REDUCTION
2011-2014

Deliverable 3.3

Prototype Basic Eco-Route Prediction

11th of December 2012

Public Document

Project acronym: REDUCTION

Project full title: Reducing Environmental Footprint based on Multi-Modal Fleet management Systems for Eco-Routing and Driver Behavior Adaptation

Work Package:	**3**
Document title:	Prototype Basic Eco-Route Prediction
Version:	1.2
Official delivery date:	2012.09.01
Actual publication date:	2012.09.01
Type of document:	Report
Nature:	Public
Authors:	Ove Andersen(AAU), Benjamin B. Krogh(AAU), Kristian Torp (AAU) Christian S. Jensen (AU), Bin Yang (AU) Nick Christiansen (FD) Niels T. Larsen (FD)
Approved by:	____________________

Version	**Date**	**Sections Affected**
0.1	2012-06-01	First version.
0.2	2012-07-13	Version circulated to WP3 group.
0.3	2012-08-14	Implemented various first round comments. Updated coverage figures Fixed TODO + bad references Add information on high-frequent GNSS with CANBus
04	2012-08-17	Comments from AU integrated.
1.0	2012-08-30	Minor changes before first public release.
1.1	2012-12-11	Updated to use new template + fixed comments from 1. Year review.
1.2	2014-01-31	Comments after second year review. • Updated with internal review comments

Executive Summary

This report describes the prototype that allows for the basic and simple eco-route prediction. First the report describes the different data-source types available to the REDUCTION project. These types consist of combinations of GNSS measurements and CANBus data. The usefulness and the availability of the different data-source types are discussed in detail. The conclusion is that at present there are a sufficient number of GNSS measurements but there is a lack of CANBus data. To alleviate this lack of data a weighted combination approach is proposed where all data available is used to its fullest. Next, algorithms on how to convert each of the different data-source types to eco-route predications are presented. Having data for all roads on a large map are unrealistic and it is therefore described how to deal with regions on a map where there is insufficient data. Finally, the results/output of the prototype is described. These outputs are; a complete digital map with both travel-time and eco-route prediction that is well-suited for a larger IT organizations and web interface that can be accessed by the general public.

Table of Contents

List of Figures

List of Tables

1. Introduction

This section identifies the high-level objectives and targets of the project in Section 1.1, followed by the objectives and contributions of work package WP3 in Section 1.2. Section 1.3 lists the objectives and contributions of this deliverable.

1.1 Project Overview

Reduction of CO2 emissions is a great challenge for the transport sector nowadays. Despite recent progress in vehicle manufacturing and fuel technology, still a significant fraction of CO2 emissions in EU cities is resulting from vehicular transportation. Therefore, additional innovative technologies are needed to address the challenge of reducing emissions. The REDUCTION project focuses on advanced ICT solutions for managing multi-modal fleets and reducing their environmental footprint. REDUCTION collects historic and real-time data about driving behavior, routing information, and emissions measurements, that are processed by advanced predictive analytics to enable fleets enhancing their current services as follows:

1) Optimizing driving behavior: supporting effective decision making for the enhancement of drivers' education and the formation of effective policies about optimal traffic operations (speeding, braking, etc.), based on the analytical results over the data that associate driving-behavior patterns with CO2 emissions;

2) Eco-routing: suggesting environmental-friendly routes and allowing multi-modal lets to reduce their overall mileage automatically; and

3) Support for multi-modality: offering a transparent way to support multiple transportation modes and enabling co-modality.

REDUCTION follows an interdisciplinary approach and brings together expertise from several communities. Its innovative, decentralized architecture allows scalability to large fleets by combining both V2V and V2I approaches. Its planned commercial exploitation, based on its proposed cutting edge technology, aims at providing a major breakthrough in the fast growing market of services for "green" fleets in EU and worldwide, and present substantial impact to the challenging environmental goals of EU.

1.2 Work Package Objectives and Tasks

The objective of WP3 is to design and develop a software prototype that can convert vehicle-related data, primarily GPS data, to metrics that capture environmental impact. The prototype must handle very large volumes of data from different types of vehicles and must efficiently compute the multi-modal eco-routes in both real-time and offline modes. In addition, the prototype must be able to report on the temporal evolution of eco-routes, e.g., due to a variety of changes in the transportation infrastructure and its use. The work package will (a) define the interfaces for how vehicles communicate with the server side and with each other; (b) invent and prototype techniques for computing eco-routes; (c) invent and prototype techniques for the validation of eco-routes; and (d) design and prototype high-performance data structures and algorithms for the handling of very large volumes of streaming data from the vehicles.

1.3 Objective of this Deliverable

This section describes the objectivities of this deliverable. First the task from the Description of Work (DOW) is listed and then the general idea on how to make eco-route predictions is presented.

1.3.1 Task from DOW

The basic techniques for computing eco-routes are invented and implemented in this task. Particular focus is given to how to "convert" the data available into existing metrics for the capture of environmental impact, e.g., greenhouse gas emissions. A central challenge is to invent techniques for evaluating and validating the results provided by the eco-routing techniques proposed. Here the use of fuel consumption and other vehicle data is expected to be used. This evaluation and validation must support the field trials carried out in WP5. The focus in this task is on ensuring that eco-routing techniques invented are conceptually sound, correct, and offer adequately and predictably accurate environmental impact estimates. Extensive experimental studies are expected where alternative routing and evaluation techniques are assessed and improved. Eco-routing is closely related to map data. It is therefore necessary to provide output formats that are easy to integrate with existing mapping solutions so that the eco-routes can be easily included in reports to transport authorities or offered as feedback to the individual drivers.

1.3.2 The General Idea

The purpose of the basic eco-route prediction prototype is to provide a basic software service that allows for the experimentation of using eco-routes in the scheduling of flex-traffic. A foundation for building the eco-route estimation software is that there is a well-tested software component that can estimate travel time. The reason that good travel-time estimation is needed is that the travel time is much simpler to verify than the eco-route estimation, i.e., a stop watch can be used. Thus if the travel-time is accurately estimated it is assumed that the eco-route is also accurately estimated.

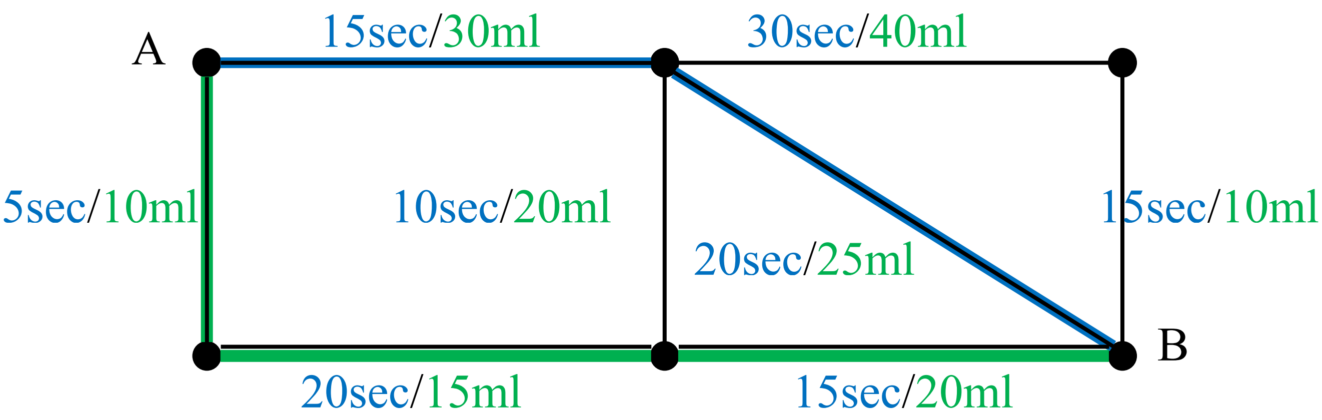


Figure 1 Basic Idea Travel-Time and Eco-route Estimation

The basic idea in both the travel-time and eco-route estimation is shown in Figure 1 where a small part of a digital map is shown. The black dots are nodes and the black lines are edges in the map as used in many map technologies, e.g., the de-facto standard Shapefile [1] . The blue text indicates the travel-time for passing the edge (a road segment on a map) and the green texts indicates the estimated fuel consumption for passing the edge. The blue lines show the fastest routes between the points A and B (35 seconds and fuel consumption of 55 ml). The green line indicates the best eco-route between the points A and B (40 seconds and 45 ml of fuel).

1.3.3 Summary Table

Objective	How Meet
Convert available data to metrics	• Convert low-frequent data to eco weights • Convert high-frequent data to eco weights • Combine weights from GNSS data with real values from CANBus data
Data foundation for entire map	• Find similar road segments • For road segments with insufficient data borrow eco weights from neighbors • Create map with time and eco-weights for all road segments
Prototype	• Implementation of basic prototype • Release prototype to users (taxi drivers) for initial feedback and evaluation

Table 1 Objectivities of Deliverable and how meet

2. Glossary

This section lists the central terms and abbreviations used in the document

Term	Description
CANBus data	Detailed information about a single vehicle, e.g., actual fuel consumption, engine RPM, and throttle position.
Flex-Traffic	On-demand collective transportation using smaller vehicles. Routes are according to passenger requests and there is no fixed timetable.
GHG	Greenhouse gas, such as CO_2 and NO_x
GNSS	Global Navigation Satellite System, e.g., GPS, GLONAS, Compass, or Galileo.
GNSS measurement	The recording of vehicle ID, longitude, latitude, speed, and other relevant data for a single vehicle using, e.g., a GPS device.
High-frequent GNSS data	GNSS data received from a vehicle with a sufficiently high frequency, e.g., every second such that the actual route followed by the vehicle can be determined.
Low-frequent GNSS data	GNSS data received from a vehicle with a low frequency, e.g., every 60 seconds such that the actual route of the vehicle cannot be determined.
Map matching	The process of mapping latitude and longitude coordinates received from a GNSS device to locations in a digital road network. Necessary due to the inherent imprecision of GNSS equipment.
Shapefile	De-facto file format for digital maps.
System	The complete software described in this document, i.e., the software prototype developed in the work package 3.
Trip	The path taken by a single vehicle in a road network. The trip consists of a start point, a destination, and planned stops along the path.
Vehicle type	The type of vehicle such as car, mini-bus, bus, truck, or train

Table 2 Terms used in Document

3. Data-Source Types

A number of different data sources are available to the REDUCTION project. In this section, these data types are classified and it is estimated how good each data-source type is in estimating both travel-time and eco-routes.

3.1 GNSS Measurements

A large set of GNSS measurements are available. These can be classified into two categories.

- Low-frequent
- High-frequent

The difference between the two categories is whether consecutive GNSS measurements from a single vehicle can be grouped into a trip or not, i.e., can the actual route that the vehicle has been driving be determined using map-matching technology [2] .

Please note that there is no fixed limit between when GNSS measurements are low-frequent or high-frequent. Experimentation using the M-GEMMA map-matching algorithm shows [3] that if there are more than 9 seconds between consecutive GNSS measurements from a vehicle this particular map-matching algorithm cannot create a trip. Since the M-GEMMA map-matching algorithm is the only open-source map-matching algorithm available (to the best of our knowledge) this is the current border between the two categories of GNSS measurements.

The minimal requirements to the GNSS measurements in the REDUCTION project are documented in [4].

3.2 CANBus Data

For CANBus data the situation is simpler than for GNSS measurement, because either a vehicle collects CANBus data as it is driving or it does not. Please note that at the-time-of-writing the exact format of the CANBus data that will be supplied to the REDUCTION project is not exactly known.

3.3 Combining GNSS Measurements and CANBus Data

The availability of GNSS measurements is independent of the availability of CANBus data; however, CANBus data without a corresponding GNSS measurement cannot be used. It is therefore possible to combine GNSS measurements with CANBus data as shown in Table 3. In this table, the expected usefulness of each data-source type for both travel-time and eco-route estimated is also listed, using the scale {very low, low, medium, high, very high}. The reason why the expected usefulness is listed here should become clear after having read Section 3.4 below.

Data-Source Type	Travel-Time	Eco-Route
Low Frequent	Medium	low
Low Frequent + CANBus	Medium	High
High Frequent	Very high	Medium
High Frequent + CANBus	Very high	Very high

Table 3 The Data-Source Types and the Expected Usefulness for Travel-time and Eco-Route Estimation

The low-frequent GNSS measurements are evaluated as medium useful as a data foundation for estimating travel-time. This does not mean that low-frequent GNSS measurements are unsuited for computing travel-time it only means that large number of GNSS measurements must be available to estimated travel-time with a sufficient accuracy.

Low-frequent GNSS measurements are considered low with respect to being used for estimating eco-routes. The reason why this data-source is being used to estimated eco-route is that the REDUCITION project as access to a very large set of these GNSS measurements.

The combination of low-frequent GNSS measurements and CANBus data is expected to be high in the usage for estimating eco-routes. The reason is that the actual fuel consumption of vehicles can be matched to road segments and aggregated as shown in Figure 1.

High-frequent GNSS measurements are expected to the best usefulness for estimating travel time. The reason is that with this data-source type it is possible to reconstruct the actual route driven by the vehicle. This allows for very precisely estimations of how long it takes to driver from Point A to Point B as shown in Figure 1 because the time used to drive blue route can be deduced from the high-frequency GNSS measurements.

High-frequent GNSS measurements are expected to be medium useful for estimating eco-route. The reason is that the speed of the vehicle can be directly read from the GNSS measurement and that the accelerations can be deduced with a high-degree of accuracy. Having both the speed and the acceleration a number of existing approaches exists for converting GNSS measurements into estimations of fuel consumption [5] [6] [7].

If CANBus data is associated with high-frequency GNSS measurements this is considered to have the very best usefulness for estimating eco-routes. The reason is that the CANBus data report this actual fuel consumption and by using the GNSS measurements the can be map-matched. This data is considered the ground-truth for fuel consumption! [4]

3.4 Data Availability

In the REDUCTION project there is only a very limited budget for installing new equipment in vehicles in the range of 5-20 vehicles. This is way too small a fleet to get sufficient data for conducting the BeKTra/FlexDanmark field trials in Task 5.2 and Task 5.5. For these field trials the REDUCTION project must therefore rely on external sources for both GNSS measurements and CANBus data.

Please note that the following discussion of data availability only considers Denmark as this is where BeKTra/FlexDanmark schedules flex-traffic. We would like to stretch that there is absolutely nothing that prevents the solutions presented to be used anywhere in the World if a digital map, GNSS measurements, and CANBus data is available.

The estimated data size available at the time of writing (2012.09.01) and the expected data sizes in a year from now (2013.09.01) are shown in Table 4 (previous scale is reused). This corresponds to the expected data size that will be available for the first and second BeKTra/FlexDanmark field trial, respectively.

Data-Source Type	2012.09.01	2013.09.01
Low Frequent	Very high	Very high
Low Frequent + CANBus	low	Medium
High Frequent	Low	Medium
High Frequent + CANBus	Very low	Low

Table 4: Expected Data Sizes available for BeKTra Field Trials now and in a Year

As can be seen a very large sets of low-frequent GNSS measurements are already available approximately 1,200 million from approximately 10,000 vehicles. The expected growth is approximately 5 million new GNSS measurements per week. It is the vehicles used to execute the schedules prepared by BeKTra/FlexDanmark that supply this data and it is therefore a solid data source. The distribution of this data for Denmark is shown in Figure 2. As can be seen from the figure there are most road segments have more than 10,000 GNSS measurements and the entire main road network is well-covered.

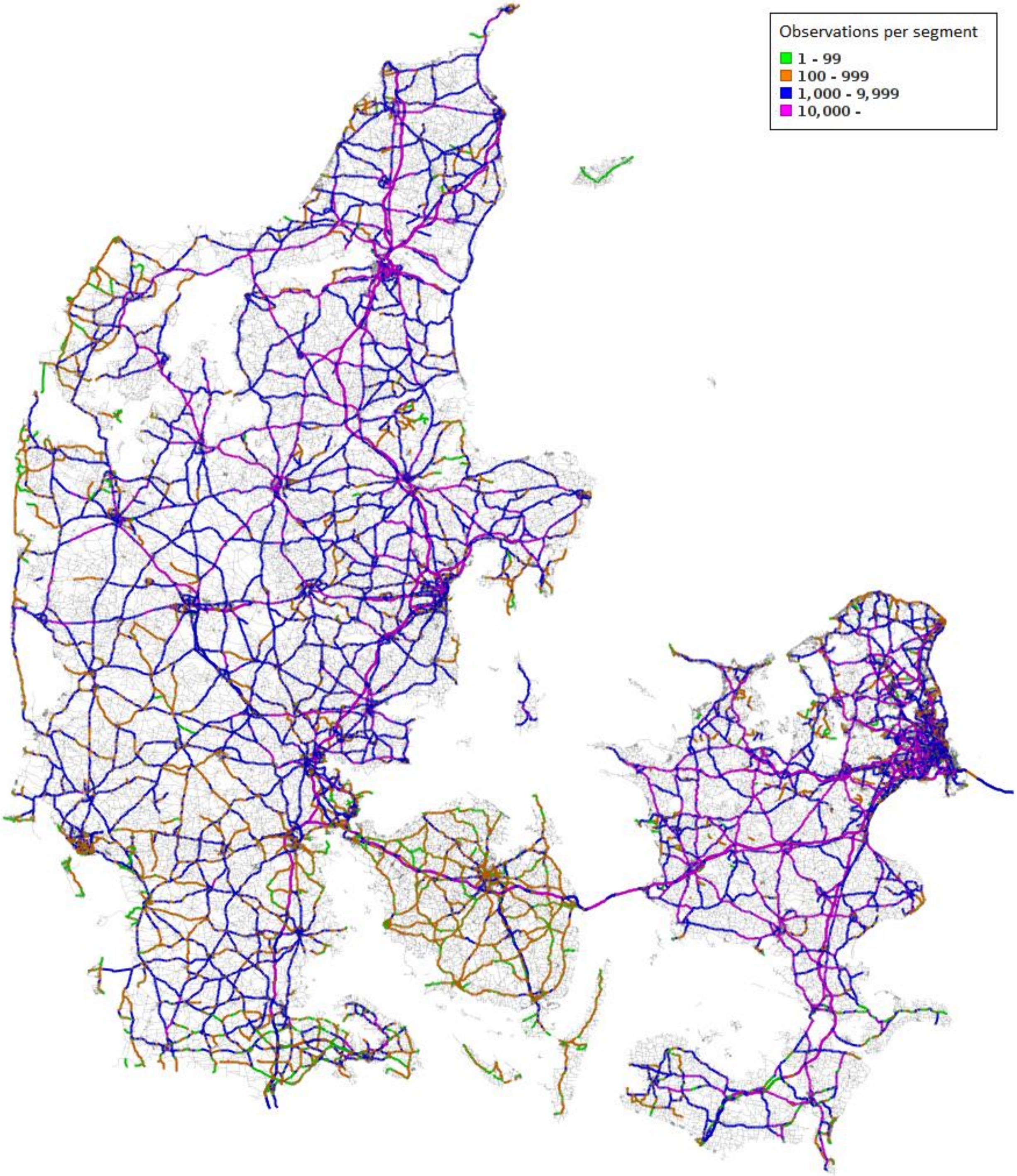


Figure 2: Distribution of Low-Frequent GNSS Data for Denmark (Main Road Network)

The data size for low-frequent GNSS measurements with CANBus data is currently low. Currently approximately 14 million GNSS measurements with corresponding CANBus data are available. This is on the main road network as shown in Figure 3. The reason that only a low number of low-frequent GNSS measurements with CANBus data are available is that the collection of this data first started in 2012 with a limited set of vehicles.

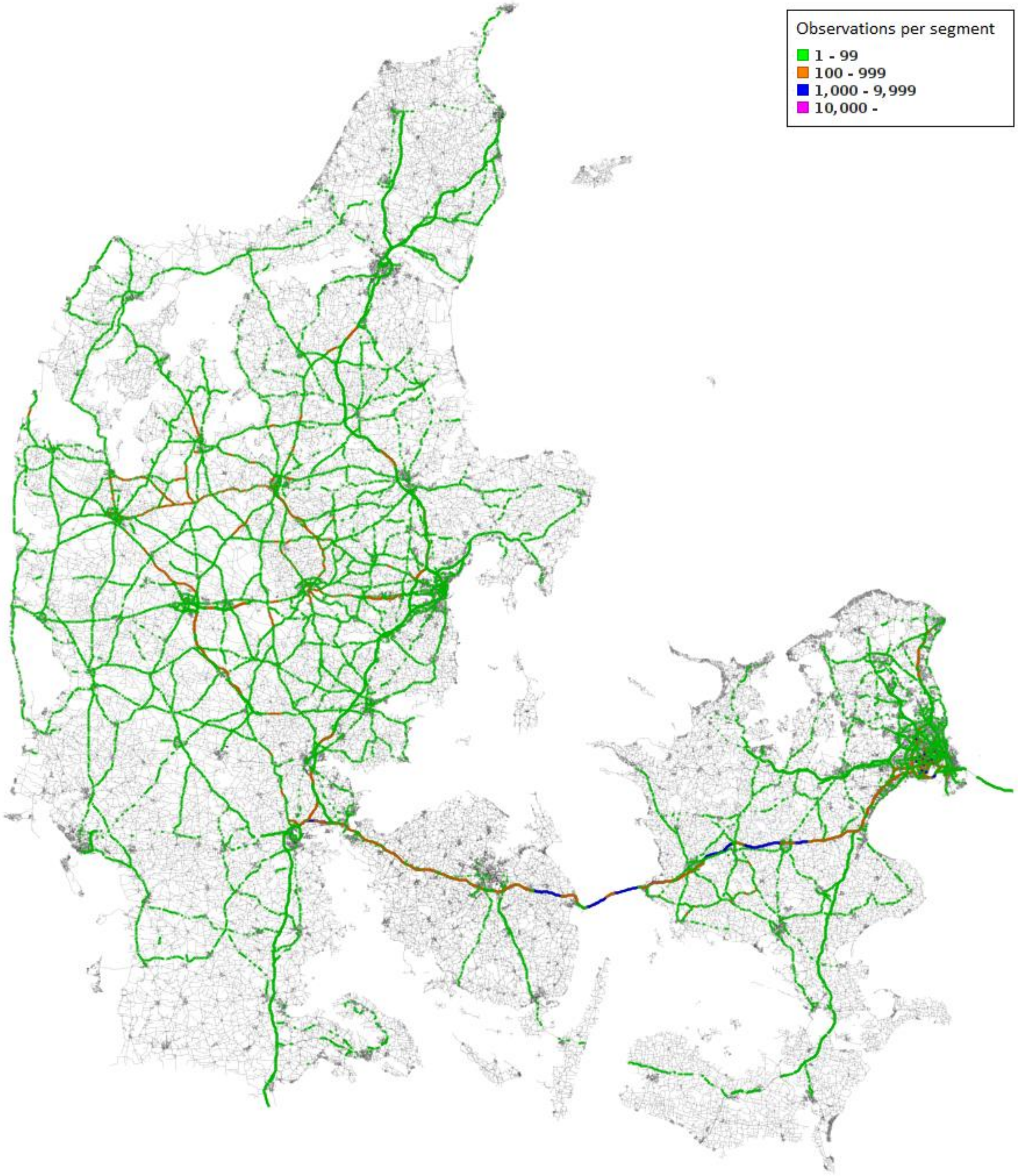


Figure 3: Distribution of Low-Frequent GNSS Data with CANBus Data (for Denmark)

The low-frequent GNSS measurements with CANBus data is supplied by a third party and is considered a stable data source with an expected growth

The rate of growth of number of low-frequent GNSS measurements with CANBus data is approximately 50,000 measurements per week.

The data size for high-frequent GNSS measurement is currently estimated to be low. The REDUCTION project has access to approximately 370 million GNSS measurements from 120 vehicles with a total of approximately 500,000 trips. This data is from a concluded project "Pay as You Speed" [8]. The distribution of high-frequency GNSS data in Denmark is shown in Figure 4.

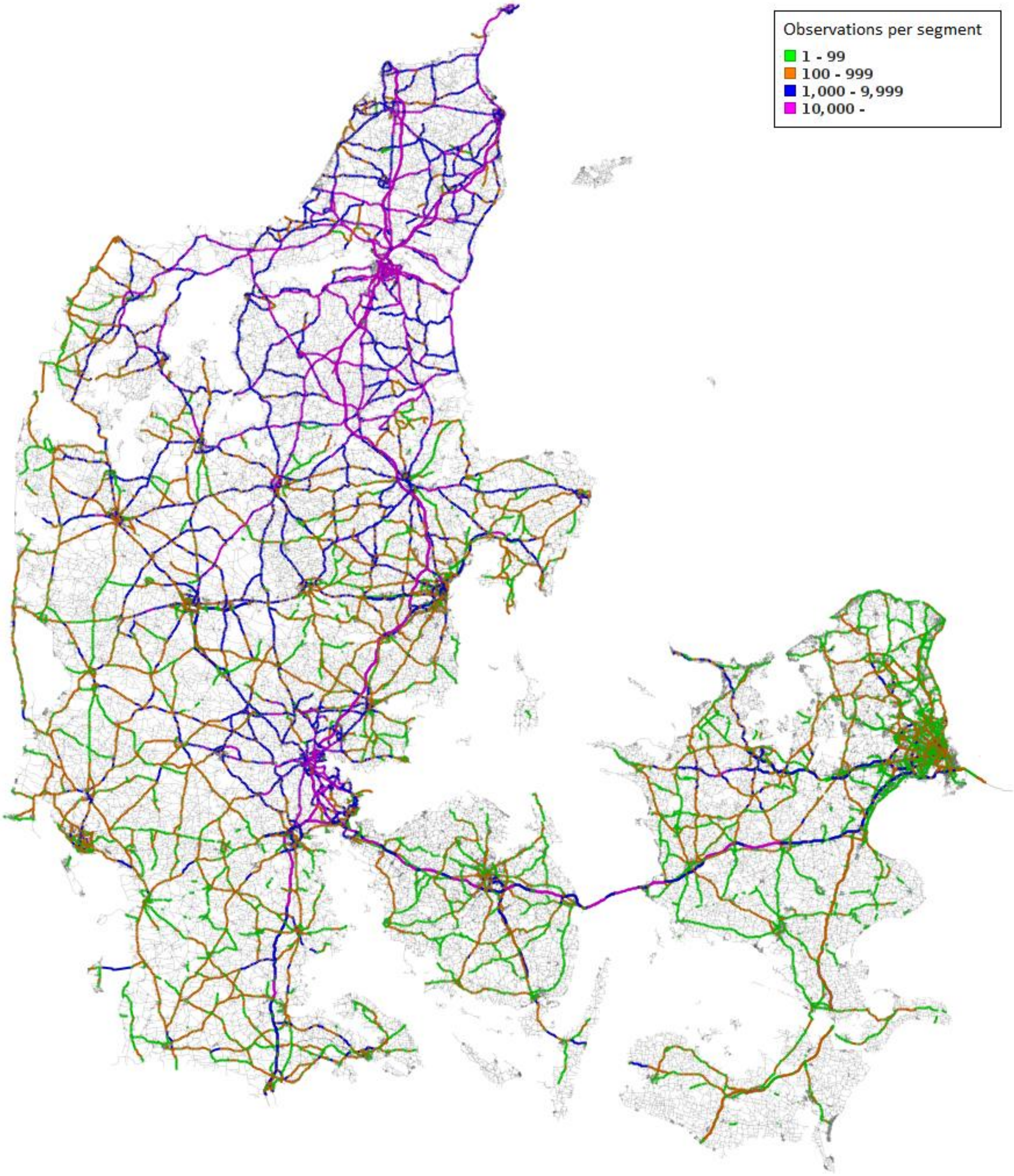


Figure 4: Distribution of High-Frequency GNSS Measurements for Denmark

Even though there are 370 million GNSS measurements are available it is considered a low number of trips where large parts of the road network are not covered at all (the gray lines in Figure 4).

Further the distribution of data is skewed with most data in North Jutland (were the “Pay as You Speed” project was mainly conducted).

A single source of high-frequency GNSS measurements [9] has been identified. The data size that can be provided is at the time of writing unclear. However, getting access to the data is possible for the BeKTra/FlexDanmark second field trial.

The data-source type high-frequent GNSS measurements with CANBus data is currently estimated as very low. A single source of providing this data has been identified. It is the same source as for the low-frequency GNSS measurement with CANBus data and the source is therefore considered very reliable. This data-source type will first be available for the second field trial. The high-frequent GNSS measurements with CANBus data that is currently available can be seen in Figure 5.

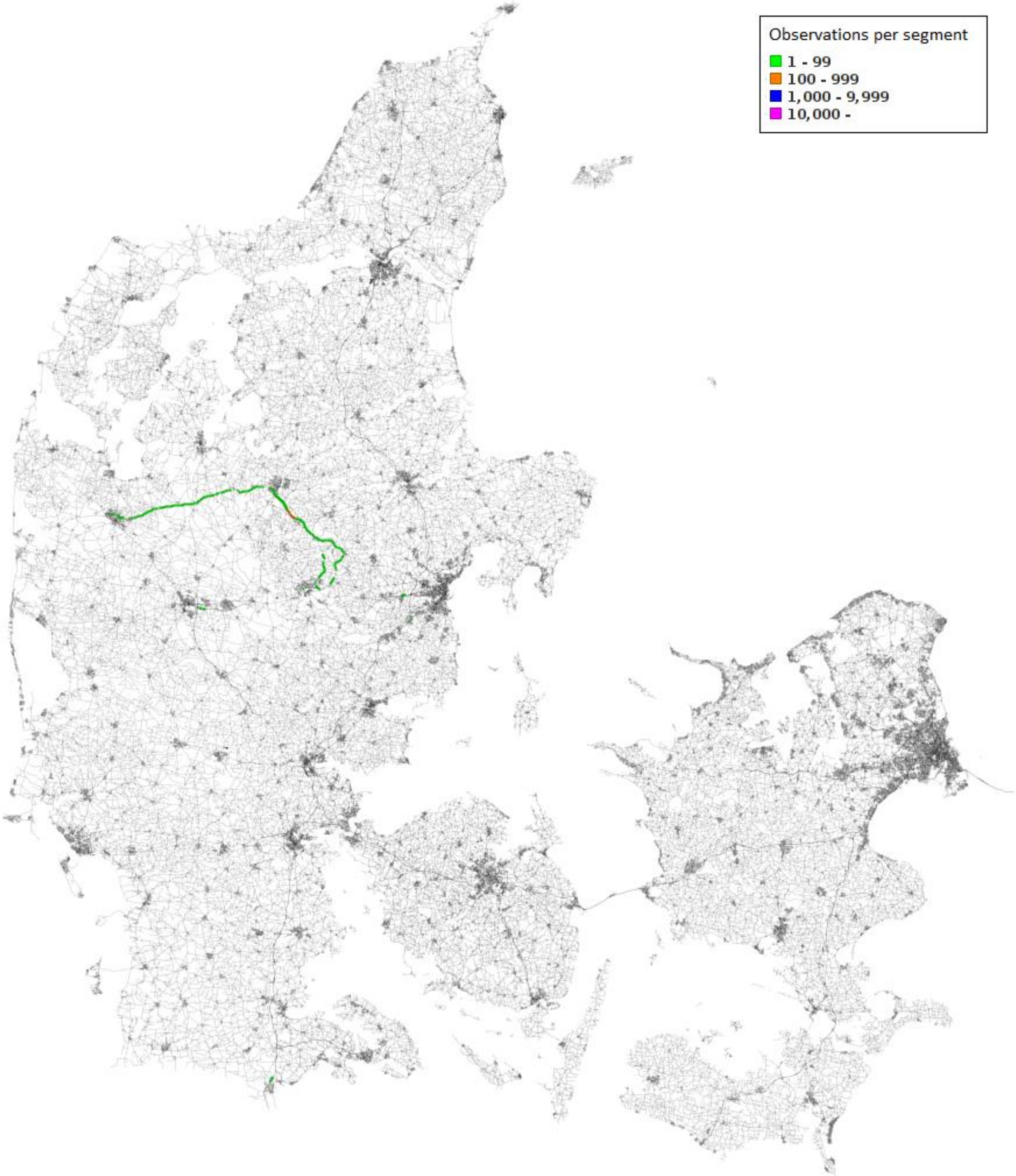


Figure 5 Distribution of High-Frequent GNSS Measurements with CANBus data for Denmark

4. Algorithms

The previous section has described the data-source types available. From here it should be clear that getting access to data is a major hurdle and that the basic eco-route approach therefor should make an effort in using all of the available data. More generally, there is a boots-trap problem when eco-routes must be estimated because it is very data intensive.

In this section, it will be described how each of the data-source types can contribute to estimating eco-routes and in the last subsection it is shown how these methods can be combined. Table 5 provided an overview of how each data-source type is converted to make the data useful for estimating eco-routes. Each of the three approaches is discussed in details in the following subsections. Please note that the same approach is used for data-source types where CANBus data is available.

Data-Source Type	Approach
Low Frequent	Time-to-Eco
Low Frequent + CANBus	CANBus-to-Eco
High Frequent	Trajectory-to-Eco
High Frequent + CANBus	CANBus-to-Eco

Table 5 Approaches used for the Different Data-Source Type

4.1 Time-to-Eco

The idea in the Time-to-Eco algorithm is to use the very large quantities of low-frequent GNSS measurements that are available to the REDUCTION project, see Table 4. This data-source type is well-suited to be used for estimating travel-time as indicated in Table 3 but not as well-suited for estimating eco-routes. However, the large data size makes this data-source type very attractable to use also eco-routing and therefore this subsection proposes an algorithm for converting travel-time to an estimation of fuel consumption.

The data that is available for converting travel-time to eco-route on a single edge (segment) as shown in Figure 1 is the following.

- The length of the segment
- The travel-time on the segment

To convert the travel-time to an eco-route the following must be estimated.

- The fuel economy of the vehicle

- The fuel consumption at given speed.

An example of a graph to convert from travel-time to fuel consumption is shown in Figure 6. How this graph is created is discussed in the next subsection. The x-axis is the speed of the vehicle in km/h the y-axis is an indicator of the fuel consumption. The y-axis has no unit [10] but is a metric such that 50 is half the fuel consumption of 100.

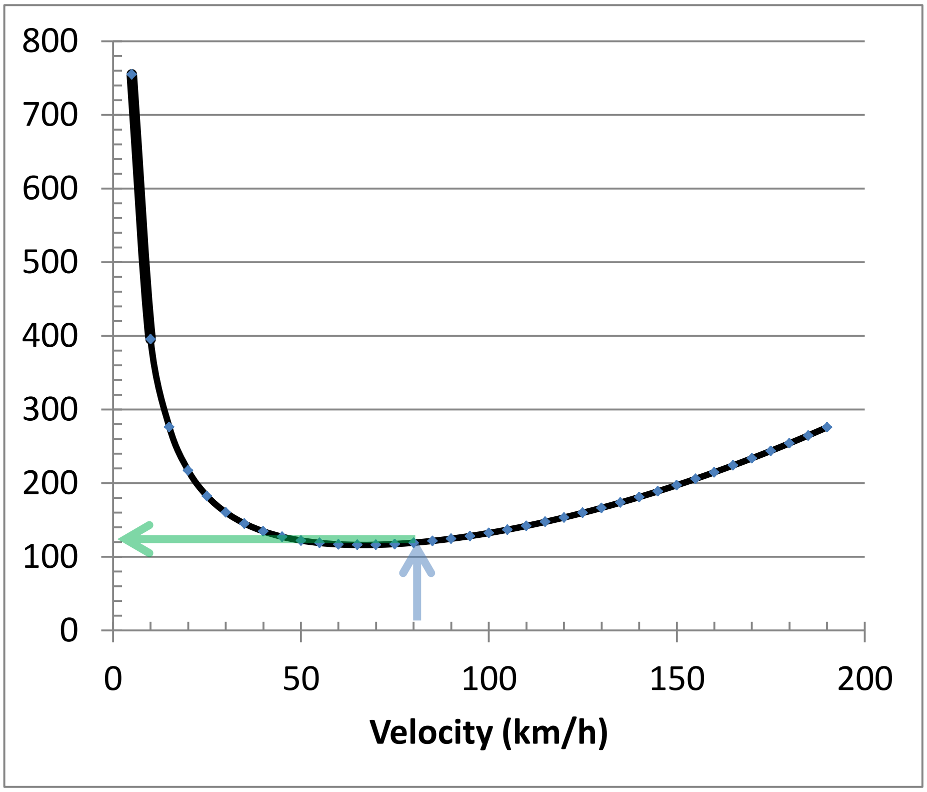


Figure 6 Converting Travel-Time to Fuel Consumption X-axis Speed, Y-Axis Fuel Consumption

In Figure 6, the thick blue and green arrows show how the conversion from travel-time to eco-route is implemented. The speed of the vehicle is known from the GNSS measurement. In the example, the vehicle is drives with 80 km/h. From the x-axis in Figure 6 it is found what the expected fuel consumption is at 80 km/h by going via first the blue arrow and then the green arrow to the y-axis that indicates the fuel consumption at this speed, which is approximately 120.

The Time-to-Eco algorithm is assumed not to be as accurate as the other algorithms described in this report and implemented in the software prototype. A main reason to the inaccuracy is that the acceleration of the vehicle is unknown. In a later section is this report it is shown how to deal with the assumed inaccuracy of Time-to-Eco algorithm by using a weighted combination.

4.2 Trajectory-to-Eco

The basic idea in the Trajectory-to-Eco approach is the base the estimation of eco-routes on the speed and acceleration of a vehicle. A small part of a trajectory for single vehicle is shown in Table 2. In this example, the GNSS measurements are sample with 1Hz, other frequencies may be used.

ID	Vehicle ID	Latitude	Longitude	Time	Velocity
1	42	57,0	9,9	12:00:00	80
2	42	57,0	9,9	12:00:01	80
3	42	57,0	9,9	12:00:02	70
4	42	57,0	9,9	12:00:03	75
5	42	57,0	9,9	12:00:04	80

Table 6 Trajectory of GNSS Measurement for a Single Vehicle

The speed of the vehicle is directly available from the data in Table 6; however the acceleration must be estimated. To do this estimation the GNSS measurements are pair wise compared. This is illustrated in Table 7 where the first row compares rows one and two from Table 6, the second row in Table 7 compares rows two and three from Table 6, and so on.

Id/Id	Duration	Velocity (km/h)	Acc. (km/h/s)
1/2	1	80	0
2/3	1	70	0
3/4	1	75	5
4/5	1	80	5

Table 7: Pairwise Delta for GNSS Measurements

Please note that in Table 7 the duration is recorded this makes it possible to handle different sampling frequencies, e.g., due to different data sources or due to drop-outs in the GNSS signals.

From Table 6 and Table 7 the speed and acceleration of the vehicle are now available and existing approaches to converting trajectory data to estimation of fuel consumptions can be used. The simplest of these approaches is documented in [10] [11] where an index of the fuel consumption is estimated as follows.

$$consumption = v * (1{,}1 * a + 9{,}81 * h + 0.132) * 0{,}000302 * v^3$$

Where v is the current speed (m/s), a is the acceleration is m/s^{2,} and h is the gradient of the road in degrees. The factor h is unknown and the formula is simplified to

$$consumption = v * (1{,}1 * a + 0.132) * 0{,}000302 * v^3$$

The index of the fuel consumption must be normalized according to the type of vehicle. The normalization is tested using three different vehicles in [10] and the generic formula below is suggested.

$$normaliseret\ consumption = \begin{cases} 0{,}264 * consumption + 1 & if\ v > 0 \\ 1 & otherwise \end{cases}$$

The result of using the normalized fuel consumption is shown in Figure 7. The graph shows the fuel consumption for driving one kilometer with constant speed for speeds between 0 to 200 km/h (the x-axis. Again the y-axis is the relative fuel consumption, i.e., without a unit. The graph shows that it is very expensive to drive one kilometer with 25 km/h compared to driving one kilometer with 80 km/h. It also shows that it becomes more and more expensive as the speed exceeds approximately 90 km/h. The graph to the right in Figure 7 shows the fuel consumption when accelerating from 50 km/h or 130 km/h. It requires much more fuel to accelerate when driving with 130 km/h than 50 km/h. We postulate that the results shown in Figure 7 are reasonable based on intuition. The purpose of the field trials is to determine the actual fuel consumption from the index of fuel consumption computed based on the trajectories driven by the vehicles.

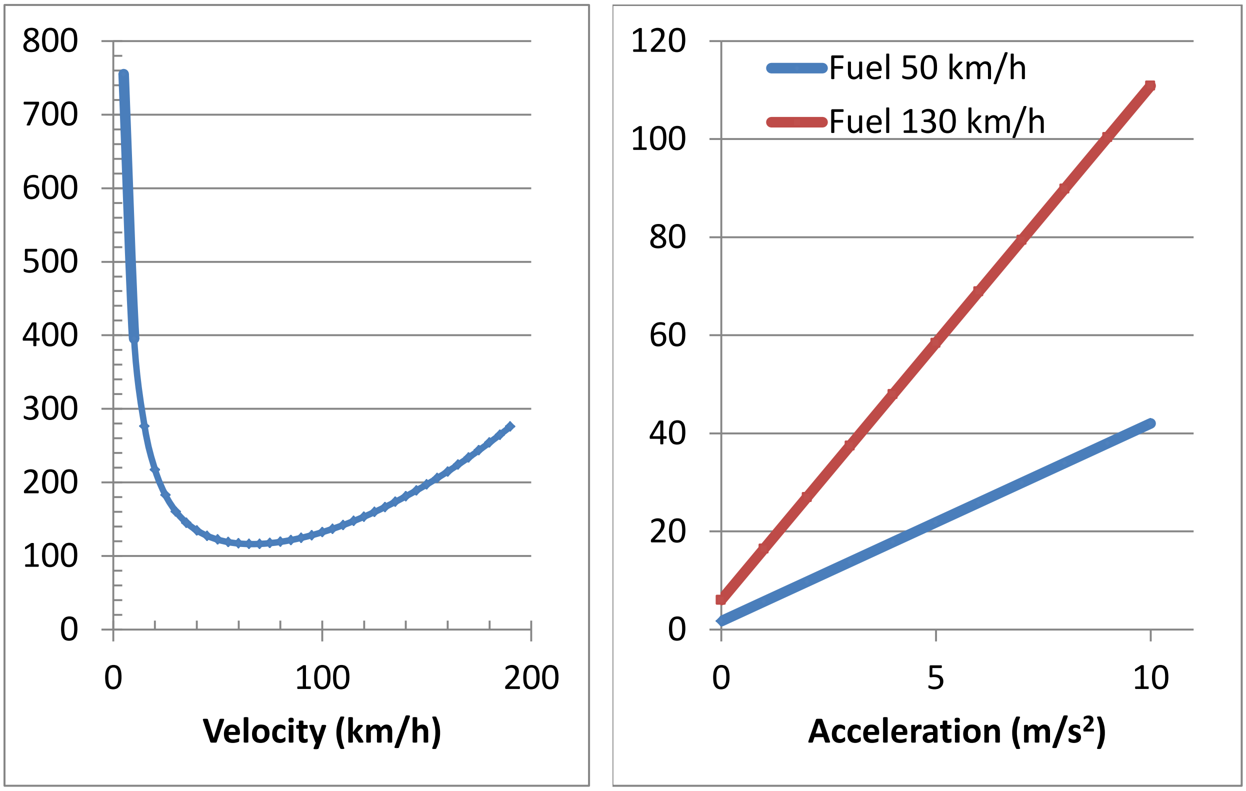


Figure 7 Results of Using Index of Fuel Consumption (both Y-axis Fuel Consumption no unit)

The field trials must determine if the index for fuel consumption is reasonable accurate. The software prototype makes it simple to experiment with all constants listed above. In addition, how to convert how to convert the index of fuel consumption to milliliter of fuel is done at part of the field trials. Both of these tasks are postponed to the field trails because it will require more CANBus data than currently available and require a lot of experimentation.

4.3 CANBus-to-Eco

The CANBus-to-Eco approach is the simplest approach for enabling a data-source type to be used for eco-route estimation because the actual fuel consumption is directly available in the CANBus data. The approach is illustrated with the following example where Table 8 lists the data foundation and Figure 8 shows how the CANBus data is map-matched to a simple map consisting of only three segments.

ID	Segment ID	Vehicle ID	Fuel (ml/l)
1	17	1111	26
2	17	1111	20
3	22	1111	15
4	38	1111	25
5	17	2222	20
6	38	2222	29

Table 8 Data Foundation CANBus data Example

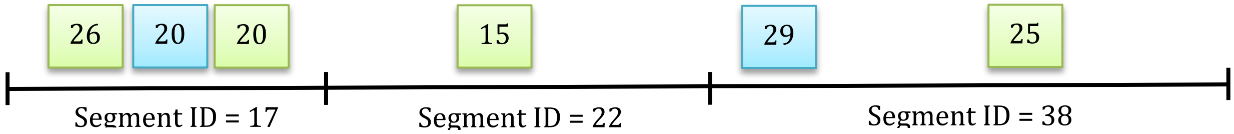


Figure 8 Input to CANBus-to-Eco Algorithm

The data in Table 8 is from two vehicles with IDs 1111 and 2222, respectively. The data from Vehicle ID 1111 is shown in green in Figure 8 and the data from Vehicle ID 2222 is shown in blue. The segment IDs used in Figure 8 are taken directly from Table 8 and it is assumed the segment IDs are founding using a map-matching algorithm. The result of the CANBus-to-Eco is shown in Figure 9.

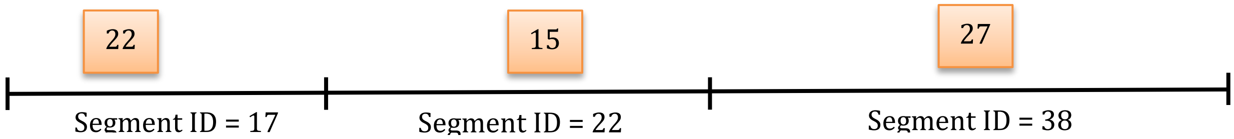


Figure 9 Result of CANBus-to-Eco Algorithm

Here the average of the CANBus data for each segment is used. Experimentation in the field trials in Task 5.2 (and Task 5.5) must determine if this is a reasonable approach or more sophisticated statistical methods must be used. The prototype allows for simple experimentation with the aggregation method used.

4.4 Weighted Combination

To boots-trap the estimation of eco-routes all data-source types available are used. Table 3 shows that the data-source types are not equally well-suited for the eco-route estimation and Table 4 shows that the data-source type least useful for eco-route estimation is the data-source type for which most data is available. In this subsection, it is described how these factors are taken into consideration to make an eco-route estimated that uses all the available data to its fullest.

To balance the usefulness and the availability of data four vectors are used where there is a value for each data-source type available (four). These vectors are the following.

- A *weight* vector that is balance the usefulness of the different data-source types.
- A *sufficient-observation* vector that states when enough data size is available for a data-source type to be considered as accurate as possible. As an example, the estimation of Trajectory-to-Eco does not become more accurate when a segment is covered by 10 trajectories.
- An *observation* vector, which is a count on the cardinality of each data-source type.
- An *eco-value* which is the eco-route estimated one a single data-source type. The values are computed using one of the three algorithms Time-to-Eco, Trajectory-to-Eco, or CANBus-to-Eco.

w_{low}	$w_{low+CAN}$	w_{high}	$w_{high+CAN}$
0.25	1.0	0.5	1.0

Table 9 Example of Weight Vector

As an example consider the weight vector in Table 9 where w_{low} is the weight for the data-source type low-frequent GNSS measurements, $w_{low+CAN}$ is the weight for low-frequent GNSS measurements with CANBus data, w_{high} is high-frequent GNSS measurements, and $w_{high+CAN}$ is high-frequent GNSS measurements with CANBus data. As can be seen from Table 9 the least weight is given to the low-frequent GNSS measurements (0.25), the double (0.50) is given to high-frequent GNSS measurements, and both data-source types with CANBus data is assigned the weight 1.0.

An example of sufficient-observation vector is shown in Table 10 where the subscripts follow the pattern used for the weight vector shown in Table 9.

m_{low}	$m_{low+CAN}$	m_{high}	$m_{high+CAN}$
100	20	50	10

Table 10 Example Sufficient-Observation Vector

The interpretation of the maximum-observation vector is that if 100 observations are available for a single segment then the eco-route estimation using the Time-to-Eco algorithm does not get more accurate if we have 101 or 20,000 low-frequent GNSS measurements. Similar, if 10 high-frequent GNSS measurements with CANBus data are available then the CANBus-to-Eco is assumed not to get more accurate.

An example of an observation vector is shown in Table 11. The interpretation of the table is that 200 low-frequent GNSS measurements are available (above the limit shown in Table 10) and 10 high-frequent GNSS measurements are available (below the limit shown Table 10). Unfortunately, no CANBus data is available for the part of the map that an eco-route is to be estimated for.

n_{low}	$m_{low+CAN}$	m_{high}	$m_{high+CAN}$
200	0	10	0

Table 11 Example of Observation Vector

An example of an eco-value vector is shown in Table 12. The eco-value e_{low} is computing using the Time-to-Eco algorithm, the eco-value e_{high} is computed using the Trajectory-to-Eco algorithm, and $e_{low+CAN}$ and $e_{high+CAN}$ are computed using the CANBus-to-Eco algorithm. In Table 12 it is assumed that the eco-values are computed based on the number of observations listed in Table 11 and therefore are the results for $e_{low+CAN}$ and $e_{high+CAN}$ not available (indicated by the dash) since the example assumes that no data is available (see Table 11).

e_{low}	$e_{low+CAN}$	e_{high}	$e_{high+CAN}$
20	-	40	-

Table 12 Example of Eco-Value Vector

An implementation of the weight approach to compute an eco-value using the Python programming language is shown in Figure 10. For the example listed about, the result is 25.71. This value indicated that the e_{low} value has a significantly higher weight than the e_{high} value. This is because that it is assumed that enough observation are available for low-frequent GNSS measurement but not for high-frequent GNSS measurements, see Table 10 and Table 11).

The reason that a vector based approach is used is that it is straight forward to extend the implementation in Figure 10 to use additional data-source types, such as information from induction loops, Bluetooth, traffic cameras, or manual eco-value estimates if such data-source types should become available to the REDUCITION project.

```python
"""
    Converting the four data-source types in the REDUCTION project to one shared
    value. The ordering in the arrays is (low_gps, low_can, high_gps, high_can)

"""
import numpy

def eco_estimate(weights, sufficient_observations, observations, eco_values):
    # Normalize the number of observations
    rel_obs = numpy.array((0.0, 0.0, 0.0, 0.0))
    for i in range(len(rel_obs)):
        rel_obs[i] = min(1, float(observations[i]) /
                            float(sufficient_observations[i]))

    I = numpy.identity(4) # 4 is the size of vectors
    woi = (rel_obs * I) * weights.transpose()
    woin = woi / woi.sum()
    left = eco_values * woin.transpose()
    return left.sum()

# Values from example
weights = numpy.array([0.25, 1.0, 0.5, 1.0])
sufficient_observations = numpy.array([100, 20, 50, 10])
observations = numpy.array([200, 0, 10, 0])
eco_values = numpy.matrix([20, 0, 40, 0])

print eco_estimate(weights, sufficient_observations, observations, eco_values)
```

Figure 10: Python Code for Implementation of Weighted Eco-Value Estimation

5. Dealing with Insufficient Data

The basic idea behind both travel-time and eco-route computation shown in Figure 1 requires that there is a travel-time and eco-route weight on each of the edges in the graph (the digital map). As shown in Figure 2, Figure 3, and Figure 4 there are not observations on all edges for all data-source types.

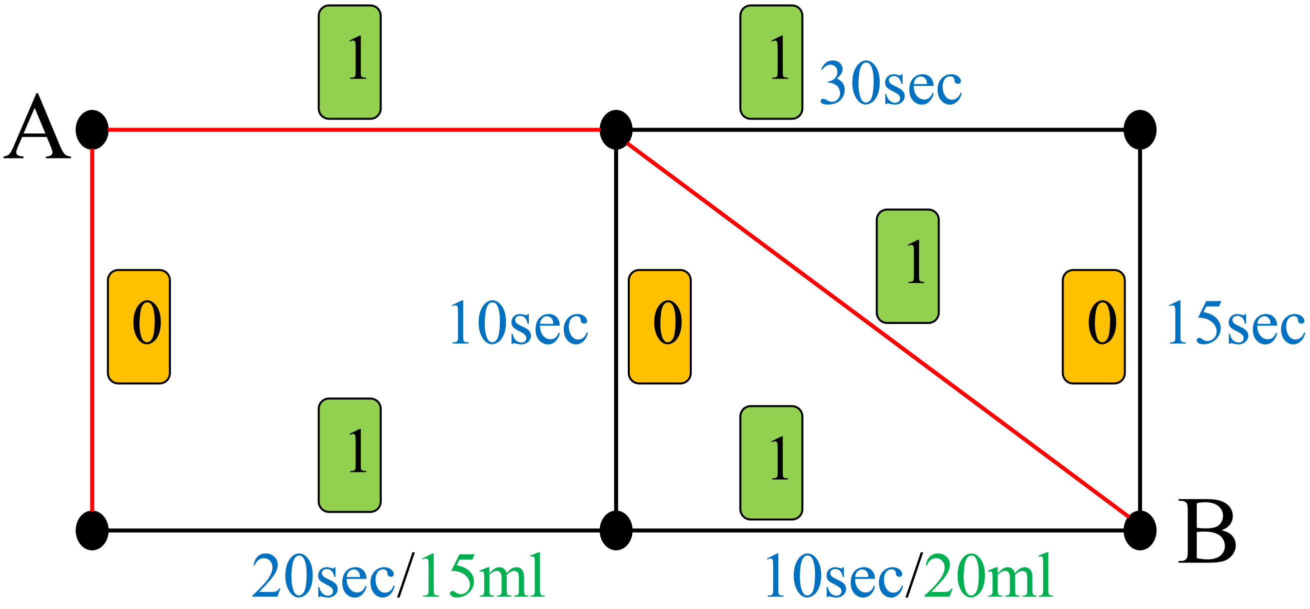


Figure 11 Digital Road Map with Categories and Insufficient Data

This subsection deals with how edges with no of few observations can be assigned a reasonable value. The problem is shown in Figure 11. The map is the same as shown in Figure 1, however here red lines indicates edges with no GNSS measurements or CANBus Data. In addition, there are edges where there are GNSS measurements but no CANBus Data. This is indicated by the missing green values.

The solution to the problem is also shown in Figure 11 where the green and orange boxes indicate segment categories. As all vertical segments (edges) are assumed to be category 0 and the rest of the segments category 1. On digital maps such as Open-Street Map [12] categories are available this is also the case for commercial maps from vendors like NAVTEQ [13] and TeleAtlas [14].

The categories are used to deal with insufficient GNSS measurements and CANBus data. In the example, in Figure 11 it is assumed that travel-time is computed on a category basis because there are GNSS measurement for both category 0 and 1. For the eco-routes there are only values for category 1 and therefore category 0 and 1 are treated as one category. It a simple aggregation is used and it further is assumed that all segments have the same length the result of dealing with the insufficient data in Figure 11 is illustrated in Figure 12 where the travel-time and eco-values on a blue background are estimated values.

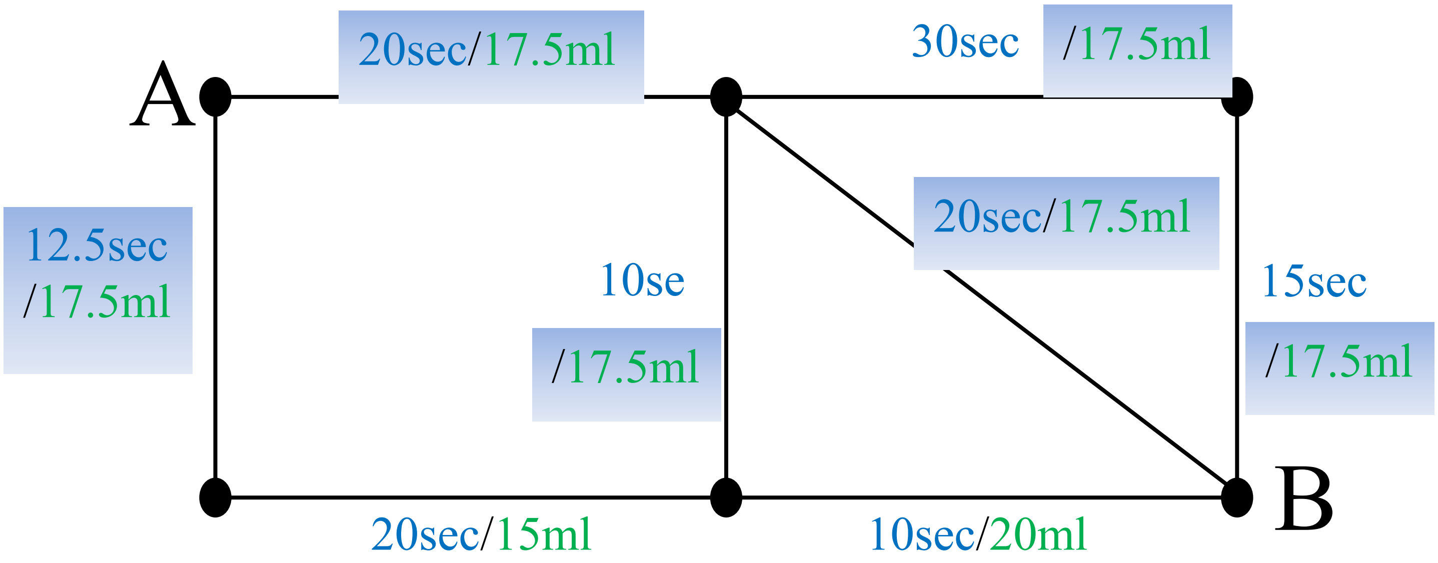


Figure 12 Result of Dealing with Insufficient Data

Please note that in Figure 12 all edges have both a travel-time and an eco-value. The basic idea introduced in Figure 1 can therefore be used.

5.1 Alternative Approaches

There are a number of alternative approaches to dealing with insufficient data values. These approaches can be classified as follows.

- Using additional (and refined) map information, this can for example only aggregating segments with the same category and the same street name or road number.
- Use neighborhood, the can for example be using only segments with the same category within a 2 kilometer radius.
- Using network-distance, this can for example be using neighbor segments or maximum the neighbor of neighbor segments.

In the prototype only the simple category approach is implemented. The field trials must determine if more sophisticated approach to dealing with insufficient data is needed.

6. Results/Output

Three types of output (or results) are available for from the prototype. These outputs are the following.

- A digital map in the Shapefile format with a travel-time and weighted eco-value for each segment for each quarter of an hour in the data.
- A simple web interface that can compute the travel time and eco-value between two points

These outputs are explained in more details in the following.

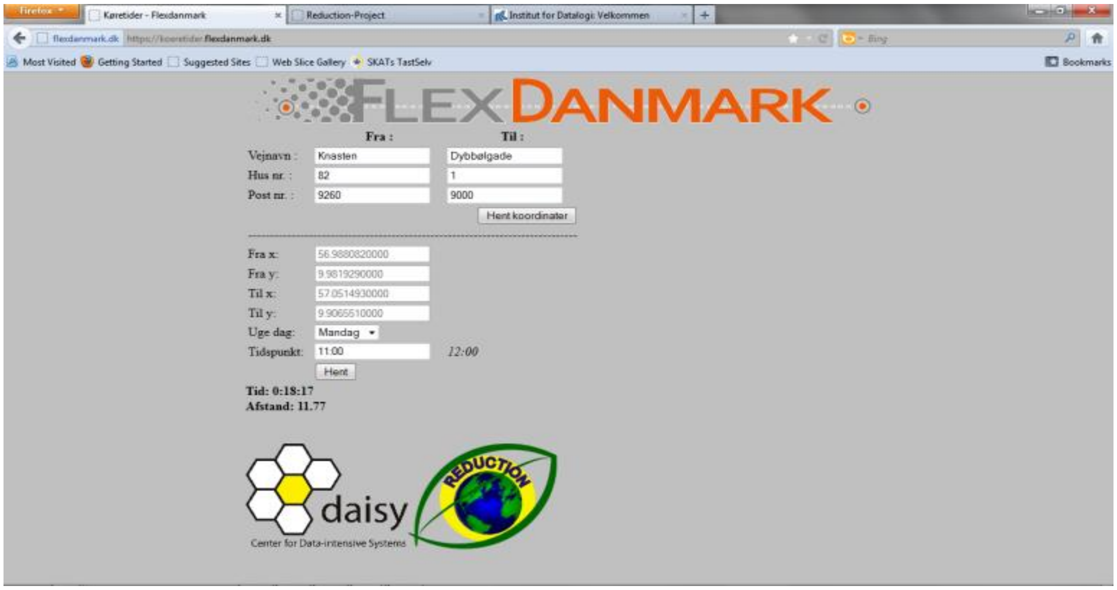
6.1 Shapefile

The Shapefile format is the defector-standard digital map technology. The benefit of using the Shapefile format is that a very large set of existing tools support this file format. The complete output of the REDUCTION project can thus very simply be used by other since it is a very well-known technology. The Shapefile format will be made available for download from the REDUCTION public web site.

The main drawback of the Shapefile format is that it requires a number of IT skills to use. To make the results of the REDUCTION project available to a broader audience a web service is created. This service is discussed next.

6.2 Web Interface/Web Service

Most persons are very familiar with using a web browser. The software prototype therefore provides a simple web interface. The user inputs two addresses and the prototype returns the distance, the travel-time, and eco-route between the two addresses.



	Fra :	Til :
Vejnavn :	Knasten	Dybbølgade
Hus nr. :	82	1
Post nr. :	9260	9000

Fra x: 56.9880820000
Fra y: 9.9819290000
Til x: 57.0514930000
Til y: 9.9065510000
Uge dag: Mandag
Tidspunkt: 11:00 *12:00*

Tid: 0:18:17
Afstand: 11.77

Figure 13 Simple Web Interface: Supply from/to Address and Results Presented

It is also possible to specify a day-of-week and time-of-day such that the information returned to the user can take rush-hour into consideration.

It is fairly simple to extend the prototype to return additional information such as the number of observations used. How to extend the web interface will depend on the results from the two BeKTra/FlexDanmark field trials.

The web interface is implemented by calling a web-service. This make the travel-time and eco-route information easy accessibly from most programming languages. Additional new features added to the web interface will also be implemented using web services such that both a GUI (web-browser) and a programmatic interface are provided.

The web site for computing travel-time between addresses is available from the following https://koeretider.flexdanmark.dk/ (In Danish). The web site allows for the computation of travel-time between any two-addresses in Denmark. The user can specify data weekday ("Uge dag" in Danish) and time-of-day ("Tidspunkt" in Danish). The web site is used by taxi companies.

7. Risks

In the implementation of the functionality described in this document the following risks can be identified. This is a subset of the risks identified in D3.1. However, the additional new risks have been added as sub-bullets.

1) Insufficient access to GNSS data
    a) Low-frequent data
    b) High-frequent data
2) Insufficient set of CANBus data
    a) Low-frequent data
    b) High-frequent data
3) Too large diversity in input data

An evaluation of the each individual risk is shown in Table 13 below. The severity and the likelihood are defined as 1 low, 2 medium, and 3 high. The impact is calculated as the product of the severity and the likelihood (severity * likelihood). The impact it the classified into three categories 1-3 is low, 4-6 is medium, and 7-9 is high.

No	Risk	Severity	Likelihood	Impact
1A	Insufficient low-frequent GNSS data	3	1	3
1B	Insufficient high-frequent GNSS data	2	2	4
2A	Insufficient low-frequent CANBus data	2	1	2
2B	Insufficient high-frequent CANBus data	1	1	1
3	Too large a diversity in input data	2	3	6

Table 13 Risk Assessment

Risk 1A: The partners in the project have themselves access to significant quantities of low-frequent GNSS measurements from for an extended period of time. Further, the partners get new data every day. Therefore this risk is assumed to be low.

Risk 1B: The partners in the project have themselves access to some high-frequent GNSS measurements. This data is from a limited geographical area. However, it is estimated that this data itself is sufficient to be able to use high-frequent GNSS data in the field trial. In addition, an effort has been done contacting a number of owners of high-frequent GNSS data. These owners have been positive providing access to their data. Therefore this risk is assumed to be medium to low.

Risk 2A: The partners in the project have themselves access to some low-frequent CANBus data. This data set is currently not sufficient. However, the project partners have been in contact with a number of taxi companies that are willing to provided low-frequent CANBus data. Have access to CANBus data is not as critical as having access to GNSS data. Therefore this risk is assumed to be low.

Risk 2B: The partners have not themselves access to high-frequent CANBus data. A single EU-project has been asked if access to their CANBus data is possible however, this has request has been declined. The project partners have established contact with a small company that can provide a limited set of high-frequent CANBus data for a limited geographical area in Denmark. The entire project can be done without access to high-frequent CANBus data but we have now established contact to a company that can supply such data. Therefore this risk is assumed to be very low.

Risk 3: The partners have access to both GNSS and CANBus data from a number of companies. These companies typically use CSV files for data transfer. Such files have no common format. It is therefore very important to be able to have a very flexible system for receiving data. Being able to read the data is naturally vital for the project. Therefore this risk is assumed to be medium.

8. Open Issues

This section lists the open issues that should be solved before or during the BeKTra/FlexDanmark field trials.

- As mentioned in the format of the CANBus data supplied to the REDUCTION project is not exactly known at the time of writing. However, it is clear that values engine RPM, throttle position (in percentage), and current fuel consumption can be supplied. The CANBus data currently available to the REDUCTION project is a number of different formats and the prototype therefore does special handling of each case. This is not a solution that scales well.

- There are a number of parameters that must be specified for the Time-to-Eco and Trajectory-to-Eco algorithms. The value of these parameters must be found as part of the experimentation in the first field trial and verified/adjusted in the second field trial.

- There are also a number of weights to be specified for the weighted approach to compute eco-routes. These weights are currently specified using a qualified guesswork. The first field trials should make it possible to find reasonable values for the weights. This included that there may be different set of vectors for each road category or even more refined for each road category in each region.

- To deal with the lack of CANBus data it has be examined if other funded projects can provide access to CANBus data. In particular, has the FP7 project *EuroFOT* [15] been studied because this project has created and collected very large numbers of both GNSS measurements and CANBus data. Unfortunately, the REDUCTION project cannot get access to this data due to “ensure privacy, respect IPR, avoid benchmarking, maintain reasonable information content” [16] page 9. If the REDUCTION project cannot supply sufficient GNSS measurements and CANBus Data before the second BeKTra/FlexDanmark field trail perusing access to the access to EuroFOT data at the REDUCTION project level can be considered.

- The proto-type presented here computes the estimated fuel consumption for driving a trip. It can be consider that the GHG emission for a trip should also be estimated. To do this a number of existing approaches can be used [17]. It is fairly straight forward to include such additional GHG emission estimates when the basic fuel consumption has been estimated. The field trials and feedback from the companies should determine if these additional estimates are needed.

It is the assessments that that all of these open issues can be resolved during the field trials.

9. Conclusion

This report has documented how a basic prototype for travel-time and eco-route has been implemented. Both of these metrics are computed using a number of different data-source types based on GNSS measurements and CANBus data. The usefulness of the different data-source types for computing travel-time and eco-routes has been discussed in details.

The data size available for the REDUCTION project for each of the data-source types has been discussed. It has been identified as a major hurdle that too little CANBus data is available. Of this reason the prototype is implemented such that all the available data is used as far as possible. The central part of the eco-route estimation is therefore a weighted approach where all data-source types available are used.

The output of the prototype is a Shapefile that is mostly suited for IT-person that wants to experiment with travel-time and eco-route. In addition, the prototype has a simple web interface where the public can query a map for the travel-time and eco-route between two addresses.

To the best of our knowledge this delivery has contributed to state-of-the-art in the following areas.

- Categorized GNSS and CANBus data an evaluated how suitable low-frequent and high-frequent data is for eco-routing.
- Proposed algorithms for converting point-based GNSS data, trip-based GNSS data and CANBus data to be useable for eco-routing.
- Suggested methods for how the different data-source types can be combined to make the most of the data available to the project.

References

[1] Environmental Systems Research Institute, Inc, "ESRI Shapefile Technical Description," 1998. [Online]. Available: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf.

[2] J. J. B. R. X. G. W. Wang, "Large-Scale Freeway Network Traffic Monitoring: A Map-Matching Algorithm Based on Low-Logging Frequency GPS Probe Data," 2011.

[3] H. C. N. P. F.C. Pereira, "An Off-line Map-matching Algorithm for Incomplete Map Databases," 2009.

[4] O. Andersen, K. B. Benjamin, T. Kristian, C. S. Jensen, B. Yang, N. Christiansen and N. T. Larsen, "Deliverable 3.1: Report Covering Requirements, Specification, Data-Flow Analysis, and the System Architecture," 2011.

[5] K. Ahn and H. Rakha, "The effects of route choice decisions on vehicle energy consumption and emissions," vol. 13, no. 3, p. 151–167, 2008.

[6] K. Ahn, H. Rakha, A. Trani and M. Van Aerde, "Estimating vehicle fuel consumption and emissions based on instantaneous speed and accelerationlevels," vol. 128, no. 2, 2002.

[7] l. Pelkmans, P. Debal, T. Hood, G. Hauser and M. Delgado, "Development of a simulation tool to calculate fuel consumption and emissions of vehicles operating in dynamic conditions," 2004.

[8] A. N. T. N. B. K. K. H. L. Lahrmann H, "Pay as You Speed, ISA with incentive for not speeding: Results and interpretation of speed data".*Accident Analysis & Prevention.*

[9] ITS Platform Project, "ITS Platform," 09 2012. [Online]. Available: http://www.itsplatform.dk/.

[10] G. Song, L. Yu and Z. Wang, "Aggregate Fuel Consumption Model of Light-Duty Vehicles for Evaluating Effectiveness of Traffic Management Strategies on Fuels," *Journal of Transportation Engineering,* vol. 135, no. 9, pp. 611-618, 2009.

[11] J. L. Jimenez-Palacios, "Understanding and Quantifying Motor Vehicle Emissions with Vehicle Specific Power and TILDAS Remote Sensing," 1999.

[12] Open-Street tMap, "Open-Street Map," 09 2012. [Online]. Available: http://www.openstreetmap.org/.

[13] NAVTEQ, "NAVTEQ," 09 2012. [Online]. Available: http://www.navteq.com/.

[14] TomTom, "TeleAtlas," 09 2012. [Online]. Available: http://www.teleatlas.com/.

[15] EuroFOT, "EuroFOT," 09 2012. [Online]. Available: http://www.eurofot-ip.eu/.

[16] C. Kessler and A. Etemad, "Delivery 6.8 FOT Data," EuroFOT project, 2012.

[17] EPA, "UNIT CONVERSIONS, EMISSIONS FACTORS, AND, OTHER REFERENCE DATA," 2004.