Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics

Evaluating the Effect of Digital Elevation Model Resolution on the Accuracy of Large Scale Road Grade Information

Geoffrey Natin
14318196
Supervisor: Dr Mike Brady

A Dissertation submitted in partial fulfilment
of the requirements for the degree of
Master in Computer Science

Submitted to the University of Dublin, Trinity College
April 11, 2019

Declaration

I, Geoffrey Natin, declare that the following dissertation, except where otherwise stated, is entirely my own work; that it has not previously been submitted as an exercise for a degree, either in Trinity College Dublin, or in any other University; and that the library may lend or copy it or any part thereof on request.

Signed: ______________________ Date: ______________________

Abstract

Road Grade is the slope of roads. Road grade is an important factor in eco-routing. Taking road grade into account when choosing a route can save up to 20% on fuel. Road grade profiles are generated using GPS points and elevation data from digital elevation models (DEMs). The resolution of the DEMs used dictates the accuracy of the generated road grade information. DEMs with a precision of up to 30m^2 are freely available[1]. More precise DEMs can be purchased from companies but are very expensive. Analysis of how much the resolution of DEMs effects the accuracy of the road grade information must be done to give insight into how accurate road grade data is when generated from DEMs of a certain resolution. This analysis will also give an idea of what resolution DEM should be used, or bought, to create road grade data of a certain accuracy. A 2m^2 DEM was downsampled to 6m^2, 10m^2, 20m^2, and 25m^2 DEMs. GPS traces collected from buses in the city of Dublin, Ireland, were combined to generate route profiles for bus routes. Road grade profiles were generated for each bus route using the DEMs of varying resolutions. The road grade profiles' accuracy was found to increase linearly as the DEMs used got more precise. The increase in accuracy afforded by using a higher resolution DEMs was found to be insignificant, with only a decrease in mean average error of less than 0.1 degrees from 26m^2 to 10m^2. The results indicate that using freely available DEMs is likely acceptable for most eco-routing solutions.

Acknowledgements

I stand humbly on the shoulders of giants. I would not have any of the knowledge necessary to complete this work were it not for the efforts of many previous researchers. Most of all were it not for the works by Jared Magrath, Tushti Singla, Mike Brady, and Aashish Arora who deserve much of the credit for any work I do in this field.

A special thanks to Mike Brady, who made this project possible, for all of his help, advice, and work from its conception to completion.

Also thanks to my close friend Luke Gilligan for his support in helping me understand concepts and tackle problems unrelated to my usual field of study throughout this project's journey.

Contents

List of Figures

Abbreviations

API	Application Programming Interface
DEM	Digital Elevation Model
DHM	Digital Height Model
DTM	Digital Terrain Model
DSM	Digital Surface Model
GPS	Global Positioning System
ITM	Irish Transverse Mercator
ITS	Intelligent Transport System
MAE	Mean Absolute Error
NASA	National Aeronautics & Space Administration
OSi	Ordnance Survey Ireland
OSNi	Ordnance Survey of Northern Ireland
OSM	Open Street Maps
RMSE	Root Mean Square Error
SRTM	Shuttle Radar Topography Mission

1 Introduction and Motivation

Road grade is the change of elevation between two points along a road divided by the horizontal distance between the two points. Road grade is an essential factor in vehicles' fuel consumption. Taking road grade into account when choosing a route can save 20% of fuel consumption[3].

Eco-routing is choosing which route a vehicle should take from their origin to destination to use the least amount of fuel. There has been an increasing amount of work on eco-routing in the past few years[4]. It is important to create accurate energy consumption models to improve the effectiveness of eco-routing. Vehicle fuel consumption models depend on road grade, and the many models which leave out the effects of road grade[4] result in incorrect estimates of energy consumption[5]. Therefore more accurate road grade information would be of benefit to the development of eco-routing solutions.

It has been shown that of the best ways to obtain road grade on a city-wide scale is to use data gathered by intelligent transport systems (ITS) along with elevation from digital elevation models (DEMs)[6]. GPS traces from repeated bus runs outline the location of roads. The repetition of the bus runs increases precision of the road grade profiles by giving a more accurate representation of the bus route[7]. DEMs map elevation values to coordinates so that elevation profiles can be generated for bus routes which are then filtered to remove anomalous points[8]. Road grade profiles for the bus routes can then be calculated.

Global DEMs that model the elevation of the entire earth are made readily available to the public[1][9]. Such DEMs have been generated by satellites and have a maximum resolution of 30m^2[10]. The accuracy of road grade is dependent on the resolution of the DEMs used to calculate it. However, precise elevation data can be difficult or expensive to obtain. There is a need to analyse the effect that DEM resolution has on road grade calculations to determine the necessity of using more precise elevation data.

In this work, we aimed to determine how much of an improvement in road grade accuracy is afforded by an increase in DEM resolution. Elevation data of various resolutions is used with Dublin Bus GPS traces to calculate road grade information for bus routes around County Dublin, Ireland. Road grade is calculated using the pipeline for road grade automation developed by Singla et al. (2018)[11]. The road grade profiles calculated using different DEMs resolutions are compared to determine how the DEM resolution affects road grade profile accuracy.

Road grade accuracy was found to scale approximately linearly with DEM resolution. However, it was shown that using higher resolution DEMs affords only a small improvement in road grade accuracy, with only a decrease of less than 0.1 degrees in mean absolute road grade error across the routes when using a DEM of $10m^2$ over a DEM of $26m^2$.

Figure 1.1: The Dublin Bus routes of which road grade profiles were generated in this work.

This dissertation adheres to the following structure:

- Chapter 2 discusses other works that have done similar investigations into the effects of DEM resolution in different application areas and state of the art in generating large scale road grade information.
- Chapter 3 introduces the concepts and background information necessary for understanding the work.
- Chapter 4 describes in detail how the experiments in this work were carried out.
- Chapter 5 presents and discusses the results of the work.
- Chapter 6 concludes the results and evaluates their significance.
- Chapter 7 suggests future work that could be done following this research.

2 State of the Art

2.1 Similar Works

Elevation data has a wide range of use cases and is essential in many fields. There have been similar pieces of work where DEMs have been downsampled to different resolutions to evaluate the effect their resolution has on calculations from various application areas.

Xun Shi et al. (2012)[12] compared how LiDAR DEMs and widely available less accurate DEMs perform in calculations of slope gradient for knowledge-based digital soil mapping. They compared how the gradients generated from a 10m^2, a 1m^2 and a 5m^2 (downsampled from the 1m^2) DEM matched the gradient generated from 159 elevation values measured in the field by soil scientists. It was found that the gradients calculated from the LiDAR data (1m^2 and 5m^2) "performed significantly better" than the 10m^2 DEM. They did not discover any considerable differences between the 1m^2 DEM and the 5m^2 downsampled from it.

Zhao et al. (2010)[13] compared how DEM accuracy and resolution affects hydrologic parameters and modelling. They also compared 1m^2 and 10m^2 LiDAR DEMs and a 10m^2 DEM that was made using aerial photogrammetry. LiDAR has been shown to give more accurate elevation values than DEMs obtained using aerial photogrammetry[14]. They used each DEM to estimate soil loss in a watershed, and a field survey showed that the DEM resolutions had "substantial influence" on a number of the different parameters used in hydrologic modelling. The 1m^2 DEM was able to find many features of interest that the DEMs of a lower resolution could not.

Gillin et al. (2015)[15] compared the accuracy of terrain analysis metrics calculated from a 1m^2 DEM and 3m^2, 5m^2 and 10m^2 DEMs (downsampled from the 1m^2 DEM) to those calculated from field measured values. They were interested in the suitability of higher resolution DEMs for terrain analysis, as less precise DEMs have mostly been used because of their widespread availability. The similarities in the watershed boundaries, topographic metrics and RMSE values that were calculated for each of the DEMs revealed that the different DEMs did not produce much of an accuracy difference. The researchers found that the DEMs produced the best estimate of gradient when their resolution reflected the field measurement scale. They suggest that, for hydrologic modelling, DEM resolution be carefully selected based on the scale of features in the area being researched that have a significant effect on hydrologic response, rather than always choosing the DEM of the highest resolution.

Ruhoff et al. (2011)[16] performed an analysis of how DEM resolution affects topographic wetness index (TWI; which is a metric vital to hydrologic processes) calculations. They compared the TWI generated from $25m^2$, $50m^2$ and $100m^2$ DEMs. They found that DEM resolution had a significant effect on TWI values, but do not claim that the values obtained with a DEM of a higher resolution are necessarily better.

2.2 Generating Large Scale Road Grade Information

There are many different tested methods for generating road grade. Magrath et al. (2017)[6] evaluated different methods of determining road grade by replicating five studies where different methods were used. The researchers conclude that two methods met all requirements that had been set out in the study, making them particularly suited to large scale analysis and estimation of road grade profiles above the other methods.

The first of the two especially suitable methods is one suggested in a study by Wood et al. (2014)[8] where elevation values from the one arc-second SRTM DEM were appended to GPS speed traces recorded by vehicles to generate elevation profiles. The elevation profiles are then put through a filtering and smoothing routine to remove anomalous elevation values.

The second method replicated a study by Boroujeni et al. (2013)[7] where repeated runs were performed on several study routes, and their data was combined to create route profiles. The route profiles generated were used to generate road grade information. It was found that at least ten runs were necessary to get accurate road grade information, due to errors in the collected GPS information during the runs.

A paper by Singla et al. (2018)[11] reports on an approach to generate large scale road grade information that is based on the two methods suggested by Magrath et al. (2017)[6] described above. The work used Dublin Bus data and SRTM elevation data to generate road grade profiles for routes around the city of Dublin, Ireland. A pipeline for automating the road grade calculations was developed, with different methods tested at each stage of the pipeline to determine the methods which would result in road grade profiles most similar to those derived from elevation values taken from the Google Elevation API[17].

This work uses the pipeline for generating large scale road grade information set out in the work by Singla et al. (2018) to determine road grade profiles for many bus routes around the city of Dublin, Ireland. The high-level overview of the pipeline for generating road grade is described in the following section.

2.3 Road Grade Automation Pipeline

The pipeline for generating large scale road grade information set out by Singla et al. (2018) is described below. A version of this pipeline is used in this project to generate road grade profiles for bus routes around the city of Dublin, Ireland.

- Extract bus GPS traces (trips).
- Segment trips by route.
- Match GPS points from trips to their nearest OSM nodes.
- Clean trip data.
- Average trips into routes.
- Interpolate elevation values for each GPS point in routes to obtain elevation profiles.
- Filter elevation profiles.
- Calculate road grade profiles from elevation profiles.

3 Background

3.1 DEMs

A digital elevation model (DEM) is a 3D representation of the surface of a terrain. It is a geographic database specifically for the elevation of terrain. DEMs usually model the surface of a planet. In the context of this project, the DEMs model the surface of the earth. For most DEMs that represent the earth, and all of those discussed in this project, this is the elevation above (or below) sea level.

DEMs and digital terrain models (DTMs) represent the bare earth, which means that they do not take into account the elevation of vegetation or human-made structures such as buildings and bridges. Digital surface models (DSMs) and digital height models (DHMs) are different types of elevation models that do capture the natural and built features on the earth's surface[18] [19].

DEMs come in the form of raster data, which are grids of pixels. Each pixel value of a DEM is referred to as a 'tile'. A tile of a DEM represents a region on the surface of the planet. The size of the tiles dictates the horizontal resolution of the DEM. A "30m DEM" has tiles that are $30m^2$; giving one elevation value to each $30m^2$.

DEM resolution can sometimes be described in "arc-seconds". An arc-second represents the distance of latitude or longitude traversed on the earth's surface while travelling one second (1/3600th of a degree), which is 30m[1]. A 1 arc-second DEM is said to have a resolution of $30m^2$[20].

There are many methods for generating DEMs. Different methods of creating DEMs are employed depending on the intended use of the DEM. Among the methods of creating DEMs are satellite mapping and LiDAR, which is discussed below. DEMs are particularly suitable for generating road grade information on a large scale[6].

3.1.1 LiDAR

LiDAR stands for Light detection and ranging[21]. It is a method of surveying by emitting pulses of light at a target and measuring the distance to the target by sensing the light's reflection. The differences in the return times of the pulses of light are used to create 3D models of the target.

Airborne LiDAR is when an aircraft uses LiDAR to create a 3D point model of a landscape[22]. LiDAR data collected this way creates very high-resolution DEMs[23]. As LiDAR DEMs are collected with aircraft that are much closer to the earth's surface than satellites, it is more challenging to create LiDAR DEMs that span a region as large as those created with satellites. LiDAR DEMs have a much higher resolution than satellite made DEMs, having resolutions of 30cms or better[24]. An example of high-density LiDAR data collected over Dublin city centre, Ireland is shown in figure 3.1. The data in visualised in figure 3.1 comes from the New York University Spatial Data Repository[25] and has 335 points per m^2.

Figure 3.1: An example of high-density LiDAR data in the city of Dublin, Ireland.

3.1.2 Coordinate Systems

A map projection is a transformation of latitude and longitude coordinates onto a plane[26]. A coordinate reference system defines a map projection and a conversion between different spatial reference systems. Irish Transverse Mercator (ITM) is the geographic coordinate system for Ireland. It was implemented jointly by the Ordnance Survey Ireland and the Ordnance Survey of Northern Ireland[27]. ITM coordinates are also known as EPSG:2157 coordinates[28]. The elevation data used in this project is defined using the ITM coordinate system. For data to be sampled from the elevation files, GPS points must be transformed from latitude and longitude into ITM coordinates. An example of coordinate systems and map projections can be seen in figure 3.3.

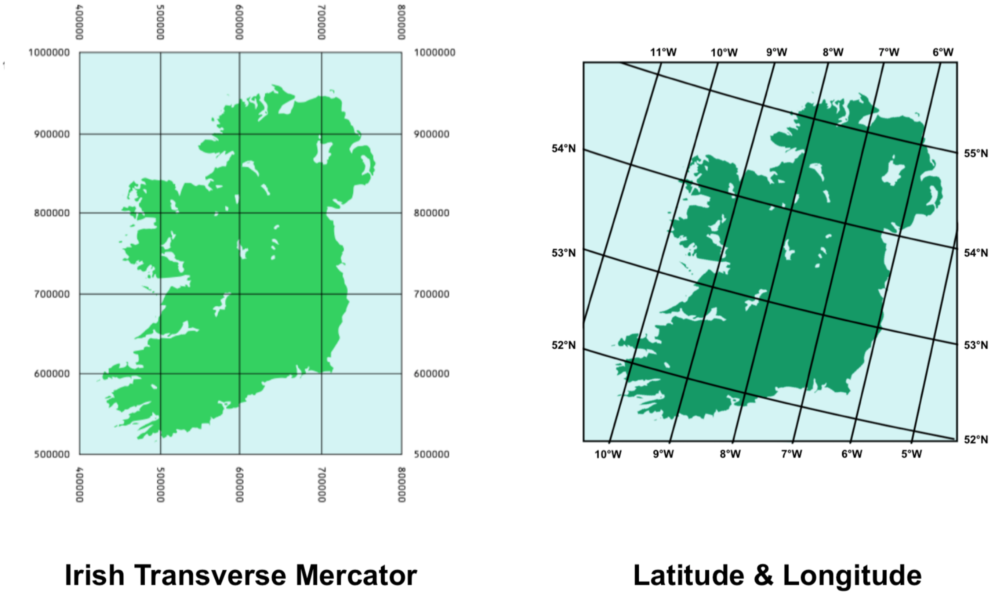


Figure 3.2: The Irish Transverse Mercator and Latitude and Longitude coordinate reference systems.

3.2 OSM

"OpenStreetMap is a free, editable map of the whole world that is being built by volunteers largely from scratch and released with an open-content license."[29]. OSM is made up of billions[30] of nodes that represent different features and amenities of the world. Nodes have tags to describe what type of feature or amenity they are. Nodes with the "highway" tag are taken to make up part of a road. In this project, OSM nodes are used to help in the cleaning process. GPS points are assigned the closest OSM "highway" node to them to determine the part of a road they are supposed to represent.

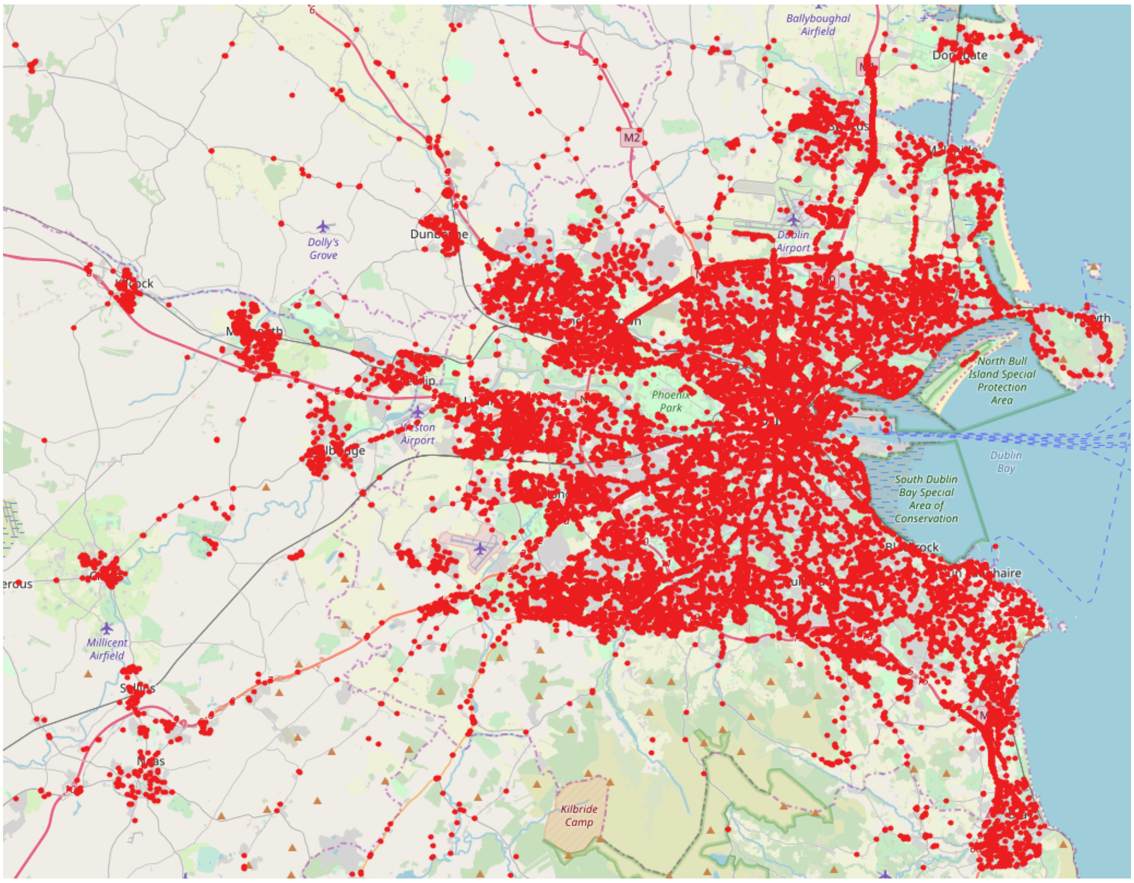


Figure 3.3: The OSM nodes tagged with "highway" in County Dublin, Ireland.

3.3 Filters

During several steps of the process of estimating road grade profiles, data is filtered to remove anomalous or erroneous points.

3.3.1 Moving Average Filter

A moving average filter is used to remove random noise from signals and is the most commonly used filter in digital signal processing[31]. It takes an array of 'n' consecutive points of a signal and returns the unweighted average value of those points.

In this project, a moving average filter is used when averaging GPS traces into a single route profile. The moving average filter is described by the following equation:

$$\mathrm{y}[i] = \frac{1}{n}\sum_{j=0}^{n-1}\mathrm{x}[i+j]$$

Where x[ ] is the input signal, y[ ] is the output signal, and *n* is the size of the filter.

3.3.2 Savitzky Golay Filter

The Savitzky Golay filter is a filter which is useful for reducing noise in signals while maintaining the shape and height of waveform peaks[32]. It was developed by Abraham Savitzky and Marcel Golay in 1964[33]. The elevation filtering stage of the road grade automation pipeline makes use of the Savitzky Golay filter.

3.3.3 Advanced Binomial Filter

Binomial filters are an approximation of Gaussian filters[34]. A binomial filter applies weighted averaging to the input signal. A filter of size *n* is defined which has *n* weights. Each value in the output signal is given the weighted average of the *n* points around its corresponding input value. The equation that defines the binomial filter is given below:

$$\mathrm{y}[i] = \frac{1}{\mathrm{w}^*}\sum_{j=1}^{j=5}\mathrm{x}[i+j-3].\mathrm{w}[i]$$

Where x[ ] is the input signal, y[ ] is the output signal, *n* is the size of the filter, w[ ] is the array of weights, and w* is the sum of the weights.

In this project, a filter of size $n = 5$ is used with the weights [1,4,6,4,1], following the work done by Singla et al. (2018)[11].

3.3.4 Combined Binomial Savitzky Golay Filter

In the combined binomial Savitzky-Golay filter, an advanced binomial filter and a Savitzky-Golay filter are applied to the input signal to produce two output signals. The output of the combined filter is the average of those two output signals. Each pair of corresponding points from the two signals are added together and divided by two to get their corresponding output value.

3.4 Interpolation

During several stages of the road grade automation pipeline, values are interpolated from data to estimate values of data which are missing or to get a more accurate estimation of what values should be at certain points.

3.4.1 Linear Interpolation

Linear interpolation is used to estimate the value of a signal at points along the signal where there is no data. The values of the data points surrounding the missing datapoint are used to estimate the missing datapoint's value. The equation for linear interpolation is given below:

$$y = y0 + (x - x0)\frac{y1 - y0}{x1 - x0}$$

Where the missing datapoint (x,y) can be found between the points (x0,y0) and (x1,y1).

3.4.2 Bilinear Interpolation

Bilinear interpolation is used to estimate the unknown value of a point using the known values of the four nearest points. The values of the four nearest points surrounding the point of interest are put into a weighted average to determine the output value of the bilinear interpolation. The weights applied to the interpolation points are inversely proportional to their distance from the point who's value is being interpolated[35]. A variation of bilinear interpolation is used in this work to estimate elevation values for GPS points on routes.

3.5 Calculations

3.5.1 Distance Calculation

There are different ways of calculating the distance between two GPS points. The most commonly used approach of calculating the distance between two GPS points is the Haversine formula[36]. The work by Singla et al. (2018)[11] used the Haversine formula to calculate the distance between points.

This piece of work uses the Vincenty formula for the distance between two points. The Haversine formula is based on the assumption that the earth is a sphere, whereas the Vincenty formulae are based on the assumption that the figure of the earth is an oblate ellipsoid[37]. This assumption results in more accurate calculations of distance[38], with the caveat of a greater calculation time[39].

3.5.2 Road Grade Calculation

Road grade is the slope of the road at a point. To create a route's road grade profile, the slope between consecutive points at regular intervals along the route is calculated. For this piece of work, road grade is calculated as the ratio of the vertical distance between two points to the horizontal distance between the two points. The formula for road grade angle in radians is given below:

$$\Theta = \arctan(\frac{\Delta\psi}{\Delta d})$$

Where Θ is the road grade angle in radians, $\Delta\psi$ is the vertical distance between the two points, and Δd is the horizontal distance between the two points.

In this piece of work, road grade is expressed in degrees. Road grade can be converted from radians to degrees with the following equation:

$$\Theta_d = \Theta_r \frac{180}{\pi}$$

Where Θ_d is the road grade angle in degrees and Θ_r is the road grade angle in degrees.

3.6 Statistical Measures of Error

The two statistical measures of error used in this work are the root mean square error and the mean absolute error. Both errors have differences in how they describe the error between predicted values and observed values. Each measure of error has advantages over the other in certain circumstances, but the use of both gives a better picture of the errors in the data[40] [41].

3.6.1 Root Mean Square Error

Root mean square error (RMSE) way of expressing the error between a set of predicted values and a set of observed values. In RMSE, outliers of a higher magnitude have more of an effect on the overall error value. This is because the squaring of the differences means that high errors affect the RMSE value significantly.

RMSE is defined as the square root of the mean of the squared differences between the corresponding elements of the predicted values and the observed values[42].

$$RMSE = \sqrt{\Sigma_{i=1}^{n}(p_i - o_i)^2/n}$$

Where n is the number of elements in the predicted and observed values, p_i is the i^{th} element of the predicted values and o_i is the i^{th} element of the observed values.

3.6.2 Mean Absolute Error

Mean absolute error (MAE) is another way of expressing the error between a set of predicted values and a set of observed values. A high MAE means that there are many errors throughout the data. Single errors with a high magnitude have less of an impact on the overall output value than in RMSE as the value of the differences between predicted and observed values are not squared before being averaged.

The absolute error between a predicted and observed value is the absolute value of the difference between them. The mean absolute error of a set of predicted values against a set of observed values is the mean of the absolute errors between each corresponding predicted and observed element of the two sets.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|p_i - o_i|$$

4 Experimental Setup

4.1 Datasets

4.1.1 Bus Data

Dublin Bus data was made freely available by Dublin City Council on data.gov.ie in June 2013[43]. Data representing trips that buses took in January 2013 was downloaded. This data is made up of many points that each represent part of a bus journey that took place during that time. A new datapoint was recorded by the buses roughly every 20 seconds. Each datapoint includes the following relevant data as well as other data not relevant to this work.

- **time_stamp**: The datetime at which the datapoint was recorded at (which allows nodes to be ordered by when then they were recorded).
- **journey_pattern_ID**: The unique identifier of the route (and inbound or outbound direction) the bus was on.
- **vehicle_journey_ID**: The unique identifier of the trip the bus was taking on the route.
- **lon_WGS84** & **lat_WGS84**: The coordinates of the bus at the time the datapoint was recorded.

4.1.2 OSM Data

OSM data was downloaded. The OSM data of interest is the ID and coordinates of the nearest OSM "highway" node of every GPS point in the Dublin Bus dataset.

OSM has an API called Overpass[44]. Users can query the Overpass API for nodes that fit certain requirements near some coordinates[45]. This query can be used to get the nearest "highway" node to the GPS coordinates of a point in the Dublin Bus data. Unfortunately, Overpass rate limits users so that they can only make a certain number of requests before having to wait a period of time until they can make their next request. This rate limiting makes it infeasible to make a request for every single node in the Dublin Bus dataset.

Users can also query the Overpass API for nodes that fit specific requirements and are located within a bounding box. For instance, all "highway" nodes found within the region Dublin Bus operates in can be downloaded within seconds. With all the nodes of interest downloaded to a local machine, the closest OSM nodes to Dublin Bus GPS points could be extracted without making further network requests. The following trivial algorithm was created to find the nearest OSM to every point in the Dublin Bus data, and to append that node's information to the Dublin Bus datapoint.

```
for point in dataset:
    closest_node = osm_nodes[0]
    for node in osm_nodes:
        if distance(point,node) < distance(point,closest_node):
            closest_node = node
    point['closest_osm_node'] = closest_node
```

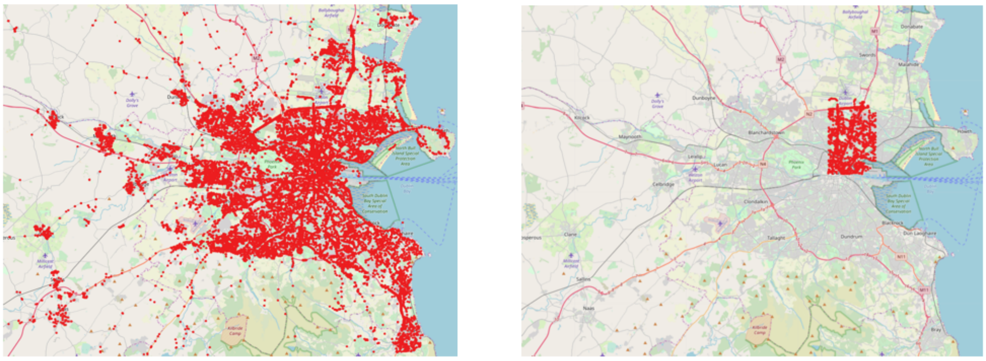
This algorithm proved too slow when iterating through each "highway" node in the entire region that Dublin Bus operates. The process of assigning OSM nodes to every point in the Dublin Bus dataset was streamlined by downloading all the OSM "highway" nodes that can be found within the region that each *route* operates. The difference in these approaches is depicted in figure 4.1. Only highway nodes which appear in the region that Dublin Bus datapoint's route operates in are considered when assigning an OSM node to the datapoint. This sped up the "point matching" stage of the pipeline considerably.

Once this process had finished, each point in the Dublin Bus dataset had been updated to have the ID and longitude and latitude coordinates of its closest OSM "highway" node.

4.1.3 Elevation Data

LiDAR elevation data was provided by OSi. It covers the entire area in which that Dublin Bus operates.

The data is 7.25GB large. It is in .xyz format. .xyz files are made of rows of three pieces of data. Each row in this dataset has an x coordinate, a y coordinate and an elevation value. The dataset uses the Irish Transverse Mercator coordinate system[28], which was implemented by OSi themselves along with OSNi[27]. The elevation values at each location are in metres, representing the elevation above sea level. The data is made up of 273 files that make up this area. Each row of the files represents the elevation value given to one 2m^2



OSM “highway” nodes in Dublin **OSM “highway” nodes in the 27B’s operating region**

Figure 4.1: The OSM highway nodes in the region that Dublin Bus operates vs the OSM highway nodes in the region that the 27B bus route operates.

area in Dublin.

The elevation data was downsampled from 2m^2 to 6m^2, 10m^2, 20m^2 and 26m^2. To reflect the elevation values of freely available global DEMs which are of a lower resolution [46], the larger tiles in the downsampled DEMs were each given the average value of the smaller 2m^2 tiles from the 2m^2 DEM to which they correspond. This downsampling technique is visualised in figure 4.2.

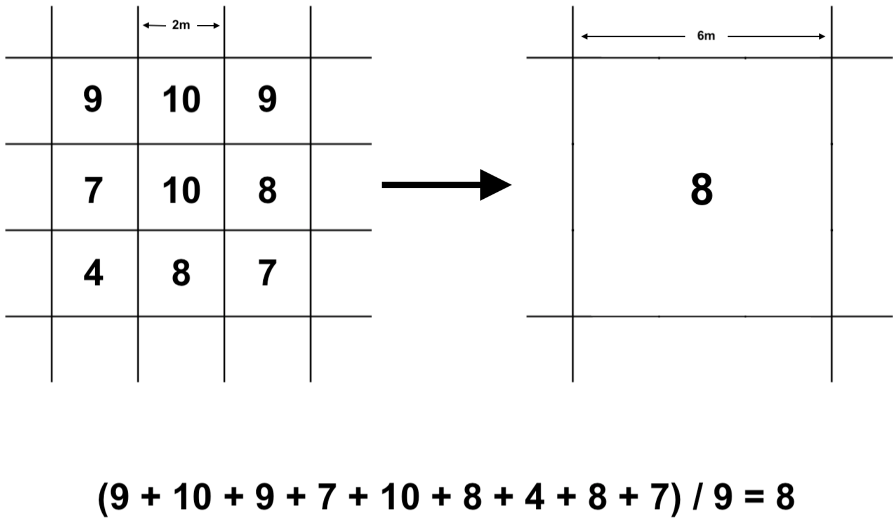


Figure 4.2: An example of the downsampling of a 2m^2 DEM to a 6m^2 DEM.

4.2 Method Overview

4.2.1 Creating Route Profiles

A route profile was created for each route by performing the following steps, which are visualized in figure 4.3:

- Get the datapoints that have been recorded for the route.
- Assign each datapoint its nearest OSM node.
- Segment the datapoints by the trip during which they were recorded.
- Clean the trip data to get rid of anomalous trips.
- Recombine trip datapoints, ordering datapoints by cumulative distance on their respective trips.
- Average datapoints into route profile.

4.2.2 Creating Road Grade Profiles

A road grade profile was created for each route by performing the following steps:

- Interpolate the elevation for each point in the route.
- Generate an elevation profile for the route using the points' cumulative distances.
- Using linear interpolation, downsample the elevation profile to points at 100m intervals.
- Filter the downsampled elevation profile with a combined binomial Savitzky-Golay filter.
- Discard and backfill points which have elevation values that differ from their old one by more than the threshold (10m).
- Filter the elevation profile once more with the combined binomial Savitzky-Golay filter.
- Generate a road grade profile from the route's elevation profile.

The above steps were performed for every route using the 2m^2 LiDAR DEM and the downsampled 6m^2, 10m^2, 20m^2 and 26m^2 DEMs.

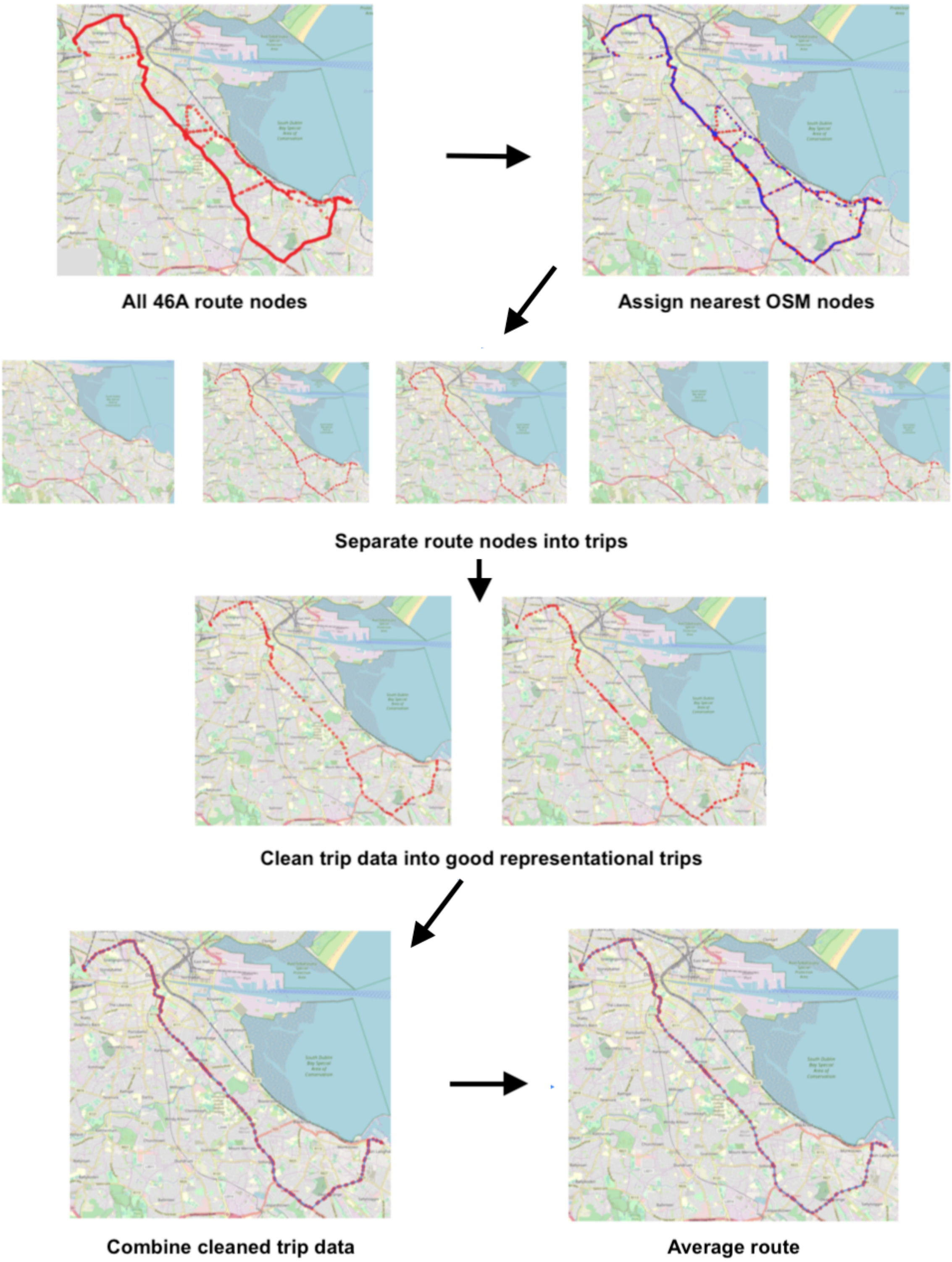


Figure 4.3: The steps involved in creating a route profile from Dublin Bus data.

4.3 Data Cleansing

Dublin Bus data is not ready to be put into the road grade calculation until it has been cleaned. Each route is represented by many trips which are one GPS trace of a journey that a bus took on a particular day. The multiple trips that represent a route in the Dublin Bus data may be very different from each other. Some trips have very few points recorded, while other trips for the same route may have many points. Buses may have been diverted for some time due to reasons such as construction work on roads. This means that for that time, the GPS traces recorded can be very different from the GPS traces of the rest of the trips that represent that route. These diversions do not represent the usual running of the route well, as they are full of GPS points that are not in the regular route. Also, some of the GPS traces that represent a route may start in the middle or may end much before the bus has reached its destination. All of these occurrences lead to anomalies and outliers in the data which must be dealt with to obtain an accurate representation of the bus route.

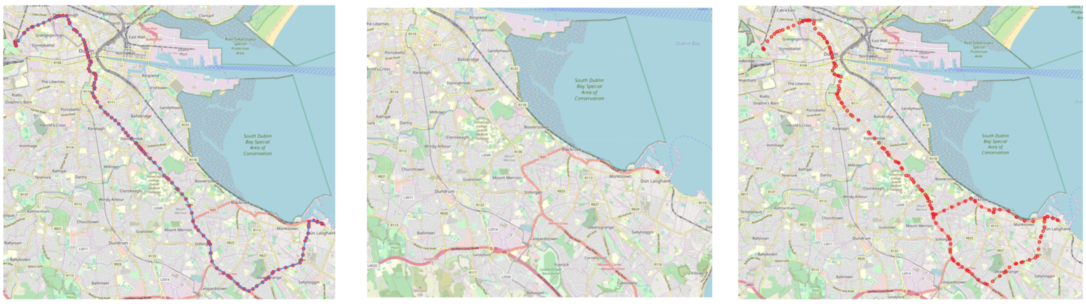


Figure 4.4: The 46A bus route profile alongside two of the recorded 46A GPS traces. One trip has only one GPS point recorded. The other trip includes a path that is not in the route profile.

Before any road grade calculations are done, the bus data is cleaned by the following steps so that the output data will represent the bus routes as accurately as possible:

For each route:

- The points that correspond to that route are extracted from the data and sorted by their time stamp.
- The points are separated into 'trips' using their **'vehicle_journey_id'** which corresponds to which bus journey on which the point was recorded.
- Any trip which is made up of less than 50 points is discarded, as the trips which most accurately represent the routes are made up of hundreds of points. GPS traces with

less than 50 points represent a version of a route which cuts many corners and misses out on critical parts of the route. These trips end up having a cumulative distance which is far too small, which would mess with the algorithm that combines trips into one single representation of the trip if they were not discarded.

- A set of OSM nodes is created from the nearest OSM node to each of the points in all the trips for the route.
- The number of times each OSM node appears as a point's nearest OSM node across all the trips in the route is recorded.
- The OSM node appearance counts from the previous step are normalised to a frequency with respect to the total number of trips for the route so that, for example, an OSM node that appeared in a third of the route's trips has a frequency of $\frac{1}{3}$.
- Any trip that has a node with a frequency of less than the threshold is discarded. This means that every trip that remains has only nodes which appear in more than the threshold number of trips. In the work done by Singla et al. (2018), the threshold value was set as 0.333333. In the Dublin Bus data used in this work, some routes would have no trips meeting the criteria set out if this threshold value was taken. Instead, the threshold value for each route was chosen as the highest threshold possible that would return some trips to represent the route. This route-specific threshold value resulted in trips that represented each route as well as possible.
- The remaining trips are considered 'good trips' that represent the route accurately.

4.4 Route Averaging

Now that each remaining GPS trace is considered as accurately representing the route, they can be combined to create a single representation of the bus route. The following steps are followed to combine GPS traces into a route profile:

- For each trip, all of its points are assigned their cumulative distance along the trip. Each point's cumulative distance is the distance between it and the previous point in the GPS trace added to the previous point's cumulative distance. The cumulative distances of points are depicted in figure 4.5.
- All points from every trip are combined into a list which is then sorted by the cumulative distances that were assigned to the points.
- Now there are many GPS points that represent the route, but many GPS points that were recorded with some error are now beside each other. If these points were used to define the route, it would be seen that the total distance of the route would be far greater than it is in reality. To get a more accurate representation of the route, the points are put through a moving average filter as described in Section 3.3.1.

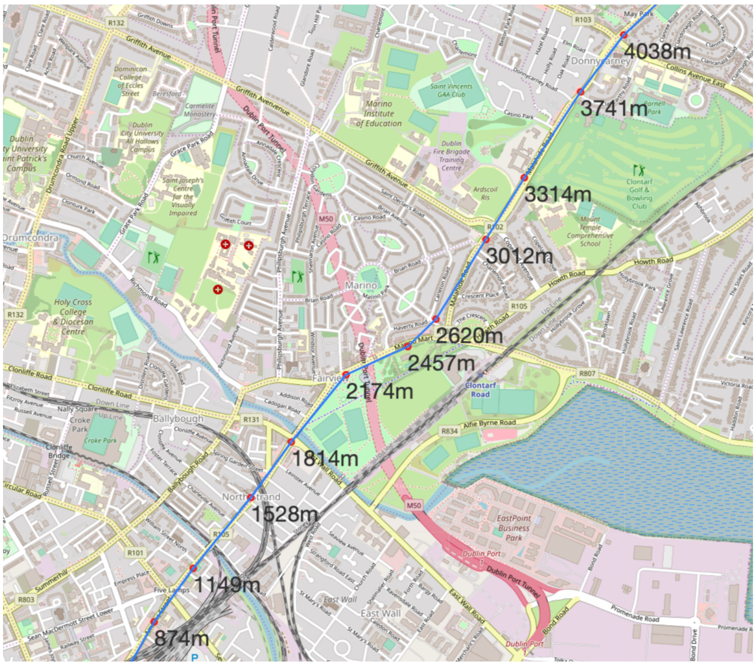


Figure 4.5: The cumulative distances of nodes along a GPS trace of the 27B bus.

Once the points from all the trips have been put through the moving average filter, a smaller number of points which more accurately represent the trip are left. An example of the points before the moving average filter is applied to the 27B bus route can be seen in figure 4.6. It can be seen between Beaumont and Coolock that the route profile goes back and forward. These errors are fixed by the averaging step.

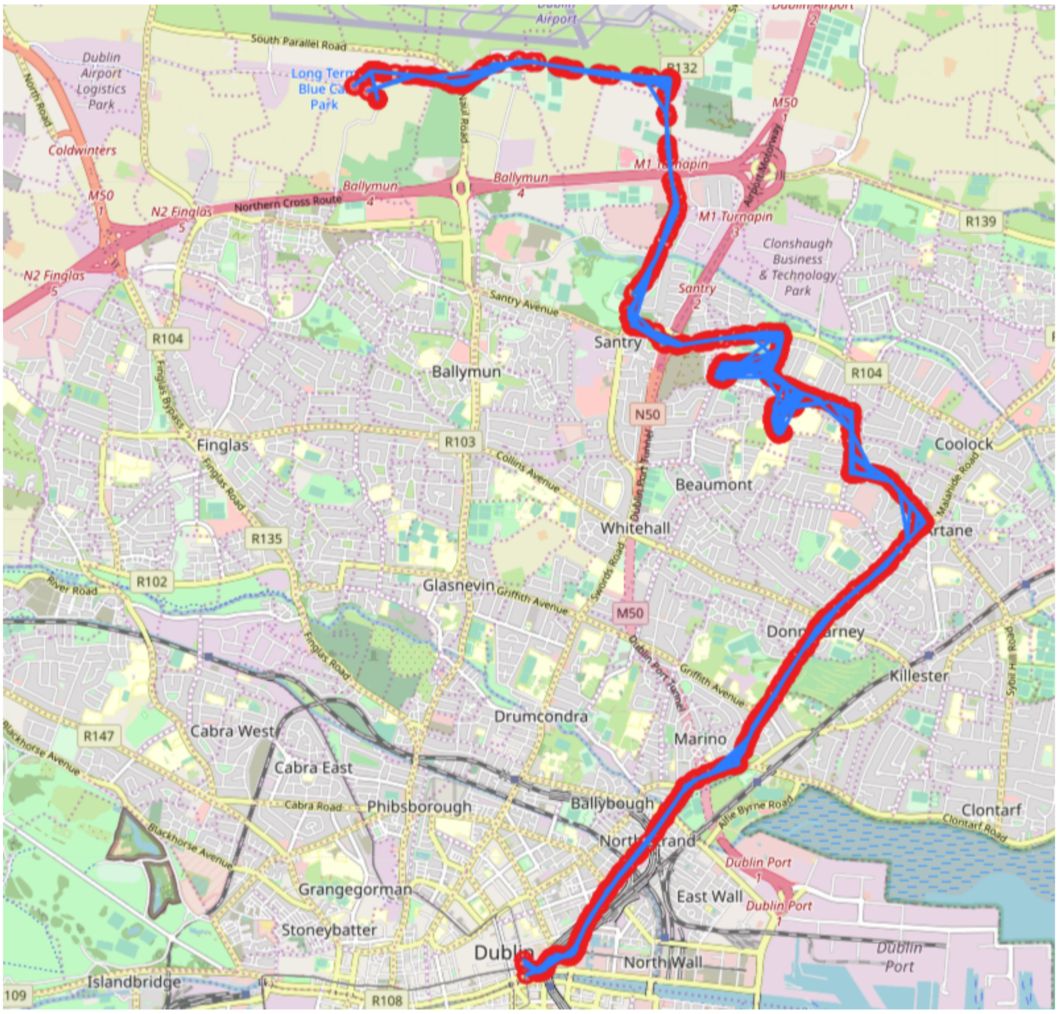


Figure 4.6: The route profile of the 27B without averaging GPS points from different trips.

An example of the points after the moving average filter is applied to the 27B bus route can be seen in figure 4.7. These are the points which are used to represent the route in the road grade calculations.

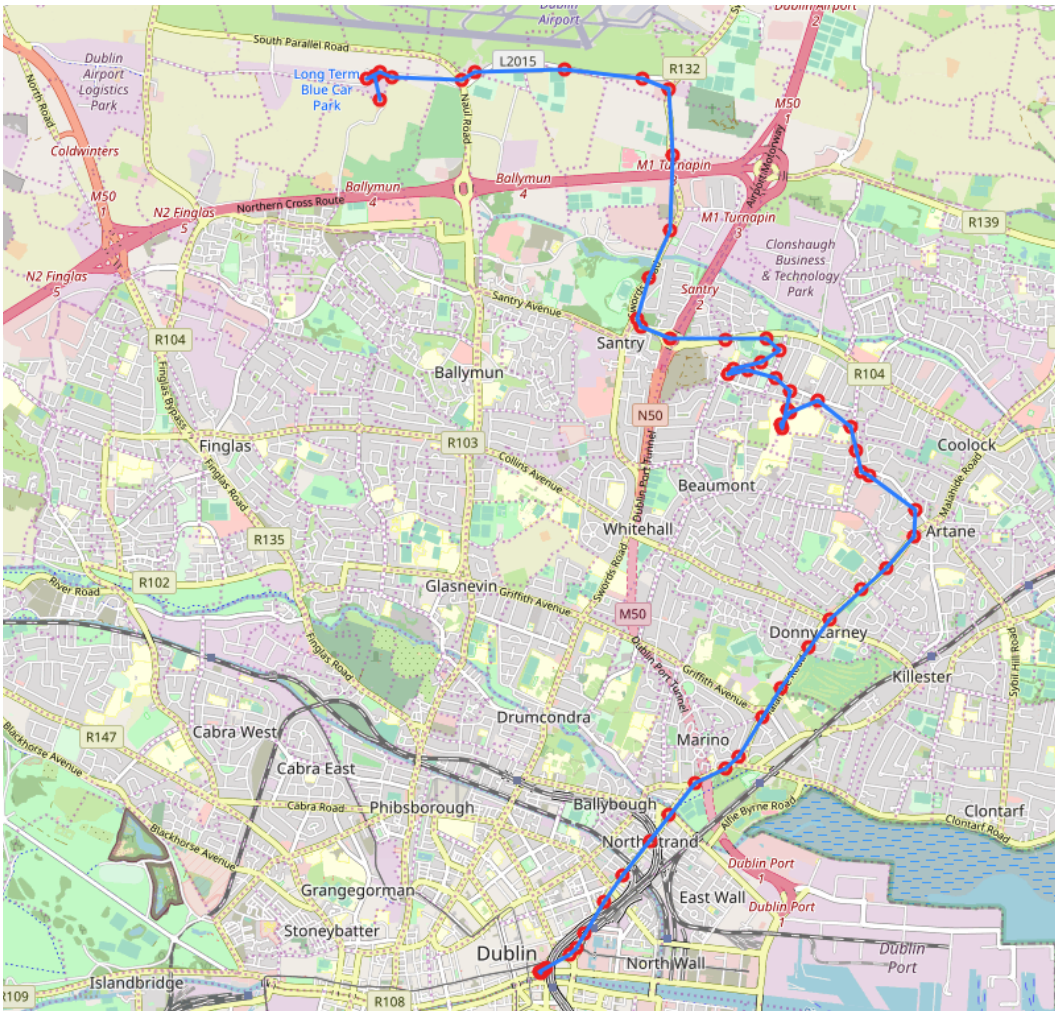


Figure 4.7: The 27B bus route after averaging the data from different trips.

4.5 Generating Elevation Profile

4.5.1 Elevation Interpolation

It has been shown that when using tiled DEMs, elevation is more accurate when interpolated from the values in the DEM[2]. Singla et al. (2018) found that Advanced Bilinear Interpolation was the best interpolation technique for achieving accurate elevation values. This piece of work also uses the Advanced Bilinear Interpolation technique to get more accurate elevation values for each GPS coordinate. The interpolation of GPS points' elevation values that is used in this work is described in the following section.

Advanced Bilinear Interpolation

Rather than assign each GPS point on the route the elevation value given to it from the DEM, elevation values were interpolated using the nearby tiles in the DEM. Figure 4.8 visualises the finding of points to be used in the interpolation. The grid on the left depicts the GPS point in the route who's elevation is to be found. The grid on the right depicts the DEM tiles which contain the interpolation points, (1,3), (2,3) & (2,4), and the interpolation points themselves (ptA, ptB, ptC & ptD).

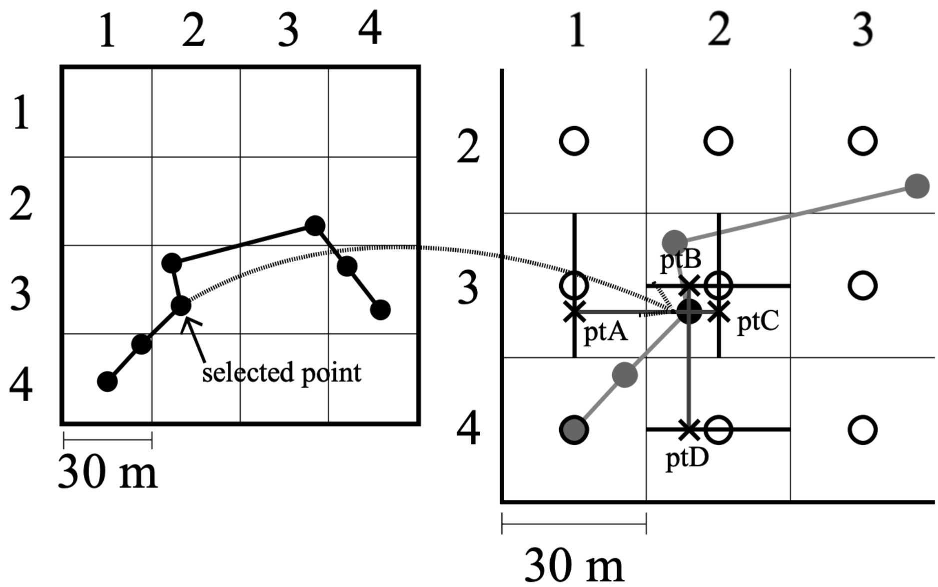


Figure 4.8: An illustration of the selection of the points to be used for interpolating GPS points' elevations (from Henriques and Bento (2013)[2]).

The advanced bilinear interpolation technique is described in the steps below:

- The point is converted to its EPSG2157 (ITM) format to get the x and y values of the point in the DEM's coordinate reference system.

- The x and y values are rounded to find the column and row number of the point's tile. These values indicate the index of the tile in the DEM.
- If the difference between the point's x value and the tile's column number is more than half of the DEM tile size, then the horizontally neighbouring tile chosen for interpolation is taken from the next column. If the difference is less than half the DEM tile size, then the horizontally neighbouring tile chosen for interpolation is taken from the previous column, as is shown in the example in figure 4.8. The same process is applied with the point's y value and the tile's row number to find the vertically neighbouring tile to be chosen for interpolation.
- The three tiles that will be used in the interpolation have now been found. In the example shown in figure 4.8, the interpolation tiles (1,3), (2,3) & (2,4).
- The coordinates of the four interpolation points are created using combinations of the x and y values of the original point and the center of the interpolation tiles:
  - ptA: (x value of horizontal neighbour's center, y value of original point)
  - ptB: (x value of original point, y value of original point's tile's center)
  - ptC: (x value of original point's tile's center, y value of original point)
  - ptD: (x value of original point, y value of vertical neighbour's center)
- The elevation values of the four interpolation points is extracted from the DEM.
- The distance between the original point and each of the four interpolation points is found.
- The weighted average of the four elevation points is found. The weight applied to each interpolation point's elevation in the average is the inverse of the distance between the interpolation point and the original point.
- The weighted average of the interpolation points is assigned as the GPS point's elevation value.

4.5.2 Elevation Extraction

The elevation files are in .xyz format. The python library 'rasterio' was used to sample the data from the elevation files. The 'rasterio' 'sample' function takes a list of coordinates to be sampled from a file and returns the elevation values for each coordinate. In the interest of efficiency, a list of coordinates to be sampled from each file must be made instead of sampling a file for one GPS point at a time. Having to create a list of elevation values needed from each file before sampling means that the steps of the elevation interpolation, as described above, are performed in a somewhat non-intuitive order; finding all the necessary information to perform the interpolation except the elevation values which will be found later.

A list of files that need to be sampled is created. A dictionary that maps EPSG coordinates to elevations is also created. For each point in the route, its four interpolation coordinates are calculated and added to the list that corresponds to the file from which their elevations will be sampled. Once the interpolation coordinates of all the points in the route have been added to a file's list, each file is opened, and all of the related coordinates' elevations are sampled. As the coordinates' elevations are sampled, they are added to the dictionary. Once there is an elevation value for every necessary coordinate, the elevation interpolation can be done for every point in the route by calculating the weighted average elevation of its interpolation coordinates.

4.5.3 Cumulative Distance

As the elevation for every point in the route is being interpolated, the cumulative distance of the point along the route is also calculated. Once each point is given a cumulative distance and elevation value, these can be plotted as (x,y) coordinates to display the elevation profile for the route. The elevation profile for the 27B bus route can be seen in figure 4.9.

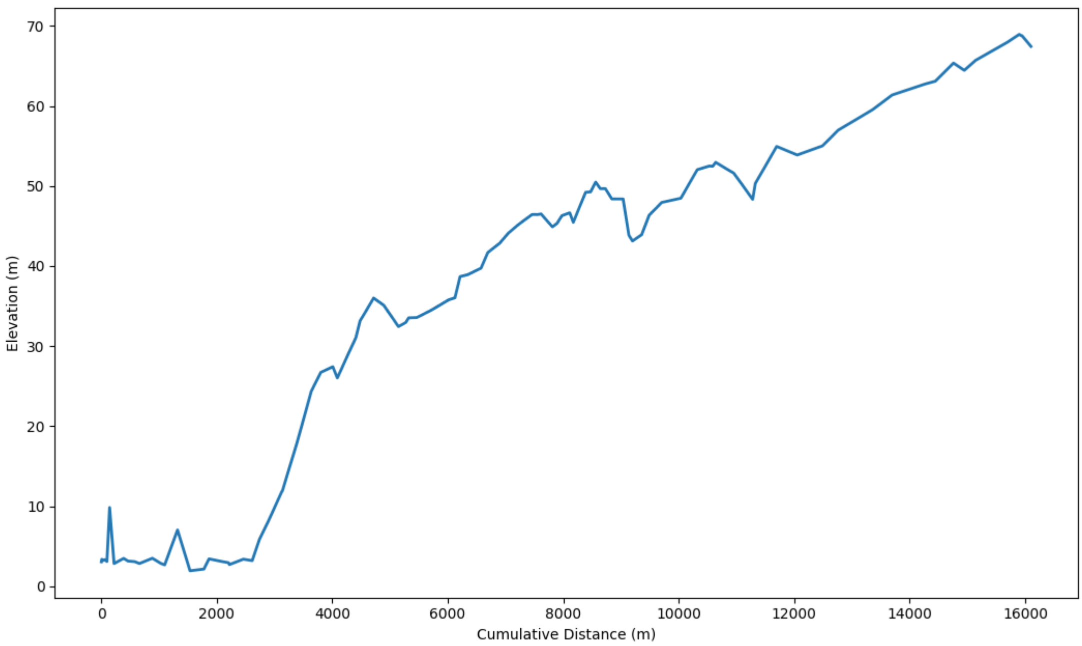


Figure 4.9: The elevation profile of the 27B bus route generated from the 2m^2 DEM.

4.6 Elevation Filtering

Once the elevation profile for a route is found, a function to approximate it is interpolated using linear interpolation. This function is used to downsample the elevation profile to points which are spaced out by an interval of 100m along the route. An example of the approximated function is shown in figure 4.10.

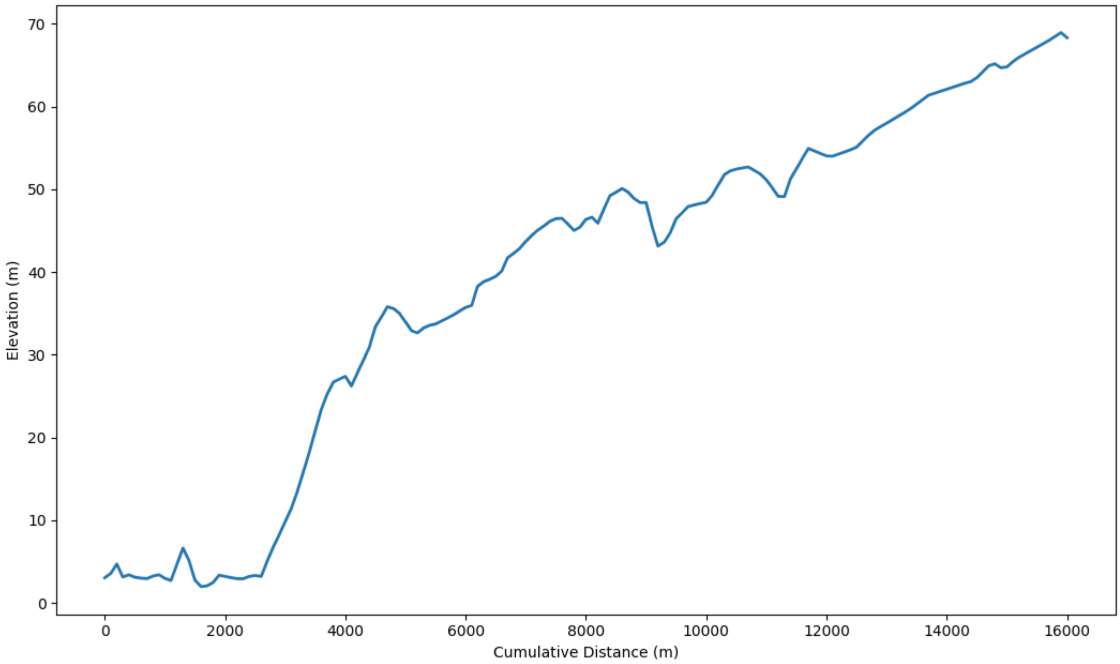


Figure 4.10: The elevation profile of the 27B bus route after being downsampled to 100m intervals.

The downsampled points are passed into the combined binomial Savitzky-Golay filter described in Chapter 3. Any point who has a filtered elevation value and original elevation value that differ by more than the threshold (10m) are discarded and backfilled using linear interpolation. After the backfilling of the elevation values, the points are once again passed through the combined binomial Savitzky-Golay filter. The filtered elevation profile for the 27B bus route can be seen in figure 4.11.

After the points have been filtered a second time, they are used to interpolate a new function to represent the elevation profile. A new elevation profile is generated by calculating the output elevation for each original point in the route using the interpolated function. The final result for the 27B bus route can be seen in figure 4.12.

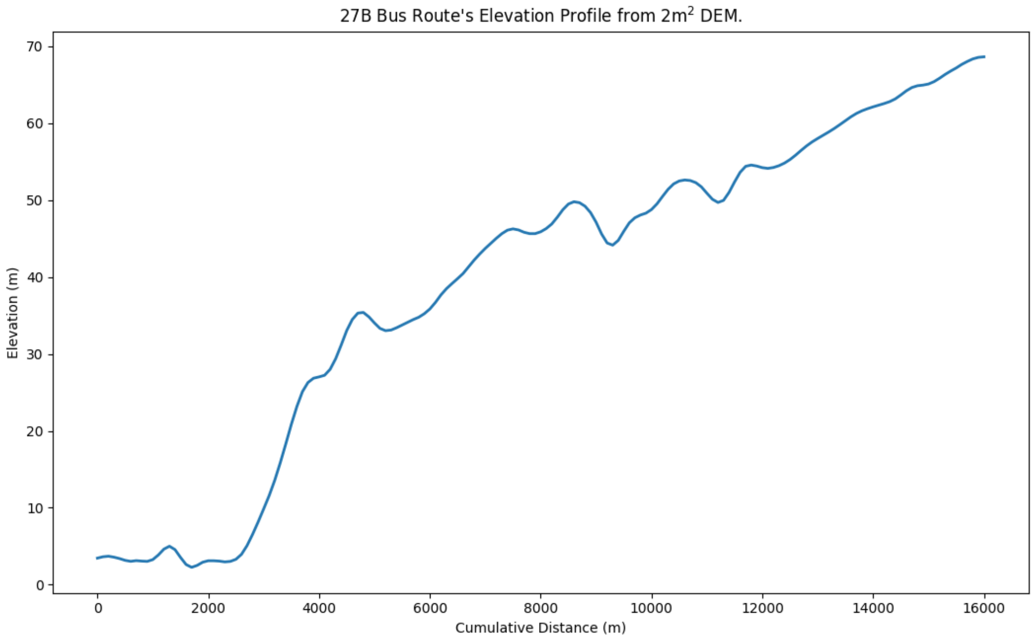


Figure 4.11: The elevation profile of the 27B bus route after being put through the combined binomial Savitzky-Golay filter.

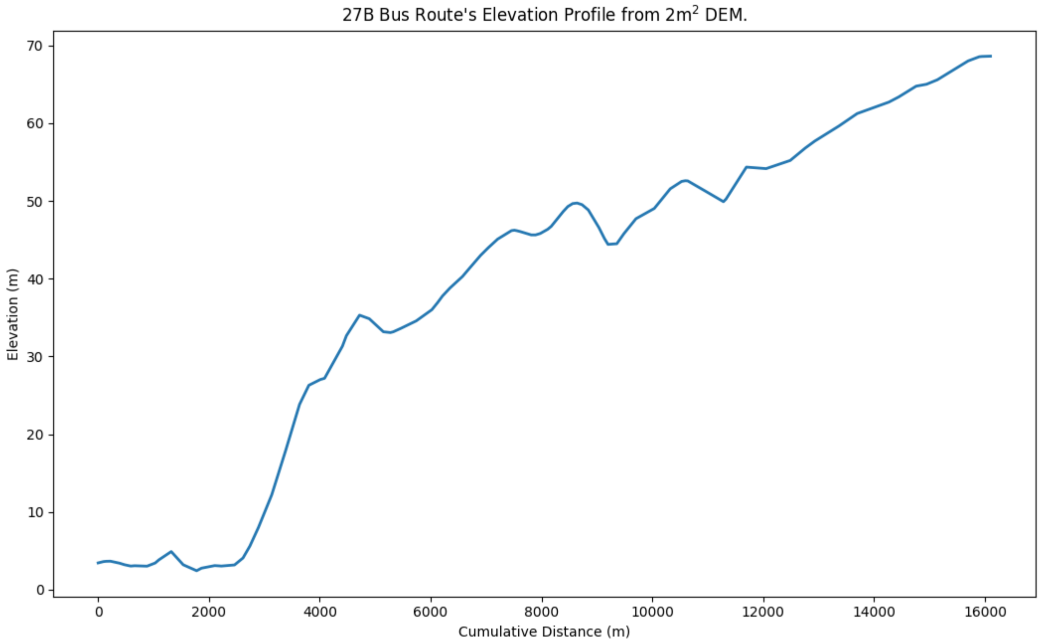


Figure 4.12: The elevation profile of the 27B bus route after being the filtering has been completed.

4.7 Road Grade Profile Generation

Once the filtered elevation profile has been obtained, the road grade profile for the route can be calculated. The road grade calculation function, as discussed in Chapter 3, is used to convert the elevation profile into a road grade profile. Road grade angles are calculated at 100m intervals along the route by interpolating the elevation profile. The road grade profile merely is the slope of the elevation profile. The road grade profile for the 27B bus route can be seen in figure 4.13.

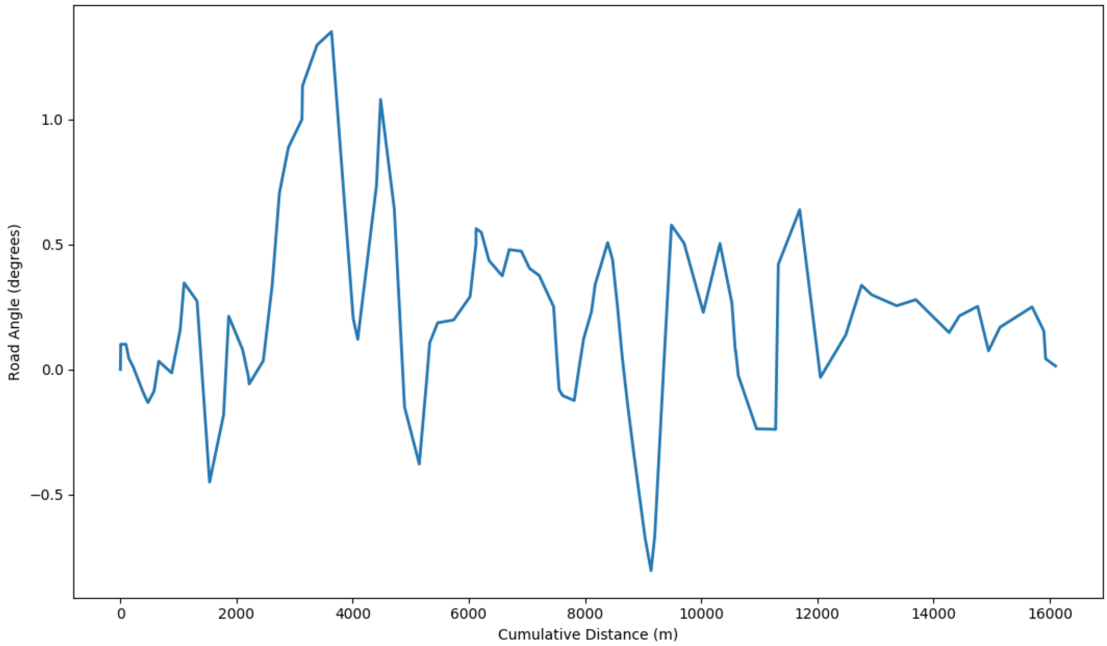


Figure 4.13: The road grade profile of the 27B bus route from the $2m^2$ DEM.

4.8 Comparing Road Grade Profiles

Road grade profiles were created for 15 bus routes using the $2m^2$, $6m^2$, $10m^2$, $20m^2$ and $26m^2$ DEMs.

Using the $2m^2$ DEM road grade profiles as ground truths, the root mean square error and mean absolute error for the $6m^2$, $10m^2$, $20m^2$ and $26m^2$ DEM road grade profiles were calculated.

5 Results and Discussion

5.1 Elevation Profiles

The elevation profiles generated for the 27B bus route are shown in figure 5.1. As a result of interpolating the elevation for each point and smoothing the elevation profiles, there is no significant difference in the elevation profiles between the elevation profile of the $2m^2$ DEM and the elevation profiles of the lower resolution DEMs.

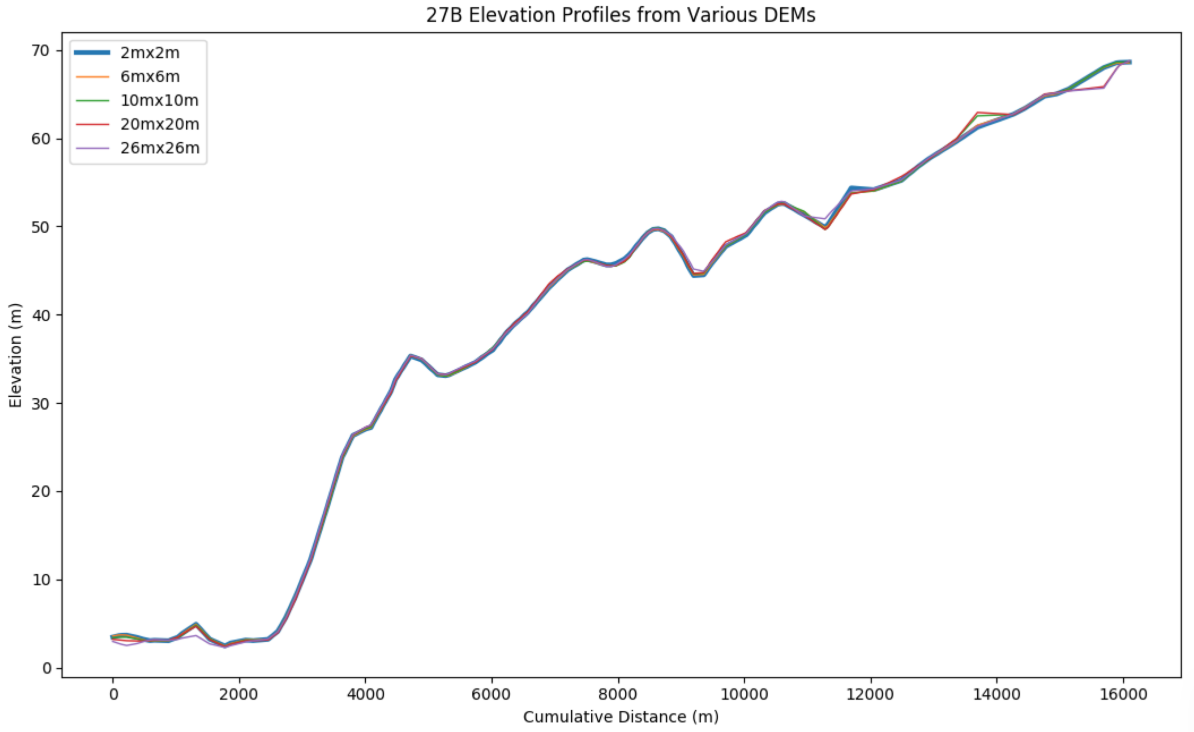


Figure 5.1: The elevation profiles generated for the 27B bus route from $2m^2$, $6m^2$, $10m^2$, $20m^2$ and $26m^2$ DEMs.

5.2 Road Grade Profiles

The road grade profiles generated for the 27B bus route are shown in figure 5.2. It can be seen that for the most part, the road grade profiles of the lower resolution DEMs match the road grade profile generated by the 2m^2 DEM.

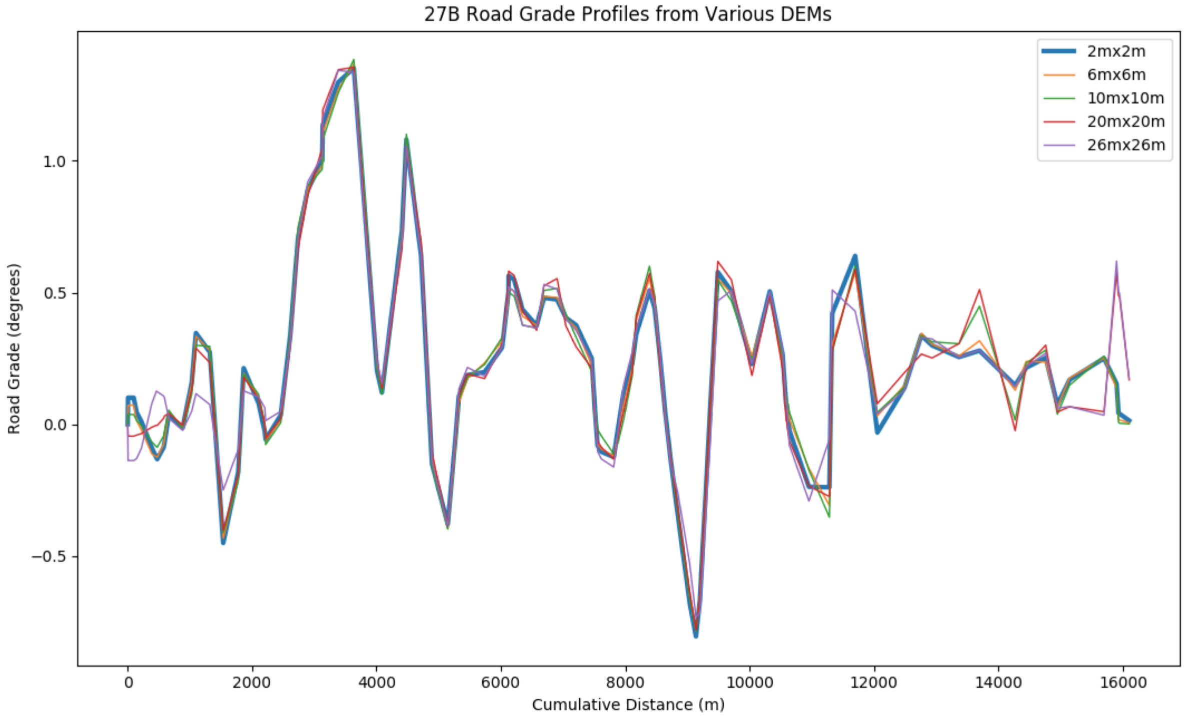


Figure 5.2: The road grade profiles generated for the 27B bus route from 2m^2, 6m^2, 10m^2, 20m^2 and 26m^2 DEMs.

5.3 Root Mean Square Error

5.3.1 27B Bus Route

The root mean square error of the road grade profiles generated from the 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile generated from the 2m^2 DEM for the 27B bus route can be seen in figure 5.3. Figure 5.3 serves not as an indication of RMSE across routes, but as an illustration of what the RMSE looks like for the road grade profiles in figure 5.2. RMSEs like those in figure 5.3 were calculated for each bus route.

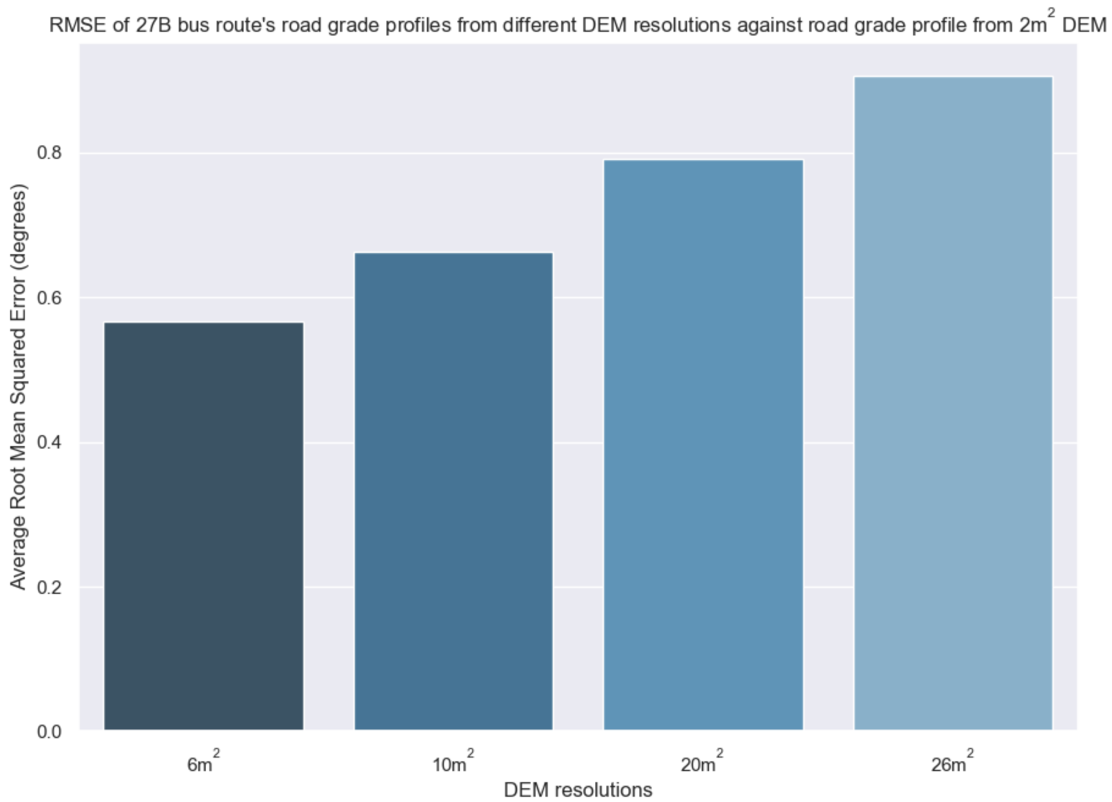


Figure 5.3: The root mean square error generated for the 27B bus route's road grade profiles from 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against its road grade profile from the 2m^2 DEM.

5.3.2 All Bus Routes

The average root mean square error of the road grade profiles generated from the 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile generated from the 2m^2 DEM for across the set of bus routes can be seen in figure 5.4.

It can be seen that the RMSE drops by no more than 0.3 degrees when increasing DEM resolution from 26m^2 to 10m^2. The road grade profile generated by the 6m^2 DEM has an

RMSE only 0.1 degrees less than that of the road grade profile generated by the $10m^2$ DEM. Overall, DEM resolution seems to have an insignificant effect on the RMSE of road grade profiles.

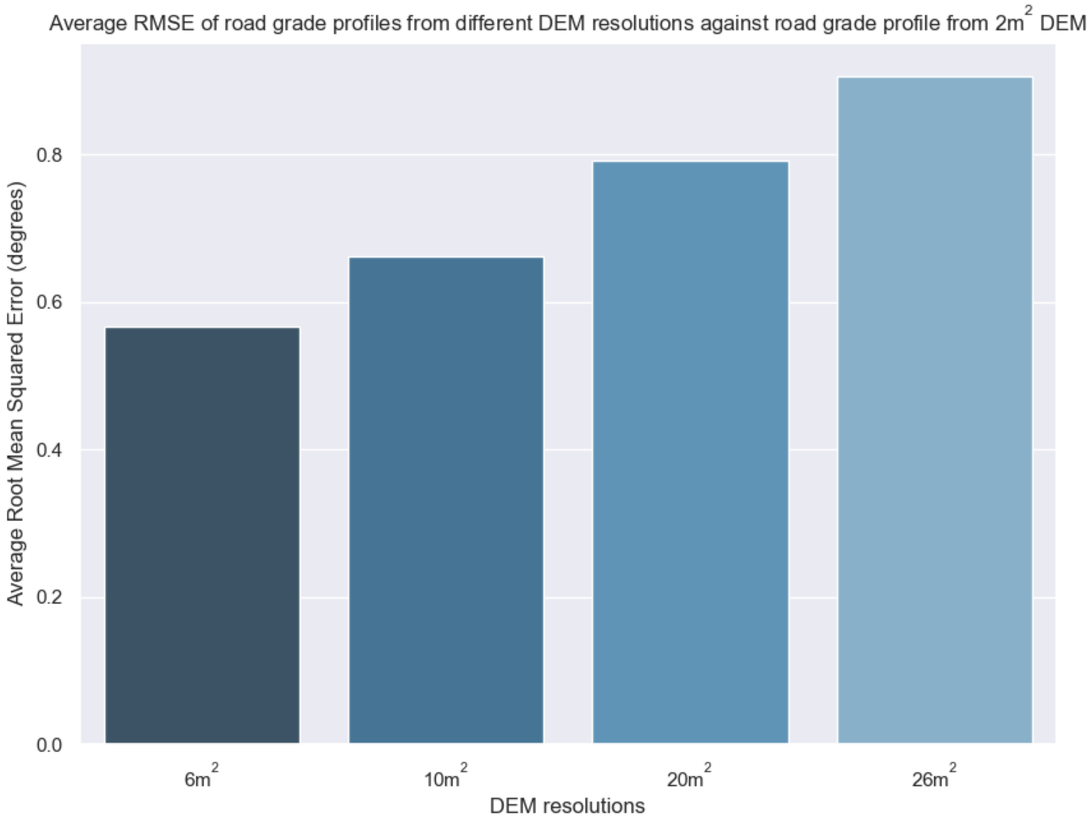


Figure 5.4: The average root mean square error across the bus routes for road grade profiles from $6m^2$, $10m^2$, $20m^2$ and $26m^2$ DEMs against the road grade profile from the $2m^2$ DEM.

5.4 Mean Absolute Error

5.4.1 27B Bus Route

The mean absolute error of the road grade profiles generated from the 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile generated from the 2m^2 DEM for the 27B bus route can be seen in figure 5.5. Figure 5.5 serves not as an indication of MAE across the routes, but as an illustration of what the MAE looks like for the road grade profiles in figure 5.2. MAEs like those in figure 5.5 were calculated for every bus route. It can be seen that the MAE of the road grade profiles scales approximately linearly with the resolution of the DEM used to generate them.

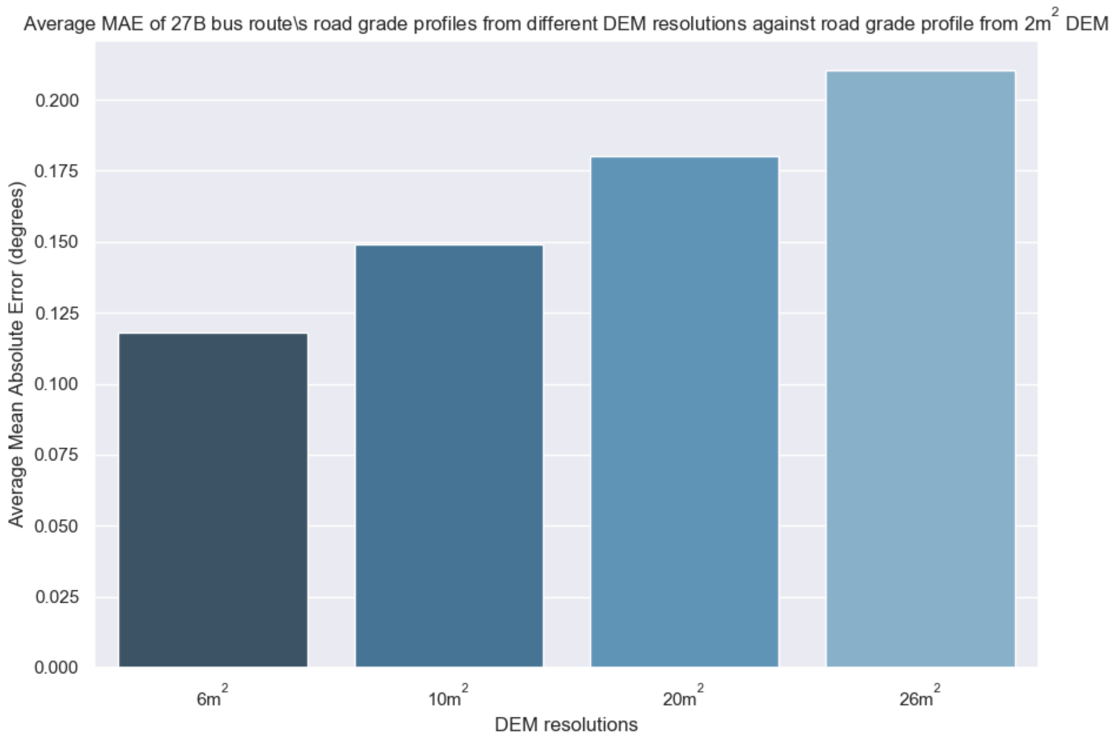


Figure 5.5: The mean absolute error generated for the 27B bus route's road grade profiles from 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile from the 2m^2 DEM.

5.4.2 All Bus Routes

The average mean average error of the road grade profiles generated from the 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile generated from the 2m^2 DEM for across the set of bus routes can be seen in figure 5.6.

It can be seen that the average MAE of the road grade profiles scales approximately linearly with the resolution of the DEM used to generate them. It can be seen that an increase of

DEM resolution from 26m^2 to 6m^2 only increased the average MAE of the road grade profiles by less than 0.1 degrees. The trend of the average MAE of road grade profiles across routes as DEM resolution gets higher is to decrease, as expected. However, the error in the road grade profiles only decreases very slightly as DEM resolution gets significantly higher.

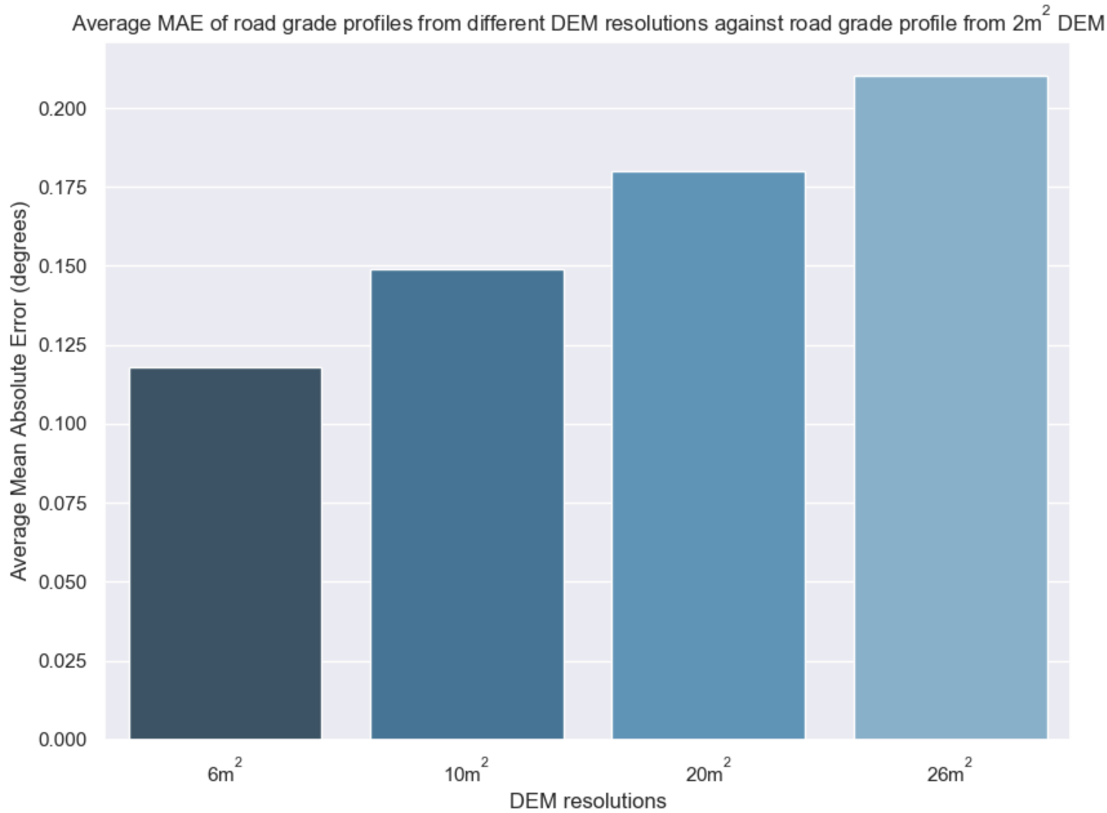


Figure 5.6: The average mean absolute error across routes for road grade profiles from 6m^2, 10m^2, 20m^2 and 26m^2 DEMs against the road grade profile from the 2m^2 DEM.

In the application of producing large scale road grade information, the results from the MAE of road grade profiles are more important than the RMSE. It is desirable that most roads are given as accurate road grade information as possible. Decreasing the MAE of the road grade information produced means that there will be fewer errors in the road grade information on average, with less emphasis on the size of the errors than in the RMSE. However, it is important to take both measures of error into account [41].

The results show that both the RMSE and MAE only decrease by less than 0.3 degrees when DEM resolution is increased from 26m^2 to 6m^2. This indicates that road grade profile information is likely already very accurate using low resolution freely available DEMs. As far as eco-routing is concerned, it may be advisable to look to improving the data used for other parameters rather than road grade, which will only be improved slightly with more accurate elevation data.

These findings contrast those of other similar works mentioned in Section 2.1, such as the works by Xun Shi et al. (2012) and Zhao et al. (2010), in that the impact of increasing DEM resolution was found to be relatively insignificant for the application area in this work.

The results may be heavily influenced by the resulting route profiles generated from cleaning the Dublin Bus data. Some GPS points of the final bus route profiles do not end up exactly on the road that they should, due to the averaging of multiple trips into one and the inconsistency of the GPS point recordings along trips. These imperfections of the route profiles may be essential in the resulting errors of the road grade profiles, and the relationship portrayed between DEM resolution and road grade profile accuracy. In Chapter 7, future work using perfect route profiles is discussed. It may also be of note that different results may arise should this experiment be performed in an area with a landscape dissimilar to that of County Dublin, Ireland.

Finally, though the small decrease in error afforded by higher resolution DEMs may be insignificant to most, it could be of significance to companies with large fleets of vehicles each of which make many journeys per day.

6 Conclusion

Road grade profiles were generated for all bus routes operating in the city of Dublin, Ireland using DEMs with resolutions of $2m^2$, $6m^2$, $10m^2$, $20m^2$ and $26m^2$. Taking the road grade profile generated by the $2m^2$ DEM as ground truth, the root mean square and mean average errors were calculated for each of the road grade profiles generated using the DEMs of lower resolutions.

It was found that generalised across all routes there was no significant trend in route mean square error in road grade profiles as the resolution of DEMs increased. The mean squared error of road grade profiles scaled approximately linearly with DEM resolution. The decrease in mean squared error across all bus routes was found to be very small.

The results indicate that using a DEM of a higher resolution rather than the freely available global DEM of the highest resolution only affords a very minimal increase in road grade profile accuracy.

7 Future Work

Research into the accuracy of these additional DEMs was not added to this work due to time constraints.

DEM using Google Elevation API

This work could be expanded on by also calculating the road grade profiles of the bus routes using a DEM created from Google Elevation API data. Google's elevation data is used by many applications, and was used as the ground truth in the research done by Singla et al. (2018), but has been shown to have some degree of error[47]. It would be of interest to see how the Google Elevation API data compares to a 2m^2 LiDAR DEM in generating road grade profiles.

SRTM-1 DEM

The road grade profiles of the bus routes could also be generated using the SRTM-1 DEM. SRTM-1 is the DEM that would be used to generate road grade data on a large scale, without acquiring data that is not freely available[1]. Because this is the actual data that is being used today, it is of interest to see how it compares to a 2m^2 LiDAR DEM in generating road grade profiles. It would also be of interest to see what differences there are between the road grade profiles generated by a 30m^2 DEM which is downsampled from a 2m^2 DEM and the SRTM-1 DEM (which also has a resolution of 30m^2).

Road grade profile generation from perfect route data

This experiment was performed using the state of the art pipeline for generating large scale road grade information. This pipeline which takes and cleans GPS traces from intelligent transport systems is prone to error, causing the eventual route profiles not to be 100% accurate. The route averaging step of the pipeline can result in some points in the route profile to be off the actual road by several metres. In the future, it may be possible to create or obtain more accurate representations of routes. In this case, it would be of interest to perform this experiment again with perfect representations of each route, where the route has as many GPS points which all occur on precisely the roads that the bus travels.

Bibliography

[1] National Aeronautics & Space Administration. U.S. Releases Enhanced Shuttle Land Elevation Data, . `https://www2.jpl.nasa.gov/srtm/`.

[2] Norberto Henriques and Carlos Bento. Integration of gps traces and digital elevation maps for improving bicycle traffic simulation behavior. In *Transportation Research Board 92nd Annual Meeting, Transportation Research Board*, 2013.

[3] Kanok Boriboonsomsin and Matthew Barth. Impacts of road grade on fuel consumption and carbon dioxide emissions evidenced by use of advanced navigation systems. *Transportation Research Record*, 2139(1):21–30, 2009.

[4] Min Zhou, Hui Jin, and Wenshuo Wang. A review of vehicle fuel consumption models to evaluate eco-driving and eco-routing. *Transportation Research Part D: Transport and Environment*, 49:203–218, 2016.

[5] Michael W Levin, Melissa Duell, and S Travis Waller. Effect of road grade on networkwide vehicle energy consumption and ecorouting. *Transportation Research Record*, 2427(1):26–33, 2014.

[6] Jared Magrath and Mike Brady. Evaluating different methods for determining road grade best suited to advanced bus transportation systems. In *2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)*, pages 108–113. IEEE, 2017.

[7] Behdad Yazdani Boroujeni, H Christopher Frey, and Gurdas Singh Sandhu. Road grade measurement using in-vehicle, stand-alone gps with barometric altimeter. *Journal of Transportation Engineering*, 139(6):605–611, 2013.

[8] Eric Wood, E Burton, A Duran, and J Gonder. Appending high-resolution elevation data to gps speed traces for vehicle energy modeling and simulation. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014.

[9] National Aeronautics & Space Administration. ASTER Global Digital Elevation Map Announcement, . `https://asterweb.jpl.nasa.gov/gdem.asp`.

[10] GISGeography. 5 Free Global DEM Data Sources – Digital Elevation Models, . https://gisgeography.com/free-global-dem-data-sources/.

[11] Tushti Singla, Mike Brady, and Jared Magrath. Development of city-wide road grade profiles utilising advanced bus transportation systems. *IFAC-PapersOnLine*, 51(9): 428–433, 2018.

[12] Xun Shi, Lyndsey Girod, Robert Long, Roger DeKett, Jessica Philippe, and Tom Burke. A comparison of lidar-based dems and usgs-sourced dems in terrain analysis for knowledge-based digital soil mapping. *Geoderma*, 170:217–226, 2012.

[13] Zhengyong Zhao, Glenn Benoy, Thien Lien Chow, Herb W Rees, Jean-Louis Daigle, and Fan-Rui Meng. Impacts of accuracy and resolution of conventional and lidar based dems on parameters used in hydrologic modeling. *Water resources management*, 24(7): 1363–1380, 2010.

[14] Paul NC Murphy, Jae Ogilvie, Fan-Rui Meng, and Paul Arp. Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification. *Hydrological Processes: An International Journal*, 22(12):1747–1754, 2008.

[15] Cody P Gillin, Scott W Bailey, Kevin J McGuire, and Stephen P Prisley. Evaluation of lidar-derived dems through terrain analysis and field comparison. *Photogrammetric Engineering & Remote Sensing*, 81(5):387–396, 2015.

[16] Anderson Luis Ruhoff, Nilza Maria Reis Castro, and Alfonso Risso. Numerical modelling of the topographic wetness index: an analysis at different scales. 2011.

[17] Google elevation api. Accessed April 2019. https://developers.google.com/maps/documentation/elevation/start.

[18] WGIV ISPRS. Characteristics of worldwide and nearly worldwide height models karsten jacobsen leibniz university hannover, institute of photogrammetry and geoinformation, germany; jacobsen@ ipi. uni-hannover. de. *BBC 26.17: 32.81 G 547*, (25):58, 2013.

[19] GISGeography. DEM, DSM & DTM Differences – A Look at Elevation Models in GIS, . https://gisgeography.com/dem-dsm-dtm-differences/.

[20] John A. Dutton David DiBiase. The nature of geographical information.

[21] Jamie Carter, Keil Schmid, Kirk Waters, L Betzhold, B Hadley, R Mataosky, and J Halleran. An introduction to lidar technology, data, and applications. *NOAA Coastal Services Center*, page 2, 2012.

[22] Xiaoye Liu. Airborne lidar for dem generation: some critical issues. *Progress in Physical Geography*, 32(1):31–49, 2008.

[23] Michael E Hodgson and Patrick Bresnahan. Accuracy of airborne lidar-derived elevation. *Photogrammetric Engineering & Remote Sensing*, 70(3):331–339, 2004.

[24] Jamie Carter, Keil Schmid, Kirk Waters, Lindy Betzhold, Brian Hadley, Rebecca Mataosky, and Jennifer Halleran. Lidar 101: An introduction to lidar technology, data, and applications. *National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center, Charleston, South Carolina, coast. noaa. gov/digitalcoast/_/pdf/lidar101. pdf. Last accessed December*, 30:2015, 2012.

[25] NYU Spatial Data Repository. 2015 Aerial Laser and Photogrammetry Survey of Dublin City Collection Record. https://doi.org/10.17609/N8MQ0N.

[26] Geokov. Map Projections - types and distortion patterns. http://geokov.com/Education/map-projection.aspx.

[27] OSi / OSNI. New Map Projections for Ireland - Consultation Paper. https://www.osi.ie/wp-content/uploads/2015/04/New-Map-Projections-for-Ireland-Consultation-Paper.pdf.

[28] epsg.io. IRENET95 / Irish Transverse Mercator. https://epsg.io/2157.

[29] Open Street Maps. About OpenStreetMap, . https://wiki.openstreetmap.org/wiki/About_OpenStreetMap.

[30] Open Street Maps. OpenStreetMap Statistics, . https://www.openstreetmap.org/stats/data_stats.html.

[31] Steven W Smith et al. The scientist and engineer's guide to digital signal processing. 1997.

[32] Ronald W Schafer et al. What is a savitzky-golay filter. *IEEE Signal processing magazine*, 28(4):111–117, 2011.

[33] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.

[34] Matthew Aubury and Wayne Luk. Binomial filters. *Journal of VLSI signal processing systems for signal, image and video technology*, 12(1):35–50, 1996.

[35] GISGeography. Bilinear Interpolation: Resample Image Cell Size with 4 Nearest Neighbors, . https://gisgeography.com/bilinear-interpolation-resampling/.

[36] Movable Type Scripts. Calculate distance, bearing and more between Latitude/Longitude points. https://www.movable-type.co.uk/scripts/latlong.html.

[37] Thaddeus Vincenty. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey review*, 23(176):88–93, 1975.

[38] Benny Dwi Kifana and Maman Abdurohman. Great circle distance methode for improving operational control system based on gps tracking system. *International Journal on Computer Science and Engineering*, 4(4):647, 2012.

[39] Hagar Mahmoud and Nadine Akkari. Shortest path calculation: a comparative study for location-based recommender system. In *2016 World Symposium on Computer Applications & Research (WSCAR)*, pages 1–5. IEEE, 2016.

[40] Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82, 2005.

[41] Tianfeng Chai and Roland R Draxler. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific model development*, 7(3):1247–1250, 2014.

[42] Anthony G Barnston. Correspondence among the correlation, rmse, and heidke forecast verification measures; refinement of the heidke score. *Weather and Forecasting*, 7(4): 699–709, 1992.

[43] DATA.GOV.IE. Dublin Bus GPS sample data from Dublin City Council . `https://data.gov.ie/dataset/dublin-bus-gps-sample-data-from-dublin-city-council-insight-project`.

[44] Open Street Maps. Overpass API - OpenStreetMap, . `https://wiki.openstreetmap.org/wiki/Overpass_API`.

[45] Open Street Maps. Overpass API/Language Guidep, . `https://wiki.openstreetmap.org/wiki/Overpass_API/Language_Guide#Special_clauses:_around_and_poly`.

[46] Dr Tom Farr & Dr Robert Crippen (NASA Jet Propulsion Laboratory). Personal Communication.

[47] Yinsong Wang, Yajie Zou, Kristian Henrickson, Yinhai Wang, Jinjun Tang, and Byung-Jung Park. Google earth elevation data extraction and accuracy assessment for transportation applications. *PloS one*, 12(4):e0175756, 2017.