

Cartografía de la vulnerabilidad a la contaminación de las aguas subterráneas empleando aprendizaje automático: cuenca hidrográfica Mayarí, Cuba

Mapping the vulnerability to groundwater contamination using machine learning: Mayarí watershed, Cuba

Camila Cruz González^{1*}, Rosa María Valcarce Ortega¹, Marina Beatriz Vega Carreño¹, Willy Roberto Rodríguez Miranda¹

¹Universidad Tecnológica de La Habana José Antonio Echeverría Calle 114 No 11901 e/ Rotonda y Ciclo Vía, Marianao, La Habana, Cuba.

*camilacrugon@civil.cujae.edu.cu, cruzgonzalezcamila5@gmail.com

Resumen

La cuenca hidrográfica Mayarí, ubicada en la región oriental de Cuba, ha sido declarada de interés nacional por el Consejo Nacional de Cuencas Hidrográficas por lo que estudios dirigidos a preservar sus recursos naturales tienen especial interés. El objetivo de este artículo es evaluar la vulnerabilidad intrínseca a la contaminación de las aguas subterráneas en la cuenca hidrográfica Mayarí por el impacto de la migración vertical de potenciales contaminantes vertidos en su superficie. Se empleó el algoritmo de aprendizaje automático k-medias sobre las variables pendiente topográfica, índice de atenuación del suelo, densidad de fallas tectónicas y formaciones geológicas. Se analizó la variación espacio-temporal de la vegetación y la salinidad de los suelos en la cuenca según el índice normalizado de vegetación (NDVI) y el índice de salinidad de suelos (NDSI) empleando imágenes satelitales Landsat de los años 2000 y 2023. Se obtuvo un mapa de vulnerabilidad intrínseca de las aguas subterráneas a la contaminación por infiltración de potenciales contaminantes superficiales a escala 1:100 000 que discrimina zonas de baja, moderada, alta y extrema vulnerabilidad. En el período analizado el índice de vegetación aumentó y la salinidad de los suelos disminuyó, lo que impacta favorablemente en la protección de los recursos hídricos. El procedimiento presentado forma parte de un sistema de alerta temprana para prevenir el posible deterioro de la calidad de las aguas subterráneas en la cuenca Mayarí. Puede ser aplicado a otras cuencas del territorio nacional y a escalas de trabajo más detalladas.

Palabras clave: vulnerabilidad de acuíferos, aprendizaje automático, cuenca Mayarí.

Abstract

The Mayarí watershed, located in the eastern region of Cuba, has been declared of national interest by the National Council for Watershed Basins; therefore, studies aimed at preserving its natural resources are of special importance. The objective of this article is to evaluate the intrinsic vulnerability of groundwater to contamination in the Mayarí watershed due to the impact of the vertical migration of potential contaminants discharged on its surface. The k-means machine learning algorithm was employed using the variables of topographic slope, soil attenuation index, density of tectonic faults, and geological formations. The spatiotemporal variation of vegetation and soil salinity in the basin was analyzed according to the Normalized Difference Vegetation Index (NDVI) and the Soil Salinity Index (NDSI), using Landsat satellite images from the years 2000 and 2023. An intrinsic vulnerability map of groundwater to contamination from the infiltration of potential surface contaminants was obtained at a scale of 1:100,000, which

distinguishes areas of low, moderate, high, and extreme vulnerability. The vegetation index increased while soil salinity decreased during the analyzed period, which positively impacts the protection of water resources. The presented procedure is part of an early warning system designed to prevent the potential deterioration of groundwater quality in the Mayarí watershed. It can be applied to other basins across the national territory and at more detailed work scales.

Keywords: aquifer vulnerability, machine learning, Mayarí watershed.

1. Introducción

Las cuencas hidrográficas son unidades territoriales que vinculan espacial y funcionalmente ecosistemas terrestres y acuáticos con múltiples actividades sociales, económicas y culturales, que se desarrollan sobre ellas. Desempeñan un rol importante en la estabilidad de los ecosistemas y requieren adecuada planificación para lograr la explotación sostenible de los valiosos recursos naturales que poseen [1]. Sin embargo, en el empeño de lograr una gestión integral sostenible de las cuencas hidrográficas, la protección a la contaminación de sus aguas subterráneas es un elemento esencial al que no siempre se brinda toda la atención necesaria.

En Cuba representa una prioridad la protección de los recursos hídricos y su gestión responsable ha experimentado una evolución significativa a lo largo del tiempo. Después del triunfo de la Revolución, el comandante en jefe Fidel Castro Ruz promovió políticas destinadas a la protección y explotación sostenible de las aguas, creando el programa "Voluntad Hidráulica" en 1963. Hoy Cuba cuenta con la Ley No. 124 de las Aguas Terrestres, que regula la gestión integrada y sostenible de este recurso para lograr su uso y preservación en armonía con el desarrollo socioeconómico, a la vez que promueve la adecuada aplicación de la ciencia y la tecnología con este propósito en total correspondencia con los Objetivos de Desarrollo Sostenible.

Los mapas de vulnerabilidad a la contaminación de los acuíferos son herramientas útiles que forman parte de un sistema de alerta temprana para prevenir el deterioro de la calidad de las aguas subterráneas y que ayudan a la toma de decisiones informadas para establecer sistemas de monitoreo de la calidad del agua y un ordenamiento territorial que favorezca la protección de este recurso [2]. El desarrollo de estos mapas ha cobrado una importancia notable en Cuba en los últimos años, como lo demuestran las investigaciones llevadas a cabo por diferentes autores [4-7].

La cuenca hidrográfica Mayarí abarca el municipio Mayarí perteneciente a la provincia de Holguín y tres municipios de la provincia de Santiago de Cuba (Figura 1). Se localiza entre las regiones Alturas de Mayarí y Montañas Nipe-Cristal. Limita al norte por la Bahía de Nipe y el Océano Atlántico, al sur por las Alturas de Mayarí, al este por las vertientes NO y SO de la Sierra de Mícara-Cristal y por el oeste por las vertientes NE y SE de la Sierra de Nipe.

La longitud del cauce principal del río Mayarí es de 107 Km, orientado de sur a norte, con un ancho medio no mayor de 20 m [8]. Está formada por 19 afluentes, de los cuales 14 corresponden al municipio Mayarí.

Las rocas que forman la porción meridional de la cuenca son areniscas, calcarenitas, tobas, margas, pedernales, basaltos; y en la porción septentrional la composición litológica es terrígeno-carbonatado con predominio de calizas, margas, arcillas, conglomerados y depósitos carbonatados-terrígenos. En los macizos de Mayarí se describen ofiolitas separadas por una falla subhorizontal y afloran rocas volcánico-sedimentarias del Arco Volcánico del Cretácico, cubiertas tansgresivamente por depósitos flyschoides y olistostrómicos de edad Maastrichtiano a Paleoceno [9].

Sus características hidrogeológicas la definen los complejos acuíferos, que son unidades litológico-faciales enmarcadas en un rango de edad cuyos límites pueden ser tectónicos, de contactos litológicos o de cambios faciales. En la cuenca hidrográfica Mayarí los sistemas acuíferos predominantes son los vinculados al ultrabásico metamorfozado e intrusivos básicos del Cretácico-Jurásico representado por la Asociación Ofiolítica, así como el complejo clástico-carbonatado del

Paleógeno-Cretácico representado por las formaciones geológicas Mucaral, San Luis, Mícaro y La Picota (Figura 2).

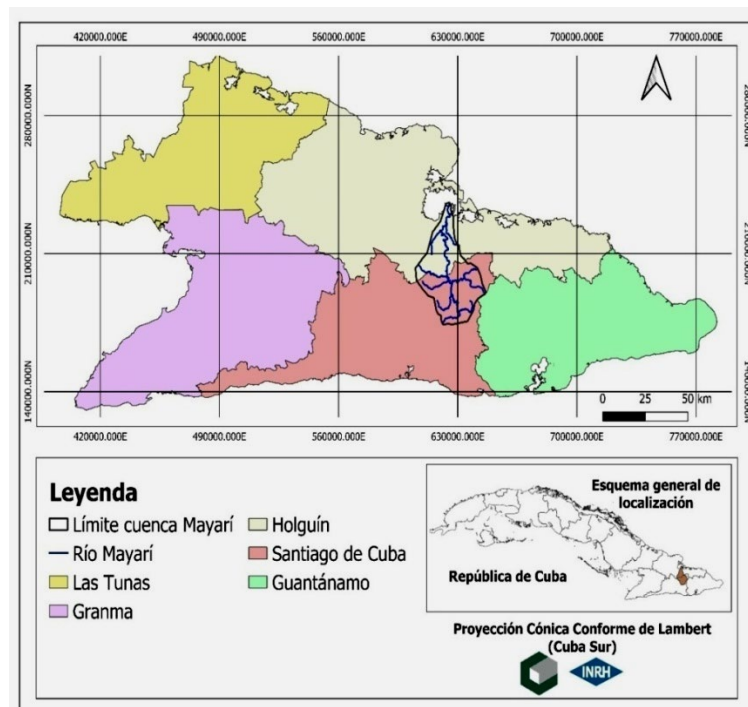


Fig.1 Ubicación geográfica de la cuenca hidrográfica Mayarí

El potencial hídrico subterráneo de la cuenca alcanza 10,7 hm³/año. Entre las principales amenazas a la calidad de las aguas subterráneas se encuentran la tala de bosques, las actividades mineras a cielo abierto, la erosión de los suelos, el deficiente tratamiento de residuales líquidos provenientes de actividades domésticas, industriales y agropecuarias [10].

El objetivo de esta investigación fue evaluar la vulnerabilidad intrínseca a la contaminación de las aguas subterráneas en la cuenca hidrográfica Mayarí por el impacto de la migración vertical de potenciales contaminantes vertidos en su superficie.

2. Materiales y Métodos

2.1. Materiales disponibles

Para evaluar la vulnerabilidad natural a la contaminación de las aguas subterráneas en la cuenca hidrográfica Mayarí se dispuso de:

- Mapa Geológico de la República de Cuba a escala 1:100 000
- Mapa Tectónico de la República de Cuba a escala 1:100 000
- Mapa de Suelos de Cuba a escala 1:100 000
- Mapa de Suelos de Cuba a escala 1:25 000
- Modelo Digital de Elevaciones (MDE) 25x25 m

2.2. Métodos empleados

El término vulnerabilidad del agua subterránea a la contaminación fue introducido por el hidrogeólogo francés J. Margat a finales de la década del 60 del siglo pasado al fundamentar que, en cierta medida, el medio físico protege al acuífero de contaminantes que pueden infiltrarse desde la superficie [11].

En la literatura se identifican dos tipos de vulnerabilidad a la contaminación de las aguas subterráneas: la vulnerabilidad intrínseca o natural, que es aquella que se evalúa sin tener en cuenta la naturaleza de los contaminantes; y la vulnerabilidad específica, referida a la susceptibilidad del agua subterránea a un contaminante o grupo de contaminantes, en función de sus características y

relaciones con los componentes de la vulnerabilidad intrínseca [12]. En esta investigación se estudió la vulnerabilidad intrínseca o natural.

Desde la década de los años 80 del pasado siglo, han sido desarrollados diferentes métodos para evaluar la vulnerabilidad intrínseca a la contaminación de las aguas subterráneas, siendo los métodos de superposición de índices ponderados los más empleados. Ejemplo de ellos son los métodos DRASTIC, GOD, SINTACS, EPIK, RISK, entre otros [13,14]. En los últimos años se han desarrollado aplicaciones de las técnicas de aprendizaje automático que están demostrando alto poder resolutivo para generar mapas de vulnerabilidad a la contaminación de los acuíferos [15, 16]. Precisamente, en esta investigación se empleó el algoritmo k-medias.

2.2.1 Algoritmo k-medias

El algoritmo K-medias forma parte del conjunto de herramientas de aprendizaje automático que clasifican como técnicas de inteligencia artificial. Es un método de reconocimiento de patrones no supervisado que permite clasificar en grupos una gran cantidad de datos, de tal manera que los elementos de cada grupo sean muy similares entre sí y a la vez sean muy diferentes a los elementos de los otros grupos. Este algoritmo permite asignar cada objeto al grupo más cercano, lo que se logra calculando una medida de similitud entre el objeto y el centroide de cada clúster. En esta investigación se utilizó como medida de similitud el coeficiente de distancia euclidiano, recomendado cuando las variables empleadas se expresan de forma cuantitativa y existe baja correlación estadística entre ellas [17].

El número de grupos y sus centroides se calculan inicialmente de forma aleatoria y después de la primera asignación de objetos a cada grupo se recalculan los centroides como la media de las variables de los puntos que le fueron asignados. Una vez actualizado el centroide de cada clúster se vuelven a reasignar los objetos al grupo más cercano. Este procedimiento se repite hasta lograr la convergencia, o sea, hasta que las asignaciones de los puntos no cambien o hasta alcanzar el número de iteraciones prefijado. Este resultado final representa el ajuste que maximiza la distancia entre los distintos grupos y minimiza la distancia intragrupo. La principal ventaja del método es su sencillez y rapidez, pero es un algoritmo significativamente sensible a los centroides que se seleccionan inicialmente de manera aleatoria. Este efecto se puede reducir incrementando el número de iteraciones del procedimiento. El algoritmo es más eficiente en la medida que las variables empleadas no sean redundantes, y es muy sensible al hecho de que las variables posean diferente rango de variación, por lo que se recomienda estandarizarlas entre 0 y 1 sustrayendo a cada variable su valor mínimo y dividiendo por su rango. Es importante destacar que el algoritmo siempre separará los objetos en grupos. El conocimiento del investigador hará posible identificar qué grupos son significativos y cuáles no, así como el significado de cada uno [7].

Las variables a emplear deben ser seleccionadas en dependencia de la tarea técnica a resolver y de los datos disponibles, velando porque no exista redundancia en la información, o sea, seleccionando variables que sean independientes entre sí. En esta investigación las variables empleadas fueron: pendiente topográfica del terreno, litología de la zona no saturada, densidad de fallas tectónicas por km² e índice de atenuación del suelo a sustancias contaminantes.

A continuación, se fundamenta la selección de las variables y se explica cómo fueron obtenidas.

Pendiente topográfica: Las áreas con pendientes más pronunciadas experimentan mayor escorrentía superficial y menor infiltración vertical de los contaminantes, lo que disminuye la vulnerabilidad a la contaminación de los acuíferos. Fue extraída del MDE con resolución espacial de 25 metros, empleando la función Slope en la herramienta *Spatial Analyst* del sistema de información geográfico QGIS.

Litología de la zona no saturada: La litología, el grado de consolidación, la posible presencia de fracturas, el desarrollo kárstico, la arcillosidad; son factores que determinan la permeabilidad de la zona no saturada favoreciendo o retardando la migración vertical de contaminantes. Se asignaron valores entre 0,4 y 1 a las diferentes litologías, los mayores valores se asignaron a las formaciones geológicas potencialmente más permeables, de la misma manera que se realiza al aplicar el método paramétrico GOD [18]. En la cuenca Mayarí, atendiendo a la composición litológica predominante

de las formaciones geológicas presentes se asignó puntuación 0,6 a las formaciones Dátil y San Luis, y al resto se asignó puntuación 1.

Densidad de fallas tectónicas por km²: Las fallas tectónicas generan fracturas en la corteza terrestre que pueden incrementar la permeabilidad secundaria de las rocas, facilitando la infiltración y migración de los contaminantes hasta alcanzar el nivel freático. La densidad de fallas por km² fue extraída del mapa tectónico empleando la herramienta *Kernel density estimation* del software QGIS.

Índice de atenuación del suelo: La capacidad del suelo para retener y degradar sustancias contaminantes depende de sus propiedades físicas y químicas, como textura, profundidad efectiva y humificación [2]. La textura influye porque si el contenido de arcilla disminuye, su capacidad de retención y degradación también. A mayor espesor del suelo mayor capacidad de retención de contaminantes. A mayor humificación mayor contenido de materia orgánica y mayor será su función protectora. La tabla 1 presenta la evaluación de cada parámetro disponible en el mapa de suelos.

Tabla 1. Puntuación de los parámetros del suelo. Tomado del Mapa de Suelos de Cuba a escala 1:25 000.

Textura	Puntuación	Profundidad efectiva	Puntuación	Humificación	Puntuación
Arcilla montmorillonita	1	Muy profundo (mayor de 100 cm)	1	Muy humificado (mayor del 6%)	1
Arcilla caolinita	2	Profundo (entre 61 y 99 cm)	2	Humificado (entre 4.1 y 6%)	2
Arcilla	3	Medianamente profundo (entre 41 y 60 cm)	3	Medianamente humificado (entre 2 y 4%)	3
Arcilla loamosa	4	Poco profundo (entre 21 y 40 cm)	4	Poco humificado (menor del 2%)	4
Arcilla arenosa	5	Muy poco profundo (entre 11 y 20 cm)	5		
Loam arcilloso	6				
Loam arcilloso arenoso	7				
Loam arenoso	8				
Arena arcillosa	9				

Para calcular el índice de atenuación del suelo esta investigación propuso la ecuación (1).

$$A = \frac{1}{Tx + PE + H} \quad (1)$$

Donde:

IA: Índice de Atenuación

Tx: Textura

PE: Profundidad Efectiva

H: Humificación

Para las variables descritas anteriormente fue calculada la matriz de correlación lineal, comprobándose que las variables son independientes entre sí y no aportan información redundante, condición necesaria para obtener los mejores resultados con el algoritmo k-medias.

Fueron obtenidos modelos de clasificación con dos, tres, cuatro, cinco y seis grupos. Para seleccionar el número óptimo de grupos se aplicó el Método del Codo, donde el número óptimo de conglomerados se define cuando incluir más clústeres no produce disminución significativa en la suma del error cuadrático dentro de los grupos [19]. Según este criterio el modelo de agrupación con cuatro clústeres resultó el óptimo. Para asociar una categoría de vulnerabilidad a cada grupo se tuvo en cuenta que los grupos serán menos vulnerables en la medida que se caractericen por presentar litologías menos permeables en la zona no saturada, menor densidad de fallas tectónicas, mayor pendiente topográfica y mayor índice de atenuación del suelo.

2.2.2 Evaluación de índices espectrales NDVI y NDSI

Con el objetivo de analizar si han existido cambios en la cobertura vegetal y la salinidad del suelo que puedan incrementar la susceptibilidad a la contaminación de las aguas subterráneas en la cuenca, se realizó el cálculo del índice de vegetación de diferencia normalizada (NDVI) y del índice de salinidad de diferencia normalizada (NDSI). Para ello, se contrastaron las imágenes satelitales Landsat 7 y 8 correspondientes a los años 2000 y 2023, en los meses de mayo y julio respectivamente.

El Índice de Vegetación de Diferencia Normalizada (NDVI por sus siglas en inglés) es una medida de vegetación sana. Evalúa el nivel de verdor y densidad de la vegetación captada en una imagen satelital. La vegetación saludable presenta un patrón de reflectancia espectral distintivo, que se puede determinar mediante la diferencia entre dos bandas: la del rojo visible y la del infrarrojo cercano. Este índice varía entre -1 y 1. Un descenso repentino de los valores de NDVI puede ser un síntoma de deterioro de la salud vegetal [20].

El Índice de Salinidad de Diferencia Normalizada (NDSI por sus siglas en inglés) es un índice utilizado para detectar y cuantificar la salinidad en cuerpos de agua, suelos y terrenos a partir de imágenes satelitales. Se basa en la diferencia entre la reflectancia en bandas espectrales sensibles a la presencia de sales disueltas en el agua o en el suelo. Los valores del NDSI pueden variar entre -1 y 1; valores altos indican mayor salinidad en el área analizada [21].

A partir de las imágenes Landsat y mediante el uso del sistema informático QGIS, fueron calculados los índices espectrales NDVI y NDSI, empleando las ecuaciones (2) y (3).

$$NDVI = \frac{\text{Infra rojo cercano} - \text{Rojo}}{\text{Infra rojo cercano} + \text{Rojo}} \quad (2)$$

$$NDSI = \frac{\text{Verde} - \text{Infra rojo de onda corta}}{\text{Verde} + \text{Infra rojo de onda corta}} \quad (3)$$

Los resultados fueron representados en mapas ráster a escala 1: 100 000 empleando el sistema QGIS.

3. Resultados y Discusión

3.1. Análisis de la vulnerabilidad a la contaminación de las aguas subterráneas en la cuenca hidrográfica Mayarí empleando el algoritmo k-medias

La tabla 2 muestra el modelo de clasificación con 4 grupos, resultado de aplicar el algoritmo k-medias en la cuenca hidrográfica Mayarí. El análisis de las características de los centroides de cada

grupo permitió asociar un nivel de vulnerabilidad a cada clúster. Se aprecia que el índice de atenuación del suelo presenta un comportamiento similar en toda la cuenca.

Los grupos 1 y 2 clasifican de alta vulnerabilidad debido a que presentan predominio de litología carbonatada con muy elevado desarrollo cárstico y moderada pendiente topográfica, pero el grupo 2 se interpreta de extrema vulnerabilidad al estar caracterizado por mayores valores de densidad de fallas tectónicas y presentar en la zona no saturada litologías aún más permeables.

Tabla 2. Resultados del modelo de clasificación con 4 grupos al aplicar el algoritmo k-medias en la cuenca hidrográfica Mayarí

	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Pendiente topográfica (%)	13,0	12,0	6,0	25,0
Índice de atenuación (multiplicado por 10 ²)	9,5	9,2	10,0	9,7
Densidad de fallas tectónicas/km ²	0,1	0,5	0,2	0,2
Litología	0,7	0,9	0,5	0,6
Clasificación de la vulnerabilidad	Alta	Extrema	Moderada	Baja

El grupo 3 presenta los menores valores de pendiente topográfica, pero litologías menos permeables y pequeña densidad de fallas tectónicas. Todo esto permite considerar que presenta vulnerabilidad moderada respecto al sistema en su conjunto. El grupo 4 contiene la pendiente topográfica más alta, pequeña densidad de fracturamiento y litología arcillosa, por lo que se clasifica de baja vulnerabilidad intrínseca a la contaminación de las aguas subterráneas.

El análisis espacial de estos resultados permitió conocer que el complejo volcánico-sedimentario ubicado al este de la cuenca, donde predomina la Formación Santo Domingo; y el complejo ofiolítico, asociado a los sistemas acuíferos ultrabásico metamorfozado e intrusivos básicos ubicados en el centro de la cuenca, presentan baja y moderada vulnerabilidad. Esta categoría se caracteriza por presentar bajos valores de densidad de fallas tectónicas, la mayor pendiente topográfica y la litología menos permeable en la zona no saturada.

La alta vulnerabilidad está asociada al complejo volcánico-sedimentario representado por la Formación Sabaneta. En esta zona se describen formaciones geológicas con predominio de litología carbonatada con desarrollo cárstico, pendiente topográfica moderada y baja densidad de fallas tectónicas. La vulnerabilidad extrema ubicada al norte, centro y sur de la cuenca, está asociada a los complejos acuíferos carbonatado-clástico del Neógeno-Oligoceno, clástico-carbonatado y carbonatado-clástico del Paleógeno-Cretácico, donde es elevada la presencia de litologías carbonatadas cársticas en las Formaciones Bitirí, Camazán, Mucará, Charco Redondo y Puerto Boniato. En estas zonas la pendiente topográfica es moderada y está presente la mayor densidad de fallas tectónicas.

La figura 2 muestra los complejos acuíferos presentes en la cuenca y el mapa de vulnerabilidad natural a la contaminación de las aguas subterráneas.

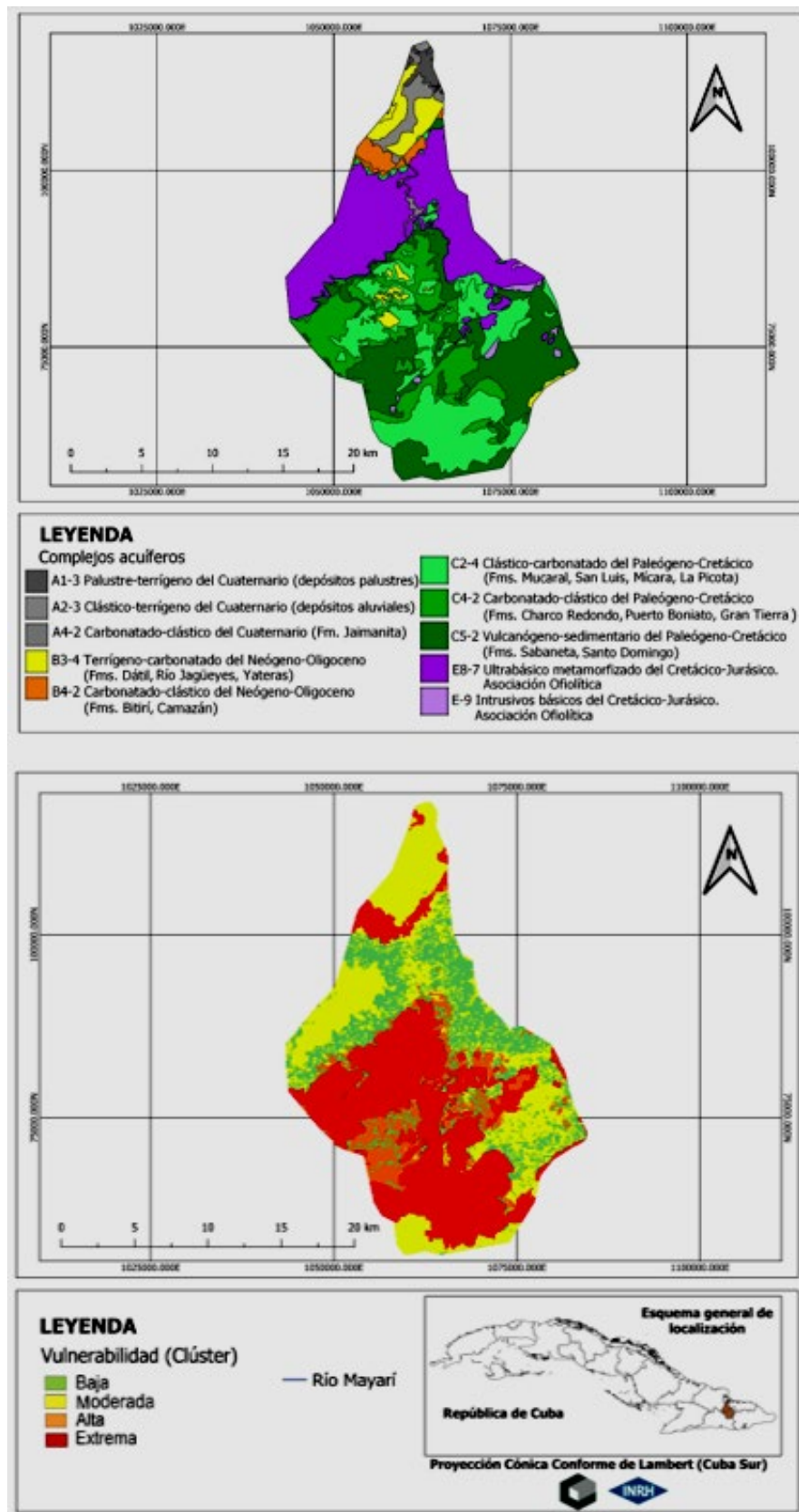


Fig. 2 Complejos acuíferos (arriba) y mapa de vulnerabilidad intrínseca a la contaminación de las aguas subterráneas de la cuenca hidrográfica Mayarí (abajo)

3.2 Variación espacio – temporal de la cobertura vegetal y la salinidad de los suelos en la cuenca hidrográfica Mayarí

La figura 3 muestra el resultado del cálculo del índice NDVI para la cuenca hidrográfica Mayarí. El color verde indica presencia de vegetación y el color rojo se asocia a áreas que no presentan vegetación. Se aprecia que existió incremento de la vegetación en el período comprendido entre los años 2000 y 2023, aunque en el extremo oeste se perciben zonas con ausencia de vegetación o con vegetación muy escasa, resultado del impacto de la actividad minera pero que actualmente se ha logrado revertir, como destacan otras investigaciones [10].

La misma figura muestra el resultado del cálculo del índice NDSI para las imágenes satelitales Landsat de mayo del año 2000 y julio del año 2023. El análisis visual permite afirmar que existió una discreta disminución de este índice con el paso del tiempo, lo que debe estar asociado a buenas prácticas en la gestión de la cuenca.

Estos resultados corroboran que en el período comprendido entre los años 2011 al 2022 la gestión responsable de la cuenca Mayarí permitió una mejor planificación territorial y manejo de sus recursos naturales. En ese período se ha reportado que las áreas afectadas por la minería pudieron ser recuperadas, se reforestaron los bosques incrementándose el patrimonio forestal, se instalaron nuevas tecnologías para el despulpe de café y se pusieron en práctica tecnologías de avanzada para el desarrollo agropecuario y el uso eficiente del agua [10].

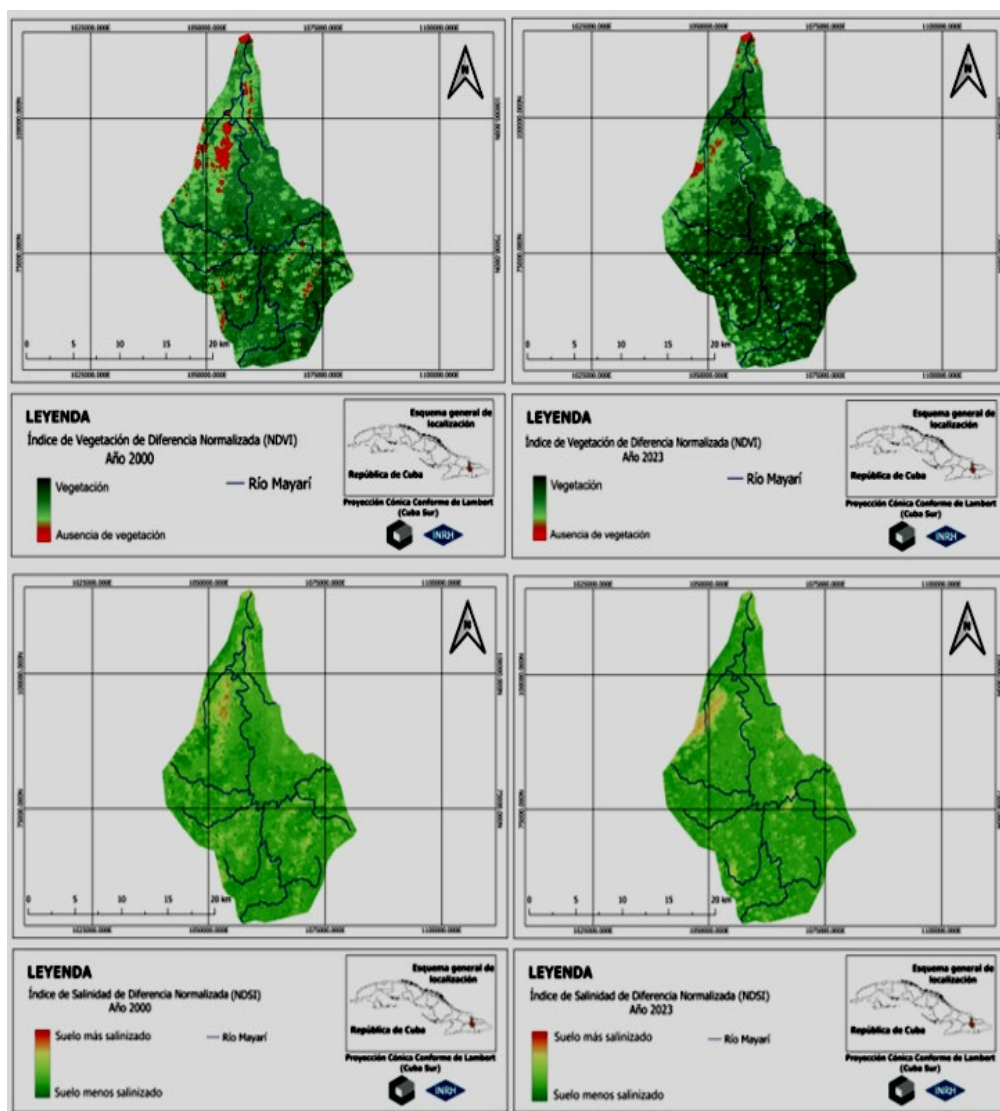


Fig. 3 Variación espacio-temporal del NDVI y del NDSI entre los años 2000 (a la izquierda) y 2023 (a la derecha) en la cuenca hidrográfica Mayarí

4. Conclusiones

El empleo del algoritmo de aprendizaje automático k- medias resultó útil y efectivo para evaluar la vulnerabilidad natural a la contaminación de las aguas subterráneas, previa selección de variables informativas y no redundantes. Este algoritmo permitió obtener un mapa de susceptibilidad a la contaminación por el efecto de la migración vertical de potenciales contaminantes depositados en la superficie del terreno a escala 1:100 000, de gran utilidad en la toma de decisiones informadas que contribuyan a establecer estrategias de manejo ambiental, proyectar mediciones de control de la calidad de las aguas subterráneas y la mejor planificación territorial para proteger este importante recurso.

El análisis de la variación espacio - temporal de la vegetación y la salinidad del suelo entre los años 2000 y 2023 mostró que, en general, el índice de vegetación aumentó y la salinidad de los suelos disminuyó, lo cual favorece la protección de la calidad de los recursos hídricos en esta zona.

El procedimiento presentado forma parte de un sistema de alerta temprana para prevenir el posible deterioro de la calidad de las aguas subterráneas en la cuenca hidrográfica Mayarí. Puede ser aplicado a otras cuencas del territorio nacional y a escalas de trabajo más detalladas.

Referencias

1. García, J., Gutiérrez, J. *La gestión de cuencas hidrográficas en Cuba*, 2015. Instituto Nacional de Recursos Hidráulicos. La Habana.
2. Foster, S., Hirata, R., Andreo, B. *The aquifer pollution vulnerability concept: aid or impediment in promoting groundwater protection?* Hydrogeology Journal [en línea], 2013. 26: DOI: <https://doi.org/10.1007/s100-40-013-1019-7>
3. García, J. A., Sacasas, C., Díaz-Deulofeu, A. *Vulnerabilidad a la contaminación en horizontes acuíferos del sur de Cuba occidental*. Ingeniería Hidráulica y Ambiental, 2018. 34 (1): p. 73-83.
4. Vázquez, B., Farfán, G. *Evaluación de la vulnerabilidad y el riesgo de contaminación de las aguas subterráneas del acuífero cársico Vento*. Infomin, 2019. 11 (1): p. 1-16.
5. Valcarce, R. M., Vega, M., Rodríguez, W., Suárez, O. *Vulnerabilidad intrínseca de las aguas subterráneas en la cuenca Almendares - Vento*. Ingeniería Hidráulica y Ambiental, 2020. 41 (2): p. 33-47.
6. Valcarce, R. M., Solís, E. *Evaluación de la vulnerabilidad intrínseca a la contaminación de las aguas subterráneas en las cuencas Cuyaguateje y Costera Sur de Pinar del Río, Cuba*. Geociencias, 2021. 40 (3): p. 751-761.
7. Valcarce, R. M., Suárez, O., Rodríguez, W., Vega, M. *Aplicación de la minería de datos a la evaluación de la vulnerabilidad de acuíferos*. Revista Cubana de Ciencias Informáticas, 2021. 15(2). Recuperado a partir de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992021000200001
8. Cotilla, M., González, E., Cañete, C., Díaz, J., Carral, R. *La red fluvial de Cuba y su interpretación morfoestructural*. Revista Geográfica [Publicación periódica en línea], 2003. Recuperado a partir de <http://www.jstor.org/stable/40992852>
9. Iturralde-Vinent, M. *Geología de las ofiolitas de Cuba*. [en línea]. Cuba: 1996 [Recuperado a partir de http://www.redciencia.cu/geobiblio/paper/1996_Iturralde_geologia%20de%20las%20ofiolitas,%20cuba.pdf
10. Ramírez, R.D., Cabrera, F.A. *Cuenca Mayarí: un caso de estudio especial en el contexto de las transformaciones ambientales*. Ciencias Holguín [Publicación periódica en línea] 2023. 29 (4): p. 60-73. Recuperado a partir de <http://www.ciencias.holguin.cu/revista/article/view/298>
11. Margat, J. *Groundwater vulnerability to contamination*, 1968. Orleans, Francia: BRGM.
12. Vrba, J., Zaporozec, A. *Guidebook on mapping ground water vulnerability*, 1994. Heise, Hannover: International Association of Hydrogeologists.
13. Vargas-Quintero, M. *Propuesta metodológica para la evaluación de la vulnerabilidad intrínseca de los acuíferos a la contaminación*, 2010. Bogotá: Ministerio de Ambiente, Vivienda y Desarrollo Territorial. Recuperado a partir de <https://archivo.minambiente.gov.co>
14. Dörfliker, N., Jauffret, D., Et Loubier, S. *Cartographie de la vulnérabilité des aquifères karstiques en Franche-Comté*, 2004. Francia: BRGM RP-53576-FR. Recuperado a partir de <https://www.google.com.cu/search?q=Cartographie+de+la+vulnérabilit%C3%A9+des+aquifères+karstiques+en+Franche+%E2%80%93+Comté%C3%A9&spell=1&sa=X&ved=2ahUKewj-176iprLqAhVHU8KHUsLBpQQkeECKAB6BAGLECo&biw=1999&bih=979>
15. Gómez – Escalonilla, V., Martínez-Santos, P.A. *Machine Learning Approach to Map the Vulnerability of Groundwater Resources to Agricultural Contamination*, 2024. Hydrology [Publicación periódica en línea]. 11(9): DOI: <https://doi.org/10.3390/hydrology11090153>
16. Liang, Y., Zhang, X., Sun, Y., Yao, L., Wu, J., Chen, S., Li, J., Wang, J. *Enhanced groundwater vulnerability assessment to nitrate contamination in Chongqing, Southwest China: Integrating novel explainable machine learning algorithms with DRASTIC-LU*, 2024. Hydrology Research. 55(6): p. 1-20. DOI: <https://doi.org/10.2166/nh.2024.036>

17. Hernández-Orallo, J., Ramirez-Quintana, M.J., Ferri-Ramírez, C. *Introducción a la minería de datos*, 2004. España: Pearson Educación.
18. Foster, S., Hirata, R., Andreo, B. *Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy*, 1987. En: W. Van Duijvenbooden y H. G. Van Waegeningh, editores. Proceedings and information n0 38 of the international conference on vulnerability of soil and groundwater to pollutants. TheHague. North Holland. p. 69-17.
19. Moya, R. *Selección del número óptimo de clusters*, 2016. Recuperado a partir de <https://jrrob.com/seleccion-del-numero-optimo-clusters>
20. García-Reyes, R.A., González, M.D., Torres-Calzado, K., Villazón-Gómez, J.A., Abellón-Molina, M.I., Velázquez-Sánchez, E.C. *Sensoramiento remoto de la salinidad en el agroecosistema Mayarí de la provincia Holguín, Cuba*, 2021. Ciencias Técnicas Agropecuarias. 30 (1): p. 26-32.
21. Aslanov I., Kholdorov S., Ochilov S., Jumanov A., Jabbarov Z., Jumaniyazov I., et al. *Evaluation of soil salinity level through using Landsat-8 OLI in Central Fergana valley, Uzbekistan.E3S*, 2021. DOI:<https://doi.org/10.1051/e3sconf/202125803012>

Conflicto de Intereses

Los autores declaran que no existe conflicto de intereses relacionado con la investigación que se presenta.

Contribución de los autores

Camila Cruz González. ORCID: <https://orcid.org/0009-0003-5910-2239>

Participó en el diseño de la investigación, el procedimiento para realizar la cartografía de vulnerabilidad a la contaminación de aguas subterráneas en la cuenca hidrográfica Mayarí y la obtención de los resultados. Participó en la redacción, revisión y edición del manuscrito.

Rosa María Valcarce Ortega. ORCID: <https://orcid.org/0000-0001-9981-6832>

Participó en el diseño de la investigación, el procedimiento para realizar la cartografía de vulnerabilidad a la contaminación de aguas subterráneas en la cuenca hidrográfica Mayarí, supervisión, evaluación de los resultados obtenidos, redacción y edición del manuscrito.

Marina Beatriz Vega Carreño. ORCID: <https://orcid.org/0000-0001-6745-5382>

Participó en el diseño de la investigación y la revisión de los resultados, así como en la edición del manuscrito.

Willy Roberto Rodríguez Miranda. ORCID: <https://orcid.org/0000-0003-2938-6472>

Participó en el diseño de la investigación, adquisición de fondos, administración del proyecto, revisión y evaluación de los resultados obtenidos.