



Universidad Nacional de Entre Ríos
FACULTAD DE INGENIERÍA

Aplicación de un modelo
epidemiológico para la predicción del
crecimiento y difusión del vector que
transmite el dengue utilizando
sensores remotos en la
localidad de Oro Verde

TRABAJO INTEGRADOR FINAL
TIC Y Geomática

Autor:
Justo Garcia
Docente a cargo:
Walter Elías

Junio 2024

Tabla de contenidos

1	Introducción	2
2	Desarrollo	3
2.1	Área de estudio	3
2.2	Obtención de imágenes	4
2.2.1	Dispositivo de sensado	4
2.2.2	Análisis visual	5
2.3	Procesamiento de imágenes	6
2.4	Descripción del modelo	7
2.5	Cálculo de índices	9
2.5.1	NDVI	9
2.5.2	NDWI	11
2.5.3	NDBI	11
2.5.4	NDBaI	12
2.5.5	NDMI	12
2.6	Clasificación de superficies	13
2.7	Campo de atracción	14
2.7.1	Identificación de viviendas	15
2.7.2	Cálculo del campo de atracción	16
2.8	Densidad Inicial	17
2.8.1	Lectura y Procesamiento de Datos	17
2.8.2	Interpolación Gaussiana	18
2.8.3	Visualización de Resultados	18
3	Conclusiones	20

1 Introducción

El presente informe detalla un proyecto realizado en el marco del cursado de la materia TIC y Geomática de la Facultad de Ingeniería de la Universidad Nacional de Entre Ríos, enfocado en el análisis de la distribución espacial de vectores del dengue utilizando tecnologías geoespaciales y análisis de imágenes satelitales. Todas las técnicas utilizadas en este proyecto tienen como objetivo extraer la información necesaria para aplicar un método numérico concreto de la bibliografía. Aunque se plantea la resolución matemática del modelo como una proyección futura, las técnicas y procedimientos de extracción, obtención y procesamiento de los datos necesarios para ello se implementan y explican detalladamente en este informe.

El dengue, una enfermedad transmitida por mosquitos del género *Aedes*, representa un desafío significativo para la salud pública. El estudio de la distribución de estos vectores es esencial para la implementación de estrategias. La combinación de datos satelitales y herramientas geoespaciales permite explorar como diferentes factores afectan la proliferación de los mosquitos.

El objetivo principal de este proyecto es aplicar los conocimientos adquiridos en la materia en una problemática real. A partir de la búsqueda bibliográfica se plantea la replicación en un área de interés diferente (Oro Verde, Entre Ríos, Argentina) de un trabajo que caracteriza y modela la distribución espacial de los vectores del dengue.

Este informe detalla el proceso desde la obtención de imágenes satelitales hasta el procesamiento geoespacial de los datos, incluyendo la delimitación del área de estudio, la extracción y análisis de índices espectrales como el NDVI y NDWI, la clasificación de superficies utilizando técnicas avanzadas en el análisis de datos y finalmente la implementación de la simulación propiamente dicha.

2 Desarrollo

En este informe se detalla el procedimiento seguido para el procesamiento de datos, el cual es crucial para su posible replicación futura. Este procedimiento incluye la obtención de imágenes, el uso de diferentes índices espectrales, y la clasificación de superficies mediante técnicas avanzadas de análisis de datos, todo orientado a extraer la información necesaria para un método numérico concreto de la bibliografía. La resolución matemática del modelo se proyecta para el futuro.

2.1 Área de estudio

Para este estudio, se utilizaron datos recolectados a partir de la utilización de ovitrampas por parte de un proyecto de la Facultad de Ingeniería de la Universidad Nacional de Entre Ríos en Oro Verde, Entre Ríos, Argentina. El área de interés se definió mediante un procesamiento en Python, utilizando las bibliotecas Folium y Pandas. Este enfoque nos permitió generar gráficos detallados que representan los puntos geográficos con información sobre la densidad de mosquitos registrada experimentalmente. En la siguiente imagen (Fig.1), se muestra la salida de este procesamiento, con un marcador representando cada ovitrampa.

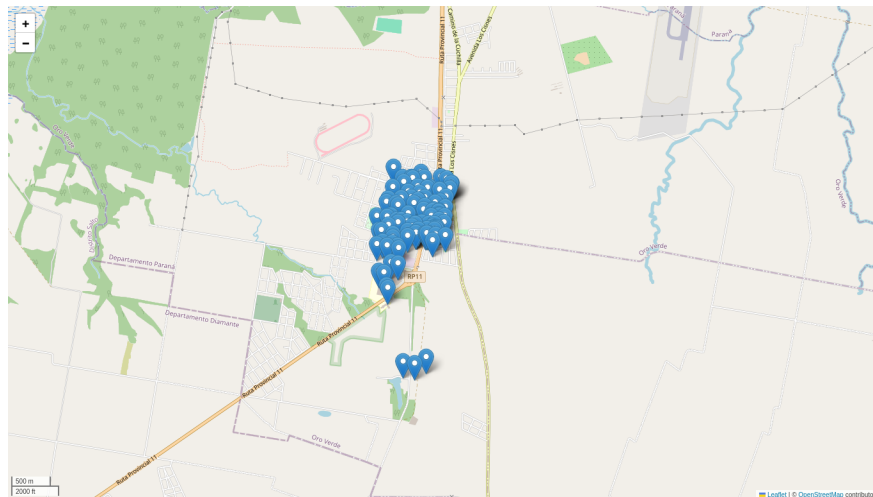


Fig. 1: Posiciones de ovitrampas en Oro Verde.

A partir de estos marcadores, se extendió el procesamiento de los datos definiendo un cuadrado que encierra a todos los puntos con un margen adicional de 1km (Fig.2). Esta

delimitación nos permitió enfocarnos en el área específica para la cuál tenemos datos, es decir, nuestra región de interés (ROI, Region of Interest).

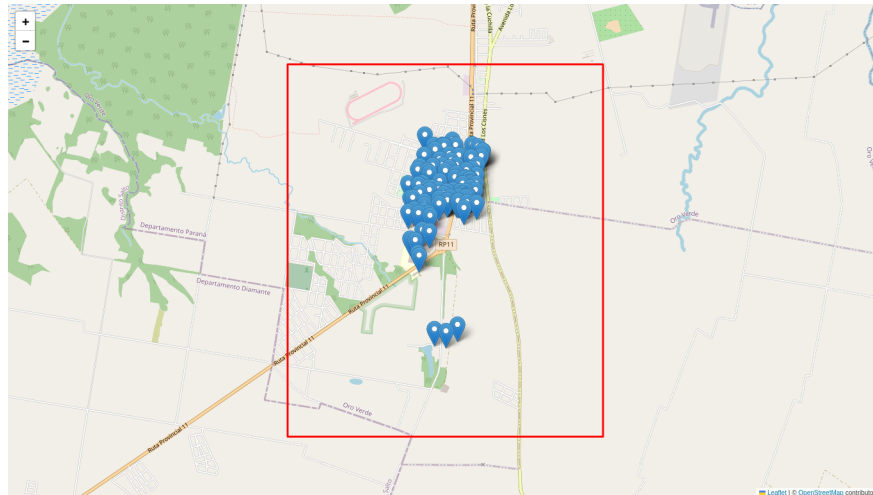


Fig. 2: Posiciones de ovitrampas en Oro Verde con ROI delimitada.

2.2 Obtención de imágenes

Para el análisis, se utilizaron imágenes satelitales provenientes de Landsat 8, las cuales se obtendrán a través del Earth Explorer del USGS (United States Geological Survey) CITA. Estas imágenes, con un nivel de análisis L2, proporcionan la información necesaria para estudiar la densidad de mosquitos en el área de interés a través de distintas técnicas.

2.2.1 Dispositivo de sensado

El Landsat 8 es un satélite de observación terrestre que forma parte del Programa Landsat, administrado por el USGS y la NASA CITA. Este satélite consta de dos sensores principales:

- OLI (Operational Land Imager)
- TIRS (Thermal Infrared Sensor)

A su vez, cada uno de estos sensores posee diversas bandas. Estas son listadas en la tabla 1.

Banda	Longitud de onda (μ)	Resolución espacial (m)
1 - Coastal/Aerosol	0.435-0.451	30
2 - Azul	0.452-0.512	30
3 - Verde	0.533-0.590	30
4 - Roja	0.636-0.673	30
5 - NIR	0.851-0.879	30
6 - SWIR-1	1.566-1.651	30
7 - SWIR-2	2.107-2.294	30
8 - Pancromático	0.503-0.676	15
9 - Cirro	1.363-1.384	30
10 - TIR-1	10.60-11.19	100
11 - TIR-2	11.50-12.51	100

Tabla 1: Tabla con información sobre las diferentes bandas que capta el Landsat 8

Para este estudio es muy relevante la diversidad de bandas espectrales, permitiendo analizar diversas características del terreno. Esta información, combinada con los datos de las ovitrampas, proporciona una visión integral y detallada del entorno.

2.2.2 Análisis visual

Como se mencionó previamente, se analizó la disponibilidad de imágenes a través de Earth-Explorer. Se seleccionó la región de interés en esta plataforma y se estableció un rango de fechas según la disponibilidad de datos, comenzando con datos de octubre de 2017. Se identificaron dos opciones de pasada del Landsat 8: una del path 227 y otra del 226, ambas para row 82. Estas opciones fueron marcadas en el mapa para evaluar la cobertura del área que proporcionaba cada una, como se muestra en Fig.3

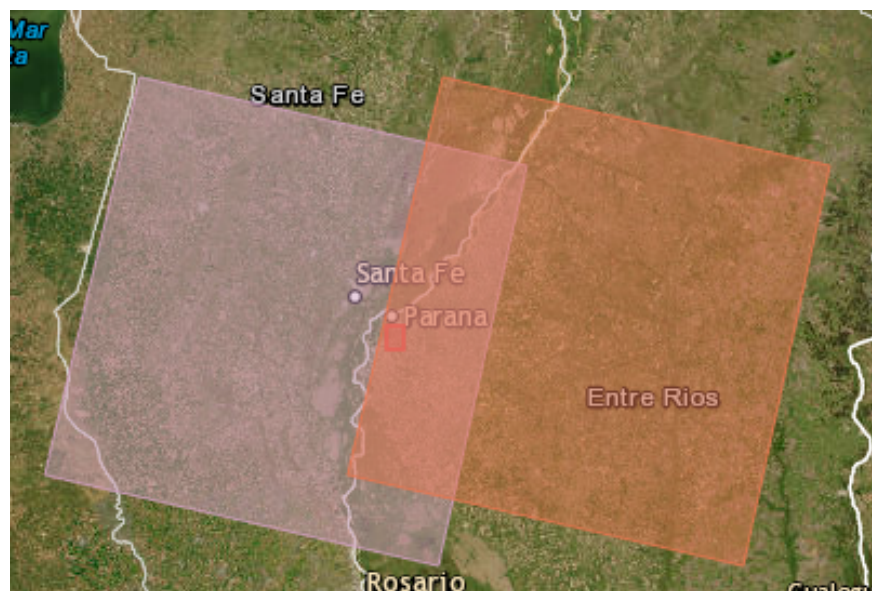


Fig. 3: Opciones de cobertura.

Finalmente, se decidió utilizar la opción de la izquierda, ya que en la opción de la derecha, Oro Verde se encuentra sobre el borde de la imagen, lo que limitaría la capacidad de ampliar la región de interés. Por lo tanto, se procedió a la descarga de las imágenes correspondientes a la fecha de interés, utilizando el path 227 y el row 82.

Una vez obtenidos los datos de sensado de interés, se procedió a realizar un análisis visual detallado de la región utilizando el software QGIS. Este proceso incluyó la superposición de bandas espectrales como capas ráster y de los datos de coordenadas de las ovitrampas como vectores.

2.3 Procesamiento de imágenes

Con el objetivo de reducir el costo computacional y evitar el desperdicio de recursos, se procedió a definir un polígono delimitado por las coordenadas del área de interés. Este proceso incluyó la conversión de las coordenadas de las ovitrampas desde EPSG:4326 a EPSG:32620, seguido de la creación de un archivo de polígono con extensión *.shp*.

Una vez que se estableció correctamente el archivo de polígono, se realizó una verificación visual utilizando QGIS para validar los cálculos (Fig.4). Posteriormente, se implementó el recorte de todos los archivos *.TIF* utilizando la herramienta *gdalwarp* de *GDAL*. REFER-

ENCIA



Fig. 4: Polígono generado.

2.4 Descripción del modelo

El modelo presentado fue inicialmente propuesto por Raffy y Tran y ha sido aplicado en otras regiones de nuestro país, como en Salta. El modelo se describe matemáticamente de la siguiente manera:

$$\frac{\partial \rho(P, t)}{\partial t} = \nabla \cdot (D_R \nabla \rho) - \nabla \cdot (\rho D_W V) - \nabla \cdot (\rho K_H \nabla H) + \alpha - \beta$$

Donde:

Símbolo	Variable	Valor
P	Densidad de mosquitos	No homogéneo
α	Tasa de nacimientos	$6(m^2/dia)$
β	Tasa de muertes	0.2
V	Velocidad Viento Superficie	No homogéneo
K_H	Tensor de atracción	100
H	Campo de atracción	No homogéneo
D_R	Tensor de difusión	No homogéneo / ver Tabla 2
D_W	Tensor de rugosidad	No homogéneo / ver Tabla 2

Desglosando:

$$\frac{\partial \rho}{\partial t}(P, t) = \text{Rozamiento}(P, t) - \text{Transporte}(P, t) - \text{Atraccion}(P, t) + \alpha(P, t) - \beta(P, t)$$

Donde:

$$\text{Rozamiento}(P, t) = \text{div}[D_R(P, t) \cdot \nabla \rho(P, t)]$$

$$\text{Transporte}(P, t) = [D_W(P, t) \vec{W}(P, t) \cdot \nabla \rho(P, t)]$$

$$\text{Atraccion}(P, t) = \text{div}[K_H \cdot \rho(P, t) \cdot \nabla H(P, t)]$$

Con:

- $\rho(P, t)$ = densidad de mosquitos
- $D_R(P, t)$ = tensor de rozamiento
- $D_W(P, t)$ = tensor de viento
- $W(P, t)$ = velocidad del viento

- K_H = constante de atracción
- $H(P, t)$ = campo de atracción
- $\alpha(P, t)$ = razón de nacimientos
- $\beta(P, t)$ = razón de muertes

La resolución de estas ecuaciones se plantea como una proyección futura. En secciones posteriores se lleva a cabo la extracción de la información necesaria para aplicar el modelo en el área de estudio, como la medición de los parámetros de rugosidad y difusión (D_W y D_R) a partir de la clasificación de superficies. La integración de estos datos permitirá la resolución del modelo, facilitando la predicción de la dinámica de las poblaciones de mosquitos en el área analizada.

2.5 Cálculo de índices

El cálculo de índices en geomática es una herramienta crucial para llevar a cabo diversos análisis y aplicaciones a partir de la transformación de la información contenida en las imágenes satelitales. Por esta razón, en este trabajo se utiliza una selección de algunos de estos índices que son relevantes para cumplir con el objetivo, a continuación se detallan aquellos utilizados.

2.5.1 NDVI

2.5.1.1 Cálculo

El NDVI (Landsat Normalized Difference Vegetation Index) es usado para cuantificar cuán verde está la vegetación y es de utilidad a la hora de comprender la densidad de la vegetación. REFERENCIA.

Su cálculo de forma generalizada se lleva a cabo con la siguiente combinación de bandas:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

En el caso específicos de bandas de Landsat 8:

$$NDVI = \frac{Banda\ 5 - Banda\ 4}{Banda\ 5 + Banda\ 4}$$

En primer instancia, se llevó a cabo el cálculo con la función de calculadora ráster de QGIS. Sin embargo, luego se procedió a llevar a cabo una implementación en Python para su cálculo con el fin de tener el código necesario para automatizar su cálculo. Se evaluó una comparativa de los resultados obtenidos y el grado de precisión del cálculo era el mismo para ambos (Fig.5)

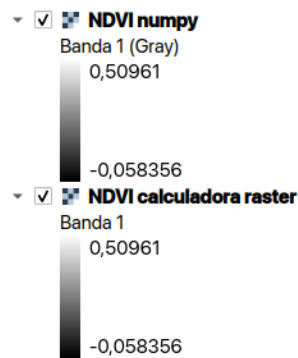


Fig. 5: Comparativa de cálculos del NDVI.

2.5.1.2 Clasificación

Para la clasificación del tipo de vegetación en base al NDVI se consideraron cuatro clases: 1) Suelo expuesto, 2) Vegetación baja, 3) Vegetación alta; las cuales se etiquetaron con el siguiente criterio (Tabla2):

Rango	Tipo
Suelo expuesto	0.0 – 0.25
Vegetación baja	0.25 – 0.4
Vegetación alta	> 0.4

Tabla 2: Criterio de clasificación en base al NDVI.

2.5.2 NDWI

El Índice Normalizado de Diferencia de Agua (NDWI) se emplea para destacar áreas de agua en una imagen captada por un satélite. Se utiliza principalmente para la detección y monitoreo de masas de agua

2.5.2.1 Cálculo

El cálculo de este índice se lleva a cabo a partir de las bandas verdes y de infrarrojos cercanos, respetando la siguiente fórmula:

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

Reformulando teniendo en cuenta la información correspondiente a las bandas del Landsat 8:

$$NDWI = \frac{Banda\ 3 - Banda\ 5}{Banda\ 3 + Banda\ 5}$$

El cálculo de este índice se realiza de la misma forma que en la sección 2.5.1.

2.5.2.2 Clasificación

A las clases descritas en 2.5.1.2, se añade una nueva que representa cuerpos de agua. Se define como cuerpos de agua aquellos puntos con un valor de *NDWI* superior a cero.

2.5.3 NDBI

2.5.3.1 Cálculo

El cálculo de este índice se lleva a cabo a partir de las bandas verdes y de infrarrojos cercanos, respetando la siguiente fórmula:

$$NDBI = \frac{SWIR1 - NIR}{SWIR1 + NIR}$$

Reformulando teniendo en cuenta la información correspondiente a las bandas del Landsat 8:

$$NDBI = \frac{Banda\ 6 - Banda\ 5}{Banda\ 6 + Banda\ 5}$$

El cálculo de este índice se realizó de la misma forma que en la sección 2.5.1.

2.5.4 NDBaI

El índice NDBaI, o Índice Normalizado de Diferencia de Bareness y Albedo Integrado combina información sobre la ausencia de vegetación (bareness) y la reflectancia de la superficie (albedo). Es útil para identificar áreas sin vegetación y evaluar la reflectancia de la superficie, proporcionando información complementaria en estudios de cobertura del suelo.

2.5.4.1 Cálculo

$$NDBaI = \frac{SWIR1 - TIRS1}{SWIR1 + TIRS1}$$

Para las bandas específicas del Landsat 8, la fórmula se ajusta a:

$$NDBaI = \frac{Banda\ 6 - Banda\ 10}{Banda\ 6 + Banda\ 10}$$

2.5.5 NDMI

El Índice Normalizado de la Diferencia de Humedad (NDMI) se utiliza para evaluar la humedad en la superficie terrestre utilizando información de las bandas de infrarrojo cercano (NIR) e infrarrojo medio (SWIR1). Es especialmente útil para detectar y monitorear la

presencia y cambios en la humedad del suelo y la vegetación, siendo aplicable en estudios ambientales y de recursos naturales.

2.5.5.1 Cálculo

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1}$$

La fórmula para calcular el NDMI en el contexto de Landsat 8 es la siguiente:

$$NDMI = \frac{Banda\ 5 - Banda\ 6}{Banda\ 5 + Banda\ 6}$$

2.6 Clasificación de superficies

Habiendo definido claramente los tipos de superficie a identificar y los criterios para determinarlos, se llevó a cabo la clasificación de los puntos del ráster siguiendo la secuencia lógica ya desarrollada de evaluación de índices. Todos los procesamiento relevantes para llevar a cabo este cometido se realizaron con Python y módulos propios de este lenguaje.

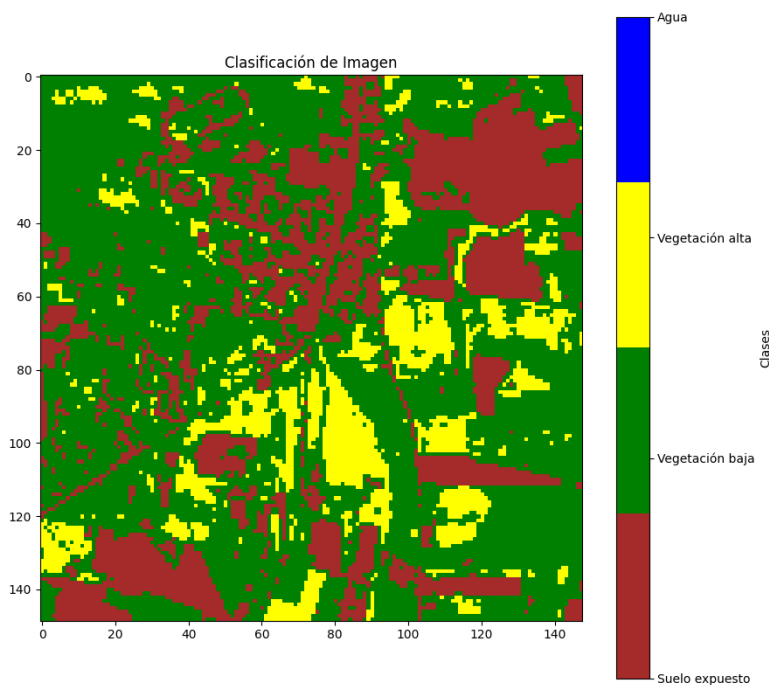
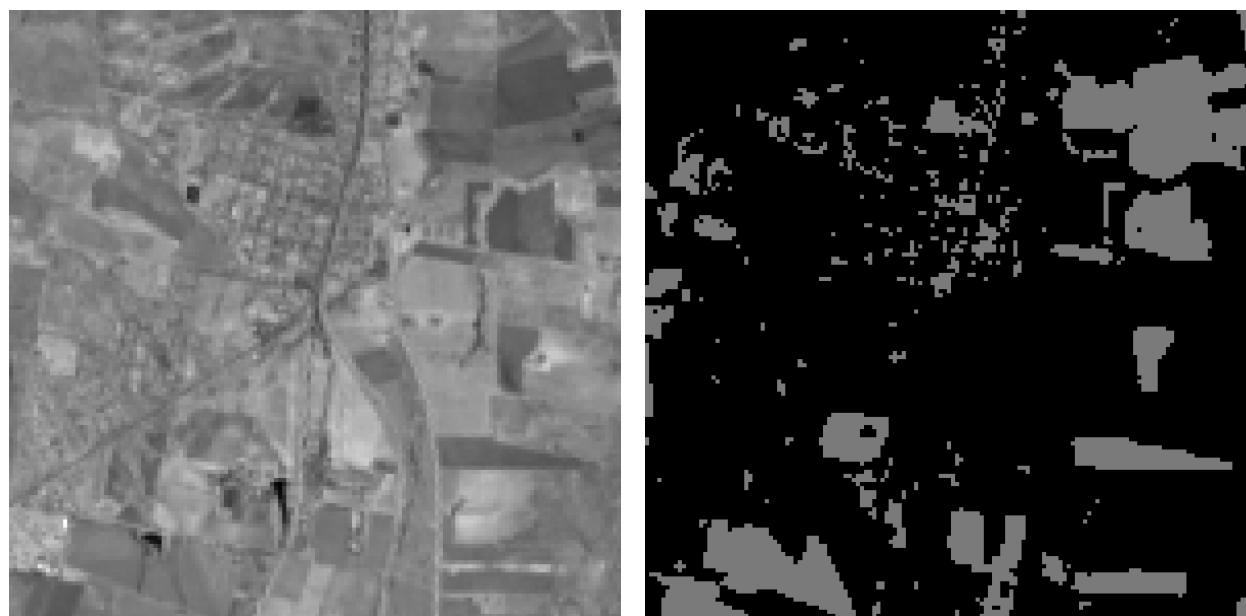


Fig. 6: Capa con la clasificación para las distintas clases.

2.7 Campo de atracción

El campo de atracción es un ráster del tamaño del área de interés, el valor de un píxel se determina por la densidad de píxeles que representan viviendas en su entorno cercado. Cuando un píxel contiene una vivienda (es decir, una fuente de hematofuentes) y está rodeado por una ventana de 90x90 metros (3 píxeles) que contiene exclusivamente viviendas el valor de atracción es máximo, estableciéndose en 1. En contraste, un píxel que no contiene viviendas y está rodeado por áreas que no incorporan hematofuentes obtiene un valor de atracción nulo. REFERENCIA



(a) Banda 5 del área de interés.

(b) *NDBI* positivo como vivienda.

Fig. 7: Comparativa de la clasificación con el área de interés.

2.7.1 Identificación de viviendas

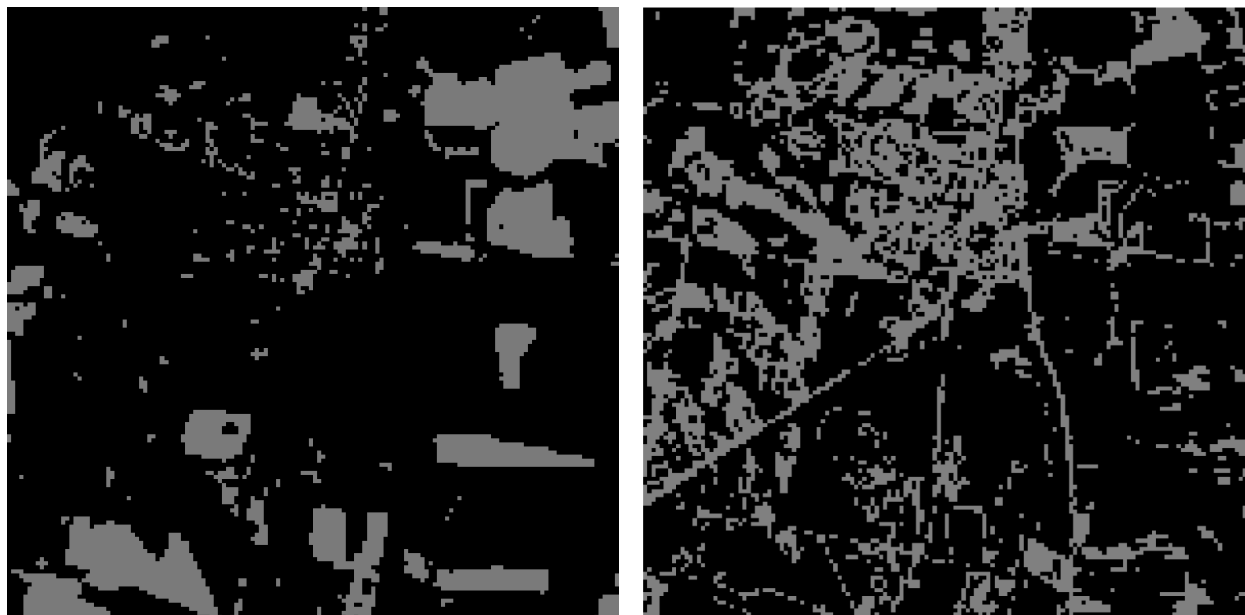
La identificación precisa de viviendas es fundamental para la determinación del campo de atracción en el análisis de datos geospaciales. Inicialmente, se empleó únicamente el NDBI como criterio para clasificar áreas con posibles viviendas.

La validación de esta clasificación se llevó a cabo mediante un análisis visual en QGIS y un conocimiento profundo del área de estudio (ver Fig.7), se observó que el enfoque no logró identificar adecuadamente las zonas urbanas completas. Además, algunas áreas rurales fueron incorrectamente clasificadas como urbanas, indicando una falta de precisión en la distinción entre áreas construidas y no construidas.

Con esta información sobre la mesa, se tomó la decisión de implementar una clasificación no supervisada utilizando el algoritmo de K-Means. Este enfoque fue elegido para mitigar posibles sesgos introducidos anteriormente y mejorar la precisión en la identificación de áreas urbanas y rurales. Se integraron diversos índices, incluyendo NDBI, NDVI, NDMI y NDBaI, con el objetivo de capturar múltiples aspectos del entorno que podrían indicar presencia de viviendas y otras estructuras construidas.

Después de calcular cada uno de los índices, se empleó el NDVI para generar una máscara

identificando así los píxeles que corresponden a áreas de vegetación.



(a) Clasificación con *NDBI*.

(b) Clasificación con K-Means.

Fig. 8: Comparativa de la clasificación con el área de interés.

Como se puede observar en la Figura 8, la clasificación resulta significativamente más precisa en comparación con el método anterior basado únicamente en *NDBI*. Esta mejora es evidente al considerar el conocimiento detallado del área de interés, donde ahora se logra una distinción más clara entre zonas urbanas y rurales. Por lo tanto, a partir de este punto, se adopta este criterio basado en la clasificación con K-Means para identificar y delimitar áreas que contienen viviendas en el estudio.

2.7.2 Cálculo del campo de atracción

Basandose en los fundamentos establecidos en la sección 2.7, se procedió a implementar el algoritmo para calcular el campo de atracción. En términos generales, este cálculo se realiza para cada punto según la siguiente fórmula:

$$Atraccion = 1 * \frac{Viviendas\ vecinas}{Vecinas}$$

Este proceso genera una matriz de números decimales que reflejan la atracción que el vector sienten hacia cada píxel debido a la presencia de hematofuentes, variando de 0 a 1 (ver Figura 9).

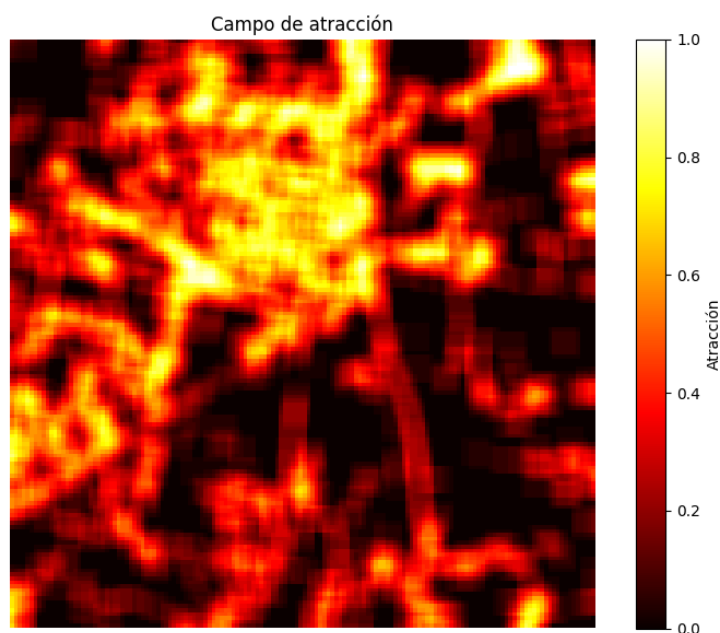


Fig. 9: Campo de atracción para la región de interés.

2.8 Densidad Inicial

En esta subsección, se detalla el procedimiento utilizado para calcular la densidad inicial de huevos en el área de interés, empleando técnicas avanzadas de procesamiento de imágenes y análisis de datos geoespaciales.

2.8.1 Lectura y Procesamiento de Datos

El primer paso consistió en la lectura y procesamiento de los datos de entrada. Se utilizaron dos tipos de datos: una imagen rasterizada en formato TIF y un archivo CSV con coordenadas y conteos de huevos.

El archivo CSV contiene datos puntuales con coordenadas y conteos de huevos, los cuales se filtraron por una fecha específica a modo de ejemplo para enfocarse en un momento

determinado del estudio, que se corresponde con aquella en la que más huevos se registraron.

2.8.2 Interpolación Gaussiana

Para calcular la densidad espacial de los huevos, se empleó una técnica de interpolación basada en una función gaussiana bidimensional. La función gaussiana es una representación matemática que describe una distribución normal en dos dimensiones, caracterizada por su forma de campana:

$$G(x, y) = \exp \left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2} \right) \quad (1)$$

Se utilizó esta función para modelar la distribución espacial de los huevos alrededor de cada punto de datos.

2.8.3 Visualización de Resultados

Finalmente, los resultados se visualizaron superponiendo el mapa de densidad calculado sobre la imagen TIF original. Esta visualización permite observar la distribución espacial de la densidad de huevos en la región de estudio (Figura 10).

Mapa de densidad de número de huevos sobre imagen TIF (Interpolación gaussiana)

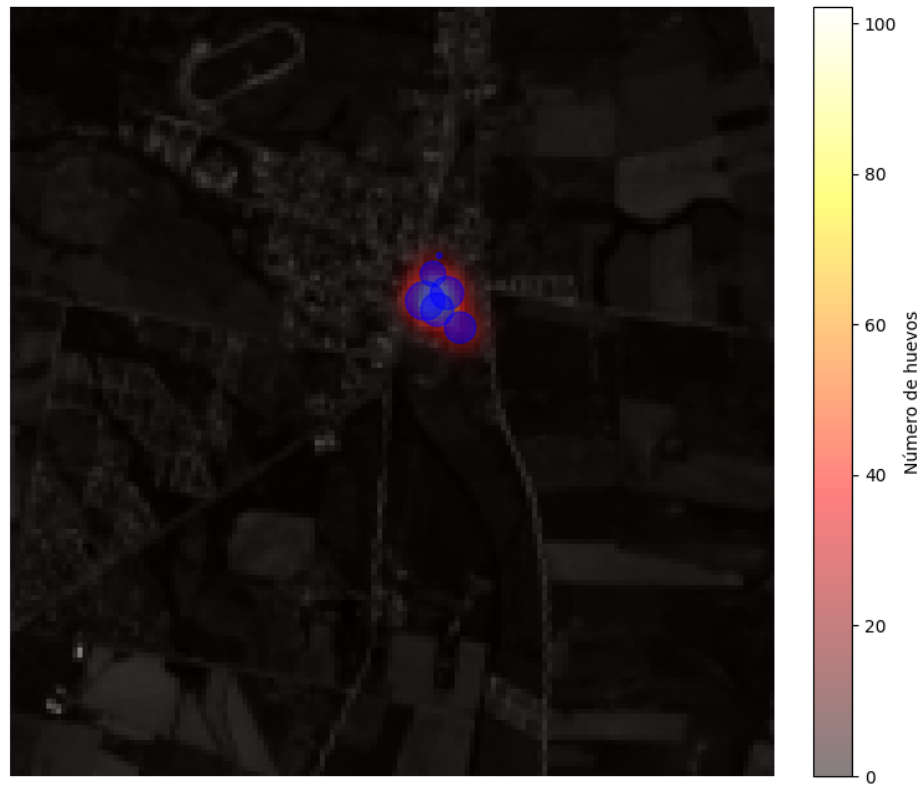


Fig. 10: Densidad inicial para la región de interés.

3 Conclusiones

3 Conclusiones