalon slave [34], and Control Circuit. PLL is used to generate two clock signals of 100 MHz and 4 MHz; 100 MHz

clock signal is used for synchronization of MLFSR, DCFIFO, on-chip RAM and HPS-To-FPGA interface [31]; and 4 MHz signal is used to synchronize external ADC chip. MLFSR is 9-bit wide, and uses primitive polynomial in  $p(x) = x^9 + x^5 + 1$  [35] to generate PN sequence. DCFIFO allows transfer of samples from ADC to on-chip RAM by using two clock domains. Avalon slave serves as gateway between FPGA-side and HPS-side, through which HPS can start a capture operation and read acquired sub-Nyquist samples. Finally, Control Circuit manages operation of AIC by using one 4-state Finite State Machine (FSM), whose ASM chart is shown in .

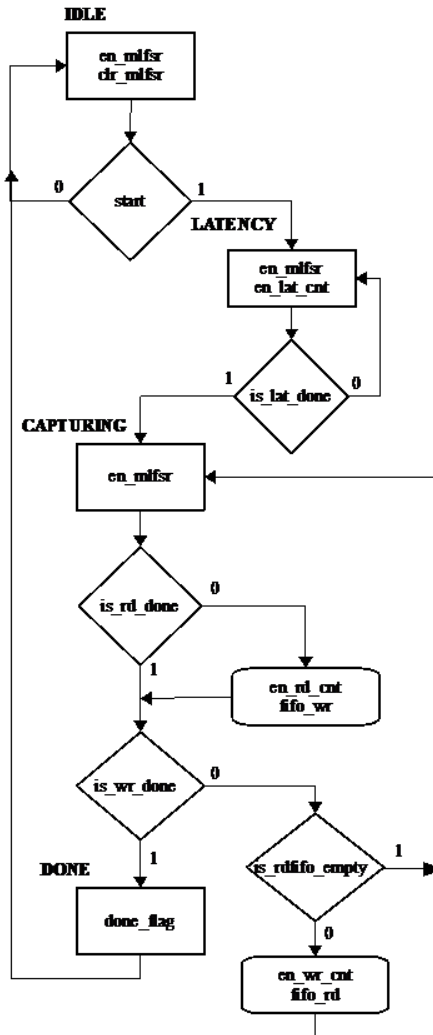


Figure 7. ASM Chart for Control FSM

Source: Prepared by the authors.



FSM operates along with three counters `lat_cnt`, `rd_cnt`, and `wr_cnt`; five flip-flop synchronizers for clock domain crossing; and one combinational circuit to control write operations of on-chip RAM. Operation of FSM is described as follows: In IDLE state, FSM waits for a capture operation, which is triggered when `start=1`. Then, it enters into Latency state, where it enables MLFSR and waits for 12-cycle latency of AD9254 ADC [36] by using `lat_cnt`: This ensures synchronization between  $p_c(t)$  signal and captured vector for proper RD operation [9] [37]. Once latency passes, FSM enters into Capture state, where samples from ADC are written into DCFIFO as long samples from DCFIFO are written into on-chip RAM: This operation is controlled by means of `rd_cnt`, `wr_cnt`, and `is_fifo_rdempty` flag from DCFIFO. Finally, capture end is signaled by `done_flag` in DONE state, after which FSM returns to IDLE state.

Regarding software for our AIC implementation, first we wrote a C program to achieve low level access to FPGA-side hardware through HPS-To-FPGA interface, and retrieve vector  $\mathbf{y}$  from RD. Then, we made a Python implementation of spectrum reconstruction algorithm Compressive Sampling Matching Pursuit (CoSaMP) to solve the CS problem in equation (1) [38].

### 1.3 Information Management Service

We designed and implemented a web application to manage the testbed data. This app allows the users to observe the captured spectra for configuring different parameters like bandwidth or central frequency, and for uploading different sensing algorithms to the embedded system.

The information management service has two main components: The web server, and the embedded sensing client. The main architecture of this system is shown in.

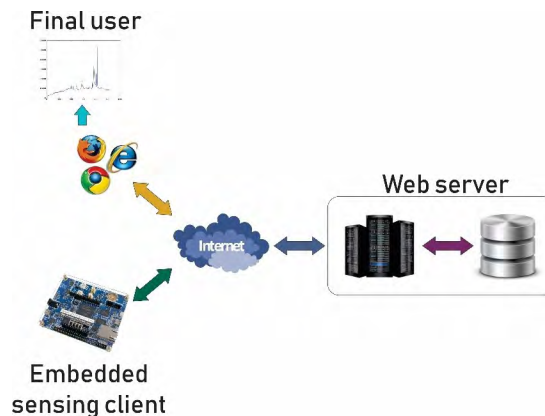


Figure 8. Global Architecture of the Information Management Service

Source: Prepared by the authors.

The functioning parameters for our system to work correctly are listed below:

- 32768-points spectra storage capacity.
- Easy visualization/download of historical data.
- Capable of uploading different sensing algorithms to the embedded sensing client.
- Ability to configure different parameters such as central frequency, bandwidth, capture period and spectral resolution.
- Compatible with the embedded system technology.

It was decided to use Python to develop both the web server and the embedded sensing client, for its compatibility with web applications and embedded systems [39]. A web server was implemented to interact with the human users and the embedded sensing client. In this regard, it was decided to use Flask [40], a Python framework used in web applications development, to implement our server.

The web server has four main tasks: Two of them are oriented to the interaction with the human client, and the others are focused on the embedded sensing client. The web server tasks overview is presented in.

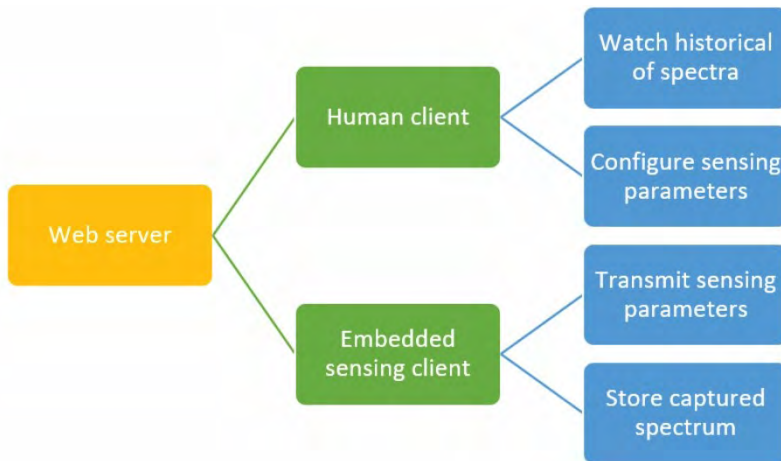


Figure 9. Web Server tasks Sorted by Type of Client

Source: Prepared by the authors.

All of the data needed by the web server are stored in a database with three tables: The first one is used to store the spectrum data, the second one is used for the configuration parameters, and the third stores the human user's information. Also, every spectrum captured by the embedded platform is stored, along with another intrinsic

parameters such as central frequency, bandwidth, frequency resolution, and timestamp. All the frequency components are stored as a text string separated by commas. shows the spectrum table of the database.

Table 1. Spectrum Table for the Web Server Database

Spectrum	
ID	INTEGER
DFT	MEDIUMTEXT
Central frequency	INTEGER
Bandwidth	INTEGER
Frequency resolution	INTEGER
Timestamp	TIMESTAMP

Source: Prepared by the authors.

The configuration parameters must stay still, unless someone changes them; for that reason, they are stored in another table of the database. These data are overwritten every time someone updates any parameter. The member “Pending updates” is used as a flag to check if the current sensing algorithm is up to date in the embedded system or not. shows the *configuration data* table of the database.

Table 2. Configuration Data Table for the Web Server Database

ConfigData	
ID	INTEGER
Central frequency	INTEGER
Bandwidth	INTEGER
Frequency resolution	FLOAT
Capture period	INTEGER
Current algorithm	MEDIUMTEXT
FFT points	INTEGER
Pending updates	MEDIUMTEXT
timestamp	TIMESTAMP

Source: Prepared by the authors.

The last table of our database is used to store basic information of the web application users. This information is used to demand a login of the user before he or she updates any configuration parameter and, therefore, to have some control over who can access the system. Stored information of human users is shown in . The login password is hashed before it is saved (using SHA-256), so the security of the users is not compromised.

Table 3. User Data Table for the Web Server Database

Spectrum	
ID	INTEGER
Username	MEDIUMTEXT
Password	MEDIUMTEXT
Email	MEDIUMTEXT

Source: Prepared by the authors.

The web server has a graphic user interface (GUI) to interact with the human clients using a web browser. It was developed using mostly HTML, but we some Flask utilities were also used to implement the web forms.

The main page of our interface allows the client to see or download every spectrum captured by the embedded sensing client. The plotted spectrum data must be selected by date, starting by the year and finishing by the hour in which they were captured. shows the GUI for the spectrum sensing visualization. This information can be accessed by anyone interested without restriction, and every single spectrum captured by the sensor can be downloaded in CSV format and imported from Matlab.

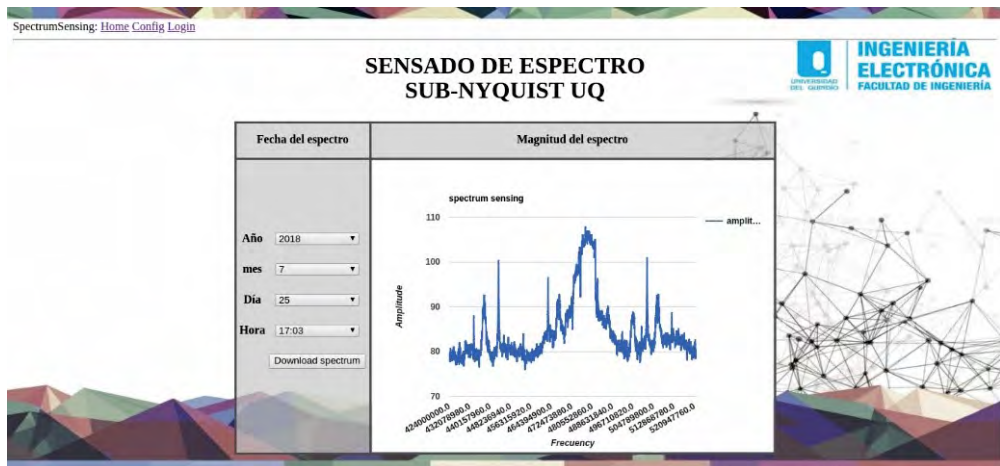


Figure 10. Spectrum Visualization GUI for the Web Application

Source: Prepared by the authors.

Other important page in our web application allows the user to change the configuration of the embedded sensing client using a simple user interface. The configuration page shown in permits the user to change the central frequency, desired bandwidth, spectral resolution, sensing period, and spectrum sensing algorithm.

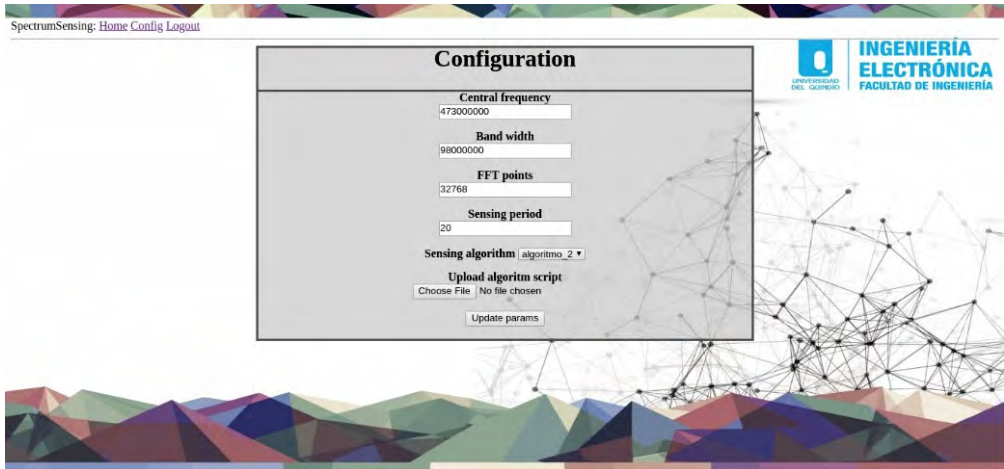


Figure 11. Spectrum Sensing Configuration GUI for the Web Application

Source: Prepared by the authors.

Every user must be logged in before he or she can change any of the parameters shown in . This action is performed in the *login page*, shown in . This page allows the user to create an account if he or she does not have one already.

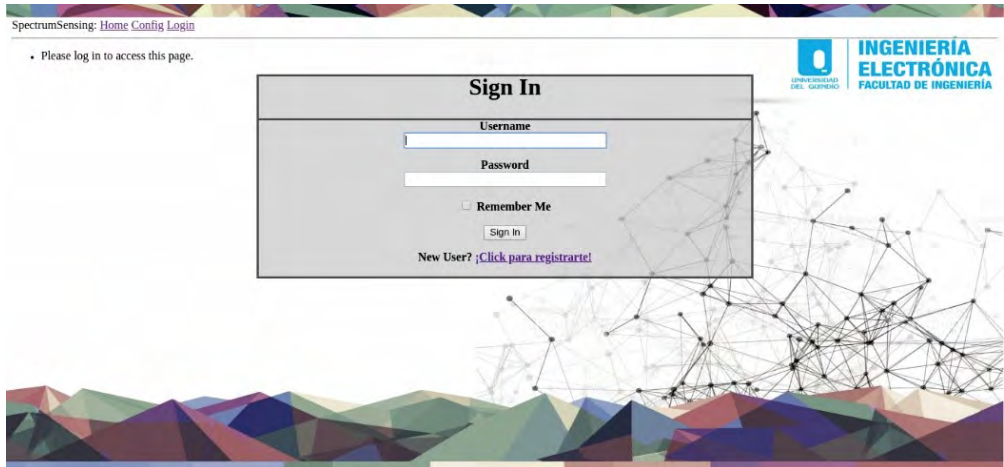


Figure 12. Login GUI for the Web Application

Source: Prepared by the authors.

Embedded platform performs spectrum sensing by means of a Python program that constantly sends every spectrum captured to the web server. This program also updates the sensing parameters and sensing algorithms using polling synchronization, which allows the embedded client to work correctly from anywhere in the world without having a public IP, just an Internet connection to communicate with the web server.

There are two periodic Python tasks running on the embedded platform using timers. The first one updates the sensing parameters, and the second one obtains the current spectrum and sends it to the web server. Both algorithms are shown in .

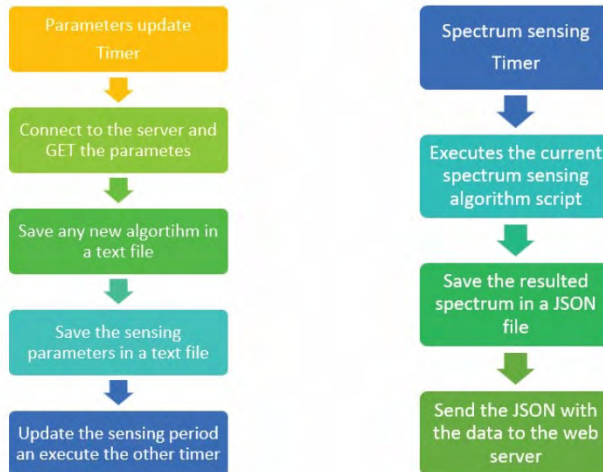


Figure 13. Flow Charts for the Embedded Platform Processes

Source: Prepared by the authors.

The *parameter update* timer is executed every 2 seconds. In every iteration, it requests the current parameters and saves them into a text file. The *spectrum sensing* timer is executed periodically, depending on the *sensing time* parameter. Each time the *spectrum sensing* timer is triggered, it asks FPGA for the current spectrum, then saves it in a JSON format, and sends it to the web server.

## 2. TESTBED VERIFICATION RESULTS

This section describes a set of tests that were performed to prove the correct operation of the built testbed and its sub-systems. First, AIC to verify the spectrum recovery of sparse signals; after that, the spectrum monitoring capabilities of testbed were verified, and finally, responsiveness of server to clients was measured.

### 2.1 AIC verification results

Designed AIC was tested by using a measurement window of 125  $\mu$ s, which leads to values for  $M$  and  $N$  of 500 and 12500, respectively. In addition, a multi-tone signal with frequencies of 4.17 MHz, 14.19 MHz, 25.02 MHz and 35.36MHz was used as a test vehicle. This signal was fed into the AIC, and then  $M=500$  samples were acquired and stored in  $y$  vector.

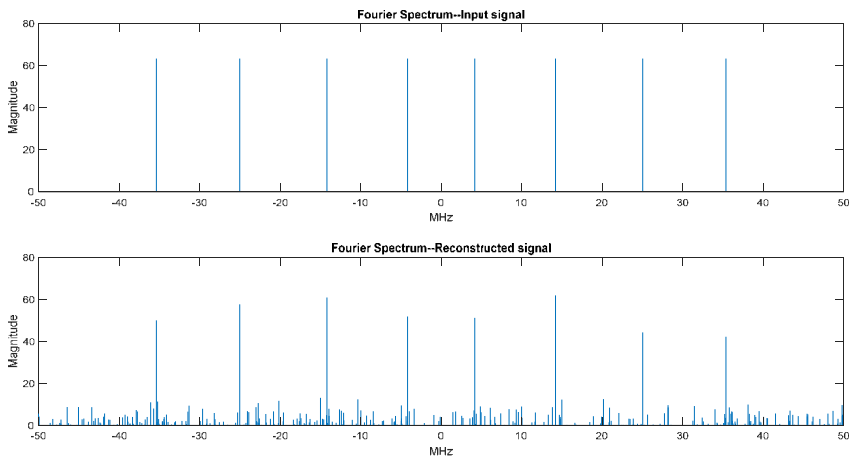


Figure 14. Spectrum of Multitone Signal and Recovered Spectrum

Source: Prepared by the authors.

Top of shows spectrum of test signal and bottom shows recovered spectrum when using CoSaMP algorithm. Results in show that CoSaMP did it well recovering support of  $\alpha$ : This can be verified when calculating average squared error, which amounts to 0.0005897302. Hence, this well-known algorithm was able to perform spectral reconstruction from samples acquired by our mixed-signal hardware.

Finally, shows synthesis results for the FPGA-side circuit in when implemented on the Cyclone V device of ADC-SoC board [30][31]. Results in demonstrate that digital hardware for designed AIC fits on a low-end SoC-FPGA chip, providing a complete low-cost sub-Nyquist sampling solution for compressible signals.

Table 4. Synthesis Results of AIC Hardware

Parameter	Value
Adaptive Logic Modules	1959 / 15880 (12 %)
Dedicated Logic Registers	2445 / 31760 (8 %)
Block Memory Bits	1628672 / 2764800 (59 %)
PLLs	1 / 5 (20 %)
Operating Frequency (MLFSR clock)	125 MHz
Operating Frequency (ADC clock)	300 MHz

Source: Prepared by the authors.

## 2.2 Spectrum monitoring capabilities

The testbed can perform spectrum monitoring around any frequency of UHF band with an instantaneous bandwidth of 100 MHz. This allowed to perform a large amount of measurements on this band; however, these results were narrowed down to demonstrate the correct operation of testbed.



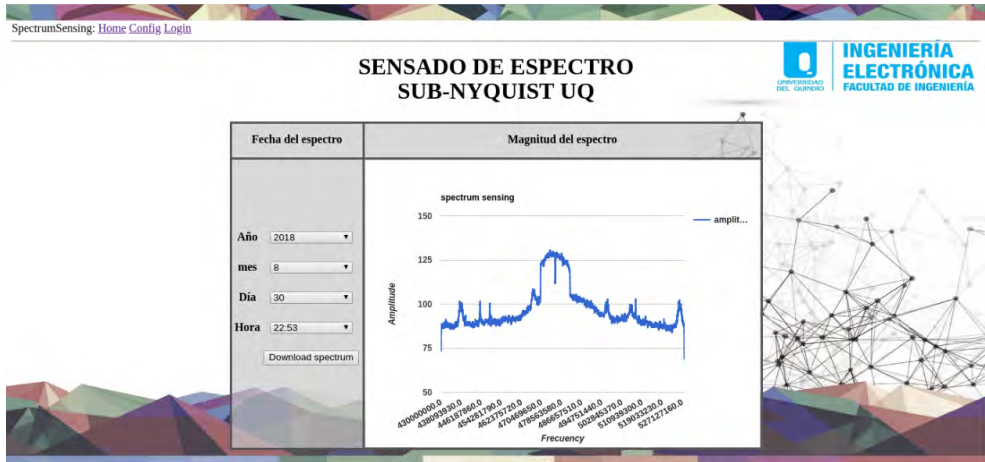


Figure 15. Spectrum for Colombian HD-TV Channels Caracol (473 MHz) and RCN (479 MHz)

Source: Prepared by the authors.

shows spectrum measurement around frequency 473 MHz, where two Colombian HD-TV channels reside. From this it is possible to observe two typical DVB-T2 spectral shapes, corresponding to TV channels Caracol (473 MHz) and RCN (479 MHz), whose spectrum occupancy could be monitored by using our testbed [41].

shows spectrum measurement around frequency 2655 MHz, where Long Term Evolution (LTE) downlink in band 7 resides [42], in this case, for carrier Claro in Colombia. The figure also depicts a typical LTE spectral shape for a channel with 15-MHz bandwidth [42].

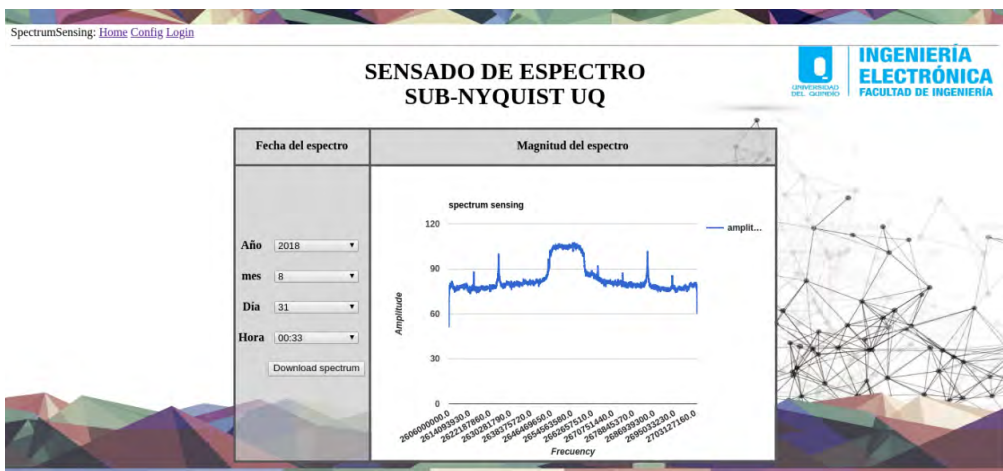


Figure 16. Spectrum for LTE Downlink in Band 7 of Carrier Claro in Colombia (2655 MHz)

Source: Prepared by the authors.



### 2.3 Web server traffic tests

The web server described in section 1 can be used by many people around the world to experiment with different sensing algorithms, and to observe the historical spectral data. Nevertheless, the web application is not aimed at handling high traffic or hundreds of requests at the same time. Some tests were performed using the Jmeter software [43] to analyze the performance of the web server when executing the following tasks:

- Watching spectral data.
- Configuring sensing parameters.
- Storing spectral data.

*Watching spectral data* is the most demanding task of the web server because it must find the desired spectrum (32.768-long) in the database according to the selected date, retrieve it to the browser, and plot it. Two experiments were performed to test the latency of this task. In the first one, there was only one client requesting the data 300 times. In the second one, 10 simultaneous clients were configured, each of them requesting 30 plots. Results of both experiments are shown in , which shows that the latency of the web server is a little higher than 8 seconds when 10 users are trying to watch the spectral data at the same time. If only one person is accessing the application, the latency is reduced to 0.86 seconds.

Table 5. Performance test for plotting spectral data.

	#samples	Avg.	Min.	Max.	Std. Dev.	Error %	Throughput	Received kB
1 client	300	860	829	1033	21.38	0.00 %	1.2/sec	2542.76
10 clients	300	8471	881	9498	875.62	0.00 %	1.1/sec	2484.51

Source: Prepared by the authors.

The performance of the web server when logging in and updating the sensing parameters for the spectrum sensing was also tested. Therefore, tests were performed for one user performing 500 operations, and 10 simultaneous users performing 100 operations each ( ).

Table 6. Performance Test for Log in and Updating Parameters

	#samples	Avg.	Min.	Max.	Std. Dev.	Error %	Throughput	Received kB
1 client Logging in	500	52	49	86	3.87	0.00 %	16.6/sec	70.73
1 client Upd. params	1000	7	6	26	1.3	0.00 %	33.2/sec	136.21
10 clients Logging in	1000	309	51	448	46.89	0.00 %	17.1/sec	73.05
10 clients Upd. params	2000	265	8	407	46.64	0.00 %	34.2/sec	140.47

Source: Prepared by the authors.

shows that this operation is a lot faster than watching spectral data, since it only takes about 52 ms to login one client, and 309 ms when ten clients are trying to do so at the same time. The times are shorter for updating parameters: Only 7 ms for updating one client, and 265 ms when ten people try to update parameters at the same time. Is worth to mention that ten people updating parameters at the same time probably would generate misbehaviors in the embedded sensing client.

The final test performed on the web server aimed to check how long did it take to store a spectrum captured by the embedded sensing client in the database. Only one experiment for this parameter was performed, since there will only be one embedded sensing client in our system. It can be concluded from , which shows the results of this experiment, that storing in database is a very fast operation for our server, since it only takes 72 ms (average) to store a 32768-long spectrum.

Table 7. Performance Test for Database Storing

	#samples	Avg.	Min.	Max.	Std. Dev.	Error %	Throughput	Received kB
1 client	500	72	34	107	5.96	0.00%	13.6/sec	2.13

Source: Prepared by the authors.

### 3. CONCLUSION

A testbed that constitutes a complete sub-Nyquist spectrum monitoring tool that can be used by researchers, spectrum management authorities, and hobbyists to carry out monitoring and research tasks was built. It is the first of its kind in Colombia, and shares unique features with a small group of tools around the world. On the other hand, the aforementioned testbed includes a sub-Nyquist AIC that can be used to recover sparse signals with a bandwidth of 100 MHz, and according to reviewed literature, it is the first one with such bandwidth which was built from off-the-shelf components. Finally, this testbed allows remote users to use custom algorithms, and to perform either spectrum monitoring or recovery of sparse signals.

### 4. ACKNOWLEDGMENT

The authors thank the Universidad del Quindío for supporting this project.

### REFERENCES

- [1] D. Cabric, I. D. O'Donnell, M. S.-W. Chen, and R. W. Brodersen, "Spectrum sharing radios", *IEEE Circuits Syst. Mag.*, vol. 6, n.º 2, pp. 30–45, 2006, DOI: 10.1109/MCAS.2006.1648988.

- 
- [2] European Telecommunications Standards Institute, “ETSI EN 300 744 Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television.” 2009.
- [3] IEEE, “IEEE 802.22-2011, Wireless Regional Area Networks (WRAN) - Specific requirements Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands”, 2011, DOI: 10.1109/IEEESTD.2011.5951707.
- [4] J. Mitola and G. Q. Maguire, “Cognitive radio: making software radios more personal”, *IEEE Pers. Commun.*, vol. 6, n.º 4, pp. 13–18, 1999, DOI: 10.1109/98.788210.
- [5] V. T. Nguyen, F. Villain, and Y. Le Guillou, “Cognitive Radio RF: Overview and Challenges”, *VLSI Des.*, vol. 2012, pp. 1-13, May 2012, DOI: 10.1155/2012/716476.
- [6] C.-P. Yen, Y. Tsai, and X. Wang, “Wideband Spectrum Sensing Based on Sub-Nyquist Sampling”, *IEEE Trans. Signal Process.*, vol. 61, n.º 12, pp. 3028–3040, Jun. 2013, DOI: 10.1109/TSP.2013.2251342.
- [7] P. Palacios-Jativa and C. Saavedra-Arancibia, “Comparison of spectrum detection methods applied to cognitive mobile radio networks”, *Rev. Cienc. E Investig.*, vol. 3, n.º 10, pp. 16–20, 2018, DOI: 10.26910/issn.2528-8083vol3iss10.2018pp16-20p.
- [8] M. M. Mabrook, G. A. Fahmy, A. I. Hussein, and M. A. Abdelghany, “Novel adaptive non-uniform sub-Nyquist sampling technique for cooperative wideband spectrum sensing”, in *2016 12th International Computer Engineering Conference (ICENCO)*, Dec. 2016, pp. 20–25, DOI: 10.1109/ICENCO.2016.7856439.
- [9] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, “Theory and Implementation of an Analog-to-Information Converter using Random Demodulation”, in *2007 IEEE International Symposium on Circuits and Systems*, May 2007, pp. 1959–1962, DOI: 10.1109/ISCAS.2007.378360.
- [10] S. Kirolos, T. Ragheb, J. Laska, M. F. Duarte, Y. Massoud, and R. G. Baraniuk, “Practical Issues in Implementing Analog-to-Information Converters”, in *2006 6th International Workshop on System on Chip for Real Time Applications*, Dec. 2006, pp. 141–146, DOI: 10.1109/IWSOC.2006.348224.
- [11] T. Ragheb, J. N. Laska, H. Nejati, S. Kirolos, R. G. Baraniuk, and Y. Massoud, “A prototype hardware for random demodulation based compressive analog-to-digital conversion”, in *Midwest Symposium on Circuits and Systems*, 2008, pp. 37-40, DOI: 10.1109/MWSCAS.2008.4616730.
- [12] D. E. Bellasi, L. Bettini, C. Benkeser, T. Burger, Q. Huang, and C. Studer, “VLSI design of a monolithic compressive-sensing wideband analog-to-information converter”, *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 3, n.º 4, pp. 552–565, 2013, DOI: 10.1109/JETCAS.2013.2284618.

- [13] M. Wakin *et al.*, “A Nonuniform Sampler for Wideband Spectrally-Sparse Environments”, *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 2, n.º 3, pp. 516–529, Sep. 2012, DOI: 10.1109/JETCAS.2012.2214635.
- [14] D. Cohen, S. Tsiper, and Y. C. Eldar, “Analog-to-Digital Cognitive Radio: Sampling, Detection, and Hardware”, *IEEE Signal Process. Mag.*, vol. 35, n.º 1, pp. 137–166, 2018, DOI: 10.1109/MSP.2017.2740966.
- [15] D. J. K. Adams, “A Practical Implementation of the Modulated Wideband Converter Compressive Sensing Receive Architecture”, Stanford University, 2016.
- [16] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, “Xampling: Analog to Digital at Sub-Nyquist Rates”, *IET Circuits, Devices & Systems*, vol. 5, n.º 1, pp. 8-20, January 2011, DOI: 10.1049/iet-cds.2010.0147.
- [17] Zhang Jingchao, Liu Peizhuo, Fu Ning, and Peng Xiyuan, “Prototype design of multicore sampling based on compressed sensing”, in *12th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 2015, pp. 1303–1308, DOI: 10.1109/ICEMI.2015.7494524.
- [18] M. Yaghoobi, B. Mulgrew, and M. E. Davies, “An efficient implementation of the low-complexity multi-coset sub-Nyquist wideband radar electronic surveillance”, in *2014 Sensor Signal Processing for Defence (SSPD)*, Sep. 2014, pp. 1-5, DOI: 10.1109/SSPD.2014.6943320.
- [19] H. Hassanieh, L. Shi, O. Abari, E. Hamed, and D. Katabi, “GHz-wide sensing and decoding using the sparse Fourier transform”, in *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, Apr. 2014, pp. 2256-2264, DOI: 10.1109/INFOCOM.2014.6848169.
- [20] A. López-Parrado and J. Velasco-Medina, “Cooperative Wideband Spectrum Sensing Based on Sub-Nyquist Sparse Fast Fourier Transform”, *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 63, n.º 1, pp. 39-43, 2016, DOI: 10.1109/TCSII.2015.2483278.
- [21] A. B. Korucu, O. Cakar, Y. K. Alp, G. Gok, and O. Arikan, “Compressive Digital Receiver: First hardware implementation results”, in *2018 26th Signal Processing and Communications Applications Conference (SIU)*, May 2018, pp. 1-4, DOI: 10.1109/SIU.2018.8404634.
- [22] C. Chang, J. Wawrzynek, and R. W. Brodersen, “BEE2: A high-end reconfigurable computing system”, *IEEE Des. Test Comput.*, vol. 22, n.º 2, pp. 114-125, Feb. 2005, DOI: 10.1109/MDT.2005.30.
- [23] a. Tkachenko, D. Cabric, and R. W. Brodersen, “Cognitive Radio Experiments using Reconfigurable BEE2”, in *2006 Fortieth Asilomar Conference on Signals, Systems and Computers*, 2006, pp. 2041-2045, DOI: 10.1109/ACSSC.2006.355125.
- [24] G. Eichinger, K. Chowdhury, and M. Leeser, “CRUSH: Cognitive radio universal software hardware”, *Proc. - 22nd Int. Conf. F. Program. Log. Appl. FPL 2012*, pp. 26-32, 2012, DOI: 10.1109/FPL.2012.6339237.

- [25] Ettus Research, “USRP Hardware Driver and USRP Manual: Table Of Contents” [internet]. Available: <https://files.ettus.com/manual/>.
- [26] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, “Xampling: Analog to Digital at Sub-Nyquist Rates”, *IET Circuits, Devices Syst.*, vol. 5, n.º 1, p. 8, 2009, DOI: 10.1049/iet-cds.2010.0147.
- [27] MP Antenna LTD, “Datasheet Super-M Ultra Base antenna.” Elyria, Ohio, p. 2, [Online]. Available: [www.mpantenna.com](http://www.mpantenna.com).
- [28] Myriad RF, “Myriad-RF Development Kit.” p. 94, 2013.
- [29] S. Kirolos *et al.*, “Analog-to-Information Conversion via Random Demodulation”, in *2006 IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*, Oct. 2006, pp. 71–74, DOI: 10.1109/DCAS.2006.321036.
- [30] Terasic Incorporated, “ADC-SoC User Manual.” 2017, [Online]. Available: [https://www.terasic.com.tw/cgi-bin/page/archive\\_download.pl?Language=English&No=1061&FID=2170285f16d7e05879607fccd2f31574](https://www.terasic.com.tw/cgi-bin/page/archive_download.pl?Language=English&No=1061&FID=2170285f16d7e05879607fccd2f31574).
- [31] Altera Corporation, “Cyclone V Hard Processor System Technical Reference Manual.” 2015, [Online]. Available: [https://www.altera.com/content/dam/altera-www/global/en\\_US/pdfs/literature/hb/cyclone-v/cv\\_5v4.pdf](https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/hb/cyclone-v/cv_5v4.pdf).
- [32] Intersil Corporation, “HFA3101 Gilbert Cell UHF Transistor Array Datasheet.” 2004, [Online]. Available: <https://www.intersil.com/content/dam/intersil/documents/hfa3/hfa3101.pdf>.
- [33] Analog Devices Incorporated, “ADA4807-1/ADA4807-2/ADA4807-4 3.1 nV/√Hz, 1 mA, 180 MHz, Rail-to-Rail Input/Output Amplifiers.” 2006, [Online]. Available: [http://www.analog.com/media/en/technical-documentation/data-sheets/ADA4807-1\\_4807-2\\_4807-4.pdf](http://www.analog.com/media/en/technical-documentation/data-sheets/ADA4807-1_4807-2_4807-4.pdf).
- [34] Altera Corporation, “Avalon Interface Specifications.” 2015, [Online]. Available: [https://www.altera.com/content/dam/altera-www/global/en\\_US/pdfs/literature/manual/mnl\\_avalon\\_spec.pdf](https://www.altera.com/content/dam/altera-www/global/en_US/pdfs/literature/manual/mnl_avalon_spec.pdf).
- [35] P. Alfke, “Application note: Efficient Shift Registers, LFSR Counters, and Long Pseudo-Random Sequence Generators.” 1996, [Online]. Available: [https://www.xilinx.com/support/documentation/application\\_notes/xapp052.pdf](https://www.xilinx.com/support/documentation/application_notes/xapp052.pdf).
- [36] Analog Devices Incorporated, “AD9254 14-Bit, 150 MSPS, 1.8 V Analog-to-Digital Converter.” 2015, [Online]. Available: <http://www.analog.com/media/en/technical-documentation/data-sheets/AD9254.pdf>.
- [37] M. Lexa, M. Davies, and J. Thompson, “Sampling Sparse Multitone Signals with a Random Demodulator”, 2010.
- [38] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples”, *Appl. Comput. Harmon. Anal.*, vol. 26, n.º 3, pp. 301–321, May 2009, DOI: 10.1016/j.acha.2008.07.002.

- [39] Stephen Cass, “The 2018 Top Programming Languages - IEEE Spectrum.” <https://spectrum.ieee.org/at-work/innovation/the-2018-top-programming-languages> (accessed Sep. 13, 2018).
- [40] A. Ronacher, “Welcome | Flask (A Python Microframework).” <http://flask.pocoo.org/> (accessed Sep. 13, 2018).
- [41] European Telecommunications Standards Institute, “EUROPEAN STANDARD Digital Video Broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2).” 2015.
- [42] European Telecommunications Standards Institute, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (3GPP TS 36.101 version 10.3.0 Release 10).” 2011.
- [43] Apache Software Foundation, “Jmeter”, 2018. <https://jmeter.apache.org/> (accessed Sep. 10, 2018).

# A Curriculum Design Approach by the Means of a General Morphological Analysis\*

Alexander Vera Tasama\*\*

Jorge Iván Marín Hurtado\*\*\*

Jaiber Evelio Cardona Aristizábal\*\*\*\*

Francisco Javier Ibargüen Ocampo\*\*\*\*\*

Recibido: 11/10/2018 • Aceptado: 14/03/2020

<https://doi.org/10.22395/rium.v19n37a3>

## Abstract

The context of conceiving, designing, implementing, and operating real-world systems and products, namely the CDIO initiative, is a framework for engineering education. It considers 12 standards, which are the reference for curriculum design and assessment. A good implementation of the CDIO standards can be considered as a multi-dimensional complex problem. In order to propose strategies for implementing the CDIO initiative in the Electronic Engineering curriculum at Universidad del Quindío, the General Morphological Analysis (GMA) was used. Some relevant dimensions of the curriculum and their values were contrasted in a cross-consistency assessment (CCA), where 8 dimensions were established, and a total of 34,560 combinations were obtained in the problem space. Through the CCA, the number of coherent combinations was significantly reduced. Finally, these combinations were analyzed to propose the corresponding strategies that are the input for the implementation of the CDIO curriculum in the Electronic Engineering program.

**Keywords:** curriculum design, engineering education, general morphological analysis (GMA), CDIO.

---

\* This paper was founded by the research project “General Morphological Analysis (GMA) applied to the CDIO-Curriculum of the Electronics Engineering Program at the Universidad del Quindío”; it was started in 2018, ID 913 – Universidad del Quindío.

\*\* Doctor en ingeniería. Profesor asistente, Grupo de Investigación en Procesamiento Digital de Señales y Procesadores (GDSPROC), Programa de Ingeniería Electrónica, Universidad del Quindío. Correo electrónico: [avera@uniquindio.edu.co](mailto:avera@uniquindio.edu.co). Orcid: <https://orcid.org/0000-0002-4183-9545>.

\*\*\* Doctor en ingeniería eléctrica y de la computación. Profesor asistente, director del Grupo de Investigación en Procesamiento Digital de Señales y Procesadores (GDSPROC), y director del Programa de Ingeniería Electrónica, Universidad del Quindío. Correo electrónico: [jorgemarin@uniquindio.edu.co](mailto:jorgemarin@uniquindio.edu.co). Orcid: <https://orcid.org/0000-0002-6334-7322>.

\*\*\*\* Doctor en Ingeniería, Profesor Asociado, Grupo de Investigación en Automatización y Máquinas de Aprendizaje (GAMA), Programa Ingeniería Electrónica, Universidad del Quindío. Correo electrónico: [jaibercardona@uniquindio.edu.co](mailto:jaibercardona@uniquindio.edu.co). Orcid: <https://orcid.org/0000-0002-1193-0612>.

\*\*\*\*\* Magister en ingeniería. Profesor asistente, Grupo de Investigación en Automatización y Máquinas de Aprendizaje (GAMA), Programa de Ingeniería Electrónica, Universidad del Quindío. Correo electrónico: [fjibarguen@uniquindio.edu.co](mailto:fjibarguen@uniquindio.edu.co). Orcid: <https://orcid.org/0000-0001-6025-8677>.

## Aproximación al diseño curricular mediante análisis morfológico general

### **Resumen**

La iniciativa CDIO —esto es, concebir, diseñar, implementar y operar sistemas y productos del mundo real— es un marco para la educación en ingeniería que considera 12 estándares como referencia para el diseño y evaluación del currículo. Respecto del diseño curricular, una buena implementación de los estándares CDIO puede considerarse un problema complejo multidimensional. Con el objetivo de proponer estrategias para implementar la iniciativa CDIO en el currículo de Ingeniería Electrónica de la Universidad del Quindío, se usó el análisis morfológico general (GMA). Así, algunas dimensiones relevantes del currículo y sus valores se contrastaron en una evaluación de consistencia cruzada (CCA) dentro de la cual se establecieron 8 dimensiones, para un total de 34.560 combinaciones en el espacio de problema. Con lo anterior, el número de combinaciones coherentes se redujo de forma significativa. Finalmente, esas combinaciones se analizaron para proponer las estrategias como insumo para la implementación del currículo CDIO en Ingeniería Electrónica.

**Palabras clave:** desarrollo curricular, educación en ingeniería, Análisis Morfológico General (GMA), CDIO.

## Aproximaçãõ ao desenho curricular mediante análise morfológica geral

### **Resumo**

A iniciativa CDIO — isto é, conceber, desenhar, implementar e operar sistemas e produtos do mundo real — é um marco para a educação em Engenharia que considera 12 padrões como referência para o desenho e a avaliação do currículo. Com relação ao desenho curricular, uma boa implementação dos padrões CDIO pode ser considerada um problema complexo multidimensional. Com o objetivo de propor estratégias para implementar a iniciativa CDIO no currículo de Engenharia Eletrônica da Universidad del Quindío, utilizou-se da análise morfológica geral (GMA). Assim, algumas dimensões relevantes do currículo e seus valores foram contrastadas numa avaliação de consistência cruzada (CCA) dentro da qual se estabeleceram oito dimensões, para um total de 34.560 combinações no espaço de problema. Com isso, o número de combinações coerentes reduziu-se de forma significativa. Finalmente, essas combinações foram analisadas para propor as estratégias como insumo para a implementação do currículo CDIO em Engenharia Eletrônica.

**Palavras-chave:** desenvolvimento curricular, educação em Engenharia, Análise Morfológica Geral (GMA), CDIO.



## INTRODUCTION

The implementation of the CDIO framework in an academic program faces several challenges. One of these is the curriculum design, since it involves a detailed analysis of the context in order to propose a design that satisfies the context requirements. To address this challenge, there is a recent interest to consider the curriculum design as a complex problem that can be aided by mathematical and computational tools [1-4]. In this paper, we use a tool called General Morphological Analysis (GMA) to deal with complex problems, aiming to perform a conceptual analysis about the implementation of a CDIO curriculum. This analysis is contextualized in the Electronic Engineering Program at Universidad del Quindío (Colombia).

A previous work was presented in the 13th International CDIO Conference at Calgary, Canada, titled “Conceptual Verification of CDIO Skills in the Electronic Engineering Curriculum at Universidad del Quindío” [3], using a different tool, the Formal Concept Analysis (FCA). In that work, we developed an attribute exploration using FCA for two different contexts: the design-implement projects, and the integration of the skill training across the courses in the curriculum.

In the CDIO framework, the curriculum is stated in the CDIO standard three (integrated curriculum) [5]. According to the CDIO guidelines, the curriculum design behaves as a complex problem as described below. The definition of a complex problem is given by Horst Rittel in [6]. Rittel and his collaborators established in the 1970s that a complex problem exists when different technical concepts and socio-political interests are involved. In the case of the curriculum design, this complexity appears due to the involved actors and interests:

- The students, and their expectations for training and future job opportunities.
- The specific disciplinary knowledge and skills in the training area, and the strategies for their effective teaching.
- The institution and program’s purposes, which are stated in the professional outcomes.
- The government requirements to meet quality standards.
- The society needs for professionals to effectively solve the problems that arise in the search for a better lifestyle.

Horst Rittel called this kind of problems as “wicked” ones, opposed to the so-called “tame problems”. Solving a wicked problem as if it were tamable constitutes an endless cycle of attempts to reach the solution. Some criteria to identify that the curriculum design is a wicked problem are the following [7]:

1. No definitive formulation of a complex problem. Since the curriculum is constantly evolving, we cannot establish something as “the definitive design”. Instead, we must use tools to carry out an exhaustive inventory of all imaginable solutions in advance.
2. Wicked problems do not have stop rules. This corresponds to the question: how do we know that we have designed the best possible curriculum? In other words, neither we can state that we have reached the final solution, nor that it is completely correct. A curriculum design ends when, subjectively, all stakeholders involved in it agree that its structure satisfies all initial requirements and needs stated at beginning.
3. Solutions to wicked problems are not true or false, but better or worse. The solution depends strongly on the stakeholders. Since the curriculum design involves different actors, each actor perceives different solutions as better or worse.
4. There is not an immediate and definitive test for validating the solution to a wicked problem. After implementing a curriculum, the consequences are observed in a virtually unlimited period of time, *i. e.*, a curriculum may have completely undesirable repercussions that may exceed the expected or desirable benefits envisioned at the beginning.
5. Each solution to a wicked problem is a one-shot operation. Since curriculum designers have no opportunity to learn by trial and error during the design process, all the attempts count in a significant way. The implementation of a new curriculum has consequences, such as leaving traces in students’ performance and their acquired skills. In addition, each attempt to reverse a decision, or to correct the unwanted consequences, raises another series of complex problems. These issues count from the point of view of the graduates.
6. Wicked problems do not have an enumerable (or feasible to describe) set of potential solutions, neither there is a well-described set of permissible operations that can be incorporated into the plan. In the case of a curriculum design, its implementation depends on the context and the disciplinary trends.
7. Each wicked problem is essentially unique, *i. e.*, it is not possible to establish categories for wicked problems. In the curriculum design, there is no solution principles that can be adapted to all contexts. It requires a process of analysis and synthesis to find out the solution type to be applied.
8. Each wicked problem can be considered as a symptom of another wicked problem. It is evident that the problem of curriculum design is immersed in the general problem of education, where a society demands competent professionals to satisfy specific needs.

9. The causes of a wicked problem can be explained in numerous ways. The choice of explanation determines the nature of problem resolution. For this reason, curriculum design has become a research in progress, where different regulations and procedures have been developed to generate the best design. However, it is also possible to propose several ways to refute the hypotheses on which the curriculum design is based on.
10. For wicked problems, the planner has no right to make mistakes. From the point of view of curriculum design, the planners are responsible for the consequences arising from the implementation of the curriculum. Whereas a hypothesis in the sciences can be refuted to find the truth, in the search for the solution of a wicked problem, the objective is not to find the truth, but to improve some environmental features. In the context of curriculum design, the environment is related to the quality of life for a given community, so that the social impact of a curriculum design is measured by how graduates contribute effectively for a better quality of life in their social context.

Among the available alternatives to address the solution of a wicked problem, we have chosen a systematic approach called General Morphological Analysis (GMA) [7]. This approach has the capability of generating a problem space, and possible solution spaces. Thus, through a process of analysis and synthesis of this solution space, it is possible to find alternative solutions, which are called morphotypes. To perform this analysis, it is necessary to use a software tool to support the entire GMA process [8]. To conduct the analysis in this paper, we designed a software tool using the LabView environment, and the information used to feed the GMA analysis (dimensions and values) is described in the next section.

## **1. STRATEGIES AND METHODS: CREATING A MORPHOLOGICAL FIELD FROM A CURRICULUM**

### **1.1 The CDIO Initiative**

The CDIO initiative is an association of more than 130 universities worldwide that have adopted a methodological model for engineering training. This model states that the newly graduated engineer must be able to conceive, design, implement, and operate products, systems, and processes in collaborative environments [9]. This training approach implies the development of disciplinary, personal, interpersonal, and professional skills that are expected for an engineer.

In 2004, the CDIO initiative adopted 12 standards to guide the principles of CDIO programs. These CDIO standards define the features that distinguish a CDIO program from traditional ones, serve as guidelines for curriculum design and evaluation of

educational programs, generate benchmarks, and provide a framework for continuous improvement.

The twelve CDIO standards address the philosophy of the program (standard 1), development of the curriculum (standards 2, 3 and 4), design-build experiences and working spaces (standards 5 and 6), teaching and learning methods (standards 7 and 8), teacher development (standards 9 and 10), and evaluation (standards 11 and 12) [10]. In this article, we make a special emphasis on the elements that describe the curriculum structure and pedagogical practices for a CDIO program. These elements are stated mainly in the following standards [10]:

- Standard 2. Learning outcomes. According to the CDIO principles, the engineer training is emphasized in learning experiences with verifiable results, which involve constant planning of goals and generation of deliverables. This training implies the development of skills in four categories: disciplinary, personal, interpersonal, and professional (*i. e.* skills to conceive, design, implement and operate systems). These skills must be explicit in the courses' syllabus as intended learning outcomes (ILOs). CDIO provides a concise description of the desired skills in the CDIO syllabus [11].
- Standard 3. Integrated Curriculum. This includes the strategies for improving the personal and interpersonal skills of the student, as well as the acquisition of product, process, and system building skills (professional skills). In this sense, the disciplinary courses must be explicitly interconnected to support the development of the learning outcomes and skills.
- Standard 5. Design-Implement Experiences. This standard refers to the training of engineers who are able find solutions to problems in context from the conceptual design (conceive) to the methodological design (implement).
- Standard 7. Integrated Learning Experiences. The disciplinary training of engineers, based on learning experiences, involves factors of diversity, topic connections, and development of professional skills.
- Standard 8. Active Learning. This standard refers to the usage of pedagogical methodologies that engage students in manipulating, analyzing, applying, and evaluating ideas, rather than passive transmission of information. Thus, active learning provides a way to acquire disciplinary knowledge simultaneously with the development of personal, interpersonal, and professional skills.
- Standard 9. Enhancement of Faculty Competence. The implementation of the CDIO framework requires the professional experience of the faculty to be prioritized not only in research, but also in the industrial sector.

- Standard 11. Learning Assessment. In the context of an engineering training based on knowledge experiences, active methodologies, and skills development, it is necessary to consider a complete assessment system suited to these requirements.

## 1.2 GMA Methodology

The GMA methodology was proposed and developed by Fritz Zwicky at the California Institute of Technology (Caltech), as a strategy to solve complex problems [8]. This methodology has been applied to different fields, such as [8] engineering and product design, general design theory and architecture, futures studies and scenario development, technology foresight, and technological forecasting, among others. In [8], the curriculum design is also considered as a potential application of GMA. Hence, this paper explores the usage of GMA for the implementation of a CDIO curriculum in an Electronic Engineering program.

The GMA methodology starts by defining the dimensions or parameters for the given problem. If dimensions are defined within the same context, we are dealing with a simplex model; by contrast, in the multiplex model the dimensions belong to different contexts or frameworks.

Next, different values or alternative solutions are proposed for each dimension. The values within a given dimension must be mutually excluding, so that choosing a particular value in a dimension is the only feasible solution for the selected dimension. Finally, a cross-consistency matrix is constructed for the problem space. In this matrix, a particular matrix element is referred to the compatibility of the combination of two values from different dimensions. If the combination of these two values is convenient, the matrix element is marked with “\_”. On the other hand, if the combination is impossible or inconvenient, the matrix element is marked with an “x”.

A feasible solution for the whole problem is such that choosing one value for each dimension produces a set of values that are mutually convenient according to the cross-consistency matrix. This solution is called “solution space” or “morphotype”, according to the GMA terminology. Besides, by fixing the value in a particular dimension, it is possible to analyze alternative solutions (or solution spaces) in the remaining dimensions based on the convenience of these combinations, according to the cross-consistency matrix.

The higher the number of dimensions, the higher the complexity of the analysis. Therefore, it is necessary to use a software tool to aid the analysis of these alternative solutions: This is one of the four necessary requirements proposed by Tom Ritchey for developing the GMA [7]. The software tool must allow us to find different solution

spaces across dimensions. If several dimensions are present, this process may be complicated, or even impossible, without a dedicated software tool. To develop a proper morphological analysis, a LabVIEW-based software [12] that uses the morphological field and the cross-consistency matrix [13] was elaborated. This allowed us to find and select different morphotypes, or configuration fields, which are solutions when only one value for each dimension is active.

### 1.3 GMA Applied to Curriculum Design

For the purposes of this work, we use the rubrics stated in the CDIO standards 2, 3, 7, 8, 9, and 11 [5] to propose the following eight dimensions in our GMA model:

- A. Type of curricular structure.
- B. Strategies to train professional skills (*i. e.* personal, professional, interpersonal, and CDIO skills) in the students.
- C. Strategies to integrate professional skills in the study plan (macro viewpoint).
- D. Development of professional skills in the faculty.
- E. Strategies to enhance faculty skills for integrated learning experiences, learning methods, and evaluation strategies.
- F. Strategies to evidence the alignment of program ILOs [14] with area-specific ILOs.
- G. Strategies for assessing professional skills development in students.
- H. Strategies to evidence how are professional skills integrated in the courses.

The organizing principle for the proposed analysis is to first consider the dimension A, since that one is important to start a curriculum design process [5]. The support for training, assessment, and integration of professional skills in the curriculum design is included in the remaining dimensions (B, C, D, E, F, G, H), attending the principle of constructive alignment [14].

Based on the previous principles, the rubrics for each CDIO standard, and the theoretical foundation given in [3, 5, 14], an expert group proposed different solution alternatives for each dimension. This group is comprised of the people in charge of the curriculum design: Professors with more than ten years of experience in education for engineering and advanced knowledge in the CDIO approach. This group obtained solution alternatives that correspond to the set of values used to build the problem space, and the cross-consistency matrix. Since the values for a given dimension must be mutually exclusive according to the GMA methodology, in some dimensions,

each proposed solution (or value) corresponds to the primary approach that should be addressed to fully satisfy the given dimension. For example, as shown below, for the dimension B (strategies to train professional skills in the students), some values (B2-teamwork and collaborative activities, B3-internships, and B4-classroom scenarios) could be complimentary solutions for this dimension, however, considering that only one value is enough to address the dimension. In other words, selecting only one value in any given dimension leads us to find a complete solution using the minimum amount of resources. The same kind of approach was conducted for all dimensions, were the expert group agreed to find values mutually exclusive or fully satisfying a given dimension. For the cross-consistency matrix design, the expert group analyzed the convenience and inconvenience for two values belonging to two different dimensions. Disagreements on the selection procedure by the expert group were solved by arguing on each viewpoint, and searching for complimentary information.

We propose the following values for each dimension:

- Dimension A
  - A1: Strictly disciplinary curriculum.
  - A2: Integrated curriculum.
  - A3: Problem-based curriculum.
  - A4: Project-based curriculum.
- Dimension B
  - B1: Integration of the high-order skills of the CDIO syllabus with the ILOs and the specific activities.
  - B2: Teamwork and collaborative activities that demand autonomous thinking for professional skills development.
  - B3: Implementation of internships in industrial environments that stimulate the development of professional skills in context.
  - B4: Classroom scenarios for the simulation of the issues commonly experienced in the industry or during the professional practice.
- Dimension C
  - C1: Constructive alignment between skills and the program's outcomes, which is reflected in the syllabus of the courses.

- C2: Formal concept analysis that determine the integration level of skills in the curriculum.
- C3: Curricular and co-curricular spaces for the development of design-implementation experiences throughout the study plan, encouraging the acquisition of professional skills supported by disciplinary knowledge.
- Dimension D
  - D1: Workshops oriented to the training of the faculty in professional skills (these can be led by an expert).
  - D2: Enhancement of professional skills in the faculty by means of business practices or internships.
  - D3: Academic spaces for the development of design-implementation experiences throughout the curriculum, that encourage the acquisition of professional skills supported by disciplinary knowledge.
- Dimension E
  - E1: Workshops focused on, firstly, training the faculty in integrated learning experiences such as case studies, role games, or active learning, and secondly, on training it in how to align ILOs, activities, and evaluation. These workshops can be led by an expert.
  - E2: Policies and administrative support for professors' participation in CDIO conferences, workshops, or meetings on active learning and assessment. This also calls for them to share their experiences with the whole faculty.
  - E3: Sharing classroom experiences by professors that implement strategies for integrated learning, active learning, and aligned assessment, contrasting them with the skills stated in the syllabus.
- Dimension F
  - F1: Development of mind maps that relate each course's learning outcomes with program and area-specific ones.
  - F2: Verification through a dynamic matrix that demonstrates how each skill is introduced, taught, and applied throughout the study plan.
  - F3: Report on how the program's learning outcomes are reflected in the activities developed for each course. This report is shared with the faculty.
  - F4: Design-implementation projects associated with a particular course. The student performance is recorded during this project, according to the topics for a given depth level.



- Dimension G
  - G1: Establishing a profile with descriptors of student behavior regarding the topic in development. Associating descriptors with the expected performance levels, and developing an observation process based on rubrics applied in the classroom.
  - G2: Conducting written evaluations in which students are asked to self-assess their performance level, associated with a specific activity.
  - G3: Conducting surveys to inquire about the perception of the required level in the development of the topic.
  - G4: Formulating course projects strictly associated with the disciplinary knowledge, and observing performance results during the project.
  
- Dimension H
  - H1: Syllabus of each course with the description of the activities aimed at developing the proposed skills.
  - H2: Multimedia records of the classroom activities that develop the chosen skills.
  - H3: Log for the classroom activities addressing the topics related to the professional skills.
  - H4: Report on classroom experiences (of pedagogical and didactic types) for discussion with the faculty.
  - H5: Presentation to stakeholders of the products made from the design-implementation experiences in the classroom.

Since our previous analysis shows that all dimensions belong to the same context, the proposed model for the curriculum analysis is a *simplex model* [15]. It is an 8-dimensional space (Table 1) which contains  $4 \times 4 \times 3 \times 3 \times 3 \times 3 \times 4 \times 4 \times 5 = 34560$  distinct formal configurations or morphotypes. To ease the analysis of this huge amount of morphotypes, only the morphotypes that are coherent with the cross-consistency matrix were considered, as described in the next section.

Table 1. Proposed 8-Dimensional Morphological Field

A	B	C	D	E	F	G	H
A1	B1	C1	D1	E1	F1	G1	H1
A2	B2	C2	D2	E2	F2	G2	H2
A3	B3	C3	D3	E3	F3	G3	H3
A4	B4				F4	G4	H4
							H5

Source: Prepared by the authors.

## 2. RESULTS AND DISCUSSION

A Cross-Consistency Assessment (CCA) [16] was performed on a cross-consistency matrix (see Table 2) by checking the connective relationship and constraints between the dimensions. This assessment process was carried out by a team with experience on designing activities to develop skills in students and the faculty.

Table 2. Cross-Consistency Matrix for the Morphological Field in Table 1. X: It is not Possible or Inappropriate. \_: It is Possible and Fully Appropriate/Optimal

1	A			B				C			D			E			F				G				
	2	3	4	1	2	3	4	1	2	3	1	2	3	1	2	3	1	2	3	4	1	2	3	4	
B	1	X	_	_	_																				
	2	X	_	_	_																				
	3	X	_	X	_																				
	4	X	_	_	_																				
C	1	X	_	_	X	_	_	_	_																
	2	_	_	_	_	_	X	X	X																
	3	X	_	_	_	_	_	_	_																
D	1	X	_	_	_	_	_	X	_	_	X	X													
	2	X	_	_	_	_	_	X	_	_	X	X													
	3	X	_	_	_	_	_	X	_	_	X	X													
E	1	X	_	_	_	_	_	X	_	_	X	X	_	X	_										
	2	X	_	_	_	_	_	X	X	_	X	X	_	_	_										
	3	X	_	_	_	_	_	X	_	_	_	X	_	X	X										
F	1	X	_	_	X	_	X	X	X	_	_	_	X	_	_	_	X	X							
	2	X	_	_	_	_	X	_	X	_	_	X	_	X	X	X	X	X	_						
	3	X	_	_	_	_	_	X	_	_	X	X	_	X	X	_	_	_							
	4	_	_	_	_	_	_	_	_	_	X	_	X	X	X	_	X	_							
G	1	X	_	_	_	_	X	_	_	_	_	_	_	_	_	X	X	X	X	_					
	2	_	_	_	_	X	_	X	_	X	X	_	X	X	X	_	X	_	X	X	_	_			
	3	_	_	_	_	_	_	_	_	X	X	X	X	X	X	_	X	_	X	X	_	_			
	4	_	_	_	_	_	X	_	_	X	_	X	X	X	_	X	_	X	X	_	X				
H	1	X	_	_	_	_	X	_	_	_	X	_	X	_	_	X	_	_	_	X	_	_	_	_	X
	2	_	_	_	_	_	X	_	X	X	X	_	X	_	_	_	_	_	X	X	_	_	_	_	X
	3	X	_	_	_	_	X	_	X	X	_	_	X	_	_	_	_	X	X	_	_	_	_	_	X
	4	_	_	_	_	_	X	X	_	_	X	_	X	_	_	_	_	X	X	_	_	_	_	_	_
	5	_	_	_	_	X	_	X	_	_	X	_	X	_	_	_	_	X	X	_	_	_	_	_	_

Source: Prepared by the authors.

The cross-consistency matrix was analyzed with the simulation software, finding that there is not a complete morphotype, *i. e.* we cannot find a solution that considers

at least one value for each dimension. In fact, three (3) dimensions, D, F, and G, are mutually incompatible. Hence, complete morphotypes can be found in the remaining seven (7) dimensions by eliminating one of these three dimensions.

The incompatibility of these three dimensions must be considered carefully, either by reviewing the values of each dimension or by reviewing the values of the cross-consistency matrix. Considering that these three dimensions show incompatibility problems among them, we decided to find the morphotypes by using only the remaining dimensions when one of them is eliminated.

When the dimension D is eliminated, the morphological field can be reduced up to 22 morphotypes, which is a reduction of 99.80 %. If the parameter F is taken out, the solution space can be reduced to 28 morphotypes (99.75 %); and when the parameter G is eliminated, the solution can be reduced to 32 formal configurations (99.72 %). According to these results, an important reduction of the problem space is achieved. However, the values of the dimensions D, F, and G should be reviewed to avoid possible overlaps between the involved strategies. The cross-consistency matrix could be also reviewed. This review is proposed as a further work.

It was found that some values of the dimensions appear in three morphotype groups, or only two in the case of the deleted dimensions. This condition shows that these values have relevance in the curriculum’s configuration. Table 3 presents the relevant values that were found for each dimension.

Table 3 shows some dimensions with only one relevant value. This result suggests that it could be the best option for the curriculum implementation. The dimensions and the associated values are C1, E1, F3, G1, and H4. In the other dimensions, two values are relevant: A2, A3, B1, B4, D1, and D3.

Table 3. Dimensions and Relevant Values

Dimension	Relevant values
A	A2, A3
B	B1, B4
C	C1
D	D1, D3
E	E1
F	F3
G	G1
H	H4

Source: Prepared by the authors.

The election of only relevant values for each dimension may help to reduce the number of options for curriculum implementation, and the searching for better solutions. In addition, the values that were not considered relevant according to this methodology imply that they do not constitute an effective opportunity for the design of a CDIO curriculum using the lowest amount of resources and strategies.

Taking into account the dimensions with incompatibilities, D, F, and G, only the first one has two values, but only D1 is compatible with the F3 and G1 values. Therefore, D3 can be discarded, and the solution space can be reduced. In the case of F3 and G1, they are incompatible according to the cross-consistency matrix. Therefore, other values must be considered to reformulate these dimensions.

According to this analysis, the kind of curricular structure, dimension A, has only two relevant values: Integrated curriculum, and problem-based curriculum. Among these values, the “strictly-disciplinary curriculum” option can be discarded because, according to the software, dimensions B, D, and E appear without values when A1 is selected. Besides, the other value of the dimension A, project-based curriculum, do not appear at any morphotype, even when dimensions D, F, or G are eliminated. This result suggests that this curriculum type is not a good option, even though it seemed promising when A4 was selected in the software. It should be noted that an integrated and a problem-based curriculum allow the integration of professional skill development.

### 3. CONCLUSIONS

In this paper, we addressed the curriculum design as a wicked problem, and analyzed different dimensions of this problem by using the general morphological analysis framework. As a work-in-progress, the results of this analysis are useful to find, and then select, the appropriate strategies to design and implement a CDIO curriculum in our engineering program. Our confidence in the results is supported by the relevance of the GMA method for design theories, strategic management, and scenario development. So, we evaluated different curriculum approaches to determine which one is suitable for the implementation of a CDIO program. From this evaluation, we concluded that an integrated (value A2) and a problem-based curricula (value A3) are the most convenient ones. Furthermore, from the morphotypes viewpoint, neither of the solution alternatives satisfies the full coverage of the whole problem dimensions.

However, three of these dimensions were identified as critical in the problem solution, since taking out at least one of them allows to obtain a problem solution covering the full scope of dimensions. These are D (development of professional skills in the faculty), F (strategies to evidence the alignment of program ILOs with specific-areas ILOs), and G (strategies to assess the development of professional skills in the

students). Finally, the developed software tool allows us to reduce the problem space, and to seek out proper solutions for curriculum design problem. As future research, we propose to adjust the values for each dimension: These can be matched to the needs of a particular program.

## REFERENCES

- [1] K. E. Willcox and L. Huang, "Mapping the CDIO curriculum with network models," in *Proceedings of the 13th International CDIO Conference*, 2017.
- [2] A. Álvarez and T. Ritchey, "Applications of General Morphological Analysis From Engineering Design to Policy Analysis," *Acta Morphol. Gen. AMG*, vol. 4, no. 1, 2015.
- [3] J. I. Marín et al., "Conceptual Verification of CDIO Skills in the Electronic Engineering Curriculum at Quindío University," in *Proceedings of the 13th International CDIO Conference*, 2017.
- [4] Australian Government - Office for Learning & Teaching, "Curriculum Design Workbench (Tool) | Assuring Learning." [Online]. Available: <http://www.assuringlearning.com/curriculum-design-workbench-tool>. [Accessed: 17-Sep-2018].
- [5] E. F. Crawley et al., *Rethinking Engineering Education*, 2014. DOI: <https://doi.org/10.1109/FIE.2017.8190506>
- [6] H. W. J. Rittel and M. M. Webber, "Dilemmas in a General Theory of Planning," *Policy Sci.*, no. 4, pp. 155-169, 1973. DOI: <https://doi.org/10.1007/BF01405730>
- [7] T. Ritchey, *Wicked problems—social messes: Decision support modelling with morphological analysis*, Berlin: Springer, 2011.
- [8] A. Álvarez and T. Ritchey, "Applications of General Morphological Analysis," *Acta Morphol. Gen.*, vol. 4, no. 1, pp. 1–40, 2015.
- [9] "CDIO Vision | Worldwide CDIO Initiative." [Online]. Available: <http://cdio.org/cdio-vision>. [Accessed: 13-Sep-2018].
- [10] "CDIO Standard 2.1 | Worldwide CDIO Initiative." [Online]. Available: <http://www.cdio.org/content/cdio-standard-21>. [Accessed: 15-Sep-2018].
- [11] "CDIO Syllabus 2.0 | Worldwide CDIO Initiative." [Online]. Available: <http://cdio.org/benefits-cdio/cdio-syllabus/cdio-syllabus-topical-form>. [Accessed: 13-Sep-2018].
- [12] National Instruments, "¿Qué es LabVIEW? - National Instruments," 2018. [Online]. Available: <http://www.ni.com/es-co/shop/labview.html>. [Accessed: 15-Sep-2018].
- [13] I. Johansen, "Scenario modelling with morphological analysis," *Technol. Forecast. Soc. Change*, vol. 126, pp. 116–125, Jan. 2018. DOI: <https://doi.org/10.1016/j.techfore.2017.05.016>

- [14] J. Biggs and C. Tang, *Teaching for Quality Learning at University*. 2007. DOI: <https://doi.org/10.1097/00005176-200304000-00028>
- [15] T. Ritchey, “General morphological analysis as a basic scientific modelling method,” *Technol. Forecast. Soc. Change*, vol. 126, no. May 2017, pp. 81–91, 2018. DOI: <https://doi.org/10.1016/j.techfore.2017.05.027>
- [16] T. Ritchey, “Principles of Cross-Consistency Assessment in General Morphological Modelling,” *Acta Morphol. Gen.*, vol. 4, no. 2, pp. 1–20, 2015.

# Evaluation of Clusters based on Systems on a Chip for High-Performance Computing: A Review\*

Melissa Johanna Aldana\*\*

Jaime Alberto Buitrago\*\*\*

Julián Esteban Gutiérrez\*\*\*\*

Recibido: 11/10/2018 • Aceptado: 14/03/2020

<https://doi.org/10.22395/rium.v19n37a4>

## Abstract

High-performance computing systems are the maximum expression in the field of processing for large amounts of data. However, their energy consumption is an aspect of great importance, which was not considered decades ago. Hence, software developers and hardware providers are obligated to approach new challenges to address energy consumption, and costs. Constructing a computational cluster with a large amount of systems on a chip can result in a powerful, ecologic platform, with the capacity to offer sufficient performance for different applications, as long as low costs and minimum energy consumption can be maintained. As a result, energy efficient hardware has an opportunity to impact upon the area of high-performance computing. This article presents a systematic review of the evaluations conducted on clusters of Systems on a Chip for High-Performance computing in the research setting.

**Keywords:** Systems on a chip; high-performance computing; clusters; benchmarks.

---

\* This review paper stems from a research project (no. 811) called “Cluster of mobile systems on a Chip for high-performance computing”, funded by the Universidad del Quindío through its Vice-rectory of Research. Start date: 2017. End date: 2018.

\*\* Magister en proyectos educativos mediados por TIC, e ingeniera de sistemas y computación. Profesora asistente del Programa de Ingeniería de Sistemas y Computación, Grupo de Investigación GRID, Universidad del Quindío. Correo electrónico: mjaldana@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-0257-4783>.

\*\*\* Ph. D. en Ingeniería e ingeniero electrónico. Profesor asistente del Programa de Ingeniería Electrónica, Grupo de Investigación Sinfoci, Universidad del Quindío. Correo electrónico: jalbertob@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-8328-8470>.

\*\*\*\* Ph. D en Ciencias de la Computación, Ingeniero de Sistemas, Profesor Titular Programa de Ingeniería de Sistemas y Computación, Grupo de Investigación GRID Universidad del Quindío. Correo electrónico: jugutier@uniquindio.edu.co. Orcid: <https://orcid.org/0000-0002-1825-0097>.

## Evaluación de clústeres basados en sistemas en un chip para computación de alto desempeño: una revisión

### **Resumen**

Los sistemas de computación de alto desempeño son la máxima expresión en el campo de procesamiento para grandes cantidades de datos. Sin embargo, su consumo de energía es un aspecto de gran importancia que no era tenido en cuenta en décadas pasadas. Por lo tanto, desarrolladores de *software* y proveedores de *hardware* están obligados a enfocarse en nuevos retos para abordar el consumo de energía y costos. Construir un clúster informático con una gran cantidad de sistemas en un chip puede dar como resultado una plataforma poderosa, ecológica y capaz de ofrecer el rendimiento suficiente para diferentes aplicaciones, siempre y cuando se puedan mantener bajos costos y el menor consumo de energía posible. Como resultado, el *hardware* eficiente en el consumo de energía tiene la oportunidad de tener un impacto en el área de la computación de alto desempeño. En este artículo se presenta una revisión sistemática para conocer las evaluaciones realizadas a clústeres de sistemas en un chip para computación de alto desempeño en el ámbito investigativo.

**Palabras clave:** sistemas en un chip; computación de alto desempeño; clústeres; *benchmarks*.

## Avaliação de clusters baseados em sistemas em um chip para a computação de alto desempenho: uma revisão

### **Resumo**

Os sistemas de computação de alto desempenho são a máxima expressão no campo de processamento para grandes quantidades de dados. No entanto, seu consumo de energia é um aspecto de grande importância que não era levado em consideração em décadas passadas. Portanto, desenvolvedores de *software* e provedores de *hardware* estão obrigados a focar-se em novos desafios para abordar o consumo de energia e custos. Construir um cluster informático com uma grande quantidade de sistemas em um chip pode dar como resultado uma plataforma poderosa, ecológica e capaz de oferecer o rendimento suficiente para diferentes aplicações, desde que possam ser mantidos baixos custos e o menor consumo de energia possível. Como resultado, o *hardware* eficiente no consumo de energia tem a oportunidade de ter um impacto na área da computação de alto desempenho. Neste artigo, apresenta-se uma revisão sistemática para conhecer as avaliações realizadas a clusters de sistemas em um chip para computação de alto desempenho no âmbito investigativo.

**Palavras-chave:** sistemas em um chip; computação de alto desempenho; clusters; *benchmarks*.



## INTRODUCTION

High-performance computing (HPC) is the concept that encompasses the principles, methods, and techniques that allow to address problems with complex computer structures, and of high requirements. The solution to those problems involves massive data sets, a large amount of variables, and complex calculation processes, which require efficient application of modern parallel computation tools [1]. In addition, the scientific challenges arising in engineering, geophysics, bioinformatics, and other types of applications of intensive computational use, require ever-higher amounts of computational calculations [2].

High-performance computing systems have shown exponential growth in computational power in recent decades, and this is due mainly to the evolution of microprocessor technology [3]. Large HPC systems are dominated by processors using mainly x86 and Power instruction sets, supplied by three providers: Intel, AMD, and IBM [3]. Performance of HPC systems has increased continuously with advances of Moore's Law and parallel processing, while energy efficiency could be considered a secondary problem [4]. By 2000, hardware manufacturers improved their processors' performance by adding optimization technology, like bifurcation prediction, predictive execution, and increased cache size, besides making the clock frequency faster. However, the downside of this situation was increased energy consumption, which obligated manufacturers to add multiple nuclei in a processor to avoid having problems due to overheating [3]. According to the aforementioned, it was clear that energy consumption was the dominant parameter in the scaling challenges to achieve better performance. It is widely accepted that future HPC systems will be limited by their energy consumption [2], and thermal problems [5]. Improvement is required in energy efficiency to achieve exascale computing (1018 FLOPS) [4-5].

Using concepts of integrated technologies, like systems on a chip (SoC, System-on-Chip), have emerged naturally to address the problem of energy consumption. The first prototypes were studied based on a vast number of microprocessors of many low-power nuclei instead of rapid complex nuclei, to comply with HPC and power consumption demands. In addition, recent progress and the availability of 64-bit Advanced RISC Machine (ARM) nuclei open new expectations for cluster development [4].

One way to address this problem of energy consumption is to replace the central processing units (CPU) of high-end servers with the low-power processors traditionally found in embedded systems. The use of integrated processors in clusters is not new: Diverse BlueGene machines use integrated PowerPC chips [7]. Likewise, low-power processors, which were originally destined for mobile devices, have had enormous improvement with respect to their computational power. Low-power modern processors,

like the most recent ARM designs, provide great performance in relation to their low energy consumption. Today, most mobile devices are themselves small supercomputers, which supply computing power due to multiple processors and sophisticated graphic processors, while maintaining the low energy consumption that is necessary for mobile applications [5].

Implementation of HPC platforms is a costly undertaking, which may be inaccessible for small and medium institutions. Projects like Mont-Blanc [<http://www.montblanc-project.eu/>] and COSA (Computing on SoC Architectures, <http://www.cosa-project.it/>) [8] have successfully demonstrated the use of SoC-based clusters for HPC systems, but the need still exists to analyze their viability. The principal deficiencies of the current evaluations of these systems are the lack of detailed information on the performance levels in distributed systems, and the comparative evaluation in large-scale applications [9].

This work analyzed publications in the area of clusters of Systems on a Chip for high-performance computing. Inspired on the PRISMA Declaration, a search and selection of investigations was conducted. Similarly, a systematic review was carried out to learn of the evaluations made of these types of systems in the research setting. This evaluation does not compare the systems to each other, since each application has its own characteristics and different configurations, which does not make it appropriate to know which is better or worse at the performance level. Likewise, it is highlighted that the objective of this article is to know the type of evaluations carried out on Systems on a Chip for high-performance computing.

## 1. METHODOLOGY

The search and systematic review used an adaptation of the PRISMA Declaration, which sought to help the authors to improve the presentation of the systematic reviews [10]. The PRISMA Declaration consists of a 27-item checklist and a four-phase flow diagram, which are the methodological route to conduct the search and systematic review. Additionally, PRISMA can be useful for the critical evaluation of published systematic reviews [11]. The following sections will present the results of the search process and systematic review. Based on the PRISMA Declaration, we first defined the theme or problem to analyze and then formulated three research questions, which are the base of the search criteria and selection of the bibliography review.

## 2. RESULTS

This work sought to identify what types of performance evaluations have been conducted on clusters based on systems on a chip for high-performance computing. Hence, it

is important to know what hardware, benchmarks, and measurement parameters have been used in implementing these types of systems.

Specifically, three questions focused on the recognition of the performance evaluations mentioned were formulated, to allow concentration in the search of the bibliography review.

- Research questions:
  - What types of devices were used in the investigations consulted?
  - What tests or benchmarks were implemented to measure cluster performance?
  - What other parameters were considered to conduct the evaluations?

## 2.1 Search sources

According to the research questions, a set of search sources was defined, which include IEEE Xplore Digital Library (Institute of Electrical and Electronics Engineers), Science Direct, Engineering Village, and Google Scholar, which contain published scientific bibliography according to the theme of interest.

## 2.2 Search criteria

In line with the search sources, the study created the permitted criteria to carry out the bibliography review according to terminology employed by experts in high-performance computing. General criteria, like date range, expressions of interest, languages selected, and expressions not permitted are listed ahead:

- Study period: 2010 onwards.
- Languages: Spanish and English.
- Expressions:
  - Measurement in system on chip and cluster
  - Measurement in cluster system on chip high performance computing
  - Benchmark in system on chip and cluster
- Expressions not permitted (the word SoC captures too much information that does not belong to the study theme):
  - Assessment System on Chip

- Metrics System on Chip
- Measurement in SoC
- Benchmark in SoC

With these terms defined, a search chain has been specified that includes adequate terminology according to the functioning of the databases (figure 1).

```
{
    {
        "System on Chip" AND "messurment" AND "cluster"
    }
    OR
    {
        "System on Chip" AND "benchmark" AND "cluster"
    }
    OR
    {
        "System on Chip" AND "cluster" AND "HPC"
    }
    OR
    {
        "System on Chip" AND "cluster" AND "High
        Performance Computing"
    }
}
```

Figure 1. Chain of Selected Search

Source: Prepared by the authors.

Additionally, the study defined the exclusion criteria of the works related:

- Publications prior to 2010.
- Duplicate publications.
- Studies in languages other than those expressed.
- Publications without metrics, or evaluations of clusters of mobile systems on a chip.

### 2.3 Systematic Review Results

Specifically, the information of interest, upon obtaining the publications selected, was based on evaluation strategies and tools, types of tests, and results of evaluations for clusters based on systems on a chip for high-performance computing. Table 1 shows the amount of works found, excluded, and selected from each source consulted to be analyzed in this review.

Table 1. Amount of Publications per Each Research Source

Source	Amount of Works	Excluded	Final amount of works
IEEE	36	25	11
Science Direct	17	14	3
Engineering Village	61	59	2
Google Scholar	30	27	3
Total	144	125	19

Source: Prepared by the authors.

According to the PRISMA Declaration, 144 studies were identified in the “*Identification*” phase, of which 40 publications duplicated in the different databases were manually excluded, thus, leaving 104 studies, which make up the “*Screening*” phase of the Prisma model. The “*Eligibility*” phase excluded 75 more publications and studies that did not contain the requirements according to the inclusion and exclusion criteria initially expressed. This analysis was first conducted by the title, followed by the abstract, and the contents of the evaluation methodologies. Finally, these 29 documents were analyzed in depth, finding that 10 of the publications contained evaluations of devices with emphasis on computer networks, which is not part of the respective analysis, thereby, considering the 19 studies that finally complied with all the criteria and the emphasis required, and which were selected to be part of this systematic review. The phases of the methodology, as well as the number of publications worked in each of them, can be observed in figure 2.

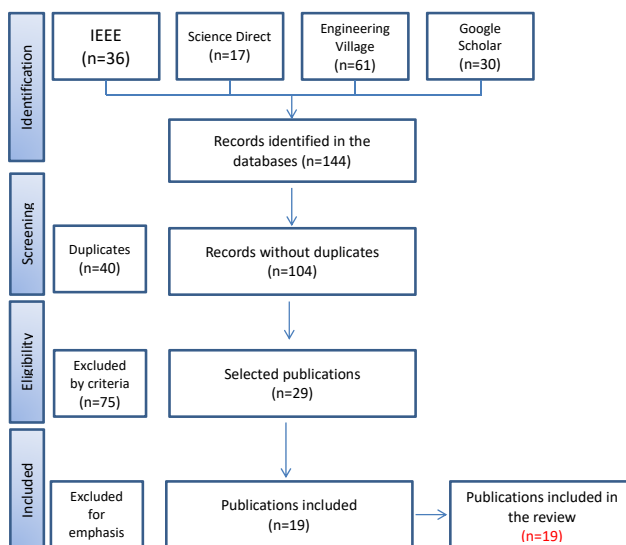


Figure 2. Flow Diagram, PRISMA Declaration

Source: Prepared by the authors.

## 2.4 Statistical information

The following presents the statistical data on the studies selected, from a general vision of the information analyzed in the evaluation of clusters based on systems on a chip for high-performance computing. Table 2 summarizes the data collected.

Table 2. Publications Selected for Systematic Review

Publication	Hardware	Benchmarks
<i>A study of big data processing constraints on a low-power Hadoop cluster [12]</i>	Cubieboard	BigDataBench. micro-benchmarks: • wordcount • sort • grep
<i>Evaluating Performance and Energy on ARM-based Clusters for High Performance Computing [13]</i>	BeagleBoard PandaBoard	High-Performance Linpack (HPL)
<i>Evaluation of Mobile ARM-Based SoCs for High Performance Computing [14]</i>	NVIDIA Jetson TK1 Odroid XU3-Lite The Parallella Board Wandboard Quad	OpenBLAS STREAM benchmark for measuring memory bandwidth EPCC OpenMP microbenchmark suite.
<i>Performance measurements and comparison between Cruz Cluster I and Cruz Cluster II [15]</i>	Raspberry Pi 2 Raspberry Pi B+	High Performance Linpack (HPL)
<i>Optimizing performance and power consumption for an ARM-based big data cluster [16]</i>	Cubieboard	JVM, CRC32, data compression
<i>Performance and Energy evaluation of Spark applications on low-power SoCs [17]</i>	Intel E5-2650 Intel i5-430M Raspberry Pi 3 Snapdragon410	Spark applications benchmark (micro-benchmarks)
<i>A Low-Cost Energy-Efficient Raspberry Pi Cluster for Data Mining Algorithms [6]</i>	Raspberry Pi 2 Intel Xeon Phi	Data mining algorithms: Association Rule Learning (A priori) and K-Means.
<i>A Raspberry Pi Cluster Instrumented for Fine-Grained Power Measurement [7]</i>	Raspberry Pi Model 2B	Cluster benchmarking: Linpack STREAM
<i>Efficiency Modeling and Analysis of 64-bit ARM Clusters for HPC [4]</i>	Juno ARM AMD Seattle Micro X-Gene	SGEMM, DGEMM [10] and HPL [11].
<i>Energy Efficiency of a Low Power Hardware Cluster for High Performance Computing [5]</i>	Odroid-C2 Odroid-XU4 Raspberry Pi 3 ASRock J3160TM-ITX	Himeno Benchmark. NASA NAS Parallel Benchmarks. MPI-based k-means Algorithm. Distributed Video Encoding.
<i>Evaluating ARM HPC Clusters for Scientific Workloads [9]</i>	ODROID-X	Benchmarks (STREAM, Sysbench, and PARSEC) and benchmarks (High Performance Linpack (HPL), NASA Advanced Supercomputing (NAS) Parallel Benchmark (NPB), and Gadget-2)

Publication	Hardware	Benchmarks
<i>Evaluating Systems on Chip through HPC bioinformatics and astrophysics applications [8]</i>	NVIDIA Jetson K1 ODROID-XU3 ODROID-XU-E ARNDALÉ OCTA Freescale IMX.6	High Performance Linpack (HPL) benchmark  GROMACS, NAMD and GADGET2
<i>Study of Raspberry Pi 2 Quad-core Cortex-A7 CPU Cluster as a Mini Supercomputer [18]</i>	Raspberry Pi 2 model B	High Performance Linpack Benchmark
<i>Supercomputing with Commodity CPUs Are Mobile SoCs Ready for HPC [19]</i>	NVIDIA Tegra 2 NVIDIA Tegra 3 Samsung Exynos 5250	11 Micro-kernels benchmarks: Vector operation Dense matrix-matrix multiplication 3D volume stencil computation 2D convolution One-dimensional Fast Fourier Transform Reduction operation Histogram calculation Generic merge sort N-body calculation Markov Chain Monte Carlo method xi. Sparse Vector-Matrix Multiplication
<i>Tibidabo - Making the case for an ARM-based HPC system [2]</i>	NVIDIA Tegra 2	SPEC CPU2006 benchmarks STREAM benchmark High-Performance Linpack (HPL)
<i>Tiny GPU Cluster for Big Spatial Data A Preliminary Performance Evaluation [20]especially for geo-referenced spatial (or geospatial)</i>	NVIDIA Tegra K1	Micro-Kernels benchmarks: taxi-nycb g10m-wwf
<i>Towards an energy efficient SoC computing cluster [21]</i>	Cubieboard 2	NAS Parallel Benchmarks
<i>Building a low consumption cluster using SBC technology [22]</i>	Raspberry PI B+	NAS Benchmarks
<i>Building and benchmarking a low power ARM cluster [3]</i>	Pandaboard ES Raspberry Pi	CoreMark STREAM benchmark Linpack High Performance Linpack (HPL) ping pong benchmark NAS parallel benchmark

Source: Prepared by the authors.

According to the information illustrated in figure 3, the date range of the publications selected is between 2012 and 2017; bear in mind that the study consultation began in 2010.

2012	2013	2014	2015	2016	2017
10.5 %	5.3 %	26.3%	10.5 %	36.8 %	10.5 %

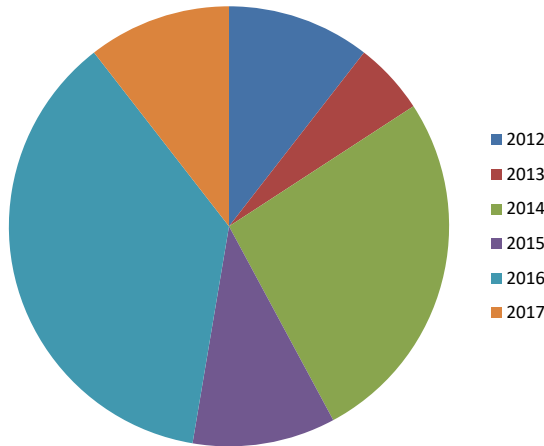


Figure 3. Percentage of Studies per Year of Publication

Source: Prepared by the authors.

The origin of the publications included in the review can be seen in figure 4. These publications are concentrated in three continents, America, Asia, and Europe, with Europe having the highest percentage of publications on the theme selected with 52 % of all the studies. It should be noted that Spain was the European country with the highest number of publications. Likewise, in the other continents, The United States, Brazil, and Thailand had the highest participation (figure 5).



Figure 4. Publications per Country

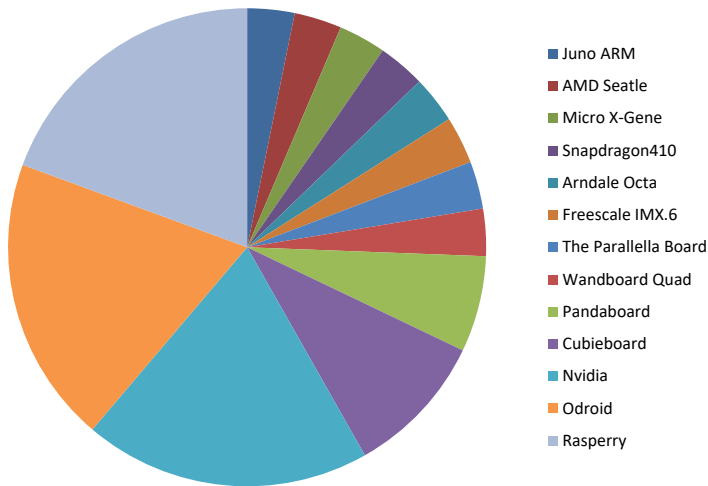
Source: Prepared by the authors.





according to the selection made in the bibliography review. These technologies represent 58.2 % of all the technologies used, according to figure 6.

Each of these systems represented 19.4 %. The remaining percentage focused on systems like Juno ARM, AMD SeattleMicro X-Gen, Snapdragon 410, Arndale Octa, Freescale IMX.6, The Parallella Board, Wandboard Quad, Pandaboard, and Cubieboard.



Juno ARM	AMD Seattle	Micro X-Gen	Snapdragon 410	Arndale Octa	Freescale IMX.6	The Parallella Board
3.2 %	3.2 %	3.2 %	3.2 %	3.2 %	3.2 %	3.2 %
Wandboard Quad	Pandaboard	Cubieboard	Nvidia	Odroid	Raspberry	
3.2 %	6.5 %	9.7 %	19.4 %	19.4 %	19.4 %	

Figure 6. Technology Used in the Publications of Clusters of Systems on a Chip

Source: Prepared by the authors.

The second question referred to What tests or benchmarks were implemented to measure cluster performance. As a result of the review, approximately 25 different performance tests were found (figure 7), highlighting the High Performance Linpack (HPL), Stream, and NAS Parallel Benchmarks tests.

Finally, the third question asked What other parameters were kept in mind to conduct the evaluations. To answer this question, it was established that the benchmarks currently used are standardized, and the measurement parameters are already identified for the comparisons made to be measured under the same conditions. Hence, no additional parameters were evidenced from those specified in the performance tests.



on the resolution of dense linear systems (generated randomly) from double-precision equations (64 bits), over a system with distributed memory [25]. The HPL provides programs to evaluate the numerical precision of the results and the time of resolution. The reported performance value depends on diverse factors, but under certain efficiency assumptions of the intercommunication network; the HPL resolution algorithm may be considered scalable, and its efficiency is maintained constant with respect to the memory used by each processor. Execution of HPL requires an implementation of MPI distributed memory and the BLAS library. The unit to present the HPL results is Flops [15].

STREAM is a benchmark that permits evaluating the efficiency of access to the principal memory through four operations over vectors and it is the standard industry benchmark to measure a system's memory bandwidth and the computation rhythm of simple vectors. The program runs four copy, scale, add, and triad operations of floating-point values over large matrices to calculate the memory bandwidth obtained in each operation in MB/s. Table 3 shows the number of flops and bytes counted every loop iteration [3].

Table 3. Operations of the STREAM benchmark

Operation	Instructions	Per iteration	
		bytes	Flops
<i>COPY</i>	$a(i) = b(i)$	16	0
<i>SCALE</i>	$a(i) = q * b(i)$	16	1
<i>SUM</i>	$a(i) = b(i) + c(i)$	24	1
<i>TRIAD</i>	$a(i) = b(i) + q * c(i)$	24	3

Source: [3].

The “NAS Parallel Benchmarks” suite is a project developed by the NASA Advanced Supercomputing (NAS) Division to evaluate the performance of highly parallel computing in supercomputers. The benchmark derives from Computational Fluid Dynamics (CFD) applications, and consists of five kernels and three pseudo-applications. Currently, NAS Benchmarks comprise 11 benchmarks, of which eight were originally available in the NPB 1 version: MultiGrid (MG), Conjugate Gradient (CG), Fast Fourier Transform (FT), Integer Sort (IS), Embarrassingly Parallel (EP), Block Tridiagonal (BT), Scalar Pentadiagonal (SP), and Lower-Upper symmetric Gauss-Seidel (LU); and three other benchmarks available from version NPB 3.1 and NPB 3.2: Unstructured Adaptive (UA), Data Cube Operator (DC), and Data Traffic (DT) [3].

According to the benchmarks presented and the works studied, these three performance tests are robust tools to evaluate cluster systems based on SoC for HPC.

To implement these tests, it is necessary to conduct a specific study in each of these tools to obtain standardized results that are comparable with the related literature.

In most of the works analyzed, it was found that the evaluation conducted made possible to know the performance and power consumed by the clusters. Performance was mainly evaluated by using performance tests already named. Power consumption is a determinant factor in HPC systems based on SoC, given that the number of instructions executed per watt consumed are related, with this being a currently recognized parameter, which was previously not relevant for the construction of clusters for HPC.

Additionally, it must be highlighted that referent measurements exist to evaluate this technology, like the Top500 and Green500 lists [26]. The former classifies supercomputers according to the performance of instructions in floating point, while the latter classifies the most energy efficient supercomputers.

The Top500 list does not have as criterion to classify big machines according to energy consumption, although it is a parameter that has always been measured. The Green500 list calculates the performance of the cluster, and captures the power consumed by the cluster during the execution of the performance tests. Thereafter, performance per watt is determined by using the following formula: Performance per watt (PPW) = Performance/power [3].

It should be stressed that when designing an HPC cluster, the main goal is to obtain a hardware configuration that offers good results in the application or sets of applications. This hardware will be evaluated through performance tests that will allow to know the efficiency of these systems. These tests must be conducted with standard benchmarks for their results to be comparable.

#### 4. CONCLUSIONS

This work presented a systematic review of the evaluations conducted on clusters of Systems on a chip for high-performance computing, using elements from the Prisma Declaration. The Declaration made possible to establish a methodological route to carry out the search and systematic review of the bibliography consulted.

As a result of the review, it was found that the performance tests most used were: High-Performance Linpack (HPL), NAS Parallel Benchmarks, and STREAM benchmark. These tests are the best known in HPC, given that they introduce standard results, which permit using comparatives. Likewise, the hardware most used were the Raspberry, Odroid, and Nvidia systems.

## REFERENCES

- [1] Schadt, E., Linderman, M., Sorenson, J. *et al.* “Computational solutions to large-scale data management and analysis,” *Nat Rev Genet*, no. 11, pp. 647–657, 2010. DOI: <https://doi.org/10.1038/nrg2857>
- [2] N. Rajovic, A. Rico, N. Puzovic, C. Adeniyi-Jones, and A. Ramírez, “Tibidabo: Making the case for an ARM-based HPC system,” *Future Generation Computer Systems*, no. 36, pp. 322–334, 2014. DOI: <https://doi.org/10.1016/j.future.2013.07.013>
- [3] N. Balakrishnan, *Building and benchmarking a low power ARM cluster*, M.S. Thesis, EPCC Edinburgh Parallel Computing Center, The University of Edinburgh, 2012. Available: <http://static.epcc.ed.ac.uk/dissertations/hpc-msc/2011-2012/Submission-1126390.pdf>
- [4] J. W. Weloli, S. Bilavarn, S. Derradji, C. Belleudy and S. Lesmanne, “Efficiency Modeling and Analysis of 64-bit ARM Clusters for HPC,” 2016 Euromicro Conference on Digital System Design (DSD), Limassol, pp. 342-347, 2016. DOI: <https://doi.org/10.1109/DSD.2016.74>
- [5] M. Görtz, R. Kühn, O. Zietek, R. Bernhard, M. Bulinski, D. Duman, B. Freisen, U. Jentsch, T. Klöppner, D. Popovic, and L. Xu, “Energy Efficiency of a Low Power Hardware Cluster for High Performance Computing,” Eibl, M. & Gaedke, M. (Hrsg.), *INFORMATIK 2017*. Gesellschaft für Informatik, Bonn, pp. 2537-2548, 2017. DOI: [https://doi.org/10.18420/in2017\\_256](https://doi.org/10.18420/in2017_256)
- [6] J. Saffran *et al.*, “A Low-Cost Energy-Efficient Raspberry Pi Cluster for Data Mining Algorithms,” in Desprez F. *et al.* (eds) *Euro-Par 2016: Parallel Processing Workshops*. Euro-Par 2016. Lecture Notes in Computer Science, vol 10104. Springer, Cham. 2017. DOI: [https://doi.org/10.1007/978-3-319-58943-5\\_63](https://doi.org/10.1007/978-3-319-58943-5_63)
- [7] M. Cloutier, C. Paradis, and V. Weaver, “A Raspberry Pi Cluster Instrumented for Fine-Grained Power Measurement,” *Electronics*, vol. 5, no. 4, p. 61, 2016. DOI: <https://doi.org/10.3390/electronics5040061>
- [8] L. Morganti, D. Cesini, and A. Ferraro, “Evaluating Systems on Chip through HPC Bioinformatic and Astrophysic Applications,” in *2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, Heraklion, pp. 541-544, 2016. DOI: <https://doi.org/10.1109/PDP.2016.82>
- [9] J. Maqbool, S. Oh, and G. C. Fox, “Evaluating ARM HPC clusters for scientific workloads,” *Concurrency Computation*, vol. 27, no. 17, pp. 5390–5410, 2015. DOI: <https://doi.org/10.1002/cpe.3602>
- [10] R. Manchado Garabito, S. Tamames Gómez, M. López González, L. Mohedano Macías, M. D’Agostino, and J. Veiga de Cabo, “Revisiones Sistemáticas Exploratorias,” *Medicina y Seguridad del Trabajo*, vol. 55, no. 215, pp. 28–51, 2009.

- 
- [11] G. Urrútia, and X. Bonfill, “Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis,” *Med. Clin. (Barc)*, vol. 135, no. 11, pp. 507–511, 2010. DOI: <https://doi.org/10.1016/j.medcli.2010.01.015>
- [12] C. Kaewkasi, and W. Srisuruk, “A study of big data processing constraints on a low-power Hadoop cluster,” 2014 International Computer Science and Engineering Conference (ICSEC), Khon Kaen, pp. 267-272, 2014. DOI: <https://doi.org/10.1109/ICSEC.2014.6978206>
- [13] E. L. Padoin, D. P. Velho, and P. O. A. Navaux, “Evaluating Performance and Energy on ARM-based Clusters for High Performance Computing,” in *41st International Conference on Parallel Processing Workshops, Pittsburgh*, 2012. DOI: <https://doi.org/10.1109/ICPPW.2012.21>
- [14] A. Selinger, K. Rupp, and S. Selberherr, “Evaluation of Mobile ARM-Based SoCs for High Performance Computing,” in *Proceedings of the 24th High Performance Computing Symposium (HPC '16). Society for Computer Simulation International*, pp. 1–7, 2016. DOI: <https://doi.org/10.22360/SpringSim.2016.HPC.022>
- [15] C. Salazar, “Medidas de rendimiento y comparación entre el Clúster Cruz I y el Clúster Cruz II,” *Revista de la Facultad de Ciencias de la UNI, Revciuni*, vol. 17, no. 1, pp. 9–16, 2014.
- [16] C. Kaewkasi and W. Srisuruk, “Optimizing performance and power consumption for an ARM-based big data cluster,” *TENCON 2014 - IEEE Region 10 Conference, Bangkok*, pp. 1-6, 2014. DOI: <https://doi.org/10.1109/TENCON.2014.7022399>
- [18] I. Stamelos, D. Soudris, and C. Kachris, “Performance and energy evaluation of spark applications on low-power SoCs,” in *2016 International Conference on Embedded Computer Systems: Architectures, Modeling and Simulation (SAMOS), Agios Konstantinos*, pp. 300-305, 2016. DOI: <https://doi.org/10.1109/SAMOS.2016.7818362>
- [19] A. Mappuji, N. Effendy, M. Mustaghfirin, F. Sondok, R. P. Yuniar and S. P. Pangesti, “Study of Raspberry Pi 2 quad-core Cortex-A7 CPU cluster as a mini supercomputer,” in *8th International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta*, 2016, pp. 1-4, 2016. DOI: <https://doi.org/10.1109/ICITEED.2016.7863250>
- [20] N. Rajovic, P. M. Carpenter, I. Gelado, N. Puzovic, A. Ramirez and M. Valero, “Supercomputing with commodity CPUs: Are mobile SoCs ready for HPC?,” in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, pp. 1-12, 2013. DOI: <https://doi.org/10.1145/2503210.2503281>
- [21] J. Zhang, S. You and L. Gruenwald, “Tiny GPU Cluster for Big Spatial Data: A Preliminary Performance Evaluation,” in *IEEE 35th International Conference on Distributed Computing Systems Workshops*, pp. 142-147, 2015. DOI: <https://doi.org/10.1109/ICDCSW.2015.33>
- [22] Z. Krpić, G. Horvat, D. Žagar and G. Martinović, “Towards an energy efficient SoC computing cluster,” 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, pp. 178-182, 2014. DOI: <https://doi.org/10.1109/MIPRO.2014.6859556>
-

- [23] L. O. Salvador, *Building a low consumption cluster using SBC technology*, B.Sc. Thesis, Ingeniería Informática, Universidad de Cantabria, 2016. Available: <http://hdl.handle.net/10902/9383>
- [24] M. Tsuji, W. T. C. Kramer and M. Sato, "A Performance Projection of Mini-Applications onto Benchmarks Toward the Performance Projection of Real-Applications," 2017 IEEE International Conference on Cluster Computing (Cluster), Honolulu, HI, pp. 826-833, 2017. DOI: <https://doi.org/10.1109/CLUSTER.2017.123>
- [25] M. Sayeed, H. Bae, Y. Zheng, B. Armstrong, R. Eigenmann and F. Saied, "Measuring High-Performance Computing with Real Applications," *Computing in Science & Engineering*, vol. 10, no. 4, pp. 60-70, 2008. DOI: <https://doi.org/10.1109/MCSE.2008.98>
- [26] A. Remy. *Solving dense linear systems on accelerated multicore architectures*, PhD thesis, Hardware Architecture, Université Paris Sud - Paris XI, 2015. Available: <https://tel.archives-ouvertes.fr/tel-01225745/document>
- [27] Top 500.org, "Top 500 The list," 2018. [Online]. Available: <https://www.top500.org/> [Accessed: 28-Jan-2018].



# Monitoreo de indicadores de valor a través de minería de datos, gestión de procesos de negocio y mejoramiento continuo con gestión del riesgo\*

Darío Enrique Soto Durán\*\*

Juan Camilo Giraldo Mejía\*\*\*

Fabio Alberto Vargas Agudelo\*\*\*\*

Jovani Jiménez Builes\*\*\*\*\*

Antonio Jesús Valderrama Jaramillo\*\*\*\*\*

Recibido: 01/07/2019 • Aceptado: 26/11/2019

<https://doi.org/10.22395/rium.v19n37a5>

## Resumen

Para las organizaciones es importante conocer el comportamiento de los procesos a través de la gestión del riesgo y la evaluación de los indicadores clave de desempeño (*Key Performance Indicators*, KPI). Un propósito del ciclo de mejoramiento continuo “planear - hacer - verificar - actuar” (PHVA) es determinar el estado de los indicadores y realizar reingeniería de procesos para lograr el cumplimiento ideal. La gestión del riesgo previene la desviación de los indicadores frente a las metas propuestas. Para determinar las variables relevantes de un proceso es necesario establecer mecanismos de monitoreo y control como forma eficiente de obtención de conocimiento basado en el uso de técnicas como las de minería de datos (MD). Estos conceptos se han articulado aquí en un modelo desarrollado para lograr el estado ideal de los KPI en una organización, y se evidencian a través de un caso de estudio aplicado en un proceso misional de una institución de educación superior, por medio del cual se describe el éxito de la teoría y el análisis planteados. Este modelo establece unos componentes a partir de la articulación de características de MD, la gestión de procesos de negocio o *Business Process Management* (BPM) y el ciclo PHVA con gestión del riesgo, de tal suerte que se hace posible un diseño ajustado a las necesidades particulares que sirve como modelo para investigaciones futuras.

**Palabras clave:** indicador clave de valor KPI; minería de datos; gestión de procesos de negocio; ciclo de mejoramiento continuo PHVA; gestión del riesgo; Norma ISO.

\* El contenido de este artículo es derivado de una investigación realizada en el Tecnológico de Antioquia por los autores, quienes forman parte del grupo de Investigación GIISTA. Esta investigación comprendió un trabajo de grado del programa de Maestría en Gestión de Tecnología de la Información de la Facultad de Ingeniería.

\*\* Ingeniero de sistemas, magíster en ciencias de la computación y Ph. D. en ingeniería de sistemas e informática. Tecnológico de Antioquia Institución Universitaria. Correo electrónico: dsoto@tdea.edu.co. Orcid: <https://orcid.org/0000-0002-6557-844X>.

\*\*\* Ingeniero de sistemas, especialista en informática, magíster en ingeniería de sistemas, y Ph. D. en ingeniería de sistemas e informática. Tecnológico de Antioquia Institución Universitaria. Email: [jgiraldo1@tdea.edu.co](mailto:jgiraldo1@tdea.edu.co), Orcid: <https://orcid.org/0000-0002-6564-3029>.

\*\*\*\* Ingeniero de sistemas, especialista en ingeniería de software, magíster en ingeniería de sistemas, y Ph. D. en ingeniería de sistemas e informática. Tecnológico de Antioquia Institución Universitaria. Email: [fvargas@tdea.edu.co](mailto:fvargas@tdea.edu.co), Orcid: <http://orcid.org/0000-0002-6921-4918>.

\*\*\*\*\* Docencia de Computadores, magíster en ingeniería de sistemas, y Ph. D. en ingeniería de sistemas e informática, Universidad Nacional de Colombia, Facultad de Minas, Medellín. Correo electrónico: [jajimen1@unal.edu.co](mailto:jajimen1@unal.edu.co). Orcid: <https://orcid.org/0000-0001-7598-7696>.

\*\*\*\*\* Ingeniero de software, magíster en gestión de tecnología de la información y especialista en administración de riesgos y seguros. Tecnológico de Antioquia Institución Universitaria. Correo electrónico: [avalderr@tdea.edu.co](mailto:avalderr@tdea.edu.co). Orcid: <https://orcid.org/0000-0002-7016-8031>.

## Monitoring Value Indicators by Applying Data Mining, Business Process Management, and Continuous Improvement with Risk Management

### **Abstract**

Recognizing the behavior of processes through risk management and the assessment of value indicators (KPI), which stands for ‘Key Performance Indicators’, is something of paramount importance for institutions. One of the purposes of the continuous PDCA improvement cycle (Plan - Do - Check - Act) is to determine the state of the indicators and carry out processes reengineering to achieve the ideal goal. Risk management comes from the deviation of the indicators with respect to the proposed goals. In order to determine the relevant variables of a process, monitoring and control mechanisms must be set as an efficient way to obtain the knowledge based on the use of data mining techniques (MD). These concepts articulate in a model that was developed to achieve the ideal condition of the KPIs within an institution, and it is evidenced through a case study applied to a missional process in an institution of higher education. The establishment of data mining, business process management (BPM), PDCA continuous improvement cycle (Plan - Do - Check - Act), and risk management characteristics was used to define the components of the model. The aforementioned allowed to create an effective model, capable of meeting the needs of this research in particular, and able to be used as a model for future research. The proposed model was applied to a specific case, which allowed to describe the success of the theory, as well as the analysis stated.

**Keywords:** Key value indicator; data mining; business process management; PDCA continuous improvement cycle; risk management; ISO Standard.

## Monitoramento de indicadores de valor por meio da mineração de dados, gestão de processos de negócio e melhoria contínua com gestão de risco

### **Resumo**

Para as organizações é importante conhecer o comportamento dos processos por meio da gestão de risco e da avaliação dos indicadores chave de desempenho (Key Performance Indicators, KPI). Um propósito do ciclo de melhoria contínua “planejar – fazer – verificar – agir” (PHVA) é determinar o estado dos indicadores e realizar reengenharia de processos para conseguir a execução ideal. A gestão de risco previne o desvio dos indicadores diante das metas propostas. Para determinar as variáveis relevantes de um processo é necessário estabelecer mecanismos de monitoramento e de controle como forma eficiente de obtenção de conhecimento baseado no uso de técnicas como as de mineração de dados (MD). Esses conceitos foram articulados aqui num modelo desenvolvido para alcançar o estado ideal dos KPI numa organização, e são evidenciados por meio de um caso de estudo aplicado num processo missional de uma instituição de educação superior, por meio do qual se descreve o sucesso da teoria e das análises propostas. Esse modelo estabelece uns componentes a partir da articulação de características de MD, a gestão de processos de negócio ou Business Process Management (BPM) e o ciclo PHVA com gestão de risco, de tal forma que se torna possível um desenho ajustado às necessidades particulares que serve como modelo para pesquisas futuras.

**Palavras-chave:** indicador chave de valor KPI; mineração de dados; gestão de processos de negócio; ciclo de melhoria contínua PHVA; gestão de risco; Norma ISO.

## INTRODUCCIÓN

Para las organizaciones es importante conocer cómo se encuentran los procesos a través de la gestión del riesgo y la evaluación de los KPI, de tal manera que sea posible aplicarles reingeniería y así establecer un ciclo de mejoramiento continuo PHVA, que mediante la gestión del riesgo determine la posibilidad de desviación por los riesgos a que se ven sujetos dichos indicadores. Los KPI suelen presentarse en un cuadro de mando integral, con clave de colores para facilitar su interpretación [1], y deben ser específicos, medibles y cuantificables [2] para evaluar el éxito de la organización [3]. Se pueden extraer por métodos o técnicas de minería de datos, los cuales acceden a las bodegas de datos o modelos analíticos y permiten analizar las variables que corresponden a los KPI [4]. Sin embargo, la carencia de reingeniería y optimización de procesos de negocio es uno de los principales problemas que afecta la competitividad de las organizaciones [5]. Una solución es hacer seguimiento con BPM, que se aplica sobre el rediseño, configuración, ejecución y monitoreo de procesos [6], con lo que se facilita la gestión empresarial con herramientas tecnológicas [7] para analizar los procesos y obtener mejora continua [8]; esto, a su turno, permite a las organizaciones hacer operaciones inteligentes [9] [10].

Las etapas que comprende el seguimiento con BPM son [11]: cinco:

- 1) *Diseño*: definición de los objetivos y estrategias.
- 2) *Modelamiento*: especificación de los usuarios y reglas para la ejecución del proceso.
- 3) *Ejecución*: entrega de la información a los usuarios involucrados.
- 4) *Monitoreo*: control del proceso para su correcta ejecución.
- 5) *Optimización*: identificación de rutas críticas, flujos a mejorar y cuellos de botella.

Algunas investigaciones relacionan este proceso de BPM con MD, para explorar las grandes bases de datos de manera automática o semiautomática. Así, es posible encontrar patrones que expliquen el comportamiento de los datos en un determinado contexto [12] y analizar el conocimiento a partir de datos almacenados [13]. Para desarrollar proyectos de MD se sigue una metodología de obtención de conocimiento conocida como *Cross Industry Standard Process for Data Mining* (CRISP-DM) [14], que consta de seis fases:

- 1) *Comprensión del negocio*: que se centra en objetivos, requerimientos y necesidades de minería de datos.

- 2) *Comprensión de los datos*: se orienta a la recopilación, exploración y comprobación de calidad de los datos.
- 3) *Preparación de los datos*: corresponde a la elección del conjunto final de datos.
- 4) *Modelado: selección* de técnicas de modelado que cumpla con los requisitos.
- 5) *Evaluación*: se evalúan los modelos para el cumplimiento de los objetivos.
- 6) *Despliegue*: se usa el conocimiento obtenido [15] [16] [17] [18].

En la literatura se encuentran, en menor cantidad, propuestas que integran BPM y MD con PHVA, hecho que plantea la necesidad de investigar y aportar propuestas que relacionen MD, BPM, Y PHVA [19]. El agregar el ciclo PHVA hace posible la planear y entender el problema, definir objetivos e identificar procesos, en aras de medir la efectividad de los cambios. Cada etapa de PHVA tiene un propósito [20]:

- *Planificar*: establecer objetivos y procesos.
- *Hacer*: implementar cambios planificados.
- *Verificar*: establecer un periodo de prueba para medir y valorar la efectividad de los cambios.
- *Actuar*: en el caso de que los resultados no se ajusten a las expectativas y objetivos predefinidos, realizar las correcciones y modificaciones necesarias.

Según Cardoso [21] es posible aprender a partir de MD encontrando reglas, usándolas para establecer la predicción del comportamiento y buscar la mejora de procesos en los cuales se presentan grandes volúmenes de datos y su análisis manual es bastante complejo [22]; con ello se optimizaría el proceso a partir de indicadores y patrones obtenidos de los datos [23]. En [24] se establece que BPM aporta de forma significativa a cumplir con los objetivos de una organización con el análisis sistemático de los procesos, sumado a la aplicación de técnicas de MD (como *Clustering* [25]) para generar conocimiento, mejorar los indicadores KPI, y apoyar en la toma de decisiones [26] [27]. Estas investigaciones han utilizado algunas técnicas de MD o BPM, o una fusión de ambas para llegar a los resultados esperados; pero se tiene la necesidad de realizar un monitoreo frecuente sobre los procesos a través de los KPI, y un análisis

de riesgo que permita aplicar PHVA sobre la cadena de valor de la organización, de tal suerte que se identifiquen amenazas o peligros que puedan surgir [28].

La gestión del riesgo, según el Sistema de Gestión ISO 31000, aporta a la valoración y tratamiento del riesgo para su respectiva evaluación, dado que tiene inmerso el ciclo PHVA. El análisis del riesgo, a su turno, se realiza a partir de la determinación de la probabilidad de ocurrencia y del efecto/consecuencia. Luego de la evaluación, se determina la priorización del riesgo, antes de la siguiente etapa —tratamiento [29]—. Empresas pertenecientes a los grupos Sura, Orbis y Aval, y la Organización Leonisa, entre otros, con buenas prácticas de gestión y que cuentan o han contado con alguno o varios sistemas de gestión ISO, han establecido sus indicadores KPI apoyados por la gestión del riesgo respecto a posibles desviaciones, con una debida “calificación de regla para relacionar indicadores respecto a sus metas”; establecen con ello su estado como “crítico”, “alerta” o “ideal”, y a partir de ello priorizan planes de acción y su respectivo seguimiento.

Este artículo presenta un modelo que integra PHVA, BPM, y MD para alcanzar el estado ideal de los KPI [30]. Su medición es la clave para dar información de gestión, reduciendo la incertidumbre y tomar decisiones adecuadas [31], aportando en la reconstrucción de procesos con MD [32] y obtención de conocimiento. El resto del artículo se organiza así: en la primera sección se establecen los materiales y métodos, que presentan la descripción del proceso para el diseño del modelo. En la segunda y tercera sección se especifica el modelo propuesto y los resultados de su aplicación en el caso de estudio, respectivamente. En la cuarta sección se plantea una discusión acerca del modelo obtenido. Por último, se presentan las conclusiones.

## 1. MATERIALES Y MÉTODOS

A partir de un proceso de revisión de literatura se identificaron las características y funcionalidades de algunos modelos que integran BPM y MD, y por ende las carencias en el monitoreo de actividades de procesos organizacionales. Como una forma de superar estas debilidades, se propone una forma de hacer seguimiento, evaluación y mejoramiento de los KPI. Para la aplicación y evaluación del modelo se tomó como caso de estudio el proceso de los semilleros de investigación del Tecnológico de Antioquia, siguiendo la metodología CRISP-DM como estándar para desarrollo de proyectos de obtención de conocimiento y MD. Consideradas las etapas de CRISP-DM se especifican los componentes del modelo y sus actividades, articuladas con MD, BPM, y PHVA, como se describe en la tabla 1.

Tabla 1. Especificación de componentes del modelo a partir de CRISP-DM

Etapas modelo propuesto	Etapas de CRISPDM
Modelado de la organización	Contexto del negocio, de los datos, y su preparación
Análisis de procesos	Modelado
Mejora de procesos	Evaluación
Realimentación	Despliegue

Fuente: elaboración propia.

### 1.1 Modelado de la organización

El modelado de la organización busca establecer la estructuración de la organización en función de los objetivos. Esta etapa comprende las siguientes actividades:

- Con MD: *comprensión del negocio* (involucra la descripción de la organización, del proceso e indicadores), *comprensión y preparación de los datos*.
- Con BPM: *diseño y modelamiento del proceso*.
- Con PHVA mediante la gestión del riesgo: *planeación*, que involucra la determinación del contexto de la organización que articula la actividad de comprensión del negocio de la MD y la identificación de las amenazas o riesgos.

### 1.2 Análisis de procesos

El análisis de procesos busca encontrar los registros o datos de los diferentes indicadores, determinándolos a partir de los datos que se hayan ingresado al sistema de registros. Consta de las siguientes actividades:

- Con MD: modelado y evaluación.
- Con BPM: ejecución.
- Con PHVA mediante la gestión del riesgo: hacer, determinando la evaluación y priorización del riesgo.

### 1.3 Mejora de procesos

En la mejora de procesos, al margen del estado en que se encuentren los KPI, se evalúa la situación de cada una de las variables frente al proceso de gestión. Comprende las siguientes actividades:

- Con MD: despliegue.
- Con BPM: monitoreo.

- Con PHVA mediante la gestión del riesgo: verificar, determinando el cumplimiento de planes de acción previamente establecidos.

### 1.4 Realimentación

La realimentación, o plan de mejora, es posterior a la evaluación; se establece en ella la optimización del KPI, con un correspondiente plan de mejora. Consta de tres actividades:

- Con MD: despliegue.
- Con BPM: optimización.
- Con PHVA mediante la gestión del riesgo: actuar, determinando nuevos planes de acción o mejoras, acorde al cumplimiento de los indicadores o nuevos riesgos.

## 2. MODELO PROPUESTO

En esta sección se presenta el modelo propuesto, integrado por componentes y actividades, especificados a partir de las etapas de la metodología CRISP-DM, el ciclo PHVA y la estrategia para BPM. En la figura 1 se representa el modelo.

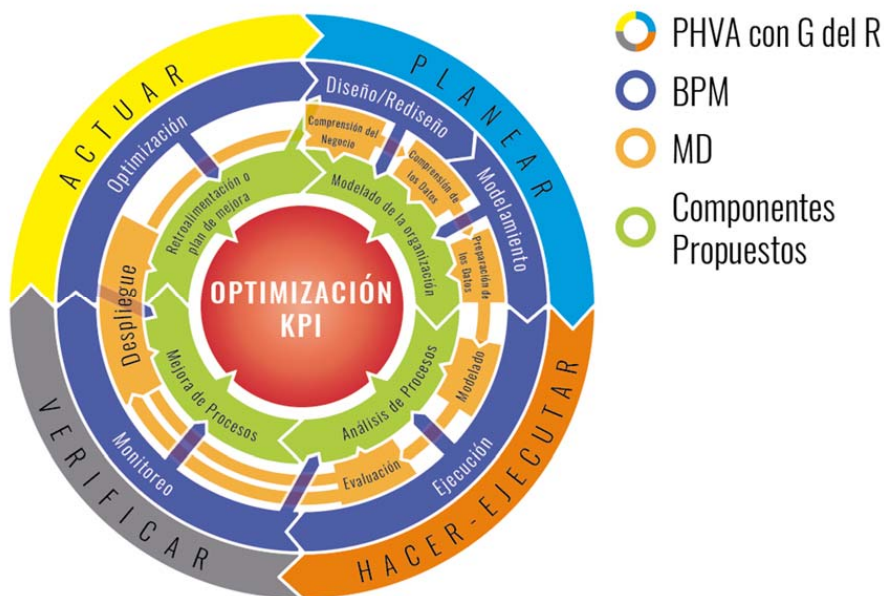


Figura 1. Modelo propuesto: integración de componentes

Fuente: elaboración propia.

## 2.2 Especificación y articulación de los componentes del modelo

A continuación se presenta la relación de los componentes propuestos para el modelo, y su articulación con MD, BPM y PHVA con gestión del riesgo.

Articulación con el componente *modelado de la organización*: *planear* permite conocer la organización, establecer los objetivos y el estado actual de los indicadores del proceso, y definir un cronograma de trabajo (figura 2).

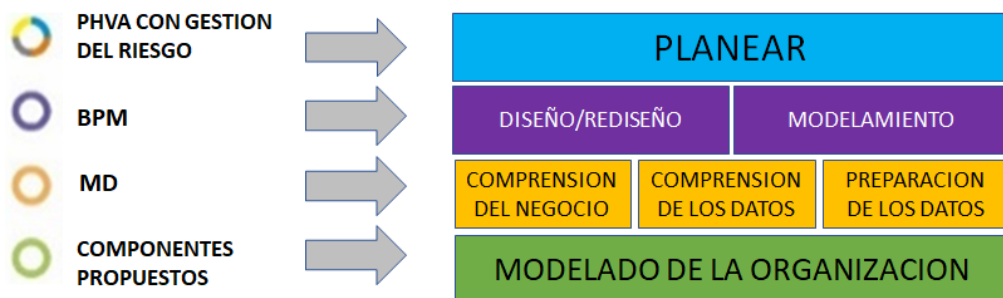


Figura 2. Articulación del componente *modelado de la organización*

Fuente: elaboración propia.

Articulación del componente *análisis del proceso*: *hacer* es el camino hacia la acción; se ejecuta el cronograma planeado y se evalúa todo el proceso en ejecución, con lo que se cumple el respectivo análisis de procesos (figura 3).

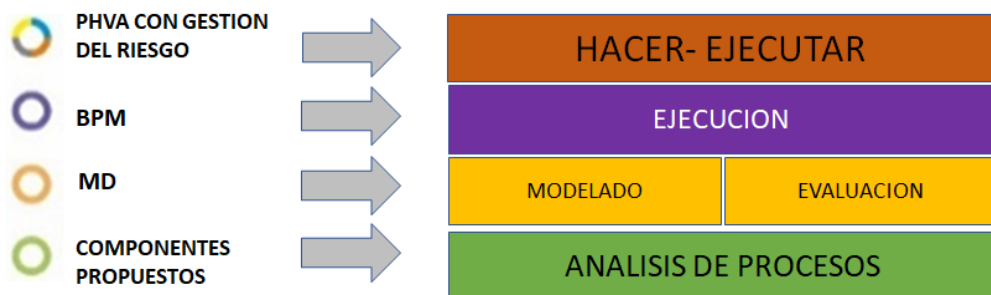


Figura 3. Articulación del componente *análisis del proceso*

Fuente: elaboración propia.

Articulación con el componente *mejora de procesos*: *verificar* es el camino hacia la revisión y monitoreo de la ejecución del cronograma planeado, demostrándose el despliegue respectivo (figura 4).



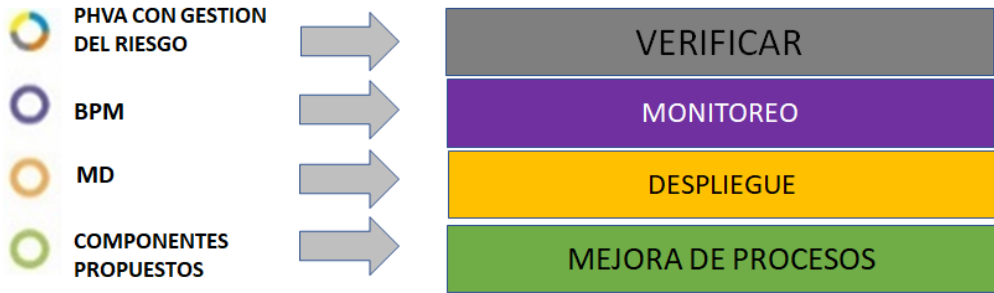


Figura 4. Articulación del componente *mejora de procesos*.

Fuente: elaboración propia.

Articulación con el componente *retroalimentación: actuar* es la última etapa de la revisión y seguimiento del proceso; en su desarrollo se toman decisiones de rediseño de procesos, se realizan acciones de despliegue, y se da paso a la retroalimentación o plan de mejora (figura 5).

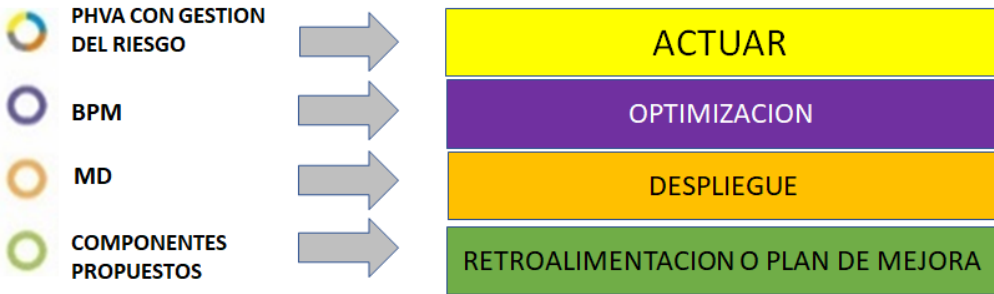


Figura 5. Articulación con componente *plan de mejora*

Fuente: elaboración propia.

### 3. CASO DE ESTUDIO

#### 3.1 Modelado de la organización

El caso de estudio para evaluar la aplicación del modelo se llevó a cabo en el Tecnológico de Antioquia Institución Universitaria - TdeA (Medellín, Colombia), al interior de la Dirección de Investigaciones de dicha institución.

##### 3.1.1 Comprensión del negocio

El proceso analizado corresponde a la planeación y seguimiento a semilleros de investigación del TdeA (figura 6).

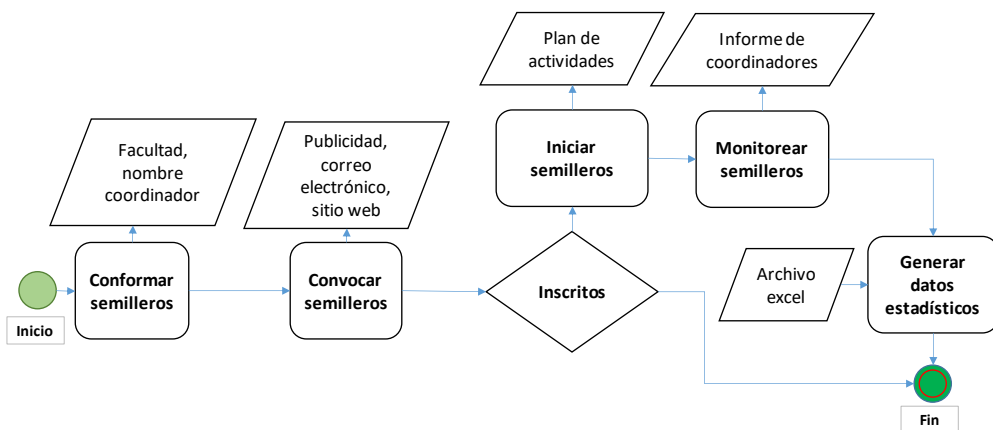


Figura 6. Flujo del proceso de semilleros de investigación TdeA

Fuente: elaboración propia.

### 3.1.2 Comprensión y preparación de los datos

*Especificación de fuentes de datos.* Los datos se encuentran en una hoja de cálculo (en la aplicación Microsoft Excel<sup>®</sup>) que contiene la información de registros que realiza la Dirección de Investigaciones. Se seleccionaron las hojas correspondientes al primer y segundo semestres académicos de los años 2016, 2017 y 2018 (6 hojas de registros en total), con la selección de campos *periodo, facultad y grupo de investigación*.

*Especificación de Indicadores.* La Dirección de Investigaciones requiere varios indicadores, entre los que se encuentran los siguientes: participación de estudiantes en semilleros de investigación; cantidad anual de semilleros; proporción anual de estudiantes por semillero; total acumulado de estudiantes en semilleros por facultades; y total acumulado de cantidad semilleros por facultades, entre otros.

### 3.1.3 Gestión del riesgo en componente planear

En la identificación del riesgo de las amenazas a evaluar, se destaca la baja vinculación de nuevos docentes con alto nivel de cualificación, la inadecuada estrategia de divulgación de semilleros y el bajo reconocimiento, apoyo económico y estímulos.

## 3.2 Análisis de procesos

El análisis realizado a partir de los indicadores respecto a la meta definida se presenta en la tabla 2. Para cada indicador de gestión de procesos se definió una regla compuesta por un valor base y un porcentaje (parámetros establecidos con anterioridad por expertos de la organización).

Tabla 2. Calificación de regla para relacionar indicadores respecto a su meta.

Rangos en valor absoluto	Rango en porcentaje sobre una meta de 5 (valor absoluto / meta)* 100	Priorización e interpretación de los rangos en valor absoluto o porcentajes
>4	> 80 % o 0,80	Ideal
<= 4 y >= 3	<=80 % o 0,80 y >= 60 % o 0,60	Alerta
<3	< 60 % o 0,60	Crítico

Fuente: elaboración propia.

### 3.2.1 Modelado

El indicador de valor KPI se determinó a partir del periodo anual, con una regla compuesta por el presupuesto o meta, el resultado obtenido, y el porcentaje de cumplimiento. Para establecer la participación de los estudiantes del TdeA en semilleros de investigación se definieron variables que indican la relación en cada año, de lo proyectado a alcanzar y lo obtenido al finalizar cada periodo académico. Estas metas son establecidas desde la Dirección de Investigaciones.

### 3.2.2 Ejecución

En la tabla 3 se especifica la regla de determinación del indicador KPI para presupuestos vs. resultados y el porcentaje de cumplimiento por cada año.

Tabla 3. Determinación del indicador KPI y meta del proceso para *participación de los estudiantes en los semilleros*

Periodo (año)	Regla del Indicador		
	Presupuesto o Meta	Resultado obtenido	Porcentaje de cumplimiento (resultado/meta) *100
2016	500	51	10,2 %
2017	500	350	70,0 %
2018	500	554	110,8 %

Fuente: elaboración propia.

*Hacer - ejecutar.* En este punto se realiza la aplicación de la metodología de gestión del riesgo, con la determinación de su análisis respectivo.

*Análisis del Riesgo.* Se establecen las probabilidades y frecuencias, con lo cual se definen los rangos y niveles. Las probabilidades establecidas permitirán determinar la frecuencia de las ocurrencias —en este caso, las causas que relacionan las acciones o actividades que corresponden a la participación de los estudiantes en los semilleros—. En la tabla 4 se presenta la probabilidad/frecuencia de la metodología de gestión de riesgo.

Tabla 4. Tabla de probabilidad / frecuencia

Tabla de Probabilidad / Frecuencia		
Descripción detallada	Rangos	Nivel
Ocurriría solo bajo circunstancias excepcionales	Raro	1
Podría ocurrir algunas veces	Improbable	2
Puede ocurrir en algún momento	Posible	3
Ocurrencia en la mayoría de las circunstancias	Probable	4
La ocurrencia se da en la mayoría de las circunstancias	Casi seguro	5

Fuente: elaboración propia con base en la Norma ISO 31000.

Los impactos de la metodología de gestión de riesgos se reflejan en lo financiero, lo operativo, la imagen organizacional y los clientes. Para cada uno de estos aspectos en las se describen los impactos, con los rangos y niveles definidos, en las tablas 5, 6, 7 y 8.

Tabla 5. Impacto en lo financiero.

Impacto en lo financiero		
Descripción detallada en millones	Rangos	Nivel
Pérdida menor a \$ 10	Insignificante	1
Pérdida mayor o igual a \$ 10 y menor a \$ 20	Menor	2
Pérdida mayor o igual a \$ 20 y menor a \$ 30	Intermedio	3
Pérdida mayor o igual a \$ 30 y menor a \$ 40	Mayor	4
Pérdida mayor o igual a \$ 40	Superior	5

Fuente: elaboración propia con base en la Norma ISO 31000.

Tabla 6. Impacto en la operación

Impacto en la operación		
Descripción detallada	Rangos	Nivel
Indisponibilidad del semillero de investigación	Insignificante	1
Indisponibilidad a nivel del Grupo de investigación	Menor	2
Indisponibilidad a nivel de Facultad	Intermedio	3
Indisponibilidad de la Dirección de Investigación	Mayor	4
Indisponibilidad a nivel externo	Superior	5

Fuente: elaboración propia con base en la Norma ISO 31000.

Tabla 7. Impacto en la imagen

Impacto en la imagen		
Descripción detallada	Rangos	Nivel
Identifica el problema de imagen en el grupo	Insignificante	1
Identifica el problema de imagen institucional	Menor	2
Identifica el problema de imagen a nivel regional	Intermedio	3
Identifica el problema de imagen a nivel nacional	Mayor	4
Identifica el problema de imagen internacionalmente	Superior	5

Fuente: elaboración propia con base en la Norma ISO 31000.

Tabla 8. Impacto en los clientes / mercado

Impacto en los clientes / mercado		
Descripción detallada	Rangos	Nivel
Pérdida menor de 5 % de estudiantes en semilleros	Insignificante	1
Pérdida hasta 10 % de estudiantes en semilleros	Menor	2
Pérdida hasta 15 % de estudiantes en semilleros	Intermedio	3
Pérdida hasta 25 % de estudiantes en semilleros	Mayor	4
Pérdida superior a 50 % de estudiantes en semilleros	Superior	5

Fuente: elaboración propia con base en la Norma ISO 31000.

### Evaluación

La evaluación del riesgo es el producto del nivel de la probabilidad/frecuencia por el nivel de cada impacto involucrado. En la tabla 9 se presenta la evaluación del riesgo a partir de la definición de valores de probabilidad/frecuencia (tabla 4) y los impactos (tablas 5, 6, 7 y 8). La evaluación del riesgo facilita la toma de decisiones en función de la manera en que se priorizan, dentro de lo cual se destaca la criticidad de cada grado de riesgo (GR).

Tabla 9. Grado de riesgo

Nivel de probabilidad \ Nivel de impacto	Nivel de impacto				
	5	4	3	2	1
5	25	20	15	10	5
4	20	16	12	8	4
3	15	12	9	6	3
2	10	8	6	4	2
1	5	4	3	2	1

Fuente: elaboración propia con base en la Norma ISO 31000.

### 3.3 Mejora de procesos

#### 3.3.1 Despliegue

Se aplicó la técnica *clustering* de MD al conjunto de registros conformado por los atributos de año; nombre de facultad; nombre del grupo; nombre del programa; nombre del semillero; semestre o periodo académico; y total de estudiantes. El algoritmo se entrenó tomando como variables de entrada los atributos anteriores, y como variable de predicción el Total Estudiantes.

*Diagrama de perfiles.* Con la aplicación de la técnica se muestran los registros agrupados en subconjuntos o subgrupos —denominados ‘clústeres’—, en un diagrama de perfiles que se presenta en la figura 7. Este diagrama muestra cuatro columnas:

variables (atributos que se eligieron para el análisis de datos, esto es, variables de predicción y entrada); estados (colores que identifican a cada uno de los datos de las variables); datos que contiene cada variable; y concentración de datos por variable (especifica los porcentajes indicados por colores).

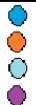

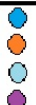
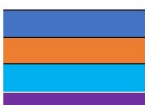










Atributos			Perfiles del Clúster
Variables	Estados		Población (Todo) Tamaño: 139
Año		2018 2017 2016 ausente	
Nombre F		FACULTAD DE INGENIERIA FACULTAD DE EDUCACIÓN Y CIENCIAS SOCIALES FACULTAD DE CIENCIAS ADMINISTRATIVAS Y ECONOMICAS FACULTAD DE DERECHO Y CIENCIAS FORENSES Otros	
Nombre G		INTEGRA OBSERVATOS GIISTA OBSERVATORIO PUBLICO Otros	
Nombre P		Ingeniería Ambiental Administración Financiera Psicología Negocios y comercio Internacional Otros	
Nombre		En familia, texto, y contexto Aplicaciones económicas en Finanzas y negocios Estudios Internacionales Desarrollo Organizacional, responsabilidad social Otros	
Semestre		2 1 Ausente	
Total Estudiantes		17,56 5,0 1,0	

Figura 7. Diagrama de perfiles de clúster

Fuente: elaboración propia.

De acuerdo con el análisis del comportamiento de las variables de la figura 7, “año” muestra la mayor concentración de datos en el 2018, y el color que relaciona es azul. Los datos relacionados con el total de estudiantes que participan en semilleros de investigación a nivel institucional indican una mayor concentración (color violeta) para el año 2017 y menor para el 2016 (color verde). Para la variable “facultad”, los colores indican el espacio en que se presentó más concentración de estudiantes mediante otro código de colores; y la variable “total estudiantes” muestra un comportamiento general, lo cual indica que entre 1 y 18 estudiantes se concentran en uno o más semilleros.

*Diagrama de características.* Este presenta un resumen del comportamiento de las variables respecto al total de registros. La probabilidad más alta se presenta en los datos del atributo “semestre” (valor 2), y se destaca el año 2018; así mismo, la facultad con más concentración de estudiantes es la de Ingeniería: entre 5 y 7 estudiantes participan en semilleros de investigación (figura 8).

Características para población (Todo)		
VARIABLES	VALORES	PROBABILIDAD
SEMESTRE	2	
ANIO	2018	
SEMESTRE	1	
NOMBRE F	FACULTAD DE INGENIERÍA	
ANIO	2017	
NOMBRE F	FACULTAD DE EDUCACIÓN Y CIENCIAS SOCIALES	
TOTAL ESTUDIANTES	5-7	
TOTAL ESTUDIANTES	2-4	
TOTAL ESTUDIANTES	8-17	
NOMBRE F	FACULTAD DE CIENCIAS ADMINISTRATIVAS	
NOMBRE G	INTEGRA	
ANIO	2016	
NOMBRE G	OBSERVATOS	
NOMBRE G	GIISTA	
NOMBRE G	OBSERVATORIO PUBLICO	
NOMBRE G	RED	
NOMBRE P	INGENIERÍA AMBIENTAL	
NOMBRE P	ADMINISTRACIÓN FINANCIERA	
NOMBRE P	PSICOLOGÍA	
NOMBRE G	SENDEROS	
NOMBRE P	NEGOCIOS Y COMERCIO INTERNACIONAL	

Figura 8. Diagrama de características de clúster

Fuente: elaboración propia.

En la figura 9 puede observarse el comportamiento de la participación de los estudiantes frente a su meta. De aquí se observa el crecimiento de la participación en semilleros entre 2016 y 2018, representado en un 10 %, 70 % y 111 %, respectivamente. Igualmente, se destaca que la meta ha estado por encima de los estudiantes participantes, excepto en el año 2018, en que se superó.

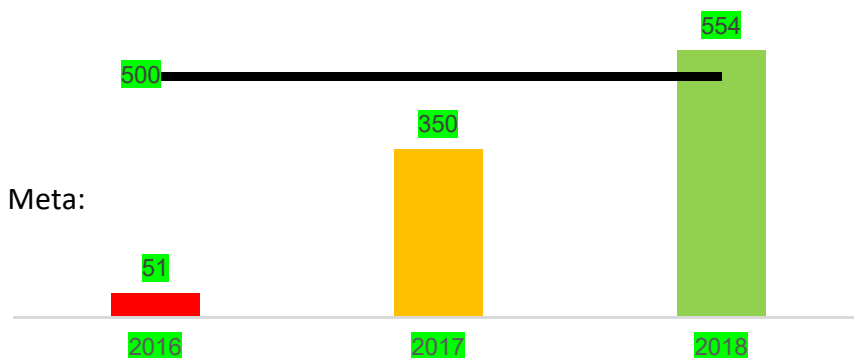


Figura 9. Total de estudiantes en semillero anual institucional vs. estudiantes participantes esperados (meta)

Fuente: elaboración propia.

*Monitoreo.* A partir de la aplicación de la metodología de gestión del riesgo se presentan los resultados de las acciones de monitoreo. En la figura 10 se evidencia un crecimiento paulatino de la cantidad de semilleros de investigación entre 2016 y 2018 —representado en 18 %, 29 % y 53 % sobre el total, respectivamente—.

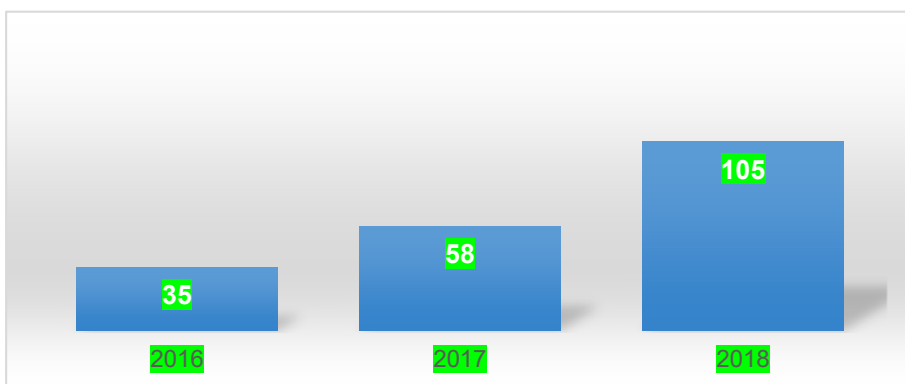


Figura 10. Cantidad de semilleros por año a nivel institucional

Fuente: elaboración propia.

En la figura 11 se observa un crecimiento paulatino en la proporción del número de estudiantes por semillero por año de 314 % y 262 % en 2017 y 2018, respectivamente, frente al 2016.



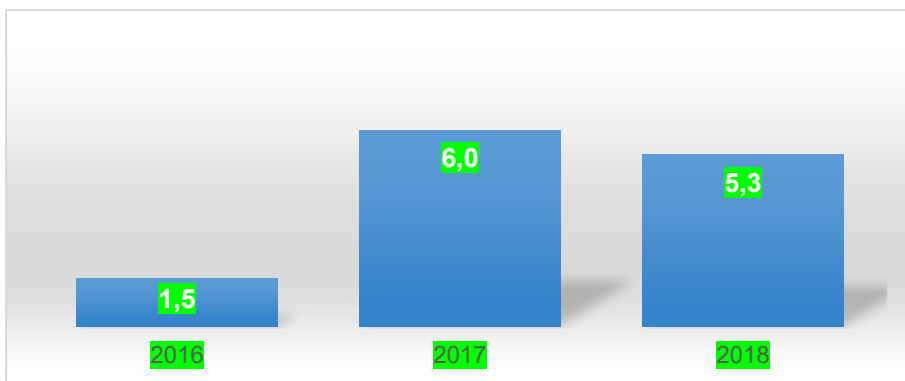


Figura 11. Proporción del número de estudiantes por semillero por año a nivel institucional

Fuente: elaboración propia.

La proporción de estudiantes en los semilleros está representada por las facultades de Ingeniería, Educación y Ciencias Sociales, Ciencias Administrativas y Económicas, y Derecho y Ciencias Forenses, y por el Departamento de Ciencias Básicas y Áreas Comunes: ostentan el 40 %, 32 %, 16 %, 6 % y 6 % en el periodo estudiado, respectivamente (figura 12).



Figura 12 Total acumulado de estudiantes en semilleros por facultad, 2016 - 2018

Fuente: elaboración propia.

De las cinco facultades de la investigación, el 72 % de los estudiantes participantes en semilleros pertenecen a las de Ingeniería, Educación y Ciencias Sociales.

*Evaluación del riesgo.* A partir de la información analizada, en las tablas 10, 11 y 12 se presenta la evaluación del riesgo de las amenazas.

Tabla 10. Gestión de riesgo por amenazas para 2016.

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica
					En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN	En Los CLIENTES / MERCADO	
Baja vinculación de nuevos docentes con alto perfil o nivel de cualificación	Probabilidad de ocurrencia en la mayoría de las circunstancias	Probable	4	Descripción	Pérdida menor a \$10 millones	Hay indisponibilidad a nivel de Dirección de Investigación / Institución	Se identifica el problema de imagen a nivel de Institución	Pérdida superior al 50% de los estudiantes participantes en semilleros.	20
				Rango	Insignificante	Mayor	Menor	Superior	
				Nivel	1	4	2	5	
				XXX	4	16	8	20	

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica
					En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN	En Los CLIENTES / MERCADO	
Inadecuada estrategia de divulgación de semilleros (poca publicidad o escasa divulgación por parte de docentes).	La expectativa de ocurrencia se da en la mayoría de circunstancias	Casi seguro	5	Descripción	Pérdida menor a \$10 millones	Hay indisponibilidad a nivel de Dirección de Investigación / Institución	Se identifica el problema de imagen a nivel de Institución	Pérdida superior al 50% de los estudiantes participantes en semilleros.	25
				Rango	Insignificante	Mayor	Menor	Superior	
				Nivel	1	4	2	5	
				XXX	5	20	10	25	

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica
					En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN	En Los CLIENTES / MERCADO	
Bajo reconocimiento, apoyo económico y estímulos (Motivación y sensibilización de los estudiantes)	Probabilidad de ocurrencia en la mayoría de las circunstancias	Probable	4	Descripción	Pérdida menor a \$10 millones	Hay indisponibilidad a nivel de Dirección de Investigación / Institución	Se identifica el problema de imagen a nivel de Institución	Pérdida superior al 50% de los estudiantes participantes en semilleros.	20
				Rango	Insignificante	Mayor	Menor	Superior	
				Nivel	1	4	2	5	
				XXX	4	16	8	20	

Fuente: elaboración propia.

Tabla 11. Gestión del riesgo por amenazas para 2017.

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica
					En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN	En Los CLIENTES / MERCADO	
Baja vinculación de nuevos docentes con alto perfil o nivel de cualificación	Puede ocurrir en algún momento	Posible	3	Descripción	Pérdida menor a \$10 millones	Hay indisponibilidad a nivel de Dirección de Investigación / Institución	Se identifica el problema de imagen a nivel de Institución	Pérdida hasta el 25% de los estudiantes participantes en semilleros.	12
				Rango	Insignificante	Mayor	Menor	Mayor	
				Nivel	1	4	2	4	
				XXX	3	12	6	12	

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica				
					Descripción	En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN		En Los CLIENTES / MERCADO			
						Rango	Insignificante	Mayor		Menor	Mayor		
Inadecuada estrategia de divulgación de semilleros (poca publicidad o escasa divulgación por parte de docentes).		Probabilidad de ocurrencia en la mayoría de las circunstancias		Probable		4		XXX	4	16	8	16	16

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica				
					Descripción	En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN		En Los CLIENTES / MERCADO			
						Rango	Insignificante	Mayor		Menor	Superior		
Bajo reconocimiento, apoyo económico y estímulos (Motivación y sensibilización de los estudiantes)		Puede ocurrir en algún momento		Posible		3		XXX	3	12	6	15	15

Fuente: elaboración propia.

Tabla 12. Gestión del riesgo por amenazas para 2018.

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica				
					Descripción	En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN		En Los CLIENTES / MERCADO			
						Rango	Insignificante	Intermedio		Menor	Intermedio		
Baja vinculación de nuevos docentes con alto perfil o nivel de cualificación		Puede ocurrir en algún momento		Posible		3		XXX	3	9	6	9	9

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"		Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica				
					Descripción	En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN		En Los CLIENTES / MERCADO			
						Rango	Insignificante	Mayor		Menor	Intermedio		
Inadecuada estrategia de divulgación de semilleros (poca publicidad o escasa divulgación por parte de docentes).		Puede ocurrir en algún momento		Posible		3		XXX	3	12	6	9	12

Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"	Probabilidad / Frecuencia			IMPACTO				Mayor grado de riesgos (GR) de la amenaza específica	
				En lo FINANCIERO	En la OPERACIÓN	En la IMAGEN	En Los CLIENTES / MERCADO		
				Descripción	Rango	Nivel			
Bajo reconocimiento, apoyo económico y estímulos (Motivación y sensibilización de los estudiantes)	Probabilidad de ocurrencia en la mayoría de las circunstancias	Probable	4	Descripción	Pérdida menor a \$10 millones	Hay indisponibilidad a nivel de Dirección de Investigación / Institución	Se identifica el problema de imagen a nivel de Institución	Pérdida hasta el 10% de los estudiantes participantes en semilleros.	16
				Rango	Insignificante	Mayor	Menor	Menor	
				Nivel	1	4	2	2	
				XXX	4	16	8	8	

Fuente: elaboración propia.

Los colores del semáforo utilizado en la determinación de GR en las tablas anteriores van de rojo (crítico) a verde (menos crítico). Se observa un decremento de los rojos y los amarillos hacia los verdes con el incremento de los años del estudio desde 2016 hasta 2018.

### 3.3 Retroalimentación o plan de mejora

#### 3.3.1 Despliegue

La retroalimentación o plan de mejora se realiza a las medidas e indicadores del periodo actual: se determina el grado de cumplimiento con rangos por valor absoluto, porcentaje referente a la meta y priorización del estado (ideal, alerta o crítico), tal como se presenta en la tabla 13.

Tabla 13. Priorización del GR.

Rango del GR en valor absoluto	Rango del GR en porcentaje sobre el valor máximo = 25 (valor absoluto / máximo) * 100	Priorización del GR
>= 20	>= 80 % o 0,80	Crítico
< 20 y >= 10	<80 % o 0,80 y >= 40 % o 0,40	Alerta
<10	< 40 % o 0,40	Ideal

Fuente: elaboración propia.

Luego de la priorización de la evaluación del riesgo por año se realizó la consolidación de las amenazas sobre el “No cumplimiento de la meta de cantidad de estudiantes en los semilleros”, dentro de lo que se destacaron los estados “crítico” y “alerta” (tablas 14, 15 y 16).

Tabla 14. Consolidado de las amenazas, año 2016.

Valor	Amenazas sobre el "No cumplimiento de la meta de cantidad de estudiantes en los semilleros"
25	Bajo reconocimiento, apoyo económico y estímulos
20	Baja vinculación de nuevos docentes con alto perfil o nivel de cualificación
16	Baja vinculación de nuevos docentes con alto perfil o nivel de cualificación
12	Poco aporte en planes institucionales, facultad, dirección de investigación
10	Bajo reconocimiento, apoyo económico y estímulos

Fuente: elaboración propia.

Tabla 15. Consolidado de las amenazas, año 2017.

Valor	Amenazas "No cumplimiento de la meta de estudiantes en semilleros"
25	*****
20	*****
16	Bajo reconocimiento, apoyo económico y estímulos
15	Bajo reconocimiento, apoyo económico y estímulos
12	Baja vinculación de nuevos docentes nivel de cualificación
10	Bajo reconocimiento, apoyo económico y estímulos

Fuente: elaboración propia.

Tabla 16. Consolidado de las amenazas, año 2018.

Valor	Amenazas "No cumplimiento de la meta de estudiantes en semilleros"
25	*****
20	*****
16	Poco aporte e impacto en los planes institucionales, de facultad, dirección de investigación, mínimo impacto en la investigación formativa.
15	Poco aporte e impacto en los planes institucionales, de facultad
12	Bajo reconocimiento, apoyo económico y estímulos. Mínimo impacto en la investigación formativa
10	*****

Fuente: elaboración propia.

*Optimización y actuar.* El tratamiento del riesgo, como optimización y actuación frente a las debilidades manifestadas durante la valoración del GR, se describe en la tabla 17. Este tratamiento implica la selección de acciones, entre las que se puede nombrar: evitar el riesgo, incrementar el riesgo como una oportunidad, retirar la fuente de riesgo, cambiar la probabilidad, cambiar las consecuencias, transferir o compartir el riesgo, y retener el riesgo como decisión informada.

Tabla 17. Definición de recomendaciones a partir de la priorización del GR, año 2018.

Grado de Riesgo (GR)		Alerta			
Amenazas sobre :		Inadecuada estrategia de publicidad y bajo voz a voz de los docentes hacia los estudiantes.			
Tratamiento	Evitar el riesgo: no iniciar o no continuar				
	Incrementar el riesgo como una oportunidad				
	Retirar la fuente del riesgo				
	Cambiar la probabilidad		X		
	Cambiar las consecuencias		X		
	transferir o compartir el riesgo				
	Retener el riesgo como decisión informada				
Plan de acción sugerido (general)	Responsable sugerido	Fecha sugerida	Recursos o Costo aprox. (\$)	Seguimiento	
Continuar con la divulgación de los semilleros en temas de proyectos por realizar, realizados y en beneficios de los mismos en aportes a la investigación	Dirección de investigación, Decanaturas y coordinadores de semilleros con apoyo de docentes de Áreas Técnicas	Semestres 1 y 2 de 2019	Menos de \$10 millones		
Grado de Riesgo (GR)		Alerta			
Amenazas sobre :		Bajo reconocimiento, apoyo económico y estímulos.			
Tratamiento	Evitar el riesgo: no iniciar o no continuar				
	Incrementar el riesgo como una oportunidad				
	Retirar la fuente del riesgo				
	Cambiar la probabilidad				
	Cambiar las consecuencias				
	transferir o compartir el riesgo				
	Retener el riesgo como decisión informada				
Plan de acción sugerido (general)	Responsable sugerido	Fecha sugerida	Recursos o Costo aprox. (\$)	Seguimiento	
Continuar con la divulgación de los reconocimientos, de apoyos económicos y de los estímulos a participantes de semilleros, ante la comunidad Institucional.	Dirección de investigación, Decanaturas y coordinadores de semilleros	Semestres 1 y 2 de 2019	Menos de \$10 millones		
Grado de Riesgo (GR)		Ideal			
Amenazas sobre :		Baja vinculación de nuevos docentes con perfiles y titulaciones más elevados.			
Tratamiento	Evitar el riesgo: no iniciar o no continuar				
	Incrementar el riesgo como una oportunidad				
	Retirar la fuente del riesgo				
	Cambiar la probabilidad		X		
	Cambiar las consecuencias				
	transferir o compartir el riesgo				
	Retener el riesgo como decisión informada				
Plan de acción sugerido (general)	Responsable sugerido	Fecha sugerida	Recursos o Costo aprox. (\$)	Seguimiento	
Continuar con la vinculación de docentes con perfiles titulaciones más elevados, como apoyo a los semilleros.	Vicerectoría académica	Semestres 1 y 2 de 2019	Lo define Vicerectoría académica y Rectoría		

Fuente: elaboración propia.

#### 4. DISCUSIÓN

Las investigaciones revisadas a manera de antecedentes permitieron identificar fortalezas en cuanto a la integración de MD y BPM para la gestión del riesgo. Sin embargo, se encontraron debilidades referidas a los temas de mejora continua con gestión del riesgo, aplicables a la determinación de indicadores y sus posibles incumplimientos. Frente a ello, en este trabajo se presenta un modelo que integra los resultados desde las áreas del ciclo PHVA, BPM y MD: aporta más integralidad en dichos resultados y brinda apoyo para la gestión del riesgo y el análisis del estado de los KPI obtenidos a partir de un modelo dimensional.

En el marco de los procesos vitales de la Dirección de Investigaciones del Tecnológico de Antioquia se estableció una escala de calificación que permite conocer el nivel de cumplimiento y posible optimización de los KPI y de los registros estadísticos, luego de un proceso de minería de datos. Para ello, se definió un listado de amenazas que pudiesen afectar de forma negativa el cumplimiento de la meta asociada a la cantidad de estudiantes en los semilleros de investigación proyectados. Al generar un listado y tablas de probabilidad - frecuencia, así como y la respectiva valoración de impactos en varios aspectos (como los que se involucran en este trabajo), se logró llegar a la valoración del riesgo, su priorización y la determinación de planes de mejora para su gestión.

#### 5. CONCLUSIONES

El modelo propuesto se presenta como una aproximación para realizar gestión del riesgo a partir del ciclo PHVA. Su valor agregado está representado por la incorporación de componentes que no solo permiten la verificación del proceso de negocio bajo análisis a través de BPM, sino que a la vez se obtienen y analizan sus indicadores, con lo que se identifica su estado a partir de probabilidades y frecuencias que se miden desde la gestión del riesgo.

La incorporación del modelo propuesto en el proceso que relaciona los semilleros de investigación de la institución estudiada ha permitido conocer el comportamiento de la participación de los estudiantes a nivel institucional. En este sentido, el proceso permitió resaltar las variables más significativas que permiten determinar por qué se ha presentado más participación en unos periodos académicos respecto a otros, y las causas que motivan dicho comportamiento. Estas últimas se analizaron a partir de los riesgos y permitieron activar elementos del modelo que están orientados a revisar el proceso con el componente BPM, y generar mejora continua a través del componente PHVA. La incorporación del modelo aplicado periodo tras periodo en la definición, valor y seguimiento de los indicadores ayuda a definir planes de acción que favorezcan

y posibiliten una mayor participación de los estudiantes en los diferentes semilleros de investigación de las facultades del Tecnológico de Antioquia.

## REFERENCIAS

- [1] “ISOTools,” [En línea]. Available: <https://www.isotools.org/soluciones/procesos/kpis-indicadores/>.
- [2] ”JLV Enterprises,” 2018. [En línea]. Available: <https://jlvbusiness.com/2017/06/19/que-es-un-kpi-y-cuales-son-sus-caracteristicas/>.
- [3] T. P. y. C. Felden, “On the restriction to numeric indicators in Performance Measurement Systems,” 15th IEEE International Enterprise Distributed Object Computing Conference Workshops, 2011.
- [4] M. A. y. J. J. Jeng, “A Tool Framework for KPI Application Development,” IEEE International Conference on e-Business Engineering, 2007.
- [5] Z. Wang, Y. Qing y Y. S., “The Research of Process Mining Assesment used in Business Intelligence,” IEEE, 11th International Conference on Computer and Information Science, 2012.
- [6] E. Souza, “Challenges in Performance Analysis in Enterprise Architectures,” *IEEE Computer Society*, 2013.
- [7] G. Xiang, “Towards the Next Generation Intelligent BPM – In the Era of Big Data,” *Springer*, 2013
- [8] A. Orellana, C. Sanchez y L. Gonzalez, “Aplicación del Modelo L\* de minería de proceso al módulo Almacén del Sistema de Información Hospitalaria alas HIS,” *13th Laccei International Conference*, 2015.
- [9] R. X. J. Y. H. L. W. Z. Wencong Lu, “Data mining-aided materials discovery and optimization.,” 2017.
- [10] A. K. M. Nidhi Tomar, “A Survey on Data Mining Optimization Techniques.,” *International Journal of Science Technology & Engineering | Volume 2 | Issue 06*, December 2015.
- [11] F. N. a. B. M. Christoph Gröger, “Data Mining-driven Manufacturing Process Optimization.,” *Proceedings of the World Congress on Engineering 2012 Vol III.*, p. Vol III., 2012.
- [12] J. I. Gaiña Yungán, “<http://dspace.uniandes.edu.ec/handle/123456789/1815>,” 12 2015. [En línea]. Available: <http://dspace.uniandes.edu.ec/handle/123456789/1815>.
- [13] J. J. P. R. A. E. A. Surelys Pérez Jiménez, “Modelo clustering para el análisis en la ejecución de procesos de negocio,” *Investigación Operacional*, pp. Vol 33, No. 3, 210-221, 2012.



- [14] R. Alcover *et al.*, “Análisis del rendimiento académico en los estudios de informática de la Universidad Politécnica de Valencia aplicando técnicas de minería de datos”, 2007 [internet]. Disponible en: <http://bioinfo.uib.es/~joemi/procJenui/Jen2007/alanal.pdf>.
- [15] “Significados”. [internet]. Disponible en: <https://www.significados.com/amenaza/>.
- [16] R. & T. J. Arun, “Analysis and Enhancement Of Process Models Using Scoring For Customer Relationship Management.,” *International Journal of Computer Science Engineering and Information Technology Research.*, pp. 6(5), 75–84. Retrieved from [http://www.tjprc.org/publishpapers/2-14-1476361205-9.IJCSEITR-Analysis and Enhancement of process model using scoring for Customer Relationship Management.pdf](http://www.tjprc.org/publishpapers/2-14-1476361205-9.IJCSEITR-Analysis%20and%20enhancement%20of%20process%20model%20using%20scoring%20for%20customer%20relationship%20management.pdf), 2016.
- [17] U. & Q. H. Shafique, “A Comparative Study of Data Mining Process Models ( KDD , CRISP-DM and SEMMA ).,” *International Journal of Innovation and Scientific Research*, pp. 12(1), 217–222. Retrieved from <http://www.ijisr.issr-journals.org/abstract.php?>, 2014.
- [18] G. Quiceno, L. Uribe y P. Noreña Cardona, “Visualizaciones analíticas para la toma de decisiones en pequeñas y medianas empresas utilizando Data Mining”, *Cuaderno Activa*, vol. 8, n.º 8, pp. 31–39, 2016.
- [19] A. Rembert y C. Ellis, “An Initial Approach to Mining Multiple Perspectives of a Business Process,” in *TAPIA '09: The Fifth Richard Tapia Celebration of Diversity in Computing Conference: Intellect, Initiatives, Insight, and Innovations*, pp. 35–40, 2010.
- [20] “ISOTools”. [internet]. Disponible en: <https://www.isotools.org/2015/02/20/en-que-consiste-el-ciclo-phva-de-mejora-continua/>
- [21] E. Cardoso, “Towards a Methodology for Goal-Oriented Enterprise Management,” en *17th IEEE International Enterprise Distributed Object Computing Conference Workshops*, 2013.
- [22] X. Gao, “Towards the Next Generation Intelligent BPM”, *Business Process Management. Lecture Notes in Computer Science*, vol. 8094, pp. 4-10, 2013.
- [23] M. A. Cetina, “Gestión de procesos con BPM,” *Tecnología, Investigación y Academia. Universidad Distrital Francisco José de Caldas*, vol. 4, n.º 2, pp.45-56, 2016.
- [24] “Sinnexus Business Intelligence Informática estratégica”. [internet]. Disponible en [https://www.sinnexus.com/business\\_intelligence/datamining.aspx](https://www.sinnexus.com/business_intelligence/datamining.aspx).
- [25] E. Blanco-Hermida, *Algoritmos de clustering y aprendizaje automático aplicados a Twitter*, Barcelona: Universidad Politécnica de Valencia, 2016.
- [26] K. R. Saylam y S. Ozgur, “Process mining in business process management: concepts and challenges,” en *2013 International Conference on Electronics, Computer and Computation (ICECCO)*, pp. 131-134, 2013.
- [27] E. Leonard Brizuela y Y. Castro Blanco, “Metodologías para desarrollar Almacén de Datos” *Revista de Arquitectura e Ingeniería, Revista de Arquitectura e Ingeniería*, vol. 7, n.º 3, pp. 1–12.

- [28] U. Shafique y H. Qaiser, “A Comparative Study of Data Mining Process Models ( KDD , CRISP-DM and SEMMA )”, *International Journal of Innovation and Scientific Research.*, vol. 12, n.º 1, 217–222.
- [29] ICONTEC, *Norma Técnica Colombiana NTC-31000. Gestión del Riesgo. Principios y Directrices*, Bogotá: Icontec, 2011.
- [30] T. Pidun y C. Felden, “On the Restriction to Numeric Indicators in Performance Measurement Systems”, 2011 IEEE 15th International Enterprise Distributed Object Computing Conference Workshops, Helsinki, 2011, pp. 96-102. DOI: 10.1109/EDOCW.2011.16.
- [31] D. Wegener y S. Rüping, “On Reusing Data Mining in Business Process – A Pattern- Based Approach”, en *International Conference on Business Process Management*, Berlín, 2011, pp. 264-276.

# Proceso y evolución de los métodos formales en la ingeniería de requisitos\*

Jorge Hernán Suaza Jiménez\*\*

Gloria Amparo Lora Patiño\*\*\*

Recibido: 01/07/2019 • Aceptado: 26/11/2019

<https://doi.org/10.22395/rium.v19n37a6>

## Resumen

La ingeniería de requisitos se considera la fase más importante del ciclo de vida de los productos de *software*, porque en ella se especifican las necesidades de los clientes y es la base para la ejecución de las demás fases de la ingeniería de *software*. Los modelos que actualmente se utilizan para realizar la elicitación de requisitos se han propuesto y documentado de forma amplia, pero se centran solo en las técnicas para coleccionar información y descuidan la actividad de documentarla de manera adecuada. Además, para estructurar la especificación de requisitos sigue en uso el lenguaje natural como forma de comunicación y comprensión con el cliente. Debido a las ambigüedades que este causa, se dificulta la interpretación; esto conlleva reprocesos en las etapas posteriores del ciclo de vida del *software*. De acuerdo con ello, es necesario que las organizaciones desarrolladoras de *software* consideren formalizar el proceso de elicitación de requisitos si desean hacer más eficiente su proceso de desarrollo. En este artículo se hace una revisión de literatura para determinar el proceso y evolución de los métodos formales desde la perspectiva de la ingeniería de requisitos.

**Palabras clave:** métodos formales; ingeniería de requisitos; calidad del software; lenguaje matemático.

---

\* Artículo derivado del proyecto de investigación “Diseño de un modelo semiformal para documentar la elicitación de requisitos”, identificado con el código interno P14117 del Instituto Tecnológico Metropolitano, Medellín. Año de inicio: 2014. Año de finalización: 2015.

\*\* Magister en ingeniería de *software* e ingeniero informático. Profesor investigador del Instituto Tecnológico Metropolitano, Facultad de Ingeniería, Grupo de Investigación Automática, Electrónica y Ciencias Computacionales. Correo electrónico: jorgesuaza@itm.edu.co. Orcid: <https://orcid.org/0000-0002-9669-8145>.

\*\*\* Magister en seguridad informática. Instituto Tecnológico Metropolitano. Correo electrónico: glorialora231287@correo.itm.edu.co. Orcid: <http://orcid.org/0000-0002-0635-7305>.

## Process and evolution of formal methods in requirements engineering

### **Abstract**

Requirements Engineering is considered the most important phase of the life cycle of software products because it specifies the needs of the customers, and it is also the basis for the execution of the other phases of software engineering. The models currently used to perform the requirements elicitation have been proposed and widely documented, but they are focused only on the techniques to collect information, disregarding the activity of properly documenting this information. Moreover, to structure the requirements specification, natural language continues to be used as a means of communication and understanding with the customer. Due to the ambiguities caused by this language, its interpretation becomes difficult, and this leads to reprocesses in the later stages of the software life cycle. According to the above, it is necessary for software development organizations to consider formalizing the process of requirements elicitation if they wish to make their development process more efficient. A literature review is carried out in this paper to determine the process and evolution of the formal methods from the requirements engineering perspective.

**Keywords:** Formal methods; requirements engineering; software quality; mathematic language.

## Processo e evolução dos métodos formais na engenharia de requisitos

### **Resumo**

A engenharia de requisitos é considerada a fase mais importante do ciclo de vida dos produtos de software, porque nela se especificam as necessidades dos clientes e é a base para a execução das demais fases da engenharia de *software*. Os modelos que atualmente são utilizados para realizar a elicitação de requisitos foram propostos e documentados de forma ampla, mas se centram só nas técnicas para coletar informação e descuidam a atividade de documentar de maneira adequada. Além disso, para estruturar a especificação de requisitos segue em uso a linguagem natural como forma de comunicação e entendimento com o cliente. Devido às ambiguidades que este causa, dificulta-se a interpretação; isso implica reprocessos nas etapas posteriores do ciclo de vida do software. De acordo com isso, é necessário que as organizações desenvolvedoras de *software* considerem formalizar o processo de elicitação de requisitos se desejam tornar mais eficiente seu processo de desenvolvimento. Neste artigo, faz-se uma revisão de literatura para determinar o processo e a evolução dos métodos formais a partir da perspectiva da engenharia de requisitos.

**Palavras-chave:** métodos formais; engenharia de requisitos; qualidade do *software*; linguagem matemática.

## INTRODUCCIÓN

El *software* se diseña y desarrolla a través de una serie de etapas conocidas como *ciclo de vida*, en el que se estructura una solución a las necesidades del cliente. Es una labor de trabajo en equipo entre los ingenieros de *software* y los interesados en el producto, y la mayor parte de los problemas en el desarrollo se relaciona con la ingeniería de requisitos —de modo específico, con la educación—. Si las necesidades del cliente no se elicitán correctamente, el desarrollo presentará problemas porque el producto no responderá a las necesidades de las partes, aunque se disponga de buenos lenguajes de especificación. Aquí radica la importancia de la educación de requisitos: en ella se genera la información necesaria para la especificación, que a su vez es la base para el diseño y el desarrollo de la solución.

Los modelos que se utilizan hoy para realizar la educación de requisitos se centran solo en las técnicas para acopiar información, pero descuidan la actividad de documentarla de forma adecuada. Además, persisten en usar, en primer orden, el lenguaje natural como forma de comunicación y comprensión con el cliente. Debido a las ambigüedades causadas por este último, la interpretación se dificulta, hecho que acarrea un incremento en costos y tiempos de entrega, además de procesos de reingeniería en el diseño y la arquitectura [1].

Desde que en la década de 1960 se promulgó la llamada “crisis del *software*” debido a los errores encontrados en la calidad, fiabilidad y seguridad de los productos de *software*, los investigadores y la industria han intentado solucionarla, hasta ahora sin éxito. Una manera de evitar estos inconvenientes radica en la formalización de requisitos: según Gore y Diallo [2], este es el proceso de describir un sistema y sus propiedades deseadas utilizando un lenguaje cuyas sintaxis y semántica se definan en términos matemáticos.

El concepto de *métodos formales* alude a técnicas y herramientas basadas en principios y postulados matemáticos que se utilizan, entre otros, para especificar, diseñar, validar y verificar sistemas de *software* y *hardware*. La especificación utilizada en los métodos formales está conformada por enunciados bien formados en una lógica matemática. La verificación formal por deducciones rigurosas sigue la misma lógica; es decir, cada paso sigue una regla de inferencia y, por lo tanto, se puede comprobar mediante un proceso mecánico [3]. Los métodos formales surgieron como puntos de vista analíticos con los que es posible verificar el desarrollo de sistemas mediante la lógica y las matemáticas, lo que aporta ventajas para mejorar la calidad de los programas y, por tanto, contribuye a la ingeniería de *software*. En la fase de la ingeniería de requisitos, especificar en términos formales es importante y requiere mayor cuidado, porque su objeto es garantizar que tanto el funcionamiento como el

desempeño del programa sean correctos bajo cualquier situación. En la educación se utilizan fórmulas matemáticas para documentar los requisitos, de tal forma que sean comprendidos por las partes involucradas y que en el diseño se armonicen de mejor forma en la estructura de la solución [4].

La importancia de los métodos formales yace en sistematizar e introducir rigor en todas las fases de desarrollo del *software*, en especial en la Ingeniería de Requisitos, con lo que sería posible evitar que se pasen por alto problemas críticos. Además, proporcionan un método estándar de trabajo a lo largo del proyecto, constituyen una base coherente entre las muchas actividades relacionadas y, al contar con mecanismos de descripción precisos y no ambiguos, proporcionan el conocimiento necesario para realizarlas con éxito [5].

Este artículo es el resultado de una revisión de literatura para determinar el proceso y evolución de los métodos formales en la elicitación de requisitos. En este sentido, se llevó a cabo una labor de consulta en diferentes bases de datos y bibliotecas digitales, encaminada a encontrar trabajos representativos que aportaran a la investigación.

## 1. MARCO TEÓRICO

Para Mathiassen y Munk-Madsen [6], las formalizaciones están relacionadas con los tipos de expresión y comportamiento del *software*. Estos autores discuten los límites a la aplicación de las formalizaciones en estos dos sentidos y los ilustran con ejemplos en el desarrollo de sistemas prácticos. También establecen que las formalizaciones son valiosas en algunas situaciones e insuficientes en otras, y plantean que la alternativa a la utilización acrítica de las formalizaciones es que los desarrolladores analicen las situaciones en las que se encuentran y, a partir de allí, elijan una combinación entre lo formal y lo informal hasta que sean capaces de comprenderlas y dominarlas.

Vilkomir *et al.* [7] afirman que los requisitos reglamentarios, a diferencia de aquellos para un sistema en particular, tienen un carácter genérico, son aplicables a una amplia gama de sistemas, y son la base para la certificación o el proceso de concesión de licencias. La formalización de los requisitos está resolviendo las inconsistencias entre ellos y la interpretación de los desarrolladores de sistemas críticos. Estos autores proponen un enfoque para hacerlo, que incluye la formalización de los requisitos reglamentarios como base para el desarrollo de métodos de evaluación del *software*. Ilustran su enfoque con ejemplos de requisitos formalizados para la protección de sistemas de control contra el acceso no autorizado y los fallos de *software* en modo común, usando la notación  $Z$  para la formalización.

Por su parte, Webel y Gotzhein [8] establecen que la necesidad de la formalización de los requisitos se debe al hecho de que se necesita una descripción precisa de ellos,

así como una amplia comprensión entre el usuario y el proveedor de los servicios de calidad. Para esto, los mecanismos que realizan estos procesos requieren descripciones más precisas y un alcance real de la formalización. Estos mecanismos suelen estar integrados en capas, por lo que se necesita más de un punto de vista sobre los requisitos. Actualmente, esta relación rigurosa de los puntos de vista no se está logrando y los métodos formales necesarios aún están por definirse. Se espera que en el futuro se involucren mejor y en mayor grado en los procesos de la ingeniería de requisitos, y que los desarrolladores tengan una mejor comprensión de la formalización.

La especificación formal es el proceso de describir un sistema y sus propiedades deseadas mediante un lenguaje con una sintaxis y una semántica definidas matemáticamente [9]. Por eso, el objetivo de Chatterjee y Johari [10] tiene que ver con la formalización automatizada de la especificación de requisitos, de forma que se puedan generar con facilidad los casos de prueba formales.

Cavada *et al.* [11] llegan a la conclusión que la ingeniería de requisitos es una de las fases más importantes del proceso de desarrollo de *software*, y que en aplicaciones críticas es importante apoyar la validación de los requisitos con técnicas formales para identificar y eliminar los defectos. Sin embargo, los requisitos se escriben a menudo en documentos textuales, y su formalización y validación no son adecuadas por la falta de conocimiento sobre los métodos formales.

Post y Hoenicke [12] evalúan una cadena de herramientas para analizar requisitos de tiempo real de modo algorítmico. De acuerdo con esta cadena de herramientas, los autores formalizan los requisitos de un sistema en lenguaje natural. Ellos experimentan la compilación automática desde fórmulas en una lógica en tiempo real. Las fórmulas se pueden comprobar de forma automática para las propiedades cuya violación indica un error en la especificación de requisitos. Presentan un estudio de viabilidad en el contexto de varios proyectos en la herramienta *Bosch*. Los resultados del estudio indican que el esfuerzo para formalizar los requisitos de tiempo real es aceptable, los algoritmos de análisis son computacionalmente factibles; y el beneficio (en cuanto a detección de errores de especificación) parece significativo.

Li *et al.* [13] concluyen que los requisitos son informales, vagos, ambiguos y, muchas veces, inalcanzables, por lo que el problema de la ingeniería de requisitos es formalizarlos y luego transformarlos a través de un proceso sistemático en una especificación formal que pueda entregarse a los diseñadores para el desarrollo. En su trabajo proponen un marco para la transformación de los requisitos informales a formales, y luego a una especificación. El marco consiste en una ontología de requisitos, un lenguaje formal de modelado de requisitos para la representación de los funcionales y no funcionales, así como un amplio conjunto de operadores de refinamiento por el

cual los requisitos se transforman de forma gradual en una especificación formal y medible. Esta propuesta incluye una metodología sistemática y herramientas de apoyo para realizar la transformación. Los resultados sugieren que la ontología y el lenguaje de modelado son adecuados para la captura de requisitos, y que la metodología es eficaz en el manejo de los requisitos en la práctica.

Al utilizar notaciones y lenguajes formales es posible realizar transformaciones de modelos automáticos [14]. A diferencia de un modelo no formal, sus notaciones son intuitivas: permite una mejor abstracción de los detalles y aplica metodologías estandarizadas y bien definidas [15], lo que permite estructurar con claridad los requisitos del sistema y generar la formalización de requisitos.

Los requisitos se consideran la parte fundamental sobre la que se pueden desarrollar las soluciones de *software*; y en la actualidad existe una creciente demanda de enfoques más rigurosos y sistemáticos para formalizarlos [16]. Serna y Serna [17] argumentan que los métodos formales se utilizan hoy para modelar complejos sistemas críticos de seguridad, pero poco se trabaja en la formalización de los requisitos desde las primeras etapas de la ingeniería de requisitos.

La representación más utilizada para formalizar los requisitos son los lenguajes de especificación formal [16]. Con este tipo de especificación, los desarrolladores comprenden mejor el sistema y eluden las ambigüedades, los flujos, las omisiones y las inconsistencias de los requisitos. Además, la especificación es un importante mecanismo de comunicación entre los clientes y los diseñadores; entre los diseñadores y ejecutores; y entre los ejecutores y los probadores [18]. Pero las especificaciones formales no son documentos que se escriben una vez, y por lo general no se logran al principio del proceso de desarrollo de *software*. Se necesita tiempo para crear una primera versión de utilidad que, luego de mucho esfuerzo y revisiones, permita desarrollar una especificación cercana a las necesidades que los clientes tienen en mente [19]; al respecto, Bollin y Rauner [20] sostienen que un buen nivel de comprensión es una cuestión clave como atributo de calidad. Para Wolf [21], si se hace uso de un lenguaje de especificación formal, el sistema puede ser descrito con precisión en cuanto a funcionalidad, concurrencia, integridad y exactitud. Esto significa que las propiedades de un sistema se pueden analizar sin tener que ejecutarlo.

## 1.1 Trabajos relacionados

A continuación se detallan los 41 trabajos que compusieron la muestra final de la revisión.

Bhavsar [22] presenta un uso de los métodos formales en la formalización de requisitos centrados en el usuario. De acuerdo con este autor, los resultados son



prometedores en cuanto al mejoramiento de la calidad del producto de *software*. Por su parte, Serna [23] hace un recorrido histórico por los métodos formales y describe sus aplicaciones y beneficios en la formalización de requisitos. Propone que la investigación en esta área se debe orientar a aplicaciones prácticas y a buscar reducir los costos de su utilización.

Bishop *et al.* [24] integran los conceptos de los métodos formales desde el inicio de la programación y, mediante una didáctica simple, muestran cómo hacer una especificación formal y qué modelos se pueden utilizar. De Sousa *et al.* [25] analizan el método *B* para la especificación de requisitos y describen los casos de uso y las propiedades de seguridad. Aunque su aplicación es sencilla, es un aporte hacia la masificación de los métodos formales en la especificación de requisitos.

Gao y Wang [26] proponen una metodología para formalizar la ingeniería de requisitos, mediante el diseño de un modelo basado en lenguaje *Z*, orientado a analizar y gestionar requisitos (MZASR): describen la especificación de requisitos de *software* por herramientas y modelos UML, y la semántica individual se establece como especificación *Z* mediante un enfoque ascendente. Por su parte, Hassan *et al.* [27] plantean el análisis y el diseño formal de la seguridad de los requisitos mediante especificaciones en *B*. Los resultados demuestran efectividad y capacidad en la producción de *software* seguro. Van der Poll [28], a su turno, describe los métodos formales más comunes y la importancia de aplicar la especificación formal en los primeros pasos del ciclo de vida; enfrenta las posiciones de los críticos y defensores de los métodos formales y presenta ejemplos de historias de éxito en la industria. Singh y Patterh [29], a su vez, utilizan la notación *Z* para formalizar los requisitos en busca de protección y seguridad contra amenazas internas: un caso de estudio de éxito.

Cimatti *et al.* [30] desarrollan una metodología y técnicas para formalizar y validar requisitos en aplicaciones críticas de seguridad, y aplican el resultado en un caso de estudio con resultados mejorados. En cambio, Serna [5] describe la aplicación de la especificación formal de requisitos y el mejoramiento que se logra en la calidad del producto final. Este autor propone la incorporación de los métodos formales en los contenidos curriculares de los programas de ciencias computacionales.

Barringer *et al.* [31] presentan un método práctico para usar especificaciones formales en pruebas de registro en el Mars Science Laboratory. Demuestran, así, las ventajas de utilizar un lenguaje de especificación formal. Por otro lado, Fernández *et al.* [32] proponen una revisión a la aplicación de los métodos formales en proyectos reales en la industria: presentan los diversos casos de éxito y describen la situación actual en México en cuanto a la aplicación de métodos formales.

Ibrahim *et al.* [33] aplican una base formal para la especificación y verificación de requisitos, dentro de lo cual presentan una definición formal y una teoría de composición formal para la Ingeniería de Requisitos y describe los resultados. Además de esto, Yan [19] describe la aplicación de la especificación formal en *Communications-Based Train Control* (CBTC), con lo que obtiene una descripción profunda y exacta del sistema, y lenguaje Z en el desarrollo. El autor concluye que al usar métodos formales se gana confianza para asegurar el sistema. Incluso Ammar y Abdallah [34], presentan un nuevo enfoque para la especificación y verificación formal basado en la reescritura de la lógica; luego, lo aplican en un caso de estudio con resultados prometedores en cuanto a mejoramiento de la calidad. Por su parte, Kaur *et al.* [35] presentan un estudio comparativo entre los métodos de especificación formal y descripción de las razones de por qué usarlos en los procesos industriales.

Sumado este trabajo a lo anterior, You *et al.* [36] muestran los logros y problemas de los métodos formales y validan su aplicación para garantizar la seguridad y estabilidad del sistema. A su turno, Barlas *et al.* [37] presentan una comparación entre la especificación de requisitos tradicional y la formal; con la segunda se alcanza mayor comprensión, se eliminan ambigüedades y se obliga a una mejor precisión de la especificación.

Al contrario de Barlas *et al.* [37], Wolff [21] utiliza la experiencia en la industria para combinar el desarrollo ágil *Scrum* con los métodos formales, pero desde una conceptualización teórica; presenta una evaluación y discusión acerca de los beneficios de utilizarlos para formalizar los requisitos. Bollin [38], desde un marco teórico, concluye que el proceso de la formalización de los requisitos es un problema de adaptación transcultural: analiza sus pros y contras, y da a conocer un modelo refinado para un proceso de desarrollo de *software* formal. Desde la experiencia, Serna y Serna [39] presentan una investigación teórica en la que se concluye que el método formal ha alcanzado significativos avances, lo que vislumbra aplicaciones ambiciosas en el futuro para apoyar rigurosos y coherentes planes de estudios en ciencias computacionales.

Atlee *et al.* [40] ofrecen una serie de recomendaciones para formalizar los requisitos, aportan al logro práctico de la selección de una técnica específica. Su objetivo es mejorar la aceptación de los métodos formales en el modelado de requisitos.

Por su parte, Serna y Serna [41] sostienen que la especificación formal todavía tiene usos limitados; la comunidad tiene una comprensión diferente acerca de su utilidad y necesidad. Estos autores sostienen que los investigadores se focalizan en la especificación escrita durante el diseño del modelo funcional y dejan a un lado cualquier tratamiento formal a los no funcionales. Chen y Ouyang [42], a su turno, analizan el dominio de los métodos formales para identificar defectos y disminuir fracasos. Aplican

la formalización en la especificación, el modelado y la verificación. De igual manera, Chan *et al.* [43] proponen una metodología de modelado de requisitos para convertir los informales en formales; con ella se busca eliminar ambigüedades y defectos, refinar y perfeccionar el sistema, y componer un medio de comunicación entre las partes interesadas. Gore y Diallo [2], por su parte, hacen un estudio de los lenguajes de especificación formal; describen los modelos que se pueden aplicar y realizan un análisis general, y realizan una experimentación con estadística de depuración en un estudio de caso.

Noaman *et al.* [44] aplican la especificación formal por medio de  $Z$  para potenciar la seguridad del sistema y reducir las amenazas. Los resultados son prometedores para ampliaciones y elaboraciones futuras. Lockhart *et al.* [45] demuestran la utilidad de los métodos formales para la especificación de requisitos en sistemas de seguridad críticos. La especificación formal reduce la ambigüedad del diseño, ayuda a probar la consistencia, e incrementa la confianza y el rendimiento. Por su parte, Wu *et al.* [46] proponen reglas de transformación desde el modelo  $B$  y describen los requisitos con máquinas abstractas. Para los autores, la especificación formal es un proceso intermedio entre la especificación de requisitos y la escritura del código.

Serna y Serna [17] realizan una revisión de la literatura, hacen un recorrido por la esencia, la función, el uso y los inconvenientes de las técnicas de especificación formal, y analizan criterios de valoración y evaluación a sus debilidades. De igual manera, Bollin y Rauner-Reithmayer [20] muestran una serie de recomendaciones para realizar especificaciones formales, e identifican que la facilidad de lectura de una especificación es un factor clave para mejorar la calidad del *software*.

Schraps y Peters [9] aplican gramática formal con anotaciones semánticas para formalizar requisitos, de tal forma que puedan ser analizados y procesados por el computador. Los requisitos formulados pueden ser analizados y procesados en las primeras etapas de desarrollo.

Serna y Serna [47] afirman que muchos de los problemas desafiantes en la construcción de sistemas requieren apoyo formal para su modelado y análisis. También sostienen que en los procesos de investigación y aplicación de las ciencias computacionales existe un número creciente de aplicaciones, pero todavía no constituyen una parte integral de los procesos formativos en pregrado y posgrado.

Azeem *et al.* [48] utilizan especificación formal  $Z$  para mejorar la calidad y confianza del sistema: muestran trabajos relacionados y aplican una propuesta en un estudio de caso en salud. Esto lo confirman Li *et al.* [49], quienes aplican *Object-Z* para describir sistemas complejos y demuestran su utilidad en un caso de estudio de

un sistema de suministro de gasolina. Los resultados obtenidos alientan el uso de la especificación formal.

Serna y Serna [50] describen aspectos relevantes de los métodos formales y realizan un marco del futuro de la investigación en el área. Los resultados invitan a la reflexión sobre este componente de las ciencias computacionales y respecto a la necesidad de tener una comunidad más amplia y sólida.

Jeffery *et al.* [51] realizan un estudio empírico para encontrar la productividad de los proyectos que utilizan métodos formales, específicamente al formalizar los requisitos. Estos autores identifican una serie de preguntas sobre los resultados en productividad de estos proyectos. A su turno, Tamrakar y Sharma [52] comparan tres métodos para especificar formalmente: *Z*, *B* y *VDM*. Usar especificaciones formales no asegura un sistema completamente correcto, pero se aumenta la confianza en él. De igual manera, Singh y Yadav [53] utilizan el método *Event-B* para escribir especificaciones formales, con el fin de garantizar la comprensión de los límites en los que un algoritmo puede ser usado. Para los autores, la especificación, la validación y la verificación formales son la clave para obtener un mejor diseño. Walter *et al.* [54], por su parte, realizan una representación de formalización de requisitos por medio de diagramas *SysML* a través de la semántica *RSL*; formalizan únicamente los requisitos no funcionales y aplican ese diseño a un caso de estudio en el área automotriz.

## 2. ANÁLISIS DE RESULTADOS

Al utilizar la especificación formal, las propiedades del sistema se describen mediante un lenguaje con una sintaxis y una semántica definidas en términos matemáticos [7]. Algunos lenguajes formales, tales como *Z*, *B* y *VDM*, se centran en especificar el comportamiento de los sistemas secuenciales, y en ellos los estados se describen en estructuras matemáticas como conjuntos, relaciones y funciones [36]; mientras que métodos como *CSP*, *CCS*, *Statecharts*, lógica temporal y autómatas se centran en la especificación de los comportamientos del sistema en términos de secuencias, árboles u órdenes parciales de eventos [9]. El uso de lenguajes formales como *Z* y *B* puede mejorar la confianza del usuario en el sistema y los impactos se esta última en su uso [48].

*Z* trabaja en alta abstracción a nivel del sistema y proporciona una sólida base para su diseño. Con el uso de este lenguaje se descubren errores en la especificación y en las fases de pruebas y mantenimiento. Esto es conveniente porque, con la ingeniería tradicional, corregir errores en etapas posteriores incrementa los costos [48]; además, *Z* es una manera de descomponer una especificación en pequeñas partes, llamadas *esquemas*. Por su parte, *B* es uno de los métodos formales más conocidos. Se basa en la lógica de primer orden, la teoría de conjuntos, la aritmética de enteros y las sustituciones

generalizadas; se utiliza para el diseño de *software* desde los requisitos funcionales y permite producir casos de prueba que demuestran la exactitud y la consistencia del modelo *software* aplicado. El objetivo de *B* es obtener un producto probado y fiable [25]; mientras que *VDM* se utiliza para demostrar la equivalencia de los conceptos del lenguaje de programación [52].

De acuerdo con Pandey y Batra [16], las empresas se enfrentan en la actual era digital al desafío de liberar proyectos de *software* de calidad a tiempo y ajustados al presupuesto; pero la realidad es que se entrega con errores, falta de funcionalidad y, a veces, con sobrecostos. Estos ocurren debido a errores en la especificación de requisitos, cuya corrección puede acarrear fuertes cargas en tiempo y dinero cuando se detectan en las fases tardías del ciclo de vida. Por lo tanto, en la especificación de requisitos se debe seleccionar una metodología formal para abordar la fiabilidad durante la ingeniería de requisitos y el diseño, porque los métodos formales permiten desarrollar los errores antes de la liberación del producto [45].

Los errores de especificación pueden ser reducidos drásticamente mediante el uso de métodos formales y, en consecuencia, el ingeniero de *software* puede crear una especificación más completa, coherente e inequívoca que con los métodos convencionales [35] [48]. Por su parte, Bollin y Rauner-Reithmayer [20] manifiestan que una buena especificación formal es sintáctica y semánticamente correcta, y permite un mapeo sin pérdidas entre todos los conceptos de la especificación y el modelo mental del sistema especificado; también agregan que debe ser completa, coherente y adecuada, y tener en cuenta que la facilidad de comprensión es un requisito esencial para decidir sobre su corrección semántica.

En resumen, para conocer qué se publica en la literatura en relación con la formalización de requisitos, en la tabla 1 se presenta el resumen de los resultados de esta investigación. El inicio de la línea de tiempo de observación se estableció en 2010. Aunque es conocido que la mayor parte del trabajo en esta área se presentó en la segunda mitad del siglo pasado, el objetivo de esta investigación fue verificar su progreso actual. Los trabajos de la muestra final se clasificaron de acuerdo con el enfoque primario presentado. Se aclara aquí que muchos autores presentan resultados en más de una tipología, es decir, su investigación puede ser teórica y, al mismo tiempo, experimental.

- *Investigación teórica*: artículos que contienen definiciones o descripciones acerca de la formalización de requisitos.
- *Investigación experimental*: demuestra resultados a partir de experimentos en laboratorio.

- *Trabajos de aplicación práctica*: describen y aplican métodos, técnicas o procedimientos de formalización de requisitos en casos de estudio.
- *Trabajos de aplicación industrial*: aquellos en los que la formalización de requisitos se aplica en casos reales.

Tabla 1. Formalización de requisitos según la literatura

Medio	Aplicación			
	Teórica	Experimental	Práctica	Industrial
Revista	66 %	63 %	20 %	23 %
Conferencia	34 %	34 %	20 %	23 %

Fuente: elaboración propia.

### 3. PROGRESO DEL TRABAJO EN LA FORMALIZACIÓN DE REQUISITOS

El progreso de la formalización de requisitos será definido como *acelerado o lento* en función del porcentaje de aplicaciones en la industria. En la figura 1 se clasifican los resultados de la investigación y se muestra su participación porcentual.

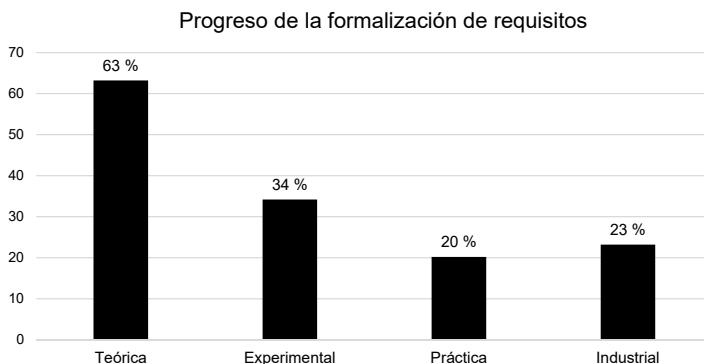


Figura 1. Clasificación y participación porcentual de los resultados de la investigación

Fuente: elaboración propia.

De acuerdo con la revisión de literatura en esta investigación, el trabajo en formalización de requisitos ha tenido un progreso lento. En la figura 1 se muestra que la mayoría de los trabajos tomados en cuenta para la presente investigación se enfocan más en definiciones teóricas, que son representadas por el 63 %, y solo el 23 % de estos en aplicación en la industria. Algunos autores han analizado este hecho y presentan sus conclusiones. Entre las causas se destacan falta de formación en la academia; falta de aceptación en la industria; ausencia de demostraciones realmente importantes de sus ventajas; costos; falta de personal capacitado; fobia de los ingenieros a las matemáticas; ausencia de deseo en la industria por experimentar con principios

que no tienen comprobación suficiente; y pérdida de interés y patrocinio respecto de la investigación en métodos formales durante la última década [23] [55] [40] [17].

Los investigadores de mediados del siglo pasado estaban convencidos de que la única manera de mejorar la calidad de los productos de *software* radicaba en la matematización de la misma ingeniería de *software* [28]; empero, conforme sus resultados comenzaron a publicarse, la industria comenzó a encontrar y poner barreras al progreso de este trabajo. Solo el *software* crítico los arropó como tabla de salvación para solucionar sus problemas de fiabilidad y seguridad, y son diversos los casos de éxito en estos desarrollos. El problema surge cuando las empresas de desarrollo de *software* comercial se enfrentan a situaciones de incumplimiento, que incrementan los costos del proceso y las obligan a dedicar menos tiempo a la estructura misma de las actividades. La decisión es disminuir la mayor cantidad de tiempo posible y, como la forma tradicional de realizar la verificación y validación es una barrera de contención al final del ciclo de vida, esta fase se recorta para encontrar algún tiempo extra [36].

El estudio sobre la formalización de requisitos se ha relegado a unos pocos investigadores que convencen a sus estudiantes de posgrado para que trabajen con ellos. Visto el mencionado inconveniente de que los profesionales temen a todo lo que tenga que ver con matemáticas, se necesita con premura que la formalización empiece a permear los procesos formativos en la academia y que los estudiantes conozcan sus ventajas y potencialidades, lo mismo que sus desventajas y problemas, de tal forma que se interesen y profundicen en su asimilación y comprensión.

De acuerdo con los autores citados en esta investigación, la formalización de requisitos se investiga de forma mayoritariamente teórica y de esta manera no se logrará convencer a la industria de su aplicación. Puede ser que la etapa de la teorización ya haya sido superada en los muchos trabajos que se presentaron el siglo pasado. La necesidad actual es masificar los métodos formales y encontrar la forma de aplicarlos en la matematización de la ingeniería de *software*, pero con costos que estén al alcance de las empresas que lo desarrollen. De esta manera será posible reavivar el interés en esta área de trabajo y encontrar el patrocinio necesario para ejecutar la investigación necesaria. Si el objetivo es superar la crisis del *software*, los métodos formales deben ser el centro de desarrollo porque sus bondades ya han sido demostradas con suficiencia, y porque el lenguaje con el que funciona el computador es matemático.

#### 4. CONCLUSIONES

Los resultados demuestran que el progreso y desarrollo de la formalización de requisitos son lentos y que, en los últimos años, el interés por darle continuidad ha disminuido. Aunque se acepta que los métodos formales son una herramienta útil y necesaria para

superar la crisis de la calidad del *software*, muy pocos investigadores quieren abordar la formalización de requisitos.

Dentro de las investigaciones obtenidas se puede evidenciar la importancia de la aplicación de los métodos formales durante el desarrollo de un producto de *software*, ya que evita ambigüedades y permite obtener productos más precisos y seguros. No obstante, cabe resaltar que en los últimos 5 años la aplicación de la formalización de requisitos en la industria ha sido muy poca por factores como desconocimiento de la formalización de requisitos por parte de los desarrolladores; búsqueda de desarrollos de *software* ágiles o hechos en el menor tiempo posible; y falta de publicaciones por parte de las empresas que aplican la formalización de requisitos en sus desarrollos.

La industria necesita demostraciones de los beneficios que representa la formalización, pero hasta ahora no tiene inclinación directa por apoyar las iniciativas que en este sentido propone la comunidad. Solo en el desarrollo de sistemas críticos se aprecia una amplia participación de las empresas de desarrollo y la industria porque su objetivo es que funcionen sin poner en riesgo la vida humana o las inversiones económicas.

Se necesita mayor diligencia de los métodos formales en los procesos de ingeniería de requisitos, pero, sobre todo, es urgente contar con el personal capacitado para aplicarla y darle continuidad.

Esta es una sociedad dependiente del *software* porque muy pocas de sus actividades están por fuera del ámbito de este desarrollo tecnológico; empero, los productos que se liberan o entregan a los usuarios aún no resultan satisfactorios en aspectos como fiabilidad y seguridad. Las matemáticas ofrecen la posibilidad de superar este problema y su representación en los métodos formales es una alternativa prometedora.

## REFERENCIAS

- [1] Å. Dahlstedt, y A. Persson, “Requirements interdependencies - Moulding the state of research into a research agenda”, en *Ninth International Workshop on Requirements Engineering Foundation for Software Quality (REFSQ 2003)*, Klagenfurt, 2003.
- [2] R. Gore, y S. Diallo, “The need for usable formal methods in verification and validation”, en *2013 Winter Simulation Conference (WSC)*, Washington, 2013. DOI: <https://doi.org/10.1109/WSC.2013.6721513>.
- [3] C. Burgess, *The Role of Formal Methods in Software Engineering Education and Industry*, Bristol: University of Bristol, 1995.
- [4] P. Zave, “Classification of research efforts in requirements engineering”, *ACM Computing Surveys*, vol. 29, n.º 4, pp. 315-321, 1997. DOI: <https://doi.org/10.1109/isre.1995.512563>.



- [5] E. Serna, “Métodos formales e Ingeniería de Software”, *Revista Virtual Universidad Católica del Norte*, n.º 30, pp. 158-184, 2010, Disponible en: <http://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/viewFile/62/129>.
- [6] L. Mathiassen y A. Munk-Madsen, “Formalization in systems development”, en *International Joint Conference on Theory and Practice of Software Development on Formal Methods and Software Development*, Berlín, 1985. DOI: [https://10.1007/3-540-15199-0\\_7](https://10.1007/3-540-15199-0_7).
- [7] S. Vilkomir, J. Bowen y A. Ghose, “Formalization and assessment of regulatory requirements for safety-critical software”, *Innovations in Systems and Software Engineering*, vol. 2, n.º 3-4, pp. 165-178, 2006. DOI: <https://10.1007/s11334-006-0006-8>.
- [8] C. Webel y R. Gotzhein, “Formalization of Network Quality-of-Service Requirements”, *Lecture Notes in Computer Science*, vol. 4574, pp. 309-324, 2007. DOI: [https://10.1007/978-3-540-73196-2\\_20](https://10.1007/978-3-540-73196-2_20).
- [9] M. Schrap y M. Peters, “Semantic annotation of a formal grammar by Semantic Patterns”, en *2014 IEEE 4th International Workshop on Requirements Patterns (RePa)*, Karlskrona, 2014. DOI: <https://10.1109/rep.2014.6894838>.
- [10] R. Chatterjee y K. Johari, “A Simplified and Corroborative Approach towards Formalization of Requirements”, *Communications in Computer and Information Science*, vol. 94, pp. 486-496, 2010. DOI: [https://10.1007/978-3-642-14834-7\\_46](https://10.1007/978-3-642-14834-7_46).
- [11] R. Cavada, A. Cimatti, A. Micheli, M. Roveri, A. Susi y S. Tonetta, “OthelloPlay – A Plug-in based tool for requirement formalization and validation”, en *The 1st workshop on Developing tools as plug-ins – TOPI’11*, Honolulu, 2011. DOI: <https://10.1145/1984708.1984728>.
- [12] A. Post y J. Hoenicke, “Formalization and Analysis of Real-Time Requirements: A Feasibility Study at BOSCH”, *Lecture Notes in Computer Science*, vol. 7152, pp. 225-240, 2012. DOI: [https://10.1007/978-3-642-27705-4\\_18](https://10.1007/978-3-642-27705-4_18).
- [13] F. Li, J. Horkoff, A. Borgida, G. Guizzardi, L. Liu y J. Mylopoulos, “From Stakeholder Requirements to Formal Specifications Through Refinement”, *Requirements Engineering: Foundation for Software Quality*, vol. 9013, pp. 164-180, 2015. DOI: [https://10.1007/978-3-319-16101-3\\_11](https://10.1007/978-3-319-16101-3_11).
- [14] H. Ehrig, G. Engels, F. Orejas y M. Wirsing, “Semi-Formal and Formal Specification Techniques for Software Systems” [internet], 2000. Disponible en: <https://pdfs.semanticscholar.org/3323/1736efa0f613ba416094e1a59262b23f8a2d.pdf>.
- [15] C. Snook y M. Butler, “UML-B: Formal Modeling and Design Aided by UML”, *ACM Transactions on Software Engineering and Methodology (TOSEM)*, vol. 15, n.º 1, pp. 92-122, 2006.
- [16] S. Pandey y M. Batra, “Formal methods in requirements phase of SDLC”, *International Journal of Computer Applications (0975 – 8887)*, vol. 70, n.º 13, pp. 7-14, 2013.
- [17] E. Serna y A. Serna, “La especificación formal en contexto: actual y futuro”, *Ingeniare – Revista Chilena de Ingeniería*, vol. 22, n.º 2, pp. 243-256, 2014. DOI: <https://10.4067/S0718-33052014000200010>.

- [18] H. Krad, “Formal methods and automation for system verification”, en *2011 Fourth International Conference on Modeling, Simulation and Applied Optimization (ICMSAO 2011)*, Kuala Lumpur, 2011. DOI: <https://10.1109/icmsao.2011.5775479>.
- [19] F. Yan, “Studying Formal Methods Applications in CBTC”, en *2011 International Conference on Management and Service Science (MASS2011)*, Bangkok, 2011. DOI: <https://10.1109/icmss.2011.5999325>.
- [20] A. Bollin y D. Rauner-Reithmayer, “Formal specification comprehension: the art of reading and writing z”, en *The 2<sup>nd</sup> FME Workshop on Formal Methods in Software Engineering - FormaliSE 2014*, Hyderabad, 2014. DOI: <https://10.1145/2593489.2593491>.
- [21] S. Wolff, “Scrum goes formal: Agile methods for safety-critical systems”, en *2012 First International Workshop on Formal Methods in Software Engineering: Rigorous and Agile Approaches (FormSERA)*, Zurich, 2012. DOI: <https://10.1109/formsera.2012.6229784>.
- [22] M. Bhavsar, “Analysis of Multiagent Based Interactive Grid Using Formal Methods - A Reliable Approach”, en *2010 3<sup>rd</sup> International Conference on Emerging Trends in Engineering and Technology*, Goa, 2010. DOI: <https://10.1109/icetet.2010.5>.
- [23] E. Serna, “Métodos formales: Perspectiva y aplicación futura”, en *III Jornadas de Investigación de la Facultad de Ingenierías*, pp. 64-68, Medellín, 2011.
- [24] M. Bishop, B. Hay y K. Nance, “Applying Formal Methods Informally”, en *2011 44<sup>th</sup> Hawaii International Conference on System Sciences*, Kauai, 2011. DOI: <https://10.1109/hicss.2011.71>.
- [25] T. de Sousa, J. Almeida, S. Viana y J. Pavón, “Automatic analysis of requirements consistency with the B method”, *ACM SIGSOFT Software Engineering Notes*, vol. 35, n.º 2, pp. 1-4, 2010. DOI: <https://10.1145/1734103.1734114>.
- [26] H. Gao y S. Wang, “Based on Formal Methods in Trustable Software Requirements Engineering”, en *2010 International Conference on Internet Technology and Applications*, Wuhan, 2010. DOI: <https://10.1109/itapp.2010.5566647>.
- [27] R. Hassan, M. Eltoweissy, S. Bohner y S. El-Kassas, “Formal analysis and design for engineering security automated derivation of formal software security specifications from goal-oriented security requirements”, *IET Software*, vol. 4, n.º 2, pp. 149-160, 2010. DOI: <https://10.1049/iet-sen.2009.0059>.
- [28] J. Van der Poll, “Formal methods in software development: A road less travelled”, *South African Computer Journal*, vol. 45, pp. 165-178, 2010. DOI: <https://10.18489/sacj.v45i0.33>.
- [29] M. Singh y M. Patterh, “Formal Specification of Common Criteria Based Access Control Policy Model”, *International Journal of Network Security*, vol. 11, n.º 3, pp. 139-148, 2010.
- [30] A. Cimatti, M. Roveri, A. Susi y S. Tonetta, “Formalization and Validation of Safety-Critical Requirements”, *Electronic Proceedings in Theoretical Computer Science*, vol. 20, pp. 68-75, 2010. DOI: <https://10.4204/eptcs.20>.

- [31] H. Barringer, A. Groce, K. Havelund y M. Smith, “An Entry Point for Formal Methods: Specification and Analysis of Event Logs”, *Electronic Proceedings in Theoretical Computer Science*, vol. 20, pp. 16-21, 2010. DOI: <https://10.4204/eptcs.20.2>.
- [32] C. Fernández *et al.*, “Métodos formales aplicados en la industria del software”, *Temas de Ciencia y Tecnología*, vol. 5, n.º 43, pp. 3-12, 2011. Disponible en: [http://www.utm.mx/edi\\_antiores/temas43/1ENSAYO\\_43\\_1-R.pdf](http://www.utm.mx/edi_antiores/temas43/1ENSAYO_43_1-R.pdf).
- [33] N. Ibrahim, V. Alagar y M. Mohammad, “Specification and Verification of Context-dependent Services”, *Electronic Proceedings in Theoretical Computer Science*, vol. 61, pp. 17-33, 2011. DOI: <https://10.4204/eptcs.61.2>.
- [34] B. Ammar y K. Abdallah, “Towards the formal specification and verification of multi-agent-based systems”, *International Journal of Computer Science Issues*, vol. 8, n.º 4, pp. 200-210, 2011.
- [35] A. Kaur, S. Gulati y S. Singh, “A comparative study of two formal specification languages: Z-notation & B-method”, en *The Second International Conference on Computational Science, Engineering and Information Technology - CCSEIT '12*, Coimbatore, 2012. DOI: <https://10.1145/2393216.2393304>.
- [36] J. You, S. Xia y J. Li, “A survey on formal methods using in software development”, en *IET International Conference on Information Science and Control Engineering 2012 (ICISCE 2012)*, Shenzhen, 2012. DOI: <https://10.1049/cp.2012.2353>.
- [37] K. Barlas, G. Koletsos y P. Stefaneas, “Extending standards with formal methods: Open Document Architecture”, en *2012 International Symposium on Innovations in Intelligent Systems and Applications*, Trabzon, 2012. DOI: <https://10.1109/INISTA.2012.6246931>.
- [38] A. Bollin, “Do you speak Z? Formal methods under the perspective of a cross-cultural adaptation problem”, en *2013 1st FME Workshop on Formal Methods in Software Engineering (FormaliSE)*, San Francisco, 2013. DOI: <https://10.1109/formalise.2013.6612271>.
- [39] E. Serna y A. Serna, “Desafíos y Oportunidades de la Investigación en Métodos Formales”, en *XII Conferencia Iberoamericana en Sistemas, Cibernética e Informática*, Orlando, 2011.
- [40] J. Atlee, S. Beidu, N. Day, F. Faghieh y P. Shaker, “Recommendations for improving the usability of formal methods for product lines”, en *2013 1st FME Workshop on Formal Methods in Software Engineering (FormaliSE)*, San Francisco, 2013. DOI: <https://10.1109/formalise.2013.6612276>.
- [41] E. Serna y A. Serna, “Especificación formal - Presente y Futuro”, en *XV International Convention and Fair Informatica*, La Habana, 2013.
- [42] X. Chen y D. Ouyang, “Research on the Implementation of Internal Control in Enterprise Information System by Domain Analysis and Formal Methods: A Case Study of Sales Activities Internal Control Under Chinese Enterprise Environment”, en *2013 Fourth World Congress on Software Engineering (WCSE 2013)*, Hong Kong, 2013. DOI: <https://10.1109/wcse.2013.31>.

- [43] L. Chan, R. Hexel y L. Wen, “Rule-Based Behaviour Engineering: Integrated, Intuitive Formal Rule Modelling”, en *2013 22nd Australian Software Engineering Conference (ASWEC' 2013)*, Melbourne, 2013. DOI: <https://10.1109/aswec.2013.13>.
- [44] M. Noaman, I. Alsmadi y A. Jaradat, “The specifications of E-Commerce Secure System using Z language”, *The Research Bulletin of Jordan ACM*, vol. II, n.º III, pp. 127-131, 2013.
- [45] J. Lockhart, C. Purdy y P. Wilsey, “Formal methods for safety critical system specification”, en *2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS)*, College Station, 2014. DOI: <https://10.1109/mwscas.2014.6908387>.
- [46] T. Wu, Y. Dong y N. Hu, “Formal Specification and Transformation Method of System Requirements from B Method to AADL Model”, en *2014 IEEE 17th International Conference on Computational Science and Engineering (CSE 2014)*, Chengdu, 2014. DOI: <https://10.1109/cse.2014.299>.
- [47] E. Serna y A. Serna, “Los Métodos Formales en Contexto”, en *XIII Conferencia Iberoamericana en Sistemas, Cibernética e Informática (CISCI 2014)*, Orlando, 2014.
- [48] M. Azeem, M. Ahsan, N. Minhas y K. Noreen, “Specification of e-Health system using Z: A motivation to formal methods”, en *International Conference for Convergence for Technology - 2014*, Pune, 2014. DOI: <https://10.1109/i2ct.2014.7092123>.
- [49] Y. Li, X. Pan, T. Hu, S. Sung y H. Yuan, “Specifying Complex Systems in Object-Z: A Case Study of Petrol Supply Systems”, *Journal of Software*, vol. 9, n.º 7, pp. 1707-1717, 2014. DOI: <https://10.4304/jsw.9.7.1707-1717>.
- [50] E. Serna y A. Serna, “Perspectiva y Aplicación de los Métodos Formales”, en *XIII Conferencia Iberoamericana en Sistemas, Cibernética e Informática*, Orlando, 2014.
- [51] R. Jeffery, M. Staples, J. Andronick, G. Klein y T. Murray, “An empirical research agenda for understanding formal methods productivity”, *Information and Software Technology*, vol. 60, pp. 102-112, 2015. DOI: <https://10.1016/j.infsof.2014.11.005>.
- [52] S. Tamrakar y A. Sharma, “Comparative Study and Performance Evaluation of Formal Specification Language based on Z, B and VDM Tools”, *International Journal of Scientific & Engineering Research*, vol. 6, n.º 9, pp. 1540-1543, 2015.
- [53] A. Singh y D. Yadav, “Formal Specification and Verification of Total Order Broadcast through Destination Agreement Using Event-B”, *International Journal of Computer Science and Information Technology*, vol. 7, n.º 5, pp. 85-95, 2015. DOI: <https://10.5121/ijcsit.2015.7506>.
- [54] S. Walter, A. Rettberg y M. Kreutz, “Towards Formalized Model-Based Requirements for a Seamless Design Approach in Safety-Critical Systems Development”, en *2015 IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing Workshops*, Auckland, 2015. DOI: <https://10.1109/isorcw.2015.51>.
- [55] A. Serna, “Los Métodos Formales en la Industria”, *Revista Antioqueña de las Ciencias Computacionales y la Ingeniería de Software (RACCIS)*, vol. 2, n.º 2, pp. 44-51, 2012, Disponible en: <http://fundacioniai.org/raccis/v2n2/n3a7.pdf>.

# Definiendo un modelo de proceso de *software* para la práctica del *modding*\*

Andrés Felipe Ceballos\*\*

Wilson Libardo Pantoja Yépez\*\*\*

Julio Ariel Hurtado\*\*\*\*

Recibido: 01/07/2019 • Aceptado: 27/11/2019

<https://doi.org/10.22395/rium.v19n37a7>

## Resumen

El *modding* es una práctica de *software* difundida principalmente en la comunidad de videojuegos, en la que se involucran grupos interdisciplinarios de *modders* con la finalidad de crear, mejorar y distribuir extensiones, llamadas *mods*, para *software* que ha sido liberado previamente. Lo anterior se realiza con la finalidad de ampliar las características de productos base, lo cual trae consigo nuevas aplicaciones, compatibilidad y extensibilidad, entre otras mejoras. Sin embargo, diferentes problemáticas, similares a las presentes en el proceso de desarrollo de *software*, se han visto involucradas dentro de esta práctica: barreras en la comunicación, ausencia de planeación, resultados inesperados y repetición de trabajo, entre otras. Así, en busca de integrar aspectos de la ingeniería de procesos para brindar soporte a estos grupos y disminuir los problemas presentes en este tipo de proyectos, se propone en el presente escrito el proceso de *software* *Kross Modding Process*, que brinda guía a los *modders* a través de información recogida en estudios previos sobre esta área.

**Palabras clave:** *modding*; proceso de *software*; ingeniería de *software*; modelo.

---

\* Este trabajo forma parte del proyecto de investigación “Diseño de juegos pervasivos basados en experiencias de aprendizaje sensible al contexto” - Dispersa de las universidades del Cauca y de Granada, financiado parcialmente por el Ministerio de Economía y Competitividad de España y la Universidad del Cauca.

\*\* Ingeniero de sistemas. Universidad del Cauca. Correo electrónico: [aceballos@unicauca.edu.co](mailto:aceballos@unicauca.edu.co). Orcid: <https://orcid.org/0000-0002-5849-7924>

\*\*\* Ingeniero de sistemas y magíster en computación. Profesor titular, Universidad del Cauca. Correo electrónico: [wpantoja@unicauca.edu.co](mailto:wpantoja@unicauca.edu.co). Orcid: <https://orcid.org/0000-0002-7235-6036>.

\*\*\*\* Ingeniero en electrónica y telecomunicaciones, y doctor en ciencias de la Computación (Universidad de Chile). Profesor titular, Universidad del Cauca. Correo electrónico: [jhurtado@dcc.uchile.cl](mailto:jhurtado@dcc.uchile.cl). Orcid: <https://orcid.org/0000-0002-2508-0962>.

## Defining a Software Process Model for Modding Practice

### **Abstract**

Modding is a software practice, mainly spread in the videogame community, in which interdisciplinary groups of modders are involved with the aim of creating, improving and distributing extensions, called mods, for previously released software. This is done with the purpose of extending the characteristics of base products; thus, bringing new applications, compatibility, and extensibility, among other improvements. However, various problems, similar to those present in the software development process, have been involved in this practice: Communication barriers, lack of planning, unexpected results, and work repetition, among others. Thus, seeking to integrate aspects of process engineering to support these groups and reduce the problems present in this type of project, the software process “Kross Modding Process” is proposed, which provides guidance to modders through information collected in previous studies on this area.

**Keywords:** Modding; software process; software engineering; model.

## Definindo um modelo de processo de *software* para a prática do modding

### **Resumo**

O modding é uma prática de *software* difundida principalmente na comunidade de videogames, no qual se envolvem grupos interdisciplinares de *modders* com a finalidade de criar, melhorar e distribuir extensões chamadas *mods*, para software que foi lançado previamente. Isso se realiza com a finalidade de ampliar as características de produtos base, o qual traz consigo novas aplicações, compatibilidade e extensibilidade, entre outras melhoras. No entanto, diferentes problemáticas, similares às presentes no processo de desenvolvimento de *software*, viram-se envolvidas dentro dessa prática: barreiras na comunicação, ausência de planejamento, resultados inesperados e repetição de trabalho, entre outras. Desse modo, em busca de integrar aspectos da engenharia de processos para dar suporte a esses grupos e diminuir os problemas presentes nesse tipo de projetos, propõe-se no presente escrito o processo de software *Kross Modding Process*, que oferece guia aos *modders* por meio de informação recolhida em estudos prévios sobre essa área.

**Palavras-chave:** modding; processo de software; engenharia de *software*; modelo.

## INTRODUCCIÓN

El *modding* es una práctica de grupos de desarrollo interdisciplinarios que buscan crear, mantener y distribuir extensiones respecto de un *software* desarrollado previamente, denominadas *mods*. A su vez, esta práctica se encuentra muy difundida en la comunidad de videojuegos: gracias a ella se obtienen productos derivados con nuevas características, tales como cambios en su comportamiento y estética. Con ello mejora la jugabilidad y se aportan nuevas historias, mecánicas y arte, entre otros aspectos [1] [2].

Mediante habilidades de exploración, transformación e innovación, los usuarios avanzados de los videojuegos establecieron esta práctica, originada como una ocupación de tiempo libre que buscaba ampliar la experiencia del producto original [3], impulsando la creación de nuevo contenido, basado en los gustos de los consumidores, mediante la experimentación y uso de herramientas dedicadas [4]. Así, esta práctica evolucionó con los usuarios y en colaboración con las casas desarrolladoras, que a su turno brindaron soporte, al tiempo que facilitaron herramientas de edición como *software development kits* (SDK) e incentivos para que los usuarios pudieran llevar a cabo sus proyectos, con lo cual se extendió la vida útil de sus productos [3] [5]. Gracias a esto se han conformado comunidades abiertas de *modders* [1], en las cuales se comparte el conocimiento, experiencia y recursos para la creación de mejores trabajos de *modding* de productos de *software*. Las webs Moddb y NexusMods, ejemplos de estas plataformas, están dedicadas a la gestión, desarrollo y distribución de *mods* [6].

El enfoque de las *mods* es conducido por el desarrollo del usuario final, también llamado *end-user development* [4], el cual permite que los creadores de contenido tomen el rol de desarrolladores, también llamados *fan-programadores* [7], para hacer uso de recursos propios como técnicas de texturizado y programación, hasta actividades de análisis como permisos, derechos de autor y requerimientos iniciales [8]. Sus actividades se realizan también de forma análoga a otros dominios de la construcción del *software* —planificación y pruebas, por ejemplo—. Así, de forma similar al desarrollo de *software*, el *modding* comparte algunas de las problemáticas comunes presentes en equipos de trabajo, como la falta de planeación; a la vez que ostenta otras más críticas, tales como actualizaciones inesperadas o daños en el núcleo del juego [1]. Sumado a lo anterior, y debido a las pocas prácticas de ingeniería en el *modding*, surgen aún más problemas propios de comunidades abiertas de trabajo, entre los cuales se encuentran los siguientes:

- Repetición del trabajo: surge en inconvenientes de coordinación entre los miembros de trabajo, y de la poca orientación hacia la definición y asignación de roles.

- Problemas de comunicación: aparecen junto a barreras como la cultura, el idioma, el horario y la geografía, entre otros.
- Resultados inesperados: una documentación inadecuada, o su ausencia, puede llevar a imprevistos, y a que su seguimiento resulte bastante complejo de realizar.

Estos problemas de comunicación y duplicación del trabajo hacen que el esfuerzo de los *modders* sea alto, a la vez que su avance vaya lento, puesto que se realizan esfuerzos adicionales en la integración de los módulos creados por cada miembro del equipo, además de posibles inconsistencias entre estos datos. Por otro lado, dada la naturaleza del *modding* como una actividad cultural y creativa [3] [9], su dinámica es conducida por prácticas empíricas, basadas en el conocimiento de los miembros del equipo de desarrollo. Sumado a lo anterior, debido al desconocimiento de productos y herramientas de soporte [10], se inician modificaciones por caminos errados que normalmente conducen a un mayor esfuerzo y, de modo eventual, a la inviabilidad de la modificación. Debido a estas razones, muchos proyectos de *modding* fracasan, con lo cual se pierden grandes esfuerzos en exploración, desarrollo y configuración [11].

Mediante la similitud de las actividades del *modding* respecto de las presentes en el desarrollo de *software*, y a través de la ingeniería de procesos de *software*, se propuso en este proyecto hacer uso de los datos recolectados a través de una revisión de la literatura científica y un estudio exploratorio previo con respecto al *modding*, para el diseño y posterior propuesta de un proceso *software* que brinde soporte a los grupos de *modders* en la realización de este tipo de proyectos.

## 1. MATERIALES Y MÉTODOS

### 1.1 Análisis desde la literatura

Algunos trabajos relacionados con respecto al *modding* identifican aspectos típicos de esta práctica: la similitud con la ingeniería de software, estudios de caso previos, guías para la realización de proyectos y la identificación de aspectos como roles y tipos de *mods*. Sin embargo, se observa la ausencia de un proceso definido que brinde soporte a la práctica del *modding* y que, a su vez, aporte a la comunidad tareas, artefactos y demás elementos necesarios para establecer un camino a la creación de *mods* de una forma productiva, al tiempo que reduzca el número de incidencias por posibles fallos dentro de su desarrollo. Así, tomando la información recolectada desde el estado del arte previo en *Entendiendo el modding como un proceso de ingeniería* [6], se logró abstraer algunos conceptos clave dentro de esta práctica, lo que permitió realizar una caracterización más precisa de la escena del *modding* en comunidades *gamer* [6]. Estos conceptos se presentan como actividades ejecutadas por los *modders* dentro



de sus proyectos creativos y se dividieron en 3 categorías: *pre-modding*, *modding* y *post-modding*.

### 1.1.1 Actividades de *pre-modding*

*Análisis.* Como una técnica basada en el desarrollo de *software* por parte del usuario final (*end-user development*) [4], los propietarios del *software* son los encargados de la generación de contenido nuevo al producto base, en forma de extensiones [12] [13]. Por lo tanto, en el *modding* se realizan actividades que constan de un análisis anterior a los cambios, con el objetivo de definir en un comienzo la meta de modificación y las herramientas y permisos disponibles, y organizar el equipo de trabajo. A continuación se identifican las actividades que componen el análisis:

- *Análisis de modificación:* dentro de esta actividad se realizan planeaciones de los cambios que se desea obtener, a la vez que se establecen las metas a conseguir en el desarrollo del proyecto. Estos análisis se evidencian en *mods* a videojuegos populares como StarCraft [11], Half Life II [14] y Minecraft [15]. En estas guías los *modders* plasman desde un comienzo lo que se desea realizar, a la vez que se definen los objetivos que se espera obtener con el proyecto.
- *Análisis del producto:* en este se incorpora una serie de observaciones de los recursos disponibles en cuanto al *software* base, al igual que herramientas para realizar la labor de modificación. Se identifican los recursos necesarios y los que se encuentran disponibles para poner en marcha el proyecto [11], además de lo que el producto ofrece para ser modificado —arquitectura, plataforma, licencias y jugabilidad, entre otros— [14] [16] [17]. En la actualidad, las empresas desarrolladoras facilitan la labor de modificar partes de sus productos mientras no se realicen violaciones a su *Contrato de licencia para el usuario final* (CLUF) y, por lo general, brindan soporte a esta práctica. En consonancia con ello, es pertinente realizar una evaluación de las licencias del *software* que será objeto de uso en el proyecto de modificación [11].

*Adquisición de herramientas especializadas.* Tanto casas desarrolladoras, que apoyan la práctica de *modding* en sus productos, como *modders* experimentados que desarrollan herramientas de terceros (*third-party software*) [13], usan *software* dedicado a esta práctica. Estas herramientas permiten realizar la labor de exploración, edición e inclusión de archivos, y empaquetado de la extensión del *software* a modificar, con el fin de mejorar la eficiencia y facilitar la labor de algunas tareas que suelen ser complejas o repetitivas dentro de estos proyectos. Se encuentran dentro de estas herramientas kits de desarrollo de software (SDK) propios del *software* base, *software* de diseño asistido por computadora (CAD) y editores especializados de imagen, mapas, texto y sonido, entre otros [14] [16] [17].

*Definición de roles.* En tanto actividad colaborativa, el *modding* suele dar lugar a la conformación de equipos de personas con roles establecidos, que desempeñarán funciones específicas dentro del desarrollo del proyecto. Las responsabilidades y habilidades de cada integrante, así como su conocimiento, experiencia y talento, moldearán su respectivo rol (Tabla 1).

Los citados roles suelen asociarse a las motivaciones [18] de cada una de las personas pertenecientes al equipo, y además suelen variar en el transcurso del desarrollo del *mod* [7]. Los siguientes son algunos ejemplos de funciones que se pueden establecer en un proyecto de *modding*.

Tabla 1. Algunas funciones identificadas en proyectos de *modding*

Artista de texturas	Artista de sonido/música	Beta tester
Fundador	Artista conceptual	Proveedor de Servidor
Líder	Programador de herramientas	Administrador del servidor
Modelador	Escritores de ficción	Mapper
Programador	Animador	Artista gráfico
Webmaster	Diseñador web	Skinner
Scripter		

Fuentes: *Battlefield 1942* [7] y *Operation Flashpoint* [18] with special focus on the forms and consequences of collaboration between hobbyists. The case discussed in the article is the shooter-game *Operation Flashpoint* (OFP).

Se observa que la cantidad de funciones identificadas dentro de un grupo de trabajo puede llegar a ser extensa y muy variada, siempre en función de las capacidades de cada miembro.

### 1.1.2 Actividades de *modding*

*Cambios en la jugabilidad.* En los apartados del *modding*, la jugabilidad es un aspecto ampliamente modificado en estas comunidades [8] [9] [11]: reúne en conjunto las reglas, el funcionamiento, el diseño, la experiencia del videojuego con respecto al usuario final y con ella sus atributos —satisfacción, inmersión y emoción— [19].

*Edición de scripts – código de alto nivel.* Esta permitirá el cambio lógico del videojuego, con lo que se abrirá la posibilidad de inclusión de nuevos comportamientos en el sistema y la modificación de los existentes [17]. En este apartado, el *modder* podrá realizar cambios en las reglas o la física del juego, o bien habilitar contenido oculto.

*Edición multimedia.* Los recursos multimedia son elementos encargados de presentar la información al jugador en respuesta a sus interacciones. La edición de estos reúne todo lo correspondiente al manejo de gráficos, video, sonido y música, que van enlazados a los cambios que provee la extensión realizada [20].

*Extensión de la vida del software base.* Mediante la construcción y liberación de contenido derivado, las extensiones de *software* brindan a los usuarios finales contenido personalizado a través de la inmersión creativa: desde conversiones y edición multimedia hasta aspectos más avanzados, tales como corrección de errores y vulnerabilidades [1]. Este enfoque permite a las empresas realizar un seguimiento de sus productos en cuanto a los deseos de sus consumidores para futuros lanzamientos y mantenimiento del *software* liberado con anterioridad [13].

### 1.1.3 Actividades post-modding

*Divulgación del mod.* Dentro de la realización de un proyecto de modificación, los equipos de *modding* buscan distribuir el contenido realizado en la comunidad *gamer*. En ella, los equipos de trabajo buscan recibir retroalimentación frente a sus proyectos y, de igual manera, notificar posibles errores encontrados para un refinamiento posterior. En la actualidad existen comunidades en internet conformadas por millones de miembros que permiten alojar estos trabajos creativos y, posteriormente, compartirlos [13].

*Mejoras posteriores.* Posterior al resultado de la retroalimentación de la comunidad, y mediante el deseo de exploración, innovación y corrección [7] [18], se realizan mejoras a los trabajos realizados por los grupos de trabajo de *modding*. En ese sentido, se mejora la calidad del producto desarrollado y, con ello, se logra una mejor aceptación en la comunidad.

## 1.2 Datos desde el estudio exploratorio

Mediante la información recogida desde el estudio exploratorio realizado en [6] se pudieron obtener los siguientes datos, clasificados en tres secciones:

- Información con respecto a los *mods*
- Información con respecto a los *modders*
- Información con respecto a la administración del *modding*

Con respecto a los mods se encontró que un 38 % de los proyectos de *modding* de la comunidad de ModDB no han sido liberados, y que la mayoría de ellos lleva más de 6 años de desarrollo. Esto revela un porcentaje alto de pérdidas en ese sentido.

La cantidad de miembros dentro de las comunidades de *modding* analizadas supera 18 millones. Estos usuarios son los encargados de aportar a la escena del *modding* con el desarrollo de proyectos, en tanto que comparten conocimiento, habilidades y experiencia dentro de sus foros. En ese sentido, los participantes crean grupos de desarrollo y pequeños estudios independientes para la liberación continua de nuevos productos.

Las comunidades de ModDB y NexusMods proveen asistencia a los usuarios que desean empezar en proyectos de *modding*. Dentro de sus servicios brindan insumos básicos para el desarrollo, como foros de ayuda, seguimiento de actividades, manejo de versiones y herramientas de administración, entre otros.

### 1.3 El *modding* y los procesos de *software*

La ingeniería de procesos es un área perteneciente a la ingeniería de *software*, encargada de definir, implementar, medir y mejorar procesos de *software* [21]. Estos últimos se definen como conjuntos de actividades, participantes, estructuras, artefactos, herramientas y metodologías, entre otros elementos, cuyo fin es definir, desarrollar y mantener un producto de *software* [22]. Gracias al uso de procesos de *software* y de lenguajes de metamodelado como *Software Process Engineering Metamodel* (SPEM) se obtienen ventajas en el desarrollo, tales como:

- Facilitar la comprensión y comunicación humana.
- Facilitar la reutilización.
- Soporte para la gestión y mejora de procesos.
- Guiar la automatización de procesos.
- Dar soporte a la ejecución automática [21].

Lo anterior trae beneficios adicionales, como disponer de información para equipos de desarrollo cambiantes y facilitar la evaluación y certificación de estándares como CCMI e ISO, lo cual trae con ello ventajas comerciales y mejoras en la calidad del producto [6]. Debido a la gran diversidad de proyectos de desarrollo de *software* provistos por la complejidad, el contexto y el equipo de desarrollo, entre otros factores, el proceso de desarrollo de *software* no es universal. Sin embargo, algunas actividades son fundamentales y se encuentran presentes en la mayoría de los modelos de desarrollo [23]:

- 1) *Especificación de software*: se definen la funcionalidad y restricciones del software.
- 2) *Diseño e implementación*: se diseña el software de acuerdo con las especificaciones recogidas.
- 3) *Validación*: se realiza un proceso de esta naturaleza para asegurarse de que el producto cumpla con lo pactado.
- 4) *Evolución*: etapas de desarrollo y mantenimiento del *software* para que se adapte.

A través de la información presentada en el estado del arte y en el estudio exploratorio en [6] se identifican conceptos y prácticas que son similares entre los grupos de desarrollo de *modding* y el proceso de desarrollo de *software*. A continuación, se presenta una tabla comparativa entre estas dos prácticas.

Tabla 2: Comparación entre el *modding* y el proceso de desarrollo de *software*

Actividades del proceso de desarrollo de <i>software</i>	Actividades en el <i>modding</i>
Especificación de <i>software</i>	Los equipos de <i>modders</i> establecen lo que se desea alcanzar con el proyecto de <i>modding</i> , así como las restricciones vigentes conforme a las licencias, permisos y recursos —humanos y de <i>software</i> — [11].
Diseño e implementación	Después de haberse establecido los objetivos en el inicio del proceso, los <i>modders</i> implementan los cambios en el sistema base a fin de desarrollar su producto [14] [16] [17].
Validación	Antes de la liberación de sus productos, los equipos de <i>modding</i> realizan la evaluación de las metas propuestas en un comienzo, para luego compartirlas en la web de la comunidad [3].
Evolución	Mediante la retroalimentación obtenida de los usuarios participantes en comunidades de <i>modding</i> , los equipos de desarrollo realizan correcciones a sus productos, así como mejoras y adaptaciones a otras plataformas [3] [24].

Fuente: [6].

Sumadas a las actividades fundamentales del proceso de desarrollo de *software*, se pudieron establecer otras similitudes con respecto al *modding* en [6], que se presentan a continuación:

- *Roles*: con el estudio exploratorio y el estado del arte se ha identificado que los miembros de los equipos de desarrollo de *modding* toman un rol dentro de sus proyectos y la comunidad con respecto a su talento, conocimiento y experiencia [7] [11] [18].
- *Actividades*: tanto en las guías presentes en el estado del arte [14] [16] [17] como en foros de las comunidades de *modding* se ha visto que los *modders* realizan pasos coherentes para desarrollar sus productos a partir del análisis, diseño, implementación y mantenimiento con la retroalimentación de los miembros de la comunidad.
- *Artefactos*: para la administración de los grupos de desarrollo de *modding*, así como para el establecimiento de tareas, se realiza el uso de artefactos de *software* mediante la documentación y la distribución de recursos multimedia dentro del equipo.
- *Productos*: dentro del estudio exploratorio de las comunidades de ambas webs se observa el lanzamiento constante de versiones del producto en función del cumplimiento de los objetivos trazados en el proyecto de *modding*.

- Equipos de desarrollo: mediante la colaboración de los miembros dentro de los equipos de *modding* se trabaja en conjunto para lograr los objetivos de desarrollo, de manera similar a las metodologías para el desarrollo de *software* [1].

Así, el *modding* abre la posibilidad de obtener, mediante la ingeniería de procesos, insumos para la mejora de esta práctica; ello involucra la definición de actividades, tareas y roles, y trae consigo la posibilidad de generar metodologías para el respaldo, de tal manera que se obtenga soporte y, con ello, mejores resultados en proyectos de *modding*.

#### 1.4 Aspectos generales del *modding*

A través de la información recolectada desde el estado del arte y el estudio exploratorio realizado en [6], se procederá aquí a identificar aspectos comunes del *modding* como base para el posterior diseño del proceso de *software*.

##### 1.4.1 Tipos de *modding*

Dentro del área del *modding* existen varios tipos de proyectos, los cuales se clasifican de acuerdo con su finalidad. En [1] se realiza una clasificación de 4 tipos de proyectos, a saber:

- *User interface*: los proyectos de este tipo se relacionan con las adaptaciones realizadas al *software* para mejorar la experiencia de los usuarios finales.
- *Game conversions*: es la forma más popular de *modding*. Se divide en dos partes, esto es, conversiones parcial y total (Tabla 1).

Tabla 1: Diferencias entre las conversiones parcial y total

Conversión parcial	Conversión total
Cambios en caracteres	La conversión total incluye los cambios de la conversión parcial, pero en este caso se realiza una reconstrucción completa del juego, que dará como resultado un juego nuevo.
Cambios en los objetos	
Niveles, mapas, terrenos...	
Reglas del juego	
Mecánicas del juego	

Fuente: elaboración propia.

- *Machinima*: generación de contenido multimedia a través del motor del videojuego que se modificará. Entre ellos se encuentran cambios en la historia del personaje, continuaciones o *spin-off* (proyecto que nace como extensión de otro anterior). Actualmente existen grandes comunidades generadoras de este tipo de contenido, que generan *fan-fiction* (relatos de ficción, obras literarias o dramáticas escritas por

Fans de algún producto) extendiendo la experiencia en el juego con sus proyectos creativos.

- *Hacking closed systems*: este estilo de modificación consiste en el uso de técnicas como la ingeniería inversa para obtener información de partes cerradas del sistema —muchas veces con el fin de obtener ventaja frente a otros usuarios—. Las prácticas de este tipo suelen estar prohibidas por las casas desarrolladoras y la comunidad *gamer* en general, ya que se violan los acuerdos de la licencia del *software* y las normas de la comunidad misma.

#### 1.4.2 Aspectos sociales del *modding*

En el trabajo de Sotamaa [18] with special focus on the forms and consequences of collaboration between hobbyists. The case discussed in the article is the shooter-game Operation Flashpoint (OFP se presenta un enfoque de la comunidad *modder* y algunos aspectos sociales en esta clase de grupos. La tabla 4 presenta una clasificación de las distintas motivaciones de los *modders* dentro del proyecto *Operation Flashpoint*.

Tabla 2: Motivaciones presentes en el *modding*

Jugabilidad	Disfrutar la experiencia del videojuego y con ello cubrir, mediante la construcción de contenido propio, la sensación de incompletitud del producto.
Hacking	Descubrir el funcionamiento interno del producto.
Investigación	Dada la naturaleza del videojuego ( <i>Operation Flashpoint</i> ), se realiza una observación desde el punto militar.
Expresión artística	La experiencia de la creación en el apartado del arte.
Cooperación	Creación de ideas con otros miembros del equipo y el entusiasmo por trabajar hacia un objetivo compartido.

Fuente: elaboración propia.

Estas motivaciones van de la mano con el análisis de Postigo [7] respecto a los temas que motivan a los *modders* y, en especial, a los *mappers*. A partir de sus creaciones es posible realizar una clasificación de tres temas centrales que los motivan: el primero, su deseo de hacer una contribución artística a las comunidades y obtener retroalimentación de ellas; el segundo, la identificación que sienten con su creación y su apropiación al producto mediante el diseño de elementos únicos que les representan algún significado fuerte; y el tercero, el efecto de su acción en el campo laboral, por cuanto cada proyecto se suma a su lista de experiencia y mejora sus oportunidades de conseguir un trabajo lucrativo en la industria del desarrollo de videojuegos.

#### 1.4.3 Licencias híbridas

En el trabajo de Scacchi [1] se muestra un enfoque de licencias híbridas en que se parte de un *software* propietario, pero con modificaciones abiertas (Figura 1). Mediante este enfoque de trabajo se permite reutilizar recursos producidos en la

comunidad de *modding*, cuestión que posibilitaría la retroalimentación y mejora en los productos que se desarrollan —siempre que se mantenga respeto por el núcleo o *software base* (privativo)—.

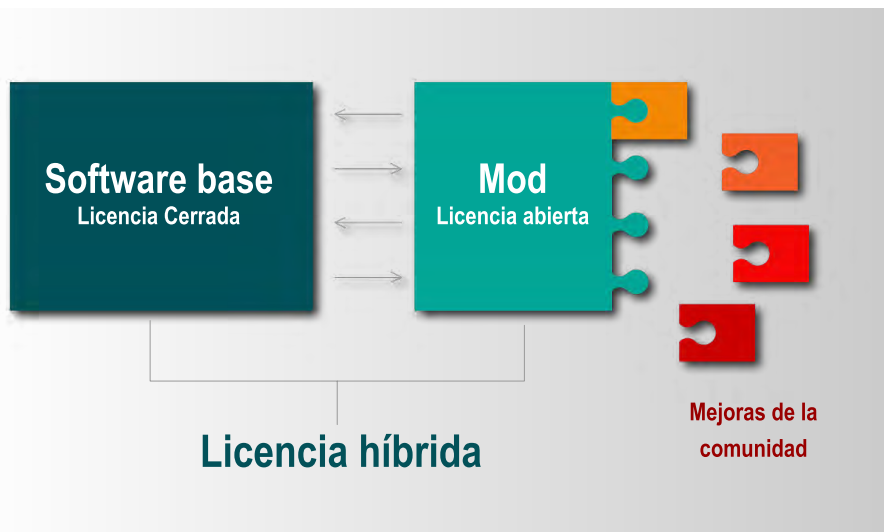


Figura 1. Enfoque de licencias híbridas

Fuente: elaboración propia.

### 1.5 Metodología de Formalización

Para lograr la formalización y diseño de “*Kross Modding Process*”, como un proceso que permita apoyar la práctica del modding mediante la colaboración, se procedió a realizar las actividades propuestas en el metaproceto definido por Ruiz [22], que se divide en tres fases:

- 1) *Planeación*: según lo descrito en el metaproceto de Ruiz, esta fase consiste en entender las necesidades de la empresa para la cual se diseñarán o mejorarán los procesos de software. Esta etapa involucra, a su turno, tres tareas principales (figura 2):
  - a. *Analizar el contexto de la empresa*: radica en conocer las metas de negocio de empresa y el modo en que se alcanzan.
  - b. *Definir y planear actividades*: se reúnen representantes de cada área, interesados y usuarios finales con fin de mostrar las ventajas de disponer un proceso definido en la empresa, establecer su alcance, identificar las principales actividades y designar a personal de cada área para funcionar como enlace de comunicación con los ingenieros de procesos.



c. *Entrenar a los stakeholders*: tras conocerse las metas y actividades, se realiza una familiarización con el proceso, el lenguaje y la herramienta de modelado para efectos de mantenimiento y evolución posteriores, mediante la participación activa de los miembros.

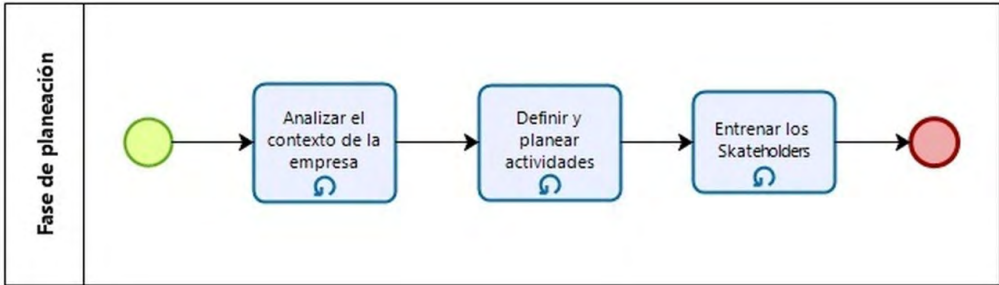


Figura 2. Tareas que componen la fase de planeación

Fuente: metaproceso de Ruiz [22].

2) *Ejecución*: esta fase del metaproceso consta de cinco tareas que permiten obtener la información de los procesos de *software* de la empresa, que luego darán pie a analizar, diseñar y especificar un prototipo mediante el lenguaje SPEM para realizar, finalmente, una validación con los interesados del proyecto, lo cual involucrará refinar u optimizar el producto final (figura 3).

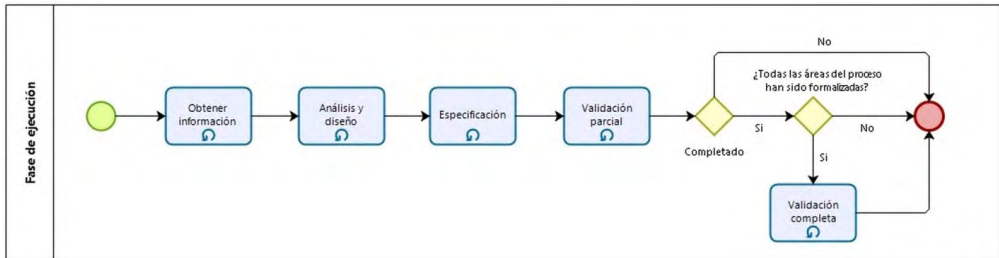


Figura 3. Fase de ejecución

Fuente: metaproceso de Ruiz [22].

3) *Entrega*: esta etapa consta de una sola tarea, llamada *despliegue del proceso*. Consiste en entregar a la empresa desarrolladora el proceso mediante un lenguaje como SPEM, desplegado en sus computadoras, en aras de tenerlo como referencia e información para los miembros para que evolucione y se adapte a las necesidades de la empresa (figura 4).

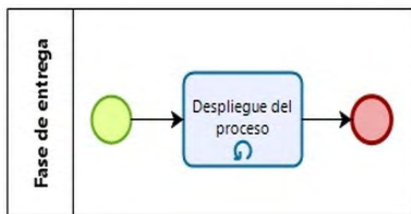


Figura 4. Tarea de despliegue

Fuente: metaproceso de Ruiz [22].

## 2. RESULTADOS

Mediante los datos recogidos en la literatura científica y el estudio exploratorio previo, sumado a la revisión de aspectos generales del *modding*, y a través de la metodología de formalización de Ruiz [22], se diseñó un proceso de *software* denominado *Kross Modding Process*, que se presenta a continuación.

### 2.1 Kross Modding Process

Este es un proceso de *software* para el desarrollo de la práctica *modding*. Está diseñado para ofrecer soporte a sus principales actividades de forma guiada, mediante tareas que abarcan la administración del proyecto, así como construcción y mejoras a través de la coordinación y colaboración de los miembros del equipo de trabajo (Figura 5).

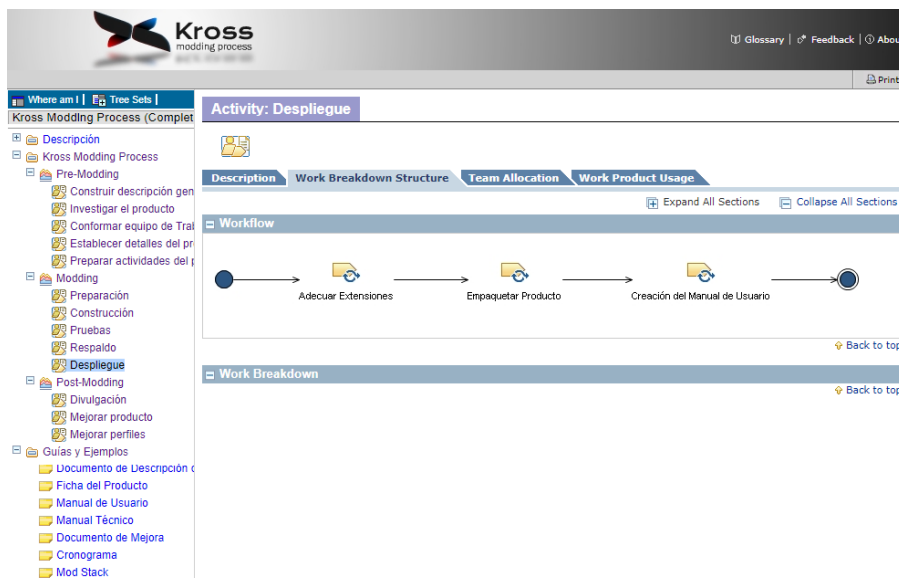


Figura 5. Versión web del proceso

Fuente: elaboración propia.

El proceso busca solucionar algunas problemáticas detectadas en la comunidad y la escena del *modding*—tales como *retrabajo*, problemas de comunicación y ausencia de definición de tareas y roles— mediante el uso de pasos coherentes para principiantes, la comunicación continua y la clasificación de la información obtenida en el proyecto.

### 2.1.1 Principios

*Kross Modding Process* se basa en cinco principios fundamentales, que buscan guiar a los participantes del proyecto mediante las buenas prácticas internas en un equipo de desarrollo (Tabla 3).

Tabla 3. Principios de *Kross Modding Process*

Principio	Descripción
1. Respeto	En la práctica del <i>modding</i> es importante considerar el respeto hacia el trabajo ajeno. Ello implica respetar dos elementos puntuales: de un lado, las licencias y las restricciones en los productos que se van a modificar, para evitar con ello el uso de herramientas que violen la propiedad intelectual de los recursos involucrados (incluido el <i>software</i> ); y de otro, el trabajo de otros <i>modders</i> dentro de la comunidad y conforme a ello realizar la respectiva acreditación cuando se usen contenidos ajenos.
2. Multidisciplinariedad	“Solo trabajarás más rápido, pero acompañado llegarás más lejos”. Las múltiples habilidades de los miembros de un equipo promueven la creación de buen contenido. Por lo tanto, es necesario contar con estos talentos y capacidades en la integración del material para lograr mejores resultados y, con ello, mejores <i>mods</i> .
3. Colaboración	La colaboración es necesaria dentro de la escena del <i>modding</i> : se requiere en este sentido el concierto de los miembros del equipo de desarrollo, así como trabajo, herramientas, contenido y opinión de terceros.
4. Modularidad	La base de un proyecto de <i>modding</i> es crear módulos cuya naturaleza es de extensión conectable al producto de <i>software</i> que será objeto de modificación, y requiere su previa adquisición para funcionar. Un <i>mod</i> no funciona como un producto aislado.
5. Extensibilidad	La naturaleza de la escena del <i>modding</i> es creativa y abierta. Se busca que los productos de <i>software</i> se expandan en características y en mejoras.

Fuente: elaboración propia.

### 2.1.2 Fases

*Kross Modding Process* consta de tres fases, que representan los estadios de un proyecto *modding* (Figura 6). A su vez, estas fases se dividen en trece actividades que parten de la investigación del producto base hasta la mejora del producto *mod* final.

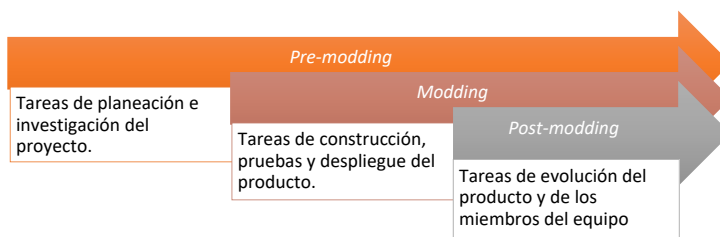


Figura 6. Fases de *Kross Modding Process*

Fuente: elaboración propia.

### 2.1.3 Roles

Mediante la abstracción de algunas funciones comunes en proyectos de *modding*, y a partir de la literatura y el estudio exploratorio, se identificaron tres roles principales, cuyas actividades y tareas se disponen en función de la capacidad del participante y el tipo de proyecto en desarrollo. A continuación, se describen los roles presentes en el proceso (Figura 7).

- *Líder del equipo*: se encarga de dirigir las actividades del equipo para conseguir buenos resultados en el desarrollo del proyecto. Además, establece roles y responsabilidades, facilita documentación y herramientas a sus miembros, fija metodologías para el desarrollo, y socializa conceptos sobre la idea de modificación.
- *Modder*: encargado de dar lugar a los cambios en el software base. Este miembro puede desempeñar uno o varios tipos de funciones, según sus habilidades y talentos (tabla 6).
- *Participante*: desempeña funciones alternas a las tareas de modificación dentro del grupo de trabajo. Con respecto a su talento y capacidad para aportar en el desarrollo del producto final, las funciones que puede desempeñar un participante pueden relacionarse con tareas como generador de retroalimentación, divulgador en redes sociales y *webmaster*, entre otras.

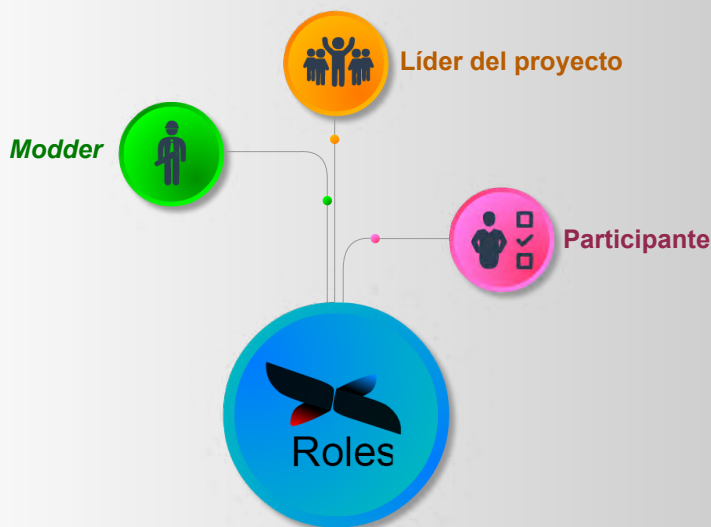


Figura 7. Roles de *Kross Modding Process*

Fuente: elaboración propia.

Sumado a lo anterior, cada miembro del equipo de desarrollo puede desempeñar múltiples funciones con respecto a su rol; en la Tabla 4 se muestran algunos ejemplos de ellas en un proyecto típico de *modding*.

Tabla 4: Algunas funciones presentes en los roles del proceso

Rol	Clasificación	Administración	Arte	Conceptos	Programación
Líder del proyecto		<ul style="list-style-type: none"> <li>• Director</li> <li>• Administrador de repositorios</li> <li>• Facilitador</li> </ul>		<ul style="list-style-type: none"> <li>• Analista</li> <li>• Planificador</li> </ul>	
Modder			<ul style="list-style-type: none"> <li>• Diseñador de interfaces</li> <li>• <i>Mapper</i></li> <li>• Texturizador</li> <li>• Artista de sonido</li> <li>• Modelador 3D/2D</li> <li>• Animador</li> </ul>	<ul style="list-style-type: none"> <li>• Creador de personajes</li> <li>• Diseñador de niveles</li> </ul>	<ul style="list-style-type: none"> <li>• Programador de herramientas</li> <li>• <i>Scripter</i></li> <li>• Editor de características</li> <li>• Programador general</li> </ul>
Participante		<ul style="list-style-type: none"> <li>• Generador de retroalimentación</li> <li>• <i>Webmaster</i></li> <li>• Divulgador en redes</li> </ul>	<ul style="list-style-type: none"> <li>• Artista conceptual</li> </ul>	<ul style="list-style-type: none"> <li>• Historiador</li> </ul>	

Fuente: elaboración propia.

#### 2.1.4 Artefactos

*Kross Modding Process* ofrece diez artefactos que brindan la documentación necesaria para el desarrollo del proyecto de *modding* (figura 8), cuya finalidad es guiar a los miembros del equipo para facilitar la colaboración; se describen a continuación.

- *Cronograma del proyecto*: se compone de actividades que se delimitan por duración y tienen fechas de inicio y fin. Este artefacto permite al grupo de desarrollo gestionar las actividades a cumplir y, con ello, administrar el tiempo y recursos para lograrlo.
- *Documento de descripción del proyecto*: recoge los datos del diseño del proyecto. En él se incluyen las ideas principal y secundarias; la información del producto que será objeto de *modding*; los datos de los miembros del equipo y sus responsabilidades; y los objetivos, enlaces de las herramientas y recursos en línea.
- *Ficha del producto*: en este documento se recoge la información obtenida sobre los criterios para la elección, licencia, permisos y prestaciones para el proyecto que posee el producto base.
- *Mod stack*: este artefacto agrupa los *modding* items según su importancia en el desarrollo del proyecto, para después priorizarlos. Queda a disposición del equipo definir los criterios de priorización.

- *Manual técnico del mod*: se incluyen aquí todos los datos correspondientes a las modificaciones que hayan resultado de las tareas de modding del proyecto. Esta información funciona como referencia para el equipo de modders, así como para pruebas técnicas del producto, corrección de errores y posible adaptación de nuevas modding tasks.
- *Incremento*: representa cada liberación previa del mod, la cual se traduce a un conjunto de modding tasks completadas en la fase de modding.
- *Mod*: es el resultado del proyecto de modding que está en ejecución. Este es una extensión de software que se conectará al software base del proyecto para funcionar. Cada vez que se completa la fase de modding, el resultado es la liberación de una nueva versión del mod (incremento).
- *Manual de usuario*: documento en el que se consolida la información necesaria para la instalación y el manejo del mod. Dentro de él se describen al usuario los pasos a seguir para poner a marchar el mod.
- *Parche*: el propósito de un parche de software es corregir errores imprevistos en el mod en tiempo de liberación que impidan su ejecución. Estos parches vienen en forma de archivos, ejecutables o comprimidos que se liberan al usuario final con su respectiva descripción.
- *Documento de mejora*: tiene como finalidad recopilar las ideas de mejora para el mod que puedan surgir gracias a la retroalimentación de los usuarios finales. Esto con el fin de ampliar la experiencia del producto, de tal suerte que sus características mejoren en desarrollos futuros para aumentar su calidad.

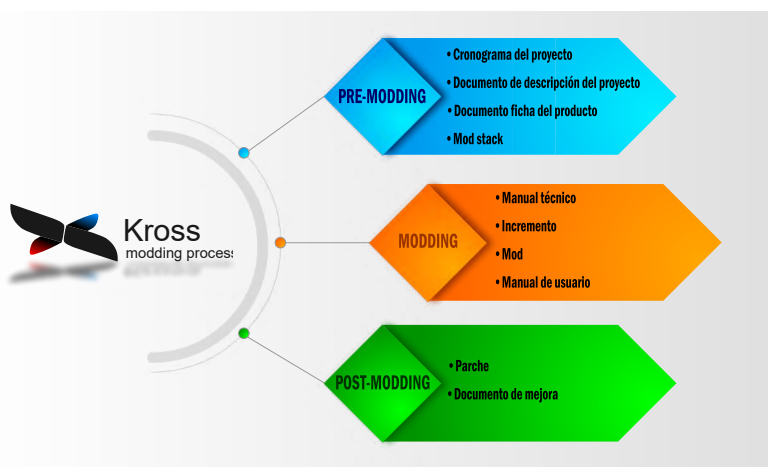


Figura 8. Artefactos de *Kross Modding Process*

Fuente: elaboración propia.

### 3. DISCUSIÓN

De acuerdo con el planteamiento del proyecto, y a través de la información recogida desde el estado del arte y el estudio exploratorio previo, *Kross Modding Process* provee una base para que grupos de desarrollo de *modding* puedan comenzar proyectos de esta naturaleza con actividades, tareas, roles e información definidos. Frente a los problemas identificados mediante esta investigación, *Kross Modding Process* permite reducir inconvenientes comunes dentro de los proyectos de *modding* como el retrabajo, problemas de comunicación y resultados inesperados a través de la documentación y definición de actividades y roles, a la par con el uso de los artefactos de soporte.

Debido a la naturaleza del *modding* y la heterogeneidad de sus proyectos, *Kross Modding Process* busca evolucionar y adaptarse a sus usuarios mediante el refinamiento, dado por experiencias reales en proyectos de *modding*. De igual manera, el proceso busca abordar un tema aún desconocido en la ingeniería de procesos, con lo que constituye un proyecto de investigación preliminar que puede conducir a futuras investigaciones en este campo (tales como la definición de metodologías enfocadas a la extensión de *software*).

### 4. CONCLUSIONES

El *modding* es una práctica en crecimiento que trae consigo ventajas y oportunidades de negocio para empresas productoras y estudios independientes de videojuegos, así como para los usuarios finales que actúan como desarrolladores: mejoran sus perfiles al tiempo que aportan a la innovación de nuevos productos en el mercado. Sin embargo, el *modding* ha sido poco abordado en el área de la ingeniería del *software* dado que su naturaleza es más creativa y empírica, y sus proyectos son multidisciplinarios. En ese sentido el *Kross Modding Process* propuesto en este trabajo busca servir de guía para los *modders* y dar solución a problemas típicos en ese tipo de proyectos.

Este proceso es el resultado de seguir una metodología investigativa empírica. Aun cuando se encuentra en mora de aplicarse en diferentes escenarios, brinda las bases iniciales para sistematizar el *modding* como un proceso de desarrollo de *mods*. Es importante aclarar que el autor principal de esta investigación es un *modder* que ha abordado esta temática durante varios años y conoce cómo funcionan las comunidades, su estilo de trabajo y los productos que desarrollan, así como las limitaciones y dificultades de esta práctica.

Como proyecto a futuro se busca realizar una evaluación, mediante varios estudios de caso, encaminada a medir aspectos como la calidad y la productividad, e identificar limitaciones, con la finalidad de refinar el proceso adaptándolo a las diferentes

necesidades de grupos independientes de *modders* y a los distintos ámbitos de trabajo por los cuales se caracterizan los proyectos de esta naturaleza.

## REFERENCIAS

- [1] W. Scacchi, “Modding as an Open Source Approach to Extending Computer Game Systems”, *Int. J. Open Source Softw. Process.*, vol. 3, n.º 3, pp. 36-47, 2011. DOI: 10.4018/jossp.2011070103.
- [2] J. Kücklich, “Precarious Playbour : Modders and the Digital Games Industry The History of Modding The Economy of Modding”, *The Fibreculture Journal*, n.º 5, pp. 1-9, 2005.
- [3] M. Trenta, “La gestión de las comunidades de modding entre explotación y participación” en *Actas – V Congreso Internacional Latino de Comunicación*, pp. 1-15, 2013.
- [4] H. Lieberman, F. Paternò y V. Wulf, “End User Development: An Emerging Paradigm”, *End User Dev.*, vol. 9, pp. 9-16, 2006. DOI: 10.1007/1-4020-5386-X.
- [5] S. Guthals, S. Foster, y L. Handley, *Minecraft modding for kids for dummies*, Hoboken: John Wiley & Sons, 2015.
- [6] A. Ceballos, L. Pantoja y J. A. Hurtado, “Entendiendo el Modding como un Proceso de Ingeniería”, en *Ciencias Computacionales*, Cartagena, pp. 175-186, 2018.
- [7] H. Postigo, “Of Mods and Modders: Chasing Down the Value of Fan-Based Digital Game Modifications”, *Games Cult.*, vol. 2, n.º 4, pp. 300-313, 2007. DOI: 10.1177/1555412007307955.
- [8] W. Scacchi, “Modding as a basis for developing game systems”, en *Proceeding of the 1st international workshop on Games and software engineering - GAS '11*, p. 5, 2011. DOI: 10.1145/1984674.1984677.
- [9] A. Unger, “Modding as Part of Game Culture”, en *Computer Games and New Media Cultures*, Dordrecht: Springer Netherlands, 2012, pp. 509-523. DOI: 10.1007/978-94-007-2777-9\_32.
- [10] G. Poderi y D. J. Hakken, “Modding a free and open source software video game: “Play testing is hard work”, *Transformative Works and Cultures*, vol. 15, 2013. DOI: 10.3983/twc.2014.0493.
- [11] D. Johnson, “StarCraft fan craft: Game mods, ownership, and totally incomplete conversions”, *Velv. Light Trap*, vol. 64, n.º 1, pp. 50-63, 2009. DOI: 10.1353/vlt.0.0041.
- [12] O. Sotamaa, “Computer game modding, intermediality and participatory culture”, *New Media*, pp. 1-26, 2003.
- [13] K. A. Moody, “Modders: Changing the game through user-generated content and online communities.”, *Diss. Abstr. Int. Sect. A Humanit. Soc. Sci.*, vol. 75, n.º 11-A(E), 2015. DOI: 10.17077/etd.5ak8cz3w.
- [14] E. Guilfoyle, *Half-Life 2 mods for dummies*, Hoboken: John Wiley & Sons, 2007.



- [15] A. Gupta y A. Gupta, *Minecraft modding with Forge*, Estados Unidos: O'Reilly Media, 2015.
- [16] E. Guilfoyle, *Quake 4 mods for dummies*, Hoboken: John Wiley & Sons, 2006.
- [17] J. Van Gumster y C. Ammann, *Farming simulator modding for dummies*, Hoboken: John Wiley & Sons, 2014.
- [18] O. Sotamaa, "When the Game Is Not Enough: Motivations and Practices Among Computer Game Modding Culture", *Games Cult.*, vol. 5, n.º 3, pp. 239-255, 2010, <https://doi.org/10.1177/1555412009359765>.
- [19] J. L. G. Sánchez, N. P. Zea y F. L. Gutiérrez, "Playability: How to identify the player experience in a video game", 2009, [https://doi.org/10.1007/978-3-642-03655-2\\_39](https://doi.org/10.1007/978-3-642-03655-2_39).
- [20] R. Hong, "Game Modding, Prosumerism and Neoliberal Labor Practices", *Int. J. Commun.*, vol. 7, pp. 984-1002, 2013.
- [21] F. Ruiz y J. Verdugo, "Guía de Uso de SPEM 2 con EPF Composer", *Composer*, vol. 3, p. 93, 2008. DOI: 10.13140/2.1.1455.9049.
- [22] P. Ruiz, A. Quispe, M. C. Bastarrica y J. A. H. Alegría, "Formalizing the Software Process in Small Companies" 2012.
- [23] P. Letelier, "Proceso de Desarrollo de Software", *Univ. Politécnica Val.*, pp. 1-14, 2003.
- [24] T. Sihvonen, *Players Unleashed! Modding the Sims and the Culture of Gaming*, Amsterdam: Amsterdam University Press, 2011. DOI: 10.5117/9789089642011.



# Construcción de un juego serio como apoyo al aprendizaje de la física cinemática\*

*Carlos Fernando Aguilar Castrillón\*\**

*Andrés Felipe Ojeda Rivera\*\*\**

*Carol Julieth Aguilar Paz\*\*\*\**

*María Isabel Vidal Caicedo\*\*\*\*\**

*Marta Cecilia Camacho Ojeda\*\*\*\*\**

*Gabriel Elías Chanchí Golondrino\*\*\*\*\**

Recibido: 01/07/2019 • Aceptado: 28/11/2019

<https://doi.org/10.22395/rium.v19n37a8>

## Resumen

En este artículo se presenta la experiencia obtenida en la construcción de un prototipo de juego serio, cuyos objetivos son apoyar el desarrollo de los cursos de física de educación media e incluir procesos lúdicos en el proceso de aprendizaje. Del mismo modo, este juego pretende servir de referencia para la construcción de juegos serios en diferentes contextos de aplicación. Esta iniciativa académica estuvo motivada por las dificultades evidenciadas en estudiantes de educación media en cuanto a la apropiación del conocimiento de física cinemática. El juego fue construido con base en un modelo en el cual se consideran elementos pedagógicos y lúdicos propios de este tipo de herramientas interactivas, entre los cuales se tienen en cuenta la definición de los objetivos de aprendizaje; el diseño del componente lúdico del juego; y el diseño de interfaces. Para el proceso de evaluación se realizó una inspección de usabilidad, conducida a partir de los diez principios heurísticos de usabilidad de Nielsen; con ello se constató que ciertas cuestiones relacionadas con los principios de visibilidad del sistema, consistencia y estándares, ayuda y documentación deberían ser objeto de mejora en una versión futura del juego serio.

**Palabras clave:** física cinemática; juegos serios; educación media; metodología de aprendizaje.

\* Artículo derivado del proyecto de investigación de convocatoria interna de la Institución Universitaria Colegio Mayor del Cauca titulado “Definición de un proceso pedagógico y tecnológico para apoyar ambientes de enseñanza aprendizaje de la física cinemática a través de un sistema soportado por videojuegos”. Entidad que financia: Institución Universitaria Colegio Mayor del Cauca. Fecha de inicio: 12 de febrero de 2018. Fecha de finalización: 12 de febrero de 2019.

\*\* Ingeniero Informático, Institución Universitaria Colegio Mayor del Cauca. Correo electrónico: carlosaguilar@unimayor.edu.co. Orcid: 0000-0002-3316-0209.

\*\*\* Ingeniero Informático, Institución Universitaria Colegio Mayor del Cauca. Correo electrónico: felip77@unimayor.edu.co. Orcid: 0000-0003-2789-6727.

\*\*\*\* Magister en Ingeniería Física. Grupo de Investigación I+D en Informática, Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca. Correo electrónico: carola@unimayor.edu.co. Orcid: <https://orcid.org/0000-0001-5886-3640>.

\*\*\*\*\* Magister en Ingeniería. Grupo de Investigación I+D en Informática, Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca, Popayán, Colombia. Correo electrónico: mvidal@unimayor.edu.co. Orcid: <https://orcid.org/0000-0002-5456-3082>.

\*\*\*\*\* Ph. D. en ciencias de la electrónica. Miembro del Grupo de Investigación I+D en Informática, Facultad de Ingeniería, Institución Universitaria Colegio Mayor del Cauca. Correo electrónico: cecamacho@unimayor.edu.co. Orcid: <https://orcid.org/0000-0003-1973-3063>.

\*\*\*\*\* Ph. D. en Ingeniería Telemática. Facultad de Ingeniería, Universidad de Cartagena. Correo electrónico: gchanchig@unicartagena.edu.co. Orcid: <https://orcid.org/0000-0002-0257-1988>.

## Construction of a Serious Game as Support for the Learning of Kinematics

### **Abstract**

The experience obtained in the construction of a serious game prototype is presented in this paper, aimed at supporting the development of physics courses in secondary education, and including game-oriented processes in the learning process. Similarly, the proposed serious game is intended to serve as a reference for the construction of serious games in different application contexts. This development is mainly driven by the difficulties evidenced in high school students regarding appropriation of knowledge of kinematics. The proposed serious game was built based on a model in which pedagogical and playful aspects of this type of interactive tools are considered, among which the definition of learning objectives, the design of the playful component of the game, and interface design are taken into account. A usability inspection was carried out for evaluation purposes, based on the ten Nielsen heuristic principles of usability, which yielded various recommendations: it would be necessary to improve a future version of the game in some aspects related to the principles of consistency and standards, help and documentation, and visibility of the system.

**Keywords:** Kinematics; middle education; learning methodology; serious games.

## Construção de um jogo sério como apoio à aprendizagem da física cinemática

### **Resumo**

Neste artigo apresenta-se a experiência obtida na construção de um protótipo de jogo sério, cujos objetivos são apoiar o desenvolvimento dos cursos de física de educação média e incluir processos lúdicos no processo de aprendizagem. Do mesmo modo, esse jogo pretende servir de referência para a construção de jogos sérios em diferentes contextos de aplicação. Essa iniciativa acadêmica esteve motivada pelas dificuldades evidenciadas em estudantes de educação média quanto à apropriação do conhecimento de física cinemática. O jogo foi construído com base num modelo no qual se consideram elementos pedagógicos e lúdicos próprios desse tipo de ferramentas interativas, entre os quais se levam em consideração a definição dos objetivos de aprendizagem; o desenho do componente lúdico do jogo; e o desenho de interfaces. Para o processo de avaliação, realizou-se uma inspeção de usabilidade, conduzida a partir dos dez princípios heurísticos de usabilidade de Nielsen; com isso se constatou que certas questões relacionadas com os princípios de visibilidade do sistema, consistência e padrões, ajuda e documentação deveriam ser objeto de melhora numa versão futura do jogo sério.

**Palavras-chave:** física cinemática; jogos sérios; educação média; metodologia de aprendizagem.

## INTRODUCCIÓN

El esquema general del aprendizaje de la física inicia por el estudio de la cinemática de una partícula: los principios de esta unidad de aprendizaje corresponden a la base para el entendimiento de conceptos avanzados. El estudio de la cinemática requiere establecer de forma efectiva los contenidos temáticos mediante el aprendizaje significativo, teniendo en cuenta además, aquellos conocimientos previos conceptualizados a través de la observación del entorno.

En su trabajo sobre la preparación de docentes en física y otras ciencias, McDermott concluye que el estudio de la cinemática debe iniciar por el desarrollo de la comprensión cualitativa a partir de la experiencia o de la observación: ha de retenerse el formalismo matemático hasta que los estudiantes adquieran cierta práctica de razonamiento cualitativo con relación al fenómeno estudiado. El objetivo final de este tipo de comprensión es buscar estrategias que permitan a los estudiantes apropiarse de conocimientos básicos y fundamentos matemáticos, a fin de que articulen las relaciones conceptuales según sus propios términos [1]. Por lo tanto, se busca establecer nuevas metodologías en las que la representación de situaciones cercanas a la realidad pueda ser la clave en el aprendizaje de la física, de tal forma que los contenidos se presenten con material potencialmente significativo; y se implementen actividades de experimentación simuladas que permitan a los estudiantes observar en tiempo real los fenómenos estudiados. En este aspecto, el uso de las TIC y, en particular, la implementación de los juegos serios, constituyen nuevas líneas de investigación en las metodologías de aprendizaje.

En trabajos realizados en investigaciones anteriores, como el desarrollado por Mayes y De Freitas [2], se ha encontrado que los juegos y simulaciones pueden mejorar el aprendizaje diferenciado, en especial a grupos de estudiantes con habilidades y enfoques de aprendizaje muy diversos. La investigación también indicó que los juegos y simulaciones pueden apoyar de manera efectiva a los alumnos con necesidades basadas en habilidades (como la alfabetización y la aritmética). Así mismo, los juegos y las simulaciones apoyan el aprendizaje formal e informal, y pueden convertirse en una forma eficaz de vincular estos procesos y acelerar los resultados esperados de aprendizaje.

En la actualidad, las investigaciones sobre diseño de juegos se han centrado en los roles de los estudiantes en el proceso de aprendizaje; la interacción entre grupos de estudiantes y profesores; los procesos de enseñanza-aprendizaje [3]; el estudio de actividades y habilidades que se derivan de la interacción de los estudiantes con la tecnología [4]; y el proceso de organización y facilitación de las fases de diseño del juego. Las prácticas y los enfoques de esta área incluyen juegos de diseño [5], actividades de diseño [6], presentación progresiva de desafíos, esquemas en el desarrollo de

habilidades de diseño de los juegos [7] y la presentación del concepto de *aprendizaje* durante el proceso de diseño [8].

En todos los casos mostrados, el diseño y uso de juegos digitales serios tiene sus bases conceptuales en las teorías del aprendizaje constructivista, en el cual el conocimiento se crea a través de la experiencia mientras se explora el mundo y se realizan actividades [9] [10]. El constructivismo enfatiza la importancia del aprendizaje para construir su propio conocimiento. Al analizar la experiencia del usuario, el aprendizaje es una actividad compleja que requiere gradualidad y necesita varios pasos que deben ser respaldados por varias herramientas y, en general, deben ser guiados por un adulto real, con el fin de que sea significativo para el alumno y se optimicen tiempo y recursos. [11]. En esta postura se concibe al sujeto como capaz de construir su propio conocimiento al interactuar con el objeto de estudio [12].

Dos métodos de aprendizaje parecen complementarios, simples y particularmente útiles para diseñar juegos serios. El primero es el *modelo del aprendizaje experiencial de Kolb*, que sistematiza el trabajo del desarrollo cognitivo sobre el centro del proceso de aprendizaje. El segundo, a su turno, se basa en la *taxonomía de dominios del aprendizaje* [13], mejorado y actualizado posteriormente en un nuevo modelo [14] que es el enfoque cognitivo más popular para la implementación de juegos serios. Los buenos juegos serios y las simulaciones deberían permitir a los usuarios realizar experiencias significativas, por lo tanto, deben estar apoyados en el modelo pedagógico de aprendizaje experiencial.

El primer modelo, basado en la teoría del aprendizaje experiencial (ELT), constituye la base de investigaciones sobre el efecto de los juegos serios en la construcción de nuevo conocimiento. Este modelo multilíneal de desarrollo se basa en la importancia de la experiencia en el proceso de aprendizaje; desde el punto de vista pedagógico, este es un método a través del cual se construye conocimiento mediante técnicas de reflexión a través de las experiencias, teniendo en cuenta los diferentes estilos individuales de aprendizaje [10]. Además, este modelo identifica dos tipos de actividades de aprendizaje: la percepción, es decir, el modo de captar nueva información; y el procesamiento, o el modo en que se procesa y transforma dicha información en algo significativo y utilizable. A lo largo del ciclo de aprendizaje se percibe y procesa información de diferentes maneras, según la etapa.

El segundo modelo se basa en la *taxonomía de dominios del aprendizaje de Bloom* y está diseñado para clasificar los distintos ámbitos del aprendizaje humano en el aspecto cognitivo; este dominio categoriza y ordena habilidades de pensamiento y objetivos. Su taxonomía sigue un proceso del pensamiento, con lo cual se plantea que no puede entenderse un concepto si no se lo recuerda primero; además, no es posible

aplicar conocimientos y conceptos si estos no se entienden. La propuesta original parte de habilidades de pensamiento de orden inferior y va hacia habilidades de pensamiento de orden superior. Bloom describe cada categoría como un sustantivo y las organiza en orden ascendente, de inferior a superior: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación. Sin embargo, en 2001, Anderson y Krathwohl modificaron el modelo usando verbos en lugar de sustantivos para cada categoría y cambiando su secuencia; por lo tanto, en orden ascendente las categorías del nuevo modelo se ordenan así: recordar, comprender, aplicar, analizar, evaluar y crear [13] [15].

El nuevo modelo se diferencia del original porque expone muy bien los componentes para que puedan ser considerados y utilizados; además, en este modelo, los procesos cognitivos pueden ser fácilmente documentados y rastreados en relación con las tareas de instrucción elegida.

El modelo general establece unos niveles de conocimiento: *de los hechos*, o conocimiento fundamental de la disciplina específica; *conceptual*, referido a los principios, teorías o estructuras del área disciplinaria particular; y *procedimental*, o métodos de investigación que ayudan al estudiante a lograr un objetivo específico. Los tres primeros niveles fueron identificados en la obra original de Bloom; mientras que en la versión revisada se añadió un cuarto nivel, denominado *metacognición*, el cual hace referencia al conocimiento estratégico o de reflexión sobre la solución de los problemas planteados.

Desde un punto de vista metodológico, para describir y analizar las características educativas de cada juego —paso fundamental para la selección del diseño—, algunos autores informan sobre su experiencia positiva mediante la taxonomía de Bloom modificada y las etapas de aprendizaje de Kolb [13]. También proponen un esquema para una integración adecuada de los juegos en la educación, con el que apoyan objetivos en diferentes pasos de una educación formal (tales como la motivación inicial, la teoría, la práctica y la verificación). Por lo tanto, los modelos anteriormente propuestos constituirán la base para la construcción del modelo de juego serio en la enseñanza de la física cinemática que se propone en este artículo.

Actualmente existen varios modelos de juegos serios propuestos para el aprendizaje de la física basados en los citados aprendizaje experiencial de Kolb y taxonomía de dominios del aprendizaje. Muñoz *et al.* [16] exploraron a este respecto la creación de un juego de física educativa con una arquitectura de juego diseñada para involucrar al estudiante mientras juega y aprende; además, proporcionaron un método de evaluación de los resultados registrados.

Otro ejemplo de juego serio, esta vez basado en el aprendizaje de la mecánica newtoniana, es propuesto por Juuti *et al.* [17]; en este trabajo se examina el impacto que tiene el uso de un ambiente de aprendizaje virtual para enseñar varios principios de la física, como la gravedad y las colisiones. Adicionalmente, Mohanty *et al.* [18] discuten el uso de varios juegos de video comerciales de la plataforma Sony PlayStation 3 para enseñar física a estudiantes de pregrado; los resultados de esta investigación indican beneficios respecto al desarrollo del pensamiento tridimensional y de habilidades de visualización y conceptualización.

Todos estos ejemplos de juegos que utilizan conceptos de física fueron la base para determinar la importancia de la representación visual de un fenómeno y sus aplicaciones cognitivas en los estudiantes. A partir de ello correspondería diseñar un modelo cuyas etapas guíen el proceso de desarrollo de juegos serios, teniendo en cuenta elementos pedagógicos y las mecánicas y dinámicas propias de ese tipo de aplicaciones, que permitan una mejor empatía en el proceso de aprendizaje de la física cinemática. Ahora bien, a pesar del aumento en el uso de este tipo de tecnologías, no existe un único modelo que permita introducir marcos fáciles para la implementación de un juego serio, orientado a un área temática específica de conocimiento [19]. En este sentido, el desafío de la aplicación de un modelo para el desarrollo de un juego serio consiste en equilibrar contenido pedagógico [20], necesidades del usuario, entretenimiento [21], implementación, evidencia y retroalimentación [22].

La construcción del juego serio propuesto como aporte en este artículo se sustenta en los modelos de aprendizaje discutidos. Se basa, entonces, en un esquema básico de construcción de juegos serios que establece cuatro dimensiones para su implementación teniendo en cuenta el esquema planteado en el modelo que se denomina *marco cuantitativo* [20]: 1) el contexto del aprendizaje; 2) las particularidades del alumno; 3) la representación del juego; y 4) las pedagogías adoptadas. Otro aspecto importante para tener en cuenta se encontró en el modelo exploratorio de aprendizaje [23], en el que la interacción social se convierte en el centro de procesos de aprendizaje más interactivos y atractivos. Por otro lado, con respecto al desarrollo de la motivación del juego, se podrían establecer algunas actividades lúdicas que faciliten el proceso de aprendizaje: al respecto, Staalduinen *et al.* [24] vinculan el diseño del juego y los resultados del aprendizaje en un marco que permite determinar la efectividad de un juego serio en el proceso de desarrollo cognitivo. Sumado a lo anterior, se tomó en cuenta el modelo de diseño centrado en la evidencia (ECD): esto es, un marco de diseño conceptual usado para recolectar datos de evaluación en muchos tipos de formatos, incluidos los juegos digitales.



El proceso metodológico para la construcción de juegos serios para el apoyo del aprendizaje de la física cinemática que se usó en este proyecto recoge cinco aspectos comunes de los modelos citados: contexto; definición y diseño conceptual; diseño de interacción e implementación; validación; y evaluación. Para el desarrollo e implementación, fue necesario contar con cinco equipos de trabajo, encargados de los aspectos pedagógicos, creativos, de diseño gráfico, desarrollo y usabilidad [25].

Definir el material educativo que los estudiantes deberán aprender mientras interactúan con un juego serio requiere establecer los resultados de aprendizaje en función de los objetivos planteados. Los resultados de aprendizaje representan metas que los alumnos deben ser capaces de alcanzar una vez completen con éxito todas las actividades, las cuales se centran en un conjunto específico de tareas que deben completarse. En este sentido, es importante que todas las actividades de aprendizaje promuevan la motivación para que los estudiantes permanezcan interesados e inmersos en el escenario del juego.

## 1. CONSTRUCCIÓN DEL JUEGO SERIO

Para el desarrollo del juego que se expone aquí se tuvieron en cuenta dos tipos de movimiento: en una dimensión, es decir, a lo largo de una línea recta; y en un plano, mediante la simplificación del movimiento parabólico. Se establecieron objetivos relacionados con los temas seleccionados del componente de cinemática del área de física. El esquema de diseño permitió componer actividades y tareas que posibilitaron obtener resultados de aprendizaje frente a los conceptos generales abordados como parte de la temática del juego. Cabe resaltar que el objetivo es presentar en este contexto un desarrollo conceptual cualitativo de los saberes relacionados con los temas, permitiendo al estudiante relacionarse mediante ayudas con las definiciones formales y leyes matemáticas establecidas para cada tema a tratar.

### 1.1 Definición de los objetivos de aprendizaje

El objetivo de aprendizaje define el contenido educativo relacionado con el conocimiento que se desea obtener por medio del cumplimiento de actividades y tareas educativas. Debido a que el conjunto de objetivos constituye una unidad de aprendizaje, cada objetivo trazado incluye la información de los elementos de competencia por alcanzar y sus actividades relacionadas. En el caso particular de esta investigación se asignaron identificadores a cada objetivo, con lo que se establecieron actividades y tareas que especificaron el contenido a desarrollar; estas guardarían relación con el diseño de las actividades lúdicas del sistema propuesto.

Se eligieron cuatro subtemas fundamentales en el estudio del movimiento de una partícula: posición, desplazamiento, movimiento rectilíneo uniforme y movimiento parabólico, y se incluyeron algunos conceptos intrínsecos como trayectoria, distancia y velocidad. De este modo se crearon los objetivos de aprendizaje, presentados en la en la tabla 1 con sus correspondientes elementos de competencia, resultados de aprendizaje, actividades y tareas.

Tabla 1. Elementos del esquema pedagógico planteado para el diseño de un juego serio enfocado a la física cinemática

Objetivo	Elemento de competencia	Resultado de aprendizaje	Actividad	Tarea
Establecer las características generales de los vectores de posición.	Construir el concepto fundamental de coordenadas de un punto y su significado vectorial relacionado con la posición de un objeto.	Usa un marco de referencia para determinar el vector posición de una partícula dada.	Ubicación de una partícula en un marco de referencia y su representación como un vector de posición.	Establecer diferentes posiciones de objetos en un marco de referencia con base en simulaciones de situaciones cotidianas.
Determinar el desplazamiento de un objeto, su trayectoria y distancia total recorrida.	Usar la magnitud fundamental longitud, en forma de camino recorrido y desplazamiento.	Identifica el vector desplazamiento resultante debido al cambio de posición de una partícula desde un punto inicial hasta uno final.	Determinar el desplazamiento experimentado por un objeto y la distancia a un punto de referencia.	Desplazar un objeto a través de un lugar, estableciendo la diferencia entre la cantidad vectorial desplazamiento y la cantidad escalar relacionada con la distancia.
Establecer las características generales del movimiento rectilíneo con velocidad constante.	Analizar el movimiento en línea recta de una partícula mediante el uso del concepto de velocidad.	Determina la velocidad a partir del desplazamiento realizado por un objeto en un tiempo determinado.	Construir el concepto: marco de referencia, sistema de coordenadas, posición, desplazamiento, distancia recorrida, trayectoria.	Mover una partícula en una trayectoria lineal. Estimar la velocidad de una partícula en un instante dado.
Identificar un movimiento parabólico en la naturaleza y las variables que intervienen en él.	Analizar el movimiento compuesto de los cuerpos en direcciones horizontal y vertical.	Identifica las componentes de un movimiento en dos dimensiones. Reconoce características de un movimiento parabólico en la naturaleza.	Se determinará el vector velocidad inicial y ángulo de lanzamiento de una esfera que describe una trayectoria parabólica producto del campo gravitatorio, la cual deberá caer en una posición determinada.	4.1 Realizar una estimación de la posición (x, y) del punto de choque de la esfera con la placa receptora. 4.2 Obtener el módulo de la velocidad inicial y el ángulo de lanzamiento

Fuente: elaboración propia.

Una vez definidos los objetivos se estableció un esquema general de aprendizaje de conceptos relacionados con la física cinemática (figura 1). Esta implementación del modelo pedagógico permitiría establecer un diseño del componente lúdico del juego, así como un esquema de evaluación que favorecería la aplicación del prototipo de juego serio diseñado en un grupo de estudiantes. En el esquema no necesariamente debe

seguirse un nivel jerárquico; en este sentido, el hecho de que algunos de ellos aparezcan en un mismo nivel no presupone ninguna secuencia organizativa.

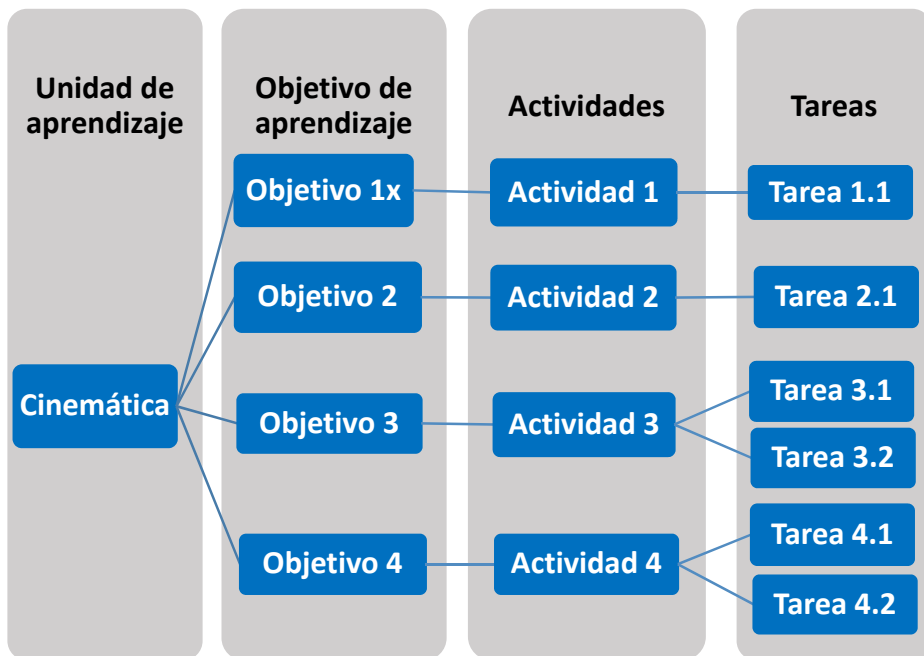


Figura 1. Planteamiento de la unidad de aprendizaje enmarcado en el modelo pedagógico propuesto para el desarrollo de un prototipo

Fuente: elaboración propia.

## 1.2 Diseño del componente lúdico

Como se expuso, el modelo propuesto consta de un diseño pedagógico que involucra objetivos educativos definidos para ciertos subtemas que componen el área de la cinemática; estarán ligados a actividades y tareas educativas de las diferentes temáticas, que a su vez están relacionadas con pruebas lúdicas que constituyen los retos a superar dentro del juego. Lo anterior se hizo de acuerdo con el diseño conceptual propuesto y las herramientas de *software* Unreal Engine 4 [26], y en articulación con el uso de las mecánicas y dinámicas propias del juego, a saber: un conjunto de conocimientos asociados a cada prueba o reto; una historia en la que se enmarca el contenido; el porcentaje de vida del jugador al finalizar cada prueba; el registro de tiempo empleado por el jugador al abordar la prueba, el puntaje obtenido de cada jugador al abordar cada reto de forma adecuada; y el porcentaje de avance en el cumplimiento de cada prueba.

Definidos los objetivos, actividades y tareas específicas, se diseñaron los subniveles, retos y fases que ayudarían a dar cumplimiento a los objetivos de aprendizaje.

En el modelo propuesto (figura 2), los subniveles del juego corresponden a los objetivos educativos; los retos serán las actividades propuestas para la adquisición del conocimiento; y las fases serán las tareas por realizar dentro del juego con el fin de dar cumplimiento al desarrollo del contenido educativo propuesto. La figura 2 muestra la representación gráfica de los elementos del diseño del videojuego educativo que sirvieron en su implementación; la línea que une los retos 2 y 3 indica que estos se relacionan con la contextualización de los conceptos *desplazamiento* y *velocidad* —lo cual es válido debido a que la velocidad está determinada por desplazamiento de una partícula—.

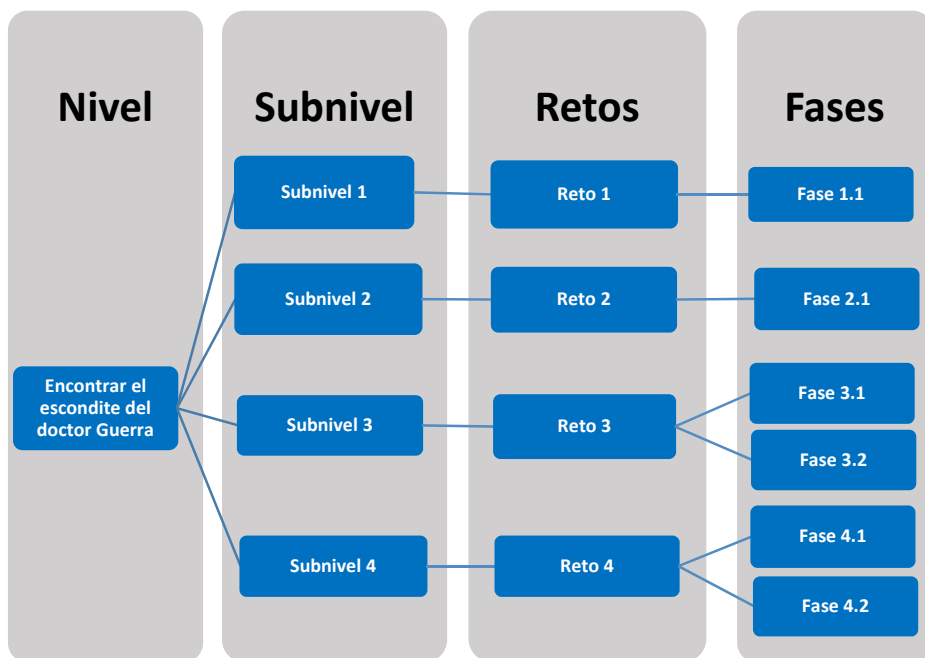


Figura 2. Diseño de la componente lúdica del videojuego educativo orientado al aprendizaje de la física cinemática

Fuente: elaboración propia.

En la tabla 2 se indica la relación directa entre los objetivos propuestos (tabla 1) y el diseño del juego (figura 2). Se especifican allí los elementos de diseño implementados en el desarrollo de los subniveles del juego, incluidos los elementos de competencia, los resultados de los niveles, los retos y las fases del prototipo.

Tabla 2. Implementación del diseño de retos y fases dentro de los subniveles del juego

Subnivel	Elemento de competencia	Resultado del nivel	Retos	Fase
1. Encontrar el mapa del laberinto y la brújula que se encuentran el pueblo de la Isla Kinematic.	Construir el concepto fundamental de coordenadas de un punto relacionado con la posición del personaje principal respecto a las coordenadas marcadas por la brújula dentro del mapa.	Ubicar la posición de la brújula y el mapa con referencia a la puerta de acceso del laberinto.	1. Encontrar la brújula y el mapa del laberinto que se encuentran en la Isla Kinematic.	1.1. Moverse a través del escenario para localizar dentro del pueblo los dos objetos colocados en ciertas posiciones, utilizándolos posteriormente para acceder por una puerta al laberinto.
2. Desplazarse con el personaje principal a través del laberinto ubicando, con la ayuda del mapa, tanto los objetos que recargan la vida del personaje como las criaturas que deberán ser enviadas a la zona de cuarentena.	Desplazarse a través del laberinto haciendo uso del mapa y la brújula, de tal modo que sea posible determinar el concepto desplazamiento del avatar a través de sus cambios sucesivos de posiciones.	Determinar las características del desplazamiento de una partícula en un plano mediante herramientas como una brújula y un mapa.	2. Encontrar los elementos de recarga de energía y criaturas que deberán ser transportadas a cuarentena.	2.1. Desplazar el personaje a través de un laberinto ubicando criaturas y elementos de ayuda mediante el mapa y la brújula.
3. Mover el personaje a través del laberinto, cambiando su velocidad para esquivar los golpes de las criaturas, de tal forma que se haga posible su teletransportación.	Identificar el cambio de velocidad del personaje al huir de las criaturas en una zona determinada.	Moverse a través del laberinto para encontrar la salida.	3. Esquivar los ataques de las criaturas y enviarlas a cuarentena.  Moverse y salir del laberinto	3.2. Recorrer el laberinto utilizando herramientas como la brújula y el mapa de posicionamiento.  3.3. Moverse en la zona de ataque evitando el contacto con las criaturas, enviándolas a la zona de cuarentena.
4. Disparar el cañón para apagar la llama encendida de la gárgola que se encuentra en la entrada del castillo.	Variar la velocidad y el ángulo de lanzamiento de un objeto con el fin de describir una trayectoria parabólica.	Apagar la llama con un objeto que describe una trayectoria parabólica, disparándolo con un cañón.	4. Disparar y apagar la llama para abrir la puerta del castillo.	4.1. Estimar la posición de la llama.  4.2. Modificar la velocidad y ángulo de lanzamiento del proyectil que logre apagar la llama.

Fuente: elaboración propia.

### 1.3 Implementación del diseño

En esta sección se presenta la implementación del modelo pedagógico propuesto en las secciones anteriores para el apoyo a los conceptos de cinemática. Este prototipo sirvió de base para realizar una primera aproximación de la implementación del modelo. Los diferentes retos del juego serio fueron implementados mediante la herramienta libre Unreal Engine 4, la cual permitió el modelado, animación y creación de gráficos tridimensionales. En el prototipo de diseño de cada nivel del juego se sugirió a los desarrolladores usar elementos gráficos siguiendo el modelo mental del usuario, es decir, aquellos que permitieran al jugador la relación entre sistema y mundo real.

Para garantizar la simplicidad y la facilidad de uso, el nivel 1 tiene cuatro subniveles. Inicialmente se presentará al usuario una introducción al juego que detalla la historia y la misión que se deberá cumplir (figura 3a); luego, aparecerá el menú principal (figura 3b), donde aparecerán las opciones que permitirán al usuario ver los controles y las opciones (figura 3c), jugar o cerrar la aplicación. Al terminar la introducción, el avatar (o personaje principal) será transportado por una máquina a la isla y deberá moverse a través del escenario con el fin de encontrar una serie de ayudas que le indicarán los retos a cumplir. El avatar deberá completar las misiones convirtiendo las criaturas a su estado natural, teletransportándolas a zonas de cuarentena mediante un rayo especial, ideando una forma de inactivar el laboratorio de nitro sin comprometer la biodiversidad de la isla, y ubicando al doctor Guerra (la figura 3 muestra la implementación de la historia dentro del juego). Los retos diseñados contendrán escenarios de aventura, estrategia y táctica.

El primer nivel del juego, diseñado para la aplicación del modelo planteado, consiste en encontrar el escondite del citado Doctor Guerra. Se diseñaron a este respecto cuatro subniveles descritos en el modelo de contenido lúdico, los cuales a su vez tienen un reto asociado que permite desarrollar las fases del juego y cumplir el objetivo de llegar al castillo donde se encuentra el villano. La realización de las tareas adaptadas en las fases del juego requerirá que el estudiante maneje el personaje para recorrer los escenarios, los cuales se describen a continuación.

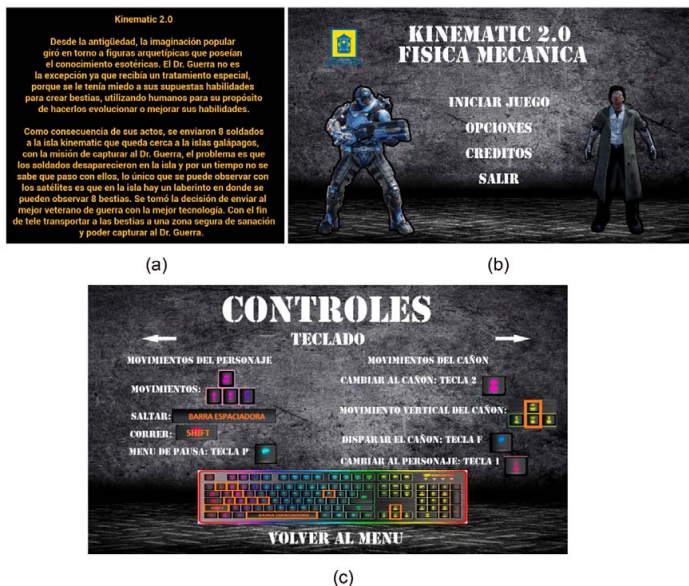


Figura 3. Esquema de inicio del juego. a); historia del juego; b) menú principal; c) controles

Fuente: elaboración propia.

### 1.3 Descripción de los niveles del juego

La aventura comienza cuando el personaje es transportado al primer escenario, donde el jugador deberá desplazarse en busca de pistas que le permitan continuar su misión (figura 4a). Al continuar, encontrará una puerta de acceso a un segundo escenario; un elemento de ayuda le indicará que debe dirigirse al pueblo de la Isla Kinematic y encontrar dos elementos —una brújula y un mapa— que permitirán abrirla. La segunda escena está relacionada con el desarrollo del concepto de posición, mediante el diseño del componente lúdico descrito en la tabla 2 (subnivel 1); la figura 4 (b) muestra el personaje principal desplazándose hacia el pueblo abandonado (en las actividades propuestas se deberá relacionar el cambio de posición del personaje y el concepto de ubicación de un punto en un plano). Finalmente, al encontrar la brújula y el mapa, el personaje deberá acceder a un tercer escenario: un laberinto. Durante la ejecución de la tercera escena se introduce el concepto de desplazamiento, debido a la necesidad del personaje de ubicar objetivos e ítems de ayuda mientras se desplaza por el laberinto (tabla 2, subnivel 2).

Los retos 2 y 3 de sus respectivos subniveles (tabla 2, subniveles 2 y 3) permitirán afianzar los conocimientos de desplazamiento y velocidad; el objetivo del avatar para finalizar esta parte de su misión será localizar las criaturas infectadas, teletransportarlas a cuarentena y salir del laberinto (figura 4c). Para acceder al escondite del villano, el personaje deberá disparar un cañón calculando de forma aproximada la velocidad y ángulo de lanzamiento de un proyectil.

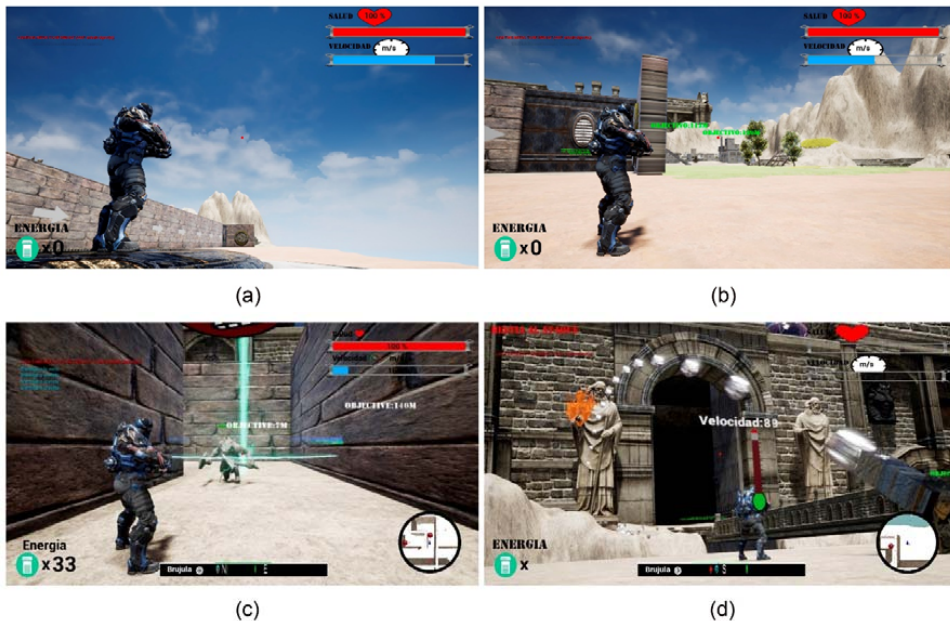


Figura 4. a) escena inicial del juego; b) personaje moviéndose hacia el pueblo de la Isla Kinematic; c) ubicación de un infectado en el laberinto; d) cañón disparando un proyectil mediante la variación de la velocidad y ángulo de lanzamiento

Fuente: elaboración propia.

La ejecución del reto final se encamina a introducir de forma general el concepto de trayectoria parabólica relacionado con el subnivel 4 (tabla 2), en tanto se permite al estudiante variar dos parámetros físicos que le permitirán cumplir su objetivo (figura 4d). Al abrirse la puerta del castillo, el personaje ingresa para capturar al Doctor Guerra. Debido a que el desarrollo del prototipo finaliza en este nivel, aquellos estudiantes que han logrado cumplir con todos los objetivos obtendrán una imagen que los acreditará como ganadores.

### 1.5. Evaluación del modelo propuesto

Se realizó un proceso de valoración basado en la técnica de grupos focales. En esta línea, se llevaron a cabo entrevistas de grupo a cinco profesionales encargados de realizar la evaluación correspondiente a la implementación desarrollada. La actividad fue dirigida por un moderador, encargado de guiar la entrevista colectiva y organizar la discusión en torno a las características y dimensiones del tema propuesto. La técnica de grupo focal se dirigió, en concreto, al cumplimiento de los principios heurísticos de Nielsen y de las pautas [27] [28]. En la tabla 3 se presenta el perfil de los evaluadores, encargados de realizar la inspección sobre el prototipo desarrollado.



Tabla 3. Expertos encargados de la evaluación

Experto	Descripción
Experto 1	Docente en el área de ingeniería: estudios de física.
Experto 2	Ingeniero en el área de ingeniería y telecomunicaciones, estudio de elementos colaborativos en ambientes de aprendizaje
Experto 3	Docente en el área de ingeniería, desarrollo en herramientas de usabilidad.
Experto 4	Docente en el área de ingeniería con profundización en programación, seguridad informática y gestión de proyectos Informáticos
Experto 5	Docente en el área de ingeniería: programación orientada a objetos.

Fuente: elaboración propia.

Los principios heurísticos de Nielsen evaluados fueron visibilidad del estado del sistema; relación entre el sistema y el mundo real; control y libertad de usuario; consistencia y estándares; prevención de errores; reconocimiento, antes que recuerdo; flexibilidad y eficiencia de uso; diseño estético y minimalista; ayuda a los usuarios a reconocer; ayuda y documentación; estos tienen por objetivo garantizar el cumplimiento de las características de usabilidad dentro de aplicaciones interactivas de propósito general [27] [28]. A partir de lo anterior, la sesión de Focus Group estuvo conducida hacia la evaluación del cumplimiento de los diez principios de usabilidad de Nielsen.

Como resultado de esta sesión se obtuvo que los principios heurísticos de visibilidad del sistema, consistencia y estándares, y ayuda y documentación fueron los menos aplicados en el prototipo. Respecto al primero, los evaluadores manifestaron la necesidad de incluir títulos en cada una de las vistas para orientar al usuario sobre el reto a superar. Así mismo, se sugirió la inclusión de un mapa, brújula, velocidad y estado de vida en la parte superior derecha de la pantalla. En cuanto al segundo, se recomendó usar los mismos tipos de fuentes, colores en los títulos y vistas de las diferentes pruebas o retos del juego. Y respecto al tercero se recomendó incluir en pantalla el modo de interacción del usuario con el prototipo de *software*, así como el manejo de ayuda contextual teórica según el reto en que el usuario se encontrara. Los resultados obtenidos en este proceso de evaluación dieron pie a realizar cambios en las diferentes vistas del juego (figura 5). En términos generales, el prototipo desarrollado cumple con las características básicas de usabilidad y accesibilidad para ser usado y apropiado por la población objetivo del presente trabajo.



Figura 5. Modificaciones realizadas con base en los resultados de la etapa de evaluación

Fuente: elaboración propia.

## 2. CONCLUSIONES Y TRABAJOS FUTUROS

- El modelo propuesto para el diseño e implementación de juegos serios toma en consideración aspectos relevantes de la pedagogía y de la usabilidad, los cuales no suelen tomarse en cuenta en el proceso de construcción de aplicaciones. Lo anterior permite que la propuesta de juego serio esté centrada en los objetivos del aprendizaje; ello daría origen a un modelo lúdico que permitiría desarrollar con facilidad los retos dentro del juego, de tal manera que se favorezca la interacción del usuario con los conceptos por aprender.
- Se establecieron herramientas de diseño para subniveles del juego serio, correspondientes a los objetivos de aprendizaje propuestos en el modelo pedagógico. Ello dio origen a una serie de criterios que facilitaron el diseño de los componentes lúdicos adecuados respecto de las actividades y tareas planteadas, mediante su implementación en retos y fases del juego.
- Se implementó el modelo considerándose el diseño conceptual propuesto en esta investigación y la herramienta de *software* Unreal Engine 4. Los escenarios desarrollados se relacionaron con los retos y fases del juego establecidos en los componentes lúdicos.
- El prototipo desarrollado está basado en el modelo propuesto; en él se toman en cuenta elementos relevantes para el funcionamiento del juego como la definición de los objetivos de aprendizaje, el diseño del componente lúdico y el diseño de interfaces. Lo anterior contribuye a que la herramienta construida sirva de apoyo

para los docentes de física de educación secundaria, favoreciendo la lúdica en el aprendizaje de conceptos que cuya asimilación plantea desafíos para los estudiantes.

- Las pruebas de usabilidad, así como la evaluación de expertos, mostraron algunas deficiencias en el diseño gráfico. Por lo tanto, fue necesario mejorar en aspectos como los movimientos del personaje y la ejecución en tiempo real de ciertas acciones.
- Como trabajo futuro se sugiere tomar como base el juego desarrollado y planear nuevas temáticas relacionadas con la física. De igual manera, y con el propósito de realizar una evaluación con un método diferente, se sugiere para valorar la jugabilidad de versiones posteriores del producto con el método EGameFlow: ello que permitirá tener una nueva visión del producto, en la cual se incluyan la jugabilidad y la inmersión del usuario en el juego.

## REFERENCIAS

- [1] L. C McDermott, “A perspective on teacher preparation in physics and other sciences: The need for special science courses for teachers”, *American Journal of Physics*, 58, 734-742, 1990. DOI: <https://doi.org/10.1119/1.16395>
- [2] T. Mayes and S. de Freitas, “Review of e-learning theories, frameworks and models”, en *Joint Information Systems Committee*, pp. 1-44, 2004. Disponible en: <https://researchrepository.murdoch.edu.au/id/eprint/32662/1/review-of-e-Learning-theories.pdf>.
- [3] E. Bonsignore, D. Hansen, A. Pellicone, “Traversing Transmedia Together: Co-designing an Educational Alternate Reality Game for Teens, With Teens”, in: *Proceedings of the The 15th International Conference on Interaction Design and Children*. ACM, New York, pp 11-24, 2016. DOI: 10.1145/2930674.2930712
- [4] J. C. Read y M. Bekker, “The Nature of Child Computer Interaction”, *BCS-HCI '11 Proc 25th BCS Conf.Human-Computer Interact*, pp. 163-170, 2011. Disponible en: <https://pdfs.semanticscholar.org/788d/10837d0c0bd31a4910fe74eb14da167928ca.pdf>.
- [5] G. Triantafyllakos, G. Palaigeorgiou y I. Tsoukalas, “Designing educational software with students through collaborative design games: The We! Design & Play framework,” *Computers & Education*, vol. 56, n° 1, pp. 227-242, 2011. DOI: 10.1016/j.compedu.2010.08.002
- [6] Y. Kafai, K. Peppler, “Developing Gaming Fluencies with Scratch”, en *Games, Learning, and Society: Learning and Meaning in the Digital Age*; C. Steinkuehler, K. Squire y S. Barab, eds., Cambridge: Cambridge University Press, pp. 355-380, 2012. DOI: 10.1017/CBO9781139031127.026.
- [7] R. Khaled, A. Vasalou, “Bridging serious games and participatory design”, *Int J Child-Computer Interact*, n.º 2, pp. 93-100, 2014. DOI: 10.1016/j.ijcci.2014.03.001.

- [8] G. Kalmpourtzis, "Developing kindergarten students' game design skills by teaching game design through organized game design interventions," *Multimedia Tools and Applications*, vol. 78, n.º 14, pp. 20485-20510, 2019. DOI: 10.1007/s11042-019-7393-y.
- [9] J. Dewey, *How we Think*, Nueva York: D.C. Heath & Co. Publishers, 1933.
- [10] D. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Nueva Jersey: Pearson Education, 2015.
- [11] J. P. Rowe, L. R. Shores, B. W. Mott y J. C. Lester, "Individual Differences in Gameplay and Learning: A Narrative-Centered Learning Perspective", en *Proceedings of the Fifth International Conference on the Foundations of Digital Games (FDG)*, Monterrey - California, 2010. DOI: 10.1145/1822348.1822371.
- [12] A. M. Al-Huneidi y J. Schreurs, "Constructivism Based Blended Learning in Higher Education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 7, n.º 1, pp. 4-9, 2012. DOI: 10.3991/ijet.v7i1.1792.
- [13] J. Baalsrud, F. Bellotti, R. Berta, M. Carvalho, A. De Gloria, E. Lavagnino, R. Nadolski y M. Ott, "Field assessment of Serious Games for Entrepreneurship in Higher Education," *Journal of Convergence Information Technology*, vol. 8, n.º 13, pp. 1-12, 2013.
- [14] L. Anderson y D. Krathwohl, *A taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy*, Nueva York: Addison Wesley Longman, 2001.
- [15] B. S. Bloom (ed.), *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Londres: Longman, pp. 201-207, 1956. DOI: 10.1177/001316445601600310.
- [16] K. Muñoz, J. Noguez, P. McKevitt, L. Neri, V. Robledo-Rella y T. Lunney, "Adding Features of Educational Games for Teaching Physics" en *39th ASEE/IEEE Frontiers in Education Conference*, University of Ulster, Magee, pp. 18-21, 2009. DOI: 10.1109/FIE.2009.5350630.
- [17] K. Juuti, L. Lavonen y V. Meisalo, "Learning Newtonian Mechanics in Virtual and Real Learning Environments in Grade 6 in a Finnish Primary School", en *Proceedings of the IASTED International Conference*, University of Helsinki, pp. 51-68, 2007.
- [18] S. Mohanty y S. Cantu, "Teaching Introductory Undergraduate Physics Using Commercial Video Games", *Physics Education*, vol. 46, n.º 5. DOI: 10.1088/0031-9120/46/5/009.
- [19] R. Amory y R. Seagram, "Educational game models: conceptualization and evaluation", *South African Journal of Higher Education*, vol. 17, n.º 2, pp. 206-217, 2003. DOI: 10.4314/sajhe.v17i2.25314.
- [20] S. de Freitas y M. Oliver, "How can exploratory learning with games and simulations within the curriculum be most effectively evaluated", *Computers and Education*, vol. 46, pp. 249-264. 2006. DOI: <https://doi.org/10.1016/j.compedu.2005.11.007>.

- [21] M. Zyda, J. Hiles, A. Mayberry, C. Wardynski, M. Capps, B. Osborn, R. Shilling, M. Robaszewski y M. Davis, “Entertainment R&D for defense”, *Computer Graphics and Applications, IEEE*, vol. 23, n.º 1, pp. 28-36, 2003. DOI: 10.1109/MCG.2003.1159611.
- [22] J. Groff, J. Clarke-Midura, E. Owen E, L. Rosenheck y M. Beall, “Better Learning in Games. A Balanced Design Lens for a New Generation of Learning Games”, en *Learning Games Network, MIT Education Arcade*, pp. 6-20, 2015.
- [23] S. de Freitas y T. Neumann, “The use of ‘exploratory learning’ for supporting immersive learning in virtual environments”, *Computers and Education*, vol. 52, n.º 2, pp. 343- 352, 2009. DOI: 10.1016/j.compedu.2008.09.010.
- [24] J. Staalduinen y S. de Freitas, “A game-based learning framework: Linking game design and learning outcomes”, en M. Khyne (ed.), *Learning to Play: Exploring the Future of Education with Video Games*, New York: Peter Lang, pp. 29-54, 2011. DOI: 20.500.11937/26899.
- [25] C. Aguilar, G. Chanchí y M. Vidal, “Definición de un proceso metodológico para la construcción de juegos serios para el apoyo de la enseñanza de la física cinemática”, *Publicaciones e Investigación*, vol. 12, n.º 1, pp. 35-48, 2018. DOI: 10.22490/25394088.2815.
- [26] “Unreal Engine 4. Epic Game” [internet]. Disponible en: <https://www.unrealengine.com/en-US/what-is-unreal-engine-4>, 2018.
- [27] J. Nielsen. “Enhancing the explanatory power Definición de un proceso metodológico para la construcción de juegos serios para el apoyo de la enseñanza de la física cinemática of usability heuristics in Proceedings of the SIGCHI”, en *Conference on Human factors in computing systems celebrating interdependence*, Nueva York, pp. 152-158, 1994. DOI: 10.1145/191666.191729.
- [28] J. Nielsen, “Usability inspection methods”, en *Conference companion on Human*, Nueva York, 1994. DOI: 10.1145/259963.260531.



# Influencia de la posición de impresión y la densidad de relleno en las propiedades mecánicas de probetas fabricadas en ABS\*

*Manuel José Carvajal Loaiza\*\**

*Pablo González Díaz\*\*\**

*Carlos Alberto Mejía Blandón\*\*\*\**

*Liliana Marcejal Bustamante Góez\*\*\*\*\**

*Junes Abdul Villarraga Ossa\*\*\*\*\**

Recibido: 28/06/2019 • Aceptado: 02/03/2020

<https://doi.org/10.22395/riium.v19n37a9>

## **Resumen**

En este artículo se investiga la influencia de la posición de impresión y la densidad del relleno sobre el módulo de elasticidad, el límite de fluencia, el esfuerzo máximo y la tenacidad de probetas elaboradas mediante modelado por deposición fundida. Se elaboraron probetas de ABS en dos posiciones diferentes *XY* y *XZ*, cuyos ángulos *raster* fueron  $-45^{\circ}/45^{\circ}$  y  $0^{\circ}/90^{\circ}$ , respectivamente, las cuales se sometieron a ensayos de tracción para calcular sus propiedades mecánicas y evidenciar la incidencia de los parámetros estudiados. Los hallazgos permitieron identificar que las propiedades mecánicas de las piezas terminadas difieren de las del material base (filamento); por lo tanto, según la aplicación final de la parte, es necesario establecer criterios de fabricación que permitan obtener piezas con mayores expectativas de vida útil para una carga dada.

**Palabras clave:** ABS; manufactura aditiva; FDM; ensayo tensión; propiedades mecánicas; impresión 3D.

---

\* Artículo Derivado de una investigación numérica y experimental acerca de los mecanismos de tenacidad de las estructuras tipo Bouligand.

\*\* Estudiante de Ingeniería Mecánica, Grupo Diseño Mecánico, Universidad de Antioquia. Correo electrónico: manuelj.carvajal@udea.edu.co. Orcid: <https://orcid.org/0000-0003-1010-9045>.

\*\*\* Estudiante de Ingeniería Mecánica, Grupo Diseño Mecánico, Universidad de Antioquia. Correo electrónico: pablo.gonzalezd@udea.edu.co. Orcid: <http://orcid.org/0000-0002-9307-883X>.

\*\*\*\* Magister en ingeniería y candidato a doctor en ingeniería de materiales. Profesor del Departamento de Ingeniería Mecánica, Grupo Diseño Mecánico, Universidad de Antioquia. Correo electrónico: carlos.mejiab@udea.edu.co. Orcid: <http://orcid.org/0000-0002-4875-5993>.

\*\*\*\*\* Universidad de Antioquia, Magister en Ingeniería Candidato a Doctorado en Ingeniería de Materiales, Profesor Departamento de Ingeniería Mecánica, Grupo Diseño Mecánico, Correo electrónico: liliana.bustamante@udea.edu.co. Orcid: <http://orcid.org/0000-0002-1150-1252>.

\*\*\*\*\* Magister en ingeniería mecánica. Doctorado en Ingeniería y Ciencia de los Materiales, Profesor del Departamento de Ingeniería Mecánica, Grupo Diseño Mecánico, Universidad de Antioquia. Correo electrónico: jun.es.villarraga@udea.edu.co. Orcid: <http://orcid.org/0000-0002-7893-5362>.

## Influence of the Printing position and the Infill Density on the Mechanical Properties of 3D Printed ABS Specimens

### **Abstract**

In this paper, the influence of the printing position and the infill density on the elastic modulus, yield strength, ultimate strength, and toughness are researched in specimens built by fused deposition modeling (FDM). ABS specimens were built in two positions,  $XY$  and  $XZ$ , in which the angle raster was  $-45^\circ/45^\circ$  and  $0^\circ/90^\circ$ , respectively. These specimens were tested to find its mechanical properties, and the studied parameters' influence was demonstrated. The discoveries allowed to identify that the mechanical properties of the final product differ from the ones in the base material (filament). In conclusion, to make parts by FDM, it is important to establish the mechanical properties of the final product, because these can differ substantially from the filament properties. Therefore, depending on the final application of the part, it will be necessary to established fabrication criteria that allow to build parts with more lifespan for a given load.

**Keywords:** ABS; additive manufacturing; FDM; tensile test; mechanical properties; 3D printing.

## Influência da posição de impressão e a densidade de enchimento nas propriedades mecânicas de provetas fabricadas em ABS

### **Resumo**

Neste artigo, pesquisa-se a influência da posição de impressão e a densidade do enchimento sobre o módulo de elasticidade, o limite de fluência, o esforço máximo e a tenacidade de provetas elaboradas mediante modelagem por deposição fundida. Elaboraram-se provetas de ABS em duas posições diferentes  $XY$  e  $XZ$ , cujos ângulos *raster* foram  $-45^\circ/45^\circ$  e  $0^\circ/90^\circ$ , respectivamente, as quais se submeteram a ensaios de tração para calcular suas propriedades mecânicas e evidenciar a incidência dos parâmetros estudados. As descobertas permitiram identificar que as propriedades mecânicas das peças terminadas diferem das do material base (filamento); portanto, segundo a aplicação final da parte, é necessário estabelecer critérios de fabricação que permitam obter peças com maiores expectativas de vida útil para uma carga dada.

**Palavras-chave:** ABS; manufatura aditiva; FDM; ensaio tensão; propriedades mecânicas; impressão 3D.



## INTRODUCCIÓN

La popularidad del uso de manufactura aditiva (AM por sus iniciales en inglés) o impresión 3D ha aumentado para la fabricación de productos personalizados en las industrias agrícola, automotriz y aeroespacial, en la ingeniería, en los sistemas biológicos y en las cadenas de suministro de alimentos [1] [2]. La AM es un proceso asistido por computador que fabrica objetos 3D físicos capa por capa, a partir de datos de un modelo sólido virtual (CAD), con materiales metálicos, plásticos, cerámicos, compuestos o biológicos. En la actualidad existen diversos sistemas disponibles de manufactura aditiva en el mercado, entre los que se encuentran modelado por deposición fundida (FDM por sus iniciales en inglés), deposición directa de metal (DMD), sinterización selectiva por láser (SLS), modelado por inyección de tinta (IJM) y estereolitografía (SLA), entre otros [3]. La adopción de estas tecnologías ha propiciado el diseño de piezas de libre acceso, sobre las cuales la información es limitada debido, por un lado, a la falta de conocimiento sobre el desempeño final de las piezas durante el servicio [4] [5] y por otro, a la falta de soporte para el *software* de diseño empleado. Las piezas construidas por AM poseen propiedades mecánicas anisotrópicas que son inherentes a este proceso de fabricación. Los datos de las propiedades reportadas por las diversas compañías que fabrican estas tecnologías no son comparables entre sí, debido a que no brindan suficiente información sobre los parámetros del proceso. Por lo tanto, el uso de partes impresas en 3D requiere el reporte público de los datos de las propiedades mecánicas finales de la pieza [6] para una aplicación específica.

Una de las tecnologías más utilizadas para la fabricación de piezas mediante AM es el FDM, ya que es una de las más económicas y posee un sistema de extrusión flexible [7]. El FDM es un proceso en el cual un filamento, generalmente de acrilonitrilo butadieno estireno (ABS) o ácido poliláctico (PLA), pasa a través de una resistencia y llega a un estado semifundido; luego, una boquilla es alimentada por este filamento y se deposita de forma entrecruzada en la pieza parcialmente construida. Los materiales extruido y recién depositado se fusionan entre sí, y luego la boquilla se mueve en el plano *XY* conforme deposita material según la geometría de la pieza. La plataforma que contiene la pieza se mueve de forma vertical en el eje *Z* para depositar una nueva capa sobre la anterior. Después de un periodo de tiempo —por lo general, varias horas—, la boquilla habrá depositado una representación física completa del archivo CAD original (figura 1) [8] [9].

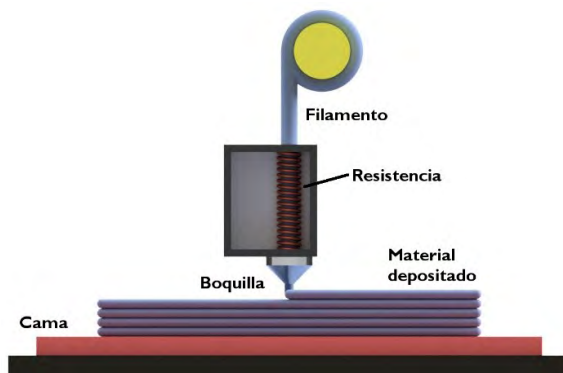


Figura 1: Esquematación de FDM

Fuente: elaboración propia.

La tecnología de FDM reviste interés para la academia y la industria debido a su capacidad de fabricar geometrías complejas a costos relativamente bajos, su facilidad de utilización y su compatibilidad con materiales reciclables [10] [11] [12]. La calidad y las propiedades mecánicas de las piezas construidas mediante FDM dependen de una selección adecuada de parámetros que intervienen en el proceso de impresión (figura 2). La incidencia de estas variables sobre el acabado y el desempeño de la pieza final requieren ser estudiados para obtener las condiciones óptimas durante el proceso de impresión.

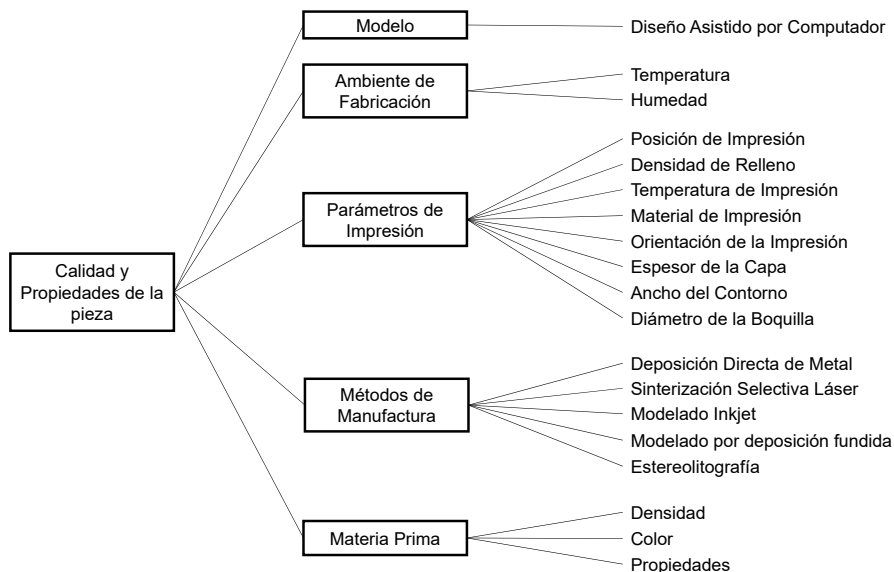


Figura 2: Parámetros que modifican la calidad y propiedades de la pieza en FDM

Fuente: adaptada de [9].

En la actualidad, el principal uso que se da a las piezas fabricadas mediante FDM corresponde a ayudas visuales y presentación de prototipos, ensambles o modelos educativos. Debido a los avances tecnológicos en la técnica de FDM y la disminución de costos, se ha incrementado el número de usuarios domésticos que utilizan esta técnica y han surgido muchos *software* de código libre para el uso de estas impresoras —Cura y Repetier-host, por nombrar solo algunos—, los cuales permiten modificar parámetros como los que se muestran en la figura 3; estos influyen en las propiedades mecánicas, el acabado y la calidad de la pieza final.

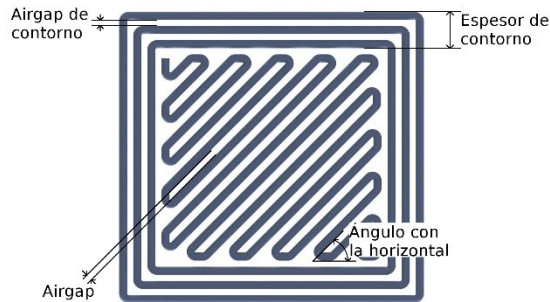


Figura 3: Algunos parámetros de impresión usados en el programa CURA

Fuente: elaboración propia.

El diseño de piezas con FDM que involucren mejores acabados y buen desempeño mecánico, y en cuya fabricación se empleen las herramientas y ventajas de la técnica al tiempo que se minimicen sus limitaciones, es una tarea que está ganando atención por parte de la comunidad científica y académica. Además, visto que la mayoría de los usuarios domésticos no tienen forma de verificar la resistencia de las piezas que se imprimen; que tampoco hay suficiente información disponible sobre las propiedades mecánicas finales y del desempeño de las piezas impresas; y que no existen investigaciones que involucren la tenacidad de las partes fabricadas por FDM, resulta pertinente analizar la influencia de algunos parámetros de impresión que intervienen en el proceso FDM en las propiedades mecánicas de las piezas finales.

Con lo anterior, este artículo describe un proceso de análisis respecto de los cambios en las propiedades mecánicas de probetas impresas en ABS sometidas a esfuerzos de tensión, al variarse los parámetros de densidad del relleno, la orientación de la pieza al imprimirla y su ángulo *raster* (ángulo de la trayectoria del filamento con respecto al eje x), lo cual implica un cambio de dirección en las fibras impresas. La orientación hace referencia a la forma en que la pieza se ubica con respecto a los ejes *X*, *Y* y *Z* de la plataforma; y la densidad de relleno se define como el porcentaje de polímero impreso al interior de la pieza con respecto al espacio total dentro de ella: 100 % indica que el relleno

de la pieza es solo polímero y no hay ningún espacio entre las fibras. La información recolectada tiene el fin de proporcionar criterios de resistencia de los productos finales a los usuarios de impresoras controladas por programas con código abierto.

## 1. MATERIALES Y MÉTODOS

Para determinar las propiedades mecánicas de las piezas impresas y su dependencia a los parámetros de impresión que es posible alterar en una impresora 3D, se analizaron la incidencia de la orientación de la pieza a imprimir y el porcentaje de relleno (densidad) en la resistencia a la tracción, el módulo de rigidez y la tenacidad. Las piezas se imprimieron en dos posiciones y orientaciones específicas,  $-45^{\circ}/45^{\circ}$  ( $XY$ ) y  $0^{\circ}/90^{\circ}$  ( $XZ$ ), (figura 4), y las densidades de relleno fueron 30 %, 50 %, 75 % y 100 %.

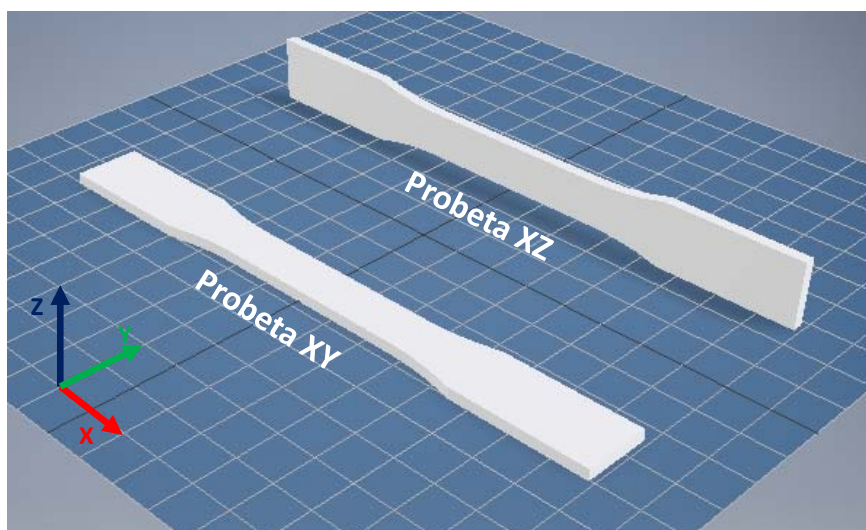


Figura 4. Modelo y posición de espécimen de probetas

Fuente: elaboración propia.

Los demás parámetros se mantuvieron constantes para todas las configuraciones de impresión. La altura de capa fue de 0,3 mm; el espesor de contorno, 1,2 mm; la velocidad de impresión, 30 mm/s; y el diámetro de boquilla, 0,4 mm (la tabla 1 muestra las configuraciones utilizadas). En total se imprimieron 101 piezas.

Se fabricó un mayor número de probetas en la posición  $XZ$  debido a que de tal modo era posible imprimir más probetas de forma simultánea. Además, se desecharon las piezas que después de haber sido elaboradas presentaban algún defecto geométrico. Para esta investigación no se tuvo en cuenta la incidencia de las condiciones ambientales en la impresión, las cuales pueden afectar levemente los resultados [13] [14].

Tabla 1. Variación de parámetros de las probetas impresas

Tipo de probeta	Porcentaje de relleno	Ángulo raster
30XY	30 %	-45°/45°
30XZ		0°/90°
50XY	50 %	-45°/45°
50XZ		0°/90°
75XY	75 %	-45°/45°
75XZ		0°/90°
100XY	100 %	-45°/45°
100XZ		0°/90°

Fuente: elaboración propia.

Se utilizó una impresora Ultimaker 2 [15] junto con el *software* de código abierto Cura 2.4, recomendado para ese modelo. El polímero usado en las impresiones fue un filamento de ABS SKU: ABS32 natural blanco, con una resistencia última a la tensión de 33 MPa, un módulo elástico de 2,2 GPa y una elongación de 6 %. Se diseñaron las probetas conforme a la norma ASTM D638 [16] in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon ( tipo I: cada espécimen tenía 3,3 mm de espesor. Luego, se hizo el modelado geométrico en el programa Autodesk Inventor 2017 (figura 4) y se exportó lo anterior al formato STL (*stereo lithographie*). La temperatura de extrusión fue de 210 °C y la cama fue precalentada a 85 °C para favorecer la adhesión del material a ella.

Los ensayos de tracción se practicaron de acuerdo con la norma ASTM D638 para probeta tipo I en una máquina Shimadzu AGX Plus; se empleó una celda de carga de 500 kg clase 1 con una resolución de 1/5000. La velocidad utilizada para el ensayo fue de 5 mm/min, y para determinar el módulo elástico del material para cada configuración de la probeta se utilizó un extensómetro axial Epsilon 3542-050M-020-ST. En la figura 5 se muestra el montaje del ensayo.

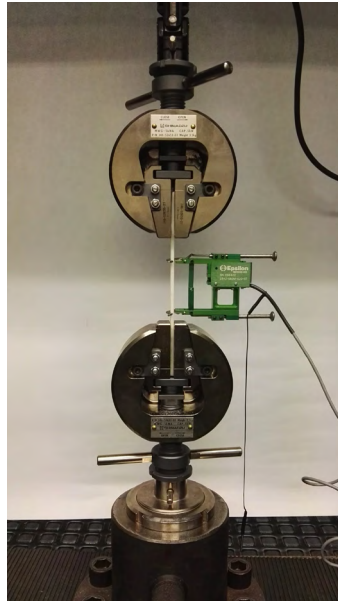


Figura 5: Montaje de ensayo de tracción

Fuente: elaboración propia.

## 2. RESULTADOS Y DISCUSIÓN

En esta investigación se analizó la variación del módulo de rigidez con respecto al porcentaje de relleno de la probeta y de la orientación en la cama de impresión; también se obtuvieron valores para los esfuerzos de fluencia y máximos en cada uno de los casos, al tiempo que se calcularon los valores de tenacidad y se relacionaron con los valores de esfuerzos.

En la figura 6 se presentan las curvas de esfuerzo-deformación para cada orientación de impresión y las densidades de relleno de 30 %, 50 %, 75 % y 100 %. La configuración de impresión *XY* se representa con líneas de color azul, mientras que *XZ* está representada con líneas rojas. En esta figura se observa que el módulo de elasticidad y la resistencia máxima son afectados por la posición de impresión de la probeta. En esta línea, las probetas fabricadas en posición *XY* presentan un rendimiento a tracción inferior, es decir, poseen un esfuerzo máximo a la tensión y un módulo de elasticidad menor que en el caso de las probetas elaboradas en posición *XZ*, lo cual concuerda con los resultados presentados en otras investigaciones [1] [8] [9] [14].

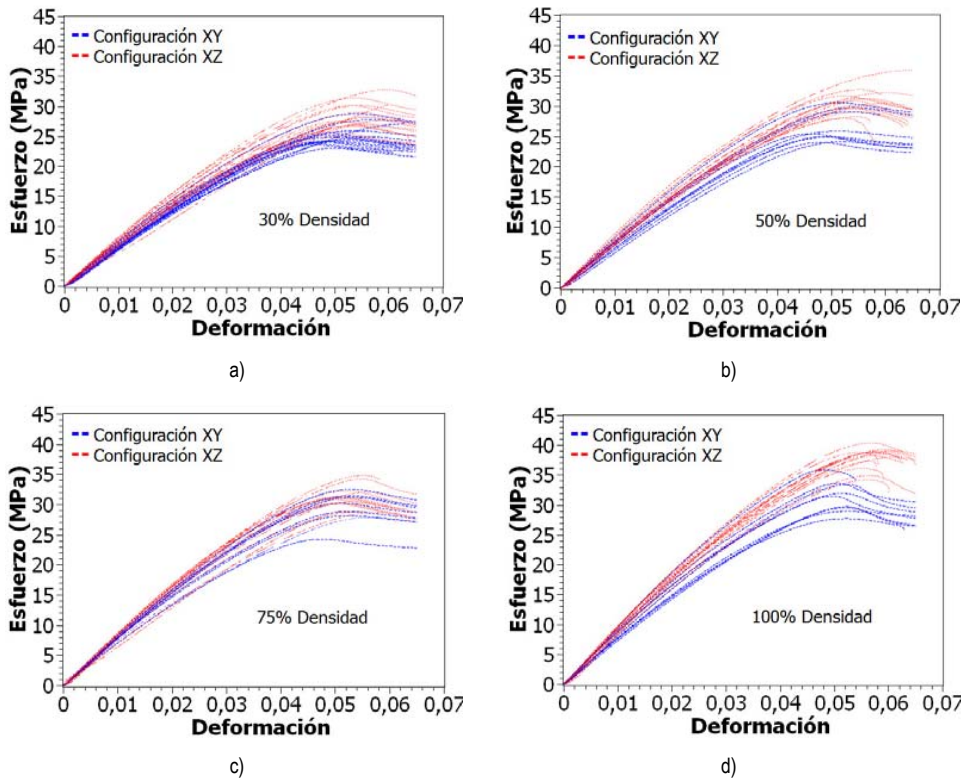


Figura 6: Gráficas esfuerzo vs deformación según densidad: a) 30 %; b) 50 %; c) 75 %; d) 100 %

Fuente: elaboración propia.

Los resultados de los módulos de elasticidad y los valores de las tenacidades se observan en la tabla 2. Respecto a las posiciones de impresión  $X$  y  $XZ$  se encontró que los valores más altos de módulo de elasticidad se presentan en la posición  $XZ$ . Además, las probetas con densidades superiores al 75 %, al margen de su posición de impresión, obtienen los valores de  $E$  más altos, por encima de 1,5 GPa, lo cual representa una disminución del 5 % con respecto al valor del módulo de elasticidad del material sin extruir; para las demás configuraciones estudiadas, esta variación llegó hasta el 14 %. Estas observaciones son similares a las reportadas por Behzad *et al.* [17], quienes realizaron un estudio con ABS variando la configuración de impresión y el número de capas. En otros estudios se han encontrado módulos de elasticidad para el ABS que varían de 1,0 a 1,9 GPa [14] [18]. Se observa entonces una tendencia al aumentar del valor del módulo con el incremento del porcentaje de relleno. En algunos casos se observó discrepancia en los resultados, pero esto pudo deberse a la falta de control sobre la temperatura ambiente del cuarto de impresión, por cuanto esta afecta la adhesión de la capa de material depositado por la boquilla [9] [19].

Tabla 2. Valores promedio del módulo de elasticidad y de la tenacidad para cada posición de fabricación y porcentaje de relleno

Posición de fabricación	Relleno de la probeta			
	30 %	50 %	75 %	100 %
$E_{XY}$ (GPa)	1,44 ± 0,04	1,39 ± 0,04	1,58 ± 0,14	1,62 ± 0,1
$E_{ZY}$ (GPa)	1,49 ± 0,04	1,48 ± 0,02	1,78 ± 0,09	1,89 ± 0,12
$T_{XY}$ (MJ/m <sup>3</sup> )	1,34 ± 0,33	1,99 ± 0,8	1,26 ± 0,12	1,31 ± 0,1
$T_{ZY}$ (MJ/m <sup>3</sup> )	1,61 ± 0,35	1,64 ± 0,52	1,24 ± 0,15	1,48 ± 0,12

Fuente: elaboración propia.

En la figura 7 se observan los valores promedio y las desviaciones estándar obtenidos para los valores de los esfuerzos de fluencia y máximos de los diferentes porcentajes de relleno y posiciones estudiadas; para calcularlos se empleó la metodología planteada por la norma ASTM D638 [13]. Se observa que la posición XY ostenta menores valores de esfuerzos máximos y de fluencia, lo cual concuerda con los resultados reportados por Sung-Hoon *et al.* [8], Wang *et al.* [20], Sood *et al.* [21] y Rayegani y Onwubolu [22] que realizaron experimentos con cambios en el ángulo *raster* y orientación, y Tymrak *et al.* [14], que experimentaron bajo condiciones ambientales reales. En estos artículos también se pudo apreciar que al aumentar el porcentaje de relleno de las probetas, los valores de los esfuerzos también aumentan. Para ninguna de las posiciones estudiadas se logró alcanzar los valores reportados para el material base (filamento) de la extrusión; como se mencionó, este es uno de los factores que causa incertidumbre frente a las piezas fabricadas mediante FDM [6], y esto se hace más notorio al fabricar piezas con un porcentaje de relleno menor.

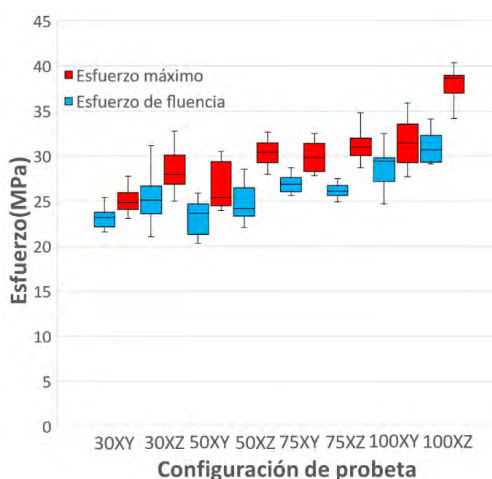


Figura 7. Esfuerzos de fluencia y máximo para cada una de las probetas ensayadas

Fuente: elaboración propia.



La relación entre la tenacidad y el módulo de elasticidad para las probetas ensayadas se puede observar en la figura 8; esta relación es importante en aplicaciones en que no solo se busca resistencia mecánica, sino la combinación de las dos propiedades. En estas relaciones se observa que, para las probetas con menor valor de módulo elástico, se obtiene más dispersión en los datos de la tenacidad, es decir, se genera mayor incertidumbre en el comportamiento global de la pieza impresa; además, se obtienen valores promedio ligeramente mayores para estas mismas probetas.

A partir de porcentajes de relleno de 75 % se puede decir que la tenacidad permanece constante, lo cual puede atribuirse a que los *airgaps* y la densidad de relleno determinan la dimensión de la zona de contacto entre los filamentos y las capas [23]. Por lo tanto, cuanto más densa la pieza, más juntos estarán los filamentos, lo que incrementa la resistencia de la unión de los filamentos y, por ende, la resistencia de la pieza final.

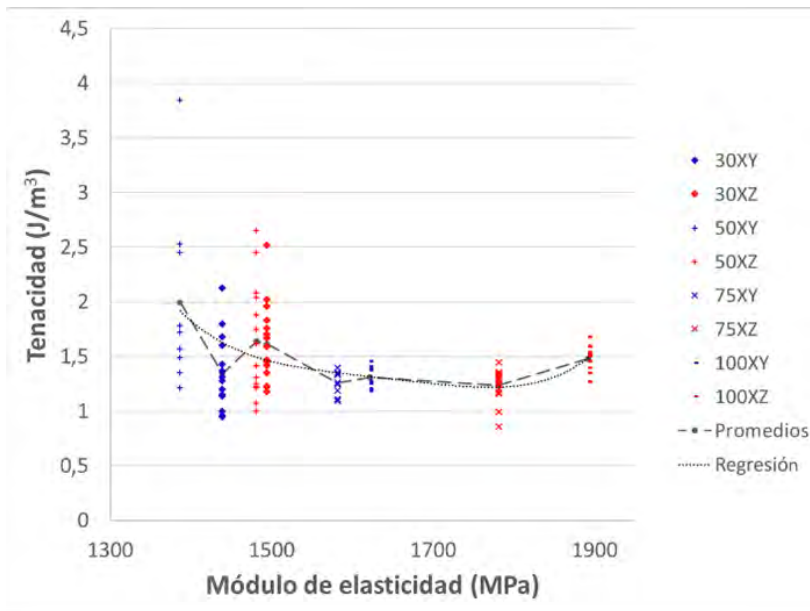


Figura 8: Relación entre la tenacidad y el módulo de elasticidad

Fuente: elaboración propia.

En la figura 9 se observan los modos de falla para las probetas: todas fallaron en la dirección transversal. Además, se percibe el efecto del blanqueamiento (*whitening*) [24] [25] [26] reportado para este tipo de materiales, causado por la deformación localizada: esta consiste en que algunas fibras se tornan de un color más claro que el mostrado por el original en la pieza fabricada. Se nota que las probetas impresas en la orientación *XZ* presentan una fractura frágil, mientras que las impresas en la orientación *XY* presentan fractura dúctil. También se observa que las probetas impresas

en  $XZ$  evidencian mecanismo de retraso al crecimiento de la grieta, lo que se traduce en un aumento de la energía de fractura. Esto se aprecia en los datos de la figura 8: se observa allí que se pueden obtener mayores valores de tenacidad para las probetas con menor densidad de relleno, efecto que puede ser ocasionado por la orientación de las fibras; y que la sección transversal efectiva de estas últimas cambia de acuerdo con la orientación de impresión, datos que guardan relación con los reportados por Sung-Hoon [8].

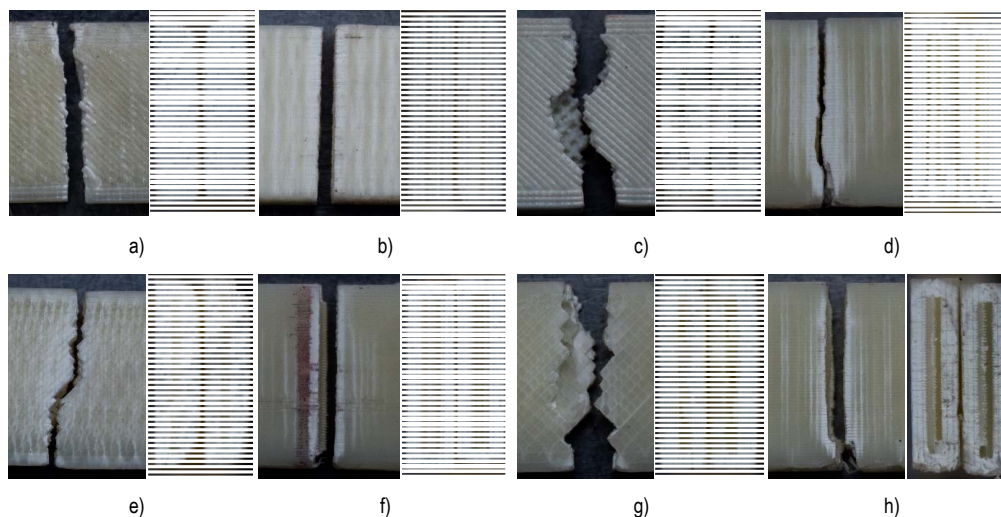


Figura 9. Modos de falla para las diferentes posiciones y densidades: a) 100 % orientación  $XY$ ; b) 100 % orientación  $XZ$ ; c) 75 % orientación  $XY$ ; d) 75 % orientación  $XZ$ ; e) 50 % orientación  $XY$ ; f) 50 % orientación  $XZ$ ; g) 30 % orientación  $XY$ ; h) 30 % orientación  $XZ$

Fuente: elaboración propia.

### 3. CONCLUSIONES

Como se ha dicho en secciones previas, el objeto de la investigación mostrada fue revisar la incidencia de la posición de impresión y la densidad de relleno en las propiedades mecánicas de ABS impreso por medio de FDM con dos tipos de orientaciones y cuatro densidades. El análisis de las curvas esfuerzo-deformación indicó que tanto la posición de fabricación como el porcentaje de relleno tienen un impacto en el rendimiento de las propiedades mecánicas de la pieza elaborada. La posición  $XZ$  y el 100 % de densidad presentaron los mayores valores para los módulos de rigidez y los esfuerzos; mientras que para valores de densidad inferiores al 50 % se obtuvieron los mayores registros de tenacidad.

Al elaborar piezas con porcentajes de relleno inferiores al 75 % las propiedades mecánicas de rigidez, límite de fluencia, límite último y tenacidad presentan variaciones excesivas (que pueden superar el 80 %), las cuales no permiten establecer con claridad criterios de diseño y de falla. Por tal motivo, es recomendable hacer pruebas para las piezas elaboradas por manufactura aditiva, a fin de establecer su comportamiento en el entorno real de trabajo.

## RECOMENDACIONES

Se debe establecer un criterio de funcionamiento principal; es decir, si se desea una pieza con una buena relación de resistencia mecánica, es aconsejable emplear densidades mínimas del 75 % y que la fuerza aplicada sea paralela a la dirección de las fibras en la pieza fabricada. Si, por el contrario, la aplicación necesita una mayor tenacidad, se recomienda emplear densidades inferiores al 50 % y que la dirección de las fibras de la pieza sea de 45° con respecto a la dirección de la carga aplicada.

## AGRADECIMIENTOS

Los autores de esta investigación agradecemos al Servicio Nacional de Aprendizaje (SENA) y su programa Red Tecnoparque - nodo Medellín; y especialmente al ingeniero Ramiro Isaza Escobar, gestor de línea ingeniería y diseño, por brindarnos su apoyo en la fabricación de las probetas. De igual modo, agradecemos a Walter Villaneda, técnico del Laboratorio de Resistencia de Materiales, por su colaboración en la realización de los ensayos, y al estudiante Juan Fernando Cucuyame por su asistencia en la adquisición de las imágenes.

## REFERENCIAS

- [1] Ö. Keleş, C. W. Blevins, K. J. Bowman y Ö. Kele, "Effect of build orientation on the mechanical reliability of 3D printed ABS", *Rapid Prototyp. J.*, vol. 23, n.º 2, pp. 320-328, 2017. DOI: 10.1108/RPJ-09-2015-0122.
- [2] W. Gao *et al.*, "The status, challenges y future of additive manufacturing in engineering", *Comput. Des.*, vol. 69, pp. 65-89, 2015. DOI: 10.1016/j.cad.2015.04.001.
- [3] Y. Huang, M. C. Leu, J. Mazumder y A. Donmez, "Additive Manufacturing: Current State, Future Potential, Gaps and Needs, and Recommendations", *J. Manuf. Sci. Eng.*, vol. 137, n.º 1, p. 014001, 2015. DOI: 10.1115/1.4028725.
- [4] D. H. Stamatis, *Failure mode and effect analysis : FMEA from theory to execution*, Milwaukee: ASQ Quality Press, 2003. DOI: 10.1080/00401706.1996.10484424.
- [5] G. S. Wasserman, *Reliability verification, testing and analysis in engineering design*, Nueva York: Marcel Dekker, 2002. DOI: 10.1201/9780203910443.

- [6] I. Gibson, D. W. Rosen y B. Stucker, *Additive Manufacturing Technologies*, Boston: Springer, 2010. DOI: 10.1595/205651315X688406.
- [7] N. Aliheidari, R. Tripuraneni, A. Ameli y S. Nadimpalli, “Fracture resistance measurement of fused deposition modeling 3D printed polymers”, *Polym. Test.*, vol. 60, pp. 94-101, Jul. 2017. DOI: 10.1016/J.POLYMERTESTING.2017.03.016.
- [8] S. Ahn, M. Montero, D. Odell, S. Roundy y P. K. Wright, “Anisotropic material properties of fused deposition modeling ABS”, *Rapid Prototyp. J.*, vol. 8, n.º 4, pp. 248-257, 2002. DOI: 10.1108/13552540210441166.
- [9] O. A. Mohamed, H. Syed, Masood y J. L. Bhowmik, “Optimization of fused deposition modeling process parameters: a review of current research and future prospects”, *Advances in Manufacturing*, vol. 3, pp. 42-53. DOI: 10.1007/s40436-014-0097-7.
- [10] K. Gnanasekaran *et al.*, “3D printing of CNT- and graphene-based conductive polymer nanocomposites by fused deposition modeling”, *Appl. Mater. Today*, vol. 9, pp. 21-28, 2017. DOI: 10.1016/J.APMT.2017.04.003.
- [11] S. Berretta, R. Davies, Y. T. Shyng, Y. Wang y O. Ghita, “Fused Deposition Modelling of high temperature polymers: Exploring CNT PEEK composites”, *Polym. Test.*, vol. 63, pp. 251-262, 2017. DOI: 10.1016/J.POLYMERTESTING.2017.08.024.
- [12] J. Torres, J. Coteló, J. Karl y A. P. Gordon, “Mechanical Property Optimization of FDM PLA in Shear with Multiple Objectives”, *JOM*, vol. 67, n.º 5, pp. 1183–1193, 2015. DOI: 10.1007/s11837-015-1367-y.
- [13] N. G. Tanikella, B. Wittbrodt y J. M. Pearce, “Tensile strength of commercial polymer materials for fused filament fabrication 3D printing”, *Addit. Manuf.*, vol. 15, pp. 40-47, 2017. DOI: 10.1016/j.addma.2017.03.005.
- [14] B. M. Tymrak, M. Kreiger y J. M. Pearce, “Mechanical properties of components fabricated with open-source 3-D printers under realistic environmental conditions”, *Mater. Des.*, vol. 58, pp. 242–246, 2014. DOI: 10.1016/j.matdes.2014.02.038.
- [15] Ultimaker, *Manual de instalación y uso Ultimaker 2+: hace que lo fácil sea incluso más fácil*, Países Bajos: Ultimaker, 2016.
- [16] ASTM International, *Standard Test Method for Tensile Properties of Plastics I*, West Conshohocken: ASTM, 2019. DOI: 10.1520/D0638-14.
- [17] B. Rankouhi, S. Javadpour, F. Delfanian y T. Letcher, “Failure Analysis and Mechanical Characterization of 3D Printed ABS With Respect to Layer Thickness and Orientation”, *J. Fail. Anal. Prev.*, vol. 16, pp. 467-481, 2016. DOI: 10.1007/s11668-016-0113-2.
- [18] A. Bellini y S. Güçeri, “Mechanical characterization of parts fabricated using fused deposition modeling”, *Rapid Prototyp. J.*, vol. 9, n.º 4, pp. 252-264, 2003. DOI: 10.1108/13552540310489631.

- [19] O. A. Mohamed, S. H. Masood y J. L. Bhowmik, “Optimization of fused deposition modeling process parameters for dimensional accuracy using I-optimality criterion”, *Measurement*, vol. 81, pp. 174–196, 2016. DOI: 10.1016/J.MEASUREMENT.2015.12.011.
- [20] C. Chung Wang, T. Lin y S. Hu, “Optimizing the rapid prototyping process by integrating the Taguchi method with the Gray relational analysis”, *Rapid Prototyp. J.*, vol. 13, n.º 5, pp. 304–315, 2007. DOI: 10.1108/13552540710824814.
- [21] A. K. Sood, R. K. K. Ohdar y S. S. S. Mahapatra, “Parametric appraisal of mechanical property of fused deposition modelling processed parts”, *Materials & Design*, vol. 31, n.º 1, pp. 287-295, 2010. DOI: 10.1016/j.matdes.2009.06.016.
- [22] F. Rayegani y G. C. Onwubolu, “Fused deposition modelling (FDM) process parameter prediction and optimization using group method for data handling (GMDH) and differential evolution (DE)”, *Int. J. Adv. Manuf. Technol.*, vol. 73, n.º 1-4, pp. 509–519, 2014. DOI: 10.1007/s00170-014-5835-2.
- [23] D. Popescu, A. Zapciu, C. Amza, F. Baciú y R. Marinescu, “Material Properties FDM process parameters influence over the mechanical properties of polymer specimens: A review”, *Polym. Test.*, vol. 69, pp. 157-166, 2018. DOI: 10.1016/j.polymertesting.2018.05.020.
- [24] A. M. Do Nald y E. J. Krame, “Plastic deformation mechanisms in poly(acrylonitrile-butadiene styrene) [ABS]”, *Journal of Materials Science*, vol. 17, pp. 1765-1772, 1982. DOI: <https://doi.org/10.1007/BF00540805>.
- [25] R. W. Truss y G. A. Chadwick, “Tensile deformation ABS polymers behaviour of ABS polymers”, *Journal of Materials Science*, vol. 11, pp. 111-117. DOI: 10.1007/BF00541081.
- [26] R. Braglia y M. SpA, “Craze formation in ABS polymers”, *Journal of Materials Science*, vol. 19, pp. 2643–2650 1984. DOI: 10.1007/BF00550821.



# Desarrollo de una aplicación para la predicción de ingredientes y recetas de cocina por medio de TensorFlow y máquinas de soporte vectorial\*

Yeny Muñoz-Castaño\*\*

Luis Castillo-Ossa\*\*\*

Omar Castrillón-Gomez\*\*\*\*

Felipe Buitrago-Carmona\*\*\*\*\*

Santiago Loaiza Giraldo\*\*\*\*\*

Recibido: 11/09/2019 • Aceptado: 13/07/2020

<https://doi.org/10.22395/rium.v19n37a10>

## Resumen

Este artículo es derivado de un proyecto de investigación en el cual se realizó el desarrollo de una aplicación para la predicción de ingredientes y recetas por medio de TensorFlow y máquinas de soporte vectorial. Se realizó un esquema de la arquitectura general, luego se desarrolló una red neuronal y después se efectuó la ejecución de la máquina de soporte vectorial. Finalmente, se realizó la integración en una aplicación que permite al usuario seleccionar imágenes de los ingredientes para su predicción y de la receta de cocina para desayunos de manera didáctica. Se concluyó que el sistema tiene un promedio de precisión del 75,8 % para las 17 categorías de ingredientes y de 71 % para el clasificador de recetas. Adicionalmente, se realizó un análisis de esta estabilidad y se observó que todos los resultados eran iguales en términos estadísticos.

**Palabras clave:** reconocimiento; imagen; *TensorFlow*; SVM; redes neuronales.

---

\* Artículo original derivado del proyecto de investigación “Prototipo computacional para la fusión y análisis de grandes volúmenes de datos en entornos IOT (internet de las cosas) a partir de técnicas de *machine learning* y arquitecturas seguras entre sensores, para caracterizar el comportamiento e interacción de los usuarios en un ecosistema de *connected home*” (código 36715), financiado por la Dirección de investigaciones de Manizales (DIMA) de la Universidad Nacional de Colombia (Sede Manizales). Periodo de ejecución: 2018 - 2019.

\*\* Ingeniera electrónica. Miembro del Grupo de Investigación Innovación y Desarrollo Tecnológico. Universidad Nacional de Colombia, Sede Manizales. Correo electrónico: yymunozc@unal.edu.co. Orcid: 0000-0002-5853-2646.

\*\*\* Ph. D. Miembro del Grupo Gitir. Facultad de ingenierías, Universidad de Caldas. Correo electrónico: luis.castillo@ucaldas.edu.co. Orcid: 0000-0002-2878-8229.

\*\*\*\* Ph. D. Miembro del Grupo de Investigación Innovación y Desarrollo Tecnológico. Universidad Nacional de Colombia, Sede Manizales. Correo electrónico: odcastrillong@unal.edu.co. Orcid: 0000-0003-3713-0696.

\*\*\*\*\* Ingeniero de sistemas y computación. Miembro del grupo de investigación Semillero de Inteligencia Artificial I3A - Grupo Gitir. Universidad de Caldas. Correo electrónico: felipe.1701416667@ucaldas.edu.co. Orcid: 0000-0002-1085-0979.

\*\*\*\*\* Ingeniero de sistemas y computación. Miembro del grupo de investigación Semillero de Inteligencia Artificial I3A - Grupo Gitir. Universidad de Caldas. Correo electrónico: santiago.1701412598@ucaldas.edu.co. Orcid: 0000-0001-6672-9804.

## Development of an Application for the Prediction of Kitchen Ingredients and Recipes through TensorFlow and Support-Vector Machines

### **Abstract**

This article is derived from a research project in which an application for the prediction of ingredients and recipes by TensorFlow and support-vector machines was developed. A scheme with general architecture was developed, then a neural network was implemented, and then, the support-vector machine was run. After that, they were integrated via an application that allows the user to select ingredients' images for their prediction and the prediction of kitchens recipe in a didactic manner. It was concluded that the system has an average precision value of 75.8% and 71% for 17 ingredients categories and recipes classifier. In addition, sensitivity testing was performed on the application resulting on statistically equivalent results.

**Keywords:** Recognition, image, TensorFlow, SVM, Neural Networks.

## Desenvolvimento de um aplicativo para a predição de ingredientes e receitas de cozinha por meio de TensorFlow e máquinas de suporte vetorial

### **Resumo**

Este artigo é derivado de um projeto de pesquisa no qual se realizou o desenvolvimento de um aplicativo para a predição de ingredientes e receitas por meio de TensorFlow e máquinas de suporte vetorial. Realizou-se um esquema da arquitetura geral, desenvolveu-se uma rede neuronal e depois efetuou-se a execução da máquina de suporte vetorial. Finalmente, realizou-se a integração num aplicativo que permite ao usuário selecionar imagens dos ingredientes para sua predição e da receita de cozinha para cafés da manhã de maneira didática. Concluiu-se que o sistema tem uma média de precisão de 75,8 % para as 17 categorias de ingredientes e de 71 % para o classificador de receitas. Adicionalmente, realizou-se uma análise dessa estabilidade e observou-se que todos os resultados eram iguais em termos estatísticos.

**Palavras-chave:** reconhecimento; imagem; TensorFlow; SVM; redes neurais.



## INTRODUCCIÓN

El reconocimiento de imágenes y objetos en nuestro entorno está en auge. Uno de los métodos más utilizados es el uso de redes neuronales, las cuales han logrado clasificar de 1,2 millones de imágenes con 1000 clases diferentes [1]; igualmente, se han utilizado para extracción de patrones característicos de imágenes [2]. Así mismo, se han implementado técnicas de *deep learning*, el cual permite que los modelos computacionales compuestos de múltiples capas de procesamiento aprendan representaciones de datos con múltiples niveles de abstracción [3]. En la misma línea, se han implementado modelos como *Bag of Words* a fin de realizar un reconocimiento con base en el contenido de las imágenes [4]; o filtros como el de Base Haar y clasificadores en cascada para la extracción de características sobre imágenes digitales, con los que se han obtenido porcentajes de detección de rostro y de ojos de 100 % y 92 %, respectivamente [5].

Las máquinas de soporte vectorial (*support-vector machines*, SVM) constituyen otra de las técnicas utilizadas a este respecto: aprenden a predecir una nueva clase de muestra a partir de ejemplos. Esta técnica está basada en la minimización de riesgo estructural (*structural risk minimization*, SRM) [6] la cual se ha utilizado para la clasificación automática de señales sísmicas [7] dentro de la identificación de fracturas naturales en pozos de yacimientos de hidrocarburos: su exactitud osciló entre 72,3 % y 82,2 % en cinco pozos evaluados del campo estudiado [8]. También se la ha usado para el reconocimiento de patrones mediante huellas dactilares de descargas parciales con filtrado de *wavelet*, con altas tasas de reconocimiento [9].

Sumado a lo anterior, se han implementado SVM junto con los métodos de Kernel para mejorar la inferencia transductiva de motores de búsqueda de texto [10]; y también se han empleado para el reconocimiento de voz, en aras de emular la toma de decisiones de los humanos (tarea de clasificación y análisis de audio realizada por los técnicos): los resultados del correspondiente estudio arrojaron un porcentaje de clasificación exitosa en comparación con una plataforma automática llamada *CheckMyRoutes* [11].

TensorFlow ha sido utilizado para la construcción de una API (*Application Programming Interface*) de Python denominada *InferPy* para el modelado probabilístico construido sobre Edward y TensorFlow [12]. Así mismo, se ha usado para acelerar la inversión sísmica geoestadística, la cual se puede incluso mejorar el modelo utilizando *graphics processing units* (GPU) [13]. También se han realizado simulaciones con *CitySim*, un simulador que combina la energía de construcción y TensorFlow, el cual permite investigar algoritmos para el ahorro de energía y respuesta a la demanda [14]. Se han usado, además, *deep convolutional generative adversarial networks* (DCGAN), es decir, redes de confrontación generativas de convolución profunda: se ha realizado una exploración de esta última a través de TensorFlow y se ha encontrado

que el modelo DCGAN mejora de forma significativa frente al modelo de generación de caras virtuales [15].

El uso de las herramientas citadas ha empezado a difundirse en empresas para optimizar sus procesos dado que son versátiles y mejoran el nivel de desarrollo. Así mismo, miles de startups e investigadores universitarios apuestan por TensorFlow como base para desarrollar sus propios modelos de inteligencia artificial [16]. Por lo tanto, es importante desde la academia incentivar el uso de estas herramientas con miras a desarrollar aplicaciones que puedan ser útiles a la sociedad y, en esta línea, que sean usadas en entornos de gran importancia en la cotidianidad, como la cocina. A este respecto, si bien en la actualidad se ha automatizado el hogar con sistemas de iluminación inteligente, calefacción y ventilación enfocados en la reducción del consumo de energía [17], es necesario hacer un mayor aporte tecnológico con el uso de estas técnicas nuevas para que la automatización mejore y, con ello, sea mucho más eficaz.

Con lo anterior, el objetivo de esta investigación fue desarrollar una aplicación para la predicción de ingredientes y recetas por medio de TensorFlow y máquinas de soporte vectorial. Para esto se realizó un esquema de la arquitectura general, luego se desarrolló una red neuronal y después se efectuó la ejecución de la máquina de soporte vectorial. Finalmente se realizó la integración en una aplicación que permitiría al usuario seleccionar imágenes de los ingredientes para la predicción de los ingredientes y la receta de manera didáctica.

Este desarrollo, realizado con TensorFlow y máquinas de soporte vectorial para uso en la cocina, permitiría reconocer los ingredientes y realizar una clasificación de recetas que se tienen en una cocina. Aunque en la fase de la investigación que describe aquí se realizaron las pruebas con imágenes tomadas de Google (210 por cada categoría de ingredientes), se espera que en la fase posterior, de entrenamiento, el sistema se implemente en un microcontrolador de bajo consumo como Raspberry Pi 3, dado que ya se ha usado con anterioridad para supervisar y controlar procesos a nivel industrial [18]; igualmente, se aspira a incrementar el número de ingredientes reconocidos (100) y el reconocimiento en recetas (20), y establecer una asociación con un servidor IoT, dado que permite la gestión de entornos domésticos [19].

## 1. MATERIALES Y MÉTODOS

Esta sección presenta la información y las descripciones cortas de los procedimientos realizados para desarrollar de una aplicación para la predicción de ingredientes y recetas por medio de TensorFlow y máquinas de soporte vectorial.

## 1.1 Paso 1: descripción de la arquitectura general

Se procedió a desarrollar un sistema con el modelo cliente - servidor para obtener la predicción de 17 categorías de ingredientes y 6 categorías de recetas (tabla 1). La parte denominada ‘cliente’, esto es, aquello conocido como *front-end*, que brindaría al usuario final una relación agradable e intuitiva con el sistema, se desarrollaría en código HTML con JavaScript para vincularla con la parte de servidor (*back-end*). En la figura 1 se observa la arquitectura mínima planteada.

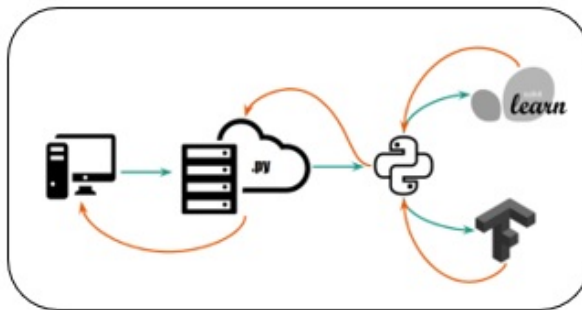


Figura 1. Descripción de la arquitectura general

Fuente: elaboración propia.

Tabla 1. Categorías generales

Ítem	Descripción
Categoría A: ingredientes	Huevos, arepas, mantequilla, chocolate, pan, cereales, café, leche, tocino, changua, tamal, papas, calentado, yuca frita, jugo de naranja, yogur, pollo.
Categoría B: recetas de desayunos	Paisa, cafetero, rolo, fitness, tolimense, americano.

Fuente: elaboración propia.

## 1.2 Paso 2: desarrollo de la red neuronal

El modelo de las redes neuronales surgió de simulaciones abstractas de sistemas nerviosos biológicos, constituidos a su vez por neuronas o nodos interconectados entre sí [20]. En un modelo de neurona artificial se tienen cuatro componentes básicos: primero, un conjunto de conexiones o sinapsis que determinan el comportamiento de la neurona; segundo, un sumador que se encarga de unir todas las entradas multiplicadas por las respectivas sinapsis; tercero, una función de activación no lineal para limitar la amplitud de la salida de la neurona; y cuarto, un umbral exterior que determina cuándo se activa la neurona. En la figura 2 se observa el modelo [21].

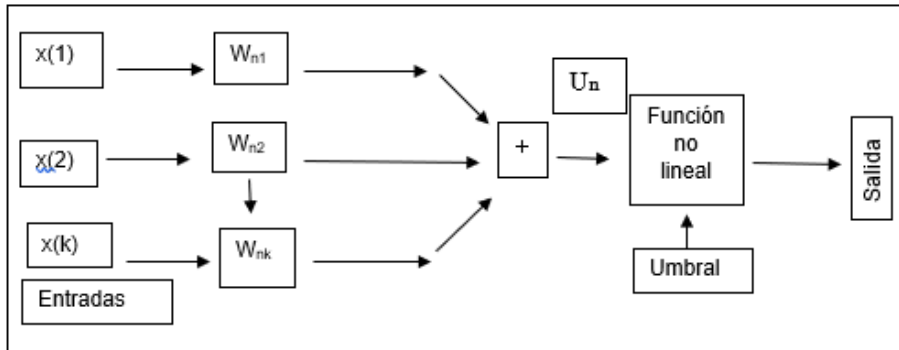


Figura 2. Modelo de red neuronal

Fuente: adaptada de [21].

En la ecuación (1) se observa la operación matemática a realizar [21, p.15]

$$"U_n = \sum_{j=1}^k W_{nj} \cdot x(j)" \tag{1}$$

Y en la ecuación (2) se observa la salida [21, p.15]

$$"salida = \rho(U_n - umbral)" \tag{2}$$

Donde  $\rho$  es una función de activación no lineal; adicionalmente, se asocia el umbral a la salida  $U_n$  mediante una entrada que vale -1 y un peso asociado. Por lo tanto, en las ecuaciones (3) y (4), se observa el cambio del umbral [21, p.15]

$$"umbral = -W_{no} \rightarrow \{U_n = \sum_{j=1}^k W_{nj} \cdot x(j) \rightarrow salida = \rho(U_n)" \tag{3}$$

O

$$"umbral = -W_{no} \rightarrow \{x(0) = 1 \rightarrow salida = \rho(U_n)" \tag{4}$$

Este modelo planteado es el más común. Empero, hay otros que no realizan un promedio de las entradas de forma directa, sino que realizan una transformación a dichas entradas —cuadrática, polinómica, esférica—; adicionalmente, también se podría tener consideración de las entradas anteriores [21].

El desarrollo de una red neural es el núcleo principal para el funcionamiento del sistema y se busca que brinde inteligencia específica al decidir sobre una imagen que está llegando como entrada de datos. Para ella se tuvieron en consideración los siguientes elementos:

- Modelo:** la red neuronal diseñada es de tipo convolucional, muy similar a un perceptrón multicapa. Este se compone de una capa de entrada, una de salida y una o más que están ocultas; la figura 3 muestra un perceptrón típico [20]. En el desarrollo de esta investigación se definió un modelo de dos capas de convolución, las cuales podrán variar internamente o aumentar en cantidad por temas de rendimiento y porcentaje de precisión en versiones futuras del sistema. Ambas capas de convolución utilizan dos dimensiones con el objetivo, ya que la teoría las propone como ideales para el trabajo con imágenes; se utilizan cinco unidades como dimensión de la ventana de convolución (alto y ancho) y un paso por esta ventana. La primera capa cuenta con una salida de 16 filtros; y la segunda, con una de 36. A ellas se les aplica un proceso de discretización basado en muestras conocido como *'Max-pooling'* encaminado a reducir la muestra de una representación de entrada (imagen, matriz de salida de capa oculta, etc.) de tal suerte que se reduzca su dimensionalidad y se posibiliten suposiciones sobre las características contenidas en las subregiones agrupadas.

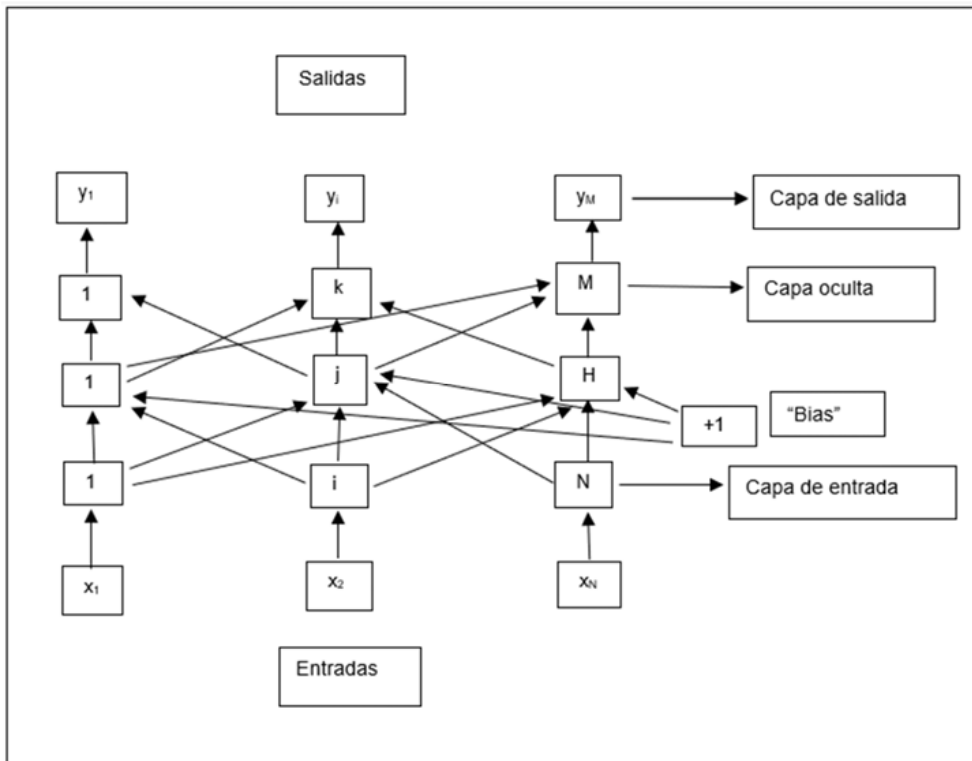


Figura 3. Modelo de perceptrón típico

Fuente: adaptada de [20].

- *Función de costo*: esta trata de determinar el error entre los valores estimado y real con el fin de optimizar los parámetros de la red neuronal. Para esta investigación se utilizó el estimador de raíz cuadrada media RMSE (ecuación 5): esta es una medida de precisión calculada como la raíz cuadrada media de los residuos, entendidos estos últimos como la diferencia entre los valores previsto (correcto) y real obtenido [22, p. 1].

$$"RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}" \quad (5)$$

- *Factor de aprendizaje*: el aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante la etapa de aprendizaje se reducen a la destrucción (el peso de la conexión toma el valor 0), modificación y creación (el peso de la conexión toma un valor distinto de 0) de conexiones entre las neuronas.

Podemos considerar que el proceso de aprendizaje ha terminado (ecuación 6) cuando los valores de los pesos permanecen estables. [23, p.1]

$$"\frac{dW_j}{dt} = 0" \quad (6)$$

En esta investigación se realizó un aprendizaje supervisado: se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida generada por el sistema y, en caso de que no coincida con la esperada, se procederá a modificar los pesos de las conexiones [23].

### 13. Paso 3: desarrollo de la máquina de soporte vectorial

La teoría de la SVM está basada en la idea de minimización de riesgo estructural (SRM). En muchas aplicaciones, las SVM han mostrado tener gran desempeño, incluso más que máquinas de aprendizaje tradicional como las redes neuronales; y han sido introducidas como herramientas poderosas para resolver problemas de clasificación [6]. Dentro de la tarea de clasificación, las SVM pertenecen a la categoría de los clasificadores lineales puesto que inducen separadores lineales o hiperplanos, ya sea en el espacio original de los ejemplos de entrada si estos son separables o cuasi separables (ruido), o en un espacio transformado (espacio de características) si los ejemplos no son separables linealmente [24].

Se define un conjunto (ecuación 7) para una clasificación binaria de ejemplos separables linealmente [24, p. 2]:

$$"S = \{(x_1, y_1); \dots; (x_n, y_n)\}" \tag{7}$$

Donde  $x_i \in R^d$  e  $y_i \in \{+1, -1\}$  [24, p. 2] se puede definir un hiperplano de separación (figura 4A) como una función lineal que es capaz de separar dicho conjunto (ecuación 8) sin error [24, p. 2].

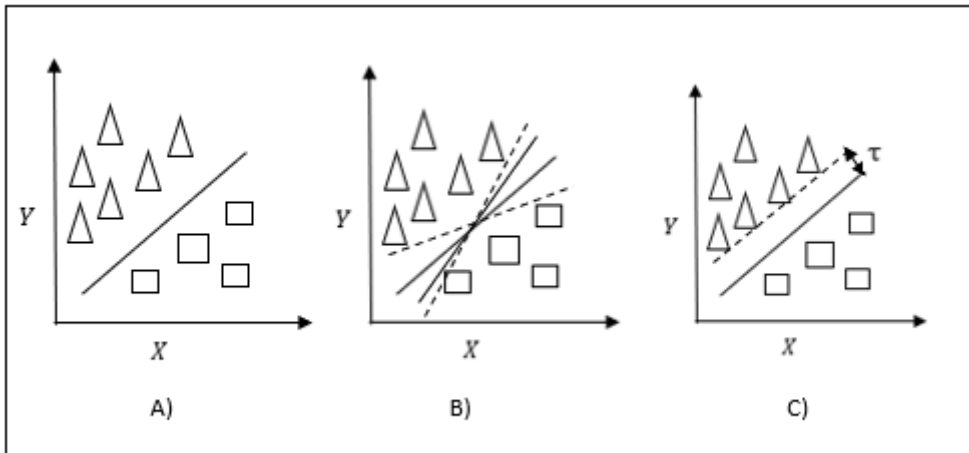


Figura 4. a) ejemplo de hiperplano de separación de dos clases; b) otros ejemplos de hiperplanos de separación; c) margen de un hiperplano de separación

Fuente: adaptada de [22].

$$"D(x) = (w_1, x_1) + \dots + (w_d, x_d) + b = \langle w, x \rangle + b" \tag{8}$$

Donde  $w$  y  $b$  son coeficientes reales [24], el hiperplano de separación cumplirá las restricciones presentadas por las ecuaciones (9), (10), (11) y (12) para todo  $x_i$  del conjunto de ejemplos [24, p. 3].

$$\langle w, x_i \rangle + b \geq 0 \text{ si } y_i = +1" \tag{9}$$

$$\langle w, x_i \rangle + b \leq 0 \text{ si } y_i = -1, "i = 1, \dots, n" \tag{10}$$

O también,

$$"y_i (\langle w, x_i \rangle + b) \geq 0, "i = 1, \dots, n" \tag{11}$$

O de forma más compacta,

$$"y_i D(x_i) \geq 0, i = 1, \dots, n" \tag{12}$$

Ahora el hiperplano que permite separar los ejemplos no es único, como se observa en la figura 4B; por lo tanto, se define el concepto de *margen* en un hiperplano de separación denotado por  $\tau$  como la mínima distancia entre dicho hiperplano y el ejemplo más cercano de cualquiera de las dos clases, como se observa en la figura 4C, si su margen de tamaño máximo se denominara óptimo.

Una propiedad del hiperplano de separación óptimo es que resulta equidistante respecto del ejemplo más cercano de cada clase: Por geometría (ecuación 13) se sabe que la distancia entre un hiperplano de separación  $D(x)$  y un ejemplo  $x'$  está dada por [24]:

$$\frac{|D(x')|}{w} \quad (13) \quad \hat{\sigma}$$

Al hacer uso de las ecuaciones (12) y (13), todos los ejemplos de entrenamiento (ecuación 14) cumplirán [24, p. 4].

$$\frac{y_i D(x_i)}{w} \geq \tau, \quad "i = 1, \dots, n". \quad (14)$$

De la ecuación (13) se deduce que encontrar el hiperplano óptimo es igual a encontrar el valor de  $w$  que maximiza el margen si se expresa la ecuación (14) de la siguiente manera [24, p. 4]:

$$y_i D(x_i) \geq w\tau, \quad "i = 1, \dots, n". \quad (15)$$

Si se fija el producto  $\tau$  y  $w = 1$ , se llega a la conclusión (ecuación 16) de que aumentar el margen es equivalente a disminuir la norma [24, p. 4].

$$\tau = \frac{1}{w} \quad (16)$$

En conclusión, un hiperplano óptimo será aquel que posea un margen máximo y un valor mínimo de  $\|w\|$ .

Con todo lo dicho, y en términos más concretos, la SVM en el caso que nos ocupa es la agrupación de algoritmos que proporcionan los medios necesarios para elaborar procedimientos de aprendizaje supervisado; y su función es determinar, a partir de los ingredientes predeterminados y seleccionados por el usuario, a qué tipo de desayuno entre los establecidos podría pertenecer. Los elementos que componen una SVM son los siguientes:

- *Kernel*: Las funciones kernel permiten convertir lo que sería un problema de clasificación no lineal en uno de clasificación lineal con un espacio de dimensión mayor; normalmente son usadas en las SVM [25]. Para la instanciación del clasificador mencionado se definió el tipo de función bajo la cual se regiría. En este



caso puntual se utilizó el tipo *radial basis function* o función de base radial, definida a partir de pruebas y comparaciones con las demás opciones como la precisa para el *software* planteado.[26]

- *Grupo de datos*: para el caso del clasificador, la fuente de datos se genera en un archivo de valores separados por comas (*comma-separated values*, CSV); contiene, en una estructura definida por cada línea, un registro posible donde cada celda será un atributo en valor binario (1 pertenece, 0 no pertenece) y la última celda por cada línea representará la clase en un número entero asociado a un diccionario de datos.

#### 1.4 Paso 4: pruebas de entrenamiento

Para el desarrollo de la red neuronal se realizó un módulo de pruebas: como entrada se proporciona una imagen; y con procedimientos de edición, almacenamiento, acceso y traspaso, esta llega como se desea a la red neuronal, la cual retornará entre ciertas categorías su decisión al respecto. Para el desarrollo de la máquina de soporte vectorial se realizó un módulo de pruebas cuya entrada correspondió a un arreglo con datos binarios por cada categoría (como se explicó): como resultado, el clasificador retornará un valor léxico de la clase asociada que representa una clase de desayuno definida, a la cual la receta podría pertenecer. Después de esto se realizan diversos procesos de entrenamiento en general, los cuales inician con muestras pequeñas (5 categorías, 10 imágenes por categoría), y luego se inician pruebas con los parámetros utilizados en la prueba de despegue con todo el grupo de datos (17 imágenes, 210 imágenes por categoría).

## 2. RESULTADOS

En primer lugar, se describirán aquí los resultados obtenidos en cuanto a los requerimientos: se observará un ejemplo de las categorías mencionadas, así como la descripción computacional de la arquitectura. En segundo lugar, se mostrarán resultados parciales de la red neuronal y las máquinas de soporte vectorial; y por último se presentará la interfaz para el usuario, así como las tablas de las predicciones obtenidas para ingredientes y recetas.

### 2.1 Pasos 1 y 2: descripción de la arquitectura general y desarrollo de la red neuronal

La programación de la arquitectura planteada se puede observar en el diagrama de la figura 5: allí se muestra que al interactuar con la interfaz, el usuario debe seguir un proceso en que ha de elegirse si se realiza el reconocimiento del ingrediente por medio de la red neuronal, o si se desea realizar la predicción de la receta a través de la SVM. Para el entrenamiento de la red neuronal, realizada en TensorFlow sobre Python, se

realiza un entrenamiento basado en épocas, cada una de las cuales tiene dos etapas: una de entrenamiento y una de pruebas. Como se estableció en secciones previas, para la red neuronal del sistema hay 210 imágenes por 17 categorías. A modo de ejemplo, en la tabla 2 se muestra una época del entrenamiento; en él, la cantidad es configurable. Finalizado esto, se obtiene un valor porcentual de efectividad con el set de datos del 87,06 %. La tabla 3 presenta los resultados del entrenamiento de la red neuronal con variaciones en los porcentajes de entrenamiento y validación.

Tabla 2. Programación, época d6e entrenamiento

```
python .train.py
#Se enmarca el proceso de TensorFlow sobre PYTHON
Using TensorFlow backend.
Epoch 1/1
2890/2890 [=====>] 2890/2890 [=====>] - 6s 32ms/step -
loss: 3.4935 - acc: 0.0412
#ETAPA DE PRUEBA, SE EVALÚAN 680 IMÁGENES POR ÉPOCA (40 POR CADA CATEGORÍA)
680/680 [=====>] 680/680 [=====>] - 1s 9ms/step
#FINALIZADO EL PROCESO DE ENTRENAMIENTO, Y EN RELACIÓN CON LAS ETAPAS DE PRUEBA
#SE OBTIENE UN VALOR PORCENTUAL DE EFECTIVIDAD DE LA RED NEURONAL CON EL SET DE DATOS
acc: 87.06%

-----
Layer (type) Output Shape Param #
-----
input_1 (InputLayer) (None, 16384) 0

-----
reshape_1 (Reshape) (None, 128, 128, 1) 0

-----
layer_conv1 (Conv2D) (None, 128, 128, 16) 416

-----
max_pooling2d_1 (MaxPooling2 (None, 64, 64, 16) 0

-----
layer_conv2 (Conv2D) (None, 64, 64, 36) 14436

-----
max_pooling2d_2 (MaxPooling2 (None, 32, 32, 36) 0

-----
flatten_1 (Flatten) (None, 36864) 0

-----
dense_1 (Dense) (None, 128) 4718720

-----
dense_2 (Dense) (None, 17) 2193
-----
Total params: 4,735,765
Trainable params: 4,735,765
Non-trainable params: 0
```

Fuente: elaboración propia.

Tabla 3. Resultados del entrenamiento de la red neuronal para predicción de ingredientes con variaciones en los porcentajes de entrenamiento y validación

Porcentaje de entrenamiento	Porcentaje de validación	Resultados
20 %	80 %	40,7 %
40 %	60 %	62,1 %
60 %	40 %	73,3 %
80 %	20 %	75,2 %

Fuente: elaboración propia.

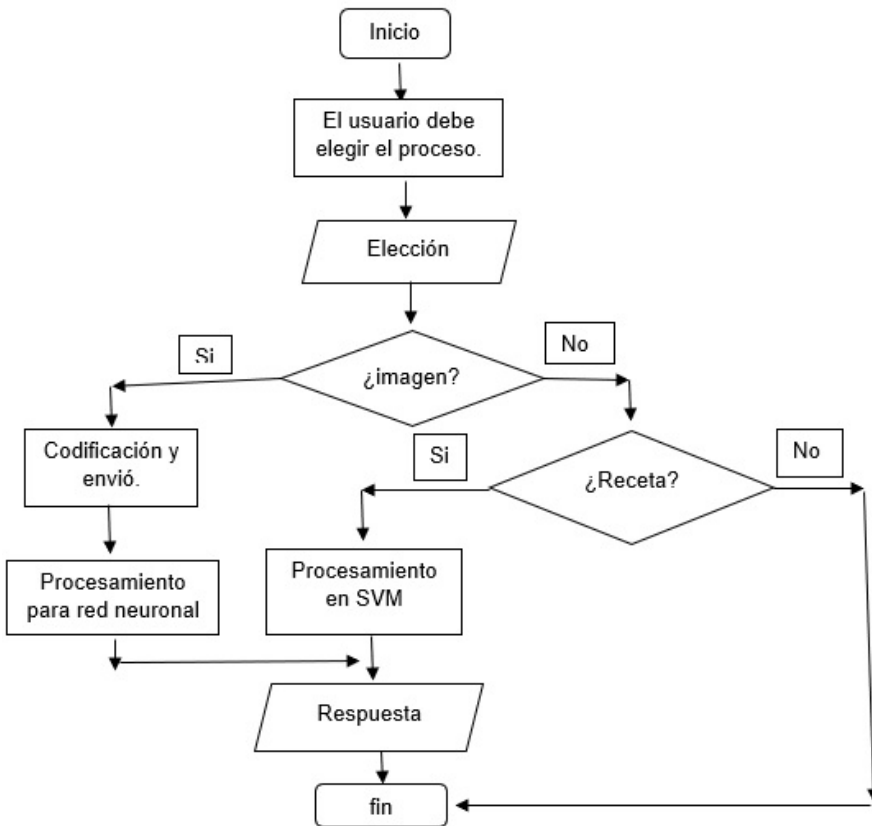


Figura 5. Programación de arquitectura general

Fuente: elaboración propia.

## 2.2 Paso 3: desarrollo de la SVM

La tabla 4 muestra la configuración de la SVM, en el marco de la cual se define un arreglo con las nomenclaturas que corresponden a los desayunos. Los índices de posición corresponden a aquellos que se disponen en el *dataset* que alimenta la SVM y retorna

la predicción realizada. La tabla 5, a su turno, muestra los resultados del entrenamiento de la SVM con variaciones en los porcentajes de entrenamiento y validación.

Tabla 4. Programación configuración máquina de soporte vectorial

```

self.desayunos = [
    "PAISA",
    "CAFETERO",
    "ROLO",
    "AMERICANO",
    "FITNESS",
    "TOLIMENSE",
]
ruta_data = ". /models/data_svc.csv"
data_set = pd.read_csv(ruta_data)
X = data_set.drop("CLASE", axis = 1)
Y = data_set["CLASE"]
self.svc_classifier = SVC(Kernel = "rbf")
self.svc_classifier.fit(x_train, y_train)

def predict(self, recipe):
    x = self.svc_classifier.predict(recipe)
    return self.desayunos[int(x)]

```

Fuente: elaboración propia.

Tabla 5. Resultados de entrenamiento con la máquina de soporte vectorial para la predicción de recetas, con variaciones en los porcentajes de entrenamiento y validación

Porcentaje de entrenamiento	Porcentaje de validación	Resultados
20%	80%	100%
40%	60%	100%
60%	40%	100%
80%	20%	100%

Fuente: elaboración propia.

### 2.3 Paso 4: pruebas de entrenamiento

En cuanto al entrenamiento se realizaron diversos procesos que iniciaron con muestras pequeñas (5 categorías, 10 imágenes por categoría), para las cuales se variaron los valores de épocas y tamaño de lote; luego, se iniciaron pruebas con los parámetros utilizados en la prueba de despegue con todo el grupo de datos (17 imágenes, 210 imágenes por categoría); y finalmente se realizaron tres pruebas más con todo el grupo de datos, cambiando los parámetros principales (épocas y tamaño del lote).

Se obtuvieron promedios de precisión del 75,8 % para el reconocimiento de ingredientes con la red neuronal (tabla 6) y de 71,3 % respecto al de recetas para el entrenamiento realizado con la SVM (tabla 7). En la figura 6 se observa la interfaz terminada: se realizó la predicción de 2 ingredientes (huevos y chocolate) y a partir de ellos se predijo que los ingredientes seleccionados conformarían un desayuno cafetero. Con el fin de obtener la estabilidad del sistema, el proceso fue adelantado 10 veces en grupos de 5, ante lo cual todos los resultados fueron iguales en términos estadísticos (tablas 8 y 9).

Tabla 6. Resultados de cada ingrediente en el entrenamiento con la red neuronal

Ítem	Descripción	Precisión
1	Huevos	87 %
2	Arepas	91 %
3	Mantequilla	79 %
4	Chocolate	74 %
5	Pan	77 %
6	Cereales	79 %
7	Café	73 %
8	Leche	70 %
9	Tocino	73 %
10	Changua	74 %
11	Tamal	70 %
12	Papas	72 %
13	Calentado	71 %
14	Yuca frita	77 %
15	Jugo de naranja	72 %
16	Yogur	68 %
17	Pollo	83 %
Promedio		75,8 %

Fuente: elaboración propia.

Tabla 7. Resultados de cada receta en el entrenamiento con la SVM

Ítem	Descripción	Precisión
1	Paísa	69 %
2	Cafetero	82 %
3	Rolo	72 %
4	Fitness	59 %
5	Tolimense	71 %
6	Americano	75 %
Promedio		71,3 %

Fuente: elaboración propia.

Tabla 8. Estabilidad del sistema - ingredientes

Técnica		Numero de intentos promedio predicción de ingredientes								
Iteración	1	2	3	4	5	6	7	8	9	10
Red neuronal (Café)	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	Ok	OK	OK	OK	OK	OK	OK	OK	OK	OK
	Ok	NG	OK	OK	OK	OK	OK	OK	OK	OK
	NG	NG	NG	NG	NG	NG	NG	NG	NG	NG
	OK	OK	OK	NG	OK	OK	OK	OK	OK	OK
	NG	OK	NG	NG	NG	NG	OK	NG	NG	NG
	NG	NG	NG	NG	NG	NG	NG	NG	NG	NG
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	NG	NG	NG	NG	NG	NG	NG	NG	NG	NG
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
Porcentajes	73,3 %	73,3 %	73,3 %	66,6 %	73,3 %	73,3 %	80 %	73,3 %	73,3 %	73,3 %

Fuente: elaboración propia.

Tabla 9. Estabilidad del sistema - recetas

Técnica		Numero de intentos promedio predicción de Recetas								
Iteración	1	2	3	4	5	6	7	8	9	10
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	NG	NG	NG	NG	NG	NG	NG	NG	NG	NG
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
	NG	NG	NG	NG	NG	NG	NG	NG	NG	NG
	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
Porcentajes	83 %	83 %	83 %	83 %	83 %	83 %	83 %	83 %	83 %	83 %

Fuente: elaboración propia.



Figura 6. Aplicación para la predicción de ingredientes y recetas

Fuente: elaboración propia.

### 3. DISCUSIÓN

El reconocimiento de imágenes se puede aplicar para predecir tareas domésticas. En esta investigación se usó para predecir los ingredientes y las recetas; sin embargo, se encontró que esta metodología se ha utilizado en otras áreas: la asistencia a personas con alguna discapacidad visual, mediante el reconocimiento de la moneda local [27]; las ventajas y aplicaciones de la domótica para ayudar personas con discapacidad [28]; y el reconocimiento de las actividades por medio de una cámara web que realiza una persona en la cocina [29]. A medida que avanza la tecnología se busca que estos procesos se realicen con procesadores de bajo costo como Raspberry [30], usado con anterioridad para supervisar y controlar procesos en el ámbito industrial [18], o el Node MCU ESP82866, empleado para predecir la escena en la cocina por medio de sensores IoT y que se asocia a un servidor IoT, dado que este hace posible la gestión de entornos domésticos [19] [31].

A la par con lo dicho, se han desarrollado investigaciones sobre robots de servicio inteligentes que operen en entornos humanos estándar y automaticen tareas comunes [32]. Cabe señalar que la navegación autónoma de robots es uno de los principales problemas que plantea su desarrollo debido a su complejidad y dinamismo [33]; por ello, se espera integrar estos sistemas de reconocimiento de alimentos al robot para que este tenga un aprendizaje autónomo y pueda desempeñar mejor sus funciones.

Los resultados obtenidos con esta investigación son similares a los obtenidos por Leal, Ochoa, y García [8], quienes usaron la técnica de SVM para identificar fracturas naturales en pozos de yacimientos de hidrocarburos con una exactitud que osciló entre 72,3 % y 82,2 % [8]; ello demuestra la gran flexibilidad que reviste el modelo.

La técnica de SVM también se ha empleado en sistemas eléctricos, enfocándose en temas como predicción del mercado eléctrico, pérdidas no técnicas de electricidad (hurto), energías alternativas y transformadores, entre otros [34]. Además, también se ha hecho uso de SVM en iniciativas de tipo social, como la identificación del grado de riesgo psicosocial en docentes de colegios públicos en Colombia [35].

En cuanto al uso de TensorFlow, se observó que este admite una variedad de aplicaciones con un enfoque en capacitación e inferencia en redes neuronales profundas [36]; y que se están desarrollando herramientas paralelas como TensorD, la cual proporciona métodos de descomposición del tensor, así como operaciones básicas del tensor para facilitar la práctica de los métodos tensoriales en visión artificial, aprendizaje profundo y otros [37]. Así entonces, se espera usar la herramienta compuesta en esta investigación para aumentar el número de ingredientes reconocidos (100) y recetas conocidas (200), e integrarla a futuro con otro sistema para el surtimiento automático de alimentos o con diversas aplicaciones en la industria alimentaria, habida cuenta de que la evolución tecnológica permite disponer de soluciones altamente flexibles, de fácil acceso y con valores de inversión más bajo [38].

#### 4. CONCLUSIONES

Los datos obtenidos permiten inferir que el sistema concebido funciona y puede implementarse en entornos reales. Se determinó igualmente que el uso de SVM junto con TensorFlow representa una gran herramienta de clasificaciones. Se aspira en este sentido a extender este modelo en una segunda fase con el uso de microprocesadores de bajo consumo como Raspberry Pi y la cámara Pi, en aras de realizar este reconocimiento en un entorno real y aumentar tanto la cantidad de ingredientes reconocidos (100) como el reconocimiento en recetas (20), a fin de que en el futuro se pueda implementar esta herramienta con un sistema para el sistema de surtimiento automático de alimentos. Solo resta apuntar que, en cuanto al análisis realizado para verificar la estabilidad, se observó que los resultados fueron iguales en términos estadísticos.

#### AGRADECIMIENTOS

Se agradece por su colaboración a las universidades Nacional de Colombia (sede Manizales) y de Caldas, a Mabe y a todos los integrantes del proyecto “Prototipo computacional para la fusión y análisis de grandes volúmenes de datos en entornos IoT (internet de las cosas) a partir de técnicas de *machine learning* y arquitecturas seguras entre sensores para caracterizar el comportamiento e interacción de los usuarios en un ecosistema de *connected home*” (código 36715).



## REFERENCIAS

- [1] A. Krizhevsky *et al.*, “ImageNet Classification with Deep Convolutional Neural Networks”, *NIPS*, vol. 1, pp. 1097-1105, 2012. DOI: 10.1201/9781420010749
- [2] G. P. García, *Reconocimiento de imágenes utilizando redes neuronales artificiales*, Proyecto fin de Máster, Universidad Complutense de Madrid, Madrid, 2013.
- [3] Y. LeCun *et al.*, “Deep learning”, *Nature Methods*, vol. 521, n.º1, pp. 436-444, 2015. DOI: 10.1038/nature14539
- [4] F. S. Pardo, “Aplicación del modelo Bag-of-Words al reconocimiento de imágenes”, proyecto de fin de carrera, Universidad Carlos III de Madrid, Madrid, 2009.
- [5] M. L. Guevara *et al.*, “Faces Detection in Digital Images Using Cascade Classifiers”, *Scientia et Technica*, vol. 38, n.º 38, pp. 1-6, 2008.
- [6] G. A. Betancourt “Las máquinas de soporte vectorial (SVMs)”, *Scientia Et Technica*, vol. 27, pp. 67-72, 2005.
- [7] J. E. Hurtado *et al.*, “Clasificación de Señales Sísmicas por Medio de Onditas y Máquinas de Soporte Vectorial”, en *Primer simposio colombiano de sismología*, Manizales, Colombia, 2002.
- [8] J. A. Leal *et al.*, “Identification of natural fractures using resistive image logs, fractal dimension and support vector machines”, *Ingeniería e Investigación*, vol. 36, n.º 3, pp. 125-132, 2016. DOI: 10.15446/ing.investig.v36n3.56198.
- [9] I. C. Guzmán *et al.*, “Wavelet denoising of partial discharge signals and their pattern classification using artificial neural networks and support vector machines”, *DYNA*, vol. 84, n.º 203, pp. 240-248, 2017. DOI: 10.15446/dyna.v84n203.63745.
- [10] J. E. Espinosa *et al.*, “Kernel Methods for improving Text Search Engines Transductive Inference by Using Support Vector Machines”, *TECCIENCIA*, vol. 12, n.º 52, pp. 51-60, 2017. DOI: 10.18180/tecciencia.2017.22.6.
- [11] J. R. Wilches *et al.*, “A VoIP classifier for carrier grade based on Support Vector Machines”, *DYNA*, vol. 84, n.º 202, pp. 75-83, 2017. DOI: 10.15446/dyna.v84n202.60975.
- [12] R. Cabañas *et al.*, “InferPy: Probabilistic modeling with TensorFlow made easy”, *Knowledge-Based Systems*, vol. 168, pp. 25-27, 2019. DOI: 10.1016/j.knosys.2018.12.030.
- [13] M. Liu y D. Grana, “Accelerating geostatistical seismic inversion using TensorFlow: A heterogeneous distributed deep learning framework”, *Computers and Geosciences*, vol. 124, pp. 37-45, 2019. DOI: 10.1016/j.cageo.2018.12.007.
- [14] J. R. Vázquez *et al.*, “Fusing TensorFlow with building energy simulation for intelligent energy management in smart cities”, *Sustainable Cities and Society*, vol. 45, pp. 243-257, 2019. DOI: 10.1016/j.scs.2018.11.021.

- [15] S. Liu *et al.*, “The research of virtual face based of deep convolutional generative adversarial networks using TensorFlow”, *Physica A: Statistical Mechanics and its applications*, vol. 521, pp. 667-580, 2019.
- [16] X. Grandio (2017, 14 de julio) “Blog: Que es TensorFlow: Aplicaciones del sistema de inteligencia artificial de Google”. [internet]. Disponible en <https://marketing4ecommerce.net>.
- [17] P. Parsch y A. Masrur, “On Reliable Communication in Transmit-only Networks for Home Automation”, *Journal of Network and Computer Applications*, vol. 101, pp. 96-110, 2017. DOI: 10.1016/j.jnca.2017.10.023.
- [18] S. A. Castro *et al.*, “Supervisión y Control Industrial a través de Teléfonos Inteligentes usando un computador de placa única Raspberry Pi”, *Inf. Tecnol*, vol. 27, n.º 2, 2016. DOI: 10.4067/S0718-07642016000200015.
- [19] J. A. Asensio *et al.*, “Emulating home automation installations through component-based web technology”, *Future Generation Computer Systems*, vol. 93, pp. 777-791, 2017. DOI: 10.1016/j.future.2017.09.062.
- [20] J. M. Marín (2014). “Introducción a las redes neuronales aplicadas”. [internet]. Disponible en <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/DM/tema3dm.pdf>.
- [21] A. Serrano, E. Soria y J. D. Martin (2009). “Redes Neuronales Artificiales”. [internet]. Disponible en [http://ocw.uv.es/ingenieria-y-arquitectura/1-2/libro\\_ocw\\_libro\\_de\\_redes.pdf](http://ocw.uv.es/ingenieria-y-arquitectura/1-2/libro_ocw_libro_de_redes.pdf).
- [22] D. Calvo (2018). “Función de coste – Redes Neuronales”. [internet]. Disponible en <http://www.diegocalvo.es/funcion-de-coste-redes-neuronales/>.
- [23] J. F. Quesada (1997). “Características de las redes neuronales”. [internet]. Disponible en <https://thales.cica.es/rd/Recursos/rd98/TecInfo/07/capitulo3.html>.
- [24] E. Carmona (2014). “Tutorial sobre Máquinas de Vectores Soporte (SVM)” [internet]. Disponible en [http://www.ia.uned.es/~ejcarmona/publicaciones/\[2013-Carmona\]%20SVM.pdf](http://www.ia.uned.es/~ejcarmona/publicaciones/[2013-Carmona]%20SVM.pdf).
- [25] J. A. Reséndiz (2006). “Las máquinas de vectores de soporte para identificación en línea”. [internet]. Disponible en <https://www.ctrl.cinvestav.mx/~yuw/pdf/MaTesJAR.pdf>.
- [26] J. Echeverri *et al.*, “Mejoramiento de imágenes usando funciones de base radial”, *Revista ingenierías universidad de Medellín*, vol. 8, n.º 15, sup., 1, pp. 21-28, 2009.
- [27] B. C. Gayón, *Desarrollo de una Aplicación para Reconocimiento de Billetes por medio de Procesamiento de Imágenes con Diversidad Visual Basada en Tecnología Android*, proyecto de fin de carrera, Universidad Libre, Bogotá, 2017.
- [28] F. J. Rodríguez, *Automatización Mediante Equipos EIB de una Cocina Adaptada Dentro del Entorno del Robot Asistencial ASIBOT*, proyecto de fin de carrera, Universidad Carlos III, Madrid, 2010.

- [29] F. J. García, *Reconocimiento de objetos en una cocina con una webcam*, proyecto de fin de carrera, Universidad Carlos III, Madrid, 2009.
- [30] R. Colombo, “*Deep Learning*” para reconocimiento de imágenes en Raspberry Pi 2, proyecto de fin de carrera, Universidad de la Laguna, Santa Cruz de Tenerife, 2016.
- [31] Y. Muñoz *et al.*, “Análisis de la escena en la cocina por medio de sensores IoT Diseñados basados en el microcontrolador node MCu ESP8266 y conectados al servidor ThingSpeak”, *Inf Tecnol*, vol. 30, pp. 173-170, 2019. DOI: 10.4067/S0718-07642019000500173.
- [32] R. B. Rusu *et al.*, “Robots in the kitchen: Exploiting ubiquitous sensing and actuation”, *Robotics and Autonomous Systems*, vol. 56, n.º 10, pp. 844-856, 2008. DOI: 10.1016/j.robot.2008.06.010.
- [33] Y. Quiñonez *et al.*, “Aplicación de técnicas evolutivas y visión por computadora para navegación autónoma de robots utilizando un TurtleBot”, *RISTI*, vol. 3, pp. 93-105, 2015. DOI: 10.17013/risti.e3.93-105.
- [34] J. Estupiñán *et al.*, “Implementación de algoritmos basados en máquinas de soporte vectorial (SVM) para sistemas eléctricos: revisión de tema”, *Tecnura*, vol. 20, n.º 10, pp. 149-170, 2016. DOI: 10.14483/udistrital.jour.tecnura.2016.2.a11.
- [35] R. Mosquera *et al.*, “Máquinas de Soporte Vectorial, Clasificador Naïve Bayes y Algoritmos Genéticos para la Predicción de Riesgos Psicosociales en Docentes de Colegios Públicos Colombianos”, *Inf. Tecnol*, vol. 29, n.º 6, pp. 153-162, 2018. DOI: 10.4067/s0718-07642018000600153.
- [36] M. Abadi *et al.*, “TensorFlow: A system for large-scale machine learning. TensorFlow: A System for Large-Scale Machine Learning”, en *XII Simposio USENIX sobre diseño e implementación de sistemas operativos OSDI'16*, Savannah, 2016.
- [37] L. Hao *et al.*, “TensorD: A tensor decomposition library in TensorFlow”, *Neurocomputing*, vol. 318, pp. 196-200, 2018. DOI: 10.1016/j.neucom.2018.08.055.
- [38] V. M. Araújo y M. P. Cota, “Software como um Serviço: Uma visão holística”, *RISTI*, n.º 19, pp. 145-157, 2016. DOI: 10.17013/risti.19.145-157.



# Current-Voltage Characteristics of a Self-Assembled DNA

Carlos José Páez-González\*\*

Jorge H Quintero-Orozco\*\*\*

Andrés Camilo García-Castro\*\*\*\*

Received: 01/12/2019 • Accepted: 18/03/2020

<https://doi.org/10.22395/rium.v19n37a11>

## Abstract

This paper shows a research on the transport properties of two-dimensional square lattice patterns built from a telomeric DNA sequence. A tight-binding model, and the recursive Green's function method were used. It is showed that the self-assembled DNA structures based on telomeric DNA strands have current-voltage (I-V) characteristics, with robust plateau structures that favor the scrutiny of DNA-lead, as well as interference effects. An increase of the current, dependent on the distance between the crosses in the self-assembled square lattice structures, is observed, which makes the system eligible for nanoelectronic applications.

**Keywords:** DNA; telomeric sequence; electronic transport; mesoscopic systems; tight-binding; Green's function.

---

\* This paper describes the result of the ongoing project "Propiedades Electrónicas y Ópticas de Nuevos Materiales 2D Determinadas por un Modelo Tight-Binding Multiorbital Optimizado" 2019 - 2020. This project is funded by the Universidad Industrial de Santander.

\*\* PhD. in Physics. Professor at the School of Physics, Ficomaco Research Group, Universidad Industrial de Santander. Email: [cjpaezg@uis.edu.co](mailto:cjpaezg@uis.edu.co). Orcid: 0000-0003-4037-064X.

\*\*\* PhD. in Physics. Professor at the School of Physics, Cimbios Research Group, Universidad Industrial de Santander. Email: [jhquinte@uis.edu.co](mailto:jhquinte@uis.edu.co). Orcid: 0000-0002-9394-4515.

\*\*\*\* PhD. in Physics. Professor at the School of Physics, Cimbios Research Group, Universidad Industrial de Santander. Email: [acgarcia@uis.edu.co](mailto:acgarcia@uis.edu.co). Orcid: <https://orcid.org/0000-0003-3379-4495>

## Curva característica de voltaje y corriente de un ADN autoensamblado

### Resumen

Este artículo muestra un trabajo en el que se investigan numéricamente las propiedades de transporte de los patrones de una red cuadrada bidimensional construida a partir de una secuencia de ADN telomérico, a través de un modelo *tight-binding* efectivo para la estructura electrónica, mientras que la corriente se obtiene dentro del marco de funciones de Green. Se muestra que las estructuras de ADN autoensambladas basadas en cadenas de ADN teloméricas tienen características de voltaje de corriente con una robusta secuencia de escalones que favorecen el control del ADN con los contactos, así como los efectos de interferencia. Se observan variaciones interesantes del mecanismo de percolación, que dependen de la competencia entre la longitud de localización y la distancia entre los cruces en las estructuras de red cuadrada bidimensional autoensambladas, las cuales hacen que el sistema sea elegible para aplicaciones nanoelectrónicas.

**Palabras clave:** ADN; secuencia telomérica; transporte electrónico; sistemas mesoscópicos; *tight-binding*; función de Green.

## Curva característica de voltagem e corrente de um DNA autoassembledo

### Resumo

Este artigo mostra um trabalho no qual se pesquisam numericamente as propriedades de transporte dos padrões de uma rede quadrada bidimensional construída a partir de uma sequência de DNA telomérico, por meio de um modelo *tight-binding* efetivo para a estrutura eletrônica, enquanto a corrente é obtida dentro do âmbito de funções de Green. Mostra-se que as estruturas de DNA autoassembledas baseadas em correntes de DNA teloméricas têm características de voltagem de corrente com uma robusta sequência de etapas que favorecem o controle do DNA com os contatos, bem como os efeitos de interferência. Observam-se variações interessantes do mecanismo de percolação, que dependem da competição entre a longitude de localização e a distância entre os cruzamentos nas estruturas de rede quadrada bidimensional autoassembledas, as quais fazem que o sistema seja elegível para aplicações nano eletrônicas.

**Palavras-chave:** DNA; sequência telomérica; transporte eletrônico; sistemas mesoscópicos; *tight-binding*; função de Green.

## INTRODUCTION

Today, there is an increasing interest in nanotechnology based in DNA, especially, after Seeman's works [1, 2]. Seeman converted the DNA molecule (unidimensional structure) into 2D and 3D complex asymmetric structures named self-assembled DNA, where single strands of DNA were hybridized to form stable crosses. This self-arrangement of branched artificial DNA is useful for designing and constructing electronic nanodevices [3 - 8] for its high nanoscale precision [9-11]. The self-assembled DNA has an accuracy at the nanoscale regime, *i. e.*, a much higher resolution than conventional lithography.

The electronic and transport properties of self-assembled DNA remain unexplored. Thus, as the electronic properties of electronic nanodevices based in the self-assembled DNA depend on their structural features [12-14], the study of these electronic and transport properties may become essential. Here, we contribute to the discussion by studying the electrical transport across of a self-assembled telomeric DNA: A real, simple, and periodic DNA sequence, rich in Guanine base pairs (TTAGGG) [15, 16].

## 1. THEORETICAL FORMULATION

In this study, we describe the transport properties of a self-assembled DNA molecule according to an heuristic tight-binding model where the self-assembled DNA molecule is coupled to two electrodes (right - R, and left - L), as shown in figure 1.

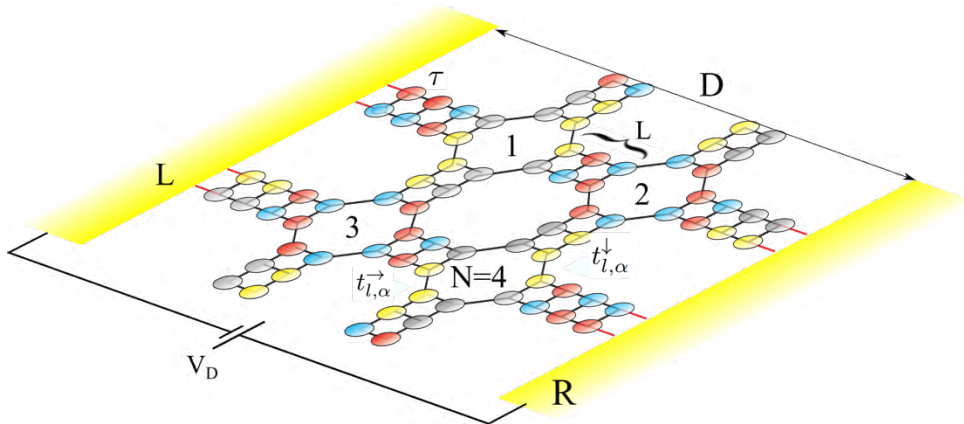


Figure 1. Schematic representation of a ladder-type model for electronic transport along self-assembled telomeric DNA between two semi-infinite source (L) and drain (R) contacts, as indicated

Source: Prepared by the authors.

The nucleotide base pairs sequence is indicated in figure 1 by circles, representing the four possible effective nucleotides. Different colors indicate different nucleotides: A (Red), T (blue), C (yellow) and G (gray). The possible electronic pathways are shown as lines.

We use a simple double chain model [17-19], where we take the correlation between pyrimidines (C, T) and purines (A, G) through base to base hopping terms. Thus, the Hamiltonian can be written as described in equation (1).

$$H_{DNA} = \sum_{l=1}^N \sum_{\alpha=1}^2 (\varepsilon_{l\alpha} |l, \alpha l, \alpha\rangle - t'_{l,\alpha} |l, \alpha l + 1, \alpha\rangle - t_{l,\alpha} |l, \alpha l, \alpha + 1\rangle) + h.c. \quad (1)$$

Here, N is the total number of base pairs (bps) in the two-dimensional self-assembled DNA,  $\alpha$  corresponds to the specific base pair that completes the correlation, and  $\varepsilon_{l\alpha}$  are the effective primary ionization energies of the base nucleotides, *i. e.*,  $\varepsilon_A = 8.24$  eV for adenine,  $\varepsilon_T = 9.14$  eV for thymine,  $\varepsilon_C = 8.87$  eV for cytosine and  $\varepsilon_G = 7.75$  eV for guanine. The base to base couplings are  $t' = 0.35$  eV between identical bases and  $t' = 0.17$  eV among different bases. Perpendicular couplings between pyrimidines or purines are  $t = 0.7$  eV. The choice of these parameters depends on the type of DNA, and its interaction with the environment. Environmental characteristics influence the load migration properties [20].

The transmission probability  $T(E)$  between the electrodes can be evaluated by equation (2):

$$T(E) = T_r [\Gamma_L G^r \Gamma_R (G^r)^\dagger], \quad (2)$$

Where  $\Gamma_{L(R)} = i (\Sigma_{L(R)} - \Sigma_{L(R)})$  is the broadening function, and is related to the lifetimes of electron in the systems.  $G^r$  is the retarded Green's function of the system which can be found from [21], according to equation (3).

$$G^r = [E - H_{DNA} - \Sigma_L - \Sigma_R]^{-1} \quad (3)$$

Here,  $E$  is a diagonal matrix, and  $\Sigma_{L(R)} = \tau_{L,(R)}^\dagger \mathbf{g}_{L,R} \tau_{L,(R)}$  are the self-energies for the source (left) and drain (right) contacts, respectively. The contact Green's function  $\mathbf{g}$ , in this case, a square lattice, was numerically estimated by using a recursive technique [22]. The electrode-molecule coupling  $\tau$  is determined by the geometry of the chemical bond. We use a constant coupling  $\tau = 0.35$  eV of similar magnitude as the inter-chain DNA couplings [23].

The current-voltage characteristic can be expressed as [24] shown in equation (4):

$$I = -\frac{2e}{h} \int_{\mu_L}^{\mu_R} T(E) dE, \quad (4)$$

Where the Fermi levels in source and drain are  $\mu_L$  and  $\mu_R = V_D + \mu_L$ , respectively.  $V_D$  is the voltage applied between the contacts. The use of this model has grown in



interest in the last decade [23] because, despite the simplified mathematics structure, it allows to study the electronic and transport properties of systems with more than 10,000 sites in contradistinction to *ab-initio* approaches.

## 2. RESULTS AND DISCUSSION

As it has been demonstrated, the transport of electrons through self-assembled DNA occurs by quantum tunneling. Thus, order or disorder in the structure play an important role in the characterization of these devices. Based on previous studies of transport in simple molecules [15], we studied the charge transport across an orderly self-assembled telomeric DNA. The telomeric sequence combines the advantages of a simple periodic structure with the richness and biological function of a real DNA sequence. Figure 2 shows the characteristic current-voltage I-V, behavior of a telomeric self-assembled DNA system, consisting of four horizontal chains and four vertical chains separated by 8 bps (red curve), 18 bps (black curve), and 23 bps (green curve).

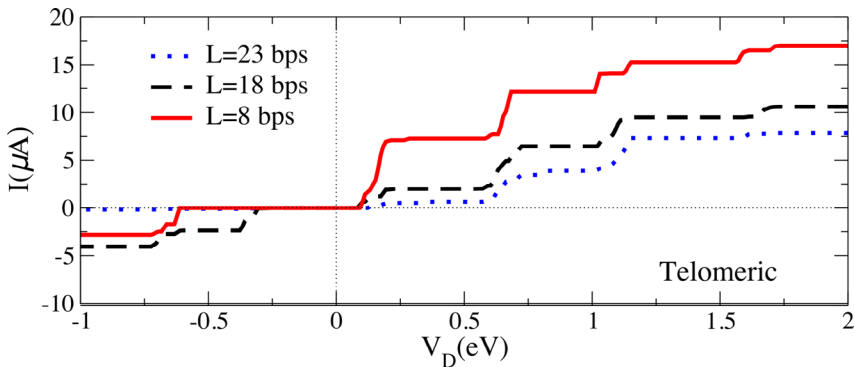


Figure 2. I-V characteristics for self-assembled telomeric DNA with different leg lengths ( $L = 8$  bps [red],  $L = 28$  bps [black], and  $L = 98$  bps [blue]) for a fixed number of crosses  $N = 4$

Source: Prepared by the authors.

The plateaus in the I-V characteristics are associated with gaps in the transmission spectrum. The increase in the two-dimensionality of the system introduced more plateaus in the I-V characteristic. This effect is caused by the increase in the number of transmission channels, which depends on the number of horizontal DNA chains linked to the contacts. Figure 2 shows that the robust effect of the plateaus on the I-V characteristic, common to the telomeric sequence, is conserved [15]. In the telomeric case, the wave function is delocalized over the entire system. For any energy, the electronic probability is higher near the contacts, and decreases towards the center of the system as the number of vertical chains is increased. An important feature is emphasized: In self-assembled telomeric DNA by increasing two-dimensionality, that is, the number of horizontal and vertical chains, the saturation current is increased. Two-dimensional

DNA models offer the opportunity to construct materials with nanometric precision, allowing them to be considered as possible components of functional devices. However, the main difficulty lies in its characterization.

Figure 3 shows the effect of the size ( $D$ ) of the auto-arrangement (square mesh) and the separation ( $L$ ) between the chains that make the auto-arrangement over the current. The red curve represents a system with  $L = 10$  bps; the black curve is a system with  $L = 20$  bps, and the blue curve is a system with  $L = 25$  bps.

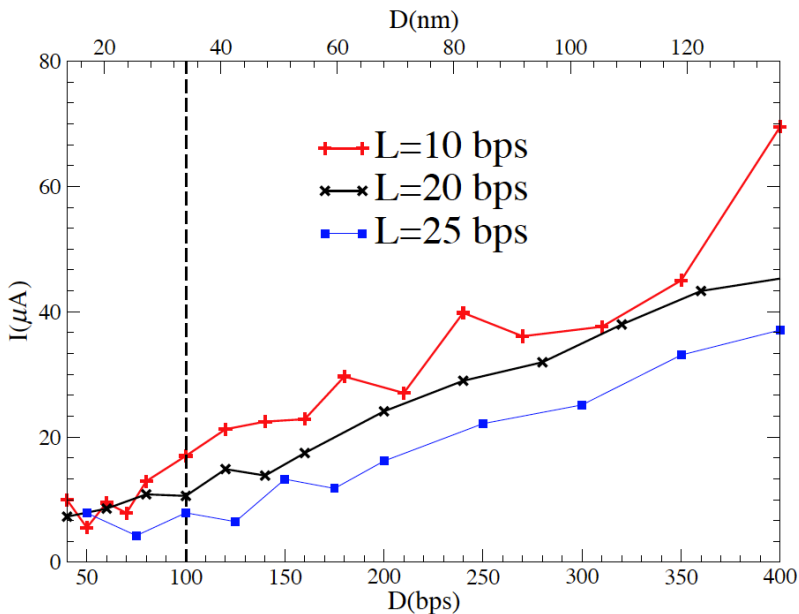


Figure 3. The current  $I$  as a function of length  $D$  at  $V_D = 1.5$  eV for telomeric at chains separations  $L=10$ ,  $L=20$ , and  $L=30$  base pairs

Source: Prepared by the authors.

For the telomeric case, regardless of the separation  $L$ , at a voltage  $V_D = 2$  eV, the current increases with increasing size  $D$ . For any size  $D$ , decreasing the separation  $L$  increases the current, due to an increase in the degree of relocation of the electronic states. We assumed that the coupled between the self-assembled telomeric DNA and the contacts is similar to a coupled between base-base pair. The DNA-contact bond can influence the I-V characteristic. Thus, we used a metallic contact. For the telomeric case, given a number of channels, if the separation between the chains ( $L$ ) increases, the current is increased. Our main contribution is to identify self-assembled telomeric DNA as candidates for the construction of 2D nanodevices such as DNA Origami, as it is possible to make more structured self-arrangements without

drastically decreasing current. Although the structures shown here are simple, their study should be the first step to denote more complex DNA devices, which could be used in a wide variety of circuits.

### 3. CONCLUSIONS

A study on the transport properties of self-assembled DNA using Green's function method was presented here. The self-assembled telomeric DNA that was the subject of this work provides a reference for the systematic investigation of possible DNA applications in nanoelectronics.

We present results on transport properties in the self-assembled Telomeric DNA system. I-V characteristic spectrum shows an increase of current with increasing systems size and a set of plateaus in the I-V characteristics, associated with gaps in the transmission spectrum, so the current is quantized. The increase in the two-dimensionality of the system introduced more plateaus in the I-V characteristic. This effect is caused by the increase in the number of transmission channels, which depend on the number of horizontal DNA chains linked to the contacts.

### ACKNOWLEDGMENTS [T2]

The authors acknowledge the financial support from VIE - Universidad Industrial de Santander, through project 2528. Additionally, all the authors have read and approved the final manuscript.

### REFERENCES

- [1] N. C. Seeman, and H. F. Sleiman, "DNA nanotechnology," *Nature Reviews Materials*, vol. 3, p. 17068, 11 2018. DOI: 10.1038/natrevmats.2017.68.
- [2] N. C. Seeman, "DNA in a material world," *Nature*, vol. 421, no. 6921, p. 427, 2003. DOI: 10.1038/nature01406.
- [3] G. M. Whitesides, and B. Grzybowski, "Self-assembly at all scales," *Science*, vol. 295, no. 5564, pp. 2418–2421, 2002. DOI: 10.1126/science.1070821.
- [4] J. R. McMillan *et al.*, "Protein Materials Engineering with DNA" *Acc. Chem. Res* vol. 52, no. 7, p. 1939, 2019. DOI: 10.1021/acs.accounts.9b00165.
- [5] J. Nangreave, D. Han, Y. Liu, and H. Yan, "DNA origami: a history and current perspective," *Current Opinion in Chemical Biology*, vol. 14, no. 5, p. 608, 2010. DOI: 10.1016/j.cbpa.2010.06.182.

- [6] A. Mangalum, M. Rahman, and M. L. Norton, "From DNA Nanotechnology to Material Systems Engineering" *advanced materials*, vol. 31, p. 1806294, 2019. DOI: 10.1002/adma.201806294.
- [7] J. Y. Kishi, T. E. Schaus, N. Gopalkrishnan, F. Xuan, and P. Yin, "Programmable autonomous synthesis of single-stranded DNA," *Nature Chemistry*, vol. 10, p. 155, 2017. DOI: <https://doi.org/10.1038/nchem.2872>
- [8] X. Liu, F. Zhang, X. Jing, M. Pan, P. Liu, W. Li, B. Zhu, J. Li, H. Chen, L. Wang, J. Lin, Y. Liu, D. Zhao, H. Yan, and C. Fan, "Complex silica composite nanomaterials templated with DNA origami," *Nature*, vol. 559, no. 7715, p. 593, 2018. DOI: <https://doi.org/10.1038/s41586-018-0332-7>
- [9] T. Lin, J. Yan, L. L. Ong, J. Robaszewski, H. D. Lu, Y. Mi, P. Yin, and B. Wei, "Hierarchical assembly of DNA nanostructures based on four-way toehold-mediated strand displacement," *Nano Letters*, vol. 18, no. 8, p. 4791, 2018. DOI: <https://doi.org/10.1021/acs.nanolett.8b01355>
- [10] D. Y. Tam, X. Zhuang, S. W. Wong, and P. K. Lo, "Photoresponsive self-assembled DNA nanomaterials: Design, working principles, and applications," *Small*, vol. 15, no. 26, p. 1805481, 2019. DOI: 10.1002/smll.201805481.
- [11] Y. Li, and R. Schulman, "DNA nanostructures that self-heal in serum," *Nano Letters*, vol. 19, no. 6, p. 3751, 2019. DOI: 10.1021/acs.nanolett.9b00888.
- [12] C. Páez, and P. Schulz, "Electronic localization at mesoscopic length scales: different definitions of localization and contact effects in a heuristic DNA model," *Eur. Phys. J. B*, vol. 86, p. 104, 2013. DOI: 10.1140/epjb/e2013-30728-9.
- [13] R. Gutiérrez, S. Mandal, and G. Cuniberti, "Quantum transport through a DNA wire in a dissipative environment," *Nano Letters*, vol. 5, no. 6, p. 1093, 2005. DOI: 10.1021/nl050623g.
- [14] C. J. Páez, and P. A. Schulz, "Delocalization of vibrational normal modes in double chains: Application to DNA systems," *Microelectronics Journal*, vol. 39, no. 11, p. 1222, 2008. DOI: 10.1016/j.mejo.2008.01.007.
- [15] C. J. Páez, P. A. Schulz, N. R. Wilson, and R. A. Roemer, "Robust signatures in the current-voltage characteristics of DNA molecules oriented between two graphene nanoribbon electrodes," *New Journal of Physics*, vol. 14, no. 9, p. 093049, 2012. DOI: 10.1088%2F1367-2630%2F14%2F9%2F093049.
- [16] C.-T. Shih, S. Roche, and R. Roemer, "Point-mutation effects on charge-transport properties of the tumor-suppressor gene p 53," *Physical review letters*, vol. 100, p. 018105, 02 2008. DOI: 10.1103/PhysRevLett.100.018105.
- [17] C. J. Páez, R. Rey-González, and P. A. Schulz, "Macroscopic localization lengths of vibrational normal modes in a heuristic DNA model," *Phys. Rev. B*, vol. 81, p. 024203, Jan 2010. DOI: 10.1103/PhysRevB.81.024203.

- [18] S. A. Wells, C.-T. Shih, and R. A. Römer, “Modelling charge transport in DNA using transfer matrices with diagonal terms,” vol. 23, p. 4138, 2009. DOI: 10.1142/S0217979209063328.
- [19] J. Rak, A. Voityuk, A. Márquez, and N. Rösch, “The effect of pyrimidine bases on the hole-transfer coupling in DNA,” vol. 106, no. 32, p. 7919, 2002. DOI: 10.1021/jp014261m.
- [20] M. P. L. Sancho, J. M. L. Sancho, J. M. L. Sancho, and J. Rubio, “Highly convergent schemes for the calculation of bulk and surface green functions,” *Journal of Physics F: Metal Physics*, vol. 15, no. 4, p. 851, 1985. DOI: 10.1088%2F0305-4608%2F15%2F4%2F009.
- [21] N. B. Larsen, H. Biebuyck, E. Delamarche, and B. Michel, “Order in microcontact printed self-assembled monolayers,” *Journal of the American Chemical Society*, vol. 119, no. 13, p. 3017, 1997. DOI: 10.1021/ja964090c.
- [22] D. Porath, A. Bezryadin, S. De Vries, and C. Dekker, “Direct measurement of electrical transport through DNA molecules,” vol. 403, p. 635, 2000. DOI: 10.1038/35001029.
- [23] H. Cohen, C. Noguez, R. Naaman, and D. Porath, “Direct measurement of electrical transport through single DNA molecules of complex sequence,” vol. 102, p. 11589, 2005. DOI: 10.1073/pnas.0505272102.
- [24] C. Lewenkopf, and E. Mucciolo, “The recursive green’s function method for graphene,” *Journal of Computational Electronics*, vol. 12, no. 203, p. 203, 2013. DOI: 10.1007/s10825-013-0458-7.



# Estado de corrosión en tuberías y accesorios metálicos expuestos al aire, basado en conjuntos difusos\*

*Fernando Gutiérrez Fique\*\**

*Milton Mena Serna\*\*\**

*César Augusto García-Ubaque\*\*\*\**

Recibido: 14/08/2019 • Aceptado: 13/04/2020

<https://doi.org/10.22395/rium.v19n37a12>

## **RESUMEN**

Los materiales metálicos usados en sistemas de distribución de agua potable presentan daños por corrosión, lo que hace necesario desarrollar métodos sencillos y accesibles, que permitan llevar a cabo medidas de tipo preventivo o correctivo con celeridad. En este sentido, con el trabajo del que se da cuenta en este artículo se pretendió construir un modelo de lógica difusa basado en datos obtenidos de inspecciones visuales del revestimiento y del grado de expansión de las apariencias de deterioro visibles (envejecimiento) de los elementos. El diseño metodológico utilizado correspondió a investigación cuantitativa, también conocida como enfoque matemático, que se caracteriza por la interpretación de datos, estadísticas, números y tablas, entre otros. Se analizaron los datos y registros de varias estaciones reductoras de presión (ERP) de las zonas 1 y 4 del acueducto y alcantarillado de Bogotá. Se utilizó Fuzzy Logic Toolbox™, complemento del software MATLAB, a partir de la selección de dos parámetros (protección/revestimiento y envejecimiento), a los cuales se les asignaron condiciones de calidad que operaron como datos de entrada del programa. Mediante la aplicación del modelo se obtuvo una superficie tridimensional con los valores de los estados de corrosión, según la calidad del revestimiento y la dimensión de envejecimiento. Se identificaron cuatro zonas, que representan los estados de corrosión de los elementos evaluados: óptimo, imperceptible, leve y grave, sin intervalo de transición. Los resultados demuestran que esta metodología provee una evaluación confiable del estado de corrosión de las estructuras metálicas evaluadas. Se concluye que la implementación del sistema de inferencia difusa proporciona una herramienta de sencilla aplicación, la cual puede ser ajustada con información recopilada en los trabajos rutinarios de mantenimiento preventivo o correctivo, y en actividades que relacionen otras variables con la corrosión.

**Palabras clave:** corrosión; inspección visual; lógica difusa; revestimiento; tuberías.

---

\*\* Ingeniero civil, Universidad Distrital Francisco José de Caldas. Correo electrónico: fgutierrezf@correo.udistrital.edu.co. Orcid: 0000-0003-3442-4285.

\*\*\* MSc., Universidad Distrital Francisco José de Caldas. Correo electrónico: mmenas@correo.udistrital.edu.co. Orcid: 0000-0002-7377-2787.

\*\*\*\* Ph. D., Universidad Distrital Francisco José de Caldas. Correo electrónico: cagarciau@udistrital.edu.co. Orcid: 0000-0002-6959-6610.

## Corrosion State in Metal Pipes and Fittings Exposed to Air Based on Fuzzy Sets

### Abstract

Metallic materials used in drinking water distribution systems show corrosion damage, which makes it necessary to develop simple and accessible methods that allow preventive and / or corrective measures to be carried out quickly. The work that is described in this paper intended to build a fuzzy logic model based on data obtained from visual inspections of the cladding, and the visible deterioration appearances degree (aging) of the elements. The methodological design used was quantitative research, also known as a mathematical approach, distinguished by the interpretation of data, statistics, numbers, and tables, among others. The data and records of various pressure reducing stations (ERPs) in zones 1 and 4 of aqueduct and sewer in Bogotá were analyzed. Fuzzy Logic Toolbox™ MATLAB software complementing was used, based on the selection of two parameters (protection, and coating and aging). Each one was assigned quality conditions that operated as the program input data. By applying the model, a three-dimensional surface was obtained with the values of the corrosion states, according to the quality of the coating and the aging dimension. Four zones were identified, which represent the corrosion states of the evaluated elements: optimal, imperceptible, light, and severe, with no transition interval. The results show that this methodology provides a reliable evaluation of the state of corrosion of the evaluated metal structures. It is concluded that the implementation of the fuzzy inference system provides a simple application tool, which can be adjusted with information collected in routine preventive or corrective maintenance work, and in activities that relate other variables to corrosion.

**Keywords:** Corrosion; visual inspection; diffuse logic; coating; pipelines.

## Estado de corrosão em encanamentos e acessórios metálicos expostos ao ar, baseado em conjuntos difusos

### Resumo

Os materiais metálicos usados em sistemas de distribuição de água potável apresentam danos por corrosão, o que torna necessário desenvolver métodos simples e acessíveis, que permitam tomar medidas de tipo preventivo ou corretivo com celeridade. Nesse sentido, com o trabalho que se realizou neste artigo pretendeu-se construir um modelo de lógica difusa baseado em dados obtidos de inspeções visuais do revestimento e do grau de expansão das aparências de deterioração visíveis (envelhecimento) dos elementos. O desenho metodológico utilizado correspondeu à pesquisa quantitativa, também conhecida como enfoque matemático, que se caracteriza pela interpretação de dados, estatísticas, números e tabelas, entre outros. Analisaram-se os dados e registros de várias estações redutoras de pressão (ERP) das zonas 1 e 4 do Sistema de Saneamento Básico de Bogotá. Utilizou-se Fuzzy Logic Toolbox™, complemento do software MATLAB, a partir da seleção de dois parâmetros (proteção/revestimento e envelhecimento), aos quais lhes atribuíram condições de qualidade que operaram como dados de entrada do programa. Mediante a aplicação do modelo, obteve-se uma superfície tridimensional com os valores dos estados de corrosão, segundo a qualidade do revestimento e a dimensão de envelhecimento. Identificaram-se quatro zonas, que representam os estados de corrosão dos elementos avaliados: ótimo, imperceptível, leve e grave, sem intervalo de transição. Os resultados demonstram que essa metodologia provê uma avaliação confiável do estado de corrosão das estruturas metálicas avaliadas. Conclui-se que a implementação do sistema de inferência difusa proporciona uma ferramenta de simples aplicação, a qual pode ser ajustada com informação recopilada nos trabalhos rotineiros de manutenção preventiva ou corretiva, e em atividades que relacionem outras variáveis com a corrosão.

Palavras-chave: corrosão; inspeção visual; lógica difusa; revestimento; encanamentos.



## INTRODUCCIÓN

Los sistemas de distribución de agua potable que utilizan tuberías y accesorios metálicos (hierro fundido, hierro dúctil y acero) expuestos al aire son susceptibles a la corrosión electroquímica debido a la interacción del metal con la lluvia, la humedad y el flujo de agua dentro de la tubería [1] [2]. Estos elementos pueden presentar múltiples apariencias de corrosión [3], por lo que se han establecido especificaciones de protección interior y exterior [4] de las tuberías metálicas para garantizar su durabilidad y la calidad del servicio de agua potable [5] [6]. Sin embargo, la aplicación incorrecta de la capa de recubrimiento permitiría la pronta aparición de la corrosión en los elementos metálicos, hecho que reduciría su vida útil y aumentaría los costos de mantenimiento.

En la actualidad no existen equipos de medición de la corrosión [7] para hacer seguimiento y evaluación del estado actual de los elementos. Por ello se utilizan y desarrollan técnicas especializadas en detectar y calificar determinados mecanismos de daño [8] [9] y suelen emplearse para todas las condiciones, habida cuenta de que cada tipología de daño y material tiene una técnica asociada disponible. Así, la inspección visual proporciona una caracterización del material [10] y, por lo tanto, puede ser utilizada como método de evaluación de corrosión externa e interna en tuberías y accesorios metálicos [11].

Por todo lo anterior, el objetivo de estudio mostrado aquí fue presentar una metodología para estimar de forma cuantitativa el estado de corrosión de las tuberías y accesorios expuestas al aire. El modelo se basa en los estados de corrosión que pueda presentar una estructura metálica corroída. Estos estados se han determinado mediante la elección de dos parámetros, clasificados de acuerdo con dos o tres condiciones de calidad. Dado que el modelo propuesto involucra conocimientos y experiencias de personas (trabajadores, expertos y equipo de trabajo) se decidió utilizar la herramienta de lógica difusa, la cual permite producir resultados exactos a partir de información imprecisa [12] [13].

## 1. METODOLOGÍA

El diseño metodológico utilizado correspondió a investigación cuantitativa, también conocida como enfoque matemático; se caracteriza por la interpretación de datos, estadísticas, números y tablas, entre otros. Así entonces, se analizaron los datos y registros de varias estaciones reductoras de presión (ERP) de las zonas 1 y 4 del acueducto y alcantarillado de Bogotá (Colombia).

## 1.1 Técnicas de inspección de tuberías actuales

Existe una variedad de técnicas destructivas y no destructivas para evaluar de forma directa alguna forma de corrosión. Los cupones y núcleos permiten conocer la integridad estructural de la tubería y determinar las propiedades del material expuesto a la corrosión. Las pruebas no destructivas (NDT) incluyen métodos ultrasónicos, electromagnéticos y radiografías que permiten localizar áreas con problemas, reducción de espesor de pared, y fugas. También existen pruebas indirectas que permiten correlacionar con la corrosión algunas propiedades de calidad de agua y del suelo (en caso de tuberías enterradas) [14]. No obstante, el procesamiento de la información y datos obtenidos con estos equipos varía ampliamente, ya que dependen de la interpretación del operador.

## 1.2 Aplicación del modelo de lógica difusa (LD)

Para la implementación de LD se utilizó Fuzzy Logic Toolbox™, complemento del *software* MATLAB [15]. La aplicación de LD requiere tres pasos: selección de parámetros y construcción de funciones de pertenencia; definición de reglas; y *defuzzificación*.

### 1.2.1 Selección de parámetros y construcción de funciones de pertenencia [T4]

Dado que múltiples factores intervienen en la corrosión, el modelo propuesto se limita a dos parámetros: protección y envejecimiento, cuya estimación no requiere ningún equipo especializado. El parámetro de protección o revestimiento cuenta con 3 condiciones de calidad en el modelo:

- a. Bueno: recubrimiento con correcta adherencia y espesor requerido.
- b. Regular: recubrimiento con adherencia incompleta y espesor mínimo.
- c. Malo: Recubrimiento sin adherencia, no cumple espesor mínimo y presenta imperfecciones como ampollas, grietas, vacíos o picadoras. También aplica cuando se usa un recubrimiento que no cumple las normas técnicas de instalación de tuberías y accesorios metálicos para acueducto. En este parámetro se usó la función de pertenencia trapezoidal (figura 1).

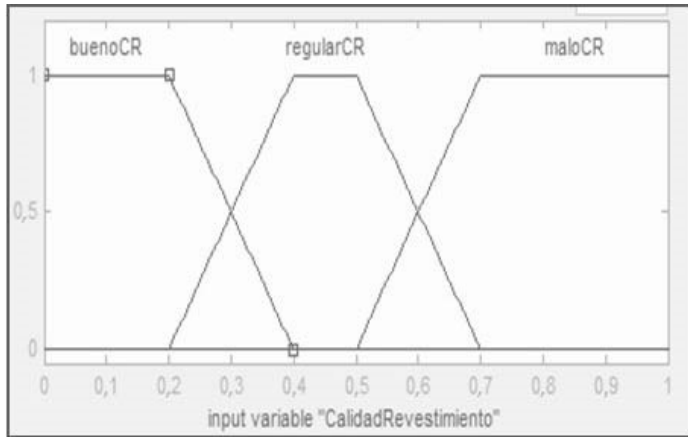


Figura 1. Función de pertenencia, parámetro 'revestimiento'

Fuente: elaboración propia.

El parámetro de envejecimiento tiene definidas dos condiciones de dimensión de las áreas de actividad de corrosión, que puede haber ocurrido o estar ocurriendo. Estas definiciones toman en cuenta la norma ASME B31G y una relación de proporción (ecuación 1).

- Alto: apariencias de corrosión localizadas en toda superficie del elemento, o pérdida de recubrimiento por corrosión.
- Bajo: apariencias de corrosión localizadas puntualmente en el elemento, sin presencia de corrosión visible.

En este parámetro también se utilizó la función de pertenencia trapezoidal como se muestra en la figura 2.

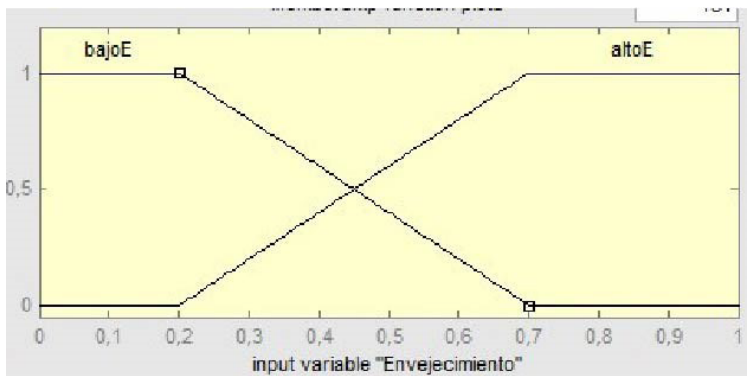


Figura 2. Función de pertenencia, parámetro 'envejecimiento'

Fuente: elaboración propia.

### 1.2.2 Definición de reglas

Para la construcción de las reglas de decisión se emplearon las combinaciones de las funciones de pertenencia de las entradas con apreciaciones basadas en la opinión de expertos, ingenieros y operarios de la empresa.

### 1.2.3 Defuzzificación

El sistema de inferencia difusa (FIS por sus iniciales en inglés) desarrollado permite obtener cualitativa y cuantitativamente el estado de corrosión de las tuberías expuestas al aire libre a partir de los parámetros y reglas definidos con anterioridad. Para este caso se empleó como método de *defuzzificación* el método del centroide, el cual obtiene un resultado único a partir de la función característica de salida [16]. La estructura básica del sistema se muestra en la figura 3.

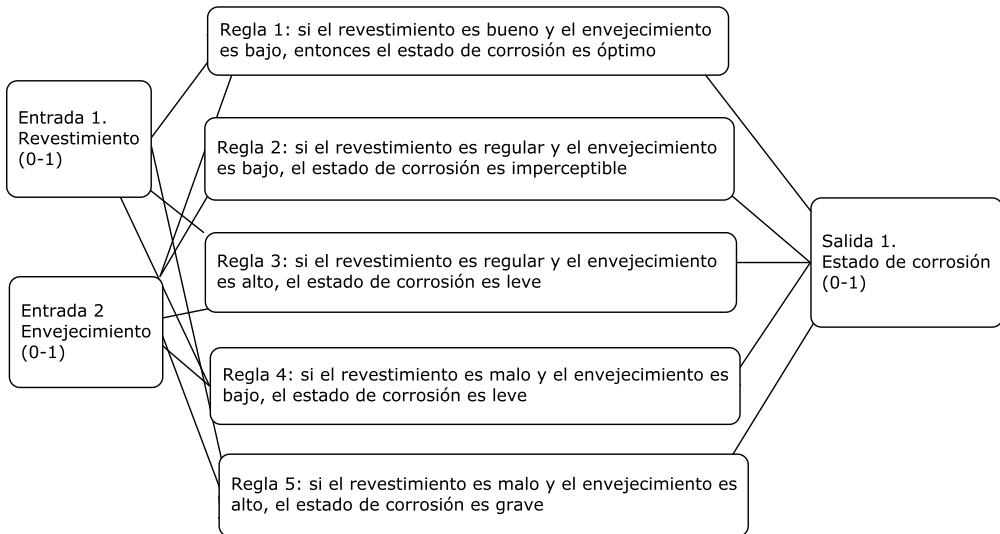


Figura 3. Esquema del sistema de lógica difusa

Fuente: elaboración propia.

Los estados de corrosión a partir de números difusos se basan en las variables lingüísticas ‘óptimo’, ‘imperceptible’, ‘leve’ y ‘grave’. Cada una de ellas fue representada por la función de pertenencia mostrada en la figura 4.

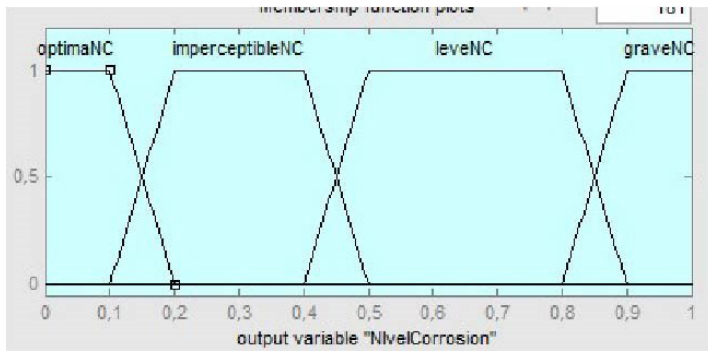


Figura 4. Función de pertenencia del parámetro de estado de corrosión

Fuente: elaboración propia.

## 2. RESULTADOS

A continuación, se muestra una superficie tridimensional con los valores de los estados de corrosión del FIS, según la calidad del revestimiento y la dimensión de corrosión (figura 5):

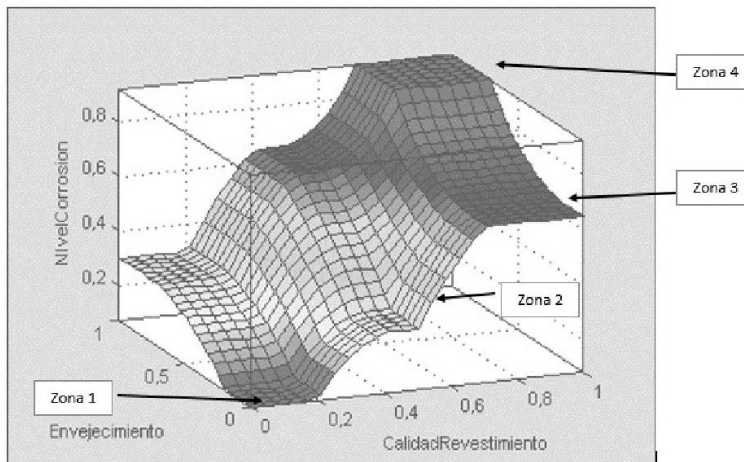


Figura 5. Visor de superficie de nivel corrosión

Fuente: elaboración propia.

A partir de esta imagen es posible hacer dos inferencias: primero, las zonas 1 - 4 representan los estados de corrosión ‘óptimo’, ‘imperceptible’, ‘leve’ y ‘grave’, respectivamente, sin intervalo de transición; y segundo, entre las zonas se encuentran los estados de corrosión con pertenencia a dos estados de corrosión. Estos datos demuestran que la metodología presentada brinda respuestas confiables del estado de corrosión de las estructuras metálicas.

En la tabla 1 se presenta el promedio de años de servicio para cada estación, de acuerdo con los datos analizados.

Tabla 1. Promedio de años de servicio por cada estado de corrosión

Estado de corrosión	Años de servicio
Óptimo	2
Imperceptible	5,8
Leve	9,6
Grave	17,1
Daño total	35

Fuente: elaboración propia.

Luego se realizó la regresión logarítmica de esta información y se obtuvo una función aproximada de los estados de corrosión en función del tiempo (figura 6).

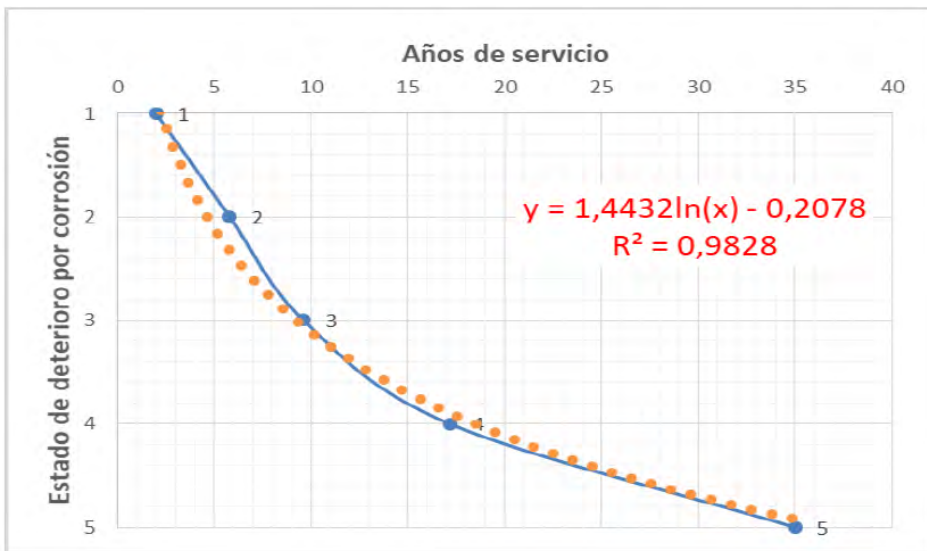


Figura 6. Línea de tendencia logarítmica

Fuente: elaboración propia.

Esta regresión tiene un coeficiente de determinación de 0,9828, valor cercano a 1, lo cual indica que la función representa el estado de corrosión en relación con los años de servicio en un 98,28 % por lo que la función obtenida es apropiada para este caso.

### 3. CONCLUSIONES

La implementación del FIS ha permitido incluir una herramienta de sencilla aplicación, que puede ajustarse si es preciso con información recopilada en los trabajos rutinarios

de mantenimiento preventivo o correctivo, y trabajos que relacionen variables con la corrosión. Además, provee información para las empresas prestadoras de servicio de agua potable, ya que permite determinar el grado de deterioro actual de las estructuras metálicas, habida cuenta de que las apreciaciones de estado de corrosión se limitan a las variables de calidad de revestimiento y de las dimensiones de las formas de corrosión presentes en la inspección visual.

La metodología propuesta se puede complementar con el análisis que se muestra en la tabla 2: esta ilustra la relación entre las variables de entrada y salida del algoritmo, lo cual permite una interpretación gráfica de los resultados antes y después del proceso del algoritmo.

Tabla 2. Interpretación gráfica del estado integral de la tubería basado en las dos variables de entrada

		Recubrimiento		
		Bueno	Regular	Malo
Envejecimiento	Bajo			
	Alto			

Fuente: elaboración propia.

De otra parte, también se propone utilizar la norma ASME B31G [17] como referencia para evaluar los distintos niveles de corrosión que puede tener el material, de tal manera que fuese posible estandarizar los tipos de corrosión (envejecimiento) que pudiera tener la tubería para aplicarlo como variable de entrada en el algoritmo. En este caso se realizan medidas directamente en la tubería y se obtiene un resultado de longitud o presión de falla, de acuerdo con cada caso. Al comparar estos resultados se puede llegar a tener una mejor aproximación del nivel alto o bajo de envejecimiento, que puede ser más preciso para utilizarlo como parámetro de entrada.

ASME B31G [17] propone tres niveles de evaluación que determinan el grado de oxidación en una pieza, lo cual permite determinar de manera analítica y cuantitativa si el envejecimiento es alto o bajo. Ello complementa el procedimiento y cumple el objetivo de tomar mediciones y decisiones de manera efectiva y rápida, con variables de entrada fáciles de medir. La evaluación nivel 0 trae consigo resultados conservadores de la integridad de la tubería respecto a la oxidación que posee; al establecer una relación entre estas dos variables se puede determinar si la tubería se ajusta al nivel bajo o alto (figura 7).

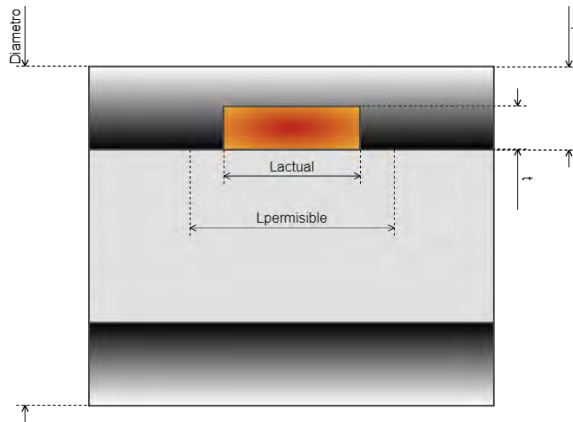


Figura 7. Características dimensionales de la tubería y la grieta

Fuente: elaboración propia.

Dado que la corrosión es un tema extenso, se podría seguir avanzando con base en estudios que permitan determinar el estado de corrosión actual con equipos sencillos. Además, se deberían realizar campañas de capacitación de corrosión en elementos metálicos al personal operativo y líderes de los procesos de mantenimiento de acueductos, a fin de brindar soluciones prácticas ante este fenómeno.

La evaluación 0 va a tomar como variables de entrada las características dimensionales de la tubería y traerá consigo, como resultado del proceso, la longitud permisible de la grieta que se ha generado. Luego de que se tenga esta longitud, se puede hacer una relación entre ambas longitudes (permisible y actual) para establecer de manera cuantitativa si el envejecimiento es alto o bajo (ecuación 1):

$$\frac{L_{actual}}{L_{permisible}} = L \quad (1)$$

Si,  $0 < L \leq 0,5 \rightarrow$  envejecimiento bajo; y si  $0,5 < L < 1 \rightarrow$  envejecimiento alto.

## REFERENCIAS

- [1] N. Asrar, B. Mackay, Ø. Birketveit, M. Stipanicev, J. Jackson, A. Jenkins, D. Melot., J. Sheie, J. Vittonato. "La corrosión: La lucha más extensa". *Oilfield Review*, vol. 28, n.º 2, pp. 36-51, 2016.
- [2] R. Galván-Martínez, D. Cabrera-de la Cruz, A. Contreras, R. Orozco-Cruz. "A novel experimental arrangement for corrosion study of X60 pipeline steel weldments at turbulent flow conditions". *Corrosion Engineering, Science and Technology*, vol 51, n.º 6, pp. 80-91, 2016. DOI: 10.1080/1478422X.2015.1124598.



- [3] W. Smith, “Corrosión”, en *Fundamentos de la ciencia e ingeniería de materiales*, Madrid: McGraw-Hill Interamericana, 1998, pp. 718-777.
- [4] American Water Works Association AWWA, *C210-15 Liquid-Epoxy Coatings and Linings for Steel Water Pipe and Fittings*. Denver: American Water Works Association, 2015.
- [5] J. Rebollo, “Técnicas de protecciones de tuberías y accesorios”, en *Montaje de redes de distribución de agua*, Bogotá: Ediciones de la U, 2014.
- [6] J. Torres, E. Meraz y L. Veleza, “Evaluación de un sistema de protección catódica de un gasoducto enterrado”, *Revista Latinoamericana de Metalurgia y Materiales*, vol. 37, n.º 1, pp.19-26, 2017.
- [7] V. Tzatchkov, V. Alcocer, V. Bourguett y F. Arreguín. “Rehabilitación de tuberías de agua potable”, en: *Avances en la hidráulica de redes de distribución de agua potable*, México D.F.: Instituto Mexicano de Tecnología del Agua, 2014, pp. 363-376.
- [8] F. Pancorbo, *Corrosión, degradación y envejecimiento de los materiales empleados en la edificación*, Bogotá: Alfaomega, 2013.
- [9] G. Macioski, D. de Souza, A. Capraro y H. Medeiros. “Analysis of steel bars corrosion as a function of the environment Ph”. *Revista ALCONPAT*, vol. 6, n.º 3, pp. 223-234, 2016. DOI: 10.21041/rav6i3.153.
- [10] M. Fontana, “Corrosion Principles”, en *Corrosion Engineering*, Singapur: McGraw-Hill, 1987, pp. 12-38.
- [11] A. Lee. “Condition Assessment Technologies for water transmission and sewage conveyance systems”, en *UBC Sustainability scholar program 2017 & Metro Vancouver*, Vancouver, 2017. [internet]. Disponible en [https://sustain.ubc.ca/sites/sustain.ubc.ca/files/Sustainability%20Scholars/2017\\_Sustainability\\_Scholars/Final\\_Reports/Condition%20Assessment%20Technologies%20for%20Water%20Transmission%20and%20Sewage%20Conveyance%20Systems\\_Lee\\_%202017%20SS.pdf](https://sustain.ubc.ca/sites/sustain.ubc.ca/files/Sustainability%20Scholars/2017_Sustainability_Scholars/Final_Reports/Condition%20Assessment%20Technologies%20for%20Water%20Transmission%20and%20Sewage%20Conveyance%20Systems_Lee_%202017%20SS.pdf).
- [12] G. Morales, *Introducción a la lógica difusa*, México D.F.: Cinvestav-IPN, 2002.
- [13] M. Biezma, D. Agudo y G. Barrón. “A Fuzzy Logic method: Predicting pipeline external corrosion rate”, *International Journal of Pressure Vessels and Piping*, vol. 163, pp. 56-62. 2018. DOI: 10.1016/j.ijpvp.2018.05.001.
- [14] J. Thomson y L. Wang, “Technologies for condition assessment”, en *Condition Assessment of Ferrous Water Transmission and Distribution Systems*, Washington: EPA, 2009, pp. 14-85.
- [15] MathWorks, *Design and simulate fuzzy logic systems*. [internet]. Disponible en <https://la.mathworks.com/products/fuzzy-logic.html>.
- [16] R. Pérez, *Procesado y optimización de Espectros Raman mediante Técnicas de Lógica Difusa: Aplicación a la identificación de Materiales Pictóricos*, disertación doctoral,

Departamento de Teoría de la Señal y Comunicaciones, Universidad Politécnica de Cataluña, Barcelona, 2005.

- [17] The American Society of Mechanical Engineers, *Manual for Determining the Remaining Strength of Corroded Pipelines B31G – 2012*, Nueva York: ASME, 2012.

# Web Service Adaptation for the Customization of Early Alerts in Agriculture\*

Oscar Ricardo Valencia Aguilar\*\*

Emmanuel Lasso\*\*\*

Juan Carlos Corrales\*\*\*\*

Received: 18/02/2019 • Accepted: 14/02/2020

<https://doi.org/10.22395/rium.v19n37a13>

## Abstract

Early warning systems are designed to inform the largest number of users, such as a country or a region, about a risky situation. However, in specific domains such as agriculture, it is commonly required for these alerts to be more specific according to the crops location and their properties. Consequently, the web services of these systems must be adapted. On the other hand, the enterprise services bus, with its mediation capabilities (such as message transformation and routing), and complex event processing with their monitoring characteristics, can be integrated to meet the adaptation requirements of web services at runtime. This paper presents an improvement for an early warning system for coffee production that, according to the area in which a crop is located and its phenology, manages the adaptation of alerts for coffee rust, based on the integration of an enterprise services bus and a complex events processing.

**Keywords:** Web service adaptation; early warning systems; enterprise service bus; complex event processing.

---

\*\* Ingeniero de Sistemas, Universidad del Cauca. Ms(c) en Ingeniería Telemática, Universidad del Cauca. Investigador, Universidad del Cauca, Grupo Ingeniería Telemática. Correo electrónico: ovalencia@unicauca.edu.co. Orcid: 0000-0002-6224-0777.

\*\*\* Ingeniero en Electrónica y Telecomunicaciones, MsC. en Ingeniería Telemática y Ph. D. (c) en Ingeniería Telemática, Universidad del Cauca. Investigador, Universidad del Cauca, Grupo de Ingeniería Telemática. Correo electrónico: eglasso@unicauca.edu.co. Orcid: 0000-0002-4950-5048.

\*\*\*\* Ingeniero en Electrónica y Telecomunicaciones y MsC. en Ingeniería Telemática, Universidad del Cauca. Ph. D. en Ciencias de la Computación, Université de Versailles Saint-Quentin-en-Yvelines. Docente de planta, Universidad del Cauca, director del Grupo de Ingeniería Telemática. Correo electrónico: jcorral@unicauca.edu.co. Orcid: 0000-0002-5608-9097.

## Adaptación de servicios web para la personalización de alertas tempranas en agricultura

### Resumen

Los sistemas de alerta temprana están diseñados para informar a un gran número de usuarios —un país o una región— sobre una situación de riesgo. Sin embargo, en dominios específicos como la agricultura suele requerirse que estos sean más específicos, de acuerdo con la ubicación del cultivo y sus propiedades. Por lo tanto, los servicios web de estos sistemas requieren adaptarse. Así mismo, el bus de servicios empresariales con sus capacidades de mediación, así como el procesamiento de eventos complejos con sus características de monitorización, pueden integrarse para cumplir con los requisitos de adaptación de los servicios web en tiempo de ejecución. Este artículo presenta la adaptación de los servicios web de acuerdo con la ubicación de un cultivo y su fenología, manejando la adaptación de alertas para la roya en el café, con base en la integración del bus de servicios empresariales y el procesamiento de eventos complejos.

**Palabras clave:** adaptación de servicios web; sistemas de alerta temprana; bus de servicios empresariales; procesamiento de eventos complejos.

## Adaptação de serviços web para a personalização de alertas precoces em agricultura\*

### Resumo

Os sistemas de alerta precoce estão desenhados para informar a um grande número de usuários —um país ou uma região— sobre uma situação de risco. No entanto, em domínios específicos como a agricultura costuma-se requerer que estes sejam mais específicos, de acordo com a localização do cultivo e suas propriedades. Portanto, os serviços web desses sistemas requerem adaptar-se. Desse modo, o *Enterprise Service Bus* com suas capacidades de mediação, bem como o processamento de eventos complexos com suas características de monitoramento podem ser integrados para cumprir os requisitos de adaptação dos serviços web em tempo de execução. Este artigo apresenta a adaptação dos serviços web de acordo com a localização de um cultivo e sua fenologia, gerenciando a adaptação de alerta para a ferrugem no café, com base na integração do *Enterprise Service Bus* e o processamento de eventos complexos.

**Palavras-chave:** adaptação de serviços web; sistemas de alerta precoce; *Enterprise Service Bus*; processamento de eventos complexos.

## INTRODUCTION

In agriculture, farmers are faced with several problems during the production of different crops [1] [2]. The diseases and pests can affect crops, decreasing the quality and quantity of production. Generally, their origin is due to the attack of fungi, bacteria, plagues, and other agents. The generation of diseases mainly depends on the physiological status of the crop, meteorological conditions, and production zones of the country, among others.

Over time, diseases have devastated crops throughout the world, causing epidemics and losses [3] [4] [5]. This implies that important measures must be taken to mitigate the effects and improve productivity through practices such as crop protection and clean technologies, among others. For this reason, different tools that allow the detection of different diseases for subsequent treatment are used. An early warning system (EWS) can perform this detection through monitoring and analysis of crops conditions, which allow to provide timely information to the farmers about possible natural, biotic, and physical threats in such a way that they can prepare and take the necessary measures according to the level of crop risk [6].

The EWSs are usually designed to inform a large number of users, *e. g.*, a large population such as a country or a region, of any risk situation. In the agriculture, these alerts should be more specific, so that they include characteristics as the location of the crop, as well as some of its properties. Taking into account the above, an improved EWS for diseases in crops is proposed in this paper for the generation of alerts and their adaptation through the integration of an enterprise services bus (ESB) and complex events processing (CEP), according to the area in which the crop is located and its phenology.

This paper is organized as follows: section 1 contains the state of the art; section 2 describes the solution proposed from the integration of the ESB and CEP to the EWS; section 3 presents the case study; section 4 presents the evaluation of the solution through a case study; section 5 describes the results obtained, and section 6 relates the conclusions and future work.

### 1. STATE OF THE ART

This section provides relevant information about studies carried out in the development of EWSs in agriculture, and personalization and adaptation of web services.

The EWSs are widely used in different domains as the detection of diseases in crops. These include key elements such as knowledge and threat mapping, monitoring and forecasting of impending events, and diffusion of alerts [7]. The correct combination

of these elements results in a complete and reliable system that provides timely and understandable information to the population. In addition, an EWS should be people-centered, allowing communities threatened by a phenomenon to be able to act in a timely and appropriate manner to reduce potential losses.

In [8], weather data is included within an EWS for *late blight* in potato crops. Eight experimental plots with weather monitoring were used. After the data were analyzed and processed, a fixed calendar was obtained, containing the alarm dates for the disease occurrence and treatments to be carried out in the crops. In [9], an EWS to mitigate the impact of cockcrowding on coffee crops is presented. This system is based on mathematical models that use meteorological data to predict the favorable conditions for the development of the disease, and recommend the fungicides application at the right time, thus preventing its development. On the other hand, weather data monitored in crops is used in [10] as support for decision making, helping producers to determine the right moment for disease control.

Some EWSs make use of web services to provide data to the user. Precisely, the web services personalization collects user information during its interaction with the system, so that it is used to provide adequate contents and services for each user, in order to improve the user experience [11]. The purpose of using web service personalization is to improve an EWS with relevant information of each user, allowing that its response can be delivered only to those potentially affected.

Furthermore, web services can be modified in such a way that they can be adjusted to different circumstances or conditions, either by the specific user needs or by the environment in which they are running [12]: this feature is called *adaptation*. In [13], an approach using the mediation characteristics of the ESB to provide a dynamic adaptation of the services is presented. In the same way, the research presented in [14] describes an ESB implementation using JBossESB, focused on the description of a services adaptation framework based on the ESB capabilities previously mentioned. The research described in [15] proposes a solution that uses the ESB mediation patterns to adapt the services to a context, making use of the CEP to analyze the events received from external resources, and to detect the relevant situations that require an adaptation.

The studies presented above address the mediation features provided by the ESB to perform the adaptation of services, and some use CEP to identify context events (e. g. climate) that are relevant to the user. Although they consider an integration with CEP, this is conceived as an external and additional component of the ESB. In fact, some studies propose the use of the ESB to integrate event-driven architectures using CEP [16].

Therefore, we propose to carry out the adaptation of web services based on an ESB-CEP integration, making use of the mediation capabilities of the ESB to combine services-oriented architectures (SOA) for the generation of intercommunicated services [17] and event-driven architectures (EDA) to design and implement applications based on events [18]. The combination of these architectures allows the integration of different applications in heterogeneous environments and, through ESB, these can be unified to operate in specific areas. For the same purpose, the CEP monitors relevant situations (events) according to the system requirements. Our approach proposes a system with the ability to react to environmental events (expected or unexpected), in order to ensure the proper operation of web services.

## 2. ESB – CEP INTEGRATION FOR WEB SERVICES ADAPTATION

Although in [14] and [15] an adaptation of the services using the mediation characteristics of the ESB is implemented, these proposals do not use CEP to detect relevant events that require adaptation. Our proposal is aimed at the web services adaptation around the *monitoring* component of an EWS, since these systems do not provide solutions for a dynamic and automatic adaptation at runtime based on users' needs [19]. This is addressed through the integration between ESB and CEP, acting as a mediator element through which all messages are transported and can be transformed, routed, and invoked, among others. Also, it provides the analysis, detection, and correlation of the messages that are exchanged by the system actors, in order to identify relevant events of interest in an application scenario.

Therefore, following the approach proposed in [14], where the adaptation is performed within the ESB using its mediation capabilities, we introduce the model presented in figure 1.

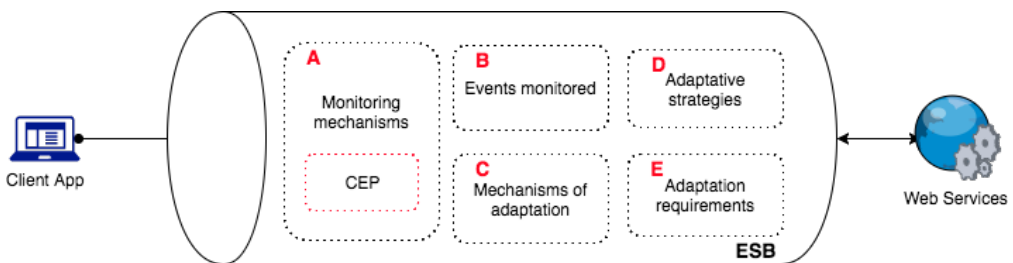


Figure 1. Model of Web Services Adaptation

Source: Prepared by the authors.

Figure 1 shows the ESB-CEP integration to perform the adaptation process, which contains the elements presented in the adaptation and monitoring framework proposed

by the S-Cube project [20]. This framework was adjusted to include the CEP within the ESB. The components are presented below:

- *Monitoring mechanisms* check if the current situation corresponds to the expected one. This component can be addressed through the complex event processing (CEP) in charge of monitoring the events of interest.
- *Monitored events* provide relevant information on the execution, evolution, or context of the services, detected through the monitoring mechanisms. These events represent the difference in the status, functionality or expected environment.
- *Adaptation requirements* represent the need to modify a service to eliminate the difference between the current and expected state. These may include requirements for reliability or functional accuracy, optimization, interoperability, and usability.
- *Adaptation mechanisms* contain the techniques and facilities provided by the ESB to carry out strategies according to the adaptation requirements.
- *Adaptation strategies* are alternatives to meet adaptation requirements according to the current situation, *e. g.*, to invoke an equivalent operation that considers new parameters. Different strategies can be used for the same situation, which requires decision mechanisms to select the most appropriate.

The general process proposed in figure 1 is executed as follows: a) relevant information is collected through monitoring mechanisms; b) critical events are detected; c) the need for adaptation is identified; d) a strategy for adaptation is identified; and e) adaptation is executed using the available adaptation mechanisms [20].

The model of adaptation described above can be applied to any environment. We applied it for the generation of early warnings for crop diseases, taking as a particular case the Coffee Rust, which will be explained in the next section.

## 2. CASE STUDY: COFFEE RUST DETECTION IN COLOMBIAN CROPS

In this section, we present an application of the proposal mentioned above, describing the problem domain and the cases of study to apply our solution.

### 2.1 Application domain

Coffee Rust is the main high-impact phytosanitary problem for coffee cultivation, caused by the pathogen *Hemileia vastratix* [21], a fungus that develops only in the living tissue of its host plant (for this case, the coffee leaves). In Colombia, this disease affects more than one million families that depend on coffee production. The consequences related



to the disease are mainly related to economic and quality losses for coffee farmers. There is, in fact, an urgent need to take preventive measures for reducing the losses caused by this disease.

Some proposals [22 - 28] have been made from the computer sciences related to the automatic identification of conditions conducive to this disease, making use of machine learning techniques. In these studies, the authors make use of the different variables involved in the development of coffee rust in crops, which have been monitored in them. In the particular case of [25], an induction of decision trees to generate a model that predicts an infection rate of coffee rust from monitored data in crops is presented. This rate indicates whether the infection level of the disease tends to remain stable or reduced, to increase moderately, or to accelerate. The results of this research can be used as the main element to alert coffee producers about favorable conditions (weather and agronomic properties) present in crops for the disease.

This last research is implemented within the AgroCloud project, which belongs to the Interinstitutional Network of Climate Change and Food Security of Colombia (RICCLISA). In this project, a web platform [29] that aims to provide monitoring services to warn users about threat situations for coffee crops located in the upper Cauca River basin has been developed. For this purpose, a set of services are deployed: Web services of climatic data (*WS\_DClima*), in charge of obtaining all the weather data that are monitored in different coffee production zones; web services for the calculation of the variables that describe the crop condition against the disease, generated from weather data (*WS\_VClimaticas*); and web services for alert management (*WS\_Alertas*), which perform the database management of generated alerts and allow for their search, update, and deletion.

However, these approximations do not consider the alert personalization, considering factors that determine how timely is its diffusion, given the different coffee production zones in the country, which have distinct production cycles, flowering and harvest dates. Thus, two case studies are presented for the personalization of early warnings by adapting web services. The first one takes the coffee crop location as an additional factor for the personalization, and the second one takes the flowering date for the same purpose.

## 2.2 Case 1: Coffee crop location

The first case is based on the generation of early warnings for coffee rust making use of weather data monitored in the area where a crop is located. From this location, a crop can be classified within a coffee zone. Additionally, the supply stores and experts available within the coffee zone are obtained.

In Colombia, there are three coffee zones divided by semesters where the main harvest is presented [30]:

- Zone 1: main harvest in the second half of the year. It includes the departments of Guajira, Magdalena, Cesar, Norte de Santander, Santander, Antioquia, Caldas, Risaralda and Meta (Figure 2).
- Zone 2: main harvest in the first half of the year. It includes the departments of Cauca and Nariño (Figure 2).
- Zone 3: main harvest in the two semesters. It includes the departments of Valle del Cauca, Huila, Tolima, Quindío, Cundinamarca and Boyacá (Figure 2).



Figure 2. Colombian Coffee Zones

Source: Adapted from [30].

### 2.3 Case 2: Coffee flowering date

The second case aims to generate early warnings for coffee rust based on weather data in the area where a crop is located, and its flowering dates are provided by the farmer. Additionally, information about the chemical control products that can be used for the proper management of the disease, as well as recommendations about the experts, are obtained. Flowering can be understood as a phenological state, directly associated

with crop production [31]. Therefore, flowering periods in coffee are directly related to outbreaks of diseases. In the case of coffee rust, if the date of flowering is known, it is possible to identify the moment in which the crop is more prone to the disease, and from this, to establish the ideal dates to start the control process.

According to [30], knowing the quantity and concentration of flowering can establish the period where the application of fungicides is more effective. Figures 3, 4 and 5 show the map of the Colombian coffee zones and the phenological behavior of crops. According to the disease and crop production cycle, the controls for the disease are defined in [30]. Therefore, it is possible to customize the recommendation according to the flowering date.

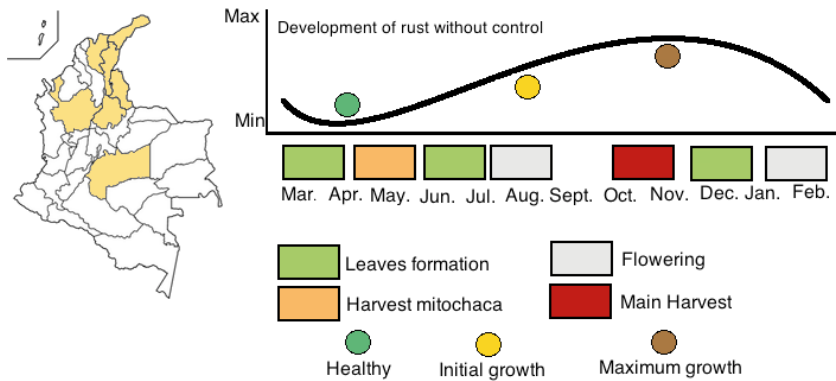


Figure 3. Coffee Zone 1 and Phenology

Source: Adapted from [30].

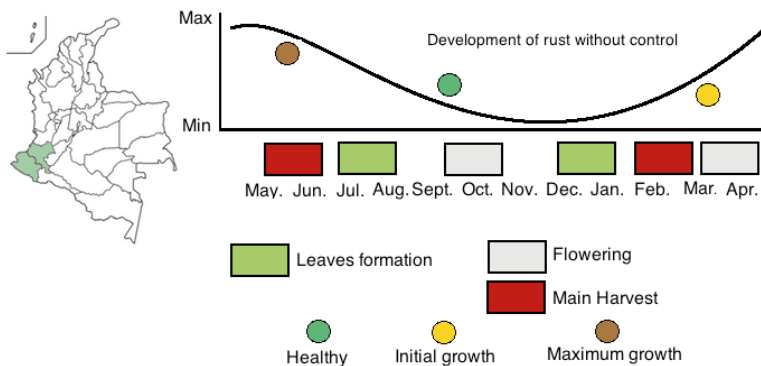


Figure 4. Coffee Zone 2 and Phenology

Source: Adapted from [30].

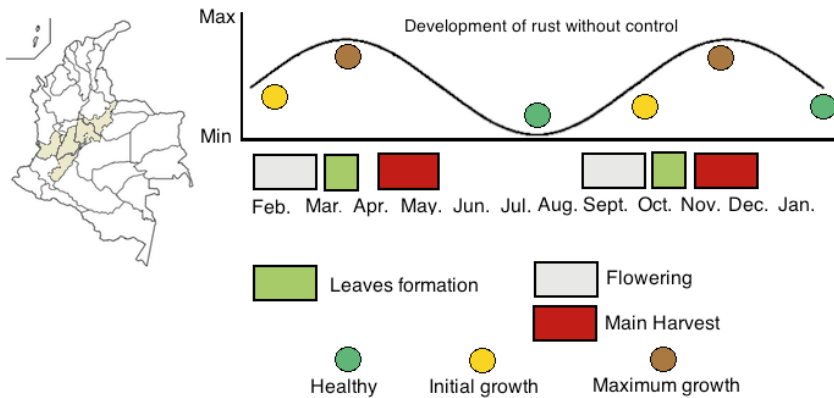


Figure 5. Coffee Zone 3 and Phenology

Source: Adapted from [30].

The adaptation proposal for each case is presented in table 1.

Table 1. Adaptation Summary per Case

Case	Adaptation
Case 1	Adding of the user's location to identify the coffee zone in which he is located.
Case 2	Adding of the flowering date to identify the critical periods of coffee rust onset.

Source: Prepared by the authors.

### 3. AGROCLOUD: WSA

This section addresses the implementation of the proposed solution for the study cases mentioned above within the AgroCloud platform [29], specifying the technology and the architecture used.

#### 3.1 Technology

Our proposal was implemented using the following technologies:

- Enterprise Service Bus (ESB): Mule ESB [32].
- Complex Event Processing Engine (CEP): Esper [33].
- Application server: Glassfish [34].
- Database management server: PostgreSQL [35].

### 3.2 Implementation of the Proposed Solution

The generation of early warnings for coffee rust is addressed from the case studies described above. The services were personalized by adapting them using the scheme proposed in figure 6, which contains the interaction of the elements related in section 2.

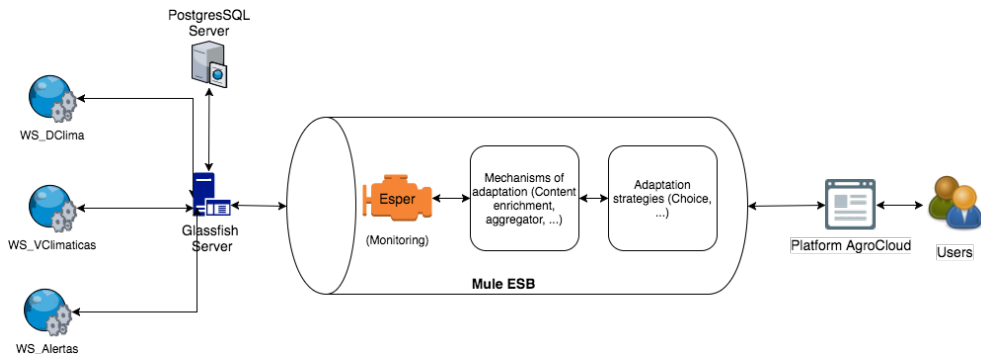


Figure 6. Adaptation of the Proposed ESB-CEP

Source: Prepared by the authors.

The components of the scheme are described below:

- *Web services*: Facilities implemented by AgroCloud to detect the favorable weather conditions for the appearance of coffee rust, and to generate the disease alerts (WS\_DCLima, WS\_VClimaticas, WS\_Alertas).
- *GlassFish server*: Contains all the services developed (weather information, alert management, database access, among others).
- *PostgreSQL server*: Persistence of the data handled by services.
- *Mule ESB*: Through this bus, all messages between services, servers, and applications are transported. Additionally, everything related to adaptation is developed in this element.
- *Esper*: CEP Engine in charge of monitoring the relevant situations presented in the case studies, in order to find all events that require an adaptation, according to the information provided by the users and their needs.
- *Mechanisms of adaptation*: Tasks available within the ESB that allow the adaptation.
- *Adaptation strategies*: Procedures to satisfy the requirements of adaptation according to the current situation.

### 3.2.1 Case Adapted 1

In order to implement the personalization of services for Case 1 (described in section 2.2) the proposed execution flow is the following:

- a) The user makes a query through the platform AgroCloud, and the location of the crop is requested or detected.
- b) Monitoring mechanisms (located on Esper CEP) process the location and recognize it as an event of interest.
- c) The need for adaptation is identified, and the mechanisms available in the ESB (content enrichment, aggregator) are determined.
- e) The most appropriate adaptation strategy to meet the requirement is identified.

### 3.2.2 Case Adapted 2

In order to implement the personalization of services for case 2 (described in section 2.3), the proposed execution flow is described as follows:

- a) The user makes a query through the platform AgroCloud, and introduces the date where the crop flowering was presented.
- b) Monitoring mechanisms (located on Esper CEP) process the date of flowering, and recognize it as an event of interest.
- c) The need for adaptation is identified, and the mechanisms available in the ESB (content enrichment, aggregator) are determined.
- d) The most appropriate adaptation strategy to meet the requirement is identified.

## 4. RESULTS

This section presents the tests performed on the proposed solution. We considered the response time of web services, the use of computational resources, the efficiency in the events of interest detection, and user acceptance testing. Each of the tests was performed considering situations with and without the proposed adaptation.

### 4.1 Use of Computational Resources and Response Time

In order to measure the use of computational resources (CPU and RAM) in the web services execution with and without performing the proposed adaptation, the VisualVM 1.3.9 tool was used. These services were deployed on a server with the following features: 16GB of RAM, Intel Xeon processor at 1.8 GHz, and Ubuntu 12.04 as

operating system. The results are shown in table 2. The use of both CPU and RAM is low and very similar in all cases, indicating that the proposed solution does not generate an overload in computational resources. On the other side, in order to measure the response time of web services with and without performing the proposed adaptation process, the SoapUI tool 5.3.0 was used to invoke each of them. The results obtained, based on 1200 invocations for each of the different web services, are presented in table 2. According to [36], the values obtained are considered as acceptable, since average response times are below 200 milliseconds.

Table 2. Use of Computational Resources

Web Service	Average percentage of CPU usage	Average RAM usage (MB)	Average time (ms)
WS_Ubication	1.58	53.20	9.97
WS_Flowering	1.72	45.28	33.07
WS_TasaRoya	1.55	50.43	8.77
WS_Products	1.40	40.27	7.8
WS_Store	1.50	37.3	7.4
WS_Experts	1.38	36.5	8.0
WS_Adaptation Case 1 (adapted for Case 1)	1.39	51.02	39.17
WS_Adaptation Case 2 (adapted for Case 2)	1.52	44.10	14.66

Source: Prepared by the authors.

## 4.2 Decision Effectiveness

In the test that was performed to the CEP to verify the effectiveness in the detection of interest events, ten cases were considered. First, 100 examples were generated, divided into true (events requiring adaptation) and negative instances (events that do not require adaptation). Subsequently, the number of instances correctly classified by CEP was observed, determining the true positives, false negatives, true negatives, and false positives. For the following cases, the number of examples was increased in 100 successively until obtaining 1000 with a random proportion of true/negative instances in each of them. As a result, the number of false negatives and false positives was zero in all cases, which means that the CEP was able to identify the need for adaptation in the instances for each of the cases tested. This indicates that the process of detecting events of interest carried out by the CEP is effective, and adequate for the proposed solution.

## 4.3 User Acceptance Testing

The user acceptance test (UAT) [37] methodology is intended to analyze the use of the solution by the user, detecting possible errors in the answers, usability, and in the

entry forms, corresponding to the problem that must be solved. From the theoretical basis of UAT, the application of this methodology is presented below:

- *Type of test:* To carry out the UAT methodology, a functional or “black box” test was carried out, in which the users interact with the inputs and outputs of the system, leaving aside the source code and internal flow of the system. This test was developed from meetings in which each user interacted with the system through a computer.
- *Users:* Within the coffee production environment, three types of users were identified, related to the treatment and control of the crop. The UAT methodology was applied to 53 users, described below:
  - *Producers:* Small farmers of medium or large sizes, according to the level of production. 5 medium producers from Cajibío (Cauca), an association of organic coffee producers (Popayán) (12 producers), and an association of coffee producers from Los Andes - Corinto (Cauca) (Asproccan) (20 producers).
  - *Agricultural extension officers (AEO):* Staff with technical knowledge in coffee production (delegates from the Federation of Coffee Growers) to advise producers on different problems surrounding the process. A technical assistance group of the National Federation of Coffee Growers (Colombia) was interviewed (15 users).
  - *Researchers:* They are usually agronomists who study the phenomena that affect coffee production, and how to solve them. An agronomist was interviewed for UAT.
- *Business requirements* provide additional recommendations from specific information entered by the user to the system, in such a way that, based on it, the user can suggest a control to be made from their location, the agricultural input stores within the location, the experts of the area, and the dates in which the control of the disease can be carried out from the date of flowering.
- *Test cases:* As already mentioned in section 3, in this case two test cases were worked on. Based on them, the results are shown below:
  - *Test Case 1*
    - Identification of the coffee zone.
    - Obtain recommendation from the supply stores available from the coffee zone.
    - Obtain recommendations from experts from the coffee area in which it is located.



○ *Test Case 2*

- Identify the critical periods for the beginning of the disease from the date of flowering.
- Obtain a recommendation of the products to be used from the date of flowering.

In addition, four questions were applied to each test case:

- Q1) Was the task completed successfully?  
 Q2) Did the user require assistance to complete the task successfully?  
 Q3) Were the information request forms satisfactorily understood?  
 Q4) Do users consider the tool useful?

The number of users who answered affirmatively is shown in table 3.

Table 3. UAT Results

Test case	Question	Producers (%)		AEO (%)		Researchers (%)	
		Yes	No	Yes	No	Yes	No
1	P1	84	16	100	0	100	0
	P2	6	94	0	100	0	100
	P3	78	22	100	0	100	0
	P4	100	0	100	0	100	0
2	P1	94	6	100	0	100	0
	P2	3	97	0	100	0	100
	P3	100	0	100	0	100	0
	P4	100	0	100	0	100	0

Source: Prepared by the authors.

## 5. CONCLUSIONS

This paper presents an adaptation proposal for web services used in early warning systems for crops diseases through a CEP-ESB integration, based on the adaptation and monitoring framework proposed by the S-Cube project.

The main contribution of our proposal is the use of the mediation capabilities provided by the ESB to carry out the adaptation process, in addition to the use of the CEP as a component in charge of monitoring events of interest that may require such adaptation. The use of computational resources (CPU and RAM) produced by the solution in its execution is among the acceptable ranges. The tests carried out through the

UAT allow us to know the level of satisfaction of the end user with the proposed system, as well as to improve the usability and interpretability of the response provided. The EWS proposed is aimed at the generation of coffee rust that considers local properties of crops. Precisely, as related in [38], to combat this disease, the systematization and dissemination of these type of computational tools should be considered from regional research platforms to producer farm levels.

As future work, it is proposed to address other adaptation requirements, as well as the implementation and testing of other ESB and CEP products. On the other hand, incorrect input data could be solved by handling exceptions in web services.

## ACKNOWLEDGEMENTS

The authors are grateful to the Universidad del Cauca and its Telematics Engineering Group (GIT), the Ministerio de Ciencia, Tecnología e Innovación (Minciencias) and the AgroCloud project of The Interinstitutional Network of Climate Change and Food Security of Colombia (RICCLISA) for supporting this research, and to the *InnovAccion Cauca* project of the Colombian Science, Technology and Innovation Fund (SGR-CTeI) for the PhD. scholarship granted to MsC. Emmanuel Lasso.

## REFERENCES

- [1] G. C. Gómez, “Desarrollos científicos de Cenicafe en la última década”, *Rev. Acad. Colombiana Cienc. Exactas Físicas Nat.*, vol. 1, n.º 30, pp. 89-100, 2005.
- [2] A. de Camargo, and A. R. Pereira, *Agrometeorology of the coffee crop*, Geneva: World Meteorology Organization, 1994.
- [3] C. De León, *Enfermedades del maíz: guía para su identificación en el campo*, México: CIMMYT, 1974.
- [4] F. Gauhl *et al.*, “Multilocal evaluation of black Sigatoka resistance in banana and plantain”, *IITA Res. Guide*, n.º 47, 1993.
- [5] G. N. Agrios, *Fitopatología*, México: Uteha/Noriega.
- [6] A. García, and D. Obín, “Sistemas de Alerta Temprana para Prevención de Enfermedades y Plagas”, 2013. [Online]. Available: <https://dialnet.unirioja.es/servlet/articulo?codigo=4934270&orden=1&info=link>
- [7] A. Wiltshire, “Developing early warning systems: A checklist”, en *Proc. 3rd Int. Conf. Early Warning (EWC)*, 2006.
- [8] J. Ocharan, “Sistemas de Alerta Temprana. Fotografía actual y retos”, *Cuad. Int. Tecnol. Para El Desarro. Hum.*, n.º 6, p. 2, 2007.

- [9] P. Flores N., J. Lerdon F., R. Bravo H., and I. Acuña, “Factibilidad de implementar pronosticadores automatizados para controlar el tizón tardío de la papa en el sur de Chile”, *Agro Sur*, vol. 36, n.º 1, pp. 37-42, 2008. DOI: 10.4206/agrosur.2008.v36n1-06.
- [10] M. Barquero Miranda, “Sistema de alerta temprana para el ojo de gallo”, *Rev. Inf. 2012*, 2012.
- [11] M. L. Gleason *et al.*, “Obtaining weather data for input to crop disease-warning systems: leaf wetness duration as a case study”, *Sci. Agric.*, vol. 65, n.º SPE, pp. 76-87, 2008. DOI: 10.1590/S0103-90162008000700013.
- [12] M. Bonett, “Personalization of Web Services: Opportunities and Challenges”, *Ariadne*, n.º 28, 2001.
- [13] V. García Gutiérrez, “Sistema para la adaptación de servicios a nivel de presentación y de navegación en portales web”, 2013.
- [14] L. González, and R. Ruggia, “Towards dynamic adaptation within an ESB-based service infrastructure layer”, en *Proceedings of the 3rd International workshop on Monitoring, Adaptation and Beyond*, 2010, pp. 40–47. <https://doi.org/10.1145/1929566.1929572>
- [15] L. González, J. L. Laborde, M. Galnares, M. Fenoglio, and R. Ruggia, “An adaptive enterprise service bus infrastructure for service based systems”, en *Service-Oriented Computing–ICSOC 2013 Workshops*, 2013, pp. 480–491. [https://doi.org/10.1007/978-3-319-06859-6\\_42](https://doi.org/10.1007/978-3-319-06859-6_42)
- [16] G. Ortiz, J. Boubeta-Puig, A. García de Prado, and I. Medina-Bulo, “Towards event-driven context-aware web services”, *Adapt. Web Serv. Modul. Reusable Softw. Dev. Tactics Solut.*, pp. 148–159, 2012.
- [17] L. González, and G. Ortiz, “An Event-Driven Integration Platform for Context-Aware Web Services.”, *J UCS*, vol. 20, n.º 8, pp. 1071–1088, 2014. <http://dx.doi.org/10.3217/jucs-020-08-1071>
- [18] L. González, and G. Ortiz, “An ESB-Based Infrastructure for Event-Driven Context-Aware Web Services”, en *Advances in Service-Oriented and Cloud Computing*, Springer, 2013, pp. 360–369. [https://doi.org/10.1007/978-3-642-45364-9\\_29](https://doi.org/10.1007/978-3-642-45364-9_29)
- [19] Service Architecture, “Service-Oriented Architecture (SOA) Definition”. [internet]. Disponible en [http://www.service-architecture.com/articles/web-services/service-oriented\\_architecture\\_soa\\_definition.html](http://www.service-architecture.com/articles/web-services/service-oriented_architecture_soa_definition.html).
- [20] S.-T. Yuan, and M.-R. Lu, “An value-centric event driven model and architecture: A case study of adaptive complement of SOA for distributed care service delivery”, *Expert Syst. Appl.*, vol. 36, n.º 2, Part 2, pp. 3671-3694, 2009. DOI: 10.1016/j.eswa.2008.02.024.
- [21] L. González, *Plataforma ESB Adaptativa para Sistemas Basados en Servicios*, Montevideo: Universidad de la República, 2011.
- [22] R. Kazhamiakin, “Adaptation and Monitoring in S-Cube: Global Vision and Roadmap”, en *Workshop on Service Monitoring, Adaptation and Beyond*, 2009, p. 67.

- [23] A. Gaitán *et al.*, “Evento de La Niña en Colombia: recomendaciones para la caficultura”, 2016. *Evento de La Niña en Colombia: Recomendaciones para la caficultura*. Centro Nacional de Investigaciones de Café (Cenicafé).
- [24] D. C. Corrales, A. Ledezma, A. J. Peña, J. Hoyos, A. Figueroa, and J. C. Corrales, “Un nuevo conjunto de datos para la detección de roya en cultivos de café Colombianos basado en clasificadores”, *Sist. Telemática*, vol. 12, n.º 29, pp. 9-23, 2014. DOI: 10.18046/syt.v12i29.1802.
- [25] D. C. Corrales, A. Figueroa, A. Ledezma, and J. C. Corrales, “An Empirical Multi-classifier for Coffee Rust Detection in Colombian Crops”, en *Computational Science and Its Applications – ICCSA 2015: 15th International Conference*, Canadá, 2015, pp. 60-74. DOI: 10.1007/978-3-319-21404-7\_5.
- [26] D. C. Corrales, A. F. Casas, A. Ledezma, and J. C. Corrales, “Two-Level Classifier Ensembles for Coffee Rust Estimation in Colombian Crops”, *Int. J. Agric. Environ. Inf. Syst. IJAEIS*, vol. 7, n.º 3, pp. 41-59, 2016. DOI: 10.4018/IJAEIS.2016070103.
- [27] E. Lasso, T. T. Thamada, C. A. A. Meira, and J. C. Corrales, “Graph Patterns as Representation of Rules Extracted from Decision Trees for Coffee Rust Detection”, en *Metadata and Semantics Research*; E. Garoufallou, R. J. Hartley, y P. Gaitanou, eds., New York: Springer, 2015, pp. 405-414. DOI: 10.1007/978-3-319-24129-6\_35.
- [28] C. A. Meira, L. H. Rodrigues, and S. A. Moraes, “Análise da epidemia da ferrugem do cafeeiro com árvore de decisão”, *Trop. Plant Pathol.*, vol. 33, n.º 2, pp. 114-124, 2008.
- [29] C. A. A. Meira, L. H. A. Rodrigues, and S. A. de Moraes, “Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente”, *Pesqui. Agropecuária Bras.*, vol. 44, pp. 233-242, 2009. DOI: 10.1590/S0100-204X2009000300003.
- [30] M. E. Cintra, C. A. A. Meira, M. C. Monard, H. A. Camargo, and L. H. A. Rodrigues, “The use of fuzzy decision trees for coffee rust warning in Brazilian crops”, en *2011 11th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2011, pp. 1347-1352.
- [31] C. A. Rivillas Osorio, C. A. Serna Giraldo, M. A. Cristancho Ardila, and A. L. Gaitán Bustamante, “La Roya del Cafeto en Colombia Impacto, manejo y costos del control”, 2011. Cenicafé, 2011. 51 p. (Boletín Técnico No. 36).
- [32] V. Ramírez *et al.*, “Variabilidad climática y la floración del café en Colombia”, 2013. Variabilidad climática y la floración del café en Colombia. Centro Nacional de Investigaciones de Café (Cenicafé).
- [33] E. Lasso, O. Valencia, D. C. Corrales, I. D. López, A. Figueroa, and J. C. Corrales, “A Cloud-Based Platform for Decision Making Support in Colombian Agriculture: A Study Case in Coffee Rust”, en *International Conference of ICT for Adapting Agriculture to Climate Change*, 2017, pp. 182-196. DOI: 10.1007/978-3-319-70187-5\_14.
- [34] D. Dossot, J. D’Emic, and V. Romero, *Mule in action*, Greenwich: Manning, 2014.

- [35] EsperTech, “Products”. [internet]. Available: <http://www.espertech.com/products/>.
- [36] A. Goncalves, *Beginning Java EE 6 with GlassFish 3*, Birmingham: Apress, 2010.
- [37] B. Momjian, *PostgreSQL: introduction and concepts*, vol. 192, New York: Addison-Wesley, 2001.
- [38] Z. Liang-Jie, *Web Services Research and Practices*, Hershey: Idea Group Inc., 2008.
- [39] R. Cimperman, *UAT defined: a guide to practical user acceptance testing*, Upper Saddle River: Addison-Wesley Professional, 2006.
- [40] ResearchGate, “PROCAGICA by Jacques Avelino - Research Project on ResearchGate”. [internet]. Available: <https://www.researchgate.net/project/PROCAGICA>.



# Prototype and Method for Crops Analysis in the Visible and Infrared Spectrum from a Multicopter Air Vehicle\*

Julián Andrés Bolaños\*\*  
Liseth Viviana Campo\*\*\*  
Juan Carlos Corrales\*\*\*\*

Received: 02/02/2018 • Accepted: 13/04/2018

<https://doi.org/10.22395/rium.v19n37a14>

## Abstract

Plant health has a direct impact on the quality and quantity of agricultural products. Due to this fact, farmers must monitor crop conditions frequently. However, the current tools for achieving this are complex and inaccessible. Therefore, this article proposes a method for the characterization of crops that allows to monitor the plants using photographs in the visible and infrared spectrum acquired from a multi-rotor air vehicle, using low-cost cameras and free use tools for designing a prototype of processing information. The characterization is performed by identifying the normalized difference vegetation index (NDVI) in the photographic mosaics of the crops. This index provides information about plant health: Consequently, it is calculated and represented on a NDVI map, where the status of a crop is analyzed. The highest values of NDVI represent healthy plants, and the lowest do so for plants with problems, water, or others. The proposed method allows the monitoring of crops in a temporary and spatial form, letting a producer to adopt measures that help the optimization of resources.

**Keywords:** Infrared spectrum; unmanned aerial vehicles; NDVI; plant health.

---

\* Trabajo derivado de la investigación de tesis de pregrado “Caracterización de cultivos agrícolas en el espectro visible e infrarrojo desde un vehículo aéreo multirrotor”, adelantada entre julio de 2015 y agosto de 2017 en la Universidad del Cauca.

\*\* Ingeniero en electrónica y telecomunicaciones, Universidad del Cauca. Miembro del Grupo de investigación en Ingeniería Telemática, FIET Sector Tulcán, Universidad del Cauca. Correo electrónico: julianbolanos@unicauca.edu.co. Orcid: 0000-0002-7562-8877.

\*\*\* Msc. y estudiante de Doctorado en Ingeniería Telemática. Investigador de la Universidad del Cauca, FIET Sector Tulcán. Correo electrónico: liscampo@unicauca.edu.co. Orcid: 0000-0002-0211-9449.

\*\*\*\* Ph. D., docente investigador de la Universidad del Cauca, FIET Sector Tulcán. Correo electrónico: jcorral@unicauca.edu.co. Orcid: 0000-0002-5608-9097.

## Prototipo y método para análisis de cultivos en el espectro visible e infrarrojo desde un vehículo aéreo multirrotor

### **Resumen**

La salud de las plantas tiene un impacto directo en la calidad y cantidad de los productos agrícolas. Debido a esto, los agricultores deben monitorear las condiciones de los cultivos con frecuencia, pero las herramientas actuales para llevar a cabo esta tarea son complejas e inaccesibles. Frente a esta situación, se propone en este artículo un método para la caracterización de cultivos que permita un monitoreo de las plantas a través fotografías en el espectro visible e infrarrojo adquiridas desde un vehículo aéreo multirrotor, mediante cámaras de bajo costo y herramientas de uso libre para el diseño de un prototipo de procesamiento de información. La caracterización se realizó mediante la identificación del índice de vegetación de diferencia normalizado (NDVI) en los mosaicos fotográficos de los cultivos. Este índice es capaz de proveer información acerca de la salud de las plantas, por lo cual se calcula y representa en un mapa NDVI en el que se analiza el estado del cultivo. Los valores más altos de NDVI representan a las plantas saludables, y los más bajos a las plantas con problemas, al agua u otros. El método propuesto permite monitorear cultivos de forma temporal y espacial, con lo cual se llevaría al productor a tomar medidas que permitan la optimización de recursos.

Palabras clave: espectro infrarrojo; vehículos aéreos no tripulados; NDVI; salud de las plantas.

## Protótipo e método para análise de cultivos no espectro visível e infravermelho a partir de um veículo aéreo multirrotor

### **Resumo**

A saúde das plantas tem um impacto direto na qualidade e quantidade dos produtos agrícolas. Devido a isso, os agricultores devem monitorar as condições dos cultivos com frequência. Diante dessa situação, propõe-se neste artigo um método para a caracterização de cultivos que permita um monitoramento das plantas por meio de fotografias no espectro visível e infravermelho adquiridas a partir de um veículo aéreo multirrotor, mediante câmeras de baixo custo e ferramentas de uso livre para o desenho de um protótipo de processamento de informação. A caracterização realizou-se mediante a identificação do índice de vegetação de diferença normalizada (NDVI) nos mosaicos fotográficos dos cultivos. Esse índice é capaz de prover informação a respeito da saúde das plantas, pelo qual se calcula e representa num mapa NDVI no qual se analisa o estado do cultivo. Os valores mais altos de NDVI representam as plantas saudáveis, e os mais baixos as plantas com problemas, a água ou outros. O método proposto permite monitorar cultivos de forma temporária e espacial, com o qual se levaria o produtor a tomar medidas que permitam a otimização de recursos.

**Palavras-chave:** espectro infravermelho; veículos aéreos não tripulados; NDVI; saúde das plantas.



## INTRODUCTION

Colombia is a rich country in raw materials, and enjoys tropical conditions to carry out agricultural practices: These conditions allow that all kinds of products grow. According to this context, studies are executed from engineering to achieve solutions that improve the production and quality of the products, which are aimed towards precision agriculture (PA), itself based on the development of excellent agricultural practices to optimize farming techniques. Precision farming has evolved over the last 10 years up to implement technological advances such as geographic information systems (GIS), global positioning systems (GPS), satellite or terrestrial sensors, software tools, and data processing, among others; these methods help the farmer to know crop conditions in detail, and to carry out actions of containment, prevention and sustenance [1] [2]. One important phase in precision agriculture is information processing, which is the basis for crop characterization, and it is basically the knowledge of their physical and physiological properties.

Among the relevant methods for crop characterization is aerial photogrammetry, which includes georeferenced mosaics, digital terrain and surface models, and vegetation indexes that correspond to mathematical combinations between the different spectral levels found in a photography [1]. Generally, photogrammetric analysis is performed on satellite images [2]; however, this technology involves high implementation costs and low resolution. This difficulty generates interest in evaluating other alternatives such as the use of unmanned aerial vehicles (UAV). A UAV has the advantage of carrying high-resolution sensors such as lasers, heat sensors, infrared cameras, multispectral cameras, RGB (visible spectrum) cameras and ultraviolet cameras, outstanding among which are cameras of visible and infrared spectrum for their price and availability. Using cameras in the visible and infrared spectrum, fundamental vegetation indexes are calculated such as NVDI (normalized difference vegetation index), which can be useful to determine the vegetative state of the plants [3]. With a correct agricultural crop characterization, it is possible to interpret the data acquired in the practices of precision agriculture, thus optimizing fertilizers application and an effective control in time and space. A correct interpretation translates into economic benefits due to fertilizers savings and a reduction of the treatments performed in the influenced areas.

The direct observation of the plants does not allow to know their state in an objective way. Due to that, it is necessary to use a system capable of recognizing different bands of the electromagnetic spectrum that reveal the crop health, and allow to monitor plant growing. These requirements can be met by performing an agricultural crop characterization using photographs in the visible and infrared spectrum. Although technologies dedicated to this task do exist in the market, their high cost does not allow them to be easily

accessible to farmers. These situations motivate the development of low-cost technology that could give benefits to the farmer, and promote agricultural activities [8] [9].

The present proposal is based on works by Bendig Juliane *et al.* [2], who studied the biomass in an agricultural crop from the vegetation indexes NDVI, SAVI, MSAVI, OSAVI, GnyLi, GRVI, MGRVI, and RGB VI, that are calculated through images in the visible and near infrared spectrum obtained from an octocopter UAV. The objective is to find relations between the physical factors of the crops and the captured images. The work developed in [4] is based on results comparisons taking into account the size of the pixels; it is sought to evaluate the accuracy of the NDVI, GNDIV, NGRDI, and TGI vegetation indexes obtained from images in the visible spectrum, and near infrared.

The method is based on free software, with the disadvantage that increasing the number of images and the size of the crop, increases exponentially the effort and work required to obtain the results. Wei, Wang *et al.* [5] used the vegetation indexes SAVI and RVI to evaluate nitrogen concentration in leaves of wheat and rice plants; the efforts made are based on analyzing the behavior of the different spectral bands. It is concluded that vegetation indexes are appropriate to determine the nitrogen absorption in the leaves. Li, Wang *et al.* [6] desire to find the height of the canopy, as well as the area of biomass within a maize crop using images obtained from a low-cost unmanned aircraft. The location is 84 km north of Beijing, China, 30 m above sea level. A flight was made at a height of 150 meters and at a temperature of 9.1° C, capturing RGB images with 80% longitudinal overlap and a 40% lateral overlap, and vegetation indexes found in the range of the visible spectrum were used. Information on canopy height and biomass were presented in a georeferenced map obtained from vegetation indexes.

The work in [7] describes algorithms for the detection of the plant fraction in a crop from multispectral images captured at different times to measure the spectral reflectance during plant growth, using the vegetation indexes NDVI, GNDVI, and VARI to later relate them with the existing vegetal fraction. The results are presented as graphs of percentage of vegetal fraction related to each one of the indexes of vegetation studied.

In the review, most of the software tools used for image processing, and subsequent obtaining of vegetation indexes, are privative products. It can also be evidenced that most of the projects found are focused on the calculation of the vegetation indexes by multispectral cameras of high cost, therefore, the use of an accessible commercial camera is proposed. Throughout this work, free-use processing tools are accessible to any farmer in the region. All the process is controlled by a software prototype that is composed of a mobile application and a web application, in order to facilitate the collection and processing of information. One of the most important approaches is obtaining a reliable NDVI from a commercial camera through the approximation of the resulting bands and

the use of suppressive filters of visible light allowing to obtain information of infrared radiation, which is useful to determine the vegetative state of plants.

The section 1 of this paper presents a brief introduction regarding vegetation indexes. Later, on section 2, the tools used are presented; sections 3 and 4 describe the processing stage, accompanied by the results, and conclusions are presented in section 5.

## 1. BACKGROUND

For a better interpretation of this paper, the concepts used in this research are described below.

### 1.1 NDVI Calculation

The visible and infrared bands of the electromagnetic spectrum are directly related to different physiological factors on plants, such as the process of photosynthesis, humidity, chlorophyll, etc. [27] [28]. Normally, the spectral reflectance in a plant varies according to the wavelength and its physiological state. Therefore, it is possible to make mathematical combinations between these spectral bands to highlight the vegetation of its surroundings. The result of this process is called Index of Vegetation, and it becomes useful to have an idea of the physiological processes of plants [3]. According to [10] [11], the NDVI is suitable to carry out a process of crops monitoring in order to determine the quality of the growth of the plants; thus, it is possible to determine affectations by pests or diseases, or if there is a problem of water stress in the plants. In [12] [13] [14], it is found that it is possible to approximate the values of NDVI using a commercial camera by removing the infrared cutting filter and adding a filter that is capable of filtering the red band or the blue band of the visible spectrum, taking advantage of this space to form an image in false color. The relationship involved for NDVI, in this case, is represented in equation (1).

$$NDVI = \frac{NIR - VISIBLE}{NIR + VISIBLE} \quad (1)$$

Where the band corresponding to the NIR represents the one that removes the special filter used, and the visible band corresponds to the remaining band. In particular, NDVI with the red band blocker filter is calculated with equation (2), and the blue band with equation (3).

$$NDVI = \frac{RED - BLUE}{RED + BLUE} \quad (2)$$

$$NDVI = \frac{BLUE - RED}{BLUE + RED} \quad (3)$$

## 2. TECHNOLOGY AND METHOD

The aircraft used was the DJI Phantom 3 Standard quadcopter, which has an approximate operating range of 500 meters and an autonomy of approximately 25 minutes, operated by an application called Litchi from a mobile device. This application allows to manually create missions from the “Waypoints” function, and to deploy them in the aircraft according to the desired configuration. The Raspberry Pi camera has a resolution of 5 Megapixels. It has two band-blocking filters: Roscolux # 19, and Roscolux # 2007. A web and mobile prototype was developed: It was responsible for the loading, processing, and subsequent deployment of results.

This article presents some experiences and considerations for precision agriculture applications using UAV platforms, low cost cameras, and free license processing tools. The experiments are focused on achieving a reliable NDVI from a commercial camera that captures infrared light. In particular, the proposed method focuses on the flight planning and image capturing by a commercial camera carried by an UAV; stitching of obtained images; loading, processing and deployment of results to obtain a reliable NDVI; and analysis and testing over time to monitor crops in a real environment.

### 2.1 Flight Planning and Image Capturing

Flight planning and image capturing are the processes by which photographic information is collected in the field; they are important to obtain quality results. The flight planning considers the prior phases to obtaining information in the field, among which there is the analysis of aspects such as target area, height, possible obstacles, and others. This project considers calculations such as maximum flying speed, distance between flight lines, and distance between captures, in order to optimize resources and improve outcomes [15] [16] [17]. The capture process is carried out by considering the information obtained in the planning process. In this work, the Mission Planner software is used, which will help to establish the routes for autonomous flights. Finally, in the camera module, repetitive captures are implemented with a certain configuration from time to time.

### 2.2 Stitching of obtained images

After the capture step, it is important to process the images obtained. The software used to do this was Image Composite Editor, which is freely used for non-commercial purposes. This process consists of forming a single image from the individual images captured at different points in the field, and preferably at the same height.

### 2.3 Loading, Processing, and Deployment of Results to Obtain a Reliable NDVI

The prototype developed considers the farmer as a business actor, whose business use case is to monitor crops. For the business use case to be fulfilled, different functional requirements were established.

### 2.4 Functional Requirements

According to the above, it is defined that the prototype must allow authentication by user name and password, allow the storage of aircraft and cameras with their respective parameters within the system, allow control of the image capture system, and provide a method for the storage of images to later process them and show the results.

#### 2.4.1 Architecture

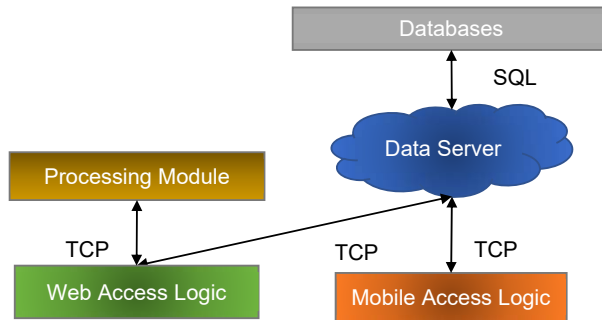


Figure 1. Architecture

Source: Prepared by the authors.

The architecture can be seen in Figure 1, above. It is composed of a database server, a web-mobile prototype, and a processing module. The web-mobile prototype controls the display of results and the upload of information, and the information processing module receives the NGB image and generates the NDVI map.

### 2.5 Implementation Prototype

#### 2.5.1 Mobile Prototype

The mobile prototype was implemented for devices with the Android Operating System, and has the logic capable of receiving and delivering information on the processes carried out by the user. The most representative interfaces can be seen below.

##### 2.5.1.1 Main View

The *main view* shows the process list, and the user can add or delete information capture processes.



Figure 2. Mobile Prototype, Main View

Source: Prepared by the authors.

### 2.5.1.2 New Process View [T5]

This view allows the user to insert new processes to the server. It consists of three steps: Basic information entry, camera choice, and aircraft choice. In step 1, the user types the name of the process, the number of weeks of the test, and the number of tasks (limited between 1 to 3). Step 2 allows to insert new aircrafts and select one, and step 3 allows to insert new cameras and select one.

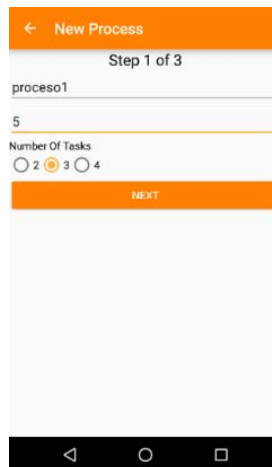


Figure 3. Mobile Prototype, New Process View

Source: Prepared by the authors.

### 2.5.1.3 Process View

This view summarizes the detailed information of the information gathering process. if the user taps on a task, the system shows its current status.



Figure 4. Mobile Prototype, Process View

Source: Prepared by the authors.

### 2.5.2 Web Prototype

The web prototype allows to control photographs loading and processing.

#### 2.5.2.1 Main View [T5]

The main view shows the processes created by the user using the mobile system.

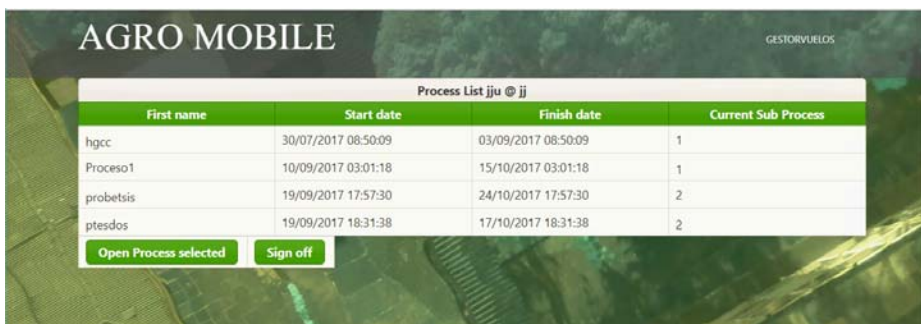


Figure 5. Web Prototype, Main View

Source: Prepared by the authors.

### 2.5.2.2 Process View

This view shows the detail of a single selected process. The server enables the option to upload photographs if the task date is in the allowed range.

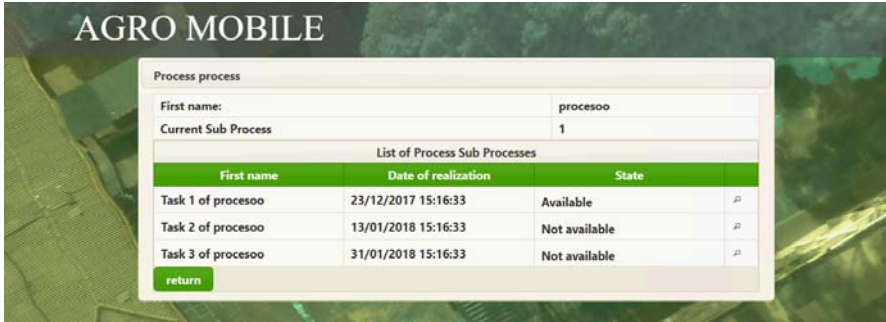


Figure 6. Web Prototype, Process View

Source: Prepared by the authors.



Figure 7. Web Prototype, Task View

Source: Prepared by the authors.

### 2.5.2.3 Load Information View

The information load view allows to establish the information of each task in the process. It receives two photographs as a reference to carry out the information processing. After uploading the photographs, they are automatically processed and referenced.



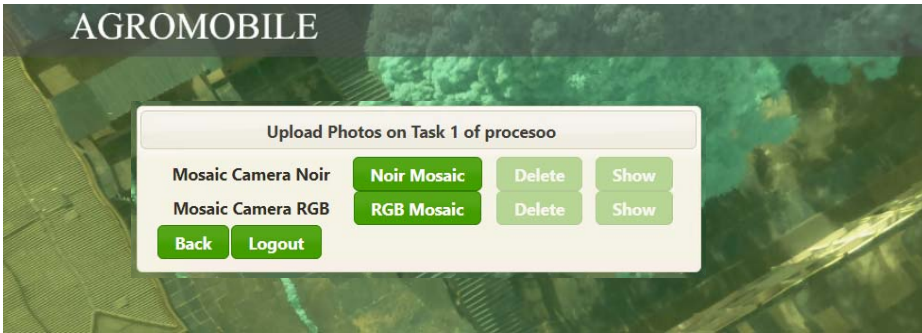


Figure 8. Web Prototype, Load Information View

Source: Prepared by the authors.

### 2.5.3 Processing Information Module-logic

Figure 9 shows the proposed diagram to obtain NDVI values. Python language was used due to its versatility, and all implementations were made using the Linux operating system. The proposed method allows to obtain as output a NDVI map which corresponds to a representation of the NDVI index in graphic form.

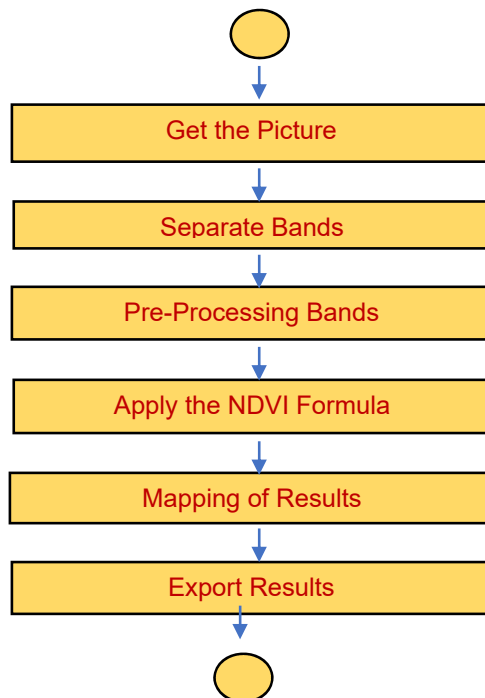


Figure 9. Image Analysis Process

Source: Prepared by the authors.

### 2.5.4 Approximation of Visible Spectrum Bands

In this work, a normal RGB camera identical to the Raspberry Pi NoIR camera is used, with the only difference being that it has the IR filter. The following tests are performed in a controlled environment to compare the bands such as the response of what should look like pure red, pure blue, and pure green. The following tests were performed on coffee plants: On the one hand, a verification and comparison of the resulting bands from the two cameras; and on the other, an approximation of the resulting visible spectrum bands in the two cameras.

#### 2.5.4.1 Verification and Comparison of the Resulting Bands of the two Cameras [T5]

The objective of this test is to establish similarities and differences between the visible band used to calculate the NDVI with the NoIR camera, and the visible band of the RGB camera.

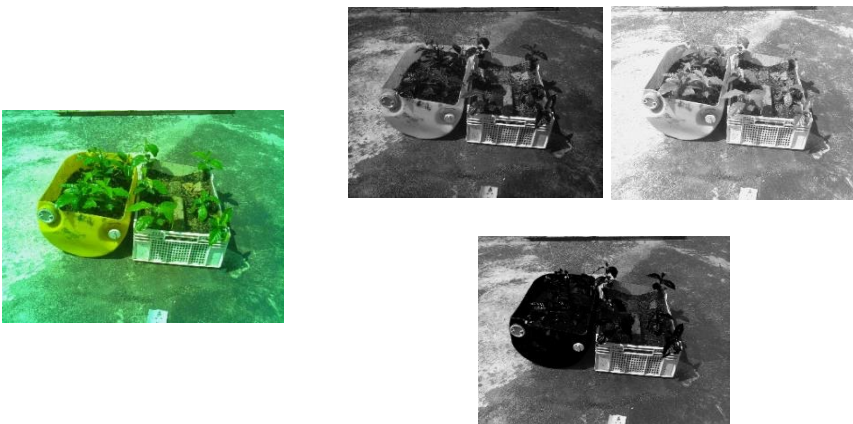


Figure 10. Original Image (RGB)

Source: Prepared by the authors.

Figure 11. Red, Green and Blue RGB image bands

Source: Prepared by the authors.

Figure 11 presents the behavior of the plants around the visible spectrum. It can be seen that in red and blue, these do not have a high reflectance. These results allow having a reference to obtain the bands that correspond to the visible spectrum when using the NoIR camera. The first step is a decomposition of the images obtained with the NoIR camera using both filters, that can be seen below.

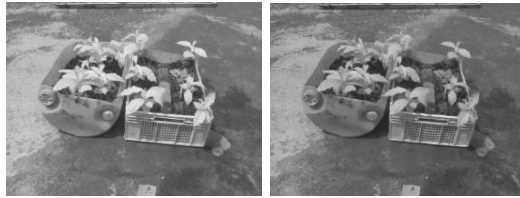


Figure 12. Image NoIR # 2007 Filter      Figure 13. Bands Image NoIR # 2007 Filter (red, green, and blue)

Source: Prepared by the authors.

Source: Prepared by the authors.

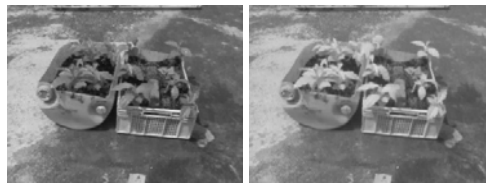


Figure 14. Image NoIR # 19 Filter      Figure 15. Bands Image NoIR # 19 filter (red, green, and blue).

Source: Prepared by the authors.

Source: Prepared by the authors.

Figure 13 and Figure 15 show that the bands used to store the information of the visible spectrum are not similar to the bands originated by the RGB image, which leads to analyze the processing of the bands to bring those results closer.

#### 2.5.4.1 Approximation of the Visible Spectrum Resulting Bands in the Two Cameras

As it can be seen, there are notable differences between the visible bands resulting from the RGB camera, and visible bands resulting from the NoIR camera dependent on the filter to be used. This is due to any used filters these bands are contaminated by infrared information, which causes this difference. Considering this information, a mathematical process is carried out to achieve an approximation between the visible

bands resulting from the NoIR camera and RGB camera. In order to approximate the bands, a mathematical manipulation of the visible bands resulting from the capture with the NoIR camera is made, based on those obtained with the RGB camera under the same conditions.

It is known that the visible band information is affected by infrared information that cannot be filtered, and from this fact it is proposed to remove that information by subtracting the pure infrared band, obtained by using the filter, to the visible band of the same photography. Specifically, the proposed equation is:

$$VISresult = A*(B*VISobtaining - C*INFRAREDObtainin) \quad (4)$$

Where VISresul is *resultant visible band*, VISobtaining is *visible band captured*, INFRAREDObtaining is *Captured infrared band*, and A, B and C are *multiplying decimal constants*.

To determine the values of A, B and C, the resulting visible band must be compared with the visible band obtained from the RGB camera. This comparison process is performed by the linear correlation index that defines the correlation between two random variables—in this case, two images—. The values of this index range from -1 to 1, of which the value of least similarity is 0, which means that the two images have nothing in common. This index can be calculated using the OpenCV functions that are already predefined in the library [18].

In general, the design of an algorithm that allows to modify those values in a determined range until obtaining a value for each one of them, that allows to calculate the maximum of possible correlation and to verify the results, is proposed. The flow diagram that was used is presented in Figure 16.

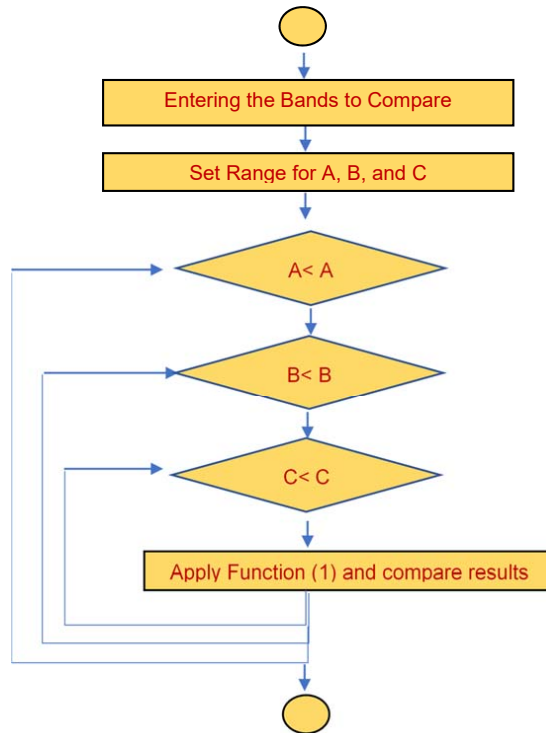


Figure 16. Flow Diagram Band Approach

Source: Prepared by the authors.

The correlation index obtained allows to define values for A, B and C suitable for each of the filter configuration used in the photographs aligned in the test of the combined bands. The process is performed with three different white balance settings.

### 3. RESULTS

#### 3.1 Nomenclature

One of the most important sources of radiation is the sun: The surfaces absorb or reflect this radiation, and that depends on its wavelengths or the characteristics of such surfaces as their composition or roughness [19] [20]. Healthy vegetation, for example, absorbs almost all the wavelengths that make up the visible spectrum; this is because of photosynthesis, but due to the molecular structure of the leaves, it completely reflects the wavelengths in the near-infrared band. Pure water reflects low energy at all wavelengths, different from ice and snow, which have a high reflectance in all bands. In the case of soil, reflected energy depends on the presence of organic matter, texture, and roughness, among others [20] [21].

The interpretation of the results shown is centered in describing the meaning of the colors that compose the resulting maps. According to [22] [23] [24] [25], NDVI is a normalized index, which means that it has infinite values between -1 and 1, in which the vigorous plants with good health reflect all the energy of the sun in the band of the near infrared, they whereas absorb about 90% in the red and blue bands for the process of photosynthesis, reflecting part of the green band, due to the fact that the human eye perceives them in that color. That leads to have a high NDVI, greater than 0.5, whereas plants with some type of anomaly or problem reflect energy equally in all the bands described, the visible spectrum being comparable with the near infrared; this means that they have a lower NDVI, between 0.1 and 0.3 [24] [25]. For this reason, in Figure 17 the continuous NDVI map is established where the NDVI values are mapped, with the blue color on the left end representing -1 and the red color on the right end representing 1.



Figure 17. Colormap

Source: Prepared by the authors.

Figure 18 shows the classification that was made to achieve an interpretation of the results shown, according to the defined nomenclature.

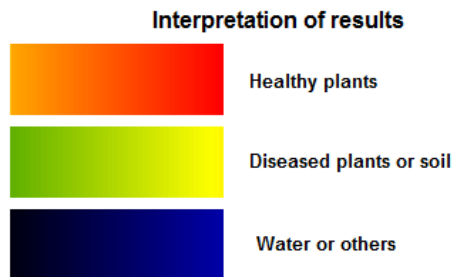


Figure 18. Interpretation of Results

Source: Prepared by the authors.

### 3.2 Captures

In order to test the proposed method, captures were made in a coffee crop in the municipality of Timbío, Department of Cauca, at 1,782 meters above sea level [26], delimiting an area of 1 hectare of flat terrain without obstacles higher than 50 meters. The objective of these studies is to demonstrate the capacity of the proposed method for calculating the NDVI that allows to carry out a monitoring of crops in a time lapse of three weeks.

The testing process begins by collecting the images in the crop, considering that to make the samples comparable, it is necessary to keep the initial configuration as much as possible. After collecting the images, the stitching process of pictures is executed, and the tool to be used is Image Composite Editor. The mosaics shown in the Figure. 19 were made using the process described in [27]. In Figure. 19, Figure. 20, and Figure. 21, three configurations established for the camera are presented using the visible band eliminator filters, whose characteristic curve can be found in [28]. In total, three captures were made, with one week of time difference between them.

### 3.2.1 Filter # 19 AWB 0.54,1

This configuration allows to capture images for the NDVI calculation using the filter # 2007.

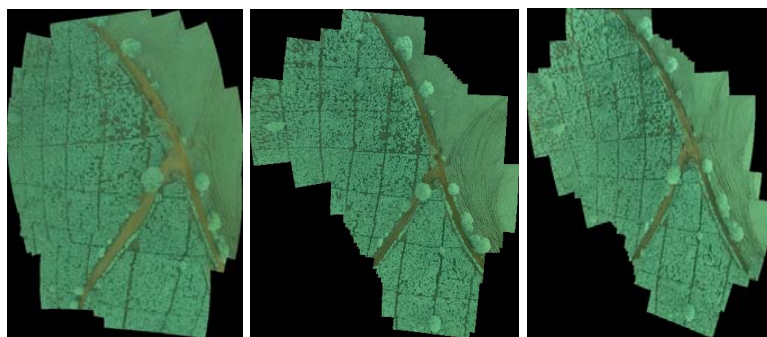


Figure 19. Mosaics with # 19 Filter and AWB 0.54,1 Configuration

Source: Prepared by the authors.

### 3.2.2 Filter # 19 AWB 1,1

This configuration is an alternative to using filter # 19, with which NDVI calculations are also performed. The process to be executed was the same as in the previous configuration.

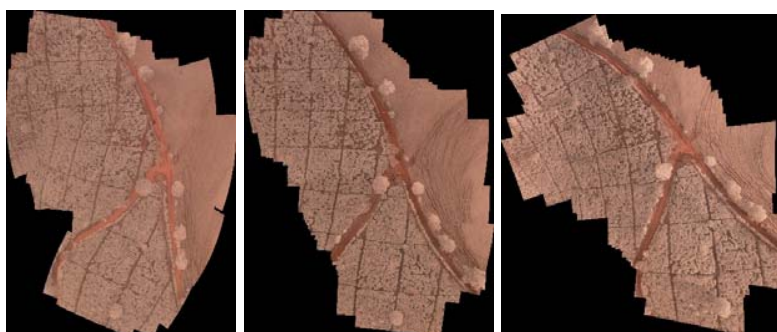


Figure 20. Mosaics with #19 Filter and AWB 1,1 Configuration

Source: Prepared by the authors.



### 3.2.3 Filter # 2007 AWB 1,1

This setting is for NDVI calculation using filter # 2007.

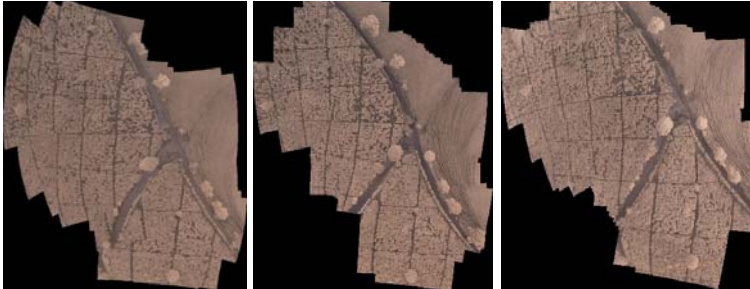


Figure 21. Mosaics #2007 Filter and AWB 1,1 Configuration

Source: Prepared by the authors.

#### 3.2.3.1 NDVI with #19 filter AWB 1,1 configuration

In Figure 22, the NDVI generated in the three corresponding tests are presented with the configuration of the filter # 19 AWB 1,1. These tests correspond to the results of each of the images captured in each of the tasks when the process is finished, and they are generated automatically.

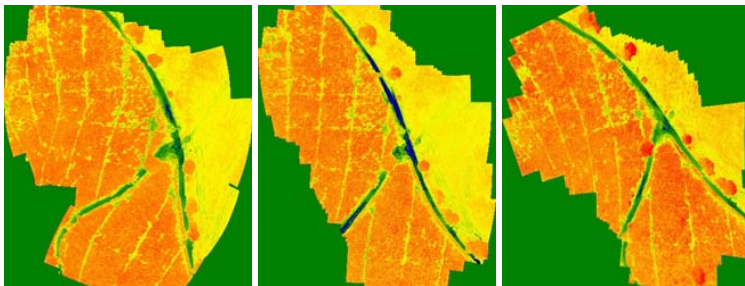


Figure 22. NDVI with #19 Filter and AWB 1,1 Configuration

Source: Prepared by the authors.

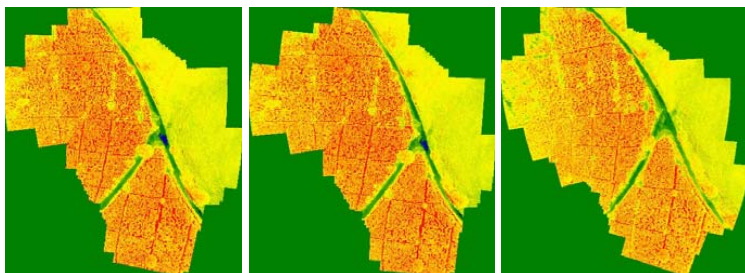


Figure 23. NDVI with #19 Filter and AWB 0.54,1 Configuration

Source: Prepared by the authors.



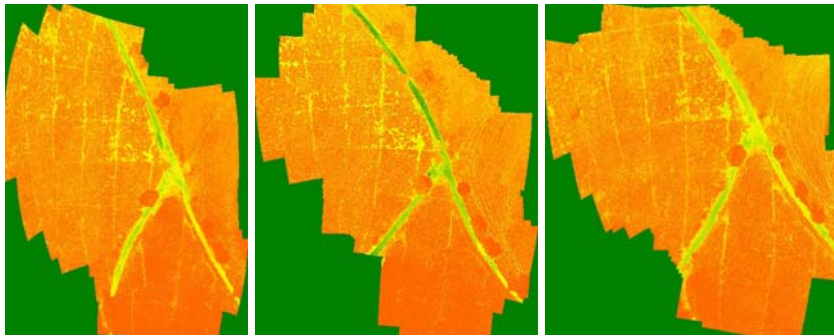


Figure 24. NDVI with #2007 Filter and AWB 1,1 Configuration

Source: Prepared by the authors.

Table 1. Red Filter 0541

Red Filter 0541	Week 1		Week 2		Week 3	
	Total	%	Total	%	Total	%
Healthy plants	12435271	51.09	16557882	30.51	20720728	35.13
Diseased plants or soil	2412707	9.91	1840675	3.39	2536176	4.30
Water or others	9490265	38.99	35866248	66.09	35728564	60.57
Total	24338243	100.00	54264805	100.0	58985468	100.00

Source: Prepared by the authors.

Table 2. Red Filter 11

Red Filter 11	Week 1		Week 2		Week 3	
	Total	%	Total	%	Total	%
Healthy plants	35518790	50.08	35185944	47.69	44526570	47.68
Diseased plants or soil	13615191	19.20	14133325	19.16	9420010	10.09
Water or others	21791232	30.72	24457095	33.15	39439314	42.23
Total	70925213	100.00	73776364	100.00	93385894	100.00

Source: Prepared by the authors.

Table 3. Blue Filter 11

Blue Filter 11	Week 1		Week 2		Week 3	
	Total	%	Total	%	Total	%
Healthy plants	27981069	54.53	27460136	46.81	26010323	53.69
Diseased plants or soil	1228673	2.39	2266056	3.86	1311104	2.71
Water or others	22101541	43.07	28942734	49.33	21126269	43.61
Total	51311283	100.00	58668926	100.00	48447696	100.00

Source: Prepared by the authors.

From the results presented in figures 22, 23 and 24, and in the tables 1, 2 and 3, it can be said that although there are differences between them, all three keep the above characteristics. However, the one that best represents the expected information is the one in Figure 22: With this configuration, a suitable representation is obtained for the values corresponding to the soil and the vegetation, achieving a good contrast between the different NDVI values. It is important to emphasize the temporal variation between the three consecutive captures at different times. Although in general terms the crop condition remained constant, a slight variation can be observed between them that may belong to changes in the physiology of the plants.

Considering the nomenclature and the analysis above, it can be affirmed that in the three defined configurations of the camera there is a relation between what is obtained and what is expected according to the literature. This mainly refers to the fact that for vegetation there are high values of NDVI that approach 1, and they are classified as healthy or vigorous plants. Furthermore, it can be appreciated that for the areas that do not correspond to the vegetation, there are low NDVI values of around 0.3 or even negative ones. As the tests were performed with one week of difference in each and the environmental conditions were constant, a great difference in time between the results is not appreciated. However, it can be seen that Figure 22, which belongs to the capture with a white balance of 1,1, represents the camera configuration that presents a superior performance compared to the others. This configuration maintains a differentiated contrast in comparison to the other captures, thus allowing to establish it as the configuration for future works.

#### 4. CONCLUSIONS

In this article, a method for the characterization of crops using a multicopter air vehicle is proposed. The analysis shows that it is possible to obtain a reliable NDVI to determine and monitor the health of plants in any type of crop, regardless of their age. The proposed method for the monitoring of crops is economically convenient, since the hardware and software processing tools used for processing are free to use, and it is possible to monitor crops over a given time interval using the proposed method for monitoring crops. It is necessary to maintain environmental conditions to obtain valid results. The NDVI map obtaining is based on the infrared information from crops, and it is necessary for the analysis of the evolution of plants individually an experimented farmer as future work. The calculated NDVI demonstrates, among other things, the relationship between plants' health and spectral reflectance: that is why the NDVI is theoretically useful to detect plants' health at any stage of the crop. This paper is a basis for future work focusing on low cost solutions.

## 5. ACKNOWLEDGEMENT

The authors are grateful to the Universidad del Cauca and its Telematics Engineering Group (GIT), the Ministerio de Ciencia, Tecnología e Innovación (Minciencias), and the AgroCloud project of The Interinstitutional Network of Climate Change and Food Security of Colombia (RICCLISA) for supporting this research, and to the “Alternativas Innovadoras de Agricultura Inteligente para sistemas productivos agrícolas del departamento del Cauca soportado en entornos de IoT-ID 4633” research project, funded by “Convocatoria 04C–2018 del Banco de Proyectos Conjuntos UEES-Sostenibilidad” of Project InnovAcción Cauca.

## REFERENCES

- [1] A. Samad, and A. Manan, “Aerial mapping using high resolution digital camera and unmanned aerial vehicle for Geographical Information System”, in *2010 6th International Colloquium on Signal Processing and its Applications (CSPA)*, 2010. DOI: 10.1109/CSPA.2010.5545303.
- [2] J. Bendig, K. Yu, H. Aasen, A. Bolten, S. Bennertz, J. Broscheit, M. Gnyp, and G. Bareth, “Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 39, pp. 79-87, 2015. DOI: 10.1016/j.jag.2015.02.012
- [3] C. Gonzaga Aguilar, “Aplicación de índices de vegetación derivados de imágenes satelitales Landsat 7 ETM+ y ASTER para la caracterización de la cobertura vegetal en la zona centro de la provincia de Loja, Ecuador,” Universidad Nacional de La Plata, 2014. [Online]. Available: [http://sedici.unlp.edu.ar/bitstream/handle/10915/34487/Documento\\_completo.pdf?sequence=1](http://sedici.unlp.edu.ar/bitstream/handle/10915/34487/Documento_completo.pdf?sequence=1).
- [4] E. R. Hunt, C. S. Daughtry, S. B. Mirsky, and W. D. Hively, “Remote sensing with simulated unmanned aircraft imagery for precision agriculture applications,” *IEEE Journal of elected Topics in Applied Earth Observations and Remote Sensing*, pp. 4566-4571, 2014. DOI: 10.1109/JSTARS.2014.2317876.
- [5] W. Wei, Y. Xia, Y.-c. TIAN, X.-j. LIU, N. Jun, W.-x. CAO, and Z. Yan, “Common spectral bands and optimum vegetation indices for monitoring leaf nitrogen accumulation in rice and wheat,” *Journal of Integrative Agriculture*, pp. 2001-2012, 2012. DOI: 10.1016/S2095-3119(12)60457-2.
- [6] W. Li, Z. Niu, H. Chen, D. Li, M. Wu and W. Zhao, “Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system,” *Ecological Indicators*, pp. 637-648, 2016. DOI: 10.1016/j.ecolind.2016.03.036.
- [7] A. Gitelson, Y. Kaufman, R. Stark, and D. Rundquist, “Novel algorithms for remote estimation of vegetation fraction,” *Remote Sensing of Environment*, pp. 76-87, 2002. DOI: 10.1016/S0034-4257(01)00289-9.

- [8] Banco Mundial, “Agricultura y desarrollo-rural,” 2016. [Online]. Available: <http://datos.bancomundial.org/tema/agricultura-y-desarrollo-rural?locations=CO>.
- [9] Banco Mundial, “Indicadores de desarrollo del mundo,” 2015. [Online]. Available: <http://wdi.worldbank.org/table/4.2>.
- [10] ArcGis, “Función NDVI,” 2016. [Online]. Available: <http://desktop.arcgis.com/es/arcmap/10.3/manage-data/raster-and-images/ndvi-function.htm>.
- [11] Universidad de Piura, “Biblioteca Universidad de Piura”. [Online]. Available: [http://www.biblioteca.udep.edu.pe/BibVirUDEP/tesis/pdf/1\\_59\\_185\\_24\\_508.pdf](http://www.biblioteca.udep.edu.pe/BibVirUDEP/tesis/pdf/1_59_185_24_508.pdf).
- [12] V. Dworak, J. Selbeck, K.-H. Dammer, M. Hoffmann, A. A. Zarezadeh, and C. Bobda, “Strategy for the development of a smart NDVI camera system for outdoor plant detection and agricultural embedded systems,” *Multidisciplinary Digital Publishing Institute*, vol. 13, no. 2, pp. 1523-1538, 2013. DOI: 10.3390/s130201523.
- [13] B. José, M. Nicolás, C. Danilo and A. Eduardo, “Multispectral NDVI aerial image system for vegetation analysis by using a consumer camera,” in *Electronics and Computing (ROPEC), 2014 IEEE International Autumn Meeting on Power*, pp. 1-6, 2014. DOI: 10.1109/ROPEC.2014.7036302.
- [14] M. Ghazal, Y. Al Khalil and H. Hajjdiab, “UAV-based remote sensing for vegetation cover estimation using NDVI imagery and level sets method,” in *2015 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 332-337, 2015. DOI: 10.1109/ISSPIT.2015.7394354.
- [15] N. R. Rodrigo Andrés, “Fotografía Aérea,” 2005. [Online]. Available: <http://cybertesis.uach.cl/tesis/uach/2005/bmfcin415f/doc/bmfcin415f.pdf>.
- [16] A. Thomas, “Aerial Photography,” University of the Western Cape. [Online]. Available: [http://www.cdioinstitute.org/papers/Day1/AERIAL%20PHOTOGRAPHY\\_Abraham\\_Thomas.pdf](http://www.cdioinstitute.org/papers/Day1/AERIAL%20PHOTOGRAPHY_Abraham_Thomas.pdf).
- [17] A. Díaz, and M. A. Cisternas, “Fotografías Aéreas Aspectos Técnicos”. [Online]. Available: <https://wikigeomatica.wikispaces.com/file/view/fotosaareas.pdf>.
- [18] OpenCV, “OpenCV Image Matching”. [Online]. Available: [http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template\\_matching/template\\_matching.html](http://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/template_matching/template_matching.html).
- [19] V. Andreo, “Remote Sensing and Geographic Information Systems in Precision Farming,” 2013. [Online]. Available: [http://aulavirtual.ig.conae.gov.ar/moodle/pluginfile.php/513/mod\\_page/content/96/seminario\\_andreo\\_2013.pdf](http://aulavirtual.ig.conae.gov.ar/moodle/pluginfile.php/513/mod_page/content/96/seminario_andreo_2013.pdf).
- [20] Humboldt State University, “Spectral Reflectance”. [Online]. Available: [http://gsp.humboldt.edu/OLM/Courses/GSP\\_216\\_Online/lesson2-1/reflectance.html](http://gsp.humboldt.edu/OLM/Courses/GSP_216_Online/lesson2-1/reflectance.html).
- [21] University of Oldenburg, “Spectral Reflectance Properties”. [Online]. Available: <http://www.seos-project.eu/modules/remotesensing/remotesensing-c01-p05.html>.

- [22] L. C. Velásquez, J. Argueta and K. Mazariegos, "Implementation of a low cost aerial vehicle for crop analysis in emerging countries," *IEEE Journal*, pp. 21-27, 2016. <https://doi.org/10.1109/GHTC.2016.7857255>
- [23] B. Ortiz, J. Shaw, and J. Fulton, "Basics of crop sensing," *Alabama Cooperative Extension System*, pp. 1-3, 2011.
- [24] A. A. Gitelson, Y. Gritz, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves," *Journal of Plant Physiology*, vol. 160, no. 3, pp. 271-282, 2003. DOI: 10.1078/0176-1617-00887.
- [25] D. M. Gates, H. J. Keegan, J. C. Schleter, and V. R. Weidner, "Spectral properties of plants," *Applied Optics*, vol. 4, no. 1, pp. 11-20, 1965. DOI: 10.1364/AO.4.000011.
- [26] Municipio de Timbío, "Información General Timbío". [Online]. Available: [http://timbio-cauca.gov.co/informacion\\_general.shtml](http://timbio-cauca.gov.co/informacion_general.shtml).
- [27] M. Research, "Image Composite Editor," 2015. [Online]. Available: <https://www.microsoft.com/en-us/research/product/computational-photography-applications/image-composite-editor/>.



## Instrucciones a autores

Quienes envíen artículos con destino a publicación en cualquiera de las revistas de la Universidad de Medellín aceptan que con su sometimiento se entiende que el artículo es inédito, de su autoría y que no ha sido propuesto para publicación en ningún otro medio simultáneamente. Además, con la aceptación del artículo se entiende la cesión de sus derechos patrimoniales a la institución y la autorizan a difundir tales artículos por cualquier medio, impreso o electrónico incluido Internet.

**Revista Ingenierías Universidad de Medellín** es una revista científica para la publicación de artículos de alta calidad, resultantes de las investigaciones en el campo de la ingeniería. La revista dará prioridad a las primeras tres tipologías de artículos que se describen a continuación y evaluará la conveniencia de aceptar no derivados de investigación de acuerdo con su relevancia y calidad.

- 1) **Artículo original.** Documento que presenta, de manera detallada, los resultados originales de proyectos terminados de investigación. La estructura generalmente utilizada contiene cinco apartes importantes: introducción, materiales y métodos, resultados, discusión y conclusiones.
- 2) **Artículo de revisión.** Documento resultado de una investigación terminada donde se analizan, sistematizan e integran los resultados de investigaciones publicadas sobre un campo en ciencia o tecnología, con el fin de dar cuenta de los avances y las tendencias de desarrollo. Se caracteriza por presentar una cuidadosa revisión bibliográfica de por lo menos 50 referencias.
- 3) **Artículo derivado de conferencia.** Son resultados de proyectos de investigación terminados o en curso en los que se presentan aportes originales y relevantes en ciencia que requieren pronta divulgación. Se elaboran a partir de ponencias en eventos científicos de calidad académica.
- 4) **Artículo no derivado de investigación.** Documento que presenta puntos de vista originales que surgen de una revisión de un tema, de la reflexión sobre un problema particular, del interés para abrir el debate o del análisis continuado de un problema de interés para la ciencia en el tiempo, para lo cual se recurre a fuentes de calidad científica.

Las personas interesadas en someter a evaluación sus escritos con destino a publicación en la *Revista Ingenierías Universidad de Medellín*, deberán cumplir las condiciones que se enuncian a continuación. Los artículos que no se ajusten a la temática o al estilo de la revista serán devueltos sin pasar a evaluación.

Para efectos de facilitar el proceso de evaluación, se prefiere la recepción de artículos en inglés. No obstante, los artículos pueden ser remitidos en inglés, español o portugués a la revista. Aquellos artículos que sean aprobados para publicación deberán ser entregados por sus autores a la revista en su versión final en inglés de buena calidad para publicación en ese idioma. La no entrega del artículo en versión final en inglés es causal de rechazo del artículo. No se aceptan artículos ya publicados o que se encuentren en proceso de publicación en otra revista.

El artículo debe ser enviado a través de nuestro sistema en línea disponible en <http://revistas.udem.edu.co/index.php/ingenierias/about/submissions#onlineSubmissions>. Para cualquier duda en el manejo de la plataforma, escribir a [revistaingenierias@udem.edu.co](mailto:revistaingenierias@udem.edu.co).

Se presentará el artículo en MS Office Word®, en tamaño carta (215 x 280 mm u 8.5 x 11 pulgadas), a doble espacio, a una columna, con márgenes de 2.5 cm en los cuatro lados, en letra Arial tamaño 12 y con una extensión máxima de 5000 palabras para los artículos derivados de investigación, y de hasta 6000 palabras para los artículos de revisión, excluyendo títulos, resúmenes y palabras clave. No se aceptan artículos de una extensión mayor. Todas las páginas del artículo deben ser numeradas consecutivamente en la esquina superior derecha.

- En el envío del artículo por la plataforma, los autores deben sugerir al menos tres pares evaluadores, dos de ellos de un país diferente del suyo. Los pares evaluadores deben ser académicos de reconocida trayectoria y tener al menos una publicación en revistas indexadas en los dos años anteriores a la fecha de sometimiento del artículo. Para hacer estas sugerencias, los autores deben indicarlo en la sección “Comentarios del autor/a” que encuentra en la plataforma durante el envío. La revista se reserva el derecho de tener en cuenta dichos evaluadores.
- Salvo los artículos de revisión, contienen al menos 50 referencias, el resto de artículos deben tener un mínimo de 15 referencias.
- En los artículos originales y de conferencias, al menos el **30%** de las citas deben ser recientes (últimos 5 años)
- Las referencias relacionadas con monografías o trabajos de grado deben ser máximo el **15%** del total de las referencias.



## SECCIONES

El artículo se debe dividir en la siguiente secuencia de secciones:

1. En la primera página el título centrado. Como el envío es por la plataforma Open Journal System de la revista, se debe garantizar el anonimato en la evaluación a doble ciego. En consecuencia, en el artículo enviado no se mencionan o incluyen a los autores. Sin embargo, en el registro de cada autor en la plataforma, se deben indicar con claridad sus títulos académicos o cargos que indiquen su autoridad en la materia, la dirección física, los teléfonos, la dirección de correo electrónico institucional y el fax de cada uno. No deben usarse direcciones de correo electrónico personales. Es obligatorio que se incluya en el registro en plataforma, el código ORCID de cada autor.
2. En un pie de página asociado al título, debe indicarse si el artículo es de investigación terminada, en curso, de revisión, de conferencia o no derivado de investigación. Debe indicarse el título de la investigación que le da origen, su tiempo de ejecución y los entes financiadores.
3. El resumen tendrá una extensión máxima de 300 palabras, e irá acompañado de 10 palabras clave al final. Las palabras clave deberán ser extraídas de un tesoro para permitir su homologación internacional.
4. En las páginas siguientes irán en su orden la introducción, los materiales y métodos, los resultados, la discusión, los agradecimientos y las referencias. Estos títulos van en mayúscula sostenida. La numeración de las secciones empieza en materiales y métodos con el 1 en arábigos. La introducción y las referencias no se numeran.

## TABLAS, FIGURAS Y ECUACIONES

- Las tablas, ecuaciones y figuras deben estar citadas en el texto; enumérelas en el orden en que aparecen citadas por primera vez; evite redundancia entre tablas, figuras y texto. La rotulación de cada tabla debe ir a la cabeza de la misma; la fuente y la simbología deben aparecer al pie de la tabla.
- Las ecuaciones deben enumerarse en arábigos entre paréntesis al margen extremo derecho de cada una. Se exige que las ecuaciones estén escritas en el editor de ecuaciones de MS Office Word® y no incrustadas como imágenes en el texto. Las ecuaciones deben ser mencionadas en el cuerpo del artículo y se debe evitar “como se muestra en la siguiente ecuación”. Todos y cada uno de los términos de las ecuaciones deben ser descritos y nombrados después de estas.

- Todas las ilustraciones, incluyendo fotos, diagramas, mapas y gráficos, se clasifican como figuras y deben estar a una resolución no inferior a 300dpi. Las figuras deben realizarse en alta calidad y resolución, con líneas y letras de suficiente tamaño, de manera que permanezcan legibles cuando se procesen para la publicación. Las fotografías deben ser de excelente calidad. Se debe poner un número y una leyenda en la parte inferior de cada figura, así como su respectiva fuente y hacer referencia a ella en el lugar apropiado del texto. En caso de que la figura sea de elaboración propia, se debe indicar así: Fuente: elaboración propia. La leyenda de cada figura o tabla tiene que ser lo suficientemente elocuente, para que la figura o tabla se pueda leer independientemente del texto.

Las figuras, en la medida de lo posible, deben estar en un formato editable que permita, por ejemplo, ajustar el tamaño o el tipo de letra.

- No se aceptan tablas incrustadas como imagen. Estas deben estar hechas en MS Office Word® o MS Office Excel®.

## **ABREVIATURAS Y ESTILO**

- No se admite el uso de pies de página en el texto, salvo los necesarios para indicar la filiación institucional de los autores y la procedencia del artículo.
- Se debe usar el sistema métrico y sus abreviaciones para todas las medidas y utilizar decimales en lugar de fracciones (por ejemplo 0,6) excepto en ecuaciones.
- No incluir saltos de página o finales de sección.
- Si se desea resaltar palabras o frases del texto no usar letra negrita sino cursiva.

## **REFERENCIAS**

Cuando se citen referencias en el texto, se deben emplear las pautas del estilo IEEE tal como se utiliza desde el software Mendeley®. Es obligatoria la inclusión de los DOI siempre que estén disponibles para cada referencia. Favor remitirse al uso de ese estándar y consultar cualquier duda con la Revista.

Por tratarse de una publicación con arbitraje, cada trabajo es revisado por un editor designado quien a su vez decide si el artículo se rechaza o si pasa a revisión de pares. Tras la revisión de pares y basado en ese concepto especializado, el editor a cargo solicita ajustes del artículo a los autores. En esta etapa el artículo puede ser rechazado si los autores exceden el tiempo de respuesta indicado por el editor. Una vez que los ajustes

al artículo son satisfactorios, el editor a cargo aprueba el artículo para publicación. El editor general es el que decide el número en el que se incluirá el artículo.

Una causal adicional de rechazo, es si en la etapa de corrección de estilo, posterior a la evaluación y corrección de autores, se encuentran problemas relevantes de sintaxis, coherencia, problemas de originalidad no detectados previamente, redacción u organización que hagan ilegible el texto o dificulten de manera evidente la comunicación que busca el artículo, así como si se detecta que el artículo o gran parte de este ya ha sido publicado por otro medio o en otra revista.

### **CORRESPONDENCIA:**

Los autores deben especificar en su registro en el OJS cuál de ellos es el autor de correspondencia. Esta será la única persona con la cual el Editor o el Comité Editorial tengan contacto para todo lo relacionado con el proceso editorial y la publicación.

**Nota: la revista no cobra honorarios para la publicación de manuscritos.**

