

PROBLEMAS QUE PRESENTAN LAS TRAZAS GNSS PROCEDENTES DE INFORMACIÓN GEOGRÁFICA VOLUNTARIA

PAULA GIL DE LA VEGA^{1*}, FRANCISCO JAVIER ARIZA LÓPEZ*, ANTONIO TOMÁS MOZAS CALVACHE*

*Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría
Universidad de Jaén. Escuela Politécnica Superior. Paraje de las Lagunillas, S/N. 23071, Jaén.
España

¹pgil@ujaen.es

RESUMEN

Las trazas GNSS de Información Geográfica Voluntaria (VGI) proceden de desplazamientos de conductores, ciclistas y viandantes voluntarios, en lugar de haber sido capturados por un levantamiento topográfico específico. La extensión, accesibilidad, disponibilidad y redundancia de estas trazas pueden permitir su uso en la producción cartográfica. Estas trazas son compartidas en diversos servidores de Internet y están afectadas por una serie de problemas que pueden limitar su uso. En este artículo se exponen los problemas más comunes que pueden afectar a las trazas GNSS procedentes de VGI y se muestran sobre un conjunto real de datos correspondientes a la "Vía Verde del Aceite" (Jaén).

Palabras clave: VGI, trazas GNSS.

PROBLEMS OF GNSS TRACKS FROM VOLUNTEERED GEOGRAPHIC INFORMATION

ABSTRACT

GNSS tracks from Volunteered Geographic Information (VGI) are tracks that come from movements of drivers, cyclists and walkers volunteers, instead of specific topographical surveys. These tracks are shared on the Internet by volunteers and are affected of a number of problems that can limit their use. In this article the most common problems that can affect GNSS tracks from VGI are shown and displayed on a real set corresponding to "Vía Verde del Aceite" (Jaén).

Keywords: VGI, GNSS tracks.

1. Introducción

La obtención de Información Geográfica (IG) ha sido tradicionalmente costosa debido a los procedimientos que se venían utilizando (levantamientos topográficos, fotogrametría, etc.) que requerían una cualificación técnica muy específica y unos equipos con un alto coste. Es por ello que la producción de la IG estaba limitada a ciertos profesionales (topógrafos, cartógrafos, geógrafos) formados en la materia y equipados con la instrumentación adecuada. La irrupción de las Tecnologías de la Información Geográfica (TIG), que incluyen, entre otras, la Cartografía digital, las Infraestructuras de Datos Espaciales (IDEs) y los Sistemas Globales de Navegación por Satélite (GNSS), han acercado la IG al público general, no sólo como consumidores, sino también como parte activa en la producción de la misma (Ruiz Almar, 2010).

La IG creada o recogida por voluntarios fue denominada "Información Geográfica Voluntaria" (VGI, del inglés "Volunteered Geographical Information") por Goodchild (2007). La VGI permite disponer de voluminosos conjuntos de IG en forma de puntos de interés, comentarios, imágenes (fotografías), trazas de navegación, etc. (Castelein, Grus, Crompvoets, & Bregt, 2010). La VGI presenta algunas ventajas frente a otro tipo de IG. En primer lugar, los costes de la información son muchos menores que los de otros procesos de producción de IG. Además, los voluntarios suelen ser personas que tienen un gran conocimiento del territorio del cual capturan la información, aportando un conocimiento local más experto. Por otro lado, el *universo de discurso*¹ es mucho más amplio que en la IG tradicional. Son los voluntarios quienes deciden qué información quieren capturar y compartir, y lo hacen atendiendo a sus preferencias.

En este artículo nos centramos en las trazas VGI que han sido capturadas mediante la utilización de equipos GNSS (teléfonos, navegadores, etc.). Las trazas GNSS 3D son elementos lineales de información geográfica formados por vértices o puntos tridimensionales. Las trazas son capturadas mediante la utilización de un equipo GNSS en el desarrollo de diferentes actividades (conducir, montar en bicicleta, caminar, etc.) sobre diversos elementos lineales (carreteras, caminos, vías de tren, carriles bici, etc.) y utilizando distintos criterios o configuraciones (tiempo, distancia, puntos críticos). Las trazas GNSS 3D, en general, contienen la identidad del elemento, su posición y el tiempo en que han sido capturadas. Información sobre velocidad, aceleración y dirección del movimiento puede extraerse fácilmente de la información anterior. En este trabajo se analizan las alteraciones en la posición de las trazas debido a los problemas que pasaremos a citar a continuación.

Sin embargo, la explotación de estos conjuntos de datos es compleja debido tanto al volumen como a la variabilidad de sus contenidos. No todos los voluntarios tienen la misma formación ni el mismo compromiso en la tarea de adquisición de IG. Los datos pueden ser erróneos (de manera o no intencionada), presentar diferentes niveles de detalle, haber sido capturados con distintos métodos (captura con equipos GNSS, digitalización sobre imágenes, etc.) y con distintos equipos, etc. Todo ello lleva a la necesidad de considerar diferentes niveles de fiabilidad.

¹ El universo de discurso es la visión del mundo real, o hipotético, que incluye todo lo que es de interés para una aplicación geoespacial concreta. Esta visión está condicionada por el proceso de abstracción que genera el mundo conceptual y por las restricciones que impone el modelo geoespacial y por los requisitos de la aplicación concreta (Ariza-López & García-Balboa, 2013).

Las trazas GNSS 3D, además de las alteraciones indicadas en el párrafo anterior, pueden estar afectadas por muchos otros factores que alteran su exactitud (tipo de levantamiento, equipo de captura, hora del día, horizonte de satélites, condiciones meteorológicas, etc.). El objetivo de este trabajo es doble. Por un lado se presentan los problemas que perturban o pueden perturbar a las trazas GNSS 3D procedentes de aportaciones voluntarias, y que afectan a la componente posicional, pudiendo invalidarlas para formar parte de procesos cartográficos como la actualización de Bases de Datos Geográficas. Por otro lado, se plantea un método que permite determinar en qué medida dicha componente se ve afectada. A modo de aplicación de lo anterior, se analizará un conjunto de trazas de la "Vía Verde del Aceite" descargadas de la web www.wikiloc.com.

El presente artículo se estructura como se muestra a continuación. Después de la presente introducción se analiza la posible problemática que puede afectar a estas trazas en la sección 2. A continuación se propone un método de análisis que permita detectar la problemática presentada en la sección anterior. La sección 4 y 5 presentan un caso real de estudio y se muestran y comentan los resultados obtenidos. Por último se presentan las conclusiones extraídas del estudio en la sección 6.

2. Problemas de trazas GNSS procedentes de VGI

Las trazas GNSS VGI pueden verse afectadas por problemas de diferente índole que pueden afectar a su usabilidad en algunos procesos cartográficos, como por ejemplo la actualización de cartografía. En este artículo, se han agrupado estos problemas en las siguientes categorías:

- Acceso y disponibilidad.
- Falta de información complementaria (metadatos).
- Calidad de los datos GNSS.

2.1. Acceso y disponibilidad:

Dos de las principales fuentes donde se pueden encontrar trazas procedentes de aportaciones voluntarias en España son: Wikiloc y Wikirutas. Las dos fuentes mencionadas tienen diferentes políticas de visualización y acceso a los datos. En Wikiloc, por ejemplo, cualquier persona tiene acceso a la visualización de la información subida, pero para poder descargar o subir trazas el usuario debe registrarse en la página. La información que sube un usuario, ya sea geográfica (trazas o puntos de interés) o de cualquier otra índole (comentarios, fotografías, etc.) no puede ser modificada por ningún otro usuario, a no ser que incumpla las condiciones legales indicadas en el sitio web. En Wikirutas, por su parte, las rutas ya están definidas y tienen asignado un nombre, siendo los usuarios los que complementan y/o corrigen la información de cada una de ellas. Esta aportación de información requiere un registro, aunque el acceso a la información es libre.

El acceso a los repositorios donde se encuentran las trazas GNSS de las fuentes mencionadas se realiza de manera individual, por lo que no puede realizarse una descarga masiva. Esto obliga a realizar tantas consultas a los repositorios donde se encuentran las trazas como trazas quieran descargarse. Esto no ocurre con todas las fuentes de datos VGI. Existen algunas (p.e. OpenStreetMap) que permiten el acceso a los datos mediante una API (p.e. utilizando el editor

JOSM), lo cual permite establecer diferentes criterios de búsqueda que agilizan la obtención de las trazas.

Por otro lado, la disponibilidad de la información es limitada. Las trazas disponibles son aquellas que los usuarios han querido compartir. No existe ningún compromiso por parte de ningún organismo de cubrir determinadas zonas, ni determinados elementos como ocurre por ejemplo en la cartografía oficial, donde están definidos explícitamente los elementos que se tienen que levantar, las zonas que se cubren y la periodicidad de actualización. Por lo tanto, la disponibilidad de IG de una zona depende del grado de implicación con la VGI que exista en dicha zona por parte de los usuarios que suelen desplazarse por ella.

2.2. Falta de información complementaria (metadatos):

Los datos procedentes de VGI no suelen contener metadatos. En la mayoría de los casos, estos datos sólo contienen la propia información geográfica (coordenadas), el tiempo en que fueron capturados y, en ocasiones, algún comentario descriptivo de la ruta. Por lo tanto, no tenemos información sobre una serie de aspectos que son relevantes cuando se trabaja con IG. Consideramos que la falta de metadatos afecta mayormente a los siguientes aspectos:

- **Dispositivos de captura:** La precisión alcanzada en la captura de trazas GNSS depende del tipo de receptor que se utilice. Mientras que los receptores conocidos como geodésicos alcanzan precisiones del orden del centímetro en la determinación de la posición, los receptores de navegación tienen una precisión de 6 a 10 metros en condiciones normales (Zhang & Sester, 2010). Éstos últimos son los que habrán sido utilizados, en la mayoría de los casos, en las bases de datos constituidas por VGI. En las especificaciones técnicas de cada equipo suelen venir sus precisiones, pero la falta de conocimiento sobre qué equipo se está utilizando no permiten conocer a priori qué precisión se puede esperar de la IG capturada. Por lo tanto, ante la posibilidad de descargar varias trazas de una misma ruta, no podemos conocer cuál de ellas ha sido capturada con el equipo de mayor precisión.
- **Método de levantamiento:** Los datos procedentes de VGI no son el resultado de un levantamiento específico de la ruta que representan. Los voluntarios no tienen compromiso de capturar una ruta determinada, sino que se desplazan libremente. Por lo tanto, no siguen los patrones que tradicionalmente han conformado los levantamientos topográficos, como puede ser por ejemplo el levantamiento del eje de la vía recorrida, o de los márgenes. Estos levantamientos, además, pueden verse afectados de diversas situaciones: paradas para hacer un descanso, desplazamientos para visitar puntos de interés a lo largo de la ruta, rutas alternativas a la "principal", comportamientos anómalos en el desplazamiento, etc. Tampoco se dispone información sobre la posición del equipo (Figura 1) cuando se capturan las trazas (salpicadero del coche, en el manillar de la bicicleta, en la muñeca, etc.). Esto afecta principalmente a la componente Z de los datos, puesto que no se conoce su altura sobre el suelo.



Figura 1: Ejemplos de diferentes localizaciones de un dispositivo GNSS.

- Configuración del dispositivo: dependiendo de la configuración del equipo de captura, las trazas estarán compuestas por más o menos vértices (o waypoints). Normalmente, los equipos de captura GNSS convencionales ofrecen dos opciones para grabar los trazas, o bien tomando vértices cada cierto tiempo, o bien hacerlo cada vez que se recorre una determinada distancia. Tanto la distancia como el tiempo, suelen ser definidos por el usuario, aunque por defecto vienen configurados para unos tiempos o distancias determinadas. Cuanto menor sea el tiempo o la distancia con la que se toman los vértices que forman la traza, mayor será el número de vértices, así como el tamaño de almacenamiento. Un número elevado de vértices permite una mejor definición de la geometría de una traza, sobre todo en zona de curvas. Sin embargo, algunos dispositivos no tienen capacidad de almacenar gran cantidad de vértices, y es por ello que a veces los usuarios se ven obligados a aumentar la distancia o el tiempo de almacenamiento. Además de la frecuencia con la que se graban los vértices que definen una traza, influye también la velocidad de desplazamiento. Por otro lado, aunque los datos hayan sido capturados con equipos GNSS, éstos no siempre tienen información sobre la coordenada Z. Por lo tanto, en muchos casos, las trazas GNSS no pueden ser utilizadas en aplicaciones 3D.
- Origen de altitudes: El elipsoide es el modelo que usan los equipos GNSS para calcular la altura elipsoidal, si bien algunos de ellos tienen un modelo sencillo del geoide que usan para calcular la altura ortométrica (altura sobre el nivel del mar). Algunos dispositivos, como los de la marca Garmin, pueden realizar una corrección de altura. En estos dispositivos, las correcciones de altura contrastan la posición horizontal (latitud/longitud) proporcionada por el GNSS con los datos de altura adquiridos mediante mediciones profesionales. Cuando se realizan correcciones a los datos de altura, cada punto de rastreo de la actividad contiene la altura del servicio web, no la altura proporcionada por el dispositivo con GNSS. Garmin Connect (la aplicación web que gestiona los datos capturados con equipos Garmin) aplica las correcciones de manera selectiva para hacer una representación más realista de la altura. Por otro lado, existen dispositivos que disponen de altímetro barométrico que mejora, a priori, la precisión de altura proporcionada por la señal GNSS, siempre y cuando las condiciones atmosféricas sean las adecuadas. La altura proporcionada por el barómetro es especialmente útil cuando no existe solución GPS 3D o para suavizar la cota derivada del GPS, que por lo general suele presentar bastante ruido. La falta de metadatos en este sentido, no permite conocer el origen de altitudes de los datos y, por lo tanto, tiene influencia en los procesos cartográficos que consideren la altitud.

3. Métodos de análisis del comportamiento posicional de las trazas GNSS 3D procedentes de VGI.

Para poner de manifiesto como los distintos problemas comentados en el apartado anterior y que afectan a la componente posicional, se plantea un método que permita el análisis del comportamiento posicional de las trazas GNSS 3D procedentes de VGI.

Cho (2012) propone dos tipos de evaluación para la VGI: evaluaciones directas (mediante comparaciones de la VGI con datos procedentes de otra cartografía, como por ejemplo la cartografía oficial), o evaluaciones indirectas (número de voluntarios de un área dada, perfil de cada voluntario, frecuencia de las actualizaciones, etc.). El análisis propuesto en este trabajo, y que se muestra a continuación, encaja parcialmente en la clasificación de evaluación directa, puesto que gran parte del análisis visual y una de las variables del análisis cuantitativo toman como referencias ortofotografías y modelos digitales del terreno procedentes de un organismo oficial.

En este trabajo se propone un análisis de la componente posicional basado en conjuntos de trazas que pertenecen a una misma entidad geográfica (p.e. una carretera, un camino, una vía, etc.), a partir del cual se puede detectar comportamientos anómalos de unas trazas con respecto al conjunto. Proponemos utilizar dos mecanismos para detectar estos comportamientos anómalos: análisis visual y análisis cuantitativo.

3.1. Análisis visual (2D y 3D):

En este primer análisis se visualizan las trazas y se comparan sus posiciones, detectando las posibles discrepancias que puedan existir.

- Análisis visual 2D: El análisis visual de la planimetría de las trazas es relativamente sencillo si se dispone, por ejemplo, de una Ortofotografía sobre la cual superponer las trazas (Figura 3, derecha). De este modo se pueden detectar desplazamientos de las trazas con respecto al elemento que representan.
- Análisis visual 3D: El análisis visual de la altimetría puede realizarse con la superposición de las trazas sobre un Modelo Digital del Terreno (MDT) (Figura 3, izquierda) a partir del cual poder observar las discrepancias con respecto al terreno. Otra opción es mediante la comparación de los perfiles longitudinales de las trazas que forman el conjunto.

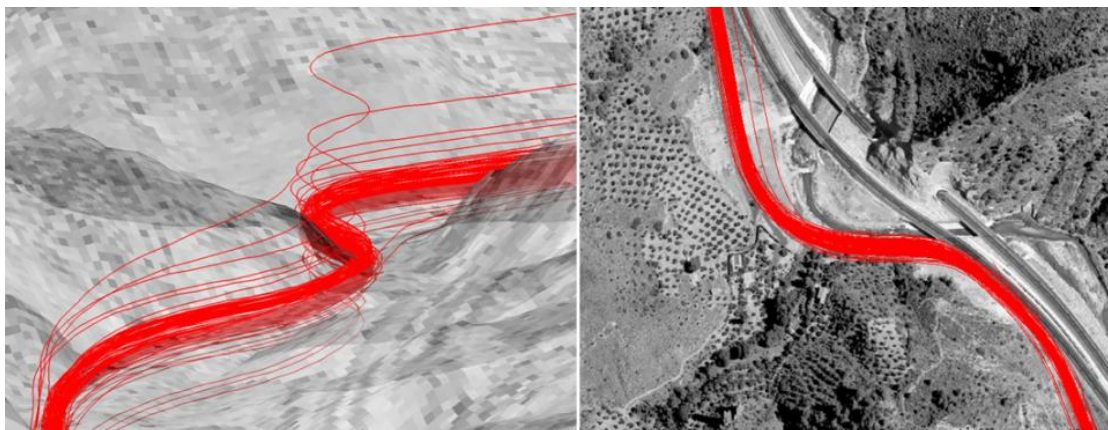


Figura 3: Visualización de trazas 3D (izquierda) y 2D (derecha).

Fuente: Ariza-López et al. (2014).

3.2. Análisis cuantitativo.

Se analizan aspectos cuantitativos del conjunto de trazas (longitud, número de vértices, etc.) que pongan de manifiesto las posibles anomalías posicionales que puedan presentar las trazas. Estas anomalías posicionales pueden denotarse como observaciones atípicas u outliers y existen numerosas variables y diversas formas de detectarlas.

Para este estudio se ha utilizado una de las representaciones gráficas propuestas por Ariza-López et al. (2014), los diagramas box-plot. Los diagramas de caja o box-plot son uno de los tipos de gráfico utilizado frecuentemente en la detección de valores atípicos. Estos diagramas incorporan en un solo gráfico medidas de posición, dispersión y observaciones atípicas. Por tanto, muestran a simple vista las principales características de la muestra. Este tipo de gráficos utilizan medidas de posición y dispersión robustas, basadas en los cuartiles.

Este tipo de diagrama se ha utilizado para analizar cuatro variables: longitud, número de vértices de las trazas, las desviaciones típicas de las distancias entre vértices consecutivos dentro de una traza y desplazamiento máximo en Z de los vértices respecto al Modelo Digital del Terreno con paso de malla de 5 m (MDT05) del Instituto Geográfico Nacional (IGN). Las dos primeras variables son dos propiedades directas de las trazas, mientras que las dos siguientes variables se calculan.

4. Caso de estudio

Para determinar los problemas anteriormente mencionados afectan a la componente posicional de las trazas GNSS VGI, se hace una descarga de trazas desde la página www.wikiloc.com. Se ha elegido Wikiloc porque en esta plataforma existen numerosas trazas de las mismas rutas capturadas por distintos usuarios, lo cual ha permitido obtener un conjunto de trazas. Otras plataformas como OpenStreetMap también tiene esta posibilidad, además de poder realizar la descarga utilizando una API, pero se ha descartado su uso porque las trazas que se descargan carecen de la coordenada Z, y en este análisis se quería trabajar con trazas 3D.

La zona elegida para llevar a cabo el análisis es un tramo de la "Vía Verde del Aceite" situada en la provincia de Jaén. En España, existen más de 2.100 kilómetros de infraestructuras ferroviarias en desuso que han sido reconvertidas en itinerarios cicloturistas y senderistas en el marco del Programa Vías Verdes, coordinado por la Fundación de los Ferrocarriles Españoles (FFE). Las principales características de las vías verdes son (Luque Valle, 2011):

- Tránsito cómodo debido a un escaso desnivel (máximo 3%), y una anchura de vía amplia y constante (2,5 metros).
- Infraestructura específica para tráfico no motorizado o tráfico lento.
- Número reducido de cruces con carreteras.
- Fáciles y agradables de recorrer, adentrándose en paisajes de gran valor natural y cultural.

Las vías verdes acogen tanto a usuarios, entendidos como personas que viven en el entorno de la vía verde y hacen uso de ella, como a visitantes, que son aquellas personas que se desplazan desde su lugar de origen. Según una encuesta realizada a los gerentes y responsables de la gestión de las vías verdes en (MERCODES, 2008), la "Vía Verde del Aceite" tiene una estimación de 30.000 personas anuales, de los cuales el 60% son usuarios locales y un 40% son visitantes. De estos usuarios, un 52,3% son caminantes, senderistas o corredores, el 45,3% son ciclistas, el 0,7% recorre la vía a caballo, el 0,9% son personas con movilidad reducida y el resto son otros usuarios (Luque Valle, 2011).



Figura 4: Itinerario de la Vía Verde del Aceite.

Fuente: www.viasverdes.com

La "Vía Verde del Aceite" se localiza entre la ciudad de Jaén y el río Guadajoz, límite entre las provincias de Jaén y Córdoba (Figura 4), a partir de donde da comienzo la conocida como "Vía Verde de la Subbética". Este recorrido tiene una longitud total de 54km, un desnivel total de 250m con una pendiente máxima del 3% y puede ser recorrido a pie, en bicicleta y a caballo, estando prohibido el acceso de vehículos a motor (salvo en determinados tramos donde hay parcelas particulares colindantes). Esta vía verde se considera, según la tipología establecida por la FFE, como una vía verde acondicionada, es decir, que ha existido una inversión pública para su recuperación. Las últimas actuaciones de adecuación de esta vía verde fueron llevadas a cabo en el segundo semestre de 2008, dando la configuración actual de esta vía.

4.1. Caracterización del conjunto de trazas descargadas.

La "Vía Verde del Aceite" es un recorrido de "ida y vuelta", aunque muchos usuarios utilizan caminos alternativos para realizar un recorrido "circular". Aunque muchos usuarios la recorren completa, sobre todo aquellos que se desplazan en bicicleta, otros recorren sólo unos determinados tramos. Para este trabajo se ha seleccionado el tramo que discurre entre las localidades de Jaén y Torredelcampo. Este es un tramo habitual entre los usuarios de la vía verde que comienzan en dichas localidades, y ha permitido obtener un conjunto numeroso de trazas GNSS (Figura 5). La longitud total del tramo seleccionado es de aproximadamente 8,4km. Se han descargado de forma individual un total de 61 trazas procedentes tanto por ciclistas como por personas que se desplazan a pie, de las cuales 24 son trazas de ida y vuelta y con fechas comprendidas entre los años 2008 y 2015. Las trazas se descargan visualizando una por una conforme aparecen en la búsqueda, y descargando aquellas que cubren el tramo de vía verde seleccionado. Según se va avanzando en la búsqueda, cada vez hay más trazas que pertenecen a otros ámbitos geográficos (p.e. otras vías verdes). Cuando dejan de aparecer trazas del tramo seleccionado, se interrumpe la búsqueda.

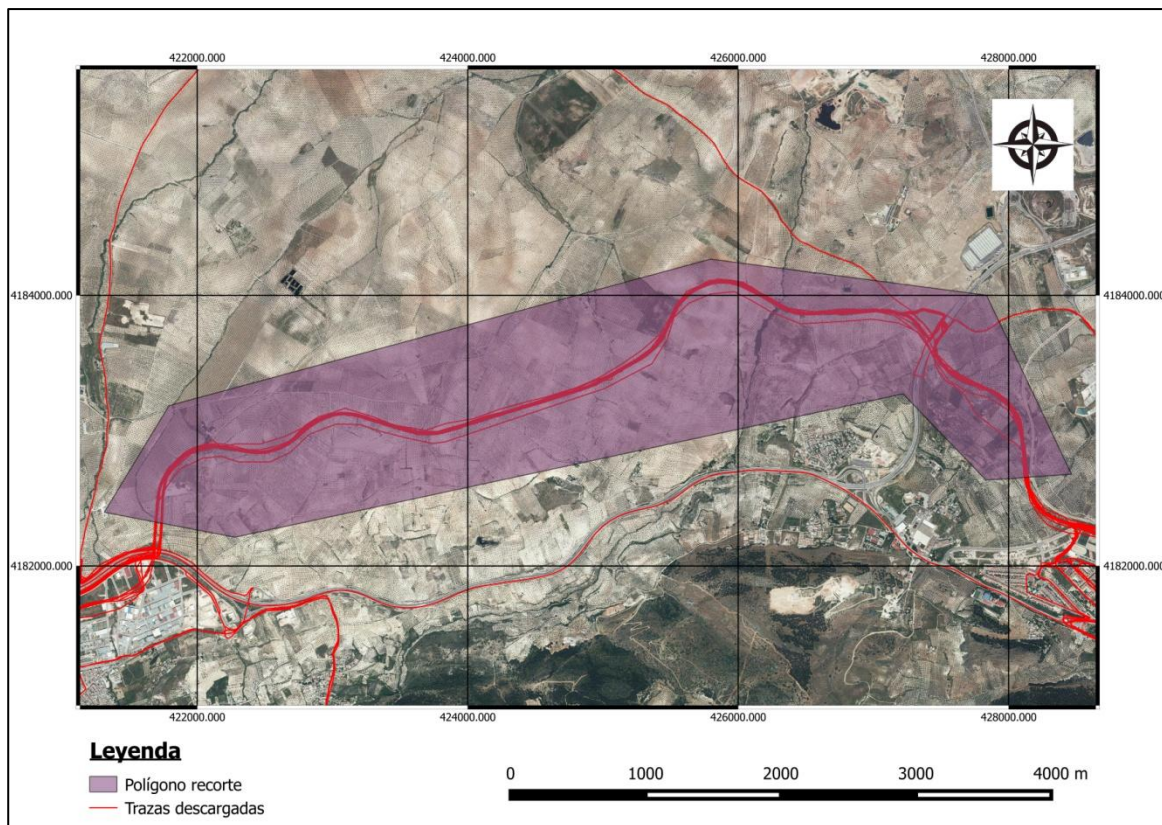


Figura 5: Trazas GNSS 3D descargadas y tramo seleccionado. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)

4.2. Tratamiento de las trazas descargadas.

Como se aprecia en la Figura 5, las trazas descargadas cubrían una longitud mayor que la del tramo seleccionado, por lo que han sido necesarias realizar una serie de operaciones (Figura 6) que permitieran extraer la parte de traza correspondiente al tramo seleccionado para este análisis. Para ello se ha desarrollado un script con una secuencia de comandos en código Python que utiliza las herramientas de ArcGIS 10 a través del paquete ArcPy.

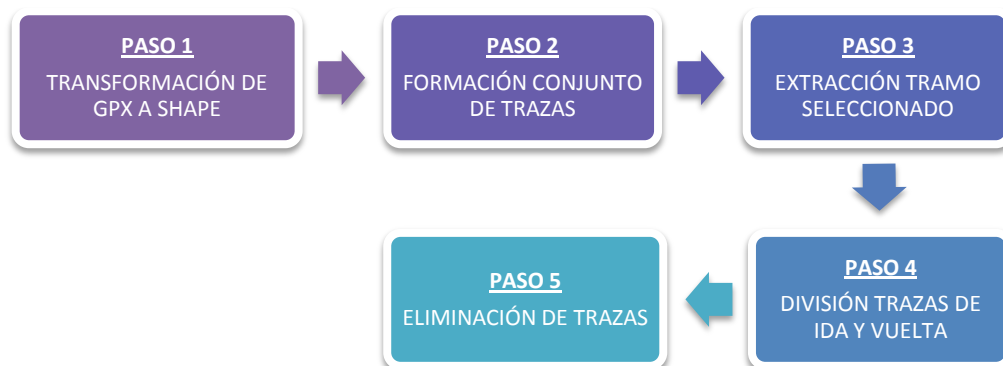


Figura 6: Tratamiento de las trazas descargadas.

En el tratamiento de las trazas descargadas se incluyen los pasos que muestra la Figura 6:

- PASO 1: Las trazas, originalmente en el formato GPX (GPS Exchange Format), se transforman a formato Shapefile (formato vectorial de almacenamiento digital donde se guarda la localización de los elementos geográficos y los atributos asociados a ellos) para poder aplicar los comandos que se muestran a continuación.
- PASO 2: Todas las trazas, que se encuentran originalmente en archivos individuales, se añaden a un único archivo.
- PASO 3: Se extraen los tramos de las trazas que están dentro de la zona seleccionada para realizar el análisis.
- PASO 4: Las trazas que son de ida y vuelta, se dividen en dos, por lo tanto, aumenta el número de trazas con respecto al conjunto de trazas original.
- PASO 5: Se realiza un proceso de depuración en el cual se eliminan algunas trazas que están afectadas de alguna de estas circunstancias:
 - no tenían información de la coordenada Z.
 - no recorren completamente el tramo seleccionado.
 - no fueron capturadas con equipos GNSS (sino que están digitalizadas directamente por los usuarios sobre una imagen).

Tras el tratamiento de las trazas descargadas descrito anteriormente, se dispone de un conjunto total de 76 trazas (Tabla 1). De las 61 trazas originales, 24 se dividen en dos en el Paso 4 al tratarse de trazas de ida y vuelta, de manera que se dispone de un conjunto de 85 trazas. En el paso 5 se ha depurado el conjunto, eliminando una traza que no contenía información Z, dos trazas que no completaban el recorrido total seleccionado y seis trazas que fueron introducidas a mano por el usuario. Finalmente, el conjunto de trazas que queda para el análisis está formado por 76 trazas.

Tabla 1: Caracterización de las trazas.

Número de trazas descargadas		61
Trazas de ida y vuelta (se desdoblán como dos trazas individuales)		24
Trazas eliminadas	No tienen información Z	1
	No recorren completamente el tramo seleccionado	2
	Digitalizadas	6
	Total trazas descartadas	-9
Número de trazas del conjunto final analizado		76

5. Resultados y discusión

En este apartado se recogen y discuten los resultados obtenidos en el análisis propuesto en el apartado 3 (análisis visual y cuantitativo), aunque en primer lugar se hace referencia al problema de acceso y disponibilidad de las trazas.

5.1. Acceso y disponibilidad.

El primer problema que se afronta para obtener el conjunto de datos descrito en el apartado 4 es el acceso a los datos. Como se ha comentado, las trazas obtenidas para el análisis, se han descargado de la www.wikiloc.com. Para realizar la descarga de las 61 trazas seleccionadas, se utiliza el buscador de texto que ofrece la página, donde se introduce el texto "Vía Verde del Aceite". Esta búsqueda devuelve 27118 resultados, sin embargo, la inmensa mayoría de ellos no pertenecen realmente a la ruta que se busca. Por un lado, algunos de los resultados, aún perteneciendo a la Vía Verde del Aceite, no transcurren por el tramo seleccionado. Por otro lado, muchas trazas no pertenecen a esta Vía Verde sino que son trazas capturadas en otros lugares. Hay que seleccionar una a una, visualizarlas sobre un mapa y, en caso de pertenecer al tramo seleccionado, descargarla para su posterior análisis. Como alternativa al buscador de Wikipedia, y para evitar que en la búsqueda aparezcan trazas que no contengan estrictamente la cadena de caracteres buscada ("Vía Verde del Aceite"), se propone el uso del buscador de Google. Éste permite una búsqueda más exhaustiva. Por ejemplo, con el buscador de Google se puede indicar que sólo busque aquellas rutas que contengan la frase completa, lo cual evita que en los resultados puedan aparecer, por ejemplo, rutas que pertenezcan a otra vía verde.

5.2. Análisis visual.

En la vista general que ofrece la Figura 5 se aprecian las discrepancias que existen entre las trazas. A continuación se muestran otras visualizaciones ampliadas de determinadas zonas que permiten detectar las anomalías de alguna de las trazas (Figuras 7-12).

Al descargar una traza de una determinada ruta, puede que el voluntario que subió la información a internet no siguiera exactamente la ruta considerada. Muchos de los voluntarios toman rutas alternativas de la ruta considerada (Figura 7.a). Los voluntarios realizan sus desplazamientos libremente y no tiene porqué seguir ninguna ruta. Por ello, en muchas ocasiones, existirán desvíos respecto a la ruta analizada, bien porque tomen por caminos que discurren de forma más o menos paralela, como es el caso que se muestra en la Figura 7.a, o bien porque deciden no ir expresamente por el camino por diversas razones (p.e. zonas de charcos o barro, alguna incidencia en la ruta, evitar aglomeraciones de gente en determinados lugares, etc.).

Gil de la Vega, P., Ariza López F. J., Mozas Calvache, A. T. (2016): "Problemas que presentan las trazas GNSS procedentes de Información Geográfica Voluntaria", *GeoFocus (Artículos)*, nº17, p. 161-184. ISSN: 1578-5157

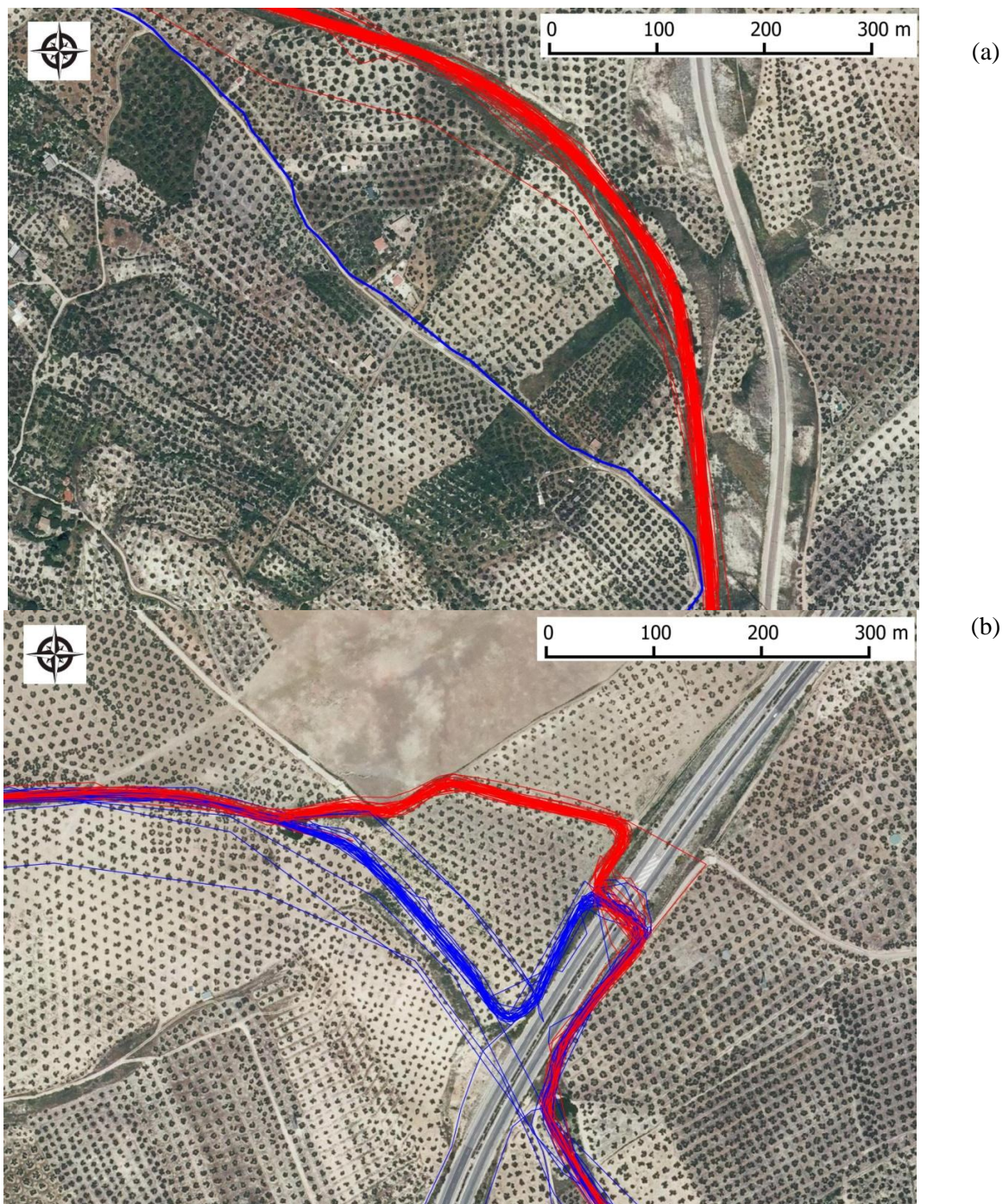
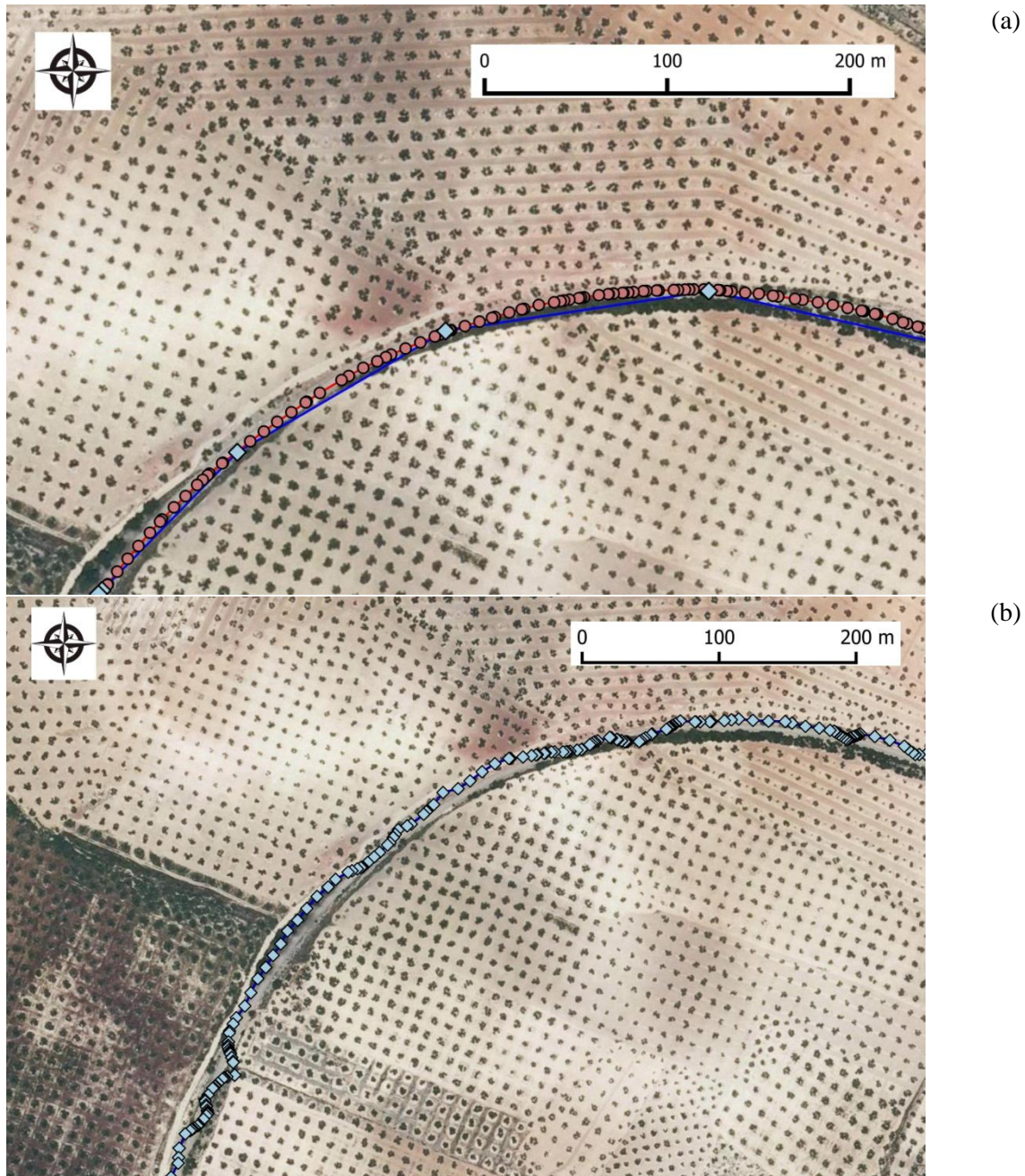


Figura 7: Rutas alternativas a la ruta considerada. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)



**Figura 8: Definición geométrica de una traza en función del número de vértices.
Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)**

En otras ocasiones, es la propia ruta la que considera diferentes alternativas (Figura 7.b). En este caso, la ruta señalada como Vía Verde del Aceite e identificada como tal en el terreno mediante carteles, es la que se encuentra más hacia el sur, sin embargo ambas alternativas se usan indistintamente por los usuarios, como se puede apreciar en la Figura 7.b. Así, el 54% de las trazas analizadas toman el recorrido más al norte marcado en color rojo, frente a un 46% que toma la otra

alternativa. La alternativa sur en dirección hacia Torredelcampo, presenta una pendiente ascendiente mayor y un suelo de gravilla que la hace más difícil de transitar, sobre todo para ciclistas, y es por ello que muchos usuarios prefieren seguir transitando por el recorrido original de la Vía Verde.

La configuración del equipo GNSS en la captura de los datos tienen una influencia directa en la geometría de la traza capturada. Cuanto mayor sea el número de vértices que definen una traza, mejor será la definición de ésta (Figura 8.a) Sin embargo, puede ocurrir que, aunque exista una densidad de vértices adecuada (Figura 8.b), éstos tengan una exactitud posicional inadecuada. Por lo tanto, la densidad de puntos no es, por sí sola, una variable que pueda ser fiable para elegir una traza frente a otra.

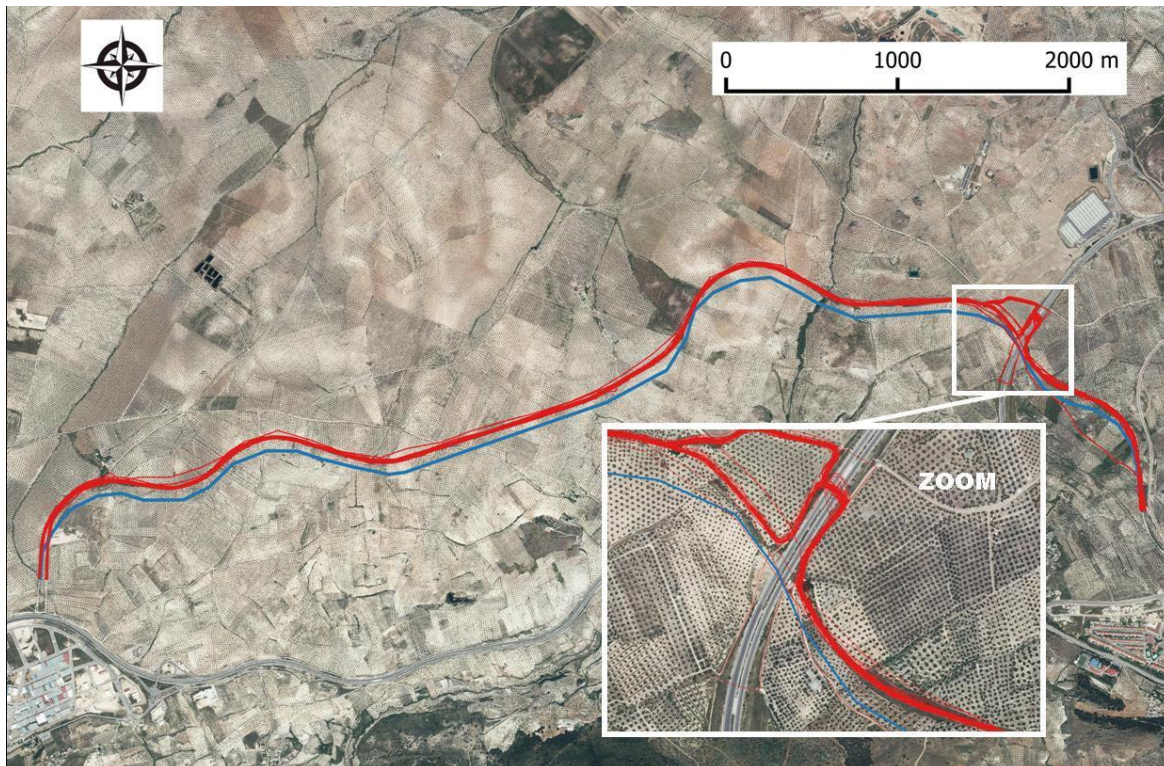


Figura 9: Traza sesgada. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)

En la Figura 9 se observa una traza con un comportamiento totalmente sesgado (azul) frente al resto de trazas (rojo). En este caso, y observando la superposición de la traza sobre la ortofotografía, parece poco probable que el voluntario haya tomado un camino alternativo, y más teniendo en cuenta que la traza, como se muestra en el zoom que se presenta en la Figura 9, cruza la autovía por un lugar donde no existe ningún paso, aunque no se podría descartar del todo pues, como ya se comentó, el voluntario se desplaza libremente por donde quiere. El error posicional podría deberse a la precisión del equipo de captura, que suele oscilar entre 6 y 10 metros. Otra posible causa puede ser que la señal recibida pueda estar afectada por las fuentes de error comentadas en el apartado 0. En este caso concreto, donde los errores se cuantifican en unos 60

metros, es probable que se haya producido una concatenación de errores que llevan a un desplazamiento de tal magnitud, o bien se debe a una diferencia del datum. Sin embargo, dicha diferencia de datum no justificaría que la traza azul "cruzara" la autovía en lugar de realizar la curva que todas las demás trazas realizan para pasar por el túnel que se encuentra en el noreste, tal y como se aprecia en el zoom de la Figura 9.

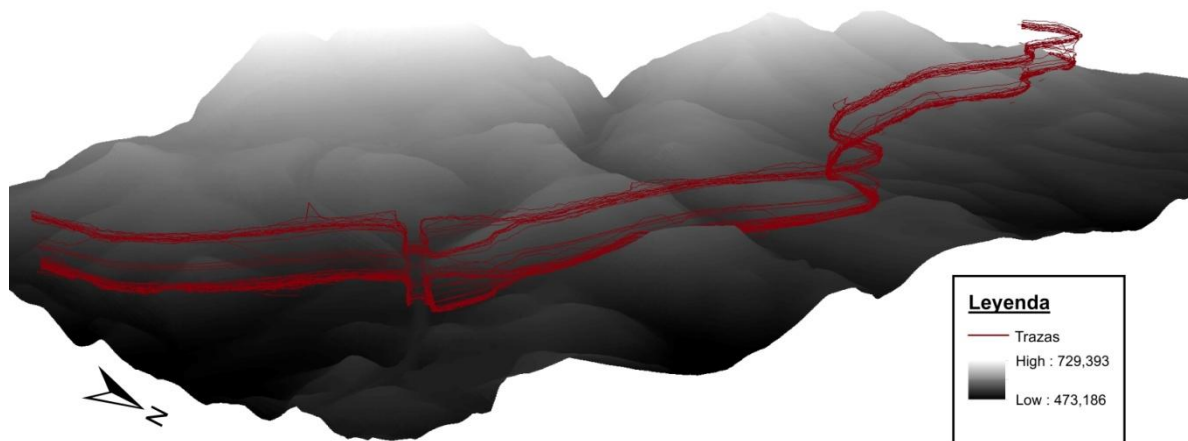


Figura 10: Visualización 3D de las trazas sobre el MDT05 (cota exagerada x5).

Otro tipo de anomalía frecuente en las trazas VGI es el que se presenta en el comportamiento en Z. Los equipos GNSS determinan la posición altimétrica con menor precisión que la posición horizontal. La precisión de la componente horizontal es aproximadamente 1,5 veces mayor que la precisión vertical (Arnold, 1998). En la Figura 10 se pueden observar las trazas sobre el MDT05. Se aprecia un desplazamiento de las trazas por encima de modelo que alcanza, en el punto más alejado, los 77m.

En la Figura 10, se observa que las trazas parecen agruparse en dos alturas diferentes, lo que puede indicar que, en algunos casos el dispositivo de captura haya registrado la altura elipsoidal y en otros se haya registrado la altura ortométrica en función de las opciones de configuración. Para comprobar esto, se realiza el cálculo promedio del desplazamiento entre los dos grupos de trazas de alturas diferentes en un punto determinado del recorrido, y se obtiene que ambos grupos están desplazados unos 55m (Figura 11). Para ese mismo punto, se determina la ondulación del geoide con la herramienta "Programa de Aplicaciones Geodésicas" disponible en el IGN, y se determina una ondulación de geoide de 50m aproximadamente. Esto permite corroborar la hipótesis de que existen dos tipos de altitudes en los datos.

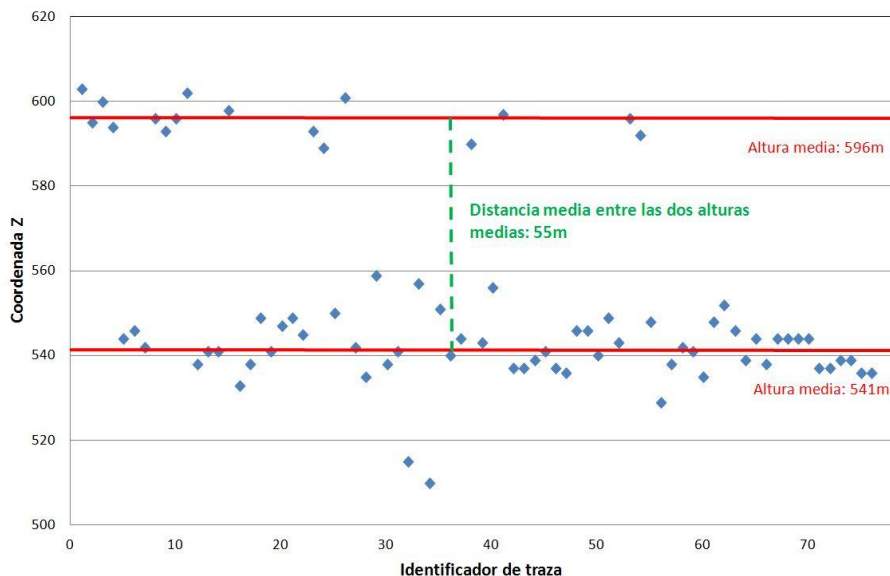


Figura 11: Gráfico de la coordenada Z de las trazas en un punto de la vía verde.

Además de los desplazamientos respecto al modelo de elevaciones, se puede observar cómo se comporta la coordenada Z a lo largo de una misma traza. En la Figura 12 se observan dos perfiles: en rojo un perfil de una traza con un comportamiento más ajustado a la realidad de la Vía Verde (sin grandes oscilaciones en Z de vértices consecutivos), y en azul otro perfil de una traza que presenta grandes variaciones en Z para vértices consecutivos. La traza azul presenta unas oscilaciones que no se corresponden con la realidad, en la que no hay cambios bruscos en la pendiente, tal y como se muestra en las características recogidas de las vías verdes en el apartado 4. Estas oscilaciones podrían tener origen, como se vio en el apartado 2, en una mala medición del altímetro barométrico por unas condiciones atmosféricas inestables, aunque la falta de información respecto al equipo de captura no permiten conocer si éste o no usa un altímetro barométrico.

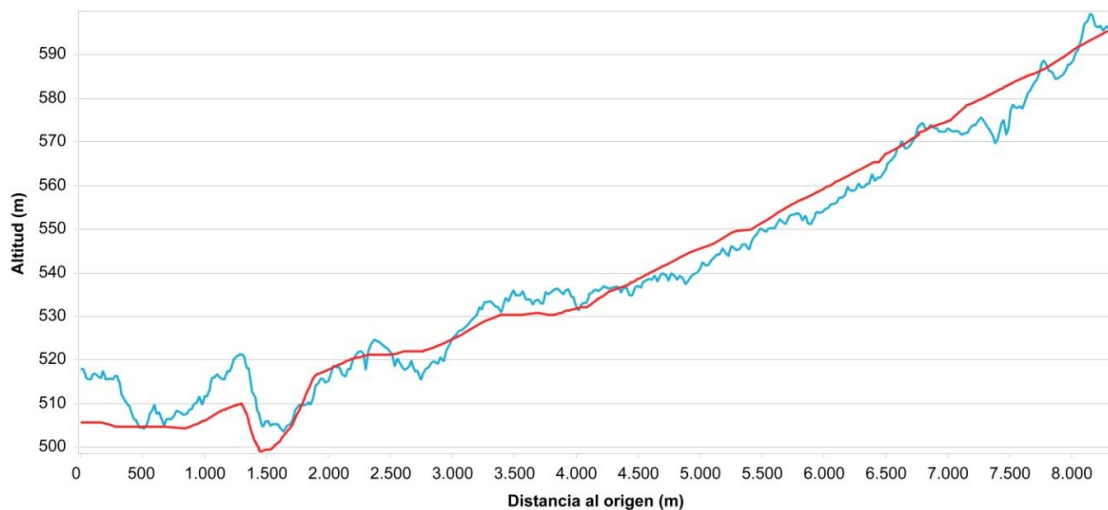


Figura 12: Perfiles de traza.

5.3. Análisis cuantitativo.

Los diagramas box-plot de las cuatro variables estudiadas (longitud, número de vértices, desviación típica de la distancia entre vértices, distancia máxima en z) se recogen en la Figura 13. En estos diagramas se representan los límites inferiores y superiores para cada variable y, en rojo, se muestran aquellos valores que se salen de estos límites y que, por lo tanto, se denotan como valores atípicos. En las cuatro variables analizadas se identifican trazas que contienen comportamientos anómalos y que, por lo tanto, muestran que ha existido algún problema en la captura de la información con los equipos GNSS.

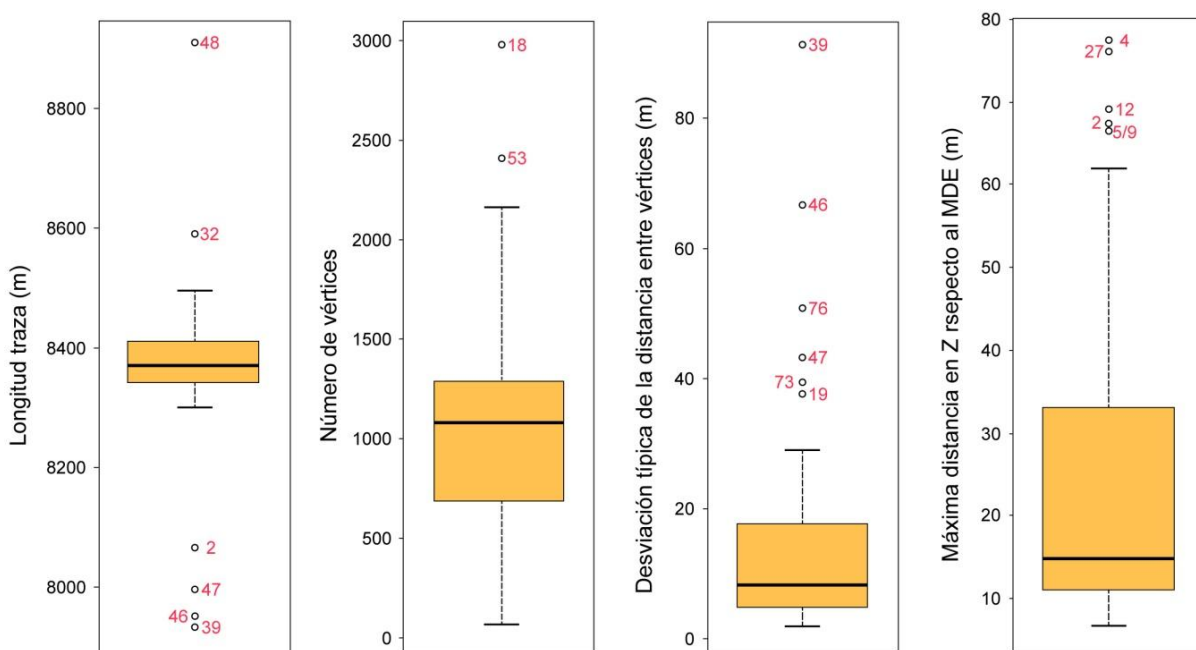


Figura 13: Diagramas box-plot.

Los valores atípicos de longitud se salen tanto del límite inferior como del superior. Los primeros presentan comportamientos como los que se observan en color azul en la Figura 14 donde las trazas cruzan la autovía en lugar de seguir el recorrido que marca la vía verde, por lo que tienen una longitud menor. En general, estas trazas, excepto la 2, están afectadas de error, como se observa en la superposición sobre la Ortofotografía que aparece en la Figura 14, ya que las trazas en todo momento van fuera de la Vía Verde. Las dos trazas que presentan valores atípicos por tener una longitud mayor a la establecida por el límite superior (en color rojo) presentan dos comportamientos diferentes. La traza 32 presenta una longitud mayor porque durante una parte del camino discurre por un lugar diferente recorriendo una mayor distancia, tal y como se observa en la Figura 14. En el caso de la traza 48, aunque en la Figura 14 no se aprecia, presenta una geometría titubeante tanto en planimetría como en altimetría que tiene como consecuencia una mayor longitud de la traza.

Con respecto a los atípicos detectados analizando la variable número de vértices, sólo existen valores por encima del límite superior. El exceso de vértices, no tiene por qué denotar una

observación atípica en la componente posicional de la traza, puesto que indica la frecuencia con la cual se capturan los vértices que forman la traza. Como se observó en la Figura 8, un mayor número de vértices pueden representar mejor la geometría, siempre y cuando estos vértices tengan una exactitud adecuada. En las trazas detectadas como atípicas, al superponerlas sobre la ortofotografía (Figura 15), no se han encontrado comportamientos anómalos. Por lo tanto, en este caso las trazas no deben considerarse atípicas por tener muchos vértices, sino que el exceso de vértices, como se observa en la Figura 15 ayuda a una mejor representación de la traza en las zonas con curvas.

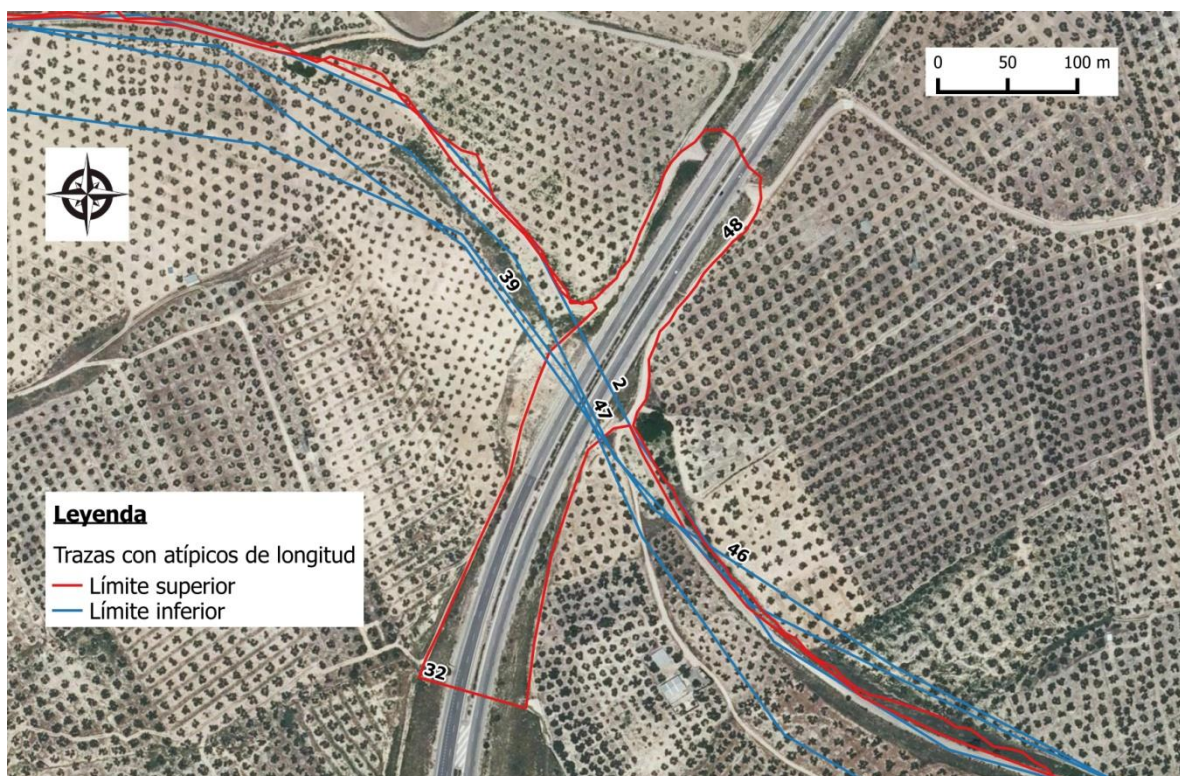


Figura 14: Representación de las trazas con atípicos de longitud. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)

De los atípicos detectados en el análisis de la variable desviación típica de las distancias entre vértices consecutivos, tres de ellos ya habían sido incluidos en las observaciones atípicas correspondientes al análisis de la longitud de las trazas (trazas 47, 46 y 39) por presentar una longitud demasiado pequeña con respecto al conjunto. Para estas tres trazas, como muestra la Figura 16, el problema de la traza no es la acumulación de vértices en determinados lugares, sino la existencia de vértices muy desplazados respecto a la posición real que representan. En las otras tres trazas detectadas (73, 76 y 20) la acumulación de vértices en determinados lugares no influye negativamente en la precisión ni en la geometría de las trazas (Figura 16). Esta variabilidad en la distancia entre vértices consecutivos suelen ser la consecuencia de paradas de los usuarios, cambios en la velocidad (p.e. en zonas de pendiente favorables) o pérdidas temporales de la señal del GNSS.

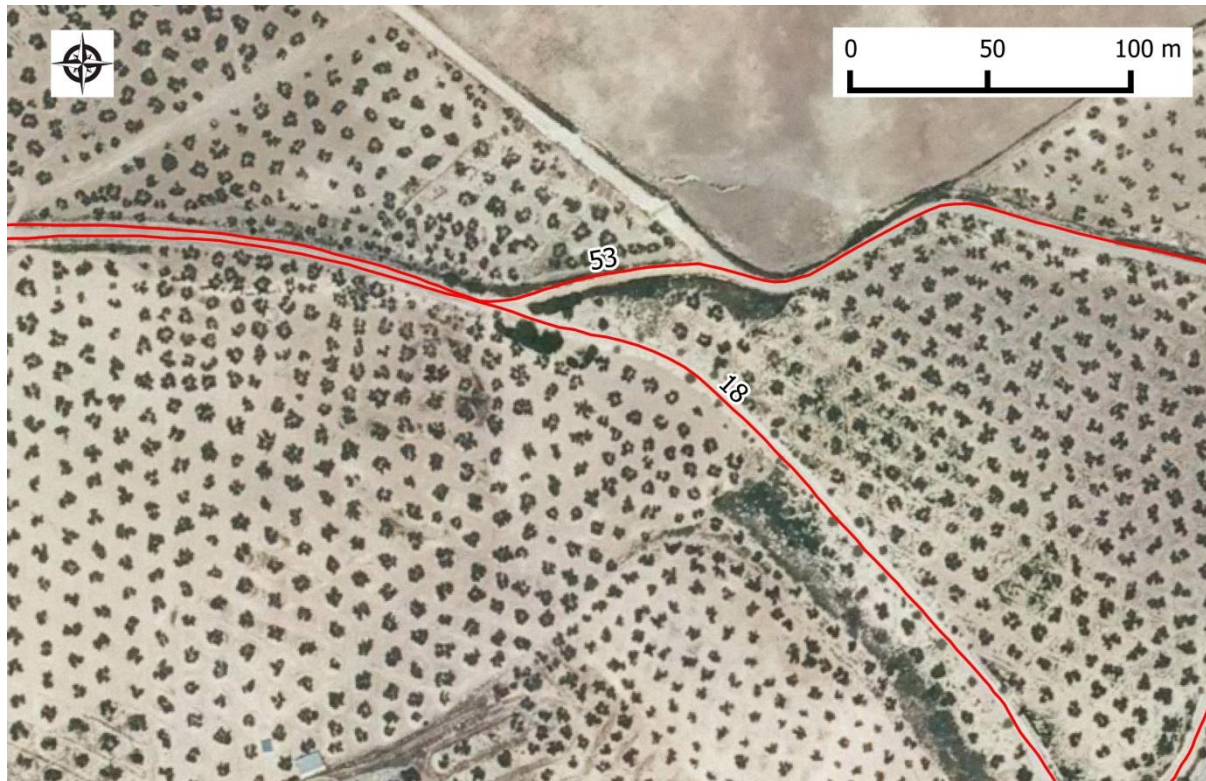


Figura 15: Representación de las trazas con atípicos de número de vértices. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)

Para el análisis de los desplazamientos en altitud, se ha tomado como “verdad terreno” la representada por el MDT05. Para cada traza se ha determinado el desplazamiento máximo con respecto al modelo por considerar que esta distancia es la que puede denotar a una traza como atípica (una distancia mínima tiene menos sentido puesto que, cuanto más se aproxime una traza a la realidad, mejor será su componente posicional). Aunque, en general, el desplazamiento es elevado (la media de las distancias máximas es de 25m), sólo 6 trazas exceden en el umbral superior establecido por el diagrama box-plot. En este caso, dado que no existen metadatos, no es posible saber qué tipo de altitud se está utilizando (elipsoidal, ortométrica), lo cual puede también justificar el excesivo desplazamiento que presentan algunas trazas.

6. Conclusiones

La VGI supone una verdadera revolución en la obtención de IG, reduciendo los costes de forma drástica. Esto permite que la Producción Cartográfica se nutra de una cantidad de IG sin precedentes hasta el momento. En este trabajo se ha demostrado que disponer de un conjunto abundante de trazas no es sencillo, pues los servidores donde se almacenan no están pensados para este tipo de uso. Además, no se debe olvidar que esta información no está tomada bajo ningún criterio de calidad dado que no suele existir finalidad cartográfica en la toma de los datos. Existen numerosos problemas que alteran la exactitud de la IG voluntaria, haciendo que se deban tomar una

Gil de la Vega, P., Ariza López F. J., Mozas Calvache, A. T. (2016): "Problemas que presentan las trazas GNSS procedentes de Información Geográfica Voluntaria", *GeoFocus (Artículos)*, nº17, p. 161-184. ISSN: 1578-5157

serie de precauciones y que se deban desarrollar algunos mecanismos que permitan detectar las anomalías y reducirlas en la medida de lo posible.

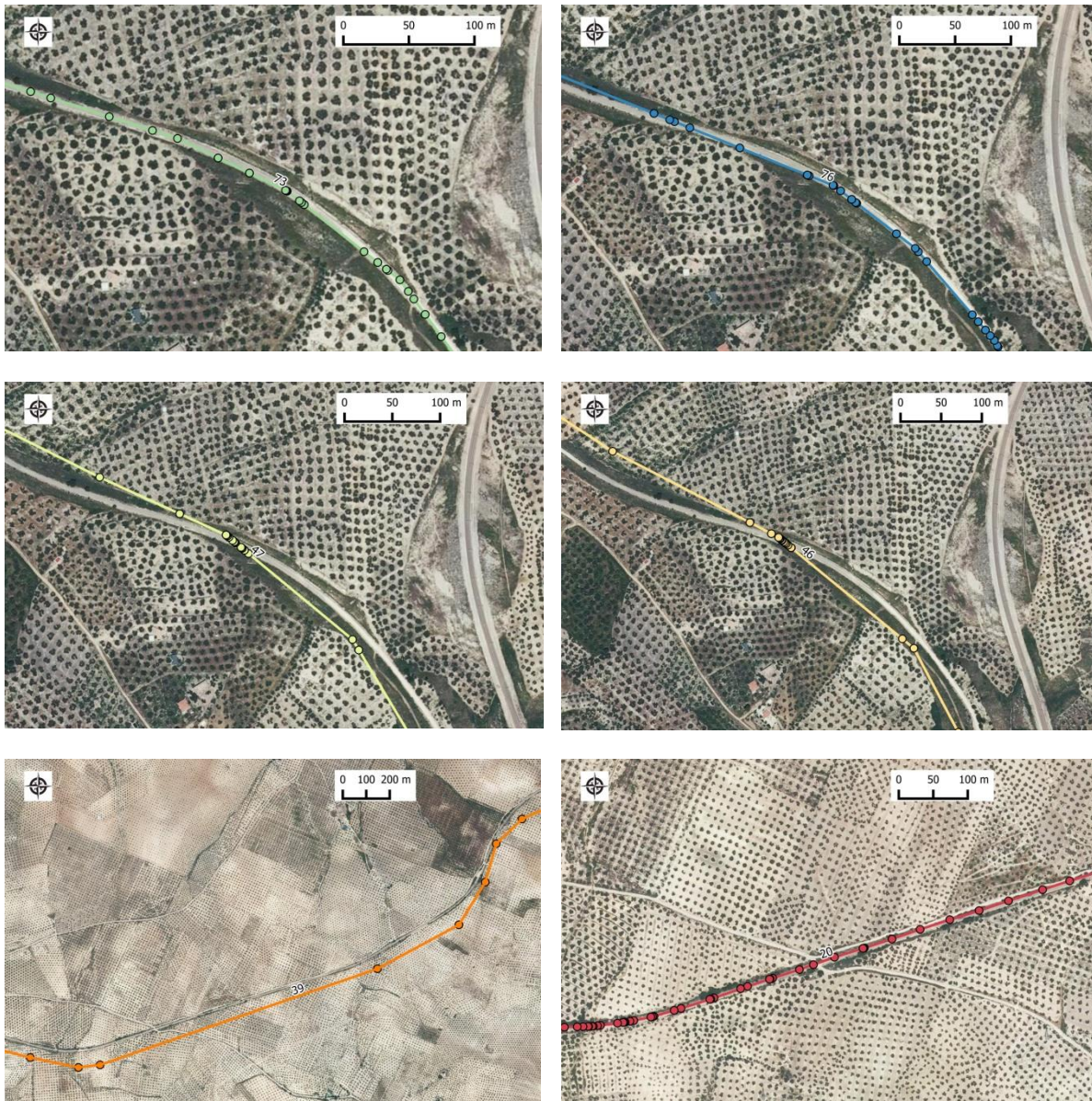


Figura 16: Representación de trazas con atípicos en la desviación típica de la distancia entre vértices. Ortofotografía de fondo: PNOA 2013 (Instituto Geográfico Nacional)

Algunos de los problemas más comunes de las trazas GNSS procedentes de VGI han sido expuestos en este artículo. Para mostrar cómo afectan estos problemas sobre la componente posicional, se han analizado un conjunto multitrizas. El análisis de este conjunto, tanto de forma visual como de forma cuantitativa, ha permitido vislumbrar cómo los problemas expuestos afectan a

las trazas, haciéndolas en muchos casos inutilizables para cualquier proceso cartográfico, tanto de manera individual (si se toma una única traza para un determinado procedimiento) como de manera colectiva (si se toman conjuntos de trazas sobre los que aplicar por ejemplo minería de datos con el fin de obtener un eje medio).

Para evitar el uso de trazas anómalas, y siempre que se disponga de un conjunto de trazas, se puede aplicar algún proceso de detección y eliminación de atípicos. En este artículo se ha probado el uso de diagramas box-plot. Se han elaborado cuatro diagramas para cuatro variables diferentes: longitud, número de vértices, desviaciones típicas de las distancias entre vértices consecutivos y desplazamiento máximo en altitud de la traza respecto al MDE, obteniendo los valores atípicos a partir de diagramas box-plot.

El hecho de que los atípicos detectados en longitud coincidan con las situaciones anómalas detectadas en el análisis visual pone de manifiesto que esta medida puede ser de gran utilidad en un proceso de detección de atípicos. Todos los atípicos detectados con esta medida lo son realmente, aunque sería necesario un análisis más profundo para determinar si todos los atípicos han sido detectados. Los atípicos detectados en el análisis del número de vértices y desviación típica de la distancia entre vértices, requiere un estudio más profundo que considere, además del número de vértices, la calidad individual de estos vértices. En la componente Z, sería necesario conocer el origen de altitudes para poder hacer una evaluación más exhaustiva. Es imprescindible dotar de un mínimo de metadatos a la VGI para que pueda ser tenida en cuenta de manera efectiva en los procesos cartográficos estándar.

Por último, para poder incorporar la evaluación de la calidad de las trazas recogida en este trabajo a la plataforma de Wikiloc, lo único que ésta permite es hacerlo mediante una valoración, que podrá ser positiva o negativa. Esta valoración sirve, entre otros parámetros, para calcular el "TrailRank", que es el valor a partir del cual se ordenan las trazas en los listados de búsqueda de Wikiloc. Sin embargo, tampoco podemos saber el alcance de esta valoración en el "TrailRank" dado que la fórmula de cálculo no es conocida.

Agradecimientos

Este trabajo ha sido financiado parcialmente por Ministerio de Economía y Competitividad del Gobierno de España y por el Fondo Europeo de Desarrollo Regional (FEDER), con la ayuda de referencia BIA2011-23271 correspondiente al proyecto de investigación "Evaluación 3D de elementos lineales de Información Geográfica (E3DLING)". Igualmente, se agradece a la Junta de Andalucía la financiación económica del Grupo de Investigación en Ingeniería Cartográfica (PAIDI-TEP-164) desde 1997 hasta la fecha.

Referencias

- Ariza-López, F. J., & García-Balboa, J. L. (2013). Evaluación de las componentes de la calidad de la información geográfica. En F. J. Ariza-López (Ed.), *Fundamentos de*

Gil de la Vega, P., Ariza López F. J., Mozas Calvache, A. T. (2016): "Problemas que presentan las trazas GNSS procedentes de Información Geográfica Voluntaria", *GeoFocus (Artículos)*, n°17, p. 161-184. ISSN: 1578-5157

evaluación de la calidad de la Información Geográfica (págs. 389-497). Jaén: Servicio de Publicaciones de la Universidad de Jaén.

- Ariza-López, F. J., Rodríguez-Avi, J., & Reinoso, G. J. (2014). An approximation to outliers in GNSS traces. *Spatial Accuracy 2014*, (págs. 186-189). Michigan.
- Arnold, J. A. (1998). Surface Transportation and Global Positioning System Improvements: L5 and DGPS. *Public Roads*, 62 (4), 1-10.
- Castelein, W., Grus, L., Crompvoets, J., & Bregt, A. (2010). A characterization of Volunteered Geographic Information. *13th AGILE International Conference on Geographic Information Science*, (págs. 1-10). Guimaraes.
- Cho, W. (2012). Data Quality Control: Crowd-Sourcing Geospatial Information. *Global Geospatial Information Management Forum*. Hangzhou.
- Goodchild, M. F. (2007). Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. *International Journal of Spatial Data Infrastructures Research*, 24-32.
- Luque Valle, P. (2011). *Análisis del modelo de uso-visita de los deportistas-turistas de las vías verdes andaluzas*. Granada: Editorial de la Universidad de Granada.
- MERCODES. (2008). *Consultoría y Asistencia para la Creación del Producto Cicloturismo en Vías Verdes*. Ministerio de Industria, Turismo y Comercio.
- Ruiz Almar, E. (2010). Consideraciones acerca de la explosión geográfica: Geografía colaborativa e información geográfica voluntaria acreditada. *GeoFocus* (10), 280-298.
- Zhang, L., & Sester, M. (2010). Incremental data acquisition from GPS-Traces. *A special joint symposium of ISPRS Technical Commission IV & Autocarto*. Florida.