



Documento de trabajo N° 1/2024

**Estimación de tiempos de viaje en Bolivia con
datos geoespaciales**

Por:

Gabriel Eduardo Weise Andrade

Estimación de tiempos de viaje en Bolivia con datos geoespaciales

Por:

Gabriel Eduardo Weise Andrade^{1*}

La Paz, Febrero 2024

Resumen: Este trabajo presenta una metodología para estimar el tiempo de viaje desde cualquier punto geográfico a un conjunto de puntos de interés. Integra datos de diversas fuentes, incluyendo carreteras, senderos, terreno y pendiente, para crear una capa de costo que refleja el tiempo necesario para atravesar cada píxel del mapa. Utiliza una implementación ponderada del algoritmo de Dijkstra para calcular la ruta de menor costo y estimar el tiempo de viaje. Los resultados se calibran y validan utilizando datos de la API de Google Maps, mostrando una correlación significativa entre ambos conjuntos de datos.

Palabras clave: Accesibilidad, tiempo de viaje, Bolivia

Clasificación JEL: R11; O18; R40

^{1*} psiport@gmail.com

1. Introducción

El acceso a diferentes lugares y servicios es una condición necesaria para el desarrollo social y económico de un país. En Bolivia, la topografía diversa y la infraestructura vial limitada presentan desafíos para la accesibilidad en muchas regiones. Este trabajo busca desarrollar una descripción del costo de movimiento que facilite la evaluación del acceso a centros económicos y recursos básicos como salud y educación.

Esta descripción se basa en la integración de datos de diversas fuentes:

- **Carreteras:** datos del Instituto Geográfico Militar (IGM) [A], OpenStreetMap (OSM) [B] y RoadDetection de Microsoft [C].
- **Senderos:** datos de Open Mapping por Facebook (OMF) [D].
- **Terreno:** datos de Digital Elevation Model (DEM) de SRTM [F] y WorldCover de la ESA [E].

El artículo se organiza de la siguiente manera:

- **Metodología:** Se describe en detalle el proceso de integración de datos, la construcción de la capa de costos de movimiento y el cálculo del tiempo de viaje.
- **Resultados:** Se presenta un mapa de calor que muestra la estimación del tiempo de viaje a las capitales departamentales.
- **Validación:** Se describe la metodología utilizada para validar los resultados obtenidos.
- **Limitaciones:** Se discuten las limitaciones del método y las posibles mejoras para futuras investigaciones.

2. Metodología

Para integrar las fuentes de datos en un solo conjunto, se inició con una fase de pre-procesamiento, es decir, limpieza de datos para eliminar errores e inconsistencias y transformarlos para que sean compatibles. Este procedimiento tiene el objetivo de armonizar los conjuntos de datos para una representación coherente del paisaje en diversos terrenos y redes de infraestructura.

Para evaluar el impacto de cada conjunto de datos en la capa final, se establecieron cuatro métricas: cobertura de la población (2012) a 1 km de la red vial, cobertura del territorio nacional a 1 km de la red vial, cantidad de tramos con datos de velocidad y cantidad de tramos eliminados durante la limpieza de datos. Estas métricas permiten analizar la contribución individual de cada fuente de datos a la representación final del paisaje y la infraestructura vial.

2.1. Desarrollo de la capa integrada de costo de movimiento

Se denomina como "costo" de viaje al tiempo estimado necesario para trasladarse a través de un área en función de diversos factores que afectan la velocidad de desplazamiento. Su asignación se basa en identificar y analizar los factores que influyen en esta velocidad, especialmente en relación con la red de carreteras y senderos que cubren la región. Estos factores incluyen el tipo y curvaturas (sinuosidad) de la carretera y el tipo y pendiente del terreno; el resultado es un mapa donde cada píxel (100x100m) cuantifica el tiempo necesario para recorrerlo.

2.1.1. Carreteras

En el análisis del tiempo de viaje, se comienza con las carreteras, la columna vertebral de la mayoría de los sistemas de transporte. Para reflejar la variedad de escenarios de viaje del mundo real, se desarrolla una capa que busca reflejar la extensión y complicaciones de la red de carreteras.

2.1.1.1. Geometría de la carretera: Razón de sinuosidad

Al evaluar los tiempos de viaje, es importante reconocer que las velocidades reales a menudo difieren de los límites identificados por varios factores, entre ellos, el material de la carretera, dato no disponible para la mayoría de los tramos en Bolivia. Para tomar en cuenta este efecto en la diferencia en la velocidad, la metodología emplea un enfoque indirecto establecido como un factor de corrección basado en la razón de sinuosidad² [1, 5].

La razón de sinuosidad se calcula dividiendo la longitud real de la carretera con la distancia en línea recta entre sus extremos. Esta métrica adquiere mayor valor en regiones con topografía diversa, donde las carreteras pueden tener que rodear obstáculos naturales como colinas o cuerpos de agua.

2.1.1.2. Conjuntos de datos

*** Datos Oficiales de Bolivia: Instituto Geográfico Militar (IGM)**

El repositorio oficial de datos sobre carreteras para Bolivia proviene del Instituto Geográfico Militar (IGM) y se constituye en la base fundamental de este trabajo. Este conjunto de datos proporciona la geometría precisa de la red vial fundamental del país, ofreciendo una imagen clara de las principales arterias de transporte.

² Las curvas que se suceden unas a otras conforman la sinuosidad de un camino.

Características Clave

- **Tipo:** IGM clasifica los caminos en primario, secundario y comunal.
- **Material:** Este conjunto de datos incluye material de construcción de las carreteras indexadas: asfalto, empedrado y tierra.

Un límite importante es la ausencia de información sobre velocidad, que es crucial para una estimación precisa del tiempo de viaje. Si bien, cubre carreteras primarias y secundarias y una amplia población, el conjunto de datos es menos detallado para las carreteras comunales y residenciales, lo que puede generar vacíos en la representación de la red vial.

* Datos Abiertos de Red de Carreteras: OpenStreetMap (OSM)

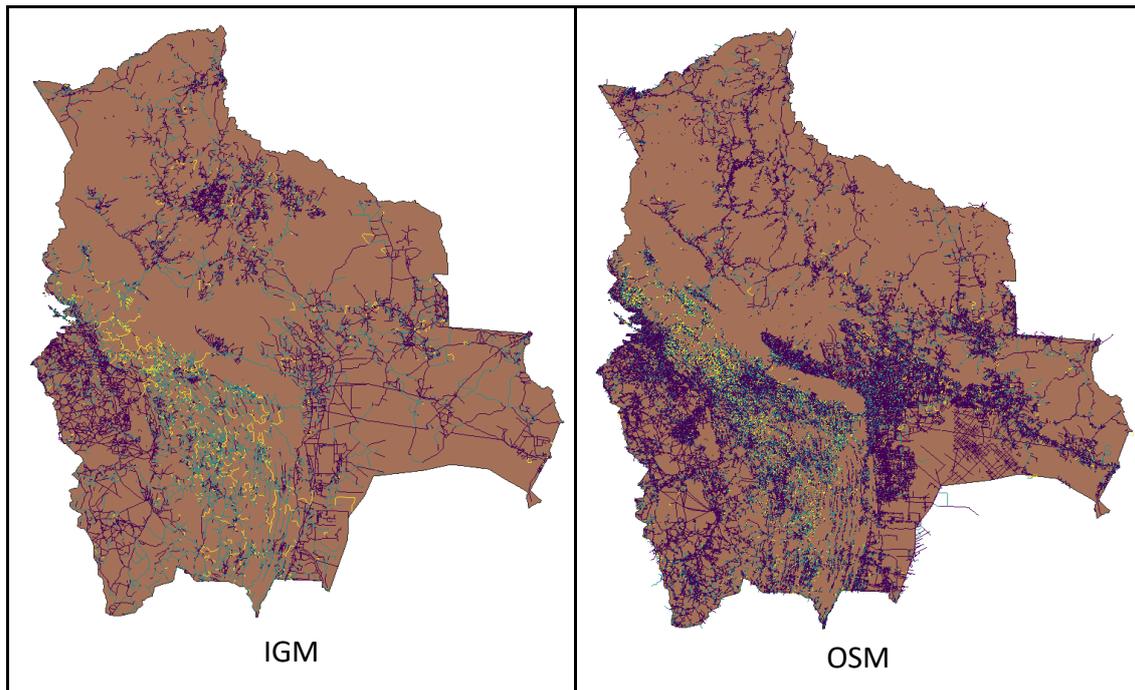
OpenStreetMap (OSM) es un conjunto de datos abiertos creado y sostenido por una comunidad de usuarios, que permite una perspectiva diversa y a menudo actualizada sobre las condiciones y tipos de carreteras. Este conjunto de datos proporciona una mayor cobertura de caminos secundarios, terciarios y vecinales.

Características Clave

- **Tipo:** OSM usa una clasificación más desagregada de los varios tipos de carretera, incluyendo, autopistas, troncales, primarias, secundarias, terciarias, vecinales, etc.
- **Velocidad:** Una característica notable de OSM es la inclusión de estimaciones de velocidades máximas para los tramos de carretera, proporcionadas por los usuarios. Estos datos son un insumo clave para la estimación del tiempo de viaje en Bolivia.

A continuación, se presenta el mapa de los datos obtenidos de ambas fuentes y su nivel de sinuosidad.

Mapa 1. Nivel de sinuosidad de las carreteras de Bolivia



Cada mapa resalta el nivel de sinuosidad de las carreteras identificadas en los datos de IGM y de OSM. El color violeta, correspondiente a carreteras con una sinuosidad en el rango de $[1, 1.2)$, indica tramos con una baja sinuosidad, es decir, trayectos relativamente rectos y directos. Las carreteras en verde tienen una sinuosidad con rango $[1.2, 1.8)$ que representa tramos con una moderada sinuosidad y aquellas en amarillo presentan una sinuosidad en el rango $[1.8, 10)$, señala tramos altamente sinuosos, caracterizados por curvas significativas y un recorrido más serpenteante.

Para armonizar los datos de OSM con el conjunto de datos de IGM se utiliza un proceso de geo intersección. Este proceso tiene dos etapas, la primera es la identificación y eliminación de duplicados o tramos de carretera que se superponen entre ambos conjuntos de datos. La segunda, compatibiliza los tipos de carretera, mapeando el tipo/material de IGM con el tipo de OSM entre las superposiciones identificadas.

Sin embargo, y a pesar de su riqueza, los datos del OSM presentan límites en su eficiencia. Depender de una comunidad de voluntarios también implica que los datos completos son escasos, especialmente en zonas remotas. De esta manera, una cantidad significativa de carreteras no están clasificadas o carecen de velocidades, esto representa un nuevo desafío para la estimación del tiempo de viaje en este trabajo.

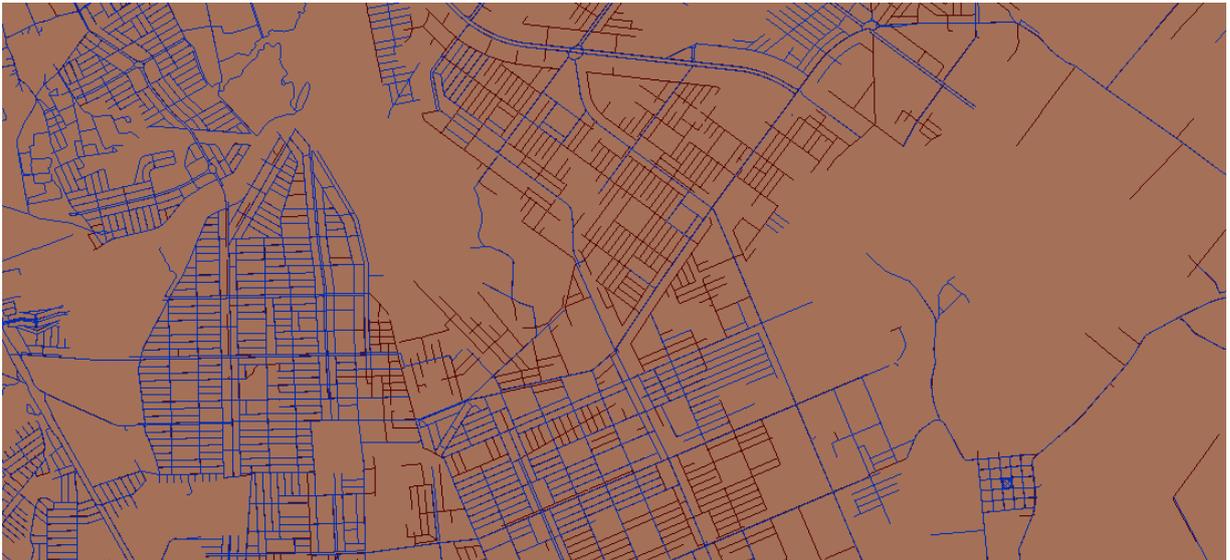
*** RoadDetection de Microsoft:**

Este conjunto de datos construido usando inteligencia artificial, entrenada con datos extensos de OSM, Bing Maps, e imágenes satelitales de Maxar (2020-2022), ofrece una actualización de amplio alcance de las redes de carreteras existentes.

Características Clave

- **Amplia cobertura:** Proporciona una amplia cobertura de las carreteras, incluida información actualizada de zonas comunitarias y residenciales nuevas o menos frecuentadas.

Mapa 2. Comparación de datos de OSM y RoadDetection



En azul OSM, en rojo RoadDetection

Para abordar la falta de información específica sobre el tipo de carretera en este conjunto de datos, se empleó una técnica de interpolación espacial, específicamente una forma de propagación de etiquetas basada en proximidad. Este proceso implica analizar las intersecciones con los datos de OSM, donde los tipos de carretera (etiquetas) de OSM se transfieren a los tramos de carretera correspondientes en el conjunto de datos de Microsoft. A partir de aquí, se propagan las etiquetas de forma recursiva hasta que no sean posibles más ediciones, lo que garantiza un enriquecimiento exhaustivo de estos datos.

2.1.1.3. Integración

* **Completando Velocidad**

Aproximadamente el 98% de los caminos carecen de información de velocidad. Para abordar este problema, se emplea un método de imputación basado en la velocidad promedio ponderada por la longitud de cada tramo, utilizando la información disponible sobre el tipo de carretera y la sinuosidad. Luego se asignan estos valores a tramos con características similares.

* Rasterizado

Para integrar los conjuntos de datos preprocesados, se genera el ráster de costos de movimiento en carreteras. La secuencia en la que se rasterizan las capas garantiza una representación coherente de la velocidad necesaria para transitar cada píxel. Se sigue una jerarquía específica al agregar las capas, estableciendo un orden de prioridad que determina la influencia de cada conjunto de datos en la formación del resultado final:

1ro. Caminos Residenciales de OSM: Estos caminos tienen mayor prioridad debido al impacto significativo en las velocidades de viaje en áreas urbanas y suburbanas. Las carreteras residenciales suelen tener límites de velocidad más bajos y mayor tráfico, que son factores cruciales en la estimación del tiempo de viaje.

2do. Carreteras de IGM: Estos datos proveen información de carreteras primarias que tienen límites de velocidad más altos.

3er. Datos de OSM: Introducen una amplia gama de tipos de carreteras y sus condiciones.

4to. Datos de MS RoadDetection: Estos datos completan cualquier vacío en las capas anteriores.

Tabla 1. Integración de los conjuntos de datos

FUENTE	% Cobertura de población a 1km de la carretera.	% Cobertura de área a 1 km de carretera.	% Datos de velocidad	% Tramos eliminados
IGM	60.7	11.0	0	0
OSM	94.9	20.2	2.2	3
MS RoadDetection	92.3	8.5	-	96
Integrado	96.7	24.4	-	-

La tabla representa las diferencias en la cobertura de datos en las cuatro fuentes de datos utilizada.

2.1.2. Más allá de las carreteras: Senderos y terreno abierto

Para comprender plenamente la movilidad en diversas geografías hay que considerar también los caminos más accidentados y el terreno abierto. Esta sección del análisis incluye caminos de tierra, senderos y otras rutas que, aunque no forman parte de la red principal de carreteras, son accesos cotidianos en el movimiento en áreas rurales o menos desarrolladas. Se examina la capacidad de transitar por estos terrenos, que adquiere mayor relevancia para regiones donde las rutas establecidas son escasas o

inexistentes, alcanzando una comprensión más completa de los tiempos de viaje y las opciones de accesibilidad en diferentes entornos.

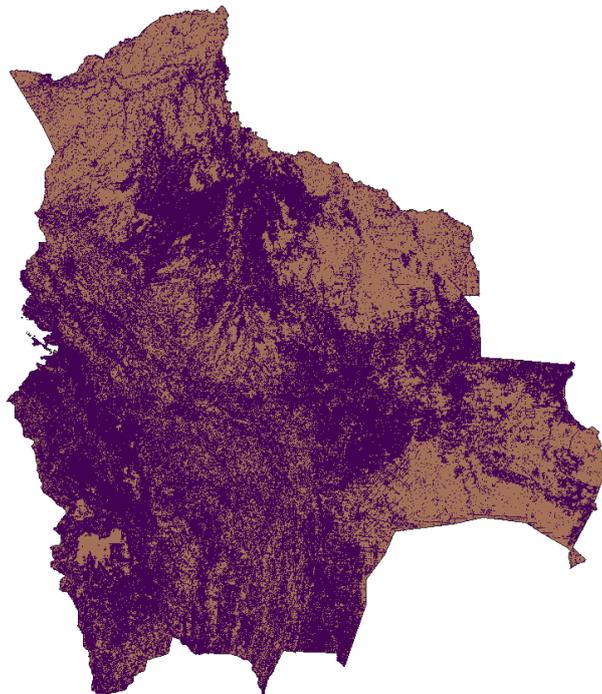
2.1.2.1. Open Mapping por Facebook (OMF)

Este conjunto de datos es fundamental para mapear rutas de viaje fuera de carretera. OMF se desarrolla usando algoritmos de aprendizaje automático entrenados con una combinación de datos OSM e imágenes satelitales de alta resolución, este diseño ofrece un nivel de detalle sin precedentes en la identificación de senderos y caminos comunales que a menudo se pasan por alto en los conjuntos de datos convencionales.

Sin embargo, un gran problema es la incidencia de falsos positivos, características naturales identificadas incorrectamente como caminos: ríos, laderas de montañas y cicatrices en el terreno. Para superarlo se aplican la siguiente solución:

- Cuerpos de agua: Se usan datos de ríos primarios [G] y secundarios [H] del IGM para identificar y eliminar imprecisiones y clasificaciones erróneas.
- Factor de penalización: Asegura que el algoritmo de asignación de costos considere a características naturales como los ríos como obstáculos significativos, favoreciendo así las rutas que los eviten en favor de caminos u otras vías de transporte planificadas.

Mapa 3. Cobertura de rutas/senderos de OMF



Este mapa refleja la amplia cobertura de rutas/senderos que reflejan los datos de OMF. El terreno cubierto se representa en color morado.

Tabla 2. Integración de los conjuntos de datos

FUENTE	% Cobertura de población a 1km de la carretera	% Cobertura de área a 1 km de carretera.	% Datos de velocidad	% Tramos eliminados
Integrado	96.7	24.4	-	-
OMF	82.0	34.0	-	2
Combinado	98.9	45.4	-	-

La tabla representa los resultados de la integración de los conjuntos de datos integrados previamente con los datos de OMF.

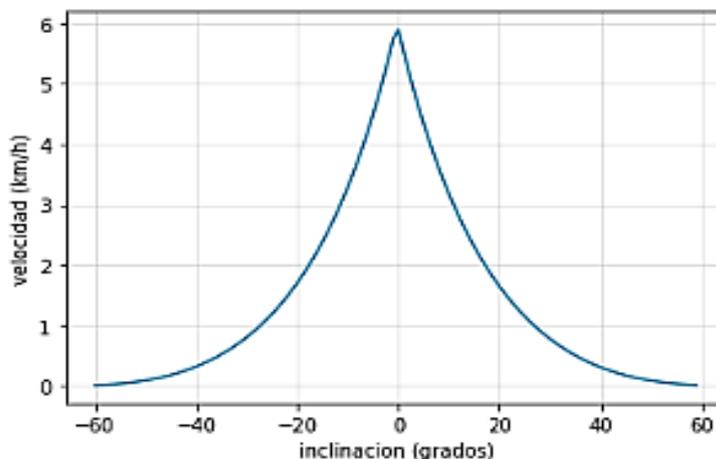
2.2. Caminos peatonales: Estimaciones basadas en elevación y terreno

La elevación afecta sustancialmente la velocidad al caminar y este efecto necesita cuantificación para un modelado realista. El enfoque implica ajustar la velocidad de los peatones según la pendiente del terreno utilizando una regla empírica conocida como Tobler's Hiking Function.

2.2.1. La función de Tobler

Esta regla [8] modela la relación entre la velocidad de caminata y la pendiente, evaluando el impacto de la elevación en el movimiento de los peatones. La función requiere calcular la pendiente promedio [3] de cada pixel a partir de datos de elevación (usando datos del `Digital Elevation Model` de SRTM). Estos datos de pendiente se aplican luego a la fórmula de Tobler para obtener la velocidad de caminata ajustada (base 5 km/h) para cada pixel según la pendiente.

Gráfico 1. Estimación de velocidad peatonal usando la función de Tobler



2.2.2. Estimación de velocidad peatonal según el terreno

Diferentes terrenos afectan la velocidad de viaje debido a sus características físicas únicas. Para abordar esto, los datos WorldCover de la ESA aportan clasificaciones detalladas de la cobertura del suelo, como bosques, pastizales y áreas urbanas.

A cada tipo de terreno se le asignan pesos en función del impacto que tienen en la velocidad de caminata derivadas de una investigación referencial [4]. Por ejemplo, el avance a través de bosques densos suele presentar una velocidad menor en contraste con la marcha a través de pastizales abiertos, debido a la mayor resistencia al movimiento y los obstáculos que se pueden presentar.

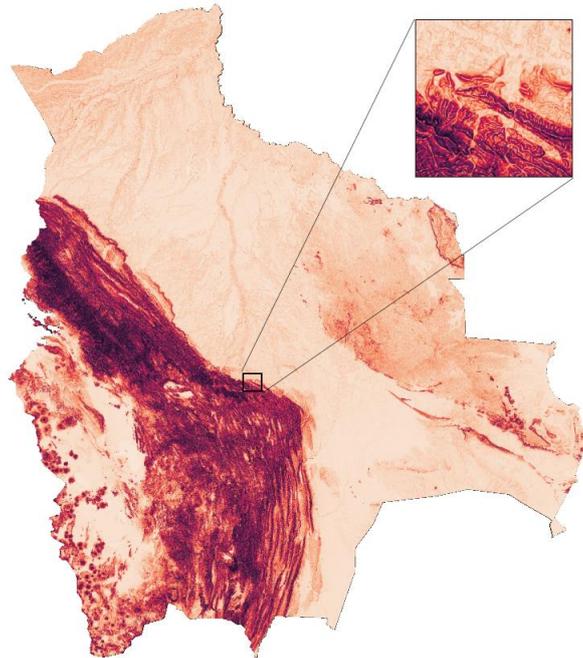
La velocidad estimada usando la función de Tobler se ajusta tomando en cuenta estos coeficientes. Esto permite una aproximación más ajustada a la realidad del desplazamiento humano en variadas condiciones ambientales.

Tabla 3: Tabla de ajustes de velocidad por tipo de terreno.

Terreno	Ajuste
Abierto	1
Pastizales	0.8
Arbustos	0.8
Tierras de cultivo	0.7
Cubierto de árboles	0.7
Líquenes/Musgo	0.7
Humedales	0.6
Nieve/Hielo	0.6

2.2.3. Integración

Mapa 4. Estimación de la velocidad de caminata ajustada por pendiente/terreno



La velocidad más lenta está representada en los tonos más oscuros.

- **Ajuste de pendiente para datos OMF:**
 - Comenzamos aplicando una velocidad base de 10 km/h a los tramos de senderos identificados en el conjunto de datos OMF.
 - Esta velocidad base se ajusta utilizando la función de Tobler para tener en cuenta el impacto de la pendiente en los tiempos de viaje. El resultado nos proporciona velocidades ajustadas a la pendiente para cada tramo.
- **Combinación de velocidades ajustadas:**
 - La capa final es una síntesis de ambos conjuntos de datos. Cuando los datos de OMF y los de terreno se superponen, los de OMF tienen prioridad.

2.3. Estimación de tiempo de viaje óptimo

Con la capa de costos ya construida, el trabajo se dirige al problema de cómo navegar eficientemente a través de cada pixel para encontrar la ruta más corta (en tiempo) hacia puntos de interés específicos. La principal dificultad está en determinar el camino óptimo que minimice el tiempo de viaje, considerando los diferentes costos de desplazamiento en cada pixel de la capa.

2.3.1. Algoritmo de Dijkstra para la Solución del Problema de la Ruta Más Corta

Este algoritmo [6], diseñado para operar en grafos ponderados, calcula caminos eficientes en redes complejas mediante la actualización iterativa de las distancias mínimas desde un nodo origen (punto de interés) hacia todos los demás, priorizando en cada paso aquellos nodos con la menor distancia acumulada.

- Se parte de un nodo origen, y se establece una distancia inicial de cero para este nodo, mientras que para todos los demás nodos la distancia inicial es infinita.
- Desde el nodo actual, se examinan todos los nodos vecinos no visitados y se calcula la distancia tentativa a cada uno sumando la distancia ya recorrida hasta el nodo actual y el costo de desplazamiento al nodo vecino.
- Una vez explorados todos los vecinos del nodo actual, se selecciona el nodo no visitado con la distancia más corta recorrida y este se convierte en el nuevo nodo actual.
- Este proceso se repite hasta que se han visitado todos los nodos.

La implementación del algoritmo aplicada en este trabajo es con la clase `mcp_geometric` de la librería `scikit-image` [7]. Se destacan dos aspectos de esta implementación:

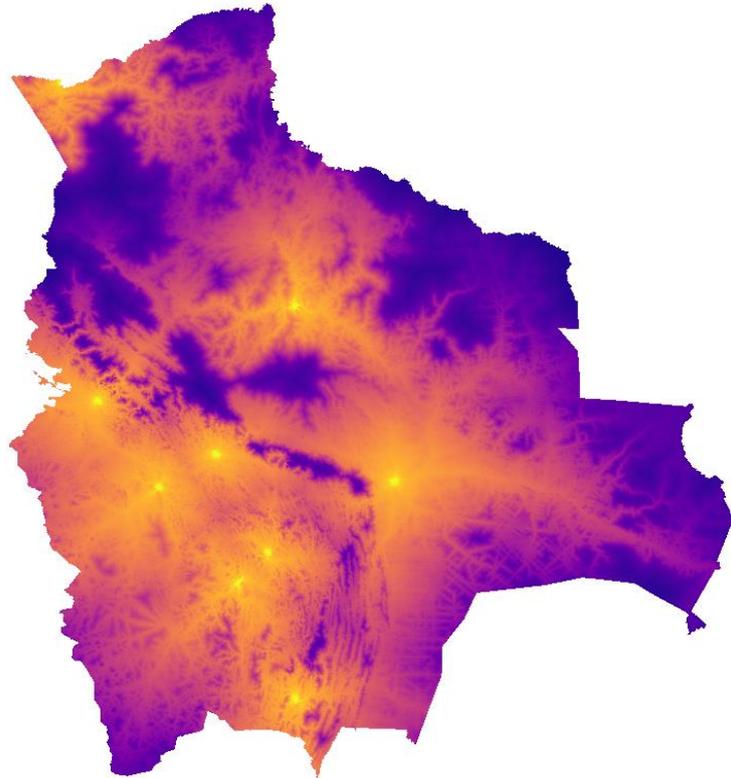
- La integración de la capa de costos como una matriz, donde cada elemento refleja el costo (en tiempo) asociado al desplazamiento a través de ese pixel en el mapa. Esta matriz no solo representa los costos, sino que también establece la conectividad entre pixels.
- Calcula el costo de movimiento entre pixels adyacentes, considerando tanto movimientos directos como diagonales:
 - Para un movimiento directo, el costo es el promedio de los costos de los pixels de inicio y destino.
 - Para un movimiento diagonal, se considera la distancia diagonal y se pondera el costo de los pixels involucrados.

Cualquier dato nulo resultado de una operación inválida es interpolado linealmente tomando en cuenta los valores de los pixels aledaños [2].

3. Resultado: Tiempo de viaje a las capitales departamentales

Todo el proceso detallado hasta este punto tiene como resultado la estimación del tiempo de viaje desde cualquier punto geográfico a cualquier capital departamental en Bolivia.

Mapa 5. Resultado de la estimación de tiempo de viaje o costo de accesibilidad³



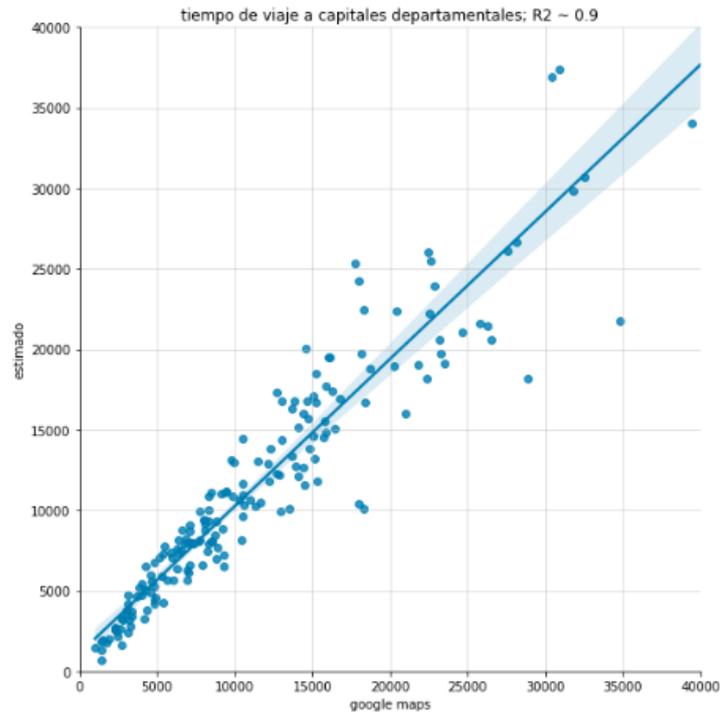
Los tonos del mapa indican la estimación del tiempo de viaje desde cualquier punto a la capital departamental más cercana.

4. Validación de los resultados obtenidos

Para realizar la validación de la estimación de tiempo de viaje establecidos en esta investigación se recurrió a los datos de la API de Google Maps, que proporciona funcionalidades de geocodificación y cálculo de rutas, incluyendo estimaciones de tiempos de viaje. Se normalizaron ambos conjuntos de datos a la misma escala temporal para que sean comparables y se aplicó una correlación entre ambos.

³ Mapa navegable en <https://psiport.users.earthengine.app/view/tiempo-viaje-sdsn>

Gráfico N° 2. Correlación entre tiempos de viaje Google Maps y los estimados para Bolivia.



El gráfico muestra una correlación entre los tiempos de viaje de Google Maps [1] (eje x) y los estimados en este trabajo (eje y), con un coeficiente R^2 cercano a 0.9. La dispersión a lo largo de la línea de regresión apunta a que, aunque hay variaciones, la mayoría de las estimaciones están alineadas con las expectativas generadas por una fuente externa y ampliamente utilizada.

5. Limitaciones

El trabajo presenta algunas limitaciones que deben considerarse al interpretar los resultados:

- La función de Tobler para estimar el efecto de la pendiente en el tiempo de viaje ni los ajustes del impacto de terreno toman en cuenta el gasto energético real de los individuos al desplazarse, tampoco se realizó un ajuste por el efecto de la altura en el oxígeno disponible.
- Los datos de OpenStreetMap (OMA), que se utilizan como base para algunas de las estimaciones, presentan niveles significativos de ruido en especial en las velocidades registradas.
- Se asume que la velocidad de movimiento en ríos primarios es equivalente a la velocidad de caminata, lo que no refleja condiciones reales de navegación, además el algoritmo de cálculo puede cambiar repetidamente entre modos de transporte (navegación, caminata).

- Solo se incluyen rutas nacionales en el cálculo del tiempo de viaje, omite rutas internacionales que, en algunos casos, pueden ofrecer alternativas más directas o rápidas.
- La resolución espacial del raster utilizado conduce a dos problemas específicos:
 - No captura curvas y giros bruscos, estos caminos pueden simplificarse de manera excesiva, resultando en representaciones mucho más rectas de lo que son en la realidad.
 - La resolución no es suficiente para realizar estimaciones de calidad en entornos urbanos, incluyendo la evaluación de diferentes modos de transporte, la dirección de las calles y obstáculos geográficos (medianos y cercanos a un camino) que impidan la transitabilidad.

Referencias

Boillat S, Sandoval Y, Patón L, Lerch L. (2016). *Modelo de accesibilidad a capitales departamentales y municipales de Bolivia*. La Paz: IIGEO / GeoBolivia

Burrough, P. A., & McDonnell, R. A. (1998). *Principles of geographical information systems*. Oxford University Press, USA.

C Horn, B.K.P., (1981). Hill shading and the reflectance map. *Proceedings of the IEEE* 69, 14–47.

De Gruchy, M., Caswell, E., and Edwards, J. (2017). *Velocity-Based Terrain Coefficients for Time-Based Models of Human Movement*. *Internet Archaeology*.

Hacar, Müslüm, Batuhan Kılıç, and Kadir Şahbaz. (2018). "Analyzing OpenStreetMap Road Data and Characterizing the Behavior of Contributors in Ankara, Turkey." *ISPRS International Journal of Geo-Information*, 7(10), 400.

Javaid, A. (2013). *Understanding Dijkstra's algorithm*. Available at SSRN 2340905.

Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu and the scikit-image contributors (2014). *scikit-image: Image processing in Python*.

Tobler W. (1993). Three Presentations on Geographical Analysis and Modelling. *National Center for Geographic Information and Analysis Technical Report*. 93(1).

Recursos

- A. Caminos IGM
<https://geo.gob.bo/download/?w=igm&l=CAMINOS Y VIA FERREA>
- B. OSM <https://download.geofabrik.de/south-america/bolivia.html> (2023/11)
- C. RoadDetection <https://github.com/microsoft/RoadDetections> (2023/03)
- D. OpenMapping <https://github.com/facebookmicrosites/Open-Mapping-At-Facebook/wiki/Available-Countries> (2022/12)
- E. WorldCover <https://esa-worldcover.org/en/data-access> (2021)
- F. DEM <https://registry.opendata.aws/terrain-tiles/> (2010)
- G. Rios Principales IGM
https://geo.gob.bo/download/?w=igm&l=RIOS_PRINCIPALES
- H. Rios Secundarios IGM
https://geo.gob.bo/download/?w=igm&l=RIOS_SECUNDARIOS
- I. Google Distance Matrix
<https://developers.google.com/maps/documentation/distance-matrix/intro>
- J. Lagos IGM <https://geo.gob.bo/download/?w=igm&l=LAGOS>