

UNIVERSIDAD SAN FRANCISCO DE QUITO

Colegio de Ciencias e Ingeniería

Millón datos, 0 información: Generación de una metodología de análisis de datos de GPS, para analizar la entrega de supercongelados en la Ciudad de Quito

Felipe Xavier Bueno Almeida
Cristina Camacho, M.Sc., Directora de Tesis

Tesis de Grado presentada como requisito
para la obtención del Título de Ingeniero Industrial

Quito, diciembre de 2014

Universidad San Francisco de Quito

Colegio de Ciencias e Ingeniería

HOJA DE APROBACIÓN DE TESIS

Millón datos, 0 información: Generación de una metodología de análisis de datos de GPS, para analizar la entrega de supercongelados en la Ciudad de Quito

Felipe Xavier Bueno Almeida

Cristina Camacho, M.Sc.
Directora de Tesis

Carlos Suárez, Ph.D.
Miembro de comité de Tesis

Danny Navarrete, M.Sc.
Miembro de comité de tesis

Ximena Córdova, Ph.D.
Decana de la Escuela de Ingeniería
Colegio de Ciencias e Ingeniería

Quito, diciembre de 2014

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído la Política de Propiedad Intelectual de la Universidad San Francisco de Quito y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo de investigación quedan sujetos a lo dispuesto en la Política.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo de investigación en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Firma:

Nombre: Felipe Xavier Bueno Almeida

C. I.: 1719210195

Fecha: Quito, diciembre de 2014

DEDICATORIA

A mi familia, mis profesores y mis amigos.

AGRADECIMIENTOS

A mi madre por nunca dejarme dar por vencido.

A mi papá, por ser siempre la motivación de ser más y confiar siempre en mí.

A mis hermanas, que son las sonrisas de mi vida.

A mis profesores por guiarme en este hermoso camino llamado educación. En especial a Ximena Córdova, Cristina Camacho y Daniel Merchán.

A mis amigos, que han estado y estarán siempre en la risas y las penas.

A Andrés Jarrín Gerente de Jarygom por abrirme las puertas de su empresa.

RESUMEN

Se recopiló una base de datos de rastreo de GPS de la flota de la empresa Jarygom, esta tiene 15 camiones y la base de datos es de 7 meses en los que se entregan helados a tiendas y autoservicios en la ciudad de Quito. Siguiendo el método científico se generó una metodología que guía en el proceso de generar la base de datos, limpiar, analizar y graficar los datos. Se programó en Matlab todos los códigos para llevar a cabo la metodología antes descrita. Se obtuvieron indicadores relacionados a distancia, tiempo de movimiento y velocidad de 1344 rutas. Se generó propuestas de mejora para la empresa, que incluyen la eliminación de ciertas rutas, creación de nuevas políticas y proyectos de mejora para el futuro. En adición a esto, se encontró evidencia que valida la importancia del estudio de Logística Urbana en la ciudad de Quito.

ABSTRACT

A GPS tracking database fleet of 15 trucks of the company Jarygom was compiled, the database is seven months long in which ice cream is delivered to shops and supermarkets in the city of Quito. Following the scientific method, a methodology to guide the process of creating the database, clean, analyze and graph the data was generated. All codes are programmed in Matlab to implement the methodology described above. Indicators related to distance, time of movement and speed of 1344 routes was obtained. Suggestions for improvements to the company, including the elimination of certain routes, creating new policies and improvement projects for the future was generated. In addition to this, evidence was found that validates the importance of the study of Urban Logistics in the city of Quito

RESUMEN.....	7
ABSTRACT	8
CAPITULO 1. INTRODUCCION	14
1.1 Antecedentes.....	14
1.1.1. Descripción de la empresa.....	14
1.1.2. Distrito Metropolitano Quito (DMQ)	15
1.2 Justificación	16
1.3. Dificultades del estudio de Logística Urbana.....	17
1.4. Objetivo General.....	18
1.5. Objetivos Específicos	18
1.6. Alcance	19
1.6.1. Factores incluidos	19
1.6.2. Factores excluidos	19
CAPITULO 2. MARCO TEORICO Y REVISION LITERARIA.....	20
2.1 Marco teórico.....	20
2.1.1. GIS y GPS	20
2.1.1.1. GIS.....	20
2.1.1.2. GPS.....	20
2.1.2. Logística	23
2.1.2.1. Redes Logísticas	24
2.1.2.2. Concepto Logística urbana	24
2.1.3 Matlab.....	25
2.1.4. Notepad++	29
2.2 Revisión de Literatura	29
2.2.1. Géneros de Literatura Incluidos en la Revisión.....	30
2.2.2. Metodología de la Revisión.....	30
2.2.3. Uso del GPS y GIS en Logística Urbana.....	31
CAPITULO 3. METODOLOGÍA Y DISEÑO DE LA INVESTIGACIÓN.....	40
3.1. Justificación de la metodología	40
3.2. Descripción de la metodología	40
3.3. Aplicación del método científico para este estudio en específico.	41
3.4. Fuentes y recolección de datos.	43
3.5. Equipo de trabajo.....	43
CAPITULO 4. ANÁLISIS DE DATOS Y PRESENTACIÓN DE RESULTADOS	44
4.1 Obtención de datos.	44
4.2 Construcción de base de datos.....	45

4.3 Lectura de Base de Datos en Matlab	48
4.4. Limpieza de datos.	50
4.4.1. Rutas fuera de Quito	52
4.4.2 Datos nocturnos	56
4.4.3. Lugares de parqueo.....	58
4.4.4 Paradas largas	62
4.5. Análisis de datos.....	65
4.5.1 Análisis realizado en Matlab	66
4.5.2 Migración archivos de Matlab a Excel	72
4.6. Validación y Análisis de resultados.	75
4.7. Esquema de metodología.....	104
4.7.1 Construir	104
4.7.2. Limpiar	107
4.7.3. Analizar	108
4.7.4. Validar	108
4.7.4. Analizar	109
CAPITULO 5. CONCLUSIONES Y RECOMENDACIONES	110
5.1 Hallazgos significativos del estudio	110
5.2. Limitaciones del estudio.....	111
5.3. Recomendaciones	111
5.4. Futuros estudios.....	112
CAPITULO 6. TRABAJOS CITADOS	113
ANEXO 1	116
ANEXO 2	124
ANEXO 3	128
ANEXO 4	137
ANEXO 5	145

Tabla 1. Resultados de la data de GPS en Madrid.....	36
Tabla 2. Resultados de la data de GPS en Soria, España.	36
Tabla 3. Diferencia de datos entre corte 1 y corte 2.	47
Tabla 4. Tipos de datos de la base de datos consolidada.....	47
Tabla 5. Latitud y longitud de vértices del perímetro urbano.	54
Tabla 6. Rutas con distancia y tiempo 0.....	77
Tabla 7. Resumen de análisis de mapeo de rutas con distancia y tiempo igual a 0.....	78
Tabla 8. Rutas con distancia menor a 1 km.	80
Tabla 9. Resumen de análisis de mapeo de rutas menores a 1 km..	81
Tabla 10. Percentiles en intervalos de 5% de rutas mayores a 1 km.	83
Tabla 11. Muestreo aleatorio de rutas entre el 0 y 15 percentil.....	84
Tabla 12. Resumen de análisis de mapeo muestra aleatoria de rutas en el 0 y 15 percentil..	85
Tabla 13. Percentiles en intervalos de 5 de los datos con distancia mayor a 12 km..	87
Tabla 14. Muestreo aleatorio de rutas entre el 85 y 100 percentil.....	88
Tabla 15. Resumen de análisis de mapeo muestra aleatoria de rutas en el 85 y 100 percentil.	89
Tabla 16. Muestreo aleatorio de rutas entre el 40 y 60 percentil.....	90
Tabla 17. Resumen de análisis de mapeo muestra aleatoria de rutas en el 40 y 60 percentil.	91
Tabla 18. Promedio y desviación estándar de métricas obtenidas.	92
Tabla 19. Coeficiente de variación. Fuente: Elaboración propia.	93
Tabla 22. Tipos de datos de la base de datos consolidada.....	107

Figura 1. Constelación de satélites de GPS	21
Figura 2. Ubicación de dispositivo GPS con 1 satélite.....	22
Figura 3. Ubicación de dispositivo GPS con 2 satélites.....	22
Figura 4. Ubicación de dispositivo GPS con 3 satélites.....	23
Figura 5. Ejemplo de la función estructura en Matlab para 1 dato con 3 campos.....	27
Figura 6. Ejemplo de código de la función estructura en Matlab.....	28
Figura 7. Ejemplo gráfico y de código de la función estructura en Matlab para 2 datos con 3 campos.	28
Figura 8. Proceso de construcción, limpieza y análisis de la base de datos.	44
Figura 9. Proceso de construcción de la base de datos.....	46
Figura 10. Código de programa para leer base de datos.....	49
Figura 11. Código de programa para equivaler fechas.....	50
Figura 12. Situaciones en las que se limpiara los datos.	52
Figura 13. Mapeo georeferenciado de rutas originales.....	53
Figura 14. División parroquial del DMQ.	53
Figura 15. Georeferenciación gráfica de límites urbanos del DMQ.	54
Figura 16. Código de programa de limpieza de datos fuera del perímetro urbano del DMQ.. ..	55
Figura 17. Mapeo de rutas después de la limpieza de rutas fuera del perímetro urbano del DMQ.....	56
Figura 18. Código de programa de limpieza de datos nocturnos.	58
Figura 19. Georeferenciación gráfica de centro de distribución y centro de operaciones... ..	59
Figura 20. Código de programa de limpieza de datos dentro de centro de distribución y centro de operaciones.	60
Figura 21. Flujograma de programa de limpieza de paradas largas.	63
Figura 22. Código de programa de limpieza de datos dentro de paradas largas.	64
Figura 23. Proceso de programa de análisis de datos.....	66
Figura 24. Código de programa de análisis y almacenamiento de datos.....	67
Figura 25. Estructura de almacenamiento de datos.. ..	71
Figura 26. Código de programa de consolidación de métricas en matrices independientes.....	72
Figura 27. Proceso de migración de datos de Matlab a Excel.....	73
Figura 28. Código de programa de escritura de matriz de Matlab a un archivo de texto....	74
Figura 29. Archivo de texto abierto en Notepad++. Fuente: Elaboración propia.	75
Figura 30. Gráfico de dispersión con líneas rectas de la distancia de las 1610 rutas.	76
Figura 31. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las 1610 rutas.	76
Figura 32. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 0 km.	79
Figura 33. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las rutas mayores a 0 km.....	80
Figura 34. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 1 km.	82
Figura 35. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las rutas mayores a 1 km.....	83
Figura 36. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 12 km.	87
Figura 37. Gráfico de dispersión con líneas rectas de tiempo de duración de las rutas mayores a 12 km.....	87
Figura 38. Gráfico de caja de las distancias.	93

Figura 39. Gráfico de caja de los tiempos de duración de las rutas..	94
Figura 40. Prueba de normalidad de tiempos de duración entre 27 y 54 km.	95
Figura 41. Prueba de normalidad de tiempos de duración entre 54 y 90 km.	96
Figura 42. Histograma de tiempos de duración entre 27 y 54 km.	97
Figura 43. Histograma de tiempos de duración entre 54 y 90 km.	97
Figura 44. Gráfico de dispersión de crecimiento de distancias..	99
Figura 45. Gráfico de dispersión de tiempos de duración.	99
Figura 46. Gráfico de residuales de residuales de distancias..	100
Figura 47. Gráfico de residuales de residuales de tiempos de duración.	101
Figura 48. Mapeo de la ruta #1316 camión 338 el 13 de junio de 2014..	102
Figura 49. Mapeo de la ruta #1323 camión 338 el 21 de junio de 2014.	103
Figura 50. Pasos principales de la metodología de análisis.	104
Figura 51. Proceso de la fase "Construir"	105
Figura 52. Encuesta de identificación de rutas.	105
Figura 53. Encuesta de identificación de tipo datos.	106
Figura 54. Pasos a seguir para consolidar la base de datos..	107
Figura 55. Flujo grama de limpieza de datos. Fuente:	108

CAPITULO 1. INTRODUCCION

En la actualidad es común la utilización de sistemas de rastreo satelital conocidos como GPS en las flotas de entrega y transporte. Estos sistemas recopilan gran cantidad de información que comúnmente no se utiliza. El siguiente proyecto planteará una metodología que permita analizar estos datos y que no se conviertan en un recurso subutilizado de la empresa. Se va a utilizar como modelo la empresa Jarygom que se encarga de entregar helado en todo el Ecuador pero nos enfocaremos en la ciudad de Quito. La metodología pretende brindar soporte al momento generar la base de datos, limpiarla, analizarla y graficar estos datos. Se utilizara como herramientas principales Matlab y Excel.

1.1 Antecedentes

1.1.1. Descripción de la empresa

Unilever adquiere la fábrica de Helados Pingüino en 1996, posterior a esto deciden enfocarse en la parte central de su negocio que son la producción y venta de helados y como consecuencia decide subcontratar la distribución de la cadena primaria (Ciudad-Ciudad) y la cadena secundaria (Ciudad-Minorista) dentro del país. Dada esta situación se crea la empresa JaryGom que se encarga de dar el servicio exclusivo de transporte de supercongelados a Unilever, con camiones de hasta un máximo de 20 pies. Al momento Unilever posee tres distribuidores regionales (Quito, Portoviejo y Machala) y su planta de producción en Guayaquil. Toda las logística entre planta de producción, centros de distribución (CD) y minoristas son llevados a cabo por JaryGom (A. Jarrín, comunicación personal, 14 de enero de 2014).

JaryGom no posee ningún programa que le ayude a diseñar sus rutas, las realiza de forma empírica. En el último año JaryGom decidió mejorar la tecnología usada en el rastreo de sus camiones para lo cual dotó de sistemas de rastreo satelital a todos sus

camiones, además esta tecnología brinda respaldo cuando hay divergencias con Unilever sobre si llegó o no llegó el camión al lugar previsto a la hora prevista, para controlar excesos de velocidad, salidas de ruta o paradas muy extensas. Sin embargo el sistema de posicionamiento global solo se lo usa de forma reactiva y de control, nadie utiliza la información que esta almacena con el fin de optimizar el trabajo de entrega que realiza la empresa. (A. Jarrín, comunicación personal, 14 de enero de 2014)

1.1.2. Distrito Metropolitano Quito (DMQ)

El DMQ es actualmente la capital económica del país, el 45% de sus ingresos proviene de la venta de bienes y servicios y asciende a 65 millones de dólares “Invierta en Quito, 2014”, adicionalmente se encuentra en el puesto 34 de las mejores ciudades para invertir en América Latina según la revista América Economía (2013). Según el INEC: “Las cifras reflejan que en Quito existen 101.937 establecimientos económicos” (2011), estas cifras reflejan que la ciudad de Quito es una metrópoli pujante a nivel nacional e internacional.

El parque automotor en la ciudad es de 410,000 vehículos (Agencia Pública de Noticias de Quito, 2012); según el diario El Telégrafo este crece a mayor velocidad que la población, exactamente un 11% para el parque automotor y un 1.6% para la población en el 2013 (El Banco Mundial, 2014), esta es una de las razones por las cuales el tráfico y todos los problemas relacionados a un alto crecimiento vehicular se han acelerado en el último tiempo (El Mercurio, 2012). Debido a esta cantidad de vehículos se ha implementado algunos correctivos que van desde el pico y placa, hasta la construcción del metro de Quito, pasando por la reducción de cupos para la importación de vehículos. Por otro lado en relación a soluciones logísticas para el sector empresarial privado por parte del

Municipio no se aprecia mayores esfuerzos, el enfoque municipal en este momento es dar mejor servicio a la parte civil de la ciudad más que a la industria misma.

1.2 Justificación

La congestión es un problema a nivel mundial, su principal manifestación es la progresiva reducción de las velocidades de circulación, aumento de los tiempos de viaje, mayor consumo de combustibles, incremento de otros costos de operación y sin duda mayor contaminación atmosférica (Bull, 2003). El presente es el primer estudio en materia de logística urbana relacionado con sistemas de monitoreo GPS que se realiza en la ciudad de Quito. Estudios que utilizan datos de GPS se han llevado a cabo en mega ciudades como Sao Paulo, Madrid, Boston, etc. Aparte de esto el Municipio del Distrito Metropolitano de Quito ha definido la movilidad como uno de sus ejes estratégicos de acción (Agencia Pública de Noticias, 2010). En la ciudad de Quito se han tomado políticas reactivas para mejorar el tráfico dentro de la ciudad como el pico y placa que prometía reducir un 20% del tráfico vehicular el problema es que en menos de dos años el parque automotor creció este porcentaje (Editorial El Comercio, 2013). Esto nos dice que las medidas restrictivas son de muy corto plazo y lo que requiere la ciudad son análisis y políticas proactivas para realmente generar soluciones válidas. Ese es el objetivo de este estudio aportar conocimiento acerca de cuál es el sentir y la realidad del sector privado en relación al tráfico. Obtener indicadores de velocidad media, tiempo de viaje de entrega promedio, entre otros.

Hay 3 puntos que justifican la realización de este proyecto

JaryGom se encuentra en un momento en el que las herramientas de monitoreo por GPS no fueron lo esperado desde el punto de vista de optimización, sin embargo son una herramienta muy valiosa ya que poseen mucha información. Por esta razón se utilizará

dicha información para llevar a cabo un proceso de obtención de métricas e indicadores sobre la compañía para a su vez volverla más óptima acorde a la actualidad.

El tráfico en la ciudad de Quito requiere rutas y soluciones óptimas para perder la menor cantidad de recursos al momento de realizar una entrega.

Entender el verdadero potencial de los datos de GPS, en la actualidad son pocos los estudios sobre la relación entre logística y GPS. Por esta razón es una propuesta innovadora utilizar este tipo de datos para generar conocimiento y así contribuir a estudios futuros sobre el uso de datos de GPS para logística.

1.3. Dificultades del estudio de Logística Urbana

Las dificultades del estudio se centran en tres perspectivas, una pública y dos privadas. La primera es que no existen políticas para mejorar la distribución en el DMQ, pero si existen medidas restrictivas como pico y placa, y restricción por tonelaje para tráfico pesado, esto causa muchos problemas entre los encargados de mantener el orden y las empresas de distribución (La hora, 2012). Frente a esto no se ve ninguna intención de crear políticas o mejoras de infraestructura que faciliten la entrega de bienes en la ciudad de Quito.

La segunda perspectiva es la de Unilever, ellos son los dueños del producto, dueños de la imagen, responsables del negocio. Por esta razón ellos están interesados en tener el nivel de servicio más alto posible al menor costo (Jarrín, 2014). Ellos tercerizan la entrega, por esta razón requieren un servicio de primera calidad, con el menor costo y sin preocuparse de satisfacer ningún otro requerimiento más que el que sus clientes estén satisfechos con la entrega (Jarrín, 2014).

El tercer actor y sin duda el más importante, es JaryGom. Ellos son los encargados de satisfacer a los otros dos actores de este problema y a ellos mismos. Su problema es

como satisfacer la necesidad de tener un muy alto nivel de servicio (perspectiva Unilever), como no infringir las leyes y políticas nulas que propone la ciudad, aparte de la ya complicada situación geográfica y de tráfico que tiene el DMQ (perspectiva pública). Por último sus necesidades son mantener los costos lo más bajo posible satisfaciendo las otras dos constantes. (Jarrín, 2014)

Partiendo de la problemática expuesta, se procede a definir el problema de estudio. Se busca encontrar una forma más eficiente de entregar súper congelados en la ciudad de Quito, en términos de tiempo, costo y aplicabilidad dentro del DMQ, usando información de GPS y modelos matemáticos de optimización.

1.4. Objetivo General

Definir una metodología para analizar datos de GPS, que permita obtener información para entender cómo se comportan las empresas de logística urbana en la ciudad de Quito.

1.5. Objetivos Específicos

Graficar los datos de GPS de forma geo referenciada y dentro de un mapa de la ciudad de Quito para tener una referencia visual de la distribución y poder analizar las particularidades de los datos.

Generar una metodología de análisis de datos que se encargue de limpiar la data en base a funcionamiento particular de la empresa.

Obtener métricas como la velocidad promedio, distancia de ruta promedio, duración de ruta promedio, para realizar análisis de la movilidad urbana de flota a analizarse.

1.6. Alcance

1.6.1. Factores incluidos

Todos los puntos de GPS dentro de la ciudad de Quito, de la red secundaria de entrega de la empresa JaryGom.

Consolidación de la base de datos incluyendo las rutas de entrega desde marzo hasta agosto 2014, de 15 camiones.

Limpieza de datos en base al funcionamiento particular de Jarygom

Análisis estadístico de datos.

Código en Matlab de todos los programas utilizados para la limpieza y análisis de datos.

Legislación sobre tráfico del DMQ.

1.6.2. Factores excluidos

Cualquier dato fuera del DMQ.

Cualquier dato fuera de los meses previstos.

Cualquier entrega que no pertenezca a la red secundaria.

Encuestas o conteo de camiones, la única información acerca del comportamiento de los vehículos proviene de los datos de GPS

Datos de fin de semana, datos de paradas prolongadas, datos nocturnos, datos fuera del DMQ.

CAPITULO 2. MARCO TEORICO Y REVISION LITERARIA

2.1 Marco teórico

2.1.1. GIS y GPS

2.1.1.1. GIS

GIS por su siglas en ingles significa “Geographical Information System”, que traducido al español quiere decir sistema de información geográfico (Coppock & Rhind, 1991). Este integra hardware, software y datos para manejar, analizar y mapear todo tipo de información geo referenciada (USGS, 2007). “GIS nos permite ver, comprender, cuestionar, interpretar y visualizar los datos de muchas maneras que revelan las relaciones, patrones y tendencias en forma de mapas, globos terráqueos, informes y gráficos.” (ESRI, 2013)

La versión computarizada del GIS se utiliza desde 1960 pero el GIS en su forma manual se lo utilizaba 100 años antes por lo menos (Coppock & Rhind, 1991). La mayoría si no son todos, de los desarrollos tempranos en GIS fueron realizados en los EEUU. Sin embargo la mayor parte del avance se lo ha realizado en los años anteriores a los 90 (Coppock & Rhind, 1991). Vale aclarar que el sistema de GPS y GIS no es lo mismo, GPS utiliza todo el concepto de GIS para referenciar sus datos pero son dos sistemas diferentes.

2.1.1.2. GPS

Breve historia del GPS:

GPS por sus siglas en ingles quiere decir “Geographical Positioning System”, es una tecnología que se inició en el año de 1960, con el uso de dos satélites de la armada americana (DANA, 1997). Sin embargo fue en 1973 cuando comenzó el desarrollo de la tecnología como la conocemos hoy, lanzando el primer satélite de GPS en 1978 y en 1993 el último de los 24 que brindan cobertura a nivel mundial, claro que solo con fines militares. (DANA, 1997)

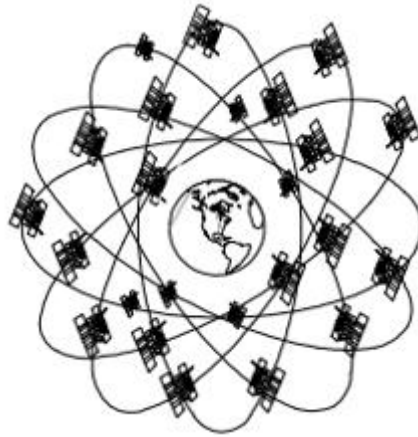


Figura 1. Constelación de satélites de GPS Fuente: (DANA, 1997)

En la imagen presentada anteriormente se pueden observar los 21 satélites operacionales que están en órbita y los 3 de repuesto, que recorren 6 planos orbitales, 55 grados de inclinación, órbitas de 20,200 kilómetros que demoran 12 horas aproximadamente (DANA, 1997).

En el año de 1987 se dio la posibilidad de usar GPS de forma comercial a civiles, con ciertas restricciones, 100 metros de precisión entre otras. (DANA, 1997), en la actualidad la restricción se ha reducido a 3 metros (National Coordination Office for Space-Based Positioning, Navigation, and Timing, 17).

Como funciona GPS:

“Un receptor GPS calcula su posición mediante la medición de la distancia entre sí mismo y tres o más satélites GPS utilizando un método llamado trilateración. En trilateración, una señal de microondas es enviada desde un satélite al receptor GPS, que mide el tiempo que tomó para la señal para llegar a él. Debido a que la señal viaja a una velocidad conocida, el receptor utiliza entonces la longitud de tiempo de viaje de la señal para calcular una gama de posibles ubicaciones circulares.” (International Biological System Information, 2003, pág. 1)

Con un satélite hay infinitas posibilidades de ubicación del dispositivo como se muestra en la siguiente figura

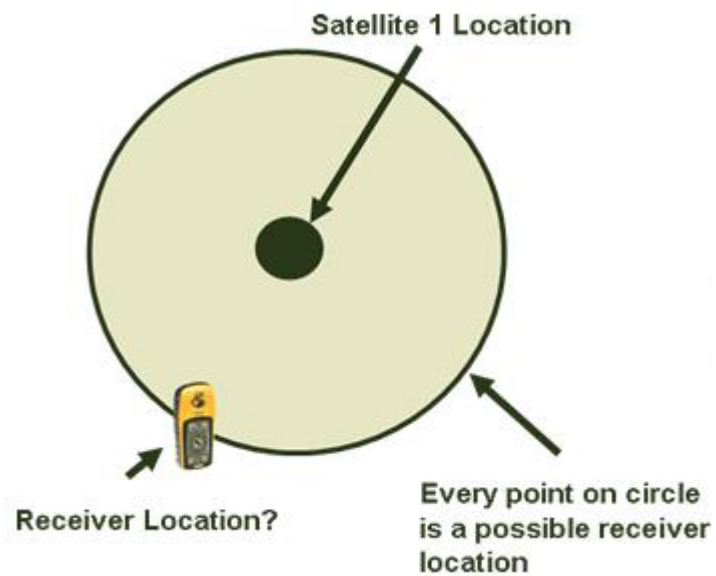


Figura 2. Ubicación de dispositivo GPS con 1 satélite. Fuente: (International Biological System Information, 2003)

Con dos satélites se obtiene dos posibles ubicaciones como muestra la siguiente figura.

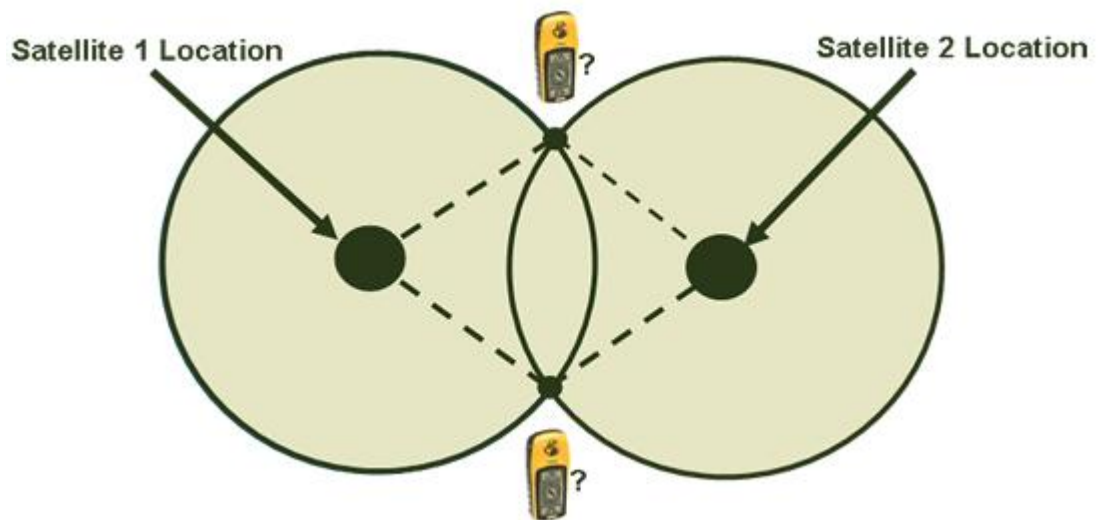


Figura 3. Ubicación de dispositivo GPS con 2 satélites. Fuente: (International Biological System Information, 2003)

Y con 3 ya se obtiene la ubicación exacta como se muestra en la siguiente imagen.

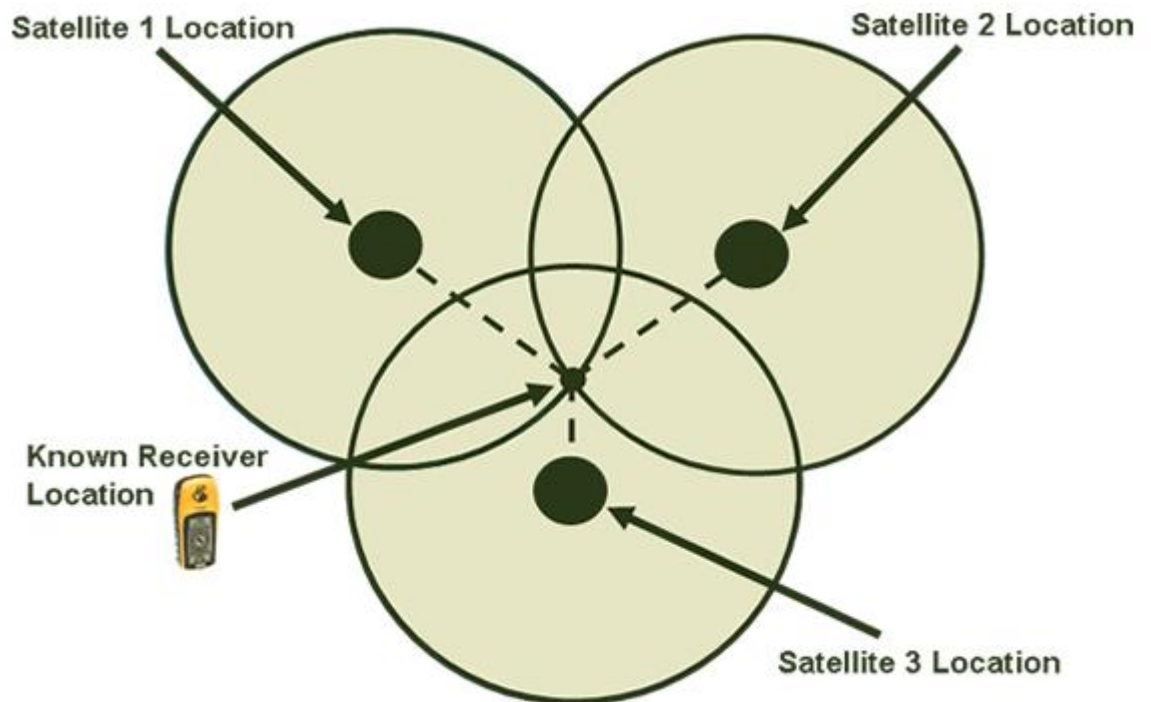


Figura 4. Ubicación de dispositivo GPS con 3 satélites. Fuente: (International Biological System Information, 2003)

Por esta razón se necesita tres satélites con línea de vista para conocer la ubicación exacta del dispositivo. (DANA, 1997)

Sistemas de monitoreo GPS, son los dispositivos que utilizan el principio de GPS para enviar información en tiempo real, sobre ubicación, velocidad entre otros (DANA, 1997). Los sistemas de monitoreo están en auge ya que cada vez el hardware es menos costoso. Vale decir que son sistemas reactivos, solo si algo sucede ellos comunican, no son preventivos. (National Coordination Office for Space-Based Positioning, Navigation, and Timing, 17)

2.1.2. Logística

Una definición simple y válida de lo que es logística se puede apreciar en la siguiente cita

“Es el proceso de planear, implementar y controlar el flujo y almacenamiento eficiente y a un costo efectivo de las materias primas, inventarios en proceso, de producto terminado e información relacionada, desde los puntos de origen hasta los de consumo; con el propósito de satisfacer las necesidades de los clientes.” (Mora Garcia, 2008, pág. 7).

Es básicamente llevar de forma eficiente de un punto A un punto B, lo que el cliente desea satisfaciendo sus requerimientos específicos (calidad, tiempos, precio, entre otros) (Ghiani, Laporte, & Musmanno, 2004).

2.1.2.1. Redes Logísticas

“El proceso de planeación de la red logística consiste en el diseño del sistema a través del cual partes y productos fluyen desde proveedores hasta los puntos de demanda” (Ghiani, Laporte, & Musmanno, 2004). Es el camino por el cual los productos van a transitar para cumplir su ciclo. El ciclo desde el punto de vista de cadena de suministro es desde que son materia prima hasta su entrega final o en el caso de una empresa de distribución de alimentos desde su CD hasta la tienda o auto mercado determinado, muchas veces no el cliente final (Mora García, 2008).

“Las redes logísticas son la estructura física de un sistema logístico. Los sistemas logísticos están determinados por tres actividades principales: procesamiento de órdenes, manejo de inventario y transporte.” (Cordova, 2013).

2.1.2.2. Concepto Logística urbana

La logística urbana se preocupa de las actividades de recolectar y entregar mercancías dentro de la ciudad y los centros de ciudad. (Directorate-General for energy and Transport, 2006). La logística urbana está compuesta por procesos de transporte, manipulación y almacenamiento de mercancías, la gestión de inventario, los residuos y

devoluciones, así como los servicios de entrega a domicilio (aunque muchos de estos no se den dentro de los límites geográficos de la ciudad si afectan a la misma).

La razón por la cual la logística urbana es razón de estudio en la actualidad es porque de la entrega de mercancías dentro de la ciudad se nutre el 18% del tráfico de las ciudades, además de que es el último eslabón de la cadena de suministro por lo que sus entregas se hacen en paquetes pequeños, lo que conlleva a que hay que recorrer más kilómetros para satisfacer la demanda (Directorate-General for energy and Transport, 2006). También como dice la ONU, desde el 2008 es el primer año que más gente va a vivir en áreas urbanas que rurales (UN, 2007). A esto se le suma que para el 2020, la distancia recorrida por los vehículos de entrega se incrementará en un 41% en la ciudad de Múnich (Ehmke, 2012).

2.1.3 Matlab

Matlab es un entorno de computación numérica, que permite analizar cualquier tipo de datos (Matworks Inc, 2014). En este proyecto toda la limpieza, el análisis inicial de los datos y la graficación de rutas será llevado a cabo en Matlab. Matlab tiene gran cantidad de funciones predefinidas que facilitan el análisis de los datos. Todas las funciones que se utilizaron en el proyecto se detallan a continuación por orden de aparición.

clc y clear: limpian cualquier operación previa que se haya realizado en el programa. Esto es una práctica que se realiza para que no haya contaminación cruzada entre las operaciones anteriores y las nuevas.

cd: “current directory”, indica el directorio donde se debe buscar los archivos que el programa desee abrir mas adelante

TXT: Indica que los datos deben ser escritos como texto

RAW: Indica que los datos son “crudos” o sin modificar

Length: Indica que el lazo o bucle “for” debe recorrer todo el “largo” o la cantidad de filas que tiene la matriz la forma que se debe utilizar es la siguiente

length(NOMBREMATRIZ)

datenum: convierte una fecha al valor entero y fracción de días tomando en cuenta una fecha fijada en este caso (0 de enero del 0000)

Vector de tiempo: provee la siguiente información

[ñ][][h][] .

Unique: Elimina todos los datos duplicados, para mantener el orden de aparición se debe utilizar el parámetro “stable”.

Find: lo que hace es devolver los valores que no son cero dentro de la matriz X. La operación “find” también permite colocar una condición como por ejemplo $I = \text{find}(X < 10)$, que lo que quiere decir es que encuentre todos los valores en “X” que sean mayores a 10 y los almacene en “I”.

Datevec: crea un vector de 6 posiciones para un valor de fecha, que se encuentra como entero y fracción de días, tomando en cuenta una fecha fijada. Para el caso de Matlab la fecha es el 0 de enero del año 0000 [ñ][][h][] .

Distance: ”. Los parámetros de la función son los siguientes “distance(lat1,lon1,lat2,lon2)”, esta función arroja dos parámetros “[arclen,az]” que permitirán en un futuro transformar la distancia a cualquier unidad. “Arclen” es la distancia de los arcos del gran círculo que conectan los pares de puntos en la superficie de una esfera y “az” es el azimuth entre los dos puntos.

deg2km: lo que esta función realiza es transformar una distancia que se encuentra en grados a km, en este caso está en el parámetro “arclen” que se explicó anteriormente,

utilizando como referencia el radio de un gran círculo que en este caso es el de la tierra, de 6971. (Mathworks, 2014).

Cat: Los parámetros de la función son los siguientes $C = \text{cat}(\text{dim}, A, B)$, el parámetro “dim” puede ser 1 o 2 lo que hace es llenar la matriz manteniendo las dimensión horizontal o vertical respectivamente, es este caso como se quiere llenar la matriz hacia “abajo” se debe mantener la dimensión horizontal. El siguiente parámetro es “A”, que en este caso es la nueva matriz que se está llenando y “B” que es el dato “i” que aprobó toda la condición antes descrita” (Mathworks, 2014).

Estructura: Para almacenar la información se utilizó una “matriz estructura”. Según Mathworks “Una estructura es un tipo de data que agrupa datos relacionados utilizando contenedores de datos denominados campos. Cada campo puede contener datos de cualquier tipo o tamaño” (2014). La siguiente figura ilustra lo que se dijo anteriormente, utilizando un ejemplo donde la estructura es paciente, los campos son nombre, cuenta y exámenes.

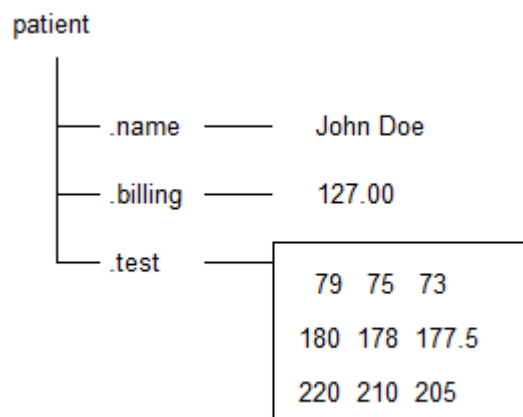


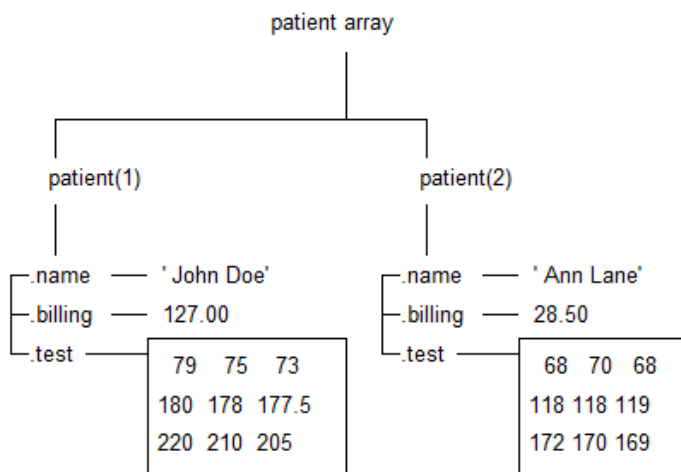
Figura 5. Ejemplo de la función estructura en Matlab para 1 dato con 3 campos. Fuente: (Matlworks Inc, 2014)

Y la forma de escribir esto en el código se ilustra en la siguiente imagen:

```
patient.name = 'John Doe';
patient.billing = 127.00;
patient.test = [79, 75, 73; 180, 178, 177.5; 220, 210, 205];
patient
```

Figura 6. Ejemplo de código de la función estructura en Matlab. Fuente (Mathworks Inc, 2014)

Si se desea aumentar otro paciente lo que se debe hacer es lo siguiente:



```
patient(2).name = 'Ann Lane';
patient(2).billing = 28.50;
patient(2).test = [68, 70, 68; 118, 118, 119; 172, 170, 169];
patient
```

Figura 7. Ejemplo gráfico y de código de la función estructura en Matlab para 2 datos con 3 campos.

Fuente: (Mathworks Inc, 2014)

Fopen: esta función permite abrir archivos que no son de Matlab para su futura modificación, los parámetros de la función son los siguientes: “fileID = fopen(filename,permission)” . Primero se escribe el nombre del archivo y la extensión que se desea en este caso es un archivo de texto y su nombre es “tiempo.txt”, el siguiente parámetro es de suma importancia porque es el que permite que se cree este archivo o si el archivo ya está creado se borre toda la información previa. Este parámetro que permite hacer lo antes descrito se identifica de la siguiente manera “ ’w’ ” (Mathworks Inc, 2014).

Fprintf: parámetros de la función “fprintf” son los siguientes:

“fprintf(fileID,formatSpec,A1,...,An)”, el primer parámetro es el nombre de identificación del archivo donde se va a escribir, en este caso es “fid”, luego viene el formato en el que se va a escribir. El formato varía dependiendo si es numérico, textual, cuantos decimales se necesita escribir. Para definir eso se utiliza la siguiente notación “%5.6f”, lo que quiere decir es que se escriba un numero con 5 dígitos y 6 dígitos después de un numero decimal, esto varía dependiendo de las necesidades. Lo siguiente es escribir el salto de línea en el que se utiliza 'n', un dato importante para futuros estudios es que si se planea abrir utilizando el “Notepad” de Windows se debe escribir '\r\n' en vez de 'n'. El siguiente parámetro es especificar qué es lo que se va a escribir.

Fclose: Cierra el archivo que Matlab modificó.

2.1.4. Notepad++

Programa para leer archivos de texto, programa de uso libre. Se utiliza este programa porque muchas veces cuando se escribe un archivo de un programa para que lo lea otro, los formatos varían. “Notepad++” facilita la comunión entre un programa y otro, por lo que simplifica el trabajo. En este caso entre Matlab y Excel (Ho, 2014).

2.2 Revisión de Literatura

El presente capítulo tiene como objetivo presentar trabajos similares que se han realizado alrededor del mundo en el área de uso de sistemas GIS con enfoque a logística urbana. Estos dos temas pocas veces han sido tratados de forma conjunta, como la siguiente cita lo dice “Otra categoría es la de recolección de datos basado en GPS. Aunque es muy popular en el transporte de personas, esta segunda categoría de métodos no ha sido todavía mucho más desarrollado para los bienes urbanos y sólo estudios preliminares se han llevado a cabo recientemente” (Gonzalez-Feliu, Pluvinet, Serouge, & Gardrat, 2013,

pág. 1). Existen investigaciones de uso de información de GPS en ciudades pero no han trascendido a buscar mejoras, es un área relativamente nueva el uso de información de GPS para entender cómo se mueven las mercancías, tráfico y vehículos en las ciudades. En adición a esto pocas veces datos provenientes de sistemas de GPS han sido usados para generar modelos de optimización. Esta integración de ramas propuesta en este proyecto es algo nuevo pero que sin duda es el futuro.

La revisión literaria inicia con una breve explicación de los géneros incluidos en la misma, para después pasar a definir la metodología utilizada en la investigación y por último hablar de las investigaciones que se encontró que aportaban con ideas importantes al proyecto.

2.2.1. Géneros de Literatura Incluidos en la Revisión

La revisión literaria se basó en artículos académicos que fueron publicados en journals peer review. La mayoría de artículos en relación a uso de GPS provienen de las conferencias de City Logistics. Cuando vamos al tema de logística urbana, las fuentes son más amplias pero un alto porcentaje proviene de estudios que se han realizado en Europa, que se podría decir que está en la primera línea en estudios en relación a logística de ciudades (Greaves & Figliozzi, 2008). Por último en relación a modelos de optimización e investigación de operaciones, las fuentes son más variadas, ya que los modelos de investigación son usados por muchas más ramas de la ciencia.

2.2.2. Metodología de la Revisión

Los temas elegidos para la revisión literaria son una selección de papeles e investigaciones que tienen relación a los temas a tratarse en el proyecto. Gran parte de estos papeles fueron sugeridos por profesores de la USFQ y estudiantes del Megacities Logistics Lab (MLL) que se encuentra en el Massachusetts Institute of Technology (MIT).

La otra parte de los papeles fueron buscados en bases de datos como ProQuest, Ebsco, Science Direct. También se utilizaron fuentes alternas de consulta como son libros y otras fuentes electrónicas (Kovács & Spens, 2011), otra fuente complementaria son tesis que tratan de temas similares.

Las palabras claves que se utilizaron (mayormente) para realizar la búsqueda, son combinaciones o palabras de los temas que interesan en la ejecución del proyecto, como por ejemplo: GPS, urban logistics, TSP, GPS logistics.

2.2.3. Uso del GPS y GIS en Logística Urbana

En el papel publicador por Michael F. Goodchild titulado SIG y Transporte: Situación y Desafíos, vemos que la aplicación de sistemas de información geográficos (SIG en español o GIS en inglés) ha sido utilizada en relación a transporte desde 1960 (2000). Aunque en relación a transporte se ha usado desde hace mucho tiempo GIS, Michael James Saunders y Antonio Nelson Rodríguez da Silva, en el papel publicado en International Journal of Sustainable Transportation hablan sobre la subutilización de estas herramientas “Reducciones dramáticas en la dependencia energética del transporte urbano aún no están siendo observadas a pesar de la variedad de herramientas de planificación urbana actualmente disponibles” (2009). Por esta razón Saunders y Rodríguez (2009) dicen que se debe utilizar herramientas modernas que permitan modelar con mayor precisión este tipo de fenómenos. “Un nuevo marco de desarrollo urbano se presenta para hacer frente a este problema que hace uso de una herramienta basada en GIS recientemente desarrollado y probado satisfactoriamente” (Saunders & Rodríguez da Silva, 2009).

A esto se suma lo que se dice en la investigación escrita por Sigrid Reiter y Anne-Francoise Marique, que habla de ciudades con bajo consumo de energía específicamente aplicada a Liege, Bélgica en el cual utilizan herramientas basadas en el uso de sistemas de

información geográfica (SIG) en combinación con un tratamiento estadístico de los criterios urbanos y de transporte (2012) el uso de GIS les permite:

“La metodología nos permite modelar la construcción y el uso de energía para el transporte en la escala de la ciudad, así como a considerar la posible evolución de la ciudad, el consumo de energía y para simular los efectos de algunas de las estrategias de renovación urbana” (Reiter & Francoise, 2012, pág. 829)

La metodología usada por Saunders & Rodríguez da Silva (2009) para encontrar el TES (Energía de especificación del transporte), que es una herramienta para medir la dependencia de energía del transporte en función de su ubicación geográfica (2009), es la siguiente:

- “1. Importación de datos de mapas
2. Adición de datos de uso de la tierra
3. Creación de la red
4. Generación de la trayectoria más corta
5. Simulación de modificaciones en el área urbana” (Saunders & Rodrigues da Silva, 2009)

La conclusión a la que se llegó al llevar a cabo la simulación en Sao Paulo, es que se encontró reducciones en el consumo de energía de hasta un 20% (Saunders & Rodrigues da Silva, 2009).

En otro estudio encontrado se puede ver como integran modelos de contaminación ambiental ya existentes con programas GIS para identificar la locación, donde los estándares de contaminación del aire y de ruido son violados (Chakraborty, Schweitzer, & Forkenbrock, 1999). Si a estos límites geográficos se los relaciona con las características raciales y económicas se llega a conclusiones de qué tipo de factores afectan a la contaminación en las ciudades (Chakraborty, Schweitzer, & Forkenbrock, 1999).

Hasta el momento se ha hablado de algunos casos de éxito sobre el uso de datos GIS y GPS en relación a logística, ahora se expondrá cuáles son los límites o las barreras que presenta el uso de datos GIS, un tema muy importante pero poco estudiado, el siguiente papel llamado *Barriers to GIS Use in Planning*, habla sobre estas barreras. Las limitaciones de los sistemas basados en GIS se dice que son tres tecnológicas (falta de programas para realizar lo requerido), factores organizacionales (falta de apoyo y comprometimiento de la dirección) y cuestiones institucionales (leyes que restringen el uso de este recurso) (Gocmen & Ventura, 2010). De estos tres limitantes el tecnológico es el que menos limita, normalmente son problemas de la organización o temas institucionales (Gocmen & Ventura, 2010), por esta razón se debe comprometer a la gerencia cuando se decide aplicar este tipo de herramientas ya que son costosas y de demorada aplicación.

Por otro lado tenemos los estudios que ya han utilizado información de GPS para su investigación. Iniciaremos hablando del papel “GPS data analysis for understanding urban goods movement” de Pascal Pluvineta, Jesus Gonzalez-Feliua, Christian Ambrosinia (2012) en el que se habla sobre la capacidad de los datos de GPS con transmisión de datos en tiempo real por medio de celulares inteligentes, de brindar información que permita caracterizar y diagnosticar el movimiento de bienes dentro de la ciudad. El estudio se lo realiza en la ciudad de Melbourne, Australia a 30 camiones (Pluvineta, Gonzalez-Feliua, & Ambrosinia, 2012). En el papel se habla de varios temas muy importantes al momento de utilizar datos de GPS, uno de ellos es cuál es el objetivo de recolectar los datos.

“Hay que señalar que los objetivos de la recogida de datos GPS no siempre son similares, y tres usos principales se pueden identificar. La primera de ellas permite el suministro de información al sondeo de vehículos para probar una tecnología o una metodología con el fin de realizar un seguimiento de los vehículos y recoger toda la

información deseada. Estas obras suelen tratar con pocos vehículos y una pequeña cantidad de datos. El segundo se refiere a la recogida de datos para la calibración del modelo y pruebas. En este caso, muy popular en la logística de la ciudad, el número de vehículos puede ser limitado y los datos obtenidos son fácilmente identificables y analizables. La tercera es menos usada y en relación con las encuestas de recopilación de datos para la caracterización UGM” (Pluineta, Gonzalez-Feliua, & Ambrosinia, 2012, pág. 451).

También habla acerca de porque el uso de dispositivos de GPS es muy superior al antiguo método de las encuestas que se entregaban a los conductores, como se dice en el papel la tasa de no respuestas y de formularios incompletos es muy alta, por esta razón casi solo el 15% de las encuestas enviadas sirven para analizar (Pluineta, Gonzalez-Feliua, & Ambrosinia, 2012). “Es importante señalar que los datos recogidos en la planificación logística de la ciudad se puede alimentar a los modelos de comportamiento para ayudar a los tomadores de decisiones públicas y privadas” (Pluineta, Gonzalez-Feliua, & Ambrosinia, 2012, pág. 452). Esto es crucial porque en base a este tipo de estudios se crean las políticas públicas y se exponen las necesidades de todos los actores. Por último este estudio arroja algunas métricas que son interesantes y dan la medida de lo que se puede lograr con un correcto manejo de los datos de GPS, por ejemplo tenemos que el viaje promedio de todos los camiones es de 184 km y dura 8.5 horas, que dos tercios viaja alrededor de una velocidad de 33 km/h y el otro tercio pasa parado (haciendo entregas, cargando combustible, etc.). Otro dato muy interesante es que el promedio de velocidad más bajo se observa de 7:30 a 9:30 y de 15:30 a 18:00, a esto se le suma que el 80% de los viajes inician solo en 8 horas de 7:30 a 15:30. Lo anotado da una interesante idea de la información que pueden brindar los datos de GPS para tomar decisiones o cambio de políticas.

Otro estudio es el realizado en la Universidad Politécnica de Madrid que se llama “A GPS Analysis for Urban Freight Distribution” escrito por Julio Comendadora, María E. López-Lambasb, Andrés Monzónb (2012). Este estudio trata sobre la integración de tres herramientas, los datos de GPS, encuestas a los conductores y formularios basados en observación. Se lo realiza con camiones de 3.5 toneladas de capacidad de carga en dos ciudades de diferente tamaño en España, Soria y Madrid, para comparar los resultados entre ciudades e intentar encontrar diferencias validas en función del tamaño de mercado. (Comendadora, López-Lambasb, & Monzónb, 2012).

“En la actualidad existe un consenso general de que las ciudades dependen en gran medida de la actuación armónica y eficiente de sus sistemas de transporte, en el que el transporte urbano de mercancías desempeña un papel clave. De hecho, los flujos de carga comprenden alrededor de una cuarta parte de tráfico urbano (Dablanc, 2007), por lo que son responsables no sólo del tráfico urbano, pero también de factores externos (por ejemplo, impactos ambientales indeseables), como la contaminación, el ruido, la congestión y los accidentes (Visser et. al, 1999). (Comendadora, López-Lambasb, & Monzónb, 2012, pág. 521).

Esto claramente nos dice lo crucial que es el estudio del tráfico de mercancías dentro de la ciudad. Este estudio arrojó varios resultados interesantes entre ellos se encuentran los siguientes el grupo más grande que identificó el estudio es la clasificación de “otros” (58% Madrid, 34% Soria), y los dos siguientes grupos más significantes son construcción y servicios (Comendadora, López-Lambasb, & Monzónb, 2012). Esto es crucial ya que cuando se trata de caracterizar cualquier fenómeno se tiene que ser muy cuidadoso de incluir las categorías más representativas para que "otros" o "no identificados" no sea la categoría principal. Otro hallazgo muy interesante es que sin importar la diferencia de tamaño de las ciudades, la cuenta de camiones en una esquina

similar en las dos ciudades resultó ser muy similar, mientras que en las zonas comerciales a las afueras de la ciudad, la cuenta de camiones si fue mucho mayor la de Madrid a la de Soria (Comendadora, López-Lambasb, & Monzónb, 2012). Estos resultados fueron de las dos primeras herramientas a usarse en este estudio (encuestas a choferes, y formularios basados en observación) los datos de GPS arrojaron los siguiente indicadores

Tabla 1. Resultados de la data de GPS en Madrid. Fuente: (Comendadora, López-Lambasb, & Monzónb, 2012, pág. 530)

Freight category	Stops/day	Stop length/day		Distance between stops/day		Use of van/day: (driving time/ time stopped)
		average (min)	st.deviation	average (km)	st.deviation	
Courier	43.6	8.8	5.3	1.8	2.7	0.6
Food	5.1	22.5	14.0	15.3	19.5	0.8
Service	8.1	11.6	1.9	10.6	3.6	2.5
Construction	12.6	23.4	17.8	12.0	11.0	1.0
Self-employed	13.9	35.3	20.2	12.9	2.7	0.7

Tabla 2. Resultados de la data de GPS en Soria, España. Fuente: (Comendadora, López-Lambasb, & Monzónb, 2012, pág. 531)

Table 3. Results of the GPS data, Soria

Freight category	Stops/day	Stop length/day		Distance between stops/day		Use of van/day: (driving time/ time stopped)
		average (min)	st.deviation	average (km)	st.deviation	
Courier	31.3	7.3	10.7	0.9	1.0	0.7
Food	10.7	10.7	8.3	20.3	23.6	1.9
Service	8.2	16.1	3.7	3.9	2.2	1.0
Construction	8.8	19.2	37.8	7.2	1.5	1.2
Self-employed	8.8	14.8	9.3	7.3	8.5	1.2

Este es el poder del GPS, brinda información más exacta y más detallada.

Otro papel muy relevante al estudio a realizarse en esta tesis es el siguiente

“Recopilación de Información de Vehículos Comerciales de datos con tecnología pasiva Sistema de Posicionamiento Global” realizado en la ciudad de Melbourne, Australia en el

año de 2008 (Greaves & Figliozi). Al inicio de este papel dice algo muy importante que alienta a realizar estudios de este tipo y es lo siguiente:

“A pesar del papel fundamental desempeñado por el transporte de carga, los movimientos de mercancías urbanas no han recibido el mismo nivel de atención que el dado a los movimientos de la gente. La modelación del transporte de carga urbano sigue siendo un campo relativamente inmaduro en que la mayoría de los modelos y aplicaciones son de naturaleza altamente agregada” (Greaves & Figliozi, 2008, pág. 158)

Dice claramente que el estudio del transporte de carga urbano es un ámbito que requiere investigación, lo que motiva a seguir con este proyecto. El papel también dice que la mayoría de esfuerzos en relación a recolectar datos de GPS se lo ha realizado básicamente en movimientos entre ciudades más no dentro de las mismas (Greaves & Figliozi, 2008). El manejo de datos de GPS no es algo sencillo como Greaves y Figliozi lo dicen “El procesamiento de datos brutos de GPS en viajes discretos para un análisis significativo presenta muchos desafíos (2008, pág. 160)”. Ellos identifican 3 problemas relacionados con el manejo de datos de GPS. El primero es el volumen de datos que existe, por ejemplo son decenas de megabytes por una semana de datos, lo que requiere de modelos y algoritmos automatizados para analizarlos. Segundo, los datos brindados por un GPS son continuos, es difícil saber cuándo un viaje termina y empieza. Tercero, es de precisión, la primera es la que se conoce como “comienzo frio” que es el tiempo en que el GPS se demora en conectarse con todos los satélites para dar una lectura correcta, en este comienzo frio el GPS puede dar datos con errores de hasta 2 km, y otro inconveniente es cuando no se tiene línea de vista ya sea por un túnel, edificios altos u otros obstáculos. (Greaves & Figliozi, 2008). El equipo que realizó este estudio uso PYTHON para el manejo de los datos.

Los resultados que arrojó este estudio son muy interesantes. El primero es que cada ruta diaria fue en promedio de 184 km y duro un promedio de 8.5 horas, de las cuales dos tercios se movió a una velocidad promedio de 33km/h y un tercio permaneció parado (haciendo entregas, almorzando, poniendo gasolina). Otro dato muy importante es el promedio de paradas por ruta, en Melbourne es de 12, y para ciudades como Londres es 6, Calgary es 5.6, pero esto es un promedio dicen ellos, por ejemplo para la entrega de bebidas suaves el promedio de paradas es de 25 (Greaves & Figliozzi, 2008).

Las conclusiones del papel dicen que el potencial de recolectar información de GPS es muy alta ya que la mayoría de vehículos comerciales ya cuentan con estos instalados, sin embargo las compañías temen compartir información porque consideran que la misma es de alta confidencialidad, pero dice que esta mentalidad tarde o temprano debe cambiar. También dice que tener los datos es solo el primer paso para obtener información. (Greaves & Figliozzi, 2008)

Otro estudio muy relevante es “Processing GPS Data from Travel Surveys”, escrito por Stopher, Jiang y Fitzgerald. Este inicia diciendo algo muy importante, y es que la capacidad de recolectar datos de sistemas de GPS es casi ilimitada sea cual sea su propósito. Si bien es relativamente sencillo importar y tener una vista de estos puntos, es extremadamente difícil, sino imposible llevar a cabo cualquier tipo de análisis de datos significativo utilizando herramientas o programas estándares (Stopher, Jiang, & Fitzgerald, 2005). Otro papel interesante en relación al uso de datos de GPS es “GPS-based Data Production in Urban Freight Distribution” que trata acerca de las capacidades de la información obtenida de datos de GPS para caracterizar el movimiento de carga urbana y diagnosticar la implementación y aplicación de la introducción de transmisión en tiempo real de datos de GPS usando celulares inteligentes (Gonzalez-Feliu, Pluvinet, Serouge, & Gardrat, 2013). El papel hace un análisis de los estudios más significativos que se han

realizado hasta el año 2013 y nunca deja de hacer hincapié en lo moderno que es utilizar GPS en temas de logística urbana, haciendo una diferencia entre dispositivos fijos y móviles. Pluvinet y sus colaboradores dicen que uno de los mayores problemas en el diagnóstico de distribución urbana es la falta de datos. Dicen que si bien la encuestas dan información, no permiten realizar un diagnóstico profundo y que los dispositivos GPS han sido utilizados en su mayoría para transporte de tipo civil mas no para logística urbana (Gonzalez-Feliu, Pluvinet, Serouge, & Gardrat, 2013). Como conclusiones más importantes dice que no importa qué tipo de dispositivo se use si fijo o móvil, los dos brindan información con un nivel de precisión muy similar. También se dice que el uso de celulares inteligentes como dispositivo reduce el costo de implementación de estas iniciativas. Sin embargo al final se dice que no importa que metodología se use para recolectar los datos, que lo importante es estandarizar la recolección y transferencia de estos datos para desarrollar herramientas de análisis unificadas que permitan universalizar el uso de información de GPS (Gonzalez-Feliu, Pluvinet, Serouge, & Gardrat, 2013).

CAPITULO 3. METODOLOGÍA Y DISEÑO DE LA INVESTIGACIÓN

3.1. Justificación de la metodología

El estudio realizado utiliza el método científico. Se investigó sobre metodologías para el uso y manejo de datos de GPS y si bien se encontró técnicas y herramientas para la limpieza y el análisis de los datos no se encontró nada similar a una metodología. Esto se debe a que el análisis de datos de GPS es algo que se está utilizando en el último tiempo, por lo que aún no se ha estandarizado casi nada en relación a esto. Como se recolectan los datos es diferente, como se los limpia, como se los analiza, como se los presenta, todo varia de estudio a estudio. Por esta razón se decidió que para este proyecto en particular se va a proceder a usar el método científico, y utilizar los estudios en relación a temas similares para nutrir el estudio.

3.2. Descripción de la metodología

El método científico partiendo de su definición etimológica, método viene de griego “methodus”, que significa camino a seguir y científico del griego “scientia” que significa conocimiento. Uniendo estas dos definiciones el método científico quiere decir camino hacia el conocimiento. Una definición un poco más desarrollada de esto es el conjunto de pasos ordenados o etapas que se debe recorrer para obtener un conocimiento que sea válido para la comunidad científica todo esto utilizando herramientas que sean fiables y validadas (Hillier & Lieberman, 2006).

El método científico es un enfoque sistemático para la investigación. Todas las ciencias incluidas las sociales utilizan este método con mínimas variaciones. (Chang, 2007). Este método según Hillier y Lieberman tiene tres pilares que son los siguientes: se

sustenta en leyes que han sido deducidas por el hombre, utiliza modelos matemáticos para establecer relaciones entre las variables y por último nunca toma referencias absolutas, por el contrario la raíz de sus referencias nacen en la observación (Hillier & Lieberman, 2006).

Los pasos del método científico varían un poco de autor a autor pero los siguientes son los que más se adecuan al estudio.

Observación y definición del problema. “El investigador debe extraer la realidad de un problema mediante la recolección de datos y revisión documental a fin de formular a una representación mediante un modelo matemático y estudiar el fenómeno de la misma manera en que éste se muestra en la realidad” (Hillier & Lieberman, 2006)

Experimentación e inducción.- Al haber ya observado, obtenido los datos e información necesaria, se procede a experimentar hasta obtener una hipótesis, una idea, algo que resulte de esta observación tan detallada que se realizó al inicio.

Validación.- “Luego se diseñan experimentos adicionales para verificar la validez de la hipótesis en tantas formas como sea posible y el proceso se inicia de nuevo” (Chang, 2007, pág. 9)

Conclusiones.- Por último se debe llegar a concluir sea cual sea el resultado de la experimentación y validación. Para poder contribuir a la comunidad científica para futuras investigaciones.

3.3. Aplicación del método científico para este estudio en específico.

Como dice Chang todas las ciencias, incluidas las sociales usan el método científico con mínimas variaciones (2007). Y este estudio no es el caso, se debe realizar una customización mínima a lo anterior propuesto para que este método sea exitoso.

Observación y definición del problema. Este paso se basara en dos partes la primera que es en relación a la empresa y la segunda que es en relación a estudios similares que se

han llevado a cabo. En relación a la empresa se enfocara en obtener la mayor cantidad de información acerca de cómo se realizan las rutas en la actualidad donde creen ellos que se encuentran las partes más sensibles del proceso. Lo segundo en relación a la empresa es como obtener, manipular y entender los datos de GPS que es en verdad la base de esta investigación. La segunda parte que es investigar acerca de estudios similares en relación a datos provenientes de dispositivos de GPS, es crucial porque es la única forma de nutrir el estudio con herramientas validadas por estudios previos. También es muy importante esta revisión de estudios similares, porque el manejo de datos de GPS para encontrar indicadores y mejorar la generación de rutas es relativamente nuevo, así que es muy importante estar al tanto de hacia dónde van orientadas estas investigaciones. Ya con la información externa e interna la formulación del problema o los problemas será más apegada a la realidad.

Experimentación e inducción.- La parte de experimentación en este proyecto es muy extensa, ya que como no existe estandarización en la recolección de datos, ni existe un programa ni metodología ya establecidos que los analiza el proceso es de prueba y error lo que toma mucho tiempo. Se codificara en Matlab utilizando ideas que provienen del primer paso. Una vez que ya se pueda convertir esos datos en información se procederá a comenzar a formular una hipótesis o en este caso una propuesta de solución, una respuesta a los problemas que se plantearon en el paso anterior.

Validación.- Una vez ya planteada la hipótesis, se procederá a la validación de la misma, hasta encontrar una solución que haya superado todas las pruebas y sea la más óptima de todas las posibles. En este proceso de validación se buscara opiniones de expertos, análisis de sensibilidad y hasta pruebas en el campo.

Conclusiones.- Las conclusiones y recomendaciones se las hará de forma tradicional.

3.4. Fuentes y recolección de datos.

La recolección de datos se realizara en su totalidad de la plataforma virtual que brinda el servicio a JaryGom, que se llama Tracklink. Tracklink da el servicio de los dispositivos y tiene un portal web en el cual se puede bajar información histórica a igual que ver los vehículos en tiempo real. Vale recalcar que la plataforma no brinda ningún tipo de indicador que pueda ayudar a mejorar la eficiencia de las rutas, la plataforma solo brinda información para tomar medidas reactivas. Por ejemplo el portal web dice la velocidad instantánea, pero no dice velocidad promedio, lo que esto quiere decir es que puede identificar si existe un exceso de velocidad pero no se puede saber si la velocidad promedio está dentro de la media o no para tomar decisiones acerca de la ruta.

3.5. Equipo de trabajo

Participan del proyecto, Felipe Xavier Bueno Almeida, estudiante egresado de Ingeniería Industrial de la USFQ como autor e investigador principal. Cristina Camacho, profesora del departamento de Ingeniería Industrial como directora de proyecto. Ximena Córdova Decana del colegio “El Politécnico” y directora de tesis sustituta. Daniel Merchán, candidato a PHD en el Instituto de Tecnología de Massachusetts como experto en el tema. En adición a esto Andrés Jarrín, Gerente de Operaciones de JaryGom como el patrocinador del proyecto.

CAPITULO 4. ANÁLISIS DE DATOS Y PRESENTACIÓN DE RESULTADOS

El siguiente capítulo intenta mostrar todo el trabajo que se realizó desde la obtención de los datos hasta el momento que se presentó las métricas. En este se detallara todo el proceso, incluyendo la explicación de los códigos y todas la metodologías que se utilizaron para obtener información validad de toda la data.

El siguiente grafico muestra parte de los procesos que se llevaron a cabo.



Figura 8. Proceso de construcción, limpieza y análisis de la base de datos. Fuente: Elaboración propia.

4.1 Obtención de datos.

Los datos fueron obtenidos de la plataforma de monitoreo de GPS, esta se llama Tracklink. La siguiente imagen es una captura de pantalla de la interfaz del usuario de la página web donde se obtuvieron los datos.

El periodo de tiempo que se estudió es desde el 15 de marzo del 2014 hasta el 25 de agosto del 2014. Dentro de este periodo se hizo dos cortes, el primer corte se obtuvo el 23 de junio del 2014 y el segundo el 26 de agosto del 2014. Se bajó la información de 15 camiones.

Existieron varios inconvenientes al momento de intentar bajar la información del servidor de Tracklink. Es válido siempre tomar en cuenta este tipo de restricciones para no perder tiempo en futuros estudios. Los problemas más importantes con los que se tropezó fueron los siguientes.

Tracklink y la mayoría de sistemas de monitoreo no poseen los servidores, ni anchos de banda lo suficientemente potentes para poder bajar grandes cantidades de

información de una sola vez. Esto se da porque pocas veces el cliente revisa la información del pasado, son sistemas que se enfocan en monitoreo de tiempo real, mas no en dar información histórica.

Como se explica en el punto anterior no era posible bajar archivos de gran tamaño del portal de Tracklink, por esta razón se decidió obtener reportes mensuales por cada uno de los camiones. Como resultado se obtuvo 105 reportes, que pertenecen a 15 camiones por 7 meses.

Los formatos de los reportes del primer corte y del segundo corte no fueron similares, las actualizaciones de software a veces cambian el orden o las unidades, es importante validar que los datos estén siempre en el mismo formato cuando se va a utilizar información obtenida en diferentes fechas como se realizó en este estudio.

Al momento de exportar a Excel los reportes estos cambiaban el orden, o las unidades en relación a lo que se visualizaba en la página web de Tracklink.

4.2 Construcción de base de datos

Con los reportes ya listos, se unificó la base de datos. Para unificar la base de datos se realizó el siguiente proceso que se muestra de forma resumida en el siguiente gráfico.

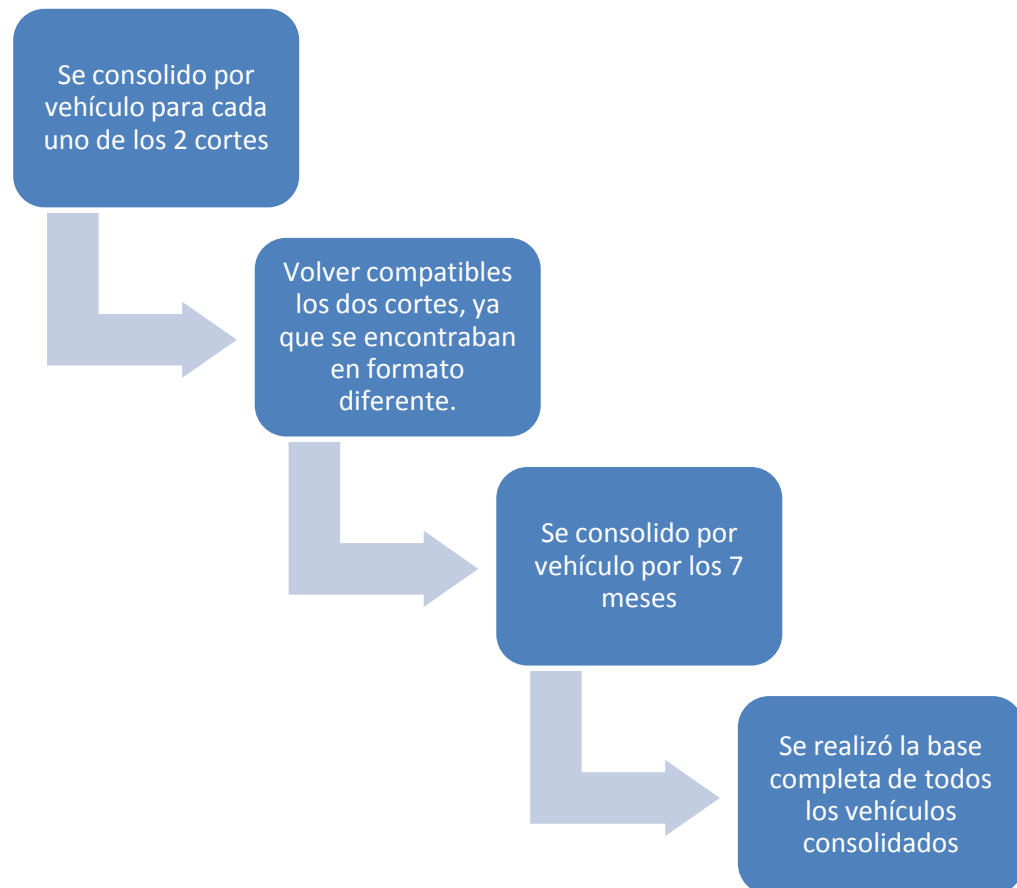


Figura 9. Proceso de construcción de la base de datos. Fuente: Elaboración propia

En adición a esto se tomaron ciertas precauciones con los datos, ya que tomo mucho tiempo obtenerlos. Se guardó un respaldo en Dropbox y en un disco duro externo de los 105 reportes. Después en cada paso que se modificaba un reporte se guardaba una memoria del mismo para tener siempre la versión anterior en caso de cometer cualquier error al momento de modificar. Esto fue de gran ayuda dado que al momento de revisar la última de las versiones de la base consolidada se encontró un problema y la solución estaba en la primera memoria. Fue un error que se cometió al inicio y salto a la vista al final sino se hubiera tenido estas memorias o respaldos de cada modificación hubiese sido imposible identificar cuando se cometió el error y como solucionarlo.

El primer paso que se llevó a cabo fue el de unificar por vehículo cada uno de los cortes para pasar de tener 105 reportes a solo 30, 15 del primer corte y 15 del segundo. En

la realización de este paso no hubo problemas significativos, ya que se manejó formatos y vehículos iguales.

Lo siguiente que se llevó a cabo fue consolidar los dos cortes. Durante la realización de este paso se hallaron algunos inconvenientes. El más significativo fue que hubo un actualización en el software de Tracklink por lo que los formatos cambiaron, se aumentó alguna información y se cambió otra, como lo muestra la siguiente tabla.

Tabla 3. Diferencia de datos entre corte 1 y corte 2. Fuente: Elaboración propia

Corte 1	Fecha	Hora	Latitud	Longitud	Velocidad	Odómetro	Distancia	GPS
Corte 2	Alias	Fecha Servidor	Fecha GPS	Latitud	Longitud	Odómetro (Km)	Velocidad	GPS

En el gráfico anterior podemos ver que los dos reportes proveen información ligeramente diferente. Entre las diferencias más importantes que se puede observar es que en el corte 1 los reportes brindaban la fecha y la hora en columnas separadas, y que para el corte número 2 horas y fecha eran un solo campo para solucionar este inconveniente se unificó los datos de los reportes del primer corte. Otra información diferente que se puede observar es que el corte número 2 provee dos fechas una del servidor y otra del GPS. Se decidió eliminar la del servidor y utilizar la del GPS aunque estos disten solo unos pocos segundos en su mayoría. Los datos del odómetro y la velocidad se eliminaron porque no estaban calibrados correctamente en estos GPS, porque los GPS eran más antiguos que el software por esta razón el hardware no proporcionaba toda la información que el software podía presentar. Se estandarizó los dos cortes como se muestra en la siguiente imagen.

Tabla 4. Tipos de datos de la base de datos consolidada. Fuente: Elaboración Propia

Alias	Código ID	Fecha	Lat	Lon	GPS
-------	-----------	-------	-----	-----	-----

La información que cada una de estas columnas da se explicará a continuación.

- Alias: nombre con el que se le conoce al vehículo dentro de la empresa.
- Código ID: Se generó un código de identificación de rutas. Este código posee tres datos que brindan dos campos de información muy importantes

para poder identificar a que ruta pertenece cada dato. El código está compuesto por [código de camión+mes+día]

- Fecha GPS: Es la fecha hora en la que se tomó el dato de GPS, está en el siguiente formato YYYY/MM/DD hh/mm/ss. Vale recalcar que esta es la hora real en la que se tomó el dato.
- Latitud: Representa la latitud en a que se tomó el dato esta en grados decimales.
- Longitud: Representa la longitud en a que se tomó el dato esta en grados decimales.
- GPS: Denota el estado del GPS, si este está en correcto, funcionamiento, guardado o dañado.

Una vez ya estandarizada la información de los dos cortes se procedió a unir los dos cortes para cada vehículo, en este ejercicio se redujo a 15 reportes de 7 meses cada uno. Lo que se realizó a continuación fue consolidar los 15 reportes en uno solo. El archivo final tuvo 424 mil datos (Alias, CódigoID, Fecha, Lat., Lon. y estado GPS) y aproximadamente un tamaño de 25 MB. El archivo de la base de datos es un archivo .xls más conocido como Excel.

4.3 Lectura de Base de Datos en Matlab

Antes de proceder con cualquier tipo de análisis era necesario que el programa que se decidió utilizar para el análisis en este caso Matlab sea capaz de leer la base de datos de forma correcta. Para realizar esto se utilizó el siguiente código.


```

1 -   clc
2 -   clear
3 -   cd C:\Users\Felipe\Documents\MATLAB\
4 -   [DATAOR, TXT, RAW]=xlsread('Consolidado2.xlsx', 'TOTAL');

```

Figura 10. Código de programa para leer base de datos. Fuente: Elaboración propia

La línea 1 y 2 solo limpian cualquier operación previa que se haya realizado en el programa. Esto es una práctica que se realiza para que no haya contaminación cruzada entre las operaciones anteriores y las nuevas.

La línea 3 utiliza la función “cd”, lo que esta hace indicar la dirección donde debe buscar todos los archivos que se van a utilizar (programas, archivos, funciones, entre otras)

Por último la línea 4 crea la matriz donde se va a guardar los datos y se provee ciertos parámetros. El primer parámetro es el nombre de esta matriz en este caso se asignó “DATAOR” que significa que estos son los datos originales. El siguiente parámetro “TXT” indica que los datos son escritos como texto y por último “RAW” indica que son datos sin modificar aun.

La matriz DATAOR está compuesta por 7 columnas y 426 631 filas en esta se encuentra la misma información que el archivo en Excel.

Al momento de ya leer los datos se encontró un problema crítico y tenía que ver con los sistemas de fechas que utiliza Excel por un lado y Matlab por otro. El sistema de Excel inicia el 1 de enero de 1900, mientras que el de Matlab inicia el 0 de enero del 0000. Para solucionar esto se debió escribir una rutina en Matlab muy simple que suma 1900 años y resta 1 día a la fecha que proporciona Excel y de esta forma se hace la conversión. El código es el siguiente

```

1 - for i=1:length(data1)
2 -
3 -     data1(i,3)=data1(i,3)+datenum([1900 0 0 0 0 0])-datenum([0 0 1 0 0 0]);
4 - end

```

Figura 11. Código de programa para equivaler fechas. Fuente: Elaboración propia.

La línea 1 inicia el “for” que va desde el dato 1 hasta el largo de la matriz DATAOR. La función “length” te permite ir hasta la última fila de la matriz sin tener que saber cuál es el número de filas exacto. Esto es muy útil cuando se utiliza varias matrices y sus tamaños son relativamente grandes.

La línea 3 lo que hace es abrir la matriz DATAOR en la fila i (cambia en función del loop for) y en la columna 3, que es donde se encuentra almacenada la fecha. Posterior a esto lo que se hace es sumar 1900 años y restar 1 día. Para esto se utiliza la función “datenum”, lo que esta hace es convertir un vector de tiempo a fracción de tiempo para poder realizar operaciones con este valor. Un vector de tiempo cuenta con 6 elementos que especifican lo siguiente año, mes, día, hora, minutos, segundos (Matlworks Inc, 2014).

El vector de tiempo provee la siguiente información

[año mes día hora minutos segundos]

Por último la fila se cierra el loop for.

Una vez realizada esta rutina los datos de Excel y de la matriz de Matlab son exactamente iguales y la base consolidada ya se encuentra lista para ser limpiada y analizada

4.4. Limpieza de datos.

Al momento de ya tener la base consolidada ya leída correctamente en Matlab, se empezó con la limpieza de datos. La mayor parte de la literatura se enfoca en problemas inherentes de los dispositivos de GPS. Como por ejemplo en el papel escrito por Gonzalez, Pluvinet, Serouge, Gardrat hablan sobre los comienzos fríos, que son el tiempo que los dispositivos se demoran en conectarse con los satélites y durante este tiempo dan

información errada (2013). Este no es el caso de los dispositivos que utiliza Jarygom, ya que estos dispositivos están constantemente enviando señales dado que su objetivo es monitorear a todo momento el estado del camión, en otras palabras no se apagan nunca. En el caso que se lleguen a apagar porque se terminó su batería interna los camiones siempre están prendidos por política de la empresa 3 minutos hasta calentar el vehículo y después salen a realizar su ruta. Según Fitzgerald, Stopher y Jiang los dispositivos demoran entre 15 y 40 segundos en conectarse con los satélites por lo que este no es un problema en este caso (2005).

Como se explicó anteriormente la mayoría de los problemas que se especifican en la literatura son relacionados directamente con el dispositivo de GPS, durante el estudio se encontró otra fuerte fuente de ruido o datos no validos que provienen del negocio mas no del dispositivo. Vale recalcar que la literatura siempre se ha enfocado en estudiar muchos tipos de industrias lo que no le permite encontrar problemas inherentes a una industria en particular dado que son muy numerosas y ese no es el objetivo de aquellos estudios. Por ejemplo Comendador, Lopez y Monzon en su estudio tienen 8 categorías de la cual en Madrid la más numerosa con un 58.12% fue no identificada (2012). Esto da la medida de porque es muy complicado entender cada giro de negocio cuando se analiza varios tipos de industrias.

En el presente estudio todos los datos provienen de una sola empresa lo que permitió encontrar varias fuentes de datos inherentes al funcionamiento de la misma, pero que no sirven para el estudio a realizarse. Los problemas que se encontraron inherentes al giro del negocio y que se va a atacar con la limpieza se presentan en la siguiente imagen.

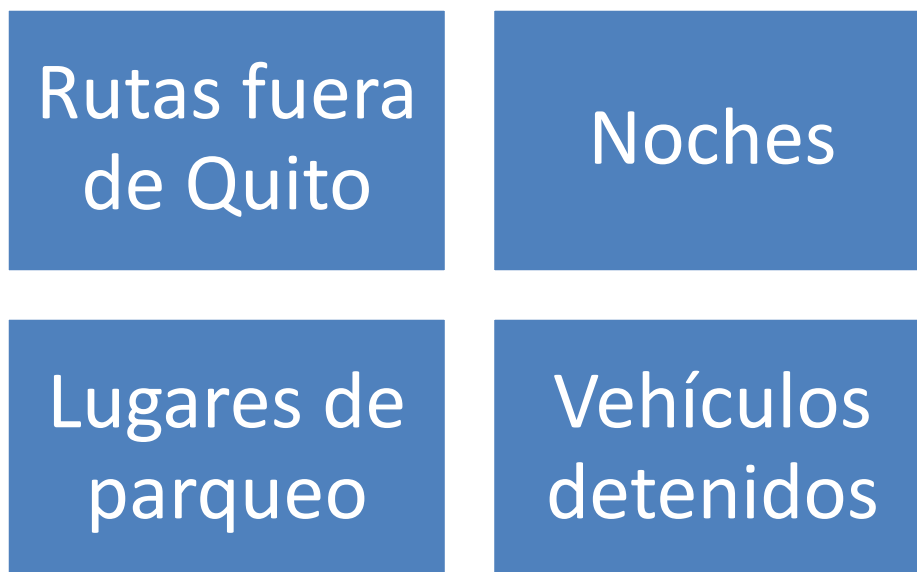


Figura 12. Situaciones en las que se limpiara los datos. Fuente: Elaboración propia.

4.4.1. Rutas fuera de Quito

Se iniciara hablando de las rutas fuera de Quito. Los datos se obtuvieron de los camiones que se encuentran en el centro de distribución (CD) de la ciudad de Quito por obvias razones, el objetivo es estudiar solo los camiones que se desplazan dentro de la capital. Al momento estudiar los datos, aparecieron problemas. Vale recalcar que la mayoría del trabajo fue prueba y error. Se probaba se encontraba un problema y se buscaba una forma de solucionarlo y se volvía a probar. El problema que se encontró es que los camiones que están en Quito de vez en cuando van a Guayaquil, Esmeraldas, Ibarra, Tulcán entre otras como lo denota el siguiente gráfico.

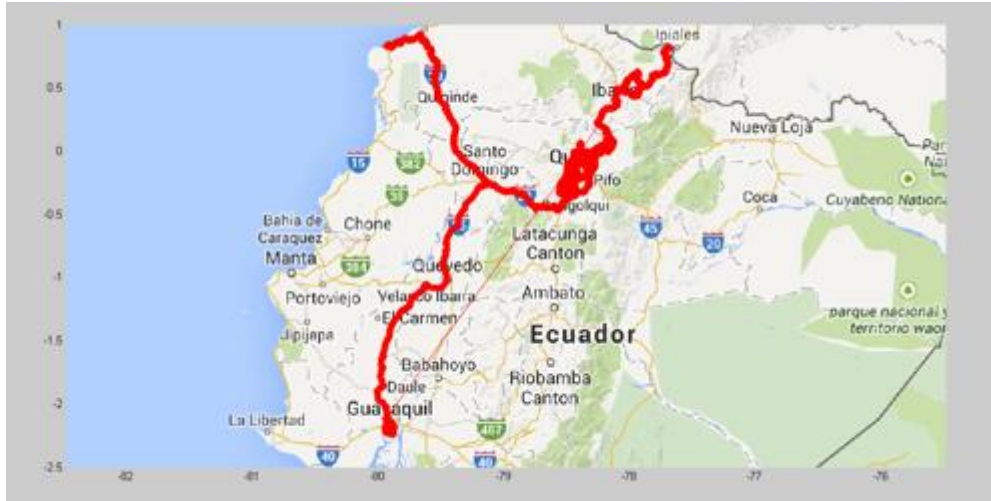


Figura 13. Mapeo georeferenciado de rutas originales. Fuente: Elaboración propia.

Como podemos ver en el gráfico los camiones no solo se mantienen dentro de Quito y esto se debe a que muchas veces hay pedidos extraordinarios, o simplemente ese camión realiza esa ruta fuera de la ciudad de forma periódica.

Una vez encontrado este problema lo primero que se realizó fue determinar el perímetro urbano sobre el que se desea trabajar. Para esto se investigó en los archivos municipales y se encontró un plano que explica claramente los límites urbanos de Quito como los muestra la siguiente imagen.

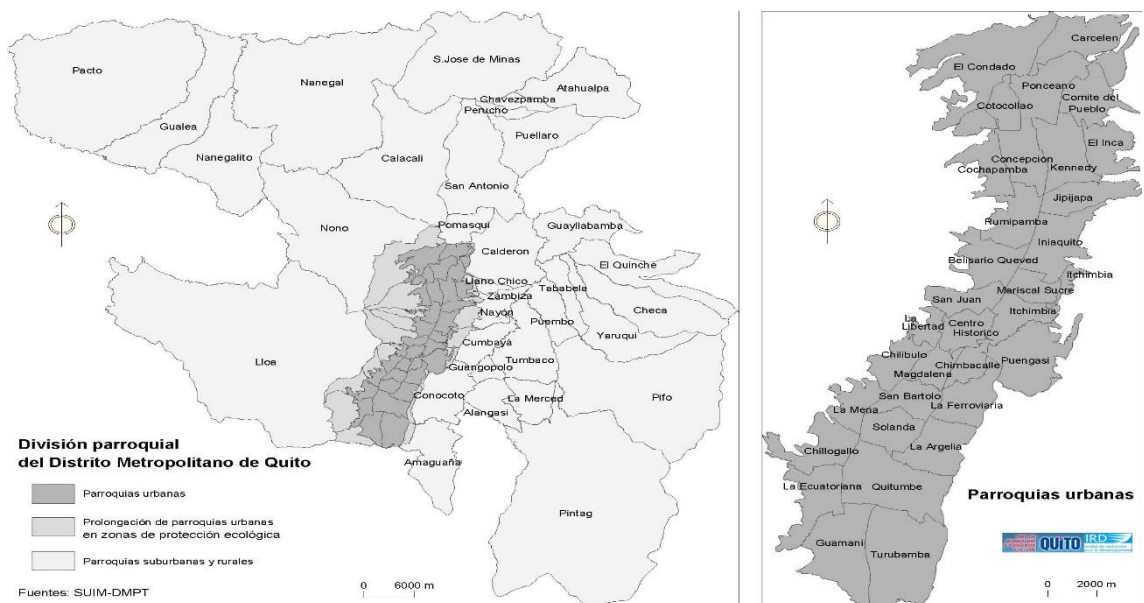


Figura 14. División parroquial del DMQ. Fuente: (Secretaría de Territorio, Habitat y Vivienda, 2014)

El siguiente paso fue georeferenciar los límites que proporciona el municipio, para esto se utilizó Google Earth como muestra la siguiente imagen.

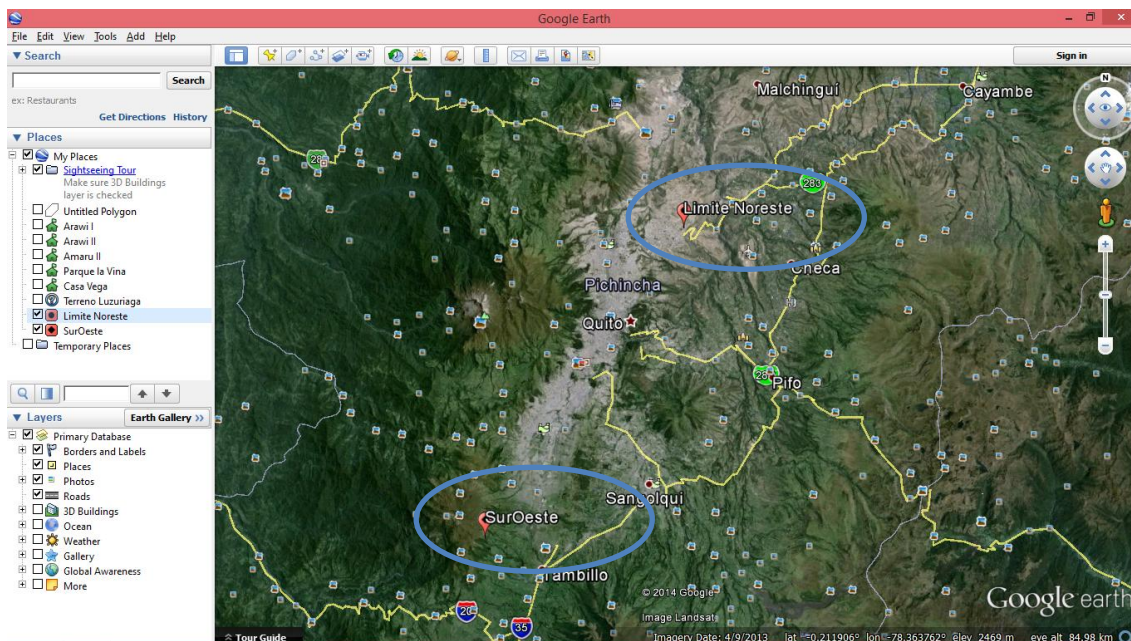


Figura 15. Georeferenciación gráfica de límites urbanos del DMQ. Fuente: Elaboración propia.

El resultado de límites se los aprecia en el siguiente cuadro:

Tabla 5. Latitud y longitud de vértices del perímetro urbano. Fuente: Elaboración propia

	Latitud	Longitud
Noreste	-0.091883°	-78.396893
Sureste	-0.377518°	-78.396893
Suroeste	-0.377518°	-78.631039
Noroeste	-0.091883°	-78.631039

Se realizó un rectángulo que incluye a todas las parroquias que se consideran urbanas según el municipio de Quito, es muy complicado poner el perímetro exacto del área urbana, porque son demasiados puntos y para luego utilizar ese perímetro se necesita de algoritmos mucho más avanzados. No se dice que no es posible pero los conocimientos para este proyecto no permitían ese tipo de manejo de información. Además como se

explicará más adelante en el programa de limpieza es mucho más simple si se utiliza un rectángulo.

Una vez ya definido el perímetro del cual se quería tener los datos se programó el código para realizar la limpieza a estos datos.

El código se muestra en la siguiente imagen:

```

1 - LON=[-78.631039 -78.396893];
2 - LAT=[-0.377518 -0.091883];
3 - perimetro=[];
4 - for k=1:length(data1)
5 -     if data1(k,5)<LON(1) || data1(k,5)>LON(2) || data1(k,4)<LAT(1) || data1(k,4)>LAT(2)
6 -         perimetro=[perimetro;data1(k,2)];
7 -     end
8 - end
9
10 - disp(['Puntos fuera de perimetro encontrados']);
11
12 - perimetro_fi=unique(perimetro);
13 - for i=1:length(perimetro_fi)
14 -     aa=find(data1(:,2)==perimetro_fi(i));
15 -     data1(aa,:)=[];
16 - end
17
18 - disp(['Dias con puntos fuera de perimetros eliminados']);

```

Figura 16. Código de programa de limpieza de datos fuera del perímetro urbano del DMQ. Fuente:

Elaboración propia.

A continuación se procederá a explicar lo que significa y lo que se busca obtener con el programa previo.

La línea 1 y 2 crea dos variables en las que se insertan los valores de longitud y latitud que se obtuvieron anteriormente.

La línea 3 crea la variable “Perímetro”, en esta variable se procederá a guardar todos los puntos que están fuera de este perímetro. El valor que se guarda es el código de identificación único que se creó inicialmente. Al tener ese código ya tenemos información de que ruta fue la que salió fuera del perímetro establecido.

Después se utilizó la función Unique, para eliminar todos los códigos duplicados ya que muchos puntos tienen el mismo código porque pertenecen a la misma ruta. Una vez ya con la lista de códigos sin duplicados esas son las rutas que se procederá a eliminar.

La función en la fila 13 tiene como objetivo eliminar todos los puntos de la matriz de datos original que su código es igual al que ya se encontró que sale del rectángulo establecido.

La limpieza dio los resultados esperados, se redujo la matriz original de 424 mil datos a 409 mil datos. Una forma más simple de entender lo que se consiguió es mirando el siguiente gráfico en la que mapea solo las rutas que respetan este perímetro.

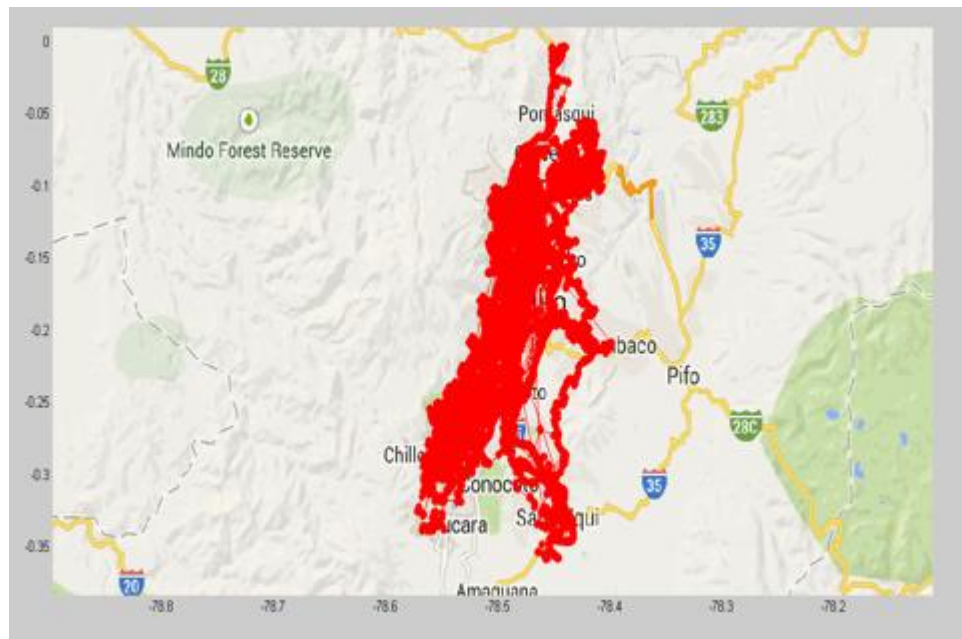


Figura 17. Mapeo de rutas después de la limpieza de rutas fuera del perímetro urbano del DMQ. Fuente: Elaboración propia.

4.4.2 Datos nocturnos

Otro problema que se encontró como se explicó anteriormente fue el que los dispositivos seguían enviando señales durante la noche. Esto causa que la base de datos se

llene de datos que no brindan ningún tipo de información válida y entorpecen el análisis de datos. Estos continuos datos que pertenecen a la noche se los atribuyo a dos causas. La primera es que dado que es un sistema de monitoreo al cliente le interesa que sus vehículos siempre reporten su ubicación así sea que no esté en movimiento, esta precaución no es estrictamente necesaria pero Jarygom solicito dado que es más fácil detectar si el dispositivo es desconectado durante la noche. Un ejemplo para explicar esta situación es si el dispositivo está configurado para dar una señal solo el momento que se mueve, consideremos que tiene un horario de 6am a 8 pm la última señal enviaría más o menos a las 10 pm y la siguiente señal a las 6 am del día siguiente. Desde el punto de vista de seguridad se podía haber desactivado el sistema a las 11 pm y en la central recién podrían darse cuenta a las 6 am que el camión no envía señales, dando valioso tiempo a los ladrones de usufructuar ilegalmente del camión. Por el otro lado si se recibe señales constantes es mucho más fácil identificar cualquier acto vandálico sobre el camión. La segunda razón por la cual existe esta gran cantidad de datos en la noche fue atribuida a que el proceso de carga de los helados a los camiones se los realiza en el transcurso de la noche desde las 11 pm hasta las 3 am, y como los camiones se tienen que mover durante este lapso los dispositivos están enviando señales continuamente, pero estas señales están dentro de un área que no brinda ningún tipo de información valida por esta razón se decidió eliminar este horario.

Para proceder a eliminar estos datos se construyó un programa en Matlab muy simple, el código es el siguiente:

```

1 - for i=1:length(data1)
2 -     date=datevec(data1(i,3));
3 -     if date(4)>=8 && date(4)<=17
4 -         data2=cat(1,data2,data1(i,:));
5 -     end
6 - end

```

Figura 18. Código de programa de limpieza de datos nocturnos. Fuente: Elaboración propia.

Lo primero que este programa realiza es abrir un lazo “for” que va desde el valor 1 hasta el último valor de toda la matriz en este caso data1. Como ya se explicó anteriormente la función “length” permite hacer eso.

Lo siguiente es crear un vector de 6 posiciones con la fecha del punto “i”, utilizando la función “datevec”.

Al momento de ya tener el vector ya se puede realizar cualquier tipo de comparaciones dado que cada posición del vector representa un valor, en este caso el que nos importa es el cuarto (4) que es representa la hora (vector tiempo)

La comparación que se va a realizar es muy simple, para esto se utilizara la función “if” que lo que va a realizar es comparar si el dato de la hora del punto “i” está dentro de la hora en la que la empresa hace entregas que son de 5am a 8 pm.

La siguiente fila lo que hace es guardar el dato que este dentro del parámetro antes enunciado.

Con este corto código se eliminó aproximadamente 88 mil datos con lo cual se pasó de tener 409 mil datos a 360 mil datos, lo que da cuenta de la significativa cantidad de datos que pertenecían a este horario que no se pretende estudiar.

4.4.3. Lugares de parqueo

El siguiente inconveniente que se presento fue que durante los fines de semana o días entre semana que los camiones no salían a hacer entregas, por diferentes razones ya

sea mantenimiento, no había suficientes entregas, o salían por cortos periodos de tiempo y pasaban la mayor parte del día parqueados. Esto es un problema similar al antes narrado acerca de las señales nocturnas, lo que sucede es que la base se sigue llenando de datos que no poseen información. Para intentar eliminar esto se definió dos lugares donde los vehículos pasan parados la mayor parte del tiempo. Para definir estos dos lugares se conversó con el gerente de operaciones Andrés Jarrín, se definió el centro de operaciones de Jarygom y el CD de Unilever. Las coordenadas geográficas se obtuvieron de Google Earth como se ve en la siguiente imagen.

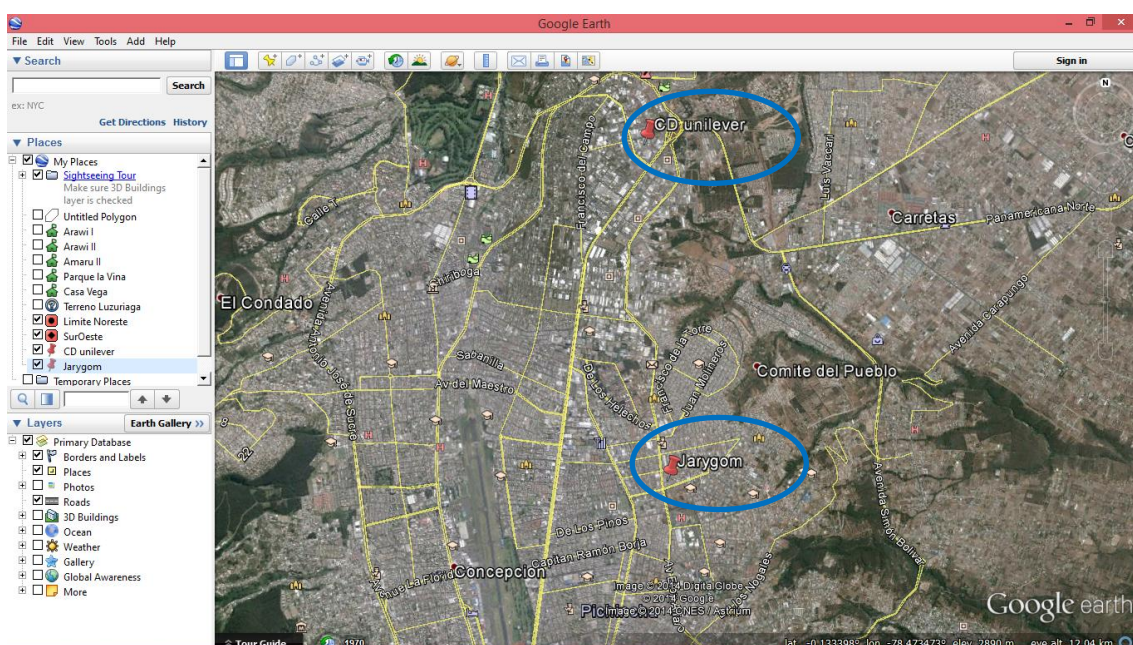


Figura 19. Georeferenciación gráfica de centro de distribución y centro de operaciones. Fuente:

Elaboración propia.

Las coordenadas son las siguientes:

Coordenadas Unilever:

-0.097467, -78.473034

Coordenadas Jarygom:

-0.132967, -78.470415

Una vez ya definido los dos lugares de parqueo se procedió a realizar un programa dentro de Matlab que encuentre todos los puntos que estén a una distancia de 30 metros de estos dos lugares. Esta distancia fue decidía en función del papel que escribe Stopher, Jiang y Fitzgerald en el que se establece que cualquier variación de 15 metros (0.00005°) en la latitud o longitud se considera el mismo lugar, como se está utilizando distancia entre dos puntos y no latitud y longitud como establece el papel se duplico este valor 30 metros (Stopher, Jiang, & Fitzgerald, 2005).

El código es el siguiente:

```

1 - for i=1:length(data2)
2 -     if deg2km(distance(-0.132967, -78.470415, data2(i,4), data2(i,5)))>=0.03 &&
3 -         deg2km(distance(-0.097467, -78.473034, data2(i,4), data2(i,5)))>=0.03
4 -             data4=cat(1,data4,data2(i,:));
5 -         end
6 -     end

```

Figura 20. Código de programa de limpieza de datos dentro de centro de distribución y centro de operaciones. Fuente: Elaboración propia

Como se explicó en el punto anterior se inicia con un lazo “for” que recorre todo el largo de la matriz en este caso es “data2”

Después utiliza la función “if” para ver si se cumple la siguiente condición que en palabras quiere decir lo siguiente, posteriormente se explicara el código. La condición es, comparar la distancia del CD y el centro de operaciones con el dato “i”, si la distancia del punto “i” con estos dos datos no es menor en ninguno de los dos casos a 30 metros el punto se guarda, si en alguno de los dos casos es menor (no puede ser los dos casos ya que la distancia entro los dos supera los 2 km), este punto se procede a desechar.

Lo siguiente explica cómo se escribió el código de esta condición que se acaba de narrar verbalmente.

Lo primero es encontrar la distancia en grados entre el punto “i” y uno de los dos lugares de parque que se estableció, para esto se utiliza la función “distance”. Los parámetros de la función son los siguientes “distance (lat1,lon1,lat2,lon2)”, en este caso los datos 1 son un lugar de parque y los datos 2 son el punto “i”, esta función arroja dos parámetros “[arclen,az]” que permitirán en un futuro transformar la distancia a cualquier unidad. “Arclen” es la distancia de los arcos del gran círculo que conectan los pares de puntos en la superficie de una esfera y “az” es el azimuth entre los dos puntos (Mathworks, 2014).

Lo siguiente que se realizó fue utilizar la función “deg2km”, lo que esta función realiza es transformar una distancia que se encuentran en grados a km, en este caso está en el parámetro “arclen” que se explicó anteriormente, utilizando como referencia el radio de un gran círculo que en este caso es el de la tierra que es de 6971 km (Mathworks, 2014).

Por último la distancia ya convertida a kilómetros se la comparo con la distancia de 30 metros (0.03 km) y para esto se utilizó el operador de comparación mayor o igual (\geq) (Cplusplus, 2014).

Dado que se quería comprobar que esta condición se cumpla para los dos lugares de parqueo se escribió dos veces la comparación antes explicada y se utilizó el operador lógico && que obliga para que sea cierta la condición se cumplan las dos condiciones en este caso que la distancia no sea menor a 30 metros en relación a ninguno de los dos lugares de parqueo (Cplusplus, 2014).

El último paso es guardar los datos que cumplieron las condiciones y para esto se utiliza la función “cat”. Los parámetros de la función son los siguientes $C = \text{cat}(\text{dim}, A, B)$, el parámetro “dim” puede ser 1 o 2 lo que hace es llenar la matriz manteniendo las dimensión horizontal o vertical respectivamente, en este caso como se quiere llenar la matriz hacia “abajo” se debe mantener la dimensión horizontal. El siguiente parámetro es

“A”, que en este caso es la nueva matriz que se está llenando y “B” que es el dato “i” que aprobó toda la condición antes descrita” (Mathworks, 2014). Vale especificar que se pone 1 en el valor de “di” porque cada dato “i” que aprueba la condición siempre debe ser el siguiente y con di 1 se logra que la matriz se llene hacia abajo y Matlab lee las matrices de arriba hacia abajo.

Con este programa se eliminaron 73 mil datos que no se requerían para el análisis, pasando de 360 mil a 287 mil datos. Siendo esta ya una base muy depurada pero sin embargo se encontró otro tipo de problemas con los estacionamientos que se explicara a continuación.

4.4.4 Paradas largas

El siguiente problema que se identificó, fue que los vehículos realizan paradas en otros lugares que no son identificables o varían, por lo que pierde lógica georeferenciar uno por uno, si se puede identificar estas paradas automáticamente. Se podría decir que es una forma diferente de solucionar el mismo problema que se presentó en el punto anterior. Un ejemplo de lugares de parqueo aleatorio es el siguiente: las mecánicas donde se arreglan los camiones. Las marcas más comunes que tiene Jarygom son Hino, Chevrolet, Mitsubishi, entre otras. Cada una de estas marcas tiene al menos 3 concesionarios diferentes en la ciudad de Quito. Esto nos da 9 lugares nuevos de parqueo, sin tomar en cuenta donde se hace arreglos de metalmecánica, o existen veces en que los camiones “duermen” en las casas de los choferes si bien es muy raro todo esto genera datos que no brindan información, porque como se explicó anteriormente los dispositivos siempre están enviando señales. Partiendo de esta premisa que muchos lugares de parqueo son aleatorios o es demasiado laborioso llegar a identificar y georeferenciarlos a todos, se tomó la decisión

de crear un programa que logre identificar estas paradas y que las elimine de la base de datos para obtener una depurada.

Para facilitar la explicación del programa se va utilizar un flujograma que explica con más claridad lo que hace el mismo, el flujograma de explicación se lo puede apreciar en la siguiente figura.

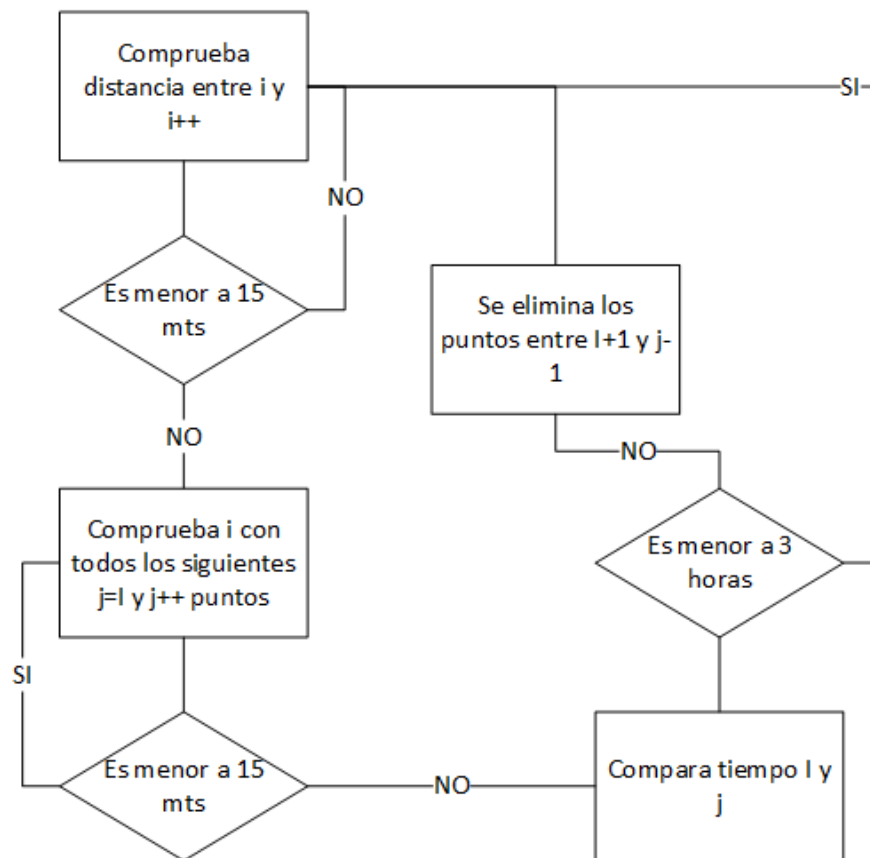


Figura 21. Flujograma de programa de limpieza de paradas largas. Fuente: Elaboración propia

El código del programa se muestra en la siguiente figura.

```

1 - for i=1:length(data3)
2 -     if deg2km(distance(data3(i,4), data3(i,5), data3(i+1,4), data3(i+1,5)))<=0.015
3 -         for j=i:i+500
4 -             if deg2km(distance(data3(i,4), data3(i,5), data3(j,4), data3(j,5)))<=0.015
5 -                 continue
6 -             else
7 -                 time=datevec(data3(j-1,3)-data3(i,3));
8 -                 if time(4)<=3
9 -                     break
10 -                else
11 -                    data3(i+1:j-1,:)=[];
12 -                    break
13 -                end
14 -            end
15 -        end
16 -        continue
17 -    end
18 - end

```

Figura 22. Código de programa de limpieza de datos dentro de paradas largas. Fuente: Elaboración propia

El programa compara dos variables conocidas que son distancia entre dos puntos y tiempo entre esos puntos, que en función de ciertos parámetros decide si se considera parada o no. El principio básico que se utiliza es que si existe un grupo de puntos que estén dentro de un radio de 15 metros y que el tiempo entre el primer punto y el último es de más de 3 horas, este cúmulo de puntos se considera una parada y se eliminan estos datos ya que no proveen información válida (Stopher, Jiang, & Fitzgerald, 2005).

El programa inicia con un lazo “for” que va a recorrer desde 1 hasta el último dato de la matriz que se intenta depurar.

Lo primero que busca que programa es la distancia entre dos pares de puntos consecutivos. Si esta distancia es mayor a 15 metros pasa al siguiente par de puntos consecutivos. Si la distancia es menor a 15 metros el programa pasa a la siguiente comparación.

La siguiente comparación se da solo si el par de puntos están a menos de 15 metros, lo que esta parte realiza es buscar cuantos puntos consecutivos siguen dentro de este radio de 15 metros.

Una vez encontrado los puntos que están dentro del radio preestablecido, se procede a comparar la diferencia de tiempo entre el primer punto y el último que están

dentro de este radio, si la diferencia es mayor a tres horas se procede a eliminar estos puntos si no lo son se procede a seguir comparando los pares de puntos consecutivos.

Un ejemplo para ilustrar la situación. El punto 1 y 2 se compara, si su distancia es mayor a 15 metros, se pasa a comparar los puntos 2 y 3, si su distancia es mayor a 15 metros se pasa a la siguiente comparación y así sucesivamente. Supongamos que la primera distancia menor a 15 metros se encuentra entre el punto 100 y 101. Lo siguiente que el programa hace es comparar la distancia entre el punto 100 y el punto 102, se supone que la distancia sigue siendo menor a 15 metros sino lo fuera volvería medir la distancia entre los pares consecutivos es decir 102 y 102, en este caso supones que fue menor. Lo siguiente es compara entre el punto 100 y el punto 103, se supone que son menores a 15 metros hasta el punto 120, y en el 121 ya la distancia vuelve a ser mayor a 15 metros. Cuando esto sucede se compara la diferencia de tiempo entre el punto 100 y el 120. Si el tiempo es menor a 3 horas se conserva los datos si es mayor se considera que es una parada larga y se los elimina. Y después se sigue comparando pares consecutivos hasta encontrar otra vez el fenómeno antes expuesto.

El programa solo encontró 1,000 puntos que cumplían estas condiciones así la base de datos paso de tener 287,000 a 286,000 datos.

Estos fueron los 4 programas que se utilizaron para depurar la base de datos. En el proceso de depuración se pasó de tener 424 mil a 286 mil, eliminado 138 mil datos que no brindan información válida para el estudio.

4.5. Análisis de datos.

Al momento de ya tener una base de datos “limpia”, el siguiente paso es analizar toda la data para obtener información, que a su vez permita tomar decisiones para mejorar la eficiencia, productividad, efectividad, dependiendo de qué área se decida atacar.

Las variables que se identificara con este programa son distancia recorrida por cada día-ruta, tiempo en el que el camión esta en movimiento. Para con esta información obtener indicadores acerca del comportamiento de toda la flota.

Lo que el programa realiza esta explicado en la siguiente figura.

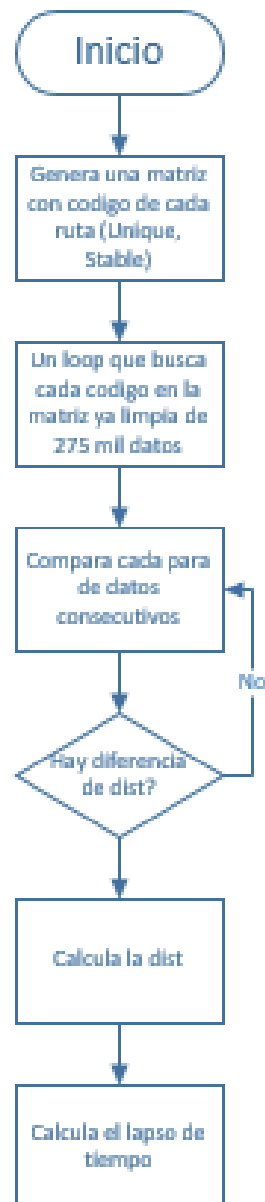


Figura 23. Proceso de programa de análisis de datos. Fuente: Elaboración propia.

4.5.1 Análisis realizado en Matlab

En términos simples lo que el programa realiza es encontrar la diferencia de distancias entre un par de puntos y si existe esta diferencia de distancias calcula el tiempo entre estos dos puntos y los va almacenando ruta por ruta. En la siguiente figura se observa el código que se utilizó, posterior al código se procederá a explicar paso a paso lo que este realiza.

```

1 -   clc
2 -   clear
3
4 -   load datafinal.mat
5 -   codigo=unique(data3(:,2),'stable');
6 -   for i=1:length(codigo)
7 -       aa=find(data3(:,2)==codigo(i));
8 -       time1=[]; dis1=[]; vel1=[];
9 -       for j=1:length(aa)-1
10 -          if data3(aa(j),4)==data3(aa(j+1),4) && data3(aa(j),5)==data3(aa(j+1),5)
11 -              else
12 -                  if data3(aa(j),3)==data3(aa(j+1),3)
13 -                      else
14 -                          timediff=etime(datevec(data3(aa(j+1),3)),datevec(data3(aa(j),3)));
15 -                          time=time+timediff/3600;
16 -                          time1=[time1;time];
17 -                          dis=deg2km(distance(data3(aa(j),4),data3(aa(j),5),data3(aa(j+1),4),data3(aa(j+1),5)));
18 -                          dis1=[dis1;dis];
19 -                          vel=dis/time;
20 -                          vel1=[vel1;vel];
21 -                      end
22 -                  end
23 -              end
24 -          S(i).tag=codigo(i);
25 -          S(i).vel=vel1;
26 -          S(i).dis=dis1;
27 -          S(i).time=time1;
28 -          S(i).meanvel=mean(S(i).vel);
29 -          S(i).sumdis=sum(S(i).dis);
30 -          S(i).sumtime=sum(S(i).time);
31 -          S(i).lon=data3(aa,5);
32 -          S(i).lat=data3(aa,4);
33 -       end

```

Figura 24. Código de programa de análisis y almacenamiento de datos. Fuente: Elaboración propia

Se inicia eliminando cualquier información previa que este en cualquiera de las matrices o la zona de trabajo a utilizarse usando las funciones “clc” y “clear”.

El siguiente paso es cargar la matriz de datos que se quiere analizar.

Con la matriz ya cargada y lista para usarse, se procede a realizar la primera operación que lo que hace es obtener una lista en orden de todas las rutas que existe en la base de datos ya depurada. Para esto se utiliza el código de identificación que se generó en la construcción de la base de datos que tiene la información de camión, mes y día. Como

todos los datos poseen este código se utiliza la función “unique”. Esta función lo que hace es eliminar datos duplicados o repetidos (Matlworks Inc, 2014). En este caso los duplicados son muy numerosos, porque si un día tiene mil datos todos estos datos tendrán el mismo código de identificación. Los parámetros de la función son la siguiente “C = unique(A)” en este caso “A” es todos los datos que se encuentran en la columna 2 de la base de datos (en este caso data3) y por eso se utiliza la identificación data3(:,2).

Aquí se identificó un problema con el código de identificación, el código del camión 335 del 31 de julio es 3350731 y el código del camión 335 del 1 de agosto es 335081. Como se puede observar el código de un día anterior es mucho mayor numéricamente hablando que el código del 1 de agosto. Al momento de aplicar por sí solo la función “unique”, este los reordena de mayor a menor, así que se desordenaban por este inconveniente en la creación de los códigos y se tenía los camiones mezclados, las fechas en desorden entre otros. Este problema causaba un caos porque no se podía concatenar ningún código de la matriz depurada con la lista de los códigos sin duplicados. Se encontraron dos soluciones para este problema, la primera era reformular el código de identificación, y luego volver a correr todos los programas de depuración, trabajo que se demora solo en computar el programa Matlab alrededor de 1 semana. La otra opción que se identificó fue el aumento del parámetro “stable” en la función unique, lo que este parámetro hace es mantener el orden original con el que van apareciendo los datos, esto solucionó el problema y permitió no tener que reformular la construcción del código y continuar el análisis de datos. Con la lista de códigos de identificación se procede al siguiente paso.

Lo siguiente es utilizar un lazo “for” que va desde 1 hasta el largo de la matriz “códigos”. Este lazo busca los datos para cada uno de los días y los almacena en una nueva

variable que en este caso se llama “aa”. Básicamente lo que hace es sacar todos los datos del camión X del día X a una nueva variable para poder analizarlos más fácilmente.

Para realizar lo antes descrito se utiliza la función “find”, los parámetros de la función son los siguientes: “I = find(X)” y lo que hace es devolver los valores que no son cero dentro de la matriz X. La operación “find” también permite colocar una condición como por ejemplo `I = find(X<10)`, que lo que quiere decir es que encuentre todos los valores en “X” que sean mayores a 10 y los almacene en “I”. Partiendo de la idea que el “find” debe cumplir una condición se escribió lo siguiente `“aa=find (data3 (:,2) ==código (i))”`, que lo que hace es buscar todos los valores en la columna 2 de la matriz “data3” sean iguales al valor que se encuentra en la matriz “código (i)”, con esto se extrae todos los datos del día “i” de la base de datos depurada.

Una vez que ya se tiene los datos del día “i” ahora se procede a analizar solo esos datos, que se guardaron en la matriz “aa” como se indicó anteriormente. Lo siguiente que se realiza es crear 3 variables `“time1= []; dis1= []; vel1= [];”` en las cuales se almacenara temporalmente las diferencias de distancia y tiempo.

Seguido se abre un lazo “for”, que va a ir desde 1 hasta la última fila de la matriz “aa”. Dentro de este lazo comienza la primera comparación entre cada par de puntos y esta es si la latitud y longitud son iguales, no hace nada si son diferentes utiliza “else” para realizar otra comparación.

La siguiente comparación es si la fecha hora de este par de puntos es igual. Esto se realiza porque si la ubicación es diferente pero la señal fue enviada en el mismo momento se atribuye a un error de calentamiento del sistema, o algún otro tipo de error y estos datos no son válidos. Si no es diferente se procede con otro “else” a realizar las siguientes operaciones.

La primera operación que se realiza es encontrar el lapso de tiempo que existe entre estos dos puntos. El lapso de tiempo se calcula utilizando la función “etime”. La función “etime” calcula el lapso de tiempo entre dos vectores. Por eso es necesario convertir cada uno de los valores de tiempo en un vector de tiempo para poder realizar operaciones con ellos.

Lo que se escribió es lo siguiente

“timediff=etime(datevec(data3(aa(j+1),3)),datevec(data3(aa(j),3)));” lo que hace es convertir en vector de tiempo el valor de la columna 3 y la fila j y j+1 en. Luego obtiene el tiempo entre estos dos puntos y se almacena en la variable “timediff”, el valor que se obtiene está dado en segundos.

Lo siguiente es convertir estos valores en horas, esto es opcional pero se decidió hacerlo así porque la distancia se obtiene en kilómetros y la velocidad se decidió que sea en km/h. Para esto se utilizó la variable “time”

Lo último que se hace en relación al tiempo es guardar cada una de estos deltas de tiempo en la matriz “time1”

Una vez terminado con el tiempo se procede a calcular las distancias entre este par de puntos, se utiliza la ya conocida función “deg2km” y se almacena estas diferencias de distancias en la matriz “dis1”

Por último se obtiene la velocidad instantánea que es la operación de dividir la distancia para el lapso de tiempo, que existe entre ese par de puntos, y se almacena en la matriz “vel1”. Acerca de este valor se hablara más adelante.

Vale recalcar que este lazo for se realiza a todos los pares de puntos consecutivos para cada día o ruta, el siguiente paso es almacenar la información que se obtiene para cada día.

Para almacenar la información se utilizó una “matriz estructura”. Su explicación y funcionamiento se detalla en el marco teórico (sección 2.1).

La estructura que se utilizó para almacenar los datos se la puede observar en la siguiente figura:

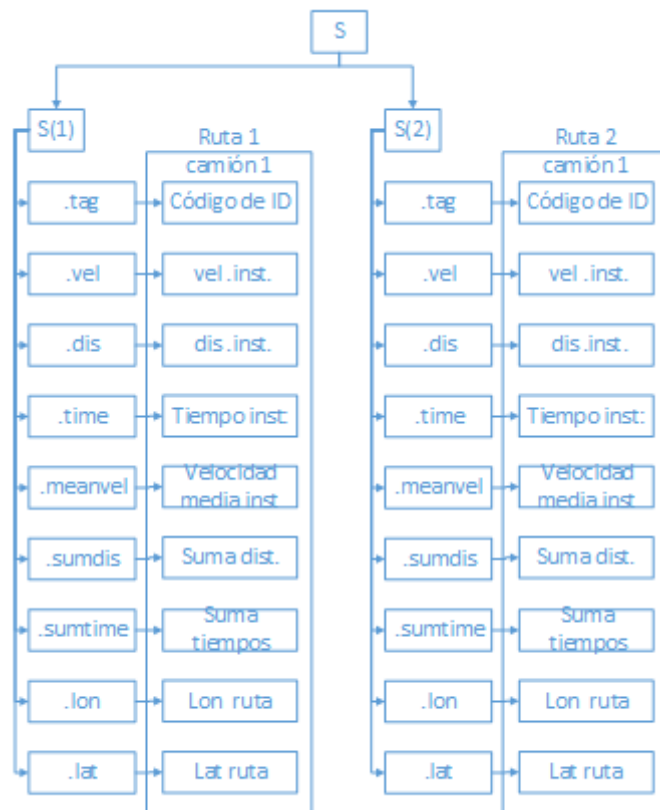


Figura 25. Estructura de almacenamiento de datos. Fuente: Elaboración Propia.

Vale recalcar que la estructura se llena por cada ruta, por esta razón esta después del lazo for que analiza cada ruta. Y posterior a esto vuelve al lazo “for” principal y vuelve a hacer todo el proceso antes descrito.

El programa tiene una segunda parte que se encarga de consolidar todos los datos de la estructura en nuevas matrices, esto se lo hace para posteriormente exportar la información a otra herramienta, dado que el único programa que entiende lo que es una estructura de este tipo es Matlab. Por esta razón se debe tener en consideración este tema

para después trabajar en otras herramientas como Excel, Minitab, etc. Esta parte se puede observar en la siguiente figura.

```

34 |
35 -   meanvelall=[];
36 -   for i=1:length(S)
37 -       meanvelall=[meanvelall,S(i).meanvel];
38 -   end
39
40 -   sumdisall=[];
41 -   for i=1:length(S)
42 -       sumdisall=[sumdisall,S(i).sumdis];
43 -   end
44 -   sumtimeall=[];
45 -   for i=1:length(S)
46 -       sumtimeall=[sumtimeall,S(i).sumtime];
47 -   end
48 -   velpro=[];
49 -   for i=1:length(sumtimeall)
50 -       velpro=[velpro,sumdisall(i)/sumtimeall(i)];
51 -   end
52 -   id=[];
53 -   for i=1:length(S)
54 -       id=[id,S(i).tag];
55 -   end

```

Figura 26. Código de programa de consolidación de métricas en matrices independientes. Fuente:

Elaboración propia

Lo que hace este programa es por medio de un lazo “for” escribir ciertos campos que se desea obtener de la estructura “S”. Para este caso se decidió obtener las velocidades promedio de las velocidades instantáneas, la suma de distancias de cada ruta y el tiempo que se demoró cada ruta.

Posterior a esto se decidió obtener la velocidad promedio de cada una de las rutas. Para esto se procedió a dividir la distancia recorrida total por ruta para el tiempo total en movimiento para esa ruta. Con esto se tiene la velocidad promedio de cada una de las rutas

Lo último que se requiere para poder exportar estos datos es identificarlos, por lo que se escribió en otra variable todos los códigos de identificación de la estructura “S”.

4.5.2 Migración archivos de Matlab a Excel

Una vez ya realizado el análisis de los datos, el siguiente paso es exportar a una herramienta que facilite su manipulación, porque Matlab si bien es un programa muy poderoso, dado que siempre toca escribir todas las operaciones en forma de código por momentos es más difícil llevar otro tipo de análisis. La siguiente figura explica brevemente el proceso que se realizó para exportar a Excel los datos que se obtuvieron en Matlab.

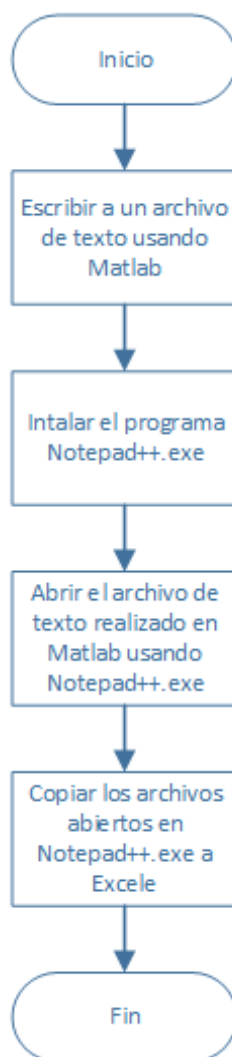


Figura 27. Proceso de migración de datos de Matlab a Excel. Fuente: Elaboración propia.

Lo primero en explicarse es el pequeño código para escribir a un archivo de texto (.txt) la información que se requiera. Se ilustrara el proceso que se utilizó para escribir los tiempos, la siguiente figura muestra el código que se utilizó.

```

fid=fopen('tiempo.txt','w');
for i=1:length(sumtimeall)
    fprintf(fid,'%5.6f\n',sumtimeall(i));
end
fclose(fid)

```

Figura 28. Código de programa de escritura de matriz de Matlab a un archivo de texto. Fuente: Elaboración propia

Lo primero es abrir este archivo utilizando la función “fopen”,

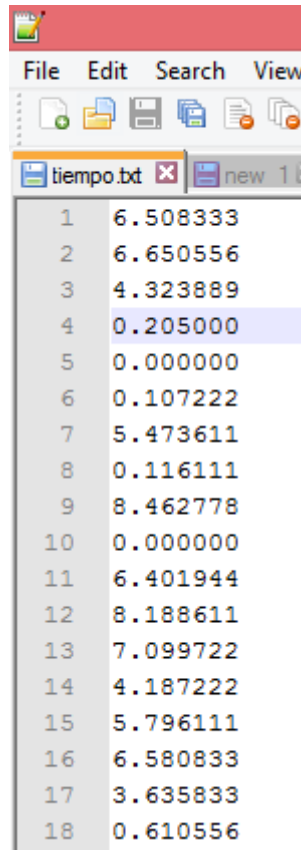
Lo siguiente es abrir un lazo “for” que va a recorrer todos los datos de la matriz “sumtimeall” en este caso. Lo siguiente es escribir cada uno de los datos presentes en la matriz “sumtimeall” y para realizar esto se utilizó la función “fprintf”.

Especificar qué es lo que se va a escribir en este caso es el valor “i” de la matriz “sumtimeall”.

Una vez que ya se escribió todos los datos requeridos el último paso es cerrar el archivo que se abrió, para esto se utiliza la función “fclose”, en la que el único parámetro que se debe especificar es el nombre del archivo que se desea cerrar en este caso en particular es “fid”.

El archivo se crea como cualquier archivo de texto en la carpeta especificada. Lo siguiente es abrirlo en este caso utilizando “Notepad++”.

Una vez abierto el archivo en la siguiente figura muestra como se ve un archivo abierto.



1	6.508333
2	6.650556
3	4.323889
4	0.205000
5	0.000000
6	0.107222
7	5.473611
8	0.116111
9	8.462778
10	0.000000
11	6.401944
12	8.188611
13	7.099722
14	4.187222
15	5.796111
16	6.580833
17	3.635833
18	0.610556

Figura 29. Archivo de texto abierto en Notepad++. Fuente: Elaboración propia.

El siguiente paso es copiar los datos a una hoja de Excel y ya se tiene la información lista para ser trabajada en otra herramienta. Este paso se repitió para copiar distancias y código de ID.

Esta fue la herramienta que se utilizó para obtener información de la base de datos depurada el siguiente paso es analizar la data que se obtuvo. Para finalizar este capítulo se obtuvo 1610 rutas que se reparten entre los 15 camiones durante los 7 meses de estudio, en el siguiente capítulo se estudiara más a fondo esta rutas.

4.6. Validación y Análisis de resultados.

Para proceder al análisis se utilizó Excel y Matlab al mismo tiempo. Se hace esto porque Excel facilita la manipulación de datos (1610 filas de datos) y Matlab permite graficar estos datos o rutas en este caso. Son herramientas complementarias para este estudio, por ejemplo en Excel se identificó anomalías o simplemente datos que requerían

más investigación. La forma más simple de ver que sucedió con esa ruta es graficarla dentro de un mapa georeferenciado y esta parte es donde Matlab es muy importante.

Lo primero fue hacer gráficos de dispersión de los tiempos de duración y la distancias para tener una mirada más global de cómo se comportan los datos. Las siguientes dos figuras nos muestran estos gráficos de dispersión.

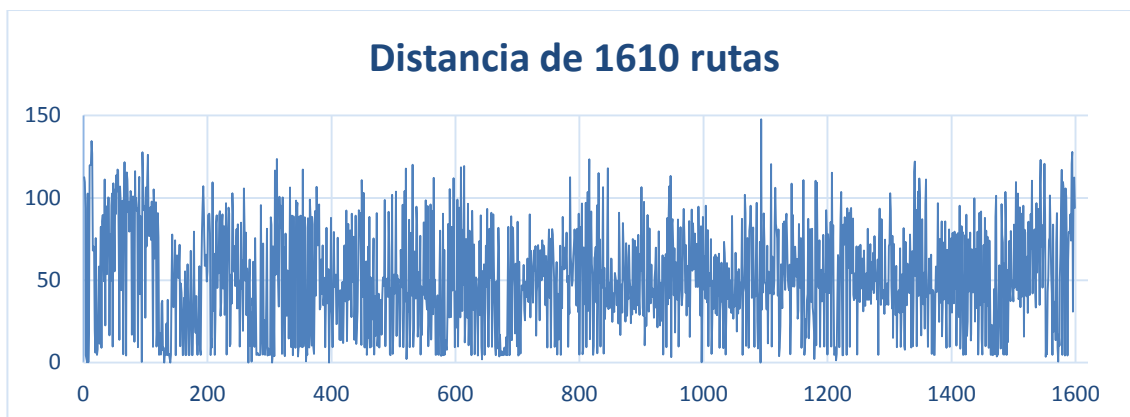


Figura 30. Gráfico de dispersión con líneas rectas de la distancia de las 1610 rutas. Fuente: Elaboración propia.

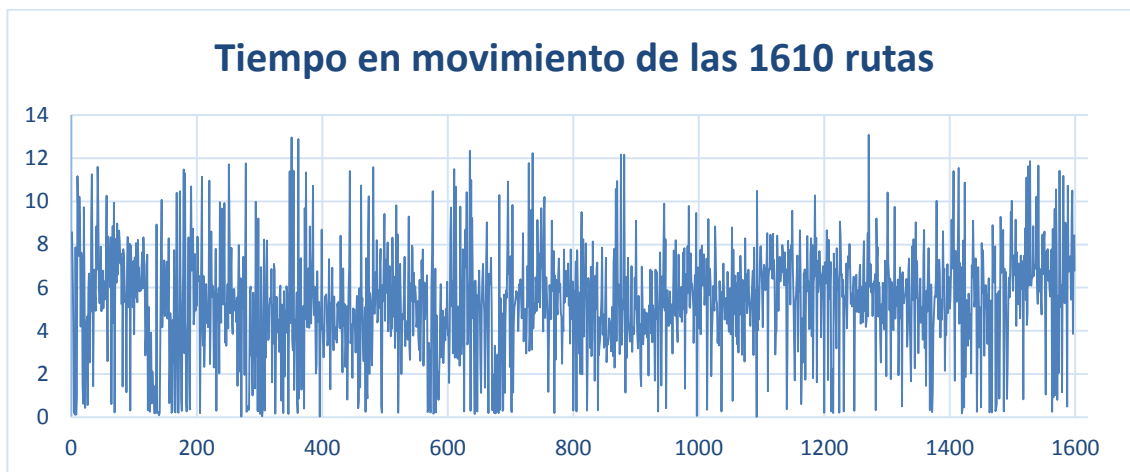


Figura 31. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las 1610 rutas. Fuente: Elaboración propia.

En este momento se presentó un inconveniente con los datos, algunos tiempos y distancias presentan valores de 0. Lo que se procedió a hacer es identificar las rutas que presentaron este problema y la siguiente tabla muestra estas rutas.

Tabla 6. Rutas con distancia y tiempo 0. Fuente: Elaboración Propia

#	Tag	DIS	Time	VEL
5	201710	0	0	#_DIV/0!
11	201720	0	0	#_DIV/0!
131	208726	0	0	#_DIV/0!
139	208820	0	0	#_DIV/0!
284	302712	0	0	#_DIV/0!
441	322614	0	0	#_DIV/0!
723	33346	0	0	#_DIV/0!
758	333518	0	0	#_DIV/0!
858	33546	0	0	#_DIV/0!
970	335821	0	0	#_DIV/0!
990	33646	0	0	#_DIV/0!
1100	336820	0	0	#_DIV/0!
1330	338629	0	0	#_DIV/0!
1590	351723	0	0	#_DIV/0!
1606	351812	0	0	#_DIV/0!

Estas son las rutas con valor de 0, hay que recalcar que todas las rutas con valor 0 tienen su valor de distancia y tiempo de ruta como 0. Se grafica estas rutas para identificar

la raíz de este inconveniente, a continuación se muestra una tabla de resumen de las 16 rutas graficadas en un mapa georeferenciado de la ciudad de Quito. Las rutas graficadas se las puede observar en el Anexo 11.

Tabla 7. Resumen de análisis de mapeo de rutas con distancia y tiempo igual a 0. Fuente Elaboración Propia.

#	Tag	Ruta SI NO	Notas
5	201710	NO	CO
11	201720	NO	CO
131	208726	NO	CO
139	208820	NO	CO
284	302712	NO	CO
441	322614	NO	CALDERON
723	33346	NO	CALDERON
758	333518	NO	CO
858	33546	NO	CD
970	335821	NO	CD
990	33646	NO	CD
1100	336820	NO	CD
1330	338629	NO	CD
1590	351723	NO	CD
1606	351812	NO	CO

Lo que se observó es que estas 16 rutas no tuvieron ningún movimiento en ese día. Los 16 días esos camiones permanecieron parqueados a las afueras del CD y del centro de operaciones. Hubo un lugar de parqueo cerca de Calderón y después de las pertinentes averiguaciones a la compañía fue un lugar donde se parquearon los camiones para un

trabajo temporal. Considerando esta situación se procede a eliminar estas 16 rutas erróneas, que no fueron eliminadas en los filtros anteriores. Este error se atribuye a que estos días esos dispositivos solo enviaron 1 señal así que al momento de realizar la limpieza todas ellas siempre comparaban pares de datos y en este caso no es posible realizarlo.

Una vez más se hace gráficos de dispersión para tiempo de ruta y distancia descartando los valores anteriores. Esto se puede observar en la siguiente figura

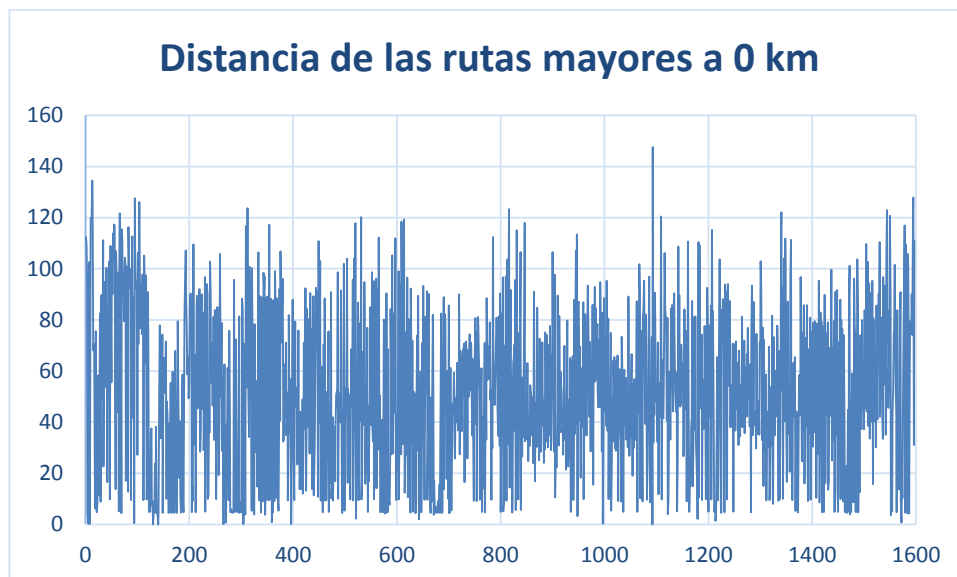


Figura 32. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 0 km. Fuente: Elaboración propia.

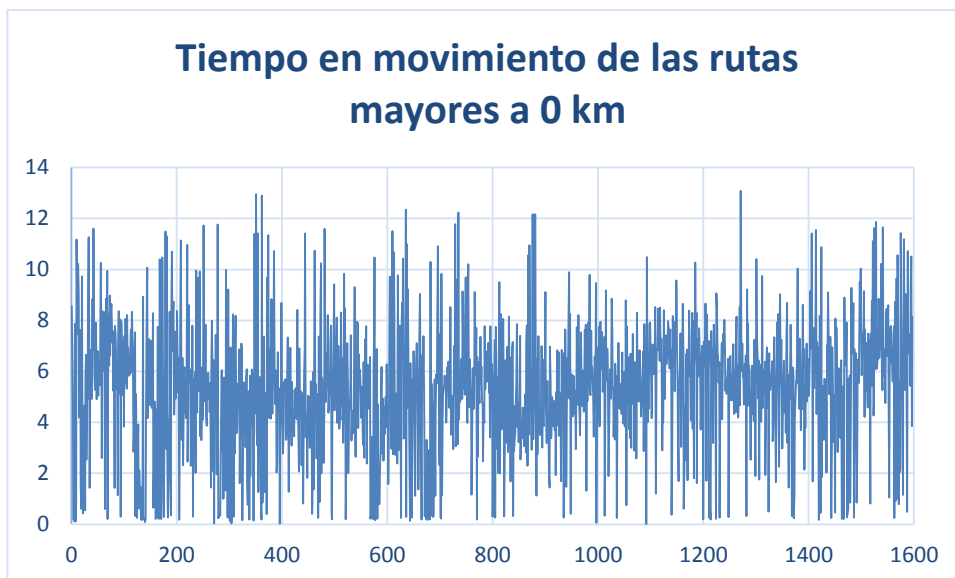


Figura 33. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las rutas mayores a 0 km.

Fuente: Elaboración propia.

En estos nuevos gráficos ya no existen datos con valores a 0, pero sin embargo hay valores muy reducidos aun. Se analiza todos los datos menores a 1 km para encontrar algún tipo de singularidad sobre estas rutas. Las rutas se muestran en la siguiente tabla.

Tabla 8. Rutas con distancia menor a 1 km. Fuente: Elaboración propia.

#	Tag	DIS	Time	VEL
1092	336810	0.000458	0.0025	0.1832
1607	351820	0.010121	0.966667	0.01047
130	208722	0.013756	1.433333	0.00959721
140	300512	0.014669	0.077778	0.18860089
396	322413	0.071334	0.023333	3.0572151
304	321512	0.104199	0.036111	2.88551965
266	302619	0.134276	5.824167	0.02305497

997	336414	0.217447	0.053611	4.05601462
8	201717	0.226564	0.110278	2.05448049
6	201715	0.233419	0.107222	2.17696928
94	20862	0.508049	0.299167	1.69821204
271	302625	0.65896	0.029444	22.3801114
1572	351630	0.669341	5.609167	0.11932984
359	32181	0.889053	5.786667	0.15363818

Se grafica las 14 rutas identificadas anteriormente y a continuación se muestra la tabla donde se resume lo observado. Las rutas mapeadas se las puede ver en el Anexo 12.

Tabla 9. Resumen de análisis de mapeo de rutas menores a 1 km. Fuente: Elaboración propia.

#	Tag	Ruta SI NO	Notas
1092	336810	NO	CD
1607	351820	NO	CO
130	208722	NO	CALDERON
140	300512	NO	NI
396	322413	NO	CD
304	321512	NO	CO
266	302619	NO	CO
997	336414	NO	CO
8	201717	NO	CD
6	201715	NO	CD
94	20862	NO	CD
271	302625	NO	CO

1572	351630	NO	CD
359	32181	NO	CO

Como se observa en la tabla anterior, todas las rutas menores a 1 km, presentan características similares. Todas están alrededor del centro de operaciones o del centro de distribución. Son distancias y rutas que no son representativas para el estudio por esta razón se elimina y se vuelve a realizar gráficos de dispersión, para la distancia y el tiempo de duración de la ruta.

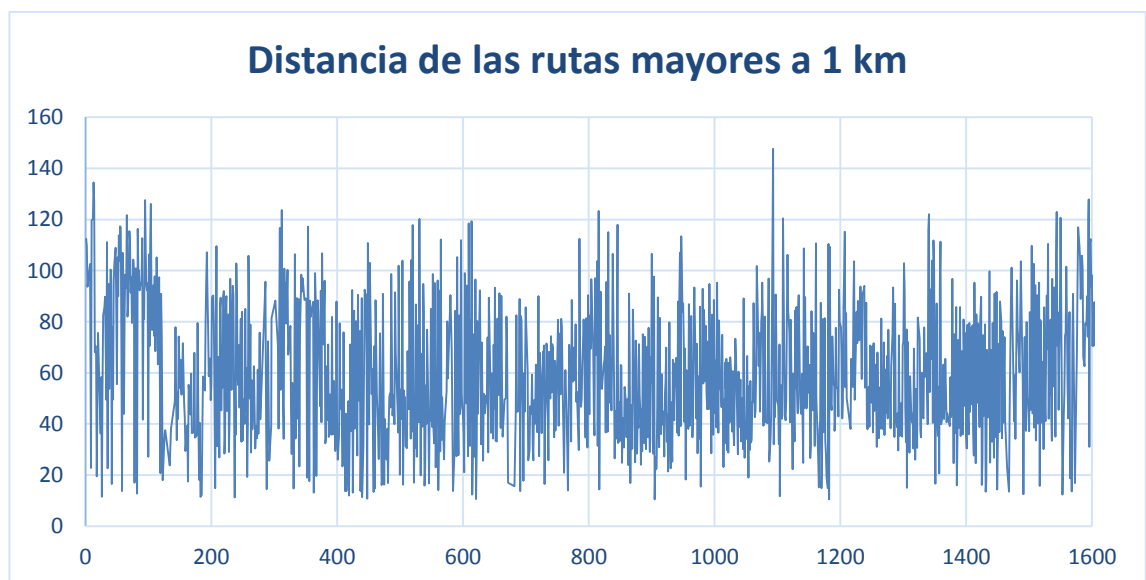


Figura 34. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 1 km. Fuente:

Elaboración propia.

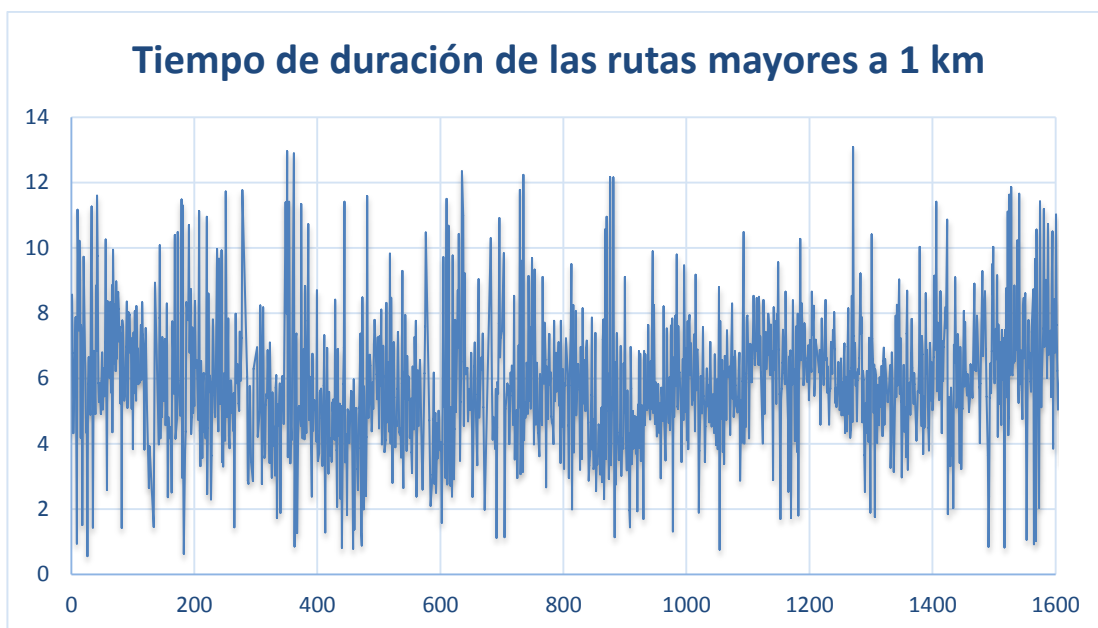


Figura 35. Gráfico de dispersión con líneas rectas de tiempo en movimiento de las rutas mayores a 1 km.

Fuente: Elaboración propia.

Dada la gran cantidad de datos se procede a realizar los siguientes análisis, prueba de normalidad, y obtener los percentiles en aumentos de 5%, para tener una idea clara del comportamiento de los datos. La siguiente tabla muestra los percentiles de distancia y el tiempo de duración de estas rutas.

Tabla 10. Percentiles en intervalos de 5% de rutas mayores a 1 km. Fuente: Elaboración propia.

Percentil	Distancia	Tiempo
5%	4.792437	0.333889
10%	9.029738	1.617222
15%	10.348173	2.773611
20%	20.874695	3.463056
25%	29.68126	3.988333
30%	34.805602	4.374722

35%	38.155366	4.688056
40%	41.769917	4.971944
45%	44.752987	5.215278
50%	48.301787	5.446111
55%	52.225194	5.756667
60%	56.975723	5.969167
65%	64.227809	6.227778
70%	69.94426	6.517778
75%	75.619043	6.832778
80%	81.372091	7.211111
85%	88.336138	7.6825
90%	93.390962	8.246111
95%	103.649215	9.505
100%	147.699647	13.089722

El primer análisis es sobre los datos que están en percentil 15%, todos los datos que su distancia es menor a 10.34 km serán los primeros en ser estudiados. Existen 238 datos dentro del 15% percentil. A estos datos se les realizar un análisis para lograr encontrar algún tipo de comportamiento peculiar. Se realizó una muestra aleatoria de 15 rutas para identificar cualquier tipo de comportamiento. De las 238 rutas se obtuvo 15 números aleatorios en este rango utilizando Excel, la siguiente tabla muestra los números de los cuales se tomara la muestra.

Tabla 11. Muestreo aleatorio de rutas entre el 0 y 15 percentil. Fuente: Elaboración propia.

Aleatorio (0-1)	Aleatorio (1-235)	#	Tag
0.613086264	145	806	333715
0.78036992	184	1312	33869
0.357924646	85	1450	340616
0.098340599	24	1225	33784
0.311685749	74	647	33062
0.171826609	41	1224	33783
0.062879159	15	1583	351716
0.216246215	51	573	324714
0.994774226	234	269	302623
0.707390875	167	585	324820
0.219339155	52	171	300721
0.78006193	185	1219	337728
0.542182671	128	372	322317
0.61049799	144	23	201813
0.673855574	159	1571	351626

A continuación se grafican las 15 rutas identificadas anteriormente y a continuación se muestra la tabla donde se resume lo observado. Para ver los mapeos de las rutas referirse al anexo 13.

Tabla 12. Resumen de análisis de mapeo muestra aleatoria de rutas en el 0 y 15 percentil. Fuente:

Elaboración propia.

#	Tag	Ruta SI NO	Notas
806	333715	NO	CD-CO

1312	33869	NO	CD-CO
1450	340616	NO	CD-CO
1225	33784	NO	CD-CO
647	33062	NO	CD-CO
1224	33783	NO	CD-CO
1583	351716	NO	CD-CO
573	324714	NO	CD-CO
269	302623	NO	CD-CO
585	324820	NO	CD-CO
171	300721	NO	CD-CO
1219	337728	NO	CD-CO
372	322317	NO	CD-CO
23	201813	NO	CD-CO
1571	351626	NO	CD-CO

Como se observa todas estas rutas son entre el centro de operación y el centro de distribución. No son rutas de entrega por lo que no se les considera útiles y se las elimina. Son distancias y rutas que no son representativas para el estudio por esta razón se elimina y se vuelve a realizar gráficos de dispersión, para la distancia y el tiempo de duración de la ruta para los datos que su distancia es mayor a 12 km.

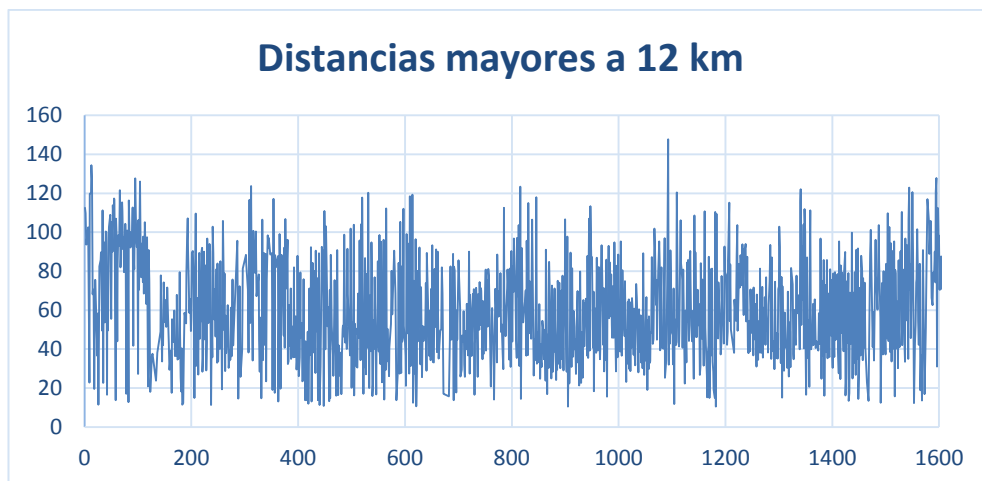


Figura 36. Gráfico de dispersión con líneas rectas de distancia de las rutas mayores a 12 km. Fuente: Elaboración propia.



Figura 37. Gráfico de dispersión con líneas rectas de tiempo de duración de las rutas mayores a 12 km. Fuente: Elaboración propia.

Se procede a calcular los percentiles con intervalos de 5% para analizar las rutas que se ubican en los extremos superiores y en entre el 40% y 60%

Tabla 13. Percentiles en intervalos de 5 de los datos con distancia mayor a 12 km. Fuente: Elaboración propia.

Percentil	Distancia	Tiempo
5%	22.8140088	2.726389
10%	28.308711	3.3964446

15%	33.0589016	3.8373336
20%	35.8793108	4.1692224
25%	38.615729	4.466111
30%	40.9907722	4.6857226
35%	43.5610366	4.898
40%	45.4284278	5.1083888
45%	48.2524214	5.3176666
50%	52.060464	5.524167
55%	55.2682846	5.7724444
60%	60.0594138	5.9416108
65%	65.429004	6.1758336
70%	70.3102938	6.437444
75%	75.529072	6.639722
80%	80.7821572	6.885556
85%	85.592332	7.1933888
90%	92.011059	7.6338886
95%	101.2817962	8.1016108
100%	123.342043	9.818056

Se analiza los datos del 15 % superior, que son todos los valores superiores a 85.59 km, se obtiene 202 rutas que están sobre este valor. Se decide tomar una muestra aleatoria de 15 rutas para identificar cualquier peculiaridad. La muestra se especifica en la siguiente tabla.

Tabla 14. Muestreo aleatorio de rutas entre el 85 y 100 percentil. Fuente: Elaboración propia

Aleatorio (0-1)	Aleatorio (1-202)	#	Tag
0.332070535	68	1584	351717
0.185456335	38	565	324628
0.087562191	18	614	330419
0.339282963	69	259	302611
0.622463721	126	1358	338729
0.602394613	122	620	330426
0.102779724	21	520	32453
0.572279132	116	810	333719
0.067171315	14	1109	337322
0.730364018	148	97	20865
0.2808263	57	96	20864
0.739717168	150	234	30258
0.809247443	164	114	208625
0.004418741	1	1093	336811
0.285585446	58	193	300823

Se grafica estas 15 rutas identificadas anteriormente y a continuación se muestra la tabla donde se resume lo observado. Para ver el mapeo de las 15 rutas se debe referir al anexo 14

Tabla 15. Resumen de análisis de mapeo muestra aleatoria de rutas en el 85 y 100 percentil. Fuente:

Elaboración propia

#	Tag	Ruta SI NO	Notas
1584	351717	SI	SUR

565	324628	SI	SUR
614	330419	SI	SUR
259	302611	SI	SUR
1358	338729	SI	SUR
620	330426	SI	SUR
520	32453	SI	SUR
810	333719	SI	CEN-NOR
1109	337322	SI	SUR
97	20865	SI	SUR
96	20864	SI	SUR
234	30258	SI	SUPERMAXI
114	208625	SI	SUR
1093	336811	SI	SUR
193	300823	SI	SUR

Como se observa en la tabla anterior las 15 rutas son de entrega, por lo que se las conserva ya que brindan información válida para el estudio.

El siguiente paso es analizar las rutas que se encuentran entre el 40 y el 60 percentil para encontrar cualquier peculiaridad, estas rutas se encuentran entre 45 y 60 km, se utiliza la misma metodología se obtendrá una muestra aleatoria de 15 rutas de las 270 totales, la muestra se encuentra en la siguiente tabla.

Tabla 16. Muestreo aleatorio de rutas entre el 40 y 60 percentil. Fuente: Elaboración propia

Aleatorio (0-1)	Aleatorio (1-270)	#	Tag
0.332070535	90	1234	337815

0.185456335	51	1029	336524
0.087562191	24	115	208626
0.339282963	92	1041	33669
0.622463721	169	781	333614
0.602394613	163	190	300819
0.102779724	28	851	335329
0.572279132	155	724	33347
0.067171315	19	1021	336515
0.730364018	198	1575	35173
0.2808263	76	196	302320
0.739717168	200	1515	351411
0.809247443	219	1497	351320
0.004418741	2	934	33575
0.285585446	78	1375	338820

Se grafica estas 15 rutas identificadas anteriormente y a continuación se muestra la tabla donde se resume lo observado. Para ver el mapeo de las 15 rutas se debe referir al anexo 15

Tabla 17. Resumen de análisis de mapeo muestra aleatoria de rutas en el 40 y 60 percentil. Fuente:

Elaboración propia

#	Tag	Ruta SI NO	Notas
1234	337815	SI	CEN-NOR
1029	336524	SI	CEN-NOR
115	208626	SI	CEN-NOR

1041	33669	SI	CEN-NOR
781	333614	SI	CEN-NOR
190	300819	SI	CEN-NOR
851	335329	SI	CEN-NOR
724	33347	SI	EXTRAORDINARIA
1021	336515	SI	CEN-NOR
1575	35173	SI	CEN-NOR
196	302320	SI	CEN-NOR
1515	351411	SI	CEN-NOR
1497	351320	SI	CEN-NOR
934	33575	SI	CEN-NOR
1375	338820	SI	CEN-NOR

Como se observa todas son rutas de entrega. Por lo que se las mantiene.

Para resumir las rutas hasta los 12 km se identificaron como viajes entre CO y CD, por eso se las elimino, por otro lado las muestras entre 40 y 60 percentil y las entre 85 y 100 percentil se las mantuvo y paso de tener 1610 rutas a 1343 rutas en la base ya validada. Lo siguiente realizado es obtener la informacion estadistica descriptiva que se observa en la siguiente tabla.

Tabla 18. Promedio y desviación estándar de métricas obtenidas. Fuente: Elaboración propia.

Métrica	Unidad	Valor
Velocidad promedio	km/h	10.61
Desviación estándar velocidad	km/h	4.34
Distancia Promedio	km	59.43

Desviación estándar distancia	km	26.51
Duración de ruta promedio	h	5.85
Desviación estándar de duración de ruta	h	2.05

Posterior a esto se realizó la prueba de coeficiente de variación para observar la homogeneidad de los datos y entender que tan bien describen estos estadísticos a las muestra. En la siguiente tabla se muestra el cálculo del CV

Tabla 19. Coeficiente de variación. Fuente: Elaboración propia.

Métrica	CV
Velocidad	41%
Distancia	45%
Duración de ruta	35%

Como se observa en la tabla los datos no son homogéneos, esto conlleva a pensar que las estimaciones antes presentadas no brindan información certera acerca de cómo se comporta la muestra. Por esta razón se realiza otro tipo de análisis descriptivo, a continuación se puede observar los gráficos de caja para las distancias y las duraciones de las rutas.

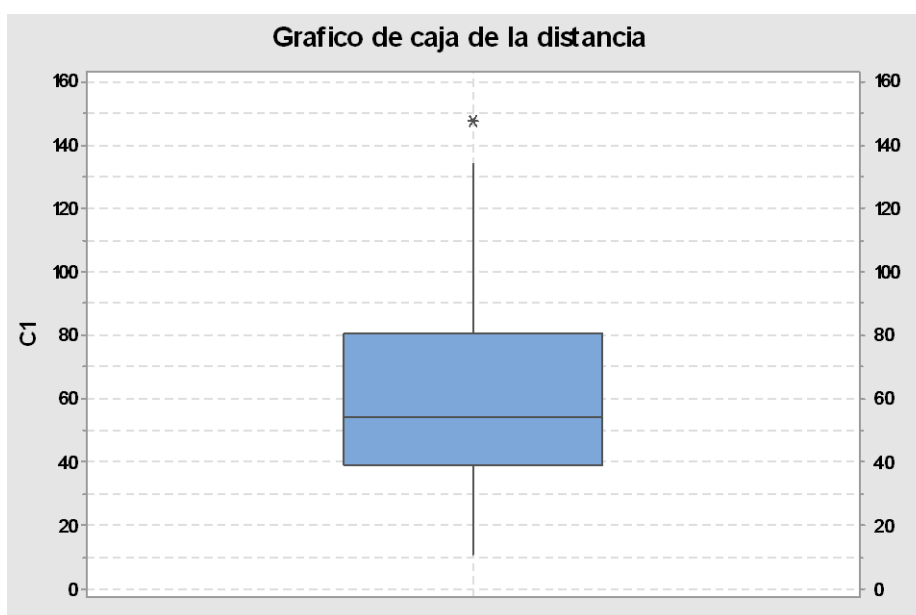


Figura 38. Gráfico de caja de las distancias. Fuente: Elaboración propia.

Se observa que entre el primer (25%) y tercer (75%) cuartil es decir el 50% de los datos se encuentran entre los 39 y 80 km de distancia aproximadamente. Con una diferencia de 41 km.

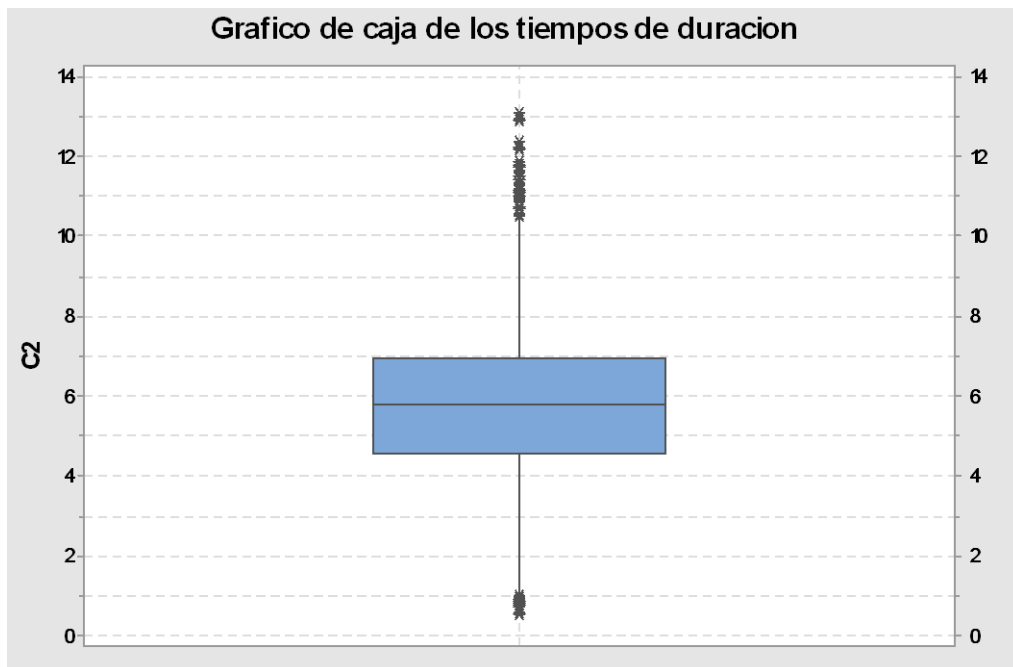


Figura 39. Gráfico de caja de los tiempos de duración de las rutas. Fuente: Elaboración propia.

Se observa que entre el primer (25%) y tercer (75%) cuartil es decir el 50% de los datos se encuentran entre los 4.58 y 6.64 horas de distancia aproximadamente. Con una diferencia de 2.36 horas.

Otro elemento que brinda información es que el 80 por ciento de los datos de distancia se encuentra entre 33 y 90 km, para el tiempo de duración están entre las 4 y 8 horas aproximadamente. También se sabe que en un rango de 27 km entre los 27 y 54 km se encuentra el 40% de las rutas, mientras que el 40 restante está entre 54 y 95 km en un rango de 40 km. El promedio de duración de las rutas entre 27 y 54 km (rutas entre el 10% y 50%) es de 5.45 horas y el promedio de duración de las rutas entre 54 y 95 km (rutas entre 50% y 90 %) es de 6.45 horas. Para inferir sobre este fenómeno se procederá a hacer

una prueba de hipótesis de dos muestras con varianza desconocida. La prueba se la realiza en Minitab y será entre la duración de las rutas que están entre los 27 y 54 km y las que están entre 54 y 95 km. Antes de realizar la prueba se hace un análisis de normalidad que se muestra a continuación

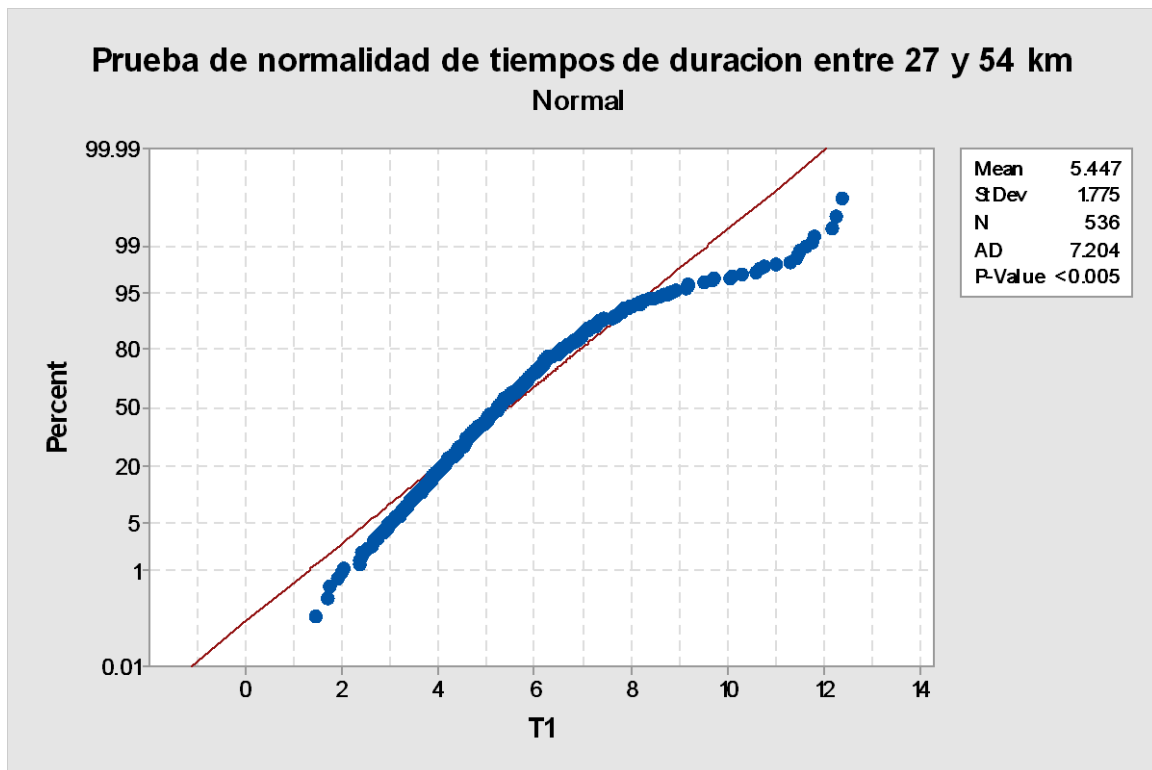


Figura 40. Prueba de normalidad de tiempos de duración entre 27 y 54 km. Fuente: Elaboración Propia.

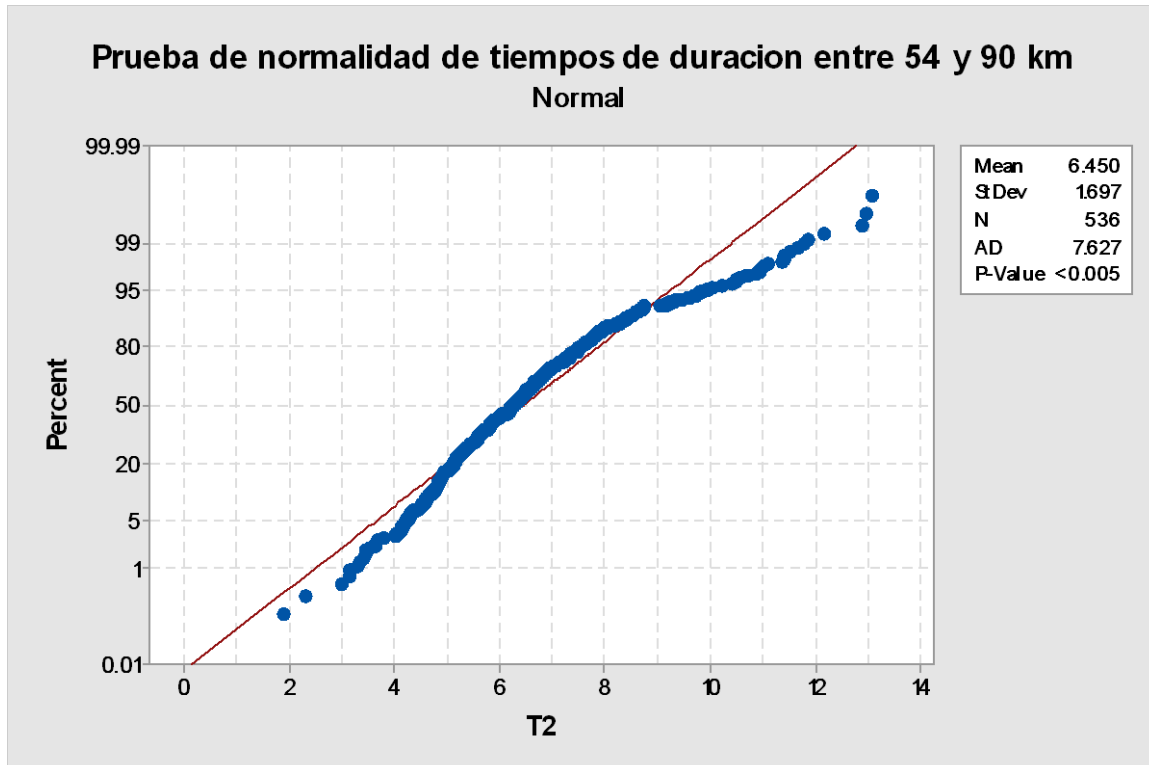


Figura 41. Prueba de normalidad de tiempos de duración entre 54 y 90 km. Fuente: Elaboración Propia.

Como se observa no se tiene una distribución normal, esto no satisface el supuesto para realizar una prueba t. Sin embargo se encontró en la literatura que se puede realizar pruebas t si los datos no son normales si se cumplen los siguientes requisitos.

- Las muestras que se comparan deben tener una distribución razonablemente simétrica
- Las muestras que se comparan deben tener un medio que está cerca del centro de la distribución;
- La distribución debe tener un solo modo (punto más alto en el histograma de frecuencias).

Si no cumple normalidad pero si cumplen las distribuciones lo antes planteado con una muestra de más de 30 datos se puede realizar una prueba t (University of Glasgow, 2014).

A continuación se muestran los histogramas para validar los anteriores supuestos.

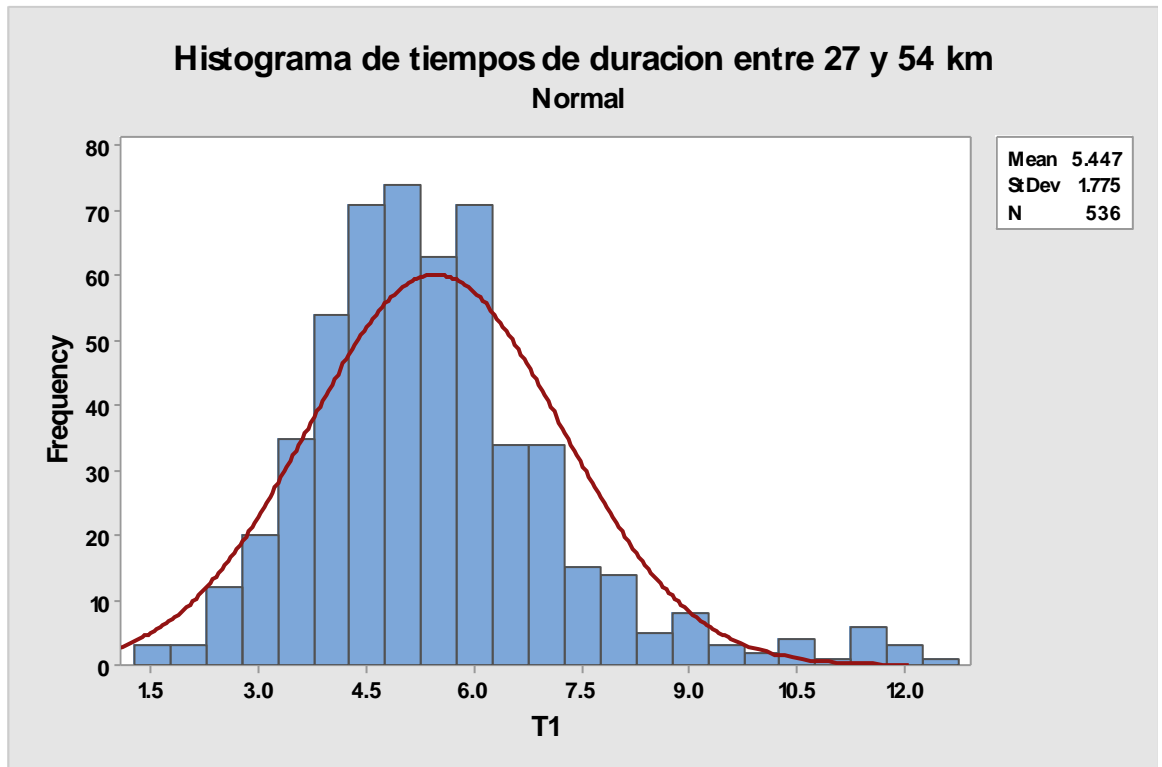


Figura 42. Histograma de tiempos de duración entre 27 y 54 km. Fuente: Elaboración propia

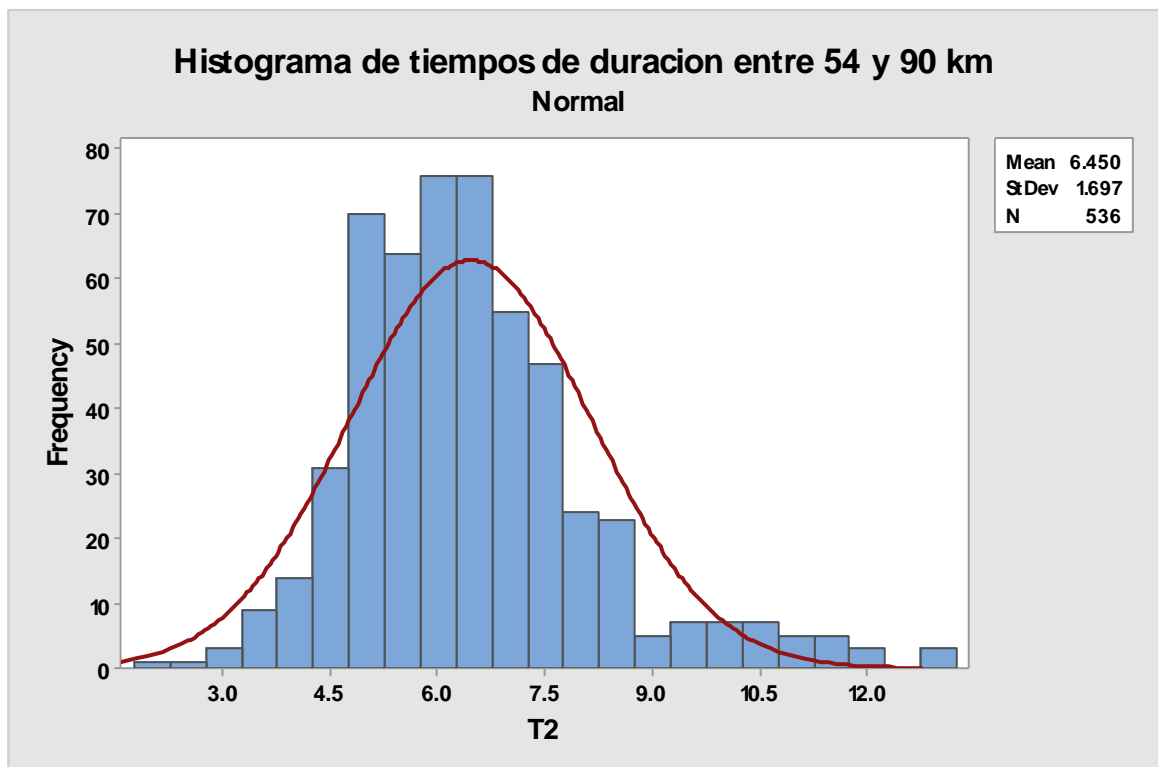


Figura 43. Histograma de tiempos de duración entre 54 y 90 km. Fuente: Elaboración propia

Como se observa en los histogramas se cumplen Los 3 supuestos expuestos anteriormente por lo que se procede a realizar la prueba t que se muestra a continuación.

La prueba de hipótesis se la observa a continuación:

Two-Sample T-Test and CI: T1, T2

Two-sample T for T1 vs T2

	N	Mean	StDev	SE Mean
T1	536	5.45	1.77	0.077
T2	536	6.45	1.70	0.073

Difference = μ (T1) - μ (T2)

Estimate for difference: -1.004

95% CI for difference: (-1.212, -0.795)

T-Test of difference = 0 (vs \neq): T-Value = -9.46 P-Value = 0.000 DF = 1070

Both use Pooled StDev = 1.7364

Puesto que el valor p es menor a alfa (0.05) se rechaza la hipótesis nula, concluyendo que si existe diferencia de medias. Estadísticamente las medias son diferentes pero se puede realizar otro tipo de análisis. La media de distancia entre 27 y 54 es de 41 km y su tiempo de duración promedio es de 5.45 horas, para el otro grupo de datos que esta entre 54 y 95 la media es de 75 km y su duración de ruta promedio es de 6.45 horas. La relación al aumento de medias entre los dos grupos pasando de 41 a 75 km es del 45% y

el aumento de los tiempos es del 15%. Existen indicios que hay una falta de proporcionalidad en relación al crecimiento de la distancias y el aumento de los tiempos de duración de las rutas. Se hace gráficos de dispersión de estos dos parámetros y se los puede ver a continuación.

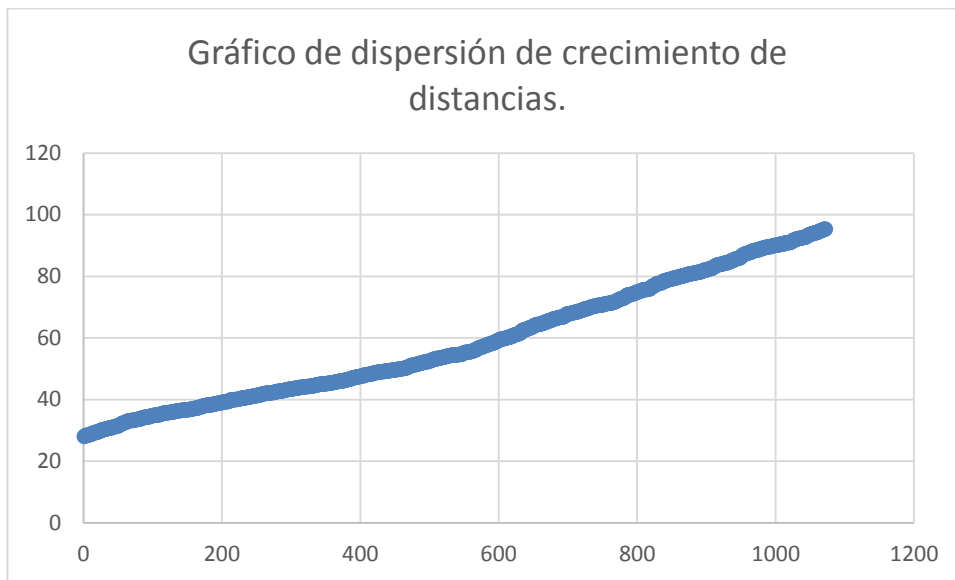


Figura 44. Gráfico de dispersión de crecimiento de distancias. Fuente: Elaboración propia.

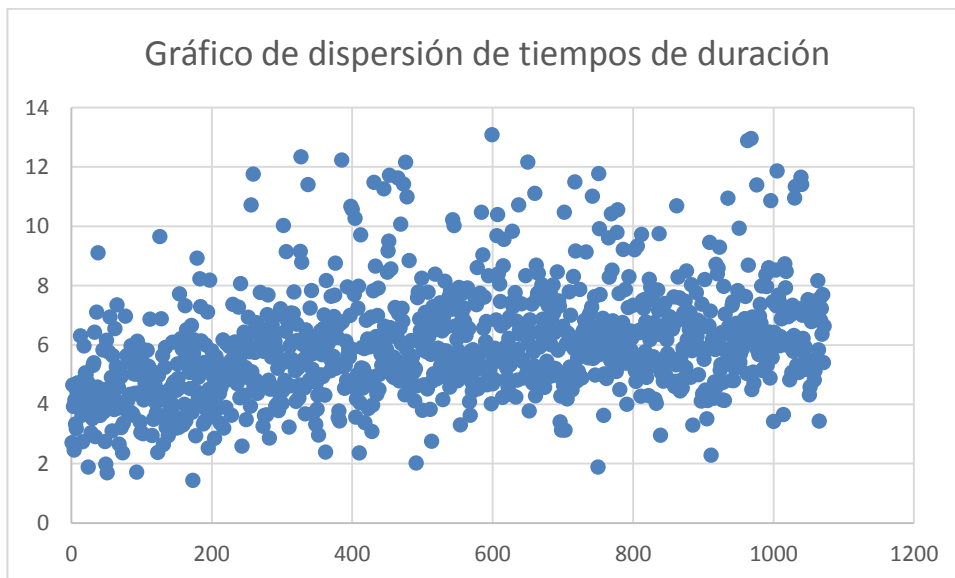


Figura 45. Gráfico de dispersión de tiempos de duración. Fuente: Elaboración propia.

Para tener una idea de cuál es pendiente de crecimiento de estas dos muestras y ver cómo crecen las distancias y como crecen los tiempos de duración se realizó una regresión. En la regresión la variable independiente para las dos muestra es el # de ruta y la variable

dependiente para un caso es la distancia y para el otro caso son los tiempos. A continuación se muestra la regresión de las distancia y se realiza un análisis de los residuales para saber si la regresión es significativa.

Model Summary

S	R-sq	R-sq (adj)	R-sq (pred)
2.45525	98.38%	98.38%	98.37%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	25.297	0.150	168.55	0.000	
x1	0.061785	0.000242	254.97	0.000	1.00

Regression Equation

$$d = 25.297 + 0.061785 x$$

Como se observa el R^2 satisface el supuesto, a continuación se muestra el análisis de residuales para concluir si la regresión es significativa o no.

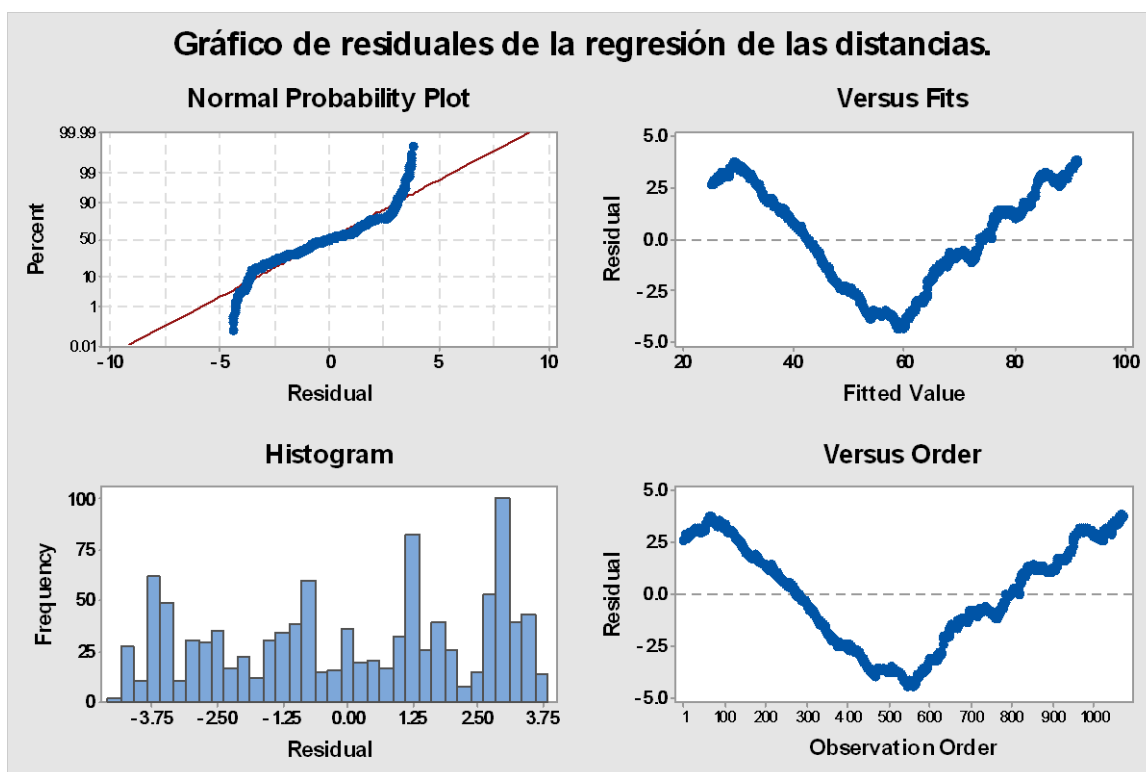


Figura 46. Gráfico de residuales de residuales de distancias. Fuente: Elaboración propia.

Como se observa en el gráfico los residuales no satisfacen los supuestos, lo que conlleva a concluir que la regresión no es significativa. Se procede a realizar el mismo análisis para los tiempos de duración.

S	R-sq	R-sq(adj)	R-sq(pred)
1.69564	12.01%	11.92%	11.70%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	4.864	0.104	46.93	0.000	
x1	0.002022	0.000167	12.08	0.000	1.00

Regression Equation

t80% = 4.864 + 0.002022 x1

Como se observa el R^2 no satisface el supuesto dado que el su valor es muy bajo, a continuación se muestra el análisis de residuales para concluir si la regresión es significativa o no.

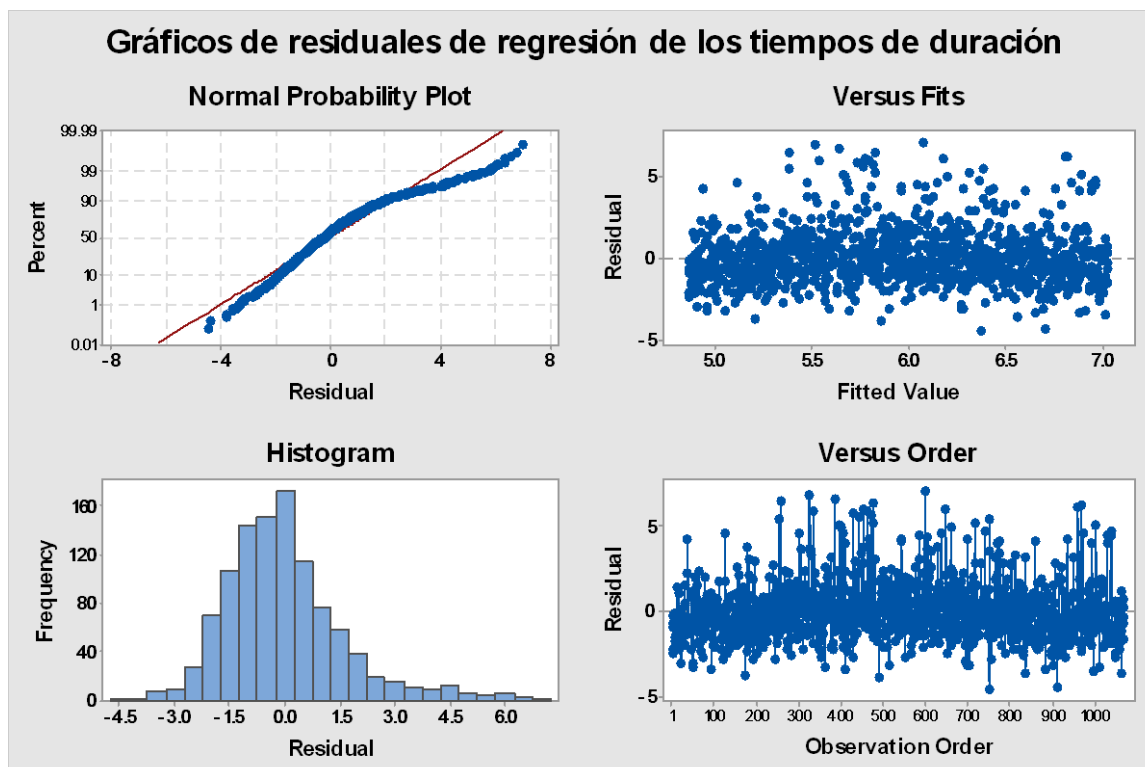


Figura 47. Gráfico de residuales de residuales de tiempos de duración. Fuente: Elaboración propia.

Como se observa en el gráfico los residuales satisfacen los supuestos, sin embargo dado que su R^2 es tan bajo se concluye que la regresión no es significativa. Dado que las regresiones no son significativas, no se puede realizar inferencias más profundas.

Si bien no se puede llegar a conclusiones acerca de esta falta de proporcionalidad se realizara comparaciones entre dos rutas que están dentro de estos dos grupos. Una ruta del centro norte de Quito que nunca sale del perímetro urbano y una ruta que va hacia el sur.

Las siguientes figuras muestran las dos rutas

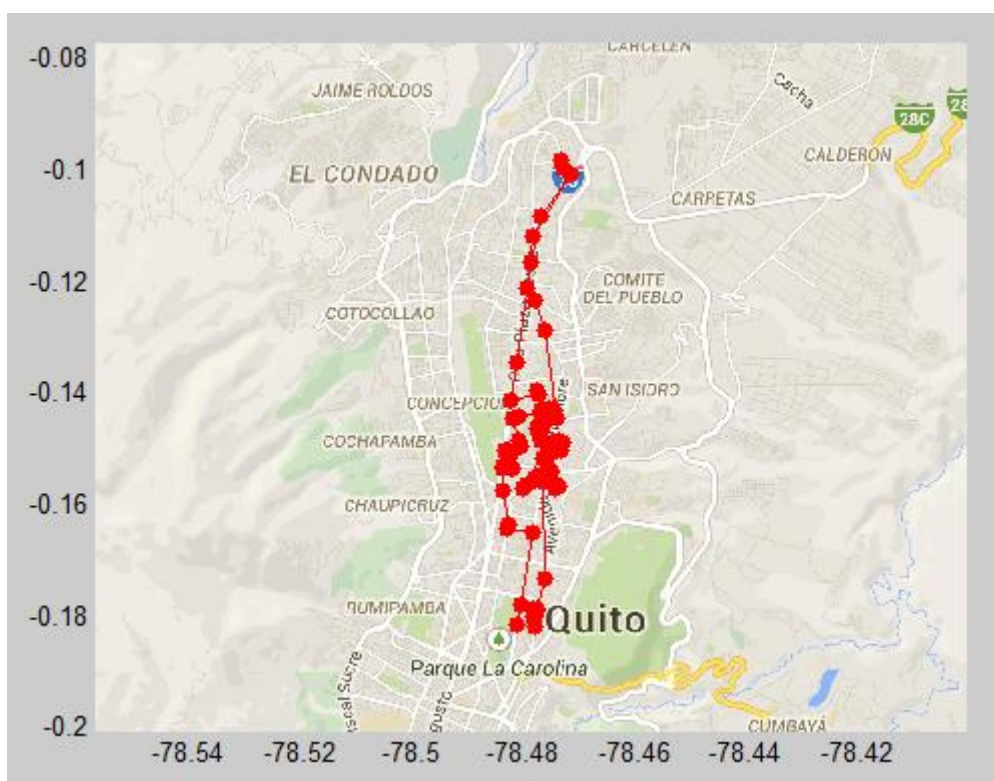


Figura 48. Mapeo de la ruta #1316 camión 338 el 13 de junio de 2014. Fuente: Elaboración propia.

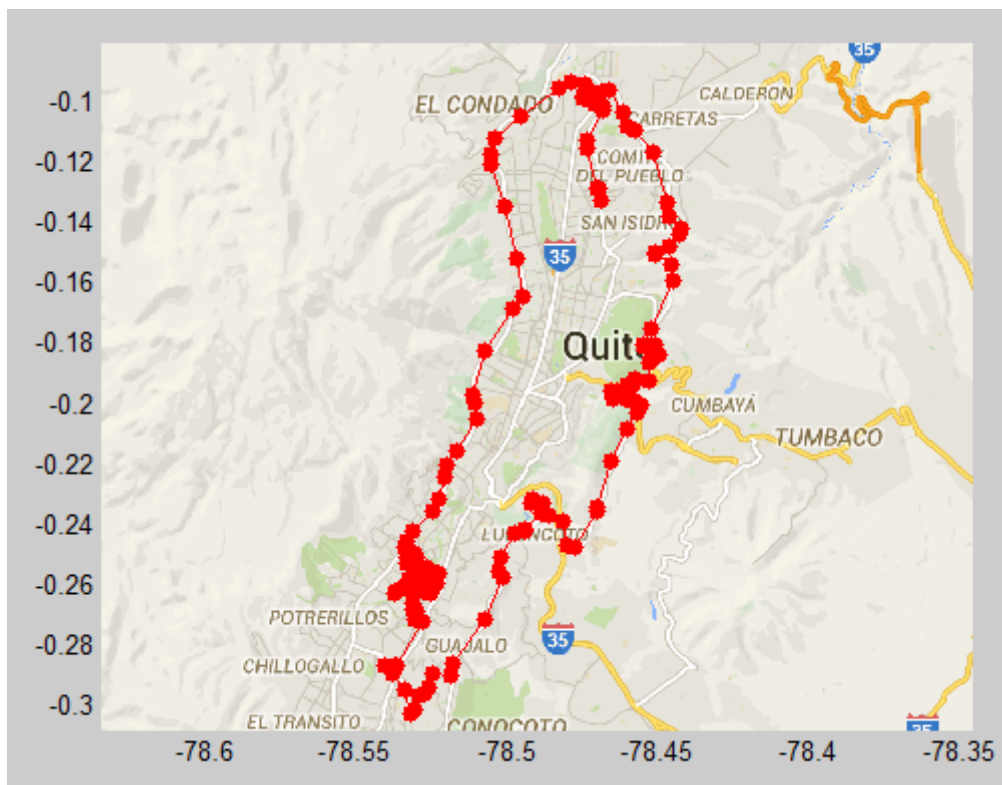


Figura 49. Mapeo de la ruta #1323 camión 338 el 21 de junio de 2014. Fuente: Elaboración propia.

La figura 44 muestra una ruta que su distancia fue de 30.37 km y duro 5.3 horas mientras que la figura 45 muestra una ruta que recorrió 81.1 km en 6.6 horas. Esto indica que la ruta que se mantuvo dentro del cordón urbano sin tomar ninguna vía rápida tuvo una duración de 1.3 horas menos que la que uso vías rápidas pero recorrió 2.33 veces más recorrido. Esto se puede explicar ya que los 50 km más de la ruta larga los realizó a una velocidad de 40 km/hora en la vía rápida lo que multiplicado por 1.3 horas da 52 km, así se puede explicar esta diferencia de distancia y de tiempo. Lo que conlleva esto es que el 80% del tiempo se demore dentro del cordón urbano. Y eso es lo que justifica la importancia de logística urbana y la última milla, porque el mayor consumo de recursos se da en esos momentos.

El segundo punto que se discutirá es como Jarygom cobra a sus clientes. La empresa cobra por km recorrido. Basandose en el análisis que se realizó en base a la figura 44 y 45 esto no es justo. Porque una ruta de 80 km no consumió 2.33 veces más recursos

que una ruta de 30 km. Esto permite inferir varias cosas o se cobra muy poco por las rutas cortas y lo justo por las largas o se cobra lo justo por las cortas y demasiado por las largas. Esto conlleva a 3 escenarios el primero es que en el promedio se equiparen las rutas largas y las cortas, algo complicado ya que como se vio no siguen una distribución normal las distancias. Lo siguiente es que la empresa esté perdiendo porque en las rutas cortas no cobra lo suficiente y la última opción es que se esté cobrando demasiado a la empresa contratante. Sea cual sea la opción tener una tarifa única solo basada en el kilometraje puede ser que no sea lo más justo, no se puede determinar para cuál de las partes, para eso se requerirá de estudios futuros.

4.7. Esquema de metodología

La metodología propuesta tiene 5 pasos principales que se pueden ver en la siguiente imagen.

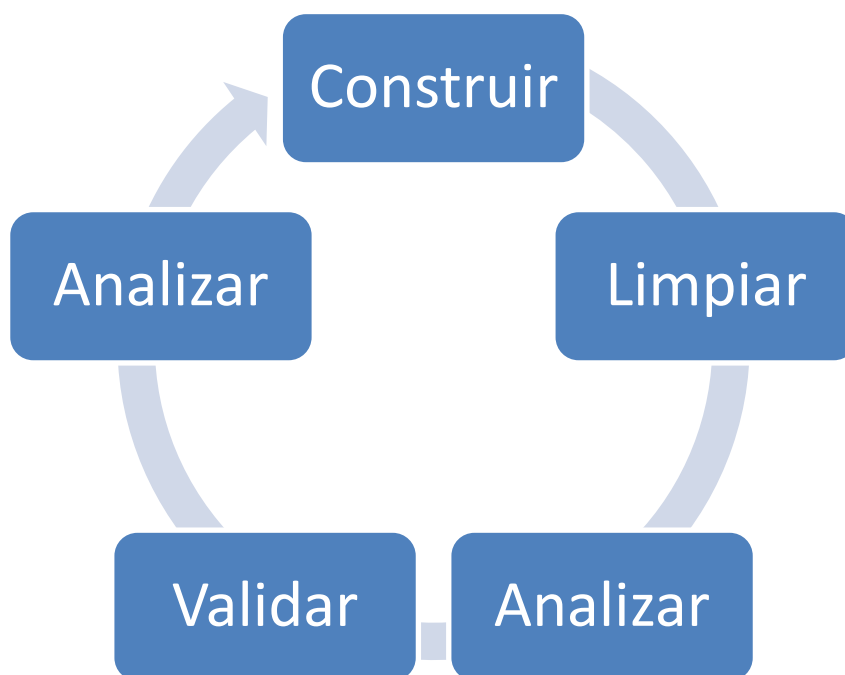


Figura 50. Pasos principales de la metodología de análisis. Fuente: Elaboración propia.

Estos 5 pasos son los que se llevaron a cabo en este análisis.

4.7.1 Construir

A continuación se ve los pasos que se deben llevar a cabo al momento de construir la base de datos. Esta se conforma de 3 pasos.



Figura 51. Proceso de la fase "Construir". Fuente: elaboración propia.

En la fase encuestar se deben llevar a cabo las siguientes encuestas que posteriormente permitirán definir como construir la base y como limpiar los datos.

Identificación de tipo de rutas a estudiarse.	
Quien define la ruta	I. Compañía: se sigue una ruta definida específicamente por la compañía. II. Conductor: el conductor de cada camión define la ruta
Periodicidad de la ruta	I. Diaria II. Semanal III. Mensual IV. No hay periodicidad
Horario de trabajo	I. 1 turno de 8 horas II. 2 turnos de 8 horas III. 24 horas
Lugares de parqueo	I. Definidos: Centros de operaciones II. Aleatorios: Casas de los choferes, parqueos en la ruta.
Camiones salen de la ciudad eventualmente.	I. Si II. No
GPS registra entregas	I. Si II. No
GPS manda señal con camión apagado	I. Si II. No

Figura 52. Encuesta de identificación de rutas. Fuente: Elaboración propia.

Identificación de tipo de datos a estudiarse.

Acceso a la plataforma de datos	I. Si II. No	
Tipo de datos.	I. CSV II. EXCEL III. ACCES IV. OTRO	Esto dependerá del programa que se esta utilizando. Pero es crucial para definir como se analizara la base de datos.
Periodicidad de los reportes	I. Diaria II. Semanal III. Mensual IV. Personalizada	Esto dependerá del programa que se esta utilizando. Pero es crucial para definir como se construirá la base de datos.
Reporte obtenido por	I. Vehículo II. Flota	Los reportes salen toda la flota junta o por cada camión
Todos los cortes mismo formato	I. Si II. No	
GPS registra con frecuencia constante	I. Si II. No	

Figura 53. Encuesta de identificación de tipo datos. Fuente: Elaboración Propia

Dependiendo si se tiene o no acceso a la plataforma, dependerá mucho que tipo de datos se va a obtener. Lo ideal es lo siguiente que se obtenga en Excel, por la periodicidad completa de estudio, por vehículo y todos los reportes o cortes en el mismo formato. Esto muy pocas veces pasa. Pero al menos se debe tener claro como vienen los datos crudos de la plataforma o la empresa.

Una vez ya con los datos se debe seguir o siguientes paso para construir la base. Si algún paso ya está hecho solo se continúa al siguiente.

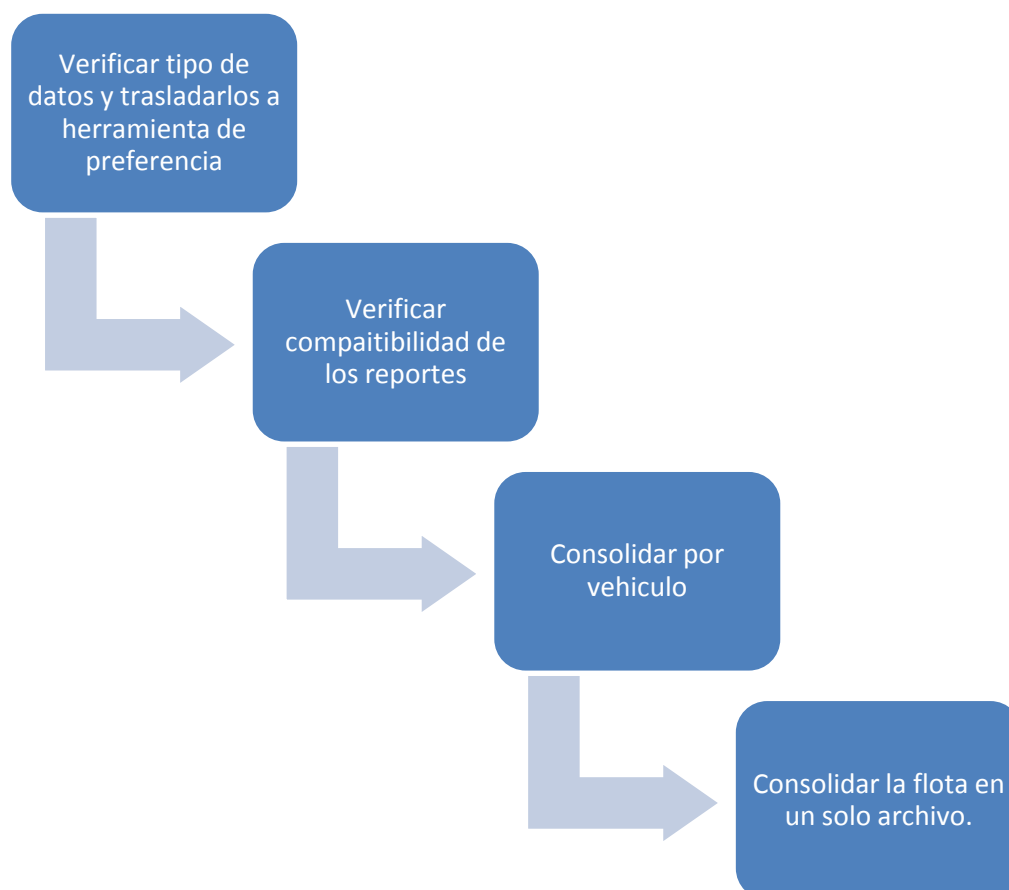


Figura 54. Pasos a seguir para consolidar la base de datos. Fuente: Elaboración propia.

Una vez realizado esto se continúa a la limpieza de los datos.

4.7.2. Limpiar

Para efectuar la limpieza de datos es crucial remitirse a la encuesta pero antes se debe observar por si sola a la base de datos y se debe dejar la siguiente información, muchas veces la base de datos posee información extra que no funciona en este tipos de análisis.

Tabla 20. Tipos de datos de la base de datos consolidada. Fuente: Elaboración Propia

Alias	Codigo ID	Fecha	Lat	Lon	GPS
-------	-----------	-------	-----	-----	-----

En la sección 4.2 se especifica en que consisten cada uno de estos datos. Lo siguiente es analizar la encuesta de tipo de datos para saber en qué aspectos se debe

enfocar la limpieza de datos. La siguiente figura muestra los pasos a seguir para la limpieza de datos.

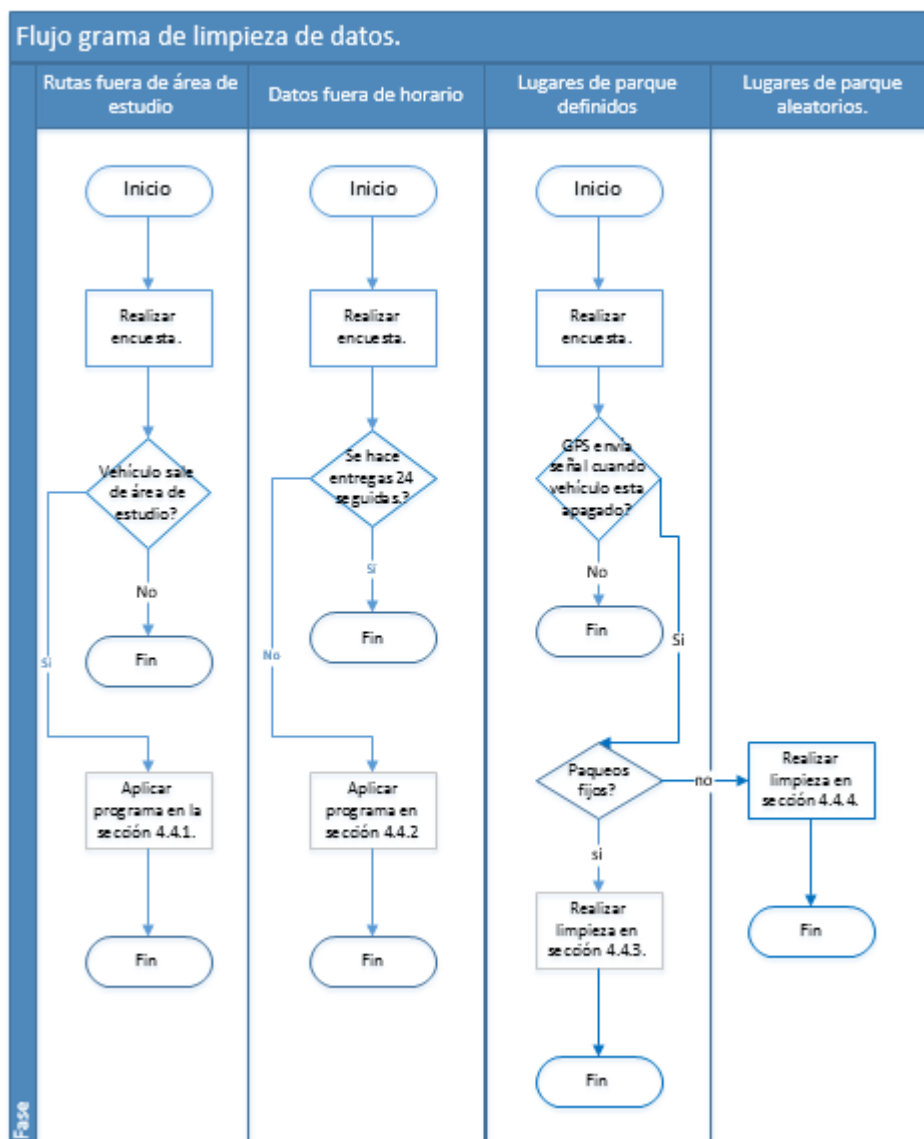


Figura 55. Flujo grama de limpieza de datos. Fuente: Elaboración propia.

4.7.3. Analizar

El siguiente paso es hacer el primer análisis, la sección 4.5 refleja lo que se hizo en este paso. Lo crucial es a la base ya limpia extraerle toda información posible para posteriormente validar este análisis.

4.7.4. Validar

Con métricas ya establecidas se debe validar las mismas, el método de validación es hacer muestreos aleatorios de las rutas y graficarlas. Aparte de eso las métricas que no hagan sentido también se debe graficar para encontrar la razón de ser de su error.

4.7.4. Analizar

Con las métricas ya validadas se procede a hacer análisis cualitativos y cuantitativos de las rutas, lo que se hizo en este paso se puede ver en la sección 4.4.6.

CAPITULO 5. CONCLUSIONES Y RECOMENDACIONES

El estudio de los datos de rastreo de GPS de la flota de entrega de helados en Quito de la empresa Jarygom completo todos los objetivos planteados al inicio. La suma de una base teórica sólida, varias iteraciones en el proceso de experimentación, el uso de un programa robusto para el análisis de datos, genero un esquema de creación, limpieza y análisis de la base de datos. Esta metodología o esquema se puede aplicar a cualquier flota que posea rastreo de GPS, y se adapta al giro del negocio del o los productos que se esté entregando, no es exclusiva para la entrega de helados ni para la ciudad de Quito. A continuación se presentan los hallazgos más significativos del estudio, sus limitaciones, recomendaciones y posibles futuros estudios partiendo del actual.

5.1 Hallazgos significativos del estudio

Establecer una metodología de análisis de datos de GPS aplicable a la entrega de productos de diferentes industrias. Que incluye la construcción de la base de datos, la limpieza de los datos que no poseen información y el análisis de los datos que se consideran válidos.

Se logró graficar las rutas en mapas actualizados y geo referenciados, que poseen la información de todas las calles de la ciudad, permiten aumentar disminuir el zoom.

Se logró manejar la herramienta de rastreo, para obtener los datos de rastreo de 15 camiones durante 7 meses. Se consolido la base de estos 15 camiones en Excel y se los migro con éxito a Matlab.

La unión de varias herramientas para el manejo de datos es lo que permite llevar a cabo la metodología propuesta. En este caso Excel, Matlab y Minitab, fueron las herramientas escogidas. Por tres razones facilidad de uso (Minitab y Excel) y por su versatilidad y capacidad de análisis de grandes matrices de datos (Matlab)

En relación a las métricas se escogió 3 cuantitativas para analizar, distancia de la ruta, tiempo de movimiento de la ruta, y velocidad promedio de la ruta. Para entender estas métricas es crucial realizar un análisis cualitativo de las rutas para lograr realmente identificar qué es lo que estas métricas significan.

Los valores de las métricas obtenidas son razonables se validó con la empresa el valor de los kilometrajes y tiempos. La empresa acepto que esas distancias son aproximadas a la realidad. Lo que no estaba al tanto era su velocidad promedio de movimiento que no se imaginaban que era tan baja.

Se encontraron 262 rutas menores a 12 kilómetros, que hacen viajes que no son de entrega sumando alrededor de 1700 km y 626 horas o 78 días. Estas rutas son absorbidas por la empresa. Generando un costo solo de mano de obra de 2730 USD

5.2. Limitaciones del estudio

Los datos de GPS no son enviados con una frecuencia constante, lo que dificulta el análisis.

No se validó la diferencia real de las métricas obtenidas con una investigación de campo dentro de un camión.

Al ser un estudio relativamente nuevo en el medio gran parte de las soluciones son en base a prueba y error ya que la literatura es muy limitada acerca de estos temas. Al ser prueba y error la mayoría de los programas de análisis el tiempo de desarrollo fue mucho más extenso del esperado, lo que limito entrar en otras áreas de estudio.

No se ligó el análisis de las rutas a otro tipo de recurso como combustible. Tampoco se relación a ubicación geográfica de los minoristas porque esta era inexistente.

5.3. Recomendaciones

Se recomienda a Jarygom eliminar los viajes entre el centro de distribución y su centro de operaciones, ya que son un desperdicio de recursos. Por política se debería prohibir ese tipo de viajes. Una posible solución que debe ser analizada más a profundidad es que si se necesita ir al centro de operaciones este se debe incluir como una parada más dentro de una ruta de entrega para que no se genere un costo extras sino se amortice esta parada en el costo de la ruta que si se cobra al contratante.

Otra recomendación es que se debe generar un método de cobro diferenciado. Posiblemente se deba mantener por distancia recorrida pero se debería hacer un ajuste a las rutas en función de su distancia. Las rutas cortas que no salen del cordón urbano deberían tener un valor mayor que las más largas.

5.4. Futuros estudios

Partiendo de este análisis nacen algunas opciones de estudio las 3 más valiosas se detallan a continuación.

Realizar un estudio que investigue si en logística urbana el consumo de combustible es proporcional a la distancia.

Diseñar un modelo de investigación de operaciones que no minimice distancia recorrida sino tome en cuenta el flujo. Esto quiere decir que trate de utilizar la mayor cantidad de vías perimetrales y minimizar el recorrido dentro de la ciudad para así minimizar el consumo de recursos y no solo la distancia.

Crear un modelo de cobranzas diferenciado que usando cercas geográficas, cobre un precio por vía perimetral y otro por vía urbana, en función del consumo de recursos de cada uno de los escenarios.

CAPITULO 6. TRABAJOS CITADOS

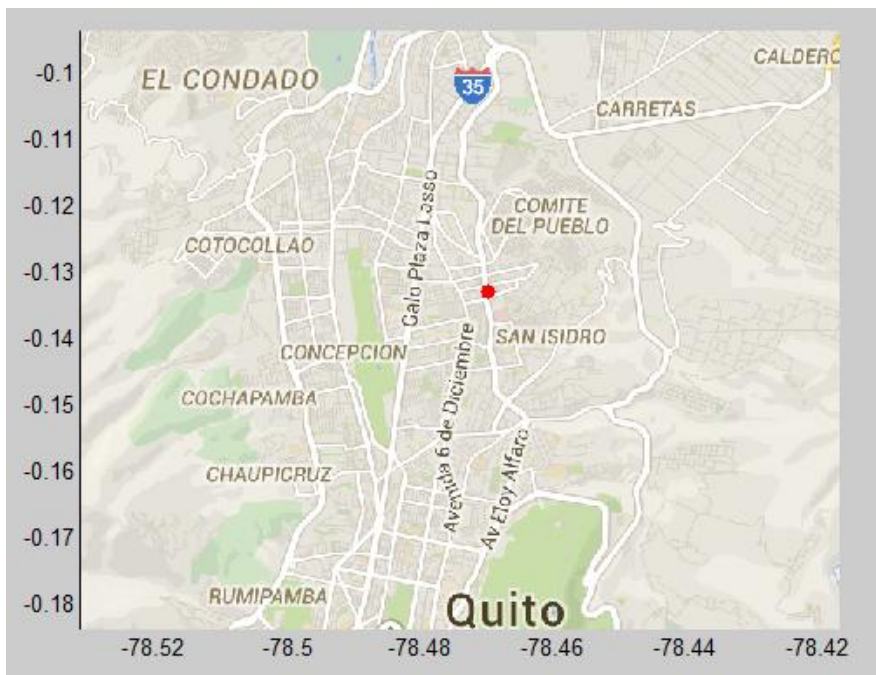
- Agencia Pública de Noticias. (14 de 04 de 2010). *Noticias Quito*. Recuperado el 6 de 8 de 2014, de http://www.noticiasquito.gob.ec/Noticias/news_user_view/municipio_de_quito_implementara_3_tipos_de_parqueaderos_cuando_se_aplique_el_pico_y_placa--139
- Agencia Pública de Noticias de Quito. (26 de Diciembre de 2012). *Movilidad*. Recuperado el 20 de Enero de 2014, de Noticias Quito: http://www.noticiasquito.gob.ec/Noticias/news_user_view/410000_vehiculos_se_revisarian_hasta_fin_de_ano--8118
- AIMMS. (2014). *AIMMS the Modeling System*. Obtenido de <http://www.aimms.com/about-us/>
- America Economía. (2013). *Las mejores ciudades para hacer negocios 2013*. Recuperado el 20 de Enero de 2014
- Boone, T., Ganesham, R., & Jarayaram, V. (s.f.). Carbon Foot Print of Products—Supply Chain.
- Bull, A. (2003). *Congestion de Transito*. Santiago de Chile: CEPAL.
- Chakraborty, J., Scheweitzer, L. A., & Forkenbrock, D. J. (1999). Using GIS to Assess the Environmental Justice Consequences of Transportation System Changes. *Transactions In GIS*, 3(3), 239.
- Chang, R. (2007). *Quimica*. Mexico DF: McGraw Hill.
- Comendadora, J., López-Lambasb, M. E., & Monzónb, A. (2012). A GPS analysis for urban freight distribution. *Procedia - Social and Behavioral Sciences*, 521 – 533.
- Coppock, J. T., & Rhind, D. W. (1991). The History of GIS. En D. W. Rhind, M. F. Goodchild, & D. J. Maguire, *Geographic Information Systems and Science* (págs. 9-20). NA: Wiley. Obtenido de <http://www.wiley.com/legacy/wileychi/gis/volumes.html>
- Cordova, J. P. (2013). *Logística Urbana en Quito: Esquema de distribución para DINADEC (Cervecería Nacional) en el Centro Histórico de Quito*. Quito: USFQ.
- Cplusplus. (2014). Obtenido de <http://www.cplusplus.com/doc/tutorial/operators/>
- DANA, P. H. (1997). Global Positioning System (GPS) Time Dissemination for Real-Time Applications. *Real-Time systems*, 12, 9-40.
- Directorate-General for energy and Transport. (2006). *Urban Freight Transport and Logistic*. Leeds: European Communities,.
- Ecker, J. G., & Kupferschmid, M. (1988). *Introduction to Operation Research*. New York: John Wiley.
- Editorial El Comercio. (4 de Mayo de 2013). EL MANEJO DEL TRÁNSITO EN QUITO. *El Comercio*. Obtenido de <http://www.elcomercio.com/opinion/editorial/manejo-del-transito-quito.html>
- Ehmke, J. F. (2012). *Integration of Information and Optimization Models for Routing in City*. Nueva York: Springer Science+Business Media. doi:10.1007/978-1-4614-3628-7_2
- El Banco Mundial. (2014). *Banco Mundial*. Obtenido de Datos: <http://datos.bancomundial.org/indicador/SP.POP.GROW>
- El Mercurio. (07 de 09 de 2012). Tránsito satura calles y avenidas. *El Mercurio*. Obtenido de http://www.elmercurio.com.ec/347908-transito-satura-calles-y-avenidas/#.VBskd_15MYM
- El Telegrafo. (14 de Agosto de 2013). *El parque automotor crece más que la población*. Recuperado el 19 de Enero de 2014, de El Telegrafo:

- <http://www.telegrafo.com.ec/noticias/quito/item/el-parque-automotor-crece-mas-que-la-poblacion.html>
- ESRI. (2013). *What is GIS*. Recuperado el 22 de Enero de 2014, de ESRI understanding the world: http://www.esri.com/what-is-gis/overview#overview_panel
- Ghiani, G., Laporte, G., & Musmanno, R. (2004). *Introduction to logistics systems planning and control*. Hoboken: J. Wiley. Recuperado el 24 de Enero de 2014
- Gocmen, A. Z., & Ventura, S. J. (2010). Barriers to GIS Use in Planning. *Journal Of The American Planning Association*, 76(2), 172-183. doi:10.1080/01944360903585060
- Gonzalez-Feliu, J., Pluvinet, P., Serouge, M., & Gardrat, M. (2013). *GPS-based data production in urban freight distribution*. Lyon: Laboratoire d'Economie des Transports,.
- Goodchild, M. (2000). GIS and Transportation: Status and Challenges. *Geoinformatica*, 4(2), 127.
- Greaves, S. P., & Figliozzi, M. A. (2008). Collecting Commercial Vehicle Tour Data with Passive Global Positioning System Technology. *Transportation Research Record: Journal of the Transportation Research Board*(2049), 158-166. doi:10.3141/2049-19
- Greco, F. (2008). *Traveling Salesman Problem*. Vienna: In-teh.
- Hemmelmayr, V., Cacchiani, V., & Tricoire, F. (2014). A set-covering based heuristic algorithm for the periodic vehicle routing problem. *Discrete Applied Mathematics*, 163(Part 1), 53-64. doi:10.1016/j.dam.2012.08.032
- Hillier, F. S., & Lieberman, G. J. (2006). *Introducción a la investigación de operaciones*. México DF: McGraw-Hill.
- Ho, D. (2014). *Notepad ++*. Obtenido de <http://notepad-plus-plus.org/>
- INEC. (2011). *En Quito existen 101.937 establecimientos económicos según el Censo Nacional Económico*. Recuperado el 2 de 5 de 2014, de http://www.inec.gob.ec/inec/index.php?option=com_content&view=article&id=292%3Aen-quito-existen-101937-establecimientos-economicos-segun-el-censo-nacional-economico&catid=68%3Aboletines&Itemid=51&lang=es
- International Biological System Information. (2003). *Introduction to Global Positioning Systems*. Obtenido de International Biological System Information: http://ibis.colostate.edu/WebContent/WS/CitSci/Tutorials_Wisconsin/Tutorial2_Static.html
- Invierta en Quito. (2010). *¿Porque invertir en Quito?, Entorno económico*. Recuperado el 20 de Enero de 2014, de http://www.inviertaenquito.com/index.php?option=com_content&view=article&id=36&Itemid=91&lang=es
- Jarrin, A. (15 de Mayo de 2014). Funcionamiento JaryGom. (F. Bueno, Entrevistador)
- Johnson, D. S., & McGeoch, L. A. (1995). *The Traveling salesman problem: A case study in local optimization*. Amherst.
- Kouki, Z., Chaar, B. F., & Ksouri, M. (2009). SP based Evolutionary optimization approach for the Vehicle Routing Problem. *AIP Conference Proceedings*, 1107(1), 373-376. doi:10.1063/1.3106506
- Kovács, G., & Spens, K. (2011). Trends and developments in humanitarian logistics a gap analysis. *Internationa Journal of Physical Distribution & Logistics Management*, 32-45.
- La hora. (31 de Mayo de 2012). 209 puntos conflictivos para el tránsito vehicular. *La hora*. Obtenido de <http://www.lahora.com.ec/index.php/noticias/show/1101338100/->

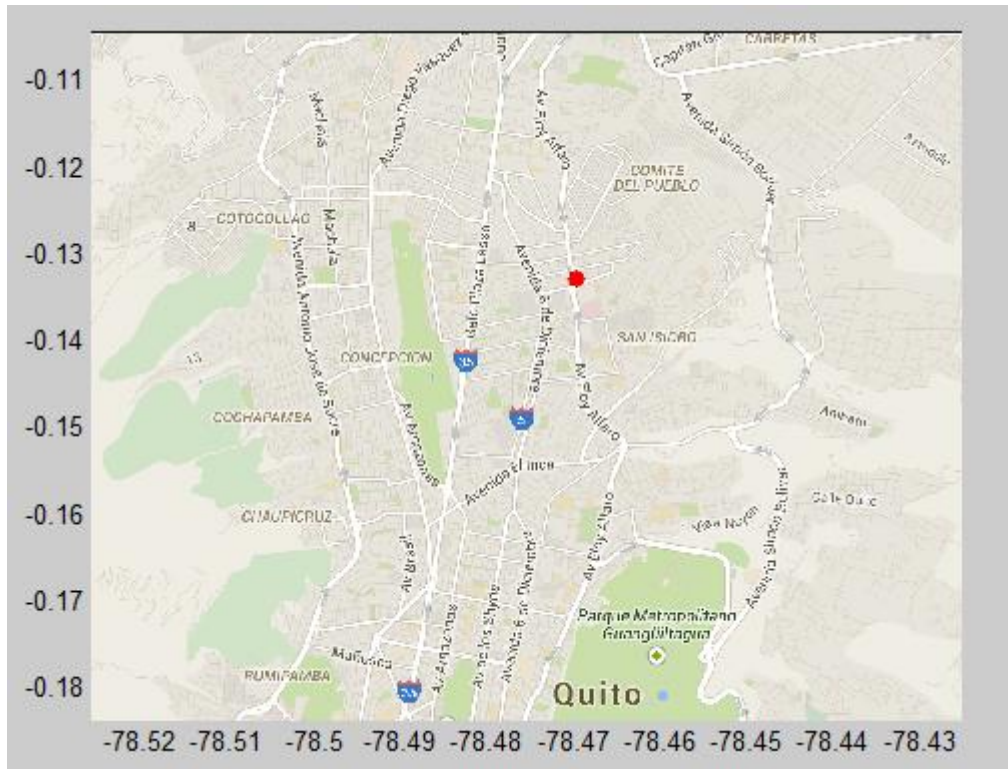
- 1/209_puntos_conflictivos_para_el_tr%C3%A1nsito_vehicular.html#.VCL7W_15MYM
- Lionheart Publishing. (Agosto de 2000). *ORMS*. Obtenido de AIMMS Review:
<http://www.orms-today.org/orms-8-00/swr.html>
- Mathworks. (2014). *Mathworks*. Obtenido de
<http://www.mathworks.com/help/map/ref/distance.html>
- Mathworks. (2014). *Mathworks*. Obtenido de
<http://www.mathworks.com/help/map/ref/deg2km.html>
- Mathworks. (2014). *Mathworks*. Obtenido de Cat:
<http://www.mathworks.com/help/matlab/ref/cat.html>
- Matworks Inc. (2014). Obtenido de
http://www.mathworks.com/index.html?s_tid=gn_logo
- MDS Transmodal Limited. (2012). *European Commission: Study on Urban Freight Transport*. EU. Obtenido de
<http://ec.europa.eu/transport/themes/urban/studies/doc/2012-04-urban-freight-transport.pdf>
- Mora Garcia, L. A. (2008). *Gestión logística integral*. Bogota: Ecoe Ediciones.
Recuperado el 22 de Enero de 2014
- National Coordination Office for Space-Based Positioning, Navigation, and Timing. (2014 de marzo de 17). *GPS.gov*. Obtenido de
<http://www.gps.gov/systems/gps/performance/accuracy/>
- Pluvineta, P., Gonzalez-Feliua, J., & Ambrosinia, C. (2012). GPS data analysis for understanding urban goods movement. *Procedia - Social and Behavioral Sciences*, 39, 450-462.
- Reiter, S., & Francoise, A.-F. (2012). Toward Low Energy Cities. *Journal Of Industrial Ecology*, 16(6), 829-838. doi:10.1111/j.1530-9290.2012.00533.x
- Sadiq, S. (2012). The Traveling Salesman Problem: Optimizing Delivery Routes Using Genetic Algorithms. *SAS Global Forum 2012*(161). Obtenido de
<http://support.sas.com/resources/papers/proceedings12/161-2012.pdf>
- Saunders, M. J., & Rodrigues da Silva, A. N. (2009). Reducing Urban Transport Energy Dependence: A New Urban Development Framework and GIS-Based Tool. *International Journal Of Sustainable Transportation*, 3(2), 71-87.
doi:10.1080/15568310701648037
- Secretaria de Territorio, Habitat y Vivienda. (2014). *Secretaria de Territorio, Habitat y Vivienda*. Obtenido de
http://sthv.quito.gob.ec/spirales/9_mapas_tematicos/9_1_cartografia_basica/9_1_1_1.html
- Stopher, P., Jiang, Q., & Fitzgerald, C. (2005). *Processing GPS Data From Travel Surveys*. Sydney: University of Sydney.
- UN. (2007). *UNFPA State of the world population 2007*. Nueva York.
- University of Glasgow. (3 de 11 de 2014). *University of Glasgow*. Obtenido de Statistics tutoria, Choosing a T test:
http://www.gla.ac.uk/sums/users/jdbmcdonald/PrePost_TTest/chooset2.html
- USGS. (13 de Febrero de 2007). *USGQ science for a changing world*. Recuperado el 7 de Mayo de 2014, de Global Geographic Information Systems:
http://webgis.wr.usgs.gov/globalgis/tutorials/what_is_gis.htm



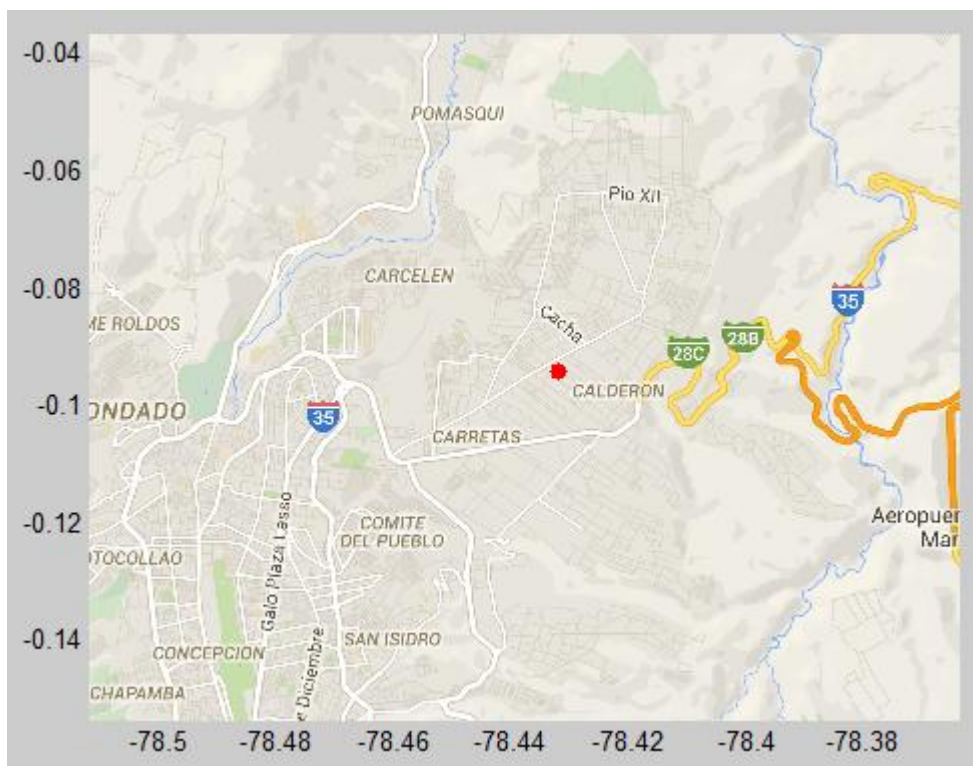
Ruta#131 Camión 208 26 julio



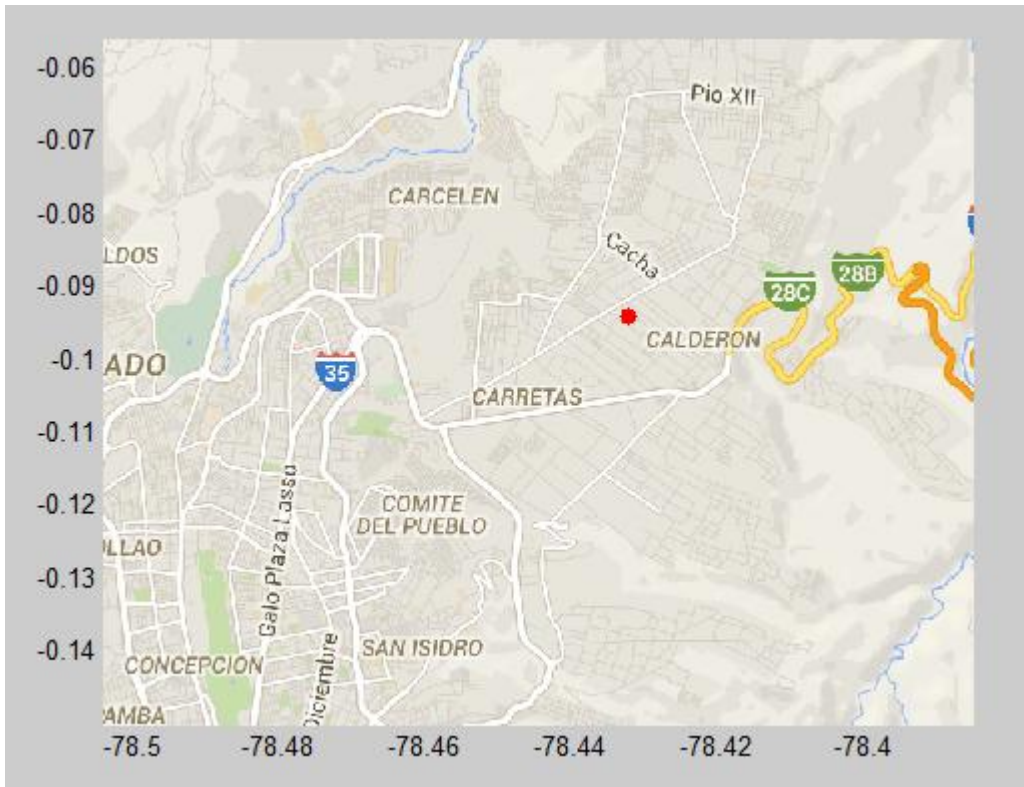
Ruta#139 Camión 208 20 agosto



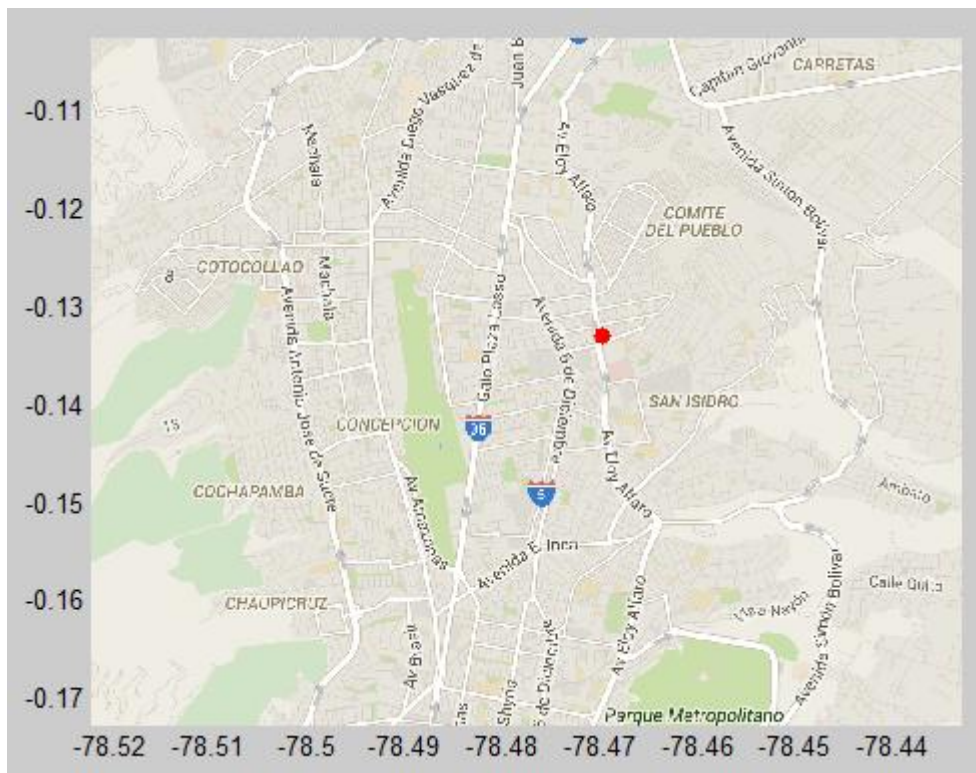
Ruta 284 Camión 302 12 de julio



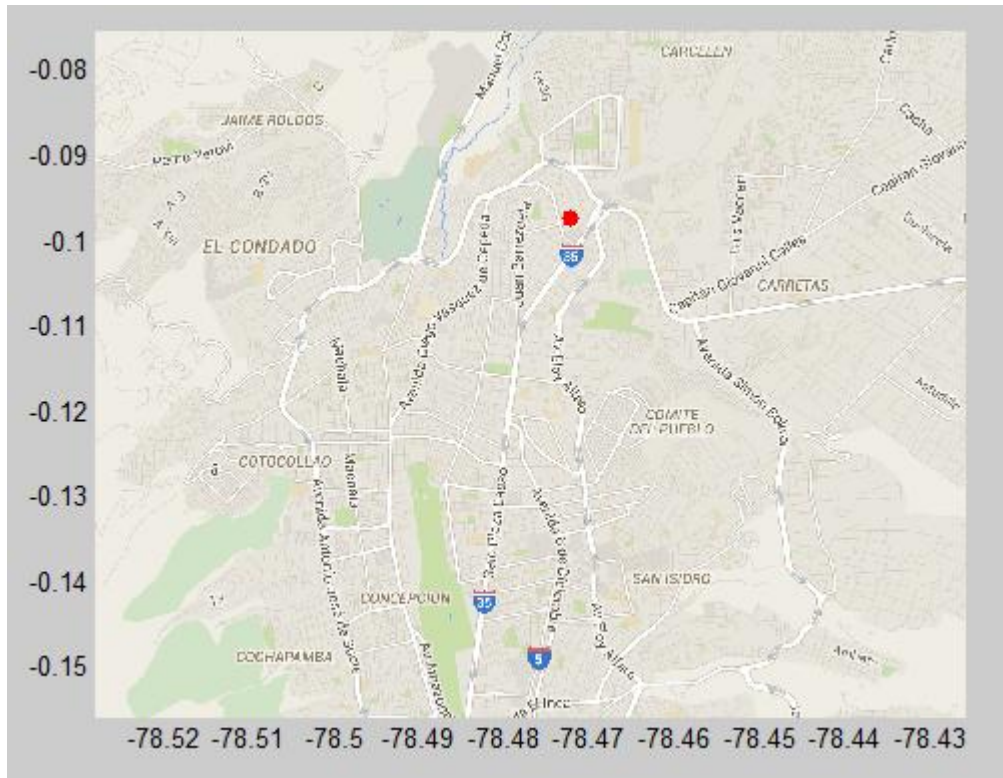
322614



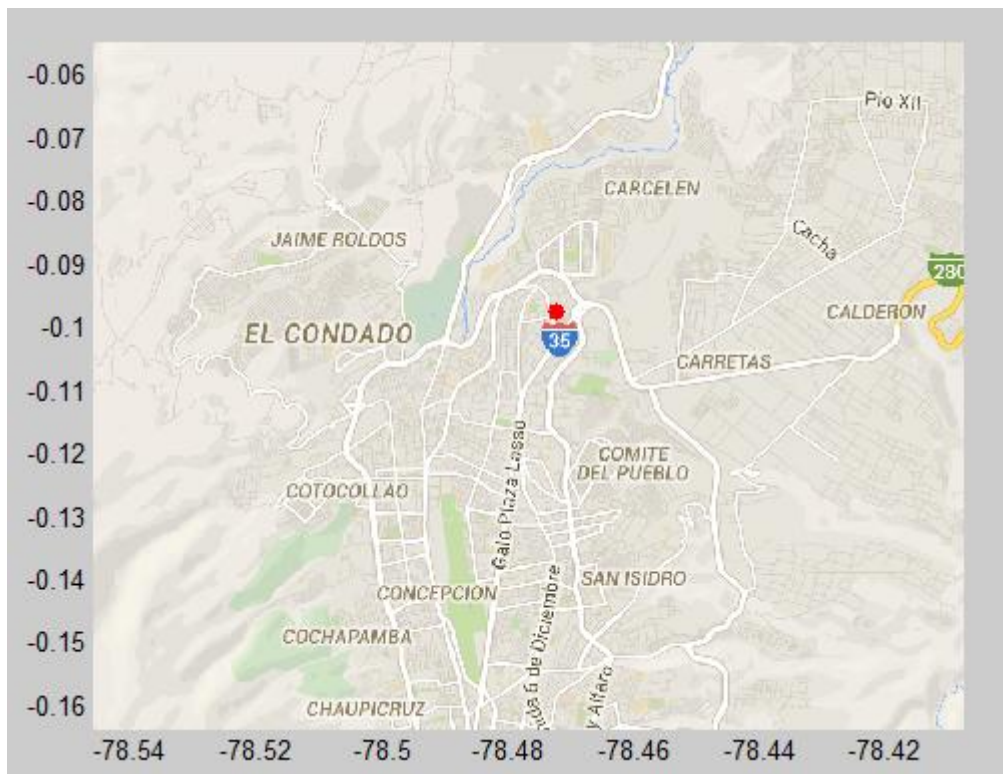
33346



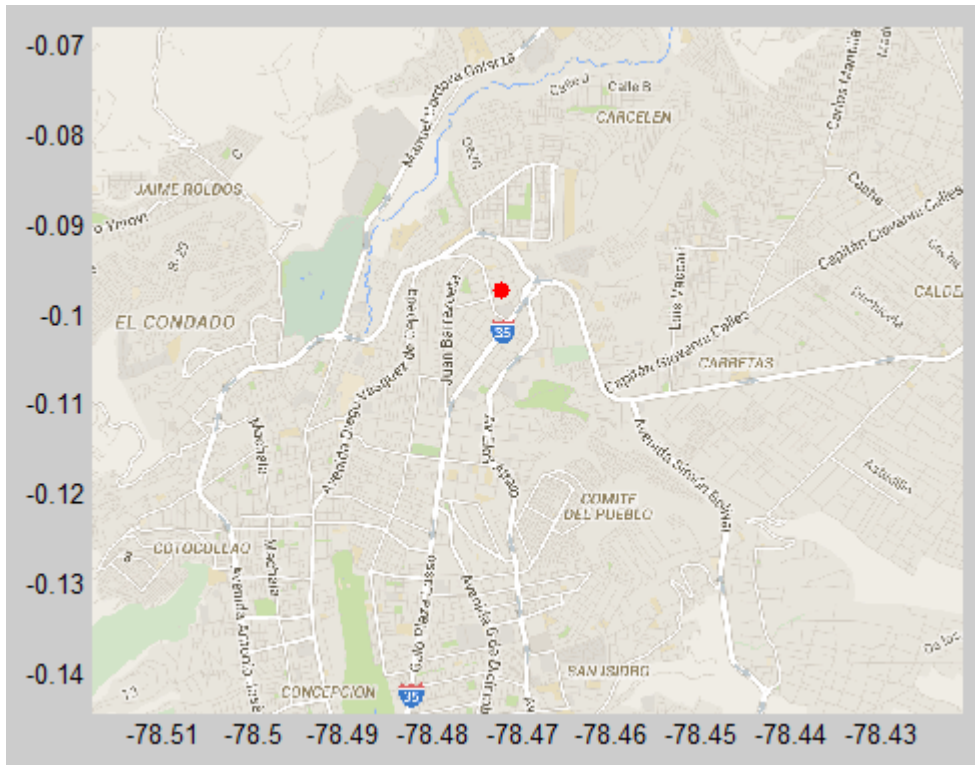
333518



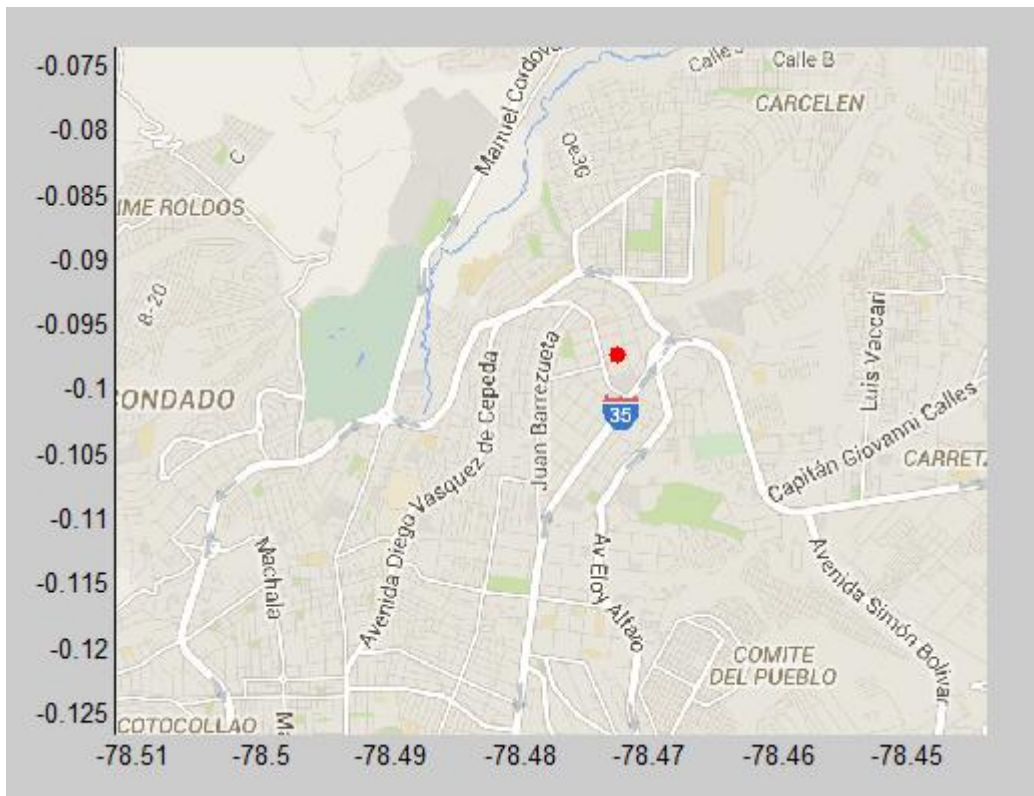
33546



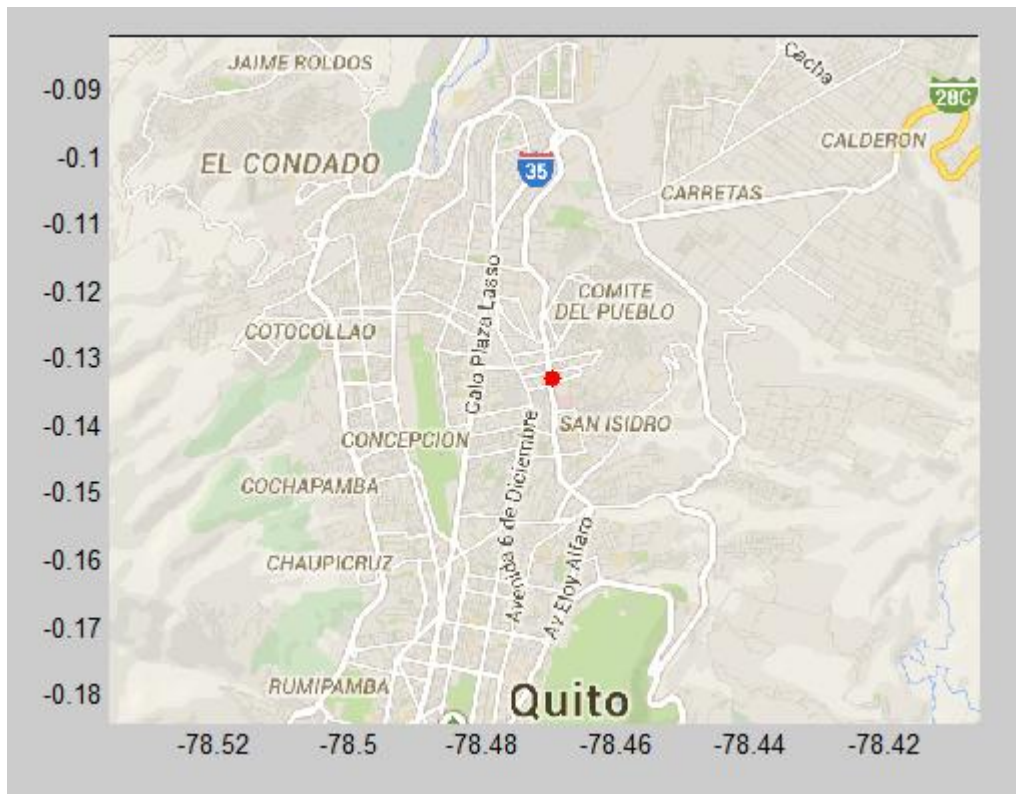
335821



33646

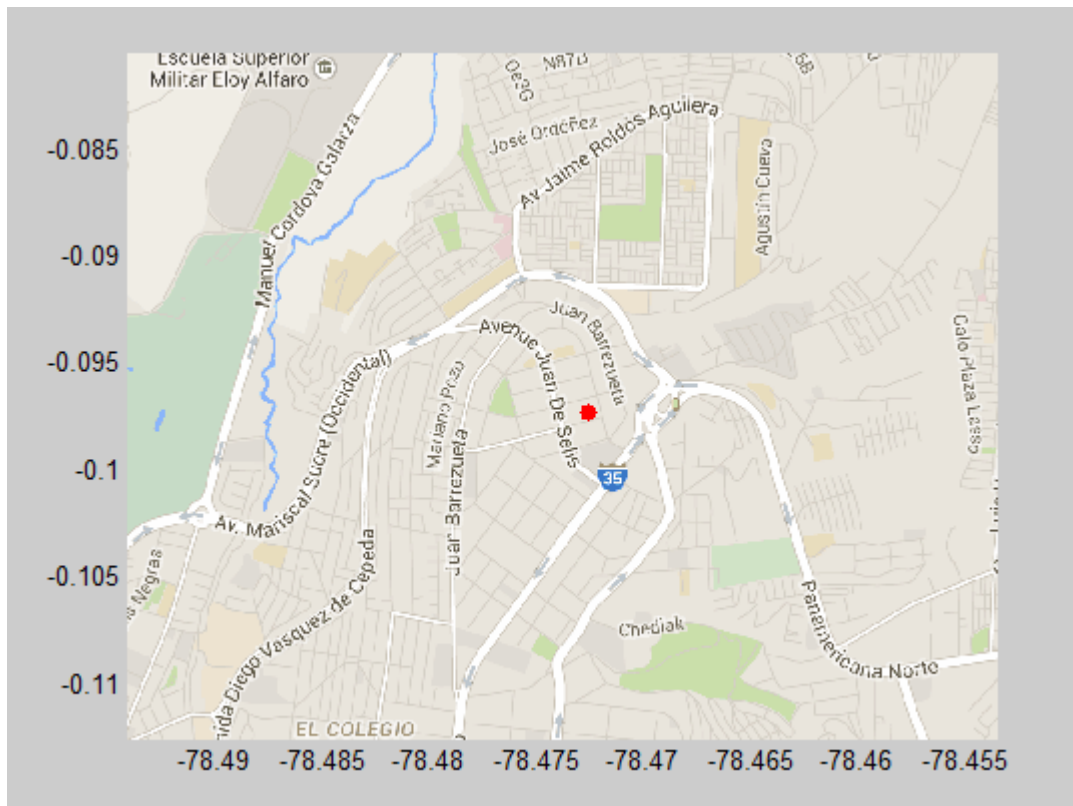


351723

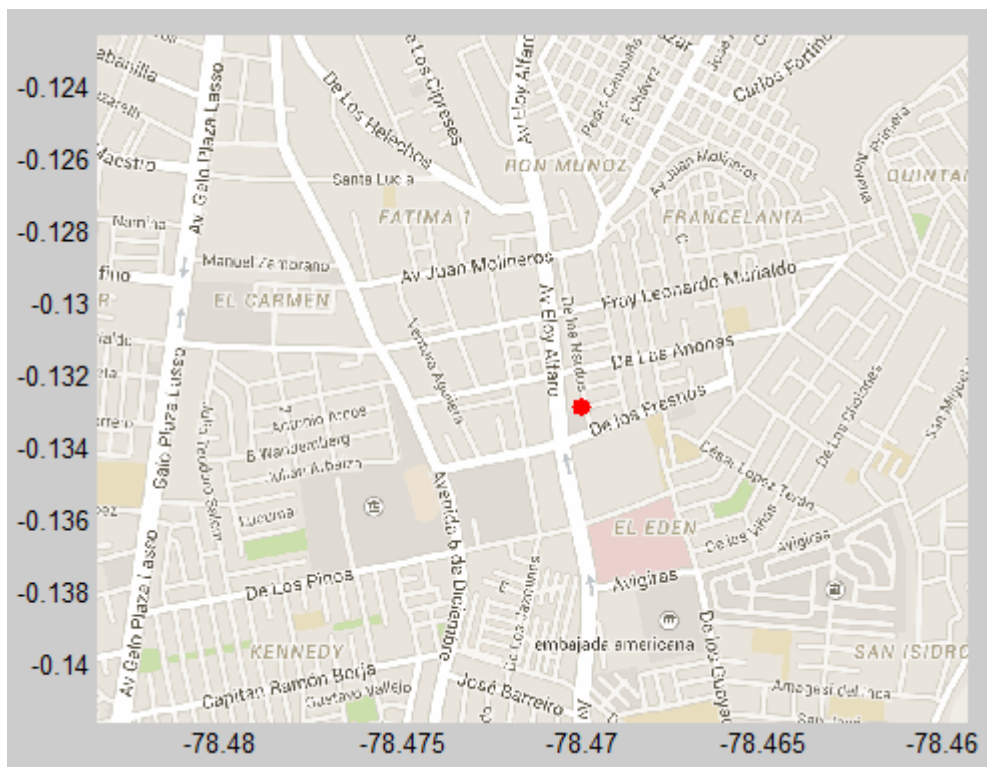


351812

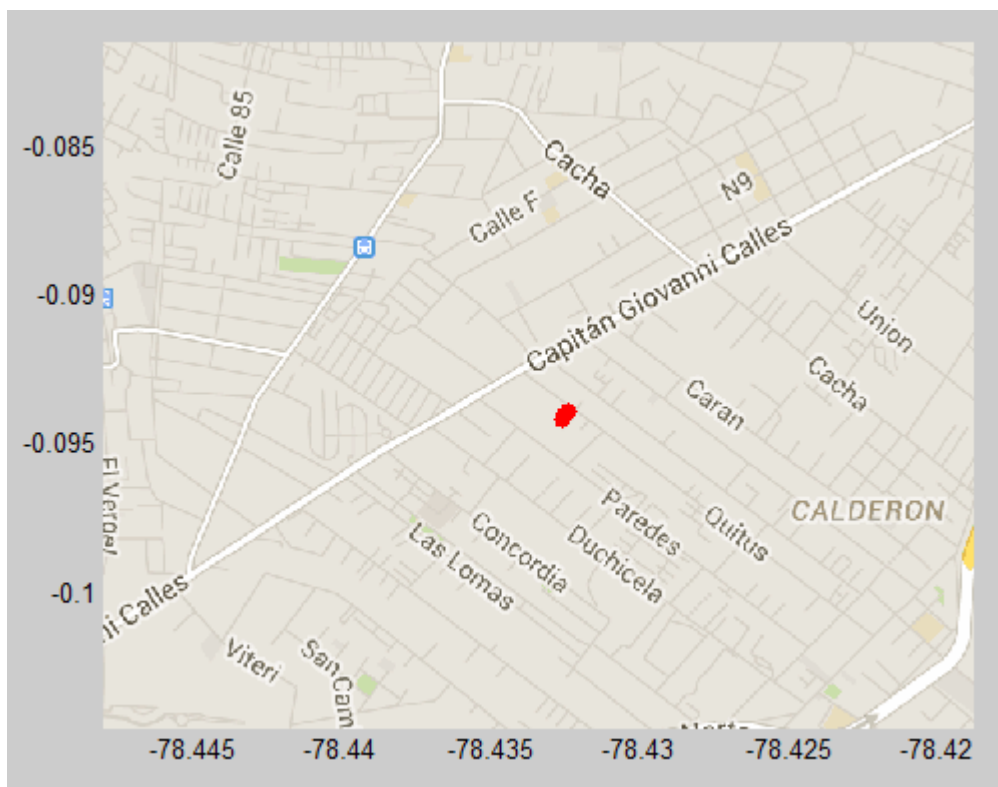
ANEXO 2



Ruta 1092 Camión 336 10 de agosto

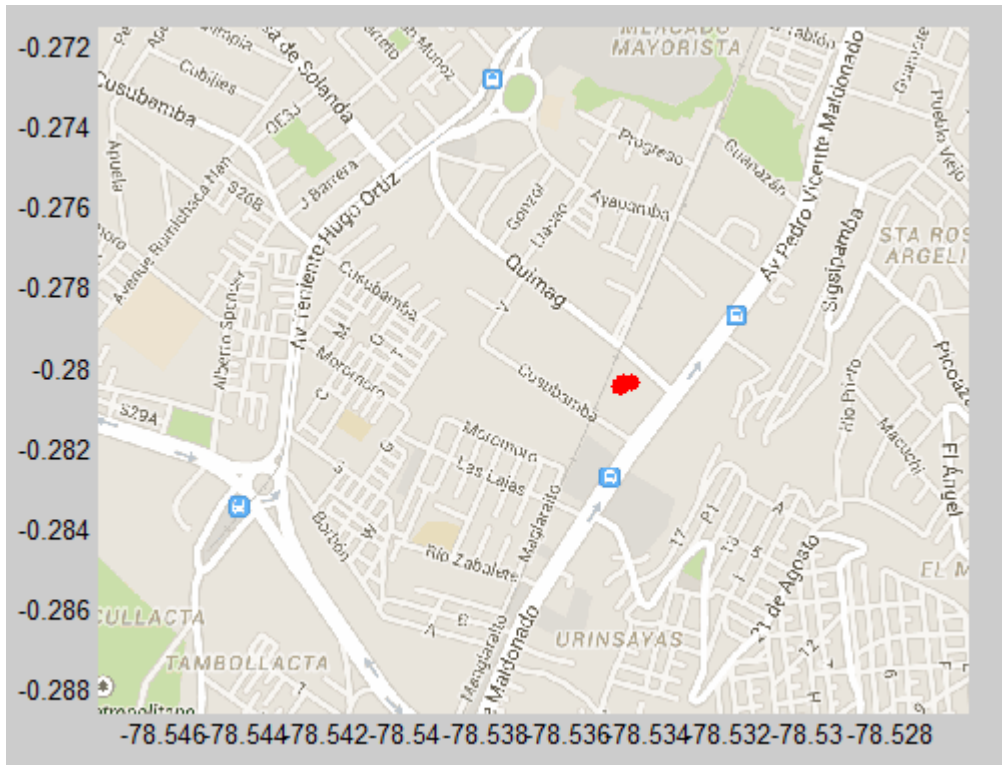


Ruta 130 208722

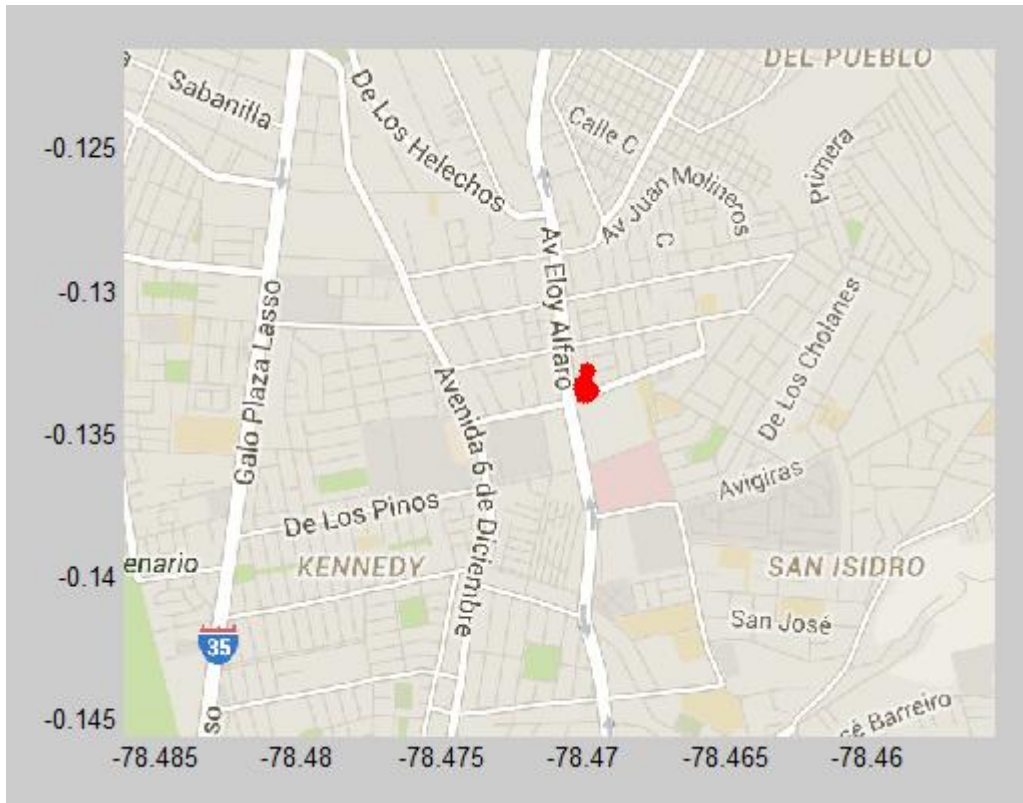


Ruta 396 322413

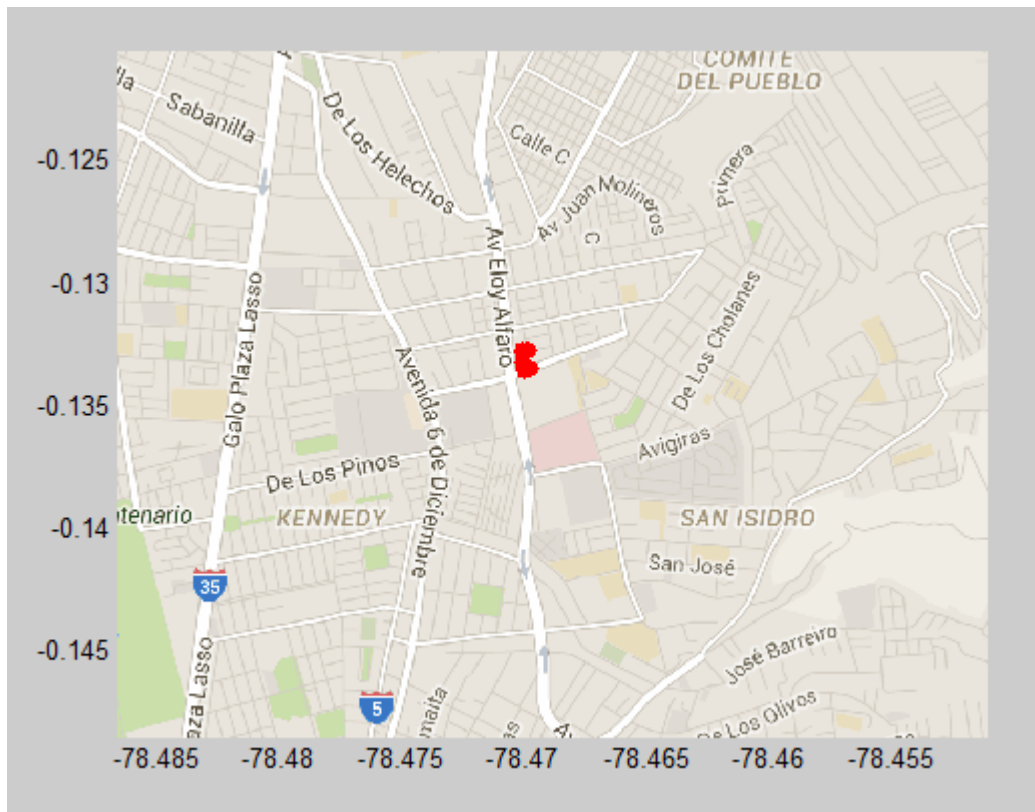
304 321512



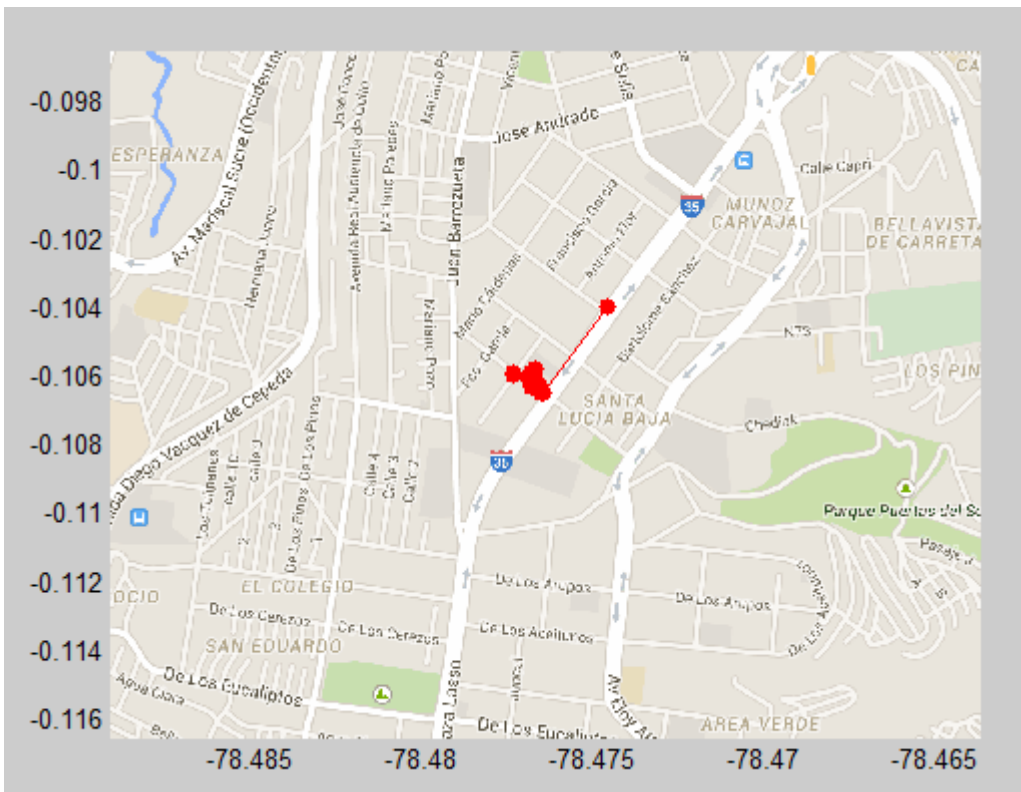
Ruta 266 302619



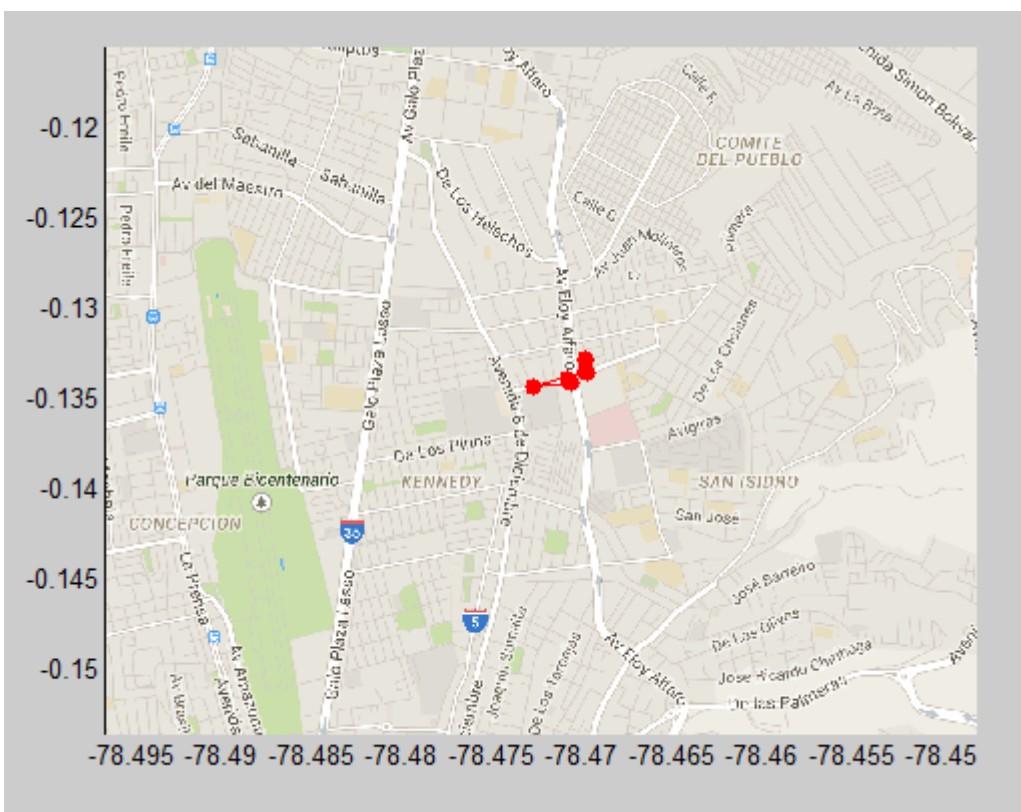
Ruta 6 201715



Ruta 8 201717

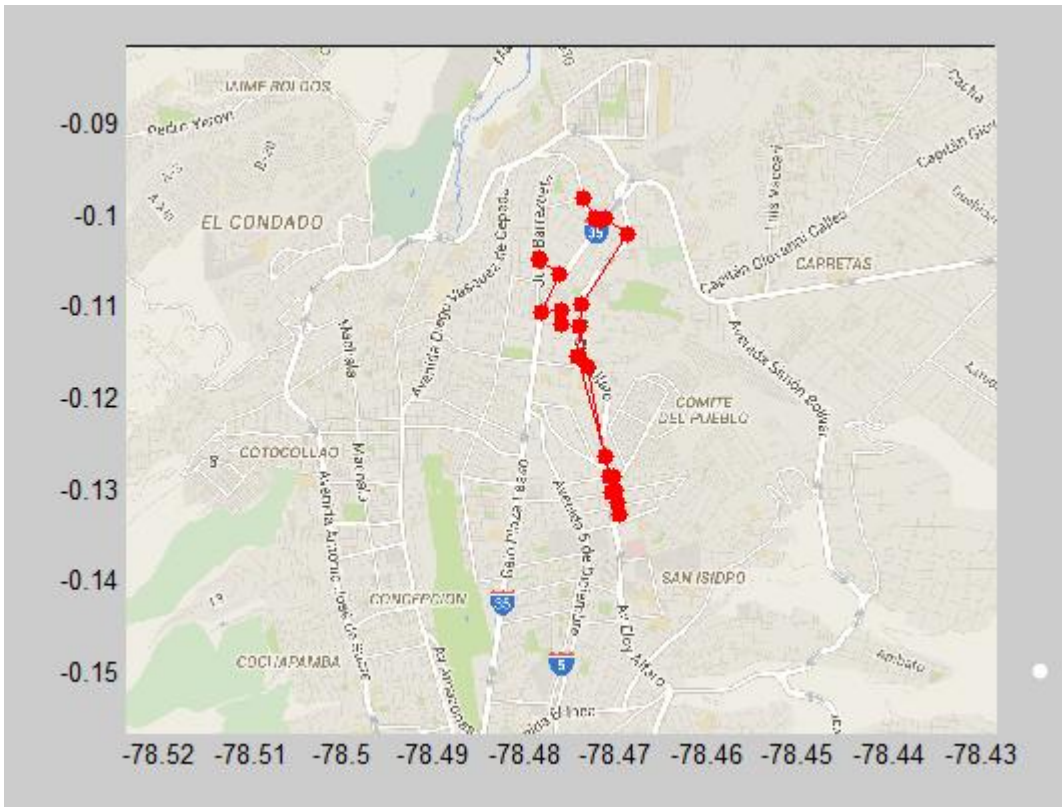


Ruta 1572 351630

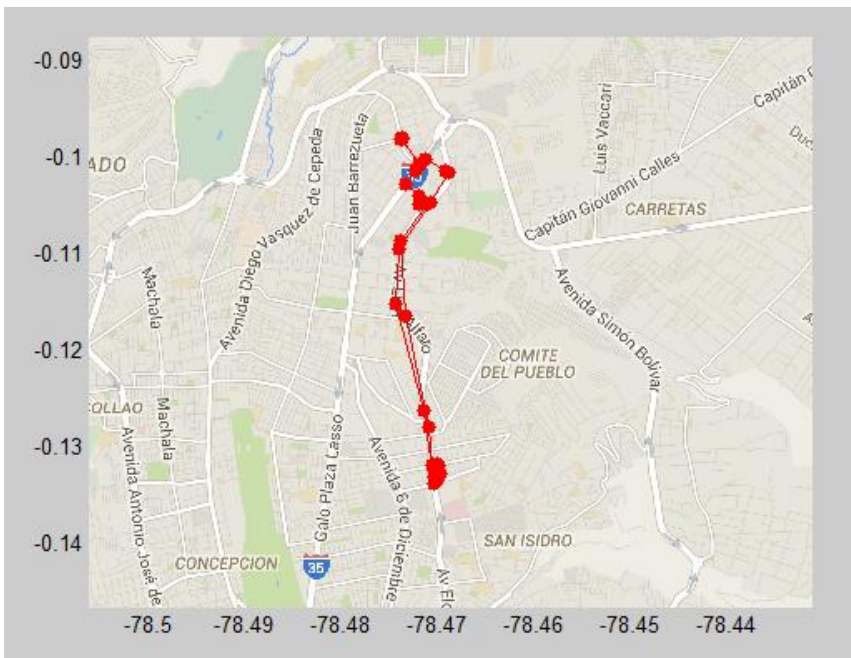


Ruta 359 32181

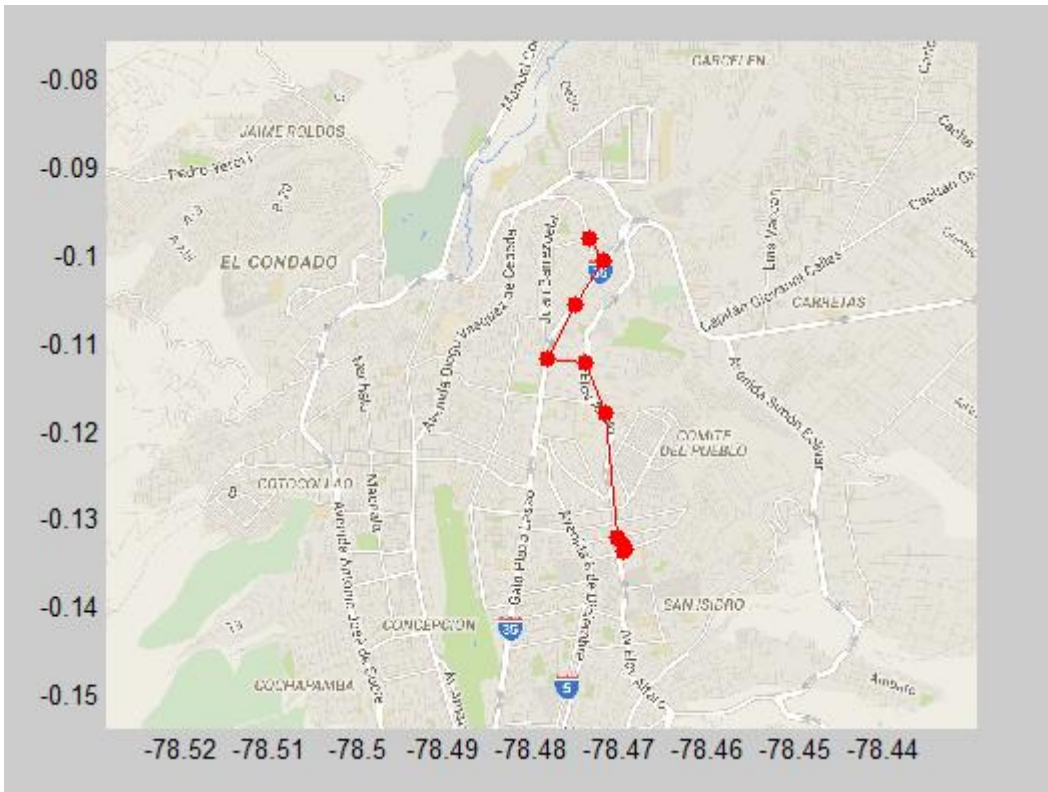
ANEXO 3



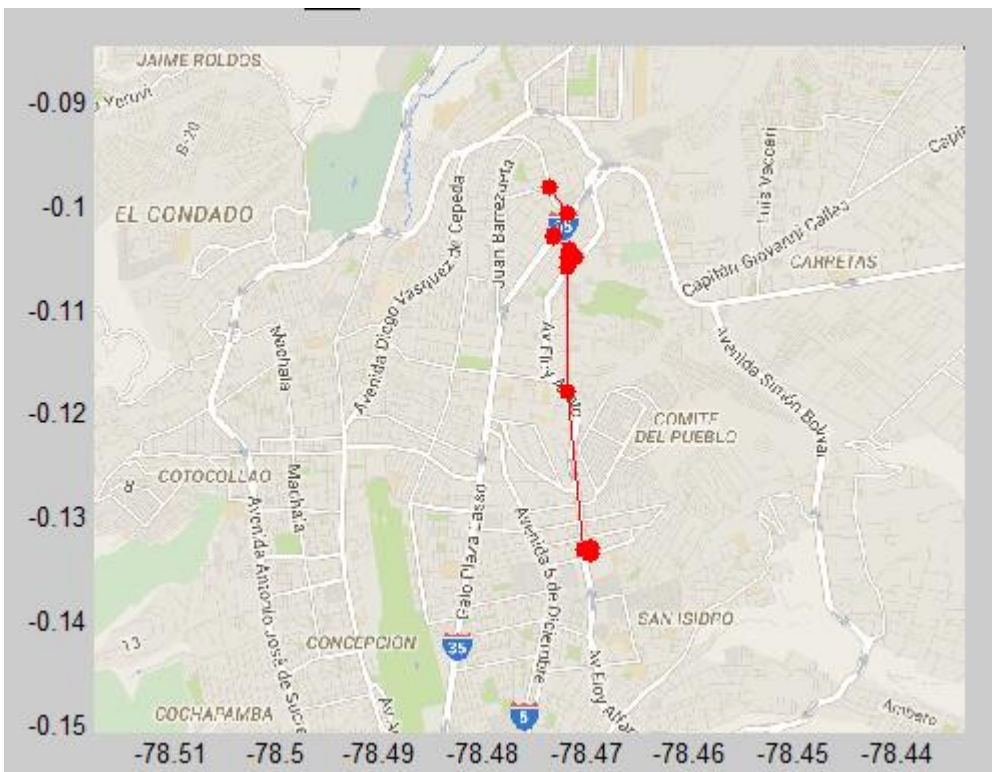
806 333715



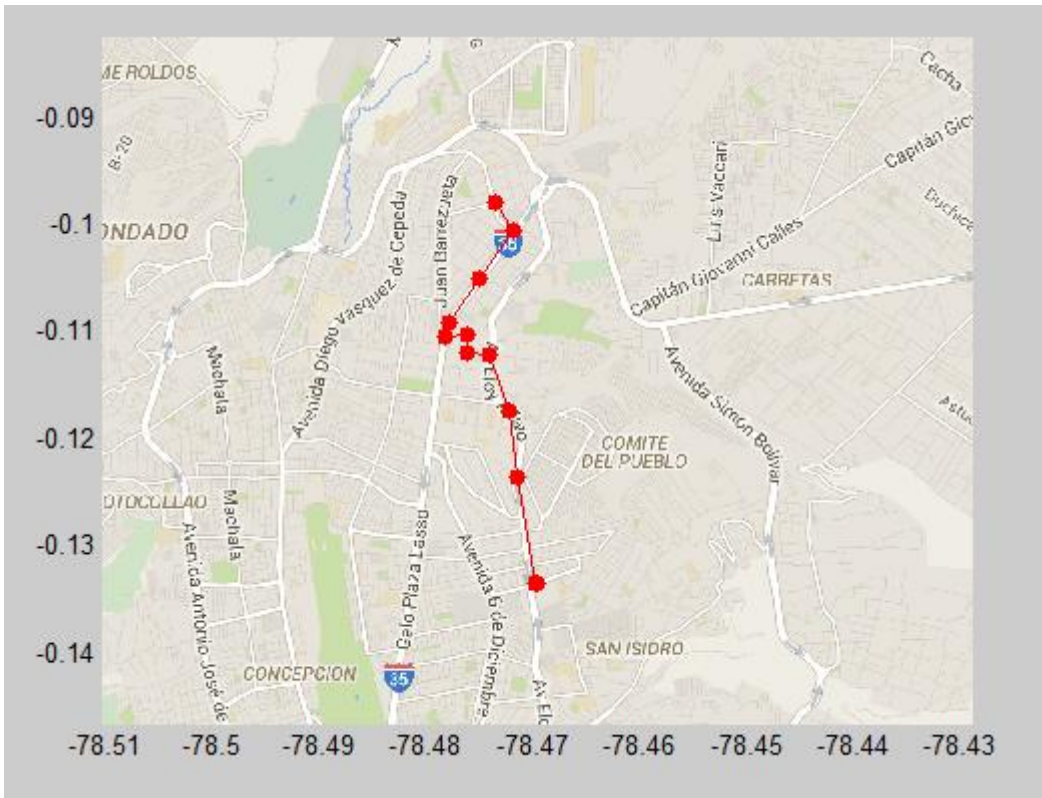
1312 33869



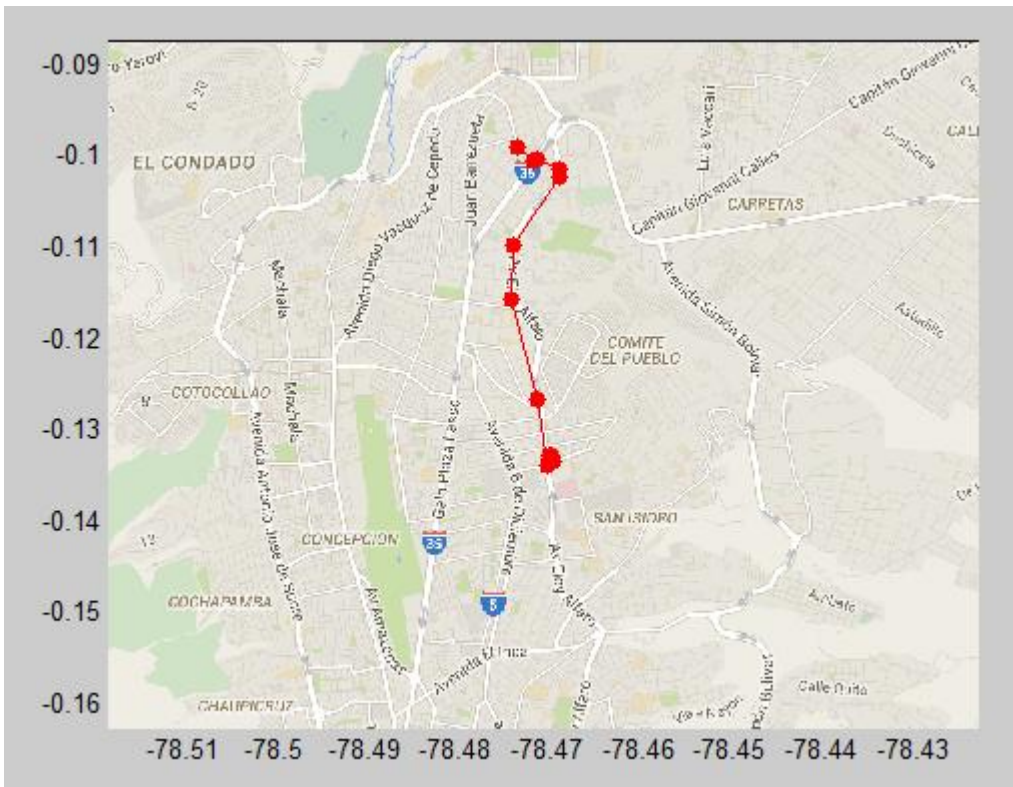
1450 340616



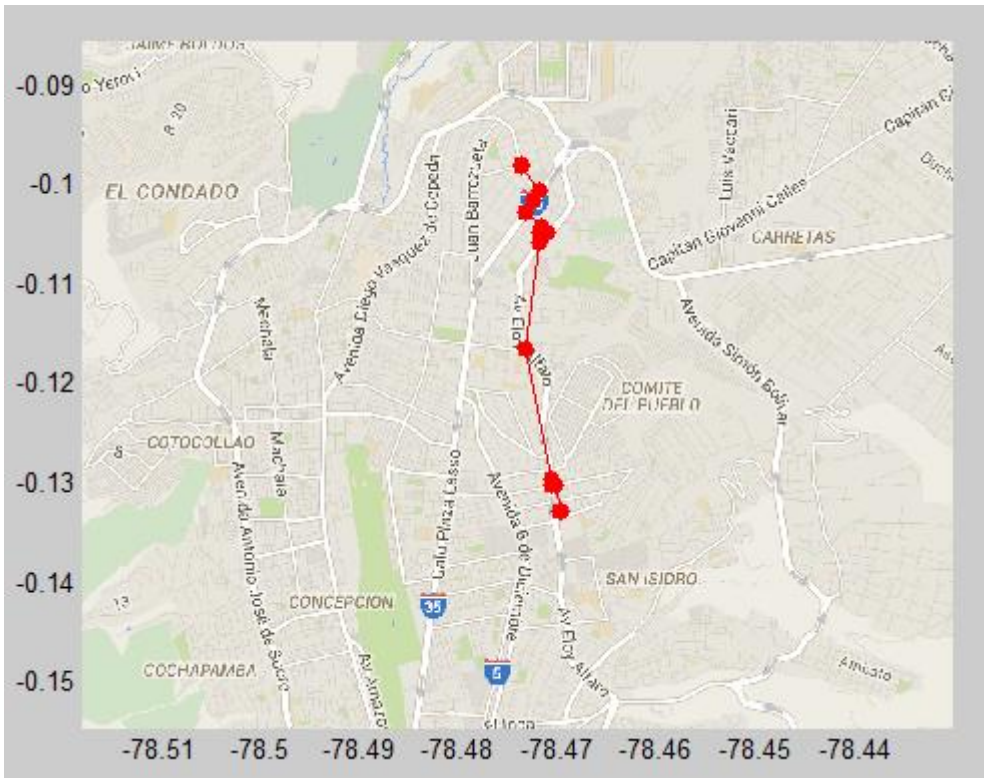
1225 33784



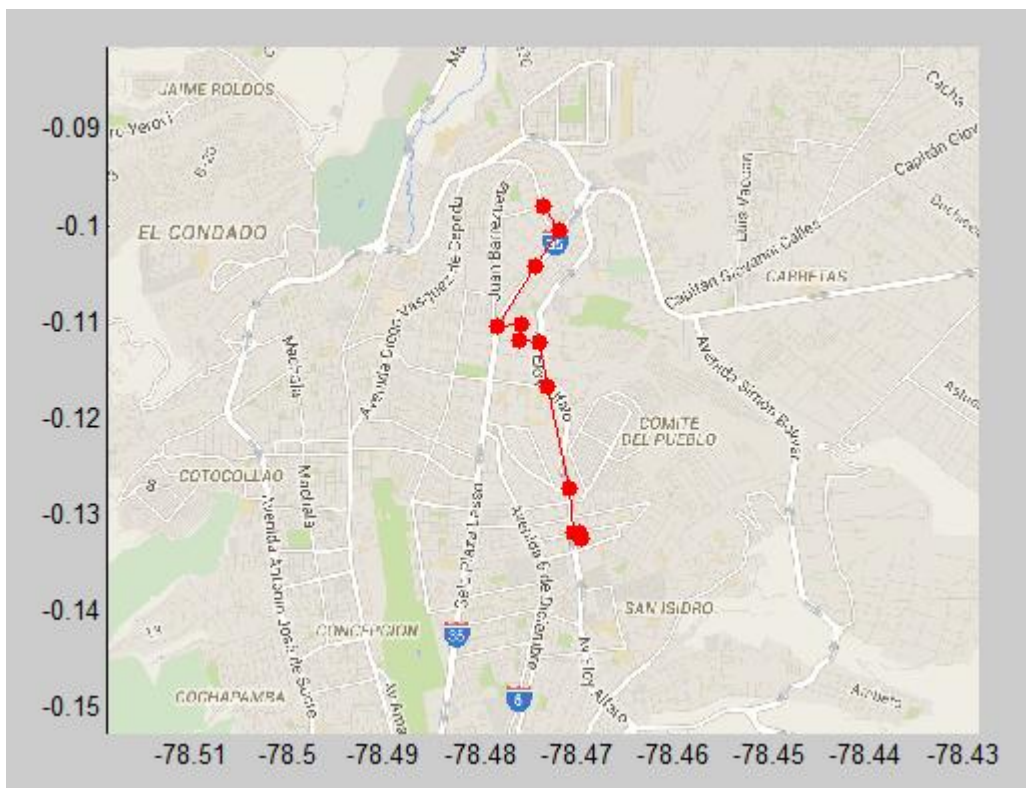
647 33062



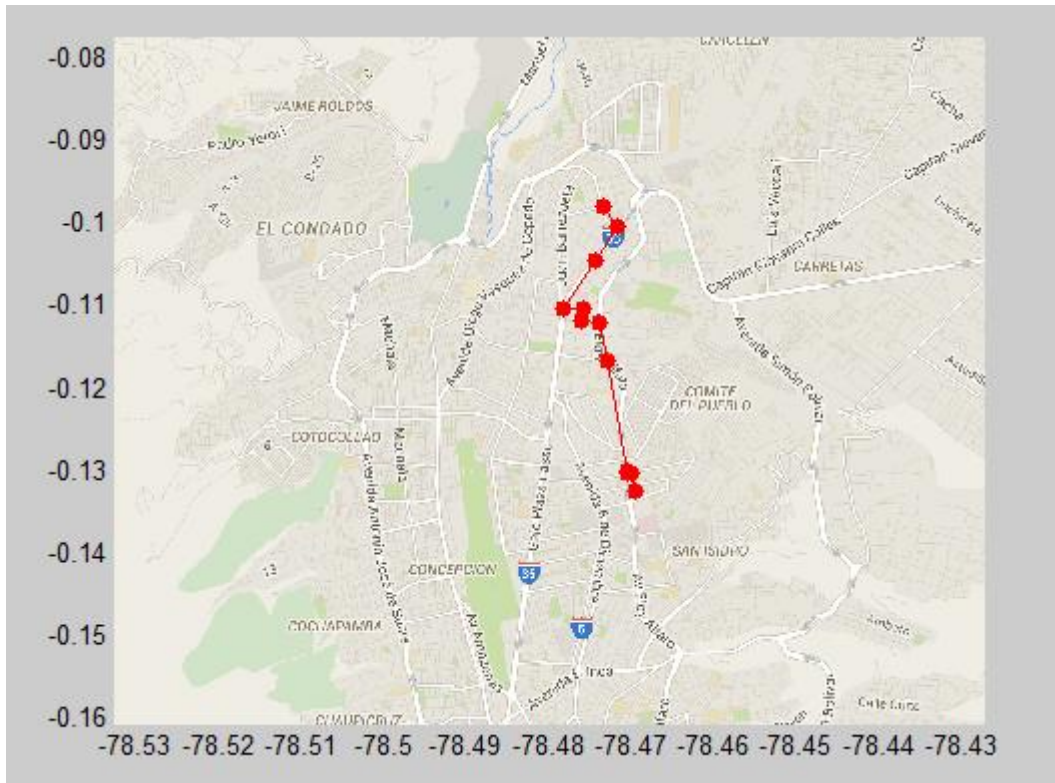
1224 33783



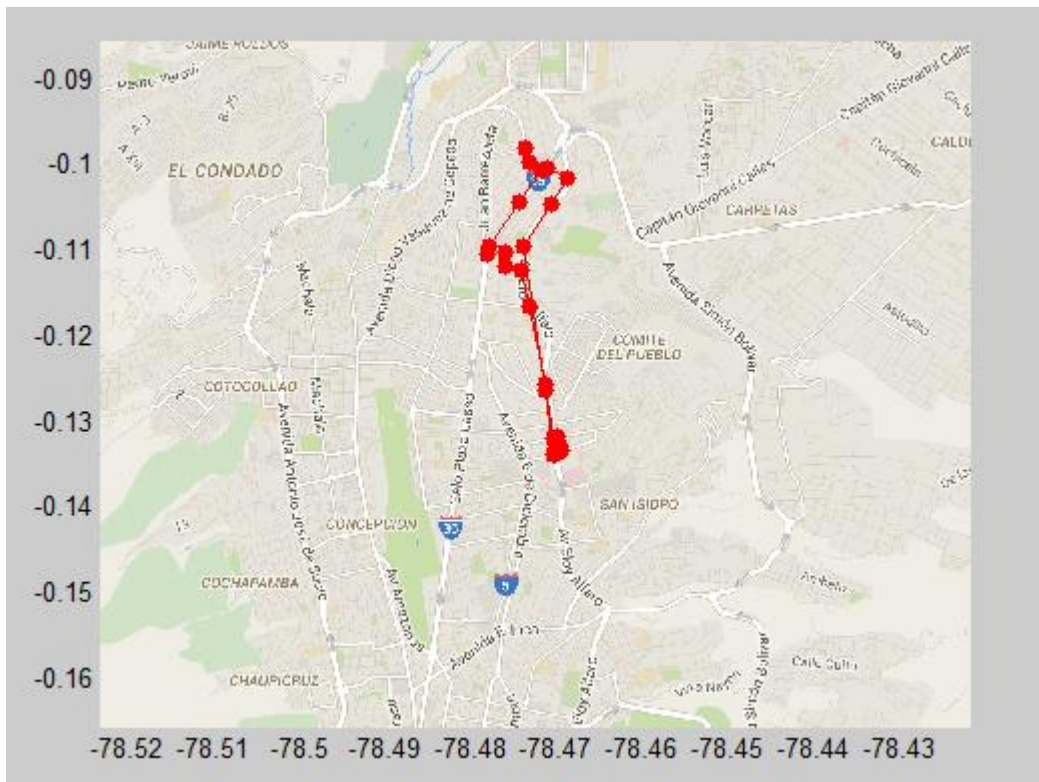
1583 351716



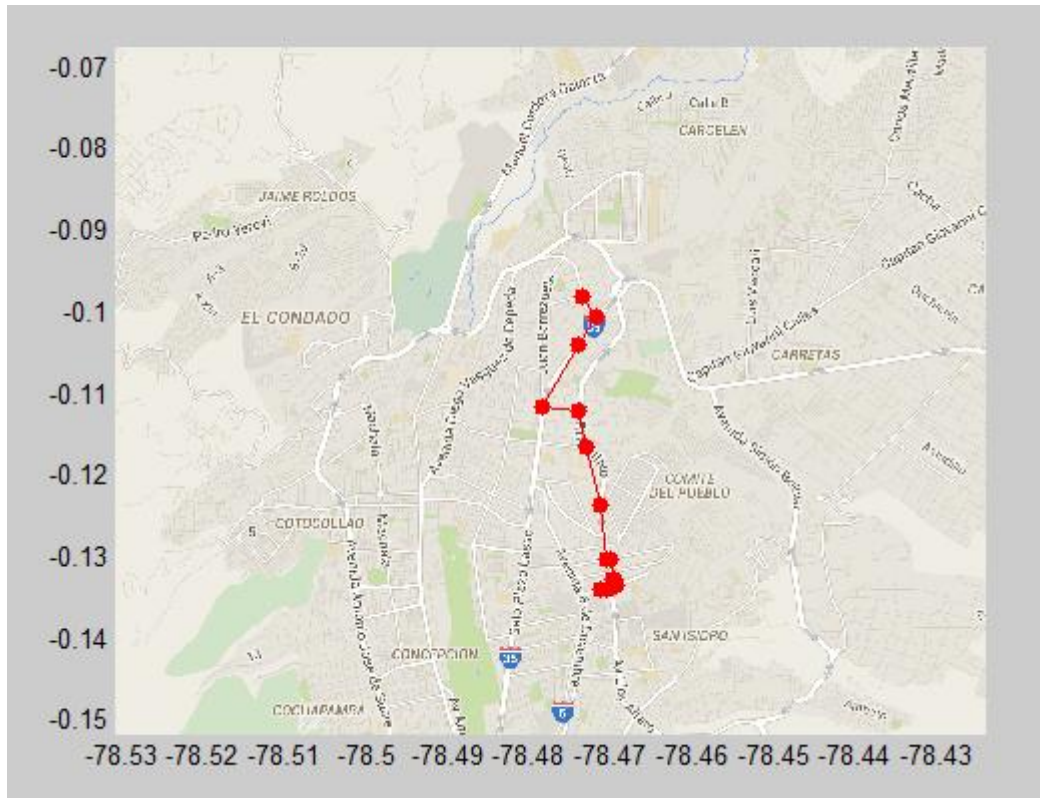
573 324714



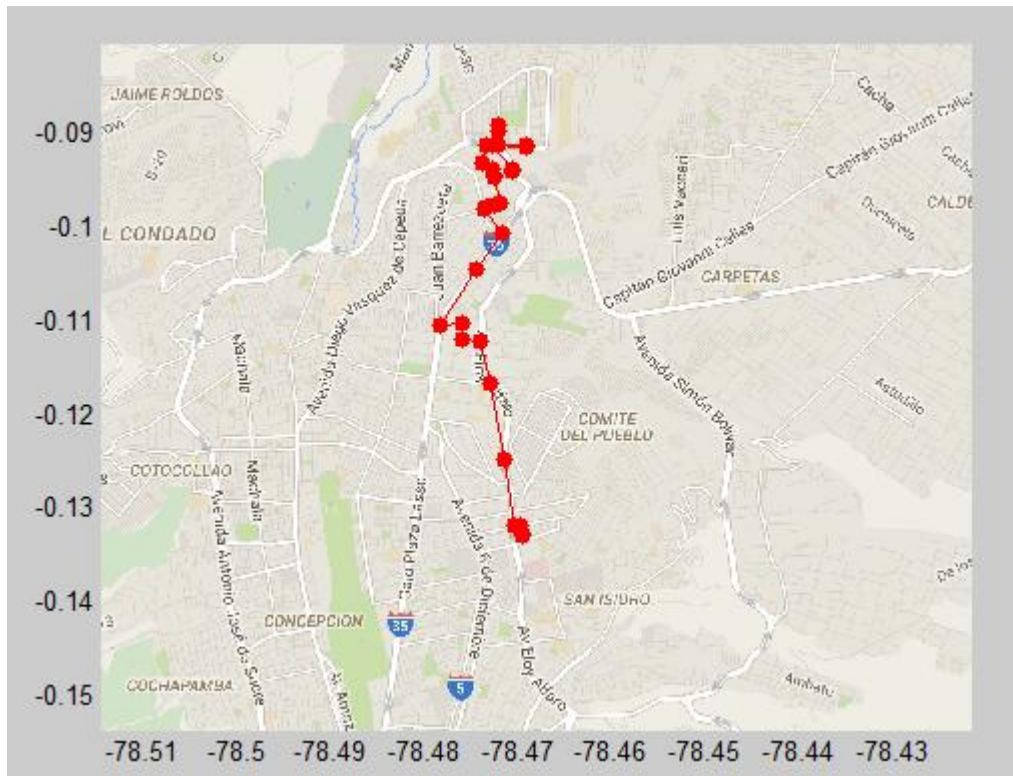
171 300721



1219 337728

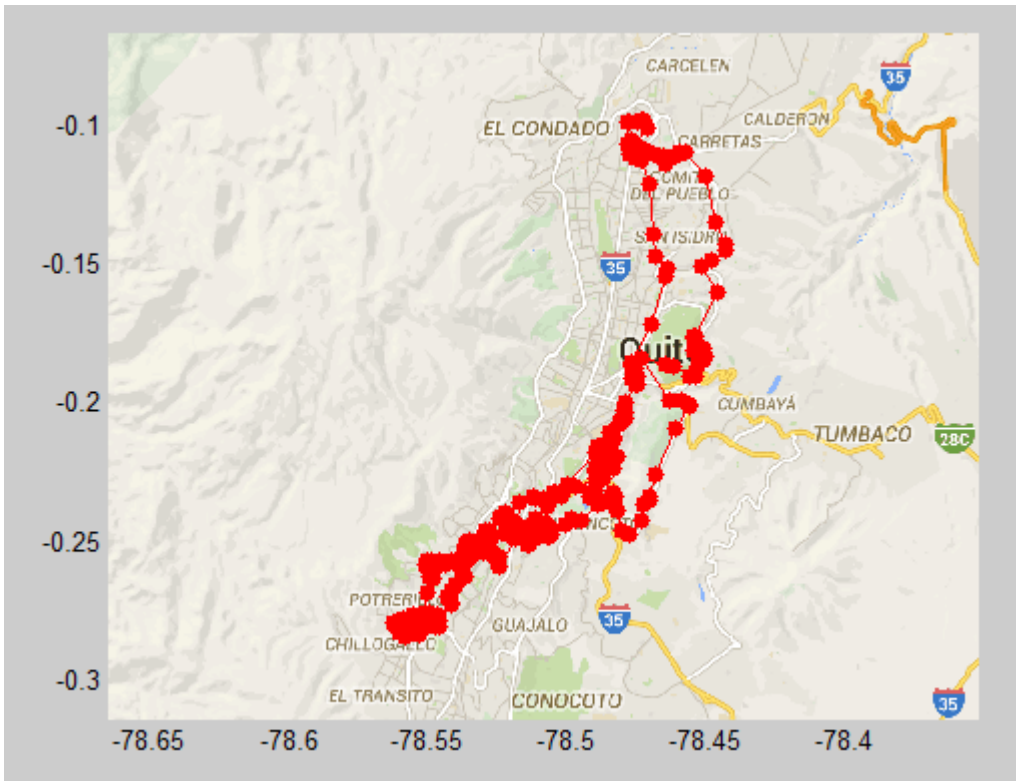


372 322317

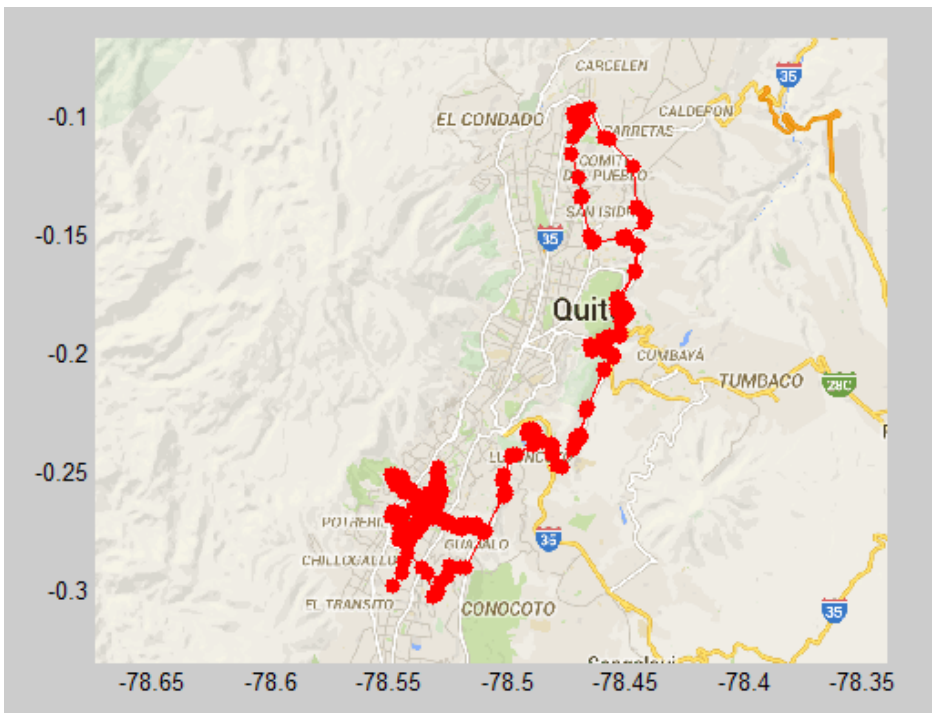


23 201813

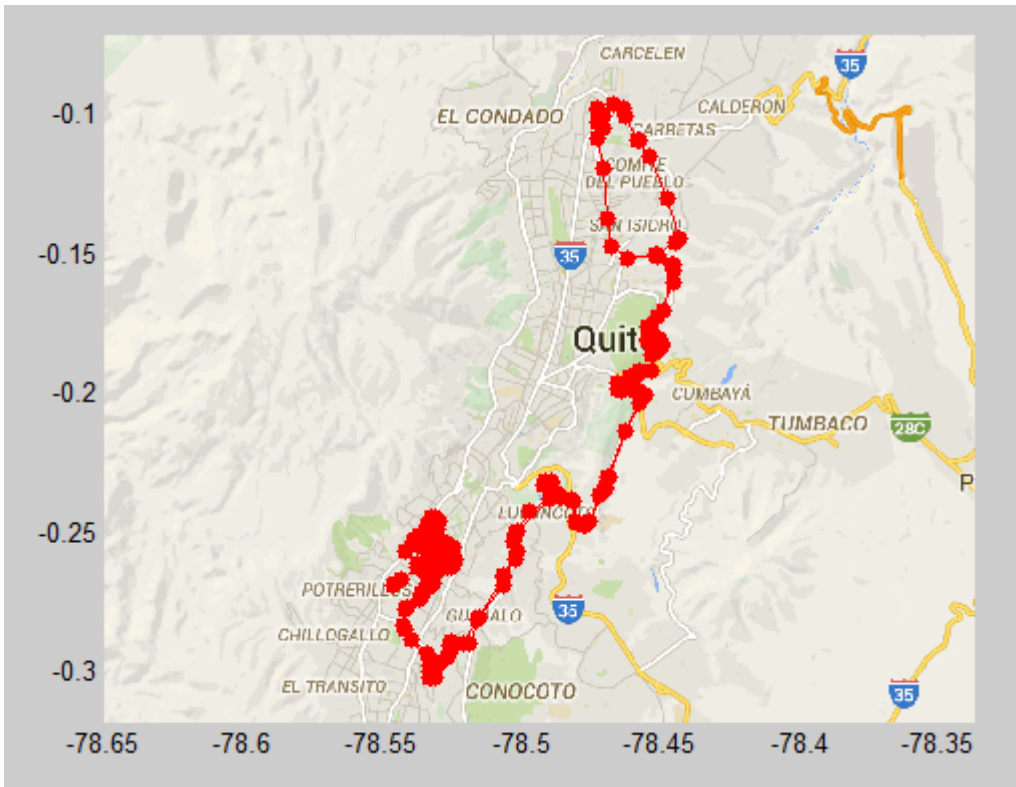
ANEXO 4



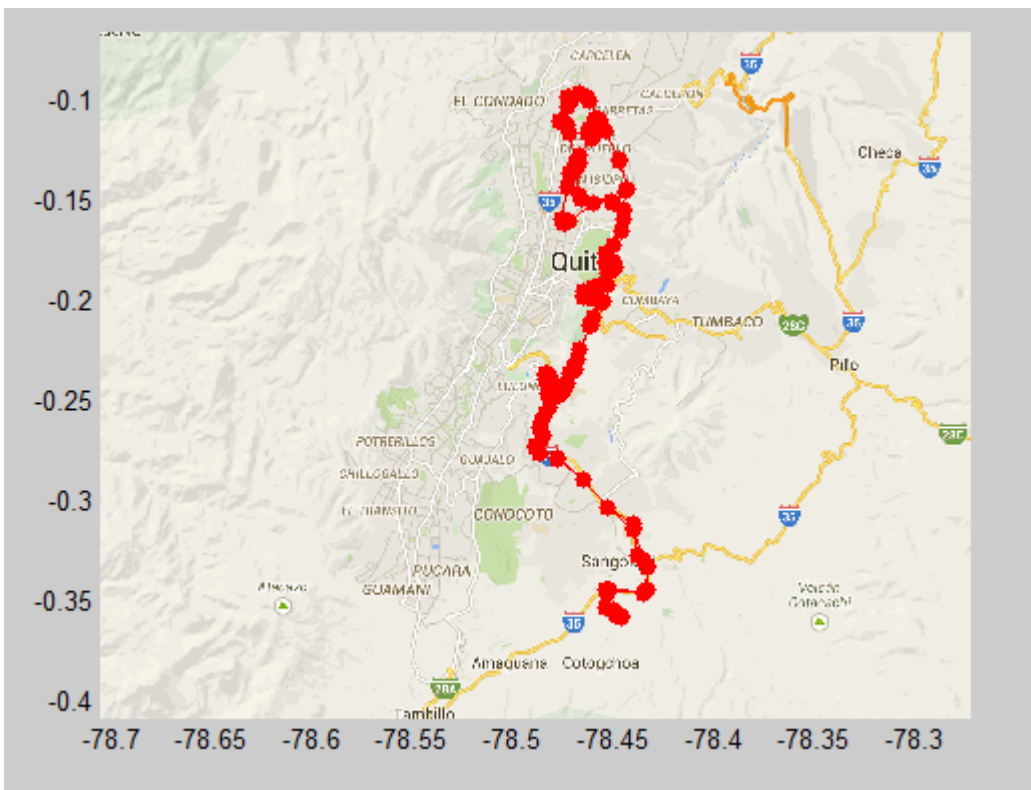
1584 351717



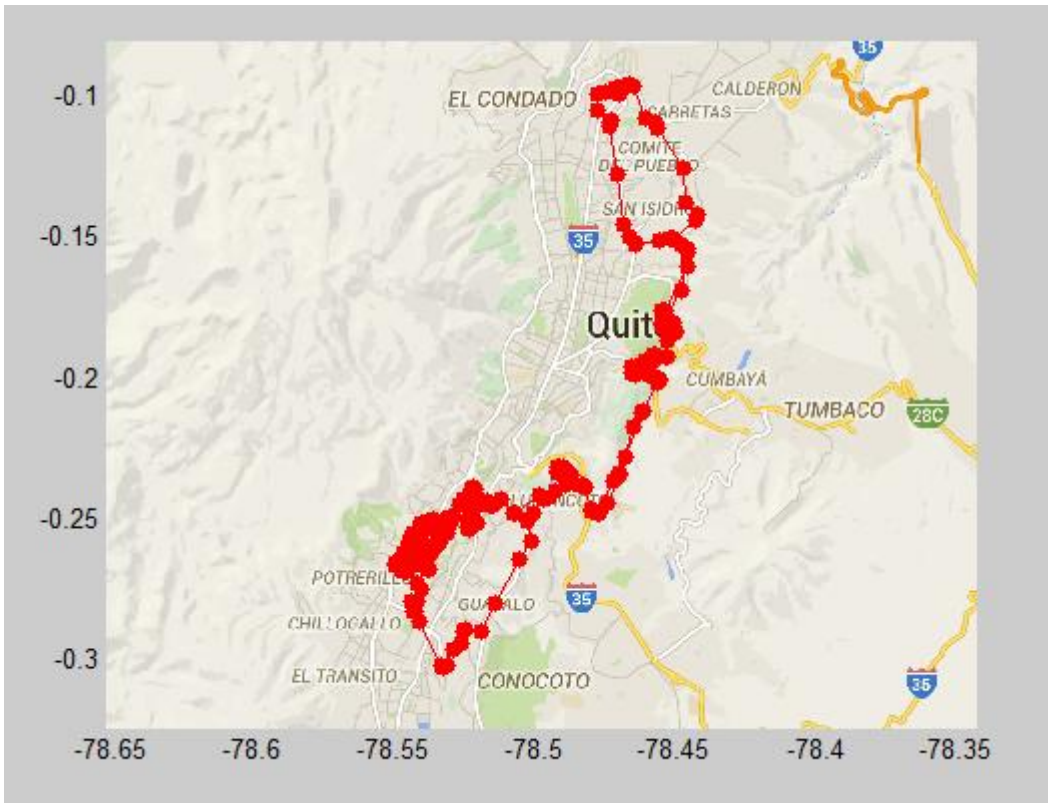
565 324628



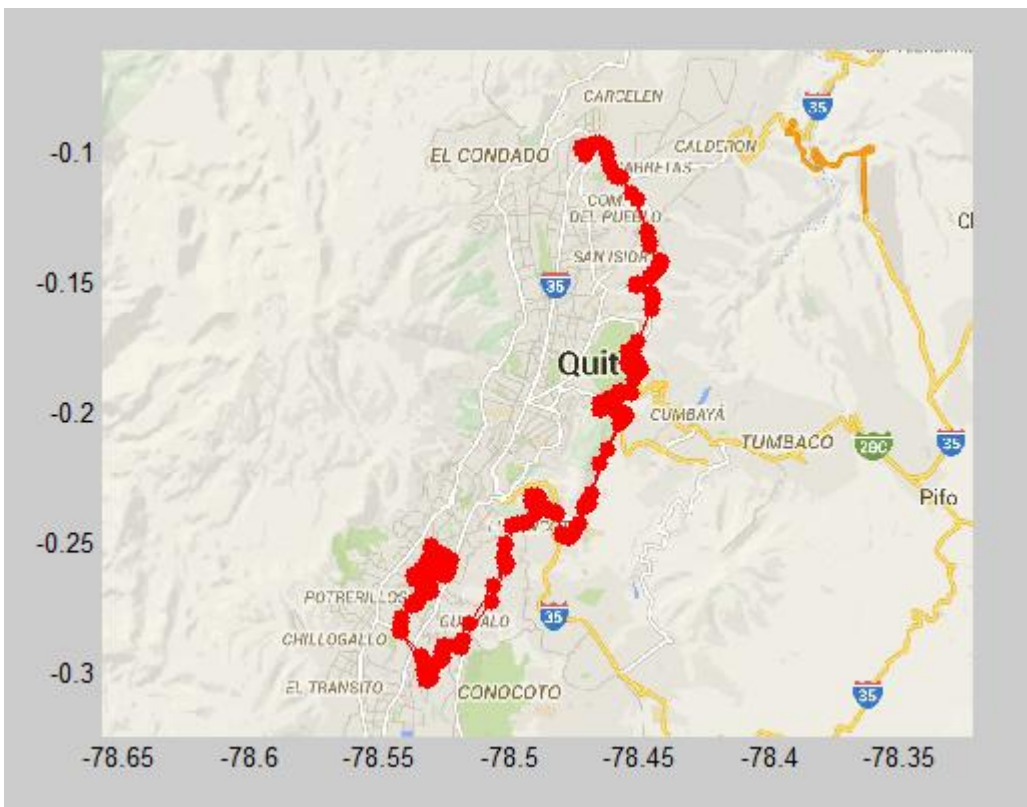
614 330419



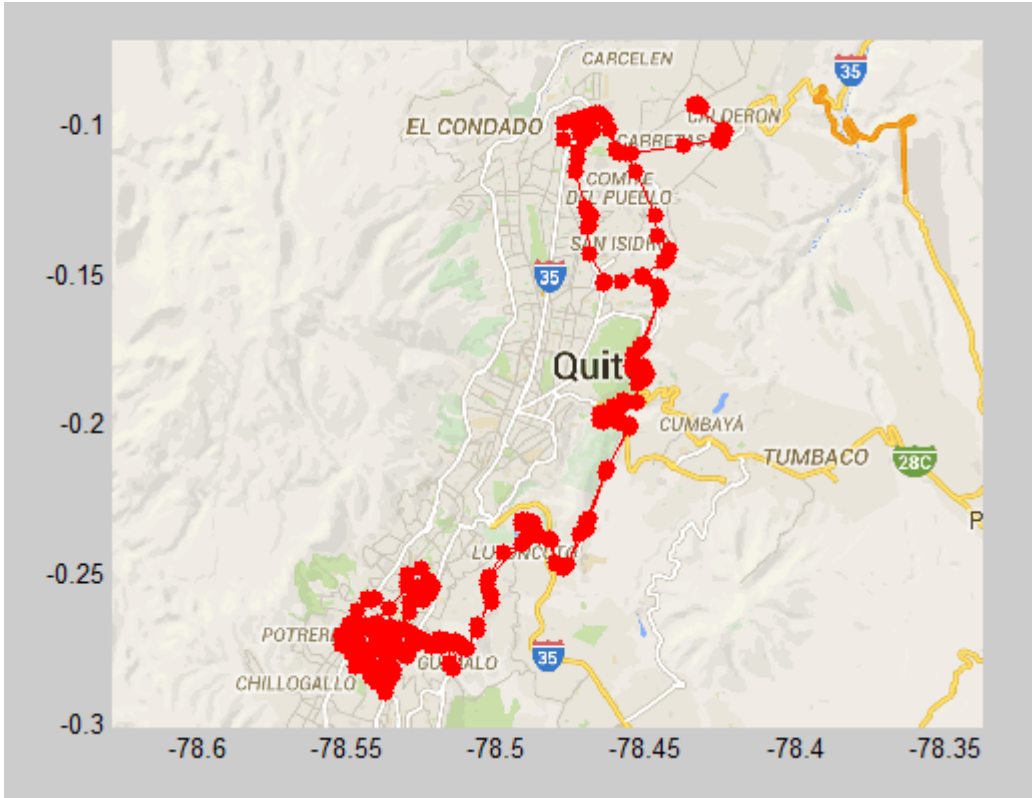
259 302611



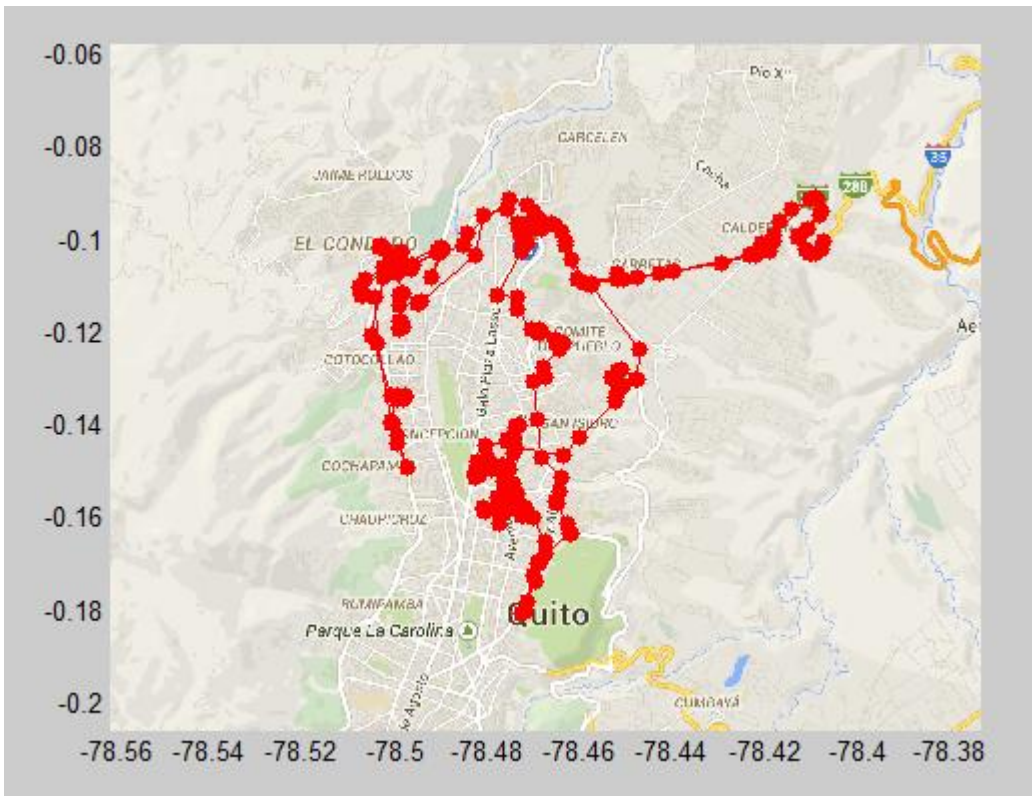
1358 338729



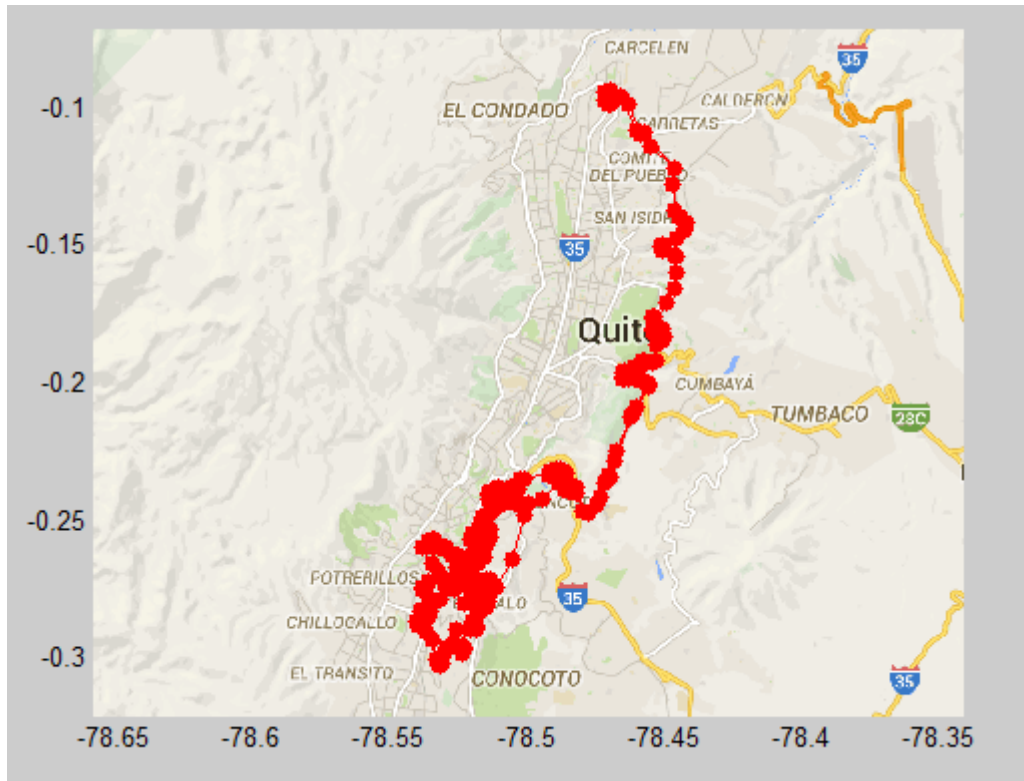
620 330426



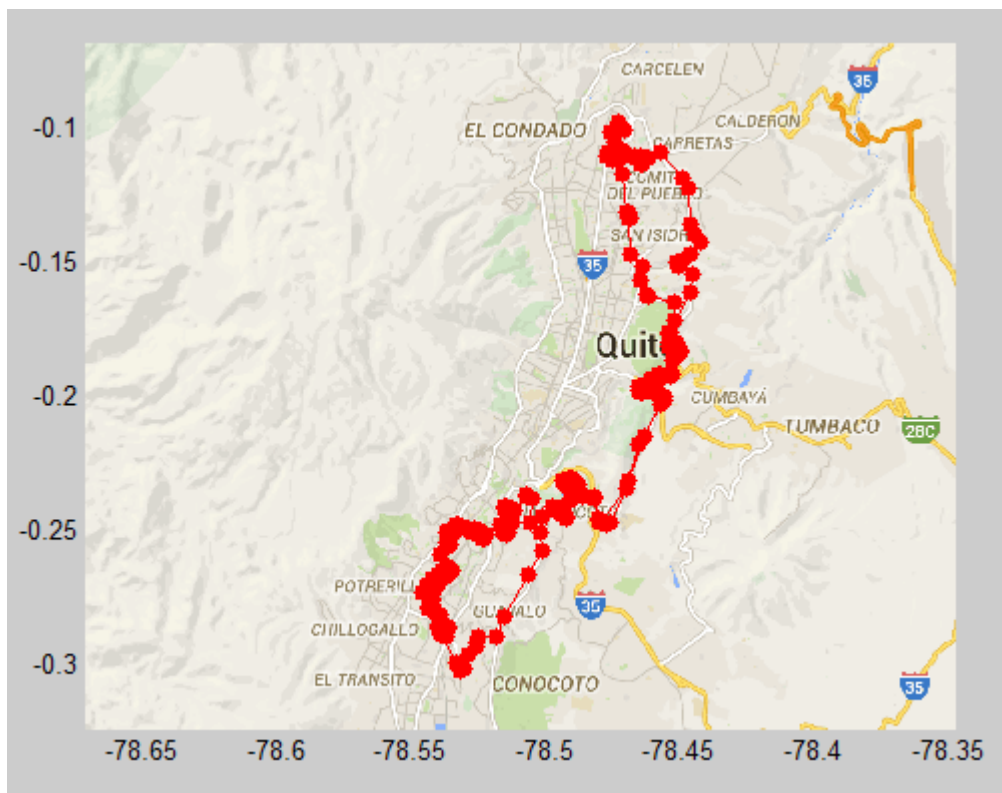
520 32453



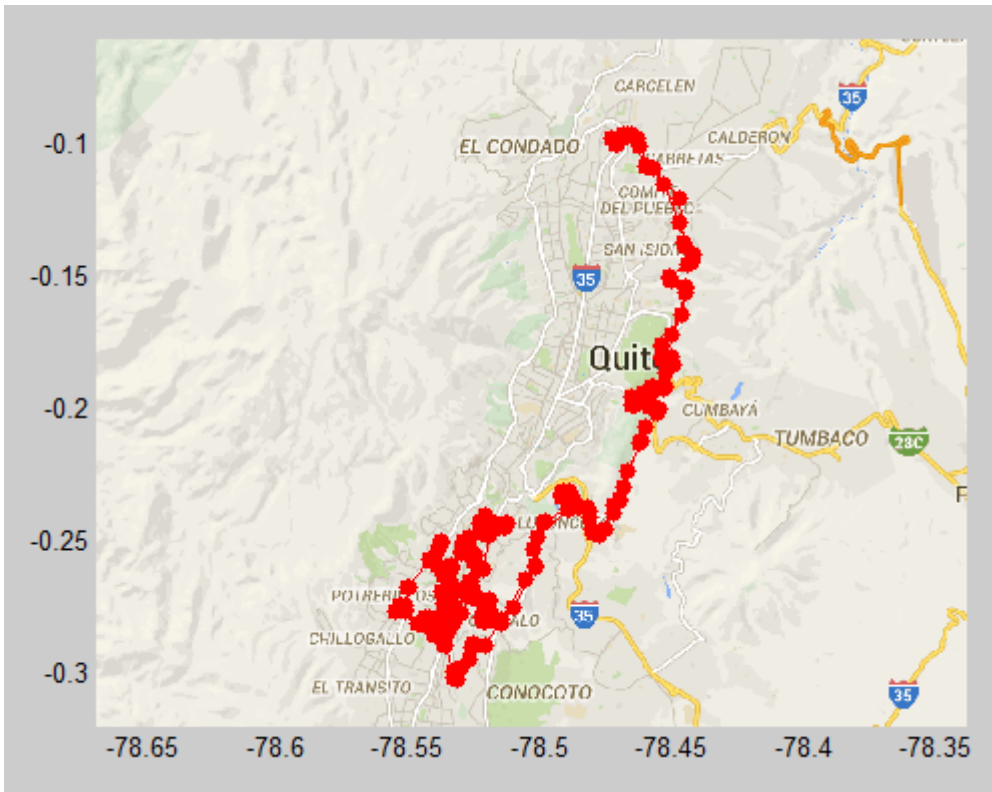
810 333719



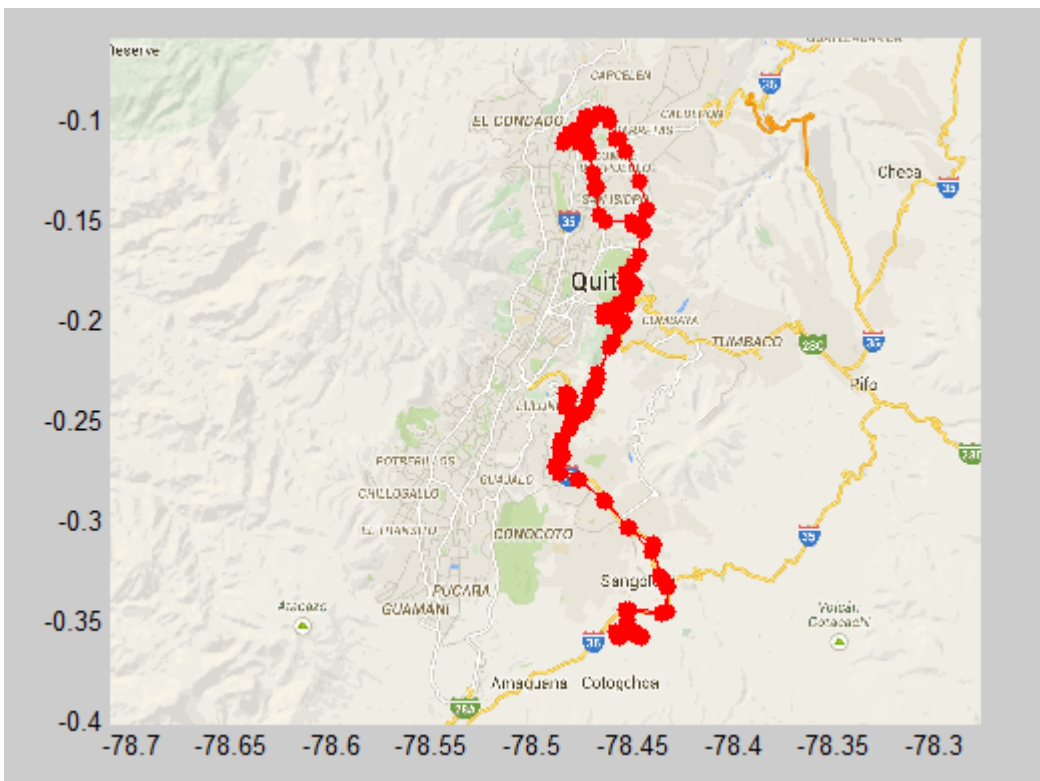
1109 337322



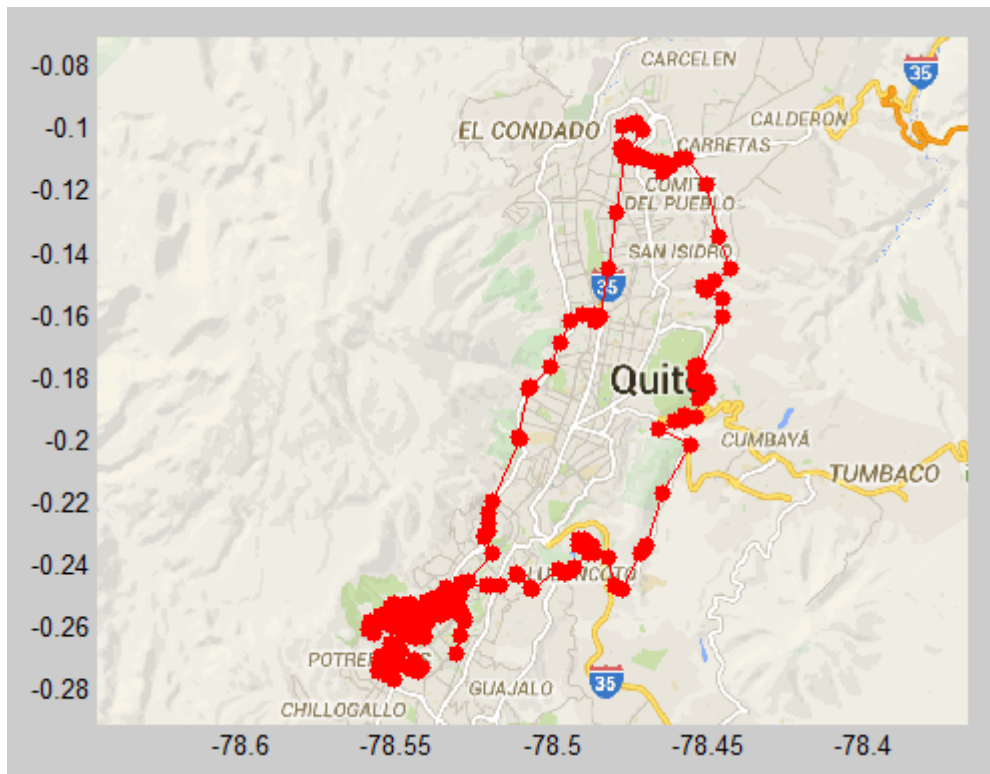
97 20865



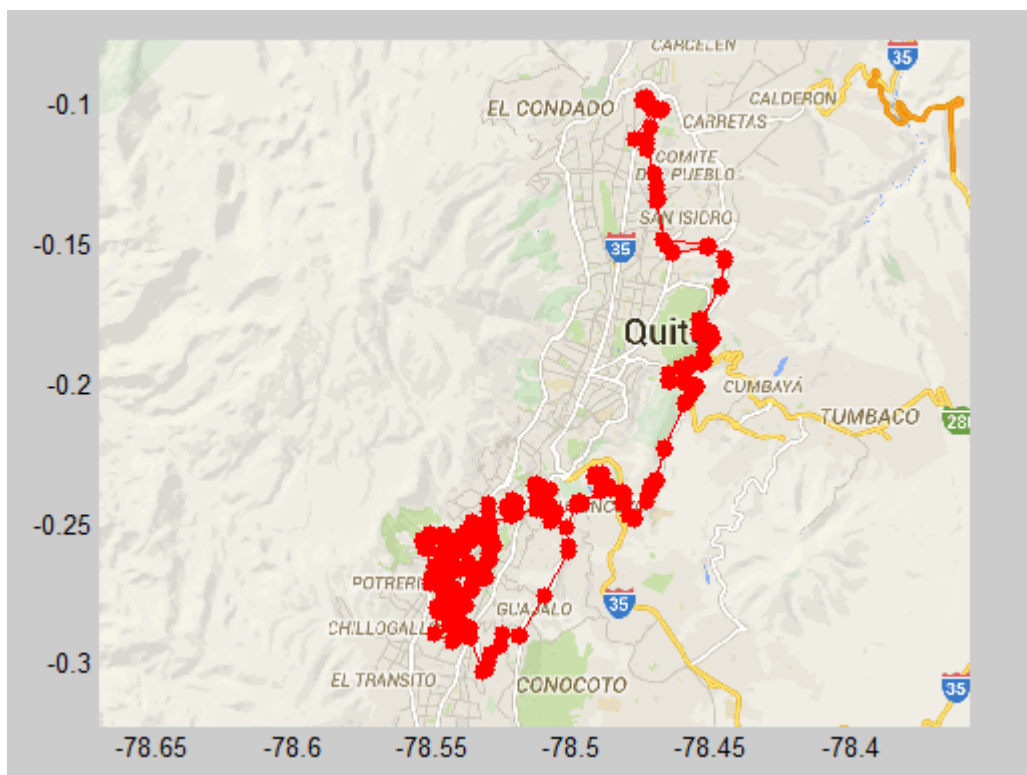
96 20864



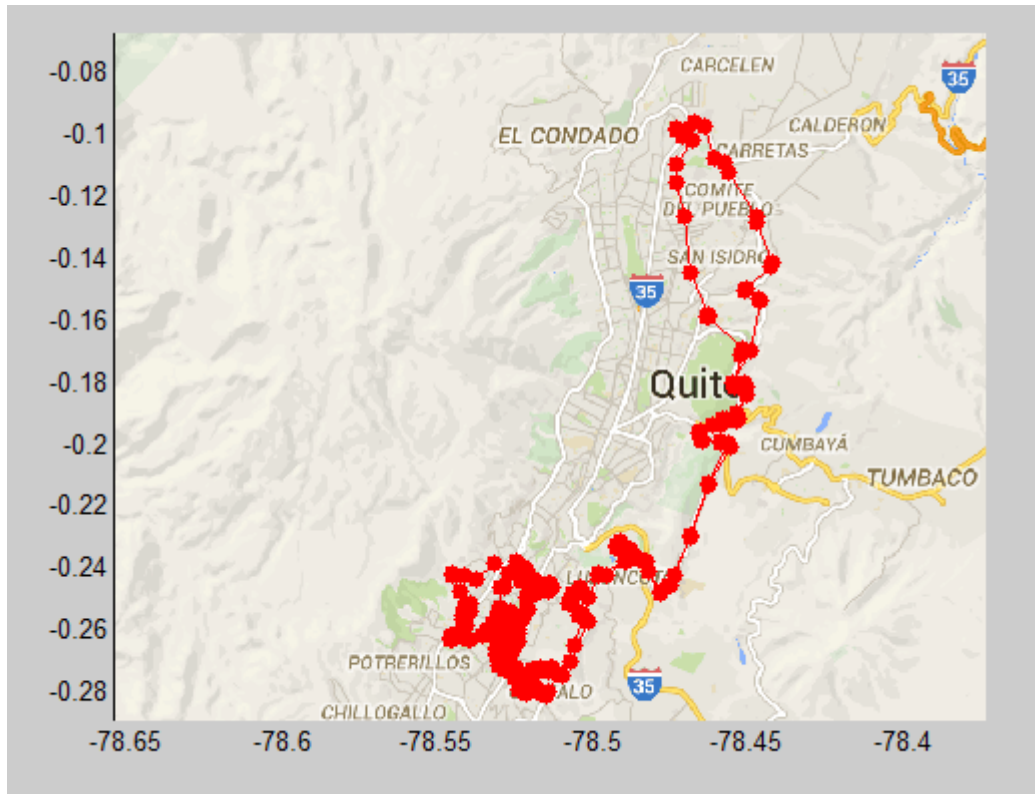
234 30258



114 208625

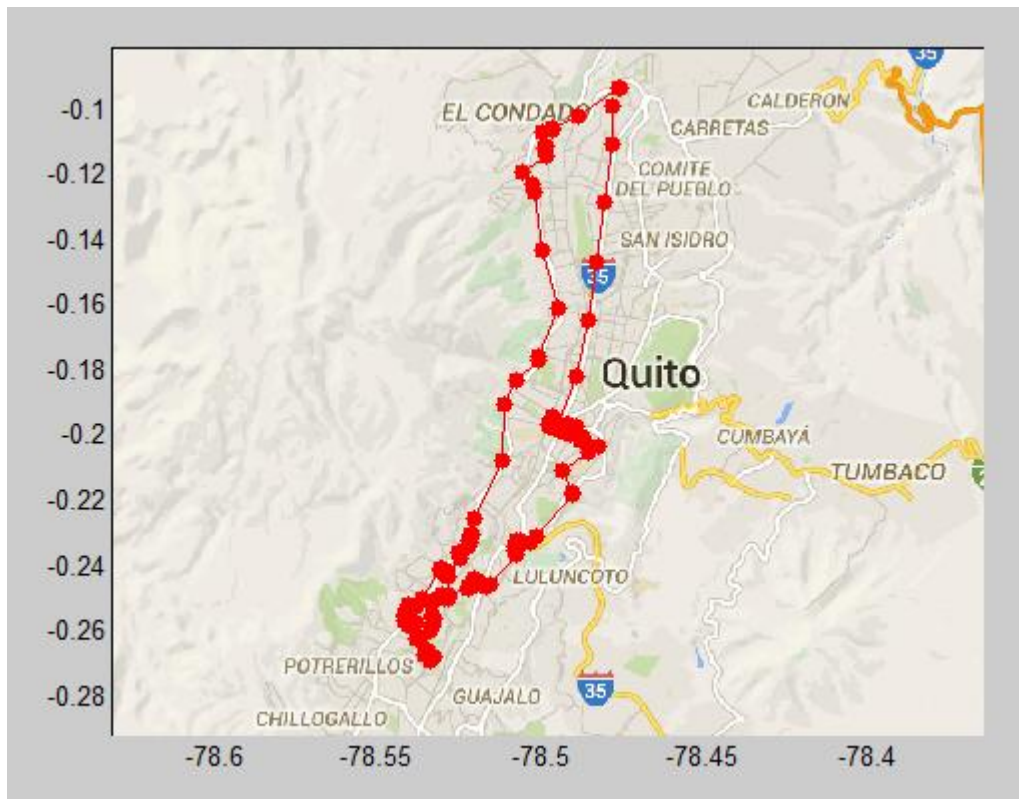


1093 336811



193 300823

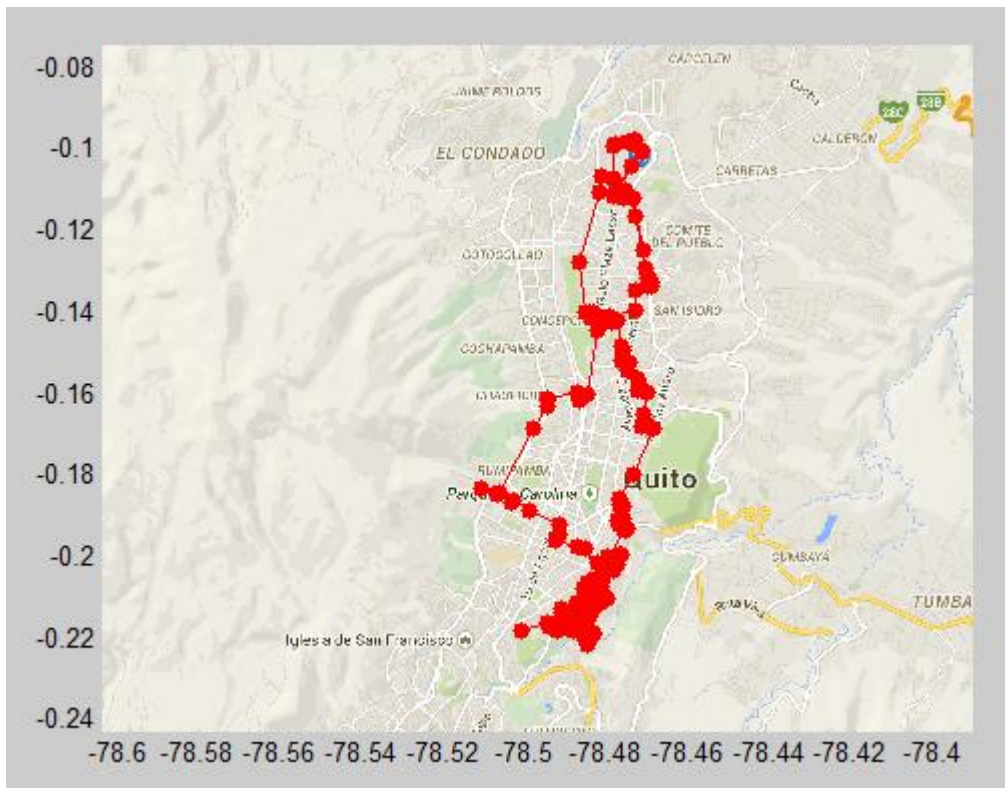
ANEXO 5



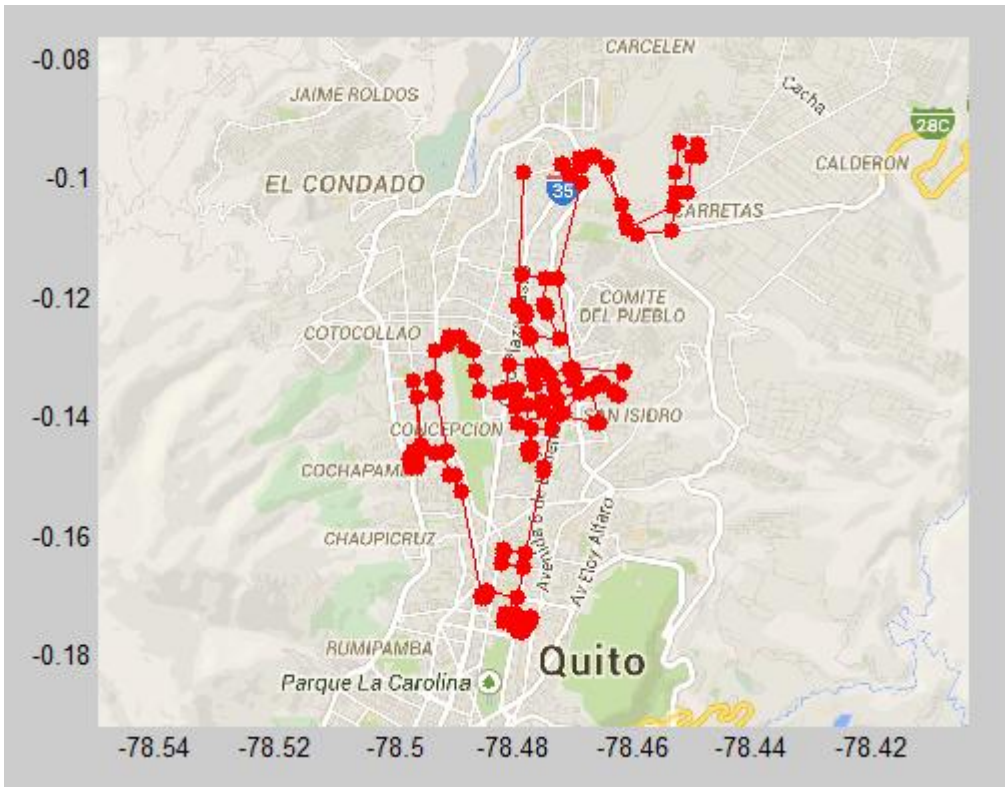
1234 337815



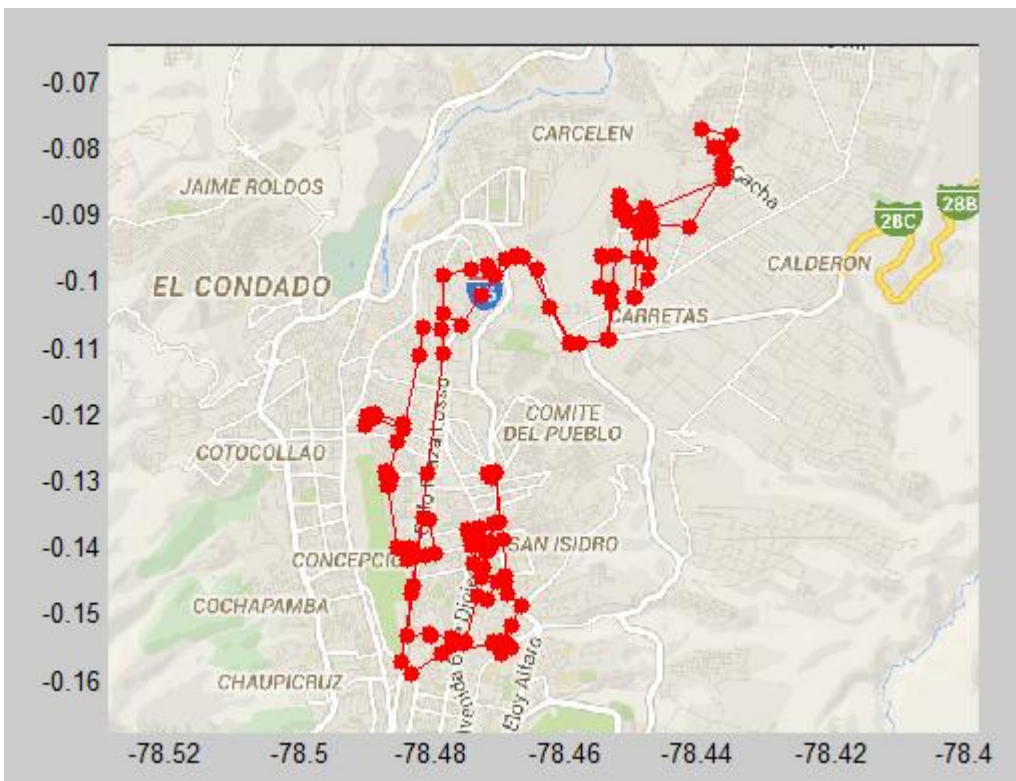
1029 336524



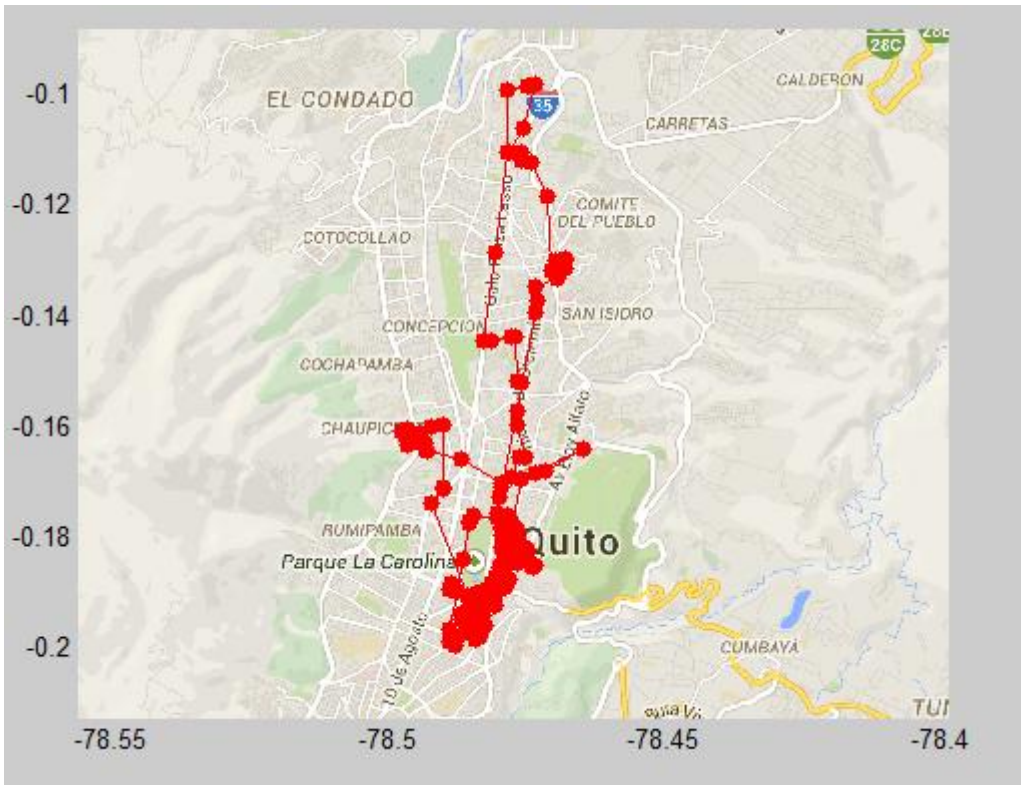
115 208626



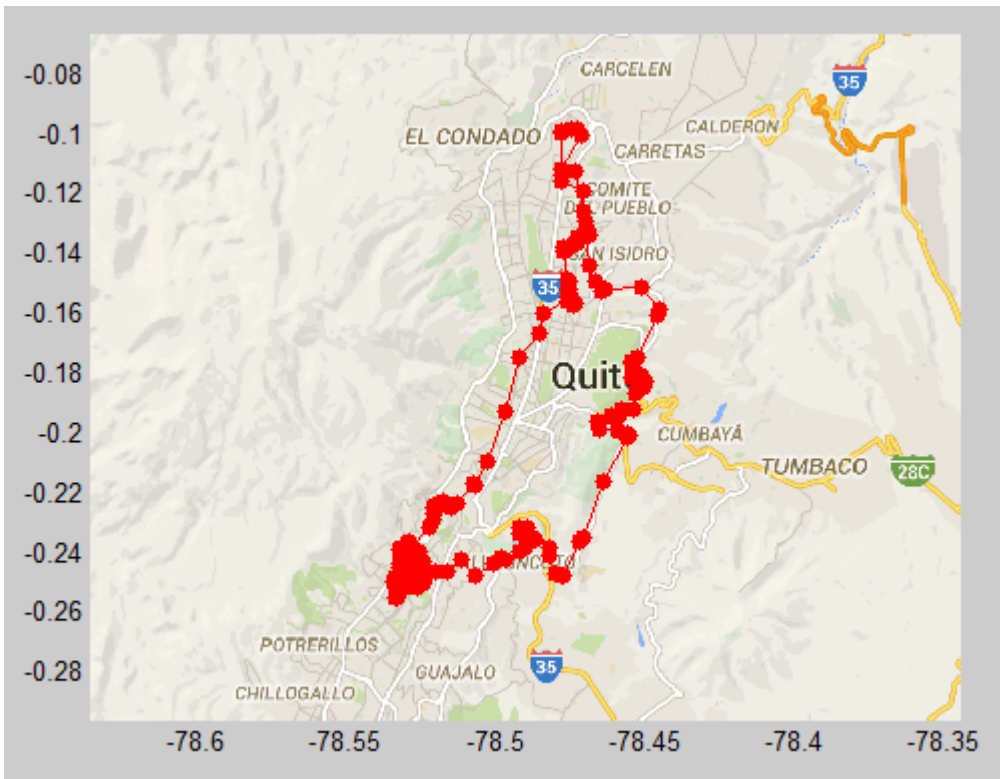
1041 33669



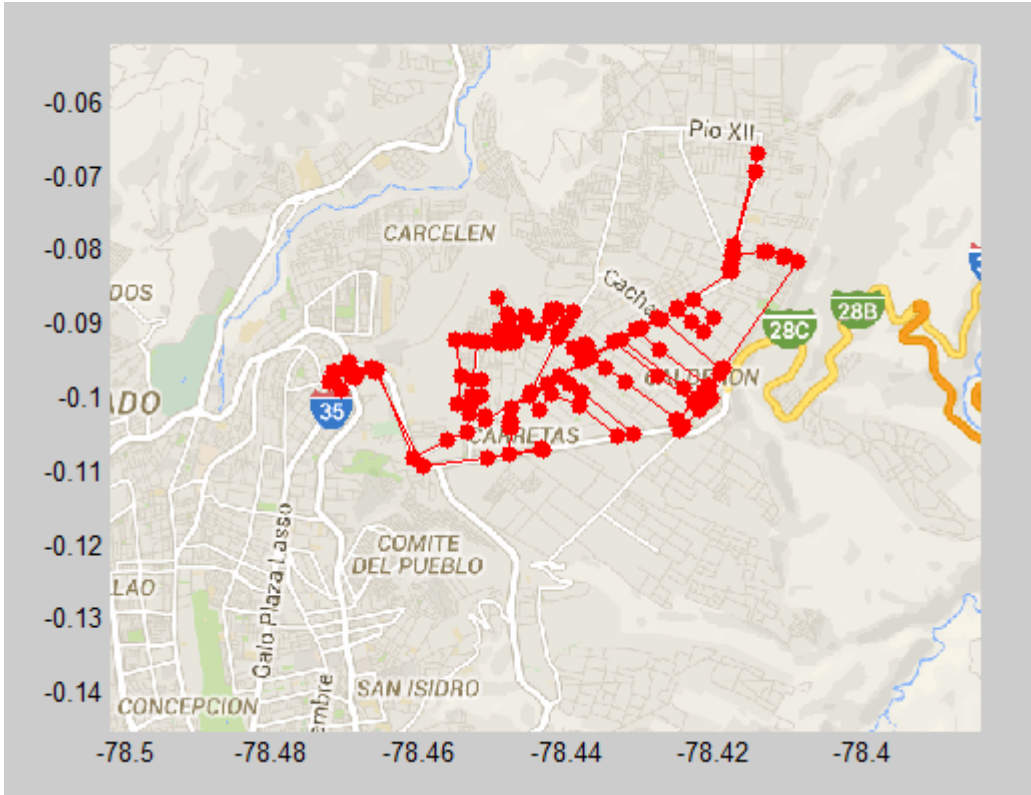
781 333614



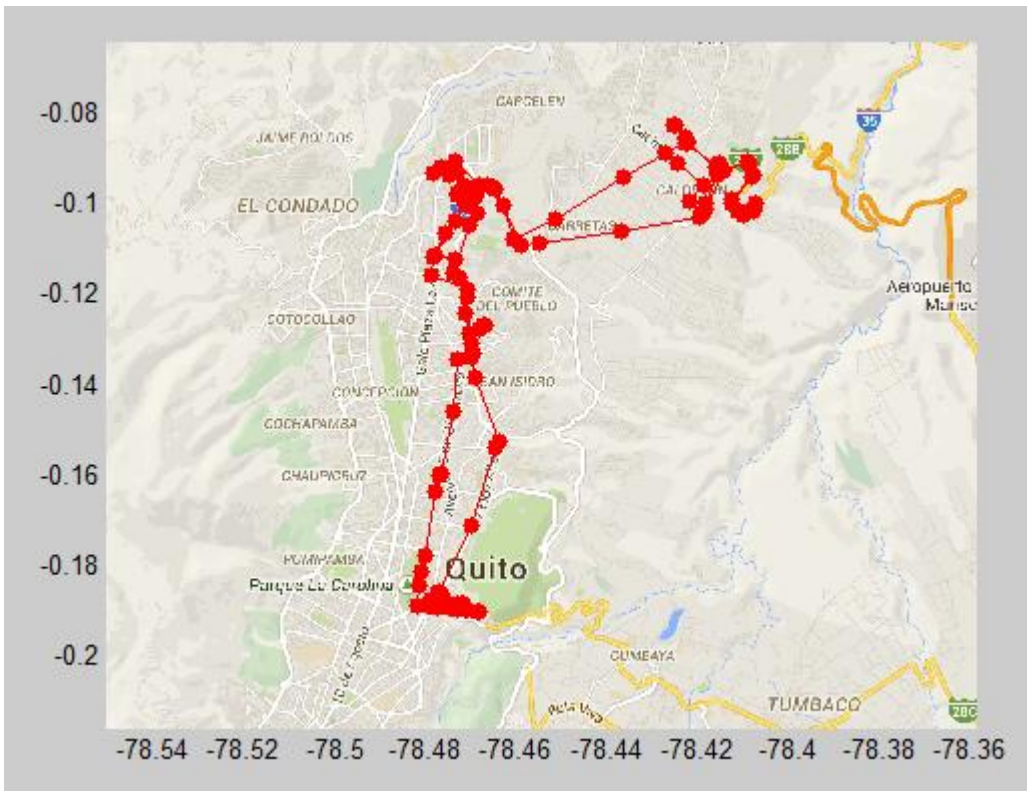
190 300819



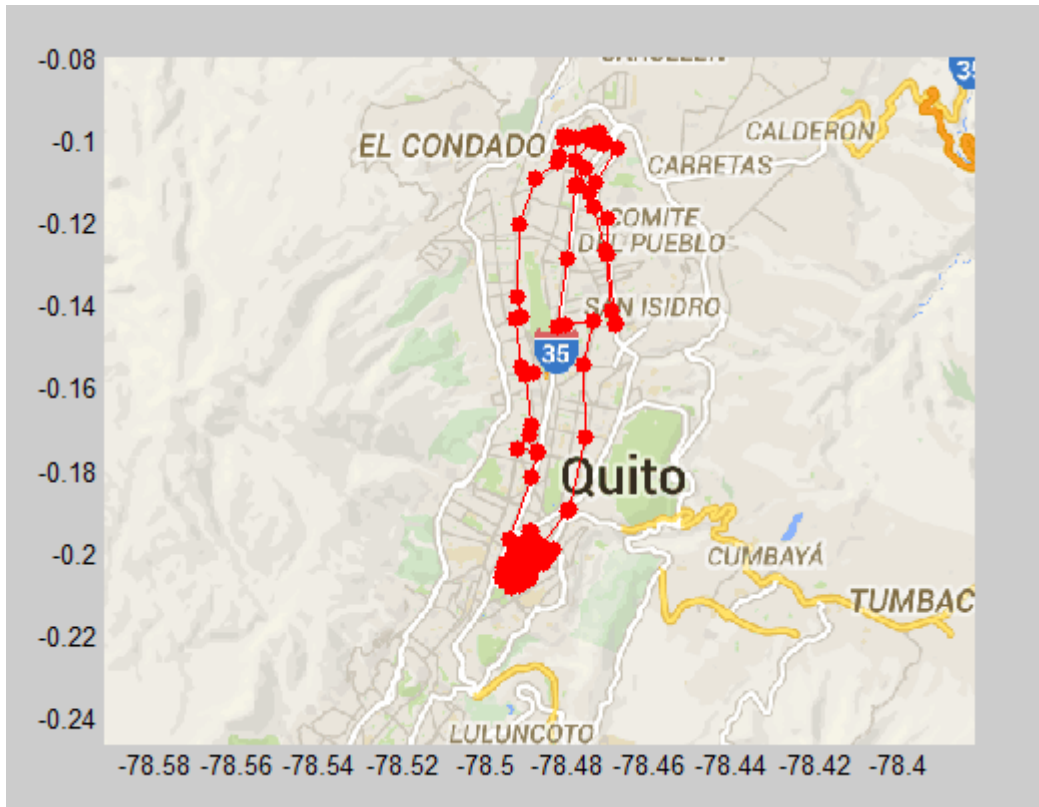
851 335329



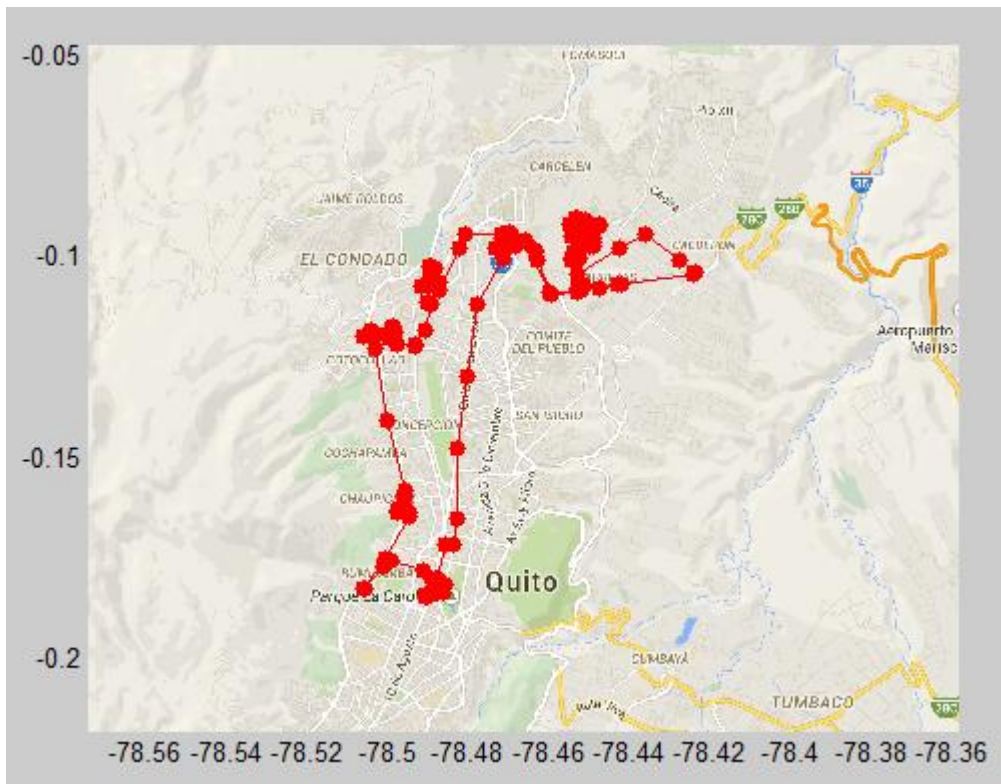
724 33347



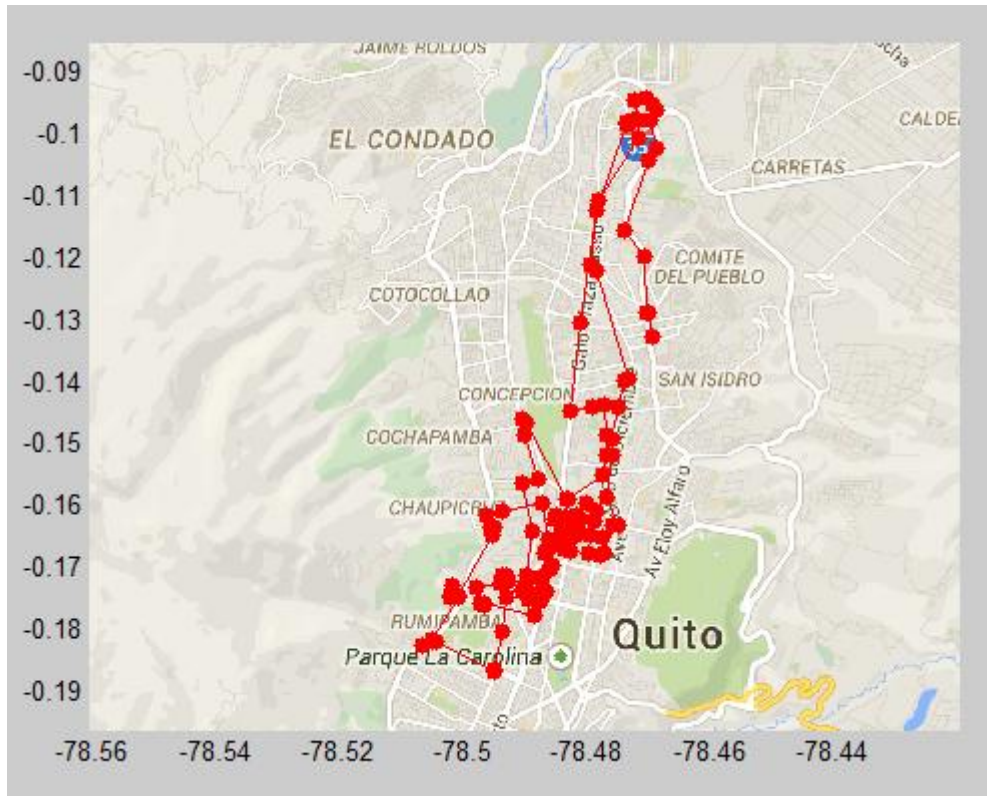
1021 336515



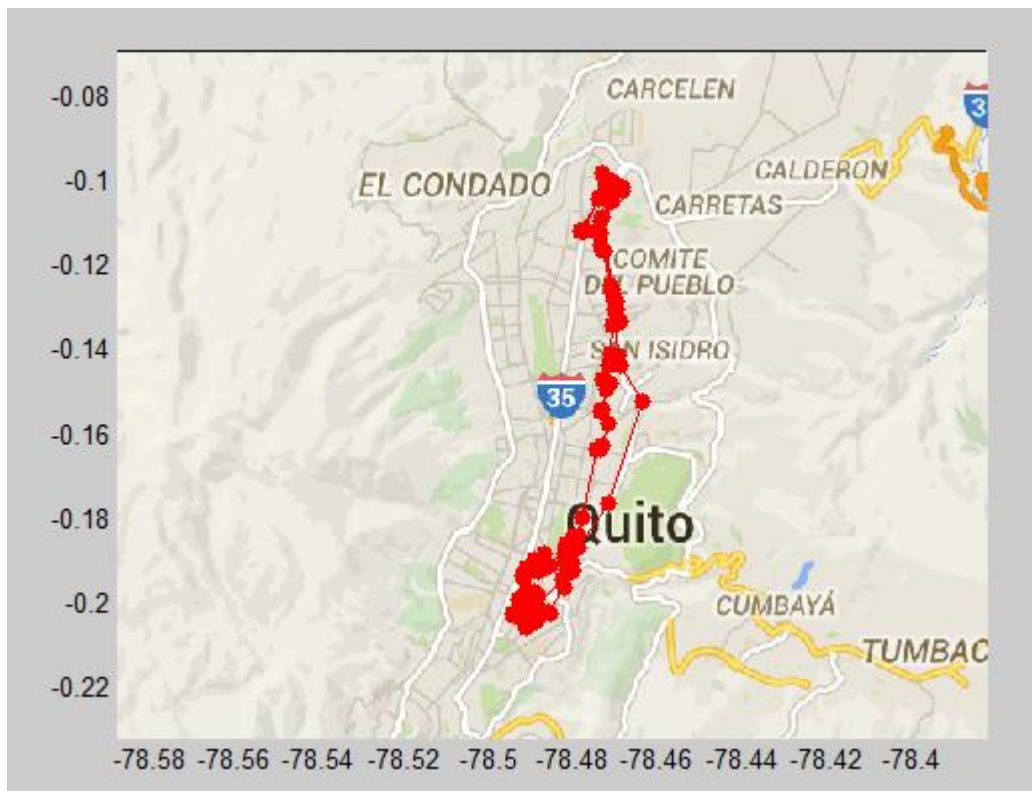
1575 35173



196 302320



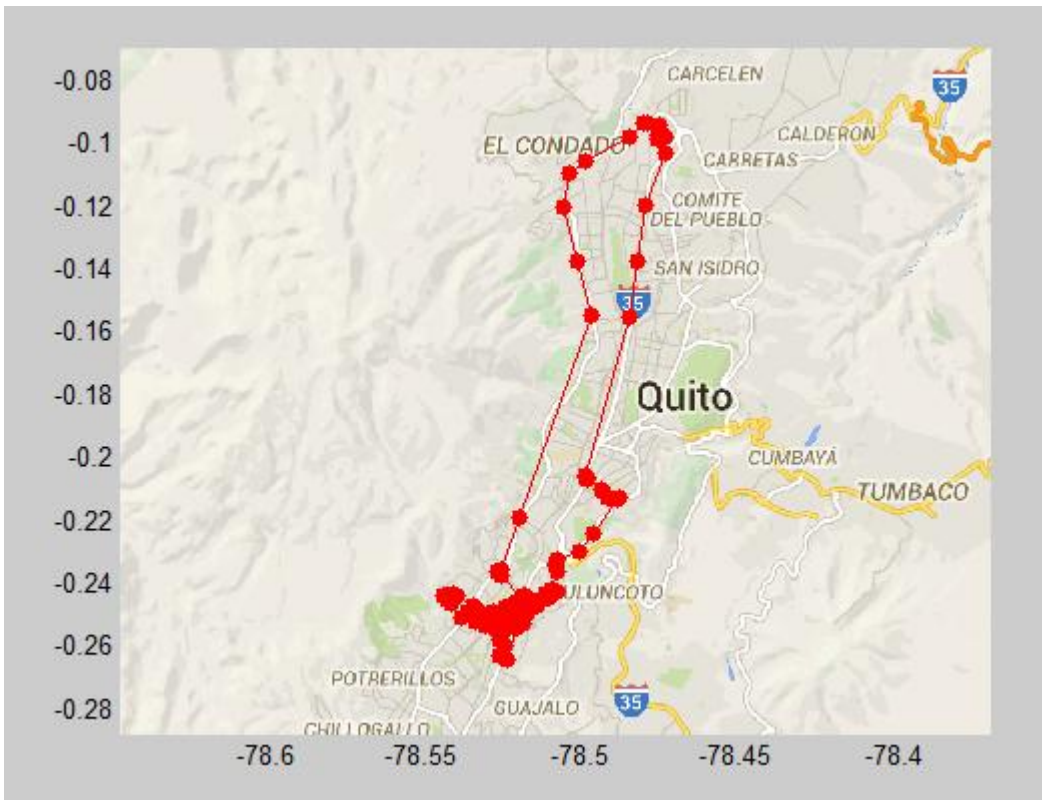
1515 351411



1497 351320



934 33575



1375 338820

