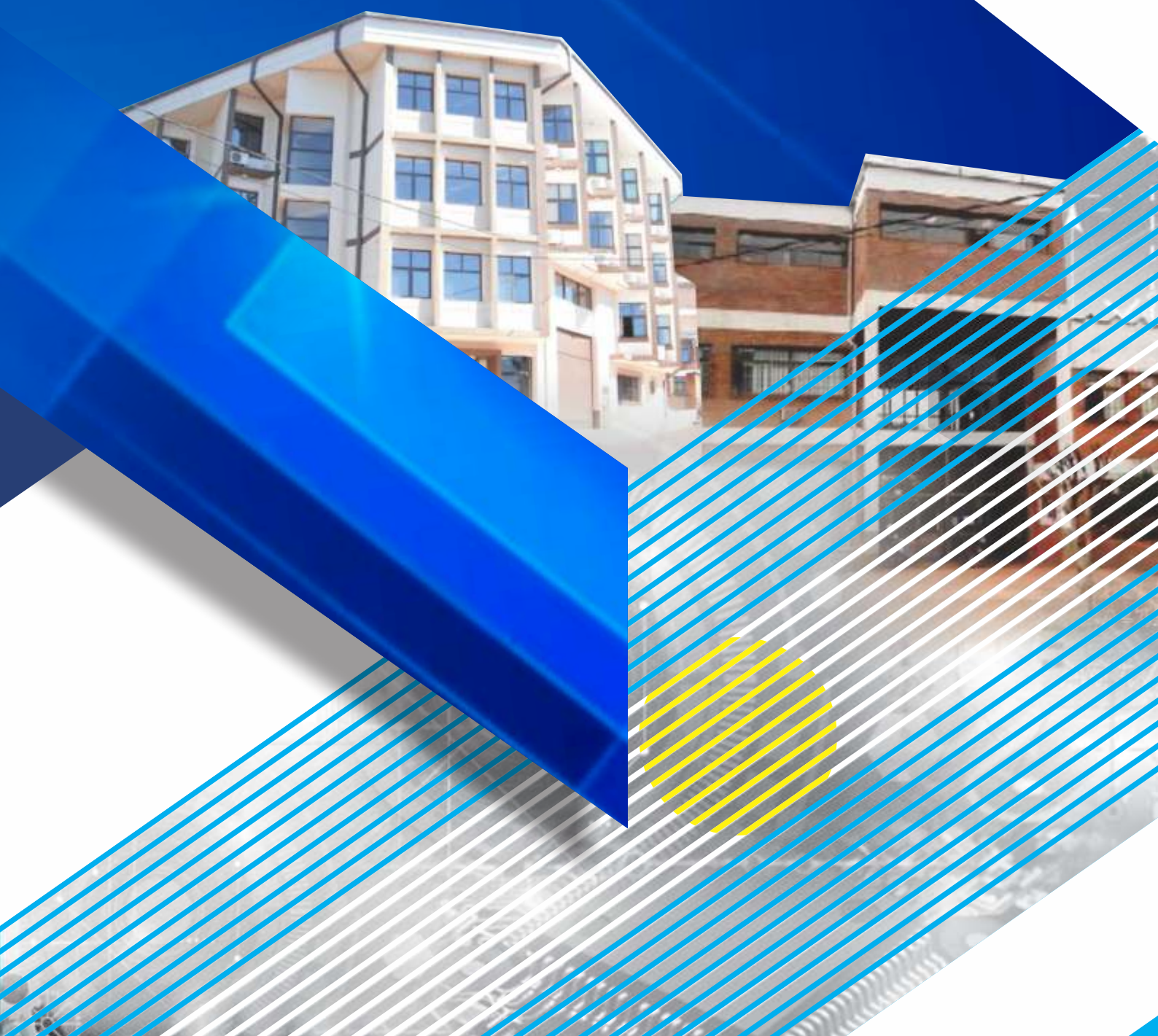


+INGENIO

Revista de Ciencia Tecnología e Innovación

Edición enero-junio 2025 Vol. 6 N° 1

ISSN 2618-5520 on line



■ Universidad Nacional de Misiones / Facultad de Ingeniería / Argentina.

FACULTAD DE INGENIERÍA UNaM

AUTORIDADES

Decana: Mtr. Ing. María Claudia Dekun
Vicedecano: Mgter. Ing. Hugo Orlando Reinert
Secretaria Académica: Arq. Hilda Marta Adelina Cámpora
Secretario de Ciencia y Técnica: Mgter. Ing. Alejandro J. Kerkhof
Secretario de Posgrado: Mgter. Ing. Carlos Rubén Brázzola
Secretario de Extensión: Mgter. Ing. Marco Paolo Bárbaro
Secretaria de Asuntos Estudiantiles: Ing. Florencia Rocío E. Lima
Secretaria Administrativa: L.G.U. y D.G. Silvina Luciana Barrios

EQUIPO EDITORIAL

DIRECCIÓN EDITORIAL

Editor en Jefe

Dr. Ing. Rubén Orlando Núñez

Co-Editores

PostDr. Ing. Mario Orlando Oliveira
PostDr. Ing. Mario José Mantulak
PostDr. Ing. Gustavo O. Bogado
PostDr. Ing. Juan Esteban Miño Valdés

DIRECCIÓN EDITORIAL

Dr. Ing. Arturo Suman Bretas. Universidad de Florida. Estados Unidos
Dr. Ing. José Renes Pinheiro. Universidad Federal de Santa María. Brasil
Dr. Ing. Daniel Gustavo Allasia Piccilli. Universidad Federal de Santa María. Brasil
Dr. Ing. Alejandro Germán Frank. Universidad Federal de Río Grande del Sur. Brasil
Dr. Ing. Gilberto Dionisio Hernández Pérez. Universidad Central de Las Villas. Cuba
Dr. Ing. Erenio Gonzalez Suarez. Universidad Central de las Villas. Cuba
Dr. Ing. Amaury Pérez Martínez. Universidad Estatal Amazónica. Ecuador.
Dr. Ing. Juan Carlos Michalus. Universidad Nacional de Misiones. Argentina
Mgter. Ing. Victor Andrés Kowalski. Universidad Nacional de Misiones. Argentina
Mgter. Arq. Delia Klees. Universidad Nacional del Nordeste. Argentina
Dr. Ing. Gustavo Aucar. Universidad Nacional del Nordeste. Argentina
Dr. Ing. Jorge Víctor Pilar. Universidad Nacional del Nordeste. Argentina
Dr. Ing. Mario De Bórtoli. Universidad Nacional del Nordeste. Argentina
Dra. Ing. Gabriela Corsano. CONICET-UTN, Universidad Nacional del Litoral. Argentina
Dra. Ing. Ada Judith Franco. Universidad Nacional de Salta. Argentina
Dra. Ing. Nora Okulik. Universidad Nacional del Chaco Austral. Argentina
Dr. Ing. Raymundo Quilez Forradellas Martínez. Universidad Nacional de Cuyo. Argentina
Mgter. Ing. Liliana Raquel Cuenca Pletsch. Universidad Tecnológica Nacional. Argentina
Dr. Ing. Oscar Möller. Universidad Nacional de Rosario. Argentina
Dr. Ing. Guillermo Rubén Bossio. Universidad Nacional de Río Cuarto. Argentina
Mgter. Ing. Raúl Barral. Universidad Nacional de Río Cuarto. Argentina
Dr. Ing. Gustavo Ernesto Maffia. Universidad Nacional de la Plata. Argentina
Dr. Alejandro Yawny. Instituto Balseiro (CAB). Argentina

Revisor de estilo para el idioma inglés

Dra. Lic. en Lengua Inglesa Gladys González Carreras. Universidad Nacional de Misiones, Facultad de Ingeniería. Misiones, Argentina.
Mgter. Lic. en Inglés Graciela Noemi Fritz. Universidad Nacional de Misiones, Facultad de Ingeniería. Misiones, Argentina.

Índice

- 05 Nota editorial**
Mgter. Ing. C. Rubén Brázzola.
- 07 Prólogo**
Esp. Ing. Edgar Adan Pereyra.
- 09 Transformaciones en las Prácticas Profesionales: Desafíos y Adaptación del Ingeniero Civil al Mundo Laboral Actual.**
Marcelo Justo Manuel Gómez, Rosanna Griselda Morán, Claudia Alejandra Pilar, Jorge Víctor Pilar.
- 18 Optimización del Análisis de Imágenes de Cámaras Trampa para Fauna: Modelos de Agrupamiento y Recuperación Basados en IA**
Axel Alfredo Skrauba, Sergio Eduardo Moya.
- 32 Los Pequeños Aprovechamientos Hidroeléctricos y la Necesidad de Especialistas.**
Ariel R. Marchegiani, Orlando A. Audisio, Víctor H. Kurtz.
- 41 Evaluación de la Generación de Residuos Per Cápita en los Barrios de Oberá, Misiones, Argentina.**
Sonia Romina Niezwida, María Fernanda Kaczynski, Mariana Gisel Jansat, Matías Ricardo Stoffel, Sonia Ester Yasinski, Juan Carlos Michalus
- 51 Afundamentos de Tensão em Sistema Elétrico de Distribuição Considerando Curto-circuitos e Geração Distribuída.**
Felipe Hernández García, Mario Orlando Oliveira, Roberto Chouhy Leborgne, Arturo Suman Bretas, Roberto José Cabral.
- 71 Clasificación Multiclase de Estados en Colmenas Mediante Bioacústica y Refinamiento de Etiquetas Centrado en Datos.**
Axel Alfredo Skrauba, Hector De Sosa, Sandro Daniel Zakowicz.

COMITE DE REVISIÓN

Mgter. Ing. Gisela Marina Alvarez
Dra. Lic. Nancy Beatriz Ganz
Esp. Ing. Alejandra María Esteban
Dra. Ing. Mauren Pomalis C. da Silva
Dr. Ing. Miguel Aguirre
Dr. Ing. Jonathan Emmanuel Ochoa Sosa
Dr. Ing. Diomar Adonis Coppetti Lima
Dr. Ing. Diego Alberto García
Mgter. Ing. Francisco Gentile
Dr. Ing. Aldo Javier Benítez
Dr. Ing. Roberto Esteban Carballo
Dr. Ing. Guilherme Phena

EQUIPO TÉCNICO

Soporte Técnico:

Lic. Miryan Rosana Puchini, Universidad Nacional de Misiones - Facultad de Ingeniería. Misiones, Argentina.

COLABORADORES

Diseño de Tapa y Formato de Publicación:

D.G. Damián Leandro Kallus, Universidad Nacional de Misiones - Facultad de Arte y Diseño. Misiones, Argentina.

Dirección de Tecnologías de la Información y Comunicaciones

A.S.C. Evaldo Meinl, Universidad Nacional de Misiones - Facultad de Ingeniería. Misiones, Argentina.
Lic. Lucas Podkowa, Universidad Nacional de Misiones - Facultad de Ingeniería. Misiones, Argentina.

EDITORIAL RESPONSABLE

Universidad Nacional de Misiones (UNaM), Facultad de Ingeniería (FI). C.P.: 3360 Juan Manuel de Rosas 325, Oberá, Misiones, Argentina. Es una publicación de la Secretaría de Ciencia y Técnica y la Secretaría de Posgrado de la FI-UNaM.



UNIVERSIDAD
NACIONAL
DE MISIONES



FACULTAD
DE INGENIERÍA
UNaM

Nota Editorial

La internacionalización de las publicaciones científicas: entre la excelencia y la equidad del conocimiento.

La ciencia, en su esencia más profunda, ha sido siempre un proyecto colectivo y universal. Desde los primeros congresos internacionales del siglo XIX hasta las actuales redes globales de investigación digital, el conocimiento científico se ha construido sobre la colaboración, el intercambio y la validación entre pares sin fronteras. Sin embargo, en las últimas décadas, la internacionalización de las publicaciones científicas ha adquirido un sentido mucho más complejo: no solo alude al diálogo global entre investigadores, sino también a los procesos institucionales, lingüísticos y económicos que determinan quién puede participar y ser visible en ese diálogo.

Por ello, el futuro de la internacionalización debería centrarse en modelos colaborativos y solidarios de publicación científica: acuerdos interinstitucionales, redes editoriales compartidas, capacitación conjunta y desarrollo de infraestructuras tecnológicas comunes. La cooperación —y no la competencia— debe ser la base de la excelencia editorial.

Una revista científica que aspire a ser internacional no debería renunciar a su identidad ni a su compromiso con su comunidad de origen. Al contrario, el desafío consiste en construir puentes entre lo local y lo global. La pertinencia temática, la sensibilidad hacia los contextos regionales y la inclusión de voces diversas fortalecen el valor universal del conocimiento científico.

Entendemos que las revistas científicas tienen hoy la responsabilidad no solo de difundir resultados de investigación, sino también de construir comunidad, formar a nuevos autores y revisores, y de actuar como mediadoras entre el conocimiento académico y la sociedad. La internacionalización debe concebirse, entonces, como una dimensión del compromiso social de la ciencia.

Desde esta revista, reafirmamos nuestro compromiso con una internacionalización basada en la calidad, la cooperación y la equidad. Aspiramos a ser un espacio de encuentro entre distintas comunidades científicas, donde el conocimiento circule libremente y donde la diversidad sea reconocida como una fuente de innovación.

Creemos que la internacionalización es, ante todo, un acto de apertura intelectual y de solidaridad global.

Mgter. Ing. C. Rubén Brázzola
Secretario de Posgrado. Facultad de Ingeniería
Universidad Nacional de Misiones

Prólogo

La revista +INGENIO cumpliendo con su compromiso de difusión de actividades realizadas por la comunidad académica y científica, presenta su nuevo número, el cual incluye seis artículos de diferentes áreas de la ingeniería, incorporando nuevamente la novedad de contar no solamente con autores internos de nuestra querida casa de estudios, sino que también investigadores ajenos a nuestra provincia y a nuestro país.

En el primer artículo, **Transformaciones en las Prácticas Profesionales: Desafíos y adaptación del Ingeniero Civil al mundo laboral actual**. Se presenta una reflexión sobre la inserción laboral de los graduados del Ingeniero Civil de la Facultad de Ingeniería de la Universidad Nacional del Nordeste analizando el perfil de estos en el contexto actual y futuro, considerando su capacidad para resolver problemas y adaptarse a cambios en el desarrollo de su profesión.

En el segundo artículo, **Optimización del Análisis de Imágenes de Cámaras Trampa para la Fauna: Modelo de Agrupamiento y Recuperación Basados en IA**, se propone el uso del modelo CLIP-ViT-B-32 de IA, cuando el mismo es utilizado para optimizar el análisis de bancos de imágenes capturados por cámaras trampa, mediante la identificación de especies de la fauna Misionera, discriminando las imágenes obtenidas por disparos provocados por el movimiento de la flora y otros factores ambientales.

El tercer artículo, **Los pequeños Aprovechamiento Hidroeléctricos (PAH) y la necesidad de especialistas**, realiza el estudio de los factores que intervienen en el proyecto y diseño de los PAH para que sean eficaces y fiables al menor costo posible, además, se plantea la importancia de la capacitación de profesionales en este campo para evaluar el potencial hidroeléctrico de una región, diseñar infraestructuras adecuadas y gestionar los recursos hídricos de manera responsable, incluyendo el estudio de impactos ambientales y sociales, garantizando que el desarrollo no afecte negativamente a los ecosistemas locales ni comunidades cercanas.

En el cuarto artículo, **Evaluación de la generación de Residuos Per Cápita en los Barrios de Oberá Misiones Argentina**, con el objetivo de determinar la cantidad de residuos producidos por habitante, por barrio y por día; se realiza un análisis de la generación per cápita del municipio de Oberá mediante un enfoque cuantitativo. Este estudio aporta información clave como punto de partida para la formulación de políticas municipales más efectivas y sostenibles para la gestión de residuos domiciliarios.

En el quinto artículo de este número, **Afundamentos de Tensão em Sistema Elétrico de Distribuição considerando Curto-circuitos e Geração Distribuída**, a partir de simulaciones en un sistema radial de 14 barras, los autores comparan metodologías clásicas de análisis y proponen índices de inmunidad basados en matrices booleanas, complementadas con representaciones gráficas de fácil interpretación. El trabajo aporta herramientas conceptuales y prácticas de alto valor para el análisis de la calidad de la energía y la toma de decisiones en redes de distribución modernas.

Finalmente, en el sexto artículo se presenta el trabajo, **Clasificación Multiclase de Estados en Colmenas Mediante Bioacústica y Refinamiento de Etiquetas Centrado en Datos**. En este trabajo se hace uso de la inteligencia artificial en una aplicación de procesamiento de audio, en esta oportunidad para el procesamiento del ruido emitido por las colmenas. Se propone una metodología de identificación y etiquetado del estado de las colmenas que minimiza la subjetividad humana.

Esp. Ing. Edgar AdanPereyra
Docente.Facultad de Ingeniería
Universidad Nacional de Misiones

Transformaciones en las Prácticas Profesionales: Desafíos y Adaptación del Ingeniero Civil al Mundo Laboral Actual

Marcelo Justo Manuel Gómez ^{1*} , Rosanna Griselda Morán ^{1,2} , Claudia Alejandra Pilar ^{1,2} , Jorge Víctor Pilar ^{1,3} 

¹ Universidad Nacional del Nordeste. Facultad de Ingeniería. Resistencia. Chaco, Argentina.

² Universidad Nacional del Nordeste. Facultad de Arquitectura y Urbanismo. Chaco, Argentina.

³ Universidad Nacional de Misiones. Facultad de Ingeniería. Misiones, Argentina.

e-mails: mgichaco@yahoo.com.ar, moranrosannag@yahoo.com.ar, capilar@yahoo.com.ar, jvpilar@gmail.com

Resumen

Desde una perspectiva histórica, la actividad del Ingeniero Civil se centra en labores en organizaciones públicas como privadas, mayormente en un contexto de empleo en relación de dependencia. En el sector público, integran equipos técnicos jerárquicos o de proyecto, mientras que en el privado son requeridos por empresas constructoras, prestadoras de servicios y grupos consultores para diversas funciones, desde representación y supervisión de obras hasta la integración de equipos multidisciplinarios. El objetivo de este estudio es reflexionar sobre la inserción laboral de los graduados, analizando el perfil profesional del Ingeniero Civil de la FIUNNE en el contexto laboral actual y futuro, destacando su capacidad para resolver problemas y adaptarse a cambios. En el actual contexto social, marcado por la demanda de soluciones sostenibles y avances tecnológicos, se requiere una actualización continua de competencias para enfrentar los desafíos del mercado laboral, el que evoluciona permanentemente. Los futuros graduados deben ser competentes y resilientes ante los cambios, desarrollando flexibilidad, capacidad de aprendizaje continuo y una mentalidad emprendedora para sobresalir en un entorno laboral en constante progreso.

Palabras Clave – Ingeniería Civil, Formación continua, Ejercicio profesional, Competencias universitarias.

Abstract

From a historical perspective, the work of Civil Engineers focuses on roles in both public and private organizations, predominantly within a context of salaried employees. In the public sector, they form part of hierarchical technical teams or project groups, while in the private sector, they are sought after by construction companies, service providers, and consulting groups for various roles, ranging from representation and supervision of construction projects to participating in multidisciplinary teams. The aim of this study is to reflect on the job placement of graduates by analyzing the professional profile of Civil Engineers from FIUNNE in both the current and future labor markets, emphasizing their problem-solving skills and adaptability to change. In today's social context, marked by the demand for sustainable solutions and technological advancements, continuous competency updates are required to meet the evolving challenges of the job market. Future graduates must be competent and resilient to changes, developing flexibility, continuous learning ability, and an entrepreneurial mindset to excel in a constantly advancing work environment.

Keywords – Civil Engineering, Continuing Education, Professional Practice, University Competencies.

1. Introducción

Desde una perspectiva histórica, la actividad principal del Ingeniero Civil ha sido desempeñarse en organizaciones públicas y privadas de diversos tamaños, generalmente bajo algún tipo de relación de dependencia.

Los servicios profesionales de los graduados son mayormente demandados por entidades gubernamentales para formar parte de sus equipos técnicos jerárquicos, trabajar en proyectos o asumir roles de supervisión. En el sector privado, son contratados por empresas de construcción

(asumiendo roles de representación o dirección de obras), en empresas prestadoras de servicios, en Grupos Consultores para la integración de equipos multidisciplinarios de ejecución de estudios y proyectos, asesoramiento y supervisiones de obras.

El graduado de la carrera de Ingeniería Civil de la Facultad de Ingeniería, Universidad Nacional del Nordeste (FIUNNE), se gradúa con un Perfil Profesional con una adecuada formación en las Ciencias Básicas y en las Ciencias de la Ingeniería [1], cabal conocimiento de tecnologías de avanzada en su especialidad, herramientas informáticas además de los procedimientos de cálculo y diseño, adquiriendo una metodología en los procesos de Investigación y Desarrollo Tecnológico. Esto le permite desarrollar metodologías intelectuales reflexivas y críticas, basadas en procesos de análisis-síntesis, razonamiento inductivo-deductivo, pensamiento creativo y espíritu de autocrítica, orientados a la construcción de conceptos, la resolución de problemas, la planificación y la toma de decisiones. De este modo, también adquiere una sólida capacidad para comunicar sus ideas de manera lógica y clara, trabajar de forma colaborativa y, especialmente, adaptarse a nuevas situaciones del ejercicio profesional y a los constantes cambios del campo disciplinar, asumiendo actitudes de compromiso, responsabilidad, iniciativa y participación.

Además, su Perfil de Formación le otorga capacidad para ser competente en: 1) Detectar, estudiar y resolver, interdisciplinariamente, problemas científicos y tecnológicos; 2) Interpretar la relación existente entre la Ciencia, la Técnica y la Naturaleza, con la Estructura Socioeconómica de todo el espectro social, para planificar su desarrollo y bienestar y mejorar la calidad de vida de los habitantes y preservar el equilibrio ecológico; 3) Propender al desarrollo regional y nacional 4) Comprender el mundo natural y social en que se desenvuelve.

En la actualidad, la creciente demanda de soluciones sostenibles, junto con el avance tecnológico que provee nuevas herramientas y acelera los procesos de producción y toma de decisiones, ha generado un impacto significativo en el mercado profesional de la Ingeniería Civil. Este escenario está además influenciado por transformaciones culturales y políticas, con efectos tecnológicos y económicos, en gran medida impulsados por la globalización.

Las transformaciones que se han dado en las últimas décadas en los mercados profesionales y en el ejercicio de la profesión, son resultado de una combinación de factores netamente tecnológicos, económicos, culturales y políticos. Una de las preocupaciones más recurrentes en materia de educación y trabajo es quizá la necesidad de construir un puente entre los espacios de formación y los de inserción laboral [2], para lo cual debemos propender a la formación de profesionales competentes, ya que los mismos se encuentran ante un reto permanente que deben enfrentar como condición sine qua non para la adaptación al nuevo modelo. Para ello, necesitan ir incorporando continuamente competencias y nuevas metodologías de aprendizajes en su formación continua para abordar los nuevos desafíos.

En forma tradicional, las universidades han sido responsables de proveer de forma inicial el conocimiento profesional, ya que durante algún tiempo fueron las únicas en desempeñar un papel de agentes de cambio y de líderes en lo relacionado con fijar los estándares bajo los cuales deberían regirse las profesiones. Sin embargo, en la última década, y más a partir de la puesta en vigor de los Tratados de Libre Comercio, las asociaciones y colegios de profesionales han tomado también esta responsabilidad, al ofrecer por diversos conductos programas de actualización de la formación profesional de sus miembros [3].

En el contexto social actual, aparecen nuevos paradigmas, como la sociedad del conocimiento, la globalización, las redes, y la actual economía que conforman un escenario particular que requiere de

nuevas formas de intercambio y de comunicación [4]. El mundo cambió y sigue cambiando, y la sociedad actual exige más a la Universidad; no sólo exige la formación profesional (el “saber”), sino también, la dotación de competencias profesionales a sus egresados (el “saber hacer”) [5]. No basta solo que sepan hacer, sino además que lo hagan en un contexto de valores éticos y morales.

El antiguo paradigma de formación de profesionales, basado en la enseñanza como simple esquema de transferencia de conocimientos que el futuro profesional oportunamente tendría que saber: abstraer, articular y aplicar eficazmente, ha ido perdiendo espacio en la realidad actual. La visión actual de la sociedad propone ver al graduado universitario como un ser competente (con un conjunto de competencias), capaz de ejercer su profesión en la realidad que lo rodea.

Los graduados de carreras de ingeniería deben tener una adecuada formación general, que les permita adquirir los nuevos conocimientos y herramientas derivados del avance de la ciencia y tecnología. Además, deberán completar y actualizar permanentemente su formación a lo largo de la vida laboral, en el marco informal o en el formal a través del posgrado o de capacitaciones (CONFEDI, 2018). Las transformaciones en las dinámicas laborales actuales requieren el desarrollo de nuevas capacidades, que a menudo resultan difusas y carecen de una codificación clara. [1].

En este sentido, el presente trabajo se propone analizar cómo las transformaciones del contexto profesional impactan en la formación del Ingeniero Civil, indagando en las nuevas competencias requeridas, los desafíos emergentes y la articulación necesaria entre los espacios formativos y el mundo laboral.

2. Desarrollo

2.1. Reflexiones sobre el contexto social actual. Desafíos y respuestas.

Por el avance tecnológico y la transformación digital, este rápido desarrollo de la tecnología ha revolucionado la forma en que nos comunicamos, accedemos a la información y realizamos diversas actividades. Esto ha llevado a un cambio fundamental en la forma en que se enseña y se aprende en las universidades. La educación en línea, el aprendizaje móvil y las herramientas digitales se han convertido en componentes esenciales del proceso educativo. Los estudiantes ahora tienen acceso a una gran cantidad de recursos informáticos, plataformas de aprendizaje interactivas y colaborativas, lo que ha generado nuevos escenarios en la enseñanza y en el aprendizaje y han permitido la producción de conocimientos emergentes en áreas relacionadas con la tecnología y la innovación.

Los cambios generados por la globalización y diversidad cultural, crearon una mayor interconexión entre las naciones y las regiones, llevando a un aumento en la diversidad cultural en las instituciones educativas. Hoy en día los programas de intercambio, enriquecen los debates, las perspectivas y la producción de conocimientos. Esta diversidad cultural estimula la generación de ideas y soluciones innovadoras, fomentando el aprendizaje intercultural y la comprensión global.

Como muy dinámicas, podemos mencionar las nuevas demandas laborales, donde el contexto social actual está experimentando cambios significativos en el mercado laboral. La aparición de profesiones emergentes y la obsolescencia de otras ha llevado a una mayor demanda de habilidades y competencias específicas. Las universidades se ven desafiadas a adaptar sus programas de formación para formar a los estudiantes en áreas emergentes, como la inteligencia artificial, la ciberseguridad, la sostenibilidad ambiental, entre otras. Esto implica una revisión constante de los planes de estudio, la incorporación de nuevas metodologías de enseñanza, la interacción con los

consejos profesionales, las cámaras de empresarios y la colaboración con la industria para garantizar la pertinencia de la formación académica.

No podemos dejar de mencionar que las crisis socioeconómicas y desigualdades sociales tienen un impacto directo en el sistema universitario. La falta de recursos puede limitar el acceso a la educación superior y perpetuar las brechas educativas en mayor escala. Las universidades se ven desafiadas a repensar sus modelos educativos y encontrar formas más inclusivas de brindar oportunidades de aprendizaje con la adopción de enfoques pedagógicos flexibles y la implementación de políticas de equidad.

El contexto social actual, marcado por los avances tecnológicos, la globalización, los cambios en las demandas laborales y las crisis socioeconómicas, genera nuevos escenarios de aprendizaje y conocimientos emergentes en el sistema universitario. Las universidades deben adaptarse a estos cambios, incorporando metodologías innovadoras, fomentando la diversidad cultural, respondiendo a las demandas del mercado laboral y abordando las desigualdades para asegurar una formación de alta calidad y pertinente.

Por tales motivos, es fundamental que las universidades promuevan la creatividad, el pensamiento crítico y la capacidad de adaptación en sus estudiantes, brindándoles herramientas para enfrentar los desafíos del mundo actual. Asimismo, deben fomentar la investigación interdisciplinaria y la colaboración entre diferentes áreas del conocimiento, ya que muchos de los problemas y oportunidades que surgen en la sociedad actual requieren soluciones integrales y multidimensionales. En este contexto, el sistema universitario tiene la responsabilidad de formar profesionales capaces de generar conocimientos emergentes, aportando así al desarrollo social, económico y cultural de las comunidades a las que sirven.

2.2. *Elementos y actores del campo laboral*

Al reflexionar críticamente sobre la configuración del campo profesional de la ingeniería civil y las prácticas laborales de sus graduados, es importante considerar el concepto de campo (sociales) propuesto por Pierre Bourdieu [6]. El campo profesional de la ingeniería civil se puede entender como un espacio social estructurado en el que los ingenieros civiles compiten por recursos y posiciones de poder dentro de la industria de la construcción y el desarrollo de infraestructuras.

Dentro del campo profesional de la ingeniería civil, se pueden identificar varios elementos y agentes que intervienen en su construcción:

- **Instituciones educativas:** Las universidades desempeñan un papel fundamental en la formación de los ingenieros civiles. A través de programas académicos y currículos específicos, estas instituciones transmiten los conocimientos técnicos y las habilidades necesarias para ejercer la profesión. Además, establecen estándares y normas de calidad que influyen en la valoración de los graduados por parte de los empleadores y en su posición en el campo.
- **Asociaciones de profesionales:** Las asociaciones y consejos profesionales, juegan un papel importante en la regulación y la representación de la profesión. Estas organizaciones establecen códigos de ética, promueven el desarrollo profesional y defienden los intereses de sus integrantes. También ofrecen oportunidades laborales y colaboración entre los profesionales del campo.
- **Empleadores y empresas:** Las empresas constructoras, las de consultoría y las entidades gubernamentales son agentes clave en el campo profesional de la ingeniería civil. Estas organizaciones contratan a los graduados y definen las condiciones laborales y prácticas en la

industria. Los empleadores también pueden influir en la especialización de los ingenieros civiles, en función de las necesidades y demandas específicas de sus proyectos.

- **Infraestructuras y proyectos:** Las infraestructuras y los proyectos de construcción representan otro elemento importante en la configuración del campo profesional de la ingeniería civil. Estos proyectos proporcionan oportunidades laborales y establecen el marco en el que los ingenieros civiles aplican sus conocimientos y habilidades. Las características y complejidades de los proyectos pueden generar tensiones y desafíos en cuanto a plazos, presupuestos, sostenibilidad y seguridad.

En el proceso de configuración del campo profesional de la ingeniería civil, también se generan tensiones significativas:

Competencia y jerarquía: La competencia por recursos y posiciones de poder en el campo de la ingeniería civil puede generar jerarquías y desigualdades entre los profesionales. Los ingenieros civiles compiten por proyectos destacados, ascensos y reconocimiento dentro de la industria, lo que puede llevar a tensiones y rivalidades con otras profesiones con la cual comparten actividades reservadas.

Cambios tecnológicos y demandas del mercado: La rápida evolución de la tecnología y las demandas cambiantes del mercado representan desafíos para los ingenieros civiles. La introducción de nuevas herramientas, metodologías y enfoques de diseño requiere una adaptación constante y el desarrollo de nuevas habilidades. Las tensiones surgen cuando los profesionales (los más antiguos) deben mantenerse actualizados, a pesar de su experiencia, y competir en un mercado laboral en constante transformación en donde el joven profesional en ciertas oportunidades lleva ventaja a pesar de la inexperiencia.

2.3. *Innovación curricular para ingenieros del futuro*

Como estrategia en la Facultad de Ingeniería de la UNNE se ha iniciado un cambio curricular en el año 2021, dado que la educación universitaria en general, y en particular la formación de ingenieros, se encuentra atravesando un proceso profundo de revisión con reflexiones profundas acerca de cómo se enseña, cómo se aprende y cómo se evalúa para formar profesionales críticos que puedan enfrentar las necesidades del mundo actual.

La Resolución 1549/21 ME [7] es el mojón que marca el inicio formal de un camino que las unidades académicas venían transitando en distintos tiempos. La necesidad de adecuar los Planes de Estudio y las planificaciones de las diferentes asignaturas a los Estándares de 2ª generación para la Acreditación de Carreras de Ingeniería sugeridos por el CONFEDI (Consejo Federal de Decanos de Ingeniería) es una realidad.

El perfil de egreso que cada Facultad establece, debe brindar claridad sobre las competencias que se esperan del profesional que se gradúa en ellas, entendiendo como competencias al conjunto de saberes (conocimientos, habilidades, actitudes y valores) que les permitirán enfrentar situaciones profesionales estructuradas o de incertidumbre.

Este nuevo paradigma propicia que las asignaturas dejen de ser una enumeración de contenidos para describirse en términos de competencias, es decir qué debe saber hacer un ingeniero al finalizar

los estudios universitarios para ejercer su profesión de manera participativa y comprometida al servicio de la sociedad.

La necesidad de encarar este proceso de cambio curricular fue presentada como consecuencia de la publicación en 2018 del Libro Rojo de CONFEDI (Consejo Federal de Decanos de Ingeniería de la República Argentina) [8], y más recientemente, en mayo de 2021 de la Resolución Ministerial RM 1549/2021 correspondiente a los estándares de acreditación de la carrera de ingeniería civil.

El Libro Rojo de CONFEDI [8] establece una “Propuesta de Estándares de Segunda Generación para la Acreditación de Carreras de Ingeniería en la República Argentina, que derivó en las recientes Resoluciones Ministeriales que aprueban los contenidos curriculares básicos, los criterios de intensidad de la formación práctica, y los estándares para la acreditación de la carrera de Ingeniería Civil entre otras.

Con los cambios vertiginosos en el ámbito tecnológico que se generan por demandas de la sociedad actual, con las nuevas necesidades de profesionales que se requieren en ingeniería civil, electromecánica y mecánica, se hace necesario establecer una conexión estrecha con los requerimientos que el medio laboral solicita, como ser: ingenieros con perfiles que se puedan adaptar con rapidez, trabajando hacia un objetivo común para el desarrollo y el progreso de la sociedad en general y la región en particular.

Hoy es necesario formar ingenieros que alcancen tanto las competencias genéricas, como las específicas de cada carrera; para lograr profesionales integrales que ofrezcan soluciones a las necesidades del medio. Este objetivo es un elemento fundamental a tener en cuenta en la elaboración de nuevos planes de estudios para dicha carrera.

Los ingenieros son entendidos como agentes de transformación de la sociedad y la propia sociedad reconoce en ellos esa contribución. El liderazgo tecnológico debe ser asumido por los ingenieros como capacidad para anticipar los cambios, planearlos, implementarlos, liderarlos, entendiéndose estos cambios no como una revolución puntual y definitiva, sino como una capacidad de innovar permanentemente.

Para garantizar el egreso de ingenieros con perfiles profesionales tan ambiciosos, surge la figura del docente como eje estructural de una ingeniería de excelencia, ya que es el actor cuyo compromiso consiste en preparar estas nuevas generaciones de profesionales capaces de ejercer su profesión al servicio de la sociedad.

Orientar las prácticas docentes hacia el aprendizaje centrado en el estudiante, que promueva la adquisición de competencias tendientes a la formación de personas responsables, con conciencia ética y solidaria, reflexivas y críticas, capaces de actuar en equipos interdisciplinarios, con claro compromiso social preocupados por contribuir a una mejor calidad de vida de la sociedad es, sin dudas, una responsabilidad inmensa.

Los ingenieros exitosos, en un futuro al que algunos llaman VICA (volátil, incierto, complejo y ambiguo) serán aquellos que posean la capacidad para reinventarse asiduamente haciendo una reingeniería de su propia ingeniería.

Si pensamos que los ingenieros deberán realizar una reingeniería de su profesión, para dar respuesta a las demandas de cada tiempo, los docentes (formadores de ingenieros que pretendemos sean competentes), también deberán acompañar y anticiparse a este proceso, haciendo una reingeniería de las prácticas docentes, reflexionando sobre sus propias prácticas, repensando las competencias requeridas para los docentes y su cambio a través del tiempo, vinculado a la formación continua inculcada por la Universidad en la actualidad.

Frente a un futuro que claramente no será una proyección lineal del presente, las prácticas educativas deben acompañar a los jóvenes en una formación centrada en cada estudiante, de manera que los prepare para el manejo de grandes cantidades de datos e información, en la autogestión de técnicas y tecnologías nuevas, propiciando su adaptación a los cambios e incertidumbre manteniendo el equilibrio emocional frente a situaciones nuevas, frecuentemente confusas y desafiantes.

2.4. *Innovación estratégica en la FIUNNE*

La Facultad de Ingeniería de la UNNE entiende que es primordial en cualquier marco formativo de la educación superior, lograr la inserción del graduado en el escenario laboral para el cual se formó y permitirle que se desarrolle plenamente de forma sostenible.

En cuanto a capacitación y formación continua, esta instancia se aplica de manera coordinada con programas de postgrados dependientes de la UNNE y ofertas de capacitación y actualización que brinda la Secretaría de Extensión y Transferencia a partir de la suscripción de acuerdos técnicos específicos con Consejos Profesionales, Organismos Gubernamentales, Cooperativas, ONG, Industrias de la Región, con el objetivo de realizar capacitaciones y articular acciones ligadas a la formación de recursos humanos profesionales y la continua actualización de sus cuadros técnicos.

Además, se implementó por Resolución N° 271/15 CD, un “Mecanismo de Seguimiento de Graduado” a partir del cual la Facultad de Ingeniería ha comenzado a desarrollar distintas acciones tendientes a producir las mejoras que garanticen alcanzar un estándar de calidad. Los objetivos propuestos son, entre otros: desarrollar líneas de acción que involucren metodologías básicas para estudios de seguimiento a los graduados que permitan a la Facultad de Ingeniería obtener información confiable y pertinente sobre la inserción y actividades que desempeñan en el ámbito profesional y determinar el grado de satisfacción y el nivel de exigencia de los empleadores en el desempeño profesional de nuestros graduados. Para llevar adelante estas acciones se creó además por Resol. N°447/16 D, el Área de Graduados.

Las Resolución N° 271/15 CD propende al fortalecimiento del contacto con los graduados estableciendo lo siguiente:

1. Creación de la Asociación de Graduados para promover la cooperación y el intercambio de conocimiento.
2. Organización de eventos que tengan por objetivo el acercamiento del graduado, el fortalecimiento de vínculos, el sentido de pertenecía y la actualización y desarrollo profesional del mismo. Ejemplos: organización de eventos relacionados con el intercambio de experiencias en la “Semana del graduado”.
3. Crear mecanismos de comunicación entre los graduados y la Facultad de Ingeniería: La Facultad promoverá y facilitará la creación de espacios para el encuentro de graduados, utilizando los aportes de las nuevas tecnologías de la información y la comunicación (Sitio Web, Redes sociales, TV, Radio, prensa escrita, otros).
4. Desarrollo de un sitio web destinado a mejorar los canales de comunicación con el graduado.
5. “Semana de las Empresas”: eventos dirigidos a empresas y a organismos públicos y privados de la región (Cámara de la Construcción, PYMES, Empresas Constructoras, otros).

6. Participación de la Asociación a través de representantes en los procesos académicos que tendrán a cargo la revisión de contenidos, metodología y bibliografía desde las visiones de adelanto científico-tecnológicos y las necesidades del sector productivo regional, provincial, nacional e internacional.

7. Fortalecimiento de la bolsa de trabajo implementada.

8. Reconocimiento a profesionales talentosos y emprendedores.

9. Difusión por diversos medios (Redes sociales, sitio web, TV, Radio, Prensa escrita, otros) de las acciones realizadas.

10. Evaluación del Sub Proyecto de Seguimiento al Graduado utilizando como instrumentos de evaluación: encuestas y entrevistas semiestructuradas a empresas pertenecientes a los sectores económicos de la región.

También ante los requerimientos en el marco de acreditación de la carrera, nuestra Facultad puso en funcionamiento a partir del año 2012, el sistema de seguimiento de graduados SIU-KOLLA que permite informar a los alumnos sobre las ofertas de posgrado, aplicar encuestas y otras novedades. Esta herramienta no es suficiente como único medio de comunicación. Se debe contar además con otros recursos apropiados dentro de las TIC'S y otras acciones y estrategias que permitan tanto la realización de las tareas necesarias para tal fin, como también la comunicación eficaz con los graduados.

Hoy en día existen redes sociales específicas para ofertas y búsquedas laborales, donde el graduado va llenando en una línea de tiempo su actividad profesional, mostrando continuamente la experiencia que va adquiriendo. Esta herramienta puede ser de utilidad para mantener un contacto y seguimiento del graduado.

3. Reflexiones finales

El contexto de la globalización y regionalización de los servicios profesionales está condicionando un nuevo modelo de profesionales para el siglo XXI, las cuales tendrán que asumir elementos que respondan a una nueva sociedad que día a día cambia. Estas transformaciones están propiciando que las profesiones incorporen, para su desarrollo, algunos aspectos como la acreditación de planes y programas de estudio de educación superior, la certificación y actualización continua de profesionales, la vinculación de los colegios y asociaciones profesionales con las instituciones de educación superior y flexibilidad para el trabajo, entre otros [3].

Por tal motivo se considera que la Facultad de Ingeniería no permanece ajena a dicho cambios y se haya encaminada a preparar a sus miembros (docentes y estudiantes) para tal reto, en la formación de los diversos campos específicos y profesionales.

Además, podemos decir que la Facultad de Ingeniería de la UNNE entiende que es primordial en cualquier marco formativo de la educación superior, lograr la inserción del graduado en el escenario laboral para el cual se formó y permitirle que se desarrolle plenamente de manera sostenible en su trayectoria profesional.

En resumen, además de ser competentes en su área de especialización, los profesionales del futuro deben ser resilientes ante los cambios. Al adquirir habilidades como la flexibilidad, el aprendizaje continuo, la mentalidad de crecimiento asociado al emprendimiento y la gestión del cambio, se encontrarán preparados para enfrentar los desafíos y aprovechar las oportunidades que se presenten en un entorno laboral en constante evolución.

Referencias

- [1] FIUNNE. (2017). Facultad de Ingeniería - UNNE. “Plan de estudios 2017”.
- [2] Panaia, M. (2006). Educación y Trabajo. Calificación y Formación Profesional. Voces en el Fénix N° 6.
RESOL-2021-1549-APN-ME en: <https://www.boletinoficial.gob.ar/detalleAviso/primera/244506/20210518>
- [3] Fernández Pérez, J. (2001) Elementos que consolidan el concepto profesión. Notas para su reflexión. Revista Electrónica de Investigación Educativa noviembre, vol. 3, número 2 Universidad Autónoma de Baja California Ensenada, México pp. 23-39.
- [4] Ceballos Acasuso Marta, Carlos A. Lovey y Gabriel C. Hruza (2020) “Caracterización de empresas del sector informacional del Chaco y la inserción laboral de graduados de Ingeniería en Sistemas de Información, de la UTN Resistencia”. En Panaia (coord.) ASALARIZACIÓN Y PROFESIONALIZACIÓN: EL DIFÍCIL EQUILIBRIO ENTRE LA AUTONOMÍA Y LA ESTABILIDAD Buenos Aires, Ed. Miño y Dávila. 201-238. ISBN 978-84-18095-49-8. Enlace:
https://drive.google.com/file/d/1E8J1pppnvdyF6NkUX2Py_UsoXniARQk9/view?usp=sharing
- [5] ASIBEI (2016) Competencias Y Perfil Del Ingeniero Iberoamericano, Formación De Profesores y Desarrollo Tecnológico E Innovación.
- [6] Bourdieu, P. (1990): Algunas propiedades de los campos. En Bourdieu, P. Sociología y cultura. México: Grijalbo.
- [7] RESOL-2021-1549-APN-ME En: <https://www.boletinoficial.gob.ar/detalleAviso/primera/244506/20210518>
- [8] CONFEDI (2018). Propuesta de normativa para la regulación de las Carreras de Ingeniería en la República Argentina. “Libro Rojo de CONFEDI”, Consejo Federal de Decanos de Ingeniería. Córdoba: Universidad FAST A Ediciones Disponible en: https://confedi.org.ar/download/documentos_confedi/LIBRO-ROJO-DE-CONFEDI-Estandares-de-Segunda-Generacion-para-Ingenieria-2018-VFPublicada.pdf

Optimización del Análisis de Imágenes de Cámaras Trampa para Fauna: Modelos de Agrupamiento y Recuperación Basados en IA

Axel Alfredo Skrauba^{1,2*} , Sergio Eduardo Moya¹ 

¹ Universidad Nacional de Misiones. Facultad de Ingeniería. Departamento de Ingeniería Electrónica y Computación (DIEC). Misiones, Argentina.

² Universidad Nacional de Misiones-CONICET. Facultad de Ingeniería. Instituto de Materiales de Misiones (IMAM) Grupo de Investigación y Desarrollo en Electrónica (GIDE). Misiones, Argentina.

e-mails: axel.skrauba@fio.unam.edu.ar, sergiomoya@fio.unam.edu.ar

Resumen

Este estudio propone el uso del modelo CLIP-ViT-B-32 para optimizar el análisis de bancos de imágenes masivos capturados por cámaras trampa, con el objetivo de filtrar falsos disparos provocados por el movimiento de la flora y otros factores ambientales. Se utiliza un enfoque multimodal que combina *embeddings* visuales y textuales para lograr una clasificación eficiente y precisa de imágenes, reduciendo significativamente el tiempo y trabajo de investigadores expertos a cargo de la clasificación de la fauna fotografiada.

Palabras Clave – Aprendizaje Profundo, Biodiversidad, Cámaras Trampa, Clustering, Conservación de Fauna, Filtrado de Imágenes, Modelo CLIP-ViT-B-32

Abstract

This study proposes the use of the CLIP-ViT-B-32 model to optimize the analysis of massive image datasets captured by camera traps, aiming to filter out false triggers caused by vegetation movement and other environmental factors. A multimodal approach is employed, combining visual and textual embeddings to achieve efficient and accurate image classification, significantly reducing the time and effort required by expert researchers responsible for classifying the photographed wildlife.

Keywords – Deep Learning, Biodiversity, Camera Traps, Clustering, Wildlife Conservation, Image Filtering, CLIP-ViT-B-32 Model.

1. Introducción

1.1 Contexto del Problema

El monitoreo de la fauna silvestre es fundamental para la conservación de la biodiversidad y el estudio de los ecosistemas a nivel global. Las cámaras trampa son equipos que, mediante un sensor de movimiento y una cámara digital, pueden capturar imágenes de todo aquello que se mueva en su área de influencia. Estos instrumentos han emergido como una herramienta indispensable para la captura automática de imágenes de vida silvestre en su hábitat natural sin intervención humana [1]. Sin embargo, estas cámaras en la gran mayoría de los casos generan un enorme volumen de fotografías, muchas de las cuales son irrelevantes para los objetivos de monitoreo debido a falsos disparos provocados por el movimiento de la flora, cambios de luz y otros factores ambientales [2].

Los falsos disparos no solo aumentan el volumen de datos a procesar, sino que también representan un desafío significativo para los investigadores, quienes deben invertir tiempo considerable en la revisión y filtrado manual de las imágenes [3]. Este proceso laborioso y costoso limita la capacidad de los científicos para enfocarse en el análisis detallado de las imágenes que realmente contienen fauna o el objetivo de estudio específico [4]. Además, la eficiencia del monitoreo se ve comprometida, lo que puede retrasar la toma de decisiones cruciales para la conservación [5].

En la última década, la implementación de técnicas de inteligencia artificial (IA) ha comenzado a transformar el análisis de imágenes de cámaras trampa, facilitando la identificación automática de animales y la filtración de datos irrelevantes [6]. Aún con estos avances, persisten desafíos significativos. Los métodos actuales a menudo requieren grandes volúmenes de datos etiquetados y pueden no generalizar bien a diferentes condiciones ambientales [7]. Además, la alta variabilidad en las condiciones de captura y la presencia de múltiples especies en entornos complejos dificultan la implementación de soluciones robustas y escalables [8].

1.2 Estado del Arte

El uso de modelos de aprendizaje profundo ha demostrado ser efectivos en la clasificación de imágenes de fauna, con redes neuronales convolucionales (CNN) liderando este avance [9]. Modelos como Inception, ResNet y EfficientNet han sido ampliamente utilizados para detectar y clasificar especies en imágenes, logrando altos niveles de precisión en condiciones controladas [10]. Sin embargo, estos modelos requieren una gran cantidad de datos de entrenamiento etiquetados, lo que no siempre es viable en entornos de campo reales en donde no se cuente o se pueda siquiera pensar en disponer de datos correctamente etiquetados, ya sea por cuestiones de costos o simplemente por ausencia de fotografías en caso de especies raras o desconocidas [11].

Una solución emergente en este contexto es el uso de modelos multimodales como CLIP (*Contrastive Language-Image Pretraining*), que combinan el procesamiento de imágenes y texto para mejorar la capacidad de los modelos de IA para entender y clasificar imágenes sin necesidad de una gran cantidad de datos etiquetados [12]. CLIP ha demostrado una capacidad notable para generalizar a nuevas tareas y entornos, superando en muchos casos a los modelos supervisados tradicionales [13]. En particular, la integración de CLIP con arquitecturas de *Transformer* como *Vision Transformer (ViT)* ha permitido una mejora significativa en la extracción de características de imágenes complejas y en la discriminación de patrones visuales [14].

El modelo CLIP-ViT-B-32, en particular, ha emergido como una herramienta prometedora para el análisis de imágenes en contextos de alta variabilidad, como los entornos de cámaras trampa. Este modelo combina la robustez del aprendizaje contrastivo con la capacidad de *ViT* para manejar datos visuales de alta complejidad, proporcionando una solución eficiente y escalable para el análisis automatizado de grandes volúmenes de datos de imágenes.

1.3 Justificación del Artículo

Dado el creciente volumen de datos generados por cámaras trampa y la necesidad urgente de mejorar la eficiencia en las tareas de monitoreo de fauna, este artículo propone explorar la implementación de CLIP-ViT-B-32 como una solución innovadora para el filtrado y análisis de imágenes. La capacidad de este modelo para generalizar a nuevas tareas sin necesidad de datos de entrenamiento extensivos lo posiciona como una herramienta clave para superar las limitaciones actuales en el procesamiento de imágenes de cámaras trampa.

El objetivo de este estudio es doble: en primer lugar, evaluar el desempeño de CLIP-ViT-B-32 en la reducción de falsos positivos y la mejora de la detección de fauna en imágenes capturadas por cámaras trampa; y, en segundo lugar, proporcionar una base para futuras investigaciones que busquen aplicar modelos multimodales en el análisis de datos ecológicos complejos, como ser fotogramas de videos de cámara trampa, detección, clasificación y agrupamiento de eventos sonoros, etc. Este enfoque no solo busca mejorar la eficiencia operativa de los investigadores especialistas en fauna,

sino que también busca contribuir significativamente a la conservación de la biodiversidad mediante una monitorización precisa y oportuna de la fauna.

2. Metodología

2.1 Conjunto de Datos

El conjunto de datos utilizado en este estudio comprende un extenso banco de imágenes capturadas por cámaras trampa en el Parque Federal Campo San Juan, ecorregión de pastizales del sur de la provincia de Misiones. Esta área presenta un desafío particular para el análisis automatizado debido a la alta incidencia de falsos disparos de la cámara causados por el movimiento de los pastos debido al viento. El conjunto de datos consta de aproximadamente 30.000 imágenes recolectadas durante agosto del 2023, abarcando diversas condiciones climáticas y horarios. Algunos ejemplos se observan en la Fig. 1, denotando la diversidad de condiciones y algunas especies capturadas.

Las imágenes fueron capturadas utilizando cuatro cámaras trampa de la marca Bushnell, cuyas imágenes fueron cedidas con fines de investigación por el equipo técnico del área protegida. Cada cámara fue configurada para activarse mediante detección de movimiento y fueron distribuidas estratégicamente para maximizar la cobertura y representatividad del ambiente natural bajo estudio.



Fig. 1. Ejemplos de imágenes de cámaras trampa sin procesar, ilustrando (a) falsos disparos causados por movimiento de pastizales, (b) fauna objetivo capturada con éxito y poca luz, (c) fauna capturada con condiciones ambientales desafiantes como niebla, y (d) falso disparo nocturno.

2.2 Preprocesamiento de Datos

El preprocesamiento de las imágenes se llevó a cabo para optimizar la calidad y consistencia de los datos de entrada para el modelo CLIP. Este proceso constó de las siguientes etapas:

- **Normalización de Tamaño:** Todas las imágenes fueron redimensionadas a una resolución uniforme de 224x224 píxeles, manteniendo la relación de aspecto original mediante relleno con negro cuando fue necesario. Esta estandarización es crucial para la compatibilidad con la arquitectura de entrada del modelo, ya que, de lo contrario, se producirían distorsiones al rescalar con un *aspect ratio* distinto al del modelo.

- **Ajuste de Contraste:** Se aplicó un algoritmo de ecualización de histograma adaptativo con contraste limitado (*CLAHE, Contrast Limited Adaptive Histogram Equalization*) para mejorar el contraste local de las imágenes, especialmente en aquellas capturadas en condiciones de baja iluminación o niebla, ídem para el amanecer o atardecer. En la Fig. 2 puede apreciarse un ejemplo de este proceso.

- **Eliminación de Metadatos:** Se removieron los metadatos incrustados por las cámaras en las imágenes.



Fig. 2. Ejemplo de ajuste de contraste adaptativo, ilustrando (a) imagen original, (b) imagen procesada.

2.3 Extracción de Características con CLIP-ViT-B-32

Para la extracción de características de las imágenes preprocesadas, se empleó el modelo CLIP-ViT-B-32. Se utilizan los pesos pre entrenados, ya que lo que interesa es poder representar al conjunto de imágenes en el espacio latente del modelo, es decir; estructurar la información. Por lo tanto, las imágenes procesadas se introdujeron en el codificador de visión del modelo y se extrajo el *embedding* de la capa final del *transformer*, resultando en un vector de características de 512 dimensiones para cada imagen. Los vectores de características resultantes se normalizaron utilizando la norma L2, para garantizar la consistencia en las comparaciones subsiguientes.

Además, la naturaleza multimodal de CLIP, que permite asociaciones entre texto e imagen, ofrece la ventaja adicional de facilitar consultas basadas en texto para el filtrado posterior de imágenes, una característica que se explorará en etapas subsiguientes del análisis. En pocas palabras, esto significa que puede mapear texto en múltiples idiomas, al mismo espacio latente que las imágenes procesadas.

2.4 Algoritmo de Clustering Basado en Densidad

Para identificar distintas especies de fauna en los datos de imágenes de cámaras trampa, se implementó un algoritmo de *clustering* basado en densidad. Este enfoque es particularmente adecuado para nuestro conjunto de datos, ya que puede descubrir *clusters* de formas arbitrarias y es robusto frente a ruido y valores atípicos, características comunes en imágenes de cámaras trampa.

El algoritmo desarrollado se basa en los principios de DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) [15], pero con modificaciones para optimizar su rendimiento en espacios de alta dimensionalidad y para manejar eficientemente grandes volúmenes de datos. El proceso consta de las siguientes etapas:

- **Cálculo de similitudes:** Para cada par de vectores de *embeddings* (e_i) y (e_j), se calcula la similitud coseno (1). Esta medida se elige por su eficacia en espacios de alta dimensionalidad y su invariancia a la escala.

$$\text{sim}(e_i, e_j) = \frac{e_i * e_j}{|e_i||e_j|} \quad (1)$$

- **Identificación de vecindarios densos:** Se define un umbral de similitud (ϵ) y un número mínimo de puntos (*MinPts*) para considerar un vecindario como denso. Para cada punto (p), se identifica su (ϵ - vecindario($N_\epsilon(p)$)), donde:

$$N_\epsilon(p) = \{q \in D | \text{sim}(p, q) \geq \epsilon\} \quad (2)$$

siendo (D) el conjunto de todos los puntos. Un punto (p) se considera como punto central si:

$$|N_\epsilon(p)| \geq \text{MinPts} \quad (3)$$

- **Formación de *clusters* iniciales:** A partir de cada punto central identificado, se forma un *cluster* inicial (C) mediante un proceso de expansión:
 - Inicializar ($C = p$)
 - Para cada punto ($q \in N_\epsilon(p)$):
 - Si (q) no ha sido visitado, marcarlo como visitado
 - Si (q) es un punto central, añadir ($N_\epsilon(p)$) a la lista de puntos por explorar
 - Si (q) no pertenece a ningún *cluster*, añadirlo a (C)

Este proceso se repite hasta que no queden puntos por explorar o se alcance un tamaño máximo de *cluster* predefinido.

- **Ordenación y filtrado de *clusters*:** Los *clusters* formados se ordenan por tamaño de forma descendente. Para resolver solapamientos y asegurar que cada punto pertenezca a un único *cluster* mayor, se aplica un proceso de filtrado:
 - Iniciar con el *cluster* más grande
 - Para cada *cluster* subsiguiente:
 - Eliminar puntos que ya pertenecen a un *cluster* más grande
 - Si el tamaño resultante es menor que (*MinPts*), descartar el *cluster*

Este algoritmo permite identificar eficientemente grupos de imágenes similares, separando las imágenes de fauna de los falsos disparos y agrupando imágenes de especies similares. Los grupos que se forman no presuponen ningún tipo de distribución (como lo harían algoritmos basados en Kmeans, por dar un ejemplo) y, a su vez, se tienen las potenciales anomalías identificadas.

2.5 Detección de Anomalías o Especies Raras

La detección de anomalías se implementa como un paso complementario al *clustering* para identificar imágenes que podrían representar especies raras o eventos inusuales. El método se basa en

la premisa de que las anomalías son puntos que no se ajustan bien a ninguno de los *clusters* identificados.

El proceso de detección de anomalías consta de los siguientes pasos:

- **Cálculo de distancias a centroides:** Para cada punto (p) no asignado a un *cluster*, se calcula su distancia al centroide más cercano:

$$d(p) = \min_{C_i \in Clusters} (1 - sim(p, centroid(C_i))) \quad (4)$$

donde ($centroid(C_i)$) es el vector promedio de todos los puntos en el *cluster* (C_i)

- **Estimación de umbral de anomalía:** Se utiliza el método de Tukey para definir un umbral de anomalía basado en la distribución de las distancias calculadas:

$$umbral = Q_3 + k(Q_3 - Q_1) \quad (5)$$

donde (Q_1) y (Q_3) son el primer y tercer cuartil de las distancias, respectivamente, y (k) es un factor de escala. Como el método de Tukey se basa en el uso de percentiles y del rango intercuartil, funciona cuando la distribución de las observaciones no necesariamente sigue una distribución Gaussiana.

- **Identificación de anomalías:** Si bien, con el algoritmo de *clustering* ya se discriminan las anomalías y pueden procesarse por separado, no quita que, por ejemplo, se deseen considerar los puntos más alejados de cada *cluster* como tentativos a revisión manual por un personal especializado. Además, según el caso, el número de anomalías puede ser elevado, y es necesario ranquearlas de alguna manera. Por lo tanto, los puntos con una distancia superior al umbral se clasifican como anomalías plausibles de revisión:

$$anomalias = p | d(p) > umbral \quad (6)$$

- **Ranking de anomalías:** Las anomalías identificadas se ordenan según su grado de "rareza", definido por su distancia normalizada al umbral:

$$rareza(p) = \frac{d(p) - umbral}{umbral} \quad (7)$$

Este enfoque permite identificar imágenes que son significativamente diferentes del contenido típico capturado por las cámaras trampa, potencialmente revelando especies raras, comportamientos inusuales o eventos de interés para los investigadores. Además, en caso de múltiples capturas de la misma especie, las mismas quedan agrupadas en su respectivo *cluster*, facilitando el etiquetado manual posterior.

La combinación del algoritmo de *clustering* basado en densidad y el método de detección de anomalías, proporciona una herramienta poderosa para el análisis automatizado de grandes conjuntos de imágenes de cámaras trampa, permitiendo tanto la categorización eficiente de imágenes comunes como la identificación de contenido potencialmente valioso pero infrecuente.

2.6 Mapeo Texto-Imagen para Filtrado de Usuario

La capacidad multimodal del modelo CLIP-ViT-B-32 se aprovechó para implementar un sistema de filtrado basado en texto, permitiendo a los usuarios buscar y clasificar imágenes utilizando descripciones textuales. Este enfoque se basa en la proyección tanto de imágenes como de texto en

un espacio de *embedding* común, facilitando la comparación directa entre consultas textuales y contenido visual. Concretamente, se utilizó CLIP-ViT-B-32-multilingual-v1 para este estudio.

El proceso de mapeo texto-imagen, en resumidas cuentas, consta de computar distancias en el espacio latente, de modo de buscar y ranquear las mejores coincidencias para los vectores estructurados de las imágenes y las consultas textuales. Consiste en los siguientes pasos:

- **Preparación de consultas:** Se define un conjunto de consultas textuales relevantes para el contexto de las cámaras trampa, incluyendo nombres de especies, descripciones de comportamientos y características ambientales (ej., “venado en pastizal”, “zorro corriendo”, “amanecer sin animales”, “ave”, “roedores”, etc.). Esto depende de lo que se desea buscar en el banco de imágenes bajo análisis.
- **Embedding de texto:** Cada consulta textual se procesa a través del codificador de texto de CLIP, generando un vector de *embeddings* de 512 dimensiones.
- **Cálculo de similitud:** Para cada consulta, se calcula la similitud coseno entre su *embedding* y los *embeddings* de todas las imágenes del conjunto de datos.
- **Ranking y filtrado:** Las imágenes se ordenan según su similitud con la consulta, permitiendo a los usuarios acceder rápidamente a las más relevantes o descartar resultados por no encontrar similitudes dentro de un umbral definido.

3. Resultados

3.1 Visualización del Espacio de Embeddings con UMAP

Se utilizó la técnica UMAP (*Uniform Manifold Approximation and Projection*) [16], para reducir la dimensionalidad de los *embeddings* desde las 512 dimensiones generados por el modelo CLIP a 2 dimensiones, permitiendo una visualización en 2D que facilita la interpretación de las relaciones entre las imágenes.

En la representación con UMAP de la Fig. 3, se observan varios *clusters* bien definidos, cada uno visualizado con un color diferente. La proximidad entre puntos implica semejanza entre las imágenes que los representan (ver los ejemplos para el *cluster* 5 en la Fig. 3, a simple vista, parecieran ser las mismas imágenes). Además, las anomalías aparecen dispersas y alejadas de los *clusters* principales, lo que sugiere su rareza y diferencia significativa respecto a las imágenes comunes. En el caso concreto del grupo de imágenes analizado en la Fig. 3, aparece un grupo pequeño con características muy próximas (señalado con la flecha roja), lo que implica alta probabilidad de haber capturado a un mismo individuo en varias fotografías.

Notar que si bien, la mayoría de las imágenes son muy parecidas (por ser tomas hechas en el mismo sitio), el modelo logra mapear las pequeñas sutilezas presentes en su vector de características, de modo de poder localizarlas con el posprocesamiento adecuado. En este caso, el pequeño grupo localizado, corresponde a una misma ave que ha sido capturada un total de 5 veces y se ejemplifica en la Fig. 4.

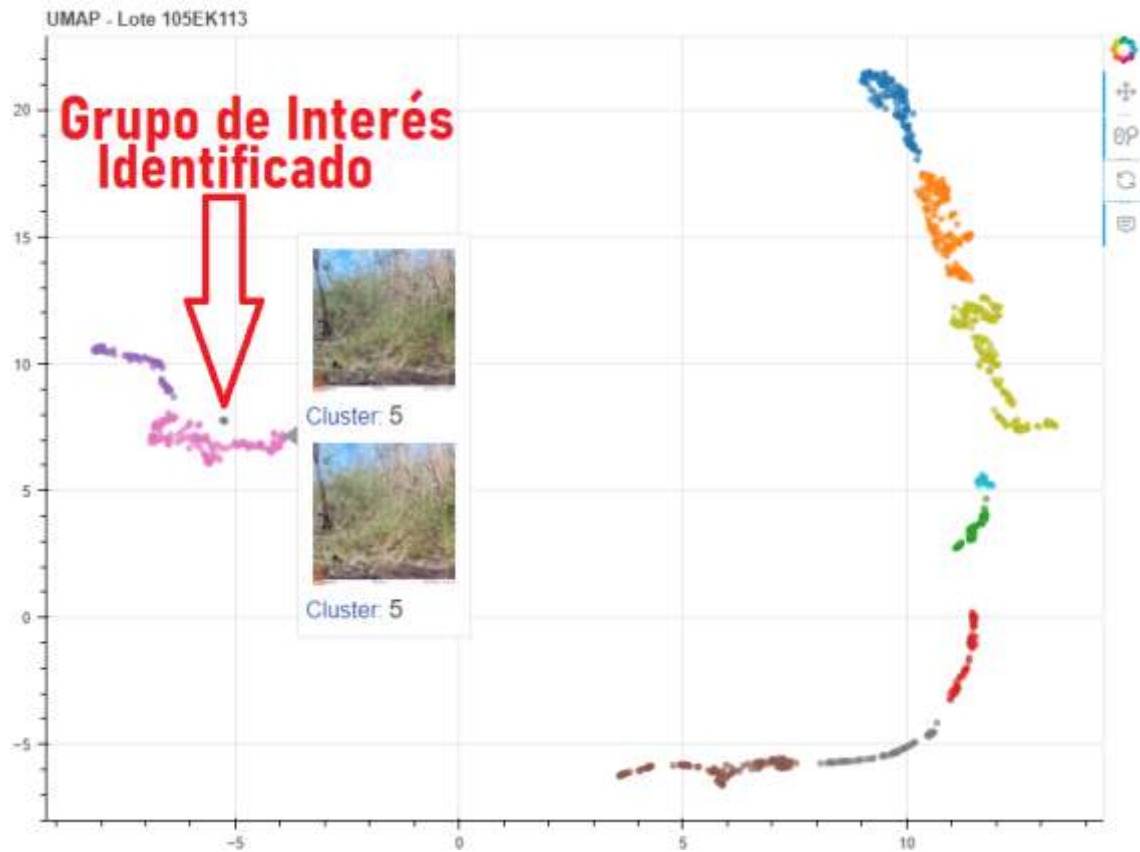


Fig. 3. Mapa de *embeddings* 2D utilizando UMAP, para un lote de 1000 imágenes.

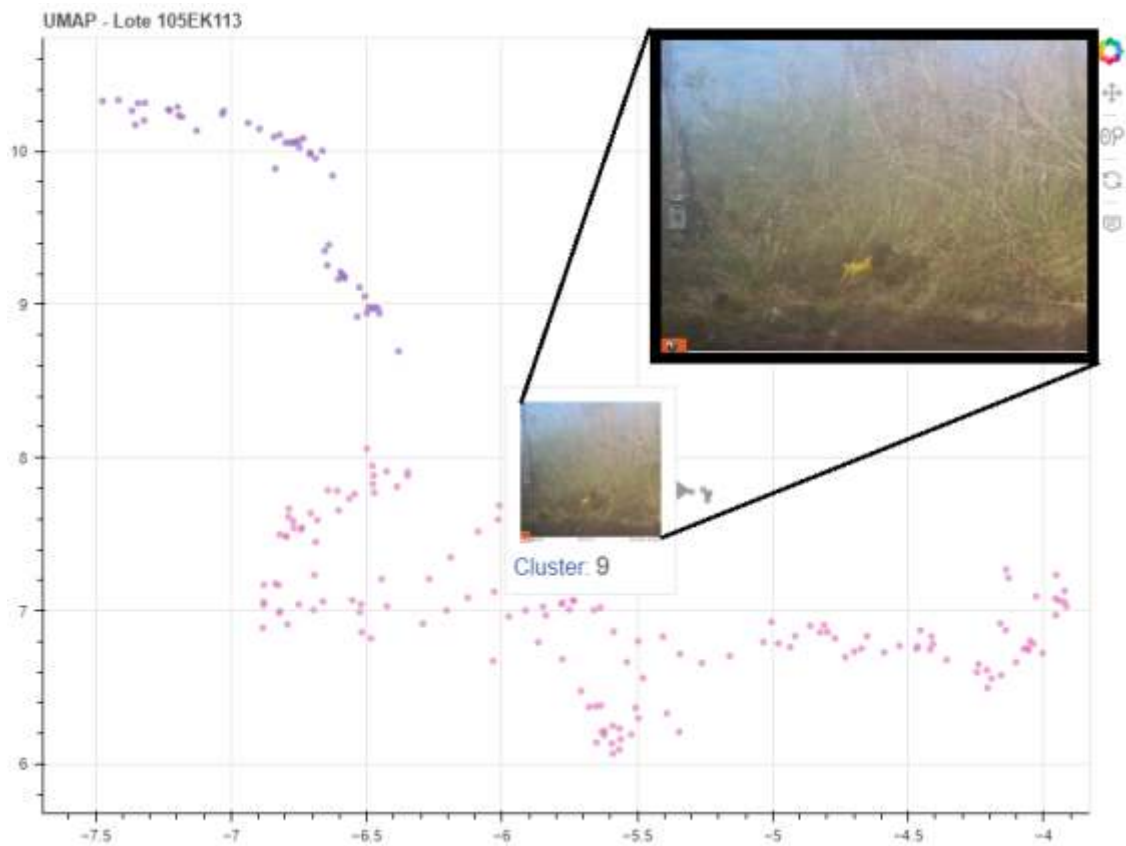


Fig. 4. Zoom sobre mapa de *embeddings* 2D de la Fig. 3, región de interés.

Cada tipo de agrupamiento puede explicarse de ser necesario, es decir, determinar las características que comparten en común. Por dar un ejemplo, en la Fig. 5 se aprecian varias muestras tanto del *cluster* 2 como del 4, mostrando que comparten las proyecciones de las sombras en el suelo, es decir, capturas que son similares en cuanto a la franja horaria de los distintos días. Se ha señalado esto en una de las 4 fotografías de ejemplo para cada *cluster*, y es sencillo de visualizar lo mismo en las demás. Hasta este punto, es evidente que, de las 1000 imágenes, si estuviéramos interesados en aves, solamente 5 eran de utilidad. Una exploración manual, puede ser costosa y extenuante.

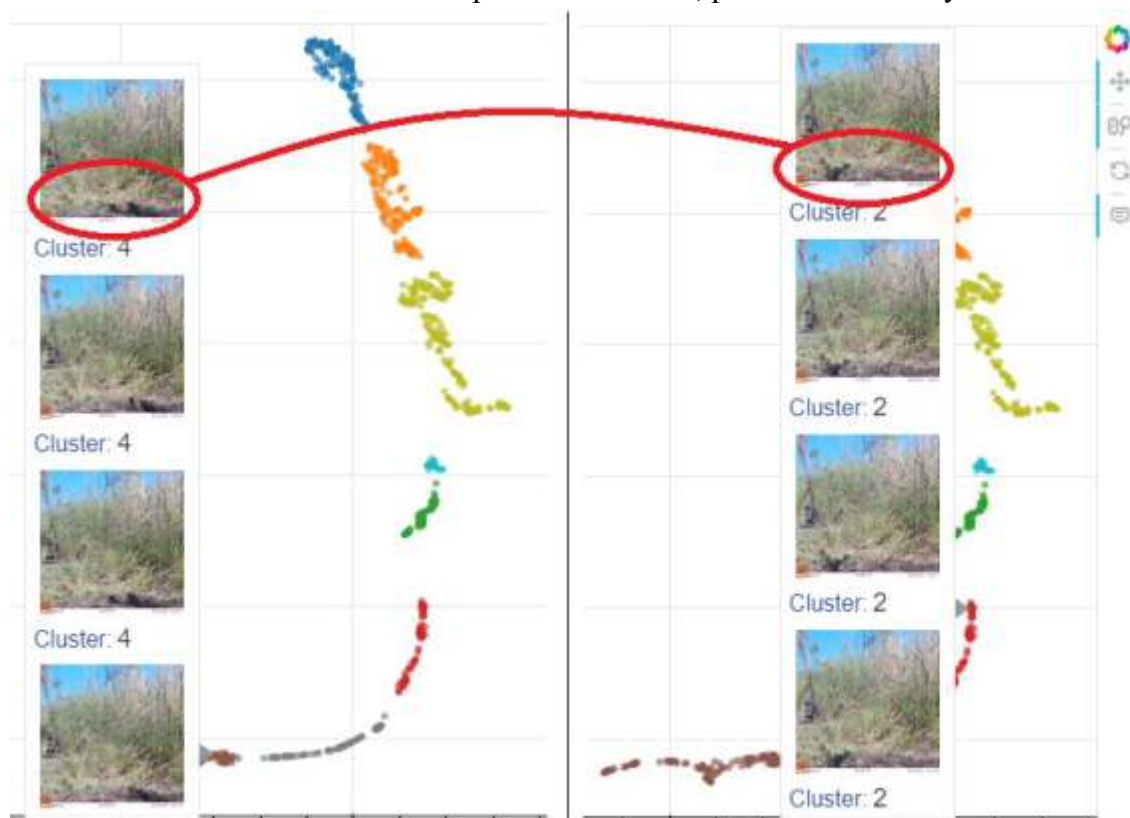


Fig. 5. Zoom sobre mapa de *embeddings* 2D de la Fig. 3, contraste de ejemplares de distintos *clusters*.

3.2 Mapeo de Texto a Imagen

El mapeo de texto a imagen es una funcionalidad crítica que permite la búsqueda y recuperación de imágenes basadas en descripciones textuales utilizando los *embeddings* generados por el modelo CLIP. Esta técnica aprovecha la capacidad del modelo para proyectar tanto texto como imágenes en un espacio de características común, donde la proximidad entre un texto y una imagen indica su similitud semántica.

Para ilustrar esta capacidad, se realizaron búsquedas utilizando términos específicos como “*bird*”, “*insect*”, “insecto”, “*butterfly*” y “mariposa”. Estos términos fueron seleccionados por su relevancia y diversidad, lo que permite evaluar la precisión del modelo al identificar imágenes correspondientes a diferentes categorías de fauna, incluyendo conceptos específicos o más abstractos (mariposa vs insecto, por ejemplo). En la Fig. 6, se presentan algunos ejemplos, denotando dónde se encuentra el objeto de interés hallado.



Fig. 6. Ejemplos de resultados de búsqueda para los términos, (a) “bird” y (b) “butterfly”.

Los resultados muestran que el modelo es capaz de recuperar imágenes altamente relevantes para los términos de búsqueda proporcionados. Por ejemplo, al buscar “bird”, el modelo identificó varias imágenes de aves en diferentes poses (las detectadas anteriormente en el *cluster* pequeño). Del mismo modo, al buscar “butterfly”, se recuperaron imágenes de mariposas en diversas posiciones y perspectivas. Para este tipo de casos, el problema se asemeja a “buscar una aguja en un pajar”, ya que, en un sublotte de 1.000 imágenes procesadas, solo 4 contenían mariposas. Además, el área ocupada por el individuo de interés es extremadamente pequeña en comparación con el paisaje general de las imágenes, lo que subraya la complejidad de la tarea. En la Fig. 7 se presentan más ejemplos para este caso. Encontrar manualmente este tipo de objetivos es altamente susceptible a errores humanos, debido al cansancio mental que implica prestar atención a detalles tan sutiles en un número tan elevado de fotografías.

Además, se destaca la capacidad del modelo para manejar la búsqueda en diferentes idiomas. Al utilizar los términos “insect” e “insecto”, el modelo proporcionó conjuntos de imágenes altamente similares, demostrando su robustez y capacidad para generalizar a través de diferentes lenguajes.

En muchos casos, si el individuo de interés tiene dimensiones pequeñas o sus colores no son perceptibles por “ojos humanos”, resulta complicado identificar elementos, incluso en las imágenes ya catalogadas como plausibles de contenerlos. En esos casos, realizar contrastes con imágenes cercanas puede facilitar la tarea. Por dar un ejemplo, en la Fig. 7(d), sobre la roca señalada, se observa una especie de chicharra que normalmente pasaría desapercibida. Sin embargo, en contraste con los ejemplos (a, b, c) de la misma figura, su presencia se vuelve evidente. Un caso similar, aunque con un contraste inverso, ocurre al localizar las mariposas señaladas.

Este mapeo de texto a imagen no solo facilita la recuperación rápida y precisa de imágenes, sino que también ofrece un enfoque poderoso para el análisis de grandes volúmenes de datos visuales, en contextos donde las anotaciones manuales no son prácticas.



Fig. 7. Comparativa de resultados para términos de búsqueda en inglés y español , (a, b) “mariposa”, (c) “insecto”, y (d) “insect”.

3.3 Eficiencia Computacional de la Implementación

La eficiencia computacional es un factor clave al implementar modelos avanzados como CLIP, especialmente en entornos con recursos limitados. Las pruebas se llevaron a cabo en el entorno de Google Colab utilizando únicamente CPU (sin GPU) para emular el rendimiento en un equipo portátil típico.

3.3.1 Hardware y Configuración

Las pruebas se realizaron en una instancia de Google Colab con la siguiente configuración:

- Procesador: Intel Xeon CPU con 2 vCPUs a 2,20 Hz.
- Memoria RAM: 13 GB.

3.3.2 Procesamiento de Imágenes

El modelo CLIP-ViT-B-32 ocupa 605 MB en memoria para almacenar los pesos de la topología. Esta carga de memoria condiciona la cantidad de imágenes que pueden ser procesadas simultáneamente en sistemas con recursos limitados. Es decir, como mínimo deben precargar en RAM los pesos del modelo más las librerías de gestión, y el espacio restante puede emplearse para precargar los tensores para las imágenes a procesar en lotes. Para simular un entorno típico de un equipo portátil, se procesaron las imágenes en lotes de 96 (32x3) para evitar sobrepasar la capacidad de la RAM (de una *notebook* estándar a la fecha, 8 GB de RAM).

- Tiempo promedio de procesamiento por lote: 42 segundos (incluyendo preprocesamiento y extracción de *embeddings*).

- Total de imágenes procesadas en las demostraciones: 1.000.
- Espacio en disco utilizado por las imágenes: ~3 GB.
- Espacio en disco utilizado por los *embeddings* generados: ~1,95 MB.

Es importante destacar que el uso de hardware con más hilos de procesamiento, mayor cantidad de RAM, o hardware de paralelización como GPUs, mejora significativamente estos tiempos.

3.3.3 Reducción de Dimensionalidad y Clustering

Para la reducción de dimensionalidad, se utilizó UMAP, que transformó los *embeddings* de 512 dimensiones a un espacio de 2 dimensiones para la visualización.

- Tiempo de procesamiento promedio de UMAP: 4 a 6 minutos (para 1.000 imágenes).

El *clustering* sobre los *embeddings* de 512 dimensiones con el algoritmo implementado fue más que satisfactorio:

- Tiempo promedio de *clustering*: 2 segundos (para 1.000 imágenes).

3.3.4 Procesamiento de Texto

El modelo CLIP-ViT-B-32-multilingual-v1 utilizado para procesar texto ocupa 539 MB en disco. Al igual que ocurre con el modelo para imágenes, esto condiciona el mínimo de RAM necesario para la carga del modelo. Aunque las pruebas se realizaron en un entorno sin GPU, los tiempos de inferencia fueron extremadamente rápidos en el procesamiento de texto, en el orden de los *ms*.

- Tiempos de inferencia para texto: Prácticamente despreciables, incluso al procesar múltiples términos de búsqueda simultáneamente.

3.3.5 Gestión de Recursos

Para optimizar el uso de recursos en sistemas con limitaciones de hardware, como podría ser un equipo portátil convencional, es viable cargar los modelos CLIP en sesiones separadas (mayor demora por carga y descarga de modelos a RAM, depende de la velocidad del disco). Esta estrategia permite gestionar la memoria RAM de manera más eficiente, asegurando que los modelos se carguen y utilicen de manera secuencial, sin exceder la capacidad disponible (si fuese necesario implementar todo el *pipeline* en un equipo de menores prestaciones a las mencionadas). Lo importante para el análisis, terminan siendo los vectores de 512 elementos, que, tanto en disco como en RAM, son ínfimos en espacio utilizado, al contrastarlos con fuente de datos original (imágenes).

En resumen, las pruebas realizadas en Google Colab no requirieron más de 8 GB de RAM en ningún momento, lo que demuestra la viabilidad de implementar este enfoque, incluso en entornos con recursos limitados.

4. Conclusiones

En este estudio se evaluó la capacidad del modelo CLIP-ViT-B-32 para extraer representaciones visuales significativas de imágenes capturadas en un entorno específico, utilizando la técnica UMAP para la reducción de dimensionalidad y visualización de los *embeddings* resultantes. Además, se realizó un análisis cualitativo de las distribuciones en el espacio reducido a 2D, identificando agrupamientos naturales y anomalías dentro de los datos.

El estudio demostró la eficacia del modelo CLIP en la búsqueda y recuperación de imágenes basadas en términos de texto. Lo que proporciona una herramienta invaluable para la exploración y análisis de grandes conjuntos de imágenes, permitiendo búsquedas semánticas precisas y eficientes

que pueden adaptarse a diferentes necesidades de investigación y aplicaciones prácticas. Esto resalta el potencial de CLIP en aplicaciones de búsqueda visual y recuperación de información multimodal, donde la precisión y la rapidez en la respuesta son esenciales.

Se evaluó la eficiencia computacional de las implementaciones en un entorno limitado, utilizando los recursos base de Google Colab. A pesar de las restricciones en hardware, el procesamiento de imágenes y texto se realizó de manera efectiva, con tiempos de procesamiento aceptables, tomando como muestras a conjuntos de 1.000 imágenes. Sin embargo, es evidente que, con mejoras en el hardware, como el uso de GPUs o sistemas con más capacidad de RAM, los tiempos podrían reducirse significativamente, haciendo la implementación más escalable y adecuada para aplicaciones en tiempo real o en grandes volúmenes de datos, si es que esto llegase a ser necesario.

Los resultados obtenidos en este estudio no solo validan la capacidad del modelo CLIP para tareas de análisis visual y multimodal, sino que también sugieren varias direcciones para futuras investigaciones. La integración de hardware más potente y la optimización de los algoritmos de reducción de dimensionalidad y *clustering* podrían mejorar aún más la precisión y eficiencia del sistema. Además, la exploración de técnicas adicionales para la detección de anomalías o la diferenciación de *clusters* podría proporcionar *insights* más detallados y valiosos en aplicaciones de monitoreo y análisis ambiental.

Actualmente, se está trabajando en el desarrollo de un software e interfaz gráfica que emplea las técnicas descritas en este estudio. El objetivo de esta herramienta es facilitar las tareas del personal dedicado a la protección, análisis e investigación de la fauna mediante cámaras trampa, proporcionándoles una herramienta eficiente y de fácil uso. Este software está siendo diseñado con un enfoque minimalista, permitiendo a los usuarios explorar grandes conjuntos de imágenes de manera intuitiva, incluyendo la capacidad de filtrar y buscar imágenes mediante términos de texto. Este desarrollo apunta a mejorar la eficiencia en la gestión y análisis de imágenes de fauna, ofreciendo una solución práctica para los desafíos cotidianos en el campo de la conservación y el monitoreo ambiental.

Agradecimientos

Este trabajo ha sido llevado a cabo gracias al aporte de imágenes de las cámaras trampa del Parque Federal Campo San Juan, quienes se vieron interesados en ayudar al desarrollo de alguna herramienta que les permita analizar sus datos. Agradecemos al intendente del Parque, Gimena Martinez y todo su equipo por la confianza en los autores de este trabajo.

Referencias

- [1] F. Rovero, F. Zimmermann, D. Bersi, y P. Meek, «“Which camera trap type and how many do I need?” A review of camera features and study designs for a range of wildlife research applications», 2013.
- [2] A. C. Burtonet al., «REVIEW: Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes», *J. Appl. Ecol.*, vol. 52, n.º 3, pp. 675-685, 2015, doi: 10.1111/1365-2664.12432.
- [3] R. Steenweg, M. Hebblewhite, J. Whittington, P. Lukacs, y K. McKelvey, «Sampling scales define occupancy and underlying occupancy-abundance relationships in animals», *Ecology*, vol. 99, n.º 1, pp. 172-183, 2018, doi: 10.1002/ecy.2054.
- [4] A. Swanson, M. Kosmala, C. Lintott, R. Simpson, A. Smith, y C. Packer, «Snapshot Serengeti-high frequency annotated camera trap images of 40 mammalian species in an African savanna», *Sci. Data*, vol. 2, n.º 1, p. 150026, jun. 2015, doi: 10.1038/sdata.2015.26.

- [5] *Camera Traps in Animal Ecology*. Accedido: 10 de junio de 2024. [En línea]. Disponible en: <https://link.springer.com/book/10.1007/978-4-431-99495-4>
- [6] M. S. Norouzzadeh *et al.*, «Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning», 15 de noviembre de 2017, *arXiv*: arXiv:1703.05830. doi: 10.48550/arXiv.1703.05830.
- [7] S. Beery, G. van Horn, y P. Perona, «Recognition in Terra Incognita», 24 de julio de 2018, *arXiv*: arXiv:1807.04975. doi: 10.48550/arXiv.1807.04975.
- [8] S. Schneider, G. W. Taylor, S. S. Linquist, y S. C. Kremer, «Past, Present, and Future Approaches Using Computer Vision for Animal Re-Identification from Camera Trap Data», 19 de noviembre de 2018, *arXiv*: arXiv:1811.07749. doi: 10.48550/arXiv.1811.07749.
- [9] G. Chen, T. X. Han, Z. He, R. Kays, y T. Forrester, «Deep convolutional neural network based species recognition for wild animal monitoring», en *2014 IEEE International Conference on Image Processing (ICIP)*, oct. 2014, pp. 858-862. doi: 10.1109/ICIP.2014.7025172.
- [10] M. A. Tabak *et al.*, «Machine learning to classify animal species in camera trap images: Applications in ecology», *Methods Ecol. Evol.*, vol. 10, n.º 4, pp. 585-590, 2019, doi: 10.1111/2041-210X.13120.
- [11] M. Willi *et al.*, «Identifying animal species in camera trap images using deep learning and citizen science», *Methods Ecol. Evol.*, vol. 10, n.º 1, pp. 80-91, 2019, doi: 10.1111/2041-210X.13099.
- [12] A. Radford *et al.*, «Learning Transferable Visual Models From Natural Language Supervision», 26 de febrero de 2021, *arXiv*: arXiv:2103.00020. Accedido: 10 de junio de 2024. [En línea]. Disponible en: <http://arxiv.org/abs/2103.00020>
- [13] A. Dosovitskiy *et al.*, «An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale», 3 de junio de 2021, *arXiv*: arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929.
- [14] X. Zhai, A. Kolesnikov, N. Houlsby, y L. Beyer, «Scaling Vision Transformers», 20 de junio de 2022, *arXiv* : arXiv:2106.04560. doi: 10.48550/arXiv.2106.04560.
- [15] M. Ester, H.-P. Kriegel, J. Sander, y X. Xu, «A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise».
- [16] L. McInnes, J. Healy, y J. Melville, «UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction».

Los Pequeños Aprovechamientos Hidroeléctricos y la Necesidad de Especialistas

Ariel R. Marchegiani ^{1*} , Orlando A. Audisio ² , Victor H. Kurtz ³ 

¹ Facultad de Ingeniería, Universidad Nacional del Comahue (UNCo), Neuquén, Argentina.

² Facultad de Ingeniería, Universidad Nacional del Comahue (UNCo), Neuquén, Argentina.

³ Facultad de Ingeniería, Universidad Nacional de Misiones (UNaM), Oberá, Misiones, Argentina.

e-mails: ariel.marchegiani@fain.uncoma.edu.ar, orlando.audisio@fain.uncoma.edu.ar, kurtzvh@fio.unam.edu.ar

Resumen

La falta de competencia, el reducido número de fabricantes nacionales y la escasa investigación en el tema lleva a una reticencia a "pensar en pequeño", junto con la situación económica del país, son algunos de los muchos factores y/o actores que hacen que la factibilidad en el desarrollo y uso de la energía hidroeléctrica a pequeña escala sea de escasa viabilidad. Es necesario un cambio de actitud, así como una mayor difusión de los conocimientos técnicos a pequeña escala, si se quiere generalizar el uso y la práctica de esta tecnología, para poder aumentar la participación de los mismos en la matriz energética nacional. En este trabajo se analizan algunos puntos que intervienen en el proyecto y diseño de un Pequeño Aprovechamiento Hidroeléctrico (PAH) y las acciones que hacen dificultosa la toma de decisiones a la hora de encarar un emprendimiento de este tipo.

Palabras Clave – Energías Renovables, Pequeñas Centrales Hidráulicas, Pequeños Aprovechamientos Hidroeléctricos, Factibilidad Técnico - económicas de PAH

Abstract

Lack of competition, reluctance to "think small" and the economic situation of the country are some of the many factors and/or involved parties that make the feasibility in the development and use of small-scale hydropower of low viability. A change of attitude is needed, as well as a greater dissemination of small-scale know-how, if the use and practice of this technology is to be generalized, in order to increase its participation in the national energy matrix. This paper analyzes some points involved in the project and design of a Small Hydroelectric Power plant (SHP) and the actions that make it difficult to make decisions at the time of undertaking this type of project.

Keywords – Renewable Energies, Small Hydroelectric Power Plants, Small Hydroelectric Developments, Technical - economic feasibility of SHP.

1. Introducción

Ya hace varias décadas se ha reconocido que los pequeños aprovechamientos hidroeléctricos (PAH) pueden complementar la producción de combustibles fósiles y proporcionar un suministro eléctrico económico y fiable a comunidades aisladas. En consecuencia, muchos países han puesto en marcha ambiciosos programas de desarrollo y, ahora que ya se han puesto en marcha algunos de estos sistemas, resulta oportuno examinar los progresos realizados y los problemas que han surgido.

Los PAH o Pequeñas Centrales Hidráulicas se clasifican según el salto o altura que aprovechan, como puede verse en la Tabla 1.

En el año 2022 el 13,9% de la demanda total de energía eléctrica argentina fue abastecida a partir de fuentes renovables. Se trata de una cifra promedio que pone de manifiesto el continuo avance del sector dentro del conjunto de la matriz energética nacional, a la que contribuye a diversificar mediante la incorporación de nuevas fuentes de generación renovable que suman capacidad en la red.

En cuanto a la generación, la tecnología eólica fue la de mayor participación, aportando el 73% de la generación eléctrica por fuentes renovables; seguida de la solar (15%), las bioenergías (6%) y los pequeños aprovechamientos hidroeléctricos (5%), según datos de la Compañía Administradora del Mercado Mayorista Eléctrico Sociedad Anónima (CAMMESA) [1]. Como puede verse, los pequeños

aprovechamientos hidroeléctricos fueron los que menos aportaron a la matriz. Es necesario analizar algunas causas de esto.

Tabla1. Clasificación de los PAH

Denominación	Rango de Potencias	Salto [m]		
	[kW]	BAJO	MEDIO	ELEVADO
MICROCENTRALES	hasta 100	< 15	15-50	> 50
MINICENTRALES	100-1000	< 20	20-100	> 100
PEQUEÑAS CENTRALES	1000-15000	< 25	23-130	> 130

De hecho, en el informe de CAMMESA se puede observar que, respecto del 2021, la generación hidroeléctrica renovable tuvo una menor participación relativa (Fig. 1) [2].

Como puede observarse, en los últimos años ha habido un aumento en la potencia instalada proveniente de las energías renovables gracias a la legislación vigente, como por ejemplo la Ley 27.424, los Certificados de Crédito Fiscal (CCF), beneficios promocionales que brinda la Subsecretaría de Energía Eléctrica para la instalación de equipos de generación distribuida, los Renovar 1, 1.5 y 2 con los cuales, por ejemplo, se construyeron varios PAH, en la provincia de Mendoza, en el Canal Cacique Guaymallén: salto N.º 6 (1,2 MW), N.º 7 (1,2 MW), N.º 8 (1,2 MW) y N.º 11 (0,51 MW). Los otros 3 construidos fueron el Triple Salto Unificado del canal San Martín en Junín (0,5 MW), entre otros. En la provincia de Neuquén se encuentra en ejecución la Pequeña Central Hidroeléctrica Nahueve (4,6 MW) (Fig. 2).

El objetivo de este trabajo es analizar algunos puntos que intervienen en el proyecto y diseño de un Pequeño Aprovechamiento Hidroeléctrico (PAH) y las acciones que hacen dificultosa la toma de decisiones a la hora de encarar un emprendimiento de este tipo y características.

2. La problemática de los PAH

Quizá la principal conclusión sea que, a diferencia de las centrales diésel, las mini centrales hidroeléctricas no pueden fabricarse en serie. Cada una debe diseñarse individualmente, ya que cada emplazamiento es único y tiene sus propios problemas peculiares. Esto ha resultado difícil porque hay pocos ingenieros con experiencia en este campo y existe poca bibliografía específica de referencia. El riesgo de esto es que muchos de los proyectos resulten poco factibles y confiables, por lo que existe el peligro de que los propietarios y los operadores se desilusionen y el interés por los proyectos pequeños desaparezca tan rápidamente como surgió. Esta cuestión se agrava cuanto menor es la potencia que puede obtenerse, concretamente cuando se trata de micro aprovechamientos (hasta 100 kW).

Sin embargo, no tiene por qué ser así: en todo el mundo hay miles de pequeñas centrales hidroeléctricas, algunas de las cuales llevan funcionando muchos años, que han contribuido enormemente al desarrollo rural. Muchas de las más exitosas fueron construidas por ingenieros locales con un presupuesto mínimo y financiación procedente de recursos nacionales, como los mencionados en párrafos anteriores. Esto contrasta con algunos de los programas de PAH financiados por organismos de crédito internacionales aplicados a proyectos diseñados por grandes empresas

consultoras internacionales. Algunas de estas empresas, hace varias décadas atrás, han reconocido los problemas específicos y para resolverlo, movilizaron a equipos de especialistas [3], pero otras, con poca experiencia previa, han elaborado planes que fueron concebidos y diseñados de manera errónea.

RESUMEN EJECUTIVO - PRINCIPALES VARIABLES MEM - ENERO A JUNIO 2022

FUENTE DE ENERGÍA	ENE - JUN 2021	ENE - JUN 2022	Variación %
BIOMASA	317	329	3.7%
EOLICO	5 939	6 961	17.2%
HIDRO <= 50MW	617	502	-18.7%
SOLAR	901	1342	49.0%
BIOGAS	163	224	37.7%
Total [GWh]	7 936	9 357	17.9%
Demanda MEM	65 945	69 464	
Renovables MEM/ Dem MEM	12.0%	13.5%	

El crecimiento de las energías renovables hizo aumentar su participación en el cubrimiento de la demanda, pasando de 12% en el 2021 a un 13.5% en el año 2022 sobre la demanda abastecida, donde en este último período se alcanzó en forma puntual casi a un 16% de participación mensual.

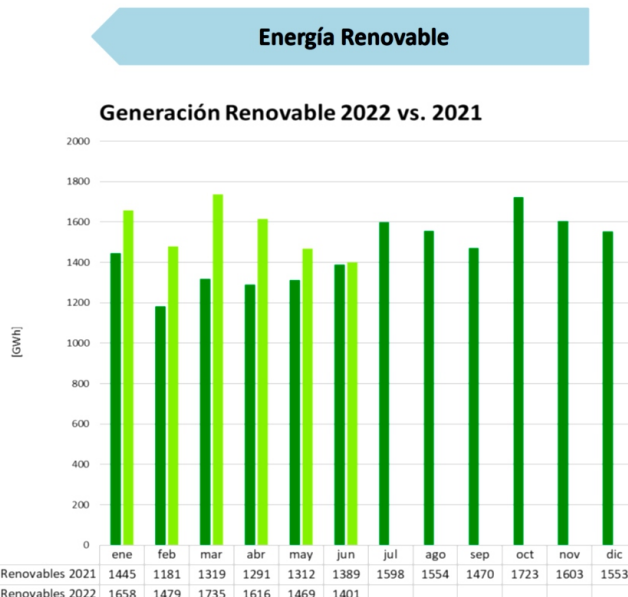


Fig. 1. Aporte de las EERR a la demanda energética Argentina periodo ene-iun 2021 – ene-iun 2022 [2].



Fig. 2. PAH Nahueve (Prov. de Neuquén) actualmente en construcción.

La clave del éxito pasa a través de realizar un estudio exhaustivo de los recursos de la zona, que indique dónde está el potencial, seguido de un estudio de viabilidad de las pequeñas centrales hidroeléctricas. Las pequeñas centrales hidroeléctricas de por sí, tienen un costo específico alto (sobre todo las de menor potencia), por lo que hay que hacer todo lo posible por extraer el último metro de altura y el último kilovatio del potencial del emplazamiento y, al mismo tiempo, seleccionar el equipo más económico y adecuado.

Desgraciadamente, muchos estudios de viabilidad suelen ser o bien muy baratos y superficiales, que aprovechan mal el potencial del emplazamiento, o bien muy caros, que, aunque optimizan la

potencia de la máquina hasta el último kilovatio, pueden ignorar las incertidumbres de la hidrología o proponer proyectos tan complejos y cómodos que, en comparación, la energía diésel resulta más ventajosa. Algunos estudios de optimización no ponderan en la medida que corresponde, las grandes variaciones estacionales del caudal de agua y utilizan curvas anuales de duración del caudal para abastecer el sistema con energía de la estación húmeda que el sistema no puede absorber, sin tener en cuenta el coste de reforzar la escasa producción de la estación seca con generación térmica. Un caso claro de esto fue el programa encarado por la provincia de Neuquén en la década del 80, que produjo un número considerable de micro centrales, de las cuales casi ninguna está en operación en la actualidad, si bien ha habido varios intentos de restauración.

En el caso de las máquinas de muy baja potencia (micro turbinas, Fig. 3), donde generalmente existe una producción “artesanal”, de turbinas que, generalmente realizan fabricantes locales no vinculados exclusivamente al tema, y por lo tanto, no siempre disponen un know-how apropiado.

Sus productos presentan defectos o malos diseños tanto en la parte hidráulica como mecánica, dejando como resultado máquinas que no tienen una eficiencia adecuada e inclusive un mal funcionamiento.

La producción de equipos que no cumplen adecuadamente con las prestaciones requeridas por los usuarios retrae el interés en las energías renovables, colocándola en un escenario más difícil de penetración [4].

2.1. Algunos problemas que pueden presentarse

Una vez identificado el sistema, puede comenzar el diseño detallado de los distintos componentes. Esto también debe hacerse con cuidado.

Para el operador, la toma de agua es la parte más crítica de todo el proyecto (Fig. 4). Incluso durante las crecidas, debe desviar el agua que normalmente va hacia las turbinas mientras el río arrastra cantos rodados, piedras, barro, troncos u hojas. Las rejillas deben poder limpiarse o auto limpiarse.

Dado que la mayoría de los proyectos pequeños implican un desvío del río en lugar de una presa, los canales, las tuberías y las compuertas suelen ser la partida más costosa del proyecto. En este sentido, deben tenerse en cuenta los canales de tierra que son los más económicos para desviar el agua de una toma a un estanque de cabecera.

Las tuberías forzadas representan un costo muy importante en el proyecto, y su diseño suele basarse en supuestos elementales de tensión, combinados con un gran factor de seguridad (es decir, un factor de ignorancia). Esto no sólo da lugar a tuberías innecesariamente pesadas, sino también a la utilización de longitudes de tubería cortas y, por tanto, a un número excesivo de juntas de obra y pedestales. Esto encarece el precio y aumenta el plazo de construcción.

El primer paso en el diseño de una central hidroeléctrica es definir el tipo y la potencia de las turbinas y los generadores. Aunque ya se ha determinado el caudal máximo que debe turbinar cada máquina, cuando hay una gran variación de altura, la determinación de los kW no resulta una definición simple. En primer lugar, para maximizar el rendimiento, hay que determinar el salto neto de diseño. Como éste es siempre inferior al salto máximo normal, la potencia nominal también será inferior a la determinada con la altura máxima. Aunque pueda parecer lógico especificar que el generador sea compatible con esta potencia nominal, la realidad indica que es recomendable que la potencia de la turbina sea más compatible con una altura de elevación cercana o igual a la altura de elevación normal máxima.

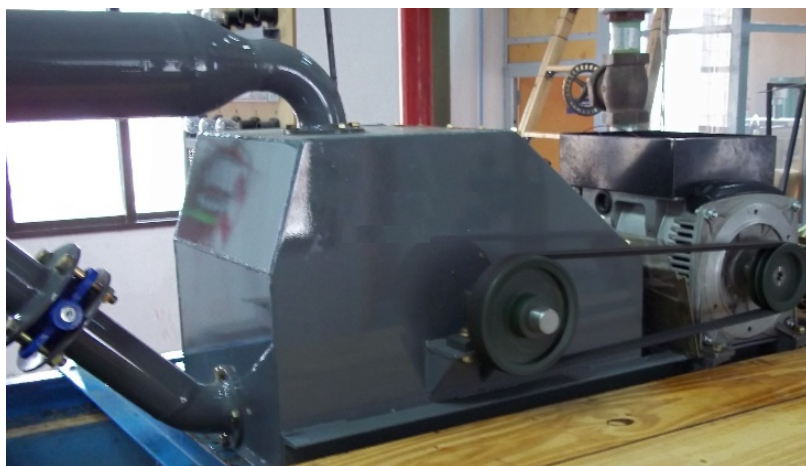


Fig. 3: Turbina Pelton de 6 kW de fabricación nacional.



Fig. 4: Obra de toma con Rejilla auto-limpiante tipo Coanda.

Otro aspecto que a menudo se descuida es el vinculado a los distintos ajustes que hay que hacer en los precios de licitación a fin de considerar las diferencias en el rendimiento de la turbina a plena carga y a carga parcial, las pérdidas de carga de la válvula de entrada, las pérdidas de salida del tubo de aspiración o descarga, la elevación del eje central y otros aspectos similares. Cuando estos factores están claramente definidos, el fabricante puede seleccionar entre su gama de modelos de turbinas y equipos complementarios, de modo de definir una combinación que ofrezca el mejor rendimiento al menor coste. Si el ingeniero consultor asume la responsabilidad de elegir el tipo de turbina, la velocidad y otros parámetros, el riesgo que corresponde al fabricante se traslada al ingeniero consultor y la competencia entre fabricantes se ve gravemente limitada. Para minimizar los costos y las complicaciones innecesarias, la especificación de las máquinas debe adecuarse al esquema en su conjunto y al sistema que suministra.

La experiencia demuestra que el fallo más común es el uso de especificaciones redactadas originalmente para máquinas grandes (es decir, de más de 50 MW), generalmente de eje vertical, y reeditadas para turbinas pequeñas sólo con cambios menores. Uno de los peores aspectos de esta práctica es que se especifican máquinas verticales donde deberían utilizarse máquinas horizontales.

Cuando el caudal es inferior a unos $12 \text{ m}^3/\text{s}$ y pueden utilizarse, por ejemplo, turbinas Francis, las máquinas y generadores horizontales son más baratos, mucho más rápidos de montar y más fáciles de mantener. Los tubos de aspiración pueden ser más sencillos y, si es necesario, se puede añadir un gran volante de inercia al extremo opuesto del generador para proporcionar una gran inercia a bajo costo (Fig. 5).

Los encargados de redactar pliegos con especificaciones deben considerar que la mayoría de los fabricantes han optimizado los diseños de sus máquinas y, además tienen estandarizados sus desarrollos. Por lo tanto, el pliego de condiciones no debe establecer requisitos absolutos sobre el diseño, sino que debe exponer claramente las condiciones de emplazamiento exigidas. Sin embargo, muchos redactan pliegos de condiciones rígidas que insisten en un cumplimiento absoluto y que no permiten a los fabricantes ofrecer lo que consideran la provisión más adecuada y óptima.

Es conveniente aprovechar la experiencia que tiene cada proveedor para los pequeños proyectos hidroeléctricos, con el fin de reducir al mínimo los plazos de entrega, costos y logísticas, de este modo resulta más factible desarrollar muchos proyectos hidroeléctricos que, de otro modo, serían antieconómicos. Además, en esta misma dirección es importante evaluar que muchos fabricantes de turbomáquinas de baja potencia son reticentes a licitar con pliegos de condiciones que contemplan detalles y/o requisitos excesivamente detallados.

Una vez definidas y seleccionadas las máquinas, la racionalidad indica que es factible avanzar y empezar el diseño de la central. Aunque los diseños suelen escalarse a partir de grandes centrales existentes, es esencial contemplar el diseño inicial con ojo crítico y eliminar las características no esenciales que aumentan costos, elevan dificultades operativas y el tiempo de construcción.

2.2. Algunos aspectos del sistema eléctrico, como ejemplo

El diseño del sistema eléctrico es otra característica que suele tomarse como referencia de los diseños de las grandes estaciones y sistemas generadores de alta potencia. Los que operan conectados en paralelo con líneas generales de distribución eléctrica comercial.



Fig. 5 PAH con turbinas tipo Francis de eje Horizontal (Brasil).

No es raro encontrar tres o cuatro turbomáquinas pequeñas con transformadores unitarios e interruptores de alta tensión conectados a sistemas de alta tensión (66 kV y superiores). Con esta disposición (Fig. 6) se necesitan muchos disyuntores de alta tensión y pequeños transformadores de

alta tensión, todos ellos caros. Una alternativa sería utilizar uno o dos transformadores elevadores con o sin conmutadores de alta tensión y conmutar a la tensión del generador.

En los sistemas pequeños no existe tal restricción y, en general, no deben utilizarse transformadores unitarios. La disposición mostrada en la Fig. 7 es más económica, más fiable y fácil de manejar que la mostrada en la Fig. 6.

Por otro lado, en el caso de que el sistema genere en forma aislada de la red eléctrica comercial. Es importante analizar minuciosamente el tipo de generador a utilizar y el equipamiento de control de velocidad (frecuencia) y tensión generada. Que son especiales y diferentes a los utilizados en los sistemas hidroeléctricos de gran porte.

Por lo tanto; si los diseños racionalizados no están bien concebidos, los esquemas resultarán malos. Si se diseñan individualmente, algunos serán buenos y otros malos. No obstante, la racionalización por sí sola puede permitir a ingenieros relativamente inexpertos diseñar y construir buenos esquemas en un tiempo relativamente corto.

La racionalización u optimización de las turbinas y generadores nunca deberían ser realizados por los encargados de elaborar el proyecto; eso es algo exclusivo que le compete y, por lo tanto, lo definen y deciden; los fabricantes y el mercado. En este sentido, aquellos que racionalicen sus diseños estandarizando con detalles, utilicen herramientas informáticas para determinar dimensiones y producir planos, y hagan uso de máquinas-herramientas de control numérico, tendrán productos con costos más bajos y, presumiblemente, una mayor cuota de mercado.

No hay que olvidar que la clave principal para reducir el costo de las máquinas es la competencia. Cuando se especifica un determinado diseño de máquina, la competencia es limitada y los precios serán elevados.

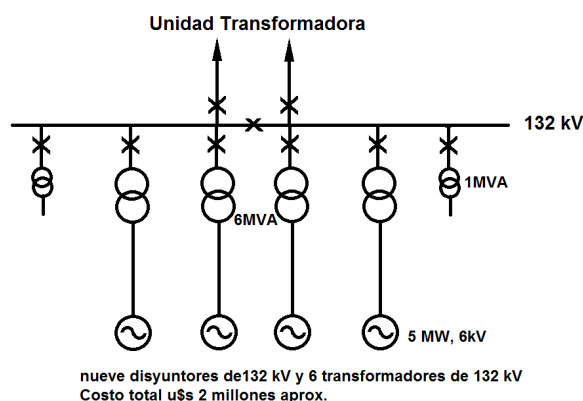


Fig. 6: Típica conexión eléctrica para 4 máquinas.

Para superar los problemas que han surgido se requiere un esfuerzo concertado y decidido por parte de las organizaciones nacionales responsables de los programas de desarrollo, los organismos internacionales de crédito, los ingenieros y los fabricantes. Las organizaciones nacionales tienen la tarea más difícil. En la mayoría de los casos, se encuentran con un campo nuevo, con poco dinero y sin experiencia que les sirva de guía. A menudo no son capaces de realizar estudios de viabilidad al nivel exigido por los organismos internacionales de crédito, por lo que muchas veces deben recurrir a organizaciones o empresas que le brinden ayuda para el diseño de ingeniería y el suministro de equipos. En repetidas ocasiones, y aunque el tipo de interés sea inferior, los diseños y el equipo suelen ser tan innecesariamente caros que los reembolsos reales son muy elevados, lo que lleva a desestimar los proyectos.

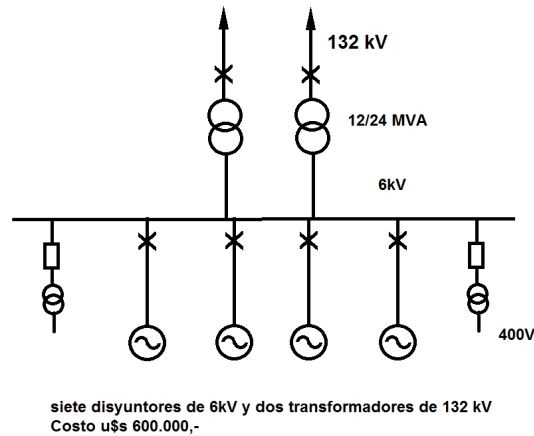


Fig. 7: Una variante de la conexión mostrada en la Fig. 6, más económica e igual de efectiva.

3. Conclusiones

Para que los proyectos de PAH sean eficaces y fiables al menor costo posible, se debe poder obtener los servicios de ingenieros y especialistas que ya hayan resuelto los problemas que probablemente se les planteen (experiencia profesional) y adquirir en el mercado los equipos que mejor se adapten a sus necesidades. Estos son los problemas que deben superarse.

Se debe tratar de sensibilizar y promover la reflexión en los diversos sectores sobre la importancia de fomentar las pequeñas centrales hidroeléctricas en el país y la importancia de considerarlas como fuentes renovables no convencionales viables y soluciones para mejorar la matriz energética nacional.

Aunque las agencias de ayuda desean realmente que su dinero se gaste bien, su burocracia y su política de selección de consultores juegan en su contra, ya que en general su experiencia previa es con grandes proyectos individuales de desarrollo que cuestan cientos de millones de dólares y que tardan muchos años en planificarse y construirse. Estos proyectos exigen el empleo de grandes organizaciones de consultoría con muchos especialistas en campos concretos que elaboran informes detallados (y caros) que cubren todos los aspectos del plan y que pueden proporcionar fácilmente los grandes equipos de ingenieros necesarios para diseñar y construir estos proyectos. Este planteamiento no es adecuado para las pequeñas centrales hidroeléctricas. El costo global y la rapidez exigen que sean diseñados por un pequeño equipo de ingenieros versátiles que puedan optimizar y racionalizar este tipo de proyectos.

Las diversas instituciones, ONG's y entes gubernamentales, deben brindar apoyo, a fin de obtener un producto de calidad, que satisfaga la demanda energética con un adecuado nivel de eficiencia. Además, las universidades deben llevar adelante algunos programas juntamente con fabricantes locales a fin de lograr diseños adecuados y eficientes, tratando de incrementar el interés en las energías renovables, y propiciar un escenario de penetración más favorable.

La formación de especialistas en el campo de los PAH asegura que se lleven a cabo implementaciones técnicas adecuadas y sostenibles. Los especialistas son fundamentales para evaluar el potencial hidroeléctrico de un área, diseñar infraestructuras adecuadas y gestionar los recursos hídricos de manera responsable. Esto incluye el estudio de impactos ambientales y sociales, garantizando que el desarrollo no afecte negativamente a los ecosistemas locales ni a las comunidades cercanas, [5], [6].

Otro aspecto relevante es la innovación tecnológica. La formación especializada permite la investigación y el desarrollo de nuevas tecnologías que pueden optimizar la eficiencia de los pequeños

aprovechamientos hidroeléctricos, reduciendo costos y aumentando su competitividad frente a otras fuentes de energía.

La capacitación de profesionales en este sector también contribuye a la creación de empleo y al desarrollo de capacidades locales, lo que fortalece las economías regionales y fomenta el compromiso comunitario con proyectos de energía renovable. En este sentido, el Laboratorio de Máquinas Hidráulicas de la U.N.Co. (Universidad Nacional del Comahue). Viniendo a cabo programas en conjunto con otras universidades como la U.Na.M. (Universidad Nacional de Misiones) y con fabricantes locales con el fin de lograr diseños adecuados y eficientes que permitan incrementar el interés en las energías renovables, y tratar de propiciar un escenario de penetración más favorable.

La Universidad Nacional del Comahue muestra su compromiso con la formación de personas y profesionales responsables que aportan soluciones sostenibles y son conscientes de las necesidades especialmente en la infraestructura del país, mediante el dictado de cursos y seminarios de los temas en cuestión.

En el contexto de plantear soluciones a la situación de posible desabastecimiento de energía eléctrica, como instituciones de educación superior, nuestro propósito debe ser, no solo de preparar adecuadamente a los profesionales que deben enfrentar estas situaciones, sino de reflexionar sobre las características, ventajas y desventajas, la factibilidad y sostenibilidad de sistemas como las pequeñas centrales hidroeléctricas de tal manera que se cuente con información relevante que nos permita tomar decisiones acertadas sobre la producción de energía y una producción de energía más amigable con el medio.

Agradecimientos

El presente trabajo se encuentra enmarcado en el Proyecto de Investigación 04/I-268 “Estudio y Desarrollo de Turbomáquinas y Sistemas Asociados Aplicados a Pequeñas Fuentes de Energía Renovable II”, aprobado por la Universidad Nacional del Comahue, (2022-2025).

Referencias

- [1] argentina.gob.ar, (enero de 2023). *Energías Renovables 2022: se cubrió el 13,9% de la demanda y se inauguraron 8 proyectos por más de 47 MW de potencia instalada* [Online]. Available: <https://www.argentina.gob.ar/noticias/energias-renovables-2022-se-cubrio-el-139-de-la-demanda-y-se-inauguraron-8-proyectos-por>.
- [2] CAMMESA, (Agosto 2022). *Resumen Ejecutivo Ene a Jun 2022 vs 2021*, Available: <https://cammesaweb.cammesa.com/download/resumen-ejecutivo-ene-a-jun-2022-vs-2021/>
- [3] Gordon, J.C., "Small Hydro Puts New Challenge to consultants", *Energy International*, Vol. 17, # 8, p.p. 19-21, Aug. 1980,
- [4] Marchegiani A. R, Audisio O. A., Kurtz V. H., “Prácticas frecuentes en el diseño y fabricación de micro turbinas hidráulicas: Estudio de casos”, *Revista + INGENIO, Revista de Ciencia Tecnología e Innovación* Vol.1 no.2, pp. 35-44, julio-diciembre 2019, ISSN 2618-5520 [online], ISSN 2618-5520 [impresa].
- [5] Kishore, T.S.; Patro, E.R.; Harish, V.S.K.V.; Haghighi, A.T., A Comprehensive Study on the Recent Progress and Trends in Development of Small Hydropower Projects. *Energies* 2021, <https://doi.org/10.3390/en14102882>, MDPI, Basel, Switzerland
- [6] Hira Singh Sachdev, Ashok Kumar Akella, Niranjana Kumar, Analysis and evaluation of small hydropower plants: A bibliographical survey, *Renewable and Sustainable Energy Reviews*, Volume 51, November 2015, p.p 1013-1022.
- [7] Liu H, Esser L, Masera, D, *World Small Hydropower Development Report*, Jul 01, 2013, TRN: CN14OA501, china, Latest edition is available from www.smallhydropowerworld.org.

Evaluación de la Generación de Residuos Per Cápite en los Barrios de Oberá, Misiones, Argentina

Sonia Romina Niezwida^{1,*} , María Fernanda Kaczynski¹ , Mariana Gisel Jansat¹ ,
Matías Ricardo Stoffel¹ , Sonia Ester Yasinski¹ , Juan Carlos Michalus¹ 

¹ Universidad Nacional de Misiones. Facultad de Ingeniería. Oberá, Misiones, Argentina.

e-mails: sonia.niezwida@fio.unam.edu.ar, kaczynskifernanda@gmail.com, mariana.jansat@gmail.com,
matiasstoffel17@gmail.com, sonia.yasinski@fio.unam.edu.ar, michalus@fio.unam.edu.ar

Resumen

Este artículo analiza la generación per cápita de residuos sólidos en un municipio de la provincia de Misiones, Argentina, mediante un enfoque cuantitativo. El objetivo ha sido determinar la cantidad de residuos producidos por habitante, por barrio y por día (kg/hab.barrio.día) en la ciudad de Oberá. Para ello, se recopilaron datos sobre el volumen de residuos recolectados, la población atendida, los recorridos, horarios y frecuencias de recolección en los distintos barrios. Posteriormente se aplicaron herramientas estadísticas—valor medio, mediana, moda y desviación estándar—para analizar la información obtenida. Con base en la densidad poblacional de cada barrio y la cantidad de residuos recolectados, se calculó la Generación Per Cápite (GPC) por barrio. Los resultados evidencian patrones diferenciados de generación de residuos entre barrios, lo que subraya la necesidad de diseñar estrategias de gestión adaptadas a las características socioeconómicas y culturales de cada sector. En este sentido, el estudio aporta información clave como punto de partida para la formulación de políticas municipales más efectivas y sostenibles para la gestión de residuos domiciliarios.

Palabras Clave – Residuos, Generación Per Cápite, Gestión, Oberá, Barrios

Abstract

This article analyzes the per capita generation of solid waste in a municipality of the Misiones Province, Argentina, using a quantitative approach. The objective was to determine the amount of waste produced per inhabitant, per neighborhood and per day (kg/inhab.neighborhood.day) in the city of Oberá. To this end, data were collected on the volume of waste gathered, the population served, as well as the collection routes, schedules, and frequencies across different neighborhoods. Statistical tools—such as mean, median, mode, and standard deviation—were then applied to analyze the information obtained. Based on the population density of each neighborhood and the amount of waste collected, the Per Capita Generation (PCG) was calculated by neighborhood. The results reveal differentiated patterns of waste generation among neighborhoods, underscoring the need to design management strategies tailored to the socioeconomic and cultural characteristics of each sector. In this regard, the study provides key information as a starting point for the formulation of more effective and sustainable municipal policies for household waste management.

Keywords – Waste, Per Capita Generation, Management, Oberá, Neighborhoods

1 Introducción

La gestión de los Residuos Sólidos Urbanos (RSU) es un desafío importante para las ciudades en crecimiento y un tema promisorio en los municipios argentinos cuyos valores de GPC se encuentran entre 0.6-1.3 según los estudios publicados [1,2,3,4,5].

El municipio de Oberá, ubicado en el centro de la provincia de Misiones, Argentina se encuentra en pleno crecimiento y muestra un marcado interés en mejorar y promover la sostenibilidad entre sus habitantes [5]. En este contexto, la municipalidad otorga una especial atención a la Gestión Integral de Residuos Sólidos Urbanos (GIRSU).

Los datos oficiales de generación de residuos en la provincia de Misiones datan del año 2010, donde se ha publicado que en los municipios de Misiones se generaron entre 0,62 (Oberá, Iguazú, Eldorado) y 0,716 (Capital) kg por habitante por día [6].

Actualmente, varios municipios de Misiones cuentan con eco puntos que son lugares para depositar materiales reciclables y puntos limpios para que los vecinos acerquen los residuos voluminosos tales como ramas y electrodomésticos en desuso. Sin duda, estos lugares son estratégicos para disminuir la cantidad de residuos a depositar en rellenos sanitarios. Sin embargo, uno de los problemas que se presenta en el municipio de Oberá para rediseñar políticas públicas y estrategias de GIRSU más eficaces, es la escasa información sobre los RSU generados por habitante por barrio y por día (kg/hab.barrio.día) en la ciudad. Por tal motivo, el objetivo de este trabajo ha sido determinar los kg de residuos/hab.barrio.día que se generan en la ciudad de Oberá, Misiones, Argentina.

2 Materiales y Métodos

Para llevar adelante la presente investigación se han realizado una serie de actividades, organizadas en las siguientes etapas:

Etapas 1: Recolección de información

La etapa inicial, denominada "Recolección de Información", ha sido fundamental para garantizar la precisión y exactitud de los datos obtenidos. A continuación, se detallan los pasos seguidos durante esta etapa:

Identificación de zonas de estudio y de recorridos de los camiones recolectores: mediante un mapa de Oberá, se delimitaron los distintos barrios que componen la ciudad y los recorridos de los camiones recolectores codificados mediante la letra C que indica "camión" y su número de interno correspondiente, por ejemplo "C12062". Asimismo, para conocer los aspectos de cada barrio se reunieron las características socioeconómicas, densidad poblacional estimada y tipos de actividad predominante.

Se muestra en la Fig. 1 el análisis del recorrido del camión "C12062" mediante flechas que indican el sentido y las zonas, perteneciente al turno noche y donde cada parada está indicada con un círculo.



Fig. 1. Recorrido del camión recolector C12062.

Dependiendo de la ruta asignada para el recorrido diario, en el municipio un mismo camión puede ser utilizado en la recolección de los residuos en los diferentes turnos (mañana, tarde, noche). Se muestra en la Fig.2 el recorridototal de un mismo camión que trabajó los 3 turnos. Los recorridos

duran entre 4-6 h dependiendo del día y la estación del año, donde en fechas festivas se debe aumentar la frecuencia de recorridos (fechas cercanas a navidad y año nuevo).



Fig. 2. Recorrido del camión recolector C12062, durante los tres turnos de trabajo.

Medición de Residuos: Se determinó la cantidad de residuo por cada recorrido y para cada camión, cuya capacidad total es de 16m³ (aproximadamente 5.000 kg, dependiendo del volumen de los residuos recolectados). La municipalidad llevó a cabo pesaje de los camiones durante 7 días consecutivos y suministró los datos al equipo de investigación, quien realizó el análisis.

En la Tabla 1 se presentan el día de recorrido, los barrios y la proporción de población atendida en función a la densidad poblacional de la zona, entre otros parámetros necesarios para el análisis de GPC para cada barrio.

Tabla 1. Registro modelo utilizado para días, superficie, barrios y población por recorrido del camión recolector.

Barrios afectados al recorrido [nombre]	Superficie del Barrio [km ²]	% del barrio atendido	Densidad poblacional Promedio [hab/km ²]	Categoría de densidad poblacional por barrio
Oberá III	0,0438	3%	671	MEDIA
Villa KindGreen	0,23425	18%		MEDIA
Villa Stembaerg	0,3828	30%		ALTA
Las Lomas	0,0924	7%		ALTA
Hultgren	0,389	31%		MEDIA
Villa Svea 2	0,131	10%		BAJA

*Frecuencia de recorrido **C12062**: 3 veces por semana; lunes, miércoles y viernes (de mañana)

Etapas 2: Análisis de la información relevada

En esta etapa, los datos obtenidos mediante un conjunto de cálculos permitieron determinar, la GPC por barrio. Para ello, se emplearon los datos de cada camión y recorrido, como así también la superficie poblacional atendida mediante las fórmulas que se presentan a continuación.

La densidad poblacional promedio se calculó utilizando la ecuación (1):

$$\text{Densidad poblacional promedio} = \frac{\text{Cantidad de habitantes}}{\text{Superficie atendida}} \quad (1)$$

Para aplicar dicha ecuación ha sido necesario estimar la población de Oberá considerando los datos del censo realizado en el año 2022 [7] obteniendo una densidad poblacional media de la localidad de $671 \text{ hab}/\text{km}^2$. Luego, se calculó el “peso determinado por barrio”, que corresponde al porcentaje de residuo generado en cada barrio durante el recorrido del camión recolector. Para este cálculo, se consideró que el vehículo recolecta residuos en varios barrios en un mismo recorrido. Por lo tanto, se realizó una aproximación porcentual basada en la superficie de cada barrio, medida en kilómetros cuadrados (Tabla 1) en relación con la superficie total del recorrido. Este porcentaje se multiplicó por los kilogramos de residuos totales recolectados en el recorrido, obteniendo así el peso por recorrido específico (2).

$$\text{Peso por recorrido específico} = \frac{\text{Superficie del barrio}}{\text{Superficie del recorrido}} * \text{Peso por recorrido total} \quad (2)$$

Con base en la clasificación de barrios proporcionada por la municipalidad y en función de su densidad poblacional, se asignó un factor específico a cada caso. Este factor indica si la densidad poblacional de cada barrio se encuentra por debajo, en la media (es decir $671 \text{ hab}/\text{km}^2$) o por encima de ella (BAJA), (MEDIA) y (ALTA) respectivamente. Posteriormente, la “Población por barrio” se calculó empleando la fórmula (3):

$$\text{Población por Barrio} = \% \text{ del Barrio atendido} * \text{Densidad Poblacional por Barrio} \quad (3)$$

Por último, se calculó la GPC de cada barrio usando (4):

$$\text{GPC por Barrio} = \frac{\text{Peso por recorrido específico}}{\text{Población por Barrio}} \quad (4)$$

En la Tabla 2 se muestran datos que derivan del recorrido del camión 12062, donde se presentan los datos necesarios para calcular la generación de residuos por persona por día por barrio (kg/hab.barrio.día).

Tabla 2. RSU recolectados por barrio.

A partir de un peso por recorrido total de 4488 kg recolectados en una semana por **C12062**. Lunes, Miércoles y viernes, por la mañana.

Barrios afectados al recorrido [nombre]	Población por barrio [hab]	Peso por Recorrido específico [kg]	Días de la semana
Oberá III	29	154	7
Villa KindGreen	157	826	7
Villa Stambaerg	514	1349	7
Las Lomas	124	326	7
Hultgren	261	1371	7
Villa Svea 2	132	462	7

Para poder analizar los datos y sacar conclusiones de la GPC por zona de la ciudad, se aplicaron herramientas estadísticas.

Para ello, se procedió a calcular el promedio ($\bar{X}$) de GPC mediante la ecuación (5):

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (5)$$

Donde:

X_i : valor individual de GPC medido en cada recorrido específico.

n : número total de recorridos realizados.

Asimismo, se calculó la mediana, usando la ecuación (6):

$$Me = \frac{X_n + 1}{2} \quad (6)$$

: mediana, es decir, el valor que divide al conjunto de datos en dos partes iguales.

X_n : valores de GPC ordenados de menor a mayor.

También se calculó el valor de la Moda y la desviación estándar mediante (7):

$$s = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} \quad (7)$$

Donde:

s : desviación estándar, mide qué tan dispersos están los valores de GPC respecto al promedio.

$\bar{X}$: promedio calculado en (5)

X_i : valor individual de GPC medido en cada recorrido específico.

n : número total de recorridos realizados.

Con base en una distribución normal y utilizando los valores estadísticos, se depuraron los datos para alcanzar un nivel de confianza del 95 %. Para ello, se eliminaron los datos atípicos, lo que permitió obtener el porcentaje de confianza deseado. Los datos depurados fueron empleados en la elaboración de las conclusiones del trabajo.

3. Resultados y discusión

3.1 Generación de residuo per cápita “cuantitativa” general

A partir de los datos proporcionados por el municipio sobre los kilogramos de residuos generados en un mes (promedio del año 2023), que ascendieron a 1.100.000 kg mensuales, y considerando los 75.000 habitantes atendidos [8] con una eficiencia de recolección 90%, se calculó la generación de residuos por habitante por día general ha sido mediante (8).

$$GPC_{general} = \frac{\text{Promedio de residuos transportados}}{\text{Población atendida} * \text{Días de recolección en el mes}} \quad (8)$$

Como resultado de lo expuesto, la GPC general basada en un cálculo de aproximación cuantitativa es de 0,66(kg/hab por día). Es importante señalar que las estimaciones de kilogramos generados, utilizadas para este cálculo, provienen de un promedio de datos recolectados durante el año 2023 por el municipio de Oberá, abarcando una estación del año correspondiente a primavera (septiembre, octubre y noviembre) [9,10].

3.2 Generación de residuo per cápita por barrio

A partir de la información proporcionada por el municipio, detallada en la Tabla 1, y la ruta de cada recorrido de los vehículos recolectores en el mapa urbano, presentada en la Fig. 1, se estimaron los kilogramos de residuos generados por habitante en cada barrio, considerando que un mismo recorrido abarca varios barrios. Asimismo, dado que el recolector cubre solo un porcentaje de cada barrio, se calculó la porción (%) de población atendida en cada recorrido del camión recolector. Una vez calculada la GPC para cada barrio, se procedió a consolidar la información y a depurar los datos. Para ello, se calcularon los principales estadísticos: media, moda, mediana y desviación estándar. Los valores resultantes sin depurar los datos se presentan en la Tabla 3:

Tabla 3. Estadísticos resultantes del cálculo antes de la depuración de los datos.

Estadístico	Valor de GPC
Media	0,6
Mediana	0,38
Moda	0,24
Desviación estándar	0,61

Se consideró una distribución normal y un nivel de confianza del 95 % en los datos recolectados, por lo que se identificaron y eliminaron los posibles datos atípicos. En este proceso, se descartó el 5 % de los datos extremos, distribuidos equitativamente entre ambas colas de la distribución (2,5 % en cada extremo). Dado que la localidad analizada cuenta con 58 barrios, el 5 % equivale a la GPC de aproximadamente 3 barrios. Considerando esta cantidad, se evaluaron los datos de GPC por barrio y se eliminaron dos valores superiores, debido a que eran significativamente altos, y un valor inferior, con el fin de mejorar la precisión de la información recopilada.

Es importante destacar que, en este tipo de casos, es común realizar una limpieza de los datos para mejorar la calidad de la información, eliminando valores extremos que podrían distorsionar los análisis estadísticos. Este procedimiento reduce el impacto de datos atípicos que podrían conducir a conclusiones erróneas. En este caso particular, al eliminar el 5 % de los datos (aproximadamente un 2,5 % en cada cola de la distribución normal), se logra trabajar con un conjunto de datos que cumple con un nivel de confianza del 95 %. Los resultados de GPC obtenidos y depurados se presentan en la Tabla 4.

Tabla 4: GPC resultantes por Barrio de Oberá.

Barrios de Oberá		Generación de Residuos per cápita [kg/hab.barrio.día]	Barrios de Oberá		Generación de Residuos per cápita [kg/hab.barrio.día]
1	Villa KindGreen	0,60	30	Barrio Ecológico	0,82
2	Villa Stemberg	0,39	31	Yerbal Viejo	0,12
3	Villa Svea 2	0,28	32	Barrio Docente	0,12
4	Villa Svea 1	0,29	33	Judiciales	0,24

5	Villa Rufft	0,31	34	Oberá VI	0,24
6	Villa Cristen	0,31	35	Oberá VII	0,24
7	Villa Cristen 2	0,41	36	Oberá V	0,24
8	Loma Pora	0,36	37	Villa Lindstron	0,24
9	Sixten Vick	0,40	38	Hedvig Andersson	0,24
10	Villa Falk	1,02	39	Odd Bothner	0,24
11	Virgen de Guadalupe	0,11	40	Otto Erik Sand	0,12
12	Punta Alta	0,29	41	Villa Shustter	0,48
13	Villa Londin	0,41	42	Caballeriza	0,22
14	Villa Erasmie	0,89	43	Oberá V	0,52
15	100 hectáreas	0,45	44	Barrio Ex 30 Viviendas	0,52
16	Villa Barreyro 2	0,32	45	Esperanza	1,01
17	Las Lomas	0,43	46	Pueblo Salto	0,35
18	Villa del Parque	1,52	47	Copisa	0,21
19	Microcentro	0,62	48	San José	0,10
20	Hultgren	0,82	49	San Miguel	0,10
21	Oberá III	0,75	50	Villa Blanquita	0,21
22	Las Lomas	0,38	51	Molinas	1,01
23	Villa Mosquere	0,24	52	Villa Shustter	1,07
24	Primeros colonos	0,30	53	Jardín Botánico	1,07
25	Villa Industrial	0,61	54	Villa Marttos	0,00
26	Villa Lorr	0,15	55	Tres Esquinas	0,38
27	Sto Cabral	0,82	56	Villa Lutz	0,37
28	Las Palmas	0,20	57	KM 0	1,23
29	El progreso	0,82	58	Villa Kleiven	0,31

Luego de la depuración de los datos, se procedió a calcular los estadísticos de la muestra, resumidos en la Tabla 5:

Tabla 5. Estadísticos resultantes del cálculo de GPC con datos depurados.

Estadístico	Valor de GPC
Media	0,51
Mediana	0,37
Moda	0,24
Desviación estándar	0,41

A partir del análisis de la información recopilada en la Tabla 4, se identificaron los barrios que generan la mayor cantidad de residuos en Oberá (Fig. 3).

En un planeamiento de nuevos recorridos, los barrios que más residuos generan (Fig. 3) deberán tener mayor frecuencia de recolección (corto plazo), además de proyección estratégica tal como eco-puntos para fomentar la economía circular (mediano plazo). La generación de residuos per cápita de Oberá resultó ser 0,51 si se toma en cuenta el significativo análisis cualitativo realizado, y 0,66 si se toman datos generales. Estas cifras pueden ser útiles para la toma de decisiones en políticas públicas relacionadas con la gestión de residuos, tales como la implementación de programas de reciclaje, compostaje o reducción de residuos por zonas o barrios con valores por arriba del promedio.

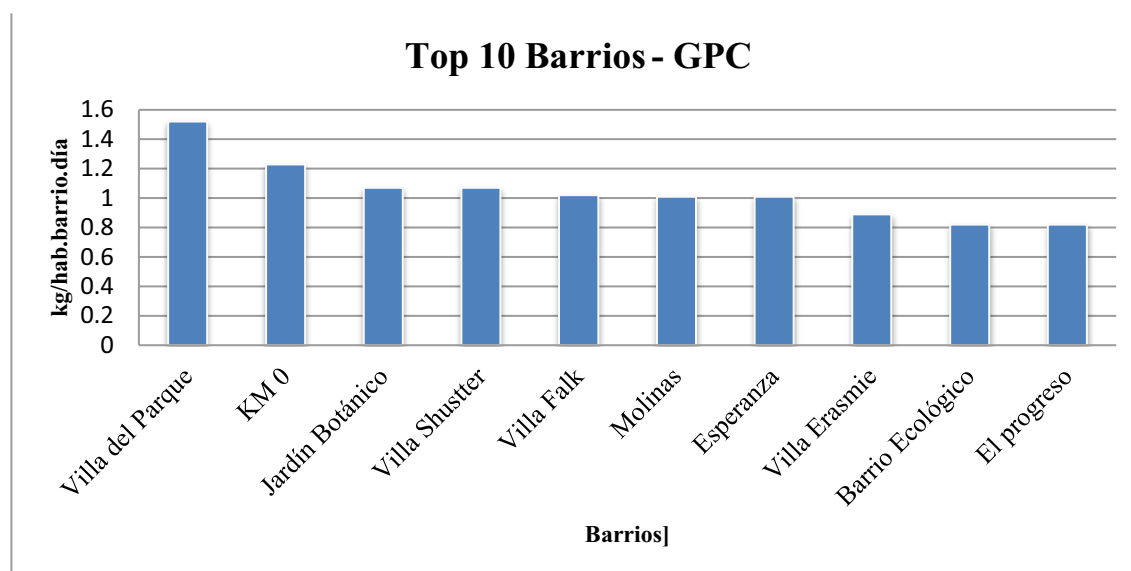


Fig.3. Barrios resultantes como mayores generadores de residuos de Oberá.

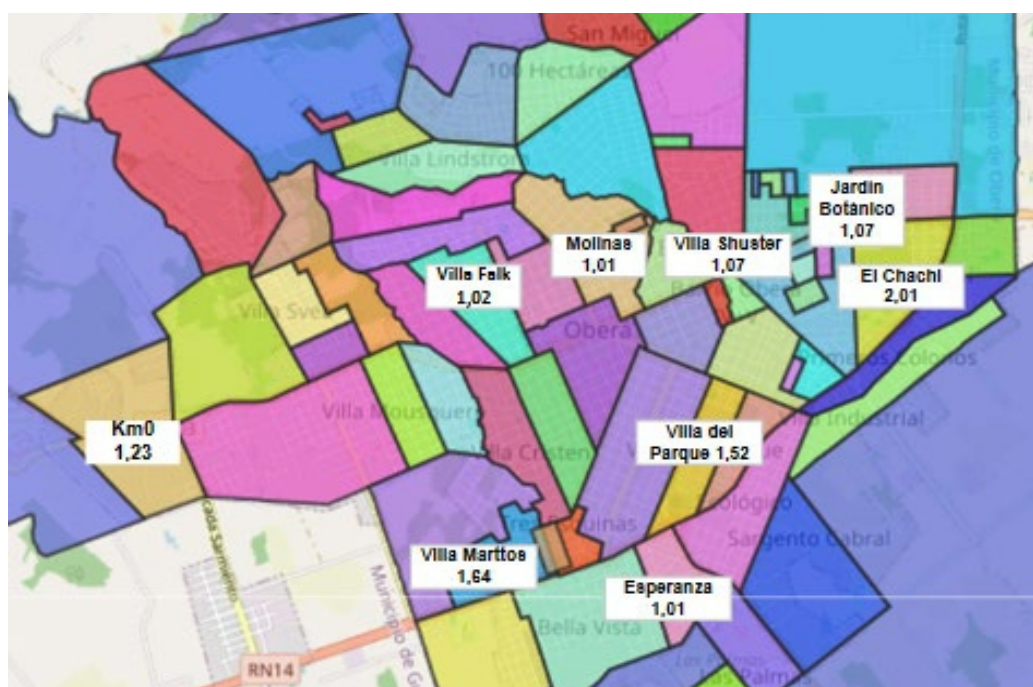


Fig.4. Barrios con mayor GPC.

Con base en los resultados, es importante destacar que la baja generación de residuos per cápita calculada en barrios como "San José" y "San Miguel", ambos con alta densidad poblacional, se debe a las dificultades de acceso que enfrentan los camiones recolectores en ciertas áreas de la zona mencionada. Esto implica que el valor cualitativo de la GPC no refleja necesariamente una menor producción de residuos, sino más bien las restricciones de acceso ocasionadas por la infraestructura vial limitada, el crecimiento no planificado de la población y la presencia de asentamientos en la zona.

La gestión de residuos, como tema de creciente relevancia, representa una problemática que requiere cambios en los comportamientos sociales y estatales [11]. En la Fig. 4 se identifican los barrios que presentan una mayor generación de residuos.

Es importante reconocer que, a pesar de los esfuerzos por asegurar la precisión y representatividad de los datos, existen inherentemente ciertas limitaciones asociadas con este tipo de estudio. Las

mediciones de residuos están sujetos a variaciones debido a factores como la estacionalidad, eventos especiales, y diferencias en los hábitos individuales y familiares. Por lo tanto, los datos presentados forman parte de un primer estudio del tema en dicha localidad y son aproximaciones [12,13,14]. Según los datos obtenidos en el estudio descrito (GPC 2023: 0,51) y en comparación al informe del año 2010 (GPC: 0,62), aunque en Oberá la población ha crecido, se denota una disminución de la generación de los residuos. Ello puede ser consecuencia de que desde hace unos años existen empresas que compran materiales residuales en Oberá y por ello, un alto porcentaje de la basura generada ya no termina en el camión recolector [15,16] sin embargo, es residuo que la sociedad genera. Sin duda hay varios actores y factores implicados en la GIRSU de una localidad [4].

Los análisis descriptivos y comparativos realizados proporcionaron información relevante para el diseño de políticas públicas y estrategias de gestión de residuos más efectivas especialmente para los recorridos de camiones recolectores de Oberá [5].

El trabajo realizado en conjunto entre la Facultad de Ingeniería Oberá de la Universidad Nacional de Misiones y la Municipalidad de la Ciudad de Oberá (Fig. 5), se presentó frente a las autoridades municipales y provinciales para dar a conocer el diagnóstico realizado que ha sido tomado como punto de partida para generar estrategias municipales sobre GIRSU donde se ha incluido optimizar el servicio de recolección de residuos sólidos urbanos, la separación el origen, el reciclaje, ecopuntos, puntos limpios y otras prácticas posibles de economía circular [17].

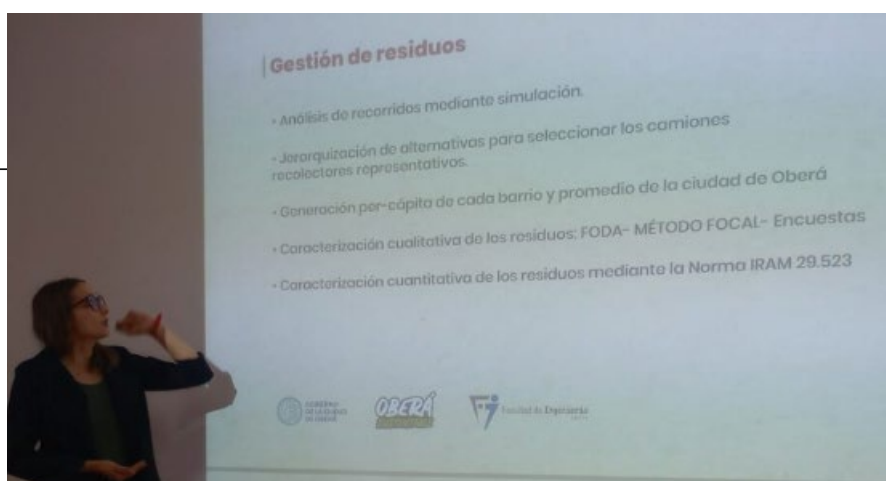


Fig.5. Presentación de resultados.

4. Conclusiones

El estudio proporcionó información sobre la GPC de RSU de la ciudad de Oberá, específicamente para residuos de tipo domiciliarios que terminan en camiones recolectores municipales y cuyo destino final es el relleno sanitario.

Los resultados evidenciaron que el nivel socioeconómico incide directamente en la generación de residuos: a mayor clase social, mayor consumo y, por ende, mayor producción de residuos. Esta observación se sustenta tanto en la información proporcionada por el área de “Acción Social” de la municipalidad como en los datos cualitativos obtenidos en el estudio.

Este tipo de análisis ofrece una primera aproximación para entender la generación de residuos per cápita en un municipio y permite identificar tendencias y áreas con necesidades específicas, como base para formular nuevos recorridos y frecuencias del camión recolector que aporten eficiencia a la gestión de residuos.

Como línea futura de investigación, se propone incorporar datos sobre la cantidad de materiales residuales recolectados por las empresas que compran residuos en Oberá. Esto permitirá reunir información para una caracterización más precisa y completa del manejo de residuos en la localidad.

Agradecimientos

Este trabajo ha sido posible gracias al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) quien financió parte de la investigación. Agradecemos profundamente a la municipalidad de la ciudad de Oberá por su colaboración y compromiso con la Facultad de Ingeniería de Oberá, de la Universidad Nacional de Misiones para realizar la investigación.

Referencias

- [1] Tchobanoglous George, K. F. (2002). *Handbook of solid waste management*. New York: McGraw-Hill. doi:10.1036/0071356231
- [2] Estrategia Nacional para la gestión integral de los residuos sólidos urbanos: ENGIRSU (2005). Disponible en: <https://www.argentina.gob.ar/interior/ambiente/control/estrategia-nacional> (27.07.2024)
- [3] Alagoz, A. Z. & Kocasoy, G. (2008). *Improvement and modification of the routing system for the health-care waste collection and transportation in Istanbul*. *Journal of Waste Management*, 28, 1461-1471.
- [4] Niezwida, S. R., Jusczysszyn, J. C. M., & Gavazzo, G. B. (2023). *Revisión bibliográfica sobre los residuos sólidos urbanos*. *InGenio Journal: La revista de la Facultad de Ciencias de la Ingeniería de la UTEQ*, 6(2), 30-39. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=9336139> (19.03.2025).
- [5] Oberá Sustentable. Sitio oficial. Disponible en: <https://obera.gob.ar/categoria/vinculacion-y-desarrollo-economico/obera-sustentable/> (20.07.2024)
- [6] Argentina.gob. Página oficial. Datos oficiales de un informe del año 2010. Disponible en: https://www.argentina.gob.ar/sites/default/files/provincia_de_misiones.pdf (27.07.2024)
- [7] Instituto Provincial de Estadísticas y Censos: IPEC. Página oficial disponible en: <https://www.ipec.misiones.gov.ar/> (27.07.2024)
- [8] INDEC. Instituto Nacional de estadísticas y censos. Recuperado de: <https://obera.gob.ar/categoria/vinculacion-y-desarrollo-economico/obera-sustentable/> (19.07.2024)
- [9] Plan GIRSU Argentina. Disponible en: <https://www.argentina.gob.ar/interior/ambiente/erradicacion-de-basurales> (15.07.2024)
- [10] Gobierno de Oberá. Informe de la SECRETARIA DE FINANZAS Y DESARROLLO ECONOMICO DIRECCION DE OBRAS Y SERVICIOS PUBLICOS: Análisis de Costos y variables de Transporte RSU.
- [11] Niezwida, S. R., & Quemel, R. P. (2023). *Una mirada sobre la gestión de residuos sólidos en municipios periféricos de Pará, Brasil y de Misiones, Argentina*. *JIDeTEV*, (13). Disponible en: <https://autoresjidetev.fio.unam.edu.ar/index.php/jidetev/article/view/316> (19.03.2025).
- [13] Jansat, M. (2024). *Gestión de residuos: Método focal, Open LCA y herramientas de ingeniería*. [Trabajo de práctica final supervisada, Universidad Nacional de Misiones]. Biblioteca Regional Oberá. Oberá, Misiones, Argentina.
- [14] Kaczynski, F. (2024). *Gestión de Residuos Sólidos Urbanos: Simulación y conclusión a partir del sistema actual y un escenario posible*. [Trabajo de práctica final supervisada, Universidad Nacional de Misiones]. Biblioteca Regional Oberá. Oberá, Misiones, Argentina.
- [15] Stoffel, M. (2024). *Gestión de Residuos Sólidos Urbanos en Oberá*. [Trabajo de práctica final supervisada, Universidad Nacional de Misiones]. Biblioteca Regional Oberá. Oberá, Misiones, Argentina.
- [16] Chiplly reciclajes. Página oficial. Disponible en: <https://chiptyreciclaje.com/> (20.07.2024)
- [17] Canal doce Misiones. *Oberá busca optimizar el servicio de recolección de residuos sólidos urbanos*. 24 de abril de 2024. <https://www.canal12misiones.com/noticias-de-misiones/sociedad/residuos-solidos-urbanos> (06.04.2025)

Afundamentos de Tensão em Sistema Elétrico de Distribuição Considerando Curto-circuitos e Geração Distribuída

Felipe Hernández García^{1*} , Mario Orlando Oliveira² , Roberto Chouhy Leborgne³ ,
Arturo Suman Bretas⁴ , Roberto José Cabral² 

¹ Universidade Federal do Rio Grande (FURG), Rio Grande, Rio Grande do Sul, Brasil.

² Universidad Nacional de Misiones (UNaM). Misiones, Argentina.

³ Universidade Federal do Rio Grande do Sul (UFRGS). Porto Alegre, Rio Grande do Sul, Brasil.

⁴ Pacific Northwest National Laboratory (PNNL). Richland, Estados Unidos de América.

e-mails: felipehernandez@furg.br, mario.oliveira@fio.unam.edu.ar, roberto.leborgne@ufrgs.br,
cabral@fio.unam.edu.ar

Resumo

Sistemas de Distribuição de Energia Elétrica estão constantemente expostos a perturbações que afetam a Qualidade da Energia fornecida. Dentre estas perturbações os afundamentos de tensão são os de maior frequência de ocorrência e os que mais impacto provocam no funcionamento de equipamentos modernos. Neste trabalho são estudados afundamentos de tensão a partir dos resultados de simulações de faltas realizadas em um sistema radial de 14 barras, operado com e sem Geração Distribuída. O sistema foi modelado no programa ATPDraw. Os resultados de correntes obtidos com o programa ATP serviram de referência para comparar resultados obtidos em uma aplicação em Matlab desenvolvida utilizando os métodos de Componentes Simétricas e de Componentes de Fases. Também servem de referência no desenvolvimento de um programa computacional próprio. O trabalho mostra que com o Método de Componentes Simétricas os resultados se afastam da referência quando se incrementa o desequilíbrio do sistema. As tensões resultantes do ATP permitiram a análise de afundamentos de tensão. A análise foi realizada a partir da elaboração de matrizes booleanas que consideram o estado de todas as barras. Dois índices relacionados com os afundamentos de tensão são propostos no trabalho, um que caracteriza o sistema como um todo e outro que utiliza como referência a barra onde esteja localizado um gerador independente. A partir destas matrizes e índices mapas de contorno e com cores são obtidos.

Palavras Chave – Afundamentos de tensão, componentes de fase, componentes simétricas, curto circuito, geração distribuída, sistema elétrico de distribuição.

Abstract

Electrical Power Distribution Systems are constantly exposed to disturbances that affect the quality of the power supplied. Among these disturbances, voltage sags are the most frequent and have the greatest impact on the operation of modern equipment. In this work, voltage sags are studied based on the results of fault simulations performed on a 14-bus radial system, operated with and without Distributed Generation. The system was modeled in the ATPDraw program. The current results obtained with the ATP program served as a reference for comparing results obtained in a Matlab application developed using the Symmetrical Components and Phase Components methods. They also serve as a reference for the development of a computer program. The work shows that with the Symmetrical Components Method, the results deviate from the reference as the system imbalance increases. The voltages resulting from the ATP allowed the analysis of voltage sags. The analysis was performed by developing Boolean matrices that consider the state of all bars. Two indexes related to voltage sags are proposed in this paper: one that characterizes the system as a whole and another that uses the node where an independent generator is located as reference. From these matrices and indexes, contour and color-coded maps are obtained.

Keywords – Distributed generation, distribution systems, phase components, short-circuits, symmetrical components, voltage sags.

1. Introdução

A Qualidade de Energia Elétrica (QEE) é de importância vital para o funcionamento eficiente, seguro e contínuo de equipamentos e sistemas que dependem do fornecimento de energia. A QEE

refere-se ao grau em que os parâmetros da energia – como tensão e frequência – permanecem dentro dos limites aceitáveis para garantir o desempenho adequado daqueles equipamentos.

Dispositivos eletrônicos modernos manifestam comportamentos fortemente não lineares provocando distorções no sinal elétrico, afetando a QEE fornecida. Exemplos destas distorções podem ser comprovados em trabalhos que analisam sistemas ou situações reais [1-3], assim como no ambiente computacional [4-6].

As distorções podem provocar aumento do consumo e até mesmo falhas em equipamentos eletroeletrônicos [7], contribuindo à redução da vida útil dos mesmos [8]. No ambiente não industrial o mal funcionamento dos equipamentos pode paralisar serviços, ou provocar perda de trabalhos e dados. Já no âmbito industrial, em processos de produção automatizados, com cargas cada vez mais exigentes quanto à QEE, desligam-se equipamentos sensíveis, paralisando-se etapas e até cadeias produtivas inteiras, com a consequente perda de materiais e produtos, e aumento dos custos [9, 10].

Dentre os distúrbios mais comuns observados nas redes elétricas encontram-se sobretensões, subtensões, harmônicas, *flickers*, surtos e afundamentos de tensão [11], sendo estes últimos os de maior frequência de ocorrência [12-14].

Os afundamentos estão caracterizados pela redução da tensão junto aos consumidores em valores que oscilam entre 10 e 90% do valor nominal, com duração entre 0,5 e 60 segundos [15]. Por causa da prevalência deste distúrbio frente aos outros e pelo impacto que provocam no funcionamento de equipamentos, os afundamentos têm recebido especial atenção de pesquisadores e de empresas do setor elétrico, seja no aspecto teórico, através do estudo do fenômeno [16, 17], seja na área tecnológica, com a instalação de dispositivos mitigadores [18,19].

Os afundamentos podem ocorrer por razões operacionais dos Sistemas Elétricos de Potência (SEP), como durante o arranque de grandes motores e a energização de transformadores e blocos de cargas. Também ocorrem por falhas em equipamentos, ou ainda por causas naturais, como os curto-circuitos motivados por descargas elétricas ou a queda de árvores sobre os condutores das redes [2, 20]. Curto-circuitos podem ocorrer envolvendo as três fases, ou duas delas, entre si ou com o neutro ou a terra. No entanto, as estatísticas revelam que os curto-circuitos monofásicos são os que se manifestam com maior frequência [12, 14].

A análise de curto-circuito é um tema relevante no estudo dos Sistemas de Distribuição de Energia Elétrica (SDEE), seja nas etapas de planejamento e expansão, seja durante a operação. Destas análises resultam, por exemplo, informações para os estudos de proteção [21, 22], para municiar os métodos de localização de faltas [23, 24] e para os estudos de QEE. Justamente, as abordagens mais comuns no estudo dos afundamentos de tensão envolvem métodos de análise de curto-circuito.

Os métodos das Componentes Simétricas (MCS) [25, 26] e das Componentes de Fase (MCF) [27, 28] continuam sendo os de maior destaque na análise de curto-circuitos. O MCS encontra sua maior utilização em sistemas trifásicos equilibrados, como os Sistemas de Transmissão de Energia Elétrica (STEE). Os SDEEs no entanto, são caracterizados pela assimetria na distribuição dos condutores e pela não transposição [29], o qual favorece o desequilíbrio. Soma-se a isto a presença de cargas monofásicas e bifásicas nas redes. Sob estas condições vários autores comentam resultados imprecisos com a utilização do MCS, particularmente quando tratam de faltas monofásicas [30, 31]. O MCF, por sua vez, revela-se como alternativa ao MCS levando em conta as possíveis assimetrias. Ao trabalhar com as fases o MCF oferece resultados para uma ampla gama de estudos dos SDEEs, contudo, sua complexidade de implementação pode ser significativa, conforme aumenta o tamanho do SDEE.

Quanto à regulamentação relacionada com a QEE, diversas normas estabelecem limites para a intensidade ou a duração dos distúrbios e inclusive para o tempo de recuperação dos mesmos [32-34]. As normas também definem índices para mensurar a performance do fornecimento, os quais contribuem para garantir que a energia atenda aos padrões de desempenho esperados.

Para afundamentos de tensão, a normativa internacional aproveita o *Índice de Frequência Média de Variação RMS no Sistema SARFI*, quantificando as reduções da tensão abaixo de determinados limites [35]. Um outro índice de alcance internacional é o *Índice de Severidade de Afundamentos de Tensão, SSI*, que considera a sensibilidade dos equipamentos perante afundamentos [36] e que se presta para que muitos pesquisadores definam novos índices nos seus trabalhos.

Por outro lado, é cada vez mais comum a presença de unidades de geração independente nas redes dos SDEEs, contribuindo para uma Geração Distribuída (GD), termo acunhado por [37]. Embora a GD possa trazer algum tipo de dificuldade operacional para as empresas responsáveis pelas redes [38, 39] é previsível a continuidade da sua expansão, sobretudo, por envolver energias renováveis, associadas à sustentabilidade e à redução de danos ambientais. Quanto à QEE, a GD pode desempenhar um papel importante, atuando como estabilizadora da tensão da rede, melhorando os parâmetros desta e inclusive a performance perante eventos causadores dos afundamentos [40, 41].

Este trabalho é dividido em duas partes. Inicialmente, são analisados os resultados obtidos através dos métodos de análise de curto circuito acima citados, aplicados a um SDEE de 14 barras de configuração radial. O SDEE foi estudado com e sem a presença da GD. Para o estudo o SDEE foi modelado no programa ATPDraw [42] e os métodos de análise foram implementados no programa Matlab [43]. O trabalho mostra a formulação matemática envolvida com estes métodos e o fluxograma de um dos algoritmos realizados. Ressalta-se que estes resultados estão sendo obtidos no marco do desenvolvimento de um programa computacional próprio para estudo de SEPs, com o Matlab e o ATP/EMTP [44] sendo utilizados para aferir os resultados dos subprogramas e algoritmos em desenvolvimento.

Na segunda parte do trabalho são analisados afundamentos de tensão, a partir do comportamento do SDEE perante os curto-circuitos simulados.

1.1. *Trabalhos precedentes com matrizes booleanas e mapas de contorno*

Análises nos SEPs envolvem uma quantidade considerável de variáveis inter-relacionadas e grande volume de dados. Uma forma de condensar os resultados é a representação matricial.

O termo *matriz de afundamentos de tensão* foi introduzido por [45] ao analisar correntes de falta simétricas. As linhas da matriz mostram a queda de tensão nas barras após faltas em uma delas (esta última localizada na diagonal principal). Através das colunas é possível identificar a situação particular de cada barra perante curto-circuitos nas outras barras. Metodologia semelhante foi utilizada por outros autores, por exemplo, em [46] organizaram em matrizes as tensões e correntes nas barras de um SDEE.

Os trabalhos de [47],[48] e [49] podem ter sido os precursores da representação gráfica. Os autores traçam curvas sobre mapas físicos ou de unifilares cercando regiões nas quais a tensão diminuirá até determinado limite quando faltas ocorram nas barras ou linhas elétricas das mesmas. Essas regiões são chamadas de áreas expostas.

Na mesma linha do tratamento gráfico em [50] acreditam ter desenvolvido o mapa de afundamentos de tensão. Os mapas são elaborados sobre o unifilar, para faixas de tensões com intervalo de 0,1 p.u. Em cada mapa são traçadas curvas que contornam as barras conforme o número

de afundamentos que sofreriam, em base anual. O trabalho foi desenvolvido para uma instalação industrial, estimando-se o número de afundamentos que sofreriam os equipamentos em uma determinada barra de acordo com os limites de funcionamento dos mesmos. Por exemplo, para um equipamento com tolerância entre 0,6 e 0,9 p.u., devem ser utilizados os mapas nessa faixa de tensões. Considerando a barra onde está colocado o equipamento se verifica se a mesma está contornada por alguma linha. No caso, toma-se o número de afundamentos correspondentes à linha. Finalmente se somam todos os afundamentos.

A representação matricial com elementos binários foi utilizada por [51] para classificar diferentes tipos de distúrbios. Já [52] a utilizaram para codificar os distúrbios.

Um exemplo de representação binária relacionada com afundamentos encontramos em [53], onde determina-se o número mínimo de medidores que devem ser instalados nos SDEEs estudados para que faltas que provoquem quedas de tensão abaixo de determinado limite sejam registradas por pelo menos um dos medidores. O propósito do estudo é também econômico, para reduzir os custos associados com os elevados preços dos medidores.

Uma matriz é elaborada, com tantas linhas quanto barras tenha o SDEE sob estudo. Cada barra contará com um medidor, de modo que é possível associar o número de linhas da matriz ao número de medidores. Faltas são simuladas, mostrando cada linha da matriz o estado do medidor correspondente respeito a observar ou não tensões abaixo do limite pré-definido. Os elementos da matriz não são formados por apenas um número, mas por um vetor de dimensão 1x4, pois a distância entre a barra do medidor e cada uma das outras barras é dividida em 4 seções, nas quais são provocadas as faltas. Cada elemento deste vetor será “1” ou “0” em correspondência de se o medidor percebe ou não a falta. Finalmente, o problema de otimização é resolvido e o número mínimo de medidores é determinado.

Essas representações são “integradas” em matrizes não numéricas, mas de tonalidades de cores. Matrizes destas características são de ordem n , com as linhas representando a cada uma das barras. As colunas coloridas revelam a tensão residual em cada barra, de 0 a 1 p.u., perante faltas que aconteçam nas outras barras [54, 55]. Assim como os mapas, estas matrizes constituem um instrumento de grande valia para a rápida identificação da situação nas diferentes barras.

Neste trabalho são utilizadas as representações matricial e binária. Cada elemento da matriz elaborada indicará se uma barra atinge (“1”), ou não (“0”), níveis de tensão abaixo de 0,9 p.u., o que seria indicativo de afundamento de tensão. A análise é realizada utilizando um SDEE de 14 barras de configuração radial, quando o mesmo opera com e sem GD. A partir destas matrizes foi possível a definição de dois índices, $IIMS_{\phi}$ e $IIMB_{\phi}$, associados à imunidade do SDEE perante o distúrbio.

O índice $IIMS_{\phi}$ está relacionado com o SDEE como um todo e pode contribuir para comparar a situação do sistema em diferentes condições, como as que resultam dos projetos de reconfiguração [56, 57] ou da introdução da GD [58, 59].

O índice $IIMB_{\phi}$ é definido para SDEEs com GD, onde a barra que tem GD é utilizada como referência. Este índice permite a elaboração de informação visual, utilizando cores para diferenciar zonas ou regiões em torno das barras.

Nas próximas seções são formulados os métodos de análise de curto-circuito, logo após apresentamos um caso de estudo, seguido da análise dos resultados das simulações realizadas e finalmente são apresentadas as conclusões mais relevantes.

2. Fundamentação Teórica

A equação (1) é a representação matricial do sistema de equações que surge ao aplicar o método das tensões nodais [60] a um SEP constituído por n barras:

$$I = Y_{barras} V \quad (1)$$

Os vetores coluna V e I representam as tensões e correntes das n barras, respectivamente, e Y_{barras} é a chamada matriz de admitâncias, a qual é de ordem n , com suas linhas representando cada uma das barras do SEP. Os elementos Y_{barra_ij} representam as admitâncias das linhas elétricas entre duas barras, i e j , por exemplo. Já os elementos da diagonal principal representam as admitâncias próprias de cada barra, determinados considerando as admitâncias de todos os componentes elétricos conectados às mesmas [61]. Para uma barra i , por exemplo:

$$Y_{barra_ii} = \sum_{k=1}^n Y_{ik} = \sum_{k=1}^n 1/Z_{ik} \quad (2)$$

com k variando conforme o número de componentes conectados à barra.

2.1. Métodos de análise

Em análises de faltas, em lugar de Y_{barras} é comum utilizar a matriz impedância de barras, Z_{barras} , obtida como:

$$Z_{barras} = Y_{barras}^{-1} = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1n} \\ Z_{21} & Z_{22} & \dots & Z_{2n} \\ \cdot & \cdot & \dots & \cdot \\ Z_{n1} & Z_{n2} & \dots & Z_{nn} \end{bmatrix} \quad (3)$$

Em SEPs equilibrados, faltas do tipo trifásico-terra em uma barra i produzirão correntes de falta de iguais magnitudes nas três fases, determinadas como:

$$(4)$$

onde V_{pf} representa a tensão pré-falta na barra e z_{ii} é o elemento ii da diagonal principal de Z_{barras} .

Entretanto, as faltas envolvendo apenas uma das fases com a terra são as de maior frequência de ocorrência nos SEPs e são as estudadas neste trabalho, e nestes casos, a determinação das correntes de faltas não é tão simples.

Os métodos de análise comentados a seguir modelam os diferentes componentes elétricos através de matrizes quadradas, de ordem 3, nas quais os elementos da diagonal principal representam às admitâncias (ou impedâncias) próprias de cada fase e os elementos fora desta representam as admitâncias (ou impedâncias) comuns, entre elas.

A equação (5), por exemplo, mostra a matriz de impedâncias de uma linha elétrica. Neste caso, o elemento kk representa a impedância própria de uma fase k e os elementos kq representam as impedâncias mútuas entre as fases k e q :

$$Z_L = \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} \quad (5)$$

Cabe destacar que quando a linha elétrica está equilibrada, isto é, transposta e constituída por condutores iguais, os elementos da diagonal principal são iguais e os elementos fora desta são todos nulos. Transformadores trifásicos e cargas balanceadas também têm matrizes diagonais, dessa característica.

Para simular correntes de falta monofásicas através do MCF [62, 27] elabora-se uma matriz de impedância de barras com a estrutura de (6), na qual cada elemento é uma submatriz de ordem 3.

$$Z_{barras} = \begin{bmatrix} Z_{11} & Z_{L12} & \dots & Z_{L1n} \\ Z_{L21} & Z_{22} & \dots & Z_{L2n} \\ \cdot & \cdot & \dots & \cdot \\ Z_{Ln1} & Z_{Ln2} & \dots & Z_{nn} \end{bmatrix} \quad (6)$$

Os elementos Z_{ii} representam as matrizes de impedância das barras, constituídas pela soma das matrizes de todos os componentes conectados às mesmas. Representam a matriz de impedância de Thévenin de cada barra. Os elementos Z_{Lij} representam as matrizes das linhas elétricas entre as barras. Como Z_{barras} é simétrica, $Z_{Lij} = Z_{Lji}$. Em caso de não existir conexão entre duas barras $Z_{Lij} = \mathbf{0}_{3 \times 3}$.

A corrente de falta na fase k de uma barra i será determinada como:

$$I_{falta_ki} = \frac{V_{pf_ki}}{Z_{qq_ki}} \quad (7)$$

onde V_{pf} é a tensão pré-falta na fase z e z_{qq} é a impedância equivalente de Thévenin vista desde essa fase, representada pelo elemento qq da submatriz ii da diagonal principal de (6).

O MCS é baseado nas relações de Fortescue [25], as quais permitem transformar um conjunto de fasores trifásicos em três conjuntos de fasores também trifásicos, conhecidos como componentes de sequência: positiva, negativa e zero. Os fasores em cada conjunto têm a mesma amplitude e nos dois primeiros conjuntos estão defasados em 120° . No terceiro conjunto os fasores não estão defasados.

Considerando os três conjuntos uma corrente trifásica é representada como:

$$\begin{aligned} I_a &= I_{a0} + I_{a1} + I_{a2} = I_{a0} + I_{a1} + I_{a2} \\ I_b &= I_{b0} + I_{b1} + I_{b2} = I_{a0} + \alpha^2 I_{a1} + \alpha I_{a2} \\ I_c &= I_{c0} + I_{c1} + I_{c2} = I_{a0} + \alpha I_{a1} + \alpha^2 I_{a2} \end{aligned} \quad (8)$$

Matricialmente:

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha^2 & \alpha \\ 1 & \alpha & \alpha^2 \end{bmatrix} \begin{bmatrix} I_{a0} \\ I_{a1} \\ I_{a2} \end{bmatrix} \quad (9)$$

ou na forma compacta:

$$I_{abc} = A I_{012} \quad (10)$$

onde I_{abc} e I_{012} representam os vetores das correntes de fase e em componentes simétricas, respectivamente, e A é a matriz de operadores α ou matriz de transformação, com α definido como:

$$\alpha = e^{j^{2/3}} = \frac{-1}{2} + j \frac{\sqrt{3}}{2} = 1 \angle 120^\circ \quad (11)$$

A matriz de transformação também pode ser utilizada para determinar as componentes de sequência das impedâncias dos componentes dos SEPs, através da expressão:

$$Z_{012} = A^{-1} Z_{abc} A \quad (12)$$

No caso de transformadores, geradores e cargas equilibradas, como resultado de (12) a matriz de sequência destas impedâncias é também uma matriz diagonal:

$$Z_{012} = \begin{bmatrix} z_0 & 0 & 0 \\ 0 & z_1 & 0 \\ 0 & 0 & z_2 \end{bmatrix} \quad (13)$$

onde z_0 , z_1 e z_2 representam as impedâncias de sequência zero, positiva e negativa, respectivamente, e são iguais.

No caso das linhas elétricas, quando as impedâncias próprias são iguais (conseguido utilizando-se condutores iguais), assim como também as mútuas (conseguido com a transposição), a matriz de impedância de sequências é também uma matriz diagonal, com:

$$z_0 = z_{ii} + 2z_{ij} \quad (14)$$

$$z_2 = z_1 = z_{ii} - z_{ij} \quad (15)$$

Entretanto, em SDEEs desequilibrados a matriz de impedância das linhas (5) não é simétrica, e menos ainda diagonal. Um procedimento comum para aproximar essas matrizes da condição de simetria consiste em transformar seus elementos para valores médios [63]:

$$z_{ii} = \frac{1}{3}(z_{aa} + z_{bb} + z_{cc}) \quad (16)$$

$$z_{ij} = \frac{1}{3}(z_{ab} + z_{bc} + z_{ca}) \quad (17)$$

A aplicação de (12) à nova matriz produz uma matriz de sequência de impedâncias com a estrutura de (13), corretamente desacoplada.

Desta forma, todos os componentes do SEP terão matrizes diagonais, sem elementos mútuos, permitindo elaborar três redes de sequência independentes, já que não existe acoplamento entre as mesmas.

A determinação da matriz de impedância de barras para cada uma das redes produzirá impedâncias equivalentes de Thévenin para cada uma delas, obtendo-se a corrente de curto-circuito (falta) em uma barra i através da seguinte equação [29]:

$$I_{falta_i} = \frac{3V_{pf_i}}{z_{0i} + z_{1i} + z_{2i}} \quad (18)$$

onde V_{pf} representa a tensão pré-falta (de linha) na barra i e z_0 , z_1 e z_2 representam as impedâncias de sequência zero, positiva e negativa vistas desde as fases da mesma, respectivamente.

Durante uma falta fase-terra a componente de sequência negativa tem menor relevância e (18) assume uma forma mais simples [64]:

$$I_{falta_ki} = \frac{V_{pf_ki}}{z_{0i} + z_{1i}} \quad (19)$$

onde V_{pf} representa a tensão pré-falta na fase k da barra i .

3. Estudo de Caso

A Fig. 1 mostra o diagrama unifilar do SDEE utilizado para este estudo, com 14 barras e de topologia semelhante a uma seção do sistema de testes de 123 barras do IEEE (2025) [65]. A Tabela 1 mostra a configuração e o comprimento dos diferentes ramais.

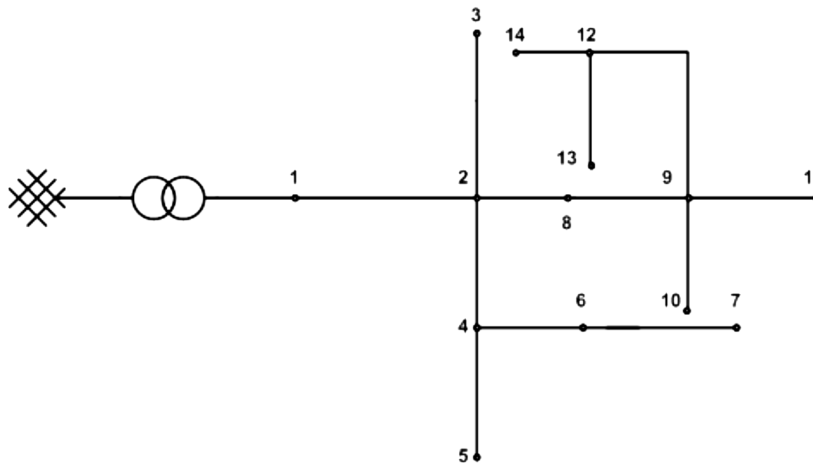


Fig. 1. Representação do SDEE em estudo.

Tabela 1. Características dos ramais da rede elétrica do SDEE

Ramal	1-2	2-3	2-4	4-5	4-6	6-7	2-8	8-9	9-10	9-11	9-12	12-13	12-14
Config*	C601	C602	C601	C602	C001	C602	C601	C601	C602	C602	C001	C602	C602
L (m)	1226	1609	1287	1287	1226	1287	966	1226	1226	1287	1448	1226	966

*Configuração. Fonte: [65].

O SDEE tem cargas em todas as barras, sendo suas impedâncias mostradas na Tabela 2.

Tabela 2. Impedâncias correspondentes às cargas nas barras do SDEE

Barra	$Z (\Omega)$
2	$28 + j18$
3	$27 + j17$
4	$22 + j15$
5	$21 + j15$
6	$20 + j15$
7	$18 + j17$
8	$21 + j15$
9	$16 + j14$
10	$16 + j13$
11	$16 + j12$
12	$14 + j12$
13	$14 + j8$
14	$13 + j10$

A subestação do SDEE está formada por um banco de três transformadores monofásicos conectados em YY, comreatância de $j0,375 \Omega$ (por fase). A alimentação deste banco é realizada por uma linha de potência de curto-circuito infinita, com tensão de linha igual a 4,16 kV e frequência de 60 Hz [65].

Foram realizadas diversas simulações de faltas monofásicas nas barras do SDEE, com e sem a presença do GD nas mesmas. Para isto o SDEE foi modelado no programa ATPDraw. As tensões e correntes obtidas com o programa ATP serviram como referência em todos os casos. As impedâncias de falta foram consideradas nulas. O MCF e o MCS foram implementados no programa Matlab, na forma de subprogramas. Na aplicação foram inseridos os valores das impedâncias das cargas e do transformador, assim como as informações dos alimentadores. A aplicação construiu as matrizes de admitância e de impedância de todos os componentes, assim como as matrizes de barras, de fases e de sequências.

As equações de Carson modificadas[63] foram utilizadas nos subprogramas para determinar as matrizes de impedâncias das linhas, considerando a configuração dos alimentadores e os tipos de condutores elétricos dos mesmos e resultaram iguais às reportadas pelo IEEE [65].

4. Resultados

4.1. Correntes de falta no SDEE

A Tabela 3 mostra as correntes de curto-circuito monofásico em cada uma das barras do SDEE obtidas através do programa ATP e na aplicação em Matlab.

Em geral, se observa uma excelente coincidência entre as correntes de referência, obtidas através do programa ATP, e as correntes obtidas através dos métodos de análise, fundamentalmente com o MCF. Com o MCS a maior diferença foi de 1,5% e foi observada na barra 9. Além da barra 1, as maiores correntes ocorrem nas barras 2, 8, 9 e 4, associadas às configurações C601.

Tabela 3. Correntes de curto-circuito no SDEE sem GD*

Barra	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$^1I_{ccATP}$	6349	2380	873	1269	726	849	569	1519	1118	678	643	739	513	537
I_{ccMCF}	6350	2380	873	1269	726	849	569	1519	1118	678	643	739	513	537
I_{ccMCS}	6349	2357	865	1252	718	844	564	1501	1101	669	635	733	508	532

*valores em Amperes (A)

A Tabela 4 mostra as correntes de curto-circuito monofásico em cada uma das barras do SDEE quando o GD está presente na barra 9.

Tabela 4. Correntes de curto-circuito no SDEE com GD*

Barra	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$^2I_{ccATP}$	6940	3073	958	1425	777	915	599	2660	4994	1259	1138	1264	710	759
I_{ccMCF}	6941	3073	958	1425	777	915	599	2660	4994	1259	1139	1264	710	759
I_{ccMCS}	6900	3040	949	1405	768	907	594	2633	4964	1249	1129	1271	709	758

*valores em Amperes (A)

Como esperado, com a GD todas as correntes aumentaram. A subestação continuou contribuindo com aproximadamente os mesmos valores de corrente, sendo a GD responsável pelo incremento observado. Nesta oportunidade, novamente se observa uma excelente coincidência entre as correntes de referência, reportadas pelo programa ATP e as que resultaram do MCF. Quanto ao MCS, quase todas as correntes foram inferiores às de referência, contudo, a diferença se manteve na faixa de 1,5%.

Novamente as maiores correntes foram registradas na barra 1 e em aquelas com a configuração C601. Novas barras mostraram valores de corrente significativos, no caso, as numeradas 10, 11 e 12, por causa da maior proximidade com a barra de GD Os incrementos nas magnitudes foram relevantes em várias barras, superando o 50%. Na barra 9 a corrente aumentou em mais de 3 vezes.

Como citado, os métodos de análise utilizados neste trabalho foram implementados no Matlab, programa de ampla utilização nos estudos dos SEPs. No entanto, o Matlab, assim como outros programas profissionais não estão disponíveis para todos os usuários. Por este motivo, empresas do setor elétrico, assim como pesquisadores do mundo todo desenvolvem programas próprios[16, 66-68].

Os resultados mostrados neste trabalho foram obtidos no marco do desenvolvimento de um programa próprio para estudos de Sistemas Elétricos (Fig. 2), com o Matlab sendo utilizado para validar os algoritmos desenvolvidos (nesta oportunidade relacionados com o MCS e o MCF para a análise de faltas).

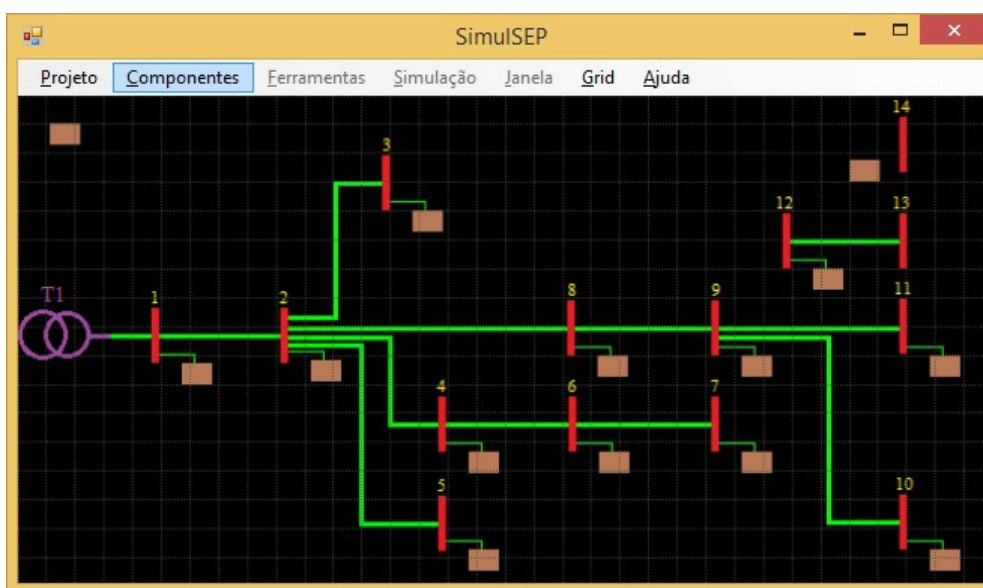


Fig. 2. Janela de programa em desenvolvimento para estudos em SEPs.

Os resultados mostrados a seguir foram perseguidos no intuito de verificar o comportamento dos métodos perante um maior nível de desequilíbrio do SDEE. Um dos testes incluiu o desbalanceamento das cargas. A Tabela 5, por exemplo, mostra as correntes de curto-circuito quando desbalanceadas as cargas da barra 7 e da barra 14, conforme a Tabela 6, estando a GD localizada na barra 9.

Novamente os resultados mostram excelente coincidência entre as correntes do programa ATP e as obtidas com o MCF, não assim quando considerado o MCS, com o qual as diferenças resultaram significativas. Isto pode ser comprovado na Fig. 3, que mostra as diferenças (Δ) relacionadas com este último método, obtidas através da seguinte equação:

$$\Delta(\%) = \left| \frac{I_{ccATP} - I_{ccMCS}}{I_{ccATP}} \right| 100 \quad (20)$$

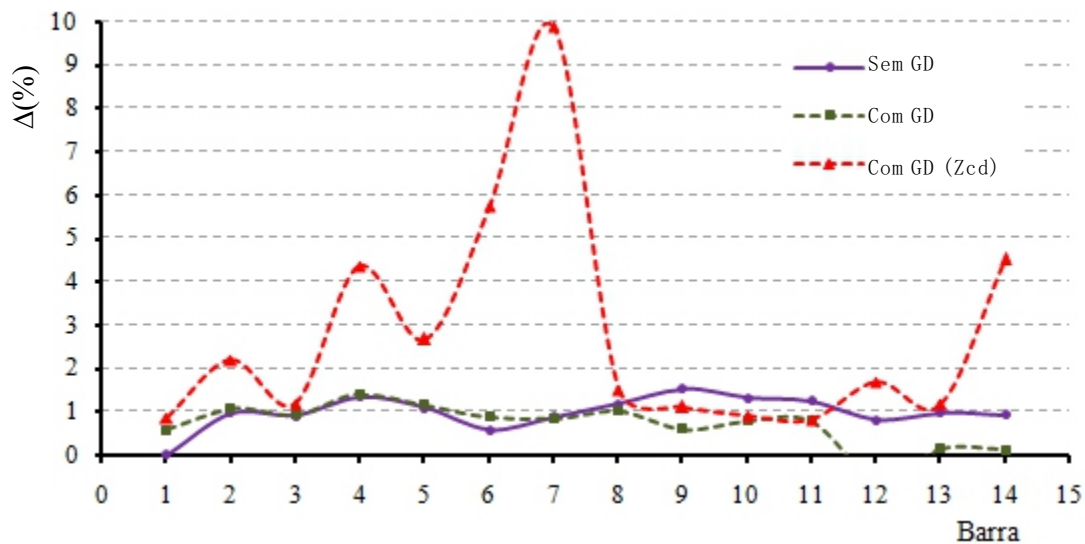
Tabela 5. Correntes de curto-circuito no SDEE com GD e cargas desbalanceadas na barra 7 e na barra 14*

Barra	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I_{ccATP}	6894	3055	926	1442	748	943	637	2598	4943	1217	1099	1309	681	842
I_{ccMCF}	6894	3055	927	1442	748	943	637	2597	4943	1217	1099	1309	681	843
I_{ccMCS}	6835	2988	915	1379	728	889	574	2559	4888	1206	1090	1287	673	804

*valores em Amperes (A)

Tabela 6. Impedâncias correspondentes a cargas desbalanceadas (Z_{cd})*

Fase	Barra	
	7	14
a	$5 + j3 \Omega$	$3 + j3 \Omega$
b	$8 + j4 \Omega$	$2 + j6 \Omega$
c	$9 + j2 \Omega$	$4 + j2 \Omega$

**Fig. 3. Diferenças relativas na determinação das correntes de curto-circuito relacionadas com o MCS.**

4.2. Afundamentos de tensão. Índices

Durante as simulações foram também registradas a tensão nas fases das diferentes barras. Nas análises que seguem foram utilizadas as tensões reportadas pelo programa ATP.

Após os curto-circuitos, quando a tensão residual na fase de alguma barra foi inferior a 0,9 p.u. considerou-se que a barra tinha sofrido um afundamento, isto é:

- 1 $\equiv$ Quando $V_{abc} \leq 0,9$ p.u. Significa barra com afundamento de tensão.
- 0 $\equiv$ Quando $V_{abc} > 0,9$ p.u. Significa barra sem afundamento de tensão.

Na seguinte *matriz binária de afundamentos de tensão de barras*, MAB_2 , cada linha representa a situação das barras perante o curto-circuito em uma delas, no caso, a representada pelo elemento i (diagonal principal). No caso particular, observa-se que, a exceção da barra 1, as tensões nas barras sofreram afundamento quando curto-circuitada qualquer uma delas. A matriz é de ordem m , com n sendo o número de barras do SDEE. Os superíndices nos elementos foram colocados para facilitar a identificação das linhas e colunas.

Sabendo a situação de cada barra é possível elaborar mapas de contorno empregando o diagrama unifilar. Por exemplo, na Fig. 4 (a) a área sombreada contorna as barras nas quais faltas não provocariam afundamentos de tensão na barra 1. Ou seja, as cargas na barra 1 seriam imunes a faltas nessas barras. A situação oposta é mostrada na Fig. 4 (b), onde são reveladas as regiões faltosas que afetariam à barra 1.

A partir da matriz MAB_2 foi elaborado o seguinte vetor, com seus elementos correspondendo à soma do número de barras sem afundamento de tensão em cada linha. Isto é equivalente a somar o número de “0” (zeros) de cada linha da matriz, isto é:

$$L^T = [0^{01} \ 0^{02} \ 1^03 \ 0^{04} \ 1^{05} \ 1^{06} \ 1^{07} \ 0^{08} \ 0^{09} \ 1^{10} \ 1^{11} \ 1^{12} \ 1^{13} \ 1^{14}] \quad (22)$$

onde:

$$L_i = \sum_{j=1}^n \begin{cases} 1 \text{ se } MAB_{2ij} = 0 \\ 0 \text{ se } MAB_{2ij} = 1 \end{cases} \quad (23)$$

$$MBA_2 = \begin{bmatrix} 1^{01} & 1^{02} & 1^{03} & 1^{04} & 1^{05} & 1^{06} & 1^{07} & 1^{08} & 1^{09} & 1^{10} & 1^{11} & 1^{12} & 1^{13} & 1^{14} \\ 1^{02} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{03} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1^{04} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{05} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{06} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{07} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1^{08} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1^{09} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{10} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{11} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{12} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{13} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{14} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (21)$$

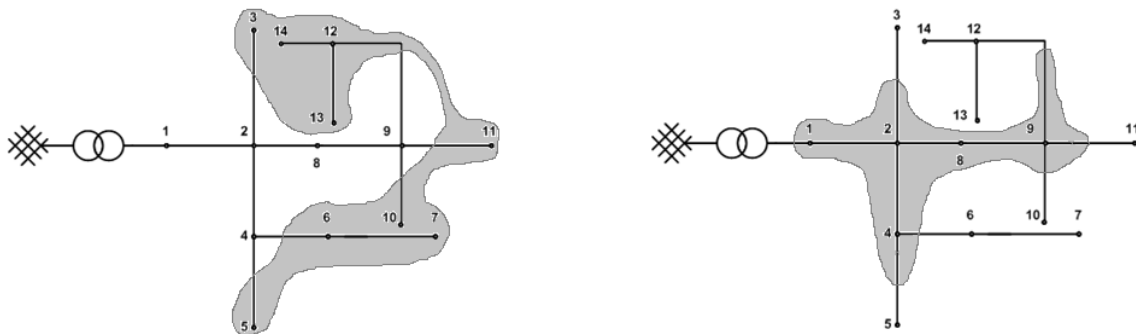


Fig. 4. Mapas de contorno relacionados com a barra 1.
a) Regiões de imunidade. b) Áreas expostas.

com j representando às colunas da matriz. A variável i representa o elemento do vetor L (que no caso

corresponde à linha i da matriz).

Assim, por exemplo, o elemento $L_3 = 1$ indica que quando da ocorrência do curto-circuito na barra 3, apenas a tensão em uma barra não sofreu afundamento. Observando a linha número 3 da matriz comprova-se que essa condição correspondeu à barra 1. Da mesma forma, o elemento $L_8 = 0$ indica que as 14 barras sofreram afundamento quando a barra 8 foi curto-circuitada. Os superíndices mostrados nos elementos do vetor L foram colocados para facilitar a identificação das linhas correspondentes na matriz MAB_2 .

A partir desta análise foi possível criar um indicador que chamamos de *Índice de Imunidade do Sistema perante Falhas Monofásicas*, $IIMS_\phi$, determinado como:

$$IIMS_\phi = \frac{N_{BSA}}{n^2} = \frac{\sum_{i=1}^n L_i}{n^2} \quad (24)$$

com:

$$IIMS_\phi \leq 1 \quad (25)$$

onde N_{BSA} representa o total de barras sem afundamentos de tensão durante os curto-circuitos (9 no caso). O divisor de (24) representa o número máximo teórico de afundamentos (196 no caso).

A determinação do indicador $IIMS_\phi$ não requer de complexos dados e seu valor dependerá das características particulares dos SDEEs, portanto, permitirá diferenciá-los. Por óbvio, quanto maior seja o $IIMS_\phi$ maior será a imunidade do SDEE perante os curto-circuitos.

Em outro cenário, foi colocado uma fonte de GD em cada uma das barras do SDEE, oportunidade em que eram provocados curto-circuitos em todas as barras. Em todos os casos foram elaboradas as matrizes binárias de afundamentos de tensão correspondentes. A matriz de (26) corresponde à situação em que a GD esteve colocada junto à barra 9. O vetor L é mostrado em (27).

Comparando (27) ou (26) com (22) ou (21) se observa que a situação perante os curto-circuitos não mudou para a barra 1. No entanto, agora com a GD, o sistema observou uma menor quantidade de afundamentos e por isso a maior quantidade de “0” (zeros) nesta última matriz (78 no total). Para analisar o impacto de uma barra em curto-circuito sobre a tensão nas outras, consideramos novamente cada uma das linhas da matriz de (26). A linha número 4, por exemplo, revela que diante de um curto-circuito na barra 4, não todas as barras são prejudicadas (especificamente, a partir da barra 9). Diferentemente, a matriz de (21) mostrou que, sem a GD, diante do mesmo fenômeno todas as barras foram afetadas. Comparação semelhante pode ser realizada com todas as outras barras.

Ao analisar as colunas da matriz de (26) uma nova informação pode ser retirada. Aproveitando que essa matriz está relacionada à GD colocada na barra 9, observemos a coluna número 9. Quando o curto circuito ocorre na barra 2, por exemplo, na barra 9 diminui a tensão provocando o afundamento ($MAB_{2_2 \times 9} = “1”$). Entretanto, considerando o curto circuito em outra barra, como a barra 4 antes comentada, ou a barra 6, percebe-se que o afundamento de tensão não se manifesta na barra 9 ($MAB_{2_4 \times 9} = MAB_{2_6 \times 9} = “0”$).

De modo geral, a coluna número 9 revela que a barra 9 sofreu afundamentos de tensão quando ela mesma e outras seis barras estiveram em curto-circuito. Curto-circuitos em uma quantidade semelhante de barras não provocaram afundamentos na barra 9. A situação particular provocada por uma barra k em curto-circuito sobre a barra 9 pode ser verificada no encontro da linha k com a coluna número 9. Uma análise semelhante pode ser realizada para as outras colunas, desta matriz, ou das que resultem de colocar GD em outras barras.

$$MBA_2 = \begin{bmatrix} 1^{01} & 1^{02} & 1^{03} & 1^{04} & 1^{05} & 1^{06} & 1^{07} & 1^{08} & 1^{09} & 1^{10} & 1^{11} & 1^{12} & 1^{13} & 1^{14} \\ 1^{02} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{03} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1^{04} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0^{05} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0^{06} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0^{07} & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1^{08} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1^{09} & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{10} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{11} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{12} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0^{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0^{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (26)$$

$$L^T = [0^{01} \quad 0^{02} \quad 7^{03} \quad 6^{04} \quad 7^{05} \quad 7^{06} \quad 8^{07} \quad 0^{08} \quad 0^{09} \quad 7^{10} \quad 7^{11} \quad 7^{12} \quad 11^{13} \quad 11^{14}] \quad (27)$$

Um novo vetor C pode ser elaborado, com seus elementos representando o número de vezes que uma barra não sofreu afundamentos de tensão, perante curto-circuitos nas outras barras, isto é:

$$C = [9^{01} \quad 5^{02} \quad 5^{03} \quad 5^{04} \quad 5^{05} \quad 5^{06} \quad 5^{07} \quad 3^{08} \quad 7^{09} \quad 7^{10} \quad 7^{11} \quad 5^{12} \quad 5^{13} \quad 5^{14}] \quad (28)$$

onde:

$$C_j = \sum_{k=1}^n \begin{cases} 1 \text{ se } MAB_{2kq} = 0 \\ 0 \text{ se } MAB_{2kq} = 1 \end{cases} \quad (29)$$

com k representando às linhas da matriz MAB_2 . A variável j representa o elemento do vetor C (que no caso corresponde à coluna j da matriz).

A análise de todas as matrizes permitiu definir um novo indicador, que chamamos de *Índice de Imunidade de Barras perante Falhas Monofásicas*, $IIMB_\phi$, válido para o sistema com GD e determinado como:

$$IIMB_{j,g} = \frac{N_{BNA}}{N_{BCA}} = \frac{(C_j)_g}{n - (C_j)_g} \quad (30)$$

onde N_{BCA} e N_{BNA} representam, respectivamente, o número de barras do SDEE que, quando curto-circuitadas, provocam, ou não, afundamentos de tensão em uma barra j , estando a GD em uma barra g .

Considerando que o curto-circuito em uma barra provoca afundamento de tensão nela mesma:

$$IIMB_\phi \leq n - 1 \quad (31)$$

4.3. Representação gráfica

Na Fig. 5 é aproveitado o diagrama unifilar do SDEE para revelar uma distribuição territorial do indicador $IIMB_\phi$. A representação é através de cores, contudo, poderia ter sido através de linhas, gerando um mapa de contorno.

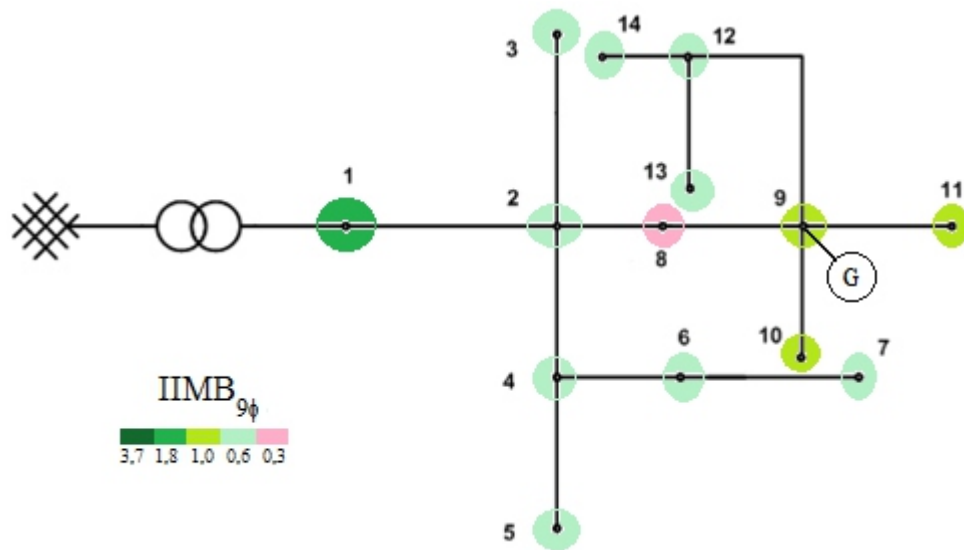


Fig. 5. Distribuição territorial do Índice de Imunidade de Barras perante Faltas Monofásicas com GD na barra 9.

Como toda representação gráfica o mapa da Fig. 5 contribui para uma rápida identificação da situação nas barras perante as faltas. Através do mapa é possível identificar a condição das cargas já posicionadas no SDEE. Por exemplo, com GD na barra 9 a carga localizada na barra 8 terá a pior condição, com $IIMB_{8,9\phi} \approx 0,3$. Para uma análise mais detalhada utilizamos a matriz MAB_2 a qual revela que tal carga se livrará dos afundamentos apenas quando as faltas aconteçam em três barras (as barras 7, 13 e 14). Já as cargas localizadas nas barras 9, 10 e 11 serão menos afetadas, com $IIMB_{\phi} = 1,0$ (a matriz MAB_2 revela que estas cargas serão afetadas por faltas em seis barras). As cargas na barra 1 teriam a melhor imunidade com $IIMB_{1,9\phi} = 1,8$.

Outra utilidade do indicador (e do mapa) é auxiliar na tomada de decisões. Por exemplo, diante de uma carga sensível localizada na barra 9 o mais conveniente seria instalar a GD na própria barra 9. De outro modo, em caso da GD já estar instalada junto a essa barra, cargas sensíveis poderiam ser instaladas nela ou nas barras 10 ou 11.

No caso de um SDEE sem GD, diante da decisão de incorporá-la, o indicador $IIMB_{\phi}$ pode revelar a melhor localização para a GD. A Fig. 6, por exemplo, revela a distribuição do indicador que se conseguiria ao instalar GD na barra 14. Nesta situação o indicador terá seu máximo valor, portanto, uma carga sensível deveria ser instalada conjuntamente nessa mesma barra para que sofra as menores consequências perante faltas nas outras barras.

De qualquer forma, as soluções requereriam de uma análise mais criteriosa, por exemplo, com esta última solução a barra 14 seria favorecida, mas as outras barras seriam fortemente prejudicadas (em comparação a solução correspondente à Fig. 5, por exemplo).

No SDEE em estudo o valor máximo teórico do indicador é 13,0, contudo, o valor máximo que se consegue é $IIMB_{14,14\phi} \approx 3,7$.

Ressalta-se que a barra 9 foi utilizada nos exemplos acima porque colocando nela GD, o SDEE experimentou o máximo valor do indicador $IIMB_{\phi}$, no caso 0,4. Para comparação, quando o SDEE trabalhou sem GD, $IIMB_{\phi} \approx 0,05$.

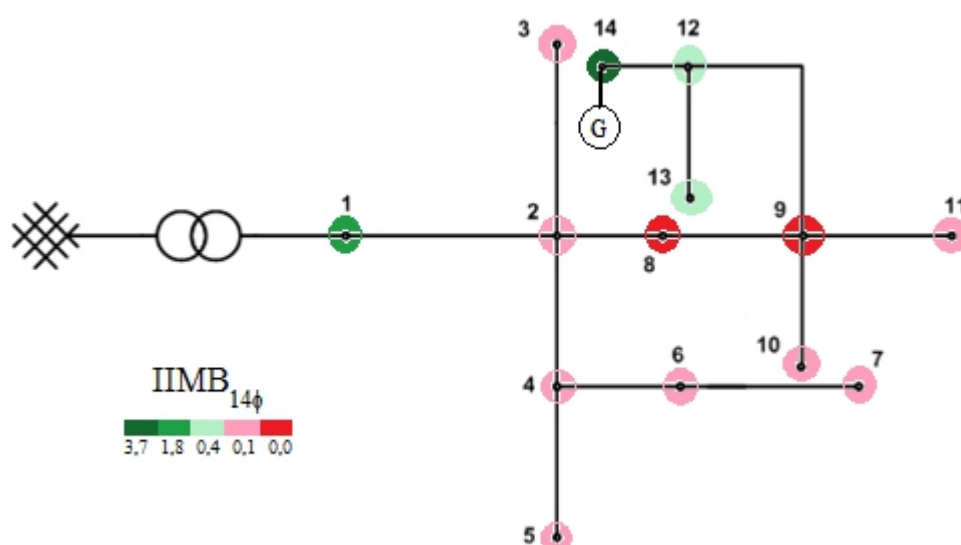


Fig. 6. Distribuição territorial do Índice de Imunidade de Barras perante Falhas Monofásicas com GD na barra 14.

5. Conclusões

Este trabalho analisa resultados de simulações de falhas monofásicas realizadas em um Sistema de Distribuição de Energia Elétrica de configuração radial (SDEE), que opera com e sem Geração Distribuída (GD). O SDEE contou com 14 barras e foi modelado no ATPDraw, sendo os dados dos seus componentes tomados de SDEEs de referência do IEEE. As tensões e correntes fornecidas pelo ATP serviram de referência nos estudos.

Em uma primeira análise as correntes do ATP foram comparadas com as obtidas através de uma aplicação em Matlab, na qual foram implementados os métodos de Componentes Simétricas (MCS) e de Componentes de Fases (MCF). O trabalho demonstra que com ambos os métodos os resultados são satisfatórios quando o SDEE opera com cargas balanceadas. Entretanto, quando algumas cargas foram desbalanceadas os resultados do MCS se afastaram dos utilizados como referência, chegando a diferença a 10% em uma das barras. Ressalta-se que esta atividade foi realizada no marco da seleção de um método de análise a ser inserido em um programa computacional próprio em desenvolvimento.

Em uma segunda análise, a partir das tensões reportadas pelo ATP foi realizado um estudo de afundamentos de tensão. Em função disso, em relação ao valor de 0,9 p.u., as tensões nas diferentes barras foram classificadas. Foram elaboradas matrizes booleanas de afundamentos de tensão a partir das quais novos instrumentos são elaborados. Por exemplo, mapas de contorno, que revelam a situação das barras perante falhas que aconteçam no SDEE.

Dois índices relacionados com os afundamentos de tensão foram obtidos com o trabalho. Os índices contabilizam os elementos das matrizes booleanas e, portanto, estão relacionados com falhas monofásicas. O Índice de Imunidade do Sistema caracteriza o SDEE integralmente e permite diferenciar SDEEs ou comparar a situação de um mesmo SDEE perante as falhas, quando modificações são projetadas ou realizadas nele.

Já o Índice de Imunidade de Barras foi definido para o SDEE trabalhando com GD e permite diferenciar as barras enquanto ao número de outras barras que as afetam. O índice facilita a elaboração de informação visual, de mapas, aproveitando o unifilar. As barras são coloridas conforme o valor do indicador facilitando a tomada de decisões a respeito da localização do gerador independente e/ou de cargas sensíveis.

Em trabalhos futuros outros níveis de tensão devem ser considerados para elaborar as matrizes, e SDEEs com maior número de barras devem ser estudados. Outros tipos de faltas também devem ser estudados. Para todos os casos a informação gráfica deve ser enriquecida, considerando a posição nas linhas elétricas até onde, ou desde onde, uma falta se propaga, tornando mais precisos os mapas de contorno.

Agradecimentos

Os autores agradecem o apoio financeiro recebido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq. Também, agradecem o apoio da Facultad de Ingeniería da Universidad Nacional de Misiones – UNaM.

Referências

- [1] B.D. Bonatto, E.A. Mertens, L.F.S. Dias, S. Nosaki, E.H. Watanabe, M. Aredes, S. Carneiro and H.W. Dommel, “Eletrônica de Potência e Qualidade da Energia Elétrica: Pesquisas Cooperativas na ELEKTRO, COPPE/UFRJ e UBC – Resultados Práticos,” in II Congresso de Inovação Tecnológica em Energia Elétrica. Salvador, Brasil, 2003.
- [2] Y. Han, Y. Feng, P. Yang, L. Xu, Y. Xu and F. Blaabjerg, “Cause, Classification of Voltage Sag, and Voltage Sag Emulators and Applications: A Comprehensive Overview,” *IEEE Access*, vol. 8, 2020. <https://doi.org/10.1109/ACCESS.2019.2958965>.
- [3] A.Stano, D.F. Bernardes, J.C.G. Rodrigues and J.M.E. Vicente, “O Impacto do Uso de Microcomputadores na Qualidade da Energia Elétrica. Instituto CGTI – Centro de Gestão de Tecnologia e Inovação”. Campinas, Brasil, 2025. Disponível em <https://www.cgti.org.br/publicacoes>.
- [4] M.V. Bhakkad and B.T. Deshmukh, “Generation of Voltage Sag for different loads and conditions using MATLAB SIMULINK,” in International Conference on Innovative Trends and Advances in Engineering and Technology. Shergaon, India, 2019. <https://doi.org/10.1109/ICITAET47105.2019.9170228>.
- [5] C.A.H. Campo, V.G. Garrido, E.C. Peñaranda and P. Guarín, “Harmonic distortion analysis and mitigation in LED public lighting systems using PID control and LC filters,” in IEEE Workshop on Power Electronics and Power Quality Applications. Cartagena, Colombia, 2025. <https://doi.org/10.1109/PEPQA66960.2025.11073514>.
- [6] Z. Reguieg, I. Bouyakoub and F. Mehedi, “Harmonic mitigation in grid-integrated renewable energy systems with nonlinear loads,” *Energy*, vol. 324, 2025. <https://doi.org/10.1016/j.energy.2025.135882>.
- [7] M.D. Teixeira, R.L. Araújo, N.S.R. Quoirin, L.M. Ardjomand, A.R. Aoki, G. Paulillo, A.P. Sgobero and I. Wunderlich, “Acceptable Power Quality Limits to Avoid Damage In Appliances,” *WSEAS Transactions on Circuits and Systems*, 4(5) junho, 2005. Disponível em <https://worldses.org/journals/circuits/odl.htm>.
- [8] S.V. Fukuda, M.D. Teixeira, A.R. Aoki, G. Paulillo, R.L. Araújo, L.M. Ardjomand, N.S.R. Quoirin, R.A. Peniche, I. Wunderlich and A.P. Sgobero, “Avaliação dos Limites de Parâmetros de Qualidade de Energia Aceitáveis para Prevenir Danos em Eletrodomésticos,” in Congresso de Inovação Tecnológica em Energia Elétrica. Florianópolis, Brasil, 2005.
- [9] V. Liubcuk, V. Radziukynas, G. Kairaitis and D. Naujokaitis, “Power Quality Impact and Its Assessment: A Review and a Survey of Lithuanian Industrial Companies,” *Inventions*, 10(30), 2025. <https://doi.org/10.3390/inventions10020030>.
- [10] A. Sharma, B.S. Rajpurohit and S.N. Singh, “A review on economics of power quality: Impact, assessment and mitigation,” *Renewable and Sustainable Energy Reviews*, vol. 88, 2018. <https://doi.org/10.1016/j.rser.2018.02.011>.
- [11] R.C. Dugan, M.F. McGranaghan, S. Santoso and H.W. Beaty, *Electrical Power Systems Quality*. 3a Ed. McGraw Hill, New York, Estados Unidos de América, 2012.

- [12] M.H.J. Bollen, M. Speychal and K. Lindén, "Estimation of Dip Frequency from Fault Statistics Including Three-Phase Characteristics," in 9th International Conference on Probabilistic Methods Applied to Power Systems. KTH, Stockholm, Sweden, 2006. <https://doi.org/10.1109/PMAPS.2006.360400>.
- [13] J.C. Gomez and M.M. Morcos, "Voltage sags and recovery time in repetitive events," *IEEE Transactions on Power Delivery*, 17(4), 2002. <https://doi.org/10.1109/TPWRD.2002.803840>.
- [14] L. Qun, "Analysis on Fault Statistics in Wenzhou Electric Power Distribution Network," in Asia-Pacific Power and Energy Engineering Conference. Wuhan, China, 2009. <https://doi.org/10.1109/APPEEC.2009.4918168>.
- [15] M.H.J. Bollen, *Solving power quality problems: voltage sags and interruptions*. New York, Estados Unidos de América: IEEE Press, 1999.
- [16] D. Gordon, P. Scott and S. Thiébaux, "SmartGridToolbox: a library for simulating modern and future electricity networks," *arXiv*, 2022. <https://doi.org/10.48550/arXiv.2110.08992>.
- [17] X. Guo, Y. Li, S. Wang, Y. Cao, M. Zhang, Y. Zhou and N. Yosuke, "A Comprehensive Weight-Based Severity Evaluation Method of Voltage Sag in Distribution Networks," *Energies*, 14(19), 2021. <https://doi.org/10.3390/en14196434>.
- [18] J.R.S. Martins, D.A. Fernandes, F.F. Costa and M.B.R. Corrêa, "Compensação de tensão trifásica em cargas sensíveis baseada em um sistema de controle repetitivo e mínimos quadrados," *Eletrônica de Potência*, 22(3), 2017. <https://doi.org/10.18618/REP.2017.3.2681>.
- [19] N. Phannil, S. Ananwattanaporn and C. Pothisarn, "Reduction of impact from voltage sags during faults in distribution system by installing neutral grounding resistor," in 13th International Electrical Engineering Congress. Hua Hin, Thailand, 2025. <https://doi.org/10.1109/IEEECON64081.2025.10987739>.
- [20] N.P. da Silva, A.C. Francisco, J.L. Kovalski and M.S. Thomaz, "Avaliação do Impacto das Descargas Atmosféricas na Qualidade de Energia Fornecida pelas Concessionárias: Estudo de Caso em uma Empresa de Distribuição de Energia do Sul do País," *Nucleus*, 7(1), 2010. <https://doi.org/10.3738/1982.2278-335>.
- [21] J.C. Ramirez, E. Gómez-Luna and J.C. Vasquez, "Technical and economic feasibility analysis to implement a solid-state transformer in local distribution systems in Colombia," *Energies*, 18 (14), 2025. <https://doi.org/10.3390/en18143723>.
- [22] J. Xie, J. Wei, H. Feng, B. Wu and L. Ran, "A short-circuit protection method for SiC MOSFET using combined Vgs and Vds detection," *IEEE Transactions on Power Electronics*, 40 (9), 2025. <https://doi.org/10.1109/TPEL.2025.3565503>.
- [23] H. Okumus, F.M. Nuroglu and S. McLoone, "Ensemble learning for short circuit fault location estimation in distribution networks," *Engineering Applications of Artificial Intelligence*, 159 (A), 2025. <https://doi.org/10.1016/j.engappai.2025.111640>.
- [24] D. Song, Z. Zhang, Y. Xing and H. Shi, "A medium-voltage side short-circuit fault location method for distribution networks based on low-voltage side phase-to-phase voltage amplitude," *Journal of Physics: Conference Series*, vol. 3043, 2025. <https://doi.org/10.1088/1742-6596/3043/1/012158>.
- [25] C.L. Fortescue, "Method of Symmetrical Co-Ordinates Applied to the Solution of Polyphase Networks," *AIEE Transactions*, vol. 37, part II, 1918. <https://doi.org/10.1109/T-AIEE.1918.4765570>.
- [26] C.F. Wagner and R.D. Evans, *Symmetrical Components as Applied to the Analysis of Unbalanced Electrical Circuits*. 1st ed., New York, Estados Unidos de América: McGraw-Hill, 1933.
- [27] M.A. Laughton, "Analysis of unbalanced polyphase networks by the method of phase coordinates. Part II: Fault analysis," *Proceedings of the Institution of Electrical Engineers*, 116 (5), 1969. <https://doi.org/10.1049/piee.1969.0160>.
- [28] E.B. Makram, M.A. Bou-Rabee and A.A. Girgis, "Three-Phase Modeling of Unbalanced Distribution Systems during Open Conductors and/or Shunt Fault Conditions Using the Bus Impedance Matrix," *Electric Power Systems Research*, 13(3), 1987. [https://doi.org/10.1016/0378-7796\(87\)90002-2](https://doi.org/10.1016/0378-7796(87)90002-2).
- [29] H. Saadat, *Power system analysis*. 2nd ed, New York, Estados Unidos de América: McGraw-Hill's Primis Custom Publishing, 2002.
- [30] S.M. Halpin and L.L. Grigsby, "Fault analysis of multi-phase unbalanced non-radial power distribution systems," in Proceedings of Industrial and Commercial Power Systems Conference, 1994. <https://doi.org/10.1109/ICPS.1994.303574>.

- [31] H. Li, J. Chen, Y. Liang, F. Liao and G. Wang, "Effects of imbalance on single-phase to ground fault characteristics in low-resistance grounded systems," *International Journal of Electrical Power and Energy Systems*, vol. 115, 2020. <https://doi.org/10.1016/j.ijepes.2019.105504>.
- [32] Institute of Electrical and Electronics Engineers– IEEE, "IEEE Std 1159-2019 – Recommended Practice for Monitoring Electric Power Quality," 2019. <https://doi.org/10.1109/IEEESTD.2019.8796486>.
- [33] International Electrotechnical Commission – IEC, "IEC Std 61000-2-2:2002 – Electromagnetic compatibility (EMC) - Part 2-2: Environment - Compatibility levels for low-frequency conducted disturbances and signalling in public low-voltage power supply systems," 2002.
- [34] Agência Nacional de Energia Elétrica– ANEEL, "Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST: Módulo 8 – Qualidade do Fornecimento de Energia Elétrica" 2023. Brasília, Brasil.
- [35] D.L. Brooks, R.C. Dugan, M. Wacławski and A. Sundaram, "Indices for assessing utility distribution system RMS variation performance," *IEEE Transactions on Power Delivery* 13(1), 1998. <https://doi.org/10.1109/61.660886>.
- [36] H. Liao, S. Abdelrahman, Y. Guo and J.V. Milanovic, "Identification of Weak Areas of Power Network Based on Exposure to Voltage Sags—Part I: Development of Sag Severity Index for Single-Event Characterization," *IEEE Transactions on Power Delivery*, 30(6), 2014. <https://doi.org/10.1109/TPWRD.2014.2362965>.
- [37] T. Ackermann, G. Andersson and L. Söder, "Distributed generation: a definition," *Electric Power Systems Research*, 57(3), 2001. [https://doi.org/10.1016/S0378-7796\(01\)00101-8](https://doi.org/10.1016/S0378-7796(01)00101-8).
- [38] M. Khodadadi and A. Askarzadeh, "A techno-enviro-economic framework for optimal operation of a battery-driven hybrid energy system with biomass: A risk-averse approach," *Energy*, vol. 320, 2025. <https://doi.org/10.1016/j.energy.2025.135273>.
- [39] N. Permal, F. Nagi, M. Marsadek, A.K. Ramasamy, N. Boopalan, K. Kumar and A.L. Balakrishna, "EV and PV penetration impact on grid with conservative voltage regulation and reactive voltage compensation," *ISA Transactions*, vol. 158, 2025. <https://doi.org/10.1016/j.isatra.2025.01.004>.
- [40] S.L.K. Bharathi, V. Thanigaivelan, S. Shenbagaraman and M.R. Babu, "Hybrid EVO-CFNN approach for improve voltage stability of distribution system performance using superconducting magnetic energy storages inter linked with wind turbine," *Journal of Energy Storage*, vol. 12, 2025. <https://doi.org/10.1016/j.est.2025.115465>.
- [41] S. Roy and R. Singh, "Distributed energy procurement with renewable energy sources in a radial distribution system," *Computers and Electrical Engineering*, 123(D), 2025. <https://doi.org/10.1016/j.compeleceng.2025.110268>.
- [42] Programa ATPDraw, 2025. <http://www.atpdraw.net>.
- [43] Programa Matlab, 2025. <http://www.mathworks.com>.
- [44] Programa EMTP/ATP, 2025. <http://www.empt.com>.
- [45] S. Quaia and F. Tosato, "The Voltage Sags Matrix: a useful tool for power quality analysis," *Electrotechnical Review*, 69(3-4), 2002. Disponível em <https://ev.fe.uni-lj.si/3-4-2002/quaia.pdf>.
- [46] G.G. Santos, A.C. dos Santos, T.L. Oliveira and C.E. Tavares, "Aplicativo computacional para cálculo de curto-circuito e estudos de afundamentos de tensão," in XIII Conferência de Estudos em Engenharia Elétrica. Uberlândia, Brasil, 2015. Disponível em <https://www.peteletricaufu.com.br/ceel/evento/artigos/>.
- [47] M.F. McGranaghan, D.R. Mueller and M.J. Samoty, "Voltage sags in industrial systems," *IEEE Transactions on Industry Applications*, 29(2), 1993. <https://doi.org/10.1109/28.216550>.
- [48] W. Liu, L.G. Tu, H.W. Hong and W. Zhong, "An integrated application for voltage sag analysis," in 20th International Conference on Power Industry Computer Application Columbus, OH, Estados Unidos de América, 1997. <https://doi.org/10.1109/PICA.1997.599386>.
- [49] M.R. Qader, M.H.J. Bollen and R.N. Allan, "Stochastic prediction of voltage sags in a large transmission system," *IEEE Transactions on Industry Applications*, 35(1), 1999. <https://doi.org/10.1109/28.740859>.
- [50] C. Behera, A. Banik, and A.K. Goswami, "A novel approach for voltage sag representation in a chemical industry: A case study," *Engineering Reports*, 2(7), 2020. <https://doi.org/10.1002/eng2.12198>.

- [51] T. Nguyen and Y. Liao, "Power quality disturbance classification utilizing S-transform and binary feature matrix method," *Electric Power Systems Research*, 79(4), 2008. <https://doi.org/10.1016/j.epsr.2008.08.007>.
- [52] S. Jashfar, S. Esmaeili, M. Zareian-Jahromi and M. Rahmanian, "Classification of power quality disturbances using S- transform and TT-transform based on the artificial neural network," *Turkish Journal of Electrical Engineering and Computer Sciences*, 21 (6), 2013. <https://doi.org/10.3906/elk-1112-51>.
- [53] E. Espinosa-Juárez and A. Hernández, "A method for voltage sag state estimation in power systems," *IEEE Transactions on Power Delivery*, 22(4), 2007. <https://doi.org/10.1109/TPWRD.2007.905587>.
- [54] G. Carpinelli, P. Caramia, C. Di Perna, P. Varilone and P. Verde, "Complete matrix formulation of fault-position method for voltage-dip characterisation," *IET Generation, Transmission and Distribution*, 1(1), 2007. <https://doi.org/10.1049/iet-gtd:20050544>.
- [55] T.R. Kempner, M. Oleskovicz and D.P.S. Gomes, "Optimal monitoring of voltage sags through simultaneous analysis of short-circuits in distribution systems," *IET Generation, Transmission and Distribution* 11(7), 2017. <https://doi.org/10.1049/iet-gtd.2016.1648>.
- [56] S. Jo, J.Y. Oh, Y.T. Yoon and Y.G. Jin, "Self -healing radial distribution network reconfiguration based on deep reinforcement learning," *Results in Engineering*, vol. 22, 2024. <https://doi.org/10.1016/j.rineng.2024.102026>.
- [57] A. Mishra, M. Tripathy and P. Ray, "A survey on different techniques for distribution network reconfiguration," *Journal of Engineering Research*, 12(1), 2024. <https://doi.org/10.1016/j.jer.2023.09.001>.
- [58] R. Fathi, B. Tousi and S. Galvani, "Allocation of renewable resources with radial distribution network reconfiguration using improved salp swarm algorithm," *Applied Soft Computing*, vol. 132, 2023. <https://doi.org/10.1016/j.asoc.2022.109828>.
- [59] R. Nebuloni, V. Ilea, C. Bovo, A. Berizzi, C. Arrigoni, F. Re and R. Bonera, "Optimal reconfiguration of radial distribution networks with renewable energy resources by considering configuration shift steps," *Sustainable Energy, Grids and Networks*, vol. 43, 2025. <https://doi.org/10.1016/j.segan.2025.101861>.
- [60] W.H. Hayt and J.E. Kemmerly, *Engineering Circuit Analysis*, 2nd ed., New York, Estados Unidos de América: McGraw-Hill, 1962.
- [61] W.D. Stevenson, *Elements of Power System Analysis*. 2nd ed., New York, Estados Unidos de América: McGraw-Hill Book Company, Inc., 1962.
- [62] M.A. Laughton, "Analysis of unbalanced polyphase networks by the method of phase coordinates. Part I: System representation in phase frame of reference," *Proceedings of the Institution of Electrical Engineers*, 115(8), 1968. <https://doi.org/10.1049/piee.1968.020>.
- [63] W.H. Kersting, *Distribution System Modeling and Analysis*. 2nd ed., Boca Raton, Florida, Estados Unidos de América: CRC Press, 2007.
- [64] S.M. Halpin and L.L. Grigsby, "A comparison of fault calculation procedures for industrial power distribution systems: the past, the present and the future," in *Proceedings of the IEEE International Conference on Industrial Technology*, 1994. <https://doi.org/10.1109/ICIT.1994.467213>.
- [65] Institute of Electrical and Electronics Engineers IEEE, "IEEE PES Test Feeder," 2025. Disponível em <https://cmte.ieee.org/pes-testfeeders/resources>.
- [66] Centro de Pesquisas de Energia Elétrica CEPEL, "ANAFAS 2.0 – Programa de Análise de Falhas Simultâneas," 2023. Disponível em <https://www.cepel.br/anafas-2>.
- [67] P. Janiga, Z. Eleschová and D. Viglaš, "Short-Circuit – Analysis and Calculation," *WSEAS Transactions on Power Systems*, vol. 9, 2014. Disponível em <https://wseas.com/journals/articles.php?id=5315>.
- [68] B.R. Sodr , B.A. Silva, A.S. Velloso, B.W. Fran a, M.Z. Fortes and G.G. Sotelo, "Desenvolvimento de um aplicativo em VBA para simula  o de curto-circuitos em rede de distribui  o utilizando o programa ANAFAS," in *Simp sio Brasileiro de Sistemas El tricos*, 2020. <https://doi.org/10.48011/sbse.v1i1.2502>.

Clasificación Multiclase de Estados en Colmenas Mediante Bioacústica y Refinamiento de Etiquetas Centrado en Datos

Axel Alfredo Skrauba ^{1,2*} , Hector De Sosa ¹ , Sandro Daniel Zakowicz ¹ 

¹ Universidad Nacional de Misiones. Facultad de Ingeniería. Departamento de Ingeniería Electrónica y Computación (DIEC). Misiones, Argentina.

² Universidad Nacional de Misiones-CONICET. Facultad de Ingeniería. Instituto de Materiales de Misiones (IMAM) Grupo de Investigación y Desarrollo en Electrónica (GIDE). Misiones, Argentina.

e-mails: axel.skrauba@fio.unam.edu.ar, hectordesosa.01@gmail.com, sandrodanielzakowicz@gmail.com

Resumen

Este estudio aborda el desafío del ruido en las etiquetas de conjuntos de datos bioacústicos para la apicultura de precisión. A diferencia de los enfoques binarios tradicionales, se propone un marco de clasificación multiclase para identificar estados complejos de la colmena (reina aceptada, rechazada, presente o ausente). La metodología integra la extracción de características mediante Coeficientes Cepstrales en las Frecuencias de Mel (MFCC) con un *pipeline* de Confident Learning (Cleanlab) para identificar y refinar etiquetas erróneas en el dataset público Smart Bee Colony Monitor. Los resultados demuestran que el saneamiento de datos mejora la precisión del modelo CatBoost del 82% al 84%, validando la utilidad de los enfoques centrados en los datos (Data-Centric AI) para mitigar la subjetividad humana en el etiquetado biológico, incluso en ausencia de validación de campo inmediata.

Palabras Clave – Abejas, Apicultura, Corrección de Etiquetas, Clasificación de Sonido, Extracción de Características, Machine Learning, Monitoreo de Colmenas, Preprocesamiento.

Abstract

This study addresses the challenge of label noise in bioacoustic datasets for precision apiculture. Departing from traditional binary approaches, it proposes a multiclass classification framework to identify complex hive states (queen accepted, rejected, present, or absent). The methodology integrates feature extraction using Mel-Frequency Cepstral Coefficients (MFCC) with a Confident Learning (Cleanlab) pipeline to detect and refine mislabeled instances in the public Smart Bee Colony Monitor dataset. The results show that data sanitization improves the accuracy of the CatBoost model from 82% to 84%, validating the usefulness of data-centric AI approaches in mitigating human subjectivity in biological labeling, even in the absence of immediate field validation.

Keywords – Bees, Beekeeping, Label Correction, Sound Classification, Feature Extraction, Machine Learning, Hive Monitoring, Preprocessing.

1. Introducción

1.1 Contexto y Motivación

La reina es fundamental para la supervivencia y productividad de una colmena, ya que es la única abeja que puede poner huevos fertilizados, lo que garantiza la continuidad de la colonia. La presencia o ausencia de la reina puede indicar el estado de salud de la colmena. Si la reina está ausente o no es aceptada, la productividad y estabilidad de la colmena se ven comprometidas.

Tradicionalmente, los apicultores inspeccionan manualmente las colmenas para verificar la presencia de la reina y el estado general de la colonia. Sin embargo, este proceso es invasivo y puede perturbar a las abejas, lo que podría llevar a un estrés innecesario en la colonia y afectar negativamente la producción de miel [1].

Para evitar la intrusión directa en las colmenas, se propone un enfoque no invasivo que utiliza el análisis del sonido del zumbido de las abejas como un indicador del estado de la colmena. Los

cambios en los patrones de zumbido pueden revelar la presencia o ausencia de la reina, así como otros estados importantes de la colmena [2], [3].

El desarrollo de un modelo de *machine learning* que pueda clasificar estos estados basados en el sonido abre la posibilidad de crear una aplicación móvil. Esta herramienta proporcionaría a los apicultores una forma sencilla y no invasiva de monitorizar la salud de sus colmenas, mejorando la gestión apícola.

1.2 *Objetivos del Estudio*

En este estudio se plantean tres objetivos principales que se centran en el desarrollo y optimización de un modelo de clasificación basado en *machine learning* para la identificación de estados en colmenas de abejas, utilizando análisis de audio. A continuación, se detallan los objetivos específicos:

1.2.1 *Desarrollar un modelo de clasificación para determinar los estados de la colmena*

El primer objetivo es diseñar y entrenar un modelo de *machine learning* capaz de clasificar los diferentes estados de una colmena en cuatro clases distintas utilizando grabaciones de audio del zumbido de las abejas. Este enfoque no invasivo tiene el potencial de revolucionar la apicultura al permitir un monitoreo continuo y en tiempo real del estado de las colmenas sin la necesidad de inspecciones manuales que podrían perturbar a las abejas. Concretamente, se abordarán distintos estados de interés en la colmena, tomando como foco a la reina.

1.2.2 *Corregir etiquetas incorrectas en los datos*

Uno de los desafíos inherentes al manejo de datos etiquetados manualmente es la posibilidad de errores en las etiquetas. Estos errores pueden afectar significativamente el rendimiento de los modelos de *machine learning*, ya que introducen ruido en el proceso de entrenamiento. Cleanlab [4], una herramienta reciente y emergente en el campo de la corrección de etiquetas, se utilizará en este estudio para identificar y corregir posibles etiquetas erróneas en el conjunto de datos de audio. Este *framework*, implementa Confident Learning, en dónde se utiliza un modelo predictivo base para estimar las probabilidades de pertenencia de cada instancia respecto a las clases objetivo.

1.2.3 *Evaluar el impacto de la corrección de etiquetas y optimización del pipeline en el rendimiento del modelo*

Finalmente, el estudio también se enfoca en evaluar cómo la corrección de etiquetas y la optimización del *pipeline*, mediante el uso del *framework* PyCaret [5], impactan el rendimiento general del modelo de clasificación. La optimización del *pipeline* es esencial para asegurar que el modelo alcance su máximo potencial en términos de precisión y generalización. Se analizarán diferentes combinaciones de parámetros y arquitecturas de modelos para determinar la configuración que ofrezca los mejores resultados. Además, se realizará un análisis comparativo del rendimiento del modelo antes y después de la aplicación de Cleanlab, lo que permitirá medir el valor añadido de la corrección de etiquetas en este contexto.

2. Revisión de Literatura

2.1 *Procesamiento de Audio en Apicultura*

El uso del sonido como herramienta para el estudio de las colmenas de abejas no es una práctica nueva, pero ha ganado relevancia significativa en las últimas décadas debido a los avances en tecnología de grabación y procesamiento de señales. Los primeros estudios se centraron en la observación del comportamiento de las abejas mediante grabaciones de audio rudimentarias, buscando correlaciones entre ciertos sonidos y el estado de la colmena. Estos métodos permitieron a los investigadores identificar patrones de zumbido que indicaban la presencia de la reina, enjambres inminentes, o estados de estrés dentro de la colonia [6].

A medida que la tecnología avanzó, también lo hicieron las aplicaciones del análisis de sonido en la apicultura. Hoy en día, se emplean técnicas sofisticadas de procesamiento de señales y *machine learning* para interpretar los datos de audio de manera más precisa y automatizada [3]. No obstante, a pesar de los avances, los métodos actuales todavía enfrentan limitaciones significativas, como la variabilidad en la calidad de las grabaciones y la necesidad de un etiquetado manual preciso de los datos, lo que introduce un nivel considerable de error humano en el proceso.

2.2 *Extracción de Características en Audio*

En el campo del procesamiento de audio, la extracción de características es una etapa crucial que determina la efectividad del modelo de clasificación. Existen varios métodos para extraer características relevantes de señales de audio, cada uno con sus propias ventajas y desventajas. Entre los métodos más comunes se encuentran los Coeficientes Cepstrales en las Frecuencias de Mel (MFCC), la transformada rápida de Fourier (FFT), y más recientemente, técnicas especializadas en la captura de descriptores en general, que se sustentan en los *embeddings* de modelos profundos, entrenados para tareas generales. En el caso concreto de audio, un ejemplo de esto sería YAMNet [7].

Los MFCCs han sido ampliamente adoptados en aplicaciones de reconocimiento de voz y se han adaptado con éxito al análisis de audio en apicultura debido a su capacidad para capturar la envolvente espectral de las señales de zumbido, que es fundamental para la diferenciación de estados en la colmena [3]. Por otro lado, la FFT proporciona una representación frecuencial directa de la señal, lo que puede ser útil para identificar patrones específicos en el zumbido de las abejas. Sin embargo, los MFCCs son preferidos en este estudio debido a su eficacia en el modelado de señales complejas como las generadas por las abejas, lo que justifica su selección como método principal para la extracción de características en este trabajo.

2.3 *Optimización de Pipelines*

La optimización de *pipelines* es un componente esencial en cualquier proyecto de inteligencia artificial, ya que permite maximizar el rendimiento del modelo mediante la selección adecuada de hiperparámetros y arquitecturas, además de factores involucrados en el preprocesamiento de los descriptores. Herramientas como PyCaret han facilitado este proceso al proporcionar una plataforma que automatiza la selección de modelos, la optimización de hiperparámetros, y la evaluación comparativa, lo que ahorra tiempo y recursos a los investigadores.

En el contexto de este estudio, PyCaret se utilizará para construir y optimizar el *pipeline* de clasificación. Estudios previos han demostrado que la optimización cuidadosa del *pipeline* puede llevar a mejoras significativas en el rendimiento de los modelos, especialmente cuando se combinan con técnicas de corrección de etiquetas como Cleanlab [4]. Estas mejoras son cruciales en aplicaciones como la apicultura, donde la precisión del modelo puede tener un impacto directo en la salud y productividad de las colmenas, y, por ende, en los ingresos de los apicultores.

2.4 Corrección de Etiquetas en Machine Learning

El problema de las etiquetas incorrectas en *datasets* es un desafío bien documentado en *machine learning*. Las etiquetas incorrectas pueden degradar severamente el rendimiento de los modelos, especialmente en aplicaciones donde la precisión es crítica. La corrección de etiquetas ha sido un área de estudio creciente, con varias metodologías propuestas para mitigar este problema. Entre estas, Cleanlab ha emergido como una herramienta potente que no solo detecta etiquetas incorrectas, sino que también sugiere correcciones basadas en la probabilidad de las clases predichas [4].

El uso de Cleanlab es particularmente relevante en este estudio debido a la naturaleza de los datos de audio, donde el etiquetado manual es propenso a errores. Dado que la mayoría de los *datasets* utilizados en estudios recientes son privados [8], [9] y solo unos pocos han sido liberados públicamente [10], [11], [12], la calidad de las etiquetas es una preocupación constante, ya que depende de los propósitos de cada autor y las técnicas empleadas. La implementación de Cleanlab permite mejorar la precisión del modelo al corregir estas etiquetas erróneas, lo que maximiza el valor de los datos disponibles y mejora los resultados de la clasificación.

3. Metodología

3.1 Conjunto de Datos

El conjunto de datos utilizado en este estudio proviene del *Smart Bee Colony Monitor Dataset*, disponible públicamente en la plataforma Kaggle [10]. Este *dataset* fue recopilado con el objetivo de fomentar el desarrollo de métodos computacionales que contribuyan a la protección de las abejas y estado de salud de las colmenas, mediante la detección remota y en tiempo real, través del análisis de audio y otras características intrínsecas. Las grabaciones de audio, que forman el núcleo de este estudio, fueron obtenidas de colmenas de abejas europeas (*Apis Mellifera*) ubicadas en California, utilizando un dispositivo personalizado.

El *dataset* contiene un total de 7100 muestras de audio, cada una de las cuales es un segmento de 60 segundos. Estas muestras derivan de un conjunto original de poco más de 1200 grabaciones largas, que fueron recortadas en los segmentos de 60 segundos para su análisis. Cada archivo de audio se asocia con una serie de metadatos que incluyen la fecha, hora, ubicación, presencia de la reina, aceptación de la reina, temperatura, humedad y condiciones climáticas. Sin embargo, para este estudio, se utilizarán solo las grabaciones de audio, prescindiendo de los datos contextuales adicionales, ya que el objetivo es desarrollar un modelo que pueda ser eventualmente implementado en una aplicación móvil, utilizando únicamente el micrófono del dispositivo como entrada.

El enfoque principal del estudio es la clasificación de los estados de la reina en la colmena, a partir de los zumbidos del enjambre.

Los estados considerados son cuatro y se catalogan como:

- 0 - reina original presente
- 1 - reina no presente
- 2 - reina presente pero rechazada
- 3 - reina presente y recién aceptada

En la Fig. 1, puede apreciarse el espectrograma de una muestra de audio para cada una de estas clases. Denotando frecuencias y patrones para cada estado de interés.

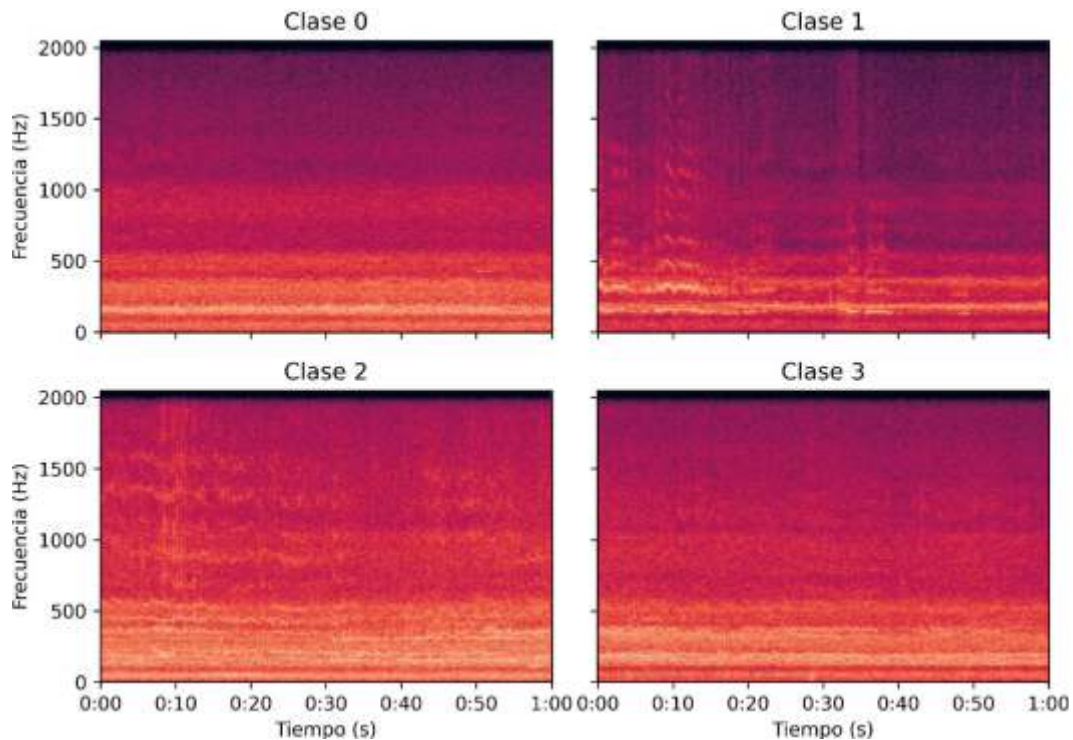


Fig. 1. Espectrogramas de los 4 estados/clases considerados.

La clasificación precisa entre estos estados es crucial para la gestión eficiente de la colmena, ya que la presencia y aceptación de la reina son indicadores fundamentales de la salud y estabilidad de la colonia. En la actualidad, la mayoría de los estudios se centra en determinar si la reina está o no en la colmena, es decir, proponen un clasificador binario [8], [13], [14]. El objetivo es poder discriminar estados más específicos, ya que la reina puede estar, pero la colonia no la acepta (caso crítico al implantar una nueva, por ejemplo). Identificar este tipo de cuestiones, sería clave para los apicultores, de modo de poder realizar las maniobras adecuadas en sus colmenas.

Una consideración importante al utilizar este conjunto de datos es el hecho de que el etiquetado de los estados de la reina fue realizado manualmente por los autores del *dataset* antes del proceso de segmentación. Esto implica que cualquier error humano cometido durante el etiquetado original se propaga a todos los segmentos derivados de esa muestra. Este es un aspecto crítico, dado el tamaño limitado del *dataset*, y subraya la importancia de utilizar técnicas avanzadas de corrección de etiquetas, para maximizar la utilidad de los datos disponibles y mejorar la precisión del modelo.

Para asegurar una evaluación robusta del modelo, se tomará en cuenta la procedencia de los fragmentos de audio al momento de dividir los datos en conjuntos de entrenamiento y prueba (70/30). Específicamente, se evitará que fragmentos derivados de la misma muestra original se

crucen entre estos conjuntos, minimizando así la posibilidad de *data leakage* y asegurando que la validación del modelo refleje de manera precisa su capacidad de generalización.

3.2 Extracción de Características

Para el análisis de los sonidos de las colmenas, se remuestrearon los audios utilizando una frecuencia de 4 kHz, seleccionada en base a estudios previos que indican que las frecuencias relevantes para la detección de estados en colmenas de abejas se encuentran típicamente en el rango de 100 Hz a 2000 Hz [3]. Esta frecuencia de remuestreo permite capturar adecuadamente las variaciones espectrales significativas sin introducir ruido de alta frecuencia, no relevante para la clasificación de los estados de la colmena y reduciendo la carga de los modelos a implementar.

Una vez remuestreados los audios, se procedió a la extracción de características utilizando los MFCCs. Para cada segmento de audio, se calcularon 50 coeficientes MFCC, utilizando un ventaneo estándar proporcionado por la biblioteca Librosa.

Los MFCCs se calcularon para ventanas de tiempo acotadas dentro de cada segmento de audio, y posteriormente, se tomó la media de cada coeficiente a lo largo del segmento. Esta aproximación produce un vector de 50 características por segmento de audio, representando la información espectral promedio del mismo. La media de coeficientes en el tiempo responde a la naturaleza del problema; se espera que las variaciones en los zumbidos no presenten cambios rápidos, lo que hace viable la simplificación del análisis espectral en ventanas temporales acotadas y mitiga efectos ambientales como impactos externos sobre la colmena.

Este enfoque, centrado en la extracción de MFCCs, resultó ser significativamente más efectivo en comparación con otros métodos de extracción de características, como los *embeddings* generados por YAMNet. La simplicidad y eficiencia computacional de los MFCCs, combinada con su efectividad en la tarea, justifican su selección como la característica principal para el modelo de clasificación de estados de la colmena.

3.3 Optimización del Pipeline de Clasificación

Para optimizar el *pipeline* de clasificación y asegurar que los modelos empleados ofrecieran un rendimiento óptimo, se evaluaron inicialmente los modelos base de PyCaret, incluyendo Logistic Regression, K Neighbors Classifier, Naive Bayes, Decision Tree Classifier, Support Vector Machines (con *kernel* lineal y radial), Random Forest, AdaBoost, Gradient Boosting Classifier, y Extra Trees Classifier, entre otros (Tabla 1).

PyCaret permitió evaluar el desempeño de cada modelo utilizando un conjunto estándar de métricas: *Accuracy*, *AUC*, *Recall*, *Precision*, *F1-Score*, *MCC* (*Matthews Correlation Coefficient*), y el tiempo de entrenamiento (TT, *Time to Train*). Estas métricas ofrecieron una visión integral del rendimiento de cada modelo, permitiendo la identificación de los más prometedores para la clasificación de los estados de la colmena.

El proceso de optimización no se limitó a los hiperparámetros de los modelos individuales, sino que se centró en la optimización global del *pipeline* de clasificación. Se llevaron a cabo experimentos que incluían la normalización de características, la aplicación de PCA (Análisis de Componentes Principales), la remoción de *outliers*, y el balanceo de clases. Cada uno de estos enfoques fue evaluado en términos de su impacto en el rendimiento global, utilizando la validación cruzada con 10 pliegues (*10-fold cross-validation*) para asegurar que los resultados fueran representativos y generalizables.

Para una visión integral del sistema desarrollado, la Fig. 2 presenta el diagrama de flujo completo. Como se puede observar, una fase crítica y distintiva de nuestra metodología es la “Etapa 2”, dedicada al refinamiento automático de las etiquetas antes de proceder al entrenamiento del modelo de clasificación final.

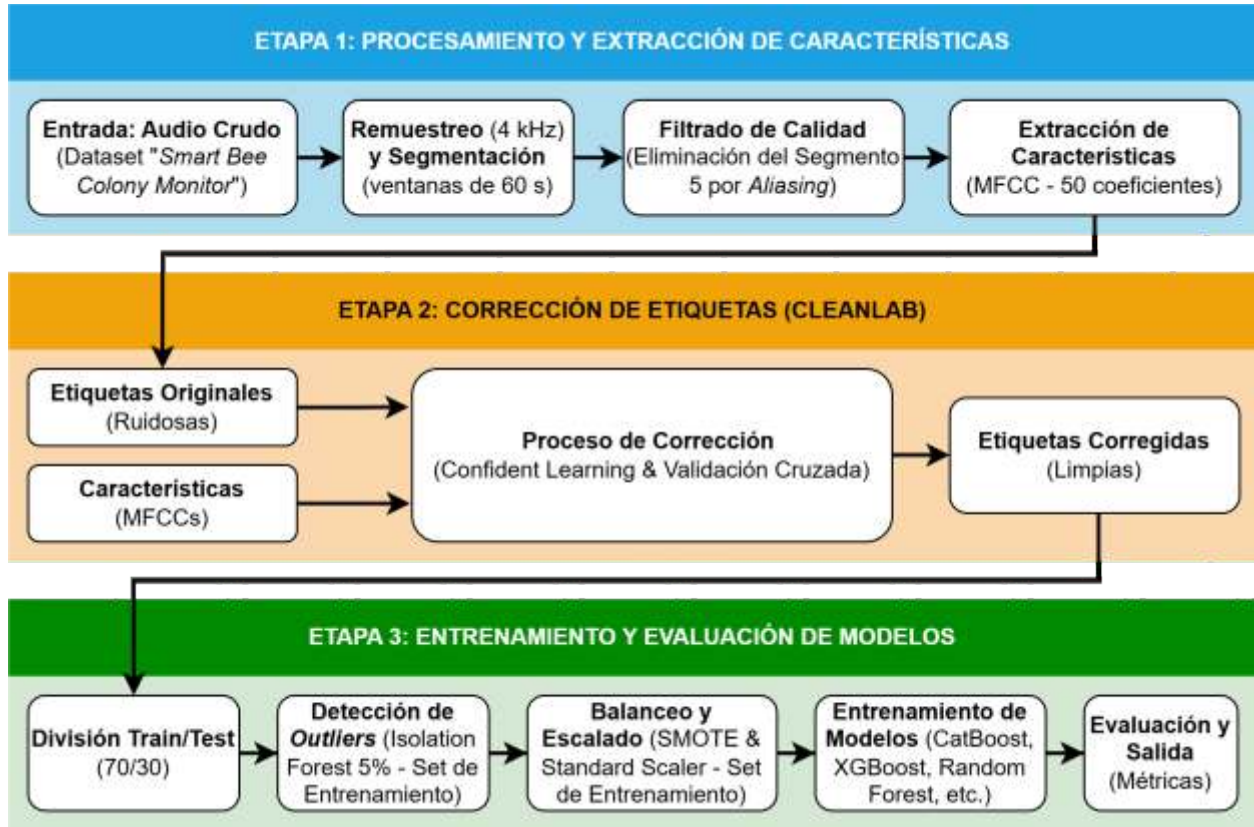


Fig. 2. Diagrama de flujo del *pipeline* metodológico propuesto.

3.4 Corrección de Etiquetas con Cleanlab

En nuestra metodología, Cleanlab se integró durante la fase del análisis exploratorio de datos. En primer lugar, se extrajeron 50 coeficientes MFCC de cada segmento de audio, que fueron utilizados como *features* para alimentar un modelo de Random Forest, utilizado como base en la implementación de Confident Learning. A través de una validación cruzada de 15 *folds*, se generaron las probabilidades de predicción para cada segmento, las cuales fueron luego analizadas por Cleanlab para identificar posibles problemas de etiquetado.

El proceso de corrección se realizó en dos niveles:

- **A nivel de segmento individual:** Se analizaron los segmentos aislados para determinar si presentaban inconsistencias que pudieran indicar errores de etiquetado debido al procesamiento o al corte de los segmentos (nativo de los autores del *dataset*).
- **A nivel de audio original:** Dado que cada audio original se segmentó en 6 partes, se implementó un criterio de corrección basado en la coincidencia de etiquetas. Si al menos tres segmentos coincidían en otra categoría diferente a la originalmente etiquetada, se consideró que el audio original había sido etiquetado erróneamente, y se procedió a corregir la etiqueta para todos sus segmentos.

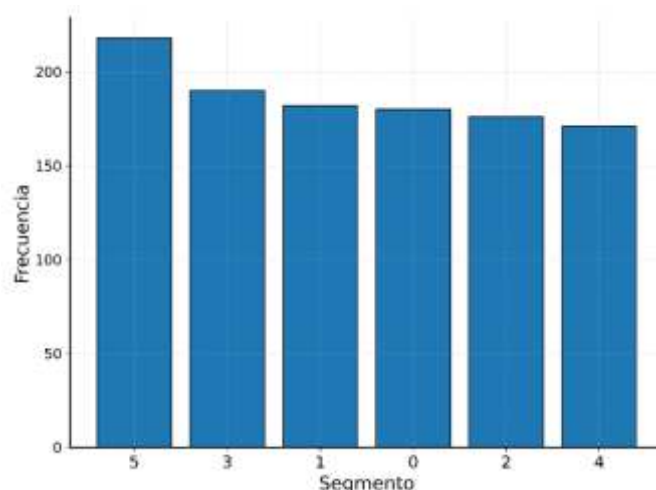


Fig. 3. Distribución de inconsistencias en los segmentos individuales.

En la Fig. 3, se observa la distribución de posibles errores de etiquetas, discriminadas por segmentos. Es de esperar, que la misma cuente con una distribución uniforme, ya que los segmentos surgen de un mismo audio de origen. Sin embargo, se denota una frecuencia superior para los segmentos número 5. En un esfuerzo por intentar entender el fenómeno, se ejemplifica en la Fig. 4, la FFT para cada uno de los segmentos de un mismo audio. Debieran ser parecidos, ya que corresponden al mismo evento dentro de la colmena, pero se denota una diferencia sustancial para el segmento 5.

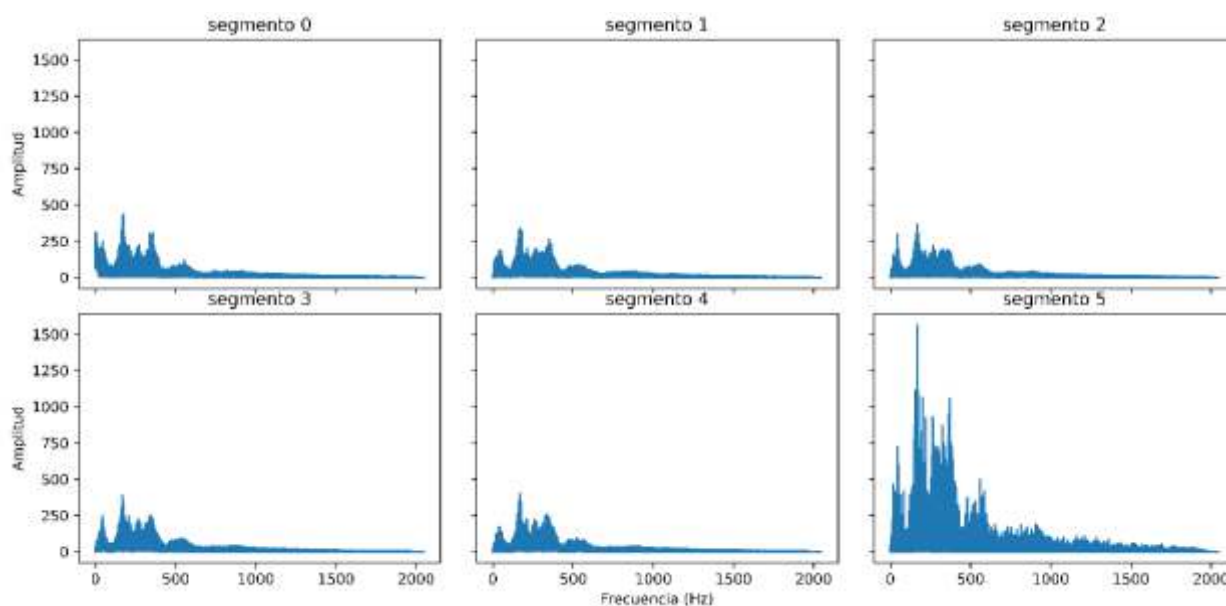


Fig. 4. FFT para los segmentos sucesivos de un mismo audio.

Este fenómeno en los segmentos 5 se observó para todos los audios, al analizarlos en detalle, es evidente que presentan efecto de *aliasing*. Se desconoce el origen de este problema debido a que no se observa en otros segmentos y proviene directamente del *dataset* original. En la Fig. 5 se observan varios ejemplos del fragmento 5 para múltiples audios y múltiples clases, todos tienen el problema de *aliasing* que se observa. Por lo tanto, se descartaron para el entrenamiento y la validación de los modelos.

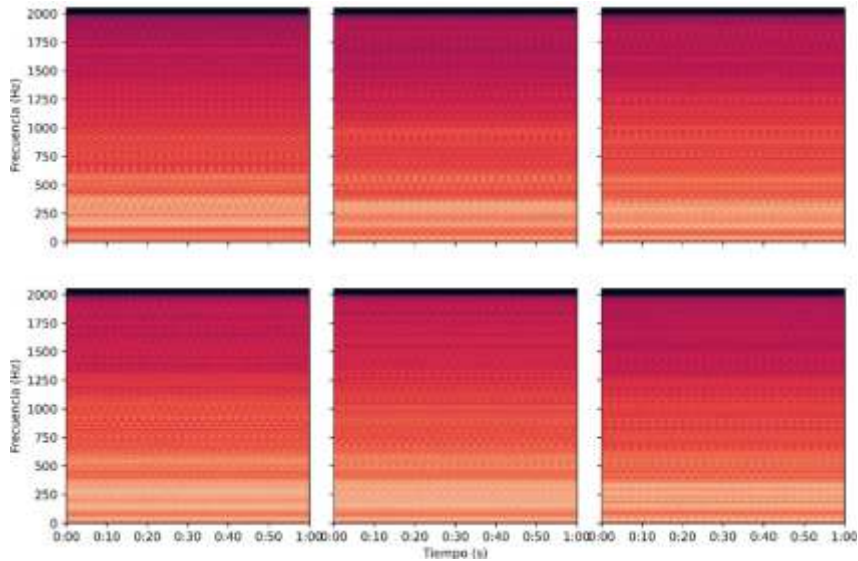


Fig. 5. Ejemplos del espectro para el segmento 5, múltiples audios y múltiples estados de la reina.

Tras aplicar Cleanlab con el criterio establecido, se redefinieron alrededor del 10% de las etiquetas de los segmentos originales. La mejora en el rendimiento de los modelos se analizará en detalle en la sección de resultados, confirmando la existencia de errores en el etiquetado inicial mediante la comprobación empírica entre el desempeño de los modelos en ambos casos (etiquetas originales y corregidas).

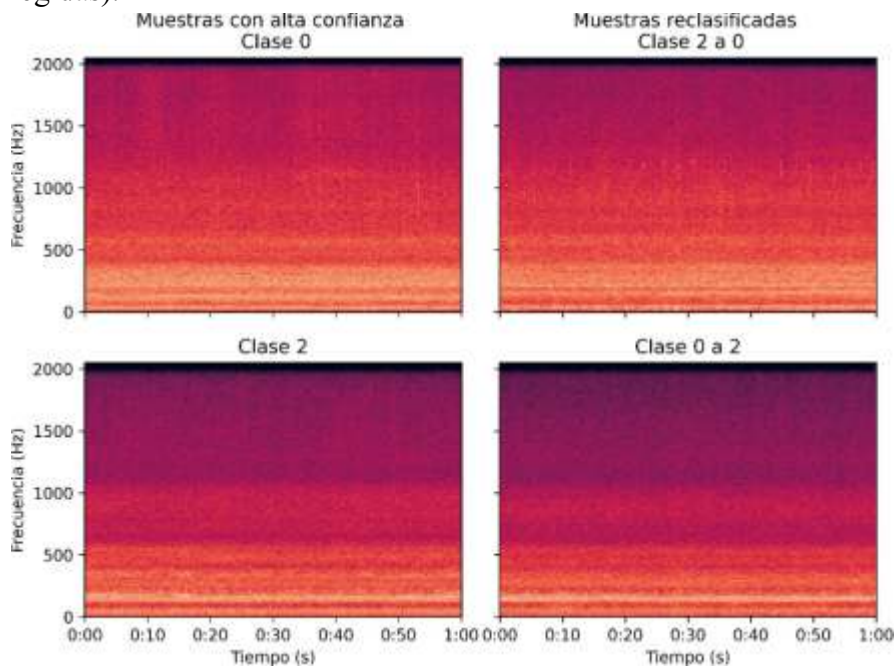


Fig. 6. Ejemplos del espectro para audios con etiquetas corregidas.

En la Fig. 6 se muestran dos casos de corrección de etiquetas. Las figuras de la izquierda (primera columna) se tomaron como referencia (alta confianza según Cleanlab) para las clases 0 y 2. En las figuras de la derecha (segunda columna) se observan espectrogramas cuyas etiquetas fueron reclasificadas, pudiendo visualizar coincidencias entre las distintas bandas en frecuencias respecto a la supuesta clase original y la corregida. En los ejemplos reclasificados, la leyenda “Clase 2 a 0” significa que; originalmente la muestra era de clase 2 pero se reclasificó como de clase 0, ídem para el otro caso.

4. Resultados y Discusión

4.1 Resultados de la Clasificación Inicial

En esta sección, se presentan los resultados obtenidos al aplicar una serie de modelos de clasificación a los datos iniciales, utilizando como características los descriptores MFCC extraídos previamente. Es importante destacar que estos resultados se obtuvieron antes de realizar cualquier optimización del *pipeline* o corrección de etiquetas, lo que nos permitirá establecer una línea base para contrastar con los resultados mejorados en etapas posteriores.

Para la evaluación inicial, se emplearon todos los modelos de clasificación rápidos disponibles en PyCaret, lo que incluye una selección estándar de algoritmos supervisados de clasificación. Los modelos se entrenaron y evaluaron utilizando los valores crudos de los 50 coeficientes MFCC, sin aplicar técnicas de preprocesamiento.

La Tabla 1 resume las métricas clave obtenidas para cada modelo:

Tabla 1. Desempeño base de los modelos, sin optimización del *pipeline*

Modelo/Métrica	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>	<i>Prec.</i>	<i>F1</i>	<i>MCC</i>	TT (S)
<i>Extra Trees Classifier</i>	0.6296	0.8198	0.6296	0.6369	0.6005	0.4210	0.0340
<i>CatBoost Classifier</i>	0.6287	0.8231	0.6287	0.6340	0.6053	0.4270	0.7090
<i>Random Forest Classifier</i>	0.6231	0.8142	0.6231	0.6264	0.5946	0.4112	0.0490
<i>Light Gradient Boosting Machine</i>	0.6212	0.8143	0.6212	0.6246	0.5971	0.4151	0.3360
<i>Extreme Gradient Boosting</i>	0.6186	0.8108	0.6186	0.6263	0.5970	0.4113	0.0520
<i>Gradient Boosting Classifier</i>	0.6071	0.0000	0.6071	0.5851	0.5598	0.3720	0.2770
<i>K Neighbors Classifier</i>	0.6052	0.7732	0.6052	0.6202	0.5868	0.3978	0.0130
<i>Ridge Classifier</i>	0.5977	0.0000	0.5977	0.4486	0.4836	0.3076	0.0060
<i>Linear Discriminant Analysis</i>	0.5965	0.0000	0.5965	0.4619	0.4876	0.3149	0.0050
<i>Logistic Regression</i>	0.5960	0.0000	0.5960	0.4663	0.4882	0.3104	0.0110
<i>Naive Bayes</i>	0.5836	0.7139	0.5836	0.5134	0.5013	0.3131	0.0060
<i>Quadratic Discriminant Analysis</i>	0.5833	0.0000	0.5833	0.5216	0.5138	0.3178	0.0060
<i>Decision Tree Classifier</i>	0.5557	0.6776	0.5557	0.5786	0.5502	0.3363	0.0080
<i>Ada Boost Classifier</i>	0.5524	0.0000	0.5524	0.5188	0.4996	0.2725	0.0270
<i>Dummy Classifier</i>	0.5100	0.5000	0.5100	0.2601	0.3445	0.0000	0.0060
<i>SVM - Linear Kernel</i>	0.4533	0.0000	0.4533	0.4294	0.4153	0.1510	0.0130

En esta evaluación inicial, el Extra Trees Classifier mostró el mejor rendimiento con una *Accuracy* de 0,6296 y un *AUC* de 0,8198. Le siguieron de cerca el CatBoost Classifier y el Random Forest Classifier con *Accuracy* de 0,6287 y 0,6231, respectivamente. Estos resultados reflejan un rendimiento modesto, lo que sugiere que el *pipeline* y los datos requieren optimización para mejorar la capacidad predictiva de los modelos, que los *features* utilizados no presentan la información necesaria o que el problema es difícil de resolver para este tipo de modelos.

Por otro lado, modelos como SVM - Linear Kernel y Dummy Classifier obtuvieron los peores resultados, con *Accuracy* de 0,4533 y 0,5100, respectivamente. Estos valores refuerzan la idea de mejorar el *pipeline* para obtener resultados más robustos.

Estos resultados iniciales proporcionan una línea base esencial, que nos permitirá evaluar de manera efectiva el impacto de las optimizaciones del *pipeline* y la corrección de etiquetas, las cuales serán discutidas en las siguientes subsecciones.

4.2 Optimización del Pipeline

Para mejorar los resultados obtenidos en la clasificación inicial, se procedió a optimizar el *pipeline* de procesamiento de los datos. La mejor configuración identificada consistió en tres pasos clave:

- **Remoción de *Outliers*:** Se empleó el algoritmo de Isolation Forest con un umbral de contaminación del 5% para eliminar muestras atípicas que podrían estar afectando negativamente la precisión del modelo. La remoción de estos *outliers* resultó en una mejora notable en la precisión global, dado que permitió que los modelos se concentraran en patrones más representativos de las clases. En particular, se observó un aumento en la precisión de los modelos más complejos, como el Random Forest y el CatBoost (Tabla 2), lo que subraya la importancia de eliminar datos ruidosos (se encontraron muestras con superposición de sonidos de maquinaria agrícola, por dar un ejemplo).
- **Balanceo de Clases:** Dado que las clases no estaban equilibradas, se utilizó la técnica SMOTE (*Synthetic Minority Over-sampling Technique*) con 5 vecinos cercanos y estrategia de muestreo automática. Este paso fue crucial para mejorar el *Recall*, especialmente en las clases minoritarias. Al generar muestras sintéticas de estas clases, se evitó que los modelos se sesgaran hacia las clases más representadas, lo que permitió una mejor generalización y un incremento en el rendimiento de modelos como el CatBoost y el Gradient Boosting, donde el *Recall* mostró una mejora significativa (Tabla 2).
- **Normalización:** Finalmente, se aplicó la normalización de los datos mediante *Standard Scaler*, que estandarizó los descriptores MFCC, asegurando que cada uno tuviera una media de 0 y una desviación estándar de 1. Esta normalización es particularmente beneficiosa para modelos que dependen de la medición de distancias entre puntos en el espacio de características, como K Neighbors Classifier (KNN), de modo de permitir explorar soluciones en igualdad de condiciones respecto a modelos no sensibles (basados en árboles, por ejemplo).

El efecto de esta etapa se observa en la Tabla 2.

Tras implementar este *pipeline*, los resultados mostraron mejoras significativas en comparación con la clasificación inicial (Tabla 1). Por ejemplo, el CatBoost Classifier, que ya era uno de los mejores en el *baseline*, mejoró su *Accuracy* de 0,6287 a 0,8222 y su *F1-Score* de 0,6053 a 0,8196. De igual manera, el Random Forest Classifier experimentó un aumento en su *Accuracy* de 0,6231 a 0,7884, demostrando que la optimización del *pipeline* permitió a los modelos capturar patrones más significativos en los datos.

Tabla 2. Desempeño de los modelos, con optimización del *pipeline*

Modelo/Métrica	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>	<i>Prec.</i>	<i>F1</i>	<i>MCC</i>	TT (S)
<i>CatBoost Classifier</i>	0.8222	0.9634	0.8222	0.8442	0.8196	0.7431	4.3480
<i>Extreme Gradient Boosting</i>	0.8042	0.9573	0.8042	0.8249	0.7997	0.7143	0.3100
<i>Quadratic Discriminant Analysis</i>	0.8038	0.0000	0.8038	0.8297	0.7916	0.7119	0.0220
<i>Light Gradient Boosting Machine</i>	0.7963	0.9534	0.7963	0.8157	0.7905	0.7010	0.5480
<i>Random Forest Classifier</i>	0.7884	0.9498	0.7884	0.8104	0.7785	0.6872	0.2180
<i>Extra Trees Classifier</i>	0.7819	0.9513	0.7819	0.8046	0.7682	0.6766	0.0690
<i>Gradient Boosting Classifier</i>	0.7515	0.0000	0.7515	0.7750	0.7475	0.6365	5.1470
<i>K Neighbors Classifier</i>	0.7474	0.8970	0.7474	0.8015	0.7474	0.6525	0.1630
<i>Logistic Regression</i>	0.7428	0.0000	0.7428	0.7838	0.7439	0.6352	0.2210
<i>SVM - Linear Kernel</i>	0.7376	0.0000	0.7376	0.7709	0.7382	0.6206	0.0300
<i>Linear Discriminant Analysis</i>	0.7337	0.0000	0.7337	0.7754	0.7357	0.6236	0.0200
<i>Ridge Classifier</i>	0.7280	0.0000	0.7280	0.7623	0.7279	0.6096	0.0210
<i>Decision Tree Classifier</i>	0.6390	0.7486	0.6390	0.6769	0.6388	0.4765	0.0610
<i>Naive Bayes</i>	0.5922	0.7805	0.5922	0.6082	0.5672	0.3740	0.0200
<i>Ada Boost Classifier</i>	0.5862	0.0000	0.5862	0.6598	0.5973	0.4171	0.2400
<i>Dummy Classifier</i>	0.1338	0.5000	0.1338	0.0179	0.0316	0.0000	0.0240

4.3 Impacto de la Corrección de Etiquetas

La corrección de etiquetas llevada a cabo, tuvo un impacto significativo en el rendimiento de los modelos evaluados. Al aplicar el *pipeline* optimizado que incluye la remoción de *outliers*, el balanceo de clases, y la normalización de los descriptores MFCC; se observaron mejoras notables en las diversas métricas de clasificación. Los resultados se muestran a continuación:

Tabla 3. Desempeño de los modelos, con optimización del *pipeline* y corrección de etiquetas

Modelo/Métrica	<i>Accuracy</i>	<i>AUC</i>	<i>Recall</i>	<i>Prec.</i>	<i>F1</i>	<i>MCC</i>	TT (S)
<i>CatBoost Classifier</i>	0.8469	0.9704	0.8469	0.8617	0.8418	0.7768	4.3150
<i>Light Gradient Boosting Machine</i>	0.8400	0.9662	0.8400	0.8529	0.8332	0.7640	0.5430
<i>Extreme Gradient Boosting</i>	0.8364	0.9677	0.8364	0.8510	0.8301	0.7593	0.2660
<i>Random Forest Classifier</i>	0.8232	0.9650	0.8232	0.8401	0.8125	0.7387	0.2120
<i>Extra Trees Classifier</i>	0.8131	0.9649	0.8131	0.8285	0.7976	0.7229	0.0670
<i>Quadratic Discriminant Analysis</i>	0.8073	0.0000	0.8073	0.8383	0.7882	0.7162	0.0230
<i>Gradient Boosting Classifier</i>	0.8069	0.0000	0.8069	0.8254	0.8021	0.7175	4.9320
<i>SVM - Linear Kernel</i>	0.7910	0.0000	0.7910	0.8139	0.7872	0.6975	0.0320
<i>Logistic Regression</i>	0.7843	0.0000	0.7843	0.8123	0.7836	0.6893	0.0350
<i>K Neighbors Classifier</i>	0.7803	0.9078	0.7803	0.8261	0.7837	0.6923	0.0280
<i>Ridge Classifier</i>	0.7663	0.0000	0.7663	0.7960	0.7615	0.6640	0.0210
<i>Linear Discriminant Analysis</i>	0.7661	0.0000	0.7661	0.8051	0.7639	0.6685	0.0220
<i>Decision Tree Classifier</i>	0.6982	0.7869	0.6982	0.7249	0.6966	0.5592	0.0630
<i>Naive Bayes</i>	0.6618	0.8290	0.6618	0.6716	0.6463	0.4872	0.0200
<i>Ada Boost Classifier</i>	0.6524	0.0000	0.6524	0.7151	0.6635	0.5085	0.2300
<i>Dummy Classifier</i>	0.1207	0.5000	0.1207	0.0146	0.0260	0.0000	0.0240

Comparando los resultados obtenidos antes (Tabla 2) y después de la corrección de etiquetas (Tabla 3), se evidencia un incremento generalizado en todas las métricas y para la mayoría de los modelos. Por ejemplo, el CatBoost Classifier, que ya mostraba un buen rendimiento en la evaluación previa, mejoró aún más su *Accuracy* de 0,8222 a 0,8469 y su MCC de 0,7431 a 0,7768. De igual manera, modelos como Light Gradient Boosting Machine y Extreme Gradient Boosting mostraron incrementos en el AUC y *F1-Score*, reflejando un mejor equilibrio entre la precisión y el *Recall* tras la corrección de etiquetas.

El impacto específico de la corrección de etiquetas fue también notable en modelos lineales y basados en árboles. El Random Forest Classifier, por ejemplo, experimentó un aumento en su *Recall* y Precisión, logrando una *F1-Score* superior, lo que vuelve a reforzar la idea de que la limpieza de etiquetas contribuyó a una mejor discriminación entre las clases.

Además, considerando ya un factor no tan exhaustivo, la corrección afectó positivamente el TT en algunos casos, lo que sugiere que la eliminación de ruido en los datos puede llevar a un proceso de entrenamiento más eficiente, aunque en otros casos, como en el Gradient Boosting Classifier, el tiempo de entrenamiento aumentó ligeramente, lo que podría estar relacionado con la necesidad de ajustar mejor el modelo a los datos, ahora más precisos.

Tomando el modelo CatBoost Classifier, que es el que presenta los mejores resultados (tanto en la Tabla 2 como en la Tabla 3), se tiene que su desempeño sobre el conjunto de pruebas, para cada una de las clases, se corresponde con lo expuesto en las matrices de confusión de la Fig. 7.

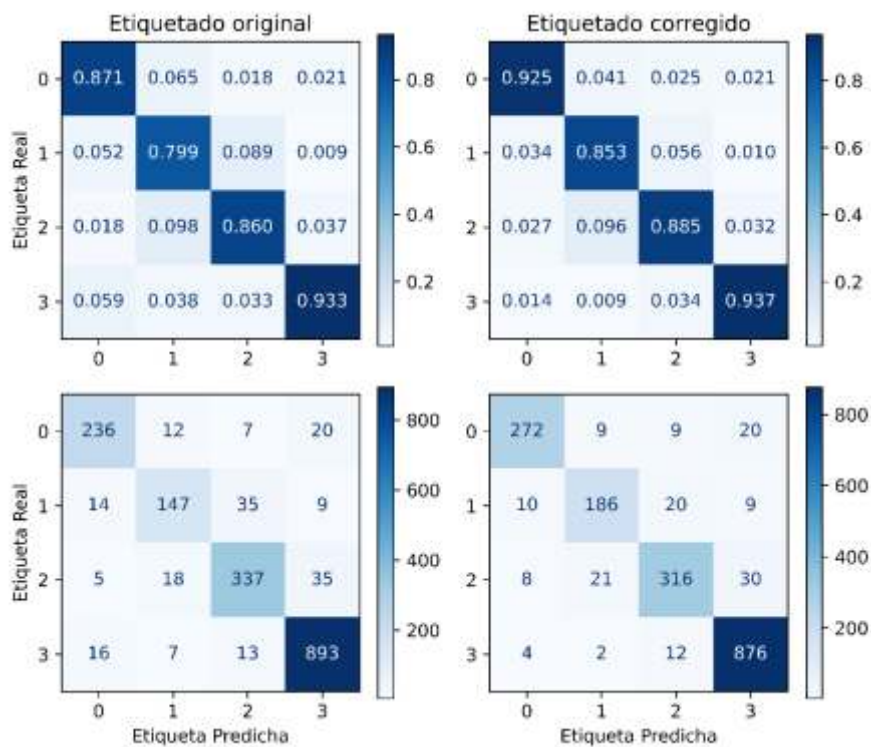


Fig. 7. Matrices de confusión sobre el set de pruebas, para etiquetas originales y etiquetas corregidas.

Las matrices de confusión de la Fig. 7 revelan que, tras la corrección de etiquetas, el modelo no solo mejora su precisión general, sino que también disminuye significativamente los errores de clasificación entre clases cercanas. Por ejemplo, el número de errores al clasificar la etiqueta 1 disminuye considerablemente, la precisión de la clase aumenta de 0,799 a 0,855 (en la primera

columna de la Fig. 7 se encuentran los valores porcentuales, mientras que, en la segunda, por cantidades de observaciones predichas). En grandes rasgos, se percibe un aumento del *Accuracy* global (sobre el set de prueba) del 0,894 al 0,915. Esta mejora sugiere que las etiquetas corregidas permiten al modelo capturar patrones más consistentes y representativos de las características subyacentes, reduciendo el ruido y los errores en las predicciones.

5. Conclusiones y Trabajo Futuro

5.1 Conclusiones Principales

Este estudio ha demostrado de manera efectiva la importancia de la optimización de *pipelines* y la corrección de etiquetas en modelos de clasificación aplicados al análisis de sonidos en la apicultura. A diferencia de trabajos previos que se centraron en la clasificación binaria del estado de las colmenas de abejas (como la presencia o ausencia de la reina), este enfoque ha abordado con éxito la clasificación multiclase, lo que permite una identificación más específica y detallada de los diferentes estados de la colmena. Esta metodología, además, se ha desarrollado utilizando datos de acceso libre, lo que subraya su potencial para ser replicada y aplicada en diferentes contextos.

La implementación de un *pipeline* optimizado, que incluye la remoción de *outliers*, el balanceo de clases y la normalización de los datos, ha resultado en una mejora sustancial en las métricas de desempeño de los modelos. Específicamente, la aplicación de la corrección de etiquetas mediante Cleanlab ha permitido un incremento significativo en la precisión global del modelo, con un *Accuracy* que aumentó del 0,894 al 0,915 en el conjunto de prueba. Estos resultados subrayan la relevancia de abordar tanto la calidad de los datos como la robustez del *pipeline* en la construcción de modelos predictivos confiables para la apicultura.

El modelo desarrollado tiene un potencial significativo para ser implementado en dispositivos portátiles, ya que su capacidad para realizar inferencias no requiere un alto poder de cómputo. Esto abre la posibilidad de integrarlo en una aplicación móvil, que podría ser una herramienta invaluable para los apicultores. Dicha aplicación permitiría a los apicultores obtener información en tiempo real sobre el estado de cada una de sus colmenas, facilitando la toma de decisiones informadas y mejorando la gestión y el bienestar de las abejas.

5.2 Limitaciones del Estudio y Desafíos Futuros

A pesar de las mejoras observadas, el enfoque presenta algunas limitaciones. La principal limitación radica en la dependencia del modelo de la calidad de los datos y las etiquetas. Si bien la corrección de etiquetas ha demostrado ser efectiva en este caso, en situaciones donde las etiquetas son inherentemente ambiguas o difíciles de definir, podría ser necesario recurrir a expertos humanos para verificar o ajustar las correcciones automatizadas.

En cuanto a desafíos futuros, una limitación a considerar es la posible sobre adaptación del modelo a las características específicas del conjunto de datos utilizado (número limitado de colmenas en una única región geográfica). Aunque los resultados obtenidos son prometedores, la generalización del modelo a diferentes poblaciones de datos o a otras especies de abejas debe ser cuidadosamente evaluada.

La metodología propuesta utiliza consistencia estadística para corregir etiquetas, asumiendo que los errores de anotación son aleatorios o sistemáticamente divergentes de la distribución principal de la clase. Sin embargo, no se ha realizado una validación de campo (*Ground Truth* biológico) para

confirmar que las muestras re-etiquetadas corresponden, en efecto, al estado biológico real de la colmena. Por tanto, la mejora en las métricas debe interpretarse como un aumento en la coherencia interna del modelo frente a los datos acústicos, un paso necesario, pero no suficiente para la certificación biológica absoluta.

El modelo actual, que ha demostrado su efectividad en la identificación de cuatro clases distintas en el análisis del sonido de colmenas, abre nuevas posibilidades para futuras investigaciones. Una de las direcciones más prometedoras sería utilizar este modelo para etiquetar automáticamente conjuntos de datos abiertos que, hasta ahora, solo han sido categorizados de manera binaria. Al aplicar esta metodología, se podría obtener información más detallada y específica sobre los estados de las colmenas en una variedad de entornos, lo que contribuiría a una comprensión más profunda y contextualizada del comportamiento de las abejas.

Además, con un volumen significativamente mayor de datos etiquetados de manera precisa, sería viable explorar la implementación de modelos basados en *deep learning*. Estos modelos, conocidos por su capacidad para captar y aprender patrones complejos, podrían mejorar aún más la precisión y la capacidad del sistema para reconocer las características distintivas de cada clase. Tal enfoque podría llevar a avances importantes en la clasificación automatizada y la monitorización en la apicultura, con aplicaciones que podrían extenderse a otras áreas del análisis de sonido en la biología.

En conclusión, este trabajo demuestra que el refinamiento algorítmico de etiquetas es una estrategia viable y costo-efectiva para mejorar modelos de bioacústica cuando la recolección de nuevos datos de campo es inviable. Proponemos este *pipeline* no como un sustituto de la validación de campo, sino como un protocolo de pre-procesamiento estándar para maximizar la fiabilidad de los datos en la apicultura de precisión.

Referencias

- [1] R. A. Martínez Sarmiento, N. C. Ortega Flórez, W. D. Maldonado Quintero, y R. E. V. R. y Otros, *Manual técnico de apicultura: abeja (Apis mellifera)*. Corporación colombiana de investigación agropecuaria - AGROSAVIA, 2012. doi: 10.21930/agrosavia.manual.2012.1.
- [2] C. Uthoff, M. N. Homsí, y M. von Bergen, «Acoustic and vibration monitoring of honeybee colonies for beekeeping-relevant aspects of presence of queen bee and swarming», *Comput. Electron. Agric.*, vol. 205, p. 107589, feb. 2023, doi: 10.1016/j.compag.2022.107589.
- [3] M. Abdollahi, P. Giovenazzo, y T. H. Falk, «Automated Beehive Acoustics Monitoring: A Comprehensive Review of the Literature and Recommendations for Future Work», *Appl. Sci.*, vol. 12, n.º 8, Art. n.º 8, ene. 2022, doi: 10.3390/app12083920.
- [4] C. G. Northcutt, L. Jiang, y I. L. Chuang, «Confident Learning: Estimating Uncertainty in Dataset Labels», 21 de agosto de 2022, *arXiv*: arXiv:1911.00068. [En línea]. Disponible en: <http://arxiv.org/abs/1911.00068>
- [5] M. Ali, *PyCaret: An open source, low-code machine learning library in Python* 2020. [En línea]. Disponible en: <https://www.pycaret.org>
- [6] D. Howard, O. Duran, G. Hunter, y K. Stebel, «Signal processing the acoustics of honeybees (APIS MELLIFERA) to identify the “queenless” state in Hives», *Proc. Inst. Acoust.*, vol. 35, pp. 290-297, ene. 2013.
- [7] S. Hershey et al., «CNN Architectures for Large-Scale Audio Classification», 10 de enero de 2017, *arXiv*: arXiv:1609.09430. doi: 10.48550/arXiv.1609.09430.
- [8] S. Ruvinga, G. Hunter, O. Duran, y J.-C. Nebel, «Identifying Queenlessness in Honeybee Hives from Audio Signals Using Machine Learning», *Electronics*, vol. 12, n.º 7, Art. n.º 7, ene. 2023, doi: 10.3390/electronics12071627.

- [9] A. Robles-Guerrero, T. Saucedo-Anaya, E. González Ramírez, y C. E. Galván-Tejada, «Frequency Analysis of Honey Bee Buzz for Automatic Recognition of Health Status: A Preliminary Study» *Res. Comput. Sci.* , vol. 142, n.º 1, pp. 89-98, dic. 2017, doi: 10.13053/rcs-142-1-9.
- [10] «Smart Bee Colony Monitor: Clips of Beehive Sounds». Accedido: 7 de agosto de 2023. [En línea]. Disponible en: <https://www.kaggle.com/datasets/annajyang/beehive-sounds/data>
- [11] A. Robles-Guerrero, T. Saucedo-Anaya, E. Gonzalez, y J. I. de la Rosa, «Queenless honeybee acoustic patterns», vol. 1, ago. 2022, doi: 10.17632/t9prmbmdfn.1.
- [12] I. Nolasco, A. Terenzi, S. Cecchi, S. Orcioni, H. L. Bear, y E. Benetos, «Audio-based identification of beehive states», 15 de febrero de 2019, *arXiv*: arXiv:1811.06330. doi: 10.48550/arXiv.1811.06330.
- [13] L. Barbisan y F. Riente, «Machine Learning Framework for the Acoustic Detection of the Queen Bee Presence», en *Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023* , Turin, Italy: European Acoustics Association, ene. 2024, pp. 4347-4350. doi: 10.61782/fa.2023.1309.
- [14] T. Cejrowski, J. Szymanski, H. Mora, y D. Gil, «Detection of the Bee Queen Presence Using Sound Analysis», 2018, pp. 297-306. doi: 10.1007/978-3-319-75420-8_28.

Es una publicación
de la Universidad Nacional de Misiones,
Facultad de Ingeniería. Argentina.



**UNIVERSIDAD
NACIONAL
DE MISIONES**



**FACULTAD
DE INGENIERÍA
UNaM**