

ISSN: en trámite

Vol. 1, Núm 2. Enero-Junio 2022

IDEAS EN CIENCIAS DE LA INGENIERÍA

Facultad de Ingeniería. Universidad Autónoma del Estado de México

Los rellenos sanitarios como
fuente de contaminación del agua subterránea
por infiltración de lixiviados:
efectos en la salud y el medio ambiente

Efectos ambientales del drenaje ácido de mina
y su tratamiento

Comparación entre MaxEnt y SVM
como clasificadores multiclase.
Caso de estudio: Río Amacuzac

Electropolimerización de polipirrol usando
electrodos de titanio

Comportamiento de algoritmos
de sobre-muestreo en Big Data





Universidad Autónoma del Estado de México

Equipo Editorial

Directora

Rosa María Valdovinos Rosas

Editora en Jefe

Ruth Hernández Pérez

Coordinadora de Publicaciones

Mercedes Lucero Chávez

Colaboradores Editoriales

Brenda Pichardo Lewenstein

Javier Salas García

Comité Editorial

Dr. Rigoberto Martínez Méndez

Dra. Lourdes Loza Hernández

Facultad de Ingeniería, Universidad Autónoma del Estado de México

Dra. Angélica Guzmán Ponce

Universidad Politécnica de Valencia, España

Editor Invitado

Dr. Roberto Alejo Eleuterio

Tecnológico Nacional de México

IDEAS EN CIENCIAS DE LA INGENIERÍA, año 3 , Vol. 2, núm. 1, enero-junio 2022, es una Publicación semestral editada por la Universidad Autónoma del Estado de México, Instituto Literario 100 Ote., Colonia Centro, Toluca, Estado de México, C.P. 50000, Tels. 7222-14-08-55, <https://hemeroteca.uaemex.mx/index.php/ideasingeneria>, ideasencienciai_fi@uaemex.mx. Directora: Rosa María Valdovinos Rosas, Facultad de Ingeniería, Ciudad Universitaria, Cerro de Coatepec s/n, Toluca, Estado de México, C.P. 50130, Tel. 7222140855. Reserva de Derechos al Uso Exclusivo no. 04-2017-072110111500-203, ISSN: en trámite, ambos otorgados por el Instituto Nacional del Derecho de Autor (INDAUTOR). Responsables de la última actualización de este número, Mercedes Lucero Chávez y Ruth Hernández Pérez, Facultad de Ingeniería. Diseño de portada y Diseño Editorial, Ruth Hernández Pérez. Fecha de última modificación: 1 de junio de 2022.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación. Se autoriza la reproducción total o parcial del contenido aquí publicado sin fines de lucro, siempre y cuando no se modifique, se cite la fuente completa y su dirección electrónica en términos de licencia aplicable.

Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional (CC BY-NC).

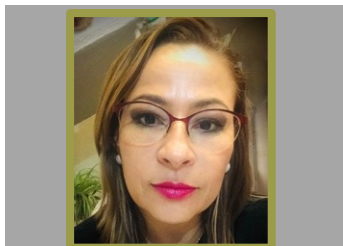
Hecho en México, Universidad Autónoma del Estado de México (UAEMéx), 2022.



CONTENIDO

	Pág.
EDITORIAL	
Rosa María Valdovinos Rosas	2
 Los rellenos sanitarios como fuente de contaminación del agua subterránea por infiltración de lixiviados: efectos en la salud y el medio ambiente. Miriam A. García-Colindres - Luis Antonio Castillo-Suárez - Carolina Álvarez Bastida - Ivonne Linares-Hernández - Verónica Martínez-Miranda	4
 Efectos ambientales del drenaje ácido de mina y su tratamiento Romel González-Hernández, Verónica Martínez-Miranda, Ivonne Linares- Hernández, Raúl Cortés-Martínez, Luis Antonio Castillo-Suárez	16
 Comparación entre MaxEnt y SVM como clasificadores multiclase. Caso de estudio: Río Amacuzac Yelena Méndez Torres, Carlos Alberto Mastachi Loza, Ricardo Arévalo Mejía, María de Lourdes Ruíz Gómez, Miguel Ángel Gómez Albores, Humberto Mejía Mojica	29
 Electropolimerización de polipirrol usando electrodos de titanio R. Ramírez, J.C. Palacios, G.J. Cruz, M.G. Olayo, E. Colín-Orozco, F.G. Flores-Nava, M. González-Torres, L.M. Gómez-Jiménez, R. Valdivia- Barrientos	42
 Comportamiento de algoritmos de sobre-muestreo en Big Data Angélica Guzmán Ponce, Cesar Ferri, Salvador Sánchez, José Raymundo Marcial-Romero	53

EDITORIAL



Rosa María Valdovinos Rosas
Universidad Autónoma del Estado de México

rvaldovinosr@uaemex.mx

Rosa María Valdovinos Rosas es Doctora en Ciencias Computacionales, miembro del Sistema Nacional de Investigadores SNI Nivel II. Miembro de la Academia Mexicana de Computación (AMEXCOMP) y de la Red de Inteligencia Computacional Aplicada (RedICA). Reguladora de la Red Tripartita de Inteligencia Artificial entre México (UAEM, ITT) y España (UJI). Participa en el fortalecimiento y consolidación de la comunidad científica mediante la dirección de tesis de nivel licenciatura y posgrado, así como en la divulgación del conocimiento y la ciencia en eventos académico-científicos a nivel nacional e internacional.

Además, cuenta con una constante publicación de artículos científicos en revistas especializadas, capítulos de libro con editoriales de prestigio y artículos de divulgación en revistas de gran alcance.

La creación y aplicación de ciencia y tecnología a la solución de problemas en México y el mundo son el pilar del desarrollo de los pueblos y un medio para disminuir la desigualdad social. El estudio de teorías y desarrollo de nuevos métodos, así como su aplicación, requieren de la interacción de múltiples disciplinas, aspecto que se ve reflejado en los trabajos inéditos que componen el presente número.

La preservación del medio ambiente y el estudio de los factores que inciden de forma negativa en él, son de gran interés en la comunidad científica a nivel mundial. Atendiendo al gran crecimiento de la población, los desechos sanitarios y la contaminación que estos producen así como su manejo, son un aspecto importante a estudiar. Al respecto, analizar el grado de contaminación de agua subterránea derivado de los rellenos sanitarios es el tema central abordado por los autores en el artículo “Los rellenos sanitarios como fuente de contaminación del agua subterránea por infiltración de lixiviados: efectos en la salud y el medioambiente” Miriam A. García-Colindres, Luis Antonio Castillo-Suárez, Carolina Álvarez Bastica, Ivonne Linares-Hernández y Verónica Martínez-Miranda.

Por otro lado, Romel González-Hernández, Verónica Martínez-Miranda, Ivonne Linares-Hernández, Raúl Cortés-Martínez, Luis Antonio Castillo-Suarez presentan el estudio “Efectos ambientales del drenaje ácido de mina y su tratamiento” en el que analizan otra forma de contaminación del agua subterránea, la producida por el drenaje de ácido de mina. Este tema es de suma importancia, dado que una de las mayores fuentes de riqueza a base de piedras preciosas y la obtención de minerales de uso común es la derivada de la industria minera. Pese a los beneficios que esta industria brinda, es importante analizar los efectos negativos en la vida acuática y los cuerpos de agua derivados del proceso de extracción.

En el artículo “Electropolimerización de polipirrol usando electrodos de titanio” R. Ramírez, J.C. Palacios, G.J. Cruz, M.G. Olayo, E. Colín-Orozco, F.G. Flores-Nava, M. González-Torres, L. M. Gómez-Jiménez, R. Valdivia Barrientos, se presenta un trabajo experimental en el que los autores muestran a detalle la sintetización de películas de polipirrol por síntesis electroquímica, reportando los resultados obtenidos del análisis elemental y análisis estructural.

La segunda parte de los artículos presentados en este número se centra en el estudio de algoritmos de *machine learning* aplicados a dos casos de estudio: el Río Amacuzac y el fenómeno del Big Data. En el primero de ellos “Comparación entre MaxEnt y SVM como clasificadores multiclase. Caso de estudio: Río Amacuzac” los autores Yelena Méndez Torres, Carlos Alberto Mastachi Loza, Ricardo Arévalo Mejía, María de Lourdes Ruíz Gómez, Miguel Ángel Gómez Albores, Humberto Mejía Mojica, realizan un estudio comparativo del rendimiento obtenido por dos de los algoritmos más populares de *machine learning* en tareas de clasificación utilizando imágenes de percepción remota del área del Río Amacuzac.

Por último, en el artículo “Comportamiento de algoritmos de sobre-muestreo en *Big Data*” de Angélica Guzmán-Ponce, César Ferri Ramírez, Jose Salvador Sánchez Garreta y J. Raymundo Marcial-Romero, se aborda un problema de gran interés en contextos del *Big Data*. La creciente generación de datos y su inherente desequilibrio han sido problemas de estudio en áreas como minería de datos y reconocimiento de patrones. No obstante, ahora en escenarios tan desafiantes como los ofrecidos por el *Big Data*, adaptar a grandes volúmenes de datos las estrategias, técnicas y métodos exitosos ampliamente utilizados y probados con pocos datos es el tema central que los autores abordan.

Como puede verse, la colección de trabajos presentados en este número brinda a nuestros lectores propuestas interesantes y ofrece una visión general de algunas áreas de la ingeniería y sus aplicaciones en la solución de problemas de actualidad.

Los rellenos sanitarios como fuente de contaminación del agua subterránea por infiltración de lixiviados: efectos en la salud y el medio ambiente

Landfills as a groundwater pollution source by infiltration of leachate: effects on health and environment

**Miriam A. García-Colindres¹, Luis Antonio Castillo-Suárez¹, Carolina Álvarez Bastida¹,
Ivonne Linares-Hernández¹, Verónica Martínez-Miranda¹**

Recibido: 7 de octubre de 2021 – Aceptado: 14 de marzo de 2022

RESUMEN

El impacto ambiental, originado por el crecimiento y concentración de la población, afecta de manera directa la cantidad de desechos sólidos urbanos que se generan, convirtiéndose en uno de los desafíos que enfrenta el mundo. Un relleno sanitario es una obra de infraestructura que involucra métodos de ingeniería para la disposición final de los RSU y de manejo especial, con el

fin de controlar, a través de la compactación e infraestructura adicionales, los impactos ambientales. Sin embargo, se ha observado que muchos de ellos funcionan como sitios controlados, causando una grave contaminación al agua subterránea por la fuga de lixiviados, producto de complejos procesos biológicos y fisicoquímicos que ocurren durante su almacenamiento y degradación. Este estudio analizó artículos científicos de diferentes países. En los

¹ Universidad Autónoma del Estado de México
Autores de correspondencia:
mmirandav@uaemex.mx, ilinaresh@uaemex.mx



diversos estudios se han detectado numerosos contaminantes del agua subterránea, así como indicadores de materia orgánica. Generalmente, se ha citado que la contaminación del agua subterránea aparece en la etapa inicial del relleno sanitario después de cinco años y alcanza su punto máximo de contaminación. Las concentraciones de los contaminantes provenientes de rellenos sanitarios disminuyeron exponencialmente a medida que aumentaba la distancia a las fuentes de agua. Por lo tanto, la contaminación del agua subterránea aparece principalmente a distancias cortas de un relleno sanitario. En este documento se describen las concentraciones máximas de los contaminantes que se han reportado, además de su impacto a la salud y ambiente.

PALABRAS CLAVE: lixiviados de rellenos sanitarios; metales; agua subterránea

ABSTRACT

The environmental impact caused by the growth and concentration of the population directly affects the amount of urban solid waste that is generated, becoming one of the challenges on the world. A sanitary landfill is

an infrastructure that involves engineering methods for the final disposal of MSW and special handling, to control, through compaction and additional infrastructure, environmental impacts. However, it has been observed that many of them function as controlled sites, causing serious contamination to groundwater due to leachate, product of complex biological and physicochemical processes that occur during storage and degradation. This study analyzed scientific articles from different countries. Numerous groundwater pollutants, as well as indicators of organic matter, have been detected in the various studies. Generally, it has been cited that groundwater contamination appears in the initial stage of the sanitary landfill after five years and reaches its maximum point of contamination. Groundwater contamination appears mainly at short distances from a sanitary landfill. This document describes the maximum concentrations of pollutants that have been reported, as well as their impact on health and the environment.

KEYWORDS: landfill leachate; metals; underground water

INTRODUCCIÓN

El agua subterránea es la principal fuente de agua dulce para la población en el mundo. Es utilizada para usos domésticos principalmente (Abiriga et al., 2021). Aproximadamente un tercio de la población en el mundo depende del agua subterránea para beber. En la mayoría de los países en desarrollo, los residuos sólidos urbanos (RSU) son almacenados de manera empírica en áreas abiertas, en sitios no controlados revestidos o no revestidos, que se asocian a fugas directas de lixiviados, lo que representan una fuente potencial de efectos adversos en las áreas circundantes que resultan en un riesgo grave para el medio ambiente (Olorunfemi et al., 2021).

El creciente aumento de la población ha generado incremento en la producción de RSU, que son depositados en espacios para su disposición, almacenamiento y reducción. Estos pueden ser ingeniería de rellenos sanitarios o tiraderos a cielo abierto, sin embargo estas prácticas pueden permitir que el lixiviado fluya directamente hacia los cuerpos de agua y el suelo más cercanos, lo que presenta riesgos para la salud humana y el medio ambiente (Wijekoon et al., 2022). Diversas investigaciones los señalan como una fuente más importante de contaminación del suelo, agua subterránea y agua superficial a través de sus lixiviados (Abiriga et al., 2020; Adamcová et al., 2017; Jayawardhana et al., 2016), así como causantes de efectos adversos para la salud humana y el medio ambiente (Abiriga et al., 2020; Aromolaran et al., 2019; Jabłońska-Trypuć et al., 2021). El lixiviado ha sido reportado como la fuente más importante de contaminación de rellenos sanitarios; puede migrar vertical y horizontalmente por advección, dispersión y dilución (Chakraborty & Kumar, 2016; Xie et al., 2016). Los lixiviados son un líquido altamente contaminado por ser una mezcla de elevadas cantidades de iones inorgánicos, compuestos orgánicos y otros elementos tóxicos como metales pesados, arsénico y nutrientes (Jayawardhana et al., 2016; Mishra et al., 2019; Przydatek, 2021). La generación de lixiviados puede variar según el tipo de desecho, el clima, el sitio y las condiciones de la superficie, lo que dificulta su tratamiento a un nivel

satisfactorio. En particular, los rellenos sanitarios municipales antiguos son asociados a una mayor toxicidad y baja biodegradabilidad (Castillo-Suárez et al., 2018).

Los factores hidrogeológicos como la lluvia, la meteorización mineral, la topografía del área de vertido y los procesos biológicos juegan un papel importante en la lixiviación de RSU por actividades antropogénicas (Mishra et al., 2019). La dispersión y difusión de los contaminantes de lixiviado en el agua subterránea dependen de las propiedades físicas y químicas de los RSU, el patrón de flujo, el nivel de infiltración de agua subterránea, el grado de hidratación residual, la tecnología utilizada en los rellenos sanitarios y el porcentaje de degradación de los residuos (Jabłońska-Trypuć et al., 2021).

La degradación de los RSU durante la temporada de lluvia, favorece la infiltración de los componentes del lixiviado con la precipitación en el agua subterránea (Aromolaran et al., 2019). Se señala que a un radio de 1000 m de la ubicación del relleno sanitario se ha encontrado una contaminación alta en el agua subterránea. También se ha observado que la concentración más alta es en un radio de 200 m (Aromolaran et al., 2019; Mishra et al., 2019). Como expresa Aromolaran y colaboradores (2019), aun cuando los rellenos sanitarios se encuentren cerrados, continuarán produciendo lixiviados y este proceso puede durar de 30 a 50 años, con un probable impacto ambiental significativo.

Estudios de investigación han demostrado que el lixiviado del rellenos sanitarios consta de compuestos químicos, en los cuales se contempla la materia orgánica disuelta (carbono orgánico total-COT; ácidos grasos volátiles-AGV; fracciones húmicas y fúlvicas; componentes inorgánicos Ca^{2+} , Mg^{2+} , NH_4^+ , Fe, Mn, Cl^- , As; metales pesados Hg, Cd, Cr, Cu, Pb) y los compuestos orgánicos xenobióticos (XOCS-benceno, tolueno, fenoles, alifáticos clorados, ftalatos) (Wijekoon et al., 2022). En los lixiviados se relaciona principalmente los XOC y los metales pesados con efectos perjudiciales en los ecosistemas y las cadenas alimentarias que provocan efectos cancerígenos, toxicidad aguda y genotoxicidad en los seres humanos, los cuales amenazan la salud humana, los servicios ecológicos y el desarrollo económico (Li et al., 2021; Wijekoon et al., 2022).

Este estudio tiene como objetivo el análisis y la revisión de la presencia de contaminantes, así como los cambios en la fisicoquímica de la calidad del agua derivado de la incorporación de contaminantes asociados a la infiltración del lixiviado de los rellenos sanitarios, así como su efecto a la salud y al ambiente.

METODOLOGÍA

La información de esta revisión se obtuvo de la base de datos *Elsevier Science Direct y Springer*. Se seleccionaron artículos de estudio donde se reportó la ubicación de un relleno sanitario o sitio de disposición de RSU a distancias circundantes de una fuente de abastecimiento y se observa alteración del agua subterránea asociada a la infiltración de lixiviados: (1) Se identificaron los principales contaminantes que se reportan en las aguas subterráneas cerca de los rellenos sanitarios. (2) Se seleccionaron los estudios donde las concentraciones de los elementos y iones superaron los límites permisibles de calidad para los estándares internacionales (WHO, 2011) poniendo en riesgo la salud y el medio ambiente. (3) Se discutieron los factores que influyen en la contaminación del agua subterránea derivado de la infiltración de los lixiviados provenientes de los RSU.

RESULTADOS

La principal preocupación ambiental en esta revisión es evidenciar las consecuencias negativas que se asocian con los lixiviados, los cuales modifican la calidad del agua subterránea. El análisis se realizó con base en los resultados reportados (Abiriga et al., 2020; Ahamad et al., 2019; Mishra et al., 2019; Talalaj, 2014; Tenodi et al., 2020) (Tabla 1). Se revela que las concentraciones de metales contaminantes del agua subterránea se encuentran en concentraciones por arriba de los límites máximos permisibles, con lo que se pone en riesgo la salud y los ecosistemas, por su grado de toxicidad, de acuerdo con la Organización Mundial de la Salud (WHO por sus siglas en inglés) (WHO, 2011). En esta revisión se analizaron las concentraciones de TOC, iones Na^+ , K^+ , Mg^{2+} , Ca^{2+} ; metales como Fe, Cr, Cd, Pb, Zn, Hg y Ni; Nutrientes NO_3^- , SO_4^{2-} ; así como As.

PARÁMETROS FISICOQUÍMICOS DEL AGUA SUBTERRÁNEA

Los valores de pH del agua subterránea cercanos a los rellenos sanitarios variaron de 4.9 a 8.8 (Tabla 1). En la misma tabla se observan que los iones mayoritarios Na^+ y Ca^{2+} son los cationes predominantes seguidos de Mg^{2+} y K^+ . En el presente estudio, la concentración de Na^+ estuvo por arriba de los límites superiores. Según las directrices de WHO (2011) el límite máximo aceptable es de 200 mg/L. La concentración de Ca^{2+} oscila entre 6.80 y 920 mg/L, lo que supera el límite aceptable de 150 mg/L prescrito para el agua de uso y consumo humano. Mg^{2+} y la concentración de K^+ varían respectivamente en el rango de 1.2 – 333.20 mg/L y 1.1 – 477.60 mg/L, concentraciones que se encuentran por encima del límite aceptable (WHO, 2011), que constituyen un desequilibrio iónico en la calidad del agua. Con lo que respecta a las concentraciones del ión Cl^- , éste es el ion más dominante entre los aniones.

La concentración de Cl^- varió de 1.6 a 2590 mg/L. Los resultados se asocian a la infiltración de los lixiviados debido a que este ión se utiliza como indicador de contaminación antropogénica, ya que es inerte y no biodegradable (Ahamad et al., 2019). En la mayoría de lo reportado se encuentra superando el límite permitido de 300 mg/L para agua de uso y consumo (Tabla 1).

TABLA 1.
RESULTADOS CONCENTRADOS DEL ANÁLISIS FÍSICOQUÍMICO EN DIFERENTES SITIOS CERCANO A RELLENOS SANITARIOS.

País	Contaminantes (Min-Max) en agua subterránea															
	pH	COT	Na ⁺	K ⁺	Mg ²⁺	Ca ²⁺	Cl ⁻	NO ₃ ⁻	As	Fe	Cr	Cd	Pb	Zn	Hg	Ni
		mg/L							µg/L							
Noruega (Abiriga et al., 2020)	4.9-8.8	0.13-47	0.09-116	1.1-86	1.2-37	6.8-83	1.6-137	0-0.16	ND	0-99,000	0.0-583	0.0-11	0-69	0-5739	0-0.145	ND
Polonia (Talakaj, 2014)	5.7-7.5	10.92-616.60	ND	ND	ND	ND	ND	ND	ND	ND	0-152	0-70	1-150	0-8893	0-12.0	ND
Serbia (Tenodi et al., 2020)	7.3-8.1	2.24-34.2	ND	ND	ND	ND	35.4-365.8	0.53-4.37	2.3-4.8	54-7910	3.37-11.0	ND	23.6-58.5	0.09-0.82	ND	5.6-55.9
India (Ahamad et al., 2019)	6.6-8.4	ND	30.8-344.98	4.9-477.6	7.57-333.20	36-920	52-2590	1.23-178.0	ND	23.3-4467.5	0-10.5	10.5-42.3	2.60-119.3	0.16-10364.5	ND	ND
Irán (Mishra et al., 2019)	6.4-7.3	ND	55.0-230	3.0-142	17.0-96.3	6.0-88.0	30.0-232	9.20-142.6	ND	0-5.11	ND	ND	ND	ND	ND	ND
WHO (2011)	6.5-8.5	ND	250	200	12	150	300	50	10	4000	50	3	10	4000	6	70
Afectaciones																
Metales	Las exposiciones a altas concentraciones de metales se asocian a graves intoxicaciones. El Cd destruye la estructura de la proteína y la función celular, daña el ADN y provoca teratogenicidad; el Pb en humanos tiene efectos subletales, el Zn la pérdida de agudeza visual y degeneración macular, el Hg inhibe el desarrollo mental y motor, causa alergias en la piel, presenta un riesgo de cáncer de pulmón, dermatitis, insuficiencia renal y gastrointestinal, además de exasperación del tracto respiratorio y los ojos (Abiriga et al., 2020; Alvarez et al., 2021).															
Nutrientes	NO ₃ ⁻ presenta riesgo de cáncer colorrectal, cáncer de vejiga y cáncer de mama, resultados reproductivos adversos, diabetes y afecciones de la tiroides en estudios epidemiológicos, el nitrato ingerido se reduce a nitrito en el tracto gastrointestinal y reduce el transporte de oxígeno a los tejidos, lo que lleva a la hipoxia (Fan, 2019; Vaiphei & Kurakalva, 2021)															
Metaloides	El As está clasificado como carcinógeno, prevalece en el ambiente natural, altamente tóxico y móvil. La exposición a As desarrolla hiperpigmentación, especialmente en el tronco y la queratosis en las palmas de las manos y las plantas de los pies (Li et al., 2021).															

CONCENTRACIÓN DE NITRATO

En general, el NO_3^- constituye una fuente de contaminación puntual antigua en el agua subterránea. En la presente revisión, la concentración de NO_3^- osciló entre 0 y 178 mg/L. El límite estándar de salud pública en el agua potable establecido es de 50 mg/L (WHO, 2011). Los resultados reportados por Ahamad y colaboradores (2019), superan el límite establecido para este anión. Típicamente, la mayor concentración de NO_3^- emana de fuentes antropogénicas, como las prácticas agrícolas intensivas (sobre aplicación de fertilizantes), la eliminación inadecuada de desechos de alcantarillado doméstico, desechos de animales y fugas de fosas sépticas (Li et al., 2020).

DISCUSIÓN

Con base en esta revisión, se observa la influencia negativa de los rellenos sanitarios de RSU, por ser una fuente antropogénica de contaminación del agua subterránea a través de la incorporación de los lixiviados, lo que pone en riesgo potencial a la salud humana. Los factores que contribuyen a la contaminación del agua subterránea cerca de un relleno sanitario son: la lluvia estacional, la edad de los rellenos sanitarios, la topografía, el modo de transporte de los lixiviados, el sistema de flujo del agua subterránea y la dispersión hidrodinámica por advección. El tipo de contaminante que altera la fisicoquímica del agua subterránea está relacionado con las características de los RSU que se disponen en el rellenos sanitarios o tiradero a cielo abierto para su disposición, almacenamiento y reducción, así como la tecnología con la que opera el sistema de tratamiento. Desde el punto de vista de Mishra y colaboradores (2018), la calidad del agua subterránea se deteriora gradualmente alrededor de los rellenos sanitarios, aun cuando han sido cerrados y el riesgo es todavía mayor a distancias cortas. La disminución de la concentración de iones inorgánicos se debe por un lado a la dilución y por otro a la precipitación de elementos solubles, mientras que la disminución de los orgánicos se debe principalmente a su biodegradación (Abiriga et al., 2020; Munagala et al., 2020; Talalaj, 2014).

CONCLUSIONES

La calidad del agua subterránea es alterada en su fisicoquímica, debido a la infiltración del lixiviado de los RSU. Esto plantea graves riesgos para la salud, lo que ha llevado a la prohibición de explotar el agua subterránea para uso humano. La presencia de las altas concentraciones de metales pesados, como el Cr, Cd, Pb, Zn, Hg en los lixiviados confirman su migración al agua subterránea, revelando la infiltración continua de la pluma de contaminación hacia los estratos del suelo. Se recomienda el uso de tecnología para la recolección y tratamiento de los lixiviados, con el fin de proteger el agua subterránea. De acuerdo con la revisión, en la mayoría de los documentos se recomienda el cierre definitivo del relleno sanitario para que se inicie el proceso de recuperación del suelo, a fin de mitigar una mayor percolación del lixiviado en agua subterránea. El monitoreo de la calidad del agua subterránea cercana a un relleno sanitario debe realizarse continuamente para evitar impactos a la salud y al ambiente, sobre todo si entra en la categoría de un sitio controlado.

AGRADECIMIENTOS

Los autores agradecen el apoyo otorgado por la SIEA/UAEMex, con No. de proyecto 6445/2022 CIB.

REFERENCIAS

- Abiriga, D., Jenkins, A., Alfsnes, K., Vestgarden, L. S. & Klempe, H. (2021). Characterisation of the bacterial microbiota of a landfill-contaminated confined aquifer undergoing intrinsic remediation. *Science of the Total Environment*, 785(3800), 0–11. <https://doi.org/10.1016/j.scitotenv.2021.147349>
- Abiriga, D., Vestgarden, L. S. & Klempe, H. (2020). Groundwater contamination from a municipal land fill: Effect of age , land fill closure, and season on groundwater chemistry. *Science of the*

- Total Environment*, 737(10), 140307. <https://doi.org/10.1016/j.scitotenv.2020.140307>
- Adamcová, D., Radziemska, M., Ridošková, A., Bartoň, S., Pelcová, P., Elbl, J., Kynický, J., Brtnický, M. & Vaverková, M. D. (2017). Environmental assessment of the effects of a municipal landfill on the content and distribution of heavy metals in *Tanacetum vulgare* L. *Chemosphere*, 185, 1011–1018. <https://doi.org/10.1016/j.chemosphere.2017.07.060>
- Ahamad, A., Raju, N. J., Madhav, S., Gossel, W. & Wycisk, P. (2019). Impact of non-engineered Bhalswa landfill on groundwater from Quaternary alluvium in Yamuna flood plain and potential human health risk, New Delhi, India. *Quaternary International*, 507, 352–369. <https://doi.org/10.1016/j.quaint.2018.06.011>
- Alvarez, C. C., Bravo Gómez, M. E. & Hernández Zavala, A. (2021). Hexavalent chromium: Regulation and health effects. *Journal of Trace Elements in Medicine and Biology*, 65, 126729. <https://doi.org/10.1016/j.jtemb.2021.126729>
- Aromolaran, O., Fagade, O. E., Aromolaran, O. K., Faleye, E. T. & Faerber, H. (2019). Assessment of groundwater pollution near Aba-Eku municipal solid waste dumpsite. *Environmental Monitoring and Assessment*, 191(12), 718. <https://doi.org/10.1007/s10661-019-7886-1>
- Castillo-Suárez, L. A., Bruno-Severo, F., Lugo-Lugo, V., Esparza-Soto, M., Martínez-Miranda, V. & Linares-Hernández, I. (2018). Peroxicoagulation and Solar Peroxicoagulation for Landfill Leachate Treatment Using a Cu–Fe System. *Water, Air, & Soil Pollution*, 229, 385. <https://doi.org/10.1007/s11270-018-4031-7>
- Chakraborty, S. & Kumar, R. N. (2016). Assessment of groundwater quality at a MSW landfill site using standard and AHP based water quality index: a case study from Ranchi, Jharkhand, India. *Environmental Monitoring and Assessment*, 188(6). <https://doi.org/10.1007/s10661-016-5336-x>
- Fan, A. M. (2019). Health, Exposure and Regulatory Implications of Nitrate and Nitrite in Drinking Water. In J. Nriagu (Ed.), *Michigan: Encyclopedia of Environmental Health* (2nd ed., pp. 417–435). Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.11837-8>
- Jabłońska-Trypuć, A., Wydro, U., Wołejko, E., Pietryczuk, A., Cudowski, A., Leszczyński, J., Rodziejewicz, J., Janczukowicz, W. & Butarewicz, A. (2021). Potential toxicity of leachate from the municipal landfill in view of the possibility of their migration to the environment through infiltration into groundwater. *Environmental Geochemistry and Health*, 43, 3683–3698. <https://doi.org/10.1007/s10653-021-00867-5>
- Jayawardhana, Y., Kumarathilaka, P., Herath, I. & Vithanage, M. (2016). Municipal Solid Waste Biochar for Prevention of Pollution From Landfill Leachate. In M. N. V. Prasad and K. Shih (Eds). *Environmental Materials and Waste* (pp. 117–148). Elsevier.

<https://doi.org/10.1016/B978-0-12-803837-6.00006-8>

- Li, L., Li, Z., Liu, D. & Song, K. (2020). Evaluation of partial nitrification efficiency as a response to cadmium concentration and microplastic polyvinylchloride abundance during landfill leachate treatment. *Chemosphere*, 247, 125903. <https://doi.org/10.1016/j.chemosphere.2020.125903>
- Li, P., Karunanidhi, D., Subramani, T. & Srinivasamoorthy, K. (2021). Sources and Consequences of Groundwater Contamination. *Archives of Environmental Contamination and Toxicology*, 80, 1–10. <https://doi.org/10.1007/s00244-020-00805-z>
- Mishra, S., Tiwary, D., Ohri, A. & Agnihotri, A. K. (2019). Impact of municipal solid waste landfill leachate on groundwater quality in Varanasi, India. *Groundwater for Sustainable Development*, 9, 100230. <https://doi.org/10.1016/j.gsd.2019.100230>
- Munagala, S., Chaitanya, D., Jagarapu, K. & Ratnamala Reddy B. S. S. (2020). Determination of water quality index for ground water near municipal dump site in Guntur. *Materials Today: Proceedings*, 33(1), 724-727. <https://doi.org/10.1016/j.matpr.2020.06.030>
- Olorunfemi, A. O., Alao-Daniel, A. B., Adesiyun, T. A. & Onah, C. E. (2021). Geochemical assessment of heavy metal impact on soil around Ewu-Elepe Dumpsite, Lagos State, Nigeria. *Ife Journal of Science*, 22(3), 119–138. <https://doi.org/10.4314/ij.s.v22i3.10>
- Przydatek, G. (2021). Using advanced statistical tools to assess the impact of a small landfill site on the aquatic environment. *Environmental Monitoring and Assessment*, 193(2), 71. <https://doi.org/10.1007/s10661-021-08850-4>
- Talalaj, I. A. (2014). Assessment of groundwater quality near the landfill site using the modified water quality index. *Environmental Monitoring and Assessment*, 186, 3673 -3683. <https://doi.org/10.1007/s10661-014-3649-1>
- Tenodi, S., Kr, D., Agbaba, J., Zrni, K., Radenovi, M., & Ubavin, D. (2020). Assessment of the environmental impact of sanitary and unsanitary parts of a municipal solid waste landfill. *Journal of Environmental Management*, 258, 110019. <https://doi.org/10.1016/j.jenvman.2019.110019>
- Vaiphei, S. P. & Kurakalva, R. M. (2021). Hydrochemical characteristics and nitrate health risk assessment of groundwater through seasonal variations from an intensive agricultural region of upper Krishna River basin, Telangana, India. *Ecotoxicology and Environmental Safety*, 213, 112073. <https://doi.org/10.1016/j.ecoenv.2021.112073>
- WHO. (2011). *Guidelines for drinking-water quality: fourth edition incorporating the first and second addenda*. World Health Organization.

- Wijekoon, P., Koliyabandara, P. A., Cooray, A. T., Lam, S. S., Athapattu, B. C. L. & Vithanage, M. (2022). Progress and prospects in mitigation of landfill leachate pollution: Risk, pollution potential, treatment and challenges. *Journal of Hazardous Materials*, 421, 126627. <https://doi.org/10.1016/j.jhazmat.2021.126627>
- Xie, H., Chen, Y., Thomas, H. R., Sedighi, M., Masum, S. A. & Ran, Q. (2016). Contaminant transport in the sub-surface soil of an uncontrolled landfill site in China: site investigation and two-dimensional numerical analysis. *Environmental Science and Pollution Research*, 23(3), 2566–2575. <https://doi.org/10.1007/s11356-015-5504-5>

Efectos ambientales del drenaje ácido de mina y su tratamiento

Environmental effects of acid mine drainage and treatment

Romel González-Hernández¹, Verónica Martínez-Miranda^{1,*}, Ivonne Linares-Hernández¹, Raúl Cortés-Martínez², Luis Antonio Castillo-Suárez¹

Recibido: 7 de octubre de 2021 – Aceptado: 14 de marzo de 2022

RESUMEN

La minería es una actividad industrial de gran importancia para el desarrollo científico, tecnológico, arquitectónico, de la salud y las telecomunicaciones. México es caracterizado por su riqueza de especies minerales con

diversidad geológica que promueve gran interés entre compañías mineras nacionales e internacionales (Ramírez Macedonio & García Castro, 2020).

Por otro lado, la minería genera grandes cantidades de desechos que no son tratados y provocan graves problemas ambientales

¹ Universidad Autónoma del Estado de México

² Universidad Michoacana de San Nicolás de Hidalgo

* Autor de correspondencia: mmirandav@uaemex.mx



(Mafra et al., 2020). Uno de ellos es el Drenaje Ácido de Mina (DAM) que es resultado de la oxidación química y bacteriana de minerales sulfurados como la pirita (FeS_2) durante la extracción de minerales metálicos y carbón.

El impacto adverso que causa el DAM sobre la vida acuática y la calidad de los cuerpos de agua es preocupante debido a la movilidad de ácidos, sulfatos y metales (Fe, Al, Mn, Zn, Cu, Pb, Hg, Cd, Ni), metaloides (As) y sólidos en suspensión (Akcil & Koldas, 2006; Alonso et al., 2020; Kefeni et al., 2018; Nqombolo et al., 2019; Rodríguez-Galán et al., 2019). Esta contaminación es un problema global debido a su toxicidad y bioacumulación (Atangana & Oberholster, 2021; Ayala & Fernández, 2019; Pan et al., 2021; Sahoo et al., 2020). El objetivo de esta revisión bibliográfica es investigar los impactos del DAM, además de los avances y propuestas para su manejo, tratamiento y mitigación.

PALABRAS CLAVE: Drenaje ácido de mina; elementos potencialmente tóxicos; tratamiento

ABSTRACT

Mining is an industrial activity of great importance for scientific and technological development, architecture, health, and telecommunications. Mexico is characterized

by its wealth of mineral species and geological diversity that promotes great interest among national and international mining companies (Ramírez Macedonio & García Castro, 2020).

On the other hand, mining generates large amount of waste that is not treated and causes serious environmental problems (Mafra et al., 2020). One of them is Acid Mine Drainage (AMD) which is the result of chemical and bacterial oxidation of sulphide minerals such as pyrite (FeS_2) during the extraction of metallic minerals and coal.

The adverse impact caused by AMD on aquatic life and the quality of water bodies is of concern due to the mobility of acids, sulfates and metals (Fe, Al, Mn, Zn, Cu, Pb, Hg, Cd, Ni), metalloids (As) and suspended solids (Akcil & Koldas, 2006; Alonso et al., 2020; Kefeni et al., 2018; Nqombolo et al., 2019; Rodríguez-Galán et al., 2019); this contamination is a global problem due to its toxicity and bioaccumulation (Atangana & Oberholster, 2021; Ayala & Fernández, 2019; Pan et al., 2021; Sahoo et al., 2020). The objective of this bibliographic revision is to investigate the impacts of AMD, the advances and the proposals for its management, treatment, and mitigation.

KEYWORDS: Acid mine drainage; potentially toxic elements; treatment.

INTRODUCCIÓN

La minería es una actividad económica muy importante como fuente de materias primas. En México está centrada en los metales industriales y preciosos (SGM, 2013), que se lleva a cabo principalmente con labores de aprovechamiento a cielo abierto. Esto produce impactos ambientales como la contaminación del agua mediante DAM, remoción de la cubierta vegetal, generación de grandes volúmenes de escombros, abatimiento de fuentes de agua, emisiones continuas de gases y polvos a la atmósfera durante la extracción y procesamiento (Lutz Ley, 2020), así como enfermedades humanas, afectación a las actividades económicas locales, marginación y pobreza (Moraes, 2000)

La pirita (FeS_2) es el mineral principal causante del DAM (Figura 1), no obstante, existen otros sulfuros metálicos como asphalerita (ZnS), galena (PbS), calcopirita (CuFeS_2), arsenopirita (FeAsS) entre otros, que contienen menores cantidades de Cd, Sn, Ag, Au, Co y Hg.

El DAM daña los ecosistemas y los cultivos alimentarios representando una amenaza a la flora, fauna y seres humanos (Pan et al., 2021). Algunos efectos a la salud por la ingesta de agua contaminada son deformaciones congénitas, mortalidad, intoxicaciones, lesiones diversas en los epitelios, trastornos durante el crecimiento y bajas tasas de reproducción, asociados a la acumulación de metales pesados y As que además causan daños hepáticos y renales (Gaikwad & Gupta, 2008).

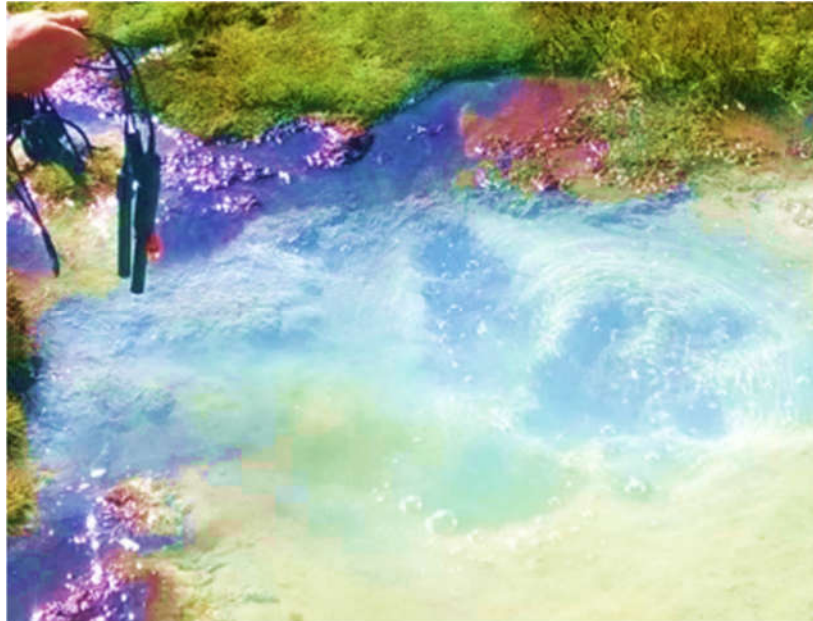


FIGURA 1.

DESCARGA DE DAM EN UN ARROYO EN MICHOACÁN

Fuente: Elaboración propia

RESIDUOS MINEROS Y SUS IMPACTOS AMBIENTALES

Durante todo el proceso de aprovechamiento de los minerales se producen desechos como relaves, lodos y aguas residuales. Durante la extracción de las rocas que contienen los minerales, el tratamiento para extraerlos y la metalurgia, se demanda un gran volumen de agua que es utilizada para reducir la generación de polvo, lavar el carbón, triturar rocas, separar floculantes, además de la extracción hidrometalúrgica. En pocas ocasiones el agua se reúsa, ya que, al contaminarse, carece de valor en la operación. Su descarga, flujo, drenaje y filtración incontrolados desde el sitio minero están asociados a altas temperaturas, sólidos suspendidos, bases, ácidos, sólidos disueltos; esto incluyendo químicos del procesamiento, metales, metaloides, sustancias radiactivas y/o sales (COCHILCO, 2020).

El Servicio Forestal de EU estimó que entre 8,000 y 16,000 km de los ríos de EUA son objeto de DAM. La Agencia de Protección Ambiental (EPA) estableció que la minería ha contaminado el 40 % de las cuencas hidrográficas de los estados del Este. En el 2011 y 2012 se reportó que la industria de las minas de metales fue la más contaminante, con el 46 % de los desechos totales. Si se suman los desechos arrojados por la industria que funde y refina, el porcentaje se eleva hasta el 55%. Sólo el 0.7 % (859,194.6 toneladas) de los desechos mineros se trataron antes de su deposición final (EPA, 2012).

En México, los periodistas reportan múltiples casos de accidentes por derrames, fractura de diques de presas de jales, volcaduras de camiones que mueven contenedores con sustancias altamente peligrosas y la contaminación de ríos, lagos y mantos freáticos por efecto de minas abandonadas. El derrame de jales de Buenavista del Cobre al Río Sonora en 2014 identificado como el de mayor impacto negativo (Cisneros, 2014), es un buen ejemplo de esto. La Procuraduría Federal de Protección al Ambiente (PROFEPA) informó que durante 2013-2015 se atendieron 44 emergencias mineras.

DISTRIBUCIÓN AMBIENTAL DE METALES Y METALOIDES DE LOS DESECHOS MINEROS

Las principales vías de transporte o vectores de los contaminantes en el ambiente son: suelo, aire, agua y biota.

Suelo

Aunque la mayoría de los procesos de transporte en el suelo son inherentemente lentos y limitados en extensión espacial (Gutiérrez-Ruiz y Moreno, 1995), en periodos largos, los contaminantes pueden ser transportados por el suelo en grandes cantidades, dando lugar a elevados riesgos para la salud humana y el ambiente local (Csavina et al., 2012).

En 1997, México identificó 59 sitios abandonados y contaminados con residuos peligrosos; para 1999 ya eran 105 sitios creciendo significativamente en tan sólo 2 años. De acuerdo con datos de PROFEPA, en el 2015 se sumaron 61 sitios, siendo más frecuentes los metales pesados en los que predominó principalmente el Pb. Los metales pesados se agrupan

en una categoría que se les denomina elementos potencialmente tóxicos (EPT) y se encuentran en residuos de jales y escorias de fundición, cuyo volumen supera las 80,000 toneladas. Los EPT más comunes son Pb, Cd, Zn, As, Se y Hg (Gutiérrez-Ruiz y Moreno, 1995; Volke et al., 2005).

Aire

La minería produce emisiones atmosféricas compuestas por polvos gruesos generados a partir de la trituración y erosión de jales, y emisión de polvos finos por fundición de vertederos de escorias. Estos polvos atmosféricos (texturas $<60\mu\text{m}$) incluyen aerosoles y en conjunto juegan un papel importante en el transporte de contaminantes de baja volatilidad, escasa solubilidad acuosa y fácil adherencia al suelo (Csavina et al., 2012).

Por otra parte, los contaminantes comúnmente asociados a polvos emitidos son As, Cd y Pb, están concentrados en la fracción de textura más fina ($<2\mu\text{m}$ y $<1\mu\text{m}$) que viajan a mayor distancia y presentan más riesgo para la salud humana que polvos gruesos (Corriveau et al., 2011; Csavina et al., 2011; Ravi et al., 2011). Los polvos finos son inhalados y depositados directamente en los pulmones a través de la boca y la nariz y pueden ser transportados a través de la sangre absorbiéndose por el sistema digestivo.

Agua

El transporte de contaminantes de las minas a través del agua ocurre en varios procesos: desde el transporte en sedimentos hasta la percolación de elementos y compuestos solubles, trasladándose desde los límites de las cuencas hidrográficas hasta las zonas más bajas, siendo de mayor riesgo por acumulación las zonas ribereñas, sistemas acuáticos endorreicos y aguas subterráneas poco profundas (Csavina et al., 2012).

Un ejemplo es la mina de cobre localizada al oeste de Montana, la cual, después de 125 años de funcionamiento, es considerada uno de los sitios altamente contaminados (*Superfund*), ya que 1,600 km de ríos de la cuenca del *Clark Fork* se contaminaron por desechos que se transportaron y redistribuyeron a 400 km de distancia del origen, según lo reportado por Moore & Luoma (1990).

ALTERNATIVAS PARA EL TRATAMIENTO DE DAM

Existen varios intentos de tratamientos para eliminar contaminantes de DAM: procesos biológicos (Magowo et al., 2020; Sahoo et al., 2020), adsorción (Ryu et al., 2020), flotación por aire (Bai et al., 2021), intercambio iónico (Park et al., 2019), ósmosis inversa (Ambiado et al., 2017), nanofiltración (Siew et al., 2020), precipitación (Menzel et al., 2021), electrocoagulación (Foudhaili et al., 2020), humedales (Zhao et al., 2020), bacterias anaeróbicas y algas (Carneiro Brandão Pereira et al., 2020). Estos tratamientos, en conjunto, retienen metales pesados y oxidan compuestos de azufre móviles (Bwapwa et al., 2017; Kaur et al., 2018; Martínez-Macias et al., 2019; Méndez-García et al., 2015; Pat-Espadas et al., 2018).

Las propuestas convencionales se centran en la neutralización del DAM mediante la adición de compuestos alcalinos como rocas calizas, cal, hidróxido de sodio, carbonato de sodio, alúmina, carbonatos de magnesio. No obstante, estos tratamientos conllevan a inversiones considerables, así como costos adicionales para reactivos y el problema que representa la eliminación de los lodos residuales (Sandlin et al., 2021).

Los tratamientos electroquímicos presentan avances prometedores; se clasifican en dos tipos: electrolíticos que requieren de una fuente externa de energía eléctrica y promueven reacciones químicas no espontáneas, y los galvánicos o voltaicos que no requieren inducción de corriente, al producir energía eléctrica a partir de las reacciones químicas espontáneas.

Actualmente los tratamientos de DAM mediante estos procesos han sido de gran utilidad debido a la posibilidad de recuperar metales disueltos. La desventaja de estos son los altos costos en el consumo de electricidad que se aplica a las celdas electrolíticas. Así, los tratamientos galvánicos podrían ser una gran opción para el tratamiento de DAM porque no requieren energía, ya que funcionan a través de procesos espontáneos. Por lo tanto, un sistema galvánico es eficiente, económico, ecológico, autónomo y portátil para operar en ubicaciones remotas (Alcalá-Delgado et al., 2018; Castillo-Suárez et al., 2019)

CONCLUSIONES

Las actividades mineras son un asunto que requiere de atención prioritaria. A través de los años ha sido cada vez más evidente la fragmentación de los ecosistemas que deterioran y reducen la flora, fauna y recursos físicos, entre ellos el suelo y el estrés hídrico, al cual son sometidas las regiones aledañas a las industrias que hacen uso de los recursos minerales. Aunado a esto, hay evidencia del deterioro de la calidad de vida de las poblaciones que se intoxican cotidianamente con el agua que beben y con los alimentos que cultivan, consumen y comercializan fuera de la región, llevando el problema a otros lugares.

[La realidad es que aun cuando las actividades mineras representan un fuerte problema ambiental, no es posible prescindir de ellas; la minería forma una parte importante del PIB y es fuente de materias primas en la industria de las telecomunicaciones, la construcción, la industria farmacéutica y los transportes, además de que es una fuente importante de empleo.

Los avances que se han generado a través de los años para tratar los residuos de la minería han contribuido de manera significativa al mejoramiento del ambiente. No se tiene la solución final, pero las investigaciones continúan y los aportes seguirán llegando a través de los años. Es de suma importancia reforzar los conocimientos científicos y tecnológicos en esta problemática como lo es el reforzamiento de las leyes en materia de protección ambiental en México y el mundo.

AGRADECIMIENTOS

Los autores agradecen el apoyo otorgado al proyecto financiado por la SIEA-UAEMEX (6449/2022 CIB).

REFERENCIAS

- Akcil, A. & Koldas, S. (2006). Acid Mine Drainage (AMD): causes, treatment and case studies. *Journal of Cleaner Production*, 14 (12-13), 1139 -114.
<https://doi.org/10.1016/j.jclepro.2004.09.006>
- Alcalá-Delgado, A. G., Lugo-Lugo, V., Linares-Hernández, I., Martínez-Miranda, V., Fuentes-Rivas, R. M. & Ureña-Nuñez, F. (2018). Industrial wastewater treated by galvanic, galvanic Fenton, and hydrogen peroxide systems. *Journal of Water Process Engineering*, 22(1–12), 1-12. <https://doi.org/10.1016/j.jwpe.2018.01.001>
- Alonso, D. L., Pérez, R., Okio, C. K. Y. A. & Castillo, E. (2020). Assessment of mining activity on arsenic contamination in surface water and sediments in southwestern area of Santurbán paramo, Colombia. *Journal of Environmental Management*, 264, 110478. <https://doi.org/10.1016/j.jenvman.2020.110478>
- Ambiado, K., Bustos, C., Schwarz, A. & Bórquez, R. (2017). Membrane technology applied to acid mine drainage from copper mining. *Water Science and Technology*, 75(3), 705-715. <https://doi.org/10.2166/wst.2016.556>
- Atangana, E. & Oberholster, P. J. (2021). Using heavy metal pollution indices to assess water quality of surface and groundwater on catchment levels in South Africa. *Journal of African Earth Sciences*, 182, 104254.
<https://doi.org/10.1016/j.jafrearsci.2021.104254>
- Ayala, J. & Fernández, B. (2019). Treatment from abandoned mine landfill leachates. Adsorption technology. *Journal of Materials Research and Technology*, 8(3), 2732–2740. <https://doi.org/10.1016/j.jmrt.2019.04.009>
- Bai, S., Bi, Y., Li, J., Yu, P., Ding, Z., Lv, C. & Wen, S. (2021). Innovative utilization of acid mine drainage (AMD): A promising activator for pyrite flotation once depressed in a high alkali solution (HAS)—Gearing towards a cleaner production concept of copper sulfide ore. *Minerals Engineering*, 170, 106997.
<https://doi.org/10.1016/j.mineng.2021.106997>
- Bwapwa, J. K., Jaiyeola, A. T. & Chetty, R. (2017). Bioremediation of acid mine drainage using algae strains: A review. *South African Journal of Chemical Engineering*, 24, 62-70. <https://doi.org/10.1016/j.sajce.2017.06.005>

- Carrillo Cisneros, E., Espinosa Guillén, S., Olguín Negrete, B. R. & Cuevas Contreras, T. (2015). Imagen del destino turístico frente a la contingencia ambiental del 2014 en el Estado de Sonora. *La Revista de la Universidad Estatal de Sonora*, 4, 9-27.
- Carneiro Brandão Pereira, T., Batista dos Santos, K., Lautert-Dutra, W., de Souza Teodoro, L., de Almeida, V. O., Weiler, J., Homrich Schneider, I. A. & Reis Bogo, M. (2020). Acid mine drainage (AMD) treatment by neutralization: Evaluation of physical-chemical performance and ecotoxicological effects on zebrafish (*Danio rerio*) development. *Chemosphere*, 253, 1126665. <https://doi.org/10.1016/j.chemosphere.2020.126665>
- Castillo-Suárez, L. A., Lugo-Lugo, V., Linares-Hernández, I., Martínez-Miranda, V., Esparza-Soto, M. & Mier-Quiroga, M. D. L. Á. (2019). Biodegradability index enhancement of landfill leachates using a Solar Galvanic-Fenton and Galvanic-Fenton system coupled to an anaerobic-aerobic bioreactor. *Solar Energy*, 188, 989-1001. <https://doi.org/10.1016/j.solener.2019.07.010>
- COCHILCO. (2020). *Proyección de consumo de agua en la minería del cobre 2020-2031*. Dirección de Estudios y Políticas Públicas, DEPP 27/20.
- Corriveau, M. C. & Jamieson, H. E. Parsons, M. B. & Hall, G. E. M. (2011). Mineralogical characterization of arsenic in gold mine tailings from three sites in Nova Scotia. *Geochemistry Exploration Environment Analysis* 11(3), 179-192. <https://doi.org/10.1144/1467-7873/09-246>
- Csavina, J., Field, J., Taylor, M. P., Gao, S., Landázuri, A., Betterton, E. A. & Sáez, A. E. (2012). A review on the importance of metals and metalloids in atmospheric dust and aerosol from mining operations. *Science of the Total Environment*, 433, 58-73. <https://doi.org/10.1016/j.scitotenv.2012.06.013>
- Csavina, J., Landázuri, A., Wonaschütz, A., Rine, K., Rheinheimer, P., Barbaris, B., Conant, W., Sáez, A. E. & Betterton, E. A. (2011). Metal and metalloid contaminants in atmospheric aerosols from mining operations. *Water, Air, and Soil Pollution*, 221(1-4). <https://doi.org/10.1007/s11270-011-0777-x>
- EPA (2011). Method 1627: Kinetic test method for the prediction of mine drainage quality. USA environmental protection agency: Washington D.C., EPA-821-R-09-002
- Foudhaili, T., Lefebvre, O., Coudert, L. & Neculita, C. M. (2020). Sulfate removal from mine drainage by electrocoagulation as a stand-alone treatment or polishing step. *Minerals Engineering*, 152, 106337. <https://doi.org/10.1016/j.mineng.2020.106337>

- Gaikwad, R. W. & Gupta, D. V. (2008). Review on removal of heavy metals from acid mine drainage. *Applied Ecology and Environmental Research*, 6(3), 81-98. https://doi.org/10.15666/aeer/0603_081098
- Gutiérrez-Ruiz M. & Moreno Turrent M. (1995). Los residuos en la minería mexicana, en F. J. Garfias Ayala y L. Barojas Weber (Eds.), *Residuos peligrosos en México* (pp. 37-44). Semarnap/INE, México D.F.
- Kaur, G., Couperthwaite, S. J., Hatton-Jones, B. W. & Millar, G. J. (2018). Alternative neutralisation materials for acid mine drainage treatment. *Journal of Water Process Engineering*, 22, 46-58. <https://doi.org/10.1016/j.jwpe.2018.01.004>
- Kefeni, K. K., Msagati, T. A. M., Nkambule, T. T. I. & Mamba, B. B. (2018). Synthesis and application of hematite nanoparticles for acid mine drainage treatment. *Journal of Environmental Chemical Engineering*, 6(2), 1865-1874. <https://doi.org/10.1016/j.jece.2018.02.037>
- Lutz Ley, A. N. (2020). Minería y seguridad hídrica en el noroeste de México: un análisis de doble exposición. *Región y Sociedad*, 32, e1295. <https://doi.org/10.22198/rys2020/32/1295>
- Mafrá, C., Bouzahzah, H., Stamenov, L. & Gaydardzhiev, S. (2020). Insights on the effect of pyrite liberation degree upon the acid mine drainage potential of sulfide flotation tailings. *Applied Geochemistry*, 123, 104774. <https://doi.org/10.1016/j.apgeochem.2020.104774>
- Magowo, W. E., Sheridan, C. & Rumbold, K. (2020). Global Co-occurrence of Acid Mine Drainage and Organic Rich Industrial and Domestic Effluent: Biological sulfate reduction as a co-treatment-option. *Journal of Water Process Engineering*, 38, 101650. <https://doi.org/10.1016/j.jwpe.2020.101650>
- Martínez-Macias, M. del R., Correa-Murrieta, M. A., Villegas-Peralta, Y., Dévora-Isiordia, G. E., Álvarez-Sánchez, J., Saldivar-Cabral, J. & Sánchez-Duarte, R. G. (2019). Uptake of copper from acid mine drainage by the microalgae *Nannochloropsis oculata*. *Environmental Science and Pollution Research*, 26(7), 6311-6318. <https://doi.org/10.1007/s11356-018-3963-1>
- Méndez-García, C., Peláez, A. I., Mesa, V., Sánchez, J., Golyshina, O. V. & Ferrer, M. (2015). Microbial diversity and metabolic networks in acid mine drainage habitats. In *Frontiers in Microbiology*, 6, 475. <https://doi.org/10.3389/fmicb.2015.00475>

- Menzel, K., Barros, L., García, A., Ruby-Figueroa, R. & Estay, H. (2021). Metal sulfide precipitation coupled with membrane filtration process for recovering copper from acid mine drainage. *Separation and Purification Technology*, 270, 118721. <https://doi.org/10.1016/j.seppur.2021.118721>
- Moore, J. N. & Luoma, S. N. (1990). Hazardous wastes from large-scale metal extraction: A case study. *Environmental Science and Technology*, 24(9), 1278-1285. <https://doi.org/10.1021/es00079a001>
- Moran, R. E. (2000). Is this number to your liking? Water quality predictions in mining impact studies en D. Saraweit, R. A. Pielke, Jr., and R. Byerly, Jr. (Eds), *Prediction: Science, Decision Making and the Future of Nature* (185-198). Island Press
- Nqombolo, A., Mpupa, A., Gugushe, A. S., Moutloali, R. M. & Nomngongo, P. N. (2019). Adsorptive removal of lead from acid mine drainage using cobalt-methylimidazolate framework as an adsorbent: kinetics, isotherm, and regeneration. *Environmental Science and Pollution Research*, 26(4), 3330-3339. <https://doi.org/10.1007/s11356-018-3868-z>
- Pan, Y., Ye, H., Li, X., Yi, X., Wen, Z., Wang, H., Lu, G. & Dang, Z. (2021). Spatial distribution characteristics of the microbial community and multi-phase distribution of toxic metals in the geochemical gradients caused by acid mine drainage, South China. *Science of the Total Environment*, 774, 145660. <https://doi.org/10.1016/j.scitotenv.2021.145660>
- Park, I., Tabelin, C. B., Jeon, S., Li, X., Seno, K., Ito, M. & Hiroyoshi, N. (2019). A review of recent strategies for acid mine drainage prevention and mine tailings recycling. *Chemosphere*, 219, 588-606. <https://doi.org/10.1016/j.chemosphere.2018.11.053>
- Pat-Espadas, A. M., Portales, R. L., Amabilis-Sosa, L. E., Gómez, G. & Vidal, G. (2018). Review of constructed wetlands for acid mine drainage treatment. *Water*, 10(11), 1685. <https://doi.org/10.3390/w10111685>
- Procuraduría Federal de Protección al Ambiente. (20 de septiembre de 2015) Atiende PROFEPA 44 emergencias ambientales en los últimos dos años. Gobierno de México. https://www.profepa.gob.mx/innovaportal/v/7576/1/mx/atiende_profepa_44_emergencias_ambientales_en_los_ultimos_dos_anos.html.
- Ramírez Macedonio, J. L. & García Castro, N. (2020). Impacto socioambiental de la minería a cielo abierto en Mezcala, Guerrero, México. *RICSH Revista Iberoamericana de las Ciencias Sociales y Humanísticas*, 9(17), 219-239. <https://doi.org/10.23913/ricsh.v9i17.195>

- Ravi, S., D'Odorico, P., Over T. M. & Zobeck T. M. (2004). On the effect of air humidity on soil susceptibility to wind erosion: The case of air-dry soils. *Geophysical Research Letters*, 31, L09501.
- Rodríguez-Galán, M., Baena-Moreno, F. M., Vázquez, S., Arroyo-Torralvo, F., Vilches, L. F. & Zhang, Z. (2019). Remediation of acid mine drainage. *Environmental Chemistry Letters*, 17(4), 1529-1538. <https://doi.org/10.1007/s10311-019-00894-w>
- Ryu, S. C., Naidu, G., Moon, H. & Vigneswaran, S. (2020). Selective copper recovery by membrane distillation and adsorption system from synthetic acid mine drainage. *Chemosphere*, 260, 127528. <https://doi.org/10.1016/j.chemosphere.2020.127528>
- Sahoo, H., Senapati, D., Thakur, I. S. & Naik, U. C. (2020). Integrated bacteria-algal bioreactor for removal of toxic metals in acid mine drainage from iron ore mines. *Bioresource Technology Reports*, 11, 100422. <https://doi.org/10.1016/j.biteb.2020.100422>
- Sandlin, W., Langman, J. & Moberly, J. (2021). A review of acid rock drainage, seasonal flux of discharge and metal concentrations, and passive treatment system limitations. *International Journal of Mining, Reclamation and Environment*, 35(1), 34-47. <https://doi.org/10.1080/17480930.2020.1728035>
- SGM (Servicio Geológico Mexicano). (20 de septiembre de 2013). Panorama minero en México: Servicio Geológico Mexicano. Gobierno de México. Consulta los panoramas mineros estatales | Servicio Geológico Mexicano | Gobierno | gob.mx (www.gob.mx)
- Siew, Y. W., Zedda, K. L. & Velizarov, S. (2020). Nanofiltration of simulated acid mine drainage: Effect of pH and membrane charge. *Applied Sciences*, 10(1), 400. <https://doi.org/10.3390/app10010400>
- Volke, T., Velasco, A. & de la Rosa, A. (2005). Suelos contaminados por metales y metaloides: muestreo y alternativas para su remediación. Secretaría del Medio Ambiente y Recursos Naturales. Instituto Nacional de Ecología. SuelosContaminadosPorMetalesYMetaloides.pdf (congreso.gob.pe)
- Zhao, Q., Huang, J. C., He, S. & Zhou, W. (2020). Enhancement of a constructed wetland water treatment system for selenium removal. *Science of the Total Environment*, 714, 136741. <https://doi.org/10.1016/j.scitotenv.2020.136741>

Comparación entre MaxEnt y SVM como clasificadores multiclase. Caso de estudio: Río Amacuzac

Comparison between MaxEnt and SVM
as multiclass classifiers.
Case Study: Amacuzac River

**Yelena Méndez Torres^{1,*}, Carlos Alberto Mastachi Loza^{1,*},
Ricardo Arévalo Mejía¹, María de Lourdes Ruíz Gómez¹, Miguel Ángel Gómez Albores¹,
Humberto Mejía Mojica²**

Recibido: 7 de octubre de 2021 – Aceptado: 14 de marzo de 2022

RESUMEN

En los últimos años, gracias a satélites como el Sentinel 2 y la información que provee, la teledetección ha sido ampliamente utilizada en el estudio de los recursos naturales. Aunado a esta tendencia, se han desarrollado y

actualizado los algoritmos que se utilizan para procesar la información disponible. Una de las principales tareas a realizar con imágenes satelitales es la clasificación de características multiclase. Para llevar a cabo esta tarea, el algoritmo más utilizado hasta el momento ha sido la Máquina de Soporte de Vectores (SVM), sin embargo, trabajos recientes han empleado el

¹ Universidad Autónoma del Estado de México

² Universidad Autónoma del Estado de Morelos

* Autores de correspondencia:
yelenamendeztorres@gmail.com
camastachil@uaemex.mx



algoritmo de Máxima Entropía (MaxEnt) con resultados prometedores.

El objetivo de la presente investigación consiste en comparar el rendimiento de ambos algoritmos en la tarea de clasificar características como: río, vegetación y arena en imágenes satelitales del área del Río Amacuzac. Para poder comparar ambos algoritmos, se propone una metodología que consiste en combinar los resultados específicos del MaxEnt, para cada clase, en un mapa general. Se utilizaron los valores por defecto de los parámetros de ambos algoritmos y se compararon los resultados para los valores de Precisión General (*Overall Accuracy*), Precisión del Productor (*Producer's Accuracy*) y el Área bajo la Curva (AUC). Se llegó a la conclusión de que el algoritmo SVM logró clasificar con mayor precisión las características de interés anteriormente mencionadas para nuestro conjunto de datos en particular.

PALABRAS CLAVE: Teledetección; MaxEnt; SVM

ABSTRACT

The environmental impact caused by the growth and concentration of the population directly affects the amount of urban solid waste that is generated,

becoming one of the challenges on the world. A sanitary landfill is an infrastructure that involves engineering methods for the final disposal of MSW and special handling, to control, through compaction and additional infrastructure, environmental impacts. However, it has been observed that many of them function as controlled sites, causing serious contamination to groundwater due to leachate, product of complex biological and physicochemical processes that occur during storage and degradation. This study analyzed scientific articles from different countries. Numerous groundwater pollutants, as well as indicators of organic matter, have been detected in the various studies. Generally, it has been cited that groundwater contamination appears in the initial stage of the sanitary landfill after five years and reaches its maximum point of contamination. Groundwater contamination appears mainly at short distances from a sanitary landfill. This document describes the maximum concentrations of pollutants that have been reported, as well as their impact on health and the environment.

KEYWORDS: remote sensing; MaxEnt; SVM

INTRODUCCIÓN

La teledetección es cada vez más utilizada como soporte para los estudios ambientales, ya que ayuda a identificar y monitorear las superficies terrestres y las condiciones del ambiente. Una de las tareas dentro de la teledetección es la clasificación multiclase de las imágenes satelitales. La clasificación multiclase (MCC) consiste en la asignación de varias etiquetas (dentro de un conjunto predefinido) a ciertas regiones de una imagen en particular, lo que la hacen una tarea más compleja que la clasificación de una sola clase (OCC) Krawczyk et al., 2018).

Para facilitar la clasificación multiclase son utilizados algoritmos de clasificación automática que reducen el tiempo y la complejidad de la tarea. Uno de los algoritmos más comúnmente utilizado en la literatura es Máquina de Soporte Vectorial (SVM) (Cortes & Vapnik, 2006), sin embargo, algunos trabajos utilizan otros como el algoritmo *Random Forest*, KNN (*K-Nearest-Neighbor*) y *MaxEnt* (*Maximum Entropy*).

El MaxEnt⁴ (Phillips et al., 2006) es particularmente popular en el modelado de distribución de especies y de nichos ambientales. Este también ha sido aplicado en diferentes estudios que incluyen mapear cobertura de tierras urbanas, formaciones de matorrales y arbustos, así como estudios regionales y continentales (Mack & Waske, 2017). Su fácil utilización, debido a los parámetros predeterminados con los que trabaja, lo hacen un algoritmo muy accesible para todos. En algunos estudios se muestra cómo el *MaxEnt* obtiene mejores resultados que el SVM en tareas de clasificación de una clase. Sin embargo, no es común verlo utilizado como un algoritmo de clasificación multiclase. Varios estudios realizan comparaciones entre *MaxEnt* y SVM para encontrar el más idóneo a la hora de clasificar (tanto para una clase como para multiclase); incluso comparan estos con otros algoritmos (Mack & Waske, 2017; Pratiwi & Widodo 2017; Chen et al., 2017; Rahimian et al., 2019; Farrell et al., 2019), sin embargo, no se ha determinado la supremacía de un algoritmo sobre otro.

Inspirados en los buenos resultados de *MaxEnt* para problemas de una sola clase y en el poco uso hasta el momento como clasificador multiclase, en esta investigación lo comparamos con el algoritmo SVM, para contribuir a la estimación de su rendimiento en este tipo de tareas. En este trabajo en específico, se compara el rendimiento de ambos algoritmos

⁴ Este algoritmo se encuentra disponible de forma gratuita en <http://www.cs.princeton.edu/schapiro/maxent/>

en la tarea de clasificar características como: río, vegetación y arena en imágenes satelitales del área del Río Amacuzac. Además de los resultados desde el punto de vista comparativo para la tarea de clasificación multiclase, las imágenes etiquetadas servirán como entradas para un estudio más general que pretende determinar la dinámica espacio temporal del río Amacuzac y sus zonas riparias.

Para la selección del área de estudio se tuvieron en cuenta algunos de los puntos de interés seleccionados por Mejía-Mojica (2012). En estos se han encontrado alteraciones en el ecosistema dentro del río Amacuzac. Teniendo en cuenta los datos disponibles y el fin de determinar la dinámica espacial del río y sus zonas riparias en investigaciones posteriores, el área de estudio quedó limitada entre las estaciones hidrométricas 18232 y 18437 (Figura 1). Ubicada en la región hidrológica 18, pertenece a la cuenca del río Balsas; abarca una superficie aproximada de 430 km² y una longitud de 50 km.

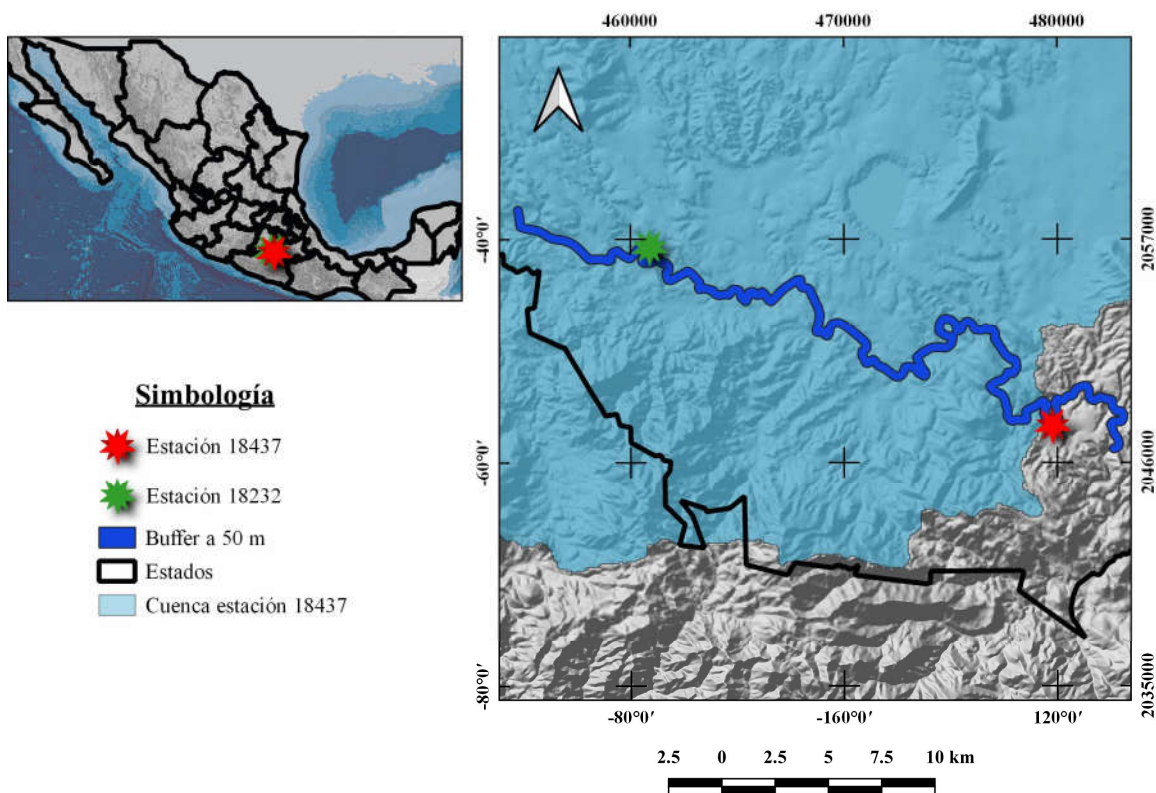


FIGURA 1.

DELIMITACIÓN DEL ÁREA DE ESTUDIO

El área comprende parte de los estados de Morelos, Guerrero y Estado de México. Geográficamente se localiza entre los paralelos 18°27' y 19°14' de latitud norte y los

meridianos $-98^{\circ}37'$ y $-100^{\circ}00'$ de longitud oeste. Se encuentra delimitada por las siguientes regiones y cuencas hidrológicas: al Norte por las Regiones Hidrológicas número 26 Pánuco y 12 Lerma-Santiago; al Sur por la cuenca hidrológica Río Bajo Atoyac; al Oeste por las cuencas hidrológicas Río Cutzamala y Río Medio Balsas; y al Este por la cuenca hidrológica Río Nexapa (DOF, 2011; Soares, 2008).

METODOLOGÍA

Procesamiento de la información

En el primer paso del proceso metodológico se obtuvieron las imágenes Sentinel 2⁵ a 10 m de resolución espacial del área de estudio entre un periodo de 2015-2021. Se realizó un filtro para descartar escenas con alta presencia de nubes sobre el área de estudio y aquellas en las que el satélite no tomó la imagen correctamente. Se procedió a la corrección atmosférica y geométrica de las imágenes satelitales con el programa SNAP (*Sentinels Application Platform*) versión 8.0 con el modelo *Sen2cor*³.

Se continuó el análisis con el Software de sistema de Información Geográfica QGIS versión 3.14. Primeramente se realizó un *buffer* a 50 m a partir del eje del río, con el cual se cubría el área de interés para la posterior investigación. Se procedió a la corrección de las nubes, fueron seleccionados los niveles de salida 2, 3, 4, 5, 6, y 8 como aquellos sin errores en la calidad. Fueron filtradas descartando aquellas que tenían una alta presencia de nubes sobre el área de estudio y aquellas en las que el satélite no tomó la imagen correctamente.

Con *Sen2Cor* se pasaron las imágenes del nivel de producto 1C de manera íntegra sin ser adulterado al 2A. Aquí se trabajó en la corrección atmosférica de los datos de entrada de Nivel 1C de la parte superior de la atmósfera.

Las salidas de Nivel 2A fueron: BOA (*Bottom-Of Atmosphere*), opcionalmente imágenes de reflectancia corregidas por terreno y nubes (cirros), mapas AOT (*Aerosol Optical Thickness*), WVP (*Water Vapour*), SLC (*Scene Classification*) e indicadores de calidad, incluidas las probabilidades de nubes y nieve. El producto de salida fue de imágenes JPEG, con tres resoluciones diferentes, 60, 20 y 10 m.

Finalmente se sumaron todas las imágenes de la banda SCL (*Scene Classification*), para determinar los píxeles en los que no se presentaba suficiente información (nubes). Se

⁵ Este algoritmo se encuentra disponible de forma gratuita en <http://www.cs.princeton.edu/schapiro/maxent/>

conformó una base de datos con áreas donde la omisión de información no fuera mayor a 5 píxeles contiguos. De esta forma, se aseguró que la disponibilidad de información fuese mayor al 95 % de la cobertura de la imagen satelital, creando así una máscara general para no cometer errores al momento de seleccionar los sitios de entrenamientos.

Para construir los sitios de entrenamiento, se determinó la cantidad de píxeles representativos de las imágenes, haciendo uso de la herramienta *Online Raosoft*⁶. Esta herramienta utiliza el tamaño total de una población para estimar el tamaño de muestra representativa. En nuestro caso, la población corresponde al total de píxeles de la imagen y la muestra a la cantidad de píxeles representativos. El valor obtenido representó el número mínimo de píxeles que se tomaron en cada imagen para determinar los sitios de entrenamiento.

Corridas de algoritmos

Se seleccionaron tres clases para los sitios de entrenamiento que se distribuyeron entre los píxeles anteriormente determinados. Las clases de interés fueron arena, vegetación y río. Para la primera se consideraron imágenes comprendidas entre los meses de noviembre a abril, periodo de lluvia. Para la vegetación y el río se seleccionaron imágenes entre los meses de mayo a octubre que corresponde al periodo de seca. De esta forma se garantizó que siempre los píxeles seleccionados correspondieran a la misma clase, aún con la dinámica temporal, pues para ello se consideró el caso más desfavorable.

Para el procesamiento con *MaxEnt* fueron introducidas las clases como datos de presencia (arena, vegetación y río) y como variables las bandas B2, B3, B4 Y B8 con una resolución de 10m y longitudes de ondas centrales respectivas de 492.4, 559.8, 664.6, 832.8. Cabe resaltar que el programa fue ejecutado con los valores predeterminados del mismo.

Como resultado de la ejecución del *MaxEnt* se obtuvieron mapas independientes para cada una de las clases. Los píxeles de cada uno de estos mapas están compuestos por valores de 0 a 1, que representan la probabilidad de encontrar la clase. Es decir, si en un pixel se encuentra el valor de 1, significa que es 100 % seguro que corresponde a la clase representada por esa imagen. Por el contrario, si se encuentra el valor de 0, significa que hay 0 % de probabilidad de encontrar ese pixel en la clase.

Como cada una de las imágenes de salida del *MaxEnt* representa una clase diferente, se determinó que se evaluaría en un mismo pixel cada una de las imágenes y se seleccionaría el de mayor probabilidad. Con estos píxeles de mayor probabilidad se creó el mapa final de clasificación.

⁶Disponible gratuitamente en <http://www.raosoft.com/samplesize.html>

Adicionalmente se utilizó el algoritmo de clasificación SVM de la caja de herramientas de Orfeo 6.7.0 en QGIS. Se le dio como entrada el compuesto de las bandas B2, B3, B4, B8 y los sitios de entrenamiento en forma de polígonos (Gualtieri, 2009). En este caso, el SVM se procesó con los valores predeterminados del módulo. Con la ejecución del algoritmo se obtuvo el mapa de clasificación para cada una de las imágenes.

Validación de resultados

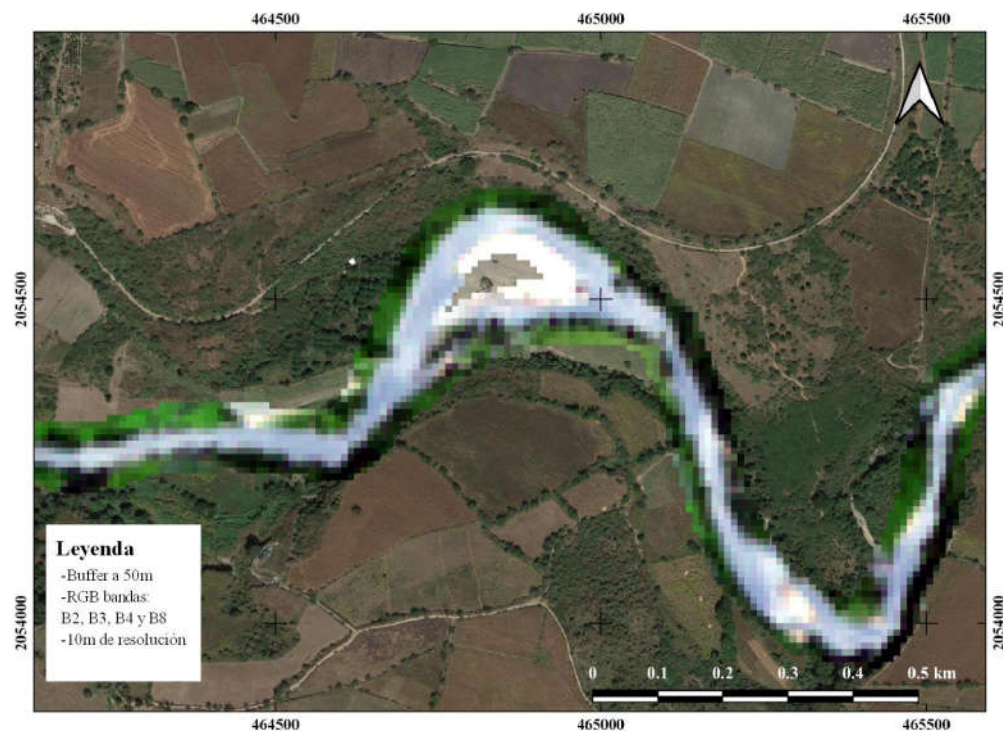
Para realizar la validación, se crearon puntos de control diferentes a los sitios de entrenamiento, pero siguiendo la metodología para determinar estos últimos. Se realizaron comparaciones entre los puntos seleccionados y cada una de las clasificaciones arrojadas por ambos algoritmos. A partir de éstas fueron creadas las Matrices de Confusión de la cual se obtuvieron los valores de *Overall Accuracy* (Stehman, 1997) y *Producer's Accuracy* (Congalton, 1991). Además, se crearon las curvas ROC y se obtuvieron los valores de AUC. Finalmente, con estos valores se logró determinar el algoritmo que clasificó mejor el área de estudio.

RESULTADOS

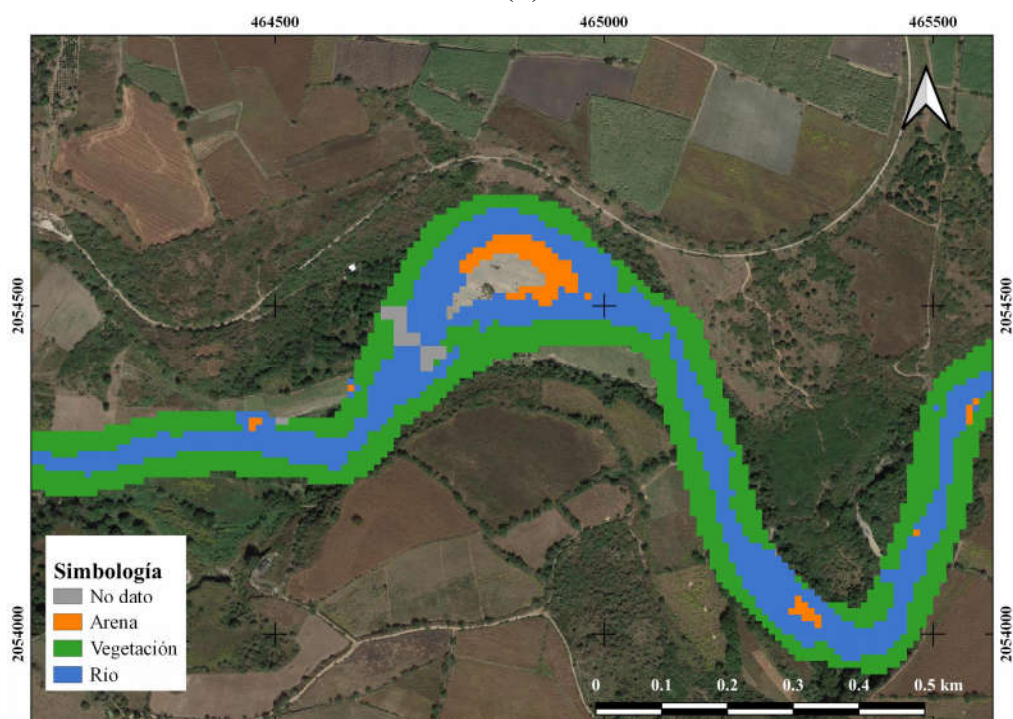
Clasificación de imágenes

Se descargó un total de 167 imágenes y luego del filtro quedaron 100 para procesar en el SNAP. Después, en la Figura 2 (a) fueron trabajadas en QGIS para realizar las correcciones geométricas y las correcciones en las nubes. El tamaño de muestra arrojado por el *Roasoft* fue de 384 píxeles, por lo que se tomaron valores por encima de este para determinar los sitios de entrenamientos.

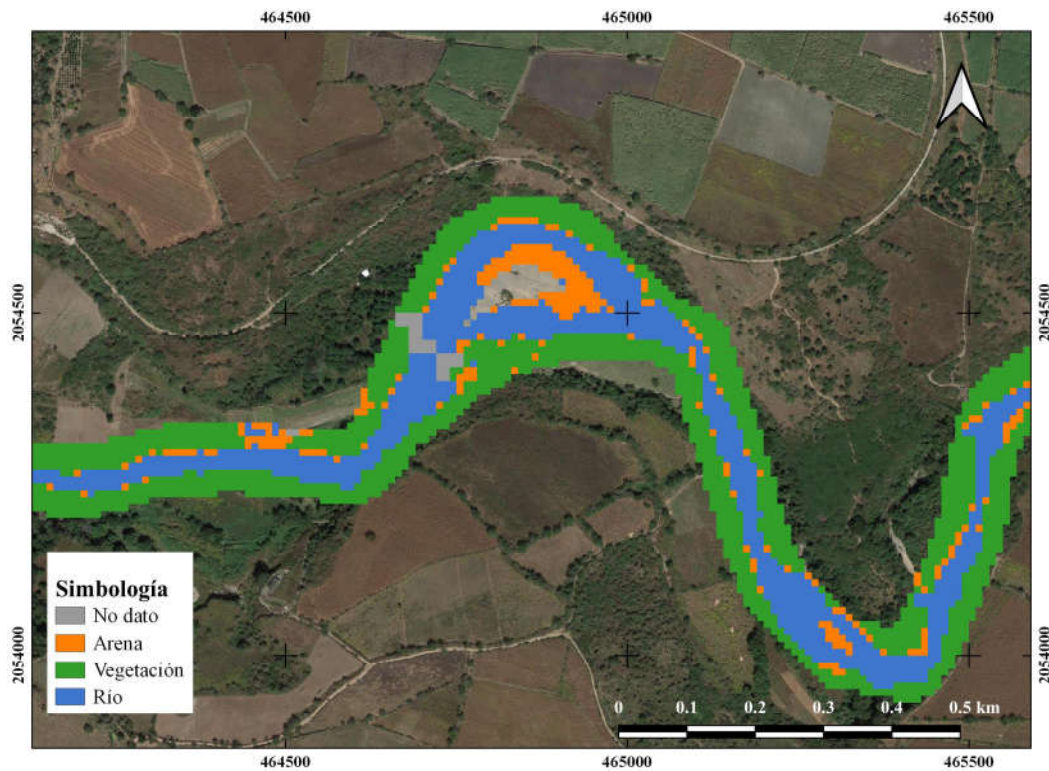
Fueron creados los sitios de entrenamiento para el *MaxEnt* y el SVM teniendo en cuenta las clases. Una vez corridos ambos algoritmos quedaron conformados los mapas de clasificación por cuenca. En la Figura 2 (b) se observa cómo el *MaxEnt* no fue capaz de identificar correctamente la arena, obteniendo mejores resultados para las otras clases. Por el contrario, en la Figura 2 (c), el SVM logró identificar un mayor grado de detalle en la arena, manteniendo resultados similares en cuanto al río y la vegetación.



(A)



(B)



(C)

FIGURA 2

(A) COMPOSICIÓN DE COLOR VERDADERO RGB BANDAS B2, B3, B4 Y B8. (B) CLASIFICACIÓN CON MAXENT,
(C) CLASIFICACIÓN CON SVM. IMAGEN 2017/10/14.

Validación de resultados

En la Figura 3 se pueden observar los resultados obtenidos a partir de matriz de confusión para ambos algoritmos. En estos se observa cómo el *MaxEnt* presenta valores de *Producer's Accuracy* inferiores al 70 % (línea naranja) para todas las clases excepto el río, mientras que el SVM también presenta valores por debajo de este umbral, pero en menor proporción. En cuanto al *Overall Accuracy* se observa que para el *MaxEnt* existen algunos valores por debajo del umbral establecido. Sin embargo, el SVM no presenta ningún valor por debajo de este umbral y la gran mayoría de sus valores están por encima del 80 %. Como resultado de la comparación de los dos algoritmos se determinó que el SVM presenta cierta superioridad ante *MaxEnt*.

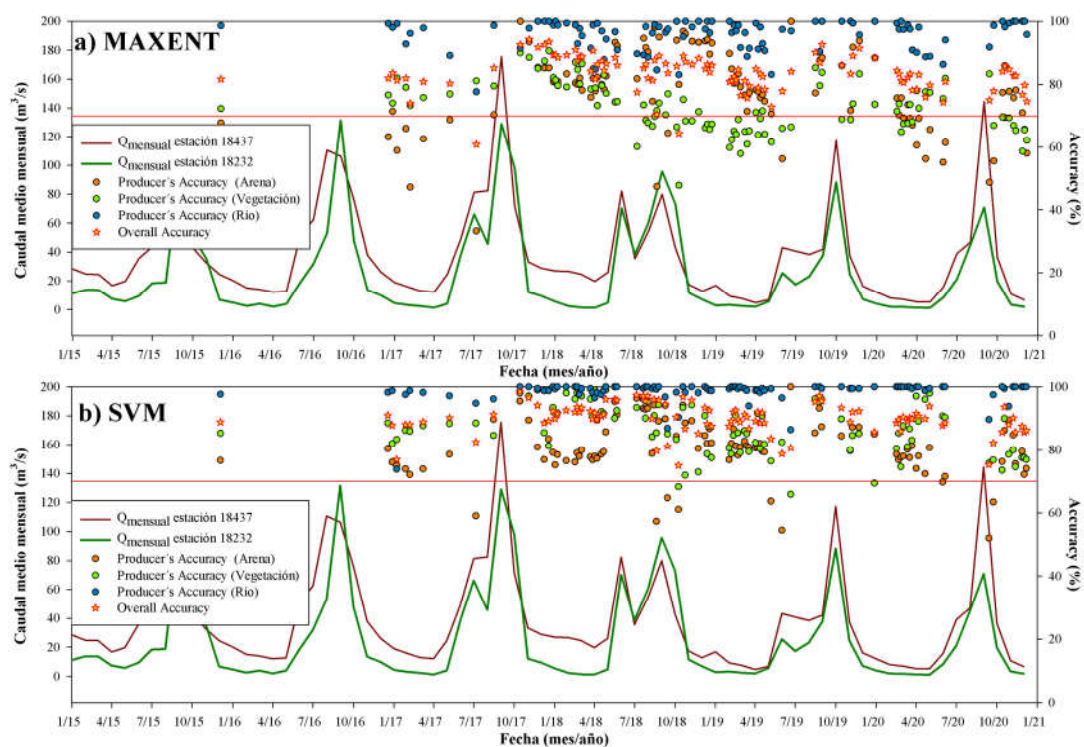


FIGURA 3.

PARÁMETROS DE CALIDAD DE CLASIFICACIÓN CON MAXENT OBTENIDOS A PARTIR DE MATRIZ DE CONFUSIÓN

Los valores arrojados por las curvas ROC, es decir los valores de AUC se muestran resumidos para ambos algoritmos en la Figura 4. En la misma se puede observar cómo la clasificación de la vegetación es la que mejor logran identificar los algoritmos, seguido por la del río y por último la de arena. Cabe resaltar que SVM presenta, para cada una de las clases, valores superiores a los reportados con el *MaxEnt*. Guiándonos por el valor límite escogido anteriormente se determinó que el algoritmo SVM muestra menos valores por debajo de éste, lo que representa una mejor clasificación. La comparación entre los algoritmos de clasificación *MaxEnt* y SVM permitió explorar el comportamiento del río desde dos puntos diferentes. Los resultados más cercanos a la realidad fueron obtenidos de la clasificación a partir del SVM.

En este estudio, las clasificaciones obtenidas por el *MaxEnt* con la metodología llevada a cabo por los autores, se presentó ligeramente inferior a la obtenida por el SVM. Cabe resaltar que en otros estudios (Li & Guo, 2010) que realizan comparaciones entre éstos, pero con una sola clase, el *MaxEnt* ha obtenido resultados más certeros. Sin embargo, esos resultados son de gran relevancia debido a los escasos estudios realizados respecto del tema,

además de que sienta las bases para determinar la dinámica espacial del río Amacuzac y sus zonas riparias.

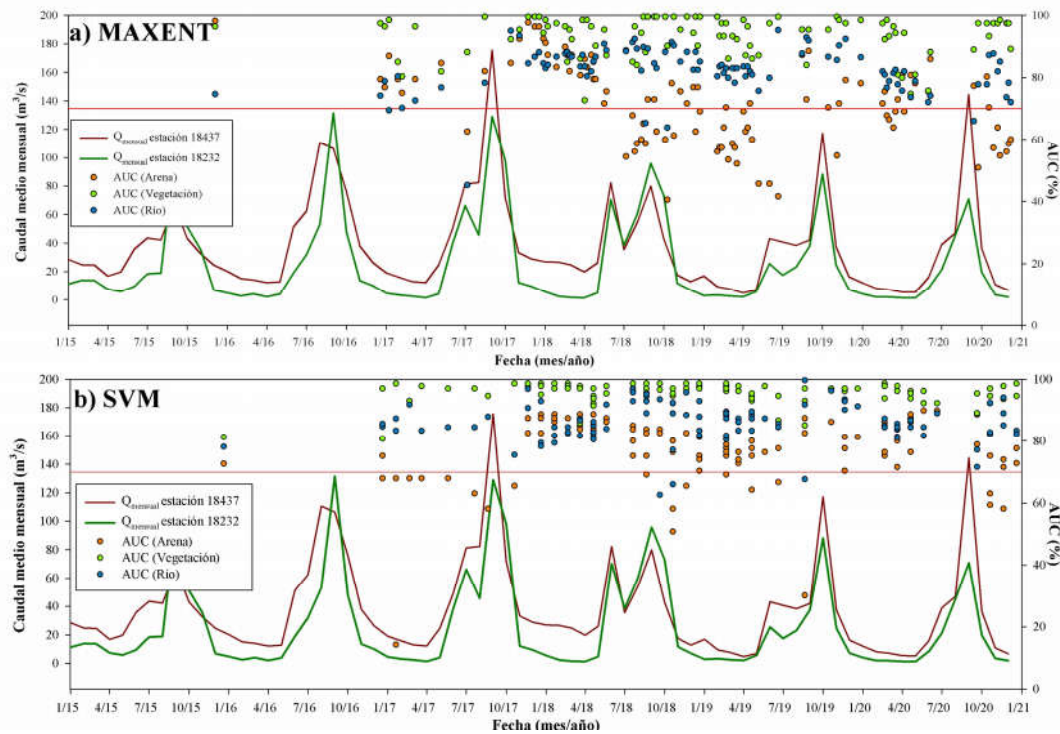


FIGURA 4

VALORES DE AUC DE LAS CLASIFICACIONES CON SVM Y MAXENT A PARTIR DE LA CURVA ROC.

CONCLUSIONES

Con el análisis de imágenes de satélites Sentinel 2 a 10m de resolución espacial, se logró obtener una buena precisión en la clasificación del área de estudio, logrando identificar cada una de las clases para la zona de interés.

Además, se comprobó que el *MaxEnt* puede ser utilizado como un clasificador multiclase con los parámetros predeterminados con los que trabaja, sin embargo, el SVM mostró una mayor precisión. Los resultados obtenidos se pueden deber únicamente a la metodología llevada a cabo, sin embargo, en futuras investigaciones se recomienda explorar los parámetros de entrada del *MaxEnt* y con ello quizás esperar resultados más acertados.

REFERENCIAS

- Chen, W., Pourghasemi, H.R., Kornejady, A. & Zhang N. (2017). Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma*, 305, 314–32. <http://doi.org/10.1016/j.geoderma.2017.06.020>
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 37(1) 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)
- Cortes, C., & Vapnik, V. (2006). Support-vector networks, *Machine learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- DOF. (2011). Acuerdo por el que se dan a conocer los estudios técnicos de aguas nacionales superficiales de la Región Hidrológica número 18 Balsas. Diario Oficial de la Federación, México. <https://dof.gob.mx>.
- Farrell, A., Wang, G., Rush, S. A., Martin, J.A., Belant, J. L., Butler, A. B., & Godwin, D. (2019). Machine learning of large-scale spatial distributions of wild turkeys with high-dimensional environmental data. *Ecology and evolution*, 9(10), 5938–5949. <https://doi.org/10.1002/ece3.5177>
- Gualtieri, J. A. (2009). The support vector machine (SVM) algorithm for supervised classification of hyperspectral remote sensing data. En G. Camps-Valls y L. Bruzzone (Eds.), *Kernel methods for remote sensing data analysis*, 49–83, Singapore, John Wiley & Sons, Ltd.
- Krawczyk, B., Galar, M., Woźniak, M., Bustince, H. & Herrera, F. (2018). Dynamic ensemble selection for multi-class classification with one-class classifiers. *Pattern Recognition*, 83, 34–51. <https://doi.org/10.1016/j.patcog.2018.05.015>
- Li, W. & Guo, Q. (2010). A maximum entropy approach to one-class classification of remote sensing imagery. *International Journal of Remote Sensing*, 1(8), 227–2235. <https://doi.org/10.1080/01431161003702245>
- Mack, B. & Waske, B. (2017). In-depth comparisons of MaxEnt, biased SVM and one-class SVM for one-class classification of remote sensing data. *Remote sensing letters*, 8(3), 290–299. <http://doi.org/10.1080/2150704x.2016.1265689>
- Mejía-Mojica, H. (2012). Estructura y dinámica de una comunidad de peces bajo la influencia de un componente invasor.
- Phillips, S. J., Anderson, R. P. & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190 (3–4), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>

- Pratiwi, Y. & Widodo, E. (2017). Comparison of Maximum Entropy and Support Vector Machine Methods for Sentiment Analysis of Peralite Product through Twitter Social Network. *International Journal of Advances in Electronics and Computer Science*, 4(12), 10–14.
- Rahimian Boogar, A., Salehi, H., Pourghasemi, H. R. & Blaschke, T. (2019). Predicting habitat suitability and conserving *Juniperus* spp. habitat using SVM and maximum entropy machine learning techniques. *Water*, 11(10), 2049. <https://doi.org/10.3390/w11102049>
- Soares, D. (2008). Recursos hídricos de la cuenca del Amacuzac. *Inventio, La génesis de la cultura universitaria en Morelos*, 4(8). 15-23.
- Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote sensing of Environment*, 62(1), 77–89. [https://doi.org/10.1016/S0034-4257\(97\)00083-7](https://doi.org/10.1016/S0034-4257(97)00083-7)

Electropolimerización de polipirrol usando electrodos de titanio

Electropolymerization of polypyrrole using titanium electrodes

Rosario Ramírez Segundo^{1,2}, J. Cuauhtémoc Palacios González¹, Guillermo Jesús Cruz², María Guadalupe Olayo González², Elena Colín Orozco¹, Fernando Gabriel Flores-Nava^{1,2}, Maribel González-Torres³, Lidia María Gómez-Jiménez^{4,*}, Ricardo Valdivia Barrientos²

Recibido: 7 de octubre de 2021 – Aceptado: 14 de marzo de 2022

RESUMEN

En este trabajo se sintetizaron películas de polipirrol (PPy) por síntesis electroquímica, usando un electrolito no convencional como yoduro de sodio (NaI) y electrodos de titanio (Ti). La película obtenida presentó una morfología rugosa con partículas sobre su

superficie. El análisis elemental indicó la presencia de C, N, O y I sobre la superficie de la película. Con la presencia de I, se demostró que el yodo se adhiere a la estructura del PPy. En el análisis estructural se observaron grupos conjugados de $-C=$ representados como $\approx C \approx$ y que no se encuentran en el pirrol, pero sí en el polipirrol, lo que indica que las

¹ Facultad de Ingeniería, Universidad Autónoma del Estado de México

² Instituto Nacional de Investigaciones Nucleares

³ División de Ingeniería Industrial, Tecnológico de Estudios Superiores de Tianguistenco

⁴ Universidad Politécnica de Oztolotepec

* Autores de correspondencia.

ecolino@uacemex.mx

lm_gomezj@hotmail.com



condiciones establecidas favorecieron el dopado del material.

PALABRAS CLAVE: Electroquímica, polipirrol, película.

ABSTRACT

The electrochemical synthesis of polypyrrole (PPy) films using a non-conventional electrolyte as sodium iodide (NaI) and titanium (Ti) electrodes is presented in this work to obtain PPy doped

with iodine. The film has a rough surface with particles. C, N, O and I were found on the surface of PPy. The % atomic of I indicates that the polymer was successfully doped. PPy presents conjugate groups $-C=$ represented as $\approx C \approx$, which are not found in pyrrole and if in polypyrroles, which indicates that the established conditions favored the doping of the material.

KEYWORDS: electrochemistry, polypyrrole, films.

INTRODUCCIÓN

Los factores que influyen en la electropolimerización de polímeros electroactivos destacan la concentración, el pH, el tiempo, la naturaleza del electrolito y de los electrodos. En particular, los electrodos tienen un rol importante, ya que son el medio por el que se transportan los electrones que generan especies oxidantes en la celda (Peña et al., 2013). El tipo de metal que se usa en la electrosíntesis influye en los procesos electroquímicos (oxidación y reducción del monómero) y en la morfología resultante del polímero. Se han usado como electrodos: acero inoxidable (inox), plata (Ag), indio-estaño (ITO) y oro (Au). Estos presentan un potencial de oxidación del pirrol de 0.75, 0.8, 0.86 y 0.65 V, respectivamente. Además, reportan morfologías granulares y en forma de coliflor, la cual es típica de los polipirroles (Saugo et al., 2018; Wu et al., 2021).

El material que adquiere interés es el uso del titanio (Ti), el cual es un metal que presenta resistencia a la corrosión, ya que, al reaccionar con el oxígeno, forma una capa superficial de óxido que funciona como protección (Lario-Femenía, et al., 2016).

Por otra parte, el polipirrol (PPy) posee una estructura conjugada de enlaces sencillos (σ) y dobles (π) a lo largo de la cadena polimérica (Lakard, 2020; Pang et al., 2020). Se ha reportado que, cuando el polímero se obtiene por vía electroquímica, puede aumentar su conductividad eléctrica hasta 10 órdenes de magnitud, pasando de tener propiedades de un material aislante a la de un semiconductor, incluso puede llegar a ser un polímero conductor (Thorat et al., 2018; Arrieta & Tarazona, 2009). Este comportamiento ocurre debido a que los residuos provenientes de la solución electrolítica actúan como agentes dopantes favoreciendo la conductividad eléctrica a lo largo de la cadena polimérica.

Dado que las propiedades eléctricas del material se favorecen, el PPy tiene aplicaciones en las áreas de energía y bioingeniería, entre otras. Para la primera, el polímero pasa a ser uno de los componentes de un supercapacitor y/o celdas solares para mejorar la absorción de la radiación electromagnética (Belkhedkar & Ubale, 2014; Patil et al., 2011; Namsheer & Chandra, 2021). En la segunda, el PPy se vuelve de particular interés debido a que se obtienen superficies ricas en grupos amina (NH), lo que al interactuar con tejidos biológicos se favorece la generación celular. Por lo anterior, el polipirrol se le ha clasificado dentro del grupo de los materiales que presentan propiedades biocompatibles (Thompson et al., 2011; Rowlands & Cooper-White, 2008). De ahí el interés de investigar la

electropolimerización del PPy usando electrodos de diferentes materiales y soluciones electrolíticas no convencionales.

Considerando lo anterior, el objetivo del trabajo consiste en la electropolimerización del PPy usando electrodos de titanio para obtener películas de PPy dopadas con yodo, que puedan aplicarse como superficies biocompatibles o como materiales que favorezcan la absorción de radiación electromagnética.

METODOLOGÍA

Materiales y equipos

- Fuente Modelo 2450 *Interactive SourceMeter® Instrument*.
- Celda electrolítica diseño propio (UAEM).
- Balanza analítica, marca *Ohaus*.
- Parrilla de calentamiento con agitación magnética, marca *StableTemp dP Cole-parmet*.
- Estufa de cultivo, marca *Thermo Scientific*.
- pHmetro, Serie 700 *Eutech Instruments Technology Made Easy*, marca OAKION®.

Reactivos

- Yoduro de sodio, *Sigma-Aldrich* de 99.5 % de pureza.
- Pirrol, *Sigma-Aldrich* de 98 % de pureza.
- Agua desionizada, resistividad de 18.2 Ω .
- Acetona de 99.5 % de pureza.
- Placa de titanio *Sigma-Aldrich*.

Preparación de soluciones

Como solución electrolítica, se usó NaI 0.1 M a pH 6, mientras que la concentración de Py fue de 0.1 M a pH 7. Es importante mencionar que cuando estuvieron en contacto las dos soluciones Py/NaI, el pH final fue de 6.68.

Electropolimerización

Se electropolimerizó PPy usando la técnica de voltamperometría cíclica en una celda electrolítica de 250 mL (Figura 1). El procedimiento consistió en agregar el electrolito NaI

en la celda electrolítica. Posteriormente, se agregó Py y se introdujeron los electrodos de titanio de forma rectangular con área de 30.2 cm^2 , como ánodo y cátodo. El potencial aplicado fue de -3 a +3 V a una velocidad constante de 20 mV/s. En total fueron 18 barridos durante 3 horas. Durante el proceso, las soluciones se mantuvieron con agitación constante.

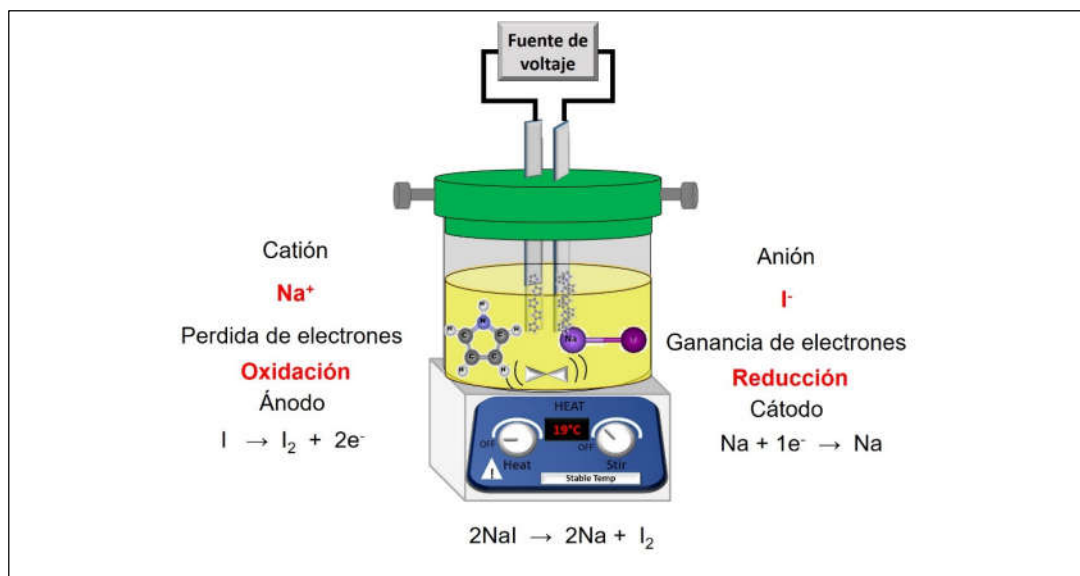


FIGURA 1.

CELDA ELECTROLÍTICA PARA LA FORMACIÓN DE LA PELÍCULA DE PPy.

Análisis de las películas de polipirrol

Se analizó la morfología de las películas de PPy con un Microscopio Electrónico de Barrido marca *Jeol IT-100*. El análisis estructural se realizó por Espectroscopia Infrarroja por Transformada de *Fourier* (FTIR), usando reflectancia total atenuada (ATR) en un equipo *Thermo Scientific Nicolet iS5*, con celda de diamante en el intervalo $550-4000 \text{ cm}^{-1}$.

RESULTADOS Y DISCUSIONES

Electropolimerización del polipirrol

En la Figura 2 se muestran los voltamperogramas que se obtuvieron en el sistema Na/Py durante 18 ciclos a velocidad constante de 20 mV/s. Se observan las corrientes de pico anódico (i_{pa}) y catódico (i_{pc}) en las curvas voltamperométricas. La i_{pa} fue de 8.56 mA con

potencial de 609 mV en el primer ciclo y se desplazó a 117.05 mA con potencial de 1552 mV en el último ciclo. Mientras que la corriente catódica se encontró en 17.66 mA, con potencial de 457 mV y se desplazó a 109.2 mA y 1544 mV.

Con lo anterior, se observa que la corriente aumenta, indicando un incremento de la carga total durante la electropolimerización, debido a que las especies en el electrolito son consumidas en la superficie del electrodo. Además, al mismo tiempo incrementa el potencial en el sistema, lo que favorece el aumento de la transferencia de electrones (Zhang et al., 2013). Este comportamiento se presenta en los polímeros conductores como el polipirrol. Por lo que el tipo de electrodo influye en los procesos de oxidación y reducción, ya que, usando electrodos de acero inoxidable, el potencial de oxidación fue de 750 mV, comparado con lo que se obtiene en este trabajo que fue a potencial de oxidación de 609 mV. Con ello se puede decir que, usando electrodos de Ti, se requiere de menor energía para oxidar al PPy, usando como electrolito a NaI. Los voltamperogramas que se presentan en la Figura 2, indican que se tiene un comportamiento reversible usando electrodos de titanio y una solución electrolítica de NaI, debido a que se observa un pico de oxidación y otro de reducción. En contraste, cuando se hace uso de sulfatos como electrolito para electropolimerizar PPy, el proceso se vuelve irreversible, ya que solo presenta un pico de oxidación (Jureviciute & Bruckenstein, 2003). Con lo anterior, el tipo de material del electrodo, la solución electrolítica y el pH influyen directamente en las propiedades finales del material.

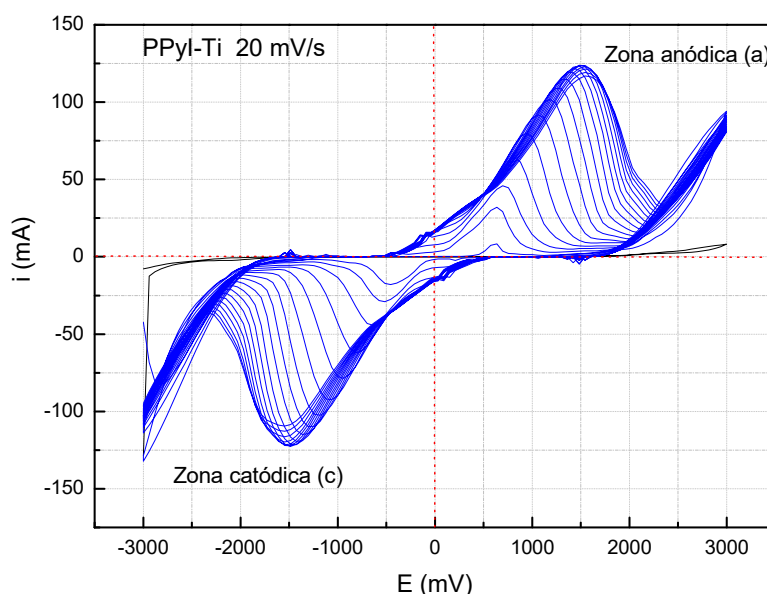


FIGURA 2.
VOLTAMPEROGRAMAS DE PPy/I DE -0.3 0.3 A 20 mV/s EN NaI 0.1M.

Análisis morfológico

En la Figura 3, se muestra la morfología de la película de PPy a 20 mV/s a 500x. Se observa que el crecimiento de la película es por capas con una superficie rugosa y granular en algunas zonas. Algunos de los trabajos que se han reportado presentan morfologías en forma de coliflor, como partículas y aglomerados (Stejskal et al., 2010). Sin embargo, el tipo de morfología que se obtenga dependerá de las condiciones de síntesis y del electrolito empleado. Con la morfología que se observa, se concluye que efectivamente el tipo de material que se usa como electrodo influye en el tipo de morfología, ya que en este trabajo, con el uso de Ti como electrodos, permitió obtener una superficie rugosa y el crecimiento fue en forma de capas pequeñas cavidades a lo largo de la película. Con este tipo de morfología probablemente podría favorecer la adsorción electromagnética, ya que en los huecos de la superficie no se reflejaría el haz como cuando se tienen superficies homogéneas lisas que hacen el efecto espejo, sino parte del haz se absorbería por las cavidades o huecos de la superficie (Niu, et al., 2019).

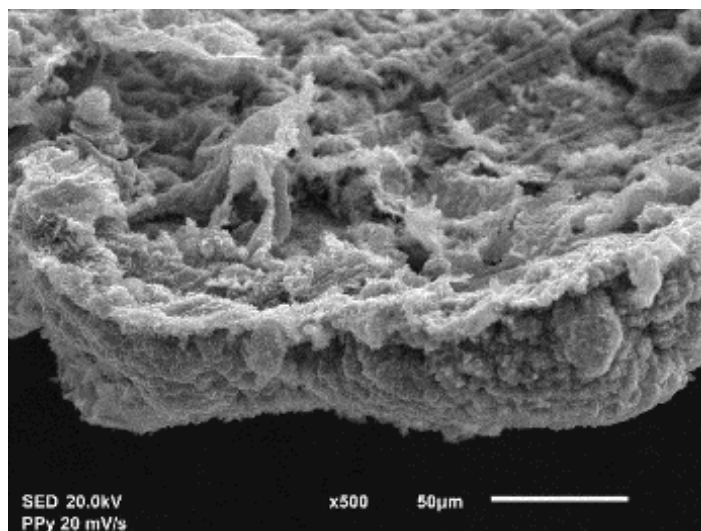


FIGURA 3.

MORFOLOGÍA DE LA PELÍCULA POLIMÉRICA DE POLIPIRROL A 500X.

Análisis estructural

En la Figura 4 se presentan los espectros de IR-ATR de pirrol (línea negra) y de polipirrol (línea azul). Los grupos funcionales corresponden al polipirrol, los cuales se encuentran en 2365 cm^{-1} indicando la presencia del doble enlace consecutivo de $\approx\text{C}\approx$ que proviene de los enlaces del Carbono α y β del anillo aromático. En la absorción de 1517 cm^{-1} , se observan

los dobles enlaces C=C y en 1410 cm^{-1} al grupo amina N-H (Niu et al., 1998; González-Torres et al., 2019). En 1290 cm^{-1} se encuentra el grupo funcional C-O, lo que indica que el polímero se deshidrogena para posteriormente oxidarse.

Finalmente, las absorciones en 1150, 1022, 664 y 594 cm^{-1} corresponden al grupo alifático C-H de la ruptura del pirrol. Con los resultados obtenidos, es importante mencionar que en el espectro de PPy no se observa la absorción del grupo amina de pirrol, debido a que probablemente se deshidrogena para formar nuevas estructuras que se unen principalmente por el N. Además, no se observan enlaces triples y la deshidrogenación de los C-H es ligera, por lo que se supone que los anillos de pirrol se conservan formando una cadena polimérica línea.

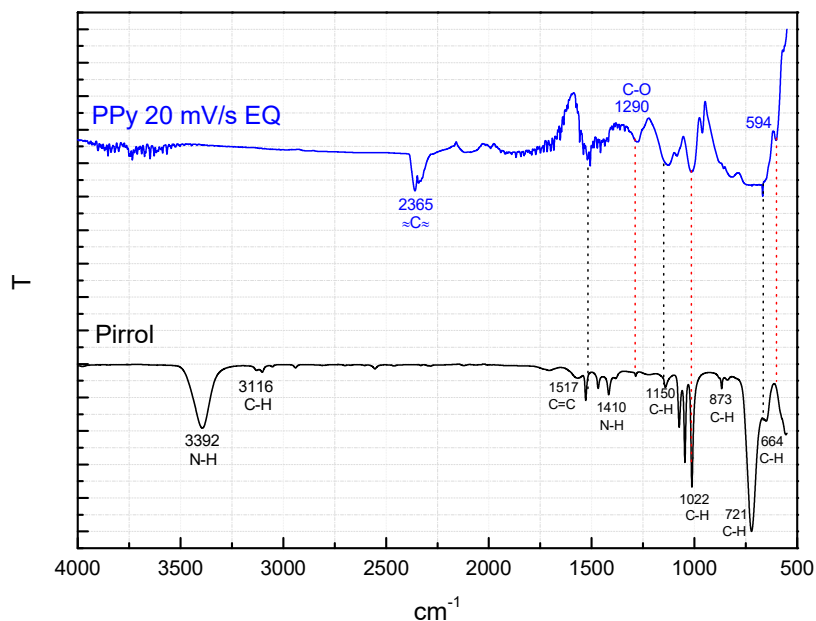


FIGURA 4.
ESPECTRO INFRARROJO DE LA PELÍCULA DE POLIPIRROL.

Mapeo Elemental

En la Figura 5 se muestra el mapeo elemental superficial de PPy en % atómico. Se encontró 78.68 % de C (rojo), 7.57 % de N (verde), 5.21 % de O (amarillo) y 8.55 % de I (morado). El C y N forman parte de la estructura química del polipirrol, mientras que el I es el residuo del electrolito NaI y O es por la oxidación del medio. Con los porcentajes obtenidos se observó que se favoreció el dopaje. Con ello se forma el compuesto a base de PPy-I, lo que indica que posiblemente se mejore la conductividad eléctrica.

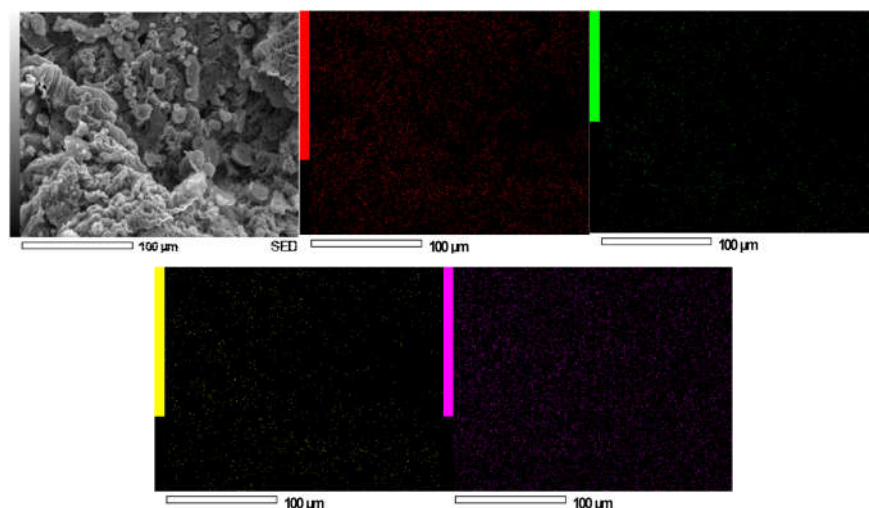


FIGURA 5.

MAPEO DE LA PELÍCULA DE PPy

CONCLUSIONES

Con los resultados obtenidos, se concluye que el efecto que provoca el uso de electrodos de diferente material influye directamente en el tipo de morfología, ya que con el uso de electrodos de Ti se obtuvo una película rugosa y granular con partículas sobre la superficie el crecimiento de una. En el mapeo elemental superficial se encontraron los elementos que corresponden al C, N, I y O. Dada la presencia de I se favorece la conductividad eléctrica del polímero. Con el análisis estructural se observó que ocurre un proceso de deshidrogenación. Los grupos funcionales que se observaron en el análisis estructural fueron $\approx C\approx$, $C=C$, $N-H$, $C-O$ y $C-H$ que corresponden a la estructura de polipirroles oxidados.

REFERENCIAS

- Arrieta, A. A. & Tarazona, R. L. (2009). Study of kinetic formation and the electrochemical behavior of polypyrrole films. *Journal of the Chilean Chemical Society*, 54(1), 14-19. <https://doi.org/10.4067/S0717-97072009000100004>
- Belkhedkar, M. R. & Ubale, A. U. (2014). Physical properties of nanostructured Mn₃O₄ thin films synthesized by SILAR method at room temperature for antibacterial application. *Journal of Molecular Structure*, 1068, 94–100. <https://doi.org/10.1016/j.molstruc.2014.03.050>
- Cruz, G. J., Morales, J. & Olayo, R. (1998). Films obtained by plasma polymerization of pyrrole. *Thin Solid Films*, 342(1-2), 119-126. [https://doi.org/10.1016/S0040-6090\(98\)01450-3](https://doi.org/10.1016/S0040-6090(98)01450-3)
- González-Torres, M., Olayo, M. G., Gómez, L. M., Morales, J., Olayo, R., Ramírez, R. Flores, F. G., Mejía-Cuero, M. R. & Cruz, G. J. (2019). Chemical interactions of heparin in porous polypyrrole, an example of drug-carrier destructive interaction. *Polymer Bulletin*, 77, 375-385. <https://doi.org/10.1007/s00289-019-02751-w>
- Jureviciute, I. & Bruckenstein, S. (2003). Electrochemical activity of chemically deposited polypyrrole films. *Journal of Solid State Electrochemistry*, 7(9), 554–560. <https://doi.org/10.1007/s10008-003-0384-x>
- Lakard, B. (2020). Electrochemical Biosensors Based on Conducting Polymers: A Review. *Applied Sciences*. 10(18), 6614. <https://doi.org/10.3390/app10186614>
- Lario-Femenía, J., Amigó Mata, A., Vicente-Escuder, A., Segovia-López, F. & Amigó, V. (2016). Desarrollo de las aleaciones de titanio y tratamientos superficiales para incrementar la vida útil de los implantes. *Revista de Metalurgia*, 52(4), e084-e096. <https://doi.org/10.3989/revmetalm.084>
- Namsheer, K. & Chandra, S. R. (2021). Conducting polymers: a comprehensive review on recent advances in synthesis, properties and applications. *RSC Advances*. 11(10), 5659-5697. <https://doi.org/10.1039/d0ra07800j>
- Niu, C., Zhu, T. & Lv, Y. (2019). Influence of Surface Morphology on Absorptivity of Light-Absorbing Materials. *International Journal of Photoenergy*, 2019, 1–9. <https://doi.org/10.1155/2019/1476217>
- Pang, A. L., Arsad, A. & Ahmadipour, M. (2020). Synthesis and factor affecting on the conductivity of polypyrrole: a short review. *Polymers for Advanced Technologies*, 32(4), 1428-1454. <https://doi.org/10.1002/pat.5201>

- Patil, U. M., Kulkarni, S. B., Jamadade, V. S. & Lokhande, C. D. (2011). Chemically synthesized hydrous RuO_2 thin films for supercapacitor application, *Journal of Alloys and Compounds*, 509(5), 1677–1682. <https://doi.org/10.1016/j.jallcom.2010.09.133>
- Peña, W.F., Villard, J.F., Ramírez, A.E. & Nikolaevish, E. (2013). Synthesis and Voltametric Characterization of Plated Electrodes Doped with Cerium, Used for Degradation of 4-Chlorophenol. *Revista Ciencia en Desarrollo*, 4(2), 49-61.
- Rowlands, A. S. & Cooper-White, J. J. (2008). Directing phenotype of vascular smooth muscle cells using electrically stimulated conducting polymer. *Biomaterials*, 29(34), 4510-4520. <https://doi.org/10.1016/j.biomaterials.2008.07.052>
- Saugo, M., Flamini, D. O. & Saidman, S. B. (2018). Formación electroquímica de películas de polipirrol sobre Nitinol a partir de soluciones de ácido sulfuro-succínico. *Revista Matéria*, 23(2), 1517-7076. <https://doi.org/10.1590/s1517-707620180002.0391>
- Stejskal, J., Sapurina, I. & Trchová, J. (2010). Polyaniline nanostructures and the role of aniline oligomers in their formation. *Progress in Polymer Science*, 35(12), 1420–1481. <https://doi.org/10.1016/j.progpolymsci.2010.07.006>
- Thompson, B. C., Moulton, S. E., Richardson, R. T. & Wallace, G. G. (2011). Effect of the dopant ani-on in polypyrrole on nerve growth and release of a neurotrophic protein. *Biomaterials*, 32(15), 3822-3831. <https://doi.org/10.1016/j.biomaterials.2011.01.053>
- Thorat, J. B., Mohite, S. V., Bagade, A. A. Shinde, T. J., Fulari, V. J. Raj-pure, K.Y. & Shine, N. S. (2018). Nanocrystalline Bi_2Te_3 thin films synthesized by electrodeposition method for photoelectrochemical application. *Material Science in Semiconductor Processing*, 79,119–126. <https://doi.org/10.1016/j.mssp.2018.02.002>
- Wu J., Wu S. & Sun W. (2021). Electropolymerization and application of polyox-ometalate-doped polypyrrole film electrodes in dye-sensitized solar cells. *Electrochemistry Communications*, 122, 106879. <https://doi.org/10.1016/j.elecom.2020.106879>
- Zhang, J., Zhang, H., Wu, J. & Zhang, J. (2013). Techniques for PEM Fuel Cell Testing and Diagnosis en Zhang J., Zhang H., Wu J., Zhang J (Eds). *Pem Fuel Cell Testing and Diagnosis*. (81–119). Elsevier.

Comportamiento de algoritmos de sobre-muestreo en Big Data

Behavior of over-sampling algorithms in Big Data

A. Guzmán-Ponce¹, César Ferri², J. S. Sánchez-Garreta¹,
J. R. Marcial-Romero³

Recibido: 17 de noviembre de 2021 – Aceptado: 4 de abril de 2022

RESUMEN

El desbalance de clases es una de las complejidades de los datos ampliamente estudiada en el campo de la ciencia de datos. A menudo dificulta el proceso de extracción de conocimiento, sesgando el aprendizaje hacia instancias de clase mayoritaria. La creciente generación de datos que estamos viviendo agrava el escenario anterior. Los desafíos en *Big Data* implica la necesidad de adaptar o crear nuevas técnicas para las restricciones de escalabilidad, dando lugar al desarrollo de técnicas que solventen el desbalance de

clases en grandes volúmenes de datos, siendo la mayoría de estas basadas en el algoritmo SMOTE, en razón de tener un mejor desempeño en conjuntos “pequeños”. En este trabajo realizamos un análisis del comportamiento de los métodos de sobre-muestreo en *Big Data*, a través de medidas de complejidad que permiten conocer las características de los conjuntos de datos procesados. Los resultados obtenidos corroboran que el problema de desbalance de clases en *Big Data* no es el único problema que debe abordarse; por otro lado, el comportamiento de SMOTE en *Big Data* no

¹ Department of Computer Languages and Systems, Institute of New Imaging Technologies, Universitat Jaume I, 12071 Castelló de la Plana, Spain

² Valencian Research Institute for Artificial Intelligence (VRAIN), Universitat Politècnica de València, Spain

³ Facultad de Ingeniería, Universidad Autónoma del Estado de México, Toluca, México.

Autor de correspondencia: aguzman@uji.es



es comparable al logrado en conjuntos de datos pequeños, debido a la presencia de redundancia por parte del proceso de interpolación.

PALABRAS CLAVE: sobre-muestreo; SMOTE; Big Data

ABSTRACT

Class imbalance is one of the data complexities widely studied in the field of classification. It often hinders the modelling process, biasing the learning towards majority-class examples. The fast-growing generation of data that we are experiencing, aggravates the former scenario. In Big Data imbalanced classification, the challenges imply both addressing the skewed class distribution, and the need to adapt or create new techniques for scalability constraints. Thus, becoming the development of

techniques that handled the class imbalance on a high amount of data, as most of these are based on the original SMOTE algorithm, by virtue of their better performance on "small" data sets. In this paper, we performed an analysis of the behaviour of over-sampling methods on Big Data, through complexity measures which allow knowing the characteristics of the processed data sets. The obtained results corroborated that the class imbalance problem on Big Data sets is not a unique problem that must be addressed; on the other hand, the behaviour for SMOTE on Big Data is not comparable to that achieved in "small data" problems, due to the presence of redundancy by the interpolation process

KEYWORDS: over-sampling; SMOTE; Big Data.

INTRODUCCIÓN

En la actualidad la recolección y el análisis de datos es de suma importancia para empresas y organizaciones, convirtiendo a los datos en un activo fundamental para la toma de decisiones. Para lograr este objetivo, los datos son usados para obtener patrones de los datos y generar un modelo que contiene el conocimiento extraído (García et al., 2020).

Con el incremento de datos provenientes de diversas fuentes con diferentes tipos de datos, a la capacidad de recopilarlos y procesarlos de manera adecuada se le conoce como *Big Data* (Luengo et al., 2020). Este concepto requiere de atención en el análisis de datos, dado que por su naturaleza los datos se encuentran bajo peculiaridades que decrementan el rendimiento de un clasificador, es decir, este tipo de algoritmos se ven afectados por la calidad de los datos. En *Big Data*, el término *Smart Data* hace referencia a datos de calidad suficiente para lograr modelos de alto rendimiento.

En una gran variedad de problemas del mundo real como nuevos descubrimientos físicos, diagnóstico de enfermedades raras, detección de fraude electrónico, entre otros, comparten la particularidad de no ser iguales en términos de tamaño de datos por clase, lo que se conduce a un escenario de desbalance (García et al., 2018). Esta situación sucede cuando una de las clases tiene un número de instancias significativamente mayor en comparación con el resto de las clases. En general, en un conjunto de dos clases, a la clase menos representada se conoce como positiva, denotada por C^+ , mientras que la clase más representada se conoce como negativa denotada por C^- (Basgall et al., 2019) (Figura 1).

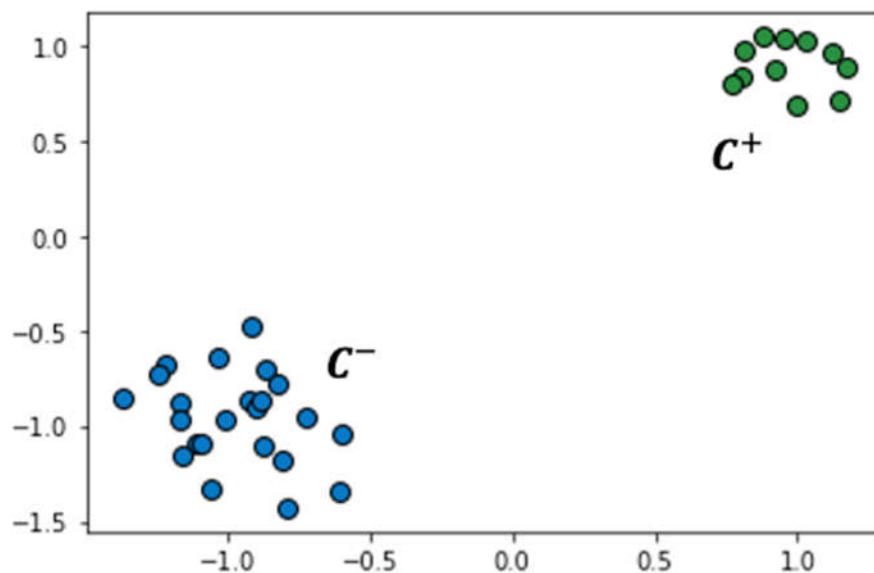


FIGURA 1.

CONJUNTO DE DATOS DESBALANCEADO

Una manera de abordar el problema de desbalance de clases es equilibrar la distribución de clases volviendo a muestrear el espacio de datos (S. García et al., 2016, Rendón et al., 2020), estas estrategias de re-muestreo son independientes del clasificador y sólo trabajan sobre el conjunto de datos, suelen dividirse en dos técnicas: sobre-muestreo y bajo-muestreo. La primera crea instancias de la clase positiva, mientras que la segunda técnica elimina instancias de la clase negativa.

El uso de métodos de sobre-muestreo se debe principalmente a disminuir la pérdida de datos en clase positiva que pueden ser útiles para generar conjuntos de datos de alta calidad, esto sucede comúnmente en escenarios con conjuntos de datos pequeños (*Small Data*) (V. García et al., 2020). Un conjunto de datos pequeños se caracteriza por su volumen generalmente limitado, siendo menor a Gigas de información, sin recopilación continua y con una variedad limitada de tipos de datos (Kitchin & Lauriault, 2015).

Sin embargo, en términos de grandes volúmenes de datos surge la cuestión ¿El sobre-muestreo mejora la calidad de los datos en escenarios de *Big Data*? Para responder esta cuestión, es necesario conocer las propiedades de los conjuntos de datos previas al tratamiento del desbalance de clases y posterior análisis una vez aplicada las técnicas de sobre-muestreo.

ESTRATEGIAS DE SOBRE-MUESTREO

A lo largo del tiempo se han propuesto un gran número de estrategias para hacer frente al desbalance de clases en escenarios *Small Data*. Sin embargo, para *Big Data* algunas investigaciones se han llevado a cabo mediante el escalado de métodos de sobre-muestreo bien conocidos (del Río et al., 2014).

El método clásico de sobre-muestreo es aleatorio (ROS) (Batista et al., 2004) el cual replica instancias tomadas aleatoriamente de la clase positiva hasta que el tamaño de la clase positiva es igual al de la clase negativa. Además de ROS, SMOTE (Gutiérrez et al., 2017) ha sido ampliamente usado, ya que funciona creando instancias a través de la interpolación de cada instancia positiva con sus vecinos.

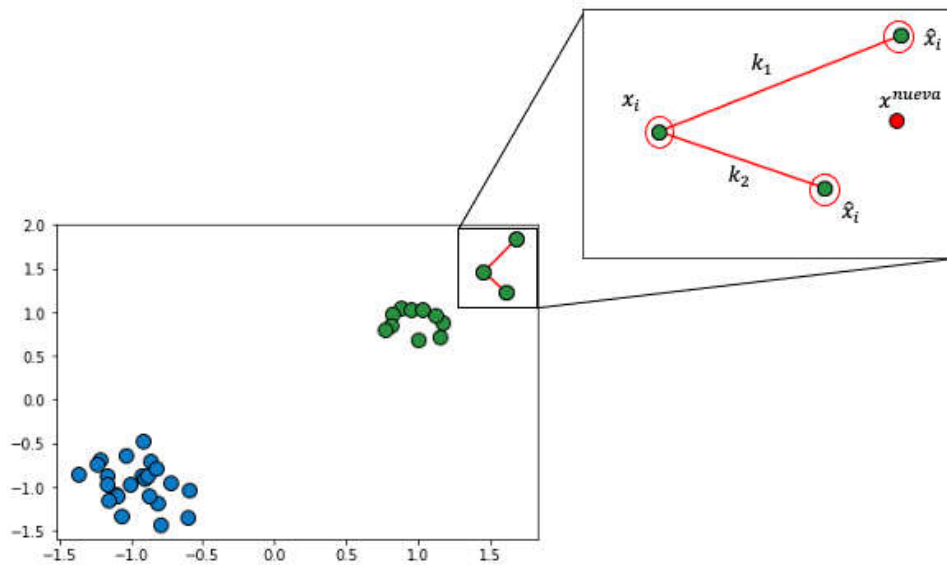


FIGURA 2.
REPRESENTACIÓN DE SMOTE

El algoritmo de SMOTE toma instancias de clase positiva y a través de sus vecinos más próximos genera nuevas instancias por interpolación. Del ejemplo de la Figura 2, la instancia x^{nueva} , se genera de combinar las características de la instancia x_i con características de sus vecinos.

En torno al método SMOTE también hay algunas propuestas que versan en el procedimiento para abordar el conjunto de datos, es decir, de manera local o global. SMOTE-MR (Basgall et al., 2018), es una versión local de SMOTE, donde para calcular los vecinos k de una instancia de la clase positiva, SMOTE-MR utiliza la misma partición a la que pertenece una instancia. Por el contrario, SMOTE-BD (Basgall et al., 2019) se desarrolla como una técnica global, donde el cálculo de k vecinos más cercanos (kNN) se basó en el método KNN-IS (Maillo et al., 2017), este algoritmo a través de múltiples reductores determina cuáles son los k vecinos más cercanos finales de cada partición.

MÉTRICAS DE CALIDAD DE LOS DATOS

Además del problema de desbalance de clases, existen otros factores que inciden en el rendimiento del clasificador, denominadas complejidades de los datos. El traslape de clases, implica regiones ambiguas del espacio de características donde la probabilidad de las clases es similar. Otro factor negativo es el ruido, lo que implica instancias mal etiquetadas, que pueden resultar en decisiones incorrectas por parte del clasificador. Con los grandes cúmulos de datos, el número de características excede significativamente el número de instancias del conjunto, lo que provoca el problema conocido como alta dimensionalidad. La Figura 3 representa estos problemas en un conjunto de dos clases.

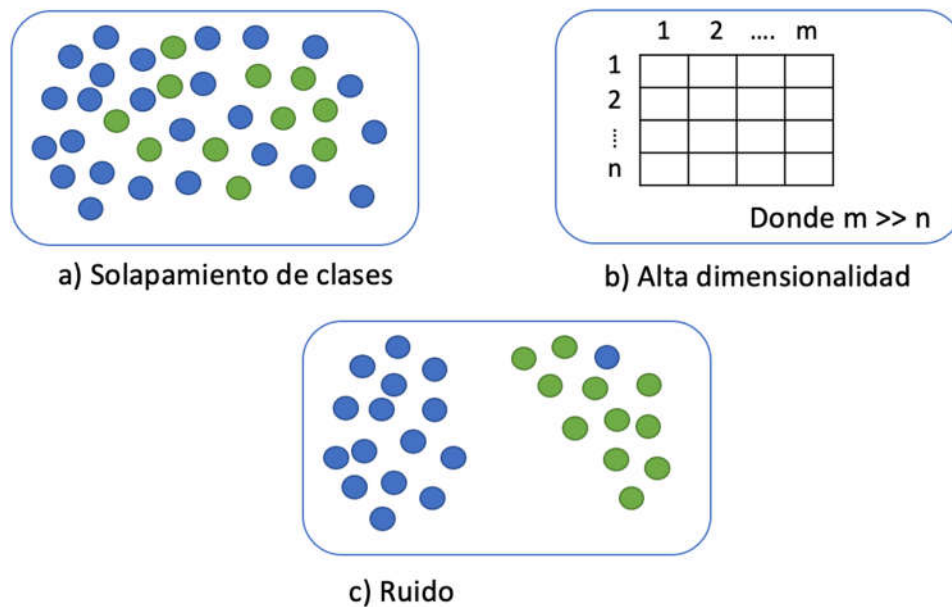


FIGURA 3.

COMPLEJIDADES DE LOS DATOS

Comúnmente el comportamiento que un conjunto de datos tiene en un clasificador es medido por el rendimiento del propio algoritmo. Sin embargo, existen algunas métricas que permiten determinar la calidad del conjunto de datos independientemente del clasificador.

Considere que un conjunto de datos CD esta conformado por n instancias, y cada instancia está compuesta por (x, y) atributos o características descritas como un arreglo $x = [x_1, \dots, x_m]$ de m características. La variable y está compuesta por n_c clases. Para determinar la presencia de traslape de clases la métrica comúnmente usada es el radio discriminante de Fisher ($F1$), (Maillo et al., 2020) escala esta métrica en escenarios de *Big Data*, la cual se define por la ecuación 1:

$$F1 = \max_{i=1}^m \frac{\sum_{j=1}^{n_c} n_{cj} (\mu_{cj}^{fi} - \mu^{fi})^2}{\sum_{j=i}^{n_c} \sum_{l=1}^{n_{cj}} (x_{li}^j - \mu_{cj}^{fi})^2} \quad (1)$$

Donde n_{cj} es el número de instancias de la clase j , μ_{cj}^{fi} es el promedio de la i -ésima característica de las instancias de la clase j , μ^{fi} es el promedio de la i -ésima característica de todas las instancias y x_{li}^j es el valor específico de la i -ésima característica para una instancia particular x .

Por otro lado, para tener una descripción general de cómo funcionan las características en conjunto, se usa la Eficiencia de funciones colectivas ($F4$). En general, cuenta el número de instancias afectadas considerando todas las características en un proceso iterativo. Esta definida por la ecuación 2:

$$F4 = \frac{n - n_o(f_{\max}(T_l))}{n} \quad (2)$$

El proceso iterativo de $f_{\max}(T_l)$ se define por la ecuación 3:

$$f_{\max}(T_i) = \{f_j | \max_{j=1}^m n_o(f_j)\} \quad (3)$$

METODOLOGÍA

En esta sección, se describe el estudio realizado para la comparación experimental del desempeño de los métodos para el manejo del desbalance de clases en *Big Data*. En general, la Figura 4 describe la Metodología seguida.

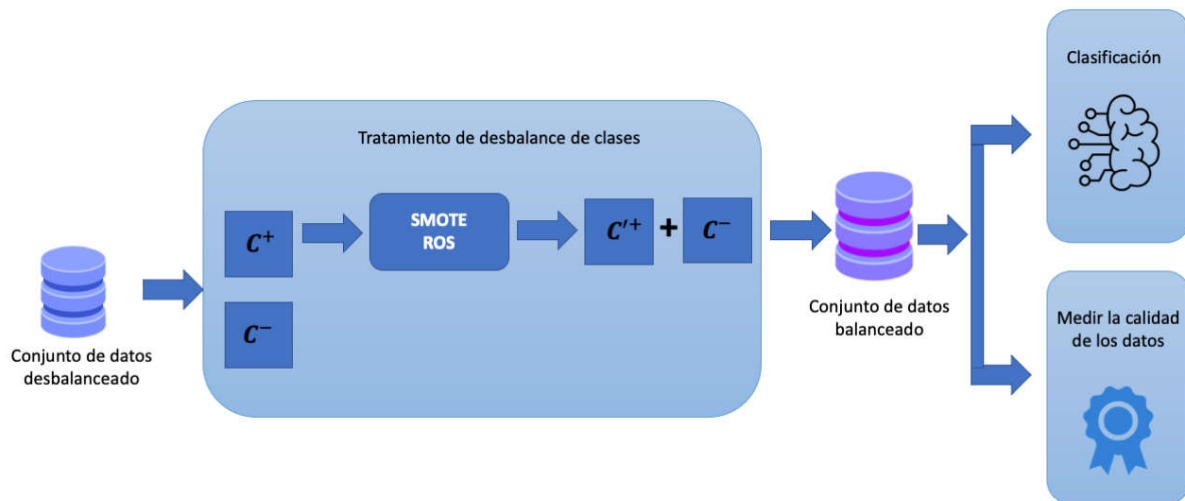


FIGURA 4.
METODOLOGÍA

En general, para grandes conjuntos de dos clases desbalanceados, se aplican las técnicas de sobre-muestreo. Una vez tratado el desbalance de clases, se evalúa el rendimiento del clasificador, además de evaluar la calidad de los conjuntos de datos generados.

CONJUNTO DE DATOS

Basamos la experimentación en 6 conjuntos de *Big Data* y 6 conjuntos de *Small Data*, todos ellos de dos clases con desbalance, tomados del repositorio UCI *Machine Learning* (Lichman, s. f.). Sus detalles se describen en la Tabla 1. Los conjuntos fueron seleccionados en función a su grado de desbalance (IR), es decir, la relación entre el número de instancias de la clase negativa y el número de instancias de la clase positiva. El grado de desbalance de estos conjuntos varía entre 2.57 y 40.

TABLA 1.
CONJUNTOS DE DATOS DESBALANCEADOS USADOS PARA BIG DATA Y SMALL DATA

	Conjunto	Total	Número de Atributos	Grado de desbalance (IR)
1	MiniBooNE	29178-74870	49	2.57
2	Susy	542434-2169739	18	4.00
3	poker	39062-410960	10	10.52
4	HEPMASS	262435-4200169	28	16.00
5	HIGGS	291417- 4663335	28	16.00
6	Covtype	16256 - 448421	54	27.58
7	vowel0	90 - 898	13	9.98
8	glass2	17 - 197	9	11.59
9	cleveland	13 - 160	13	12.31
10	glass4	13 - 201	9	15.47
11	yeast5	44 - 1440	8	32.73
12	abalone	58 - 2280	8	39.31

Se usa la validación cruzada en 5 particiones, donde el 80% de las instancias se dedicaron a entrenamiento y el 20% restante a pruebas. La Tabla 2 muestra el promedio de ejecutar los algoritmos.

TRATAMIENTO DE DESBALANCE

Con el fin de comparar el comportamiento de rendimiento de los métodos de sobre-muestreo, el estudio experimental se aplicarán los métodos de balanceo descritos en la Sección de Estrategias de sobre-muestreo, estos métodos son ROS, SMOTE-MR (*S-MR*), SMOTE-BD (*S-DB*). Sin embargo, la mayoría necesita algunos parámetros libres, por lo que las especificaciones técnicas del método de sobre-muestreo son: el porcentaje promedio de sobre-muestreo fue de 100, usamos particiones de 128, los k vecinos más cercanos para SMOTE fueron de 5, usando la distancia euclidiana.

MÉTRICAS DE EVALUACIÓN

Con el fin de describir el rendimiento del clasificador, se ejecutó el árbol de decisión de la *APIMLlib Spark*. Para un problema de dos clases, la matriz de confusión es usada para obtener las métricas de calidad. Así, las métricas que miden el rendimiento están dadas por la cantidad de instancias de cada clase que se clasificaron correctamente, nombradas como verdadero-positivo (TP) y verdadero-negativo (TN), mientras que la cantidad de casos positivos y negativos las instancias negativas mal clasificadas se denominan falso-positivo (FP) y *falso-negativo* (FN), respectivamente. Para determinar el rendimiento del algoritmo de aprendizaje en conjuntos de datos desbalanceados, comúnmente se emplea la media geométrica (Ecuación 4)

$$G_{mean} = \sqrt{(TP/(TP + FN)) \quad (TN / (TN + FP))} \quad (4)$$

Se evalúa la calidad de los datos mediante la determinación de redundancia (métrica F1) y la eficiencia de características (métrica F4).

INFRAESTRUCTURA

Manejar enormes cantidades de datos requiere procesamiento de alto rendimiento y potencia de almacenamiento. Todos los experimentos se han llevado a cabo en el servidor *falco* del grupo de trabajo VRain, con un microprocesador Intel(R) Core(TM) i9-10920X CPU @ 3.50GHz y 132 GB de RAM. En cuanto al software se utilizó Apache Spark 2.2.0.

RESULTADOS

El objetivo de las técnicas de sobre-muestreo es disminuir el sesgo de aprendizaje, mediante la creación de nuevas instancias. La Tabla 2 reporta la media geométrica promedio del esquema de validación cruzada de cinco veces cada conjunto de datos, por método de sobre-muestreo. Adicionalmente, se proporciona el rango promedio de *Friedman* para pequeños y

grandes conjuntos. Como base de comparación, se incluyen los resultados de los conjuntos de datos sin tratar el desbalance.

TABLA 2.
RESULTADOS OBTENIDOS CON LOS ALGORITMOS DE SOBRE-MUESTREO

Conjunto	Referencia	ROS	S-MR	S-DB
MiniBooNE	85.2	87.9	87.9	88.1
Susy	68.8	76.8	76.3	76.3
Poker	17.0	58.1	59.7	53.9
HEPMAS	71.9	83.3	66.8	65.5
HIGGS	11.6	66.0	64.4	64.9
Covtype	72.8	93.2	92.6	93.0
Promedio de rank	3.7	1.4	2.5	2.4
Vowel0	92.3	95.3	96.2	97
Glass2	28.9	51.5	77.1	70.6
Cleveland	41.1	51.4	75.6	75
Glass4	83.1	85.4	88.7	88.7
Yeast5	82.1	92.1	89.3	91.8
Abalone	22	80.3	79.4	80.4
Promedio de rank	4	2.5	1.9	1.6

Los resultados de la Tabla 2 señalan que para conjuntos de *Big Data*, ROS tiene un mejor rendimiento al tener el mejor rango promedio de Friedman en comparación con las propuestas basadas en SMOTE. Este comportamiento es diferente en conjuntos de datos pequeños, como se observa, el rango promedio obtenido por SMOTE-DB es el mejor.

Centrándose en el conjunto poker, el bajo rendimiento en SMOTE sugiere que es causado por el procedimiento de interpolación, dado que SMOTE crea instancias sintéticas enfocadas en el espacio de características, y este conjunto de datos tienen características

discretas, las nuevas instancias no brindan suficiente conocimiento sobre los grandes conjuntos de datos.

La medida F1 determinar si las clases están superpuestas, en la Figura 5 se observa, que los valores más bajos se presentan en los conjuntos de referencia, lo que incurre en traslape de clases. De manera específica, para conjuntos de datos pequeños con técnicas de SMOTE, se observa que los valores obtenidos son mayores. En contraste con los conjuntos de *Big Data*, en todos los métodos de sobre-muestreo la medida F1 se mantiene bien ajustada y el valor resultante es bajo, lo que sugiere que la complejidad incurre en la existencia de redundancia, aunque se generen nuevas instancias por interpolación.

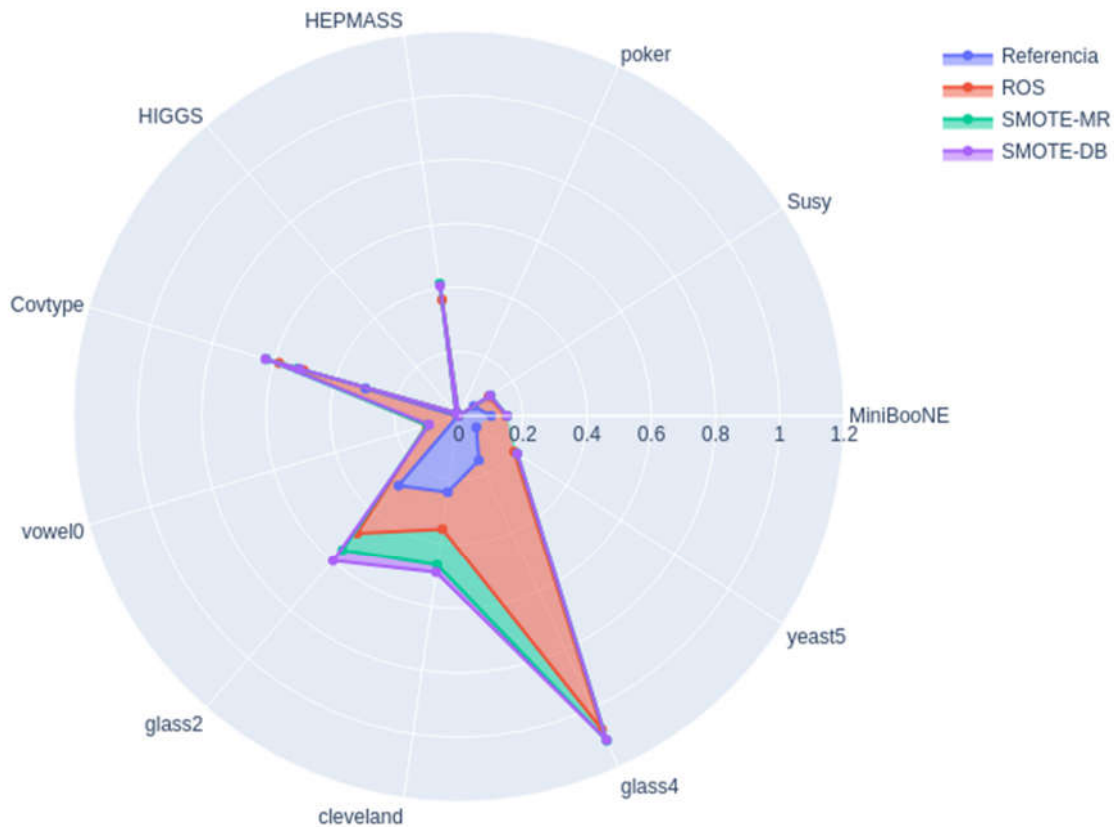


FIGURA 5.
RELACIÓN MÁXIMA DE DISCRIMINACIÓN (MÉTRICA F1)

De hecho, estos resultados evidencian que el uso de la distancia euclidiana puede no ser una norma adecuada para los conjuntos de datos utilizados, debido a la interpolación. El término de proximidad se define como todas las instancias son aproximadamente equidistantes entre sí, y la distancia euclidiana asume que todos los atributos son igualmente importantes (Maldonado et al., 2019), en este sentido, la similitud del coseno puede conducir a un mejor resultado debido a algunas propiedades de las instancias que hacen que los pesos sean mayores sin diferencia.

Podemos destacar que para conjuntos de *Big Data*, un método de sobre-muestreo basado en el espacio de características no conduce a mejores resultados, de hecho, es necesario considerar el espacio de datos en general. Dado que manejamos una gran cantidad de datos, seguimos enfrentando problemas como el traslape de clases, donde es evidente que la frontera de solución no está clara dada la redundancia de instancias que puede tener *Big Data*.

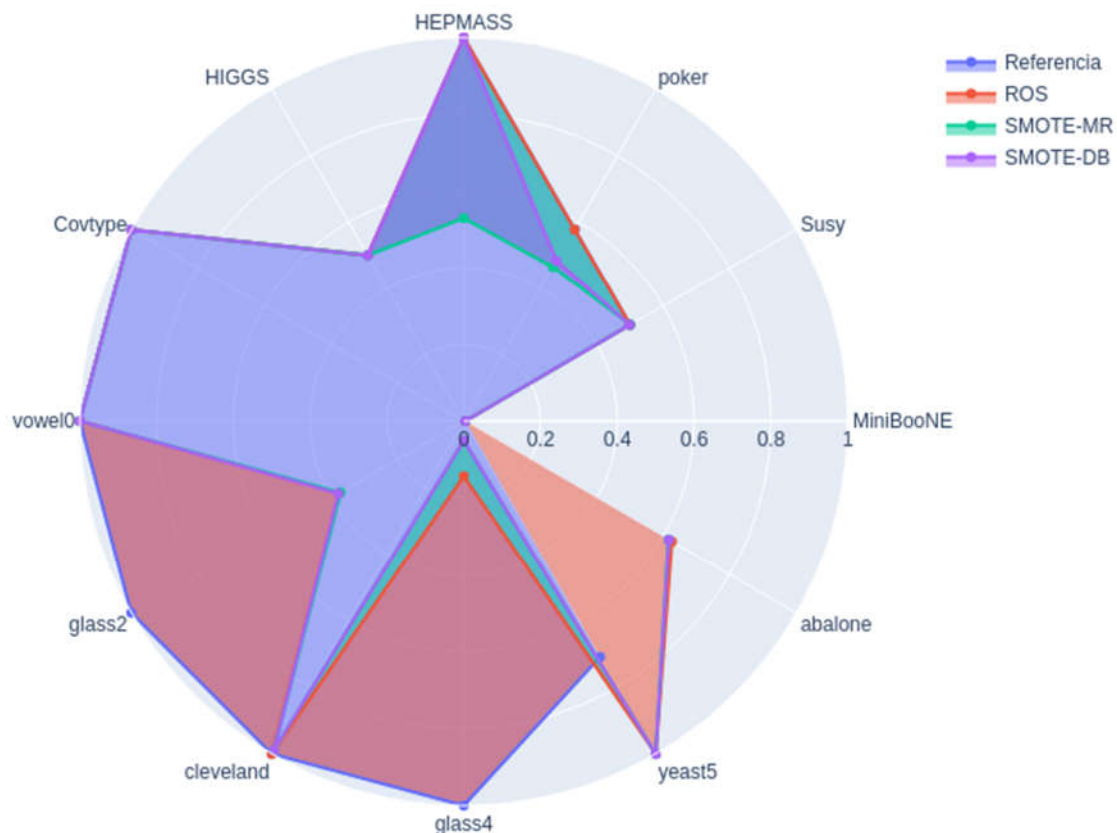


FIGURA 6.

EFICIENCIA DE CARACTERÍSTICA COLECTIVA (MÉTRICA F4)

La medida F4 obtiene una descripción general de cómo funcionan juntas las características. De este modo, los valores más bajos de F4 indican que es posible discriminar más instancias y, por lo tanto, que el problema es más simple. De la Figura 6, los valores obtenidos por el conjunto base, podemos deducir que los conjuntos de datos en ambos escenarios (*Big Data* y *Small Data*) no es posible discriminar instancias, debido a que todas ellas obtuvieron los valores más altos en la medida F4. Para *glass4* las técnicas de SMOTE reducen la complejidad, este comportamiento sugiere que la generación de instancias sintéticas mantiene características discriminatorias de las originales, no así para el método aleatorio.

Los resultados obtenidos en ambos escenarios sugieren que, en contexto de *Big Data*, las características intrínsecas de los datos se conservan. En la mayoría de los conjuntos de datos aplicando SMOTE sufren de traslape de clases, debido a la distribución de las instancias sintéticas. De hecho, esto expondría que el espacio de muestra en los conjuntos de *Big Data* es aparentemente pequeño en comparación con la cantidad de instancias en cada conjunto de datos.

CONCLUSIONES

El desbalance de clases sigue siendo uno de los problemas más comunes y relevantes, no sólo en conjuntos de datos "pequeños" sino también en conjuntos de *Big Data*. Cabe señalar que en este último caso se agrava especialmente ya que se trata de una gran cantidad de datos generado con frecuencia. Realizamos una serie de experimentos para comprobar el comportamiento de métodos de sobre-muestreo en contextos de *Big Data* y para conjuntos de datos pequeños, como ROS y algunas variantes de SMOTE.

Para examinar el comportamiento del clasificador, se aplicó un árbol de decisión, en un total de 12 conjuntos de datos de dos clases desbalanceados. Los resultados han demostrado que los métodos basados en SMOTE obtienen un rendimiento más bajo que ROS en escenarios de *Big Data*, mientras que en conjuntos de datos pequeños, SMOTE presenta un mejor comportamiento.

El rendimiento aparente en *Big Data* de los métodos de sobre-muestreo basados en SMOTE proviene del hecho de que esta técnica genera instancias sintéticas de acuerdo con el vecindario y la cobertura probable del espacio muestral. Los resultados sugiere que usar únicamente SMOTE para escenarios de *Big Data* no es una alternativa viable, ya que los resultados obtenidos no abordaron una mejora en el desempeño, a pesar de utilizar un método de agrupamiento para abordar la creación de datos sintéticos. Adicionalmente, se ha corroborado que el desbalance de clases no es un problema en sí mismo en los conjuntos de *Big Data*, por lo que es necesario lidiar con otras complejidades, debido a la presencia de redundancia, ruido y traslape de clases. Esta deducción se ha guiado por las métricas de complejidad F1 y F4.

Como trabajo futuro, la línea abierta se centra en considerar las regiones dispersas o las regiones densas para decidir la creación de nuevas instancias en SMOTE. Otra línea abierta involucra la reducción de características previas al uso de SMOTE, así como implementar otras métricas de distancia, como la similitud del coseno. Finalmente, la presencia de otras complejidades de datos en *Big Data* debe gestionarse antes de manejar el desbalance de clases, de este modo, buscar la generación de propuestas híbridas.

AGRADECIMIENTOS

Angélica Guzmán-Ponce contó con el apoyo del contrato postdoctoral Margarita Salas MGS/2021/23(UP2021-021) financiado por la Unión Europea-NextGenerationEU.

REFERENCIAS

- Basgall, M. J., Hasperué, W., Naiouf, M., Fernández, A. & Herrera, F. (2018). SMOTE-BD: An Exact and Scalable Oversampling Method for Imbalanced Classification in Big Data. *Journal of Computer Science and Technology*, 18(03), 23–28. <https://doi.org/10.24215/16666038.18.e23>
- Basgall, M. J., Hasperué, W., Naiouf, M., Fernández, A. & Herrera, F. (2019). An Analysis of Local and Global Solutions to Address Big Data Imbalanced Classification: A Case Study with SMOTE Preprocessing. *Communications in Computer and Information Science*, 75-85. https://doi.org/10.1007/978-3-030-27713-0_7
- Batista, G. E. A. P. A., Prati, R. C. & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1), 20-29. <https://doi.org/10.1145/1007730.1007735>
- del Río, S., López, V., Benítez, J. M. & Herrera, F. (2014). On the use of MapReduce for imbalanced big data using Random Forest. *Information Sciences*, 285, 112-137. <https://doi.org/10.1016/j.ins.2014.03.043>
- García, S., Galar, M., Prati, R. C., Krawczyk, B. & Herrera, F. (2018). *Learning from Imbalanced Data Sets*. Springer. <https://doi.org/10.1007/978-3-319-98074-4>
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M. & Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(9), 1–22. <https://doi.org/10.1186/s41044-016-0014-0>
- García, V., Sánchez, J., Marqués, A., Florencia, R. & Rivera, G. (2020). Understanding the apparent superiority of over-sampling through an analysis of local information for class-imbalanced data. *Expert Systems with Applications*, 158, 113026. <https://doi.org/10.1016/j.eswa.2019.113026>
- Gutiérrez, P. D., Lastra, M., Benítez, J. M. & Herrera, F. (2017). SMOTE-GPU: Big Data preprocessing on commodity hardware for imbalanced classification. *Progress in Artificial*

- Intelligence, 6(4), 347-354. <https://doi.org/10.1007/s13748-017-0128-2>
- Kitchin, R. & Lauriault, T. P. (2015). Small data in the era of big data. *GeoJournal*, 80(4), 463-475. <https://doi.org/10.1007/s10708-014-9601-7>
- Lichman, M. (s. f.). UCI Machine Learning Repository. Recuperado 5 de octubre de 2022, de <https://archive.ics.uci.edu/ml/index.php>
- Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S. & Herrera, F. (2020). *Big Data Preprocessing*. Cham, Switzerland: Springer International Publishing. <https://doi.org/10.1007/978-3-030-39105-8>
- Maillo, J., Ramírez, S., Triguero, I. & Herrera, F. (2017). kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. *Knowledge-Based Systems*, 117, 3-15. <https://doi.org/10.1016/j.knosys.2016.06.012>
- Maillo, J., Triguero, I. & Herrera, F. (2020). Redundancy and Complexity Metrics for Big Data Classification: Towards Smart Data. *IEEE Access*, 8, 87918-87928. <https://doi.org/10.1109/access.2020.2991800>
- Maldonado, S., López, J. & Vairetti, C. (2019). An alternative SMOTE oversampling strategy for high-dimensional datasets. *Applied Soft Computing*, 76, 380-389. <https://doi.org/10.1016/j.asoc.2018.12.024>
- Rendón, E., Alejo, R., Castorena, C., Isidro-Ortega, F. J. & Granda-Gutiérrez, E. E. (2020). Data Sampling Methods to Deal With the Big Data Multi-Class Imbalance Problem. *Applied Sciences*, 10(4), 1276. <https://doi.org/10.3390/app10041276>



Ideas en Ciencias de la Ingeniería es una revista de carácter académico y científico de la Facultad de Ingeniería de la Universidad Autónoma del Estado de México. Su publicación digital es semestral. El objetivo de la revista es comunicar temas en el ámbito de la Ciencia y el Desarrollo Tecnológico para un público especializado en Ingeniería.

Se **invita a publicar** artículos que pertenezcan a la siguiente cobertura temática: Ingeniería y tecnología civil en arquitectura, mecánica, electrónica, en computación y Sistemas energéticos sustentables.

La revista *Ideas en Ciencias de la Ingeniería* tiene el compromiso de publicar artículos con calidad en los idiomas español y/o inglés. Los artículos que se publiquen deberán ser originales e inéditos. Éstos se someterán a revisión bajo la modalidad de pares ciegos, por lo que deben ser originales e inéditos y no deben ser postulados para ningún otro tipo de publicación.

Instrucciones para autores e información general:

<https://ideasencienciasingenieria.uaemex.mx/>



Redes sociales

<https://www.facebook.com/profile.php?id=100086458095640>

