Bachelor Thesis:

Data fusion for instantaneous travel time estimation

Loop detector data and ETC data

Author:
Do, M.

University of Twente
Enschede (The Netherlands), July 3rd 2009

Internship at the University of Tokyo
Kuwahara Lab., March 17th 2009 till June 5th 2009

Supervisors:
Marc Miska
Rattaphol Pueboobpaphan

Preface

As part of my study program Civil Engineering and Management at the University of Twente, I conducted this internship research at Kuwahara Lab. (University of Tokyo). This research was carried out in a period of about 12 weeks, starting in early March till the end of May.

Since I had a little delay in my study, I was considering doing an internship abroad. My origin is from East-Asia, so this part of the world has always been interesting to me. After my second year of my Bachelor study I knew for sure that traffic is the direction I want to go to for my Master.

After some thinking I decided to search for a traffic related internship abroad. After talking to a few people on my university I came in contact with Mr. Bart van Arem. He happened to know someone in Japan (Mr. Marc Miska) and quickly arrangements started for an internship at Kuwahara Lab (University of Tokyo).

Before going to Japan, I had to decide on a research topic and prepare a research plan. At my own university Mr. Rattaphol Pueboobpaphan assisted me with this and from Tokyo Mr. Marc Miska was available.

The first topics that came up were related to signalized intersections. But as it turned out that the traffic control signals in Europe and Japan were very different. My research objective to improve actuated traffic controls changed into travel time estimation using loop detector data and ETC data. Although traffic signal controls in Japan not actuated, the high-tech of Japan can be found in the ETC systems that have been running there for years already.

I really enjoyed my stay in Japan. The public transport system is amazing, I was able to visit many place in Japan thanks to that amazing system. My favorite place in Japan is probably Kizaki Lake, located in Nagano prefecture. During my stay in Japan I really got into the Japanese culture, people are very kind and polite, food is amazing, and life is busy.

As for my internship, Kuwahara Lab was a very pleasant place to work at. Not only people start to work after 10:00 AM, it is located near Shibuya which is very convenient. During my internship I learned many things, for example how to work with huge amounts of data and how to eat lunch within 30 minutes. Seriously, people there don’t waste much time on eating!

I would like to thank Mr. van Arem, Mr. Rattaphol Pueboobpaphan, and Mr. Marc Miska for making this internship possible and for the support during my internship. Thanks to Mr. Masao Kuwahara and Mr. Babak Mehran for supporting my research. Special thanks to Mrs. Kiyoko Morimoto for the administrative work and support you managed for me. Actually, thank you to whole Kuwahara Lab for a great time there. Finally thank you Mrs. Ellen van Oosterzee for the administrative work in the Netherlands for me.

Michael Do
‘s-Hertogenbosch, July 3rd 2009

Summary

With the emergence of Advanced Traveler Information Systems (ATIS), it is possible to provide various kinds of information to road users. Travel time is one of the most understood measures for road users. By providing reliable travel time estimates it is possible to influence road users' route choice and travel behavior, hence improving the performance of traffic networks.

The goal of this research is to develop a data fusion between loop detector data and ETC (Electronic Toll Collection) data to make more accurate real-time (instantaneous) travel time estimates on expressways. Unlike previous attempt of data fusion, this research will not use historical and statistical analyses for data fusion. By relying on statistical methods, the models fail to take traffic engineering principles into account. And by using historical data the developed models can only be applied at locations where historical data is available. Problem here is when there is a change in the traffic network, it needs to be examined if historical data before the change can still be used as input for the model.

Loop detectors are the most common vehicle detectors for freeway traffic, these sensors continuously measure traffic speed and flow. This makes detectors very suitable for instantaneous travel time estimation, providing expected travel time to vehicles entering the expressway. But loop data does not provide an accurate image of the traffic conditions. This is because the detectors only collect data at point-locations and not over the entire road.

ETC data on the other hand gives measured travel times over the entire road, vehicles's location and times are being registered when they enter and leave a toll area. Disadvantage of this data is that it becomes available after that the travel time has been realized, while the goal is to provide estimations to vehicles at the beginning of their travel.

The study area for this research is the metropolitan expressway (MEX), route #4, leading from Takaido towards the Tokyo ring (Miyake-zaka Junction). Length of the area is about 14 km. Since the detector placement in this study area is very dense, about every 100 meters, for the data fusion not all detector data will be used. Only data from 4 sections will be used, this will make the research more representative for the European and American road conditions (concerning detector density).

The Miyake-zaka Junction connects route #4 with the ring-road in Tokyo, during peak hours this ring is heavily congested. Travel time over this route in normal (free-flow) condition is about 6 minutes, while during congestion the travel time can exceed the 20 minutes. The further away from the ring, the less the congestion gets. This makes route #4 an interesting study area. Since it goes towards a congested area, there will be various traffic conditions on the route.

For this area aggregated loop data (speed, flow, and occupancy) for each segment for every 5 minutes is available. Data from each segment is aggregated from three dual-loop detectors. Pulse data from the individual detectors and data per lane was not available. As for ETC data, entering and exiting time and locations for individual vehicles were registered. All data was from the period of July 1st 2006 till July 7th 2006, ETC market penetration at this period was about 60%.

To evaluate how accurate estimates could get based on loop data only, a time slice model was examined. This model is more suited for historical travel time analyses, because for each segment this model determines a vehicle's entering time and based and that data from the corresponding time-interval us used for estimating travel times. By using the data corresponding to the same time-interval a vehicle is traversing a segment, this model takes speed variations over time into account. In case this model is applied for real-time applications, a delay has to be taken into account. Just like ETC data, this model gives travel times after the actual travel time has been realized.

Throughout this research several fusion concepts were examined. The first one examined was a model running two models parallel, the Extrapolation and the Nam and Drew. By integrating ETC data previous time-intervals were evaluated and based on the previous intervals an estimate for the

current interval would be calculated with the estimates of the Extrapolation model and the Nam and Drew model.

The corrections for this model are illustrated in Table 1. Parts of the travel time estimates graphs and ETC graphs are plotted, identification of the situation, error determination, and correction are demonstrated. The yellow dotted line is the ETC data, the green line is the Nam and Drew model, the red line is the Extrapolation model, the blue dot is the corrected estimation.

1. Rule #1, the last two intervals with ETC data available were overestimated by one model and underestimated by the other model. The travel time estimate for the current interval is assumed to be in between of the two models. The model with the lowest output will be corrected upwards based on errors in previous intervals.
2. Rule #2, the last two intervals with ETC data available were underestimated by both models. Current estimate is assumed to be underestimated and the Extrapolation model will be corrected upwards based on errors in previous intervals.
3. Rule #2, the last two intervals with ETC data available were overestimated by both models. Current estimate is assumed to be overestimated and the Extrapolation model will be corrected downwards based on errors in previous intervals.

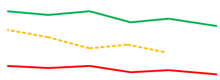
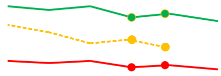
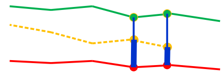
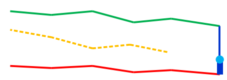
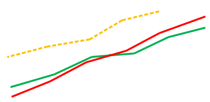
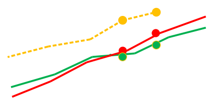
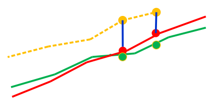
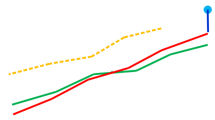
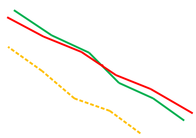
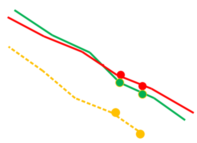
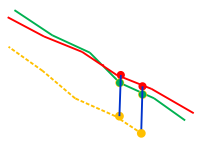
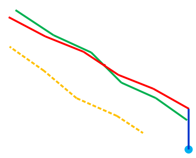
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				
#2				
#3				

Table 1 – Illustrations of corrections for first fusion model

The second concept only uses one existing estimate model as basis, the Extrapolation model. The ETC data is used to evaluate the error in previous time intervals. Based on the current travel time estimate trend, either ascending or descending, travel time would be corrected assuming that the previous error is still present in the current interval.

Illustrations of the correction methods of the second model are shown in Table 2. The yellow dotted line is the ETC data, the red line is the Extrapolation model, the blue dot is the corrected estimation.

1. The last two estimates by the Extrapolation model are ascending. Current estimate is assumed to be underestimated and will be corrected upwards based on errors in previous intervals.
2. The last two estimates by the Extrapolation model are descending. Current estimate is assumed to be overestimated and will be corrected downwards based on errors in previous intervals.

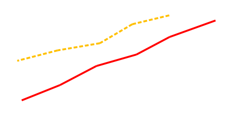
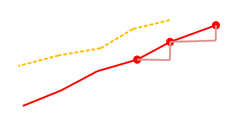
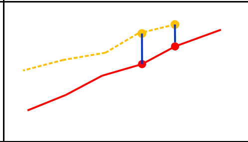
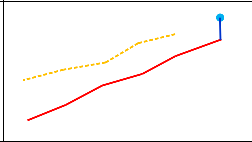
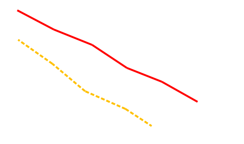
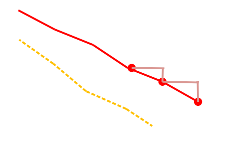
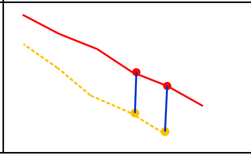
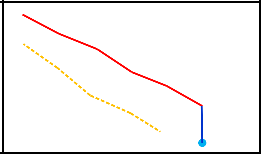
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				
#2				

Table 2 – Illustrations of corrections for second fusion model

The last concept examined in this research is very similar to the second. A moving average on the Extrapolation model was introduced to stabilize the output, which is used to identify traffic conditions. Without the moving average, the output was too instable to be used for identifying traffic conditions.

Because of time constrains only one correction rule was made for this model, see Table 3 for illustration.

1. The last two estimates by the Extrapolation model were first ascending followed and than descending. Current estimate is assumed to be overestimated and will be corrected downwards based on errors in previous intervals.

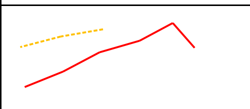
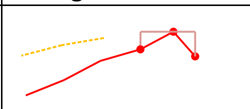
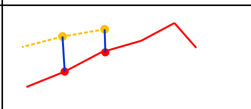
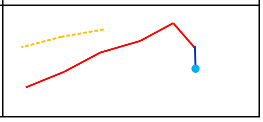
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				

Table 3 – Illustrations of correction for third fusion model

Out of the three examined concept, the first and the last concepts are successful fusions. The first method was only tested with all loop detectors as input for the instantaneous model. It quickly turned out that running two models in parallel complicates the model a lot. And because of time constrains this model was not further examined. Average error was decreased by only a few seconds.

The second model turned out to be an unsuccessful fusion. For the second model only data from four detectors was used. This resulted in very varying output from the Extrapolation model. The varying output was not suitable for traffic condition identification, which is the reason why this model didn't improve travel time estimates.

The third model is a further developed version of the second model. By introducing a moving average, the output of the Extrapolation model was stabilized and became suitable for traffic condition identification. It turned out that applying the moving average improved travel time estimates already. Because of time constrains, only one condition and correction rule was completed for this model. Estimates for this model are expected to become more accurate when the condition and correction rules are further developed. For now average error is decreased by about 10 seconds.

For further research the condition and correction rules need to be developed further, for example taking more intervals into account. This research has only demonstrated a fusion method that can be successful. Travel time estimations by instantaneous models depending on loop data clearly have systematic errors, correcting these errors without statistical methods is possible.

Table of Contents

1 Introduction

With the emergence of Advanced Traveler Information Systems (ATIS), it is possible to provide various kinds of information to road users. Travel time is one of the most understood measures for road users. By providing reliable travel time estimates it is possible to influence road users' route choice and travel behavior, hence improving the performance of traffic networks.

As pointed out by Van Hinsbergen & Van Lint (2008), a vast amount of models are available for short term travel time prediction. Selecting the most reliable and accurate model for one particular scientific or commercial application is impossible. All models have their own characteristics and perform better in certain situations.

Because of the widespread deployment of loop detectors, most travel time estimation algorithms only have detector data as input. Although detectors continuously collect data, they do not provide an accurate image of the traffic conditions on the road. This is because detectors only collect data at point-locations and not over the entire road.

ETC (Electronic Toll Collection) data on the other hand gives measured travel times over the entire road. But this data arrives too late. By the time travel time is measured, traffic conditions have most likely changed already. By comparing ETC measured travel time with the estimates of the loop detectors, it is possible to evaluate and correct travel time estimations made with loop data in real-time.

The goal of this research is to develop a data fusion between loop detector data and ETC (Electronic Toll Collection) data to make more accurate real-time travel time estimates on expressways. Unlike previous attempt of data fusion, this research will not use historical and statistical analyses for data fusion. By relying on statistical methods, the models fail to take traffic engineering principles into account. And by using historical data the developed models will be too specific for certain situations, which means they can't just be applied anywhere where wanted.

This report is divided into four parts: Introduction, Design approach, Fusion concept and Conclusions. The first part introduces the subject, explains the goal of this research, and briefly describes some existing models for real-time travel time estimations. The second part will describe the research methodology, study area and available data. In the third part some fusion techniques examined during this research will be demonstrated. Finally conclusions and further research recommends will be made.

1.1 Assignment

As part of my study program Civil Engineering and Management at the University of Twente, I conducted this internship research at Kuwahara Lab. (University of Tokyo). This research was carried out in a period of about 12 weeks, starting in early March till the end of May.

At the beginning the goal of this research was to improve travel time estimates by fusing probe vehicle data with loop detector data. ETC data would be used to obtain actually travel time for evaluation purposes. Upon arrival it turned out that there was no probe data available yet, so the topic changed into data fusion of loop data and ETC data. For the change from probe data to ETC data no major changes were required in the research since both types of data are similar.

The advantage of using ETC data instead of probe data is that with ETC data more types of vehicles and driving behaviors are captured. The probe data that would have been used for this research is coming from taxi's, this is a rather selected population of all vehicles on the road. It's discussible if this data is representative for all vehicles on the road and suitable for data fusion.

Another advantage of using ETC data is that this kind of data is more common. Around the world investments are being done in ETC and license plate recognition systems, either to collect toll fees or road taxes. So regardless of the situation this kind of data will be available. Probe data on the other hand needs extra investments and the amount of data will be very limited compared to ETC data.

1.1.1 Research Goal

The goal of this research if to develop an instantaneous travel time estimation model for expressways using loop data and ETC data. Instantaneous means that vehicles entering the expressway will be provided with an expected travel time.

While loop detectors (explained later in 1.2.1 - Loop detectors and ETC) continuously provide data, they do not give an accurate image of the traffic conditions. Although at any moment (upon present) traffic speed and flow is known, this concerns point-locations only. ETC data on the other hand is very accurate, but will only be available after the actual travel time is realized. This is too late, since the aim is to provide expected travel times at the beginning of a travel. In this research an attempt will be made to combine these two types of data, making use of the continuously available loop data and the accuracy of ETC data.

1.1.2 Research objective

The study area for this research is the metropolitan expressway (MEX), route #4, leading from Takaido towards the Tokyo ring (Miyake-zaka Junction). A more detailed description of the study area will come later (2.1 - Study area). Since the detector placement in this study area is very dense, about every 100 meters, for the data fusion not all detector data will be used. Only data from 4 sections will be used, this will make the situation resemble more like the European and American expressways.

The objective of this research is to:

"Maintain the same travel time estimate accuracy while using fewer detectors by integrating ETC data."

Because of the very dense detector placement travel time estimates are very accurate already and improvements are difficult to realize. That is way an attempt will be made to make estimates with the same accuracy with fewer detectors. ETC is basically the income for the MEX, so this data will always be available. By using this data for data fusion with fewer detectors, costs will be saved. Fewer detectors will mean less maintenance, running costs, and data storage.

1.1.3 Research tasks

There are a number of important tasks that needed to be done for this research, which are listed below:

- Sort data and import data into MatLab.
- Obtain travel times from ETC data.
- Determine current travel time estimate accuracy for existing models.
- Determine accuracy when the situation resembles the European and American situation.
- Exploring possible fusion methodologies.

For this research a program called MatLab was used to perform all calculations. All available data first needed to be sorted and imported. The loop data was stored in csv-files, which needed to be transferred over to MatLab-files. Also columns and rows needed to be re-done. ETC data had to be collected from a SQL-database and also stored into MatLab-files.

Once all the data was imported, a script was written to calculate average travel times from the ETC data. These average travel times were considered as the actual travel times and used to evaluate estimates.

Than some existing travel time estimation models were written in MatLab and evaluated. The best instantaneous model was selected as reference for the new model. And because there is no limit to improvement, a historical estimation model was used as limit till how accurate estimations can go. This was done once with all detector data and once with a situation resembling the European and American situation.

After defining the minimum and maximum accuracies for the new model, the development began. This was basically a trail-and-error process, which will be described in 3 - Fusion concept.

1.2 Literature review

For this research a small selection of existing travel time estimation models has been made, which will be described briefly below. The first three models are instantaneous travel time estimation models using loop detector data only. The thirds model is for historical analyses, for this model it is required that all loop data is available (not suited for real-time estimations). This model has been selected to see how accurate travel time can be estimated based on loop data only. The last two models are samples of existing models that fuse different types of models and data for travel time estimations.

Although the last two models are very interesting for this research, these models will only be described briefly. Because these two models use statistical and historical analyses, there is no need to go into details. As noted before, the aim of this research is to develop a data fusion without the use of statistical and historical analyses.

1.2.1 Loop detectors and ETC

Loop detectors are sensors that continuously measure traffic speed and flow and are most common for monitoring expressways. There are several different types, but the most common one is the inductive loop detector (shown in Figure 1).

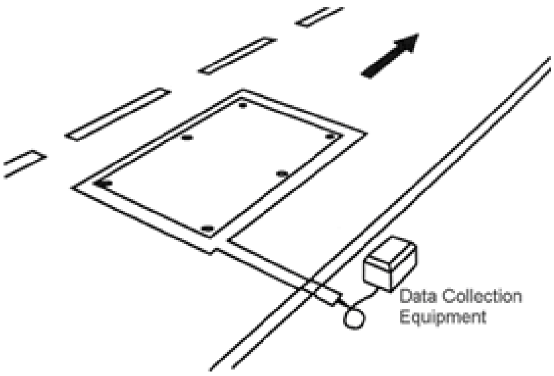


Figure 1 – Inductive loop detector

In the road there is an inductive loop, once a vehicle passes over the loop there will be a flux change in the magnetic field. Based on the flux change a sensor senses whether there is a vehicle above it or not. With this single loop it is only possible to determine the number of vehicle passing (flow) and the fraction of time a vehicle is above the sensor (occupancy). With the following equation it is possible to estimate the speed of vehicles passing over the sensor:

$$V = \frac{L \times q}{occ}$$

V is the estimated speed, L the assumed average vehicle length, q the number of passing vehicles, and occ the occupancy. By using the flow value and occupancy value of one time-interval, the estimated speed can be calculated for that same interval. In this situation the speed is estimated because of the value L, which is estimated. (Jain, M. and Coifman, B., 2005)

In case of the use of dual loop detectors, two inductive loop detectors closely behind each other, speeds can be determined by dividing the distance between the two detectors by the time difference the detectors sensed a vehicle. So by using two detectors near each other the actual speed of vehicles can be measured and no average vehicle length needs to be assumed. Disadvantage is that dual loop detectors are more expensive to deploy and maintain.

There are several variations on the loop detector. Instead of inductive loops other detection methods are available, for example ultrasonic sensors. But the methodology remains the same. The reason why there are variations on the loop detector is because inductive loops have difficulty detecting slow moving vehicles. Flux changes only happen in short fractions of time, by using alternative sensors slow moving vehicles can be detected without problems.

ETC stands for Electronic Toll Collection. Vehicles are equipped with a small transmitter than can communicate wirelessly with a toll gate. When a vehicle passes a toll gate to enter the expressway, time and location for that vehicle are registered. The same happens when the vehicle leaves an expressway. It is only possible to enter and leave an expressway through a toll gate. By filtering the data on entering and exiting location, time, and vehicle-id, travel times can be obtained from individual vehicles for specific origin-destination pairs.

1.2.2 Extrapolation speed based model and Midpoint model

In the Extrapolation speed based model an entire road is divided into segments. At both ends of a segment there is a detector present (see Figure 2). The average speed for each segment is calculated using the following equation:

$$V_{average} = 2/([1/V_1] + [1/V_2])$$

V_1 is the measured speed at the beginning of the segment and V_2 is the measured speed at the ending of the segment. The travel time is calculated by dividing the segment's length by the average speed. (Ying Liu et al, 2006) By summing the separate segments a travel time can be estimated for the entire road. Because collecting data in real-time can result in very varying travel time estimations, results can be stabilized by using the average speeds of the past three minutes.

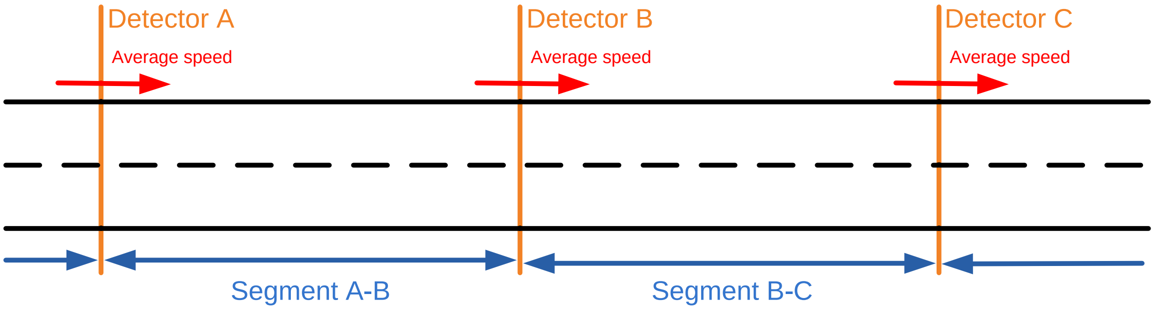


Figure 2 – Extrapolation Speed Based

Although there are many equations available for calculating average speeds on a segment, only the above mentioned equation will be used. According to research by Ying Liu et al (2006) this equation is suitable for varying traffic conditions and distances between detectors. This equation gave relatively reliable estimates for varying scenarios that were considered in their research.

Drawback for the Extrapolation model is that one detector is being used for two estimates. When there is a small distortion near one detector, this will affect the estimates of two road segments while this doesn't necessary need to happen. Another segments placement possibility is shown in Figure 3, this is the Midpoint model. Here each segment has only one detector in the middle and measured values by a detector are assumed to be the same over the whole segment. (Sirisha M. et al, 2006)

For this research both segment placements have been examined and it turned out that the Extrapolation speed based model gave more accurate results compared to the Midpoint model. Throughout this research the Midpoint model will not be mentioned anymore. Not only the Extrapolation model is more accurate, also both models are very similar. A simple comparison between these two models can be found in Appendix A. Based on the results of this comparison it is clear that it is not needed to include the Midpoint model in this research.

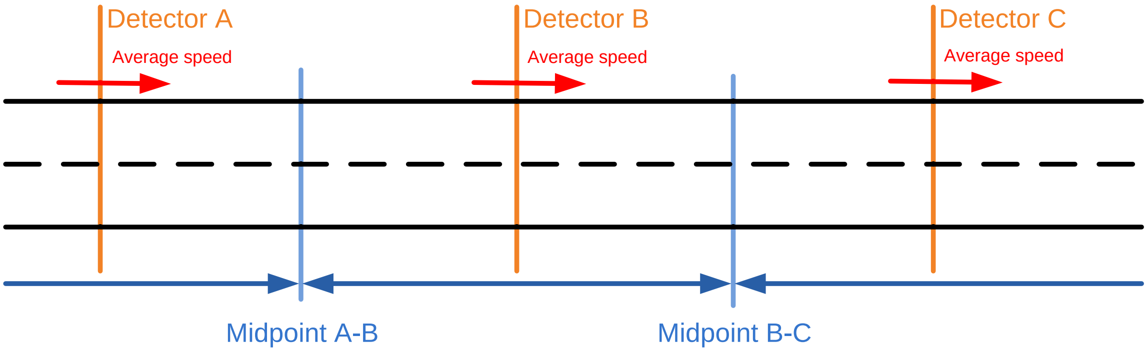


Figure 3 – Midpoint Influence Area

1.2.3 Nam and Drew dynamics model

This model has the same segment placement as the model discussed above (see Figure 4). But the detectors for this algorithm measure traffic flow instead of traffic speed. Based on the measured traffic flows the density on each segment is calculated with the following equitation:

$$k_{(t)} = \frac{Q_{in,(t)} - Q_{out,(t)}}{\Delta x}$$

In this equation $k_{(t)}$ is the density on a segment during interval t, $Q_{in,(t)}$ is the cumulative number of vehicles entering the segment during interval t, $Q_{out,(t)}$ the cumulative number of vehicles leaving the segment during the same interval, and Δx the length of the segment. With the measured traffic flow and estimated density the travel time for each segment is calculated with the following equation:

$$tt_{(t)} = \frac{\Delta x}{2} \cdot \frac{k_{(t)} + k_{(t-1)}}{q_{out,(t)}}$$

Here $tt_{(t)}$ is the estimated travel time for a segment at interval t, calculated with the density of the same interval ($k_{(t)}$), the density from the interval prior to the current interval ($k_{(t-1)}$), and the segment length (Δx). (Nam and Drew, 1998)

Although the original Nam and Drew model suggested two different formulas (one for free flow conditions and one for congested conditions), in this research only on formula will be used in this algorithm. According to research by Lelitha D. et al (2009) the use of two different formulas was unnecessary. The research also revealed that the consistent use of one formula (for congested conditions) resulted in a better estimated travel time for varying traffic flow conditions.

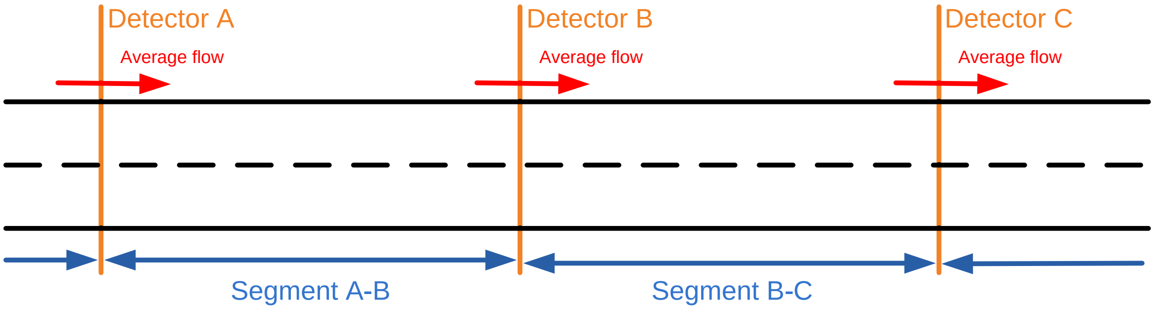


Figure 4 – Dynamics Model by Nam and Drew

Unlikely the first model, the second one considered in this research is not based on average speeds. Instead it uses traffic flows and densities to estimate travel times. The advantage of this feature is that the model is unaffected by the fact that "time mean speed" and "space mean speed" are not the same. Speed based algorithms depend on measured speeds from detectors, which is time mean speed. Average speeds measured at one location over a period of time. For speed based models it is better to use space mean speed, which is traveled distance divided by travel time from all vehicles over a road segment. These parameters may be the same when traffic conditions are homogenous, but when traffic conditions approach congestion time mean speed exceeds space mean speed. This will eventually result in underestimated travel times. (Van Lint, J.W.C., and Van der Zijpp, N.J., 2003)

By using densities on segments to estimate travel time, it is expected that this model will give accurate estimates under varying traffic conditions. One important note is that the density is determined by cumulative values, disadvantage of this is that measure errors are not averaged out. Research by Oh, J. et al. (2003) pointed out that detectors tend to undercount the real amount of vehicles passing by. This means that over time the error in the determined density by detectors increases. As alternative Oh, J. et al. suggest the uses of the following equation to determine densities:

$$k = \frac{o \cdot L}{g}$$

The density (k) is calculated with multiplying the segment's length (L) with the average occupancy of the upstream and downstream detector (o). This divided by the average vehicle length (g) will give the average density on the segment. Occupancy is the fraction of time a detector senses vehicles above it.

1.2.4 Time slice model

This third model is more suited for historical travel time analyses. In case it is applied for real-time applications, a delay has to be taken into account. Unlike the previously discussed models, the time slice model doesn't use all segment data from the same time-interval to estimate travel times. Instead, it determines when a vehicle enters each segment and uses the most up-to-date data available. By using the data corresponding to the same time-interval a vehicle is traversing a segment, this model takes speed variations over time into account. (Ruimin Li et al, 2006)

The equations used for calculating travel times for each segment can be the same as the previously mentioned models. The only difference is that this model uses data from different time-intervals to estimate travel times. For this research the equation from the Extrapolation speed based model is used for the time slice model. For example a vehicle enters segment 1 at time is t, the average speed on segment 1 is calculated as follow:

$$V_{average\ (1,t)} = 2/([1/V_{1(t)}] + [1/V_{2(t)}])$$

Again V_1 is the measured speed at the beginning of the segment and V_2 is the measured speed at the ending of the segment. The (t) determines data from which time-interval will be used.

With the average speed the travel time for the first part of a vehicle's trajectory (segment 1) can be calculated, t(k). This travel time will be used to determine the average speed on segment 2:

$$V_{average\ (2,t)} = 2/([1/V_{1(t+t(k))}] + [1/V_{2(t+t(k))}])$$

This will continue on till the destination is reached. In real-time situations the travel time for a vehicle entering at time is t cannot be given since the data at moment t+t(k) is not available yet. In real time applications the delay of this model is equal to the travel time.

Basically ETC data gives measured travel time with the same delay (after that the actual travel time is realized), the only difference is that ETC data is unrelated to loop detector data. The travel times provided by the Time slice model are obtained using loop data. As mentioned before the Time slice model will be used to evaluate how accurate travel time estimates can be based on loop data only.

1.2.5 Bayesian combination model

Although this in not a data fusion model, it does have some interesting aspects for this research. The framework of this model is given in Figure 5. (Van Hinsbergen & Van Lint, 2008)

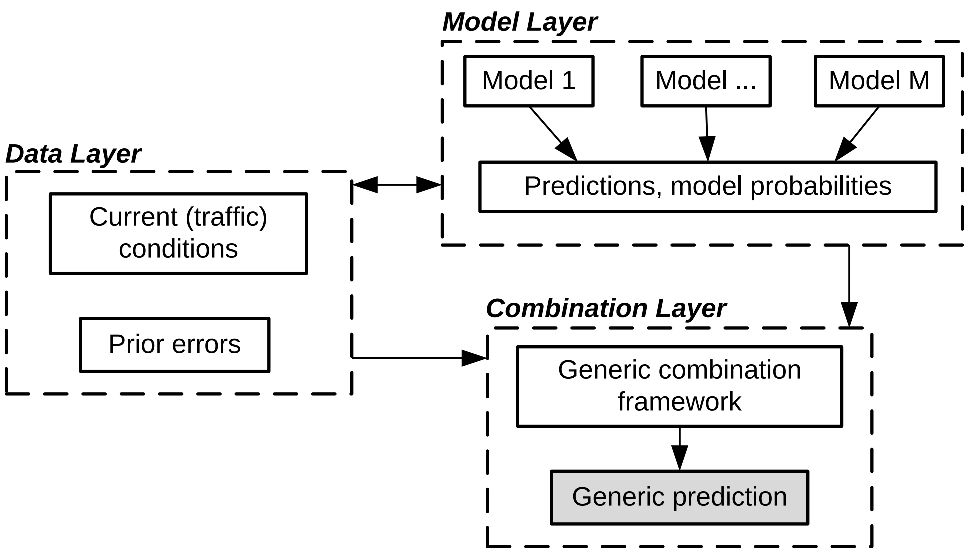


Figure 5 – Framework Bayesian combination model

This model is divided into three layers. The first layer consists of existing travel time estimation models. In the Bayesian model several estimation models run simultaneously. In the literature there are numerous models available, each of these models have their own characteristics and perform better in certain situation. Since it's impossible to select the most reliable and accurate model, the Bayesian model runs several models parallel and averages between these models based on probabilities.

The second layer (data layer) is the input for the Bayesian model. Probabilities for each model have been determined based on historical research (prior errors). Real-time data is used to determine the current traffic conditions, in this model only loop detector data is used for real-time data collection. But different types of data can be used to determine traffic conditions, although this was not mentioned.

The last layer is the combination layer, here is defined how the outputs of the individual models are handled. In the research by Van Hinsbergen & Van Lint (2008) two fusion strategies were examined.

1. *Winner Takes All*
   In this strategy the models are ranked based on their probabilities. The model with the highest probability will be selected as output for the Bayesian model.
2. *Weighted Linear Combination*
   Here probabilities are used as a weight. All M models' predictions are used, but multiplied by factors that add up to one. A weighted linear combination was examined, probabilities were normalized and used as a weight for the models.

Both combination strategies turned out to be successful, with the second strategy even better than the first one. In the research by Van Hinsbergen & Van Lint (2008) only two models were used for fusion. They recommended to increase the number and diversity of models. This would make the Bayesian model more robust and always provide the most accurate travel time prediction.

Another important note was that the probabilities weren't always right about the individual models. A way to overcome this is to introduce prior knowledge about the models' performances. For example, model 1 always outperforms model 2 in dissolving traffic conditions. By introducing prior knowledge they want to prevent probabilities to worsen accuracy in certain situations.

1.2.6 Dempster-Shafer data fusion model

This last model is an example of data fusion. Again a statistical fusion method and historical data is used. Actually this model is quite similar to the above mentioned model, but here probabilities are assigned to data sources. The research by Nour-Eddin El Faouzi et al (2009) focused on the data fusion of loop detector data and ETC data.

The first step of making their model was to define four travel time hypotheses, these were as follow:

1. $h_1 = \{TT_{(t)} \; such \; that \; TT_{(t)} \leq 1.1 \times TT_{ff}\}$
2. $h_2 = \{TT_{(t)} \; such \; that \; 1.1 < TT_{(t)}/TT_{ff} \leq 1.3\}$
3. $h_2 = \{TT_{(t)} \; such \; that \; 1.3 < TT_{(t)}/TT_{ff} \leq 1.5\}$
4. $h_1 = \{TT_{(t)} \; such \; that \; TT_{(t)} > 1.5 \times TT_{ff}\}$

Then for each data source probabilities were assigned for each travel time hypotheses. This was done by making a confusion matrix for each source. The probability for each data source for each hypothesis is found by normalizing the confusion matrices. Basically the probability values are chances that a source output is correct. For example, when a source output is that the travel time corresponds to hypotheses 1, the probability value is the chance that that is correct.

The research area for their research was a 7 km section of the AREA motorway in the Rhône-Alpes region of France. Data was available from seven days in 2003, which was used to compute the confusion matrices. And data from five days in 2004, this data was used to evaluate their model. The framework of their model is shown in Figure 6.

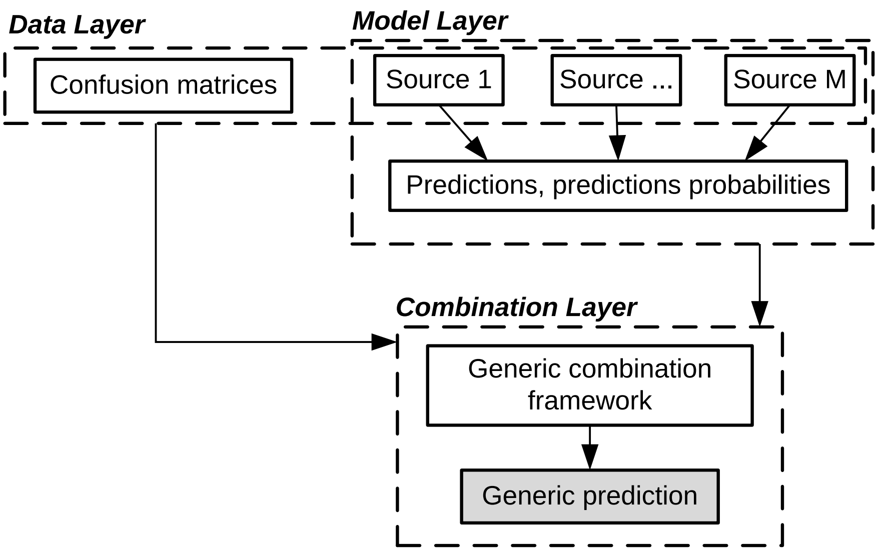


Figure 6 – Framework Dempster-Shafer data fusion model

Instead of running several models parallel, as the Bayesian model, the model layer here consists of several travel time estimates from different data sources. With the confusion matrices from the data layer probabilities are assigned to the estimates (data sources) in the model layer. In the combination layer first the hypothesis with the highest probability is selected, than the estimates that meet the criteria of the hypothesis will be selected. With the probabilities as weight an average can be calculated, which will be the output of the model.

In their research the fusion results were disappointing, there are two reasons for this. First the data source ETC always arrives with a delay, this resulted in the loop data almost always outperforming the ETC source. The second reason is because of a structural change in the motorway ETC deployment policy that resulted in an increase in market penetration of electronic toll tags. But their research did propose a methodology of fusing different types of traffic data. Any data can be used as input for this model as long as probabilities can be assigned to the sources.

2 Design approach

As mentioned before, this research's aim is to develop a travel time estimation model using loop data and ETC data. Because the research area has a very dense detector placement, an attempt will be made to use fewer detectors while maintaining the same accuracy in travel time estimations. This will also make this research more representative for the European and American detector placements.

Existing fusion models are all based on statistical methods. This doesn't always improve the travel time accuracy, since the errors are not randomly distributed but systematically. As mentioned in the paper by Van Hinsbergen & Van Lint (2008), sometimes it's needed to introduce prior knowledge of how estimate errors are because statistical methods fail to improve accuracy in certain situations. In this research no historical and statistical methods are used for the data fusion.

2.1 Study area

The study area for this research is the metropolitan expressway (MEX), route #4, leading from Takaido towards the Tokyo ring (Miyake-zaka Junction). The length of the whole area is about 14 km, with two lanes in each direction. A simplified map of the study is given in Figure 7. The ETC-gates in the area are marked with their number in a circle. Length of each segment is given in meters and detectors are marked with a blue line.

Actually for each segment data such as average speed and flow are measured with about three detectors, usually two at the beginning and one at the end of each segment. But because the data of the individual detectors is not available, it is assumed that the data of each segment is from one hypothetical detector at the middle of each segment (the blue lines).

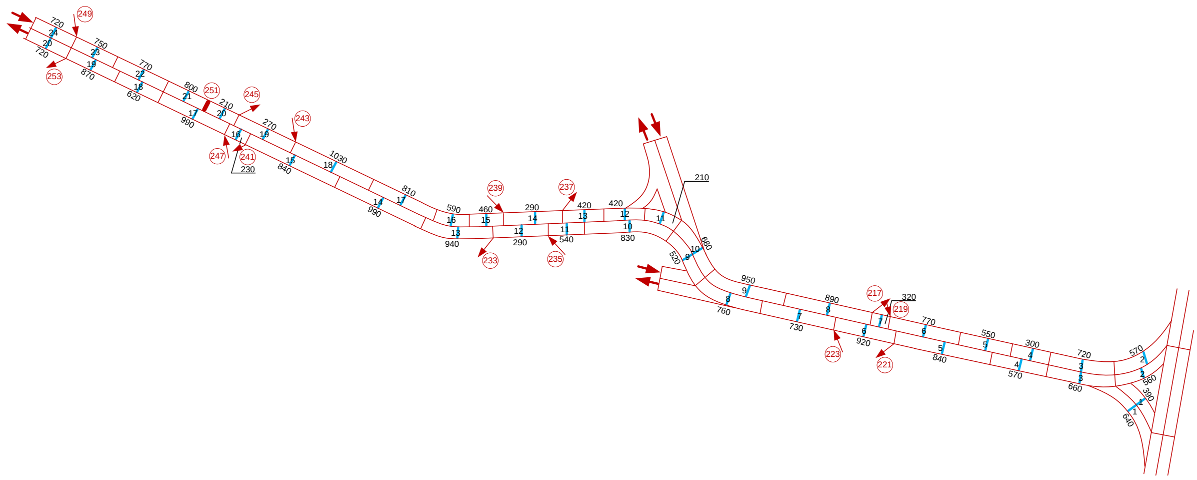


Figure 7 – Simplified map of study area

The Miyake-zaka Junction connects route #4 with the ring-road in Tokyo, during peak hours this ring is heavily congested. The further away from the ring, the less the congestion gets. This makes route #4 an interesting study area. Since it goes towards a congested area, there will be various traffic conditions on the route.

To resemble the European and American detector placement about 70% of all the detectors have been dropped out from the research area (see Figure 8). Because of time constrains only the direction towards Tokyo was examined. More precisely traffic with the origin ETC-gate 251 towards either the destination ETC-gate 237 or 217. Since data from ETC-gate 249 is not available, the longest route that can be examined is from gate 251 to 217. Maximum allowed speed on the MEX is 80 km/h. A complete map of the study area can be found in Appendix B.

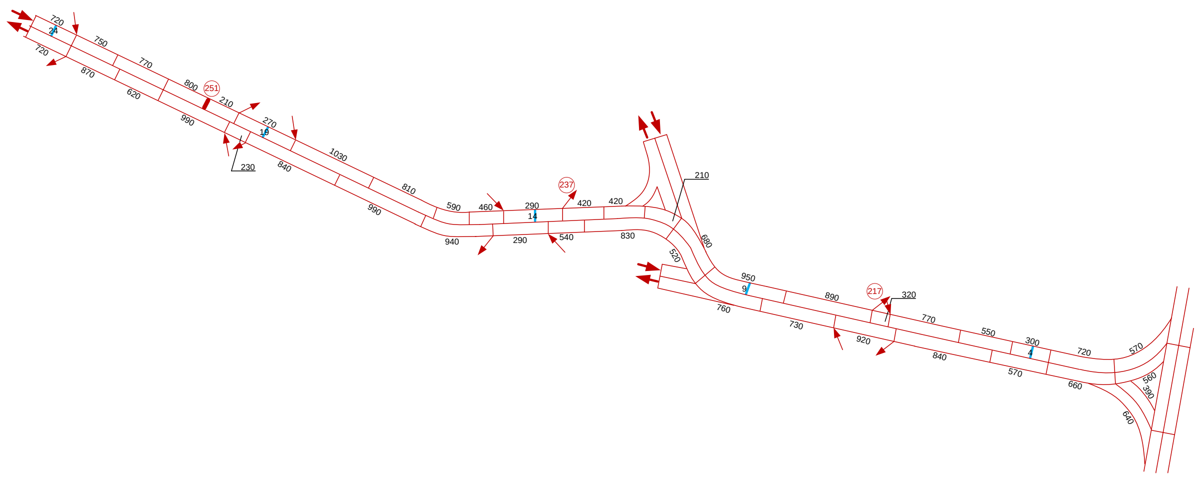


Figure 8 – Data locations used for research

2.2 Research Data

For this research two kinds of data were available, loop detector data and ETC data. The loop data was collected by ultra-sonic sensors and not by inductive loop detectors as what is common in Europe and America. Advantage of the ultra-sonic sensors is that they don't have any complications with detecting slow moving vehicles, unlikely inductive loop detectors. Furthermore all detectors are dual loops, which mean the measured speeds can be assumed to be the actual speeds. There was no need to estimate the average vehicle length for determining the speeds.

For each segment aggregated loop data was available with a five minutes update interval. Although data from each segment came from several detectors, during this research it is assumed that the data came from one detector at the middle of each segment. No issues are expected from this assumption, since the length of each segment is relatively short and traffic conditions can be assumed to be the same over the whole segment.

The second data source for this research is ETC data. Both the ETC data and loop data are from the period July 1, 2006 till July 7, 2006. During this period the ETC market penetration was about 60%. From the ETC data it is possible to determine when and where each vehicle entered and left the research area. Based on the enter time and exit time the travel time of each vehicle can be obtained.

No errors are expected in the ETC data, although some vehicles showed an exceptional long travel time. Based on the average travel time each five minutes, these exceptional vehicles were filtered out. First the ETC data was divided into five minutes intervals and for each interval average travel times are calculated. After the first calculation vehicles with an exceptional long travel time, more than 50% off from the calculated average were discarded. Again the average travel time is calculated, but without the discarded vehicles.

The travel times obtained from the ETC data are assumed to be the actual travel times. This will be the data to compare all estimates against. Although the ETC data will also be used for making travel time estimates, the data will still be a valid source for comparison. This is because there is a little delay between the data used for estimations and for comparison, illustrated in Figure 9.

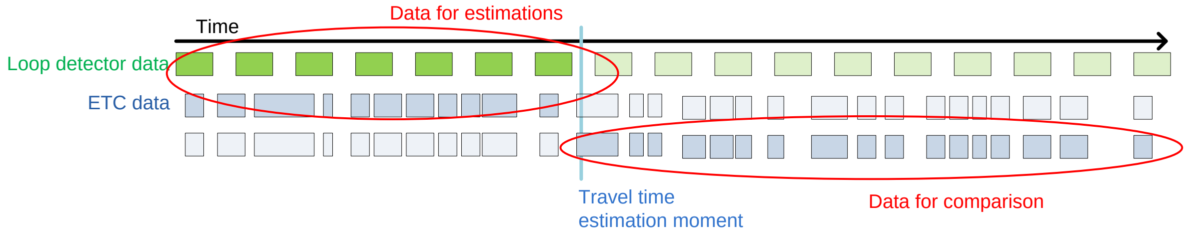


Figure 9 – Delay between estimation data and comparison data

Since in real-time travel time estimation vehicles get the travel time at the beginning of their journey, the actual travel time is yet to be realized. This means the ETC data used for comparison is not available for fusion, in Figure 9 on the left side of the “travel time estimation point”. On the right side is all the data that is available for comparison. So for each moment in time the data used for estimation is unrelated to the actual travel time (comparison) data.

During this research calculations were done in MatLab and results are stored in Excel. The advantage of using MatLab is that you have access to the workspace. All variables and temporarily results can be accessed in the workplace and be saved for later use. This means the calculations performed by MatLab are very transparent. And being able to save the temporarily results makes it possible to write small scripts. Instead of writing a whole script for travel time calculations and accuracy analyses, calculations steps and data sorting steps can be written in different scripts. This helps to keep the scripts simple and easy to work with. After running each script the results can be saved before running the next script. This means each time one script is being edited, only re-running one set of calculations in the corresponding script is needed. All calculations in previous scripts are stored in the temporarily results and can be loaded back into the workplace. Loading back old data is much faster than re-running all calculations, which can be very time consuming.

By storing the results in another program besides MatLab, it helps keeping everything organized. Since both programs work with different file-types and formats, storing the results happens manually. This does slow down the work, but paying extra attention to the handling and storing the results you will be able to keep track of everything. Also Excel has a more user-friendly interface to work with while making graphs.

2.3 Travel time estimates

First task in the research was to evaluate the current situation. All detector data (shown in Figure 7) are used as input for the Extrapolation speed based model and the Nam and Drew dynamics model. To evaluate how accurate estimates could get based on loop data only, a time slice model was examined. As it turned out that the accuracy of the current situation was very difficult to improve, a situation more like in Europe and America was made for further research. For the situation with fewer detectors the accuracy of ravel time estimates needed to be evaluated as well. And as part of the fusion method, which will be explained later (3 - Fusion concept), the accuracy of the estimates over time had to be examined.

2.3.1 Travel time estimates for the current situation

The travel time estimates according to the Nam and Drew model and the Extrapolation model are shown in Figure 10, respectively a blue line and a pink line. The yellow dotted line represents the actual travel times obtained from the ETC data.

It turned out that the Extrapolation model gives more accurate results than the Nam and Drew model. This is in conflict with the results of Lelitha D. et al (2009), according to their research the Nam and Drew model should perform better. There are two aspects that probably contributed to these conflicting results. First of all, in this research dual loop detectors are used for the collection of loop data. By introducing an estimated vehicle length, in this research an unnecessary variable with error has been introduced for travel time estimations. Another aspect that could have let to conflicting results is the very dense detector placement. This results in very accurate travel time estimation for the Extrapolation model, which is very difficult to improve.

Furthermore out of these analyses it turned out that the Nam and Drew model behaves differently from the Extrapolation model. Based on this finding the first fusion strategy was developed, which will be described later in “2.3.1 - Travel time estimates for the current situation”.

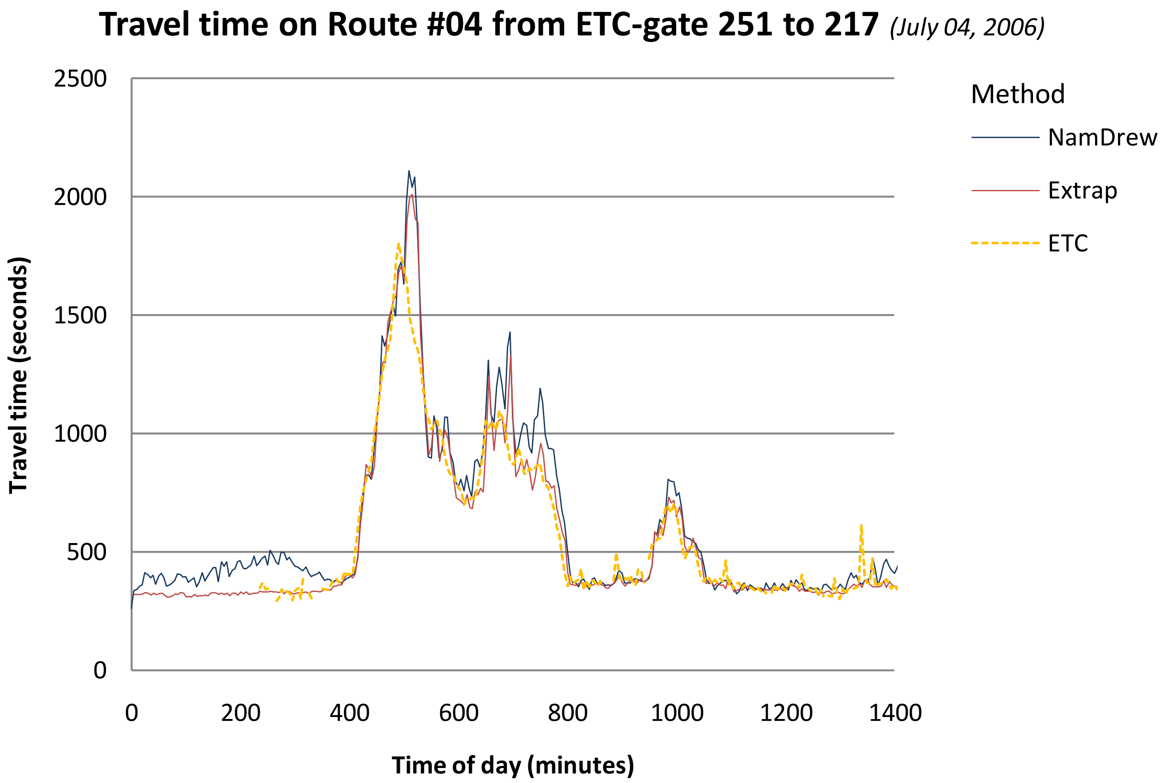


Figure 10 – Travel time estimates for the current situation

For the comparison of the accuracies of the different models, average overestimation and average underestimation were determined for each model. For each interval it was determined whether the model overestimated or underestimated the travel time according to the ETC data. This way it was possible to keep overestimations and underestimations separate. By summing all overestimations and dividing it by the number of times travel time was overestimated, an average overestimation was calculated. The same procedure goes for the underestimation. Results of the comparison between the Nam and Drew model and the Extrapolation model are shown in Table 4.

Nam and Drew dynamics model (with all loop detectors)			**Extrapolation speed based model** (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
-	-	-	-	-	-
85.30071	65.3429	75.32179	86.32575	47.437	66.8814
70.49105	47.9351	59.21305	50.41537	37.1489	43.78214
115.3225	57.393	86.35779	81.27327	50.0184	65.64581
138.3732	95.589	116.9811	109.3794	74.1254	91.75241
67.09878	51.0848	59.09177	54.148	35.9148	45.03138
124.8942	83.3378	104.116	93.97934	70.6825	82.33094
	∑ =	**501.0815**		∑ =	**395.4241**

Table 4 – Estimate error comparison between Nam and Drew model and Extrapolation model (from ETC-gate 251 to 217)

Average absolute error is the average of the average overestimation and average underestimation. The first row in Table 4 is empty because this is an exceptional day (July 1, 2006). In Appendix C the travel times for this day are shown and will make clear why this day is not included in the above table.

Based on the results in the above table it is clear that the Extrapolation speed based model performs better than the Nam and Drew dynamics model. On all six days the average absolute error of the Extrapolation model is lower than that of the Nam and Drew model.

2.3.2 Travel time estimates according to the Time slice model

The second travel time estimate analysis was done for the Time slice model. Travel times from this (historical based) model were the most accurate of all models considered in this research. The graph of this model is shown in Figure 11.

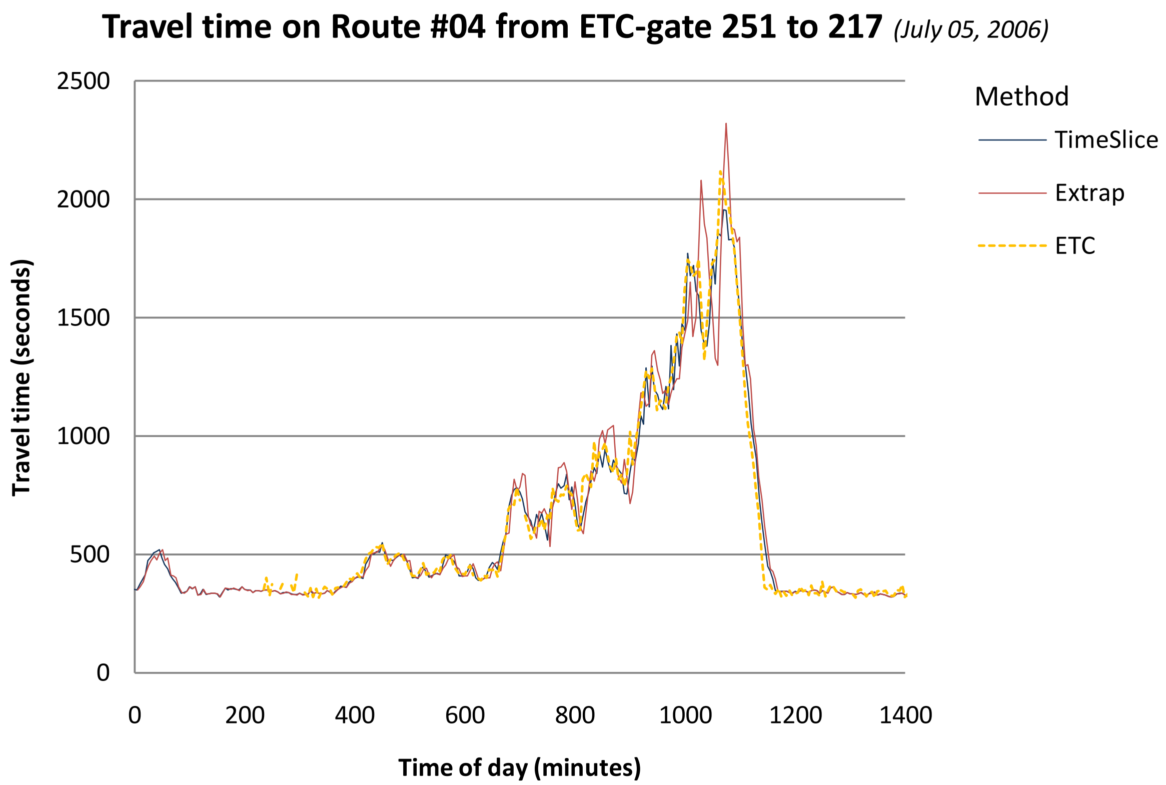


Figure 11 – Travel time estimates by the Time slice model and the Extrapolation speed based model

In the above figure the blue line represents the travel times according to the Time slice model. The pink line represents the Extrapolation model. It is clear that the Time slice model more accurately follow the actual travel times obtained from the ETC data. In Table 5 the results for the seven examined days are shown.

Time slice model (with all loop detectors)			**Extrapolation speed based model** (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
33.53924	39.7033	36.62129	129.1858	72.7797	100.9828
17.93999	30.5059	24.22292	86.32575	47.437	66.8814
27.17834	23.438	25.30816	50.41537	37.1489	43.78214
23.24303	34.0556	28.6493	81.27327	50.0184	65.64581
34.15193	43.3475	38.74971	109.3794	74.1254	91.75241
25.58925	25.0761	25.33269	54.148	35.9148	45.03138
26.54383	36.5633	31.55354	93.97934	70.6825	82.33094
	$\sum =$	**210.4376**		$\sum =$	**496.4069**

Table 5 – Estimate error comparison between Time slice model and Extrapolation model (from ETC-gate 251 to 217)

To more clearly demonstrate the accuracy of the Time slice model, also the errors of the Time slice model and the Extrapolation model are plotted in one graph, see Figure 12. In this figure for each time-interval the error of each model's estimate is plotted, the blue line corresponding to the time slice error and the pink line to the Extrapolation error. Clearly the time slice outperforms the Extrapolation model.

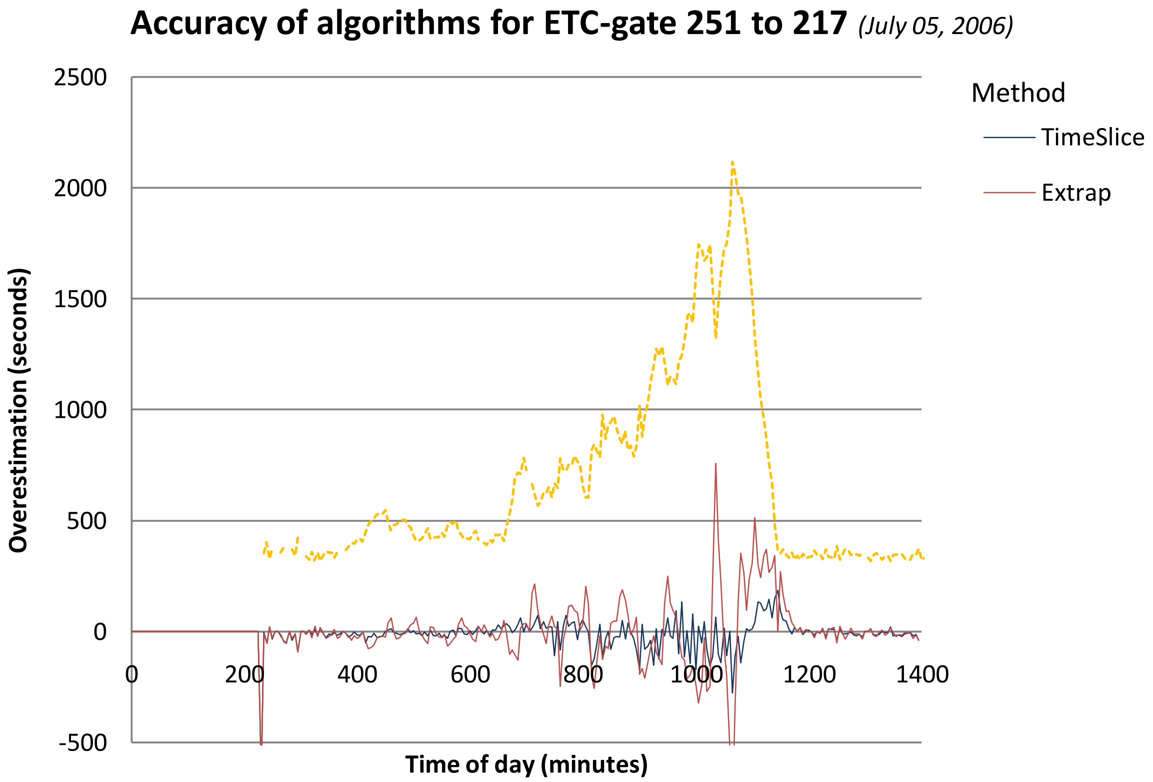


Figure 12 – Estimate errors by the Time slice model and the Extrapolation model

Since the improvement of travel time estimates can go on forever, the accuracy of the Time Slice model will be seen as the limit of how accurate travel time estimations can be. During this research the accuracy of the Extrapolation model with the very dense detector placement will be used as reference. This is the accuracy that needs to be maintained when fewer detectors are used. In case improvements can go beyond this accuracy, the accuracy of the Time slice model is what will be tried to be achieved.

2.3.3 Travel time estimates for the situation with fewer detectors

Now that the current situation is clear, analyses of the situation with fewer detectors can begin. As stated before, about 70% of all detectors will be dropped to resemble the European and American situation (Figure 8). Travel time estimates will be analyzed for the Extrapolation model and the Time slice model. The results are shown in Table 6 and Table 7.

Extrapolation speed based model (with 30% of all loop detectors)			**Extrapolation speed based model** (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
104.9583	162.861	133.9099	129.1858	72.7797	100.9828
56.46005	70.3793	63.41968	86.32575	47.437	66.8814
73.84489	25.3089	49.57688	50.41537	37.1489	43.78214
87.79284	95.1859	91.48939	81.27327	50.0184	65.64581
110.9346	116.031	113.4826	109.3794	74.1254	91.75241
93.85418	52.5319	73.19304	54.148	35.9148	45.03138
117.525	117.347	117.4359	93.97934	70.6825	82.33094
	∑ =	**642.5074**		∑ =	**496.4069**

Table 6 – Travel time estimate accuracy for the Extrapolation model with fewer detectors (from ETC-gate 251 to 217)

Time slice model (with 30% of all loop detectors)			Time slice model (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
45.34729	114.068	79.7075	33.53924	39.7033	36.62129
46.81081	63.4606	55.1357	17.93999	30.5059	24.22292
64.57156	21.8015	43.18654	27.17834	23.438	25.30816
48.63072	63.7232	56.17694	23.24303	34.0556	28.6493
75.51659	107.841	91.67896	34.15193	43.3475	38.74971
75.63694	40.695	58.16597	25.58925	25.0761	25.33269
81.09806	95.823	88.46053	26.54383	36.5633	31.55354
	∑ =	**472.5121**		∑ =	**210.4376**

Table 7 – Travel time estimate accuracy for the Time slice model with fewer detectors (from ETC-gate 251 to 217)

As expected, both models perform less accurate with fewer detectors. Interesting to see is that the accuracy of the Time slice model with fewer detectors is better than the Extrapolation model with the very dense detector placement.

2.3.4 Travel time estimate accuracy over time

All that's left before data fusion can start is to analyze the travel time estimate error over time. For this similar graphs as Figure 12 are used. For each time-interval the estimates are compared to the actual travel time. The error can then be plotted into a graph, by keeping the actual travel time (the yellow dotted line) in the same graph it will be clear at what traffic conditions the models will fail and how the errors are. In Figure 13 it is clear that when travel time increases the models underestimate travel times, while at decreasing travel times the models overestimate. This error pattern was present in all examined days, see Appendix D for the rest of the graphs.

The figure clearly shows that there is a correlation between estimate error and traffic condition. Which means that estimate errors can be corrected with certain correction-rules for certain traffic conditions. A statistical correction method should be avoided, since these errors are not randomly distributed.

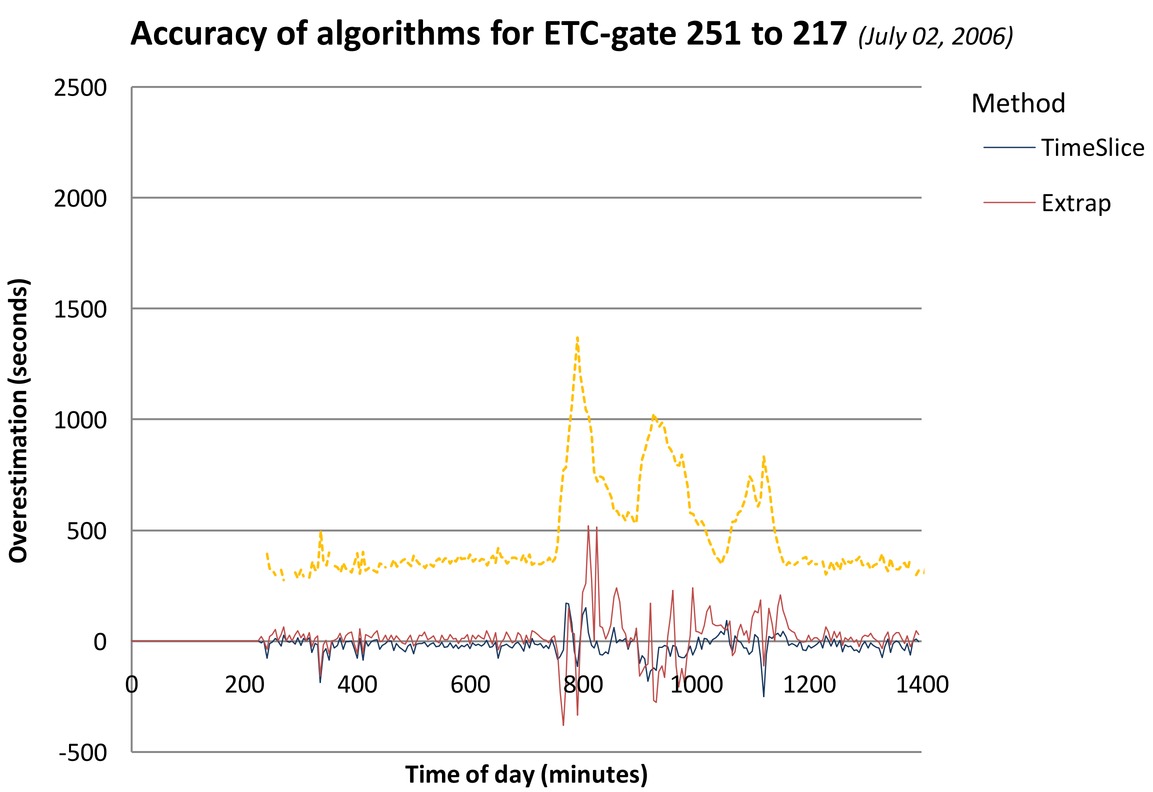


Figure 13 – Travel time estimate errors over time for situation with fewer detectors

2.4 Discussion existing models

Before starting on the development of the new travel time estimation model the current situation was evaluated. Two instantaneous travel time estimation models, Nam and Drew model and the Extrapolation model, were examined. Out of these two models, the Extrapolation model gave the most accurate estimates. So this model is selected as reference for the evaluation of the new model. The fusion model is considered successful if the travel time estimates are more accurate than the Extrapolation model.

Because the study area has a very dense detector placement, the situation was also examined with only the use of 30% of all detectors. By using only 30% of all detectors the situation resembled the European and American detector placements. With the very dense detector placement not only would this research be very specific for only this study area, also improving travel time estimates will be difficult. As shown in the results above, travel time estimates with all detectors are quite accurate already.

Besides the instantaneous models, one historical based model was examined (the Time slice model). This was mainly done to see how accurate travel time estimates with loop detectors only could be. Also this model was examined with all detectors and with the use of only 30% of all detectors. As expected, this model gave the most accurate results out of all considered models.

Since the improvement of travel time estimates can go on forever, the accuracy of the Time Slice model will be seen as the limit of how accurate travel time estimations can be. If the fusion model's accuracy is better than the Time slice model, the fusion model will be considered as finished. Making instantaneous travel time estimation models more accurate than historical models is unrealistic.

So with the Extrapolation model as minimum accuracy and Time slice model as maximum accuracy, the boundaries for the fusion model are set. The next step is to examine the systematical errors by the instantaneous models, which are clearly demonstrated in Figure 13 and Appendix D.

When travel time increases travel times are underestimated, and with decreasing travel times the models overestimate travel times. Another interesting behavior is that during free-flow conditions, the actual travel time is usually between the estimates of the Nam and Drew model and the Extrapolation model. At congested periods this behavior is not visible.

So besides increasing and decreasing travel time situations, also free-flow traffic and congested traffic are situations that need to be treated differently.

3 Fusion concept

The fusion model developed in this research uses an existing travel time estimation model as basis. By evaluating the previous time-intervals' estimates with the incoming ETC data, the current estimate by the existing model would be corrected to a more accurate estimate.

Throughout this research several fusion concepts were examined. The first one examined was a model running two models parallel, the Extrapolation and the Nam and Drew. By integrating ETC data previous time-intervals were evaluated and based on the previous intervals an estimate for the current interval would be calculated with the estimates of the Extrapolation model and the Nam and Drew model. This concept was tested on the situation with very dense detector placement. It turned out that this concept would become very complicated to be successful and another fusion concept was examined.

The second concept only uses one existing estimate model as basis, the Extrapolation model. The ETC data is used to evaluate the error in previous time intervals. Based on the current travel time estimate trend, either ascending or descending, travel time would be corrected assuming that the previous error is still present in the current interval. This concept was tested on the situation with fewer detectors, by this time it was clear that improving the situation with very dense detector placement is almost impossible. Results for this concept were still varying, so a third concept was developed.

This last concept was to introduce some boundaries to where correction would be applied. It turned out that travel times in free flow conditions should not be improved. These were already so accurate, that any adjustments to them resulted in random improvements. Another thing that was introduced into this concept was the moving average, this was needed to smoothen the travel time estimate by the Extrapolation model. The estimates were too unstable to be used to identify traffic conditions, by averaging the last two intervals the travel time estimates became more accurate and graph used for identifying traffic conditions became more stable. The correction rules were made more specific and this turned out to be a successful data fusion model.

One issue that occurred during the research is the real-time filtering of ETC data.ETC data is not perfect and always consist of some vehicles that spend an unusual long time on the expressway. Possible reasons are car breakdown or accidents. For the fusion model these entries have to be filtered out. For the models below ETC data was compared to estimates by the Extrapolation model, any entry that exceeds this estimate by 80% or more was discarded.

3.1 Adaptations during the research

In this section the examined fusion concepts will be described and evaluated. Also some important changes that were made in the fusion concept will be explained here as well.

3.1.1 Corrections on two models running parallel based on previous errors

The first fusion concept was based on the results of the analyses of the accuracies of the current situation. It seemed that the Nam and Drew model overall overestimated the travel time, while the Extrapolation model overall underestimated the travel time (see Figure 10).

By averaging between these two models, travel time estimates were expected to improve. The correction rules and traffic condition identifications were as follow:

1. In case the last two evaluated intervals each time one model overestimated the travel time and one model underestimated the travel time. It is assumed that in the current interval the actual travel time is between the two estimates of the two models.

   By determining the difference between the two estimates of the last evaluated interval, let this be Δm. And by determining how much travel time was underestimated by the lowest

travel time estimate, let this be Δu. A ratio (Δu /Δm) can be obtained to use to calculate the estimate for the current interval.
For the current interval the difference between the two models is determined, let this be ΔM. And by taking the lowest travel time estimate and adding ΔM×(Δu /Δm) a travel time output for the fusion model is obtained.

2. In case the last two evaluated intervals each time both models underestimated the travel time. It is assumed that the travel time in the current situation is ascending and thus the travel time is being underestimated.

   In this case the average error of the last two evaluated intervals is added to the current estimate by the Extrapolation model, thus the output of the fusion model.

3. In case the last two evaluated intervals each time both models overestimated the travel time. It is assumed that the travel time in the current situation is descending and thus the travel time is being overestimated.

   In this case the average error of the last two evaluated intervals is deducted from the current estimate by the Extrapolation model, thus the output of the fusion model.

4. For the rest of the situations the travel time estimates by the Extrapolation model are used as output of the fusion model.

With last two evaluated intervals, it means the last two intervals where ETC data is available for. In real-time applications this usually is a few intervals back, depending on the delay with which ETC data arrives.

Illustrations of the correction methods are given in Table 8. Parts of the travel time estimates graphs and ETC graphs are plotted, identification of the situation, error determination, and correction are demonstrated. The yellow dotted line is the ETC data, the green line is the Nam and Drew model, the red line is the Extrapolation model, the blue dot is the corrected estimation.

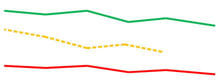
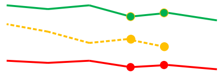
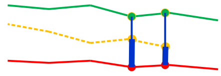
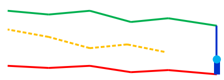
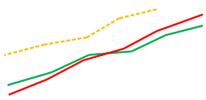
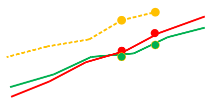
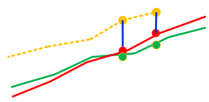
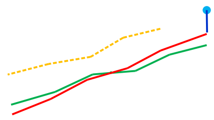
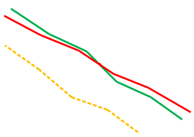
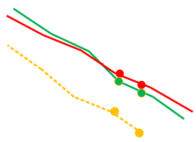
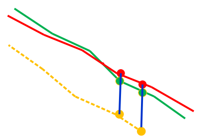
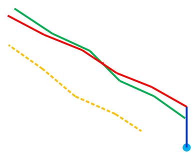
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				
#2				
#3				

Table 8 – Illustrations of corrections for first fusion model

Results of this data fusion concept are minimal, see Table 9. Although travel time estimates improved in most situations (5 out of the 7 cases), it is not much. To see more accurately where travel time estimates improved, errors over time were plotted (see Figure 14 for an example).

First fusion model (with all loop detectors)			**Extrapolation speed based model** (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
123.8824	74.502	99.19218	129.1858	72.7797	100.9828
84.58546	49.0246	66.80503	86.32575	47.437	66.8814
48.9654	37.5799	43.27266	50.41537	37.1489	43.78214
85.60895	48.2331	66.92104	81.27327	50.0184	65.64581
100.9127	76.4633	88.68801	109.3794	74.1254	91.75241
51.71613	35.6665	43.69129	54.148	35.9148	45.03138
93.04104	72.6042	82.8226	93.97934	70.6825	82.33094
	∑ =	**491.3928**		∑ =	**496.4069**

Table 9 – Travel time estimate accuracy for the First fusion model (from ETC-gate 251 to 217)

In the figure below the error of the first fusion model is plotted in blue. For comparison the error of the Extrapolation model is plotted in the same figure (pink). Both graphs seem to be identical, this is because the defined condition identifications are too specific. Only a few times the estimates were corrected, this explains the very little improvements.

In order to achieve better improvements, the identification rules need to be broader. With the current method of identifying the traffic conditions it is very difficult to make the rules broader. Not only two models are running parallel, also the data used for identification has a relatively long delay (equal to the travel time). For the next fusion concept only one model will be used as basis and data with a smaller delay is used for identification of the traffic condition.

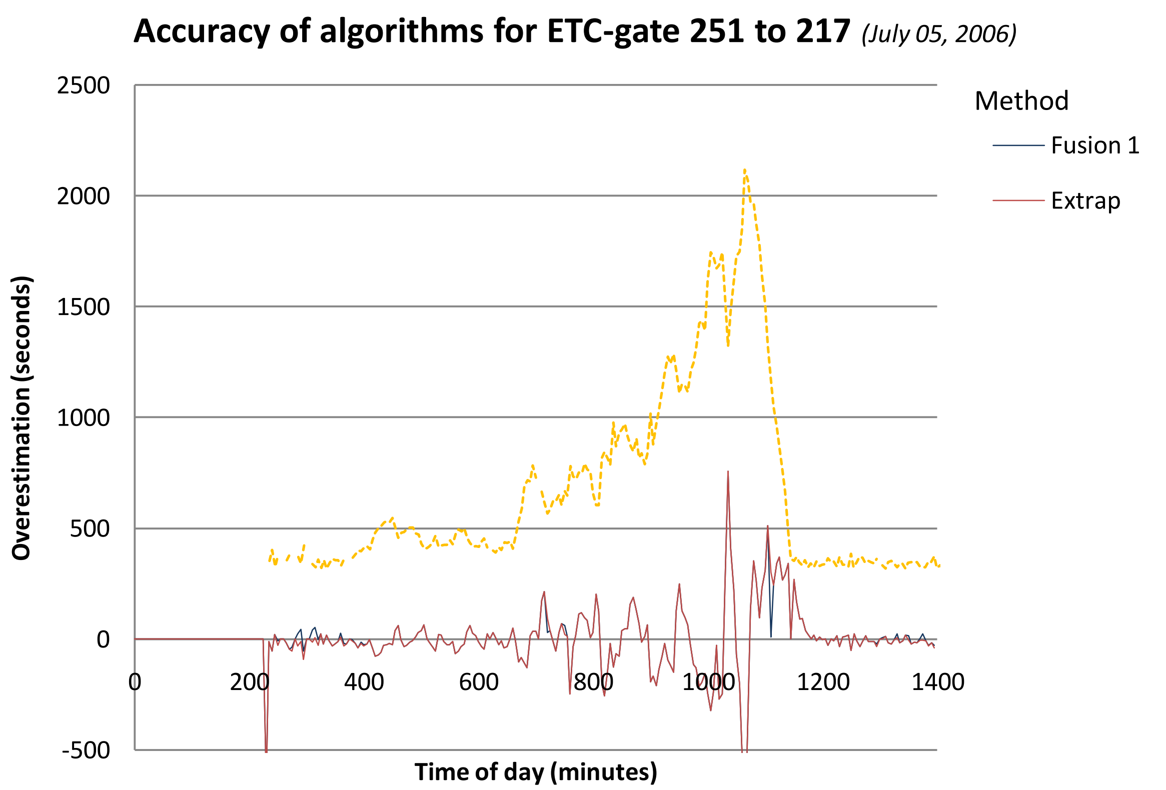


Figure 14 – Travel time estimate errors over time for first fusion model

3.1.2 Corrections on one model based on current trend

Since the first fusion model quickly became very complicated, this second model starts very simple. For the identification of the traffic condition only the last three estimates of the Extrapolation model are used. As for the correction of the estimates, the last two evaluated intervals are used to determine the error and this error is assumed to be the same in the current interval.

Also for further research the research area has been edited to resemble the European and American situations. It is assumed that trying to achieve improvements with the very dense detector placement is not worth the trouble.

The correction rules and traffic condition identifications were as follow:

1. In case the last two estimates by the Extrapolation model are ascending, travel time between two intervals is considered ascending when the increase is more than 20%. Than it is assumed that the travel time in the current situation is ascending and thus the travel time is being underestimated.

   In this case the average error of the last two evaluated intervals is added to the current estimate by the Extrapolation model, thus the output of the fusion model.

2. In case the last two estimates by the Extrapolation model are descending, travel time between two intervals is considered descending when the decrease is more than 20%. Than it is assumed that the travel time in the current situation is descending and thus the travel time is being overestimated.

   In this case the average error of the last two evaluated intervals is deducted from the current estimate by the Extrapolation model, thus the output of the fusion model.

3. For the rest of the situations no corrections will be done.

Illustrations of the above described correction methods are shown in Table 10. The yellow dotted line is the ETC data, the red line is the Extrapolation model, the blue dot is the corrected estimation.

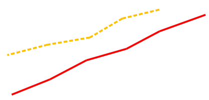
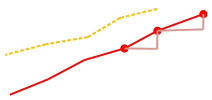
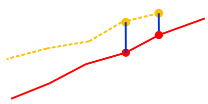
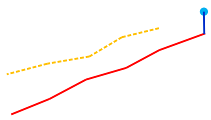
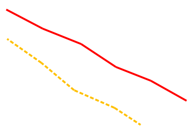
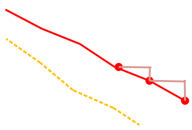
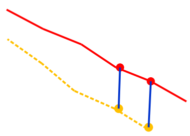
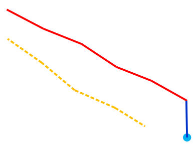
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				
#2				

Table 10 – Illustrations of corrections for second fusion model

The results for this fusion concept turned out to be unsuccessful. In Table 11 the accuracies of this model are shown, only for one out of the seven cases the travel time estimates improved. To investigate why the travel time estimates didn't improve the errors over time were plotted, one case is shown in Figure 14.

Second fusion model (with 30% of all loop detectors)			Extrapolation speed based model (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
105.1136	170.564	137.8386	104.9583	162.861	133.9099
60.69916	73.0866	66.8929	56.46005	70.3793	63.41968
73.54949	25.8932	49.72133	73.84489	25.3089	49.57688
82.2951	103.267	92.78099	87.79284	95.1859	91.48939
114.6733	118.513	116.5931	110.9346	116.031	113.4826
91.8775	51.5757	71.72658	93.85418	52.5319	73.19304
120.3491	119.775	120.0621	117.525	117.347	117.4359
	∑ =	**655.6156**		**∑ =**	**642.5074**

Table 11 – Travel time estimate accuracy for the Second fusion model (from ETC-gate 251 to 217)

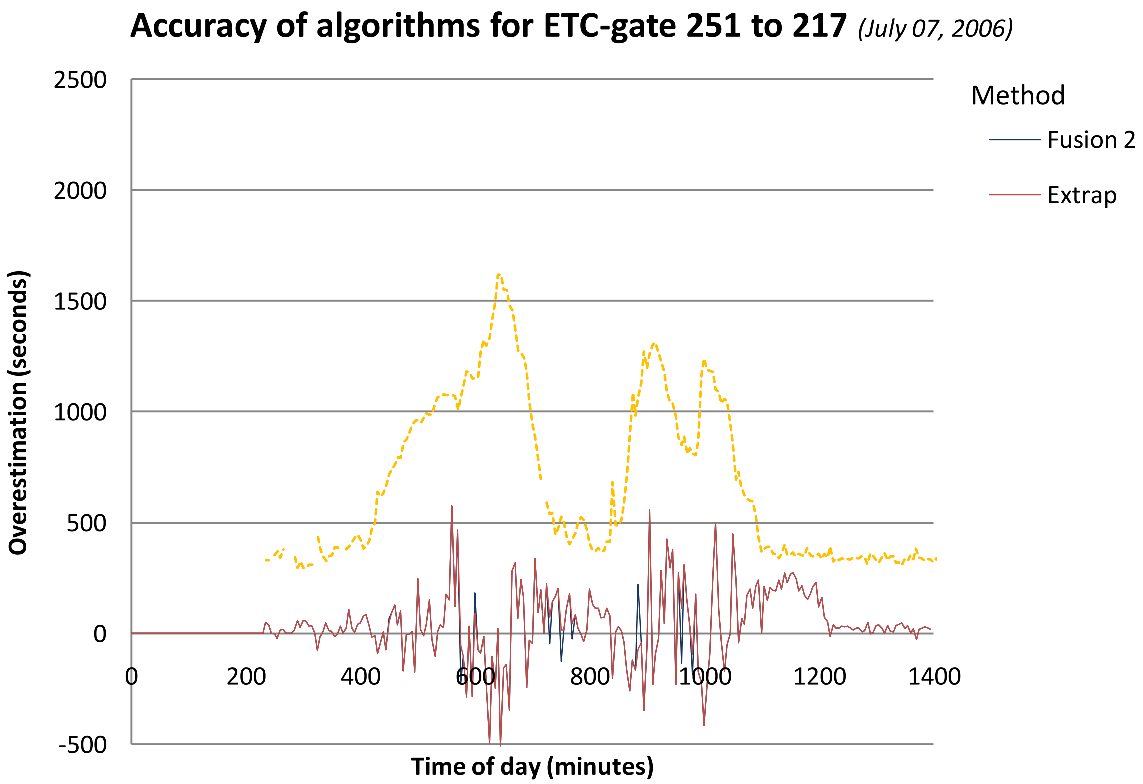


Figure 15 – Travel time estimate errors over time for second fusion model

As can be seen in the figure above, in some situations the travel time does get improved. But there are also situations where travel time estimations get worse. In the condition identification rules the ascending and descending slope values (above set to 20%) have been changed to investigate if it's possible to filter out the wrong situations. Even variations of the above mentioned rules were tested. But all this had very little effect on which situations would get "corrected".

The reason why this fusion failed is because the identification rules are not specific enough, they can't isolate the situations that can be corrected. In this case it's not the rules that are at fault, but the loop detector data. As shown in Figure 16, the estimates provided by the Extrapolation model are too instable. (This is the result of using fewer detectors.) Because the data is not accurate enough the rules can't identify the situations that can be improved.

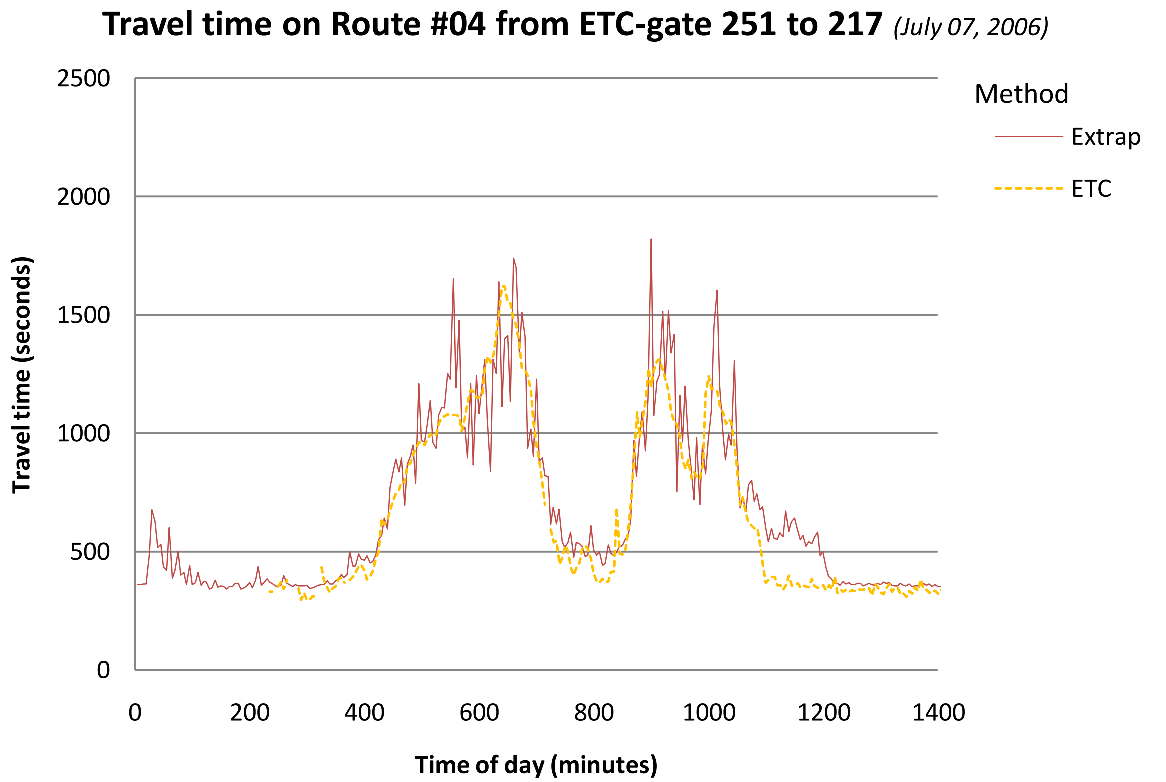


Figure 16 – Travel time estimates by the Extrapolation model (in the situation of fewer detectors)

3.1.3 Introducing moving average to stabilize estimates

In order to make the loop data suitable for identifying traffic conditions, a moving average is introduced. By constantly averaging between values of the current interval and the prior interval, the Extrapolation model's values can be stabilized. Besides the output becoming stable, it also turned out that travel time estimates become more accurate.

For this research two boundaries have been set for applying the moving average. The first boundary is to only use data from the current interval and previous interval. Although averaging between more values resulted in an even more stable output, the delay also increased. Since data with as little delay as possible is preferred, only an average of two intervals is chosen.

The second boundary is to only apply the moving average when the travel time estimate by the Extrapolation model is larger than [1.2 ×Freeflow travel time]. This is to keep the already very accurate travel time estimates untouched, adjusting these estimates resulted in larger errors.

Extrapolation speed based model (+averaging) (with 30% of all loop detectors)			**Extrapolation speed based model** (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
92.95813	137.439	115.1988	104.9583	162.861	133.9099
56.40062	67.5037	61.95215	56.46005	70.3793	63.41968
72.35246	26.4291	49.3908	73.84489	25.3089	49.57688
85.43915	79.3431	82.39111	87.79284	95.1859	91.48939
100.5511	111.871	106.2111	110.9346	116.031	113.4826
91.8912	48.3003	70.09576	93.85418	52.5319	73.19304
113.3327	111.803	112.5678	117.525	117.347	117.4359
	∑ =	**597.8075**		∑ =	**642.5074**

Table 12 – Travel time estimate accuracy when applying moving average (from ETC-gate 251 to 217)

Extrapolation speed based model (+averaging) (with 30% of all loop detectors)			**Extrapolation speed based model** (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
46.79964	59.9822	53.39094	55.2352	69.0866	62.16091
20.83835	27.393	24.11568	19.38976	24.7646	22.07715
14.79507	15.7634	15.27925	15.34376	14.8	15.07188
34.971	28.0749	31.52294	25.81295	29.4115	27.61221
33.33044	11.6814	22.50593	35.8659	14.7118	25.28882
26.14564	20.9081	23.52687	27.57223	26.3659	26.96905
32.50082	51.3333	41.91706	29.26043	55.1148	42.18759
	∑ =	**212.2587**		∑ =	**221.3676**

Table 13 – Travel time estimate accuracy when applying moving average (from ETC-gate 251 to 237)

In Table 12 and Table 13 the Extrapolation model is compared to the Extrapolation model with moving average. In all cases the travel time estimate improved. In Figure 17 a sample of the stabilized output is shown, this is the same data as shown in Figure 16. The moving average application has also been investigated for different detector densities, more about this in Appendix E.

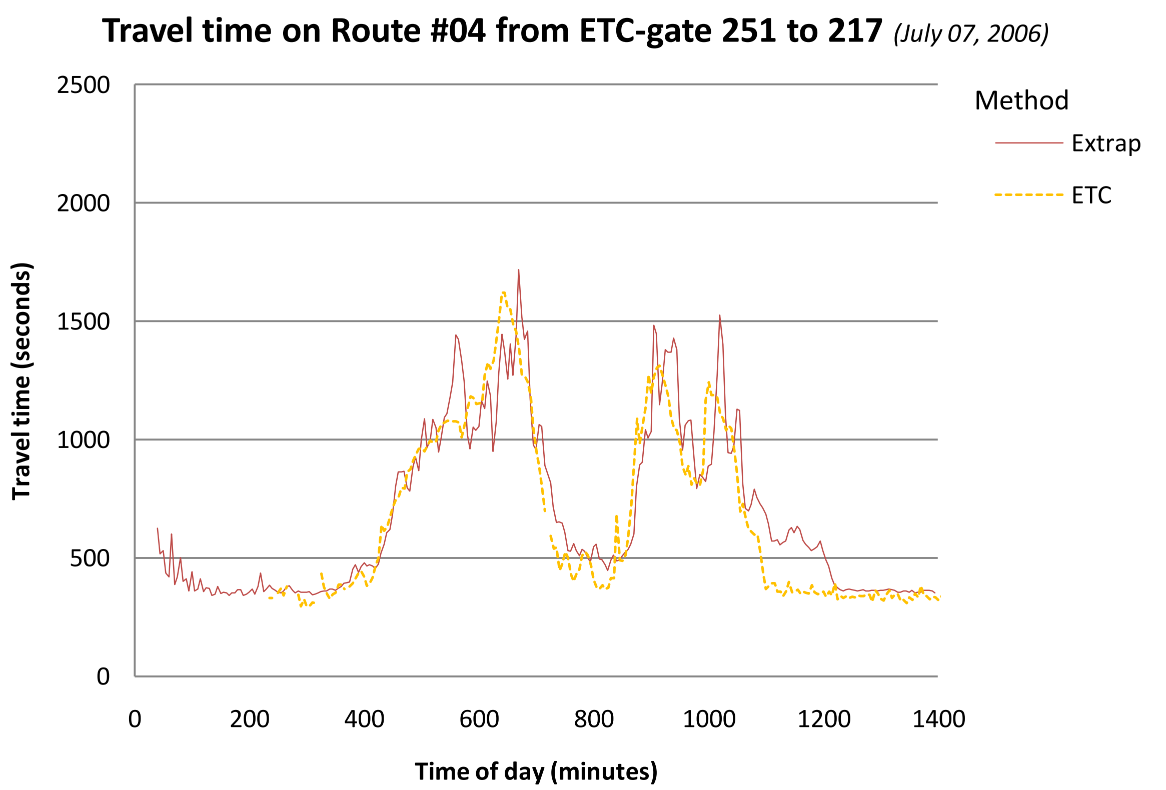


Figure 17 – Travel time estimates by the Extrapolation model with averaging (in the situation of fewer detectors)

3.1.4 Corrections on one model with moving average based on current trend

With the moving average applications it is expected that the Extrapolation model’s output is stable enough for identifying traffic conditions. Because of time constrains it was not possible to fully investigate this third fusion model. Basically for this model the same conditions and correction rules were written as for the second data fusion model. Still these rules seemed to be unable to identify the right situations for correction.

One condition and correction rule that was written for this third model that is successful is as follow:

1. In case the last two estimates by the Extrapolation model are first ascending and then descending, criteria to meet is a 20% change or more compared to the previous interval. Than it is assumed that the travel time in the current situation is overestimated.

   In this case the average error of the last two evaluated intervals is deducted from the current estimate by the Extrapolation model, thus the output of the fusion model.

The above described correction method is illustrated in Table 14. The yellow dotted line is the ETC data, the red line is the Extrapolation model, the blue dot is the corrected estimation.

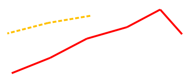
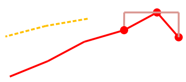
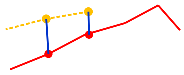
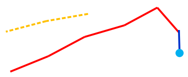
Correction rule	Situation	Recognize situation	Determine error	Correction
#1				

Table 14 – Illustrations of correction for third fusion model

The reason for this rule is that the all peaks of the travel time estimation graph by the Extrapolation model are overestimated. An after the peaks there usually is a descending part. Accuracy results of this (still under construction) third model are shown in Table 15 and Table 16.

The green values in the tables are the values for the current situation with very dense detector placement, these are the values that need to be maintained. For now the model is not capable of achieving that level of accuracy, it is possible that with more condition and correction rules this level can be reached. Further research is required to investigate this.

Third fusion model (with 30% of all loop detectors)			**Extrapolation speed based model** (+averaging) (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
93.4612	131.599	112.5299	92.95813	137.439	115.1988
57.1456	64.0745	60.61005	56.40062	67.5037	61.95215
69.15854	29.6859	49.42222	72.35246	26.4291	49.3908
85.91212	75.7059	80.8090	85.43915	79.3431	82.39111
100.5495	105.675	103.1121	100.5511	111.871	106.2111
89.55234	44.1939	66.87313	91.8912	48.3003	70.09576
113.881	105.816	109.8487	113.3327	111.803	112.5678
	∑ =	**583.2050**		∑ =	**597.8075**
		496.4069			

Table 15 – Travel time estimate accuracy the Third fusion model (from ETC-gate 251 to 217)

Third fusion model (with 30% of all loop detectors)			**Extrapolation speed based model** (+averaging) (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
46.67961	60.5047	53.59215	46.79964	59.9822	53.39094
20.48801	26.3895	23.43878	20.83835	27.393	24.11568
14.48537	14.7637	14.62453	14.79507	15.7634	15.27925
36.70476	26.4531	31.57895	34.971	28.0749	31.52294
33.52366	12.022	22.77284	33.33044	11.6814	22.50593
26.98931	20.5219	23.7556	26.14564	20.9081	23.52687
32.81215	48.0271	40.41962	32.50082	51.3333	41.91706
	∑ =	**210.1825**		∑ =	**212.2587**
		248.6903			

Table 16 – Travel time estimate accuracy the Third fusion model (from ETC-gate 251 to 237)

In Table 16 the travel time estimate improved compared to the current situation, this improvement is probably a coincidence. Quick analyses of more cases turned out that this was the only case that showed an improvement.

3.2 Discussion new models

In the first fusion model, traffic conditions were identified by looking at time-intervals where ETC data was available. This seemed to work, since the first model was a successful data fusion. But this method of identifying traffic conditions is not suited for small traffic variations, because there is a relatively long delay before a condition has been recognized.

The use of multiple models for travel time estimation and correction turned out to be very complicated. There are just too many variables to work with, which is not convenient for examining a new fusion methodology. For further research it might be interesting to run multiple models to identify very specific situations, but for this research such level of detail is not needed yet.

Last note for this first model, is that this was done with all detectors. Because of the very dense detector placements, travel time estimates were very accurate already and didn't leave much space over for improvement.

For the second fusion model only 30% of all detectors were used. This created a situation with more room for improvement. And the situation would resemble the more common European and American detector placements.

To minimize the delay in identifying traffic conditions, the last three estimates by the Extrapolation model were used for identification. This meant that more up-to-date data would be used, but it turned out to be an unsuccessful fusion. The use of 30% of all detectors resulted in a very varying estimate by the Extrapolation model, which was not suitable for identifying traffic conditions.

So this simpler model with only the Extrapolation model as base was not able to improve travel time estimates.

In the last model a moving average was introduced on the Extrapolation model, this to stabilize the output and making it suitable for indentifying traffic conditions. And in this model a criteria was introduced, so that free-flow condition remained untouched. Since improving travel time estimates for this conditions is unnecessary, if not impossible.

Because of time constrains only one correction rule was made for this model, so the improvement of this model is minimal. But this model is a successful fusion and further development will most likely result in more accuracy.

4 Conclusions and recommendations

In this research several fusion concepts have been examined. Unlike previous data fusion attempts, here no statistical or historical method is used. This means that the developed model can likely be applied anywhere without the need of prior research. Further research, by applying the new model on another study area, is required to confirm this.

Out of the tree examined fusion concepts, the first and the third models are successful data fusions. In both cases the improvements can be much more if the condition and correction rules are more developed.

The concept of running several models parallel and using ETC data to average or correct estimates is a somewhat difficult to work with model. Since there are many variables that can be used to identify traffic conditions and determine errors. Also the usage of errors of previous time-intervals is questionable, since this data comes with a relatively long delay (equal to the travel time).

Just like the third model it is better to use the output of an instantaneous model to determine traffic conditions. If possible the use of another data source for traffic condition identification is preferred, such as camera detection or probe vehicles. Basically travel times can be estimated very accurately if the correct traffic condition is timely identified. Based on the traffic condition pre-defined correction rules can be used to correct systematic errors.

For further research more detailed condition and correction rules need to be developed, for example using more variable to identify traffic conditions. This research has only demonstrated a fusion method that can be successful. Travel time estimations by instantaneous models depending on loop data clearly have systematic errors (Figure 13), correcting these errors without statistical methods is possible. And this research has pointed out that identifying traffic conditions using loop detector data needs some averaging over time (about 10 minutes).

As extra this research showed that the application of moving average can improve travel time estimations depending on how dense the detector placement is.

One last issue that needs to be pointed out is the use of ETC data. Some filtering of unusually long travel times is needed. For the models above ETC data was filtered by comparison with the Extrapolation model. Any ETC entry that exceeds the Extrapolation model by 80% or more was discarded.

For further research a better filter method should be developed, since this 80% was just used to quickly get started with the research. A possible real-time filtering method is by running a Time slice model in real-time. By comparing incoming ETC data with the time slice, it is quite easy to determine whether a vehicle was exceptionally slow or not. Both travel times should arrive with similar delays, so there will be no extra delay for filtering the ETC data.

References

Jain, M. and Coifman, B. (2005). Improved Speed Estimates from Freeway Traffic Detectors. [Electronic version]
Dept. of Civil and Environmental Engineering and Geodetic Science, The Ohio State University
Retrieved on February 17, 2009 from:
http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=JTPEDI000131000007000483000001&idtype=cvips&gifs=yes

Lelitha D. Vanajakshi, Billy M. Williams, and Laurence R. Rilett (2009). Improved Flow-Based Travel Time Estimation Method from Point Detector Data for Freeways. [Electronic version]
J. Transp. Engrg. Volume 135, Issue 1, pp. 26-36 (January 2009)
Retrieved on February 05, 2009 from:
http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=JTPEDI000135000001000026000001&idtype=cvips&gifs=yes

N. Zou, J Wang, G. Chang, J. Paracha (2008). A Hybrid Model for Reliable Travel Time Estimation on a Freeway with Sparsely Distributed Detectors. [Electronic version]
Department of Civil and Environmental Engineering, University of Maryland
Retrieved on February 25, 2009 from:
http://attap.umd.edu/bbs/data/publications/Travel_Time_Estimation_World_Congress.pdf

Nam, D. H., and Drew, D. R. (1998). Analyzing freeway traffic under congestion: Traffic dynamics approach. [Electronic version]
J. Transp. Engrg. Volume 124, Issue 3, pp. 208-212 (May/June 1998)
Retrieved on February 19, 2009 from:
http://scitation.aip.org/getpdf/servlet/GetPDFServlet?filetype=pdf&id=JTPEDI000124000003000208000001&idtype=cvips&prog=normal

Nour-Eddin El Faouzi, Lawrence A. Klein, and Olivier De Mouzon (2009). Improving Travel Time Estimates from Loop and Toll Collection Data with Dempster-Shafer Data Fusion. [Electronic version]
Laboratoire d'Ingénierie Circulation Transports (LICIT), INRETS-ENTPE
Retrieved on April 15, 2009 from:
By e-mail from Masao Kuwahara

Oh, J., Jayakrishnan, R., and Recker, W. (2003). Section Travel Time Estimation from Point Detection Data. [Electronic version]
Institute of Transportation Studies, University of California
Retrieved on February 21, 2009 from:
http://www.its.uci.edu/its/publications/papers/CTSS/UCI-ITS-TS-WP-02-14.pdf

Ruimin Li, Geoffrey Rose, and Majid Sarvi (2006). Evaluation of Speed-Based Travel Time Estimation Models. [Electronic version]
J. Transp. Engrg. Volume 132, Issue 7, pp. 540-547 (July 2006)
Retrieved on April 3, 2009 from:
By e-mail from Babak Mehran

Sirisha M. Kothuri, Kristin Tufte, Soyoung Ahn and Robert L. Bertini (2006). Development of an ITS data archive application for improving freeway travel time estimation. [Electronic version] *Department of Civil and Environmental Engineering, Portland State University* Retrieved on February 05, 2009 from: http://ieeexplore.ieee.org/search/searchresult.jsp?queryText=Kothuri%2C+S.M.&coll2=ieeecnfs&coll3=ieecnfs&history=yes&reqloc=others&scope=metadata&imageField2.x=7&imageField2.y=12

Van Hinsbergen & Van Lint (2008). Bayesian combination of travel time prediction models. [Electronic version] *Faculty of Civil Engineering and Geosciences, Delft University of Technology* Retrieved on April 20, 2009 from: TRB 2008 Annual Meeting CD-ROM

Van Lint, J.W.C., and Van der Zijpp, N.J. (2003). An improved travel-time estimation algorithm using dual loop detectors. [Electronic version] *Faculty of Civil Engineering and Geosciences, Delft University of Technology* Retrieved on February 20, 2009 from: http://www.ltrc.lsu.edu/TRB_82/TRB2003-000822.pdf

Ying Liu, Xiaorong Lai, Gang-Len Chang (2006). Detector placement strategies for freeway travel time estimation. [Electronic version] *Department of Civil and Environmental Engineering, University of Maryland* Retrieved on February 05, 2009 from: http://ieeexplore.ieee.org/search/searchresult.jsp?queryText=(detector+placement+strategies+for+freeway+travel+time+estimation+%3Cin%3E+metadata)+%3Cand%3E+(11180+%3Cin%3E+punumber)&coll2=ieeecnfs&coll3=ieecnfs&history=yes&reqloc=others&scope=metadata&imageField2.x=6&imageField2.y=8

Appendix A – Comparison of Extrapolation and Midpoint models

For the comparison of the Extrapolation speed based model with the Midpoint model a route from ETC-gate 251 to ETC-gate 217 has been used. The travel times estimated by the two models for July 1 2006 are shown in Figure 18, the error of both models are shown in Figure 19.

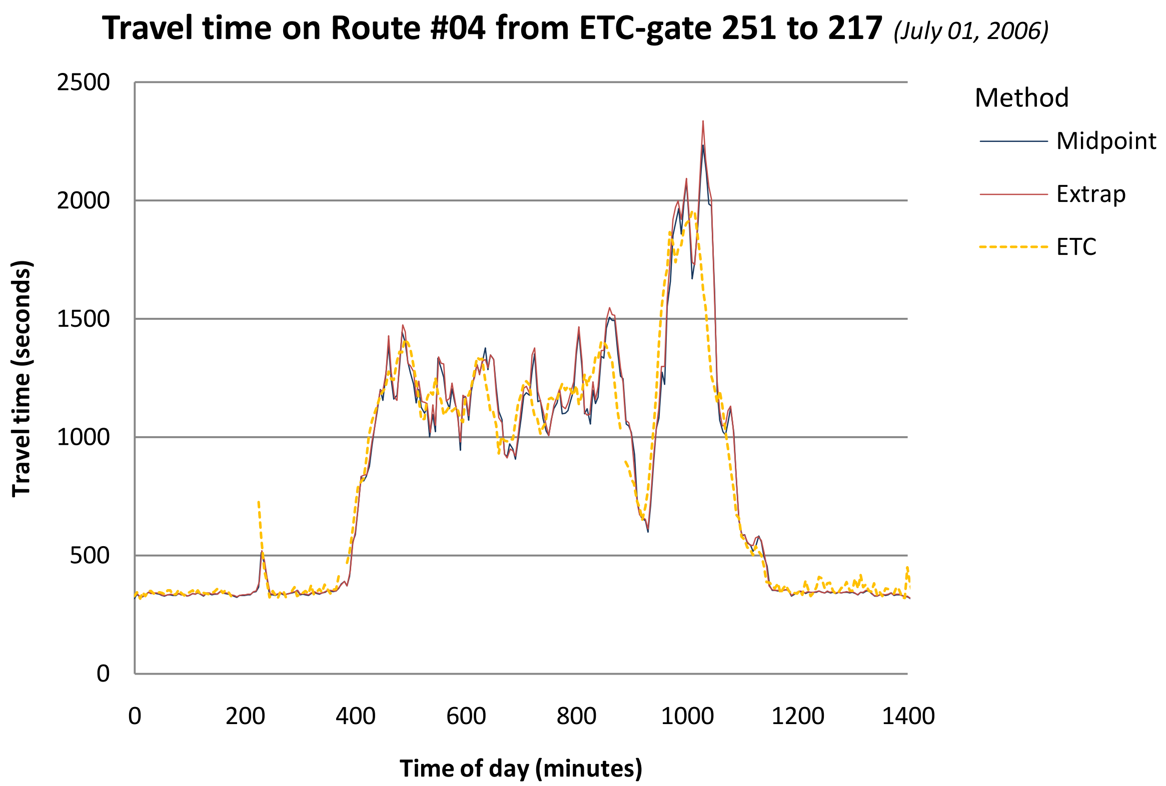


Figure 18 – Travel time comparison between Extrapolation speed based and Midpoint

In the above figure the blue line is represents the travel times estimated using the Midpoint model. The pink line represents the estimated times by the Extrapolation speed based model. The yellow dotted line is the actual travel time obtained from the ETC data.

The figure below shows the errors of both the models. For each time-interval the error is determined and plotted. Values above zero are overestimations and below zero are underestimations. Again the yellow dotted line is the actual travel time.

It is clear that both models behave very similar with similar errors. To be sure that the shown results are not coincidence, comparison has been done for data sets from July 1 till July 7 (2006). On all seven days the results of both models are very similar. In Table 17 the errors of both models for the seven days are shown.

Based on these results it is clear that only one of these models needs to be considered in the research. Since the Extrapolation speed based is more accurate, this model will be kept in the research.

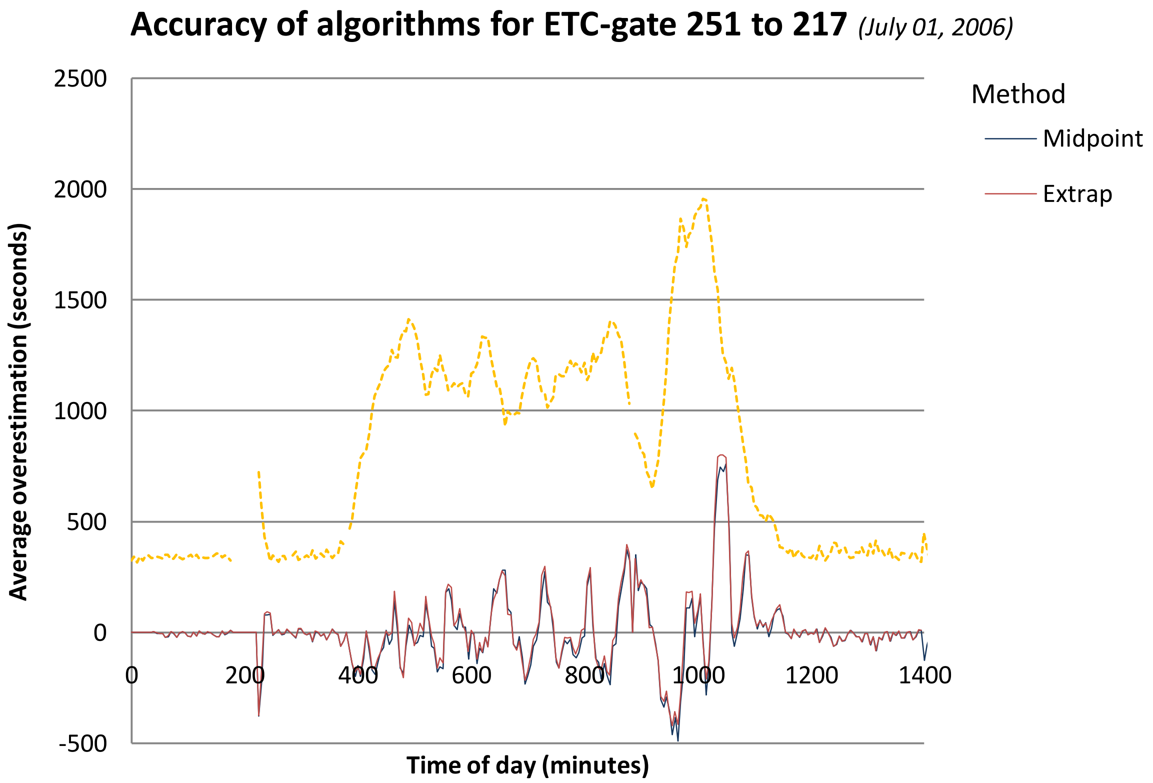


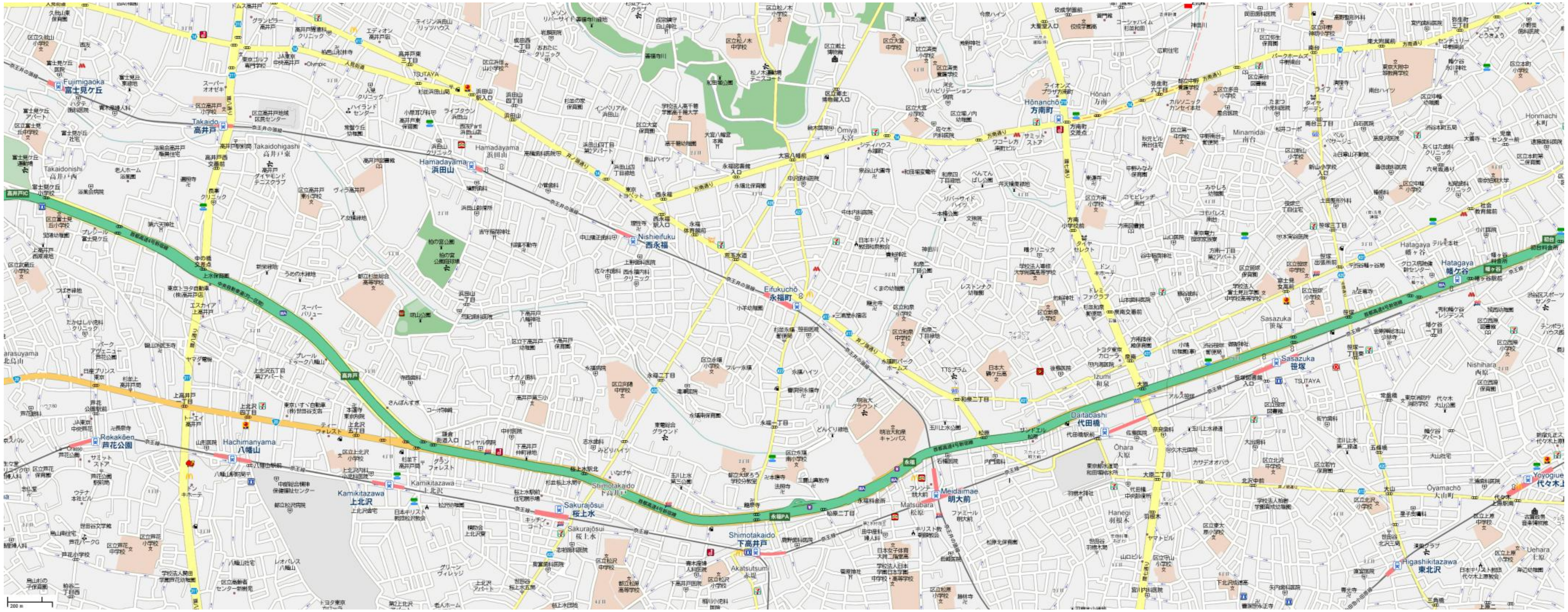
Figure 19 – Estimate error comparison between Extrapolation speed based and Midpoint

Extrapolation speed based model (original)			**Midpoint model** (original)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
129.1859	72.7797	100.9828	128.1331	77.6715	102.9023
86.32575	47.437	66.8814	90.3093	50.209	70.25917
50.41537	37.1489	43.78214	48.11783	39.6836	43.90072
81.27327	50.0184	65.64581	83.1873	54.9012	69.04423
109.3794	74.1254	91.75241	97.02928	81.4695	89.24937
54.148	35.9148	45.03138	53.92504	41.2007	47.56289
93.97934	70.6825	82.33094	92.69785	71.3806	82.03923
	∑ =	**496.4069**		∑ =	**504.9579**

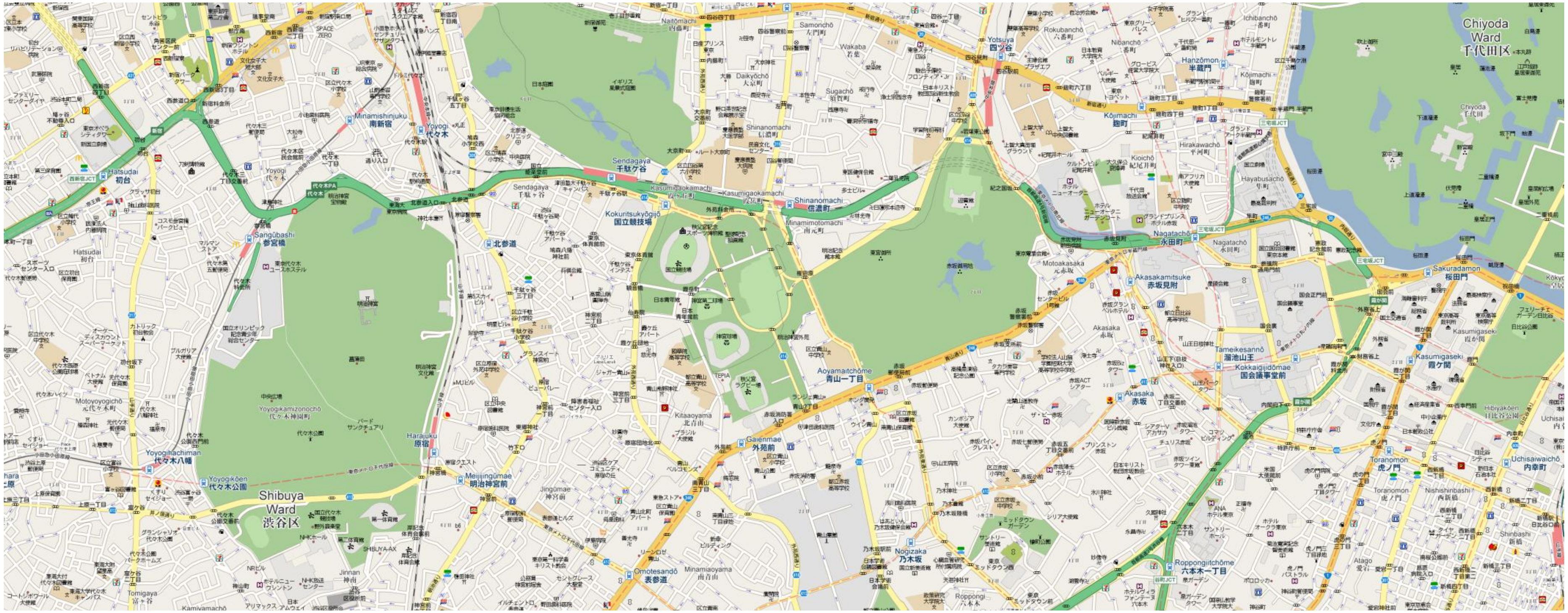
Table 17 – Estimate error comparison between Extrapolation speed based and Midpoint (from ETC-gate 251 to 217)

In the above table all errors are given in seconds. In the last row the sum of all average absolute errors of all seven days is shown. The extrapolation model overall seems more accurate than the Midpoint model.

Appendix B – Map of study area



Part 1 of 2 (left-side)



Part 2 of 2 (right-side)

Appendix C – Unrealistic behavior by Nam and Drew model

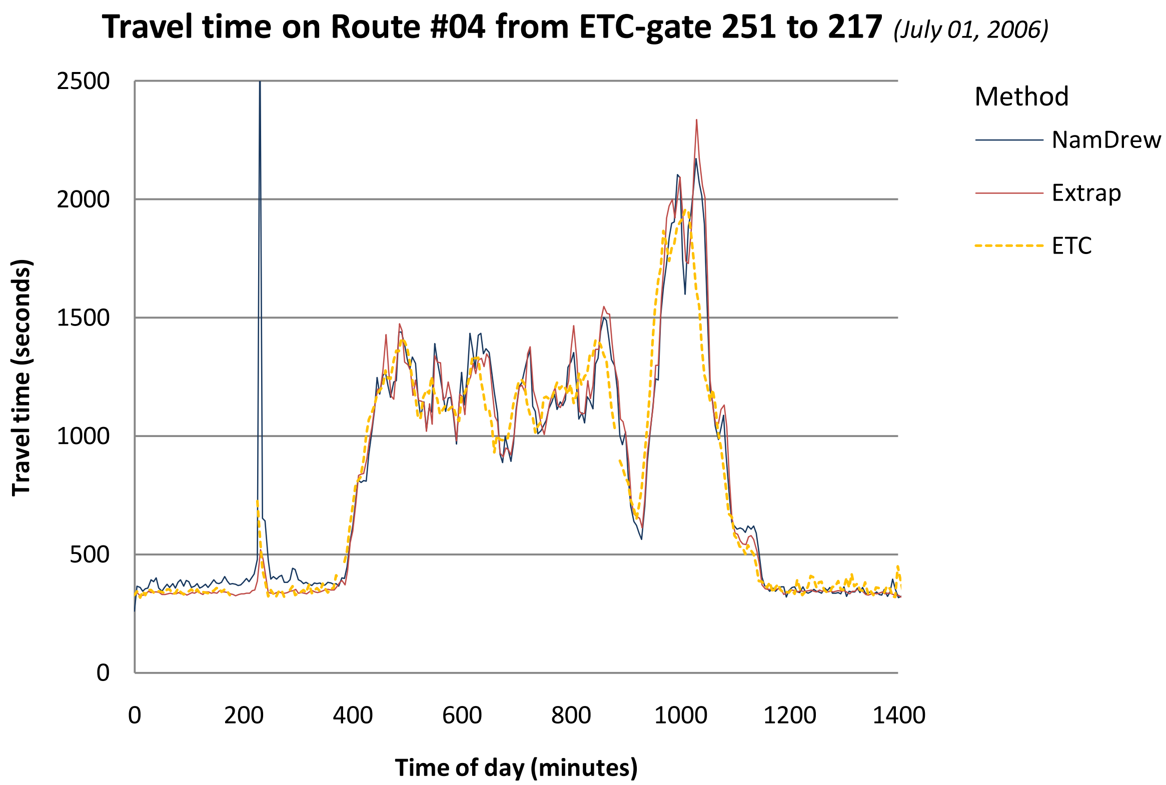
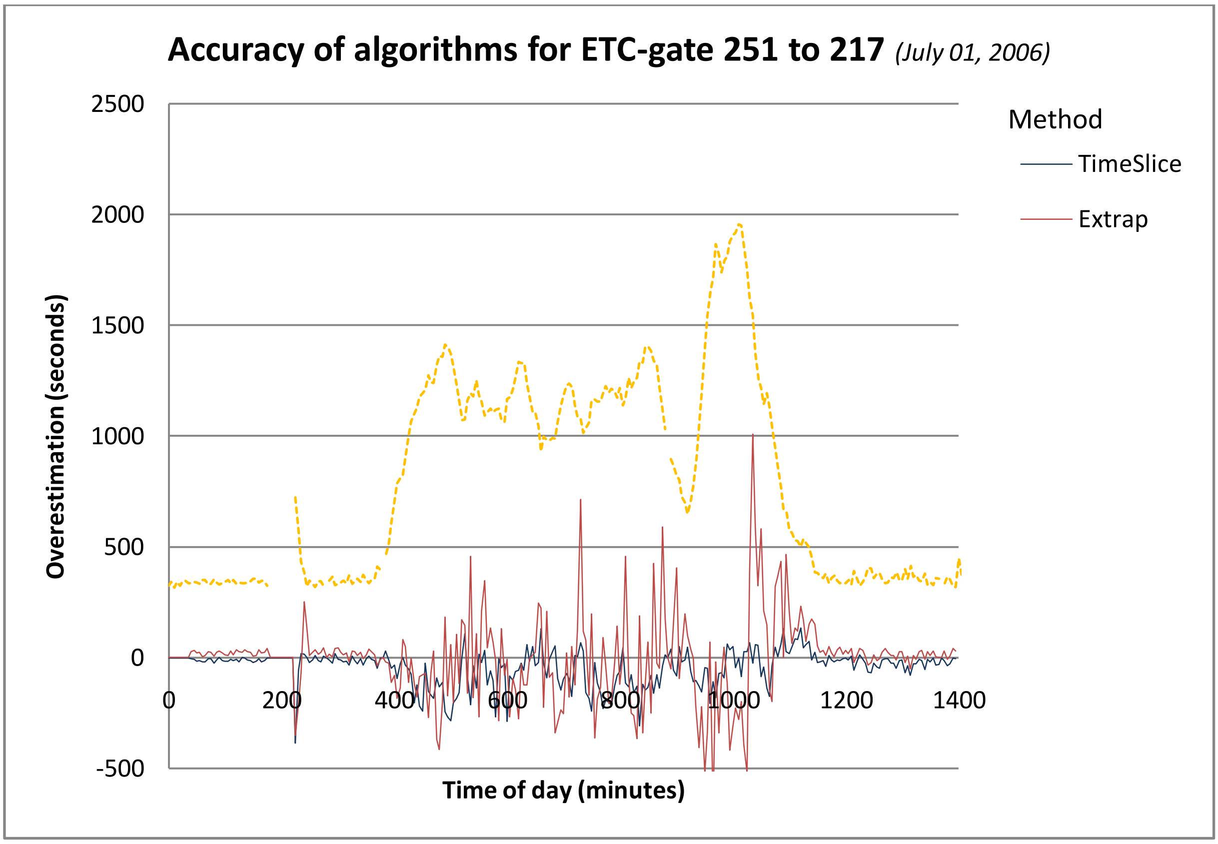


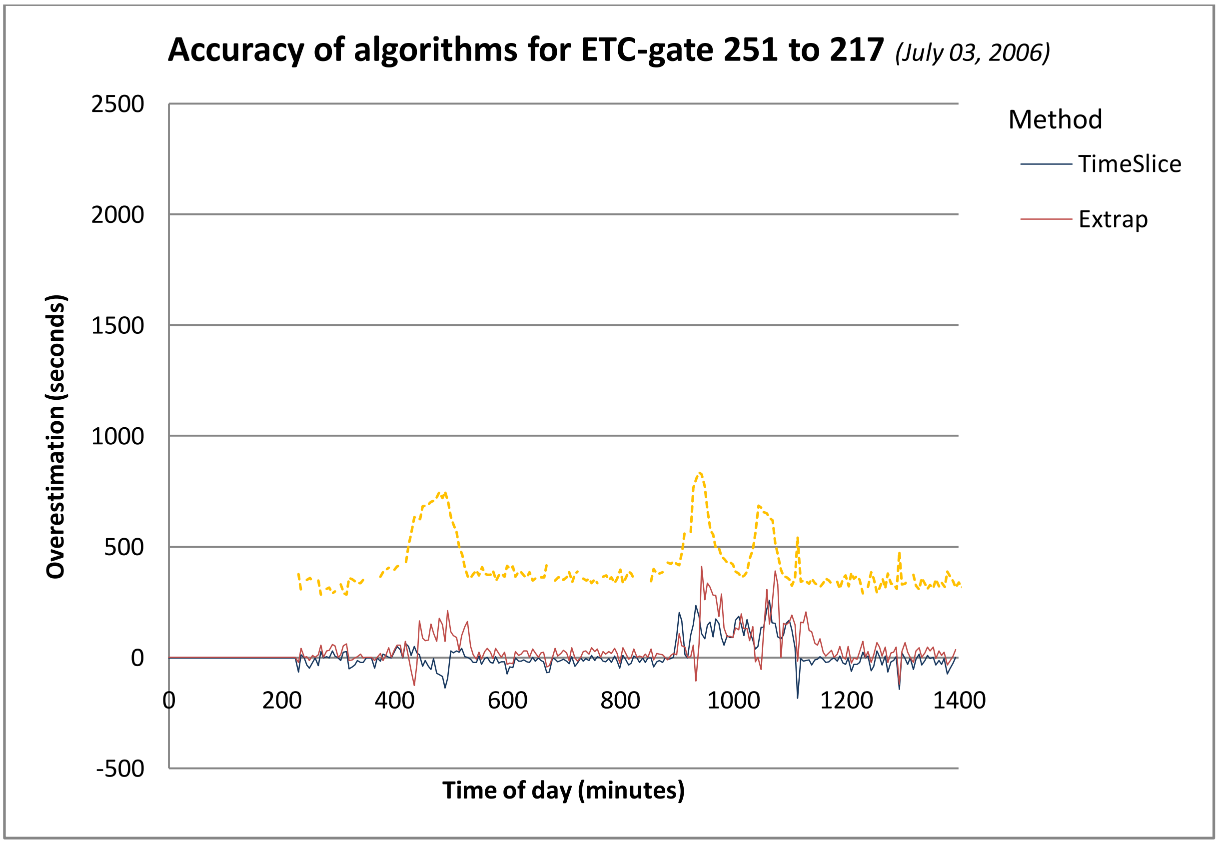
Figure 20 – Unrealistic behavior by Nam and Drew dynamics model

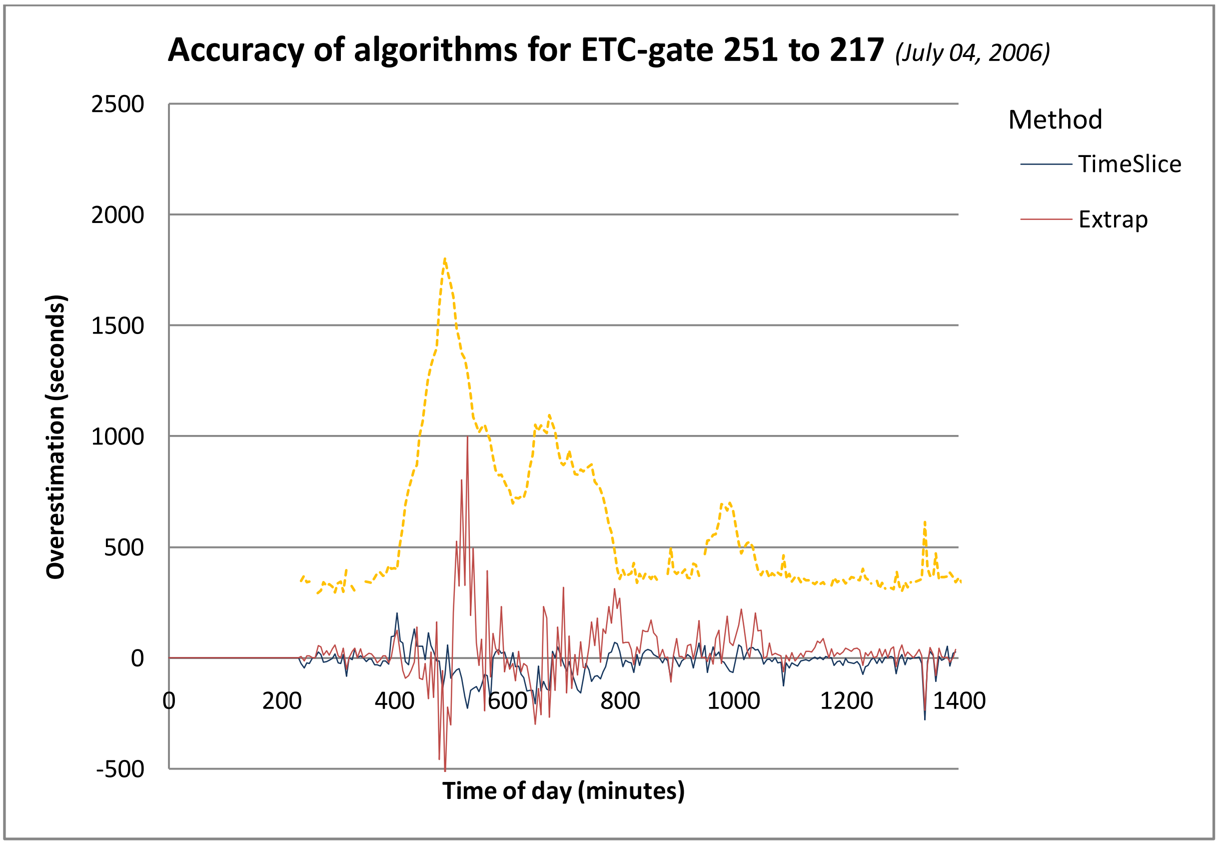
In Figure 20 the travel times on July 1, 2006 according to the Nam and Drew model and the Extrapolation model are shown. Again the yellow dotted line is the actual travel time obtained from the ETC data. At around 4:00 AM (±240 minutes) an accident happened, for this period no ETC data is available. But the travel time estimates by the Nam and Drew model are rather exceptional.

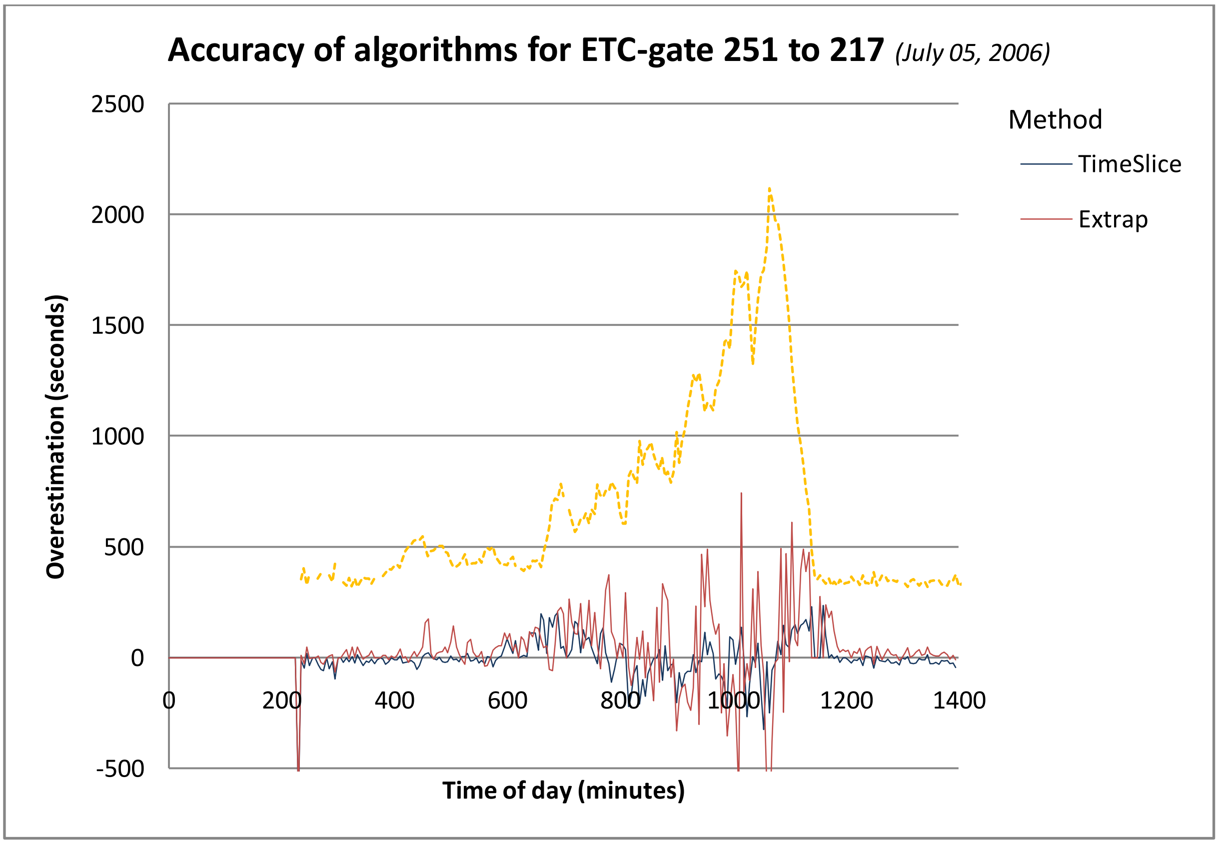
For this reason the data for this day is no included in Table 4 of this report.

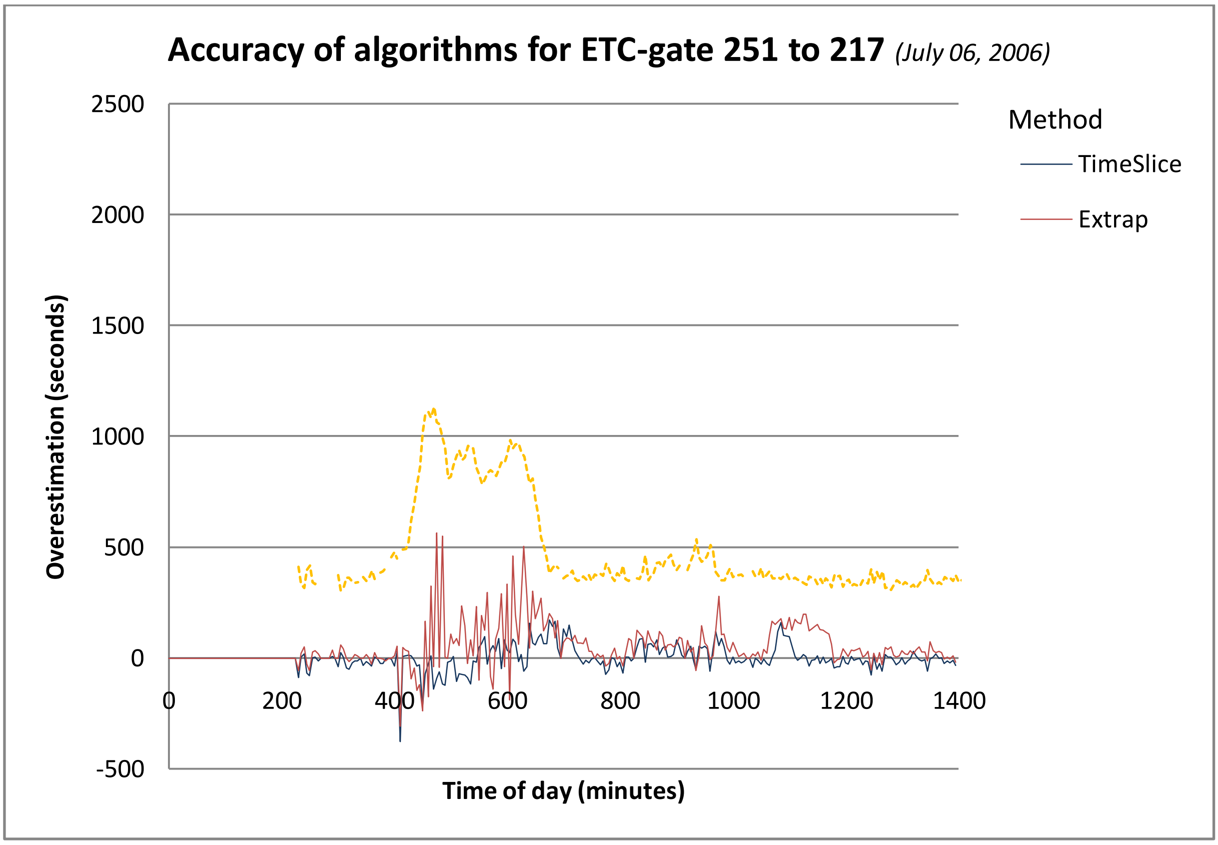
Appendix D – Remaining error over time graphs

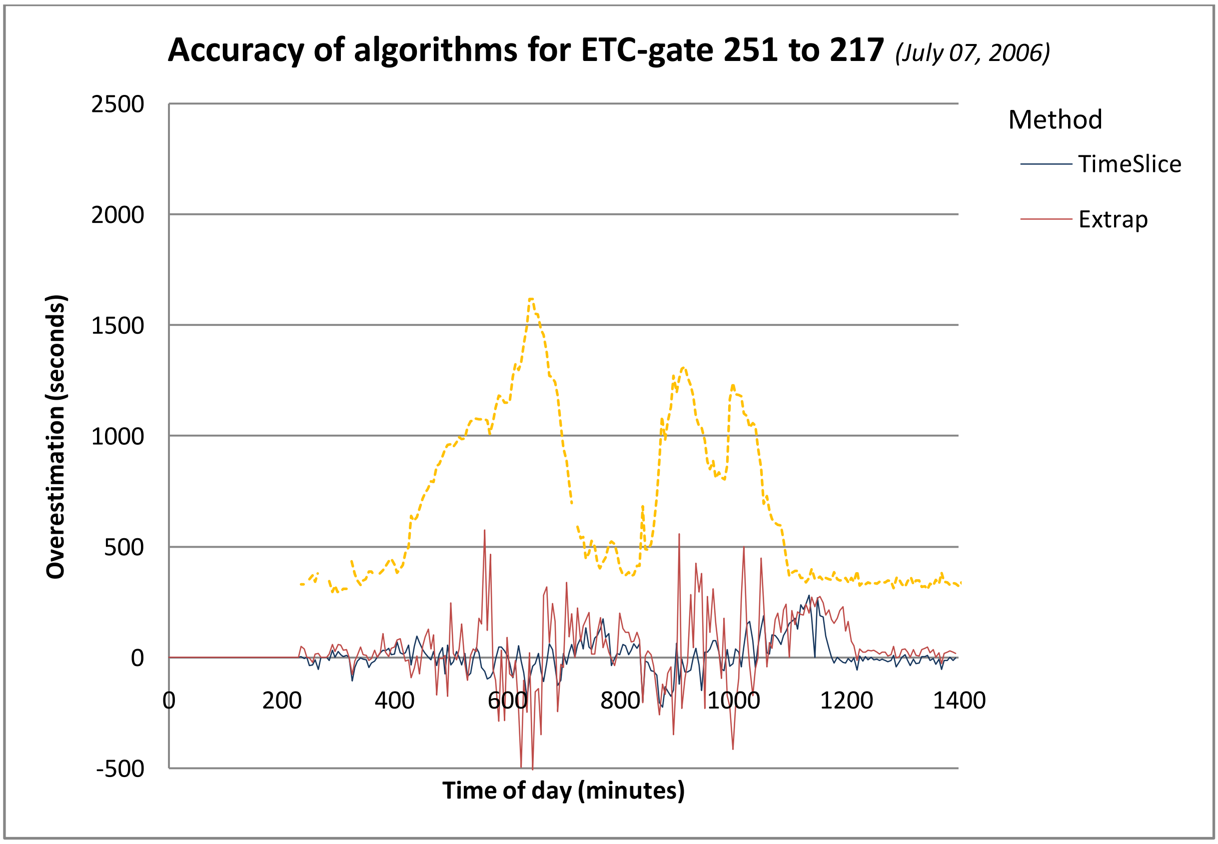





Accuracy of algorithms for ETC-gate 251 to 217 (July 04, 2006)
Overestimation (seconds)
2500
2000
1500
1000
500
0
-500
0
200
400
600
800
1000
1200
1400
Time of day (minutes)
Method
TimeSlice
Extrap


Accuracy of algorithms for ETC-gate 251 to 217 (July 05, 2006)
Overestimation (seconds)
2500
2000
1500
1000
500
0
-500
0
200
400
600
800
1000
1200
1400
Time of day (minutes)
Method
TimeSlice
Extrap


Accuracy of algorithms for ETC-gate 251 to 217 (July 06, 2006)
Overestimation (seconds)
2500
2000
1500
1000
500
0
-500
0
200
400
600
800
1000
1200
1400
Time of day (minutes)
Method
TimeSlice
Extrap


Accuracy of algorithms for ETC-gate 251 to 217 (July 07, 2006)
Overestimation (seconds)
2500
2000
1500
1000
500
0
-500
0
200
400
600
800
1000
1200
1400
Time of day (minutes)
Method
TimeSlice
Extrap

Appendix E – Moving average in relation to detector placement

Since it turned out that the application of moving average improved travel time estimates for the situation where about 70% of the detectors were dropped out. It was quickly examined if this Extrapolation model with averaging also improved estimates for more dense detector placements.

The considered scenarios were, keep 100% of all detectors and keep 50% of all detectors. Results of this quick check are shown in the tables below (14, 15, and 16). It turns out that after dropping out more than 50% of all detectors this "new" model starts to show improvements in estimates.

Extrapolation speed based model (+averaging) (with all loop detectors)			**Extrapolation speed based model** (with all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
137.3156	79.3774	108.3465	129.1859	72.7797	100.9828
79.87185	52.3103	66.09108	86.32575	47.437	66.8814
54.51507	39.6341	47.07456	50.41537	37.1489	43.78214
87.09521	53.6365	70.36588	81.27327	50.0184	65.64581
110.6183	79.0027	94.81054	109.3794	74.1254	91.75241
58.69122	38.1371	48.41416	54.148	35.9148	45.03138
102.2022	77.3971	89.79965	93.97934	70.6825	82.33094
	∑ =	**524.9024**		∑ =	**496.4069**

Table 18 – Travel time estimate accuracy when applying moving average (from ETC-gate 251 to 217, keep 100%)

Extrapolation speed based model (+averaging) (with 50% of all loop detectors)			**Extrapolation speed based model** (with 50% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
144.029	99.4448	121.7369	136.6092	102.782	119.6954
84.89793	50.1193	67.50863	61.66607	47.7188	54.69245
86.68495	30.7861	58.73554	78.86058	31.3595	55.11006
81.2006	59.4317	70.31616	86.28192	60.157	73.21947
114.7654	91.9083	103.3369	115.8032	86.7372	101.2702
69.42074	43.6511	56.53594	70.84424	41.8718	56.35803
127.4619	75.8314	101.6466	110.2141	79.3546	94.78437
	∑ =	**579.8167**		∑ =	**555.1300**

Table 19 – Travel time estimate accuracy when applying moving average (from ETC-gate 251 to 217, keep 50%)

Extrapolation speed based model (+averaging) (with 30% of all loop detectors)			**Extrapolation speed based model** (with 30% of all loop detectors)		
Average overestimation	*Average underestimation*	*Average absolute error*	*Average overestimation*	*Average underestimation*	*Average absolute error*
92.95813	137.439	115.1988	104.9583	162.861	133.9099
56.40062	67.5037	61.95215	56.46005	70.3793	63.41968
72.35246	26.4291	49.3908	73.84489	25.3089	49.57688
85.43915	79.3431	82.39111	87.79284	95.1859	91.48939
100.5511	111.871	106.2111	110.9346	116.031	113.4826
91.8912	48.3003	70.09576	93.85418	52.5319	73.19304
113.3327	111.803	112.5678	117.525	117.347	117.4359
	∑ =	**597.8075**		∑ =	**642.5074**

Table 20 – Travel time estimate accuracy when applying moving average (from ETC-gate 251 to 217, keep 30%)