

Datos Geoespaciales, Sensores Remotos y SIG

Data Science (CC3084) - 2026

Datos Geoespaciales, Sensores Remotos y SIG

Semestre 02, 2026

Introducción

Casi todo dato lleva pegada una pregunta implícita: ¿dónde ocurrió? Una venta ocurre en una tienda, un caso de dengue en una aldea, una deforestación en una parcela concreta. Cuando un dato incluye su ubicación en la superficie de la Tierra, decimos que es un dato geoespacial.

Trabajar con la dimensión espacial no es solo agregar dos columnas de latitud y longitud. La ubicación cambia las reglas del juego: los datos cercanos se parecen entre sí, la Tierra es curva y los mapas mienten, y las mismas observaciones dan resultados distintos según cómo se agrupen en el espacio.

La primera ley de la geografía

Waldo Tobler la formuló así: "todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las lejanas".

Es la idea que sostiene casi todo el análisis espacial:

- El precio de una casa se parece al de la casa de al lado.
- La temperatura de hoy en un punto se parece a la de un kilómetro más allá.
- Un brote de enfermedad tiende a aparecer en vecindarios contiguos.

Esta dependencia entre observaciones cercanas se llama autocorrelación espacial, y rompe un supuesto básico de la estadística clásica: que las observaciones son independientes.

Tres grandes bloques de esta lección

- Datos geoespaciales: cómo se representa la ubicación y en qué formatos vive.
- Sensores remotos: cómo se capturan datos de la Tierra desde satélites, aviones y drones.
- SIG: los sistemas y las operaciones para almacenar, consultar y analizar todo lo anterior.

Anatomía de un dato geoespacial

Un registro geoespacial tiene tres componentes:

- Componente espacial: dónde está. Una geometría (punto, línea, polígono) o una celda de una cuadrícula.
- Componente temporal: cuándo. Muchos fenómenos geográficos cambian, y la misma parcela no es igual en enero que en junio.
- Atributos: qué hay ahí. Población, temperatura, tipo de cultivo, número de casos.

Un dato con las tres dimensiones se llama espacio-temporal, y es el caso más común en la práctica: series de tiempo por ubicación.

Datos discretos y datos continuos

- Fenómenos discretos: objetos con bordes definidos. Un poste, una carretera, un municipio. Se representan bien como vectores.
- Fenómenos continuos: magnitudes que existen en todo punto del espacio. Elevación, temperatura, precipitación, reflectancia. Se representan bien como rásteres.

Elegir mal la representación es una fuente frecuente de errores: forzar un fenómeno continuo a polígonos crea fronteras artificiales.

Coordenadas, datums y proyecciones

La Tierra no es una esfera perfecta ni una superficie plana, y esa incomodidad se cuela en todo dato geoespacial.

- Geoide: la forma real de la Tierra, definida por su campo gravitatorio. Es irregular.
- Elipsoide: una aproximación matemática al geoide, más simple de calcular.
- Datum: el elipsoide más su anclaje al planeta. Define el origen desde el cual se miden las coordenadas. El más usado es WGS84, el del GPS.

Dos coordenadas idénticas en datums distintos pueden apuntar a lugares separados por cientos de metros.

Coordenadas geográficas

- Latitud: ángulo hacia el norte o el sur desde el ecuador, de -90° a 90° .

- Longitud: ángulo hacia el este o el oeste desde el meridiano de Greenwich, de -180° a 180° .
- Se expresan en grados, no en metros. Un grado de latitud son unos 111 km en cualquier parte, pero un grado de longitud mide 111 km en el ecuador y 0 km en los polos.

Consecuencia práctica: calcular distancias o áreas restando grados es incorrecto. Para eso hay que proyectar o usar fórmulas sobre la esfera.

Proyecciones

Una proyección es la receta para pasar de la superficie curva de la Tierra a un plano. Toda proyección distorsiona algo; no existe una perfecta.

- Conformes: preservan los ángulos y las formas locales, pero deforman áreas. Mercator es el ejemplo clásico: Groenlandia parece del tamaño de África, cuando es catorce veces más pequeña.
- Equivalentes: preservan las áreas, pero deforman las formas. Útiles para mapas temáticos donde importa comparar superficies.
- Equidistantes: preservan las distancias, pero solo desde ciertos puntos o a lo largo de ciertas líneas.

Reglas prácticas:

- Para medir distancias y áreas, usa una proyección métrica local, no grados.
- UTM divide el mundo en 60 husos de 6° de ancho y da coordenadas en metros dentro de cada uno. Guatemala cae en los husos 15 y 16 norte.
- Para mapas web, la proyección de facto es Web Mercator (EPSG:3857), buena para navegar y mala para comparar áreas.

Códigos EPSG

Cada sistema de referencia de coordenadas (CRS) tiene un código EPSG que lo identifica sin ambigüedad:

- EPSG:4326 — WGS84 en grados de latitud y longitud.
- EPSG:3857 — Web Mercator, en metros, la de los mapas de la web.
- EPSG:32615 / 32616 — UTM zonas 15N y 16N sobre WGS84, en metros.

La regla número uno al combinar capas: verificar que todas estén en el mismo CRS. Superponer datos en CRS distintos produce mapas que se ven bien y análisis que están mal.

Modelos de datos: vectorial y ráster

Hay dos maneras de representar el mundo, y casi todo dato geoespacial es una u otra.

Modelo vectorial

Representa el espacio con geometrías definidas por coordenadas.

- Punto: una ubicación sin dimensión. Un hospital, un árbol, un evento de GPS.
- Línea: una secuencia de puntos conectados. Una carretera, un río, una ruta de bus.
- Polígono: una línea cerrada que encierra un área. Un municipio, una parcela, un lago.
- Multi-geometrías: un mismo registro puede tener varias partes; un archipiélago es un multipolígono.

Cada geometría lleva una tabla de atributos asociada. Es el modelo natural para objetos discretos y para datos administrativos.

Modelo ráster

Representa el espacio como una cuadrícula regular de celdas, cada una con un valor.

- Es esencialmente una matriz, igual que una imagen digital.
- La resolución espacial es el tamaño de la celda en el terreno: 10 m, 30 m, 250 m.
- Cada capa (banda) es una matriz distinta sobre la misma cuadrícula.

Es el modelo natural para fenómenos continuos y para todo lo que viene de sensores remotos.

Cuándo usar cada uno

- Vectorial: fronteras, redes, ubicaciones puntuales, datos censales. Preciso, ligero, fácil de consultar por atributos.
- Ráster: imágenes satelitales, elevación, clima, cualquier superficie continua. Fácil de operar celda por celda, pesado en almacenamiento.

En la práctica se combinan constantemente: recortar un ráster de NDVI con el polígono de un municipio y sacar el promedio es una operación cotidiana.

Formatos y estándares

Formatos vectoriales

- Shapefile: el formato histórico de ESRI. No es un archivo sino varios (.shp, .shx, .dbf, .prj) que deben viajar juntos. Limita los nombres de campo a 10 caracteres. Sigue vivo por inercia.

- GeoJSON: JSON con geometrías. Legible, ideal para la web y para APIs. Se asume EPSG:4326.
- GeoPackage (.gpkg): base SQLite con capas vectoriales y ráster en un solo archivo. Es el reemplazo moderno recomendado del shapefile.
- KML / KMZ: el formato de Google Earth, orientado a visualización.
- WKT y WKB: representaciones de texto y binaria de una geometría, usadas dentro de bases de datos.

Formatos ráster

- GeoTIFF: un TIFF con metadatos de georreferenciación incrustados. Es el estándar de facto.
- COG (Cloud Optimized GeoTIFF): un GeoTIFF organizado internamente para que se pueda leer por pedazos desde la nube sin descargar todo el archivo.
- NetCDF y HDF: formatos multidimensionales para datos climáticos y series temporales; guardan cubos de datos (x, y, tiempo, variable).
- Zarr: formato pensado para cubos de datos enormes en almacenamiento en la nube.

Estándares y servicios

El Open Geospatial Consortium (OGC) define estándares para que los sistemas se entiendan entre sí:

- WMS: servicio que entrega mapas ya renderizados como imágenes.
- WFS: servicio que entrega las geometrías vectoriales, no la imagen.
- WCS: servicio que entrega datos ráster crudos.
- STAC: catálogo estándar para buscar imágenes satelitales por fecha, área y sensor.

Sensores remotos

Sensor remoto es cualquier instrumento que mide algo sin estar en contacto con el objeto medido. En la práctica: observar la Tierra desde el aire o el espacio.

La cadena es siempre la misma: una fuente de energía ilumina la superficie, la superficie refleja o emite radiación, el sensor la capta, y el resultado se convierte en una matriz de números georreferenciada.

Plataformas

- Satélites: cobertura global y revisita periódica. Órbita polar heliosíncrona para observación de la Tierra; órbita geoestacionaria para meteorología.
- Aviones: mayor resolución, vuelo bajo demanda, cobertura limitada.
- Drones: resolución de centímetros, control total de cuándo se vuela, cobertura de pocas hectáreas. Es la vía más accesible para proyectos locales.

Sensores pasivos y activos

- Pasivos: miden energía que no producen ellos. Dependen del Sol o de la emisión térmica de la superficie. Las cámaras multiespectrales y térmicas son pasivas. No ven de noche (salvo las térmicas) y las nubes las bloquean.
- Activos: emiten su propia energía y miden el retorno. Funcionan de noche y, según la longitud de onda, atraviesan nubes. LiDAR y radar son activos.

El espectro electromagnético

Los sensores no se limitan a lo que ve el ojo humano. Cada región del espectro revela información distinta:

- Visible (azul, verde, rojo): lo que vemos; útil para color natural y para el análisis de vegetación en el rojo.
- Infrarrojo cercano (NIR): la vegetación sana lo refleja intensamente. Es la banda estrella para agricultura.
- Infrarrojo de onda corta (SWIR): sensible a la humedad del suelo y la vegetación; sirve para detectar incendios y minerales.
- Infrarrojo térmico: mide temperatura de superficie; islas de calor urbanas, estrés hídrico.
- Microondas: las usa el radar; atraviesan nubes y lluvia.

Firmas espectrales

Cada material refleja distinta proporción de energía en cada longitud de onda. Ese patrón es su firma espectral y funciona como una huella digital:

- Vegetación sana: refleja poco en el rojo (la clorofila lo absorbe) y muchísimo en el NIR.
- Vegetación estresada o seca: sube la reflectancia en el rojo y baja en el NIR.
- Agua: absorbe casi todo el NIR, por eso se ve muy oscura en esa banda.
- Suelo desnudo: reflectancia moderada y creciente hacia el infrarrojo.

Clasificar cobertura terrestre consiste, en el fondo, en distinguir firmas espectrales.

Las cuatro resoluciones

Toda imagen satelital se describe con cuatro resoluciones, y siempre hay que negociar entre ellas:

- Resolución espacial: el tamaño del píxel en el terreno. 30 m en Landsat, 10 m en Sentinel-2, 250 m en MODIS, menos de 1 m en satélites comerciales.

- Resolución espectral: cuántas bandas y qué tan angostas. Un sensor hiperespectral tiene cientos de bandas estrechas.
- Resolución temporal: cada cuánto el satélite vuelve a pasar por el mismo lugar. 16 días en Landsat, 5 días en Sentinel-2, diaria en MODIS.
- Resolución radiométrica: cuántos niveles de intensidad distingue el sensor. 8 bits son 256 niveles; 12 bits son 4096.

El compromiso es físico: más resolución espacial implica menos área cubierta por escena y, en general, menor frecuencia de revisita.

Misiones y sensores frecuentes

- Landsat (NASA/USGS): serie continua desde 1972, 30 m, gratuita. Insustituible para estudios de cambio a largo plazo.
- Sentinel-2 (ESA): 10–20 m, revisita de 5 días, 13 bandas, gratuita. El caballo de batalla actual.
- Sentinel-1 (ESA): radar SAR, ve a través de nubes; ideal para inundaciones y zonas tropicales.
- MODIS y VIIRS: resolución gruesa pero cobertura diaria; clima, incendios, monitoreo global.
- PlanetScope, WorldView: comerciales, resolución de metros o submétrica, revisita casi diaria.

LiDAR

LiDAR emite pulsos láser y mide el tiempo que tardan en volver, produciendo una nube de puntos en tres dimensiones.

- Genera modelos de elevación muy precisos.

- Distingue el modelo digital de superficie (DSM, incluye árboles y edificios) del modelo digital del terreno (DTM, el suelo desnudo).
- Restando ambos se obtiene la altura de la vegetación o de las construcciones.
- Se usa en forestal, arqueología (encontró ciudades mayas bajo el dosel del Petén), urbanismo y modelado de inundaciones.

Radar y SAR

El radar de apertura sintética (SAR) emite microondas y mide su retorno.

- Atraviesa nubes, humo y lluvia, y funciona de noche.
- Es la única opción confiable en regiones tropicales con nubosidad permanente.
- Mide rugosidad y humedad más que color; su interpretación es distinta a la de una imagen óptica.
- La interferometría (InSAR) detecta deformaciones del terreno de milímetros: hundimientos, deslizamientos, actividad volcánica.

Preprocesamiento de imágenes satelitales

Una imagen recién bajada no está lista para analizarse. Los pasos habituales:

- Corrección radiométrica: convertir los números crudos del sensor a radiancia y luego a reflectancia física, para que dos fechas sean comparables.
- Corrección atmosférica: descontar el efecto de la atmósfera. La diferencia entre un producto de nivel 1 (reflectancia en el tope de la atmósfera) y nivel 2 (reflectancia de superficie).
- Corrección geométrica y ortorrectificación: alinear la imagen con el terreno real, corrigiendo la distorsión que introduce el relieve.

- Enmascarado de nubes: identificar y descartar píxeles con nube y sombra de nube, usando la banda de calidad que traen los productos.
- Composición temporal: combinar varias fechas tomando, por ejemplo, la mediana de cada píxel para obtener una imagen sin nubes.
- Mosaico y recorte: unir escenas vecinas y recortar al área de interés.
- Remuestreo: llevar capas de distinta resolución a una cuadrícula común.
Vecino más cercano para datos categóricos, bilineal o cúbico para continuos.

Igual que en cualquier proyecto de ciencia de datos, esta etapa consume la mayor parte del tiempo.

Índices espectrales

Un índice espectral combina bandas para resaltar un fenómeno concreto. Casi siempre es una diferencia normalizada, que produce valores acotados y menos sensibles a la iluminación.

NDVI

El índice de vegetación de diferencia normalizada es el más usado en teledetección:

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

- Va de -1 a 1 .
- Valores negativos: agua, nieve, nubes.
- Cerca de 0 : suelo desnudo, roca, área construida.
- De 0.2 a 0.5 : vegetación escasa o pastizales.
- Arriba de 0.6 : vegetación densa y sana.

Su lógica es directa: la vegetación sana absorbe rojo y refleja NIR, así que la diferencia se dispara.

Otros índices frecuentes

- NDWI: usa verde y NIR para detectar cuerpos de agua y humedad.
- NDBI: usa SWIR y NIR para resaltar áreas construidas.
- EVI: una versión del NDVI corregida para el efecto del suelo y la atmósfera; no se satura en vegetación muy densa.
- NBR: usa NIR y SWIR para medir severidad de incendios; la diferencia antes y después del fuego cuantifica el área quemada.

Para un proyecto de ciencia de datos, estos índices son variables predictoras listas para usar.

Sistemas de Información Geográfica

Un SIG es un sistema para capturar, almacenar, consultar, analizar y visualizar datos con referencia geográfica. No es solo un programa de mapas: es una base de datos donde la ubicación es un campo de primera clase.

Los cinco componentes

- Hardware: computadoras, GPS, servidores.
- Software: QGIS, ArcGIS, PostGIS, librerías de Python.
- Datos: la parte más cara y más valiosa del sistema.
- Personas: quienes diseñan el análisis e interpretan los resultados.
- Métodos: los procedimientos y flujos de trabajo documentados.

El modelo de capas

Un SIG organiza el territorio en capas temáticas superpuestas sobre la misma zona: hidrografía, vías, uso del suelo, población, parcelas. Cada capa es independiente y se combina con las demás cuando el análisis lo pide. Esa separación es lo que permite responder preguntas que ninguna capa contesta por sí sola.

Operaciones espaciales fundamentales

- Consulta por atributo: seleccionar registros por sus valores, como en SQL.
- Consulta espacial: seleccionar por relación geométrica. Contiene, intersecta, está dentro de, toca, cruza.
- Buffer: generar un área a cierta distancia de una geometría. La zona de 500 m alrededor de un río.
- Overlay: combinar capas. Intersección, unión, diferencia, recorte.
- Clip: recortar una capa con la forma de otra.
- Disolución: fusionar geometrías vecinas que comparten un atributo, por ejemplo unir municipios en departamentos.
- Unión espacial (spatial join): pegarle a cada punto los atributos del polígono que lo contiene. Es la operación más común al preparar datos.
- Geocodificación: convertir una dirección de texto en coordenadas. La inversa es la geocodificación reversa.
- Análisis de redes: rutas más cortas, áreas de servicio, asignación de demanda a instalaciones.

Álgebra de mapas

Con rásteres, el análisis se vuelve aritmética entre matrices:

- Operaciones locales: celda por celda, entre una o varias capas. Así se calcula el NDVI.
- Operaciones focales: el valor de una celda se calcula a partir de su vecindad. Suavizados, pendientes, detección de bordes.
- Operaciones zonales: se resume una capa dentro de las zonas que define otra. El NDVI promedio por municipio.
- Operaciones globales: usan todo el ráster. Distancia al elemento más cercano, cuencas visuales.

Un ejemplo de análisis encadenado

Dónde ubicar un nuevo centro de salud:

1. Partir de la capa de población por sector censal.
2. Generar buffers de 30 minutos alrededor de los centros existentes con análisis de redes.
3. Restar esas áreas para quedarse con la población no cubierta.
4. Intersectar con la capa de vías para exigir accesibilidad.
5. Excluir zonas de pendiente alta y áreas protegidas.
6. Ordenar los polígonos candidatos por población servida.

Ninguna de esas capas responde la pregunta sola; la respuesta sale de combinarlas.

Análisis y estadística espacial

Autocorrelación espacial

Si los valores altos tienden a estar cerca de otros valores altos, hay autocorrelación espacial positiva. Si los valores altos se rodean de bajos, es negativa. El índice de Moran la resume en un número:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

donde w_{ij} es el peso que expresa qué tan vecinas son las unidades i y j .

- Cerca de 1: patrón agrupado.
- Cerca de 0: patrón aleatorio.
- Cerca de -1 : patrón disperso, como un tablero de ajedrez.

Esto importa porque casi toda la estadística clásica supone observaciones independientes. Con autocorrelación, los errores estándar salen subestimados y se declaran significativos efectos que no lo son.

Detección de agrupamientos

- Getis-Ord G_i^* : identifica puntos calientes y fríos estadísticamente significativos, no solo valores altos.
- Moran local (LISA): descompone el índice global y señala dónde está cada tipo de agrupamiento.
- DBSCAN: agrupa puntos por densidad sin fijar de antemano el número de grupos, y marca el ruido. Muy usado con datos de GPS.
- Estimación de densidad de kernel: convierte una nube de puntos en una superficie continua de intensidad.

Interpolación

A partir de mediciones en pocos puntos se estima el valor en todo el espacio:

- IDW (distancia inversa ponderada): cada punto conocido influye en proporción inversa a su distancia. Simple e intuitivo.
- Kriging: modela explícitamente la estructura de la autocorrelación mediante un variograma y produce, además de la estimación, su incertidumbre.
- Vecino natural y splines: alternativas que suavizan de forma distinta.

Trampas del análisis espacial

- MAUP (problema de la unidad de área modificable): los resultados cambian según el tamaño y el trazo de las unidades espaciales. Los mismos datos agregados por departamento o por municipio pueden dar conclusiones opuestas.
- Falacia ecológica: concluir algo sobre individuos a partir de agregados por zona. Que un municipio tenga alto ingreso y alta mortalidad no implica que los ricos mueran más.
- Efecto de borde: las unidades del borde del área de estudio tienen menos vecinos, lo que sesga los estadísticos de vecindad.
- Sesgo de cobertura: las zonas pobres o rurales suelen estar peor mapeadas, y esa ausencia se confunde con ausencia del fenómeno.

Machine learning con datos geoespaciales

Clasificación de cobertura terrestre

Es la tarea clásica: asignar a cada píxel una clase de uso del suelo (bosque, agua, cultivo, urbano, suelo desnudo).

- Supervisada: se etiquetan a mano regiones de entrenamiento y se ajusta un clasificador. Random Forest y gradient boosting funcionan muy bien con

bandas e índices como variables.

- No supervisada: se agrupan los píxeles por similitud espectral con k-means y luego se interpreta cada grupo.
- Las variables son las bandas, los índices espectrales, la textura y, muy útil, la serie temporal completa del año: los cultivos se distinguen por su fenología, no por una sola fecha.

Deep learning sobre imágenes satelitales

Todo lo visto en la lección de reconocimiento de imágenes aplica aquí, con matices:

- Las CNN clasifican parches de imagen y detectan objetos: edificios, embarcaciones, piscinas, deforestación.
- U-Net y otras arquitecturas encoder-decoder hacen segmentación semántica píxel a píxel; es el estándar para mapear cobertura y extraer huellas de edificios.
- La diferencia con las fotos comunes: hay más de tres canales, los valores no están en $[0, 255]$, y el contexto geográfico y temporal importa.

Variables espaciales derivadas

El feature engineering espacial suele mejorar más un modelo que cambiar de algoritmo:

- Distancia a la carretera, al río, al mercado o al centro urbano más cercano.
- Densidad de puntos de interés en un radio dado.
- Estadísticos de la vecindad: promedio del vecindario, pendiente, elevación.
- Variables espectrales agregadas por zona: NDVI promedio y su variación en el tiempo.

Validación con datos espaciales

Aquí está el error más frecuente y más caro:

- Una partición aleatoria de entrenamiento y prueba deja puntos vecinos en ambos conjuntos. Como los vecinos se parecen, el modelo aparenta un desempeño excelente que no se sostiene en territorio nuevo. Es fuga de información espacial.
- La solución es la validación cruzada espacial: separar por bloques geográficos, por regiones o por conglomerados, de modo que el conjunto de prueba esté espacialmente alejado del de entrenamiento.
- Si además hay dimensión temporal, la partición debe respetar el tiempo: entrenar con años pasados y evaluar con años futuros.

Herramientas

Python

- GeoPandas: extiende pandas con una columna de geometría. Permite unión espacial, buffers, reproyección y overlays con sintaxis de DataFrame.
- Shapely: la geometría de fondo; operaciones sobre puntos, líneas y polígonos.
- Fiona y pyproj: lectura y escritura de formatos vectoriales, y transformación entre CRS.
- Rasterio: leer y escribir rásteres como arreglos de NumPy.
- rioxarray y xarray: cubos de datos multidimensionales, ideales para series de imágenes.
- GDAL/OGR: la librería que está debajo de casi todo el ecosistema.
- Folium, Kepler.gl y plotly: mapas interactivos.

- PySAL: estadística espacial, incluido el índice de Moran y Getis-Ord.
- Google Earth Engine: catálogo planetario de imágenes con cómputo en la nube; evita descargar terabytes.

Escritorio y bases de datos

- QGIS: SIG de escritorio libre y completo, con consola de Python.
- ArcGIS Pro: la alternativa comercial dominante en instituciones.
- PostGIS: extensión espacial de PostgreSQL. Permite consultas SQL con predicados geométricos e índices espaciales; es la forma seria de guardar datos geoespaciales.
- SpatiaLite: el equivalente ligero sobre SQLite.

Visualización de datos geoespaciales

El mapa es el resultado y también la herramienta de exploración. Unas reglas que evitan la mayoría de los mapas engañosos:

- Normaliza siempre. Un mapa de conteos absolutos por municipio es, en el fondo, un mapa de dónde vive la gente. Usa tasas, densidades o porcentajes.
- Elige bien la clasificación. Intervalos iguales, cuantiles y cortes naturales (Jenks) producen mapas distintos con los mismos datos. Declara cuál usaste.
- Usa paletas adecuadas: secuenciales para magnitudes, divergentes cuando hay un punto medio significativo, cualitativas para categorías. Nunca arcoíris para datos continuos.
- Cuidado con la proyección. Web Mercator infla las latitudes altas; para comparar áreas usa una proyección equivalente.
- Considera cartogramas o mapas de hexágonos cuando el tamaño de las unidades distorsione la lectura.

- Un mapa de calor de puntos muestra densidad, no intensidad del fenómeno; se presta a confusión.

Tipos de mapa más usados: coropletas, símbolos proporcionales, puntos, densidad de kernel, isolíneas y mapas de flujo.

Aplicaciones

- Agricultura de precisión: monitoreo de cultivos con NDVI, estimación de rendimiento, detección temprana de estrés hídrico, aplicación variable de insumos.
- Gestión de desastres: mapeo de inundaciones con SAR, evaluación de daños tras un sismo, modelado de flujos de lahares y zonas de riesgo volcánico.
- Ambiente y clima: deforestación, cambio de cobertura, islas de calor urbanas, calidad del aire, monitoreo de glaciares.
- Salud pública: mapeo de brotes, análisis de accesibilidad a servicios de salud, identificación de zonas desatendidas.
- Planificación urbana: crecimiento de la mancha urbana, catastro, transporte público, ubicación óptima de servicios.
- Logística y retail: optimización de rutas, definición de áreas de cobertura, selección de sitios para nuevas tiendas.
- Seguridad alimentaria: sistemas de alerta temprana que combinan lluvia, NDVI y precios.

En Guatemala son casos vivos el monitoreo de la deforestación en el Petén, la vigilancia del complejo volcánico de Fuego y Pacaya, el mapeo de riesgo por deslizamientos en la cuenca del Amatitlán y el LiDAR arqueológico que reveló miles de estructuras mayas bajo la selva.

Consideraciones éticas

La ubicación es un dato personal particularmente sensible.

- Reidentificación: cuatro puntos espacio-temporales bastan para identificar de forma única a la mayoría de las personas en un conjunto de datos de movilidad. Anonimizar quitando el nombre no es suficiente.
- Agregación y desplazamiento: publicar datos sensibles exige agregar a unidades suficientemente grandes o desplazar las coordenadas de forma controlada. Publicar la ubicación exacta de un caso de enfermedad o de una vivienda es revelar a una persona.
- Vigilancia: el rastreo masivo de ubicación plantea problemas de derechos que la técnica no resuelve.
- Sesgo de datos: los mapas colaborativos están mucho más completos en zonas ricas y urbanas. Un modelo entrenado sobre esa cobertura desigual perpetúa la desigualdad.
- Poder del mapa: todo mapa es una decisión sobre qué se muestra y qué se omite. Las fronteras, los nombres y las categorías de un mapa nunca son neutrales.
- Consentimiento: la ubicación se recolecta con frecuencia de forma pasiva desde teléfonos y aplicaciones, sin que el usuario entienda el alcance de lo que cede.

Cierre

- Un dato geoespacial une ubicación, tiempo y atributos, y su dimensión espacial cambia las reglas del análisis.

- Toda coordenada vive dentro de un datum y un CRS; combinar capas sin verificarlos es la primera fuente de errores.
- Vectorial para objetos discretos, ráster para fenómenos continuos; en la práctica se combinan.
- Los sensores remotos entregan una vista periódica y consistente de todo el planeta, con un compromiso permanente entre resolución espacial, espectral, temporal y radiométrica.
- Los índices espectrales convierten bandas crudas en variables directamente útiles para un modelo.
- Un SIG no es un visor de mapas: es un motor de consulta y análisis donde la ubicación es un campo de primera clase.
- La autocorrelación espacial rompe el supuesto de independencia; la validación cruzada espacial no es opcional.
- La ubicación identifica personas: el análisis geoespacial exige cuidado ético desde el diseño.