

# **Modelling Emission of Pollutants from transportation**

**Using mobile sensing data**

Anders Lehmann

July 2014

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Problem Statement . . . . .	1
1.2	Methodology . . . . .	2
1.2.1	Secondary research . . . . .	2
1.2.2	Data analysis . . . . .	2
1.2.3	Field trials . . . . .	2
1.3	Contributions . . . . .	2
1.4	Reading guide . . . . .	3
<b>2</b>	<b>Climate change policy development</b>	<b>4</b>
2.1	Climate change organisations . . . . .	4
2.2	Climate change mitigation . . . . .	4
2.3	Climate change adaptation . . . . .	5
<b>3</b>	<b>CO<sub>2</sub> Modelling</b>	<b>6</b>
3.1	IPCC methodology . . . . .	7
3.1.1	IPCC Tier 1 model . . . . .	7
3.1.2	IPCC Tier 2 model . . . . .	8
3.1.3	IPCC Tier 3 model . . . . .	8
3.2	Modelling emissions from electric vehicles . . . . .	10
<b>4</b>	<b>Air pollution dispersion Modelling</b>	<b>11</b>
4.1	Eulerian method . . . . .	12
4.2	Lagrange . . . . .	12
4.3	DASK-GAS . . . . .	12
4.4	OSPM . . . . .	13
<b>5</b>	<b>Internet of Things</b>	<b>15</b>
5.1	Linked Data . . . . .	15
5.2	Semantic annotated data . . . . .	15
5.3	Linked Devices . . . . .	16
5.4	Linked Data as an Enterprise Architecture strategy . . . . .	16
5.5	Scientific contributions . . . . .	17
<b>6</b>	<b>My Contributions</b>	<b>18</b>
6.1	Simple emission model . . . . .	18
6.1.1	Mobile sensing . . . . .	18
6.1.2	Tier 2 derived model . . . . .	19
6.2	Enhanced emission model . . . . .	20
6.2.1	Determine Idle engine revolutions . . . . .	21
6.2.2	Determine acceleration and turning . . . . .	23
6.2.3	Combine turning information with GPS . . . . .	24

<b>7</b>	<b>Future Research</b>	<b>25</b>
7.1	Hypotheses . . . . .	25
7.2	Improving estimation of emissions for single trips . . . . .	25
7.2.1	Idle emissions . . . . .	25
7.2.2	Emissions when accelerating . . . . .	26
7.2.3	Emissions when decelerating . . . . .	26
7.2.4	Emissions when cruising . . . . .	26
7.2.5	Papers . . . . .	26
7.3	Correlation of trips . . . . .	27
7.4	Improve dispersion models . . . . .	27
7.5	Impact . . . . .	27
7.5.1	Papers . . . . .	28
7.6	Field trials . . . . .	28
7.6.1	Sønderborg . . . . .	28
7.6.2	Herning . . . . .	28
7.6.3	Elbil parat . . . . .	29
7.6.4	Papers . . . . .	29
7.7	Visiting researcher . . . . .	30
<b>8</b>	<b>Conclusion</b>	<b>31</b>
	<b>Acknowledgement</b>	<b>32</b>

# 1 Introduction

During the last 40 years an overwhelming amount of evidence for Anthropogenic Global Warming has been gathered [38]. This evidence has led to international agreement on certain limits for the emission of gases that affect the global warming, most notably  $CO_2$ . The gases that affect the climate through warming of the atmosphere, are commonly known as radiative forcing gases or climate forcing gases. In order to reach these goals and keep the level of  $CO_2$  within the agreed limits, technological solutions need to be developed, partly to develop cleaner energy production technology, and partly to develop more energy efficient ways of living. Previously the focus has been on developing more efficient machines, under the assumption that the usage patterns of the energy consuming machines were optimal.

This phd study is undertaken as part of the three year research project EcoSense<sup>1</sup> which is aimed at analysing user behaviours, through the use of mobile sensing with smartphones, and if possible give advice to change user behaviour to be more resource efficient. EcoSense is not the only project trying to use mobile sensing [27] [5]. My part of EcoSense is to develop, implement and evaluate models for emissions. In this report I will mostly refer to  $CO_2$  emissions as this is the most important climate forcing gas. To measure the total impact on climate forcing, the contributions of other gases has to be taken into account. This is done by converting the contributions into  $CO_2$  equivalents - that is the amount of  $CO_2$  giving the same amount of radiative forcing.

When modelling emissions of climate gases from passenger transportation, techniques for modelling emissions of air pollutants are employed, which enables me to also study air pollution from passenger transport. Air pollution has different aspects, compared to climate science, mainly because the pollutants are considered due to their toxicity properties (towards humans) instead of their radiative forcing properties.

Due to the complexity and multifaceted nature of emission modelling the study has become a multidisciplinary study. That is, I have used methods, knowledge and technologies from different scientific fields such as computer science, environmental science, climate science, political science and engineering science.

## 1.1 Problem Statement

One driving goal for the project, is to be able to improve the accuracy and ease of use of green accounting for companies. Today it is a tedious and time-consuming task to estimate the amount of emissions a company is responsible for. One of the cumbersome areas of green accounting is with regard to emissions due to transport. There is a need for methods and models that will account for the actual emissions due to the day to day business of the company and not accounting done from mean values. In the study the focus has been on passenger cars, since passenger cars is responsible for about 60% of the emission of green house gases from transportation in Denmark [29]. The methods studied will be transferrable to other vehicle types.

---

<sup>1</sup><http://ecosense.au.dk>

In the EcoSense project the emphasis is on climate gases, but since emissions from passenger transport also contains other pollutants with no or small climate forcing properties, the methods for modelling the emissions of climate gases can also be used for modelling the emissions of pollutants. Thus I will investigate how the mobile sensing data of passenger transport can be exploited to improve modelling of air pollution.

Another focus of the study is to use the amount of data collected to be able to help inform and inspire traffic planning.

## **1.2 Methodology**

In the following I will discuss the methods used for estimating emission of climate forcing gases from transportation and proposals for improving these methods.

As the study is inspired of a number of different fields, spanning from Climate Change policy over atmospheric chemistry to data science and signal analysis, I have adapted a number of different methods for solving the problems.

### **1.2.1 Secondary research**

To find suitable models for emissions of Climate forcing gases from traffic a number of emission accounting methodologies has been investigated through a systematic literature search and appraised for the suitability of the methods in the EcoSense project.

### **1.2.2 Data analysis**

The data that is gathered from smartphones, is essentially timestamped values from the sensors in the phones. To analyse the data it is first converted into a time series format before a number of different methods are applied in order to find and present interesting patterns. For analysing GPS traces the distance travelled is calculated by applying an algorithm taking the curvature of the earth into account, thus employing methods from geometry.

For analysing accelerometer data algorithms from signal analysis, are employed to explore the data and extract information on engine speed and vehicle activity.

### **1.2.3 Field trials**

The models developed in the first part of the phd study are going to be applied to field trials in different cities in Denmark, through smartphone applications designed for each trial. These field trials will be used as an evaluation of the used methods.

## **1.3 Contributions**

To get information on the emissions from a single trip in a passenger car, the most precise way would be to measure the emissions from the tailpipe [10]. This is a bit impractical as it involves instrumenting the vehicle.

As a more crude estimate, the methods from IPCC inventory guidelines can be used. These methods are designed to make the task of creating national climate forcing gas emission inventories, a practical endeavour. In between these two methods there is a gap, which I will try to fill, without extra instrumentation of the vehicles.

My contribution to the project is threefold:

1. To create and improve emission models for individual trips.
2. To create models for aggregation and correlations of multiple trips.
3. To improve models for the dispersion of emissions.

By succeeding in these goals the impact would be:

1. Improved accuracy of green accounting.
2. Providing hints to increase efficiency of transportation.

## **1.4 Reading guide**

In the first part of this report ( section 2 to section 5 ) I will go through the related work I have found, and give a background to understand the methods used. The second part of the report is describing my contributions and my planned future work.

Section 2 is an overview of the international institutions put into place to oversee and develop Climate Change policies. It is important to keep in mind that the problems of Climate Change mitigation and adaptation are global problems, thus has to be solved in an international setting.

Section 3 describes the work done so far to understand and implement models for estimating the emission of climate forcing gases from traffic. The methodology decided by IPCC is described and the way that the principles has been adopted to single trip estimation is elaborated.

Section 4 is about ways to model the movement of airborne pollutants. Originally these methods were developed for environmental planning purposes, but in this study it will be combined with the detailed knowledge of the emission pattern from traffic.

Section 5 concerns ways connecting and relating data from sensors on different devices.

In the section 6, I have made a detailed description of the contributions I have made in the study to date.

The section 7 is a plan for the remaining period of the study.

## 2 Climate change policy development

### 2.1 Climate change organisations

It has been decided in the United Nations to create two bodies for devising policies for managing the anthropogenic global warming. The main political body is the United Nation Framework Convention for Climate Change (UNFCCC) [37]. The name refers to both a convention created in 1994, and a secretariat that services the operation of the Convention. This body is tasked with creating policy initiatives and negotiate treaties with the members of the UN. A visible part of these negotiations are the COP (Conference of parties) meetings held annually. Since 154 nations have ratified the convention, the negotiations are very difficult, and progress in for instance agreeing on firm goals and plans to reach the goals is slow [32].

To support the UNFCCC with scientific knowledge the Intergovernmental Panel on Climate Change (IPCC) was created. The panel gathers scientific evidence into reports on different aspects of climate change, the root causes, the consequences and ways to mitigate and adapt.

Even though the panel is a scientific panel, it is also a part of a political process. This means that the reports made by IPCC, especially the "summary for policy makers" are also heavily negotiated by the participating government officials.

The latest reports from IPCC are the AR-5 (fifth assessment report) series .

The IPCC reports are divided into sections, done by different workgroups. Workgroup 1 delivers a scientific background section containing the latest scientific knowledge about emissions, measured consequences, melting of Icecaps etc. As a new feature in the assessment report a number of scenarios are envisioned, to estimate the consequences of different policies. The scenarios are called "Representative Concentration Pathways", and are named by the expected rise in global mean temperature. Four scenarios are envisioned in the report: RCP2.5, RCP4.5, RCP6, RCP8.5 [41]. These scenarios are then used as a guiding principle for the rest of the sections in the AR-5.

The workgroup 1 report is generally viewed as a high quality and reliable source of knowledge.

Workgroup 2 delivers a report on "Impacts, Adaptation and Vulnerability".

Workgroup 3 is about mitigation of climate change.

### 2.2 Climate change mitigation

From the start of the debate about how to combat Anthropogenic Global Warming, there have been two competing approaches. The mitigation approach, where an effort is made for reducing the emissions of climate gases, is the approach which has received the most attention[31]. The EcoSense project is part of the mindset behind this approach. Whereas the mitigation approach previously has focused on creating more efficient machines, to either produce energy with less emissions or produce machines that produce more useful work per energy unit[7], EcoSense focuses on how the machines are used in combination. So instead of focusing on a single energy consuming item, we are trying to

analyse how the energy consumption changes as the individual items work together. In other words we are developing tools for studying network effects in the transportation sector, as well as in other sectors.

### **2.3 Climate change adaptation**

The second approach is called the Adaptation approach. Since it is probable that there will be significant changes in the climate, even if we succeed in keeping the  $CO_2$  emissions within the agreed limits, we have to invent ways to adapt to these changes. The Adaptation approach has not received much attention or research funding, but projects like EcoSense will also have an impact in determining how to adapt to changes in weather patterns. It seems that IPCC in the fifth report from Working Group 2 is beginning to give more attention the Adaptation approach, as a consequence of realising that some level of Climate Change is inevitable [21][40].



### 3 $CO_2$ Modelling

In this section the IPCC [43] methodology for calculation of emission of climate forcing gases from transportation is described. The focus of the IPCC approach is to offer methods for building national inventories for climate gas emissions. This is an attempt to summarise the key methodologies, and I will use this as a basis for developing models for single trip emissions.

Group 1	
Carbon monoxide	CO
Nitrogen oxides	(NO <sub>x</sub> : NO and NO <sub>2</sub> )
Volatile organic compounds	(VOCs)
Methane	(CH <sub>4</sub> )
Non-methane VOCs	(NMVOCs)
Nitrous oxide	(N <sub>2</sub> O)
Ammonia	(NH <sub>3</sub> )
Particulate matter	(PM)
Group 2	
Carbon dioxide	(CO <sub>2</sub> )
Sulphur dioxide	(SO <sub>2</sub> )
Lead	(Pb)
Arsenic	(As)
Cadmium	(Cd)
Chromium	(Cr)
Copper	(Cu)
Mercury	(Hg)
Nickel	(Ni)
Selenium	(Se)
Zinc	(Zn)
Group 3	
Polycyclic aromatic hydrocarbons	(PAHs)
Persistent organic pollutants	(POPs)
Polychlorinated dibenzo dioxins	(PCCDs)
Polychlorinated dibenzo furans	(PCDFs)
Group 4	
Alkanes	(C <sub>n</sub> H <sub>2n+2</sub> )
Alkenes	(C <sub>n</sub> H <sub>2n</sub> )
Alkynes	(C <sub>n</sub> H <sub>2n-2</sub> )
Aldehydes	(C <sub>n</sub> H <sub>2n</sub> O)
Ketones	(C <sub>n</sub> H <sub>2n</sub> O)
Cycloalkanes	(C <sub>n</sub> H <sub>2n</sub> )
Aromatic compounds	-

Table 1: IPCC considered species

### 3.1 IPCC methodology

The IPCC has made a number of reports on how to calculate emissions of climate forcing gases and pollutants. The gases that IPCC is describing methods for are divided into four groups (see table 1):

Group 1 are pollutants where there exists detailed methodology for estimating the emission, from activity data, such as driving conditions and engine conditions.

Group 2 are pollutants which can be estimated from fuel consumption, when there is a direct connection between the burning of fuel and the emission. The estimates of emissions of the Group 2 pollutants are regarded to be as precise as the estimations for the Group 1 pollutants, even if the methodology differs. Group 3 and Group 4 are pollutants, where no detailed methodology exist for estimating the emission, so a simple method is used to calculate the emission.

The fourth group of pollutants are species, where the emission is calculated as a fraction of the Non Methane Volatile Organic Compound (NMVOC).

When considering Climate forcing gases the most important gases are in Group 1 and 2, thus this is what the focus has been on in this phd study, the most important being  $CO_2$ ,  $CH_4$  and  $NO_2$ . The amount of  $CO_2$  emissions is by far the largest, but since the radiative forcing of  $CH_4$  is 21 times higher than  $CO_2$  and the radiative forcing of  $NO_2$  is 206 times higher (per molecule), they contribute significantly to the global warming.

The way emission inventories are created are described in the Guidelines from IPCC and extended in EMEP/EEA Guidelines[30]. In the guidelines three different methods are described, each method more accurate than the previous.

#### 3.1.1 IPCC Tier 1 model

The Tier 1 method is based on numbers for national sales of hydrocarbons (Gasoline, Diesel, Natural gas etc.). These numbers are readily available for most countries and are converted into emission inventories by multiplying emissions factors (grams of the specie pr kilogram of fuel) for each type of fuel. The Tier 1 method is the simplest, but also most crude way of estimating the national emissions, as it does not account for import or export of fuel, and that not all emissions are directly related to fuel consumption. The only data points needed are: Volume of sales of the different fuel types and the emission factors for the different fuel types.

$$E_p^{CALC} = \sum_i V_{sales,i} * e_{i,p} \quad (1)$$

Here the  $E_p^{CALC}$  is the calculated emission, where  $i$  is the fuel type,  $V_{sales,i}$  is the sales for fuel type  $i$  and  $e_i$  is the emission factor for the fuel type.  $p$  is the pollutant, for which we are finding the emission.

This method is inadequate to use to estimate, single trip emissions.

### 3.1.2 IPCC Tier 2 model

In the Tier 2 method the emission inventories are estimated by estimating the traffic volumes for different categories of vehicles, and multiplying emission factors (gram per kilometre) for each category. The vehicles are divided into six main categories: Passenger Car, Light Duty vehicles, Heavy Duty vehicles, Buses, Mopeds and Motorcycles. For each of the main categories, a subdivision is made, to accommodate for different emission characteristics stemming from pollution regulation, fuel type and engine size. For instance, in Europe, passenger gasoline cars are subdivided into 13 different types, according to the legislation governing allowed emissions. These regulations have been changed and tightened 13 times since the first emission control legislation was ratified in the early nineties. For each vehicle category and vehicle type and legislation class, activity data has to be obtained. The activity data consist of the number of vehicles, and the number of kilometres they drive pr year, for each class. The IPCC has generated tables of emission factors (as g/km) for each class of vehicles. By multiplying these emission factors with the estimated kilometres and number of vehicles in the class, the total emission of a pollutant can be estimated for the vehicle class. The total annual emission from transport can then be calculated as the sum of all the vehicle classes.

$$E_p^{CALC} = \sum_v \sum_i D_{v,i} * e_{i,v,p} \quad (2)$$

Here the  $E_p^{CALC}$  is the calculated emission, where  $i$  is the fuel type,  $D_{v,i}$  is the estimated distance travelled vehicle type  $v$  and fuel type  $i$  and  $e_{i,v,p}$  is the emission factor for the fuel type.  $p$  is the pollutant, for which we are finding the emission.

The Tier 2 model contains data that can be used for a simple emission model for single trip emissions.

### 3.1.3 IPCC Tier 3 model

The Tier 3 method is based on the Tier 2 methods and improve on the estimated emission, by also considering the velocity distribution of the travelled distances, and by considering the effects of cold-starts and vehicle age on the total emissions.

There are two ways proposed to calculate the effects of speed on exhaust emissions. Either by dividing the travelled distance into road types with different speed characteristics, i.e. urban, rural and highway. In this case the total emission for a vehicle class will be calculated as the sum of the product of travelled distance on a road type and the emission factor for that road type and vehicle type.

The other method uses a measured speed to emission curve and a speed distribution function to estimate the emission. For some pollutants an emission factor function is given, but for pollutants that are directly linked to fuel consumption (i.e.  $CO_2$ ) a fuel consumption function of speed is given (for instance equation 4). For  $CO_2$  the relationship between emission and fuel consumption is given by the formula 3.

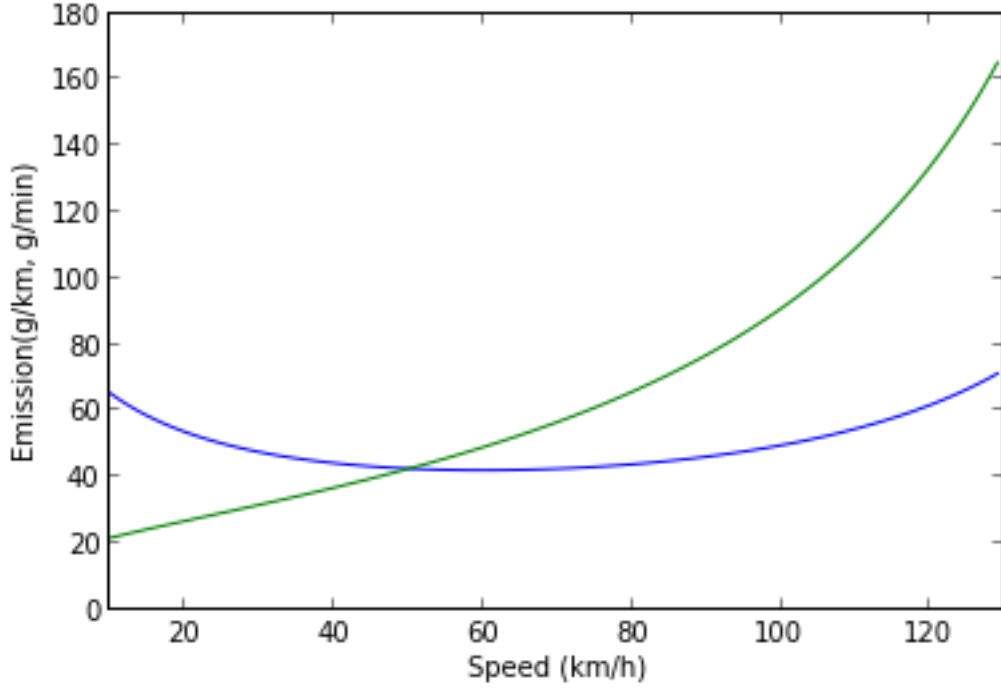


Figure 1: **Emission factor as a function of speed. g/km blue. g/min green**

$$E_{co_2}^{CALC}, km = 44,011 * \frac{FC^{CALC}}{12,011 + 1,008 * r_{H:C} + 16,000 * r_{O:C}} \quad (3)$$

This formula is based on the assumption that all fuel is consumed, i.e. that each carbon atom in the fuel is oxidised into  $CO_2$ , which has the molar weight of 44,001. The numerator  $FC^{CALC}$  is the calculated fuel consumption factor, and this is divided with the molar weight of carbon plus the molar weight of hydrogen multiplied with the ratio of hydrogen to carbon,  $r_{H:C}$ , plus the molar weight of oxygen multiplied with the ratio of oxygen to carbon,  $r_{O:C}$ , in the fuel. There are tables for different ratios  $r_{H:C}$  and  $r_{O:C}$  for different fuel types, given in the COPERT IV methodology guidebook [30].

As an example, passenger vehicles that are regulated by EURO 1 the fuel consumption factor curve is given by equation 4. The coefficients  $a$  to  $e$  are given for the different engine sizes and pollution regulations.

$$FC = \frac{a + c * V + e * V^2}{1 + b * V + d * V^2} \quad (4)$$

An emission factor (in g/km) versus speed curve can be seen in figure 1 for a vehicle with engine size less than 1,4l [26], as the blue curve. From the curve it can be seen that the emission factor is larger for low speeds and for high speeds and lowest at moderate

speeds (app. 60 km/h). The higher emission factor for low speeds is higher because the engine is underused and therefore not very many kilometres are gained. For higher speeds the opposing forces from wind and friction in bearing, force the engine to work harder, and therefore has a higher emission factor. If instead we look at the emission per minute, by multiplying the emission curve with the speed, we get the green curve. From the green curve it can be seen that the emission per time unit is a monotonic growing function, which resembles a second order polynomial. The force of the wind on an object is proportional to the square of the speed.

The emission is calculated as the integral of the speed distribution multiplied by the emission function as seen in equation 5.

$$e_{i,k,r} = \int e(V) * f_{k,r}(V) dV \quad (5)$$

( $i$  is the pollutant for which the emission is calculated,  $k$  is the vehicle type,  $r$  is the fuel type,  $e(V)$  is the emission function of speed, and  $f_{k,r}$  is the speed distribution function for the vehicle type  $k$  and fuel type  $r$ ).

The data needed to use the Tier 3 method is quite extensive. For each of the 13 vehicle classes (divided by fuel type, vehicle size) activity data for mileage for urban, rural and highway travel, hot start/cold start, as well as data for the number of vehicles in each emission regulation is needed. Fortunately the program COPERT IV contains the emission factors, and country specific activity data can be downloaded to use for evaluating national inventories.

This model will form the basis of a complex model for emissions of single trips.

### 3.2 Modelling emissions from electric vehicles

We can calculate the emissions for electric cars by using near real time emission data from Energinet.dk, under the assumption that electric cars will be charged with electricity from the public grid. There is for now no way to detect if the charging of electric vehicles are not done through the public electric grid, and not for instance by private solar plants.

The dataset from Energinet.dk contains production data for the large power plants, the distributed power plants, wind farms and solar. For each kind of producer, the production is given for east and west Denmark, and the import/export of power to the neighbouring countries is also given. There is data for the temperature and windspeed (for a single location), and an estimated emission factor in  $gCO_2/kWh$ . The resolution of the data is 5 minutes.

## 4 Air pollution dispersion Modelling

This section gives an overview of different ways for calculation of how the concentration of pollutants change over time and space. Typically we investigate how pollutants move from a pollution source with wind or through diffusion through the area of interest, the general term used is advection. The concentrations of the species (each kind of pollutant is regarded as a specie) change partly by dilution (the plume widens), and partly by chemical reactions in the atmosphere. The chemical reactions are dependent on different factors, such as light and temperature. In the section on OSPM the effect of the landscape (here a street canyon) is discussed.

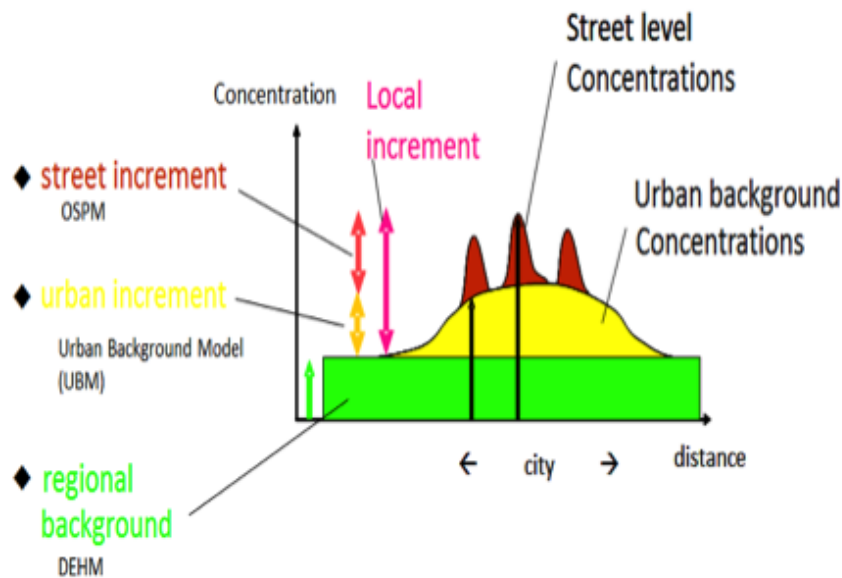


Figure 2: **Different contributions to air pollution.**[19].

To model the concentration of pollutants at a specific place, the researcher has to consider the different sources of the pollution as pollution is transported over wide ranges. To model the long range transport of pollutants coarse models are used. This will contribute with a regional background pollution concentration. For a city more fine grained models (Urban Background Models) are used to calculate the contribution from the city, which leads to calculation of a city increment of emissions over the regional background concentration. And when street level pollution are the target of the calculation, a model for the pollution contribution and distribution for the street is needed, which considers local emissions and local conditions to calculate the street level increment combined to the regional background concentrations and urban background concentrations (see figure 2).

## 4.1 Eulerian method

The Eulerian method is also called a grid model. For the area that we want to investigate, the volume of air above the area, is divided into boxes. In at least one box there should be a source of pollution. The concentration of the pollutants in the box can be calculated as the amount of the species coming into the box (either through the walls of the box, from chemical reactions or from sources located inside the box) minus the amount of the species lost (through the walls, by deposition or by chemical reactions).

For each time step the concentration of each specie is calculated by solving an ordinary differential equation (ODE) and the calculated concentrations are used to determine the start conditions for the next time step, as well as the boundary conditions for the surrounding boxes [45].

Examples of grid models are Danish Eulerian Hemisphere Model (DEHM)[8], which calculates concentration of 63 different species and 120 chemical reactions in different scales (size of a single grid) (150 km x 150 km, 50 km x 50 km and 17 km x 17 km) [11].

## 4.2 Lagrange

In the Lagrange method a parcel of air is considered. The movement of the parcel is simulated by considering the local meteorological conditions, and the statistical mechanics of the particles in the parcel. By simulating many traces of parcels the advection of the pollutants from a source can be visualised. In this method for calculating the advection of pollution the observer follows the advection from the perspective of the pollution, instead of from a fixed position [39]. Inside the parcel the reactions of the gases are modelled while the parcel is moving.

The modelling can also be performed backwards to ascertain where the air pollution at a given position is originating from.

## 4.3 DASK-GAS

DANish Solver for chemical Kinetic - Gear, Analysis, Solver (DASK-GAS)<sup>2</sup> is a chemical compiler and solver. The idea of DASK is to automate the code generation for handling chemical reactions. In DASK you give the chemical reactions and their reaction rates as input, and the program transforms the input into fortran code.

The reactions are given as the chemical compound names in a reaction schema, input reagents on the left side and resulting compounds on the right side. The temperature dependency of the reaction rate is given as a Fortran function, and if the reaction is driven by photolysis it is marked in the data file. For each of the reactions Fortran code is generated to calculate loss and gain of the species. The generated Fortran code is called inside the ODE solver, so that it is easy to add new chemistry [16].

The setup of the chemical reactions used in the simulation is called a mechanism, a number of standardised mechanisms exist, such as European Monitoring and Evaluation

---

<sup>2</sup>unpublished solver from Prof. Allan Gross

Programme, EMEP, Regional Atmospheric Chemistry Mechanism, RACM and Regional Acid Deposition Model Mechanism, RADM2 [12].

Beside the chemical compiler, there is also an easy way of setting up scenarios for the solver to model.

The specification of photolysis takes into account the amount of sunlight available in the period and place which is simulated. This is done by listing values for the sunlight for each time interval within the simulation period in a separate data file.

#### 4.4 OSPM

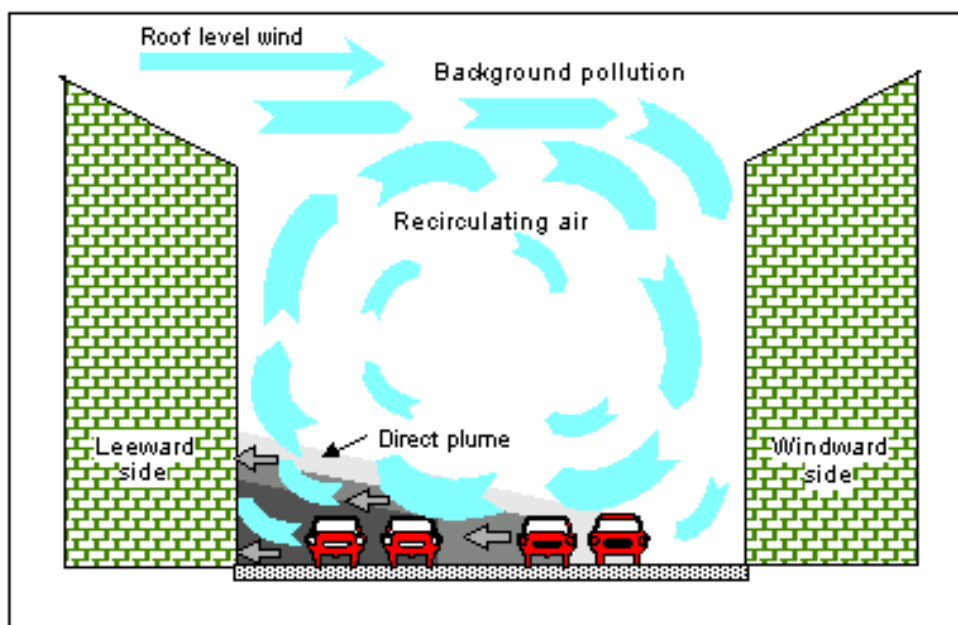


Figure 3: **Street canyon model. The turbulence drives pollution towards the leeward side of the street<sup>3</sup>.**

The Operational Street Pollution Model <sup>4</sup> is a model for calculation of pollution in urban environments. Due to turbulent wind conditions in urban street canyons the pollutants do not mix well with the surrounding air, there will be a tendency to have the heavier pollutants to concentrate on leeward side of the street canyon (see figure 3). The OSPM model considers the effects of street geometry, wind speed, emission factors and atmospheric chemistry (i.e. the  $NO - NO_2 - O_3$  cycle), as well as meteorological data, and background pollution stemming city contributions and from regional background due to long range air transport. The COPERT IV model for activity based emission

<sup>3</sup><http://www.au.dk/ospm>

<sup>4</sup><http://envs.au.dk/videnudveksling/luft/model/ospm/>



models is built in to OSPM, as an example of the extensive data used to model the street level pollution concentrations.

To use the program, detailed information on the geometry of the street and surroundings has to be given. The height of the surrounding buildings, the angle of the street with respect to the wind and the distance to intersections of street corners has to be input to the program.

The program uses emission data for background concentrations from European Monitoring and Evaluation Programme (EMEP) , as input , but actual traffic volume for the considered street has to be given [2].

## 5 Internet of Things

A centrepiece in the EcoSense project is the mobile sensing approach. Mobile sensing means that data is acquired from potentially a huge amount of distributed devices. To manage this large amount of data, and the participating devices, the technologies in Internet of Things can be applied.

The vision of Internet of Things is to connect devices, across application, geographic and company boundaries. To accomplish this vision, a common language and framework had to be agreed upon. The World Wide Web Consortium (W3C) created the Resource Description framework (RDF)<sup>5</sup> in 1999, to serve this vision. The definition of RDF have spurred the creation of query languages (SPARQL [33]), semantic frameworks (OWL [25], SKOS<sup>6</sup>), IDE's (Protégé<sup>7</sup>, TopBraid composer<sup>8</sup>) and implementations specialised RDF capable databases and frameworks (SESAME<sup>9</sup>, Virtuoso<sup>10</sup>, RDFLIB<sup>11</sup>).

### 5.1 Linked Data

The main idea in Linked Data is that every data element has an URI (Universal Resource Identifier) [4], so that in principle each data element can be looked up via the internet. URI's does not have to be dereferenceable though, as the main purpose is to be identifiers of data.

The second idea is to use a simple, but expressive, data representation and modelling language. RDF is deceptively simple at first look. The data is modelled as a graph. The nodes in the graph are data resources (URI's), and the edges are described with triples of subject (left node), predicate (the type of edge) and object (the right node).

The complexity arises from the idea that the modelled data should be consistent under the semantics, given by the semantics of the predicates and the semantics of the types system defined for literals. One of the critiques of RDF is that it is possible, and even easy, to create inconsistencies when modelling data [18]. The answer from the RDF proponents is that the restrictions put on the semantics to ensure consistency, is too restrictive and makes the data modelling hard and unnatural, and also does not offer other qualities as decidability (there are certain versions that do offer decidability, but others do not). OWL is based on Description Logics [17].

### 5.2 Semantic annotated data

In the EcoSense project, a driving idea is that by the integration of data from many distributed and mobile devices, new knowledge can be gained. To alleviate the problem of not guaranteeing consistency, the Web Ontology Language (OWL) was conceived to

---

<sup>5</sup><http://www.w3c.org/rdf>

<sup>6</sup><http://www.w3.org/TR/skos-reference/>

<sup>7</sup><http://protege.stanford.edu>

<sup>8</sup><http://www.topquadrant.com/tools/ide-topbraid-composer-maestro-edition/>

<sup>9</sup><http://www.openrdf.org>

<sup>10</sup><http://virtuoso.openlinksw.com>

<sup>11</sup><http://rdflib.readthedocs.org/en/latest/>

add subject, predicates and object types with defined and consistent semantics, so as to enable inference on data modelled with OWL. Inference is the ability to learn new facts from the available data and a data model. OWL is now in version two, which has tried to bridge the gap towards the RDF proponents by defining a number of profiles, with different tradeoffs between modelling flexibility, decidability and performance.

### 5.3 Linked Devices

For sensor devices specialised protocols for accessing and delivering data, has been developed. The Semantic Sensor Network (SSN) <sup>12</sup> proposed a way to semantically annotate sensors in a sensor network using OWL. One of the remaining problems for SSN is to find ways to cooperate with the Open Geospatial Consortium (OGC) and their Sensor Web Enablement (SWE), which is more focused on XML representations i.e. Sensor Observation Service (SOS) [14], a proposal for sensors as service endpoints, the Sensor Model Language (SensorML) [36], and a host of other protocols.

### 5.4 Linked Data as an Enterprise Architecture strategy

In the previous sections we saw that it was possible to link devices and create applications utilising the linked data both from sensors, stationary or mobile, and big existing linked data archives. In this section I will argue that the idea of using a flexible and standardised data description language, can be seen as an attempt to use ideas from Enterprise Architecture [42]. One of the powerful ideas in Enterprise Architecture is to be able to align the IT development with the strategy of the business. To translate this idea into a research organisation, which can be compared to a virtual enterprise [6], strategy of the organisation is to be on one hand agile in order to be able to attract research funding, and on the other hand be able to reuse the tools for doing research. The tools for doing research are among others the datasets acquired, the literature for finding references, the laboratories, IT networks and databases.

One proposal taken from Enterprise architecture is to describe the situation as it is currently, with respect to IT resources, processes and knowledge, and then compare that with what can be envisioned in light of the business strategy. When the picture of the situation ‘as-is’, and the picture of what is needed in the ‘to-be’ scenario [23], the task is to design the transition of the infrastructure and manpower into the ‘to-be’ scenario.

For the research organisation (as for the Business), it can be difficult to tell what the ‘to-be’ scenario should be. But there are a few items, that can be trusted to be important over an extended period. The ability to access the datasets that the research organisation has collected, and the possibility of combining data from different datasets, could be an important focus point.

Since semantic annotated data in the form of RDF or OWL, promises a standardised way of annotating data, where the semantics of the data is defined in the same language, it might be a basis for evolving research organisations by promoting reuse and cross functional research.

---

<sup>12</sup><http://www.w3.org/2005/Incubator/ssn/ssnx/ssn>

## 5.5 Scientific contributions

I coauthored a paper on testing of Semantic Services: "Test Driven Life Cycle Management for Internet of Things based Services : a Semantic Approach" [35]. In this paper we investigate the how to annotate tests of service endpoint for Internet of Things connected services. The goal is to have a testing infrastructure based on OWL annotation of services, that can cope with rapid changes in a service, to safeguard the quality of services through the lifecycle of the service.

I am working on a paper for "Journal of Enterprise Architecture" with Torben Tambo, AU Herning, to appear Q4 2014. The paper will describe how Enterprise Architecture, can be used to analyse the strategy and goals of a research project, such as EcoSense.

## 6 My Contributions

The main idea in my work so far has to be able to model the emission of climate forcing gases from single trips in passenger cars. In this section I provide a description of the contributions and results I have made towards this goal.

### 6.1 Simple emission model

In this section I will condense the contributions I have made to the EcoSense project. The emission factors for combustion engines, as used in the national emission inventories, are based on measured emissions from standardised test runs [1], and are used for a bottom up estimate of the total national emissions, by estimating the number of kilometres the total national fleet has driven [29]. This estimate is compared to an estimated emission calculated from the amount of gasoline and diesel sold nationally, and the difference between these two estimates leads to a correction factor for the total national emission.

I have developed and implemented a simple model for estimating the emissions from single trips. The trips are divided into sections by the transportation mode detection algorithm. For each transportation mode an emission factor is calculated and multiplied with the length of the section. The total emission for a trip is calculated by accumulating the emissions from the sections. The emission factors for cars, trucks, busses and trains are derived from the reporting of the national emission inventories to UNFCCC under the Kyoto Protocol.

#### 6.1.1 Mobile sensing

When using smartphones as mobile sensing devices, a number of different sensors can be used. In the following the builtin accelerometer and the global positioning sensor are considered. The accelerometer sensor, delivers a timestamped accelerometer measurement, consisting of a reading of the 3 axis accelerometer. The accelerometer is a very low power device and can thus be sampled more frequently ( around 200 Hz ) [22].

The GPS measurements consists of longitude, latitude, altitude, speed, bearing, number of satellites visible, accuracy and a timestamp of the measurement. The GPS sensor contains a radio receiver, and complex logic to decode the radio signals from multiple satellites and calculate a position from the data, thus it uses more electric power. To reduce the power drain on the smartphone battery, the GPS can be turned off between readings of the position.

When calculating the distance travelled between two GPS measurements a Haversine function is used. This function takes the curvature of the earth into account, and therefore gives a more precise value for the distance. The resolution of the test data, does not warrant the use of the Haversine approach, since it is sampled every second, but was chosen in anticipation on more crudely sampled GPS data, due to power conservation considerations.

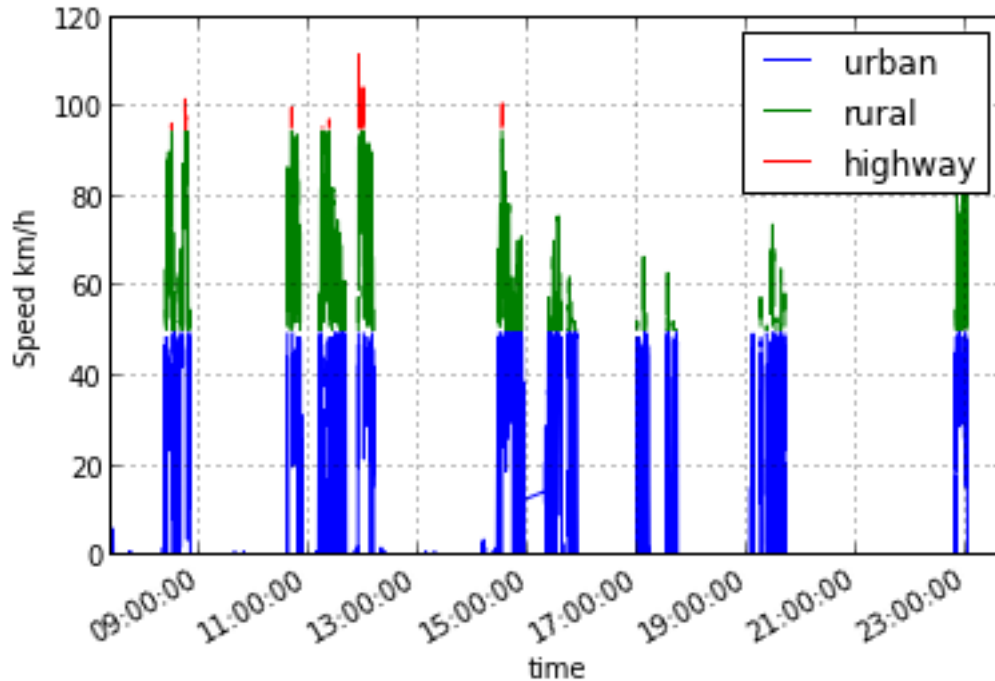


Figure 4: Dividing driving into urban, rural and highway

### 6.1.2 Tier 2 derived model

The emissions are estimated for 3 different driving patterns: Urban, with many stops and low speed. Road driving, with few stops and moderate speeds, and lastly highway, with no stops and high speed. These three modes of road traffic can be distinguished by analysing GPS. In figure 4 an example of a speed curve for a series of trips and how the different driving types the trips are divided into. The distance travelled for each road type can be found by adding the distance between sampling points for points in each driving type. These distances are multiplied with the emission factors from the Tier 2 methodology.

In figure 5 a closer look on figure 4 is shown. From the figure it can be seen that the division with fixed speed limits delivers a quite crude division and gaps where the speed changes across a boundary between road types.

To explore the data, and develop the implementation IPython Notebooks <sup>13</sup> with the numerical and scientific packages NumPy <sup>14</sup>, Scipy <sup>15</sup> and Pandas <sup>16</sup>.

<sup>13</sup><http://ipython.org/notebook.html>

<sup>14</sup><http://www.numpy.org>

<sup>15</sup><http://scipy.org>

<sup>16</sup><http://pandas.pydata.org>

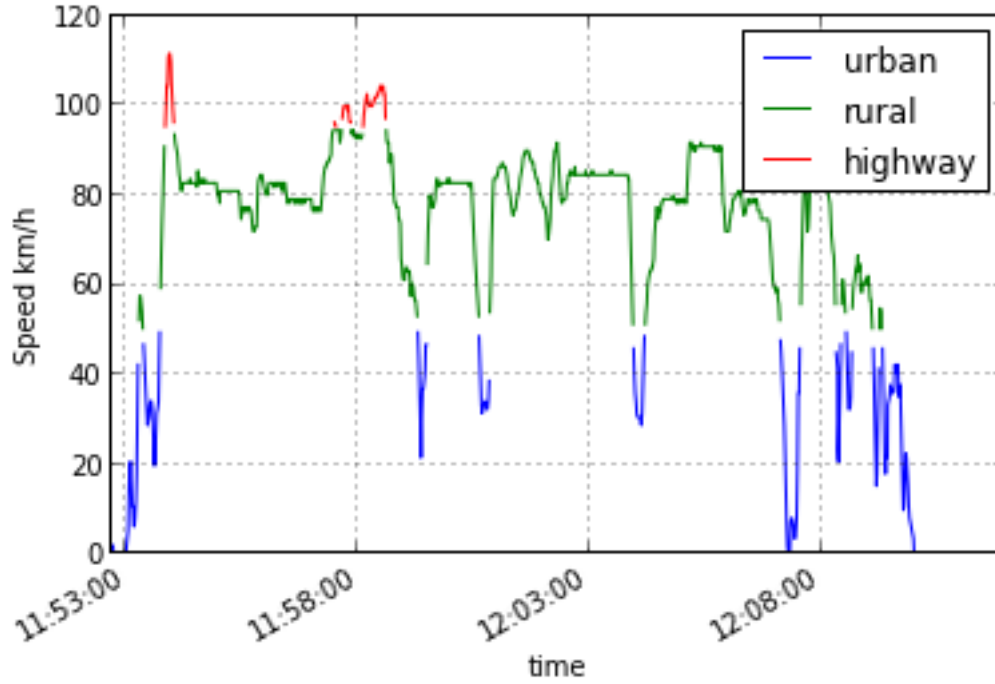


Figure 5: A closer look on figure 4

## 6.2 Enhanced emission model

To create a more accurate estimate of the emission for a single trip by car, a more complex model is proposed. From the sensed data it is possible to get information on the speed of the car at various points of the trip, and the number of stops and starts. This information can be used to further segment the trips into idle, accelerating, cruising and braking; each of these segments would have a different emission profile.

The data sensed from the smartphone can also be used to determine if the engine in the car is cold, if for instance it is the start of the trip. When the engine is cold the combustion and the catalytic converter are less effective, and thus more emissions of Carbon Monoxide, Nitrous Oxide and Volatile Organic Compounds will be higher and the emission factor should reflect that. The determination of the impact of cold starts, demands a quite complex model, taking into account the outside temperature and models for engine warm up [30].

The enhanced emission model is a work in progress. In the following I will go through some of the important parts of the model.

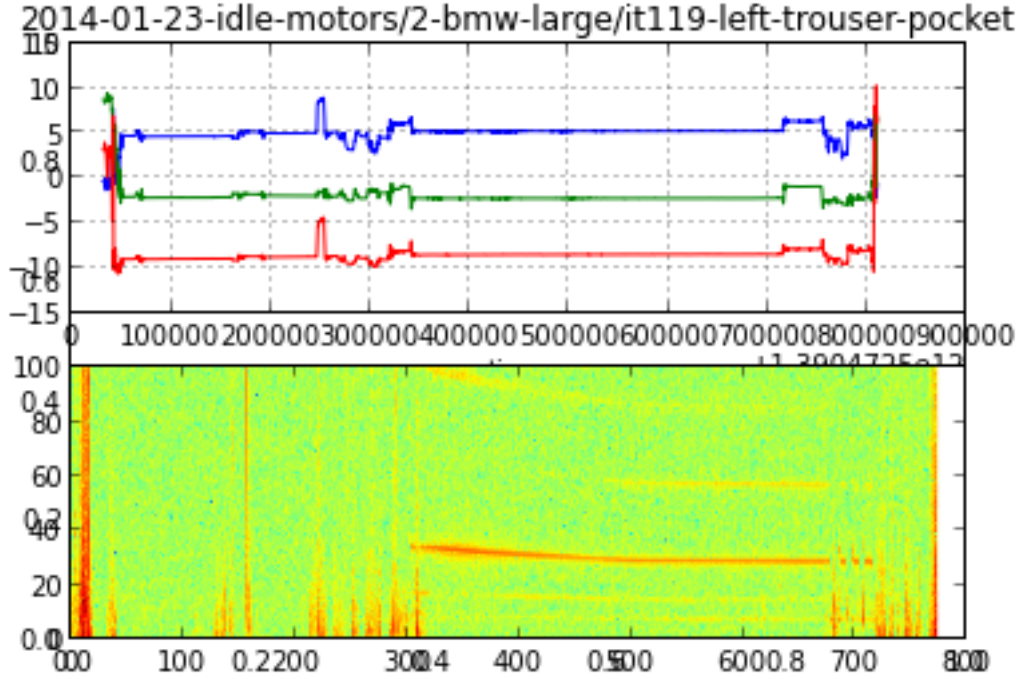


Figure 6: **Engine revolutions.** The top figure shows the three accelerometer signals. The bottom figure is the spectrogram of the size of the acceleration vector

### 6.2.1 Determine Idle engine revolutions

By analysing the frequency spectrum of the accelerometer data, information on the engine revolutions can be retrieved [44]. By employing spectrograms (shows the time evolution of spectral features) changes in engine speed can be visualised. An example (see figure 7) is the idle speed of a cold motor, where the ECU (Electronic Control Unit) of the vehicle, measures the temperature of the motor and as long as the motor is too cold, sets the idle speed at a higher value. The lowering of the idle speed as the engine grows warmer, is clearly visible in the spectrogram. The spectrogram is calculated by calculating the size of the accelerometer vector, i.e.:

$$a = \sqrt{x^2 + y^2 + z^2}$$

By assuming that the smartphone is kept still or only moved slowly, the gravity can be considered constant. The gravity will then be present as a signal with frequency equal to zero, which can easily be removed, as it is the zero'th element in the array returned by the Fast Fourier Transform (FFT). So even though the gravity is a very large signal and the signal stemming from vibrations in the engine are quite faint, it is possible to filter the gravity out [13].



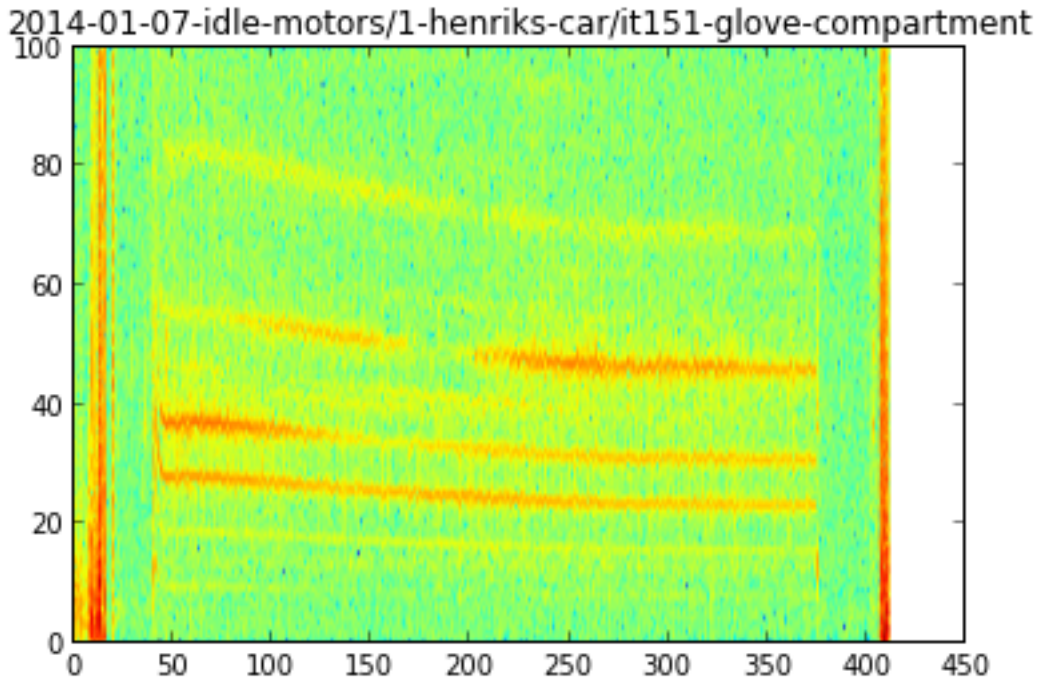


Figure 7: **figure showing rotational speed of an cold engine. As the engine heats up the engine speed lowers. The yellow and red colours indicates larger powers in that frequency range.**

The spectrogram shows the time evolution of the frequency spectrum of a signal, so the y-axis shows the frequencies, and the x-axis shows the time. The size of individual frequencies is shown by using different colours. A constant frequency will result in a horizontal line in the spectrogram, and slowly varying frequencies will result in curves in the spectrogram. The usefulness of the spectrogram is to show signals with few frequencies that change slowly, and not to ascertain the exact size of individual frequencies. The spectrogram can therefore be said to be a qualitative display technique.

An example of a spectrogram of a car in idle mode is shown in figure 6. The slowly decreasing red line coincides, timing wise, with the engine starting. As the engine warms up the engine speed can be seen to slow. The actual frequency of the red line indicates that it probably is a harmonic of the basic frequency.

During driving, the vibrations from of the moving car, that is the vibrations from the tires moving over the road and vibrations from the suspension system, drowns the signal of the engine rotational speed.

Another effect that contributes to the loss of a clear signal for the rotational speed of the engine is due to the low sampling frequency of the accelerometer. The sampling frequency of the accelerometer is approximately 200 Hz. This means that the data from

the accelerometer only can represent signals with frequency below 100 Hz, due to the Nyquist criteria. This maximum frequency of 100 Hz corresponds to a engine speed of 6000 rpm. Thus the sampling frequency only allows for accurately representation of frequencies below 6000 rpm, and if higher vibrational frequencies exist these frequencies will be folded back into the spectrum from 0 to 6000 rpm, a process called aliasing [24]. Since combustion engines often employ more than one cylinder, and since the process of combustion takes place as a violent process in the cylinder, it is reasonable to expect harmonics of the engine speed to be created. When the engine is in idle mode a rotation of 900 - 1000 rpm (15 - 16,7 Hz) is normal and the sampling of the accelerometer will allow the correctly represent up to the sixth harmonic, but when the vehicle is driving engine rotation speed between 2000 - 3000 rpm would be appropriate, thus only allowing to represent third or only second harmonics. Any higher order harmonics will be aliased back into the spectrum and contribute to the noisy picture.

As a result of vibrational noise from the moving vehicle and aliasing of engine vibration (and aliasing of the vibrational noise for that matter) the signal from the engine rotation is lost in the noise. In signal analysis, the way to prevent aliasing is to apply a low pass filter to remove or dampen frequencies above half the Nyquist frequency. But in our application it is not possible to ensure that the smartphone containing the accelerometer is adequately shielded from high frequency vibrations.

One method used for exploration of the data from the accelerometer, was to convert the signal into an audio file and listen to different types of transportation. By playing the audio file at a higher sampling frequency than the accelerometer sampling frequency, low frequency tones are more easily recognised as their frequency will be multiplied by the increase in sampling frequency. The time used for listening is shortened with the same factor. A factor of 5 was used to examining the data. Since the human hearing is quite adept in separating single frequencies, it is a good way to quickly ascertain if patterns are available in the data.

From the datasets used in this study, it can be concluded, that it is possible to detect idle mode, for vehicles with combustion engines. Other types of vehicles, like electric vehicles or vehicles that stop their engine when the vehicle stops, would not need an idle mode detection.

### **6.2.2 Determine acceleration and turning**

In order to distinguish between the four modes of driving (idling, accelerating, cruising, decelerating), which have different emission characteristics, we can use the accelerometer data to determine the acceleration patterns in the horizontal plane.

The data from the accelerometer is 3 values for each measurement, which are the 3 axis from the accelerometer. The direction of movement of the phone can be determined by analysing these 3 signals. First the direction of gravity is determined. Since it is possible for the phone to be moved freely in all directions and/or flipped with the movement of the person wearing the phone, the gravity detection needs to be dynamic [13]. When the direction of gravity has been determined, it is possible to detect acceleration in the horizontal plane. This allows us to consider distinguishing turning from linear ac-

celeration, which is important because a linear acceleration will have a different emission factor than turning.

### **6.2.3 Combine turning information with GPS**

By combining the turning information, with GPS localisation and map data, the accuracy of the positioning of the car can be improved. Improving the positioning accuracy is important for improving the accuracy of the calculation of trip length and segment of trips length.

The more accurate trip emission estimate would be of interest for green accounting purposes.

The combination of turning information with GPS location will also improve the detection of acceleration and deceleration driving modes, and thus improve the emission estimate.

To validate and compare the models the exact emissions can be measured by inserting a sensor into the exhaust pipe of the vehicle under test. By doing test runs in different kinds of traffic, and with different vehicles the results of the sensor can be compared with the results of the models see section 7.

By combining the GPS data and the accelerometer data, we can detect cruising mode, and maybe acceleration/deceleration.

## 7 Future Research

In this section the planned work for the remainder of the phd study is discussed. The focus of the section is the planned scientific contributions.

### 7.1 Hypotheses

In this section I will present the hypotheses, that I will investigate in the remaining study period. The overarching idea is still to use data from smartphones to

1. Idle mode of passenger cars can be robustly detected by using mobile sensing.
2. Cruising of passenger cars can be robustly detected by using mobile sensing.
3. Acceleration and deceleration can be robustly detected by using mobile sensing.
4. Single trip estimation of airborne climate gasses can be made more accurately than IPCC methodology.
5. Correlation of multiple trips by different facets (time, location, transport mode) can improve transport planning.
6. Local and Regional emission inventories can be created by using measurements from a few representative users.
7. Models for pollutants dispersion can be improved by using mobile sensing of passenger transport.

### 7.2 Improving estimation of emissions for single trips

As outlined in section 3, the current modelling focus is to create and maintain national inventories, which leads to a focus on mean values for emission factors, mean driving patterns for the vehicle fleet and mean trip patterns. To be able to provide personalised information on specific transportation behaviour, there is a need to provide more detailed models for the emission of single trips. In this section an outline for possible algorithms for reaching that goal is presented.

One proposal is to divide a trip into four different types of driving : Idle, accelerate, cruise and decelerate. For each type of driving a emission profile can be derived and thus the emission for each type can be determined. The total emission for a trip can then be determined as the sum of the emissions for each type.

#### 7.2.1 Idle emissions

Detection of idle situations can be done with a combination of GPS data and accelerometer data. The GPS data can be used to estimate the speed of the vehicle, and the accelerometer can be used to measure the engine speed to confirm that we are in idle mode. There are some literature about measuring emissions from idling [9][34], but it

might be necessary to update with new measurements. A possible source for idle emission data could be the approval data for Danish biannual vehicle inspections, since part of the inspection is a measurement of the contents of the exhaust in idle mode.

### **7.2.2 Emissions when accelerating**

The horizontal acceleration can be determined by finding the direction of gravity in respect to the device, through a variety of methods. These methods will have to be evaluated to find a suitable solution for the application at hand. When the gravity direction has been determined, the horizontal acceleration will be either close to zero, when cruising or idle, or have a significant value due to acceleration, turning or deceleration. It is believed that it will be possible to provide a stable algorithm for detecting horizontal acceleration and distinguish between turning and acceleration.

To model the emission from an accelerating vehicle, information, such as engine size and vehicle weight, is needed. This information has to be given as input to the model by the user, or be inferred from the Transportation Mode Detection part of EcoSense.

Another input to the model could be the road grade, since the engine will have to work harder, thus emitting more pollutants, if the vehicle is going uphill [20]. By using the GPS data to get information on the position, the road grade can be gleaned from a digital road network. By fusing the information from these different sources the emission modelling can be further improved.

### **7.2.3 Emissions when decelerating**

When decelerating, there are a couple of different situations to be aware of. The simplest situation is when the vehicle is braking using the mechanical brake. In this situation the engine will typically be in idle mode and the results from section 7.2.1 can be reused. If the vehicle incorporate regenerative braking, motor braking or automatic transmission the situation is more complex. The proposal is to first ascertain if deceleration can be detected and then in an first approximation use the results from section 7.2.1

### **7.2.4 Emissions when cruising**

In the study of emission models, models for speed dependency of emissions have been found. These models can be used as is if we can determine that we are moving at a constant speed. These models are described in section 3. The models are developed for emission modelling programs such as COPERT IV, which was developed as part of the EU project ARTEMIS. The models used in COPERT IV can be used to assign a speed dependent emission factor to specific vehicle types, engine sizes and fuel types.

### **7.2.5 Papers**

A position paper on the state of the art of modelling emission from mobile sensing is planned.

More specialised papers detailing and evaluating the improvements made in this study is also planned.

### **7.3 Correlation of trips**

In order to be able to spot inefficiencies in transportation patterns, a way to group, aggregate and correlate different trips is needed. The grouping of trips could be by persons, time of day, seasonal or geographic.

The aggregation could be looking for all trips at specific location in a certain time period. Correlation is useful for finding trips which follow a certain route.

In order to solve these problems efficiently, some heuristics may be useful. If a digital road network is available for the area under consideration, each trip can be converted into a subgraph of the digital road network, under the assumption that vehicles travels along the roads. By having the trips as a graph instead of a time series of GPS locations, will simplify the task of correlation, thus good and efficient algorithms to convert GPS traces to road network graphs is needed. The problem is then transformed in to a graph search problem, for which well understood algorithms exists [28].

### **7.4 Improve dispersion models**

The modelling done in programs like DASK and OSPM can be further improve, by adding more calculation of chemical reactions between pollutants, or adding more usable interfaces. DASK is run from the command line, which is very flexible, but also demands somewhat experienced users. OSPM is a windows program with a complex graphical user interface, and can be crashed quite easily, without any error messages to inform the user of what went wrong.

### **7.5 Impact**

One of the envisioned impacts of this research is to enable a lightweight approach to managing the task of green accounting for transportation. By using the data from a company's smartphones a comprehensive and detailed picture of the emission profile of the company's transportation needs can be constructed. This can lead to better and more efficient use of the internal transportation, incentive programs for car sharing among the employees or other emission lowering activities, and thus lower the emissions of the company. As long as the emission profile is unknown or only known with a large uncertainty, it can be hard to find the areas, where a targeted effort will have the largest outcome, or even see the purpose of doing anything at all.

Another possible impact of the phd study is in traffic planning. Today the traffic planning is based on counting of vehicles at specific points, done by laying out tubes on the road and counting the number of passages over the tube. In this work it is possible to tell not only the count of vehicles at a specific point, but also where they are coming from and where they are going. This more detailed view of the traffic flow might give traffic planers a new tool to control the capacity of the traffic network or to find new

opportunities for public transport, as in [3] where phone call records are used to track mobility and guide planning of bus routes.

The possibility to use only a small sample to describe a local traffic system, can also be applied to improve local air pollution modelling. Today traffic is input as yearly mean values, and modelled with daily fluctuation distribution (to model rush hours). More precise description of local and instantaneous traffic patterns can improve the detailed air pollution description, as well as improve the overall emission inventory.

### **7.5.1 Papers**

A paper documenting the approach and detailing the possible benefits in terms of finding opportunities for minimising the emission of pollutants through better route planning. I will attend a course at DTU Transport during the autumn concerning "Route Choice Models", to prepare for this.

## **7.6 Field trials**

Among the partners in the EcoSense project are some municipalities. The interest from the municipalities in the EcoSense project are among other, to be test sites for the results coming out of the project. In return the research partners get real world data to exploit in future research.

### **7.6.1 Sønderborg**

In the municipality of Sønderborg, there has been a long tradition for doing projects with focus to mitigate the threat of Climate Change through minimising the emission of climate forcing gasses. The task of finding ways of minimising the emission for Sønderborg, has been put to an organisation called "Project Zero". "Project Zero" initiates and run campaigns for raising the awareness of climate gas emissions, as well as measure the effects of these campaigns. "Project Zero" has campaigns that target citizens and other campaigns that target the local industry.

As "Project Zero" is a partner in EcoSense, we have the possibility to make mobile applications that underpin or are a part of the campaigns "Project Zero" runs.

### **7.6.2 Herning**

In Herning there is a project called "Herning cykler til månen", Herning bikes to the moon. The goal of the project is to get the citizens of the municipality to bike the distance to the moon. To measure the distance travelled on bike, participants can download a mobile application, developed in the EcoSense project. The app sends data to the EcoSense servers, and we thus get the possibility to look into real world transportation. The application will log the data for all transportation, not only for cycling, but the cycling data will be filtered out to serve the goal of the application.

For research purposes we can see if it is possible to find patterns in the traffic (both bicycle and automobile), although the data will be biased towards people who use bikes.

To get information on congestion in traffic it seems that you only need a very small sample of the total number of vehicles [15], so even if the data is biased towards bicycle users, there might still be opportunities for interesting discoveries.

The data obtained from the experiment, can be converted into input data to the pollution modelling program OSPM, to estimate the pollution level at different points in the city. These pollution estimates can then be evaluated by actual measurement. If combined with traffic count data, done by ordinary methods, comparisons and validation of the approach can be appraised.

Another opportunity for discovery in this data set is concerning bicycle paths. Herning contains many bicycle paths and bicycle lanes. From the data we might be able to tell the municipality which bicycle paths are used and create the data to visualise the use patterns.

Further experiments that can be conducted is to perform surveys of the users of the application, since an email address is needed to enable the application.

An air pollution measurement station is in development, by students and faculty members of AU Herning. It is planned to measure the air pollution at different points in Herning. The measurements will be used in conjunction with the planned modelling in the OSPM system.

### **7.6.3 Elbil parat**

EcoSense is conducting a small pilot project with its partner Insero Business Solutions. By downloading an EcoSense made application, people with an interest in trying a electric vehicle can get a picture of their daily transportation need. This can help the volunteer to see which part of their daily transportation could be performed by an electric car, and which part should be covered by other forms of transportation. When the transportation need has been analysed, the volunteers has the opportunity to lend an electric vehicle for a limited period.

The data gained from this experiment, is probably biased towards people interested in electric vehicles, but it is not clear if the traffic needs of this group of people is significantly different from other groups. Of course the people that lend an electric vehicle will have needs that can be served by a small electric vehicle.

### **7.6.4 Papers**

It would be reasonable to document the results of this real world (or at least outside the laboratory) trials in a number of papers. Each trial has different goals but share a number of methods. It seems reasonable to expect at least one paper for each facet of each of the trials. The danger here is that only the Herning trial has started with only a small number of participants. It is expected that further promotion of the project by the municipality will increase the attendance of people in the project, but it is still unknown if a reasonable number of participants will get involved and use the app, so that interesting data can be obtained.



## **7.7 Visiting researcher**

During the last 2 years of the study I plan to visit an university, to participate in the research as it is performed there. I plan to have the stay at another research group during the teaching free period in the summer of 2015. I am currently actively looking for suitable research groups in other transportation research, air pollution modelling or geographic data research.

## 8 Conclusion

In this multidisciplinary study I have learned from many different fields. I have learned from social and political scientists about how to assess climate change policy and the processes in the United Nations organisations devoted to climate change policy design. From chemistry I have learned about atmospheric chemistry and how to model dispersion of airborne pollutants. I have used signal analysis, and knowledge on digital signal processing to interpret measured data. From Environmental science I have learned detailed methodology for estimating emission of pollutants and from computer science I have learned about mobile sensing, database design and Enterprise Architecture. The first 2 years of the phd study has primarily been used to learn about the fields, and find ways to apply the knowledge in the context of the EcoSense project. The main scientific contributions are

1. Coauthoring a paper on testing of Semantic services
2. Development and implementation of a simple model for estimation of single trip emissions
3. Development of more precise model for emission estimation, using actual vehicle speed, instead road type, as emission factor

The future work as described in section 7, will be focused on filling the gap identified in the methodology for estimating the emission of airborne pollutants from traffic, and evaluating the above listed contributions. The gap exists between the precise methods of measuring the emissions directly from the tailpipe, by instrumenting the vehicle with extra sensors, and the use of average values for activity data and engine data. With mobile sensing the activity data can be more accurate than the mean values estimated for activity data in COPERT IV and IPCC methodologies.

The future improvement in single trip estimation of emission of pollutants, will be based on more precise measurements of the four driving modes : Idle, acceleration, cruising and deceleration. To get more precise information on these driving modes, different combinations of the data from the different sensors will be used.

The search for route patterns will be performed, when data from the field trials, will be available. To prepare for this work, I will participate in the course "Route choice models" at DTU Transport in fall of 2014.

Extensions to the pollution dispersion models, with respect to new chemical reactions, is planned for the coming study period.

To sum up the the possible impact of this phd study:

1. Improve green accounting for passenger transportation.
2. Facilitate building of regional and local emission inventories.
3. Improve the calculation of the impact on Climate Change from passenger transportation.

4. Enable air pollution guided traffic planning, due to better description of local air pollution.

I find that the possibility to be able to create emission inventories for a municipality or a company , even if there is no complete coverage of mobile sensing devices, to be quite interesting, and I look forward to further improve the methods describe in this progress report.

## **Acknowledgement**

I would like to thank my supervisors Niels Olof Bouvin and Allan Gross, as well as Henrik Blunck for the interest, comments, suggestions and great engagement in my phd. study.

A big thank you goes to Markus Wüstenberg, for providing the rich dataset, the study is based on.

I would also like to thank my family for their patience and support.

This work has been supported by the Danish Council for Strategic Research as part of the EcoSense project (11-115331).

## References

- [1] M. André. The ARTEMIS European driving cycles for measuring car pollutant emissions. *The Science of the total environment*, 334-335:73–84, Dec. 2004.
- [2] R. Berkowicz. OSPMa parameterised street pollution model. In *Urban Air Quality: Measurement, Modelling and Management*, pages 323–331. Springer, 2000.
- [3] M. Berlingerio and F. Calabrese. AllAboard: a system for exploring urban mobility and optimizing public transport using cellphone data. In *Machine Learning and Knowledge Discovery in Databases*, pages 663–666. 2013.
- [4] T. Berners-Lee, R. Fielding, and L. Masinter. Uniform Resource Identifiers (URI): Generic Syntax. *IETF Request for Comments*, 1998.
- [5] F. Calabrese, K. Kloeckl, and C. Ratti. Wikicity: Real-time location-sensitive tools for the city. *Handbook of research on urban informatics: The practice and promise of the real-time city*, pages 1–16, 2008.
- [6] L. Camarinha-Matos and H. Afsarmanesh. Towards an architecture for virtual enterprises. *Journal of intelligent Manufacturing*, 1998.
- [7] D. Chakraborty and J. Roy. Energy and carbon footprint: numbers matter in low energy and low carbon choices. *Current Opinion in Environmental Sustainability*, 5(2):237–243, June 2013.
- [8] J. Christensen. The Danish Eulerian hemispheric modelA three-dimensional air pollution model used for the Arctic. *Atmospheric Environment*, 31(24):4169–4191, 1997.
- [9] P. de Haan and M. Keller. Emission factors for passenger cars: application of instantaneous emission modeling. *Atmospheric Environment*, 34(27):4629–4638, Jan. 2000.
- [10] H. Frey and A. Unal. Use of on-board tailpipe emissions measurements for development of mobile source emission factors. *Proceedings of US Environmental Protection Agency Emission Inventory Conference*, pages 5–7, 2002.
- [11] L. M. Frohn, J. H. Christensen, and J. Brandt. Development of a High-Resolution Nested Air Pollution Model. *Journal of Computational Physics*, 179(1):68–94, June 2002.
- [12] A. Gross and W. Stockwell. Comparison of the EMEP, RADM2 and RACM mechanisms. *Journal of Atmospheric Chemistry*, (Mcm):151–170, 2003.
- [13] S. Hemminki, P. Nurmi, and S. Tarkoma. Accelerometer-based transportation mode detection on smartphones. *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems - SenSys '13*, pages 1–14, 2013.

- [14] C. a. Henson, J. K. Pschorr, A. P. Sheth, and K. Thirunarayan. SemSOS: Semantic sensor Observation Service. *2009 International Symposium on Collaborative Technologies and Systems*, pages 44–53, 2009.
- [15] J. Herrera, D. Work, and R. Herring. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*, 2010.
- [16] O. Hertel, R. Berkowicz, J. Christensen, and O. Hov. Test of two numerical schemes for use in atmospheric transport-chemistry models. *Atmospheric Environment. Part A. General Topics*, 27(16):2591–2611, 1993.
- [17] I. Horrocks, P. F. Patel-Schneider, D. L. McGuinness, and C. A. Welty. OWL: a Description Logic Based Ontology Language for the Semantic Web. *The Description Logic Handbook*, pages 450–486, 2007.
- [18] I. Horrocks, P. F. Patel-Schneider, and F. van Harmelen. From SHIQ and RDF to OWL: the making of a Web Ontology Language. *Web Semantics: Science, Services and Agents on the World Wide Web*, 1(1):7–26, Dec. 2003.
- [19] S. Jensen, M. Ketzel, J. Brandt, M. Plejdrup, O.-K. Nielsen, M. Winther, O. Evdokimova, and A. Gross. Manual for THOR-AirPAS - air pollution assessment (To be published). Technical report.
- [20] M. Kaul, B. Yang, and C. S. Jensen. Building Accurate 3D Spatial Networks to Enable Next Generation Intelligent Transportation Systems. *2013 IEEE 14th International Conference on Mobile Data Management*, pages 137–146, June 2013.
- [21] P. Kelly and W. Adger. Theory and practice in assessing vulnerability to climate change and Facilitating adaptation. *Climatic change*, pages 2–5, 2000.
- [22] M. B. Kjærgaard, J. Langdal, T. Godsk, and T. Toftkjær. EnTracked : Energy-Efficient Robust Position Tracking for Mobile Devices. *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 221–234, 2009.
- [23] A. Lapkin, P. Allega, B. Burke, and B. Burton. Gartner Clarifies the Definition of the Term Enterprise Architecture. *Report ID G*, (August), 2008.
- [24] R. G. Lyons. *Understanding digital signal processing*. Pearson Education, 2010.
- [25] D. L. McGuinness, F. Van Harmelen, and Others. OWL web ontology language overview. *W3C recommendation*, 10(10):2004, 2004.
- [26] G. Mellios, S. Hausberger, and M. Keller. *Parameterisation of fuel consumption and CO2 emissions of passenger cars and light commercial vehicles for modelling purposes*. JRC EU Publications Office, 2011.

- [27] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda. {PEIR}, the Personal Environmental Impact Report, As a Platform for Participatory Sensing Systems Research. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*, {MobiSys} '09, pages 55–68, New York, {NY}, {USA}, 2009. ACM.
- [28] M. E. J. Newman. The structure and function of complex networks. *SIAM review*, 45(2):167–256, 2003.
- [29] O.-K. Nielsen, M. Plejdrup, M. Winther, M. Nielsen, S. Gyldenkærne, M. Mikkelsen, R. Albrektsen, M. Thomsen, K. Hjelgaard, L. Hoffmann, P. Fauser, H. Bruun, V. Johannsen, T. Nord-Larsen, L. Vesterdal, I. Møller, O. Caspersen, E. Rasmussen, S. Petersen, L. Baunbæk, and M. Hansen. Denmark’s National Inventory Report 2014. (101), 2014.
- [30] L. Ntziachristos, Z. Samaras, C. Kouridis, D. Hassel, I. McCrae, J. Hickman, K.-h. Zierock, M. Andre, M. Winther, N. Gorissen, P. Boulter, R. Joumard, R. Rijkeboer, S. Geivanidis, and S. Hausberger. *EMEP/EEA emission inventory guidebook 2009, updated May 2012 1*. Number May. European Environment Agency (EEA), 2012.
- [31] R. Pielke, G. Prins, S. Rayner, and D. Sarewitz. Lifting the taboo on adaptation. *Nature*, 445(February):597—598, 2007.
- [32] R. a. Pielke. Rethinking the role of adaptation in climate policy. *Global Environmental Change*, 8(2):159–170, June 1998.
- [33] E. Prud’hommeaux, A. Seaborne, and Others. SPARQL query language for RDF. *W3C recommendation*, 15, 2008.
- [34] M. Pujadas, L. Núñez, J. Plaza, J. C. Bezares, and J. M. Fernández. Comparison between experimental and calculated vehicle idle emission factors for Madrid fleet. *The Science of the total environment*, 334-335:133–40, Dec. 2004.
- [35] E. S. Reetz, K. Daniel, T. Ralf, and A. Lehmann. Test Driven Life Cycle Management for Internet of Things based Services : a Semantic Approach. In *VALID 2012, The Fourth International Conference on Advances in System Testing and Validation Lifecycle*, pages 21–27, 2012.
- [36] D. Russomanno, C. Kothari, and O. Thomas. Building a Sensor Ontology: A Practical Approach Leveraging ISO and OGC Models. *IC-AI*, 2005.
- [37] E. Schipper and I. Burton. Understanding adaptation: origins, concepts, practice and policy. *The Earthscan reader on adaptation to climate change.*, (19), 2009.
- [38] L. Schipper, I. Burton, and Others. Earthscan Reader on Adaptation to Climate Change. 2008.

- [39] J. H. Seinfeld and S. N. Pandis. *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons Inc., 1997.
- [40] J. Smithers and B. Smit. Human adaptation to Climatic Variability and Change. *Global Environmental Change*, 1997.
- [41] T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P. M. Midgley, and Others. Climate Change 2013. The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change-Abstract for decision-makers. Technical report, Intergovernmental Panel on Climate Change-IPCC, Geneva, 2013.
- [42] T. Tamm and P. Seddon. How does enterprise architecture add value to organisations. *Communications of the Association for Information Systems*, 28(March 2011):141–168, 2011.
- [43] D. Waldron and Others. IPCC Guidelines for National Greenhouse Gas Inventories—Chapter 3: Mobile Combustion, Volume 2: Energy. *Institute for Global Environmental Strategies, Hayama, Kanagawa, Japan*, 2006.
- [44] M. Wüstenberg, H. Blunck, K. Grønbæk, and M. B. Kjærgaard. Distinguishing Electric Vehicles from Fossil-Fueled Vehicles with Mobile Sensing. *To appear at MDM'14*, 2014.
- [45] Z. Zlatev. *Computer treatment of large air pollution models*. Kluwer Academic Publishers, 1995.